Capacity-Achieving Codes for Channels with Memory with Maximum-Likelihood Decoding

UNIVERSITY OF MICHIGAN

Jung Hyun Bae and Achilleas Anastasopoulos

Electrical Engineering and Computer Science Department, University of Michigan, Ann Arbor, MI

Abstract

Codes on sparse graphs have been shown to achieve remarkable performance in point-to-point channels with low decoding complexity. Most of the results in this area are based on experimental evidence and/or approximate analysis. The question of whether codes on sparse graphs can achieve the capacity of noisy channels with iterative decoding is still open, and has only been conclusively and positively answered for the binary erasure channel. On the other hand, codes on sparse graphs have been proven to achieve the capacity of memoryless, binary-input, output-symmetric channels with finite graphical complexity per information bit when maximum likelihood (ML) decoding is performed. In this presentation, we consider transmission over finite-state channels (FSCs). We will construct a simple quantization scheme that when applied to ensembles of codes on sparse graphs induces a Markov distribution on the transmitted sequence. By deriving average error probability bounds for these quantized code ensembles, we prove that they can achieve the information rates corresponding to the induced Markov distribution, and thus approach the FSC capacity.

Finite-state channel and its capacity

Let $\{S_n\}_{n=1}^{\infty}$ with $s_n \in \mathscr{S} = \{1,2,\ldots,K\}$ be a sequence representing the channel states st time n. It is assumed that the state space $\mathscr{S}$ corresponds to K different MBIOS channels. Let $\{X_n\}_{n=1}^{\infty}$ be the random process representing the channel input sequence, where $x_n \in \mathscr{X} = \{0,1\}$. Let $\{Y_n\}_{n=1}^{\infty}$ be the random process representing the channel output sequence, where $y_n \in \mathscr{Y}$ and $\mathscr{Y}$ is assumed to be a discrete subset of the real line symmetric around zero. Let $P(\mathbf{x}^N,\mathbf{y}^N)$ be the joint pmf of $\mathbf{X}^N$ and $\mathbf{Y}^N$, where $\mathbf{x}^N$ denotes the length-N vector $(x_1,\ldots,x_N)$. Then,

$$P(\mathbf{x}^N,\mathbf{s}^N,\mathbf{y}^N) = Q(\mathbf{y}^N|\mathbf{x}^N,\mathbf{s}^N)P(\mathbf{s}^N|\mathbf{x}^N)P(\mathbf{x}^N) \tag{1a}$$

$$= P(\mathbf{x}^N)P(s_1)\prod_{n=1}^{N} Q(y_n|x_n,s_n)\prod_{n=1}^{N-1} P(s_{n+1}|s_n,x_n). \tag{1b}$$

For any sequence of real-valued random variables $(Z_1,Z_2,Z_3,\ldots)$, define the *limit inferior in probability* $p-\liminf_{N\to\infty} Z_N$ as

$$p-\liminf_{N\to\infty} Z_N \triangleq \sup\{\alpha| \lim_{N\to\infty} Pr[Z_N<\alpha]=0\}. \tag{2}$$

Then, the capacity of the aforementioned FSC when no channel state information is available at the transmitter and the receiver is defined as

$$C \triangleq \sup_{\mathbf{X}} \underline{I}(\mathbf{X};\mathbf{Y}), \tag{3}$$

where

$$\underline{I}(\mathbf{X};\mathbf{Y}) = p-\liminf_{N\to\infty} \frac{1}{N}\log_2\frac{Q(\mathbf{y}^N|\mathbf{x}^N)}{P(\mathbf{y}^N)}. \tag{4}$$

Since the focus of this work is not in finding this maximizing input distribution, in the following we will be interested in the maximum achievable rate for a certain sequence of input pmfs $P \triangleq \{P(\mathbf{x}^N)\}_N$, defined as

$$C_P \triangleq \underline{I}(\mathbf{X};\mathbf{Y}), \tag{5}$$

where $I(\mathbf{X};\mathbf{Y})$ is computed under $\{P(\mathbf{x}^N)\}_N$ using (1).

Construction of quantized code ensembles

In this section we construct quantized code ensembles that induces a Markov distribution on transmitted sequence.

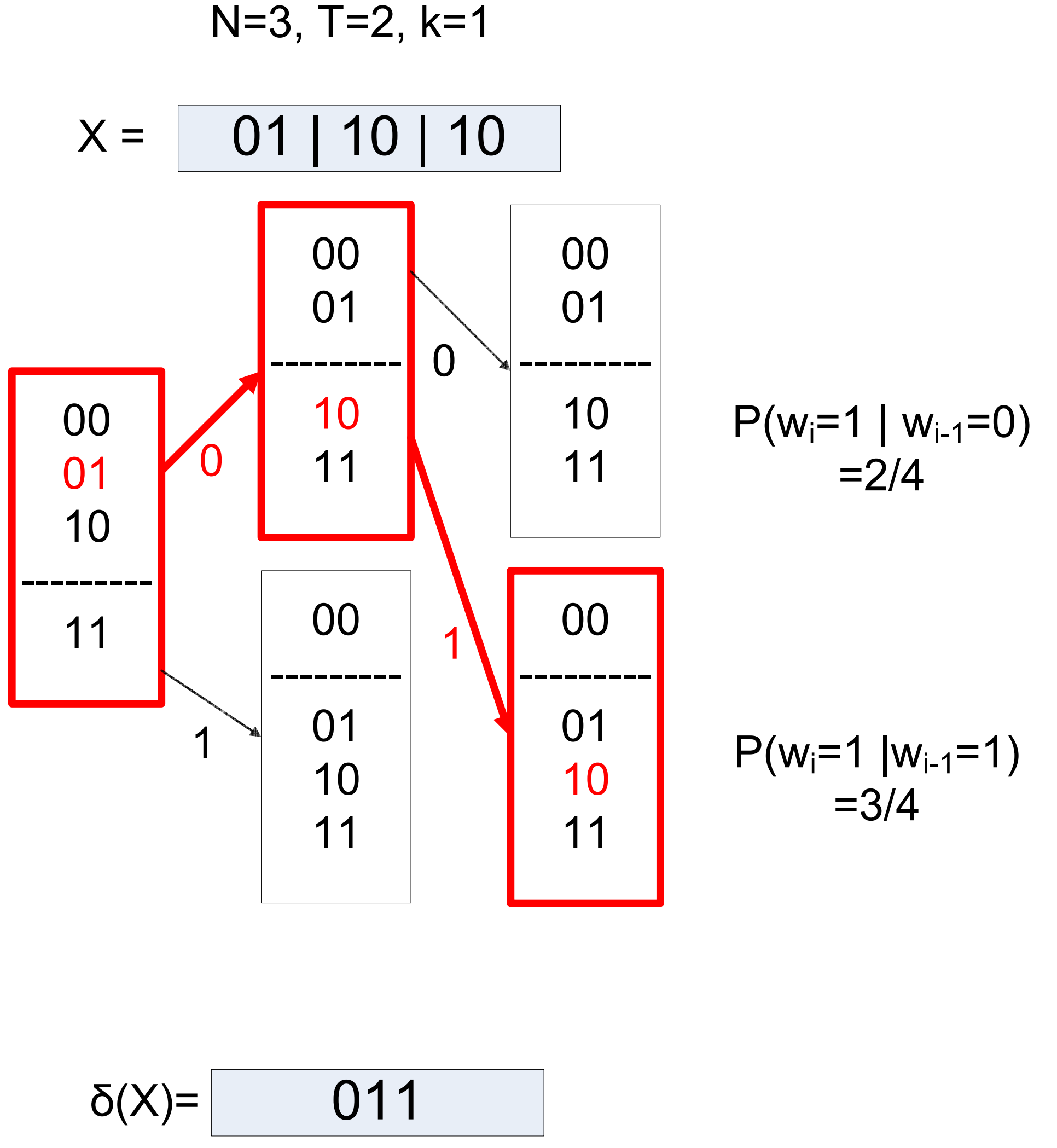
Definition 1. *Consider a sequence* $\mathbf{x}^{NT}$ *and some arbitrary function* $f:\{0,1\}^T\times\{0,1\}^k\to\{0,1\}$. *An order-k Markov quantizer (denoted by MQ-k) is a mapping* $\delta:\{0,1\}^{NT}\to\{0,1\}^N$ *with* $\delta(\mathbf{x}^{NT})=\mathbf{w}^N$, *with the following structure.*

$$w_n = f(\mathbf{x}_{(n-1)T+1}^{nT},\mathbf{w}_{n-k}^{n-1}), \qquad n=1,2,\ldots,N. \tag{6}$$

Consider now a pmf of a k-th order stationary Markov process $P(\mathbf{w}^N)=\prod_{n=1}^{N}P(w_n|\mathbf{w}_{n-k}^{n-1})$ *for a binary sequence of length N. An order-k Markov quantizer with respect to P (denoted by MQ-k-P) is an MQ-k satisfying*

$$\frac{|\{\mathbf{x}_{(n-1)T+1}^{nT}|f(\mathbf{x}_{(n-1)T+1}^{nT},\mathbf{w}_{n-k}^{n-1})=0\}|}{2^T} = P\left(0\mid\mathbf{w}_{n-k}^{n-1}\right), \quad \forall\mathbf{w}_{n-k}^{n-1} \text{ and } n=1,2,\ldots,N \tag{7}$$

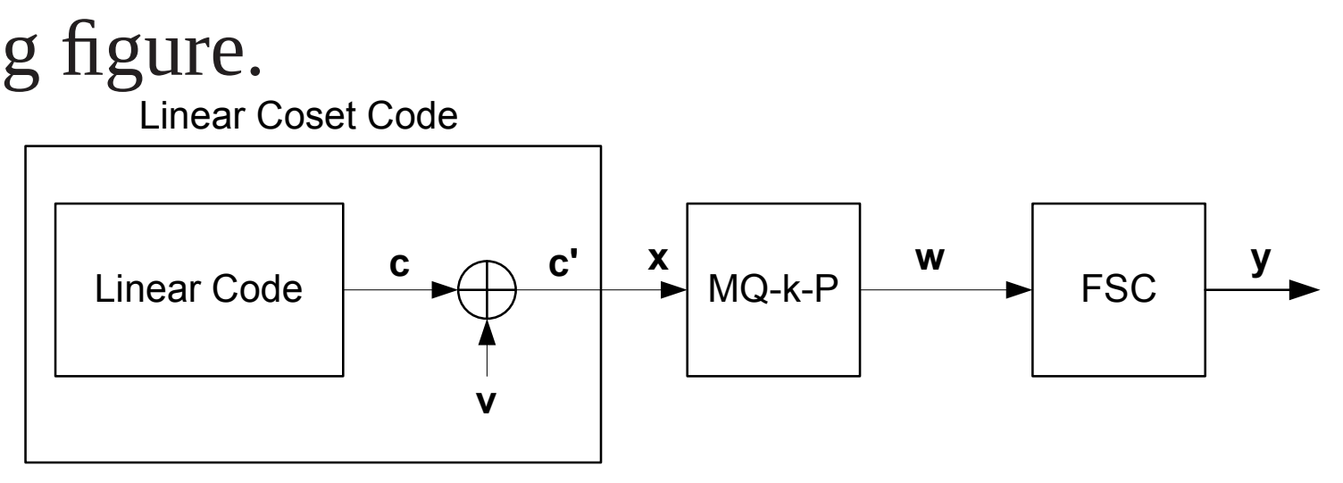
More descriptively, an MQ$-k$ partitions a length-NT binary sequence into N blocks of length T each, and then quantizes each block into a bit using a mapping that depends also on the k previously produced bits, thus producing a length-N binary sequence. This is shown in the following figure.



It can be shown that an MQ-k-P outputs a sequence with Markov distribution P if a input sequence of the MQ-k-P has i.u.d. (identical and uniform distribution). This motivates the following definition.

Definition 2. *Consider a code ensemble* $\mathscr{C}$. *The coset ensemble* $\mathscr{C}'$ *generated by* $\mathscr{C}$ *is defined as the ensemble generated from* $\mathscr{C}$ *by including, for each code* $C\in\mathscr{C}$, *codes of the form* $C'=\{\mathbf{c}\oplus\mathbf{v}|\mathbf{c}\in C\}$ *for all possible vectors* $\mathbf{v}$. *The measure on* $\mathscr{C}'$ *is the product of the uniform measure over all vectors* $\mathbf{v}$ *and the measure of the original ensemble* $\mathscr{C}$.

The block diagram of the proposed scheme is shown in the following figure.



Capacity achievability

In the following we will show that specific LDPC-like coset codes achieve the capacity of FSCs. We consider the following three ensembles with increasing degree of sophistication.

Ensemble A is the regular Gallager (N,d_v,d_c) LDPC ensemblewith column and row degrees d_v and d_c, respectively.

Ensemble B is the punctured LDPC ensemble resulting from puncturing a sufficiently low-rate (N,d_v,d_c) Gallager ensemble to achieve the specified rate.

Ensemble C is the LDPC-GM ensemble consisting of a a serial concatenation of an outer (N,d_v,d_c) Gallager LDPC code and an inner rate-one regular LDGM code with row and column degrees equal to d_c.

We are now ready to present an upper-bound on the error probability of quantized linear coset codes over FSCs.

Definition 3. *An MQ-k is called "robust" if the all-zeros block of length T,* $\mathbf{0}^T$, *and the all-ones block of length T,* $\mathbf{1}^T$, *are quantized to different values regardless of the quantizer memory, i.e.,*

$$\forall\mathbf{w}^k\in\{0,1\}^k: \quad f(\mathbf{0}^T,\mathbf{w}^k)\neq f(\mathbf{1}^T,\mathbf{w}^k), \tag{8}$$

In order to utilize the techniques for bounding the error probability, a special kind of linear code ensemble is considered. Let Π_N denote the set of all permutations of N numbers. We define a *permutation-invariant* code ensemble as follows.

Definition 4. *Let* $\mathscr{C}$ *be an ensemble of length-N block codes. We say that* $\mathscr{C}$ *is a* permutation-invariant *ensemble if for all permutations* $\pi\in\Pi_N$ *and for all codes* $C\in\mathscr{C}$, *it is true that* $\pi(C)\in\mathscr{C}$ *and the codes* $\pi(C)$ *are selected with the same probability. Here* $\pi(C)$ *denotes the codebook constructed by permuting the order of the symbols in all the codewords of C according to* π.

We now state an error probability upper bound for Markov-quantized ensembles. For a given ensemble $\mathscr{C}$ of codes with length N, we denote the average weight enumerator of the ensemble by $\overline{A}_l, l=0,1,\ldots,N$.

Proposition 1. *Consider a permutation-invariant ensemble* $\mathscr{C}$ *of binary linear codes with M codewords of length NT, rate R/T, and average weight enumerator* $\overline{A}_l$. *Let* $\mathscr{C}'$ *be the coset ensemble generated by* $\mathscr{C}$. *A code from* $\mathscr{C}'$ *is quantized using a "robust" MQ-k-P, and transmitted over an FSC. Let* $U\subseteq\{1,2,\ldots,NT\}$ *and* $NT\in U$. *Then, for any* $\varepsilon>0$, *the average error probability with ML decoding, is upper-bounded as*

$$\overline{P}_{e|m} \leq \sum_{l\in U\setminus\{NT\}}\overline{A}_l+\overline{A}_{NT}D_1^N+Pr[(\mathbf{X}^N,\mathbf{Y}^N)\notin T_N] + 2^{-N(C_P-R-\frac{\log_2\alpha}{N}-\varepsilon)} \tag{9}$$

where

$$D_1 \triangleq \frac{2^T-1+D}{2^T}, \tag{10}$$

$$D \triangleq \min_{s_0}\max_{s}\left[Q_s^+\sqrt{\frac{Q_{s_0}^-}{Q_{s_0}^+}}+Q_s^-\sqrt{\frac{Q_{s_0}^+}{Q_{s_0}^-}}\right], \tag{11}$$

$$Q_s^+ = \frac{1}{2}Q(0|1,s)+\sum_{y>0}Q(y|1,s), \tag{12a}$$

$$Q_s^- = \frac{1}{2}Q(0|1,s)+\sum_{y<0}Q(y|1,s), \tag{12b}$$

$$\alpha = \max_{l\in U^c}\frac{\overline{A}_l}{M-1}\frac{2^{NT}}{\binom{NT}{l}}, \tag{13}$$

$$T_N = \left\{(\mathbf{x}^N,\mathbf{y}^N)\in\mathscr{X}^N\times\mathscr{Y}^N|\frac{1}{N}\log_2\frac{Q(\mathbf{y}^N|\mathbf{x}^N)}{P(\mathbf{y}^N)}>C_P-\varepsilon\right\}. \tag{14}$$

By using the above bound, we can show that coset code ensembles generated by the ensembles A,B and C in conjunction with an MQ-k-P achieve C_P. Since a sequence of stationary ergodic Markov sources asymptotically achieves the capacity of FSCs as the order k goes to infinity , a sequence of quantized coset code ensembles asymptotically achieves the capacity of FSCs for a large enough T.