The Compound MAC With Common Message and Partial Channel State Information

Moritz Wiese, Holger Boche, Igor Bjelaković

Heinrich-Hertz-Lehrstuhl für Informationstheorie und Theoretische Informationstechnik

Technische Universität Berlin

{moritz.wiese, holger.boche, igor.bjelakovic}@mk.tu-berlin.de

Abstract—We characterize the capacity region of the compound Discrete Memoryless Multiple Access Channel, where both transmitters have an additional common message. The channel state information is as follows: for each transmitter, there is a finite partition of the set of channels. Each transmitter knows which element of his partition the channel actually used belongs to. The capacity region is not affected by the amount of channel state information at the receiver, which may be arbitrary.

I. INTRODUCTION

The capacity region of the Discrete Memoryless Multiple Access Channel (DM-MAC) with common message has first been characterized by Slepian and Wolf in [5]. For the compound DM-MAC with common message, where the class of channels consists of two elements and where the channel is known at the receiver only, the capacity region has been obtained in [3]. In this paper, we find the capacity region of the compound DM-MAC with common message, where the number of channels is arbitrary. Various degrees of channel state information (CSI) are assumed at the transmitter (CSIT) and at the receiver (CSIR). For each sender, there is a finite partition of the set of channels. Each sender has partial CSIT in that he knows which element of his partition the channel actually used for transmission belongs to. The receiver may have any amount of CSIR between full channel knowledge and none at all. It will be seen that CSIR has no influence on the capacity region. The above definition of CSI includes scenarios with asymmetric distribution of CSI between all nodes as well as complete lack of CSI. By letting the partitions defining CSIT become finer and finer, one sees how the capacity region scales with increasing CSIT.

As shown in [6] and [7], the achievability of the capacity region of the DM-MAC with conferencing encoders, where the transmitters are able to cooperate over noiseless channels with finite capacities, can be obtained via the capacity region of the DM-MAC with common message. The result shown in the present paper is the first step towards a characterization of the capacity region of the compound DM-MAC with conferencing encoders and with partial CSI. Such a characterization would constitute a theoretical analysis of the downlink performance that can be achieved using cooperating base stations with imperfect and possibly different CSIT, as will be incorporated in future wireless systems.

The rest of the paper is organized as follows: in Section II, we describe the channel model and state the theorem
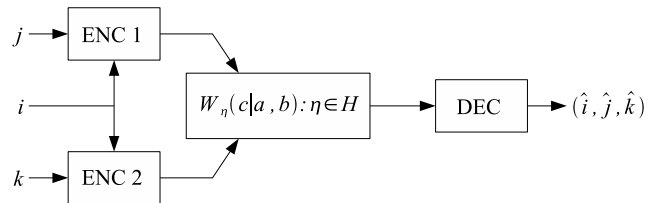

Fig. 1. The compound MAC with common message

characterizing the capacity region of the compound DM-MAC with common message and partial CSI. Section III proves the achievability of the asserted capacity region. In Section IV, a weak converse is proved. A numerical example is given in Section V. In Section VI, we discuss the result and conclude the paper. Several auxiliary lemmata are collected in the Appendix.

Some notation: for a positive integer m, define $[1, m] := \{1, \ldots, m\}$. For real numbers x, y, define $x \vee y := \max(x, y)$. For a finite set G, let G^c denote the complement of G. Let $\mathcal{P}(G)$ denote the set of probability measures on G. For $p \in \mathcal{P}(G)$, p^n denotes the n-fold tensor product of p on the set G^n.

II. CHANNEL MODEL AND THEOREM

Let A, B, C be finite sets and let H be an arbitrary set. For $\eta \in H$, let W_η be a stochastic matrix with entries $W_\eta(c|a, b)$, where $(a, b, c) \in A \times B \times C$. This defines a compound DM-MAC with input alphabets A for transmitter 1 and B for transmitter 2 and output alphabet C. For input words $x \in A^n$ and $y \in B^n$, the probability that the corresponding output equals $z \in C^n$ if the channel realization is $\eta \in H$ is

$$W_\eta^n(z|x, y) := \prod_{i=1}^n W_\eta(z_i|x_i, y_i).$$

Let T_1, T_2 be finite sets and let the surjective CSIT-functions $t_\nu : H \to T_\nu$, $\nu = 1, 2$, be given. These functions induce two finite partitions of the set H. For $\tau = (\tau_1, \tau_2) \in T_1 \times T_2$ we set $H_\tau := H_{\tau_1 \tau_2} := t_1^{-1}(\tau_1) \cap t_2^{-1}(\tau_2)$. Let another function $r : H \to R$ be given, where R is an arbitrary set. This is the CSIR-function. If channel η is used for transmission, sender ν knows $t_\nu(\eta)$ and the receiver knows $r(\eta)$. We denote the

compound DM-MAC $\{W_\eta : \eta \in H\}$ together with the CSI-functions t_1, t_2, r by $\mathcal{W}(H, t_1, t_2, r)$.

Now fix a compound DM-MAC $\mathcal{W}(H, t_1, t_2, r)$. Let M_0, M_1, M_2 be positive integers and set $R_\nu := (1/n) \log M_\nu$ for $\nu = 0, 1, 2$. A code(n, R_0, R_1, R_2) for $\mathcal{W}(H, t_1, t_2, r)$ is a triple $(f_1, f_2, (\mathcal{C}_\rho)_{\rho \in R})$. Here,

$$f_1 : [1, M_0] \times [1, M_1] \times T_1 \to A^n,$$
$$f_2 : [1, M_0] \times [1, M_2] \times T_2 \to B^n$$

are encoding functions. Each $\mathcal{C}_\rho$ is a collection $\{C_{ijk}^\rho : (i, j, k) \in [1, M_0] \times [1, M_1] \times [1, M_2]\}$ of disjoint subsets of C^n. $\mathcal{C}_\rho$ is the family of decoding sets used by the receiver if it has CSIR ρ. If $\eta \in H_{\tau_1 \tau_2}$, where $(\tau_1, \tau_2) \in T_1 \times T_2$, and if the message triple (i, ℓ_1, ℓ_2) is to be transmitted, sender ν uses the codeword $f_\nu(i, \ell_\nu, \tau_\nu)$, for $\nu = 1, 2$. This means that the codeword chosen for transmission depends on the CSIT. We also write

$$f_1(i, \ell_1, \tau_1) = x_{i\ell_1}^{\tau_1},$$
$$f_2(i, \ell_2, \tau_2) = y_{i\ell_2}^{\tau_2}.$$

If the receiver knows that the index of the channel used for transmission is contained in $r^{-1}(\rho)$ and if the received element $z \in C^n$ is contained in the set C_{ijk}^ρ, the receiver decides that the message triple (i, j, k) has been sent.

The code(n, R_0, R_1, R_2) is a code$(n, R_0, R_1, R_2, \lambda)$ if

$$\frac{1}{M_0 M_1 M_2} \sum_{i,j,k} W_\eta^n((C_{ijk}^\rho)^c | x_{ij}^{\tau_1}, y_{ik}^{\tau_2}) \leq \lambda$$

for every $(\tau_1, \tau_2, \rho) \in T_1 \times T_2 \times R$ and every $\eta \in H_{\tau_1 \tau_2} \cap r^{-1}(\rho)$, i. e. if the average error is smaller than λ for all $\eta \in H$.

A rate triple (R_0, R_1, R_2) is achievable if for every $\lambda > 0$ and for every $\varepsilon > 0$, there is an n_0 such that for every $n \geq n_0$, there is a code$(n, R_0', R_1', R_2', \lambda)$ with $R_\nu' \geq R_\nu - \varepsilon$ for every $\nu = 0, 1, 2$. Denote the set of rate triples which are achievable by $\mathcal{C}$.

Next, we introduce the notation used for the parametrization of the capacity region. Let Π_0 be the set of pairs (D, p_0), where D is a finite set and $p_0 \in \mathcal{P}(D)$. Further, for a finite set D, let $\Pi(D)$ be the set of pairs (p_1, p_2), where p_ν is a mapping from T_ν to the set of stochastic matrices with input alphabet D and output alphabet A ($\nu = 1$) or B ($\nu = 2$).

Let $(D, p_0) \in \Pi_0$, $(p_1, p_2) \in \Pi(D)$, $\tau := (\tau_1, \tau_2) \in T_1 \times T_2$, and $\eta \in H_\tau$. For $\nu = 1, 2$, we write $p_\nu(\tau_\nu) = p_{\nu \tau_\nu}$. Define the probability distribution r_η on $D \times A \times B \times C$ by

$$r_\eta(d, a, b, c) = p_0(d) p_{1\tau_1}(a|d) p_{2\tau_2}(b|d) W_\eta(c|a, b). \quad (1)$$

Let the quadruple of random variables $(U, X_{\tau_1}, Y_{\tau_2}, Z_\eta)$ take values in $D \times A \times B \times C$ with joint probability r_η. Set

$$\mathcal{R}(p_0, p_1, p_2, \tau_1, \tau_2, \eta) := \{(R_0, R_1, R_2) \in \mathbb{R}_{\geq 0}^3 :$$
$$R_1 \leq I(Z_\eta; X_{\tau_1} | Y_{\tau_2}, U),$$
$$R_2 \leq I(Z_\eta; Y_{\tau_2} | X_{\tau_1}, U),$$
$$R_1 + R_2 \leq I(Z_\eta; X_{\tau_1}, Y_{\tau_2} | U),$$
$$R_0 + R_1 + R_2 \leq I(Z_\eta; X_{\tau_1}, Y_{\tau_2})\}.$$

Further, we define the set

$$\mathcal{C}^* = \left(\bigcup_{(D, p_0) \in \Pi_0} \bigcup_{(p_1, p_2) \in \Pi(D)} \right.$$
$$\left. \bigcap_{(\tau_1, \tau_2) \in T_1 \times T_2} \bigcap_{\eta \in H_{\tau_1 \tau_2}} \mathcal{R}(p_0, p_1, p_2, \tau_1, \tau_2, \eta) \right).$$

Theorem 1. *For the capacity region $\mathcal{C}$ of the DM-MAC $\mathcal{W}(H, t_1, t_2, r)$, one has*

$$\mathcal{C} = \mathcal{C}^*.$$

$\mathcal{C}$ is convex. It can be assumed for the sets D in the definition of $\mathcal{C}^*$ that $|D| \leq \min\{|A||B| + 2, |C| + 3\}$.

Remark 1. Note that the capacity region is independent of r, the CSIR function. This is implied by the fact that for any fixed pair of CSIT functions (t_1, t_2), the inner bound for the case of no CSIR from Section III coincides with the outer bound for the case of full CSIR from Section IV. Further note that the result of [3] describes the case $|H| = 2$ with no CSIT and full CSIR.

Remark 2. The convexity of $\mathcal{C}^*$ is seen easily using the convexity properties of mutual information. The bound on the cardinality of D follows as in [6, Appendix B].

III. ACHIEVABILITY

As noted in Remark 1, it suffices to prove the achievability part of Theorem 1 for the case of no CSIR. As r is trivial in this case, we leave it away completely to simplify notation.

First assume that the number of channels is finite. To prove the achievability of the rates in $\mathcal{C}^*$, we use the random coding method. This means that we randomly generate a code. Then we show that if its rate triple is within $\mathcal{C}^*$, the probability that its average error is exponentially small in blocklength is greater than zero. This implies the existence of a deterministic code with exponentially small average error probability.

For some $(D, p_0) \in \Pi_0$ and $(p_1, p_2) \in \Pi(D)$, choose a triple (R_0, R_1, R_2) contained in the interior of

$$\bigcap_{(\tau_1, \tau_2) \in T_1 \times T_2} \bigcap_{\eta \in H_{\tau_1 \tau_2}} \mathcal{R}(p_0, p_1, p_2, \tau_1, \tau_2, \eta). \quad (2)$$

Let $\lambda > 0$, and without loss of generality, let $\varepsilon < \min\{R_0, R_1, R_2\}$. Choose positive integers n, M_0, M_1, M_2 with

$$R_\nu - \frac{\varepsilon}{2} > R_\nu' := \frac{1}{n} \log M_\nu \geq R_\nu - \varepsilon, \quad \nu = 0, 1, 2. \quad (3)$$

(This is possible for all large n.) We now define the probability space on which codes(n, R_0', R_1', R_2') are generated randomly. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space on which for each $i \in [1, M_0]$, the family of random variables

$$\{U_i, X_{ij}^{\tau_1}, Y_{ik}^{\tau_2} : j \in [1, M_1], \tau_1 \in T_1, k \in [1, M_2], \tau_2 \in T_2\}$$
$$(4)$$

is defined. These families are independent for different i. For every i, the U_i is distributed on D^n according to p_0^n. Further, the $X_{ij}^{\tau_1}$ take values in A^n, the $Y_{ik}^{\tau_2}$ take values in B^n, and

$$
\mathbb{P}\big[\, X_{ij}^{\tau_1} = x_{ij}^{\tau_1}, Y_{ik}^{\tau_2} = y_{ik}^{\tau_2} \text{ for all } i,j,k,\tau_1,\tau_2 \,\big|\, U_i = u \big]
$$
$$
= \prod_{j=1}^{M_1} \prod_{\tau_1 \in T_1} p_{1\tau_1}^n(x_{ij}^{\tau_1}|u) \prod_{k=1}^{M_2} \prod_{\tau_2 \in T_2} p_{2\tau_2}^n(y_{ik}^{\tau_2}|u).
$$

The pair $(X_{ij}^{\tau_1}, Y_{ik}^{\tau_2})$ is the pair of random codewords sent if the message triple (i,j,k) is to be transmitted and if transmitter ν knows that the channel is contained in $t_\nu^{-1}(\tau_\nu)$ for $\nu = 1, 2$.

Before we can define the corresponding decoding sets, we need to introduce some more notation. Let G be a finite set and let $x = (x_1, \ldots, x_n) \in G^n$. The type of x then is the probability measure $p_x \in \mathcal{P}(G)$ given by

$$
p_x(g) := \frac{1}{n} |\{i : x_i = g\}|.
$$

For $p \in \mathcal{P}(G)$ and $\delta > 0$, let $T_{p,\delta}^n$ be the subset of $x \in G^n$ such that $|p_x(g) - p(g)| \leq \delta$ for all $g \in G$ and such that $p(g) = 0$ implies $p_x(g) = 0$. (For more on types etc., see [2].)

Now, for the message triple (i,j,k) and for $\eta \in H_{\tau_1\tau_2}$, define the set

$$
E_{ijk}^\eta := \{z \in C^n : (U_i, X_{ij}^{\tau_1}, Y_{ik}^{\tau_2}, z) \in T_{r_\eta,\delta}^n\},
$$

where the probability distribution r_η is defined as in (1). Then, the decoding set for the message triple (i,j,k) is defined as

$$
C_{ijk} := \left(\bigcup_{\eta \in H} E_{ijk}^\eta \right) \setminus \left(\bigcup_{\substack{(i',j',k') \\ \neq (i,j,k)}} \bigcup_{\eta \in H} E_{i'j'k'}^\eta \right). \tag{5}
$$

Note that the C_{ijk} depend on the families (4), but not on η.

Up to now, we have defined the random code. In the following, we bound the expected average error from above. Once that is done, it will be easy to infer the existence of a good code. We define the random variable P_e^η depending on the families (4) and giving the average probability of error when channel $\eta \in H_{\tau_1\tau_2}$ is used,

$$
P_e^\eta := \frac{1}{M_0 M_1 M_2} \sum_{i,j,k} W_\eta^n \left(C_{ijk}^c | X_{ij}^{\tau_1}, Y_{ik}^{\tau_2} \right).
$$

Denote by $\mathbb{E}$ the expectation with respect to $\mathbb{P}$. By (5) and using the symmetry in (i,j,k), the mean average probability of making an error if channel $\eta \in H_{\tau_1\tau_2}$ is used can be estimated as follows:

$$
\mathbb{E}[P_e^\eta] = \mathbb{E}\left[\frac{1}{M_0 M_1 M_2} \sum_{i,j,k} W_\eta^n \left(C_{ijk}^c | X_{ij}^{\tau_1}, Y_{ik}^{\tau_2} \right) \right]
$$
$$
= \mathbb{E}\left[W_\eta^n \left(C_{111}^c | X_{11}^{\tau_1}, Y_{11}^{\tau_2} \right) \right]
$$
$$
\leq \mathbb{E}\left[W_\eta^n \left((E_{111}^\eta)^c | X_{11}^{\tau_1}, Y_{11}^{\tau_2} \right) \right] \tag{6}
$$
$$
+ \sum_{\substack{i,j,k: \\ i \neq 1}} \sum_{\tilde{\eta} \in H} \mathbb{E}\left[W_\eta^n \left(E_{ijk}^{\tilde{\eta}} | X_{11}^{\tau_1}, Y_{11}^{\tau_2} \right) \right] \tag{7}
$$
$$
+ \sum_{\substack{j,k: \\ j \neq 1 \neq k}} \sum_{\tilde{\eta} \in H} \mathbb{E}\left[W_\eta^n \left(E_{1jk}^{\tilde{\eta}} | X_{11}^{\tau_1}, Y_{11}^{\tau_2} \right) \right] \tag{8}
$$
$$
+ \sum_{j \neq 1} \sum_{\tilde{\eta} \in H} \mathbb{E}\left[W_\eta^n \left(E_{1j1}^{\tilde{\eta}} | X_{11}^{\tau_1}, Y_{11}^{\tau_2} \right) \right] \tag{9}
$$
$$
+ \sum_{k \neq 1} \sum_{\tilde{\eta} \in H} \mathbb{E}\left[W_\eta^n \left(E_{11k}^{\tilde{\eta}} | X_{11}^{\tau_1}, Y_{11}^{\tau_2} \right) \right]. \tag{10}
$$

For each $\tau = (\tau_1, \tau_2)$ and for each $\eta \in H_\tau$, define a quadruple of generic random variables $(U, X_{\tau_1}, Y_{\tau_2}, Z_\eta)$ with values in $D \times A \times B \times C$ and joint distribution r_η. We use the following notation: if G is the Cartesian product of the elements of a subset of the set $\{A, B, C, D\}$, denote the projection onto the corresponding components by π_G. The marginal of r_η on G then is $r_\eta \circ \pi_G^{-1}$. We use analogous notation for r_η^n. For bounding (6), note that Lemma 3 (Appendix) implies

$$
\mathbb{E}\left[W_\eta^n \left((E_{111}^\eta)^c | X_{11}^{\tau_1}, Y_{11}^{\tau_2} \right) \right]
$$
$$
= 1 - r_\eta^n(T_{r_\eta,\delta}^n)
$$
$$
\leq (n+1)^{|A||B||C||D|} 2^{-nc\delta^2}.
$$

Set $\tilde{\tau} := (\tilde{\tau}_1, \tilde{\tau}_2)$. For bounding the term in (7) corresponding to $\tilde{\eta} \in H_{\tilde{\tau}_1 \tilde{\tau}_2}$, write

$$
\mathbb{E}\left[W_\eta^n \left(E_{ijk}^{\tilde{\eta}} | X_{11}^{\tau_1}, Y_{11}^{\tau_2} \right) \right] = \sum_{(u,x,y,z) \in T_{r_{\tilde{\eta}},\delta}^n}
$$
$$
r_{\tilde{\eta}}^n \left(\pi_{D^n \times A^n \times B^n}^{-1}(u,x,y) \right) r_\eta^n \left(\pi_{C^n}^{-1}(z) \right).
$$

$(u,x,y,z) \in T_{r_{\tilde{\eta}},\delta}^n$ implies $(u,x,y) \in T_{r_{\tilde{\eta}} \circ \pi_{D \times A \times B}^{-1}, |C|\delta}^n$ and $z \in T_{r_{\tilde{\eta}} \circ \pi_C^{-1}, |D||A||B|\delta}^n$. Thus, by Lemma 1 a) (Appendix),

$$
r_{\tilde{\eta}}^n \left(\pi_{D^n \times A^n \times B^n}^{-1}(u,x,y) \right) \leq 2^{-n\left(H(U, X_{\tilde{\tau}_1}, Y_{\tilde{\tau}_2}) - \zeta_{11} \right)}
$$

and

$$
r_\eta^n \left(\pi_{C^n}^{-1}(z) \right) \leq 2^{-n\left(H(Z_{\tilde{\eta}}) - \zeta_{12} \right)}.
$$

Lemma 2 (Appendix) implies

$$
|T_{r_{\tilde{\eta}},\delta}^n| \leq (n+1)^{|D||A||B||C|} 2^{n(H(U, X_{\tilde{\tau}_1}, Y_{\tilde{\tau}_2}, Z_{\tilde{\eta}}) + \zeta_{13})}.
$$

Thus, each term in (7) with $\tilde{\eta} \in H_{\tilde{\tau}_1 \tilde{\tau}_2}$ can be bounded by

$$
(n+1)^{|D||A||B||C|}
$$
$$
2^{-n\left(H(U, X_{\tilde{\tau}_1}, Y_{\tilde{\tau}_2}) + H(Z_{\tilde{\eta}}) - H(U, X_{\tilde{\tau}_1}, Y_{\tilde{\tau}_2}, Z_{\tilde{\eta}}) - \zeta_1 \right)}
$$
$$
= (n+1)^{|D||A||B||C|} 2^{-n\left(I(Z_{\tilde{\eta}}; X_{\tilde{\tau}_1}, Y_{\tilde{\tau}_2}) - \zeta_1 \right)}, \tag{11}
$$

where $\zeta_1 = \zeta_1(|A|, |B|, |C|, |D|, \delta)$ tends to zero as δ tends to zero. Equality (11) follows from

$$I(Z_{\tilde{\eta}}; U, X_{\tilde{\tau}_1}, Y_{\tilde{\tau}_2}) = I(Z_{\tilde{\eta}}; X_{\tilde{\tau}_1}, Y_{\tilde{\tau}_2}),$$

which is fulfilled because the sequence $(U, (X_{\tilde{\tau}_1}, Y_{\tilde{\tau}_2}), Z_{\tilde{\eta}})$ forms a Markov chain.

The terms (8)-(10) can be bounded in a similar fashion. In the calculations necessary to do so, instead of $r_\eta^n(\pi_{C^n}^{-1}(z))$, there appear conditional probabilities of z, which require using Lemma 1 b). Collecting the bounds thus obtained, and recalling the choice of M_0, M_1, M_2 in (3), one sees that for the mean average probability of error when channel η is used,

$$\mathbb{E}[P_e^\eta] \leq (n+1)^{|D||A||B||C|}(2^{-nc\delta^2}$$
$$+ |H|2^{-n(\min_{\tilde{\tau}_1, \tilde{\tau}_2} \min_{\tilde{\eta} \in H_{\tilde{\tau}_1 \tilde{\tau}_2}} I(Z_{\tilde{\eta}}; X_{\tilde{\tau}_1}, Y_{\tilde{\tau}_2}) - R_0 - R_1 - R_2 + \frac{3\varepsilon}{2} - \zeta_1)}$$
$$+ |H|2^{-n(\min_{\tilde{\tau}_1, \tilde{\tau}_2} \min_{\tilde{\eta} \in H_{\tilde{\tau}_1 \tilde{\tau}_2}} I(Z_{\tilde{\eta}}; X_{\tilde{\tau}_1}, Y_{\tilde{\tau}_2} | U) - R_1 - R_2 + \varepsilon - \zeta_2)}$$
$$+ |H|2^{-n(\min_{\tilde{\tau}_1, \tilde{\tau}_2} \min_{\tilde{\eta} \in H_{\tilde{\tau}_1 \tilde{\tau}_2}} I(Z_{\tilde{\eta}}; X_{\tilde{\tau}_1} | Y_{\tilde{\tau}_2}, U) - R_1 + \frac{\varepsilon}{2} - \zeta_3 \vee \zeta_2)}$$
$$+ |H|2^{-n(\min_{\tilde{\tau}_1, \tilde{\tau}_2} \min_{\tilde{\eta} \in H_{\tilde{\tau}_1 \tilde{\tau}_2}} I(Z_{\tilde{\eta}}; Y_{\tilde{\tau}_2} | X_{\tilde{\tau}_1}, U) - R_2 + \frac{\varepsilon}{2} - \zeta_4 \vee \zeta_2)}).$$

Here, $\zeta_2, \zeta_3, \zeta_4$ have the same properties as ζ_1. Now we are in a position which allows us to choose δ and n such that from the ensemble of random codes, we can extract one deterministic code whose average error probability is exponentially small in blocklength. First, choose δ small enough such that the above sum decreases in n. This is possible by the choice of (R_0, R_1, R_2). Hence there is a $\zeta > 0$ such that

$$\mathbb{E}[P_e^\eta] \leq 2^{-n\zeta} \leq \lambda$$

for all $n \geq n_0(\zeta, \lambda)$. For $\eta \in H_{\tau_1 \tau_2}$, $0 < \tilde{\zeta} < \zeta$, set

$$A_\eta := \left\{ \omega \in \Omega : P_e^\eta(\omega) \leq 2^{-n\tilde{\zeta}} \right\}.$$

This is the set of randomly generated codes with rate triple (R_0', R_1', R_2') and average probability of error smaller than $2^{-n\tilde{\zeta}}$ if channel η is used. As

$$\mathbb{P}[A_\eta^c] \leq 2^{n\tilde{\zeta}} \mathbb{E}\left[P_e^\eta(\{U_i\}, \{X_{ij}^{\tau_1}\}, \{Y_{ik}^{\tau_2}\})\right]$$

by the Markov inequality, one has the inequality

$$\mathbb{P}\left[\bigcap_{\eta \in H} A_\eta\right] \geq 1 - \sum_{\eta \in H} \mathbb{P}[A_\eta^c]$$
$$\geq 1 - 2^{n\tilde{\zeta}} \sum_{\eta \in H} \mathbb{E}[P_e^\eta] \geq 1 - |H|2^{-n(\zeta - \tilde{\zeta})} > 0,$$

for n large enough. Any code contained in the non-empty intersection $\bigcap_{\eta \in H} A_\eta$ is a code$(n, R_0', R_1', R_2', \lambda)$. Thus the rate triple (R_0, R_1, R_2) is achievable, and as the set of achievable rates is closed, the whole set (2) is achievable. Varying $(D, p_0) \in \Pi_0$ and $(p_1, p_2) \in \Pi(D)$ concludes the proof of the achievability part of Theorem 1 for no CSIR and finite $|H|$.

In the case of an arbitrary number of channels, the achievability can be established using a refinement of the argument of [1]. Here, it is crucial that T_1 and T_2 are finite sets and that the average error probability can be made to decrease exponentially in blocklength for the finitely many channels case.

IV. CONVERSE

In this section, we prove a weak converse for Theorem 1. For this, we need to show that for every $\varepsilon > 0$ there is a $\lambda(\varepsilon)$ such that the rate triple of every code$(n, R_1', R_2', R_3', \lambda)$ is within distance ε from $\mathcal{C}^*$ if $\lambda \leq \lambda(\varepsilon)$ and n is large. We assume full CSIR (cf. Remark 1). That means that we can set $R = H$ and $r = $ identity. Instead of $r(\eta)$, we just write η.

Let $\varepsilon > 0$. Let $n \in \mathbb{N}$ and let a code$(n, R_0', R_1', R_2', \lambda)$ for the compound DM-MAC $\mathcal{W}(H, t_2, t_2)$ be given. Define (M_0, M_1, M_2) by $R_\nu' = (1/n)\log M_\nu$ for $\nu = 0, 1, 2$. The code has the form

$$\left(\left(\{x_{ij}^{\tau_1}\}\right)_{\tau_1 \in T_1}, \left(\{y_{ik}^{\tau_2}\}\right)_{\tau_2 \in T_2}, \left(\{C_{ijk}^\eta\}\right)_{\eta \in H} \right), \quad (12)$$

where $x_{ij}^{\tau_1} \in A^n$, $y_{ik}^{\tau_2} \in B^n$, $C_{ijk}^\eta \subset C^n$, and $C_{ijk}^\eta \cap C_{i'j'k'}^\eta = \emptyset$ if $(i, j, k) \neq (i', j', k')$, and where in each of the sets in (12), i, j, k vary over $[1, M_0] \times [1, M_1] \times [1, M_2]$.

We need to define an appropriate probability space again. Here, the probability space reflects the fact that the performance criterion of average error corresponds to a uniform distribution of the messages. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space on which the following random variables are defined: let (V_0, V_1, V_2) be a triple of random variables uniformly distributed in $[1, M_0] \times [1, M_1] \times [1, M_2]$. Given (V_0, V_1, V_2) and $\tau := (\tau_1, \tau_2) = (t_1(\eta), t_2(\eta))$, let (X^{τ_1}, Y^{τ_2}) be the pair of random variables on $A^n \times B^n$ that deterministically chooses the pair of codewords encoding the message triple (V_0, V_1, V_2) with channel knowledge τ, so $(X^{\tau_1}, Y^{\tau_2}) = (x_{V_0 V_1}^{\tau_1}, y_{V_0 V_2}^{\tau_2})$. Finally, given (X^{τ_1}, Y^{τ_2}), let Z^η be distributed on C^n according to

$$\mathbb{P}[Z^\eta = z | X^{\tau_1} = x, Y^{\tau_2} = y, V_0 = i, V_1 = j, V_2 = k]$$
$$= W_\eta^n(z|x, y),$$

where $x \in A^n$ and $y \in B^n$.

We present the complete calculations only for the upper bound on R_1'. The bounds on R_2', $R_1' + R_2'$, and on $R_0' + R_1' + R_2'$ can be derived in a similar fashion. Fix a $\tau = (\tau_1, \tau_2) \in T_1 \times T_2$. For every $\eta \in H_{\tau_1 \tau_2}$, define a channel $W_{\eta,1}^n$ with input alphabet $[1, M_1]$ and with output alphabet C^n by setting

$$W_{\eta,1}^n(z|j) := \mathbb{P}[Z^\eta | V_1 = j].$$

Then, consider the average error probability λ_1^η defined by

$$\frac{1}{M_1} \sum_j W_{\eta,1}^n\left(\bigcup_{i,k} C_{ijk}^\eta \middle| j\right) = 1 - \lambda_1^\eta.$$

Fano's inequality (Theorem 7.4.1 in [8]) implies

$$H(V_1 | Z^\eta, V_0, V_2) \leq H(V_1 | Z^\eta) \leq \lambda_1^\eta \log(M_1 - 1) + h(\lambda_1^\eta), \quad (13)$$

where h denotes binary entropy. The uniform distribution of the messages and (13) imply

$$\log M_1 \leq h(\lambda_1^\eta) + \lambda_1^\eta \log M_1 + I(Z^\eta; X^{\tau_1} | V_0, Y^{\tau_2}). \quad (14)$$

The next step is to obtain a single-letter description from (14) by considering the codewords and the output coordinate-wise:

$$X_{\tau_1} = (X_1^{\tau_1}, \ldots, X_n^{\tau_1}),$$
$$Y_{\tau_2} = (Y_1^{\tau_2}, \ldots, Y_n^{\tau_2}),$$
$$Z^\eta = (Z_1^\eta, \ldots, Z_n^\eta),$$
$$Z_{[1,m]}^\eta = (Z_1^\eta, \ldots, Z_m^\eta), \quad (m \le n);$$

Consider the mutual information term in (14). By the chain rule for entropy, it can be expanded as

$$I(Z^\eta; X^{\tau_1}|V_0, Y^{\tau_2}) = \sum_{m=1}^{n} \Big(H(Z_m^\eta|V_0, Y^{\tau_2}, Z_{[1,m-1]}^\eta) $$
$$- H(Z_m^\eta|X_{\tau_1}, V_0, Y^{\tau_2}, Z_{[1,m-1]}^\eta)\Big). \quad (15)$$

As $(T_0, Y_m^{\tau_2})$ is a function of $(T_0, Y^{\tau_2}, Z_{[1,m-1]}^\eta)$, one has

$$H(Z_m^\eta|V_0, Y^{\tau_2}, Z_{[1,m-1]}^\eta) \le H(Z_m^\eta|V_0, Y_m^{\tau_2}). \quad (16)$$

Further, note that given X^{τ_1} and Y^{τ_2}, the random variables Z_m^η and $Z_{[1,m-1]}^\eta$ are independent. Thus

$$H(Z_m^\eta|X^{\tau_1}, V_0, Y^{\tau_2}, Z_{[1,m-1]}^\eta)$$
$$= H(Z_m^\eta|X_m^{\tau_1}, V_0, Y_m^{\tau_2}). \quad (17)$$

(16) and (17) imply

$$\frac{1}{n}I(Z^\eta; X^{\tau_1}|V_0, Y^{\tau_2}) \le \frac{1}{n}\sum_{m=1}^{n} I(Z_m^\eta; X_m^{\tau_1}|V_0, Y_m^{\tau_2}). \quad (18)$$

Define the following $(D, p_0) \in \Pi_0$ and $(p_1, p_2) \in \Pi(D)$: let

$$D := [1, n] \times [1, M_0];$$
$$p_0(m, i) := \frac{1}{nM_0};$$
$$p_{1\tau_1}(a|m, i) := \mathbb{P}[X_m^{\tau_1} = a|V_0 = i];$$
$$p_{2\tau_2}(b|m, i) := \mathbb{P}[Y_m^{\tau_2} = b|V_0 = i].$$

For $\eta \in H_\tau$, let $U, X_{\tau_1}, Y_{\tau_2}, Z_\eta$ be random variables defined on a probability space $(\Omega', \mathcal{B}, \mathbb{Q})$ and taking values in D, A, B, and C, respectively, such that

$$\mathbb{Q}[U = (m, i), X_{\tau_1} = a, Y_{\tau_2} = b, Z_\eta = c]$$
$$= p_0(m, i)p_{1\tau_1}(a|m, i)p_{2\tau_2}(b|m, i)W_\eta(c|a, b). \quad (19)$$

Then, by (18) and (19),

$$\frac{1}{n}I(Z^\eta; X^{\tau_1}|V_0, Y^{\tau_2}) = I(Z_\eta; X_{\tau_1}|Y_{\tau_2}, U). \quad (20)$$

Now observe that $\lambda_1^\eta \le \lambda$. From the definition of the M_ν and from (14) and (20),

$$R_1' \le \frac{h(\lambda)}{n} + \lambda R_1' + I(Z_\eta; X_{\tau_1}|Y_{\tau_2}, U) + \varepsilon.$$

Now one sees easily that (R_0', R_1', R_2') is close to $\mathcal{C}^*$ for large n: assume that $n \ge (\log 2)/\varepsilon$. For this n, assume that $\lambda < \varepsilon/(R_0' + R_1' + R_2')$. Imitating the above procedure for R_2',
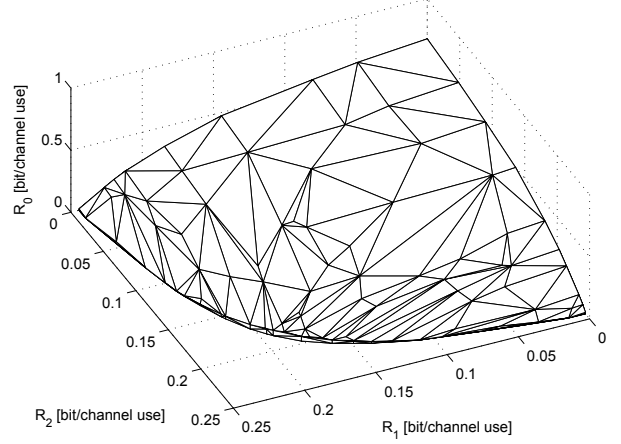

Fig. 2. The capacity region of the compound channel $\{W_1, W_2\}$ with common message.

$R_1' + R_2'$, and $R_0' + R_1' + R_2'$, one sees that all of the following equations must hold:

$$R_1' \le I(Z_\eta; X_{\tau_1}|Y_{\tau_2}, U) + 3\varepsilon; \quad (21)$$
$$R_2' \le I(Z_\eta; Y_{\tau_2}|X_{\tau_1}, U) + 3\varepsilon; \quad (22)$$
$$R_1' + R_2' \le I(Z_\eta; X_{\tau_1}, Y_{\tau_2}|U) + 4\varepsilon; \quad (23)$$
$$R_0' + R_1' + R_2' \le I(Z_\eta; X_{\tau_1}, Y_{\tau_2}) + 5\varepsilon. \quad (24)$$

(21)-(24) must hold for all $\eta \in H_{\tau_1\tau_2}$, and for all $(\tau_1, \tau_2) \in T_1 \times T_2$. Thus the triple (R_0', R_1', R_2') must be within distance 3ε of the set

$$\bigcap_{(\tau_1, \tau_2)} \bigcap_{\eta \in H_{\tau_1\tau_2}} \mathcal{R}(p_0, p_1, p_2, \tau_1, \tau_2, \eta).$$

Choosing $p_{1\tau_1}$ and $p_{2\tau_2}$ for every $(\tau_1\tau_2)$ defines a pair of functions $(p_1, p_2) \in \Pi(D)$. As $(D, p_0) \in \Pi_0$, one obtains that the rate triple (R_0', R_1', R_2') must be within distance 3ε from the asserted region, thus proving the weak converse for Theorem 1 in the case of full CSIR. Together with Remark 1, this completes the proof of Theorem 1.

V. EXAMPLE

Here we give an example of a compound DM-MAC consisting of two channels, i.e. $H = \{1, 2\}$. For the input and output alphabets we set $A = B = C = \{0, 1\}$. Writing the input combination (a, b) in row $2a + b + 1$, the transmission matrices are

$$W_1 = \begin{pmatrix} .9 & .1 \\ .4 & .6 \\ .6 & .4 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad W_2 = \begin{pmatrix} .9 & .1 \\ .6 & .4 \\ .4 & .6 \\ 0 & 1 \end{pmatrix}.$$

We assume no CSI.

The rate region of the compound channel $\{W_1, W_2\}$ is shown in Figure 2. Its projection into the (R_1, R_2)-plane equals the capacity region of the compound DM-MAC
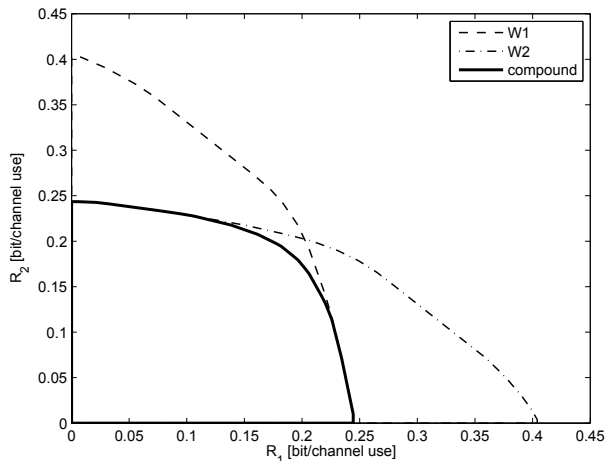
Fig. 3. The (R_1, R_2)-projections of the capacity regions of the constituent channels W_1 and W_2 and of the compound channel $\{W_1, W_2\}$.

$\{W_1, W_2\}$ without common message. Figure 3 shows this projection together with the projections of the capacity regions of the constituent DM-MACs W_1 and W_2. Note that the projection corresponding to the compound channel is strictly contained in the intersection of the projections of the constituent channels' rate regions. This fact carries over to the complete rate regions, but cannot be illustrated well.

VI. DISCUSSION

We have characterized the capacity region of the compound DM-MAC with common message and with partial CSI. Partial CSI here means for the senders that each of them knows which element of a finite partition of the set of channels the channel actually used belongs to. CSIR can be arbitrary. This definition can incorporate asymmetric CSIT. Further, by gradually refining the partitions of the set of channels, one sees how the capacity region scales with increasing CSIT. CSIR has no influence on the achievable rate region.

Apart from the relevance this result has on its own, it is a first step towards the capacity region of the compound DM-MAC with partial CSI and with conferencing encoders. In the conferencing encoders case, the transmitters can communicate via a noise-free finite-capacity link in order to coordinate their transmission to some extent. The resulting capacity region lies between the one obtained with full cooperation and the region where the transmitters cannot cooperate at all. The characterization of this region will be a contribution to the analysis of future wireless systems which incorporate base station cooperation via backbones with high, but finite capacities.

APPENDIX

Lemma 1. *a) Let A be a finite set. Let $p, \tilde{p} \in \mathcal{P}(A)$. Let $0 < \delta < 1/(2|A|)$. Then, for all $n \in \mathbb{N}$, for every $x \in T_{\tilde{p},\delta}^n$,*

$$p^n(x) \leq 2^{-n(H(\tilde{p}) - \phi(|A|,\delta))}.$$

ϕ *is a universal function (i. e. independent of everything), positive if $|A| \geq 1$ and $0 < \delta < 1$, and $\lim_{\delta \to 0} \phi(|A|, \delta) = 0$ for all values of $|A|$.*

b) Let A, B be finite sets. Let $p \in \mathcal{P}(A)$ and $W, \tilde{W}$ stochastic matrices with input alphabet A and output alphabet B. Let $0 < \delta < 1/(2|A||B|)$. Let $\tilde{r} \in \mathcal{P}(A \times B)$ be the joint distribution corresponding to p and $\tilde{W}$. Then, for all $n \in \mathbb{N}$, for all $(x, y) \in T_{\tilde{r}, \delta}^n$,

$$W^n(y|x) \leq 2^{-n(H(\tilde{W}|p) - \psi(\delta, |A|, |B|))}.$$

ψ *is a universal function (i. e. independent of everything), positive if $|A|, |B| \geq 1$, $0 < \delta < 1$, and $\lim_{\delta \to 0} \psi(\delta, |A|, |B|) = 0$ for arbitrary $|A|, |B|$.*

Proof: From [2], Lemma I.2.6 and I.2.7. ∎

Lemma 2. *Let A be a finite set. Let $p \in \mathcal{P}(A)$. Let $0 < \delta < 1/(2|A|)$. Then*

$$|T_{p,\delta}^n| \leq (n+1)^{|A|} 2^{n(H(p) + \phi(|A|,\delta))}.$$

for the same ϕ as in Lemma 1 a).

Proof: Like the proof of [2], Lemma I.2.13. ∎

Lemma 3 ([4], Lemma III.1.3)**.** *Let A be a finite set and let $p \in \mathcal{P}(A)$. Then, there is a universal constant $c > 0$ such that*

$$p^n((T_{p,\delta}^n)^c) \leq (n+1)^{|A|} 2^{-nc\delta^2}.$$

REFERENCES

[1] D. Blackwell, L. Breiman, and A. J. Thomasian, "The Capacity of a Class of Channels," *Ann. Math. Statist.*, vol. 30, no. 4, pp. 1229–1241, 1959.

[2] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, New York: Academic, 1982.

[3] I. Marić, R. D. Yates, and G. Kramer, "Capacity of Interference Channels With Partial Transmitter Cooperation," *IEEE Trans. Inf. Theory*, vol. 53,, no. 10, pp. 3536–3548, 2007.

[4] P. C. Shields, *The Ergodic Theory of Discrete Sample Paths*, Providence: American Mathematical Society, 1996.

[5] D. Slepian and J. K. Wolf, "A Coding Theorem for Multiple Access Channels With Correlated Sources," *Bell Syst. Tech. J.*, vol. 52, pp. 1037–1076, 1973

[6] F. M. J. Willems, "Information-Theoretical Results for the Discrete Memoryless Multiple Access Channel," Ph.D. dissertation, Katholieke Universiteit Leuven, Belgium, 1982.

[7] F. M. J. Willems, "The discrete memoryless multiple channel with partially cooperating encoders," *IEEE Trans. Inf. Theory*, vol. 29, no. 3, pp. 441–445, 1983.

[8] J. Wolfowitz, *Coding Theorems of Information Theory*, 3rd edition, Berlin, Heidelberg: Springer, 1978.