

Sequential Random Coding Error Exponents for Multiple Access Channels

Cheng Chang
Electrical Engineering
and Computer Science
University of California, Berkeley
Berkeley, CA 94720
Email: cchang@eecs.berkeley.edu

Anant Sahai
Electrical Engineering
and Computer Science
University of California, Berkeley
Berkeley, CA 94720
Email: sahai@eecs.berkeley.edu

Abstract—Discrete-time memoryless multiple access channels (MACs) are a useful abstraction of the uplink for many centralized wireless systems. They capture the issues involved with many different users wanting to simultaneously send information to a single site. Traditional MAC analysis proceeds in the context of block codes with the messages being known in advance by the encoders. Instead, we look at a sequential setting. Each user’s message evolves in real-time as bits stream in to the encoders. In this context, we look at the probability of error not at the block level, but at the bit-level. Furthermore, in place of block-length, we look at the delay between when the bit arrives at the encoder and when it is decoded by the central decoder.

The sequential random coding error exponent is studied and shown to be positive in the whole achievable rate region for multiple access channels. Furthermore, we show that this exponent can be achieved in a delay-universal or “anytime” fashion in that the encoder does not have to specify the target delay. The choice of decoding delay is left up to the decoder which is free to vary this on an application specific basis — the longer it is willing to wait, the lower the probability of bit error will be.

I. INTRODUCTION

In the point-to-point communication scenario, there are many types of codes. The block-coding paradigm has a semi-infinite sequence of messages, each of which is quite large. Each of these messages is assumed to be known to the encoders at the beginning of the epoch, and the decoder is assumed to produce an estimate for it by the end of an epoch. The next message is considered in the next epoch. The delay in this context is determined by the epoch size or block-length. The sequential-coding paradigm also has a semi-infinite sequence of messages, but each of these is assumed to be quite small. These small messages become available to the encoder as time evolves, and are used to generate channel input symbols causally. There is no *a priori* choice of an epoch or block-length. Instead, the decoder decodes estimates of the messages as time goes on, but does so with some delay.

While the block-coding paradigm has certainly attracted more academic interest, there are many interesting features in the sequential picture. Convolutional and tree codes represent the most well known cases of sequential codes, though these techniques can also be used to construct block-codes. In [1], Forney reviews how the probability of error varies with delay

in the point-to-point settings as the decoder is forced to make a bit-decision at a particular delay. The relevant exponent turns out to be the random block-coding error exponent, and in the high-rate regime, [2] tells us that no code can achieve a higher exponent with delay.¹ In [4], Sahai further identifies the delay-universal or “anytime” variation on sequential codes in which the choice of the delay is left entirely up to the decoder. In [5], anytime communication problems are shown to be intimately connected to problems of automatic control. Recently in [6], we brought anytime ideas to distributed source-coding in the Slepian-Wolf context.

MAC channels [7] are interesting for distributed wireless communication systems, and yet, the study of such channels has focused almost entirely on the block-coding case. The capacity region and random coding error exponents for block coding are explained in [8],[9],[7] and [10] respectively. In [11], the error exponent for tree coding is studied, but the decoding is considered in the block style. The encoding consists of two parts, encoded information bits and a tail to achieve more reliability. The error exponent is defined as $\frac{\log_2(P_e)}{N_t}$, where N_t is the length of the tail. Bounded delay decoding had not been considered.

We study the sequential communication problem for multiple access channel and derive the random coding error exponents for the problem. In this paper, we study the random coding error exponents for multiple access channel in the sense of delay-universal, or “anytime,” reliability. The probability of error is required to go to zero exponentially with delay, where the delay is chosen entirely at the decoder.

The outline of the paper is as follows. In Section II we will first describe the model of the multiple access channel and sequential channel coding. Then in Section III we will derive the random coding bound on sequential coding for multiple access channel.

¹In [2], Pinsker also tries to claim that the same bound holds with feedback, but we have recently found that he is wrong! This is easiest to see using an erasure channel example, and the full story of this will be told in [3]. It turns out that block-length and delay interact very differently with feedback when it comes to the probability of error!

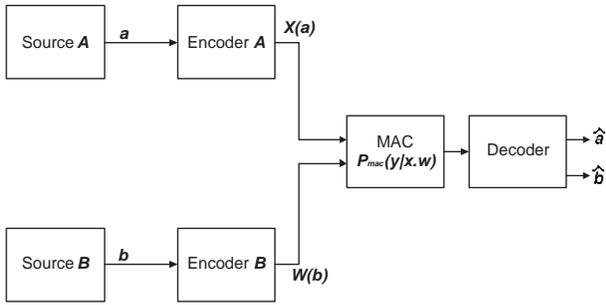


Fig. 1. Model of multiple access channels

II. PROBLEM SETUP

A. Multiple Access Channel

As shown in Fig 1, there are two independent encoders for two independent information sources \mathcal{A} and \mathcal{B} . The multiple access channel is characterized by a transition probability matrix $P_{mac}(y|x,w)$ where $y \in \mathcal{Y}$ is the channel output, $x \in \mathcal{X}$ is the input from the encoder for information source \mathcal{A} , and $w \in \mathcal{W}$ is the input from the encoder for information source \mathcal{B} .

Theorem 1: Achievable capacity region [8],[7]:The achievable rate region \mathcal{R} is the convex hull of the set of rate pairs (R_A, R_B) , for some distributions Q_A, Q_B , satisfy each of the inequalities

$$\begin{aligned} R_A + R_B &\leq I(X, W; Y) \\ 0 \leq R_A &\leq I(X; Y|W) \\ 0 \leq R_B &\leq I(W; Y|X) \end{aligned}$$

where $P(y) = \sum_{x,w} Q_A(x)Q_B(w)P_{mac}(y|x,w)$, $P(y|x) = \sum_x Q_A(x)P_{mac}(y|x,w)$ and $P(y|w) = \sum_w Q_B(w)P_{mac}(y|x,w)$.

The decoding error probability can be arbitrarily close to 0 in the capacity region.

B. Random Coding Error Exponents

The random block-coding error exponent of the multiple access channel is studied in [10]. Consider an ensemble of (n, m, l) where encoder 1 has m equal-probable messages, encoder 2 has l messages, all with n channel uses. The inputs to the channel $\{\vec{x}_1, \dots, \vec{x}_m\}$ are chosen according to $Q_A(x_1^n) = \prod_{i=1}^n Q_A(x_i)$ and $\{\vec{w}_1, \dots, \vec{w}_l\}$ are independently chosen according to $Q_B(w_1^n) = \prod_{i=1}^n Q_B(w_i)$. Write $Q(x, w) = Q_A(x)Q_B(w)$.

Theorem 2: [10] The expected probability of making a decoding error for either block over the ensemble satisfies $\bar{P}_e \leq P_{e_A} + P_{e_B} + P_{e_{AB}}$. Where

$$P_{e_i} \leq 2^{-n[-\rho_i R_i + E_{0_i}(\rho_i, Q)]} \quad (1)$$

$$\forall \rho_i \in [0, 1], i = A, B, AB, R_A = \frac{\log_2(m)}{n}, R_B = \frac{\log_2(l)}{n}$$

$$R_{AB} = R_A + R_B.$$

$$\begin{aligned} E_{0_A}(\rho_A, Q) &= -\log_2 \sum_{y,w} Q_B(w) \left[\sum_x Q_A(x) P_{mac}(y|x,w)^{\frac{1}{1+\rho_A}} \right]^{1+\rho_A} \\ E_{0_B}(\rho_B, Q) &= -\log_2 \sum_{y,x} Q_A(x) \left[\sum_w Q_B(w) P_{mac}(y|x,w)^{\frac{1}{1+\rho_B}} \right]^{1+\rho_B} \\ E_{0_{AB}}(\rho_{AB}, Q) &= -\log_2 \sum_y \left[\sum_{x,w} Q_A(x) Q_B(w) P_{mac}(y|x,w)^{\frac{1}{1+\rho_{AB}}} \right]^{1+\rho_{AB}} \end{aligned}$$

We've silently changed from nats used in [10] to bits.

C. Sequential Channel Coding

The model of sequential channel coding is as follows. For rate 1, the information bits $a_i \in \{0, 1\}$ come to the encoder at time i , $i = 1, 2, \dots, n, \dots$. The encoder must send out an encoded symbol right away, that is only allowed to depend on $a_j, j \leq i$, in a causal fashion. The decoder can choose a decision time and hence delay for a certain information bit, and the expected probability of decoding error must be bounded by an exponential function of the delay.

In order to achieve any rational rate $R = \frac{A}{C}$, we encode A information bits into C channel uses.

Definition 1: (A, C, Q_A, \mathcal{X}) random sequential coding scheme: For a distribution Q_A defined on the alphabet \mathcal{X} , and integers $A, C \in \mathcal{N}$. The encoder \mathcal{E} is a sequence of maps $\{\mathcal{E}_j\}, j = 1, 2, \dots$. The outputs of \mathcal{E}_j are the outputs of the encoder from time $(j-1)C + 1$ to jC . Where

$$\begin{aligned} \mathcal{E}_j &: \{0, 1\}^{jA} \times [0, 1]^C \longrightarrow \mathcal{X}^C \\ \mathcal{E}_j(a_1^{jA}, \lambda(j)_1^C) &= x_{(j-1)C+1}^{jC} \end{aligned}$$

Where $\lambda(j)_i, i = 1, 2, \dots, C; j = 1, 2, \dots$ are iid random variables uniformly distributed in $[0, 1]$. And

$$x_{(j-1)C+i} = \mathcal{X}_k, i = 1, 2, \dots, C$$

If $\sum_{l=1}^{k-1} Q_A(\mathcal{X}_l) < \lambda(j)_i < \sum_{l=1}^k Q_A(\mathcal{X}_l)$

The series of random series $\lambda^{(j)}_1, j = 1, 2, \dots$ are known at the encoder and the decoder.

By the construction of the codebook, it's obvious that the encoder is *casual*, i.e. for any two information sequences a_1^{nA} and \tilde{a}_1^{mA} , if for some $k \leq \min(n, m)$ s.t. $a_1^{kA} = \tilde{a}_1^{kA}$, then the first kC outputs of the encoder for both information sequence are the same.

The rate of the encoder in bits is $R_A = \frac{A}{C}$.

Forney showed in [1] that the random coding error exponent defined in [12] can be achieved in the sequential setup by forcing the decoder to give its best decision with delay $(n-i)C$ channel uses. The probability of decoding error of the i 'th information source block after nC channel uses is

$$P_{e_n}(i) \leq K 2^{-(n-i)C E_r(R)} \quad (2)$$

Where K is a constant and $E_r(R)$ is the random coding bound defined in [12].

D. Encoding and Decoding

The encoders and decoder work as following. Encoder \mathcal{E}_A uses an (A, C, Q_A, \mathcal{X}) random sequential scheme, meanwhile encoder \mathcal{E}_B uses an (B, C, Q_B, \mathcal{W}) random sequential scheme, where $R_A = \frac{A}{C}, R_B = \frac{B}{C}, A, B, C \in \mathcal{N}$. After sending n blocks of information bits, where encoder \mathcal{E}_A encodes nA information bits into nC inputs to the multiple access channel and encoder \mathcal{E}_B encodes nB information bits into nC inputs to the multiple access channel. The decoder receives a sequence y_1, \dots, y_{nC} . And the decoded information sequence pair $(\hat{a}_1^{nA}, \hat{b}_1^{nB})$ are the sequence pair with the maximum posterior probability. Formally,

$$(\hat{a}_1^{nA}, \hat{b}_1^{nB}) = \arg \max_{(s_1^{nA}, t_1^{nB})} P_{mac}(y_1^{nC} | \mathcal{E}_A(s_1^{nA}), \mathcal{E}_B(t_1^{nB}))$$

Where $s_1^{nA} \in \{0, 1\}^{nA}, t_1^{nB} \in \{0, 1\}^{nB}$ are two binary information sequences, and $\mathcal{E}_A(s_1^{nA}) \in \mathcal{X}^{nC}, \mathcal{E}_B(t_1^{nB}) \in \mathcal{W}^{nC}$ are the inputs to the channel. Notice that each time, the decoder re-estimates all the information bits. We will show that the probability of decoding error of the information bits decays exponentially with delay.

III. RANDOM CODING BOUND

In this section we derive a bound on the error probability for the sequential multiple access channel using the randomized encoders in Def 1. We are interested in the error probability of an information bit given some decoding delay. We state the main result of this paper in Theorem 3 and the proof goes as follows. After nC channel uses, depending on the first wrong decoded block of source \mathcal{A} and \mathcal{B} , we have $(n+1)^2$ disjoint error events. We will bound the probability of each error event using the random coding argument and bound the probability of decoding error of a particular block by a summation of some of the error probabilities.

A. Error Probability of Decoding Information Source \mathcal{A}

After nC channel uses, we have the ML decoded information bits \hat{a}_1^{nA} and \hat{b}_1^{nB} .

Theorem 3: After nC channel uses, write the error probability of decoding the j 'th block of source \mathcal{A} as $P_n(\{\hat{a}_{jA+1}^{(j+1)A} \neq a_{jA+1}^{(j+1)A}\})$.

$$P_n(\{\hat{a}_{jA+1}^{(j+1)A} \neq a_{jA+1}^{(j+1)A}\}) \leq \left(\frac{d}{1-2^{-dCE^*}} + \frac{2}{(1-2^{-dCE^*})^2} \right) 2^{-dCE^*} \quad (3)$$

Where $dC = (n+1-j)C$ is the decoding delay and

$$E^* = \min \left\{ \inf_{\alpha_1 \in [0,1]} \left\{ \sup_{\rho_1 \in [0,1]} \{ \alpha_1(-\rho_1 R_A + E_{0A}(Q, \rho_1)) + (1-\alpha_1)(-\rho_1 R_{AB} + E_{0AB}(Q, \rho_1)) \} \right\}, \inf_{\alpha_2 \in [0,1]} \left\{ \sup_{\rho_2 \in [0,1]} \{ \alpha_2(-\rho_2 R_B + E_{0B}(Q, \rho_2)) + (1-\alpha_2)(-\rho_2 R_{AB} + E_{0AB}(Q, \rho_2)) \} \right\} \right\} \quad (4)$$

Where $Q = Q_A Q_B$ and E^* is positive in the inner region of the capacity region.

By letting $\alpha_i = 0, 1$, we can easily see that $E^* \leq \min_{i=A,B,AB} \{ \sup_{\rho \in [0,1]} \{-\rho R_i + E_{0i}(Q, \rho)\} \}$, where $E_{0i}(Q, \rho)$ is defined in Theorem 2.

We need some preparation before proving the theorem.

Proposition 1: Partition of $\{0, 1\}^{nA} \times \{0, 1\}^{nB}$

Given the information sequence pair (a_1^{nA}, b_1^{nB}) we can partition the set $\{0, 1\}^{nA} \times \{0, 1\}^{nB}$ into $(n+1)^2$ subsets.

For $1 \leq j, k \leq n$

$$F_n(j, k, (a_1^{nA}, b_1^{nB})) = \{ (s_1^{nA}, t_1^{nB}) \in \{0, 1\}^{nA} \times \{0, 1\}^{nB} | s_1^{(j-1)A} = a_1^{(j-1)A}, s_{(j-1)A+1}^{jA} \neq a_{(j-1)A+1}^{jA}, t_1^{(k-1)B} = b_1^{(k-1)B}, t_{(k-1)B+1}^{kB} \neq b_{(k-1)B+1}^{kB} \}$$

For $1 \leq j \leq n$

$$F_n(j, n+1, (a_1^{nA}, b_1^{nB})) = \{ (s_1^{nA}, t_1^{nB}) \in \{0, 1\}^{nA} \times \{0, 1\}^{nB} | s_1^{(j-1)A} = a_1^{(j-1)A}, s_{(j-1)A+1}^{jA} \neq a_{(j-1)A+1}^{jA}, t_1^{nB} = b_1^{nB} \}$$

For $1 \leq k \leq n$

$$F_n(n+1, k, (a_1^{nA}, b_1^{nB})) = \{ (s_1^{nA}, t_1^{nB}) \in \{0, 1\}^{nA} \times \{0, 1\}^{nB} | s_1^{nA} = a_1^{nA}, t_1^{(k-1)B} = b_1^{(k-1)B}, t_{(k-1)B+1}^{kB} \neq b_{(k-1)B+1}^{kB} \}$$

And $F_n(n+1, n+1, (a_1^{nA}, b_1^{nB})) = \{(a_1^{nA}, b_1^{nB})\}$.

We use the convention that if $i_1 < i_2$, $a_{i_2}^{i_1}$ is an empty sequence. F is a partition of $\{0, 1\}^{nA} \times \{0, 1\}^{nB}$ because if $(j_1, k_1) \neq (j_2, k_2)$,

$$F_n(j_1, k_1, (a_1^{nA}, b_1^{nB})) \cap F_n(j_2, k_2, (a_1^{nA}, b_1^{nB})) = \emptyset \text{ and } \bigcup_{1 \leq j, k \leq n+1} F_n(j, k, (a_1^{nA}, b_1^{nB})) = \{0, 1\}^{nA} \times \{0, 1\}^{nB}$$

Example 1: $F_n(n+1, n+1, (a_1^{nA}, b_1^{nB}))$: in this example, $A = 1, B = 1, n = 2$, $(a_1^{nA}, b_1^{nB}) = (00, 00)$ we simply write $F_n(j, k, (a_1^{nA}, b_1^{nB}))$ as $F_{j,k}$.

$$F_{1,1} = \{(10, 10), (10, 11), (11, 10), (11, 11)\}.$$

$$F_{1,2} = \{(10, 01), (11, 01)\}.$$

$$F_{2,1} = \{(01, 10), (01, 11)\}.$$

$$F_{2,2} = \{(01, 01)\}.$$

$$F_{1,3} = \{(10, 00), (11, 00)\}.$$

$$F_{3,1} = \{(00, 10), (00, 11)\}.$$

$$F_{2,3} = \{(01, 00)\}. F_{3,2} = \{(00, 01)\}.$$

$$F_{3,3} = \{(00, 00)\}.$$

Definition 2: Error Event $E_n(j, k, (a_1^{nA}, b_1^{nB}))$:

$$E_n(j, k, (a_1^{nA}, b_1^{nB})) = \{(\hat{a}_1^{nA}, \hat{b}_1^{nB}) \in F_n(j, k, (a_1^{nA}, b_1^{nB}))\}$$

We call $E_n(j, k, (a_1^{nA}, b_1^{nB}))$ the (j, k) 'th error event.

More specifically, given information sequence pair (a_1^{nA}, b_1^{nB}) , the decoded information sequence pair is $(\hat{a}_1^{nA}, \hat{b}_1^{nB})$. We define $E_n(j, k, (a_1^{nA}, b_1^{nB}))$, $1 \leq j, k \leq n+1$ as following.

First, $E_n(j, k, (a_1^{nA}, b_1^{nB}))$, $1 \leq j, k \leq n$, is the following error event, after nC channel uses the first decoding error for information source \mathcal{A} is block j , and the first decoding error for information source \mathcal{B} is k .

Secondly if the decoder does not make any decoding errors for \mathcal{A} , but the first decoding error for \mathcal{B} is at the k 'th block, $1 \leq k \leq n$ then the error event is $E_n(n+1, k, (a_1^{nA}, b_1^{nB}))$, similarly for $E_n(j, n+1, (a_1^{nA}, b_1^{nB}))$.

Finally the event of making no decoding errors after nC channel uses is precisely $E_n(n+1, n+1, (a_1^{nA}, b_1^{nB}))$.

Now we use the random coding bound argument to give an upper bound on the probability of $E_n(j, k, (a_1^{nA}, b_1^{nB}))$. Without loss of generality we assume $j \leq k$. The argument we use is very similar to the derivation of the random coding error exponent in [12] and [10]. The error probability is an expectation taken over all the channel realizations and all the randomness in the code.

Lemma 1: Random coding bound on the (j, k) 'th error event. Encoder \mathcal{E}_A uses an $\mathcal{E}(A, C, Q_A, \mathcal{X})$ random sequential coding scheme, meanwhile \mathcal{E}_B independently uses an $\mathcal{E}(B, C, Q_B, \mathcal{W})$ random sequential coding scheme. \forall information sequence pair (a_1^{nA}, b_1^{nB}) , $1 \leq j \leq k \leq n+1$.

$$P(E_n(j, k, (a_1^{nA}, b_1^{nB}))) \leq 2^{-(k-j)C(-\rho R_A + E_{0A}(Q, \rho)) - (n+1-k)C(-\rho R_{AB} + E_{0AB}(Q, \rho))} \quad (5)$$

Where $Q = Q_A Q_B$, $E_{0A}(Q, \rho)$ and $E_{0AB}(Q, \rho)$ are defined in Theorem 2.

Proof: The proof here is similar to the derivation of the random coding bound on the block coding error probability in [12].

The probability of $E_n(j, k, (a_1^{nA}, b_1^{nB}))$ is upper bounded by the probability of the existence of a sequence pair $(s_1^{nA}, t_1^{nB}) \in F_n(j, k, a_1^{nA}, b_1^{nB})$, s.t.

$$P_{mac}(y_1^{nC} | \mathcal{E}_A(s_1^{nA}), \mathcal{E}_B(t_1^{nB})) \geq P_{mac}(y_1^{nC} | \mathcal{E}_A(a_1^{nA}), \mathcal{E}_B(b_1^{nB}))$$

Write $\mathcal{E}_A(a_1^{nA})$ as x_1^{nC} , $\mathcal{E}_B(b_1^{nB})$ as w_1^{nC} , $\mathcal{E}_A(s_1^{nA})$ as \bar{x}_1^{nC} and $\mathcal{E}_B(t_1^{nB})$ as \bar{w}_1^{nC} . Since $(s_1^{nA}, t_1^{nB}) \in F_n(j, k, a_1^{nA}, b_1^{nB})$, we have $\bar{x}_1^{(j-1)C} = x_1^{(j-1)C}$, $\bar{w}_1^{(k-1)C} = w_1^{(k-1)C}$. So

$$\begin{aligned} & P(E_n(j, k, (a_1^{nA}, b_1^{nB}))) \\ & \leq P(\exists(s_1^{nA}, t_1^{nB}) \in F_n(j, k, a_1^{nA}, b_1^{nB}), \\ & \text{s.t. } P_{mac}(y_1^{nC} | \bar{x}_1^{nC}, \bar{w}_1^{nC}) \geq P_{mac}(y_1^{nC} | x_1^{nC}, w_1^{nC})) \\ & = \sum_{x_1^{nC}} \sum_{w_1^{nC}} \sum_{y_1^{nC}} Q_A(x_1^{nC}) Q_B(w_1^{nC}) P_{mac}(y_1^{nC} | x_1^{nC}, w_1^{nC}) \\ & P(\exists(s_1^{nA}, t_1^{nB}) \in F_n(j, k, a_1^{nA}, b_1^{nB}), \\ & \text{s.t. } P_{mac}(y_1^{nC} | \bar{x}_1^{nC}, \bar{w}_1^{nC}) \geq P_{mac}(y_1^{nC} | x_1^{nC}, w_1^{nC}) \\ & | x_1^{nC}, w_1^{nC}, y_1^{nC}) \end{aligned} \quad (6)$$

First we bound the conditional probability

$$\begin{aligned} & P(\exists(s_1^{nA}, t_1^{nB}) \in F_n(j, k, a_1^{nA}, b_1^{nB}), \\ & \text{s.t. } P_{mac}(y_1^{nC} | \bar{x}_1^{nC}, \bar{w}_1^{nC}) \geq P_{mac}(y_1^{nC} | x_1^{nC}, w_1^{nC}) \\ & | x_1^{nC}, w_1^{nC}, y_1^{nC}) \\ & \leq \left[\sum_{(s_1^{nA}, t_1^{nB}) \in F_n(j, k, a_1^{nA}, b_1^{nB})} \right. \\ & \left. P(P_{mac}(y_1^{nC} | \bar{x}_1^{nC}, \bar{w}_1^{nC}) \geq P_{mac}(y_1^{nC} | x_1^{nC}, w_1^{nC}) \right. \\ & \left. | x_1^{nC}, w_1^{nC}, y_1^{nC}) \right]^\rho, \forall \rho \in [0, 1] \end{aligned} \quad (7)$$

$\forall (s_1^{nA}, t_1^{nB}) \in F_n(j, k, a_1^{nA}, b_1^{nB})$, by noticing that $\bar{x}_1^{(j-1)C} = x_1^{(j-1)C}$, $\bar{w}_1^{(k-1)C} = w_1^{(k-1)C}$ and the memorylessness of the channel. We have

$$\begin{aligned} & P_{mac}(y_1^{nC} | \bar{x}_1^{nC}, \bar{w}_1^{nC}) = \\ & P_{mac}(y_1^{(j-1)C} | x_1^{(j-1)C}, w_1^{(j-1)C}) \\ & P_{mac}(y_{(j-1)C+1}^{(k-1)C} | \bar{x}_{(j-1)C+1}^{(k-1)C}, w_{(j-1)C+1}^{(k-1)C}) \\ & P_{mac}(y_{(k-1)C+1}^{nC} | \bar{x}_{(k-1)C+1}^{nC}, \bar{w}_{(k-1)C+1}^{nC}) \end{aligned} \quad (8)$$

Now:

$$\begin{aligned} & P(P_{mac}(y_1^{nC} | \bar{x}_1^{nC}, \bar{w}_1^{nC}) \geq P_{mac}(y_1^{nC} | x_1^{nC}, w_1^{nC}) \\ & | x_1^{nC}, w_1^{nC}, y_1^{nC}) \\ & = \sum_{(\bar{x}_1^{nC}, \bar{w}_1^{nC}) : P_{mac}(y_1^{nC} | \bar{x}_1^{nC}, \bar{w}_1^{nC}) \geq P_{mac}(y_1^{nC} | x_1^{nC}, w_1^{nC})} \\ & Q_A(\bar{x}_{(j-1)C+1}^{nC}) Q_B(\bar{w}_{(k-1)C+1}^{nC}) \\ & \leq \sum_{(\bar{x}_{(j-1)C+1}^{nC}, \bar{w}_{(k-1)C+1}^{nC})} Q_A(\bar{x}_{(j-1)C+1}^{nC}) Q_B(\bar{w}_{(k-1)C+1}^{nC}) \\ & \frac{P_{mac}(y_1^{nC} | \bar{x}_1^{nC}, \bar{w}_1^{nC})^s}{P_{mac}(y_1^{nC} | x_1^{nC}, w_1^{nC})^s} \\ & = \sum_{(\bar{x}_{(j-1)C+1}^{nC}, \bar{w}_{(k-1)C+1}^{nC})} Q_A(\bar{x}_{(j-1)C+1}^{nC}) Q_B(\bar{w}_{(k-1)C+1}^{nC}) \\ & \frac{P_{mac}(y_{(j-1)C+1}^{(k-1)C} | \bar{x}_{(j-1)C+1}^{(k-1)C}, w_{(j-1)C+1}^{(k-1)C})^s}{P_{mac}(y_{(j-1)C+1}^{(k-1)C} | x_{(j-1)C+1}^{(k-1)C}, w_{(j-1)C+1}^{(k-1)C})^s} \\ & \frac{P_{mac}(y_{(k-1)C+1}^{nC} | \bar{x}_{(k-1)C+1}^{nC}, \bar{w}_{(k-1)C+1}^{nC})^s}{P_{mac}(y_{(k-1)C+1}^{nC} | x_{(k-1)C+1}^{nC}, w_{(k-1)C+1}^{nC})^s} \end{aligned} \quad (9)$$

$\forall s > 0$.

The size of $F_n(j, k, a_1^{nA}, b_1^{nB})$ can be bounded as

$$\begin{aligned} & |F_n(j, k, a_1^{nA}, b_1^{nB})| \leq 2^{(n+1-k)(A+B)} 2^{(k-j)A} \\ & = 2^{(n+1-k)C R_{AB}} 2^{(k-j)C R_A} = M \end{aligned} \quad (10)$$

(6) Substitute Eqn. 9 into Eqn. 7, and using the union bound

argument, we get:

$$\begin{aligned}
& P(\exists(s_1^{nA}, t_1^{nB}) \in F_n(j, k, a_1^{nA}, b_1^{nB}), \\
& \text{s.t. } P_{\text{mac}}(y_1^{nC} | \bar{x}_1^{nC}, \bar{w}_1^{nC}) \geq P_{\text{mac}}(y_1^{nC} | x_1^{nC}, w_1^{nC}) \\
& | x_1^{nC}, w_1^{nC}, y_1^{nC}) \\
& \leq [M \sum_{(\bar{x}_{(j-1)C+1}^{nC}, \bar{w}_{(k-1)C+1}^{nC})} Q_A(\bar{x}_{(j-1)C+1}^{nC}) Q_B(\bar{w}_{(k-1)C+1}^{nC}) \\
& \frac{P_{\text{mac}}(y_{(j-1)C+1}^{(k-1)C} | \bar{x}_{(j-1)C+1}^{(k-1)C}, w_{(j-1)C+1}^{(k-1)C})^s}{P_{\text{mac}}(y_{(j-1)C+1}^{(k-1)C} | x_{(j-1)C+1}^{(k-1)C}, w_{(j-1)C+1}^{(k-1)C})^s} \\
& \frac{P_{\text{mac}}(y_{(k-1)C+1}^{nC} | \bar{x}_{(k-1)C+1}^{nC}, \bar{w}_{(k-1)C+1}^{nC})^s}{P_{\text{mac}}(y_{(k-1)C+1}^{nC} | x_{(k-1)C+1}^{nC}, w_{(k-1)C+1}^{nC})^s}]^\rho \quad (11)
\end{aligned}$$

Substitute Eqn. 11 into Eqn. 6. And by noticing the memorylessness of the channel and the fact that $Q_A(x_1^{nC}) = Q_A(x_1^{(j-1)C}) Q_A(x_{(j-1)C+1}^{nC})$, etc. We have:

$$\begin{aligned}
& P(E_n(j, k, (a_1^{nA}, b_1^{nB}))) \\
& \leq \sum_{x_{(j-1)C+1}^{nC}} \sum_{w_{(j-1)C+1}^{nC}} \sum_{y_{(j-1)C+1}^{nC}} \\
& Q_A(x_{(j-1)C+1}^{nC}) Q_B(w_{(j-1)C+1}^{nC}) \\
& P_{\text{mac}}(y_{(j-1)C+1}^{nC} | x_{(j-1)C+1}^{nC}, w_{(j-1)C+1}^{nC}) \\
& [M \sum_{(\bar{x}_{(j-1)C+1}^{nC}, \bar{w}_{(k-1)C+1}^{nC})} Q_A(\bar{x}_{(j-1)C+1}^{nC}) Q_B(\bar{w}_{(k-1)C+1}^{nC}) \\
& \frac{P_{\text{mac}}(y_{(j-1)C+1}^{(k-1)C} | \bar{x}_{(j-1)C+1}^{(k-1)C}, w_{(j-1)C+1}^{(k-1)C})^s}{P_{\text{mac}}(y_{(j-1)C+1}^{(k-1)C} | x_{(j-1)C+1}^{(k-1)C}, w_{(j-1)C+1}^{(k-1)C})^s} \\
& \frac{P_{\text{mac}}(y_{(k-1)C+1}^{nC} | \bar{x}_{(k-1)C+1}^{nC}, \bar{w}_{(k-1)C+1}^{nC})^s}{P_{\text{mac}}(y_{(k-1)C+1}^{nC} | x_{(k-1)C+1}^{nC}, w_{(k-1)C+1}^{nC})^s}]^\rho \quad (12)
\end{aligned}$$

By letting $s = \frac{1}{1+\rho}$, and noticing the fact that \bar{x} and \bar{w} are dummy variables.

$$\begin{aligned}
& P(E_n(j, k, (a_1^{nA}, b_1^{nB}))) \\
& \leq M^\rho \sum_{y_{(k-1)C+1}^{nC}} \\
& \left\{ \sum_{x_{(k-1)C+1}^{nC}} \sum_{w_{(k-1)C+1}^{nC}} Q_A(x_{(k-1)C+1}^{nC}) Q_B(w_{(k-1)C+1}^{nC}) \right. \\
& \left. P_{\text{mac}}(y_{(k-1)C+1}^{nC} | x_{(k-1)C+1}^{nC}, w_{(k-1)C+1}^{nC})^{\frac{1}{1+\rho}} \right\}^{1+\rho} \\
& \sum_{y_{(j-1)C+1}^{(k-1)C}} \sum_{w_{(j-1)C+1}^{(k-1)C}} Q_B(w_{(j-1)C+1}^{(k-1)C}) \\
& \left\{ \sum_{x_{(j-1)C+1}^{(k-1)C}} Q_A(x_{(j-1)C+1}^{(k-1)C}) \right. \\
& \left. P_{\text{mac}}(y_{(j-1)C+1}^{(k-1)C} | x_{(j-1)C+1}^{(k-1)C}, w_{(j-1)C+1}^{(k-1)C})^{\frac{1}{1+\rho}} \right\}^{1+\rho} \\
& = 2^{-(k-j)C(-\rho R_A + E_{0A}(Q, \rho)) - (n+1-k)C(-\rho R_{AB} + E_{0AB}(Q, \rho))} \quad (13)
\end{aligned}$$

The last equality is true because $Q_A(x_{i_1}^{i_2}) = \prod_{i=i_1}^{i=i_2} Q_A(x_i)$, etc.

Similarly if $1 \leq k \leq j \leq n+1$,

$$\begin{aligned}
& P(E_n(j, k, (a_1^{nA}, b_1^{nB}))) \leq \\
& 2^{-(j-k)C(-\rho R_B + E_{0B}(Q, \rho)) - (n+1-j)C(-\rho R_{AB} + E_{0AB}(Q, \rho))} \quad \square
\end{aligned}$$

Corollary 1: An upper bound on $P(E_n(j, k, (a_1^{nA}, b_1^{nB})))$

$$P(E_n(j, k, (a_1^{nA}, b_1^{nB}))) \leq 2^{-(n+1-\min(j,k))CE^*} \quad (14)$$

Where E^* is the error exponent defined in Theorem 3.

Proof: Without loss of generality, we assume $j \leq k$. Then from Lemma 1, we know that $\forall \rho$

$$\begin{aligned}
& P(E_n(j, k, (a_1^{nA}, b_1^{nB}))) \leq \\
& 2^{-(k-j)C(-\rho R_B + E_{0B}(Q, \rho)) - (n+1-k)C(-\rho R_{AB} + E_{0AB}(Q, \rho))} \\
& = 2^{-(n+1-j)C(\alpha(-\rho R_B + E_{0B}(Q, \rho)) + (1-\alpha)(-\rho R_{AB} + E_{0AB}(Q, \rho)))}
\end{aligned}$$

Where $\alpha = \frac{k-j}{n+1-j} \in [0, 1]$. So

$$\begin{aligned}
& \alpha(-\rho R_B + E_{0B}(Q, \rho)) + (1-\alpha)(-\rho R_{AB} + E_{0AB}(Q, \rho)) \\
& \geq \inf_{\alpha_1 \in [0, 1]} (\sup_{\rho \in [0, 1]} (\alpha_1(-\rho R_A + E_{0A}(Q, \rho)) \\
& + (1-\alpha_1)(-\rho R_{AB} + E_{0AB}(Q, \rho)))) \\
& \geq E^*
\end{aligned}$$

Similarly for $j > k$. Thus we proved that

$$P(E_n(j, k, (a_1^{nA}, b_1^{nB}))) \leq 2^{-(n+1-\min(j,k))CE^*} \quad \square \quad (15)$$

Now we are ready for the proof of Theorem 3.

Proof:

The probability of making a decoding error at time nC (after nC channel uses) on the j 'th information block $a_{jA+1}^{(j+1)A}$ is upper bounded by making a decoding error on *any* block with block number not larger than j .

$$\begin{aligned}
& P_n(\{\hat{a}_{jA+1}^{(j+1)A} \neq a_{jA+1}^{(j+1)A}\}) \leq P(\bigcup_{i=1}^j \{\hat{a}_{iA+1}^{(i+1)A} \neq a_{iA+1}^{(i+1)A}\}) \\
& \leq \sum_{i=1}^j \sum_{h=1}^{n+1} P(E_n(i, h, (a_1^{nA}, b_1^{nB}))) \\
& \leq \sum_{i=1}^j \sum_{h=1}^{n+1} 2^{-(n+1-\min(i,h))CE^*} \\
& = \sum_{i=1}^j \sum_{h=j+1}^{n+1} 2^{-(n+1-i)CE^*} + \\
& \sum_{i=1}^j (2j+1-2i) 2^{-(n+1-i)CE^*} \\
& \leq \frac{n+1-j}{1-2^{-CE^*}} 2^{-(n+1-j)CE^*} + \\
& 2 \times 2^{-(n+1-j)CE^*} \sum_{i=1}^j (j+1-i) 2^{-(j-i)CE} \\
& \leq \left(\frac{n+1-j}{1-2^{-CE^*}} + \frac{2}{(1-2^{-CE^*})^2} \right) 2^{-(n+1-j)CE^*} \\
& = \left(\frac{d}{1-2^{-CE^*}} + \frac{2}{(1-2^{-CE^*})^2} \right) 2^{-dCE^*}
\end{aligned}$$

Where $dC = (n + 1 - j)C$ is the decoding delay.

Now we sketch the proof of why $E^* > 0$ everywhere in the interior of the achievable region of Theorem 1. First, $\forall i = A, B, AB$

$$-\rho R_i + E_{0i}(\rho, Q)|_{\rho=0} = 0$$

Secondly for (R_A, R_B) in the interior of \mathcal{R} defined in Theorem 1.

$$\frac{\partial}{\partial \rho} (-\rho R_i + E_{0i}(\rho, Q))|_{\rho=0} > 0 \quad (16)$$

Thus $\exists \rho_i \in [0, 1]$, s.t. $E_{0i}(\rho, Q) > 0, \forall \rho \in [0, \rho_i]$, so $\forall \rho \in [0, \min_{i=A,B,AB}(\rho_i)]$

$$E^* \geq \min_{i=A,B,AB} (-\rho R_i + E_{0i}(\rho, Q)) > 0 \quad (17)$$

□

IV. AN EXAMPLE

Consider an adder channel followed by a symmetric channel as shown in Fig. 2. We plot the error exponents E^* for $\epsilon = 0.1$ in Fig. 3. The sequential error exponents are positive in the whole capacity region. The value of α which minimizes the error exponents in Eqn. 4 is not always 0 or 1. If $\alpha = 0, 1$, the sequential error exponent is the same as one of block coding error exponents in [10]. However, on the boundary region of \mathcal{R}_1 and \mathcal{R}_3 , \mathcal{R}_2 and \mathcal{R}_3 in Fig.8 in [10], the value of α which minimizes Eqn. 4 is not 0 or 1, thus we have a different error exponent bound for sequential random coding as compared to block random coding.

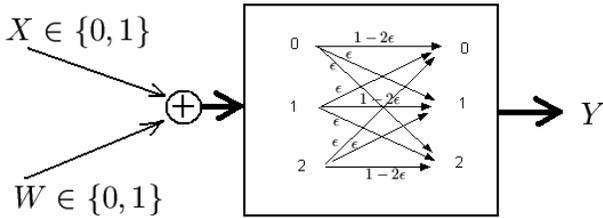


Fig. 2. A Multiple Access Channel

V. CONCLUSIONS AND FUTURE WORK

We studied sequential channel coding for multiple access channels. By applying a variation of Gallager's random coding scheme, we achieved positive sequential random coding error exponent for the whole multiple access channel rate region. This exponent measured the probability of bit-error with bit-delay rather than block-error with block-length. Furthermore, the code was "anytime" or delay universal in that the decoder can decide on a delay without telling the encoder what it is. The decision is exponentially more reliable for a longer delay. The analyzed random coding scheme is consisted of a sequential encoder and an ML decoder. ML decoding is computationally heavy, a practically approximate decoding scheme is desirable. The main difficulty is the growing complexity

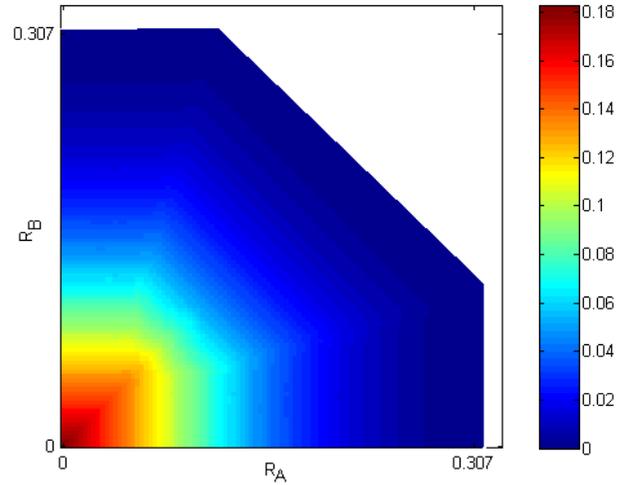


Fig. 3. Sequential Error Exponents of the MAC in Fig. 2 with $\epsilon = 0.1$

of the encoding as time goes on. To address this, we accept a certain maximum tolerable delay (or equivalently, a certain small enough error probability) beyond which we are no longer interested in correcting errors with additional waiting. This can then be realized using a (possibly time-varying) convolutional code with a long enough constraint length. Also, we believe that sequential decoding [13] might extend to such contexts and furthermore, that similar sequential random coding bounds can be achieved for other multi-terminal problems such as the broadcast channel.

REFERENCES

- [1] G. Forney, "Convolutional codes ii. maximum-likelihood decoding," *Inform. and Control*, vol. 25, pp. 222–266, 1974.
- [2] M. Pinsker, "Bounds of the probability and of the number of correctable errors for nonblock codes," *Translation from Problemy Peredachi Informatsii*, vol. 3, pp. 44–55, 1967.
- [3] A. Sahai, T. Simsek, and P. Varaiya, "Why block length is not the same as delay," *In preparation*.
- [4] A. Sahai, *Anytime Information Theory*, PhD thesis. Massachusetts Institute Technology, 2001.
- [5] A. Sahai and S. Mitter, "The necessity and sufficiency of anytime capacity for stabilization of a linear system over a noisy communication link: Part i: scalar systems," *IEEE Transactions on Information Theory*, March 2005, submitted for publication.
- [6] S. Draper, C. Chang, and A. Sahai, "Sequential random binning for streaming distributed source coding," *Submitted to ISIT*, 2005.
- [7] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: John Wiley and Sons Inc., 1991.
- [8] H. Liao, "A coding theorem for multiple access communications," *ISIT*, 1972.
- [9] R. Ahlswede, "Multi-way communication channels," *ISIT*, 1971.
- [10] R. Gallager, "A perspective on multiaccess channels," *IEEE Transactions on Information Theory*, vol. 31, 1985.
- [11] R. Peterson and D. Costello, "Error probability and free distance bounds for two-user tree codes on multiple-access channels," *IEEE Transactions on Information Theory*, vol. 26, 1980.
- [12] R. Gallager, *Information Theory and Reliable Communication*. John Wiley & Sons.
- [13] G. Forney, "Convolutional codes iii. sequential decoding," *Inform. and Control*, vol. 25, pp. 267–297, 1974.