Cooperative Multiple Access Encoding with States Available at One Transmitter

Anelia Somekh-Baruch
Princeton University
Princeton, NJ, USA
anelia@princeton.edu

Shlomo Shamai (Shitz)
Technion I.I.T.
Haifa, Israel
sshlomo@ee.technion.ac.il

Sergio Verdú
Princeton University
Princeton, NJ, USA
verdu@princeton.edu

***Abstract*— We generalize the Gel'fand-Pinsker model to encompass the setup of a memoryless multiple-access channel. According to this setup, only one of the encoders knows the state of the channel (non-causally), which is also unknown to the receiver. Two independent messages are transmitted: a common message and a message transmitted by the informed encoder. We find explicit characterizations of the capacity region with both non-causal and causal state information. Further, we apply the general formula to the Gaussian case with non-causal channel state information, under an individual power constraint as well as a sum power constraint. In this case, the capacity region is achievable by a generalized writing-on-dirty-paper scheme.**

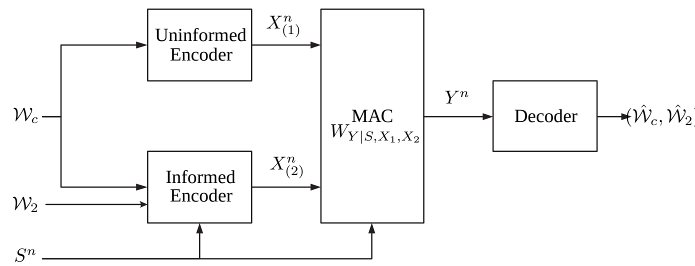

Fig. 1. Asymmetric state-dependent MAC with a common message.

I. Introduction

The capacity of state-dependent channels has become a widely investigated research area. The framework of channel states available at the transmitter dates back to Shannon [1], who characterized the capacity of a state-dependent memoryless channel whose states are i.i.d. and available causally to the transmitter. In their celebrated paper [2], Gel'fand and Pinsker (GP) established a single-letter formula for the capacity of the same channel under the conceptually different setup where the transmitter observes the channel states non-causally. The main tool to prove achievability in this setup is the binning encoding principle [2]. Costa [3] applied GP's result to the Gaussian case with two additive Gaussian noise sources, one of which, the interference, takes the role of the channel state. Costa originated the term "writing on dirty paper" which stands for an application of GP's binning encoding scheme that adapts the transmitted signal to the channel state sequence rather than attempting to cancel it. This results in a surprising conclusion: the capacity of the channel without interference can be attained even though the interference is not known to the receiver. For extensions of this principle and other related work see [4] and references therein.

Much research has been devoted to applications of these channel models, for example, watermarking, (see [5] and references therein), multi-input-multi-output (MIMO) broadcast channels, [6], where dirty-paper coding happens to be a central ingredient in achieving the capacity region, and cooperative networks [7].

In [4] we introduced the setup of cooperative encoding over the asymmetric (channel states known to one encoder only) GP MAC, with a single common message source fed to both encoders. We characterized the capacity of this channel both for the general finite input alphabet case and the Gaussian case. We also briefly discussed a generalization of this setup to include an additional message source that is fed to the informed user only. In this paper, we perform this generalization. The results presented in this paper have applications relating to cognitive radio (see [8] and references therein), watermarking, and other scenarios of cooperative communication, e.g., [9].

II. Problem Setup

A stationary memoryless state-dependent multiple-access channel is defined by a distribution Q_S on the set $\mathcal{S}$ and the channel conditional probability distribution $W_{Y|S,X_1,X_2}$ from $\mathcal{S} \times \mathcal{X}_1 \times \mathcal{X}_2$ to $\mathcal{Y}$. Let $X^n_{(1)} = (X_1(1), ..., X_1(n))$ and $X^n_{(2)} = (X_2(1), ..., X_2(n))$ designate the inputs of transmitters 1 and 2 to the channel, respectively. The output of the channel is denoted by Y^n. The symbols $S_i, X_1(i), X_2(i)$ and Y_i represent the channel state, the channel inputs produced by two distinct encoders, and the channel output, at time index i, respectively. We assume that the channel states S^n are i.i.d., each distributed according to Q_S. As can be seen in Figure 1, the setup we consider is asymmetric in the sense that only encoder 2 is informed of the channel states, while neither encoder 1 nor the decoder know the channel states. Unlike the ordinary MAC with partially known state information (which was considered in [10], [11], where inner and outer bounds on the capacity region are derived) we allow a common message source fed to both encoders and an independent message that is to be transmitted by the informed encoder. When encoder 2 observes the CSI *non-causally*, we shall refer to this channel as a Generalized Gel'fand-Pinsker (GGP) channel, when encoder

2 observes the states *causally*, the channel will be referred to as an asymmetric causal state–dependent channel.

The common message, $\mathcal{W}_c$, and the private message, $\mathcal{W}_2$, are independent random variables uniformly distributed over the sets $\{1, ..., M_c\}$ and $\{1, ..., M_2\}$, respectively, where $M_c = \lfloor e^{nR_c} \rfloor$ and $M_2 = \lfloor e^{nR_2} \rfloor$. An (e^{R_c}, e^{R_2}, n)-code for the GGP channel consists of two encoders $\varphi_n^{(1)}, \varphi_n^{(2)}$ and a decoder ψ_n: the first encoder, unaware of the CSI, is a mapping

$$\varphi_n^{(1)} : \{1, ..., M_c\} \rightarrow \mathcal{X}_1^n. \tag{1}$$

The second encoder observes the CSI non-causally and is defined by a mapping

$$\varphi_n^{(2)} : \{1, ..., M_c\} \times \{1, ..., M_2\} \times \mathcal{S}^n \rightarrow \mathcal{X}_2^n. \tag{2}$$

The decoder is a mapping

$$\psi_n : \mathcal{Y}^n \rightarrow \{1, ..., M_c\} \times \{1, ..., M_2\}. \tag{3}$$

An (e^{R_c}, e^{R_2}, n)-code for the asymmetric causal state-dependent channel is defined similarly with the exception that the second encoder is defined by a sequence of mappings

$$\varphi_{n,i}^{(2)} : \{1, ..., M_c\} \times \{1, ..., M_2\} \times \mathcal{S}^i \rightarrow \mathcal{X}_2 \tag{4}$$

where $i = 1, ..., n$, and $X_2(i) = \varphi_{n,i}^{(2)}(\mathcal{W}_c, \mathcal{W}_2, S^i)$.

An (ϵ, n, R_c, R_2)-code for the GGP channel is a code $(\varphi_n^{(1)}, \varphi_n^{(2)}, \psi_n)$ having average probability of error not exceeding ϵ, i.e., $\Pr((\mathcal{W}_c, \mathcal{W}_2) \neq \psi_n(Y_1^n)) \leq \epsilon$. A rate pair (R_c, R_2) is said to be achievable if there exists a sequence of $(\epsilon_n, n, R_c, R_2)$-codes with $\lim_{n\rightarrow\infty} \epsilon_n = 0$. The capacity region of the GGP channel is defined as the closure of the set of achievable rate-pairs. The definitions of an (ϵ, n, R_c, R_2)-code, an achievable rate pair and the capacity region of the asymmetric causal state-dependent channel are similar.

Due to space limitations, the proofs of the results of this paper are omitted and can be found in [12].

III. Capacity Region - Finite Input Alphabet GGP Channel

In this section it is assumed that the alphabets $\mathcal{S}, \mathcal{X}_1, \mathcal{X}_2$ are finite.

Theorem 1 *The capacity region of the GGP channel, $\mathcal{C}$, is the closure of the set of all rate–pairs, (R_c, R_2), satisfying*

$$\begin{aligned} R_2 &\leq I(U; Y|X_1) - I(U; S|X_1) \\ R_c + R_2 &\leq I(U, X_1; Y) - I(U, X_1; S), \end{aligned} \tag{5}$$

for some joint measure $P_{S,X_1,U,X_2,Y}$ having the form

$$P_{S,X_1,U,X_2,Y} = Q_S P_{X_1} P_{U|S,X_1} P_{X_2|S,X_1,U} W_{Y|S,X_1,X_2}, \tag{6}$$

where $|\mathcal{U}| \leq |\mathcal{S}| \cdot |\mathcal{X}_1| \cdot |\mathcal{X}_2|$.

Theorem 1 continues to hold if in (6) we replace $P_{X_2|S,X_1,U}$ by $P_{X_2|S,U}$ with slightly larger $|\mathcal{U}|$.

The proof of Theorem 1 is an immediate extension of the proof of Theorem 1 in [4]. In particular, the achievability part analyzes the error probability of a coding scheme described below (after Corollary 1).

It is noted that Theorem 1 remains intact if we allow for feedback to the informed encoder, i.e., if, before producing the i-th channel input symbol, the informed encoder observes the previous channel outputs, Y^{i-1}, that is, the informed encoder is actually a sequence of mappings $\varphi_n^{(2)} = \{\varphi_n^{(2,i)}\}_{i=1}^n$ with $\varphi_n^{(2,i)} : \{1, ..., M_c\} \times \{1, ..., M_2\} \times \mathcal{S}^n \times \mathcal{Y}^{i-1} \rightarrow \mathcal{X}_2$.

We now specialize Theorem 1 to the important case where only the common message is transmitted.

Corollary 1 *The common message capacity of the finite input alphabet GGP channel is given by*

$$C = \max\left[I(U, X_1; Y) - I(U, X_1; S)\right], \tag{7}$$

where the maximum is over all the joint measures $P_{S,X_1,U,X_2,Y}$ having the form (6) where $|\mathcal{U}| \leq |\mathcal{S}| \cdot |\mathcal{X}_1| \cdot |\mathcal{X}_2|$.

We note that there exists a maximizing measure with X_2 that is equal to a deterministic function of (S, X_1, U).

The achievability part of Theorem 1 is based on the following random coding scheme:

Generation of Codebooks:

Fix a measure $P_{S,X_1,U,X_2,Y}$ satisfying (6), and set

$$\begin{aligned} R_2^* &= I(U; Y|X_1) - I(U; S|X_1) \\ R_c^* &= I(X_1; Y) - I(X_1; S) = I(X_1; Y). \end{aligned} \tag{8}$$

Denote $M_1 = e^{n[R_c^* - \epsilon]}$, $M_2 = e^{n[R_2^* - \epsilon]}$, and $J = e^{n[I(U;S|X_1) + 2\epsilon]}$. First, M_1 i.i.d. n-vectors, $\{\mathbf{x}_\ell\}_{\ell=1}^{M_1}$, are drawn, each with i.i.d. components subject to P_{X_1}. The ordered collection of the drawn vectors constitutes the codebook used by the uninformed encoder.

For each codeword, $\mathbf{x}_\ell$, one draws $M_2 \times J$ auxiliary n-vectors denoted $\{\mathbf{u}_{\ell,k,j}\}$, $k = 1, ..., M_2$, $j = 1, ..., J$, independently and with i.i.d. components given $\mathbf{x}_\ell$ subject to the conditional measure $P_{U|X_1}$ induced by $\mathcal{Q}_S P_{X_1} P_{U,X_2|X_1,S}$. Hence, each codeword in the uninformed user's codebook is associated with a codebook of auxiliary codewords.

Encoding: To transmit ℓ, the uninformed encoder transmits the vector $\mathbf{x}_\ell$. Transmission of k is done by the informed encoder who searches for the lowest $j_0 \in \{1, ..., J\}$ such that $\mathbf{u}_{\ell,k,j_0}$ is jointly typical with $(\mathbf{x}_\ell, \mathbf{s})$. Denote this j by $j(\mathbf{s}, \ell, k)$. If such j_0 is not found, or if the observed state sequence $\mathbf{s}$ is non-typical, an error is declared and $j(\mathbf{s}, \ell, k)$ is set to $j = 1$.

Finally, the output of the second (informed) encoder is an n-vector $\tilde{\mathbf{x}}$ that is drawn i.i.d. conditionally given $(\mathbf{s}, \mathbf{u}_{\ell,k,j(\mathbf{s},\ell,k)}, \mathbf{x}_\ell)$ (using conditional measure $P_{X_2|S,X_1,U}$ induced by $\mathcal{Q}_S P_{X_1} P_{U,X_2|X_1,S}$).

Decoding: Upon observing $\mathbf{y}$, the decoder searches for a pair of indices, $(\hat{\ell}, \hat{k})$, such that $\mathbf{x}_{\hat{\ell}}, \mathbf{u}_{\hat{\ell},\hat{k},j}$ are jointly typical with $\mathbf{y}$ and outputs them. If there is no such pair, or it is not unique, an error is declared.

The analysis of the probability of error of the above described scheme (see [12]) establishes the achievability of (R_c^*, R_2^*). It is easy to realize that if a rate-pair (R_c, R_2)

is achievable, then so is $(R_c+R_2,0)$. This proves that also $(R_c^*+R_2^*,0)$ is achievable and thus the entire trapezoid (5) is an achievable region by time-sharing arguments.

When the common message capacity is concerned, the above encoding scheme can be applied by attributing the bits assigned to $\mathcal{W}_2$ in the above described scheme, to the common message $\mathcal{W}_c$. The output of decoder is the pair $\hat{m}=(\hat{\ell},\hat{k})$.

Although the capacity region of the GGP channel is characterized in Theorem 1, we next state an outer for it. The reason for that is that this outer bound is achievable in the Gaussian case, and hence is useful in the proof of the converse part of the Gaussian coding theorem. The outer bound is a generalization of the trivial upper bound $\max_{P_{X|S}} I(X;Y|S)$ on the capacity of the ordinary single-user GP channel.

Theorem 2 *The closure of the convex hull of the set of rate pairs satisfying*

$$\begin{aligned} R_2 &\leq I(X_2;Y|S,X_1) \\ R_c+R_2 &\leq I(X_1,X_2;Y|S)-I(S;X_1|Y) \end{aligned} \tag{9}$$

for some measure $P_{S,X_1,X_2,Y}=Q_S P_{X_1} P_{X_2|S,X_1} W_{Y|S,X_1,X_2}$ is an outer bound on the capacity region of the GGP channel.

A sub-class of GGP channels that will be of special interest is the following. A *memoryless parallel channel with non-causal asymmetric side information* is a GGP channel with $Y=(Y_1,Y_2)$ and $W_{Y_1,Y_2|S,X_1,X_2}=W_{Y_1|X_1,S}W_{Y_2|X_2,S}$. In words, this is a GGP channel with two outputs $Y_1(1),...,Y_1(n)$ and $Y_2(1),...,Y_2(n)$ that are both observed by the receiver. If, in addition, one has $W_{Y_2|X_2,S}=W_{Y_2|X_2}$ we shall say that the parallel channel is degenerate.

In the following theorem, we find the capacity region of the degenerate parallel GGP channel, for which we establish the fact that the CSI does not help.

Theorem 3 *The capacity region of the degenerate parallel GGP channel is equal to the capacity region obtained without transmitter CSI, i.e., the union of rate-pairs (R_c,R_2) s.t.*

$$\begin{aligned} R_2 &\leq C_2 \\ R_c+R_2 &\leq C_1+C_2, \end{aligned} \tag{10}$$

where C_1 is the capacity of the channel $W_{Y_1|X_1,S}$ obtained without transmitter CSI, and C_2 is the capacity of the channel $W_{Y_2|X_2}$.

IV. The Causal Asymmetric State-Dependent Channel

Theorem 4 *The capacity region of the finite input alphabet causal asymmetric state-dependent channel is given by the closure of the set of rate pairs (R_c,R_2) satisfying*

$$\begin{aligned} R_2 &\leq I(U;Y|X_1) \\ R_c+R_2 &\leq I(U,X_1;Y), \end{aligned} \tag{11}$$

for some joint measure $P_{S,X_1,U,X_2,Y}$ having the form

$$P_{S,X_1,U,X_2,Y}=Q_S P_{X_1,U} P_{X_2|S,X_1,U} W_{Y|S,X_1,X_2}, \tag{12}$$

where $\mathcal{U}|\leq|\mathcal{S}|\cdot|\mathcal{X}_1|\cdot|\mathcal{X}_2|+1$.

The expression for the capacity region of Theorem 4 can be interpreted as a special case of Theorem 1, where U is independent of S. Specializing Theorem 4 to the case where there is only a transmission of a common message, we get the following.

Corollary 2 *The common message capacity of the finite input alphabet causal asymmetric state-dependent channel is given by*

$$\max I(U,X_1;Y), \tag{13}$$

where the maximum is over all the joint measures $P_{S,X_1,U,X_2,Y}$ having the form (12) and $|\mathcal{U}|\leq|\mathcal{S}|\cdot|\mathcal{X}_1|\cdot|\mathcal{X}_2|+1$.

We note that an alternative expression for the common message capacity in the causal case is with $I(U;Y)$ replacing $I(U,X_1;Y)$ in (13) and where X_1 is deterministic given U with slightly larger $|\mathcal{U}|$. As a side note, we mention that if the CSI is available to both of the encoders, the single-letter expression for the common message capacity is deduced as a direct application Shannon's formula [1] for a channel with input alphabet $\mathcal{X}_1\times\mathcal{X}_2$.

V. The Gaussian GGP Channel

A. Channel Model

The Gaussian GGP channel is given by

$$Y_i = X_1(i)+X_2(i)+S_i+N_i. \tag{14}$$

As before, the message $\mathcal{W}_c$ is available to both encoders, and only the second encoder knows the realization of the interference S^n (non-causally), and the message $\mathcal{W}_2$ to be transmitted. The noise processes, S^n and N^n, are assumed to be zero-mean Gaussian i.i.d. with $E(S_i^2)=Q$ and $E(N_i^2)=N$. The process N^n is independent of $\left(X^n_{(1)},X^n_{(2)},S^n\right)$. Several power constraints can be considered: a) Individual power constraints: $\frac{1}{n}\sum_{i=1}^n X_1^2(i)\leq P_1$, $\frac{1}{n}\sum_{i=1}^n X_2^2(i)\leq P_2$. b) Sum power constraint: $\frac{1}{n}\sum_{i=1}^n X_1^2(i)+\frac{1}{n}\sum_{i=1}^n X_2^2(i)\leq P$. c) Total received power power constraint: $\frac{1}{n}\sum_{i=1}^n\left(X_1(i)+X_2(i)\right)^2\leq P$. When c) is concerned, it is evident that all the power should be assigned to the informed encoder, and the problem degenerates to the ordinary "dirty paper" Costa setup [3], where the informed transmitter can assign bits of the transmitted information to either $\mathcal{W}_c$ or $\mathcal{W}_2$. Consequently, the capacity region in this case is a triangle whose vertex (R_c,R_2) points are $(0,0)$, $\left(0,\frac{1}{2}\log\left(1+\frac{P}{N}\right)\right)$, and $\left(\frac{1}{2}\log\left(1+\frac{P}{N}\right),0\right)$.

We are interested in finding the capacity regions for the individual power constraints and the sum power constraint, denoted $\mathcal{C}(P_1,P_2,Q,N)$ and $\mathcal{C}(P,Q,N)$, respectively.

B. Capacity Region under Individual Power Constraints

First, we consider the Gaussian degenerate parallel channel with non-causal asymmetric CSI, which is a GGP channel whose i-th output is given by $Y_i=(Y_1(i),Y_2(i))$ with $Y_1(i)=$

$X_1(i)+S_i$ and $Y_2(i)=X_2(i)+N_i$, where $X_1(i), X_1(i), S_i, N_i$ are defined in (14).

Theorem 5 *The capacity region of the Gaussian degenerate parallel channel with non-causal asymmetric CSI under individual power constraints is given by the set of rate pairs (R_c, R_2) satisfying*

$$R_2 \leq \frac{1}{2}\log\left(1+\frac{P_2}{N}\right) \quad (15)$$

$$R_c + R_2 \leq \frac{1}{2}\log\left(1+\frac{P_1}{Q}\right)+\frac{1}{2}\log\left(1+\frac{P_2}{N}\right). \quad (16)$$

Theorem 5 obviously provides a simple outer bound on $\mathcal{C}(P_1,P_2,Q,N)$, as the decoder has more information, $(Y_1(i), Y_2(i))$ rather than $Y(i)=Y_1(i)+Y_2(i)$. In the sequel, we establish the tightness of this bound for a certain range of rates. The following theorem provides an explicit expression for $\mathcal{C}(P_1,P_2,Q,N)$.

Theorem 6 *Let* $\Delta_{min} = \min\left\{0, 1-\frac{P_1(P_2+N)^2}{P_2Q(P_1+Q)}\right\}$, *and let* $R(\Delta) = \max_{\rho\in[-\sqrt{1-\Delta},0]} \frac{1}{2}\log\left(1+\frac{\left(\sqrt{P_1}+\sqrt{1-\Delta-\rho^2}\sqrt{P_2}\right)^2}{P_2\Delta+\left(\sqrt{Q}+\rho\sqrt{P_2}\right)^2+N}\right)$. $\mathcal{C}(P_1,P_2,Q,N)$ *is equal to the union over $\Delta\in[\Delta_{min},1]$ of the rate pairs satisfying satisfying*

$$R_2 \leq \frac{1}{2}\log\left(1+\frac{P_2\Delta}{N}\right)$$
$$R_c+R_2 \leq \frac{1}{2}\log\left(1+\frac{P_2\Delta}{N}\right)+R(\Delta) \quad (17)$$

where if $\Delta_{min} > 0$, $R(\Delta_{min})$ is achieved by $\rho = -\frac{P_1(P_2+N)}{\sqrt{QP_2}(P_1+Q)}$, *and for $\Delta > \Delta_{min}$, it is achieved with either $\rho=-\sqrt{1-\Delta}$, $\rho=0$ or any real root of $g_\Delta(\rho)$ that satisfies $\rho\in[-\sqrt{1-\Delta},0]$, where*

$$\begin{aligned} g_\Delta(\rho) = &-P_2(P_1+Q)\rho^4 - 2\sqrt{QP_2}(P_2+Q+N+P_1)\rho^3 \\ &+ \left[P_2(1-\Delta)(P_1-2Q)-(P_2+Q+N)^2-P_1Q\right]\rho^2 \\ &+ 2\sqrt{P_2Q}(1-\Delta)\left[P_1-(P_2+Q+N)\right]\rho \\ &+ (1-\Delta)Q\left[P_1-P_2(1-\Delta)\right]. \end{aligned} \quad (18)$$

The proof is based on showing that for the Gaussian channel in (5), one can restrict attention to jointly Gaussian (S, X_1, X_2), and an optimal choice for U is $U = X_2+\alpha_{opt}S$

$$\text{with } \alpha_{opt} = \frac{P_2P_1Q - P_1\sigma_{2s}^2 - P_1N\sigma_{2s} - \sigma_{12}^2Q}{P_2P_1Q + P_1NQ - P_1\sigma_{2s}^2 - \sigma_{12}^2Q}, \quad (19)$$

where different values of $\sigma_{12}=E(X_1X_2)$ and $\sigma_{2s}=E(X_2S)$ are chosen to achieve different points that lie in (or, on the border of) the capacity region. The allowable values for the covariances, σ_{12} and σ_{2s} are such that the resulting covariance matrix of (X_1,X_2,S), satisfies the nonnegative-definiteness condition, expressed in terms of the correlation coefficients $\rho_{12}=\frac{\sigma_{12}}{\sqrt{P_1P_2}}$, $\rho_{12}=\frac{\sigma_{2s}}{\sqrt{P_2Q}}$

$$\rho_{12}^2+\rho_{2s}^2 \leq 1. \quad (20)$$

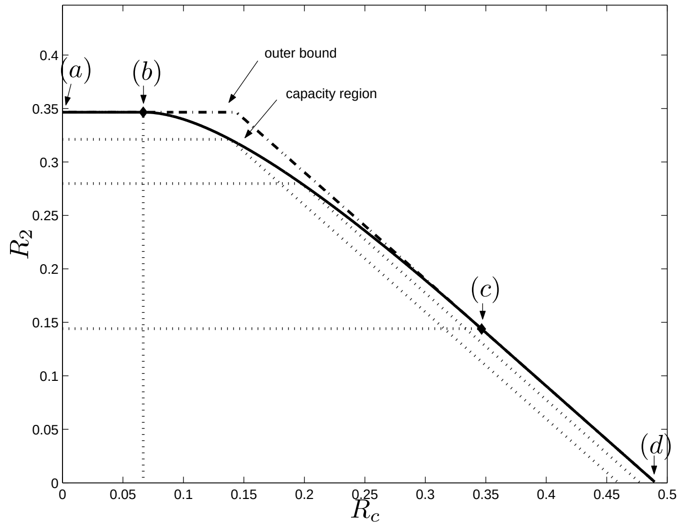

Fig. 2. $\mathcal{C}\left(\frac{1}{2},1,\frac{3}{2},1\right)$ and outer bound.

For reasons that will become clear in the sequel, we introduce the following terminology.

Definition 1 *The set of parameters P_1,P_2,Q,N such that $\frac{P_1(P_2+N)^2}{P_1+Q} \geq P_2Q$ will be referred to as the silent regime and its complement will be referred to as the active regime.*

The following proposition simplifies the capacity region expression for certain ranges of rates. It indicates a range of rates for which the outer bound given in Theorem 5 is tight.

Proposition 1 *1) For any P_1,P_2,Q,N, the segment connecting the following points (a) and (b) in the R_c-R_2 plane lies on the boundary of $\mathcal{C}(P_1,P_2,Q,N)$*

(a) $\quad (R_c,R_2) = \left(0, \frac{1}{2}\log\left(1+\frac{P_2}{N}\right)\right)$

(b) $\quad (R_c,R_2) = \left(\frac{1}{2}\log\left(1+\frac{P_1}{Q+P_2+N}\right), \frac{1}{2}\log\left(1+\frac{P_2}{N}\right)\right)$.

2) If P_1,P_2,Q,N lie in the active regime, the segment connecting the following points (c) and (d) also lies on the boundary of $\mathcal{C}(P_1,P_2,Q,N)$

(c) $\quad (R_c,R_2) = \left(\frac{1}{2}\log\left(\frac{(P_1+Q)^2}{A}\right), \frac{1}{2}\log\left(\frac{(P_2+N)A}{(P_1+Q)NQ}\right)\right)$,

(d) $\quad (R_c,R_2) = \left(\frac{1}{2}\log\left(1+\frac{P_1}{Q}\right)+\frac{1}{2}\log\left(1+\frac{P_2}{N}\right), 0\right)$.

where $A = Q(P_1+Q)-P_1(P_2+N)$.

The capacity region for $(P_1,P_2,Q,N)=(\frac{1}{2},1,\frac{3}{2},1)$ which lies in the active regime, as well as the points $(a),(b),(c),(d)$, and the outer bound of Corollary 5, are plotted in Figure 2. The dotted trapezoids express achievable regions attained by choosing specific Δ values in (17). Note that the segment connecting (c) and (d) meets the R_c axis at -45^o.

C. The Common Message Capacity with Individual Power Constraints

This subsection is devoted to specializing the results to the common message capacity, $C(P_1,P_2,Q,N)$.

Theorem 7 $C(P_1, P_2, Q, N) =$

$$\begin{cases} \frac{1}{2}\log\left(1+\frac{P_1}{Q}\right)+\frac{1}{2}\log\left(1+\frac{P_2}{N}\right) & \text{if } \frac{P_1(P_2+N)^2}{(P_1+Q)} \leq P_2 Q \\ \max_{\rho\in[-1,0]} \frac{1}{2}\log\left(1+\frac{\left(\sqrt{P_1}+\sqrt{P_2}\sqrt{1-\rho^2}\right)^2}{\left(\sqrt{Q}+\sqrt{P_2}\cdot\rho\right)^2+N}\right) & \text{otherwise} \end{cases} \quad (21)$$

where the maximization over ρ can be limited to either $\rho = -1$, $\rho = 0$ or a real root ρ of $g_0(\rho)$ (see (18)) s.t. $\rho \in [-1, 0]$.

Discussion: In the silent regime, the optimal values of σ_{12} and σ_{2s} as far as the $C(P_1, P_2, Q, N)$ is concerned, are such that inequality (20) is met with equality, i.e., $\rho_{12}^2 + \rho_{2s}^2 = 1$. This is equivalent to $X_2 = \frac{\sigma_{12}}{P_1}X_1 + \frac{\sigma_{2s}}{Q}S$. This implies that in the silent regime the optimal value of α (19) and U are $\alpha_{opt}^{silent} = -\frac{\sigma_{2s}}{Q}$, $U_{opt}^{silent} = X_2 - \frac{\sigma_{2s}}{Q}S = \frac{\sigma_{12}}{P_1}X_1$. It is easy to verify that the simpler selection of $U_{opt}^{silent} = 0$ yields the same achievable rate and hence is also optimal. Consequently, in the silent regime, in order to achieve $C(P_1, P_2, Q, N)$, the informed encoder can put all its power into decreasing the interference and enhancing the signal of the uninformed encoder. No power is devoted to the transmission of additional information, and hence, we refer to this region as silent. Since $U_{opt}^{silent} = 0$, no binning is needed. Given a message m to be transmitted, which corresponds to the codeword $\mathbf{x}_m = (x_m(1), x_m(2), ..., x_m(n))$ of the uninformed encoder, and a state-sequence $\mathbf{s}$, the informed encoder simply transmits at time index i, $\tilde{x}_i = x_m(i)\frac{\sigma_{12}}{P_1} + s_i\frac{\sigma_{2s}}{Q}$, where $\sigma_{2s} = \sqrt{P_2 Q}\cdot\rho$, with ρ being the maximizer in (21) and $\rho_{12} = \sqrt{1-\rho_{2s}^2}$.

In the active regime, the maximizing (X_1, X_2, S) is Gaussian with $\sigma_{12}^{active} = -\sigma_{2s}^{active} = \frac{P_1(P_2+N)}{P_1+Q}$. Thus, α_{opt} (see (19)) is given by $\alpha_{opt}^{active} = \frac{P_2}{P_2+N}$ which is equal to the optimal α in Costa's setup [3] when the uninformed user is not present. This results in a surprising phenomenon which happens only in the active regime. The highest achievable common message rate is $\frac{1}{2}\log\left(1+\frac{P_1}{Q}\right)+\frac{1}{2}\log\left(1+\frac{P_2}{N}\right)$, i.e., the upper bound on $C(P_1, P_2, Q, N)$ deduced from Theorem 5 is achievable.

D. Sum Power Constraints

Next, we state a closed form characterization of the capacity region under a sum power constraint. We denote by ζ the portion of the power that is used by the informed user.

Theorem 8 *The capacity region of the Gaussian GGP channel under sum power constraints, $\mathcal{C}(P, Q, N)$, is given by*

$$\mathcal{C}(P, Q, N) = \cup_{\zeta\in[0,1]}\mathcal{C}((1-\zeta)P, \zeta P, Q, N).$$

The following theorem gives the common message capacity under sum power constraints, $C(P, Q, N)$.

Theorem 9

$$C(P,Q,N) = \begin{cases} \frac{1}{2}\log\left(1+\frac{P}{N}\right) & \text{if } N+P \leq Q \\ \frac{1}{2}\log\frac{(Q+P+N)^2}{4QN} & \text{if } Q_0 \leq Q \leq N+P \\ \max_{0\leq\zeta\leq1} R(\zeta, P, Q, N) & \text{otherwise} \end{cases},$$

where $Q_0 = \frac{1}{3}(N-P) + \frac{2}{3}\sqrt{N^2+NP+P^2}$ and $R(\zeta, P, Q, N) =$

$$\max_{\rho\in[-1,0]} \frac{1}{2}\log\left(1+\frac{\left(\sqrt{(1-\zeta)P}+\sqrt{\zeta P}\sqrt{1-\rho^2}\right)^2}{\left(\sqrt{Q}+\sqrt{\zeta P}\cdot\rho\right)^2+N}\right).$$

The power allocation that achieves $C(P, Q, N)$ is

$$\zeta_{opt}(P,Q,N) = \begin{cases} 1 & \text{if } N+P \leq Q \\ \frac{P+Q-N}{2P} & \text{if } Q_0 \leq Q \leq N+P \\ \operatorname{argmax}_{0\leq\zeta\leq1} R(\zeta, P, Q, N) & \text{otherwise} \end{cases}.$$

Theorem 9 enables to simplify the expression for $\mathcal{C}(P, Q, N)$ as follows: (i) Whenever $Q \geq N + P$, the capacity region is not affected by Q, since the best strategy is to let the informed user use all the power. This degenerates to a single user Costa channel, where the transmitted information bits can be divided between $\mathcal{W}_2$ and $\mathcal{W}_c$. (ii) Whenever $Q_0 \leq Q \leq N + P$, the border of $\mathcal{C}(P, Q, N)$ contains the line segment between the points (R_c, R_2) given by $\left(\frac{(P+Q+N)^2}{4QN}, 0\right)$ and $\left(\frac{1}{2}\log\left(\frac{P+Q+N}{3Q-P-N}\right), \frac{1}{2}\log\left(\frac{(3Q-P-N)(P+Q+N)}{4QN}\right)\right)$.

ACKNOWLEDGMENT

This research was supported by a Marie Curie International Fellowship within the 6th European Community Framework Programme. The work of S. Shamai and S. Verdú has also been supported by the Binational US-Israel Scientific Foundation, Grant 2004140.

REFERENCES

[1] C. E. Shannon, "Channels with side information at the transmitter," *IBM J. Res. Develop.*, pp. 289–293, 1958.

[2] S. Gelfand and M. Pinsker, "Coding for channels with random parameters," *Problems of control and information theory*, vol. 9, no. 1, pp. 19–31, 1980.

[3] M. Costa, "Writing on dirty paper," *IEEE Transactions on Information Theory*, vol. IT-29, pp. 439–441, May 1983.

[4] A. Somekh-Baruch, S. Shamai (Shitz), and S. Verdú, "Cooperative encoding with asymmetric state information at the transmitters," in *Proc. Allerton Conf. Communications, Control, and Computing*, (Monticello, IL), September 2006.

[5] N. Merhav, "On joint coding for watermarking and encryption," *IEEE Transactions on Information Theory*, vol. 52, pp. 190–205, January 2006.

[6] G. Caire and S. Shamai (Shitz), "On the achievable throughput of a multi-antenna Gaussian broadcast channel," *IEEE Transactions on Information Theory*, vol. 49, pp. 1691–1706", July 2003.

[7] A. Høst-Madsen, "On the capacity of cooperative diversity in slow fading channels," in *Proc. Allerton Conf. Communications, Control, and Computing*, (Monticello, IL), October 2002.

[8] N. Devroye, P. Mitran, and V. Tarokh, "Limits on communications in a cognitive radio channel," *IEEE Communications Magazine*, vol. 44, pp. 44–49, June 2006.

[9] A. Høst-Madsen, "Capacity bounds for cooperative diversity," *IEEE Transactions on Information Theory*, vol. 52, pp. 1522–144, April 2006.

[10] S. Kotagiri and J. N. Laneman, "Achievable rates for multiple access channels with state information known at one encoder," in *Proc. Allerton Conf. Communications, Control, and Computing*, (Monticello, IL), October 2004.

[11] S. Kotagiri and J. N. Laneman, "Multiple access channels with state information known to some encoders." preprint, 2006.

[12] A. Somekh-Baruch, S. Shamai (Shitz), and S. Verdú, "Cooperative multiple access encoding with states available at one transmitter." submitted to *IEEE Transactions on Information Theory*, January 2007.