Transactions Letters

A Short-Coding Error Parameter for Channels with Block Interference

Shaul Laufer and Jakov Snyders

Abstract—For a class of channels, called block interference channels, a paradoxical situation prevails, namely, the capacity increases but the cutoff rate decreases with an increase of the block interference length. Also, despite the large capacity, there is a degradation in the performance of practical coding schemes when the block interference length is excessive. We introduce a short-coding error parameter (SCEP), whereby a bound on the average probability of decoding error is expressed for codes with length shorter than the block interference length. This bound is tighter than the bound based on the (conventional) cutoff rate. The SCEP is independent of the block interference length.

I. INTRODUCTION

THE cutoff rate of a channel is widely accepted (see Massey [1], [2], Viterbi [3]) to be the upper limit of code rates for which practical and reliable coding schemes exist, whereas the capacity is the theoretical upper limit. However, for a class of channels, termed block interference channels (BIC), McEliece and Stark [4] demonstrated a paradoxical situation: with an increase of the block interference length the capacity increases but the cutoff rate decreases. Also, there is a degradation in the performance of practical coding schemes when the block interference length is excessive. The block interference length of the BIC represents a certain type of memory-length. Hereafter we shall use the shorter term memory-length.

In this letter we introduce a (so called) short-coding error parameter (SCEP), which provides a bound on the average probability of decoding error for codes with length shorter than the memory-length. This bound is tighter than the bound based on the conventional cutoff rate. Also, the SCEP is independent of the channel’s memory-length, provided the codeword length is shorter than the memory-length. Suitably modified versions of this parameter seem to be useful for evaluating the performance of communication systems transmitting over channels with long memory-length, such as coded slow frequency-hopping systems subject to partial band jamming. Another potential area of application is multiple-access communications with long time slots.

The definition and properties of the SCEP are presented in Section II. In Section III the SCEP is exhibited for a simple example. We conclude with a discussion in Section IV.

II. A SHORT-CODING ERROR PARAMETER

In this section we consider the BIC as modeled by McEliece and Stark [4]. It is composed of a collection \( \{ \Delta_s \} \) of memoryless channels with noise severity level represented by the index \( s \). All component channels have the same discrete input alphabet \( A \) and discrete output alphabet \( B \). The index parameter \( s \) lies in a set \( \Omega \) upon which a probability distribution \( P(s) \) has been defined. Let \( S_1, S_2, S_3, \ldots \) be a sequence of independent identically distributed \( \Omega \)-valued random variables, with distribution \( P(s) \). For any \( m \geq 1 \) the channels \( \Delta^m \) and \( \Delta^{m'} \) are defined as follows (see Fig. 1). The channel segments the input sequence into blocks of \( m \) consecutive letters. The \( k \)-th block is transmitted over a component channel \( \Delta_s \) determined by \( S_k \). The channel is called BIC without side information (SI), and is denoted by \( \Delta^m \), if there is no \( k \) for which the value of \( S_k \) is conveyed to the receiver. If the channel provides to the receiver the value of \( S_k \) for all \( k \), it is called BIC with SI and is denoted \( \Delta^{m'} \).

McEliece and Stark [4] viewed the above BIC’s as discrete memoryless channels (DMC) with extended alphabets:
(A^m, B^m) for \Delta^m, (A^m, B^m \times \Omega) for \Delta^m. They calculated under a compatibility assumption the capacity and the cutoff rate and found the following: \( \overline{C} \) is independent of \( m \),

\[
C(m) \leq \overline{C}, \quad \lim_{m \to \infty} C(m) = \overline{C},
\]

where

\[
\overline{R}_o(m) \leq E(R_{o,s}),
\]

\[
\lim_{m \to \infty} \overline{R}_o(m) = \min_s R_{o,s}, \text{ for finite number of component channels,}
\]

and

\[
\overline{R}_o(m) \geq R_c(m)
\]

where \( R_{o,s} \) is the cutoff rate of the component channel \( \Delta_s, \overline{C} \) and \( \overline{R}_o(m) \) correspond to \( \Delta^m \), \( C(m) \) and \( R_c(m) \) correspond to \( \Delta^m \). The compatibility assumption, which will be adopted here also, is that the capacity (cutoff rate) of all component channels is obtained by the same input probability distribution.

The paradox revealed by (1)-(4), namely, that the capacity tends to increase with the memory-length \( m \) while the cutoff rate tends to decrease with \( m \), led McEliece and Stark [4] to the conclusion that \( C \) is the appropriate measure for assessing the quality of a BIC while \( R_c \) is an inverse measure of the coding delay rather than a measure of the coding complexity.

However, one may argue that whereas the capacity of a BIC is an appropriate measure for the channel's ability to transmit information reliably with the aid of very long codes, the cutoff rate is unreliable since in its evaluation the memory-length is interpreted to be the logarithm of the alphabet size. That is, the alphabet size increases exponentially with memory-length. The cutoff rate is a parameter which is a result of applying the union bound to the probability of error. The union bound is known to be loose for large alphabet sizes. This may explain the peculiar behavior and the unreliability of the cutoff rate.

The conventional performance measures \( C \) and \( R_c \) depend on the coding channel solely. Therefore alphabet extension is valid for viewing a BIC as a DMC without paying attention to specific code parameters, e.g., the size of the alphabet from which the code letters are selected and the codeword length. However, in order to obtain a bound on the probability of error (and consequently a performance measure) for short codes, specific code parameters should be taken into consideration. Then, however, the channel cannot be viewed as a DMC.

We proceed to introduce a SCEP \( R_c \) where \( S \), which stands for "short," indicates that the codeword length \( N_c \), measured in the original input channel alphabet \( A \) letters, is shorter than the BCC memory-length \( m \). For simplicity of analysis we shall impose the following further restrictions:

a) \( N_c \mid m \), i.e., \( N_c \) divides the block interference length \( m \),

b) the codewords are aligned with the change of channel state. These conditions are met in some practical applications, such as slow frequency hopping communications employed in an environment of fading and partial band jamming. In such applications \( m \) is usually large, allowing selection of a code from a large variety of codes for which \( N_c \mid m \) is fulfilled.

Furthermore, for a large \( m \) relaxation of the restrictions is expected to have a negligible effect.

The ensemble average probability \( \rho_e \) of decoding error for a memoryless channel, using maximum-likelihood decoding, satisfies [5, pp. 131-139]

\[
\rho_e \leq 2^{-N_c[R_{o,s} - R]},
\]

where \( R \) is the code rate. Here \( N_c \) is the code length and \( R_{o,s} \), measured in bits per channel use, is given by \( R_{o,s} = \sup_{F_X} E_0(1, F_X) \) where \( E_0(\xi, F_X) \) is the Gallager function and \( F_X \) is the input distribution. Therefore the ensemble average probability of decoding error \( P_e(s) \) at a component channel is bounded by

\[
P_e(s) \leq \left\{ \begin{array}{ll} 2^{-N_c[R_{o,s} - R]}, & \text{if } R \leq R_{o,s} \\ 1, & \text{otherwise} \end{array} \right.
\]

where \( R_{o,s} \) is the cutoff rate of the DMC component \( \Delta_s \). This bound is known to be tight for \( m \geq N_c \gg 1 \). Taking the expectation of \( P_e(s) \) over \( \Omega \) for obtaining the average probability of error \( P_e \), we have

\[
P_e = E\{P_e(s)\} \leq E\left\{ 2^{-N_c[R_{o,s} - R]} \right\}
\]

\[
= 2^{-N_c[R_{o,s} - R]} \quad \text{for } N_c \mid m
\]

where the equality is due to the following definition of \( R_c(N_c) \)

\[
R_c(N_c) = -\frac{1}{N_c} \log_2 E\left\{ 2^{-N_c R_{o,s}} \right\} \quad \text{for } N_c \mid m.
\]

We emphasize that \( R_c(N_c) \), despite (8), can not be regarded as a cutoff rate since it varies with \( N_c \). Notice also that in order for (8) and (9) to hold we require \( N_c \mid m \), which implies \( N_c \leq m \). Nevertheless, \( R_c(N_c) \) proves to be a useful parameter for bounding \( P_e \), as demonstrated in the sequel.

In (7) we used \( P_e(s) \leq 2^{-N_c[R_{o,s} - R]} \) for all \( R \), which is a less tight bound then (6). In fact, we allow \( R \) to exceed \( \min_{s} R_{o,s} \), provided \( R \leq R_c(N_c) \). In this case the code rate \( R \) and the cutoff rate \( R_{o,s} \) of a specific component channel \( \Delta_s \), satisfy \( R_{o,s} \leq R \leq R_c(N_c) \), then the contribution to the upper bound on \( P_e \) is at least \( P_e(s) \). Still, according to (8), the upper bound on \( P_e \) does not exceed 1. We remark that if we have used (6) for bounding \( P_e \) then we would have obtained a tighter bound then (7), but that bound would be expressed with a parameter that depends on \( R \) also.

McEliece and Stark [4] derived the following expression for \( \overline{R}_o(m) \)

\[
\overline{R}_o(m) = \frac{1}{m} \log_2 E\left\{ 2^{-m R_{o,s}} \right\}.
\]

Rearrangement of (10) and (9) yields

\[
\overline{R}_o(m) = -\log_2 \left[ E\left\{ \left( 2^{-R_{o,s}} \right)^m \right\} \right]^{1/m},
\]

\[
R_c(N_c) = -\log_2 \left[ E\left\{ \left( 2^{-R_{o,s}} \right)^{N_c} \right\} \right]^{1/N_c}.
\]

By applying a standard inequality [5, p. 523]

\[
\left[ E(a^r) \right]^{1/r} \leq \left[ E(a^t) \right]^{1/t} \text{ if } 0 < r < t, a_t > 0,
\]

...
to these expressions we deduce the following properties of $\bar{R}_o(m)$ and $R_c(N_c)$:

$$\bar{R}_o(m') \leq \bar{R}_o(m), \quad \text{if } 0 < m < m'. \quad (12)$$

and

$$R_c(N_c') \leq R_c(N_c), \quad \text{if } 0 < N_c < N_c'. \quad (13)$$

Furthermore, as $0 < N_c \leq m$,

$$\bar{R}_o(m) \leq R_c(N_c) \quad (14)$$

with equality for $N_c = m$.

By (7) it is clear that the bound on $P_e$ based on $R_c(N_c)$ is monotonically (though not exponentially) decreasing with $N_c$ (for $N_c \leq m$), although $R_c(N_c)$ is monotonically non-increasing (13) with $N_c$. In order to provide a bound on the ensemble average probability of decoding error, both for $N_c < N$, and $m \leq N_c$, a unified performance measure $R_u(m, N_c)$ may be defined as follows:

$$R_u(m, N_c) = -\frac{1}{\min(m, N_c)} \log_2 \mathbb{E}\left\{2^{-\min(m, N_c)R_o} \right\} \quad (15)$$

The bound based on $R_u(m, N_c)$ is tighter than the bound based on $\bar{R}_o(m)$ solely, when $N_c < m$.

III. AN EXAMPLE

Consider the following two-state BIC:

$$P(s) = \begin{cases} 1 - \rho, & \text{if } s = 0, \\ \rho, & \text{if } s = 1 \end{cases} \quad (16)$$

and

$$R_o,s = \begin{cases} 1, & \text{if } s = 0, \\ 0, & \text{if } s = 1 \end{cases} \quad (17)$$

where $0 \leq \rho \leq 1$. Then

$$R_u(m, N_c) = -\frac{1}{\min(m, N_c)} \log_2 \left\{(1-\rho)2^{-\min(m, N_c)} + \rho \right\}. \quad (18)$$

The behavior of the unified measure as a function of the codeword length $N_c$ is demonstrated in Fig. 2 for $m = 72$ and three values of $\rho$. Also, $\bar{R}_o(m = 72)$ is depicted as reference. We have evaluated $R_u(m, N_c)$ also for $m = 720$. Remarkably, the results are qualitatively similar for both values of $m$.

In this example we considered a rather simple situation, for which the strategy of error detection might prove to be more practical than application of error correction. However, for a general BIC error correction appears to be the preferable approach.

IV. SUMMARY

The paradox revealed by the behavior of the conventional performance measures for channels with block memory, as well as the capacity’s well-known deficiency of failing to reflect the performance when customary short block codes are used, led us to search for a new performance measure. We introduced a SCEP which, based on its properties, appears to be a preferable performance measure for the BIC since it provides a tighter bound on the average probability of error than the bound based on the conventional cutoff rate. Also, this parameter is independent of the channel’s memory-length when the codeword length is shorter than the memory-length.

We have not found a computational complexity meaning of $R_o$. Such meaning is usually attributed to $\bar{R}_o$, and is justified within the context of sequential decoding. However, a computational complexity meaning related to $R_u$ has not generally been established [6]. Furthermore, $R_u$ seems to reflect the complexity of a decoder which operates on the extended alphabet, not on the original alphabet. Accordingly, the complexity increases (the cutoff rate decreases) with increasing memory-length.

ACKNOWLEDGMENT

The authors wish to thank S. Shamai (Shitz) and A. Reichman for helpful and stimulating discussions and the referees for their constructive comments.

REFERENCES