The Multiple-Access Channel with Entangled Transmitters

Uzi Pereg, *Member, IEEE*, Christian Deppe, *Member, IEEE*, and Holger Boche, *Fellow, IEEE*

Abstract

Communication over a classical multiple-access channel (MAC) with entanglement resources is considered, whereby two transmitters share entanglement resources a priori before communication begins. Leditzky et al. (2020) presented an example of a classical MAC, defined in terms of a pseudo telepathy game, such that the sum rate with entangled transmitters is strictly higher than the best achievable sum rate without such resources. Here, we establish inner and outer bounds on the capacity region for the *general* MAC with entangled transmitters, and show that the previous result can be obtained as a special case. It has long been known that the capacity region of the classical MAC under a message-average error criterion can be strictly larger than with a maximal error criterion (Dueck, 1978). We observe that given entanglement resources, the regions coincide. Furthermore, we address the combined setting of entanglement resources and conferencing, where the transmitters can also communicate with each other over rate-limited links. Using superdense coding, entanglement can double the conferencing rate.

I. INTRODUCTION

Cooperation in quantum communication networks has gained considerable attention recently driven by both experimental advancements and theoretical discoveries [1–4]. The multiple-access channel (MAC) is a fundamental model for network communication. Quantum resources in communication over the MAC are considered in the literature in various settings. Winter [5] derived a regularized characterization for the classical capacity region of the quantum MAC. Furthermore, the authors of the present paper [6] considered the quantum MAC with cribbing encoders, whereby Transmitter 2 has access to (part of) the environment of Transmitter 1. Boche and Nötzel [7] studied the cooperation setting of a classical-quantum MAC with conferencing encoders, where the encoders exchange messages between them in a constant rate (see also [8]). Hsieh et al. [9] and Shi et al. [10] addressed the model where each transmitter shares entanglement resources with the receiver independently.

Leditzky et al. [11] presented an example of a classical MAC such that sharing entanglement between transmitters strictly increases the achievable sum rate (see also [12–14]). Recently, the classical upper bound has been improved showing that the sum rate increases from at most 3.02 to 3.17 bits per transmission [15]. The channel construction in [11] is based on a pseudo-telepathy game [16] where quantum strategies guarantee a certain win and outperform classical strategies. Additional observations are developed in [17] as well. Fawzi and Fermé [18] have also established separation in a more fundamental example, the binary adder channel, when the transmitters are provided with non-singaling correlations [19]. By analyzing the zero-error capacity region, it was shown that the sum rate increases from 1.5 without correlation resources, to 1.5425 bits per transmission in this case [18]. We have shown that the dual property does not hold for the broadcast channel, i.e., entanglement between receivers does not increase achievable rates [20].

The potential benefits of the sixth generation of cellular networks (6G) are significant, with anticipated improvements in latency, resilience, computation power, and trustworthiness for future communication systems, such as the tactile internet [21], which involves not only data transfer but also the control of physical and virtual objects. Resilience plays a very important role in achieving trustworthiness [22]. Quantum resources are expected to play a crucial role in achieving these gains, as highlighted by [23] and [21]. By enabling cooperation between trusted hardware and software components, future communication systems could isolate untrusted components and substantially reduce the attack surface of the communication system [21, 24]. In particular, cooperation in the form of conferencing encoders can significantly increase the resilience of the communication system to jamming attacks. In [25], MACs were investigated for which without conferencing no communication is possible in the case of adversely jamming attacks, but even with arbitrarily small positive conferencing rates reliable communication between the sender and the receiver is guaranteed. Conferencing can thus prevent denial-of-service attacks for these channels. For a complete characterisation of these scenarios, see [25, 26]. Quantum resources and cooperation also offer additional advantages, such as improved performance gains for communication tasks and the reduction of the attack surface, making them highly promising for 6G networks, as discussed in [27] and [28]. Investigating communication with cooperation in the form of entanglement resources could lead to more efficient protocols for future applications, making this an interesting avenue for further research. Cooperation and entanglement are important techniques to meet the very demanding requirements of 6G and future communication systems. The potential of these techniques for 6G has not really been explored in the literature until now. This paper aims to initialize this research direction. We only consider senders and receivers that are trustworthy, i.e. that adhere to the specified communication protocol. Especially in complexity theory, other interactive communication scenarios
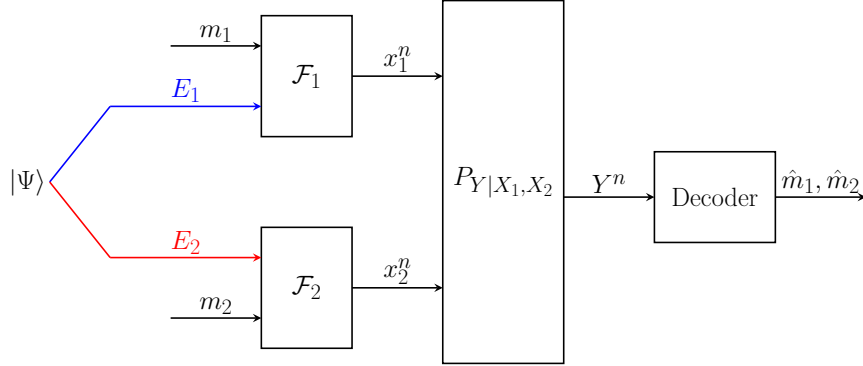
Fig. 1. The classical multiple-access channel $P_{Y|X_1,X_2}$ with pre-shared entanglement resources between the transmitters. The entanglement resources (quantum systems) of Transmitter 1 and Transmitter 2 are marked in red and blue, respectively.

have also been examined with regard to their complexity behavior. This research direction is also interesting for communication with manipulative communication parties and should also be investigated for 6G applications in the future.

Here, we consider communication over a two-user classical MAC with entanglement resources shared between the transmitters, a priori before communication begins. We establish inner and outer bounds on the capacity region for the *general* MAC with entangled transmitters (see Figure 1), and show that the result by Leditzky et al. [11] can be obtained as a special case. We also point out the following change of behavior. In general, achievable communication rates may also depend on the error criterion. Without entanglement resources, Dueck [29] showed that the relaxation of a message-average error criterion can lead to strictly higher achievable rates, when compared with a maximal error criterion. Here, however, we show that the capacity region with entangled transmitters remains the same, whether we consider a message-average or a maximal error criterion. Furthermore, we consider the combined setting of entanglement resources together with conferencing between the transmitters (see Figures 2 and 3). In this setting, in addition to pre-shared entanglement resources, the transmitters can now communicate with each other over either classical or quantum noise-free links at a limited rate. In the classical conferencing setting, we derive inner and outer bounds on the capacity region of the MAC with entangled transmitters. In the quantum conferencing setting, we establish an achievable region and show that entanglement can double the conferencing rate through superdense coding.

The remainder of this paper is organized as follows. Section II reviews related work on the MAC with classical correlation between the transmitters. Section III includes basic definitions and the model description. In Section IV, we present our main results on the classical MAC with entangled transmitters, where the entanglement resources can be either unlimited (Section IV-A), or at a limited rate (IV-B). We examine the example by Leditzky et al. [11] in Section IV-C, and consider the effect of classical and quantum conferencing on the capacity region in Section V. We prove the main results in Sections VI-A-IX. In Section VI-A, we show that our formulas are exhausted with finite-cardinality auxiliary variables. In Sections VII and VIII, we prove the capacity theorems for entangled transmitters, and in Section IX, for both entangled and conferencing transmitters. In Section X, we summarize and discuss open directions, and in the appendix, we show that the capacity regions associated with the average and maximal error criteria are identical.

II. Related Work

We briefly review the related work on the multiple-access channel (MAC) with classically correlated transmitters. In communication, shared entanglement is often thought of as the quantum version of common randomness (CR), which is a random key that is shared between the two parties. Indeed, generating CR is one of the most fundamental uses of entanglement. The MAC setting provides yet another demonstration of how entanglement is much stronger than the classical CR correlation. We will come back to this point later on (see Remark 7).

A. Basic Definitions for Classical Systems

We use the following notation conventions for classical systems. Script letters $\mathcal{X}, \mathcal{Y}, ...$ are used for finite sets. Lowercase letters $x, y, \ldots$ represent constants and values of classical random variables, and uppercase letters $X, Y, \ldots$ represent classical random variables. We use $x^r = (x[i])_{i=1}^r$ to denote a sequence of letters from $\mathcal{X}$, where r is a positive integer. The distribution of a discrete random variable X is specified by a probability mass function (pmf) $p_X(x)$ over a finite set $\mathcal{X}$. The respective

i.i.d. distribution is denoted by p_X^n, namely $p_X^n(x^n) = \prod_{i=1}^n p_X(x[i])$. A random sequence X^n and its distribution $p_{X^n}(x^n)$ are defined accordingly. The typical set $\mathcal{A}_\delta^{(n)}(p_X)$ is the set of sequences $x^n \in \mathcal{X}^n$ such that $\left| p_X(a) - \frac{1}{n}N(a|x^n) \right| \le \delta \cdot p_X(a)$, for every $a \in \mathcal{X}$, where $N(a|x^n)$ is the number of occurrences of the letter $a \in \mathcal{X}$ in the sequence x^n. The definition is extended to a joint distribution $p_{X,Y}$ in a natural manner.

A discrete memoryless multiple-access channel (MAC) $(\mathcal{X}_1, \mathcal{X}_2, P_{Y|X_1,X_2}, \mathcal{Y})$ consists of finite input alphabets $\mathcal{X}_1$, $\mathcal{X}_2$, a finite output alphabet $\mathcal{Y}$, and a collection of conditional probability mass functions $P_{Y|X_1,X_2}(\cdot|x_1,x_2)$ on $\mathcal{Y}$, for every $(x_1, x_2) \in \mathcal{X}_1 \times \mathcal{X}_2$. Having assumed that the channel is memoryless, if the inputs $x_1^n = (x_1[i])_{i=1}^n$ and $x_2^n = (x_2[i])_{i=1}^n$ are sent through n channel uses, then the output is distributed according to $P_{Y|X_1,X_2}^n(y^n|x_1^n, x_2^n) = \prod_{i=1}^n P_{Y|X_1,X_2}(y[i]|x_1[i], x_2[i])$. The transmitters and the receiver are often called Alice 1, Alice 2, and Bob.

B. Coding with Common Randomness

We begin with the definition of a code for the classical MAC $P_{Y|X_1,X_2}$ with classical CR between the transmitters. We denote the random element that is shared between them by z.

Definition 1. A $(2^{nR_1}, 2^{nR_2}, n)$ code for the classical MAC $P_{Y|X_1,X_2}$ with classical CR between the transmitters consists of the following:

- a key set $\mathcal{Z}$ and a probability distribution p_Z on $\mathcal{Z}$;
- two message sets $[1 : 2^{nR_1}]$ and $[1 : 2^{nR_2}]$, assuming 2^{nR_k} is an integer;
- a pair of encoding functions, $\mathsf{e}_k : [1 : 2^{nR_k}] \times \mathcal{Z} \to \mathcal{X}_k^n$ for $k \in \{1, 2\}$; and
- a decoding function $g : \mathcal{Y}^n \to [1 : 2^{nR_1}] \times [1 : 2^{nR_2}]$.

We denote the code by $\mathscr{C} = (P_Z, \mathsf{e}_1, \mathsf{e}_2, g)$.

The communication scheme is depicted in Figure 1. The transmitters share the classical key $z \sim p_Z$. Alice k chooses a message m_k from the message set, $[1 : 2^{nR_k}]$, for $k = 1, 2$. To send the message $m_1 \in [1 : 2^{nR_1}]$, Alice 1 transmits $x_1^n = \mathsf{e}_1(m_1, z)$ through n uses of the classical MAC $P_{Y|X_1,X_2}$. Similarly, Alice 2 transmits $x_2^n = \mathsf{e}_2(m_2, z)$. Bob receives the channel output y^n, and estimates the message pair as $(\hat{m}_1, \hat{m}_2) = g(y^n)$.

The conditional probability of error of the code $\mathscr{C} = (P_Z, \mathsf{e}_1, \mathsf{e}_2, g)$ is

$$P_e^{(n)}(\mathscr{C}|m_1, m_2) = \sum_{z \in \mathcal{Z}} p_Z(z) \sum_{y^n : g(y^n) \ne (m_1, m_2)} P_{Y|X_1,X_2}^n(y^n|\mathsf{e}_1(m_1, z), \mathsf{e}_2(m_2, z)) \tag{1}$$

for $(m_1, m_2) \in [1 : 2^{nR_1}] \times [1 : 2^{nR_2}]$. Hence, the maximal probability of error is

$$P_e^{(n)}(\mathscr{C}) \equiv \max_{m_1, m_2} P_e^{(n)}(\mathscr{C}|m_1, m_2). \tag{2}$$

A rate pair (R_1, R_2) is called achievable with classical CR between the transmitters if for every $\varepsilon > 0$ and sufficiently large n, there exists a $(2^{nR_1}, 2^{nR_2}, n)$ code such that the maximal probability of error is bounded by $P_e^{(n)}(\mathscr{C}) \le \varepsilon$. The capacity region $\mathcal{C}_{\mathrm{CRT}}(P_{Y|X_1,X_2})$ is defined as the closure of the set of achievable pairs (R_1, R_2), where the subscript 'CRT' indicates CR resources between the transmitters.

Remark 1. In general, achievable communication rates may also depend on the error criterion. In particular, we say that a rate pair (R_1, R_2) is achievable with a *message-average* error criterion if for every $\varepsilon > 0$ and sufficiently large n, there exists a $(2^{nR_1}, 2^{nR_2}, n)$ code such that $\overline{P}_e^{(n)}(\mathscr{C}) \le \varepsilon$, where $\overline{P}_e^{(n)}(\mathscr{C})$ denotes the message-average error probability:

$$\overline{P}_e^{(n)}(\mathscr{C}) \equiv \frac{1}{2^{n(R_1+R_2)}} \sum_{m_1=1}^{2^{nR_1}} \sum_{m_2=1}^{2^{nR_2}} P_e^{(n)}(\mathscr{C}|m_1, m_2). \tag{3}$$

The capacity region with a message-average error criterion is defined accordingly. Without assisting resources, the capacity region with a message-average error criterion can be strictly larger than with a maximal error criterion [29] [30, Sec. 2.2]. For a general MAC without assisting resources, the capacity region under a maximal error criterion is unknown. Nevertheless, Cai [30] has shown that given classical CR between the transmitters, this is not the case. That is, the capacity region with classical CR between the transmitters remains the same, whether we consider a message-average or a maximal error criterion.

C. Related Work

The capacity theorem for classically correlated transmitters is given below.

Theorem 1 (see also [30]). The capacity region of a classical MAC $P_{Y|X_1,X_2}$ with classical CR between the transmitters is given by

$$\mathcal{C}_{\mathrm{CRT}}(P_{Y|X_1,X_2}) = \bigcup_{p_{V_0} p_{X_1|V_0} p_{X_2|V_0}} \left\{ \begin{array}{rl} (R_1, R_2) : R_1 & \le I(X_1; Y|X_2 V_0) \\ R_2 & \le I(X_2; Y|X_1 V_0) \\ R_1 + R_2 & \le I(X_1 X_2; Y|V_0) \end{array} \right\} \tag{4}$$

where the union on the right-hand side is over the set of all classical auxiliary variables $V_0 \sim p_{V_0}$ and all classical channels $p_{X_k|V_0}$ for $k \in \{1,2\}$.

We note that given such an auxiliary variable V_0 and encoding channels $p_{X_k|V_0}$, the variables $V_0 \leftarrow\!\!\!\circ (X_1, X_2) \circ\!\!\!\rightarrow Y$ form a Markov chain:

$$p_{V_0, X_1, X_2, Y}(v_0, x_1, x_2, y) = p_{V_0}(v_0) \cdot p_{X_1|V_0}(x_1|v_0)p_{X_2|V_0}(x_2|v_0) \cdot P_{Y|X_1, X_2}(y|x_1, x_2). \tag{5}$$

The auxiliary variable V_0 in the characterization above can thus be interpreted as the CR key that is shared between the transmitters. From a mathematical perspective, the union over p_{V_0} can be viewed as a convex hull operation, hence the capacity region is convex [31].

Remark 2. Consider a message-average error criterion, instead of the maximal error criterion in Definition 1 (see Remark 1). Then, the convex hull above can also be achieved without CR through operational time sharing [31, Sec. 15.3.3]. Thereby, the capacity region with uncorrelated transmitters, i.e., without resources, is *the same* as $\mathcal{C}_{\text{CRT}}(P_{Y|X_1, X_2})$. That is, classical correlation cannot increase the achievable rates, while entanglement can. The advantage of entanglement will be demonstrated in Subsection IV-C, based on the examples by Leditzky et al. [11].

III. DEFINITIONS

A. Basic Definitions

The classical MAC $P_{Y|X_1, X_2}$ is defined as in Subsection II-A. We use the same notation for classical systems as well. In addition, we introduce the following notation for quantum systems. The state of a quantum system A is a density operator ρ on the Hilbert space $\mathcal{H}_A$. The set of all density operators acting on $\mathcal{H}_A$ is denoted by $\mathscr{D}(\mathcal{H}_A)$. Measurement distributions are described in terms of a positive operator-valued measure (POVM), i.e., a set of positive semi-denfinite operators $\{K_x\}_{x \in \mathcal{X}}$ such that $\sum_x K_x = \mathbb{1}$. According to the Born rule, if the system is in state ρ, then the probability to measure x is $p_X(x) = \text{Tr}(K_x \rho)$. A quantum channel $\mathcal{P}_{A \to B}$ is a completely-positive trace-preserving (cptp) linear map from $\mathscr{D}(\mathcal{H}_A)$ to $\mathscr{D}(\mathcal{H}_B)$. A measurement channel is a quantum-classical channel $\mathcal{K}_{A \to X}$, mapping $\rho \mapsto \mathcal{K}_{A \to X}(\rho) = \sum_{x \in \mathcal{X}} \text{Tr}(K_x \rho) |x\rangle\langle x|$, where $\{K_x\}_{x \in \mathcal{X}}$ forms a POVM [32, Def. 4.6.7].

B. Coding with Entangled Transmitters

We consider coding for the classical MAC $P_{Y|X_1, X_2}$ with entanglement resources between the transmitters. We denote the entanglement resources of Transmitter 1 and Transmitter 2 by E_1 and E_2, respectively. Let $\mathcal{H}_{E_1 E_2} \equiv \mathcal{H}_{E_1} \otimes \mathcal{H}_{E_2}$ denote the corresponding Hilbert space.

Definition 2. A $(2^{nR_1}, 2^{nR_2}, n)$ code for the classical MAC $P_{Y|X_1, X_2}$ with entangled transmitters consists of the following:
- a bipartite state $\Psi_{E_1 E_2} \in \mathscr{D}(\mathcal{H}_{E_1 E_2})$ that is shared between the transmitters.
- two message sets $[1 : 2^{nR_1}]$ and $[1 : 2^{nR_2}]$, assuming 2^{nR_k} is an integer;
- two collections of encoding POVMs $\mathcal{F}_1^{(m_1)} = \{F_{x_1^n}^{(m_1)}\}_{x_1^n \in \mathcal{X}_1^n}$ and $\mathcal{F}_2^{(m_2)} = \{F_{x_2^n}^{(m_2)}\}_{x_2^n \in \mathcal{X}_2^n}$ on E_1 and E_2, respectively, one pair of POVMs for each message pair $(m_1, m_2) \in [1 : 2^{nR_1}] \times [1 : 2^{nR_2}]$; and
- a decoding function $g : \mathcal{Y}^n \to [1 : 2^{nR_1}] \times [1 : 2^{nR_2}]$.

We denote the code by $\mathscr{C} = (\Psi, \mathcal{F}_1, \mathcal{F}_2, g)$.

The communication scheme is depicted in Figure 1. The transmitters share the entangled pair $E_1 E_2$. Alice k chooses a message m_k from the message set, $[1 : 2^{nR_k}]$, for $k = 1, 2$. To send the message $m_1 \in [1 : 2^{nR_1}]$, Alice 1 applies the encoding measurement $\mathcal{F}_1^{(m_1)}$ to her share of the entanglement resource, E_1, and obtains a measurement outcome $x_1^n \in \mathcal{X}_1^n$. She sends x_1^n through n uses of the classical MAC $P_{Y|X_1, X_2}$. Similarly, Alice 2 measures the system E_2 using the measurement $\mathcal{F}_2^{(m_2)}$, and transmits the outcome x_2^n. The joint input distribution is thus

$$f(x_1^n, x_2^n | m_1, m_2) = \text{Tr} \left[\left(F_{x_1^n}^{(m_1)} \otimes F_{x_2^n}^{(m_2)} \right) \Psi_{E_1 E_2} \right]. \tag{6}$$

Bob receives the channel output y^n, and estimates the message pair as $(\hat{m}_1, \hat{m}_2) = g(y^n)$.

The conditional probability of error of the code $\mathscr{C} = (\Psi, \mathcal{F}_1, \mathcal{F}_2, g)$ is

$$P_e^{(n)}(\mathscr{C}|m_1, m_2) = \sum_{y^n : g(y^n) \neq (m_1, m_2)} \left[\sum_{(x_1^n, x_2^n) \in \mathcal{X}_1^n \times \mathcal{X}_2^n} f(x_1^n, x_2^n | m_1, m_2) P_{Y|X_1, X_2}^n(y^n | x_1^n, x_2^n) \right] \tag{7}$$

for $(m_1, m_2) \in [1 : 2^{nR_1}] \times [1 : 2^{nR_2}]$. Hence, the maximal probability of error is

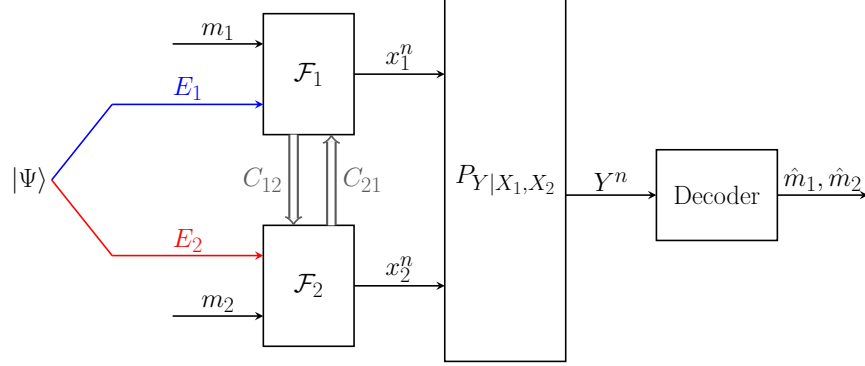$$P_e^{(n)}(\mathscr{C}) \equiv \max_{m_1, m_2} P_e^{(n)}(\mathscr{C}|m_1, m_2). \tag{8}$$

Fig. 2. The classical multiple-access channel $P_{Y|X_1,X_2}$ with pre-shared entanglement resources and classical conferencing between the transmitters. The entanglement resources of Transmitter 1 and Transmitter 2 are marked in blue and red, respectively. The conferencing links are indicated by double-line arrows.

A rate pair (R_1, R_2) is called achievable with entangled transmitters if for every $\varepsilon > 0$ and sufficiently large n, there exists a $(2^{nR_1}, 2^{nR_2}, n)$ code such that the maximal probability of error is bounded by $P_e^{(n)}(\mathscr{C}) \leq \varepsilon$. The capacity region $\mathcal{C}_{\mathrm{ET}}(P_{Y|X_1,X_2})$ of the classical MAC with entangled transmitters is defined as the closure of the set of achievable pairs (R_1, R_2), where the subscript 'ET' indicates entanglement resources between the transmitters.

Remark 3. As pointed out in Remark 1, achievable communication rates may also depend on the error criterion, and the capacity region under a maximal error criterion is unknown for a general MAC without assisting resources. We also mentioned in Remark 2 that Cai [30] has shown that this is not the case when the transmitters share classical correlation. Similarly, the capacity region with entangled transmitters remains the same, whether we consider a message-average or a maximal error criterion. For completeness, we provide the proof for this property in the appendix.

C. Classical Conferencing Cooperation

Now, we consider coding for the classical MAC $P_{Y|X_1,X_2}$ with both entanglement resources and classical conferencing between the transmitters. As before, the transmitters share entanglement resources on a product Hilbert space $\mathcal{H}_{E_1 E_2} \equiv \mathcal{H}_{E_1} \otimes \mathcal{H}_{E_2}$. Now, suppose that given the messages, the transmitters may communicate with each other over noise-free classical links at a limited rate. The link from Transmitter 1 to Transmitter 2 has a classical capacity of C_{12}, and from Transmitter 2 to Transmitter 1 has a classical capacity of C_{21}, where C_{12} and C_{21} are fixed. In this case, we say that the MAC has a classical (C_{12}, C_{21})-permissible conference between the transmitters.

Definition 3. A $(2^{nR_1}, 2^{nR_2}, n)$ code for the classical MAC $P_{Y|X_1,X_2}$ with entangled transmitters and classical conferencing consists of the following:
- a bipartite state $\Psi_{E_1 E_2} \in \mathscr{D}(\mathcal{H}_{E_1 E_2})$ that is shared between the transmitters.
- two message sets $[1 : 2^{nR_1}]$ and $[1 : 2^{nR_2}]$,
- two sets of T classical conferencing functions

$$s_{1\to2}^{(t)} : [1 : 2^{nR_1}] \times \left(\Omega_{2\to1}^{(1)} \times \cdots \times \Omega_{2\to1}^{(t-1)} \right) \to \Omega_{1\to2}^{(t)}, \tag{9}$$

$$s_{2\to1}^{(t)} : [1 : 2^{nR_2}] \times \left(\Omega_{1\to2}^{(1)} \times \cdots \times \Omega_{1\to2}^{(t-1)} \right) \to \Omega_{2\to1}^{(t)} \tag{10}$$

for $t \in [1 : T]$, where $\Omega_{1\to2}^{(t)}$ and $\Omega_{2\to1}^{(t)}$ are finite conferencing alphabets, and $\Omega_{1\to2}^{(0)} = \Omega_{2\to1}^{(0)} = \emptyset$.
- two collections of encoding POVMs $\mathcal{F}_1^{(m_1, \omega_{2\to1}^T)} = \{F_{x_1^n}^{(m_1, \omega_{2\to1}^T)}\}_{x_1^n \in \mathcal{X}_1^n}$ and $\mathcal{F}_2^{(m_2, \omega_{1\to2}^T)} = \{F_{x_2^n}^{(m_2, \omega_{1\to2}^T)}\}_{x_2^n \in \mathcal{X}_2^n}$ on E_1 and E_2, respectively; and
- a decoding function $g : \mathcal{Y}^n \to [1 : 2^{nR_1}] \times [1 : 2^{nR_2}]$.

We denote the code by $\mathscr{C} = (\Psi, s_{1\to2}, s_{2\to1}, \mathcal{F}_1, \mathcal{F}_2, g)$.

The communication scheme is depicted in Figure 2. The transmitters share the entangled pair $E_1 E_2$. Alice k chooses a message m_k from the message set, $[1 : 2^{nR_k}]$, for $k = 1, 2$. Then, Alice 1 and Alice 2 perform a classical conferencing protocol as follows. In round $t = 1$, Alice 1 sends a conferencing message $\omega_{1\to2}[1] = s_{1\to2}^{(1)}(m_1)$ to Alice 2. At the same

time, Alice 2 sends a conferencing message $\omega_{2\to1}[1] = s^{(1)}_{2\to1}(m_2)$ to Alice 1. Next, in round $t \in [2 : T]$, Alice 1 sends a conferencing message $\omega_{1\to2}[t] = s^{(t)}_{1\to2}(m_1, \omega_{2\to1}[1], \ldots, \omega_{2\to1}[t-1])$ to Alice 2, meanwhile Alice 2 sends a conferencing message $\omega_{2\to1}[t] = s^{(t)}_{2\to1}(m_2, \omega_{1\to2}[1], \ldots, \omega_{1\to2}[t-1])$ to Alice 1.

To send the message $m_1 \in [1 : 2^{nR_1}]$ to the receiver, Alice 1 applies the encoding measurement $\mathcal{F}_1^{(m_1, \omega^T_{2\to1})}$ to her share of the entanglement resource, E_1, and obtains a measurement outcome $x_1^n \in \mathcal{X}_1^n$. She sends x_1^n through n uses of the classical MAC $P_{Y|X_1,X_2}$. Similarly, Alice 2 measures the system E_2 using the measurement $\mathcal{F}_2^{(m_2, \omega^T_{1\to2})}$, and transmits the outcome x_2^n. The joint input distribution given the information messages and the conferencing message is then

$$f(x_1^n, x_2^n | m_1, m_2, \omega^T_{1\to2}, \omega^T_{2\to1}) = \text{Tr}\left[\left(F_{x_1^n}^{(m_1, \omega^T_{2\to1})} \otimes F_{x_2^n}^{(m_2, \omega^T_{1\to2})} \right) \Psi_{E_1 E_2} \right]. \tag{11}$$

Note that the measurement of Alice 1 depends on the received conferencing messages $\omega^T_{2\to1}$, which in turn, depends on the message m_2 of Alice 2. Similarly, Alice 2's measurement depends on m_1 through message $\omega^T_{1\to2}$. Therefore, the correlation between the transmitters is aided by both the shared entanglement and classical conferencing. Bob receives the channel output y^n, and estimates the message pair as $(\hat{m}_1, \hat{m}_2) = g(y^n)$.

Let C_{12} and C_{21} be fixed non-negative constants. The code is (C_{12}, C_{21})-permissible if

$$\left| \overline{\Omega}^T_{1\to2} \right| \leq 2^{nC_{12}} \text{ , and } \left| \overline{\Omega}^T_{2\to1} \right| \leq 2^{nC_{21}} \tag{12}$$

where $\overline{\Omega}^T_{k\to j} = \Omega^{(1)}_{k\to j} \times \cdots \times \Omega^{(T)}_{k\to j}$ is the overall conference message set from Alice k to Alice j. We denote the T-sequences of conferencing messages, sent from Alice 1 to Alice 2, and from Alice 2 to Alice 1, by $\omega^T_{1\to2} \equiv \overline{s}^T_{1\to2}(m_1, m_2)$ and $\omega^T_{2\to1} \equiv \overline{s}^T_{2\to1}(m_1, m_2)$, respectively, where

$$\overline{s}^T_{1\to2}(m_1, m_2) = \left[s^{(t)}_{1\to2}\left(m_1, s^{(t-1)}_{2\to1} \right) \right]^T_{t=1}, \tag{13}$$

and

$$\overline{s}^T_{2\to1}(m_1, m_2) = \left[s^{(t)}_{2\to1}\left(m_2, s^{(t-1)}_{1\to2} \right) \right]^T_{t=1}. \tag{14}$$

The conditional probability of error of the code $\mathscr{C} = (\Psi, s_{1\to2}, s_{2\to1}, \mathcal{F}_1, \mathcal{F}_2, g)$ is

$$P_e^{(n)}(\mathscr{C}|m_1, m_2) = \sum_{y^n : g(y^n) \neq (m_1, m_2)} \sum_{(x_1^n, x_2^n) \in \mathcal{X}_1^n \times \mathcal{X}_2^n} f(x_1^n, x_2^n | m_1, m_2, \overline{s}^T_{1\to2}(m_1, m_2), \overline{s}^T_{2\to1}(m_1, m_2))$$
$$\cdot P_{Y|X_1,X_2}^n(y^n | x_1^n, x_2^n) \quad \text{(15)}$$

for $(m_1, m_2) \in [1 : 2^{nR_1}] \times [1 : 2^{nR_2}]$. The maximal probability of error, $P_e^{(n)}(\mathscr{C})$, is defined as before (see (8)).

A rate pair (R_1, R_2) is called achievable with entangled transmitters and classical conferencing if for every $\varepsilon > 0$ and sufficiently large n, there exists a (C_{12}, C_{21})-permissible code such that $P_e^{(n)}(\mathscr{C}) \leq \varepsilon$. The capacity region $\mathcal{C}_{\text{ET-C}}(P_{Y|X_1,X_2}, C_{12}, C_{21})$ is defined as in Section III-B, where the subscript indicates entanglement resources and classical conferencing between the transmitters.

D. Quantum Conferencing with Entanglement

Next, we consider coding for the classical MAC $P_{Y|X_1,X_2}$ with quantum conferencing, while entanglement resources are available at both the conferencing stage and the transmission. Here, we denote the entanglement resources that are being used during the conferencing protocol by $\bar{E}_1$ and $\bar{E}_2$, and during the transmission by E_1 and E_2, corresponding to Transmitter 1 and Transmitter 2, respectively. Hence, the transmitters share entanglement resources on a product Hilbert space $\mathcal{H}_{E_1 \bar{E}_1 E_2 \bar{E}_2}$. Now, suppose that given the messages, the transmitters may communicate with each other over noise-free quantum links at a limited rate with entanglement assistance. The link from Transmitter 1 to Transmitter 2 has a quantum capacity of Q_{12}, and from Transmitter 2 to Transmitter 1 has a quantum capacity of Q_{21}, where Q_{12} and Q_{21} are fixed. In this case, we say that the MAC has a (Q_{12}, Q_{21})-permissible conference between the transmitters.

To simplify the model, we consider a single round of quantum conferencing.

Definition 4. A $(2^{nR_1}, 2^{nR_2}, n)$ code for the classical MAC $P_{Y|X_1,X_2}$ with entanglement-assisted conferencing transmitters consists of the following:

- a bipartite state $\Psi_{E_1 E_2 \bar{E}_1 \bar{E}_2} \in \mathcal{D}(\mathcal{H}_{E_1 E_2 \bar{E}_1 \bar{E}_2})$ that is shared between the transmitters.
- two message sets $[1 : 2^{nR_1}]$ and $[1 : 2^{nR_2}]$,
- two conferencing quantum channels $\mathcal{S}^{(m_1)}_{1\to2} : \mathcal{D}(\mathcal{H}_{\bar{E}_1}) \to \mathcal{D}(\mathcal{H}_{D_{1\to2}})$ and $\mathcal{S}^{(m_2)}_{2\to1} : \mathcal{D}(\mathcal{H}_{\bar{E}_2}) \to \mathcal{D}(\mathcal{H}_{D_{2\to1}})$,
- two collections of encoding POVMs $\mathcal{F}_1^{(m_1)} = \{F_{x_1^n}^{(m_1)}\}_{x_1^n \in \mathcal{X}_1^n}$ and $\mathcal{F}_2^{(m_2)} = \{F_{x_2^n}^{(m_2)}\}_{x_2^n \in \mathcal{X}_2^n}$ on $E_1 D_{2\to1}$ and $E_2 D_{1\to2}$, respectively; and
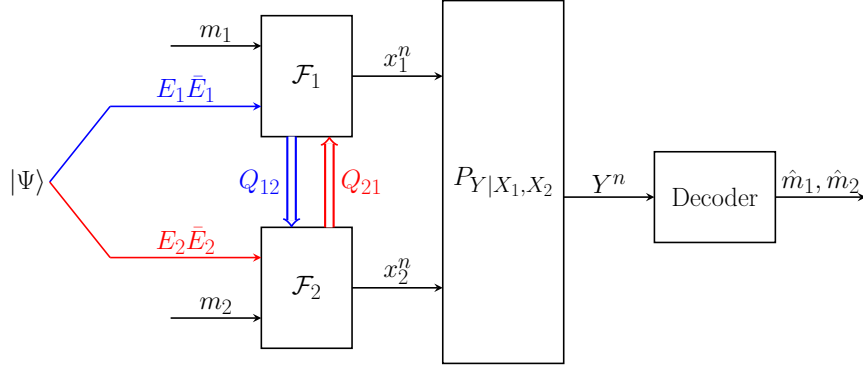
Fig. 3. The classical multiple-access channel $P_{Y|X_1,X_2}$ with pre-shared entanglement resources and quantum conferencing between the transmitters. The quantum systems of Transmitter 1 and Transmitter 2 are marked in blue and red, respectively.

- a decoding function $g : \mathcal{Y}^n \to [1 : 2^{nR_1}] \times [1 : 2^{nR_2}]$.

We denote the code by $\mathscr{C} = (\Psi, \mathcal{S}_{1\to2}, \mathcal{S}_{2\to1}, \mathcal{F}_1, \mathcal{F}_2, g)$.

The communication scheme is depicted in Figure 3. The transmitters share the entangled pair of $E_1\bar{E}_1$ and $E_2\bar{E}_2$. Alice k chooses a message m_k from the message set, $[1 : 2^{nR_k}]$, for $k = 1, 2$. Then, Alice 1 and Alice 2 perform a quantum conferencing protocol as follows. Alice 1 applies the quantum conferencing channel $\mathcal{S}_{1\to2}^{(m_1)}$ to her entangled resource $\bar{E}_1$, and sends the output system $D_{1\to2}$ to Alice 2 through her quantum conferencing link. At the same time, Alice 2 applies $\mathcal{S}_{2\to1}^{(m_2)}$ to $\bar{E}_2$, and sends the output $D_{2\to1}$ to Alice 1. At the end of the quantum conferencing protocol, we have a joint state

$$\sigma_{E_1 E_2 D_{1\to2} D_{2\to1}} = (\mathrm{id}_{E_1 E_2} \otimes \mathcal{S}_{1\to2}^{(m_1)} \otimes \mathcal{S}_{2\to1}^{(m_2)})(\Psi_{E_1 E_2 \bar{E}_1 \bar{E}_2}) \tag{16}$$

Notice that the conferencing protocol exploits the entanglement resources $\bar{E}_1$ and $\bar{E}_2$ to encode $D_{1\to2}$ and $D_{2\to1}$, respectively. The remaining resources, E_1 and E_2, will be used in the next stage.

To send the message $m_1 \in [1 : 2^{nR_1}]$ to the receiver, Alice 1 performs a joint measurement $\mathcal{F}_1^{(m_1)}$ on the entangled resource, E_1, together with the quantum conference system $D_{2\to1}$, and obtains a measurement outcome $x_1^n \in \mathcal{X}_1^n$. She sends x_1^n through n uses of the classical MAC $P_{Y|X_1,X_2}$. Similarly, Alice 2 performs a measurement $\mathcal{F}_2^{(m_2)}$ on E_2 and $D_{1\to2}$, and transmits the outcome x_2^n. The joint input distribution is then

$$f(x_1^n, x_2^n | m_1, m_2) = \mathrm{Tr}\left[\left(F_{x_1^n}^{(m_1)} \otimes F_{x_2^n}^{(m_2)}\right) \sigma_{E_1 E_2 D_{1\to2} D_{2\to1}}\right] . \tag{17}$$

Note that the measurement of Alice 1 includes the conferencing system $D_{2\to1}$, which in turn, is correlated with the message m_2 of Alice 2. Similarly, Alice 2's measurement includes a conferencing system that is correlated with m_1. Therefore, the correlation between the transmitters is aided by both the shared entanglement and quantum conferencing. Bob decodes as usual and estimates $(\hat{m}_1, \hat{m}_2) = g(y^n)$.

Let Q_{12} and Q_{21} be fixed non-negative constants. The code is (Q_{12}, Q_{21})-permissible if

$$\dim(\mathcal{H}_{D_{1\to2}}) \leq 2^{nQ_{12}} , \text{ and } \dim(\mathcal{H}_{D_{2\to1}}) \leq 2^{nQ_{21}} . \tag{18}$$

The maximal probability of error, $P_e^{(n)}(\mathscr{C})$, is defined as before (see (8)).

A rate pair (R_1, R_2) is called achievable with entangled conferencing transmitters if for every $\varepsilon > 0$ and sufficiently large n, there exists a (Q_{12}, Q_{21})-permissible code such that $P_e^{(n)}(\mathscr{C}) \leq \varepsilon$. The capacity region $\mathcal{C}_{\text{ET-Q}}(P_{Y|X_1,X_2}, Q_{12}, Q_{21})$ of the classical MAC with entangled conferencing transmitters is defined as the closure of the set of achievable pairs (R_1, R_2), where the subscript 'ET-Q' indicates entanglement resources and quantum conferencing between the transmitters.

Remark 4. We focus here on two conferencing models. In the discussion section, we will consider other variations (see Subsection X-C).

IV. MAIN RESULTS — ENTANGLED TRANSMITTERS

A. Communication With Unlimited Entanglement Resources

We state our results on the classical MAC $P_{Y|X_1,X_2}$ with entanglement resources between the transmitters. We present inner and outer bounds on the capacity region, as well as a regularized characterization.

1) Inner Bound and Regularized Characterization: Define the rate region $\mathcal{R}_{\mathrm{ET}}(P_{Y|X_1,X_2})$ as follows,

$$\mathcal{R}_{\mathrm{ET}}(P_{Y|X_1,X_2}) = \bigcup_{p_{V_0}p_{V_1|V_0}p_{V_2|V_0},\, \varphi_{A_1A_2},\, \mathcal{L}_1\otimes\mathcal{L}_2} \left\{ \begin{array}{rl} (R_1,R_2): \ R_1 & \leq I(V_1;Y|V_0V_2) \\ R_2 & \leq I(V_2;Y|V_0V_1) \\ R_1+R_2 & \leq I(V_1V_2;Y|V_0) \end{array} \right\}. \tag{19}$$

The union on the right-hand side of (19) is over the set of all bipartite states $\varphi_{A_1A_2} \in \mathscr{D}(\mathcal{H}_{A_1} \otimes \mathcal{H}_{A_2})$, classical auxiliary variables $(V_0,V_1,V_2) \sim p_{V_0}p_{V_1|V_0}p_{V_2|V_0}$, and collection of POVMs $\mathcal{L}_k(v_0,v_k) = \{L_k(x_k|v_0,v_k)\}_{x_k\in\mathcal{X}_k}$, for $v_0 \in \mathcal{V}_0$, $v_k \in \mathcal{V}_k$, $k=1,2$. Note that the classical variables $V_1 \ominus V_0 \ominus V_2$ form a Markov chain.

Given such a state, variables, and POVMs, the joint distribution of (V_0,V_1,V_2,X_1,X_2,Y) is

$$p_{V_0,V_1,V_2,X_1,X_2,Y}(v_0,v_1,v_2,x_1,x_2,y) = p_{V_0}(v_0)p_{V_1|V_0}(v_1|v_0)p_{V_2|V_0}(v_2|v_0)$$
$$\cdot \mathrm{Tr}\left[(L_1(x_1|v_0,v_1) \otimes L_2(x_2|v_0,v_2))\, \varphi_{A_1A_2}\right] \cdot P_{Y|X_1,X_2}(y|x_1,x_2). \tag{20}$$

The subscript 'ET' stands for entangled transmitters.

Before we state the achievability theorem, we give the following lemma. The property stated below in Lemma 2 can simplify the computation of the region $\mathcal{R}_{\mathrm{ET}}(P_{Y|X_1,X_2})$ for a given channel.

Lemma 2. The union of the inner bound in (19) is exhausted by auxiliary variables V_0, V_1, V_2 with $|\mathcal{V}_0| \leq 3$, $|\mathcal{V}_k| \leq 3(|\mathcal{X}_1||\mathcal{X}_2|+2)$, $k=1,2$, and pure states $\varphi_{A_1A_2} \equiv |\phi_{A_1A_2}\rangle\!\langle\phi_{A_1A_2}|$.

The proof of Lemma 2 is based on the Fenchel-Eggleston-Carathéodory theorem [33] and the perturbation method [34, 35]. The details are given in Section VI-A.

Remark 5. In this section, we consider the capacity region given unlimited entanglement resources between the transmitters. We note that the dimension of the entangled systems A_1 and A_2 in the rate formula (19) is unbounded. As the rate formula $\mathcal{R}_{\mathrm{ET}}(P_{Y|X_1,X_2})$ involves a union over the *closure* of the set of finite-dimensional quantum correlations, we include correlations that are obtained in the limit of $\dim(\mathcal{H}_{A_k}) \to \infty$. This family of correlations is often denoted in the literature by C_{qa} [36]. As we will discuss in Section X, there are cases where the union cannot be exhausted with finite-dimensional entangled systems [11, 36] (see Example 2 below).

Next, we state our achievability result and the regularized characterization.

Theorem 3. The capacity region of a classical MAC $P_{Y|X_1,X_2}$ with entangled transmitters satisfies

$$\mathcal{C}_{\mathrm{ET}}(P_{Y|X_1,X_2}) \supseteq \mathcal{R}_{\mathrm{ET}}(P_{Y|X_1,X_2}). \tag{21}$$

Furthermore, the capacity region is given by

$$\mathcal{C}_{\mathrm{ET}}(P_{Y|X_1,X_2}) = \bigcup_{\ell=1}^{\infty} \frac{1}{\ell} \mathcal{R}_{\mathrm{ET}}(P_{Y|X_1,X_2}^{\ell}) \tag{22}$$

where $P_{Y|X_1,X_2}^{\ell}$ is the ℓ-fold product channel (see notation in Subsection II-A).

The proof of Theorem 3 is given in Section VII. In Subsection IV-C below, we will review two examples by Leditzky et al. [11] that show that the capacity region with entangled transmitters can be strictly larger than without such correlation resources.

Remark 6. In the basic point-to-point communication problem, Bennett et al. [37, 38] showed that entanglement assistance between the transmitter and the *receiver* leads to a characterization that is easy to compute. In other words, in Bennett et al.'s model [38], introducing entanglement resources transforms the capacity evaluation from an uncomputable task to an optimization that can be easily performed, numerically [39, Remark 5]. Unfortunately, this is not the case in the present work. Clearly, our achievable region has a single-letter form with respect to the channel dependency. However, there is no upper bound on the necessary dimension of the auxiliary systems A_1 and A_2 in Theorem 3 (see Remark 5). In principle, one can compute achievable rates by choosing an arbitrary dimension, but the best inner bound cannot be computed with absolute precision in general. In the discussion section, we affirm that this is an inherent limitation to our problem (see Section X). Furthermore, even without entanglement assistance, there exists MACs such that the computation of the capacity region is NP-hard [11].

Remark 7. In communication, shared entanglement is often thought of as the quantum version of common randomness (CR), which is a random key that is shared between the two parties, see Section II. Recall that we denote the capacity region with CR-assisted transmitters by $\mathcal{C}_{\mathrm{CRT}}(P_{Y|X_1,X_2})$. Then, observe that if we choose the measurements in (19) such that $X_k = V_k$ for $k=1,2$, we obtain the capacity region with CR assistance:

$$\mathcal{C}_{\mathrm{ET}}(P_{Y|X_1,X_2}) \supseteq \bigcup_{p_{V_0}p_{X_1|V_0}p_{X_2|V_0}} \left\{ \begin{array}{rl} (R_1,R_2): \ R_1 & \leq I(X_1;Y|X_2V_0) \\ R_2 & \leq I(X_2;Y|X_1V_0) \\ R_1+R_2 & \leq I(X_1X_2;Y|V_0) \end{array} \right\} = \mathcal{C}_{\mathrm{CRT}}(P_{Y|X_1,X_2}). \tag{23}$$

TABLE I
THE MAGIC SQUARE GAME: DETERMINISTIC STRATEGIES

0	0	0
0	1	1
1	0	?

where the last equality holds by Theorem 1 (see [30]). In the examples by Leditzky et al. [11], there is strict inclusion, i.e., $\mathcal{C}_{\text{CRT}}(P_{Y|X_1,X_2}) \subsetneq \not\supseteq \mathcal{C}_{\text{ET}}(P_{Y|X_1,X_2})$. We will show the examples in Subsection IV-C.

2) Outer Bound: Next, we establish an outer bound. Define the rate region

$$\mathcal{O}_{\text{ET}}(P_{Y|X_1,X_2}) = \bigcup_{p_{V_0V_1V_2}, \varphi_{A_1A_2}, \mathcal{L}_1\otimes\mathcal{L}_2} \left\{ \begin{array}{rl} (R_1,R_2) : R_1 & \leq I(V_1;Y|V_0V_2) \\ R_2 & \leq I(V_2;Y|V_0V_1) \\ R_1 + R_2 & \leq I(V_1V_2;Y|V_0) \end{array} \right\} \tag{24}$$

The difference between the formula given above and the inner bound in (19) is in the joint distribution of the classical auxiliary variables. Specifically, in (19), we limit the joint distribution such that $V_1 \multimap V_0 \multimap V_2$ form a Markov chain. Here, we remove this restriction and consider any joint distribution. The union of the outer bound in (24) is exhausted by auxiliary variables and pure states as in Lemma 2 following the same considerations. The derivation of those properties is given in Section VI-B. The lack of a dimension bound for the bipartite state extends to the outer bound as well (see Remark 5).

Theorem 4. The capacity region of a classical MAC $P_{Y|X_1,X_2}$ with entangled transmitters satisfies

$$\mathcal{C}_{\text{ET}}(P_{Y|X_1,X_2}) \subseteq \mathcal{O}_{\text{ET}}(P_{Y|X_1,X_2}). \tag{25}$$

B. Rate-Limited Entanglement Resources

Now, we consider a setting where the entanglement resources are limited. A code with rate-limited entanglement resources is defined such that the encoders have access to Hilbert spaces which grow as a function of the channel uses in a rate limited fashion. The precise definition is given below.

Definition 5. A $(2^{nR_1}, 2^{nR_2}, n)$ code for the classical MAC $P_{Y|X_1,X_2}$ with an entanglement rate θ_E between the transmitters consists of the following:

- a bipartite state $\Psi_{E_1E_2}$ that is shared between the transmitters, with $\dim(\mathcal{H}_{E_k}) \leq 2^{n\theta_E}$ for $k = 1, 2$.
- two message sets $[1 : 2^{nR_1}]$ and $[1 : 2^{nR_2}]$, assuming 2^{nR_k} is an integer;
- a pair of encoding POVMs $\mathcal{F}_1^{(m_1)} = \{F_{x_1^n}^{(m_1)}\}_{x_1^n \in \mathcal{X}_1^n}$ and $\mathcal{F}_2^{(m_2)} = \{F_{x_2^n}^{(m_2)}\}_{x_2^n \in \mathcal{X}_2^n}$ on E_1 and E_2, respectively; and
- a decoding function $g : \mathcal{Y}^n \to [1 : 2^{nR_1}] \times [1 : 2^{nR_2}]$.

The communication scheme is as in Subsection III-B (see Figure 1). A rate pair (R_1, R_2) is said to be achievable with an entanglement rate θ_E if for every $\varepsilon > 0$ and sufficiently large n, there exists a $(2^{nR_1}, 2^{nR_2}, n)$ code such that the transmitters share entanglement resources at a rate θ_E and the maximal probability of error is bounded by $P_e^{(n)}(\mathscr{C}) \leq \varepsilon$. The capacity region $\mathcal{C}_{\text{ET}}(P_{Y|X_1,X_2}, \theta_E)$ with rate-limited entanglement between the transmitters is defined as the set of achievable pairs (R_1, R_2), with an entanglement rate θ_E between the transmitters.

We note that by definition, the capacity region with unlimited entanglement between the transmitters is $\mathcal{C}_{\text{ET}}(\cdot) \equiv \mathcal{C}_{\text{ET}}(\cdot, +\infty)$. Therefore, for every entanglement rate $\theta_E > 0$, we have $\mathcal{C}_{\text{ET}}(\cdot, \theta_E) \subseteq \mathcal{C}_{\text{ET}}(\cdot)$.

Based on the achievability proof in Section VII and Lemma 2, we obtain the following consequence.

Corollary 5. Let $\theta_E > 0$ be a given entanglement rate. A rate pair (R_1, R_2) is achievable with entanglement at rate θ_E between the transmitters if

$$\begin{array}{rl} R_1 & \leq I(V_1;Y|V_0V_2) \\ R_2 & \leq I(V_2;Y|V_0V_1) \\ R_1 + R_2 & \leq I(V_1V_2;Y|V_0) \end{array} \tag{26}$$

for a pure state $|\phi_{A_1A_2}\rangle$ with entropy

$$H(A_1)_\phi = H(A_2)_\phi \leq \theta_E, \tag{27}$$

some classical variables $(V_0, V_1, V_2) \sim p_{V_0}p_{V_1|V_0}p_{V_2|V_0}$, and conditional measurement $\mathcal{L}_k(v_0, v_k) = \{L_k(x_k|v_0, v_k)\}_{x_k \in \mathcal{X}_k}$, for $v_0 \in \mathcal{V}_0$, $v_k \in \mathcal{V}_k$, $k = 1, 2$.

TABLE II
THE MAGIC SQUARE GAME: QUANTUM STRATEGIES

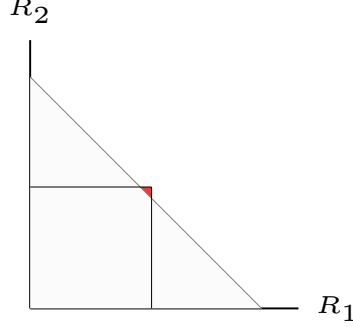| $X \otimes \mathbb{1}$	$X \otimes X$	$\mathbb{1} \otimes X$
$-X \otimes Z$	$Y \otimes Y$	$-Z \otimes X$
$\mathbb{1} \otimes Z$	$Z \otimes Z$	$Z \otimes \mathbb{1}$

Fig. 4. Rate bounds for the magic-square multiple-access channel $P_{Y|X_1,X_2}$ in Example 1. The gray triangle indicates the outer bound for the capacity region $\mathcal{C}_{\mathrm{CRT}}(P_{Y|X_1,X_2})$, with classical CR between the transmitters. The square indicate is the capacity region $\mathcal{C}_{\mathrm{ET}}(P_{Y|X_1,X_2})$, with entangled transmitters. The red area represents rate pairs that can be achieved with entanglement, but cannot be achieved with classical CR.

C. Pseudo-Telepathy Examples

As an example, we consider two channels that were introduced by Leditzky et al. [11].

Example 1 (Magic-square channel). The channel is defined in terms of a pseudo-telepathy game, i.e., a non-local game such that quantum strategies outperform classical strategies and guarantee winning with certainty. In the magic square game, a referee selects one out of nine cells uniformly at random. Suppose that the referee selected (r, c). Then, the referee informs Player 1 of the row index r, and Player 2 of the column index c. Each player fills three bits in the respective row r and column c. In order to win the game, they need to satisfy three requirements: they agree on the bit value in (r, c), the row r has even parity, and the column c has odd parity. Tables I and II demonstrate a classical strategy and a quantum strategy, respectively. If the game is limited to a classical deterministic strategy, then winning is impossible. An attempt towards winning the game is shown in Table I. Furthermore, using randomized strategies, the probability of winning is at most $\frac{8}{9}$. On the other hand, it can be shown that the following quantum strategy wins the game with probability 1:

- Before the game begins, prepare the state

$$\left| \phi_{A_1' A_1'' A_2' A_2''} \right\rangle = \frac{1}{2} \left(|00\rangle |11\rangle + |11\rangle |00\rangle - |01\rangle |10\rangle - |10\rangle |01\rangle \right) . \tag{28}$$

  Send A_1', A_1'' to Player 1, and A_2', A_2'' to Player 2.
- Play the game using the following strategy. Having received the row and column indices, each player measures the observables in Table II simultaneously, and inserts the measurement outcomes into the corresponding cells, where X, Y, and Z are the Pauli operators. The observables can be measured simultaneously because the three operators in each row and each column commute. For example, if the referee selected $r = 1$ and $c = 2$, then Player 1 measures $X \otimes \mathbb{1}$, $X \otimes X$, and $\mathbb{1} \otimes X$, whereas Player 2 measures $X \otimes X$, $Y \otimes Y$, and $Z \otimes Z$.

Leditzky et al. [11] defined a MAC such that the channel is ideal for input strategies that win the game, and pure-noise otherwise. The precise definition is given below.

In the general description of the non-local game, a referee selects two questions $q_1 \in \mathcal{Q}_1$ and $q_2 \in \mathcal{Q}_2$ uniformly at random, and sends the respective question to each player. The players choose their respective answers $a_1 \in \mathcal{A}_1$ and $a_2 \in \mathcal{A}_2$ using either deterministic or random strategies $f_k : \mathcal{Q}_k \to \mathcal{A}_k$, for $k = 1, 2$. The game is won if $(q_1, q_2, a_1, a_2) \in \mathscr{G}$, where $\mathscr{G}$ is the winning set. The CHSH game and magic square game are special cases of this description. Leditzky et al. [11] defined a classical MAC $P_{Y|X_1,X_2}$, with

$$\mathcal{X}_k = \mathcal{Q}_k \times \mathcal{A}_k \,, \ k = 1, 2 \tag{29}$$
$$\mathcal{Y} = \mathcal{Q}_1 \times \mathcal{Q}_2 \tag{30}$$

such that

$$P_{Y|X_1,X_2}\big((\hat{q}_1, \hat{q}_2) \big| q_1, a_1, q_2, a_2\big) = \begin{cases} \delta_{q_1,\hat{q}_1} \delta_{q_2,\hat{q}_2} & \text{if } (q_1, q_2, a_1, a_2) \in \mathscr{G} \,, \\ \frac{1}{|\mathcal{Q}_1||\mathcal{Q}_2|} & \text{otherwise.} \end{cases} \tag{31}$$

In words, if the inputs $X_1 = (q_1, a_1)$ and $X_2 = (q_2, a_2)$ win the game, then the decoder receives the question pair precisely, i.e., $Y = (q_1, q_2)$ with probability 1. Otherwise, if the game is lost, then the output Y is uniformly distributed over the question set.

Formally, the magic square game is specified by questions (q_1, q_2) from $\{1, 2, 3\} \times \{1, 2, 3\}$, answers a_1, a_2 from $\mathcal{A}_1 = \mathcal{A}_2 = \{0, 1\}^3$, and the winning set $\mathcal{G} = \{ (q_1, q_2, (a_1[j], a_2[j])_{j=1,2,3}) \in \mathcal{Q}_1 \times \mathcal{Q}_2 \times \mathcal{A}_1 \times \mathcal{A}_2 : a_1[q_1] = a_2[q_2], a_1[1] + a_1[2] + a_1[3] \mod 2 = 0, \text{ and } a_2[1] + a_2[2] + a_2[3] \mod 2 = 1 \}$. Seshadri et al. [15] showed that without assisting resources and with a message-average error criterion, the sum-rate is bounded by $R_1 + R_2 \le 3.02$. Therefore, this holds for classical CR between the transmitters as well by Theorem 1, for either maximal or message-average error criterion (see Remark 2). That is, the capacity region with classical CR between the transmitters satisfies

$$\mathcal{C}_{\mathrm{CRT}}(P_{Y|X_1, X_2}) \subseteq \{(R_1, R_2) : R_1 + R_2 \le 3.02\}. \tag{32}$$

The outer bound is depicted in Figure 4 as a gray triangle.

Based on [11], the sum rate $R_1 + R_2 = 2\log(3) \approx 3.17$ is achievable with entangled transmitters. This can also be obtained as a consequence of our result. By Theorem 3, the capacity region with entangled transmitters is given by

$$\mathcal{C}_{\mathrm{ET}}(P_{Y|X_1, X_2}) = \left\{ (R_1, R_2) : \begin{array}{l} R_1 \le \log(3) \\ R_2 \le \log(3) \end{array} \right\}. \tag{33}$$

The capacity region of the MAC with entangled transmitters is the square region in Figure 4. The figure demonstrates the violation of the classical upper bound using entanglement, as the red area represents rate pairs that can be achieved with entanglement, but cannot be achieved with classical CR.

The converse part is immediate since the set on the right-hand side of (33) is the capacity region of the noiseless MAC $\widetilde{P}_{Y|X_1, X_2}((\hat{q}_1, \hat{q}_2) | q_1, a_1, q_2, a_2) = \delta_{q_1, \hat{q}_1} \delta_{q_2, \hat{q}_2}$ (with or without entanglement resources). As for the direct part, consider the region in (19). We choose a bipartite state, classical variables, and POVMs as follows. Let the bipartite state $|\phi_{A_1 A_2}\rangle$ be as in (28), where $A_k \equiv A'_k A''_k$ for $k = 1, 2$. Furthremore, we let $\mathcal{V}_0 = \emptyset$ and set the distribution p_{V_k} to be uniform over $\mathcal{V}_k = \{1, 2, 3\}$, for $k = 1, 2$. Thus, the random pair (V_1, V_2) is distributed as the referee's questions. Given V_k, Alice k performs a measurement $\mathcal{L}_k(V_k)$ as follows. Alice 1 measures the observables in row number $r = V_1$ in Table II, obtains a random triplet $W_1 \equiv (a_1[j])_{j=1,2,3}$, and transmits $X_1 = (V_1, W_1)$. Similarly, Alice 2 measures the observables in column number $c = V_2$, obtains $W_2 \equiv (a_2[j])_{j=1,2,3}$, and transmits $X_2 = (V_2, W_2)$. Since (V_1, V_2, W_1, W_2) win the game, we have $Y = (V_1, V_2)$ with probability 1, hence $I(V_1 V_2; Y) = H(V_1 V_2) = 2\log(3)$, $I(V_1; Y|V_2) = H(V_1) = \log(3)$, and $I(V_2; Y|V_1) = H(V_2) = \log(3)$. This requires an entanglement rate of $\theta_E = 2$ qubit pairs per transmission.

Example 2 (Slofstra-Vidick channel). Another example by Leditzky et al. [11] is defined in terms of the following non-local game. Consider a linear system of equations, $\mathbf{H}\mathbf{x} = \mathbf{b}$, over $\mathrm{GF}(2)$, with fixed $\mathbf{H}$ and $\mathbf{b}$, where $\mathbf{H}$ is a binary matrix of size $K \times N$, and $\mathbf{b}$ is a binary vector of length K. That is, the system consists of K equations of the form

$$\mathbf{h}_k \cdot \mathbf{x} = b_k \tag{34}$$

for $k \in \{1, \ldots, K\}$, where $\mathbf{h}_k$ is the kth row in $\mathbf{H}$.

In the linear-system game [40], a referee selects two indices, an equation index $k \in \{1, \ldots, K\}$ and a variable index $j \in \{1, \ldots, N\}$, uniformly at random. Then, the referee informs Player 1 of the index k, and Player 2 of the index j. Player 1 responds with a vector $\mathbf{a}_1 \in \{0, 1\}^N$, which is interpreted as an assignment for $\mathbf{x}$, and Player 2 responds with a single bit $a_2 \in \{0, 1\}$ that is interpreted as the value of $\mathbf{x}[j]$. In order to win the game, they need to satisfy two requirements:
1) The response of Player 1 satisfies the kth equation. That is, (34) holds for $\mathbf{x} = \mathbf{a}_1$.
2) Their responses satisfy $\mathbf{H}[k, j] \cdot (\mathbf{a}_1[j] - a_2) = 0$. That is, either the jth variable is not included in the kth equation (i.e., $\mathbf{H}[k, j] = 0$), or the players must agree on the value of $\mathbf{x}[j]$ (namely, $\mathbf{a}_1[j] = a_2$).

Suppose that the players share entangled systems of dimension d_E. Slofstra and Vidick [40] showed that quantum strategies outperform classical strategies in this game. However, the minimal entanglement dimension $d_{E,\min}$ that is required in order to win the game with a probability of $1 - e^{-T}$ is bounded by

$$Ce^{T/6} \le d_{E,\min} \le C'e^{T/2} \tag{35}$$

for all $T > 0$, where C, C' are positive constants. See further details in [40, Th. 1.1]. It follows that the game can be won with certainty for $d_E \to \infty$, and cannot be won with certainty if the entanglement dimension is bounded.

Leditzky et al. [11] considered a classical MAC $P_{Y|X_1, X_2}$ defined as in (31), where $\mathcal{G}$ is the winning set for the Slofstra-Vidick game [40], as described above, where $\mathcal{Q}_1 = \{1, \ldots, K\}$, $\mathcal{Q}_2 = \{1, \ldots, N\}$, $\mathcal{A}_1 = \{0, 1\}^N$, and $\mathcal{A}_2 = \{0, 1\}$. As before, if the inputs $X_1 = (q_1, \mathbf{a}_1)$ and $X_2 = (q_2, a_2)$ win the game, then the decoder receives the question pair precisely, i.e., $Y = (q_1, q_2)$ with probability 1. Otherwise, if the game is lost, then the output Y is uniformly distributed over the question set. The capacity region of the classical MAC with unlimited entanglement resources between the transmitters is given by [11]

$$\mathcal{C}_{\mathrm{ET}}(P_{Y|X_1, X_2}) = \left\{ (R_1, R_2) : \begin{array}{l} R_1 \le \log(K) \\ R_2 \le \log(N) \end{array} \right\}. \tag{36}$$

This result can also be obtained from Theorem 3 following similar arguments as in Example 1.

Furthermore, we obtain an achievable rate region with entanglement resources at a limited rate. By Corollary 5, we have that a rate pair (R_1, R_2) is achievable given entanglement resources at a limited rate of $\theta_E = \frac{1}{2}T + \log(C')$, if

$$R_1 \leq \left(1 - 2e^{-T}\right)\log(K) - 2h_2(e^{-T}), \tag{37}$$

$$R_2 \leq \left(1 - 2e^{-T}\right)\log(N) - 2h_2(e^{-T}) \tag{38}$$

for all $T > 0$, where $h_2(p) = -p\log(p) - (1-p)\log(1-p)$ is the binary entropy function over $(0,1)$. To show this, we use the Slofstra-Vidick bound in (35), along with entropy continuity bounds (see [41] [32, Sec. 11.10]).

V. Main Results — Entangled and Conferencing Transmitters

A. Classical Conferencing

In this section, we consider the combination of entanglement resources and classical conferencing between the transmitters (see Figure 2). In addition to pre-shared entanglement resources, the transmitters can now communicate with each other over noise-free links at fixed rates, with a classical capacity of C_{12} from Transmitter 1 to Transmitter 2, and C_{21} from Transmitter 2 to Transmitter 1.

Define the following rate region,

$$\mathcal{R}_{\text{ET-C}}(P_{Y|X_1,X_2}, C_{12}, C_{21}) = \bigcup_{p_{V_0}p_{V_1|V_0}p_{V_2|V_0},\,\varphi_{A_1A_2},\,\mathcal{L}_1\otimes\mathcal{L}_2} \left\{ \begin{array}{rl} (R_1, R_2) : R_1 & \leq I(V_1; Y|V_0 V_2) + C_{12} \\ R_2 & \leq I(V_2; Y|V_0 V_1) + C_{21} \\ R_1 + R_2 & \leq I(V_1 V_2; Y|V_0) + C_{12} + C_{21} \\ R_1 + R_2 & \leq I(V_1 V_2; Y) \end{array} \right\}. \tag{39}$$

The union on the right-hand side of (39) is the same as in (19). Similarly, we define the region $\mathcal{R}_{\text{ET-C}}(P_{Y|X_1,X_2}, C_{12}, C_{21})$ with a union over all joint distributions p_{V_0,V_1,V_2}.

Theorem 6. The capacity region of a classical MAC $P_{Y|X_1,X_2}$ with entangled transmitters and classical conferencing satisfies

$$\mathcal{C}_{\text{ET-C}}(P_{Y|X_1,X_2}, C_{12}, C_{21}) \supseteq \mathcal{R}_{\text{ET-C}}(P_{Y|X_1,X_2}, C_{12}, C_{21}), \tag{40}$$

and

$$\mathcal{C}_{\text{ET-C}}(P_{Y|X_1,X_2}, C_{12}, C_{21}) \subseteq \mathcal{O}_{\text{ET-C}}(P_{Y|X_1,X_2}, C_{12}, C_{21}). \tag{41}$$

Furthermore, the capacity region is given by

$$\mathcal{C}_{\text{ET-C}}(P_{Y|X_1,X_2}, C_{12}, C_{21}) = \bigcup_{\ell=1}^{\infty} \frac{1}{\ell} \mathcal{R}_{\text{ET-C}}(P_{Y|X_1,X_2}^{\ell}, C_{12}, C_{21}), \tag{42}$$

The proof of Theorem 6 is given in Section IX.

Remark 8. If the conferencing rates are sufficiently high, then we obtain the full-cooperation region,

$$\mathcal{C}_{\text{ET-C}}(P_{Y|X_1,X_2}, C_{12}, C_{21}) = \{(R_1, R_2) : R_1 + R_2 \leq C_{\text{joint}}\} \tag{43}$$

where $C_{\text{joint}} \equiv \max_{p_{X_1 X_2}} I(X_1 X_2; Y)$. In particular, this holds if both conferencing rates are higher than C_{joint}. In this case, we effectively have a single-user channel from (X_1, X_2) to Y. On the other hand, if the conferencing rates are zero, i.e., $C_{12} = C_{21} = 0$, then we recover the achievable region without conferencing, as in Theorem 3. That is, $\mathcal{C}_{\text{ET-C}}(P_{Y|X_1,X_2}, 0, 0) = \mathcal{C}_{\text{ET}}(P_{Y|X_1,X_2})$.

Remark 9. We have compared the effect of entanglement and CR on communication performance in Remark 7. Both entanglement and CR are static resources of *non-signaling correlation*, which cannot be used in order to send information. Conferencing, however, is a *dynamic* resource of cooperation, which can be used by itself to send information from one transmitter to the other. In this sense, conferencing is stronger than entanglement. Thereby, one may ask, for example, is conferencing at a low rate better than entanglement at a high rate? While this question may seem unfair, it can be justified by the delay that is caused by using conferencing. Based on Theorem 6, the answer is no. It can be easily seen that when the values of C_{12} and C_{21} are very low, the rate increase is negligible as well. Whereas, we have seen that sufficient entanglement without conferencing can lead to a significant increase.

Remark 10. Conferencing is achieved through interaction between the two transmitters. As already mentioned, we only consider trustworthy senders here. For future communication systems, especially for 6G, it is also very important to examine communication protocols for non-trustworthy communication parties. The complexity theory shows that performance behavior, which is also very interesting for such communication parties, can be increased. In the last 40 years, enormous progress has been made in complexity theory, which should also be of interest for communication systems. In particular, the work for prover and verifier interaction is very relevant. Techniques from the coding theory were also often used here in finding of corresponding

protocols. For example, it was shown in [42] that through interaction between a prover and a verifier with polynomial time, and the verifier using randomness exactly then all languages from PSPACE (polynomial space Turing machines) are decidable. In an interaction between two provers and a verifier with polynomial time, all languages from NEXP, i.e. Turing machines with non-deterministic exponential computing time, can be decided using randomness [43]. If the performance of the verifier is described as a poly-logarithmic use of random bits with a constant number of information bits greater than or equal to 3, then in an optimal interaction with a prover the verifier can already decide all languages from NP, i.e. non- deterministic polynomial time [44, 45]. Of course, much more complex communication scenarios occur in real communication systems than those considered in [42–45], but this research direction is nevertheless particularly interesting with regard to trustworthiness for 6G. It is fascinating to see that one of the central problems in mathematics, (Millennium prize [46]), can be represented in an equivalent form by a simple interactive communication task (see [46–48] for an overview and status of the problems). For the case with entanglement assistance, this is open. Another very interesting property of interactive communication in complexity theory is: In the communication scenario with one prover and two verifiers, there is a zero-knowledge interactive proof if and only if a one-way function exists [49, 50]. It should be noted here that the existence of a one-way function as a central cryptographic primitive cannot be proved so far. In complexity theory, also for certain interactive communication scenarios, the type of interactivity, that is, how many rounds of interaction are necessary to solve certain tasks, is interesting. In [42, 43] there is no bound on the number of rounds of interaction necessary. In [51] it has now been shown for two prover and two verifier interactive systems that any language from PSPACE can be decided with a one round protocol. For the MAC with conferencing encoders and also the AV-MAC with conferencing encoders, one round of conferencing is sufficient for each. On the other hand, the existence of interactive zero-knowledge proofs for the two prover and one verifier scenario can be directly proved [52]. That is, no cryptographic assumptions are needed. This is another interesting example of the special potential of interactive and cooperative communication scenarios.

B. Quantum Conferencing

In this section, we consider the MAC with entanglement and quantum conferencing between the transmitters (see Figure 3). The transmitters can now communicate with each other over noise-free quantum links at fixed rates, with a quantum capacity of Q_{12} from Transmitter 1 to Transmitter 2, and Q_{21} from Transmitter 2 to Transmitter 1. Furthermore, the transmitters share entanglement resources that can be used at both the conferencing and the encoding stages. We establish an inner bound on the capacity region. Define

$$
\mathcal{R}_{\text{ET-Q}}(P_{Y|X_1,X_2}, Q_{12}, Q_{21}) = \bigcup_{p_{V_0} p_{V_1|V_0} p_{V_2|V_0}, \varphi_{A_1 A_2}, \mathcal{L}_1 \otimes \mathcal{L}_2} \left\{ \begin{array}{rl} (R_1, R_2) : R_1 & \leq I(V_1; Y|V_0 V_2) + 2Q_{12} \\ R_2 & \leq I(V_2; Y|V_0 V_1) + 2Q_{21} \\ R_1 + R_2 & \leq I(V_1 V_2; Y|V_0) + 2(Q_{12} + Q_{21}) \\ R_1 + R_2 & \leq I(V_1 V_2; Y) \end{array} \right\}.
$$
(44)

The union on the right-hand side of (39) is the same as in (19).

Theorem 7. The capacity region of a classical MAC $P_{Y|X_1,X_2}$ with entangled transmitters and quantum conferencing satisfies

$$
\mathcal{C}_{\text{ET-Q}}(P_{Y|X_1,X_2}, C_{12}, C_{21}) \supseteq \mathcal{R}_{\text{ET-Q}}(P_{Y|X_1,X_2}, C_{12}, C_{21}).
$$
(45)

The proof follows similar lines as for Theorem 6 in Appendix IX. Here, however, each encoder uses superdense coding in order to increase the rate of her conferencing message. Alice 1 sends nR_1' bits to Alice 2 through a quantum transmission of $\frac{1}{2}nR_1'$ qubits via her quantum conferencing link, using $\frac{1}{2}nR_1'$ entangled pairs. Similarly, Alice 2 sends nR_2' bits to Alice 2 by sending $\frac{1}{2}nR_2'$ qubits along with $\frac{1}{2}nR_2'$ entangled pairs. This is permissible if the common information rates are each bounded by *twice* the respective conferencing capacity, i.e., $R_1' \leq 2Q_{12}$ and $R_2' \leq 2Q_{21}$. The rest of the proof is identical to that of Theorem 6. Hence, The details are omitted.

C. Example

To demonstrate our results on the classical MAC with entangled transmitters and conferencing, we revisit the magic-square channel from Subsection IV-C.

Example 1 (Continued). Consider the magic-square MAC from Example 1, where the channel output is a noiseless copy of the referee's chosen cell if the inputs win the game, and it is pure noise otherwise.

Figure 5 shows achievable rate regions in the three settings that we have considered. We have seen that since the game can be won with probability 1 given entanglement resources, the capacity region with entangled transmitters, without conferencing, is given by [11]

$$
\mathcal{C}_{\text{ET}}(P_{Y|X_1,X_2}) = \left\{ \begin{array}{rl} (R_1, R_2) : & R_1 \leq \log(3) \\ & R_2 \leq \log(3) \end{array} \right\}.
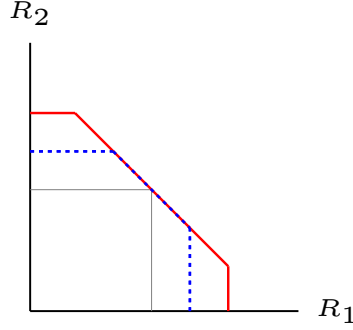$$
(46)

Fig. 5. Achievable rate regions for the magic-square multiple-access channel $P_{Y|X_1,X_2}$ with entangled transmitters and, either classical or quantum, conferencing, as we in Example 1. The inner square indicates the capacity region $\mathcal{C}_{\mathrm{ET}}(P_{Y|X_1,X_2})$ with entangled transmitters. The pentagon below the blue dashed line is an achievable region for the MAC with entangled transmitters and classical conferencing, for $C_{12} = C_{21} = 0.5$. The pentagon below the red solid line is achievable when the conferencing links are quantum as well, for $Q_{12} = Q_{21} = 0.5$.

The capacity region of the MAC with entangled transmitters and without conferencing is the square region in Figure 5.

Now, consider conferencing of classical messages between the transmitters, at rates $C_{12} = C_{21} = 0.5$. In this case, we obtain the following achievable region

$$\mathcal{R}_{\mathrm{ET\text{-}C}}(P_{Y|X_1,X_2}) \supseteq \left\{ \begin{array}{rl} (R_1, R_2) : & R_1 \leq \log(3) + 0.5 \\ & R_2 \leq \log(3) + 0.5 \\ R_1 + R_2 & \leq 2\log(3) \end{array} \right\}. \tag{47}$$

If the conferencing links are quantum, however, at the same rates $Q_{12} = Q_{21} = 0.5$, then we have an even larger achievable region,

$$\mathcal{R}_{\mathrm{ET\text{-}Q}}(P_{Y|X_1,X_2}) \supseteq \left\{ \begin{array}{rl} (R_1, R_2) : & R_1 \leq \log(3) + 1 \\ & R_2 \leq \log(3) + 1 \\ R_1 + R_2 & \leq 2\log(3) \end{array} \right\}. \tag{48}$$

The bounds are depicted in Figure 5. The pentagons, below the blue dashed line and below the red solid line, indicate the achievable regions on the right hand side of (47) and (48), respectively. We show achievability based on Theorem 6 and Theorem 7, respectively, using the same choice of a bipartite state, classical variables, and POVMs, as in the first part of Example 1. Specifically, let the bipartite state $|\phi_{A_1 A_2}\rangle$ be as in (28), where $A_k \equiv A'_k A''_k$ for $k = 1, 2$. Furthermore, we let $\mathcal{V}_0 = \emptyset$ and set the distribution p_{V_k} to be uniform over $\mathcal{V}_k = \{1, 2, 3\}$, for $k = 1, 2$, as the referee's questions. Given V_1, Alice 1 measures the observables in row number $r = V_1$ in Table II, obtains a random triplet $W_1 \equiv (a_1[j])_{j=1,2,3}$, and transmits $X_1 = (V_1, W_1)$. Similarly, Alice 2 measures the observables in column number $c = V_2$, obtains $W_2 \equiv (a_2[j])_{j=1,2,3}$, and transmits $X_2 = (V_2, W_2)$. Since (V_1, V_2, W_1, W_2) win the game, we have $Y = (V_1, V_2)$ with probability 1, hence $I(V_1 V_2; Y) = H(V_1 V_2) = 2\log(3)$, $I(V_1; Y|V_2) = H(V_1) = \log(3)$, and $I(V_2; Y|V_1) = H(V_2) = \log(3)$.

VI. PURIFICATION AND CARDINALITY

In this section, we show that the inner and outer regions can be exhausted with pure states and bounded cardinality.

A. Inner Bound

We prove Lemma 2 and show that the region $\mathcal{R}_{\mathrm{ET}}(P_{Y|X_1,X_2})$ can be exhausted with pure states and bounded cardinality.

1) Purification: First, we consider a given Hilbert space $\mathcal{H}_{A_1} \otimes \mathcal{H}_{A_2}$, and show that the union over bipartite states $\varphi_{A_1 A_2}$ is exhausted by pure states. Fix a Hilbert space $\mathcal{H}_{A_1} \otimes \mathcal{H}_{A_2}$. Consider a given distribution $\{p_{V_1|V_0}(\cdot|v_0) p_{V_2|V_0}(\cdot|v_0)\}$, a bipartite state $\varphi_{A_1 A_2}$, and measurements $\mathcal{L}_1(v_0, v_1)$ and $\mathcal{L}_2(v_0, v_2)$ on A_1 and A_2, respectively. Let $\mathfrak{R}(\varphi, \mathcal{L}_1, \mathcal{L}_2)$ denote the associated rate region,

$$\mathfrak{R}(\varphi, \mathcal{L}_1, \mathcal{L}_2) = \left\{ \begin{array}{rl} (R_1, R_2) : R_1 & \leq I(V_1; Y|V_0 V_2) \\ R_2 & \leq I(V_2; Y|V_0 V_1) \\ R_1 + R_2 & \leq I(V_1 V_2; Y|V_0) \end{array} \right\}. \tag{49}$$

Given a joint state $\varphi_{A_1 A_2}$, consider a spectral decomposition,

$$\varphi_{A_1 A_2} = \sum_{z \in \mathcal{Z}} p_Z(z) |\phi_z\rangle\langle\phi_z|_{A_1 A_2}, \tag{50}$$

where p_Z is a probability distribution, such that the pure states, $|\phi_z\rangle_{A_1 A_2}$, $z \in \mathcal{Z}$, form an orthonormal basis for $\mathcal{H}_{A_1} \otimes \mathcal{H}_{A_2}$. Then, $\varphi_{A_1 A_2}$ has the following purification,

$$|\psi_{A_1 A_2 E_1 E_2}\rangle = \sum_{z \in \mathcal{Z}} \sqrt{p_Z(z)} |\phi_z\rangle_{A_1 A_2} \otimes |z\rangle_{E_1} \otimes |z\rangle_{E_2} . \tag{51}$$

Now, we consider rate region $\mathfrak{R}(\psi, \mathcal{L}_1', \mathcal{L}_2')$ that is associated with the following choice of state, distribution, and measurements. Set the distribution $\{p_{V_1|V_0}(\cdot|v_0) p_{V_2|V_0}(\cdot|v_0)\}$ as before. Let Alice 1 and Alice 2 share the state $|\psi_{A_1 E_1 A_2 E_2}\rangle$, and suppose that Alice 1 performs a measurement on (A_1, E_1), and Alice 2 on (A_2, E_2), with the following POVMs,

$$L_1'(x_1, z|v_0, v_1) = L(x_1|v_0, v_1) \otimes |z\rangle\langle z|_{E_1} , \tag{52}$$
$$L_2'(x_2, z|v_0, v_2) = L(x_2|v_0, v_2) \otimes |z\rangle\langle z|_{E_2} , \tag{53}$$

for $(v_0, v_1, v_2, x_1, x_2, z) \in \mathcal{V}_0 \times \mathcal{V}_1 \times \mathcal{V}_2 \times \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Z}$. The corresponding input distribution $p_{X_1, X_2|V_0, V_1, V_2}'$ is given by

$$\begin{aligned}
p_{X_1, X_2|V_0, V_1, V_2}'(x_1, x_2|v_0, v_1, v_2) &= \sum_{z \in \mathcal{Z}} p_Z(z) \mathrm{Tr} \left[(L_1(x_1|v_0, v_1) \otimes L_2(x_2|v_0, v_2)) |\phi_z\rangle\langle\phi_z|_{A_1 A_2} \right] \\
&= \mathrm{Tr} \left[(L_1(x_1|v_0, v_1) \otimes L_2(x_2|v_0, v_2)) \left(\sum_{z \in \mathcal{Z}} p_Z(z) |\phi_z\rangle\langle\phi_z|_{A_1 A_2} \right) \right] \\
&= \mathrm{Tr} \left[(L_1(x_1|v_0, v_1) \otimes L_2(x_2|v_0, v_2)) \varphi_{A_1 A_2} \right] \\
&= p_{X_1, X_2|V_0, V_1, V_2}(x_1, x_2|v_0, v_1, v_2)
\end{aligned} \tag{54}$$

where the third equality follows from (50), and the last from (20). Therefore, $\mathfrak{R}(\psi, \mathcal{L}_1', \mathcal{L}_2') = \mathfrak{R}(\varphi, \mathcal{L}_1, \mathcal{L}_2)$. We deduce that the entire region $\mathcal{R}_{\mathrm{ET}}(P_{Y|X_1, X_2})$, as in (19), can be obtained from pure states.

2) Cardinality Bounds: Now, we bound the cardinality of the alphabet $\mathcal{V}_0$. Consider a given distribution $\{p_{V_1|V_0}(\cdot|v_0) p_{V_2|V_0}(\cdot|v_0)\}$, a joint state $\varphi_{A_1 A_2}$, and POVM collections, $\mathcal{L}_1(v_0, v_1)$ and $\mathcal{L}_2(v_0, v_2)$. Define a map $\tau_0 : \mathcal{V}_0 \to \mathbb{R}^3$ by

$$\tau_0(v_0) = \left(I(V_1; Y|V_2, V_0 = v_0), \, I(V_2; Y|V_1, V_0 = v_0), \, I(V_1, V_2; Y|V_0 = v_0) \right). \tag{55}$$

The map τ_0 can be extended to a map that acts on probability distributions as follows,

$$T_0 \, : \, p_{V_0}(\cdot) \mapsto \sum_{v_0 \in \mathcal{V}_0} p_{V_0}(v_0) \tau_0(v_0) = \left(I(V_1; Y|V_0 V_2), \, I(V_2; Y|V_0 V_1), \, I(V_1 V_2; Y|V_0) \right). \tag{56}$$

According to the Fenchel-Eggleston-Carathéodory theorem [33], any point in the convex closure of a connected compact set within $\mathbb{R}^d$ belongs to the convex hull of d points in the set. Since the map T_0 is linear, it maps the set of distributions on $\mathcal{V}_0$ to a connected compact set in $\mathbb{R}^3$. Thus, for every p_{V_0}, there exists a probability distribution $p_{\bar{V}_0}$ on a subset $\bar{\mathcal{V}}_0 \subseteq \mathcal{V}_0$ of size 3, such that $T_0(p_{\bar{V}_0}) = T_0(p_{V_0})$. We deduce that alphabet size can be restricted to $|\mathcal{V}_0| \leq 3$, while preserving $I(V_1; Y|V_0 V_2)$, $I(V_2; Y|V_0 V_1)$, and $I(V_1 V_2; Y|V_0)$.

Next, we bound the alphabet size for the auxiliary variables V_1. A standard application of the Fenchel-Eggleston-Carathéodory theorem does not work here, as it results in a cardinality bound of $|\mathcal{V}_2| \cdot 3(|\mathcal{X}_1||\mathcal{X}_2| + 1$ on V_1. Thereby, we use the perturbation method [34, 35].

Fix a joint state $\varphi_{A_1 A_2}$, POVM collections, $V_0 = v_0$, and $p_{V_2|V_0}(\cdot|v_0)$. Since our rate region is a convex set, it can be characterized by determining the supporting Hyperplanes. Furthermore, due to the pentagonal form of each region in (49), we may focus on the corner points and consider maximizing the following function,

$$G_\lambda(p_{V_1|V_0}) = (1 - \lambda) I(V_1; Y|V_2) + \lambda I(V_2; Y)] \tag{57}$$

for $\lambda \in \left[\frac{1}{2}, 1 \right]$.

To simplify notation, we omit the conditioning on $V_0 = v_0$. Let $(V_1, V_2) \sim p_{V_1} p_{V_2}$, and consider the joint distribution

$$Q(v_1, v_2, x_1, x_2, y) \equiv p_{V_1}(v_1) p_{V_2}(v_2) \mathrm{Tr} \left[(L_1(x_1|v_1) \otimes L_2(x_2|v_2)) \varphi_{A_1 A_2} \right] P_{Y|X_1, X_2}(y|x_1, x_2) . \tag{58}$$

Then, consider the perturbed distribution,

$$Q_\epsilon(v_1) \equiv p_{V_1}(v_1)[1 + \epsilon f(v_1)] \tag{59}$$

and

$$Q_\epsilon(v_1, v_2, x_1, x_2, y) \equiv Q_\epsilon(v_1) p_{V_2}(v_2) \mathrm{Tr} \left[(L_1(x_1|v_0, v_1) \otimes L_2(x_2|v_0, v_2)) \varphi_{A_1 A_2} \right] P_{Y|X_1, X_2}(y|x_1, x_2) , \tag{60}$$

where $f(\cdot)$ are real-valued coefficients that satisfies $1 + \epsilon f(v_1) \geq 0$ for all $v_1 \in \mathcal{V}_1$. We note that f may depend on v_0 as well, since we have conditioned on $V_0 = v_0$.

We also require that the coefficients $f(\cdot)$ satisfy

$$\mathbb{E}f(V_1) = \sum_{v_1 \in \mathcal{V}_1} Q(v_1)f(v_1) = 0 \tag{61a}$$

and

$$\mathbb{E}[f(V_1)|X_1 = x_1, X_2 = x_2] = \sum_{(v_1,v_2) \in \mathcal{V}_1 \times \mathcal{V}_2} Q(v_1, v_2|x_1, x_2)f(v_1) = 0\,, \text{ for } (x_1, x_2) \in \mathcal{X}_1 \times \mathcal{X}_2\,. \tag{61b}$$

There exists a nonzero solution for this linear equation system, provided that

$$|\mathcal{V}_1| \geq |\mathcal{X}_1||\mathcal{X}_2| + 1 \tag{62}$$

since (61) consists of $|\mathcal{X}_1||\mathcal{X}_2| + 1$ constraints.

Observe that (61a) guarantees that Q_ϵ forms a probability distribution as well. Furthermore, by (61b), $Q_\epsilon(x_1, x_2) = Q(x_1, x_2)$, hence the marginal distribution of (X_1, X_2, Y) is preserved. Now, let $(\bar{V}_1, \bar{V}_2, \bar{X}_1, \bar{X}_2, \bar{Y}) \sim Q_\epsilon$ be a random tuple that is distributed according to the perturbed distribution. Then,

$$\begin{aligned} I(\bar{V}_1; \bar{Y}|\bar{V}_2) &= H(\bar{V}_1, \bar{V}_2) - H(\bar{Y}, \bar{V}_1, \bar{V}_2) - H(\bar{V}_2) + H(\bar{Y}, \bar{V}_2) \\ &= H(V_1, V_2) - H(Y, V_1, V_2) - H(V_2) + H(\bar{Y}, \bar{V}_2) + \epsilon[H_f(V_1, V_2) - H_f(Y, V_1, V_2)]\,, \end{aligned} \tag{63}$$

and

$$I(\bar{V}_2; \bar{Y}) = H(Y) + H(V_2) - H(\bar{Y}, \bar{V}_2)\,, \tag{64}$$

where

$$H_f(V_1, V_2) = -\sum_{v_1, v_2} Q(v_1, v_2)f(v_1)\log Q(v_1, v_2)\,, \tag{65}$$

$$H_f(Y, V_1, V_2) = -\sum_{v_1, v_2, y} Q(v_1, v_2, y)f(v_1)\log Q(v_1, v_2, y)\,. \tag{66}$$

If $p_{V_1} \equiv Q_0$ attains the maximum of the function G_λ (see (57)), then the derivatives satisfy

$$\frac{\partial}{\partial\epsilon}[(1-\lambda)I(\bar{V}_1; \bar{Y}|\bar{V}_2) + \lambda I(\bar{V}_2; \bar{Y})]\Big|_{\epsilon=0} = 0\,, \tag{67}$$

and

$$\frac{\partial^2}{\partial\epsilon^2}[(1-\lambda)I(\bar{V}_1; \bar{Y}|\bar{V}_2) + \lambda I(\bar{V}_2; \bar{Y})]\Big|_{\epsilon=0} \leq 0\,. \tag{68}$$

The latter reduces to

$$-(2\lambda - 1)\frac{\partial^2}{\partial\epsilon^2}H(\bar{Y}, \bar{V}_2)\Big|_{\epsilon=0} \leq 0\,, \tag{69}$$

which is equivalent to

$$\mathbb{E}\left([\mathbb{E}(f(V_1)|V_2, Y)]^2\right) \leq 0\,, \tag{70}$$

by [34, Lemma 2, part 2]. In particular, this holds with equality if

$$\mathbb{E}(f(V_1)|V_2 = v_2, Y = y) = 0\,, \tag{71}$$

for all (v_2, y) in the support of $Q(v_2, y)$, which implies

$$Q_\epsilon(v_2, y) = Q(v_2, y)\,. \tag{72}$$

Thus,

$$I(\bar{V}_1; \bar{Y}|\bar{V}_2) = I(V_1; Y|V_2) + \epsilon[H_f(V_1, V_2) - H_f(Y, V_1, V_2)]\,, \tag{73}$$

and

$$I(\bar{V}_2; \bar{Y}) = I(V_2; Y)\,. \tag{74}$$

By (67), we now have

$$\frac{\partial}{\partial\epsilon}[(1-\lambda)I(\bar{V}_1; \bar{Y}|\bar{V}_2) + \lambda I(\bar{V}_2; \bar{Y})]\Big|_{\epsilon=0} = (1-\lambda)(H_f(V_1, V_2) - H_f(Y, V_1, V_2)) = 0\,. \tag{75}$$

We deduce that the perturbed distribution attains the same (maximal) value of the function G_λ. Therefore, we can restrict the union to $|\mathcal{V}_1| = |\mathcal{X}_1||\mathcal{X}_2| + 1$. By symmetry, this holds for the cardinality of V_2 as well.

B. Outer Bound

Next, we show that union for the outer bound $\mathcal{O}_{\mathrm{ET}}(P_{Y|X_1,X_2})$ can be restricted as well. Based on exactly the same arguments, the union can be exhausted with pure states and with $|\mathcal{V}_0| \leq 3$.

We bound the alphabet size for the auxiliary variable V_1, using similar perturbation arguments. Here, however, we fix $p_{V_2|V_0,V_1}$ as well. As before, consider maximizing the following function,

$$\tilde{G}_\lambda(p_{V_1|V_0}) = (1-\lambda)I(V_1;Y|V_2) + \lambda I(V_2;Y)], \tag{76}$$

for $\lambda \in [0,1]$.

To simplify notation, we omit the conditioning on $V_0 = v_0$. Let $(V_1, V_2) \sim p_{V_1}p_{V_2|V_1}$, and consider the joint distribution

$$\tilde{Q}(v_1,v_2,x_1,x_2,y) \equiv p_{V_1}(v_1)p_{V_2|V_1}(v_2|v_1)\mathrm{Tr}\left[(L_1(x_1|v_1)\otimes L_2(x_2|v_2))\varphi_{A_1A_2}\right]P_{Y|X_1,X_2}(y|x_1,x_2). \tag{77}$$

Then, consider the perturbed distribution,

$$\tilde{Q}_\epsilon(v_1,v_2) = Q(v_1,v_2)[1+\epsilon\tilde{f}(v_1)], \tag{78}$$

and

$$\tilde{Q}_\epsilon(v_1,v_2,x_1,x_2,y) \equiv \tilde{Q}_\epsilon(v_1,v_2)\mathrm{Tr}\left[(L_1(x_1|v_1)\otimes L_2(x_2|v_2))\varphi_{A_1A_2}\right]P_{Y|X_1,X_2}(y|x_1,x_2), \tag{79}$$

where $\tilde{f}(\cdot)$ is a real-valued function that satisfies $1+\epsilon\tilde{f}(v_1) \geq 0$. We also require that $\tilde{f}$ satisfies the equation system (61), which guarantees that $\tilde{Q}_\epsilon(x_1,x_2) = Q(x_1,x_2)$. There exists a nonzero perturbation function that satisfies this system of $|\mathcal{X}_1||\mathcal{X}_2|+1$ linear equations, provided that $|\mathcal{V}_1| \geq |\mathcal{X}_1||\mathcal{X}_2|+1$.

Hence, for $(\check{V}_1,\check{V}_2,\check{X}_1,\check{X}_2,\check{Y}) \sim \tilde{Q}_\epsilon$,

$$I(\check{V}_1;\check{Y}|\check{V}_2) = H(\check{V}_1,\check{V}_2) - H(\check{Y},\check{V}_1,\check{V}_2) - H(\check{V}_2) + H(\check{Y},\check{V}_2) \tag{80}$$

$$= H(V_1,V_2) - H(Y,V_1,V_2) - H(V_2) + H(\check{Y},\check{V}_2) + \epsilon[H_{\tilde{f}}(V_1,V_2) - H_{\tilde{f}}(Y,V_1,V_2)] \tag{81}$$

and

$$I(\check{V}_2;\check{Y}) = H(\check{Y}) + H(\check{V}_2) - H(\check{Y},\check{V}_2) \tag{82}$$

$$= H(Y) + H(V_2) - H(\check{Y},\check{V}_2) \tag{83}$$

where the perturbed entropies $H_{\tilde{f}}$ are defined as in (65)-(66). If $p_{V_1} \equiv \tilde{Q}_0$ attains the maximum of $\tilde{G}_\lambda$, then

$$\frac{\partial^2}{\partial\epsilon^2}[(1-\lambda)I(\check{V}_1;\check{Y}|\check{V}_2) + \lambda I(\check{V}_2;\check{Y})]\Big|_{\epsilon=0} = -(2\lambda-1)\frac{\partial^2}{\partial\epsilon^2}H(\check{Y},\check{V}_2)\Big|_{\epsilon=0} \leq 0, \tag{84}$$

which is equivalent to

$$\mathbb{E}\left([\mathbb{E}(\tilde{f}(V_1)|V_2,Y)]^2\right) \leq 0 \tag{85}$$

by [34, Lemma 2, part 2]. In particular, this holds with equality if $\mathbb{E}(\tilde{f}(V_1)|V_2=v_2,Y=y) = 0$, in which case

$$\tilde{Q}_\epsilon(v_2,y) = \tilde{Q}(v_2,y). \tag{86}$$

It follows that

$$I(\check{V}_1;\check{Y}|\check{V}_2) = I(V_1;Y|V_2) + \epsilon[H_{\tilde{f}}(V_1,V_2) - H_{\tilde{f}}(Y,V_1,V_2)], \tag{87}$$

and

$$I(\check{V}_2;\check{Y}) = I(V_2;Y). \tag{88}$$

Since $\tilde{Q}_0$ attains the maximum of $\tilde{G}_\lambda$, we also have

$$\frac{\partial}{\partial\epsilon}[(1-\lambda)I(\check{V}_1;\check{Y}|\check{V}_2) + \lambda I(\check{V}_2;\check{Y})]\Big|_{\epsilon=0} = (1-\lambda)(H_{\tilde{f}}(V_1,V_2) - H_{\tilde{f}}(Y,V_1,V_2)) = 0. \tag{89}$$

We deduce that the perturbed distribution attains the maximal value of the function $\tilde{G}_\lambda$. Therefore, we can restrict the union to $|\mathcal{V}_1| = |\mathcal{V}_2| = |\mathcal{X}_1||\mathcal{X}_2|+1$. This completes the proof for the cardinality bounds. $\qquad\square$

VII. PROOF OF THEOREM 3

Consider the classical MAC $P_{Y|X_1,X_2}$ with entanglement resources between the transmitters (see Figure 1).

A. Achievability Proof

We show that for every $\zeta_1, \zeta_2, \varepsilon_0 > 0$, there exists a $(2^{n(R_1 - \zeta_1)}, 2^{n(R_2 - \zeta_2)}, n)$ code for $P_{Y|X_1, X_2}$ with entangled transmitters and ε_0-error, provided that $(R_1, R_2) \in \mathcal{R}_{\text{ET}}(P_{Y|X_1, X_2})$. To prove achievability, we use coded time sharing [53, Sec. 4.5.3].

Fix a joint distribution $p_{V_0} p_{V_1|V_0} p_{V_2|V_0}$, a bipartite state $\varphi_{A_1 A_2}$, and collection of POVMs $\mathcal{L}_k(v_0, v_k) = \{L_k(x_k|v_0, v_k)\}$ for $k = 1, 2$. Suppose that Alice 1 and Alice 2 share an n copies of the bipartite state,

$$\varphi_{A_1^n A_2^n} \equiv \varphi_{A_1 A_2}^{\otimes n}. \tag{90}$$

The code construction, encoding with shared entanglement, and decoding procedures are described below.

1) Code Construction: Select a random time-sharing sequence v_0^n, according to an i.i.d. distribution, $\prod_{i=1}^n p_{V_0}(v_{0,i})$. Furthermore, select 2^{nR_1} conditionally independent sequences, $v_1^n(m_1)$, $m_1 \in [1 : 2^{nR_1}]$, each distributed as $\prod_{i=1}^n p_{V_1|V_0}(v_1[i] | v_0[i])$. In a similar manner, select 2^{nR_2} sequences, $v_2^n(m_2)$, according to $\prod_{i=1}^n p_{V_2|V_0}(v_2[i] | v_0[i])$.

The auxiliary codebooks above are revealed to Alice 1, Alice 2, and Bob.

2) Encoder k: Given the message $m_k \in [1 : 2^{nR_k}]$ and the codebooks above, perform the measurement $\bigotimes_{i=1}^n \left(\mathcal{L}_k(v_0[i], v_k[i](m_k)) \right)$ on the entangled system A_k^n, and transmit the measurement outcome x_k^n through the channel, for $k = 1, 2$.

This yields the following input distribution,

$$f(x_1^n, x_2^n | m_1, m_2) = \text{Tr}\left[\left(L_1^n\left(x_1^n | v_0^n, v_1^n(m_1)\right) \otimes L_2^n\left(x_2^n | v_0^n, v_2^n(m_2)\right) \right) \varphi_{A_1^n A_2^n} \right]$$
$$= \prod_{i=1}^n \text{Tr}\left[\left(L_1\left(x_1[i] | v_0[i], v_1(m_1)[i]\right) \otimes L_2\left(x_2[i] | v_0[i], v_2(m_2)[i]\right) \right) \varphi_{A_1 A_2} \right], \tag{91}$$

where we use the short notations $L_k^n\left(x_k^n | v_0^n, v_k^n\right) \equiv \bigotimes_{i=1}^n L_k\left(x_k[i] | v_0[i], v_k[i]\right)$, for $k = 1, 2$.

3) Decoder: Let $\delta > 0$ be arbitrarily small. Find a unique pair $(\hat{m}_1, \hat{m}_2)$ such that $(v_0^n, v_1^n(\hat{m}_1), v_2^n(\hat{m}_2), y^n) \in \mathcal{A}_\delta^{(n)}(p_{V_0, V_1, V_2, Y})$, where the marginal distribution $p_{V_0, V_1, V_2, Y}$ is induced by the following joint distribution,

$$p_{V_0, V_1, V_2, X_1, X_2, Y}(v_0, v_1, v_2, x_1, x_2, y) = p_{V_0}(v_0) p_{V_1|V_0}(v_1|v_0) p_{V_2|V_0}(v_2|v_0)$$
$$\cdot \text{Tr}\left[(L_1(x_1|v_0, v_1) \otimes L_2(x_2|v_0, v_2)) \varphi_{A_1 A_2} \right] \cdot P_{Y|X_1, X_2}(y|x_1, x_2). \tag{92}$$

If there is no such pair $(\hat{m}_1, \hat{m}_2)$, or more than one, declare an error.

4) Analysis of Probability of Error: We use the notation $\varepsilon_i(\delta)$, $i = 1, 2, \ldots$, for terms that tend to zero as $\delta \to 0$. At first, suppose that the messages are chosen at random according to a uniform distribution. By symmetry, we may assume without loss of generality that the transmitters send the messages $M_1 = M_2 = 1$. Consider the following error events,

$$\mathcal{E}_0 = \{ (V_0^n, V_1^n(1), V_2^n(1), Y^n) \notin \mathcal{A}_{\delta_1}^{(n)}(p_{V_0, V_1, V_2, Y}) \} \tag{93}$$
$$\mathcal{E}_1 = \{ (V_0^n, V_1^n(m_1), V_2^n(1), Y^n) \in \mathcal{A}_\delta^{(n)}(p_{V_0, V_1, V_2, Y}), \text{ for some } m_1 \neq 1 \} \tag{94}$$
$$\mathcal{E}_2 = \{ (V_0^n, V_1^n(1), V_2^n(m_2), Y^n) \in \mathcal{A}_\delta^{(n)}(p_{V_0, V_1, V_2, Y}), \text{ for some } m_2 \neq 1 \} \tag{95}$$
$$\mathcal{E}_3 = \{ (V_0^n, V_1^n(m_1), V_2^n(m_2), Y^n) \in \mathcal{A}_\delta^{(n)}(p_{V_0, V_1, V_2, Y}), \text{ for some } m_1 \neq 1 \text{ and } m_2 \neq 1 \} \tag{96}$$

with $\delta_1 \equiv \delta/(2|\mathcal{V}_0||\mathcal{V}_1||\mathcal{V}_2|)$. By the union of events bound, the expected probability of error is bounded by

$$\mathbb{E}\left[P_e^{(n)}(\mathcal{C}|1, 1) \right] \leq \text{Pr}(\mathcal{E}_0) + \text{Pr}(\mathcal{E}_1 \mid \mathcal{E}_0^c) + \text{Pr}(\mathcal{E}_2 \mid \mathcal{E}_0^c) + \text{Pr}(\mathcal{E}_3 \mid \mathcal{E}_0^c) \tag{97}$$

where the expectation on the left-hand side is with respect to the random auxiliary codebooks, and the conditioning on $M_1 = M_2 = 1$ is omitted from the right-hand side for convenience of notation. Observe that the (classical) codewords V_0^n, $V_1^n(1)$, $V_2^n(1)$, channel inputs X_1^n, X_2^n, and channel output Y^n, are jointly i.i.d. according to $p_{V_0, V_1, V_2, X_1, X_2, Y}$, as in (92). Hence, by the weak law of large numbers, the first probability term, $\text{Pr}(\mathcal{E}_0)$, tends to zero as $n \to \infty$ [54] [55, Th. 1.1].

As for the second error term, we have by the union bound:

$$\text{Pr}(\mathcal{E}_1 \mid \mathcal{E}_0^c) \leq \sum_{m_1 \neq 1} \text{Pr}\left((V_0^n, V_1^n(m_1), V_2^n(1), Y^n) \in \mathcal{A}_\delta^{(n)}(p_{V_0, V_1, V_2, Y}) \mid \mathcal{E}_0^c \right). \tag{98}$$

Given $\mathcal{E}_0^c$, it follows that $(V_0^n, V_2^n(1)) \in \mathcal{A}^\delta(p_{V_0, V_2})$. Thus, each summand is bounded by

$$\sum_{(v_0^n, v_2^n) \in \mathcal{A}_\delta^{(n)}(p_{V_0, V_2})} p_{V_0, V_2}^n(v_0^n, v_2^n) \left[\sum_{v_1^n, y^n \,:\, (v_0^n, v_1^n, v_2^n, y^n) \in \mathcal{A}_\delta^{(n)}(p_{V_0, V_1, V_2, Y})} p_{V_1|V_0}^n(v_1^n|v_0^n) \cdot p_{Y^n|V_0^n, V_2^n}(y^n|v_0^n, v_2^n) \right] \tag{99}$$

since for every $m_1 \neq 1$, the codeword $V_1^n(m_1)$ is conditionally independent of the sequence pair $(V_2^n(1), Y^n)$, given $V_0^n = v_0^n$. Then, by standard method-of-types arguments, the sum within the square brackets is bounded by $2^{-n(I(V_1;Y|V_0V_2)-\varepsilon_1(\delta))}$ [54] [55, Th. 1.3]. Hence, by (98),

$$\Pr\left(\mathscr{E}_1 \mid \mathscr{E}_0^c\right) \leq 2^{-n[I(V_1;Y|V_0V_2)-R_1-\varepsilon_1(\delta)]}. \tag{100}$$

Thereby, the term $\Pr\left(\mathscr{E}_1 \mid \mathscr{E}_0^c\right)$ tends to zero as $n \to \infty$, provided that

$$R_1 < I(V_1;Y|V_0V_2) - \varepsilon_1(\delta). \tag{101}$$

Following similar arguments, the probability terms $\Pr\left(\mathscr{E}_2 \mid \mathscr{E}_0^c\right)$ and $\Pr\left(\mathscr{E}_3 \mid \mathscr{E}_0^c\right)$ also tend to zero, provided that

$$R_2 < I(V_2;Y|V_0V_1) - \varepsilon_2(\delta) \tag{102}$$

and

$$R_1 + R_2 < I(V_1V_2;Y|V_0) - \varepsilon_3(\delta). \tag{103}$$

We conclude that the average probability of error, $\mathbb{E}\left[\overline{P}_e^{(n)}(\mathscr{C})\right]$, averaged over the messages and the class of random codebooks above, tends to zero as $n \to \infty$ (see (3)). Therefore, there must exist a $(2^{nR_1}, 2^{nR_2}, n)$ code such that the message-average error probability, as defined in (3), is bounded by $\overline{P}_e^{(n)}(\mathscr{C}) < \frac{1}{9}\varepsilon_0$, for a sufficiently large n. Based on the arguments in [30], it follows that the rate pair (R_1, R_2) is also achievable with a maximal error criterion, i.e., such that $P_e^{(n)}(\mathscr{C}) \leq \varepsilon_0$. For completeness, we show this in the appendix. To show achievability of the regularized formula, apply the coding scheme above to the product channel $P_{Y|X_1,X_2}^\ell$. This completes the achievability proof. $\qquad\square$

B. Converse Proof

Consider the classical MAC with entanglement resources between the transmitters. We now show the converse part for the regularized characterization. Suppose that Alice 1 and Alice 2 share a bipartite state $\Psi_{E_1 E_2}$. Then, Alice k chooses a message M_k uniformly at random, hence the message pair (M_1, M_2) is uniformly distributed over $[1:2^{nR_1}] \times [1:2^{nR_2}]$. She encodes her message by performing a measurement $\mathcal{F}_k^{(M_k)} = \{F_{x_k^n}^{(M_k)}\}$ on her share of the entangled resources, E_k, and sends the measurement outcome X_k^n over the channel. Bob receives the output Y^n and finds an estimate $(\widehat{M}_1, \widehat{M}_2) = g(Y^n)$ of the message pair.

Now, consider a sequence of codes $(\Psi_n, \mathcal{F}_{1n}, \mathcal{F}_{2n}, g_n)$ such that the average probability of error tends to zero, hence the error probabilities $\Pr\left(\widehat{M}_1 \neq M_1 | M_2\right)$, $\Pr\left(\widehat{M}_2 \neq M_2 | M_1\right)$, and $\Pr\left((\widehat{M}_1, \widehat{M}_2) \neq (M_1, M_2)\right)$, are bounded by some α_n which tends to zero as $n \to \infty$. By Fano's inequality [31], it follows that

$$H(M_1|\widehat{M}_1, M_2) \leq n\varepsilon_{1n}, \tag{104}$$
$$H(M_2|\widehat{M}_2, M_1) \leq n\varepsilon_{2n}, \tag{105}$$
$$H(M_1, M_2|\widehat{M}_1, \widehat{M}_2) \leq n\varepsilon_{3n} \tag{106}$$

where ε_{kn} tend to zero as $n \to \infty$.

Hence,

$$\begin{aligned}
nR_1 = H(M_1|M_2) &= I(M_1; \widehat{M}_1|M_2) + H(M_1|\widehat{M}_1 M_2) \\
&\leq I(M_1; \widehat{M}_1|M_2) + n\varepsilon_{1n} \\
&\leq I(M_1; Y^n|M_2) + n\varepsilon_{1n}
\end{aligned} \tag{107}$$

where the last inequality follows from the data processing inequality. Following similar arguments, we also have

$$nR_2 \leq I(M_2; Y^n|M_1) + n\varepsilon_{2n} \tag{108}$$

and

$$n(R_1 + R_2) \leq I(M_1 M_2; Y^n) + n\varepsilon_{3n} \tag{109}$$

The proof follows by identifying V_1 and V_2 as M_1 and M_2, respectively, and taking $V_0 = \emptyset$.

VIII. PROOF OF THEOREM 4 (OUTER BOUND)

We prove the outer bound on the capacity region of the classical MAC with entangled transmitters. Consider the rate bound in (107). Applying the chain rule, we can rewrite this as

$$n(R_1 - \varepsilon_{1n}) \le \sum_{i=1}^{n} I\left(M_1; Y[i] \,\middle|\, Y^{i-1}, M_2\right). \tag{110}$$

where $Y^{i-1} \equiv Y[1], \ldots, Y[i-1]$, for $i = 2, \ldots, n$, and $Y^0 \equiv \emptyset$. Define $V_0[i] \equiv Y^{i-1}$, $V_1[i] = (M_1, Y^{i-1})$, and $V_2[i] = (M_2, Y^{i-1})$, for $i = 1, \ldots, n$. Notice that $V_1[i]$ and $V_2[i]$ are *not* conditionally independent given $V_0[i]$, because conditioning on the output may introduce statistical dependence between the random messages.

Furthermore, since Alice 1 performs a measurement that depends on her message M_1 alone, and similarly for Alice 2, the channel inputs $X_1[i]$ and $X_2[i]$ can also be obtained as the output of a measurement as defined below. Consider a measurement channel

$$\widetilde{\mathcal{F}}_{E_1 \to X_1}^{(i,m_1)} \otimes \widetilde{\mathcal{F}}_{E_2 \to X_2}^{(i,m_2)} \tag{111}$$

such that

$$\widetilde{\mathcal{F}}_{E_1 \to X_1}^{(i,m_1)}(Q') = \sum_{a_1 \in \mathcal{X}_1} \left[\sum_{x_1^n \in \mathcal{X}_1^n \,:\, x_1[i]=a_1} \mathrm{Tr}\left(F_{x_1^n}^{(m_1)} Q' \right) \right] |a_1\rangle\langle a_1| \tag{112}$$

$$\widetilde{\mathcal{F}}_{E_2 \to X_2}^{(i,m_2)}(Q'') = \sum_{a_2 \in \mathcal{X}_2} \left[\sum_{x_2^n \in \mathcal{X}_2^n \,:\, x_2[i]=a_2} \mathrm{Tr}\left(F_{x_2^n}^{(m_2)} Q'' \right) \right] |a_2\rangle\langle a_2| \tag{113}$$

for every pair of operators Q' and Q'' on $\mathcal{H}_{E_1}$ and $\mathcal{H}_{E_2}$, respectively. Let $\Psi_{E_1 E_2} = \sum_{j,\ell} Q'_j \otimes Q''_\ell$ be an arbitrary decomposition of the bipartite state. Hence, by linearity,

$$\left(\widetilde{\mathcal{F}}_{E_1 \to X_1}^{(i,m_1)} \otimes \widetilde{\mathcal{F}}_{E_2 \to X_2}^{(i,m_2)} \right)(\Psi_{E_1 E_2})$$

$$= \sum_{t,r} \left(\widetilde{\mathcal{F}}_{E_1 \to X_1}^{(i,m_1)} \otimes \widetilde{\mathcal{F}}_{E_2 \to X_2}^{(i,m_2)} \right)(Q'_j \otimes Q''_\ell)$$

$$= \sum_{j,\ell} \sum_{a_1 \in \mathcal{X}_1} \left[\sum_{x_1^n \in \mathcal{X}_1^n \,:\, x_1[i]=a_1} \mathrm{Tr}\left(F_{x_1^n}^{(m_1)} Q'_j \right) \right] |a_1\rangle\langle a_1| \otimes \sum_{a_2 \in \mathcal{X}_2} \left[\sum_{x_2^n \in \mathcal{X}_2^n \,:\, x_2[i]=a_2} \mathrm{Tr}\left(F_{x_2^n}^{(m_2)} Q''_\ell \right) \right] |a_2\rangle\langle a_2|$$

$$= \sum_{(a_1,a_2) \in \mathcal{X}_1 \times \mathcal{X}_2} \left[\sum_{\substack{(x_1^n, x_2^n) \in \mathcal{X}_1^n \times \mathcal{X}_2^n \,: \\ x_1[i]=a_1 \,,\, x_2[i]=a_2}} \sum_{j,\ell} \mathrm{Tr}\left((F_{x_1^n}^{(m_1)} \otimes F_{x_2^n}^{(m_2)})(Q'_j \otimes Q''_\ell) \right) \right] |a_1, a_2\rangle\langle a_1, a_2|$$

$$= \sum_{(a_1,a_2) \in \mathcal{X}_1 \times \mathcal{X}_2} \left[\sum_{\substack{(x_1^n, x_2^n) \in \mathcal{X}_1^n \times \mathcal{X}_2^n \,: \\ x_1[i]=a_1 \,,\, x_2[i]=a_2}} \mathrm{Tr}\left((F_{x_1^n}^{(m_1)} \otimes F_{x_2^n}^{(m_2)})\Psi_{E_1 E_2} \right) \right] |a_1, a_2\rangle\langle a_1, a_2| \tag{114}$$

as expected. We deduce that the channel inputs $X_1[i]$ and $X_2[i]$ can be obtained from a product of measurements of the form $L_1(x_1|i, V_0[i], V_1[i])$ and $L_2(x_2|i, V_0[i], V_2[i])$, as required.

Then, we can rewrite (110) as

$$R_1 - \varepsilon_{1n} \le \frac{1}{n} \sum_{i=1}^{n} I\left(V_1[i]; Y[i] \,\middle|\, V_0[i], V_2[i] \right)$$

$$= I\left(V_1[J]; Y[J] \,\middle|\, V_0[J], V_2[J], J \right) \tag{115}$$

where the index J is drawn uniformly at random from $\{1, \ldots, n\}$, and it is uncorrelated with the previous systems. Following the same considerations,

$$R_2 - \varepsilon_{2n} \le I\left(V_2[J]; Y[J] \,\middle|\, V_0[J], V_1[J], J \right) \tag{116}$$

and

$$R_1 + R_2 - \varepsilon_{3n} \le I\left(V_1[J] V_2[J]; Y[J] \,\middle|\, V_0[J], J \right). \tag{117}$$

The proof follows by defining $V_0 \equiv (J, V_0[J])$, $V_k \equiv V_k[J]$, for $k = 1, 2$, then $X_k \equiv X_k[J]$, and $Y \equiv Y[J]$. $\square$

IX. PROOF OF THEOREM 6 (CLASSICAL CONFERENCING)

Consider the classical MAC $P_{Y|X_1,X_2}$ with entanglement resources and classical conferencing between the transmitters (see Figure 2).

A. Achievability Proof

We show that for every $\delta_1, \delta_2, \varepsilon_0 > 0$, there exists $(2^{n(R_1-\delta_1)}, 2^{n(R_2-\delta_2)}, n)$ code that has an ε_0-error, provided that $(R_1, R_2) \in \mathcal{R}_{\text{ET-C}}(P_{Y|X_1,X_2}, C_{12}, C_{21})$. We use the classical rate splitting method due to Willems [56].

As in the previous section, fix a joint distribution $p_{V_0} p_{V_1|V_0} p_{V_2|V_0}$, a bipartite state $\varphi_{A_1 A_2}$, and collection of POVMs $\mathcal{L}_k(v_0, v_k)$ for $k = 1, 2$. Suppose that Alice 1 and Alice 2 share $\varphi_{A_1^n A_2^n} \equiv \varphi_{A_1 A_2}^{\otimes n}$.

The code construction, conferencing protocol, encoding with shared entanglement, and decoding procedures are given below.

1) Code Construction: Split the rates of each user into $R_k = R'_k + R''_k$ for $k = 1, 2$. We think of R'_k as the rate of *common* information that is shared between the transmitters through conferencing, and R''_k as the rate of the *private* information of each transmitter.

Select $2^{n(R'_1 + R'_2)}$ sequences $v_0^n(m'_1, m'_2)$, for $m'_k \in [1 : 2^{nR'_k}]$, independently at random, each i.i.d. $\sim p_{V_0}$. For every (m'_1, m'_2), select $2^{nR''_1}$ conditionally independent sequences, $v_1^n(m'_1, m'_2, m''_1)$, $m''_1 \in [1 : 2^{nR''_1}]$, each distributed as $\prod_{i=1}^n p_{V_1|V_0}(v_1[i] | v_0(m'_1, m'_2)[i])$. In a similar manner, select $2^{nR''_2}$ sequences, $v_2^n(m'_1, m'_2, m''_2)$, according to $\prod_{i=1}^n p_{V_2|V_0}(v_2[i] | v_0(m'_1, m'_2)[i])$. The auxiliary codebooks are revealed to all parties.

2) Conferencing: The encoders share the messages m'_1 and m'_2 between them through a single round of conferencing $(T = 1)$. This is permissible if the common information rates are each bounded by the respective conferencing capacity, i.e.,

$$R'_1 \leq C_{12}, \ R'_2 \leq C_{21} \tag{118}$$

as this satisfies the conferencing limits in (12).

3) Encoder 1: Given the message (m'_1, m''_1), the conferencing message m'_2, and the codebooks above, perform the measurement $\bigotimes_{i=1}^n \left(\mathcal{L}_1(v_0(m'_1, m'_2)[i], v_1(m'_1, m'_2, m''_1)[i]) \right)$ on the entangled system A_1^n, and transmit the measurement outcome x_1^n through the channel.

4) Encoder 2: Given the message (m'_2, m''_2), the conferencing message m'_1, and the codebooks above, perform the measurement $\bigotimes_{i=1}^n \left(\mathcal{L}_2(v_0(m'_1, m'_2)[i], v_2(m'_1, m'_2, m''_2)[i]) \right)$ on the entangled system A_2^n, and transmit the measurement outcome x_2^n through the channel.

The resulting input distribution is given by

$$f(x_1^n, x_2^n | m'_1, m''_1, m'_2, m''_2) = \prod_{i=1}^n \text{Tr} \left[\left(L_1(x_1[i] | v_0[i], v_1[i]) \otimes L_2(x_2[i] | v_0[i], v_2[i]) \right) \varphi_{A_1 A_2} \right] \tag{119}$$

where $v_0^n \equiv v_0^n(m'_1, m'_2)$ and $v_k^n \equiv v_k^n(m'_1, m'_2, m''_k)$ for $k = 1, 2$.

5) Decoder: In a similar manner as in the previous proof, find a unique tuple $(\hat{m}'_1, \hat{m}''_1, \hat{m}'_2, \hat{m}''_2)$ such that $(v_0^n(\hat{m}'_1, \hat{m}'_2), v_1^n(\hat{m}'_1, \hat{m}'_2, \hat{m}''_1), v_2^n(m'_1, m'_2, \hat{m}''_2), y^n) \in \mathcal{A}_\delta^{(n)}(p_{V_0, V_1, V_2, Y})$. As before, the joint distribution is

$$p_{V_0, V_1, V_2, X_1, X_2, Y}(v_0, v_1, v_2, x_1, x_2, y) = p_{V_0}(v_0) p_{V_1|V_0}(v_1|v_0) p_{V_2|V_0}(v_2|v_0)$$
$$\cdot \text{Tr} \left[(L_1(x_1|v_0, v_1) \otimes L_2(x_2|v_0, v_2)) \varphi_{A_1 A_2} \right] \cdot P_{Y|X_1, X_2}(y|x_1, x_2). \tag{120}$$

If there is none, or more than one such pair, declare an error.

6) Analysis of Probability of Error: At first, suppose that the messages are uniformly distributed. By symmetry, we may assume without loss of generality that the transmitters send $M'_1 = M''_1 = M'_2 = M''_2 = 1$.

We consider the error events below, whereby the correct messages do not satisfy the decoding rule (the event $\mathscr{E}_0$ below), or alternatively, the decoder has ambiguity for the common information alone ($\mathscr{E}'_0$), for one of the private messages ($\mathscr{E}_1$ and $\mathscr{E}_2$),

both private messages ($\mathscr{E}_3$), the common information and one private message ($\mathscr{E}_4$ and $\mathscr{E}_5$), or all messages ($\mathscr{E}_6$):

$$\mathscr{E}_0 = \{(V_0^n(1,1), V_1^n(1,1,1), V_2^n(1,1,1), Y^n) \notin \mathcal{A}_{\delta_1}^{(n)}(p_{V_0,V_1,V_2,Y})\} \tag{121}$$

$$\mathscr{E}_0' = \{(V_0^n(m_1',m_2'), V_1^n(m_1',m_2',1), V_2^n(m_1',m_2',1), Y^n) \in \mathcal{A}_{\delta}^{(n)}(p_{V_0,V_1,V_2,Y}), \text{ for some } (m_1',m_2') \neq (1,1)\} \tag{122}$$

$$\mathscr{E}_1 = \{(V_0^n(1,1), V_1^n(1,1,m_1''), V_2^n(1,1,1), Y^n) \in \mathcal{A}_{\delta}^{(n)}(p_{V_0,V_1,V_2,Y}), \text{ for some } m_1'' \neq 1\} \tag{123}$$

$$\mathscr{E}_2 = \{(V_0^n(1,1), V_1^n(1,1,1), V_2^n(1,1,m_2''), Y^n) \in \mathcal{A}_{\delta}^{(n)}(p_{V_0,V_1,V_2,Y}), \text{ for some } m_2'' \neq 1\} \tag{124}$$

$$\mathscr{E}_3 = \{(V_0^n(1,1), V_1^n(1,1,m_1''), V_2^n(1,1,m_2''), Y^n) \in \mathcal{A}_{\delta}^{(n)}(p_{V_0,V_1,V_2,Y}), \text{ for some } m_1'' \neq 1 \text{ and } m_2'' \neq 1\} \tag{125}$$

$$\mathscr{E}_4 = \{(V_0^n(m_1',m_2'), V_1^n(m_1',m_2',m_1''), V_2^n(m_1',m_2',1), Y^n) \in \mathcal{A}_{\delta}^{(n)}(p_{V_0,V_1,V_2,Y}), \text{ for some } (m_1',m_2') \neq (1,1) \text{ and } m_1'' \neq 1\} \tag{126}$$

$$\mathscr{E}_5 = \{(V_0^n(m_1',m_2'), V_1^n(m_1',m_2',1), V_2^n(m_1',m_2',m_2''), Y^n) \in \mathcal{A}_{\delta}^{(n)}(p_{V_0,V_1,V_2,Y}), \text{ for some } (m_1',m_2') \neq (1,1) \text{ and } m_2'' \neq 1\} \tag{127}$$

$$\mathscr{E}_6 = \{(V_0^n(m_1',m_2'), V_1^n(m_1',m_2',m_1''), V_2^n(m_1',m_2',m_2''), Y^n) \in \mathcal{A}_{\delta}^{(n)}(p_{V_0,V_1,V_2,Y}), \text{ for some } (m_1',m_2') \neq (1,1), m_1'' \neq 1,$$
$$\text{and } m_2'' \neq 1\} \tag{128}$$

with $\delta_1 \equiv \delta/(2|\mathcal{V}_0||\mathcal{V}_1||\mathcal{V}_2|)$. By the union of events bound, the expected probability of error is bounded by

$$\mathbb{E}\left[P_e^{(n)}(\mathscr{C}|1,1)\right] \leq \Pr(\mathscr{E}_0) + \Pr(\mathscr{E}_0') + \Pr(\mathscr{E}_1 \mid \mathscr{E}_0^c)$$
$$+ \Pr(\mathscr{E}_2 \mid \mathscr{E}_0^c) + \Pr(\mathscr{E}_3 \mid \mathscr{E}_0^c) + \Pr(\mathscr{E}_4) \tag{129}$$

where the expectation is with respect to the random auxiliary codebooks, and the conditioning on $M_1' = M_1'' = M_2' = M_2'' = 1$ is omitted from the right-hand side for convenience of notation. Observe that the codewords $V_0^n(1,1)$, $V_1^n(1,1,1)$, $V_2^n(1,1,1)$, channel inputs X_1^n, X_2^n, and channel output Y^n, are jointly i.i.d. according to $p_{V_0,V_1,V_2,X_1,X_2,Y}$, as in (120). As in the proof of Theorem 3, we observe that $\Pr(\mathscr{E}_0)$, tends to zero as $n \to \infty$ by the weak law of large numbers.

Furthermore, by the union bound, the second error term is bounded by

$$\Pr(\mathscr{E}_0') \leq \sum_{(m_1',m_2') \neq (1,1)} \Pr\left((V_0^n(m_1',m_2'), V_1^n(m_1',m_2',1), V_2^n(m_1',m_2',1), Y^n) \in \mathcal{A}_{\delta}^{(n)}(p_{V_1,V_2,Y})\right). \tag{130}$$

By standard method-of-types arguments, each term is then bounded by $2^{-n(I(V_0V_1V_2;Y)-\varepsilon_0(\delta))}$ [54] [55, Th. 1.3]. Since $V_0 \multimap (V_1,V_2) \multimap Y$ form a Markov chain, we have $I(V_0V_1V_2;Y) = I(V_1V_2;Y)$. It follows that

$$\Pr(\mathscr{E}_0') \leq 2^{-n[I(V_1V_2;Y)-R_1'-R_2'-\varepsilon_0(\delta)]}. \tag{131}$$

Thereby, the term $\Pr(\mathscr{E}_0')$ tends to zero as $n \to \infty$, provided that

$$R_1' + R_2' < I(V_1V_2;Y) - \varepsilon_0(\delta). \tag{132}$$

Now, consider the third error term:

$$\Pr(\mathscr{E}_1 \mid \mathscr{E}_0^c) \leq \sum_{m_1'' \neq 1} \Pr\left((V_0^n(1,1), V_1^n(1,1,m_1''), V_2^n(1,1,1), Y^n) \in \mathcal{A}_{\delta}^{(n)}(p_{V_0,V_1,V_2,Y})\right). \tag{133}$$

Given $\mathscr{E}_0^c$, it follows that $(V_0^n(1,1), V_2^n(1,1,1)) \in \mathcal{A}^{\delta}(p_{V_0V_2})$. Thus, each summand is bounded by

$$\sum_{(v_0^n,v_2^n) \in \mathcal{A}_{\delta}^{(n)}(p_{V_0,V_2})} p_{V_0,V_2}^n(v_0^n,v_2^n) \left[\sum_{v_1^n,y^n \,:\, (v_0^n,v_1^n,v_2^n,y^n) \in \mathcal{A}_{\delta}^{(n)}(p_{V_0,V_1,V_2,Y})} p_{V_1|V_0}^n(v_1^n|v_0^n) \cdot p_{Y^n|V_0^n,V_2^n}(y^n|v_0^n,v_2^n) \right] \tag{134}$$

since for every $m_1'' \neq 1$, the codeword $V_1^n(1,1,m_1'')$ is conditionally independent of the sequence pair $(V_2^n(1,1,1), Y^n)$, given $V_0^n(1,1) = v_0^n$. As the sum within the square brackets is bounded by $2^{-n(I(V_1;Y|V_0V_2)-\varepsilon_1(\delta))}$ (see [54] [55, Th. 1.3]), we have

$$\Pr(\mathscr{E}_1 \mid \mathscr{E}_0^c) \leq 2^{-n[I(V_1;Y|V_0V_2)-R_1''-\varepsilon_1(\delta)]}. \tag{135}$$

Thereby, the term $\Pr(\mathscr{E}_1 \mid \mathscr{E}_0^c)$ tends to zero as $n \to \infty$, provided that $R_1'' < I(V_1;Y|V_0V_2) - \varepsilon_1(\delta)$. Since $R_1 = R_1' + R_1''$, this requires

$$R_1 < I(V_1;Y|V_0V_2) + R_1' - \varepsilon_1(\delta). \tag{136}$$

Recall that by (118), $R_1' \leq C_{12}$.

Following similar arguments, the probability terms $\Pr(\mathscr{E}_2 \mid \mathscr{E}_0^c)$ and $\Pr(\mathscr{E}_3 \mid \mathscr{E}_0^c)$ also tend to zero, provided that

$$R_2 < I(V_2;Y|V_0V_1) + R_2' - \varepsilon_2(\delta) \tag{137}$$

and

$$R_1 + R_2 < I(V_1V_2; Y|V_0) + R'_1 + R'_2 - \varepsilon_3(\delta). \tag{138}$$

By (118), $R'_1 \leq C_{12}$ and $R'_2 \leq C_{21}$. Similarly, the remaining error terms, $\Pr\left(\mathscr{E}''_j \mid \mathscr{E}^c_0\right)$, $j \in \{4,5,6\}$, tend to zero if $R'_1 + R'_2 + R''_1 < I(V_1V_2; Y) - \varepsilon_4(\delta)$, $R'_1 + R'_2 + R''_2 < I(V_1V_2; Y) - \varepsilon_5(\delta)$, and $R'_1 + R'_2 + R''_1 + R''_2 < I(V_1V_2; Y) - \varepsilon_6(\delta)$. Thus, it suffices to require

$$R_1 + R_2 < I(V_1V_2; Y) - \max_{j \in \{4,5,6\}} \varepsilon_j(\delta). \tag{139}$$

We conclude that the average probability of error, $\mathbb{E}\left[\overline{P}^{(n)}_e(\mathscr{C})\right]$, averaged over the messages and the class of random codebooks above, tends to zero as $n \to \infty$ (see (3)). Therefore, there must exist a $(2^{nR_1}, 2^{nR_2}, n)$ code such that the message-average error probability vanishes. Based on the arguments in the appendix, it follows that the rate pair (R_1, R_2) is also achievable with a maximal error criterion, i.e., such that $P^{(n)}_e(\mathscr{C}) \leq \varepsilon_0$. Achievability for the regularized formula follows by applying the coding scheme above to the product MAC $P^\ell_{Y|X_1,X_2}$. This completes the achievability proof.

B. Converse Proof

Consider the classical MAC with entanglement resources and classical conferencing between the transmitters. We now show the converse part. Suppose that Alice 1 and Alice 2 share a bipartite state $\Psi_{E_1E_2}$. Then, Alice k chooses a message M_k, uniformly at random from $[1 : 2^{nR_k}]$, for $k = 1, 2$. They perform a conferencing protocol, whereby Alice 1 sends $\omega^T_{1 \to 2} = s^T_{1 \to 2}(M_1, M_2)$ to Alice 2, and Alice 2 sends $\omega^T_{2 \to 1} = s^T_{2 \to 1}(M_1, M_2)$ to Alice 1. Next, Alice 1 encodes her message M_1 by performing a measurement $\mathcal{F}^{(M_1, \omega^T_{2 \to 1})}_1 = \{F^{(M_1, \omega^T_{2 \to 1})}_{x^n_1}\}$ on her share of the entangled resources, E_1, and transmits the measurement outcome X^n_1 over the channel. In the same manner, Alice 2 performs an encoding measurement $\mathcal{F}^{(M_2, \omega^T_{1 \to 2})}_2$ that produces her transmission, X^n_2. Bob receives the output Y^n and finds an estimate $(\widehat{M}_1, \widehat{M}_2) = g(Y^n)$ of the message pair.

Now, consider a sequence of (C_{12}, C_{21})-permissible codes $(\Psi_n, s_{1 \to 2(n)}, s_{2 \to 1(n)}, \mathcal{F}_{1(n)}, \mathcal{F}_{2(n)}, g_n)$ such that the average probability of error tends to zero, hence the error probabilities $\Pr\left(\widehat{M}_1 \neq M_1|M_2\right)$, $\Pr\left(\widehat{M}_2 \neq M_2|M_1\right)$, and $\Pr\left((\widehat{M}_1, \widehat{M}_2) \neq (M_1, M_2)\right)$, are bounded by some α_n which tends to zero as $n \to \infty$. By Fano's inequality [31],

$$H(M_1|\widehat{M}_1, M_2) \leq n\varepsilon_{1n}, \tag{140}$$

$$H(M_2|\widehat{M}_2, M_1) \leq n\varepsilon_{2n}, \tag{141}$$

$$H(M_1, M_2|\widehat{M}_1, \widehat{M}_2) \leq n\varepsilon_{3n} \tag{142}$$

where ε_{kn} tend to zero as $n \to \infty$. Since conditioning cannot increase entropy, we also have

$$H(M_1|\widehat{M}_1, M_2, \omega^T_{1 \to 2}, \omega^T_{2 \to 1}) \leq n\varepsilon_{1n}, \tag{143}$$

$$H(M_2|\widehat{M}_2, M_1, \omega^T_{1 \to 2}, \omega^T_{2 \to 1}) \leq n\varepsilon_{2n}, \tag{144}$$

$$H(M_1, M_2|\widehat{M}_1, \widehat{M}_2, \omega^T_{1 \to 2}, \omega^T_{2 \to 1}) \leq n\varepsilon_{3n}. \tag{145}$$

Now,

$$\begin{aligned}
nR_1 = H(M_1|M_2) &= I(M_1; \widehat{M}_1\omega^T_{1 \to 2}, \omega^T_{2 \to 1}|M_2) + H(M_1|\widehat{M}_1 M_2 \omega^T_{1 \to 2} \omega^T_{2 \to 1}) \\
&\leq I(M_1; \widehat{M}_1\omega^T_{1 \to 2}\omega^T_{2 \to 1}|m_2) + n\varepsilon_{1n} \\
&\leq I(M_1; Y^n\omega^T_{1 \to 2}\omega^T_{2 \to 1}|M_2) + n\varepsilon_{1n}
\end{aligned} \tag{146}$$

where the last inequality follows from the data processing inequality. Applying the chain rule, we can rewrite this as

$$\begin{aligned}
n(R_1 - \varepsilon_{1n}) &\leq I(M_1; \omega^T_{1 \to 2}\omega^T_{2 \to 1}|M_2) + I(M_1; Y^n|M_2, \omega^T_{1 \to 2}\omega^T_{2 \to 1}) \\
&= \sum_{t=1}^{T} I\left(M_1; \omega_{1 \to 2}[t]\omega_{2 \to 1}[t] \mid M_2, \omega^{t-1}_{1 \to 2}\omega^{t-1}_{2 \to 1}\right) + I(M_1; Y^n|M_2, \omega^T_{1 \to 2}\omega^T_{2 \to 1})
\end{aligned} \tag{147}$$

Since $\omega_{2 \to 1}[t]$ is a deterministic function of $(M_2, \omega^{t-1}_{1 \to 2})$, this reduces to

$$\begin{aligned}
n(R_1 - \varepsilon_{1n}) &\leq \sum_{t=1}^{T} I\left(M_1; \omega_{1 \to 2}[t] \mid M_2, \omega^{t-1}_{1 \to 2}\omega^{t-1}_{2 \to 1}\right) + \sum_{i=1}^{n} I(M_1; Y^n|M_2, \omega^T_{1 \to 2}\omega^T_{2 \to 1}) \\
&\leq nC_{12} + \sum_{i=1}^{n} I(M_1; Y^n|M_2, \omega^T_{1 \to 2}\omega^T_{2 \to 1})
\end{aligned} \tag{148}$$

as $\sum_{t=1}^{T} \log |\Omega_{12}^{(t)}| \leq nC_{12}$ for a (C_{12}, C_{21})-permissible code (see (12)). The converse part for the regularized capacity formula follows by identifying V_0, V_1, and V_2 with $\omega_{1\to2}^T \omega_{2\to1}^T$, M_1, and M_2. Then, by [56, Eq. (20)], V_1 and V_2 are statistically independent given V_0.

As for the single-letter outer bound, we now have

$$n(R_1 - \varepsilon_{1n}) \leq nC_{12} + \sum_{i=1}^{n} I\left(M_1; Y[i] \,\big|\, Y^{i-1}, \omega_{1\to2}^T, \omega_{1\to2}^T \omega_{2\to1}^T, M_2\right) \tag{149}$$

Then, define $V_0[i] \equiv (Y^{i-1}, \omega_{1\to2}^T \omega_{2\to1}^T)$, $V_1[i] = (M_1, Y^{i-1}, \omega_{1\to2}^T \omega_{2\to1}^T)$, and $V_2[i] = (M_2, Y^{i-1}, \omega_{1\to2}^T \omega_{2\to1}^T)$, for $i = 1, \ldots, n$.

Furthermore, since Alice 1 performs a measurement that depends on her message m_1 and $\omega_{2\to1}^T$ alone, and similarly for Alice 2, the channel inputs $X_1[i]$ and $X_2[i]$ can also be obtained as the output of a measurement as defined below. Consider a measurement channel

$$\widetilde{\mathcal{F}}_{E_1\to X_1}^{(i,m_1)} \otimes \widetilde{\mathcal{F}}_{E_2\to X_2}^{(i,m_2)} \tag{150}$$

such that

$$\widetilde{\mathcal{F}}_{E_1\to X_1}^{(i,m_1,\omega_{2\to1}^T)}(Q') = \sum_{a_1\in\mathcal{X}_1} \left[\sum_{x_1^n\in\mathcal{X}_1^n \,:\, x_1[i]=a_1} \mathrm{Tr}\left(F_{x_1^n}^{(m_1,\omega_{2\to1}^T)} Q' \right) \right] |a_1\rangle\langle a_1| \tag{151}$$

$$\widetilde{\mathcal{F}}_{E_2\to X_2}^{(i,m_2,\omega_{1\to2}^T)}(Q'') = \sum_{a_2\in\mathcal{X}_2} \left[\sum_{x_2^n\in\mathcal{X}_2^n \,:\, x_2[i]=a_2} \mathrm{Tr}\left(F_{x_2^n}^{(m_2,\omega_{1\to2}^T)} Q'' \right) \right] |a_2\rangle\langle a_2| \tag{152}$$

for every pair of operators Q' and Q'' on $\mathcal{H}_{E_1}$ and $\mathcal{H}_{E_2}$, respectively. Let $\Psi_{E_1E_2} = \sum_{j,\ell} Q_j' \otimes Q_\ell''$ be an arbitrary decomposition of the bipartite state. Hence, following the same steps as in (114), we have

$$\left(\widetilde{\mathcal{F}}_{E_1\to X_1}^{(i,m_1)} \otimes \widetilde{\mathcal{F}}_{E_2\to X_2}^{(i,m_2)}\right)(\Psi_{E_1E_2}) = \sum_{(a_1,a_2)\in\mathcal{X}_1\times\mathcal{X}_2} \left[\sum_{\substack{(x_1^n,x_2^n)\in\mathcal{X}_1^n\times\mathcal{X}_2^n: \\ x_1[i]=a_1 \,,\, x_2[i]=a_2}} \mathrm{Tr}\left((F_{x_1^n}^{(m_1,\omega_{2\to1}^T)} \otimes F_{x_2^n}^{(m_2,\omega_{1\to2}^T)})\Psi_{E_1E_2} \right) \right] |a_1,a_2\rangle\langle a_1,a_2| \tag{153}$$

as expected. We deduce that the channel inputs $X_1[i]$ and $X_2[i]$ can be obtained from a product of measurements of the form $L_1(x_1|i, V_0[i], V_1[i])$ and $L_2(x_2|i, V_0[i], V_2[i])$, as required.

Then, we can rewrite (149) as

$$R_1 - \varepsilon_{1n} \leq \frac{1}{n}\sum_{i=1}^{n} I\left(V_1[i]; Y[i] \,\big|\, V_0[i], V_2[i]\right) + C_{12}$$
$$= I\left(V_1[J]; Y[J] \,\big|\, V_0[J], V_2[J], J|V_0[J], J\right) + C_{12} \tag{154}$$

where the index J is drawn uniformly at random from $\{1, \ldots, n\}$, and it is uncorrelated with the previous systems. Following the same considerations,

$$R_2 - \varepsilon_{2n} \leq I\left(V_2[J]; Y[J] \,\big|\, V_0[J], V_1[J], J\right) + C_{21} \tag{155}$$

and

$$R_1 + R_2 - \varepsilon_{3n} \leq I\left(V_1[J]V_2[J]; Y[J] \,\big|\, V_0[J], J\right) + C_{12} + C_{21}. \tag{156}$$

In addition, by (142),

$$R_1 + R_2 - \varepsilon_{3n} \leq \frac{1}{n}I(m_1m_2; \hat{m}_1\hat{m}_2)$$
$$\leq \frac{1}{n}I(m_1m_2; Y^n)$$
$$= \frac{1}{n}\sum_{i=1}^{n} I(m_1m_2; Y[i] \,|\, Y^{i-1})$$
$$\leq I(V_1[J]V_2[J]J; Y[J]) \tag{157}$$

The proof follows by defining $V_k \equiv (J, V_k[J])$, for $k = 0, 1, 2$, then $X_k \equiv X_k[J]$, and $Y \equiv Y[J]$. $\qquad\square$

X. Summary and Discussion

A. Summary

To summarize, we have considered communication over a two-user classical multiple-access channel (MAC) $P_{Y|X_1,X_2}$ with entanglement resources shared between the transmitters a priori, before communication begins, as illustrated in Figure 1. Our main results are inner and outer bounds, as well as a regularized characterization, for the capacity region of the *general* MAC with entangled transmitters. We have also shown that the examples by Leditzky et al. [11] are a special case and follow from our main result in Theorem 3.

In addition, we observe the following change of behavior. In general, achievable communication rates may also depend on the error criterion. For a classical MAC without entanglement resources, it has long been known that the relaxation to a message-average error criterion can lead to strictly higher achievable rates, when compared with a maximal error criterion [29] (see Remark 3). Here, however, the capacity region with entangled transmitters remains the same, whether we consider a message-average or a maximal error criterion.

The auxiliary systems A_1 and A_2 in our characterization have unbounded dimensions, as mentioned in Remarks 5-6. Although one can always compute an achievable region by simply choosing the dimension of A_1 and A_2, the optimal rates cannot be computed exactly in general. The lack of a dimension bound for the reference system plagues many other quantum models in network information theory, such as the MAC with entangled transmitter and receiver [9], the broadcast channel (see Discussion section in [57]), wiretap channel [58, Remark 5], and squashed entanglement [59, Section 1]. Here, the problem is inherent to the communication scenario. The examples due to Leditzky et al. [11] demonstrate that for some channels, the capacity region can be achieved with a finite entanglement rate (see Example 1), while for other channels, achieving the full capacity region requires an infinite amount of entanglement resources (see Example 2).

B. Unbounded Dimensions

In the resource theory literature, the set of conditional distributions $p_{X_1,X_2|V_1,V_2}$ arising from all choices of finite dimensional Hilbert space $\mathcal{H}_{A_1}$ and $\mathcal{H}_{A_2}$, all measurements $\mathcal{L}_1(v_1) \otimes \mathcal{L}_2(v_2)$, and all $|\phi_{A_1 A_2}\rangle$, is called the family of quantum spatial correlation matrices, and it is denoted by [60]

$$C_{qs}(|\mathcal{X}_1|, |\mathcal{X}_2|).$$

Furthermore, the quantum commuting family, $C_{qc}(|\mathcal{X}_1|, |\mathcal{X}_2|)$, is the set of commuting quantum correlation matrices, i.e., such that the measurement operators for Alice 1 and Alice 2 commute, and the quantum approximate family $C_{qa}(|\mathcal{X}_1|, |\mathcal{X}_2|)$ is the closure of the quantum spatial family, $C_{qa}(|\mathcal{X}_1|, |\mathcal{X}_2|) \equiv \overline{C_{qs}(|\mathcal{X}_1|, |\mathcal{X}_2|)}$. The quantum commuting family $C_{qc}(|\mathcal{X}_1|, |\mathcal{X}_2|)$ is a closed set as well. We note that if the systems were classical, then the three classes would have been identical based on the Fenchel-Eggleston-Carathéodory theorem [33] (see Section VI-A2). However, in the quantum case, the relation between those classes is not so obvious.

There are different conjectures related to the Tsirelson problem [61]; the strong conjecture states that the commuting and spatial families are identical, i.e., $C_{qs} = C_{qc}$; the weak version that the commuting and approximate families are identical, i.e., $C_{qc} = C_{qa}$; and the middle states that the spatial and approximate families are identical, i.e., $C_{qs} = C_{qa}$. The weak Tsirelson conjecure is equivalent to Conne's embedding conjecture [62]. The Tsirelson conjecture is closely related to the study of nonlocal quantum games: $C_{qs} \neq C_{qa}$ is true if and only if there exists a nonlocal game that can be won with certainty using a limit of finite-dimensional entanglement resources, but cannot be won with certainty for a bounded dimension.

Slofstra [63] has shown that indeed there exists a game that can only be won in the limit of finite-dimensional entangelemnt, hence C_{qs} is not closed, and both the strong and middle Tsirelson conjectures are false ($C_{qs} \neq C_{qc}$ and $C_{qs} \neq C_{qa}$). More recently, Ji et al. [64] constructed a game such that the entangled value differs from its commuting-operator value, proving that the weak Tsirelson conjecture is also false.

As the dimension of our entangled systems A_1 and A_2 in the rate formulas (19) and (24) is unbounded, the formulas for $\mathcal{R}_{\mathrm{ET}}(P_{Y|X_1,X_2})$ and $\mathcal{O}_{\mathrm{ET}}(P_{Y|X_1,X_2})$ involve a union over the approximate family C_{qa}. Since $C_{qs} \neq C_{qa}$, it follows that the union cannot necessarily be exhausted by finite-dimensional bipartite states.

C. Open Problems

There are many variations of the MAC setting with entangled and conferencing settings, as illustrated in Table III. In particular, the conferencing transmitters may send either classical or quantum messages to one another, the conferencing protocol may consume pre-shared entanglement, and the encoding procedure may also use pre-shared entanglement, in addition to the conferencing outcome. We refer to each case in Table III below:

- Case 1 is the classical setting, without cooperation resources (see [54, Chap. 14]). The classical setting of Case 3, i.e., classical conferencing without entanglement, was originally studied by Willems [56].
- Case 2 corresponds to our first result in this work, on the MAC with entanglement resources between the transmitters, without conferencing. We have established in Theorem 3 that the capacity region is given by the regularization of the region $\mathcal{R}_{\mathrm{ET}}(P_{Y|X_1,X_2})$.

TABLE III
MAC Settings with Entanglement and Conferencing Transmitters

Case	Conferencing Stage			Encoding Stage	Capacity Formula
	Classical Link	Quantum Link	Entanglement?	Entanglement?	
1	No Conferencing			No	[54, Chap. 14]
2	No Conferencing			Yes	$\mathcal{R}_{\text{ET}}$ (see (19))
3	✓		No	No	[56]
4	✓		No	Yes	$\mathcal{R}_{\text{ET-C}}$ (see (39))
5	✓		Yes	No	Open
6	✓		Yes	Yes	$\mathcal{R}_{\text{ET-C}}$ (see (39))
7		✓	No	No	Open
8		✓	No	Yes	Open
9		✓	Yes	No	Open
10		✓	Yes	Yes	Open

- In the present work, we have introduced the setting in Case 4, whereby the transmitters perform a classical conferencing protocol and then encode using pre-shared entanglement. See Subsection III-C. We have established in Theorem 6 that the capacity region is given by the regularization of the region $\mathcal{R}_{\text{ET-C}}(P_{Y|X_1,X_2}, C_{12}, C_{21})$.
- In Case 5, entanglement is only available at the conferencing stage. In the encoding stage, the transmitters do not use entanglement resources. Nevertheless, the transmitters can perform *teleportation* in the conferencing stage in order to generate entanglement resources to be used in the encoding stage. Although, this could come at the expense of trading information between the transmitters.
- In Case 6, the transmitters may use pre-shared entanglement in the encoding stage as well. As the conferencing link is classical and free of noise, entanglement resources at the conferencing stage do not increase the achievable rates, and the capacity region in Case 6 is the same as in Case 4.
- The capacity region in the quantum conferencing settings remains open. In Subsection III-D, we introduced the setting in Case 10, whereby the transmitters have quantum conferencing links, and they use pre-shared entanglement in both the conferencing and encoding stages. We have shown that the transmitters can then use the entanglement in order to double the conferencing rate using superdense coding (see Theorem 7).
- We conjecture that the capacity region in Case 8 is the same as in Case 4, i.e., if the transmitters have unlimited entanglement resources at the encoding stage, but none at the conferencing stage, then it makes no difference whether the conferecing link is classical or quantum.
- Notice that in Case 7, the encoders do not have access to entanglement resources. Nevertheless, quantum conferencing can be used in order to *generate* entanglement resources. Yet, this could come at the expense of trading information between the transmitters. The same principle holds in Case 9. The optimum in those cases remains open as well.

ACKNOWLEDGMENT

APPENDIX
MAXIMAL ERROR CRITERION

We show here that if a rate pair (R_1, R_2) is achievable with a message-average error criterion, then it is also achievable with a maximal error criterion (see (8), *cf.* (3)). As pointed out in Remark 3, this property does *not* hold for a classical MAC without entanglement resources [29] [30, Sec. 2.2]. Nonetheless, Cai [30] considered the classical MAC without entanglement resources, and showed that the capacity region of the classical MAC with a message-average error criterion is also achievable

with a maximal error criterion, if the encoders are provided with a random key. Since the entanglement resources in our model can also be used in order to generate a random key, the arguments by Cai [30] extend to our model as well. For completeness, we give the details below. It can easily be seen from the calculations below that the new codes only require rate-constrained entanglement resources.

We note that if $(R_1, R_2) = (0,0)$ is the best achievable pair with a message-average error criterion, then the capacity region with a maximal error criterion is $\mathcal{C}_{ET}(P_{Y|X_1,X_2}) = \{(0,0)\}$, and there is nothing to show. Hence, we may assume without loss of generality that $R_1 > 0$ is achievable with entangled transmitters subject to a message-average error criterion. Let $\mathscr{C}$ be a code with entangled transmitters such that the message-average error probability is bounded by

$$\overline{P}_e^{(n)}(\mathscr{C}) \equiv \frac{1}{2^{n(R_1+R_2)}} \sum_{m_1=1}^{2^{nR_1}} \sum_{m_2=1}^{2^{nR_2}} P_e^{(n)}(\mathscr{C}|m_1, m_2) < \gamma \tag{158}$$

for large n, where γ is arbitrarily small. For every $m_2 \in [1 : 2^{nR_2}]$, define the semi-average probability of error by

$$e(m_2) \equiv \frac{1}{2^{nR_1}} \sum_{m_1=1}^{2^{nR_1}} P_e^{(n)}(\mathscr{C}|m_1, m_2). \tag{159}$$

Consider the following subset of "good" messages for User 2,

$$\mathcal{M}_2' = \left\{ m_2 \in [1 : 2^{nR_2}] \ : \ e(m_2) < 2 \cdot \gamma \right\}. \tag{160}$$

Observe that by (158), the average value of $e(m_2)$ is bounded by

$$\frac{1}{2^{nR_2}} \sum_{m_2=1}^{2^{nR_2}} e(m_2) < \gamma. \tag{161}$$

This, in turn, implies that the subset $\mathcal{M}_2'$ is at least as large as half the original set, i.e.,

$$|\mathcal{M}_2'| \geq \frac{1}{2} \cdot 2^{nR_2} \tag{162}$$

$$= 2^{n\left(R_2 - \frac{1}{n}\right)}. \tag{163}$$

Otherwise, the average value of $e(m_2)$ would have been greater than or equal to γ, in contradiction to (161). Therefore, User 2 can throw away the messages outside $\mathcal{M}_2'$, and transmit at a rate of $R_2' = R_2 - \frac{1}{n}$, arbitrarily close to R_2.

We now construct a new code that is reliable under the maximal error criterion. In this construction, Encoder 1 first performs a measurement on an ancilla of dimension n^2 to obtain a uniformly distributed key $L_1 \in [1 : n^2]$. Then, the transmission consists of two consecutive blocks. In the first block, Encoder 1 sends the key L_1 to the receiver using a code based on a *message-average* error criterion. At the same time, Encoder 2 sends $L_2 = 1$. Since the key set size is sub-exponential and we have assumed that User 1 can communicate at a positive rate, the key can be sent using a code of length $\nu = o(n)$. As for the second block, given a key outcome L_1, Encoder 1 chooses a permutation π_{L_1} on the message set $[1 : 2^{nR_1}]$, and encodes by applying the encoding map $\mathcal{F}_1^{(\tilde{m}_1)}$ with $\tilde{m}_1 \equiv \pi_{L_1}(m_1)$. At the same time, Encoder 2 encodes using $\mathcal{F}_2^{(m_2)}$, as in the original code. Bob receives the output sequence $(\bar{Y}^\nu, Y^n)$ of length $\nu + n$, and decodes as follows. He uses the first part, $\bar{Y}^\nu$, in order to find an estimate $\hat{L}_1$ and $\hat{L}_2 = 1$ for the keys. Then, Bob declares his estimation for the messages as $(\hat{m}_1, \hat{m}_2) = (\pi_{\hat{L}_1}^{-1} \times \mathbb{1}_2) g(Y^n)$, using the output Y^n of the second block, where $\mathbb{1}_2$ is the identity permutation on $\mathcal{M}_2'$. We denote the code that is used in the second block by $\pi_L(\mathscr{C}) = (\Psi, \pi\mathcal{F}_1, \mathcal{F}_2, \pi^{-1}g)$.

Having assumed that the key L_1 is uniformly distributed, the probability of decoding the wrong key is the semi-average probability of error for the first block. That is,

$$\Pr\left(\hat{L}_1 \neq L_1\right) = \frac{1}{n^2} \sum_{\ell=1}^{n^2} P_e^{(\nu)}(\mathscr{C}_1|\ell, 1) < 2\gamma. \tag{164}$$

Now, consider the second block. Let $\mathfrak{P}$ denote the permutation group on the message set $[1 : 2^{nR_1}]$ of User 1. Furthermore, let $\Pi_1, \ldots, \Pi_{n^2}$ be an i.i.d. sequence of random permutations, each uniformly distributed over $\mathfrak{P}$. Then, for a given m_1,

$$\Pr(\Pi_\ell(m_1) = \tilde{m}_1) = \frac{(2^{nR_1} - 1)!}{(2^{nR_1})!} = \frac{1}{2^{nR_1}} \tag{165}$$

for all $\tilde{m}_1 \in [1 : 2^{nR_1}]$ and $\ell \in [1 : n^2]$. Thus, for every message pair $(m_1, m_2) \in [1 : 2^{nR_1}] \times \mathcal{M}'_2$,

$$\mathbb{E}\left[P_e^{(n)}(\Pi_\ell(\mathscr{C})|m_1, m_2)\right] = \sum_{\tilde{m}_1} \Pr(\Pi_\ell(m_1) = \tilde{m}_1) \cdot P_e^{(n)}(\mathscr{C}|\tilde{m}_1, m_2) \tag{166}$$

$$= \frac{1}{2^{nR_1}} \sum_{\tilde{m}_1} P_e^{(n)}(\mathscr{C}|\tilde{m}_1, m_2) \tag{167}$$

$$= e(m_2) \tag{168}$$

$$< 2\gamma \tag{169}$$

where the first equality follows from the construction of the code $\pi_\ell(\mathscr{C})$, the second equality is due to (165), the third is due to (159), and the last inequality follows from the definition of $\mathcal{M}'_2$ in (160). Hence, based on the Chernoff bound, we have

$$\Pr\left(\frac{1}{n^2} \sum_{\ell=1}^{n^2} P_e^{(n)}(\Pi_\ell(\mathscr{C})|m_1, m_2) > 7\gamma\right) \leq e^{-\gamma \cdot n^2} \tag{170}$$

(see [30, Lemma 3.1], taking $L \leftarrow n^2$, $\alpha \leftarrow 2\gamma$, $\beta \leftarrow 7\gamma$). This means that the probability that the random code $\Pi_L(\mathscr{C})$ has an error higher than 7γ tends to zero in a super-exponential rate. Thus, by the union bound,

$$\Pr\left(\exists (m_1, m_2) : \frac{1}{n^2} \sum_{\ell=1}^{n^2} P_e^{(n)}(\Pi(\mathscr{C})|m_1, m_2) > 7\gamma\right) \leq 2^{n(R_1 + R'_2)} \cdot e^{-\gamma \cdot n^2} \tag{171}$$

$$\leq e^{-\frac{1}{2}\gamma \cdot n^2} \tag{172}$$

for sufficiently large n. Therefore, there exists a realization $(\pi_1, \ldots, \pi_{n^2})$, such that

$$P_e^{(n)}(\pi_L(\mathscr{C})|m_1, m_2) = \frac{1}{n^2} \sum_{\ell=1}^{n^2} P_e^{(n)}(\pi_\ell(\mathscr{C})|m_1, m_2) \leq 7\gamma \tag{173}$$

for all $(m_1, m_2) \in [1 : 2^{nR_1}] \times \mathcal{M}'_2$. In other words, the maximal error probability for the second block is bounded by 7γ. By choosing $\gamma = \frac{1}{9}\varepsilon_0$, we have that the maximal error probability of the overall code of length $n + o(n)$ is bounded by ε_0, as we wanted to show. By increasing the code length n, the overall transmission rates $\frac{n}{n+o(n)}R_1$ and $\frac{n}{n+o(n)}\left(R_2 - \frac{1}{n}\right)$ can be made arbitrarily close to R_1 and R_2, respectively. $\qquad\square$

REFERENCES

[1] P. van Loock, W. Alt, C. Becher, O. Benson, H. Boche, C. Deppe, J. Eschner, S. Höfling, D. Meschede, and P. Michler, "Extending quantum links: Modules for fiber-and memory-based quantum repeaters," *Adv. Quantum Techno.*, no. 3, p. 1900141, 2020.

[2] U. Pereg, "Communication over quantum channels with parameter estimation," *IEEE Trans. Inf. Theory*, vol. 68, no. 1, pp. 359–383, 2022.

[3] U. Pereg, C. Deppe, and H. Boche, "Communication with unreliable entanglement assistance," *IEEE Trans. Inf. Theory*, 2023.

[4] R. Bassoli, H. Boche, C. Deppe, R. Ferrara, F. H. P. Fitzek, G. Janßen, and S. Saeedinaeeni, "Quantum communication networks," in *Ser. Found. Sign. Proc. Commun. Netw.* Springer, 2021, vol. 23.

[5] A. Winter, "The capacity of the quantum multiple-access channel," *IEEE Trans. Inf. Theory*, vol. 47, no. 7, pp. 3059–3065, Nov 2001.

[6] U. Pereg, C. Deppe, and H. Boche, "The quantum multiple-access channel with cribbing encoders," *IEEE Trans. Inf. Theory*, vol. 68, no. 6, pp. 3965–3988, 2022.

[7] H. Boche and J. Nötzel, "The classical-quantum multiple access channel with conferencing encoders and with common messages," *Quantum Info. Proc.*, vol. 13, no. 12, pp. 2595 –2617, 2014.

[8] S. Diadamo and H. Boche, "The simultaneous identification capacity of the classical -quantum multiple access channel with stochastic encoders for transmission," `arXiv:1903.03395`, 2019.

[9] M. Hsieh, I. Devetak, and A. Winter, "Entanglement-assisted capacity of quantum multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 7, pp. 3078–3090, July 2008.

[10] H. Shi, M. H. Hsieh, S. Guha, Z. Zhang, and Q. Zhuang, "Entanglement-assisted capacity regions and protocol designs for quantum multiple-access channels," *npj Quantum Information*, vol. 7, no. 1, pp. 1 –9, 2021.

[11] F. Leditzky, M. A. Alhejji, J. Levin, and G. Smith, "Playing games with multiple access channels," *Nature Commun.*, vol. 11, no. 1, pp. 1 –5, 2020.

[12] J. Nötzel, "Entanglement-enabled communication," *J. Select Areas Inf. Theory*, vol. 1, no. 2, pp. 401–415, 2020.

[13] J. Nötzel and S. DiAdamo, "Entanglement-enabled communication for the internet of things," in *Proc. Int'l Conf. Comput. Inf. Telecomm. Syst. (CITS'2020)*, 2020.

[14] J. Yun, A. Rai, and J. Bae, "Non-local and quantum advantages in network coding for multiple access channels," `arXiv:2304.10792`, 2023.

[15] A. Seshadri, F. Leditzky, V. Siddhu, and G. Smith, "On the separation of correlation-assisted sum capacities of multiple access channels," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT'2022)*, June 2022, pp. 2756–2761.

[16] G. Brassard, A. Broadbent, and A. Tapp, "Quantum pseudo-telepathy," *Found. Phys.*, vol. 35, no. 11, pp. 1877 –1907, 2005.

[17] B. Doolittle, E. Chitambar, and F. Leditzky, "Nonclassical behaviors of two-sender multiple access channels," in *APS March Meeting Abstracts 2022, abstract id.G00.182*, vol. 2022, 2022.

[18] O. Fawzi and P. Fermé, "Beating the sum-rate capacity of the binary adder channel with non-signaling correlations," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT'2022)*, 2022, pp. 2750–2755.

[19] Y. Quek and P. W. Shor, "Quantum and superquantum enhancements to two-sender, two-receiver channels," *Phys. Rev. A*, vol. 95, no. 5, p. 052329, 2017.

[20] U. Pereg, C. Deppe, and H. Boche, "Quantum broadcast channels with cooperating decoders: An information-theoretic perspective on quantum repeaters," *J. Math. Phys.*, vol. 62, no. 6, p. 062204, 2021.

[21] G. P. Fettweis and H. Boche, "6G: The personal tactile internet - and open questions for information theory," *IEEE BITS Info. Th. Magazine*, 2021.

[22] ——, "6g: The personal tactile internet—and open questions for information theory," *IEEE BITS the Information Theory Magazine*, vol. 1, no. 1, pp. 71–82, 2021.

[23] S. Dang, O. Amin, and M. S. Shihada, B.and Alouini, "What should 6G be?" *Nature Electronics*, vol. 3, no. 1, pp. 20 –29, 2020.

[24] G. P. Fettweis and H. Boche, "On 6G and trustworthiness," *Communications of the ACM*, vol. 65, no. 4, pp. 48 –49, 2022.

[25] H. B. M. Wiese, "The arbitrarily varying multiple-access channel with conferencing encoders," *IEEE Transactions on Information Theory*, vol. 59, no. 3, pp. 1405–1416, Mar 2013.

[26] R. S. H. Boche, "Arbitrarily varying multiple access channels with conferencing encoders: List decoding and finite coordination resources," *Advances in Mathematics of Communications*, vol. 10, no. 2, pp. 333–354, 2016.

[27] F. Tariq, M. R. A. Khandaker, K. K. Wong, M. A. Imran, M. Bennis, and M. Debbah, "A speculative study on 6G," *IEEE Wireless Comm.*, vol. 27, no. 4, pp. 118–125, 2020.

[28] F. Fitzek and H. Boche, "6G-life: Digital transformation and sovereignty of future communication networks," *IEEE Network*, pp. 3–4, Nov/Dec 2021.

[29] G. Dueck, "Maximal error capacity regions are smaller than average error capacity regions for multi-user channels," *Prob. Inf. Theo.*, vol. 7, no. 1, pp. 11–19, 1978.

[30] N. Cai, "The maximum error probability criterion, random encoder, and feedback, in multiple input channels," *Entropy*, vol. 16, no. 3, pp. 1211–1242, 2014.

[31] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Wiley, 2006.

[32] M. M. Wilde, *Quantum Information Theory*, 2nd ed. Cambridge University Press, 2017.

[33] H. G. Eggleston, "Convexity," *J. London Math. Society*, vol. 1, no. 1, pp. 183 –186, 1966.

[34] A. A. Gohari and V. Anantharam, "Evaluation of marton's inner bound for the general broadcast channel," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 608–619, 2012.

[35] M. R. Grace and S. Guha, "Perturbation theory for quantum information," in *Proc. IEEE Inf. Theory Workshop (ITW'2022)*, 2022, pp. 500–505.

[36] W. Slofstra, "Tsirelson's problem and an embedding theorem for groups arising from non-local games," *J. American Math. Society*, vol. 33, no. 1, pp. 1 –56, 2020.

[37] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal, "Entanglement-assisted classical capacity of noisy quantum channels," *Phys. Rev. Lett.*, vol. 83, no. 15, p. 3081, Oct 1999.

[38] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal, "Entanglement-assisted capacity of a quantum channel and the reverse shannon theorem," *IEEE Trans. Inf. Theory*, vol. 48, no. 10, pp. 2637–2655, Oct 2002.

[39] U. Pereg, C. Deppe, and H. Boche, "Quantum channel state masking," *IEEE Trans. Inf. Theory*, vol. 67, no. 4, pp. 2245–2268, 2021.

[40] W. Slofstra and T. Vidick, "Entanglement in non-local games and the hyperlinear profile of groups," in *Ann. Henri Poincaré*, vol. 19, no. 10, 2018, pp. 2979 –3005.

[41] M. A. Alhejji and G. Smith, "A tight uniform continuity bound for equivocation," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT'2020)*, 2020, pp. 2270–2274.

[42] A. Shamir, "Ip=pspace," in *31st Annual Symposium on Foundations of Computer Science, St. Louis, Missouri, USA, October 22-24, 1990, Volume I.* IEEE Computer Society, 1990, pp. 11–15.

[43] L. Babai, L. Fortnow, and C. Lund, "Addendum to non-deterministic exponential time has two-prover interactive protocols,"

Comput. Complex., vol. 2, p. 374, 1992.

[44] S. Arora and S. Safra, "Probabilistic checking of proofs: A new characterization of NP," *J. ACM*, vol. 45, no. 1, pp. 70–122, 1998.

[45] I. Dinur, "The PCP theorem by gap amplification," *J. ACM*, vol. 54, no. 3, p. 12, 2007.

[46] S. Cook, "The p versus np problem," *The millennium prize problems*, 02 2001.

[47] M. Sudan, "Probabilistically checkable proofs," *Commun. ACM*, vol. 52, no. 3, pp. 76–84, 2009.

[48] L. Fortnow, "The status of the P versus NP problem," *Commun. ACM*, vol. 52, no. 9, pp. 78–86, 2009.

[49] O. Goldreich, S. Micali, and A. Wigderson, "Proofs that yield nothing but their validity and a methodology of cryptographic protocol design," in *Providing Sound Foundations for Cryptography: On the Work of Shafi Goldwasser and Silvio Micali*, O. Goldreich, Ed. ACM, 2019, pp. 285–306.

[50] R. Ostrovsky and A. Wigderson, "One-way fuctions are essential for non-trivial zero-knowledge," in *Second Israel Symposium on Theory of Computing Systems, ISTCS 1993, Natanya, Israel, June 7-9, 1993, Proceedings*. IEEE Computer Society, 1993, pp. 3–17.

[51] J. Cai, A. Condon, and R. J. Lipton, "PSPACE is provable by two provers in one round," in *Proceedings of the Sixth Annual Structure in Complexity Theory Conference, Chicago, Illinois, USA, June 30 - July 3, 1991*. IEEE Computer Society, 1991, pp. 110–115.

[52] M. Ben-Or, S. Goldwasser, J. Kilian, and A. Wigderson, "Multi-prover interactive proofs: How to remove intractability assumptions," in *Proceedings of the 20th Annual ACM Symposium on Theory of Computing, May 2-4, 1988, Chicago, Illinois, USA*, J. Simon, Ed. ACM, 1988, pp. 113–131.

[53] A. El Gamal and Y. Kim, *Network Information Theory*. Cambridge University Press, 2011.

[54] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, 2nd ed. Cambridge University Press, 2011.

[55] G. Kramer, "Topics in multi-user information theory," *Foundations and Trends in Communications and Information Theory*, vol. 4, no. 4 -5, pp. 265 –444, 2008.

[56] F. Willems, "The discrete memoryless multiple access channel with partially cooperating encoders (corresp.)," *IEEE Trans. Inf. Theory*, vol. 29, no. 3, pp. 441–445, 1983.

[57] F. Dupuis, P. Hayden, and K. Li, "A father protocol for quantum broadcast channels," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2946–2956, June 2010.

[58] H. Qi, K. Sharma, and M. M. Wilde, "Entanglement-assisted private communication over quantum broadcast channels," *J. Phys. A: Math. and Theo.*, vol. 51, no. 37, p. 374001, 2018.

[59] K. Li and A. Winter, "Relative entropy and squashed entanglement," *Commun. Math. Phys.*, vol. 326, no. 1, pp. 63 –80, Jan 2014.

[60] K. J. Dykema and V. Paulsen, "Synchronous correlation matrices and connes' embedding conjecture," *J. Math. Phys.*, vol. 57, no. 1, p. 015214, 2016.

[61] B. S. Tsirelson, "Some results and problems on quantum bell-type inequalities," *Hadronic J. Supp.*, vol. 8, no. 4, pp. 329 –345, 1993.

[62] N. Ozawa, "About the connes embedding conjecture: algebraic approaches," *Japanese J. Math.*, vol. 8, no. 1, pp. 147 –183, 2013.

[63] W. Slofstra, "The set of quantum correlations is not closed," in *Forum of Mathematics, Pi*, vol. 7. Cambridge University Press, 2019, p. e1.

[64] Z. Ji, A. Natarajan, T. Vidick, J. Wright, and H. Yuen, "MIP*= RE," *Commun. ACM*, vol. 64, no. 11, pp. 131 –138, 2021.