Bounds on Mutual Information Rates of Noisy Channels with Timing Errors

Wei Zeng[†], Jitender Tokas[‡], Ravi Motwani[*], Aleksandar Kavčić[†]
[†] DEAS, Harvard University, Cambridge, MA 02138, USA;
[‡] Data Storage Institute, Singapore; [*] Broadcom Corporation, USA

***Abstract*— In this paper, we consider the information rates of baseband linear filter channels with timing errors. We assume that the timing error can be modeled as a discrete-valued Markov process. Based on this assumption, we turn the channel into a finite-state machine model and show that the output sequence is an ergodic and asymptotically stationary hidden Markov process. We then propose Monte-Carlo methods to compute the upper and lower bounds on the information rate by utilizing the entropy ergodic theorem.**

I. Introduction

Communication and data storage systems typically suffer from synchronization errors. Virtually all timing recovery methods at the receiver produce synchronization errors. Such synchronization errors are negligible in most conventional receivers, where the timing recovery units operate at very high signal-to-noise ratios (SNRs). With the advent of powerful iteratively decodable codes, decoders are capable of operating at low SNRs. However, in these low SNR regions, conventional timing recovery units fail, and new timing recovery schemes are needed to benefit from the power of the channel codes.

It is therefore important to evaluate the theoretical limits of transmission rates for channels with timing errors. Most earlier work has been devoted to computing the capacity of special types of channels with timing error, namely, the insertion/deletion channels. Both analytical bounds [1], [2], [3], [4] and simulation-based methods [5], [6] have been proposed. In [7], Dobrushin proved Shannon's theorem for memoryless channels with synchronization errors. The results are given in terms of the supremum of the mutual information rate. However, an analytic (or even numeric) evaluation of this information rate (and its supremum) is still an open problem.

In this paper, we study a more general case than the insertion/deletion channel. In our model, the timing error can be a quantized fraction of the symbol interval. We assume that the timing error process is a discrete-valued Markov chain, and study the information rate of baseband linear filter channels corrupted by such timing errors and additive Gaussian noise. Direct computation of the information rate for channels with memory is difficult, even if the input symbols are assumed to be independent and uniformly distributed (i.u.d.). Practical simulation-based methods have been proposed [8], [9], [10] to calculate information rates for finite-state intersymbol interference (ISI) channels. These methods utilize the entropy ergodic theorem [11], [12]. We propose simulation-based bounds on the information rates for ISI channels corrupted by timing errors and additive Gaussian noise.

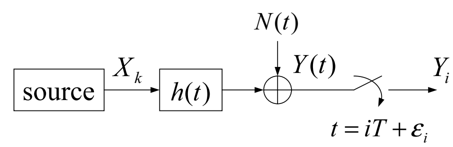

Fig. 1. A simple diagram of the channel and signal model.

II. Source and Channel Model

We consider a simple channel with intersymbol interference (ISI) as shown in Fig. 1. Denote by X_k the binary antipodal channel input symbol ($X_k \in \{-1,+1\}$) at time $k \in \mathbb{Z}$. We only consider the case where the random process $\{X_k\}$ is a *stationary ergodic* Markov chain of finite order ν. The received waveform $Y(t)$ is assumed to be of the following form

$$Y(t) \quad = \quad \sum_k X_k h(t-kT) + N(t), \tag{1}$$

where T is the symbol interval and $N(t)$ is additive Gaussian noise that is independent of the input. We assume that the baseband channel response function $h(t)$ is a *finite support function* that satisfies

$$h(t) = 0 \qquad \text{for } |t| \geq qT. \tag{2}$$

We denote the support interval of $h(t)$ by $(-qT, qT)$. Denote by Y_i the i-th sample at the receiver,

$$\begin{aligned} Y_i \quad &= \quad Y(iT+\mathcal{E}_i) = \sum_{k=-\infty}^{+\infty} X_k \cdot h(iT-kT+\mathcal{E}_i) + N_i \\ &= \sum_{k=i-q+\lceil \frac{\mathcal{E}_i}{T} \rceil}^{i+q+\lfloor \frac{\mathcal{E}_i}{T} \rfloor} X_k \cdot h(iT-kT+\mathcal{E}_i) + N_i, \end{aligned} \tag{3}$$

where $\mathcal{E}_i$ is the timing error. For simplicity, we shall assume that N_i are independent and identically distributed (i.i.d.) Gaussian random variables, with mean 0 and variance σ^2, shortly denoted by $N_i \sim \mathcal{N}(0,\sigma^2)$.

The timing error process $\{\mathcal{E}_i\}$ is independent of the input $\{X_k\}$ and noise $\{N_i\}$. We assume $\mathcal{E}_i$ to be a discrete-time, discrete-valued random variable that can take one of countably many values $\frac{jT}{Q}$, where j is an arbitrary integer and Q is a fixed positive interger. Clearly, Q is the number of quantization levels in each symbol interval T. We further assume that the process $\{\mathcal{E}_i\}$ is *slowly* varying with time, and can be represented by the following random walk process

$$\mathcal{E}_{i+1} \quad = \quad \mathcal{E}_i + \Delta_{i+1}, \tag{4}$$

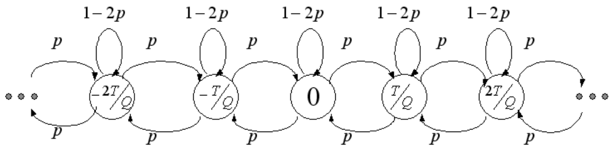

Fig. 2. The state transition diagram for the Markov timing error $\mathcal{E}_i$.

where

$$P(\Delta_i = \delta_i) \quad = \quad \begin{cases} p & \text{if } \delta_i = \pm\frac{T}{Q} \\ 1-2p & \text{if } \delta_i = 0 \end{cases} \quad . \tag{5}$$

The slowly timing-varying assumption is satisfied if $p \ll 1$. The initial value of this random process is $\mathcal{E}_0 = 0$. In practical systems, this is generally achieved by using a preamble in front of each block of data symbols. The increments of the timing error Δ_i are assumed to be i.i.d. and independent of all previous samples Y_j and previous timing errors $\mathcal{E}_j$, for $j < i$. Fig. 2 shows the state transition diagram for $\mathcal{E}_i$. More complicated Markov models than this first-order Markov chain can be adopted without changing the nature of the problem.

III. The finite-state machine channel

In this section, we reformulate the channel model into a finite-state machine channel. We also show that the output sequence $\{Y_i\}$ is a hidden Markov process, that is asymptotically stationary (as $i \to \infty$) and ergodic.

Assume that the i-th sample falls inside the k-th symbol interval, for some $k \in \mathbb{Z}^+$, then

$$iT + \mathcal{E}_i = (k-1)T + \frac{M_i + 1}{Q}T, \tag{6}$$

where $M_i \in \{0, 1, \ldots, Q-1\}$. Note that the sequence $\{M_i\}$ itself is a first order Markov chain. From (4) and (5), we have

$$\begin{aligned} &P_{M_i|M_{i-1}}(m_i|m_{i-1}) \\ &\quad = \begin{cases} p & \text{if } m_i = (m_{i-1} \pm 1) \bmod Q, \\ 1-2p. & \text{if } m_i = m_{i-1}. \end{cases} \end{aligned} \tag{7}$$

The timing instant $iT + \mathcal{E}_i$ is determined by m_i, which is the realization of M_i, as seen in (6). Further, since we assume that the input process $\{X_k\}$ is a Markov chain of finite order ν, it is clear from (3) that the channel output Y_i depends on $\kappa = \max(\nu, 2q)$ input values $X_{i+q+\lfloor\mathcal{E}_i/T\rfloor-\kappa}^{i+q+\lfloor\mathcal{E}_i/T\rfloor}$. Without loss of generality, we will assume that $\nu \le 2q$, so that Y_i only depends on $X_{i-q+\lceil\mathcal{E}_i/T\rceil}^{i+q+\lfloor\mathcal{E}_i/T\rfloor}$. Denote the realizations of these $2q$ channel input symbols by $a_1^{2q}(i) \stackrel{\triangle}{=} x_{i-q+\lceil\mathcal{E}_i/T\rceil}^{i+q+\lfloor\mathcal{E}_i/T\rfloor} \in \{-1,1\}^{2q}$. With this notation, we are ready to define the *channel state* for the i-th channel-output as $s_i = \left(m_i, a_1^{2q}(i)\right)$, where $m_i \in \{0, 1, \ldots, Q-1\}$ and $a_1^{2q}(i) \in \{-1,1\}^{2q}$. We denote by $\mathbb{S}$ the set of all possible states, i.e., $\mathbb{S} = \left\{(m, a_1^{2q}) : m \in \{0, 1, \ldots, Q-1\} \text{ and } a_1^{2q} \in \{-1,1\}^{2q}\right\}$. The number of elements in $\mathbb{S}$ is $|\mathbb{S}| = 2^{2q}Q$. Note that for a specific sample Y_i, it is not important to know the time indices of the input symbols $X_{i-q+\lceil\mathcal{E}_i/T\rceil}^{i+q+\lfloor\mathcal{E}_i/T\rfloor}$, but rather we only need to know the value of the binary vector $a_1^{2q}(i)$.

It is not hard to verify that the channel state sequence S_i, with realizations $s_i = \left(m_i, a_1^{2q}(i)\right)$ for $i \in \{0, 1, \ldots\}$, also forms a first order Markov chain with a finite number of states

$$P(S_i|S_{i-1}, S_{i-2}, \ldots) = P(S_i|S_{i-1}). \tag{8}$$

Given a channel state $s_i = \left(m_i, a_1^{2q}(i)\right)$, we get from (3) that the i-th channel output

$$Y_i(s_i) = \sum_{k=1}^{2q} a_k(i) h\left(\left[q - k + \frac{m_i+1}{Q}\right]T\right) + N_i \tag{9}$$

is a Gaussian random variable $Y_i(s_i) \sim \mathcal{N}(\mu_i, \sigma^2)$, where

$$\mu_i = \sum_{k=1}^{2q} a_k(i) h\left(\left[q - k + \frac{m_i+1}{Q}\right]T\right). \tag{10}$$

Hence, $\{Y_i\}$ is a hidden Markov process with $\{S_i\}$ being the embedded Markov chain.

In order to show that $\{Y_i\}$ is asymptotically stationary and ergodic, we need to examine the properties of the embedded hidden Markov chain $\{S_i\}$. We prove the following lemma.

Lemma 1: The state sequence $\{S_i\}$ is an ergodic finite-state Markov chain that has a unique steady state distribution, and is asymptotically stationary for any given initial state.

Proof: We first show that any two states, $s \in \mathbb{S}$ and $s' \in \mathbb{S}$, *communicate* [13]. Let $s = (m, a_1^{2q})$ and $s' = (m', b_1^{2q})$, where $0 \le m \le m' \le Q-1$, and a_i, b_i are binary symbols. From (7), we can always start from state s and take $m' - m$ transitions to reach a state $s'' = (m', c_1^{2q})$ for some binary vector c_1^{2q}. Similary, we can start from state s'' and go to state s' in at most $2q$ transitions by keeping m' unchanged and sequentially inputting the binary symbols $b_1, b_2, \ldots, b_{2q}$. Thus state s' is *accessible* from state s. Similarly, state s is also accessible from s'. Therefore we know that all the states in $\mathbb{S}$ communicate, i.e., there is only one class of states in $\mathbb{S}$. Since there is only one class of states in the set $\mathbb{S}$ and the Markov chain has finite number of states, we conclude that all the states are *recurrent* [13].

Next we show that each state in $\mathbb{S}$ is *aperiodic*. We can simply consider the state $s = (Q-1, a_1^{2q})$, where $a_1 = a_2 = \cdots = a_{2q} = 1$. Obviously, we can reach state s from state s itself using a single state transition, which shows that the state s is an *aperiodic* state. Because we have only one class of states, all the states in $\mathbb{S}$ are aperiodic [13].

Since every state in $\mathbb{S}$ is both recurrent and aperiodic, we conclude that this finite state Markov chain $\{S_i\}$ is *ergodic* and furthermore, it will converge to a *unique* steady state distribution, irrespective of the initial state [13], i.e., the state sequence is *asymptotically stationary*.

Since the statistical properties of the hidden Markov process are inherited from similar properties of the underlying Markov chain, we therefore conclude that the sampled sequence $\{Y_i\}$ is ergodic and asymptotically stationary. ∎

Using similar arguments as for Lemma 1, we can also prove the following:

Lemma 2: The sequence $\{M_i\}$, as defined in (7), is an ergodic and asymptotically stationary finite-state Markov chain. Furthermore, it has a unique steady state distribution $\pi_M(k) = \frac{1}{Q}$, where $k \in \{0, 1, \ldots, Q-1\}$.

IV. Mutual information rates

It was shown by Dobrushin [7], [5] that Shannon's fundamental theorem is true for the class of memoryless channels with timing errors[1], and the channel capacity is

$$C = \alpha \cdot \sup_{p(X_1^\infty)} \lim_{\ell\to\infty} \frac{1}{\ell} I\left(X_1^\infty; Y_1^\ell\right), \tag{11}$$

where the supremum is taken over all stationary and ergodic processes X_1^∞, and α is the expected number of received symbols *per transmitted symbol*. That is, if we denote by L_n the number of samples within the first n symbol intervals, then

$$\alpha = \lim_{n\to\infty} \frac{E(L_n)}{n}. \tag{12}$$

In this section, we study the information rate of the channel with timing errors as defined in Sec. II, i.e., we wish to evaluate the following term

$$\begin{aligned} I(X,Y) &= \lim_{\ell\to\infty} \frac{\alpha}{\ell} I\left(X_1^\infty, Y_1^\ell\right) \\ &= \lim_{\ell\to\infty} \frac{\alpha}{\ell} \left[H\left(Y_1^\ell\right) - H\left(Y_1^\ell | X_1^\infty\right)\right], \end{aligned} \tag{13}$$

which can be interpreted as the information rate *per transmitted symbol* for a given stationary and ergodic process X_1^∞. We propose upper and lower bounds for the information rate using simulation-based Monte-Carlo methods.

A. Computation of α

Since L_n is defined as the number of samples within the first n symbol intervals, we have

$$L_n T + \mathcal{E}_{L_n} \le nT \tag{14}$$

$$(L_n+1)T + \mathcal{E}_{L_n+1} > nT. \tag{15}$$

From (14) and (15), we have

$$\frac{n-1}{n} - \frac{\sum_{i=1}^{L_n+1} \Delta_i}{nT} < \frac{L_n}{n} \le 1 - \frac{\sum_{i=1}^{L_n} \Delta_i}{nT}. \tag{16}$$

By taking the the expected values of all the terms in inequality (16), and letting $n \to \infty$, we have

$$1 \le \lim_{n\to\infty} \frac{E[L_n]}{n} \le 1. \tag{17}$$

Therefore, from (12) we have $\alpha = 1$.

B. Entropy rate of the channel output sequence $\frac{1}{\ell} H(Y_1^\ell)$

From the Shannon-McMillan-Breiman theorem [14], for discrete stationary and ergodic random process

$$\lim_{\ell\to\infty} -\frac{1}{\ell} \log Pr\left(Y_1^\ell\right) = H(Y) \tag{18}$$

with probability one (almost surely). Similar results have been extended to differential entropy rate by Barron [11] for continuous stationary and ergodic random process. We have shown in Sec. III that the received sequence $\{Y_\ell\}$ is ergodic and asymptotically stationary, thus we can evaluate the entropy rate from a long sample sequence y_1^ℓ.

Recently, methods to compute the entropy rate based on (18) have been independently reported by Arnold and Loeliger [8], Pfister et al. [9], and Sharma and Singh [10]. The methods essentially consist of sampling long input and output sequences and computing $P(Y_1^\ell)$ by the forward recursion of sum-product algorithm [15]. Given a sample sequence (realization) y_1^ℓ of the hidden Markov process $\{Y_i\}$, we have

$$P\left(y_1^\ell\right) = \sum_{s_1^\ell \in \mathbb{S}^\ell} P\left(s_1^\ell, y_1^\ell\right). \tag{19}$$

We define $\mu(s_i)$ to be the *accumulated metric* of the state s_i corresponding to the i-th sample y_i. Using the Markovian property in (8) and (9), we have

$$\begin{aligned} \mu(s_i) &= \sum_{s_1^{i-1} \in \mathbb{S}^{i-1}} P\left(s_1^i, y_1^i\right) \\ &= \sum_{s_{i-1}\in\mathbb{S}} \mu\left(s_{i-1}\right) \cdot P\left(s_i, y_i | s_{i-1}\right). \end{aligned} \tag{20}$$

We can now recursively compute $\mu(s_i)$ using the sum-product algorithm (BCJR) [15], and (19) gives

$$P(y_1^\ell) = \sum_{s_\ell\in\mathbb{S}} \mu(s_\ell). \tag{21}$$

The estimated entropy rate of the sample sequence is then determined by

$$\hat{H}(Y) = \lim_{\ell\to\infty} \frac{1}{\ell} \log \sum_{s_\ell\in\mathbb{S}} \mu(s_\ell). \tag{22}$$

C. Bounding the Conditional Entropy Rate

Next, we wish to bound $H(Y_1^\ell|X_1^\infty)$. Notice that given the input sequence, the uncertainty in the channel output sequence Y_1^ℓ comes not only from the additive Gaussian noise, but also from the timing error sequence $\mathcal{E}_i$. Therefore, we write $H(Y_1^\ell|X_1^\infty)$ as follows:

$$\begin{aligned} &H\left(Y_1^\ell|X_1^\infty\right) \\ &= H\left(Y_1^\ell|X_1^\infty, \mathcal{E}_1^\ell\right) + I\left(Y_1^\ell; \mathcal{E}_1^\ell | X_1^\infty\right) \\ &= H\left(Y_1^\ell|X_1^\infty, \mathcal{E}_1^\ell\right) + H\left(\mathcal{E}_1^\ell|X_1^\infty\right) - H\left(\mathcal{E}_1^\ell|X_1^\infty, Y_1^\ell\right). \end{aligned} \tag{23}$$

Now we analyze the three terms in (23), separately. Since the additive noise N_i is i.i.d. Gaussian, we have

$$H\left(Y_1^\ell|X_1^\infty, \mathcal{E}_1^\ell\right) = \frac{\ell}{2} \log 2\pi e\sigma^2, \tag{24}$$

where σ^2 is the variance of the noise N_i.

From the random walk assumption given by (4) and (5) and the fact that the timing error process is independent of the inputs, we get

$$\begin{aligned} H(\mathcal{E}_1^\ell|X_1^\infty) &= H(\mathcal{E}_1^\ell) \\ &= H(\mathcal{E}_1) + \sum_{k=2}^{\ell} H(\mathcal{E}_k|\mathcal{E}_1^{k-1}) \end{aligned}$$

[1]The channel we are considering has finite memory due to the intersymbol interference. The complete proof in [7] was given for the memoryless channel. However, as it was stated in [7], the memoryless assumption is not essential.

$$= H(\mathcal{E}_1) + \sum_{k=2}^{\ell} H(\mathcal{E}_k|\mathcal{E}_{k-1}). \quad (25)$$

Since $\mathcal{E}_k$ takes only three values given $\mathcal{E}_{k-1}$, we have

$$\begin{aligned} & H(\mathcal{E}_k|\mathcal{E}_{k-1}) \\ &= \sum_{\varepsilon} P(\mathcal{E}_{k-1}=\varepsilon) H(\mathcal{E}_k|\mathcal{E}_{k-1}=\varepsilon) \\ &= \sum_{\varepsilon} P(\mathcal{E}_{k-1}=\varepsilon)\left[2p\log\frac{1}{p} + (1-2p)\log\frac{1}{1-2p}\right] \\ &= 2p\log\frac{1}{p} + (1-2p)\log\frac{1}{1-2p}. \quad (26) \end{aligned}$$

If we also assume $\mathcal{E}_0 = 0$, that is, the system starts from perfect timing, then (25) simplifies to

$$H(\mathcal{E}_1^{\ell}|X_1^{\infty}) = \ell \cdot \left[2p\log\frac{1}{p} + (1-2p)\log\frac{1}{1-2p}\right]. \quad (27)$$

We do not have a method to compute the term $H(\mathcal{E}_1^{\ell}|X_1^{\infty}, Y_1^{\ell})$ in (23), but we can recursively bound it as follows

$$\begin{aligned} & H(\mathcal{E}_1^{\ell}|X_1^{\infty}, Y_1^{\ell}) \\ &= H\left(\mathcal{E}_1^{\ell-1}|X_1^{\infty}, Y_1^{\ell}\right) + H\left(\mathcal{E}_{\ell}|X_1^{\infty}, Y_1^{\ell}, \mathcal{E}_1^{\ell-1}\right) \\ &\le H\left(\mathcal{E}_1^{\ell-1}|X_1^{\infty}, Y_1^{\ell-1}\right) + H\left(\mathcal{E}_{\ell}|X_1^{\infty}, Y_1^{\ell}, \mathcal{E}_1^{\ell-1}\right) \quad (28) \end{aligned}$$

The inequality comes from the fact that conditioning does not increase entropy. Note that the recursion in (28) gives an upper bound on $H(\mathcal{E}_1^{\ell}|X_1^{\infty}, Y_1^{\ell})$, and thus leads to an upper bound to the mutual information $I(X_1^{\infty}; Y_1^{\ell})$. By using the random-walk assumption in (4) and (5) as well as the Bayes rule, we can show (proof provided in Appendix) that

$$H\left(\mathcal{E}_{\ell}|X_1^{\infty}, Y_1^{\ell}, \mathcal{E}_1^{\ell-1}\right) = H\left(\mathcal{E}_{\ell}|X_1^{\infty}, Y_{\ell}, \mathcal{E}_{\ell-1}\right). \quad (29)$$

Thus (28) becomes

$$\begin{aligned} & H\left(\mathcal{E}_1^{\ell}|X_1^{\infty}, Y_1^{\ell}\right) \\ &\le H\left(\mathcal{E}_1^{\ell-1}|X_1^{\infty}, Y_1^{\ell-1}\right) + H\left(\mathcal{E}_{\ell}|X_1^{\infty}, Y_{\ell}, \mathcal{E}_{\ell-1}\right). \quad (30) \end{aligned}$$

The above inequality implies that $H(\mathcal{E}_1^{\ell}|X_1^{\infty}, Y_1^{\ell})$ can be recursively bounded if we can compute $H(\mathcal{E}_{\ell}|X_1^{\infty}, Y_{\ell}, \mathcal{E}_{\ell-1})$. We next prove the following lemma.

Lemma 3: Given a stationary input sequence $\{X_k\}$, and the timing error process $\{\mathcal{E}_k\}$ as defined by (4) and (5), with channel outputs specified by (3), the conditional entropy $H(\mathcal{E}_i|X_1^{\infty}, Y_i, \mathcal{E}_{i-1})$ converges, as $i \to \infty$. Furthermore,

$$\lim_{i\to\infty} H\left(\mathcal{E}_i|X_1^{\infty}, Y_i, \mathcal{E}_{i-1}\right) = H\left(M \,\middle|\, X_1^{2q}, Y, M'\right), \quad (31)$$

where $M' \in \{0, 1, \ldots, Q-1\}$ is an equiprobable random variable, with probability mass function $\pi_M(k) = \frac{1}{Q}$, and

$$P_{M|M'}(m|m') = \begin{cases} p & \text{if } m = (m' \pm 1) \bmod Q \\ 1-2p & \text{if } m = m' \end{cases}, \quad (32)$$

$$Y = \sum_{k=1}^{2q} X_k \cdot h\left(qT - kT + \frac{M+1}{Q}T\right) + N,$$

and $N \sim \mathcal{N}(0, \sigma^2)$ is additive Gaussian noise.

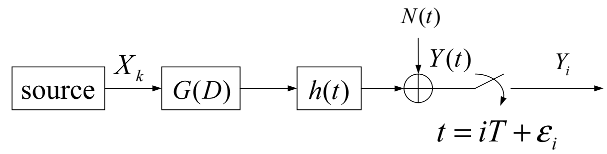

Fig. 3. The block diagram used for simulation. $G(D) = 1 - D^2$.

Proof: First, by using the finite-support assumption in (2) and (3), we have

$$H(\mathcal{E}_i|X_1^{\infty}, Y_i, \mathcal{E}_{i-1}) = H\left(\mathcal{E}_i \,\middle|\, X_{i-q+\lceil \frac{\mathcal{E}_i}{T}\rceil}^{i+q+\lfloor \frac{\mathcal{E}_i}{T}\rfloor}, Y_i, \mathcal{E}_{i-1}\right). \quad (33)$$

Next, we notice that the i-th output Y_i is only determined by the $2q$ binary symbols and the value of M_i as shown in (9), thus we can write

$$H(\mathcal{E}_i|X_1^{\infty}, Y_i, \mathcal{E}_{i-1}) = H\left(M_i \,\middle|\, X_{i-q+\lceil \frac{\mathcal{E}_i}{T}\rceil}^{i+q+\lfloor \frac{\mathcal{E}_i}{T}\rfloor}, Y_i, M_{i-1}\right).$$

Finally, since the input is stationary and by Lemma 2 the states M_i and M_{i-1} converge to a steady state, whose steady-state distribution is determined by (32), we can ignore the index i when $i \to \infty$. This proves Lemma 3. ■

We notice that $H(M|X_1^{2q}, Y, M')$ can be computed either analytically, or numerically by using the law of large numbers and averaging over a large number of simulations. By substituting (24), (27), (30) and (31) into (23), we have

$$\begin{aligned} & \lim_{\ell\to\infty}\left[\frac{1}{\ell}H(Y_1^{\ell}|X_1^{\infty})\right] \ge H_{LB}(Y|X) \\ &\triangleq \frac{1}{2}\log 2\pi e\sigma^2 + \left[2p\log\frac{1}{p} + (1-2p)\log\frac{1}{1-2p}\right] \\ &\quad -H\left(M \,\middle|\, X_1^{2q}, Y, M'\right). \quad (34) \end{aligned}$$

This lower bound on the conditional entropy rate will lead to an upper bound on the mutual information rate when substituted into (13).

To obtain a lower bound on the mutual information rate, we will need an upper bound on the conditional entropy rate. For this we resort to the reduced-state trellis techniques [16], [6]. We lower bound the information rate by keeping only a fixed number of states with the largest accumulated metric and discarding the rest at each trellis section in the forward sum-product iteration.

V. Simulation Results

To assess the proposed information rate bounds in Section IV, we consider the following simulation setup. The data symbols are first passed through the filter $G(D) = 1 - D^2$, as shown by Fig. 3. We assume that $h(t)$ is a truncated $sinc$ function of the form $h(t) = sinc(\frac{t}{T})\left[u(t+T) - u(t-T)\right]$, where $u(t)$ is the unit step function. If there is no timing error, the channel is equivalent to the $PR4$ channel [17]. The timing error injected into the system is a random walk process given by (4) and (5).

Fig. 4 shows the upper and lower bounds on the i.u.d. information rate, i.e., the information rate when the input symbols are independent and uniformly distributed (i.u.d.). The timing

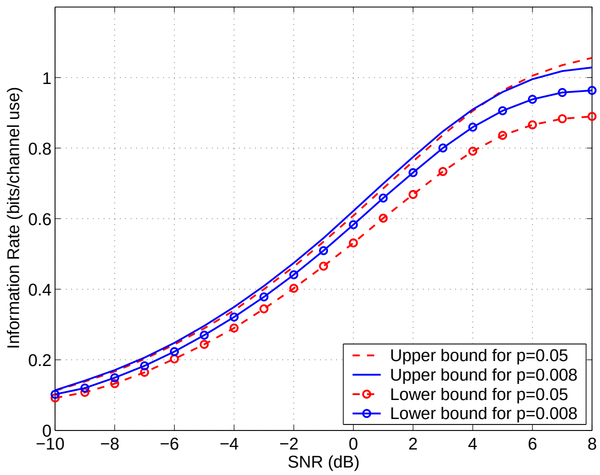

Fig. 4. Upper and lower bounds on the i.u.d. information rate. $G(D) = 1 - D^2$, and $Q = 10$.

error transition probability in (5) is set to $p = 0.05$ and $p = 0.008$, respectively. The number of quantization levels is $Q = 10$. The reduced-state lower bounds [16], [6] are plotted by keeping 40 "surviving" states in each trellis section. We notice that the upper bounds exceed 1 (bits/channel use) at high SNRs, due to the fact that the "almost" noiseless sample Y_i at high SNRs contains more information about the previous timing error $\mathcal{E}_{i-1}$ than at lower SNRs. In the derivation of (28), we dropped the conditioning on Y_ℓ, which leads to a looser bound. Similar arguments can be made to explain the fact that the upper bound for $p = 0.05$ slightly surpasses the upper bound for $p = 0.008$ in the high SNR region.

VI. Conclusions

In this paper, we studied the information rate for the base-band linear filter channels with timing errors. We considered the case where the timing error can be modelled as a discrete-valued Markov chain. Thus the channel is turned into a finite-state machine (FSM) model and the channel outputs are shown to form an ergodic and asymptotically stationary hidden Markov process. We proposed simulation-based upper and lower bounds on the information rates by utilizing the entropy ergodic theorem. The bounds are relatively tight in the low and medium SNR regions.

Appendix

In this appendix, we prove equation (29),

$$H(\mathcal{E}_n|X_1^\infty, Y_1^n, \mathcal{E}_1^{n-1}) = H(\mathcal{E}_n|X_1^\infty, Y_n, \mathcal{E}_{n-1}).$$

Proof: From the Bayes rule

$$\begin{aligned}
&P(\mathcal{E}_n|X_1^\infty, Y_1^n, \mathcal{E}_1^{n-1}) \\
&= \frac{P(X_1^\infty, Y_1^n, \mathcal{E}_1^n)}{P(X_1^\infty, Y_1^n, \mathcal{E}_1^{n-1})} \\
&= \frac{P(X_1^\infty, Y_n, \mathcal{E}_{n-1}^n) \cdot P(Y_1^{n-1}, \mathcal{E}_1^{n-2}|X_1^\infty, Y_n, \mathcal{E}_{n-1}^n)}{P(X_1^\infty, Y_n, \mathcal{E}_{n-1}) \cdot P(Y_1^{n-1}, \mathcal{E}_1^{n-2}|X_1^\infty, Y_n, \mathcal{E}_{n-1})} \\
&= P(\mathcal{E}_n|X_1^\infty, Y_n, \mathcal{E}_{n-1}) \frac{P(\mathcal{E}_1^{n-2}|X_1^\infty, Y_n, \mathcal{E}_{n-1}^n)}{P(\mathcal{E}_1^{n-2}|X_1^\infty, Y_n, \mathcal{E}_{n-1})} \\
&\quad \cdot \frac{P(Y_1^{n-1}|X_1^\infty, Y_n, \mathcal{E}_1^n)}{P(Y_1^{n-1}|X_1^\infty, Y_n, \mathcal{E}_1^{n-1})} \\
&= P(\mathcal{E}_n|X_1^\infty, Y_n, \mathcal{E}_{n-1}) \frac{P(\mathcal{E}_1^{n-2}|X_1^\infty, Y_n, \mathcal{E}_{n-1}^n)}{P(\mathcal{E}_1^{n-2}|X_1^\infty, Y_n, \mathcal{E}_{n-1})}. \qquad (35)
\end{aligned}$$

The last equality is based on the fact that, given the input process X_1^∞ and the timing error sequence $\mathcal{E}_1^{n-1}$, the first $n-1$ channel outputs Y_1^{n-1} are independent of Y_n and $\mathcal{E}_n$. By using the Bayes rule again,

$$\begin{aligned}
&P(\mathcal{E}_1^{n-2}|X_1^\infty, Y_n, \mathcal{E}_{n-1}^n) \\
&= \frac{P(\mathcal{E}_1^{n-2}, \mathcal{E}_n|X_1^\infty, Y_n, \mathcal{E}_{n-1})}{P(\mathcal{E}_n|X_1^\infty, Y_n, \mathcal{E}_{n-1})} \\
&= P(\mathcal{E}_1^{n-2}|X_1^\infty, Y_n, \mathcal{E}_{n-1}) \frac{P(\mathcal{E}_n|X_1^\infty, Y_n, \mathcal{E}_1^{n-1})}{P(\mathcal{E}_n|X_1^\infty, Y_n, \mathcal{E}_{n-1})} \\
&= P(\mathcal{E}_1^{n-2}|X_1^\infty, Y_n, \mathcal{E}_{n-1}). \qquad (36)
\end{aligned}$$

The last equality is based on the fact that $\mathcal{E}_k$ is a first order Markov chain as given by (4). Substituting (36) into (35), we obtain $H(\mathcal{E}_n|X_1^\infty, Y_1^n, \mathcal{E}_1^{n-1}) = H(\mathcal{E}_n|X_1^\infty, Y_n, \mathcal{E}_{n-1})$. ■

References

[1] R. G. Gallager, "Sequential decoding for binary channels with noise and synchronization errors," 1961, unpublished Lincoln Lab. report 25 G-2.

[2] J. D. Ullman, "On the capabilities of codes to correct synchronization errors," *IEEE Trans. Information Theory*, vol. 13, no. 5, pp. 95–105, January 1967.

[3] S. N. Diggavi and M. Grossglauser, "On transmission over deletion channels," *Allerton Conference 2001, Monticello, Illinois*, October 2001.

[4] E. Drinea and M. Mitzenmacher, "On lower bounds for the capacity of deletion channels," in *Proc. of the IEEE International Symposium on Information Theory*, Chicago, Illinois, June 2004, p. 227.

[5] R. L. Dobrushin, "The computation on a computer of the channel capacity of a line with symbols drop-out," *Problems of Information Transmission*, vol. 4, no. 3, pp. 92–95, 1968.

[6] A. Kavcic and R. Motwani, "Insertion/deletion channels: Reduced-state lower bounds on channel capacities," in *Proc. IEEE International Sympsium on Information Theory 2004*, Chicago, Illinois, June 2004, p. 229.

[7] R. L. Dobrushin, "Shannon's theorems for channels with synchronization errors," *Problems of Information Transmission*, vol. 3, no. 4, pp. 18–36, 1967.

[8] D. Arnold and H.-A. Loeliger, "On the information rate of binary-input channels with memory," in *Proc. IEEE Int. Conf. Comm. 2001*, Helsinki, Finland, June 2001, pp. 2692–2695.

[9] H. D. Pfister, J. B. Soriage, and P. H. Siegel, "On the achievable information rates of finite state ISI channels," in *Proc. IEEE Globecom 2001*, San Antonio, TX, November 2001, pp. 2992–2996.

[10] V. Sharma and S. K. Singh, "Entropy and channel capacity in the regenerative setup with applications to Markov channels," in *Proc. IEEE ISIT 2001*, Washington DC, USA, June 2001, p. 283.

[11] A. R. Barron, "The strong ergodic theorem for densities: Generalized Shannon-McMillan-Breiman theorem," *The Annals of Probability*, vol. 13, no. 4, pp. 1292–1303, 1985.

[12] L. Breiman, "The individual ergodic theorem of information theory," *The Annals of Math. Statist.*, vol. 28, pp. 809–811, 1957.

[13] R. G. Gallager, *Discrete Stochastic Processes*. Massachusetts, USA: Kluwer Academic Publishers, 1996.

[14] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. USA: John Wiley & Sons, 1991.

[15] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Information Theory*, vol. 20, pp. 284–287, March 1974.

[16] D. Arnold, A. Kavcic, H.-A. Loeliger, P. Vontobel, and W. Zeng, "Simulation based computation of information rates: Upper and lower bounds by reduced-state techniques," in *Proc. IEEE International Symposium on Information Theory 2003*, Yokohama, Japan, June 2003, p. 119.

[17] P. Kabal and S. Pasupathy, "Partial-response signaling," *IEEE Trans. Communications*, vol. 23, no. 9, pp. 921–934, September 1975.