# Matched Information Rate Codes for Binary ISI channels[1]

Xiao Ma, Nedeljko Varnica and Aleksandar Kavčić
DEAS/Harvard University, Cambridge, MA 02138
{xiaoma,varnica,kavcic}@hrl.harvard.edu

*Abstract* — **We propose a coding/decoding strategy to approach the channel capacities for binary intersymbol interference (ISI) channels. The proposed codes are serially concatenated codes: inner *matched information rate codes* and outer irregular low-density parity-check (LDPC) codes. The whole system is iteratively decodable.**

## I. SUMMARY

Binary ISI channel models are appropriate models for information storage systems [1]. The behavior of such a channel can be represented by a trellis [2]. Here we describe a more general time-invariant trellis model. At time $t \geq 0$, the trellis has $S$ states, which are indexed by $\mathcal{S} \stackrel{\text{def}}{=} \{0, 1, \cdots, S-1\}$. Exactly $2^k$ branches emanate from each state. A branch at time $t > 0$ is determined by a four-tuple $b_t = (s_{t-1}, x_t, y_t, s_t)$. Here, the two symbols $s_{t-1}$ and $s_t$ denote the two states connected by this branch; the symbol $x_t \in \{0,1\}^k$ denotes a binary input vector; and the symbol $y_t \in R^n$ denotes a real-valued output vector. We assume that $y_t$ and $s_t$ are determined uniquely by $s_{t-1}$ and $x_t$. We assume that the trellis represents an indecomposable finite state machine. Throughout this paper, we assume that the initial state $s_0$ is given. The trellis can be considered as an encoder (with rate $k/n$) which transforms a (binary) sequence $x_1^N$ into a (real-valued) sequence $y_1^N$. Assume that $y_1^N$ is transmitted through a known memoryless channel and noisy observation $z_1^N$ is received. Denote by $q_{si}$ the conditional probability $P_{X_t|S_{t-1}}(i|s)$, where $i$ is the decimal representation of $x_t$. Denote by $Q = (q_{si})$ the collection of these conditional probabilities and assume $q_{si} > 0$.

The (average) mutual information [3] between $X_1^N$ and $Z_1^N$ is a function of $Q$, $I_N(Q) \stackrel{\text{def}}{=} I(X_1^N; Z_1^N)$. We can prove that

*Theorem 1:* $I(Q) \stackrel{\text{def}}{=} \lim_{N\to\infty} \frac{1}{nN} I_N(Q)$ exists.

*Theorem 2:* $I(Q)$ is achievable.

If all entries of $Q$ are equal to $2^{-k}$, $I(Q)$ is called i.u.d. (independent uniformly distributed) information rate.

*Corollary:* The i.u.d. information rate can be achieved by a linear (coset) code.

For a given $Q$, the rate $I(Q)$ can be estimated [4][5]. Furthermore, an iterative algorithm [6] has been proposed to *maximize* $I(Q)$ over all valid matrices $Q$. The challenging problem is to design practical codes to achieve $I(Q)$ and hence to approach the channel capacity. A general method to design serially concatenated codes is proposed.

The inner codes are trellis codes that transform typical i.u.d. sequences into sequences that match the distribution matrix $Q$. The criterion is to *minimize* $|I_S - I(Q)|$ over all trellis codes under certain complexity constraints. Here we use $I_S$ to denote the i.u.d. information rate of the superchannel (the concatenation: trellis code + channel). We call such a code *a matched information rate (MIR) code*. The construction procedure is 1) split the target rate $r$ as $r = r_{in} r_{out}$; 2) determine the inner trellis code state space; 3) determine the frequency for each possible noiseless output vector; 4) determine the output vector for each branch and determine its ending state; 5) determine the input vector for each branch. We have constructed design rules for completing all 5 steps of the procedure. Using these rules, for the dicode channel and the target rate $r = 1/2$, we have constructed a 10-state MIR code with $r_{in} = k/n = 2/3$. The i.u.d. information rate $I_S$ is plotted in Figure 1.

According to the Corollary, there must exist at least one linear (coset) code that approaches $I_S$ and hence $I(Q)$. For practical reasons, we utilize low-density parity-check (LDPC) codes as outer codes, which can be optimized by a method similar to that of [7]. From Figure 1, we see that the threshold of the optimized LDPC code surpasses the i.u.d. rate of the channel.

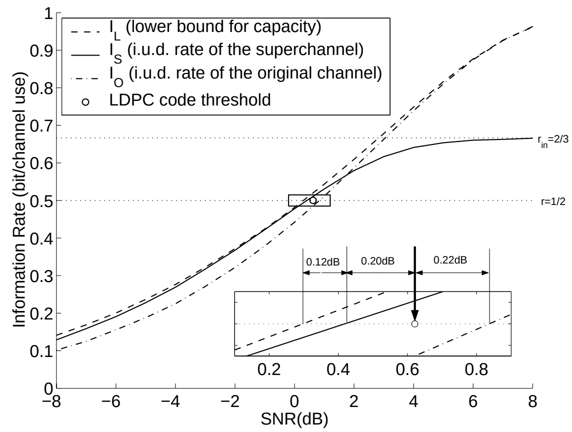


Figure 1: Asymptotic performance of the MIR code for the $1 - D$ channel. These three curves $I_L$, $I_S$, and $I_O$ cross the rate 0.5 bits/channel-use at 0.30 dB, 0.42 dB and 0.84 dB, respectively. We believe that the 0.12 dB loss between the $I_L$ and $I_S$ can be recovered by using a more complicated MIR code. Combined with outer linear (coset) codes, the asymptotic coding gain over $I_O$ is 0.42 dB. The point 'o' shown in this figure is the threshold location of the optimized outer irregular LDPC code.

[1] This work was supported by the National Science Foundation under Grant No. CCR-0118701.

## REFERENCES

[1] K. A. S. Immink, P. H. Siegel and J. K. Wolf, "Codes for digital recorders," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2260-2299, October 1998.

[2] G. D. Forney, "Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference," *IEEE Trans. Inform. Theory*, vol. 18, pp. 363-378, March 1972.

[3] R. G. Gallager, *Information Theory and Reliable Communication*. New York: John Wiley and Sons, Inc., 1968.

[4] D. Arnold and H.-A. Loeliger, "On the information rate of binary-input channels with memory," in *Proceedings IEEE International Conference on Communications 2001*, (Helsinki, Finland), June 2001.

[5] H. D. Pfister, J. B. Soriaga, and P. H. Siegel, "On the achievable information rates of finite state ISI channels", in *Proceedings IEEE Global Communication Conference 2001*, (San Antonio, Texas), November 2001.

[6] A. Kavčić, "On the capacity of Markov sources over noisy channels," in *Proceedings IEEE Global Communication Conference 2001*, (San Antonio, Texas), November 2001.

[7] N. Varnica and A. Kavčić, "Optimized LDPC codes for partial response channels," to appear in *ISIT'02*.