

# The Capacity of Injective Semi-Deterministic Two-Way Channels

Anas Chaaban\*, Lav R. Varshney<sup>†</sup>, and Mohamed-Slim Alouini\*

\* CEMSE Division  
King Abdullah University of Science and Technology  
Thuwal, Saudi Arabia  
Email: {anas.chaaban,slim.alouini}@kaust.edu.sa

<sup>†</sup> Department of Electrical and Computer Engineering  
University of Illinois at Urbana-Champaign  
Urbana, IL, USA  
Email: varshney@illinois.edu

**Abstract**—The capacity region of the class of injective semi-deterministic two-way channels (TWCs) is investigated in this paper. To characterize this capacity, two conditions under which Shannon’s bounds on the capacity region of TWCs are tight are first given. Using those conditions, it is shown that the capacity of this class of TWCs is characterized by the rectangle formed by the one-way capacities. This proves that adaptation is not needed for this class. This class encompasses, among others, all memoryless additive channels with input-independent noise, and hence, adaptation is useless for all such channels. This also shows that there exist continuous additive TWCs not of the exponential family type for which adaptation is not necessary. An example of a Cauchy TWC is given, and its capacity is characterized in closed form under a logarithmic constraint. Finally, the impact of the dependence of the noise on the inputs is discussed, and it is shown that adaptation may still be useless in such cases.

## I. INTRODUCTION

Recent developments in full-duplex communications [1] have increased the interest in studying two-way channels (TWCs) [2], [3] and extensions thereof [4]–[8]. A TWC models a scenario where two nodes communicate with each other in both directions, i.e., each node is both a message source and destination. This channel was first studied by Shannon in [9] where capacity inner and outer bounds were derived, and the capacity region was established precisely for some special cases.

Studying Shannon’s bounds for the TWC reveals an important aspect of coding for these channels, namely *adaptation*. The inner bound is achieved using nonadaptive encoding, i.e., where the transmit signal at each node at a given time instant is constructed from the message only, and is independent of the received signals at that node up to that time instant. This particular encoding makes the transmit signals of the two nodes independent. The outer bound, on the other hand, allows those signals to be dependent. Thus, the bounds do not generally coincide, and several works have focused on tightening those bounds [10], [11] (cf. [12, Ch. 17]). However, if they do coincide for a given channel, this implies that adaptive encoding is not useful in this particular channel.

This issue has raised interest in investigating channels where Shannon’s bounds coincide, particularly since nonadaptive encoding is simpler to implement than adaptive encoding. For example, in practical full-duplex communication systems,

a nonadaptive scheme consists of using one-way codes for encoding, and applying self-interference cancellation before decoding the desired signals. Special interest has been on additive noisy TWCs since they often model practical communication scenarios. For instance, in [13], a new inner bound on the capacity of the TWC was developed, and the capacity of the Gaussian TWC was characterized and shown to be achievable without adaptation. This result has been extended in [2] to exponential-family TWCs, a type that subsumes the Gaussian TWC. This extension is based on a *saddle point* theorem which holds for this family [14], [15]. The impact of adaptation in TWCs with more than two terminals has been studied in [5], [6]. The capacity of the finite-field additive noise TWC with memory was characterized in [3].

Among continuous-alphabet memoryless additive TWCs, the exponential family is the most general class of channels for which it is known that adaptation is not needed for optimality. This raises the following interesting question: Does this class encompass all continuous-alphabet memoryless additive TWCs for which adaptation is not necessary? To answer this question, we note that this class belongs to a more general class of TWCs, the class of *injective semi-deterministic (ISD) TWCs*. Thus, we study this more general class.

We first give two sufficient conditions under which Shannon’s bounds coincide (Sec. III). These conditions are then used to characterize the capacity region of the class of ISD TWCs with input-independent noise in Sec. IV. This class is different from the two classes whose capacities were characterized in [10], which have injective and deterministic properties. For this class, we show that the capacity region is rectangular, characterized by the one-way capacities, and hence adaptation is ineffective. It follows from this result that the capacity of all memoryless additive TWCs with input-independent noise is rectangular defined by the one-way capacities, and that adaptation is ineffective. Applying this result to continuous TWCs immediately answers the original question; there are channels not of the exponential family for which adaptation is ineffective. We give the Cauchy TWC as an example, and characterize its capacity in closed form under logarithmic constraints. Finally, we comment on channels with input-dependent noise, for some of which adaptation might still be unnecessary for optimality. We note that the results do not

require saddle point admissibility as in [2].

## II. SYSTEM MODEL AND PRELIMINARIES

### A. Model

A memoryless TWC models two nodes communicating with each other, where each node is both a source and a destination. It is described by input alphabets  $\mathcal{X}_1$  and  $\mathcal{X}_2$ , output alphabets  $\mathcal{Y}_1$  and  $\mathcal{Y}_2$ , and a transition probability  $P_{Y_1, Y_2 | X_1, X_2}$  (Fig. 1).

Node 1 wants to send a message  $W_1$  to node 2, and node 2 wants to send a message  $W_2$  to node 1. The messages are independent and uniformly distributed over  $\mathcal{W}_1 = \{1, \dots, 2^{nR_1}\}$  and  $\mathcal{W}_2 = \{1, \dots, 2^{nR_2}\}$ , respectively. Communication takes place over  $n$  transmissions, where in transmission time  $t$ , node 1 uses an encoding function  $\mathcal{E}_{1,t} : \mathcal{W}_1 \times \mathcal{Y}^{t-1} \mapsto \mathcal{X}_1$  to obtain  $X_1(t) = \mathcal{E}_{1,t}(W_1, Y_1(1), \dots, Y_1(t-1))$  which is sent to node 2. This type of encoding is known as adaptive encoding, contrary to restricted (nonadaptive) encoding where  $X_1(t) = \mathcal{E}_{1,t}(W_1)$ . A similar operation is done at node 2. On the other hand, node 1 uses a decoding function  $\mathcal{D}_1 : \mathcal{W}_1 \times \mathcal{Y}^n \mapsto \mathcal{W}_1$  to decode  $\hat{W}_1 = \mathcal{D}_1(W_1, Y_1(1), \dots, Y_1(n))$ , and similarly, node 2 decodes  $\hat{W}_2 = \mathcal{D}_2(W_2, Y_2(1), \dots, Y_2(n))$ . The collection of the message sets, encoders, and decoders is known as a code for the TWC. The rates  $R_1$  and  $R_2$  are said to be achievable if there exist a sequence of codes so that  $P_{e,n} \rightarrow 0$  as  $n \rightarrow \infty$ , where  $P_{e,n}$  is the average probability of decoding error where  $W_i \neq \hat{W}_i$  for  $i \in \{1, 2\}$ . The closure of the convex hull of the set of all achievable rate pairs  $(R_1, R_2)$  is the capacity region  $\mathcal{C}$  of the TWC, and is the subject of investigation of this paper.

Continuous-alphabet TWCs are defined similarly, with  $\mathcal{X}_i, \mathcal{Y}_i \in \mathbb{R}$ ,  $i \in \{1, 2\}$ .

### B. Shannon's Bounds

Let  $(X_1, X_2)$  be distributed over  $\mathcal{X}_1 \times \mathcal{X}_2$  according to a distribution  $P_{X_1, X_2}$ . Define  $\mathcal{P}$  as the collection of all distributions  $P_{X_1, X_2}$ . Then, the capacity of the TWC is outer-bounded by [9]

$$\mathcal{C} \subseteq \mathcal{R}_o \triangleq \overline{\text{CH}} \left( \bigcup_{P_{X_1, X_2} \in \mathcal{P}} \mathcal{R}(P_{X_1, X_2}) \right) \quad (1)$$

where  $\overline{\text{CH}}(\cdot)$  is the closure of the convex-hull, and  $\mathcal{R}(P_{X_1, X_2})$  is the set of  $(R_1, R_2) \in \mathbb{R}_+^2$  with

$$R_1 \leq I(X_1; Y_2 | X_2), \text{ and } R_2 \leq I(X_2; Y_1 | X_1), \quad (2)$$

with  $(X_1, X_2)$  distributed according to  $P_{X_1, X_2}$ .

On the other hand, define  $\mathcal{P}_p$  as the collection of all distributions that factor as  $P_{X_1, X_2} = P_{X_1} P_{X_2}$ , the product of the marginals of  $X_1$  and  $X_2$ . Then, the capacity of the TWC is inner-bounded by [9]

$$\mathcal{C} \supseteq \mathcal{R}_i \triangleq \overline{\text{CH}} \left( \bigcup_{P_{X_1, X_2} \in \mathcal{P}_p} \mathcal{R}(P_{X_1, X_2}) \right). \quad (3)$$

If the two bounds coincide, then the capacity achieving distributions of  $(X_1, X_2)$  (forming the boundary of  $\mathcal{C}$ ) are product distributions. This means that adaptation is not needed

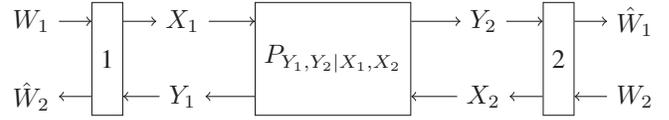


Fig. 1: A memoryless two-way channel model.

to achieve the capacity, which is interesting since nonadaptive coding is simpler than adaptive coding. In what follows, we prove that Shannon's bounds coincide for a class of TWCs, referred to as injective semi-deterministic (ISD) and defined precisely in the sequel.

### III. CONDITIONS FOR THE OPTIMALITY OF $\mathcal{R}_i$

Sufficient conditions under which the outer and inner bounds coincide have been given by Shannon in [9]. While Shannon's conditions are given in terms of the conditional probability  $P_{Y_1, Y_2 | X_1, X_2}$ , we provide conditions in terms of the conditional entropies, which are simpler and suffice for the purposes of this paper.

*Theorem 1:* The bounds  $\mathcal{R}_o$  and  $\mathcal{R}_i$  coincide if for all  $P_{X_1, X_2} = P_{X_2} P_{X_1 | X_2}$ , the following holds for  $i \in \{1, 2\}$ :

- c.1)  $H(Y_i | X_1, X_2)$  is invariant with respect to  $P_{X_1 | X_2}$ , and
- c.2)  $H(Y_i | X_i) \leq H(\bar{Y}_i | \bar{X}_i)$ , where  $\bar{Y}_i$  is the channel output corresponding to independent inputs  $\bar{X}_1$  and  $\bar{X}_2$  distributed according to  $P_{\bar{X}_1}$  and  $P_{\bar{X}_2} = P_{X_2}$ , respectively, for some  $P_{\bar{X}_1}$ .

*Proof:* To show that the bounds coincide under conditions c.1 and c.2, consider a distribution  $P_{X_1, X_2} = P_{X_2} P_{X_1 | X_2}$ , with mutual information  $I(X_1; Y_2 | X_2)$  equal to  $H(Y_2 | X_2) - H(Y_2 | X_1, X_2)$ . Under condition c.1, we have

$$\begin{aligned} H(Y_2 | X_1, X_2) &= \mathbb{E}_{X_2} \mathbb{E}_{X_1 | X_2} [H(Y_2 | X_1 = x_1, X_2 = x_2)] \\ &= \mathbb{E}_{\bar{X}_2} \mathbb{E}_{\bar{X}_1} [H(\bar{Y}_2 | \bar{X}_1 = \bar{x}_1, \bar{X}_2 = \bar{x}_2)] \\ &= H(\bar{Y}_2 | \bar{X}_1, \bar{X}_2), \end{aligned} \quad (4)$$

since  $\bar{X}_2$  has the same distribution as  $X_2$ , and since  $H(Y_2 | X_1, X_2)$  is invariant with respect to  $P_{X_1 | X_2}$ . Combining this with condition c.2 leads to

$$I(X_1; Y_2 | X_2) = H(Y_2 | X_2) - H(\bar{Y}_2 | \bar{X}_1, \bar{X}_2) \quad (5)$$

$$\leq H(\bar{Y}_2 | \bar{X}_2) - H(\bar{Y}_2 | \bar{X}_1, \bar{X}_2) \quad (6)$$

$$= I(\bar{X}_1; \bar{Y}_2 | \bar{X}_2). \quad (7)$$

Similarly,  $I(X_2; Y_1 | X_1) \leq I(\bar{X}_2; \bar{Y}_1 | \bar{X}_1)$ . Thus, the rates  $R_1$  and  $R_2$  corresponding to any distribution  $P_{X_1, X_2}$  are jointly maximized by the rates corresponding to a product distribution  $P_{\bar{X}_1} P_{\bar{X}_2}$ . Hence,  $\mathcal{R}_o \subseteq \mathcal{R}_i$ , and hence, the two coincide which concludes the proof. ■

We note that there are two main differences between these conditions and Shannon's symmetry conditions [9], which also imply that the inner and outer bounds coincide. First, c.1 demands invariance of  $H(Y_i | X_1, X_2)$  with respect to  $P_{X_1 | X_2}$ , whereas Shannon's conditions demand invariance of  $P_{Y_1, Y_2 | X_1, X_2}$  under permutations of the labels of  $x_1$  (or  $x_2$ ),  $y_1$ , and  $y_2$ . In this sense, c.1 is stricter. On the other hand,

Shannon's conditions demand the invariance of  $P_{Y_i|X_i}$  under the same permutation of labels, whereas c.2 demands the existence of an entropy maximizing  $P_{X_1}$  for a given  $P_{X_2}$ .

Interestingly, the exemplary channels given by Shannon where adaptation is not necessary, namely, the examples in [9, Fig. 2, Fig. 4, & Tab. II], satisfy both Shannon's symmetry conditions and our conditions in Theorem 1.<sup>1</sup> Note that these conditions do not necessarily imply that the capacity region is rectangular. Indeed, while  $\mathcal{R}(P_{X_1, X_2})$  is rectangular, the union thereof need not be so. A simple counterexample is Shannon's example in [9, Tab. II].

Conditions c.1 and c.2 can be generalized to continuous-alphabet channels by replacing entropy with differential entropy, and using the standard discretization procedure [12, Sec. 3.4]. These conditions are used next to establish the capacity of the class of ISD TWC with input-independent noise.

#### IV. INJECTIVE SEMI-DETERMINISTIC CHANNELS WITH INPUT-INDEPENDENT NOISE

For  $i, j \in \{1, 2\}$ ,  $i \neq j$ , consider the functions

$$g_i : \mathcal{X}_j \times \mathcal{Z}_i \mapsto \mathcal{T}_i, \text{ and } f_i : \mathcal{X}_i \times \mathcal{T}_i \mapsto \mathcal{Y}_i, \quad (8)$$

for some sets  $\mathcal{Z}_i$  and  $\mathcal{T}_i$ . Further, assume that  $f_i(X_i, T_i)$  is injective in  $T_i$ , i.e., for every  $x_i \in \mathcal{X}_i$ ,  $f_i(x_i, t_i)$  is one-to-one in  $t_i \in \mathcal{T}_i$ , and that  $g_i(X_j, Z_i)$  is injective in  $Z_i$ , i.e., for every  $x_j \in \mathcal{X}_j$ ,  $g_i(x_j, z_i)$  is one-to-one in  $z_i \in \mathcal{Z}_i$ .<sup>2</sup> We define ISD TWCs with input-independent noise as follows.

*Definition 1 (ISD TWC):* The injective semi-deterministic TWC with input-independent noise is one with

$$Y_i = f_i(X_i, T_i), \text{ and } T_i = g_i(X_j, Z_i), \quad (9)$$

$i, j \in \{1, 2\}$ ,  $i \neq j$ , where  $Z_1 \in \mathcal{Z}_1$  and  $Z_2 \in \mathcal{Z}_2$  are (possibly dependent) random variables independent of  $X_1$  and  $X_2$  (Fig. 2).

For this class, we have the following statement.

*Theorem 2:* For the class of ISD TWCs with input-independent noise, conditions c.1 and c.2 are satisfied and the capacity region is given by  $\mathcal{C} = \mathcal{R}_i$ .

*Proof:* We have

$$H(Y_1|X_1, X_2) = H(f_1(X_1, g_1(X_2, Z_1))|X_1, X_2) \quad (10)$$

$$= H(g_1(X_2, Z_1)|X_1, X_2) \quad (11)$$

$$= H(Z_1|X_1, X_2) \quad (12)$$

$$= H(Z_1), \quad (13)$$

which follows from the injectivity of  $f_1$  and  $g_1$ , and the independence of  $Z_1$  and  $(X_1, X_2)$ . This is independent of  $P_{X_1|X_2}$ . Similarly,  $H(Y_2|X_1, X_2) = H(Z_2)$ , and hence, c.1 is satisfied. On the other hand,

$$H(Y_1|X_1) = H(f_1(X_1, g_1(X_2, Z_1))|X_1) \quad (14)$$

$$= H(g_1(X_2, Z_1)|X_1) \quad (15)$$

$$\leq H(g_1(X_2, Z_1)), \quad (16)$$

<sup>1</sup>It would be interesting to find channels that satisfy c.1 and c.2 or Shannon's symmetry conditions, but not both, showing that the conditions do not generally coincide.

<sup>2</sup>Related conditions for one-way channels with feedback were given in [16].

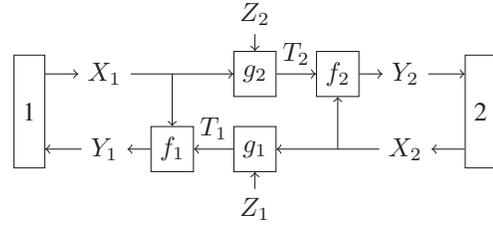


Fig. 2: A semi-deterministic TWC.

again by the injectivity of  $f_1$ . This upper bound is equal to the entropy of the output  $Y_1$  when the inputs are independent. Similarly,  $H(Y_2|X_2)$  is upper-bounded by  $H(g_2(X_1, Z_2))$ , the entropy of  $Y_2$  when  $X_1$  and  $X_2$  are independent. Thus, c.2 is also satisfied. Thus the bounds  $\mathcal{R}_i$  and  $\mathcal{R}_o$  coincide by Thm. 1 and the statement follows. ■

As a corollary of Thm. 2, we have the following.

*Corollary 1:* For the class of ISD TWCs with input-independent noise, the capacity region is rectangular given by the set of  $(R_1, R_2) \in \mathbb{R}_+^2$  so that, for  $i, j \in \{1, 2\}$ ,  $i \neq j$ :

$$R_i \leq \max_{P_{X_i}} H(g_j(X_i, Z_j)) - H(Z_j) \quad (17)$$

*Proof:* The proof follows directly from the fact that  $H(Y_i|X_1, X_2)$  is independent of  $P_{X_1, X_2}$ , and that  $H(Y_i|X_i) \leq H(g_i(X_j, Z_i))$  which is maximized by the same  $P_{X_j}$  independent of  $P_{X_i}$ , for  $i, j \in \{1, 2\}$ ,  $i \neq j$ . ■

The same result can be shown for continuous channels. This class subsumes additive channels with input-independent noise, leading to the following.

*Corollary 2:* The capacity region of an additive TWC with input-independent noise, where  $Y_i = a_i X_i + X_j + Z_i$ ,  $i \in \{1, 2\}$ ,  $i \neq j$ , is given by the rectangle defined by

$$0 \leq R_i \leq \max_{P_{X_i}} I(X_i; X_i + Z_j). \quad (18)$$

This statement follows since this channel can be obtained by replacing  $f_i(X_i, T_i)$  and  $g_i(X_j, Z_i)$  by  $a_i X_i + T_i$  and  $X_j + Z_i$ , respectively. Both  $f_i$  and  $g_i$  are injective given  $X_i$  and  $X_j$ , respectively. Corollary 2 reproduces [3, Remark 2] if we set  $a_i = 1$  and  $\mathcal{X}_i = \mathcal{Z}_i = \{0, 1, \dots, q-1\}$  for  $i \in \{1, 2\}$ , and use modulo- $q$  addition.

The class in Def. 1 subsumes other channels such as multiplicative channels with input-independent noise, where  $Y_i = X_i \cdot X_j \cdot Z_i$  such that  $0 \notin \mathcal{X}_i \cup \mathcal{Z}_i$ . In this case,  $f_i(X_i, T_i)$  and  $g_i(X_j, Z_i)$  are given by  $X_i \cdot T_i$  and  $X_j \cdot Z_i$ , which are injective given  $X_i$  and  $X_j$ , respectively. This leads to the following corollary.

*Corollary 3:* The capacity region of a multiplicative TWC with input-independent noise, where  $Y_i = X_i \cdot X_j \cdot Z_i$ ,  $i \in \{1, 2\}$ ,  $i \neq j$ , so that  $0 \notin \mathcal{X}_i \cup \mathcal{Z}_i$ , is given by the rectangle defined by

$$0 \leq R_i \leq \max_{P_{X_i}} I(X_i; X_i \cdot Z_j). \quad (19)$$

The binary multiplier channel [9] for which the bounds do not coincide does not belong to this class because  $0 \in \mathcal{X}_i$  (contrary

to Corollary 3), and thus, the channel is not injective. Capacity outer bounds for this channel tighter than Shannon's have been given [10].

These corollaries show some interesting features of Theorem 2. First, it can be used to characterize the capacity of an additive channel with input-independent noise without resorting to *convolutional divisibility* as in [2]. Note that convolutional divisibility is known to hold for the exponential-family [15], but might not hold for some other distributions. Second, while convolutional divisibility might be useful for additive channels, Theorem 2 is useful for nonadditive channels in addition to additive ones. Third, Theorem 2 shows that, among additive continuous-alphabet TWC with input-independent noise, the exponential family TWC for which adaptation is useless [2], [13] is not the only channel with this property. Next, we give examples of additive TWC not of the exponential family for which Shannon's bounds coincide.

### A. Cauchy Two-way Channel Capacity

Before we discuss the Cauchy TWC, let us briefly introduce the Cauchy one-way channel.

1) *Cauchy One-way Channel*: This channel has input  $X \in \mathbb{R}$ , noise  $Z \in \mathbb{R}$ , and output  $Y = X + Z$ , where noise is distributed as

$$P_Z(z) = \left[ \pi \gamma_i \left( 1 + \left( \frac{z}{\gamma} \right)^2 \right) \right]^{-1}, \quad (20)$$

denoted  $P_{\text{Cauchy}}(0, \gamma)$  henceforth. Here,  $\gamma$  is the dispersion parameter of the Cauchy distributed noise, whose location parameter is equal to zero without loss of generality. This channel models situations with impulsive noise or interference caused by Poisson distributed interferers, where the second moment of noise can be infinite and not suitable as a power measure [17]. In this case, the constraint

$$C(A, \gamma) \triangleq \mathbb{E}_X \left[ \log \left[ \left[ \frac{A + \gamma}{A} \right]^2 + \left[ \frac{X}{A} \right]^2 \right] \right] \leq \log(4), \quad (21)$$

for some  $A \geq \gamma$ , can be viewed as a power constraint. In particular, under this constraint, there exists  $k \in [\gamma, A]$  so that  $C(k, \gamma) = \log(4)$ , and the quantity  $k - \gamma$  can be viewed as a measure of power of  $X$ , while the quantity  $k$  can be viewed as the power of  $Y$  as discussed in [17]. The capacity of this channel is given by  $\log\left(\frac{A}{\gamma}\right)$ .

Note that this channel is not of the exponential family. Next, we define the Cauchy TWC and characterize its capacity.

2) *Cauchy Two-way Channel*: This is defined as a channel where, for  $i, j \in \{1, 2\}$ ,  $i \neq j$ ,

$$Y_i = a_i X_i + X_j + Z_i, \quad (22)$$

with  $X_i, Z_i \in \mathbb{R}$ , and where noise  $Z_i$  follows a Cauchy distribution  $P_{\text{Cauchy}}(0, \gamma_i)$ , and is independent of  $X_1$  and  $X_2$ . The channel input  $X_i$  must satisfy a logarithmic constraint

$$C(A_i, \gamma_j) \leq \log(4), \quad (23)$$

where  $C(A_i, \gamma_j)$  is as defined in (21).

This channel is an ISD TWC with input-independent noise, for which capacity can be characterized as in Corollary 2. In fact, its capacity can be characterized in closed form as given in the following theorem.

*Theorem 3*: The capacity region of the Cauchy TWC is given by the set of  $(R_1, R_2) \in \mathbb{R}_+^2$  so that  $R_i \leq \log\left(\frac{A_i}{\gamma_j}\right)$ ,  $i \neq j$ , and is achieved by  $X_i$  distributed according to  $P_{\text{Cauchy}}(0, A_i - \gamma_j)$ .

*Proof*: Since this channel belongs to the ISD class, its capacity is given by  $\mathcal{R}_i$ , described by  $0 \leq R_i \leq \max_{P_{X_i}} h(Y_j) - h(Z_j)$ ,  $i \neq j$  (cf. Corollary 1). The maximization here is subject to (23). Consider an input  $X_j$  so that

$$\mathbb{E}_{X_j} \left[ \log \left( \left( \frac{k_j + \gamma_i}{k_j} \right)^2 + \left( \frac{X_j}{k_j} \right)^2 \right) \right] = \log(4), \quad (24)$$

for some  $k_j \in [\gamma_i, A_j]$ . Then, we have [17]

$$\mathbb{E}_{Y_i} \left[ \log \left( 1 + \left( \frac{Y_i}{k_j} \right)^2 \right) \right] = \log(4). \quad (25)$$

Under this constraint, the entropy of  $Y_i$  is maximized when  $Y_i$  is distributed according to  $P_{\text{Cauchy}}(0, k_j)$ , which is possible if  $X_j$  is distributed according to  $P_{\text{Cauchy}}(0, k_j - \gamma_i)$  and satisfies (24). The entropy of a Cauchy random variable  $P_{\text{Cauchy}}(0, \mu)$  is  $\log(4\pi\mu)$ . Thus, the rate  $R_j$  satisfies

$$R_j \leq \max_{P_{X_j}} h(Y_i) - h(Z_i) \quad (26)$$

$$= h(Y_i) |_{X_j \sim P_{\text{Cauchy}}(0, k_j - \gamma_i)} - h(Z_i) \quad (27)$$

$$= \log\left(\frac{k_j}{\gamma_i}\right). \quad (28)$$

Since this is increasing in  $k_j$ , it is maximized by  $k_j = A_j$  leading to the desired result. ■

This channel provides an interesting example of a TWC not of the exponential family, where adaptation is not necessary. Note that this result is obtained without using a saddle point argument as in [2]. In fact, saddle point admissibility as defined in [2], [15] is a strong condition which is not required in our problem. Conditions c.1 and c.2 suffice for specifying the capacity region. On the other hand, the fact that the Cauchy input is max-entropic is a key element for evaluating the region in closed form. Another key element is the independence between the noise and the inputs. The lack of this independence does not imply the necessity of adaptation as explained next.

### B. Input-dependent Noise

While channels with input-dependent noise do not belong to the class defined in Def. 1, they might still satisfy c.1 and c.2. An example is the channel in [9, Tab II] which is ISD but with input-dependent noise. Namely, in this example, the inputs and outputs are binary, with

$$Y_1 = X_2, \quad \text{and} \quad Y_2 = \begin{cases} X_1 & \text{if } X_2 = 0, \\ N_2 & \text{if } X_2 = 1, \end{cases} \quad (29)$$

where  $N_2 \in \{0, 1\}$  is Bern(1/2) (Bernoulli distributed). This can be modeled as an ISD channel where

$$g_1(X_2, Z_1) = X_2, \quad g_2(X_1, Z_2) = X_1 + Z_2, \quad (30)$$

$$f_1(T_1, X_1) = T_1, \quad f_2(X_2, T_2) = T_2, \quad (31)$$

$\mathcal{Z}_1 = \{0\}$ ,  $\mathcal{Z}_2 = \{0, 1\}$ ,  $Z_2 = X_2(N_2 + X_1)$ , and where additions are in the binary field. Here, noise is input-dependent, yet, c.1 and c.2 are satisfied, and adaptation is not necessary.

Another example of this sort with practical relevance is an additive TWC with input-dependent Gaussian noise, with

$$Y_i = a_i X_i + X_j + \sqrt{X_j} \tilde{Z}_i + \hat{Z}_i, \quad (32)$$

$i, j \in \{1, 2\}$ ,  $i \neq j$ , where  $X_i, \tilde{Z}_i, \hat{Z}_i \in \mathbb{R}$ ,  $\tilde{Z}_i$  and  $\hat{Z}_i$  are Gaussian noises with zero mean and variances  $\hat{\sigma}_i^2$  and  $\tilde{\sigma}_i^2$ , respectively, independent of  $X_1$  and  $X_2$ , and  $X_i$  satisfies a cost constraint  $\mathbb{E}[c(X_i)] \leq P_i$  for some cost function  $c(\cdot)$ . This channel is not covered by Han's result [13] where input-independent noise was assumed. It models a two-way optical wireless channel where the dependence of the noise on the input is due to the random nature of photon emission [18].<sup>3</sup> This channel can be modeled as an ISD channel with  $Z_i = \sqrt{X_j} \tilde{Z}_i + \hat{Z}_i$ , and  $Y_i = f(X_i, T_i) = a_i X_i + T_i$  and  $T_i = g_i(X_j, Z_i) = X_j + Z_i$  which are injective in  $T_i$  and  $Z_i$  respectively. The main difference with Def. 1 is that noise  $Z_i$  is input-dependent. Nevertheless, the capacity of this channel can be determined. In particular,

$$I(X_1; Y_2 | X_2) = h(Y_2 | X_2) - h(\sqrt{X_1} \tilde{Z}_2 + \hat{Z}_2 | X_1) \quad (33)$$

$$\leq h(X_1 + \sqrt{X_1} \tilde{Z}_2 + \hat{Z}_2) - h(\sqrt{X_1} \tilde{Z}_2 + \hat{Z}_2 | X_1) \quad (34)$$

$$\leq \max_{P_{X_1}} \left[ h(X_1 + \sqrt{X_1} \tilde{Z}_2 + \hat{Z}_2) - \mathbb{E}_{X_1} \left[ \frac{1}{2} \log(2\pi e(x_1 \tilde{\sigma}_2^2 + \hat{\sigma}_2^2)) \right] \right] \triangleq \bar{C}_1, \quad (35)$$

and similarly  $I(X_2; Y_1 | X_1) \leq \bar{C}_2$  with  $\bar{C}_2$  defined similarly with interchanged indexes ( $1 \leftrightarrow 2$ ). But these upper bounds are achievable using independent  $X_1$  and  $X_2$ . Thus, the capacity of this channel is given by the rectangular region defined by

$$0 \leq R_1 \leq \bar{C}_1, \quad \text{and} \quad 0 \leq R_2 \leq \bar{C}_2, \quad (36)$$

which is achievable without adaptation.

## V. CONCLUSION

We have studied the capacity of the class of injective semi-deterministic two-way channels with input-independent noise. For this class, we have characterized the capacity region, which is achieved by independent coding, i.e., adaptation is not necessary. Based on this result, we conclude that the additive exponential family type two-way channel is not the only continuous additive two-way channel for which adaptation is not necessary, but this is the case for a more general class of

<sup>3</sup>Here,  $a_i X_i$  is caused by light scattered towards the detector. Generally, this leads to an additional noise term  $\sqrt{X_i} \tilde{Z}'_i$  with  $\tilde{Z}'_i$  Gaussian. However, if this scattering is weak, then the associated noise term is negligible.

additive two-way channels. As an example, we used this result to characterize the capacity of the Cauchy two-way channel with a logarithmic constraint. We also gave an example of a two-way channel with input-dependent noise for which adaptation is not necessary. This implies that this class can be further generalized to encompass such channels, leading to a more general statement. It remains to note that the conditions given in this paper under which capacity is characterized are still sufficient but not necessary.

## ACKNOWLEDGEMENT

The authors would like to acknowledge the organizers of the 2016 KAUST-NSF Conference on Electronic Materials, Devices, and Systems for a Sustainable Future, where the initial conversations on this topic began.

## REFERENCES

- [1] A. Sabharwal, P. Schniter, D. Guo, D. W. Bliss, S. Rangarajan, and R. Wichman, "In-band full-duplex wireless: Challenges and opportunities," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 9, pp. 1637–1652, Sept. 2014.
- [2] L. R. Varshney, "Two way communication over exponential family type channels," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, July 2013, pp. 2795–2799.
- [3] L. Song, F. Alajaji, and T. Linder, "Adaptation is useless for two discrete additive-noise two-way channels," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, July 2016, pp. 1854–1858.
- [4] L. Ong, G. Lechner, S. J. Johnson, and C. M. Kellett, "The three-user finite-field multi-way relay channel with correlated source," *IEEE Trans. Commun.*, vol. 61, no. 8, pp. 3125–3135, Aug. 2013.
- [5] A. Chaaban, H. Maier, A. Sezgin, and R. Mathar, "Three-way channels with multiple unicast sessions: Capacity approximation via network transformation," *IEEE Trans. Inf. Theory*, vol. 62, no. 12, pp. 7086–7102, Dec. 2016.
- [6] Z. Cheng and N. Devroye, "Two-way networks: When adaptation is useless," *IEEE Trans. Inf. Theory*, vol. 60, no. 3, pp. 1793–1813, March 2014.
- [7] Y. Song, N. Devroye, H.-R. Shao, and C. Ngo, "Lattice coding for the two-way two-relay channel," *submitted to IEEE Trans. Inf. Theory*, arXiv:1212.1198 [cs.IT], Dec. 2012.
- [8] A. Chaaban and A. Sezgin, "Multi-way communications: An information theoretic perspective," *Found. Trends Commun. Inf. Theory*, vol. 12, no. 3-4, pp. 185–371, 2015.
- [9] C. E. Shannon, "Two-way communication channels," in *Proc. of 4th Berkeley Symp. Math. Stat. Probab.*, vol. 1, 1961, pp. 611–644.
- [10] A. P. Hekstra and F. M. J. Willems, "Dependence balance bounds for single output two-way channels," *IEEE Trans. Inf. Theory*, vol. 35, no. 1, pp. 44–53, Jan. 1989.
- [11] J. Schalkwijk, "The binary multiplying channel—A coding scheme that operates beyond Shannon's inner bound region," *IEEE Trans. Inf. Theory*, vol. 28, no. 1, pp. 107–110, Jan. 1982.
- [12] A. E. Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge University Press, 2011.
- [13] T. S. Han, "A general coding scheme for the two-way channel," *IEEE Trans. Inf. Theory*, vol. 30, no. 1, pp. 35–44, Jan. 1984.
- [14] S. Verdú, "The exponential distribution in information theory," *Probl. Inf. Transm.*, vol. 32, no. 1, pp. 100–111, Jan.-Mar. 1996.
- [15] T. P. Coleman and M. Raginsky, "Mutual information saddle points in channels of exponential family type," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Austin, TX, Jun. 2010, pp. 1355–1359.
- [16] F. Alajaji and T. Fuja, "Effect of feedback on the capacity of discrete additive channels with memory," in *Proc. IEEE Int. Symp. Inf. Theory*, June 1994, p. 464.
- [17] J. Fahs and I. Abou-Faycal, "A Cauchy input achieves the capacity of a Cauchy channel under a logarithmic constraint," in *Proc. IEEE Int. Symp. Inf. Theory*, June 2014, pp. 3077–3081.
- [18] S. M. Moser, "Capacity results of an optical intensity channel with input-dependent Gaussian noise," *IEEE Trans. Inf. Theory*, vol. 58, no. 1, pp. 207–223, Jan. 2012.