Notes on Performance Bounds (v1.5)

© *K.M. Chugg October 5, 2015*

1 Basic Bounds for Generic M-ary Decision Rules

Consider an arbitrary deterministic decision rule for deciding among M hypotheses $\mathcal{H}_m$ for $m = 0, 1, \ldots M-1$ based on some measurement $\mathbf{z} \in \mathcal{Z}$. Such a rule introduces a partition of the observation space

$$\text{decide } \mathcal{H}_m \quad \Longleftrightarrow \quad \mathbf{z} \in \mathcal{Z}_m \tag{1}$$

where $\mathcal{Z}_m \bigcap \mathcal{Z}_j = \emptyset$ and $\bigcup_{m=0}^{M-1} \mathcal{Z}_m = \mathcal{X}$. As an example, MAP detection for the M-ary problem results in

$$\mathcal{Z}_m = \{\mathbf{z} : f_{\mathbf{z}(u)}(\mathbf{z}|\mathcal{H}_m)P(\mathcal{H}_m) > f_{\mathbf{z}(u)}(\mathbf{z}|\mathcal{H}_j)P(\mathcal{H}_j) \; \forall \; j \neq m\} \tag{2}$$

The conditional probability of error is given by

$$P(\mathcal{E}|\mathcal{H}_m) = \Pr\{\mathbf{z}(u) \notin \mathcal{Z}_m|\mathcal{H}_m\} = \Pr\{\mathbf{z}(u) \in \mathcal{Z}_m^c|\mathcal{H}_m\} \tag{3}$$

Many useful bounds can be constructed by expressing $\mathcal{Z}_m^c$ in specific ways. For example, it is clear that

$$\mathcal{Z}_m^c = \mathcal{Z} - \mathcal{Z}_m = \bigcup_{j=0, j\neq m}^{M-1} \mathcal{X}_j \tag{4}$$

An expression that is typically more useful is obtained by constructing $\mathcal{Z}_m$ from pairwise decision regions $\mathcal{Z}_m^{PW}(j)$, defined as the region where $\mathcal{H}_m$ would be selected over $\mathcal{H}_j$ in a pairwise (binary) decision. For example, in the case of MAP M-ary decisions,

$$\mathcal{Z}_m^{PW}(j) = \{\mathbf{z} : f_{\mathbf{z}(u)}(\mathbf{z}|\mathcal{H}_m)P(\mathcal{H}_m) > f_{\mathbf{z}(u)}(\mathbf{z}|\mathcal{H}_j)P(\mathcal{H}_j)\} \tag{5}$$

Comparing the definitions of the global decision region and the pairwise regions (i.e., (2) and (5)) it is apparent that

$$\mathcal{Z}_m = \bigcap_{j=0, j\neq m}^{M-1} \mathcal{Z}_m^{PW}(j) \tag{6}$$

The compliment of this region is obtained by applying DeMorgan's Law yielding

$$\mathcal{Z}_m^c = \bigcup_{j=0, j\neq m}^{M-1} \left[\mathcal{Z}_m^{PW}(j)\right]^c = \bigcup_{j=0, j\neq m}^{M-1} \mathcal{Z}_j^{PW}(m) \tag{7}$$

where the fact that $\left[\mathcal{Z}_m^{PW}(j)\right]^c = \mathcal{Z}_j^{PW}(m)$ has been used.

Bounds can be constructed using simple Union Bounds and related techniques. Specifically, let $\{A_i\}$ be a set of events, then it follows that

$$\max_i P(A_i) \leq P(\bigcup_i A_i) \leq \sum_i P(A_i) \tag{8}$$

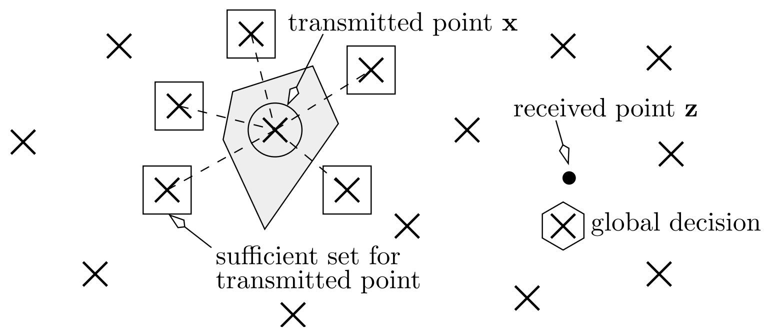

Figure 1: An example of a sufficient set for a given signal. Note that a global error cannot occur without a sufficient pairwise error event occurring.

where the lower bound is constructed by obtaining the largest lower-bound from a family (i.e., $P(A_i)$ is a lower bound for each value of i). Upper and lower bounds for the conditional error probability can then be constructed using either the union in (4) or (7). However, evaluation of $\Pr\{\mathbf{z}(u) \in \mathcal{Z}_m^c | \mathcal{H}_m\}$ is typically difficult (i.e., if it can be obtained, then often times an exact expression for the error probability can be obtained), so a bound constructed from pairwise errors is more generally applicable. In this case, applying (8) using the the expression in (7) yields

$$\max_j P_{PW}(j|\mathcal{H}_m) \leq P(\mathcal{E}|\mathcal{H}_m) \leq \sum_{j=0, j\neq m}^{M-1} P_{PW}(j|\mathcal{H}_m) \tag{9}$$

where

$$P_{PW}(j|\mathcal{H}_m) = \Pr\left\{\mathbf{z} \in \mathcal{Z}_j^{PW}(m) | \mathcal{H}_m\right\} \tag{10}$$

In practice the expression for $\mathcal{Z}_m^c$ in (7) is overly conservative owing to the fact that a subset of terms in the union may actually fully define the compliment of the decision region. Suppose that a set $\mathcal{N}_m \subset \{0, 1, \ldots M-1\}$ defines the region $\mathcal{Z}_m^c$ in the sense that

$$\mathcal{Z}_m^c = \bigcup_{j=0, j\neq m}^{M-1} \mathcal{Z}_j^{PW}(m) = \bigcup_{j \in \mathcal{N}_m}^{M-1} \mathcal{Z}_j^{PW}(m) \tag{11}$$

We refer to any such set as a sufficient set and the corresponding pairwise error events as *sufficient Pairwise Error Events*. A key property of such sets is that if a global error occurs, then some sufficient pairwise error event must have occurred. One may view the sufficient set as a set of "nearest neighbors" that determine the decision region; this concept is illustrated in Figure 1 for the case of minimum distance decisions. As a result, the upper bound in (9) can be tightened by replacing the sum over $j \neq m$ with the sum over $j \in \mathcal{N}_m$.

Bounds on the unconditional error probability can then be obtained by averaging these conditional bounds using the fact that $P(\mathcal{E}) = \sum_{m=0}^{M-1} P(\mathcal{E}|\mathcal{H}_m)P(\mathcal{H}_m)$

$$\sum_{m=0}^{M-1} P(\mathcal{H}_m)\left[\max_j P_{PW}(j|\mathcal{H}_m)\right] \leq P(\mathcal{E}) \leq \sum_{m=0}^{M-1} P(\mathcal{H}_m) \sum_{j \in \mathcal{N}_m}^{M-1} P_{PW}(j|\mathcal{H}_m) \tag{12}$$

1.1 Special Cases for AWGN Channels

A common special case for the application of the bounds developed above is that of a-priori equally-likely signaling over an AWGN where $\mathbf{z}(u) = \mathbf{s}_m + \mathbf{w}(u)$ under $\mathcal{H}_m$ with $\mathbf{w}(u)$ being AWGN. In this case, the MAP detection rule is the Minimum Distance rule and the pairwise error is

$$P_{PW}(j|\mathcal{H}_m) = \mathrm{Q}\left(\sqrt{\frac{d^2(j,m)}{2N_0}}\right) \quad d^2(j,m) = \|\mathbf{s}_j - \mathbf{s}_m\|^2 \tag{13}$$

In this case, the bound simplifies to

$$\frac{1}{M}\sum_{m=0}^{M-1}\mathrm{Q}\left(\sqrt{\frac{d^2_{\min}(m)}{2N_0}}\right) \leq P(\mathcal{E}) \leq \frac{1}{M}\sum_{m=0}^{M-1}\sum_{j\in\mathcal{N}_m}\mathrm{Q}\left(\sqrt{\frac{d^2(j,m)}{2N_0}}\right) \tag{14}$$

where

$$d^2_{\min}(m) = \min_{j\neq m} d^2(j,m) \tag{15}$$

A simple set of bounds can be obtained in terms of the global minimum distance

$$d^2_{\min} = \min_{m} d^2_{\min}(m) \tag{16}$$

Specifically, it is straightforward to show that

$$\frac{1}{M}\mathrm{Q}\left(\sqrt{\frac{d^2_{\min}}{2N_0}}\right) \leq P(\mathcal{E}) \leq (M-1)\mathrm{Q}\left(\sqrt{\frac{d^2_{\min}}{2N_0}}\right) \tag{17}$$

which implies that at high SNR, the error probability must decay proportionally to $\mathrm{Q}\left(\sqrt{\frac{d^2_{\min}}{2N_0}}\right)$ – i.e., the error probability of a binary test with only the nearest neighbor.

The last simple set of bounds is pessimistic (loose) in the sense that the upper bound was constructed using all other hypotheses as the sufficient set (i.e., hence the $(M-1)$ term) and the lower bound was constructed by taking only a single term from the sum. With some book-keeping, these bounds can be improved. Specifically, consider the *distance spectrum* of the signal set – i.e., the values of $d(m,j)$ that can occur for the specific set of signals. Order these distances $d_1 = d_{\min} < d_2 < d_3 \ldots$. Let $N_i(\{\mathcal{N}_m\})$ be the number of times that d_i occurs in a listing of all sufficient pairwise error events. Let K_i be the number of hypotheses for which $d_{\min}(m) = d_i$. The basic bound in (14) then simplifies to

$$\sum_i \frac{K_i}{M}\mathrm{Q}\left(\sqrt{\frac{d_i^2}{2N_0}}\right) \leq P(\mathcal{E}) \leq \sum_i \frac{N_i(\{\mathcal{N}_m\})}{M}\mathrm{Q}\left(\sqrt{\frac{d_i^2}{2N_0}}\right) \tag{18}$$

Note that one could use any single term from the above lower bound as a lower bound that may be much easier to compute and only slightly looser. For example, this yields a tighter lower bound in $d_{\min}$ of the form

$$P(\mathcal{E}) \geq \frac{K_1}{M}\mathrm{Q}\left(\sqrt{\frac{d^2_{\min}}{2N_0}}\right) \tag{19}$$

where K_1 is the number of times $d_{\min}$ occurs in the sum of the lower bound in (14). Note that, depending on the SNR, this may not be the best single-term lower bound. For example, the analogous bound based on d_2 may be larger at low SNR if $K_2 > K_1$. This motivates the lower bound

$$\max_i \frac{K_i}{M} \mathrm{Q}\left(\sqrt{\frac{d_i^2}{2N_0}}\right) \leq P(\mathcal{E}) \tag{20}$$

At moderate to high SNR, the upper bound is dominated by the $d_{\min}$ term so that

$$P(\mathcal{E}) \cong \frac{N_1(\{\mathcal{N}_m\})}{M} \mathrm{Q}\left(\sqrt{\frac{d_{\min}^2}{2N_0}}\right) \tag{21}$$

1.2 Conversion of M-ary Bounds to Bit Error Probability Bounds

The M hypotheses often correspond to $M = 2^q$ different binary (q-bit) words that label each signal point (one-to-one). As such, the issue of the probability of bit error arises. Let b_0, b_1, … b_{q-1} be a bit label. Any M-ary rule induces a rule for deciding between $b_i = 0$ and $b_i = 1$, so it is reasonable to consider the probability that the i^{th} bit is decided in error. Specifically, let $\mathcal{B}_i$ be the event that the M-ary rule yields a bit error at location i. Recalling that $\mathcal{E}$ is the event that a "symbol" error occurs (i.e., an error in the M-ary rule), then we have

$$P(\mathcal{B}_i) = \frac{P(\mathcal{B}_i|\mathcal{E})P(\mathcal{E})}{P(\mathcal{E}|\mathcal{B}_i)} = P(\mathcal{B}_i|\mathcal{E})P(\mathcal{E}) \tag{22}$$

since $P(\mathcal{E}|\mathcal{B}_i) = 1$ (i.e., an error in one bit location yields a symbol error). It is non-trivial to evaluate $P(\mathcal{B}_i|\mathcal{E})$, but we can consider the average bit error probability (i.e., averaged over bit locations)

$$P_b = \frac{1}{q}\sum_{i=0}^{q-1} P(\mathcal{B}_i) = \frac{1}{q}\sum_{i=0}^{q-1} P(\mathcal{B}_i|\mathcal{E})P(\mathcal{E}) \tag{23}$$

When a symbol error is made, at least one of the bits must be in error – this yields a lower bound. The upper bound s results from noting that $P(\mathcal{B}_i|\mathcal{E}) \leq 1$. This yields

$$\frac{1}{q}P(\mathcal{E}) \leq P_b \leq P(\mathcal{E}) \tag{24}$$

This result can then be combined with any set of bounds for $P(\mathcal{E})$ yielding

$$\frac{1}{q}B_L(\mathcal{E}) \leq P_b \leq B_U(\mathcal{E}) \tag{25}$$

where $B_L(\mathcal{E})$ and $B_U(\mathcal{E})$ are upper and lower bounds on $P(\mathcal{E})$, respectively. Note that this approach is sufficiently tight only when q is small (i.e., small number of bits mapped to a symbol.

2 Lower Bounds via Side Information

One tool that can be used to obtain a simple lower bound for any problem is the use of side information. This is often described as a "genie" that aids a receiver. The reasoning is that the optimal genie-aided receiver must perform at least as well as the non-aided receiver. Thus, by choosing the genie's rules carefully, one can obtain a good lower bound that is easy to evaluate.

To formalize this notion, define a random vector $\mathbf{v}(u)$ that is the side information. Let us focus on the case where a lower bound is desired for $P(\mathcal{E})$ for an M-ary decision when the MAP M-ary decision rule is used. In that case, the decision rule of a genie-aided receiver is

$$\max_m f_{\mathbf{z}(u),\mathbf{v}(u)}(\mathbf{z},\mathbf{v}|\mathcal{H}_m)P(\mathcal{H}_m) \iff \max_m f_{\mathbf{z}(u)|}(\mathbf{z}|\mathcal{H}_m)P_{\mathbf{v}(u)}(\mathbf{v}|\mathcal{H}_m)P(\mathcal{H}_m) \tag{26}$$

where it has been assumed that $\mathbf{z}(u)$ is independent of $\mathbf{v}(u)$ given the hypothesis. Consider the special case where the genie provides the receiver with the index of the correct hypothesis and some other hypothesis with a predetermined probability. For example, if $\mathcal{H}_m$ is the true hypothesis, the genie will provide the receiver with $\mathbf{v} = \{m,j\}$ with probability $P_{\mathbf{v}(u)}(\mathbf{v} = \{m,j\}|\mathcal{H}_m)$. The genie never gives incorrect information – e.g., the genie will never reveal $\mathbf{v} = \{3,4\}$ if $\mathcal{H}_2$ is correct. Notice that the error probability given the genie's side information is very similar to a pairwise error probability for the MAP detector.

We define a special type of side information scheme as *Pairwise Uniform Revelation of Side Information (PURSI)*. In a PURSI scheme the side information is a pair of indices with

$$P_{\mathbf{v}(u)}(\mathbf{v} = \{m,j\}|\mathcal{H}_m) = P_{\mathbf{v}(u)}(\mathbf{v} = \{m,j\}|\mathcal{H}_j) \tag{27}$$

Thus, in a PURSI scheme, given that $\mathcal{H}_m$ is true, the side information will be $\mathbf{v} = \{m,j\}$ and the decision in (26) will reduces to

$$f_{\mathbf{z}(u)|}(\mathbf{z}|\mathcal{H}_m)P(\mathcal{H}_m) \underset{\mathcal{H}_j}{\overset{\mathcal{H}_m}{\gtrless}} f_{\mathbf{z}(u)|}(\mathbf{z}|\mathcal{H}_j)P(\mathcal{H}_j) \tag{28}$$

This follows since $P_{\mathbf{v}(u)}(\mathbf{v} = \{m,j\}|\mathcal{H}_i)$ is zero if i is not j or m and it is equal for $i = m$ and $i = j$. Thus, the conditional probability of error for a PURSI scheme is

$$P(\mathcal{E}|\mathbf{v} = \{m,j\},\mathcal{H}_m) = P_{PW}(j|\mathcal{H}_m) \tag{29}$$

A primary example of how this may be used is to reveal pairs that correspond to signals at a given distance, thus obtaining a bound similar to that in (20).

The primary use of side information is for obtaining lower bounds on optimal bit detectors. Bounds on bit error probability based on bounds for $P(\mathcal{E})$ may not be useful when M is large. One may also be interested in the bit error probability $P(\mathcal{B}_i)$ associated with the MAP detector for b_i. Note that upper bounds of the form in (25) are valid upper bounds for $P(\mathcal{B}_i)$ for the MAP bit detector since the latter obtains the minimum bit error probability. However, a lower bound of the form in (25) is not a valid lower bound for the bit error probability of the MAP bit detector. In general, performance analysis for optimal bit detectors is more difficult that for the M-ary detectors.

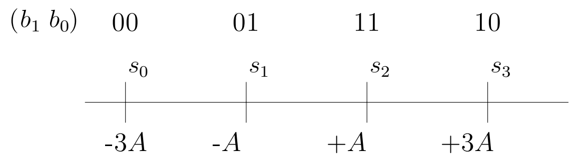

Figure 2: The 4-PAM signal set with bit labels.

3 A Detailed Example

Consider the simple $M = 4$ Pulse Amplitude Modulation (PAM) signal with $s_0 = -3A$, $s_1 = -A$, $s_2 = +A$, and $s_3 = +3A$, where $A > 0$. In addition, assume that two bits, b_0 and b_1 are mapped onto these signals using a Gray mapping. Assume that all bits, and therefore symbols, are equally likely. This is illustrated in Figure 2. In this section we apply all of the previous development to this problem. First, let us consider the MAP symbol and bit detectors based on the observation $z(u) = s_m + w(u)$ under $\mathcal{H}_m$, with $w(u)$ a mean-zero Gaussian random variable with variance σ^2.

3.1 Optimal Symbol Detector and Exact Performance

The MAP symbol detector is a simple minimum distance rule in this case. The decision regions for the MAP symbol detector are illustrated in Figure 3. We can find exact expressions for the probability of symbol error and the bit error probabilities in this simple case. In particular, it is straightforward to verify that

$$P(\mathcal{E}|\mathcal{H}_0) = \mathrm{Q}(A/\sigma) \tag{30a}$$
$$P(\mathcal{E}|\mathcal{H}_1) = 2\mathrm{Q}(A/\sigma) \tag{30b}$$
$$P(\mathcal{E}|\mathcal{H}_2) = 2\mathrm{Q}(A/\sigma) \tag{30c}$$
$$P(\mathcal{E}|\mathcal{H}_3) = \mathrm{Q}(A/\sigma) \tag{30d}$$

Averaging over the equal prior probabilities yields

$$P(\mathcal{E}) = \frac{3}{2}\mathrm{Q}(A/\sigma) \tag{31}$$

We can find the probability of bit error for each bit location for this optimal symbol detector as well. Specifically, for b_0 we have the decision rule implied by the MAP symbol detection rule as illustrated in Figure 4 with $T = 2A$.

$$P(\mathcal{B}_0|\mathcal{H}_0) = \mathrm{Q}(A/\sigma) - \mathrm{Q}(5A/\sigma) \tag{32a}$$
$$P(\mathcal{B}_0|\mathcal{H}_1) = \mathrm{Q}(A/\sigma) + \mathrm{Q}(3A/\sigma) \tag{32b}$$
$$P(\mathcal{B}_0|\mathcal{H}_2) = \mathrm{Q}(A/\sigma) + \mathrm{Q}(3A/\sigma) \tag{32c}$$
$$P(\mathcal{B}_0|\mathcal{H}_3) = \mathrm{Q}(A/\sigma) - \mathrm{Q}(5A/\sigma) \tag{32d}$$

Again, averaging over the four equally-likely hypotheses, we have

$$P(\mathcal{B}_0) = \mathrm{Q}(A/\sigma) + \frac{1}{2}\mathrm{Q}(3A/\sigma) - \frac{1}{2}\mathrm{Q}(5A/\sigma) \tag{33}$$

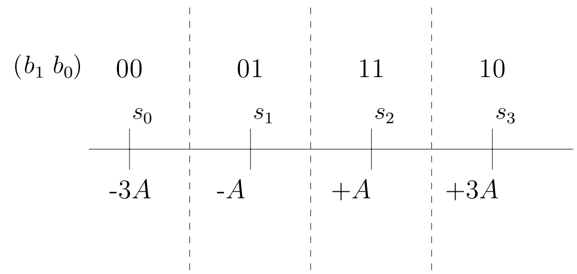

Figure 3: The decision regions for MAP symbol detection.

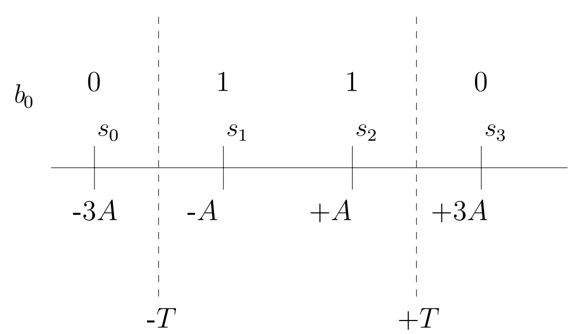

Figure 4: The decision regions for b_0 implied by the MAP symbol detection rule when $T = 2A$.

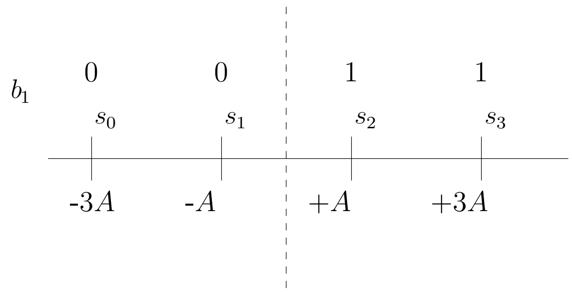

Figure 5: The decision regions for b_1 implied by the MAP symbol detection rule.

The error probability for the bit b_1 is easier to compute. Specifically, the decision region is that shown in Figure 5, which implies

$$P(\mathcal{B}_1|\mathcal{H}_0) = P(\mathcal{B}_1|\mathcal{H}_3) = \mathrm{Q}(3A/\sigma) \tag{34a}$$

$$P(\mathcal{B}_1|\mathcal{H}_1) = P(\mathcal{B}_1|\mathcal{H}_2) = \mathrm{Q}(A/\sigma) \tag{34b}$$

It follows that the average error probability on b_1 is

$$P(\mathcal{B}_1) = \frac{1}{2}\mathrm{Q}(A/\sigma) + \frac{1}{2}\mathrm{Q}(3A/\sigma) \tag{35}$$

The average bit error probability, as defined in (23), for the MAP symbol detector is

$$P_b = \frac{3}{4}\mathrm{Q}(A/\sigma) + \frac{1}{2}\mathrm{Q}(3A/\sigma) - \frac{1}{4}\mathrm{Q}(5A/\sigma) \tag{36}$$

3.2 Optimal Bit Detector

First, consider the MAP detector for bit b_1. Applying the notion of average likelihood and averaging out the effects of b_0, we obtain the following rule

$$\Lambda(z) = \frac{\exp\left[\frac{-1}{2\sigma^2}(z-3A)^2\right] + \exp\left[\frac{-1}{2\sigma^2}(z-A)^2\right]}{\exp\left[\frac{-1}{2\sigma^2}(z+3A)^2\right] + \exp\left[\frac{-1}{2\sigma^2}(z+A)^2\right]} \underset{\mathcal{H}_{b_1=0}}{\overset{\mathcal{H}_{b_1=1}}{\gtrless}} 1 \tag{37}$$

While this expression looks quite complicated, it has the property that $\Lambda(-z) = [\Lambda(z)]^{-1}$. It follows that the observation $z = 0$ must correspond to a decision boundary (i.e., $\Lambda(0) = 1$). Moreover, this property implies that the rule must be the same on a give side of $z = 0$. It is simple to verify that this yields the following rule

$$z \underset{\mathcal{H}_{b_1=0}}{\overset{\mathcal{H}_{b_1=1}}{\gtrless}} 0 \tag{38}$$

Note that this is the same rule implied for b_1 by the MAP symbol decision regions illustrated in Figure 3.

For the bit label b_0, the MAP decisions rule is different than that implied by the MAP symbol detection rule. Similar to the above we have

$$\Lambda(z) = \frac{\exp\left[\frac{-1}{2\sigma^2}(z+A)^2\right] + \exp\left[\frac{-1}{2\sigma^2}(z-A)^2\right]}{\exp\left[\frac{-1}{2\sigma^2}(z+3A)^2\right] + \exp\left[\frac{-1}{2\sigma^2}(z-3A)^2\right]} \underset{\mathcal{H}_{b_0=0}}{\overset{\mathcal{H}_{b_0=1}}{\gtrless}} 1 \tag{39}$$

This can be simplified to

$$\frac{\cosh\left(\frac{3A}{\sigma^2}|z|\right)}{\cosh\left(\frac{A}{\sigma^2}|z|\right)} \underset{\mathcal{H}_{b_1=1}}{\overset{\mathcal{H}_{b_1=0}}{\gtrless}} \exp\left[\frac{4A^2}{\sigma^2}\right] \tag{40}$$

It is straightforward to verify that $g(x) = \cosh(3x)/\cosh(x)$ is a strictly monotonic function on $x > 0$. As a result, we can define the inverse function $g^{-1}(y)$ which maps $y > 0$ to x such that $y = g(x)$. Using this fact, we obtain a simple form for that test

$$|z| \underset{\mathcal{H}_{b_0=1}}{\overset{\mathcal{H}_{b_0=0}}{\gtrless}} T = \sigma\left(\frac{A}{\sigma}\right)^{-1} g^{-1}\left(\exp\left[\frac{4A^2}{\sigma^2}\right]\right) \tag{41}$$

Notice that this test is similar to the test implied by the MAP symbol detector and shown in Figure 4. However, the value of T for the MAP symbol detection rule is $T = 2A$. For this optimal rule for b_0, T is a function of the parameters A and σ^2. However, at moderate to high SNR, we expect the rules to be approximately the same. In fact, note that $g(x) \sim e^{2x}$ so that $T \sim 2A$. This motivates writing the optimal bit detection rule in terms of a threshold centered around $2A$. Specifically, define ϵ by the relation $T = 2A + \epsilon$, where T is defined in (41).

The exact probability of error for the MAP bit detector can also be computed. In the case of b_1, since the rule is the same as that implied by the MAP symbol detector, the performance is also the same (i.e., as given in (35)). For the MAP detector of b_0, the performance can also be determined in terms of ϵ. This is very similar to the $P(\mathcal{B}_0)$ analysis for the symbol detector $\Longrightarrow$

$$P(\mathcal{B}_0|\mathcal{H}_0) = P(\mathcal{B}_0|\mathcal{H}_3) = \mathrm{Q}([A-\epsilon]/\sigma) - \mathrm{Q}([5A+\epsilon]/\sigma) \tag{42}$$

$$P(\mathcal{B}_0|\mathcal{H}_1) = P(\mathcal{B}_0|\mathcal{H}_2) = \mathrm{Q}([A+\epsilon]/\sigma) + \mathrm{Q}([3A+\epsilon]/\sigma) \tag{43}$$

Averaging over the a-priori statistics yields

$$P(\mathcal{B}_0) = \frac{1}{2}\left[\mathrm{Q}([A-\epsilon]/\sigma) + \mathrm{Q}([A+\epsilon]/\sigma)\right] + \frac{1}{2}\mathrm{Q}([3A+\epsilon]/\sigma) - \frac{1}{2}\mathrm{Q}([5A+\epsilon]/\sigma) \tag{44}$$

Notice that this expression for $P(\mathcal{B}_0)$ for the minimum bit error probability receiver reduces to that in (33) as $\epsilon \to 0$. Also, note that ϵ/σ is only a function of A/σ (as opposed to A and σ separately). A plot of ϵ/σ is given in Figure 6. Notice that ϵ/σ tends toward a value of 1 for $A/\sigma^2 \to 0$ and tends toward zero for large A/σ.

3.3 Bounds for the MAP Symbol Detector

The bounds in (14) result in the following upper and lower bounds:

$$\mathrm{Q}(A/\sigma) \leq P(\mathcal{E}) \leq \frac{3}{2}\mathrm{Q}(A/\sigma) \tag{45}$$

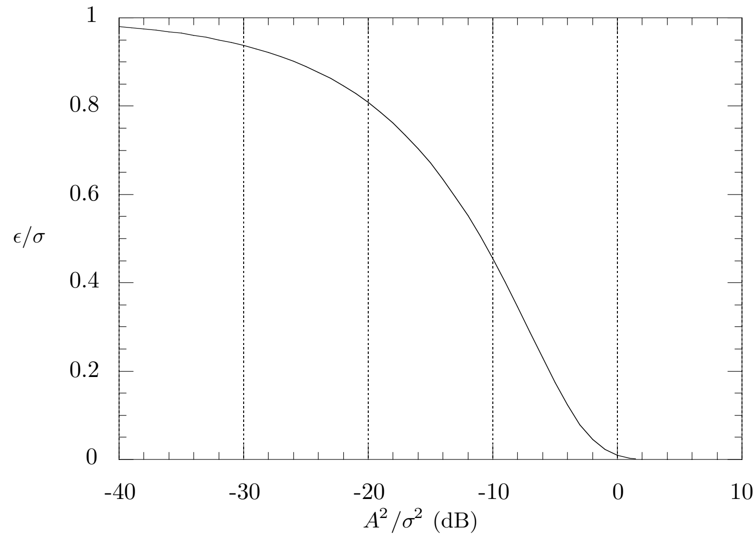

Figure 6: The deviation of the optimal threshold for detection of b_0 from that associated with the MAP symbol detector. Note that ϵ/σ is approximately zero over the entire range of useful SNR.

The lower bound follows from the fact that each signal has a neighbor at the distance $d_{\min} = 2A$. It turns out that, for this special case, evaluation of the upperbound in (14) based on pairwise error events is the same as the exact expression. This is because the complement of each global decision region is defined by two disjoint pairwise decision regions.

The method of side information can be illustrated by considering several examples. Let's consider three different side information schemes – i.e., genie-A, genie-B and genie-C. The first genie, genie-A, uses the side information scheme defined by the conditional statistics

$$\mathcal{H}_0 : \ \mathbf{v} = \{s_0, s_1\} \qquad \text{with probability 1} \tag{46}$$

$$\mathcal{H}_1 : \ \mathbf{v} = \{s_0, s_1\} \qquad \text{with probability 1} \tag{47}$$

$$\mathcal{H}_2 : \ \mathbf{v} = \{s_2, s_3\} \qquad \text{with probability 1} \tag{48}$$

$$\mathcal{H}_3 : \ \mathbf{v} = \{s_2, s_3\} \qquad \text{with probability 1} \tag{49}$$

This defines a PURSI scheme since $P(\mathbf{v}(u) = \{s_0, s_1\}|\mathcal{H}_0) = P(\mathbf{v}(u) = \{s_0, s_1\}|\mathcal{H}_1) = 1$ and $P(\mathbf{v}(u) = \{s_2, s_3\}|\mathcal{H}_2) = P(\mathbf{v}(u) = \{s_2, s_3\}|\mathcal{H}_3) = 1$. When the genie reveals $\mathbf{v} = \{s_0, s_1\}$, the optimal symbol detector executes the test

$$f_{\mathbf{z}(u)|}(\mathbf{z}|\mathcal{H}_0)P(\mathbf{v}(u) = \{s_0, s_1\}|\mathcal{H}_0)P(\mathcal{H}_0) \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\gtrless}} f_{\mathbf{z}(u)|}(\mathbf{z}|\mathcal{H}_1)P(\mathbf{v}(u) = \{s_0, s_1\}|\mathcal{H}_1)P(\mathcal{H}_1) \tag{50}$$

which, because of the PURSI property and the equal a-priori probabilities, is a minimum distance decision. Note that the receiver performs this pairwise decision because $P(\mathbf{v}(u) = \{s_0, s_1\}|\mathcal{H}_m) = 0$ for $m = 2, 3$. It follows that the optimal receiver aided by genie-A has $P(\mathcal{E}|\mathcal{H}_0) = P(\mathcal{E}|\mathcal{H}_1) = \mathrm{Q}(A/\sigma)$. The same argument yields for $P(\mathcal{E}|\mathcal{H}_2) = P(\mathcal{E}|\mathcal{H}_3) = \mathrm{Q}(A/\sigma)$. So, this side information scheme yields the bound in (45).

Consider genie-B which uses the scheme

$$\mathcal{H}_0: \ \mathbf{v} = \{s_0\} \qquad \text{with probability 1} \tag{51}$$

$$\mathcal{H}_1: \ \mathbf{v} = \{s_0, s_1\} \qquad \text{with probability 1} \tag{52}$$

$$\mathcal{H}_2: \ \mathbf{v} = \{s_2, s_3\} \qquad \text{with probability 1} \tag{53}$$

$$\mathcal{H}_3: \ \mathbf{v} = \{s_3\} \qquad \text{with probability 1} \tag{54}$$

It is straightforward to verify that this is also a PURSI scheme. Since the receiver knows the correct hypothesis when genie-B reveals either $\mathbf{v} = \{s_0\}$ or $\mathbf{v} = \{s_3\}$, errors are only made under the other two values for the side information. This leads to the lower bound $P(\mathcal{E}) \geq \frac{1}{2}\mathrm{Q}(A/\sigma)$, which is smaller by the factor of 1/2 than the bound obtained with genie-A. This illustrates the fact that, the more information the genie provides, the less tight the lower bound will be.

Consider a non-PURSI genie-C defined by

$$\mathcal{H}_0: \ \mathbf{v} = \{s_0, s_1\} \qquad \text{with probability 1} \tag{55}$$

$$\mathcal{H}_1: \ \mathbf{v} = \{s_0, s_1\} \quad \text{or } \{s_1, s_2\} \quad \text{with probability 1/2 each} \tag{56}$$

$$\mathcal{H}_2: \ \mathbf{v} = \{s_1, s_2\} \quad \text{or } \{s_2, s_3\} \quad \text{with probability 1/2 each} \tag{57}$$

$$\mathcal{H}_3: \ \mathbf{v} = \{s_2, s_3\} \qquad \text{with probability 1} \tag{58}$$

The fact that this genie is non-PURSI does not preclude our ability to construct a valid lower bound; it's just more difficult. Specifically, conditioned on $\mathcal{H}_0$, we will be given side information $\mathbf{v} = \{s_0, s_1\}$ which is only possible if $\mathcal{H}_0$ or $\mathcal{H}_1$ is true. However, $P(\mathbf{v}(u) = \{s_0, s_1\}|\mathcal{H}_0) = 1$, while $P(\mathbf{v}(u) = \{s_0, s_1\}|\mathcal{H}_1) = 1/2$. It follows that the test conducted is

$$f_{\mathbf{z}(u)|}(\mathbf{z}|\mathcal{H}_0)(1) \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\gtrless}} f_{\mathbf{z}(u)|}(\mathbf{z}|\mathcal{H}_1)(1/2) \tag{59}$$

Note that this rule is equivalent to a binary test with $\pi_0 = 2\pi_1$. As a result, conditioned on $\mathcal{H}_0$ we have

$$P(\mathcal{E}|\mathcal{H}_0) = \mathrm{Q}\left(\frac{A}{\sigma} + \left[\frac{A}{\sigma}\right]^{-1} \ln(2)\right) \tag{60}$$

To find $P(\mathcal{E}|\mathcal{H}_1)$, consider the cases of $\mathbf{v} = \{s_0, s_1\}$ and $\mathbf{v} = \{s_1, s_2\}$ separately. First consider $\mathbf{v} = \{s_0, s_1\}$, which is similar to the above development

$$P(\mathcal{E}|\mathcal{H}_1, \mathbf{v} = \{s_0, s_1\}) = \mathrm{Q}\left(\frac{A}{\sigma} - \left[\frac{A}{\sigma}\right]^{-1} \ln(2)\right) \tag{61}$$

Since $P(\mathbf{v}(u) = \{s_1, s_2\}|\mathcal{H}_1) = P(\mathbf{v}(u) = \{s_1, s_2\}|\mathcal{H}_2) = 1/2$, the associated test is a minimum distance test between s_1 and s_2 so that

$$P(\mathcal{E}|\mathcal{H}_1, \mathbf{v} = \{s_1, s_2\}) = \mathrm{Q}(A/\sigma) \tag{62}$$

It follows that

$$P(\mathcal{E}|\mathcal{H}_1) = \frac{1}{2}\mathrm{Q}\left(\frac{A}{\sigma} - \left[\frac{A}{\sigma}\right]^{-1} \ln(2)\right) + \frac{1}{2}\mathrm{Q}(A/\sigma) \tag{63}$$

Using the same type of development you can show that $P(\mathcal{E}|\mathcal{H}_3) = P(\mathcal{E}|\mathcal{H}_0)$ and $P(\mathcal{E}|\mathcal{H}_1) = P(\mathcal{E}|\mathcal{H}_2)$. The lower bound obtained is therefore that of this genie-C aided receiver

$$P(\mathcal{E}) \geq \frac{1}{2}\mathrm{Q}\left(\frac{A}{\sigma} + \left[\frac{A}{\sigma}\right]^{-1}\ln(2)\right) + \frac{1}{4}\mathrm{Q}\left(\frac{A}{\sigma} - \left[\frac{A}{\sigma}\right]^{-1}\ln(2)\right) + \frac{1}{4}\mathrm{Q}(A/\sigma) \tag{64}$$

This bound can also be written as

$$\frac{3}{4}\left[\frac{2}{3}\mathrm{Q}\left(\frac{A}{\sigma} + \left[\frac{A}{\sigma}\right]^{-1}\ln(2)\right) + \frac{1}{3}\mathrm{Q}\left(\frac{A}{\sigma} - \left[\frac{A}{\sigma}\right]^{-1}\ln(2)\right)\right] + \frac{1}{4}\mathrm{Q}(A/\sigma) \leq \mathrm{Q}(A/\sigma) \tag{65}$$

In other words, the bound obtained using genie-C is not as good as the bound obtained using genie-A. This illustrates another property of the method of side information: a tighter upperbound will generally be obtained when the genie's side information is revealed in a PURSI manner.

3.4 Bounds for the MAP Bit Detector

Note that the expressions for $P(\mathcal{B}_0)$, $P(\mathcal{B}_1)$, and P_b for the MAP symbol detector in Section 3.1 are upper bounds for the performance of the optimal bit detector (why?). We can use side information techniques to obtain lower bounds for the optimal bit detector too (which also serve as lower bounds on bit error probabilities for all detectors, including the MAP symbol detector). One technique is to use a PURSI scheme where, for each hypothesis the correct signal and a signal that differs in the location of interest is revealed.

For example, consider the bit b_0. A lower bound on $P(\mathcal{B}_0)$ can be found using the side information scheme of genie-A in Section 3.3. Note that this PURSI scheme has the property that each pair revealed differs in b_0. This is a good property for finding a large lower bound on $P(\mathcal{B}_0)$. Note that, for the genie-A-aided detector, MAP symbol detection and MAP detection of b_0 are the same. It follows that a lower bound for *any* receiver on $P(\mathcal{B}_0)$ is

$$P(\mathcal{B}_0) \geq \mathrm{Q}(A/\sigma) \tag{66}$$

For bit b_1, we know the exact value of $P(\mathcal{B}_1)$, but if one were to apply the PURSI method to obtain a lower bound, the method of genie-B from Section 3.3 provides the bound $P(\mathcal{B}_0) \geq \frac{1}{2}\mathrm{Q}(A/\sigma)$. Another PURSI scheme is defined by

$$\mathcal{H}_0 : \ \mathbf{v} = \{s_0, s_3\} \qquad \text{with probability 1} \tag{67}$$
$$\mathcal{H}_1 : \ \mathbf{v} = \{s_1, s_2\} \qquad \text{with probability 1} \tag{68}$$
$$\mathcal{H}_2 : \ \mathbf{v} = \{s_1, s_2\} \qquad \text{with probability 1} \tag{69}$$
$$\mathcal{H}_3 : \ \mathbf{v} = \{s_0, s_3\} \qquad \text{with probability 1} \tag{70}$$

It can be shown that this yields the lower bound for $P(\mathcal{B}_1)$ for any receiver of

$$P(\mathcal{B}_1) \geq \frac{1}{2}\mathrm{Q}(A/\sigma) + \frac{1}{2}\mathrm{Q}(3A/\sigma) \tag{71}$$

Once again, this illustrates that tighter bounds are obtained by revealing less information. In this case, the lower bound coincides with the exact expression.

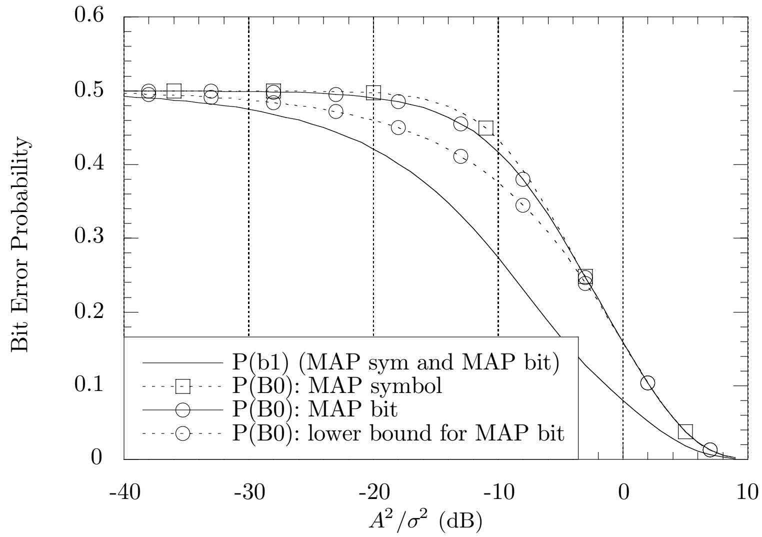

Figure 7: The bit error probability: $P(\mathcal{B}_1)$ for the optimal bit/symbol detector (i.e., see (35)), and $P(\mathcal{B}_1)$ for the optimal bit detector (i.e., see (41)), symbol detector (i.e., see (33)), and the lower bound in (66).

Combining the lower bounds for $P(\mathcal{B}_1)$ and $P(\mathcal{B}_0)$ in (71) and (66), respectively, we obtain the lower bound for the bit error probability of any receiver

$$P_b \geq \frac{3}{4}\mathrm{Q}(A/\sigma) + \frac{1}{4}\mathrm{Q}(3A/\sigma) \tag{72}$$

This toy example was intended to provide insight into the bounding techniques described. Nearly all of the bounds are useless since we already have exact expressions for the desired error probabilities. One possible exception is the bound in (66), which is a useful expression when compared against the exact expression in (44) which requires evaluation of the parameter ϵ (i.e., the $g^{-1}(\cdot)$ function). In Figure 7, $P(\mathcal{B}_1)$ is plotted along with $P(\mathcal{B}_0)$ for the optimal detector, the symbol detector, and the associated lower bound.