Transactions Letters

On the Low-Rate Shannon Limit for Binary Intersymbol Interference Channels

Joseph B. Soriaga, *Student Member, IEEE*, Henry D. Pfister, *Student Member, IEEE*, and Paul H. Siegel, *Fellow, IEEE*

***Abstract*—For a discrete-time, binary-input Gaussian channel with finite intersymbol interference, we prove that reliable communication can be achieved if and only if $E_b/N_0 > \log 2/G_{\rm opt}$, for some constant $G_{\rm opt}$ that depends on the channel. To determine this constant, we consider the finite-state machine which represents the output sequences of the channel filter when driven by binary inputs.We then define $G_{\rm opt}$ as the maximum output power achieved by a simple cycle in this graph, and show that no other cycle or asymptotically long sequence can achieve an output power greater than this.We provide examples where the binary input constraint leads to a suboptimality, and other cases where binary signaling is just as effective as real signaling at very low signal-to-noise ratios.**

***Index Terms*—Information rates, intersymbol interference (ISI), magnetic recording, modulation coding.**

I. Introduction

FOR the discrete-time, additive white Gaussian noise (AWGN) channel with noise variance σ^2 and a unit power input constraint, Shannon proved that reliable communication cannot be achieved unless the code rate R is less than

$$C_{\rm AWGN}\left(\frac{1}{\sigma^2}\right) = \frac{1}{2}\log_2\left(1+\frac{1}{\sigma^2}\right). \quad (1)$$

Now, if we consider the signal-to-noise ratio (SNR) *per information bit*, defined as $E_b/N_0 = 1/(2R\sigma^2)$, then this condition requires that E_b/N_0 be greater than

$$\gamma_{\rm AWGN}(\sigma^2) = \frac{1}{2C_{\rm AWGN}\left(\frac{1}{\sigma^2}\right)\sigma^2}. \quad (2)$$

By letting the noise power $\sigma^2 \to \infty$, we find that $C_{\rm AWGN} \to 0$, but that $\gamma_{\rm AWGN}(\sigma^2)$ decreases monotonically to the limit $\log 2 \approx -1.59$ dB. Hence, we refer to this as the *low-rate Shannon limit*, because reliable communication at *any* nonzero rate is possible if and only if $E_b/N_0 > \log 2$.[1] Interestingly, this limit can even be achieved with binary signaling (e.g., see McEliece [3, Problem 4.14]).

Paper approved by H. Leib, the Editor for Communication and Information Theory of the IEEE Communications Society. Manuscript received 10/7/02. This work was supported in part by the University of California, Office of the President, in cooperation with Marvell Semiconductor, Inc., under MICRO Grant 00-090, and in part by the Center for Magnetic Recording Research, University of California, San Diego. This paper was presented in part at the Workshop on Information, Coding, and Mathematics in honor of Robert J. McEliece, Pasadena, CA, May 2002.

J. B. Soriaga and P. H. Siegel are with the Department of Electrical and Computer Engineering, University of California, San Diego, La Jolla, CA 92093-0407 USA (email: jsoriaga@ucsd.edu; psiegel@ucsd.edu).

H. D. Pfister was with the Department of Electrical and Computer Engineering, University of California, San Diego, La Jolla, CA 92093-0407 USA. He is now with Qualcomm, Inc., San Diego, CA 92121 USA (email: hpfister@qualcomm.com).

Digital Object Identifier 10.1109/TCOMM.2003.820724

[1]Recall that an identical result holds for the AWGN waveform channel, in the limit of infinite signal bandwidth [1, eq. 8.2.10]. The low-rate Shannon limit is also related to the *capacity per unit cost* defined by Verdú [2].

For the discrete-time, finite intersymbol interference (ISI) channel with AWGN, described by the equation

$$r_k = \sum_{i=0}^{\nu} h_i x_{k-i} + n_k$$

where x_k is the input, r_k the noisy output, $\{h_0, \ldots, h_\nu\}$ the channel impulse response (also denoted by $h(D) = \sum_{i=0}^{\nu} h_i D^i$), and n_k the AWGN with variance σ^2, it is well known that the capacity can be achieved by "waterfilling," i.e., by distributing the input power at frequencies where the response is highest [1], [4]. Accordingly, (as we will also show), if $G_{\max}$ is the peak power gain (or minimum loss) of the filter, then one should expect the low-rate Shannon limit to be $\log 2/G_{\max}$. Naturally, one is then curious whether this limit changes when a binary input constraint is introduced, i.e., when $x_k \in \{\pm 1\}$. Although tighter bounds on the channel capacity in this case have recently been discovered [5]–[8], this particular question has remained unanswered.

In this letter, we prove that the low-rate Shannon limit for binary-input ISI channels equals $\log 2/G_{\rm opt}$, where $G_{\rm opt}$ is a constant depending on the channel impulse response. More specifically, as discussed in Section II, if a finite-state machine is used to represent the output sequences of the channel filter when driven with binary inputs, then $G_{\rm opt}$ is the maximum power achieved by a simple cycle. It then follows that no cycle or asymptotically long sequence can have output power greater than $G_{\rm opt}$. In Section III, we use these results to conclude that the low-rate Shannon limit is at least $\log 2/G_{\rm opt}$. To achieve this value, we then introduce a concatenated coding scheme consisting of a random outer code and an inner modulation code based on a simple cycle with output power $G_{\rm opt}$. Finally, we note how these arguments can be extended to continuous-alphabet, power-constrained ISI channels.

II. Output Power of State Sequences

Let us represent the input and channel filter output sequences with a finite-state machine, where the current state $s_i = (x_{i-\nu}, \ldots, x_{i-1})$. Note that the *channel filter output*, which we denote y_k, is different from the channel output $r_k = y_k + n_k$. Now, we can label each edge (s_{i-1}, s_i) in the finite-state machine with an input $x(s_{i-1}, s_i)$, a channel filter output $y(s_{i-1}, s_i)$, or functions thereof. For a state sequence $\mathbf{s}$, the length $l(\mathbf{s})$ equals the number of edges, and the output power is defined as

$$G(\mathbf{s}) = \frac{1}{l(\mathbf{s})} \sum_{i=1}^{l(\mathbf{s})} |y(s_{i-1}, s_i)|^2.$$

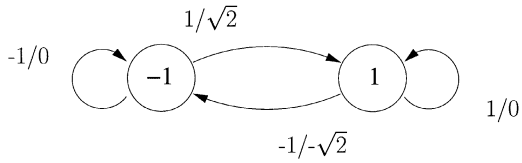

Fig. 1. State graph for the dicode channel $h(D) = (1/\sqrt{2})(1 - D)$, with labels for input and channel filter output. The only optimal simple cycle, starting from state -1, has input sequence $(1, -1)$ and $G_{\rm opt} = 2$.

As an example, Fig. 1 shows the state graph for the dicode channel, with $h(D) = (1/\sqrt{2})(1 - D)$.

Definition 1: A simple cycle, i.e., a cycle in which all edges are distinct, is *optimal* if it achieves the maximum power over all simple cycles.

In graph theory, cycles which achieve this maximum power are also known as *maximum mean-weight cycles*, and can be found using dynamic programming techniques, e.g., Karp's algorithm [9]. The next two lemmas, which also arise in the performance analysis of digital systems (e.g., see [10]), will demonstrate the importance of optimal simple cycles.

Lemma 1: Let $G_{\rm opt}$ be the average power of an optimal simple cycle. For any finite-length cycle $\mathbf{c}$, $G(\mathbf{c}) \leq G_{\rm opt}$.

Proof: Consider the case when $\mathbf{c}$ is not simple. Specifically, for some i and j, $0 < i < j < L = l(\mathbf{c})$, and some state a, we have $c_i = c_j = a$. We can then divide $\mathbf{c}$ into three segments, where the middle segment $\mathbf{u}_1$ is a cycle from time index i to j, i.e., $\mathbf{u}_1 = (c_i, \ldots, c_j)$. The remaining two segments can be connected to form another cycle, $\mathbf{u}_2 = (c_0, \ldots, c_i, c_{j+1}, \ldots, c_L)$.

The output power of $\mathbf{c}$ can now be described in terms of the output powers of $\mathbf{u}_1$ and $\mathbf{u}_2$

$$G(\mathbf{c}) = \frac{l(\mathbf{u}_1)}{l(\mathbf{c})} G(\mathbf{u}_1) + \frac{l(\mathbf{u}_2)}{l(\mathbf{c})} G(\mathbf{u}_2). \tag{3}$$

Notice that $G(\mathbf{c})$ is a weighted average between $G(\mathbf{u}_1)$ and $G(\mathbf{u}_2)$, because $l(\mathbf{u}_1) + l(\mathbf{u}_2) = l(\mathbf{c})$. Consequently, either $G(\mathbf{u}_1) > G(\mathbf{c})$ or $G(\mathbf{u}_2) > G(\mathbf{c})$ or $G(\mathbf{u}_1) = G(\mathbf{u}_2) = G(\mathbf{c})$.

So let us set $\mathbf{c}'$ to the derived cycle with larger output power, which is at least as large as $G(\mathbf{c})$. If we continue this decomposition on $\mathbf{c}'$, etc., always choosing the derived cycle with larger output power, we will eventually obtain a simple cycle $\mathbf{c}_s$ with $G(\mathbf{c}) \leq G(\mathbf{c}') \leq G(\mathbf{c}_s) \leq G_{\rm opt}$. □

Lemma 2: For any $\epsilon > 0$, there exists an n_0 such that, $G(\mathbf{s}) < G_{\rm opt} + \epsilon$ for any sequence $\mathbf{s}$ of length $n > n_0$.

Proof: We can always concatenate another segment $\mathbf{u}$ to the end of $\mathbf{s}$, so that the combined sequence becomes a cycle (this is also referred to as "terminating"). Since the state is determined by the last ν inputs, we can choose this $\mathbf{u}$ such that $l(\mathbf{u}) \leq \nu$. From Lemma 1, we know that this newly formed cycle cannot have a larger average power than $G_{\rm opt}$. Using an expansion similar to (3) we can express this as

$$\frac{l(\mathbf{u})}{l(\mathbf{u}) + l(\mathbf{s})} G(\mathbf{u}) + \frac{l(\mathbf{s})}{l(\mathbf{u}) + l(\mathbf{s})} G(\mathbf{s}) \leq G_{\rm opt}.$$

Rearranging terms

$$\begin{aligned} l(\mathbf{s})G(\mathbf{s}) &\leq \left(l(\mathbf{u}) + l(\mathbf{s})\right) G_{\rm opt} - l(\mathbf{u})G(\mathbf{u}) \\ &= l(\mathbf{s})G_{\rm opt} + l(\mathbf{u}) \left(G_{\rm opt} - G(\mathbf{u})\right) \\ &\leq l(\mathbf{s})G_{\rm opt} + \nu(G_{\rm opt} - G_{\rm min}). \end{aligned}$$

TABLE I
OPTIMAL SIMPLE CYCLES FOR SOME ISI CHANNELS

Channel	G_{max}	G_{opt}	Gap (dB)	Input
PR4	2	2	0	+,+,-,-
EPR4	64/27	2	0.75	+,+,-,- +,+,+,-,- +,+,+,-,-,-
E^{2}PR4	27/10	24/10	0.51	+,+,+,-,-,-

PR4: $h(D) = \frac{1}{\sqrt{2}}(1 - D^2)$
EPR4: $h(D) = \frac{1}{2}(1 - D)(1 + D)^2$
E^{2}PR4: $h(D) = \frac{1}{\sqrt{10}}(1 - D)(1 + D)^3$

In the last inequality, $G_{\rm min}$ is the smallest output power among all edges. Dividing through by $l(\mathbf{s})$ gives

$$G(\mathbf{s}) \leq G_{\rm opt} + \frac{\nu}{l(\mathbf{s})}(G_{\rm opt} - G_{\rm min}).$$

The rightmost term is always positive, but it can be made arbitrarily small by making $\mathbf{s}$ sufficiently long. □

III. Low-Rate Shannon Limit

A. Binary Input ISI Channels

We now state the main result of this paper.

Theorem 1: For a binary-input, finite ISI channel, consider the finite-state machine that represents output sequences from the channel filter. Reliable communication can be achieved if and only if $E_b/N_0 > \log 2/G_{\rm opt}$, where $G_{\rm opt}$ is the maximum output power over all simple cycles in this state graph.

Before we proceed with the proof, let us consider some examples. For the dicode channel with $G_{\rm max} = 2$ (Fig. 1), an optimal simple cycle has $G_{\rm opt} = G_{\rm max}$, and inputs $(1, -1)$. However, Table I shows some channels with $G_{\rm opt} < G_{\rm max}$, and hence binary signaling can impose a penalty. Also, unlike PR4 and E^{2}PR4, EPR4 has several optimal simple cycles.

Proof: Since our channel is an indecomposable finite-state channel,[2] we know that its capacity $C_{\rm BISI}(1/\sigma^2)$ exists, and that reliable communication is possible if and only if $R < C_{\rm BISI}(1/\sigma^2)$ [1, p. 178]. Of course, this condition is the same as requiring that E_b/N_0 be greater than

$$\gamma_{\rm BISI}(\sigma^2) = \frac{1}{2C_{\rm BISI}\left(\frac{1}{\sigma^2}\right)\sigma^2}.$$

Therefore, we only need to prove that

$$\inf_{\sigma^2 > 0} \gamma_{\rm BISI}(\sigma^2) = \frac{\log 2}{G_{\rm opt}} \tag{4}$$

We do this by showing that $\inf \gamma_{\rm BISI}(\sigma^2) \geq \log 2/G_{\rm opt}$, and that $\overline{\lim}_{\sigma^2 \to \infty} \gamma_{\rm BISI}(\sigma^2) \leq \log 2/G_{\rm opt}$.

(Lower Bound) As a consequence of Lemma 2, we can never make our codelength arbitrarily long while maintaining an output power greater than $G_{\rm opt}$. If we consider the channel filter as part of the source, then we equivalently have an AWGN channel with power constraint $P \leq G_{\rm opt}$, and an additional input constraint. Thus, $C_{\rm BISI}(1/\sigma^2) \leq C_{\rm AWGN}(G_{\rm opt}/\sigma^2)$.

[2]Although Gallager only addresses channels with discrete alphabets, his results can be extended to our case.

This also implies that

$$\inf_{\sigma^2>0} \gamma_{\rm BISI}(\sigma^2) \geq \inf_{\sigma^2>0} \frac{1}{2\sigma^2 C_{\rm AWGN}\left(\frac{G_{\rm opt}}{\sigma^2}\right)} = \inf_{\sigma^2>0} \frac{1}{G_{\rm opt}} \gamma_{\rm AWGN}\left(\frac{\sigma^2}{G_{\rm opt}}\right) \qquad (5)$$

which equals $\log 2/G_{\rm opt}$, because $\gamma_{\rm AWGN}(\sigma^2)$ decreases monotonically to $\log 2$.

(Upper Bound) Let us now consider a concatenated coding system which achieves $E_b/N_0 > \log 2/G_{\rm opt}$. First, let $a_k \in \{\pm 1\}$ denote the data bit to be modulated, and $\mathbf{c}_{\rm opt}$ denote an optimal simple cycle. It has a starting state $s_{\rm opt}$, input sequence $\mathbf{x}_{\rm opt} = (x_{{\rm opt},1}, \ldots, x_{{\rm opt},L})$, and output sequence $\mathbf{y}_{\rm opt}$. Also, let $\mathbf{u} = (u_1, \ldots, u_\nu)$ be a length ν input sequence which puts the channel in state $s_{\rm opt}$.

The inner modulation code is then defined as follows. To transmit a_k, we send

$$a_k \mathbf{x} = a_k \cdot (u_1, \ldots, u_\nu, x_{{\rm opt},1}, \ldots, x_{{\rm opt},L}). \qquad (6)$$

That is, we force the channel into state $s_{\rm opt}$ (or $-s_{\rm opt}$), and then transmit $a_k \mathbf{x}_{\rm opt}$. This results in the channel output, $\mathbf{r}_k = a_k \mathbf{y}_k + \mathbf{n}_k$, where the components of $\mathbf{n}_k$ are AWGN terms with variance σ^2. Also, the latter L symbols of $\mathbf{y}_k$ are always equal to $\mathbf{y}_{\rm opt}$. The first ν symbols, however, depend on a_{k-1}, which determines the channel state prior to sending (6).

To demodulate, we use the vector,

$$\mathbf{y}^* = \frac{1}{\sqrt{LG_{\rm opt}}} (\underbrace{0, \ldots, 0}_{\nu \text{ zeros}}, y_{{\rm opt},1}, \ldots, y_{{\rm opt},L}) \qquad (7)$$

with the received sequence to form $r_k = \mathbf{r}_k (\mathbf{y}^*)^T = \sqrt{LG_{\rm opt}} a_k + n'_k$. As a result, n'_k is an AWGN noise term with variance σ^2. Notice that we discarded information about the forcing sequence $\mathbf{u}$.

Clearly, this inner modulation scheme results in a binary-input AWGN channel with capacity $C_b(LG_{\rm opt}/\sigma^2)$. This implies that the overall capacity of the *concatenated system* is $(1/(L+\nu)) C_b(LG_{\rm opt}/\sigma^2)$, and thus, we can achieve reliable communication at any E_b/N_0 greater than

$$\lim_{\sigma^2 \to \infty} \frac{L+\nu}{2C_b\left(\frac{LG_{\rm opt}}{\sigma^2}\right)\sigma^2} = \lim_{\sigma^2 \to \infty} \frac{L+\nu}{LG_{\rm opt}} \gamma_b \left(\frac{\sigma^2}{LG_{\rm opt}}\right) \qquad (8)$$

where $\gamma_b(\sigma^2) = 1/(2C_b(1/\sigma^2)\sigma^2)$. But, according to McEliece [3, Problem 4.14], the low-rate Shannon limit of the AWGN channel can be achieved with binary signaling, i.e., $\gamma_b(\sigma^2) \to \log 2$ as $\sigma^2 \to \infty$. Consequently, the limit in (8) equals $(1 + \nu/L) \log 2/G_{\rm opt}$.

Furthermore, if we replace $\mathbf{x}_{\rm opt}$ in (6) with m repetitions of $\mathbf{x}_{\rm opt}$, and correspondingly extend $\mathbf{y}^*$ in (7), then our concatenated system can achieve reliable communication for any $E_b/N_0 > (1 + \nu/mL) \log 2/G_{\rm opt}$.

Of course, the capacity of the concatenated system lower bounds $C_{\rm BISI}(1/\sigma^2)$, by the data processing theorem [1, p. 80]. Hence, we can conclude that, for any $m > 0$

$$\overline{\lim_{\sigma^2 \to \infty}} \gamma_{\rm BISI}(\sigma^2) \leq \left(1 + \frac{\nu}{mL}\right) \frac{\log 2}{G_{\rm opt}}.$$

If we combine this with (5), then the conclusion (4) follows when $m \to \infty$. Of course, this also implies that $\lim_{\sigma^2 \to \infty} \gamma_{\rm BISI}(\sigma^2) = \log 2/G_{\rm opt}$. □

B. Extensions to Larger Alphabets

Theorem 1 can easily be extended to finite input alphabets of the form $\mathbb{A}_M = \{-a_M, \ldots, -a_1, a_1, \ldots, a_M\}$, because symmetric alphabets allow us to construct an inner (binary) modulation system which can achieve $\log 2/G_{\rm opt}$. Of course, $G_{\rm opt}$ now corresponds to the finite-state machine for a channel filter driven by symbols in $\mathbb{A}_M$.

In the case of real-valued signaling with unit power, the output power of the channel filter can never exceed $G_{\max}$. Hence, by reasoning similar to that in the proof of Theorem 1, we find that reliable communication on ISI channels always requires $E_b/N_0 \geq \log 2/G_{\max}$. As for the inner modulation code, rather than using long truncations of a periodic binary sequence, we can instead truncate $\sqrt{2}\cos(\omega_{\max} n)$, where $\omega_{\max} \in [-\pi, \pi]$ is a peak frequency of the channel filter. Since this sequence asymptotically realizes an output power of $G_{\max}$, our concatenated system will achieve reliable communication for any $E_b/N_0 > \log 2/G_{\max}$.

IV. Conclusion

We prove that reliable communication on a discrete-time, binary-input finite ISI channel can be achieved if and only if $E_b/N_0 > \log 2/G_{\rm opt}$, where $G_{\rm opt}$ is the maximum output power of a simple cycle in the finite-state machine that represents the channel filter output sequences. We begin by showing that no cycle and no asymptotically long sequence can have an output power greater than $G_{\rm opt}$, and from this result we prove that the low-rate Shannon limit must be at least $\log 2/G_{\rm opt}$. We then provide a semiconstructive method to achieve the limit, using a modulator/demodulator based on simple cycles with output power $G_{\rm opt}$. We show by example that binary signaling may or may not incur a penalty. Finally, we explain how the proof may be extended to continuous-alphabet, power-constrained ISI channels, for which the low-rate Shannon limit is $\log 2/G_{\max}$.

Acknowledgment

The authors would like to thank the anonymous reviewers for pointing out [2].

References

[1] R. G. Gallager, *Information Theory and Reliable Communication*. New York, USA: Wiley, 1968.
[2] S. Verdú, "On the channel capacity per unit cost," *IEEE Trans. Inform. Theory*, vol. 36, pp. 1361–1391, Sept. 1990.
[3] R. J. McEliece, *The Theory of Information and Coding (in Encyclopedia of Mathematics)*. Reading, MA: Addison-Wesley, 1977.
[4] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
[5] D. Arnold and H. Loeliger, "On the information rate of binary-input channels with memory," in *Proc. IEEE Int. Conf. Communications*, Helsinki, Finland, June 2001, pp. 2692–2695.
[6] H. D. Pfister, J. B. Soriaga, and P. H. Siegel, "On the achievable information rates of finite state ISI channels," in *Proc. IEEE Global Telecommunications Conf.*, San Antonio, TX, Nov. 2001, pp. 2992–2996.
[7] A. Kavčić, "On the capacity of Markov sources over noisy channels," in *Proc. IEEE Global Telecommunications Conf.*, San Antonio, TX, Nov. 2001, pp. 2997–3001.
[8] P. O. Vontobel and D. M. Arnold, "An upper bound on the capacity of channels with memory and constraint input," in *Proc. IEEE Information Theory Workshop*, Cairns, Australia, Sept. 2001, pp. 147–149.
[9] R. M. Karp, "A characterization of the minimum cycle mean in a digraph," *Discr. Math.*, vol. 23, pp. 309–311, 1978.
[10] A. Mathur, A. Dasdan, and R. K. Gupta, "Rate analysis of embedded systems," *ACM Trans. Design Automat. Electron. Syst. 3*, vol. 3, 1998.