Channel Code Design with Causal Side Information at the Encoder

Hamid Farmanbar, Shahab Oveis Gharan, and Amir K. Khandani
Coding and Signal Transmission Laboratory
Department of Electrical and Computer Engineering
University of Waterloo
Waterloo, Ontario, N2L 3G1
Email: {hamid,shahab,khandani}@cst.uwaterloo.ca

***Abstract*— The problem of channel code design for the M-ary input AWGN channel with additive discrete interference where the sequence of i.i.d. interference symbols is known causally at the encoder is considered. The code design criterion at high SNR is derived by defining a new distance measure between the input symbols of the Shannon's *associated* channel. For the case of binary-input channel, it is shown that it is sufficient to use only two symbols of the associated channel in the encoding as long as the distance spectrum of the code is concerned. This reduces the problem of code design for the binary-input AWGN channel with known interference to the design of binary codes for the AWGN channel with maximum Hamming distance.**

I. Introduction

Information transmission over channels with known interference at the transmitter has recently found applications in various communication problems such as digital watermarking [1] and broadcast schemes [2]. A remarkable result on such channels was obtained by Costa who showed that the capacity of additive white Gaussian noise (AWGN) channel with additive Gaussian i.i.d. interference, where the sequence of interference symbols is known non-causally at the transmitter, is the same as the capacity of AWGN channel [3]. Therefore, the interference does not incur any loss in the capacity. This result was extended to arbitrary (random or deterministic) interference in [4] by using a scheme based on multi-dimensional lattice quantization. The result obtained by Costa does not hold for the case that the sequence of interference symbols is known causally at the transmitter.

Following Costa's "Writing on Dirty Paper" famous title [3], when the interference is known non-causally at the transmitter, the channel is refereed to as "dirty paper" channel.

Recently, dirty paper coding has emerged as a building block in multiuser communication. In particular, there has been considerable research studying the application of dirty paper coding to broadcast over multiple-input multiple-output (MIMO) channels [2].

These developments motivate finding realizable dirty paper coding techniques. Building upon [4], Erez and ten Brink [5] proposed a practical code design based on vector quantization via trellis shaping and using powerful channel codes. Due to complexity of implementation, their scheme uses the knowledge of interference up to six future symbols rather than the whole interference sequence. Bennatan *et al.* [6] gave another design based on superposition coding and successive cancelation decoding. Their design uses a trellis coded quantizer with memory nine and LDPC as channel code.

The schemes that use the interference (or the host signal in watermarking applications) sequence up to the current symbol can be used as a low-complexity solutions for the dirty paper problem. For example, in [1], scalar lattice quantization is proposed for data-hiding even though in that context, the host signal in clearly known non-causally.

In this paper, we consider the problem of code design for the M-ary input AWGN channel with additive causally-known discrete interference. The discrete model for interference is more appropriate in many applications. For example, in the MIMO broadcast channel, the interference caused by the other users is discrete rather than continuous.

Our design does not rely on the suboptimal (in terms of capacity) scheme of scalar lattice quantization for the causally known interference. Instead, we consider code design for the Shannon's *associated* channel with all possible mappings from the interference alphabet to the channel input alphabet. Another distinction between our work and the related research in the field is that we consider a finite channel input alphabet rather than a continuous one.

This paper is organized as follows. In the next section, we summarize Shannon's work on channels with causal side information at the transmitter. In section III, we introduce the channel model. In section IV, we derive the code design criterion for AWGN channel with causally-known discrete interference at the encoder. In section V, we consider channels with binary input for which we show that the design criterion derived in section IV reduces to maximizing the Hamming distance. We conclude this paper in section VI.

II. Channels with Side Information at the Transmitter

Channels with known interference at the transmitter are special case of channels with side information at the transmitter which were considered by Shannon [7] in the causal

[1]This work was supported by the Nortel Networks, the Natural Sciences and Engineering Research Council of Canada (NSERC), and the Ontario Center of Excellence (OCE).

1-4244-0769-9/07/$20.00 © 2007 IEEE.

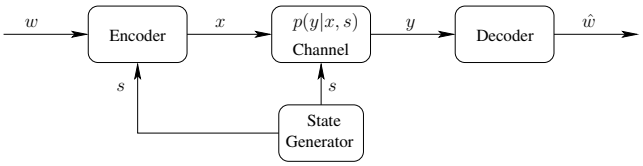

Fig. 1. SD-DMC with state information at the encoder.

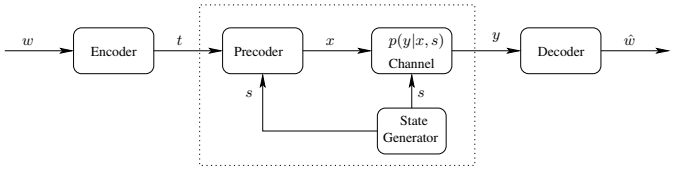

Fig. 2. The associated regular DMC.

knowledge setting and by Gel'fand and Pinsker [8] in the non-causal knowledge setting.

Shannon considered a discrete memoryless channel (DMC) whose transition matrix depends on the channel state. A state-dependent discrete memoryless channel (SD-DMC) is defined by a finite input alphabet $\mathcal{X}$, a finite output alphabet $\mathcal{Y}$, and transition probabilities $p(y|x,s)$, where the state s takes on values in a finite alphabet $\mathcal{S}$. The block diagram of a state-dependent channel with state information at the encoder is shown in fig. 1.

In the causal knowledge setting, the encoder maps a message w into $\mathcal{X}^n$ using functions

$$x_i = f_i\left(w, s_1, \ldots, s_i\right), \quad 1 \leq i \leq n. \tag{1}$$

Shannon showed that it is sufficient to consider the coding schemes that use only the current state symbol in the encoding process to achieve the capacity of an SD-DMC with i.i.d. state sequence known causally at the encoder [7].

The SD-DMC can be used in the way shown in fig. 2 to transmit information. A precoder is added in front of the SD-DMC. A message w is mapped into $\mathcal{T}^n$, where $\mathcal{T}$ is a new alphabet. The output of the precoder ranges over $\mathcal{X}$ and depends on the current interference symbol. The regular (without state) channel from T to Y is defined by the transition probabilities

$$p(y|t) = \sum_{s\in\mathcal{S}} p(s)p(y|x = t(s), s), \tag{2}$$

where $p(s)$ is the probability of the state s. The DMC defined in (2) is called the *associated* channel. The codes for the associated channel describe the codes for the SD-DMC that use only the current state symbols in the encoding operation. In order to describe all coding schemes for the SD-DMC that use only the current state symbol in the encoding process, $\mathcal{T}$ must include all functions from the state alphabet to the input alphabet of the state-dependent channel. There are a total of $|\mathcal{X}|^{|\mathcal{S}|}$ of such functions, where $|.|$ denotes the cardinality of a set. Any of the functions can be represented by a $|\mathcal{S}|$-tuple $(x_1, x_2, \ldots, x_{|\mathcal{S}|})$ composed of elements of $\mathcal{X}$, implying that the value of the function at state s is $x_s, s = 1, 2, \ldots, |\mathcal{S}|$.

III. The Channel Model

We consider data transmission over the channel

$$Y = X + S + N, \tag{3}$$

where X is the channel input, which takes on values in a fixed real constellation $\mathcal{X}$, Y is the channel output, N is additive white Gaussian noise with power σ^2, and the interference S is a discrete random variable that takes on values in a real finite set $\mathcal{S}$. The sequence of i.i.d. interference symbols is known causally at the encoder.

The above channel can be considered as a special case of the state-dependent channel considered by Shannon with one exception, that the channel output alphabet is continuous. In our case, the likelihood function $f_{Y|X,S}(y|x,s)$ is used instead of the transition probabilities. We denote the input to the associated channel by T, which can be considered as a function from $\mathcal{S}$ to $\mathcal{X}$. We denote the cardinality of $\mathcal{X}$ and $\mathcal{S}$ by M and Q, respectively. Then the cardinality of $\mathcal{T}$ will be M^Q, which is the number all possible functions from $\mathcal{S}$ to $\mathcal{X}$.

The likelihood function for the associated channel is given by

$$\begin{aligned} f_{Y|T}(y|t) &= \sum_{s\in\mathcal{S}} p(s) f_{Y|X,S}(y|t(s), s) \\ &= \sum_{s\in\mathcal{S}} p(s) f_N(y - t(s) - s), \end{aligned} \tag{4}$$

where $p(s)$ is the probability of the interference symbol s and f_N denotes the pdf of the Gaussian noise N.

IV. The Code Design Criterion

Any coding scheme for the associated channel defined by (4) translates to a coding scheme for the actual channel defined by $f_{Y|X,S}(y|x,s)$. We use the pairwise error probability (PEP) approach to derive the code design criterion at high SNR. Suppose that the messages w_1 and w_2 are encoded to $t_1^n \equiv t_1 t_2 \ldots t_n$ and $r_1^n \equiv r_1 r_2 \ldots r_n$, respectively, where t_i's and r_i's belong to the alphabet $\mathcal{T}$. Using maximum likelihood decoding, the probability of the event that message w_2 is decoded given message w_1 was sent is given by

$$\begin{aligned} \Pr\{\mathrm{e}|w_1\} &= \sum_{s_1^n} p(s_1^n) \Pr\{\mathrm{e}|w_1, s_1^n\} \\ &= \sum_{s_1^n} p(s_1^n) \Pr\left\{f(y_1^n|t_1^n) < f(y_1^n|r_1^n)|w_1, s_1^n\right\}, \\ &= \sum_{s_1^n} p(s_1^n) \Pr\left\{\prod_{i=1}^n f(y_i|t_i) < \prod_{i=1}^n f(y_i|r_i)|w_1, s_1^n\right\} \\ &= \sum_{s_1^n} p(s_1^n) \Pr\left\{\prod_{i=1}^n \sum_{s\in\mathcal{S}} p(s) f_N(y_i - t_i(s) - s) < \right. \\ &\qquad \left. \prod_{i=1}^n \sum_{s\in\mathcal{S}} p(s) f_N(y_i - r_i(s) - s)|w_1, s_1^n\right\}. \end{aligned} \tag{5}$$

where $s_1^n \equiv s_1 \cdots s_n$ denotes the interference sequence during the transmission. It can be shown that the value of PEP at high

SNR is given by

$$\text{PEP} \propto Q\left(\frac{\sqrt{\sum_{i=1}^{n} d_{SI}^2(t_i, r_i)}}{2\sigma}\right), \tag{6}$$

where $d_{SI}(t, r)$, the distance between two input symbols of the associated channel t and r, is defined as

$$d_{SI}(t, r) = \min_{s_1, s_2 \in \mathcal{S}} |t(s_1) + s_1 - r(s_2) - s_2|. \tag{7}$$

It is worth mentioning that the distance measure defined in (7) does not satisfy the triangle inequality. For example, consider a channel with $\mathcal{X} = \mathcal{S} = \{-1, +1\}$. Then the four symbols of the associated channel can be represented as $t = (-1, -1), u = (-1, +1), v = (+1, -1), w = (-1, -1)$. It is easy to check that the distances between all pairs of the symbols are zero except for $d_{SI}(u, v)$ which is 2. Therefore, the triangle inequality does not hold for $d_{SI}(t, u)$, $d_{SI}(t, v)$, and $d_{SI}(u, v)$.

V. The Binary Channel

Any code designed for the regular associated channel translates to a code for the actual channel with known interference at the encoder. The alphabet size of the associated channel is M^Q. However, we might not need to use all the symbols of the alphabet in the encoding scheme as long as the distance spectrum of the code is concerned.

For example, we consider the case where $M = 2$, i. e., when the channel accepts binary input. Then the size of $\mathcal{T}$ will be 2^Q. Let t and t' be two symbols in $\mathcal{T}$ with the maximum distance among all pairs of symbols in $\mathcal{T}$. It can be shown that there exist two symbols in $\mathcal{T}$ with nonzero distance [9]. Therefore, $d_{SI}(t, t') > 0$. Since $d_{SI}(t, t') > 0$, we have $t(s) \neq t'(s), \forall s \in \mathcal{S}$, otherwise, from (7), $d_{SI}(t, t') = 0$. We choose an arbitrary interference symbol $s_0 \in \mathcal{S}$ to partition $\mathcal{T}$ as follows. We put $r \in \mathcal{T}$ in $\mathcal{T}_1$ if $r(s_0) = t(s_0)$, otherwise (i.e., $r(s_0) = t'(s_0)$) put r in $\mathcal{T}_2$. Note that the distance between any two symbols in $\mathcal{T}_j$ is zero, $j = 1, 2$.

Suppose that a codebook is designed for the binary channel with codewords composed of elements of $\mathcal{T}$. We construct a new codebook from the current one by replacing the elements of the codewords that belong to $\mathcal{T}_1$ by t and replacing the elements of the codewords that belong to $\mathcal{T}_2$ by t'. Since the codewords of the new codebook are composed of just two elements, we may call the new code a binary code.

Theorem 1: The distance spectrum of the binary code constructed by the procedure described above is at least as good as the distance spectrum of the old code.

Proof: Consider any two codewords $(t_1, \ldots, t_n)$ and $(r_1, \ldots, r_n)$ from the old codebook, where $t_i, r_i \in \mathcal{T}$. The squared distance between the two codewords is equal to $\sum_{i=1}^{n} d_{SI}^2(t_i, r_i)$. For any $i \in \{1, 2, \ldots, n\}$, we consider two cases:

Case 1: t_i and r_i belong to the same partition. Then $d_{SI}(t_i, r_i) = 0$, so the replacement will not change the distance.

Case 2: t_i and r_i belong to different partitions. Then since $d_{SI}(t_i, r_i) \leq d_{SI}(t, t')$, the replacement will not decrease the distance. ■

According to theorem 1, as long as the distance spectrum of the code in concerned, it is sufficient to use just two symbols of $\mathcal{T}$ with maximum distance, namely t and t', in the encoding for a binary channel. Naturally, we can define the Hamming distance between any two codewords, which is the number of positions at which two codewords are different. Consider two codewords $c_1 = (t_1, \ldots, t_n)$ and $c_2 = (r_1, \ldots, r_n)$ with elements from the binary set $\{t, t'\}$. The squared distance between these codewords is given by

$$\sum_{i=1}^{n} d_{SI}^2(t_i, r_i) = d_{SI}^2(t, t') d_H(c_1, c_2), \tag{8}$$

where $d_H(c_1, c_2)$ is the Hamming distance between c_1 and c_2. Therefore, the problem of designing codes for the binary channel where the interference sequence is known causally at the encoder reduces to the design of codes for the interference-free binary-input AWGN channel. The only difference is that the coding is over the set $\{t, t'\}$ rather than $\{0, 1\}$.

If we were to use a binary code for the interference-free binary channel with the input alphabet $\mathcal{X} = \{x, x'\}$, then the Euclidean distance between any two codewords c_1 and c_2 of length n for the interference-free channel would be

$$d_E^2(c_1, c_2) = (x - x')^2 d_H(c_1, c_2), \tag{9}$$

where d_E denotes the Euclidean distance.

Using (8) and (9), we can compare the performance of a zero-one binary code for the binary channel with causal side information at the encoder with the same zero-one binary code for the interference-free binary channel. In the case of channel with side information, zero and one are mapped to t and t', and in the case of the interference-free channel, zero and one are mapped to x and x', respectively. Note that t and t' are functions from the interference alphabet $\mathcal{S}$ to the channel input alphabet $\mathcal{X} = \{x, x'\}$.

It is clear from (7) that

$$d_{SI}(t, t') \leq |x - x'|. \tag{10}$$

Therefore, using (8) and (9), the distance spectrum of the code for the interference-free channel is at least as good as the distance-spectrum of the code for the channel with known interference at the encoder. Of course, this is not surprising. However, it is interesting to search for the conditions that (10) is satisfied with equality.

If (10) is satisfied with equality, the distance spectrum of the two codes will be the same. In particular, the slope of error probability curves at high SNR (which corresponds to the minimum distance of the codebook) with maximum likelihood decoding will be the same for the two cases. In other words, if (10) is satisfied with equality, the knowledge of interference at the encoder enables us to achieve the same performance as the interference-free case at high SNR.

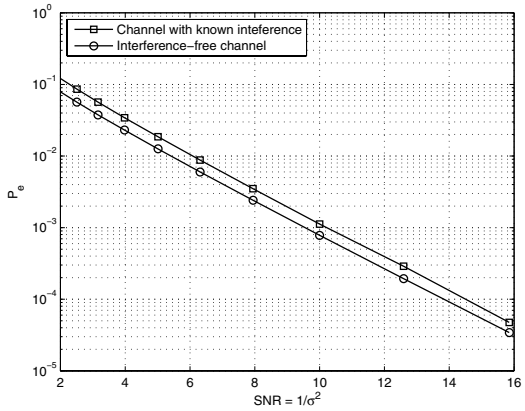

Fig. 3. Error probability vs. SNR for the binary input AWGN channel with/without (known) interference. $\mathcal{X} = \mathcal{S} = \{-1, +1\}$.

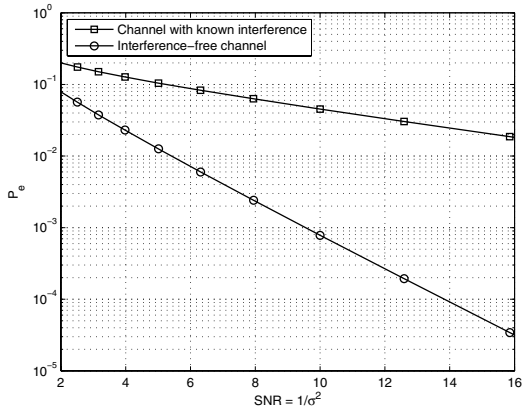

Fig. 4. Error probability vs. SNR for the binary input AWGN channel with/without (known) interference. $\mathcal{X} = \{-1, +1\}, \mathcal{S} = \{-1, 0, +1\}$.

Theorem 2: $d_{SI}(t, t') = |x - x'|$ if and only if

$$\min_{s_1 \neq s_2 \in \mathcal{S}} |s_1 - s_2| \geq |x - x'|. \tag{11}$$

Proof: If $\min |s_1 - s_2| \geq |x - x'|$, then we may take $t = (x, x', x, \ldots)$ and $t' = (x', x, x', \ldots)$. Then it is easy to check that $d_{SI}(t, t') = |x - x'|$.

For the other direction, suppose that $\min |s_1 - s_2| < |x - x'|$. We will show that $d_{SI}(t, t') < |x - x'|$. Suppose that $s, s' \in \mathcal{S}$ achieve the minimum of $|s_1 - s_2|$ and t_1 and t_2 arbitrary elements of $\mathcal{T}$. Then, we consider two possibilities:

Case 1: $t_1(s) = t_1(s') = x$ and $t_2(s) = t_2(s') = x'$. Then $|t_1(s') + s' - t_2(s) - s| < |x - x'|$.

Case 2: $t_1(s) = x, t_1(s') = x'$ and $t_2(s) = x', t_2(s') = x$. Then $|t_1(s) + s - t_2(s') - s'| < |x - x'|$. ∎

As an example, consider a binary channel with $\mathcal{X} = \mathcal{S} = \{-1, +1\}$ and with equiprobable interference symbols. The two symbols with the maximum distance in the input alphabet of the associated channel are $t = (-1, +1), t' = (+1, -1)$. We have simulated the error probability performance of the above channel without error control coding and with maximum likelihood decoding. The error probability vs. SNR for the above channel is plotted in fig. 3. The error probability curve for the interference-free binary channel with $\mathcal{X} = \{-1, +1\}$ is plotted for comparison. For the interference-free channel, $P_e = Q(\frac{1}{\sigma})$. It is easy to check that for this example, $d_{SI}(t, t') = |x - x'| = 2$. As it can be seen, the curves have the same slopes as expected at high SNR. Note that the SNR coordinate in fig. 3 is not in dB. Note that if the interference were not known at the encoder, the error probability curve would reach an error floor of $\frac{1}{4}$.

Another example is illustrated in fig. 4. For this example, $\mathcal{X} = \{-1, +1\}, \mathcal{S} = \{-1, 0, +1\}$. We can find by inspection two symbols of the associated channel input alphabet with the maximum distance as $t = (-1, -1, +1), t' = (+1, +1, -1)$. Here, we have $d_{SI}(t, t') = 1 < |x - x'| = 2$. Therefore, the error probability curve for the channel with known interference at the encoder does not decay as fast as the error probability curve for the interference-free channel.

VI. Conclusion

In this paper, we derived the code design criterion at high SNR for the M-ary input AWGN channel with additive Q-level interference, where the sequence of interference symbols is known causally at the encoder. The code design is over an input alphabet $\mathcal{T}$ of size M^Q. We defined a new distance measure between the symbols of $\mathcal{T}$. The performance of the codes at high SNR is governed by the minimum distance between the codewords with elements from $\mathcal{T}$. We may not need to use all symbols of $\mathcal{T}$ in the encoding. In particular, we showed that for the case $M = 2$, as long as the distance spectrum of the code is concerned, we just need to use two symbols of $\mathcal{T}$ with the maximum distance among all pairs of symbols. This reduces the code design problem for our channel to code design for regular binary-input AWGN channels which has been well researched in the past fifty years.

References

[1] B. Chen and G. W. Wornell, "Quantization index modulation: A class of provably good methods for digital watermarking and information embedding," *IEEE Trans. Inform. Theory*, vol. 47, no. 4, pp. 1423-1443, May 2001.

[2] G. Caire and S. Shamai,"On achievable throughput of a multiple antenna Gaussian broadcast channel," *IEEE Trans. Inform. Theory*, vol. 49, no. 7, pp. 1691-1706, Jul. 2003.

[3] M. H. M. Costa, "Writing on dirty paper," *IEEE Trans. Inform. Theory*, vol. IT-29, no. 3, pp. 439-441, May 1983.

[4] U. Erez, S. Shamai, and R. Zamir, "Capacity and lattice strategies for canceling known interference," *IEEE Trans. Inform. Theory*, vol. 51, no. 11, pp. 3820-3833, Nov. 2005.

[5] U. Erez, and S. ten Brink, "A close-to-capacity dirty paper coding scheme," *IEEE Trans. Inform. Theory*, vol. 51, no. 10, pp. 3417-3432, Oct. 2005.

[6] A. Bennatan, D. Burshtein, G. Caire, and S. Shamai, "Superposition coding for side-information channels," *IEEE Trans. Inform. Theory*, vol. 52, no. 5, pp. 1872-1889, May 2006.

[7] C. E. Shannon, "Channels with side information at the transmitter," *IBM Journal of Research and Development*, vol. 2, pp. 289-293, Oct. 1958.

[8] S. Gel'fand and M. Pinsker, "Coding for channel with random parameters," *Problems of Control and Information Theory*, vol. 9, no. 1, pp. 19-31, Jan. 1980.

[9] H. Farmanbar and A. K. Khandani, "Precoding for the AWGN channel with discrete interference," *Submitted to IEEE Trans. Inform. Theory*, March 2007, available at http://arxiv.org/.