Error Probability in the Block-Erasure Channel

Albert Guillén i Fàbregas
Institute for Telecommunications Research
University of South Australia
SPRI Building - Mawson Lakes Blvd.
Mawson Lakes SA 5095, Australia
e-mail : albert.guillen@unisa.edu.au

Abstract

We study an M-ary block-erasure channel with B blocks, where with probability ϵ a block of L coded symbols is erased. We study the behavior of the error probability of coded systems over such channels, and show that, if the code is diversitywise maximum-distance separable, its word error probability is equal to the outage probability, which admits a very simple expression. This paper is intended to complement the error probability analysis in Lapidoth's paper [1] and shed some light on the design of coding schemes for nonergodic channels.

1 Introduction and Channel Model

The block-erasure channel is a very simplified model of a fading channel where parts of the codeword are completely erased by a deep fade of the channel. This channel corresponds to the large signal-to-noise ratio (SNR) regime of the block-fading channel [2], and its interest lies on its simplicity and non-ergodicity, typical from many real wireless communication systems, such as orthogonal frequency division multiplexing (OFDM) or frequency-hopped systems. Coding for the block-erasure channel with convolutional codes has been studied in some detail in [1]. In this paper we study the problem of coding over the block-erasure channel and we show that the word and bit error probabilities admit very simple expressions. This paper complements the anaylsis in [1] for convolutional codes.

We study a block-erasure channel with B blocks. With probability ϵ a block of L symbols is completely erased and with probability $1-\epsilon$ a block of L coded symbols is received correctly (noiseless sub-channel), independently from block to block. Consider the transmission of an M-ary code $\mathcal{C}$ of length $N = BL$ and rate $R = \frac{K}{N}$ bits per channel use, where $K = \log_2 |\mathcal{C}|$. Also, let $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_B) \in \{0, 1, \ldots, M-1\}^N$ be the codewords of $\mathcal{C}$. We denote erasures by "?". The block-erasure channel is illustrated in Figure 1.

Define the erasure pattern vector $\mathbf{e} = (e_1, e_2, \ldots, e_B) \in \{0, 1\}^B$, whose b-th component is $e_b = 1$ if the block is erased and $e_b = 0$ otherwise. Thus, $P(e_b = 1) = \epsilon$ and $P(e_b = 0) = 1-\epsilon$, namely, the components of the erasure pattern $\mathbf{e}$ are i.i.d. Bernoulli random variables (with success probability ϵ). We assume that the receiver has channel state information (CSI), i.e., the receiver knows the erasure pattern $\mathbf{e}$.

2 Error Probability Analysis

In this section we define the *word* and *bit* error probabilities of coded schemes over the block-erasure channel described in the previous section. We also discuss the information theoretic limits of the channel.

We define the word error probability as the probability of decoding in favor of a codeword $\hat{\mathbf{x}}$ when codeword $\mathbf{x}$ was transmitted, averaged over all possible transmitted codewords $\mathbf{x} \in \mathcal{C}$

$$P_e^w(\epsilon) = \frac{1}{|\mathcal{C}|} \sum_{\hat{\mathbf{x}} \neq \mathbf{x}} \Pr\{\hat{\mathbf{x}} \neq \mathbf{x}\}. \tag{1}$$

We further consider linear codes only, and thus, the error probability does not depend on the transmitted codeword. We then assume the transmission of the all-zero codeword, i.e., $\mathbf{x} = (0, \ldots, 0)$. Consider the *maximum likelihood* (ML) decoder,

$$\hat{\mathbf{x}} = \arg \max_{\mathbf{x} \in \mathcal{C}} p(\mathbf{y}|\mathbf{x}) \tag{2}$$

and define the subsets

$$\mathcal{C}(\mathbf{e}) = \{\mathbf{x} \in \mathcal{C} \;\mid\; \text{if } \; e_b = 0, \;\; \mathbf{x}_b = (0, \ldots, 0) \;\; \forall b \in (1, \ldots, B)\} \tag{3}$$

as the subset of codewords that differ *only* in the erased symbols. Obviously, the transmitted codeword belongs to $\mathcal{C}(\mathbf{e})$ and by definition $|\mathcal{C}(\mathbf{e})| \geq 1, \forall \mathbf{e} \in \mathbb{F}_2^B$. In words, $\mathcal{C}(\mathbf{e})$ is the set of codewords that, once erased by a given erasure pattern, look identical to the receiver. In such a case, the ML decoder will resolve the ties evenly, and will make an error with probability [1]

$$P_e^w(\epsilon|\mathbf{e}) = 1 - \frac{1}{|\mathcal{C}(\mathbf{e})|} \tag{4}$$

which implies that

$$P_e^w(\epsilon) = \mathbb{E}[P_e^w(\epsilon|\mathbf{e})] = \mathbb{E}\left[1 - \frac{1}{|\mathcal{C}(\mathbf{e})|}\right]. \tag{5}$$

We remark that the only source of error (randomness) in the decoding process is essentially how the ML decoder resolves the ties between the equaly likely candidates in $\mathcal{C}(\mathbf{e})$.

We further define the average bit error probability as

$$P_e^b(\epsilon) = \frac{1}{K} \sum_{k=1}^{K} P_{e,k}(\epsilon) \tag{6}$$

where $P_{e,k}(\epsilon)$ is the probability of error of the k-th information bit.

Definition 1 *The block-diversity of a code is defined as*

$$\delta = \min_{\substack{\mathbf{x} \in \mathcal{C} \\ \mathbf{x} \neq \mathbf{0}}} |\{b \in (1, \ldots, B) \;\mid\; \mathbf{x}_b \neq \mathbf{0}\}|. \tag{7}$$

In words, δ represents the limit number erased blocks that $\mathcal{C}$ can tolerate. Specifically, if $\delta \leq \sum_{b=1}^{B} e_b$, $|\mathcal{C}(\mathbf{e})| > 1$ and the ML decoder will make an error with probability $1 - \frac{1}{|\mathcal{C}(\mathbf{e})|}$. Obviously, $\delta \leq B$. If $\delta = B$ we say that $\mathcal{C}$ has *full diversity*. The definition of δ shows that it corresponds to the minimum distance of a code of length B constructed over an alphabet of size M^L. Therefore by using the Singleton bound we get [2]

$$\delta \leq \delta_B \tag{8}$$

where

$$\delta_B \triangleq 1 + \left\lfloor B\left(1 - \frac{R}{\log_2 M}\right)\right\rfloor \tag{9}$$

The Singleton bound states that given B, R and M, the block diversity cannot be larger than (9), and thus full diversity is only guaranteed if $\frac{R}{\log_2 M} \leq \frac{1}{B}$.

Definition 2 *A code $\mathcal{C}$ is diversitywise maximum-distance separable (MDS) if it meets the Singleton bound with equality, i.e., $\delta = \delta_B$.*

In the following we elaborate on the optimality of MDS codes over the block-erasure channel.

2.1 Outage Probability

Similarly to other non-ergodic channels, the block-erasure channel has zero capacity in the strict Shannon sense, since, in this case, with probability ϵ^B all channels are erased and reliable communication is not possible. We define the *information outage probability* as the probability that the transmission rate R is not supported by a given channel realization,

$$P_{\text{out}}(\epsilon) \triangleq \Pr\{I(\mathbf{e}) < R\} \tag{10}$$

where $I(\mathbf{e})$ denotes the *instantaneous* mutual information between the input and output of the channel for a given erasure pattern. In such nonergodic channels, $P_{\text{out}}(\epsilon)$ is then the *best possible* word error probability[1]. In our case,

$$I(\mathbf{e}) = \frac{1}{B}\sum_{b=1}^{B} \bar{e}_b \log_2 M \quad \text{(bits per channel use)} \tag{11}$$

where $\bar{e}_b$ denotes the binary complement of e_b. This comes from the fact that the block-erasure channel is nothing but a set of B parallel channels (used an equal fraction of the time), each conveying either $\log_2 M$ bits per channel use if $e_b = 0$ or 0 if $e_b = 1$. We have the following result

Proposition 1 *Consider the transmission M-ary codes over the block-erasure channel. Then,*

$$\lim_{\epsilon \to 0} P_{\text{out}}(\epsilon) = \binom{B}{\lceil \frac{BR}{\log_2 M} \rceil - 1} \epsilon^{\delta_B} \tag{12}$$

[1]Remark that this is only true for large block length. In general, Fano's inequality gives [3, 4]

$$P_e(\epsilon) \geq \mathbb{E}\left[\left|1 - \frac{I(\mathbf{e})}{R} - \frac{1}{BLR}\right|_+\right]$$

where $|x|_+ = \max\{0, x\}$.

Proof. We can write the outage probability as

$$\begin{aligned}
P_{\text{out}}(\epsilon) &= \Pr\{I(\mathbf{e}) < R\} \\
&= \Pr\left\{\frac{1}{B}\sum_{b=1}^{B} \bar{e}_b \log_2 M < R\right\} \\
&= \Pr\left\{\sum_{b=1}^{B} \bar{e}_b < \frac{BR}{\log_2 M}\right\} \\
&= \Pr\left\{A \leq \left\lceil \frac{BR}{\log_2 M}\right\rceil - 1\right\} \\
&= \sum_{k=0}^{\left\lceil \frac{BR}{\log_2 M}\right\rceil - 1} \binom{B}{k} (1-\epsilon)^k \epsilon^{B-k}
\end{aligned} \tag{13}$$

where $A \triangleq \sum_{b=1}^{B} \bar{e}_b$ is a binomial random variable with success probability $1-\epsilon$. We have quite trivially expressed the outage probability as the c.d.f. of a binomial random variable. Therefore we clearly get that

$$\begin{aligned}
\lim_{\epsilon\to 0} P_{\text{out}}(\epsilon) &= \binom{B}{\left\lceil \frac{BR}{\log_2 M}\right\rceil - 1} \epsilon^{B-\left\lceil \frac{BR}{\log_2 M}\right\rceil+1} \\
&= \binom{B}{\left\lceil \frac{BR}{\log_2 M}\right\rceil - 1} \epsilon^{B+\left\lfloor -\frac{BR}{\log_2 M}\right\rfloor+1} \\
&= \binom{B}{\left\lceil \frac{BR}{\log_2 M}\right\rceil - 1} \epsilon^{1+\left\lfloor B\left(1-\frac{R}{\log_2 M}\right)\right\rfloor}
\end{aligned} \tag{14}$$

which shows the result. □

Remark 1 *The outage probability has slope δ_B for low ϵ in a log-log scale, and asymptotic coding gain $\binom{B}{\left\lceil \frac{BR}{\log_2 M}\right\rceil - 1}$. Thus it clearly corresponds to the high SNR regime of a block-fading channel [5].*

Remark 2 *If $\frac{R}{\log_2 M} = \frac{1}{B}$ (full diversity), $P_{\text{out}}(\epsilon) = \epsilon^B$, i.e., the probability that the rate R is not supported by the channel is equal to the probability of having all the blocks erased.*

Figure 2 shows the outage probability and the asymptotic limit (12) for $R = \frac{1}{2}$ binary codes ($M = 2$) over a block erasure channel with $B = 2, 4$ and 8 blocks.

2.2 Word Errors

The previous result proves the optimality (in diversity only) of designing MDS codes for such channels. In order to achieve the optimal performance, we start from

$$P_e^w(\epsilon|\mathbf{e}) = 1 - \frac{1}{|\mathcal{C}(\mathbf{e})|}. \tag{15}$$

For any code (with a given block diversity $\delta \le \delta_B$), since $|\mathcal{C}(\mathbf{e})| \le |\mathcal{C}|$ we can trivially upper-bound (15) (for large block length) as

$$P_e^w(\epsilon|\mathbf{e}) \le \begin{cases} 1 & \text{if } \sum_{b=1}^B e_b \ge \delta \\ 0 & \text{if } \sum_{b=1}^B e_b < \delta \end{cases} \tag{16}$$

which leads to

$$\begin{aligned} P_e^w(\epsilon) &\le \Pr\left\{\sum_{b=1}^B e_b \ge \delta\right\} = \Pr\left\{\sum_{b=1}^B \bar{e}_b \le B-\delta\right\} \\ &= \sum_{k=0}^{B-\delta} \binom{B}{k}(1-\epsilon)^k \epsilon^{B-k}. \end{aligned} \tag{17}$$

In general (17) is not necessairly tight. However (and possibly surprisingly), if $\mathcal{C}$ is MDS, i.e., $\delta = \delta_B$, the bound is tight since (17) coincides with $P_{\text{out}}(\epsilon)$. In other words, if $\mathcal{C}$ is MDS, its word error probability is given by the outage probability, since the decoder will decode correctly under all erasure patterns such that $\sum_{b=1}^B e_b < \delta_B$.

Figure 3 confirms the above discussion. We have plotted the outage probability and the limiting behavior (12), as well as the word error rate (WER) simulations for the $(23, 33)_8$ and $(133, 171)_8$ convolutional codes with $L = 25$ (circles/crosses) and $L = 2500$ (diamonds/squares) respectively. As we observe, the simulated WER of the different codes for the different block lengths is the same and matches perfectly with the outage probability.

Remark 3 *This result can be a priori surprising, since it characterizes the performance of* ***any*** *MDS code of any (sufficiently large) block length over the block erasure channel. A posteriori, the result seems rather obvious, since it clearly follows as an artifact from the channel model and the definition of MDS codes. We should not then be misled by this, since in realistic non-ergodic block-fading noisy channels MDS codes are necessary, but not sufficient to approach the outage probability [6]. For example, the WER of convolutional codes in the block-fading channel increases with the block length, while the WER of concatenated MDS codes (as the blockwise concatenated codes of [5, 6] or the parallel turbo-codes of [7]) is given by the distribution of the decoding threshold [8].*

2.3 Bit Errors

In this section we show that diversitywise MDS codes are also optimal for the bit error probability. We start with a very simple upperbound

$$P_e^b(\epsilon|\mathbf{e}) \le \begin{cases} \frac{1}{2} & \text{if } \sum_{b=1}^B e_b \ge \delta \\ 0 & \text{if } \sum_{b=1}^B e_b < \delta \end{cases} \tag{18}$$

which yields that

$$P_e^b(\epsilon) = \mathbb{E}[P_e^b(\epsilon|\mathbf{e})] \le \frac{1}{2}\Pr\left\{\sum_{b=1}^B e_b \ge \delta\right\} = \frac{1}{2}P_e^w(\epsilon). \tag{19}$$

By using the bit-error version of Fano's inequality with [9, Th. 4.3.2] (using the fact that the encoder's inputs are bits) we can lowerbound the bit-error probability and get that[2]

$$P_e^b(\epsilon|\mathbf{e}) \geq h^{-1}\left(\left|1 - \frac{I(\mathbf{e})}{R}\right|_+\right) \tag{20}$$

where $h(p) = p\log_2\frac{1}{p} + (1-p)\log_2\frac{1}{1-p}$ is the binary entropy function and $p = h^{-1}(x)$ denotes the probability p for which $h(p) = x$. Therefore, we get that

$$\begin{aligned} P_e^b(\epsilon) &= \mathbb{E}[P_e^b(\epsilon|\mathbf{e})] \geq \mathbb{E}\left[h^{-1}\left(\left|1 - \frac{I(\mathbf{e})}{R}\right|_+\right)\right] \\ &= \mathbb{E}\left[h^{-1}\left(\left|1 - \frac{\log_2 M \sum_{b=1}^B \bar{e}_b}{BR}\right|_+\right)\right] \\ &= \sum_{k=0}^{\left\lceil\frac{BR}{\log_2 M}\right\rceil - 1} h^{-1}\left(\left|1 - \frac{\log_2 M}{BR}k\right|_+\right)\binom{B}{k}(1-\epsilon)^k\epsilon^{B-k} \end{aligned} \tag{21}$$

since

$$h^{-1}\left(\left|1 - \frac{\log_2 M \sum_{b=1}^B \bar{e}_b}{BR}\right|_+\right) = 0 \quad \text{when} \sum_{b=1}^B \bar{e}_b \geq \frac{BR}{\log_2 M}. \tag{22}$$

Therefore, the bit-error probability has the same slope, namely, the Singleton bound δ_B and this slope is again achievable with MDS codes. We can also lowerbound $P_e^b(\epsilon|\mathbf{e})$ as,

$$P_e^b(\epsilon|\mathbf{e}) \geq \begin{cases} \frac{1}{2}\frac{B-\sum_{b=1}^B \bar{e}_b}{B} & \text{if } \sum_{b=1}^B \bar{e}_b < \frac{BR}{\log_2 M} \\ 0 & \text{otherwise} \end{cases} \tag{23}$$

which yields to

$$P_e^b(\epsilon) = \mathbb{E}[P_e^b(\epsilon|\mathbf{e})] \geq \sum_{k=0}^{\left\lceil\frac{BR}{\log_2 M}\right\rceil - 1} \frac{1}{2}\frac{B - \sum_{b=1}^B \bar{e}_b}{B}\binom{B}{k}(1-\epsilon)^k\epsilon^{B-k} \tag{24}$$

since the best the decoder can do in case of an outage event to guess a fraction $\frac{B-\sum_{b=1}^B e_b}{B}$ of the bits and correct all the others. Depending on the structure of the code, the decoder will do worse than that. For example, in the ML decoder of a convolutional code will choose a wrong path through the trellis, which will yield more errors in the bits corresponding to the non-erased blocks.

Remark 4 *The maximum exponent of ϵ in the BER expression for the $(23,33)_8$ code with periodic interleaving in [1, pp. 1470]*

$$P_e^b(\epsilon) = 23.5\epsilon^5(1-\epsilon)^3 + 13.5\epsilon^6(1-\epsilon)^2 + 4\epsilon^7(1-\epsilon) + 0.5\epsilon^8 \tag{25}$$

should not come as a surprise, since the code with periodic interleaving is MDS, and thus, $\delta_B = 5$ (the results in [1] are plotted in a linear scale for ϵ and the effect of the slope is not evident). It should also be clear that a random interleaver yields a non-MDS convolutional code, and hence its error probability has a worse slope.

[2]This lowerbound was also obtained in [10] for the multiple-antenna case.

Remark 5 *The upperbound* (19) *and the lowerbounds* (21) *and* (24) *coincide for full diversity codes.*

Figure 4 shows several bounds and simulations of the bit-error probability. As we see, the difference between the two lowerbounds is quite remarkable, which indicates that (21) might not be achievable in general.

3 Conclusions

A rather simple analysis reveals the usefulness of diversitywise MDS codes for the non-ergodic block-erasure channel. We show that these codes are optimal in this channel and we derive the expresisons of their frame error rate as well as tight bounds on their bit-error rate.

Acknowledgment

I am grateful to Joseph Boutros for his interest and comments on an earlier version of the paper.

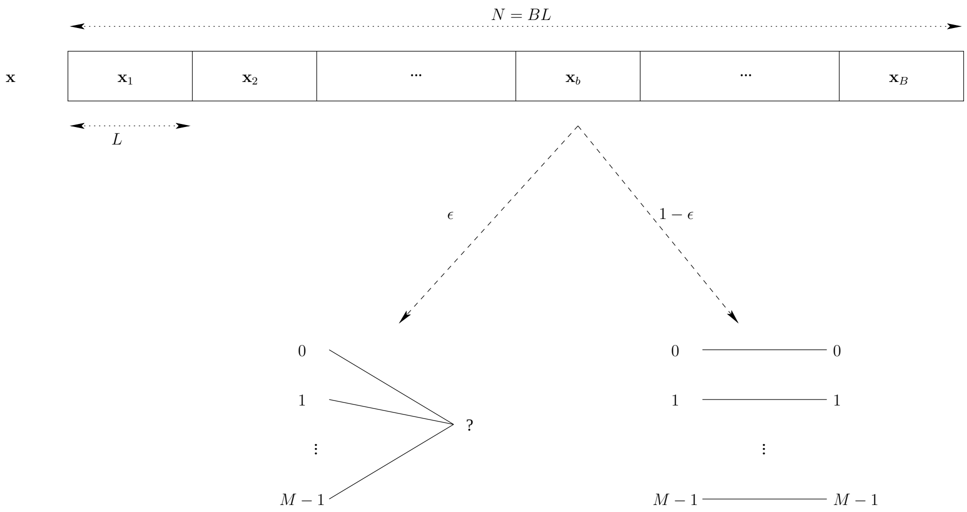

Figure 1: The block-erasure channel with B blocks. The symbols of block b, $b = 1, \ldots, B$ are erased with probability ϵ. The symbols of block b, $b = 1, \ldots, B$ are received correctly (noiseless channel) with probability $1 - \epsilon$.

References

[1] A. Lapidoth, "The performance of convolutional codes on the block erasure channel using various finite interleaving techniques," *IEEE Trans. on Inform. Theory*, vol. 40, no. 5, pp. 1459–1473, Sept. 1994.

[2] R. Knopp and P. Humblet, "On coding for block fading channels," *IEEE Trans. on Inform. Theory*, vol. 46, no. 1, pp. 1643–1646, July 1999.

[3] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, Wiley Series in Telecommunications, 1991.

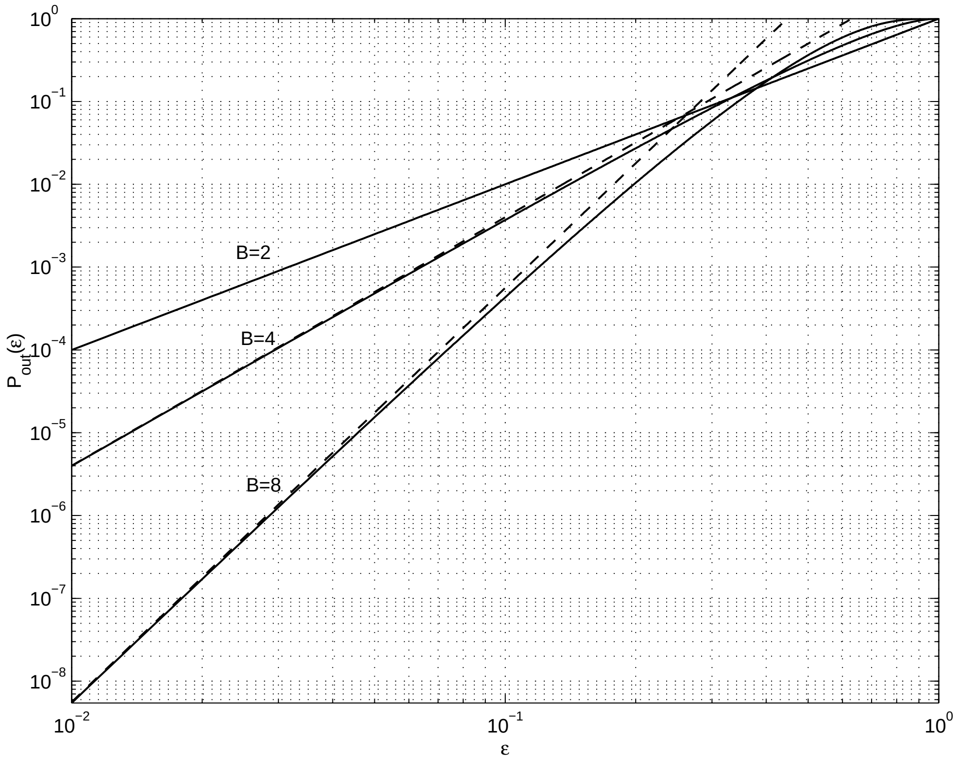

Figure 2: Outage probability (continuous lines) and Eq. (12) (dashed lines) in a log-log scale in a block erasure channel with $B = 2$, $B = 4$ and $B = 8$ blocks for $R = \frac{1}{2}$ and $M = 2$. The Singleton bound gives $\delta_B = 2, 3$ and 5 respectively.

[4] L. Zheng and D. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple antenna channels," *to appear IEEE Trans. on Inform. Theory*, 2002.

[5] A. Guillén i Fàbregas and G. Caire, "Turbo-like codes over the block-fading channel," in *2004 International Symposium on Information Theory, Chicago, IL*, June-July 2004.

[6] A. Guillén i Fàbregas, *Concatenated Codes for Block-Fading Channels*, Ph.D. thesis, Ecole Polytechnique Fédérale de Lausanne (EPFL) and Institut Eurécom, 2004.

[7] J. Boutros, E. Calvanese Strinati and A. Guillén i Fàbregas, "Turbo code design for block-fading channels," in *42nd Annual Allerton Conference on Communication, Control and Computing, Allerton, IL*, Sept.-Oct. 2004.

[8] J. Boutros, A. Guillén i Fàbregas and E. Calvanese Strinati, "Analysis of coding on nonergodic fading fading channels," in *43rd Annual Allerton Conference on Communication, Control and Computing, Allerton, IL*, Sept. 2005.

[9] R. G. Gallager, *Reliable Communication and Information Theory*, John Wiley and Sons, 1968.

[10] G. Caire and H. El Gamal, "A lower bound on BER for MIMO quasistatic fading channels," *private communication*, Jul. 2003.

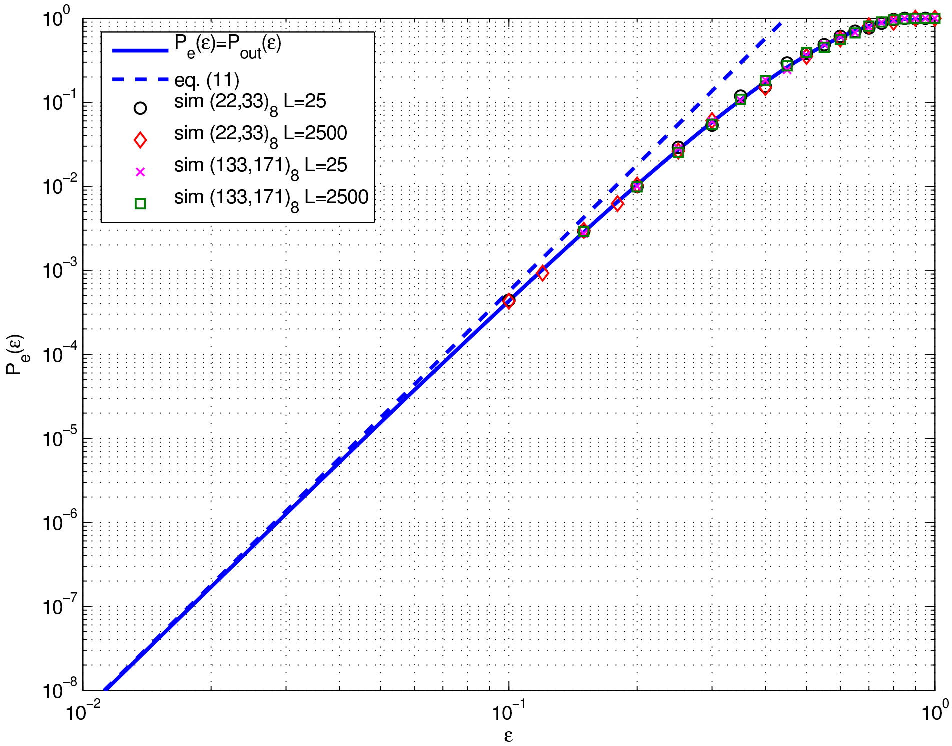

Figure 3: Outage probability (continuous lines), Eq. (12) (dashed lines) and simulations with the $(23, 33)_8$ and $(133, 171)_8$ convolutional codes with $L = 25$ (corresponds to 100 information bits per codeword) and $L = 2500$ (corresponds to 10000 information bits per codeword) in log-log scale in a block erasure channel with $B = 8$ blocks.

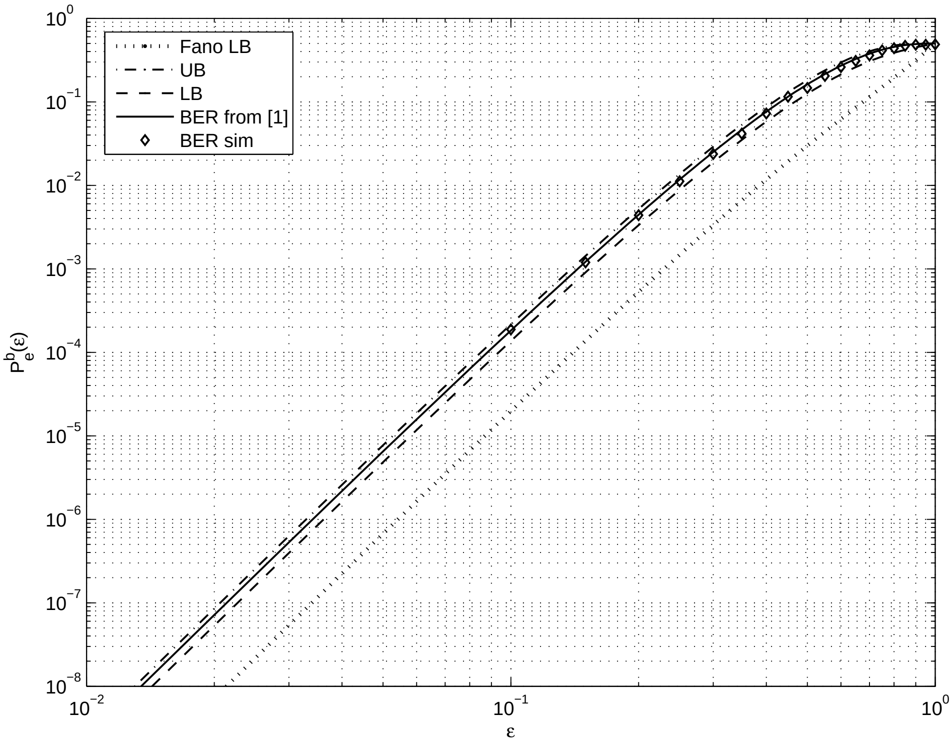

Figure 4: Bit-error probability in log-log scale in a block erasure channel with $B = 8$ blocks for $R = 1/2$. We show the Fano lowerbound (21) (dotted line), the upperbound (19) (dash-dotted line), the lowerbound (24) (dotted line), the analytical BER expression for the $(23, 33)_8$ code from [1] (continuous line) and the BER simulation (diamonds).