ISIT 2009, Seoul, Korea, June 28 - July 3, 2009

On the Tightness of Marton's Regions for Semi-Additive Broadcast Channels[1,2]

Eli Haim
Dept. Electrical Engineering - Systems,
Tel-Aviv University,
Tel-Aviv, Israel
Email: elih@eng.tau.ac.il

Ram Zamir
Dept. Electrical Engineering - Systems,
Tel-Aviv University,
Tel-Aviv, Israel
Email: zamir@eng.tau.ac.il

***Abstract*—We study cost constrained side-information channels, where the cost function depends on a state which is known only to the encoder. In the additive noise case, we bound the capacity loss due to not knowing the cost state at the decoder and show that it is small under various assumptions, and goes to zero in the limit of weak noise. This model plays an important role in the (non-degraded) broadcast channel. In the semi-additive noise case, we bound the gap between the best known single letter achievable region and the true capacity region, using tools developed for the first problem. In the limit of weak noise, we show that the bounds coincide, thus we get the complete characterization of the capacity region.**

***Index Terms*—Side-information channels, Broadcast channels.**

I. Introduction

Channel coding where the encoder has access (non-causally) to channel state side information, plays an important role in many different problems. It's capacity region was found by Gelfand and Pinsker [2].

An important special case of the this problem is when the channel state side information constrains the channel inputs, but does not affect the channel transition probabilities. This case models many problems. One is the information embedding problem [3], where the encoder should encode information over a "host" signal such that the transmitted signal is not too far from the original one. A different type of such problem is the case where the state constrains the possible channel input alphabet [4]. A classical example for that is writing to a memory with defective cells [5][6], where the information whether a memory cell is writable or not is available at the encoder but not at the decoder. Some channels where the channel transitions are affected by the state side information can be reduced to this model, for example *writing on dirty paper* [7], viewed as information embedding.

This problem is conceptually similar to source coding when the distortion measure is state-dependent and the state is known only to the encoder [8] [4]. The latter problem can be solved using interpolation-based coding (in the case of a difference distortion measures which is weighted by the side information state) and a similar coding technique can be used here.

We study the capacity loss due to not knowing the side-information at the decoder and derive bounds for the additive noise case. We then apply these observations and techniques to explore the gap between the best known achievable region and the true capacity region of the (non-degraded) broadcast channel [9].

Transmission over a broadcast channel can be viewed under the model of channel state side information, such that the state only constrains the input. A broadcast channel is a channel with a single input and two or more outputs. Information is sent through the channel, such that each output is decoded separately, i.e., the decoder of each channel is unaware of the other channel output, so different receivers cannot cooperate. The objective is to transmit a different private message to each of the receivers. Generally, there may be a common message to groups of terminals, but in this paper we assume that it is not required. The role of the side information problem can be seen by the following scheme: information is sent to the first terminal, and the encoded word constitutes the "channel state" that constrains the encoding to the second terminal.

The capacity region of the broadcast channel is still an open problem in the general (non-degraded) case. However achievable regions [9] and outer regions [9][10][11] to the capacity region are known. We call the gap between the best known achievable region and the capacity region, the *unresolved gap*.

Using the similarity between side-information channels and broadcast channels, similar codebooks are constructed, a similar bounding technique is used in both cases, and similar bounds can be derived.

In section II we describe the channel models that are studied in this paper, and discuss their similarity. Section III describes channels where the channel output is known to the encoder up to an additive noise component. We derive general upper bounds for the capacity loss due to not knowing the side information at the decoder, and show that the loss is asymptotically zero for weak noise. In this section we introduce two bounds which are used for the capacity loss in side information channel, and in the next section for the unresolved gap in the broadcast channel. Section IV explores

[1]This research was partially supported by the Bi-National Science Foundation, Grant 2004398.

[2]The results in this paper were partially derived and presented in [1].

broadcast channels which consist of a general channel in one branch, and a deterministic function followed by an additive noise channel at the second branch. We show upper bounds on the gap between the best known single letter achievable region and the true capacity region, and show that this gap vanishes for weak noise. Section V presents further results that appear in [12].

II. Problem Models and Their Similarity

A. Cost State Side Information Problem

Consider the channel $P_{Y|X}$ under the constraint $E[c(X;S)] \leq A$, where X is the channel input, Y is the channel output, $P_{Y|X}$ is the channel probability transition function, S is the cost state side-information, $c(x;s)$ is the channel input cost function which depends on the channel input and the cost state and A is the average cost constraint. We assume that S is known (non-causally) only to the encoder. Therefore the triple $S_1^n \leftrightarrow X_1^n \leftrightarrow Y_1^n$ forms a Markov chain, where n is the codeword length. We call this channel a *cost side information* (or cost-SI) channel. This channel is a special case of the side-information channel considered by Gelfand and Pinsker [2], where the state S affects the channel transition distribution. Nevertheless the capacity formula still applies so, denoted by C_{enc}, it is given by [3]:

$$C_{\text{enc}}(A) = \sup_{P_{U,X|S}:E[c(X;S) \leq A]} I(Y;U) - I(U;S), \quad (1)$$

where U is an auxiliary random variable.

The cost-SI channel is formally equivalent to the *information embedding* problem [3]. In the latter problem S is a signal which is known to the encoder, and information should be embedded over it such that the average distortion of the signal is bounded by a given constraint A. Therefore S is called the host signal. The distortion measure is a function of both X and S: $c(x;s)$, and it measures the similarity between the host signal and the encoder output signal (e.g. $(x-s)^2)$). In our problem the emphasis is on the cost of using the symbol x at state s, $c(x;s)$, therefore S may not be a host signal, but a weighting state for the cost (e.g. $w(s)x^2$). This difference in emphasis leads to new observations, coding techniques and explicit capacity expressions. The cost-SI channel was first introduced in [4] and further explored in [13].

We compare this problem to the case where the side information is known to both the encoder the decoder. We call that *the fully aware system*, and denote it's capacity by C_{both}. When S is known at both sides, we can use separate codebooks for different states of S, therefore the capacity is given by [14] [3]:

$$C_{\text{both}}(A) = \sup_{P_{X|S}:E[c(X;S) \leq A]} I(Y;X|S). \quad (2)$$

We explore the capacity loss due to not knowing the side-information at the decoder

$$Loss(A) = C_{\text{both}}(A) - C_{\text{enc}}(A). \quad (3)$$

Some channels where the transition distribution depends on S, i.e. $P_{Y|X,S}$, fall under the category of cost-SI channels. One example is the "Writing on Dirty Paper" problem [7], there $Y = X + S + Z$ which when viewed as an information embedding problem has the form of a cost-SI channel [3], [15]. Detailed explanation is given in Section III.

In the discrete case, when the channel is deterministic, $Y = f(X)$ and S constrains the channel input alphabet, there is no capacity loss due to not knowing the cost side information at the decoder [3]. The same holds for the general side-information problem, where the encoder knows the complete channel state, i.e. the channel is of the form $Y = f(X,S)$, when S is given at the encoder. This fact can be seen by substituting $U = Y$ in the C_{enc} expression in (1).

An example for this model, is coding to memory with defective cells [5][6]. In this problem the channel state constrains the input alphabet according to the following states: stuck-at 0, stuck-at 1 or $\{0,1\}$. This side information contains the complete channel state, and is known only to the encoder. Reducing this problem to a cost-SI channel is done by adding a cost function and a constraint which will restrict the channel input alphabet in the case of stuck-at state. There is no capacity loss due to not knowing the side information state at the decoder [5][6]. A coding scheme that achieves the capacity is based on the cosets of a "good" binary erasure correction (near MDS) code.

In this paper we focus on channels which are known to the encoder up to an additive noise component, i.e. of the form $Y = f(X,S) + Z$, with input cost function which depends on S and an average cost constraint A. We show that these channels can be reduced to the cost-SI model. As we shall see in the next section, in the noisy case the capacity loss is not zero, yet it often can be bounded by a small term.

B. Broadcast Channels

A broadcast channel has a single input and two outputs (see [16]). Information is sent through the channel, such that each output is decoded separately, i.e. the decoder of one channel output is unaware of the other channel output. A broadcast channel can be defined by $P_{Y_1,Y_2|X}$, where X is the channel input and Y_1, Y_2 are two outputs of the channel. The capacity region of the broadcast channel is yet unknown, except for special cases e.g. the degraded case, therefore we consider inner regions and outer regions to the capacity region.

The best known single-letter achievable region was given by Marton in [9]. In this paper we use an weaker inner region (without the common part), also given by Marton [9]:

$$\begin{aligned} \mathfrak{R}_0^{\text{M}} = \{(R_1,R_2) | \exists U_1, U_2 : & \\ & (U_1,U_2) \leftrightarrow X \leftrightarrow Y_1, \\ & (U_1,U_2) \leftrightarrow X \leftrightarrow Y_2, \\ & 0 \leq R_1 \leq I(Y_1;U_1), \\ & 0 \leq R_2 \leq I(Y_2;U_2), \\ & R_1 + R_2 \leq I(Y_1;U_1) + I(Y_2;U_2) - I(U_1;U_2)\} \end{aligned} \quad (4)$$

Channel with cost-SI known at the encoder plays an important role in Marton's achievable region. Assume a given probability distribution $P_{U_1,U_2,X}$. At the corner point, information is encoded freely to the first terminal in maximum rate, $R_1 = I(Y_1;U_1)$. This constitute the "channel state" that constrains the encoding to the second terminal. Therefore $R_2 = I(Y_2;U_2) - I(U_1;U_2)$ which is similar to the expression in (1), where the information to the first terminal U_1 acts as S, the "side information" that constrains the encoding to the second terminal.

We explore the gap between this region and the capacity region:

$$GAP = \max_{(R_1,R_2)\in\mathcal{C}} \min_{(r_1,r_2)\in\mathfrak{R}_0^{\text{M}}} \max\{R_1 - r_1, R_2 - r_2\} \quad (5)$$

where $\mathcal{C}$ is the broadcast channel capacity region, i.e. for each point in the true capacity region $\mathcal{C}$ we look for the point in $\mathfrak{R}_0^{\text{M}}$ which is closest in both rate components.

We bound this gap by bounding the gap between Marton's achievable region (4) and an outer bound to the capacity region due to Marton [9]:

$$\begin{aligned}\mathfrak{R}_{\text{out}}^{\text{K-M}} \quad = \quad & \{(R_1,R_2)|\exists U_1 : \\ & U_1 \leftrightarrow X \leftrightarrow Y_1, \\ & U_1 \leftrightarrow X \leftrightarrow Y_2, \\ & 0 \le R_1 \le I(Y_1;U_1), \\ & 0 \le R_2 \le I(Y_2;X), \\ & R_1 + R_2 \le I(Y_1;U_1) + I(Y_2;X|U_1)\ \} . \quad (6)\end{aligned}$$

The fully aware channel (see Section II-A) plays an important role in this region. Information is encoded to the first terminal freely. As in $\mathfrak{R}_0^{\text{M}}$, this constitute the "channel state" that constrains the encoding to the second terminal. However, here it is assumed that the decoder of the second terminal knows the information which was encoded for the first terminal, which plays the role of the "channel state". The rate for the first terminal is $R_1 = I(Y_1;U_1)$, and the rate for the second terminal is $R_2 = I(Y_2;X|U_1)$ which is similar to (2).

The similarity described above between the broadcast channel and the side-information channel, leads to similarity of the gap between the achievable region and the outer region of broadcast channel, to the capacity-loss in the side-information channel due to not knowing the side information at the decoder. Both expressions are of the form:

$$I(Y;X|S) - [I(Y;U) - I(U;S)]. \quad (7)$$

In this paper we focus on semi-additive noise channels, i.e. channels where one branch is general, while the second branch has the form of $Y_2 = f(X) + Z$, where f is a deterministic function and Z is the noise. Using similar bounding technique for the expression (7) in Lemma 2, we get similar bounds for the gap between the regions of this broadcast channel and for the capacity loss in side information channel.

A special case of a broadcast channel is the (discrete) deterministic broadcast channel (DBC) (see [16]). In this case Marton [9] and Pinsker [17] found the capacity region, and Marton showed that it equals $\mathfrak{R}_0^{\text{M}}$ and $\mathfrak{R}_{\text{out}}^{\text{K-M}}$.

A simple but not trivial example to a DBC is the Blackwell channel (see [16]). Gelfand [18] found the capacity region of this channel. His coding scheme is similar to one which is used for the binary case of writing on memory with defective cells (see [4]).

This work strengthens this result of the noiseless case, by showing that when the noise is weak, the inner region and the outer region to the capacity region coincide, thus characterizing the capacity region in this case. Thus the gap between the regions is a continuous function near the point where the channel is noiseless.

III. Additive Noise Channels with Cost Side Information

In this section we consider an important special case of side information channels, where the encoder knows everything about the channel up to an additive noise component, i.e.

$$Y = f(X,S) + Z \quad (8)$$

with the following constraint:

$$E[c(X;S)] \le A, \quad (9)$$

where X is the channel input, S is the channel state side information, $c(x;s)$ is the channel input cost function which depends on the state, A is the average cost constraint on the channel input and Z is an additive noise which is independent of (X,S). We assume that S is known (non-causally) only to the encoder. We call f the channel transfer function.

Lemma 1. *The channel model given in* (8)-(9) *is equivalent to the following cost-SI model:*

$$Y = \tilde{X} + Z \quad (10)$$

under the constraint $E[\tilde{c}(\tilde{X};S)] \le A$*, where* $\tilde{c}(\tilde{x};s)$ *is some modified cost function.*

Proof: Set $\tilde{X} = f(X,S)$ and modify the cost function

$$\tilde{c}(\tilde{x},s) = \begin{cases} \min\limits_{x:f(x,s)=\tilde{x}} c(x;s) & \exists x : f(x,s) = \tilde{x} \\ \infty & \forall x : f(x,s) \ne \tilde{x} \end{cases}, \quad (11)$$

■

Theorem 1 (Upper Bound on the Capacity Loss)**.** *For the channel given in* (8)-(9)*, the capacity loss (which is given in* (3)*) due to not knowing the side information at the decoder too is upper bounded by:*

$$Loss(A) \le I(Z - \tilde{Z};\tilde{Z}) \quad (12)$$

where $\tilde{Z}$ *has the same probability distribution function as* Z*, and* $\tilde{Z}, Z$ *are independent R.V.s.*

For example, if Z is a Gaussian noise, then the loss is at most $\frac{1}{2}$ bit. An additional example, if $Z \sim$ Bernulli(p), then the loss is at most $H(p * p) - H(p)$, which is upper bounded by $0.2144\ldots$ bits (where $p*p=2p(1-p)$, and the maximal bound is at $p=0.121\ldots$).

The theorem is proved using the following lemma, which will also be used for the theorem which bounds the unresolved gap for broadcast channels.

Lemma 2. *Consider random variables X,S,Z, such that Z is independent of (X,S), and denote $Y \triangleq X+Z$. Let $\tilde{Z}$ be a random variable with the probability distribution function of Z, and independent of (X,S,Z). Denote $U \triangleq X+\tilde{Z}$. Then:*

$$I(Y;X|S)-[I(Y;U)-I(U;S)] \le I(Z-\tilde{Z};\tilde{Z}) \quad (13)$$

Proof:

$$\begin{aligned} & I(Y;X|S)-[I(Y;U)-I(U;S)] && (14)\\ &= I(X+Z;X,S)-I(X+Z;X+\tilde{Z}) && (15)\\ &= I(X+Z;X)-I(X+Z;X+\tilde{Z}) && (16)\\ &= I(X+Z;X+\tilde{Z}|\tilde{Z})-I(X+Z;X+\tilde{Z}) && (17)\\ &= I(Z-\tilde{Z};\tilde{Z}|X+\tilde{Z}) && (18)\\ &= I(Z-\tilde{Z};\tilde{Z})-I(Z-\tilde{Z};X+\tilde{Z}) && (19)\\ &\le I(Z-\tilde{Z};\tilde{Z}) && (20) \end{aligned}$$

■

proof of Theorem 1: Using Lemma 1, without loss of generality we can consider only the case where $f(X,S)=X$.

Set the auxiliary R.V. to be $U \triangleq X+\tilde{Z}$, where channel input X has the probability distribution that achieves the capacity $C_{\text{both}}(A)$ in the fully-aware case, and $\tilde{Z}$ has the same probability distribution as Z and is independent of (X,S,Z).

Denote by r_{enc} the achievable rate using these U,X in the system where S is known only at the encoder. Then by using Lemma 2 we get: $Loss(A) \le C_{\text{both}}(A)-r_{\text{enc}}(A) \le I(Z-\tilde{Z};\tilde{Z})$. ■

Let us define:

$$h_{\max}(A|S) \triangleq \max_{P_{X|S}:E[c(X,S)]\le A} h(X|S). \quad (21)$$

Theorem 2 (Asymptotic Zero Loss)**.** *Consider the channel given in* (10)*, i.e., $Y=X+Z$ under the constraint $E[c(X,S)] \le A$. In the limit of weak noise, we look at a sequence of channels $Y=X+Z_n$, where $\{Z_n\}_{n=1}^{\infty}$ is a sequence of RVs such that $E(Z_n^2) \underset{n\to\infty}{\longrightarrow} 0$. Assume that the conditional maximum-entropy* (21) *exists and is continuous in the parameter A (see [19]). In addition, assume that $E[c(X+Z_n,S)] \underset{n\to\infty}{\longrightarrow} E[c(X,S)]$ uniformly over all X such that $E[c(X,S)] \le A$. Then the capacity loss (which is given in* (3)*) due to not knowing the side information at the decoder too is asymptotically zero for weak noise:*

$$\lim_{n\to\infty} Loss(A)=0, \quad (22)$$

Remark 1.
1) *This result agrees with the fact that in the noiseless side-information problem, there is no capacity loss due to not knowing the side information at the decoder too [3]. Therefore the capacity as a function of the noise power is continuous near zero.*
2) *This result applies to the model* (8)-(9) *under the transformation in Lemma 1.*
3) *A simple example where uniform convergence holds is for $c(x,s)=g(s)|x|^r$, and r is integer.*

proof of Theorem 2: Denote the achieving RV of $h_{\max}(A|S)$ by X^*. Denote the fully aware system capacity achieving distribution, $C_{\text{both}}^{(n)}$, by $X_n^{(b)}$. Define $A_n \triangleq E[c(X_n^{(b)}+Z_n,S)]$. Denote the capacity of the encoder side information system by $C_{\text{enc}}^{(n)}$.Then:

$$\begin{aligned} C_{\text{both}}^{(n)}-C_{\text{enc}}^{(n)} &\le [h(X_n^{(b)}+Z_n|S)-h(Z_n)] && (23)\\ &\quad -[h(X^*|S)-h(X^*|X^*+Z_n)] \\ &\le h_{\max}(A_n|S)-h_{\max}(A|S) \underset{n\to\infty}{\longrightarrow} 0, && (24) \end{aligned}$$

where the convergence results from the continuity conditions in the theorem. ■

Example 1. *The "Writing on Dirty Paper" problem [7] is as follow: the channel is $Y=X+S+Z$, where X is the channel input, S is a channel interference which is known non-causally at the encoder, and Z is a Gaussian noise, independent of X,S. The channel has a power constraint: $E(X^2)\le P$. It is well known [7] that there is no rate-loss due to not knowing S at the decoder. In this example we generalize the problem such that the cost function is weighted according to a state Q. The motivation is from the information embedding problem, where the distortion measure function of each sample may be different, according to it's importance. This problem was proposed in [4], where preliminary results were presented, and further explored in [13]. The channel is $Y=X+Z$ with the constraint is $E[Q\cdot(X-S)^2]\le D$. According to Theorem 2, the capacity loss due to not having S,Q at the decoder is asymptotically zero as the noise gets small ($\sigma_Z^2 \to 0$). This result was shown in [13], and the result here generalizes it. In addition, according to Theorem 1, the capacity loss due to not having S,Q at the decoder is at most $\frac{1}{2}$ bit.*

IV. Broadcast Channel with One Additive Noise Component

We consider the broadcast channel where the transition probabilities to the first terminal are given by:

$$P_{Y_1|X} \quad (25)$$

and the output to the second terminal is given by:

$$Y_2=f(X)+Z, \quad (26)$$

where Z is an additive noise which is independent of X.

Theorem 3 (Upper Bound on the Gap)**.** *For the broadcast channel given in* (25)-(26)*, the gap between $\mathfrak{R}_0^{M}$ and the capacity region (which is given at* (5)*) is upper bounded by:*

$$GAP \le I(Z-\tilde{Z};\tilde{Z}), \quad (27)$$

where $\tilde{Z}$ has the same probability distribution function as Z, and $\tilde{Z},Z$ are independent R.V.s.

Proof: Let $(R_1, R_2) \in \mathfrak{R}^{\text{K-M}}_{\text{out}}$ be a given point. Without loss of generality we can assume that the point is on the (non-trivial) boundary of $\mathfrak{R}^{\text{K-M}}_{\text{out}}$. Therefore there exists a $P_{U_1,X}$ which achieves (R_1, R_2) point according to (6).

Define $U_2 = f(X) + \tilde{Z}$, where $\tilde{Z}$ has the same probability distribution as Z, and is independent of (X, Z). The sum-rate gap is given by the following expression: $I(Y_2; f(X)|U_1) - [I(Y_2; U_2) - I(U_2; U_1)]$. Lemma 2, where X there has the role of $f(X)$ here, shows that this expression is upper bounded by $I(Z - \tilde{Z}; \tilde{Z})$.

The gap between the rates for the second terminal is given by the following expression: $I(Y_2; X) - I(Y_2; U_2)$, which is equivalent to (16), and therefore also upper bounded by $I(Z - \tilde{Z}; \tilde{Z})$. ■

Theorem 4 (Asymptotically Zero Gap)**.** *Consider the broadcast channel given in* (25)-(26)*, In the limit of weak noise, we look at a sequence of channels $Y_2 = f(X) + Z_n$, where $\{Z_n\}_{n=1}^{\infty}$ is a sequence of RVs such that $E(Z_n^2) \underset{n\to\infty}{\longrightarrow} 0$. Let $(R_1, R_2) \in \mathfrak{R}^{\text{K-M}}_{\text{out}}$ be a given point on the (non-trivial) boundary of $\mathfrak{R}^{\text{K-M}}_{\text{out}}$. For every n denote the (R_1, R_2) achieving prior by $P_{U_{1,n},X_n}$. Assume that that there exist a RV pair (X, U_1) such that $(X_n, U_{1,n}) \underset{n\to\infty}{\overset{D}{\longrightarrow}} (X, U_1)$. In addition, for $\{(X_n, U_{1,n})\}_{n=1}^{\infty}$ and (X, U_1) we assume that their p.d.f.s are uniformly bounded and that there exists some $k > 1$ such that their k-th moments are uniformly bounded. Then the distance of (R_1, R_2) from $\mathfrak{R}^{\text{M}}_0$ goes to zero when $n \to \infty$. If this property holds for every point (R_1, R_2) on the boundary of $\mathfrak{R}^{\text{K-M}}_{\text{out}}$ then the gap between $\mathfrak{R}^{\text{M}}_0$ and the capacity region (which is given at* (5)*) is asymptotically zero in weak noise:*

$$GAP \underset{n\to\infty}{\longrightarrow} 0. \tag{28}$$

Remark 2. *An immediate conclusion from the this theorem is that in small noise regime, $\mathfrak{R}^{\text{M}}_0$ is practically the capacity region. This result shows that the unresolved gap, as a function of the noise power, is a continuous function near the point where the channel is semi-noiseless. This result in the discrete case agrees with the fact that in semi-deterministic broadcast channel $\mathfrak{R}^{\text{M}}_0 = \mathfrak{R}^{\text{K-M}}_{\text{out}}$ [9] [20].*

Proof: Define $U_{2,n} = f(X_n)$. The sum-rate gap is given by the following expression:

$$\begin{aligned}
&I(Y_{2,n}; f(X_n)|U_{1,n}) - [I(Y_{2,n}; U_{2,n}) - I(U_{2,n}; U_{1,n})] \\
&= I(U_{1,n}; Z_n|f(X_n) + Z_n) \\
&= I(f(X_n) + Z_n; Z_n|U_{1,n}) - I(f(X_n) + Z_n; Z_n) \\
&\le I(f(X_n) + Z_n; Z_n|U_{1,n})
\end{aligned}$$

This mutual information goes to zero as $n \to \infty$ [21].

The gap between the rates for the second terminal is given by the following expression: $I(Y_{2,n}; X_n) - I(Y_{2,n}; U_{2,n}) = I(Y_{2,n}; X_n|U_{2,n}) = 0$. ■

V. Further Results

A stronger connection between the capacity loss in side-information channels and the unresolved gap in broadcast channel can be established [12] using a tighter outer bound to the capacity region of the broadcast channel, which appears in [10]. Using this connection, bounds on the unresolved gap in broadcast channels can be derived directly from bounds on the capacity loss in side-information channels [12].

An additional result, which is already known, however directly results from the similarity to side-information problem, is that in Marton's achievable region $\mathfrak{R}^{\text{M}}_0$, it is enough to consider only a deterministic functions $X = f(U_1, U_2)$ as the channel input [12] (using [3, Lemma 2]).

References

[1] E. Haim and R. Zamir, "Broadcast channels and input-cost side information," in *IEEE 24th Convention of Electrical and Electronics Engineers in Israel*, November 2006.

[2] S. I. Gelfand and M. S. Pinsker, "Coding for channel with random parameters," *Probl. Contr. Inform. Theory*, vol. 9, No. 1, pp. 19–31, Jan.-Mar. 1980.

[3] R. J. Barron, B. Chen, and G. W. Wornell, "The duality between information embedding and source coding with side information and some applications," *IEEE Trans. Information Theory*, vol. IT-49, pp. 1159–1180, May 2003.

[4] E. Haim and R. Zamir, "Quantization with variable resolution and coding for deterministic broadcast channels," in *Allerton Conference on Communication, Control and Computing*, Oct. 2005.

[5] A. V. Kuznetsov and B. S. Tsybakov, "Coding in memory with defective cells," *Problemy Pered. Inform. (Problems of Inform. Trans.)*, vol. 10, No. 2, pp. 52–60, 1974.

[6] C. Heegard and A. E. Gamal, "On the capacity of computer memory with defects," *IEEE Trans. Information Theory*, vol. IT-29, pp. 731–739, Sept. 1983.

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

[8] E. Martinian, G. W. Wornell, and R. Zamir, "Source coding with distortion side information," *IEEE Trans. Information Theory*, vol. IT-54, pp. 4638–4665, Oct. 2008.

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

[10] C. Nair and A. E. Gamal, "An outer bound to the capacity region of the broadcast channel," in *Proc. Int. Symp. Info. Theory*, July 2006.

[11] Y. Liang, G. Kramer, and S. Shamai, "Capacity outer bounds for broadcast channels," in *Proceedings of IEEE Information Theory Workshop, ITW 2008, Porto, Portugal*, May 2008.

[12] E. Haim, "Input cost side-information and broadcast channels," Master's thesis, Tel-Aviv University, October 2007.

[13] A. Khisti, E. Martinian, and G. Wornell, "Information embedding with distortion side information," in *Proc. Int. Symp. Info. Theory*, July 2006.

[14] J. Wolfowitz, *Coding Theorems of Information Theory*. New York: Springer-Verlag, 1964.

[15] A. S. Cohen and A. Lapidoth, "The Gaussian watermarking game," *IEEE Trans. Information Theory*, vol. IT-48, pp. 1639–1667, June 2002.

[16] T. M. Cover, "Comments on broadcast channels," *IEEE Trans. Information Theory*, vol. IT-44, pp. 2524–2530, October 1998.

[17] M. S. Pinsker, "Capacity of noiseless broadcast channels," *Problemy Pered. Inform. (Problems of Inform. Trans.)*, vol. 14, No. 2, pp. 28–34, Apr.-June 1978.

[18] S. I. Gelfand, "Capacity of one broadcast channel," *Problemy Pered. Inform. (Problems of Inform. Trans.)*, vol. 13, No. 3, pp. 106–108, July-Sept. 1977.

[19] T. Linder and R. Zamir, "On the asymptotic tightness of the Shannon lower bound," *IEEE Trans. Information Theory*, vol. IT-40, pp. 2026–2031, Nov. 1994.

[20] S. I. Gelfand and M. S. Pinsker, "Capacity of a broadcast channel with one deterministic component," *Problemy Pered. Inform. (Problems of Inform. Trans.)*, vol. 16, No. 1, pp. 24–34, Jan.-Mar. 1980.

[21] M. Godavari and A. Hero, "Convergence of differential entropies," *IEEE Trans. Information Theory*, vol. IT-50, pp. 171–176, Jan. 2004.