On the Weakest Resource for Coordination in AV-MACs with Conferencing Encoders

Moritz Wiese, Holger Boche
Lehrstuhl für Theoretische Informationstechnik, Technische Universität München, Munich, Germany
{wiese,boche}@tum.de

***Abstract*—If the senders and the receiver of an Arbitrarily Varying Multiple-Access Channel (AV-MAC) have access to the outputs of discrete correlated memoryless sources, the same rate region is achievable as if common randomness were available. This reduces the necessary amount of cooperation in an AV-MAC considerably. Moreover, to transmit blocklength-n words, no more than order $\log n$ source outputs are required.**

I. Introduction

Arbitrarily Varying Channels (AVCs) and Arbitrarily Varying Multiple Access Channels (AV-MACs) are examples of channels where deterministic coding must not be confused with random coding. Ahlswede observed that in AVCs there is a dichotomy: the deterministic capacity of an AVC, i.e. the capacity achievable with deterministic coding, *either* equals the random capacity, i.e. the capacity achievable with random codes, *or* it equals zero [1]. Csiszár and Narayan showed that the discriminating property is *symmetrizability* as introduced by Ericson [6]: the deterministic capacity of an AVC equals zero if and only if the channel is symmetrizable [5].

If the senders of an AV-MAC are allowed to do Willems conferencing at positive rates C_1, C_2, the situation is rather close. Willems conferencing was introduced by Willems in [8], [9]. This iterative way of encoder cooperation interpolates between the traditional non-cooperative case ($C_1 = C_2 = 0$) and the case where the two senders can be treated as one ($C_1 = C_2 = \infty$). The symmetrizability condition applied in this case is Ericson's if the senders of the AV-MAC are considered to be one. One obtains a dichotomy similar to the AVC case: symmetrizability means that no message can be transmitted deterministically, otherwise deterministic coding achieves the same capacity region as random coding [7].

What does random coding mean? It means that both the sender(s) and the receiver know the outcome of a single random experiment which then is used to select the codewords and decoding set for a given message (pair). The random experiment is distributed according to a finite uniform distribution on a set with arbitrary cardinality. For the cases described above, one can show that this cardinality can be chosen to be quadratic in the blocklength of the random code. But still, an application of random codes requires a lot of coordination within the channel which in reality will be hard to establish. Thus the question comes naturally whether less "coordinated randomness" still is able to achieve the random coding capacity or what it achieves instead.

For AVCs, this question was answered by Ahlswede and Cai. If there is a Discrete Memoryless Multiple Source (DMMS) with two stochastically dependent components, the first of which is known to the sender and the second of which is known to the receiver, then using this less coordinated type of randomness is sufficient to achieve the random capacity. This holds no matter whether or not the AVC is symmetrizable. Thus one gets rid of the Ahlswede dichotomy at the expense of allowing some randomness, which however does not have to be coordinated as strongly as in what we have called random coding so far [2]. To distinguish the two types of randomness, we call the type of randomness described in the first paragraphs *common randomness*, whereas in the second case, we say that sender and receiver have *access to correlated sources*.

The analogy to the AVC suggests that for AV-MACs, one might get rid of the dichotomy observed for deterministic coding with the help of correlated sources. It is shown below that this is indeed so and that the complete common randomness capacity region is achievable, both where the senders have a common source correlated with the receiver's source, and where there are three sources, one for each channel terminal. It turns out that this can even be realized in a causal manner.

The core of the proof is to show that with these models, some pair of positive rates is achievable. This can be done by reducing the multi-sender setting to the single-sender setting from [2] via conferencing. Once the existence of a pair of positive rates has been established, one can apply Ahlswede's elimination technique developed in [1]. This is a derandomization technique whose goal is to show the achievability of the random coding capacity region. The underlying idea is that the positive-rate code is used as a prefix code to transmit the random index of any common randomness code.

For the AV-MAC, the application of the elimination technique has interesting consequences: one needs even less coordination of randomness than in the case of the AVC with correlated sources. In the latter case, the sender and the receiver need n independent samples of their respective sources. This result is used to show that a positive rate is achievable in the AV-MAC with correlated sources. But the correlated multiple-access codes thus obtained only enter the code construction in the elimination technique as prefix codes for random codes. As the blocklength of these prefix codes grows logarithmically in the blocklength of the corresponding random code, only order $\log n$ source outputs are necessary to form a correlated sources code with blocklength n.

Suppose that a multiple access channel with conferencing encoders is disturbed by a jammer. Information-theoretically, this generates an AV-MAC. If only deterministic coding is possible and the AV-MAC is symmetrizable, then the jammer can completely prevent any message transmission from the senders to the receiver. However, if the senders and the receiver can observe the output of correlated sources, then the jammer loses its power: transmission is possible at all rate pairs contained in the common randomness capacity region.

Notation: For a positive integer M, we write $[M] := \{1, \ldots, M\}$.

II. Definitions

Fix real numbers $C_1, C_2 > 0$. Let $\mathcal{W}$ be a set of stochastic matrices with input alphabet $\mathcal{X}_1 \times \mathcal{X}_2$ and output alphabet $\mathcal{Y}$. $\mathcal{W}$ determines an *Arbitrarily Varying Multiple-Access Channel (AV-MAC)*. At each time instant, this channel arbitrarily assumes one of the states contained in $\mathcal{W}$. A Willems conference is an iterative protocol to exchange information between the senders. In the first time slot, each sender sends some information to the other. In the subsequent time slots, they again send information taking into account the information from the other sender obtained so far. The process terminates after a fixed number I of iterations. Such a protocol can be described as a pair of functions

$$(c_1, c_2) : A_1 \times A_2 \longrightarrow \mathcal{K}_1 \times \mathcal{K}_2,$$

where A_1, A_2 represent the respective sender's knowledge at the outset of conferencing and $\mathcal{K}_\nu = \mathcal{K}_{\nu 1} \times \ldots \times \mathcal{K}_{\nu I}$ for $\nu = 1, 2$. Details can be found in [7], [9]. A Willems conference has conferencing capacities C_1, C_2 at blocklength n if

$$\frac{1}{n} \log|\mathcal{K}_\nu| \leq C_\nu.$$

In this case, we call it an (n, C_1, C_2)-Willems conference. The input to the conference may include any knowledge the senders have. We use the notation

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

A. Separate source for each node

Let $(U_1^1, U_1^2, V_1), (U_2^1, U_2^2, V_2), \ldots$ be a Discrete Memoryless Multiple Source (DMMS) with three components. We denote a generic triple by (U^1, U^2, V). It attains values in the finite set $\mathcal{U}^1 \times \mathcal{U}^2 \times \mathcal{V}$.

Definition 1 (Model 1): Let ℓ, n, M_1, M_2 be positive integers. An (ℓ, n, M_1, M_2)-code$_1$ is a 5-tuple $(c_1, c_2, f_1, f_2, \phi)$, where

1) (c_1, c_2) is an (n, C_1, C_2)-Willems conference
$$(c_1, c_2) : ([M_1] \times \mathcal{U}^{1\ell}) \times ([M_2] \times \mathcal{U}^{2\ell}) \longrightarrow \mathcal{K}_1 \times \mathcal{K}_2$$
for finite sets $\mathcal{K}_1, \mathcal{K}_2$,
2) the *encoding functions* f_1, f_2 satisfy for $\nu = 1, 2$
$$f_\nu : [M_\nu] \times \mathcal{U}^{\nu\ell} \times \mathcal{K}_{\bar{\nu}} \longrightarrow \mathcal{X}_\nu^n,$$
3) the *decoding function* ϕ satisfies
$$\phi : \mathcal{Y}^n \times \mathcal{V}^\ell \longrightarrow [M_1] \times [M_2].$$

Thus we allow conferencing also to concern a number ℓ of outputs of the DMMS. In [2], only $\ell = n$ is considered. In our case, it is necessary to allow for $\ell \neq n$. Any (ℓ, n, M_1, M_2)-code$_1$ $(c_1, c_2, f_1, f_2, \phi)$ gives rise to a system

$$\{(g^{1n}_{m_1m_2}(u^{1\ell}, u^{2\ell}), g^{2n}_{m_1m_2}(u^{1\ell}, u^{2\ell}), D_{m_1m_2}(v^\ell))_{m_1,m_2=1}^{M_1,M_2} : \\ u^{1\ell} \in \mathcal{U}^{1\ell}, u^{2\ell} \in \mathcal{U}^{2\ell}, v^\ell \in \mathcal{V}^\ell\},$$

where

$$g^{\nu n}_{m_1m_2}(u^{1\ell}, u^{2\ell}) = f_\nu(m_\nu, u^{\nu\ell}, c_{\bar{\nu}}(m_1, u^{1\ell}, m_2, u^{2\ell})),$$
$$D_{m_1m_2}(v^\ell) = \{y^n \in \mathcal{Y}^n : \phi(y^n, v^\ell) = (m_1, m_2)\}.$$

Definition 2 (Average error): For $\lambda > 0$, an (ℓ, n, M_1, M_2)-code$_1$ $(c_1, c_2, f_1, f_2, \phi)$ is an $(\ell, n, M_1, M_2, \lambda)$-code$_1$ if

$$\frac{1}{M_1M_2} \sum_{m_1,m_2=1}^{M_1,M_2} \sum_{u^{1\ell}, u^{2\ell}, v^\ell} P^\ell_{U^1U^2V}(u^{1\ell}, u^{2\ell}, v^\ell) \cdot \\ \cdot W^n(D_{m_1m_2}(v^\ell) | g^{1n}_{m_1m_2}(u^{1\ell}, u^{2\ell}), g^{2n}_{m_1m_2}(u^{1\ell}, u^{2\ell}), s^n) \\ > 1 - \lambda$$

for every $s^n \in \mathcal{S}^n$, if $\mathcal{W} = \{W(\cdot|\cdot, \cdot, s) : s \in \mathcal{S}\}$.

B. One source for both senders

Let $(U_1, V_1), (U_2, V_2), \ldots$ be a DMMS with two components. We denote a generic pair by (U, V). It attains values in the finite set $\mathcal{U} \times \mathcal{V}$. The Willems conference can be restricted to only concern the messages because we will see that this already is optimal.

Definition 3 (Model 2): Let ℓ, n, M_1, M_2 be positive integers. An (ℓ, n, M_1, M_2)-code$_2$ is a 5-tuple $(c_1, c_2, f_1, f_2, \phi)$, where

1) (c_1, c_2) is an (n, C_1, C_2)-Willems conference
$$(c_1, c_2) : [M_1] \times [M_2] \longrightarrow \mathcal{K}_1 \times \mathcal{K}_2$$
for finite sets $\mathcal{K}_1, \mathcal{K}_2$,
2) the *encoding functions* f_1, f_2 satisfy for $\nu = 1, 2$
$$f_\nu : [M_\nu] \times \mathcal{U}^\ell \times \mathcal{K}_{\bar{\nu}} \longrightarrow \mathcal{X}_\nu^n,$$
3) the *decoding function* ϕ satisfies
$$\phi : \mathcal{Y}^n \times \mathcal{V}^\ell \longrightarrow [M_1] \times [M_2].$$

Codewords and decoding sets are defined in the same way as in Model 1, the average error and $(\ell, n, M_1, M_2, \lambda)$-codes$_2$ are defined analogously.

C. Capacity regions

Definition 4: Let $\mu \in \{1, 2\}$. A pair (R_1, R_2) of nonnegative real numbers is called a *rate pair achievable by model* μ if for every $\varepsilon, \lambda > 0$ and sufficiently large n there are ℓ, M_1, M_2 and an $(\ell, n, M_1, M_2, \lambda)$-code$_\mu$ with

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

The set of all achievable rate pairs is called the *capacity region of model* μ and denoted by $\mathcal{C}_\mu$.

D. Deterministic and Common Randomness Codes

All of the above models generalize the concept of deterministic coding. Deterministic codes are also the components of codes which apply common randomness shared by senders and receiver. The latter generalize the two above models. The question answered in this work will be where $\mathcal{C}_\mu$ lies between the deterministic and common randomness capacity regions.

Definition 5 (Deterministic codes): Let n, M_1, M_2 be positive integers. An (n, M_1, M_2)-code$_d$ is a 5-tuple $(c_1, c_2, f_1, f_2, \phi)$, where

1) (c_1, c_2) is a (n, C_1, C_2)-Willems conference
$$(c_1, c_2) : [M_1] \times [M_2] \longrightarrow \mathcal{K}_1 \times \mathcal{K}_2$$
for finite sets $\mathcal{K}_1, \mathcal{K}_2$,
2) the *encoding functions* f_1, f_2 satisfy for $\nu = 1, 2$
$$f_\nu : [M_\nu] \times \mathcal{K}_{\bar{\nu}} \longrightarrow \mathcal{X}_\nu^n,$$
3) the *decoding function* ϕ satisfies
$$\phi : \mathcal{Y}^n \longrightarrow [M_1] \times [M_2].$$

The average error is defined analogously to the average error of a code$_\mu$.

Definition 6 (Common randomness codes): Let n, M_1, M_2 be positive integers. An (n, M_1, M_2)-code$_r$ is a family
$$\{(c_1(\gamma), c_2(\gamma), f_1(\gamma), f_2(\gamma), \phi(\gamma)) : \gamma \in \Gamma\}, \tag{1}$$
Γ a finite set, of (n, M_1, M_2)-codes$_d$. $|\Gamma|$ is called the *amount of common randomness*.

For $\lambda \geq 0$, an (n, M_1, M_2)-code$_r$ is an (n, M_1, M_2, λ)-code$_r$ if
$$\begin{aligned}\frac{1}{M_1 M_2} &\sum_{m_1, m_2 = 1}^{M_1, M_2} \frac{1}{|\Gamma|} \sum_{\gamma \in \Gamma} \\ &W^n(D_{m_1 m_2}(\gamma) | g^{1n}_{m_1 m_2}(\gamma), g^{2n}_{m_1 m_2}(\gamma), s^n) \\ &> 1 - \gamma,\end{aligned}$$
where $g^{\nu n}_{m_1 m_2}(\gamma)$, $\nu = 1, 2$, and $D_{m_1 m_2}(\gamma)$ are the codewords and the decoding set, respectively, for message pair (m_1, m_2) in the γ-th deterministic component of the (n, M_1, M_2)-code$_r$.

The achievability of a rate pair by codes$_d$ or codes$_r$, respectively, is defined analogously to Definition 4. In contrast to Models 1-2, to form a code$_r$, the senders and the receiver jointly have to know the outcome of a uniform random experiment on a finite set whose size can be chosen. Note that using a uniform distribution on a sufficiently large set can be used to approximate any discrete probability distribution. The continuity of the average error of codes$_r$ in the distribution of the common randomness thus implies that common randomness codes generalize those from Models 1 and 2.

Definition 7:

1) A pair (R_1, R_2) which is achievable by codes$_d$ is called *deterministically achievable*. The set of all such rate pairs is called the *deterministic capacity region* and denoted by $\mathcal{C}_d$.
2) A pair (R_1, R_2) which is achievable by codes$_r$ is called *achievable with common randomness*. The set of all such rate pairs is called the *common randomness capacity region* and denoted by $\mathcal{C}_r$.

We note the obvious relations
$$\mathcal{C}_d \subset \mathcal{C}_1 \cap \mathcal{C}_2 \subset \mathcal{C}_1 \cup \mathcal{C}_2 \subset \mathcal{C}_r. \tag{2}$$

$\mathcal{C}_r$ depends on C_1, C_2 in a continuous manner. In Model 1, conferencing may also concern the outputs of U^1 and U^2. Such conferencing will be necessary to achieve $\mathcal{C}_r$. However, this has to be done in such a way that one does not lose rate compared to $\mathcal{C}_r$. The actual form of $\mathcal{C}_r$ is as follows. Let $\overline{\mathcal{W}}$ be the convex closure of $\mathcal{W}$ and let $\overline{\mathcal{S}}$ be an index set for $\overline{\mathcal{W}}$, so $\overline{\mathcal{W}} = \{W_{\bar{s}}(\cdot|\cdot,\cdot,\bar{s}) : \bar{s} \in \overline{\mathcal{S}}\}$. Then
$$\mathcal{C}_r = closure\left(conv\left(\bigcup_p \mathcal{R}(p)\right)\right).$$

Here, p ranges over the families of random variables $p = (Z, X_1, X_2, \{Y_{\bar{s}}\}_{\bar{s} \in \overline{\mathcal{S}}})$ where Z attains values on an arbitrary finite set, (X_1, X_2) on $\mathcal{X}_1 \times \mathcal{X}_2$ and $Y_{\bar{s}}$ on $\mathcal{Y}$ for every $\bar{s} \in \overline{\mathcal{S}}$. Further, for every $\bar{s} \in \overline{\mathcal{S}}$, the sequence $Z, (X_1, X_2), Y_{\bar{s}}$ forms a Markov chain, X_1, X_2 are independent conditional on Z, and the distribution of $Y_{\bar{s}}$ conditional on (X_1, X_2) equals $W_{\bar{s}}$. The sets $\mathcal{R}(p)$ can be written as $\bigcap_{\bar{s} \in \overline{\mathcal{S}}} \mathcal{R}(p, \bar{s})$. The set $\mathcal{R}(p, \bar{s})$ consists of those pairs (R_1, R_2) of nonnegative real numbers satisfying
$$R_1 \leq I(Y_{\bar{s}} \wedge X_1 | X_2 Z) + C_1, \tag{3a}$$
$$R_2 \leq I(Y_{\bar{s}} \wedge X_2 | X_1 Z) + C_2, \tag{3b}$$
$$\begin{aligned}R_1 + R_2 \leq \min\{&I(Y_{\bar{s}} \wedge X_1 X_2 | Z) + C_1 + C_2, \\ &I(Y_{\bar{s}} \wedge X_1 X_2)\}.\end{aligned} \tag{3c}$$

For a better understanding of the discussions, we also give the definition of symmetrizability.

Definition 8: An AV-MAC is called *jointly symmetrizable* if there is a stochastic matrix σ with inputs from $\mathcal{X}_1 \times \mathcal{X}_2$ and outputs from $\mathcal{S}$ such that for every choice $x_1, x_1' \in \mathcal{X}_1$, $x_2, x_2' \in \mathcal{X}_2$ and $y \in \mathcal{Y}$,
$$\sum_{s \in \mathcal{S}} W(y|x_1, x_2, s)\sigma(s|x_1', x_2') = \sum_{s \in \mathcal{S}} W(y|x_1', x_2', s)\sigma(s|x_1, x_2).$$

Remark 1: Two other, "marginal" symmetrizability conditions exist for AV-MACs. They are important in the description of the capacity region of the AV-MAC without conferencing. If both of these conditions are satisfied, the AV-MAC is useless with deterministic coding. However, the complete deterministic capacity region of the AV-MAC without conferencing encoders is not yet completely known. Details can be found in [3], [7]. With $C_1, C_2 > 0$, these marginal symmetrizability conditions do not matter any more. Thus conferencing changes the structure of AV-MACs into the direction of single-sender AVCs.

III. Main Theorem

Recall the dichotomy derived in [7] for $\mathcal{C}_d$: it equals $\mathcal{C}_r$ if $\mathcal{W}$ is not jointly symmetrizable, otherwise it equals $\{(0, 0)\}$. The question is whether common randomness in all generality is necessary to achieve $\mathcal{C}_r$ or whether, and if so how far, this condition can be relaxed. Models 1 and 2 describe a setting where the senders and the receiver only have access to the outputs of given correlated sources.

Theorem 1: Correlated sources at senders and receiver are sufficient to achieve $\mathcal{C}_r$. More precisely:

1) $\mathcal{C}_1 = \mathcal{C}_r$ if $I(U_1 U_2 \wedge V) > 0$.
2) $\mathcal{C}_2 = \mathcal{C}_r$ if $I(U \wedge V) > 0$.

From (3a)-(3c) it is intuitively clear that using conferencing for other purposes than to exchange information about messages could reduce the achievable rates. Hence in the proof of the theorem one has to be careful to keep the conferencing not concerned with messages as short as possible.

IV. Proof of Main Theorem

A. Achieving a positive rate

This is the central part of the proof of the theorem. We may assume that $\mathcal{C}_r \neq \{(0,0)\}$, otherwise the theorem is trivial. First we treat Model 2. We formulate the corresponding result as a lemma.

Lemma 1: Set $\alpha := C_1/(C_1+C_2)$. Then there exists an $R \in (0, \min\{C_1, C_2\})$ with $(\alpha R, (1-\alpha)R) \in \mathcal{C}_2$.

Proof: Consider $\mathcal{W}$ as a single-sender Arbitrarily Varying Channel (AVC) as in [2] with input alphabet $\mathcal{X}_1 \times \mathcal{X}_2$ and output alphabet $\mathcal{Y}$. Using $I(U \wedge V) > 0$, it is proved in [2] that with the help of the sequence of random variables $(U_1, V_1), (U_2, V_2), \ldots$, this AVC achieves a positive rate R which we can without loss of generality assume to be strictly smaller than $\min\{C_1, C_2\}$. More precisely, for arbitrary $\varepsilon, \lambda > 0$ and n sufficiently large there exists a system

$$\{(\tilde{g}_m^n(u^n), \tilde{D}_m(v^n))_{m=1}^M : u^n \in \mathcal{U}^n, v^n \in \mathcal{V}^n\}, \tag{4a}$$

$$\tilde{g}_m^n(u^n) \in (\mathcal{X}_1 \times \mathcal{X}_2)^n \text{ for } u^n \in \mathcal{U}^n, \tag{4b}$$

$$\tilde{D}_m(v^n) \subset \mathcal{Y}^n \text{ for } v^n \in \mathcal{V}^n \tag{4c}$$

$$\text{with } \tilde{D}_m(v^n) \cap \tilde{D}_{m'}(v^n) = \varnothing \text{ for } m \neq m'$$

where M is a positive integer satisfying

$$\min\{C_1, C_2\} \geq \frac{1}{n}\log M \geq R - \varepsilon \tag{5}$$

and where

$$\frac{1}{M}\sum_{m=1}^{M}\sum_{u^n,v^n} P_{UV}^n(u^n,v^n)\cdot W^n(\tilde{D}_m(v^n)|\tilde{g}_m^n(u^n),s^n) > 1-\lambda.$$

By enlarging n if necessary, one can find positive integers M_1, M_2 which satisfy $M/2 \leq M_1 M_2 \leq M$ and

$$\frac{1}{n}\log M_\nu \geq \frac{C_\nu}{C_1+C_2}R - \varepsilon, \qquad \nu = 1,2. \tag{6}$$

Thus the system (4) has a subsystem

$$\{(g_{m_1m_2}^{1n}(u^n), g_{m_1m_2}^{2n}(u^n), D_{m_1m_2}(v^n))_{m_1,m_2=1}^{M_1,M_2} : \\ u^n \in \mathcal{U}^n, v^n \in \mathcal{V}^n\},$$

where $g_{m_1m_2}^{\nu n}(u^n)$ is the $\mathcal{X}^{\nu n}$-component of $\tilde{g}_{m_1m_2}^n(u^n)$. This gives us the possibility of defining an (n, n, M_1, M_2)-code$_2$. We set $\mathcal{K}_\nu := [M_\nu]$ and define one-shot conferencing functions c_ν as the identity on $[M_\nu]$, for $\nu = 1, 2$. This conference is admissible because of (5) and $R < \min\{C_1, C_2\}$. The encoding functions f_1, f_2 are defined as

$$f_\nu(m_1, m_2, u^n) := g_{m_1m_2}^{\nu n}(u^n), \qquad \nu = 1,2.$$

The decoding sets are obvious. The average error of the resulting (n, n, M_1, M_2)-code$_2$ $(c_1, c_2, f_1, f_2, \phi)$ applied to the channel with state sequence $s^n \in \mathcal{S}^n$ satisfies

$$\begin{aligned}
&\frac{1}{M_1M_2}\sum_{m_1,m_2=1}^{M_1,M_2}\sum_{u^n,v^n} P_{UV}^n(u^n,v^n)\cdot \\
&\quad\cdot W^n(D_{m_1m_2}(v^n)^c|g_{m_1m_2}^{1n}(u^n), g_{m_1m_2}^{2n}(u^n), s^n) \\
&\leq \frac{M}{M_1M_2}\cdot \\
&\quad\cdot \frac{1}{M}\sum_{m=1}^{M}\sum_{u^n,v^n} P_{UV}^n(u^n,v^n)\cdot W^n(\tilde{D}(v^n)^c|\tilde{g}_m^n(u^n),s^n) \\
&\leq 2\lambda.
\end{aligned}$$

This completes the proof of the lemma. ■

Note that conferencing is essential to be able to apply the single-user AVC theory of [2] to the AV-MAC. To prove a statement analogous to Lemma 1 for Model 1, the idea is that the senders first have a conference about the outputs of their respective sources. Then we can refer to Lemma 1. Even though the source outputs the senders have a conference about are generated previous to the application of the code$_2$ from Lemma 1, the average error of the concatenation is small due to the memorylessness of the source.

Lemma 2: Set $\alpha := C_1/(C_1+C_2)$ and define

$$\beta := \begin{cases} \left(\frac{1}{\min\{C_1,C_2\}} + 2\right)^{-1} & \text{if } \min\{C_1,C_2\} < 1, \\ 1/2 & \text{if } \min\{C_1,C_2\} \geq 1. \end{cases}$$

Then there exists an $R \in (0, \min\{C_1, C_2\})$ with $\beta R \cdot (\alpha, 1-\alpha) \in \mathcal{C}_1$.

Proof: As noted in [2] (details in [4]) we may assume that $|\mathcal{U}^\nu| = 2$, otherwise the senders can apply binary functions b_1, b_2 such that $(b_1(U^1), (b_2(U^2))$ and V are still stochastically dependent. Let $\varepsilon, \lambda > 0$. By Lemma 1 there is a rate $0 < R < \min\{C_1, C_2\}$ such that for sufficiently large n one can find an $(n, n, M_1, M_2, \lambda)$-code$_2$ $(c_1, c_2, f_1, f_2, \phi)$ satisfying

$$\frac{1}{n}\log M_1 \geq \alpha R - \varepsilon, \qquad \frac{1}{n}\log M_2 \geq (1-\alpha)R - \varepsilon.$$

If $\min\{C_1, C_2\} < 1$, set $n' := \lceil n/\min\{C_1, C_2\}\rceil$, otherwise let $n' = n$. We define an $(n, n', 1, 1)$-code$_1$ as follows: for $\nu = 1, 2$, let c_ν^* be the identity on $\mathcal{U}^{\nu n}$. (c_1^*, c_2^*) is an admissible Willems conferencing protocol at blocklength n'. The encoding and decoding functions of the prefix $(n, n', 1, 1)$-code$_1$ are determined by the senders' unique codewords $g_*^{1n'}, g_*^{2n'}$ and the complete set $\mathcal{Y}^{n'}$ as decoding set.

To enable transmission at a positive rate in Model 2, we concatenate the $(n, n', 1, 1)$-code$_1$ and the (n, n, M_1, M_2)-code$_2$. Assume transmission starts at time instant 1. Using the prefix $(n, n', 1, 1)$-code$_1$, the senders exchange their first

n source outputs u^{1n}, u^{2n}. When the (n, n, M_1, M_2)-code$_2$ is applied after this, the senders and the receiver do not use the source outputs at times $n'+1, \ldots, n'+n$ to form their codewords and decoding sets, but those exchanged earlier. This forms a $(2n, n'+n, M_1, M_2)$-code$_1$ which is admissible because the concatenation of admissible conferencing protocols is admissible. As the sources are memoryless, the average error of the $(2n, n'+n, M_1, M_2)$-code$_1$ is upper-bounded by λ. Finally, we note that

$$\frac{1}{n'+n}\log M_1$$
$$\geq \alpha(R-\varepsilon)\cdot\begin{cases}\left(\frac{1}{\min\{C_1,C_2\}}+\frac{1}{n}+1\right)^{-1}, & \min\{C_1,C_2\}<1,\\ 1/2, & \min\{C_1,C_2\}\geq 1.\end{cases}$$

An analogous statement holds for M_2, with α replaced by $1-\alpha$. This completes the proof. ■

B. Ahlswede's elimination technique

In Lemmas 1 and 2 it was shown that a positive rate is achievable by both models. Now we can establish the equality $\mathcal{C}_\mu = \mathcal{C}_r$. This is done using Ahlswede's elimination technique from [1] which was already applied in [7]. It builds on the concatenation of the positive-rate achieving code$_\mu$ with a code$_r$.

Lemma 3 ([7], Lemma 15): $\mathcal{C}_r$ is achievable by codes$_r$ whose amount of common randomness is quadratic in blocklength.

Now for any $\lambda > 0$ take an (n, M_1, M_2, λ)-code$_r$ given by a family as in (1), where $\Gamma = [n^2]$. The goal is to construct a code$_\mu$ with approximately the same message sets and average error. To do this, one of the senders, say the first one, first chooses an element of γ uniformly at random. Using the positive-rate code$_\mu$ whose existence we are assuming at the moment, the outcome of this experiment can be transmitted to the sender at vanishing rate cost. (Note that by the construction of the code$_\mu$, the second encoder is also informed about the outcome of the random experiment.) More precisely, assume that $(R, R) \in \mathcal{C}_\mu$ for some $R > 0$. Then if n is large, there is an ℓ and an $(\ell, n', n^2, n^2, \lambda)$-code$_\mu$ $(c_1^*, c_2^*, f_1^*, f_2^*, \phi^*)$ with $n' = \lceil (4\log n)/R \rceil$, because this implies $(2\log n)/n' \leq R/2$. For a given $\gamma \in [n^2]$, the first sender can now send this index to the receiver, the other sender transmits an arbitrary message.

Once all parties have been informed about the value γ, the (n, M_1, M_2)-code$_d$ given by $(c_1(\gamma), c_2(\gamma), f_1(\gamma), f_2(\gamma), \phi(\gamma))$ can be applied. This concatenation defines an $(\ell, n'+n, n^2M_1, n^2M_2)$-code$_\mu$ with small average error. It is admissible because for $\nu = 1, 2$

$$\frac{\log(|\mathcal{K}_\nu^*||\mathcal{K}_\nu(\gamma)|)}{n'+n} \leq \frac{n'}{n'+n}C_\nu + \frac{n}{n'+n}C_\nu = C_\nu,$$

where $\mathcal{K}_\nu^*$ denotes the range of c_ν^* and $\mathcal{K}_\nu(\gamma)$ denotes the range of $c_\nu(\gamma)$. We also obtain for sufficiently large n

$$\frac{\log(n^2M_\nu)}{n'+n} \geq \frac{n}{n'+n}(R-\varepsilon) \geq R - 2\varepsilon.$$

The average error of the concatenation is upper-bounded by 2λ. Instead of proving this we just state Ahlswede's Innerproduct lemma [1] and appeal to [1] and [7] for details of the application ([1] shows and actually develops the elimination technique in the case of single-sender arbitrarily varying channels, [7] shows it in the case of AV-MACs).

Lemma 4 (Innerproduct lemma, [1]): Let $(\alpha_1, \ldots, \alpha_N)$ and $(\beta_1, \ldots, \beta_N)$ be two vectors with $0 \leq \alpha_m, \beta_m \leq 1$ for $m = 1, \ldots, N$ which for some $\lambda \in (0, 1)$ satisfy

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

then

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

Thus we have obtained an $(\ell, n'+n, n^2M_1, n^2M_2, 2\lambda)$-code$_\mu$, which shows that all rates achievable by codes$_r$ are achievable by codes$_\mu$. This means $\mathcal{C}_r \subset \mathcal{C}_\mu$. In (2) we noted that the converse holds trivially, so we have equality.

Corollary 1: For $\mu = 1, 2$, $\mathcal{C}_\mu$ is achievable using (ℓ, n, M_1, M_2)-codes$_\mu$ with $\ell = O(\log n)$.

Proof: The codes$_\mu$ constructed above using Ahlswede's elimination technique only require the knowledge of the correlated sources in the first $a \log n$ time slots, where n is the blocklength and a is a constant. ■

Corollary 1 should be compared to the single-sender setting of [2], where the construction requires the observation of n source outcomes.

Remark 2: Note that the above construction is causal as far as the DMMS is concerned. No output of the DMMS is used at the senders or the receiver before it has been realized.

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 and N. Cai, "Correlated sources help transmission over an arbitrarily varying channel," *IEEE Trans. Inf. Theory*, vol. 43, no. 4, pp. 1254–1255, 1997.

[3] ——, "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, 1999.

[4] H. Boche and J. Nötzel, "Arbitrarily small amounts of correlation for arbitrarily varying quantum channels," 2013, available at http://arxiv.org/abs/1301.6063.

[5] 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.

[6] 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.

[7] M. Wiese and H. Boche, "The arbitrarily varying multiple-access channel with conferencing encoders," *IEEE Trans. Inf. Theory*, vol. 59, no. 3, pp. 1405–1416, 2013.

[8] F. M. J. Willems, "Informationtheoretical results for the discrete memoryless multiple access channel," Ph.D. dissertation, Katholieke Universiteit Leuven, Belgium, 1982.

[9] ——, "The discrete memoryless multiple access channel with partially cooperating encoders," *IEEE Trans. Inf. Theory*, vol. IT-29, no. 3, pp. 441–445, 1983.