




# Joint Identification and Sensing for Discrete Memoryless Channels

Wafa Labidi <sup>1,†</sup>, Yaning Zhao <sup>2,\*,†</sup>, Christian Deppe <sup>2,†</sup> and Holger Boche <sup>1,†,‡,§,||</sup>

- <sup>1</sup> Chair of Theoretical Information Technology, Technical University of Munich, 80333 Munich, Germany; wafa.labidi@tum.de (W.L.); boche@tum.de (H.B.)
- <sup>2</sup> Institute for Communications Technology, Technische Universität Braunschweig, 38106 Brunswick, Germany; christian.deppe@tu-bs.de
- \* Correspondence: yaning.zhao@tu-bs.de
- † Current address: BMBF Research Hub 6G-Life, 80333 Munich, Germany.
- ‡ Current address: Cyber Security in the Age of Large-Scale Adversaries–Exzellenzcluster, Ruhr-Universität Bochum, 44780 Bochum, Germany.
- § Current address: Munich Center for Quantum Science and Technology (MCQST), 80799 Munich, Germany.
- || Current address: Munich Quantum Valley (MQV), 80799 München, Germany.

**Abstract:** In the identification (ID) scheme proposed by Ahlswede and Dueck, the receiver’s goal is simply to verify whether a specific message of interest was sent. Unlike Shannon’s transmission codes, which aim for message decoding, ID codes for a discrete memoryless channel (DMC) are far more efficient; their size grows doubly exponentially with the blocklength when randomized encoding is used. This indicates that when the receiver’s objective does not require decoding, the ID paradigm is significantly more efficient than traditional Shannon transmission in terms of both energy consumption and hardware complexity. Further benefits of ID schemes can be realized by leveraging additional resources such as feedback. In this work, we address the problem of joint ID and channel state estimation over a DMC with independent and identically distributed (i.i.d.) state sequences. State estimation functions as the sensing mechanism of the model. Specifically, the sender transmits an ID message over the DMC while simultaneously estimating the channel state through strictly causal observations of the channel output. Importantly, the random channel state is unknown to both the sender and the receiver. For this system model, we present a complete characterization of the ID capacity–distortion function.



Academic Editors: Shlomo Shamai (Shitz), Mehrasa Ahmadipour and Michele Wigger

Received: 19 November 2024

Revised: 23 December 2024

Accepted: 24 December 2024

Published: 27 December 2024

**Citation:** Labidi, W.; Zhao, Y.; Deppe, C.; Boche, H. Joint Identification and Sensing for Discrete Memoryless Channels. *Entropy* **2025**, *27*, 12. <https://doi.org/10.3390/e27010012>

**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** joint identification and sensing; message identification; information theory

## 1. Introduction

The identification (ID) scheme suggested by Ahlswede and Dueck [1] in 1989 is conceptually different from the classical message transmission scheme proposed by Shannon [2]. In classical message transmission, the encoder transmits a message over a noisy channel; at the receiver side, the aim of the decoder is to output an estimation of this message based on the channel observation. In the ID paradigm, however, the encoder sends an ID message (also called the identity) over a noisy channel, and the decoder aims to check whether or not a specific ID message of special interest to the receiver has been sent. Obviously, the sender has no prior knowledge of the specific ID message that the receiver is interested in. Ahlswede and Dueck demonstrated that in the theory of ID [1], if randomized encoding is used, then the size of ID codes for discrete memoryless channels (DMCs) grows doubly exponentially fast with the blocklength. If only deterministic encoding is allowed, then the number of identities that can be identified over a DMC scales exponentially with the blocklength. Nevertheless, the rate is still more significant than the transmission rate in the exponential scale, as shown in [3,4].

New applications in modern communications demand high reliability and latency requirements, including machine-to-machine and human-to-machine systems, digital watermarking [5–7], industry 4.0 [8], and 6G communication systems [9,10]. The aforementioned requirements are crucial for achieving trustworthiness [11]. For this purpose, the necessary latency resilience and data security requirements must be embedded in the physical domain. In this situation, the classical Shannon message transmission is limited and an ID scheme can achieve a better scaling behavior in terms of necessary energy and needed hardware components. It has been proved that information-theoretic security can be integrated into the ID scheme without paying an extra price for secrecy [12,13]. Further gains within the ID paradigm can be achieved by taking advantage of additional resources such as quantum entanglement, common randomness (CR), and feedback. In contrast to the classical Shannon message transmission, feedback can increase the ID capacity of a DMC [14]. Furthermore, it has been shown in [15] that the ID capacity of Gaussian channels with noiseless feedback is infinite. This holds to both rate definitions  $\frac{1}{n} \log M$  (as defined by Shannon for classical transmission) and  $\frac{1}{n} \log \log M$  (as defined by Ahlswede and Dueck for ID over DMCs). Interestingly, the authors of [15] showed that the ID capacity with noiseless feedback remains infinite regardless of the scaling used for the rate, e.g., double exponential, triple exponential, etc. In addition, the resource CR allows for a considerable increase in the ID capacity of channels [6,16,17]. The aforementioned communication scenarios emphasize that the ID capacity has completely different behavior than Shannon's capacity.

A key technology within 6G communication systems is the joint design of radio communication and sensor technology [11]. This enables the realization of revolutionary end-user applications [18]. Joint communication and radar/radio sensing (JCAS) means that sensing and communication are jointly designed based on sharing the same bandwidth. Sensing and communication systems are usually designed separately, meaning that resources are dedicated to either sensing or data communications. The joint sensing and communication approach is a solution that can overcome the limitations of separation-based approaches. Recent works [19–22] have explored JCAS and showed that this approach can improve spectrum efficiency while minimizing hardware costs. For instance, the fundamental limits of joint sensing and communication for a point-to-point channel were studied in [23]. In this case, the transmitter wishes to simultaneously send a message to the receiver and sense its channel state through a strictly causal feedback link. Motivated by the drastic effects of feedback on the ID capacity [15], this work investigates joint ID and sensing. To the best of our knowledge, the problem of joint ID and sensing has not been treated in the literature yet. We study the problem of joint ID and channel state estimation over a DMC with i.i.d. state sequences. The sender simultaneously sends an ID message over the DMC with a random state and estimates the channel state via a strictly causal channel output. The random channel state is available to neither the sender nor the receiver. We consider the ID capacity–distortion tradeoff as a performance metric. This metric is analogous to the one studied in [24], and is defined as the supremum of all achievable ID rates such that some distortion constraint on state sensing is fulfilled. This model was motivated by the problem of adaptive and sequential optimization of the beamforming vectors during the initial access phase of communication [25]. We establish a lower bound on the ID capacity–distortion tradeoff. In addition, we show that in our communication setup, sensing can be viewed as an additional resource that increases the ID capacity.

**Outline:**The remainder of this paper is organized as follows. Section 2 introduces the system models, reviews key definitions related to identification (ID), and presents the main results, including a complete characterization of the ID capacity–distortion function. Section 3 provides detailed proofs of these main results. In Section 4, we explore an alternative more flexible distortion constraint, namely, the average distortion, and establish a lower

bound on the corresponding ID capacity–distortion function. Finally, Section 5 concludes the paper with a discussion of the results and potential directions for future research.

Notation: The distribution of an RV  $X$  is denoted by  $P_X$ ; for a finite set  $\mathcal{X}$ , we denote the set of probability distributions on  $\mathcal{X}$  by  $\mathcal{P}(\mathcal{X})$  and the cardinality of  $\mathcal{X}$  by  $|\mathcal{X}|$ . If  $X$  is a RV with distribution  $P_X$ , we denote the Shannon entropy of  $X$  by  $H(P_X)$ , the expectation of  $X$  by  $\mathbb{E}(X)$ , and the variance of  $X$  by  $\text{Var}[X]$ . If  $X$  and  $Y$  are two RVs with probability distributions  $P_X$  and  $P_Y$ , then the mutual information between  $X$  and  $Y$  is denoted by  $I(X; Y)$ . Finally,  $\mathcal{X}^c$  denotes the complement of  $\mathcal{X}$ ,  $\mathcal{X} - \mathcal{Y}$  denotes the difference set, and all logarithms and information quantities are taken to base 2.

## 2. System Models and Main Results

Consider a discrete memoryless channel with random state  $(\mathcal{X} \times \mathcal{S}, W_S(y|x, s), \mathcal{Y})$  consisting of a finite input alphabet  $\mathcal{X}$ , finite output alphabet  $\mathcal{Y}$ , finite state set  $\mathcal{S}$ , and pmf  $W(y|x, s)$  on  $\mathcal{Y}$ . The channel is memoryless, i.e., if the input sequence  $x^n \in \mathcal{X}^n$  is sent and the sequence state is  $s^n \in \mathcal{S}^n$ , then the probability of a sequence  $y^n \in \mathcal{Y}^n$  being received is provided by

$$W_S^n(y^n|x^n, s^n) = \prod_{i=1}^n W_S(y_i|x_i, s_i). \tag{1}$$

The state sequence  $(S_1, S_2, \dots, S_n)$  is i.i.d. according to the distribution  $P_S$ . We assume that the input  $X_i$  and state  $S_i$  are statistically independent for all  $i \in \{1, 2, \dots, n\}$ . In our setup depicted in Figure 1, we assume that the channel state is known to neither the sender nor the receiver.

In the sequelae, we distinguish three scenarios:

1. Randomized ID over the state-dependent channel  $W_S$ , as depicted in Figure 1,
2. Deterministic or randomized ID over the state-dependent channel  $W_S$  in the presence of noiseless feedback between the sender and the receiver, as depicted in Figure 2,
3. Joint deterministic or randomized ID and sensing, in which the sender wishes to simultaneously send an identity to the receiver and sense the channel state sequence based on the output of the noiseless feedback link, as depicted in Figure 3.

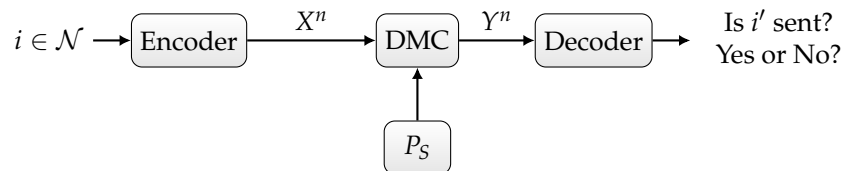


Figure 1. Discrete memoryless channel with random state.

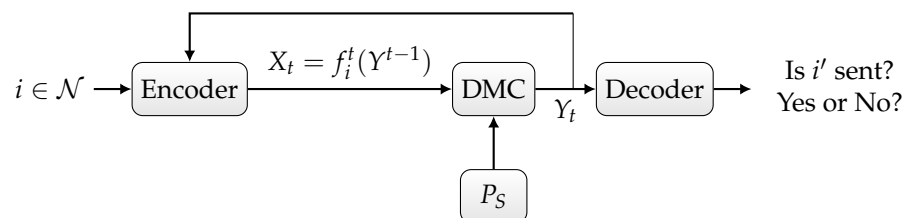


Figure 2. Discrete memoryless channel with random state and noiseless feedback.

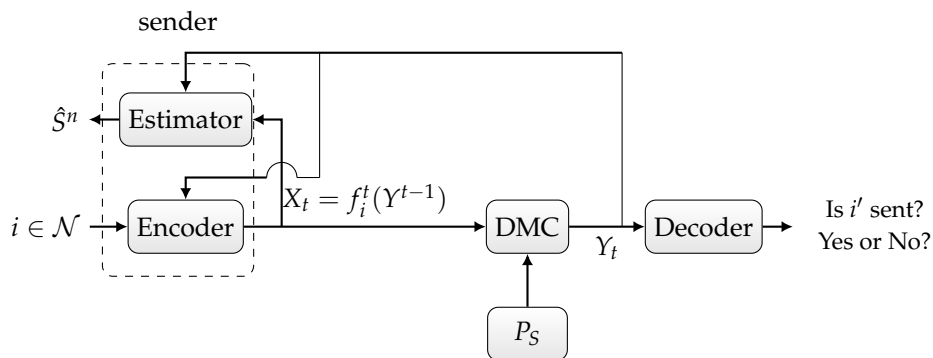


Figure 3. State-dependent channel with noiseless feedback.

First, we define randomized ID codes for the state-dependent channel defined above.

**Definition 1.** An  $(n, N, \lambda_1, \lambda_2)$  randomized ID code with  $\lambda_1 + \lambda_2 < 1$  for channel  $W_S$  is a family of pairs  $\{(Q(\cdot|i), \mathcal{D}_i(s^n))_{s^n \in \mathcal{S}^n}, i = 1, \dots, N\}$  with

$$Q(\cdot|i) \in \mathcal{P}(\mathcal{X}^n), \mathcal{D}_i(s^n) \in \mathcal{Y}^n, \forall s^n \in \mathcal{S}^n, \forall i = 1, \dots, N, \tag{2}$$

such that the errors of the first kind and the second kind are bounded as follows:

$$\sum_{s^n \in \mathcal{S}^n} P_S^n(s^n) \sum_{x^n \in \mathcal{X}^n} Q(x^n|i) W_S^n(\mathcal{D}_i(s^n)^c | x^n, s^n) \leq \lambda_1, \forall i, \tag{3}$$

$$\sum_{s^n \in \mathcal{S}^n} P_S^n(s^n) \sum_{x^n \in \mathcal{X}^n} Q(x^n|i) W_S^n(\mathcal{D}_j(s^n) | x^n, s^n) \leq \lambda_2, \forall i \neq j. \tag{4}$$

In the following, we define the achievable ID rate and ID capacity for our system model.

**Definition 2.**

1. The rate  $R$  of a randomized  $(n, N, \lambda_1, \lambda_2)$  ID code for the channel  $W_S$  is  $R = \frac{\log \log(N)}{n}$  bits.
2. The ID rate  $R$  for  $W_S$  is said to be achievable if for  $\lambda \in (0, \frac{1}{2})$  there exists an  $n_0(\lambda)$  such that for all  $n \geq n_0(\lambda)$  there exists an  $(n, 2^{2^{nR}}, \lambda, \lambda)$  randomized ID code for  $W_S$ .
3. The randomized ID capacity  $C_{ID}(W_S)$  of the channel  $W_S$  is the supremum of all achievable rates.

The following theorem characterizes the randomized ID capacity of the state-dependent channel  $W_S$  when the state information is known to neither the sender nor the receiver.

**Theorem 1.** The randomized ID capacity of the channel  $W_S$  is provided by

$$C_{ID}(W_S) = C(W_S) = \max_{P_X \in \mathcal{P}(\mathcal{X})} I(X; Y), \tag{5}$$

where  $C(W_S)$  denotes the Shannon transmission capacity of  $W_S$ .

**Proof.** The proof of Theorem 1 follows from Theorem 6.6.4 of [26] and Equation (7.2) of [27]. Because the channel  $W_S$  satisfies the strong converse property [26], the randomized ID capacity of  $W_S$  coincides with its Shannon transmission capacity determined in [27].  $\square$

Now, we consider the second scenario depicted in Figure 2. Let further  $\tilde{Y}^n = (\tilde{Y}_1, \dots, \tilde{Y}_n) \in \mathcal{Y}^n$  denote the output of the noiseless backward (feedback) channel:

$$\tilde{Y}_t = Y_t, \quad \forall t \in \{1, \dots, n\}. \tag{6}$$

In the following, we define a deterministic and randomized ID feedback code for the state-dependent channel  $W_S$ .

**Definition 3.** An  $(n, N, \lambda_1, \lambda_2)$  deterministic ID feedback code  $\{(f_i, \mathcal{D}_i(s^n))_{s^n \in \mathcal{S}^n}, i = 1, \dots, N\}$  with  $\lambda_1 + \lambda_2 < 1$  for the channel  $W_S$  is characterized as follows. The sender wants to send an ID message  $i \in \mathcal{N} := \{1, \dots, N\}$  that is encoded by the vector-valued function

$$f_i = [f_i^1, f_i^2, \dots, f_i^n], \tag{7}$$

where  $f_i^1 \in \mathcal{X}$  and  $f_i^t: \mathcal{Y}^{t-1} \rightarrow \mathcal{X}$  for  $t \in \{2, \dots, n\}$ . At  $t = 1$ , the sender sends  $f_i^1$ . At  $t \in \{2, \dots, n\}$ , the sender sends  $f_i^t(Y_1, \dots, Y^{t-1})$ . The decoding sets  $\mathcal{D}_i(s^n) \subset \mathcal{Y}^n, \forall i \in \{1, \dots, N\}$ , and  $\forall s^n \in \mathcal{S}^n$  should satisfy the following inequalities:

$$\sum_{s^n \in \mathcal{S}^n} P_S^n(s^n) W_S^n(\mathcal{D}_i(s^n)^c | f_i, s^n) \leq \lambda_1 \quad \forall i, \tag{8}$$

$$\sum_{s^n \in \mathcal{S}^n} P_S^n(s^n) W_S^n(\mathcal{D}_j(s^n) | f_i, s^n) \leq \lambda_2 \quad \forall i \neq j. \tag{9}$$

**Definition 4.** An  $(n, N, \lambda_1, \lambda_2)$  randomized ID feedback code

$$\{(Q_F(\cdot|i), \mathcal{D}_i(s^n))_{s^n \in \mathcal{S}^n}, i = 1, \dots, N\}$$

with  $\lambda_1 + \lambda_2 < 1$  for the channel  $W_S$  is characterized as follows. The sender wants to send an ID message  $i \in \mathcal{N} := \{1, \dots, N\}$  that is encoded by the probability distribution

$$Q_F(\cdot|i) \in \mathcal{P}(\mathcal{F}^n), \tag{10}$$

where  $Q_F(\cdot|i)$  denotes the set of all  $n$ -length functions  $f$  as  $\mathcal{F}^n$ . The decoding sets  $\mathcal{D}_i(s^n) \subset \mathcal{Y}^n, \forall i \in \{1, \dots, N\}$ , and  $\forall s^n \in \mathcal{S}^n$  should satisfy the following inequalities:

$$\sum_{s^n \in \mathcal{S}^n} P_S^n(s^n) \sum_{f \in \mathcal{F}^n} Q_F(f|i) W_S^n(\mathcal{D}_i(s^n)^c | f, s^n) \leq \lambda_1 \quad \forall i, \tag{11}$$

$$\sum_{s^n \in \mathcal{S}^n} P_S^n(s^n) \sum_{f \in \mathcal{F}^n} Q_F(f|i) W_S^n(\mathcal{D}_j(s^n) | f, s^n) \leq \lambda_2 \quad \forall i \neq j. \tag{12}$$

**Definition 5.**

1. The rate  $R$  of a (deterministic/randomized)  $(n, N, \lambda_1, \lambda_2)$  ID feedback code for the channel  $W_S$  is  $R = \frac{\log \log(N)}{n}$  bits.
2. The (deterministic/randomized) ID feedback rate  $R$  for  $W_S$  is said to be achievable if for  $\lambda \in (0, \frac{1}{2})$  there exists an  $n_0(\lambda)$  such that for all  $n \geq n_0(\lambda)$  there exists a (deterministic/randomized)  $(n, 2^{2^{nR}}, \lambda, \lambda)$  ID feedback code for  $W_S$ .
3. The (deterministic/randomized) ID feedback capacity  $C_{IDf}^d(W_S)/C_{IDf}^r(W_S)$  of the channel  $W_S$  is the supremum of all achievable rates.

It has been demonstrated in [26] that noise increases the ID capacity of the DMC in the case of feedback. Intuitively, noise is considered a source of randomness, i.e., a random experiment for which outcome is provided to the sender and receiver via the feedback channel. Thus, adding a perfect feedback link enables the realization of a correlated random experiment between the sender and the receiver. The size of this random experiment can be used to compute the growth of the ID rate. This result has been further emphasized in [15,28], where it has been shown that the ID capacity of the Gaussian channel with

noiseless feedback is infinite. This is because the authors of [15,28] provided a coding scheme that generates infinite common randomness between the sender and the receiver. Here, we want to investigate the effect of feedback on the ID capacity of the system model depicted in Figure 2. Theorem 2 characterizes the ID feedback capacity of the state-dependent channel  $W_S$  with noiseless feedback. The proof of Theorem 2 is provided in Section 3.

**Theorem 2.** *If  $C(W_S) > 0$ , then the deterministic ID feedback capacity of  $W_S$  is provided by*

$$C_{IDf}^d(W_S) = \max_{x \in \mathcal{X}} H(\mathbb{E}[W_S(\cdot|x, S)]). \tag{13}$$

**Theorem 3.** *If  $C(W_S) > 0$ , then the randomized ID feedback capacity of  $W_S$  is provided by*

$$C_{IDf}^r(W_S) = \max_{P \in \mathcal{P}(\mathcal{X})} H\left(\sum_{x \in \mathcal{X}} P(x) \mathbb{E}[W_S(\cdot|x, S)]\right). \tag{14}$$

**Remark 1.** *It can be shown that the same ID feedback capacity formula holds if the channel state is known to either the sender or the receiver. This is because we achieve the same amount of common randomness as in the scenario depicted in Figure 2. Intuitively, the channel state is an additional source of randomness that we can take advantage of.*

Now, we consider the third scenario depicted in Figure 3, where we want to jointly identify and sense the channel state. The sender comprises an encoder that sends a symbol  $x_t = f_i^t(y^{t-1})$  for each identity  $i \in \{1, \dots, N\}$  and delayed feedback output  $y^{t-1} \in \mathcal{Y}^{t-1}$  along with a state estimator that outputs an estimation sequence  $\hat{s}^n \in \mathcal{S}^n$  based on the feedback output and input sequence. We define the per-symbol distortion as follows:

$$d_t = \mathbb{E}[d(S_t, \hat{S}_t)], \tag{15}$$

where  $d: \mathcal{S} \times \mathcal{S} \rightarrow [0, +\infty)$  is a distortion function and the expectation is over the joint distribution of  $(S_t, \hat{S}_t)$  conditioned by the ID message  $i \in \{1, \dots, N\}$ .

**Definition 6.**

1. *An ID rate–distortion pair  $(R, D)$  for  $W_S$  is said to be achievable if for every  $\lambda \in (0, \frac{1}{2})$  there exists an  $n_0(\lambda)$  such that for all  $n \geq n_0(\lambda)$  there exists an  $(n, 2^{2nR}, \lambda, \lambda)$  (deterministic/randomized) ID code for  $W_S$  and if  $d_t \leq D$  for all  $t = 1, \dots, n$ .*
2. *The deterministic ID capacity–distortion function  $C_{ID}^d(D)$  is defined as the supremum of  $R$  such that  $(R, D)$  is achievable.*

Without loss of generality, we choose the following deterministic estimation function  $h^*$ :

$$\hat{s} = h^*(x, y) = \min_{h: \mathcal{X} \times \mathcal{Y} \rightarrow \mathcal{S}} \mathbb{E}[d(S, h(X, Y)) | X = x, Y = y], \tag{16}$$

where  $h: \mathcal{X} \times \mathcal{Y} \rightarrow \mathcal{S}$  is an estimator that maps a channel input–feedback output pair to a channel state. If there exist several functions  $h^*(\cdot, \cdot)$ , we choose one randomly. We define the minimal distortion function for each input symbol  $x \in \mathcal{X}$  as in [29]:

$$d^*(x) = \mathbb{E}_{SY}[d(S, h^*(X, Y)) | X = x] \tag{17}$$

and the minimal distortion function for each input distribution  $P \in \mathcal{P}(\mathcal{X})$  as

$$d^*(P) = \sum_{x \in \mathcal{X}} d^*(x). \tag{18}$$

In the following, we establish the ID capacity–distortion function defined above.

**Theorem 4.** *The deterministic ID capacity–distortion function of the state-dependent channel  $W_S$  depicted in Figure 3 is provided by*

$$C_{ID}^d(D) = \max_{x \in \mathcal{X}_D} H(\mathbb{E}[W_S(\cdot|x, S)]), \tag{19}$$

where the set  $\mathcal{X}_D$  is provided by

$$\mathcal{X}_D = \{x \in \mathcal{X}, \quad d^*(x) \leq D\}. \tag{20}$$

We now turn our attention to a randomized encoder. In the following, we derive the ID capacity–distortion function of the state-dependent channel  $W_S$  under the assumption of randomized encoding.

**Theorem 5.** *The randomized ID capacity–distortion function of the state-dependent channel  $W_S$  is provided by*

$$C_{ID}^r(D) = \max_{P \in \mathcal{P}_D} H\left(\sum_{x \in \mathcal{X}} P(x) \mathbb{E}[W_S(\cdot|x, S)]\right), \tag{21}$$

where the set  $\mathcal{P}_D$  is provided by

$$\mathcal{P}_D = \{P \in \mathcal{P}(\mathcal{X}), \quad d^*(P) \leq D\}. \tag{22}$$

**Remark 2.** *Randomized encoding achieves higher rates than deterministic encoding. This is because we are combining two sources of randomness: local randomness used for the encoding, and shared randomness generated via the noiseless feedback link. The result is analogous to randomized ID over DMCs in the presence of noiseless feedback, as studied in [14].*

### 3. Proof of the Main Results

In this section, we provide the proofs of Theorems 2–5.

#### 3.1. Direct Proof of Theorem 2

**Proof.** We consider an average channel  $W_{S,\text{avg}}$  provided by

$$W_{S,\text{avg}}(y|x) = \sum_{s \in \mathcal{S}} P_S(s) W_S(y|x, s), \quad \forall x \in \mathcal{X}, \quad \forall y \in \mathcal{Y}. \tag{23}$$

The DMC  $W_{S,\text{avg}}$  is obtained by averaging the DMCs  $W_S$  over the state. Now, it suffices to show that  $R = \max_{x \in \mathcal{X}} H(\mathbb{E}[W_S(\cdot|x, S)])$  is an achievable ID feedback rate for the average channel  $W_{S,\text{avg}}$ . The deterministic ID feedback capacity of the average channel  $C_{IDf}^d(W_{S,\text{avg}})$



can be determined by applying Theorem 1 of [28] on  $W_{S,\text{avg}}$ . If the transmission capacity  $C(W_{S,\text{avg}})$  of  $W_{S,\text{avg}}$  is positive, we have

$$C_{IDf}^d(W_{S,\text{avg}}) \leq \max_{x \in \mathcal{X}} H(W_{S,\text{avg}}(\cdot|x)) \tag{24}$$

$$= \max_{x \in \mathcal{X}} H(\mathbb{E}[W_S(\cdot|x, S)]). \tag{25}$$

This completes the direct proof of Theorem 2.  $\square$

### 3.2. Converse Proof of Theorem 2

**Proof.** For the converse proof, we use the techniques of [14] for deterministic ID over DMCs with noiseless feedback. We first extend Lemma 3 of [14] (image size for a deterministic feedback strategy) to the deterministic ID feedback code for  $W_S$  described in Definition 3.  $\square$

**Lemma 1.** For any feedback strategy  $f = [f^1, f^2, \dots, f^n]$  and any  $\mu \in (0, 1)$ , we have

$$\min_{\mathcal{E}_1 \subset \mathcal{Y}^n: \mathbb{E}_{S^n} [W_S^n(\mathcal{E}_1|f, S^n)] \geq 1-\mu} |\mathcal{E}_1| \leq K_1, \tag{26}$$

where  $K_1$  is provided by

$$K_1 = 2^{n \max_{x \in \mathcal{X}} H(\mathbb{E}[W_S(\cdot|x)]) + \alpha\sqrt{n}} = 2^{nH(\mathbb{E}[W_S(\cdot|x^*)]) + \alpha\sqrt{n}}, \tag{27}$$

where

$$\alpha = \sqrt{\frac{\beta}{\mu}}, \quad \beta = \max(\log^2(3), \log^2(|\mathcal{Y}|)). \tag{28}$$

**Proof.** We use a similar idea as for the proof of Lemma 3 of [14]. Let  $\mathcal{E}_1^* \subset \mathcal{Y}^n$  be defined as follows:

$$\mathcal{E}_1^* = \left\{ y^n \in \mathcal{Y}^n, \quad -\log \mathbb{E}_{S^n} [W_S^n(y^n|f, S^n)] \leq \log K_1 \right\}. \tag{29}$$

It then follows from the definition of  $\mathcal{E}_1^*$  that  $|\mathcal{E}_1^*| \leq K_1$ . It remains to show that

$$\mathbb{E}_{S^n} [W_S^n(y^n|f, S^n)] \geq 1 - \mu.$$

For  $t = 1, 2, \dots, n$  and a fixed feedback strategy  $f = [f^1, \dots, f^t] \in \mathcal{X}$ , we have

$$\Pr\{Y^t = y^t\} = \sum_{s^n \in \mathcal{S}^n} P_S^n(s^n) W^n(y^t|f, s^n) \tag{30}$$

$$= \mathbb{E}_{S^n} [W_S^n(y^t|f, S^n)], \quad \forall y^t \in \mathcal{Y}^t. \tag{31}$$

Let the RV  $Z_t$  be defined as follows:

$$Z_t = -\log \mathbb{E}_{S_t} [W_S(Y_t|f^t(Y^{(t-1)}), S_t)]. \tag{32}$$



We have

$$\mathbb{E}_{S^n} [W_S^n(\mathcal{E}_1^* | f, S^n)] \tag{33}$$

$$= \Pr \left\{ -\log \mathbb{E}_{S^n} [W_S^n(y^n | f, S^n)] \leq \log K_1 \right\} \tag{34}$$

$$= \Pr \left\{ -\log \left( \sum_{s^n \in \mathcal{S}^n} P_S^n(s^n) \prod_{t=1}^n W_S(y_t | f^t, s_t) \right) \leq \log K_1 \right\} \tag{35}$$

$$= \Pr \left\{ -\log \left( \sum_{s_1 \in \mathcal{S}} \sum_{s_2 \in \mathcal{S}} \cdots \sum_{s_n \in \mathcal{S}} \prod_{t=1}^n P_S(s_t) W_S(y_t | f^t, s_t) \right) \leq \log K_1 \right\} \tag{36}$$

$$= \Pr \left\{ -\log \left( \sum_{s_1 \in \mathcal{S}} P_S(s_1) W_S(y_1 | f^1, s_1) \right. \right. \tag{37}$$

$$\left. \cdot \sum_{s_2 \in \mathcal{S}} P_S(s_2) W_S(y_2 | f^2, s_2) \cdots \sum_{s_n \in \mathcal{S}} P_S(s_n) W_S(y_n | f^n, s_n) \right) \leq \log K_1 \left. \right\} \tag{38}$$

$$= \Pr \left\{ \sum_{t=1}^n -\log \mathbb{E}_{S_t} [W_S(Y_t | f^t(Y^{t-1}), S_t)] \leq \log K_1 \right\} \tag{39}$$

$$= \Pr \left\{ \sum_{t=1}^n Z_t \leq \log K_1 \right\}. \tag{40}$$

Now, we want to establish a lower bound on  $\mathbb{E}_{S^n} [W_S^n(\mathcal{E}^* | f, S^n)]$ . It suffices to find a lower bound on the expression in (40). Let the RV  $U_t$  be defined as follows:

$$U_t = Z_t - \mathbb{E}[Z_t | Y^{t-1}], \quad t \in \{1, 2, \dots, n\}. \tag{41}$$

It can be shown that

$$\mathbb{E}[U_t | Y^{t-1}] = \mathbb{E}_Y [Z_t - \mathbb{E}_Y [Z_t | Y^{t-1}] | Y^{t-1}] \tag{42}$$

$$= \mathbb{E}_Y [Z_t | Y^{t-1}] - \mathbb{E}_Y [Z_t | Y^{t-1}] \tag{43}$$

$$= 0. \tag{44}$$

Furthermore, we have

$$\mathbb{E}[U_t] = \mathbb{E}_Y [Z_t - \mathbb{E}_Y [Z_t | Y^{t-1}]] \tag{45}$$

$$= \mathbb{E}_Y [Z_t] - \mathbb{E}_Y [\mathbb{E}_Y [Z_t | Y^{t-1}]] \tag{46}$$

$$= \mathbb{E}_Y [Z_t] - \mathbb{E}_Y [Z_t] \tag{47}$$

$$= 0. \tag{48}$$

It can be shown that for  $y^{t-1} \in \mathcal{Y}^{t-1}$ ,

$$\mathbb{E}[Z_t | y^{t-1}] = \sum_{y_t \in \mathcal{Y}} \left( -\mathbb{E}_{S_t} [W_S(y_t | f^t(y^{t-1}), S_t)] \right. \tag{49}$$

$$\left. \log \mathbb{E}_{S_t} [W_S(y_t | f^t(y^{t-1}), S_t)] \right) \tag{50}$$

$$\leq \max_{x \in \mathcal{X}} H \left( \mathbb{E}_S [W_S(\cdot | x, S)] \right) \tag{51}$$

$$= H \left( \mathbb{E}_S [W_S(\cdot | x^*, S)] \right). \tag{52}$$

It follows from the definition of the RV  $U_t$  in (41) that

$$\mathbb{E}_{S^n} [W_S^n(\mathcal{E}^* | \mathbf{f}, S^n)] = \Pr \left\{ \sum_{t=1}^n (U_t + \mathbb{E}[Z_t | Y^{t-1}]) \leq \log K_1 \right\} \tag{53}$$

$$= \Pr \left\{ \sum_{t=1}^n U_t \leq \log K_1 - \sum_{t=1}^n \mathbb{E}[Z_t | Y^{t-1}] \right\} \tag{54}$$

$$\stackrel{(a)}{=} \Pr \left\{ \sum_{t=1}^n U_t \leq nH(\mathbb{E}_S [W_S(\cdot | x^*, S)]) \right. \tag{55}$$

$$\left. + \alpha\sqrt{n} - \sum_{t=1}^n \mathbb{E}[Z_t | Y^{t-1}] \right\} \tag{56}$$

$$\stackrel{(b)}{\geq} \Pr \left\{ \sum_{t=1}^n U_t \leq \alpha\sqrt{n} \right\}, \tag{57}$$

where (a) follows from the definition of  $K_1$  in (27) and (b) follows from (52).

It can be verified that

$$\text{Var}[U_t] \leq \beta, \quad \forall t = 1, 2, \dots, n. \tag{58}$$

Therefore, we can apply Chebyshev’s inequality and obtain

$$\mathbb{E}_{S^n} [W_S^n(\mathcal{E}^* | \mathbf{f}, S^n)] \geq \Pr \left\{ \sum_{t=1}^n U_t \leq \alpha\sqrt{n} \right\} \tag{59}$$

$$\stackrel{(a)}{\geq} 1 - \mu, \tag{60}$$

where (a) follows the definition of  $\beta$  in (28). This completes the proof of Lemma 1.  $\square$

We establish an upper bound on the deterministic ID feedback rate for the channel model  $W_S$  using Lemma 1. Let  $\{(\mathbf{f}_i, \mathcal{D}_i(s^n))_{s^n \in \mathcal{S}^n}, i = 1, \dots, N\}$  be an  $(n, N, \lambda, \lambda)$  deterministic ID feedback code for channel  $W_S$  with  $\lambda \in (0, \frac{1}{2})$ , and let  $\mu \in (0, 1)$  be chosen such that

$$1 - \mu - \lambda < \frac{1}{2}. \tag{61}$$

For each feedback strategy  $\mathbf{f}_i$ , we define the set  $\mathcal{E}_i$  that satisfies (26). For  $i \in \{1, 2, \dots, N\}$ , we have

$$\mathbb{E}_{S^n} [W_S^n(\mathcal{D}_i(s^n) \cap \mathcal{E}_i | \mathbf{f}_i, S^n)] = 1 - \mathbb{E}_{S^n} [W_S^n((\mathcal{D}_i(s^n) \cap \mathcal{E}_i)^c | \mathbf{f}_i, S^n)] \tag{62}$$

$$= 1 - \mathbb{E}_{S^n} [W_S^n((\mathcal{D}_i(s^n))^c \cup \mathcal{E}_i^c | \mathbf{f}_i, S^n)] \tag{63}$$

$$\stackrel{(a)}{\geq} 1 - \mathbb{E}_{S^n} [W_S^n((\mathcal{D}_i(s^n))^c | \mathbf{f}_i, S^n)] - \mathbb{E}_{S^n} [W_S^n(\mathcal{E}_i^c | \mathbf{f}_i, S^n)] \tag{64}$$

$$\stackrel{(b)}{\geq} 1 - \lambda - \mu \tag{65}$$

$$\stackrel{(c)}{>} \frac{1}{2}, \tag{66}$$

where (a) follows from the union bound, (b) follows from the definition of the  $(n, N, \lambda, \lambda)$  ID feedback code and from (26), and (c) follows from (61). Similarly, from the definition of the ID feedback code

$$\{(\mathbf{f}_i, \mathcal{D}_i(s^n))_{s^n \in \mathcal{S}^n}, i = 1, \dots, N\},$$

from (26) and (61), for  $i \in \{1, 2, \dots, N\}$  and  $i \neq j$  it follows that

$$\mathbb{E}_{S^n} [W_S^n(\mathcal{D}_j(s^n) \cap \mathcal{E}_i | \mathbf{f}_i, S^n)] < \frac{1}{2}. \tag{67}$$

As the error of the second kind  $\lambda$  is smaller than  $\frac{1}{2}$ , all the sets  $\mathcal{D}_i(s^n) \cap \mathcal{E}_i$  are distinct. Therefore, any  $(n, N, \lambda, \lambda)$  deterministic ID feedback code  $\{(f_i, \mathcal{D}_i(s^n))_{s^n \in \mathcal{S}^n}, i = 1, \dots, N\}$  for channel  $W_S$  with  $\lambda \in (0, \frac{1}{2})$  has an associated  $(n, N, \lambda', \lambda')$  deterministic ID feedback code  $\{(f_i, \mathcal{D}_i(s^n))_{s^n \in \mathcal{S}^n} \cap \mathcal{E}_i, i = 1, \dots, N\}$ , where  $\lambda' \in (0, \frac{1}{2})$  and the set  $\mathcal{E}_i$  satisfies (26)  $\forall i = 1, \dots, N$ . Thus, per Lemma 1, the cardinality  $N$  of the deterministic ID feedback code is upper-bounded as follows:

$$N \leq \sum_{k=0}^{K_1} \binom{|\mathcal{Y}|^n}{k} \leq (|\mathcal{Y}|^n)^{K_1} \tag{68}$$

$$= 2^n \log |\mathcal{Y}| 2^{nH(\mathbb{E}[W_S(\cdot|x^*, S)]) + \alpha\sqrt{n}}. \tag{69}$$

This completes the converse proof of Theorem 2.

### 3.3. Direct Proof of Theorem 3

**Proof.** Similarly, we can consider the average channel  $W_{S,avg}$  defined in (23). It is sufficient to show that  $R = \max_{P \in \mathcal{P}(\mathcal{X})} H(\sum_{x \in \mathcal{X}} P(x) W_{S,avg}(\cdot|x))$  is an achievable randomized ID feedback rate for the average channel  $W_{S,avg}$ . The randomized ID feedback capacity of the average channel  $C_{IDf}^r(W_{S,avg})$  can be obtained by applying Theorem 2 of [28] on  $W_{S,avg}$ . If the transmission capacity  $C(W_{S,avg})$  of  $W_{S,avg}$  is positive, then we have

$$C_{IDf}^r(W_{S,avg}) \leq \max_{P \in \mathcal{P}(\mathcal{X})} H\left(\sum_{x \in \mathcal{X}} P(x) W_{S,avg}(\cdot|x)\right). \tag{70}$$

This completes the direct proof of Theorem 3.  $\square$

### 3.4. Converse Proof of Theorem 3

**Proof.** We first extend Lemma 4 of [14] (image size for a randomized feedback strategy) to the randomized ID feedback code for  $W_S$ .  $\square$

**Lemma 2.** For any randomized feedback strategy  $Q_F(\cdot)$  over all  $n$ -length feedback encoding sets  $\mathcal{F}_n$  and any  $\mu \in (0, 1)$ ,

$$\min_{\mathcal{E}_2 \subset \mathcal{Y}^n: \sum_{f \in \mathcal{F}_n} Q_F(f) \mathbb{E}_{S^n} [W_S^n(\mathcal{E}_2|f, S^n)] \geq 1 - \mu} |\mathcal{E}_2| \leq K_2, \tag{71}$$

where  $K_2$  is provided by

$$K_2 = 2^{n \max_{P \in \mathcal{P}(\mathcal{X})} H(\sum_{x \in \mathcal{X}} P(x) \mathbb{E}[W_S(\cdot|x, S)]) + \alpha\sqrt{n}} = 2^{nH(\sum_{x \in \mathcal{X}} P^*(x) \mathbb{E}[W_S(\cdot|x, S)]) + \alpha\sqrt{n}}, \tag{72}$$

where  $\alpha = \sqrt{\frac{\beta}{\mu}}$ ,  $\beta = \max \log^2 3, \log^2 |\mathcal{Y}|$ .

**Proof.** We use a similar idea as for the proof of Lemma 4 in [14]. We define the set  $\mathcal{E}_2^* \subset \mathcal{Y}^n$  as follows:

$$\mathcal{E}_2^* = \left\{ y^n \in \mathcal{Y}^n, -\log \sum_{f \in \mathcal{F}_n} Q_F(f) \mathbb{E}_{S^n} [W_S(y^n|f, S^n)] \leq \log K_2 \right\}. \tag{73}$$

From the definition of  $\mathcal{E}_2^*$ , we have

$$|\mathcal{E}_2^*| \leq K_2.$$

$\square$

Then, it is sufficient to show that  $\sum_{f \in \mathcal{F}_n} Q_F(f) \mathbb{E}_{S^n} [W_S^n(\mathcal{E}_2 | f, S^n)] \geq 1 - \mu$ . Let the RV  $Z_t$  be defined as follows:

$$Z_t = -\log \left( \sum_{f^t \in \mathcal{F}^t} Q^t(f^t) \mathbb{E}_{S_t} [W_S^n(Y_t | f^t(Y^{t-1}), S_t)] \right) \tag{74}$$

where  $Q_F(f) = \prod_{t=1}^n Q^t(f^t)$  and  $\mathcal{F}^t$  is the set of all mapping  $\mathcal{Y}^{t-1} \mapsto \mathcal{X}$ . We have

$$\sum_{f \in \mathcal{F}_n} Q_F(f) \mathbb{E}_{S^n} [W_S^n(\mathcal{E}_2 | f, S^n)] \tag{75}$$

$$= Pr \left\{ -\log \left( \sum_{f \in \mathcal{F}_n} Q_F(f) \mathbb{E}_{S^n} [W_S(y^n | f, S^n)] \right) \leq \log K_2 \right\} \tag{76}$$

$$= Pr \left\{ -\log \left( \sum_{f \in \mathcal{F}_n} Q_F(f) \prod_{t=1}^n \mathbb{E}_{S_t} [W_S(Y_t | f^t(Y^{t-1}), S_t)] \right) \leq \log K_2 \right\} \tag{77}$$

$$= Pr \left\{ -\log \left( \sum_{f \in \mathcal{F}_n} \prod_{t=1}^n Q^t(f^t) \prod_{t=1}^n \mathbb{E}_{S_t} [W_S(Y_t | f^t(Y^{t-1}), S_t)] \right) \leq \log K_2 \right\} \tag{78}$$

$$= Pr \left\{ -\log \left( \prod_{t=1}^n \sum_{f^t \in \mathcal{F}^t} Q^t(f^t) \mathbb{E}_{S_t} [W_S(Y_t | f^t(Y^{t-1}), S_t)] \right) \leq \log K_2 \right\} \tag{79}$$

$$= Pr \left\{ \sum_{t=1}^n Z_t \leq \log K_2 \right\}. \tag{80}$$

Now, for any  $y^{t-1} \in \mathcal{Y}^{t-1}$ , we consider

$$\mathbb{E} [Z_t | y^{t-1}] \tag{81}$$

$$= \sum_{y_t \in \mathcal{Y}} \left( - \sum_{f^t \in \mathcal{F}^t} Q^t(f^t) \mathbb{E}_{S_t} [W_S^n(y_t | f^t(y^{t-1}), S_t)] \log \left( \sum_{f^t \in \mathcal{F}^t} Q^t(f^t) \mathbb{E}_{S_t} [W_S^n(y_t | f^t(y^{t-1}), S_t)] \right) \right) \tag{82}$$

$$\leq H \left( \sum_{x \in \mathcal{X}} P^*(x) \mathbb{E} [W_S^n(\cdot | x, S)] \right). \tag{83}$$

Similarly, for all  $t \in \{1, 2, \dots, n\}$ , we define the RV  $U_t = Z_t - \mathbb{E} [Z_t | Y^{t-1}]$ . It has been shown in (44) and (48) that  $\mathbb{E} [U_t | Y^{t-1}] = 0$  and  $\mathbb{E} [U_t] = 0$ . Combining (80) and (83), we have

$$\sum_{f \in \mathcal{F}_n} Q_F(f) \mathbb{E}_{S^n} [W_S^n(\mathcal{E}_2 | f, S^n)] \tag{84}$$

$$= Pr \left\{ \sum_{t=1}^n (U_t + \mathbb{E} [Z_t | Y^{t-1}]) \leq \log K_2 \right\} \tag{85}$$

$$= Pr \left\{ \sum_{t=1}^n U_t \leq nH \left( \sum_{x \in \mathcal{X}} P^*(x) \mathbb{E} [W_S(\cdot | x, S)] \right) + \alpha \sqrt{n} - \sum_{t=1}^n \mathbb{E} [Z_t | Y^{t-1}] \right\} \tag{86}$$

$$\geq Pr \left\{ \sum_{t=1}^n U_t \leq \alpha \sqrt{n} \right\}. \tag{87}$$

Assuming  $Var[U_t] \leq \beta$  for all  $t = 1, 2, \dots, n$ , we can apply Chebyshev’s inequality to obtain

$$\sum_{f \in \mathcal{F}_n} Q_F(f) \mathbb{E}_{S^n} [W_S^n(\mathcal{E}_2 | f, S^n)] \geq 1 - \frac{\beta}{\alpha^2} \tag{88}$$

$$= 1 - \mu. \tag{89}$$

Replacing  $K_1$  in the converse proof of Theorem 2 with the corresponding  $K_2$  as outlined in Lemma 2 completes the converse proof of Theorem 3.

### 3.5. Direct Proof of Theorem 4

#### 3.5.1. Coding Scheme

**Proof.** To some extent, we use the same coding scheme elaborated in [14]. We choose the blocklength as  $m = n + \lceil \sqrt{n} \rceil$ . Let  $x^*$  be some symbol in  $\mathcal{X}_D$ . Regardless of which identity  $i \in \{1, \dots, N\}$  we want to identify, the sender first sends the sequence  $x^{*n} = (x^*, x^*, \dots, x^*) \in \mathcal{X}_D^n$  over the state-dependent channel  $W_S$ . The received sequence  $y^n \in \mathcal{Y}^n$  becomes known to the sender (estimator and encoder) via the noiseless feedback link. The feedback provides the sender and receiver with knowledge of the outcome of the correlated random experiment  $(\mathcal{Y}^n, \mathbb{E}_{S^n} [W_S^n(\cdot | x^{*n}, S^n)])$ .  $\square$

#### 3.5.2. Common Randomness Generation

We want to generate uniform common randomness, as it is the most convenient form of common randomness [17]; therefore, we convert our correlated random experiment  $(\mathcal{Y}^n, \mathbb{E}_{S^n} [W_S^n(\cdot | x^{*n}, S^n)])$  to a uniform one  $(\mathcal{T}^n, \mathbb{E}_{S^n} [W_S^n(\cdot | x^{*n}, S^n)])$ . For  $\epsilon > 0$ , the set  $\mathcal{T}^n$  is provided by

$$\mathcal{T}^n = \bigcup_{V_S^a: \|V_S^a - W_S^a\| \leq \epsilon} \mathcal{T}_{V_S^a}^n(x^{*n}), \tag{90}$$

where  $\|V_S^a - W_S^a\|$  is defined as follows.

**Definition 7.** Let  $\mathcal{W}$  be the set of stochastic matrices  $W: \mathcal{X} \rightarrow \mathcal{Y}$ . Let  $W \in \mathcal{W}$  such that

$$P_{XY}(x, y) = P_X(x)W(y|x), \quad \forall x \in \mathcal{X}, \forall y \in \mathcal{Y}. \tag{91}$$

For  $V, V' \in \mathcal{W}$ , the distance  $|V - V'|$  is defined as

$$\|V - V'\| = \max_{x \in \mathcal{X}, y \in \mathcal{Y}} |V(y|x) - V'(y|x)|. \tag{92}$$

Here,  $V_S^a$  denotes the average channel defined by

$$V_S^a = \sum_{s \in \mathcal{S}} P_S(s) V_S(\cdot | \cdot, s). \tag{93}$$

We introduce the following lemmas.

**Lemma 3 ([30]).** Let  $(x^n, y^n)$  be emitted by the DMS  $P_{XY}(\cdot, \cdot) = W(\cdot | \cdot)P_X(\cdot)$  and let  $V \in \mathcal{W}$  such that  $|V - W| \leq \epsilon$ . Then, for every  $\epsilon > 0$  there exist  $\delta' > 0$  and  $n_0(\epsilon)$  such that for  $n \geq n_0(\epsilon)$  we have

$$\sum_{y^n \in \mathcal{T}_V^n(x^n)} W^n(y^n | x^n) \geq 1 - 2^{-n\delta'}. \tag{94}$$

**Lemma 4** ([30]). Let  $(x^n, y^n)$  be emitted by the DMS  $P_{XY}(\cdot, \cdot) = W(\cdot|\cdot)P_X(\cdot)$  and let  $V \in \mathcal{W}$ . For every  $\epsilon > 0$ , there exist a  $c(\epsilon) > 0$  and  $n_0(\epsilon)$  such that for  $n \geq n_0(\epsilon)$  we have

1.  $|\bigcup_{V: \|V-W\| \leq \epsilon} \mathcal{T}_V^n(x^n)| \geq 2^{n(H(W|P_X)-c(\epsilon))}$ ,
2.  $|\bigcup_{V: \|V-W\| \leq \epsilon} \mathcal{T}_V^n(x^n)| \leq 2^{n(H(W|P_X)+c(\epsilon))}$ ,
3. If  $\|V - W\| \leq \epsilon$ ,  $\mathcal{T}_V^n(x^n) \neq \emptyset$  and  $c(\epsilon) \rightarrow 0$  if  $\epsilon \rightarrow 0$ , then
 
$$|\mathcal{T}_V^n(x^n)| \geq 2^{n(H(W|P_X)-c(\epsilon))}. \tag{95}$$

**Lemma 5** ([31]). Let  $\{X_i\}$  be i.i.d. RVs taking values in  $[0, 1]$  with mean  $\mu$ ; then,  $\forall c > 0$  with  $p = \mu + c \leq 1$  we have

$$\Pr\{\bar{X}_n - \mu \geq c\} \leq \exp(-nD(p||\mu)) \leq \exp(-2nc^2), \tag{96}$$

where  $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ .

For arbitrary  $D_{min} \leq D_1 \leq D_2$ , we define  $\mathcal{X}_{D_1}$  and  $\mathcal{X}_{D_2}$  by

$$\mathcal{X}_{D_1} = \{x \in \mathcal{X}, d^*(x) \leq D_1\}, \tag{97}$$

$$\mathcal{X}_{D_2} = \{x \in \mathcal{X}, d^*(x) \leq D_2\}. \tag{98}$$

It is clear that  $g(D)$  is a non-decreasing function, because  $\mathcal{X}_{D_1} \subseteq \mathcal{X}_{D_2}$  for arbitrary  $D_1 \leq D_2$ . Letting  $\mu \in (0, 1)$ , we have

$$g(\mu D_1 + (1 - \mu)D_2) = \max_{x \in \mathcal{X}_{\mu D_1 + (1-\mu)D_2}} H(\mathbb{E}[W_S(\cdot|x, S)]). \tag{99}$$

It follows from Lemma 4 that  $\mathbb{E}_{S^n} [W_S^n(\cdot|x^{*n}, S^n)]$  is essentially uniform on  $\mathcal{T}^n$ . Let the set  $\mathcal{Y}^n - \mathcal{T}^n$  be denoted by  $\mathcal{E}^*$ . Per Lemma 3, we have

$$\Pr\{Y^n \in \mathcal{E}^* | X^n = (x^{*n}, x^{*n}, \dots, x^{*n})\} \tag{100}$$

$$= 1 - \Pr\{Y^n \notin \mathcal{T}^n | X^n = (x^{*n}, x^{*n}, \dots, x^{*n})\} \tag{101}$$

$$\leq 2^{-n\delta'}. \tag{102}$$

As mentioned earlier, we have  $|\mathcal{T}^n| \approx 2^{n\mathbb{E}_S [W_S(\cdot|x^{*n}, S^n)]}$ . This quantity is the size of the correlated random experiment  $(\mathcal{T}^n, \mathbb{E}_{S^n} [W_S^n(\cdot|x^{*n}, S^n)])$ , which determines the growth of the ID rate. Let  $\mathcal{C} = \{(u_j, \mathcal{D}_j), j = 1, \dots, M\}$  be an  $(\lceil \sqrt{n} \rceil, M, 2^{-\sqrt{n}\delta})$  code, where  $u_j \in \mathcal{X}_D^{\lceil \sqrt{n} \rceil}$  for each  $j = 1, \dots, M$ . We concatenate the sequence  $x^{*n} = (x^*, x^*, \dots, x^*) \in \mathcal{X}_D^n$  and the transmission code  $\mathcal{C}$  to build an  $(m, N, \lambda_1, \lambda_2)$  ID feedback code  $\mathcal{C}' = \{(f_i, \mathcal{D}'_i), i = 1, \dots, N\}$  for  $W_S$ . We now have  $\lambda_1, \lambda_2 < \lambda, \lambda \in (0, \frac{1}{2})$ . The concatenation is performed using the coloring functions  $\{F_i, i = 1, \dots, N\}$ . We choose a suitable set of coloring functions  $\{F_i, i = 1, \dots, N\}$  at random. Every coloring function  $F_i: \mathcal{T}^n \rightarrow \{1, \dots, M\}$  corresponds to an ID message  $i$  and maps each element  $y^n \in \mathcal{T}^n$  to an element  $F_i(y^n)$  in a smaller set  $\{1, \dots, M\}$ . After  $y^n \in \mathcal{T}^n$  has been received by the sender (encoder and estimator) via the noiseless feedback channel, if  $i \in \{1, \dots, N\}$  is available, then the encoder sends  $u_{F_i(y^n)}$ . Note that we define an encoding strategy  $f_i = [f_i^1, \dots, f_i^m] \in \mathcal{F}_m$  for each coloring function  $F_i$ , as presented in Definition 4. If  $y^n \notin \mathcal{T}^n$ , then an error is declared. This error probability goes to zero as  $n$  goes to infinity, as computed in (102). For a fixed

family of maps  $\{F_i, i = 1, \dots, N\}$  and for each  $i \in \{1, \dots, N\}$ , we define the decoding sets  $\mathcal{D}(F_i) = \cup_{y^n \in \mathcal{T}^n} \{y^n\} \times \mathcal{D}_{F_i}(y^n)$ .

### 3.5.3. Error Analysis

Next, we analyze the maximal error performance of the deterministic ID feedback code. For our analysis of the error of the first kind, we choose a fixed set  $\{F_i, i = 1, \dots, N\}$ . The error of the first kind is upper-bounded by

$$\mathbb{E}_{S^m} [W_S^m(\mathcal{D}'_i | \mathbf{f}_i, S^m)] \tag{103}$$

$$= \mathbb{E}_{S^m} [W_S^m(\mathcal{D}(F_i)^c | \mathbf{f}_i, S^m)] \tag{104}$$

$$\stackrel{(a)}{\leq} \sum_{s^m \in \mathcal{S}^m} P_S^m(s^m) (W_S^m((\mathcal{T}^n)^c | x^{*n}, s^n) + W_S^{\lceil \sqrt{n} \rceil}(\mathcal{D}_{F_i}(y^n) | \mathbf{u}_{F_i}(y^n), s^{\lceil \sqrt{n} \rceil})) \tag{105}$$

$$\stackrel{(b)}{\leq} 2^{-n\delta} + 2^{-\sqrt{n}\delta'}, \tag{106}$$

where (a) follows from the memorylessness property of the channel and the union bound, while (b) follows from Lemma 3 and the definition of the transmission code  $\mathcal{C}$ .

In order to achieve a small error of the second kind, we choose suitable maps  $\{F_i, i = 1, \dots, N\}$  randomly. For  $i \in \{1, \dots, N\}$ ,  $y^n \in \mathcal{T}^n$ , let  $\bar{F}_i(y^n)$  be independent RVs such that

$$\Pr\{\bar{F}_i(y^n) = j\} = \frac{1}{M}, \quad j \in \{1, \dots, M\}. \tag{107}$$

Let  $F_1$  be a realization of  $\bar{F}_1$ . For each  $y^n \in \mathcal{T}^n$ , we define the RVs  $\psi_{y^n} = \psi_{y^n}(\bar{F}_2)$  analogously to Section IV of [14]:

$$\psi_{y^n} = \psi_{y^n}(\bar{F}_2) = \begin{cases} 1, & \text{if } F_1(y^n) = \bar{F}_2(y^n) \\ 0, & \text{otherwise.} \end{cases} \tag{108}$$

The  $\psi_{y^n}$  are also independent for every  $y^n \in \mathcal{T}^n$ . The expectation of  $\psi_{y^n}$  is computed as follows:

$$\mathbb{E}[\psi_{y^n}] = \Pr\{F_1(y^n) = \bar{F}_2(y^n)\} = \frac{1}{M}. \tag{109}$$

Because the  $\psi_{y^n}$  are i.i.d. for all  $y^n \in \mathcal{T}^n$ , we can apply Hoeffding's inequality Lemma 5 to obtain the following Lemma.

**Lemma 6.** For  $\lambda \in (0, 1)$ ,  $\frac{1}{M} < \lambda$  for each channel  $V_S^a$  with  $\|V_S^a - W_S^a\| \leq \epsilon$ , while for each  $n \geq n_0(\epsilon)$  we have

$$\Pr\left\{ \sum_{y^n \in \mathcal{T}^n} \psi_{y^n} > |\mathcal{T}_{V_S^a}^n(x^{*n})| \cdot \lambda \right\} \leq 2^{-|\mathcal{T}^n| \cdot \lambda \sqrt{n} \epsilon}. \tag{110}$$

We can derive an upper bound on the error of the second kind for those values of  $\bar{F}_2$  satisfying Lemma 6:



$$\mathbb{E}_{S^m} [W_S^m(\mathcal{D}(\bar{F}_2)|f_1, S^m)] \tag{111}$$

$$= \sum_{s^m \in \mathcal{S}^m} P_S^m(s^m) W_S^m(\mathcal{D}(\bar{F}_2)|f_1, s^m) \tag{112}$$

$$= \sum_{s^m \in \mathcal{S}^m} P_S^m(s^m) W_S^m\left(\mathcal{D}(\bar{F}_2) \cap ((\mathcal{T}^n \times \mathcal{Y}^{\lceil \sqrt{n} \rceil}) \cup (\mathcal{T}^n \times \mathcal{Y}^{\lceil \sqrt{n} \rceil})^c) | f_1, s^m\right) \tag{113}$$

$$\stackrel{(a)}{\leq} \sum_{s^m \in \mathcal{S}^m} P_S^m(s^m) W_S^m((\mathcal{T}^n \times \mathcal{Y}^{\lceil \sqrt{n} \rceil})^c | f_1, s^m) \tag{114}$$

$$+ \sum_{s^m \in \mathcal{S}^m} P_S^m(s^m) W_S^m(\mathcal{D}(\bar{F}_2) \cap (\mathcal{T}^n \times \mathcal{Y}^{\lceil \sqrt{n} \rceil}) | f_1, s^m) \tag{115}$$

$$\stackrel{(b)}{\leq} \sum_{s^n \in \mathcal{S}^n} P_S^n(s^n) W_S^n((\mathcal{T}^n)^c | x^{*n}, s^n) \tag{116}$$

$$+ \sum_{s^m \in \mathcal{S}^m} P_S^m(s^m) \left( \sum_{\substack{y^n \in \mathcal{T}^n \\ F_1(y^n) \neq \bar{F}_2(y^n)}} W_S^n(y^n | x^{*n}, s^n) \cdot W_S^{\lceil \sqrt{n} \rceil}(y^{\lceil \sqrt{n} \rceil} | \mathbf{u}_{F_1(y^n)}, s^{\lceil \sqrt{n} \rceil}) \right) \tag{117}$$

$$+ \sum_{\substack{y^n \in \mathcal{T}^n \\ F_1(y^n) = \bar{F}_2(y^n)}} W_S^n(y^n | x^{*n}, s^n) \tag{118}$$

$$\leq 2^{-n\delta'} + 2^{-\sqrt{n}\delta} + \sum_{s^n \in \mathcal{S}^n} P_S^n(s^n) \sum_{\substack{y^n \in \mathcal{T}^n \\ F_1(y^n) = \bar{F}_2(y^n)}} W_S^n(y^n | x^{*n}, s^n) \tag{119}$$

$$\stackrel{(c)}{\leq} 2^{-n\delta'} + 2^{-\sqrt{n}\delta} \tag{120}$$

$$+ \sum_{V_S^a: |V_S^a - W_S^a| \leq \epsilon} (W_S^a)^n(\mathcal{T}_{V_S^a}^n(x^{*n}) | x^{*n}) \cdot \left( \sum_{y^n \in \mathcal{T}_{V_S^a}^n(x^{*n})} \psi_{y^n} \right) \cdot |\mathcal{T}_{V_S^a}^n(x^{*n})|^{-1} \tag{121}$$

$$\stackrel{(d)}{\leq} 2^{-n\delta'} + 2^{-\sqrt{n}\delta} + \lambda \tag{122}$$

where (a) follows from the union bound, (b) follows from the memorylessness property of the channel and the union bound, (c) follows from Lemma 3 along with the definition of the transmission code  $\mathcal{C}$  and the definition of the set  $\mathcal{T}^n$  in (90), and (d) follows from Lemma 6.

We repeatedly perform the same analysis of the error of the second kind for all pairs  $(i_1, i_2) \in \{1, \dots, N\}^2$ ,  $i_1 \neq i_2$ . For simplicity of notation, we denote the error of the second kind between the pair  $(i_1, i_2)$  by  $\mu_2^{(i_1, i_2)}$ . We have

$$\Pr\{\mathcal{C}' \text{ is not an } (n, N, \lambda_1, \lambda_2) \text{ code}\} \tag{123}$$

$$= \Pr\left\{ \bigcup_{\substack{i_1, i_2 \in \{1, \dots, N\} \\ i_1 \neq i_2}} \mu_2^{(i_1, i_2)} \geq \lambda_2 \right\} \tag{124}$$

$$= \Pr\left\{ \bigcup_{\substack{i_1, i_2 \in \{1, \dots, N\} \\ i_1 \neq i_2}} \mu_2^{(i_1, i_2)} \geq \lambda + 2^{-n\delta'} + 2^{-\sqrt{n}\delta} \right\} \tag{125}$$

$$\stackrel{(a)}{\leq} N \cdot (N - 1) \cdot 2^{-|\mathcal{T}^n| \cdot \lambda \cdot \sqrt{n}\epsilon}, \tag{126}$$

where (a) follows from the union bound, Equation (122) and Lemma 6. It is verifiable that we can construct an ID feedback code for  $W_S$  with cardinality  $N$  satisfying

$$N \geq (n + 1)^{-2|\mathcal{X}||\mathcal{Y}|} \cdot 2^{|\mathcal{T}^n| \cdot \lambda \sqrt{n} \epsilon} \tag{127}$$

and with an error of the second kind upper-bounded as in (122).

The next step in the proof is dedicated to the state estimator. The per-symbol distortion defined in (15) can be rewritten as follows:

$$d_t = \mathbb{E}[d(S_t, \hat{S}_t)] \tag{128}$$

$$= \mathbb{E}[\mathbb{E}[d(S_t, \hat{S}_t) | X_t, Y_t]] \tag{129}$$

$$= \sum_{(x,y) \in \mathcal{X} \times \mathcal{Y}} P_{X_t Y_t}(x, y) \sum_{s_t \in \mathcal{S}} P_{S_t | X_t Y_t}(s_t | x, y) \sum_{\hat{s}_t \in \mathcal{S}} P_{\hat{S}_t | X_t Y_t}(\hat{s}_t | x, y) d(s, \hat{s}) \tag{130}$$

$$= \sum_{(x,y) \in \mathcal{X} \times \mathcal{Y}} P_{X_t Y_t}(x, y) \sum_{s_t \in \mathcal{S}} P_{S_t | X_t Y_t}(s_t | x, y) d(s, h^*(x, y)) \tag{131}$$

$$= \sum_{x \in \mathcal{X}} P_{X_t}(x) \mathbb{E}_{S_t Y_t}[d(S_t, h^*(X_t, Y_t)) | X_t = x] \tag{132}$$

$$= \sum_{x \in \mathcal{X}} P_{X_t}(x) d^*(x) \tag{133}$$

$$\leq D. \tag{134}$$

This completes the direct proof of Theorem 4.

### 3.6. Converse Proof of Theorem 4

**Proof.** For the converse proof, we use the techniques for deterministic ID for  $W_S$  as described in Section 3.2. We first extend Lemma 1 to the joint deterministic ID and sensing problem.  $\square$

**Lemma 7.** For any feedback strategy  $f_D = [f_D^1, f_D^2, \dots, f_D^n]$  which satisfies the per-symbol distortion constraint as described in (17), i.e., for all  $t \in \{1, 2, \dots, n\}$ ,  $d^*(f_D^t) \leq D$  and for any  $\mu \in (0, 1)$ , we have

$$\min_{\mathcal{E}_3 \subset \mathcal{Y}^n: \mathbb{E}_{S^n}[W_S^n(\mathcal{E}_3 | f_D, S^n)] \geq 1 - \mu} |\mathcal{E}_3| \leq K_3, \tag{135}$$

where  $K_3$  is provided by

$$K_3 = 2^{n \max_{x \in \mathcal{X}_D} H(\mathbb{E}[W_S(\cdot | x)]) + \alpha \sqrt{n}} = 2^{n H(\mathbb{E}[W_S(\cdot | x^*)]) + \alpha \sqrt{n}}, \tag{136}$$

where

$$\alpha = \sqrt{\frac{\beta}{\mu}}, \quad \beta = \max(\log^2(3), \log^2(|\mathcal{Y}|)). \tag{137}$$

**Proof.** Let  $\mathcal{E}_3^* \subset \mathcal{Y}^n$  be defined as follows:

$$\mathcal{E}_3^* = \{y^n \in \mathcal{Y}^n, -\log \mathbb{E}_{S^n}[W_S^n(f_D, S^n)] \leq K_3\}. \tag{138}$$

Define an RV  $Z_t = -\log \mathbb{E}_{S_t}[W_S(Y_t | f_D^t(Y^{t-1}), S_t)]$ . Per (40), we have

$$\mathbb{E}_{S^n}[W_S^n(\mathcal{E}_3^* | f_D, S^n)] = Pr \left\{ \sum_{t=1}^n Z_t \leq \log K_3 \right\}. \tag{139}$$

Similarly, for all  $t \in \{1, 2, \dots, n\}$ , we define an RV  $U_t = Z_t - \mathbb{E}[Z_t|Y^{t-1}]$ . It has been shown in (44) and (48) that  $\mathbb{E}[U_t|Y^{t-1}] = 0$  and  $\mathbb{E}[U_t] = 0$ .

Moreover, for all  $t \in \{1, 2, \dots, n\}$  and for all  $y^{t-1} \in \mathcal{Y}^{t-1}$ , we have

$$\mathbb{E}[Z_t|y^{t-1}] = \sum_{y_t \in \mathcal{Y}} \left( -\mathbb{E}_{S_t} [W_S(Y_t|f_D^t(Y^{t-1}), S_t)] \log \mathbb{E}_{S_t} [W_S(Y_t|f_D^t(Y^{t-1}), S_t)] \right) \tag{140}$$

$$\leq \max_{x \in \mathcal{X}_D} H(\mathbb{E}_S[W_S(\cdot|x, S)]) \tag{141}$$

$$= H(\mathbb{E}_S[W_S(\cdot|x^*, S)]). \tag{142}$$

By combining (139) and (142), we have

$$\mathbb{E}_{S^n} [W_S^n(f_D, S^n)] \geq Pr \left\{ \sum_{t=1}^n U_t \leq \alpha \sqrt{n} \right\} \tag{143}$$

$$\geq 1 - \mu. \tag{144}$$

This completes the proof of Lemma 7  $\square$

The subsequent steps in the proof are identical to Section 3.2.

### 3.7. Direct Proof of Theorem 5

**Proof.** For the direct proof of this theorem, we follow a code construction similar to that presented in [14], with one key difference, namely, that we optimize only over input distributions that satisfy the per-symbol constraint  $d^*(P) \leq D$ .  $\square$

#### 3.7.1. Coding Scheme

We construct a randomized ID code with blocklength  $m = n + \sqrt{n}$  by concatenating two transmission codes, which is described in detail later. The first  $n$  symbols are allocated for generation of common randomness. We employ a distribution

$$P^* = \arg \max_{P \in \mathcal{P}(\mathcal{X})} H\left(\sum_{x \in \mathcal{X}} P(x) \mathbb{E}[W_S(\cdot|x, S)]\right),$$

where the maximization is performed over distributions subject to the constraint  $\mathcal{P}_D = \{P \in \mathcal{P}(\mathcal{X}) | d^*(P) \leq D\}$ . Regardless of which identity  $i \in \{1, \dots, N\}$  we want to identify, the sender first sends  $n$  symbols with respect to the distribution  $P^{*n} \in \mathcal{P}_D^n$  over the state-dependent channel  $W_S$ . The received sequence  $y^n \in \mathcal{Y}^n$  becomes known to the sender (estimator and encoder) via the noiseless feedback link. The feedback provides the sender and the receiver with knowledge of the outcome of the correlated random experiment  $\left(\mathcal{Y}^n, \sum_{x^n \in \mathcal{X}^n} P^{*n}(x^n) \mathbb{E}_{S^n} [W_S^n(\cdot|x^n, S^n)]\right)$ .

#### 3.7.2. Common Randomness Generation

Similar to the deterministic coding scheme, we want to generate uniform common randomness. Therefore, we convert our correlated random experiment

$$\left(\mathcal{Y}^n, \sum_{x^n \in \mathcal{X}^n} P^{*n}(x^n) \mathbb{E}_{S^n} [W_S^n(\cdot|x^n, S^n)]\right)$$

to a uniform one  $\left(\mathcal{T}^n, \sum_{x^n \in \mathcal{X}^n} P^{*n}(x^n) \mathbb{E}_{S^n} [W_S^n(\cdot | x^n, S^n)]\right)$ . For  $\epsilon > 0$ , the set  $\mathcal{T}^n$  is provided by

$$\mathcal{T}^n = \bigcup_{V_S^n: \|V_S^n - W_S^n\| \leq \epsilon} \sum_{x \in \mathcal{X}} P^{*n}(x^n) \mathcal{T}_{V_S^n}^n(x^n). \tag{145}$$

Per Lemma 4, we can obtain the following corollary.

**Corollary 1.** *Let  $(x^n, y^n)$  be emitted by the DMS  $P_{XY}(\cdot, \cdot) = W(\cdot | \cdot) P_X(\cdot)$  and let  $V \in \mathcal{W}$ . For every  $\epsilon > 0$ , there exist a  $c(\epsilon) > 0$  and  $n_0(\epsilon)$  such that for  $n \geq n_0(\epsilon)$  we have:*

1.  $\left| \bigcup_{V: \|V-W\| \leq \epsilon} \mathcal{T}_V^n(x^n) \right| \geq 2^{n(H(\sum_{x \in \mathcal{X}} P_X(x)W(\cdot|x)) - c(\epsilon))}$ ,
2.  $\left| \bigcup_{V: \|V-W\| \leq \epsilon} \mathcal{T}_V^n(x^n) \right| \leq 2^{n(H(\sum_{x \in \mathcal{X}} P_X(x)W(\cdot|x)) + c(\epsilon))}$ ,
3. *If  $\|V - W\| \leq \epsilon$ ,  $\mathcal{T}_V^n(x^n) \neq \emptyset$  and  $c(\epsilon) \rightarrow 0$  if  $\epsilon \rightarrow 0$ , then*

$$|\mathcal{T}_V^n(x^n)| \geq 2^{n(H(\sum_{x \in \mathcal{X}} P_X(x)W(\cdot|x)) - c(\epsilon))}. \tag{146}$$

Therefore, we have  $|\mathcal{T}^n| \approx 2^{nH(\sum_{x \in \mathcal{X}} P(x) \mathbb{E}_S [W(\cdot|x, S)])}$ . The sender generates randomness according to the random experiment  $(\mathcal{T}^n, \sum_{x^n \in \mathcal{X}^n} P^{*n}(x^n) \mathbb{E}_{S^n} [W_S^n(\cdot | x^n, S^n)])$ . Asymptotically, the error probability of  $y^n \notin \mathcal{T}^n$  goes to zero. Similar to the deterministic scheme, we prepare the coloring functions  $\{F_i, i = 1, \dots, N\}$ . The last  $\lceil \sqrt{n} \rceil$  symbols are used to transmit  $F_i(y^n)$  using a standard  $(\lceil \sqrt{n} \rceil, M, 2^{-n\sqrt{n}\delta})$  transmission code  $C = \{(u_j, \mathcal{D}_j), j = 1, \dots, M\}$ , where  $u_j \in \mathcal{X}_D^{\lceil \sqrt{n} \rceil}$  for each  $j = 1, \dots, M$ . The probability distribution for encoding is defined as  $Q_F(\cdot | i) = P^{*n} \times \mathbb{I}\{x_n^m = F_i(y^n)\}$  and the decoding region is provided by  $\mathcal{D}(F_i) = \bigcup_{y^n \in \mathcal{T}^n} \{y^n\} \times \mathcal{D}(F_i(y^n))$ .

### 3.7.3. Error Analysis

Subsequently, for all  $i = 1, \dots, N$ , the error of the first kind  $P_{e,1}(i)$  is upper-bounded by

$$P_{e,1}(i) \tag{147}$$

$$= \sum_{f \in \mathcal{F}^m} Q_F(f|i) \mathbb{E}_{S^m} [W_S^m(\mathcal{D}(F_i) | f, S^m)] \tag{148}$$

$$= \sum_{s^m \in \mathcal{S}^m} P_S^m(s^m) \sum_{f \in \mathcal{F}^m} Q_F(f|i) W_S^m \left( \left( \bigcup_{y^n \in \mathcal{T}^n} y^n \times \mathcal{D}_{F_i}(y^n) \right)^c \middle| f, S^m \right) \tag{149}$$

$$\stackrel{(a)}{\leq} \sum_{s^m \in \mathcal{S}^m} P_S^m(s^m) \left( \sum_{x^n \in \mathcal{X}^n} P^{*n}(x^n) W_S^n(\mathcal{T}^c | x^n, S^n) + W_S^{\lceil \sqrt{n} \rceil}(\mathcal{D}_{F_i}^c | u_{F_i}(y^n), S^{\lceil \sqrt{n} \rceil}) \right) \tag{150}$$

$$\stackrel{(b)}{\leq} 2^{-n\delta} + 2^{-\sqrt{n}\delta'}, \tag{151}$$

where (a) follows from the union bound and (b) follows from Corollary 1. Furthermore, for all  $i, j = 1, \dots, N$  with  $i \neq j$ , the probability of the error of the second kind  $P_{e,2}(i, j)$  should be asymptotically upper-bounded by  $\lambda$ . Without loss of generality, we fix  $i = 1, j = 2$  and examine the error probability  $P_{e,2}(1, 2)$ :

$$P_{e,2}(1,2) \tag{152}$$

$$= \sum_{f \in \mathcal{F}^m} Q(f|i=1) \mathbb{E}_{S^m} [W_S^m(\mathcal{D}(\bar{F}_2)|f, S^m)] \tag{153}$$

$$= \sum_{s^m \in \mathcal{S}^m} P_S^m(s^m) \sum_{f \in \mathcal{F}^m} Q(f|i=1) W_S^m(\mathcal{D}(\bar{F}_2)|f, s^m) \tag{154}$$

$$= \sum_{s^m \in \mathcal{S}^m} P_S^m(s^m) \sum_{f \in \mathcal{F}^m} Q(f|i=1) \cdot \tag{155}$$

$$\cdot W_S^m \left( \mathcal{D}(\bar{F}_2) \cap ((\mathcal{T}^m \times \mathcal{Y}^{\lceil \sqrt{n} \rceil}) \cup (\mathcal{T}^m \times \mathcal{Y}^{\lceil \sqrt{n} \rceil})^c) | f, s^m \right) \tag{156}$$

$$\stackrel{(a)}{\leq} \sum_{s^m \in \mathcal{S}^m} P_S^m(s^m) \sum_{f \in \mathcal{F}^m} Q(f|i=1) W_S^m((\mathcal{T}^m \times \mathcal{Y}^{\lceil \sqrt{n} \rceil})^c | f, s^m) \tag{157}$$

$$+ \sum_{s^m \in \mathcal{S}^m} P_S^m(s^m) \sum_{f \in \mathcal{F}^m} Q(f|i=1) W_S^m(\mathcal{D}(\bar{F}_2) \cap (\mathcal{T}^m \times \mathcal{Y}^{\lceil \sqrt{n} \rceil}) | f, s^m) \tag{158}$$

$$\stackrel{(b)}{\leq} \sum_{s^m \in \mathcal{S}^m} P_S^m(s^m) \sum_{x^n \in \mathcal{X}^n} P^{*n}(x^n) W_S^n((\mathcal{T}^n)^c | x^n, s^n) W_S^{\lceil \sqrt{n} \rceil}(\mathcal{D}_{F_i(y^n)}^c | \mathbf{u}_{F_i(y^n)}, s^{\lceil \sqrt{n} \rceil}) \tag{159}$$

$$+ \sum_{s^m \in \mathcal{S}^m} P_S^m(s^m) \left( \sum_{\substack{y^n \in \mathcal{T}^n \\ F_1(y^n) \neq \bar{F}_2(y^n)}} \sum_{x^n \in \mathcal{X}^n} P^{*n}(x^n) W_S^n(y^n | x^n, s^n) \cdot \tag{160}$$

$$\cdot W_S^{\lceil \sqrt{n} \rceil}(y^{\lceil \sqrt{n} \rceil} | \mathbf{u}_{F_1(y^n)}, s^{\lceil \sqrt{n} \rceil}) + \sum_{\substack{y^n \in \mathcal{T}^n \\ F_1(y^n) = \bar{F}_2(y^n)}} \sum_{x^n \in \mathcal{X}^n} P^{*n}(x^n) W_S^n(y^n | x^n, s^n) \right) \tag{161}$$

$$\leq 2^{-n\delta'} + 2^{-\sqrt{n}\delta} + \sum_{s^n \in \mathcal{S}^n} P_S^n(s^n) \sum_{\substack{y^n \in \mathcal{T}^n \\ F_1(y^n) = \bar{F}_2(y^n)}} \sum_{x^n \in \mathcal{X}^n} P^{*n}(x^n) W_S^n(y^n | x^n, s^n) \tag{162}$$

$$\stackrel{(c)}{\leq} 2^{-n\delta'} + 2^+ \sum_{x^n \in \mathcal{X}^n} P^{*n}(x^n) \sum_{V_S^a: |V_S^a - W_S^a| \leq \epsilon} (W_S^a)^n(\mathcal{T}_{V_S^a}^n(x^n) | x^n) \cdot \tag{163}$$

$$\cdot \left( \sum_{y^n \in \mathcal{T}_{V_S^a}^n(x^n)} \psi_{y^n} \right) \cdot |\mathcal{T}_{V_S^a}^n(x^n)|^{-1} \tag{164}$$

$$\stackrel{(d)}{\leq} 2^{-n\delta'} + 2^{-\sqrt{n}\delta} + \lambda \tag{165}$$

where (a) follows from the union bound, (b) follows from the memoryless channel and the union bound, (c) follows from Lemma 3 along with the definition of the transmission code  $\mathcal{C}$  and the definition of the set  $\mathcal{T}^n$  in (90), and (d) follows from Corollary 1.

We repeatedly perform the same analysis of the error of the second kind for all pairs  $(i_1, i_2) \in \{1, \dots, N\}^2, i_1 \neq i_2$ . It is verifiable that we can construct a randomized ID feedback code for  $W_S$  with cardinality  $N$  satisfying

$$N \geq (n + 1)^{-2|\mathcal{X}||\mathcal{Y}|} \cdot 2^{|\mathcal{T}^m| \cdot \lambda \sqrt{n}\epsilon} \tag{166}$$

and with errors of the first and second kind that are upper-bounded as in (11) and (12), respectively.

Finally the state estimator is checked. The per-symbol distortion defined in (15) can be rewritten as follows:

$$d_t = \mathbb{E}[d(S_t, \hat{S}_t)] \tag{167}$$

$$= \mathbb{E}[\mathbb{E}[d(S_t, \hat{S}_t)|X_t, Y_t]] \tag{168}$$

$$= \sum_{(x,y) \in \mathcal{X} \times \mathcal{Y}} P_{X_t Y_t}(x, y) \sum_{s_t \in \mathcal{S}} P_{S_t|X_t Y_t}(s|x, y) \sum_{\hat{s}_t \in \mathcal{S}} P_{\hat{S}_t|X_t Y_t}(\hat{s}|x, y) d(s, \hat{s}) \tag{169}$$

$$= \sum_{(x,y) \in \mathcal{X} \times \mathcal{Y}} P_{X_t Y_t}(x, y) \sum_{s_t \in \mathcal{S}} P_{S_t|X_t Y_t}(s|x, y) d(s, h^*(x, y)) \tag{170}$$

$$= \sum_{x \in \mathcal{X}} P_{X_t}(x) \mathbb{E}_{S_t, Y_t}[d(S_t, h^*(X_t, Y_t))|X_t = x] \tag{171}$$

$$= d^*(P) \tag{172}$$

$$\leq D. \tag{173}$$

This completes the direct proof of Theorem 5.

### 3.8. Converse Proof of Theorem 5

**Proof.** We first extend Lemma 2 to the joint randomized ID and sensing problem.  $\square$

**Lemma 8.** For any randomized feedback strategy  $Q_D(f) = \prod_{t=1}^n Q_D^t(f^t)$  over all  $n$ -length feedback encoding sets  $\mathcal{F}_n$  which satisfies the per symbol distortion constraint as described in (142), i.e., for all  $t \in \{1, 2, \dots, n\}$ ,  $d^*(Q_D^t(f^t)) \leq D$  and for any  $\mu \in (0, 1)$ , we have

$$\min_{\mathcal{E}_4 \subset \mathcal{Y}^n: \sum_{f \in \mathcal{F}_n} Q_D(f) \mathbb{E}_{S^n}[W_S^n(\mathcal{E}_4|f, S^n)] \geq 1 - \mu} |\mathcal{E}_4| \leq K_4, \tag{174}$$

where  $K_4$  is provided by

$$K_4 = 2^{n \max_{P \in \mathcal{P}^D} H(\sum_{x \in \mathcal{X}} P(x) \mathbb{E}[W_S(\cdot|x, S)]) + \alpha \sqrt{n}} = 2^{n H(\sum_{x \in \mathcal{X}} P^*(x) \mathbb{E}[W_S(\cdot|x, S)]) + \alpha \sqrt{n}}, \tag{175}$$

where  $\alpha = \sqrt{\frac{\beta}{\mu}}$ ,  $\beta = \max \log^2 3, \log^2 |\mathcal{Y}|$ .

**Proof.** Define a set  $\mathcal{E}_4^* \subset \mathcal{Y}^n$  as follows:

$$\mathcal{E}_4^* = \left\{ y^n \in \mathcal{Y}^n, -\log \sum_{f \in \mathcal{F}_n} Q_D(f) \mathbb{E}_{S^n}[W_S(y^n|f, S^n)] \leq \log K_4 \right\}. \tag{176}$$

Define an auxiliary RV  $Z_t$  as follows:

$$Z_t = -\log \left( \sum_{f^t \in \mathcal{F}^t} Q_D^t(f^t) \mathbb{E}_{S_t}[W_S^n(Y_t|f^t(Y^{t-1}), S_t)] \right). \tag{177}$$

Then, per (80), we have

$$\sum_{f \in \mathcal{F}_n} Q_D(f) \mathbb{E}[W_S^n(\mathcal{E}_4|f, S^n)] = Pr \left\{ \sum_{t=1}^n Z_t \leq \log K_4 \right\}. \tag{178}$$

Similarly, for any  $y^{t-1} \in \mathcal{Y}^{t-1}$ , we can examine

$$\mathbb{E}\left[Z_t|y^{t-1}\right] \tag{179}$$

$$= \sum_{y_t \in \mathcal{Y}} \left( - \sum_{f^t \in \mathcal{F}^t} Q_D^t(f^t) \mathbb{E}_{S_t} [W_S^n(y_t|f^t(y^{t-1}), S_t)] \log \left( \sum_{f^t \in \mathcal{F}^t} Q_D^t(f^t) \mathbb{E}_{S_t} [W_S^n(y_t|f^t(y^{t-1}), S_t)] \right) \right) \tag{180}$$

$$\leq \max_{P \in \mathcal{P}_D} H \left( \sum_{x \in \mathcal{X}} P(x) \mathbb{E}[W_S(\cdot|x, S)] \right) \tag{181}$$

$$= H \left( \sum_{x \in \mathcal{X}} P^*(x) \mathbb{E}[W_S^n(\cdot|x, S)] \right). \tag{182}$$

By applying Chebyshev’s inequality, we complete the proof of Lemma 8. □

The subsequent steps in the proof are the same as in Section 3.6.

### 4. Average Distortion

In addition to the per-symbol distortion constraint, an alternative and more flexible distortion constraint is the average distortion. This approach is valuable because it relaxes the per-symbol fidelity requirement, allowing for minor variations in individual symbol quality as long as the overall average distortion remains below a specified threshold. As defined in [32], the average distortion for a sequence of symbols is provided by

$$\bar{d}^n = \mathbb{E}_{S^n, \hat{S}^n} [d(S^n, \hat{S}^n)] \tag{183}$$

$$= \frac{1}{n} \sum_{t=1}^n \mathbb{E}_{S_t, \hat{S}_t} [d(S_t, \hat{S}_t)]. \tag{184}$$

This metric captures the average quality of the reconstructed sequence, making it suitable for applications where consistent strict fidelity for each symbol is not essential but where the overall fidelity of the transmission needs to remain within acceptable limits.

In the case of a deterministic ID code, the average distortion can be expressed in a more detailed form as

$$\bar{d}^n = \frac{1}{N} \sum_{i=1}^N \mathbb{E}_{S^n Y^n} \left[ \frac{1}{n} \sum_{t=1}^n d(S_t, \hat{S}_t) | X^n = f_i \right]. \tag{185}$$

Using the code construction method from Section 3.5 along with the minimum distortion condition defined in (17), we propose the following theorem, which provides a lower bound on the deterministic ID capacity–distortion function for a state-dependent channel  $W_S$  under an average distortion constraint.

**Theorem 6.** *The deterministic ID capacity–distortion function with average distortion constraint  $\bar{d}^n \leq D$  of the state-dependent channel  $W_S$  is lower-bounded as follows:*

$$C_{ID,avg}^d(D) \geq \max_{x \in \mathcal{X}_D} H(\mathbb{E}[W_S(\cdot|x, S)]), \tag{186}$$

where the set  $\mathcal{X}_D$  is provided by

$$\mathcal{X}_D = \{x \in \mathcal{X}, d^*(x) \leq D\}. \tag{187}$$

Despite the practical implications of this result, proving a converse theorem for this bound remains an open problem.



## 5. Conclusions and Discussion

In this work, we have studied the problem of joint ID and channel state estimation over a DMC with i.i.d. state sequences where the sender simultaneously sends an identity and senses the state via a strictly causal channel output. After establishing the capacity on the corresponding ID capacity–distortion function, it emerges that sensing can increase the ID capacity. In the proof of our theorem, we noticed that the generation of common randomness is a key tool for achieving a high ID rate. The common randomness generation is helped by feedback. The ID rate can be further increased by adding local randomness at the sender.

Our framework closely mirrors the one described in [23], with the key distinction being that we utilize an identification scheme instead of the classical transmission scheme. We want to simultaneously identify the sent message and estimate the channel's state. As noted in the results of [23], the capacity–distortion function is consistently smaller than the transmission capacity of the state-dependent DMC except when the distortion is infinite. This observation aligns with expectations for the message transmission scheme, as the optimization is performed over a constrained input set defined by the distortion function. However, this does not directly apply to the ID scheme. An interesting aspect is that the capacity–distortion function for the deterministic encoding case scales double exponentially with the blocklength, as highlighted in Theorem 4. However, the ID capacity of the state-dependent DMC with deterministic encoding scales only exponentially with the blocklength. This is because feedback significantly enhances the ID capacity, enabling double-exponential growth of the ID capacity for the state-dependent DMC, as established in Theorem 2. This contrasts sharply with the message transmission scheme, where feedback does not increase the capacity of a DMC. Introducing an estimator into our framework naturally reduces the ID capacity compared to the scenario with feedback but without an estimator. This reduction occurs because the optimization is performed over a constrained input set defined by the distortion function. Nevertheless, the capacity–distortion function remains higher than in the case without feedback and without an estimator. This difference underscores a unique characteristic of the ID scheme, highlighting its distinct scaling behavior and potential advantages in certain scenarios.

We consider two cases, namely, deterministic and randomized identification. For a transmission system without sensing, it was shown in [1,3] that the number of messages grows exponentially, i.e.,  $N = 2^{nC_{ID}^d}$ .

Remarkably, Theorem 4 demonstrates that by incorporating sensing, the growth rate of the number of messages becomes double exponential ( $N = 2^{2^{nC_{ID}^d(D)}}$ ); this result is notable, and closely parallels the findings on identification with feedback in [14].

In the case of randomized identification, Theorem 5 shows that the capacity is also improved by incorporating sensing. However, in both the deterministic and randomized settings, the scaling remains double exponential.

One application of message identification is in control and alarm systems [10,33]. For instance, it has been shown that identification codes can be used for status monitoring in digital twins [34]. In this context, our results demonstrate that incorporating a sensing process can significantly enhance the capacity.

Another potential application of our framework is molecular communication, where nanomachines use identification codes to determine when to perform specific actions such as drug delivery [35]. In this context, sensing the position of the nanomachines can enhance the communication rate. For such scenarios, it is also essential to explore alternative channel models such as the Poisson channel.

Furthermore, it is clear that in other applications it would be necessary to consider different distortion functions.

In the future, it would be interesting to apply the method proposed in this paper to other distortion functions. Furthermore, in practical scenarios, there are models where the sensing is performed either additionally or exclusively by the receiver. This suggests the need to study a wider variety of system models. For wireless communications, the Gaussian channel is more practical and widely applicable. Therefore, it would be valuable to extend our results to the Gaussian case (JIDAS scheme with a Gaussian channel as the forward channel). It has been shown in [15,28] that the ID capacity of a Gaussian channel with noiseless feedback is infinite. Interestingly, the ID capacity of a Gaussian channel with noiseless feedback remains infinite regardless of the scaling used for the rate, e.g., double exponential, triple exponential, etc. By introducing an estimator, we conjecture that the same results will hold, leading to an infinite capacity–distortion function. Thus, considering scenarios with noisy feedback is more practical for future research.

**Author Contributions:** Writing—original draft preparation, W.L. and Y.Z.; supervision, C.D. and H.B. All authors have read and agreed to the published version of the manuscript.

**Funding:** The authors acknowledge financial support by the Federal Ministry of Education and Research of Germany (BMBF) through the “Souverän. Digital. Vernetzt.” programme, joint project “6G-life”, project identification number 16KISK002. H. Boche and W. Labidi were further supported in part by the BMBF within the national initiative on Post-Shannon Communication (NewCom) under Grant 16KIS1003K and within the national initiative on molecular communication (IoBNT) under grant 16KIS1988. C. Deppe was further supported in part by the BMBF within the national initiative on Post-Shannon Communication (NewCom) under Grant 16KIS1005. C. Deppe, W. Labidi and Y. Zhao were supported by the DFG within the projects DE1915/2-1 and BO 1734/38-1.

**Institutional Review Board Statement:** Not applicable.

**Data Availability Statement:** The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

**Acknowledgments:** This work has been presented in part at the IEEE International Symposium on Information Theory 2023 (ISIT 2023) [32].

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

## References

1. Ahlswede, R.; Dueck, G. Identification via channels. *IEEE Trans. Inf. Theory* **1989**, *35*, 15–29. [[CrossRef](#)]
2. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [[CrossRef](#)]
3. Salariseddigh, M.J.; Pereg, U.; Boche, H.; Deppe, C. Deterministic Identification over Channels with Power Constraints. *IEEE Trans. Inf. Theory* **2022**, *68*, 1–24. [[CrossRef](#)]
4. Ahlswede, R.; Cai, N. Identification without randomization. *IEEE Trans. Inf. Theory* **1999**, *45*, 2636–2642. [[CrossRef](#)]
5. Moulin, P. The role of information theory in watermarking and its application to image watermarking. *Signal Process.* **2001**, *81*, 1121–1139. [[CrossRef](#)]
6. Ahlswede, R., Watermarking Identification Codes with Related Topics on Common Randomness. In *Identification and Other Probabilistic Models: Rudolf Ahlswede’s Lectures on Information Theory 6*; Ahlswede, A., Althöfer, I., Deppe, C., Tamm, U., Eds.; Springer International Publishing: Cham, Switzerland, 2021; pp. 271–325. [[CrossRef](#)]
7. Steinberg, Y.; Merhav, N. Identification in the presence of side information with application to watermarking. *IEEE Trans. Inf. Theory* **2001**, *47*, 1410–1422. [[CrossRef](#)]
8. Lu, Y. Industry 4.0: A survey on technologies, applications and open research issues. *J. Ind. Inf. Integr.* **2017**, *6*, 1–10. [[CrossRef](#)]
9. Fettweis, G.; Boche, H. 6G: The Personal Tactile Internet—And Open Questions for Information Theory. *IEEE BITS Inf. Theory Mag.* **2021**, *1*, 71–82. [[CrossRef](#)]
10. Cabrera, J.; Boche, H.; Deppe, C.; Schaefer, R.F.; Scheunert, C.; Fitzek, F. 6G and the Post-Shannon-Theory. In *Shaping Future 6G Networks: Needs, Impacts and Technologies*; Bertin, E., Crespi, N., Magedanz, T., Eds.; Wiley-Blackwell: Oxford, UK, 2021. [[CrossRef](#)]
11. Fettweis, G.; Boche, H. On 6G and trustworthiness. *Commun. ACM* **2022**, *65*, 48–49. [[CrossRef](#)]

12. Ahlswede, R.; Zhang, Z. New directions in the theory of identification via channels. *IEEE Trans. Inf. Theory* **1995**, *41*, 1040–1050. [[CrossRef](#)]
13. Labidi, W.; Deppe, C.; Boche, H. Secure Identification for Gaussian Channels. In Proceedings of the ICASSP 2020—2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Barcelona, Spain, 4–8 May 2020; pp. 2872–2876. [[CrossRef](#)]
14. Ahlswede, R.; Dueck, G. Identification in the presence of feedback—a discovery of new capacity formulas. *IEEE Trans. Inf. Theory* **1989**, *35*, 30–36. [[CrossRef](#)]
15. Labidi, W.; Boche, H.; Deppe, C.; Wiese, M. Identification over the Gaussian Channel in the presence of feedback. In Proceedings of the 2021 IEEE International Symposium on Information Theory (ISIT), Melbourne, Australia, 12–20 July 2021; pp. 278–283. [[CrossRef](#)]
16. Ahlswede, R. General theory of information transfer: Updated. *Discret. Appl. Math.* **2008**, *156*, 1348–1388. [[CrossRef](#)]
17. Ahlswede, R.; Csiszar, I. Common randomness in information theory and cryptography. II. CR capacity. *IEEE Trans. Inf. Theory* **1998**, *44*, 225–240. [[CrossRef](#)]
18. *Proceedings of the 1st IEEE International Online Symposium on Joint Communications and Sensing*; IEEE: Piscataway, NJ, USA, 2021. <https://www.proceedings.com/58212.html>.
19. Sturm, C.; Wiesbeck, W. Waveform Design and Signal Processing Aspects for Fusion of Wireless Communications and Radar Sensing. *Proc. IEEE* **2011**, *99*, 1236–1259. [[CrossRef](#)]
20. Bliss, D.W. Cooperative radar and communications signaling: The estimation and information theory odd couple. In Proceedings of the 2014 IEEE Radar Conference, Cincinnati, OH, USA, 9–23 May 2014; pp. 50–55. [[CrossRef](#)]
21. Bica, M.; Huang, K.W.; Mitra, U.; Koivunen, V. Opportunistic Radar Waveform Design in Joint Radar and Cellular Communication Systems. In Proceedings of the 2015 IEEE Global Communications Conference (GLOBECOM), San Diego, CA, USA, 6–10 December 2015; pp. 1–7. [[CrossRef](#)]
22. Huang, K.W.; Bică, M.; Mitra, U.; Koivunen, V. Radar waveform design in spectrum sharing environment: Coexistence and cognition. In Proceedings of the 2015 IEEE Radar Conference (RadarCon), Arlington, VA, USA, 10–15 May 2015; pp. 1698–1703. [[CrossRef](#)]
23. Kobayashi, M.; Caire, G.; Kramer, G. Joint State Sensing and Communication: Optimal Tradeoff for a Memoryless Case. In Proceedings of the 2018 IEEE International Symposium on Information Theory (ISIT), Vail, CO, USA, 17–22 June 2018; pp. 111–115. [[CrossRef](#)]
24. Choudhuri, C.; Kim, Y.H.; Mitra, U. Causal State Communication. *IEEE Trans. Inf. Theory* **2013**, *59*, 3709–3719. [[CrossRef](#)]
25. Chiu, S.E.; Ronquillo, N.; Javidi, T. Active Learning and CSI Acquisition for mmWave Initial Alignment. *IEEE J. Sel. Areas Commun.* **2019**, *37*, 2474–2489. [[CrossRef](#)]
26. Han, T.S. *Information-Spectrum Methods in Information Theory; Stochastic Modelling and Applied Probability*; Springer: Berlin/Heidelberg, Germany, 2014. [[CrossRef](#)]
27. El Gamal, A.; Kim, Y.H. *Network Information Theory*; Cambridge University Press: Cambridge, UK, 2011. [[CrossRef](#)]
28. Wiese, M.; Labidi, W.; Deppe, C.; Boche, H. Identification over Additive Noise Channels in the Presence of Feedback. *IEEE Trans. Inf. Theory* **2022**, *69*, 6811–6821. [[CrossRef](#)]
29. Zhang, W.; Vedantam, S.; Mitra, U. Joint Transmission and State Estimation: A Constrained Channel Coding Approach. *IEEE Trans. Inf. Theory* **2011**, *57*, 7084–7095. [[CrossRef](#)]
30. Csiszár, I.; Körner, J. *Information Theory: Coding Theorems for Discrete Memoryless Systems*; Cambridge University Press: Cambridge, UK, 2011. [[CrossRef](#)]
31. Hoeffding, W. On Probabilities of Large Deviations. In *The Collected Works of Wassily Hoeffding*; Springer: New York, NY, USA, 1994; pp. 473–490. [[CrossRef](#)]
32. Labidi, W.; Deppe, C.; Boche, H. Joint identification and sensing for discrete memoryless channels. In Proceedings of the 2023 IEEE International Symposium on Information Theory (ISIT), Taipei, Taiwan, 25–30 June 2023; pp. 442–447. [[CrossRef](#)]
33. Bringer, J.; Chabanne, H. Code Reverse Engineering Problem for Identification Codes. *IEEE Trans. Inf. Theory* **2012**, *58*, 2406–2412. [[CrossRef](#)]
34. von Lengerke, C.; Cabrera, J.A.; Fitzek, F.H.P. Identification Codes for Increased Reliability in Digital Twin Applications over Noisy Channels. In Proceedings of the 2023 IEEE International Conference on Metaverse Computing, Networking and Applications (MetaCom), Kyoto, Japan, 26–28 June 2023; pp. 550–557. [[CrossRef](#)]
35. Labidi, W.; Deppe, C.; Boche, H. Information-Theoretical Analysis of Event-Triggered Molecular Communication. In Proceedings of the European Wireless 2024, Brno, Czech Republic, 9–11 September 2024.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.