ECEN 5682 Theory and Practice of Error Control Codes

Block Code Performance

Peter Mathys

University of Colorado

Spring 2007

Performance Measures

"Goodness" Criteria for Block Codes. How does one select the "best" code for a particular application? This is not an easy question to answer in general because many factors, some of which are only marginally related to coding theory (e.g. the wordlength used on a VLSI chip), need to be considered. But since error control codes are used to control the effect of transmission or channel errors, the *probability of error after encoding and decoding with a code* $\mathcal{C}$ is certainly one of the major, if not the most important, criterion for "goodness". Depending on whether a code is used for error detection and/or error correction, we are interested in the *probability of an undetected error* $P_u(\mathcal{E})$, or in the *probability of a decoding error* $P(\mathcal{E})$.

Probability of Undetected Error

Theorem: Let $\mathcal{C}$ be any binary code with minimum (Hamming) distance $d_{\min}$. Then the probability of undetected error, $P_u(\mathcal{E})$, on a memoryless binary symmetric channel (BSC) with transition probability ϵ satisfies

$$P_u(\mathcal{E}) \leq \sum_{w=d_{\min}}^{n} \binom{n}{w} \epsilon^w (1-\epsilon)^{n-w} .$$

Proof: Since the code has minimum distance $d_{\min}$, it can detect all error patterns of (Hamming) weight $d_{\min} - 1$ or less. QED

Note: The above theorem also holds for memoryless q-ary symmetric channels (QSC) with transition probabilities (x denotes input, y denotes output of channel)

$$P\{y{=}\beta | x{=}\alpha\} = \begin{cases} 1-\epsilon , & \text{if } \alpha = \beta,\ \alpha, \beta \in \{0, 1, \ldots q-1\} , \\ \frac{\epsilon}{q-1} , & \text{if } \alpha \neq \beta,\ \alpha, \beta \in \{0, 1, \ldots q-1\} . \end{cases}$$

Note: Usually, many more error patterns can be detected than those indicated by the above theorem. A q-ary linear (n, k) code, for example, can detect $q^n - q^k$ error patterns.

Definition: Let A_w be the number of codewords of a linear (n, k) code with (Hamming) weight w. Then $\{A_0, A_1, \ldots, A_n\}$ is called the *weight distribution* of the code.

Example: Code #1 (binary $(5, 2, 3)$ code) with generator matrix

$$\mathbf{G} = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 \end{bmatrix},$$

has codewords

$$\mathcal{C} = \{00000, 01011, 10101, 11101\},$$

and thus weight distribution

$$\{A_0 = 1,\ A_3 = 2,\ A_4 = 1, \quad A_w = 0,\ w \neq 0, 3, 4\}.$$

Theorem: Let $\mathcal{C}$ be a binary linear (n, k) code with weight distribution $\{A_0, A_1, \ldots, A_n\}$ which is used on a binary symmetric channel (BSC) with transition probability ϵ. Then the probability of undetected error, $P_u(\mathcal{E})$, is given by

$$P_u(\mathcal{E}) = \sum_{w=d_{\min}}^{n} A_w \, \epsilon^w \, (1-\epsilon)^{n-w} \, .$$

Note that if $\mathcal{C}$ has minimum distance $d_{\min}$, then $A_1, A_2, \ldots, A_{d_{\min}-1}$ are all zero.

Proof: Every error pattern that is equal to a codeword cannot be detected when a linear code is used.

Note: For a q-ary symmetric channel (QSC) with transition probabilities as given earlier, this theorem becomes

$$P_u(\mathcal{E}) = \sum_{w=d_{\min}}^{n} A_w \left(\frac{\epsilon}{q-1}\right)^w (1-\epsilon)^{n-w} \, .$$

Examples: The weight distributions $\{A_w\}$ of a few small binary codes are:

- $(7, 4, 3)$ Hamming code:

$$\{A_0 = 1, A_3 = 7, A_4 = 7, A_7 = 1\} .$$

- $(15, 11, 3)$ Hamming code:

$$\{A_0 = 1, A_3 = 35, A_4 = 105, A_5 = 168, A_6 = 280, A_7 = 435, A_8 = 435,$$
$$A_9 = 280, A_{10} = 168, A_{11} = 105, A_{12} = 35, A_{15} = 1\} .$$

- $(31, 26, 3)$ Hamming code:

$$\{A_0 = 1, A_3 = 155, A_4 = 1085, A_5 = 5208, A_6 = 22568, A_7 = 82615, A_8 = 247845,$$
$$A_9 = 628680, A_{10} = 1383096, A_{11} = 2648919, A_{12} = 4414865, A_{13} = 6440560,$$
$$A_{14} = 8280720, A_{15} = 9398115, A_{16} = 9398115, A_{17} = 8280720, A_{18} = 6440560,$$
$$A_{19} = 4414865, A_{20} = 2648919, A_{21} = 1383096, A_{22} = 628680, A_{23} = 247845,$$
$$A_{24} = 82615, A_{25} = 22568, A_{26} = 5208, A_{27} = 1085, A_{28} = 155, A_{31} = 1\} .$$

Examples: (contd.)

- $(31, 21, 5)$ BCH code:

$$\{A_0 = 1, A_5 = 186, A_6 = 806, A_7 = 2635, A_8 = 7905, A_9 = 18910, A_{10} = 41602, A_{11} = 85560, A_{12} = 142600, A_{13} = 195300, A_{14} = 251100, A_{15} = 301971, A_{16} = 301971, A_{17} = 251100, A_{18} = 195300, A_{19} = 142600, A_{20} = 85560, A_{21} = 41602, A_{22} = 18910, A_{23} = 7905, A_{24} = 2635, A_{25} = 806, A_{26} = 186, A_{31} = 1\} .$$

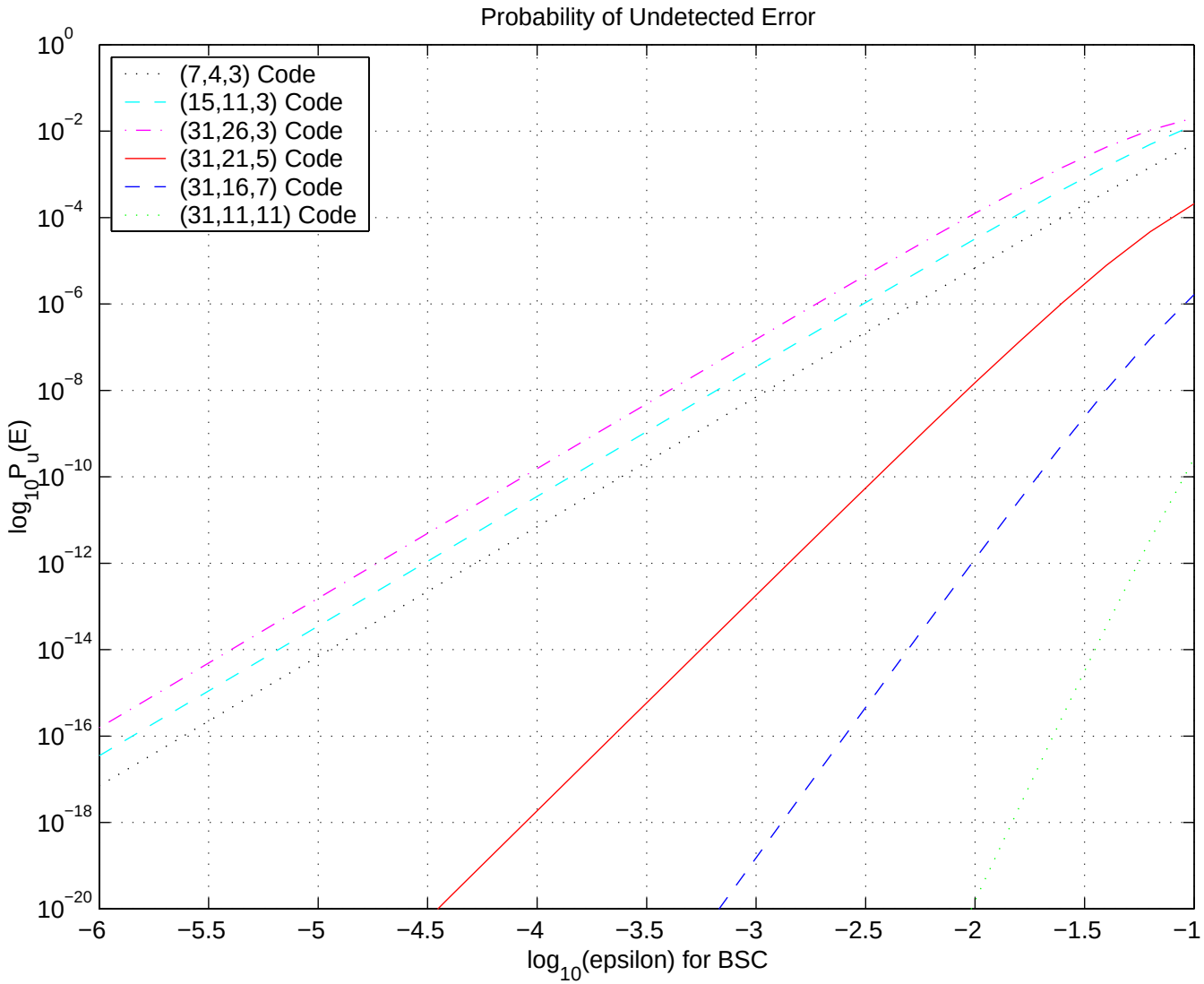
- $(31, 16, 7)$ BCH code:

$$\{A_0 = 1, A_7 = 155, A_8 = 465, A_{11} = 5208, A_{12} = 8680, A_{15} = 18259, A_{16} = 18259, A_{19} = 8680, A_{20} = 5208, A_{23} = 465, A_{24} = 155, A_{31} = 1\} .$$

- $(31, 11, 11)$ BCH code:

$$\{A_0 = 1, A_{11} = 186, A_{12} = 310, A_{15} = 527, A_{16} = 527, A_{19} = 310, A_{20} = 186, A_{31} = 1\} .$$

The (exact) probability of undetected error for these six codes, when used on a memoryless BSC with transition probability ϵ, is shown in the following graph.



The upper three curves show that if $d_{\min}$ is fixed and the blocklength n is increased, then the probability of undetected error $P_u(\mathcal{E})$ increases. The lower three curves show that if n is fixed and $d_{\min}$ is increased, then $P_u(\mathcal{E})$ decreases quite rapidly. But note that the code rate $R = k/n$, which does not show up explicitly in the above graph, also needs to be taken into account when designing a system that uses error control. A lower code rate means in general that either data can be sent less rapidly over a given channel, or the channel bandwidth needs to be increased.

Definition: Let $\{A_0, A_1, \ldots, A_n\}$ be the weight distribution of a linear (n, k) code $\mathcal{C}$ and let $\{B_0, B_1, \ldots, B_n\}$ be the weight distribution of its dual code $\mathcal{C}^\perp$. Define the *weight distribution polynomials* of $\mathcal{C}$ and $\mathcal{C}^\perp$ as

$$A(z) = \sum_{w=0}^{n} A_w \, z^w \quad \text{and} \quad B(z) = \sum_{w=0}^{n} B_w \, z^w \, .$$

Example: Code #1 with weight distribution $\{A_0{=}1, A_3{=}2, A_4{=}1\}$ has weight distribution polynomial $A(z) = 1 + 2\,z^3 + z^4$.

Theorem: **MacWilliams Identity for Binary Codes.** The weight distribution polynomials, $A(z)$ and $B(z)$, of a binary linear (n, k) code $\mathcal{C}$ and its dual code $\mathcal{C}^\perp$ are related by

$$A(z) = 2^{-(n-k)}\,(1+z)^n\,B\left(\frac{1-z}{1+z}\right)\,.$$

Theorem: **MacWilliams Identity for q-ary Codes.** The weight distribution polynomials, $A(z)$ and $B(z)$, of a q-ary linear (n, k) code $\mathcal{C}$ and its dual code $\mathcal{C}^\perp$ are related by

$$A(z) = q^{-(n-k)}\,(1+(q-1)z)^n\,B\left(\frac{1-z}{1+(q-1)z}\right)\,.$$

Proof: See F.J. MacWilliams, "A Theorem on the Distribution of Weights in a Systematic Code," *Bell Syst. Tech. J.*, vol. 42, 1963, pp. 79–94.

Example: Consider the binary $(5, 3)$ code with generator matrix

$$\mathbf{G} = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{bmatrix} .$$

Its dual code is code #1, the $(5, 2)$ code with $B(z) = 1 + 2\,z^3 + z^4$ given in a previous example. Using the MacWilliams identity, one finds that

$$\begin{aligned} A(z) &= 2^{-2}\,(1+z)^5 \left(1 + 2\left(\tfrac{1-z}{1+z}\right)^3 + \left(\tfrac{1-z}{1+z}\right)^4\right) \\ &= \tfrac{1}{4}\left((1+z)^5 + 2\,(1-z)^3\,(1+z)^2 + (1-z)^4\,(1+z)\right) \\ &= 1 + 2\,z^2 + 4\,z^3 + z^4 \,, \end{aligned}$$

and thus the weight distribution of this $(5, 3)$ code is

$$\{A_0 = 1,\ A_2 = 2,\ A_3 = 4,\ A_4 = 1,\ \ A_i = 0,\ i \neq 0, 2, 3, 4\} .$$

Theorem: In terms of the weight distribution polynomials $A(z)$ (for code $\mathcal{C}$) and $B(z)$ (for code $\mathcal{C}^{\perp}$), the probability of undetected error for a linear binary (n, k) code $\mathcal{C}$ used on a BSC with transition probability ϵ, can be expressed as

$$P_u(\mathcal{E}) = (1-\epsilon)^n \left(A\left(\frac{\epsilon}{1-\epsilon}\right) - 1 \right) = 2^{-(n-k)} B(1-2\epsilon) - (1-\epsilon)^n .$$

Proof: For the first equality note that

$$P_u(\mathcal{E}) = \sum_{w=d_{\min}}^{n} A_w \epsilon^w (1-\epsilon)^{n-w} = (1-\epsilon)^n \sum_{w=d_{\min}}^{n} A_w \left(\frac{\epsilon}{1-\epsilon}\right)^w = (1-\epsilon)^n \left(A\left(\frac{\epsilon}{1-\epsilon}\right) - 1 \right).$$

For the second equality use MacWilliams' identity

$$A(z) = 2^{-(n-k)} (1+z)^n B\left(\frac{1-z}{1+z}\right),$$

for binary (n, k) codes.

Probability of Decoding Error

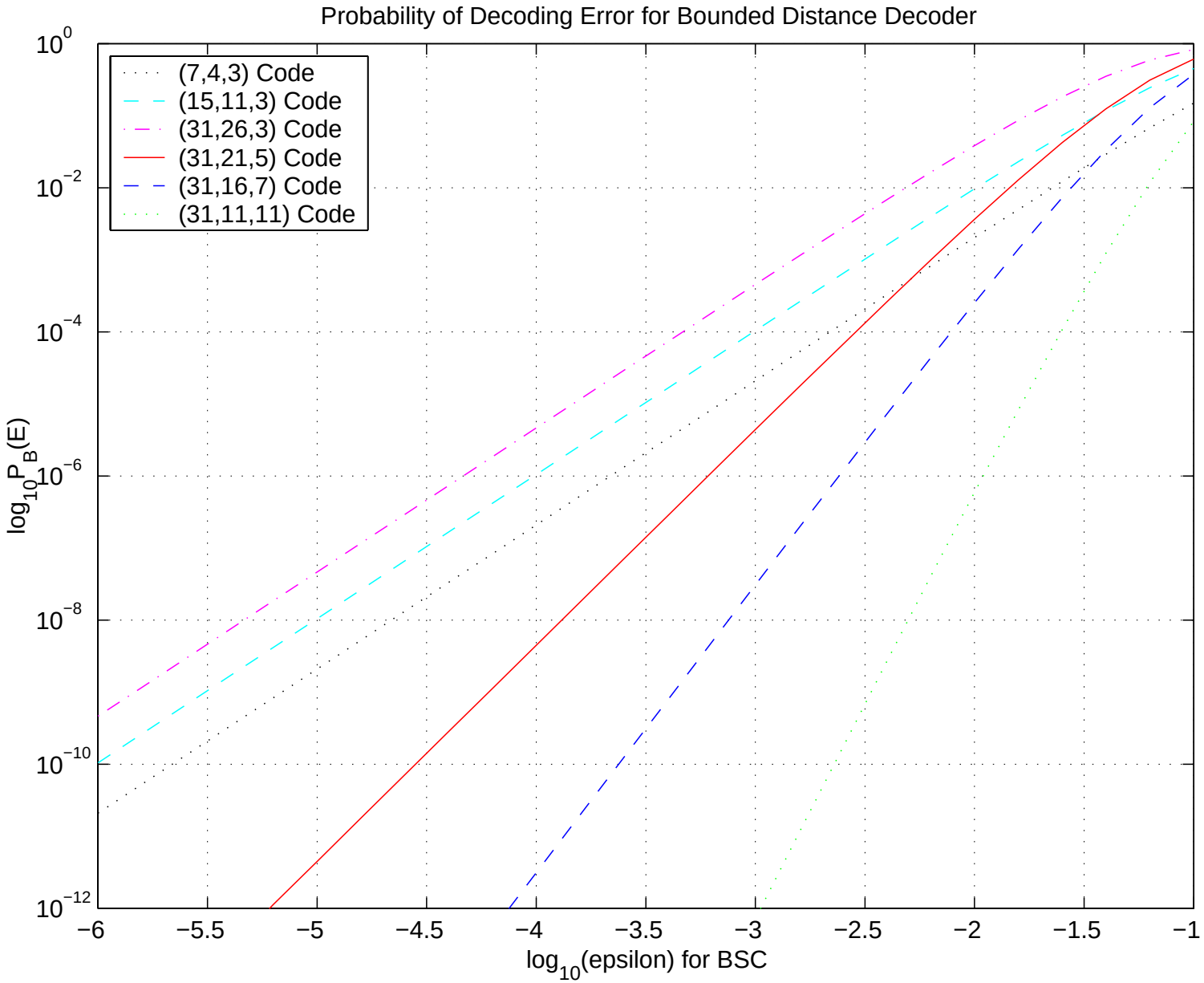
Theorem: Let $\mathcal{C}$ be any binary code with minimum (Hamming) distance $d_{\min}$ and let $t = \lfloor (d_{\min} - 1)/2 \rfloor$. Then the probability of a (block) decoding error, $P_B(\mathcal{E})$, on a memoryless BSC with transition probability ϵ is upper bounded by

$$P_B(\mathcal{E}) \leq \sum_{w=t+1}^{n} \binom{n}{w} \epsilon^w (1-\epsilon)^{n-w} = 1 - \underbrace{\sum_{w=0}^{t} \binom{n}{w} \epsilon^w (1-\epsilon)^{n-w}}_{= P_B(\mathcal{C})}.$$

Proof: Any code with minimum distance $d_{\min}$ can correct all error patterns of $t = \lfloor (d_{\min} - 1)/2 \rfloor$ or less (random) errors. QED

Example: Consider bounded distance decoding of all error patterns up to t errors, $t = \lfloor (d_{\min} - 1)/2 \rfloor$, using binary codes with the following parameters $(7, 4, 3)$, $(15, 11, 3)$, $(31, 26, 3)$, $(31, 21, 5)$, $(31, 16, 7)$, and $(31, 11, 11)$. Upper bounds on the probability of decoding error $P_B(\mathcal{E})$ when using a BSC with transition probability ϵ are given in the graph on the next slide. Note that the first three codes are Hamming codes and since these are perfect codes, the $P_B(\mathcal{E})$ curves are exact rather than bounds.

As was the case for error detection, the upper three curves show that if $d_{\min}$ is kept fixed and n is increased, then the probability of error is increased. Conversely, if the blocklength n is kept fixed and $d_{\min}$ is increased, then $P_B(\mathcal{E})$ decreases quite rapidly as can be seen from the lower three curves.


Probability of Decoding Error for Bounded Distance Decoder
(7,4,3) Code
(15,11,3) Code
(31,26,3) Code
(31,21,5) Code
(31,16,7) Code
(31,11,11) Code
10^{0}
10^{-2}
10^{-4}
10^{-6}
10^{-8}
10^{-10}
10^{-12}
$\log_{10} P_B(E)$
−6
−5.5
−5
−4.5
−4
−3.5
−3
−2.5
−2
−1.5
−1
$\log_{10}$(epsilon) for BSC

Note: In most cases a t-error-correcting code can also correct many patterns of $t+1$ or even more errors. A q-ary linear (n,k) code, for example, can correct $q^{n-k}-1$ error patterns (including those with t or less errors).

The bound on the probability of decoding error given above assumes a bounded distance decoder that only corrects up to t errors. If a complete decoder is used, one should in most cases (except for perfect codes) be able to do better. Let $\mathbf{X}$ and $\mathbf{Y}$ be random vectors of length n. For any given channel model, let $\mathbf{X}$ denote the channel input and $\mathbf{Y}$ denote the channel output. A particular channel model is specified by giving the conditional probability mass function (pmf) $p_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x})$. Assume that the code $\mathcal{C} = \{\mathbf{c}_i\}$, $i = 0, 1, \ldots, M-1$, is used and the transmission of a codeword from $\mathcal{C}$ over the specified channel results in the received n-tuple $\mathbf{v}$.

The decoder then needs to look at the *a posteriori* probabilities

$$p_{\mathbf{X}|\mathbf{Y}}(\mathbf{c}_i|\mathbf{v}) = \frac{p_{\mathbf{XY}}(\mathbf{c}_i\mathbf{v})}{p_{\mathbf{Y}}(\mathbf{v})} = \frac{p_{\mathbf{Y}|\mathbf{X}}(\mathbf{v}|\mathbf{c}_i)\, p_{\mathbf{X}}(\mathbf{c}_i)}{p_{\mathbf{Y}}(\mathbf{v})} ,$$

for $i = 0, 1, \ldots, M-1$.

Definition: For a given code $\mathcal{C} = \{\mathbf{c}_0, \mathbf{c}_1, \ldots, \mathbf{c}_{M-1}\}$, a priori pmf $p_{\mathbf{X}}(\mathbf{c})$, channel model $p_{\mathbf{Y}|\mathbf{X}}(\mathbf{v}|\mathbf{c})$ and received n-tuple $\mathbf{v}$, a *maximum a posteriori (MAP) decoder* outputs estimate $\hat{\mathbf{c}} = \mathbf{c}_i$ iff i is the index which maximizes the expression

$$p_{\mathbf{Y}|\mathbf{X}}(\mathbf{v}|\mathbf{c}_i)\, p_{\mathbf{X}}(\mathbf{c}_i) .$$

If there is more than one index that maximizes this expression, one of the maximizing indexes can be chosen at random without loss of optimality.

Definition: For a given code $\mathcal{C} = \{\mathbf{c}_0, \mathbf{c}_1, \ldots, \mathbf{c}_{M-1}\}$, channel model $p_{\mathbf{Y}|\mathbf{X}}(\mathbf{v}|\mathbf{c})$ and received n-tuple $\mathbf{v}$, a *maximum likelihood (ML) decoder* outputs estimate $\hat{\mathbf{c}} = \mathbf{c}_i$ iff i is the index which maximizes the expression

$$p_{\mathbf{Y}|\mathbf{X}}(\mathbf{v}|\mathbf{c}_i) \, .$$

If there is more than one index that maximizes this expression, one of the maximizing indexes can be chosen at random without loss of optimality.

Note: If an additive error model is used, then the conditional probability $p_{\mathbf{Y}|\mathbf{X}}(\mathbf{v}|\mathbf{c})$ is equal to the probability $p_{\mathbf{E}}(\mathbf{v}-\mathbf{c})$ that the error pattern $\mathbf{e}=\mathbf{v}-\mathbf{c}$ occurs. If the channel model is such that lower (Hamming) weight error patterns are more likely to occur than higher weight patterns, then the ML decoder outputs $\hat{\mathbf{c}}=\mathbf{c}_i$ iff i is the index that minimizes the (Hamming) distance $d_H(\mathbf{v},\mathbf{c}_i)$. If there are several indexes that minimize this distance, then one of these indexes can be chosen at random without loss of optimality.

Note: If $p_{\mathbf{X}}(\mathbf{c}_i)=1/M$ for all $i=0,1,\ldots,M-1$ (uniform distribution of codewords), then the MAP decoder and the ML decoder are the same.

For ML decoders an upper bound for the block error probability $P_B(\mathcal{E})$ can be derived as follows. Define

$P_d(\mathcal{E}) = P\{\text{Decoding error between two codewords distance } d \text{ apart}\}.$

Then, assuming a linear code with M equally likely codewords, one can use a union bound (i.e., the probability of a union of events is upper bounded by the sum of the probabilities of the events) to write

$$P_B(\mathcal{E}) \leq \sum_{m=1}^{M-1} P_{w_m}(\mathcal{E}) \,,$$

where w_m, $m = 1, 2, \ldots, M-1$, are the weights of all nonzero codewords of the code.

Using the weight distribution $\{A_w\}$ of the code one can sum over all weights and multiply by the number of codewords with each weight, rather than sum over all codewords individually, and thus

$$P_B(\mathcal{E}) \leq \sum_{w=d_{\min}}^{n} A_w \, P_w(\mathcal{E}) \, .$$

The probability $P_d(\mathcal{E})$ of a decoding error between two codewords distance d apart is the probability that an error pattern occurs with $d/2$ or more errors in the d positions in which the two codewords differ. To be precise in the case when d is an even integer, only one half of the cases when exactly $d/2$ errors occur causes an error on the average. Thus, one obtains the following upper bound (or exact expression when d is an odd integer) for $P_d(\mathcal{E})$

$$P_d(\mathcal{E}) \leq \sum_{e=\lceil d/2 \rceil}^{d} \binom{d}{e} \epsilon^e \, (1-\epsilon)^{d-e} \, .$$

Substituting this in the union bound above the following theorem is proved.

Theorem: Let $\mathcal{C}$ be a binary block code with blocklength n, minimum (Hamming) distance $d_{\min}$, and weight distribution $\{A_w\}$. Then, using a ML decoder, the *probability of a (block) decoding error* $P_B(\mathcal{E})$ on a memoryless BSC with transition probability ϵ is upper bounded by

$$P_B(\mathcal{E}) \leq \sum_{w=d_{\min}}^{n} A_w \sum_{e=\lceil w/2 \rceil}^{w} \binom{w}{e} \epsilon^e (1-\epsilon)^{w-e} .$$

A somewhat weaker bound, based on the so called Bhattacharyya distance between two codewords, which is

$$d_B = -0.5 \log(4\epsilon(1-\epsilon))$$

for the BSC with transition probability ϵ, is given in the following theorem.

Theorem: Let $\mathcal{C}$ be a binary block code with blocklength n, minimum (Hamming) distance $d_{\min}$, and weight distribution $\{A_w\}$. Then, using a ML decoder, the *probability of a (block) decoding error* $P_B(\mathcal{E})$ on a memoryless BSC with transition probability ϵ is upper bounded by

$$P_B(\mathcal{E}) \leq \frac{1}{2} \sum_{w=d_{\min}}^{n} A_w \left[4\epsilon\,(1-\epsilon)\right]^{w/2}.$$

Proof: The probability $P_d(\mathcal{E})$ of a decoding error between two codewords (Hamming) distance d apart can be upper bounded as follows

$$P_d(\mathcal{E}) \leq \sum_{e=\lceil d/2 \rceil}^{d} \binom{d}{e} \epsilon^e (1-\epsilon)^{d-e} = (1-\epsilon)^d \sum_{e=\lceil d/2 \rceil}^{d} \binom{d}{e} \left(\frac{\epsilon}{1-\epsilon}\right)^e$$

$$\leq (1-\epsilon)^d \left(\frac{\epsilon}{1-\epsilon}\right)^{d/2} \sum_{e=\lceil d/2 \rceil}^{d} \binom{d}{e} = \epsilon^{d/2} (1-\epsilon)^{d/2} \underbrace{\sum_{e=\lceil d/2 \rceil}^{d} \binom{d}{e}}_{\leq 2^d/2}$$

$$\leq \frac{1}{2} 2^d \epsilon^{d/2} (1-\epsilon)^{d/2} = \frac{1}{2} \left[4\epsilon(1-\epsilon)\right]^{d/2}.$$

Substituting this in the union bound for $P_B(\mathcal{E})$ completes the proof of the theorem. QED

Suppose now that it is desired to compare the "goodness" of two codes with different k. The probability of a block error is then not really the best measure. Rather, one should compare the *probability of a data symbol error* $P_s(\mathcal{E})$ or, in the binary case, the *probability of a data bit error* $P_b(\mathcal{E})$. Because a block error will affect between 1 and k data symbols, the probability of a symbol error can be bounded as

$$\frac{1}{k} P_B(\mathcal{E}) \leq P_s(\mathcal{E}) \leq P_B(\mathcal{E}) ,$$

but for large k there is quite a gap between the upper and the lower bounds. Another approach is to extend the concept of the weight distribution of a code so that it also contains information about the weight of the data symbols.

Definition: Let $A(w, i)$ be the number of codewords of a linear (n,k) code with total weight w and data weight i (both using Hamming weight). Then the set $\{A(w, i)\}$, $w = 0, 1, \ldots, n$, $i = 0, 1, \ldots, k$, is called the *extended weight distribution* of the code.

If $\{A(w, i)\}$ is known, then it is easy to obtain $\{A_w\}$ using

$$A_w = \sum_{i=0}^{k} A(w, i) \,, \qquad w = 0, 1, \ldots, n \,.$$

Substituting this expression in the union bound for the probability of a block error yields

$$P_B(\mathcal{E}) \leq \sum_{w=d_{\min}}^{n} \sum_{i=0}^{k} A(w, i)\, P_w(\mathcal{E}) \,.$$

To obtain an expression for the probability of a symbol error, multiply $A(w, i)$ by the number i of data symbol errors in the formula above and then divide by k to obtain

$$P_s(\mathcal{E}) \leq \frac{1}{k} \sum_{w=d_{\min}}^{n} \sum_{i=1}^{k} i\, A(w, i)\, P_w(\mathcal{E}) \,.$$

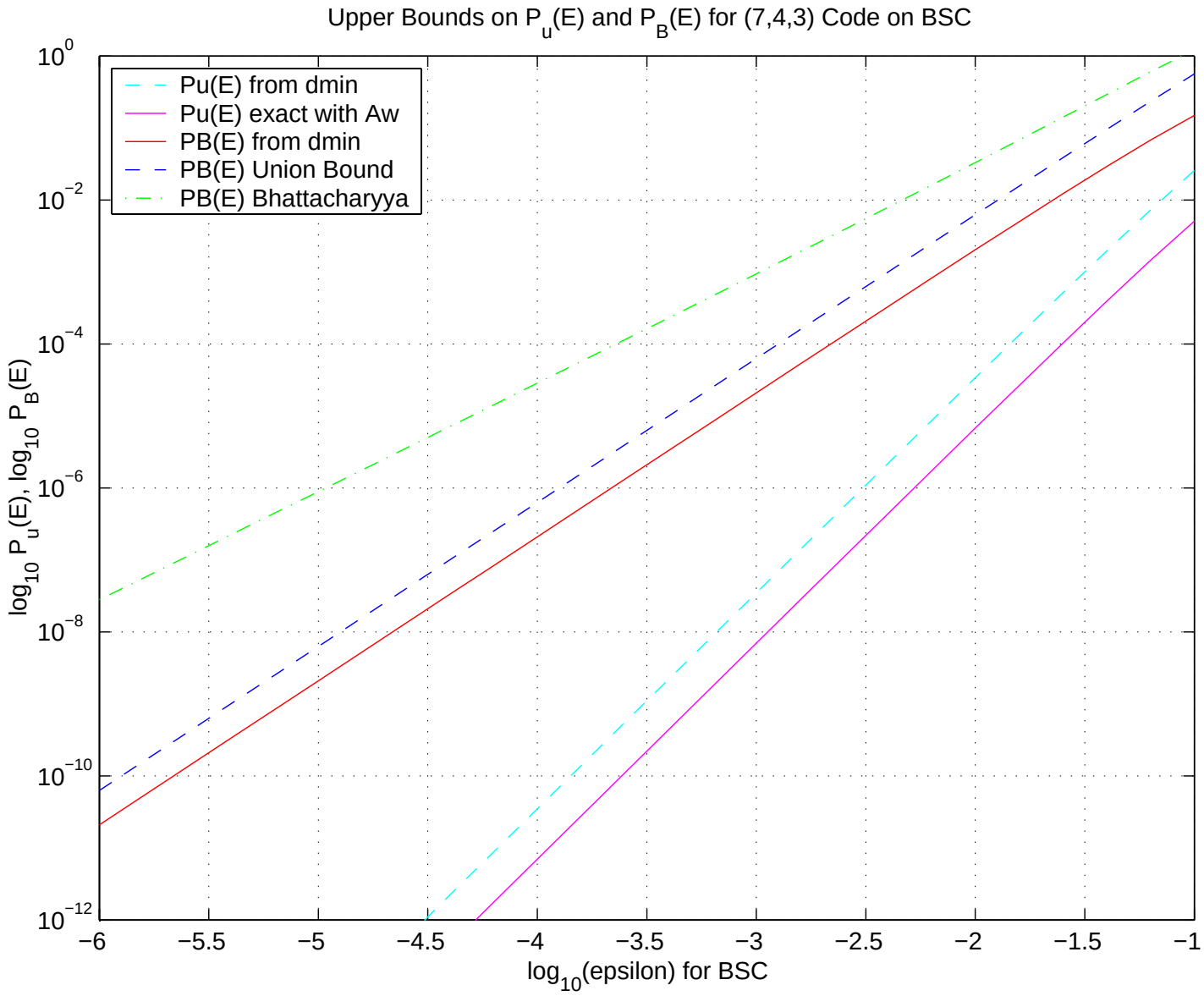
Substituting the expression for $P_w(\mathcal{E})$ on a BSC with transition probability ϵ thus proves the following theorem.

Theorem: Let $\mathcal{C}$ be a binary $(n, k, d_{\min})$ block code with extended weight distribution $\{A(w, i)\}$. Using a ML decoder, the *probability of a bit error* $P_b(\mathcal{E})$ on a memoryless BSC with transition probability ϵ is then upper bounded by

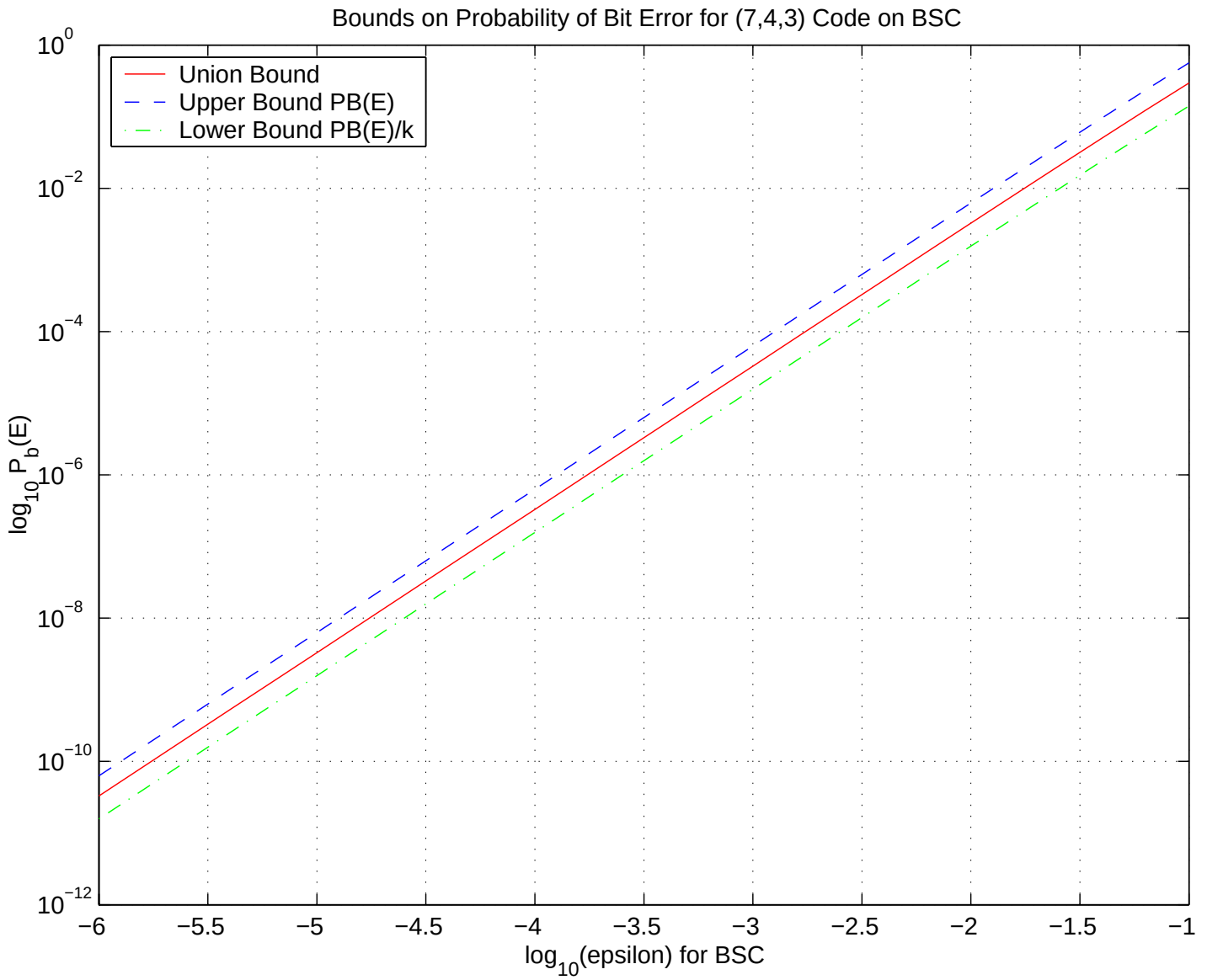
$$P_b(\mathcal{E}) \leq \frac{1}{k} \sum_{w=d_{\min}}^{n} \sum_{i=1}^{k} i\, A(w, i) \sum_{e=\lceil w/2 \rceil}^{w} \binom{w}{e} \epsilon^e (1-\epsilon)^{w-e} \,.$$

Example: Bounds on the probability of block error for a binary $(7, 4, 3)$ Hamming code on a memoryless BSC with transition probability ϵ are shown in the graph on the next page.

Clearly, the upper bound on $P_u(\mathcal{E})$ based on $d_{\min}$ is quite a bit weaker than the exact value which uses the weight distribution $\{A_w\}$ of the code. Because Hamming codes are perfect codes, the value of the block error probability $P_B(\mathcal{E})$ based on $d_{\min}$ is exact for both bounded distance and ML decoders. Note that the simpler bound for $P_B(\mathcal{E})$ of a ML decoder which uses the Bhattacharyya distance is quite loose, mostly because the $(7, 4, 3)$ code has both even and odd weight codewords.

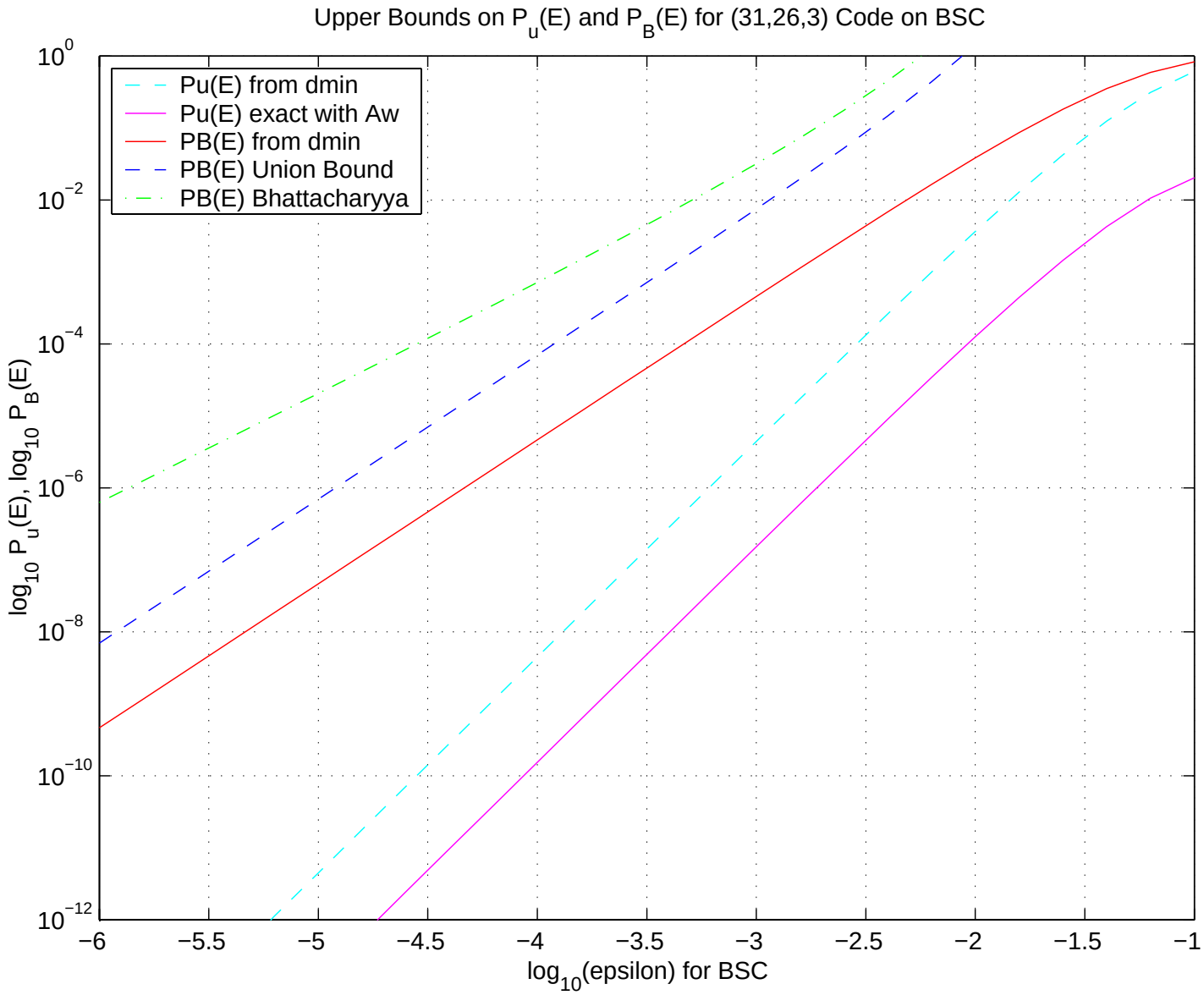

Upper Bounds on P_u(E) and P_B(E) for (7,4,3) Code on BSC
Pu(E) from dmin
Pu(E) exact with Aw
PB(E) from dmin
PB(E) Union Bound
PB(E) Bhattacharyya
log10 Pu(E), log10 PB(E)
log10(epsilon) for BSC
10^0
10^-2
10^-4
10^-6
10^-8
10^-10
10^-12
−6
−5.5
−5
−4.5
−4
−3.5
−3
−2.5
−2
−1.5
−1

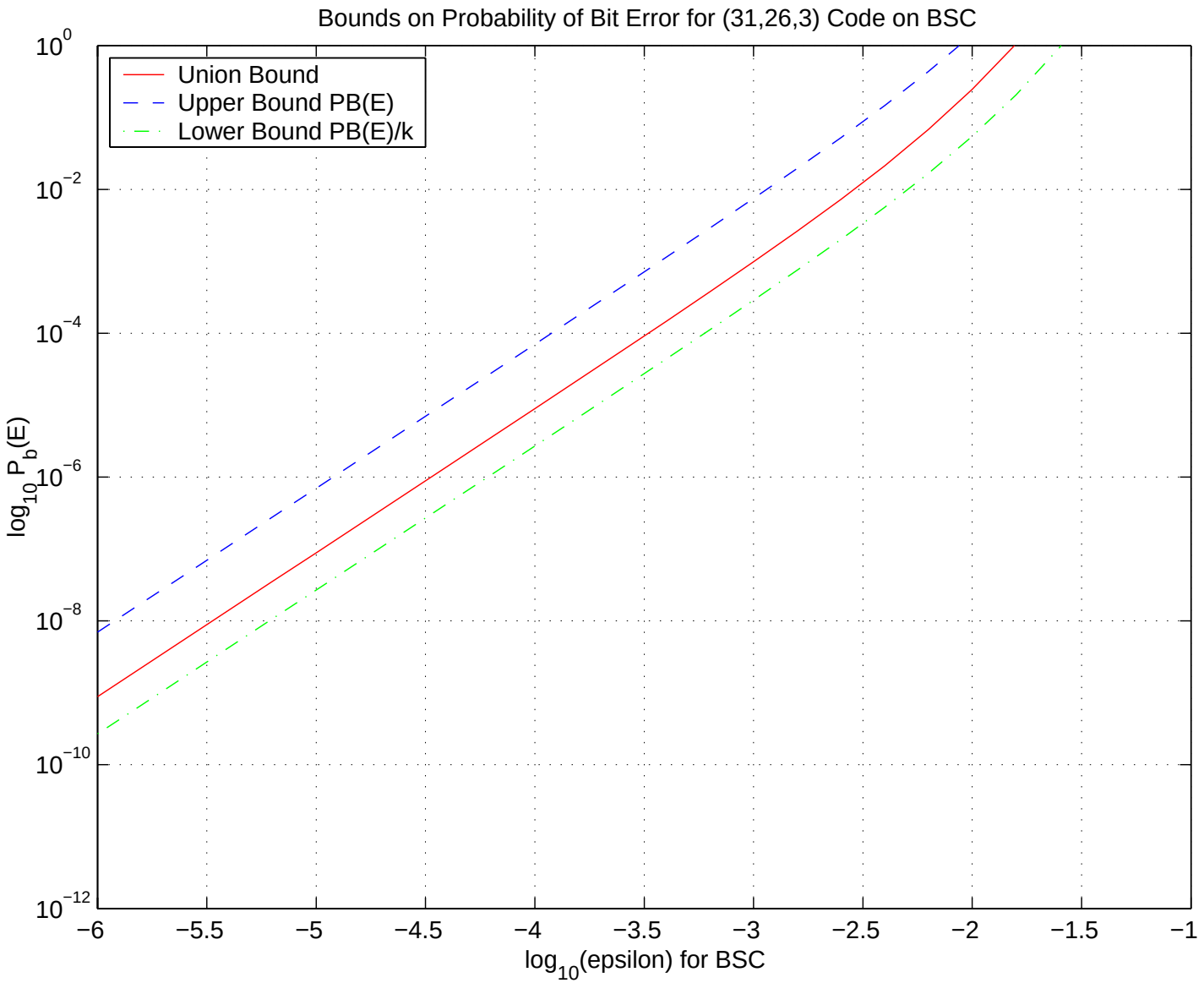
Bounds on the probability of bit error are shown below.



The upper and lower bounds, which use the more complicated but also more tight bound on $P_B(\mathcal{E})$ for ML decoding, are not too far apart in this case because k is rather small.

Example: Bounds for the block and bit error probabilities of a binary $(31, 26, 3)$ Hamming code on a memoryless BSC with transition probability ϵ are shown in the next two figures. Similar comments as for the $(7, 4, 3)$ code apply. Generally, the bounds that were loose already for the $(7, 4, 3)$ code become even more loose as the blocklength n is increased (and $d_{\min}$ is kept fixed). In part this is due to the fact that Hamming codes are perfect codes, and in part it comes from the increase in the data length k.

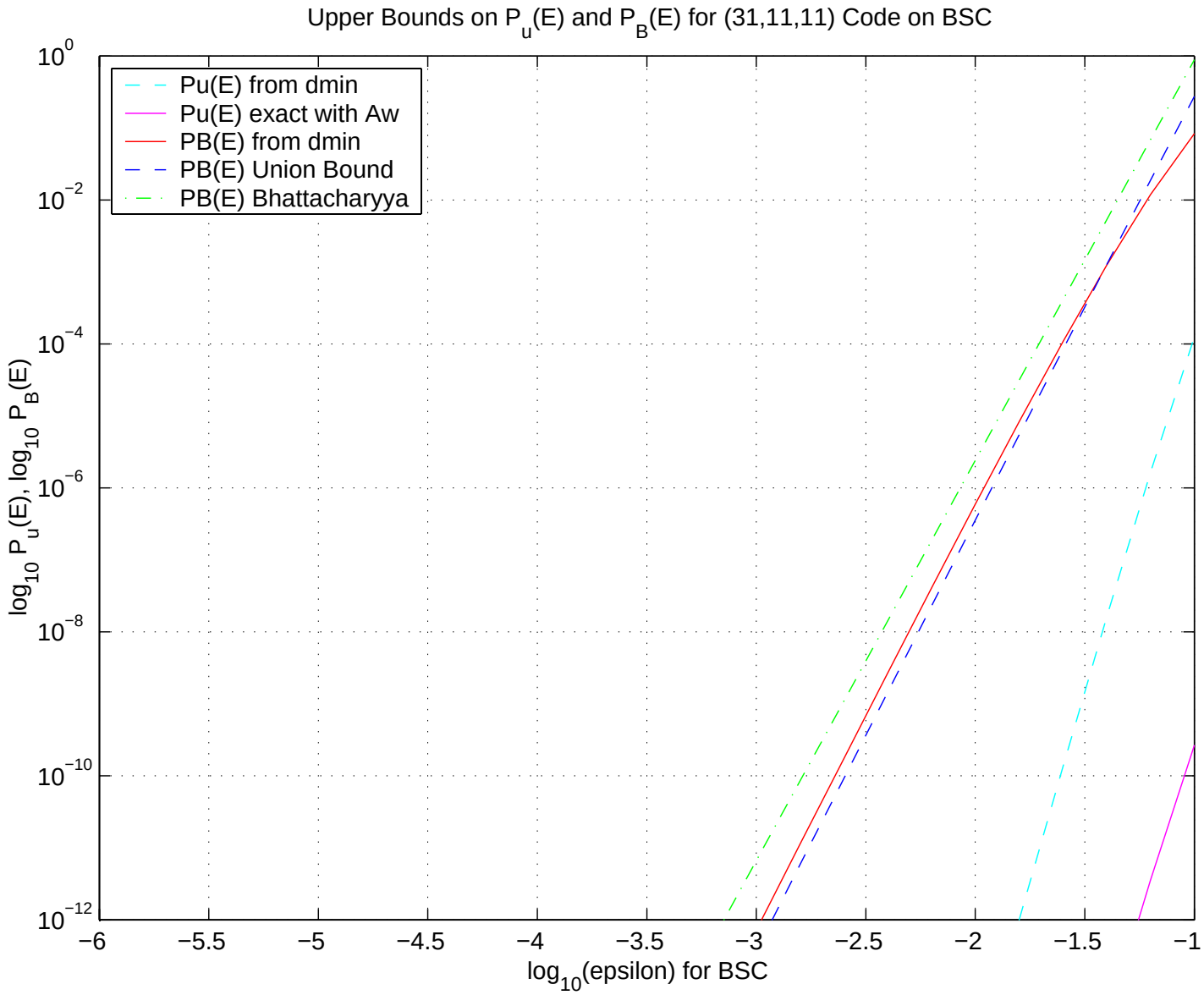

Upper Bounds on P_u(E) and P_B(E) for (31,26,3) Code on BSC
Pu(E) from dmin
Pu(E) exact with Aw
PB(E) from dmin
PB(E) Union Bound
PB(E) Bhattacharyya
log10 Pu(E), log10 PB(E)
log10(epsilon) for BSC
10^0
10^-2
10^-4
10^-6
10^-8
10^-10
10^-12
−6
−5.5
−5
−4.5
−4
−3.5
−3
−2.5
−2
−1.5
−1

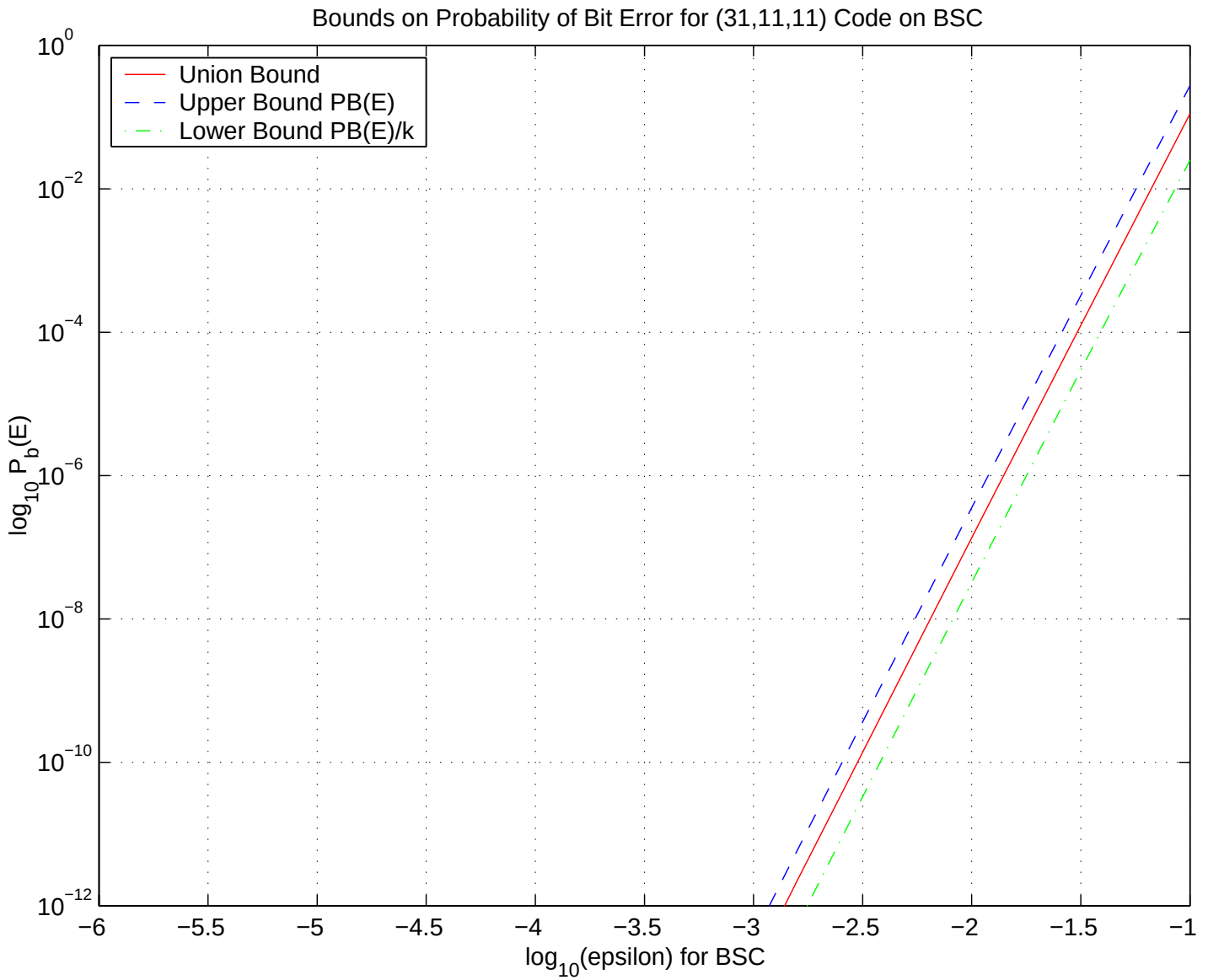

Bounds on Probability of Bit Error for (31,26,3) Code on BSC
Union Bound
Upper Bound PB(E)
Lower Bound PB(E)/k
10^{0}
10^{-2}
10^{-4}
10^{-6}
10^{-8}
10^{-10}
10^{-12}
$\log_{10} P_b(E)$
−6
−5.5
−5
−4.5
−4
−3.5
−3
−2.5
−2
−1.5
−1
$\log_{10}$(epsilon) for BSC

Example: Bounds for the block and bit error probabilities of a binary $(31, 11, 11)$ BCH code on a memoryless BSC with transition probability ϵ are shown in the next two graphs.

Because of the larger $d_{\min}$, the error probabilities at the decoder output now decrease much more rapidly as ϵ decreases. More interestingly, the bound on $P_B(\mathcal{E})$ for a ML decoder is now below the bound for bounded distance decoding in the main region of interest (i.e., $P_B(\mathcal{E}) < 10^{-4}$). This is a direct consequence of the fact that the $(31, 11, 11)$ code is not a perfect code. Note that the bound on $P_u(\mathcal{E})$ based on $d_{\min}$ is quite loose and essentially useless.

The upper and lower bounds on the probability of bit error, which are based on $P_B(\mathcal{E})$ for a ML decoder, are not too far apart because k is relatively small.


Upper Bounds on P_u(E) and P_B(E) for (31,11,11) Code on BSC
Pu(E) from dmin
Pu(E) exact with Aw
PB(E) from dmin
PB(E) Union Bound
PB(E) Bhattacharyya
log10 Pu(E), log10 PB(E)
log10(epsilon) for BSC
10^0
10^-2
10^-4
10^-6
10^-8
10^-10
10^-12
−6
−5.5
−5
−4.5
−4
−3.5
−3
−2.5
−2
−1.5
−1


Bounds on Probability of Bit Error for (31,11,11) Code on BSC
Union Bound
Upper Bound PB(E)
Lower Bound PB(E)/k
10^{0}
10^{-2}
10^{-4}
10^{-6}
10^{-8}
10^{-10}
10^{-12}
$\log_{10} P_b(E)$
−6
−5.5
−5
−4.5
−4
−3.5
−3
−2.5
−2
−1.5
−1
$\log_{10}$(epsilon) for BSC

The bounds considered up to now were based on the BSC channel model and thus assumed that hard decisions are made by the receiver before decoding. To get an idea of the performance improvements that are possible if the receiver uses **soft decisions**, consider binary antipodal signaling with message bits $m_0 \rightarrow -\sqrt{E_b}$ and $m_1 \rightarrow +\sqrt{E_b}$ over an additive white Gaussian noise (AWGN) channel with two-sided noise power spectral density (PSD) $\mathcal{N}_0/2$. Assume that the receiver uses a matched filter (MF) and that both E_b and $\mathcal{N}_0$ are measured at the output of the MF at the optimum sampling time instant at which the SNR $E_b/\mathcal{N}_0$ is maximized.

Theorem: The probability ϵ of an uncoded bit error for binary antipodal signaling over an AWGN channel with MF receiver and decision rule $\hat{m} = m_0$ iff the MF output $b < 0$ (at the optimum sampling time instant) is given by

$$\epsilon = \frac{1}{2}\text{erfc}\left(\sqrt{\frac{E_b}{\mathcal{N}_0}}\right), \qquad \text{where} \quad \text{erfc}(x) = \frac{2}{\sqrt{\pi}}\int_x^\infty e^{-\mu^2}\, d\mu\,.$$

Proof: The output b of the matched filter at the optimum sampling time instant (which maximizes the SNR $E_b/\mathcal{N}_0$) is Gaussian with variance $\sigma_b^2 = \mathcal{N}_0/2$ and mean either $-\sqrt{E_b}$ or $+\sqrt{E_b}$, i.e.,

$$f_b(\beta|m_0) = \frac{e^{-(\beta+\sqrt{E_b})^2/\mathcal{N}_0}}{\sqrt{\pi\mathcal{N}_0}}, \qquad \text{or} \qquad f_b(\beta|m_1) = \frac{e^{-(\beta-\sqrt{E_b})^2/\mathcal{N}_0}}{\sqrt{\pi\mathcal{N}_0}}.$$

Setting the decision threshold at $\beta = 0$ yields

$$P(\mathcal{E}|m_0) = \frac{1}{\sqrt{\pi\mathcal{N}_0}} \int_0^\infty e^{-(\beta+\sqrt{E_b})^2/\mathcal{N}_0}\, d\beta = \frac{1}{\sqrt{\pi}} \int_{\sqrt{E_b/\mathcal{N}_0}}^\infty e^{-\mu^2}\, d\mu = \frac{1}{2}\text{erfc}\left(\sqrt{\frac{E_b}{\mathcal{N}_0}}\right).$$

A similar computation yields $P(\mathcal{E}|m_1) = P(\mathcal{E}|m_0)$ and thus

$$\epsilon = P(\mathcal{E}|m_0)\,P(m_0) + P(\mathcal{E}|m_1)\,P(m_1) = \frac{1}{2}\text{erfc}\left(\sqrt{\frac{E_b}{\mathcal{N}_0}}\right).$$

QED

Now consider the case where message m_i is encoded into codeword $\mathbf{c}_i = (c_{i0}, c_{i1}, \ldots, c_{i\,n-1})$ and then transmitted using antipodal signaling such that the receiver (after sampling at the output of the MF) sees

$$m_i \;\rightarrow\; \mathbf{c}_i \quad\longrightarrow\quad (2c_{i0}-1, 2c_{i1}-1, \ldots, 2c_{i\,n-1}-1)\sqrt{E_b}\,.$$

Thus, if the receiver has to distinguish between $m_i \rightarrow \mathbf{c}_i$ and $m_j \rightarrow \mathbf{c}_j$ and $d(\mathbf{c}_i, \mathbf{c}_j) = d$, then this is equivalent to d uncoded decoding decisions (assuming a memoryless channel). This proves the following

Theorem: The probability $P_d(\mathcal{E})$ of a decoding error between two codewords (Hamming) distance d apart for binary antipodal signaling over an AWGN channel with MF receiver and ML decision rule is given by

$$P_d(\mathcal{E}) = \frac{1}{2}\mathrm{erfc}\left(\sqrt{\frac{d\,E_b}{\mathcal{N}_0}}\right), \qquad \text{where} \quad \mathrm{erfc}(x) = \frac{2}{\sqrt{\pi}}\int_x^\infty e^{-\mu^2}\,d\mu.$$

Combining the union bound with the result for $P_d(\mathcal{E})$ from the previous page results in the following theorem.

Theorem: **Probability of Block Error for Soft Decisions.** Let $\mathcal{C}$ be a linear binary $(n, k, d_{\min})$ block code with weight distribution $\{A_w\}$. Then, using antipodal signaling over an AWGN channel and (soft-decision) ML decoding with a matched filter at the receiver, the *probability of a block decoding error* $P_B(\mathcal{E})$ is upper bounded by

$$P_B(\mathcal{E}) \leq \frac{1}{2} \sum_{w=d_{\min}}^{n} A_w \operatorname{erfc}\left(\sqrt{\frac{w\, E_b}{\mathcal{N}_0}}\right),$$

where E_b is the bit energy of the code bits and $\mathcal{N}_0/2$ is the noise power (or noise variance), both at the output of the MF at the optimum sampling time instant.

Replacing $\{A_w\}$ by $\{A(w,i)\}$ finally yields the following theorem.

Theorem: **Probability of Bit Error for Soft Decisions.** Let $\mathcal{C}$ be a linear binary $(n,k,d_{\min})$ block code with extended weight distribution $\{A(w,i)\}$. Then, using antipodal signaling over an AWGN channel and (soft-decision) ML decoding with a matched filter at the receiver, the *probability of a bit decoding error* $P_b(\mathcal{E})$ is upper bounded by

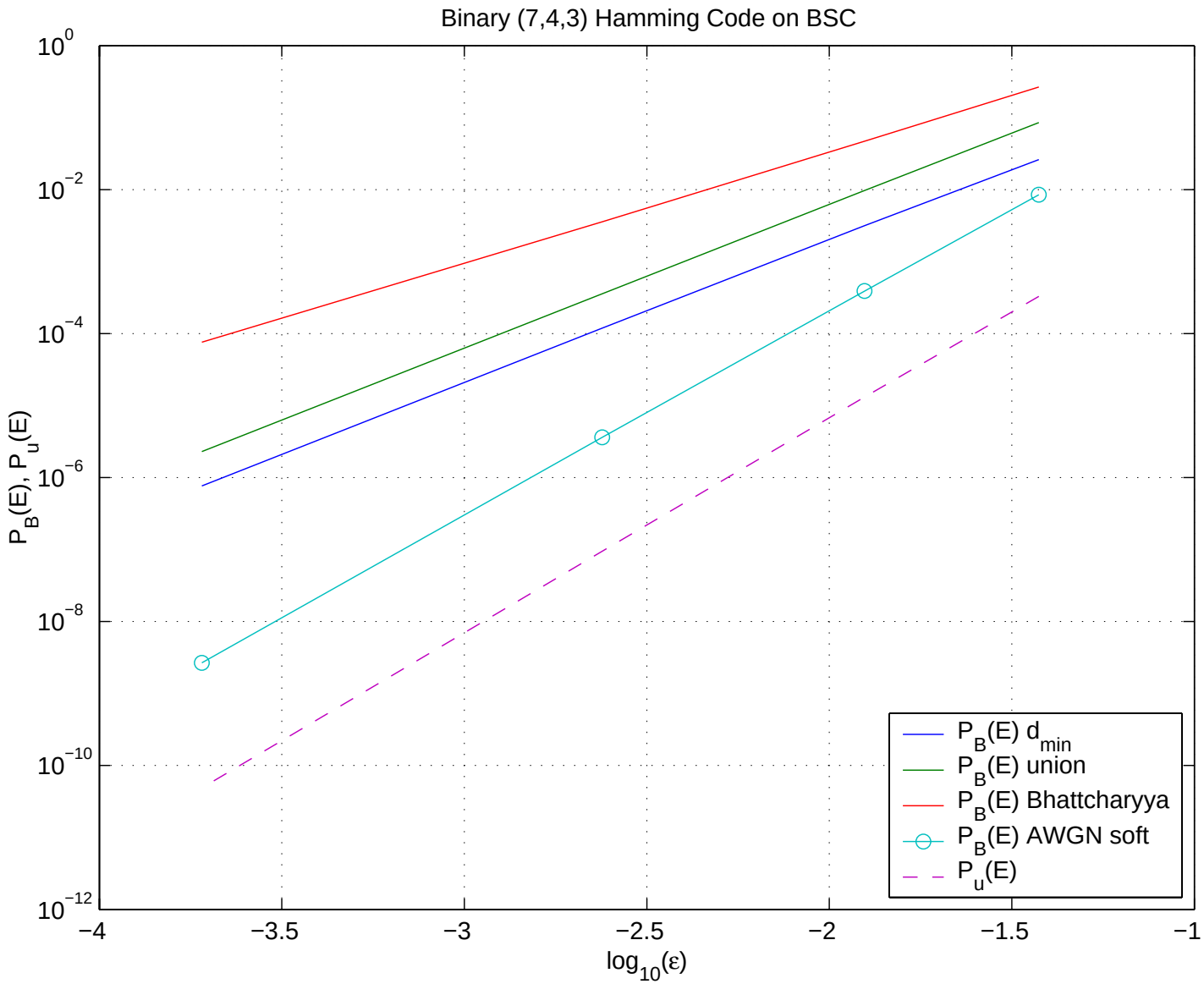
$$P_b(\mathcal{E}) \leq \frac{1}{2k} \sum_{w=d_{\min}}^{n} \operatorname{erfc}\left(\sqrt{\frac{w\,E_b}{\mathcal{N}_0}}\right) \sum_{i=1}^{k} i\,A(w,i)\,,$$

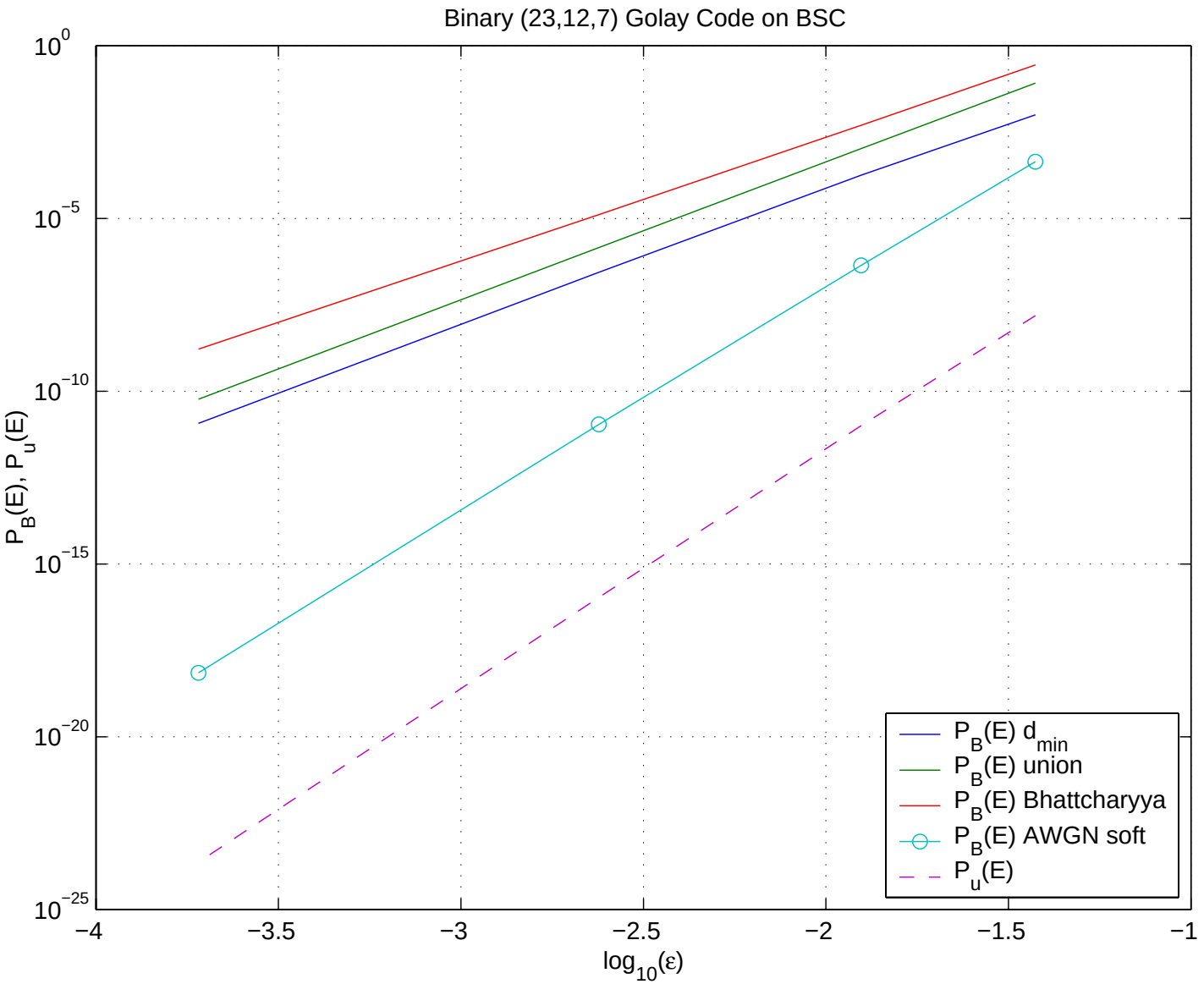
where E_b is the bit energy of the code bits and $\mathcal{N}_0/2$ is the noise power (or noise variance), both at the output of the MF at the optimum sampling time instant.

Example: Bounds on the probabilities of block error for hard and soft decisions of the binary (7,4,3) Hamming code and the binary (23,12,7) Golay code are shown on the following two pages. For the soft decision error probabilities the SNR $E_b/\mathcal{N}_0$ was converted to ϵ of a BSC using

$$\epsilon = \frac{1}{2}\text{erfc}\left(\sqrt{\frac{E_b}{\mathcal{N}_0}}\right).$$

For comparison purposes, the probability of undetected error was also included in the graphs.


Binary (7,4,3) Hamming Code on BSC
10^{0}
10^{-2}
10^{-4}
10^{-6}
10^{-8}
10^{-10}
10^{-12}
$P_B(E), P_u(E)$
−4
−3.5
−3
−2.5
−2
−1.5
−1
$\log_{10}(\varepsilon)$
$P_B(E)$ d_{min}
$P_B(E)$ union
$P_B(E)$ Bhattcharyya
$P_B(E)$ AWGN soft
$P_u(E)$


Binary (23,12,7) Golay Code on BSC
10^{0}
10^{-5}
10^{-10}
10^{-15}
10^{-20}
10^{-25}
$P_B(E), P_u(E)$
−4
−3.5
−3
−2.5
−2
−1.5
−1
$\log_{10}(\varepsilon)$
$P_B(E)$ d_{min}
$P_B(E)$ union
$P_B(E)$ Bhattcharyya
$P_B(E)$ AWGN soft
$P_u(E)$