# On the Rate-Limited Gelfand-Pinsker Problem

Ravi Tandon Sennur Ulukus
Department of Electrical and Computer Engineering
University of Maryland, College Park, MD 20742
*ravit@umd.edu* *ulukus@umd.edu*

***Abstract*—We study a rate-limited version of the well known problem of coding for channels with random parameters which was studied by Gelfand and Pinsker [1]. In particular, we consider a state-dependent channel when the transmitter is supplied with the state information at a rate $R_e$. We obtain a new upper bound on the capacity, $\mathcal{C}(R_e)$, for this channel. We explicitly evaluate this upper bound for the rate-limited dirty paper coding (DPC) problem and show that it strictly improves upon the DPC capacity for certain values of $R_e$.**

## I. Introduction

The study of state-dependent channels was initiated by Shannon in [2] where the channel state information (CSI) is assumed to be available at the transmitter in a causal fashion. Shannon derived the capacity of this channel by showing that it is equal to the capacity of another discrete memoryless channel with the same output alphabet and an enlarged input alphabet of size $|\mathcal{X}|^{|\mathcal{T}|}$, where $|\mathcal{T}|$ is the size of the state alphabet.

The case of non-causal CSI at the transmitter was first considered by Kuznetsov and Tsybakov [3] where achievable rates were provided, although capacity was not found. Gelfand and Pinsker derived the capacity of the state-dependent channel with non-causal CSI at the transmitter in their landmark paper [1]. The result of [1] was used by Costa [4] to evaluate the capacity of a channel with input power constraint and when the channel is an additive Gaussian state channel corrupted with independent additive Gaussian noise. This problem is commonly referred to as the dirty paper coding (DPC) problem and has received much attention recently.

Heegard and El Gamal [5] studied state-dependent channels with various modifications regarding the rate-limited knowledge of the state at both the transmitter and the receiver. For the general case when the transmitter is supplied the state information at a rate $R_e$ and the receiver is supplied the state information at a rate $R_d$, an achievable rate was obtained in [5] as a function of $(R_e, R_d)$. So far, for all the cases where the capacity has been established, the achievable rate proposed by Heegard and El Gamal has turned out to be optimal [6]. The two seemingly simple cases are still open: 1) When $R_e = 0$ and we wish to determine the capacity as a function of $R_d$. This corresponds to rate-limited CSI at the receiver and no CSI at the transmitter. Ahlswede and Han [7] obtained an achievable rate for this channel and conjectured it to be the capacity. 2) When $R_d = 0$ and we wish to characterize the capacity as a function of $R_e$. The achievable rates for this case can be obtained via [5]. In this paper, we provide a new upper bound on the capacity of this second channel model. We explicitly evaluate our upper bound for the rate-limited DPC problem. We show that for a certain range of values of $R_e$, our upper bound strictly improves upon the trivial upper bound of DPC capacity obtained by Costa [4].

This work was supported by NSF Grants CCF 04-47613, CCF 05-14846, CNS 07-16311 and CCF 07-29127.

## II. Channel Model

A discrete memoryless state-dependent channel is defined by a channel input alphabet $\mathcal{X}$, a state alphabet $\mathcal{T}$, a channel output alphabet $\mathcal{Y}$ and a transition probability function $p(y|x,t)$ defined for every pair $(x,t) \in \mathcal{X} \times \mathcal{T}$. It is also assumed that the transmitter is supplied with the state information at a rate $R_e$ (see Figure 1).

An $(n, M, P_e)$ code for this channel is defined by a state encoding function, $f_s : \mathcal{T}^n \rightarrow J$, a channel encoding function, $f_e : W \times J \rightarrow \mathcal{X}^n$ and a decoding function, $g : \mathcal{Y}^n \rightarrow W$. The transmitter produces a message $W$ which is uniformly distributed on the set $\{1, \ldots, M\}$ and communicates it in $n$ channel uses. The average probability of error is defined as $P_e = \Pr[\hat{W} \neq W]$. A rate $R$ is said to be achievable for this channel if for any $\epsilon > 0$, there exists an $(n, M, P_e)$ code such that $R \leq \log(M)/n$ and $P_e < \epsilon$ for sufficiently large $n$. The capacity of this channel is the supremum of all achievable rates $R$.

## III. A New Upper Bound

We now present the main result of this paper, which is a new upper bound on the capacity of state-dependent channels with rate-limited state information at the transmitter:

$$\mathcal{UB}(R_e) = \sup_{T \rightarrow V \rightarrow (U,X): I(T;V) \leq R_e} I(U;Y) \tag{1}$$

We will now present the proof of the upper bound. We start by obtaining an upper bound on $R$ as,

$$\begin{aligned} nR &= H(W) && (2)\\ &= I(W;Y^n) + H(W|Y^n) && (3)\\ &\leq I(W;Y^n) + n\epsilon_n && (4)\\ &= \sum_{i=1}^{n} I(W;Y_i|Y^{i-1}) + n\epsilon_n && (5)\\ &= \sum_{i=1}^{n} I(W,Y^{i-1};Y_i) - \sum_{i=1}^{n} I(Y_i;Y^{i-1}) + n\epsilon_n && (6) \end{aligned}$$

where (4) follows from Fano's inequality [8]. Moreover, we also have the following condition from the fact that the state

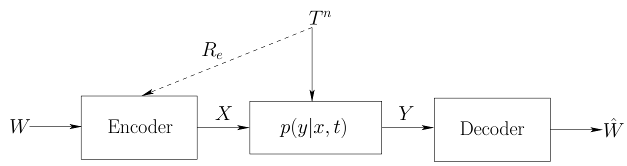


Fig. 1. The state-dependent channel with rate-limited state information at the transmitter.

information is available to the encoder at a rate $R_e$,

$$nR_e \geq H(J) \quad (7)$$
$$\geq I(J;T^n) \quad (8)$$
$$= \sum_{i=1}^{n} I(J;T_i|T^{i-1}) \quad (9)$$
$$= \sum_{i=1}^{n} I(J,T^{i-1};T_i) \quad (10)$$

where (10) follows from the fact that $T_i$s are i.i.d. Finally, we note the following Markov chain,

$$T_i \to (J,T^{i-1}) \to (W,Y^{i-1},X_i) \quad (11)$$

We now define

$$U_i = (W,Y^{i-1}), \quad V_i = (J,T^{i-1}) \quad (12)$$

Returning to (6), we have

$$nR \leq \sum_{i=1}^{n} I(W,Y^{i-1};Y_i) - \sum_{i=1}^{n} I(Y_i;Y^{i-1}) + n\epsilon_n \quad (13)$$
$$\leq \sum_{i=1}^{n} I(W,Y^{i-1};Y_i) + n\epsilon_n \quad (14)$$
$$= nI(U_Q;Y_Q|Q) + n\epsilon_n \quad (15)$$
$$\leq nI(U_Q,Q;Y_Q) + n\epsilon_n \quad (16)$$
$$= nI(U;Y) + n\epsilon_n \quad (17)$$

and returning to (10), we have

$$nR_e \geq \sum_{i=1}^{n} I(J,T^{i-1};T_i) \quad (18)$$
$$= nI(V_Q;T_Q|Q) \quad (19)$$
$$= nI(V_Q,Q;T_Q) \quad (20)$$
$$= nI(V;T) \quad (21)$$

where (21) follows from the fact that $T_i$s are i.i.d. and therefore $T_Q$ is independent of $Q$, where $Q$ is uniformly distributed over $\{1,\ldots,n\}$ and is independent of all other random variables, and we have defined $U=(Q,U_Q)$, $V=(Q,V_Q)$, $Y=Y_Q$, $X=X_Q$ and $T=T_Q$ and $T \to V \to (U,X)$ is a Markov chain.

We now combine (17) and (21) to express our upper bound on the capacity of the state-dependent channel with rate-limited state information at the transmitter as,

$$\mathcal{UB}(R_e) = \max_{p(v|t),p(u,x|v):I(T;V)\leq R_e} I(U;Y) \quad (22)$$

On the other hand, Heegard and El Gamal proposed the following achievable rates for this channel, which can be obtained from [5] by substituting $S_0=c$, $S_d=c$ and $S_e=V$, where $c$ is a constant,

$$\mathcal{LB}(R_e) = \max_{p(v|t),p(u,x|v):I(T;V)\leq R_e} I(U;Y) - I(U;V) \quad (23)$$

We will now show that our upper bound obtained in (22) matches (23) and yields the capacity for two special classes of state-dependent channels.

## IV. The Modulo-Additive State Channel

For the case when $Y = X \oplus N$, and $T = N \oplus \tilde{N}$, and $|\mathcal{X}| = |\mathcal{Y}| = \mathcal{K}$, we can further upper bound our upper bound to obtain an upper bound for this class of channels which was also obtained in [9], as follows,

$$\mathcal{C}(R_e) \leq \max I(U;Y) \quad (24)$$
$$\leq \max \log(\mathcal{K}) - H(Y|U) \quad (25)$$
$$\leq \max \log(\mathcal{K}) - H(Y|X,U) \quad (26)$$
$$= \max \log(\mathcal{K}) - H(N|X,U) \quad (27)$$
$$\leq \max \log(\mathcal{K}) - H(N|V) \quad (28)$$
$$= \log(\mathcal{K}) - \min_{p(v|t):I(T;V)\leq R_e} H(N|V) \quad (29)$$

where (28) follows from the Markov chain $N \to T \to V \to (U,X)$ which implies $I(N;U,X) \leq I(N;V)$, which in turn implies $H(N|X,U) \geq H(N|V)$.

For the case when $X,Y$ and $T$ are binary, this bound becomes

$$\mathcal{C}(R_e) \leq 1 - \min_{p(v|t):I(T;V)\leq R_e} H(N|V) \quad (30)$$

where $N \to T \to V$ forms a Markov chain. It was shown in [9] that the above upper bound is tight and matches the achievable rate of [5] for the case when $T \sim$ Ber$(1/2)$.

## V. Capacity Result for a Symmetric Binary Erasure Channel with Two States

We will show that for a particular binary input state-dependent channel with two states, our upper bound yields the capacity. The state $T$ is binary with $\Pr(T=0)=1/2$. The channel input $X$ is binary and channel output $Y$ is ternary. For channel states $T=0,1$, the transition probabilities, $p(y|x,t)$, are as shown in Figure 2. Also note that this class of channels does not fall in the category of modulo-additive channels. We start by further upper bounding $\mathcal{UB}(R_e)$ as follows,

$$\mathcal{UB}(R_e) = \max I(U;Y) \quad (31)$$
$$\leq \max h(\epsilon) + \epsilon - H(Y|U) \quad (32)$$
$$\leq \max h(\epsilon) + \epsilon - H(Y|V,U,X) \quad (33)$$
$$= \max h(\epsilon) + \epsilon - H(\tilde{U}|V) \quad (34)$$
$$= h(\epsilon) + \epsilon - \inf H(\tilde{U}|V) \quad (35)$$

where (32) follows from the fact that $H(Y) \leq h(\epsilon)+\epsilon$, (33) follows from the fact that conditioning reduces entropy and (34) follows from easily verifying the following,

$$H(Y|X,V,U) = H(\tilde{U}|V) \quad (36)$$

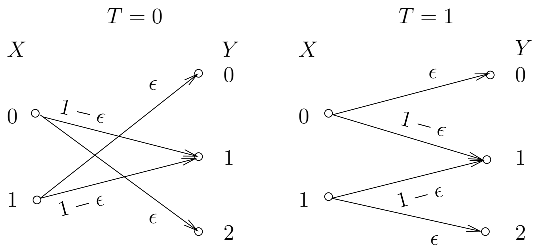


Fig. 2. A symmetric binary erasure channel with two states.

where $\tilde{U}$ is a random variable with $|\tilde{\mathcal{U}}| = 3$ and $p(\tilde{u}|t)$, expressed as a stochastic matrix $G$ as,

$$G = \begin{pmatrix} \epsilon & 1-\epsilon & 0 \\ 0 & 1-\epsilon & \epsilon \end{pmatrix} \tag{37}$$

and the random variables $(\tilde{U}, T, V)$ satisfy the Markov chain $\tilde{U} \rightarrow T \rightarrow V$ by construction. Using [10], we can explicitly evaluate (35) to arrive at

$$\mathcal{UB}(R_e) \leq \min(\epsilon R_e, \epsilon) \tag{38}$$

We now evaluate (23) by setting $|\mathcal{V}| = 2$ and selecting $\Pr(V = 1|T = 0) = \Pr(V = 0|T = 1) = \mu$, such that $I(T; V) \leq R_e$. We set $|\mathcal{U}| = 2$, with $U$ being selected as uniformly distributed on $\{0, 1\}$ and independent of $V$, i.e., $I(U; V) = 0$. We finally select $X$ as a deterministic function of $(U, V)$ as follows,

$$X = U \oplus V \tag{39}$$

For this selection of random variables, it is straightforward to show that,

$$\mathcal{LB}(R_e) \geq \min(\epsilon R_e, \epsilon) \tag{40}$$

and hence we have the capacity expression as

$$\mathcal{C}(R_e) = \min(\epsilon R_e, \epsilon) \tag{41}$$

## VI. Rate-Limited Dirty Paper Coding

We will now provide an upper bound for the case when the forward channel is an additive Gaussian noise channel and the channel states are also additive and Gaussian (see Figure 3). In particular, the channel is described as

$$Y = X + T + Z \tag{42}$$

where the channel input $X$ is subject to an average power constraint $P$, the channel state $T$ and the channel input $X$ are independent of $Z$, where $Z$ is a zero-mean, Gaussian random variable with variance $\sigma_Z^2$. Moreover, the state random variable $T$ is a zero-mean Gaussian random variable with variance $\sigma_T^2$. The capacity of this channel is known when the state sequence is non-causally known at the transmitter. This result was obtained by Costa in [4] and the capacity was found to be

$$\mathcal{C}_{DPC} = \frac{1}{2}\log\left(1 + \frac{P}{\sigma_Z^2}\right) \tag{43}$$

We will provide an upper bound for the case when the transmitter is supplied information about the channel state $T$ at a rate of $R_e$. It is clear that when $R_e \rightarrow \infty$, this situation

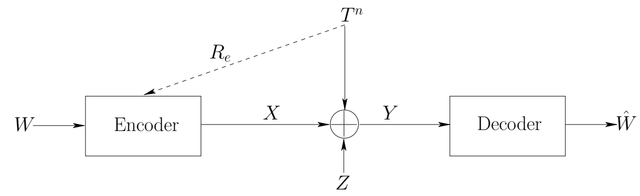


Fig. 3. The rate-limited DPC channel model.

corresponds to the setting of [4] and we have

$$\mathcal{C}(\infty) = \mathcal{C}_{DPC} = \frac{1}{2}\log\left(1 + \frac{P}{\sigma_Z^2}\right) \tag{44}$$

On the other hand, when $R_e = 0$, we know that

$$\mathcal{C}(0) = \frac{1}{2}\log\left(1 + \frac{P}{\sigma_Z^2 + \sigma_T^2}\right) \tag{45}$$

which is the capacity of a channel with total Gaussian noise $T+Z$, i.e., when there is no state information at the transmitter and the state random variable $T$ acts as additional additive Gaussian noise besides $Z$.

Capacity of the rate-limited dirty paper channel, i.e., $\mathcal{C}(R_e)$ is not known for $0 < R_e < \infty$. Trivial lower/upper bounds for any $0 < R_e < \infty$ are

$$\frac{1}{2}\log\left(1 + \frac{P}{\sigma_Z^2 + \sigma_T^2}\right) \leq \mathcal{C}(R_e) \leq \frac{1}{2}\log\left(1 + \frac{P}{\sigma_Z^2}\right) \tag{46}$$

We will show that a strengthened version of our upper bound is strictly less than $\mathcal{C}_{DPC}$ for certain values of $R_e$. We start by obtaining an upper bound on $R$ as,

$$\begin{aligned} nR &= H(W) && (47)\\ &= I(W; Y^n, J) + H(W|Y^n, J) && (48)\\ &\leq I(W; Y^n, J) + n\epsilon_n && (49)\\ &= I(W; Y^n|J) + n\epsilon_n && (50)\\ &= h(Y^n|J) - h(Y^n|W, J) + n\epsilon_n && (51) \end{aligned}$$

where (49) follows from Fano's inequality [8] and (50) follows from the fact that the message $W$ and the random variable $J$ are independent. The main idea behind this strengthened upper bound is to consider a larger quantity $I(W; Y^n, J)$ in (49) as opposed to $I(W; Y^n)$ in (4). This approach will permit us to invoke the Markov chain $X^n \rightarrow J \rightarrow T^n$ which will subsequently yield an improved upper bound.

Returning to (51), we will separately obtain an upper bound on $h(Y^n|J)$ and a lower bound on $h(Y^n|W, J)$. We start by considering the first term in (51),

$$\begin{aligned} h(Y^n|J) &= \sum_{i=1}^{n} h(Y_i|J, Y^{i-1}) && (52)\\ &\leq \sum_{i=1}^{n} h(Y_i|J) && (53)\\ &\leq \frac{n}{2}\log\left((2\pi e)(P + \sigma_T^2 + \sigma_Z^2)\right) && (54) \end{aligned}$$

where (53) follows from the fact that conditioning reduces entropy and by dropping $Y^{i-1}$ from the conditioning, and (54) follows from the following sequence of inequalities,

$$\sum_{i=1}^{n} h(Y_i|J)$$
$$\leq \sum_{i=1}^{n} \frac{1}{2}\log(2\pi e \text{Var}(Y_i|J)) \tag{55}$$
$$= \sum_{i=1}^{n} \frac{1}{2}\log(2\pi e(\text{Var}(X_i|J) + \text{Var}(T_i|J) + \text{Var}(Z_i|J))) \tag{56}$$
$$\leq \sum_{i=1}^{n} \frac{1}{2}\log(2\pi e(\text{Var}(X_i) + \text{Var}(T_i) + \text{Var}(Z_i))) \tag{57}$$
$$= \sum_{i=1}^{n} \frac{1}{2}\log(2\pi e(\text{Var}(X_i) + \sigma_T^2 + \sigma_Z^2)) \tag{58}$$
$$\leq \frac{n}{2}\log\left((2\pi e)(P + \sigma_T^2 + \sigma_Z^2)\right) \tag{59}$$

where (55) follows from the maximum entropy theorem [8], (56) follows from the fact that $Z^n$ is independent of $(X^n, T^n, J)$ and the Markov chain $X_i \rightarrow J \rightarrow T_i$, which also implies that $\text{Cov}(X_i, T_i|J) = 0$ for all $i = 1, \ldots, n$, (57) follows from the fact that expected conditional variance is upper bounded by unconditional variance, (58) follows from the fact that $\text{Var}(T_i) = \sigma_T^2$ and $\text{Var}(Z_i) = \sigma_Z^2$ for all $i = 1, \ldots, n$ and (59) follows from the concavity of log function and the average input power constraint $P$.

We now consider the second term in (51) and obtain a lower bound as,

$$h(Y^n|W, J) \geq h(Y^n|X^n, W, J) \tag{60}$$
$$= h(T^n + Z^n|X^n, W, J) \tag{61}$$
$$= h(T^n + Z^n|J) \tag{62}$$
$$\geq \frac{n}{2}\log\left(e^{\frac{2}{n}h(T^n|J)} + 2\pi e\sigma_Z^2\right) \tag{63}$$
$$\geq \frac{n}{2}\log\left((2\pi e)(\sigma_T^2 e^{-2R_e} + \sigma_Z^2)\right) \tag{64}$$

where (60) follows from the fact that conditioning reduces entropy, (62) follows from the Markov chain $T^n \rightarrow J \rightarrow (X^n, W)$ and (63) follows from the vector entropy power inequality (EPI) [8]. Finally, (64) follows from the following,

$$nR_e \geq H(J) \tag{65}$$
$$\geq I(J; T^n) \tag{66}$$
$$= h(T^n) - h(T^n|J) \tag{67}$$

which yields

$$h(T^n|J) \geq \frac{n}{2}\left(\log(2\pi e\sigma_T^2) - 2R_e\right) \tag{68}$$

and we substitute (68) in (63) to arrive at (64).

We now substitute (54) and (64) in (51) to finally arrive at our upper bound,

$$\mathcal{UB}(R_e) = \frac{1}{2}\log\left(\frac{P + \sigma_T^2 + \sigma_Z^2}{\sigma_Z^2 + \sigma_T^2 e^{-2R_e}}\right) \tag{69}$$

When $R_e = 0$, our upper bound is clearly optimal,

$$\mathcal{UB}(0) = \frac{1}{2}\log\left(\frac{P + \sigma_T^2 + \sigma_Z^2}{\sigma_Z^2 + \sigma_T^2}\right) \tag{70}$$
$$= \mathcal{C}(0) \tag{71}$$

On the other hand, our upper bound is strictly smaller than the DPC upper bound, $\mathcal{C}(\infty)$, for $0 < R_e < R_e^c$, where

$$R_e^c = \frac{1}{2}\log\left(1 + \frac{P}{\sigma_Z^2}\right) \tag{72}$$

For $R_e \geq R_e^c$, the DPC upper bound is strictly smaller than our upper bound. Therefore, we take the smaller of these two bounds and obtain a compact expression for the upper bound as,

$$\mathcal{UB}(R_e) = \begin{cases} \frac{1}{2}\log\left(\frac{P+\sigma_T^2+\sigma_Z^2}{\sigma_Z^2+\sigma_T^2 e^{-2R_e}}\right), & 0 \leq R_e < R_e^c; \\ \frac{1}{2}\log\left(1 + \frac{P}{\sigma_Z^2}\right), & R_e^c \leq R_e < \infty. \end{cases} \tag{73}$$

We now obtain achievable rates for rate-limited DPC. In particular, we will obtain a potentially sub-optimal evaluation of the following achievable rate given in [5].

$$\mathcal{LB}(R_e) = \max_{p(v|t), p(u,x|v): I(T;V) \leq R_e} I(U; Y) - I(U; V) \tag{74}$$

The main idea behind this achievable scheme is a combination of rate-distortion type coding [8] and Gelfand-Pinsker binning [1]. We select the following auxiliary random variables,

$$V = T + \tilde{N} \tag{75}$$

where $\tilde{N}$ is a zero-mean Gaussian random variable with variance $\sigma_{\tilde{N}}^2$ and is independent of $T$. Here, $\tilde{N}$ can be interpreted as the compression noise. From the constraint $I(T; V) \leq R_e$, we have

$$I(T; V) = I(T; T + \tilde{N}) \tag{76}$$
$$= \frac{1}{2}\log\left(1 + \frac{\sigma_T^2}{\sigma_{\tilde{N}}^2}\right) \leq R_e \tag{77}$$

From (77), we obtain a constraint on the variance $\sigma_{\tilde{N}}^2$ as,

$$\sigma_{\tilde{N}}^2 \geq \frac{\sigma_T^2}{e^{2R_e} - 1} \tag{78}$$

Next, we select $X$ as a zero-mean Gaussian random variable with variance $P$, which is independent of $V$. We select the random variable $U$ as

$$U = X + \alpha V \tag{79}$$

We are now ready to evaluate the achievable rates for this selection of random variables $(V, X, U)$. So far, we have not specified $\alpha$. We will later optimize $\alpha$, as a function of $R_e$, to obtain the best possible achievable rate for this selection of auxiliary random variables. We start by simplifying the expression in (74),

$$I(U; Y) - I(U; V) = h(U|V) - h(U|Y) \tag{80}$$

We first consider

$$h(U|V) = h(X + \alpha V|V) \quad (81)$$
$$= h(X|V) \quad (82)$$
$$= h(X) \quad (83)$$
$$= \frac{1}{2}\log(2\pi e P) \quad (84)$$

where (83) follows since $X$ and $V$ are selected to be independent. Now consider

$$h(U|Y) = h(X + \alpha V|X + T + Z) \quad (85)$$
$$= h(X + \alpha(T + \tilde{N})|X + T + Z) \quad (86)$$
$$= \frac{1}{2}\log\left((2\pi e)\frac{P\sigma_Z^2 + \mu(\alpha, R_e)}{P + \sigma_T^2 + \sigma_Z^2}\right) \quad (87)$$

where

$$\mu(\alpha, R_e) = \alpha^2\sigma_Z^2\sigma_T^2 + (1-\alpha)^2 P\sigma_T^2 + \frac{\alpha^2\sigma_T^2(P + \sigma_T^2 + \sigma_Z^2)}{e^{2R_e} - 1} \quad (88)$$

Combining (84) and (87) and substituting in (80) we obtain an achievable rate as a function of $\alpha$, for any $R_e$ as,

$$\mathcal{LB}(R_e, \alpha) = \frac{1}{2}\log\left(\frac{P(P + \sigma_T^2 + \sigma_Z^2)}{P\sigma_Z^2 + \mu(\alpha, R_e)}\right) \quad (89)$$

Next, we optimize the above achievable rate with respect to $\alpha$. This is equivalent to minimizing $\mu(\alpha, R_e)$. We first note that $\mu(\alpha, R_e)$ is convex in $\alpha$ and therefore, the minimum of $\mu(\alpha, R_e)$ is obtainded at $\alpha^*(R_e)$ where $d\mu(\alpha^*, R_e)/d\alpha = 0$. We therefore have the following

$$\alpha^*(R_e) = \frac{P}{P + \sigma_Z^2 + \frac{P + \sigma_T^2 + \sigma_Z^2}{e^{2R_e} - 1}} \quad (90)$$

We substitute (90) in (89) to obtain a closed form expression for the achievable rate as follows

$$\mathcal{LB}(R_e) = \frac{1}{2}\log\left(\frac{P + \sigma_T^2 e^{-2R_e} + \sigma_Z^2}{\sigma_Z^2 + \sigma_T^2 e^{-2R_e}}\right) \quad (91)$$

We now consider the two extreme cases for the values of $R_e$. If $R_e = 0$, then from (90), the optimal selection of $\alpha$ is

$$\alpha^*(0) = 0 \quad (92)$$

and the achievable rate is

$$\mathcal{LB}(0) = \frac{1}{2}\log\left(1 + \frac{P}{\sigma_T^2 + \sigma_Z^2}\right) \quad (93)$$

which yields the capacity $\mathcal{C}(0)$. If $R_e = \infty$, then the optimal selection of $\alpha$ is

$$\alpha^*(\infty) = \frac{P}{P + \sigma_Z^2} \quad (94)$$

and the achievable rate is

$$\mathcal{LB}(\infty) = \frac{1}{2}\log\left(1 + \frac{P}{\sigma_Z^2}\right) \quad (95)$$

which yields the DPC capacity $\mathcal{C}_{DPC}$. We should remark here that this $\alpha^*(\infty)$ is the same selection used by Costa in [4] to obtain the DPC capacity.

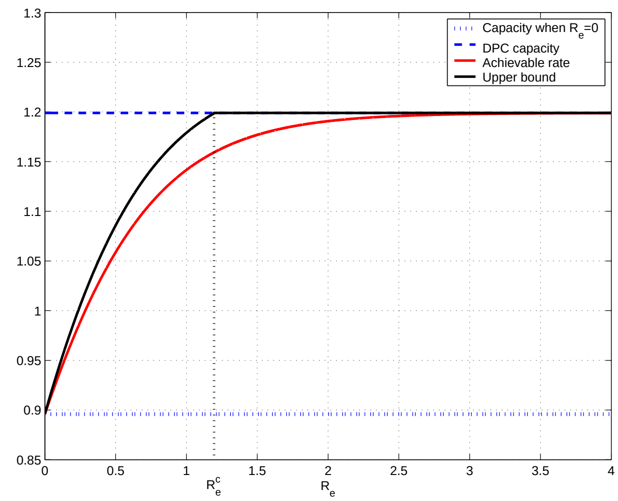


Fig. 4. Illustration of bounds when $P = 10$, $\sigma_T^2 = \sigma_Z^2 = 1$.

Figure 4 shows our upper bound in (73), the achievable rate in (91), the DPC upper bound in (43) and the capacity when $R_e = 0$ in (45) for the case when $P = 10$, $\sigma_Z^2 = \sigma_T^2 = 1$.

## VII. Conclusions

We obtained a new upper bound on the capacity of state-dependent channels with rate-limited state information at the transmitter. We showed that our bound matches the upper bound obtained in [9] for modulo-additive state channels. We also showed that our upper bound yields the capacity for a new class of state-dependent channels. Furthermore, we evaluated our upper bound for the rate-limited DPC problem. We showed that for all finite values of $(P, \sigma_Z^2, \sigma_T^2)$, our upper bound is strictly less than the trivial DPC upper bound for a certain range of $R_e$. We also provided a potentially sub-optimal evaluation of the achievable rates [5] for the rate-limited DPC problem.

## References


[1] S. I. Gelfand and M. S. Pinsker. Coding for channel with random parameters. *Problems of Control and Information Theory*, 9(1):19–31, 1980.

[2] C. E. Shannon. Channels with side information at the transmitter. *IBM Journal Research and Development*, 2:289–293, 1958.

[3] N. V. Kuznetsov and B. S. Tsybakov. Coding in memories with defective cells. *Problemy peredachi informatsii*, 10(2):52–60, 1974.

[4] M. H. M. Costa. Writing on dirty paper. *IEEE Trans. on Information Theory*, 29(3):439–441, May 1983.

[5] C. Heegard and A. El Gamal. On the capacity of computer memory with defects. *IEEE Trans. on Information Theory*, 29(5):731–739, Sep. 1983.

[6] G. Keshet, Y. Steinberg, and N. Merhav. Channel coding in the presence of side information: subject review. *Foundations and Trends in Communications and Information Theory, NOW Publishers*, June 2008.

[7] R. Ahlswede and T. S. Han. On source coding with side information via a multiple-access channel and related problems in multi-user information theory. *IEEE Trans. on Information Theory*, 29(3):396–412, May 1983.

[8] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. New York:Wiley, 1991.

[9] M. Aleksic, P. Razaghi, and W. Yu. Capacity of a class of modulo-sum relay channels. *To appear in IEEE Trans. on Information Theory*.

[10] H. S. Witsenhausen and A. D. Wyner. A conditional entropy bound for a pair of discrete random variables. *IEEE Trans. on Information Theory*, 21(5):493–501, September 1975.