

# The Capacity Region of the Arbitrarily Varying MAC: With and Without Constraints

Uzi Pereg and Yossef Steinberg

Department of Electrical Engineering, Technion, Haifa 32000, Israel.

Email: uzipereg@campus.technion.ac.il, ysteinbe@ee.technion.ac.il

## Abstract

We determine both the random code capacity region and the deterministic code capacity region of the arbitrarily varying multiple access channel (AVMAC) under input and state constraints. The underlying assumption is that zero-rate transmission can be arbitrary in the deterministic setting as well, where there is no shared randomness. As opposed to the random code capacity region, the deterministic code capacity region can be non convex. For the AVMAC without constraints, the characterization due to Ahlswede and Cai is complete except for two cases, pointed out in the literature as an open problem. The missing piece is obtained as a special case of our results.

## Index Terms

Arbitrarily varying channel, multiple access, minimax, deterministic code, symmetrizability, random code, input and state constraints.

## I. INTRODUCTION

The arbitrarily varying multiple access channel (AVMAC) without constraints was first considered by Jahn [33, 34], to describe a communication network with unknown statistics, that may change over time. It is especially relevant to uplink communication in the presence of an adversary, or a *jammer*, attempting to disrupt communication. Another scenario is that one of multiple users becomes adversarial and attacks the other users [43, 44]. Jahn established the ‘divided-randomness capacity region’ [33, 34], namely the capacity region achieved when each encoder shares randomness with the decoder independently, and showed that the AVMAC inherits some of the properties of its single user counterpart. In particular, the divided-randomness capacity region is not necessarily achievable using deterministic codes [8]. Furthermore, Jahn showed that the deterministic code capacity region either coincides with the divided-randomness capacity region or else, it has an empty interior [33, 34]. This phenomenon is an analogue of Ahlswede’s dichotomy property [1]. Therefore, in order to calculate the deterministic code capacity region, it is essential to confirm that the capacity region has a non empty interior, *i.e.* positive rates are achievable using deterministic codes. Gubner [24] presented three computable conditions which are necessary for this to hold, and conjectured that they are also sufficient. Then, Ahlswede and Cai [3] confirmed Gubner’s conjecture [24], implying that Gubner’s conditions are both necessary and sufficient for a non empty capacity region. As Wiese and Boche recognized, the case where exactly one of the users has zero capacity has remained an open problem [48, Remark 9], until now.

Furthermore, constraints are known to have a drastic effect on the behavior of the single user AVC [17], while the effect on the AVMAC has never been established. Csiszár and Narayan [17] considered the single user AVC when input and state constraints are imposed on the user and the jammer, respectively. Such constraints are often due to power limitations of the transmitter and the jamming signal. Not only the constrained setting provokes serious technical difficulties analytically, but also, as shown in [17], there is a significant effect on the behavior of the deterministic code capacity. Specifically, it is shown in [17] that dichotomy in the notion of [1] no longer holds when state constraints are imposed on the jammer. That is, the deterministic code capacity can be lower than the random code capacity, and yet non-zero. As for the AVMAC under constraints, Gubner and Hughes [27] determined the divided-randomness capacity region. Results on the Gaussian AVMAC were recently presented in a talk [31]. Solved examples can be found in [25, 26] as well.

Other relevant settings include the AVMAC with conferencing encoders [48, 47, 9], list codes [40, 9, 11], fading [45, 10], and an eavesdropper [29, 5, 13]. Among the models of channel uncertainty are also the compound multiple access channel [16, 38, 49, 51, 39] and the random parameter multiple access with side information [19, 12, 46, 36, 37]. After the publication of this work, Sangwan *et al.* [43] considered a multiple access channel with three users, where one of the users is possibly adversarial, yet the identity of the jamming user is not known in advance (see also [44]).

In this work, we consider the AVMAC when input and state constraints are imposed on the users and the jammer, respectively. We give full characterization for both the random code capacity region and the deterministic code capacity region. The underlying assumption is that zero-rate transmission can be arbitrary in the deterministic setting as well, where there is no shared randomness. In particular, the encoder can simulate a random transmission, as long as there is no shared randomness

between the parties. This assumption is generally considered to be natural for real-life communication systems (see *e.g.* remark in [16, p. 282]), whereas shared randomness is often impractical [34, 6]. Nonetheless, the analysis without shared randomness is a lot more challenging. When state constraints are imposed, the operational time sharing argument does not apply to the AVMAC. Roughly speaking, using a code over a part of the blocklength effectively increases the state constraint and loosens the restriction on the jammer over this period of time. Thus, it becomes essential to replace the operational time sharing argument with coded time sharing [28]. Our decoder is then a coded time sharing variant of Ahlswede and Cai's decoding rule [3], while the time sharing sequence is deterministic and known to the jammer as well. Yet, a fundamental difference between our coding scheme and the one in [3] arises from the dichotomy discrepancy, since Ahlswede and Cai only showed achievability of positive rates  $R_1 = R_2 = \varepsilon > 0$ , proving that the capacity region has a non-empty interior. However, for the AVMAC under constraints, dichotomy does not apply and achievability of positive rates is insufficient. Hence, in our problem, proving achievability is more demanding. Hereby, the codebooks construction and the analysis are based on generalization of the techniques by Csiszár and Narayan [17], along with the insights of Ahlswede and Cai [3]. The converse proof involves observations by Gubner [24] as well. As a special case, we obtain a full characterization of the capacity region of the AVMAC without constraints, filling the gap left by Ahlswede and Cai [3].

## II. DEFINITIONS AND PREVIOUS RESULTS

We use the following notation conventions throughout. Calligraphic letters  $\mathcal{X}, \mathcal{S}, \mathcal{Y}, \dots$  are used for finite sets. Lowercase letters  $x, s, y, \dots$  stand for constants and values of random variables, and uppercase letters  $X, S, Y, \dots$  stand for random variables. The distribution of a random variable  $X$  is specified by a probability mass function (pmf)  $P_X(x) = p(x)$  over a finite set  $\mathcal{X}$ . The set of all pmfs over  $\mathcal{X}$  is denoted by  $\mathcal{P}(\mathcal{X})$ . We use  $x^j = (x_1, x_2, \dots, x_j)$  to denote a sequence of letters from  $\mathcal{X}$ . A random sequence  $X^n$  and its distribution  $P_{X^n}(x^n)$  are defined accordingly. The type  $\hat{P}_{x^n}$  of a given sequence  $x^n$  is defined as the empirical distribution  $\hat{P}_{x^n}(a) = N(a|x^n)/n$  for  $a \in \mathcal{X}$ , where  $N(a|x^n)$  is the number of occurrences of the symbol  $a$  in the sequence  $x^n$ . A type class is denoted by  $\mathcal{T}^n(\hat{P}) = \{x^n : \hat{P}_{x^n} = \hat{P}\}$ . For a pair of integers  $i$  and  $j$ ,  $1 \leq i \leq j$ , we define the discrete interval  $[i : j] = \{i, i+1, \dots, j\}$ . In the continuous case, we use the cumulative distribution function  $F_Z(z) = \Pr(Z \leq z)$  for  $z \in \mathbb{R}$ , or alternatively, the probability density function (pdf)  $f_Z(z)$ , when it exists.

### A. Channel Description

A state-dependent discrete memoryless multiple access channel (MAC)  $(\mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{S}, W_{Y|X_1, X_2, S}, \mathcal{Y})$  consists of finite input alphabets  $\mathcal{X}_1$  and  $\mathcal{X}_2$ , state alphabet  $\mathcal{S}$ , output alphabet  $\mathcal{Y}$ , and a conditional pmf  $W_{Y|X_1, X_2, S}$  over  $\mathcal{Y}$ . The channel is memoryless without feedback, and therefore  $W_{Y^n|X_1^n, X_2^n, S^n}(y^n|x_1^n, x_2^n, s^n) = \prod_{i=1}^n W_{Y|X_1, X_2, S}(y_i|x_{1,i}, x_{2,i}, s_i)$ . The AVMAC is a MAC with a state sequence of unknown distribution, not necessarily independent nor stationary. That is,  $S^n \sim q(s^n)$  with an unknown joint pmf  $q(s^n)$  over  $\mathcal{S}^n$ . In particular,  $q(s^n)$  could give mass 1 to some state sequence  $s^n$ . The AVMAC is denoted by  $\mathcal{A} = \{W_{Y|X_1, X_2, S}\}$ .

The compound MAC is used as a tool in the analysis. Different models of compound MACs are described in the literature [16, 38]. Here, the compound MAC is a channel with a discrete memoryless state, where the state distribution  $q(s)$  is not known in exact, but rather belongs to a family of distributions  $\mathcal{Q}$ , with  $\mathcal{Q} \subseteq \mathcal{P}(\mathcal{S})$ . That is, the state sequence  $S^n$  is independent and identically distributed (i.i.d.) according to  $q(s)$ , for some pmf  $q \in \mathcal{Q}$ . We note that this differs from the classical definition of the compound channel, as in [16], where the state is fixed throughout the transmission. The compound MAC is denoted by  $\mathcal{A}^{\mathcal{Q}}$ .

### B. Coding

We introduce some preliminary definitions, starting with the definitions of a deterministic code and a random code for the AVMAC  $\mathcal{A}$  under input and state constraints.

**Definition 1 (Code).** A  $(2^{nR_1}, 2^{nR_2}, n)$  code for the AVMAC  $\mathcal{A}$  consists of the following; two message sets  $[1 : 2^{nR_1}]$  and  $[1 : 2^{nR_2}]$ , where  $2^{nR_1}$  and  $2^{nR_2}$  are assumed to be integers, two encoding functions  $f_1 : [1 : 2^{nR_1}] \rightarrow \mathcal{X}_1^n$  and  $f_2 : [1 : 2^{nR_2}] \rightarrow \mathcal{X}_2^n$ , and a decoding function  $g : \mathcal{Y}^n \rightarrow [1 : 2^{nR_1}] \times [1 : 2^{nR_2}]$ .

Given a pair of messages  $m_1 \in [1 : 2^{nR_1}]$  and  $m_2 \in [1 : 2^{nR_2}]$ , Encoder  $k$  transmits the codeword  $x_k^n = f_k(m_k)$ , for  $k = 1, 2$ . The decoder receives the channel output  $y^n$ , and finds an estimate of the message pair  $(\hat{m}_1, \hat{m}_2) = g(y^n)$ . We denote the code by  $\mathcal{C} = (f_1(\cdot), f_2(\cdot), g(\cdot))$ .

We proceed now to coding schemes when using stochastic-encoders stochastic-decoder pairs with common randomness. We distinguish between two classes; random codes [48] and divided-randomness codes [34].

**Definition 2 (Random code).** A  $(2^{nR_1}, 2^{nR_2}, n)$  random code for the AVMAC  $\mathcal{A}$  consists of a collection of  $(2^{nR_1}, 2^{nR_2}, n)$  codes  $\{\mathcal{C}_\gamma = (f_{1,\gamma}, f_{2,\gamma}, g_\gamma)\}_{\gamma \in \Gamma}$ , along with a probability distribution  $\mu(\gamma)$  over the code collection  $\Gamma$ . We denote such a code by  $\mathcal{C}^\Gamma = (\mu, \Gamma, \{\mathcal{C}_\gamma\}_{\gamma \in \Gamma})$ .

**Definition 3 (Divided-randomness code).** A  $(2^{nR_1}, 2^{nR_2}, n)$  divided-randomness code for the AVMAC  $\mathcal{A}$  is a random code, where the random element consists of two components, *i.e.*  $\gamma = (\gamma_1, \gamma_2)$ , one at each encoder. The components are drawn

according to a product distribution  $\mu(\gamma_1, \gamma_2) = \mu_1(\gamma_1)\mu_2(\gamma_2)$  over  $\Gamma_1 \times \Gamma_2$ . Then, User 1 sends  $x_1^n = f_{1,\gamma_1}(m_1)$ , User 2 sends  $x_2^n = f_{2,\gamma_2}(m_2)$ , and upon receiving the channel output  $y^n$ , the receiver applies a decoding mapping  $g_{\gamma_1, \gamma_2}$ . We denote such a code by  $\mathcal{C}^{\Gamma_1 \times \Gamma_2}$ .

The general random code in Definition 2 can thus be thought of as a divided-randomness code where statistical dependence between  $\gamma_1$  and  $\gamma_2$  is viable.

*Remark 1.* Our underlying assumption is that zero-rate transmission can be arbitrary in the deterministic setting as well, as long as there is no shared randomness. In particular, if User 1 has zero capacity while User 2 transmits at a positive rate, then Encoder 1 may transmit a random sequence at zero rate, *i.e.*  $x_1^n = f_1(\sigma)$  where  $\sigma \in [1 : 2^{n\varepsilon}]$  is a random parameter, which is not known to the other encoder, the decoder, nor the jammer, and the decoder is not required to recover the value of  $\sigma$ . This means that the encoders have access to  $n\varepsilon$  random bits, where  $\varepsilon > 0$  is arbitrarily small. Since the randomness is local, such an assumption is generally considered to be reasonable (see *e.g.* remark in [16, p. 282]). On the other hand, in the codes in Definitions 2 and 3, there is shared randomness between the encoders and the decoder, which is often impractical [34, 6].

One may also consider the AVMAC with stochastic encoders, *i.e.* when Encoder  $k$  transmits  $x_k^n = f_k(m_k, \sigma_k)$ , for  $m_k \in [1 : 2^{nR_k}]$ ,  $k = 1, 2$ , where  $\sigma_1$  and  $\sigma_2$  the random parameters. Then, our results apply to the stochastic encoder capacity region.

### C. Input and State Constraints

Next, we consider input constraints and state constraint, imposed on the encoders and the jammer, respectively. We note that the constraints specifications are known to both users and the jammer in this model. Let  $\phi_k : \mathcal{X}_k \rightarrow [0, \infty)$ ,  $k = 1, 2$ , and  $l : \mathcal{S} \rightarrow [0, \infty)$  be some given bounded functions, and define

$$\phi_k^n(x_k^n) = \frac{1}{n} \sum_{i=1}^n \phi_k(x_{k,i}), \quad k = 1, 2, \quad (1)$$

$$l^n(s^n) = \frac{1}{n} \sum_{i=1}^n l(s_i). \quad (2)$$

Let  $\Omega_1 > 0$ ,  $\Omega_2 > 0$ , and  $\Lambda > 0$ . Below, we specify the input constraints  $(\Omega_1, \Omega_2)$  and state constraint  $\Lambda$ , corresponding to the functions  $\phi_1^n(x_1^n)$ ,  $\phi_2^n(x_2^n)$ , and  $l^n(s^n)$ , respectively, for the AVMAC and the compound MAC.

Given input constraints  $(\Omega_1, \Omega_2)$ , the encoding functions need to satisfy

$$\phi_k^n(f_k(m_k)) \leq \Omega_k, \quad \text{for all } m_k \in [1 : 2^{nR_k}], \quad k = 1, 2. \quad (3)$$

That is, the inputs satisfy  $\phi_1^n(X_1^n) \leq \Omega_1$  and  $\phi_2^n(X_2^n) \leq \Omega_2$  with probability 1. Moving to the state constraint  $\Lambda$ , we have different definitions for the AVMAC and for the compound MAC.

The compound MAC has a constraint on average, with a memoryless state such that  $\mathbb{E}_q l(S) \leq \Lambda$ , while the AVMAC has an almost-surely constraint, with a non-stationary state sequence such that  $l^n(S^n) \leq \Lambda$  with probability 1. Explicitly, we say that a compound MAC  $\mathcal{A}^{\mathcal{Q}}$  is under a state constraint  $\Lambda$ , if the set  $\mathcal{Q}$  of state distributions is limited to  $\mathcal{Q} \subseteq \overline{\mathcal{P}}_{\Lambda}(\mathcal{S})$ , where

$$\overline{\mathcal{P}}_{\Lambda}(\mathcal{S}) \triangleq \{q(s) \in \mathcal{P}(\mathcal{S}) : \mathbb{E}_q l(S) \leq \Lambda\}. \quad (4)$$

As for the AVMAC  $\mathcal{A}$ , it is now assumed that the joint distribution of the state sequence is limited to  $q(s^n) \in \mathcal{P}_{\Lambda}(S^n)$ , where

$$\mathcal{P}_{\Lambda}(S^n) \triangleq \{q(s^n) \in \mathcal{P}(S^n) : q(s^n) = 0 \text{ if } l^n(s^n) > \Lambda\}. \quad (5)$$

This includes the case of a deterministic unknown state sequence, *i.e.* when  $q$  gives probability 1 to a particular  $s^n \in \mathcal{S}^n$  with  $l^n(s^n) \leq \Lambda$ .

We may assume without loss of generality that  $0 \leq \Omega_k \leq \phi_{k,max}$ ,  $k = 1, 2$ , and  $0 \leq \Lambda \leq l_{max}$ , where  $\phi_{k,max} = \max_{x_k \in \mathcal{X}_k} \phi_k(x_k)$ ,  $k = 1, 2$ , and  $l_{max} = \max_{s \in \mathcal{S}} l(s)$ . It is also assumed that for some  $a \in \mathcal{X}_1$ ,  $b \in \mathcal{X}_2$ , and  $s_0 \in \mathcal{S}$ ,  $\phi_1(a) = \phi_2(b) = l(s_0) = 0$ .

### D. Capacity Region Under Constraints

We move to the definition of achievable rate pairs and the capacity region of the AVMAC  $\mathcal{A}$  under input and state constraints. Deterministic codes and random codes over the AVMAC  $\mathcal{A}$  are defined as in Definition 1 and Definition 2, respectively, with the additional constraint (3) on the codebook.

Define the conditional probability of error of a code  $\mathcal{C}$  given a state sequence  $s^n \in \mathcal{S}^n$  by

$$P_{e|s^n}^{(n)}(\mathcal{C}) \triangleq \frac{1}{2^{n(R_1+R_2)}} \sum_{m_1=1}^{2^{nR_1}} \sum_{m_2=1}^{2^{nR_2}} \sum_{y^n: g(y^n) \neq (m_1, m_2)} W_{Y^n|X_1^n, X_2^n, S^n}(y^n | f_1(m_1), f_2(m_2), s^n). \quad (6a)$$

Now, define the average probability of error of  $\mathcal{C}$  for some distribution  $q(s^n) \in \mathcal{P}(\mathcal{S}^n)$ ,

$$P_e^{(n)}(q, \mathcal{C}) \triangleq \sum_{s^n \in \mathcal{S}^n} q(s^n) \cdot P_{e|s^n}^{(n)}(\mathcal{C}). \quad (6b)$$

*Definition 4* (Achievable rate pair and capacity region under constraints). A code  $\mathcal{C} = (f_1, f_2, g)$  is called a  $(2^{nR_1}, 2^{nR_2}, n, \varepsilon)$  code for the AVMAC  $\mathcal{A}$  under input constraints  $(\Omega_1, \Omega_2)$  and state constraint  $\Lambda$ , when (3) is satisfied and

$$P_e^{(n)}(q, \mathcal{C}) \leq \varepsilon, \quad \text{for all } q \in \mathcal{P}_\Lambda(\mathcal{S}^n), \quad (7)$$

or, equivalently,  $P_{e|s^n}^{(n)}(\mathcal{C}) \leq \varepsilon$  for all  $s^n \in \mathcal{S}^n$  with  $l^n(s^n) \leq \Lambda$ .

We say that a rate pair  $(R_1, R_2)$  is achievable, under input constraints  $(\Omega_1, \Omega_2)$  and state constraint  $\Lambda$ , if for every  $\varepsilon > 0$  and sufficiently large  $n$ , there exists a  $(2^{nR_1}, 2^{nR_2}, n, \varepsilon)$  code for the AVMAC  $\mathcal{A}$  under input constraints  $(\Omega_1, \Omega_2)$  and state constraint  $\Lambda$ . The operational capacity region is defined as the closure of the set of achievable rate pairs, and it is denoted by  $\mathbb{C}(\mathcal{A})$ . We use the term ‘capacity region’ referring to this operational meaning, and in some places we call it the deterministic code capacity region in order to emphasize that achievability is measured with respect to deterministic codes.

Analogously to the deterministic case, a  $(2^{nR_1}, 2^{nR_2}, n, \varepsilon)$  random code  $\mathcal{C}^\Gamma = (\mu, \Gamma, \{\mathcal{C}_\gamma\}_{\gamma \in \Gamma})$  for the AVMAC  $\mathcal{A}$ , under input constraints  $(\Omega_1, \Omega_2)$  and state constraint  $\Lambda$ , satisfies the requirements

$$\sum_{\gamma \in \Gamma} \mu(\gamma) \phi_k^n(f_{k,\gamma}^n(m_k)) \leq \Omega_k, \quad \text{for all } m_k \in [1 : 2^{nR_k}], \quad k = 1, 2, \quad (8a)$$

and

$$P_e^{(n)}(q, \mathcal{C}^\Gamma) \triangleq \sum_{\gamma \in \Gamma} \mu(\gamma) P_e^{(n)}(q, \mathcal{C}_\gamma) \leq \varepsilon, \quad \text{for all } q \in \mathcal{P}_\Lambda(\mathcal{S}^n). \quad (8b)$$

The capacity region achieved by random codes is then denoted by  $\mathbb{C}^*(\mathcal{A})$ , and it is referred to as the *random code capacity region*. In addition, a  $(2^{nR_1}, 2^{nR_2}, n, \varepsilon)$  divided-randomness code  $\mathcal{C}^{\Gamma_1 \times \Gamma_2}$  satisfies the requirements

$$\sum_{\gamma_k} \mu_k(\gamma_k) \phi_k^n(f_{k,\gamma_k}^n(m_k)) \leq \Omega_k, \quad \text{for all } m_k \in [1 : 2^{nR_k}], \quad k = 1, 2, \quad (9a)$$

and

$$P_e^{(n)}(q, \mathcal{C}^{\Gamma_1 \times \Gamma_2}) \triangleq \sum_{\gamma_1, \gamma_2} \mu_1(\gamma_1) \mu_2(\gamma_2) P_e^{(n)}(q, \mathcal{C}_{\gamma_1, \gamma_2}) \leq \varepsilon, \quad \text{for all } q \in \mathcal{P}_\Lambda(\mathcal{S}^n). \quad (9b)$$

The capacity region achieved by divided-randomness codes is then denoted by  $\mathbb{C}^{**}(\mathcal{A})$ , and it is referred to as the *divided-randomness capacity region*.

Note that based on the definitions above,

$$\mathbb{C}(\mathcal{A}) \subseteq \mathbb{C}^{**}(\mathcal{A}) \subseteq \mathbb{C}^*(\mathcal{A}). \quad (10)$$

The definitions above are naturally extended to the compound MAC, under input constraints  $(\Omega_1, \Omega_2)$  and state constraint  $\Lambda$ , by limiting the requirements (3), (7) and (8) to i.i.d. state distributions  $q \in \mathcal{Q}$ . The respective deterministic code capacity region, random code capacity region, and divided-randomness capacity region  $\mathbb{C}(\mathcal{A}^\mathcal{Q})$ ,  $\mathbb{C}^*(\mathcal{A}^\mathcal{Q})$  and  $\mathbb{C}^{**}(\mathcal{A}^\mathcal{Q})$  are defined accordingly.

### E. Related Work

1) *Without Constraints*: In this subsection, we briefly review known results for the case where there are no constraints. Denote the deterministic code capacity region and the divided-randomness capacity regions of the AVMAC free of constraints by  $\mathbb{C}(\mathcal{A}_{\text{free}})$  and  $\mathbb{C}^{**}(\mathcal{A}_{\text{free}})$ , respectively. We note that this is a special case of the AVMAC under constraints, with  $\Omega_1 \geq \phi_{1,max}$ ,  $\Omega_2 \geq \phi_{2,max}$ , and  $\Lambda \geq l_{max}$ .

We cite the divided-randomness capacity theorem of the AVMAC free of constraints, due to Jahn [33]. Let

$$\mathbb{C}^{**}(\mathcal{A}_{\text{free}}) = \bigcup_{P_U P_{X_1|U} P_{X_2|U}} \left\{ (R_1, R_2) : \begin{array}{l} R_1 \leq \min_{q(s|u)} I_q(X_1; Y|X_2, U), \\ R_2 \leq \min_{q(s|u)} I_q(X_2; Y|X_1, U), \\ R_1 + R_2 \leq \min_{q(s|u)} I_q(X_1, X_2; Y|U) \end{array} \right\}, \quad (11)$$

with  $(U, X_1, X_2, S) \sim P_U(u) P_{X_1|U}(x_1|u) P_{X_2|U}(x_2|u) q(s|u)$ .

*Theorem 1* (see [33, Theorem 1a]). The divided-randomness capacity region of an AVMAC, free of constraints, is given by

$$\mathbb{C}^{**}(\mathcal{A}_{\text{free}}) = \mathbb{C}^*(\mathcal{A}_{\text{free}}). \quad (12)$$

*Remark 2.* Originally, the random code capacity region is expressed in [33, Theorem 1] as a closed convex hull of a union of regions. Achievability of (11) is established through time sharing.

Now, we move to the deterministic code capacity region.

*Theorem 2* (Ahlswede’s Dichotomy [33, 34]). The capacity region of an AVMAC, free of constraints, either coincides with the divided-randomness capacity region or else, it has an empty interior. That is,  $\mathbb{C}(\mathcal{A}_{\text{free}}) = \mathbb{C}^{\star\star}(\mathcal{A}_{\text{free}})$  or else,  $\text{int}(\mathbb{C}(\mathcal{A}_{\text{free}})) = \emptyset$ .

Necessary and sufficient conditions for the capacity region to have a non-empty interior were established by Gubner [24, 23] and Ahlswede and Cai [3] in terms of the following definition.

*Definition 5.* [24, 23, 22] A state-dependent MAC  $W_{Y|X_1, X_2, S}$  is said to be

1) *symmetrizable-* $\mathcal{X}_1 \times \mathcal{X}_2$  if for some conditional distribution  $J(s|x_1, x_2)$ ,

$$\sum_{s \in \mathcal{S}} W_{Y|X_1, X_2, S}(y|x_1, x_2, s) J(s|\tilde{x}_1, \tilde{x}_2) = \sum_{s \in \mathcal{S}} W_{Y|X_1, X_2, S}(y|\tilde{x}_1, \tilde{x}_2, s) J(s|x_1, x_2),$$

$$\forall x_1, \tilde{x}_1 \in \mathcal{X}_1, x_2, \tilde{x}_2 \in \mathcal{X}_2, y \in \mathcal{Y}. \quad (13)$$

Equivalently, the channel  $\widetilde{W}(y|x_1, x_2, \tilde{x}_1, \tilde{x}_2) = \sum_{s \in \mathcal{S}} W_{Y|X, S}(y|x_1, x_2, s) J(s|\tilde{x}_1, \tilde{x}_2)$  is symmetric with respect to  $(x_1, x_2)$  and  $(\tilde{x}_1, \tilde{x}_2)$ .

2) *symmetrizable-* $\mathcal{X}_1|\mathcal{X}_2$  if for some conditional distribution  $J_1(s|x_1)$ ,

$$\sum_{s \in \mathcal{S}} W_{Y|X_1, X_2, S}(y|x_1, x_2, s) J_1(s|\tilde{x}_1) = \sum_{s \in \mathcal{S}} W_{Y|X_1, X_2, S}(y|\tilde{x}_1, x_2, s) J_1(s|x_1), \quad \forall x_1, \tilde{x}_1 \in \mathcal{X}_1, x_2 \in \mathcal{X}_2, y \in \mathcal{Y}. \quad (14)$$

3) *symmetrizable-* $\mathcal{X}_2|\mathcal{X}_1$  if for some conditional distribution  $J_2(s|x_2)$ ,

$$\sum_{s \in \mathcal{S}} W_{Y|X_1, X_2, S}(y|x_1, x_2, s) J_2(s|\tilde{x}_2) = \sum_{s \in \mathcal{S}} W_{Y|X_1, X_2, S}(y|x_1, \tilde{x}_2, s) J_2(s|x_2), \quad \forall x_1 \in \mathcal{X}_1, x_2, \tilde{x}_2 \in \mathcal{X}_2, y \in \mathcal{Y}. \quad (15)$$

We say that the AVMAC  $\mathcal{A}$  is *symmetrizable-* $\mathcal{X}_1 \times \mathcal{X}_2$  if the corresponding state-dependent MAC  $W_{Y|X_1, X_2, S}$  is *symmetrizable-* $\mathcal{X}_1 \times \mathcal{X}_2$ , and similarly for *symmetrizable-* $\mathcal{X}_1|\mathcal{X}_2$  and *symmetrizable-* $\mathcal{X}_2|\mathcal{X}_1$ .

*Example 1.* [24] Consider an adder channel specified by  $Y = X_1 + X_2 + S$ , where  $\mathcal{X}_1 = \mathcal{X}_2 = \{0, 1\}$ . For  $\mathcal{S} = \{0, 1, 2\}$ , the AVMAC satisfies the conditions in Definition 5, as (13)-(15) hold with  $J(s|x_1, x_2) = \delta(s - x_1 - x_2)$ ,  $J_1(s|x_1) = \delta(s - x_1)$ ,  $J_2(s|x_2) = \delta(s - x_2)$ , where  $\delta(u)$  is the Kronecker delta function, *i.e.*  $\delta(u) = 1$  for  $u = 0$ , and  $\delta(u) = 0$  otherwise. On the other hand, it is shown in [24] that for  $\mathcal{S} = \{0, 1\}$ , the AVMAC is *symmetrizable-* $\mathcal{X}_1|\mathcal{X}_2$  and *symmetrizable-* $\mathcal{X}_2|\mathcal{X}_1$ , but *non-symmetrizable-* $\mathcal{X}_1 \times \mathcal{X}_2$ .

Ahlswede and Cai [3] showed by example that it is also possible that an AVMAC satisfies the first condition in Definition 5 but does not satisfy the other two.

*Example 2.* [3] Consider a binary MAC, with  $\mathcal{X}_1 = \mathcal{X}_2 = \mathcal{S} = \mathcal{Y} = \{0, 1\}$ , specified by the following. For  $s = 0$ ,

$$W_{Y|X_1, X_2, S}(\cdot|0, 0, 0) = W_{Y|X_1, X_2, S}(\cdot|1, 1, 0) = (1, 0),$$

$$W_{Y|X_1, X_2, S}(\cdot|1, 0, 0) = W_{Y|X_1, X_2, S}(\cdot|0, 1, 0) = \left(\frac{1}{2}, \frac{1}{2}\right), \quad (16)$$

and for  $s = 1$ ,

$$W_{Y|X_1, X_2, S}(\cdot|0, 0, 1) = W_{Y|X_1, X_2, S}(\cdot|1, 1, 1) = \left(\frac{1}{2}, \frac{1}{2}\right),$$

$$W_{Y|X_1, X_2, S}(\cdot|1, 0, 1) = W_{Y|X_1, X_2, S}(\cdot|0, 1, 1) = (0, 1). \quad (17)$$

Then it is shown in [3] that  $W_{Y|X_1, X_2, S}$  is *symmetrizable-* $\mathcal{X}_1 \times \mathcal{X}_2$ , as (13) holds for  $J(s|x_1, x_2) = 1$  for  $(x_1 = x_2, s = 0)$  or  $(x_1 \neq x_2, s = 1)$ , and  $J(s|x_1, x_2) = 0$  otherwise. On the other hand, plugging  $x_2 = 0$  in (14) yields  $J_1(s|x_1) = \delta(s - x_1)$ , while plugging  $x_2 = 1$  in (14) yields  $J_1(s|x_1) = \delta(s - (1 - x_1))$ . This means that fixing  $x_2$ , both marginals  $W_{Y|X_1, X_2, S}(\cdot|\cdot, 0, \cdot)$  and  $W_{Y|X_1, X_2, S}(\cdot|\cdot, 1, \cdot)$  of User 1 are *symmetrizable* in the single-user sense, *i.e.* as in [17, Definition 2]. However, the AVMAC is *not symmetrizable-* $\mathcal{X}_1|\mathcal{X}_2$ , because there is no  $J_1(s|x_1)$  which *symmetrizes* both marginals at the same time. In a similar manner, (15) implies a contradiction as well. Therefore, the AVMAC is *symmetrizable-* $\mathcal{X}_1 \times \mathcal{X}_2$ , but *non-symmetrizable-* $\mathcal{X}_2|\mathcal{X}_1$  and *non-symmetrizable-* $\mathcal{X}_1|\mathcal{X}_2$ .

Intuitively, *symmetrizable-* $\mathcal{X}_1 \times \mathcal{X}_2$  identifies a poor channel, where the jammer can impinge the communication scheme by randomizing the state sequence  $S^n$  according to  $J^n(s^n|\tilde{x}_1^n, \tilde{x}_2^n) = \prod_{i=1}^n J(s_i|\tilde{x}_{1,i}, \tilde{x}_{2,i})$ , for some codewords  $\tilde{x}_1^n$  and  $\tilde{x}_2^n$  in the codebooks of User 1 and User 2, respectively. Suppose that the transmitted codewords are  $x_1^n$  and  $x_2^n$ . The codewords  $\tilde{x}_1^n$  and  $\tilde{x}_2^n$  can be thought of as impostors transmitted by the jammer. Now, since the ‘‘average channel’’  $\widetilde{W}$  is symmetric with respect to  $(x_1^n, x_2^n)$  and  $(\tilde{x}_1^n, \tilde{x}_2^n)$ , the codeword pairs appear to the receiver as equally likely. Similarly, if the AVMAC is

symmetrizable- $\mathcal{X}_1|\mathcal{X}_2$ , then the decoder confuses between  $(x_1^n, x_2^n)$  and  $(\tilde{x}_1^n, x_2^n)$ , and if the AVMAC is symmetrizable- $\mathcal{X}_2|\mathcal{X}_1$ , then the decoder confuses between  $(x_1^n, x_2^n)$  and  $(x_1^n, \tilde{x}_2^n)$ , where  $\tilde{x}_1$  and  $\tilde{x}_2$  are the codewords chosen by the jammer. Indeed, by [24], if one of the conditions in Definition 5 holds, then the capacity region of the AVMAC free of constraints has an empty interior. This means that the capacity region  $\mathbb{C}(\mathcal{A}_{\text{free}})$  is either an interval or  $\{(0, 0)\}$ , *i.e.* one of the users or both have zero capacity.

Ahlsweide and Cai [3] proved that the three types of non-symmetrizability are not only a necessary condition for a non-empty capacity region, but they are sufficient conditions as well. This yields the following theorem.

*Theorem 3* (see [24, 23][3, Theorem 1]). An AVMAC free of constraints has a capacity region with a non-empty interior, *i.e.*  $\text{int}(\mathbb{C}(\mathcal{A}_{\text{free}})) \neq \emptyset$ , if and only if it is non-symmetrizable- $\mathcal{X}_1 \times \mathcal{X}_2$ , non-symmetrizable- $\mathcal{X}_1|\mathcal{X}_2$ , and non-symmetrizable- $\mathcal{X}_2|\mathcal{X}_1$ .

The following theorem combines the results by Gubner [24, 23] and Ahlsweide and Cai [3]. This statement was also given by Boche and Wiese [48, 47], who considered the AVMAC with conferencing encoders.

*Theorem 4* (see [48, Theorem 8]). There are four scenarios for the capacity region of the AVMAC free of constraints:

a) If  $W_{Y|X_1, X_2, S}$  is not symmetrizable- $\mathcal{X}_1 \times \mathcal{X}_2$ , - $\mathcal{X}_1|\mathcal{X}_2$ , nor - $\mathcal{X}_2|\mathcal{X}_1$ , then

$$\mathbb{C}(\mathcal{A}_{\text{free}}) = \mathbb{C}^{*\star}(\mathcal{A}_{\text{free}}). \quad (18)$$

b) If  $W_{Y|X_1, X_2, S}$  is not symmetrizable- $\mathcal{X}_1 \times \mathcal{X}_2$  nor - $\mathcal{X}_2|\mathcal{X}_1$ , but symmetrizable- $\mathcal{X}_1|\mathcal{X}_2$ , then

$$\mathbb{C}(\mathcal{A}_{\text{free}}) \subseteq \left\{ (0, R_2) : R_2 \leq \min_{q(s)} \max_{P_{X_1} P_{X_2}} I_q(X_2; Y|X_1) \right\}. \quad (19)$$

c) If  $W_{Y|X_1, X_2, S}$  is not symmetrizable- $\mathcal{X}_1 \times \mathcal{X}_2$  nor - $\mathcal{X}_1|\mathcal{X}_2$ , but symmetrizable- $\mathcal{X}_2|\mathcal{X}_1$ , then

$$\mathbb{C}(\mathcal{A}_{\text{free}}) \subseteq \left\{ (R_1, 0) : R_1 \leq \min_{q(s)} \max_{P_{X_1} P_{X_2}} I_q(X_1; Y|X_2) \right\}. \quad (20)$$

d) In all other cases,

$$\mathbb{C}(\mathcal{A}_{\text{free}}) = \{(0, 0)\}. \quad (21)$$

*Remark 3.* Observe that in Case b) and Case c) of Theorem 4, the characterization is incomplete. As pointed out by Wiese and Boche, this has remained an open problem for nearly 20 years [48, Remark 9] (see also [47, Remark 5.6]). At first glance, it may appear as if achievability in Cases b) and c) immediately follows from the capacity theorem of the single user AVC [17]. Consider Case c), and denote the channel from  $X_1$  to  $Y$ , for a fixed  $x_2 \in \mathcal{X}_2$ , by  $W_{Y|X_1, S}^{(x_2)} = W_{Y|X_1, X_2, S}(\cdot|\cdot, x_2, \cdot)$ . Then, based on the results by Csiszár and Narayan for the single user AVC [17], if  $W_{Y|X_1, S}^{(x_2)}$  is non-symmetrizable- $\mathcal{X}_1$  for some  $x_2 \in \mathcal{X}_2$ , then the capacity of User 1 is positive. Furthermore, if  $W_{Y|X_1, S}^{(x_2)}$  is non-symmetrizable for all  $x_2 \in \mathcal{X}_2$ , then User 1 can achieve every rate  $R_1 < \min_{q(s)} \max_{P_{X_1} P_{X_2}} I_q(X_1; Y|X_2)$ . However, in Case c), knowing that the AVMAC is non-symmetrizable- $\mathcal{X}_1|\mathcal{X}_2$  does not guarantee that  $W_{Y|X_1, S}^{(x_2)}$  is non-symmetrizable for all  $x_2 \in \mathcal{X}_2$ . Actually, it only guarantees that the channels  $W_{Y|X_1, S}^{(x_2)}$ ,  $x_2 \in \mathcal{X}_2$ , are not all symmetrized by a single  $J_1(s|x_1)$  (see Example 2). Therefore, it is not immediately clear whether the conditions in Cases b) and c) are sufficient for achievability. We are going to fill this gap and show that (19) and (20) hold with equality.

*Remark 4.* The dichotomy property in Theorem 2 was proved using Ahlsweide's Elimination Technique [1], where the encoder transmits the random elements  $\gamma_1$  and  $\gamma_2$  over a negligible portion of the blocklength. The Elimination Technique only works without state constraints [17], since positive capacity under a state constraint does not guarantee reliable transmission over a fraction of the blocklength. Moreover, as Csiszár and Narayan demonstrated in the single user setting, the dichotomy property does not hold when state constraints are imposed on the jammer. That is, the deterministic code capacity can be lower than the capacity with shared randomness, even if positive rates are achievable with deterministic codes. This demonstrates the significant effect that constraints have on the behavior of the deterministic code capacity of arbitrarily varying channels.

2) *Divided-Randomness Capacity Region:* Gubner and Hughes [27] considered the AVMAC under input and state constraints, and determined the divided-randomness capacity region, *i.e.* assuming each encoder shares an independent random element with the decoder. Their result is given below. Define

$$\mathbb{C}^{*\star}(\mathcal{A}) = \bigcup_{\substack{P_U P_{X_1|U} P_{X_2|U} : \\ \mathbb{E}\phi_k(X_k) \leq \Omega_k, k=1,2.}} \left\{ (R_1, R_2) : \begin{array}{l} R_1 \leq \min_{q(s|u) : \mathbb{E}_q I(S) \leq \Lambda} I_q(X_1; Y|X_2, U), \\ R_2 \leq \min_{q(s|u) : \mathbb{E}_q I(S) \leq \Lambda} I_q(X_2; Y|X_1, U), \\ R_1 + R_2 \leq \min_{q(s|u) : \mathbb{E}_q I(S) \leq \Lambda} I_q(X_1, X_2; Y|U) \end{array} \right\}, \quad (22)$$

with  $(U, X_1, X_2, S) \sim P_U(u)P_{X_1|U}(x_1|u)P_{X_2|U}(x_2|u)q(s|u)$ . It is shown in [27] that the region above is not necessarily convex. In Remark 7, we discuss the interpretation of this property and the connection to the statistical independence between the variables  $S$  and  $U$  above.

*Theorem 5* (see [27]). The divided-randomness capacity region of the AVMAC under input constraints  $(\Omega_1, \Omega_2)$  and state constraint  $\Lambda$  is given by

$$\mathbb{C}^{\star\star}(\mathcal{A}) = \mathbb{C}^{\star}(\mathcal{A}). \quad (23)$$

In the next sections, we determine both the random code capacity region and the deterministic code capacity region.

### III. MAIN RESULTS – RANDOM CODE CAPACITY REGION

In this section, we establish the random code capacity region of the AVMAC under input and state constraints. To this end, we first give an auxiliary result on the compound MAC.

#### A. The Compound MAC

We begin with the capacity theorem for the compound MAC  $\mathcal{A}^{\mathcal{Q}}$  under input constraints  $(\Omega_1, \Omega_2)$  and state constraint  $\Lambda$ . This is an auxiliary result, obtained by a simple extension of related work (see [38]). Let

$$\mathbb{C}(\mathcal{A}^{\mathcal{Q}}) = \bigcup_{\substack{P_U P_{X_1|U} P_{X_2|U} : \\ \mathbb{E}\phi_k(X_k) \leq \Omega_k, k=1,2.}} \left\{ (R_1, R_2) : \begin{array}{l} R_1 \leq \inf_{q \in \mathcal{Q}} I_q(X_1; Y|X_2, U), \\ R_2 \leq \inf_{q \in \mathcal{Q}} I_q(X_2; Y|X_1, U), \\ R_1 + R_2 \leq \inf_{q \in \mathcal{Q}} I_q(X_1, X_2; Y|U) \end{array} \right\}. \quad (24)$$

with  $(U, X_1, X_2, S) \sim P_U(u)P_{X_1|U}(x_1|u)P_{X_2|U}(x_2|u)q(s)$ .

*Lemma 6.* The capacity region of the compound MAC  $\mathcal{A}^{\mathcal{Q}}$  is given by

$$\mathbb{C}(\mathcal{A}^{\mathcal{Q}}) = \mathbb{C}(\mathcal{A}^{\mathcal{Q}}), \quad (25)$$

and it is identical to the divided-randomness capacity region and the random code capacity region, *i.e.*  $\mathbb{C}^{\star}(\mathcal{A}^{\mathcal{Q}}) = \mathbb{C}^{\star\star}(\mathcal{A}^{\mathcal{Q}}) = \mathbb{C}(\mathcal{A}^{\mathcal{Q}})$ .

The proof of Lemma 6 is given in Appendix A.

*Remark 5.* Regardless of the statement in Lemma 6, the capacity region of the compound MAC must be convex, due to the operational time sharing argument. That is, if  $(R_{1,u}, R_{2,u})$ ,  $u \in \mathcal{U}$ , are achievable rate pairs, then any convex combination

$$\left( \sum_{u \in \mathcal{U}} \theta_u R_{1,u}, \sum_{u \in \mathcal{U}} \theta_u R_{2,u} \right) \quad (26)$$

is achievable, for  $\theta_u \geq 0$ ,  $\sum_{u \in \mathcal{U}} \theta_u = 1$ . To achieve this rate pair, one can employ a sequence of consecutive codes that achieve  $(R_{1,u}, R_{2,u})$ , such that  $\theta_u$  is the fraction of the corresponding code length from the total blocklength (see [14, Section 15.3.3]). In the classical setting, the random variable  $U$  is referred to as the time sharing variable, since  $P_U(u)$  can be interpreted as the coefficient  $\theta_u$  in the convex combination. We explain below why this interpretation is lacking in the case of the compound MAC. Furthermore, we will see that operational time sharing is impossible for the AVMAC, yet the convexity of  $\mathbb{C}(\mathcal{A}^{\mathcal{Q}})$  will play a role (see the remarks below Theorem 7).

For the classical MAC, achievability of the capacity region can be established by first considering independent inputs  $(X_1, X_2) \sim P_{X_1}(x_1)P_{X_2}(x_2)$ , and then generalizing to  $P_U(u)P_{X_1|U}(x_1|u)P_{X_2|U}(x_2|u)$  through the operational time sharing argument. However, for the compound MAC, a straightforward application of the operational time sharing argument is insufficient, because

$$\inf_{q \in \mathcal{Q}} I_q(X_1, X_2; Y|U) \geq \sum_{u \in \mathcal{U}} p(u) \cdot \inf_{q \in \mathcal{Q}} I_q(X_1, X_2; Y|U = u), \quad (27)$$

in general, and similarly for  $I_q(X_1; Y|X_2, U)$  and  $I_q(X_2; Y|X_1, U)$ . Hence, achievability of the convex combination in the RHS of (27) does *not* immediately imply achievability of the LHS.

We deduce that the external variable  $U$  may not represent the *operational* time sharing strategy as in the classical sense. Nevertheless, we associate  $U$  with *coded* time sharing [28] [21, Section 4.5.3]. Specifically, to prove Lemma 6, we use a coded time sharing scheme, where a time sharing sequence  $U^n$  is generated, and then a single codebook is selected accordingly.

## B. The AVMAC

We determine the random code capacity region of the AVMAC under input constraints  $(\Omega_1, \Omega_2)$  and state constraint  $\Lambda$ . As opposed to Theorem 5 by [27], considering divided-randomness coding, we address the case where the three parties, two encoders and decoder, share randomness together. The random code derivation is based on our result on the compound MAC and a simple extension of Ahlswede's RT.

Define  $C^*(\mathcal{A}) \triangleq C(\mathcal{A}^{\mathcal{Q}})|_{\mathcal{Q}=\overline{\mathcal{P}}_{\Lambda}(S)}$ , i.e.

$$C^*(\mathcal{A}) = \bigcup_{\substack{P_U P_{X_1|U} P_{X_2|U} : \\ \mathbb{E}\phi_k(X_k) \leq \Omega_k, k=1,2.}} \left\{ (R_1, R_2) : \begin{array}{l} R_1 \leq \min_{q(s) : \mathbb{E}_q l(S) \leq \Lambda} I_q(X_1; Y|X_2, U), \\ R_2 \leq \min_{q(s) : \mathbb{E}_q l(S) \leq \Lambda} I_q(X_2; Y|X_1, U), \\ R_1 + R_2 \leq \min_{q(s) : \mathbb{E}_q l(S) \leq \Lambda} I_q(X_1, X_2; Y|U) \end{array} \right\}. \quad (28)$$

with  $(U, X_1, X_2, S) \sim P_U(u)P_{X_1|U}(x_1|u)P_{X_2|U}(x_2|u)q(s)$ . Notice the resemblance between the random code capacity region formula (28) and the divided-randomness capacity region formula (22), as the only difference between the formulas is that the state  $S$  and the "time-sharing" variable  $U$  are statistically independent in the random code case, while  $S$  and  $U$  are dependent in the divided-randomness case. An intuitive interpretation is given in the remark.

*Theorem 7.* The random code capacity region of the AVMAC  $\mathcal{A}$  under input constraints  $(\Omega_1, \Omega_2)$  and state constraint  $\Lambda$  is given by

$$C^*(\mathcal{A}) = C^*(\mathcal{A}). \quad (29)$$

The proof of Theorem 7 is given in Appendix B. The proof is based on the aforementioned result on the compound MAC and an extension of Ahlswede's Robustification Technique [2]. Essentially, we use a reliable code for the compound MAC to construct a random code for the AVMAC by applying random permutations to each codeword symbols.

*Remark 6.* As opposed to the compound channel, the operational time sharing argument is not eligible for the AVC under a state constraint, even in the single user case, as we explain below. Suppose that two codebooks are used in time sharing, one of rate  $R'$  and length  $\theta n$ , and one of rate  $R''$  and length  $(1 - \theta)n$ , for  $0 < \theta < 1$ , where  $R'$  and  $R''$  are both achievable for the AVC under a state constraint  $\Lambda$ . Due to the state constraint, it is guaranteed that  $\sum_{i=1}^n l(S_i) \leq n\Lambda$  with probability 1. However, the jammer is entitled to concentrate the jamming power on the first  $\theta n$  symbols, in which case,  $\sum_{i=1}^{\theta n} l(S_i) = n\Lambda$ . If the jammer does so, then the first code, of rate  $R'$  and length  $\theta n$ , needs to be robust against a state sequence  $S^{\theta n}$  with the following cost,

$$\frac{1}{\theta n} \sum_{i=1}^{\theta n} l(S_i) = \frac{\Lambda}{\theta} > \Lambda \quad \text{a.s.} \quad (30)$$

Therefore, in order to achieve the rate  $R = \theta R' + (1 - \theta)R''$  with operational time sharing, the first coding rate  $R'$  needs to be achievable for an AVC under a state constraint  $\Lambda/\theta$ , and the second coding rate  $R''$  needs to be achievable for a state constraint  $\Lambda/(1 - \theta)$ , which is not guaranteed. Hence, operational time sharing is not eligible for neither the single user AVC, nor the AVMAC (see also [27]).

Nevertheless, we observe that using the proof technique in Appendix B, one can devise a reliable coding scheme with "shuffled time sharing". That is, instead of using the codes of rates  $R'$  and  $R''$  above consecutively, random interleaving of the codes can be realized. Intuitively, the users are thus able to carry out a time sharing protocol of which the jammer is oblivious, which explains the statistical independence between the state  $S$  and the time sharing variable  $U$  in (28).

*Remark 7.* As mentioned above, the difference between the formulas given for the random code capacity region and for the divided-randomness capacity region is the statistical independence between the state and the time sharing variable. Now, we discuss the implications in terms of the convexity of the regions.

By Theorem 7 and Lemma 6, we have that the random code capacity region of the AVMAC under input and state constraints is the same as that of the compound MAC with  $\mathcal{Q} = \overline{\mathcal{P}}_{\Lambda}(S)$ . As a consequence, we have that the random code capacity region of the AVMAC is convex. It can also be verified directly that the set in the RHS of (28) is convex, as  $(U, S) \sim P_U(u)q(s)$ .

On the other hand, Gubner and Hughes [27] demonstrated that the divided-randomness capacity region given by (22) is *not* necessarily convex, as  $(U, S) \sim P_U(u)q(s|u)$  (see Section IV in [27]). In their setting, the encoders have statistically independent random elements  $\gamma_1$  and  $\gamma_2$  (see Definition 3). Gubner and Hughes attribute the non-convexity to the preclusion of operational time sharing [27]. It is further mentioned in [27] that there are other instances of non-convex capacity regions in the literature, such as the asynchronous MAC [32, 42], where the users' timeframes do not synchronize hence time sharing does not work either. We observe that the "shuffled time sharing" mentioned in the previous remark could only work if the shared randomness element is exploited for the coordination between the users. Whereas, in a scenario where the random elements are independent, as in [27], such coordination is impossible.

In our setting, the random elements are not independent, as  $\gamma_1 = \gamma_2 = \gamma$  (cf. Definition 2 and Definition 3). In the remark that follows Definition 2 in [27], it is stated without proof that removing the restriction of independence could result in a



strictly larger capacity region. Indeed, our result above that  $\mathbb{C}^*(\mathcal{A})$  is convex implies that for the erasure AVMAC in [27, Section IV], our random code capacity region  $\mathbb{C}^*(\mathcal{A})$  must be strictly larger than the non-convex divided-randomness capacity region. Therefore, we have now validated the assertion by Gubner and Hughes [27]. Comparing (22) and (28), we infer that the conditioning of the state distribution on  $U$  may lead to a strictly smaller region. Intuitively, knowing the time sharing protocol helps the jammer reduce the coding rates for such a channel.

#### IV. MAIN RESULTS – DETERMINISTIC CODE CAPACITY REGION

The principal result of this paper is the deterministic code capacity theorem, *i.e.* without shared randomness. The deterministic code derivation is independent of our previous results, and the analysis modifies the techniques of Csiszár and Narayan [17], and merges their ideas with those of Ahlswede and Cai [3].

Before we state the capacity theorem, we give the following definitions. Given an input distribution  $P_{X_1, X_2} \in \mathcal{P}(\mathcal{X}_1 \times \mathcal{X}_2)$ , consider the average state costs below,

$$\Psi(P_{X_1, X_2}) = \min_{\text{symm. } J} \sum_{k=1,2} \sum_{x_k \in \mathcal{X}_k} \sum_{s \in \mathcal{S}} P_{X_1, X_2}(x_1, x_2) J(s|x_1, x_2) l(s), \quad (31a)$$

$$\Psi_1(P_{X_1}) = \min_{\text{symm. } J_1} \sum_{x_1 \in \mathcal{X}_1} \sum_{s \in \mathcal{S}} P_{X_1}(x_1) J_1(s|x_1) l(s), \quad (31b)$$

$$\Psi_2(P_{X_2}) = \min_{\text{symm. } J_2} \sum_{x_2 \in \mathcal{X}_2} \sum_{s \in \mathcal{S}} P_{X_2}(x_2) J_2(s|x_2) l(s), \quad (31c)$$

where the minimizations are over  $J(s|x_1, x_2)$ ,  $J_1(s|x_1)$ , and  $J_2(s|x_2)$ , which satisfy the symmetrizing conditions in (13), (14), and (15), respectively. We use the convention that a minimum over an empty set is  $+\infty$ . Then, for every  $P_{U, X_1, X_2} \in \mathcal{P}(\mathcal{U} \times \mathcal{X}_1 \times \mathcal{X}_2)$ , define

$$\tilde{\Lambda}(P_{U, X_1, X_2}) = \sum_{u \in \mathcal{U}} P_U(u) \Psi(P_{X_1, X_2|U=u}), \quad (32a)$$

$$\tilde{\Lambda}_1(P_{U, X_1}) = \sum_{u \in \mathcal{U}} P_U(u) \Psi_1(P_{X_1|U=u}), \quad (32b)$$

$$\tilde{\Lambda}_2(P_{U, X_2}) = \sum_{u \in \mathcal{U}} P_U(u) \Psi_2(P_{X_2|U=u}). \quad (32c)$$

Intuitively,  $\min\{\tilde{\Lambda}(P_{U, X_1, X_2}), \tilde{\Lambda}_1(P_{U, X_1}), \tilde{\Lambda}_2(P_{U, X_2})\}$  is the minimal average state cost which the jammer has to pay to symmetrize the channel, for a given inputs distribution  $P_{X_1, X_2|U}$ , where symmetrizing refers to using a conditional distribution that satisfies either one of the symmetrizability conditions in Definition 5. If this minimal state cost violates the state constraint  $\Lambda$ , then the jammer is prohibited from symmetrizing the channel.

*Remark 8.* The minimal average state cost can be expressed more explicitly as

$$\tilde{\Lambda}(P_{U, X_1, X_2}) = \min_{\text{symm. } \{J_u\}} \sum_{u, x_1, x_2, s} P_U(u) P_{X_1, X_2|U}(x_1, x_2|u) J_u(s|x_1, x_2) l(s), \quad (33)$$

with minimization over a set of distribution  $\{J_u\}_{u \in \mathcal{U}}$ , where each distribution  $J_u(s|x_1, x_2)$  symmetrizes  $\mathcal{X}_1 \times \mathcal{X}_2$  the AVMAC. Notice that the state distributions are indexed by the time sharing variable. This has the interpretation of a jamming scheme that varies over time in accordance with the time sharing sequence chosen by the users.

We have defined  $\tilde{\Lambda}(P_{U, X_1, X_2})$ ,  $\tilde{\Lambda}_1(P_{U, X_1})$  and  $\tilde{\Lambda}_2(P_{U, X_2})$  in (32) as the minimal average state costs which the jammer has to pay to symmetrize the channel, for a given input distribution  $P_{U, X_1, X_2}$ . Intuitively, the users are interested in restricting the jammer by increasing those costs as much as possible, hence the following quantities represent the best thresholds the users can obtain,

$$L^* \triangleq \max_{P_{U, X_1, X_2} : \mathbb{E}\phi_k(X_k) \leq \Omega_k, k=1,2} \tilde{\Lambda}(P_{U, X_1, X_2}), \quad (34)$$

$$L_1^* \triangleq \max_{P_{U, X_1} : \mathbb{E}\phi_1(X_1) \leq \Omega_1} \tilde{\Lambda}_1(P_{U, X_1}), \quad (35)$$

$$L_2^* \triangleq \max_{P_{U, X_2} : \mathbb{E}\phi_2(X_2) \leq \Omega_2} \tilde{\Lambda}_2(P_{U, X_2}), \quad (36)$$

We note that  $L^*$ ,  $L_1^*$  and  $L_2^*$  depend on the input constraints  $(\Omega_1, \Omega_2)$ , the MAC  $W_{Y|X_1, X_2, S}$ , and the state cost function  $l : \mathcal{S} \rightarrow [0, \infty)$ , but they do not depend on the state constraint  $\Lambda$ .

*Definition 6.* Define the rate region  $\mathbb{C}(\mathcal{A})$  as follows.

a) If  $L^* > \Lambda$ ,  $L_1^* > \Lambda$ , and  $L_2^* > \Lambda$ , then

$$\mathbb{C}(\mathcal{A}) = \bigcup_{\overline{\mathcal{P}}_{\Omega_1, \Omega_2, \Lambda}(\mathcal{U} \times \mathcal{X}_1 \times \mathcal{X}_2)} \left\{ (R_1, R_2) : \begin{array}{l} R_1 \leq \min_{q(s|u) : \mathbb{E}_q l(S) \leq \Lambda} I_q(X_1; Y|X_2, U), \\ R_2 \leq \min_{q(s|u) : \mathbb{E}_q l(S) \leq \Lambda} I_q(X_2; Y|X_1, U), \\ R_1 + R_2 \leq \min_{q(s|u) : \mathbb{E}_q l(S) \leq \Lambda} I_q(X_1, X_2; Y|U) \end{array} \right\}, \quad (37a)$$

where

$$\overline{\mathcal{P}}_{\Omega_1, \Omega_2, \Lambda}(\mathcal{U} \times \mathcal{X}_1 \times \mathcal{X}_2) = \{P_U P_{X_1|U} P_{X_2|U} : \mathbb{E} \phi_k(X_k) \leq \Omega_k, k = 1, 2, \text{ and } \min\{\tilde{\Lambda}(P_{U, X_1, X_2}), \tilde{\Lambda}_1(P_{U, X_1}), \tilde{\Lambda}_2(P_{U, X_2})\} \geq \Lambda\}. \quad (37b)$$

b) If  $L^* > \Lambda$ ,  $L_2^* > \Lambda$ , but  $L_1^* \leq \Lambda$ , then

$$\mathbb{C}(\mathcal{A}) = \{(0, R_2) : R_2 \leq \min_{q(s) : \mathbb{E}_q l(S) \leq \Lambda} \max_{P_{X_1} P_{X_2} : \mathbb{E} \phi_k(X_k) \leq \Omega_k, k=1,2, \min\{\tilde{\Lambda}(P_{X_1} P_{X_2}), \tilde{\Lambda}_2(P_{X_2})\} \geq \Lambda} I_q(X_2; Y|X_1)\}. \quad (37c)$$

c) If  $L^* > \Lambda$ ,  $L_1^* > \Lambda$ , but  $L_2^* \leq \Lambda$ , then

$$\mathbb{C}(\mathcal{A}) = \{(R_1, 0) : R_1 \leq \min_{q(s) : \mathbb{E}_q l(S) \leq \Lambda} \max_{P_{X_1} P_{X_2} : \mathbb{E} \phi_k(X_k) \leq \Omega_k, k=1,2, \min\{\tilde{\Lambda}(P_{X_1} P_{X_2}), \tilde{\Lambda}_1(P_{X_1})\} \geq \Lambda} I_q(X_1; Y|X_2)\}. \quad (37d)$$

d) Otherwise, if  $L^* \leq \Lambda$ , or if both  $L_1^* \leq \Lambda$  and  $L_2^* \leq \Lambda$ , then

$$\mathbb{C}(\mathcal{A}) = \{(0, 0)\}. \quad (37e)$$

*Remark 9.* Observe that the optimization set of the input distribution  $P_U P_{X_1|U} P_{X_2|U}$  in (37) is a subset of the corresponding set in (22), for the divided-randomness capacity region, hence  $\mathbb{C}(\mathcal{A}) \subseteq \mathbb{C}^{**}(\mathcal{A})$ . Furthermore, if the AVMAC is non-symmetrizable in the sense of neither  $\mathcal{X}_1 \times \mathcal{X}_2$ ,  $\mathcal{X}_1|\mathcal{X}_2$ , nor  $\mathcal{X}_2|\mathcal{X}_1$ , then  $\tilde{\Lambda}(P_{U, X_1, X_2}) = \tilde{\Lambda}_1(P_{U, X_1}) = \tilde{\Lambda}_2(P_{U, X_2}) = +\infty$ , in which case we have that  $\mathbb{C}(\mathcal{A}) = \mathbb{C}^{**}(\mathcal{A})$ .

*Theorem 8.* Assume that  $L^* \neq \Lambda$ ,  $L_1^* \neq \Lambda$  and  $L_2^* \neq \Lambda$ . Then, the capacity region of the AVMAC  $\mathcal{A}$  under input constraints  $(\Omega_1, \Omega_2)$  and state constraint  $\Lambda$  is given by

$$\mathbb{C}(\mathcal{A}) = \mathbb{C}(\mathcal{A}). \quad (38)$$

Furthermore, if  $\mathcal{A}$  is non-symmetrizable- $\mathcal{X}_1 \times \mathcal{X}_2$ , non-symmetrizable- $\mathcal{X}_1|\mathcal{X}_2$ , and non-symmetrizable- $\mathcal{X}_2|\mathcal{X}_1$ , then the capacity region coincides with the divided-randomness capacity region, i.e.  $\mathbb{C}(\mathcal{A}) = \mathbb{C}^{**}(\mathcal{A}) = \mathbb{C}^{**}(\mathcal{A})$ .

The proof of Theorem 8 is given in Appendix E. The second part of the theorem follows from Remark 9 and (38). The proof does not use our results on the random code capacity region of the AVMAC and on the compound MAC, and it is independent of the divided-randomness analysis by Gubner and Hughes [27]. The analysis, however, makes use of the properties established for the decoding rule and codebooks specified below, in Subsections IV-A and IV-B. As mentioned, coded time sharing is an essential replacement for the classical operational time sharing argument, which cannot be applied to the AVMAC under constraints (see Remark 6). Hence, our analysis combines our coded time sharing variant of the decoder by Ahlswede and Cai [3], with our generalization of the codebook generated by Csiszár and Narayan [17]. The converse proof further uses Gubner's observations in [24].

*Remark 10.* As explained in Remark 3 for the AVMAC without constraints, the case where one of the users has zero capacity does not immediately follow from the results on the single user AVC. The reason behind this is that Gubner's second and third conditions are stronger than single-user symmetrizability, as defined in [17, Definition 2]. In particular, if the AVMAC is symmetrized- $\mathcal{X}_1|\mathcal{X}_2$  by  $J_1(s|x_1)$ , then the marginal AVC  $W_{Y|X_1, S}$  is also symmetrized by  $J_1(s|x_1)$ , but the other direction is not true. In the constrained setting, this means that the minimal state cost  $\tilde{\Lambda}_1(P_{X_1})$  for symmetrizability- $\mathcal{X}_1|\mathcal{X}_2$  can be higher than the minimal state cost  $\tilde{\Lambda}_0(P_{X_1})$  in [17, Equation (2.13)], for symmetrizability of the marginal AVC  $W_{Y|X_1, S}$ . In Example 2, we have seen that the AVMAC is non-symmetrizable- $\mathcal{X}_1|\mathcal{X}_2$ , even though the marginal  $W_{Y|X_1, S}$  could be symmetrizable, in which case,  $\tilde{\Lambda}_1(P_{X_1}) = +\infty$  but  $\tilde{\Lambda}_0(P_{X_1}) \leq l_{max} < \infty$ . Therefore, even if  $L_1^* > \Lambda$ , it is not guaranteed that User 1 can achieve a positive rate.

*Remark 11.* The boundary case where either  $L^* = \Lambda$  or  $L_k^* = \Lambda$ ,  $k = 1, 2$ , remains unsolved. Even in the single user setting, say  $\mathcal{X}_2 = \emptyset$ , the case of  $L_1^* = \Lambda$  is an open problem (see [17]), although it is conjectured in [17] that the capacity is zero in this case. Similarly, we conjecture that the capacity region is  $\mathbb{C}(\mathcal{A}) = \mathbb{C}(\mathcal{A})$  for all values of  $L^*$ ,  $L_1^*$  and  $L_2^*$ . There are special cases where we can prove that this holds, given in the corollary below. The corollary generalizes the remark following Theorem 3 in [17].

*Corollary 9.* Let  $\mathcal{A}$  be an AVMAC under input constraints  $(\Omega_1, \Omega_2)$  and state constraint  $\Lambda$ , where  $\mathcal{A}$  is symmetrizable- $\mathcal{X}_1 \times \mathcal{X}_2$ , symmetrizable- $\mathcal{X}_1 | \mathcal{X}_2$ , and symmetrizable- $\mathcal{X}_2 | \mathcal{X}_1$ . If the symmetrizability equations (13), (14), and (15) are only satisfied by conditional distributions  $J(s|x_1, x_2)$ ,  $J_1(s|x_1)$   $J_2(s|x_2)$  which are 0-1 laws, then

$$\mathbb{C}(\mathcal{A}) = \mathbb{C}(\mathcal{A}). \quad (39)$$

The proof of Corollary 9 is given in Appendix F. In particular, we note that the condition of 0-1 laws in Corollary 9 holds when the output  $Y$  is a deterministic function of  $X_1$ ,  $X_2$  and  $S$ .

#### A. Decoding Rule

We specify the decoding rule and state the corresponding properties, which are used in the analysis. As mentioned in Remark 6 above, we cannot use the operational time sharing argument for the AVMAC under constraints, and therefore, we use coded time sharing [28] [21, Section 4.5.3]. Our decoder is similar to that of Ahlswede and Cai [3], and the codebooks are generated based on the techniques of Csiszár and Narayan [17], along with the insights of Ahlswede and Cai [3]. We note that since the code is deterministic, the time sharing sequence is also deterministic, and it is known to the encoders, the decoder, and the jammer as well.

A fundamental difference between our coding scheme and the one in [3] arises from Ahlswede's dichotomy property. Specifically, Ahlswede and Cai only showed achievability of positive rates  $R_1 = R_2 = \varepsilon > 0$ , proving that the capacity region has a non-empty interior. According to the dichotomy result in Theorem 2 by Jahn [33, 34], this implies that the capacity region of the AVMAC free of constraints is the same as the random code capacity region, which was also determined in [33, 34]. However, for the AVMAC under constraints, dichotomy does not apply and achievability of positive rates is insufficient. Thereby, our proof is a lot more involved than the one in [3].

To specify the decoding rule, we define the decoding sets  $\mathcal{D}(m_1, m_2) \subseteq \mathcal{Y}^n$ , for  $(m_1, m_2) \in [1 : 2^{nR_1}] \times [1 : 2^{nR_2}]$ , such that  $g(y^n) = (m_1, m_2)$  iff  $y^n \in \mathcal{D}(m_1, m_2)$ .

*Definition 7 (Decoder).* Given the codebooks  $\{f_k(m_k)\}_{m_k \in [1:2^{nR_k}]}$ ,  $k = 1, 2$ , and a time sharing sequence  $u^n$ , declare that  $y^n \in \mathcal{D}(m_1, m_2)$  if there exists  $s^n \in \mathcal{S}^n$  with  $l^n(s^n) \leq \Lambda$  such that the following hold.

1) For  $(U, X_1, X_2, S, Y)$  which is distributed according to the joint type  $\hat{P}_{u^n, f_1(m_1), f_2(m_2), s^n, y^n}$ , we have that

$$D(P_{U, X_1, X_2, S, Y} || P_U \times P_{X_1|U} \times P_{X_2|U} \times P_{S|U} \times W_{Y|X_1, X_2, S}) \leq \eta. \quad (40)$$

2) a) For every  $\tilde{m}_1 \neq m_1$  and  $\tilde{m}_2 \neq m_2$  such that for some  $\tilde{s}^n \in \mathcal{S}^n$  with  $l^n(\tilde{s}^n) \leq \Lambda$ ,

$$D(P_{U, \tilde{X}_1, \tilde{X}_2, \tilde{S}, Y} || P_U \times P_{\tilde{X}_1|U} \times P_{\tilde{X}_2|U} \times P_{\tilde{S}|U} \times W_{Y|X_1, X_2, S}) \leq \eta, \quad (41)$$

where  $(U, \tilde{X}_1, \tilde{X}_2, \tilde{S}, Y) \sim \hat{P}_{u^n, f_1(\tilde{m}_1), f_2(\tilde{m}_2), \tilde{s}^n, y^n}$ , we have that

$$I(X_1, X_2, Y; \tilde{X}_1, \tilde{X}_2 | U, S) \leq \eta. \quad (42)$$

b) For every  $\tilde{m}_1 \neq m_1$  such that for some  $\tilde{s}^n \in \mathcal{S}^n$  with  $l^n(\tilde{s}^n) \leq \Lambda$ ,

$$D(P_{U, \tilde{X}_1, X_2, \tilde{S}, Y} || P_U \times P_{\tilde{X}_1|U} \times P_{X_2|U} \times P_{\tilde{S}|U} \times W_{Y|X_1, X_2, S}) \leq \eta, \quad (43)$$

where  $(U, \tilde{X}_1, X_2, \tilde{S}, Y) \sim \hat{P}_{u^n, f_1(\tilde{m}_1), f_2(m_2), \tilde{s}^n, y^n}$ , we have that

$$I(X_1, X_2, Y; \tilde{X}_1 | U, S) \leq \eta_1. \quad (44)$$

c) For every  $\tilde{m}_2 \neq m_2$  such that for some  $\tilde{s}^n \in \mathcal{S}^n$  with  $l^n(\tilde{s}^n) \leq \Lambda$ ,

$$D(P_{U, X_1, \tilde{X}_2, \tilde{S}, Y} || P_U \times P_{X_1|U} \times P_{\tilde{X}_2|U} \times P_{\tilde{S}|U} \times W_{Y|X_1, X_2, S}) \leq \eta, \quad (45)$$

where  $(U, X_1, \tilde{X}_2, \tilde{S}, Y) \sim \hat{P}_{u^n, f_1(m_1), f_2(\tilde{m}_2), \tilde{s}^n, y^n}$ , we have that

$$I(X_1, X_2, Y; \tilde{X}_2 | U, S) \leq \eta_2. \quad (46)$$

We note that in Definition 7, the variables  $U, X_1, X_2, \tilde{X}_1, \tilde{X}_2, S, \tilde{S}, Y$  are dummy random variables, distributed according to the joint type of  $(u^n, f_1(m_1), f_2(m_2), f_1(\tilde{m}_1), f_2(\tilde{m}_2), s^n, \tilde{s}^n, y^n)$ , where  $u^n$  is a given time sharing sequence,  $f_1(m_1), f_2(m_2)$  are "tested" codewords,  $f_1(\tilde{m}_1), f_2(\tilde{m}_2)$  are competing codewords,  $s^n$  is a "tested" state sequence,  $\tilde{s}^n$  is a competing state sequence, and  $y^n$  is the received sequence. None of the sequences are random here. The Markov relation  $U \ominus (X_1, X_2, S) \ominus Y$  may not hold for those dummy variables, and we may have that the conditional type  $P_{Y|X_1, X_2, S}$  differs from the actual channel  $W_{Y|X_1, X_2, S}$ . Therefore, the divergences and mutual informations in Definition 7 could be positive.

For the definition above to be proper, we need to verify that the decoding sets are disjoint, as stated in the following lemma.

*Lemma 10* (Decoding Disambiguity). Let  $u^n$  be a given time sharing sequence, and denote by  $P_U$  its type, that is,  $P_U = \hat{P}_{u^n}$ . Suppose that in each codebook, all codewords have the same conditional type, i.e.  $\hat{P}_{f_1(m_1)|u^n} = P_{X_1|U}$  and  $\hat{P}_{f_2(m_2)|u^n} = P_{X_2|U}$  for all  $(m_1, m_2)$ . Assume that for some  $\delta, \delta_k > 0$ ,  $P_U(u) \geq \delta$ ,  $P_{X_k|U}(x_k|u) \geq \delta_k \forall x_k \in \mathcal{X}_k$ ,  $u \in \mathcal{U}$ ,  $k = 1, 2$ , and also

$$\min \left\{ \tilde{\Lambda}(P_{U, X_1, X_2}), \tilde{\Lambda}_1(P_{U, X_1}), \tilde{\Lambda}_2(P_{U, X_2}) \right\} > \Lambda. \quad (47)$$

Then, for sufficiently small  $\eta, \eta_1, \eta_2 > 0$ ,

$$\mathcal{D}(m_1, m_2) \neq \mathcal{D}(\tilde{m}_1, \tilde{m}_2), \text{ for all } (m_1, m_2) \neq (\tilde{m}_1, \tilde{m}_2). \quad (48)$$

Specifically,

1) Conditions 1) and 2a) of the decoding rule, with  $\tilde{\Lambda}(P_{U, X_1, X_2}) > \Lambda$ , imply that for sufficiently small  $\eta$ ,

$$\mathcal{D}(m_1, m_2) \cap \mathcal{D}(\tilde{m}_1, \tilde{m}_2) = \emptyset, \text{ for } m_1 \neq \tilde{m}_1 \text{ and } m_2 \neq \tilde{m}_2. \quad (49)$$

2) Conditions 1) and 2b) of the decoding rule, with  $\tilde{\Lambda}_1(P_{U, X_1}) > \Lambda$ , imply that for sufficiently small  $\eta$  and  $\eta_1$ ,

$$\mathcal{D}(m_1, m_2) \cap \mathcal{D}(\tilde{m}_1, m_2) = \emptyset, \text{ for } m_1 \neq \tilde{m}_1. \quad (50)$$

3) Conditions 1) and 2c) of the decoding rule, with  $\tilde{\Lambda}_2(P_{U, X_2}) > \Lambda$ , imply that for sufficiently small  $\eta$  and  $\eta_2$ ,

$$\mathcal{D}(m_1, m_2) \cap \mathcal{D}(m_1, \tilde{m}_2) = \emptyset, \text{ for } m_2 \neq \tilde{m}_2. \quad (51)$$

The proof of Lemma 10 is given in Appendix C.

## B. Codebooks

While the decoding rule above is similar to that of Ahlswede and Cai [3], here we prove a generalization of a lemma by Csiszár and Narayan [17], in order to generate proper codebooks.

*Lemma 11* (Codebooks Generation). For every  $\varepsilon > 0$ , sufficiently large  $n$ , rates  $R_k \geq \varepsilon$  and types  $P_U$  and  $P_k = P_{X_k|U}$ ,  $k = 1, 2$ , there exist a time sharing sequence  $u^n \in \mathcal{T}^n(P_U)$ , and codebooks,  $\{(x_1^n(m_1), x_2^n(m_2)) : m_k \in [1 : 2^{nR_k}], k = 1, 2\}$  of type  $P_1 \times P_2$ , such that for every  $a_1^n \in \mathcal{X}_1^n$ ,  $a_2^n \in \mathcal{X}_2^n$ ,  $s^n \in \mathcal{S}^n$  with  $l^n(s^n) \leq \Lambda$ , and every joint type  $P_{U, X_1, X_2, \tilde{X}_1, \tilde{X}_2, S}$  with  $P_{X_1, X_2|U} = P_{\tilde{X}_1, \tilde{X}_2|U} = P_1 \times P_2$ , the following hold.

1) *Joint Typicality*

$$|\{(\tilde{m}_1, \tilde{m}_2) : (u^n, a_1^n, a_2^n, x_1^n(\tilde{m}_1), x_2^n(\tilde{m}_2), s^n) \in \mathcal{T}^n(P_{U, X_1, X_2, \tilde{X}_1, \tilde{X}_2, S})\}| \leq 2^{n([R_1 + R_2 - I(\tilde{X}_1, \tilde{X}_2; X_1, X_2, S|U)]_+ + \varepsilon)}, \quad (52)$$

$$|\{(m_1, m_2) : (u^n, x_1^n(m_1), x_2^n(m_2), s^n) \in \mathcal{T}^n(P_{U, X_1, X_2, S})\}| \leq 2^{n(R_1 + R_2 - \frac{\varepsilon}{2})}, \text{ if } I(X_1, X_2; S|U) > \varepsilon, \quad (53)$$

and

$$\begin{aligned} & |\{(m_1, m_2) : (u^n, x_1^n(m_1), x_2^n(m_2), x_1^n(\tilde{m}_1), x_2^n(\tilde{m}_2), s^n) \in \mathcal{T}^n(P_{U, X_1, X_2, \tilde{X}_1, \tilde{X}_2, S}), \text{ for some } \tilde{m}_1 \neq m_1, \tilde{m}_2 \neq m_2\}| \\ & \leq 2^{n(R_1 + R_2 - \frac{\varepsilon}{2})}, \text{ if } I(X_1, X_2; \tilde{X}_1, \tilde{X}_2, S|U) - \left[ R_1 + R_2 - I(\tilde{X}_1, \tilde{X}_2; S|U) \right]_+ > \varepsilon. \end{aligned} \quad (54)$$

2) *Conditional Typicality Given  $m_2$*

$$|\{\tilde{m}_1 : (u^n, a_1^n, a_2^n, x_1^n(\tilde{m}_1), s^n) \in \mathcal{T}^n(P_{U, X_1, X_2, \tilde{X}_1, S})\}| \leq 2^{n([R_1 - I(\tilde{X}_1; X_1, X_2, S|U)]_+ + \varepsilon)}, \quad (55)$$

and

$$\begin{aligned} & |\{m_1 : (u^n, x_1^n(m_1), a_2^n, x_1^n(\tilde{m}_1), s^n) \in \mathcal{T}^n(P_{U, X_1, X_2, \tilde{X}_1, S}) \text{ for some } \tilde{m}_1 \neq m_1\}| \leq 2^{n(R_1 - \frac{\varepsilon}{2})}, \\ & \text{if } I(X_1, X_2; \tilde{X}_1, S|U) - \left[ R_1 - I(\tilde{X}_1; S|U) \right]_+ > \varepsilon. \end{aligned} \quad (56)$$

3) *Conditional Typicality Given  $m_1$*

$$|\{\tilde{m}_2 : (u^n, a_1^n, a_2^n, x_2^n(\tilde{m}_2), s^n) \in \mathcal{T}^n(P_{U, X_1, X_2, \tilde{X}_2, S})\}| \leq 2^{n([R_2 - I(\tilde{X}_2; X_1, X_2, S|U)]_+ + \varepsilon)}, \quad (57)$$

and

$$\begin{aligned} & |\{m_2 : (u^n, a_1^n, x_2^n(m_2), x_2^n(\tilde{m}_2), s^n) \in \mathcal{T}^n(P_{U, X_1, X_2, \tilde{X}_2, S}) \text{ for some } \tilde{m}_2 \neq m_2\}| \leq 2^{n(R_2 - \frac{\varepsilon}{2})} \\ & \text{, if } I(X_1, X_2; \tilde{X}_2, S|U) - \left[ R_2 - I(\tilde{X}_2; S|U) \right]_+ > \varepsilon. \end{aligned} \quad (58)$$

The proof of Lemma 11 is given in Appendix D.

### C. Examples

To illustrate our results, we give the following examples.

*Example 3.* (see [27]) In the first example, we use Corollary 9 and previous results by Gubner and Hughes [27] to show that the deterministic code capacity region can be non convex. Consider the state dependent erasure MAC, specified

$$Y = \begin{cases} X_1 + X_2 & X_1 \cdot X_2 = S = 0, \\ r & \text{otherwise} \end{cases}, \quad (59)$$

where  $S = \{0, 1\}$ ,  $\mathcal{X}_1 = \mathcal{X}_2 = \{0, 1, \dots, r-1\}$ , and  $\mathcal{Y} = \{0, 1, \dots, r\}$ , with  $r \geq 2$ . Consider the AVMAC under a state constraint  $\frac{1}{n} \sum_{i=1}^n s_i \leq \Lambda$ , for  $0 < \Lambda \leq 1$ , and with inactive input constraints, *i.e.*  $\Omega_k = \phi_{k, \max}$  for  $k = 1, 2$ .

It can be readily verified that the symmetrizability conditions in Definition 5 hold when the distributions  $J(s|x_1, x_2)$  and  $J_k(s|x_k)$  assign unit probability to  $S = 1$ . Then,  $\tilde{\Lambda}(P_{U, X_1, X_2}) = \tilde{\Lambda}_k(P_{U, X_k}) = 1$  for all  $P_{U, X_1, X_2}$ , hence  $L^* = L_k^* = 1$  for  $k = 1, 2$ . By Corollary 9, we have that without a state constraint, *i.e.* for  $\Lambda = 1$ , the deterministic code capacity region is

$$\mathbb{C}(\mathcal{A}) = \{(0, 0)\}. \quad (60)$$

Whereas, if  $\Lambda < 1$ , then  $\Lambda$  is strictly less than  $L^* = L_k^* = \tilde{\Lambda}(P_{U, X_1, X_2}) = \tilde{\Lambda}_k(P_{U, X_k}) = 1$ , for all  $P_{U, X_1, X_2}$ ,  $k = 1, 2$ . Hence, by Corollary 9,

$$\mathbb{C}(\mathcal{A}) = \bigcup_{P_U P_{X_1|U} P_{X_2|U}} \left\{ (R_1, R_2) : \begin{array}{l} R_1 \leq \min_{q(s|u): \mathbb{E}S \leq \Lambda} I_q(X_1; Y|X_2, U), \\ R_2 \leq \min_{q(s|u): \mathbb{E}S \leq \Lambda} I_q(X_2; Y|X_1, U), \\ R_1 + R_2 \leq \min_{q(s|u): \mathbb{E}S \leq \Lambda} I_q(X_1, X_2; Y|U) \end{array} \right\}. \quad (61)$$

Now, based on Theorem 5 (see [27]), we deduce that the deterministic code capacity region is the same as the divided-randomness capacity region, *i.e.*  $\mathbb{C}(\mathcal{A}) = \mathbb{C}^*(\mathcal{A})$ . For the latter, Gubner and Hughes derive inner and outer bounds that are close enough in order to establish that  $\mathbb{C}^*(\mathcal{A})$  is non convex for high values of  $r \geq 2$ . This, in turn, implies that the deterministic code capacity region is also non convex in general.

*Example 4.* Let  $\mathcal{A}$  be an AVMAC which consists of independent binary symmetric channels. Specifically, let the state and the output be pairs as well, *i.e.*  $S = (S_1, S_2)$  and  $Y = (Y_1, Y_2)$ , such that

$$\begin{aligned} Y_1 &= X_1 + S_1 \pmod{2}, \\ Y_2 &= X_2 + S_2 \pmod{2}, \end{aligned} \quad (62)$$

where  $X_1, X_2, S_1, S_2, Y_1$  and  $Y_2$  are binary. Suppose that the input and state cost functions are Hamming weights, *i.e.*

$$\phi_1(x_1) = x_1, \phi_2(x_2) = x_2, l(s) = s_1 + s_2, \quad (63)$$

while the constraints  $\Omega_1, \Omega_2$  and  $\Lambda$  are in the interval  $(0, 1]$ .

First, we use Theorem 7 to show that the random code capacity is given by

$$\mathbb{C}^*(\mathcal{A}) = \left\{ (R_1, R_2) : \begin{array}{l} R_1 \leq h(\omega_1 * \lambda) - h(\lambda), \\ R_2 \leq h(\omega_2 * \lambda) - h(\lambda) \end{array} \right\}, \quad (64)$$

where

$$\omega_1 = \min\left(\Omega_1, \frac{1}{2}\right), \omega_2 = \min\left(\Omega_2, \frac{1}{2}\right), \lambda = \min\left(\Lambda, \frac{1}{2}\right). \quad (65)$$

In particular, if  $\Lambda \geq \frac{1}{2}$ , then the random code capacity region is  $\mathbb{C}^*(\mathcal{A}) = \{(0, 0)\}$ .

It can further be seen that the binary AVMAC is symmetrizable- $\mathcal{X}_1 \times \mathcal{X}_2$ , symmetrizable- $\mathcal{X}_1|\mathcal{X}_2$ , and symmetrizable- $\mathcal{X}_2|\mathcal{X}_1$ . In particular, the symmetrizability equations (13), (14), and (15) only hold with the 0-1 laws  $J(s|x_1, x_2) = \delta(s_1 - x_1)\delta(s_2 - x_2)$ ,  $J_1(s|x_1) = \delta(s_1 - x_1)\delta(s_2 - x'_2)$ ,  $J_2(s|x_2) = \delta(s_1 - x'_1)\delta(s_2 - x_2)$ , for arbitrary  $x'_1, x'_2 \in \{0, 1\}$ , where  $\delta(u) = 1$  for  $u = 0$ , and  $\delta(u) = 0$  otherwise.

Then, we use Corollary 9 to show that the capacity region is given by the following. If  $\Omega_1 > \Lambda$  and  $\Omega_2 > \Lambda$ , then

$$\mathbb{C}(\mathcal{A}) = \mathbb{C}^*(\mathcal{A}) = \left\{ (R_1, R_2) : \begin{array}{l} R_1 \leq h(\omega_1 * \lambda) - h(\lambda), \\ R_2 \leq h(\omega_2 * \lambda) - h(\lambda) \end{array} \right\}, \quad (66)$$

where  $h(t) = -t \log t - (1-t) \log(1-t)$  for  $0 < t < 1$ , and  $\alpha * \beta = (1-\alpha)\beta + \alpha(1-\beta)$ . If  $\Omega_1 \leq \Lambda$  and  $\Omega_2 > \Lambda$ , then

$$\mathbb{C}(\mathcal{A}) = \{(0, R_2) : R_2 \leq h(\omega_2 * \lambda) - h(\lambda)\}. \quad (67)$$

If  $\Omega_1 > \Lambda$  and  $\Omega_2 \leq \Lambda$ , then

$$\mathbb{C}(\mathcal{A}) = \{(R_1, 0) : R_1 \leq h(\omega_1 * \lambda) - h(\lambda)\}. \quad (68)$$

Otherwise, if  $\Omega_1 \leq \Lambda$  and  $\Omega_2 \leq \Lambda$ , then

$$\mathbb{C}(\mathcal{A}) = \{(0, 0)\}. \quad (69)$$

The analysis is given in Appendix G.

We observe that the deterministic code capacity region and the random code capacity region are the same, only if  $\Lambda \geq \frac{1}{2}$  or both input constraints are higher than the state constraints. In all other cases, the deterministic code capacity is strictly included within the random code capacity region, *i.e.*  $\mathbb{C}(\mathcal{A}) \subset \mathbb{C}^*(\mathcal{A})$ .

*Example 5.* Consider the Gaussian AVMAC, specified by

$$Y = X_1 + X_2 + S + Z, \quad (70)$$

with  $Z \sim \mathcal{N}(0, \sigma^2)$ , where the transmitters and the jammer have the power constraints  $\frac{1}{n} \sum_{i=1}^n X_{k,i}^2 \leq \Omega_k$ , for  $k = 1, 2$ , and  $\frac{1}{n} \sum_{i=1}^n S_i^2 \leq \Lambda$ . This channel was treated independently by Hosseinigoki and Kosut [31], using the packing lemmas from [30]. Here, we use our results on the general AVMAC.

Although we previously assumed that the input, state and output alphabets are finite, our results can be extended to the continuous case as well, using standard discretization techniques [7, 1, 15] [21, Section 3.4.1]. First, we use Theorem 7 to show that the random code capacity region is

$$\mathbb{C}^*(\mathcal{A}) = \left\{ (R_1, R_2) : \begin{array}{l} R_1 \leq \frac{1}{2} \log \left( 1 + \frac{\Omega_1}{\Lambda + \sigma^2} \right), \\ R_2 \leq \frac{1}{2} \log \left( 1 + \frac{\Omega_2}{\Lambda + \sigma^2} \right), \\ R_1 + R_2 \leq \frac{1}{2} \log \left( 1 + \frac{\Omega_1 + \Omega_2}{\Lambda + \sigma^2} \right) \end{array} \right\}. \quad (71)$$

Then, we use Theorem 8 to show that the capacity region is given by the following. If  $\Omega_1 > \Lambda$  and  $\Omega_2 > \Lambda$ , then

$$\mathbb{C}(\mathcal{A}) = \mathbb{C}^*(\mathcal{A}) = \left\{ (R_1, R_2) : \begin{array}{l} R_1 \leq \frac{1}{2} \log \left( 1 + \frac{\Omega_1}{\Lambda + \sigma^2} \right), \\ R_2 \leq \frac{1}{2} \log \left( 1 + \frac{\Omega_2}{\Lambda + \sigma^2} \right), \\ R_1 + R_2 \leq \frac{1}{2} \log \left( 1 + \frac{\Omega_1 + \Omega_2}{\Lambda + \sigma^2} \right) \end{array} \right\}. \quad (72)$$

If  $\Omega_1 \leq \Lambda$  and  $\Omega_2 > \Lambda$ , then

$$\mathbb{C}(\mathcal{A}) = \left\{ (0, R_2) : R_2 \leq \frac{1}{2} \log \left( 1 + \frac{\Omega_2}{\Lambda + \sigma^2} \right) \right\}. \quad (73)$$

If  $\Omega_1 > \Lambda$  and  $\Omega_2 \leq \Lambda$ , then

$$\mathbb{C}(\mathcal{A}) = \left\{ (R_1, 0) : R_1 \leq \frac{1}{2} \log \left( 1 + \frac{\Omega_1}{\Lambda + \sigma^2} \right) \right\}. \quad (74)$$

Otherwise, if  $\Omega_1 \leq \Lambda$  and  $\Omega_2 \leq \Lambda$ , then

$$\mathbb{C}(\mathcal{A}) = \{(0, 0)\}. \quad (75)$$

The analysis is given in Appendix H. We observe that the deterministic code capacity region and the random code capacity region are the same, only if both input constraints are higher than the state constraints. In all other cases, the deterministic code capacity is strictly included within the random code capacity region.

1) *Without Constraints:* We have seen in Subsection II-E1 that Gubner [24] and Ahlswede and Cai [3] determined the capacity region in all but two cases, where User 1 has zero capacity and User 2 has positive capacity, and vice versa (see Remark 3). In this subsection, we give full characterization of the capacity region of the AVMAC without constraints, closing the gap in the results by Ahlswede and Cai [3].

*Theorem 12.* The capacity region of the AVMAC free of constraints is given by the following.

a) If  $W_{Y|X_1, X_2, S}$  is not symmetrizable- $\mathcal{X}_1 \times \mathcal{X}_2$ ,  $-\mathcal{X}_1|\mathcal{X}_2$ , nor  $-\mathcal{X}_2|\mathcal{X}_1$ , then

$$\mathbb{C}(\mathcal{A}_{\text{free}}) = \mathbb{C}^{**}(\mathcal{A}_{\text{free}}). \quad (76)$$

b) If  $W_{Y|X_1, X_2, S}$  is not symmetrizable- $\mathcal{X}_1 \times \mathcal{X}_2$  nor  $-\mathcal{X}_2|\mathcal{X}_1$ , but symmetrizable- $\mathcal{X}_1|\mathcal{X}_2$ , then

$$\mathbb{C}(\mathcal{A}_{\text{free}}) = \left\{ (0, R_2) : R_2 \leq \min_{q(s)} \max_{p(x_1)p(x_2)} I_q(X_2; Y|X_1) \right\}. \quad (77)$$

c) If  $W_{Y|X_1, X_2, S}$  is not symmetrizable- $\mathcal{X}_1 \times \mathcal{X}_2$  nor  $-\mathcal{X}_1|\mathcal{X}_2$ , but symmetrizable- $\mathcal{X}_2|\mathcal{X}_1$ , then

$$\mathbb{C}(\mathcal{A}_{\text{free}}) = \left\{ (R_1, 0) : R_1 \leq \min_{q(s)} \max_{p(x_1)p(x_2)} I_q(X_1; Y|X_2) \right\}. \quad (78)$$

d) In all other cases,

$$\mathbb{C}(\mathcal{A}_{\text{free}}) = \{(0, 0)\}. \quad (79)$$

Theorem 12 is a direct consequence of our previous results in the presence of constraints, since plugging  $\Omega_k > \phi_{k, \max}$ ,  $k = 1, 2$ , and  $\Lambda > l_{\max}$  yields the capacity region without constraints. In particular, if the AVMAC free of constraints is non-symmetrizable- $\mathcal{X}_1 \times \mathcal{X}_2$ , then  $L^* = +\infty > \Lambda$ . Otherwise, if the AVMAC free of constraints is symmetrizable- $\mathcal{X}_1 \times \mathcal{X}_2$ , then  $L^* \leq l_{\max} < \Lambda$ . A similar argument holds for  $L_1^*$ ,  $L_2^*$  and non-symmetrizability- $\mathcal{X}_1|\mathcal{X}_2$ ,  $-\mathcal{X}_2|\mathcal{X}_1$ , respectively. Thus, Theorem 12 follows from Theorem 8 and Definition 6. The theorem above completes the partial characterization in Cases b) and c) in Theorem 4.

## APPENDIX A PROOF OF LEMMA 6

Consider the compound MAC  $\mathcal{A}^{\mathcal{Q}}$  under input constraints  $(\Omega_1, \Omega_2)$  and state constraint  $\Lambda$ . To prove the direct part, we construct a code based on simultaneous decoding with respect to a state type which is “close” to some  $q \in \mathcal{Q}$ . The converse part follows by standard arguments.

### A. Achievability Proof

Let  $\varepsilon, \delta > 0$  be arbitrarily small. We use the following notation. Basic method of types concepts are defined as in [16, Chapter 2]; including the definition of a type  $\hat{P}_{x^n}$  of a sequence  $x^n$ ; a joint type  $\hat{P}_{x^n, y^n}$  and a conditional type  $\hat{P}_{x^n|y^n}$  of a pair of sequences  $(x^n, y^n)$ ; and a  $\delta$ -typical set  $\mathcal{A}^\delta(P_{X,Y})$  with respect to a distribution  $P_{X,Y}(x, y)$ . We also define a set of state types  $\hat{\mathcal{Q}}_n$  by

$$\hat{\mathcal{Q}}_n = \{\hat{P}_{s^n} : s^n \in \mathcal{A}^{\delta_1}(q) \text{ for some } q \in \overline{\mathcal{P}}_\Lambda(\mathcal{S})\}, \quad (80)$$

where

$$\delta_1 \triangleq \frac{\delta}{2 \cdot |\mathcal{S}|}. \quad (81)$$

Namely,  $\hat{\mathcal{Q}}_n$  is the set of types that are  $\delta_1$ -close to some state distribution  $q(s)$  in  $\overline{\mathcal{P}}_\Lambda(\mathcal{S})$ . Then, fix  $P_U P_{X_1|U} P_{X_2|U}$  such that  $\mathbb{E}\phi_k(X_k) \leq \Omega_k - \varepsilon$ , for  $k = 1, 2$ .

*Codebook Generation:* Generate a random time sharing sequence  $u^n \sim \prod_{i=1}^n P_U(u_i)$ . Then, generate  $2^{nR_k}$  conditionally independent sequences  $x_k^n(m_k)$ ,  $m_k \in [1 : 2^{nR_k}]$ , at random, each according to  $\prod_{i=1}^n P_{X_k|U}(x_{k,i}|u_i)$ , for  $k = 1, 2$ . Reveal the sequence  $u^n$  and the codebooks  $\{x_1^n(m_1)\}$  and  $\{x_2^n(m_2)\}$  to the encoders and the decoder.

*Encoding:* To send  $(m_1, m_2)$ , Encoder  $k$  transmits  $x_k^n(m_k)$ , provided that

$$\phi_k(x_k^n(m_k)) \leq \Omega_k, \text{ for } k = 1, 2. \quad (82)$$

Otherwise, repeatedly send the symbol  $a_k$  with  $\phi_k(a_k) = 0$ .

*Decoding:* For every state distribution  $q(s)$ , define

$$P_{Y|X_1, X_2}^q(y|x_1, x_2) = \sum_{s \in \mathcal{S}} q(s) W_{Y|X_1, X_2, S}(y|x_1, x_2, s). \quad (83)$$

As  $y^n$  is received, the decoder finds a unique pair  $(\hat{m}_1, \hat{m}_2) \in [1 : 2^{nR_1}] \times [1 : 2^{nR_2}]$  such that  $(u^n, x_1^n(\hat{m}_1), x_2^n(\hat{m}_2), y^n) \in \mathcal{A}^\delta(P_U P_{X_1|U} P_{X_2|U} P_{Y|X_1, X_2}^q)$  for some type  $q \in \hat{\mathcal{Q}}_n$ . If there is none, or more than one such pair, declare an error.

*Analysis of Probability of Error:* Assume without loss of generality that the users sent the messages  $m_1 = m_2 = 1$ . Let  $q(s) \in \mathcal{Q}$  denote the *actual* state distribution chosen by the jammer. The error event is within the union of the following events,

$$\mathcal{E}_1 = \{(U^n, X_1^n(1), X_2^n(1)) \notin \mathcal{A}^{\delta/3}(P_U P_{X_1|U} P_{X_2|U})\}, \quad (84)$$

$$\mathcal{E}_2 = \{(U^n, X_1^n(1), X_2^n(1), Y^n) \notin \mathcal{A}^\delta(P_U P_{X_1|U} P_{X_2|U} P_{Y|X_1, X_2}^{q'})\}, \text{ for all } q' \in \hat{\mathcal{Q}}_n\}, \quad (85)$$

$$\mathcal{E}_3 = \{(U^n, X_1^n(m_1), X_2^n(m_2), Y^n) \in \mathcal{A}^\delta(P_U P_{X_1|U} P_{X_2|U} P_{Y|X_1, X_2}^{q'})\}, \text{ for some } m_1 \neq 1, m_2 \neq 1, q' \in \hat{\mathcal{Q}}_n\}, \quad (86)$$

$$\mathcal{E}_4 = \{(U^n, X_1^n(m_1), X_2^n(1), Y^n) \in \mathcal{A}^\delta(P_U P_{X_1|U} P_{X_2|U} P_{Y|X_1, X_2}^{q'})\}, \text{ for some } m_1 \neq 1, q' \in \hat{\mathcal{Q}}_n\}, \quad (87)$$

$$\mathcal{E}_5 = \{(U^n, X_1^n(1), X_2^n(m_2), Y^n) \in \mathcal{A}^\delta(P_U P_{X_1|U} P_{X_2|U} P_{Y|X_1, X_2}^{q'})\}, \text{ for some } m_2 \neq 1, q' \in \hat{\mathcal{Q}}_n\}. \quad (88)$$

The probability of error is then bounded by

$$P_e^{(n)}(q, \mathcal{C}) \leq \Pr(\mathcal{E}_1) + \Pr(\mathcal{E}_2 | \mathcal{E}_1^c) + \Pr(\mathcal{E}_3 | \mathcal{E}_2^c) + \Pr(\mathcal{E}_4 | \mathcal{E}_3^c) + \Pr(\mathcal{E}_5 | \mathcal{E}_4^c), \quad (89)$$

where the conditioning on  $(M_1, M_2) = (1, 1)$  is omitted for convenience of notation. The first term in the RHS of (89) tends to zero exponentially as  $n \rightarrow \infty$ , by the law of large numbers and Chernoff's bound (see e.g. [35, Theorem 1.1]). Now, given

that the event  $\mathcal{E}_1^c$  occurs, we have that  $X_1^n(1)$  and  $X_2^n(1)$  satisfy the input constraints (82), for sufficiently small  $\delta > 0$ , and are thus the channel inputs.

Moving to the second term, suppose that

$$(U^n, X_1^n(1), X_2^n(1), Y^n) \in \mathcal{A}^{\delta/2}(P_U P_{X_1|U} P_{X_2|U} P_{Y|X_1, X_2}^q). \quad (90)$$

Then, for sufficiently large  $n$ , there is a type  $q'(s)$  such that  $|q'(s) - q(s)| \leq \delta_1$ , for all  $s \in \mathcal{S}$ , hence,  $q' \in \hat{\mathcal{Q}}_n$  (see definition in (80)), and

$$|P_{Y|X_1, X_2}^{q'}(y|x_1, x_2) - P_{Y|X_1, X_2}^q(y|x_1, x_2)| \leq |\mathcal{S}| \cdot \delta_1 = \frac{\delta}{2}, x_k \in \mathcal{X}_k, y \in \mathcal{Y}, \quad (91)$$

for  $\delta_1 = \delta/2|\mathcal{S}|$  (see (83)), hence  $\mathcal{E}_2$  does not hold. It follows by contradiction that

$$\Pr(\mathcal{E}_2 | \mathcal{E}_1^c) \leq \Pr\left((U^n, X_1^n(1), X_2^n(1), Y^n) \notin \mathcal{A}^{\delta/2}(P_U P_{Y|U}^q) | \mathcal{E}_1^c\right), \quad (92)$$

which tends to zero exponentially as  $n \rightarrow \infty$  by the law of large numbers and Chernoff's bound.

As for the third term in the RHS of (89), by the union of events bound and the fact that the number of type classes in  $\mathcal{S}^n$  is bounded by  $(n+1)^{|\mathcal{S}|}$ , we have that

$$\begin{aligned} \Pr(\mathcal{E}_3 | \mathcal{E}_2^c) &\leq (n+1)^{|\mathcal{S}|} \cdot \sup_{q' \in \hat{\mathcal{Q}}_n} \\ &\Pr\left((U^n, X_1^n(m_1), X_2^n(m_2), Y^n) \in \mathcal{A}^\delta(P_U P_{X_1|U} P_{X_2|U} P_{Y|X_1, X_2}^{q'}), \text{ for some } m_1 \neq 1, m_2 \neq 1 | \mathcal{E}_1^c\right) \\ &\leq (n+1)^{|\mathcal{S}|} \cdot 2^{n(R_1+R_2)} \cdot \sup_{q' \in \hat{\mathcal{Q}}_n} \left[ \sum_{u^n, x_1^n, x_2^n} P_{U^n}(u^n) P_{X_1^n|U^n}(x_1^n|u^n) P_{X_2^n|U^n}(x_2^n|u^n) \right. \\ &\quad \cdot \left. \sum_{y^n : (u^n, x_1^n, x_2^n, y^n) \in \mathcal{A}^\delta(P_U P_{X_1|U} P_{X_2|U} P_{Y|X_1, X_2}^{q'})} P_{Y^n}^q(y^n|u^n) \right], \quad (93) \end{aligned}$$

where we have defined  $P_{Y|U}^q(y|u) = \sum_{x_1, x_2, s \in \mathcal{S}} P_{X_1|U}(x_1|u) P_{X_2|U}(x_2|u) q(s) W_{Y|X_1, X_2, \mathcal{S}}(y|x_1, x_2, s)$ . This follows since  $X_1^n(m_1)$  and  $X_2^n(m_2)$  are independent of  $Y^n$  for every  $m_1 \neq 1$  and  $m_2 \neq 1$ . Let  $y^n$  satisfy  $(u^n, x_1^n, x_2^n, y^n) \in \mathcal{A}^\delta(P_U P_{X_1|U} P_{X_2|U} P_{Y|X_1, X_2}^{q'})$ . Then,  $(u^n, y^n) \in \mathcal{A}^{\delta_2}(P_{U, Y}^{q'})$  with  $\delta_2 \triangleq |\mathcal{X}_1| |\mathcal{X}_2| \cdot \delta$ . By Lemmas 2.6 and 2.7 in [16],

$$P_{Y^n|U^n}^q(y^n|u^n) = 2^{-n(H(\hat{P}_{y^n|u^n}) + D(\hat{P}_{y^n|u^n} || P_{Y|U}^{q'}))} \leq 2^{-nH(\hat{P}_{y^n|u^n})} \leq 2^{-n(H_{q'}(Y|U) - \varepsilon_1(\delta))}, \quad (94)$$

where  $\varepsilon_1(\delta) \rightarrow 0$  as  $\delta \rightarrow 0$ . Therefore, by (93)–(94), along with [16, Lemma 2.13],

$$\Pr(\mathcal{E}_3 | \mathcal{E}_2^c) \leq (n+1)^{|\mathcal{S}|} \cdot \sup_{q' \in \hat{\mathcal{Q}}_n} 2^{-n[I_{q'}(X_1, X_2; Y|U) - R_1 - R_2 - \varepsilon_2(\delta)]}, \quad (95)$$

with  $\varepsilon_2(\delta) \rightarrow 0$  as  $\delta \rightarrow 0$ . The RHS of (95) tends to zero exponentially as  $n \rightarrow \infty$ , provided that

$$R_1 + R_2 < \inf_{q' \in \mathcal{Q}} I_{q'}(X_1, X_2; Y|U) - \varepsilon_2(\delta). \quad (96)$$

By similar considerations, the fourth term is bounded by  $\Pr(\mathcal{E}_4 | \mathcal{E}_2^c) \leq (n+1)^{|\mathcal{S}|} \cdot \sup_{q' \in \hat{\mathcal{Q}}_n} 2^{-n[I_{q'}(X_1; Y|X_2, U) - R_1 - \varepsilon_3(\delta)]}$ , with  $\varepsilon_3(\delta) \rightarrow 0$  as  $\delta \rightarrow 0$ . This bound tends to zero exponentially as  $n \rightarrow \infty$ , provided that  $R_1 < \inf_{q' \in \mathcal{Q}} I_{q'}(X_1; Y|X_2, U) - \varepsilon_3(\delta)$ . By symmetry, we have that  $\Pr(\mathcal{E}_5 | \mathcal{E}_2^c)$  tends to zero as well, provided that  $R_2 < \inf_{q' \in \mathcal{Q}} I_{q'}(X_2; Y|X_1, U) - \varepsilon_3(\delta)$ .

We conclude that the probability of error, averaged over the class of the codebooks, exponentially decays to zero as  $n \rightarrow \infty$ . Therefore, there must exist a  $(2^{nR_1}, 2^{nR_2}, n, \varepsilon)$  deterministic code, for a sufficiently large  $n$ .  $\square$

### B. Converse Proof

The converse part follows from the same arguments as in the converse proof of the classical MAC [4] (see also [14, Section 15.3.4]). Since the deterministic code capacity region is always bounded by the random code capacity region, we consider a sequence of  $(2^{nR_1}, 2^{nR_2}, n, \alpha_n)$  random codes, where  $\alpha_n \rightarrow 0$  as  $n \rightarrow \infty$ . Then, let  $X_1^n = f_{1, \gamma}^n(M_1)$  and  $X_2^n = f_{2, \gamma}^n(M_2)$  be the channel input sequences, and  $Y^n$  be the corresponding output sequence, where  $\gamma \in \Gamma$  is the random element shared between the encoders and the decoder. For every  $q \in \mathcal{Q}$ , we have by Fano's inequality that  $H_q(M_1, M_2 | Y^n, \gamma) \leq n\varepsilon_n$ , hence  $H_q(M_1 | M_2, Y^n, \gamma) \leq n\varepsilon_n$  and  $H_q(M_2 | M_1, Y^n, \gamma) \leq n\varepsilon_n$ , where  $\varepsilon_n \rightarrow 0$  as  $n \rightarrow \infty$ . Since

$$n(R_1 + R_2) = H(M_1, M_2 | \gamma) = I_q(M_1, M_2; Y^n | \gamma) + H(M_1, M_2 | Y^n, \gamma), \quad (97)$$

$$nR_1 = H(M_1 | M_2, \gamma) = I_q(M_1; Y^n | M_2, \gamma) + H(M_1 | M_2, Y^n, \gamma), \quad (98)$$

$$nR_2 = H(M_2 | M_1, \gamma) = I_q(M_2; Y^n | M_1, \gamma) + H(M_2 | M_1, Y^n, \gamma), \quad (99)$$



it follows that

$$nR_1 \leq I_q(M_1; Y^n | M_2, \gamma) + n\varepsilon_n = \sum_{i=1}^n I_q(M_1; Y_i | M_2, Y^{i-1}, \gamma) + n\varepsilon_n, \quad (100)$$

$$nR_2 \leq I_q(M_2; Y^n | M_1, \gamma) + n\varepsilon_n = \sum_{i=1}^n I_q(M_2; Y_i | M_1, Y^{i-1}, \gamma) + n\varepsilon_n, \quad (101)$$

$$n(R_1 + R_2) \leq I_q(M_1, M_2; Y^n | \gamma) + n\varepsilon_n = \sum_{i=1}^n I_q(M_1, M_2; Y_i | Y^{i-1}, \gamma) + n\varepsilon_n. \quad (102)$$

As  $X_k^n = f_{k,\gamma}(M_k)$ , this yields

$$R_1 \leq \frac{1}{n} \sum_{i=1}^n I_q(X_{1,i}, M_1; Y^n | X_{2,i}, M_2, Y^{i-1}, \gamma) + \varepsilon_n \quad (103)$$

$$R_2 \leq \frac{1}{n} \sum_{i=1}^n I_q(X_{2,i}, M_2; Y^n | X_{1,i}, M_1, Y^{i-1}, \gamma) + \varepsilon_n \quad (104)$$

$$R_1 + R_2 \leq \frac{1}{n} \sum_{i=1}^n I_q(X_{1,i}, X_{2,i}, M_1, M_2; Y_i | Y^{i-1}, \gamma) + \varepsilon_n. \quad (105)$$

Then, since  $(\gamma, M_1, M_2, Y^{i-1}) \ominus (X_{1,i}, X_{2,i}) \ominus Y_i$  form a Markov chain, we have that for every  $q \in \mathcal{Q}$ ,

$$R_1 \leq I_q(X_{1,T}; Y_T | X_{2,T}, T, \gamma) + \varepsilon_n, \quad (106)$$

$$R_2 \leq I_q(X_{2,T}; Y_T | X_{1,T}, T, \gamma) + \varepsilon_n, \quad (107)$$

$$R_1 + R_2 \leq I_q(X_{1,T}, X_{2,T}; Y_T | T, \gamma) + \varepsilon_n, \quad (108)$$

where  $T$  is a random variable which is uniformly distributed over  $[1 : n]$ , and independent of  $(\gamma, X_1^n, X_2^n, S^n, Y^n)$ . Defining  $X_1 = X_{1,T}$ ,  $X_2 = X_{2,T}$ ,  $Y = Y_T$ , and  $U = (T, \gamma)$ , it follows that

$$R_1 \leq \inf_{q \in \mathcal{Q}} I_q(X_1; Y | X_2, U) + \varepsilon_n, \quad (109)$$

$$R_2 \leq \inf_{q \in \mathcal{Q}} I_q(X_2; Y | X_1, U) + \varepsilon_n, \quad (110)$$

$$R_1 + R_2 \leq \inf_{q \in \mathcal{Q}} I_q(X_1, X_2; Y | U) + \varepsilon_n. \quad (111)$$

As  $X_1$  and  $X_2$  are conditionally independent given  $U$ , this completes the proof of the converse part.  $\square$

## APPENDIX B PROOF OF THEOREM 7

Consider the AVMAC  $\mathcal{A}$  under input constraints  $(\Omega_1, \Omega_2)$  and state constraint  $\Lambda$ .

### A. Achievability Proof

To prove the random code capacity theorem for the AVMAC, we use our result on the compound MAC along with a simple extension of Ahlswede's Robustification Technique (RT). We begin with a lemma from [41], based on Ahlswede's RT [2].

*Lemma 13* (Ahlswede's RT [2] [41, Lemma 9]). Let  $h : \mathcal{S}^n \rightarrow [0, 1]$  be a given function. If, for some fixed  $\alpha_n \in (0, 1)$ , and for all  $q^n(s^n) = \prod_{i=1}^n q(s_i)$ , with  $q \in \overline{\mathcal{P}}_\Lambda(\mathcal{S})$ ,

$$\sum_{s^n \in \mathcal{S}^n} q^n(s^n) h(s^n) \leq \alpha_n, \quad (112)$$

then,

$$\frac{1}{n!} \sum_{\pi \in \Pi_n} h(\pi s^n) \leq \beta_n, \quad \text{for all } s^n \in \mathcal{S}^n \text{ such that } l^n(s^n) \leq \Lambda, \quad (113)$$

where  $\Pi_n$  is the set of all  $n$ -tuple permutations  $\pi : \mathcal{S}^n \rightarrow \mathcal{S}^n$ , and  $\beta_n = (n+1)^{|\mathcal{S}|} \cdot \alpha_n$ .

Let  $(R_1, R_2) \in C^*(\mathcal{A})$ . At first, we consider the compound MAC under input constraints  $(\Omega_1, \Omega_2)$ , with  $\mathcal{Q} = \overline{\mathcal{P}}_\Lambda(\mathcal{S})$ . According to Lemma 6, for some  $\theta > 0$  and sufficiently large  $n$ , there exists a  $(2^{nR_1}, 2^{nR_2}, n)$  code  $\mathcal{C} = (f_1^n(m_1), f_2^n(m_2), g(y^n))$  for the compound MAC  $\mathcal{A}^{\overline{\mathcal{P}}_\Lambda(\mathcal{S})}$  such that

$$\phi_k^n(f_k(m_k)) \leq \Omega_k, \quad \text{for all } m_k \in [1 : 2^{nR_k}], k = 1, 2, \quad (114)$$

and

$$P_e^{(n)}(q, \mathcal{C}) = \sum_{s^n \in \mathcal{S}^n} q(s^n) \cdot P_{e|s^n}^{(n)}(\mathcal{C}) \leq e^{-2\theta n}, \quad (115)$$

for all i.i.d. state distributions  $q(s^n) = \prod_{i=1}^n q(s_i)$ , with  $q \in \overline{\mathcal{P}}_\Lambda(\mathcal{S})$ .

Therefore, by Lemma 13, taking  $h_0(s^n) = P_{e|s^n}^{(n)}(\mathcal{C})$  and  $\alpha_n = e^{-2\theta n}$ , we have that for a sufficiently large  $n$ ,

$$\frac{1}{n!} \sum_{\pi \in \Pi_n} \mathbb{E} P_{e|\pi s^n}^{(n)}(\mathcal{C}) \leq (n+1)^{|\mathcal{S}|} e^{-2\theta n} \leq e^{-\theta n}, \quad (116)$$

for all  $s^n \in \mathcal{S}^n$  with  $l^n(s^n) \leq \Lambda$ , where the sum is over the set of all  $n$ -tuple permutations.

On the other hand, for every  $\pi \in \Pi_n$ ,

$$\begin{aligned} P_{e|\pi s^n}^{(n)}(\mathcal{C}) &\stackrel{(a)}{=} \frac{1}{2^{n(R_1+R_2)}} \sum_{m_1, m_2} \sum_{y^n: g(y^n) \neq (m_1, m_2)} W_{Y^n|X_1^n, X_2^n, S^n}(y^n | f_1(m_1), f_2(m_2), \pi s^n) \\ &\stackrel{(b)}{=} \frac{1}{2^{n(R_1+R_2)}} \sum_{m_1, m_2} \sum_{y^n: g(\pi y^n) \neq (m_1, m_2)} W_{Y^n|X_1^n, X_2^n, S^n}(\pi y^n | f_1(m_1), f_2(m_2), \pi s^n) \\ &\stackrel{(c)}{=} \frac{1}{2^{n(R_1+R_2)}} \sum_{m_1, m_2} \sum_{y^n: g(\pi y^n) \neq (m_1, m_2)} W_{Y^n|X_1^n, X_2^n, S^n}(y^n | \pi^{-1} f_1(m_1), \pi^{-1} f_2(m_2), s^n), \end{aligned} \quad (117)$$

where (a) is obtained by plugging  $\pi s^n$  in (6a); in (b) we simply change the order of summation over  $y^n$ ; and (c) holds because the channel is memoryless.

Then, consider the  $(2^{nR_1}, 2^{nR_2}, n)$  random code  $\mathcal{C}^\Pi$ , specified by

$$f_{1,\pi}(m_1) = \pi^{-1} f_1(m_1), \quad f_{2,\pi}(m_2) = \pi^{-1} f_2(m_2), \quad g_\pi(y^n) = g(\pi y^n), \quad (118)$$

with a uniform distribution  $\mu(\pi) = \frac{1}{|\Pi_n|} = \frac{1}{n!}$  for  $\pi \in \Pi_n$ . As the inputs cost is additive (see (1)), the permutation does not affect the costs of the codewords, hence the random code satisfies the input constraints  $(\Omega_1, \Omega_2)$ . From (117), we see that  $P_{e|s^n}^{(n)}(\mathcal{C}^\Pi) = \sum_{\pi \in \Pi_n} \mu(\pi) \cdot \mathbb{E} P_{e|\pi s^n}^{(n)}(\mathcal{C})$ , for all  $s^n \in \mathcal{S}^n$  with  $l^n(s^n) \leq \Lambda$ . Therefore, together with (116), we have that the probability of error of the random code  $\mathcal{C}^\Pi$  is bounded by  $P_e^{(n)}(q, \mathcal{C}^\Pi) \leq e^{-\theta n}$ , for every  $q(s^n) \in \mathcal{P}_\Lambda(\mathcal{S}^n)$ . It follows that  $\mathcal{C}^\Pi$  is a  $(2^{nR}, n, e^{-\theta n})$  random code for the AVMAC  $\mathcal{A}$  under input constraints  $(\Omega_1, \Omega_2)$  and state constraint  $\Lambda$ .  $\square$

## B. Converse Proof

Assume to the contrary that there exists an achievable rate pair

$$(R_1, R_2) \notin \mathbb{C}(\mathcal{A}^{\mathcal{Q}}) \Big|_{\mathcal{Q}=\mathcal{P}_{\Lambda-\delta}}, \quad (119)$$

using random codes over the AVMAC  $\mathcal{A}$  under input constraints  $(\Omega_1, \Omega_2)$  and state constraint  $\Lambda$ , where  $\delta > 0$  is arbitrarily small. That is, for every  $\varepsilon > 0$  and sufficiently large  $n$ , there exists a  $(2^{nR_1}, 2^{nR_2}, n)$  random code  $\mathcal{C}^\Gamma = (\mu, \Gamma, \{\mathcal{C}_\gamma\}_{\gamma \in \Gamma})$  for the AVMAC  $\mathcal{A}$ , such that  $\sum_{\gamma \in \Gamma} \mu(\gamma) \phi_k^n(f_{k,\gamma}(m_k)) \leq \Omega_k$ , for  $k = 1, 2$ , and

$$P_e^{(n)}(q, \mathcal{C}^\Gamma) \leq \varepsilon, \quad (120)$$

for all  $m_1 \in [1 : 2^{nR_1}]$ ,  $m_2 \in [1 : 2^{nR_2}]$ , and  $q(s^n) \in \mathcal{P}_\Lambda(\mathcal{S}^n)$ . In particular, for distributions  $q(\cdot)$  which give mass 1 to some sequence  $s^n \in \mathcal{S}^n$  with  $l^n(s^n) \leq \Lambda$ , we have that  $P_{e|s^n}^{(n)}(\mathcal{C}^\Gamma) \leq \varepsilon$ .

Consider using the random code  $\mathcal{C}^\Gamma$  over the compound MAC  $\mathcal{W}^{\overline{\mathcal{P}}_{\Lambda-\delta}(\mathcal{S})}$  under input constraints  $(\Omega_1, \Omega_2)$ . Let  $\overline{q}(s) \in \overline{\mathcal{P}}_{\Lambda-\delta}(\mathcal{S})$  be a given state distribution. Then, define a sequence of i.i.d. random variables  $\overline{S}_1, \dots, \overline{S}_n \sim \overline{q}(s)$ . Letting  $\overline{q}^n(s^n) \triangleq \prod_{i=1}^n \overline{q}(s_i)$ , the probability of error is bounded by

$$P_e^{(n)}(\overline{q}, \mathcal{C}^\Gamma) \leq \sum_{s^n: l^n(s^n) \leq \Lambda} \overline{q}^n(s^n) P_{e|s^n}^{(n)}(\mathcal{C}^\Gamma) + \Pr(l^n(\overline{S}^n) > \Lambda). \quad (121)$$

The first sum is bounded by (120), and the second term vanishes by the law of large numbers, since  $\overline{q}(s) \in \overline{\mathcal{P}}_{\Lambda-\delta}(\mathcal{S})$ . It follows that the random code  $\mathcal{C}^\Gamma$  achieves a rate pair  $(R_1, R_2)$  as in (119) over the compound MAC  $\mathcal{A}^{\overline{\mathcal{P}}_{\Lambda-\delta}(\mathcal{S})}$  under input constraints  $(\Omega_1, \Omega_2)$ , for an arbitrarily small  $\delta > 0$ , in contradiction to Lemma 6. We deduce that the assumption is false, and  $\mathbb{C}^*(\mathcal{A}) \subseteq \mathbb{C}(\mathcal{A}^{\mathcal{Q}}) \Big|_{\mathcal{Q}=\mathcal{P}_\Lambda} = \mathbb{C}^*(\mathcal{A})$ .  $\square$

APPENDIX C  
PROOF OF LEMMA 10

Consider the AVMAC  $\mathcal{A}$  under input constraints  $(\Omega_1, \Omega_2)$  and state constraint  $\Lambda$ . We accommodate the proof of Lemma 1 in [3] to the case where there are state constraints. We prove part 2 of the lemma, and the rest follows by similar considerations.

Our first step is to extend an auxiliary lemma by Csis ar and Narayan [17, Lemma A2]. Fix  $P_U, P_{X_k|U}$ , for  $k = 1, 2$ , as in Lemma 10.

*Lemma 14.* For every pair of conditional state distributions  $Q(s|x_1, u)$  and  $Q'(s|x_1, u)$  such that

$$\max \left\{ \sum_{u, x_1, s} P_U(u) P_{X_1|U}(x_1|u) Q(s|x_1, u) l(s), \sum_{u, x_1, s} P_U(u) P_{X_1|U}(x_1|u) Q'(s|x_1, u) l(s) \right\} < \tilde{\Lambda}_1(P_{U, X_1}), \quad (122)$$

there exists  $\xi > 0$  such that

$$\max_{u, x_1, \tilde{x}_1, x_2, y} \left| \sum_s Q(s|\tilde{x}_1, u) W_{Y|X_1, X_2, S}(y|x_1, x_2, s) - \sum_s Q'(s|x_1, u) W_{Y|X_1, X_2, S}(y|\tilde{x}_1, x_2, s) \right| \geq \xi. \quad (123)$$

*Proof of Lemma 14.* Assume to the contrary that the LHS in (123) is zero, and define

$$Q_A(s|x_1, u) = \frac{1}{2} (Q(s|x_1, u) + Q'(s|x_1, u)). \quad (124)$$

By symmetry,

$$\begin{aligned} 0 &= \max_{u, x_1, \tilde{x}_1, x_2, y} \left| \sum_s Q(s|\tilde{x}_1, u) W_{Y|X_1, X_2, S}(y|x_1, x_2, s) - \sum_s Q'(s|x_1, u) W_{Y|X_1, X_2, S}(y|\tilde{x}_1, x_2, s) \right| \\ &= \frac{1}{2} \max_{u, x_1, \tilde{x}_1, x_2, y} \left| \sum_s Q(s|\tilde{x}_1, u) W_{Y|X_1, X_2, S}(y|x_1, x_2, s) - \sum_s Q'(s|x_1, u) W_{Y|X_1, X_2, S}(y|\tilde{x}_1, x_2, s) \right| \\ &\quad + \frac{1}{2} \max_{u, x_1, \tilde{x}_1, x_2, y} \left| \sum_s Q'(s|\tilde{x}_1, u) W_{Y|X_1, X_2, S}(y|x_1, x_2, s) - \sum_s Q(s|x_1, u) W_{Y|X_1, X_2, S}(y|\tilde{x}_1, x_2, s) \right| \\ &\geq \max_{u, x_1, \tilde{x}_1, x_2, y} \left| \sum_s Q_A(s|x_1, u) W_{Y|X_1, X_2, S}(y|\tilde{x}_1, x_2, s) - \sum_s Q_A(s|\tilde{x}_1, u) W_{Y|X_1, X_2, S}(y|x_1, x_2, s) \right|, \end{aligned} \quad (125)$$

where the last line follows from the triangle inequality. Then, it follows that

$$\sum_{s \in \mathcal{S}} Q_A(s|x_1, u) W_{Y|X_1, X_2, S}(y|\tilde{x}_1, x_2, s) = \sum_{s \in \mathcal{S}} Q_A(s|\tilde{x}_1, u) W_{Y|X_1, X_2, S}(y|x_1, x_2, s), \quad (126)$$

for all  $u \in \mathcal{U}$ ,  $x_1, \tilde{x}_1 \in \mathcal{X}_1$ ,  $x_2 \in \mathcal{X}_2$ , and  $y \in \mathcal{Y}$ . In other words,  $J_{1,u} \equiv Q_A(\cdot|\cdot, u)$  symmetrizes  $\mathcal{X}_1|\mathcal{X}_2$  the AVMAC, for all  $u \in \mathcal{U}$ .

Next, recall from Remark 8 that the minimal state cost in (32b) can be written as

$$\tilde{\Lambda}_1(P_{U, X_1}) = \min_{\text{symm. } \{J_{1,u}\}} \sum_{u \in \mathcal{U}} \sum_{x_1 \in \mathcal{X}_1} \sum_{s \in \mathcal{S}} P_U(u) P_{X_1|U}(x_1|u) J_{1,u}(s|x_1) l(s), \quad (127)$$

where the minimization is over the set of distributions  $\{J_{1,u}\}_{u \in \mathcal{U}}$ , such that each  $J_{1,u}(s|x_1)$  satisfies (14), for  $u \in \mathcal{U}$ . Nevertheless, by (122),

$$\sum_{u, x_1, s} P_U(u) P_{X_1|U}(x_1|u) Q_A(s|x_1, u) l(s) < \tilde{\Lambda}_1(P_{U, X_1}). \quad (128)$$

This is a contradiction, since we have seen above that the distributions  $J_{1,u} \equiv Q_A(\cdot|\cdot, u)$  symmetrize  $\mathcal{X}_1|\mathcal{X}_2$  the AVMAC, for all  $u \in \mathcal{U}$ . It follows that the LHS of (123) must be positive. This completes the proof of the auxiliary Lemma.  $\square$

We move to the main part of the proof. Notice that while the parameters  $\eta$ ,  $\eta_1$  and  $\eta_2$  can be chosen freely, the rest of the parameters,  $\xi$ ,  $\delta_0$ ,  $\delta_1$ , and  $\delta_2$ , depend on  $P_U$  and  $P_{X_k|U}$ . To show that (50) holds for sufficiently small  $\eta$  and  $\eta_1$ , assume to the contrary that there exists  $y^n$  in  $\mathcal{D}(m_1, m_2) \cap \mathcal{D}(\check{m}_1, m_2) \neq \emptyset$  for some  $\check{m}_1 \neq m_1$ . By the assumption in the lemma, the codewords  $\{f_k(m_k)\}_{m_k \in [1:2^{nR_k}]}$  in Codebook  $k$  have the same conditional type, given the time sharing sequence  $u^n$ , for  $k = 1, 2$ . In particular,  $P_{\check{X}_1|U} = P_{X_1|U} = P_1$  and  $P_{\check{X}_2|U} = P_{X_2|U} = P_2$ .

By Condition 1) of the decoding rule for  $\mathcal{D}(m_1, m_2)$ ,

$$\begin{aligned} &D(P_{U, X_1, X_2, S, Y} || P_U \times P_{X_1|U} \times P_{X_2|U} \times P_{S|U} \times W_{Y|X_1, X_2, S}) \\ &= \sum_{u, x_1, x_2, s, y} P_{U, X_1, X_2, S, Y}(u, x_1, x_2, s, y) \cdot \log \frac{P_{U, X_1, X_2, S, Y}(u, x_1, x_2, s, y)}{P_U(u) P_1(x_1|u) P_2(x_2|u) P_{S|U}(s|u) W_{Y|X_1, X_2, S}(y|x_1, x_2, s)} \leq \eta, \end{aligned} \quad (129)$$

and by Condition 2b) of the decoding rule for  $\mathcal{D}(m_1, m_2)$ ,

$$I(X_1, X_2, Y; \tilde{X}_1|U, S) = \sum_{u, x_1, \tilde{x}_1, x_2, s, y} P_{U, X_1, \tilde{X}_1, X_2, S, Y}(u, x_1, \tilde{x}_1, x_2, s, y) \cdot \log \frac{P_{\tilde{X}_1|U, X_1, X_2, S, Y}(\tilde{x}_1|u, x_1, x_2, s, y)}{P_{\tilde{X}_1|U, S}(\tilde{x}_1|u, s)} \leq \eta_1, \quad (130)$$

where  $U, X_1, \tilde{X}_1, X_2, S, Y$  are distributed according to the joint type of  $u^n, f_1(m_1), f_1(\tilde{m}_1), f_2(m_2), s^n$ , and  $y^n$ , for some  $s^n \in \mathcal{S}^n$  with  $l^n(s^n) \leq \Lambda$ . Adding (129) and (130) yields

$$\sum_{u, x_1, \tilde{x}_1, x_2, s, y} P_{U, X_1, \tilde{X}_1, X_2, S, Y}(u, x_1, \tilde{x}_1, x_2, s, y) \cdot \log \frac{P_{U, X_1, \tilde{X}_1, X_2, S, Y}(u, x_1, \tilde{x}_1, x_2, s, y)}{P_U(u)P_1(x_1|u)P_{\tilde{X}_1, S|U}(\tilde{x}_1, s|u)P_2(x_2|u)W_{Y|X_1, X_2, S}(y|x_1, x_2, s)} \leq \eta + \eta_1. \quad (131)$$

That is,  $D(P_{U, X_1, \tilde{X}_1, X_2, S, Y} || P_U \times P_1 \times P_1 \times P_{S|U, \tilde{X}_1} \times P_2 \times W_{Y|X_1, X_2, S}) \leq \eta + \eta_1$ . Therefore, by the log-sum inequality (see e.g. [14, Theorem 2.7.1]),

$$\begin{aligned} & D(P_{U, X_1, \tilde{X}_1, X_2, Y} || P_U \times P_1 \times P_1 \times P_2 \times V_{Y|U, X_1, \tilde{X}_1, X_2}) \\ & \leq D(P_{U, X_1, \tilde{X}_1, X_2, S, Y} || P_U \times P_1 \times P_1 \times P_{S|U, \tilde{X}_1} \times P_2 \times W_{Y|X_1, X_2, S}) \leq \eta + \eta_1, \end{aligned} \quad (132)$$

where  $V_{Y|U, X_1, \tilde{X}_1, X_2}(y|u, x_1, \tilde{x}_1, x_2) = \sum_{s \in \mathcal{S}} W_{Y|X_1, X_2, S}(y|x_1, x_2, s)P_{S|U, \tilde{X}_1}(s|u, \tilde{x}_1)$ . Then, by Pinsker's inequality (see e.g. [16, Problem 3.18]),

$$\sum_{u, x_1, \tilde{x}_1, x_2, y} |P_{U, X_1, \tilde{X}_1, X_2, Y}(u, x_1, \tilde{x}_1, x_2, y) - P_U(u)P_1(x_1|u)P_1(\tilde{x}_1|u)P_2(x_2|u)V_{Y|U, X_1, \tilde{X}_1, X_2}(y|u, x_1, \tilde{x}_1, x_2)| \leq c\sqrt{\eta + \eta_1}, \quad (133)$$

where  $c > 0$  is a constant.

Similarly, following our assumption that the decoding rules for  $\mathcal{D}(\tilde{m}_1, m_2)$  hold as well, Conditions 1) and 2b) claim that

$$D(P_{U, \tilde{X}_1, X_2, S, Y} || P_U \times P_{\tilde{X}_1|U} \times P_{X_2|U} \times P_{\tilde{S}|U} \times W_{Y|X_1, X_2, S}) \leq \eta, \quad (134)$$

and

$$I(\tilde{X}_1, X_2, Y; X_1|U, \tilde{S}) \leq \eta_1, \quad (135)$$

where  $U, X_1, \tilde{X}_1, X_2, \tilde{S}, Y$  are distributed according to the joint type of  $u^n, f_1(m_1), f_1(\tilde{m}_1), f_2(m_2), \tilde{s}^n$ , and  $y^n$ , for some  $\tilde{s}^n \in \mathcal{S}^n$  with  $l^n(\tilde{s}^n) \leq \Lambda$ . Here,  $\tilde{X}_1$  and  $\tilde{S}$  have switched places with  $X$  and  $S$ , respectively, since  $f_1(\tilde{m}_1)$  and  $\tilde{s}^n$  are the tested codeword and state sequence, while  $f_1(m_1)$  and  $s^n$  are the competing ones. By the same arguments that led to (133), it follows that

$$\sum_{u, x_1, \tilde{x}_1, x_2, y} |P_{U, X_1, \tilde{X}_1, X_2, Y}(u, x_1, \tilde{x}_1, x_2, s) - P_U(u)P_1(x_1|u)P_1(\tilde{x}_1|u)P_2(x_2|u)V'_{Y|U, X_1, \tilde{X}_1, X_2}(y|u, x_1, \tilde{x}_1, x_2)| \leq c\sqrt{\eta + \eta_1}, \quad (136)$$

where  $V'_{Y|U, X_1, \tilde{X}_1, X_2}(y|u, x_1, \tilde{x}_1, x_2) = \sum_{s \in \mathcal{S}} W_{Y|X_1, X_2, S}(y|\tilde{x}_1, x_2, s)P_{\tilde{S}|U, X_1}(s|u, x_1)$ . Since  $P_U(u) \geq \delta_0$  and  $P_{X_k|U}(x_k|u) \geq \delta_k$ , for all  $u \in \mathcal{U}$  and  $x_k \in \mathcal{X}_k$ , for  $k = 1, 2$ , we have by (133) and (136) that

$$\max_{u, x_1, \tilde{x}_1, x_2, y} \left| V_{Y|U, X_1, \tilde{X}_1, X_2}(y|u, x_1, \tilde{x}_1, x_2) - V'_{Y|U, X_1, \tilde{X}_1, X_2}(y|u, x_1, \tilde{x}_1, x_2) \right| \leq \frac{2c\sqrt{\eta + \eta_1}}{\delta_0 \delta_1^2 \delta_2}, \quad (137)$$

Equivalently, the above can be expressed as

$$\begin{aligned} & \max_{u, x_1, \tilde{x}_1, x_2, y} \left| \sum_s P_{S|U, \tilde{X}_1}(s|u, \tilde{x}_1)W_{Y|X_1, X_2, S}(y|x_1, x_2) - \sum_s P_{\tilde{S}|U, X_1}(s|u, x_1)W_{Y|X_1, X_2, S}(y|\tilde{x}_1, x_2, s) \right| \\ & \leq \frac{2c\sqrt{\eta + \eta_1}}{\delta_0 \delta_1^2 \delta_2}, \end{aligned} \quad (138)$$

Now, we show that the state distributions  $Q = P_{S|U, \tilde{X}_1}$  and  $Q' = P_{\tilde{S}|U, X_1}$  satisfy the conditions of Lemma 14. Indeed,

$$\begin{aligned}
& \max \left\{ \sum_{u, \tilde{x}_1, s} P_U(u) P_1(\tilde{x}_1|u) Q(s|u, \tilde{x}_1) l(s), \sum_{u, x_1, s} P_U(u) P_1(x_1|u) Q'(s|u, x_1) l(s) \right\} \\
&= \max \left\{ \sum_{u, \tilde{x}_1, s} P_U(u) P_1(\tilde{x}_1|u) P_{S|U, \tilde{X}_1}(s|u, \tilde{x}_1) l(s), \sum_{u, x_1, s} P_U(u) P_1(x_1|u) P_{\tilde{S}|U, X_1}(s|u, x_1) l(s) \right\} \\
&= \max \left\{ \sum_s P_S(s) l(s), \sum_s P_{\tilde{S}}(s) l(s) \right\} \\
&= \max \{ l^n(s^n), l^n(\tilde{s}^n) \} \leq \Lambda < \tilde{\Lambda}_1(P_{U, X_1}), \tag{139}
\end{aligned}$$

where the last inequality is due to (47). Thus, there exists  $\xi > 0$  such that (123) holds with  $Q = P_{S|U, \tilde{X}_1}$  and  $Q' = P_{\tilde{S}|U, X_1}$ , which contradicts (138), if we choose  $\eta$  and  $\eta_1$  to be sufficiently small such that  $\frac{2c\sqrt{\eta+\eta_1}}{\delta_0\delta_1^2\delta_2} < \xi$ .  $\square$

#### APPENDIX D PROOF OF LEMMA 11

Fix a sequence  $u^n \in \mathcal{U}^n$  of type  $P_U$ . Let  $\bar{Z}^n(m) = (Z_1^n(m_1), Z_2^n(m_2))$ ,  $m \in [1 : 2^{n(R_1+R_2)}]$ , be independent sequence pairs, uniformly distributed over the conditional type classes  $\mathcal{T}^n(P_1)$  and  $\mathcal{T}^n(P_2)$ , where we have assigned an index  $m \in [1 : 2^{n(R_1+R_2)}]$  to each message pair  $(m_1, m_2) \in [1 : 2^{nR_1}] \times [1 : 2^{nR_2}]$ . Fix  $a_1^n \in \mathcal{X}_1^n$ ,  $a_2^n \in \mathcal{X}_2^n$ , and  $s^n \in \mathcal{S}^n$ , and consider a joint type  $P_{U, X_1, X_2, \tilde{X}_1, \tilde{X}_2, S}$ , such that  $P_{X_1, X_2|U} = P_{\tilde{X}_1, \tilde{X}_2|U} = P_1 \times P_2$ , i.e.

$$P_{X_1, X_2|U}(x_1, x_2|u) = P_{\tilde{X}_1, \tilde{X}_2|U}(x_1, x_2|u) = P_1(x_1|u) \cdot P_2(x_2|u). \tag{140}$$

We intend to show that  $\{Z_1^n(m_1)\}$  and  $\{Z_2^n(m_2)\}$  satisfy each of the desired properties with double exponential high probability  $(1 - e^{-2^{Bn}})$ ,  $B > 0$ , implying that there exist deterministic codebooks that satisfy (52)-(58) simultaneously. This will only be shown for the properties in parts 1 and 2, since part 3 is symmetric with part 2.

We will use the following large deviations result by Csiszár and Narayan [17].

*Lemma 15* (see [17, Lemma A1]). Let  $\alpha, \beta \in [0, 1]$ , and consider a sequence of random vectors  $V^n(m)$ , and functions  $\varphi_m : \mathcal{X}^{nm} \rightarrow [0, 1]$ , for  $m \in [1 : M]$ . If

$$\mathbb{E}(\varphi_m(V^n(1) \dots, V^n(m)) | V^n(1) \dots, V^n(m-1)) \leq \alpha \text{ a.s., for } m \in [1 : M], \tag{141}$$

then

$$\Pr \left( \sum_{m=1}^M \varphi_m(V^n(1) \dots, V^n(m)) > M\beta \right) \leq \exp\{-M(\beta - \alpha \log e)\}. \tag{142}$$

*Part 1*

To show that (52) holds, consider the indicator function

$$\varphi_m(\bar{Z}^n(1), \dots, \bar{Z}^n(m)) = \begin{cases} 1 & \text{if } (u^n, a_1^n, a_2^n, Z_1^n(m_1), Z_2^n(m_2), s^n) \in \mathcal{T}^n(P_{U, X_1, X_2, \tilde{X}_1, \tilde{X}_2, S}) \\ 0 & \text{otherwise} \end{cases} \tag{143}$$

where  $\bar{Z}^n(m) = (Z_1^n(m_1), Z_2^n(m_2))$  as defined above. By standard type class considerations (see e.g. [35, Theorem 1.3]), we have that

$$\mathbb{E}[\varphi_m(\bar{Z}^n(1), \dots, \bar{Z}^n(m)) | \bar{Z}^n(1), \dots, \bar{Z}^n(m-1)] \leq 2^{-n(I(\tilde{X}_1, \tilde{X}_2; U, X_1, X_2, S) - \frac{\epsilon}{4})} \tag{144}$$

$$\leq 2^{-n(I(\tilde{X}_1, \tilde{X}_2; X_1, X_2, S|U) - \frac{\epsilon}{4})}, \tag{145}$$

where the last inequality holds since  $I(\tilde{X}_1, \tilde{X}_2; U, X_1, X_2, S) \geq I(\tilde{X}_1, \tilde{X}_2; X_1, X_2, S|U)$ .

Next, we use Lemma 15, and plug

$$\begin{aligned}
(V(1), \dots, V(M)) &\leftarrow (\bar{Z}^n(1), \dots, \bar{Z}^n(2^{n(R_1+R_2)})), \quad M = 2^{n(R_1+R_2)}, \\
\alpha &= 2^{-n(I(\tilde{X}_1, \tilde{X}_2; X_1, X_2, S|U) - \frac{\epsilon}{4})}, \\
\beta &= 2^{n([R_1+R_2 - I(\tilde{X}_1, \tilde{X}_2; X_1, X_2, S|U)]_+ - (R_1+R_2) + \epsilon)}. \tag{146}
\end{aligned}$$

For sufficiently large  $n$ , we have that  $M(\beta - \alpha \log e) \geq 2^{n\epsilon/2}$ . Hence, by Lemma 15,

$$\Pr \left( \sum_{m=1}^{2^{n(R_1+R_2)}} \varphi_m(\bar{Z}^n(1), \dots, \bar{Z}^n(2^{n(R_1+R_2)})) > 2^n \left( [R_1+R_2 - I(\tilde{X}_1, \tilde{X}_2; X_1, X_2, S|U)]_+ + \epsilon \right) \right) \leq e^{-2^{n\epsilon/2}}. \quad (147)$$

The double exponential decay of the probability above implies that there exist codebooks which satisfy (52).

Similarly, to show (53), we replace the indicator of the type  $P_{U, X_1, X_2, \tilde{X}_1, \tilde{X}_2, S}$  in (143) by an indicator of the type  $P_{U, \tilde{X}_1, \tilde{X}_2, S}$ , and rewrite (144) with  $I(\tilde{X}_1, \tilde{X}_2; S|U)$ , to obtain

$$\Pr \left( |\{(\tilde{m}_1, \tilde{m}_2) : (u^n, Z_1^n(\tilde{m}_1), Z_2^n(\tilde{m}_2), s^n) \in \mathcal{T}^n(P_{U, \tilde{X}_1, \tilde{X}_2, S})\}| > 2^n \left( [R_1+R_2 - I(\tilde{X}_1, \tilde{X}_2; S|U)]_+ + \epsilon_1 \right) \right) < e^{-2^{n\epsilon_1/2}}, \quad (148)$$

where  $\epsilon_1 > 0$  is arbitrarily small. If  $I(\tilde{X}_1, \tilde{X}_2; S|U) > \epsilon$  and  $R_1 + R_2 \geq \epsilon$ , then choosing  $\epsilon_1 = \frac{\epsilon}{2}$ , we have that

$$\left[ R_1 + R_2 - I(\tilde{X}_1, \tilde{X}_2; S|U) \right]_+ + \epsilon_1 \leq R_1 + R_2 - \frac{\epsilon}{2}, \quad (149)$$

hence,

$$\Pr \left( |\{(\tilde{m}_1, \tilde{m}_2) : (u^n, Z_1^n(\tilde{m}_1), Z_2^n(\tilde{m}_2), s^n) \in \mathcal{T}^n(P_{U, \tilde{X}_1, \tilde{X}_2, S})\}| > 2^{n(R_1+R_2-\frac{\epsilon}{2})} \right) < e^{-2^{n\epsilon/4}}. \quad (150)$$

It remains to show that (54) holds. Assume that

$$I(X_1, X_2; \tilde{X}_1, \tilde{X}_2, S|U) - \left[ R_1 + R_2 - I(\tilde{X}_1, \tilde{X}_2; S|U) \right]_+ > \epsilon. \quad (151)$$

Let  $\mathcal{A}_m$  denote the set of indices  $\tilde{m} < m$  such that  $(u^n, \bar{Z}^n(\tilde{m}), s^n) \in \mathcal{T}^n(P_{U, \tilde{X}_1, \tilde{X}_2, S})$ , provided that their number does not exceed  $2^n \left( [R_1+R_2 - I(\tilde{X}_1, \tilde{X}_2; S|U)]_+ + \frac{\epsilon}{8} \right)$ ; else, let  $\mathcal{A}_m = \emptyset$ . Also, let

$$\psi_m(\bar{Z}^n(1), \dots, \bar{Z}^n(m)) = \begin{cases} 1 & \text{if } (u^n, \bar{Z}^n(m), \bar{Z}^n(\tilde{m}), s^n) \in \mathcal{T}^n(P_{U, X_1, X_2, \tilde{X}_1, \tilde{X}_2, S}) \\ & \text{for some } \tilde{m} \in \mathcal{A}_m, \\ 0 & \text{otherwise.} \end{cases} \quad (152)$$

Then, choosing  $\epsilon_1 = \frac{\epsilon}{8}$  in (148) yields

$$\Pr \left( \sum_{m=1}^{2^{n(R_1+R_2)}} \psi_m(\bar{Z}^n(1), \dots, \bar{Z}^n(m)) \neq |\{m : (u^n, \bar{Z}^n(m), \bar{Z}^n(\tilde{m}), s^n) \in \mathcal{T}^n(P_{U, X_1, X_2, \tilde{X}_1, \tilde{X}_2, S}) \text{ for some } \tilde{m} < m\}| \right) < e^{-2^{n\epsilon/16}}. \quad (153)$$

Therefore, instead of bounding the set of message pairs, it is sufficient to consider the sum  $\sum \psi_m(\bar{Z}^n(1), \dots, \bar{Z}^n(m))$ . Furthermore, by standard type class considerations (see e.g. [35, Theorem 1.3]), we have that

$$\begin{aligned} \mathbb{E} \left( \psi_m(\bar{Z}^n(1), \dots, \bar{Z}^n(m)) \middle| \bar{Z}^n(1), \dots, \bar{Z}^n(m-1) \right) &\leq |\mathcal{A}_m| \cdot 2^{-n(I(X_1, X_2; \tilde{X}_1, \tilde{X}_2, S|U) - \frac{\epsilon}{8})} \\ &\leq 2^n \left( [R_1+R_2 - I(\tilde{X}_1, \tilde{X}_2; S|U)]_+ - I(X_1, X_2; \tilde{X}_1, \tilde{X}_2, S|U) + \frac{\epsilon}{4} \right) < 2^{-3n\epsilon/4}, \end{aligned} \quad (154)$$

where the last inequality is due to (151). Thus, by Lemma 15,

$$\Pr \left( \sum_{m=1}^{2^{n(R_1+R_2)}} \psi_m(\bar{Z}^n(1), \dots, \bar{Z}^n(m)) > 2^{n(R_1+R_2-\frac{\epsilon}{2})} \right) < e^{-2^{n(R_1+R_2-\frac{3\epsilon}{4})}} \leq e^{-2^{n\epsilon/4}}, \quad (155)$$

as we have assumed that  $R_1 + R_2 \geq \epsilon$ . Equations (153) and (155) imply that the property in (54) holds with double exponential probability  $1 - e^{-2^{E_1 n}}$ , where  $E_1 > 0$ .

## Part 2

Fix  $m_2 \in [1 : 2^{nR_2}]$  and  $Z_2^n(m_2) = z_2^n \in \mathcal{T}^n(P_2)$ . To show that (55) holds, consider the indicator

$$\varphi_{m_1}(Z_1^n(1), \dots, Z_1^n(m_1)) = \begin{cases} 1 & \text{if } (u^n, a_1^n, z_2^n, Z_1^n(m_1), s^n) \in \mathcal{T}^n(P_{U, X_1, X_2, \tilde{X}_1, S}) \\ 0 & \text{otherwise} \end{cases} \quad (156)$$

By standard type class considerations (see e.g. [35, Theorem 1.3]), we have that

$$\mathbb{E} \left[ \varphi_{m_1}(Z_1^n(1), \dots, Z_1^n(m_1)) \middle| Z_1^n(1), \dots, Z_1^n(m_1-1) \right] \leq 2^{-n(I(\tilde{X}_1; X_1, X_2, S|U) - \frac{\epsilon}{4})}. \quad (157)$$

Next, we use Lemma 15, and plug

$$\begin{aligned} (V(1), \dots, V(M)) &\leftarrow (Z_1^n(1), \dots, Z_1^n(2^{nR_1})), \quad M = 2^{nR_1}, \\ \alpha &= 2^{-n(I(\tilde{X}_1; X_1, X_2, S|U) - \frac{\varepsilon}{4})}, \\ \beta &= 2^n([R_1 - I(\tilde{X}_1; X_1, X_2, S|U)]_+ - R_1 + \varepsilon). \end{aligned} \quad (158)$$

For sufficiently large  $n$ , we have that  $M(\beta - \alpha \log e) \geq 2^{n\varepsilon/2}$ . Hence, by Lemma 15,

$$\Pr \left( \sum_{m_1=1}^{2^{nR_1}} \varphi_{m_1}(Z_1^n(1), \dots, Z_1^n(2^{nR_1})) > 2^n([R_1 - I(\tilde{X}_1; X_1, X_2, S|U)]_+ + \varepsilon) \right) \leq e^{-2^{n\varepsilon/2}}. \quad (159)$$

By the symmetry between  $m_1$  and  $\tilde{m}_1$  in the derivation above, the double exponential decay of the probability in (159) implies that there exist codebooks which satisfy (55).

Next, we show that (56) holds. Replacing the indicator of the type  $P_{X_1, X_2, \tilde{X}_1, S}$  in (156) with an indicator of the type  $P_{\tilde{X}_1, S}$  yields

$$\Pr \left( |\{\tilde{m}_1 : (u^n, x_1^n(\tilde{m}_1), s^n) \in \mathcal{T}^n(P_{U, \tilde{X}_1, S})\}| > 2^n([R_1 - I(\tilde{X}_1; S|U)]_+ + \varepsilon_2) \right) < e^{-2^{n\varepsilon_2/2}}, \quad (160)$$

where  $\varepsilon_2 > 0$  is arbitrarily small. Assume that

$$I(X_1, X_2; \tilde{X}_1, S|U) - [R_1 - I(\tilde{X}_1; S|U)]_+ > \varepsilon. \quad (161)$$

Let  $\mathcal{A}_{m_1}$  denote the set of indices  $\tilde{m}_1 < m_1$  such that  $(Z_1^n(\tilde{m}_1), s^n) \in \mathcal{T}^n(P_{\tilde{X}_1, S})$ , provided that their number does not exceed  $2^n([R_1 - I(\tilde{X}_1; S)]_+ + \frac{\varepsilon}{8})$ ; else, let  $\mathcal{A}_{m_1} = \emptyset$ . Also, let

$$\psi_{m_1}(Z_1^n(1), \dots, Z_1^n(m_1)) = \begin{cases} 1 & \text{if } (u^n, Z_1^n(m_1), z_2^n, Z_1^n(\tilde{m}_1), s^n) \in \mathcal{T}^n(P_{U, X_1, X_2, \tilde{X}_1, S}), \text{ for some } \tilde{m}_1 \in \mathcal{A}_{m_1}, \\ 0 & \text{otherwise.} \end{cases} \quad (162)$$

Then, choosing  $\varepsilon_2 = \frac{\varepsilon}{8}$  in (160) yields

$$\Pr \left( \sum_{m_1=1}^{2^{nR_1}} \psi_{m_1}(Z_1^n(1), \dots, Z_1^n(m_1)) \neq |\{m_1 : (u^n, Z_1^n(m_1), z_2^n, Z_1^n(\tilde{m}_1), s^n) \in \mathcal{T}^n(P_{U, X_1, X_2, \tilde{X}_1, S}) \text{ for some } \tilde{m}_1 < m_1\}| \right) < e^{-2^{n\varepsilon/16}}. \quad (163)$$

Therefore, instead of bounding the set of messages, it is sufficient to consider the sum  $\sum \psi_{m_1}(Z_1^n(1), \dots, Z_1^n(m_1))$ . Furthermore, by standard type class considerations (see e.g. [35, Theorem 1.3]), we have that

$$\begin{aligned} \mathbb{E} \left( \psi_{m_1}(Z_1^n(1), \dots, Z_1^n(m_1)) \middle| Z_1^n(1), \dots, Z_1^n(m_1 - 1) \right) &\leq |\mathcal{A}_{m_1}| \cdot 2^{-n(I(X_1, X_2; \tilde{X}_1, S|U) + \frac{\varepsilon}{8})} \\ &\leq 2^n([R_1 - I(\tilde{X}_1; S|U)]_+ - I(X_1, X_2; \tilde{X}_1, S|U) + \frac{\varepsilon}{4}) < 2^{-3n\varepsilon/4}, \end{aligned} \quad (164)$$

where the last inequality is due to (161). Thus, by Lemma 15,

$$\Pr \left( \sum_{m_1=1}^{2^{nR_1}} \psi_{m_1}(Z_1^n(1), \dots, Z_1^n(m_1)) > 2^{n(R_1 - \frac{\varepsilon}{2})} \right) < e^{-2^{n(R_1 - \frac{3\varepsilon}{4})}} \leq e^{-2^{n\varepsilon/4}}, \quad (165)$$

as we have assumed that  $R_1 + R_2 \geq \varepsilon$ . The double exponential bounds in (163) and (165) imply that there exists codebooks that satisfy (56) as well.  $\square$

## APPENDIX E PROOF OF THEOREM 8

To prove the theorem, we consider each case in the definition of the region  $\mathcal{C}(\mathcal{A})$  separately (see Definition 6). Case A requires most of the effort, as the other cases follow from similar, yet simpler, considerations.

*Case A*

Suppose that  $L^* > \Lambda$ ,  $L_1^* > \Lambda$  and  $L_2^* > \Lambda$ .

*Achievability Proof:* Let  $\varepsilon > 0$  be chosen later, and  $u^n \in \mathcal{U}^n$  be a sequence in the type class of  $P_U$ , such that  $P_U(u) > 0 \forall u \in \mathcal{U}$ . For  $k = 1, 2$ , let  $P_{X_k|U}$  be a conditional type over  $\mathcal{X}_k$ , for which  $P_{X_k|U}(x_k|u) > 0 \forall x_k \in \mathcal{X}_k, u \in \mathcal{U}, \mathbb{E}\phi_k(X_k) \leq \Omega_k$ , with

$$\tilde{\Lambda}_k(P_{U, X_k}) > \Lambda, \quad (166)$$

and

$$\tilde{\Lambda}(P_{U, X_1, X_2}) > \Lambda. \quad (167)$$

Furthermore, choose  $\eta, \eta_1, \eta_2 > 0$  accordingly to be sufficiently small, such that Lemma 10 guarantees that the decoder in Definition 7 is well defined. Now, Lemma 11 assures that there are codebooks,  $\{x_1^n(m_1) : m_1 \in [1 : 2^{nR_1}]\}$  of type  $P_{X_1|U}$ , and  $\{x_2^n(m_2) : m_2 \in [1 : 2^{nR_2}]\}$  of type  $P_{X_2|U}$ , which satisfy (52)-(58). Consider the following coding scheme.

*Encoding:* To send  $m_k \in [1 : 2^{nR_k}]$ , Encoder  $k$  transmits  $x_k^n(m_k)$ , for  $k = 1, 2$ .

*Decoding:* Find a unique message pair  $(\hat{m}_1, \hat{m}_2)$  such that the received sequence  $y^n$  belongs to  $\mathcal{D}(\hat{m}_1, \hat{m}_2)$ , as in Definition 7. If there is none, declare an error. Lemma 10 guarantees that there cannot be two message pairs for which this holds.

*Analysis of Probability of Error:* Fix  $s^n \in \mathcal{S}^n$  with  $l^n(s^n) \leq \Lambda$ , let  $q = P_{S|U}$  denote the conditional type of  $s^n$  given  $u^n$ , and let  $(M_1, M_2)$  denote the transmitted message pair. Consider the error events

$$\mathcal{E}_1 = \{D(P_{U, X_1, X_2, S, Y} || P_U \times P_{X_1|U} \times P_{X_2|U} \times P_{S|U} \times W_{Y|X_1, X_2, S}) > \eta\} \quad (168)$$

$$\mathcal{E}_{2a} = \{\text{Condition 2a) of the decoding rule is violated}\} \quad (169)$$

$$\mathcal{E}_{2b} = \{\text{Condition 2b) of the decoding rule is violated}\} \quad (170)$$

$$\mathcal{E}_{2c} = \{\text{Condition 2c) of the decoding rule is violated}\} \quad (171)$$

and

$$\mathcal{F}_1 = \{I_q(U, X_1, X_2; S) > \varepsilon\}, \quad (172)$$

$$\mathcal{F}_2 = \{I_q(X_1, X_2; \tilde{X}_1, \tilde{X}_2, S|U) > [R_1 + R_2 - I(\tilde{X}_1, \tilde{X}_2; S|U)]_+ + \varepsilon, \text{ for some } \tilde{m}_1 \neq M_1 \text{ and } \tilde{m}_2 \neq M_2\}, \quad (173)$$

$$\mathcal{F}_3 = \{I_q(X_1, X_2; \tilde{X}_1, S|U) > [R_1 - I(\tilde{X}_1; S|U)]_+ + \varepsilon, \text{ for some } \tilde{m}_1 \neq M_1\}, \quad (174)$$

$$\mathcal{F}_4 = \{I_q(X_1, X_2; \tilde{X}_2, S|U) > [R_2 - I(\tilde{X}_2; S|U)]_+ + \varepsilon, \text{ for some } \tilde{m}_2 \neq M_2\}, \quad (175)$$

where  $(U, X_1, X_2, \tilde{X}_1, \tilde{X}_2, S)$  are dummy random variables, which are distributed as the joint type of  $(x_1^n(M_1), x_2^n(M_2), x_1^n(\tilde{m}_1), x_2^n(\tilde{m}_2), s^n)$ . By the union of events bound,

$$\begin{aligned} P_{e|s^n}^{(n)}(\mathcal{C}) &\leq \Pr(\mathcal{F}_1) + \Pr(\mathcal{F}_2) + \Pr(\mathcal{F}_3) + \Pr(\mathcal{F}_4) \\ &\quad + \Pr(\mathcal{E}_1 \cap \mathcal{F}_1^c) + \Pr(\mathcal{E}_{2a} \cap \mathcal{E}_1^c \cap \mathcal{F}_2^c) + \Pr(\mathcal{E}_{2b} \cap \mathcal{E}_1^c \cap \mathcal{F}_3^c) + \Pr(\mathcal{E}_{2c} \cap \mathcal{E}_1^c \cap \mathcal{F}_4^c), \end{aligned} \quad (176)$$

where the conditioning on  $s^n$  is omitted for convenience of notation. Based on Lemma 11, the probabilities of the events  $\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3$ , and  $\mathcal{F}_4$ , tend to zero as  $n \rightarrow \infty$ , by (53), (54), (56), and (58), respectively.

Now, suppose that Condition 1) of the decoding rule is violated. Observe that the event  $\mathcal{E}_1 \cap \mathcal{F}_1^c$  implies that

$$\begin{aligned} &D(P_{U, X_1, X_2, S, Y} || P_{U, X_1, X_2, S} \times W_{Y|X_1, X_2, S}) \\ &= D(P_{U, X_1, X_2, S, Y} || P_U \times P_{X_1|U} \times P_{X_2|U} \times P_{S|U} \times W_{Y|X_1, X_2, S}) - I(X_1, X_2; S|U) > \eta - \varepsilon. \end{aligned} \quad (177)$$

Then, by standard large deviations considerations (see e.g. [14, pp. 362–364]),

$$\Pr(\mathcal{E}_1 \cap \mathcal{F}_1^c) \leq \max_{P_{U, X_1, X_2, S, Y} : \mathcal{E}_1 \cap \mathcal{F}_1^c \text{ holds}} 2^{-n(D(P_{U, X_1, X_2, S, Y} || P_{U, X_1, X_2, S} \times W_{Y|X_1, X_2, S}) - \varepsilon)} < 2^{-n(\eta - 2\varepsilon)}, \quad (178)$$

which tends to zero as  $n \rightarrow \infty$ , for sufficiently small  $\varepsilon > 0$ , with  $\varepsilon < \frac{1}{2}\eta$ .

Moving to Condition 2a) of the decoding rule, let  $\mathcal{D}_{2a}$  denote the set of joint types  $P_{U, X_1, X_2, \tilde{X}_1, \tilde{X}_2, S}$  such that

$$D(P_{U, X_1, X_2, S, Y} || P_U \times P_{X_1|U} \times P_{X_2|U} \times P_{S|U} \times W_{Y|X_1, X_2, S}) \leq \eta, \quad (179)$$

$$D(P_{U, \tilde{X}_1, \tilde{X}_2, \tilde{S}, Y} || P_U \times P_{\tilde{X}_1|U} \times P_{\tilde{X}_2|U} \times P_{\tilde{S}|U} \times W_{Y|X_1, X_2, S}) \leq \eta, \text{ for some } \tilde{S} \sim \tilde{q}(s|u), \quad (180)$$

$$I_q(X_1, X_2, Y; \tilde{X}_1, \tilde{X}_2|U, S) > \eta. \quad (181)$$



Observe that the event  $\mathcal{E}_1^c$  implies that (179) holds. Also, when the event  $\mathcal{E}_{2a}$  occurs, *i.e.* Condition 2a) of the decoding rule is violated, then there exist  $\tilde{m}_1 \neq m_1$  and  $\tilde{m}_2 \neq m_2$  such that (180) and (181) hold for some  $\tilde{s}^n \in \mathcal{S}^n$  with  $l^n(\tilde{s}^n) \leq \Lambda$ . Then, by standard type class considerations (see *e.g.* [35, Theorem 1.3]),

$$\begin{aligned} & \Pr(\mathcal{E}_{2a} \cap \mathcal{E}_1^c \cap \mathcal{F}_2^c \mid M_1 = m_1, M_2 = m_2) \\ & \leq \sum_{\substack{P_{U, X_1, X_2, \tilde{X}_1, \tilde{X}_2, S} \in \mathcal{D}_{2a}: \\ \mathcal{F}_2^c \text{ holds}}} |\{(\tilde{m}_1, \tilde{m}_2) : (u^n, x_1^n(m_1), x_2^n(m_2), x_1^n(\tilde{m}_1), x_2^n(\tilde{m}_2), s^n) \in \mathcal{T}^n(P_{U, X_1, X_2, \tilde{X}_1, \tilde{X}_2, S})\}| \\ & \quad \times 2^{-n(I_q(\tilde{X}_1, \tilde{X}_2; Y|U, X_1, X_2, S) - \varepsilon)}, \end{aligned} \quad (182)$$

for every given  $m_1 \in [1 : 2^{nR_1}]$  and  $m_2 \in [1 : 2^{nR_2}]$ . Hence, by (52),

$$\Pr(\mathcal{E}_{2a} \cap \mathcal{E}_1^c \cap \mathcal{F}_2^c) \leq \sum_{\substack{P_{U, X_1, X_2, \tilde{X}_1, \tilde{X}_2, S} \in \mathcal{D}_{2a}: \\ \mathcal{F}_2^c \text{ holds}}} 2^{-n(I_q(\tilde{X}_1, \tilde{X}_2; Y|U, X_1, X_2, S) - [R_1 + R_2 - I_q(\tilde{X}_1, \tilde{X}_2; X_1, X_2, S|U)]_+ - 2\varepsilon)}. \quad (183)$$

To further bound  $\Pr(\mathcal{E}_{2a} \cap \mathcal{E}_1^c \cap \mathcal{F}_2^c)$ , consider the following cases. Suppose that  $R_1 + R_2 \leq I_q(\tilde{X}_1, \tilde{X}_2; S|U)$ . Then, given  $\mathcal{F}_2^c$ , we have that

$$I_q(X_1, X_2; \tilde{X}_1, \tilde{X}_2|U, S) \leq I_q(X_1, X_2; \tilde{X}_1, \tilde{X}_2, S|U) \leq \varepsilon. \quad (184)$$

By (181), it then follows that

$$\begin{aligned} I_q(\tilde{X}_1, \tilde{X}_2; Y|U, X_1, X_2, S) &= I_q(\tilde{X}_1, \tilde{X}_2; X_1, X_2, Y|U, S) - I_q(\tilde{X}_1, \tilde{X}_2; X_1, X_2|U, S) \\ &\geq \eta - \varepsilon. \end{aligned} \quad (185)$$

Returning to (183), we note that since the number of types is polynomial in  $n$ , the cardinality of the set of types  $\mathcal{D}_{2a}$  can be bounded by  $2^{n\varepsilon}$ , for sufficiently large  $n$ . Hence, by (183) and (185), we have that  $\Pr(\mathcal{E}_{2a} \cap \mathcal{E}_1^c \cap \mathcal{F}_2^c) \leq 2^{-n(\eta - 4\varepsilon)}$ , which tends to zero as  $n \rightarrow \infty$ , for  $\varepsilon < \frac{1}{4}\eta$ .

Otherwise, if  $R_1 + R_2 > I_q(\tilde{X}_1, \tilde{X}_2; S|U)$ , then given  $\mathcal{F}_2^c$ ,

$$\begin{aligned} R_1 + R_2 &> I_q(X_1, X_2; \tilde{X}_1, \tilde{X}_2, S|U) + I(\tilde{X}_1, \tilde{X}_2; S|U) - \varepsilon \\ &= I_q(\tilde{X}_1, \tilde{X}_2; X_1, X_2, S|U) + I(X_1, X_2; S|U) - \varepsilon \\ &\geq I_q(\tilde{X}_1, \tilde{X}_2; X_1, X_2, S|U) - \varepsilon. \end{aligned} \quad (186)$$

Thus,

$$\left[ R_1 + R_2 - I_q(\tilde{X}_1, \tilde{X}_2; X_1, X_2, S|U) \right]_+ \leq R_1 + R_2 - I_q(\tilde{X}_1, \tilde{X}_2; X_1, X_2, S|U) + \varepsilon. \quad (187)$$

Hence, by (183) we have that

$$\begin{aligned} \Pr(\mathcal{E}_{2a} \cap \mathcal{F}_2^c) &\leq \sum_{\substack{P_{U, X_1, X_2, \tilde{X}_1, \tilde{X}_2, S} \in \mathcal{D}_{2a}: \\ \mathcal{F}_2^c \text{ holds}}} 2^{-n(I(\tilde{X}_1, \tilde{X}_2; X_1, X_2, S, Y|U) - R_1 - R_2 - 3\varepsilon)} \\ &\leq \sum_{\substack{P_{U, X_1, X_2, \tilde{X}_1, \tilde{X}_2, S} \in \mathcal{D}_{2a}: \\ \mathcal{F}_2^c \text{ holds}}} 2^{-n(I_q(\tilde{X}_1, \tilde{X}_2; Y|U) - R_1 - R_2 - 3\varepsilon)}. \end{aligned} \quad (188)$$

For  $P_{U, X_1, X_2, \tilde{X}_1, \tilde{X}_2, S} \in \mathcal{D}_{2a}$ , we have by (180) that  $P_{\tilde{X}_1, \tilde{X}_2, \tilde{S}, Y|U}$  is arbitrarily close to some  $P_{X_1, X_2, \tilde{S}, \tilde{Y}|U}$ , where

$$P_{X_1, X_2, \tilde{S}, \tilde{Y}|U}(x_1, x_2, s, y|u) = P_{X_1|U}(x_1|u)P_{X_2|U}(x_2|u)\tilde{q}(s|u)W_{Y|X_1, X_2, S}(y|x_1, x_2, s), \quad (189)$$

if  $\eta > 0$  is sufficiently small. In which case,

$$I_q(\tilde{X}_1, \tilde{X}_2; Y|U) \geq I_{\tilde{q}}(X_1, X_2; Y|U) - \delta, \quad (190)$$

where  $\delta > 0$  is arbitrarily small. Therefore, provided that

$$R_1 + R_2 < \min_{q(s|u): \mathbb{E}_q l(S) \leq \Lambda} I_q(X_1, X_2; Y|U) - \delta - 5\varepsilon, \quad (191)$$

we have that  $\Pr(\mathcal{E}_{2a} \cap \mathcal{F}_2^c) \leq 2^{-n(I_q(\tilde{X}_1, \tilde{X}_2; Y|U) - R_1 - R_2 - 4\varepsilon)}$  tends to zero as  $n \rightarrow \infty$ .

Next, consider Condition 2b) of the decoding rule, and let  $\mathcal{D}_{2b}$  denote the set of joint types  $P_{U,X_1,X_2,\tilde{X}_1,S}$  such that

$$D(P_{U,X_1,X_2,S,Y} \| P_U \times P_{X_1|U} \times P_{X_2|U} \times P_{S|U} \times W_{Y|X_1,X_2,S}) \leq \eta, \quad (192)$$

$$D(P_{U,\tilde{X}_1,X_2,\tilde{S},Y} \| P_U \times P_{\tilde{X}_1|U} \times P_{X_2|U} \times P_{\tilde{S}|U} \times W_{Y|X_1,X_2,S}) \leq \eta, \text{ for some } \tilde{S} \sim \tilde{q}(s|u) \quad (193)$$

$$I_q(X_1, X_2, Y; \tilde{X}_1|U, S) > \eta. \quad (194)$$

Observe that when the event  $\mathcal{E}_{2b}$  occurs, *i.e.* Condition 2b) of the decoding rule is violated, then there exists  $\tilde{m}_1 \neq m_1$  such that (193) and (194) hold for some  $\tilde{s}^n \in \mathcal{S}^n$  with  $l^n(\tilde{s}^n) \leq \Lambda$ . Then, by standard type class considerations (see *e.g.* [35, Theorem 1.3]),

$$\begin{aligned} & \Pr(\mathcal{E}_{2b} \cap \mathcal{E}_1^c \cap \mathcal{F}_3^c \mid M_1 = m_1, M_2 = m_2) \\ & \leq \sum_{\substack{P_{U,X_1,X_2,\tilde{X}_1,S} \in \mathcal{D}_{2b}: \\ \mathcal{F}_3^c \text{ holds}}} |\{\tilde{m}_1 : (u^n, x_1^n(m_1), x_2^n(m_2), x_1^n(\tilde{m}_1), s^n) \in \mathcal{T}^n(P_{U,X_1,X_2,\tilde{X}_1,S})\}| \cdot 2^{-n(I_q(\tilde{X}_1; Y|U, X_1, X_2, S) - \varepsilon)}, \end{aligned} \quad (195)$$

for every given  $m_1 \in [1 : 2^{nR_1}]$  and  $m_2 \in [1 : 2^{nR_2}]$ . Hence, by (55),

$$\Pr(\mathcal{E}_{2b} \cap \mathcal{E}_1^c \cap \mathcal{F}_3^c) \leq \sum_{\substack{P_{U,X_1,X_2,\tilde{X}_1,S} \in \mathcal{D}_{2b}: \\ \mathcal{F}_3^c \text{ holds}}} 2^{-n(I_q(\tilde{X}_1; Y|U, X_1, X_2, S) - [R_1 - I_q(\tilde{X}_1; X_1, X_2, S|U)]_+ - 2\varepsilon)}. \quad (196)$$

If  $R_1 \leq I_q(\tilde{X}_1; S|U)$ , then given  $\mathcal{F}_3^c$ , we have that

$$I_q(X_1, X_2; \tilde{X}_1|U, S) \leq I_q(X_1, X_2; \tilde{X}_1, S|U) \leq \varepsilon. \quad (197)$$

Hence, by (194),

$$I_q(\tilde{X}_1; Y|U, X_1, X_2, S) = I_q(\tilde{X}_1; X_1, X_2, Y|U, S) - I_q(\tilde{X}_1; X_1, X_2|U, S) \geq \eta - \varepsilon. \quad (198)$$

Returning to (196), we note that since the number of types is polynomial in the sequence length,  $|\mathcal{D}_{2b}| \leq 2^{n\varepsilon}$  for sufficiently large  $n$ . Hence, by (196) and (198), we have that  $\Pr(\mathcal{E}_{2b} \cap \mathcal{F}_3^c) \leq 2^{-n(\eta - 4\varepsilon)}$ , which tends to zero as  $n \rightarrow \infty$ , for  $\varepsilon < \frac{1}{4}\eta$ .

Otherwise, if  $R_1 > I_q(\tilde{X}_1; S|U)$ , then given  $\mathcal{F}_3^c$ ,

$$\begin{aligned} R_1 & > I_q(X_1, X_2; \tilde{X}_1, S|U) + I(\tilde{X}_1; S|U) - \varepsilon \\ & = I_q(\tilde{X}_1; X_1, X_2, S|U) + I(X_1, X_2; S|U) - \varepsilon \\ & \geq I_q(\tilde{X}_1; X_1, X_2, S|U) - \varepsilon. \end{aligned} \quad (199)$$

Thus,

$$\left[ R_1 - I_q(\tilde{X}_1; X_1, X_2, S|U) \right]_+ \leq R_1 - I_q(\tilde{X}_1; X_1, X_2, S|U) + \varepsilon. \quad (200)$$

Hence, by (196) we have that

$$\begin{aligned} \Pr(\mathcal{E}_{2b} \cap \mathcal{E}_1^c \cap \mathcal{F}_3^c) & \leq \sum_{\substack{P_{U,X_1,X_2,\tilde{X}_1,S} \in \mathcal{D}_{2b}: \\ \mathcal{F}_3^c \text{ holds}}} 2^{-n(I_q(\tilde{X}_1; X_1, X_2, S, Y|U) - R_1 - 3\varepsilon)} \\ & \leq \sum_{\substack{P_{U,X_1,X_2,\tilde{X}_1,S} \in \mathcal{D}_{2b}: \\ \mathcal{F}_3^c \text{ holds}}} 2^{-n(I_q(\tilde{X}_1; Y|X_2, U) - R_1 - 3\varepsilon)}, \end{aligned} \quad (201)$$

where the last inequality holds since

$$\begin{aligned} I_q(\tilde{X}_1; X_1, X_2, S, Y|U) & = I_q(\tilde{X}_1; Y|X_2, U) + I_q(\tilde{X}_1; X_2|U) + I_q(\tilde{X}_1; X_1, S|X_2, Y, U) \\ & \geq I_q(\tilde{X}_1; Y|X_2, U). \end{aligned} \quad (202)$$

For  $P_{U,X_1,X_2,\tilde{X}_1,S} \in \mathcal{D}_{2b}$ , we have by (193) that  $P_{\tilde{X}_1,X_2,\tilde{S},Y|U}$  is arbitrarily close to some  $P_{X_1,X_2,\tilde{S},\tilde{Y}|U}$ , where

$$P_{X_1,X_2,\tilde{S},\tilde{Y}|U}(x_1, x_2, s, y|u) = P_{X_1|U}(x_1|u)P_{X_2|U}(x_2|u)\tilde{q}(s|u)W_{Y|X_1,X_2,S}(y|x_1, x_2, s), \quad (203)$$

if  $\eta > 0$  is sufficiently small. In which case,

$$I_q(\tilde{X}_1; Y|X_2, U) \geq I_q(X_1; Y|X_2, U) - \delta_1, \quad (204)$$

where  $\delta_1 > 0$  is arbitrarily small. Therefore, provided that

$$R_1 < \min_{q(s|u): \mathbb{E}_q l(S) \leq \Lambda} I_q(X_1; Y|X_2, U) - \delta_1 - 5\varepsilon \quad (205)$$

we have that  $\Pr(\mathcal{E}_{2b} \cap \mathcal{F}_3^c) \leq 2^{-n(I_q(\tilde{X}_1; Y|X_2, U) - R_1 - 4\varepsilon)}$  tends to zero as  $n \rightarrow \infty$ .

In same manner, it can be shown that  $\Pr(\mathcal{E}_{2c} \cap \mathcal{F}_4^c)$  tends to zero as  $n \rightarrow \infty$ , provided that

$$R_2 < \min_{q(s|u): \mathbb{E}_q l(S) \leq \Lambda} I_q(X_2; Y|X_1, U) - \delta_2 - 5\varepsilon, \quad (206)$$

where  $\delta_2 > 0$  is arbitrarily small.  $\square$

*Converse Proof:* We will use the following lemma, based on the observations of Gubner [24].

**Lemma 16.** Consider the AVMAC free of state constraints, and let  $\mathcal{C} = (f_1, f_2, g)$  be a  $(2^{nR_1}, 2^{nR_2}, n)$  deterministic code.

1) Suppose that  $W_{Y|X_1, X_2, S}$  is symmetrizable- $\mathcal{X}_1 \times \mathcal{X}_2$ , and let  $J_i(s|x_1, x_2)$ ,  $i \in [1 : n]$ , be a set of conditional state distributions that satisfy (13). If  $R_1 + R_2 > 0$ , then

$$P_e^{(n)}(\tilde{q}, \mathcal{C}) \geq \frac{1}{4}, \quad (207)$$

for

$$\tilde{q}(s^n) = \frac{1}{2^{n(R_1+R_2)}} \sum_{m_1=1}^{2^{nR_1}} \sum_{m_2=1}^{2^{nR_2}} J^n(s^n|f_1(m_1), f_2(m_2)), \quad (208)$$

where  $J^n(s^n|x_1^n, x_2^n) = \prod_{i=1}^n J_i(s_i|x_{1,i}, x_{2,i})$ .

2) Suppose that  $W_{Y|X_1, X_2, S}$  is symmetrizable- $\mathcal{X}_1|\mathcal{X}_2$ , and let  $J_{1,i}(s|x_1)$ ,  $i \in [1 : n]$ , be a set of conditional state distributions that satisfy (14). If  $R_1 > 0$ , then

$$P_e^{(n)}(\tilde{q}_1, \mathcal{C}) \geq \frac{1}{4}, \quad (209)$$

for

$$\tilde{q}_1(s^n) = \frac{1}{2^{nR_1}} \sum_{m_1=1}^{2^{nR_1}} J_1^n(s^n|f_1(m_1)), \quad (210)$$

where  $J_1^n(s^n|x_1^n) = \prod_{i=1}^n J_{1,i}(s_i|x_{1,i})$ .

3) Suppose that  $W_{Y|X_1, X_2, S}$  is symmetrizable- $\mathcal{X}_2|\mathcal{X}_1$ , and let  $J_{2,i}(s|x_2)$ ,  $i \in [1 : n]$ , be a set of conditional state distributions that satisfy (15). If  $R_2 > 0$ , then

$$P_e^{(n)}(\tilde{q}_2, \mathcal{C}) \geq \frac{1}{4}, \quad (211)$$

for

$$\tilde{q}_2(s^n) = \frac{1}{2^{nR_2}} \sum_{m_2=1}^{2^{nR_2}} J_2^n(s^n|f_2(m_2)), \quad (212)$$

where  $J_2^n(s^n|x_2^n) = \prod_{i=1}^n J_{2,i}(s_i|x_{2,i})$ .

For completeness, we give the proof below.

*Proof of Lemma 16.* Denote the codebooks size by  $M_k = 2^{nR_k}$ ,  $k = 1, 2$ ,  $M = 2^{n(R_1+R_2)}$ , and the codewords by  $x_k^n(m_k) = f_k(m_k)$ ,  $k = 1, 2$ .

Under the conditions of part 1,

$$\begin{aligned} P_e^{(n)}(\tilde{q}, \mathcal{C}) &= \sum_{s^n \in \mathcal{S}^n} q(s^n) \frac{1}{M} \sum_{m_1, m_2} \sum_{y^n: g(y^n) \neq (m_1, m_2)} W^n(y^n|x_1^n(m_1), x_2^n(m_2), s^n) \\ &= \frac{1}{M^2} \sum_{\tilde{m}_1, \tilde{m}_2} \sum_{s^n \in \mathcal{S}^n} J^n(s^n|x_1^n(\tilde{m}_1), x_2^n(\tilde{m}_2)) \sum_{m_1, m_2} \sum_{y^n: g(y^n) \neq (m_1, m_2)} W^n(y^n|x_1^n(m_1), x_2^n(m_2), s^n) \end{aligned} \quad (213)$$

where have defined  $W^n \equiv W_{Y^n|X_1^n, X_2^n, S^n}$  for short notation. By switching between the summation indices  $(m_1, m_2)$  and  $(\tilde{m}_1, \tilde{m}_2)$ , we obtain

$$\begin{aligned} P_e^{(n)}(\tilde{q}, \mathcal{E}) &= \frac{1}{2M^2} \sum_{m_1, m_2, \tilde{m}_1, \tilde{m}_2} \sum_{y^n : g(y^n) \neq (m_1, m_2)} \sum_{s^n \in \mathcal{S}^n} W^n(y^n | x_1^n(m_1), x_2^n(m_2), s^n) J^n(s^n | x_1^n(\tilde{m}_1), x_2^n(\tilde{m}_2)) \\ &+ \frac{1}{2M^2} \sum_{m_1, m_2, \tilde{m}_1, \tilde{m}_2} \sum_{y^n : g(y^n) \neq (\tilde{m}_1, \tilde{m}_2)} \sum_{s^n \in \mathcal{S}^n} W^n(y^n | x_1^n(\tilde{m}_1), x_2^n(\tilde{m}_2), s^n) J^n(s^n | x_1^n(m_1), x_2^n(m_2)). \end{aligned} \quad (214)$$

Now, as the channel is memoryless,

$$\begin{aligned} &\sum_{s^n \in \mathcal{S}^n} W^n(y^n | x_1^n(\tilde{m}_1), x_2^n(\tilde{m}_2), s^n) J^n(s^n | x_1^n(m_1), x_2^n(m_2)) \\ &= \prod_{i=1}^n \sum_{s_i \in \mathcal{S}} W_{Y|X_1, X_2, S}(y_i | x_{1,i}(\tilde{m}_1), x_{2,i}(\tilde{m}_2), s_i) J_i(s_i | x_{1,i}(m_1), x_{2,i}(m_2)) \\ &= \prod_{i=1}^n \sum_{s_i \in \mathcal{S}} W_{Y|X_1, X_2, S}(y_i | x_{1,i}(m_1), x_{2,i}(m_2), s_i) J_i(s_i | x_{1,i}(\tilde{m}_1), x_{2,i}(\tilde{m}_2)) \\ &= \sum_{s^n \in \mathcal{S}^n} W^n(y^n | x_1^n(m_1), x_2^n(m_2), s^n) J^n(s^n | x_1^n(\tilde{m}_1), x_2^n(\tilde{m}_2)), \end{aligned} \quad (215)$$

where the second equality is due to (13). Therefore,

$$\begin{aligned} P_e^{(n)}(\tilde{q}, \mathcal{E}) &\geq \frac{1}{2M^2} \sum_{(\tilde{m}_1, \tilde{m}_2) \neq (m_1, m_2)} \sum_{s^n \in \mathcal{S}^n} \left[ \sum_{y^n : g(y^n) \neq (m_1, m_2)} W^n(y^n | x_1^n(m_1), x_2^n(m_2), s^n) J^n(s^n | x_1^n(\tilde{m}_1), x_2^n(\tilde{m}_2)) \right. \\ &+ \left. \sum_{y^n : g(y^n) \neq (\tilde{m}_1, \tilde{m}_2)} W^n(y^n | x_1^n(m_1), x_2^n(m_2), s^n) J^n(s^n | x_1^n(\tilde{m}_1), x_2^n(\tilde{m}_2)) \right] \\ &\geq \frac{1}{2M^2} \sum_{(\tilde{m}_1, \tilde{m}_2) \neq (m_1, m_2)} \sum_{s^n \in \mathcal{S}^n} \sum_{y^n \in \mathcal{Y}^n} W^n(y^n | x_1^n(m_1), x_2^n(m_2), s^n) J^n(s^n | x_1^n(\tilde{m}_1), x_2^n(\tilde{m}_2)) \\ &= \frac{M(M-1)}{2M^2} = \frac{1}{2} \left(1 - \frac{1}{M}\right). \end{aligned} \quad (216)$$

Assuming the sum rate is positive, we have that  $M \geq 2$ , hence  $P_e^{(n)}(\tilde{q}, \mathcal{E}) \geq \frac{1}{4}$ .

Under the conditions of part 2,

$$P_e^{(n)}(\tilde{q}_1, \mathcal{E}) = \frac{1}{M_1^2 M_2} \sum_{\tilde{m}_1} \sum_{s^n \in \mathcal{S}^n} J_1^n(s^n | x_1^n(\tilde{m}_1)) \sum_{m_1, m_2} \sum_{y^n : g(y^n) \neq (m_1, m_2)} W^n(y^n | x_1^n(m_1), x_2^n(m_2), s^n). \quad (217)$$

Then, switching between  $m_1$  and  $\tilde{m}_1$  yields

$$\begin{aligned} P_e^{(n)}(\tilde{q}_1, \mathcal{E}) &= \frac{1}{2M_1^2 M_2} \sum_{m_1, m_2, \tilde{m}_1} \sum_{y^n : g(y^n) \neq (m_1, m_2)} \sum_{s^n \in \mathcal{S}^n} W^n(y^n | x_1^n(m_1), x_2^n(m_2), s^n) J_1^n(s^n | x_1^n(\tilde{m}_1)) \\ &+ \frac{1}{2M_1^2 M_2} \sum_{m_1, m_2, \tilde{m}_1} \sum_{y^n : g(y^n) \neq (\tilde{m}_1, m_2)} \sum_{s^n \in \mathcal{S}^n} W^n(y^n | x_1^n(\tilde{m}_1), x_2^n(m_2), s^n) J_1^n(s^n | x_1^n(m_1)). \end{aligned} \quad (218)$$

Now, as the channel is memoryless, we have by (14) that

$$\begin{aligned} &\sum_{s^n \in \mathcal{S}^n} W^n(y^n | x_1^n(\tilde{m}_1), x_2^n(m_2), s^n) J_1^n(s^n | x_1^n(m_1)) \\ &= \sum_{s^n \in \mathcal{S}^n} W^n(y^n | x_1^n(m_1), x_2^n(m_2), s^n) J_1^n(s^n | x_1^n(\tilde{m}_1)). \end{aligned} \quad (219)$$

Therefore,

$$\begin{aligned}
P_e^{(n)}(\tilde{q}_1, \mathcal{C}) &\geq \frac{1}{2M_1^2 M_2} \sum_{\tilde{m}_1 \neq m_1, m_2} \sum_{s^n \in \mathcal{S}^n} \left[ \sum_{y^n : g(y^n) \neq (m_1, m_2)} W^n(y^n | x_1^n(m_1), x_2^n(m_2), s^n) J_1^n(s^n | x_1^n(\tilde{m}_1)) \right. \\
&\quad \left. + \sum_{y^n : g(y^n) \neq (\tilde{m}_1, m_2)} W^n(y^n | x_1^n(m_1), x_2^n(m_2), s^n) J_1^n(s^n | x_1^n(\tilde{m}_1)) \right] \\
&\geq \frac{1}{2M_1^2 M_2} \sum_{\tilde{m}_1 \neq m_1, m_2} \sum_{s^n \in \mathcal{S}^n} \sum_{y^n \in \mathcal{Y}^n} W^n(y^n | x_1^n(m_1), x_2^n(m_2), s^n) J_1^n(s^n | x_1^n(\tilde{m}_1)) \\
&= \frac{M_1(M_1 - 1)M_2}{2M_1^2} \geq \frac{1}{2} \left( 1 - \frac{1}{M_1} \right). \tag{220}
\end{aligned}$$

Since User 1 has a positive rate,  $R_1 > 0$ , the corresponding codebook has size  $M_1 \geq 2$ , hence  $P_e^{(n)}(q, \mathcal{C}) \geq \frac{1}{4}$ .

The proof of part 3 is similar, and thus omitted.  $\square$

Now, we are in position to prove the converse part of Theorem 8. Consider a sequence of  $(2^{nR_1}, 2^{nR_2}, n, \alpha_n)$  deterministic codes  $\mathcal{C}_n$  over the AVMAC under input constraints  $(\Omega_1, \Omega_2)$  and state constraint  $\Lambda$ , where  $\alpha_n \rightarrow 0$  as  $n \rightarrow \infty$ . In particular, we have that the conditional probability of error given a state sequence  $s^n$  is bounded by

$$P_{e|s^n}^{(n)}(\mathcal{C}_n) \leq \alpha_n, \text{ for } s^n \in \mathcal{S}^n \text{ with } l^n(s^n) \leq \Lambda. \tag{221}$$

For simplicity, we assume that both  $R_1 > 0$  and  $R_2 > 0$ , but the proof can be easily modified elsewhere.

First, we show that

$$R_1 \leq \min_{q(s|u) : E_{q_l(S)} \leq \Lambda} I_q(X_1; Y | X_2, U) + \varepsilon_n, \tag{222}$$

$$R_2 \leq \min_{q(s|u) : E_{q_l(S)} \leq \Lambda} I_q(X_2; Y | X_1, U) + \varepsilon_n, \tag{223}$$

$$R_1 + R_2 \leq \min_{q(s|u) : E_{q_l(S)} \leq \Lambda} I_q(X_1, X_2; Y | U) + \varepsilon_n, \tag{224}$$

where  $\varepsilon_n > 0$  tends to zero as  $n \rightarrow \infty$ . To this end, consider using the same code in the following setting. Consider a different channel model, with an average state constraint. Specifically, consider a MAC where the jammer selects an independent state sequence at random,  $\bar{S}^n \sim \prod_{i=1}^n \bar{q}_i(z_i)$ , under the *average* state constraint  $\frac{1}{n} \sum_{i=1}^n \mathbb{E}l(S_i) \leq \Lambda - \delta$ . Here, there is no state constraint with probability 1, as the jammer may select a sequence  $\bar{S}^n$  with  $l^n(\bar{S}^n) > \Lambda$ . We claim that the code sequence of the constrained AVMAC achieves the same rate pair  $(R_1, R_2)$  over the “new” MAC  $W_{Y|X_1, X_2, \bar{S}}$ , which is governed by the state sequence  $\bar{S}^n$ , under an average constraint. Indeed, using the code  $\mathcal{C}_n$  over the MAC  $W_{Y|X_1, X_2, \bar{S}}$ , the probability of error is given by

$$\begin{aligned}
P_e^{(n)}(\bar{q}, \mathcal{C}_n) &= \sum_{s^n \in \mathcal{S}^n} \bar{q}^n(s^n) P_{e|s^n}^{(n)}(\mathcal{C}_n) \\
&\leq \sum_{s^n \in \mathcal{S}^n : l^n(s^n) \leq \Lambda} \bar{q}^n(s^n) P_{e|s^n}^{(n)}(\mathcal{C}_n) + \Pr\left(l^n(\bar{S}^n) \geq \Lambda\right). \tag{225}
\end{aligned}$$

By (221), we have that the sum in the RHS is bounded by  $\alpha_n$ , hence tends to zero as  $n \rightarrow \infty$ . As for the second term,

$$\Pr\left(l^n(\bar{S}^n) \geq \Lambda\right) \leq \Pr\left(\frac{1}{n} \sum_{i=1}^n (l(\bar{S}_i) - \mathbb{E}l(\bar{S}_i)) \geq \delta\right) \leq \frac{\sum_{i=1}^n \text{Var}(l(\bar{S}_i))}{n^2 \delta^2} \leq \frac{l_{\max}^2}{n \delta^2} \tag{226}$$

where the first inequality holds since  $\frac{1}{n} \sum_{i=1}^n \mathbb{E}l(\bar{S}_i) \leq \Lambda - \delta$ , and the second is due to Chebyshev’s inequality. Thus, we have by (225) that the probability of error tends to zero as  $n \rightarrow \infty$ , when using the code  $\mathcal{C}_n$  over the MAC governed by  $\bar{S}^n$ .

Therefore, it suffices to prove the converse part for the MAC  $W_{Y|X_1, X_2, \bar{S}}$  governed by the state sequence  $\bar{S}^n \sim \bar{q}^n(s^n) = \prod_{i=1}^n \bar{q}_i(s_i)$ . Then, let  $X_1^n = f_1^n(M_1)$  and  $X_2^n = f_2^n(M_2)$  be the channel input sequences, and  $Y^n$  be the corresponding output sequence. Fano’s inequality implies that for every jamming strategy  $\bar{q}^n(s^n)$ ,

$$R_1 \leq \frac{1}{n} \sum_{i=1}^n I_{\bar{q}_i}(X_{1,i}; Y_i | X_{2,i}) + \varepsilon_n \tag{227}$$

$$R_2 \leq \frac{1}{n} \sum_{i=1}^n I_{\bar{q}_i}(X_{2,i}; Y_i | X_{1,i}) + \varepsilon_n \tag{228}$$

$$R_1 + R_2 \leq \frac{1}{n} \sum_{i=1}^n I_{\bar{q}_i}(X_{1,i}, X_{2,i}; Y_i) + \varepsilon_n. \tag{229}$$

(see [14, Section 15.3.4]). Let  $U$  be a random variable which is uniformly distributed over  $[1 : n]$ , and independent of  $(X_1^n, X_2^n, S^n, Y^n)$ . Then, the bounds can be expressed as

$$R_1 \leq I_q(X_{1,U}; Y_U | X_{2,U}, U) + \varepsilon_n, \quad (230)$$

$$R_2 \leq I_q(X_{2,U}; Y_U | X_{1,U}, U) + \varepsilon_n, \quad (231)$$

$$R_1 + R_2 \leq I_q(X_{1,U}, X_{2,U}; Y_U | U) + \varepsilon_n, \quad (232)$$

where we have defined  $q(s|u) = \bar{q}_u$  for  $u \in [1 : n]$ . Then, the bounds (230)-(232) hold for every conditional state distribution  $q(s|u)$  such that  $\mathbb{E}l(S_U) \leq \Lambda$ . Thus, the bounds in (222)-(224) follow by defining

$$X_1 = X_{1,U}, \quad X_2 = X_{2,U}, \quad \text{and } Y = Y_U, \quad (233)$$

Note that  $X_1$  and  $X_2$  are conditionally independent given  $U$ , as required.

Returning to the original AVMAC, we now show that  $\tilde{\Lambda}(P_{U, X_1, X_2}) \geq \Lambda$ . If the AVMAC is non-symmetrizable- $\mathcal{X}_1 \times \mathcal{X}_2$ , then  $\tilde{\Lambda}(P_{U, X_1, X_2}) = +\infty$ , and there is nothing to show. Hence, consider the case where the AVMAC is symmetrizable- $\mathcal{X}_1 \times \mathcal{X}_2$ . Assume to the contrary that  $\tilde{\Lambda}(P_{U, X_1, X_2}) < \Lambda$ . Based on Remark 8, and our definition of the external variable  $U$ , this means that there exist conditional state distributions  $J_i(s|x_1, x_2)$ ,  $i \in [1 : n]$ , which symmetrize- $\mathcal{X}_1 \times \mathcal{X}_2$  the AVMAC, such that

$$\tilde{\Lambda}(P_{U, X_1, X_2}) = \frac{1}{n} \sum_{i=1}^n \sum_{x_{1,i}, x_{2,i}, s_i} P_{X_{1,i}, X_{2,i}}(x_{1,i}, x_{2,i}) J_i(s_i | x_{1,i}, x_{2,i}) l(s_i) \leq \Lambda. \quad (234)$$

Now, consider the following jamming strategy. First, the jammer selects from the codebooks a pair of codewords  $(\tilde{X}_1^n, \tilde{X}_2^n)$  uniformly at random. Then, the jammer selects a sequence  $\tilde{S}^n$  at random, according to the conditional distribution

$$\Pr(\tilde{S}^n = s^n | \tilde{X}_1 = x_1^n, \tilde{X}_2 = x_2^n) = J^n(s^n | x_1^n, x_2^n) \triangleq \prod_{i=1}^n J_i(s_i | x_{1,i}, x_{2,i}). \quad (235)$$

At last, if  $l^n(\tilde{S}^n) \leq \Lambda$ , the jammer chooses the state sequence to be  $S^n = \tilde{S}^n$ . Otherwise, the jammer chooses  $S^n$  to be some sequence of zero cost. Such jamming strategy satisfies the state constraint  $\Lambda$  with probability 1.

To contradict our assumption that  $\tilde{\Lambda}(P_{U, X_1, X_2}) < \Lambda$ , we first show that  $\mathbb{E}l^n(\tilde{S}^n) = \tilde{\Lambda}(P_{U, X_1, X_2})$ . Observe that for every  $(x_1^n, x_2^n) \in \mathcal{X}_1^n \times \mathcal{X}_2^n$ ,

$$\mathbb{E}(l^n(\tilde{S}^n) | \tilde{X}_1^n = x_1^n, \tilde{X}_2^n = x_2^n) = \frac{1}{n} \sum_{i=1}^n \sum_{s \in \mathcal{S}} l(s) J_i(s | x_{1,i}, x_{2,i}). \quad (236)$$

Since  $(\tilde{X}_1^n, \tilde{X}_2^n)$  are distributed as  $(X_1^n, X_2^n)$ , we obtain

$$\begin{aligned} \mathbb{E}l^n(\tilde{S}^n) &= \sum_{s \in \mathcal{S}} l(s) \cdot \frac{1}{n} \sum_{i=1}^n \mathbb{E}J_i(s | X_{1,i}, X_{2,i}) \\ &= \sum_{u, x_1, x_2, s} P_U(u) P_{X_1, X_2 | U}(x_1, x_2 | u) J_u(s | x_1, x_2) l(s) \\ &= \tilde{\Lambda}(P_{X_1, X_2}) < \Lambda. \end{aligned} \quad (237)$$

Thus, by Chebyshev's inequality we have that for sufficiently large  $n$ ,

$$\Pr(l^n(\tilde{S}^n) > \Lambda) \leq \delta_0, \quad (238)$$

where  $\delta_0 > 0$  is arbitrarily small. Now, on the one hand, the probability of error is bounded by

$$\begin{aligned} P_e^{(n)}(q, \mathcal{C}_n) &\geq \Pr(g(Y^n) \neq (M_1, M_2), l^n(\tilde{S}^n) \leq \Lambda) \\ &= \sum_{s^n : l^n(s^n) \leq \Lambda} \tilde{q}(s^n) P_{e|s^n}^{(n)}(\mathcal{C}_n), \end{aligned} \quad (239)$$

where  $\tilde{q}(s^n)$  is as defined in (208). On the other hand, the sequence  $\tilde{S}^n$  can be thought of as the state sequence of an AVMAC without a state constraint, hence, by part 1 of Lemma 16,

$$\begin{aligned} \frac{1}{4} \leq P_e^{(n)}(\tilde{q}, \mathcal{C}_n) &\leq \sum_{s^n : l^n(s^n) \leq \Lambda} \tilde{q}(s^n) P_{e|s^n}^{(n)}(\mathcal{C}_n) + \Pr(l^n(\tilde{S}^n) > \Lambda) \\ &\leq \sum_{s^n : l^n(s^n) \leq \Lambda} \tilde{q}(s^n) P_{e|s^n}^{(n)}(\mathcal{C}_n) + \delta_0. \end{aligned} \quad (240)$$

Thus, by (239)-(240), the probability of error is bounded by  $P_e^{(n)}(q, \mathcal{C}_n) \geq \frac{1}{4} - \delta_0$ . As this cannot be the case for a code with vanishing probability of error, we deduce that the assumption is false, *i.e.*  $\tilde{\Lambda}(P_{U, X_1, X_2}) \geq \Lambda$ .

It remains to show that  $\tilde{\Lambda}_1(P_{U, X_1}) \geq \Lambda$  and  $\tilde{\Lambda}_2(P_{U, X_2}) \geq \Lambda$ . Due to the symmetry, it suffices to show this for User 1. We only need to consider an AVMAC which is symmetrizable- $\mathcal{X}_1|\mathcal{X}_2$ , as otherwise,  $\tilde{\Lambda}_1(P_{X_1}) = +\infty$ . Then, assume to the contrary that  $\tilde{\Lambda}_1(P_{X_1}) < \Lambda$ , and let  $J_{1,i}(s|x_1)$ ,  $i \in [1 : n]$ , be the symmetrizing distributions that satisfy (14) and achieves the minimum in (32b), *i.e.*

$$\tilde{\Lambda}_1(P_{U, X_1}) = \sum_{u, x_1, s} P_U(u) P_{X_1|U}(x_1|u) J_{1,u}(s|x_1) l(s) < \Lambda. \quad (241)$$

Consider a jamming strategy, where the jammer first selects a codeword  $\tilde{X}_1^n$  from the codebook of User 1, uniformly at random. Then, the jammer selects a sequence  $\tilde{S}_1^n$  at random, according to the conditional distribution

$$\Pr\left(\tilde{S}_1^n = s^n \mid \tilde{X}_1^n = x_1^n\right) = J_1^n(s^n|x_1^n) \triangleq \prod_{i=1}^n J_{1,i}(s_i|x_{1,i}). \quad (242)$$

At last, if  $l^n(\tilde{S}_1^n) \leq \Lambda$ , the jammer chooses the state sequence to be  $S^n = \tilde{S}_1^n$ . Otherwise, the jammer chooses  $S^n$  to be some sequence of zero cost.

To contradict our assumption that  $\tilde{\Lambda}_1(P_{U, X_1}) < \Lambda$ , we first show that  $\mathbb{E}l^n(\tilde{S}_1^n) = \tilde{\Lambda}_1(P_{U, X_1})$ . Observe that for every  $x_1^n \in \mathcal{X}_1^n$ ,

$$\mathbb{E}\left(l^n(\tilde{S}_1^n) \mid \tilde{X}_1^n = x_1^n\right) = \frac{1}{n} \sum_{i=1}^n \sum_{s \in \mathcal{S}} l(s) J_{1,i}(s|x_{1,i}). \quad (243)$$

Since  $\tilde{X}_1^n$  is distributed as  $X_1^n$ , we obtain

$$\mathbb{E}l^n(\tilde{S}_1^n) = \sum_{u, s \in \mathcal{S}} P_U(u) P_{X_1|U}(x_1|u) J_{1,u}(s|x_1) l(s) = \tilde{\Lambda}_1(P_{U, X_1}) < \Lambda. \quad (244)$$

where the last equality is due to (241). Next, the probability of error is bounded by

$$\begin{aligned} P_e^{(n)}(q, \mathcal{C}_n) &\geq \Pr\left(g(Y^n) \neq (M_1, M_2), l^n(\tilde{S}_1^n) \leq \Lambda\right) \\ &= \sum_{s^n : l^n(s^n) \leq \Lambda} \tilde{q}_1(s^n) P_{e|s^n}^{(n)}(\mathcal{C}_n), \end{aligned} \quad (245)$$

where  $\tilde{q}_1(s^n)$  is as defined in (210). On the other hand, the sequence  $\tilde{S}_1^n$  can be thought of as the state sequence of an AVMAC without a state constraint, hence, by part 2 of Lemma 16,

$$\begin{aligned} \frac{1}{4} &\leq P_e^{(n)}(\tilde{q}_1, \mathcal{C}_n) \leq \sum_{s^n : l^n(s^n) \leq \Lambda} \tilde{q}_1(s^n) P_{e|s^n}^{(n)}(\mathcal{C}_n) + \Pr\left(l^n(\tilde{S}_1^n) > \Lambda\right) \\ &\leq \sum_{s^n : l^n(s^n) \leq \Lambda} \tilde{q}_1(s^n) P_{e|s^n}^{(n)}(\mathcal{C}_n) + \delta_1, \end{aligned} \quad (246)$$

where the last line is due to (244) and Chebyshev's inequality, with arbitrarily small  $\delta_1 > 0$ . Thus, (245)-(246) imply that  $P_e^{(n)}(q, \mathcal{C}_n) \geq \frac{1}{4} - \delta_1$ , which cannot hold for a code with vanishing probability of error. We deduce that the assumption is false, *i.e.*  $\tilde{\Lambda}_1(P_{U, X_1}) \geq \Lambda$ . This completes the converse proof.  $\square$

### Case B and Case C

Before we begin, we note that Cases B-D can also be proved by directly adjusting the techniques of Csiszár and Narayan for the single user AVC [17]. Although, as explained in Remark 3, it is not an immediate consequence. Thereby, it is easier for us to use our previous derivations instead.

The proof follows similar arguments as in Case A, and thus we only give the outline. Since Case B and Case C in Definition 6 are symmetric, we only treat the former. Suppose that  $L^* > \Lambda$  and  $L_2^* > \Lambda$ , but  $L_1^* < \Lambda$ . For the direct part, we can use the same coding scheme as in Case A with the following changes. First, coded time sharing is no longer necessary, hence we take  $U = \emptyset$ . Then, let  $P_{X_1}$  and  $P_{X_2}$  be types, such that  $\mathbb{E}\phi_k(X_k) \leq \Omega_k$ , for  $k = 1, 2$ ,  $\tilde{\Lambda}_2(P_{X_2}) > \Lambda$  and  $\tilde{\Lambda}(P_{X_1, X_2}) > \Lambda$ . As User 1 transmits at zero rate, we can discard of Condition 2b) of the decoding rule (see Definition 7). Nevertheless, given our assumption in Remark 1, Encoder 1 may use "local randomness" and send a sequence  $x_1^n = f_1(\sigma)$ , where  $\sigma \in [1 : 2^{nR_1}]$  is drawn uniformly at random, with  $R_1 = \varepsilon$ . Upon receiving  $y^n \in \mathcal{Y}^n$ , the decoder declares its estimation  $g(y^n) = m_2$  iff there exists  $\sigma$  such that  $y^n \in \mathcal{D}(\sigma, m_2)$ , where the decoding sets  $\mathcal{D}(\sigma, m_2) \subseteq \mathcal{Y}^n$  are as in Definition 7. The message of User 2 is still decoded uniquely, since the only part of Lemma 10 that depends on  $\tilde{\Lambda}_1(P_{X_1})$  is part 2, which is no longer necessary.

The analysis of the probability of error remains exactly the same, except that the error event  $\mathcal{E}_{2b}$  can be ignored. It follows that the probability of error tends to zero, provided that

$$\begin{aligned} R_1 + R_2 &< \min_{q(s): \mathbb{E}_q l(S) \leq \Lambda} I_q(X_1, X_2; Y) - \delta - 5\varepsilon, \\ R_2 &< \min_{q(s): \mathbb{E}_q l(S) \leq \Lambda} I_q(X_2; Y|X_1) - \delta_2 - 5\varepsilon. \end{aligned} \quad (247)$$

Since  $R_1 = \varepsilon$ , and  $I_q(X_1, X_2; Y) \geq I_q(X_2; Y|X_1)$ , the first inequality is inactive, and the direct part follows.

The converse part also follows from the converse proof for Case A. It was shown that if the jammer selects the state sequence to be

$$S^n = \begin{cases} \tilde{S}_1^n & \text{if } l^n(\tilde{S}_1^n) \leq \Lambda, \\ (s_0, \dots, s_0) & \text{otherwise} \end{cases}, \quad (248)$$

for  $\tilde{S}_1^n \sim \tilde{q}_1(s^n)$  as in (210), and  $s_0 \in \mathcal{S}$  with  $l(s_0) = 0$ , then the probability of error is lower bounded by  $P_e^{(n)}(q, \mathcal{C}_n) \geq \frac{1}{4} - \delta_1$  for  $R_1 > 0$ , hence User 1 cannot achieve positive rates. As for User 2, we have by (223) that

$$R_2 \leq \min_{q(s|u): \mathbb{E}_q l(S) \leq \Lambda - \delta} I_q(X_2; Y|X_1, U) + \varepsilon_n. \quad (249)$$

Then, observe that  $I_q(X_2; Y|X_1, U) \leq I_q(X_2; Y|X_1)$ , since  $U \ominus (X_1, X_2) \ominus Y$  form a Markov chain, and conditioning reduces entropy (see *e.g.* [14, Theorem 2.6.5]). By the same considerations as in Case A,  $\tilde{\Lambda}_2(P_{X_2}) \geq \Lambda$ , and the converse part follows.  $\square$

#### Case D

Suppose that  $L^* < \Lambda$ . It was shown in the converse proof for Case A, that if the jammer selects the state sequence

$$S^n = \begin{cases} \tilde{S}^n & \text{if } l^n(\tilde{S}^n) \leq \Lambda, \\ (s_0, \dots, s_0) & \text{otherwise} \end{cases}, \quad (250)$$

for  $\tilde{S}^n \sim \tilde{q}(s^n)$  as in (208), and  $s_0 \in \mathcal{S}$  with  $l(s_0) = 0$ , then the probability of error is lower bounded by  $P_e^{(n)}(q, \mathcal{C}_n) \geq \frac{1}{4} - \delta_0$  for  $R_1 + R_2 > 0$ . Thus, positive rates cannot be achieved.

Now, suppose that both  $L_1^* < \Lambda$  and  $L_2^* < \Lambda$ . We have already seen in the proof of Case B, that  $L_1^* < \Lambda$  implies that User 1 cannot achieve  $R_1 > 0$ , and by symmetry,  $L_2^* < \Lambda$  implies that User 2 cannot achieve  $R_2 > 0$ . Therefore, if  $L^* < \Lambda$ , or both  $L_1^* < \Lambda$  and  $L_2^* < \Lambda$ , then the deterministic code capacity region is  $\{(0, 0)\}$ , as we were set to prove. This concludes the proof of Theorem 8.  $\square$

## APPENDIX F PROOF OF COROLLARY 9

Assume that the AVMAC  $\mathcal{A}$  satisfies the conditions of Corollary 9. Looking into the converse proof of Theorem 8 in Appendix E above, the following addition suffices. We show that for every code  $\mathcal{C}_n$  as in Appendix E,  $\tilde{\Lambda}(P_{U, X_1, X_2}) = \Lambda$  implies that  $R_1 + R_2 = 0$ ,  $\tilde{\Lambda}_1(P_{U, X_1}) = \Lambda$  implies that  $R_1 = 0$ , and  $\tilde{\Lambda}_2(P_{U, X_2}) = \Lambda$  implies that  $R_2 = 0$ . Since there is only a polynomial number of types, we may consider  $P_{U, X_1, X_2}$  to be the joint type of  $(u^n, f_1(m_1), f_2(m_2))$ , for all  $m_1$  and  $m_2$  (see [16, Problem 6.19]).

Suppose that  $\tilde{\Lambda}(P_{U, X_1, X_2}) = \Lambda$ , assume to the contrary that  $R_1 + R_2 > 0$ , and let  $J_u(s|x_1, x_2)$  be distributions that achieve the minimum in (32a), *i.e.*

$$\tilde{\Lambda}(P_{U, X_1, X_2}) = \sum_{u, x_1, x_2, s} P_U(u) P_{X_1, X_2|U}(x_1, x_2|u) J_u(s|x_1, x_2) l(s) = \Lambda. \quad (251)$$

Based on the condition of the corollary, we may assume that  $J_u(s|x_1, x_2)$  is a 0-1 law, *i.e.*

$$J_u(s|x_1, x_2) = \begin{cases} 1 & \text{if } s = G_u(x_1, x_2), \\ 0 & \text{otherwise} \end{cases}, \quad (252)$$

for some deterministic function  $G_u : \mathcal{X}_1 \times \mathcal{X}_2 \rightarrow \mathcal{S}$ . Thus, by (251),

$$\mathbb{E}l(G_U(X_1, X_2)) = \sum_{u, x_1, x_2, s} P_U(u) P_{X_1, X_2|U}(x_1, x_2|u) J_u(s|x_1, x_2) l(s) = \Lambda. \quad (253)$$

Recall that we have defined  $U$  in the converse proof as a uniformly distributed variable over  $\mathcal{U} = [1 : n]$ . Now, consider the following jamming strategy. First, the jammer selects from the codebooks a pair of codewords  $(\tilde{X}_1^n, \tilde{X}_2^n)$  uniformly at random.



Then, given  $\tilde{X}_1^n = x_1^n$  and  $\tilde{X}_2^n = x_2^n$ , the jammer chooses the state sequence  $S^n = (G_i(x_{1,i}, x_{2,i}))_{i=1}^n$ . Observe that given pair of codewords,  $\tilde{X}_1^n = x_1^n$  and  $\tilde{X}_2^n = x_2^n$ ,

$$l^n(S^n) = \frac{1}{n} \sum_{i=1}^n l(G_i(x_{1,i}, x_{2,i})) = \mathbb{E}l(G_U(X_1, X_2)) = \Lambda, \quad (254)$$

where the last equality is due to (253). Thus, the state sequence satisfies the state constraint. Now, observe that the jamming strategy  $S^n = \left(G_i(\tilde{X}_{1,i}, \tilde{X}_{2,i})\right)_{i=1}^n$  is equivalent to  $S^n \sim \tilde{q}(s^n)$  as in (208). Thus, by part 1 of Lemma 16, we have that  $P_e^{(n)}(\tilde{q}, \mathcal{C}_n) \geq \frac{1}{4}$ , hence both users cannot achieve a positive rate.

Next, consider the case where  $\tilde{\Lambda}_1(P_{U, X_1}) = \Lambda$ . Assume to the contrary that  $R_1 > 0$ , and let  $J_{1,u}(s|x_1)$  be distributions that achieves the minimum in (32b), *i.e.*

$$\tilde{\Lambda}_1(P_{U, X_1}) = \sum_{u, x_1, s} P_U(u) P_{X_1|U}(x_1|u) J_{1,u}(s|x_1) l(s) = \Lambda. \quad (255)$$

By assumption, every  $J_{1,u}(s|x_1)$  has a 0-1 law,

$$J_{1,u}(s|x_1) = \begin{cases} 1 & \text{if } s = G_{1,u}(x_1), \\ 0 & \text{otherwise} \end{cases}, \quad (256)$$

for some deterministic function  $G_{1,u} : \mathcal{X}_1 \rightarrow \mathcal{S}$ . Thus, by (255),

$$\mathbb{E}l(G_{1,U}(X_1)) = \sum_{u, x_1} P_U(u) P_{X_1|U}(x_1|u) J_{1,u}(s|x_1) l(s) = \Lambda. \quad (257)$$

Now, suppose the jammer selects from the codebook of User 1, a codeword  $\tilde{X}_1^n$  uniformly at random. Then, given  $\tilde{X}_1^n = x_1^n$ , the jammer chooses the state sequence  $S^n = (G_{1,i}(x_{1,i}))_{i=1}^n$ . Hence, For every given codeword  $\tilde{X}_1^n = x_1^n$ ,

$$l^n(S^n) = \frac{1}{n} \sum_{i=1}^n l(G_{1,i}(x_{1,i})) = \mathbb{E}l(G_{1,U}(X_1)). \quad (258)$$

Thus, by (257), we have that  $l^n(S^n) = \Lambda$  with probability 1. This means that the state sequence satisfies the state constraint. Now, observe that the jamming strategy  $S^n = \left(G_{1,i}(\tilde{X}_{1,i})\right)_{i=1}^n$  is equivalent to  $S^n \sim \tilde{q}_1(s^n)$  as in (210). Thus, by part 2 of Lemma 16, we have that  $P_e^{(n)}(\tilde{q}_1, \mathcal{C}_n) \geq \frac{1}{4}$ , hence  $R_1 = 0$ . By symmetry, we have that  $\tilde{\Lambda}_2(P_{X_2}) = \Lambda$  implies that  $R_2 = 0$ .  $\square$

#### APPENDIX G ANALYSIS OF EXAMPLE 4

Let  $\mathcal{A}$  be the arbitrarily varying binary symmetric MAC, with two independent binary symmetric channels, as in Example 4.

We begin with the random code capacity region. To show achievability, set  $U = \emptyset$ ,  $X_1 \sim \text{Bernoulli}(\omega_1)$ ,  $X_2 \sim \text{Bernoulli}(\omega_2)$ , and observe that

$$\begin{aligned} I_q(X_1; Y_1, Y_2|X_2) &\geq I_q(X_1; Y_1|X_2) \stackrel{(a)}{=} I_q(X_1; Y_1), \\ I_q(X_2; Y_1, Y_2|X_1) &\geq I_q(X_2; Y_2|X_1) \stackrel{(b)}{=} I_q(X_2; Y_2), \end{aligned} \quad (259)$$

and

$$\begin{aligned} I_q(X_1, X_2; Y_1, Y_2) &= H(X_1) + H(X_2) - H_q(X_1, X_2|Y_1, Y_2) \\ &\stackrel{(c)}{\geq} H(X_1) + H(X_2) - H_q(X_1|Y_1, Y_2) - H_q(X_2|Y_1, Y_2) \\ &\stackrel{(d)}{\geq} I_q(X_1; Y_1) + I_q(X_2; Y_2), \end{aligned} \quad (260)$$

where (a) holds since  $X_2$  is independent of  $(X_1, S_1, Y_1)$ ; (b) holds since  $X_1$  is independent of  $(X_2, S_2, Y_2)$ ; (c) is due to the independence bound on entropy [14, Theorem 2.6.6.]; and (d) holds since conditioning reduces entropy [14, Theorem 2.6.5.]. Therefore, based on Theorem 7,  $(R_1, R_2)$  is achievable for

$$R_1 \leq \min_{q(s_1): \mathbb{E}S_1 \leq \Lambda} I_q(X_1; Y_1) = \min_{0 \leq q_1 \leq \Lambda} [h(\omega_1 * q_1) - h(q_1)], \quad (261)$$

$$R_2 \leq \min_{q(s_2): \mathbb{E}S_2 \leq \Lambda} I_q(X_2; Y_2) = \min_{0 \leq q_2 \leq \Lambda} [h(\omega_2 * q_2) - h(q_2)]. \quad (262)$$

Since  $h(\omega * t) - h(t)$  is a convex- $\cup$  function over  $0 \leq t \leq 1$  with minimum at  $t = \frac{1}{2}$ , we have that

$$\mathbb{C}^*(\mathcal{A}) \supseteq \left\{ (R_1, R_2) : \begin{array}{l} R_1 \leq h(\omega_1 * \lambda) - h(\lambda), \\ R_2 \leq h(\omega_2 * \lambda) - h(\lambda) \end{array} \right\}, \quad (263)$$

which completes the achievability proof. To prove the converse part, we observe that the rate of User 1 is bounded by

$$\begin{aligned} & \min_{q(s_1, s_2): \mathbb{E}S_1 + \mathbb{E}S_2 \leq \Lambda} I_q(X_1; Y_1, Y_2 | X_2, U) \leq I_q(X_1; Y_1, Y_2 | X_2, U) \Big|_{\substack{S_1 \sim \text{Bernoulli}(\lambda) \\ S_2 = 0}} \\ & = I_q(X_1; Y_1 | X_2, U) \Big|_{\substack{S_1 \sim \text{Bernoulli}(\lambda) \\ S_2 = 0}} = h(p_1 * \lambda) - h(\lambda) \leq h(\omega_1 * \lambda) - h(\lambda), \end{aligned} \quad (264)$$

with  $X_1 \sim \text{Bernoulli}(p_1)$ , for  $0 \leq p_1 \leq \Omega_1$ , where the first equality holds since  $S_2 = 0$  implies that  $Y_2 = X_2$ , and the last inequality holds since  $h(\alpha * t)$  is a concave- $\cap$  function over  $0 \leq t \leq 1$  with maximum at  $t = \frac{1}{2}$ . Similarly, the rate of User 2 is bounded by

$$\min_{q(s_1, s_2): \mathbb{E}S_1 + \mathbb{E}S_2 \leq \Lambda} I_q(X_2; Y_1, Y_2 | X_1, U) \leq h(\omega_2 * \lambda) - h(\lambda). \quad (265)$$

This proves that the random code capacity region of the AVMAC in Example 4 is given by (64).

Moving to the deterministic code capacity region, we first compute  $L^*$ ,  $L_1^*$  and  $L_2^*$ . For every  $P_{X_1, X_2}$ ,

$$\Psi(P_{X_1, X_2}) = \min_{0 \leq \alpha_1, \alpha_2 \leq 1} (\alpha_1 * p_1 + \alpha_2 * p_2) = \min(p_1, 1 - p_1) + \min(p_2, 1 - p_2), \quad (266)$$

$$\Psi_1(p_1) = \min_{0 \leq \alpha_1 \leq 1} \alpha_1 * p_1 = \min(p_1, 1 - p_1), \quad (267)$$

$$\Psi_2(p_2) = \min_{0 \leq \alpha_2 \leq 1} \alpha_2 * p_2 = \min(p_2, 1 - p_2), \quad (268)$$

where we have used the notation  $p_1 = P_{X_1}(1) = 1 - P_{X_1}(0)$  and  $p_2 = P_{X_2}(1) = 1 - P_{X_2}(0)$ . Therefore,

$$L^* = \max_{\substack{0 \leq p_1 \leq \Omega_1, \\ 0 \leq p_2 \leq \Omega_2}} \Psi(P_{X_1, X_2}) = \omega_1 + \omega_2, \quad (269)$$

$$L_1^* = \max_{0 \leq p_1 \leq \Omega_1} \Psi_1(p_1) = \omega_1, \quad (270)$$

$$L_2^* = \max_{0 \leq p_2 \leq \Omega_2} \Psi_2(p_2) = \omega_2. \quad (271)$$

Observe that based on the result on the random code capacity region, we have that for  $\Lambda \geq \frac{1}{2}$ , or equivalently,  $\lambda = \frac{1}{2}$ , the capacity region is given by  $\mathbb{C}(\mathcal{A}) = \mathbb{C}^*(\mathcal{A}) = \{(0, 0)\}$ , in agreement with (66)-(69). Henceforth, assume that  $\Lambda < \frac{1}{2}$ .

The first case to consider is  $\Omega_1 > \Lambda$  and  $\Omega_2 > \Lambda$ . The converse part is immediate, since the deterministic code capacity region is always bounded by the random code capacity region. As for the direct part, we are going to show that under the assumption that  $\Lambda < \frac{1}{2}$ , we have that  $L^*$ ,  $L_1^*$  and  $L_2^*$  are greater than  $\Lambda$ . Indeed, if  $\Omega_k \geq \frac{1}{2}$ , then  $\omega_k = \frac{1}{2} > \Lambda$ , for  $k = 1, 2$ . Otherwise, if  $\Omega_k < \frac{1}{2}$ , then  $\omega_k = \Omega_k > \Lambda$ , for  $k = 1, 2$ . Therefore, in both cases, we have by (269)-(271) that  $L^* > \Lambda$ ,  $L_1^* > \Lambda$ , and  $L_2^* > \Lambda$ , which corresponds to Case a) in Definition 6. It is further inferred that taking  $p_1 = \omega_1$  and  $p_2 = \omega_2$ , we have that  $\tilde{\Lambda}(p_1, p_2) = \omega_1 + \omega_2 > \Lambda$  and  $\tilde{\Lambda}_k(p_k) = \omega_k > \Lambda$  for  $k = 1, 2$ . It follows that this inputs distribution is legitimate, in the sense that it belongs to the optimization set  $\overline{\mathcal{P}}_{\Omega_1, \Omega_2, \Lambda}(\mathcal{U} \times \mathcal{X}_1 \times \mathcal{X}_2)$  (see definition in (37b)). This completes the achievability proof for the first case, because we have already seen in the achievability proof of the random code capacity region above, that the inputs distribution  $p_1 = \omega_1$  and  $p_2 = \omega_2$  achieves the region in the RHS of (66), with  $U = \emptyset$ .

Next, we consider the second case,  $\Omega_1 \leq \Lambda$  and  $\Omega_2 > \Lambda$ . Assuming that  $\Lambda < \frac{1}{2}$ , we have that  $\Omega_1 < \frac{1}{2}$ , hence  $L_1^* = \omega_1 = \Omega_1 \leq \Lambda$ . As for  $L_2^*$ , if  $\Omega_2 \geq \frac{1}{2}$ , then  $\omega_2 = \frac{1}{2} > \Lambda$ , and if  $\Omega_2 < \frac{1}{2}$ , then  $\omega_2 = \Omega_2 > \Lambda$ . It follows that  $L^* = \omega_1 + \omega_2 > \Lambda$  and  $L_2^* = \omega_2 > \Lambda$ , but  $L_1 \leq \Lambda$ , which corresponds to Case b) in Definition 6. Furthermore, the input distributions  $p_1 = \omega_1$  and  $p_2 = \omega_2$  belong to the maximization set in (37a), as  $\tilde{\Lambda}(p_1 p_2) = \omega_1 + \omega_2 > \Lambda$  and  $\tilde{\Lambda}_2(p_2) = \omega_2 > \Lambda$ . Thus, User 2 can achieve rates below

$$\begin{aligned} & \min_{q(s_1, s_2): \mathbb{E}S_1 + \mathbb{E}S_2 \leq \Lambda} I_q(X_2; Y_1, Y_2 | X_1) \geq \min_{q(s_1, s_2): \mathbb{E}S_1 + \mathbb{E}S_2 \leq \Lambda} I_q(X_2; Y_2 | X_1) \\ & = \min_{0 \leq q_2 \leq \Lambda} [h(\omega_2 * q_2) - h(q_2)] = h(\omega_2 * \lambda) - h(\lambda). \end{aligned} \quad (272)$$

It is also the highest rate achievable for User 2, since

$$\begin{aligned} R_2 & \leq \min_{q(s_1, s_2): \mathbb{E}S_1 + \mathbb{E}S_2 \leq \Lambda} I_q(X_2; Y_1, Y_2 | X_1) \leq I_q(X_2; Y_1, Y_2 | X_1) \Big|_{\substack{S_1 = 0 \\ S_2 \sim \text{Bernoulli}(\lambda)}} \\ & = I_q(X_2; Y_2) \Big|_{\substack{S_1 = 0 \\ S_2 \sim \text{Bernoulli}(\lambda)}} \leq \max_{0 \leq p_2 \leq \Omega_2} h(p_2 * \lambda) - h(\lambda) = h(\omega_2 * \lambda) - h(\lambda), \end{aligned} \quad (273)$$

following the same considerations as in the derivation of the random code capacity region above. The third case,  $\Omega_1 > \Lambda$  and  $\Omega_2 \leq \Lambda$ , follows by symmetry.

In the fourth case,  $\Omega_1 < \Lambda$  and  $\Omega_2 < \Lambda$ , we have that

$$\begin{aligned} L_1^* &= \omega_1 \leq \Omega_1 < \Lambda, \\ L_2^* &= \omega_2 \leq \Omega_2 < \Lambda, \end{aligned} \quad (274)$$

as in Case d) in Definition 6. Thus, the capacity region is  $\mathbb{C}(\mathcal{A}) = \{(0, 0)\}$ .  $\square$

#### APPENDIX H ANALYSIS OF EXAMPLE 5

Deriving the random code capacity region is straightforward. To show achievability, we set  $U = \emptyset$ ,  $X_1 \sim \mathcal{N}(0, \Omega_1)$ , and  $X_2 \sim \mathcal{N}(0, \Omega_2)$ . Then, since Gaussian noise is known to be the worst additive noise under variance constraint [20, Lemma II.2], we have that

$$\min_{F_S: \mathbb{E}S^2 \leq \Lambda} I(X_1; Y|X_2) = \frac{1}{2} \log \left( 1 + \frac{\Omega_1}{\Lambda + \sigma^2} \right), \quad (275)$$

$$\min_{F_S: \mathbb{E}S^2 \leq \Lambda} I(X_2; Y|X_1) = \frac{1}{2} \log \left( 1 + \frac{\Omega_2}{\Lambda + \sigma^2} \right), \quad (276)$$

$$\min_{F_S: \mathbb{E}S^2 \leq \Lambda} I(X_1, X_2; Y) = \frac{1}{2} \log \left( 1 + \frac{\Omega_1 + \Omega_2}{\Lambda + \sigma^2} \right). \quad (277)$$

This proves achievability. To prove the converse part, observe that the rate of User 1 is bounded by

$$\min_{F_S: \mathbb{E}S^2 \leq \Lambda} I(X_1; Y|X_2, U) \leq I(X_1; Y|X_2, U) \Big|_{S \sim \mathcal{N}(0, \Lambda)} \leq \frac{1}{2} \log \left( 1 + \frac{\Omega_1}{\Lambda + \sigma^2} \right), \quad (278)$$

where the last inequality follows as in the converse proof of the classical Gaussian MAC  $Y = X_1 + X_2 + \tilde{Z}$ , with  $\tilde{Z} \sim \mathcal{N}(0, \Lambda + \sigma^2)$  [50]. The bounds on the rate of User 2 and on the sum rate are proved in the same manner. We have thus determined the random code capacity region.

We move to the deterministic code capacity region, as in Theorem 8. First, we calculate the thresholds  $L^*$ ,  $L_1^*$ , and  $L_2^*$ . Based on Definition 5, the Gaussian AVMAC is symmetrized- $\mathcal{X}_1 \times \mathcal{X}_2$  by a conditional pdf  $\varphi(s|x_1, x_2)$  if

$$\int_{-\infty}^{\infty} \varphi(s|\tilde{x}_1, \tilde{x}_2) f_Z(y - x_1 - x_2 - s) ds = \int_{-\infty}^{\infty} \varphi(s|x_1, x_2) f_Z(y - \tilde{x}_1 - \tilde{x}_2 - s) ds, \quad \forall x_1, x_2, \tilde{x}_1, \tilde{x}_2, y \in \mathbb{R}, \quad (279)$$

where  $f_Z(z) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-z^2/2\sigma^2}$ . In particular, observe that (279) holds for  $\varphi(s|x_1, x_2) = \delta(s - x_1 - x_2)$ , where  $\delta(\cdot)$  is the Dirac delta function. In other words, the channel is symmetrized by a distribution  $\varphi(s|x_1, x_2)$  which gives probability 1 to  $S = x_1 + x_2$ . The minimal state cost  $\tilde{\Lambda}(F_{X_1} F_{X_2})$  for the jammer to symmetrize- $\mathcal{X}_1 \times \mathcal{X}_2$ , for the input distribution  $f_{X_1}(x_1) f_{X_2}(x_2)$ , is the continuous version of (32a),

$$\tilde{\Lambda}(F_{X_1} F_{X_2}) = \min \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X_1}(x_1) f_{X_2}(x_2) \varphi(s|x_1, x_2) s^2 ds dx_1 dx_2, \quad (280)$$

where the minimization is over all conditional pdfs  $\varphi(s|x_1, x_2)$  that symmetrize- $\mathcal{X}_1 \times \mathcal{X}_2$  the channel, that is, satisfy (279). Similar expressions can be written for the individual state costs  $\tilde{\Lambda}_1(F_{X_1})$  and  $\tilde{\Lambda}_2(F_{X_2})$  as the continuous versions of (32b) and (32c), respectively. The following lemma states that the minimal state cost for joint symmetrizability is the total input power, and the minimal state cost for individual symmetrizability is the input power of the corresponding transmitter.

*Lemma 17.* For zero mean random variables  $X_1$  and  $X_2$ ,

$$\Psi(F_{X_1} F_{X_2}) = \mathbb{E}X_1^2 + \mathbb{E}X_2^2, \quad (281)$$

$$\Psi_1(F_{X_1}) = \mathbb{E}X_1^2, \quad (282)$$

$$\Psi_2(F_{X_2}) = \mathbb{E}X_2^2. \quad (283)$$

*Proof of Lemma 17.* First, we evaluate  $\Psi(F_{X_1} F_{X_2})$ . Observe that by (280), the Gaussian AVMAC is symmetrized- $\mathcal{X}_1 \times \mathcal{X}_2$  by a conditional pdf  $\varphi_{x_1, x_2}(s) = \varphi(s|x_1, x_2)$  if

$$\int_{-\infty}^{\infty} \varphi_{0,0}(s) f_Z(y - x_1 - x_2 - s) ds = \int_{-\infty}^{\infty} \varphi_{x_1, x_2}(s) f_Z(y - s) ds, \quad (284)$$

for all  $x_1, x_2, y \in \mathbb{R}$ . By substituting  $z = y - x_1 - x_2 - s$  in the LHS, and  $\bar{z} = y - s$  in the RHS, this is equivalent to

$$\int_{-\infty}^{\infty} \varphi_{0,0}(y - x_1 - x_2 - z) f_Z(z) dz = \int_{-\infty}^{\infty} \varphi_{x_1, x_2}(y - \bar{z}) f_Z(\bar{z}) d\bar{z}. \quad (285)$$

For every  $x_1, x_2 \in \mathbb{R}$ , define the random variable  $\bar{S}(x_1, x_2) \sim \varphi_{x_1, x_2}$ . We note that the RHS is the convolution of the pdfs of the random variables  $Z$  and  $\bar{S}(x_1, x_2)$ , while the LHS is the convolution of the pdfs of the random variables  $Z$  and  $\bar{S}(0, 0) + x_1 + x_2$ . This is not surprising since the channel output  $Y$  is a sum of independent random variables, and thus the pdf of  $Y$  is a convolution of pdfs. It follows that  $\varphi_{0,0}(y - x_1 - x_2) = \varphi_{x_1, x_2}(y)$ , and by plugging  $s$  instead of  $y$ , we have that  $\varphi_{x_1, x_2}$  symmetrizes  $\mathcal{X}_1 \times \mathcal{X}_2$  the Gaussian AVMAC if and only if

$$\varphi_{x_1, x_2}(s) = \varphi_{0,0}(s - x_1 - x_2). \quad (286)$$

Then, the corresponding state cost satisfies

$$\begin{aligned} & \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X_1}(x_1) f_{X_2}(x_2) \varphi_{x_1, x_2}(s) s^2 dx_1 dx_2 ds \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X_1}(x_1) f_{X_2}(x_2) \varphi_{0,0}(s - x_1 - x_2) s^2 dx_1 dx_2 ds \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X_1}(x_1) f_{X_2}(x_2) \varphi_{0,0}(a) (a + x_1 + x_2)^2 da dx_1 dx_2 \\ &= \int_{-\infty}^{\infty} \left[ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_1 + x_2 + a)^2 f_{X_1}(x_1) f_{X_2}(x_2) dx_1 dx_2 \right] \varphi_{0,0}(a) da \end{aligned} \quad (287)$$

where the second equality follows by the integral substitution of  $a = s - x_1 - x_2$ . Observe that the bracketed integral can be expressed as  $\mathbb{E}[(X_1 + X_2 + a)^2] = \mathbb{E}X_1^2 + \mathbb{E}X_2^2 + a^2$ . Thus,

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_1 + x_2 + a)^2 f_{X_1}(x_1) f_{X_2}(x_2) dx_1 dx_2 \geq \mathbb{E}X_1^2 + \mathbb{E}X_2^2. \quad (288)$$

The last inequality holds for any  $\varphi_{x_1, x_2}$  which symmetrizes  $\mathcal{X}_1 \times \mathcal{X}_2$  the channel. Now, observe that (286) holds for  $\hat{\varphi}_{x_1, x_2}(s) = \delta(s - x_1 - x_2)$ , where  $\delta(\cdot)$  is the Dirac delta function, hence  $\hat{\varphi}_{x_1, x_2}$  symmetrizes  $\mathcal{X}_1 \times \mathcal{X}_2$  the channel. In addition, since  $\hat{\varphi}_{0,0}$  gives probability 1 to  $S = 0$ , we have that (288) holds with equality for  $\hat{\varphi}_{x_1, x_2}$ , and thus,  $\Psi(F_{X_1} F_{X_2}) = \mathbb{E}X_1^2 + \mathbb{E}X_2^2$ .

Next, consider  $\Psi_1(F_{X_1})$ . The Gaussian AVMAC is symmetrized  $\mathcal{X}_1 | \mathcal{X}_2$  by a conditional pdf  $\varphi'_{x_1}(s) = \varphi_1(s | x_1)$  if

$$\int_{-\infty}^{\infty} \varphi'_0(s) f_Z(y - x_1 - x_2 - s) ds = \int_{-\infty}^{\infty} \varphi'_{x_1}(s) f_Z(y - x_2 - s) ds, \quad (289)$$

for all  $x_1, x_2, y \in \mathbb{R}$ . By substituting  $y' = y - x_2$ ,  $z = y' - x_1 - s$  in the LHS, and  $\bar{z} = y' - s$  in the RHS, this is equivalent to

$$\int_{-\infty}^{\infty} \varphi'_0(y' - x_1 - z) f_Z(z) dz = \int_{-\infty}^{\infty} \varphi'_{x_1}(y' - \bar{z}) f_Z(\bar{z}) d\bar{z}. \quad (290)$$

As earlier, it follows that  $\varphi'_{x_1}$  symmetrizes  $\mathcal{X}_1 | \mathcal{X}_2$  if and only if  $\varphi'_{x_1}(s) = \varphi'_0(s - x_1)$ . By similar derivation as in (287), the corresponding state cost satisfies

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X_1}(x_1) \varphi'_{x_1}(s) s^2 dx_1 ds = \mathbb{E}X_1^2 + \int_{-\infty}^{\infty} a'^2 \varphi'_0(a') da' \geq \mathbb{E}X_1^2, \quad (291)$$

with equality for  $\varphi'_{x_1}(s) = \delta(s - x_1)$ . Hence,  $\tilde{\Lambda}_1(F_{X_1}) = \mathbb{E}X_1^2$ , and by symmetry,  $\Psi_2(F_{X_2}) = \mathbb{E}X_2^2$ . This completes the proof of Lemma 17.  $\square$

Going forward with the derivation of the deterministic code capacity region, we have by Lemma 17 that the thresholds defined in (34)-(36) are given by

$$L^* = \max_{F_{X_1} F_{X_2} : \mathbb{E}X_1^2 \leq \Omega_1, \mathbb{E}X_2^2 \leq \Omega_2} \tilde{\Lambda}(F_{U, X_1, X_2}) = \Omega_1 + \Omega_2, \quad (292)$$

$$L_1^* = \max_{F_{X_1} : \mathbb{E}X_1^2 \leq \Omega_1} \tilde{\Lambda}_1(F_{U, X_1}) = \Omega_1, \quad (293)$$

$$L_2^* = \max_{F_{X_2} : \mathbb{E}X_2^2 \leq \Omega_2} \tilde{\Lambda}_2(F_{U, X_2}) = \Omega_2. \quad (294)$$

We can now complete the derivation by applying Theorem 8 to cases where  $\Omega_k \neq \Lambda$  for  $k = 1, 2$ .

If  $\Omega_1 > \Lambda$  and  $\Omega_2 > \Lambda$ , then  $L^* > \Lambda$ ,  $L_1^* > \Lambda$ , and  $L_2^* > \Lambda$ , which corresponds to Case a) in Definition 6. We have seen that the random code capacity region is achieved with the input distribution specified by  $U = \emptyset$ ,  $X_1 \sim \mathcal{N}(0, \Omega_1)$ , and  $X_2 \sim \mathcal{N}(0, \Omega_2)$ , which is in the set  $\overline{\mathcal{P}}_{\Omega_1, \Omega_2, \Lambda}(\mathcal{U} \times \mathcal{X}_1 \times \mathcal{X}_2)$  (see (37b)). It follows that the capacity region is the same as the random code capacity region, *i.e.*  $\mathbb{C}(\mathcal{A}) = \mathbb{C}^*(\mathcal{A})$ , as in (72).

For the Gaussian AVMAC, as opposed to the scenario discussed in Remark 10, the cases where one of the users has zero capacity can be derived from the single user results. This occurs as the minimal state cost  $\tilde{\Lambda}_1(F_{U, X_1})$  for symmetrizability  $\mathcal{X}_1$ , given in Lemma 17, is the same as the minimal state cost for single user symmetrizability of the Gaussian AVC (see [17]).

Now, based on Csiszár and Narayan's results on the single user Gaussian AVC [18], we have the following. If  $\Omega_1 \leq \Lambda$  and  $\Omega_2 > \Lambda$ , then the individual capacities of User 1 and User 2 are  $C_1 = 0$  and  $C_2 = \frac{1}{2} \log \left( 1 + \frac{\Omega_2}{\Lambda + \sigma^2} \right)$ , respectively, which implies (73). Similarly, if  $\Omega_1 > \Lambda$  and  $\Omega_2 \leq \Lambda$ , then the individual capacities are  $C_1 = \frac{1}{2} \log \left( 1 + \frac{\Omega_1}{\Lambda + \sigma^2} \right)$  and  $C_2 = 0$ , which results in (74). If  $\Omega_1 \leq \Lambda$  and  $\Omega_2 \leq \Lambda$ , then  $C_1 = C_2 = 0$ , hence, (75) follows.  $\square$

#### ACKNOWLEDGMENT

We gratefully thank Vinod M. Prabhakaran (Tata Institute of Fundamental Research) and Yiqi Chen (Shanghai Jiao Tong University) for useful discussions and valuable comments.

#### REFERENCES

- [1] R. Ahlswede. Elimination of correlation in random codes for arbitrarily varying channels. *Z. Wahrscheinlichkeitstheorie Verw. Gebiete*, 44(2):159–175, Jun 1978.
- [2] R. Ahlswede. Arbitrarily varying channels with states sequence known to the sender. *IEEE Trans. Inform. Theory*, 32(5):621–629, Sep 1986.
- [3] R. Ahlswede and N. Cai. Arbitrarily varying multiple-access channels. i. ericson's symmetrizability is adequate, gubner's conjecture is true. *IEEE Trans. Inform. Theory*, 45(2):742–749, Mar 1999.
- [4] R. Ahlswede et al. The capacity region of a channel with two senders and two receivers. *Ann. Probab.*, 2(5):805–814, 1974.
- [5] K. S. K. Arumugam and M. R. Bloch. Keyless covert communication over multiple-access channels. In *Proc. IEEE Int'l Symp. Inform. Theory (ISIT'2016)*, pages 2229–2233, Barcelona, Spain, July 2016.
- [6] C. H. Bennett, G. Brassard, C. Crepeau, and U. M. Maurer. Generalized privacy amplification. *IEEE Trans. Inform. Theory*, 41(6):1915–1923, Nov 1995.
- [7] D. Blackwell, L. Breiman, and A. J. Thomasian. The capacity of a class of channels. *Ann. Math. Statist.*, 30(4):1229–1241, Dec 1959.
- [8] D. Blackwell, L. Breiman, and A. J. Thomasian. The capacities of certain channel classes under random coding. *Ann. Math. Statist.*, 31(3):558–567, Sep 1960.
- [9] H. Boche and R. F. Schaefer. List decoding for arbitrarily varying multiple access channels with conferencing encoders. In *Proc. IEEE Int'l Conf. Commun. (ICC'2014)*, pages 1934–1940, June 2014.
- [10] A. J. Budkuley. Jamming in fading multiple access channels. In *Proc. National Conf. Commun. (NCC'2015)*, pages 1–6, Mumbai, India, Feb 2015.
- [11] N. Cai. List decoding for arbitrarily varying multiple access channel revisited: List configuration and symmetrizability. *IEEE Trans. Inform. Theory*, 62(11):6095–6110, Nov 2016.
- [12] Y. Cernal and Y. Steinberg. The multiple-access channel with partial state information at the encoders. *IEEE Trans. Inform. Theory*, 51(11):3992–4003, Nov 2005.
- [13] R. A. Chou and A. Yener. The gaussian multiple access wiretap channel when the eavesdropper can arbitrarily jam. In *Proc. IEEE Int'l Symp. Inform. Theory (ISIT'2017)*, pages 1958–1962, Aachen, Germany, June 2017.
- [14] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Wiley, 2 edition, 2006.
- [15] I. Csiszár. Arbitrarily varying channels with general alphabets and states. *IEEE Trans. Inform. Theory*, 38(6):1725–1742, Nov 1992.
- [16] I. Csiszár and J. Körner. *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Cambridge University Press, 2 edition, 2011.
- [17] I. Csiszár and P. Narayan. The capacity of the arbitrarily varying channel revisited: positivity, constraints. *IEEE Trans. Inform. Theory*, 34(2):181–193, Mar 1988.
- [18] I. Csiszár and P. Narayan. Capacity of the gaussian arbitrarily varying channel. *IEEE Transactions on Information Theory*, 37(1):18–26, Jan 1991.
- [19] A. Das and P. Narayan. Capacities of time-varying multiple-access channels with side information. In *Proc. IEEE Int'l Symp. Inform. Theory (ISIT'2000)*, page 349, Sorrento, Italy, June 2000.
- [20] S. N. Diggavi and T. M. Cover. The worst additive noise under a covariance constraint. *IEEE Trans. Inform. Theory*, 47(7):3072–3081, Nov 2001.
- [21] A. El Gamal and Y. Kim. *Network Information Theory*. Cambridge University Press, 2011.
- [22] T. Ericson. Exponential error bounds for random codes in the arbitrarily varying channel. *IEEE Trans. Inform. Theory*, 31(1):42–48, Jan 1985.
- [23] J. A. Gubner. *Deterministic Codes for Arbitrarily Varying Multiple-Access Channels*. Ph.d. thesis, University of Maryland, 1988.
- [24] J. A. Gubner. On the deterministic-code capacity of the multiple-access arbitrarily varying channel. *IEEE Trans. Inform. Theory*, 36(2):262–275, March 1990.
- [25] J. A. Gubner. State constraints for the multiple-access arbitrarily varying channel. *IEEE Transactions on Information Theory*, 37(1):27–35, Jan 1991.

- [26] J. A. Gubner. On the capacity region of the discrete additive multiple-access arbitrarily varying channel. *IEEE Trans. Inform. Theory*, 38(4):1344–1347, July 1992.
- [27] J. A. Gubner and B. L. Hughes. Nonconvexity of the capacity region of the multiple-access arbitrarily varying channel subject to constraints. *IEEE Trans. Inform. Theory*, 41(1):3–13, Jan 1995.
- [28] T. Han and K. Kobayashi. A new achievable rate region for the interference channel. *IEEE Trans. Inform. Theory*, 27(1):49–60, Jan 1981.
- [29] X. He, A. Khisti, and A. Yener. MIMO multiple access channel with an arbitrarily varying eavesdropper: secrecy degrees of freedom. *IEEE Trans. Inform. Theory*, 59(8):4733–4745, Aug 2013.
- [30] F. Hosseinigoki and O. Kosut. The gaussian interference channel in the presence of malicious jammers. arXiv: 1712.04133, December 2017. URL <https://arxiv.org/pdf/1712.04133.pdf>.
- [31] F. Hosseinigoki and O. Kosut. Packing lemmas for gaussian jamming networks. In *a talk given in Inform. Theory Appl. Workshop (ITA'2019)*, San Diego, California, February 2019.
- [32] J. Hui and P. Humblet. The capacity region of the totally asynchronous multiple-access channel. *IEEE Trans. Inform. Theory*, 31(2):207–216, March 1985.
- [33] J. H. Jahn. Coding of arbitrarily varying multi-user channels. *Dept. Stat., Stanford Univ., Stanford, CA, Tech. Rep.*, (37), March 1979.
- [34] J. H. Jahn. Coding of arbitrarily varying multiuser channels. *IEEE Trans. Inform. Theory*, 27(2):212–226, Mar 1981.
- [35] G. Kramer. Topics in multi-user information theory. *Foundations and Trends<sup>®</sup> in Communications and Information Theory*, 4(4–5):265–444, 2008.
- [36] A. Lapidoth and Y. Steinberg. The multiple-access channel with causal side information: common state. *IEEE Trans. Inform. Theory*, 59(1):32–50, Jan 2013.
- [37] A. Lapidoth and Y. Steinberg. The multiple-access channel with causal side information: double state. *IEEE Trans. Inform. Theory*, 59(3):1379–1393, March 2013.
- [38] I. Maric, R. D. Yates, and G. Kramer. The discrete memoryless compound multiple access channel with conferencing encoders. In *Proc. IEEE Int'l Symp. Inform. Theory (ISIT'2005)*, pages 407–410, Adelaide, Australia, Sept 2005.
- [39] D. Mitra, H. Asnani, and S. B. R. Pillai. On the sum-capacity of compound MAC models with distributed CSI and unknown fading statistics. In *Conf. Inform. Sciences and Syst. (CISS'2019)*, pages 1–6, Baltimore, Maryland, March 2019.
- [40] S. Nitinawarat. On the deterministic code capacity region of an arbitrarily varying multiple-access channel under list decoding. *IEEE Trans. Inform. Theory*, 59(5):2683–2693, May 2013.
- [41] U. Pereg and Y. Steinberg. The arbitrarily varying channel under constraints with side information at the encoder. *IEEE Trans. Inform. Theory*, 65(2):861–887, Feb 2019.
- [42] G. S. Poltyrev. Coding in an asynchronous multiple-access channel. *Prob. Pered. Inform.*, 19(3):12–21, 1983.
- [43] N. Sangwan, M. Bakshi, B. K. Dey, and V. M. Prabhakaran. Byzantine multiple access. arXiv: 1904.11925, April 2019. URL <https://arxiv.org/pdf/1904.11925.pdf>.
- [44] N. Sangwan, M. Bakshi, B. K. Dey, and V. M. Prabhakaran. Multiple access channels with adversarial users. In *accepted to Proc. IEEE Int'l Symp. Inform. Theory (ISIT'2019)*, Paris, France, July 2019.
- [45] S. Shafiee and S. Ulukus. Capacity of multiple access channels with correlated jamming. In *Proc. IEEE Conf. Military Commun. (MILCOM'2005)*, pages 218–224 Vol. 1, Oct 2005.
- [46] A. Somekh-Baruch, S. Shamai, and S. Verdú. Cooperative multiple-access encoding with states available at one transmitter. *IEEE Trans. Inform. Theory*, 54(10):4448–4469, Oct 2008.
- [47] M. Wiese. *Multiple Access Channels with Cooperating Encoders*. Ph.D. thesis, Technische Universität München, 2013.
- [48] M. Wiese and H. Boche. The arbitrarily varying multiple-access channel with conferencing encoders. *IEEE Trans. Inform. Theory*, 59(3):1405–1416, March 2013.
- [49] M. Wiese, H. Boche, I. Bjelakovic, and V. Jungnickel. The compound multiple access channel with partially cooperating encoders. *IEEE Trans. Inform. Theory*, 57(5):3045–3066, May 2011.
- [50] A. Wyner. Recent results in the Shannon theory. *IEEE Trans. Inform. Theory*, 20(1):2–10, January 1974.
- [51] H. Zivari-Fard, B. Akhbari, M. Ahmadian-Attari, and M. R. Aref. Compound multiple access channel with confidential messages. In *Proc. IEEE Int'l Conf. Commun. (ICC'2019)*, pages 1922–1927, Sydney, Australia, June 2014.