

Achievable Rates for Noisy Channels With Synchronization Errors

Mojtaba Rahmati, *Student Member, IEEE*, and Tolga M. Duman, *Fellow, IEEE*

Abstract—We develop several lower bounds on the capacity of binary input symmetric output channels with synchronization errors, which also suffer from other types of impairments such as substitutions, erasures, additive white Gaussian noise (AWGN), etc. More precisely, we show that if a channel suffering from synchronization errors as well as other type of impairments can be decomposed into a cascade of two component channels where the first one is another channel with synchronization errors and the second one is a memoryless channel (with no synchronization errors), a lower bound on the capacity of the original channel in terms of the capacity of the component synchronization error channel can be derived. A primary application of our results is that we can employ any lower bound derived on the capacity of the component synchronization error channel to find lower bounds on the capacity of the (original) noisy channel with synchronization errors. We apply the general ideas to several specific classes of channels such as synchronization error channels with erasures and substitutions, with symmetric q -ary outputs and with AWGN explicitly, and obtain easy-to-compute bounds. We illustrate that, with our approach, it is possible to derive tighter capacity lower bounds compared to the currently available bounds in the literature for certain classes of channels, e.g., deletion/substitution channels and deletion/AWGN channels (for certain signal-to-noise ratio (SNR) ranges).

Index Terms—Synchronization errors, insertion/deletion channels, channel capacity, achievable rates.

I. INTRODUCTION

DEPENDING on the transmission medium and the particular design, different limiting factors degrade the performance of a general communication system. For instance, imperfect alignment of the transmitter and receiver clocks may be one such factor resulting in a synchronization error channel modeled typically through insertion and/or deletion of symbols. Other factors include the effects of additive noise at the receiver among other impairments. The main objective of this paper is

Manuscript received May 7, 2014; revised September 21, 2014; accepted September 23, 2014. Date of publication March 10, 2014; date of current version November 18, 2014. This work was supported in part by the National Science Foundation under Contract NSF-TF 0830611. The work of T. M. Duman was supported by EC Marie Curie Career Integration under Grant PCIG12-GA-2012-334213. This paper was presented in part at the 2012 IEEE International Symposium on Information Theory (ISIT). The editor coordinating the review of this paper and approving it for publication was G. Colavolpe.

M. Rahmati was with the School of Electrical, Computer and Energy Engineering (ECEE), Arizona State University, Tempe, AZ 85287-5706 USA. He is now with Samsung Modem R&D, San Diego, CA 92121 USA (e-mail: mojtaba@asu.edu).

T. M. Duman is with the Department of Electrical and Electronics Engineering, Bilkent University, Ankara 06800, Turkey, on leave from the School of Electrical, Computer and Energy Engineering (ECEE), Arizona State University, Tempe, AZ 85287-5706 USA (e-mail: duman@ee.bilkent.edu.tr).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TCOMM.2014.2361330

to study the combined effects of the synchronization errors and additive noise type impairments, and in particular, to “decouple” the effects of the synchronization errors from (some of) the additive noise impairments to obtain expressions relating the channel capacity of the combined model and the component synchronization error channel.

We focus on achievable rates for channels which can be considered as a concatenation of two independent channels where the first one is a binary channel suffering from synchronization errors and the second one is either a memoryless binary input symmetric q -ary output (BSQO) channel or a binary input AWGN (BI-AWGN) channel. For instance, the first channel can be a binary insertion/deletion channel and the second one can be a binary symmetric channel (BSC) or a ternary output substitution/erasure channel. Specifically, we obtain achievable rates for the concatenated channel in terms of the capacity of the component synchronization error channel by lower bounding the information rate of the concatenated channel for input distributions which achieve the capacity of the component synchronization error channel and the statistical parameters of the memoryless channel. The lower bounds are derived without the use of the exact capacity achieving input distribution of the component synchronization error channel, hence any existing lower bound on the capacity (of the component synchronization error channel) can be employed to obtain an achievable rate characterization for the original channel model of interest.

By channels with synchronization errors we refer to the binary memoryless channels with synchronization errors as defined by Dobrushin in [1] where every transmitted bit is independently replaced with a random number of symbols (possibly the empty string, i.e., a deletion event is also allowed), and the transmitter and receiver have no information about the position and/or the pattern of the insertions/deletions. While different specific models on channels with synchronization errors are considered in the literature, commonly used models include insertion/deletion channels such as Gallager’s insertion/deletion channel [2], the sticky channel [3] and the segmented insertion/deletion channel [4].

Dobrushin [1] proved that Shannon’s theorem applies for discrete memoryless channels with synchronization errors by demonstrating that information stability holds. That is, for the capacity of the synchronization error channel, C_s , we can write $C_s = \lim_{N \rightarrow \infty} \max_{P(\mathbf{X})} \frac{1}{N} I(\mathbf{X}; \mathbf{Y})$ where \mathbf{X} and \mathbf{Y} are the transmitted and received sequences, respectively, and N is the length of the transmitted sequence. Therefore, the information and transmission capacities of the memoryless channels with synchronization errors are equal and we can employ any lower

bound on the information capacity as a lower bound on the transmission capacity of a channel with synchronization errors. In [5], the author extends Dobrushin's result on discrete memoryless channels with synchronization errors to the case of continuous output memoryless channels with synchronization errors.

There are many papers deriving upper and/or lower bounds on the capacity of the insertion/deletion channels, e.g., see [6]–[12]; however, only a very few results exist for insertion/deletion channels with substitution errors, e.g., [2], [13], [14] or in the presence of AWGN, e.g., [15], [16]. Our interest is on the latter, in fact, on more general models incorporating erasures as well as q -ary channel outputs.

Let us review some of the existing relevant results on insertion/deletion channels in a bit more detail. In [2], Gallager considers a channel model with substitution and insertion/deletion errors (sub/ins/del) where each bit gets deleted with probability p_d , replaced by two random bits with probability p_i , correctly received with probability $p_c = (1 - p_d - p_i)(1 - p_s)$, and changed with probability $p_f = (1 - p_d - p_i)p_s$, and derives a lower bound on the channel capacity (in bits/use) given by

$$C \geq 1 + p_d \log p_d + p_i \log p_i + p_c \log p_c + p_f \log p_f, \quad (1)$$

where $\log(\cdot)$ is in base 2. In [13], several upper and lower bounds are developed on the capacity of the ins/del/sub channel, where a genie-aided decoder is employed that is supplied with side information about some suitably selected random processes giving rise to an auxiliary memoryless channel in such a way that the Blahut-Arimoto algorithm (BAA) can be employed to obtain upper bounds on the capacity of the original channel. Furthermore, it is shown that by subtracting some quantity from the derived upper bounds which is, roughly speaking, more than the extra information provided by the side information, lower bounds on the capacity can also be derived. In [16], Monte Carlo simulation based results are used to estimate the information rates of different insertion and/or deletion channels even in the presence of intersymbol interference (ISI) in addition to AWGN with independent uniformly distributed (i.u.d.) input sequences. In [15], the synchronization errors are modeled as a Markov process and simulations are used to compute achievable information rates of an ISI channel with synchronization errors in the presence of AWGN. In [11], Rahmati and Duman compute analytical lower bounds on the capacity of the independent and identically distributed (i.i.d.) del/sub and i.i.d. del/AWGN channels, by lower bounding the mutual information rate between the transmitted and received sequences for i.u.d. input sequences focusing on small deletion probabilities.

The paper is organized as follows. In Section II, we formally give the models for binary input symmetric q -ary output channels with synchronization errors and BI-AWGN channels with synchronization errors. In Section III, we give two lemmas which will be useful in the proof of the result on BSQO channels with synchronization errors. In Section IV, we first present the main results on BSQO channels for arbitrary values of q , and then, we focus on the specific case of substitution/erasure/synchronization error channel which is a

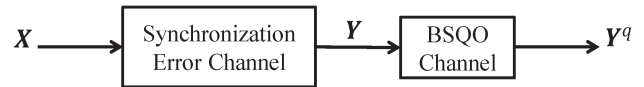


Fig. 1. Binary input symmetric q -ary output channel with synchronization errors.

binary input symmetric ternary output channel. In Section V, we lower bound the capacity of a synchronization error channel with AWGN (abbreviated as AWGN/synch channel) in terms of the capacity of the component synchronization error channel. We present several numerical examples illustrating the derived results in Section VI. Finally, we conclude the paper in Section VII.

II. CHANNEL MODELS

A general memoryless channel with synchronization errors [1] is defined via a stochastic matrix $\{p(y_i|x_i), y_i \in \mathcal{Y}, x_i \in \mathcal{X}\}$ where \mathcal{X} is the input alphabet (e.g., for a binary input channel $\mathcal{X} = \{0, 1\}$), and \mathcal{Y} is the (possibly empty) set of output symbols, $0 \leq p(y_i|x_i) \leq 1$, and $\sum_{y_i \in \mathcal{Y}} p(y_i|x) = 1$. As a particular instance of this channel, if $p(y_i = \emptyset|x_i) = p_d$ (\emptyset denoting the null string) and $p(y_i = x_i) = 1 - p_d$, we obtain an i.i.d. deletion channel.

A. Concatenation of a Synchronization Error Channel With a Memoryless BSQO Channel

To model synchronization errors as well as other types of impairments (e.g., substitution or erasure errors), we consider the concatenation of an arbitrary binary input binary output synchronization error channel with a memoryless BSQO channel. We refer to such a channel as a BSQO channel with synchronization errors. The concatenation of the two independent channels is depicted in Fig. 1 in which \mathbf{X} and \mathbf{Y} represent the input and output sequences of the component synchronization error channel, respectively, and \mathbf{Y} and $\mathbf{Y}^{(q)}$ represent the input and output sequences of the BSQO channel, respectively. By a symmetric channel we refer to the definition given in [17, p. 94], i.e., a channel is symmetric if by dividing the columns of the transition matrix into sub-matrices, in each sub-matrix, each row is a permutation of any other row and each column is a permutation of any other column. For example, a channel with independent substitution, erasure and synchronization errors (sub/ers/synch channel) can be considered as a concatenation of a synchronization error channel with input sequence \mathbf{X} and output sequence \mathbf{Y} and a substitution/erasure channel (binary input ternary output channel) with input sequence \mathbf{Y} and output sequence $\mathbf{Y}^{(3)}$. In such a model, each bit is independently flipped with probability p_s or erased with probability p_e , as illustrated in Fig. 2.

Another example is a binary input symmetric quaternary output channel with synchronization errors which can be decomposed into two independent channels such that the first one is a memoryless synchronization error channel and the second one is a memoryless binary input symmetric quaternary output channel illustrated in Fig. 3.

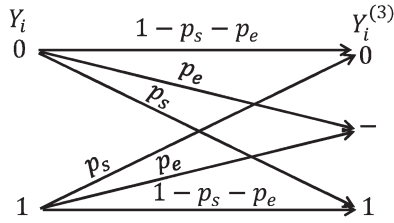


Fig. 2. Input-output relation in the substitution/erasure channel ($P(Y_i^{(3)}|Y_i)$ for all $1 \leq i \leq |\mathbf{Y}|$).

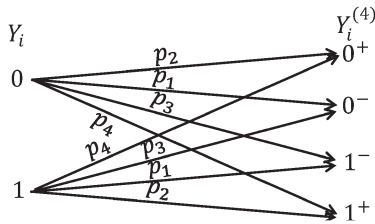


Fig. 3. Input-output relation in the binary input quaternary output channel ($P(Y_i^{(4)}|Y_i)$ for all $1 \leq i \leq |\mathbf{Y}|$).

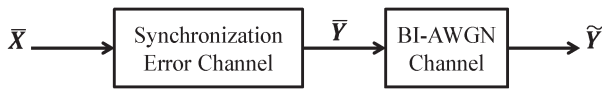


Fig. 4. AWGN channel with synchronization errors.

B. Concatenation of a Synchronization Error Channel With a BI-AWGN Channel

To model the effect of both synchronization errors and AWGN, we can consider concatenation of a binary input binary output synchronization error channel with a BI-AWGN channel. More precisely, in the overall channel, bits are transmitted using binary phase shift keying (BPSK) and the received signal contains AWGN in addition to the synchronization errors. Fig. 4 illustrates the overall channel model which we refer to as BI-AWGN channel with synchronization errors. We use $\bar{\mathbf{X}}$ to denote the input sequence to the first channel which is a BPSK modulated version of the binary input sequence \mathbf{X} , i.e., $\bar{X}_i = 1 - 2X_i$ and $\bar{\mathbf{Y}}$ to denote the output sequence of the first channel and input to the second one. $\tilde{\mathbf{Y}}$ is the output sequence of the second channel that is the noisy version of $\bar{\mathbf{Y}}$, i.e.,

$$\tilde{Y}_i = \bar{Y}_i + Z_i,$$

where Z_i 's are i.i.d. zero mean Gaussian random variables with variance σ^2 , and \tilde{Y}_i and \bar{Y}_i are the i^{th} received and transmitted bits of the second channel, respectively.

C. Simple Example of a Synchronization Error Channel Decomposition Into Two Independent Channels

The procedure used in finding the capacity bounds in this paper can be employed for any channel which can be decomposed into two independent channels such that the first one is a memoryless synchronization error channel and the second one is a symmetric memoryless channel with no effect on the length of the input sequence. Therefore, if we can decompose a given synchronization error channel into two channels with

described properties, we can derive lower bounds on the capacity of the concatenated channel. The decomposition of the original synchronization error channel into a well characterized synchronization error channel and a memoryless channel could be done in such a way that lower bounding the capacity of the component synchronization error channel can be simpler than the capacity analysis of the original synchronization error channel. In Table I, we provide an example of a hypothetical channel with synchronization errors that can be decomposed into a different synchronization error channel and a memoryless binary symmetric channel (BSC). In Table II, the two channels used in the decomposition are given.

III. ENTROPY BOUNDS FOR BINARY INPUT q -ARY OUTPUT CHANNELS WITH SYNCHRONIZATION ERRORS

In the following two lemmas, we provide a lower bound on the output entropy and an upper bound on the conditional output entropy of the binary input q -ary output channel in terms of the corresponding output entropies of the component synchronization error channel, respectively. Note that the following two lemmas hold for any binary input q -ary output channel with synchronization errors regardless of any symmetry.

Lemma 1: In any binary input q -ary output channel with synchronization errors and for all non-negative integer values of q , we have

$$H(\mathbf{Y}^{(q)}) - H(\mathbf{Y}) \geq -E_M \left\{ \log \left(\sum_{\mathbf{y}^{(q)} \neq \mathbf{y}} \sum_{p(\mathbf{y}^{(q)}|\mathbf{y}, M) p(\mathbf{y}^{(q)}|M)} \right) \right\},$$

where M is the random variable denoting the length of the received sequence, \mathbf{Y} denotes the output sequence of the component synchronization error channel and the input sequence of the binary input q -ary output channel, and $\mathbf{Y}^{(q)}$ denotes the output sequence of the binary input q -ary output channel.

Proof: By using two different expansions of $H(\mathbf{Y}^{(q)}, M)$, we have

$$\begin{aligned} H(\mathbf{Y}^{(q)}, M) &= H(\mathbf{Y}^{(q)}) + H(M|\mathbf{Y}^{(q)}) \\ &= H(\mathbf{Y}^{(q)}|M) + H(M). \end{aligned}$$

Since by knowing $\mathbf{Y}^{(q)}$, random variable M is also known, i.e., $H(M|\mathbf{Y}^{(q)}) = 0$, we can write

$$H(\mathbf{Y}^{(q)}) = H(\mathbf{Y}^{(q)}|M) + H(M),$$

Using the same approach for $H(\mathbf{Y})$ yields

$$H(\mathbf{Y}) = H(\mathbf{Y}|M) + H(M).$$

Finally, we can write

$$\begin{aligned} H(\mathbf{Y}^{(q)}) - H(\mathbf{Y}) &= H(\mathbf{Y}^{(q)}|M) - H(\mathbf{Y}|M) \\ &= \sum_m p(m) \left[H(\mathbf{Y}^{(q)}|m) - H(\mathbf{Y}|m) \right], \end{aligned} \tag{2}$$

TABLE I
TRANSITION PROBABILITIES OF THE EXAMPLE SYNCHRONIZATION ERROR CHANNEL

X_j	$P(Y_j X_j)$							
	$Y_j = 0$		$Y_j = 1$		$Y_j = 00$	$Y_j = 01$	$Y_j = 10$	$Y_j = 11$
0	$(\frac{1}{2} - (\alpha + \beta)) (1 + \sqrt{\frac{\alpha - \beta}{\alpha + \beta}})$	$(\frac{1}{2} - (\alpha + \beta)) (1 - \sqrt{\frac{\alpha - \beta}{\alpha + \beta}})$	$(\frac{1}{2} - (\alpha + \beta)) (1 - \sqrt{\frac{\alpha - \beta}{\alpha + \beta}})$	$(\frac{1}{2} - (\alpha + \beta)) (1 + \sqrt{\frac{\alpha - \beta}{\alpha + \beta}})$	α	β	β	α
1	$(\frac{1}{2} - (\alpha + \beta)) (1 - \sqrt{\frac{\alpha - \beta}{\alpha + \beta}})$	$(\frac{1}{2} - (\alpha + \beta)) (1 + \sqrt{\frac{\alpha - \beta}{\alpha + \beta}})$	$(\frac{1}{2} - (\alpha + \beta)) (1 + \sqrt{\frac{\alpha - \beta}{\alpha + \beta}})$	$(\frac{1}{2} - (\alpha + \beta)) (1 - \sqrt{\frac{\alpha - \beta}{\alpha + \beta}})$	α	β	β	α

TABLE II
TRANSITION PROBABILITIES OF TWO COMPONENT CHANNELS GIVING RISE TO THE SYNCHRONIZATION ERROR CHANNEL IN TABLE I

X_j	$P(Z_j X_j)$				$P(Y_j Z_j)$	
	$Z_j = 0$	$Z_j = 1$	$Z_j = 00$	$Z_j = 11$	$Y_j = 0$	$Y_j = 1$
0	$1 - 2(\alpha + \beta)$	0	$\alpha + \beta$	$\alpha + \beta$	$0.5 + 0.5\sqrt{\frac{\alpha - \beta}{\alpha + \beta}}$	$0.5 - 0.5\sqrt{\frac{\alpha - \beta}{\alpha + \beta}}$
1	0	$1 - 2(\alpha + \beta)$	$\alpha + \beta$	$\alpha + \beta$	$0.5 - 0.5\sqrt{\frac{\alpha - \beta}{\alpha + \beta}}$	$0.5 + 0.5\sqrt{\frac{\alpha - \beta}{\alpha + \beta}}$

where $p(m) = P(M = m)$. Using the definition of the entropy, we obtain

$$\begin{aligned} & H(\mathbf{Y}^{(q)}|m) - H(\mathbf{Y}|m) \\ &= E_{\mathbf{Y}^{(q)}} \left\{ -\log(p(\mathbf{Y}^{(q)})) \middle| m \right\} - E_{\mathbf{Y}} \left\{ -\log(p(\mathbf{Y})) \middle| m \right\} \\ &= E_{(\mathbf{Y}, \mathbf{Y}^{(q)})} \left\{ -\log \left(\frac{p(\mathbf{Y}^{(q)})}{p(\mathbf{Y})} \right) \middle| m \right\} \\ &= -\sum_{\mathbf{y}^{(q)}} \sum_{\mathbf{y}: p(\mathbf{y}) \neq 0} p(\mathbf{y}^{(q)}|\mathbf{y}, m) p(\mathbf{y}|m) \log \left(\frac{p(\mathbf{y}^{(q)}|m)}{p(\mathbf{y}|m)} \right), \end{aligned}$$

where $E_{\mathbf{Z}}\{\cdot\}$ denotes the expected value with respect to the random variable \mathbf{Z} . Since $-\log(x)$ is a convex function of x , by applying Jensen's inequality we arrive at

$$\begin{aligned} & H(\mathbf{Y}^{(q)}|m) - H(\mathbf{Y}|m) \\ &\geq -\log \left(\sum_{\mathbf{y}^{(q)}} \sum_{\mathbf{y}: p(\mathbf{y}) \neq 0} p(\mathbf{y}^{(q)}|\mathbf{y}, m) p(\mathbf{y}|m) \frac{p(\mathbf{y}^{(q)}|m)}{p(\mathbf{y}|m)} \right) \\ &= -\log \left(\sum_{\mathbf{y}^{(q)}} \sum_{\mathbf{y}: p(\mathbf{y}) \neq 0} p(\mathbf{y}^{(q)}|\mathbf{y}, m) p(\mathbf{y}^{(q)}|m) \right). \quad (3) \end{aligned}$$

By substituting this result into (2), the proof follows. ■

Lemma 2: In any binary input q -ary output channel with synchronization errors and for any input distribution, we have (for all $j \in \{1, \dots, M\}$)

$$H(\mathbf{Y}^{(q)}|\mathbf{X}) \leq H(\mathbf{Y}|\mathbf{X}) + E\{M\}H(Y_j^{(q)}|Y_j),$$

where Y_j denotes the j -th output bit of the synchronization error channel and j -th input bit of the binary input q -ary output channel, and $Y_j^{(q)}$ denotes the output symbol of the binary input q -ary output channel corresponding to the input bit Y_j .

Proof: For the conditional output entropy, one way of decomposition yields

$$\begin{aligned} H(\mathbf{Y}^{(q)}, \mathbf{Y}|\mathbf{X}) &= H(\mathbf{Y}|\mathbf{X}) + H(\mathbf{Y}^{(q)}|\mathbf{Y}, \mathbf{X}) \\ &= H(\mathbf{Y}|\mathbf{X}) + H(\mathbf{Y}^{(q)}|\mathbf{Y}), \quad (4) \end{aligned}$$

where the last equality follows since $\mathbf{X} \rightarrow \mathbf{Y} \rightarrow \mathbf{Y}^{(q)}$ form a Markov chain. Another way of decomposition for $H(\mathbf{Y}^{(q)}, \mathbf{Y}|\mathbf{X})$ is

$$H(\mathbf{Y}^{(q)}, \mathbf{Y}|\mathbf{X}) = H(\mathbf{Y}^{(q)}|\mathbf{X}) + H(\mathbf{Y}|\mathbf{Y}^{(q)}, \mathbf{X}).$$

From the above two different decompositions, we obtain

$$\begin{aligned} H(\mathbf{Y}^{(q)}|\mathbf{X}) &= H(\mathbf{Y}|\mathbf{X}) + H(\mathbf{Y}^{(q)}|\mathbf{Y}) - H(\mathbf{Y}|\mathbf{X}, \mathbf{Y}^{(q)}) \\ &\leq H(\mathbf{Y}|\mathbf{X}) + H(\mathbf{Y}^{(q)}|\mathbf{Y}). \end{aligned}$$

Since by knowing \mathbf{Y} , M is also known, we can write

$$H(\mathbf{Y}^{(q)}|\mathbf{Y}) = H(\mathbf{Y}^{(q)}|M, \mathbf{Y}).$$

Furthermore, since the second channel is memoryless, we obtain

$$\begin{aligned} H(\mathbf{Y}^{(q)}|\mathbf{Y}, M) &= \sum_m p(m)H(\mathbf{Y}^{(q)}|\mathbf{Y}, M = m) \\ &= \sum_m p(m)mH(Y_j^{(q)}|Y_j) \\ &= E\{M\}H(Y_j^{(q)}|Y_j), \quad (5) \end{aligned}$$

which concludes the proof. ■

Combining the results of Lemmas 1 and 2 yields

$$\begin{aligned} I(\mathbf{X}; \mathbf{Y}^q) &\geq I(\mathbf{X}; \mathbf{Y}) - E\{M\}H(Y_j^{(q)}|Y_j) \\ &\quad - E_M \left\{ \log \left(\sum_{\mathbf{y}^{(q)}} \sum_{\mathbf{y}: p(\mathbf{y}) \neq 0} p(\mathbf{y}^{(q)}|\mathbf{y}, M) p(\mathbf{y}^{(q)}|M) \right) \right\}, \end{aligned}$$

which is a lower bound on the mutual information between the transmitted and received sequences of the concatenated channel $I(\mathbf{X}; \mathbf{Y}^q)$ in terms of the mutual information between the transmitted and received sequences of the component synchronization error channel $I(\mathbf{X}; \mathbf{Y})$.

TABLE III
TRANSITION PROBABILITIES FOR A BINARY
INPUT 5-ARY OUTPUT CHANNEL

	$P(Y_j^{(5)} \bar{Y}_j)$				
Y_j	$Y_j^{(5)} = -2$	$Y_j^{(5)} = -1$	$Y_j^{(5)} = 0$	$Y_j^{(5)} = 1$	$Y_j^{(5)} = 2$
-1	p_2	p_1	p_0	p_{-1}	p_{-2}
1	p_{-2}	p_{-1}	p_0	p_1	p_2

TABLE IV
TRANSITION PROBABILITIES FOR A BINARY INPUT
SYMMETRIC 6-ARY OUTPUT CHANNEL

	$P(Y_j^{(6)} \bar{Y}_j)$					
Y_j	$Y_j^{(6)} = -3$	$Y_j^{(6)} = -2$	$Y_j^{(6)} = -1$	$Y_j^{(6)} = 1$	$Y_j^{(6)} = 2$	$Y_j^{(6)} = 3$
-1	p_3	p_2	p_1	p_{-1}	p_{-2}	p_{-3}
1	p_{-3}	p_{-2}	p_{-1}	p_1	p_2	p_3

IV. ACHIEVABLE RATES OVER BSQO CHANNELS WITH SYNCHRONIZATION ERRORS

In this section, we focus on BSQO channels with synchronization errors (as introduced in Section II-A) and provide lower bounds on their capacity. We first develop the result for an arbitrary value of q , and then specialize it for sub/ers/synch channels.

A. Binary Input Symmetric q -ary Output Channel With Synchronization Errors

The following theorem gives a lower bound on the capacity of the BSQO channel with synchronization errors for an arbitrary value of q . The transition probability values $P(Y_j^{(q)} = k|\bar{Y}_j = b)$ for different values of $b \in \{-1, 1\}$ and $k \in \mathcal{K}_q$ are represented by $P(Y_j^{(q)} = k|Y_j = b) = p_{k \times b}$ where

$$\mathcal{K}_q = \begin{cases} \left\{ -\frac{q-1}{2}, \dots, -1, 0, 1, \dots, \frac{q-1}{2} \right\}, & \text{odd } q \\ \left\{ -\frac{q}{2}, \dots, -1, 1, \dots, \frac{q}{2} \right\}, & \text{even } q. \end{cases}$$

For instance, Tables III and IV show transition probabilities for a binary input 5-ary output and binary input 6-ary output channels, respectively.

Theorem 1: The capacity of the BSQO channel with synchronization errors C_{Q_s} can be lower bounded by

$$C_{Q_s} \geq C_s - r(H(\mathcal{P}_q) + F(\mathcal{P}_q)) \quad (6)$$

in which

$$\mathcal{P}_q = \begin{cases} \left\{ p_{-\frac{q-1}{2}}, \dots, p_{\frac{q-1}{2}} \right\}, & \text{odd } q \\ \left\{ p_{-\frac{q}{2}}, \dots, p_{-1}, p_1, \dots, p_{\frac{q}{2}} \right\}, & \text{even } q \end{cases}$$

$$F(\mathcal{P}_q) = \begin{cases} \log \left(2p_0^2 + \sum_{k=1}^{\frac{q-1}{2}} (p_k + p_{-k})^2 \right), & \text{odd } q \\ \log \left(\sum_{k=1}^{\frac{q}{2}} (p_k + p_{-k})^2 \right), & \text{even } q \end{cases}$$

C_s denotes the capacity of the binary input synchronization error channel, and $r = \lim_{n \rightarrow \infty} \frac{E\{M\}}{n}$ with n and M denoting the length of the transmitted and received sequences, respectively.

To prove Theorem 1, we need the following three lemmas. In the first two lemmas, we give lower bounds on the output

entropy of the binary input q -ary output channels with synchronization errors in terms of the output entropy of the component binary input binary output synchronization error channels for arbitrary odd and even values of q , respectively. In the third one we give an upper bound on the conditional output entropy of the binary input q -ary output channel with synchronization errors in terms of the conditional output entropy of the component synchronization error channel.

Lemma 3: For a BSQO channel with synchronization errors, for any input distribution and any odd q , we have

$$H(\mathbf{Y}^{(q)}) \geq H(\mathbf{Y}) - E\{M\} \log \left(2p_0^2 + \sum_{k=1}^{\frac{q-1}{2}} (p_k + p_{-k})^2 \right),$$

where \mathbf{Y} denotes the output sequence of the component synchronization error channel and input sequence of the BSQO channel, $\mathbf{Y}^{(q)}$ denotes the output sequence of the BSQO channel, and M is the length of $\mathbf{Y}^{(q)}$.

Proof: From Lemma 1, we have

$$H(\mathbf{Y}^{(q)}) - H(\mathbf{Y}) \geq -E_M \left\{ \log \left(\sum_{\mathbf{y}^{(q)}} \sum_{\mathbf{y}, p(\mathbf{y}) \neq 0} p(\mathbf{y}^{(q)}|\mathbf{y}, M) p(\mathbf{y}^{(q)}|M) \right) \right\}. \quad (7)$$

For $p(\mathbf{y}^{(q)}|\mathbf{y}, M = m)$, we can write

$$\begin{aligned} p(\mathbf{y}^{(q)}|\mathbf{y}, M = m) &= \prod_{i=1}^m p(Y_i^{(q)}|Y_i) \\ &= \prod_{k=-\frac{q-1}{2}}^{\frac{q-1}{2}} p_k^{l_k}, \end{aligned} \quad (8)$$

where $l_k = \sum_{i=1}^m \delta(y_i y_i^{(q)} - k)$ in which $\delta(\cdot)$ denotes the Kronecker delta function. For example, in a binary input 5-ary output channel we have $p(-1102|1111) = p_{-1}p_1p_0p_2$. Let us define $d(\mathbf{y}, \mathbf{y}^{(q)})_{b \rightarrow k} = \sum_{i=1}^m \delta(y_i - b) \delta(y_i^{(q)} - k)$, $j_k = \sum_{i=1}^m \delta(y_i^{(q)} - k)$ and $m_k(\mathbf{y}^{(q)}) = j_k + j_{-k}$. For a fixed output sequence $\mathbf{y}^{(q)}$ of length m with j_k occurrence of $y_i^{(q)} = k$ and j_{-k} occurrence of $y_i^{(q)} = -k$, there are $2^{j_0} \prod_{k=1}^{\frac{q-1}{2}} \binom{j_k}{i_k} \binom{j_{-k}}{i_{-k}}$ possibilities for \mathbf{y} such that $i_k = d(\mathbf{y}, \mathbf{y}^{(q)})_{1 \rightarrow k}$ and $i_{-k} = d(\mathbf{y}, \mathbf{y}^{(q)})_{1 \rightarrow -k}$, therefore we can write

$$\begin{aligned} &\sum_{\mathbf{y}, p(\mathbf{y}) \neq 0} p(\mathbf{y}^{(q)}|\mathbf{y}, m) \\ &\leq 2^{j_0} p_0^{j_0} \prod_{k=1}^{\frac{q-1}{2}} \sum_{i_k=0}^{j_k} \binom{j_k}{i_k} p_k^{i_k} p_{-k}^{j_k - i_k} \sum_{i_{-k}=0}^{j_{-k}} \binom{j_{-k}}{i_{-k}} p_{-k}^{i_{-k}} p_k^{j_{-k} - i_{-k}} \\ &= 2^{j_0} p_0^{j_0} \prod_{k=1}^{\frac{q-1}{2}} (p_k + p_{-k})^{j_k + j_{-k}} \\ &= (2p_0)^{m_0(\mathbf{y}^{(q)})} \prod_{k=1}^{\frac{q-1}{2}} (p_k + p_{-k})^{m_k(\mathbf{y}^{(q)})}. \end{aligned} \quad (9)$$

Note that in deriving the inequality in (3), the summation is taken over the values of \mathbf{y} with $p(\mathbf{y}) \neq 0$. However, in (9) the summation is taken over all possible values of \mathbf{y} of length m (over all m -tuples), i.e., over all \mathbf{y} regardless of whether $p(\mathbf{y}) = 0$ or $p(\mathbf{y}) \neq 0$, which results in the upper bound in (9). Furthermore, by defining $\mathcal{M}_{m,q} = \{m_0, \dots, m_{\frac{q-1}{2}} | m_0 + \dots + m_{\frac{q-1}{2}} = m\}$, for odd q , we can write

$$\begin{aligned} & \sum_{\mathbf{y}^{(q)}} p(\mathbf{y}^{(q)} | m) \sum_{\mathbf{y}, p(\mathbf{y}) \neq 0} p(\mathbf{y}^{(q)} | \mathbf{y}, m) \\ & \leq \sum_{\mathcal{M}_{m,q}} \left(\sum_{\mathbf{y}^{(q)}, m_k(\mathbf{y}^{(q)}) = m_k} p(\mathbf{y}^{(q)} | m) \right. \\ & \quad \left. \times (2p_0)^{m_0} \prod_{k=1}^{\frac{q-1}{2}} (p_k + p_{-k})^{m_k} \right) \\ & = \sum_{\mathcal{M}_{m,q}} \binom{m}{m_0, \dots, m_{\frac{q-1}{2}}} p_0^{m_0} \prod_{l=1}^{\frac{q-1}{2}} (p_l + p_{-l})^{m_l} \\ & \quad \times (2p_0)^{m_0} \prod_{k=1}^{\frac{q-1}{2}} (p_k + p_{-k})^{m_k} \\ & = \left(2p_0^2 + \sum_{k=1}^{\frac{q-1}{2}} (p_k + p_{-k})^2 \right)^m. \end{aligned} \quad (10)$$

By substituting the result of (10) into the result of Lemma 1, we arrive at

$$H(\mathbf{Y}^{(q)}) \geq H(\mathbf{Y}) - E\{M\} \log \left(2p_0^2 + \sum_{k=1}^{\frac{q-1}{2}} (p_k + p_{-k})^2 \right),$$

which concludes the proof. \blacksquare

It is also worth noting as a side note that any capacity achieving input distribution over a discrete memoryless channel results in strictly positive output probabilities for possible output sequences of the channel [17, p. 95]. In other words, for any given memoryless channel $\mathbf{X} \rightarrow \mathbf{Y}$, for the capacity achieving input distribution $P(\mathbf{X})$, if there is an input \mathbf{x} such that $P(\mathbf{y}|\mathbf{x}) \neq 0$, then $P(\mathbf{y}) \neq 0$ as well. Therefore, for any component synchronization error channel in which $\mathcal{Y} = \{(y_1, \dots, y_m) | 0 \leq m \leq M, y_i \in \{0, 1\}\}$, if there is an input sequence \mathbf{x} with $P(\mathbf{y}|\mathbf{x}) \neq 0$, then $p(\mathbf{y}^{(q)}) > 0$ for all $\mathbf{y}^{(q)}$ of length m and all possible m . Hence, the bounds in (9) and (10) hold with equalities for these cases. For instance, i.i.d. deletion and random insertion channels satisfy this condition.

Lemma 4: For a BSQO channel with synchronization errors, for any input distribution and any even q , we have

$$H(\mathbf{Y}^{(q)}) \geq H(\mathbf{Y}) - E\{M\} \log \left(\sum_{k=1}^{\frac{q}{2}} (p_k + p_{-k})^2 \right).$$

Proof: For even values of q , we have

$$p(\mathbf{y}^{(q)} | \mathbf{y}, M = m) = \prod_{k=1}^{\frac{q}{2}} p_k^{l_k} p_{-k}^{l_{-k}}, \quad (11)$$

where $l_k = \sum_{i=1}^m \delta(y_i y_i^{(q)} - k)$ and $l_{-k} = \sum_{i=1}^m \delta(y_i y_i^{(q)} + k)$. Furthermore, for a fixed output sequence $\mathbf{y}^{(q)}$ of length m with j_k symbols of k , there are $\prod_{k=1}^{\frac{q}{2}} \binom{j_k}{i_k} \binom{j_{-k}}{i_{-k}}$ possibilities for \mathbf{y} with $i_k = d(\mathbf{y}, \mathbf{y}^{(q)})_{1 \rightarrow k}$ and $i_{-k} = d(\mathbf{y}, \mathbf{y}^{(q)})_{1 \rightarrow -k}$. Therefore, we can write

$$\begin{aligned} & \sum_{\mathbf{y}, p(\mathbf{y}) \neq 0} p(\mathbf{y}^{(q)} | \mathbf{y}, m) \\ & \leq \prod_{k=1}^{\frac{q}{2}} \sum_{i_k=0}^{j_k} \binom{j_k}{i_k} p_k^{i_k} p_{-k}^{j_k - i_k} \sum_{i_{-k}=0}^{j_{-k}} \binom{j_{-k}}{i_{-k}} p_{-k}^{i_{-k}} p_k^{j_{-k} - i_{-k}} \\ & = \prod_{k=1}^{\frac{q}{2}} (p_k + p_{-k})^{j_k + j_{-k}} \\ & = \prod_{k=1}^{\frac{q}{2}} (p_k + p_{-k})^{m_k(\mathbf{y}^{(q)})}. \end{aligned} \quad (12)$$

If we define $\mathcal{M}_{m,q} = \{m_0, \dots, m_{\frac{q}{2}} | m_0 + \dots + m_{\frac{q}{2}} = m\}$, for even q , by taking the summation over all the possibilities of $\mathbf{y}^{(q)}$ in (12), we obtain

$$\begin{aligned} & \sum_{\mathbf{y}^{(q)}} p(\mathbf{y}^{(q)} | m) \sum_{\mathbf{y}, p(\mathbf{y}) \neq 0} p(\mathbf{y}^{(q)} | \mathbf{y}, m) \\ & \leq \sum_{\mathcal{M}_{m,q}} \left(\sum_{\mathbf{y}^{(q)}, m_k(\mathbf{y}^{(q)}) = m_k} p(\mathbf{y}^{(q)} | m) \prod_{k=1}^{\frac{q}{2}} (p_k + p_{-k})^{m_k} \right) \\ & = \sum_{\mathcal{M}_{m,q}} \binom{m}{m_1, \dots, m_{\frac{q}{2}}} \prod_{l=1}^{\frac{q}{2}} (p_l + p_{-l})^{m_l} \prod_{k=1}^{\frac{q}{2}} (p_k + p_{-k})^{m_k} \\ & = \left(\sum_{k=1}^{\frac{q}{2}} (p_k + p_{-k})^2 \right)^m. \end{aligned} \quad (13)$$

By substituting the result of (13) into (7), we obtain

$$\begin{aligned} H(\mathbf{Y}^{(q)}) - H(\mathbf{Y}) & \geq -\log \left(\sum_{k=1}^{\frac{q}{2}} (p_k + p_{-k})^2 \right) \sum_m m p(m) \\ & = -E\{M\} \log \left(\sum_{k=1}^{\frac{q}{2}} (p_k + p_{-k})^2 \right). \end{aligned}$$

which concludes the proof. \blacksquare

Lemma 5: For a BSQO channel with synchronization errors, for any input distribution, we have

$$H(\mathbf{Y}^{(q)} | \mathbf{X}) \leq H(\mathbf{Y} | \mathbf{X}) + E\{M\} H(\mathcal{P}_q).$$

Proof: By using the result of Lemma 2, we can write

$$\begin{aligned} H(\mathbf{Y}^{(q)} | \mathbf{X}) & \leq E\{M\} H(\mathbf{Y}_j^{(q)} | \mathbf{Y}_j) + H(\mathbf{Y} | \mathbf{X}) \\ & = E\{M\} H(\mathcal{P}_q) + H(\mathbf{Y} | \mathbf{X}), \end{aligned}$$

which concludes the proof. \blacksquare

We can now complete the proof of the main theorem of this section.

Proof of Theorem 1: Let \mathbf{X} be distributed according to $P(\mathbf{X})$ for both $\mathbf{X} \rightarrow \mathbf{Y}$ and $\mathbf{X} \rightarrow \mathbf{Y}^{(q)}$ channels. Combining the results of Lemmas 4 and 5, we obtain

$$I(\mathbf{X}; \mathbf{Y}^{(q)}) \geq I(\mathbf{X}; \mathbf{Y}) - E\{M\} (H(\mathcal{P}_q) + F(\mathcal{P}_q)).$$

Furthermore, considering $P(\mathbf{X})$ as the capacity achieving input distribution of the channel $\mathbf{X} \rightarrow \mathbf{Y}$ and using the fact that M is independent of the input distribution $P(\mathbf{X})$, we can write

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} I(\mathbf{X}; \mathbf{Y}^{(q)}) &\geq \lim_{n \rightarrow \infty} \frac{1}{n} I(\mathbf{X}; \mathbf{Y}) - r (H(\mathcal{P}_q) + F(\mathcal{P}_q)) \\ &= C_s - r (H(\mathcal{P}_q) + F(\mathcal{P}_q)). \end{aligned}$$

Hence for C_{Q_s} , we have

$$\begin{aligned} C_{Q_s} &\geq \lim_{n \rightarrow \infty} \frac{1}{n} I(\mathbf{X}; \mathbf{Y}^{(q)}) \\ &\geq C_s - r (H(\mathcal{P}_q) + F(\mathcal{P}_q)), \end{aligned}$$

concluding the proof. \blacksquare

B. Substitution/Erasure Channels With Synchronization Errors

We now consider sub/ers channel as a special case of binary input symmetric q -ary output channel with $q = 3$. From the result of Theorem 1, we obtain the following corollary.

Corollary 1: The capacity of the sub/ers/synch channel C_{ses} can be lower bounded by

$$C_{ses} \geq C_s - r [H(p_s, p_e, 1 - p_s - p_e) + \log((1 - p_e)^2 + 2p_e^2)], \quad (14)$$

where C_s denotes the capacity of the synchronization error channel.

Since we have considered the general synchronization error channel model of Dobrushin [1], the lower bound (14) holds for many different models on channels with synchronization errors. A popular model for channels with synchronization errors is the Gallager's ins/del model.¹ If we employ the Gallager's model for which the parameter r is given by

$$\begin{aligned} r &= \lim_{n \rightarrow \infty} \frac{E\{M\}}{n} \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} E\{|s|\} \\ &= 1 - p_d + p_i, \end{aligned} \quad (15)$$

where $|s|$ denotes the length of the output sequence in one use of the ins/del channel, and the equality results since the channel is memoryless. By utilizing the result of (15) in (14), we obtain the following two corollaries.

¹In fact, Gallager's model in general refers to a channel with insertion, deletion and substitution errors, but with Gallager's ins/del model we refer to the case with $p_s = 0$ (i.e., substitution error probability being zero).

Corollary 2: The capacity of the sub/ers/ins/del channel C_{seid} is lower bounded by

$$\begin{aligned} C_{seid} &\geq C_{id} - (1 - p_d + p_i) \\ &\quad \times [H(p_s, p_e, 1 - p_s - p_e) + \log((1 - p_e)^2 + 2p_e^2)], \end{aligned} \quad (16)$$

where C_{id} denotes the capacity of an insertion/deletion channel with parameters p_d and p_i .

Taking $p_e = 0$ in this channel model gives the ins/del/sub channel, hence we have the following simple result.

Corollary 3: The capacity of the ins/del/sub channel C_{ids} can be lower bounded by

$$C_{ids} \geq C_{id} - (1 - p_d + p_i) H_b(p_s). \quad (17)$$

V. ACHIEVABLE RATES OVER BI-AWGN CHANNELS WITH SYNCHRONIZATION ERRORS

In this section, we consider a binary synchronization error channel in the presence of AWGN as defined in Section II-B, and we derive a lower bound on its capacity. The main result is given in the following theorem.

Theorem 2: Let C_s denote the capacity of the synchronization error channel, then for the capacity of the AWGN/synch channel C_{As} , we have

$$\begin{aligned} C_{As} &\geq C_s - r \left[\log(e) \left(\frac{2}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}} - \frac{2}{\sigma^2} Q\left(\frac{1}{\sigma}\right) \right) \right. \\ &\quad \left. + \log\left(1 + Q\left(\frac{1}{\sigma}\right) e^{\frac{4}{\sigma^2}} Q\left(\frac{3}{\sigma}\right)\right) \right]. \end{aligned} \quad (18)$$

where $Q(\cdot)$ is the right tail probability of the standard normal distribution.

Proof: To prove the theorem, we first define an appropriate symmetric non-uniform quantizer with $2W$ quantization levels. Then, by letting W go to infinity and employing the result of Theorem 1, we complete the proof. Noting that any processing on the output sequence of a channel results in a lower bound on the capacity of the channel, it is guaranteed that our approach will result in a lower bound on the capacity of the synch/AWGN channel as desired.

By utilizing a symmetric quantizer with $2W$ quantization levels on the output symbols \tilde{Y}_j , for the transition probabilities of the resulting binary input symmetric $2W$ -ary output channel, we have

$$\begin{aligned} p_w &= P(Y^{(2W)} = bw | \bar{X}_j = b) \\ &= \begin{cases} P(t_{w-1} < \tilde{Y}_j < t_w | \bar{X}_j = 1), & 0 < w \leq W \\ P(t_{-w} < \tilde{Y}_j < t_{1-w} | \bar{X}_j = -1), & -W \leq w < 0 \end{cases}, \end{aligned}$$

where $t_{-w} = -t_w$, $t_0 = 0$ and $t_{w-1} < t_w$ for $w = \{1, \dots, W\}$. We choose the quantization step sizes, i.e., $\Delta_w = t_w - t_{w-1}$ for $w = \{1, \dots, W\}$, to satisfy $p_1 = p_2 = \dots = p_W$. Note that due to the symmetry of the quantizer, we have $\Delta_{-w} = \Delta_w$ (as illustrated in Fig. 5). Furthermore, by defining $P = Q\left(\frac{1}{\sigma}\right)$,

we have $\sum_{w=1}^W p_{-w} = P$ and $\sum_{w=1}^W p_w = 1 - P$ which results in $p_w = \frac{1-P}{W}$ for $w = \{1, \dots, W\}$.

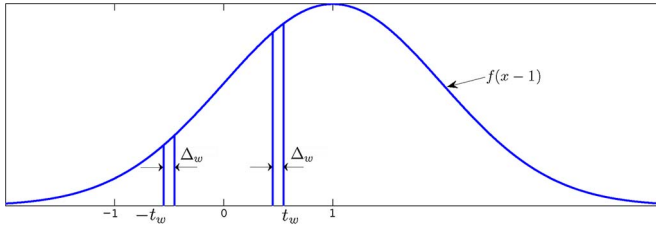


Fig. 5. Symmetric non-uniform quantizer step sizes.

Using the result of Theorem 1, to derive a lower bound on the capacity of the channel with $2W$ -level quantized outputs, we need to obtain $H(\mathcal{P}_{2W}) + \log\left(\sum_{w=1}^W (p_w + p_{-w})^2\right)$ in which $\mathcal{P}_{2W} = \{p_{-W}, \dots, p_{-1}, p_1, \dots, p_W\}$. It can be shown that (see Appendix A)

$$\begin{aligned} & \lim_{W \rightarrow \infty} \left(H(\mathcal{P}_{2W}) + \log\left(\sum_{w=1}^W (p_w + p_{-w})^2\right) \right) \\ &= \log(e) \left(\frac{2e^{-\frac{1}{2\sigma^2}}}{\sqrt{2\pi}\sigma} - \frac{2}{\sigma^2} P \right) + \log\left(1 + P + e^{\frac{4}{\sigma^2}} Q\left(\frac{3}{\sigma}\right)\right) \end{aligned}$$

Finally, by substituting this result into (6), the proof follows. ■

One can also consider the use of a uniform quantizer which after some computations gives

$$C_{As} \geq C_s - r \log\left(\sqrt{\frac{e}{2}}(1 + e^{-\frac{1}{\sigma^2}})\right). \quad (19)$$

However, the symmetric non-uniform quantizer considered in Theorem 2 provides a tighter lower bound on the capacity of the AWGN/synch channel compared to the one in (19). For instance, even for $\sigma = 0$, i.e., the noiseless scenario, unlike the expression in (18), (19) does not match with the trivial result $C_{As} = C_s$.

VI. NUMERICAL EXAMPLES

In this section, we give several numerical examples of the lower bounds on the capacity of the ins/del/sub and del/AWGN channels and compare them with the existing ones in the literature.

A. Insertion/Deletion/Substitution Channel

In Table V, we compare the lower bound on the capacity of the ins/del/sub channel (17) with the existing lower bounds in [2], [13] for several values of p_d , p_i and p_s . We employ the lower bound derived in [7] as the lower bound on the capacity of the deletion channel and the lower bound in [13] as the lower bound on the capacity of the ins/del channel in (17). Note that the Gallager's model in [2] by parameters p_d , p_i and p_c can be considered as concatenation of an ins/del channel with parameters p_d and p_i , and a BSC channel with cross error probability of p_s where p_s is the solution of $p_c = (1 - p_s)(1 - p_d - p_i)$.

TABLE V
COMPARISON OF THE LOWER BOUND DERIVED ON THE CAPACITY OF THE INS/DEL/SUB CHANNEL WITH THE EXISTING LOWER AND UPPER BOUNDS (BOLDFACE VALUES SHOW THE BEST BOUNDS)

p_d	p_i	p_s	LB from [2]	LB (17)	LB from [13]	UB from [13]
0.001	0.00	0.001	0.9772	0.9775	0.9773	0.9856
0.001	0.00	0.01	0.9079	0.9082	0.9081	0.9163
0.001	0.00	0.1	0.5201	0.5204	0.5210	0.5292
0.01	0.00	0.001	0.9079	0.9107	0.9091	0.9586
0.01	0.00	0.01	0.839	0.842	0.842	0.886
0.01	0.00	0.10	0.454	0.458	0.466	0.510
0.10	0.000	0.001	0.5207	0.5514	0.5346	0.7300
0.10	0.00	0.01	0.458	0.489	0.492	0.644
0.10	0.00	0.10	0.108	0.140	0.211	0.363
0.10	0.10	0.001	0.0689	0.1678	0.1761	0.4504
0.10	0.10	0.01	0.013	0.0984	0.139	0.438

The advantage of the lower bound (17) is in using the tightest lower bound on the capacity of the ins/del channel (lower bound from [13]) in lower bounding the capacity of the overall channel, i.e., the information rate of the overall channel is lower bounded for the input distribution which results in the tightest lower bound on the capacity of the ins/del channel. We observe that for $p_i = 0$, a fixed p_d and small values of p_s , the lower bound (17) improves the lower bound given in [13]. This is not unexpected, because for small values of p_s the input distribution achieving the capacity of the i.i.d. deletion channel is not far from the optimal input distribution of the del/sub channel. We also observe that the lower bound (17) outperforms the lower bound given in [2]. However, for the case $p_i \neq 0$ it does not improve the result given in [13]. The reason is that we substituted the result from [13] as a lower bound on the capacity of the ins/del in (17) to obtain a lower bound on the capacity of the ins/del/sub channel.

B. Insertion/Deletion/AWGN Channel

We now give several numerical examples of the lower bound (18) on the capacity of the ins/del/AWGN channel and compare them with existing results. For comparison purposes, we consider the simulation based bound of [16] obtained for i.u.d. input sequences through Monte-Carlo simulations.

Fig. 6 shows a comparison of the lower bound on the capacity of the del/AWGN channel in (18) with the results in [16]. We observe from Fig. 6 that for small values of σ^2 , the lower bound (18) improves the one given in [16]. This is because in deriving the capacity lower bound (18) for the del/AWGN channel, we employ the tightest lower bound available on the capacity of the deletion channel while the result in [16] uses i.u.d. inputs. Furthermore, we observe that the lower bound (18) is not close to the simulation based results of [16] for large σ^2 values and small deletion probabilities. This is not unexpected, because in [16], the achievable information rate for i.u.d. input sequences are obtained (through lengthy Monte-Carlo simulations) and i.u.d. inputs are close to optimal. We further note that the procedure employed in [16] is only useful for computing capacity lower bounds for small values of deletion probabilities, e.g., $p_d \leq 0.1$, while the lower bound in (18) holds for the entire range of deletion probabilities by employing any lower bound on the capacity of the deletion channel in lower bounding the capacity of the deletion/AWGN channel.

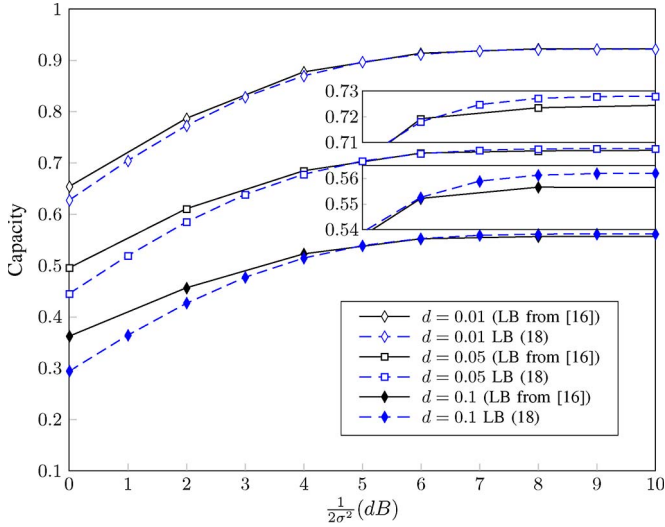


Fig. 6. Comparison between the lower bound (18) with the lower bound in [16] versus SNR for different deletion probabilities.

VII. SUMMARY AND CONCLUSION

In this paper, we presented several lower bounds on the capacity of binary input symmetric output channels with synchronization errors in addition to substitutions, erasures or AWGN. We showed that the capacity of any channel with synchronization errors which can be considered as a cascade of two channels (where only the first one suffers from synchronization errors and the second one is a memoryless channel) can be lower bounded in terms of the capacity of the first channel and the parameters of the second channel. We considered two classes of channels: binary input symmetric q -ary output channels (e.g., for $q = 3$ a binary input channel with substitutions and erasures) with synchronization errors and BI-AWGN channels with synchronization errors. We gave the first lower bound on the capacity of substitution/erasure channel with synchronization errors and the first analytical result on the capacity of BI-AWGN channel with synchronization errors. We also demonstrated that the lower bounds developed on the capacity of the del/AWGN channel for small σ^2 values and the del/sub channel for small values of p_s improve the existing ones.

APPENDIX A

PART OF THE PROOF OF THEOREM 2

To calculate $H(\mathcal{P}_{2W}) + \log\left(\sum_{w=1}^W (p_w + p_{-w})^2\right)$, we first compute $H_W = H(\mathcal{P}_{2W}) - \log(W)$, then $\log\left(\sum_{w=1}^W (p_w + p_{-w})^2\right) + \log(W)$. For H_W , we have

$$\begin{aligned} H_W &= -\sum_{w=1}^W p_w \log(p_w) - \sum_{w=1}^W p_{-w} \log(p_{-w}) - \log(W) \\ &= -(1-P) \log(1-P) - \sum_{w=1}^W p_{-w} \log(W p_{-w}). \end{aligned} \quad (20)$$

To calculate $-\sum_{w=1}^W p_{-w} \log(W p_{-w})$, we first derive a relation between p_w and p_{-w} by using the fact that $\Delta_w = \Delta_{-w}$. For

large W and $w = \{1, \dots, W\}$, we have $p_w \cong f(1-t_w)\Delta_w$ and $p_{-w} \cong f(1+t_w)\Delta_w$, where $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}$. Furthermore, since $p_w = \frac{1-P}{W}$ for $w = \{1, \dots, W\}$ and $\frac{f(1+t_w)}{f(1-t_w)} = e^{-\frac{2t_w}{\sigma^2}}$, we can write

$$\begin{aligned} p_{-w} &\cong \frac{f(1+t_w)}{f(1-t_w)} p_w \\ &= \frac{1-P}{W} e^{-\frac{2t_w}{\sigma^2}}, \end{aligned}$$

with the understanding that the approximation becomes exact as $W \rightarrow \infty$. By using this result, we obtain

$$\begin{aligned} \lim_{W \rightarrow \infty} -\sum_{w=1}^W p_{-w} \log(W p_{-w}) &= \lim_{W \rightarrow \infty} -\sum_{w=1}^W p_{-w} \left(\log(1-P) + \log\left(e^{-\frac{2t_w}{\sigma^2}}\right) \right) \\ &= -P \log(1-P) - \lim_{W \rightarrow \infty} \sum_{w=1}^W p_{-w} \log\left(e^{-\frac{2t_w}{\sigma^2}}\right), \end{aligned} \quad (21)$$

where we used the fact that $\sum_{w=1}^W p_{-w} = P$. Furthermore, for

$$\begin{aligned} \lim_{W \rightarrow \infty} -\sum_{w=1}^W p_{-w} \log\left(e^{-\frac{2t_w}{\sigma^2}}\right), \text{ we can write} \\ \lim_{W \rightarrow \infty} -\sum_{w=1}^W p_{-w} \log\left(e^{-\frac{2t_w}{\sigma^2}}\right) &= \lim_{W \rightarrow \infty} \log(e) \sum_{w=1}^W f(1+t_w)\Delta_w \frac{2t_w}{\sigma^2} \\ &= \log(e) \int_0^\infty f(1+t) \frac{2t}{\sigma^2} dt \\ &= \log(e) \frac{2}{\sigma^2} \int_0^\infty \frac{t}{\sqrt{2\pi}\sigma} e^{-\frac{(t+1)^2}{2\sigma^2}} dt \\ &= \log(e) \left(\frac{2}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}} - \frac{2}{\sigma^2} P \right). \end{aligned} \quad (22)$$

By substituting (22) and (21) into (20), we obtain

$$\lim_{W \rightarrow \infty} H_W = -\log(1-P) + \log(e) \left(\frac{2}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}} - \frac{2}{\sigma^2} P \right). \quad (23)$$

At this point, we only need to obtain the exact value of $\sum_{w=1}^W (p_w + p_{-w})^2$, where we have

$$\begin{aligned} \sum_{w=1}^W (p_w + p_{-w})^2 &= \sum_{w=1}^W (p_w^2 + 2p_w p_{-w} + p_{-w}^2) \\ &= (1-P)^2 + 2P(1-P) + \sum_{w=1}^W W p_{-w}^2. \end{aligned} \quad (24)$$

Furthermore, if we let W go to infinity, for $\sum_{w=1}^W Wp_{-w}^2$, we can write

$$\begin{aligned} & \lim_{W \rightarrow \infty} \sum_{w=1}^W Wp_{-w}^2 \\ &= \lim_{W \rightarrow \infty} \sum_{w=1}^W Wf(1+t_w)\Delta_w \frac{f(1+t_w)}{f(1-t_w)} p_w \\ &= \lim_{W \rightarrow \infty} (1-P) \sum_{w=1}^W \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(t_w+1)^2}{2\sigma^2}} e^{-\frac{2t_w}{\sigma^2}} \Delta_w \\ &= (1-P) \int_0^\infty \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(t+1)^2}{2\sigma^2}} e^{-\frac{2t}{\sigma^2}} dt \\ &= (1-P) \int_0^\infty \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(t+3)^2-8}{2\sigma^2}} dt \\ &= (1-P)e^{\frac{4}{\sigma^2}} Q\left(\frac{3}{\sigma}\right). \end{aligned} \tag{25}$$

Using the results of (25) and (23), we obtain

$$\begin{aligned} & \lim_{W \rightarrow \infty} \left(H(\mathcal{P}_{2W}) + \log \left(\sum_{w=1}^W (p_w + p_{-w})^2 \right) \right) \\ &= \log(e) \left(\frac{2e^{-\frac{1}{2\sigma^2}}}{\sqrt{2\pi\sigma}} - \frac{2}{\sigma^2} P \right) + \log \left(1 + P + e^{\frac{4}{\sigma^2}} Q\left(\frac{3}{\sigma}\right) \right). \end{aligned}$$

REFERENCES

[1] R. L. Dobrushin, "Shannon's theorems for channels with synchronization errors," *Probl. Inf. Transmiss.*, vol. 3, no. 4, pp. 11–26, 1967.
 [2] R. G. Gallager, "Sequential decoding for binary channels with noise and synchronization errors," MIT Lincoln Lab., Lexington, MA, USA, Tech. Rep., 1961, Group Report.
 [3] M. Mitzenmacher, "Capacity bounds for sticky channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 1, pp. 72–77, Jan. 2008.
 [4] Z. Liu and M. Mitzenmacher, "Codes for deletion and insertion channels with segmented errors," *IEEE Trans. Inf. Theory*, vol. 56, no. 1, pp. 224–232, Jan. 2010.
 [5] S. Z. Stambler, "Memoryless channels with synchronization errors: The general case," *Probl. Inf. Transmiss.*, vol. 6, no. 3, pp. 223–229, 1970.
 [6] E. Drinea and M. Mitzenmacher, "Improved lower bounds for the capacity of i.i.d. deletion and duplication channels," *IEEE Trans. Inf. Theory*, vol. 53, no. 8, pp. 2693–2714, Aug. 2007.
 [7] A. Kirsch and E. Drinea, "Directly lower bounding the information capacity for channels with i.i.d. deletions and duplications," *IEEE Trans. Inf. Theory*, vol. 56, no. 1, pp. 86–102, Jan. 2010.
 [8] D. Fertonani and T. M. Duman, "Novel bounds on the capacity of the binary deletion channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2753–2765, Jun. 2010.
 [9] Y. Kanoria and A. Montanari, "On the deletion channel with small deletion probability," in *Proc. IEEE ISIT*, Jun. 2010, pp. 1002–1006.

[10] A. Kalai, M. Mitzenmacher, and M. Sudan, "Tight asymptotic bounds for the deletion channel with small deletion probabilities," in *Proc. IEEE ISIT*, Jun. 2010, pp. 997–1001.
 [11] M. Rahmati and T. M. Duman, "Bounds on the capacity of random insertion and deletion-additive noise channels," *IEEE Trans. Inf. Theory*, vol. 59, no. 9, pp. 5534–5546, Sep. 2013.
 [12] R. Venkataramanan, S. Tatikonda, and K. Ramchandran, "Achievable rates for channels with deletions and insertions," *IEEE Trans. Inf. Theory*, vol. 59, no. 11, pp. 6990–7013, Nov. 2013.
 [13] D. Fertonani, T. M. Duman, and M. F. Erden, "Bounds on the capacity of channels with insertions, deletions and substitutions," *IEEE Trans. Commun.*, vol. 59, no. 1, pp. 2–6, Jan. 2011.
 [14] H. Mercier, V. Tarokh, and F. Labeau, "Bounds on the capacity of discrete memoryless channels corrupted by synchronization and substitution errors," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4306–4330, Jul. 2012.
 [15] W. Zeng, J. Tokas, R. Motwani, and A. Kavcic, "Bounds on mutual information rates of noisy channels with timing errors," in *Proc. IEEE ISIT*, Sep. 2005, pp. 709–713.
 [16] J. Hu, T. M. Duman, M. F. Erden, and A. Kavcic, "Achievable information rates for channels with insertions, deletions and intersymbol interference with i.i.d. inputs," *IEEE Trans. Commun.*, vol. 58, no. 4, pp. 1102–1111, Apr. 2010.
 [17] R. G. Gallager, *Information Theory and Reliable Communication*. New York, NY, USA: Wiley, 1968.



Mojtaba Rahmati (S'11) received the B.S. degree in electrical engineering from the University of Tehran, Tehran, Iran, in 2007; the M.S. degree in telecommunication systems from Sharif University of Technology, Tehran, in 2009; and the Ph.D. degree in electrical engineering from Arizona State University, Tempe, AZ, USA, in 2013. He is currently a Research Engineer with Samsung Modem R&D, San Diego, CA, USA. His research interests include wireless and mobile communication, digital communications, digital signal processing, and information

theory.



Tolga M. Duman (S'95–M'98–SM'03–F'11) received the B.S. degree from Bilkent University, Ankara, Turkey, in 1993 and the M.S. and Ph.D. degrees from Northeastern University, Boston, MA, USA, in 1995 and 1998, respectively, all in electrical engineering. He is currently a Professor with the Department of Electrical and Electronics Engineering, Bilkent University, Ankara, Turkey, and is on leave from the School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe, AZ, USA. Prior to joining Bilkent University

in September 2012, he has been with the Department of Electrical Engineering, Arizona State University first as an Assistant Professor (1998–2004), then as an Associate Professor (2004–2008), and currently as a Professor (August 2008–present). His current research interests include communication and signal processing systems, particularly on wireless and mobile communications, coding/modulation, coding for wireless communications, data storage systems, and underwater acoustic communications.

Dr. Duman is a Fellow of IEEE, a recipient of the National Science Foundation CAREER Award, and an IEEE Third Millennium medal. He served as an Editor for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS (2003–2008), IEEE TRANSACTIONS ON COMMUNICATIONS (2007–2012), and IEEE ONLINE JOURNAL OF SURVEYS AND TUTORIALS (2002–2007). He is currently the Coding and Communication Theory Area Editor for IEEE TRANSACTIONS ON COMMUNICATIONS (2011–present) and an Editor for *Elsevier Physical Communications Journal* (2010–present).