

Coding of Streaming Sources for the Bidirectional Broadcast Channel

(Invited Paper)

Tobias J. Oechtering, Vishwambhar Rathi
 School of Electrical Engineering and the ACCESS Linnaeus Center
 Royal Institute of Technology (KTH), Stockholm, Sweden

Abstract—The bidirectional broadcast channel denotes a broadcast channel with two receivers where each receiver knows the message intended for the other. We consider streaming sources where messages of each user arrive as bit-stream to the encoder. Thus our main quantity of interest is the bit error probability. By extending Sahai’s argument to bidirectional broadcast channel, we show that under maximum-likelihood decoding the bit error probability decays exponentially in delay with positive exponent for all the rate pairs inside the capacity region. We also show existence of deterministic codes which achieve exponentially decaying bit error probability with delay.

I. INTRODUCTION

In this work we consider sequential coding for the bidirectional broadcast channel. The bidirectional broadcast channel is a two-terminal broadcast channel where the receiving terminals know the message intended for the other terminal. Originally, it describes the second phase of a bidirectional decode-and-forward relaying protocol where two terminals want to communicate with each other using the help of a half-duplex relay node. In the first phase terminals send their messages to the relay node, which decodes the messages. It therefore is described by the classical multiple access channel. In the second phase the relay broadcasts a re-encoded message which allows both terminals to decode the other message using its own source message as side information. Bidirectional relaying using a decode-and-forward strategy based on superposition coding has been introduced in [1]. The XOR-coding on the decoded data for bidirectional relaying has been studied first in [2–5]. The optimal coding strategy combines channel coding for the broadcast channel and the network coding idea. The capacity region of the corresponding bidirectional broadcast channel is independently derived in [6–8].

Theorem 1 ([6]): The capacity region \mathcal{C}_{BC} of the discrete bidirectional memoryless broadcast channel is the convex set of all rate pairs $[R_1, R_2]$ such that

$$0 \leq R_1 \leq I(X; Y), \quad 0 \leq R_2 \leq I(X; Z),$$

for some input probability mass function $P_X(x)$ and given channel transition probability $P_{YZ|X}(y, z|x)$.

The information flow in a bidirectional relaying protocol corresponds to the famous butterfly network[9], which is one

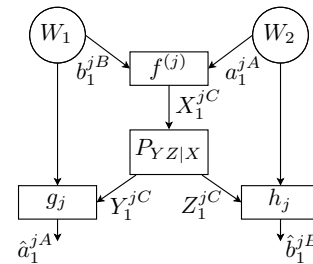


Fig. 1. The system diagram of the bidirectional broadcast channel at time instant j with streaming sources W_1 and W_2 which emitted source sequences b_1^{jB} and a_1^{jA} so far. The messages are mapped onto the channel input sequence X_1^{jC} by the encoder $f^{(j)}$. The sequences Y_1^{jC} and Z_1^{jC} denote the channel output of the memoryless broadcast channel $P_{YZ|X}$. Both decoder g_j and h_j make an estimate of the unknown source sequences \hat{a}_1^{jA} and \hat{b}_1^{jB} using the channel output and its side-information. The illustration shows that the information flow in the bidirectional broadcast channel corresponds to the famous butterfly network.

of the smallest networks to benefit from network coding. An alternative proof of the coding theorem for the bidirectional broadcast channel has been given in [10]. This proof is based on Gallager’s error exponent approach [11, 12].

The bidirectional broadcast channel models the communication scenario where each receiver can observe one source and wants to know the other source. In this work we are interested streaming sources, which means that both sources steadily emit information while the transmitter sends what it has received so far. This is in contrast to the classical block coding structure where the transmitter knows the whole message before the transmission starts. Our treatment of the bidirectional channel is motivated from the following networked control setup. Consider two separated plants which should be stabilized. The random process at a plant is described by the corresponding streaming source. Each controller is lumped together with the decoder and its control policy is assumed to require the knowledge of both the plant states for its control action. Each controller knows the state of its own plant and infers the state of other plant with the help of the relay.

Recently, there is an increasing interest in such problems where communication is considered in the context of a control setup. In [13, 14], Sahai and Mitter introduced the concept of anytime reliability for communication of streaming sources, which can be used to derive necessary and sufficient conditions for stabilizing an unstable plant over a noisy communication

link. The anytime coding strategy is first extended to multi-terminal problems in [15,16], where Chang and Sahai study the sequential channel coding for the multiple-access channel and degraded broadcast channel. In [17] the ideas are extended to a distributed source coding problem.

In this paper we extend the sequential coding results from Sahai and Chang to the bidirectional communication problem. The paper is organized in the following manner. In Section II we formally define the sequential setup for the bidirectional broadcast channel. We extend the sequential coding results for the degraded broadcast from Chang and Sahai [16] to the bidirectional broadcast channel in Section III. In Section IV we extend Sahai's arguments used for the point-to-point channel to show that the rates inside the capacity region with positive error exponents can be achieved using deterministic codes [13]. Finally, we conclude in Section V with some discussion and future work.

II. PROBLEM SETUP

We consider sequential encoding and anytime decoding of streaming sources when transmitted over the bidirectional broadcast channel [13,14].

A. Bidirectional Broadcast Channel

We consider a two terminal *discrete memoryless broadcast channel*. The channel input is defined on the finite input alphabet \mathcal{X} . The output at terminal 1 and 2 is defined on the finite output alphabets \mathcal{Y} and \mathcal{Z} respectively. Then the broadcast channel is defined by a collection of conditional probability mass function $P_{YZ|X}(y, z|x)$ for all $(x, y, z) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$. Further, we assume that the channel input does not depend the previous outputs so that the memoryless property implies that for any time instant n we have

$$P(y^n, z^n | x^n) = \prod_{i=1}^n P_{YZ|X}(y_i, z_i | x_i),$$

with the notation $x_m^n := (x_m, x_{m+1}, \dots, x_n)$ with the common convention $x^n = x_1^n$ if $m = 1$. Since the decoder at terminal 1 and 2 do not cooperate, we define the marginal conditional probability mass functions

$$P_1(y^n | x^n) = \prod_{i=1}^n \sum_{z_i \in \mathcal{Z}} P_{YZ|X}(y_i, z_i | x_i),$$

$$P_2(z^n | x^n) = \prod_{i=1}^n \sum_{y_i \in \mathcal{Y}} P_{YZ|X}(y_i, z_i | x_i).$$

Let W_1 and W_2 denote the random information sources known at the transmitter. In addition, W_1 is known at terminal 1 and W_2 is known at terminal 2. The *bidirectional* communication task for the transmitter is to inform both receiving terminals about their unknown source. In this work we consider streaming sources, which means that the source sequences are not known before the transmission takes place. Furthermore, both source emit source bits at a different rate than the channel is used. Formally, we assume that there exist integers A , B , and C such that in C channel uses

the source W_2 emits a source sequence of length A with letters $a_j \in \{0, 1\}$ and source W_1 emits a source sequence of length B with letters $b_j \in \{0, 1\}$. Accordingly, the sequential encoding strategy produces channel input sequences of length C . For the following definitions we closely follow [15,16].

B. Sequential Encoding

Let the input alphabet be given by $\mathcal{X} := \{x_k\}_{k=1}^{|\mathcal{X}|}$. For integers A , B , and C and a probability mass function P_X defined on the input alphabet \mathcal{X} , a random sequential encoding scheme is defined by a sequences of encoders $\{f_j\}, j = 1, 2, \dots$ where,

$$f_j : \{0, 1\}^{jA} \times \{0, 1\}^{jB} \times [0, 1]^C \rightarrow \mathcal{X}^C,$$

$$f_j(a_1^{jA}, b_1^{jB}, \lambda_1^C(j, a_1^{jA}, b_1^{jB})) = x_{(j-1)C+1}^{jC}.$$

The random variables $\lambda_i(j, a^{jA}, b^{jB}), i = 1, 2, \dots, C, a_1^{jA} \in \{0, 1\}^{jA}, b_1^{jB} \in \{0, 1\}^{jB}$, and $j = 1, 2, \dots$ provide the common randomness and are independent and uniformly distributed on the interval $[0, 1]$. The realizations of the random variables are known at both decoders and encoder and are used to generate a random code. Let $\mathcal{I}_1 := [0, P_X(x_1)]$ and $\mathcal{I}_k := [\sum_{i=1}^{k-1} P_X(x_i), \sum_{i=1}^k P_X(x_i)], k = 2, \dots, |\mathcal{X}|$. such that $\bigcup_{k=1}^{|\mathcal{X}|} \mathcal{I}_k = [0, 1]$. Then i -th element of the encoder output sequence f_j for the messages a_1^{jA} and b_1^{jB} is given by

$$x_{(j-1)C+i} = x_k, \quad \text{if } \lambda_i(j, a^{jA}, b^{jB}) \in \mathcal{I}_k.$$

Further, let $f^{(j)} : \{0, 1\}^{jA} \times \{0, 1\}^{jB} \rightarrow \mathcal{X}^{jC}$ denote the collection of the encoders $f^{(j)} := (f_1, f_2, \dots, f_j), j = 1, 2, \dots$. Note that in order to simplify notation, we drop the argument corresponding to common randomness in the definition of $f^{(j)}$. The encoder has the rate pair $R_1 := A/C$ and $R_2 := B/C$ in bits/channel use.

Since for two message sequence pairs (a_1^{nA}, b_1^{nB}) and $(\tilde{a}_1^{nA}, \tilde{b}_1^{nB})$ with $a_1^{kA} = \tilde{a}_1^{kA}$ and $b_1^{kB} = \tilde{b}_1^{kB}$ for some $k \leq \min\{n, m\}$ the first kC encoder outputs are the same, the definition of the sequential encoder is causal and consistent.

C. Anytime Decoding

For the bidirectional broadcast channel at time jC the decoder of terminal 1 has received the sequence y_1^{jC} and knows the input sequence b_1^{jB} . Similarly, at time jC the decoder of terminal 2 has received the sequence z_1^{jC} and knows the input sequence a_1^{jA} .

The anytime decoding concept means that at any time instant the decoders produce estimates of whole sequences and therefore are allowed to correct past errors. Accordingly, the decoders are defined by sequence $\{g_j\}$ at the receiving terminal 1 with

$$g_j : \{0, 1\}^{jB} \times \mathcal{Y}^{jC} \rightarrow \{0, 1\}^{jA},$$

$$g_j(b_1^{jB}, y_1^{jC}) = \hat{a}_1^{jA}$$

and sequence $\{h_j\}$ at the receiving terminal 2 with

$$h_j : \{0, 1\}^{jA} \times \mathcal{Z}^{jC} \rightarrow \{0, 1\}^{jB},$$

$$h_j(a_1^{jA}, z_1^{jC}) = \hat{b}_1^{jB},$$

$j = 1, 2, \dots$, where \hat{a}_1^{jA} and \hat{b}_1^{jB} denote the decoded information sequences based on the received channel outputs and side-information so far. Without taking complexity issues into account we consider maximum likelihood decoders at both receivers at any time instant jC , $j = 1, 2, \dots$. In more detail, if the transmitter has sent the codeword $x_1^{jC} = f^{(j)}(a_1^{jA}, b_1^{jB})$ for the source sequences a_1^{jA} and b_1^{jB} and terminal 1 has received the sequence y_1^{jC} , then the decoder g_j decides for the source sequence \hat{a}_1^{jA} if

$$P_1(y_1^{jC} | f^{(j)}(\hat{a}_1^{jA}, b_1^{jB})) \geq \max_{\tilde{a}_1^{jA} \in \{0,1\}^{jA}} P_1(y_1^{jC} | f^{(j)}(\tilde{a}_1^{jA}, b_1^{jB})).$$

Similarly, terminal 2 has received sequence z_1^{jC} , then the decoder h_j decides for the source sequence \hat{b}_1^{jB} if

$$P_2(z_1^{jC} | f^{(j)}(a_1^{jA}, \hat{b}_1^{jB})) \geq \max_{\tilde{b}_1^{jB} \in \{0,1\}^{jB}} P_2(z_1^{jC} | f^{(j)}(a_1^{jA}, \tilde{b}_1^{jB})).$$

If the maximum probability is achieved with multiple source sequences, then the decoders randomly decide between them.

III. UPPER BOUND TO ERROR PROBABILITY

For the derivation of the upper bound to the decoding error probability we closely follow the derivation of [15, 16]. In [18] Forney provides an upper bound on the probability of decoding error for the point-to-point channel where the decoder is forced to provide a decision after a certain delay. The result is proved using Gallager's error exponent [11]. In [13] Sahai used this concept in the sequential coding setup.

Theorem 2: Let \hat{a}_1^{nA} and \hat{b}_1^{nB} be the maximum likelihood decoded sequences at terminal 1 and 2 of the source sequences a_1^{nA} and b_1^{nB} after nC channel uses. Then the probability of decoding error of the j -th block can be upper bounded as follows

$$\begin{aligned} \mathbb{P}\{\hat{a}_{jA+1}^{(j+1)A} \neq a_{jA+1}^{(j+1)A}\} &\leq \mathbb{P}\{\hat{a}_1^{(j+1)A} \neq a_1^{(j+1)A}\} \\ &\leq K_1 2^{-dC E_{r,1}(R_1, P_X)} \\ \mathbb{P}\{\hat{b}_{jB+1}^{(j+1)B} \neq b_{jB+1}^{(j+1)B}\} &\leq \mathbb{P}\{\hat{b}_1^{(j+1)B} \neq b_1^{(j+1)B}\} \\ &\leq K_2 2^{-dC E_{r,2}(R_2, P_X)} \end{aligned}$$

with decoding delay $dC = (n - j)C$, constant $K_k = (1 - 2^{-C E_{r,k}(R_k, P_X)})^{-1}$, $k = 1, 2$, and Gallager's random error exponents

$$E_{r,k}(R_k, P_X) := \max_{0 \leq \rho_k \leq 1} (E_{0,k}(\rho_k, P_X) - \rho_k R_k), \quad k = 1, 2$$

$$E_{0,1}(\rho_1, P_X) := -\log_2 \sum_{y \in \mathcal{Y}} \left(\sum_{x \in \mathcal{X}} P_X(x) P_1(y|x)^{\frac{1}{1+\rho_1}} \right)^{1+\rho_1}$$

$$E_{0,2}(\rho_2, P_X) := -\log_2 \sum_{z \in \mathcal{Z}} \left(\sum_{x \in \mathcal{X}} P_X(x) P_2(z|x)^{\frac{1}{1+\rho_2}} \right)^{1+\rho_2}$$

where P_X is the distribution used for the sequential code.

Remark: From [10, 11] we know that for any rate pair in the interior of the capacity region \mathcal{C}_{BC} we can find a distribution P_X such that the error exponents are positive.

Proof: The proof of the theorem, although notationally cumbersome, is simpler than for the degraded broadcast channel case [16] and can be deduced from [13, 16, 18] which

all follow [12]. We go through the proof for decoder 1 for completeness, the analysis for decoder 2 works likewise.

First let us define the set of erroneous suffixes

$$F_{n,j}(a_1^{nA}) := \{s_1^{nA} \in \{0,1\}^{nA} : s_1^{jA} = a_1^{jA}, \\ s_{jA+1}^{nA} \neq a_{jA+1}^{nA}\},$$

for $0 < j < n$ and $F_{n,0}(a_1^{nA}) := \{0,1\}^{nA} \setminus \{a_1^{nA}\}$. $F_{n,n}(a_1^{nA}) := \{a_1^{nA}\}$. By definition we have $F_{n,j}(a_1^{nA}) \subseteq F_{n,k}(a_1^{nA})$ for $0 \leq k \leq j < n$.

Let \hat{a}_1^{nA} be the decoded source sequence at terminal 1. Then the error event

$$E_{n,j}(a_1^{nA}, b_1^{nB}) := \{\hat{a}_1^{nA} \in F_{n,j}(a_1^{nA})\}$$

for $0 \leq j < n$ describes the case when we have the source sequences a_1^{nA} and b_1^{nB} and the decoder at terminal 1 correctly decodes the first j blocks and makes an decoding error in the remaining sequence. By definition we have $E_{n,j}(a_1^{nA}, b_1^{nB}) \subseteq E_{n,k}(a_1^{nA}, b_1^{nB})$ for $0 \leq k \leq j < n$.

Let $x_1^{nC} = f^{(n)}(a_1^{nA}, b_1^{nB})$, then the probability of an error event at terminal 1 using the maximum likelihood decoder is given by

$$\begin{aligned} \mathbb{P}\{E_{n,j}(a_1^{nA}, b_1^{nB})\} &\leq \mathbb{P}\{\exists \hat{a}_1^{nA} \in F_{n,j}(a_1^{nA}) : \\ &P_1(y_1^{nC} | f^{(n)}(\hat{a}_1^{nA}, b_1^{nB})) \geq P_1(y_1^{nC} | x_1^{nC})\} \\ &= \sum_{x_1^{nC}, y_1^{nC}} P_X(x_1^{nC}) P_1(y_1^{nC} | x_1^{nC}) \mathbb{P}\{\exists \hat{a}_1^{nA} \in F_{n,j}(a_1^{nA}) : \\ &P_1(y_1^{nC} | f^{(n)}(\hat{a}_1^{nA}, b_1^{nB})) \geq P_1(y_1^{nC} | x_1^{nC}) | x_1^{nC}, y_1^{nC}\} \\ &\leq \sum_{x_1^{nC}, y_1^{nC}} P_X(x_1^{nC}) P_1(y_1^{nC} | x_1^{nC}) \left[\sum_{\hat{a}_1^{nA} \in F_{n,j}(a_1^{nA})} \right. \\ &\left. \mathbb{P}\{P_1(y_1^{nC} | f^{(n)}(\hat{a}_1^{nA}, b_1^{nB})) \geq P_1(y_1^{nC} | x_1^{nC}) | x_1^{nC}, y_1^{nC}\} \right]^{\rho_1} \end{aligned} \quad (1)$$

for all $\rho_1 \in [0, 1]$. For an \hat{a}_1^{nA} let \hat{x}_1^{nC} denote $f^{(n)}(\hat{a}_1^{nA}, b_1^{nB})$. Since $\hat{a}_1^{nA} \in F_{n,j}(a_1^{nA})$ we have $\hat{a}_1^{jA} = a_1^{jA}$ and $\hat{a}_{jA+1}^{nA} \neq a_{jA+1}^{nA}$ so that we have

$$\hat{x}_1^{jC} = x_1^{jC}, \quad \hat{x}_{jC+1}^{nC} \neq x_{jC+1}^{nC}.$$

From the memoryless property of the channel we have

$$P_1(y_1^{nC} | \hat{x}_1^{nC}) = P_1(y_1^{jC} | x_1^{jC}) P_1(y_{jC+1}^{nC} | \hat{x}_{jC+1}^{nC}).$$

Accordingly, for $\hat{a}_1^{nA} \in F_{n,j}(a_1^{nA})$ we only have to take the last $n - j$ blocks into account

$$\begin{aligned} \mathbb{P}\{P_1(y_1^{nC} | f^{(n)}(\hat{a}_1^{nA}, b_1^{nB})) \geq P_1(y_1^{nC} | x_1^{nC}) | x_1^{nC}, y_1^{nC}\} \\ = \sum_{\hat{x}_{jC+1}^{nC} : P_1(y_{jC+1}^{nC} | \hat{x}_{jC+1}^{nC}) \geq P_1(y_{jC+1}^{nC} | x_{jC+1}^{nC})} P_X(\hat{x}_{jC+1}^{nC}) \\ \leq \sum_{\hat{x}_{jC+1}^{nC} \in \mathcal{X}^{(n-j)C}} P_X(\hat{x}_{jC+1}^{nC}) \frac{P_1(y_{jC+1}^{nC} | \hat{x}_{jC+1}^{nC})^{\rho_1}}{P_1(y_{jC+1}^{nC} | x_{jC+1}^{nC})^{\rho_1}} \end{aligned}$$

for all $s > 0$. For $s = \frac{1}{1+\rho_1}$ and with $|F_{n,j}(a_1^{nA})| = 2^{(n-j)A}$ we can proceed on inequality (1) as follows

$$\begin{aligned}
\mathbb{P}\{E_{n,j}(a_1^{nA}, b_1^{nB})\} &\leq \sum_{x_{jC+1}^{nC}, y_{jC+1}^{nC}} P_X(x_{jC+1}^{nC}) P_1(y_{jC+1}^{nC} | x_{jC+1}^{nC}) \\
&\quad \left[2^{(n-j)A} \sum_{\hat{x}_{jC+1}^{nC}} P_X(\hat{x}_{jC+1}^{nC}) \frac{P_1(y_{jC+1}^{nC} | \hat{x}_{jC+1}^{nC})^{\frac{1}{1+\rho_1}}}{P_1(y_{jC+1}^{nC} | x_{jC+1}^{nC})^{\frac{1}{1+\rho_1}}} \right]^{\rho_1} \\
&= 2^{\rho_1(n-j)A} \sum_{y_{jC+1}^{nC}} \left[\sum_{x_{jC+1}^{nC}} P_X(x_{jC+1}^{nC}) \frac{P_1(y_{jC+1}^{nC} | x_{jC+1}^{nC})^{\frac{\rho_1}{1+\rho_1}}}{P_1(y_{jC+1}^{nC} | x_{jC+1}^{nC})^{\frac{1}{1+\rho_1}}} \right]^{\rho_1} \\
&\quad \left[\sum_{\hat{x}_{jC+1}^{nC}} P_X(\hat{x}_{jC+1}^{nC}) P_1(y_{jC+1}^{nC} | \hat{x}_{jC+1}^{nC})^{\frac{1}{1+\rho_1}} \right]^{\rho_1} \\
&= 2^{\rho_1(n-j)A} \times \\
&\quad \sum_{y_{jC+1}^{nC}} \left[\sum_{x_{jC+1}^{nC}} P_X(x_{jC+1}^{nC}) P_1(y_{jC+1}^{nC} | x_{jC+1}^{nC})^{\frac{1}{1+\rho_1}} \right]^{1+\rho_1} \\
&= 2^{\rho_1(n-j)CR_1} \left(\sum_{y \in \mathcal{Y}} \left[\sum_{x \in \mathcal{X}} P_X(x) P_1(y|x)^{\frac{1}{1+\rho_1}} \right]^{1+\rho_1} \right)^{(n-j)C} \\
&= 2^{\rho_1(n-j)CR_1} 2^{-(n-j)CE_{0,1}(\rho_1, P_X)},
\end{aligned}$$

which can be optimized with respect to ρ_1 such that we have

$$\mathbb{P}\{E_{n,j}(a_1^{nA}, b_1^{nB})\} \leq 2^{-(n-j)CE_{r,1}(R_1, P_X)}.$$

Finally, the probability to make a decoding error in the j -th block after receiving nC symbols is upper bounded by the probability of a decoding error in any of the first $j+1$ blocks which leads to the desired bound

$$\begin{aligned}
\mathbb{P}\{\hat{a}_{jA+1}^{(j+1)A} \neq a_{jA+1}^{(j+1)A}\} &\leq \mathbb{P}\left\{ \bigcup_{i=0}^j \{\hat{a}_{iA+1}^{(i+1)A} \neq a_{iA+1}^{(i+1)A}\} \right\} \\
&\leq \sum_{i=0}^j \mathbb{P}\{E_{n,i}(a_1^{nA}, b_1^{nB})\} \leq \sum_{i=0}^j 2^{-(n-i)CE_{r,1}(R_1, P_X)} \\
&\leq \frac{2^{-(n-j)CE_{r,1}(R_1, P_X)}}{1 - 2^{-CE_{r,1}(R_1, P_X)}} = K_1 2^{-(n-j)CE_{r,1}(R_1, P_X)}
\end{aligned}$$

using $\sum_{i=0}^j r^{n-i} = \frac{r^{n-j} - r^{n+1}}{1-r} \leq \frac{r^{n-j}}{1-r}$ for $r \geq 0$. ■

In the next section we show existence of deterministic codes whose error probability is upper bounded by bounds given in Theorem 2.

IV. EXISTENCE OF DETERMINISTIC CODES

In order to show existence of deterministic codes achieving same error performance as that of the codes with common randomness, we extend Sahai's analysis for point-to-point channel to the bidirectional setting [13]. Note that a code with common randomness is equivalent to an ensemble of codes with some probability measure. Thus the performance of a code with common randomness is equal to the average performance of the corresponding ensemble of codes with respect to an appropriate probability measure.

In the following analysis, we will encounter randomness because of the channel and also because of random code selection from the ensemble of codes. We denote the probability

measure corresponding to the random code selection by \mathbf{P} . As in Section III, we denote average over the ensemble of codes of the probability of an event due to channel randomness by $\mathbb{P}(\cdot)$. For a given encoder \mathcal{F} , the probability of an event due to channel randomness is denoted by $\mathbb{P}_{\mathcal{F}}(\cdot)$. In order to further clarify the notation, consider the error event in decoding the j th block at terminal one i.e. $\{\hat{a}_{jA+1}^{(j+1)A} \neq a_{jA+1}^{(j+1)A}\}$. Then,

$$\mathbb{P}\{\hat{a}_{jA+1}^{(j+1)A} \neq a_{jA+1}^{(j+1)A}\} = E_{\{\mathcal{F}\}} \left(\mathbb{P}_{\mathcal{F}} \left\{ \hat{a}_{jA+1}^{(j+1)A} \neq a_{jA+1}^{(j+1)A} \right\} \right).$$

In the following lemma, we show that average error probability over any set of encoders with positive probability has same behavior as that of the average over the whole ensemble of encoders.

Lemma 3: Consider transmission over bidirectional broadcast channel with sequential encoding and anytime decoding as defined in Section II. Let \mathbf{P} be the probability measure defined on the set of encoders $\{\mathcal{F}\}$ with rate pair (R_1, R_2) . The rates are rational numbers $R_1 = A/C$, $R_2 = B/C$. If the average error probability for decoders 1 and 2 decay exponentially with delay d as

$$\begin{aligned}
\mathbb{P}\{\hat{a}_{jA+1}^{(j+1)A} \neq a_{jA+1}^{(j+1)A}\} &\leq K_1 2^{-dC\alpha_1}, \\
\mathbb{P}\{\hat{b}_{jB+1}^{(j+1)B} \neq b_{jB+1}^{(j+1)B}\} &\leq K_2 2^{-dC\alpha_2},
\end{aligned}$$

then the average error probability over any set with strictly positive probability decays with delay with the same exponent.

Proof: We prove the lemma by contradiction. Let $\{\mathcal{H}\} \subseteq \{\mathcal{F}\}$ and $\mathbf{P}(\{\mathcal{H}\}) > 0$. Assume that there exists $K'_1 > 0$ such that

$$E_{\{\mathcal{H}\}} \left(\mathbb{P}_{\mathcal{F}} \left\{ \hat{a}_{jA+1}^{(j+1)A} \neq a_{jA+1}^{(j+1)A} \right\} \right) \geq K'_1 2^{-dC\alpha_1}. \quad (2)$$

for all $d \geq 0$. Then

$$\begin{aligned}
\mathbb{P}\{\hat{a}_{jA+1}^{(j+1)A} \neq a_{jA+1}^{(j+1)A}\} &= \\
&\mathbf{P}(\{\mathcal{H}\}) E_{\{\mathcal{H}\}} \left(\mathbb{P}_{\mathcal{F}} \left\{ \hat{a}_{jA+1}^{(j+1)A} \neq a_{jA+1}^{(j+1)A} \right\} \right) + \\
&(1 - \mathbf{P}(\{\mathcal{H}\})) E_{\{\bar{\mathcal{H}}\}} \left(\mathbb{P}_{\mathcal{F}} \left\{ \hat{a}_{jA+1}^{(j+1)A} \neq a_{jA+1}^{(j+1)A} \right\} \right).
\end{aligned}$$

We lower bound the average by using the assumption given in (2) and obtain

$$\mathbb{P}\{\hat{a}_{jA+1}^{(j+1)A} \neq a_{jA+1}^{(j+1)A}\} \geq \mathbf{P}(\{\mathcal{H}\}) K'_1 2^{-dC\alpha_1}.$$

We arrive at the contradiction by noting that $\mathbf{P}(\{\mathcal{H}\}) K'_1 > 0$. Note that the same line of reasoning also applies for the second decoder. This completes the proof. ■

The previous lemma shows that there can not be a "large" set of encoders whose performance is worse than the average. However it does not say anything about the performance of a deterministic code. The next lemma deals with this issue.

Lemma 4: Consider the set of encoders $\{\mathcal{F}\}$ with the probability measure \mathbf{P} for the bidirectional broadcast channel for the rate pair $R_1 = A/C$ and $R_2 = B/C$. Suppose that the average probability of error is upper bounded as

$$\mathbb{P}\{\hat{a}_{jA+1}^{(j+1)A} \neq a_{jA+1}^{(j+1)A}\} \leq K_{1,j} 2^{-dC\alpha_1}, \quad (3)$$

$$\mathbb{P}\{\hat{b}_{jB+1}^{(j+1)B} \neq b_{jB+1}^{(j+1)B}\} \leq K_{2,j} 2^{-dC\alpha_2}, \quad (4)$$

for all $d \geq 0$. Then for almost every deterministic encoder \mathcal{F} , every $\epsilon > 0$, there are positive constants $K_{1,j}^{\epsilon,\mathcal{F}}$ and $K_{2,j}^{\epsilon,\mathcal{F}}$ such that

$$\begin{aligned}\mathbb{P}_{\mathcal{F}}\{\hat{a}_{jA+1}^{(j+1)A} \neq a_{jA+1}^{(j+1)A}\} &\leq K_{1,j}^{\epsilon,\mathcal{F}} 2^{-dC(\alpha_1-\epsilon)}, \\ \mathbb{P}_{\mathcal{F}}\{\hat{b}_{jB+1}^{(j+1)B} \neq b_{jB+1}^{(j+1)B}\} &\leq K_{2,j}^{\epsilon,\mathcal{F}} 2^{-dC(\alpha_2-\epsilon)}.\end{aligned}$$

In addition, $K_{1,j}^{\epsilon,\mathcal{F}}$ and $K_{2,j}^{\epsilon,\mathcal{F}}$ can not be too large in the following probabilistic sense,

$$\begin{aligned}\mathbf{P}\left(\left\{K_{1,j}^{\epsilon,\mathcal{F}} > K_1\right\} \cup \left\{K_{2,j}^{\epsilon,\mathcal{F}} > K_2\right\}\right) &\leq \\ &\frac{1}{1-2^{-C\epsilon}} \left(\frac{K_{1,j}^{\epsilon,\mathcal{F}}}{K_1^{\frac{\alpha_1}{\alpha_1-\epsilon}}} + \frac{K_{2,j}^{\epsilon,\mathcal{F}}}{K_2^{\frac{\alpha_2}{\alpha_2-\epsilon}}} \right). \quad (5)\end{aligned}$$

Proof: For all $\epsilon > 0$, $K_1 > 1$, and $K_2 > 1$, let $\mathcal{H}_{(1,\epsilon,K_1,j)}^d$ be the set of encoders for which

$$\mathbb{P}\{\hat{a}_{jA+1}^{(j+1)A} \neq a_{jA+1}^{(j+1)A}\} \geq K_1 2^{-dC(\alpha_1-\epsilon)}.$$

Similarly $\mathcal{H}_{(2,\epsilon,K_2,j)}^d$ be the set of encoders for which the error probability satisfies:

$$\mathbb{P}\{\hat{b}_{jB+1}^{(j+1)B} \neq b_{jB+1}^{(j+1)B}\} \geq K_2 2^{-dC(\alpha_2-\epsilon)}.$$

Note that as probability is always upper bounded by 1, we have $\mathbf{P}\left(\mathcal{H}_{(1,\epsilon,K_1,j)}^d\right) = 0$ for $d < \frac{\log_2(K_1)}{C(\alpha_1-\epsilon)}$ and $\mathbf{P}\left(\mathcal{H}_{(2,\epsilon,K_2,j)}^d\right) = 0$ for $d < \frac{\log_2(K_2)}{C(\alpha_2-\epsilon)}$. We now upper bound the probability of $\mathcal{H}_{(1,\epsilon,K_1,j)}^d$ and $\mathcal{H}_{(2,\epsilon,K_2,j)}^d$ using Markov's inequality. Using the assumptions (3) and (4) in the Markov's inequality, we obtain

$$\mathbf{P}\left(\mathcal{H}_{(1,\epsilon,K_1,j)}^d\right) \leq \frac{K_{1,j}^{\epsilon,\mathcal{F}} 2^{-dC\epsilon}}{K_1}, \quad (6)$$

$$\mathbf{P}\left(\mathcal{H}_{(2,\epsilon,K_2,j)}^d\right) \leq \frac{K_{2,j}^{\epsilon,\mathcal{F}} 2^{-dC\epsilon}}{K_2}. \quad (7)$$

Let $\mathcal{H}_{(\epsilon,K_1,K_2,j)}^d$ be the set of encoders whose probability of error for the first decoder goes above $K_1 2^{-dC(\alpha_1-\epsilon)}$ or the probability of error for the second decoder goes above $K_2 2^{-dC(\alpha_2-\epsilon)}$ for the decoding of j^{th} block for any $d \geq 0$. This implies

$$\mathcal{H}_{(\epsilon,K_1,K_2,j)}^d = \bigcup_{d \geq 0} \left(\mathcal{H}_{(1,\epsilon,K_1,j)}^d \cup \mathcal{H}_{(2,\epsilon,K_2,j)}^d \right).$$

Using (6), (7), and the union bound, we obtain

$$\begin{aligned}\mathbf{P}(\mathcal{H}_{(\epsilon,K_1,K_2,j)}^d) &\leq \sum_{d=0}^{\infty} \left\{ \mathbf{P}\left(\mathcal{H}_{(1,\epsilon,K_1,j)}^d\right) + \right. \\ &\quad \left. \mathbf{P}\left(\mathcal{H}_{(2,\epsilon,K_2,j)}^d\right) \right\} \\ &= \sum_{d=\frac{\log_2(K_1)}{C(\alpha_1-\epsilon)}}^{\infty} \mathbf{P}\left(\mathcal{H}_{(1,\epsilon,K_1,j)}^d\right) + \\ &\quad \sum_{d=\frac{\log_2(K_2)}{C(\alpha_2-\epsilon)}}^{\infty} \mathbf{P}\left(\mathcal{H}_{(2,\epsilon,K_2,j)}^d\right)\end{aligned}$$

$$\begin{aligned}&\leq \sum_{d=\frac{\log_2(K_1)}{C(\alpha_1-\epsilon)}}^{\infty} \frac{K_{1,j}^{\epsilon,\mathcal{F}} 2^{-dC\epsilon}}{K_1} + \\ &\quad \sum_{d=\frac{\log_2(K_2)}{C(\alpha_2-\epsilon)}}^{\infty} \frac{K_{2,j}^{\epsilon,\mathcal{F}} 2^{-dC\epsilon}}{K_2} \\ &\leq \frac{K_{1,j}^{\epsilon,\mathcal{F}} 2^{-\epsilon \left(\frac{\log_2(K_1)}{\alpha_1-\epsilon} \right)}}{K_1 (1-2^{-C\epsilon})} + \\ &\quad \frac{K_{2,j}^{\epsilon,\mathcal{F}} 2^{-\epsilon \left(\frac{\log_2(K_2)}{\alpha_2-\epsilon} \right)}}{K_2 (1-2^{-C\epsilon})}, \\ &= \frac{1}{1-2^{-C\epsilon}} \left(\frac{K_{1,j}^{\epsilon,\mathcal{F}}}{K_1^{\frac{\alpha_1}{\alpha_1-\epsilon}}} + \frac{K_{2,j}^{\epsilon,\mathcal{F}}}{K_2^{\frac{\alpha_2}{\alpha_2-\epsilon}}} \right).\end{aligned}$$

This can be made as small as desired by choosing K_1 and K_2 large enough. Hence we prove the lemma. \blacksquare

In the following theorem we prove a uniform bound on the constants $K_{1,j}^{\epsilon,\mathcal{F}}$ and $K_{2,j}^{\epsilon,\mathcal{F}}$ (over j).

Theorem 5: Let $\{\mathcal{F}\}$ be the set of deterministic encoders with rate pair $R_1 = A/C$ and $R_2 = B/C$. The encoders are generated with probability measure \mathbf{P} induced by the sequential encoding described in Subsection II-B. For all $d > 0$ and all $j \geq 0$, let the average error probability (with respect to \mathbf{P}) for the two decoders is given by

$$\mathbb{P}\{\hat{a}_{jA+1}^{(j+1)A} \neq a_{jA+1}^{(j+1)A}\} \leq K_1 2^{-dC\alpha_1}, \quad (8)$$

$$\mathbb{P}\{\hat{b}_{jB+1}^{(j+1)B} \neq b_{jB+1}^{(j+1)B}\} \leq K_2 2^{-dC\alpha_2}, \quad (9)$$

for some positive constants K_1 and K_2 which are independent of j and d . Then for every $\epsilon > 0$, $j \geq 0$, and $d > 0$, the error probability for almost every deterministic encoder \mathcal{F} satisfies,

$$\mathbb{P}_{\mathcal{F}}\{\hat{a}_{jA+1}^{(j+1)A} \neq a_{jA+1}^{(j+1)A}\} \leq K_1^{\mathcal{F}} 2^{-dC(\alpha_1-\epsilon)}, \quad (10)$$

$$\mathbb{P}_{\mathcal{F}}\{\hat{b}_{jB+1}^{(j+1)B} \neq b_{jB+1}^{(j+1)B}\} \leq K_2^{\mathcal{F}} 2^{-dC(\alpha_2-\epsilon)}, \quad (11)$$

where $K_1^{\mathcal{F}}$ and $K_2^{\mathcal{F}}$ depends only on the encoder \mathcal{F} and is independent of j and d .

Proof: Assume that the encoder has received bits a_1^{jA} (resp. b_1^{jB}) from the first (resp. second) source. Then define $\mathcal{F}(a_1^{jA}, b_1^{jB})$ to be the sequence of encoding functions starting from the $(j+1)^{\text{th}}$ block. The encoders $\mathcal{F}(a_1^{jA}, b_1^{jB})$ is independent of $\mathcal{F}(\tilde{a}_1^{jA}, \tilde{b}_1^{jB})$ if $a_1^{jA} \neq \tilde{a}_1^{jA}$ or $b_1^{jA} \neq \tilde{b}_1^{jA}$. In addition the encoders belonging to $\mathcal{F}(a_1^{jA}, b_1^{jB})$ are drawn from the same distribution as the actual encoders.

We now consider the first decoder. The following reasoning also follows exactly for the second decoder. Let $0 < \epsilon < \min(\alpha_1/2, \alpha_2/2)$. Then from Lemma 4 we know that there is a positive constant ν_1 independent of j such that

$$\mathbf{P}\left(K_{1,j}^{\epsilon,\mathcal{F}} > K_1\right) \leq \nu_1 K_1^{-\left(\frac{\alpha_1}{\alpha_1-\epsilon}\right)}.$$

From the arguments of [13], we conclude that $K_{1,j}^{\epsilon,\mathcal{F}}$ has finite mean.

For a given encoder \mathcal{F} , let $P(1, \mathcal{F}, j, d)$ be the probability that the first decoder makes an error in any of the block

numbered less than or equal to j after receiving the n^{th} block i.e. nC channel output, where $d = n - j$. More formally,

$$P(1, \mathcal{F}, j, d) = E_{\{a_1^{nA}, b_1^{nB}\}} \left[\mathbb{P}_{\mathcal{F}} \left(\hat{a}_1^{jA} \neq a_1^{jA} \right) \right].$$

As in [13], we now show by induction over j that there is a positive constant $L_j^{\mathcal{F}}$ such that $P(1, \mathcal{F}, j, d) \leq L_j^{\mathcal{F}} 2^{-dC(\alpha_1 - \epsilon)}$. This is satisfied for $j = 1$ from Lemma 4 with $L_1^{\mathcal{F}} = K_{1,1}^{\epsilon, \mathcal{F}}$. For general j we write

$$P(1, \mathcal{F}, j+1, d) \leq P(1, \mathcal{F}, j, d+1) + E_{\{a_1^{jA}, b_1^{jB}\}} \left[\mathbb{P} \left\{ a_{jA+1}^{(j+1)A} \neq \hat{a}_{jA+1}^{(j+1)A} \mid a_1^{jA} = \hat{a}_1^{jA} \right\} \right].$$

Note that the second term correspond to the error probability of the encoder $\mathcal{F}^{(a_1^{jA}, b_1^{jB})}$. Using the induction hypothesis and error probability bound for the encoder $\mathcal{F}^{(a_1^{jA}, b_1^{jB})}$, we obtain

$$P(1, \mathcal{F}, j+1, d) \leq L_j^{\mathcal{F}} 2^{-(d+1)C(\alpha_1 - \epsilon)} + E_{\{a_1^{jA}, b_1^{jB}\}} \left[P(1, \mathcal{F}^{(a_1^{jA}, b_1^{jB})}, j+1, d) \right].$$

Using Lemma 4, this can be upper bounded as

$$P(1, \mathcal{F}, j+1, d) \leq 2^{-dC(\alpha_1 - \epsilon)} \times \left(L_j^{\mathcal{F}} 2^{-C(\alpha_1 - \epsilon)} + E_{\{a_1^{jA}, b_1^{jB}\}} \left[K_{1,1}^{\epsilon, \mathcal{F}^{(a_1^{jA}, b_1^{jB})}} \right] \right).$$

This gives the following recursion for almost every encoder \mathcal{F} ,

$$L_{j+1}^{\mathcal{F}} = L_j^{\mathcal{F}} 2^{-C(\alpha_1 - \epsilon)} + \frac{1}{2^{j(A+B)}} \sum_{a_1^{jA} \in \{0,1\}^{jA}} \sum_{b_1^{jB} \in \{0,1\}^{jB}} K_{1,1}^{\epsilon, \mathcal{F}^{(a_1^{jA}, b_1^{jB})}}.$$

As was discussed earlier, the above empirical average converges towards the mean \bar{K}_1 . By using the generalization of the Chebychev's inequality given in [13] we know that $\exists B > 0$ such that

$$\mathbf{P} \left(\frac{1}{2^{j(A+B)}} \sum_{\substack{a_1^{jA} \in \{0,1\}^{jA} \\ b_1^{jB} \in \{0,1\}^{jB}}} K_{1,1}^{\epsilon, \mathcal{F}^{(a_1^{jA}, b_1^{jB})}} \geq \bar{K}_1 + \mu \right) \leq \frac{B}{\mu^{\frac{\alpha_1}{\alpha_1 - \epsilon}}} 2^{-\frac{\epsilon(A+B)j}{\alpha_1 - \epsilon}}.$$

Let \mathcal{M}_j^{μ} be the set of encoders for which the average exceeds $\bar{K}_1 + \mu$. Let $\mathcal{M}^{\mu} = \cup_{j \geq 1} \mathcal{M}_j^{\mu}$. Then

$$\begin{aligned} \mathbf{P}(\bar{\mathcal{M}}^{\mu}) &\geq 1 - \sum_{j=1}^{\infty} \mathbf{P}(\mathcal{M}_j^{\mu}) \\ &\geq 1 - \sum_{j=1}^{\infty} \frac{B}{\mu^{\frac{\alpha_1}{\alpha_1 - \epsilon}}} 2^{-\frac{\epsilon(A+B)j}{\alpha_1 - \epsilon}} \\ &= 1 - \frac{B}{\mu^{\frac{\alpha_1}{\alpha_1 - \epsilon}}} \frac{1}{\left(2^{\frac{\epsilon(A+B)}{\alpha_1 - \epsilon}} - 1 \right)}. \end{aligned}$$

Define $\bar{\mathcal{M}} = \cup_{\mu \geq 0} \bar{\mathcal{M}}^{\mu}$. This implies

$$\mathbf{P}(\bar{\mathcal{M}}) \geq \sup_{\mu \geq 0} \left(1 - \frac{B}{\mu^{\frac{\alpha_1}{\alpha_1 - \epsilon}}} \frac{1}{\left(2^{\frac{\epsilon(A+B)}{\alpha_1 - \epsilon}} - 1 \right)} \right) = 1.$$

Thus for almost every encoder \mathcal{F} , there is a constant $K_1^{\epsilon, \mathcal{F}}$ such that $P(1, \mathcal{F}, j, d) \leq K_1^{\epsilon, \mathcal{F}} 2^{-Cd(\alpha_1 - \epsilon)}$. As $P(1, \mathcal{F}, j, d)$ is an upper bound on the error probability in the j^{th} block, we show that (10) holds for almost every encoder. By the same reasoning for the second decoder, it can be shown that (11) holds for almost every encoder. By noting that the union of probability of two set of encoders \mathcal{E}_1 and \mathcal{E}_2 is lower bounded by

$$\mathbf{P}(\mathcal{E}_1 \cup \mathcal{E}_2) \geq 1 - \mathbf{P}(\bar{\mathcal{E}}_1) - \mathbf{P}(\bar{\mathcal{E}}_2),$$

we show that (10) and (11) together holds for almost every encoder. \blacksquare

V. CONCLUSION

In this paper we considered transmission of streaming sources over the bidirectional broadcast channel using sequential encoding and anytime decoding [13]. We showed that for encoding scheme using common randomness the error probability decays exponential in delay with positive error exponent for all the rate pair inside the capacity region. By extending Sahai's arguments from point to point channel to the bidirectional broadcast channel, we showed that there exist deterministic codes which have the same error probability performance as the codes with common randomness.

The bidirectional broadcast channel corresponds to the second phase of a bidirectional decode-and-forward relaying protocol with a half-duplex relay node. The first phase is the multiple access phase where both the sources communicate to the relay. This study is our first step towards a complete understanding of the error probability performance of half-duplex decode-and-forward relaying for streaming sources where we also take into account the error probability of the multiple access phase.

REFERENCES

- [1] B. Rankov and A. Wittneben, "Spectral Efficient Protocols for Half-Duplex Fading Relay Channels," *IEEE J. Select. Areas Commun.*, vol. 25, no. 2, pp. 379–389, Feb. 2007.
- [2] P. Larsson, N. Johansson, and K.-E. Sunell, "Coded Bi-directional Relaying," in *Proc. 5th Scandinavian Workshop on Ad Hoc Networks (ADHOC'05)*, Stockholm, Sweden, May 2005.
- [3] Y. Wu, P. Chou, and S.-Y. Kung, "Information Exchange in Wireless Networks with Network Coding and Physical-Layer Broadcast," in *Proc. of 39th Annual Conf. on Information Sciences and Systems*, Mar. 2005.
- [4] C. Fragouli, J.-Y. L. Boudec, and J. Widmer, "Network coding: An instant primer," *SIGCOMM Comput. Commun. Rev.*, vol. 36, no. 1, pp. 63–68, 2006.
- [5] P. Popovski and H. Yomo, "The antipackets can increase the achievable throughput of a wireless multihop network," in *Proc. IEEE Int. Conf. on Communications*, Istanbul, Turkey, Jun. 2006, pp. 3885–3890.
- [6] T. J. Oechtering, C. Schnurr, I. Bjelakovic, and H. Boche, "Broadcast Capacity Region of Two-Phase Bidirectional Relaying," *IEEE Trans. Inf. Theory*, vol. 54, no. 1, pp. 454–458, Jan. 2008.
- [7] S. J. Kim, P. Mitran, and V. Tarokh, "Performance Bounds for Bi-Directional Coded Cooperation Protocols," *IEEE Trans. Inf. Theory*, vol. 54, no. 11, pp. 5235–5241, Nov. 2008.

- [8] L.-L. Xie, "Network Coding and Random Binning for Multi-User Channels," in *Proc. 10th Canadian Workshop on Inf. Theory*, Edmonton, Alberta, Canada, Jun. 2007, pp. 85–88.
- [9] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network Information Flow," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, 2000.
- [10] T. J. Oechtering and M. Skoglund, "Upper bound to error probability for coding on bidirectional broadcast channels," in *Proc. Int. Conf. on Telecommunications*, April 2010, accepted.
- [11] R. G. Gallager, *Information Theory and Reliable Communication*. New York: John Wiley & Sons, Inc., 1968.
- [12] —, "A Simple Derivation of the Coding Theorem and Some Applications," *IEEE Trans. Inf. Theory*, vol. 18, no. 1, pp. 3–18, 1965.
- [13] A. Sahai, "Any-Time Information Theory," Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA, 2001.
- [14] A. Sahai and S. K. Mitter, "The necessity and sufficiency of anytime capacity for stabilization of a linear system over a noisy communication link Part I: scalar systems," *IEEE Trans. Inf. Theory*, vol. 52, no. 8, pp. 3369–3395, Aug. 2006.
- [15] C. Chang and A. Sahai, "Sequential random coding error exponents for multiple access channels," in *Wireless Networks, Communications and Mobile Computing, 2005 International Conference on*, vol. 2, June 2005, pp. 1581 – 1586.
- [16] —, "Sequential random coding error exponents for degraded broadcast channels," in *Allerton Conference on Control, Communications, and Computation*, Monticello, IL, 2005.
- [17] S. C. Draper, C. Chang, and A. Sahai, "Sequential Random Binning for Streaming Distributed Source Coding," in *Proc. Int. Symp. on Inf. Theory*, Adelaide, Australia, Sep. 2005, pp. 1396–1400.
- [18] G. D. Forney, "Convolutional Codes II. Maximum-Likelihood Decoding," *Information and Control*, vol. 25, pp. 222–266, 1974.