The Arbitrarily Varying Multiple-Access Channel With Conferencing Encoders

Moritz Wiese, *Student Member, IEEE*, and Holger Boche, *Fellow, IEEE*

***Abstract*—We derive the capacity region of arbitrarily varying multiple-access channels (AV-MACs) with conferencing encoders for both deterministic and random coding. For a complete description, it is sufficient that one conferencing capacity is positive. We obtain a dichotomy: either the channel's deterministic capacity region is zero or it equals the 2-D random coding region. We determine exactly when either case holds. We also discuss the benefits of conferencing. We give the example of an AV-MAC which does not achieve any nonzero rate pair without encoder cooperation, but the 2-D random coding capacity region if conferencing is possible. Unlike compound multiple-access channels, arbitrarily varying multiple-access channels may exhibit a discontinuous increase of the capacity region when conferencing in at least one direction is enabled.**

***Index Terms*—Arbitrarily varying channels (AVCs), base-station cooperation, channel uncertainty, compound channels, conferencing encoders.**

I. Introduction

MULTIPLE-ACCESS Channels (MACs) and similar multisender channels with conferencing encoders have attracted attention recently due to the inclusion of base-station cooperation methods in standards for future wireless systems [6], [13], [15], [18]. The original conferencing protocol for the discrete memoryless MAC is due to Willems [19], [20]. The conferencing MAC with imperfect channel state information was modeled as a compound MAC with conferencing encoders and considered in [17]; a different model for channel state uncertainty is given in [14].

This paper covers a very high degree of channel uncertainty in MACs: the channel states may vary arbitrarily over time. The task is to use coding to enable reliable communication for every possible state sequence. The corresponding information-theoretic channel model is the arbitrarily varying MAC (AV-MAC). The random coding capacity region of the AV-MAC without encoder cooperation was determined in [12]. Building on this result, the deterministic coding capacity region of some AV-MACs without cooperation was determined in [4]. In general, it is still open. We will use the "robustification" and "elimination of correlation" techniques developed by Ahlswede in [1] and [2], and partly already used in [12] in a multiuser setting, in order to characterize both the deterministic and random coding capacity regions of any AV-MAC with conferencing encoders, i.e., of any AV-MAC where encoding is done using a Willems conference as in [19] and [20] with at least one positive conferencing capacity. Thus, none of the techniques we apply in this paper is completely new, but in contrast to the nonconferencing situation, they allow for the complete solution of the problems considered here. The rather general "robustification" technique establishes the random coding capacity region of the AV-MAC with conferencing encoders. Both single- and multi-user arbitrarily varying channels (AVCs) are special in that random coding as commonly used in information theory does not yield the same results as deterministic coding. This shows that common randomness shared at the senders and the receiver is an important additional resource. There is a dichotomy: either reliable communication at any nonzero rate pair is impossible with the application of deterministic codes, or the deterministic capacity region coincides with the random coding capacity region, which then is 2-D. In the latter case, one needs the nonstandard "elimination of correlation" [1] for derandomization. It is a two-step protocol which achieves the random coding capacity region if this is possible.

The combination of the elimination technique with conferencing proves to be very fruitful. Here lies the main difference between the AV-MAC with and without conferencing. One can show that there exist channels which only achieve the zero rate pair without transmitter cooperation, but where derandomization using the elimination technique is possible if the transmitters may have a conference. The reason for this is symmetrizability. This can be interpreted in terms of an adversary knowing the channel input symbols and randomizing over the channel states. There are three kinds of symmetrizability for MACs. The capacity region of the AV-MAC without conferencing equals $\{(0,0)\}$ if all three symmetrizability conditions are satisfied. In contrast, the elimination of correlation technique works if the AV-MAC with Willems conferencing encoders does not satisfy the conditions for the first of the three kinds of symmetrizabilities. The two others do not matter. By conferencing with at least one positive conferencing capacity, the AV-MAC gets closer to a single-sender AVC where only one symmetrizability condition exists [8]. This induced change of the channel structure is also reflected in the counterintuitive fact that conferencing with rates tending to zero in blocklength can enlarge the capacity region. The adversary interpretation of symmetrizability highlights the importance of the AV-MAC for the theory of information-theoretic secrecy: if a channel is symmetrizable, an adversary can completely prevent communication.

This paper is organized as follows: Section II is devoted to the formalization of the channel model and the coding problems.

Manuscript received February 07, 2011; revised August 28, 2012; accepted November 06, 2012. Date of publication December 20, 2012; date of current version February 12, 2013. This work was supported in part by the German Ministry of Education and Research under Grant 01BQ1050 and in part by the Deutsche Forschungsgemeinschaft Communication in Interference Limited Networks Project. This paper was presented in part at the 2011 IEEE International Symposium on Information Theory.

The authors are with the Lehrstuhl für Theoretische Informationstechnik, Technische Universität München, Munich 80333, Germany (e-mail: wiese@tum.de; boche@tum.de).

Communicated by Y. Steinberg, Associate Editor At Large.

Digital Object Identifier 10.1109/TIT.2012.2229779

We present the main theorems and several auxiliary coding results. The direct parts of the random and deterministic coding theorems are solved in Section III. Section IV gives the converses of the random and deterministic AV-MAC coding theorems. Section V concludes this paper with a discussion. In particular, the gains of conferencing are analyzed there.

A. Notation

In the information-theoretic setting, we also use the terms "encoders" for the senders and "decoder" for the receiver. For any positive integer m, we write $[1,m]$ for the set $\{1,\dots,m\}$. For a set $A \subset \mathcal{X}$, we denote its complement by $A^c := \mathcal{X}\backslash A$. For real numbers x and y, we set $x \wedge y := \min(x,y)$ and $x \vee y := \max\{x,y\}$. $\mathcal{P}(\mathcal{X})$ denotes the set of probability measures on the discrete set $\mathcal{X}$.

II. Problem Setting

A. Main Coding Problems

Let $\mathcal{X},\mathcal{Y},\mathcal{Z}$ be finite alphabets, and let $\mathcal{S}$ be another finite set. For every $s \in \mathcal{S}$, let a stochastic matrix

$$W(z|x,y|s): \quad (x,y,z) \in \mathcal{X}\times\mathcal{Y}\times\mathcal{Z}$$

be given with inputs from $\mathcal{X}\times\mathcal{Y}$ and outputs from $\mathcal{Z}$. $\mathcal{S}$ is to be interpreted as the set of channel states. We set

$$\mathcal{W} := \{W(\,\cdot\,|\,\cdot,\,\cdot\,|s) : s \in \mathcal{S}\}.$$

We assume that the channel state varies arbitrarily from channel use to channel use. Given words $\mathbf{x} = (x_1,\dots,x_n) \in \mathcal{X}^n$, $\mathbf{y} = (y_1,\dots,y_n) \in \mathcal{Y}^n$, and $\mathbf{z} = (z_1,\dots,z_n) \in \mathcal{Z}^n$, the probability that $\mathbf{z}$ is received upon transmission of $\mathbf{x}$ and $\mathbf{y}$ depends on the sequence $\mathbf{s} \in \mathcal{S}^n$ of channel states attained during the transmission. It equals

$$W^n(\mathbf{z}|\mathbf{x},\mathbf{y}|\mathbf{s}) := \prod_{m=1}^{n} W(z_m|x_m,y_m|s_m). \tag{1}$$

Definition 1: The set of stochastic matrices

$$\{W^n(\,\cdot\,|\,\cdot,\,\cdot\,|\mathbf{s}) : \mathbf{s} \in \mathcal{S}^n, n = 1,2,\dots\}$$

is called the AV-MAC determined by $\mathcal{W}$.

In the traditional noncooperative encoding schemes used for MACs, none of the senders has any information about the other sender's message. The goal here is to characterize the capacity region of the AV-MAC achievable when limited information can be exchanged between the encoders. We use Willems conferencing for this exchange [19], [20]. If the encoders' message sets are $[1,M_1]$ and $[1,M_2]$, respectively, then this can be described as follows. Let positive integers V_1 and V_2 be given which can be written as products

$$V_\nu = V_{\nu,1}\cdots V_{\nu,I}$$

for some positive integer I which does not depend on ν. A pair (c_1,c_2) of Willems conferencing functions is determined in an iterative manner via sequences of functions $c_{1,1},\dots,c_{1,I}$ and $c_{2,1},\dots,c_{2,I}$. The function $c_{1,i}$ describes what encoder 1 tells the other encoder in the ith conferencing iteration given the knowledge accumulated so far at encoder 1. Thus, in general, using the notation

$$\bar{\nu} := \begin{cases} 1 & \text{if } \nu = 2 \\ 2 & \text{if } \nu = 1 \end{cases}$$

these functions satisfy for $\nu = 1,2$ and $i = 2,\dots,I$

$$\begin{aligned} c_{\nu,1} &: [1,M_\nu] \to [1,V_{\nu,1}] \\ c_{\nu,i} &: [1,M_\nu]\times[1,V_{\bar{\nu},1}]\times\ldots\times[1,V_{\bar{\nu},i-1}] \to [1,V_{\nu,i}]. \end{aligned}$$

These functions recursively define other functions

$$\begin{aligned} c^*_{\nu,1} &: [1,M_\nu] \to [1,V_{\nu,1}], \\ c^*_{\nu,i} &: [1,M_1]\times[1,M_2] \to [1,V_{\nu,i}] \end{aligned}$$

by

$$\begin{aligned} c^*_{1,1}(j) &= c_{1,1}(j) \\ c^*_{2,1}(k) &= c_{2,1}(k) \\ c^*_{1,i}(j,k) &= c_{1,i}\big(j, c^*_{2,1}(k),\dots,c^*_{2,i-1}(j,k)\big) \\ c^*_{2,i}(j,k) &= c_{2,i}\big(k, c^*_{1,1}(j),\dots,c^*_{1,i-1}(j,k)\big). \end{aligned}$$

Then, we set

$$c_\nu(j,k) := \big((c^*_{\nu,1}(j,k),\dots,c^*_{\nu,I}(j,k)\big).$$

Observe that given a message pair (j,k), the conferencing outcome $(c_1(j,k),c_2(j,k))$ is known at both transmitters. If all conferencing protocols were allowed, the encoders could inform each other precisely about their messages, so this would turn the MAC into a single-sender channel. Thus for nonnegative numbers C_1,C_2, if conferencing is used in a blocklength-n code, Willems introduces the restrictions

$$\frac{1}{n}\log V_\nu \le C_\nu. \tag{2}$$

C_1,C_2 are called the conferencing capacities. Having introduced Willems conferencing, we can now define the codes we are going to consider.

Definition 2:

1) Let n,M_1,M_2 be positive integers and $C_1,C_2 \ge 0$. A deterministic $\mathrm{code}_{\mathrm{CONF}}(n,M_1,M_2,C_1,C_2)$ with blocklength n, codelength pair (M_1,M_2), and conferencing capacities C_1,C_2 is given by functions c_1,c_2,f_1,f_2,Φ. Here, (c_1,c_2) is a Willems conferencing protocol satisfying (2). f_1,f_2 are the encoding functions

$$\begin{aligned} f_1 &: [1,M_1]\times[1,V_2] \to \mathcal{X}^n \\ f_2 &: [1,M_2]\times[1,V_1] \to \mathcal{Y}^n. \end{aligned}$$

The *decoding function* Φ is a function

$$\Phi : \mathcal{Z}^n \to [1,M_1]\times[1,M_2].$$

2) A random $\mathrm{code}_{\mathrm{CONF}}(n,M_1,M_2,C_1,C_2)$ with blocklength n, codelength pair (M_1,M_2), and confer-

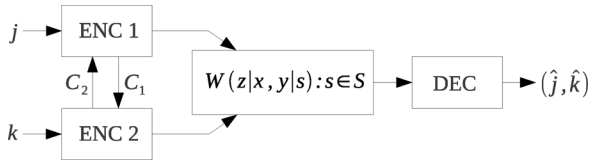

Fig. 1. AV-MAC $\mathcal{W}$ with conferencing encoders.

encing capacities C_1, C_2 is a pair (C, G), where $C = \{C(\gamma) : \gamma \in \Gamma\}$ is a finite family of deterministic $\text{codes}_{\text{CONF}}(n, M_1, M_2, C_1, C_2)$, and where G is a random variable taking values in Γ.

Note that a $\text{code}_{\text{CONF}}(n, M_1, M_2, 0, 0)$ is a traditional MAC code without conferencing. An AV-MAC together with the aforementioned coding procedure is called an AV-MAC with conferencing encoders (see Fig. 1). A $\text{code}_{\text{CONF}}$ (n, M_1, M_2, C_1, C_2) defined by $(c_1, c_2, f_1, f_2, \Phi)$ gives rise to a family

$$\{(\mathbf{x}_{jk}, \mathbf{y}_{jk}, F_{jk}) : (j,k) \in [1, M_1] \times [1, M_2]\} \tag{3}$$

where

$$\begin{aligned} \mathbf{x}_{jk} &:= f_1(j, c_2(j,k)) \in \mathcal{X}^n \\ \mathbf{y}_{jk} &:= f_2(k, c_1(j,k)) \in \mathcal{Y}^n \\ F_{jk} &:= \Phi^{-1}\{(j,k)\} \subset \mathcal{Z}^n. \end{aligned}$$

If the message pair (j,k) is present at the senders, the *codewords* $\mathbf{x}_{jk}$ and $\mathbf{y}_{jk}$ are sent. The decoding sets $\{F_{jk} : (j,k) \in [1, M_1] \times [1, M_2]\}$ give a partition of $\mathcal{Z}^n$ which, just like Φ, assigns to every channel output $\mathbf{z} \in \mathcal{Z}^n$ a message pair which the receiver will decide for upon reception of $\mathbf{z}$.

Note that every family (3), where the F_{jk} are disjoint, together with a Willems conferencing protocol (c_1, c_2) satisfying (2) defines a $\text{code}_{\text{CONF}}(n, M_1, M_2, C_1, C_2)$ if

$$\mathbf{x}_{jk} = \mathbf{x}_{jk'} \qquad \text{if } c_2(j,k) = c_2(j,k') \tag{4}$$

$$\mathbf{y}_{jk} = \mathbf{y}_{j'k} \qquad \text{if } c_1(j,k) = c_1(j',k). \tag{5}$$

Thus, a $\text{code}_{\text{CONF}}$ can equivalently be defined by a family (3) together with a conferencing protocol (c_1, c_2) such that (4) and (5) are satisfied. We will often refer to a $\text{code}_{\text{CONF}}$ using the description (3), and usually without specifying the corresponding conferencing protocol by just assuming that there is one.

The first example of this convention is encountered in our definition of the average error, where the explicit form of the conferencing protocol is irrelevant.

Definition 3:

1) A $\text{code}_{\text{CONF}}(n, M_1, M_2, C_1, C_2)$ defining a family (3) has an average error probability less than $\lambda \in (0,1)$ if

$$\frac{1}{M_1 M_2} \sum_{j,k} W^n(F_{jk}^c | \mathbf{x}_{jk}, \mathbf{y}_{jk} | \mathbf{s}) \le \lambda \quad \text{for all } \mathbf{s} \in \mathcal{S}^n.$$

2) Let a random $\text{code}_{\text{CONF}}(n, M_1, M_2, C_1, C_2)$ with the form (C, G) be given. Assume that the deterministic $\text{code}_{\text{CONF}}$ $C(\gamma)$ has the form

$$\{(\mathbf{x}_{jk}^\gamma, \mathbf{y}_{jk}^\gamma, F_{jk}^\gamma) : (j,k) \in [1, M_1] \times [1, M_2]\}.$$

Then, for any $\mathbf{s} \in \mathcal{S}^n$, define

$$P_e(C(\gamma)|\mathbf{s}) := \frac{1}{M_1 M_2} \sum_{j,k} W^n((F_{jk}^\gamma)^c | \mathbf{x}_{jk}^\gamma, \mathbf{y}_{jk}^\gamma | \mathbf{s}) \tag{6}$$

to be the average error incurred by $C(\gamma)$ under channel conditions $\mathbf{s}$. Assume that G has distribution p_G. We say that the random $\text{code}_{\text{CONF}}$ defined by (C, G) has an average error smaller than $\lambda \in (0,1)$ if

$$\sum_{\gamma \in \Gamma} P_e(C(\gamma)|\mathbf{s}) p_G(\gamma) \le \lambda \qquad \text{for every } \mathbf{s} \in \mathcal{S}^n.$$

This means that uniformly for every interfering sequence, transmission using this code is reliable up to the average error level λ. The possible state sequences are not weighted by any probability measure. One can interpret this in a communication setting with an adversary who knows which words $\mathbf{x}, \mathbf{y}$ are input into the channel by the senders and then can choose any state sequence $\mathbf{s} \in \mathcal{S}^n$ in order to obstruct the transmission of $\mathbf{x}$ and $\mathbf{y}$. The goal of the encoders then is to enable reliable communication no matter what sequence $\mathbf{s}$ the bad guy might use.

The concept of achievability of a rate pair is the usual one except that conferencing $\text{codes}_{\text{CONF}}$ are allowed for code construction.

Definition 4: A rate pair (R_1, R_2) is achievable by the AV-MAC with conferencing encoders and conferencing capacities C_1, C_2 under deterministic/random coding if for every $\lambda \in (0,1)$ and for every $\varepsilon > 0$, for n sufficiently large, there is a deterministic/random $\text{code}_{\text{CONF}}(n, M_1, M_2, C_1, C_2)$ with

$$\frac{1}{n} \log M_\nu \ge R_\nu - \varepsilon \quad (\nu = 1, 2)$$

and with an average error smaller than λ. The set of achievable rates under deterministic/random coding is called the deterministic/random capacity region of the AV-MAC with conferencing encoders and conferencing capacities C_1, C_2 and denoted by $\mathcal{C}_d(\mathcal{S}, C_1, C_2)$ (for deterministic coding) and $\mathcal{C}_r(\mathcal{S}, C_1, C_2)$ (for random coding).

We can now formulate the coding problems which are at the center of this study.

Characterize the deterministic/random capacity regions $\mathcal{C}_d(\mathcal{S}, C_1, C_2)$ and $\mathcal{C}_r(\mathcal{S}, C_1, C_2)$ of the AV-MAC with conferencing capacities C_1, C_2.

Of course, the main focus is on the deterministic capacity region $\mathcal{C}_d(\mathcal{S}, C_1, C_2)$ as the random capacity region $\mathcal{C}_r(\mathcal{S}, C_1, C_2)$ requires common randomness shared at the encoders and the receiver. For both $\mathcal{C}_d(\mathcal{S}, C_1, C_2)$ and $\mathcal{C}_r(\mathcal{S}, C_1, C_2)$, we need to consider the convex hull $\overline{\mathcal{W}}$ of $\mathcal{W}$. It is parameterized by the set of probability distributions $\mathcal{P}(\mathcal{S})$ on $\mathcal{S}$, so one can regard $\mathcal{P}(\mathcal{S})$ as its "state space." The stochastic matrix from $\overline{\mathcal{W}}$ assigned to the "state" $q \in \mathcal{P}(\mathcal{S})$ is the matrix with inputs from $\mathcal{X} \times \mathcal{Y}$ and outputs from $\mathcal{Z}$ having the form

$$W(z|x,y|q) := \sum_{s \in \mathcal{S}} W(z|x,y|s) q(s), \qquad (x,y,z) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}.$$

We have $\mathcal{W} \subset \overline{\mathcal{W}}$ by identifying $s \in \mathcal{S}$ with the Dirac measure $\delta_s \in \mathcal{P}(\mathcal{S})$, so that $W(\cdot\,|\,\cdot,\cdot\,|s) = W(\cdot\,|\,\cdot,\cdot\,|\delta_s)$.

Next, we define a set of rates $\mathcal{C}^*(\mathcal{S}, C_1, C_2)$. Let Π be the set consisting of probability distributions $p \in \mathcal{P}(\mathcal{U} \times \mathcal{X} \times \mathcal{Y})$, where $\mathcal{U}$ ranges over the finite subsets of the integers and where p has the form

$$p(u,x,y) = p_0(u)p_1(x|u)p_2(y|u).$$

To each $p \in \Pi$ and $q \in \mathcal{P}(\mathcal{S})$, one can associate a generic random vector (U, X, Y, Z_q) with distribution

$$p_q(u,x,y,z) = p(u,x,y)W(z|x,y|q). \tag{7}$$

In this way, every $p \in \Pi$ and $q \in \mathcal{P}(\mathcal{S})$ define a set $\mathcal{R}(p,q,C_1,C_2)$ consisting of those pairs (R_1, R_2) of nonnegative real numbers which satisfy

$$\begin{aligned} R_1 &\le I(Z_q; X|Y,U) + C_1 \\ R_2 &\le I(Z_q; Y|X,U) + C_2 \\ R_1 + R_2 &\le (I(Z_q; X,Y|U) + C_1 + C_2) \wedge I(Z_q; X, Y). \end{aligned}$$

Then, set

$$\mathcal{C}^*(\mathcal{S}, C_1, C_2) := \bigcup_{p \in \Pi} \bigcap_{q \in \mathcal{P}(\mathcal{S})} \mathcal{R}(p,q,C_1,C_2).$$

Theorem 5: For the AV-MAC determined by $\mathcal{W}$ with conferencing capacities $C_1, C_2 \ge 0$, we have

$$\mathcal{C}_r(\mathcal{S}, C_1, C_2) = \mathcal{C}^*(\mathcal{S}, C_1, C_2).$$

More precisely, for every $(R_1, R_2) \in \mathcal{C}^*(\mathcal{S}, C_1, C_2)$ and every $\varepsilon > 0$, there is a $\zeta > 0$ and a sequence (C_n, G_n) of random $\mathrm{codes}_{\mathrm{CONF}}(n, M_1^{(n)}, M_2^{(n)}, C_1, C_2)$ with an average error at most $2^{-n\zeta}$ such that

$$\frac{1}{n} \log M_\nu^{(n)} \ge R_\nu - \varepsilon \quad (\nu = 1, 2).$$

Additionally, the (C_n, G_n) can be chosen such that for every n, the constituent deterministic $\mathrm{codes}_{\mathrm{CONF}}$ share the same noniterative Willems conferencing protocol $(c_1^{(n)}, c_2^{(n)})$ given by

$$c_\nu^{(n)} : [1, M_\nu^{(n)}] \to [1, V_\nu^{(n)}] \qquad (\nu = 1, 2). \tag{8}$$

Remark 1: The simple form (8) of conferencing means that no complicated conferencing protocol needs to be designed.

Remark 2: $\mathcal{C}^*(\mathcal{S}, C_1, C_2)$ was analyzed in [17]. It is convex and the auxiliary sets $\mathcal{U}$ can be restricted to have cardinality at most $(|\mathcal{X}||\mathcal{Y}| + 2) \wedge (|\mathcal{Z}| + 3)$. Moreover, one can determine finite C_1, C_2 such that

1) the full-cooperation sum rate, or
2) the full-cooperation capacity region

are achievable. The first statement can be phrased as

$$\begin{aligned} &\max_{p \in \Pi} \min_{q \in \mathcal{P}(\mathcal{S})} \big\{ (I(Z_q; X, Y|U) + C_1 + C_2) \wedge I(Z_q; X, Y) \big\} \\ &\quad = \max_{p \in \Pi} \min_{q \in \mathcal{P}(\mathcal{S})} I(Z_q; X, Y) =: C_\infty. \end{aligned}$$

If for $p \in \Pi$ we set $\mathcal{Q}_p := \{q \in \mathcal{P}(\mathcal{S}) : I(Z_q \wedge X, Y) = C_\infty\}$, then a simple calculation shows that the aforementioned condition is satisfied if

$$C_1 + C_2 \ge C_\infty - \min_{p \in \Pi} \max_{q \in \mathcal{Q}_p} I(Z \wedge X, Y|U).$$

Statement 2) is valid if both $R_1 = C_\infty$ and $R_2 = C_\infty$ are possible. For this, one needs

$$\begin{aligned} C_1 &\ge C_\infty - \max_{p \in \Pi} \min_{q \in \mathcal{Q}_p} I(Z_q \wedge X|Y,U) \\ C_2 &\ge C_\infty - \max_{p \in \Pi} \min_{q \in \mathcal{Q}_p} I(Z_q \wedge Y|X,U). \end{aligned}$$

Remark 3: Theorem 5 has a weak converse.

Determining the general deterministic capacity region $\mathcal{C}_d(\mathcal{S}, C_1, C_2)$ is more complex. We give the solution in Theorem 7 below for $C_1 \vee C_2 > 0$. For $C_1 = C_2 = 0$, a partial solution is given in [4] and [10]–[12]. The relation between the cases $C_1 = C_2 = 0$ and $C_1 \vee C_2 > 0$ is discussed in detail in Section V.

Definition 6 ([11]):

1) $\mathcal{W}$ is called $(\mathcal{X}, \mathcal{Y})$-*symmetrizable* if there is a stochastic matrix

$$\sigma(s|x,y): \quad (x,y,s) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{S}$$

such that for every $z \in \mathcal{Z}$ and $x, x' \in \mathcal{X}$ and $y, y' \in \mathcal{Y}$

$$\sum_s W(z|x,y|s)\sigma(s|x',y') = \sum_s W(z|x',y'|s)\sigma(s|x,y).$$

2) $\mathcal{W}$ is called $\mathcal{X}$-*symmetrizable* if there is a stochastic matrix

$$\sigma_1(s|x): \quad (x,s) \in \mathcal{X} \times \mathcal{S}$$

such that for every $z \in \mathcal{Z}$ and $x, x' \in \mathcal{X}$ and $y \in \mathcal{Y}$

$$\sum_s W(z|x,y|s)\sigma_1(s|x') = \sum_s W(z|x',y|s)\sigma_1(s|x).$$

3) $\mathcal{W}$ is called $\mathcal{Y}$-*symmetrizable* if there is a stochastic matrix

$$\sigma_2(s|y): \quad (y,s) \in \mathcal{Y} \times \mathcal{S}$$

such that for every $z \in \mathcal{Z}$ and $x \in \mathcal{X}$ and $y, y' \in \mathcal{Y}$

$$\sum_s W(z|x,y|s)\sigma_2(s|y') = \sum_s W(z|x,y'|s)\sigma_2(s|y).$$

Theorem 7: For the deterministic capacity region of the AV-MAC determined by $\mathcal{W}$ with conferencing capacities $C_1 \vee C_2 > 0$, we have

$$\mathcal{C}_d(\mathcal{S}, C_1, C_2) = \mathcal{C}^*(\mathcal{S}, C_1, C_2)$$

if $\mathcal{W}$ is not $(\mathcal{X}, \mathcal{Y})$-symmetrizable and

$$\mathcal{C}_d(\mathcal{S}, C_1, C_2) = \{(0,0)\}$$

if $\mathcal{W}$ is $(\mathcal{X}, \mathcal{Y})$-symmetrizable. As for $\mathcal{C}_r(\mathcal{S}, C_1, C_2)$, the Willems conferencing protocols can be assumed to have the simple noniterative form (8).

Remark 4: If $\mathcal{W}$ is not $(\mathcal{X},\mathcal{Y})$-symmetrizable, then $\mathcal{C}_d(\mathcal{S},C_1,C_2)=\mathcal{C}^*(\mathcal{S},C_1,C_2)$ is at least 1-D. As $C_1\vee C_2>0$, in order to show this, it clearly suffices to check that

$$\max_{p\in\Pi}\min_{q\in\mathcal{P}(\mathcal{S})} I(Z_q;X,Y)>0 \tag{9}$$

if $\mathcal{W}$ is not symmetrizable. If (9) were violated, then by [7, Lemma 1.3.2] there would be a $q\in\mathcal{P}(\mathcal{S})$ such that

$$W(z|x,y|q)=W(z|x',y'|q)$$

for all $x,x'\in\mathcal{X},y,y'\in\mathcal{Y},z\in\mathcal{Z}$. Thus, $\mathcal{W}$ would be $(\mathcal{X},\mathcal{Y})$-symmetrizable using the stochastic matrix

$$\sigma(s|x,y)=q(s),\quad (x,y,s)\in\mathcal{X}\times\mathcal{Y}\times\mathcal{S}.$$

But this would contradict our assumption, so (9) must hold.

Remark 5: One can regard symmetrizability as the single-letterization of the adversary interpretation of the AV-MAC given previously. There, a complete input word pair has to be known to the adversary who can then choose the state sequence. In the definition of $(\mathcal{X},\mathcal{Y})$-symmetrizability, the stochastic matrix $\sigma:\mathcal{X}\to\mathcal{S}$ means that given a *letter* $x\in\mathcal{X}$, the adversary chooses a *random* state $s\in\mathcal{S}$. If $\mathcal{W}$ is $(\mathcal{X},\mathcal{Y})$-symmetrizable, the adversary can thus produce a useless single-state MAC $\tilde{W}:(\mathcal{X}\times\mathcal{Y})^2\to\mathcal{Z}$ defined by

$$\tilde{W}(z|x,y,x',y')=\sum_{s\in\mathcal{S}}W(z|x,y|s)\sigma(s|x',y').$$

$\tilde{W}$ is useless because it is symmetric in (x,y) and (x',y'). Thus, for word pairs $(\mathbf{x},\mathbf{y})$ and $(\mathbf{x}',\mathbf{y}')$, the receiver cannot decide which of the pairs was input into the channel by the senders and which was induced by the adversary's random state choice.

Remark 6: The aforementioned adversary interpretation of symmetrizability makes AV-MACs relevant for information-theoretic secrecy. Clearly, we do not say anything about the decodability of communication taking place in an AV-MAC for nonlegitimate listeners. However, reliable communication can be completely prevented in the case the AV-MAC is symmetrizable. A discussion of the single-sender arbitrarily varying wiretap channel can be found in [5].

Remark 7: By the definition of Willems conferencing, setting $C_1=C_2=0$ yields the traditional MAC coding, i.e., no conferencing at all is allowed. An inspection of the elimination technique applied in Section III-D shows that actually it suffices to have conferencing with $V_1=n^2$, so $C_1=(2\log n)/n$ (or, by symmetry, $V_2=n^2$). Using conferencing with this nonconstant rate tending to zero in non-$(\mathcal{X},\mathcal{Y})$-symmetrizable AV-MACs yields the capacity region $\mathcal{C}^*(\mathcal{S},0,0)$.

Remark 8: If $\mathcal{W}$ is $(\mathcal{X},\mathcal{Y})$-symmetrizable, then Theorem 7 almost has a strong converse: it is possible to show that every code that encodes more than one message incurs an average error at least 1/4. If $\mathcal{W}$ is not $(\mathcal{X},\mathcal{Y})$-symmetrizable, then we have a weak converse.

Theorem 7 does not carry over to the case $C_1=C_2=0$, which is the traditional AV-MAC with noncooperative coding. To our knowledge, the full characterization of the deterministic capacity region $\mathcal{C}_d(\mathcal{S},0,0)$ of the AV-MAC without cooperation is still open. We summarize here what has been found out in [4], [10], [11], and [12]. For notation, observe that

$$\begin{aligned}&\max_{p\in\Pi}\inf_{q\in\mathcal{P}(\mathcal{S})} I(Z_q;X|Y,U)\\ &=\max_{p\in\Pi}\inf_{q\in\mathcal{P}(\mathcal{S})} I(Z_q;X|Y)\\ &=\max_{y\in\mathcal{Y}}\max_{r\in\mathcal{P}(\mathcal{X})}\inf_{q\in\mathcal{P}(\mathcal{S})} I(Z_q;X|Y=y)\end{aligned}$$

where in the last term, the random vector (X,Z_q) has the distribution $r(x)W(z|x,y|q)$.

Theorem 8:

1) If $\mathcal{W}$ is neither $(\mathcal{X},\mathcal{Y})$- nor $\mathcal{X}$- nor $\mathcal{Y}$-symmetrizable, then $\mathcal{C}_d(\mathcal{S},0,0)=\mathcal{C}^*(\mathcal{S},0,0)$ and $\mathcal{C}^*(\mathcal{S},0,0)$ has nonempty interior.
2) If $\mathcal{W}$ is neither $(\mathcal{X},\mathcal{Y})$- nor $\mathcal{X}$-symmetrizable, but $\mathcal{Y}$-symmetrizable, then
$$\mathcal{C}_d(\mathcal{S},0,0)\subset[0,\max_{y\in\mathcal{Y}}\max_{r\in\mathcal{P}(\mathcal{X})}\inf_{q\in\mathcal{P}(\mathcal{S})} I(Z_q;X|Y=y)]\times\{0\}.$$
3) If $\mathcal{W}$ is neither $(\mathcal{X},\mathcal{Y})$- nor $\mathcal{Y}$-symmetrizable, but $\mathcal{X}$-symmetrizable, then
$$\mathcal{C}_d(\mathcal{S},0,0)\subset\{0\}\times[0,\max_{x\in\mathcal{X}}\max_{r\in\mathcal{P}(\mathcal{Y})}\inf_{q\in\mathcal{P}(\mathcal{S})} I(Z_q;Y|X=x)].$$
4) If $\mathcal{W}$ is $(\mathcal{X},\mathcal{Y})$-symmetrizable, then $\mathcal{C}_d(\mathcal{S},0,0)=\{(0,0)\}$.

In particular, if $\mathcal{W}$ is both $\mathcal{X}$- and $\mathcal{Y}$-symmetrizable, then $\mathcal{C}_d(\mathcal{S},0,0)=\{(0,0)\}$.

Remark 9: Point 1) from Theorem 8 is due to [4] and [12]. The other points are due to [10] and [11]. The precise characterization of $\mathcal{C}_d(\mathcal{S},0,0)$ in points 2) and 3) is still open.

Remark 10: The relation between the three kinds of symmetrizability from Definition 6 is treated in Section V. There, we provide the example of an AV-MAC which is both $\mathcal{X}$- and $\mathcal{Y}$-symmetrizable but not $(\mathcal{X},\mathcal{Y})$-symmetrizable.

B. Related Coding Results

The set $\overline{\mathcal{W}}$ also determines a compound MAC. This channel differs from the AV-MAC in that it does not change its state during the transmission of a codeword, only constant state sequences are possible. Thus, the probability that $\mathbf{z}\in\mathcal{Z}^n$ is received given the transmission of words $\mathbf{x}\in\mathcal{X}^n,\mathbf{y}\in\mathcal{Y}^n$ only depends on the given state $q\in\mathcal{P}(\mathcal{S})$. It equals

$$W^n(\mathbf{z}|\mathbf{x},\mathbf{y}|\mathbf{q}_q):=\prod_{m=1}^n W(z_m|x_m,y_m|q) \tag{10}$$

where we denote elements of $\mathcal{P}(\mathcal{S})^n$ by $\mathbf{q}$ and set $\mathbf{q}_q:=(q,\ldots,q)\in\mathcal{P}(\mathcal{S})^n$.

Definition 9: The set of stochastic matrices

$$\{W^n(\cdot\,|\,\cdot,\cdot|\mathbf{q}_q):q\in\mathcal{P}(\mathcal{S}),n=1,2,\ldots\}$$

is called the compound MAC determined by $\overline{\mathcal{W}}$.

One uses the same deterministic $\mathrm{codes}_{\mathrm{CONF}}$ as for the AV-MAC. Let

$$\{(\mathbf{x}_{jk}, \mathbf{y}_{jk}, F_{jk}) : (j,k) \in [1, M_1] \times [1, M_2]\}$$

be such a $\mathrm{code}_{\mathrm{CONF}}$. It has an average error less than λ for the compound MAC determined by $\overline{\mathcal{W}}$ if

$$\frac{1}{M_1 M_2} \sum_{j,k} W^n(F_{jk}^c | \mathbf{x}_{jk}, \mathbf{y}_{jk} | \mathbf{q}_q) \leq \lambda$$

for every $q \in \mathcal{P}(\mathcal{S})$. Using this average error criterion, the concept of achievability and the definition of the capacity region is analogous to that for deterministic coding for AV-MACs.

Remark 11: Comparing the error criteria for the AV-MAC and the compound MAC from the adversary perspective, one observes that the AV-MAC yields a significantly more robust performance. Theorem 7 describes the region achievable if transmission is reliable for every possible sequence the adversary might choose, whereas Theorem 10 describes the region which is achievable if the adversary is restricted to constant state sequences.

In [17], the following theorem was proved.

Theorem 10: The capacity region of the compound MAC determined by $\overline{\mathcal{W}}$ with conferencing capacities $C_1, C_2 \geq 0$ equals $\mathcal{C}^*(\mathcal{S}, C_1, C_2)$. More precisely, for every achievable rate pair $(R_1, R_2) \in \mathcal{C}^*(\mathcal{S}, C_1, C_2)$ and every $\varepsilon > 0$, there is a $\zeta > 0$ and a sequence of $\mathrm{codes}_{\mathrm{CONF}}$ $(n, M_1^{(n)}, M_2^{(n)}, C_1, C_2)$ with an average error at most $2^{-n\zeta}$ and

$$\frac{1}{n} \log M_\nu^{(n)} \geq R_\nu - \varepsilon \quad (\nu = 1, 2).$$

These $\mathrm{codes}_{\mathrm{CONF}}$ can be chosen such that their conferencing protocols have the form (8).

Finally, we have to recall the definition and a corollary of the deterministic coding result for single-user AVCs. Let $\mathcal{A}$ be a finite input alphabet, $\mathcal{B}$ a finite output alphabet, and $\mathcal{S}$ a finite state set. Let a family

$$\mathcal{H} := \{H(\cdot | \cdot | s) : s \in \mathcal{S})\}$$

of stochastic matrices

$$H(b|a|s) : \quad (a,b) \in \mathcal{A} \times \mathcal{B}$$

be given. As for AV-MACs, every state sequence $\mathbf{s} = (s_1, \ldots, s_n) \in \mathcal{S}^n$ determines a new stochastic matrix

$$H^n(\mathbf{b}|\mathbf{a}|\mathbf{s}) := \prod_{m=1}^{n} H(b_m|a_m|s_m) : \quad (\mathbf{a}, \mathbf{b}) \in \mathcal{A}^n \times \mathcal{B}^n.$$

Definition 11: The set of stochastic matrices

$$\{H^n(\cdot | \cdot | \mathbf{s}) : \mathbf{s} \in \mathcal{S}^n, n = 1, 2, \ldots\}$$

is called the AVC determined by $\mathcal{H}$.

The admissible codes are classical single-user codes as used for discrete memoryless channels. If such a code with blocklength n and codelength M has the form

$$\{(\mathbf{a}_\ell, F_\ell) : \ell \in [1, M]\}$$

then the average error incurred by this code is smaller than $\lambda \in (0,1)$ if

$$\frac{1}{M} \sum_{\ell} H^n(F_\ell^c | \mathbf{a}_\ell | \mathbf{s}) \leq \lambda \qquad \text{for all } \mathbf{s} \in \mathcal{S}^n.$$

Then, it is obvious what is meant by "achievable rates" and "capacity" for $\mathcal{H}$. The capacity of single-user AVCs, which was determined in [8], exhibits a dichotomy similar to the one claimed in Theorem 7. It is described by the original symmetrizability concept from [9].

Definition 12: $\mathcal{H}$ is called *symmetrizable* if there is a stochastic matrix

$$\sigma(s|a) : \quad (a,s) \in \mathcal{A} \times \mathcal{S}$$

such that for every $b \in \mathcal{B}$ and $a, a' \in \mathcal{A}$

$$\sum_s H(b|a|s)\sigma(s|a') = \sum_s H(b|a'|s)\sigma(s|a).$$

Remark 12: Clearly, the $(\mathcal{X}, \mathcal{Y})$-symmetrizability of the MAC $\mathcal{W}$ means nothing but symmetrizability of $\mathcal{W}$ when considered as a set of stochastic matrices with inputs from the alphabet $\mathcal{A} = \mathcal{X} \times \mathcal{Y}$.

Theorem 13 ([8], Th. 1): The deterministic capacity of the single-user AVC determined by $\mathcal{H}$ is positive if and only if $\mathcal{H}$ is not symmetrizable. If $\mathcal{H}$ is symmetrizable, then every code with at least two codewords incurs an average error at least 1/4.

III. Direct Parts

We derive the direct part of Theorem 5 from Theorem 10 in Sections III-A and III-B. Then, if $\mathcal{W}$ is not $(\mathcal{X}, \mathcal{Y})$-symmetrizable, we derandomize in Sections III-C and III-D to obtain the direct part of Theorem 7.

A. From Compound to Arbitrarily Varying

Here, we prove the direct part of Theorem 5. We use Ahlswede's "robustification lemma." Let S_n be the symmetric group (the group of permutations) on the set $[1, n]$. S_n operates on $\mathcal{S}^n$ by $\pi(\mathbf{s}) := (s_{\pi(1)}, \ldots, s_{\pi(n)})$ for any $\pi \in S_n$ and $\mathbf{s} = (s_1, \ldots, s_n) \in \mathcal{S}^n$. Further recall the notation $\mathbf{q}_q$ defined in (10).

Lemma 14 ([3, Lemma RT]): If $h : \mathcal{S}^n \to [0,1]$ satisfies for a $\lambda \in (0,1)$ and for all $q \in \mathcal{P}(\mathcal{S})$ the inequality

$$\sum_{\mathbf{s} \in \mathcal{S}^n} h(\mathbf{s}) \mathbf{q}_q(\mathbf{s}) \geq 1 - \lambda \tag{11}$$

then it also satisfies the inequality

$$\frac{1}{n!}\sum_{\pi\in S_n} h(\pi(\mathbf{s})) \geq 1-(n+1)^{|\mathcal{S}|}\lambda \qquad \text{for all } \mathbf{s}\in\mathcal{S}^n.$$

Now let $(R_1,R_2)\in\mathcal{C}^*(\mathcal{S},C_1,C_2)$. Theorem 10 states that for any $\varepsilon>0$, there is a $\zeta>0$ such that for sufficiently large n there is a $\mathrm{code}_{\mathrm{CONF}}(n,M_1,M_2,C_1,C_2)$ with an average error at most $2^{-n\zeta}$ and satisfying

$$\frac{1}{n}\log M_\nu \geq R_\nu-\varepsilon \quad (\nu=1,2).$$

Writing this $\mathrm{code}_{\mathrm{CONF}}$ in the form (3), this means for every $q\in\mathcal{P}(\mathcal{S})$ that

$$\frac{1}{M_1M_2}\sum_{j,k} W^n(F_{jk}|\mathbf{x}_{jk},\mathbf{y}_{jk}|\mathbf{q}_q)\geq 1-2^{-n\zeta}. \tag{12}$$

We would like to apply Lemma 14 with $\lambda=2^{-n\zeta}$ to the function $h:\mathcal{S}^n\to[0,1]$ defined by

$$h(\mathbf{s}):=\frac{1}{M_1M_2}\sum_{j,k} W^n(F_{jk}|\mathbf{x}_{jk},\mathbf{y}_{jk}|\mathbf{s}).$$

Thus, we need to show that h satisfies (11). Let $q\in\mathcal{P}(\mathcal{S})$. By (12), one obtains

$$\begin{aligned}
&\sum_{\mathbf{s}\in\mathcal{S}^n} h(\mathbf{s})q^n(\mathbf{s})\\
&=\frac{1}{M_1M_2}\sum_{j,k}\sum_{\mathbf{z}\in F_{jk}}\sum_{\mathbf{s}\in\mathcal{S}^n} W^n(\mathbf{z}|\mathbf{x}_{jk},\mathbf{y}_{jk}|\mathbf{s})\mathbf{q}_q(\mathbf{s})\\
&=\frac{1}{M_1M_2}\sum_{j,k}\sum_{\mathbf{z}\in F_{jk}} W^n(\mathbf{z}|\mathbf{x}_{jk},\mathbf{y}_{jk}|\mathbf{q}_q)\\
&\geq 1-2^{-n\zeta}
\end{aligned}$$

and (11) is satisfied. Applying Lemma 14, one obtains

$$\frac{1}{n!}\sum_{\pi\in S_n} h(\pi(\mathbf{s}))\geq 1-(n+1)^{|\mathcal{S}|}2^{-n\zeta} \qquad \text{for all } \mathbf{s}\in\mathcal{S}^n. \tag{13}$$

Recall that π^{-1} also is an element of S_n. Writing $\pi^{-1}(F_{jk})=\{\pi^{-1}(\mathbf{z}):\mathbf{z}\in F_{jk}\}$, the left-hand side of (13) equals

$$\begin{aligned}
&\frac{1}{n!}\sum_{\pi\in S_n}\left(\frac{1}{M_1M_2}\sum_{j,k} W^n(F_{jk}|\mathbf{x}_{jk},\mathbf{y}_{jk}|\pi(\mathbf{s}))\right)\\
&=\frac{1}{n!}\sum_{\pi\in S_n}\left(\frac{1}{M_1M_2}\times\right.\\
&\quad\left.\times\sum_{j,k} W^n\left(\pi^{-1}(F_{jk})|\pi^{-1}(\mathbf{x}_{jk}),\pi^{-1}(\mathbf{y}_{jk})|\mathbf{s}\right)\right). \qquad (14)
\end{aligned}$$

Because of the bijectivity of π^{-1}, the family of sets $\{\pi^{-1}(F_{jk}):(j,k)\in[1,M_1]\times[1,M_2]\}$ is disjoint. Thus, (14) is the average error expression of the random code (C,G) when applied in the AV-MAC determined by $\mathcal{W}$, where G is uniformly distributed on $\Gamma:=S_n$ and where for every $\pi\in S_n$

$$\begin{aligned}
C(\pi):=\{&\left(\pi^{-1}(\mathbf{x}_{jk}),\pi^{-1}(\mathbf{y}_{jk}),\pi^{-1}(F_{jk})\right):\\
&(j,k)\in[1,M_1]\times[1,M_2]\}.
\end{aligned}$$

The conferencing protocol remains the same for every $C(\pi)$, as the conference only concerns the messages and not the codewords. By (13), the average error of this random code is less than $(n+1)^{|\mathcal{S}|}2^{-n\zeta}$; hence, it tends to zero exponentially. Thus, $\mathcal{C}^*(\mathcal{S},C_1,C_2)\subset\mathcal{C}_r(\mathcal{S},C_1,C_2)$, which proves the direct part of Theorem 5.

B. Bounding the Amount of Correlation

As a first derandomization step to proving the direct part of Theorem 7, we have to show the following lemma.

Lemma 15: To every random $\mathrm{code}_{\mathrm{CONF}}(n,M_1,M_2,C_1,C_2)$ with average error at most λ which is given as a pair (C,G), there exists a random $\mathrm{code}_{\mathrm{CONF}}(n,M_1,M_2,C_1,C_2)$ with an average error smaller than 3λ given as a pair (C',G') where $C'\subset C$ and $|C'|=n^2$ and where G' is uniformly distributed on $[1,n^2]$.

For the proof of Lemma 15, we need a simple result from [12, Sec. IV].

Lemma 16: Let N i.i.d. random variables $T_1,\ldots,T_N$ with values in $[0,1]$ and underlying probability measure $\mathbb{P}$ be given. Let $\bar{\lambda}>0$. Denote by $\mathbb{E}$ the expectation corresponding to $\mathbb{P}$. Then

$$\mathbb{P}\left[\frac{1}{N}\sum_{m=1}^{N} T_m>\bar{\lambda}\right]\leq\exp\left(-(\bar{\lambda}-e\mathbb{E}[T_1])N\right).$$

Proof of Lemma 15: Let a random $\mathrm{code}_{\mathrm{CONF}}$ (C,G) with blocklength n and average error smaller than λ. Recalling our notation (6), the fact that (C,G) has an average error less than λ can be stated as

$$\mathbb{E}[P_e(C(G)|\mathbf{s})]\leq\lambda \quad \text{for every } \mathbf{s}\in\mathcal{S}^n.$$

Let $G_1,\ldots,G_{n^2}$ be independent copies of G. This induces a family of n^2 independent copies of (C,G). The goal is to show

$$\mathbb{P}\left[\frac{1}{n^2}\sum_{m=1}^{n^2} P_e(C(G_m)|\mathbf{s})\leq 3\lambda \text{ for all } \mathbf{s}\in\mathcal{S}^n\right]>0. \tag{15}$$

Given (15), there is a realization $(\gamma_1,\ldots,\gamma_{n^2})$ of $(G_1,\ldots,G_{n^2})$ such that

$$\frac{1}{n^2}\sum_{m=1}^{n^2} P_e(C(\gamma_m)|\mathbf{s})\leq 3\lambda \tag{16}$$

for every $\mathbf{s}\in\mathcal{S}^n$. Then, one defines a random code (C',G') by setting

$$C':=\{C(\gamma_m):m\in[1,n^2]\}$$

and by taking G' to be uniformly distributed on $[1, n^2]$. The expression (16) then is nothing but the statement that the average error of the random code (C', G') is smaller than 3λ, and we are done.

It remains to prove (15). $\mathcal{S}$ is finite by assumption, so $|\mathcal{S}^n|$ grows exponentially with blocklength. Hence, it suffices to show that

$$\mathbb{P}\left[\frac{1}{n^2}\sum_m P_e(C(G_m)|\mathbf{s}) > 3\lambda\right] \tag{17}$$

is superexponentially small uniformly in $\mathbf{s} \in \mathcal{S}^n$. Let us fix an $\mathbf{s} \in \mathcal{S}^n$. The G_m are i.i.d. copies of G, so by Lemma 16, the term (17) is smaller than

$$\exp\bigl(-(3\lambda - e\,\mathbb{E}[P_e(C(G)|\mathbf{s})])\,n^2\bigr). \tag{18}$$

By assumption

$$\mathbb{E}[P_e(C(G)|\mathbf{s})] \le \lambda$$

so the exponent in (18) is negative. This gives the desired superexponential bound on (17). ■

Remark 13: Note that we cannot require the $\mathrm{codes}_{\mathrm{CONF}}$ with at most n^2 values of G to have an exponentially small probability of error. This is due to the fact that the exponent in (18) must not decrease exponentially in order for the proof to work. Thus, there is a tradeoff between the error probability and the number of deterministic component $\mathrm{codes}_{\mathrm{CONF}}$ of the random $\mathrm{codes}_{\mathrm{CONF}}$ used to achieve the random capacity region of the AV-MAC with conferencing encoders.

C. Positive Rate

In the second derandomization step, we show that if $\mathcal{W}$ is not $(\mathcal{X}, \mathcal{Y})$-symmetrizable and $C_1 > 0$ or $C_2 > 0$, then the encoder with the positive conferencing capacity achieves a positive rate by deterministic coding. Without loss of generality, we may assume that $C_1 > 0$.

Theorem 17: Let $C_1 > 0$. If $\mathcal{W}$ is not $(\mathcal{X}, \mathcal{Y})$-symmetrizable, then there exists an R with $0 < R < C_1$ such that the rate pair $(R, 0) \in \mathcal{C}^*(\mathcal{S}, C_1, C_2)$ is deterministically achievable using $\mathrm{codes}_{\mathrm{CONF}}$ with conferencing capacity pair $(C_1, 0)$ such that the conferencing function c_1 is the identity on the message set.

Proof: By Remark 12 and Theorem 13, $\mathcal{W}$ considered as a single-user AVC with vector inputs from the alphabet $\mathcal{X} \times \mathcal{Y}$ has positive capacity. The idea of the proof is to construct from a code for this single-user AVC a $\mathrm{code}_{\mathrm{CONF}}$ such that the first transmitter achieves a positive rate. There is a positive rate $R < C_1$ which is deterministically achievable by the single-user AVC determined by $\mathcal{W}$. This means that for every $\lambda \in (0, 1)$ and every $\varepsilon > 0$, for n large enough, there is a single-user code

$$\{(\mathbf{x}_\ell, \mathbf{y}_\ell, F_\ell) : \ell \in [1, M_1]\}$$

for $\mathcal{W}$ with

$$2^{n(R-\varepsilon/2)} \le M_1 \le 2^{nR}$$

and with

$$\frac{1}{M_1}\sum_{\ell=1}^{M_1} W^n(F_\ell^c|\mathbf{x}_\ell, \mathbf{y}_\ell|\mathbf{s}) \le \lambda \qquad \text{for all } \mathbf{s} \in \mathcal{S}^n.$$

By setting c_1 to be the identity on $[1, M_1]$, this code becomes a $\mathrm{code}_{\mathrm{CONF}}(n, M_1, 1, C_1, 0)$. This is allowed because $\log M_1 \le nR \le nC_1$. The encoding and decoding functions are defined in the obvious way. Thus, the positive rate pair $(R, 0)$ is achievable. ■

D. From Random to Deterministic

Finally, we can show that if $\mathcal{W}$ is not $(\mathcal{X}, \mathcal{Y})$-symmetrizable, then $\mathcal{C}^*(\mathcal{S}, C_1, C_2) \subset \mathcal{C}_d(\mathcal{S}, C_1, C_2)$. To do so, we follow Ahlswede's "Elimination Technique" [1], whose idea is to use random codes and to replace the randomness needed there by a prefix code with small blocklength which encodes the set of constituent deterministic codes. We again assume that $C_1 > 0$.

Theorem 17 implies that there is a $0 < R < C_1$ such that for any $\varepsilon \in (0, R)$ and any $\lambda \in (0, 1)$, if n is large, there is a $\mathrm{code}_{\mathrm{CONF}}$

$$\{(\mathbf{x}_\gamma^*, \mathbf{y}_\gamma^*, F_\gamma^*) : \gamma \in [1, n^2]\}$$

with blocklength m

$$\frac{2}{R}\log n \le m \le \frac{2}{R-\varepsilon}\log n \tag{19}$$

with codelength pair $(n^2, 0)$ and with average error smaller than λ. Further, the conferencing function c_1^* is the identity on the set $[1, n^2]$.

For any $0 < \delta < C_1$, let $(R_1, R_2) \in \mathcal{R}(p, q, C_1 - \delta, C_2)$ for some $p \in \Pi$ and all $q \in \mathcal{P}(\mathcal{S})$. By Theorem 5, this rate pair is achievable with conferencing capacities $C_1 - \delta$ and C_2 under random coding. For every $\varepsilon > 0$, $\lambda \in (0, 1)$, and large n, this implies the existence of a random $\mathrm{code}_{\mathrm{CONF}}(n, M_1, M_2, C_1 - \delta, C_2)$ defined by a pair (C, G) such that

$$\frac{1}{n}\log M_\nu \ge R_\nu - \varepsilon$$

and with an average error smaller than λ. Again by Theorem 5, we may further assume that the deterministic component $\mathrm{codes}_{\mathrm{CONF}}$ $C(\gamma)$ of C share the same conferencing protocol (c_1, c_2). As we do not need an exponential decrease of the average error, we may by Lemma 15 assume that G is uniformly distributed on $\Gamma = [1, n^2]$. Let every deterministic component $\mathrm{code}_{\mathrm{CONF}}$ $C(\gamma)$ be given as

$$\{(\mathbf{x}_{jk}^\gamma, \mathbf{y}_{jk}^\gamma, F_{jk}^\gamma) : (j, k) \in [1, M_1] \times [1, M_2]\}.$$

We now construct a deterministic $\mathrm{code}_{\mathrm{CONF}}$ $(\tilde{c}_1, \tilde{c}_2, \tilde{f}_1, \tilde{f}_2, \tilde{\Phi})$ with blocklength $m + n$, message sets $[1, n^2] \times [1, M_1]$ and $[1, M_2]$ (yielding the codelength pair $(n^2 M_1, M_2)$), conferencing capacities C_1, C_2, and average error smaller than 2λ. It is defined via concatenation. We define the conferencing functions to be

$$\begin{aligned}\tilde{c}_1(\gamma, j) &:= (\gamma, c_1(j)) \in [1, n^2] \times [1, V_1]\\ \tilde{c}_2(k) &:= c_2(k) \in [1, V_2].\end{aligned}$$

Note that $(\tilde{c}_1, \tilde{c}_2)$ has the form (8). It is a permissible conferencing protocol if

$$\frac{1}{2\,R^{-1}\log n + n}\log n \le \delta$$

because then

$$\frac{1}{m+n}\log(n^2 V_1) \le \frac{1}{2\,R^{-1}\log n + n}\log n + \frac{1}{n}\log V_1 \le C_1.$$

If the encoders have the messages (γ, j) and k, respectively, they use the codewords

$$\left(\mathbf{x}^*_\gamma, \mathbf{x}^\gamma_{jk}\right) \in \mathcal{X}^{m+n} \quad \text{and} \quad \left(\mathbf{y}^*_\gamma, \mathbf{y}^\gamma_{jk}\right) \in \mathcal{Y}^{m+n}.$$

Together with the conferencing protocol $(\tilde{c}_1, \tilde{c}_2)$ defined previously, this fixes encoding functions f_1 and f_2, as (4) and (5) are satisfied. The decoding set of the $\mathrm{code}_{\mathrm{CONF}}$ deciding for the pair $\big((\gamma, j), k\big)$ is defined to be $F^*_\gamma \times F^\gamma_{jk} \subset \mathcal{Z}^{m+n}$. Thus, the deterministic $\mathrm{code}_{\mathrm{CONF}}$ achieving the rate pair $(R, 0)$ is used as a prefix code which distinguishes the deterministic component $\mathrm{codes}_{\mathrm{CONF}}$ of the random $\mathrm{code}_{\mathrm{CONF}}$. In this way, derandomization can be seen as a two-step protocol. Setting $a := 2/(R-\varepsilon)$, the rates of the new code are

$$\frac{1}{m+n}\log(nM_\nu) \ge \frac{1}{a\frac{\log n}{n}+1}\cdot\frac{1}{n}\log M_\nu \ge R_\nu - 2\varepsilon$$

where the second inequality holds for all n large enough such that

$$\frac{1}{a\frac{\log n}{n}+1} \ge \frac{R_\nu - 2\varepsilon}{R_\nu - \varepsilon}.$$

The randomness of the random code is needed in the estimation of the average error incurred by this coding procedure. Recall Ahlswede's Innerproduct Lemma [1].

Lemma 18: Let $(\alpha_1, \ldots, \alpha_N)$ and $(\beta_1, \ldots, \beta_N)$ be two vectors with $0 \le \alpha_m, \beta_m \le 1$ for $m = 1, \ldots, N$ which for some $\lambda \in (0,1)$ satisfy

$$\frac{1}{N}\sum_{m=1}^{N}\beta_m \ge 1-\lambda, \qquad \frac{1}{N}\sum_{m=1}^{N}\alpha_m \ge 1-\lambda \tag{20}$$

then

$$\frac{1}{N}\sum_{m=1}^{N}\alpha_m\beta_m \ge 1-2\lambda.$$

We use this lemma with $N = n^2$ and replace the index m by $\gamma \in [1, n^2]$. Fix an $\mathbf{s} \in \mathcal{S}^n$ and set

$$\alpha_\gamma = W^m(F^*_\gamma|\mathbf{x}^*_\gamma, \mathbf{y}^*_\gamma|\mathbf{s})$$
$$\beta_\gamma = \frac{1}{M_1 M_2}\sum_{j,k} W^n\big(F^\gamma_{jk}|\mathbf{x}^\gamma_{jk}, \mathbf{y}^\gamma_{jk}|\mathbf{s}\big).$$

Then, the conditions in (20) are satisfied because both the deterministic prefix code $(c^*_1, c^*_2, f^*_1, f^*_2, \Phi^*)$ and the random code (C, G) with constituent codes $(c_1, c_2, f^\gamma_1, f^\gamma_2, \Phi^\gamma)$ have an average error smaller than λ. Lemma 18 now implies that the $\mathrm{code}_{\mathrm{CONF}}$ $(\tilde{c}_1, \tilde{c}_2, \tilde{f}_1, \tilde{f}_2, \tilde{\Phi})$ constructed previously has an average error probability smaller than 2λ.

This shows that the rate pair (R_1, R_2) is achievable for $\mathcal{W}$ with conferencing capacities C_1, C_2. Consequently, one obtains

$$\bigcup_{\delta>0}\mathcal{C}^*(\mathcal{S}, C_1-\delta, C_2) \subset \mathcal{C}_d(\mathcal{S}, C_1, C_2).$$

As the capacity region is closed, $\mathcal{C}^*(\mathcal{S}, C_1, C_2)$, which is the closure of the set on the left-hand side, is contained in $\mathcal{C}_d(\mathcal{S}, C_1, C_2)$ as well. This proves the direct part of Theorem 7.

IV. Converses for the AV-MAC With Conferencing Encoders

Here, we prove the converses claimed in Remark 3 and 8. Recall that a weak converse means that, depending on the situation, any deterministic or random $\mathrm{code}_{\mathrm{CONF}}(n, M_1, M_2, C_1, C_2)$ such that the real 2-D vector $((1/n)\log M_1, (1/n)\log M_2)$ is at least distance $\varepsilon > 0$ from the achievable rate region incurs an average error at least $\lambda(\varepsilon) > 0$ if n is large.

A. Random Coding

Here, we prove the weak converse for Theorem 5 (see Remark 3). The idea of the proof is to reduce it to the weak converse for the compound MAC with conferencing encoders defined by $\overline{\mathcal{W}}$ where the use of random codes is allowed. This is proved in the Appendix.

Every $\mathbf{q} \in \mathcal{P}(\mathcal{S})^n$ induces a product measure via $\mathbf{q}(\mathbf{s}) = q_1(s_1)\cdots q_n(s_n)$. The notation (10) carries over to these general $\mathbf{q}$. Further, recall the notation introduced in (6). We generalize this notation by setting

$$P_e(C(\gamma)|\mathbf{q}) := \frac{1}{M_1 M_2}\sum_{j,k} W^n((F^\gamma_{jk})^c|\mathbf{x}^\gamma_{jk}, \mathbf{y}^\gamma_{jk}|\mathbf{q}).$$

The following lemma is a generalized version of Lemma 2.6.3 in [7].

Lemma 19: For any random $\mathrm{code}_{\mathrm{CONF}}$ which is defined by (C, G) and whose components have the form

$$\{(\mathbf{x}^\gamma_{jk}, \mathbf{y}^\gamma_{jk}, F^\gamma_{jk}) : (j,k) \in [1, M_1] \times [1, M_2]\}$$

one has

$$\sup_{\mathbf{s}\in\mathcal{S}^n}\sum_{\gamma\in\Gamma} P_e(C(\gamma)|\mathbf{s})p_G(\gamma) = \sup_{\mathbf{q}\in\mathcal{P}(\mathcal{S})^n}\sum_{\gamma\in\Gamma} P_e(C(\gamma)|\mathbf{q})p_G(\gamma).$$

Proof: The direction "$\le$" is clear. In order to prove "$\ge$," let $\mathbf{q} \in \mathcal{P}(\mathcal{S})^n$. Clearly

$$W^n(\mathbf{z}|\mathbf{x}, \mathbf{y}|\mathbf{q}) = \sum_{\mathbf{s}\in\mathcal{S}^n}\mathbf{q}(\mathbf{s})W^n(\mathbf{z}|\mathbf{x}, \mathbf{y}|\mathbf{s}).$$

Thus

$$\sum_{\gamma\in\Gamma} P_e(C(\gamma)|\mathbf{q})p_G(\gamma) = \sum_{\mathbf{s}\in\mathcal{S}^n}\mathbf{q}(\mathbf{s})\sum_{\gamma\in\Gamma} P_e(C(\gamma)|\mathbf{s})p_G(\gamma) \le \sup_{\mathbf{s}\in\mathcal{S}^n}\sum_{\gamma\in\Gamma} P_e(C(\gamma)|\mathbf{s})p_G(\gamma).$$

Upon taking the supremum over $\mathbf{q} \in \mathcal{P}(\mathcal{S})^n$ on the left-hand side, the lemma is proved. ■

Now let a random $\text{code}_{\text{CONF}}$ (n, M_1, M_2, C_1, C_2) be given defined by (C, G) and with average error at most λ. Assume that the pair $((1/n)\log M_1, (1/n)\log M_2)$ is at distance at least ε from $\mathcal{C}^*(\mathcal{S}, C_1, C_2)$. Because of Lemma 19

$$\lambda_0 := \sup_{q\in\mathcal{P}(\mathcal{S})} \sum_{\gamma\in\Gamma} P_e(C(\gamma)|\mathbf{q}_q) p_G(\gamma) \leq \lambda. \tag{21}$$

Thus, the random $\text{code}_{\text{CONF}}$ (C, G) has an average error at most λ for the compound MAC with conferencing encoders defined by $\overline{\mathcal{W}}$. But the weak converse for the compound MAC with conferencing encoders and random coding, which is proved in the Appendix, implies that (21) can only hold if $\lambda \geq \lambda_0 \geq \lambda(\varepsilon) > 0$. This concludes the weak converse for the AV-MAC with conferencing encoders using random $\text{codes}_{\text{CONF}}$, and Theorem 5 is proved.

B. Deterministic Coding

1) If $\mathcal{W}$ is $(\mathcal{X}, \mathcal{Y})$-Symmetrizable: If $\mathcal{W}$ is $(\mathcal{X}, \mathcal{Y})$-symmetrizable, then by Remark 12, it is also symmetrizable if considered as a single-user AVC with input alphabet $\mathcal{X} \times \mathcal{Y}$. Thus, Theorem 13 implies that any single-user code with at least two codewords incurs an average error greater than 1/4. Finally, note that every $\text{code}_{\text{CONF}}$ for the AV-MAC with conferencing encoders determined by $\mathcal{W}$ also is a code for the single-user AVC determined by $\mathcal{W}$, so this carries over to the multiuser situation. This proves Theorem 7 if $\mathcal{W}$ is $(\mathcal{X}, \mathcal{Y})$-symmetrizable.

2) If $\mathcal{W}$ is Not $(\mathcal{X}, \mathcal{Y})$-Symmetrizable: We show that the weak converse for the compound MAC determined by $\overline{\mathcal{W}}$ implies the weak converse for the AV-MAC determined by $\mathcal{W}$. Let a $\text{code}_{\text{CONF}}$ (n, M_1, M_2, C_1, C_2) be given. If the rate pair $((1/n)M_1, (1/n)M_2)$ is at least distance ε away from $\mathcal{C}^*(\mathcal{S}, C_1, C_2)$ and if n is sufficiently large, then there is a $q \in \mathcal{P}(\mathcal{S})$ such that

$$\frac{1}{M_1 M_2} \sum_{j,k} W^n(F_{jk}^c | \mathbf{x}_{jk}, \mathbf{y}_{jk} | \mathbf{q}_q) \geq \lambda(\varepsilon)$$

for some $\lambda(\varepsilon) > 0$ because of the weak converse for the compound MAC. Lemma 19 now implies that

$$\sup_{\mathbf{s}\in\mathcal{S}^n} \frac{1}{M_1 M_2} \sum_{j,k} W^n(F_{jk}^c | \mathbf{x}_{jk}, \mathbf{y}_{jk} | \mathbf{s}) \geq \lambda(\varepsilon)$$

must hold. Thus, the proof of Theorem 7 is complete.

V. Discussion and Conclusion

The goal of this paper was to characterize the capacity region of an AV-MAC whose encoders may exchange limited information about their messages. This topic is motivated by the increasing interest of cooperative networks which are subject to exterior interference. For example, spectrum sharing has been discussed for inclusion into future wireless system standards. We saw above that the AV-MAC can be interpreted as a channel suffering from attacks by an adversary who may choose the state sequence given the channel inputs. The reliability requirements for AV-MACs are very strict—coding is done such that the average error is small for every possible state sequence. The resulting capacity region is the same as that for the conferencing compound MAC determined by the convex hull of the set of channel matrices of the original AV-MAC *if* the latter is not $(\mathcal{X}, \mathcal{Y})$-symmetrizable. Otherwise, the AV-MAC is useless. In contrast, for AV-MACs without conferencing, the complete characterization of the deterministic capacity region is still open.

The dichotomy in the form of the deterministic capacity regions of AV-MACs does not occur if random coding is used. However, using a random code requires both the senders and the receiver to have access to a common source of randomness. Thus, random coding is usually only used as a mathematical tool for finding good deterministic codes. It is well known that derandomization is no problem for compound channels (including discrete memoryless channels as a special case)—this is nothing but the well-known "random coding method." It builds on the fact that the finite number of channel states does not increase with blocklength. This is not so in the case of AVCs. The number of states per channel use remains constant, but the number of states per transmission of a codeword increases *exponentially* in blocklength. This is the reason why the deterministic capacity region of AV-MACs may be strictly contained in the random coding capacity region. In fact, if derandomization is not possible, then no positive rates are achievable at all.

In contrast to the derandomization technique used for simpler channels, Ahlswede's elimination technique gives rise to a two-step protocol. In order to approximate a given achievable rate pair (R_1, R_2), one only needs the constituent deterministic codes of a random code whose rate pair approximates (R_1, R_2). The randomness of the random code is used in the average error estimate. (On the other hand, this shows how much weaker the average error criterion is compared with the maximal error requirement—the randomized part can be "hidden" in the average error.)

It is noteworthy that for the AV-MAC, the conferencing protocols needed to achieve any rate pair within the capacity region remain as simple as for the compound MAC with conferencing encoders. There are no iterative steps, so the implementation of such a conference is straightforward.

Finally, we would like to analyze the benefits of Willems conferencing. We compare the gains obtained in AV-MACs to the gains obtained in compound MACs. For both compound and AV-MACs, conferencing may help to achieve positive rates where only the rate pair $(0, 0)$ is achievable without transmitter cooperation. This effect is similar to the "superactivation" of quantum channels as observed in [16], where it was shown that there are pairs of quantum channels with zero quantum capacity each which achieve positive rates when used together.

Every compound MAC with conferencing capacities $C_1 \vee C_2 > 0$ and

$$\max_{p\in\Pi} \min_{q\in\mathcal{P}(\mathcal{S})} I(Z_q; X, Y) > 0 \tag{22}$$

has at least 1-D capacity region. If (22) is not satisfied, then the capacity region equals $\{(0, 0)\}$. No matter what dimension the corresponding $\mathcal{C}^*(\mathcal{S}, 0, 0)$ has, the gains of conferencing are continuous in C_1, C_2, in particular in $(C_1, C_2) = (0, 0)$. This is

in contrast to the AV-MAC. The changes in the deterministic capacity region $\mathcal{C}_d(\mathcal{S}, C_1, C_2)$ are continuous in all (C_1, C_2) with $C_1 \vee C_2 > 0$ because either $\mathcal{C}_d(\mathcal{S}, C_1, C_2) = \mathcal{C}^*(\mathcal{S}, C_1, C_2)$ for all (C_1, C_2) with $C_1 \vee C_2 > 0$ or $\mathcal{C}_d(\mathcal{S}, C_1, C_2) = \{(0,0)\}$ for all C_1, C_2. However, there may be a discontinuity in $(C_1, C_2) = (0,0)$.

This corresponds to the two roles conferencing plays in AV-MACs. The "traditional" role is to generate a common message and to use the coding result for the (compound) MAC with common message to enlarge the capacity region. For AV-MACs, it does even more—it changes the channel structure. Recall Remark 7. For a conferencing rate pair with $C_1 \vee C_2 = (2\log n)/n$, the capacity region of the compound MAC stays as it is. Under the conditions that $\mathcal{W}$ is not $(\mathcal{X}, \mathcal{Y})$-symmetrizable and $\mathcal{C}_d(\mathcal{S}, 0, 0) \neq \mathcal{C}^*(\mathcal{S}, 0, 0)$, though, we can strictly enlarge the capacity region of the AV-MAC with this kind of conferencing.

General conditions for $\mathcal{C}_d(\mathcal{S}, 0, 0) \neq \mathcal{C}^*(\mathcal{S}, 0, 0)$ to hold cannot be given because an exact characterization of $\mathcal{C}_d(\mathcal{S}, 0, 0)$ is generally unavailable. We certainly know by Theorem 8 that if $\mathcal{C}_d(\mathcal{S}, 0, 0)$ is 2-D, then $\mathcal{C}_d(\mathcal{S}, 0, 0) = \mathcal{C}^*(\mathcal{S}, 0, 0)$. We can further say that if in addition to not being $(\mathcal{X}, \mathcal{Y})$-symmetrizable, $\mathcal{W}$ is both $\mathcal{X}$- and $\mathcal{Y}$-symmetrizable, then $\mathcal{C}_d(\mathcal{S}, 0, 0) = \{(0,0)\}$, again by Theorem 8. This is a situation where already the conferencing from Remark 7 helps. With the same argumentation as in Remark 4, it can easily be seen that

$$\max_{p \in \Pi} \min_{q \in \mathcal{P}(\mathcal{S})} I(Z_q; X, Y|U) > 0$$

so $\mathcal{C}^*(\mathcal{S}, 0, 0)$ is at least 1-D. Thus, there is a discontinuity in $(C_1, C_2) = (0,0)$ in this case. Gubner [11] has found the example of a $\mathcal{W}$ which is both $\mathcal{X}$- and $\mathcal{Y}$-symmetrizable, but not $(\mathcal{X}, \mathcal{Y})$-symmetrizable.

Example 1: Let $\mathcal{X} = \mathcal{Y} = \mathcal{S} = \{0,1\}$ and $\mathcal{Z} = \{0,1,2,3\}$. For $s \in \mathcal{S}$, set

$$W(z|x, y|s) = \delta(z - x - y - s)$$

where $\delta(t) = 1$ if $t = 0$ and $\delta(t) = 0$ else. An equivalent description of this is

$$z = x + y + s.$$

Gubner shows that $\mathcal{W}$ is not $(\mathcal{X}, \mathcal{Y})$-symmetrizable, but that it is both $\mathcal{X}$- and $\mathcal{Y}$-symmetrizable. Thus, this channel is useless if coding is done without conferencing, even though the interfering signal is only added to the sum of the transmitters' signals—the reliable transmission of messages through the channel is completely prevented. This shows that even the structure of rather simple AV-MACs can be changed by conferencing so as to produce discontinuous jumps at $(C_1, C_2) = (0,0)$.

APPENDIX

Here, we prove the weak converse for the compound MAC with conferencing encoders defined by $\overline{\mathcal{W}}$ and with random coding. Let a random $\mathrm{code}_{\mathrm{CONF}}(n, M_1, M_2, C_1, C_2)$ be given which is defined by the pair (C, G). Let this code have average error at most λ. Denote the conferencing function of the deterministic component $\mathrm{code}_{\mathrm{CONF}}$ with index γ by c_γ. The set c_γ maps into is denoted by $[1, V_\gamma]$.

We assume that the pair $((1/n)\log M_1, (1/n)\log M_2)$ is at least distance ε away from $\mathcal{C}^*(\mathcal{S}, C_1, C_2)$. As all norms are equivalent on the plane, we can without loss of generality work with the ℓ^1-norm. That means that we assume that

$$\sup_{(R_1,R_2) \in \mathcal{C}^*(\mathcal{S}, C_1, C_2)} \left\{ \left| \frac{1}{n}\log M_1 - R_1 \right| + \left| \frac{1}{n}\log M_2 - R_2 \right| \right\} \geq \varepsilon.$$

This statement is equivalent to the fact that for every $p \in \Pi$, there is some $q \in \mathcal{P}(\mathcal{S})$ such that one of the following inequalities holds:

$$\frac{1}{n}\log M_1 \geq C_1 + I(Z_q; X|Y, U) + \varepsilon \tag{23}$$

$$\frac{1}{n}\log M_2 \geq C_2 + I(Z_q; Y|X, U) + \varepsilon \tag{24}$$

$$\frac{1}{n}\log M_1 M_2 \tag{25}$$

$$\geq \{C_1 + C_2 + I(Z_q; X, Y|U) \wedge I(Z_q; X, Y|U)\} + \varepsilon. \tag{26}$$

Our goal is to mainly use arguments already known from the weak converse for deterministic coding, so that we can refer to [17]. From the random $\mathrm{code}_{\mathrm{CONF}}$, we define several random variables in addition to G:

1) the pair (T_1, T_2), which is uniformly distributed on $[1, M_1] \times [1, M_2]$ and independent of G;
2) the pair $(\tilde{U}_1, \tilde{U}_2) := (c_1^G(T_1, T_2), c_2^G(T_1, T_2))$ taking values in $[1, V_1] \times [1, V_2]$, where for $\nu = 1, 2$ we define $V_\nu = \max_{\gamma \in \Gamma} V_\nu^\gamma$;
3) $\tilde{X} := \mathbf{x}^G_{T_1 T_2}$, $\tilde{Y} := \mathbf{y}^G_{T_1 T_2}$;
4) a random variable $\tilde{Z} \in \mathcal{Z}^n$ which satisfies

$$\mathbb{P}[\tilde{Z} = \mathbf{z}|\tilde{X} = \mathbf{x}, \tilde{Y} = \mathbf{y}, \tilde{U}_1 = v_1, \tilde{U}_2 = v_2 \quad T_1 = j, T_2 = k, G = \gamma] = W^n(\mathbf{z}|\mathbf{x}, \mathbf{y}).$$

Every $\gamma \in \Gamma$ corresponds to a deterministic code $C(\gamma)$ with average error at most λ_γ. For each of these codes, we can proceed as in [17]. That means that we first apply Fano's inequality and then obtain single-letter bounds on the code rates. More precisely, writing $\mathcal{U} = [1, n] \times [1, V_1] \times [1, V_2]$, we can construct for each γ a probability distribution $p(u, x, y|\gamma)$ on $\mathcal{U} \times \mathcal{X} \times \mathcal{Y}$ which is contained in Π. This is due to the fact proved in [19] that conditional on γ and $(\tilde{U}_1, \tilde{U}_2)$, the random variables $\tilde{X}$ and $\tilde{Y}$ are independent. Thus, we have

$$p(u, x, y|\gamma) = p_0(u|\gamma) p_1(x|u, \gamma) p_2(y|u, \gamma).$$

Further, for each $q \in \mathcal{P}(\mathcal{S})$, we construct the random vector (U, X, Y, Z_q) which together with G has the distribution

$$\mathbb{P}[Z_q = z, Y = y, X = x, U = u, G = \gamma] = W(z|x, y|q) p(u, x, y|\gamma) p_G(\gamma). \tag{27}$$

By construction, this random vector satisfies for every $q \in \mathcal{P}(\mathcal{S})$

$$\frac{1}{n}\log M_1 \leq C_1 + I(Z_q; X|Y,U,G=\gamma) + \frac{1}{n}\Delta_\gamma$$
$$\frac{1}{n}\log M_2 \leq C_2 + I(Z_q; Y|X,U,G=\gamma) + \frac{1}{n}\Delta_\gamma$$
$$\frac{1}{n}\log M_1 M_2 \leq \{(C_1 + C_2 + I(Z_q; X,Y|U,G=\gamma)) \wedge I(Z_q; X,Y|G=\gamma)\} + \frac{1}{n}\Delta_\gamma$$

where

$$\Delta_\gamma := 2h(2\lambda_\gamma) + 4\lambda_\gamma \log M_1 M_2.$$

Next, we take the expectation over G. Using the concavity of h and (27), this can be transformed into

$$\frac{1}{n}\log M_1 \leq C_1 + I(Z_q; X|Y,U,G) + \frac{1}{n}\Delta \quad (28)$$
$$\frac{1}{n}\log M_2 \leq C_2 + I(Z_q; Y|X,U,G) + \frac{1}{n}\Delta \quad (29)$$
$$\frac{1}{n}\log M_1 M_2 \leq \{(C_1 + C_2 + I(Z_q; X,Y|U,G)) \wedge I(Z_q; X,Y|G)\} + \frac{1}{n}\Delta \quad (30)$$

with

$$\Delta := 2h(2\lambda) + 4\lambda \log M_1 M_2.$$

As $I(Z_q; X,Y|G) = H(Z_q|G) - H(Z_q|X,Y)$, the concavity of entropy implies that the bound in (30) is relaxed if one replaces $I(Z_q; X,Y|G)$ by $I(Z_q; X,Y)$. We now set $\hat{U} := (U,G)$ and observe that the distribution of $(\hat{U}, X, Y)$ is contained in Π. Comparing the resulting set of inequalities with a valid one among (23)–(26) and using the same simple arguments as in [17], we can now show that $\lambda \geq \lambda(\varepsilon) > 0$. This finishes the proof of the weak converse.

ACKNOWLEDGMENT

The authors would like to thank Frans Willems for the fruitful discussions about this work at the Banff workshop "Interactive Information Theory" in January 2012. They would also like to thank the associate editor Yossef Steinberg for his valuable comments given during the review process of the paper.

REFERENCES

[1] R. Ahlswede, "Elimination of correlation in random codes for arbitrarily varying channels," *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, vol. 44, pp. 159–175, 1978.
[2] R. Ahlswede, "Coloring hypergraphs: A new approach to multi-user source coding—II," *J. Comb. Inf. Syst. Sci.*, vol. 5, no. 3, pp. 220–268, 1980.
[3] R. Ahlswede, "Arbitrarily varying channels with states sequence known to the sender," *IEEE Trans. Inf. Theory*, vol. IT-32, no. 5, pp. 621–629, Sep. 1986.
[4] R. Ahlswede and N. Cai, "Arbitrarily varying multiple-access channels Part I—Ericson's symmetrizability is adequate, Gubner's conjecture is true," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 742–749, Mar. 1999.
[5] I. Bjelaković, H. Boche, and J. Sommerfeld, Capacity results for arbitrarily varying wiretap channels 2012 [Online]. Available: http://arxiv.org/abs/1209.6325
[6] S. Bross, A. Lapidoth, and M. Wigger, "The Gaussian MAC with conferencing encoders," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2008, pp. 2702–2706.
[7] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2011.
[8] I. Csiszar and P. Narayan, "The capacity of the arbitrarily varying channel revisited: Positivity, constraints," *IEEE Trans. Inf. Theory*, vol. 34, no. 2, pp. 181–193, Mar. 1988.
[9] T. Ericson, "Exponential error bounds for random codes in the arbitrarily varying channel," *IEEE Trans. Inf. Theory*, vol. 31, no. 1, pp. 42–48, Jan. 1985.
[10] J. Gubner, "Deterministic codes for arbitrarily varying multiple-access channels," Ph.D. dissertation, Univ. Maryland, College Park, 1988.
[11] J. Gubner, "On the deterministic-code capacity of the multiple-access arbitrarily varying channel," *IEEE Trans. Inf. Theory*, vol. 36, no. 2, pp. 262–275, Mar. 1990.
[12] J.-H. Jahn, "Coding of arbitrarily varying multiuser channels," *IEEE Trans. Inf. Theory*, vol. 27, no. 2, pp. 212–226, Mar. 1981.
[13] I. Maric, R. Yates, and G. Kramer, "Capacity of interference channels with partial transmitter cooperation," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3536–3548, Oct. 2007.
[14] H. Permuter, S. Shamai, and A. Somekh-Baruch, "Message and state cooperation in multiple access channels," *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6379–6396, Oct. 2011.
[15] O. Simeone, D. Gunduz, H. Poor, A. Goldsmith, and S. Shamai, "Compound multiple-access channels with partial cooperation," *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2425–2441, Jun. 2009.
[16] G. Smith and J. Yard, "Quantum communication with zero-capacity channels," *Science*, vol. 321, no. 5897, pp. 1812–1815, 2008.
[17] M. Wiese, H. Boche, I. Bjelaković, and V. Jungnickel, "The compound multiple access channel with partially cooperating encoders," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 3045–3066, May 2011.
[18] M. A. Wigger, "Cooperation on the multiple-access channel," Ph.D. dissertation, ETH Zürich, Zürich, Switzerland, 2008.
[19] F. M. J. Willems, "Information theoretical results for the discrete memoryless multiple access channel," Ph.D. dissertation, Katholieke Universiteit Leuven, Leuven, Belgium, 1982.
[20] F. M. J. Willems, "The discrete memoryless multiple access channel with partially cooperating encoders," *IEEE Trans. Inf. Theory*, vol. IT-29, no. 3, pp. 441–445, May 1983.

Moritz Wiese (S'09) received the Dipl.-Math. degree in mathematics from the university of Bonn, Germany, in 2007. He has been pursuing the PhD degree since then. From 2007 to 2010, he was a research assistant at the Heinrich-Hertz-Lehrstuhl für Mobilkommunikation, Technische Universität Berlin, Germany. Since 2010, he is a research and teaching assistant at the Lehrstuhl für Theoretische Informationstechnik, Technische Universität München, Munich, Germany.

Holger Boche (M'04–SM'07–F'11) received the Dipl.-Ing. and Dr.-Ing. degrees in electrical engineering from the Technische Universität Dresden, Dresden, Germany, in 1990 and 1994, respectively. He graduated in mathematics from the Technische Universität Dresden in 1992. From 1994 to 1997, he did postgraduate studies in mathematics at the Friedrich-Schiller Universität Jena, Jena, Germany. He received his Dr. rer. nat. degree in pure mathematics from the Technische Universität Berlin, Berlin, Germany, in 1998. In 1997, he joined the Heinrich-Hertz-Institut (HHI) für Nachrichtentechnik Berlin, Berlin, Germany. Starting in 2002, he was a Full Professor for mobile communication networks with the Institute for Communications Systems, Technische Universität Berlin. In 2003, he became Director of the Fraunhofer German-Sino Lab for Mobile Communications, Berlin, Germany, and in 2004 he became the Director of the Fraunhofer Institute for Telecommunications (HHI), Berlin, Germany. Since October 2010 he has been with the Institute of Theoretical Information Technology and Full Professor at the Technische Universität München, Munich, Germany. He was a Visiting Professor with the ETH Zurich, Zurich, Switzerland, during the 2004 and 2006 Winter terms, and with KTH Stockholm, Stockholm, Sweden, during the 2005 Summer term. Prof. Boche is a Member of IEEE Signal Processing Society SPCOM and SPTM Technical Committee. He was elected a Member of the German Academy of Sciences (Leopoldina) in 2008 and of the Berlin Brandenburg Academy of Sciences and Humanities in 2009. He received the Research Award "Technische Kommunikation" from the Alcatel SEL Foundation in October 2003, the "Innovation Award" from the Vodafone Foundation in June 2006, and the Gottfried Wilhelm Leibniz Prize from the Deutsche Forschungsgemeinschaft (German Research Foundation) in 2008. He was co-recipient of the 2006 IEEE Signal Processing Society Best Paper Award and recipient of the 2007 IEEE Signal Processing Society Best Paper Award.