University of
Waterloo

Coding over an Erasure Channel with a Large Alphabet Size

Shervan Fashandi, Shahab Oveis Gharan and Amir K. Khandani

Electrical and Computer Engineering Department
University of Waterloo, Waterloo, ON, Canada
Email:{sfashand,shahab, khandani}@cst.uwaterloo.ca

Technical Report UW-E&CE#2008-06

Mar 19, 2008

Coding over an Erasure Channel with a Large Alphabet Size

Shervan Fashandi, Shahab Oveis Gharan and Amir K. Khandani
ECE Dept., University of Waterloo, Waterloo, ON, Canada, N2L3G1
email: {sfashand,shahab,khandani}@cst.uwaterloo.ca

Abstract

An erasure channel with a fixed alphabet size q, where $q \gg 1$, is studied . It is proved that over any erasure channel (with or without memory), *Maximum Distance Separable* (MDS) codes achieve the minimum probability of error (assuming maximum likelihood decoding). Assuming a memoryless erasure channel, the error exponent of MDS codes are compared with that of random codes. It is shown that the envelopes of these two exponents are identical for rates above the critical rate. Noting the optimality of MDS codes, it is concluded that random coding is exponentially optimal as long as the block size N satisfies $N < q + 1$. [1]

I. Introduction

Erasure channels with large alphabet sizes have recently received significant attention in networking applications. Different erasure channel models are adopted to study the performance of end-to-end connections over the Internet [1], [2]. In such models, each packet is seen as a $q = 2^b$-ary symbol where b is the packet length in bits. In this work, a memoryless erasure channel with a fixed but large alphabet size is considered. The error probability over this channel (assuming maximum-likelihood decoding) for MDS and random codebooks are compared and shown to be exponentially identical for rates above the critical rate.

Shannon [3] was the first who observed that the error probability for maximum likelihood decoding of a random code ($P_{E,ML}^{rand}$) can be upper-bounded by an exponentially decaying function with respect to the code block length N. This exponent is positive as long as the rate stays below the channel capacity $R < C$. Following this result, tighter bounds were proposed in the literature [4], [5]. Interestingly, this upper-bound on $P_{E,ML}^{rand}$ remains valid regardless of the alphabet size q, even in the case where q is larger than the block size N (e.g. see the steps of the proofs in [6]). There is also a lower-bound on the probability of error

[1] Financial support provided by Nortel and the corresponding matching funds by the Natural Sciences and Engineering Research Council of Canada (NSERC), and Ontario Centres of Excellence (OCE) are gratefully acknowledged.

using random coding which is known as the sphere packing bound [7]. For channels with a relatively small alphabet size ($q \ll N$), both the sphere packing bound and the random coding upper-bound on the error probability are exponentially tight for rates above the critical rate. However, the sphere packing bound is not tight if the alphabet size, q, is comparable to the coding block length N. For rates below the critical rate, modifications of random coding are proposed to achieve tighter bounds [8].

Maximum Distance Separable (MDS) [9] codes are optimum in the sense that they achieve the largest possible minimum distance, d_{min}, among all block codes of the same size. Indeed, any $[N, K]$ MDS code can be successfully decoded from any subset of its coded symbols of size K or more. This property makes MDS codes suitable for use over the erasure channels like the Internet [1], [2], [10]. However, the practical encoding-decoding algorithms for such codes have quadratic time complexity in terms of the code block length [11]. Theoretically, more efficient ($O\left(N \log^2 N\right)$) MDS codes can be constructed based on evaluating and interpolating polynomials over specially chosen finite fields using Discrete Fourier Transform [12]. However, in practice these methods can not compete with the quadratic methods except for extremely large block sizes. Recently, a family of almost-MDS codes with low encoding-decoding complexity (linear in length) is proposed and shown to provide a practical alternative for coding over the erasure channels like the Internet [13]. In these codes, any subset of symbols of size $K(1+\epsilon)$ is sufficient to recover the original K symbols with high probability [13]. Digital Fountain codes, based on the idea of almost-MDS codes, are proposed for information multicasting to many users over an erasure channel [14], [15].

In this work, a memoryless erasure channel with a fixed but large alphabet size is studied. First, it is proved that MDS block codes offer the minimum probability of decoding error over any erasure channel. Then, error exponents of MDS and random codes for a memoryless erasure channel are analyzed and shown to be identical at rates above the critical rate. Combining the two results, we conclude that random codes are exponentially as good as MDS codes (exponentially optimal) over a wide range of rates.

The rest of this paper is organized as follows. In section II, the erasure channel model is introduced, and the assumption of large alphabet sizes is justified. Section III proves that MDS codes are optimum over any erasure channel. Error exponents of MDS codes and random codes over a memoryless erasure channel are compared in section IV. Finally, section V concludes the paper.

II. Erasure Channel Model

The memoryless erasure channel studied in this work has the alphabet size q and the erasure probability π (see Fig. 1). The alphabet size q is assumed to be fixed and large, i.e., $q \gg 1$. Note that all the known MDS codes have alphabets of a large size (growing at least linearly with the block length N). Indeed, a conjecture on MDS codes states that for every linear $[N, K]$ MDS code over the Galois field $\mathbb{F}_q$, if

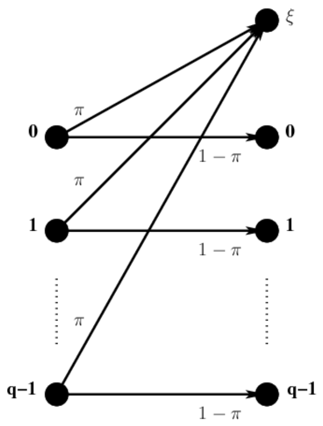


Fig. 1. Erasure memoryless channel model with the alphabet size q, probability of erasure π, and the erasure symbol ξ.

$1 < K < q$, then $N \leq q+1$, except when q is even and $K = 3$ or $K = q-1$, for which $N \leq q+2$ [16]. To have a feasible MDS code over a channel with the alphabet size q, the block size N should satisfy $N \leq q+1$. Thus, throughout this paper, wherever we refer to '*large block sizes*', it means N can grow large as long as it satisfies the constraint $N \leq q+1$.

The described channel model occurs in many practical scenarios such as the Internet. From an end to end protocol's perspective, performance of the lower layers in the protocol stack can be modeled as a random *channel* called an *Internet channel*. Since each packet usually includes an internal error detection mechanism (for instance a Cyclic Redundancy Check), the Internet channel can be modeled as an erasure channel with packets as symbols [17]. If each packet contains b bits, the corresponding channel will have an alphabet size of $q = 2^b$ which is huge for typical packet sizes. Therefore, in practical networking applications, the block size is usually much smaller than the alphabet size. Algebraic computations over Galois fields $\mathbb{F}_q$ of such large cardinalities is now practically feasible with the increasing processing power of electronic circuits. Note that network coding schemes, recently proposed and applied for content distribution over large networks, have a comparable computational complexity [18]–[20].

III. Optimality of MDS Codes over Erasure Channels

Maximum Distance Separable (MDS) codes are optimum in the sense of achieving the largest possible minimum distance, d_{min}, among all block codes of the same size [9]. The following proposition shows that MDS codes are also optimum over any erasure channel in the sense of achieving the minimum probability of decoding error.

Proposition I. Consider an erasure channel (memoryless or with memory) with the input vector $\mathbf{x} \in \mathcal{X}^N$, $|\mathcal{X}| = q$, the output vector $\mathbf{y} \in (\mathcal{X} \cup \{\xi\})^N$, and the transition probability $p(\mathbf{y}|\mathbf{x})$ satisfying:

1) $p\left(y_j \notin \{x_j, \xi\} |\, x_j\right) = 0, \ \forall\ j.$
2) Defining the erasure identifier vector $\mathbf{e}$ as

$$e_j = \begin{cases} 1 & y_j = \xi \\ 0 & \text{otherwise} \end{cases}$$

$p(\mathbf{e}|\mathbf{x})$ is independent of $\mathbf{x}$.

Consider a block code of size $[N, K]$ (with equiprobable codewords) over an erasure channel and an optimum (maximum likelihood) decoder. The code has the minimum probability of decoding error among all the $[N, K]$ block codes *if* it is *Maximum Distance Separable* (MDS).

Proof. Consider a $[N, K, d]$ codebook $\mathcal{C}$ with the q-ary codewords of length N, number of code-words q^K, and minimum distance d. The distance between two codewords is defined as the number of positions in which the corresponding symbols are different (Hamming distance). A codeword $\mathbf{x} \in \mathcal{C}$ is transmitted and a vector $\mathbf{y} \in \left(\mathcal{X} \cup \{\xi\}\right)^N$ is received. The number of erased symbols is equal to the Hamming weight of $\mathbf{e}$ denoted by $w(\mathbf{e})$. An error occurs if the decoder decides for a codeword different from $\mathbf{x}$. Let us assume that the probability of having a specific erasure pattern $\mathbf{e}$ is $\mathbb{P}\{\mathbf{e}\}$ which is independent of the transmitted codeword (depends only on the channel). We assume a specific erasure vector $\mathbf{e}$ of weight m. The decoder decodes the transmitted codeword based on the $N - m$ correctly received symbols. We partition the code-book, $\mathcal{C}$, into q^{N-m} bins, each bin representing a specific received vector satisfying the erasure pattern $\mathbf{e}$. The number of codewords in the i'th bin is denoted by $b_{\mathbf{e}}(i)$ for $i = 1, ..., q^{N-m}$. Knowing the erasure vector $\mathbf{e}$ and the received vector $\mathbf{y}$, the decoder selects the bin i corresponding to $\mathbf{y}$. The set of possible transmitted codewords is equal to the set of codewords in bin i (all the codewords in bin i are equiprobable to be transmitted). If $b_{\mathbf{e}}(i) = 1$, the transmitted codeword $\mathbf{x}$ can be decoded with no ambiguity. Otherwise, the optimum decoder randomly selects one of the $b_{\mathbf{e}}(i) > 1$ codewords in the bin. Thus, the probability of error is $1 - \frac{1}{b_{\mathbf{e}}(i)}$ when bin i is selected. Bin i is selected if one of the codewords it contains is transmitted. Hence, probability of selecting bin i is equal to $\frac{b_{\mathbf{e}}(i)}{q^K}$. Based on the above arguments, probability of decoding error for the maximum likelihood decoder of any codebook,$\mathcal{C}$, is equal to

$$\begin{aligned}
P_{E,ML}^{\mathcal{C}} &\overset{(a)}{=} \sum_{m=d}^{N} \sum_{\mathbf{e}:w(\mathbf{e})=m} \mathbb{P}\{\mathbf{e}\}\mathbb{P}\{\text{error}|\mathbf{e}\} \\
&= \sum_{m=d}^{N} \sum_{\mathbf{e}:w(\mathbf{e})=m} \mathbb{P}\{\mathbf{e}\} \sum_{i=1,\ b_{\mathbf{e}}(i)>0}^{q^{N-m}} \left(1-\frac{1}{b_{\mathbf{e}}(i)}\right)\frac{b_{\mathbf{e}}(i)}{q^K} \\
&\overset{(b)}{=} \sum_{m=d}^{N} \sum_{\mathbf{e}:w(\mathbf{e})=m} \mathbb{P}\{\mathbf{e}\} \left(1-\frac{b_{\mathbf{e}}^{+}}{q^K}\right) \\
&\geq \sum_{m=d}^{N} \sum_{\mathbf{e}:w(\mathbf{e})=m} \mathbb{P}\{\mathbf{e}\} \left(1-\frac{q^{N-m}}{q^K}\right) \qquad (1)
\end{aligned}$$

where $b_{\mathbf{e}}^{+}$ indicates the number of bins containing one or more codewords. (a) follows from the fact that the transmitted codeword can be uniquely decoded if the number of erasures in the channel is less than the minimum distance of the codebook and (b) follows from the fact that $\sum_{i=1}^{q^{N-m}} b_{\mathbf{e}}(i) = q^K$.

According to (1), $P_{E,ML}^{\mathcal{C}}$ is minimized for a code-book $\mathcal{C}$ *if* two conditions are satisfied. First, the minimum distance of $\mathcal{C}$ should achieve the maximum possible value, i.e., $d = N - K + 1$. Second, we should have $b_{\mathbf{e}}^{+} = q^{N-m}$ for all possible erasure vectors $\mathbf{e}$ with any weight $d \leq m \leq N$. Any MDS code satisfies the first condition by definition. Moreover, it is easy to show that for any MDS code, we have $b_{\mathbf{e}}(i) = q^{K-N+m}$. We first prove this for the case of $m = N - K$. Consider the bins of an MDS code for any arbitrary erasure pattern $\mathbf{e}, w(\mathbf{e}) = N - K$. From the fact that $d = N - K + 1$ and $\sum_{i=1}^{q^K} b_{\mathbf{e}}(i) = q^K$, it is concluded that each bin contains exactly one codeword. Therefore, there exists only one codeword which matches any K correctly received symbols. Now, consider any general erasure pattern $\mathbf{e}, w(\mathbf{e}) = m > N - K$. For the i'th bin, concatenating any $K - N + m$ arbitrary symbols to the $N - m$ correctly received symbols results in a distinct codeword of the MDS codebook. Having q^{K-N+m} possibilities to expand the received $N - m$ symbols to K symbols, we have $b_{\mathbf{e}}(i) = q^{K-N+m}$. This completes the proof.

A. MDS codes with Suboptimal Decoding

In the proof of proposition I, it is assumed that the received codewords are decoded based on maximum likelihood decoding which is optimum in this case. However, in many practical cases, MDS codes are decoded by simpler decoders [21]. Such suboptimal decoders can perfectly reconstruct the codewords of a $[N, K]$ codebook if they receive K or more symbols correctly. In case more than $N-K$ symbols are erased, a decoding error occurs. Let $P_{E,sub}^{MDS}$ denote the probability of this event. $P_{E,sub}^{MDS}$ is obviously different from the decoding error probability of the maximum likelihood decoder denoted by $P_{E,ML}^{MDS}$. Theoretically, an optimum maximum likelihood decoder of an MDS code may still decode the original codeword correctly with a positive, but small probability, if it receives less than K symbols. More precisely, according to the proof of Proposition I, such a decoder is able to correctly decode an MDS code over $\mathbb{F}_q$ with the probability of $\frac{1}{q^i}$ after receiving $K-i$ correct symbols. Of course, for Galois fields with large cardinality, this probability is usually negligible. The relationship between $P_{E,sub}^{MDS}$ and $P_{E,ML}^{MDS}$ can be summarized as follows

$$\begin{aligned} P_{E,ML}^{MDS} &= P_{E,sub}^{MDS} - \sum_{i=1}^{K} \frac{\mathbb{P}\{K-i \text{ symbols received correctly}\}}{q^i} \\ &\geq P_{E,sub}^{MDS} - \sum_{i=1}^{K} \frac{\mathbb{P}\{K-i \text{ symbols received correctly}\}}{q} \\ &= P_{E,sub}^{MDS}\left(1-\frac{1}{q}\right). \end{aligned} \tag{2}$$

Hence, $P_{E,ML}^{MDS}$ is bounded as

$$P_{E,sub}^{MDS}\left(1-\frac{1}{q}\right) \leq P_{E,ML}^{MDS} \leq P_{E,sub}^{MDS}. \tag{3}$$

IV. Error Exponents of MDS and Random Codes

A. Error Exponent of MDS Codes over a Memoryless Erasure Channel

Consider a block code of size $[N,K]$ over the memoryless erasure channel of Fig. 1. Let $\alpha = \frac{N-K}{N}$ define the coding overhead. For a q-ary $[N,K]$ code, the rate per symbol, R, is equal to

$$R = \frac{K}{N}\log q = (1-\alpha)\log q. \tag{4}$$

In a block code of length N, the number of lost symbols would be $\sum_{i=1}^{N} e_i$ where e_i is defined in Proposition I. Thus, the probability of decoding error for the suboptimal decoder of subsection III-A can be written as

$$P_{E,sub}^{MDS} = \mathbb{P}\left\{\frac{1}{N}\sum_{i=1}^{N} e_i > \alpha\right\} = \sum_{i=0}^{K-1} P_i \tag{5}$$

where P_i denotes the probability that i packets are received correctly. Since e_i's are i.i.d random variables with Bernoulli distribution, we have $P_i = (1-\pi)^i \pi^{N-i}\binom{N}{i}$. It is easy to see that

$$\frac{P_i}{P_{i-1}} = \frac{(N-i+1)(1-\pi)}{i\pi} > 1 \quad \text{for } i = 1, \cdots, K-1$$

if $\alpha = \frac{N-K}{N} > \pi$. According to equation (4), the condition $\alpha > \pi$ can be rewritten as $R < (1-\pi)\log q = C$ where C is the capacity of the memoryless erasure channel. Therefore, the summation terms in equation (5) are always increasing, and the largest term is the last one. Now, we can bound $P_{E,sub}^{MDS}$ as $P_{K-1} \leq P_{E,sub}^{MDS} \leq KP_{K-1}$. The term $\binom{N}{K-1}$ in P_{K-1} can be bounded using the fact that for any $N > K > 0$, we have [22]

$$\frac{1}{N+1}e^{NH\left(\frac{K}{N}\right)} \leq \binom{N}{K} \leq e^{NH\left(\frac{K}{N}\right)}$$

where the entropy, $H\left(\frac{K}{N}\right)$, is computed in nats. Thus, $P_{E,sub}^{MDS}$ is bounded as

$$\frac{\pi(1-\alpha)Ne^{-Nu(\alpha)}}{(1-\pi)(N+1)(\alpha N+1)} \leq P_{E,sub}^{MDS} \leq \frac{\pi(1-\alpha)^2N^2e^{-Nu(\alpha)}}{(1-\pi)(\alpha N+1)} \tag{6}$$

where $u(\alpha)$ is defined as

$$u(\alpha) = \begin{cases} 0 & \text{for } \alpha \leq \pi \\ \\ \alpha \log\left(\dfrac{\alpha(1-\pi)}{\pi(1-\alpha)}\right) & \\ -\log\left(\dfrac{1-\pi}{1-\alpha}\right) & \text{for } \pi < \alpha \leq 1. \end{cases} \tag{7}$$

with the $\log$ functions computed in the Neperian base.

Using equation (4), the MDS coding error exponent, $u(.)$, can be expressed in terms of R instead of α. In (4), K should be an integer, and we should have $q+1 \geq N$ for a feasible MDS code. Thus, the finest resolution of rates achievable by a single MDS codebook would be $R = \frac{i}{q+1}\log q$ for $i = 1, 2, \ldots, q$. Of course, it is also possible to achieve the rates in the intervals $\frac{i}{q+1}\log q < R < \frac{i+1}{q+1}\log q$ by time sharing between two MDS codebooks of sizes $[q+1, i]$ and $[q+1, i+1]$. However, in such cases, the smaller error exponent belonging to the codebook of the size $[q+1, i+1]$ dominates. Therefore, $u(R)$ will have a stepwise shape of the form

$$u(R) = \begin{cases} 0 & \text{for } 1-\pi \leq \tilde{r} \\ \\ -\tilde{r}\log\dfrac{(1-\pi)(1-\tilde{r})}{\tilde{r}\pi} & \\ -\log\dfrac{\pi}{1-\tilde{r}} & \text{for } 0 < \tilde{r} \leq 1-\pi \end{cases} \tag{8}$$

where $\tilde{r}$ is defined as

$$\tilde{r} = \frac{1}{q+1}\left\lceil\frac{(q+1)R}{\log q}\right\rceil \tag{9}$$

B. Random Coding Error Exponent of a Memoryless Erasure Channel

It is interesting to compare the error exponent in (8) with the random coding error exponent as described in [6]. This exponent, $E_r(R)$, can be written as

$$E_r(R) = \max_{0\leq\rho\leq 1}\left\{-\rho R + \max_{\mathbf{Q}} E_0(\rho, \mathbf{Q})\right\} \tag{10}$$

where $\mathbf{Q}$ is the input distribution, and $E_0(\rho, \mathbf{Q})$ equals

$$E_0(\rho, \mathbf{Q}) = -\log\left(\sum_{j=0}^{q}\left[\sum_{k=0}^{q-1} Q(k)P(j|k)^{\dfrac{1}{1+\rho}}\right]^{1+\rho}\right). \tag{11}$$

Due to the symmetry of the channel transition probabilities, the uniform distribution maximizes (10) over all possible input distributions. Therefore, $E_0(\rho, \mathbf{Q})$ can be simplified as

$$E_0(\rho, \mathbf{Q}) = -\log\left(\frac{1-\pi}{q^{\rho}} + \pi\right). \tag{12}$$

Solving the maximization (10), gives us $E_r(R)$ as

$$E_r(R) = \begin{cases} -\log\dfrac{1-\pi+\pi q}{q} - r\log q \\ \qquad\qquad \text{for } 0 \leq r \leq \dfrac{R_c}{\log q} \\ \\ -r\log\dfrac{(1-\pi)(1-r)}{r\pi} - \log\dfrac{\pi}{1-r} \\ \qquad\qquad \text{for } \dfrac{R_c}{\log q} \leq r \leq 1-\pi \end{cases} \tag{13}$$

where $r = \frac{R}{\log q}$, and $R_c = \frac{1-\pi}{1-\pi+\pi q}\log q$ are the normalized and the critical rates, respectively.

Comparing (8) and (13), we observe that the MDS codes and the random codes perform exponentially the same for rates between the critical rate and the capacity. However, for the region below the critical rate, where the error exponent of the random code decays linearly with R, MDS codes achieve a larger error exponent. It is worth noting that this interval is negligible for large alphabet sizes. Moreover, the stepwise graph of $u(R)$ meets its envelope as the steps are very small for large values of q.

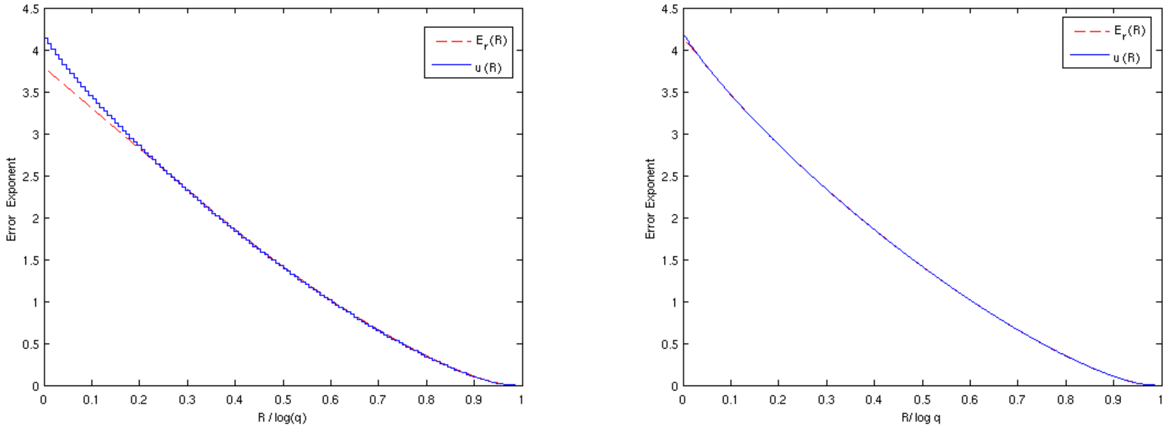
Figure 2 depicts the error exponents of random codes and MDS codes for the alphabet sizes of $q = 128$ and $q = 1024$ over an erasure channel with $\pi = 0.015$. As observed in Fig. 2(a), $u(R)$ can be approximated by its envelope very closely even for a relatively small alphabet size ($q = 128$). For a larger alphabet size (Fig. 2(b)), the graph of $u(R)$ almost coincides its envelope which equals $E_r(R)$ for the region above the critical rate. Moreover, as observed in Fig. 2(b), the region where MDS codes outperform random codes becomes very small even for moderate values of alphabet size ($q = 1024$).

C. Exponential Optimality of Random Coding

Using the sphere packing bound, it is shown that random coding is exponentially optimal for the rates above the critical rate over channels with relatively small alphabet sizes ($q \ll N$) [7]. However, the sphere packing bound is not tight for the channels whose alphabet size, q, is comparable to the block length. Here, based on Proposition I and the results of section IV, we prove the exponential optimality of random coding for erasure channels satisfying $q + 1 > N$.

Decoding error probability for a random codebook with the maximum-likelihood decoding can be upper bounded as

$$P_{E,ML}^{rand} \overset{(a)}{\leq} e^{-NE_r(R)} \overset{(b)}{\doteq} e^{-Nu(R)} \tag{14}$$



(a)

(b)

Fig. 2. Error exponents of random coding ($E_r(R)$) and MDS coding ($u(R)$) for a memoryless erasure channel with $\pi = 0.015$, and (a): $q = 128$, (b): $q = 1024$.

where (a) follows from [6], and (b) is valid only for rates above the critical rate according to (8) and (13).

We can also lower bound $P^{rand}_{E,ML}$ as

$$\begin{aligned} P^{rand}_{E,ML} &\overset{(a)}{\geq} P^{MDS}_{E,ML} \\ &\overset{(b)}{\geq} \left(1 - \frac{1}{q}\right) P^{MDS}_{E,sub} \\ &\overset{(c)}{\geq} \frac{\left(1 - \frac{1}{q}\right) \pi r N e^{-Nu(R)}}{(1-\pi)\,(N+1)\,((1-r)\,N+1)} \end{aligned} \tag{15}$$

where (a) follows from Proposition I, (b) from inequality (3), and (c) from inequality (6).

Combining (14) and (15) guarantees that both the upper-bound and the lower-bound on $P^{rand}_{E,ML}$ are exponentially tight, and the decaying exponent of $P^{rand}_{E,ML}$ versus N is indeed $u(R)$. Moreover, we can write

$$P^{MDS}_{E,ML} \overset{(a)}{\leq} P^{rand}_{E,ML} \overset{(b)}{\leq} \frac{(1-\pi)(N+1)(N-rN+1)}{\left(1 - \frac{1}{q}\right) \pi r N} P^{MDS}_{E,ML} \tag{16}$$

where (a) follows from Proposition I, and (b) results from inequalities (14) and (15). Since the coefficient of $P^{MDS}_{E,ML}$ in inequality (16) does not include any exponential terms, it can be concluded that for rates above the critical rate, random codes perform exponentially the same as MDS codes, which are already shown to be optimum.

V. Conclusion

Performance of random codes and MDS codes over an erasure channel with a fixed but large alphabet size is analyzed. It is shown that MDS codes minimize the probability of decoding error (using maximum-likelihood decoding), and any erasure channel (with or without memory). Then, the decoding error probability of MDS and random codes are bounded by exponential terms, and the corresponding exponents are compared. It is observed that the error exponents are identical over a wide range of rates. Knowing MDS codes are optimum, it is concluded that random coding is exponentially optimal over a memoryless erasure channel as long as the code block length, N, does not exceed the alphabet size of the channel by more than one.

References

[1] W. T. Tan and A. Zakhor, "Video Multicast Using Layered FEC and Scalable Compression," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, no. 3, pp. 373–386, 2001.

[2] L. Dairaine, L. Lancrica, J. Lacan, and J. Fimes, "Content-Access QoS in Peer-to-Peer Networks Using a Fast MDS Erasure Code," *Elsevier Computer Communications*, vol. 28, no. 15, pp. 1778–1790, 2005.

[3] C. E. Shannon, "A Mathematical Theory of Communications," *Bell Systems Technical Journal*, vol. 27, pp. 379–423,623–656, 1948.

[4] P. Elias, "Coding for Noisy Channels," *IRE Convention Record*, vol. 4, pp. 37–46, 1955.

[5] C. E. Shannon, R. G. Gallager, and E. R. Berlekamp, "Lower bounds to Error Probability for Coding on Discrete Memoryless Channels," *Information and Control*, vol. 10, pp. 65–103,522–552, 1967.

[6] R. G. Gallager, *Information Theory and Reliable Communication*, 1st ed. New York, NY, USA: John Wiley & Sons, 1968, pp. 135–144.

[7] ——, *Information Theory and Reliable Communication*, 1st ed. New York, NY, USA: John Wiley & Sons, 1968, pp. 157–158.

[8] G. Forney, "Exponential Error Bounds for Erasure, List, and Decision Feedback Schemes," *IEEE Transactions on Information Theory*, vol. 14, no. 2, pp. 206–220, 1968.

[9] Ron M. Roth, *Introduction to Coding Theory*, 1st ed. Cambridge University Press, 2006, pp. 333–351.

[10] X. H. Peng, "Erasure-control Coding for Distributed Networks," *IEE Proceedings on Communications*, vol. 152, pp. 1075 – 1080, 2005.

[11] N. Alon, J. Edmonds, and M. Luby, "Linear Time Erasure Codes with Nearly Optimal Recovery," in *IEEE Symposium on Foundations of Computer Science, Proc. IEEE Vol. 3*, 1995, pp. 512–519.

[12] J. Justesen , "On the complexity of decoding Reed-Solomon codes," *IEEE transactions on information theory*, vol. 22, no. 2, pp. 237–238, 1993.

[13] M. G. Luby, M. Mitzenmacher, M. A. Shokrollahi, and D. A. Spielman, "Efficient Erasure Correcting Codes," *IEEE Transactions on Information Theory*, vol. 47, no. 2, pp. 569–584, 2001.

[14] M. G. Luby, "LT Codes," in *IEEE Symposium on the Foundations of Computer Science (FOCS)*, 2002, pp. 271–280.

[15] A. Shokrollahi, "Raptor Codes," *IEEE Transactions on Information Theory*, vol. 52, no. 6, pp. 2551–2567, 2006.

[16] J. L. Walker, "A New Approach to the Main Conjecture on Algebraic-geometric MDS Codes," *Journal of Designs, Codes and Cryptography*, vol. 9, no. 1, pp. 115– 120, 1996.

[17] S. Fashandi, S. Oveisgharan, and A.K. Khandani, "Path Diversity in Packet Switched Networks: Performance Analysis and Rate Allocation," in *IEEE Global Telecommunications Conference, GLOBECOM '07*, 2007, pp. 1840–1844.

[18] R. Koetter and M. Medard , "An algebraic approach to network coding," *IEEE transactions on Networking*, vol. 11, no. 5, pp. 782–795, 2003.

[19] P. A. Chou, Y. Wu, and K. Jain, "Practical Network Coding ," in *51st Allerton Conference on Communication, Control and Computing*, 2003.

[20] C. Gkantsidis and P. R. Rodriguez, "Network coding for large scale content distribution," in *IEEE INFOCOM, Proc. IEEE Vol. 4*, 2005, pp. 2235–2245.

[21] Ron M. Roth, *Introduction to Coding Theory*, 1st ed. Cambridge University Press, 2006, pp. 183–204.

[22] T. Cover and J. Thomas, *Elements of Information Theory*, 1st ed. New York: Wiley, 2006, pp. 284–285.