# The Capacity Region of $\mathfrak{p}$-Transmitter/qu-Receiver Multiple-Access Channels with Common Information 

A. Haghi Student Member, IEEE, R. Khosravi-Farsani Student Member, IEEE, M. R. Aref, and F. Marvasti Senior Member, IEEE


#### Abstract

This paper investigates the capacity problem for some multiple-access scenarios with cooperative transmitters. First, a general Multiple-Access Channel (MAC) with common information, i.e., a scenario where $\boldsymbol{p}$ transmitters send private messages and also a common message to $q$ receivers and each receiver decodes all of the messages, is considered. The capacity region of the discrete memoryless channel is characterized. Then, the general Gaussian fading MAC with common information wherein partial Channel State Information (CSI) is available at the transmitters (CSIT) and perfect CSI is available at the receivers (CSIR) is investigated. A coding theorem is proved for this model that yields an exact characterization of the throughput capacity region. Finally, a two-transmitter/one-receiver Gaussian fading MAC with conferencing encoders with partial CSIT and perfect CSIR is studied and its capacity region is determined. For the Gaussian fading models with CSIR only (transmitters have no access to CSIT), some numerical examples and simulation results are provided for Rayleigh fading.


Index Terms- Multiple access channel (MAC); cooperative encoders; conferencing; Gaussian fading channels; channel state information (CSI).

## I. INTRODUCTION

The classical Multiple Access Channel (MAC) first considered by Shannon [1] models a communication system in which separated transmitters send independent messages to a receiver. The capacity region of the classical MAC was derived by Ahlswede [2] and Liao [3]. The MAC with cooperative encoders was first investigated by Slepian and Wolf [4], (see also [5]) wherein the capacity region of a discrete MAC with common information was established. In a MAC with common information, the transmitters send a common message (in addition to their respective private messages) to the receiver cooperatively. The concept of cooperation between transmitters in a multiple-access scenario and its effect on the capacity region has been widely studied. For a comprehensive review of the existing results for the two-transmitter/one-receive MAC with cooperative encoders, see [6] and the literature therein. In another model, the transmitters of a MAC, where a private message is associated to each one, cooperate with each other via conferencing that yields a larger capacity region. This model sometimes is referred to as MAC with partially cooperating encoders. The notion of conferencing between transmitters of MAC was first introduced by Willems [7] wherein the capacity region of a discrete memoryless two-transmitter/one-receiver MAC with conferencing encoders was established. In this scenario, encoders are connected by communication links of finite capacity, which permit both encoders to communicate over noise-free bit-pipes of given capacity. The capacity region of the Gaussian two-transmitter/one-receiver MAC with conferencing encoders was characterized in [8].

[^0]Some parts of this paper were presented at the IEEE International Symposium on Information Theory (ISIT' 10 ), Austin, Texas, U.S.A, Jun. 2010, [21].

Cooperation between various nodes of a communication system via common message and conferencing was also studied for other models. Specifically, in [9] a three-transmitter/one-receiver Gaussian MAC with encoders, partially cooperating over a ring of finite-capacity unidirectional links, was studied. In [10], the capacity region of the compound MAC (a two-transmitter/tworeceiver MAC) with a common message, also with conferencing encoders was derived. The authors in [10] also studied the interference channel with common information and also with conferencing encoders; the capacity region was established for some special cases. A two-user broadcast channel where the decoders cooperate via conferencing links was investigated in [11]. The compound MAC with a common message and conferencing decoders and also the compound MAC when both of the encoders and decoders cooperate via conferencing links was considered in [12]; the capacity region was established for some cases, specially, the physically degraded channel.

Another important feature in communication systems, specifically in mobile wireless communications, is the fading environment. The classical Gaussian fading MAC with Additive White Gaussian Noise (AWGN) was investigated in [13]-[15]. The authors in [13] and [14] characterized the capacity region of the Gaussian fading MAC where only the receiver can track the time-varying channel, i.e., the receiver has access to the Channel State Information (CSI) perfectly while the transmitters have no access to CSI. In [15] the Gaussian fading MAC with perfect CSI at the transmitters and the receiver was studied and the throughput capacities, the delay limited capacities, and the associated optimal resource allocation were obtained.

In this paper, we aim at studying the above features, i.e., cooperation among transmitters and fading environment, in a MAC, simultaneously. To this end, we consider three multiple-access scenarios and provide an exact characterization of the capacity region for each one. First, we consider a $p$-transmitter $/ q$-reciever MAC coined as General MAC (GMAC) with a common message. In this network, each transmitter sends a private message over the channel as well as all transmitters send a common message cooperatively to the receivers. Each receiver decodes all transmitted messages. The capacity region of the discrete GMAC without common message was previously obtained in [16]. Here, we extend the result of [4] to the case of $\mathcal{p}$ transmitters and $\mathcal{q}$, receivers and characterize the capacity region of the discrete GMAC with a common message using the superposition coding technique. Then, we will consider the Gaussian fading GMAC with a common message. We show that the encoding and decoding schemes used to derive the capacity region of the discrete GMAC with a common message can constitute a frame to describe the signaling procedure for the Gaussian fading channel. In fading networks, based on the availability of CSI at different nodes (transmitters and receivers), the capacity region differs from one model to another. In a wide range of models for wireless communications, it is common to measure the state of the channel at the destinations perfectly, and then feed it to the transmitters through perfect or noisy feedback links. In our model for the Gaussian fading GMAC with a common message, we consider the channel with partial CSI at the transmitters (CSIT) and perfect CSI at the receivers (CSIR). Here, partial CSIT is modeled as the knowledge of a deterministic (potentially discrete-valued) function of CSI, where this deterministic function may vary from one transmitter to another. This model, which had been previously studied for the classical Gaussian fading MAC in [17], contains important benefits; it considers the general CSIT with full control over how much CSI is available. Specifically, it unifies the characterization of the two models in which the transmitters have perfect or no knowledge of CSI [14], [15]. Furthermore, since each transmitter is associated to its own deterministic function, the model also includes the scenario in which not all transmitters have access to CSI, which is more realistic. Under such a scenario, we prove a coding theorem for the Gaussian fading GMAC with a common message yielding an exact characterization of its throughput capacity region. It should be mentioned that the authors in [18] also studied the Gaussian fading MAC with a common message. The model studied in [18] is a special case of our derivation; it only considers a two-transmitter/one-receiver MAC $(p=2$, and $q=1)$ with perfect CSI at both transmitters. Also, the techniques used in [18] to derive the capacity region are different from ours, yielding different formulation of the capacity region. However, we show that the result of [18] and ours in the respective special case are indeed equivalent. This yields a new proof for the result of [18].

Finally, as a third scenario, we study a two-transmitter/one-receiver Gaussian fading MAC with conferencing encoders with partial CSIT (the same model as before) and perfect CSIR, wherein it is assumed that the transmitters have access to CSIT after the conferencing. We determine the throughput capacity region of this channel.

The rest of the paper is organized as follows: In Section II, preliminaries and channel models are introduced. The main results are presented in Section III; in Subsection III-A, the GMAC with a common message and in Subsection III-B, the Gaussian fading MAC with conferencing encoders are investigated. In Section IV, numerical examples and simulation results are provided for the Gaussian Rayleigh fading. Finally, the paper is concluded in Section V.

## II. Preliminaries and definitions

In this paper, we use the following notations: random variables (r.v.) are denoted by upper case letters (e.g., $X$ ) and lower case letters are used to show their realizations (e.g., $x$ ). The probability distribution function (p.d.f.) of a r.v. $X$ with alphabet set $\mathcal{X}$, is denoted by $P_{X}(x)$ where $x \in \mathcal{X} ; P_{X \mid Y}(x \mid y)$ denotes the conditional p.d.f. of $X$ given $Y$. A sequence of r.v.'s $\left\langle X_{t}, \ldots, X_{n}\right\rangle$ with the same alphabet set $\mathcal{X}$ is denoted by $X_{t}^{n}$ and its realization is denoted by $\left\langle x_{t}, \ldots, x_{n}\right\rangle, x_{i} \in \mathcal{X}, i=t, \ldots, n$, where for the case of $X_{1}^{n}$ the subscript is dropped, occasionally. The set of all $\epsilon$-typical $n$-sequences $x^{n}$ with respect to the p.d.f. $P_{X}(x)$, as defined in [19, Ch. 3], is denoted by $A_{\epsilon}^{n}\left(P_{X}\right)$. The notation $\mathbb{E}[$.] indicates the expectation operator, where sometimes to be more precise we use $\mathbb{E}_{X}[$.$] to denote expectation with respect to the distribution of the r.v. X$. We also use $\boldsymbol{h}($.$) to represent the differential entropy.$ The set of real numbers and also the set of all nonnegative real numbers is denoted by $\mathbb{R}$ and $\mathbb{R}_{+}$, respectively. Finally, $\mathbb{C}(x) \equiv$ $\frac{1}{2} \log (1+x)$.


Figure 1. The GMAC with a common message.

## A) General Multiple Access Channel

A discrete memoryless $\mathfrak{p}$-transmitter/ $q$-receiver GMAC $\left\{x_{1}, x_{2}, \ldots, x_{p}, \mathbb{P}\left(y_{1}, \ldots, y_{q} \mid x_{1}, \ldots, x_{p}\right), y_{1}, y_{2}, \ldots, y_{q}\right\}$, is a channel with finite input alphabets $X_{i}, i=1, \ldots, p$, finite output alphabets $\mathcal{Y}_{j}, j=1, \ldots, q$, and a conditional p.d.f. $\mathbb{P}\left(y_{1}, \ldots, y_{q} \mid x_{1}, \ldots, x_{\mathfrak{p}}\right)$ that describes the relation between the inputs and outputs of the channel. The channel is assumed to be memoryless, i.e.

$$
\begin{equation*}
\operatorname{Pr}\left(y_{1}^{n}, \ldots, y_{q}^{n} \mid x_{1}^{n}, \ldots, x_{p}^{n}\right)=\prod_{t=1}^{n} \mathbb{P}\left(y_{1, t}, \ldots, y_{q, t} \mid x_{1, t}, \ldots, x_{p}, t\right), \quad n \geq 1 \tag{1}
\end{equation*}
$$

where $\operatorname{Pr}(A \mid B)$ is the probability of the event $A$ given $B$. Fig. 1 illustrates the channel model. In this paper, we consider a communication setting in which each transmitter sends a private message (unknown to other transmitters) over the channel, as well as all transmitters send a common message to the receivers cooperatively. Both of the common massage and the private messages are decoded at each of the receivers.
For the GMAC described by (1) with a common message, a length- $n$ code $\mathfrak{C}_{n}$ with a common message set $\mathcal{W}_{0}=\left\{1, \ldots, 2^{n R_{0}}\right\}$ and private message sets $\mathcal{W}_{i}=\left\{1, \ldots, 2^{n R_{i}}\right\}, i=1, \ldots, p$, is defined by a set of encoder functions $\left\{\mathfrak{E}_{i}(.)\right\}_{i=1}^{p}$ as:

$$
\mathfrak{E}_{i}: \mathcal{W}_{0} \times \mathcal{W}_{i} \rightarrow X_{i}^{n}, \quad i=1, \ldots, p
$$

and a set of decoder functions $\left\{\mathfrak{D}_{j}(.)\right\}_{j=1}^{q}$ as:

$$
\mathfrak{D}_{j}: \mathcal{Y}_{j}^{n} \rightarrow \mathcal{W}_{0} \times \mathcal{W}_{1} \times \ldots \times \mathcal{W}_{p}
$$

The rate of the code is the $(p+1)$-tuple $\left(R_{0}, R_{1}, \ldots, R_{\mathcal{p}}\right)$. The average error probability of decoding at the $j^{\text {th }}$ receiver $P_{e, j}^{\mathbb{C}_{n}}$, $j=1, \ldots, q$, is given by:

$$
\begin{equation*}
P_{e, j}^{\mathbb{C}_{n}}=\frac{1}{2^{n\left(\sum_{i=0}^{p} R_{i}\right)}} \sum_{\bar{w}_{0}, \bar{w}_{1}, \ldots, \bar{w}_{\mathfrak{p}}} \operatorname{Pr}\left(\left\{\mathfrak{D}_{j}\left(Y_{j}^{n}\right) \neq\left(\bar{w}_{0}, \bar{w}_{1}, \ldots, \bar{w}_{\mathfrak{p}}\right)\right\} \mid\left(\bar{w}_{0}, \bar{w}_{1}, \ldots, \bar{w}_{\mathfrak{p}}\right) \text { is sent }\right) \tag{2}
\end{equation*}
$$

and the total average error probability of the code $P_{e}^{\mathbb{®}_{n}}$ is given by:

$$
\begin{equation*}
P_{e}^{\mathbb{C}_{n}}=\frac{1}{2^{n\left(\sum_{i=0}^{p} R_{i}\right)}} \sum_{\bar{w}_{0}, \bar{w}_{1}, \ldots, \bar{w}_{\mathfrak{p}}} \operatorname{Pr}\left(\left\{\bigcup_{j=1}^{n}\left\{\mathfrak{D}_{j}\left(Y_{j}^{n}\right) \neq\left(\bar{w}_{0}, \bar{w}_{1}, \ldots, \bar{w}_{\mathfrak{p}}\right)\right\}\right\} \mid\left(\bar{w}_{0}, \bar{w}_{1}, \ldots, \bar{w}_{\mathfrak{p}}\right) \text { is sent }\right) \tag{3}
\end{equation*}
$$

Note that $P_{e}^{\mathbb{C}_{n}} \leq \sum_{j=1}^{q} P_{e, j}^{\mathbb{C}_{n}}$. The $(\mathcal{p}+1)$-tuple $\left(R_{0}, R_{1}, \ldots, R_{\mathfrak{p}}\right)$ of nonnegative real numbers is said to be achievable if for every $\epsilon>0$ and for all sufficiently large $n$ there exists a length- $n$ code $\mathfrak{C}_{n}$ such that $P_{e}^{\mathbb{C}_{n}}<\epsilon$. The closure of the set of all achievable rates is the capacity region.

In addition to the previous communication setting in which the input and output alphabets of the channel are of finite size, we also investigate the Gaussian fading GMAC with a common message. The channel is formulated as:

$$
\begin{equation*}
Y_{j, t}=\sum_{i=1}^{p} S_{j i, t} X_{i, t}+Z_{j, t}, \quad t \geq 1, \quad j=1, \ldots, q \tag{4}
\end{equation*}
$$

where $\left\{Y_{j, t}\right\}_{t=1}^{\infty},\left\{X_{i, t}\right\}_{t=1}^{\infty}$ are the $\mathbb{R}$-valued received signal at the $j^{t h}$ receiver and the $\mathbb{R}$-valued transmitted signal of the $i^{\text {th }}$ transmitter, respectively; $\left\{S_{j i, t}\right\}_{t=1}^{\infty}$ is the $\mathbb{R}_{+}$-valued fading coefficient experienced by the signal of the $i^{\text {th }}$ transmitter at the $j^{\text {th }}$ receiver, and $\left\{Z_{j, t}\right\}_{t=1}^{\infty}$ is independent identically distributed (i.i.d.) Additive White Gaussian Noise (AWGN) process at the $j^{t h}$ channel with zero mean and variance $\sigma_{j}^{2}$. The fading processes $\left\{S_{j i, t}\right\}_{t=1}^{\infty}, i=1, \ldots, p, j=1, \ldots, q$, are assumed to be jointly stationary and ergodic but not necessarily independent of each other. For the system given in (4) we define the state process of the channel as the following $q \times p$ matrix:

$$
S_{t}=\left[\begin{array}{ccc}
S_{11, t} & \cdots & S_{1 p, t}  \tag{5}\\
\vdots & \ddots & \vdots \\
S_{q 1, t} & \cdots & S_{q p, t}
\end{array}\right]_{q, \times p}, \quad t \geq 1
$$

Furthermore, we assume that the state process of the channel $\left\{S_{t}\right\}_{t=1}^{\infty}$ is independent of the AWGN process $\left\{Z_{j, t}\right\}_{t=1}^{\infty}, j=1, \ldots, q$.
Partial CSIT and perfect CSIR: In this paper, we consider the channel model in (4) with partial CSIT; the $i^{\text {th }}$ transmitter at each time instant knows a version of the current CSI determined by a deterministic function of the state process of the channel. In other words, the CSI available at the $i^{t h}$ transmitter is given by $\xi_{i}:\left[\mathbb{R}_{+}\right]_{q \times p} \rightarrow \varepsilon_{i}, i=1, \ldots, p$, where $\left[\mathbb{R}_{+}\right]_{q \times p}$ is the space of all $q \times p$ matrices with $\mathbb{R}_{+}$-valued entries and $\mathcal{E}_{i}$ is an arbitrary set (maybe finite) associated to the $i^{\text {th }}$ transmitter. Thus, for the channel in (4) the $i^{\text {th }}$ encoder consists of a set of functions $\left\{\mathfrak{E}_{i, t}\right\}_{t=1}^{n}$ where:

$$
\mathfrak{E}_{i, t}: \mathcal{W}_{0} \times \mathcal{W}_{i} \times \varepsilon_{i} \rightarrow \mathbb{R}, \quad t=1, \ldots, n, \quad i=1, \ldots, p
$$

and the input symbol of the channel due to the $i^{\text {th }}$ transmitter at each time instant $t$, is given by $X_{i, t}=\mathfrak{F}_{i, t}\left(W_{0}, W_{i}, E_{i, t}\right)$, where $E_{i, t}=\xi_{i}\left(S_{t}\right) \in \mathcal{E}_{i}$. Furthermore, we assume that the receivers have access to perfect CSI, i.e., the $j^{\text {th }}$ decoder function $j=$ $1, \ldots, q$, is given by:

$$
\mathfrak{D}_{j}: \mathcal{Y}_{j}^{n} \times\left(\left[\mathbb{R}_{+}\right]_{q \times p}\right)^{n} \rightarrow \mathcal{W}_{0} \times \mathcal{W}_{1} \times \ldots \times \mathcal{W}_{p}
$$

which estimates the messages as:

$$
\left(\widehat{W}_{0}{ }^{(j)}, \widehat{W}_{1}^{(j)}, \ldots, \widehat{W}_{p}^{(j)}\right)=\mathfrak{D}_{j}\left(Y_{j}^{n}, S^{n}\right)
$$

where $S^{n}=\left(S_{1}, S_{2}, \ldots, S_{n}\right)$ is the state process of the channel.
An average power constraint is imposed on the codewords of each encoder. Thus, the $i^{t h}$ transmitter is subjected to an average power constraint $P_{i} \in \mathbb{R}_{+}$, in the following way:

$$
\begin{equation*}
\frac{1}{n} \mathbb{E}\left[\sum_{t=1}^{n}\left(X_{i, t}\left(W_{0}, W_{i}, \xi_{i}\left(S_{t}\right)\right)\right)^{2}\right] \leq P_{i}, \quad i=1, \ldots, p \tag{6}
\end{equation*}
$$

## B) Multiple Access Channels with Conferencing Encoders



Figure 2. The MAC with conferencing encoders.
A discrete two-transmitter/one-receiver MAC with conferencing encoders $\left\{X_{1}, X_{2}, \mathbb{P}\left(y \mid x_{1}, x_{2}\right), \mathcal{y}, C_{12}, C_{21}\right\}$, is a channel with two (finite) input alphabet sets $\mathcal{X}_{1}, X_{2}$, an (finite) output alphabet set and a conditional probability distribution $\mathbb{P}\left(y \mid x_{1}, x_{2}\right)$ which characterizes the channel. In addition, encoders are connected to each other with two links of capacities $C_{12}, C_{21}$. Fig. 2 depicts the channel model.

A length- $n$ code $\mathfrak{C}\left(n, K, C_{12}, C_{21}\right)$ for the MAC with conferencing encoders has two message sets $\mathcal{W}_{i}=\left\{1, \ldots, 2^{n R_{i}}\right\}, i=1,2$ : before transmitting of messages over the channel, the two encoders hold a conference, i.e., the code $\mathfrak{C}\left(n, K, C_{12}, C_{21}\right)$ consists of two sets of $K$ communicating functions $\left\{k_{i, 1}, \ldots, k_{i, K}\right\}, i=1,2$, and two sets of (finite) conferencing alphabets $\left\{\mathcal{V}_{i, 1}, \ldots, \nu_{i, K}\right\}, i=$ 1,2. Each communicating function $k_{i, k}, i=1,2$, for $k=1, \ldots, K$, maps the message $W_{i}$ and the sequence of the previously received symbols from the other transmitter into the $k^{\text {th }}$ symbol $V_{i, k} \in \mathcal{V}_{i, k}$. In notations we have:

$$
\begin{array}{ll}
k_{1, k}: \mathcal{W}_{1} \times \mathcal{V}_{2}^{k-1} \rightarrow \mathcal{V}_{1, k}, & V_{1, k}=k_{1, k}\left(W_{1}, V_{2}^{k-1}\right) \\
k_{2, k}: \mathcal{W}_{2} \times \mathcal{V}_{1}^{k-1} \rightarrow \mathcal{V}_{2, k}, & V_{2, k}=k_{2, k}\left(W_{2}, V_{1}^{k-1}\right) \tag{7}
\end{array}
$$

A conference is said to be ( $C_{12}, C_{21}$ )-permissible [7] if the sets of communicating functions are such that:

$$
\begin{equation*}
\sum_{k=1}^{K} \log \left\|\nu_{1, k}\right\| \leq n C_{12}, \quad \sum_{k=1}^{K} \log \left\|\mathcal{V}_{2, k}\right\| \leq n C_{21} \tag{8}
\end{equation*}
$$

where $\left\|\nu_{i, k}\right\|$ denotes the cardinality of $\nu_{i, k}$.

After the conferencing, transmitter 1 knows the sequence $V_{2}^{K}=\left(V_{2,1}, \ldots, V_{2, K}\right)$ and transmitter 2 knows the sequence $V_{1}^{K}=$ $\left(V_{1,1}, \ldots, V_{1, K}\right)$. Furthermore, the code $\mathfrak{C}\left(n, K, C_{12}, C_{21}\right)$ has two encoder functions $\left\{\mathfrak{E}_{1}, \mathfrak{E}_{2}\right\}$ described as:

$$
\mathfrak{E}_{1}: \mathcal{W}_{1} \times \mathcal{V}_{2}^{K} \rightarrow X_{1}^{n}, \quad \mathfrak{E}_{2}: \mathcal{W}_{2} \times \mathcal{V}_{1}^{K} \rightarrow X_{2}^{n}
$$

also the decoder is given by:

$$
\mathfrak{D}: \mathcal{Y}^{n} \rightarrow \mathcal{W}_{1} \times \mathcal{W}_{2}
$$

The rate of the code $\mathfrak{C}\left(n, K, C_{12}, C_{21}\right)$ is the pair $\left(R_{1}, R_{2}\right)$ and the error probability of the code $P_{e}^{\mathbb{(}\left(n, K, C_{12}, C_{21}\right)}$ is defined as:

$$
\begin{equation*}
P_{e}^{\mathbb{C}\left(n, K, C_{12}, C_{21}\right)}=\frac{1}{2^{n\left(R_{1}+R_{2}\right)}} \sum_{\bar{w}_{1}, \bar{w}_{2}} \operatorname{Pr}\left(\left\{\mathcal{D}\left(Y^{n}\right) \neq\left(\bar{w}_{1}, \bar{w}_{2}\right)\right\} \mid\left(\bar{w}_{1}, \bar{w}_{2}\right) \text { is sent }\right) . \tag{9}
\end{equation*}
$$

A rate pair $\left(R_{1}, R_{2}\right)$ of nonnegative numbers is said to be achievable for the MAC with conferencing encoders if for every $\epsilon>0$ and for all sufficiently large $n$, there exists a code $\mathfrak{C}\left(n, K, C_{12}, C_{21}\right)$ such that the resulting conference of the code is $\left(C_{12}, C_{21}\right)$ permissible and $P_{e}^{\llbracket\left(n, K, C_{12}, C_{21}\right)}<\epsilon$. The capacity region of the MAC with conferencing encoders is the closure of all achievable rate pairs.

The MAC with conferencing encoders was first introduced in [7] where the capacity region of the discrete channel model was derived. In this paper, we investigate the Gaussian fading MAC with conferencing encoders in which the channel is formulated as:

$$
\begin{equation*}
Y_{t}=S_{1, t} X_{1, t}+S_{2, t} X_{2, t}+Z_{t}, \quad t \geq 1 \tag{10}
\end{equation*}
$$

where $\left\{Y_{t}\right\}_{t=1}^{\infty},\left\{X_{i, t}\right\}_{t=1}^{\infty}, i=1,2$, are the $\mathbb{R}$-valued received signal and the $\mathbb{R}$-valued transmitted signal of the $i^{t h}$ transmitter, respectively; $\left\{S_{i, t}\right\}_{t=1}^{\infty}, i=1,2$, is the $\mathbb{R}_{+}$-valued fading coefficient experienced by the signal of the $i^{\text {th }}$ transmitter, and $\left\{Z_{t}\right\}_{t=1}^{\infty}$ is i.i.d. AWGN process with zero mean and variance $\sigma^{2}$. The fading processes $\left\{S_{i, t}\right\}_{t=1}^{\infty}, i=1,2$, are assumed to be jointly stationary and ergodic but not necessarily independent of each other. For the channel given by (10) we define the state process of the channel as $S_{t}=\left(S_{1, t}, S_{2, t}\right), t \geq 1$. The definition of conference for the channel in (10) is the same as discrete channel and is given by (7) and (8). Similar to the communication system in (4), we consider the Gaussian fading MAC with conferencing encoders with partial CSIT and perfect CSIR, where CSIT is known to the encoders after the conferencing. At each time instant, CSIT at the $i^{\text {th }}$ transmitter, $i=1,2$, is a deterministic function of the current state of channel. Let $\xi_{i}: \mathbb{R}_{+}^{2} \rightarrow \varepsilon_{i}$ be the mapping that describes CSIT at the $i^{\text {th }}$ transmitter, where $\mathcal{E}_{i}$ is an arbitrary set (maybe finite) associated to the $i^{\text {th }}$ transmitter, $i=1,2$. Assume that the information that transmitter 1 obtains after the conferencing is $V_{2}^{K}=\left(V_{2,1}, \ldots, V_{2, K}\right)$ and for transmitter 2 is $V_{1}^{K}=$ $\left(V_{1,1}, \ldots, V_{1, K}\right)$. By these assumptions the $i^{t h}$ encoder consists of a set of functions $\left\{\mathfrak{E}_{i, t}\right\}_{t=1}^{n}, i=1,2$, where for $t=1, \ldots, n$ :

$$
\left\{\begin{array}{l}
\mathfrak{E}_{1, t}: \mathcal{W}_{1} \times V_{2}^{K} \times \mathcal{E}_{1} \rightarrow \mathbb{R}, \\
\mathfrak{E}_{2, t}: \mathcal{W}_{2} \times V_{1}^{K} \times \mathcal{E}_{2} \rightarrow \mathbb{R}
\end{array}\right.
$$

Thus, the input symbol of the channel due to transmitter 1 at each time instant $t$ is given by $X_{1, t}=\mathfrak{E}_{1, t}\left(W_{1}, V_{2}^{n}, E_{1, t}\right)$, where $E_{1, t}=\xi_{1}\left(S_{t}\right) \in \mathcal{E}_{1}$, (respectively for transmitter $2, X_{2, t}=\mathfrak{E}_{2, t}\left(W_{2}, V_{1}^{n}, E_{2, t}\right)$, where $\left.E_{2, t}=\xi_{2}\left(S_{t}\right) \in \mathcal{E}_{2}\right)$. We also assume that the receiver knows perfect CSI, i.e., the decoder function is given by:

$$
\mathfrak{D}: \mathcal{Y}^{n} \times\left(\mathbb{R}_{+}^{2}\right)^{n} \rightarrow \mathcal{W}_{1} \times \mathcal{W}_{2}
$$

which estimates the messages as:

$$
\left(\widehat{W}_{1}, \widehat{W}_{2}\right)=\mathfrak{D}\left(Y^{n}, S^{n}\right)
$$

where $S^{n}=\left(S_{1}, S_{2}, \ldots, S_{n}\right)$ is the state process of the channel.

Finally, we note that an average power constraint is also imposed on the codewords of each encoder and the $i^{\text {th }}$ transmitter is subjected to an average power constraint $P_{i} \in \mathbb{R}_{+}, i=1,2$. Precisely speaking, for the codewords of transmitter 1 :

$$
\begin{equation*}
\frac{1}{n} \mathbb{E}\left[\sum_{t=1}^{n}\left(X_{1, t}\left(W_{1}, V_{2}^{n}, \xi_{1}\left(S_{t}\right)\right)\right)^{2}\right] \leq P_{1} \tag{11}
\end{equation*}
$$

and to obtain the power constraint on the codewords of transmitter 2 one can exchange the indices 1 and 2 in (11).
In the next section, we state our main results for the GMAC with a common message and also for the Gaussian fading MAC with conferencing encoders.

## III. Main Results

## III-A) General multiple access channel with a common message

The main result of this subsection is given in Theorem 1 where we provide an exact characterization of the capacity region for the Gaussian fading GMAC defined by (4) with common message. Firstly, we derive the capacity region of the discrete memoryless GMAC (1) with a common message using the superposition coding technique. As we see later, the achievability proof of the capacity region for this channel is useful to obtain the capacity region for the Gaussian fading GMAC (4) with common message.

Proposition 1: The capacity region of the discrete memoryless GMAC with $\mathcal{p}$ transmitters and $\mathcal{q}$, receivers and with common information denoted by $\mathcal{C}_{c m}^{D-G M A C}$, is given by:

$$
\mathcal{C}_{c m}^{D-G M A C}=\bigcup\left\{\begin{array}{l}
\left(R_{0}, R_{1}, R_{2}, \ldots, R_{\mathfrak{p}}\right) \in \mathbb{R}_{+}^{p+1}  \tag{12}\\
\forall \Lambda \subseteq\{1, \ldots, \mathfrak{p}\}, \forall j \in\{1, \ldots, q\} \\
\sum_{k \in \Lambda} R_{k} \leq I\left(\left\{X_{k}\right\}_{k \in \Lambda} ; Y_{j} \mid\left\{X_{k}\right\}_{k \notin \Lambda}, U\right) \\
, \\
\sum_{k=0}^{p} R_{k} \leq I\left(X_{1}, \ldots, X_{p} ; Y_{j}\right)
\end{array}\right\}
$$

where $U$ is an auxiliary random variable and the joint p.d.f. of the r.v.'s $\left\langle U, X_{1}, \ldots, X_{p}, Y_{1}, \ldots, Y_{q}\right\rangle$ is given by:

$$
\begin{equation*}
P_{U X_{1} \ldots X_{\mathcal{p}} Y_{1} \ldots Y_{q}}\left(u, x_{1}, \ldots, x_{p}, y_{1}, \ldots, y_{q}\right)=P_{U}(u) \prod_{i=1}^{p} P_{X_{i} \mid U}\left(x_{i} \mid u\right) \mathbb{P}\left(y_{1}, \ldots, y_{q} \mid x_{1}, \ldots, x_{p}\right) \tag{13}
\end{equation*}
$$

Proof of Proposition 1: This proposition is an extension of the result of [4] to the case of $p$ transmitters and $q$ receivers. In Appendix I, we provide a complete proof of the direct part (achievability proof); similar encoding and decoding schemes are exploited to obtain the capacity region of the Gaussian fading GMAC (4) with common message, as shown in the next theorem. The proof of the converse part will be omitted for brevity.

## Remarks:

1) The cardinality of the auxiliary r.v. $U$ is bounded above as $\|U\| \leq \prod_{i=1}^{p}\left\|X_{i}\right\|+q \times 2^{p}-1$. This can be proved through the standard methods using the Support lemma [20, p. 310].
2) By setting $p=2$ and $q=1$, the result of Proposition 1 reduces to the capacity region of a two user discrete MAC with a common message obtained in [4].
3) By setting $p=q=2$, the result of Proposition 1 reduces to the capacity region of a compound MAC with a common message derived in [10].
4) By setting $U \equiv \emptyset$ and $R_{0}=0$, the result of Proposition 1 reduces to the capacity region of a GMAC without common message derived in [16].

Now, we consider the Gaussian fading GMAC (4) with common message. In the following theorem, we characterize the throughput capacity region [15] of this channel.

Theorem 1: Consider the Gaussian fading GMAC in (4) with a common message, with partial CSIT and perfect CSIR. The capacity region denoted by $\mathcal{C}_{c m}^{G f-G M A C}$, is given by:

$$
\mathcal{C}_{c m}^{G f-G M A C}=\bigcup_{\substack{\left\{\varphi_{i}(.): \mathbb{E}\left[\varphi_{i}\left(E_{j}\right)\right] \leq P_{i}\right\}_{i=1}^{j}  \tag{14}\\
\left\{e_{i}(.): 0 \leq \varrho_{i}(.) \leq 1\right\}_{i=1}^{p}}}\left\{\begin{array}{l}
\left(R_{0}, R_{1}, R_{2}, \ldots, R_{\mathcal{p}}\right) \in \mathbb{R}_{+}^{p+1} \\
\forall \Lambda \subseteq\{1, \ldots, p\}, \forall j \in\{1, \ldots, q\}: \\
\sum_{k \in \Lambda} R_{k} \leq \mathbb{E}_{S}\left[\mathbb{C}\left(\frac{\sum_{k \in \Lambda}\left(\left(s_{j k}\right)^{2} \varphi_{k}\left(E_{k}\right)\left(1-\varrho_{k}^{2}\left(E_{k}\right)\right)\right)}{\sigma_{j}^{2}}\right)\right] \\
\sum_{k=0}^{p} R_{k} \leq \mathbb{E}_{S}\left[\mathbb{C}\left(\frac{\sum_{k=1}^{p}\left(s_{j k}\right)^{2} \varphi_{k}\left(E_{k}\right)}{+2 \sum_{1 \leq k<i \leq \mathcal{p}}\left(s_{j k} S_{j i} e_{k}\left(E_{k}\right) e_{i}\left(E_{i}\right) \sqrt{\left.\varphi_{k}\left(E_{k}\right) \varphi_{i}\left(E_{i}\right)\right)}\right.} \sigma_{j}^{2}\right)\right]
\end{array}\right\}
$$

where $S \in\left[\mathbb{R}_{+}\right]_{q \times p}$ is a r.v. representing the state process of the channel (the matrix of fading coefficients) and $S_{i j} \in \mathbb{R}_{+}, i=$ $1, \ldots, p, j=1, \ldots, q$, is a r.v. representing the corresponding entry of $S, E_{i} \in \mathcal{E}_{i}, i=1, \ldots, p$, is a r.v. representing the CSIT available at the $i^{\text {th }}$ transmitter which is a deterministic function of $S: E_{i}=\xi_{i}(S)$. In addition, the function $\varphi_{i}: \mathcal{E}_{i} \rightarrow \mathbb{R}_{+}$satisfying the following constraint:

$$
\begin{equation*}
\mathbb{E}\left[\varphi_{i}\left(E_{i}\right)\right] \leq P_{i}, \quad i=1, \ldots, p \tag{15}
\end{equation*}
$$

denotes the power allocation policy for the $i^{\text {th }}$ transmitter and $\varrho_{i}: \mathcal{E}_{i} \rightarrow[0,1], i=1, \ldots, p$, is an arbitrary (limited) deterministic function that takes its values from the interval $[0,1]$.

## Remarks:

1) By setting $q=1$, and $R_{0}=0, \varrho_{k}(.) \equiv 0, k=1, \ldots, p$, and also $E_{k} \equiv S, k=1, \ldots, p$, (perfect CSIT), then the rate region (14) reduces to the capacity region of a $\mathfrak{p}$-user Gaussian fading MAC without common message with perfect CSIT and also perfect CSIR, obtained previously in [15, Part I, Th. 2.1].
2) We observe that the throughput capacity region of the Gaussian fading GMAC with common message given in (14) depends only on the stationary (first order) distribution of the joint fading processes, i.e., $P_{S}(s)$ where $S$ is the matrix of fading coefficients and not on the correlation structure (memory of the state process). A same observation could be found in [15, Part I, Th. 2.1] for the capacity of Gaussian fading MAC without common message, as indicated in [15, Part II].
3) As we see from (14), the improvement in the capacity because of having partial knowledge of the CSI at the transmitters comes from the ability to allocate powers (according to the functions $\varphi_{i}: \mathcal{E}_{i} \rightarrow \mathbb{R}^{+}, i=1, \ldots, p$, with constraints (15)) and also from the functions $\varrho_{i}: \varepsilon_{i} \rightarrow[0,1], i=1, \ldots, p$. The functions $\varrho_{i}($.$) appeared in the capacity formulation (14)$ arise from the existence of the common message in the system which causes that the transmitted signals to be correlated. Precisely speaking, by (24) the transmitted signals of $i^{t h}$ and $j^{t h}$ encoders are correlated according to a correlation coefficient denoted by $\Upsilon_{i, j}, i, j=1, \ldots, p$, which is given as follows:

$$
\begin{equation*}
r_{i, j} \triangleq \frac{\mathbb{E}\left[X_{i} X_{j}\right]}{\sqrt{\mathbb{E}\left[X_{i}^{2}\right] \mathbb{E}\left[X_{j}^{2}\right]}}=\frac{\mathbb{E}\left[\sqrt{\varphi_{i}\left(E_{i}\right) \varphi_{j}\left(E_{j}\right)} \varrho_{i}\left(E_{i}\right) \varrho_{j}\left(E_{j}\right)\right]}{\sqrt{\mathbb{E}\left[\varphi_{i}\left(E_{i}\right)\right] \mathbb{E}\left[\varphi_{j}\left(E_{j}\right)\right]}} \tag{16}
\end{equation*}
$$

As mentioned before, for the case where there is no common message in the system these functions disappear, i.e., $\varrho_{k}(.) \equiv 0, k=1, \ldots, p$; therefore, the improvement in the capacity comes only from the ability to allocate powers [15, Part II, Th. 2.1] and the transmitted signals are uncorrelated.
4) Lagrangian characterization of the capacity region: One can solve for the boundary surface of the rate region (14) to explicitly characterize the capacity region. Due to the convexity of the capacity region the boundary surface of $\mathcal{C}_{c m}^{G f-G M A C}$ given by (14), is the closure of all points $\boldsymbol{R}^{*}=\left(R_{0}^{*}, R_{1}^{*}, R_{2}^{*}, \ldots, R_{\mathfrak{p}}^{*}\right)$ such that $\boldsymbol{R}^{*}$ is a solution to the following optimization problem:

$$
\begin{equation*}
\max _{\left(R_{0}, R_{1}, R_{2}, \ldots, R_{\mathfrak{p}}\right)} \sum_{i=0}^{p} \mu_{i} R_{i}, \quad \text { subject to }\left(R_{0}, R_{1}, R_{2}, \ldots, R_{\mathfrak{p}}\right) \in \mathcal{C}_{c m}^{G f-G M A C} \tag{17}
\end{equation*}
$$

For some $\left(\mu_{0}, \mu_{1}, \mu_{2}, \ldots, \mu_{\mathfrak{p}}\right) \in \mathbb{R}_{+}^{\mathfrak{p + 1}}$. Now consider the following set:

$$
\begin{equation*}
\partial \triangleq\left\{(\boldsymbol{R}, \boldsymbol{P}): \boldsymbol{P} \in \mathbb{R}_{+}^{p}, \boldsymbol{R} \in \mathcal{C}_{c m}^{G f-G M A C}(\boldsymbol{P})\right\} \tag{18}
\end{equation*}
$$

where $\mathcal{C}_{c m}^{G f-G M A C}(\boldsymbol{P})$ represents the capacity assuming that the power constraint associated to the $i^{t h}$ transmitter is given by the $i^{t h}$ entry of $\boldsymbol{P}$. By concavity of the logarithm function, it can be shown that $\partial$ is a convex set. Therefore, there exist Lagrange multipliers $\left(\lambda_{1}, \lambda_{2}, \ldots, \lambda_{\mathfrak{p}}\right) \in \mathbb{R}_{+}^{\mathfrak{p}}$, such that $\boldsymbol{R}^{*}$ is a solution to the following equivalent optimization problem:

$$
\begin{equation*}
\max _{(R, P) \in \partial} \sum_{i=0}^{p} \mu_{i} R_{i}-\sum_{i=1}^{p} \lambda_{i} P_{i} \tag{19}
\end{equation*}
$$

As a fact, the capacity region (14) remains unchanged if we replace the power constraint inequalities in (15) with equality. Furthermore, the $\mathcal{C}_{c m}^{G f-G M A C}$ given by (14) is a union taken over the set of all functions $\left\{\varphi_{i}():. \mathbb{E}\left[\varphi_{i}\left(E_{i}\right)\right]=\right.$ $\left.P_{i}\right\}_{i=1}^{\mathfrak{p}}$ and $\left\{\varrho_{i}(.): 0 \leq \varrho_{i}(.) \leq 1\right\}_{i=1}^{p}$. Thus we can rewrite (19) as an optimization over the set of such functions, as follows:

$$
\begin{gather*}
\max _{\left\{\begin{array}{l}
\boldsymbol{R},\left\{\varphi_{i}(.): 0 \leq \varphi_{i}(.)\right\}_{i=1}^{p} \\
\left\{\varrho_{i}(.): 0 \leq \varrho_{i}(.) \leq 1\right\}_{i=1}^{p}
\end{array}\right.} \sum_{i=0}^{p} \mu_{i} R_{i}-\sum_{i=1}^{p} \lambda_{i} \mathbb{E}\left[\varphi_{i}\left(E_{i}\right)\right], \\
\text { subject to } \boldsymbol{R} \in \mathcal{C}_{c m}^{G f-G M A C}\left(\left\{\varphi_{i}(.)\right\}_{i=1}^{p},\left\{\varrho_{i}(.): 0 \leq \varrho_{i}(.) \leq 1\right\}_{i=1}^{p}\right) \tag{20}
\end{gather*}
$$

where $\mathcal{C}_{c m}^{G f-G M A C}\left(\left\{\varphi_{i}(.)\right\}_{i=1}^{p},\left\{\varrho_{i}(.): 0 \leq \varrho_{i}(.) \leq 1\right\}_{i=1}^{p}\right)$ is the capacity where we fix $\varphi_{i}($.$) -functions and \varrho_{i}($.$) -$ functions to be as in the arguments. Solving this optimization problem is not the subject of this paper and we relegate it to future work.

Proof of Theorem 1: We prove this theorem in two parts; first we prove the achievability of (14) through a random coding technique and then we provide a converse theorem.

Achievability part: It is worth noting that for the case where there is no common message in the system, the optimal signaling for the channel is rather trivial (see [15, App. A, p. 2812]). However, the same does not hold for the channel with common message. Here, we prove the achievability scheme through a random coding argument where the encoding and decoding procedure is based on the superposition technique. Let $\left\langle U, V_{1}, \ldots, V_{p}\right\rangle$ be a sequence of auxiliary r.v.'s with alphabets $U \in U$ and $V_{i} \in \mathcal{V}_{i}, i=$ $1, \ldots, p$. First, we show that for the Gaussian fading GMAC (4) with a common message, with partial CSIT and perfect CSIR, the following rate region is achievable:
where:

$$
\begin{equation*}
Y_{j}=\sum_{k=1}^{p} S_{j k} X_{k}+Z_{j}, \quad j=1, \ldots, q \tag{22}
\end{equation*}
$$

and the joint p.d.f. of the r.v.'s $\left\langle S, E_{1}, \ldots, E_{\mathfrak{p}}, U, V_{1}, \ldots, V_{\mathfrak{p}}\right\rangle$ is given by:

$$
\begin{equation*}
P_{S E_{1} \ldots E_{\mathfrak{p}} U V_{1} \ldots V_{\mathfrak{p}}}\left(s, e_{1}, \ldots, e_{\mathfrak{p}}, u, v_{1}, \ldots, v_{\mathfrak{p}}\right)=P_{S E_{1} \ldots E_{\mathfrak{p}}}\left(s, e_{1}, \ldots, e_{\mathfrak{p}}\right) P_{U}(u) \prod_{i=1}^{p} P_{V_{i} \mid U}\left(v_{i} \mid u\right) . \tag{23}
\end{equation*}
$$

Furthermore, $\left\{f_{i}(.): \mathcal{V}_{i} \times U \times \mathcal{E}_{i} \rightarrow \mathbb{R}\right\}_{i=1}^{p}$ is a set of deterministic functions such that $X_{i}=f_{i}\left(V_{i}, U, E_{i}\right), i=1, \ldots, p$, satisfies the power constraint policy of the $i^{\text {th }}$ transmitter: $\mathbb{E}\left[X_{i}^{2}\right] \leq P_{i}$.

To prove this, we use a random codebook generation technique as follows:

1) Fix the distribution $P_{U V_{1} \ldots v_{\mathcal{p}}}\left(u, v_{1}, \ldots, v_{\mathcal{p}}\right)=P_{U}(u) \prod_{i=1}^{p} P_{V_{i} \mid U}\left(v_{i} \mid u\right)$ and the set of deterministic functions $\left\{f_{i}():. v_{i} \times\right.$ $\left.U \times \mathcal{E}_{i} \rightarrow \mathbb{R}\right\}_{i=1}^{p}$ such that $\mathbb{E}\left[\left(f_{i}\left(V_{i}, U, E_{i}\right)\right)^{2}\right] \leq P_{i}, i=1, \ldots, p$.
2) Generate at random $2^{n R_{0}}$ i.i.d. sequences $u^{n}$ according to $P\left(u^{n}\right)=\prod_{t=1}^{n} P_{U}\left(u_{t}\right)$. Label these sequences $u^{n}\left(w_{0}\right)$, where $w_{0} \in\left\{1, \ldots, 2^{n R_{0}}\right\}$.
3) For each $u^{n}\left(w_{0}\right)$, generate $2^{n R_{i}}, i=\{1,2, \ldots, p\}$, sequences $v_{i}^{n}$ according to the p.d.f. $P\left(v_{i}^{n} \mid u^{n}\left(w_{0}\right)\right)=\prod_{t=1}^{n} P_{V_{i} \mid U}\left(v_{i, t} \mid u_{t}\right)$. Label these sequences $v_{i}^{n}\left(w_{0}, w_{i}\right)$, where $w_{i} \in\left\{1, \ldots, 2^{n R_{i}}\right\}$.

Encoding: To send a common message $w_{0}$ and a private message $w_{i}$, at each time instant $t, 1 \leq t \leq n$, the $i^{\text {th }}$ encoder assuming observation of $e_{i, t} \in \mathcal{E}_{i}$ as CSIT, sends $x_{i, t}=f_{i}\left(v_{i, t}\left(w_{0}, w_{i}\right), u_{t}\left(w_{0}\right), e_{i, t}\right)$ over the channel.

Decoding at the $j^{\text {th }}$ receiver, $j \in\{1,2, \ldots, q\}$ : The $j^{\text {th }}$ receiver assuming reception of the signal sequence $y_{j}^{n}$ and tracking the state process of the channel perfectly as $s^{n}$, tries to find a unique $\left(\widehat{w}_{0}, \widehat{w}_{1}, \ldots, \widehat{w}_{p}\right)$ such that $\left(u^{n}\left(\widehat{w}_{0}\right), v_{1}^{n}\left(\widehat{w}_{0}, \widehat{w}_{1}\right), \ldots, v_{p}^{n}\left(\widehat{w}_{0}, \widehat{w}_{p}\right), y_{j}^{n}, s^{n}\right) \in A_{\epsilon}^{n}\left(P_{S U V_{1} \ldots V_{p} Y_{j}}\right)$.

By analysis of the error probability that is similar to the proof of direct part of Proposition 1 (given in Appendix I), and noting that the state process is stationary, the channel is memoryless and the codewords $\left\langle U^{n}, V_{1}^{n}, \ldots, V_{p}^{n}\right\rangle$ were generated i.i.d. and independent of the state process $S^{n}$, one can see that the probability of error tends to zero provided that the rate tuple $\left(R_{0}, R_{1}, R_{2}, \ldots, R_{p}\right)$ belongs to the rate region (21).

Now, to prove the achievability of (14) let $\left\langle U, V_{1}, V_{2}, \ldots, V_{p}\right\rangle$ be a sequence of Gaussian-distributed r.v.'s with zero mean and unit variance, independent of each other and also independent of the state matrix $S$;
In addition, let $\left\{\varphi_{i}: \varepsilon_{i} \rightarrow \mathbb{R}_{+}\right\}_{i=1}^{p}$ be a set of power allocation policy functions which satisfy the power constraints in (15) and $\left\{\varrho_{i}: \mathcal{E}_{i} \rightarrow[0,1]\right\}_{i=1}^{p}$ be a set of arbitrary deterministic functions with range $[0,1]$. Define the r.v.'s $\left\langle X_{1}, X_{2}, \ldots, X_{p}\right\rangle$ as:

$$
\begin{equation*}
X_{i} \triangleq \sqrt{\varphi_{i}\left(E_{i}\right)}\left(\varrho_{i}\left(E_{i}\right) U+\sqrt{1-\varrho_{i}^{2}\left(E_{i}\right)} V_{i}\right), \quad i=1, \ldots, p \tag{24}
\end{equation*}
$$

Note that in definition (24), $\left\langle E_{1}, \ldots, E_{p}\right\rangle$ is the sequence of CSITs at the transmitters with joint p.d.f. determined by $P_{S E_{1} \ldots E_{\mathfrak{p}}}\left(s, e_{1}, \ldots, e_{\mathfrak{p}}\right)$. One can easily check that the r.v.'s $\left\langle X_{1}, X_{2}, \ldots, X_{\mathfrak{p}}\right\rangle$ defined in (24) satisfy:

$$
\begin{equation*}
\mathbb{E}_{S}\left[X_{i}^{2}\right]=\mathbb{E}_{S}\left[\varphi_{i}\left(E_{i}\right)\right] \leq P_{i}, \quad i=1, \ldots, p \tag{25}
\end{equation*}
$$

Thus by substituting $\left\langle U, V_{1}, V_{2}, \ldots, V_{p}\right\rangle$ and $\left\langle X_{1}, X_{2}, \ldots, X_{p}\right\rangle$ as defined in (24) in the rate region given by (21) we have:
$\forall \Lambda \subseteq\{1, \ldots, p\}, \forall j \in\{1, \ldots, q\}:$
$I\left(\left\{V_{k}\right\}_{k \in \Lambda} ; Y_{j} \mid\left\{V_{k}\right\}_{k \notin \Lambda}, U, S\right) \stackrel{(a)}{=} \boldsymbol{h}\left(Y_{j} \mid\left\{V_{k}\right\}_{k \notin \Lambda},\left\{X_{k}\right\}_{k \notin \Lambda}, U, S\right)-\boldsymbol{h}\left(Y_{j} \mid\left\{V_{k}\right\}_{k=1}^{p},\left\{X_{k}\right\}_{k=1}^{p}, U, S\right)$
$=\boldsymbol{h}\left(\sum_{k=1}^{\mathcal{p}} S_{j k} X_{k}+Z_{j} \mid\left\{V_{k}\right\}_{k \notin \Lambda},\left\{X_{k}\right\}_{k \notin \Lambda}, U, S\right)-\boldsymbol{h}\left(\sum_{k=1}^{p} S_{j k} X_{k}+Z_{j} \mid\left\{V_{k}\right\}_{k=1}^{p},\left\{X_{k}\right\}_{k=1}^{p}, U, S\right)$
$\stackrel{(b)}{=} \boldsymbol{h}\left(\sum_{k \in \Lambda} S_{j k} X_{k}+Z_{j} \mid U, S\right)-\frac{1}{2} \log \left(2 \pi e \sigma_{j}^{2}\right)$
$=\boldsymbol{h}\left(\sum_{k \in \Lambda} S_{j k} \sqrt{\varphi_{k}\left(E_{k}\right)}\left(\varrho_{k}\left(E_{k}\right) U+\sqrt{1-\varrho_{k}^{2}\left(E_{k}\right)} V_{k}\right)+Z_{j} \mid U, S\right)-\frac{1}{2} \log \left(2 \pi e \sigma_{j}^{2}\right)$
$=\boldsymbol{h}\left(\sum_{k \in \Lambda} S_{j k} \sqrt{\varphi_{k}\left(E_{k}\right)}\left(\sqrt{1-\varrho_{k}^{2}\left(E_{k}\right)} V_{k}\right) V_{k}+Z_{j} \mid S\right)-\frac{1}{2} \log \left(2 \pi e \sigma_{j}^{2}\right)$
$=\mathbb{E}_{S}\left[\boldsymbol{h}\left(\sum_{k \in \Lambda} s_{j k} \sqrt{\varphi_{k}\left(\xi_{k}(s)\right)}\left(\sqrt{1-\varrho_{k}^{2}\left(\xi_{k}(s)\right)} V_{k}\right)+Z_{j} \mid S=s\right)\right]-\frac{1}{2} \log \left(2 \pi e \sigma_{j}^{2}\right)$
$=\mathbb{E}_{S}\left[\mathbb{C}\left(\frac{\sum_{k \in \Lambda}\left(\left(s_{j k}\right)^{2} \varphi_{k}\left(E_{k}\right)\left(1-\varrho_{k}^{2}\left(E_{k}\right)\right)\right)}{\sigma_{j}^{2}}\right)\right]$
where the equality (a) is due to the fact that $X_{k}$ is known given $\left(V_{k}, U, E_{k}=\xi_{k}(S)\right)$ and (b) is because from (24) $\left(\left\{V_{k}\right\}_{k \notin \Lambda},\left\{X_{k}\right\}_{k \notin \Lambda}\right) \rightarrow U, S \rightarrow\left\{X_{k}\right\}_{k \in \Lambda}$ forms a Markov chain. Furthermore, we can write:
$I\left(U, V_{1}, \ldots, V_{p} ; Y_{j} \mid S\right)=\boldsymbol{h}\left(\sum_{k=1}^{p} S_{j k} X_{k}+Z_{j} \mid S\right)-\boldsymbol{h}\left(\sum_{k=1}^{p} S_{j k} X_{k}+Z_{j} \mid U,\left\{V_{k}\right\}_{k=1}^{\mathfrak{p}},\left\{X_{k}\right\}_{k=1}^{p}, S\right)$
$=\boldsymbol{h}\left(\sum_{k=1}^{p} S_{j k} \sqrt{\varphi_{k}\left(E_{k}\right)}\left(\varrho_{k}\left(E_{k}\right) U+\sqrt{1-\varrho_{k}^{2}\left(E_{k}\right)} V_{k}\right)+Z_{j} \mid S\right)-\frac{1}{2} \log \left(2 \pi e \sigma_{j}^{2}\right)$
$=\mathbb{E}_{S}\left[\boldsymbol{h}\left(\sum_{k=1}^{p} s_{j k} \sqrt{\varphi_{k}\left(\xi_{k}(s)\right)}\left(\varrho_{k}\left(\xi_{k}(s)\right) U+\sqrt{1-\varrho_{k}^{2}\left(\xi_{k}(s)\right)} V_{k}\right)+Z_{j} \mid S=s\right)\right]-\frac{1}{2} \log \left(2 \pi e \sigma_{j}^{2}\right)$
$=\mathbb{E}_{S}\left[\mathbb{C}\left(\frac{\sum_{k=1}^{p}\left(S_{j k}\right)^{2} \varphi_{k}\left(E_{k}\right)}{+2 \sum_{1 \leq k<i \leq p}\left(S_{j k} S_{j i} \varphi_{k}\left(E_{k}\right) \varrho_{i}\left(E_{i}\right) \sqrt{\varphi_{k}\left(E_{k}\right) \varphi_{i}\left(E_{i}\right)}\right)} \sigma_{j}^{2}\right)\right]$

Now, by substituting (26) and (27) in the rate region (21) we obtain the achievability of (14).
Converse part: To prove the converse part we first derive an outer bound on $\mathcal{C}_{c m}^{G f-G M A C}$ in terms of mutual information functions, in the following lemma.

Lemma 1: $\mathcal{C}_{c m}^{G f-G M A C}$ is outer bounded by:

$$
\mathcal{C}_{c m}^{G f-G M A C} \subseteq \bigcup\left\{\begin{array}{l}
\left(R_{0}, R_{1}, R_{2}, \ldots, R_{p}\right) \in \mathbb{R}_{+}^{p+1}  \tag{28}\\
\forall \Lambda \subseteq\{1, \ldots, p\}, \forall j \in\{1, \ldots, q\}, \\
\sum_{k \in \Lambda} R_{k} \leq I\left(\left\{X_{k}\right\}_{k \in \Lambda} ; Y_{j} \mid\left\{X_{k}\right\}_{k \notin \Lambda}, U, S\right) \\
\sum_{k=0}^{p} R_{k} \leq I\left(X_{1}, \ldots, X_{p} ; Y_{j} \mid S\right)
\end{array}\right\}
$$

over all joint p.d.f. of the form:

$$
\begin{equation*}
P_{S E_{1} \ldots E_{\mathcal{p}} U X_{1} \ldots X_{\mathcal{p}}}\left(s, e_{1}, \ldots, e_{\mathfrak{p}}, u, x_{1}, \ldots, x_{\mathfrak{p}}\right)=P_{S E_{1} \ldots E_{\mathfrak{p}}}\left(s, e_{1}, \ldots, e_{\mathfrak{p}}\right) P_{U}(u) \prod_{i=1}^{p} P_{X_{i} \mid U E_{i}}\left(x_{i} \mid u, e_{i}\right) \tag{29}
\end{equation*}
$$

that satisfy the power constraints: $\mathbb{E}\left[X_{i}^{2}\right] \leq P_{i}, i=1, \ldots, p$; where $S \in\left[\mathbb{R}_{+}\right]_{q \times p}$ is a r.v. representing the state process of the channel defined in (5) and $E_{i} \in \mathcal{E}_{i}$ is the r.v. representing the CSIT available at the $i^{t h}$ transmitter that is (by definition of the channel) a deterministic function of $S: E_{i}=\xi_{i}(S), i=1, \ldots, p$. We also note that $P_{S E_{1} \ldots E_{p}}\left(s, e_{1}, \ldots, e_{p}\right)$ is obtained from the (first order) distribution of the channel state process, i.e., $P_{S}(s)$, which describes the behavior of the (matrix) fading coefficients.

## Proof: See Appendix II.

Now, to see that the rate region (14) is also an outer bound for the capacity region of the underlying channel, i.e., $\mathcal{C}_{c m}^{G f-G M A C}$, it is sufficient to show that is a superset of (28). However, we prove a stronger result that is the rate region described by (28) is equivalent to that one given by (14). First notice that one can define the sequence of r.v.'s $\left\langle X_{1}, X_{2}, \ldots, X_{\mathcal{p}}\right\rangle$ as in (24) where $\left\langle U, V_{1}, V_{2}, \ldots, V_{p}\right\rangle$ is a sequence of Gaussian-distributed r.v.'s with zero mean and unit variance, independent of each other and also independent of the state matrix $S$, to satisfy the joint p.d.f. (29) and also the power constraints $\mathbb{E}_{S}\left[X_{i}^{2}\right] \leq P_{i}, i=1, \ldots, p$. Thus by substituting $\left\langle U, X_{1}, X_{2}, \ldots, X_{p}\right\rangle$ as defined in (24) in the rate region described by (28) one can see that (14) results. Conversely, we show that the rate region (28) does not exceed any points of (14); first, we state some useful technical lemmas from probability theory in the following:
For arbitrary r.v.'s $X$ and $Y, \operatorname{VAR}(X \mid Y=y)$ is defined as:

$$
\begin{equation*}
\operatorname{VAR}(X \mid Y=y) \triangleq \mathbb{E}\left[X^{2} \mid Y=y\right]-(\mathbb{E}[X \mid Y=y])^{2} \tag{30}
\end{equation*}
$$

in fact, it is the variance of $X$ with respect to the distribution $P_{X \mid Y}(x \mid y)$ and therefore a positive quantity.
Lemma 2: Consider three r.v.'s $X, Y$ and $Z$ :
I) Assume that $Y$ and $Z$ are independent. The following holds:

$$
\begin{equation*}
\mathbb{E}_{Y}\left[\mathbb{E}\left[X^{2} \mid Y, Z=z\right]\right]=\mathbb{E}\left[X^{2} \mid Z=z\right] \tag{31}
\end{equation*}
$$

II) Assume that $X$ and $Z$ are $\mathbb{R}$-valued and $X \rightarrow Y \rightarrow Z$ forms a Markov chain. Then:

$$
\begin{equation*}
\operatorname{VAR}(X+Z \mid Y=y)=\operatorname{VAR}(X \mid Y=y)+\operatorname{VAR}(Z \mid Y=y) \tag{32}
\end{equation*}
$$

Proof: By direct computation.
Lemma 3: Consider the r.v.'s $\left\langle S, E_{1}, E_{2}, U, X_{1}, X_{2}\right\rangle$ with the joint p.d.f. given as follows:

$$
\begin{equation*}
P_{S E_{1} E_{2} U X_{1} X_{2}}\left(s, e_{1}, e_{2}, u, x_{1}, x_{2}\right)=P_{S E_{1} E_{2}}\left(s, e_{1}, e_{2}\right) P_{U}(u) \prod_{i=1}^{2} P_{X_{i} \mid U E_{i}}\left(x_{i} \mid u, e_{i}\right) \tag{33}
\end{equation*}
$$

where $X_{1}$ and $X_{2}$ are $\mathbb{R}$-valued and there exist deterministic functions $\xi_{i}($.$) such that: E_{i}=\xi_{i}(S), i=1,2$. Then, the following inequality holds:

$$
\begin{equation*}
\mathbb{E}\left[X_{1} X_{2} \mid S=s\right] \leq \sqrt{\mathbb{E}_{U}\left[\left(\mathbb{E}\left[X_{1} \mid U, E_{1}=\xi_{1}(s)\right]\right)^{2}\right]} \sqrt{\mathbb{E}_{U}\left[\left(\mathbb{E}\left[X_{2} \mid U, E_{2}=\xi_{2}(s)\right]\right)^{2}\right]} \tag{34}
\end{equation*}
$$

Proof: Since $S$ is independent of $U$, we can write:
$\mathbb{E}\left[X_{1} X_{2} \mid S=s\right]=\mathbb{E}_{U}\left[\mathbb{E}\left[X_{1} X_{2} \mid U, S=s\right]\right]$
$\stackrel{(a)}{=} \mathbb{E}_{U}\left[\mathbb{E}\left[X_{1} \mid U, S=s\right] \mathbb{E}\left[X_{2} \mid U, S=s\right]\right]$
$\stackrel{(b)}{=} \mathbb{E}_{U}\left[\mathbb{E}\left[X_{1} \mid U, E_{1}=\xi_{1}(s)\right] \mathbb{E}\left[X_{2} \mid U, E_{2}=\xi_{2}(s)\right]\right]$
$\stackrel{(c)}{\leq} \sqrt{\mathbb{E}_{U}\left[\left(\mathbb{E}\left[X_{1} \mid U, E_{1}=\xi_{1}(s)\right]\right)^{2}\right]} \sqrt{\mathbb{E}_{U}\left[\left(\mathbb{E}\left[X_{2} \mid U, E_{2}=\xi_{2}(s)\right]\right)^{2}\right]}$
where (a) is true because $X_{2} \rightarrow(U, S) \rightarrow X_{1}$ forms a Markov chain, (b) is true since $S \rightarrow\left(U, E_{1}=\xi_{1}(S)\right) \rightarrow X_{1}$ and also $S \rightarrow\left(U, E_{2}=\xi_{2}(S)\right) \rightarrow X_{2}$ form Markov chains, and finally (c) is due to the Cauchy-Schwarz inequality.

Now, consider the rate region described in (28). For $i=1, \ldots, \mathcal{p}$, define deterministic functions $\varphi_{i}():. \mathcal{E}_{i} \rightarrow \mathbb{R}^{+}$such that:

$$
\begin{equation*}
\forall e_{i} \in \mathcal{E}_{i}: \quad \varphi_{i}\left(e_{i}\right) \triangleq \mathbb{E}\left[X_{i}^{2} \mid E_{i}=e_{i}\right] \tag{36}
\end{equation*}
$$

and also deterministic functions $\varrho_{i}():. \varepsilon_{i} \rightarrow[0,1]$ as follows:

$$
\begin{equation*}
\forall e_{i} \in \mathcal{E}_{i}: \quad \varrho_{i}\left(e_{i}\right) \triangleq \sqrt{\frac{\mathbb{E}_{U}\left[\left(\mathbb{E}\left[X_{i} \mid U, E_{i}=e_{i}\right]\right)^{2}\right]}{\varphi_{i}\left(e_{i}\right)}} \tag{37}
\end{equation*}
$$

Notice that:

$$
\begin{equation*}
\mathbb{E}\left[X_{i}^{2} \mid U=u, E_{i}=e_{i}\right]-\left(\mathbb{E}\left[X_{i} \mid U=u, E_{i}=e_{i}\right]\right)^{2}=\operatorname{VAR}\left(X_{i} \mid U=u, E_{i}=e_{i}\right) \geq 0 \tag{38}
\end{equation*}
$$

which yields:

$$
\begin{equation*}
\mathbb{E}\left[X_{i}^{2} \mid E_{i}=e_{i}\right]-\mathbb{E}_{U}\left[\left(\mathbb{E}\left[X_{i} \mid U, E_{i}=e_{i}\right]\right)^{2}\right] \stackrel{(a)}{=} \mathbb{E}_{U}\left[\mathbb{E}\left[X_{i}^{2} \mid U, E_{i}=e_{i}\right]-\left(\mathbb{E}\left[X_{i} \mid U, E_{i}=e_{i}\right]\right)^{2}\right] \geq 0 \tag{39}
\end{equation*}
$$

where (a) is obtained from Part I of Lemma 2. Therefore, for each $e_{i} \in \varepsilon_{i}, \varrho_{i}\left(e_{i}\right)$ belongs to the interval [ 0,1$]$ and the notation $\varrho_{i}():. \varepsilon_{i} \rightarrow[0,1]$ is well-defined. Furthermore, from (36) we have:

$$
\begin{equation*}
\mathbb{E}\left[\varphi_{i}\left(E_{i}\right)\right]=\mathbb{E}\left[\mathbb{E}\left[X_{i}^{2} \mid E_{i}\right]\right]=\mathbb{E}\left[X_{i}^{2}\right] \leq P_{i}, \quad i=1, \ldots, p \tag{40}
\end{equation*}
$$

Now for the mutual information function terms in the rate region (28) we can proceed as:
$\forall \Lambda \subseteq\{1, \ldots, p\}, \forall j \in\{1, \ldots, q\}$,
$I\left(\left\{X_{k}\right\}_{k \in \Lambda} ; Y_{j} \mid\left\{X_{k}\right\}_{k \notin \Lambda}, U, S\right)=\boldsymbol{h}\left(\sum_{k=1}^{p} S_{j k} X_{k}+Z_{j} \mid\left\{X_{k}\right\}_{k \notin \Lambda}, U, S\right)-\boldsymbol{h}\left(\sum_{k=1}^{p} S_{j k} X_{k}+Z_{j} \mid\left\{X_{k}\right\}_{k=1}^{p}, U, S\right)$
$\stackrel{(a)}{=} \boldsymbol{h}\left(\sum_{k \in \Lambda} S_{j k} X_{k}+Z_{j} \mid U, S\right)-\frac{1}{2} \log \left(2 \pi e \sigma_{j}^{2}\right)$
$\stackrel{(b)}{=} \mathbb{E}_{S}\left[\mathbb{E}_{U}\left[\boldsymbol{h}\left(\sum_{k \in \Lambda} S_{j k} X_{k}+Z_{j} \mid U=u, S=s\right)\right]\right]-\frac{1}{2} \log \left(2 \pi e \sigma_{j}^{2}\right)$
$\leq \mathbb{E}_{S}\left[\mathbb{E}_{U}\left[\frac{1}{2} \log \left(2 \pi e V A R\left(\sum_{k \in \Lambda} S_{j k} X_{k}+Z_{j} \mid U=u, S=s\right)\right)\right]\right]-\frac{1}{2} \log \left(2 \pi e \sigma_{j}^{2}\right)$
$=\mathbb{E}_{S}\left[\mathbb{E}_{U}\left[\frac{1}{2} \log \left(2 \pi e\left(\sum_{k \in \Lambda}\left(S_{j k}\right)^{2} \operatorname{VAR}\left(X_{k} \mid U=u, S=s\right)+\sigma_{j}^{2}\right)\right)\right]\right]-\frac{1}{2} \log \left(2 \pi e \sigma_{j}^{2}\right)$
$\stackrel{(c)}{=} \mathbb{E}_{S}\left[\mathbb{E}_{U}\left[\mathbb{C}\left(\frac{\sum_{k \in \Lambda}\left(s_{j k}\right)^{2} V A R\left(X_{k} \mid U=u, E_{k}=e_{k}\right)}{\sigma_{j}^{2}}\right)\right]\right]$
$=\mathbb{E}_{S}\left[\mathbb{E}_{U}\left[\mathbb{C}\left(\frac{\sum_{k \in \Lambda}\left(S_{j k}\right)^{2}\left(\mathbb{E}\left[X_{k}^{2} \mid U=u, E_{k}=e_{k}\right]-\left(\mathbb{E}\left[X_{i} \mid U=u, E_{i}=e_{i}\right]\right)^{2}\right)}{\sigma_{j}^{2}}\right)\right]\right]$
$\stackrel{(d)}{\leq} \mathbb{E}_{S}\left[\mathbb{C}\left(\frac{\sum_{k \in \Lambda}\left(s_{j k}\right)^{2} \mathbb{E}_{U}\left[\mathbb{E}\left[X_{k}^{2} \mid U=u, E_{k}=e_{k}\right]-\left(\mathbb{E}\left[X_{i} \mid U=u, E_{i}=e_{i}\right]\right)^{2}\right]}{\sigma_{j}^{2}}\right)\right]$
$\stackrel{(f)}{=} \mathbb{E}_{S}\left[\mathbb{C}\left(\frac{\sum_{k \in \Lambda}\left(\left(s_{k j}\right)^{2} \varphi_{k}\left(E_{k}\right)\left(1-\varrho_{k}^{2}\left(E_{k}\right)\right)\right)}{\sigma_{j}^{2}}\right)\right]$
where (a) is true since from the joint p.d.f. (29), $\left\{X_{k}\right\}_{k \in \Lambda} \rightarrow(U, S) \rightarrow\left\{X_{k}\right\}_{k \notin \Lambda}$ forms a Markov chain, (b) is true since $U$ and $S$ are independent of each other, (c) is due to the fact that from (29) $S \rightarrow\left(U, E_{k}\right) \rightarrow X_{k}$ forms a Markov chain, (d) is due to Jensen's inequality and (f) is obtained from definitions (36) and (37). Moreover, for the bounds on the sum rate we have:
$I\left(X_{1}, \ldots, X_{p} ; Y_{j} \mid S\right)=\boldsymbol{h}\left(\sum_{k=1}^{p} S_{j k} X_{k}+Z_{j} \mid S\right)-\boldsymbol{h}\left(\sum_{k=1}^{p} S_{j k} X_{k}+Z_{j} \mid\left\{X_{k}\right\}_{k=1}^{p}, S\right)$
$=\boldsymbol{h}\left(\sum_{k=1}^{p} S_{j k} X_{k}+Z_{j} \mid S\right)-\frac{1}{2} \log \left(2 \pi e \sigma_{j}^{2}\right)$
$=\mathbb{E}_{S}\left[\boldsymbol{h}\left(\sum_{k=1}^{p} S_{j k} X_{k}+Z_{j} \mid S=s\right)\right]-\frac{1}{2} \log \left(2 \pi e \sigma_{j}^{2}\right)$
$\leq \mathbb{E}_{S}\left[\frac{1}{2} \log \left(2 \pi e \mathbb{E}\left[\left(\sum_{k=1}^{p} S_{j k} X_{k}+Z_{j}\right)^{2} \mid S=s\right]\right)\right]-\frac{1}{2} \log \left(2 \pi e \sigma_{j}^{2}\right)$
$=\mathbb{E}_{S}\left[\mathbb{C}\left(\frac{\sum_{k=1}^{p}\left(s_{j k}\right)^{2} \mathbb{E}\left[X_{k}^{2} \mid S=s\right]}{+2 \sum_{1 \leq k<i \leq p}\left(s_{j k} S_{j i} \mathbb{E}\left[X_{k} X_{i} \mid S=s\right]\right)} \sigma_{j}^{2}\right)\right]$
$\stackrel{(a)}{\leq} \mathbb{E}_{S}\left[\mathbb{C}\left(\frac{\sum_{k=1}^{p}\left(S_{j k}\right)^{2} \mathbb{E}\left[X_{k}^{2} \mid E_{k}=\xi_{k}(s)\right]}{+2 \Sigma_{1 \leq k<i \leq p}\left(s_{j k} s_{j i} \sqrt{\mathbb{E}_{U}\left[\left(\mathbb{E}\left[X_{k} \mid U, E_{k}=\xi_{k}(s)\right]\right)^{2}\right]} \sqrt{\mathbb{E}_{U}\left[\left(\mathbb{E}\left[X_{i} \mid U, E_{i}=\xi_{i}(s)\right]\right)^{2}\right]}\right)} \sigma_{j}^{\sigma_{j}^{2}}\right)\right]$
$\stackrel{(b)}{=} \mathbb{E}_{S}\left[\mathbb{C}\left(\frac{\sum_{k=1}^{p}\left(S_{j k}\right)^{2} \varphi_{k}\left(E_{k}\right)}{+2 \sum_{1 \leq k<i \leq p}\left(S_{j k} S_{j i} e_{k}\left(E_{k}\right) \varrho_{i}\left(E_{i}\right) \sqrt{\left.\varphi_{k}\left(E_{k}\right) \varphi_{i}\left(E_{i}\right)\right)}\right.} \begin{array}{|l}\sigma_{j}^{2}\end{array}\right)\right]$
where (a) is obtained from Lemma 3 and also because by the joint p.d.f. (29) $S \rightarrow E_{k} \rightarrow X_{k}$ forms a Markov chain. The equality (b) is obtained from (36) and (37). Now by substituting (41) and (42) in (28) we obtain the fact that no point outside the rate region (14) is achievable. This completes the proof of Theorem 2.

As mentioned in introduction, the capacity region of the Gaussian fading two transmitter/one receiver MAC with common message where perfect CSI is available at both transmitters and the receiver was previously characterized in [18]. It is clear that the result of [18] is a special case of Theorem 1. On the other hand, it can be verified that the techniques used in [18] to derive its result are different from ours, yielding different formulation of the capacity region. Nonetheless, it is expected that these two formulations are equivalent. Let us consider the two-transmitter/one-receiver Gaussian fading MAC with a common message formulated as (10), where for notational convenience we set $\sigma^{2} \equiv 1$. Moreover, assume that both transmitters and also the receiver have access to perfect CSI. Using Theorem 1, the capacity region under these conditions is given by:

$$
\bigcup_{\substack{\left.\varphi_{i}(\cdot): \mathbb{R}+\rightarrow \mathbb{R}_{+}  \tag{43}\\
e_{i}(\cdot): \mathbb{R}_{+}^{+}+0,0,1\right] \\
i=1,2}}\left\{\begin{array}{l}
\left(R_{0}, R_{1}, R_{2}\right) \in \mathbb{R}_{+}^{3} \\
R_{1} \leq \mathbb{E}_{S}\left[\mathbb{C}\left(\left(S_{1}\right)^{2} \varphi_{1}(S)\left(1-\varrho_{1}^{2}(S)\right)\right)\right] \\
R_{2} \leq \mathbb{E}_{S}\left[\mathbb{C}\left(\left(S_{2}\right)^{2} \varphi_{2}(S)\left(1-\varrho_{2}^{2}(S)\right)\right)\right] \\
R_{1}+R_{2} \leq \mathbb{E}_{S}\left[\mathbb{C}\left(\sum_{k=1}^{2}\left(S_{k}\right)^{2} \varphi_{k}(S)\left(1-\varrho_{k}^{2}(S)\right)\right)\right] \\
R_{0}+R_{1}+R_{2} \leq \mathbb{E}_{S}\left[\mathbb{C}\binom{\left(S_{1}\right)^{2} \varphi_{1}(S)+\left(S_{2}\right)^{2} \varphi_{2}(S)}{+2 S_{1} S_{2} \varrho_{1}(S) \varrho_{2}(S) \sqrt{\varphi_{1}(S) \varphi_{2}(S)}}\right]
\end{array}\right\}
$$

where $S=\left(S_{1}, S_{2}\right) \in \mathbb{R}_{+}^{2}$ is a r.v. representing the state process of the channel, $\varphi_{i}():. \mathbb{R}_{+}^{2} \rightarrow \mathbb{R}_{+}$is the power allocation policy for the $i^{\text {th }}$ transmitter satisfying the following constraint:

$$
\mathbb{E}\left[\varphi_{i}(S)\right] \leq P_{i}, \quad i=1,2
$$

and $\varrho_{i}():. \mathbb{R}_{+}^{2} \rightarrow[0,1]$ is an arbitrary deterministic function. This is a consequence of Theorem 1 by setting $p \equiv 2, q \equiv 1$, and $E_{i} \equiv S=\left(S_{1}, S_{2}\right), i=1,2,($ perfect CSIT) in (14).
Now, let us bring out the formulation of [18] for the capacity region of the underlying channel, i.e., the two-transmitter/onereceiver Gaussian fading MAC with a common message with perfect CSIT and perfect CSIR. Using our notations, this formulation is given by:

$$
\bigcup_{\substack{\varrho(.): \mathbb{R}_{+}^{2} \rightarrow[0,1]  \tag{45}\\
p_{i}(.): \mathbb{R}_{+}^{2} \rightarrow \mathbb{R}_{+} \\
i=0,1,2}}\left\{\begin{array}{l}
\left(R_{0}, R_{1}, R_{2}\right) \in \mathbb{R}_{+}^{3} \\
R_{1} \leq \mathbb{E}_{S}\left[\mathbb{C}\left(p_{1}(H)\right)\right] \\
R_{2} \leq \mathbb{E}_{S}\left[\mathbb{C}\left(p_{2}(H)\right)\right] \\
R_{1}+R_{2} \leq \mathbb{E}_{S}\left[\mathbb{C}\left(p_{1}(H)+p_{2}(H)\right)\right] \\
R_{0}+R_{1}+R_{2} \leq \mathbb{E}_{S}\left[\mathbb{C}\left(p_{1}(H)+p_{2}(H)+p_{0}(H)\right)\right]
\end{array}\right\}
$$

where $S=\left(S_{1}, S_{2}\right) \in \mathbb{R}_{+}^{2}, H=\left(S_{1}^{2}, S_{2}^{2}\right) \in \mathbb{R}_{+}^{2}$, and the functions $\varrho():. \mathbb{R}_{+}^{2} \rightarrow[0,1]$ and $p_{i}():. \mathbb{R}_{+}^{2} \rightarrow \mathbb{R}_{+}, i=0,1,2$, satisfy the following conditions:

$$
\left\{\begin{array}{l}
\mathbb{E}_{S}\left[\frac{p_{1}(H)}{S_{1}^{2}}+\frac{\varrho(H)^{2} p_{0}(H)}{S_{1}^{2}}\right] \leq P_{1}  \tag{46}\\
\mathbb{E}_{S}\left[\frac{p_{2}(H)}{S_{2}^{2}}+\frac{(1-\varrho(H))^{2} p_{0}(H)}{S_{2}^{2}}\right] \leq P_{2}
\end{array}\right.
$$

As we see, the formulation (43) is completely different from (45). Next, we prove that these two formulations describe the same rate region.

Proposition 2: The rate region described by (43) is equivalent to (45)-(46).
Proof of Proposition 2: First consider the rate region (43). Let $\varphi_{i}():. \mathbb{R}_{+}^{2} \rightarrow \mathbb{R}_{+}, i=1,2$, be two arbitrary deterministic functions satisfying $\mathbb{E}\left[\varphi_{i}(S)\right] \leq P_{i}, i=1,2$. Also, assume that $\varrho_{i}():. \mathbb{R}_{+}^{2} \rightarrow[0,1], i=1,2$, are two arbitrary deterministic functions. Define the functions $p_{i}():. \mathbb{R}_{+}^{2} \rightarrow \mathbb{R}_{+}, i=0,1,2$, and $\varrho():. \mathbb{R}_{+}^{2} \rightarrow[0,1]$ as follows:

$$
\left\{\begin{array}{l}
p_{0}\left(H=\left(S_{1}^{2}, S_{2}^{2}\right)\right) \triangleq\left(S_{1} \varrho_{1}(S) \sqrt{\varphi_{1}(S)}+S_{2} \varrho_{2}(S) \sqrt{\varphi_{2}(S)}\right)^{2}  \tag{47}\\
p_{1}\left(H=\left(S_{1}^{2}, S_{2}^{2}\right)\right) \triangleq\left(S_{1}\right)^{2} \varphi_{1}(S)\left(1-\varrho_{1}^{2}(S)\right) \\
p_{2}\left(H=\left(S_{1}^{2}, S_{2}^{2}\right)\right) \triangleq\left(S_{2}\right)^{2} \varphi_{2}(S)\left(1-\varrho_{2}^{2}(S)\right) \\
\varrho\left(H=\left(S_{1}^{2}, S_{2}^{2}\right)\right) \triangleq \frac{S_{1} \varrho_{1}(S) \sqrt{\varphi_{1}(S)}}{S_{1} \varrho_{1}(S) \sqrt{\varphi_{1}(S)}+S_{2} \varrho_{2}(S) \sqrt{\varphi_{2}(S)}}
\end{array}\right.
$$

Note that $S=\left(S_{1}, S_{2}\right)=\left(\sqrt{S_{1}^{2}}, \sqrt{S_{2}^{2}}\right)$. It can be easily verified that the functions $p_{i}(),. i=0,1,2$, and $\varrho($.$) defined by (47) satisfy$ (46). Now by substituting $p_{i}(),. i=0,1,2$, and $\varrho($.$) as defined by (47), in (45) we obtain that the rate region (43) is a subset of$ (45). Conversely, consider the rate region (45). Let $p_{i}():. \mathbb{R}_{+}^{2} \rightarrow \mathbb{R}_{+}, i=0,1,2$, and $\varrho():. \mathbb{R}_{+}^{2} \rightarrow[0,1]$ be arbitrary deterministic functions that satisfy (46). Define the functions $\varphi_{i}():. \mathbb{R}_{+}^{2} \rightarrow \mathbb{R}_{+}, i=1,2$, and $\varrho_{i}():. \mathbb{R}_{+}^{2} \rightarrow[0,1], i=1,2$, as follows:

$$
\left\{\begin{array}{l}
\varphi_{1}\left(S=\left(S_{1}, S_{2}\right)\right) \triangleq \frac{p_{1}\left(S_{1}^{2}, S_{2}^{2}\right)}{S_{1}^{2}}+\frac{\varrho\left(S_{1}^{2}, S_{2}^{2}\right)^{2} p_{0}\left(S_{1}^{2}, S_{2}^{2}\right)}{S_{1}^{2}}  \tag{48}\\
\varphi_{2}\left(S=\left(S_{1}, S_{2}\right)\right) \triangleq \frac{p_{2}\left(S_{1}^{2}, S_{2}^{2}\right)}{S_{2}^{2}}+\frac{\left(1-\varrho\left(S_{1}^{2}, S_{2}^{2}\right)\right)^{2} p_{0}\left(S_{1}^{2}, S_{2}^{2}\right)}{S_{2}^{2}} \\
\varrho_{1}\left(S=\left(S_{1}, S_{2}\right)\right) \triangleq \frac{\varrho\left(S_{1}^{2}, S_{2}^{2}\right) \sqrt{p_{0}\left(S_{1}^{2}, S_{2}^{2}\right)}}{\sqrt{p_{1}\left(S_{1}^{2}, S_{2}^{2}\right)+\varrho\left(S_{1}^{2}, S_{2}^{2}\right)^{2} p_{0}\left(S_{1}^{2}, S_{2}^{2}\right)}} \\
\varrho_{2}\left(S=\left(S_{1}, S_{2}\right)\right) \triangleq \frac{\left(1-\varrho\left(S_{1}^{2}, S_{2}^{2}\right)\right) \sqrt{p_{0}\left(S_{1}^{2}, S_{2}^{2}\right)}}{\sqrt{p_{2}\left(S_{1}^{2}, S_{2}^{2}\right)+\left(1-\varrho\left(S_{1}^{2}, S_{2}^{2}\right)\right)^{2} p_{0}\left(S_{1}^{2}, S_{2}^{2}\right)}}
\end{array}\right.
$$

From (46), it is easily seen that $\mathbb{E}\left[\varphi_{i}(S)\right] \leq P_{i}, i=1,2$. Thus, we can substitute $\varphi_{i}():. \mathbb{R}_{+}^{2} \rightarrow \mathbb{R}_{+}, i=1,2$, and $\varrho_{i}():. \mathbb{R}_{+}^{2} \rightarrow$ $[0,1], i=1,2$, as defined by (48), in (43) and then find (45), verifying that (45) is a subset of (43). This completes the proof.

In fact, Theorem 1 (specialized for the case of a two-transmitter/one-receiver MAC with perfect CSI at both transmitters and the receiver) together with Proposition 2 result in an alternative proof of the capacity region for the Gaussian fading MAC with a common message studied in [18].

## III-B) Two-transmitter one-receiver multiple access channel with conferencing encoders

In this section, we consider the MAC with conferencing encoders. The capacity region of a discrete two-user MAC with conferencing encoders was obtained in [7]. In the sequence, we first state the result of Willems [7] and then derive the capacity region of a two-user Gaussian fading MAC with conferencing encoders.

Proposition 3 [7]: Consider a discrete two-user memoryless MAC with encoders connected by communication links of capacities $C_{12}$ and $C_{21}$, as depicted in Fig. 2. The capacity region denoted by $\mathcal{C}_{\text {conf }}^{D-M A C}\left(C_{12}, C_{21}\right)$, is given by:

$$
\mathcal{C}_{\text {conf }}^{\text {D-MAC }}\left(C_{12}, C_{21}\right)=\bigcup_{P_{U} P_{X_{1} \mid U}{ }^{P} X_{2} \mid U}\left\{\begin{array}{l}
\left(R_{1}, R_{2}\right) \in \mathbb{R}_{+}^{2}  \tag{49}\\
R_{1} \leq I\left(X_{1} ; Y \mid X_{2}, U\right)+C_{12} \\
R_{2} \leq I\left(X_{2} ; Y \mid X_{1}, U\right)+C_{21} \\
R_{1}+R_{2} \leq I\left(X_{1}, X_{2} ; Y \mid U\right)+C_{12}+C_{21} \\
R_{1}+R_{2} \leq I\left(X_{1}, X_{2} ; Y\right)
\end{array}\right\}
$$

To prove the achievability of the rate region in (49) for the two-user MAC with conferencing encoders, Willems [7] used a twouser discrete MAC with a common message. The capacity region of such a channel had already been obtained in [4] by Slepian and Wolf, (see also Remark 3 of Theorem 1). In the following theorem, we exploit the technique proposed by Willems [7] to obtain the capacity region of the two-user Gaussian fading MAC with conferencing encoders.

Theorem 2: The capacity region of the two-user Gaussian fading MAC (10) with conferencing encoders connected to each other with links of capacities $C_{12}$ and $C_{21}$, denoted by $\mathcal{C}_{\text {conf }}^{G f-M A C}\left(C_{12}, C_{21}\right)$, when the transmitters have access to partial CSIT and the receiver has access to perfect CSIR, is given by:

$$
\mathcal{C}_{\text {conf }}^{G f-M A C}\left(C_{12}, C_{21}\right)=\bigcup_{\substack{\varphi_{1}(.), \varphi_{2}(.)}}^{\varrho_{1}(.), e_{2}(.)}\left\{\begin{array}{l}
\left(R_{1}, R_{2}\right) \in \mathbb{R}_{+}^{2}  \tag{50}\\
R_{1} \leq \mathbb{E}_{S}\left[\mathbb{C}\left(\frac{\left(S_{1}\right)^{2} \varphi_{1}\left(E_{1}\right)\left(1-\varrho_{1}^{2}\left(E_{1}\right)\right)}{\sigma^{2}}\right)\right]+C_{12} \\
R_{2} \leq \mathbb{E}_{S}\left[\mathbb{C}\left(\frac{\left(S_{2}\right)^{2} \varphi_{2}\left(E_{2}\right)\left(1-\varrho_{2}^{2}\left(E_{2}\right)\right)}{\sigma^{2}}\right)\right]+C_{21} \\
R_{1}+R_{2} \leq \mathbb{E}_{S}\left[\mathbb{C}\left(\frac{\sum_{k=1}^{2}\left(S_{k}\right)^{2} \varphi_{k}\left(E_{k}\right)\left(1-\varrho_{k}^{2}\left(E_{k}\right)\right)}{\sigma^{2}}\right)\right]+C_{12}+C_{21} \\
R_{1}+R_{2} \leq \mathbb{E}_{S}\left[\mathbb{C}\left(\frac{\left(2 S_{1} S_{2} \varrho_{1}\left(E_{1}\right) \varrho_{2}\left(E_{2}\right) \sqrt{\varphi_{1}\left(E_{1}\right) \varphi_{2}\left(E_{2}\right)}\right.}{\sigma^{2}}\right)\right]
\end{array}\right\}
$$

where $S=\left(S_{1}, S_{2}\right) \in\left(\mathbb{R}^{+}\right)^{2}$ is a r.v. representing the state process of the channel and $E_{i} \in \mathcal{E}_{i}$, is a r.v. representing CSIT available at the $i^{\text {th }}$ transmitter, $i=1,2$, which is a deterministic function of $S$, i.e. $E_{i}=\xi_{i}(S)$. Furthermore, $\varphi_{i}():. \mathcal{E}_{i} \rightarrow \mathbb{R}^{+}$is the power allocation policy for the $i^{\text {th }}$ transmitter, that satisfies the following constraint:

$$
\begin{equation*}
\mathbb{E}\left[\varphi_{i}\left(E_{i}\right)\right] \leq P_{i}, \quad i=1,2 \tag{51}
\end{equation*}
$$

and $\varrho_{i}():. \varepsilon_{i} \rightarrow[0,1]$ is an arbitrary (limited) deterministic function that takes its values from the interval [ 0,1$]$.
Proof of Theorem 2: To prove the achievability of (50), we directly apply Willems' approach [7] for the discrete model to the Gaussian fading channel. Consider a block length- $n$ code with $2^{n R_{1}}$ messages $w_{1}$ for transmitter 1 and $2^{n R_{2}}$ messages $w_{2}$ for transmitter 2. Partition the set $\left\{1, \ldots, 2^{n R_{1}}\right\}$ into $2^{n R_{12}}$ cells each containing $2^{n\left(R_{1}-R_{12}\right)}$ elements. We label the cells $c_{1} \in$ $\left\{1, \ldots, 2^{n R_{12}}\right\}$ and the elements inside each cell $t_{1} \in\left\{1, \ldots, 2^{n\left(R_{1}-R_{12}\right)}\right\}$. Let $c_{1}\left(w_{1}\right)=s_{1}$, if $w_{1}$ is inside the cell $s_{1}$. A similar partitioning is done for the message set $\left\{1, \ldots, 2^{n R_{2}}\right\}$. We define $R_{12}=\min \left\{R_{1}, C_{12}\right\}$ and $R_{21}=\min \left\{R_{2}, C_{21}\right\}$. Now since $R_{12} \leq C_{12}$ and $R_{21} \leq C_{21}$, it is possible for encoder 1 to send $c_{1}\left(w_{1}\right) \in\left\{1, \ldots, 2^{n R_{12}}\right\}$ to encoder 2 and for encoder 2 to send $c_{2}\left(w_{2}\right) \in\left\{1, \ldots, 2^{n R_{21}}\right\}$ to encoder 1, by holding a ( $C_{12}, C_{21}$ )-permissible conference. Thus, one can define a common message for the encoders as:

$$
\dot{w}_{0}=\left(c_{1}\left(w_{1}\right), c_{2}\left(w_{2}\right)\right) .
$$

We also note that $t_{1}$ and $t_{2}$ are unknown to encoders 2 and 1 , respectively. Therefore, $\dot{w}_{0} \in\left\{1, \ldots, 2^{n\left(R_{12}+R_{21}\right)}\right\}$ can be considered as a common message and $t_{1} \in\left\{1, \ldots, 2^{n\left(R_{1}-R_{12}\right)}\right\}$ and $t_{2} \in\left\{1, \ldots, 2^{n\left(R_{2}-R_{21}\right)}\right\}$ as private messages. Now, from the result of Theorem 1 specialized for the case of a two-user Gaussian fading MAC with a common message, we conclude that ( $R_{1}, R_{2}$ ) is achievable for the two-user Gaussian fading MAC with conferencing encoders if the rates $R_{1}, R_{2}, R_{12}, R_{21}$, satisfy the following:

$$
\left\{\begin{array}{l}
0 \leq R_{1}-R_{12} \leq \mathbb{E}_{S}\left[\mathbb{C}\left(\frac{\left(S_{1}\right)^{2} \varphi_{1}\left(E_{1}\right)\left(1-\varrho_{1}^{2}\left(E_{1}\right)\right)}{\sigma^{2}}\right)\right]  \tag{52}\\
0 \leq R_{2}-R_{21} \leq \mathbb{E}_{S}\left[\mathbb{C}\left(\frac{\left(S_{2}\right)^{2} \varphi_{2}\left(E_{2}\right)\left(1-\varrho_{2}^{2}\left(E_{2}\right)\right)}{\sigma^{2}}\right)\right] \\
R_{1}+R_{2}-\left(R_{12}+R_{21}\right) \leq \mathbb{E}_{S}\left[\mathbb{C}\left(\frac{\Sigma_{k=1}^{2}\left(S_{k}\right)^{2} \varphi_{k}\left(E_{k}\right)\left(1-\varrho_{k}^{2}\left(E_{k}\right)\right)}{\sigma^{2}}\right)\right] \\
R_{1}+R_{2} \leq \mathbb{E}_{S}\left[\mathbb{C}\left(\frac{\left(2 S_{1} S_{2} \varrho_{1}\left(E_{1}\right) \varrho_{2}\left(E_{2}\right) \sqrt{\varphi_{1}\left(E_{1}\right) \varphi_{2}\left(E_{2}\right)}\right.}{\sigma^{2}}\right)\right]
\end{array}\right\}
$$

for some power allocation policy functions $\varphi_{i}(),. i=1,2$, satisfying (51) and also two arbitrary deterministic functions $\varrho_{i}: \varepsilon_{i} \rightarrow$ $[0,1], i=1,2$. On the other hand, by definition of $R_{12}$ and $R_{21}$, one can easily see that the rates $R_{1}, R_{2}, R_{12}, R_{21}$, satisfy (52), if and only if $\left(R_{1}, R_{2}\right) \in \mathcal{C}_{\text {conf }}^{G f-M A C}\left(C_{12}, C_{21}\right)$, where $\mathcal{C}_{\text {conf }}^{G f-M A C}\left(C_{12}, C_{21}\right)$ is given by (50). This proves the achievability of (50). For the converse part similar to Theorem 1, we first derive an outer bound on $\mathcal{C}_{\text {conf }}^{G f-M A C}\left(C_{12}, C_{21}\right)$ in terms of mutual information functions, in the following lemma.

Lemma 4: The capacity region of the two-user Gaussian fading MAC with conferencing encoders, i.e., $\mathcal{C}_{\text {conf }}^{G f-M A C}\left(C_{12}, C_{21}\right)$, is outer bounded by:

$$
\mathcal{C}_{\text {conf }}^{G f-M A C}\left(C_{12}, C_{21}\right) \subseteq \bigcup_{\substack{P_{U} P_{X_{1}}| | E_{1} P_{X_{2} \mid U E_{2}}  \tag{53}\\
\mathbb{E}\left[X_{i}^{2}\right] \leq P_{i}, i=1,2}}\left\{\begin{array}{l}
\left(R_{1}, R_{2}\right) \in \mathbb{R}_{+}^{2} \\
R_{1} \leq I\left(X_{1} ; Y \mid X_{2}, U, S\right)+C_{12} \\
R_{2} \leq I\left(X_{2} ; Y \mid X_{1}, U, S\right)+C_{12} \\
R_{1}+R_{2} \leq I\left(X_{1}, X_{2} ; Y \mid U, S\right)+C_{12}+C_{21} \\
R_{1}+R_{2} \leq I\left(X_{1}, X_{2} ; Y \mid S\right)
\end{array}\right\}
$$

where $S=\left(S_{1}, S_{2}\right) \in\left(\mathbb{R}^{+}\right)^{2}$ is a r.v. representing the state process of the channel and $E_{i} \in \mathcal{E}_{i}$, is a r.v. representing CSIT available at the $i^{\text {th }}$ transmitter, which is a deterministic function of $S: E_{i}=\xi_{i}(S), i=1,2$.

## Proof: See Appendix III.

Now, note that the mutual information functions in the rate region described by (53) and also the joint p.d.f. over which the union in (53) is taken are exactly the same as (28), (29), respectively, when specialized for the case of a two-user MAC. Thus, one can proceed as in proof of Theorem 1 to optimize the rate region given in (53) using suitable joint p.d.f.'s, and show that the rate region given in (53) is equivalent to $\mathcal{C}_{\text {conf }}^{G f-M A C}\left(C_{12}, C_{21}\right)$ given by (50). This completes the proof of Theorem 2 .

## IV. NUMERICAL EXAMPLES AND SIMULATIONS

In this section, we provide some numerical results for the Gaussian fading two transmitter/one receiver MACs studied in the previous section. We consider the Gaussian fading channel with no CSIT (the transmitters have no knowledge of CSI) and perfect CSIR. At first, we briefly review the results obtained in previous section under these assumptions, in the following.

Corollary: Consider a two-user Gaussian fading MAC (10) with no CSIT and perfect CSIR:
I) The capacity region of the channel with a common message, denoted by $\mathcal{C}_{c m}^{n o-C S I T}$, is given by:

$$
\mathcal{C}_{c m}^{\text {no-CSIT }}=\bigcup_{0 \leq \varrho_{1}, \varrho_{2} \leq 1}\left\{\begin{array}{l}
\left(R_{0}, R_{1}, R_{2}\right) \in \mathbb{R}_{+}^{3}  \tag{54}\\
R_{1} \leq \mathbb{E}_{S}\left[\mathbb{C}\left(\frac{\left(S_{1}\right)^{2} P_{1}\left(1-\varrho_{1}^{2}\right)}{\sigma^{2}}\right)\right] \\
R_{2} \leq \mathbb{E}_{S}\left[\mathbb{C}\left(\frac{\left(S_{2}\right)^{2} P_{2}\left(1-\varrho_{2}^{2}\right)}{\sigma^{2}}\right)\right] \\
R_{1}+R_{2} \leq \mathbb{E}_{S}\left[\mathbb{C}\left(\frac{\left(S_{1}\right)^{2} P_{1}\left(1-\varrho_{1}^{2}\right)+\left(S_{2}\right)^{2} P_{2}\left(1-\varrho_{2}^{2}\right)}{\sigma^{2}}\right)\right] \\
R_{0}+R_{1}+R_{2} \leq \mathbb{E}_{S}\left[\mathbb{C}\left(\frac{\left(S_{1}\right)^{2} P_{1}+\left(S_{2}\right)^{2} P_{2}+2 S_{1} S_{2} \varrho_{1} \varrho_{2} \sqrt{P_{1} P_{2}}}{\sigma^{2}}\right)\right]
\end{array}\right\}
$$

II) The capacity region of the channel model with connected encoders by links of capacities $C_{12}$ and $C_{21}$, denoted by $\mathcal{C}_{\text {conf }}^{\text {no-CSIT }}\left(C_{12}, C_{21}\right)$, is given by:

$$
\mathcal{C}_{\text {conf }}^{\text {no-CSIT }}\left(C_{12}, C_{21}\right)=\bigcup_{0 \leq \varrho_{1}, \varrho_{2} \leq 1}\left\{\begin{array}{l}
\left(R_{0}, R_{1}, R_{2}\right) \in \mathbb{R}_{+}^{3}  \tag{55}\\
R_{1} \leq \mathbb{E}_{S}\left[\mathbb{C}\left(\frac{\left(S_{1}\right)^{2} P_{1}\left(1-\varrho_{1}^{2}\right)}{\sigma^{2}}\right)\right]+C_{12} \\
R_{2} \leq \mathbb{E}_{S}\left[\mathbb{C}\left(\frac{\left(S_{2}\right)^{2} P_{2}\left(1-\varrho_{2}^{2}\right)}{\sigma^{2}}\right)\right]+C_{21} \\
R_{1}+R_{2} \leq \mathbb{E}_{S}\left[\mathbb{C}\left(\frac{\left(S_{1}\right)^{2} P_{1}\left(1-\varrho_{1}^{2}\right)+\left(S_{2}\right)^{2} P_{2}\left(1-\varrho_{2}^{2}\right)}{\sigma^{2}}\right)\right]+C_{12}+C_{21} \\
R_{1}+R_{2} \leq \mathbb{E}_{S}\left[\mathbb{C}\left(\frac{\left(S_{1}\right)^{2} P_{1}+\left(S_{2}\right)^{2} P_{2}+2 S_{1} S_{2} \varrho_{1} \varrho_{2} \sqrt{P_{1} P_{2}}}{\sigma^{2}}\right)\right]
\end{array}\right\}
$$

Proof: Part I is obtained from Theorem 1 by assuming a two-user MAC and also setting $E_{k} \equiv \emptyset, k=1,2$. For part II set $E_{k} \equiv$ $\emptyset, k=1,2$, in $\mathcal{C}_{\text {conf }}^{G f-M A C}\left(C_{12}, C_{21}\right)$ given by (50).

Now, we examine a few implications of our results for the channel model in (10), for a Rayleigh fading environment. Assume that the fading processes $\left\{S_{1, t}\right\}_{t=1}^{\infty}$ and $\left\{S_{2, t}\right\}_{t=1}^{\infty}$ in (10) are independent of each other and $S_{k, t}, k=1,2$, is a Rayleigh-distributed r.v. with a p.d.f. given by:

$$
\begin{equation*}
P_{S_{k}}(s)=2 s e^{-s^{2}}, \quad s \geq 0 \tag{56}
\end{equation*}
$$

In the following, our discussion is concerned to the systems with capacities given by (54) and (55), where the state process of the channel is described by (56).


Figure 3. The capacity region of the two-user Gaussian fading MAC with common message, $\mathcal{C}_{c m}^{n o-C S I T}$. The sum power for all plots is fixed and is equal to

$$
P_{1}+P_{2}=23.01 \mathrm{~dB}, \text { and also } P_{z}=0 \mathrm{~dB}
$$

We have plotted the capacity region of the two-user Gaussian MAC with common message, $\mathcal{C}_{c m}^{\text {no-CSIT }}$, in Fig. 3, under different values of the power ratio of transmitters $\frac{P_{1}}{P_{2}}$ (boundaries in each plane with solid lines). The sum power of the two transmitters, i.e., $P_{1}+P_{2}$ is fixed for all plots. As we see from Fig. 3, when both transmitters have the same power, i.e., $\frac{P_{1}}{P_{2}}=1$, the maximum rate of $R_{0}$ is attained. This can be justified as follows. It could be easily verified from (54) that the maximum achievable rate of the common message $R_{0}$ is given by:

$$
\begin{equation*}
\max _{R_{0} \in \mathcal{C}_{c m}^{o b}-C S I T} R_{0}=\mathbb{E}_{S}\left[\mathbb{C}\left(\frac{\left(S_{1}\right)^{2} P_{1}+\left(S_{2}\right)^{2} P_{2}+2 S_{1} s_{2} \sqrt{P_{1} P_{2}}}{\sigma^{2}}\right)\right] \tag{57}
\end{equation*}
$$

On the other hand, for fixed value of $P_{1}+P_{2}$, the geometric average $\sqrt{P_{1} P_{2}}$ is maximum when $P_{1}=P_{2}$, (note that $S_{1}$ and $S_{2}$ are positive-valued i.i.d. r.v.'s). Therefore, when the powers $P_{1}$ and $P_{2}$ become far from each other, the maximum achievable value of the rate $R_{0}$, i.e., the cut-off point on the $R_{0}$ axis, decreases. Note that for values of $\frac{P_{1}}{P_{2}}$ that are greater than 10 , the variation of the power $P_{1}$ is less than 10 percent, therefore the increment in the maximum achievable rate of $R_{1}$, i.e., the cut-off point on the $R_{1}$-axis, is not significant, as can be seen from Fig 3.


Figure 4. The capacity region of the two-user Gaussian fading MAC with conferencing encoders, $\mathcal{C}_{\text {conf }}^{\text {no-CSIT }}\left(C_{12}, C_{21}\right)$, for $C_{12}=C_{21} \triangleq C$, and $P_{1}=P_{2} \triangleq P=$ 20 dB .

In Fig. 4, the capacity region of the two-user Gaussian fading MAC with conferencing encoders, i.e., $\mathcal{C}_{\text {conf }}^{\text {no-CSIT }}\left(C_{12}, C_{21}\right)$ is plotted for the case $C_{12}=C_{21}=C$, and $P_{1}=P_{2}=P$ (boundaries with solid lines). When there is no conferencing, i.e. $C=0$, the capacity region forms a pentagon. It is clear from the figure that as $C$ increases beyond zero the capacity region enlarges. However, no further improvement is possible for values of $C$ such that $C \geq \mathbb{E}_{S}\left[\mathbb{C}\left(\frac{\left(S_{1}\right)^{2} P_{1}+\left(S_{2}\right)^{2} P_{2}+2 S_{1} S_{2} \sqrt{P_{1} P_{2}}}{\sigma^{2}}\right)\right]$, (in Fig. 4, for $C \geq 4.04(\mathrm{bps} / \mathrm{Hz}))$. In fact, for values of $C$ equal or greater than $\mathbb{E}_{S}\left[\mathbb{C}\left(\frac{\left(S_{1}\right)^{2} P_{1}+\left(S_{2}\right)^{2} P_{2}+2 S_{1} S_{2} \sqrt{P_{1} P_{2}}}{\sigma^{2}}\right)\right]$, the capacity region forms a triangle whose boundary lines are given by $R_{1}=0, R_{2}=0$, and also:

$$
\begin{equation*}
R_{1}+R_{2}=\mathbb{E}_{S}\left[\mathbb{C}\left(\frac{\left(S_{1}\right)^{2} P_{1}+\left(S_{2}\right)^{2} P_{2}+2 S_{1} S_{2} \sqrt{P_{1} P_{2}}}{\sigma^{2}}\right)\right] \tag{58}
\end{equation*}
$$

In this case, the conferencing rate $C$ is high enough such that each encoder can perfectly communicate the intended message to the other encoder by holding a conference before transmission, hence the channel reduces to a MAC in which both of the transmitters would send a same message $W=\left(W_{1}, W_{2}\right)$. Therefore, the capacity region is given by (58) which is the same as the maximum achievable rate of $R_{0}$ for the Gaussian fading MAC with common message given by (57), and no further improvement can be obtained by increasing the rate of conferencing.


Figure 5. The capacity region of the two-user Gaussian fading MAC with one encoder communicating to the other, $\mathcal{C}_{\text {conf }}^{\text {no-CSIT }}\left(0, C_{21}\right)$, for $P_{1}=23.01 \mathrm{~dB}$, $P_{2}=20 \mathrm{~dB}$.

Now consider the case where only one of the encoders is connected to the other via a link of given capacity. In Fig. 5, the capacity region of the two-user Gaussian fading MAC with one encoder communicating to the other, i.e., $\mathcal{C}_{\text {conf }}^{\text {no-CSIT }}\left(0, C_{21}\right)$, has been plotted. Therefore, in this situation the second encoder can communicate to the first encoder via a link of capacity $C_{21}$. It is clear from the figure that as the rate of conferencing $C_{21}$ increases, the capacity region enlarges, however, since only the second encoder can cooperate its message with the first one via conferencing, the maximum achievable rate of $R_{1}$ (the cut-off point at $R_{1}$-axis), does not change. In Fig. 5 it has been assumed that the associated power of the second encoder is half of that one for the first encoder, i.e. $P_{2}=\frac{1}{2} P_{1}$. This causes the capacity region with no conferencing $\left(C_{21}=0\right)$ to be an asymmetric region (the region with blue boundary in Fig. 5), such that the value of the cut-off point on the $R_{2}$-axis is less than that one on the $R_{1}$-axis. As we see from Fig. 5, by holding a conference with the rate of $C_{21}=0.47(\mathrm{bps} / \mathrm{Hz})$, (the region with green boundary), however, the second encoder can cooperate some part of its message to the first one, thus the maximum achievable rate of $R_{2}$ is the same as that one for $R_{1}$, i.e., the cut-off points on both axes are the same. This means that conferencing can compensate for the lack of power. In general, one can see that when the second transmitter is restricted by $P_{2}=\alpha P_{1}$ where $0<\alpha<1$, a conferencing link of capacity $C_{21}(\alpha)=\mathbb{E}_{S}\left[\mathbb{C}\left(\frac{\left(S_{1}\right)^{2} \bar{\alpha} P_{1}}{\sigma^{2}+\left(S_{2}\right)^{2} \alpha P_{2}}\right)\right]$ could compensate for the lack of power, depicted by $C_{21}(0.5)=0.47(\mathrm{bps} / \mathrm{Hz})$ in Fig. 5. Moreover, as in previous setup, for values of $C_{21}$ greater than $\mathbb{E}_{S}\left[\mathbb{C}\left(\frac{\left(S_{1}\right)^{2} P_{1}+\left(S_{2}\right)^{2} P_{2}+2 S_{1} S_{2} \sqrt{P_{1} P_{2}}}{\sigma^{2}}\right)\right]$, (in Fig. 5 for $C_{21} \geq 3.81$ $(\mathrm{bps} / \mathrm{Hz})$ ), no further improvement of the capacity region is possible. In fact, for values of $C_{21}$ equal or greater than $\mathbb{E}_{S}\left[\mathbb{C}\left(\frac{\left(S_{1}\right)^{2} P_{1}+\left(S_{2}\right)^{2} P_{2}+2 S_{1} S_{2} \sqrt{P_{1} P_{2}}}{\sigma^{2}}\right)\right]$, by holding a conference, the second transmitter can cooperate perfectly with the first one, and the channel acts as a MAC with degraded message sets: a common message is sent by both transmitters and a private message by the first transmitter. Therefore, the shape of the capacity region is similar to the capacity region of MAC with common message in ( $R_{0}, R_{1}$ )-plane, depicted in Fig .3.

## V. CONCLUSION

In this paper, we characterized the capacity region for some multiple-access scenarios with cooperative transmitters. Firstly, the general MAC with common information was considered. The capacity region of the discrete memoryless channel was characterized. Then, the general Gaussian fading MAC with common information with partial CSI at the transmitters and perfect CSI at the receivers was investigated. A coding theorem was proved for this model yielding an exact characterization of the throughput capacity region. Finally, the capacity region of a two-transmitter/one-receiver Gaussian fading MAC with conferencing encoders with partial CSIT and perfect CSIR was determined. For the Gaussian fading systems in a Rayleigh fading environment, some numerical examples and simulations were provided and the effect of conferencing on the improvement of the capacity region was explored. In future work, we will study the problem of optimal resource allocation for the Gaussian fading channels with partial CSIT and perfect CSIR that were considered in this paper.

## Acknowledgment

We would like to thank R. Bayat, M. Moghimi, the associate editor, and the anonymous reviewers for their helpful comments.

## Appendix I

## Proof of Theorem 1

To prove achievability of (12) a length $-n$ random codebook $\mathfrak{C}_{n}$ is constructed as follows:

1) Fix the distribution $P_{U X_{1} \ldots X_{\mathcal{p}}}\left(u, x_{1}, \ldots, x_{\mathfrak{p}}\right)=P_{U}(u) \prod_{i=1}^{p} P_{X_{i} \mid U}\left(x_{i} \mid u\right)$.
2) Generate randomly $2^{n R_{0}}$ i.i.d. sequences $u^{n}$ according to $P\left(u^{n}\right)=\prod_{t=1}^{n} P_{U}\left(u_{t}\right)$. Label these sequences $u^{n}\left(w_{0}\right)$, where $w_{0} \in\left\{1, \ldots, 2^{n R_{0}}\right\}$.
3) For each $u^{n}\left(w_{0}\right)$, generate $2^{n R_{i}}, i=\{1,2, \ldots, p\}$, sequences $x_{i}^{n}$ according to the p.d.f. $P\left(x_{i}^{n} \mid u^{n}\left(w_{0}\right)\right)=\prod_{t=1}^{n} P_{X_{i} \mid U}\left(x_{i, t} \mid u_{t}\right)$. Label these sequences $x_{i}^{n}\left(w_{0}, w_{i}\right)$, where $w_{i} \in\left\{1, \ldots, 2^{n R_{i}}\right\}$.

Encoding: To send a common message $w_{0}$ and a private message $w_{i}$, the $i^{\text {th }}$ encoder sends the codeword $x_{i}^{n}\left(w_{0}, w_{i}\right)$ over the channel.

Decoding at the $j^{\text {th }}$ receiver, $j \in\{1,2, \ldots, q\}$ : After receiving $y_{j}^{n}$, the $j^{\text {th }}$ receiver tries to find a unique $\left(\widehat{w}_{0}, \widehat{w}_{1}, \ldots, \widehat{w}_{p}\right)$ such that $\left(u^{n}\left(\widehat{w}_{0}\right), x_{1}^{n}\left(\widehat{w}_{0}, \widehat{w}_{1}\right), \ldots, x_{p}^{n}\left(\widehat{w}_{0}, \widehat{w}_{\mathfrak{p}}\right), y_{j}^{n}\right) \in A_{\epsilon}^{n}\left(P_{U X_{1} \ldots X_{p} Y_{j}}\right)$.

Analysis of probability of error: From (3) we have: $P_{e}^{\mathbb{C}_{n}} \leq P_{e, 1}^{\mathbb{C}_{n}}+\cdots+P_{e, q}^{\mathbb{C}_{n}}$, where $P_{e, j}^{\mathbb{C}_{n}}$ is given by (2). Now we compute each $P_{e, j}^{\mathbb{C}_{n}}$ (error probability at the $j^{t h}$ receiver). Define the events:

$$
E_{i_{0}, i_{1}, \ldots, i_{\mathfrak{p}}}^{(j)} \triangleq\left\{\left(U^{n}\left(i_{0}\right), X_{1}^{n}\left(i_{0}, i_{1}\right), \ldots, X_{\mathfrak{p}}^{n}\left(i_{0}, i_{\mathfrak{p}}\right), Y_{j}^{n}\right) \in A_{\epsilon}^{n}\left(P_{U X_{1} \ldots X_{\mathcal{p}} Y_{j}}\right)\right\}, \quad j=1, \ldots, q .
$$

By symmetry of the random code construction, the conditional probability of error does not depend on the choice of indices. Thus, the conditional probability of error is the same as the unconditional probability of error, such that without loss of generality we can assume that $\left(\bar{w}_{0}, \bar{w}_{1}, \ldots, \bar{w}_{\mathfrak{p}}\right)$ is sent. Then we have:

$$
\begin{aligned}
& P_{e, j}^{\mathbb{C}_{n}}=\operatorname{Pr}\left\{\left(E_{\bar{w}_{0}, \bar{w}_{1}, \ldots \bar{w}_{\mathcal{p}}}^{(j)}\right)^{c} \cup\left\{\bigcup_{\substack{\left(i_{0}, i_{1}, \ldots, i_{p}\right) \neq 7 \\
\left(\bar{w}_{0}, \bar{w}_{1}, \ldots, \bar{w}_{\mathcal{p}}\right)}} E_{i_{0}, i_{1}, \ldots, i_{\mathcal{p}}}^{(j)}\right\}\right\}
\end{aligned}
$$

from the Asymptotic Equipartition Property (AEP) theorem [19, Ch. 3], $\operatorname{Pr}\left\{\left(E_{\bar{w}_{0}, \bar{w}_{1}, \ldots, \bar{w}_{p}}^{(j)}\right)^{c}\right\} \rightarrow 0$ as $n \rightarrow \infty$ and $\epsilon \rightarrow 0$. Furthermore, from [19, Th. 15.2.1 and 15.2.3] one can see:

$$
\begin{equation*}
\operatorname{Pr}\left\{\bigcup_{\substack{\left(i_{0}, i_{1}, \ldots, i_{p}\right) \\ i_{0} \neq \bar{w}_{0}}} E_{i_{0}, i_{1}, \ldots, i_{p}}^{(j)}\right\} \rightarrow 0 \tag{60}
\end{equation*}
$$

Provided that $0 \leq \sum_{k=0}^{p} R_{k} \leq I\left(U, X_{1}, \ldots, X_{p} ; Y_{j}\right), n \rightarrow \infty$ and $\epsilon \rightarrow 0$. From this and by considering that $I\left(U, X_{1}, \ldots, X_{p} ; Y_{j}\right)=$ $I\left(X_{1}, \ldots, X_{p} ; Y_{j}\right)$ due to the Markov structure $U \rightarrow X_{1}, \ldots, X_{p} \rightarrow Y_{j}$, the last bound in (12) is obtained. Again, from [19, Th. 15.2.1 and 15.2.3] we can see :

Provided that $0 \leq \sum_{k \in \Lambda} R_{k} \leq I\left(\left\{X_{k}\right\}_{k \in \Lambda} ; Y_{j} \mid\left\{X_{k}\right\}_{k \notin \Lambda}, U\right), n \rightarrow \infty$ and $\epsilon \rightarrow 0$. The above bounds show that the average probability of error is arbitrarily small provided that $\left(R_{0}, R_{1}, R_{2}, \ldots, R_{p}\right)$ belongs to the rate region (12), as desired.

## Appendix II

## Proof of Lemma 1

Consider a length $n$ code with average error probability $P_{e}^{(n)}$, where $P_{e}^{(n)} \rightarrow 0$ as $n \rightarrow \infty$. Let $\Lambda$ be a subset of $\{1, \ldots, p\}$. Define new r.v.'s $\widetilde{U}_{t}, t=1, \ldots, n$, as: $\widetilde{U}_{t} \triangleq W_{0}$.

By Fano's inequality, for $j=1, \ldots, q$, we have:

$$
\begin{align*}
& n \sum_{k \in \Lambda} R_{k}=H\left(\left\{W_{k}\right\}_{k \in \Lambda}\right)=I\left(\left\{W_{k}\right\}_{k \in \Lambda} ; Y_{j}^{n}, S^{n}\right)+H\left(\left\{W_{k}\right\}_{k \in \Lambda} \mid Y_{j}^{n}, S^{n}\right) \\
& \leq I\left(\left\{W_{k}\right\}_{k \in \Lambda} ; Y_{j}^{n}, S^{n}\right)+n \epsilon_{n}^{(j, \Lambda)} \leq I\left(\left\{W_{k}\right\}_{k \in \Lambda} ; Y_{j}^{n}, S^{n}, W_{0},\left\{W_{k}\right\}_{k \notin \Lambda}\right)+n \epsilon_{n}^{(j, \Lambda)} \\
& = \\
& =I\left(\left\{W_{k}\right\}_{k \in \Lambda} ; Y_{j}^{n} \mid S^{n}, W_{0},\left\{W_{k}\right\}_{k \notin \Lambda}\right)+n \epsilon_{n}^{(j, \Lambda)} \\
& =\sum_{t=1}^{n} I\left(\left\{W_{k}\right\}_{k \in \Lambda} ; Y_{j, t} \mid S^{n}, W_{0},\left\{W_{k}\right\}_{k \notin \Lambda}, Y_{j}^{t-1}\right)+n \epsilon_{n}^{(j, \Lambda)} \\
& =\sum_{t=1}^{n}\left[\boldsymbol{h}\left(Y_{j, t} \mid S^{n}, W_{0},\left\{W_{k}\right\}_{k \notin \Lambda}, Y_{j}^{t-1}\right)-\boldsymbol{h}\left(Y_{j, t} \mid S^{n}, W_{0},\left\{W_{k}\right\}_{k \notin \Lambda},\left\{W_{k}\right\}_{k \in L}, Y_{j}^{t-1}\right)\right]+n \epsilon_{n}^{(j, \Lambda)} \\
& \stackrel{(a)}{=} \sum_{t=1}^{n}\left[\boldsymbol{h}\left(Y_{j, t} \mid S^{n}, W_{0},\left\{W_{k}\right\}_{k \notin \Lambda}, Y_{j}^{t-1},\left\{X_{k, t}\right\}_{k \notin \Lambda}\right)-\boldsymbol{h}\left(Y_{j, t} \mid S^{n}, W_{0},\left\{W_{k}\right\}_{k \notin \Lambda},\left\{W_{k}\right\}_{k \in \Lambda}, Y_{j}^{t-1},\left\{X_{k, t}\right\}_{k \notin \Lambda^{\prime}}\left\{X_{k, t}\right\}_{k \in \Lambda}\right)\right]+n \epsilon_{n}^{(j, \Lambda)} \\
& \text { (b) }  \tag{62}\\
& =\sum_{t=1}^{n}\left[\boldsymbol{h}\left(Y_{j, t} \mid \widetilde{U}_{t},\left\{X_{k, t}\right\}_{k \notin \Lambda^{\prime}}, S_{t}\right)-\boldsymbol{h}\left(Y_{j, t} \mid \widetilde{U}_{t},\left\{X_{k, t}\right\}_{k \notin \Lambda^{\prime}},\left\{X_{k, t}\right\}_{k \in \Lambda^{\prime}}, S_{t}\right)\right]+n \epsilon_{n}^{(j, \Lambda)} \\
& =\sum_{t=1}^{n} I\left(\left\{X_{k, t}\right\}_{k \in \Lambda} ; Y_{j, t} \mid \widetilde{U}_{t},\left\{X_{k, t}\right\}_{k \notin \Lambda}, S_{t}\right)+n \epsilon_{n}^{(j, \Lambda)}
\end{align*}
$$

where $\epsilon_{n}^{(j, \Lambda)} \rightarrow 0$ as $n \rightarrow \infty$. The equality in (a) is due to the fact that $X_{k, t}$ is a deterministic function of $\left(W_{0}, W_{k}, E_{k}=\xi_{k}\left(S_{t}\right)\right)$, the inequality in (b) holds because conditioning does not increase entropy and also $Y_{j, t}$ conditioned on $\left(\left\{X_{k, t}\right\}_{k=1}^{p}, S_{t}\right)$ is independent of $\left(S^{t-1}, S_{t}^{n}, W_{0},\left\{W_{k}\right\}_{k \notin \Lambda},\left\{W_{k}\right\}_{k \in \Lambda}, Y_{j}^{t-1}\right)$. For the sum rate (last bound in (28)), again by Fano’s inequality we have:
$n \sum_{k=0}^{p} R_{k}=H\left(W_{0}, W_{1}, \ldots, W_{\mathfrak{p}}\right)=I\left(W_{0}, W_{1}, \ldots, W_{\mathfrak{p}} ; Y_{j}^{n}, S^{n}\right)+H\left(W_{0}, W_{1}, \ldots, W_{\mathfrak{p}} \mid Y_{j}^{n}, S^{n}\right)$
$\leq I\left(W_{0}, W_{1}, \ldots, W_{p} ; Y_{j}^{n}, S^{n}\right)+n \epsilon_{n}^{(j)}$
$=\sum_{t=1}^{n} I\left(W_{0}, W_{1}, \ldots, W_{p} ; Y_{j, t} \mid Y_{j}^{t-1}, S^{n}\right)+n \epsilon_{n}^{(j)}$
$=\sum_{t=1}^{n}\left[\boldsymbol{h}\left(Y_{j, t} \mid Y_{j}^{t-1}, S^{n}\right)-\boldsymbol{h}\left(Y_{j, t} \mid W_{0}, W_{1}, \ldots, W_{p}, Y_{j}^{t-1}, S^{n}\right)\right]+n \epsilon_{n}^{(j)}$
$\stackrel{(a)}{=} \sum_{t=1}^{n}\left[\boldsymbol{h}\left(Y_{j, t} \mid Y_{j}^{t-1}, S^{n}\right)-\boldsymbol{h}\left(Y_{j, t} \mid W_{0}, W_{1}, \ldots, W_{p}, Y_{j}^{t-1}, S^{n}, X_{1, t}, \ldots, X_{p, t}\right)\right]+n \epsilon_{n}^{(j)}$
$\stackrel{(b)}{\leq} \sum_{t=1}^{n}\left[\boldsymbol{h}\left(Y_{j, t} \mid S_{t}\right)-\boldsymbol{h}\left(Y_{j, t} \mid X_{1, t}, \ldots, X_{p, t}, S_{t}\right)\right]+n \epsilon_{n}^{(j)}$
$=\sum_{t=1}^{n} I\left(X_{1, t}, \ldots, X_{p, t} ; Y_{j, t} \mid S_{t}\right)+n \epsilon_{n}^{(j)}$
where $\epsilon_{n}^{(j)} \rightarrow 0$ as $n \rightarrow \infty$. The equality in (a) is due to the fact that $\left(X_{1, t}, \ldots, X_{p, t}\right)$ is a deterministic function of $\left(W_{0}, W_{1}, \ldots, W_{p}, E_{1}, \ldots, E_{p}\right)$ and the inequality in (b) holds because conditioning does not increase entropy and also $\left(W_{0}, W_{1}, \ldots, W_{p}, Y_{j}^{t-1}, S^{t-1}, S_{t}^{n}\right) \rightarrow\left(S_{t}, X_{1, t}, \ldots, X_{p, t}\right) \rightarrow Y_{j, t}$ forms a Markov chain. Note that for $i=1, \ldots, p$, we have:

$$
\begin{equation*}
\operatorname{Pr}\left(x_{i, t} \mid w_{0}, e_{i, t},\left\{x_{k, t}\right\}_{k=1, k \neq i}^{p}\right)=\sum_{w_{i}} \operatorname{Pr}\left(x_{i, t}, w_{i} \mid w_{0}, e_{i, t},\left\{x_{k, t}\right\}_{k=1, k \neq i}^{p}\right)=\sum_{w_{i}} \operatorname{Pr}\left(w_{i}\right) \operatorname{Pr}\left(x_{i, t} \mid w_{0}, e_{i, t}, w_{i}\right) \tag{64}
\end{equation*}
$$

Thus, the following Markov relationship holds:

$$
\begin{equation*}
\left\{X_{k, t}\right\}_{k=1, k \neq i}^{\mathfrak{p}} \rightarrow\left(\widetilde{U}_{t}, E_{i, t}\right) \rightarrow X_{i, t}, \quad i=1, \ldots, p \tag{65}
\end{equation*}
$$

In this step, we introduce a time-sharing r.v. $Q$ uniformly distributed over the set $\{1, \ldots, n\}$ and independent of other r.v.'s. Define $X_{i} \triangleq X_{i, Q}$ and $E_{i} \triangleq E_{i, Q}$ for $i=1, \ldots, p, Y_{j} \triangleq Y_{j, Q}$ and $Z_{j} \triangleq Z_{j, Q}$ for $j=1, \ldots, p, S_{j i} \triangleq S_{j i, Q}$ for $i=1, \ldots, p, j=1, . ., q, \widetilde{U} \triangleq$ $\widetilde{U}_{Q}$, and also the state matrix $S \triangleq S_{Q}$. We Note that:

$$
\begin{equation*}
Y_{j}=\sum_{i=1}^{p} S_{j i} X_{i}+Z_{j}, \quad j=1, \ldots, q \tag{66}
\end{equation*}
$$

where $Z_{j}, j=1, \ldots, q$, is a zero mean Gaussian r.v. with variance $\sigma_{j}^{2}$, which is independent of $Q, X_{i}, i=1, \ldots, p$, and also the state matrix $S$. Therefore, from (62) and (63) we can write:

$$
\left\{\begin{array}{l}
\sum_{k \in \Lambda} R_{k} \leq I\left(\left\{X_{k}\right\}_{k \in \Lambda} ; Y_{j} \mid \widetilde{U}_{Q}, Q,\left\{X_{k}\right\}_{k \notin \Lambda}, S\right)+\epsilon_{n}^{(j, \Lambda)}  \tag{67}\\
\sum_{k=0}^{p} R_{k} \leq I\left(X_{1}, \ldots, X_{p} ; Y_{j} \mid S, Q\right)+\epsilon_{n}^{(j)} \leq I\left(Q, X_{1}, \ldots, X_{p} ; Y_{j} \mid S\right)+\epsilon_{n}^{(j)}=I\left(X_{1}, \ldots, X_{p} ; Y_{j} \mid S\right)+\epsilon_{n}^{(j)}
\end{array}\right.
$$

where the last equality is due to that $Q \rightarrow S, X_{1}, \ldots, X_{p} \rightarrow Y_{j}$ forms a Markov chain, as arises from (66). On the other hand, from the Markov relationships in (65), the following Markov chains result:

$$
\begin{equation*}
\left\{X_{k}\right\}_{k=1, k \neq i}^{p} \rightarrow\left(Q, \widetilde{U}, E_{i}\right) \rightarrow X_{i}, \quad i=1, \ldots, p \tag{68}
\end{equation*}
$$

Now by redefining $U \triangleq(Q, \widetilde{U})$ and by letting $n$ tends to infinity in (67), we obtain:

$$
\left\{\begin{array}{l}
\sum_{k \in \Lambda} R_{k} \leq I\left(\left\{X_{k}\right\}_{k \in \Lambda} ; Y_{j} \mid U,\left\{X_{k}\right\}_{k \notin \Lambda}, S\right)  \tag{69}\\
\sum_{k=0}^{p} R_{k} \leq I\left(X_{1}, \ldots, X_{p} ; Y_{j} \mid S\right)
\end{array}\right.
$$

where the joint p.d.f. of the r.v.'s $\left\{X_{i}\right\}_{i=1}^{p},\left\{E_{i}\right\}_{i=1}^{p}, U$ satisfies (29). Furthermore, by definition of the code, the transmitted codewords should satisfy the power constraints in (6). Thus, the bounds in (69) should be considered under the following power constraints:

$$
\begin{equation*}
\mathbb{E}\left[X_{i}^{2}\right] \leq P_{i}, \quad i=1, \ldots, p \tag{70}
\end{equation*}
$$

This completes the proof.

## Appendix III

## Proof of Lemma 4

Consider a length- $n$ code with average error probability $P_{e}^{(n)}$, for the two-user Gaussian fading MAC with conferencing encoders, such that $P_{e}^{(n)} \rightarrow 0$ as $n \rightarrow \infty$. Define new r.v.'s $\widetilde{U}_{t}, t=1, \ldots, n$, as: $\widetilde{U}_{t} \triangleq\left(V_{1}^{K}, V_{2}^{K}\right)$.

Using Fano's inequality we can write:
$n R_{1} \leq H\left(W_{1}\right)=I\left(W_{1} ; Y^{n}, S^{n}\right)+H\left(W_{1} \mid Y^{n}, S^{n}\right)$
$\leq I\left(W_{1} ; Y^{n}, S^{n}\right)+n \epsilon_{n}^{(1)} \leq I\left(W_{1} ; Y^{n}, S^{n}, W_{2}\right)+n \epsilon_{n}^{(1)}$
$=I\left(W_{1} ; Y^{n}, S^{n} \mid W_{2}\right)+n \epsilon_{n}^{(1)}$
$\leq I\left(W_{1} ; Y^{n}, S^{n}, V_{1}^{K}, V_{2}^{K} \mid W_{2}\right)+n \epsilon_{n}^{(1)}$
$=I\left(W_{1} ; Y^{n}, S^{n} \mid V_{1}^{K}, V_{2}^{K}, W_{2}\right)+I\left(W_{1} ; V_{1}^{K}, V_{2}^{K} \mid W_{2}\right)+n \epsilon_{n}^{(1)}$
where $\epsilon_{n}^{(1)} \rightarrow 0$ as $n \rightarrow 0$. Now, for the first term in (71) we have:
$I\left(W_{1} ; Y^{n}, S^{n} \mid V_{1}^{K}, V_{2}^{K}, W_{2}\right)=I\left(W_{1} ; Y^{n} \mid V_{1}^{K}, V_{2}^{K}, W_{2}, S^{n}\right)$
$=\sum_{t=1}^{n} I\left(W_{1} ; Y_{t} \mid V_{1}^{K}, V_{2}^{K}, W_{2}, S^{n}, Y^{t-1}\right)$
$=\sum_{t=1}^{n} \boldsymbol{h}\left(Y_{t} \mid V_{1}^{K}, V_{2}^{K}, W_{2}, S^{n}, Y^{t-1}\right)-\boldsymbol{h}\left(Y_{t} \mid V_{1}^{K}, V_{2}^{K}, W_{2}, S^{n}, Y^{t-1}, W_{1}\right)$
$\stackrel{(a)}{=} \sum_{t=1}^{n} \boldsymbol{h}\left(Y_{t} \mid V_{1}^{K}, V_{2}^{K}, W_{2}, S^{n}, Y^{t-1}, X_{2, t}\right)-\boldsymbol{h}\left(Y_{t} \mid V_{1}^{K}, V_{2}^{K}, W_{2}, S^{n}, Y^{t-1}, W_{1}, X_{1, t}, X_{2, t}\right)$
$\stackrel{(b)}{\leq} \sum_{t=1}^{n} \boldsymbol{h}\left(Y_{t} \mid \widetilde{U}_{t}, X_{2, t}, S_{t}\right)-\boldsymbol{h}\left(Y_{t} \mid \widetilde{U}_{t}, X_{1, t}, X_{2, t}, S_{t}\right)$
$=\sum_{t=1}^{n} I\left(X_{1, t} ; Y_{t} \mid \widetilde{U}_{t}, X_{2, t}, S_{t}\right)$
where the equality (a) is due to that $X_{2, t}$ is a deterministic function of $\left(W_{2}, V_{2}^{K}, E_{2, t}=\xi_{2}\left(S_{t}\right)\right)$, and the inequality (b) is because conditioning does not increase entropy and also $Y_{t}$ conditioned on $\left(X_{1, t}, X_{2, t}, S_{t}\right)$ is independent of $\left(V_{1}^{K}, V_{2}^{K}, W_{2}, S^{t-1}, S_{t+1}^{n}, Y^{t-1}, W_{1}\right)$. Furthermore, by exactly the same procedure as in [7, p. 443], we have:
$I\left(W_{1} ; V_{1}^{K}, V_{2}^{K} \mid W_{2}\right) \leq n C_{12}$

Combining (71)-(73), we obtain:
$n R_{1} \leq \sum_{t=1}^{n} I\left(X_{1, t} ; Y_{t} \mid \widetilde{U}_{t}, X_{2, t}, S_{t}\right)+n C_{12}+n \epsilon_{n}^{(1)}$

Similarly one can show that:
$n R_{2} \leq \sum_{t=1}^{n} I\left(X_{2, t} ; Y_{t} \mid \widetilde{U}_{t}, X_{1, t}, S_{t}\right)+n C_{21}+n \epsilon_{n}^{(2)}$
where $\epsilon_{n}^{(2)} \rightarrow 0$ as $n \rightarrow 0$. For the sum rate, we can write:
$n\left(R_{1}+R_{2}\right) \leq H\left(W_{1}, W_{2}\right)=I\left(W_{1}, W_{2} ; Y^{n}, S^{n}\right)+H\left(W_{1}, W_{2} \mid Y^{n}, S^{n}\right)$
$\leq I\left(W_{1}, W_{2} ; Y^{n}, S^{n}\right)+n \epsilon_{n}^{(3)}$
$\leq I\left(W_{1}, W_{2} ; Y^{n}, S^{n}, V_{1}^{K}, V_{2}^{K}\right)+n \epsilon_{n}^{(3)}$
$=I\left(W_{1}, W_{2} ; Y^{n}, S^{n} \mid V_{1}^{K}, V_{2}^{K}\right)+I\left(W_{1}, W_{2} ; V_{1}^{K}, V_{2}^{K}\right)+n \epsilon_{n}^{(3)}$
where $\epsilon_{n}^{(3)} \rightarrow 0$ as $n \rightarrow 0$. Now, observe that:
$I\left(W_{1}, W_{2} ; Y^{n}, S^{n} \mid V_{1}^{K}, V_{2}^{K}\right)=I\left(W_{1}, W_{2} ; Y^{n} \mid V_{1}^{K}, V_{2}^{K}, S^{n}\right)$
$=\sum_{t=1}^{n} I\left(W_{1}, W_{2} ; Y_{t} \mid V_{1}^{K}, V_{2}^{K}, S^{n}, Y^{t-1}\right)$
$=\sum_{t=1}^{n} \boldsymbol{h}\left(Y_{t} \mid V_{1}^{K}, V_{2}^{K}, S^{n}, Y^{t-1}\right)-\boldsymbol{h}\left(Y_{t} \mid V_{1}^{K}, V_{2}^{K}, S^{n}, Y^{t-1}, W_{1}, W_{2}\right)$
$=\sum_{t=1}^{n} \boldsymbol{h}\left(Y_{t} \mid V_{1}^{K}, V_{2}^{K}, S^{n}, Y^{t-1}\right)-\boldsymbol{h}\left(Y_{t} \mid V_{1}^{K}, V_{2}^{K}, S^{n}, Y^{t-1}, W_{1}, W_{2}, X_{1, t}, X_{2, t}\right)$
$\leq \sum_{t=1}^{n} \boldsymbol{h}\left(Y_{t} \mid \widetilde{U}_{t}, S_{t}\right)-\boldsymbol{h}\left(Y_{t} \mid \widetilde{U}_{t}, X_{1, t}, X_{2, t}, S_{t}\right)$
$=\sum_{t=1}^{n} I\left(X_{1, t}, X_{2, t} ; Y_{t} \mid \widetilde{U}_{t}, S_{t}\right)$
and again similar to [7, p. 443]
$I\left(W_{1}, W_{2} ; V_{1}^{K}, V_{2}^{K}\right) \leq n\left(C_{12}+C_{21}\right)$

Therefore, by combining (76)-(78) we obtain:
$n\left(R_{1}+R_{2}\right) \leq \sum_{t=1}^{n} I\left(X_{1, t}, X_{2, t} ; Y_{t} \mid \widetilde{U}_{t}, S_{t}\right)+n\left(C_{12}+C_{21}\right)+n \epsilon_{n}^{(3)}$

To derive the last bound on the sum rate in (53) we proceed as:

$$
\begin{align*}
& n\left(R_{1}+R_{2}\right) \leq I\left(W_{1}, W_{2} ; Y^{n}, S^{n}\right)+n \epsilon_{n}^{(3)} \\
& =\sum_{t=1}^{n} I\left(W_{1}, W_{2} ; Y_{t} \mid S^{n}, Y^{t-1}\right)+n \epsilon_{n}^{(3)} \\
& =\sum_{t=1}^{n} \boldsymbol{h}\left(Y_{t} \mid S^{n}, Y^{t-1}\right)-\boldsymbol{h}\left(Y_{t} \mid S^{n}, Y^{t-1}, W_{1}, W_{2}\right)+n \epsilon_{n}^{(3)} \\
& =\sum_{t=1}^{n} \boldsymbol{h}\left(Y_{t} \mid S^{n}, Y^{t-1}\right)-\boldsymbol{h}\left(Y_{t} \mid S^{n}, Y^{t-1}, W_{1}, W_{2}, X_{1, t}, X_{2, t}\right)+n \epsilon_{n}^{(3)} \\
& \leq \sum_{t=1}^{n} \boldsymbol{h}\left(Y_{t} \mid S_{t}\right)-\boldsymbol{h}\left(Y_{t} \mid X_{1, t}, X_{2, t}, S_{t}\right)+n \epsilon_{n}^{(3)} \\
& =\sum_{t=1}^{n} I\left(X_{1, t}, X_{2, t} ; Y_{t} \mid S_{t}\right)+n \epsilon_{n}^{(3)} \tag{80}
\end{align*}
$$

By combining (74), (75), (79) and (80), applying a time-sharing argument similar to Appendix I, and then letting $n$ tends to infinity, the bounds in (53) are derived, where the fact that the underlying signaling satisfies the power constraints in (11) should also be considered. Finally, considering the code construction defined in Section II, one can verify that the following Markov relationships hold:

$$
\begin{align*}
& X_{2, t} \rightarrow\left(\widetilde{U}_{t}, E_{1, t}\right) \rightarrow X_{1, t} \\
& X_{1, t} \rightarrow\left(\widetilde{U}_{t}, E_{2, t}\right) \rightarrow X_{2, t} \tag{81}
\end{align*}
$$

The above Markov chains justify the distribution over which the union in (53) is taken. This completes the proof.

## References

[1] C. E. Shannon, "Two-way comunication channels," in Proceedings of $4{ }^{\text {th }}$ Berkely Symposium on Mathematical Statistics and Probability, (J. Neyman, ed.), pp.611-644, Berkeley, CA: University California Press, 1961.
[2] R. Ahlswede, "Multi-way comunication channels," in Proceedings of $2^{\text {nd }}$ International Symposium on Information Theory (1971), pp. 23-52, Tsahkadsor, Armenian S.S.R.: Publishing House of the Hungarian Academy of Sciences, 1973.
[3] H. Liao, "A coding theorem for multiple access communications," in Proceedings of IEEE International Symposium on Information Theory, Asilomar, CA, 1972.
[4] D. Slepian and J. K. Wolf, "A coding theorem for multiple access channels with correlated sources," Bell Syst. Tech. J., vol. 52, pp 1037-1076, 1973.
[5] F. M. J. Willems, "Information-Theoretical Results for the Discrete Memoryless Multiple Access Channel," Ph.D. dissertation, Katholieke Universiteit Leuven, Haverlee, Belgium, 1982.
[6] M. A. Wigger, Cooperation on the Multiple-Access Channel, A. Lapidoth, Ed. Konstanz, Switzerland: Hartung-Gorre Verlag, 2008, vol. 3, ETH Series in Information Theory and its Applications.
[7] F. M. J.Willems, "The discrete memoryless multiple channel with partially cooperating encoders," IEEE Trans. Inf. Theory, vol. 29, no. 3, pp. 441-445, May 1983.
[8] S. I. Bross, A. Lapidoth, and M. A. Wigger, "The Gaussian MAC with conferencing encoders," in Proc. IEEE Int. Symp. Information Theory (ISIT 2008), Toronto, ON, Canada, Jul. 2008, pp. 2702-2706.
[9] O. Simeone, O. Somekh, G. Kramer, H. V. Poor, and S. Shamai (Shitz), "Three-user Gaussian multiple access channel with partially cooperating encoders," in Proc. Asilomar Conf. Signals, Systems and computers, Monterey, CA, Oct. 2008.
[10] I. Maric, R. Yates, and G. Kramer, "Capacity of interference channels with partial transmitter cooperation," IEEE Trans. Inf. Theory, vol. 53, no. 10, pp. 3536-3548, Oct. 2007.
[11] R. Dabora and S. Servetto, "Broadcast channels with cooperating decoders." IEEE Trans. Inf. Theory, vol. 52, no. 12, pp. 5438-5454, Dec. 2006.
[12] O. Simeone, D. Gündüz, H. Vincent Poor, A. J. Goldsmith and S. Shamai, "Compound Multiple-Access Channels With Partial Cooperation." IEEE Trans. Inf. Theory, vol. 55, no. 6, pp. 2425-2441, June 2009.
[13] R. Gallager, "An inequality on the capacity region of multiaccess fading channels," in Communications and Cryptography-Two Sides of One Tapestry. Boston, MA: Kluwer, 1994, pp. 129-139.
[14] S. Shamai and A. D. Wyner, "Information theoretic considerations for symmetric, cellular, multiple-access fading channels-Part I," IEEE Trans. Inform. Theory, vol. 43, pp. 1877-1894, Nov. 1997.
[15] D. Tse and S. Hanly, "Multiaccess fading channels-Parts I: Polymatroidal structure, optimal resource allocation and throughput capacities; Part II: Delaylimited capacities," IEEE Trans. Inform. Theory, vol. 44, pp. 2796-2831, Nov. 1998.
[16] M. L. Ulrey, "The capacity region of a channel with $s$ senders and $r$ receivers," IC 29, pp 185-203, 1975.
[17] A.Das and P.Narayan, "Capacity of time-varying multiple access channels with side information," IEEE Trans. Inform. Theory, vol. 48, pp. 4-25, Jan. 2002.
[18] N. Liu, and S. Ulukus, "Capacity Region and Optimum Power Control Strategies for Fading Gaussian Multiple Access Channels with Common Data," IEEE Trans. on Communications, vol. 54, no. 10, pp. 1815-1826, Oct. 2006.
[19] T. M. Cover and J. A. Thomas, Elements of Information Theory. USA: Wiley, 2006.
[20] I. Csiszar and J. Korner, Information Theory. Coding Theorems for Discrete Memoryless Systems. London, U.K.: Academic, 1981.
[21] A .Haghi, R. Khoravi-Farsani, M. R. Aref and F. Marvasti "The Capacity Region of Fading Multiple Access Channels with Cooperative Encoders and Partial CSIT." in Proc. IEEE Int. Symp. Information Theory (ISIT 2010), Austin, Texas, U.S.A., pp. 485-489, Jun. 2010.


[^0]:    A. Haghi was with the Information Systems and Security Lab. (ISSL), EE Dept., Sharif University of Technology, Tehran, Iran. He is now with the ECE Dept., University of Waterloo, Waterloo, Canada. (e-mail: ali_haghi@alum.sharif.ir). R. Khosravi-Farsani is a research scientist with the Advanced Communications Research Institute (ACRI), Sharif University of Technology, Tehran, Iran. (e-mail: reza_khosravi@alum.sharif.ir). M. R. Aref is the Director of ISSL-ACRI and a professor at the EE Dept, Sharif University of Technology, Tehran, Iran (e-mail: aref@sharif.ir). F. Marvasti is the Director of ACRI and a professor at the EE Dept., Sharif University of Technology, Tehran, Iran (e-mail: marvasti@sharif.ir).

