Channel Capacity with Side Information - A Unified View

Syed Ali Jafar
Department of Electrical Engineering and Computer Science
University of California Irvine, CA 92697 USA
E-mail : syed@uci.edu

***Abstract*— We identify the common underlying form of the capacity expression that is applicable to both cases where causal or non-causal side information is made available to the transmitter. Using this common form we find that the capacity with causal and non-causal side information are identical when all the transmitter side information is also made available to the receiver. A genie-aided outerbound is developed that states that when a genie provides one bit of side information to the receiver the resulting capacity improvement can not be more than a bit. Combining these two results we are able to bound the relative capacity advantage of non-causal side information over causal side information for both single user as well as multiple user communication scenarios. Applications of these capacity bounds are demonstrated through examples of random access channels. It is also shown that even one bit of side information at the transmitter can result in unbounded capacity improvement.**

I. Introduction

The theory of communication with side information originated as a dichotomy between the two cases of non-causal side information and causal side information. Initial motivation for non-causal side information came from data storage and data-hiding whereas causal side information in the form of channel knowledge was considered more relevant for wireless communications. However, recent focus on multiuser communications in ad-hoc, cellular, and cognitive radio systems brings together causal and non-causal side information. Causal side information captures the temporal variations of the channel that may only be tracked based on past observations. Non-causal side information is inherent in the prior knowledge of some common information that forms the basis for cooperation as well as the frequency selectivity of the channel that may be tracked at the transmitters and the receivers. In this work we focus on both causal and non-causal side information, the relationship between them, and their extensions to cooperative communications. Our interest is in general capacity expressions. Specialized expressions for AWGN or fading channels may be obtained as special cases from these general expressions, subject to input distribution optimizations. In particular we are interested in the following questions:

- What is the connection between the capacity expressions with causal and non-causal side information at the transmitter?
- What is the maximum possible capacity improvement with a limited amount of side information at the transmitter/receiver?
- What is the relative capacity advantage of non-causal side information over causal side information?

II. Background

For non-causal side information at the transmitter, the single user capacity is known to be [1] [2] [3]:

$$C^{\text{non-causal}} = \max_{\mathcal{P}_{\text{non-causal}}} I(U;Y) - I(U;S_T) \tag{1}$$

where $\mathcal{P}_{\text{non-causal}} = \{P(U,X|S_T) = P(U|S_T)P(X|U,S_T)\}$. Comparing this to the case where no side information is available ($S_T = \phi$),

$$C = \max_{P(U,X)} I(U;Y) = \max_{P(X)} I(X;Y), \tag{2}$$

note that the availability of side information at the transmitter is helpful in that the transmitter can match its input to the channel information by picking the input alphabet U, X conditioned on S_T, as opposed to (2) where the input can not be matched to the channel state. However, the benefit of matching the input to the channel state comes with the cost of the subtractive term in (1), i.e., $I(U;S_T)$ which can be interpreted as the overhead required to communicate to the receiver, the adaptation to the channel state at the transmitter. For the case where the side information is available at the transmitter only causally, the capacity expression has been found by Shannon as [4]

$$C^{\text{causal}} = \max_{P(t)} I(T;Y) \tag{3}$$

where T is an extended alphabet of mappings from the channel state to the input alphabet.

The capacity expressions (1), (2), (3) explicitly account for side information at the transmitter. Side information at the receiver, S_R, is easily incorporated into the same expressions by replacing Y with (Y, S_R) in the corresponding expressions. Recall that unlike transmitter side information, for the receiver side information it does not matter whether it is obtained causally or non-causally.

We start by relating the causal and non-causal cases.

III. Relating Capacity with Causal and Non-causal Side Information

Comparing the non-causal case (1) with the causal case (3) the two capacity expressions are in different forms so that their relationship is not obvious. We make the relationship clearly apparent by representing the causal case in a different form, comparable to (1)

$$
\begin{aligned}
C^{\text{non-causal}} &= \max_{\mathcal{P}_{\text{non-causal}}} I(U;Y,S_R) - I(U;S_T),\\
C^{\text{causal}} &= \max_{\mathcal{P}_{\text{causal}}} I(U;Y,S_R) - I(U;S_T),
\end{aligned}
$$

with

$$
\begin{aligned}
\mathcal{P}_{\text{non-causal}} &= \{P(U,X|S_T) = P(U|S_T)P(X|U,S_T)\}\\
\mathcal{P}_{\text{causal}} &= \{P(U,X|S_T) = P(U)P(X|U,S_T)\}
\end{aligned}
$$

In the non-causal case, the choice of U can be made conditional on the channel state S_T. In the causal case U is picked independent of S_T. This makes the subtractive term equal to zero for the causal case. In both cases, it suffices for the optimal input symbol X to be just a function of U, S_T, i.e. $P(X|U,S_T)$ is either 0 or 1.

Interestingly, the capacity expression (1) has been shown [5], [6] to be the common form of single user capacity for all four cases of non-causal side information as well as the corresponding cases for rate-distortion. In other words, for the capacity problem, whether the non-causal side information is available at the transmitter, the receiver, both, or neither, the capacity expression has the common form $I(U;Y,S_R) - I(U;S_T)$. The only difference is in the constraints on the distribution of the auxiliary random variable U, the input alphabet X and the state variable S. Thus, combining the results of [5], [6] with the common expression obtained above we find that the expression $I(U;Y,S_R) - I(U;S_T)$ is indeed the common expression for not only all cases of non-causal side information but also for causal side information as well.

Investigating the relationship between causal and non-causal information further, we extend a result from [7] previously reported for the case of causal side-information at the transmitter. It is shown that if $S_T = f(S_R)$, i.e., if the side-information at the transmitter is a deterministic function of the side-information at the receiver then the auxiliary random variable U is not needed and coding can be performed directly on the input alphabet X. The relationship between causal and non-causal side information takes the following form in this context.

Theorem: (Relationship between causal and non-causal side information capacity) *If the side-information at the transmitter is a deterministic function of the side-information at the receiver, i.e., if $S_T = f(S_R)$, then capacity with causal side information is equal to the capacity with non-causal side information.* Capacity achieving codes, in both cases, can be constructed directly on the input alphabet and the auxiliary random variable U is not required.

Proof: $C^{\text{non-causal}}$

$$
\begin{aligned}
&= \max_{P(U,X|S_T)=P(U|S_T)P(X|U,S_T)} I(U;Y,S_R) - I(U;S_T)\\
&= \max_{P(U|S_T)P(X|U,S_T)} I(U;S_R) + I(U;Y|S_R) - I(U;S_T)\\
&= \max_{P(U|S_T)P(X|U,S_T)} I(U;S_R,S_T) + I(U;Y|S_R) - I(U;S_T)\\
&= \max_{P(U|S_T)P(X|U,S_T)} I(U;S_R|S_T) + I(U;Y|S_R)\\
&= \max_{P(U|S_T)P(X|U,S_T)} I(U;Y|S_R)\\
&= \max_{P(X|S_T)} I(X;Y|S_R)\\
&= C^{\text{causal}}
\end{aligned}
$$

where the last equality follows from the results of [7]. Thus, if the side information at the transmitter is a deterministic function of the side information at the receiver, then capacity with non-causal side information is the same as the capacity with causal side-information. ■

Next we investigate the value of side information through capacity bounds.

IV. Capacity Bounds with Genie Bits

Transmitter and receiver side information are different in their potential capacity advantages. The following example points out this distinction in two user multiple access and broadcast channel scenarios.

Theorem: *The capacity improvement due to the availability of receiver side information in a multiple access channel is always limited by the amount of the side information itself, whereas it is possible for even one bit of causal transmitter side information on a broadcast channel to result in an unbounded capacity improvement.*

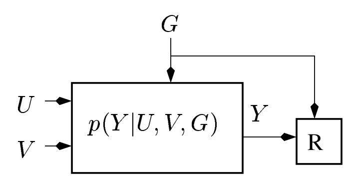

Fig. 2(a): Multiple Access Channel: One genie bit cannot improve sum capacity by more than one bit.

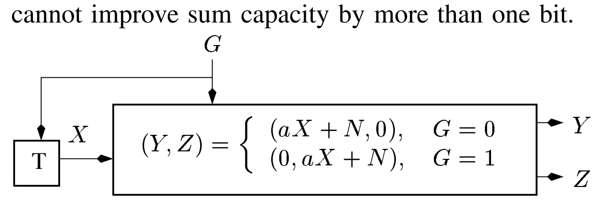

Fig. 2(b): Broadcast Channel Example: One genie bit can improve capacity by an unbounded amount.

To understand this result consider, first a multiple access channel, with two users sending independent symbols U and V respectively and suppose in addition to the signals received through the communication channel Y, there is an additional information G made available to the receiver by a genie every channel use. The information provided by the genie may be correlated, uncorrelated, causal or non-causal, precise or statistical information about the user's own channel coefficients, the interference or the users' message itself. It is easily seen that with this side information G available to the receiver, the sum capacity benefit is no more than the entropy rate of the side information $H(\mathcal{G}) = \lim_{N\to\infty} \frac{1}{N} H(G_1, G_2, \cdots, G_N)$, as C_G

$$\begin{aligned}
&= \sup_{p(U)p(V)} \lim_{N\to\infty} \frac{I(U^N, V^N; Y^N, G^N)}{N} \\
&= \sup_{p(U)p(V)} \lim_{N\to\infty} \frac{I(U^N, V^N; Y^N) + I(U^N, V^N; G^N|Y^N)}{N} \\
&= C + \Delta C
\end{aligned}$$

where C_G is the sum capacity with the side information provided by the genie, C is the sum capacity without the side information and ΔC, the capacity improvement, is bounded by the entropy rate $H(\mathcal{G})$. In other words, if the genie provides one bit of side information per channel use, the capacity benefit $\Delta C = C_G - C$ can not be more than 1 bit, regardless of the kind of side information.

To contrast the previous result with the value of transmitter side information we look at the natural reciprocal of the multiple access channel, i.e., a broadcast channel. Interestingly, it is easy to construct examples where one bit of information made available to the transmitter in a broadcast channel results in an unbounded increase in the sum capacity of a broadcast channel. For example, consider the AWGN broadcast channel with a switch as shown in Fig. 1. The switch randomly changes state every symbol, so either user 1 or user 2 sees the signal $Y = aX + N$ while the other user sees $Y = 0$. With no side information about the switch state at the transmitter the users are statistically equivalent, they are both capable of decoding any message that either one can decode, and therefore the sum capacity is bounded by the single user capacity. The single user capacity of user 1 (or user 2) in this case is $\frac{1}{2}\log(1+a^2)$, where the factor half appears because the channel is active only half the time (assume noise and transmit power are normalized to unity). However, clearly if the switch state is made available to the transmitter as a genie bit, the total capacity would be $C_V = \frac{1}{2}\log(1+a^2) + \frac{1}{2}\log(1+a^2)$. Therefore, one bit of genie information per channel use in this simple example leads to a sum capacity improvement of $C_V - C = \frac{1}{2}\log(1+a^2)$ which is unbounded as a can be chosen to be arbitrary large.

The preceding discussion motivates further investigations into the capacity benefits of side information in wireless networks.

V. Advantage of Non-Causal Side Information over Causal Side Information

For simplicity of exposition we start with a single user channel, and generalize the results to multiple access and broadcast channels in subsequent sections. From the preceding sections we have the following two results.

- If the transmitter side information is also available to the receiver, capacity with causal and non-causal side information is identical.
- If a Genie provides a bit of side information to the receiver it can not improve capacity by more than a bit.

Combining these two results, suppose the transmitter side information is made available to the receiver by a Genie. This requires $H(S_T|S_R)$ Genie bits and therefore cannot improve capacity by more than $H(S_T|S_R)$ bits. Using the results stated above, we have:

$$C^{\text{non-causal}}(S_T, S_R) - C^{\text{causal}}(S_T, S_R) \quad \leq \quad H(S_T|S_R)$$

A. Example: Random Access

Consider a single user DMC characterized by $P(Y|X)$ and with a capacity C_0 when the input is directly controlled by a transmitter. Now, suppose instead that the transmitter is only able to access the channel (control the input) in a random manner as:

$$X = X_T S + X_r(1 - S),$$

where $S \in \{0, 1\}$ is a switch state that determines when the transmitter can access the channel with the symbol X_T, and X_r is a randomly generated input. Suppose the state S is known to the transmitter and *not known* to the receiver. Such a channel is relevant to cognitive communication scenarios [8] and is also similar to the "memory with stuck-at defects" problem considered in [3]. Clearly, if the switch state is provided to the receiver by a Genie the resulting capacity is $C(S, S) = \text{Prob}(S = 1)C_0 = \overline{S}C_0$. Since the extra information provided by the genie is only one bit we have

$$\overline{S}C_0 \quad \geq \quad C^{\text{non-causal}}(S, \phi) \geq \overline{S}C_0 - 1$$

Interestingly, the capacity with causal side information in this case is the same as the capacity with no side information. This is easily seen as follows:

$$X_T = \begin{cases} f_0(U), S = 0, \\ f_1(U), S = 1 \end{cases} \equiv f_1(U), S = 0, 1.$$

In other words, the choice of input symbol does not matter when the switch is open ($S = 0$). Thus, we have

$$C^{\text{causal}}(S, \phi) = C(\phi, \phi) \geq C(S, S) - 1.$$

The effect of memory in side information is also revealed by this example. Suppose the switch changes state in a block static model, i.e. it retains its state for N symbols and then changes to an i.i.d. realization. In this case, the Genie only needs to provide one bit to the receiver every N channel uses and the bounds are tighter.

$$\overline{S}C_0 = C(S,S) \geq C^{\text{non-causal}}(S,\phi) \geq C^{\text{causal}}(S,\phi) = C(\phi,\phi) \geq \overline{S}C_0 - \frac{1}{N}$$

VI. MAC with Side Information

Achievable regions with causal side information are straightforward to obtain because the codewords can always be constructed on mappings from the side information to the channel input alphabet. In the multiple access channel, the two transmitters have (possibly correlated) side information S_{T1}, S_{T2} respectively, and the common receiver has side information S_R. The characterization of the achievable region with side information is the same as without side information, with the codes defined on auxiliary random variables that are independent of the side information, and the actual channel input symbols chosen as a function of the auxiliary random variable and the instantaneous side information. Thus the following achievable region is obtained.

$$\begin{aligned} R_1 &\leq I(U_1; Y, S_R | U_2) \\ R_2 &\leq I(U_2; Y, S_R | U_1) \\ R_1 + R_2 &\leq I(U_1, U_2; Y, S_R) \end{aligned}$$

where U_1, U_2 are mutually independent as well as independent of the side information and the channel inputs are given by $X_1 = f_1(U_1, S_{T1}), X_2 = f_2(U_2, S_{T2})$.

While a full converse is not known, upperbounds for the MAC with causal side information are obtained in [9] in terms of the capacity achieved with transmitter cooperation. Here, we prove a tight converse for the sum rate.

$$\begin{aligned} n(R_1 + R_2) &\leq I(W_1, W_2; Y^n, S_R^n) + n\epsilon \\ &\leq \sum_{i=1}^{n} I(W_1, W_2, S_{T1}^{i-1}, S_{T2}^{i-1}; Y_i, S_{R,i} | Y^{i-1}, S_R^{i-1}) + n\epsilon \\ &\leq \textstyle\sum_{i=1}^{n} H(Y_i, S_{R,i}) - H(Y_i, S_{R,i} | W_1, S_{T1}^{i-1}, W_2, S_{T2}^{i-1}, Y^{i-1}, S_R^{i-1}) + n\epsilon \\ &= \sum_{i=1}^{n} I(U_{1,i}, U_{2,i}; Y_i, S_{R,i}) + n\epsilon \end{aligned}$$

where $U_{1,i} = W_1, S_{T1}^{i-1}$ and $U_{2,i} = W_2, S_{T2}^{i-1}$ are independent of $S_{T1,i}, S_{T2,i}$. The third inequality applies because the current output $Y_i, S_{R,i}$ is independent of all past outputs Y^{i-1}, S_R^{i-1}, given $W_1, S_{T1}^{i-1}, W_2, S_{T2}^{i-1}$ which determines all past inputs. Note that this converse applies for any correlation between the side information at the two transmitters.

For non-causal side information, an achievable region is readily obtained when the side information at the two transmitters is independent. However, to the best of the author's knowledge a converse has not been shown even for independent side information[1]. Correlation of the side information makes even the achievable region non-trivial, as the coding of correlated side information at the two transmitters must be exploited.

For our purpose however, we show that with any kind of correlated side information at the transmitters, if all the transmitter side information is also made available to the receiver, i.e. $(S_{T1}, S_{T2}) = f(S_R)$,then the MAC capacity region with causal side information is identical to the capacity region with non-causal side information. In both cases, the capacity region is given by the convex hull of the following rate pairs:

$$\begin{aligned} R_1 &\leq I(X_1; Y | S_R X_2) \\ R_2 &\leq I(X_2; Y | S_R, X_1) \\ R_1 + R_2 &\leq I(X_1, X_2; Y | S_R)\} \end{aligned}$$

for all $P(X_1|S_{T1}), P(X_2|S_{T2})$. Achievability is easily established with multiplexed codebooks. A sketch of the proof of converse is as follows.

$$\begin{aligned} nR_1 &\leq I(W_1; Y^n, S_R^n | W_2) + n\epsilon \\ &= \sum_{i=1}^{n} I(W_1; Y_i | W_2, Y^{i-1}, S_R^n) + n\epsilon \\ &\leq \textstyle\sum_{i=1}^{n} H(Y_i | S_{R,i}, X_{2,i}) - H(Y_i | W_2, W_1, S_R^n, Y^{i-1}, X_{1,i}, X_{2,i}) \\ &= \sum_{i=1}^{n} H(Y_i | S_{R,i}, X_{2,i}) - H(Y_i | X_{1,i}, X_{2,i}, S_{R,i}) + n\epsilon \\ &= \sum_{i=1}^{n} I(X_{1,i}; Y_i | X_{2,i}, S_{R,i}) + n\epsilon \end{aligned}$$

The converse for R_2 follows similarly. Finally, for the sum rate we have:

$$\begin{aligned} n(R_1 + R_2) &\leq I(W_1, W_2; Y^n | S_R^n) + n\epsilon \\ &= \textstyle\sum_{i=1}^{n} H(Y_i | S_R^n) - H(Y_i | W_1, W_2, Y^{i-1}, S_R^n) + n\epsilon \\ &\leq \textstyle\sum_{i=1}^{n} H(Y_i | S_{R,i}) - H(Y_i | X_{1,i}, X_{2,i}, S_{R,i}) + n\epsilon \\ &\leq \sum_{i=1}^{n} I(X_{1,i}, X_{2,i}; Y_i | S_{R,i}) + n\epsilon \end{aligned}$$

A. Advantage of Non-causal Side Information over Causal Side Information

We compare the causal and non-causal capacity regions in terms of the sum rate point C_Σ. Similar to the single user case, we have shown that the sum capacities (and the entire capacity regions) are identical when S_{T1}, S_{T2} are also available to the receiver. To make this information available to the receiver requires $H(S_{T1}, S_{T2}|S_R)$ Genie bits per symbol. And because we have shown for the multiple access channel that Genie bits can not improve capacity by more than their own entropy, we have the following result: $C_\Sigma(S_{T1}, S_{T2}, (S_R, S_{T1}, S_{T2})) \geq C_\Sigma^{\text{non-causal}}(S_{T1}, S_{T2}, S_R) \geq C_\Sigma^{\text{causal}}(S_{T1}, S_{T2}, S_R) \geq C_\Sigma(S_{T1}, S_{T2}, (S_R, S_{T1}, S_{T2})) - H(S_{T1}, S_{T2}|S_R)$.

B. Example: Multiple Users Random Access Channel

Consider the random access channel described before, except now the channel input is controlled by two users as:

$$X = X_{T1} S + X_{T2}(1 - S).$$

Thus, a scheduler randomly allows user 1 or user 2 to access the channel in bursts of N symbols. As before, the switch state S is known to the transmitters but not to the receiver. If a Genie provides S to the receiver, the sum capacity is C_0. Using the same arguments as in the single user example, we have:

$$\begin{aligned} C_0 &= C_\Sigma(S, S, S) \\ &\geq C_\Sigma^{\text{non-causal}}(S, S, \phi) \\ &\geq C_\Sigma^{\text{causal}}(S, S, \phi) \\ &= C_\Sigma(\phi, \phi, \phi) \\ &\geq C_\Sigma(S, S, S) - \frac{1}{N} H(S) = C_0 - 1/N. \end{aligned}$$

[1]The problem with extending the single user approach appears to be that including Y^{i-1} into the auxiliary random variables makes them correlated.

VII. THE BROADCAST CHANNEL WITH SIDE INFORMATION

In the general broadcast channel, the transmitter has side information S_T and the two receivers have side information S_{R1} and S_{R2}. The capacity region of the general broadcast channel is not known even without side information. The best known achievable region is due to Marton [10] with auxiliary variables U, V and W. A binning approach can be used to incorporate side information into the achievable region. A limited extension has been reported by the author in [12] and a general extension with a complete proof is provided in an independent parallel work [11] as well. To avoid repetition, we provide only a brief intuitive explanation of the resulting expressions in this paper.

For simplicity of exposition let us consider a special achievable region with only auxiliary random variables U, V. Furthermore, since the receiver side information can be assumed to be a part of the received symbols Y_1 and Y_2 we will simplify the notation by not considering receiver side information explicitly,

$$\begin{aligned} C \;\supseteq \{ & (R_1, R_2) : R_1 \leq I(U;Y_1) - I(U;S) \\ & R_2 \leq I(V;Y_2) - I(V;S) \\ & R_1+R_2 \leq I(U;Y_1)+I(V;Y_2)-I(U;S)-I(V;S)-I(U;V|S) \\ & \qquad = I(U;Y_1)+I(V;Y_2)-I(U;V)-I(U,V;S)\} \end{aligned}$$

for all joint distributions of U, V, X conditioned on S.

The individual rate constraints are directly recognizable from the single user capacity expression with non-causal side information. The sum rate constraint can also be readily recognized as follows. The probability that an independently generated U sequence, an independently generated V sequence, and an independently generated S_T sequence are jointly typical is bounded below by $2^{-n[I(U;V)+I(U,V;S_T)-\delta(\epsilon)]} = 2^{-n[I(U;S_T)+I(V;S_T)+I(U;V|S_T)-\delta(\epsilon)]}$ where $\delta(\epsilon)$ goes to zero as ϵ goes to zero. This explains the subtractive terms in the sum rate constraint. The same idea carries over to the general achievable region with auxiliary random variables U, V and W.

With causal side information, the precise multiuser capacity regions for the physically degraded broadcast channel is shown to be [9]:

$$\begin{aligned} C \supseteq \{(R_1, R_2) \quad : \quad & R_1 \leq I(U;Y_1) \\ & R_2 \leq I(V;Y_2|U)\} \end{aligned}$$

for all distributions $P(U, V, X|S) = P(U)P(V|U)P(X|V, S)$.

A. Genie Bounds and the Relative Advantage of Non-causal Side Information over Causal Side Information

In order to provide the transmitter side information to the receivers, the Genie must provide $H(S_T|S_{R1})$ bits to receiver 1 and $H(S_T|S_{R2})$ bits to receiver 2. Therefore, using similar arguments as before, we can argue that the maximum achievable sum rate with non-causal transmitter side information can not exceed that with causal side information by more than $H(S_T|S_{R1}) + H(S_T|S_{R2})$ bits.

B. Example: Broadcast Channel

Consider a two user fading broadcast channel where the users' channels experience i.i.d. block fading (block size N) given by h_1 and h_2 respectively.

$$\begin{aligned} Y_1 &= h_1 X + n_1 \\ Y_2 &= h_2 X + n_2 \end{aligned}$$

where $n_1, n_2 \sim \mathcal{N}(0, 1)$ are additive white Gaussian noise terms and a transmit power constraint of P is in place. It is assumed that each user knows only his own channel, while the transmitter knows only the side information variable S_T ($S_T = 1$ if $h_1 > h_2$ and 0 otherwise). While in a degraded broadcast channel it is well known that the sum rate maximizing policy is to transmit only to the strongest user, the problem faced here is that a receiver does not know when it has the stronger channel compared to the other receiver's channel. Clearly, if a Genie provides S_T to both receivers then the sum capacity of this channel is $C_0 = C_\Sigma(S_T, (h_1, S_T), (h_2, S_T)) = \mathrm{E}\log(1 + P\max(|h_1|^2, |h_2|^2))$. Since the Genie's bits can not increase capacity by more than 2 bits (one bit to each receiver), the sum capacity (with causal or non-causal side information) can be bounded as:

$$C_0 \geq C_\Sigma^{\text{non-causal}}(S_T, h_1, h_2) \geq C_\Sigma^{\text{causal}}(S_T, h_1, h_2) \geq C_0 - 2/N.$$

REFERENCES

[1] S. I. Gelfand and M. Pinsker, "Coding for channel with random parameters," in *Problems of Control and Information Theory*, pp. 19–31, 1980.

[2] A. Kusnetsov and B. Tsybakov, "Coding in a memory with defective cells," *Prob. Pered. Inform.*, vol. 10, pp. 52–60, April-June 1974.

[3] C. Heegard and A. El Gamal, "On the capacity of computer memories with defects," *IEEE Trans. Inform. Theory*, vol. 29, pp. 731–739, Sep. 1983.

[4] C. E. Shannon, "Channels with side information at the transmitter," in *IBMJ. Res. Develop*, pp. 289–293, 1958.

[5] T. M. Cover and M. Chiang, "Duality between channel capacity and rate distortion with two-sided state information," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1629–1638, June 2002.

[6] M. Chiang and T. Cover, "Unified duality between channel capacity and rate distortion with state information," in *Proceedings of IEEE Int. Symp. Inform. Theory*, p. 301, June 2001.

[7] G. Caire and S. Shamai, "On the capacity of some channels with channel state information," *IEEE Trans. Inform. Theory*, vol. 45, pp. 2007–2019, Sept 1999.

[8] S. Jafar and S. Srinivasa, "Capacity limits of cognitive radio with distributed and dynamic spectral activity," *cs.IT/0509077*, 2005. Download available at http://arxiv.org.

[9] S. Sigurjonsson and Y. Kim, "On multiple user channels with causal state information at the transmitters," *cs.IT/0508096*, 2005. Download available at http://arxiv.org.

[10] K. Marton, "A coding theorem for the discrete memoryless broadcast channel," *IEEE Trans. Inform. Theory*, vol. 25, pp. 306–311, May 1979.

[11] Y. Steinberg and S. Shamai, "Achievable rates for the broadcast channel with state known at the transmitter," in *Proceedings of IEEE Int. Symp. Inform. Theory*, 2005.

[12] S. Jafar, "Degrees of freedom in distributed mimo communication," *IEEE Communication Theory Workshop*, June 2005.