### CONVOLUTIONAL CODING OPTIONS FOR MFSK/FH SIGNALING ON RAYLEIGH FADING AND PARTIAL BAND INTERFERENCE CHANNELS

#### Authors

P. J. Crepeau and D. L. Tate

#### Report Date

November 22, 1982

#### Distribution Statement

Approved for public release; distribution unlimited.

#### Abstract

Convolutional coding offers substantial improvement in the bit error probability performance of MFSK/FH modulation when applied to Rayleigh fading or partial band Gaussian interference channels. Using a 4-ary modulation with rate-1/2 codes as a special example, several convolutional coding options are considered, including binary, dual-2-, and triple-2 codes. Union-Cherno upper bounds reveal that improvements of up to 35 dB (in E_b/N_0 reduction) are achievable at a 0.1% bit error probability in comparison to an uncoded system.
## CONTENTS

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>2.</td>
<td>PERFORMANCE OF UNCODED MFSK/FH ON RAYLEIGH FADING AND PARTIAL BAND GAUSSIAN CHANNELS</td>
<td>1</td>
</tr>
<tr>
<td>3.</td>
<td>CODES AND THEIR WEIGHT STRUCTURE</td>
<td>6</td>
</tr>
<tr>
<td>4.</td>
<td>PERFORMANCE BOUNDS</td>
<td>12</td>
</tr>
<tr>
<td>5.</td>
<td>DECODER COMPLEXITY</td>
<td>17</td>
</tr>
<tr>
<td>6.</td>
<td>RESULTS AND CONCLUSIONS</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>REFERENCES</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>APPENDIX A - COMPUTATION OF CODE WEIGHT STRUCTURE</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>APPENDIX B - CHERNOFF BOUND FOR BINARY FSK ERROR PROBABILITY ON A RAYLEIGH FADING CHANNEL</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>APPENDIX C - CHERNOFF BOUND FOR BINARY FSK ERROR PROBABILITY ON A PARTIAL BAND GAUSSIAN CHANNEL</td>
<td>27</td>
</tr>
</tbody>
</table>
CONVOLUTIONAL CODING OPTIONS FOR MFSK/FH SIGNALING ON RALEIGH FADING AND PARTIAL BAND INTERFERENCE CHANNELS

1. INTRODUCTION

It is well-known that forward error correction coding offers significant performance advantages when applied to the reception of signals in coherent additive white Gaussian noise (AWGN) channels. For example, with a constraint length 7, rate 1/2, convolutional code with soft decision decoding, the $E_b/N_0$ required for a $10^{-5}$ bit error probability is reduced by about 5 dB relative to the uncoded system. In 1975, Viterbi and Jacobs [1] showed that much larger coding gains are achievable on noncoherent channels with either amplitude fading or additive partial band Gaussian interference. In fact, even with modest coding protection, coding gains greater than 30 dB are possible for fading or worst case partial band interference channels.

In this report we investigate several convolutional coding options which are applicable to the noncoherent fading or partial band Gaussian channel. These results are useful, for example, on an HF channel where multiple ionospheric reflections lead to a loss of phase information as well as introducing a fading amplitude characteristic. In other instances, channels exist with partial band interference caused either by unregulated other-users or by intentional jammers.

The modulation technique considered here is multiple frequency shift keying with frequency hopping (MFSK/FH). With orthogonal tone spacings, MFSK is an optimal M-ary modulation for noncoherent channels. Frequency hopping and (possibly) interleaving are assumed to be used to make the channel memoryless, that is, to produce statistical independence from symbol to symbol in a sequence of received symbols. In the following section the uncoded performance of this modulation technique is given for the two channel models in question.

2. PERFORMANCE OF UNCODED MFSK/FH ON RAYLEIGH FADING AND PARTIAL BAND GAUSSIAN CHANNELS

For orthogonal MFSK the M-ary signal alphabet is chosen as the set of tones:

$$s_i(t) = \sqrt{2E_s} \cos(2\pi f_s t + \theta), 0 \leq t \leq T_s$$

where

- $E_s$ is the received symbol energy,
- $T_s$ is the symbol duration,
- and $\theta$ is a uniformly distributed random phase.

The frequencies are chosen to insure mutual orthogonality of the signals, so that the tone spacings are

\[ f_{i+1} - f_i = \frac{j}{T_s}, \quad i = 1, 2, \ldots, M-1 \] (2)

For a symbol to carry \( k \) bits of information, we use \( M = 2^k \), and it follows that

\[ E_b = \frac{E_s}{b} \] (3)

where \( E_b \) is the received energy per bit.

For a Rayleigh fading channel when signal \( s_i(t) \) is sent, the received signal is

\[ r(t) = r \sqrt{\frac{2E_s}{T_s}} \cos(2\pi f_i t + \theta) + n(t), \quad 0 \leq t \leq T_s \] (4)

where \( n(t) \) is an AWGN process with two-sided power spectral density \( \frac{N_0}{2} \) Watts per Hertz, and \( r \) is a Rayleigh random variable with normalized probability density function

\[ p(r) = 2r e^{-r^2}, \quad r > 0 \] (5)

With this normalization, \( r^2 E_s \) is the received energy random variable whose average value \( E_s \) is \( E_s \).

The uncoded performance of MFSK in a memoryless Rayleigh fading channel is well-known [2]. The symbol error probability \( P_s \) is given by

\[ P_s = \sum_{n=1}^{M-1} \binom{M-1}{n} (-1)^{n+1} \frac{1}{E_s} \frac{E_s}{N_0} \] (6)

Furthermore, with orthogonal signaling, bit and symbol error probabilities are related by

\[ P_b = \frac{M}{2(M-1)} P_s \] (7)
Thus from (6), (7), and (3) we have

\[ P_b = \frac{M}{2(M-1)} \sum_{n=1}^{M-1} \binom{M-1}{n} (-1)^{n+1} \frac{1}{(n+1) + nk} \frac{E_b}{N_0} \]  

(8)

The curves representing (8) are plotted in Figure 1 for \( M=2, 4, \ldots, 32 \). We see that for the case of \( M=2 \), an \( E_b/N_0 \) of approximately 50 dB is required to achieve a bit error probability of \( 10^{-5} \). Increasing \( M \) offers little improvement; in fact it has been shown [2] that as \( M \rightarrow \infty \), the improvement over the binary case is only \( 2/\ln 2 \) (4.6 dB).

For uncoded MFSK/FH on a partial band Gaussian interference channels the result is similarly poor. The partial band channel is characterized by constant density AWGN over a fraction of the total hopping transmission bandwidth. Thus, the noise spectrum has density \( N_0/\rho \) over a fraction \( \rho \) of the band (where \( 0 < \rho < 1 \)) and is zero elsewhere (over a fraction \( 1-\rho \) of the band). We assume that the \( M \) candidate tone slots for each symbol are either all interfered with or they are all noise free. As a worst-case condition, we consider only the situation where the parameter \( \rho \) is chosen so as to maximize the resulting probability of error.

For uncoded MFSK/FH in a worst case partial band Gaussian channel the results (over the range of interest) have been found in [3] as

\[ P_b = \frac{b}{E_b N_0} \]  

(9)

where the numerator \( b \) is a constant depending on the parameter \( M \), given as follows:

\[
\begin{array}{c|c}
M & b \\
2 & .37 \\
4 & .23 \\
8 & .20 \\
16 & .18 \\
32 & .17 \\
\end{array}
\]

The curves representing (9) are plotted in Figure 2. The results are about 4 dB better than the corresponding Rayleigh fading curves, but the performance is again poor because the dependence of \( P_b \) on \( E_b/N_0 \) is inverse linear (rather than exponential as in the AWGN channel). For each of these two channels we shall see that coding can provide a substantial performance improvement.
Fig. 1 — Error rates for M-ary orthogonal modulation on Rayleigh fading channels
Fig. 2 — Error rates for M-ary orthogonal modulation on partial band Gaussian channels
3. CODES AND THEIR WEIGHT STRUCTURE

In this report we consider the performance of various convolutional codes on M-ary orthogonal channels with Rayleigh fading or partial band Gaussian interference. The general block diagram for this system is shown in Figure 3. We restrict our consideration to convolutional codes, and assume that soft decision Viterbi decoding is used throughout.

![General modulation/coding configuration](image)

Fig. 3 – General modulation/coding configuration

The specific case of rate-1/2 codes with 4-ary orthogonal modulation shall be considered in detail. The techniques used, however, are applicable to all M-ary modulations and for various code rates. The codes considered fall into two groups: binary convolutional codes (that accept single bit inputs), and quaternary convolutional codes (that accept pairs of bits at the input). The former are designated by their constraint length K and we shall consider the cases K=3,4,5,6, and 7. From the latter class we shall investigate the dual-2 and triple-2 codes, which have equivalent binary constraint lengths K=4 and K=6 respectively.

These seven encoders are shown in Figures 4 through 10. All tap connections (code generators) are optimized so that the codes will have maximum (or nearly maximum) distance properties with respect to 4-ary channels [4,5].
Fig. 4 – K = 7 encoder

Fig. 5 – K = 6 encoder
Fig. 6 – Triple-2 encoder

Fig. 7 – $K = 5$ encoder

Fig. 8 – $K = 4$ encoder
Fig. 9 – Dual-2 encoder

Fig. 10 – K = 3 encoder
In evaluating performance of a convolutional code, the only code property of importance is its weight structure. The code weight structure gives the number of decoded errors associated with erroneous decoding paths of various distances, starting with the code's minimum distance. Because of the linear (group) property of convolutional encoders, all error calculations may be based on the assumption that the code has an input which is the all-zero binary sequence.

In computing distance it must be kept in mind that we are dealing with an M-ary orthogonal (equidistant) signal set. A 4-ary modulator performs a mapping of code bit pairs into channel signals as follows:

\[
\begin{align*}
00 & \rightarrow s_0 \\
01 & \rightarrow s_1 \\
10 & \rightarrow s_2 \\
11 & \rightarrow s_3
\end{align*}
\]

Since \(s_1, s_2,\) and \(s_3\) are equidistant from \(s_0\) in the Euclidean signal space of the channel, we see that these signals must be treated as equal adversaries to the correct signal \(s_0\) at the demodulator output. Distance from the all-\(s_0\) sequence of signals for this channel is the number of non-\(s_0\) signals produced by the modulator as it maps a sequence of encoder output code bits to a sequence of modulator output signals. The total number of 1's at the encoder input which produce all of the sequences of distance \(d\) is called the weight of the code at distance \(d\). A list of the weights, starting at the shortest (minimum) distance is called the weight structure of the code.

The search procedure for the code weight structure is described in Appendix A. A table of the code weight structure for the four shortest distances for each code considered is given in Table I.
<table>
<thead>
<tr>
<th>d</th>
<th>N_d</th>
<th>d</th>
<th>N_d</th>
<th>d</th>
<th>N_d</th>
<th>d</th>
<th>N_d</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>7</td>
<td>6</td>
<td>9</td>
<td>6</td>
<td>12</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>8</td>
<td>39</td>
<td>7</td>
<td>14</td>
<td>7</td>
<td>40</td>
<td>6</td>
<td>15</td>
</tr>
<tr>
<td>9</td>
<td>104</td>
<td>8</td>
<td>62</td>
<td>8</td>
<td>144</td>
<td>7</td>
<td>22</td>
</tr>
<tr>
<td>10</td>
<td>352</td>
<td>9</td>
<td>212</td>
<td>9</td>
<td>488</td>
<td>8</td>
<td>196</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>d</th>
<th>N_d</th>
<th>d</th>
<th>N_d</th>
<th>d</th>
<th>N_d</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>14</td>
<td>5</td>
<td>16</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>21</td>
<td>6</td>
<td>56</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>7</td>
<td>44</td>
<td>7</td>
<td>176</td>
<td>6</td>
<td>56</td>
</tr>
</tbody>
</table>

Table I. Weight Structure (First Four Entries) For Codes On 4-ary Channels.
4. **PERFORMANCE BOUNDS**

Upper bounds on the error probability performance of convolutional codes with MFSK systems can be found by using the union-Chernoff bound technique [6]. Results obtained from this method are usually about 1 dB higher than the exact result for bit error probabilities less than $10^{-3}$. In the union-Chernoff bound technique, we consider all candidate error paths remerging at a node on the correct (all-zero) path of the decoding trellis (state vs. time diagram). This is typically illustrated in Figure 11.
A union bound is employed because the events corresponding to choosing the various incorrect paths are not mutually exclusive. The calculation is facilitated if we overbound the probability of deciding on an incorrect path by

\[ P(A_1+A_2+A_3+\ldots) \leq P(A_1)+P(A_2)+P(A_3)+\ldots \]  

(10)

where \( A_1, A_2, A_3, \ldots \) are the error events corresponding to individual incorrect paths.

The importance of the Chernoff bound lies in the fact that the Chernoff bound is itself a moment generating function. (See Appendices B and C.) If the probability of error for a single use of a memoryless channel has a Chernoff bound \( B \), then the Chernoff bound for \( n \) uses of the channel is \( B^n \). In the decoding trellis, the probability of choosing an incorrect path of distance \( d \) is therefore Chernoff bounded by \( B^d \). All error paths of distance \( d \) have a collective total of \( N_d \) ones in the decoded bit sequence, where \( N_d \) is the code weight as defined in the preceding section. The union-Chernoff bound on the decoded bit error probability is given by

\[ P_b \leq \frac{1}{2} \sum_{d=d_{\text{min}}}^{\infty} N_d B^d \]  

(11-a)

for codes with binary inputs (\( K = 3 \) through \( K = 7 \)) and by

\[ P_b \leq \frac{1}{2} \sum_{d=d_{\text{min}}}^{\infty} \frac{N_d}{2} B^d \]  

(11-b)

for codes with quaternary inputs (dual-2 and triple-2). In equations (11), the factor of \( 1/2 \) in front of the summation is an additional tightening factor which can be applied to Chernoff bounds under fairly general conditions [7]. In equation (11-b), the code weight \( N_d \) is divided by a normalizing factor of 2 to take into account the fact that inputs to the dual-2 and triple-2 code are pairs of bits [8].

Equations (11) have been plotted in Figs. (12) and (13) using the Chernoff bound \( B \) derived in Appendices B and C for the Rayleigh fading and partial band Gaussian channels. The values of \( N_d \) are those given in the preceding section. All calculations are truncated after the first four terms. Furthermore, in the partial band case, it is assumed that all symbols on an error path must be jammed in order to produce an error. This is the "jammer state known" assumption, and it is
Fig. 12 — Union-Chernoff bound for Rayleigh fading channels
Fig. 13 — Union-Chernoff bound for partial band Gaussian noise channels
frequently used in analyses [1].

By comparing Figures (12) and (13) with Figures (1) and (2) we see that the coding improvement (with soft decision coding) at $10^{-5}$ error probability ranges between 27 dB and 35 dB from the weakest to the strongest codes. In the next section we consider the decoder complexity for the codes considered.
5. **DECODER COMPLEXITY**

The most important complexity consideration in a maximum likelihood decoder (employing the Viterbi algorithm) is the number of pairwise comparisons required in the decoding algorithm. This dictates the maximum speed for serial processing, or computational complexity for parallel processing.

For a binary convolutional encoder of constraint length $K$, the number of encoder states is $2^{K-1}$ and at each state a single comparison is required for each input bit. For dual-2 and triple-2 encoders of equivalent (binary) constraint length $K=4$ and $K=6$ respectively, the number of states is $2^{K-2}$, but at each state three comparisons are required to determine the largest of four competing path metrics. Since the encoder accepts two-bit inputs, the number of comparisons at each node per input bit is $3/2$.

In Table II the number of comparisons per bit is presented for each code considered in this report. Using this as a simple complexity measure, we see that the complexity of the decoder increases as the performance of the coding system improves. (The single exception to this is the triple-2 code which slightly outperforms the $K=6$ code, although having fewer comparisons per bit).

<table>
<thead>
<tr>
<th>Code</th>
<th># states</th>
<th>$#_{\text{comp./bit/state}}$</th>
<th>$#_{\text{comp./bit}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K=7$</td>
<td>64</td>
<td>1</td>
<td>64</td>
</tr>
<tr>
<td>$K=6$</td>
<td>32</td>
<td>1</td>
<td>32</td>
</tr>
<tr>
<td>Triple-2</td>
<td>16</td>
<td>$3/2$</td>
<td>24</td>
</tr>
<tr>
<td>$K=5$</td>
<td>16</td>
<td>1</td>
<td>16</td>
</tr>
<tr>
<td>$K=4$</td>
<td>8</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>Dual-2</td>
<td>4</td>
<td>$3/2$</td>
<td>6</td>
</tr>
<tr>
<td>$K=3$</td>
<td>4</td>
<td>1</td>
<td>4</td>
</tr>
</tbody>
</table>

*Table II. Decoder Complexity Measure (comparisons per bit)*
6. RESULTS AND CONCLUSIONS

In this report we have demonstrated the importance of coding in transmission of MFSK/FH signals on Rayleigh fading or partial band Gaussian interference channels. For a probability of error of $10^{-5}$ the coding gain of a $K=7$, rate-$1/2$ code on either of these channels is approximately 35 dB. For the weakest code considered ($K=3$), the coding gain is nearly 27 dB.

Using a simple measure of decoder complexity (comparisons per bit) it was found that complexity may be reduced with a resulting loss in performance. One case of special interest is the triple-2 code which has performance (at $10^{-5}$) which is within 1.5 dB of the best code considered ($K=7$ binary) but has complexity which is less than one half of that of the $K=7$ code.

Finally, it might be pointed out that the performance results for combined Rayleigh fading, worst case partial band Gaussian noise channels are implicit in the results given. It has been shown [9] that the worst case Gaussian jammer for a Rayleigh fading channel is a broadband jammer, so that the Rayleigh fading results are also applicable to the combined case.
REFERENCES


APPENDIX A. COMPUTATION OF CODE WEIGHT STRUCTURE

The weight structure of a convolutional code is readily available from the generating function for any given code. However, direct calculation of the generating function becomes difficult for codes having constraint lengths of five or more, since the number of states grows exponentially with the constraint length. An alternate approach is to implement the encoder and calculate the weights associated with the various distance error paths. The path errors are characterized by the codeword distances and the weights are determined from the number of input "ones" associated with each codeword distance. This approach is equivalent to evaluating the generating function at some selected finite sequence length.

A computer program was developed to compute the code weights. It does so by implementing the coder with a K-stage shift register and 2n mod-2 adders where n is the number of 4-ary output symbols of the encoder. The shift register is shifted by 1(or 2) bits at a time to allow binary (or quaternary) input data to be used. The mod-2 adder outputs are arranged in groups of 2 bits to provide n 4-ary output symbols. This program computes distances and weights of each error path for constraint lengths up to K=15 and input data sequence lengths up to 31 bits.

The convolutional encoder is specified by selecting the shift register length, the input symbol radix and the output symbol description. Each output symbol is determined from the output symbol radix and the bits in the shift register with tap connections which are used in each mod-2 summation. The tap specification is used to generate a mask which is logically "and-ed" with the data and mod-2 added to produce the appropriate bit in the output. The output distances are accumulated as each data sequence is shifted through to produce the codeword distance for that sequence. The number of "ones" in the data sequence is added to the weight accumulator for the corresponding codeword distance to determine the total weight. The program uses a modified exhaustive search of all possible patterns of the specified sequence length. Path errors begin when the path selected by the decoder differs from the correct (all-zero) path, so that sequences which begin with all "zeros" will have already been counted when the coder used the same data sequence with the leading "zeros" removed. For this reason, these data sequences (which are multiples of the shift register radix) are omitted. Another pattern can exist in the data sequence which does not contribute to the result, namely any data pattern which causes the shift register to enter the all "zeros" state. This can occur for data sequence lengths which exceed the constraint length of the decoder, and in essence cause the decoder to leave the correct path and return to it more than once. This is not a single path error and these data
are not included in the weight calculation.

After calculating the weights for each code word distance in the selected sequence length, the program lists the weights and input data bit patterns associated with all distances found. The truncated generating function can then be determined directly from these data. Table A-1 shows the results of this procedure for a K=7, R=1/2 code using a data sequence length of 18 bits. This code produces one quaternary output symbol for each binary input symbol. The taps used for each digit of the output symbol are listed in binary form and represent an encoder structure as shown in Figure 4 (in body of report). From Table A-1 we see that this encoder has weights 7, 39, 104, and 352 for distances 7, 8, 9, and 10 respectively, as shown in Table I of the report.*

* This code weight structure replaces a previously published code weight structure [4] which was found to contain numerical errors.
SEQUENCE LENGTH = 18 BITS  
NUMBER OF WEIGHTS LISTED = 4  
SHIFT REGISTER LENGTH = 7 BITS, RADIX = 2  
NUMBER OF OUTPUT SYMBOLS = 1, RADIX = 4  
SYMBOL # 1:  
  TAP # 1 = 0111111  
  TAP # 2 = 1101101

<table>
<thead>
<tr>
<th>DISTANCE</th>
<th>NUMBER OF PATHS</th>
<th>TOTAL NUMBER OF ONES (WEIGHT)</th>
</tr>
</thead>
<tbody>
<tr>
<td>7 BIT PATTERNS</td>
<td>1</td>
<td>11</td>
</tr>
<tr>
<td>8 DISTANCE</td>
<td>NUMBER OF PATHS</td>
<td>TOTAL NUMBER OF ONES (WEIGHT)</td>
</tr>
<tr>
<td>9 BIT PATTERNS</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>1111111</td>
<td>110111</td>
<td>110111</td>
</tr>
<tr>
<td>1000111</td>
<td>101111</td>
<td>101111</td>
</tr>
<tr>
<td>1011101</td>
<td>110111</td>
<td>110111</td>
</tr>
<tr>
<td>10111011</td>
<td>11001111</td>
<td>11001111</td>
</tr>
<tr>
<td>10 DISTANCE</td>
<td>NUMBER OF PATHS</td>
<td>TOTAL NUMBER OF ONES (WEIGHT)</td>
</tr>
<tr>
<td>11 BIT PATTERNS</td>
<td>19</td>
<td>62</td>
</tr>
</tbody>
</table>

Table A-1 Weight structure for K=7 code.
APPENDIX B: CHERNOFF BOUND FOR BINARY FSK ERROR PROBABILITY ON A RAYLEIGH FADING CHANNEL

The optimal receiver for noncoherent binary FSK on a Rayleigh fading channel consists of two pairs of matched filters, one pair matched to each frequency tone. In Figure B-1, we see that each matched filter pair consists of one filter matched to the in-phase component and one filter matched to the quadrature component. The output of each matched filter pair is a pair of i.i.d. random variables with probability density $G(0, \sigma^2)$. For the "correct" filter pair (assuming signal 1 is sent) the outputs $x_1$ and $y_1$ have variance $\sigma_1^2 = \frac{N_0 + E}{2}$, and for the "incorrect" filter with outputs $x_2$ and $y_2$, the output variance is $\sigma_2^2 = \frac{N_0}{2}$. (Here, $E$ is the average energy of the received signal).

Fig. B-1 – Optimal receiver for noncoherent BFSK
For the receiver configuration shown in Figure B-1, the output statistic is \( w \), the difference between the squares of the envelopes of each filter pair output. If signal 1 is sent, an error will occur if \( w > 0 \). The resulting probability of error \( P_e \) is \( P_r \{w > 0\} \) so that

\[
P_e = \int_{0}^{\infty} p(w) \, dw \tag{B-1}
\]

where \( p(w) \) is the probability density function for the output random variable \( w \). To obtain a Chernoff bound on \( P_e \) we overbound the unit step \( u(w) = 1 \) (for \( w > 0 \)) by the exponential function \( e^{\lambda w}, -\infty < w < \infty \), where \( \lambda \) is a free parameter, \( \lambda > 0 \). This is shown in Figure B-2.

Fig. B-2 — Exponential overbound of unit step indicator function
Thus we may bound the error probability by a function of \( \lambda \),

\[
P_e \leq B(\lambda) = \int_{-\infty}^{\infty} e^{\lambda w} p(w) \, dw \tag{B-2}
\]

The right-hand side of Eq. (B-2) is the moment generating function for the random variable \( w \). The fact that the error bound is a moment generating function is the reason why the error bound for \( N \) uses of a memoryless channel is simply the \( N \)th power of the bound for a single use. Tightening this bound by selecting the free parameter \( \lambda \) yields the Chernoff bound \( B \):

\[
B = \min_{\lambda > 0} B(\lambda) = \min_{\lambda > 0} \int_{-\infty}^{\infty} e^{\lambda w} p(w) \, dw = \min_{\lambda > 0} e^{\lambda w} \tag{B-3}
\]

where the wiggly overbar indicates the expectation operator.

For this particular channel, we may perform this averaging over the individual (independent) components of

\[
w = x_2^2 + y_2^2 - x_1^2 - y_1^2.
\]

The random variable \( x_1 \) is \( \mathcal{N}(0, \sigma_1^2) \), so that

\[
e^{-\lambda x_1^2} = \frac{1}{\sqrt{2\pi\sigma_1^2}} \int_{-\infty}^{\infty} e^{-\lambda x_1^2} e^{-2\sigma_1^2 x_1^2} \, dx_1 = \frac{1}{\sqrt{1+2\sigma_1^2 \lambda}} \tag{B-4}
\]

Similarly

\[
e^{-\lambda y_1^2} = \frac{1}{\sqrt{1+2\sigma_1^2 \lambda}} \tag{B-5}
\]

and

\[
e^{\lambda x_2^2} = e^{\lambda y_2^2} = \frac{1}{\sqrt{1-2\sigma_2^2 \lambda}} \tag{B-6}
\]
Thus
\[
B(\lambda) = e^{\lambda w} = \left[ \frac{1}{1+2\sigma_1^2 \lambda} \right] \left[ \frac{1}{1-2\sigma_1^2 \lambda} \right]
\]
\[
= \left[ \frac{1}{1+(E+N_0)\lambda} \right] \left[ \frac{1}{1-N_0 \lambda} \right] \tag{B-7}
\]

Letting \( \lambda N_0 = \lambda' \) and \( 1 + \frac{E}{N_0} = \beta \)

We find
\[
B(\lambda) = \frac{1}{(1+\beta\lambda')(1-\lambda')} = \frac{1}{1+(\beta-1)\lambda' - \beta\lambda'^2} \tag{B-8}
\]

Differentiating the denominator with respect to \( \lambda' \) and equating to zero, we find
\[
\lambda' = \frac{\beta-1}{2\beta} = \frac{E/N_0}{2(1+E/N_0)} \tag{B-9}
\]

Applying this to \( B(\lambda) \), we obtain the Chernoff bound for the Rayleigh fading channel.

\[
P_e \leq B = \min_{\lambda > 0} B(\lambda) = \frac{1+E/N_0}{\left(1+E/N_0\right)^2} = \frac{4(1+E/N_0)}{(2+E/N_0)^2} \tag{B-10}
\]
APPENDIX C: CHERNOFF BOUND FOR BINARY FSK ERROR PROBABILITY ON A PARTIAL BAND GAUSSIAN CHANNEL

The optimal receiver for noncoherent binary FSK on a partial band Gaussian channel is the same as shown in Figure B-1. For partial band Gaussian jamming, there is a probability \( \rho \) of hopping into a jammed portion of the band where the noise density is \( N_0/\rho \); otherwise, the transmission is noise free. During jamming, the (independent) output statistics (for the receiver in Figure B-1) have probability density functions as follows:

\[
\begin{align*}
    x_1 & : G\left(\sqrt{E} \cos \theta, \frac{N_0}{2\rho}\right) \\
    y_1 & : G\left(-\sqrt{E} \sin \theta, \frac{N_0}{2\rho}\right) \\
    x_2 & : G\left(0, \frac{N_0}{2\rho}\right) \\
    y_2 & : G\left(0, \frac{N_0}{2\rho}\right)
\end{align*}
\]

Following the approach used in Appendix B, the Chernoff bound for the probability of error (with worst case jamming) is

\[
B = \max_{0<\rho\leq1} \min_{\lambda>0} \left| \rho e^{\lambda (x_2^2+y_2^2-x_1^2-y_1^2)} \right| 
\]

(C-1)

Letting \( N_0' = \frac{N_0}{\rho} \), we get

\[
\begin{align*}
    e^{-\lambda x_1^2} & = \frac{1}{\sqrt{1+\lambda N_0'}} \quad e^{-\frac{E \cos^2 \theta}{1+\lambda N_0'}} \\
    e^{-\lambda y_1^2} & = \frac{1}{\sqrt{1+\lambda N_0'}} \quad e^{-\frac{E \sin^2 \theta}{1+\lambda N_0'}}
\end{align*}
\]

(C-2)  
(C-3)
Letting $\lambda' = \lambda N_o$, we find

$$e^{\lambda(x_2^2 + y_2^2 - x_1^2 - y_1^2)} = \frac{1}{\sqrt{1-\lambda N_o'}} e^{\lambda' \frac{E}{N_o}}$$

Thus

$$P_e \leq B = \max_{0 < \rho \leq 1} \min_{\lambda' > 0} \left\{ \rho \frac{1}{1-\lambda'^2} e^{1+\lambda'} \rho \frac{E}{N_o} \right\}$$

Taking $\frac{d}{d\rho}$ and equating to zero, we find

$$\rho = \frac{1}{1+\lambda'} \frac{E}{N_o}$$

so that

$$P_e \leq B = \min_{\lambda' > 0} \frac{1}{\frac{E}{N_o}} e^{-\frac{1}{\lambda' (1-\lambda')}}$$

Taking $\frac{d}{d\lambda'}$, and equating to zero we find

$$\lambda' = \frac{1}{2}$$
Thus, the Chernoff bound for the partial band Gaussian channel is

\[ P_e \leq B = \frac{4e^{-1}}{E/N_0} \]  

This bound is effective in the range \( E/N_0 > 3 \), which is the range of interest in this application. For \( E/N_0 \leq 3 \), broadband jamming (\( \rho = 1 \)) is the worst case jamming.