

On the Capacity of Compound State-Dependent Channels with States Known at the Transmitter

Pablo Piantanida

Department of Telecommunications, SUPELEC
 Plateau de Moulon, 91192 Gif-sur-Yvette, France
 Email: pablo.piantanida@supelec.fr

Shlomo Shamai (Shitz)

Department of Electrical Engineering, Technion - Israel Institute of Technology
 Technion city, Haifa 32000, Israel
 Email: sshlomo@ee.technion.ac.il

Abstract

This paper investigates the capacity of compound state-dependent channels with non-causal state information available at only the transmitter. A new lower bound on the capacity of this class of channels is derived. This bound is shown to be tight for the special case of compound channels with stochastic degraded components, yielding the full characterization of the capacity. Specific results are derived for the compound Gaussian Dirty-Paper (GDP) channel. This model consists of an additive white Gaussian noise (AWGN) channel corrupted by an additive Gaussian interfering signal, known at the transmitter only, where the input and the state signals are affected by fading coefficients whose realizations are unknown at the transmitter. Our bounds are shown to be tight for specific cases. Applications of these results arise in a variety of wireless scenarios as multicast channels, cognitive radio and problems with interference cancellation.

I. INTRODUCTION

In the recent years, intensive research addressing theoretical and practical aspects was undertaken on communications over channels controlled by random parameters, namely states. Gel'fand and Pinsker [1] derived the capacity expression for discrete memoryless channels (DMCs), where the i.i.d. state sequence is known at the transmitter before the start of the communication, but not at the receiver. This scenario

is known as state-dependent DMCs with non-causal state information. Costa [2] considered the case of an additive white Gaussian noise (AWGN) channel corrupted by an additive Gaussian interference which is available at the transmitter only. He showed that choosing an appropriate probability distribution (PD) for the auxiliary random variable (RV) and the state, referred to as *Dirty-Paper Coding* (DPC), there is no loss in capacity if the interference is known only to the encoder. This result has gained considerable attention because of its potential use to mitigate the interference effects in multi-user scenarios.

In this work we focus on the compound state-dependent channel with non-causal state information at the transmitter. This channel arises in scenarios where there is uncertainty on the channel statistic. In this model, the conditional PD of the channel is parameterized by θ , which belongs to an arbitrary set Θ and remains constant during the communication. Whereas, neither the sender nor the receiver are cognizant of the realization θ that governs the communication. This problem was initially investigated in [3], where lower and upper bounds on the capacity were derived. In [4], this problem is identified as being equivalent to the *common-message* broadcast channel (BC) with non-causal state information at the transmitter. Moreover in [5], this is recognized to be the *multicast channel*. Results were obtained for AWGN and binary channels, where a transmitter sends a common message to multiple receivers and each of them experiences an additive interference available at the transmitter only. These channels are of great interest because of their role in multi-user channels and in particular, for the emerging field of cognitive radios. Recent work in [6] investigated the capacity of this framework, which is essentially related to the problem considered here when the cognitive user is unaware of the channel path gains. Broadcast channels with imperfect channel knowledge are also instances of this class of channels (cf. [7] and [8]).

In prior work [9], [10], it was claimed that a strong converse establishes the optimality of the lower bound first derived in [3]. In this paper we will demonstrate that this is not the case in general. In fact, the rate expression (2) that was conjectured to be optimal for the general compound channel with states corresponds to the natural extension of the capacity expression obtained by Gel'fand and Pinsker's [1] to the compound setting case. Here we establish a new lower bound on the capacity of this class of channels that can outperform the previous lower bound. This bound is based on a non-conventional approach [11], [12] via a broadcasting strategy that allows the encoder to adapt the auxiliary RVs to each of possible channel outcomes (or each of different users in the multicast setting). Finally, we specialize this bound to the compound Gaussian Dirty-Paper (GDP) channel and derive an upper bound which is tight for some compound models. Furthermore, we show that our lower bound is tight for the compound channel with stochastic degraded components. Recent independent efforts deriving similar results are reported

in [13], where explicit examples demonstrate also that the rate in expression (2) can be surpassed. The organization of this paper is as follows. Definitions and main results are stated in Section II, while the proof outline and an application to the compound GDP channel are given in Sections III and IV.

II. PROBLEM STATEMENT AND MAIN RESULTS

In this section, we introduce main definitions, formalize the problem and present lower and upper bounds on the capacity.

A. Definitions and Problem Statement

We begin with the description of an arbitrary family of channels with discrete input $x \in \mathcal{X}$, discrete state $s \in \mathcal{S}$ and discrete output $y \in \mathcal{Y}$, which is characterized by a set of conditional probability distributions (PDs) $\mathcal{W}_\Theta = \{W_\theta : \mathcal{X} \times \mathcal{S} \mapsto \mathcal{Y}\}_{\theta \in \Theta}$, indexed by $\theta \in \Theta$, where Θ is the set of indexes (assumed to be finite). The transition PD of the n -memoryless extension with inputs $\mathbf{x} = (x_1, \dots, x_n)$, states $\mathbf{s} = (s_1, \dots, s_n)$ and outputs $\mathbf{y} = (y_1, \dots, y_n)$ is given by

$$W_\theta^n(\mathbf{y}|\mathbf{x}, \mathbf{s}) = \prod_{i=1}^n W_\theta(y_i|x_i, s_i). \quad (1)$$

The sequence of states \mathbf{s} is assumed to be drawn i.i.d. with PD P_S . The encoder is assumed to know the sequence of states before the transmission starts, but the decoder does not know it. Whereas, neither the sender nor the receiver are cognizant of the realization of θ that governs the communication. The channel states change from letter to letter following the PD P_S , but $\theta \in \Theta$ should not change during the communication. This scenario is known as compound DMCs with non-causal state information at the transmitter. We argue the capacity is not increased if the decoder is aware of the index $\theta \in \Theta$.

Definition 2.1 (Code): A code for this channel consists of two mappings, the encoder mapping $\{\varphi : \mathcal{M}_n \times \mathcal{S}^n \mapsto \mathcal{X}^n\}$ and the decoder mapping $\{\psi : \mathcal{Y}^n \mapsto \mathcal{M}_n\}$ for some finite set of integers $\mathcal{M}_n = \{1, \dots, M_n\}$. The encoding function $\{\varphi\}$ maps the corresponding message $m \in \mathcal{M}_n$ and the states \mathcal{S}^n into \mathcal{X}^n and the decoding function $\{\psi\}$ maps \mathcal{Y}^n into \mathcal{M}_n . In presence of feedback, where the past of the channel outputs are available at the transmitter, the encoder mappings are given by $\{\varphi_i : \mathcal{M}_n \times \mathcal{Y}^{i-1} \times \mathcal{S}^n \mapsto \mathcal{X}\}_{i=1}^n$. An n -length block code for simultaneous DMCs $\{W_\theta^n : \mathcal{X}^n \times \mathcal{S}^n \mapsto \mathcal{Y}^n, \theta \in \Theta\}_{n=1}^\infty$ consists on a common code (φ, ψ) for the set of channels $\mathcal{W}_\Theta^n = \{W_\theta^n\}_{\theta \in \Theta}$. The rate of such code is $n^{-1} \log M_n$ and its error probability associated to the message $m \in \mathcal{M}_n$ is defined as

$$e_m^{(n)}(W_\theta^n, \varphi, \psi|\mathbf{s}) = W_\theta^n\left(\bigcup_{m' \neq m} \psi^{-1}(m') \mid \varphi(m, \mathbf{s}), \mathbf{s}\right),$$

for $\theta \in \Theta$ and $\mathbf{s} \in \mathcal{S}^n$. The maximum of average error probability (over all messages) is defined as

$$\bar{e}_{\max}^{(n)}(\mathcal{W}_{\Theta}^n, \varphi, \psi) = \max_{m \in \mathcal{M}} \max_{\theta \in \Theta} \sum_{\mathbf{s} \in \mathcal{S}^n} P_S^n(\mathbf{s}) e_m^{(n)}(W_{\theta}^n, \varphi, \psi | \mathbf{s}).$$

An n -length block code for the simultaneous DMCs \mathcal{W}_{Θ}^n whose maximum of average error probability (2.1) satisfies $\bar{e}_{\max}^{(n)}(\mathcal{W}_{\Theta}^n, \varphi, \psi) \leq \epsilon$ will be called an (n, ϵ) -code.

Definition 2.2 (Achievable rate and capacity): Given $0 < \epsilon, \gamma < 1$, a non-negative number R is an ϵ -achievable rate for the compound channel \mathcal{W}_{Θ} if for every sufficiently large n there exist (n, ϵ) -codes of rate $n^{-1} \log M_n \geq R - \gamma$. Then, R is an achievable rate if it is ϵ -achievable for every $0 < \epsilon < 1$. The supremum of ϵ -achievable rates is called the ϵ -capacity C_{ϵ} while the supremum of achievable rates is called the capacity.

In the remainder of this section we state lower and upper bounds on the capacity of the general compound DMC (1).

B. Lower Bounds on the Capacity

The following achievable rate, first found in [3], corresponds to the straightforward extension of the Gel'fand and Pinsker's capacity [1] to the compound setting case.

Theorem 2.3: A lower bound on the capacity of the compound DMC $\{W_{\theta} : \mathcal{X} \times \mathcal{S} \mapsto \mathcal{Y}\}_{\theta \in \Theta}$ with states non-causally known only at the transmitter is given by

$$R = \sup_{P_{XU|S} \in \mathcal{Q}} \min_{\theta \in \Theta} \{I(U; Y_{\theta}) - I(U; S)\}, \quad (2)$$

where $U \oplus (X, S) \oplus Y_{\theta}$ for all $\theta \in \Theta$ and the set of admissible input PDs is defined as follows $\mathcal{Q} = \{P_{XU|S} \in \mathcal{P}(\mathcal{X} \times \mathcal{U}) : P_{XU|S} = P_{X|US}P_{U|S}, \|\mathcal{U}\| \leq \|\mathcal{X}\| + \|\mathcal{S}\| + \|\Theta\|\}$.

Notice that if the encoder is unaware of the states, i.e., (X, U) must be independent of S , expression (2) reduces to the capacity of standard compound DMCs [14].

We next state a new achievable rate that improves (2). For sake of clarity, we first consider the case of two components $\Theta = \{1, 2\}$ and then we generalize this to an arbitrary set Θ .

Theorem 2.4: A lower bound on the capacity of the compound DMC $\{W_1, W_2 : \mathcal{X} \times \mathcal{S} \mapsto \mathcal{Y}\}$ with states non-causally known only at the transmitter is given by

$$\begin{aligned} R = \sup \min \{ & I(U, V_1; Y_1) - I(U, V_1; S), \\ & I(U, V_2; Y_2) - I(U, V_2; S), \\ & \frac{1}{2} [I(U, V_1; Y_1) - I(U, V_1; S) + \\ & I(U, V_2; Y_2) - I(U, V_2; S) - I(V_1; V_2 | U, S)] \}, \end{aligned} \quad (3)$$

where the supremum is taken over the set of all joint PDs $P_{XUV_1V_2|S} = P_{X|UV_1V_2S}P_{UV_1V_2|S}$ that satisfy $(U, V_1, V_2) \ominus (X, S) \ominus (Y_1, Y_2)$ form a Markov chain.

Remark 2.5: Expression (3) can be reduced to (2) by setting $V_1 = V_2 = U$. In contrast, there is no possible choice for (X, U) in theorem 2.3 yielding the rate (3). This observation implies that expression (2) cannot be optimal for the general compound DMC (1). Furthermore, we shall see (section IV) that for the compound GDP channel the RVs (V_1, V_2) are indeed needed.

Theorem 2.6: A lower bound on the capacity of the compound DMC $\{W_\theta : \mathcal{X} \times \mathcal{S} \mapsto \mathcal{Y}\}_{\theta \in \Theta}$ with general components $\Theta = \{1, \dots, K\}$ and states non-causally known only at the transmitter is given by

$$R = \sup \min_{\mathcal{K} \subseteq \Theta} \frac{1}{\|\mathcal{K}\|} \left[\sum_{k \in \mathcal{K}} I(U, V_k; Y_k) - \|\mathcal{K}\| I(U; S) + H(\{V_t | t \in \mathcal{K}\} | U, S) - \sum_{k \in \mathcal{K}} H(V_k | U) \right], \quad (4)$$

where the supremum is taken over the set of all joint PDs $P_{XUV_1 \dots V_K | S} = P_{X|UV_1 \dots V_K S} P_{UV_1 \dots V_K | S}$ satisfying $(U, V_1, \dots, V_K) \ominus (X, S) \ominus (Y_1, \dots, Y_K)$ form a Markov chain.

Observe that the rate (4) reduces to the rate (3) for the case of $K = 2$. The proofs of these theorems are sketched in Section III. We next state capacity results for some special cases.

C. Capacity of Some Compound Channels

Definition 2.7 (Degraded components): Let $\{W_1, W_2 : \mathcal{X} \times \mathcal{S} \mapsto \mathcal{Y}\}$ be a compound DMC with components $\Theta = \{1, 2\}$. It is said to be a *stochastically degraded* [15] if there exists some stochastic mapping $\{\tilde{W} : \mathcal{Y} \mapsto \mathcal{Y}\}$ such that $W_2(y_2 | x, s) = \sum_{y_1 \in \mathcal{Y}} W_1(y_1 | x, s) \tilde{W}(y_2 | y_1)$, for all $y_2 \in \mathcal{Y}$ and every pair $(x, s) \in \mathcal{X} \times \mathcal{S}$. This shall be denoted by $W_2 \preceq W_1$ (i.e. W_2 is a degraded version of the channel W_1).

Theorem 2.8 (degraded components): The capacity of the compound DMC $\{W_\theta : \mathcal{X} \times \mathcal{S} \mapsto \mathcal{Y}\}_{\theta \in \Theta}$ with components $\Theta = \{1, \dots, K\}$ where $W_K \preceq W_{K-1} \preceq \dots \preceq W_1$ and states non-causally known only at the transmitter is given by

$$C_\Theta = \sup_{P_{XV_1 \dots V_K | S} \in \mathcal{Q}_D} \min_{\theta \in \Theta} \{I(V_\theta; Y_\theta) - I(V_\theta; S)\}, \quad (5)$$

where the set of admissible input PDs is defined by

$$\mathcal{Q}_D = \{P_{XV_1 \dots V_K | S} \in \mathcal{P}(\mathcal{X} \times \mathcal{V}_1 \times \dots \times \mathcal{V}_K) : P_{XV_1 \dots V_K | S} = P_{X|SV_1} P_{V_1|SV_2} \dots P_{V_{K-1}|SV_K} P_{V_K|S}, (V_1, \dots, V_K) \ominus (X, S) \ominus (Y_1, \dots, Y_K)\}.$$

Proof: For the case $\|\Theta\| = 2$, the direct proof follows by choosing $V_2 = U$ in theorem 2.4 and the converse proof follows by linking together the outputs (Y_1, \dots, Y_θ) . Whereas this proof procedure easily extends to an arbitrary set Θ . ■

Theorem 2.9 (feedback): The capacity of the compound DMC \mathcal{W}_Θ with states non-causally known only at the transmitter and feedback is given by

$$C_{\text{FB}} = \min_{\theta \in \Theta} \sup_{P_{XU_\theta|S} \in \mathcal{Q}} \{I(U_\theta; Y_\theta) - I(U_\theta; S)\}. \quad (6)$$

This theorem easily follows from [1] and by observing that the encoder is able to estimate the channel from the feedback.

III. SKETCH OF PROOF OF THEOREMS 2.4 AND 2.6

Notation: $\mathcal{P}(\mathcal{X})$ denotes the set of all atomic PDs on \mathcal{X} with finite number of atoms. The n -th Cartesian power is defined as the sample space of $\mathbf{X} = (X_1, \dots, X_n)$, with $P_{\mathbf{X}}^n$ -PD determined in terms of the n -th Cartesian power of P_X . The cardinality of an alphabet is denoted by $\|\cdot\|$. For every $\delta > 0$, we denote δ -typical and conditional δ -typical sets by $\mathcal{T}_{[X]_\delta}^n$ and $\mathcal{T}_{[Y|X]_\delta}^n(\mathbf{x})$, respectively.

Proof: We first provide details of the proof of theorem 2.4 where the general idea is as follows. We encode the message m into a RV U , by using superposition and Marton coding we allow partial (or indirect [12]) decoding of U via two other RVs, namely (V_1, V_2) . Hence receiver Y_1 indirectly decodes U via V_1 while receiver Y_2 indirectly decodes U via V_2 .

Code Generation: Let $T_0 \geq R$ and $S_i \geq 0$ with $i = 1, 2$. Fix a PD of the require form $P_{XUV_1V_2|S} = P_{X|UV_1V_2S}P_{UV_1V_2|S}$ satisfying $(U, V_1, V_2) \oplus (X, S) \oplus (Y_1, Y_2)$ form a Markov chain. Randomly and independently generate $\lfloor 2^{nT_0} \rfloor$ sequences $\mathbf{u}(t_0)$ form $\mathcal{T}_{[U]_\delta}^n$ indexed by $t_0 \in \{1, \dots, \lfloor 2^{nT_0} \rfloor\}$. Randomly partition the $\lfloor 2^{nT_0} \rfloor$ sequences into $\lfloor 2^{nR} \rfloor$ equal size bins. For each $\mathbf{u}(t_0)$, randomly and independent generate: (i) $\lfloor 2^{nT_1} \rfloor$ sequences $\mathbf{v}_1(t_0, t_1)$ indexed by $t_1 \in \{1, \dots, \lfloor 2^{nT_1} \rfloor\}$, each distributed uniformly over the set $\mathcal{T}_{[V_1|U]_\delta}^n(\mathbf{u}(t_0))$, (ii) $\lfloor 2^{nT_2} \rfloor$ sequences $\mathbf{v}_2(t_0, t_2)$ indexed by $t_2 \in \{1, \dots, \lfloor 2^{nT_2} \rfloor\}$, each distributed uniformly over the set $\mathcal{T}_{[V_2|U]_\delta}^n(\mathbf{u}(t_0))$.

Encoding: To send a message $m \in \{1, \dots, \lfloor 2^{nR} \rfloor\}$, choose an index $t_0^* \in \{1, \dots, \lfloor 2^{nT_0} \rfloor\}$ from the bin m such that $\mathbf{u}(t_0^*)$ and \mathbf{s} are jointly typical, and choose indices $t_1^* \in \{1, \dots, \lfloor 2^{nT_1} \rfloor\}$ and $t_2^* \in \{1, \dots, \lfloor 2^{nT_2} \rfloor\}$ such that $\mathbf{v}_1(t_0^*, t_1^*)$ and \mathbf{s} are jointly typical, $\mathbf{v}_2(t_0^*, t_2^*)$ and \mathbf{s} are jointly typical and the pair $(\mathbf{v}_1(t_0^*, t_1^*), \mathbf{v}_2(t_0^*, t_2^*))$ is jointly typical with high probability. Then send the codeword \mathbf{x} distributed uniformly over the set $\mathcal{T}_{[X|UV_1V_2S]_\delta}^n(\mathbf{u}(t_0^*), \mathbf{v}_1(t_0^*, t_1^*), \mathbf{v}_2(t_0^*, t_2^*), \mathbf{s})$. To ensure the success of this coding,

we require that

$$\begin{aligned}
T_0 - R &> S_0, \quad T_1 \geq S_1 \quad \text{and} \quad T_2 \geq S_2 \\
S_0 &> I(U; S), \\
S_1 + S_2 &> I(V_1; V_2|U) + I(V_1, V_2; S|U), \\
S_1 &> I(V_1; S|U), \quad S_2 > I(V_2; S|U).
\end{aligned} \tag{7}$$

Decoding: Receiver Y_1 finds t_0 and thus the message m via indirect decoding of $\mathbf{u}(t_0)$ based on $\mathbf{v}_1(t_0, t_1)$. Hence receiver Y_1 declares that $t_0 \in \{1, \dots, \lfloor 2^{nT_0} \rfloor\}$ is sent if it is the unique index such that $\mathbf{v}_1(t_0, t_1)$ and \mathbf{y}_1 are jointly typical $(\mathbf{u}(t_0), \mathbf{v}_1(t_0, t_1)) \in \mathcal{T}_{[UV_1]_\delta}^n$ for some $t_1 \in \{1, \dots, \lfloor 2^{nT_1} \rfloor\}$. This can be achieved with small probability of error provided

$$T_0 + T_1 < I(U, V_1; Y_1). \tag{8}$$

Notice that here receiver Y_1 cannot correctly decode $\mathbf{v}_1(t_0, t_1)$. Similarly, receiver Y_2 finds t_0 and thus the message m via indirect decoding of $\mathbf{u}(t_0)$ based on $\mathbf{v}_2(t_0, t_2)$. Hence receiver Y_2 declares that $t_0 \in \{1, \dots, \lfloor 2^{nT_0} \rfloor\}$ is sent if it is the unique index such that $\mathbf{v}_2(t_0, t_2)$ and \mathbf{y}_2 are jointly typical $(\mathbf{u}(t_0), \mathbf{v}_2(t_0, t_2)) \in \mathcal{T}_{[UV_2]_\delta}^n$ for some $t_2 \in \{1, \dots, \lfloor 2^{nT_2} \rfloor\}$. This can be achieved with small probability of error provided

$$T_0 + T_2 < I(U, V_2; Y_2). \tag{9}$$

Observe that receiver Y_2 cannot correctly decode $\mathbf{v}_2(t_0, t_2)$. By applying the Fourier-Motzkin procedure to eliminate $(T_i, S_i)_{\{i=0,1,2\}}$ from (7)-(9), we obtain the following inequalities:

$$\begin{aligned}
R &\leq I(U, V_1; Y_1) - I(U, V_1; S), \\
R &\leq I(U, V_2; Y_2) - I(U, V_2; S), \\
2R &\leq I(U, V_1; Y_1) + I(U, V_2; Y_2) - 2I(U; S) \\
&\quad - I(V_1; V_2|U) - I(V_1, V_2; S|U).
\end{aligned} \tag{10}$$

This concludes the proof of the rate (3). We now provide details on the proof of the extended rate (4). The code generation, encoding and decoding remain very similar to the previous. Encoding succeeds with high probability as long as

$$\begin{aligned}
T_0 - R &> S_0, \quad T_k \geq S_k \quad \text{for all } k = \{1, \dots, K\} \\
\sum_{k \in \mathcal{K}} S_k &> \sum_{k \in \mathcal{K}} H(V_k|U) - H(\{V_t | t \in \mathcal{K}\}|U, S),
\end{aligned} \tag{11}$$

for every subset $\mathcal{K} \subseteq \{1, \dots, K\}$ and $S_k \geq 0$ for all $k = \{1, \dots, K\}$. Decoding succeeds with high probability if

$$T_0 + T_k < I(U, V_k; Y_k), \quad k = \{1, \dots, K\}. \quad (12)$$

By combining expressions (11) and (12), applying Fourier-Motzkin procedure, it is not difficult to show the rate (4). ■

IV. APPLICATION EXAMPLE: COMPOUND GAUSSIAN DIRTY-PAPER CHANNEL

In this section, we begin by introducing the compound Gaussian Dirty-Paper (GDP) channel and then present lower and upper bounds on its capacity. Consider a compound memoryless GDP channel whose output is given by

$$\{\mathbf{Y}_{(\beta, \theta)} = \beta \cdot \mathbf{X} + \theta \cdot \mathbf{S} + \mathbf{Z}\}_{(\beta, \theta) \in \Theta}, \quad (13)$$

where $\mathbf{X} = (X_1, \dots, X_n)$ is the channel input and \mathbf{S} is a white Gaussian interference (known to the transmitter only) of power Q and independent of the sequence \mathbf{Z} of white Gaussian noise of power N . The inputs must satisfy a limited-power constraint P (often $\ll Q$), which takes the form $\mathbb{E} [\sum_{i=1}^n X_i^2(m, \mathbf{S})] \leq nP$, where the expectation is taken over the ensemble of messages and the interference sequence.

We focus on the case $\beta = \beta_0 = 1$, where the transmitter is unaware of $\theta \in \Theta$, assumed to take values from a set of real numbers, namely $\Theta \triangleq \{\theta_1, \dots, \theta_{\|\Theta\|}\}$. The fading coefficient θ remains fixed throughout a transmission.

A. Lower and Upper Bounds on the Capacity

Lemma 4.1 (Lower bound): A lower bound on the capacity of the compound GDP channel (13) is given by

$$C_{\Theta}(P) \geq \max_{(P_C, P_{\Delta}): P_C \geq 0, P_{\Delta} \geq 0, P_C + P_{\Delta} \leq P} R_{-}^{\Theta}(P_C, P_{\Delta}), \quad (14)$$

where $\theta_{\min} \triangleq \min\{\theta : \theta \in \Theta\}$, $\theta_{\max} \triangleq \max\{\theta : \theta \in \Theta\}$ and

$$\begin{aligned} R_{-}^{\Theta}(P_C, P_{\Delta}) &= \frac{1}{2\|\Theta\|} \log \left(1 + \frac{P_{\Delta}}{N} \right) \\ &+ \begin{cases} \frac{1}{2} \log \left(1 + \frac{P_C}{P_{\Delta} + N + \theta_{\min}^2 Q} \right) & \text{if } |\theta_{\min}| = |\theta_{\max}| \\ \frac{1}{2} \log \left[1 + \frac{P_C(1 - \epsilon_{\Theta})}{P_{\Delta} + N + \epsilon_{\Theta} P_C} \right] & \text{if } |\theta_{\min}| \neq |\theta_{\max}|, \end{cases} \end{aligned} \quad (15)$$

and the *mismatch factor* $0 \leq \epsilon_\Theta \leq 1$ is defined as

$$\epsilon_\Theta \triangleq \frac{1}{(\theta_{\min} + \theta_{\max})^2} \left[\sqrt{\theta_{\max}^2 + \frac{T}{Q}} - \sqrt{\theta_{\min}^2 + \frac{T}{Q}} \right]^2$$

with $T = P_C + P_\Delta + N$. Optimizing expression (15) over P_C and P_Δ subject to $P_C + P_\Delta \leq P$ yields the rate:

$$R_-^\Theta(P) = \begin{cases} \frac{1}{2} \log \left[1 + \frac{P(1 - \epsilon_\Theta^*)}{N + \epsilon_\Theta^* P} \right], & \text{if } \epsilon_\Theta^* < \frac{N(\|\Theta\| - 1)}{P + \|\Theta\|N} \\ \frac{1}{2\|\Theta\|} \log \left[\frac{(P + N)}{\|\Theta\|N(1 - \epsilon_\Theta^*)} \left(\frac{\|\Theta\| - 1}{\|\Theta\|\epsilon_\Theta^*} \right)^{\|\Theta\| - 1} \right], & \text{if } \frac{N(\|\Theta\| - 1)}{P + \|\Theta\|N} \leq \epsilon_\Theta^* < \frac{\|\Theta\| - 1}{\|\Theta\|} \\ \frac{1}{2\|\Theta\|} \log \left(1 + \frac{P}{N} \right), & \text{if } \epsilon_\Theta^* \geq \frac{\|\Theta\| - 1}{\|\Theta\|} \end{cases} \quad (16)$$

where

$$\epsilon_\Theta^* \triangleq \frac{1}{(\theta_{\min} + \theta_{\max})^2} \left[\sqrt{\theta_{\max}^2 + \frac{P + N}{Q}} - \sqrt{\theta_{\min}^2 + \frac{P + N}{Q}} \right]^2,$$

for $|\theta_{\min}| \neq |\theta_{\max}|$. For the case $|\theta_{\min}| = |\theta_{\max}|$,

$$R_-^\Theta(P) = \begin{cases} \frac{1}{2} \log \left(1 + \frac{P}{N + \theta_{\min}^2 Q} \right), & \text{if } \frac{\theta_{\min}^2}{(\|\Theta\| - 1)} < \frac{N}{Q} \\ \frac{1}{2\|\Theta\|} \log \left[\frac{(P + N + \theta_{\min}^2 Q)^{\|\Theta\|}}{\|\Theta\|N} \left(\frac{\|\Theta\| - 1}{\|\Theta\|\theta_{\min}^2 Q} \right)^{M-1} \right], & \text{if } \frac{N}{Q} \leq \frac{\theta_{\min}^2}{(\|\Theta\| - 1)} < \frac{(P + N)}{Q} \\ \frac{1}{2\|\Theta\|} \log \left(1 + \frac{P}{N} \right), & \text{if } \frac{\theta_{\min}^2}{(\|\Theta\| - 1)} \geq \frac{(P + N)}{Q}. \end{cases} \quad (17)$$

Lemma 4.2 (Upper bound): An upper bound on the capacity of the compound GDP channel is given by

$$C_\Theta(P) \leq R_+^\Theta(P) \triangleq \max_{\rho \in [-1, 1]} \min \left\{ \frac{1}{2} \log \left[1 + \frac{P(1 - \rho^2)}{N} \right], \right. \\ \left. \frac{1}{4} \log \left[\frac{P + N + \theta_{\max}^2 Q + 2\theta_{\max} \rho \sqrt{PQ}}{\sqrt{(\theta_{\max} - \theta_{\min})^2 NQ}} \right] \right. \\ \left. + \frac{1}{4} \log \left[\frac{P + N + \theta_{\min}^2 Q + 2\theta_{\min} \rho \sqrt{PQ}}{\sqrt{(\theta_{\max} - \theta_{\min})^2 NQ}} \right] \right\}, \quad (18)$$

with $\theta_{\min} \triangleq \min\{\theta : \theta \in \Theta\}$ and $\theta_{\max} \triangleq \max\{\theta : \theta \in \Theta\}$.

The proof of lemma 4.1 is sketched below while the proof of lemma 4.2 follows similar to [5]. Observe that the *mismatch factor* introduces the capacity loss due to the uncertainty at the encoder on the value

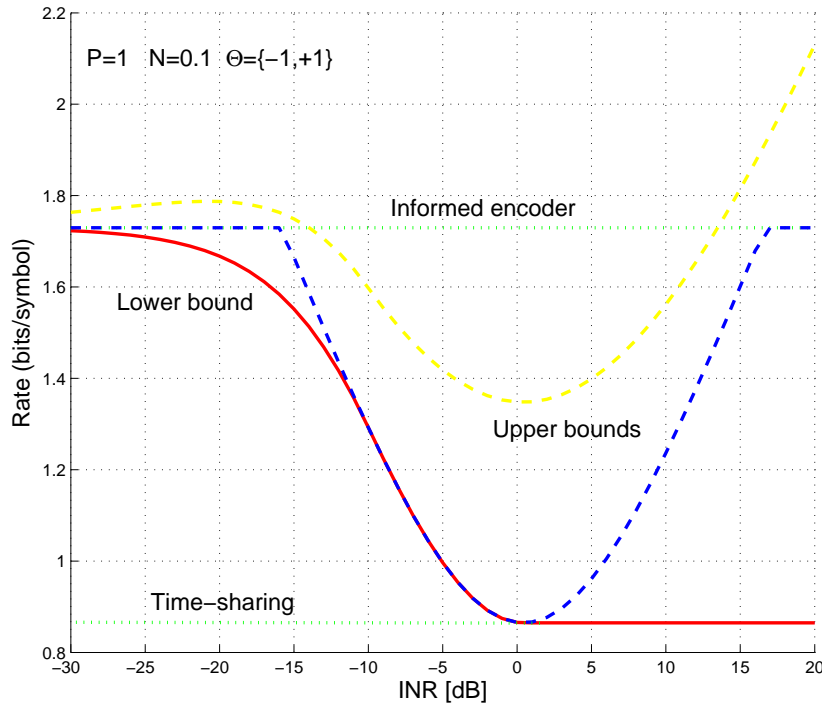


Fig. 1. Lower and upper bounds on the capacity of compound GDP channel $\Theta = \{-1, +1\}$, for $P = 1$ and $N = 0.1$, as a function of $\text{INR} = Q/N$.

of θ . Hence for scenarios where the mismatch factor is smaller, e.g. $\theta_{\min} \approx \theta_{\max}$ or $(P + N) \gg Q$, expression (15) becomes closer to the capacity when the encoder and the decoder are both aware of the channel index θ controlling the communication.

For $\|\Theta\| = 2$, the lower bound (16) provides significant gains compared to the previous bound [5]. Although the bound (18) is not tight in general, notice that it is a sharper bound than those derived by previous results in [3], [5], [16] and it is tight for some special sets Θ as shown in Figure 1.

Coding strategy: Notice that when $\epsilon_{\Theta}^* < N(\|\Theta\| - 1)/(P + \|\Theta\|N)$ the best encoder strategy is implementing a DPC to mitigate the common part of the interfering signal and hence the remainder part is treated as additional noise. In contrast to this, if $\epsilon_{\Theta}^* \geq (\|\Theta\| - 1)/\|\Theta\|$ the best encoder strategy becomes to use time-sharing to mitigate (completely) the interference. This is obtained by allowing the encoder and the decoder to have access to a source of common randomness (e.g. a dither sequence [7]), which is not available if X is restricted to be a deterministic mapping. Otherwise, when $N(\|\Theta\| - 1)/(P + \|\Theta\|N) \leq \epsilon_{\Theta}^* < (\|\Theta\| - 1)/\|\Theta\|$ the encoder combines both strategies by using superposition coding. *Asymptotic*

analysis: In the limit of high SNR (fixed N and Q , $P_C + P_\Delta \rightarrow \infty$) the mismatch factor ϵ_Θ vanishes to zero and thus the rate expression (15) coincides with its natural upper bound given by the case when the encoder is informed with θ , which establishes the optimality of the lower bound in the high SNR limit.

In the limit when $Q \rightarrow \infty$ (for $N, P, \|\Theta\|$ fixed) the mismatch factor becomes

$$\epsilon_\infty^* \triangleq \lim_{Q \rightarrow \infty} \epsilon_\Theta^* = \left(\frac{\theta_{\max} - \theta_{\min}}{\theta_{\max} + \theta_{\min}} \right)^2.$$

Furthermore, when $\|\Theta\| \gg 1$ (for N, P, Q fixed) the lower bound in (16) reduces to

$$\lim_{\|\Theta\| \rightarrow \infty} R_-^\Theta(P) = \frac{1}{2} \log \left[1 + \frac{P(1 - \epsilon_\Theta^*)}{N + \epsilon_\Theta^* P} \right], \quad (19)$$

while for $\theta_{\max} \gg \theta_{\min}$ (for N, P, Q, K fixed) (16) writes

$$\lim_{\theta_{\max}/\theta_{\min} \rightarrow \infty} R_-^\Theta(P) = \frac{1}{2\|\Theta\|} \log \left(1 + \frac{P}{N} \right). \quad (20)$$

The scenarios (19) and (20), i.e. $\epsilon_\infty^* \approx 1$ and $\|\Theta\| \gg 1$, yield the most important loss of degrees of freedom.

B. Sketch of Proof of theorem 4.1

Coding scheme: The encoder splits the information $m = (m_0, m_1)$, namely common information m_0 and private information m_1 . Then it divides the power P into $\{P_C, P_1, \dots, P_{\|\Theta\|}\}$. The encoder sends m_0 using a standard DPC \mathbf{U} , sampled of length n i.i.d. from a PD $P_{U|S} = \mathcal{N}(\alpha_c S, P_C)$, applied to the interference \mathbf{S} and treats the remainder interference as noise. Whereas m_1 is sent using time-sharing via $\|\Theta\|$ different DPCs $\{\mathbf{V}_1, \dots, \mathbf{V}_K\}$, sampled i.i.d. of lengths $\{\lfloor n\lambda_1 \rfloor, \dots, \lfloor n\lambda_K \rfloor\}$ from PDs $P_{V_k|US} = \mathcal{N}(\alpha_k(\theta_k - \alpha_c)S + U, P_k)$, applied once to each of interferences $\{\theta_1 \mathbf{S}, \dots, \theta_K \mathbf{S}\}$. Send $\mathbf{X} = \mathbf{X}_C + \mathbf{X}_D$ with $\mathbf{X}_C = \mathbf{U} - \alpha_c \mathbf{S}$ and $\mathbf{X}_D = [\mathbf{X}_1 \dots \mathbf{X}_K]$, where $\mathbf{X}_k = \mathbf{V}_k - \alpha_k(\theta_k - \alpha_c) \mathbf{S} - \mathbf{U}$. By substituting this in (4), it is not difficult to show that

$$R_-^\Theta(P) = \max_{(\alpha_c, \underline{\alpha}) \in \mathbb{R}^{\|\Theta\|+1}} \min_{k \in \{1, \dots, \|\Theta\|\}} \left\{ I(U^{(\alpha_c)}; Y_k) - I(U^{(\alpha_c)}; S) + \lambda_k \left[I(V_k^{(\alpha_k)}; Y_k | U^{(\alpha_c)}) - I(V_k^{(\alpha_k)}; S | U^{(\alpha_c)}) \right] \right\}.$$

V. SUMMARY AND DISCUSSION

We have investigated the compound state-dependent DMC with non-causal state information at the transmitter but not at the receiver. Some references [9], [10] and conjectures on the capacity of these channels [3] have lent support to the general belief that the natural extension (2) of the Gel'fand and Pinsker's capacity [1] to the compound setting case is indeed optimal. This paper shows that this is not

in general the case. We found that the capacity of the general compound DMC can be strictly larger than the straightforward extension of the Gel'fand and Pinsker's capacity. We derived a new lower bound on the capacity and showed that it is tight for the compound channel with degraded components. It would be of interest to determine whether the result here (3) is strictly better than the common rate result reported in [4, eq. (45)] and further explore the optimality of this result for channels with semi-deterministic and less noisy components.

The compound Gaussian Dirty-Paper channel that consists of an AWGN channel with an additive interference, where the input and the state signals are affected by fading coefficients whose realizations are unknown at the transmitter, was also considered. We derived lower and upper bounds on the capacity of this channel that are tight for some special cases.

ACKNOWLEDGMENT

This research is supported by the FP7 Network of Excellence in Wireless COMmunications NEW-COM++.

REFERENCES

- [1] S. I. Gel'fand and M. S. Pinsker, "Coding for channel with random parameters," *Problems of Control and IT.*, vol. 9, pp. 19–31, 1980.
- [2] M. Costa, "Writing on dirty paper (corresp.)," *Information Theory, IEEE Transactions on*, vol. 29, no. 3, pp. 439–441, May 1983.
- [3] P. Mitran, N. Devroye, and V. Tarokh, "On compound channels with side information at the transmitter," *Information Theory, IEEE Transactions on*, vol. 52, no. 4, pp. 1745–1755, April 2006.
- [4] Y. Steinberg and S. Shamai, "Achievable rates for the broadcast channel with states known at the transmitter," *Information Theory, 2005. Proceedings. International Symposium on*, pp. 2184–2188, Sept. 2005.
- [5] A. Khisti, U. Erez, A. Lapidoth, and G. Wornell, "Carbon copying onto dirty paper," *Information Theory, IEEE Transactions on*, vol. 53, no. 5, pp. 1814–1827, May 2007.
- [6] I. Maric, A. Goldsmith, G. Kramer, and S. Shamai (Shitz), "On the capacity of interference channels with one cooperating transmitter," *European Trans. on Telecommunications, Special issue: "New Directions in Information Theory"*, vol. 19, no. 4, pp. 329 – 495, June 2008.
- [7] U. Erez, S. Shamai, and R. Zamir, "Capacity and lattice strategies for canceling known interference," *Information Theory, IEEE Transactions on*, vol. 51, no. 11, pp. 3820–3833, Nov. 2005.
- [8] H. Weingarten, S. Shamai (shitz), and G. Kramer, "On the compound mimo broadcast channel," in *Information Theory and Applications (ITA 2007), UCSD, San Diego, USA.*, January 29-February 2 2007.
- [9] P. Piantanida and S. Shamai (Shitz), "Capacity of compound state-dependent channels with states known at the transmitter," in *IEEE Symposium on Information Theory (ISIT 2009), Korea*, June 2009.
- [10] P. Moulin and Y. Wang, "Capacity and random-coding exponents for channel coding with side information," *Information Theory, IEEE Transactions on*, vol. 53, no. 4, pp. 1326–1347, April 2007.

- [11] C. Nair and A. El Gamal, "The capacity region of a class of three-receiver broadcast channels with degraded message sets," *Information Theory, IEEE Trans. on*, vol. 55, no. 10, pp. 4479–4493, Oct. 2009.
- [12] Y.-K. Chia and A. El Gamal, "3-receiver broadcast channels with common and confidential messages," in *Proceedings of International Symposium on Information Theory*, June 2009, pp. 1849–1853.
- [13] N. Chandra, A. El Gamal, and Y.-K. Chia, "An achievability scheme for the compound channel with state noncausally available at the encoder," in <http://arxiv.org/abs/1004.3427>, April 2010.
- [14] A. Lapidoth and P. Narayan, "Reliable communication under channel uncertainty," *Information Theory, IEEE Transactions on*, vol. 44, pp. 2148–2177, October 1998.
- [15] Y. Steinberg, "Coding for the degraded broadcast channel with random parameters, with causal and noncausal side information," *Information Theory, IEEE Trans. on*, vol. 51, no. 8, pp. 2867–2877, aug. 2005.
- [16] P. Grover and A. Sahai, "Writing on rayleigh faded dirt: a computable upper bound to the outage capacity," in *Information Theory, 2007. ISIT 2007. IEEE International Symposium on*, June 2007, pp. 2166–2170.