# Trellis-Coded Permutation Modulation for Improved Performance of Narrowband Noncoherent FSK

By

Xu Lin

Thesis submitted to the Faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

IN

ELECTRICAL ENGINEERING

APPROVED:

Brian D. Woerner, Chair

Jeffrey H. Reed

F. Gail Gray

May, 1996

Blacksburg, Virginia

**Keywords:** Coding, Modulation, Permutation, Trellis, Noncoherent

c 2

LD
5655
V855
1996
L567
c.2

# TRELLIS-CODED PERMUTATION MODULATION FOR IMPROVED PERFORMANCE OF NARROWBAND NONCOHERENT FSK

By

Xu Lin

Committee Chairman: Dr. Brian D. Woerner

(ABSTRACT)

Noncoherent modulation is an important technique in wireless communication systems. Although noncoherent modulation usually does not perform as well as its coherent counterpart, it is practical and useful in some applications, such as paging systems. In this thesis we investigate ways to improve the performance of noncoherent FSK in narrowband channels.

It is shown that FSK permutation modulation has better spectral efficiency than conventional FSK modulation, but with the tradeoff on reduced energy efficiency. To overcome this problem, we apply trellis-coded modulation (TCM) , which is a combined technology of convolutional coding and modulation, to FSK permutation. TCM was originally designed for coherent modulation. The application of TCM to permutation modulation retains the fundamental concepts of TCM. The simulation results show that trellis-coded permutation modulation provides a better combination of energy efficiency and spectral efficiency than conventional FSK noncoherent demodulation.

# Acknowledgments

My sincere thanks to my advisor, Dr. Brian Woerner, for his understanding and patience, for his constant support and encouragement. His knowledge and guidance are essential to the completion of this thesis. Without his help, I might not be able to achieve this final goal in school. I feel so fortunate to have him as my advisor.

I would like to thank my committee member Dr. Jeffrey H. Reed and Dr. F. Gail Gray, for their time and energy during the process of constructing this thesis. Their comments and guidance are invaluable to my thesis. I would also like to thank all the faculty, staff and students at MPRG for making it an excellent place for doing research.

My biggest thanks to my parents for their endless love, encouragement, inspiration and support; for their significant sacrifices during my study in school. They are the greatest parents one can ever ask for. Thanks also to my best friend Munling, for her support and help during the editing of this thesis.

My special thanks to my husband, Weimin, for his love, understanding and support. Without his encouragement, I might not be able to go to school and finish the program in this relatively short period of time. My final thanks to my son Kevin, for his love and for his sacrifices during my course of study.

# Table of Contents


Abstract ..... ii

Acknowledgments ..... iii

List of Figures ..... vi

List of Tables ..... ix

1

1.1 Background ..... 1

1.2 Permutation Modulation ..... 2

1.3 Convolutional Coding ..... 4

1.4 Trellis-Coded Modulation ..... 7

1.5 Outline of the Thesis ..... 9

**Chapter 2 Introduction to Noncoherent Modulation** ..... 11

2.1 Introduction ..... 11

2.2 Performance of Noncoherent FSK in AWGN ..... 11

2.2.1 Binary FSK ..... 12

2.2.2 M-ary FSK ..... 18

2.3 Simulation of M-ary Noncoherent FSK ..... 20

2.3.1 Complex Envelope Representation of Bandpass Signals ..... 21

2.3.2 Representation of the M-ary Noncoherent FSK Signals ..... 22

2.3.3 Simulation of M-ary FSK Signals ..... 23

2.4 Conclusion ..... 28

**Chapter 3 FSK Permutation Modulation** ..... 29

3.1 General Description ..... 29

3.2 Simulation of Several PFSK Cases ..... 33

3.3 Comparison of Different PFSK ..... 38

3.4 Conclusion ..... 40

**Chapter 4 Introduction to Trellis-Coded Modulation ..... 41**

4.1 Introduction ..... 41

4.2 Fundamentals of TCM ..... 42

4.2.1 Signal Constellation ..... 42

4.2.2 Set Partitioning ..... 44

4.3 Trellis-Coded Modulation ..... 46

4.4 Decoding of TCM ..... 51

4.4.1 The Branch Metric ..... 52

4.4.2 Viterbi Algorithm ..... 55

4.5 Conclusion ..... 57

**Chapter 5 Trellis Coding for Permutation Modulation ..... 59**

5.1 Introduction ..... 59

5.2 Signal Set of Permutation Modulation ..... 60

5.3 Set Partitioning of Permutation Modulation ..... 64

5.3.1 $\binom{6}{3}$ Permutation Modulation ..... 64

5.3.2 $\binom{8}{4}$ Permutation Modulation ..... 66

5.4 Selection of Partitions by Convolutional Codes ..... 70

5.5 Decoding of TCPM ..... 73

5.6 Simulation of TCPM ..... 74

5.7 Conclusion ..... 82

**Chapter 6 Conclusion and Future Work ..... 84**

6.1 Conclusion ..... 84

6.2 Future Works ..... 85

**Reference ..... 86**

Vita ..... 89

# List of Figures


Figure 1-1 Illustration of Permutation Modulation ..... 4
Figure 1-2 A Rate-1/2 Linear Convolutional Encoder ..... 5
Figure 1-3 Trellis Representation for 1/2 Convolutional Encoder ..... 6
Figure 1-4 Example Signal Constellation and Their Minimum Euclidean Distances 7

Figure 2-1 Noncoherent FSK Receiver ..... 13
Figure 2-2 Structure of Optimum Noncoherent FSK Receiver ..... 16
Figure 2-3 Probability of Bit Error vs. Eb/No for Coherent and Noncoherent Detection of FSK ..... 18
Figure 2-4 Probability of Bit Error vs. Eb/No for Noncoherent M-ary FSK ..... 20
Figure 2-5 BER vs. Eb/No for Binary Noncoherent FSK ..... 26
Figure 2-6 BER vs. Eb/No for 4-ary Noncoherent FSK ..... 26
Figure 2-7 BER vs. Eb/No for 8-ary Noncoherent FSK ..... 27
Figure 2-8 BER vs. Eb/No for 16-ary Noncoherent FSK ..... 27

Figure 3-1 Transmission System for $\binom{m}{2}$ PFSK ..... 31
Figure 3-2 Error Probability for $\binom{6}{3}$ PFSK ..... 37
Figure 3-3 Error Probability for $\binom{8}{4}$ PFSK ..... 38
Figure 3-4 Performance Comparison of PFSK Alphabets ..... 39

Figure 4-1 Several Signal Constellations ..... 43
Figure 4-2 Set Partitioning of an 8-PSK Signal Set ..... 45
Figure 4-3 An Ungerboeck Encoder ..... 47

| | | |
|---|---|---|
| Figure 4-4 | Partitions for the Ungerboeck Encoder in Figure 4-5 | 48 |
| Figure 4-5 | Four State Ungerboeck Encoder | 48 |
| Figure 4-6 | Trellis Diagram for the Uncoded and Coded System | 49 |
| Figure 4-7 | Original Trellis Diagram | 56 |
| Figure 4-8 | The Viterbi Algorithm Iteratively Finds the Path with Minimum Path Metric | 57 |
| Figure 5-1 | Signal Patterns of $\binom{6}{3}$ FSK Permutation Modulation | 60 |
| Figure 5-2 | Signal Groups with Distance of 6 | 64 |
| Figure 5-3 | Signal Groups with Distance of 2 | 65 |
| Figure 5-4 | Set Partitioning of $\binom{6}{3}$ Permutation Modulation | 65 |
| Figure 5-5 | Groups with Distance of 8 | 68 |
| Figure 5-6 | Groups with Distance of 4 | 69 |
| Figure 5-7 | Set Partitioning of $\binom{8}{4}$ Permutation Modulation | 69 |
| Figure 5-8 | Groups with Distance of 2 | 69 |
| Figure 5-9 | Eight-State Ungerboeck Encoder | 70 |
| Figure 5-10 | Selection of $\binom{6}{3}$ Partitions by a 8-State Convolutional Encoder | 71 |
| Figure 5-11 | Set Partitioning of $\binom{8}{4}$ Signal Set Associated with Encoder in Figure 5-8 | 72 |
| Figure 5-12 | Block Diagram for Simulation of Trellis-Coded Permutation Modulation | 75 |
| Figure 5-13 | Assignment of the Subsets on the Trellis Diagram of the Encoder | 77 |
| Figure 5-14 | BER vs. Eb/No for Trellis-Coded $\binom{6}{3}$ Permutation Modulation | 80 |

Figure 5-15 BER vs. Eb/No for Trellis-Coded $\binom{8}{4}$ Permutation Modulation ..... 81

Figure 5-16 New Energy Efficiency & Spectral Efficiency Plot ..... 82

# List of Tables


Table 3-1 Spectral Efficiency of PFSK ..... 30

Table 3-2 Mapping of Each of the 16 Symbols to the Transmitted Characters ..... 35

Table 5-1 Signal Sets of $\binom{6}{3}$ FSK Permutation Modulation ..... 61

Table 5-2 Signal Patterns of $\binom{8}{4}$ FSK Permutation Modulation ..... 62

Table 5-3 Modulated Signal Patterns of Coded Bits ..... 78

# Chapter 1 Introduction

## 1.1 Background

In a digital communication system, both coherent and noncoherent techniques may be used to demodulate signals. It is well known that coherent demodulation has the advantage of having higher energy efficiency than noncoherent demodulation, i.e., for a given bit error rate coherent demodulation requires less transmitter power than noncoherent demodulation. However, coherent receivers are relatively complex and expensive since they require coherent detection, and sometimes it is not convenient or is extremely difficult (e.g., when jamming or fading is present) to establish a reference signal at the receiver which is coherent with the received carrier. In this case noncoherent digital modulation is practical and useful. One of the most commonly used modulation formats in digital communication systems that uses noncoherent demodulation is Frequency Shift Keying (FSK) technique.

FSK is used in a number of military communication applications. Frequency-hop spread spectrum (FH) is one of the examples [22]. Because of the difficulty of building truly coherent frequency synthesizers, most FH systems use noncoherent M-ary FSK modulation. FSK is also commonly utilized in commercial paging systems. Since the bandwidth resource in that particular application area is very limited. Therefore, improving spectral efficiency for paging systems is an important issue.

An FSK signal is generated by shifting the carrier by a certain amount to reflect the digital information that is being transmitted. A typical binary FSK signal is represented as follows:

$$\left.\begin{aligned} s_1(t) &= A\cos(2\pi f_c t) \\ s_2(t) &= A\cos[2\pi(f_c + \Delta f)t] \end{aligned}\right\} \quad 0 \le t \le T_b, \tag{1-1}$$

where $T_b$ is the duration of one bit,

and $\Delta f = 1/T_b$ is the minimum frequency spacing for noncoherent orthogonal signals, $s_1(t)$ and $s_2(t)$ are, thus, said to be orthogonal since

$$\int_0^{T_b} s_1(t)s_2(t)dt = 0 \tag{1-2}$$

We consider orthogonal signals for this research because they have the best performance.

The FSK signal can be characterized as one of the two different types, depending on the method used to generate the FSK signal. One type is called discontinuous-phase FSK. It is generated by switching the transmitter output line between waveforms that are discontinuous at the switching times. The other type is generated by feeding the data signal into a frequency modulator. It is called continuous-phase FSK [6]. Minimum-shift keying (MSK) is a special case of continuous-phase FSK, with the minimum modulation index that produces orthogonal signaling. MSK is a bandwidth conservation technique. Gaussian Filtered MSK (GMSK) modulation is widely used in wireless standards such as the Global System for Mobile Communications (GSM) and the Digital European Cordless Telephone (DECT). It has the advantage of being a constant-amplitude signal and, consequently, can be amplified with Class C amplifiers without distortion [6].

To evaluate the performance of a modulation scheme energy efficiency and spectral efficiency are used. Energy efficiency is defined as energy per bit required for a given error probability, while spectral efficiency is given by the number of bits per second of data that can be supported per Hertz of bandwidth.

## 1.2 Permutation Modulation

There is a need for noncoherent modulation techniques which increase the spectral efficiency in order to accommodate the increasingly limited bandwidth

resources in the wireless area. Permutation modulation is one of the transmission schemes which can improve spectral efficiency. It is first introduced by David Slepian in [14], where the general permutation modulation was described. What we are interested here is a subclass of the general permutation modulation, the so-called FSK permutation modulation (PFSK). It applies the general permutation modulation to a multifrequency modulation scheme. The energy is transmitted simultaneously on $n$ frequencies out of $m$, instead of just on one frequency which is the case of M-ary FSK. The whole signal set is sent by transmitting the different combinations of the carrier frequencies.

It is known that the bandwidth is proportional to the number of carrier frequencies times the frequency separation, which is $1/T_b$ for noncoherent orthogonal signals with minimum frequency spacing. As a result, the bandwidth requirement is determined by the number of carrier frequencies.

It can be shown that FSK permutation modulation increases the spectral efficiency over conventional FSK. For the same signal set and same symbol length, less bandwidth is needed if using PFSK than using the traditional M-ary FSK technique. As an example, consider the 16-ary FSK scheme, where 4 information bits are sent per symbol and 16 carrier frequencies are required. However, only 6 carrier frequencies are sufficient if using $\binom{6}{3}$ FSK permutation modulation. Since in this case, we propose that during each symbol interval, a combination of three carriers are ON while the other three are OFF, see Figure 1-1. The total number of distinct combinations is

$$\binom{6}{3} = \frac{6!}{3!(6-3)!} = 20 \tag{1-3}$$

which is sufficient for transmitting a signal with 4 bits per symbol. It is easily seen that using FSK permutation modulation improves the performance in terms of spectral efficiency.

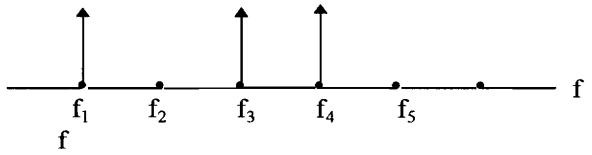



Figure 1-1. Illustration of Permutation Modulation

FSK permutation modulation has the same features as the general permutation, such as constant energy for each symbol and efficient spectrum utilization. The latter feature is significant for bandlimited applications, such as the paging industry where limited bandwidth is a big concern.

## 1.3 Convolutional Coding

It is shown in [4] that with permutation modulation the energy efficiency decreases. For a given bit error rate permutation modulation needs more power than the traditional FSK modulation. A tradeoff between spectral efficiency and energy efficiency exists.

In communications, error-correction coding is a powerful technique for achieving desired bit error rates at reduced transmitter power levels. Error-correction codes are used to format the transmitted information so as to increase its immunity to noise. It is accomplished by inserting controlled redundancy into the transmitted data stream. In this way it allows the receiver to detect and possibly correct errors [5]. The two most popularly used channel codes are block codes and convolutional codes. We'll focus on convolutional codes in this thesis because of their use in trellis-coded modulation.

In convolution coding, the redundancy is introduced into a data stream through the use of a linear shift register. A typical rate-1/2 linear convolutional encoder is shown in Fig. 1-2.

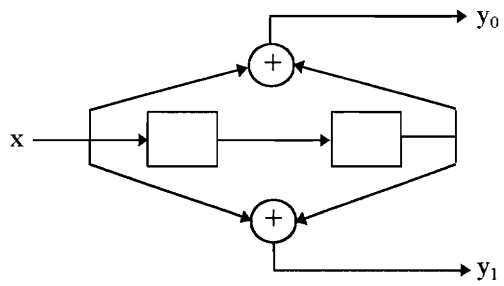



Figure 1-2. A Rate -1/2 Linear Convolutional Encoder [14]

A convolutional code is defined by the number of stages in the shift register, the number of outputs, and the connections between the shift register and the modulo-2 adders [14]. The state of the encoder is defined to be the contents of the shift register and is completely determined by the previous information bit inputs. The encoder of Figure 1-2 has four possible states corresponding to all possible contents of the binary two-states shift register. A convolutional code is fully defined by a set of so called generator polynomials $g_j^{(i)}$ (i = 0, 1, 2, ...; j = 0, 1, 2,...). $g_j^{(i)}$ is obtained for the ith output of an encoder by applying a single 1 at the jth input followed by a string of zeros [5]. Strings of zeros are applied to all other inputs. The polynomial for the encoder in Figure 1-2 are as follows

$$g^{(0)} = (111)$$
$$g^{(1)} = (101) \qquad (1\text{-}4)$$

A convolutional coder can be represented by a state transition diagram and by a trellis diagram, which is the basis of the trellis coded modulation. Figure 1-3 illustrates the trellis diagram for the code of Figure 1-2.

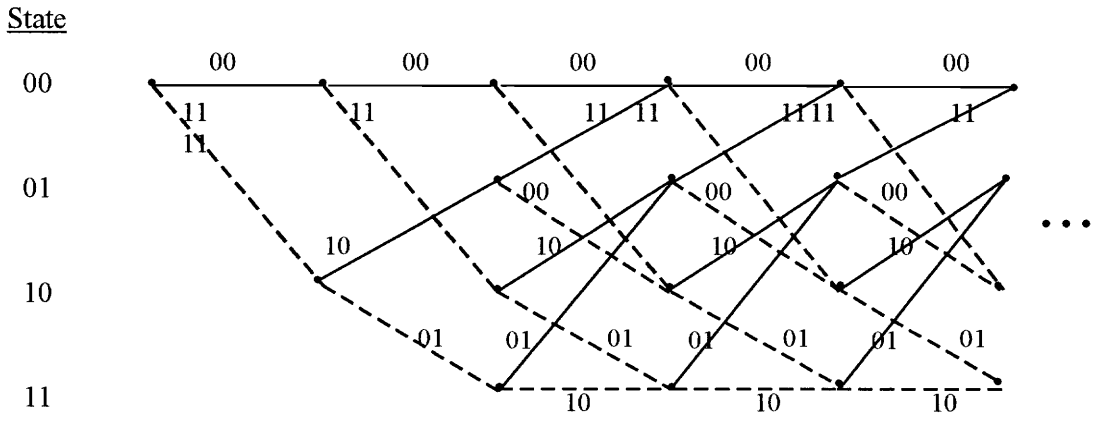



Fig 1-3. Trellis Representation for 1/2 Convolutional Encoder [14]

The encoder represented by Figure 1-2 has four states which are represented by 00, 01, 10 and 11. The branches of the trellis diagram are labeled with the output bits corresponding to the associated state transitions. If the first encoder input were 1, the encoder would move along the dashed branch out of state 00 arriving at state 10. In this case the output is 11. The second encoder input causes the encoder to move to the right one more branch and to output the associated branch label. This process continues as long as desired. Every code word in a convolutional code is associated with an unique path. The decoder is based on this fact. The most popular algorithm for decoding convolutional codes is the Viterbi algorithm. This is an elegant method for performing maximum likelihood decoding. The function of a maximum-likelihood decoder is to find the code sequence which was most likely to have been transmitted given the received channel output sequence. More details will be given in Chapter 4.

## 1.4 Trellis-Coded Modulation

In classical digital communication systems, the implementation of modulation and error-correction coding are done separately. A conventional multilevel modulator maps m binary symbols (bits) into one of $M = 2^m$ possible transmitted signals, and the demodulator recovers the m bits by making an independent M-ary nearest-neighbor decision on each signal received [1]. If all signals are equally likely to be transmitted, the maximum likelihood receiver selects the signal that is closest in Euclidean distance to the received signal. Euclidean distance is defined as the distance between a pair of distinct signals in a constellation diagram. Minimum Euclidean distance is the shortest distance between any pair of signals in the constellation diagram. Figure 1-4 shows some commonly used QAM and PSK constellations and their minimum Euclidean distance.

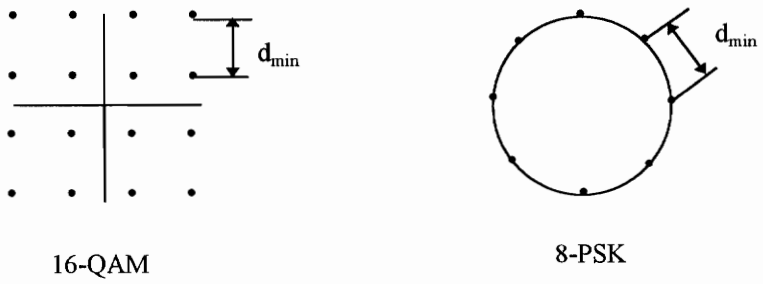



Figure 1-4. Example Signal Constellations and Their Minimum Euclidean Distances

As mentioned in 1.3, error-correction coding is accomplished by appending redundant check symbols. Redundancy is measured in terms of the code rate R, the ratio of the number of information bits to the total number of transmitted bits. If the input information transmission rate is the same in the coded system as the uncoded system, then the symbol transmission rate for the coded system must be $R^{-1}$ times that of the uncoded system ($R<1$). This requires more bandwidth, which is sometimes difficult to achieve. If bandwidth is not allowed to expand, the information rate must be reduced. One possibility to compensate for the rate loss is enlarging the size of the modulation signal set. This alternative leads to the use of nonbinary modulation ($M>2$). However, when modulation and error-correction coding are performed in the classical independent way, unsatisfactory results are expected. With an expanded modulation signal set, the signal spacing becomes smaller, i.e. the Euclidean distance is reduced. Since the demodulator determines a transmitted signal which has the shortest Euclidean distance with the received signal, it is expected that the error rate at the demodulator of the coded system exceeds that of the uncoded system. As a result, error performance may not be improved even with the coded system.

Trellis-coded modulation (TCM), introduced by Ungerboeck [1], integrates coding and modulation which provides a powerful, additional alternative. TCM uses convolutional codes and multidimensional signal constellations to provide reliable, high data rate communication over bandwidth-limited channel.

TCM schemes employ redundant nonbinary modulation in combination with a finite-state encoder which governs the selection of modulation signals to generate coded signal sequences. In the receiver, the received signals are decoded by a soft-decision maximum-likelihood sequence decoder [1]. Ungerboeck has shown that a simple four-state TCM scheme can improve the performance of digital transmission against additive noise by 3 dB, compared to conventional uncoded modulation. With more complex TCM schemes, the coding gain can reach 6dB or more. These gains are obtained without bandwidth expansion or reduction of the effective information rate as required by traditional error-correction schemes.

Ungerboeck's original TCM scheme was designed to work with coherent modulation. As we mentioned before, in some wireless areas especially in the paging industry, noncoherent modulation is important. This motivates us to explore the application of Ungerboeck's TCM to noncoherent modulation. In this thesis we explore a form of trellis-coded permutation modulation, which combines convolutional coding and permutation modulation. The essential idea in Ungerboeck's version of TCM has been employed in our research. Signal-set expansion is used to provide redundancy for coding, and signal-mapping functions are designed jointly so as to maximize the minimum Euclidean distance between coded signals, at the same information rate, bandwidth, and signal power. Several examples, such as trellis-coded $\binom{6}{3}$ and $\binom{8}{4}$ permutation, are examined and the simulation results are compared with uncoded FSK schemes and permutation modulation schemes. It is shown that using this new approach, modest performance improvements may be possible, and it is possible to achieve some combination of energy and spectral efficiency which are not otherwise achievable using noncoherent modulation. This new technology might be possibly applied to paging systems.

## 1.5 Outline of the Thesis

The rest of this thesis is organized as follows. Chapter 2 reviews the principles of noncoherent modulation. In particular, binary FSK and M-ary FSK are studied, followed by the simulation results for FSK. Permutation modulation is discussed in Chapter 3. Examples for $\binom{6}{3}$ and $\binom{8}{4}$ permutation are studied and simulation and analysis results in AWGN are provided. An energy and spectral efficiency graph is also

presented in this chapter. In Chapter 4 trellis-coded modulation is discussed in some detail. An example of Ungerboeck's TCM schemes is described. Chapter 5 develops the idea of trellis-coded permutation modulation. Some example codes are studied and simulation results are presented. In addition, a new graph of spectral efficiency and energy efficiency is provided. Chapter 6 concludes the thesis.

# Chapter 2 Introduction to Noncoherent Modulation

## 2.1 Introduction

Noncoherent modulation schemes do not require a coherent phase reference at the receiver to demodulate the signal. Although noncoherent schemes perform worse in terms of error probability for a given signal-to-noise ratio than do their coherent counterparts, there are situations where noncoherent schemes may be employed instead of a coherent signaling method. For example, the channel may not allow the use of a coherent signaling scheme, or the carrier phase may be unknown at the receiver with no attempt made to estimate its value, or the receiver may be much simpler since the carrier reference required in the coherent system need not be generated. Therefore, noncoherent modulation is an important technique used in communication systems.

For FSK signals which are sufficiently spaced in the frequency domain, it is possible to demodulate with no phase information. This is called a noncoherent FSK modulation scheme. FSK modulation with noncoherent reception is one of the most frequently used noncoherent modulation scheme in communication systems. It is cheap to build and robust for wireless applications. In this chapter, the performance of binary FSK and M-ary FSK in terms of error probability in the presence of additive white Gausssian noise is evaluated and compared with their coherent counterparts. It will be shown that noncoherent FSK requires only slightly higher $E_b/N_o$ than that for coherent FSK. This explains why in practice, many FSK receivers use noncoherent detection. Simulation results for M-ary FSK are also provided, which are matched to theoretical results.

## 2.2 Performance of Noncoherent FSK in AWGN

### 2.2.1 Binary FSK

Consider binary FSK with two signals:

$$\begin{aligned} s_1(t) &= \sqrt{2P}\cos(2\pi f_1 t + \phi_1) \\ &= \sqrt{2\mathrm{P}}\cos(2\pi f_1 t)\cos(\phi_1) - \sqrt{2P}\sin(2\pi f_1 t)\sin(\phi_1) \end{aligned} \tag{2-1}$$

$$\begin{aligned} s_2(t) &= \sqrt{2P}\cos(2\pi f_2 t + \phi_2) \\ &= \sqrt{2\mathrm{P}}\cos(2\pi f_2 t)\cos(\phi_2) - \sqrt{2P}\sin(2\pi f_2 t)\sin(\phi_2) \end{aligned} \tag{2-2}$$

where $f_1$ and $f_2$ are the two carrier frequencies which are assumed to be spaced far enough to make sure that $s_1(t)$ and $s_2(t)$ are orthogonal, $\phi_1$ and $\phi_2$ are unknown random phases with uniform distribution:

$$p_{\phi_1}(x) = p_{\phi_2}(x) = \begin{cases} 1/2\pi & x \in [0,\ 2\pi] \\ 0, & \text{else} \end{cases} \tag{2-3}$$

It is apparent that no matter what the two phases are, the signals can be expressed as a linear combination of the four basis functions:

$$\begin{aligned} f_{1c}(t) &= \sqrt{2/T}\cos(2\pi f_1 t), \quad \mathrm{f}_{1s}(t) = -\sqrt{2/T}\sin(2\pi f_1 t) \\ f_{2c}(t) &= \sqrt{2/T}\cos(2\pi f_2 t), \quad \mathrm{f}_{2s}(t) = -\sqrt{2/T}\sin(2\pi f_2 t) \end{aligned} \tag{2-4}$$

Therefore, the signals in (2-1) and (2-2) have the representation:

$$s_1(t) = \sqrt{E_b}\cos(\phi_1) f_{1c}(t) + \sqrt{E_b}\sin(\phi_1) f_{1s}(t) \tag{2-5}$$

$$s_2(t) = \sqrt{E_b}\cos(\phi_{21}) f_{2c}(t) + \sqrt{E_b}\sin(\phi_2) f_{2s}(t) \tag{2-6}$$

It is seen that the correlation with these basis functions are sufficient statistics for deciding which signal is more likely transmitted. The structure of a receiver with this idea is shown in Fig. 2-1.

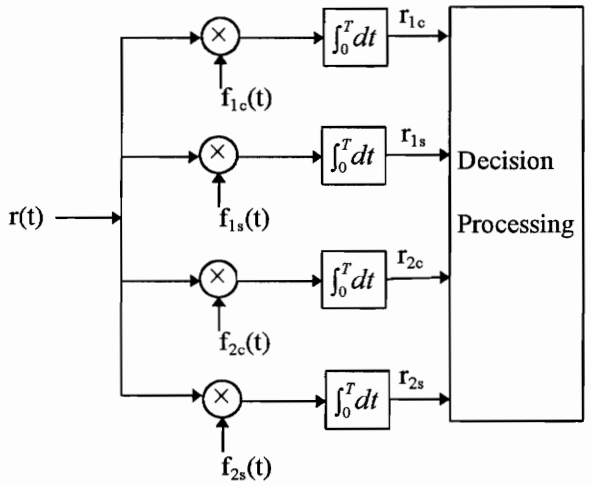



Figure 2-1. Noncoherent FSK Receiver

where $r_{1c}(t)$, $r_{1s}(t)$, $r_{2c}(t)$ and $r_{2s}(t)$ are the decision statistics for the in-phase and quadrature components of $s_1(t)$ and $s_2(t)$ respectively.

According to MAP decision rule, the receiver will decide signal 1 is transmitted if

$$P[s1]p(r/s_1) \geq P[s2/]p(r/s_2) \tag{2-7}$$

where $p(r/s_1)$ and $p(r/s_2)$ are the conditional pdfs of the decision statistics given by $s_1$ or $s_2$ are sent, $P_r(s_1)$ and $P_r(s_2)$ are the source statistics or a prior statistics, which are considered to be equally likely, that is

$$P[s_1] = P[s_2] \tag{2-8}$$

(2-7) is therefore becomes

$$p(r/s_1) = p(r/s_2) \tag{2-9}$$

which is the maximum likelihood decision rule (ML rule).

It is shown [7] that the conditional pdf of the received signal is

$$p(r/s_1,\phi_1) = \frac{1}{\pi N_0}\exp\left[-\frac{(r_{1c}-\sqrt{E}\cos\phi_1)^2+(r_{1s}-\sqrt{E}\sin\phi_1)^2}{N_0}\right]\times\frac{1}{\pi N_0}\exp\left[-\frac{{r_{2c}}^2+{r_{2s}}^2}{N_0}\right] \tag{2-10}$$

Similarly,

$$p(r/s_2,\phi_2) = \frac{1}{\pi N_0}\exp\left[-\frac{(r_{2c}-\sqrt{E}\cos\phi_2)^2+(r_{2s}-\sqrt{E}\sin\phi_2)^2}{N_0}\right]\times\frac{1}{\pi N_0}\exp\left[-\frac{{r_{1c}}^2+{r_{1s}}^2}{N_0}\right] \tag{2-11}$$

We need to evaluate (2-9) for the ML decision. It is equivalent to evaluate

$$\frac{1}{2\pi}\int_0^{2\pi} p(r/s_1,\phi_1)d\phi_1 \geq \frac{1}{2\pi}\int_0^{2\pi} p(r/s_2,\phi_2)d\phi_2 \tag{2-12}$$

Equivalently, (2-12) can be expressed as

$$\frac{1}{2\pi}\int_0^{2\pi}\exp\left[\frac{2\sqrt{E}r_{1c}\cos\phi_1+2\sqrt{E}r_{1s}\sin\phi_1}{N_0}\right]\geq\frac{1}{2\pi}\int_0^{2\pi}\exp\left[\frac{2\sqrt{E}r_{2c}\cos\phi_2+2\sqrt{E}r_{2s}\sin\phi_2}{N_0}\right] \tag{2-13}$$

By definition

$$\frac{1}{2\pi}\int_0^{2\pi}\exp\left[\frac{2\sqrt{E}r_{1c}\cos\phi_1+2\sqrt{E}r_{1s}\sin\phi_1}{N_0}\right]=I_0\left(\frac{2\sqrt{E(r_{1c}^{\ 2}+r_{1s}^{\ 2})}}{N_0}\right) \tag{2-14}$$

where $I_0$ (.) is a modified Bessel function of the zero-th order.

Therefore out ML decision rule becomes choosing signal 1 if

$$I_0\left(\frac{2\sqrt{E(r_{1c}^{\ 2}+r_{1s}^{\ 2})}}{N_0}\right)\geq I_0\left(\frac{2\sqrt{E(r_{2c}^{\ 2}+r_{2s}^{\ 2}}}{N_0}\right) \tag{2-15}$$

It is known that Bessel function is monotonically increasing. That is, if (2-5) is true the following inequality exists:

$$\sqrt{r_{1c}^{\ 2}+r_{1s}^{\ 2}}\geq\sqrt{r_{2c}^{\ 2}+r_{2s}^{\ 2}} \tag{2-16}$$

which tells us the maximum likelihood decision rule is to compare the energy in the two frequency bands and pick the larger one.

Based on the ML decision rule, the optimum noncoherent FSK receiver is illustrated in Fig. 2-2.

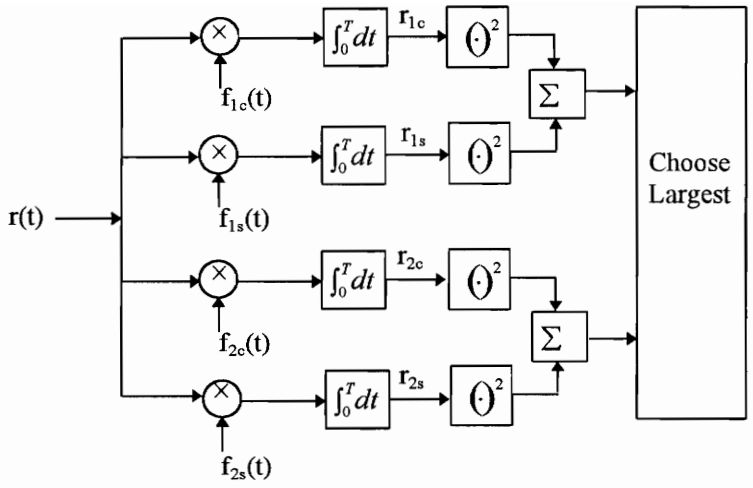



Figure 2-2. Structure of Optimum Noncoherent FSK Receiver

Assuming signal 2 was sent, under the assumption that white Gaussian noise is added in the channel, $r_{1c}$ and $r_{1s}$ have Gaussian distributions (no signal ), then $\sqrt{r_{1c}^2 + r_{1s}^2}$ has a Rayleigh distribution. Let $r_1 = \sqrt{r_{1c}^2 + r_{1s}^2}$

$$p(r_1 / s_2) = \begin{cases} \frac{r_1}{\sigma^2} e^{-r_1^2/(2\sigma^2)}, & r_1 \geq 0 \\ 0, & r_1 < 0 \end{cases} \tag{2-17}$$

where $\sigma^2 = N_0 B$. On the other hand, $\sqrt{r_{2c}^2 + r_{2s}^2}$ has a Rician distribution if signal 2 was sent [6]. Let $r_2 = \sqrt{r_{2c}^2 + r_{2s}^2}$

$$p(r_2 / s_2) = \begin{cases} \dfrac{r_2}{\sigma^2} e^{-(r_2^2+E)/(2\sigma^2)} I_0\left(\dfrac{r_2\sqrt{E}}{\sigma^2}\right), & r_2 \geq 0 \\ 0, & r_2 < 0 \end{cases} \tag{2-18}$$

An error is made if $\sqrt{{r_{1c}}^2 + {r_{1s}}^2} \geq \sqrt{{r_{2c}}^2 + {r_{2s}}^2}$ i.e. $r_1 > r_2$, given signal 2 was sent. Therefore, the error probability

$$\begin{aligned} P_b &= P_r(r_1 > r_2 / s_2) \\ &= \int_0^\infty p(r_2 / s_2)\left[\int p(r_1 / s_2) dr_1\right] dr_2 \end{aligned} \tag{2-19}$$

Substituting (2-17) and (2-18) into (2-19) and evaluating the integral by using the integral table in [6], the bit error rate (BER) of binary FSK for noncoherent detection is

$$P_b = \frac{1}{2} e^{-(1/2)(E_b/N_0)} \tag{2-20}$$

For comparison, BER for coherent detection of binary FSK is [6]

$$P_b = Q\left(\sqrt{\frac{E_b}{N_0}}\right) \tag{2-21}$$

A plot of (2-20) and (2-21) is given in Fig. 2-3. It is seen that noncoherent FSK requires, at most, only 1 dB more $E_b/N_0$ than that for coherent FSK if $P_b$ is $10^{-4}$ or less. The noncoherent FSK demodulator is considerably easier to build since the coherent reference signals do not have to be generated. Therefore, in practice, most FSK receivers use noncoherent detection. (Differentially coherent reception is popular for GMSK)

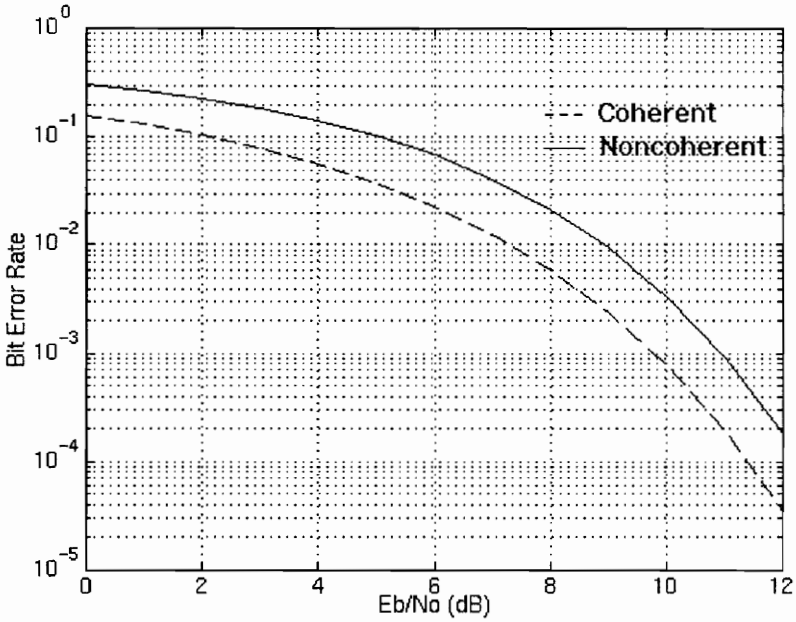



Figure 2-3. Probability of Bit Error Versus $E_b/N_0$ for Coherent and Noncoherent Detection of FSK

## 2.2.2 M-ary FSK

FSK signal waveforms are commonly employed for M-ary orthogonal signaling with noncoherent detection. An orthogonality of the FSK waveforms requires a frequency separation of $\Delta f = 1/T$, $T$ is duration of signaling interval. Consequently the channel bandwidth required for transmission is

$$B = M\Delta f = \frac{M}{T} \tag{2-13}$$

and the corresponding transmission rate $R$ is

$$R = \frac{1}{T}\log_2 M \tag{2-14}$$

Thus the bandwidth efficiency is

$$\frac{R}{B} = \frac{\log_2 M}{M} \tag{2-15}$$

The optimum noncoherent receiver for M-ary FSK can be obtained by the extension of that for binary FSK which is shown in Fig. 2-2. The error probability for noncoherent detection of M-ary FSK in AWGN is derived in [7]. The probability of symbol error is

$$P_s = \sum_{n=0}^{M-1} (-1)^{n+1} \binom{M-1}{n} \frac{1}{n+1} e^{-n k \gamma_b/(n+1)} \tag{2-16}$$

where $\gamma_b = E_b/N_0$. Converting the probability of a symbol error in (2-16) into an equivalent probability of a binary digit error, by means of the formula given below [7]

$$P_b = \frac{2^{k-1}}{2^k - 1} P_s \tag{2-17}$$

where $k = \log_2 M$, the number of bits per symbol. We obtain the BER for noncoherent M-ary FSK

$$P_b = \frac{2^{k-1}}{2^k - 1} \sum_{n=0}^{M-1} (-1)^{n+1} \binom{M-1}{n} \frac{1}{n+1} e^{-n k \gamma_b/(n+1)} \tag{2-18}$$

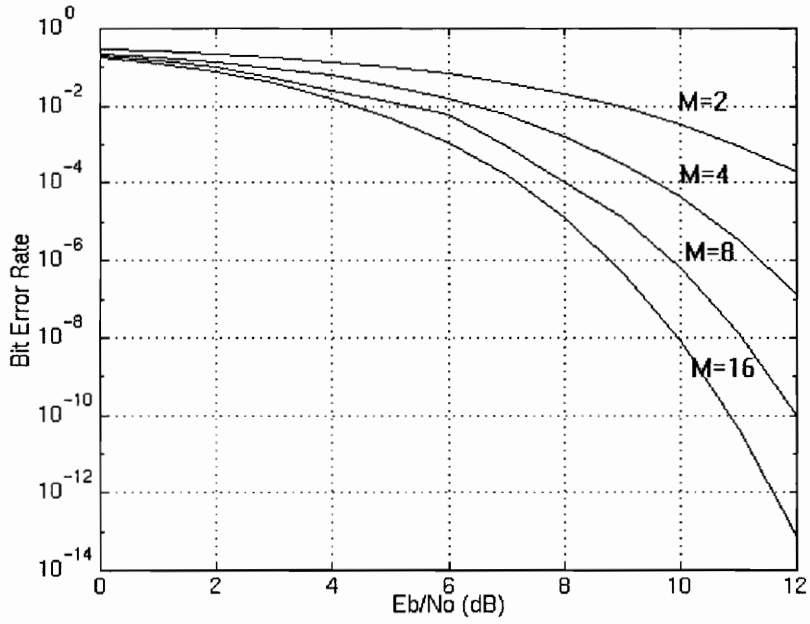



Figure 2-4. Probability of Bit Error Performance vs $E_b/N_0$ for Noncoherent M-ary FSK (M=2, 4, 8, 16)

The bit error probability is plotted in Fig. 2-4 for M = 2, 4, 8, 16. It is noted that as M increases, the SNR ($\gamma_b$) per bit required to achieve a given performance decreases. That is, the energy efficiency increases as M increases. It can be shown that $P_b \rightarrow 0$ as $M \rightarrow \infty$, provided $E_b > \ln 2$ [7].

## 2.3 Simulation of M-ary Noncoherent FSK

### 2.3.1 Complex Envelope Representation of Bandpass Signals

The fundamentals of simulation of digital communication system are the sampling theory and the complex envelop technique to represent signals and noise in a proper way [6]. Since we ultimately need to sample these functions, we wish to do so in the most efficient way possible. It is known that the bandpass modulated signals and bandpass systems can, with minor restrictions, be analyzed and simulated as if they were low-pass. The technique used in the implementation of this idea is the so-called complex envelop method [8].

Any carrier modulated signal $x(t)$ can be represented as

$$\begin{aligned} x(t) &= r(t)\cos\left[2\pi f_c t + \varphi(t)\right] \\ &= \mathrm{Re}\left[r(t)e^{j(2\pi f_c + \varphi(t)}\right] = \mathrm{Re}\left[r(t)e^{j\varphi(t)}e^{j2\pi f_c t}\right. \end{aligned} \tag{2-19}$$

where $r(t)$ is the amplitude modulation, $\varphi(t)$ is the phase modulation of the signal, and $f_c$ is the carrier frequency. The signal

$$v(t) = r(t)e^{j\varphi(t)} \tag{2-20}$$

evidently contains all of the information-related variations, and is of a low-pass nature. It is often called the complex low-pass equivalent or the complex envelope of the signal. If the bandwidth of $x(t)$ is B, the narrowband condition $B << f_c$ is frequently satisfied.

Under this condition, one can show that bandpass filtering of $x(t)$ can be evaluated through and equivalent low-pass filtering operation with $v(t)$ as an input.

In order to find the low-pass equivalent modulated signals, let's begin with the modulated signal in the inphase-quadrature form.

$$x(t) = p(t)\cos 2\pi f_c t - q(t)\sin 2\pi f_c t \tag{2-21}$$

where $p(t)$ and $q(t)$ are the inphase and quadrature parts of the modulated signal $x(t)$, respectively. It is shown in [8] that if the bandwidth B of both the inphase and quadrature terms are such that B << $f_c$, then the Hibert transform of $x(t)$ is

$$\hat{x}(t) = p(t)sn2\pi f_c t + q(t)x\cos 2\pi f_c t \tag{2-22}$$

The analytic signal of $x(t)$ is therefore

$$z_x(t) = x(t) + j\hat{x}(t) = [p(t) + jq(t)]e^{j2\pi f_c t} \tag{2-23}$$

The modulated signal is then

$$x(t) = \mathrm{Re}[z_x(t)] = \mathrm{Re}[\tilde{x}_L(t)e^{j2\pi f_c t}] \tag{2-24}$$

The low-pass equivalent signal or the complex envelope is then defined as

$$\tilde{x}_L(t) = z_x(t)e^{-j2\pi f_c t} = p(t) + jq(t) \tag{2-25}$$

## 2.3.2 Representation of the M-ary Noncoherent FSK Signals

A M-ary FSK signal can be represented as

$$s_i = A\cos\left[2\pi(f_c + I_i\Delta f)t + \theta_i\right] \quad i = 1,\ 2,\ldots,\ M \tag{2-26}$$

where $\Delta f = 1/T$, minimum frequency separation in order to meet the orthagonality the requirement for noncoherent detection. $I_i = (2i\text{-}1\text{-}M)/2$, making sure that the minimum frequency spacing is $\Delta f$ . $\theta_m$ is the phase of each signal in signal set. It is a random variable which is assumed to be uniformly distributed.

The complex envelope of $s_i$ is obtained by evaluating (2-5). $s_i$ in (2-26) can be also represented in the following form.

$$\begin{aligned} s_i &= A\cos(2\pi I_i\Delta f t + \theta_i)\cos 2\pi f_c t - A\sin(2\pi I_i\Delta f t + \theta_i)\sin 2\pi f_c t \\ &= fI_i\cos 2\pi f_c t - fQ_i\sin 2\pi f_c t \end{aligned} \tag{2-27}$$

where

$$\begin{aligned} fI_i &= A\cos(2\pi I_i\Delta f t + \theta_i) = A\cos\left[\pi(2i - 1 - M) + \theta_i\right. \\ fQ_i &= A\sin(2\pi I_i\Delta f t + \theta_i) = A\sin\left[\pi(2i - 1 - M) + \theta_i\right. \end{aligned} \tag{2-28}$$

are the inphase and quadrature terms of the modulated signal respectively. Therefore, the complex envelope representation of the noncoherent M-ary FSK is

$$s_{iL} = fI_i + jfQ_i \tag{2-29}$$

In our simulation to evaluate the performance of M-ary noncoherent FSK in AWGN, $fI_i$ and $fQ_i$ are generated to represent the modulated FSK signals.

### 2.3.3 Simulation of the M-ary FSK Signal

To verify the analytical results which are obtained in 2.2, simulation has been done to test the performance of M-ary noncoherent FSK in terms of error probability in AWGN. The simulation is carried out using MATLAB version 4.2. The procedure of the simulation is as follows.

1. Generate a symbol s in the range of [1, M].
2. Generate random phase θ which is uniformly distributed in [0, 2π].
3. Construct the complex envelope of the signal by using (2-28).

$$\begin{aligned} fI_i &= A\cos\left[\pi(2i-1-M)+\theta_i\right] \\ fQ_i &= A\sin\left[\pi(2i-1-M)+\theta_i\right] \end{aligned} \tag{2-30}$$

where $fI_i$ and $fQ_i$ is the inphase and quadrature terms of the ith signal. A is chosen to be 1. That is, the magnitude of a symbol is 1.

4. Generate noise according to $E_b/N_0$, i.e. signal-to-noise radio.

$$\frac{E_b}{N_0} = \frac{0.5A^2T_b}{\sigma^2(k/T_b)} = \frac{0.5}{\sigma^2 k} \tag{2-31}$$

where $k = log_2 M$, number of bits per symbol.

Therefore, the noise deviation

$$\sigma = \sqrt{\frac{0.5}{m(E_b/N_0)}} \tag{2-32}$$

The change of $E_b/N_0$ is achieved by scaling the noise variance and keeping bit energy $E_b$ constant.

5. Compute the received signal.

$$\begin{aligned} rI_i &= fI_i + nI_i \\ rQ_i &= fI_i + nQ_i \end{aligned} \tag{2-33}$$

where $nI_i$ and $nQ_i$ are the inphase and quadrature terms of the noise added to the ith signal and $rI_i$ and $rQ_i$ are the inphase and quadrature parts of the ith received signals.

6. Calculate the decision variables using noncoherent detection

$$z_i = rI_i^2 + rQ_i^2 \quad i = 1,\ 2,\ldots,\ M \tag{2-34}$$

7. Determine the received symbol in terms of maximum likelihood estimation.

$$s_est = k$$

where

$$z_k = \max(z_i) \quad i = 1,\ 2,\ \ldots,\ M$$

8. Update the bit error statistics.

In order to have confidence in simulation results, 100 independent error events are accumulated before calculate the BER. The results with M = 2, 4, 8, 16 are plotted in Fig. 2-5, Fig. 2-6, Fig. 2-7 and Fig. 2-8 respectively. The comparison with their analytical results are also given. It is seen that the simulation results agree with their theoretical counterparts.

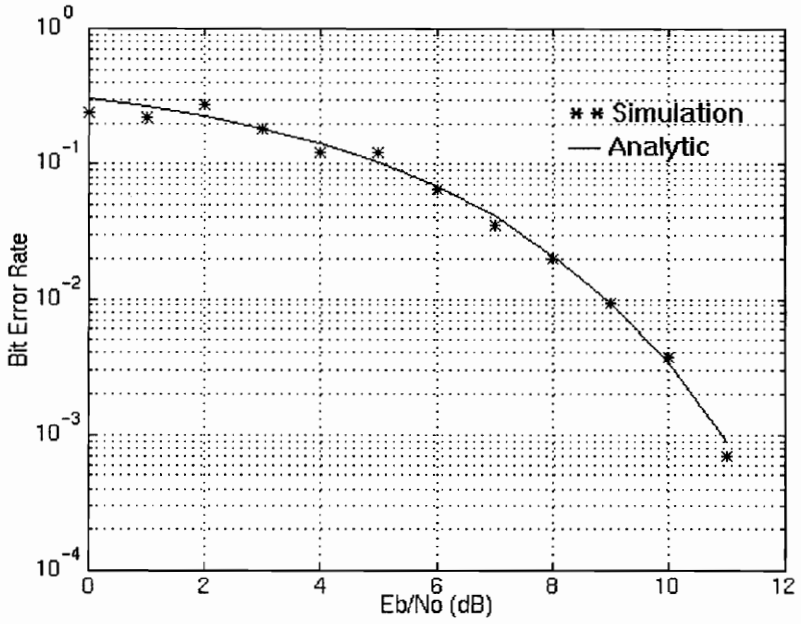



Figure 2-5. BER vs Eb/No for Binary Noncoherent FSK

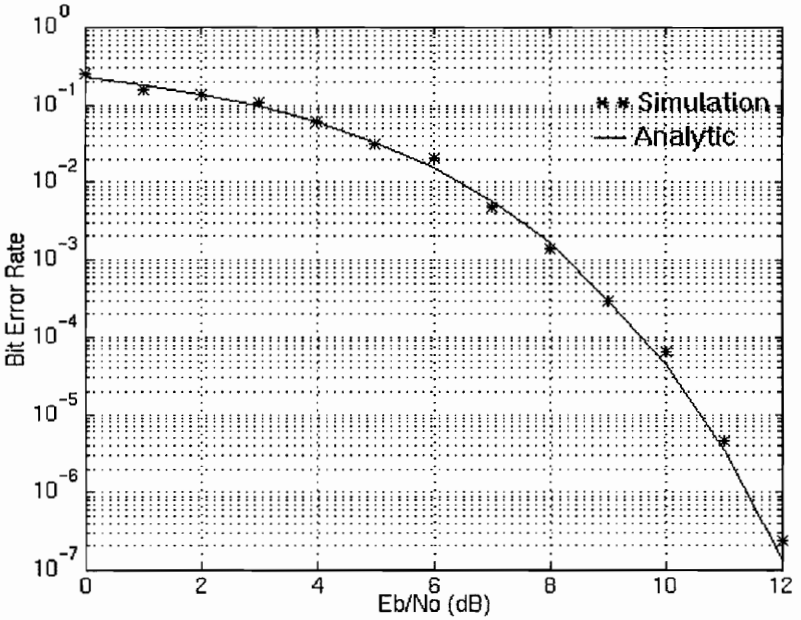



Figure 2-6. BER vs Eb/No for 4-ary Noncoherent FSK

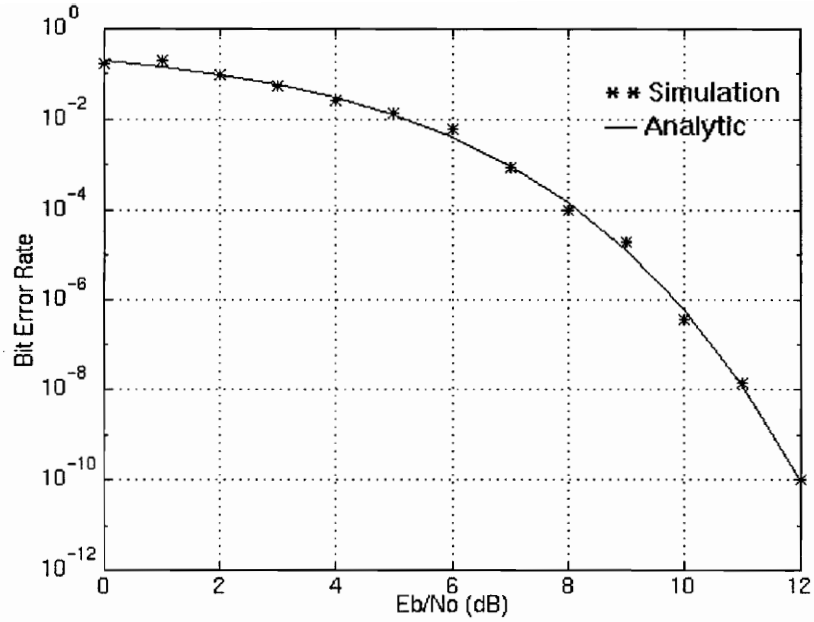



Figure 2-7. BER vs Eb/No for 8-ary Noncoherent FSK

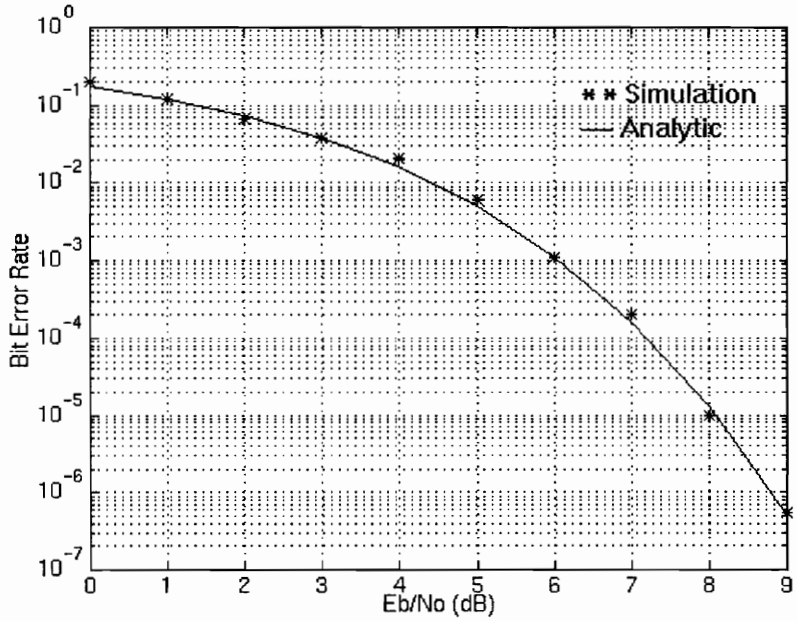



Figure 2-8. BER vs Eb/No for 16-ary Noncoherent FSK

## 2.4 Conclusion

In this chapter, the basic concepts of noncoherent modulation scheme are summarized. The emphasis is on the noncoherent M-ary FSK modulation. It is shown that the main difference between coherent FSK and noncoherent FSK modulation is the criterion on the demodulator. For coherent modulation, the demodulation is usually done by finding the signal which is the closest to the received signal, while for noncoherent FSK, the demodulation is based on energy detection.

Although the performance of noncoherent FSK is not as good as coherent FSK, in some applications, noncoherent FSK is practical and useful. In fact, noncoherent FSK modulation is one of the most widely used modulation schemes in communication systems, particularly when low cost and robust performance are critical such as in paging systems.

# Chapter 3 FSK Permutation Modulation

## 3.1 General Description

FSK permutation modulation (PFSK) is a subclass of the very general class of permutation modulations introduced by David Slepian [14]. PFSK applies the general permutation modulation to a multifrequency modulation scheme in which energy is transmitted simultaneously on n frequencies out of m, thus conveying $\log_2\binom{m}{n}$ bits of information per signal waveform. Binary FSK and M-ary FSK modulation are special cases of FSK permutation modulation.

Slepian pointed out the following advantages that the general class of permutation modulations have. First, the maximum-likelihood receiver for permutation modulation is algebraic in nature, easy to instrument, and does not require local generation of the possible sent messages; second, the probability of incorrect decoding is the same for each sent message; and third, each code word requires the same energy for transmission [4]. In addition, PFSK modulations also have the advantage (as do the FSK modulations ) of not requiring a threshold receiver detector.

Considering a $\binom{m}{n}$ FSK permutation scheme, the total number of carrier frequencies is $m$, the number of distinct combinations of $n$ out of $m$ is

$$\binom{m}{n} = \frac{m!}{n!(m-n)!} \tag{3-1}$$

and the number of bits which can be transmitted per signal waveform is $k = \log_2\binom{m}{n}$.

The spectral efficiency is thus

$$\frac{k}{m} = \frac{\log_2 \binom{m}{n}}{m} \tag{3-2}$$

Table 3-1 illustrates the spectral efficiency for several PFSK schemes. It is apparent that using PFSK modulation increases the spectral efficiency substantially.

Table 3-1. Spectral Efficiency of PFSK

| $\binom{m}{n}$ | $\frac{k}{m}$ | |
|---|---|---|
| $\binom{2}{1}$ | 0.5 | (BFSK) |
| $\binom{4}{1}$ | 0.5 | (4-ary FSK) |
| $\binom{8}{1}$ | 0.375 | (8-ary FSK) |
| $\binom{4}{2}$ | 0.646 | |
| $\binom{6}{3}$ | 0.72 | |
| $\binom{8}{4}$ | 0.77 | |

PFSK transmission operates in a manner shown for the $\binom{m}{2}$ alphabet in Fig. 3-1. One of the $\binom{m}{n}$ characters is the input to the transmitter: The signal output is n simultaneous pulses of energy, one pulse on each of m distinct frequencies, lasting for T seconds. White Gaussian noise is added in the channel. Filters matched to pulse shape, are tuned to each of the m possible frequencies. The filter outputs are envelope-detected and all $m$ envelope samples are intercompared at the end of the pulse period. The largest $n$ of these outputs determine the transmitted symbol.

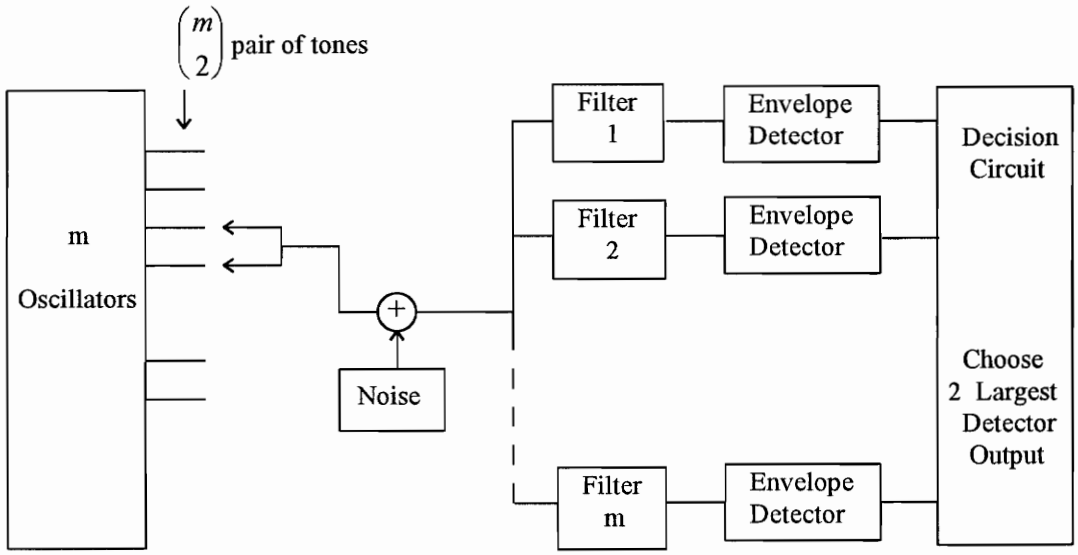



Figure 3-1. Transmission System for $\binom{m}{2}$ PFSK [10]

The error performance of PFSK is studied by Schneider [10] where the error probability is derived. For $\binom{m}{n}$ PFSK, the symbol error probability

$$P_s = n\sum_{r=1}^{m-n}(-1)^{r+1}\binom{m-n}{r}\int_0^{\infty}[Q(\sqrt{2R},\ s)]^{n-1} \times sI_0(s\sqrt{2R})\exp\left[-\frac{(r+1)s^2+2R}{2}\right]ds \qquad (3\text{-}3)$$

where

$I_0$ is the modified Bessel function

$R$ is $E_s/N_0$

$E_s$ is the received signal energy at each of the $n$ transmitted frequencies

$N_0$ is the noise density, in watts per Hz

Q (.) is the Q function

A closed form expressions for the cases of $n = 1$ was found by Reiger [11].

$$P_s(n=1) = \frac{1}{m}\sum_{r=2}^{m}(-1)^r\binom{m}{r}\exp\left[-R(1-\frac{1}{r})\right] \qquad (3\text{-}4)$$

which can be shown equivalent to $P_s$ in (2-16 ) for M-ary FSK .

Closed form expression for cased of $n > 2$ are not known; however, an asymptotic form for large $R$ is obtained following arguments by Helstrom [12] for approximating the Q function.

$$P_s \approx \frac{n}{m-n+1}\sum_{r=2}^{m-n+1}(-1)^r\binom{m-n+1}{r}\exp\left[-R\left(1-\frac{1}{r}\right)\right] \qquad (3\text{-}5)$$

or taking the predominant first term of equation (3-5) we have

$$P_s \approx \frac{n(m-n)}{2}\exp\left(-\frac{R}{2}\right) \tag{3-6}$$

The probability of bit error is also derived by Schneider

$$P_b = 1-(1-P_s)^{1/k} \approx \frac{1}{k}P_s \tag{3-7}$$

where $k = \log_2\binom{m}{n}$

## 3.2 Simulation of Several PFSK Cases

To better understand FSK permutation modulation scheme, several PFSK cases are simulated, such as $\binom{6}{3}$ and $\binom{8}{4}$ permutation modulations. Description is only given for the $\binom{6}{3}$ case, because the $\binom{8}{4}$ scheme is very similar to $\binom{6}{3}$ one.

For $\binom{6}{3}$ permutation modulation, we propose that during each symbol interval a combination of three carriers are ON while the other three are OFF. The total number of distinct symbols is

$$\binom{6}{3} = \frac{6!}{3!(6-3)!} = 20$$

which gives $\log_2 20 = 4.32$ bits per symbol.

We use 16 out of 20 symbols for data transmission., which corresponds to 4 bits per symbol. The actual mapping of each symbol to the transmitting character is shown in Table 3-2. The mapping is randomly done.

The simulation procedure is described in the following.

1. Generate a symbol in the range of [0, 15].
2. Generate random phase θ which is uniformly distributed in [0, 2π].
3. Determine the corresponding signal pattern by looking it up in Table 3-2 and construct the complex envelope of the selected signal, similar to the procedure in Chap 2.

$$fI_i = \begin{cases} \cos[\pi(2i-1-6)+\theta], & i \in A_i \\ 0, & i \notin A_i \end{cases}$$
$$fQ_i = \begin{cases} \sin[\pi(2i-1-6)+\theta], & i \in A_i \\ 0, & i \notin A_i \end{cases} \qquad (3\text{-}8)$$

4. Generate white Gaussian noise according to $E_b/N_0$

Let $E_b$ = energy per bit

$E_s$ = energy per symbol

$E_t$ = energy per tone

They relate each other by (3-9)

$$4\,E_b = E_s = 3\,E_t \qquad (3\text{-}9)$$

which yields

$$E_b = 3/4\,E_t \qquad (3\text{-}10)$$

Table 3-2. Mapping of Each of the 16 Symbols to the Transmitted Characters

| Symbol | Patterns | | | | | |
|---|---|---|---|---|---|---|
| | $f_1$ | $f_2$ | $f_3$ | $f_4$ | $f_5$ | $f_6$ |
| 0000 | 0 | 0 | 0 | 1 | 1 | 1 |
| 0001 | 0 | 0 | 1 | 0 | 1 | 1 |
| 0010 | 0 | 0 | 1 | 1 | 0 | 1 |
| 0011 | 0 | 0 | 1 | 1 | 1 | 0 |
| 0100 | 0 | 1 | 0 | 0 | 1 | 1 |
| 0101 | 0 | 1 | 0 | 1 | 0 | 1 |
| 0110 | 0 | 1 | 0 | 1 | 1 | 0 |
| 0111 | 0 | 1 | 1 | 0 | 0 | 1 |
| 1000 | 0 | 1 | 1 | 0 | 1 | 0 |
| 1001 | 0 | 1 | 1 | 1 | 0 | 0 |
| 1010 | 1 | 0 | 0 | 0 | 1 | 1 |
| 1011 | 1 | 0 | 0 | 1 | 0 | 1 |
| 1100 | 1 | 0 | 0 | 1 | 1 | 0 |
| 1101 | 1 | 0 | 1 | 0 | 0 | 1 |
| 1110 | 1 | 0 | 1 | 0 | 1 | 0 |
| 1111 | 1 | 0 | 1 | 1 | 0 | 0 |

Therefore

$$\frac{E_b}{N_0} = \frac{3 \times 0.5}{4 \times \sigma^2} \tag{3-11}$$

The noise deviation is thus

$$\sigma = \sqrt{\frac{3 \times 0.5}{4 \times (E_b / N_0)}} \tag{3-12}$$

The change of $E_b/N_0$ is obtained by scaling the noise variance and keeping bit energy $E_b$ constant.

5. Compute the received signal.

$$\begin{aligned} rI_i &= fI_i + nI_i \\ rQ_i &= fQ_i + nQ_i \end{aligned} \qquad i = 0, 1, \ldots, 15 \tag{3-13}$$

6. Calculate a decision statistic for each of 16 symbols

$$z_i = \sum_{i \in A_i} (rI_i^2 + rQ_i^2) \qquad i = 1, 2, \ldots, 16 \tag{3-14}$$

where $A_i$ is the set of tones which are on for symbol $i$.

7. Determine the received symbol in terms of maximum likelihood estimation.

$$s_est = k$$

where

$$z_k = \max(z_i) \qquad i = 0,\ 1,\ \ldots,\ 15 \tag{3-15}$$

8. Update the error statistics.

The simulation results for $\binom{6}{3}$ and $\binom{8}{4}$ are shown in Fig. 3-2 and Fig. 3-3 respectively. For comparison, plots of (3-7) for $\binom{6}{3}$ and $\binom{8}{4}$ are also displayed.

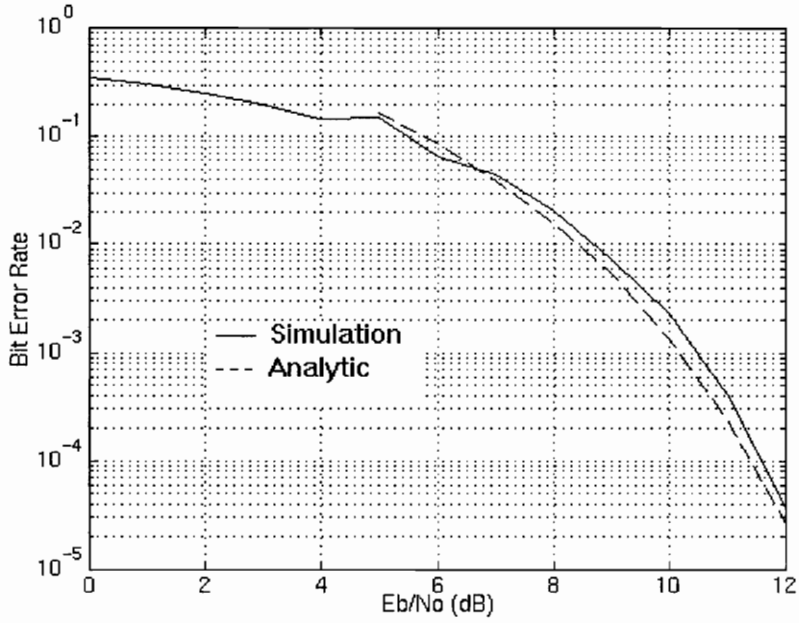



Figure 3-2. Error Probability for $\binom{6}{3}$ PFSK

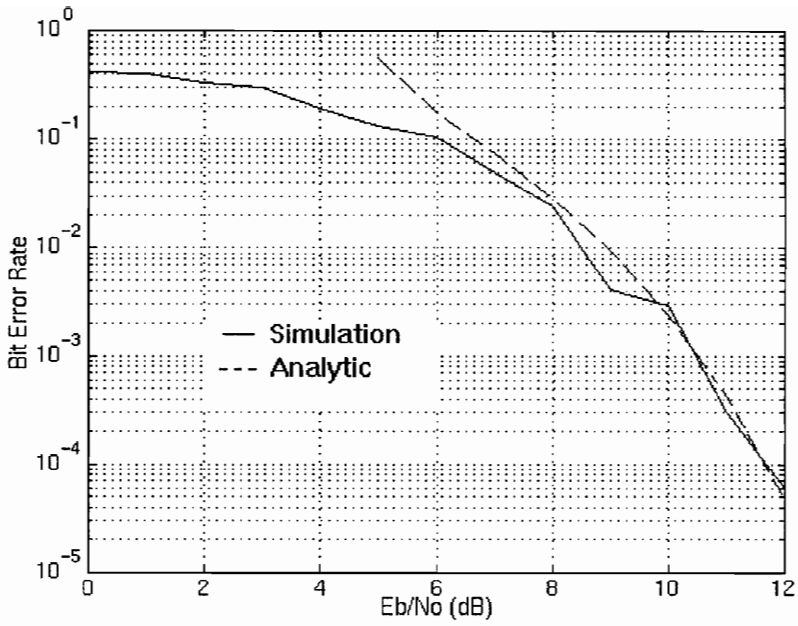



Figure 3-3. Error Probability for $\binom{8}{4}$ PFSK

## 3.3 Comparison of Different PFSK

The performance of PFSK schemes and M-ary FSK are compared in this section. Since M-ary FSK is a special case of $\binom{m}{n}$ PFSK with n = 1, the comparison is done within PFSK for n = 1 and n > 1. Two criteria are used: energy efficiency and spectral efficiency. The spectral efficiency for $\binom{m}{n}$ PFSK was already shown before, which is

$$\frac{k}{m} = \frac{\log_2 \binom{m}{n}}{m}$$

The energy efficiency is based on energy per bit required for an equivalent error rate, i.e., bit error rate.

Fig. 3-4 shows paired values of spectral efficiency and $E_b/N_0$ at a bit error rate of $10^{-4}$, for illustrative alphabets. It is also called energy efficiency and spectral efficiency plot. It is desirable for a modulation scheme to be further towards the right and bottom of the plot. It shows that with permutation modulation the spectral efficiency increases, but, on the other hand, the energy efficiency decreases.

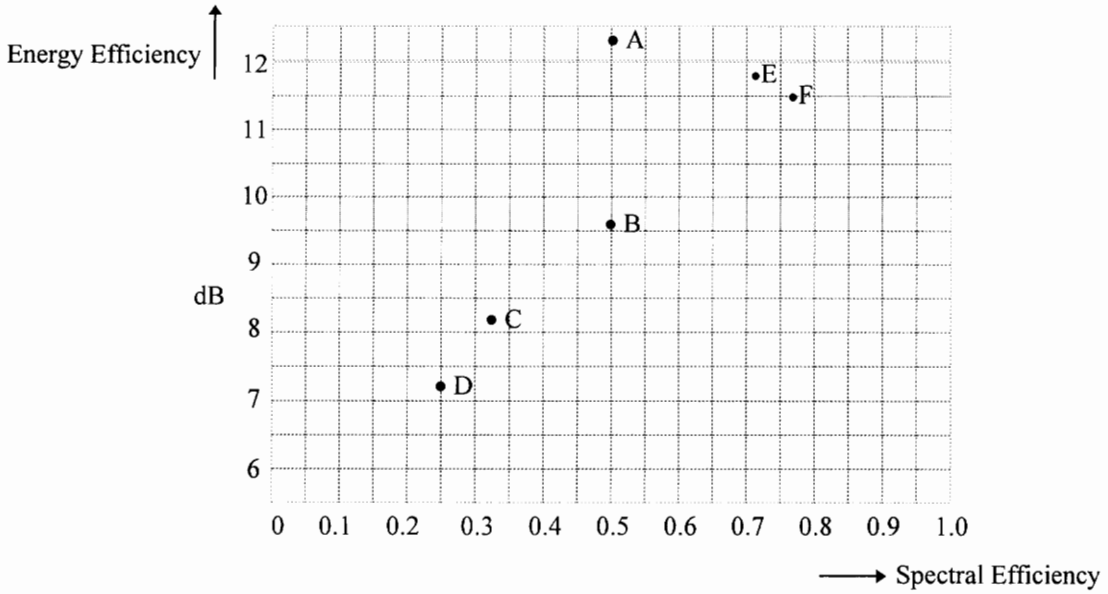



Figure 3-4. Performance Comparison of PFSK Alphabets

| Symbol | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| Alphabet | $\binom{2}{1}$ | $\binom{4}{1}$ | $\binom{8}{1}$ | $\binom{16}{1}$ | $\binom{6}{3}$ | $\binom{8}{4}$ |

## 3.4 Conclusion

In this chapter, we introduced FSK permutation modulation, a subclass of general permutation modulation. The major advantage of using FSK permutation modulation over conventional M-ary FSK is the spectral efficiency can be increased. This feature may be interested by bandwidth-limited applications (e.g. paging industry). It is shown, however, that the increased spectral efficiency is obtained by trading off its energy efficiency. This drawback can be improved by the application of the so called trellis coding, which will be discussed in the following chapters.

# Chapter 4 Introduction to Trellis-Coded Modulation

## 4.1 Introduction

Trellis-coded modulation (TCM) is introduced in this chapter. TCM is a combined coding and modulation scheme for improving reliability of a digital transmission system without increasing the transmitted power or the required bandwidth [13]. In a power-limited environment, the system performance can be improved by the use of error-correction codes, which increases the power efficiency by adding redundant bits to the transmitted symbol sequence. This procedure requires the modulator to operate at a higher data rate, if the information rate is to retain the same, and hence needs a larger bandwidth. In a bandwidth-limited environment, the efficiency in bandwidth utilization can be increased by using higher order modulation schemes, i.e., 8-PSK instead of 4-PSK. However, a larger signal power would be needed to maintain the same signal separation. Therefore an unchanged error probability is expected if in a both power-limited and bandwidth-limited environment.

The disappointing situation mentioned above results from the fact that the error correction coding and modulation are done separately. One solution to this problem is to combine the coding and the modulation into the so called trellis-coded modulation (TCM). TCM is introduced by Ungerboeck on his classic 1982 paper "Channel coding with multilevel/phase signals " [1]. It integrates the choice of a higher order modulation scheme with that of a convolutional code, while the receiver, instead of performing demodulation and decoding in two separate steps, combines the two operations into one. TCM provides coding gain without bandwidth expansion or reduction of the effective information rate as required by traditional error correction schemes. It is well-suited for both power-limited and bandwidth-limited applications.

Since TCM was introduced in the early 80s, it has found various applications in different areas. The most immediate application for TCM was in the area of digital data transmission over a standard telephone line [5]. With TCM technology it is possible to increase the data rate substantially. CCITT Study Group XVII has recently approved a 9.6/14.4-Kbps trellis coded modem for use with both intercomputer and FAX communication [19]. It is also noted that TCM has had a substantial impact on other bandwidth-limited applications, particularly in cellular mobile radio and satellite communications [16], [20].

In this chapter, the basic concepts of TCM are explained. The emphasis is placed on three key concepts: expanded signal constellations, signal constellation partitioning and the selection of partitions by convolutional encoders, followed by an example of Ungerboeck's TCM scheme. Then, the decoding of TCM is introduced with emphasis on the definition of a metric and the application of the Viterbi algorithm.

## 4.2 Fundamentals of TCM

### 4.2.1 Signal Constellation

It is assumed that a discrete-time, continuous-amplitude model is used for the transmission of data on the additive white Gaussian noise channel. In this communication model, the messages to be delivered to the receiver are represented by points, or vector, in an N-dimensional Euclidean space $\mathbf{R}^N$, called the signal space. The geometric representation of the signal space is a signal constellation. For example, the constellations of 4-PSK and 8-PSK signals are illustrated in Fig. 4-1.

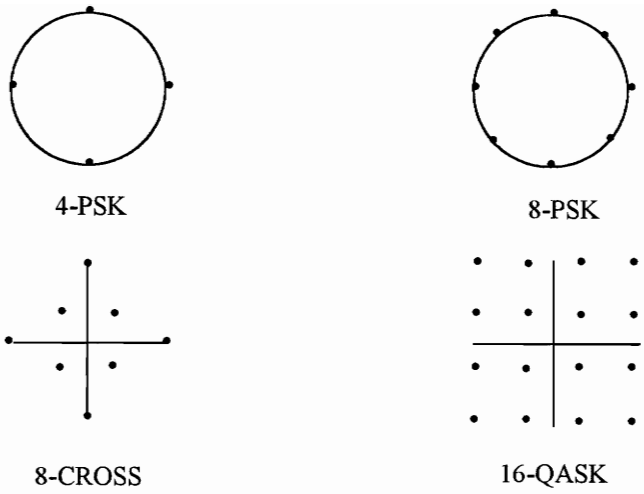



Figure 4-1. Several Signal Constellations

The position of a signal on the diagram is proportional to its magnitude and its relative phase or, equivalently, to the square root of the transmitted or received signal energy. The distance between signal points is called the Euclidean distance. The shortest Euclidean distance between any pair of signals in the diagram is called the minimum Euclidean distance.

$$d_{\min} = \min_{i \neq j} \left\| x_i - x_j \right\| \quad i, \mathrm{j} = 1,\ 2,\ \ldots,\ M$$

In the cases of 4-PSK and 8-PSK, assuming that the signals have unity energy (i.e., the circle on which the signals rest has unit radius), the minimum Euclidean distance of 4-PSK and 8-PSK are $\sqrt{2}$ and $2\sin(\pi/8) = 0.765$, respectively.

When a vector **x** is transmitted, the received signal is represented by the vector

$$\boldsymbol{z} = \boldsymbol{x} + \boldsymbol{n}$$

where **n** is a noise vector whose components are independent Gaussian random variables with mean zero and the same variance $N_0/2$. The vector **x** is chosen from the constellation $\Omega$ consisting of M signal vectors. The average square length [13]

$$E = \frac{1}{M}\sum_{x\in\Omega}\|x\|^2$$

will be referred to as the average signal energy.

If all the signals are equally likely to be transmitted, the maximum likelihood receiver selects the signal that is closest in Euclidean distance to the received signal. It is shown that the error probability of a modulation scheme is upper bounded by a decreasing function of its minimum Euclidean distance [13].

$$P_e \le \frac{M-1}{2} erfc\left(\frac{d_{\min}}{2\sqrt{N_0}}\right) \tag{4-1}$$

With this model, the problem of designing a good communication system is that of choosing a set of vector signals such that the minimum Euclidean distance is maximized, once the quantities M, N, and E are given [13].

## 4.2.2 Set Partitioning

The key to the integrated modulation and coding approach is to invent an effective method for mapping the coded bits into signal points such that the minimum Euclidean distance is maximized. Such a method was developed by Ungerboeck, based on the principle of mapping by set partitioning. Set partitioning is accomplished by successively partitioning of the M-ary constellation into subconstellations, having successively larger minimum distances.

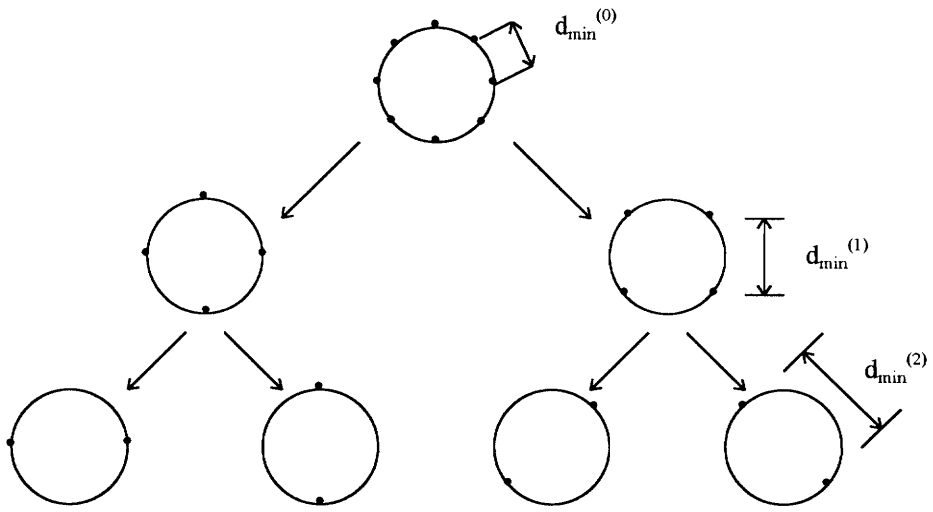



Figure 4-2. Set Partitioning of an 8-PSK Signal Set [5]

Let's look at an example of the partitioning of an 8-PSK in Fig. 4-2. In the eight-phase signal set the signal points are located on a circle of radius 1 (normalized) and have a minimum distance separation of

$$d_{\min}^{(0)} = 2\sin\frac{\pi}{8} = 0.765$$

In the first partitioning the eight points are subdivided into two subsets of four points each, such that the minimum distance between points increases to

$$d_{\min}^{(1)} = \sqrt{2}$$

In the second level of partitioning, each of the two subsets is subdivided into two subsets of two points, such that the minimum distance increases to

$$d_{\min}^{(2)} = 2$$

In this example, the partitioning was carried out to the limit where each subset contains two single points. In general, the degree to which the signal is partitioned depends on the characteristics of the code.

## 4.3 Trellis-Coded Modulation

In the last section, the fundamentals of TCM were discussed. The actual implementation of TCM is introduced in this section. It is mentioned above, that TCM integrates convolutional coding with modulation. Ungerboeck introduced the redundancy required for error control without increasing the signal bandwidth through a three step procedure [2].

Ungerboeck encoding is performed as follows. Suppose that the uncoded system uses a $2^m$-ary signal constellation.

1. Every m source bits are encoded by an *m/(m+1)* encoder to add one redundant bit.
2. Expand the signal constellation from $2^m$ to $2^{m+1}$ signals.
3. Use the *(m+1)*-bits encoded source blocks to select signals from the expanded constellation.

In this way, a $2^{m+1}$-ary constellation is being used to transmit information at a rate of m bits/sec/Hz. The symbol transmission rate for the coded system is the same as that of the uncoded system, therefore the Nyquist bandwidth is not increased.

The genius of Ungerboeck's system lies in the manner by which the m information bits are mapped onto the $2^{m+1}$ signals in the expanded constellation. This

mapping is achieved through set partitioning [5]. An Ungerboeck encoder is shown in Fig. 4-3 [5]. A total of m information bits $x_1, x_2, \ldots, x_m$ are taken from the data stream for mapping onto a signal from the $2^{m+1}$-ary constellation. Among m information bits, k of them are encoded using a rate $k/(k+1)$ convolutional encoder. The resulting $(k+1)$ bits are used to select one of the $2^{k+1}$ partitions of the $2^{m+1}$-ary signal constellation at the $(k+1)$st level of the constellation's partition tree. The remaining $(m-k)$ information bits are then used to select a signal z within the specified partition.

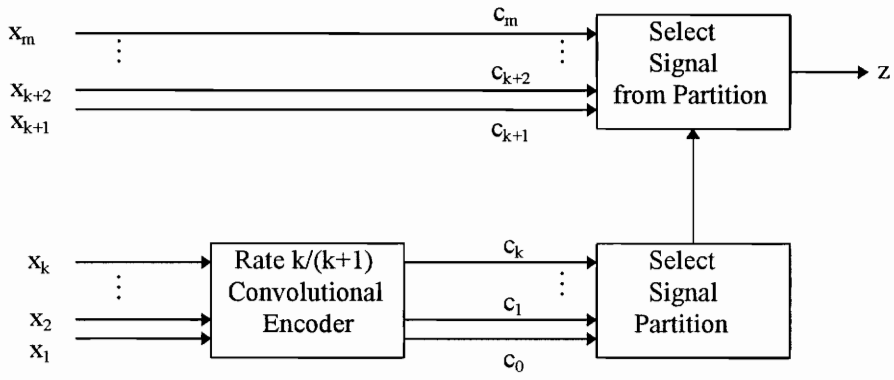



Figure 4-3. An Ungerboeck Encoder [5]

To better understand the trellis-coded modulation scheme, let's consider one of Ungerboeck's simple examples, which uses the 8-PSK signal constellation in conjunction with trellis codes. Uncoded 4-PSK is used as a reference in measuring coding gain. The two constellations are shown in Fig. 4-4, assuming normalized average signal energy. Since a rate-1/2 convolutional encoder is to be used, the 8-PSK constellation is partitioned into four subconstellations {D0, D1, D2, D3}. Each subconstellations is assigned a label that corresponds to a unique two-bit output from the convolutional encoder.

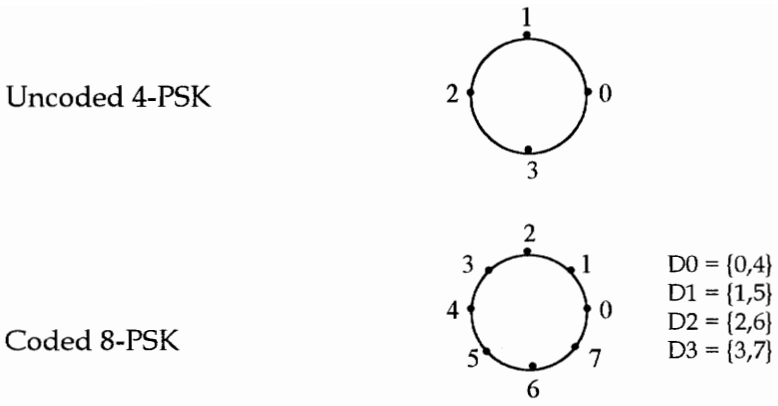



Figure 4-4. Partitions for the Ungerboeck Encoder in Figure 4-5 [5]

The encoder with four states is shown in Fig. 4-5. The output of the encoder selects one of the four partitions of the 8-PSK constellation. The uncoded bit $x_1$ is then used to choose one of the two signals in the selected partition.

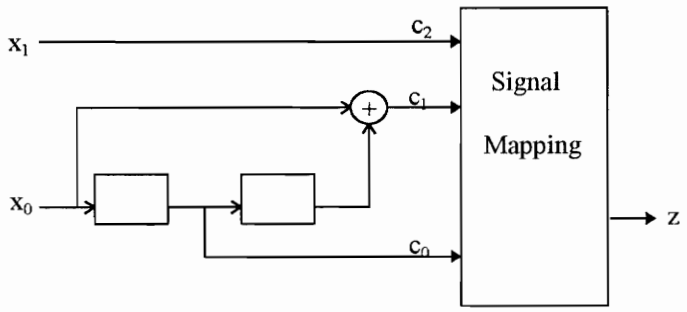
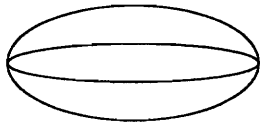


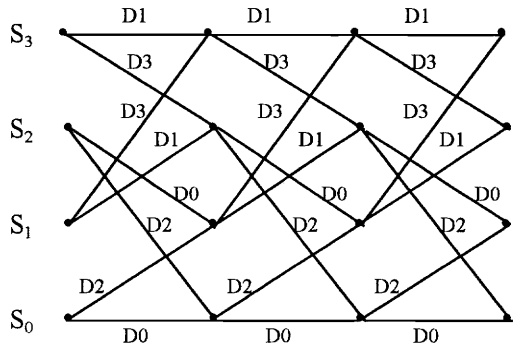

Figure 4-5. Four State Ungerboeck Encoder [5]

Consider the uncoded and coded cases through the use of trellis diagrams. In the uncoded case, the signal corresponds to a one-state trellis with four parallel state transitions as shown in Fig. 4-6a. The two source bits are taken each time and mapped

onto one of the four signals in the 4-PSK constellation. Since there is no memory in this system, any of the four signals can be selected at any symbol transmission time slot.

For the coded 8-PSK case, the four-state trellis is shown in Fig. 4-6b. Note that the states of the trellis diagram correspond to the contents of the memory elements in the convolutional encoder. Each branch in the trellis associates with one of the four subconstellations D0, D1, D2, D3 from which signals may be selected for transmission, given a specific state transition. For instance, when the convolutional encoder moves from state $S_0$ (00) to state $S_1$ (01), only a signal from partition D2 may be selected for transmission.






(a) One-State Trellis (b) Four-State Trellis

Figure 4-6. Trellis Diagrams for Uncoded and Coded Systems

The probability of symbol error for transmission over a noisy channel is a function of the minimum Euclidean distance between pairs of distinct signal sequences. The minimum Euclidean distance in the trellis code is called the free Euclidean distance and denoted as $d_{free}$. In the uncoded 4-PSK case, two sequences of symbols may differ in only one coordinate, and any values can be taken by that coordinate. The minimum free

distance is thus the minimum distance for the 4-PSK constellation: $d_{free/uncoded} = \sqrt{2}$. In the case of coded 8-PSK, we have a more interesting situation. Fig. 4-6b shows one way in which two sequences diverge from one state and remerge at the same state after more than one transition. This pair of sequences is one of the closest pairs of sequences for this encoder. The Euclidean distance between this pair is the minimum free distance for nonparallel transition.

$$\begin{aligned} d_{free/nonparallel} &= \sqrt{d^2_{\min}(D0, D2) + d^2_{\min}(D0, D1) + d^2_{\min}(D0, D2)} \\ &= \sqrt{2 + \left(2\sin\frac{\pi}{8}\right)^2 + 2} \\ &= 2.14 \end{aligned}$$

On the other hand, there is more than one signal associated with each branch in the trellis. These signal sequences differs by a single parallel transition. The minimum free distance for parallel transitions is the minimum distance for the subconstellation on each branch. For the example under consideration, the minimum free distance between parallel transitions is the shortest distance among D0, D1, D2, and D3. Thus, $d_{free/parallel} = 2$. Hence, the minimum free Euclidean distance in the four-state trellis is $d_{free/coded} = 2$.

The performance improvement provided by the coded system relative to the uncoded is usually measured in terms of the asymptotic coding gain $\gamma$. It is defined as follows [5].

$$\gamma = \frac{\left(\frac{S_{uncoded}}{d^2_{free/uncoded}}\right)}{\left(\frac{S_{coded}}{d^2_{free/coded}}\right)} \qquad (4\text{-}2)$$

where $S_{uncoded}$ denotes the normalized average received energy for the uncoded system and $S_{coded}$ is the normalized average received energy for the coded system. In the coded 8-PSK example the uncoded and coded constellations have the same average energy

(assumed to be unity). The minimum squared free distance for the uncoded case is 2, while that for the coded case is 4. The asymptotic coding gain is thus $\gamma = 2$ (3.01dB). Note that this coding gain is achieved by a very simple four-state encoder. Ungerboeck has shown that with more complex encoder structure, the coding gain can reach 6 dB or more.

Ungerboeck found that several principals were useful for the construction of good TCM codes. Although other researchers have found optimum codes recently through computer search [15]. Ungerboeck's design rules are summarized as follows [5].

a. Parallel transitions (when they occur) are assigned to signal points separated by the maximum Euclidean distance.
b. The transition originating from and merging into any state is assigned the signals from the preceding partitions.
c. All signals should occur with equal frequency.

The first of the above rules help maximize the Euclidean distance between signals assigned to parallel transitions. The second rule maximized the distance between nonparallel paths at their first and last branches. The last rule guarantees that the trellis code will have a regular structure.

## 4.4 Decoding of TCM

The decoding of TCM is based on the Viterbi algorithm, which is an optimum algorithm in the sense of maximum likelihood detection [23]. If the TCM signal is described using a trellis, whose branches are associated with transitions between encoder states and with signals transmitted over the channel, the task of the TCM decoder is to estimate the path that the encoded signal sequence traverses through the

trellis. This is done by associating with each branch of the trellis a number, called the branch metric, and looking for the path whose total metric is minimum. This path corresponds to the estimate of transmitted signal sequence. Thus the decoding is done by two procedures. First of all, the definition of a branch metric, and its computation based on the observed values of the received signal. Second, the evaluation of the minimum-metric path.

### 4.4.1 The Branch Metric

Consider a sequence of M-ary symbols to be transmitted

$$\boldsymbol{x} = (x_0, x_1, \ldots, x_{K-1}) = \sum_{k=0}^{K-1} x_k \tag{4-3}$$

where each $x_i$ can take on M values. These symbols are used to modulate a signal g(t), which is sent through the channel. Therefore, the transmitted signal can be given as

$$s(t) = \sum_{k=0}^{K-1} x_k g(t - kT) \tag{4-4}$$

where T denotes the symbol duration (equivalently, 1/T denotes the baud rate). If the channel effect is to add the random noise n(t) to the transmitted signal, the received signal can be written in the form

$$\begin{aligned} y(t) &= s(t) + n(t) \\ &= \sum_{k=0}^{K-1} x_k g(t - kT) + n(t) \end{aligned} \tag{4-5}$$

If we consider only signal samples instead of continuous signals, we may rewrite y(t) as

$$y(t) = \sum_{k=0}^{K-1} y_k \tag{4-6}$$

The task of the demodulator is to process the observed signal y(t) in order to produce and estimate $\hat{x}$ of the transmitted symbol sequence **x**.

$$\hat{x} = (\hat{x}_0, \hat{x}_1, \ldots, \hat{x}_{K-1}) \tag{4-7}$$

We are interested in optimum decoding, where by "optimum" we mean the error probability is minimized. Note that the probability of receiving the output y given that the input was **x** is ([13])

$$P[y / x] = \prod_{k=1}^{K-1} P[\, y_k / x_k] \tag{4-8}$$

Given **y**, the path most likely to have been followed through the trellis by the encoder is the path whose sequence maximizes $P[y / x]$. That is the decoder chooses $\hat{x}$ if

$$P[y(t) / \hat{x}] = \max_{all\ x} P[y(t) / x] \tag{4-9}$$

The function P[ y(t) / $x$] is the metric used to compare the sequence $\hat{x}$ and the received sequence **y**.

Since the logarithm is a monotonically increasing function of an increasing argument, the decoder could also use the metric $\log[P[y(t) / \hat{x}]$ rather than $P[y(t) / \hat{x}]$. In this way, the decoder would calculate $\log[P[y(t) / \hat{x}]$ for all paths and would choose the path with the largest value as the decoder output path. Taking the

logarithm converts the product of (4-8) to a summation. Therefore, the decoding metric is

$$\log P[(\boldsymbol{y}/\boldsymbol{x})] = \sum_{k=1}^{K-1} \log[P(y_k / x_k)] \tag{4-10}$$

Recall that the demodulation is done by using the minimum-distance rule, that is, by looking for the value of $x_k$ such that the Euclidean distance

$$\|y_k - x_k g_k\|^2$$

between the observed signal $y_k$ and the candidate signal $x_k g_k$ is a minimum. Thus, the probability of received signal $y_k$ given by transmitted signal $x_k$ is a function of their Euclidean distance. Therefore, the following equation exists.

$$\sum_{k=1}^{K-1} \log[P(y_k / x_k)] = C - \sum_{k=0}^{K-1} \|y_k - x_k g_k\|^2 \tag{4-11}$$

where C is a constant that can be disregarded in the maximization. In conclusion, maximum-likelihood demodulation of a sequence of signal $x_k$, k = 0, 1, ..., K-1, is equivalent to finding the sequences with the smallest path metric associated with $\mathbf{y}$. The metric is defined as

$$m[\boldsymbol{y}/\boldsymbol{x}] = \sum_{k=0}^{K-1} \|y_k - x_k g_k\|^2 \tag{4-12}$$

As a summary, for coherent modulation in Gaussian noise, the best choice of metric is the squared Euclidean distance. For noncoherent modulation, another metric may lead to superior performance, which will be discussed in the next chapter.

### 4.4.2 Viterbi Algorithm

The Viterbi algorithm was proposed for decoding convolutional codes. Its application to decode TCM is based on the trellis diagram representing the TCM encoder. For simplicity, assume that there is only one signal associated with each branch of the trellis, i.e., no parallel transitions. In this case, we have a one-to-one correspondence between the sequence of source symbols and the path through the trellis. The symbol $s_k$, or, equivalently, a branch of the trellis, is transmitted every T seconds.

The Viterbi algorithm finds the path with the minimum path metric by sequentially moving through the trellis stage by stage. In every stage, i.e., from stage k to stage (k+1), the receiver observes the sample $y_k$ received in that interval and computes the metric $\|y_k - x_k g_k\|^2$ associated with all the branches [13]. The path metric for a particular path through a given node is the sum of the partial path metrics for the portion of the path to the left of the node and the portion to the right of the node. Among the possible paths to the left, the demodulator will prefer the one with the smallest partial path metric, called the survivor path for that node [13]. We discard all partial paths to the left except for the survivor path and we store the survivor path and its partial metric for each node.

At each stage of the trellis we do not know which node the optimum path passes through, so we keep one survivor path for each and every node. To get this survivor path for a given node at stage k, we look at all the branches leading to that node from stage k-1, and at the partial path metrics of the nodes from which they stem. The partial path metric for each of those paths is determined by summing the partial path metric of the survivor at time k-1 and the current branch metric. The survivor path at that node is chosen as the path with the smallest path metric.

We illustrate the Viterbi algorithm by using a simple example. Consider the trellis shown in Fig. 4-7, which is the original trellis diagram, where the branch metric for each branch is shown. Fig. 4-8 shows how Viterbi algorithm works stage by stage. In Fig. 4-8, the survivor paths (solid line) at each node, and the partial path metric of each survivor path (number within the parenthesis) are shown. It is seen that the Viterbi algorithm iteratively searches the path with the minimum path metric, see Fig. 4-8 e.

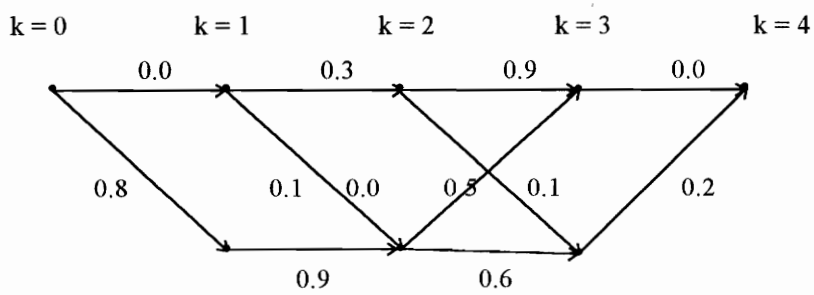
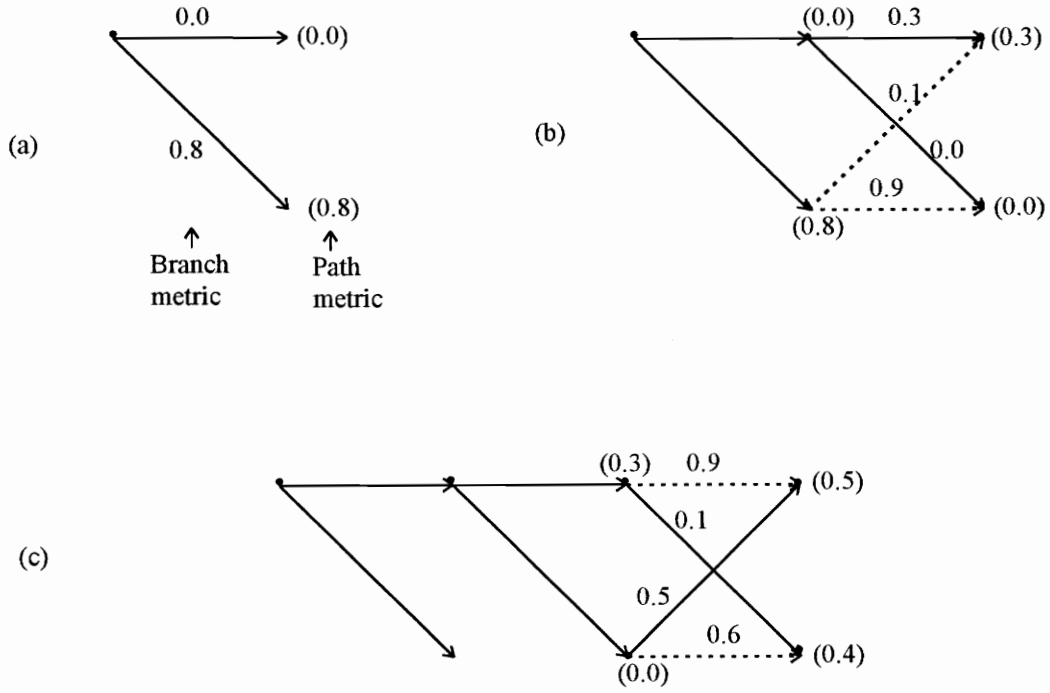



Figure 4-7. Original Trellis Diagram



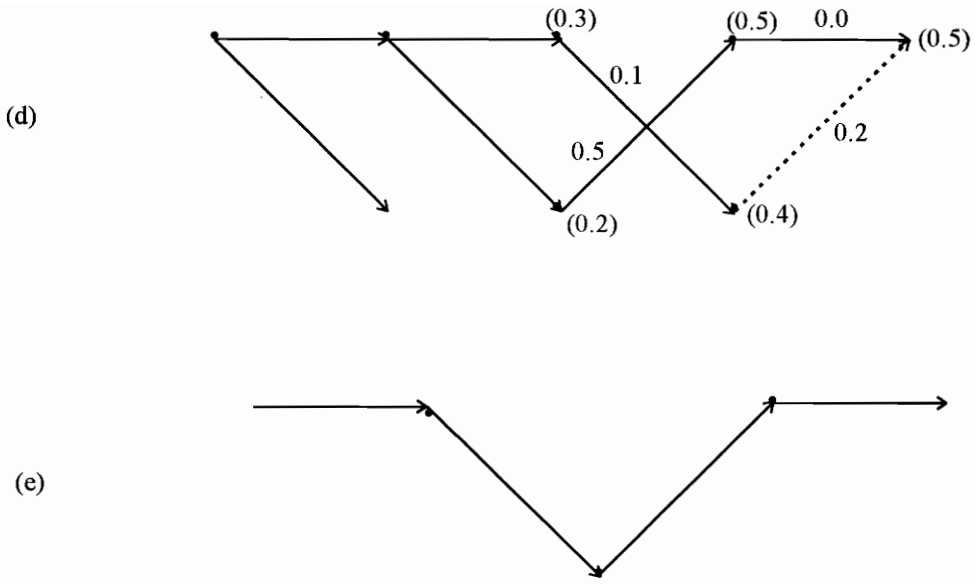



Figure 4-8. The Viterbi algorithm iteratively finds the path with the minimum path metric [13]

In the presence of parallel transitions, i.e., more than one signal is associated with the transition from one trellis state to the next, a small modification should be made in the algorithm. In this situation, the decoder compares, at each branch in the trellis, the received signal to each of the signals allowed for that branch. The branch metric with the smallest value is saved in the memory and is used thereafter for that branch, and the Viterbi algorithm can proceed as in the case of nonparallel transitions.

## 4.5 Conclusion

In this chapter, we have summarized the basic principles of TCM for use in coherent communication systems. These codes have proven useful for a number of important applications, including telephone modem, fixed-point to fixed-point microwave and are being considered for use in satellite communication systems [24]. However, in some instances, such as mobile paging applications, coherent reception may not be economically possible. For this reason, in subsequent chapters, we explore a novel technique for using trellis coded modulation with noncoherent reception.

# Chapter 5 Trellis Coding for Permutation Modulation

## 5.1 Introduction

We have discussed a noncoherent modulation scheme, FSK permutation modulation in chapter 3. It has been shown that using FSK permutation modulation can increase spectral efficiency at the expense of energy efficiency. In this chapter we explore the application of TCM to permutation modulation, in order to increase the energy efficiency of the system without trading of its spectral efficiency. We call this new technology trellis-coded permutation modulation (TCPM).

Basically, the same concepts as in TCM are applied to TCPM. Set partitioning and selection of partitions by convolutional encoders are also the keys in TCPM. However, because of the characteristics of noncoherent modulation the signal distance and the branch metric in TCPM are defined in a different way. We studied a few examples, such as trellis-coding for $\binom{6}{3}$ permutation modulation and trellis-coding for $\binom{8}{4}$ permutation modulation. The simulation results show that modest performance improvement has been achieved with the TCPM systems.

We begin this chapter by describing set partitioning for permutation modulation, followed by the selection of convolutional codes. Then the decoding of TCPM codes is discussed. The emphasis is placed on the definition of the new metric. In the next section the simulations of the two TCPM examples were outlined. Finally, we draw the plot of energy and spectral efficiency again and compare with the plot which was shown in Chapter 3.

## 5.2 Signal Set of Permutation Modulation

It is discussed in Chapter 3 that with $\binom{m}{n}$ FSK permutation modulation, $n$ out of $m$ carrier frequencies are ON at each time while the rest ($m$-$n$) frequencies are OFF, see Fig. 5-1 for $\binom{6}{3}$ FSK permutation modulation. The total distinct combination of signal patterns is $\log_2\binom{m}{n}$. The distance between any pair of signals is defined as the number of frequencies which are not in the same state. For example, the distance between signal pattern of A and B in Figure 5-1 is 2, while the distance between A and C is 4.

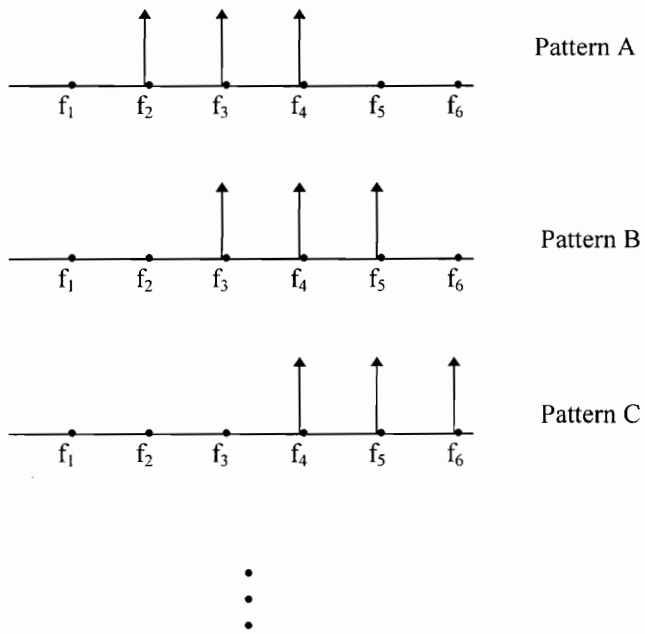



Figure 5-1. Signal Patterns of $\binom{6}{3}$ FSK Permutation Modulation

If we symbolize the ON frequencies as 1 and the OFF frequencies as 0, we can have a simpler representation of the signal patterns of the permutation modulation. Table 5-1 shows all the signal characters of $\binom{6}{3}$ FSK permutation modulation, where at

each symbol interval, 3 out of 6 frequencies are ON and the remaining 3 frequencies are OFF.

Table 5-1. Signal Sets of $\binom{6}{3}$ FSK Permutation Modulation

| Signal Characters | Patterns | | | | | |
|---|---|---|---|---|---|---|
| | $f_1$ | $f_2$ | $f_3$ | $f_4$ | $f_5$ | $f_6$ |
| P1 | 0 | 0 | 0 | 1 | 1 | 1 |
| P2 | 0 | 0 | 1 | 0 | 1 | 1 |
| P3 | 0 | 0 | 1 | 1 | 0 | 1 |
| P4 | 0 | 0 | 1 | 1 | 1 | 0 |
| P5 | 0 | 1 | 0 | 0 | 1 | 1 |
| P6 | 0 | 1 | 0 | 1 | 0 | 1 |
| P7 | 0 | 1 | 0 | 1 | 1 | 0 |
| P8 | 0 | 1 | 1 | 0 | 0 | 1 |
| P9 | 0 | 1 | 1 | 0 | 1 | 0 |
| P10 | 0 | 1 | 1 | 1 | 0 | 0 |
| P11 | 1 | 0 | 0 | 0 | 1 | 1 |
| P12 | 1 | 0 | 0 | 1 | 0 | 1 |
| P13 | 1 | 0 | 0 | 1 | 1 | 0 |
| P14 | 1 | 0 | 1 | 0 | 0 | 1 |
| P15 | 1 | 0 | 1 | 0 | 1 | 0 |
| P16 | 1 | 0 | 1 | 1 | 0 | 0 |
| P17 | 1 | 1 | 0 | 0 | 0 | 1 |
| P18 | 1 | 1 | 0 | 0 | 1 | 0 |
| P19 | 1 | 1 | 0 | 1 | 0 | 0 |
| P20 | 1 | 1 | 1 | 0 | 0 | 0 |

As another example, signal sets of $\binom{8}{4}$ FSK permutation modulation is displayed on Table 5-2. Note that in this case 4 frequencies are ON and the remaining 4 frequencies are OFF at each transmission interval.

Table 5-2. Signal Patterns of $\binom{8}{4}$ FSK Permutation Modulation

| Signal Characters | Patterns | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | $f_1$ | $f_2$ | $f_3$ | $f_4$ | $f_5$ | $f_6$ | $f_7$ | $f_8$ |
| P1 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 |
| P2 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 1 |
| P3 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 1 |
| P4 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 1 |
| P5 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 0 |
| P6 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 1 |
| P7 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 1 |
| P8 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 |
| P9 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 0 |
| P10 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 |
| P11 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 1 |
| P12 | 0 | 0 | 1 | 1 | 0 | 1 | 1 | 0 |
| P13 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 1 |
| P14 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 0 |
| P15 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 0 |
| P16 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 1 |
| P17 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 1 |
| P18 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 1 |
| P19 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 0 |
| P20 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 1 |
| P21 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 |
| P22 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 0 |
| P23 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 1 |
| P24 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 0 |
| P25 | 0 | 1 | 0 | 1 | 1 | 1 | 0 | 0 |
| P26 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 1 |
| P27 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 1 |
| P28 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 0 |
| P29 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 1 |
| P30 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 0 |
| P31 | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 0 |
| P32 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 1 |
| P33 | 0 | 1 | 1 | 1 | 0 | 0 | 1 | 0 |
| P34 | 0 | 1 | 1 | 1 | 0 | 1 | 0 | 0 |

(continued)

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| P35 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 |
| P36 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 1 |
| P37 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 1 |
| P38 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 1 |
| P39 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 0 |
| P40 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 1 |
| P41 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 1 |
| P42 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 0 |
| P43 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 1 |
| P44 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 0 |
| P45 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 0 |
| P46 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 1 |
| P47 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 1 |
| P48 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 0 |
| P49 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 |
| P50 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 |
| P51 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 0 |
| P52 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 1 |
| P53 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 0 |
| P54 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 0 |
| P55 | 1 | 0 | 1 | 1 | 1 | 0 | 0 | 0 |
| P56 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 |
| P57 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 1 |
| P58 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 0 |
| P59 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 1 |
| P60 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 0 |
| P61 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 |
| P62 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 1 |
| P63 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 0 |
| P64 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 |
| P65 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 0 |
| P66 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 1 |
| P67 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 |
| P68 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 0 |
| P69 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 |
| P70 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 |

## 5.3 Set Partitioning of Permutation Modulation

### 5.3.1 $\binom{6}{3}$ Permutation Modulation

It is shown in Table 5-1 that the number of signal characters in $\binom{6}{3}$ permutation modulation is 20. We use only 16 out of 20 to transmit a 4 bits/symbol signal. It is noticed that the distance between any pair of signals is 2, 4 or 6. We define the minimum distance of a permutation modulation scheme to the shortest distance between any pair of signal characters within the signal set. Therefore the minimum distance of this permutation modulation case is 2. We group those with distance of 6 and have the following eight groups which are labeled as C1, C2, ..., C8.

$$C1 = \begin{Bmatrix} 000111 \\ 111000 \end{Bmatrix} = \{P1, P20\} \qquad C5 = \begin{Bmatrix} 001101 \\ 110010 \end{Bmatrix} = \{P3, P18\}$$

$$C2 = \begin{Bmatrix} 011100 \\ 100011 \end{Bmatrix} = \{P10, P11\} \qquad C6 = \begin{Bmatrix} 011001 \\ 100110 \end{Bmatrix} = \{P5, P16\}$$

$$C3 = \begin{Bmatrix} 101010 \\ 010101 \end{Bmatrix} = \{P15, P6\} \qquad C7 = \begin{Bmatrix} 011100 \\ 100011 \end{Bmatrix} = \{P8, P13\}$$

$$C4 = \begin{Bmatrix} 110001 \\ 001110 \end{Bmatrix} = \{P17, P4\} \qquad C8 = \begin{Bmatrix} 110100 \\ 001011 \end{Bmatrix} = \{P19, P2\}$$

Figure 5-2. Signal Groups with Distance of 6.

If we group C1 through C4 and C5 through C8 together, forming new groups B1 and B2, as in Figure 5-3, we find that any pair of the 4 signal characters in the first row or any signal characters in the second row in both B1 and B2 has distance of 4. This means that any one signal character in B1 and B2 has distance of 4 or more to four other signal characters within the same group, although the minimum distance in B1 and B2 is still 2.

$$B1 = \begin{Bmatrix} 000111 & 011100 & 101010 & 110001 \\ 111000 & 100011 & 010101 & 001110 \end{Bmatrix} = \{C1, C2, C3, C4\}$$

$$B2 = \begin{Bmatrix} 001101 & 011001 & 011100 & 110100 \\ 110010 & 100110 & 100011 & 001011 \end{Bmatrix} = \{C5, C6, C7, C8\}$$

Figure 5-3. Signal Groups with Distance of 2

The set partitioning of $\binom{6}{3}$ permutation modulation, which is illustrated in Fig. 5-4, is actually the reverse procedure of the above grouping. It shows that the original signal set A is partitioned into subsets B1 and B2, which are further partitioned into subsets C1, C2, C3, C4 and C5, C6, C7, C8 respectively. It is seen that the minimum distance from level B to level C increases from 2 to 6. Unfortunately, the minimum distance does not change from level A to level B.

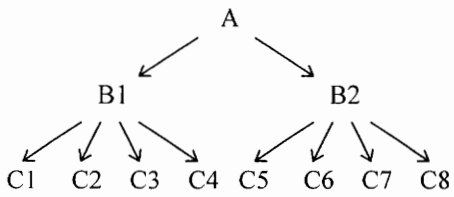



Figure 5-4. Set Partitioning of $\binom{6}{3}$ Permutation Modulation

### 5.3.2 $\binom{8}{4}$ Permutation Modulation

It is seen from Table 5-2 that we have much more signal patterns in the case of $\binom{8}{4}$ permutation modulation. Consequently, more interesting situations occur in the set partitioning operation. We shall consider 64 out of 70 total signal characters in the signal set, which corresponds to 6 bits per symbol.

The same procedure as in the case of $\binom{6}{3}$ permutation modulation is followed to achieve the set partitioning. We begin by grouping those signal characters with the largest distance which is 8. Obviously, there are 35 groups available, but only 32 of them are considered. They are labeled as D1, D2, ..., D32.

$$D1 = \begin{Bmatrix} 00001111 \\ 11110000 \end{Bmatrix} = \{P1, P70\} \qquad D17 = \begin{Bmatrix} 11011000 \\ 00100111 \end{Bmatrix} = \{P65, P6\}$$

$$D2 = \begin{Bmatrix} 11001100 \\ 00110011 \end{Bmatrix} = \{P61, P10\} \qquad D18 = \begin{Bmatrix} 11100001 \\ 00011110 \end{Bmatrix} = \{P66, P5\}$$

$$D3 = \begin{Bmatrix} 10101010 \\ 01010101 \end{Bmatrix} = \{P50, P21\} \qquad D19 = \begin{Bmatrix} 10010101 \\ 01101010 \end{Bmatrix} = \{P41, P30\}$$

$$D4 = \begin{Bmatrix} 11000011 \\ 00111100 \end{Bmatrix} = \{P56, P15\} \qquad D20 = \begin{Bmatrix} 10110010 \\ 01001101 \end{Bmatrix} = \{P53, P18\}$$

$$D5 = \begin{Bmatrix} 11101000 \\ 00010111 \end{Bmatrix} = \{P69, P2\} \qquad D21 = \begin{Bmatrix} 10101100 \\ 01010011 \end{Bmatrix} = \{P51, P20\}$$

$$D6 = \begin{Bmatrix} 10110001 \\ 01001110 \end{Bmatrix} = \{P52, P19\} \qquad D22 = \begin{Bmatrix} 10000111 \\ 01111000 \end{Bmatrix} = \{P36, P35\}$$

$$D7 = \begin{Bmatrix} 11010100 \\ 00101011 \end{Bmatrix} = \{P64, P7\} \qquad D23 = \begin{Bmatrix} 11001010 \\ 00110101 \end{Bmatrix} = \{P60, P11\}$$

$$D8 = \begin{Bmatrix} 01011001 \\ 10100110 \end{Bmatrix} = \{P23, P48\} \qquad D24 = \begin{Bmatrix} 10011001 \\ 01100110 \end{Bmatrix} = \{P43, P29\}$$

$$D9 = \begin{Bmatrix} 10100011 \\ 01011100 \end{Bmatrix} = \{P46, P25\} \qquad D25 = \begin{Bmatrix} 00110110 \\ 11001001 \end{Bmatrix} = \{P12, P59\}$$

$$D10 = \begin{Bmatrix} 11010010 \\ 00101101 \end{Bmatrix} = \{P63, P8\} \qquad D26 = \begin{Bmatrix} 00111001 \\ 11000110 \end{Bmatrix} = \{P13, P58\}$$

$$D11 = \begin{Bmatrix} 01000111 \\ 10111000 \end{Bmatrix} = \{P16, P55\} \qquad D27 = \begin{Bmatrix} 01100011 \\ 10011100 \end{Bmatrix} = \{P26, P45\}$$

$$D12 = \begin{Bmatrix} 00011011 \\ 11100100 \end{Bmatrix} = \{P3, P68\} \qquad D28 = \begin{Bmatrix} 01011010 \\ 10100101 \end{Bmatrix} = \{P24, P47\}$$

$$D13 = \begin{Bmatrix} 10110100 \\ 01001011 \end{Bmatrix} = \{P54, P17\} \qquad D29 = \begin{Bmatrix} 01110001 \\ 10001110 \end{Bmatrix} = \{P32, P38\}$$

$$D14 = \begin{Bmatrix} 01100101 \\ 10011010 \end{Bmatrix} = \{P27, P44\} \qquad D30 = \begin{Bmatrix} 00111010 \\ 11000101 \end{Bmatrix} = \{P14, P57\}$$

$$D15 = \begin{Bmatrix} 00011101 \\ 11100010 \end{Bmatrix} = \{P4, P67\} \qquad D31 = \begin{Bmatrix} 01010110 \\ 10101001 \end{Bmatrix} = \{P22, P49\}$$

$$D16 = \begin{Bmatrix} 00101110 \\ 11010001 \end{Bmatrix} = \{P9, P62\} \qquad D32 = \begin{Bmatrix} 01101100 \\ 10010011 \end{Bmatrix} = \{P31, P40\}$$

Figure 5-5. Groups with Distance of 8.

It is seen that there are two signal characters in each group. The minimum distance is simply the distance between two signals, which is 8.

Next, we group the consecutive four groups as follows.

$$C1 = \begin{Bmatrix} 00001111 & 11001100 & 10101010 & 11000011 \\ 11110000 & 11001100 & 01010101 & 00111100 \end{Bmatrix} = \{D1, D2, D3, D4\}$$

$$C2 = \begin{Bmatrix} 11101000 & 10110001 & 11010100 & 01011001 \\ 00010111 & 01001110 & 00101011 & 10100110 \end{Bmatrix} = \{D5, D6, D7, D8\}$$

$$C3 = \begin{Bmatrix} 10100011 & 11010010 & 01000111 & 00011011 \\ 01011100 & 00101101 & 10111000 & 11100100 \end{Bmatrix} = \{D9, D10, D11, D12\}$$

$$C4 = \begin{Bmatrix} 10110100 & 01100101 & 00011101 & 00101110 \\ 01001011 & 10011010 & 11100010 & 11010001 \end{Bmatrix} = \{D13, D14, D15, D16\}$$

$$C5 = \begin{Bmatrix} 11011000 & 11100001 & 10010101 & 10110010 \\ 00100111 & 00011110 & 01101010 & 01001101 \end{Bmatrix} = \{D17, D18, D19, D20\}$$

$$C6 = \begin{Bmatrix} 10101100 & 10000111 & 11001010 & 10011001 \\ 01010011 & 01111000 & 00110101 & 01100110 \end{Bmatrix} = \{D21, D22, D23, D24\}$$

$$C7 = \begin{Bmatrix} 00110110 & 00111001 & 01100011 & 01011010 \\ 11001001 & 11000110 & 10011100 & 10100101 \end{Bmatrix} = \{D25, D26, D27, D28\}$$

$$C8 = \begin{Bmatrix} 01110001 & 00111010 & 01010110 & 01101100 \\ 10001110 & 11000101 & 10101001 & 10010011 \end{Bmatrix} = \{D29, D30, D31, D32\}$$

Figure 5-6. Groups with Distance of 4

It can be shown that the distance between any pair of signal characters within the same group is at least 4. Thus, the minimum distance of C1, C2, ..., C8 is 4.

Again, the set partitioning is the reverse procedure of the grouping. The set partitioning of $\binom{8}{4}$ permutation modulation is shown in Fig. 5-7,

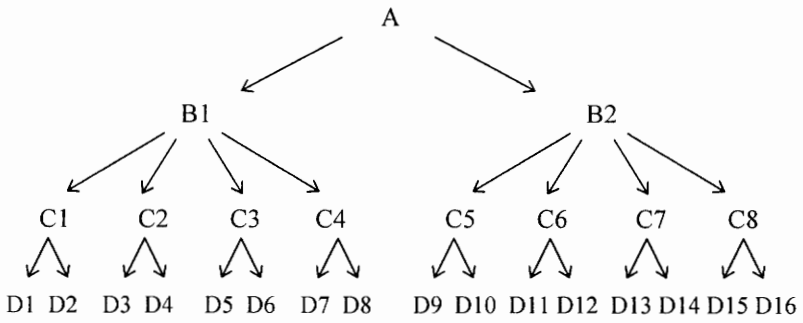



Figure 5-7. Set Partitioning of $\binom{8}{4}$ Permutation Modulation

where the original set is partitioned into B1 and B2, which are defined as

$$B1 = \{C1, C2, C3, C4\}$$
$$B2 = \{C5, C6, C7, C8\}$$

Figure 5-8. Groups with Distance of 2

It is noticed that the minimum distance of B1 and B2 is equal to 2, the same as that of the original set A. The two subset is in turn partitioned to from C1, C2, C3, C4 and C5, C6, C7, C8 respectively, the minimum distance is increased to 4. The last level of partitioning results in the minimum distance equal to 8.

## 5.4 Selection of Partitions by Convolutional Codes

As we mentioned earlier, the purpose of set partitioning is to have subsets which have a larger minimum distance than the original set. Then these subsets are mapped into different branches in a trellis diagram. The rule of the mapping is based on Ungerboeck design rule which is described in Chapter 4. The goal is to get the largest possible minimum free distance. Recall that the minimum free distance ( $d_{free}$ ) is defined as the minimum Euclidean distance between a pair of valid, distinct signal sequences.

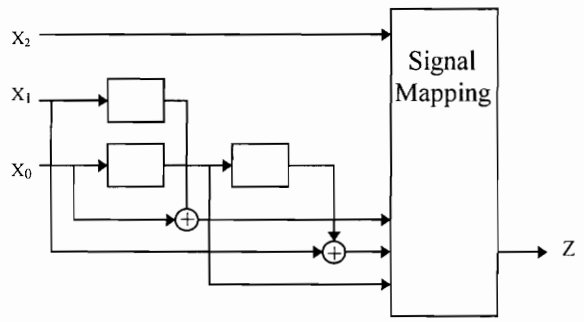



Figure 5-9. Eight-State Ungerboeck Encoder [5]

Figure 5-9 illustrates how the partitions of $\binom{6}{3}$ permutation modulation are mapped into the trellis diagram. The trellis in Figure 5-9 is the representation of a rate-2/3 with 8-state Ungerboeck encoder shown in Figure 5-8.

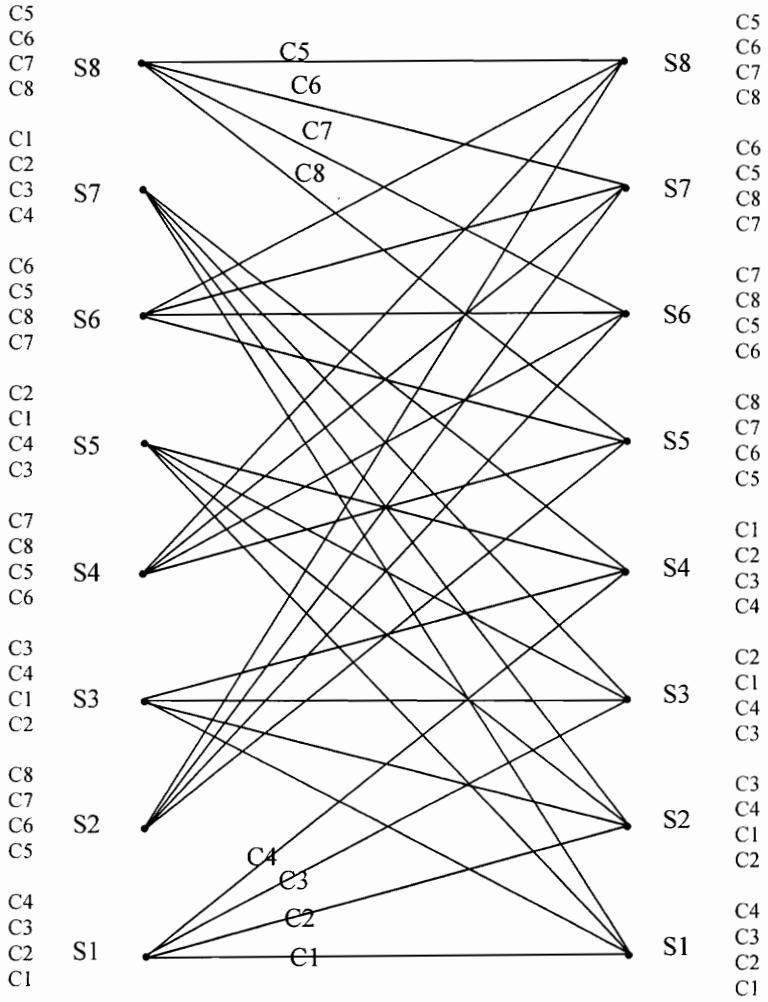



Figure 5-10. Selection of $\binom{6}{3}$ Partitions by a 8-State Convolutional Encoder

The output of the convolutional encoder selects one of the eight partitions (C1 through C8) of the $\binom{6}{3}$ permutation signal set. The uncoded bit $x_2$ is then used to choose one of the two signals in the selected partition. Ungerboeck design rules have been followed in assigning signals to the trellis branches.

It is noted that there are two signals associated with each branch in the trellis. It is called parallel transition. According to Ungerboeck design rule, parallel transitions are assigned to signals separated by the maximum Euclidean distance. Consequently, each branch in the trellis is assigned to one of the subsets with the maximum distance of 6, i.e., C1, ..., C8. The numbers next to the subsets are the output bits of the encoder. The branches emitting from and merging into any state should be assigned the signals from the preceding partitions, i.e., B1 or B2.

The selection of codes for $\binom{8}{4}$ permutation modulation is very similar to the case of $\binom{6}{3}$ permutation. If we consider the same rate-2/3 convolutional encoder in Figure 5-8, the $\binom{8}{4}$ permutation signal set should be partitioned into 8 subsets (recall that the degree to which the signal set is partitioned depends on the characteristics of the code), shown in Figure 5-10.

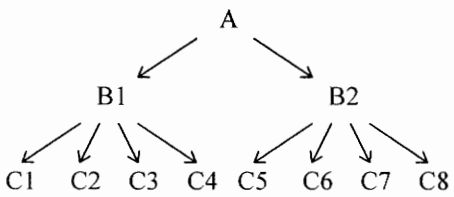



Figure 5-11. Set Partitioning of $\binom{8}{4}$ Signal Set Associated with Encoder in Figure 5-8

It is known that 3 uncoded bits are needed in this case so that the output consists of 6 bits. The 3 coded output bits are used to select one of the eight partitions of the signal set shown at the bottom of Figure 5-10. The remaining three uncoded bits are used to select one of the eight signals in the selected partition. Consequently, the trellis diagram with each branch assigned a proper partition looks exactly the same as in Figure 5-8. But the meaning of the partitions is different.

## 5.5 Decoding of TCPM

As we described in Chapter 4 that the goal of decoding of TCM is to estimate the path that the encoded signal sequence traversed through the trellis. This is done by associating with each branch of the trellis a metric, and looking for the path whose total metric is the smallest [13]. The decoding of TCPM is very similar to that of TCM. The only difference is the definition of branch metric. We have shown in the previous chapter that the metric for coherent modulation is the squared Euclidean distance. Because of the characteristics of the noncoherent modulation, a new metric is required in the decoding of TCPM.

It is shown in Chapter 2 that for noncoherent M-ary FSK signal, the optimum decision rule in the demodulator is to compare the energy in the M frequency bands and pick the largest one. Accordingly, the decision rule for FSK permutation modulation is to compare the energy in all the different signal patterns and pick the largest. This tells us that the branch metric in TCPM should be the energy of all the candidate signals.

For a specific branch which associates to a received signal $y_k$

$$y_k = \left[ y_{k,1}, y_{k,2}, \ldots, y_{k,m} \right] \tag{5-1}$$

where $y_{i,k}$ represents the value of $y_k$ on the ith frequency.

Then the metric of this branch is defined as

$$m\ (y_k) = \max(\|A_i\|^2) = \max(\sum_{i \in A_i} \|y_{k,i}\|^2), \qquad i = 1,\ 2,\ \ldots,\ N \tag{5-2}$$

where $N$ is the number of patterns which are allowed for that branch

$A_i$ is the set of frequencies which are on for pattern $i$.

Therefore, the decoding of TCPM is performed in two steps. First, at each branch in the trellis, the decoder compares the energy among all the signals allowed for that branch. The identity with the largest value is saved in memory, and the branch is labeled with that value which is the branch metric. Then, the Viterbi algorithm is applied to the trellis, in exactly the same way as in TCM decoding, but instead of finding the path with the smallest total metric, we are now interested in the path with the largest total path metric

## 5.6 Simulation of TCPM

In the previous sections, TCPM is discussed through the examples of trellis coded $\binom{6}{3}$ and $\binom{8}{4}$ permutation modulation. The simulations of these two examples are presented in this section. The design procedures are outlined in details and the simulation results are also provided.

Figure 5-11 shows the block diagram of the simulation.

1. Generate random numbers.

For $\binom{6}{3}$ permutation, we generated a sequence of signals with 3 bits per symbol, while for $\binom{8}{4}$ permutation, a sequence of signals with 5 bits per symbol is generated.

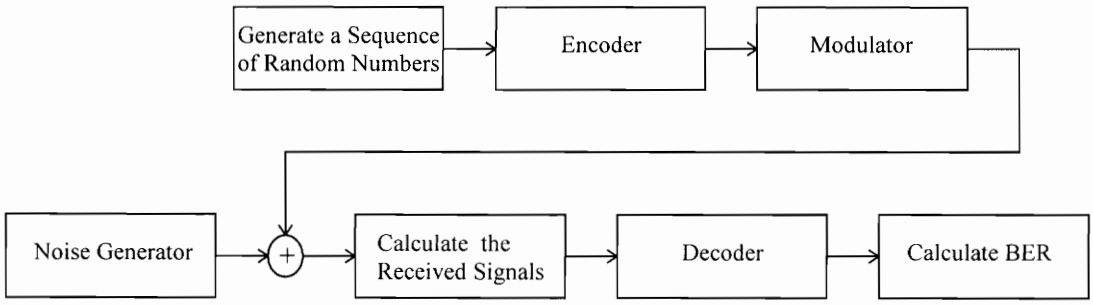



Figure 5-12. Block Diagram for Simulation of Trellis-Coded Permutation Modulation

2. Encode the generated symbols

Two different convolutional encoders are applied to both $\binom{6}{3}$ and $\binom{8}{4}$ cases. They are chosen from the so called best known codes with maximum free distance [5]. One is rate-2/3 with 8-state encoder. The impulse response of the encoder is

$$\begin{aligned} g_0^{(0)} &= 100 \quad g_0^{(1)} = 010 \quad g_0^{(2)} = 110 \\ g_1^{(0)} &= 001 \quad g_1^{(1)} = 100 \quad g_1^{(2)} = 111 \end{aligned} \tag{5-3}$$

Another encoder is rate-3/4 with 32 state. The impulse response of the encoder is

$$\begin{aligned} g_0^{(0)} &= 100 \quad g_0^{(1)} = 010 \quad g_0^{(2)} = 010 \quad g_0^{(3)} = 110 \\ g_1^{(0)} &= 001 \quad g_1^{(1)} = 110 \quad g_1^{(2)} = 000 \quad g_1^{(3)} = 111 \\ g_2^{(0)} &= 000 \quad g_2^{(1)} = 010 \quad g_2^{(2)} = 101 \quad g_2^{(3)} = 101 \end{aligned} \tag{5-4}$$

Note that the impulse response describes the relationship among the input and output signal, as well as the state in an elegant way. It is a complete representation of an encoder. The outputs of the encoder depends on the inputs and the state of the encoder. For a given state, any input signal corresponds to an unique output signal.

Assume the initial state of the encoder is zero, the encoding the input sequence is implemented by the following equation [5].

$$y_j^{(i)} = \sum_{t=0}^{k-1}\left(\sum_{l=0}^{m} x_{j-l}^{(t)} g_{t,l}^{(i)}\right) \tag{5-5}$$

Where $y^{(i)}_j$ denotes the $j$th bit of the ith output of the encoder, $x^{(i)}_l$ is the $l$th bit of the ith input of the encoder, k is the number of inputs, m is the number of the memory elements.

3. Modulate the outputs of the encoder

Actually, the modulation and encoding are done in an integrated way, which is the core of TCPM system. To demonstrate the modulation, we use rate-2/3 convolutional code with $\binom{6}{3}$ permutation as an example. The other combinations of encoder and permutation modulation work in pretty much the same way.

We redraw the trellis diagram with the assignment of the permutation modulation from Figure 5-9, in Figure 5-12.

Figure 5-12 shows the assignment of subset to each branch. It is known that each branch also corresponds to a particular output of the encoder. Thus, every output of the encoder is associated with a unique subset, which the modulation is based. The outputs that each branch corresponds are shown in Figure 5-12. Table 5-3 summarizes the modulation of $\binom{6}{3}$ permutation with the rate-2/3 convolutional encoder.

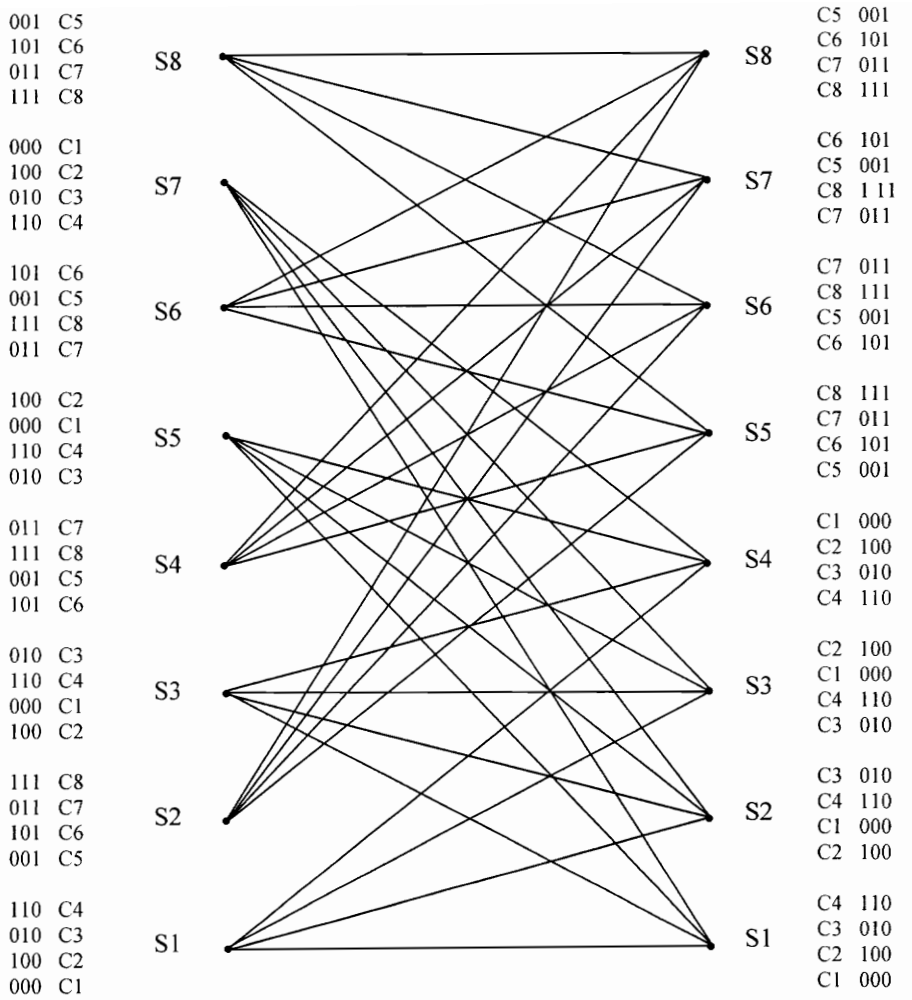



Figure 5-13. Assignment of the Subsets on the Trellis Diagram of a Rate - 2/3 with 8 State Convolutional Encoder

Table 5-3. Modulated Signal Pattern of Coded Bits

| Encoder Outputs | | | | Modulated Signal Patterns | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $y_3$ | $y_2$ | $y_1$ | $y_0$ | $f_1$ | $f_2$ | $f_3$ | $f_4$ | $f_5$ | $f_6$ |
| 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 |
| 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 1 |
| 1 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 1 | 0 |
| 0 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 |
| 1 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 1 |
| 0 | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 1 |
| 1 | 0 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 |
| 0 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 |
| 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 1 |
| 0 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 0 | 0 |
| 1 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 1 |
| 0 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 1 |
| 1 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 1 |
| 0 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 1 |
| 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 0 |

Once the modulated signal patterns are determined, the transmitted signals are generated in the same way as in Chapter 3. We use complex envelope to represent the modulated signal. For those frequencies which are ON, the in-phase and quadrature representation of the signal is as follows.

$$\begin{aligned} fI_i &= \cos\left[\pi(2i-1-16)+\theta_i\right] \\ fQ_i &= \sin\left[\pi(2i-1-16)+\theta_i\right] \end{aligned} \qquad i = 1, 2, \ldots, 6 \tag{5-4}$$

4. Generate noise

Again, AWGN is assumed to be the channel model. The generation of white Gaussian noise is in the same way as in Chapter 3. The calculation of noise variance σ with respect to the signal-to-noise ratio $E_b/N_0$ is by the following equation.

$$\sigma = \sqrt{\frac{0.5 \times 3}{6(E_b / N_0)}} \tag{5-5}$$

5. Calculate the received signal and the decision variable

The received signal is simply the transmitted signal plus noise.

$$\begin{aligned} rI &= fI + nI \\ rQ &= fQ + nQ \end{aligned} \tag{5-6}$$

where $nI$ and $rQ$ are the in-phase and quadrature parts of the noise, $rI$ and $rQ$ are the in-phase and quadrature of the received signal.

The decision variable $z$ is the energy of the received signal.

$$z = rI^2 + rQ^2 \tag{5-7}$$

6. Decoding of the received signal

The decoding of the received signal is based on the trellis diagram shown in Figure 5-9. At each stage, the decoder computes the partial metric for all the paths entering a state by adding the branch metric entering that state to the metric of the connecting survivor at the preceding stage. For each state, store the path with the largest metric and its metric and eliminate all other paths. If the partial metric ties occur, there is no meaningful decision can be made regarding the best path segment. In this case the decoder can select one path arbitrarily. This is done for every state in each stage until the input sequence finishes. At the end, the decoder determines the estimates of the

transmitted data by reversing the direction in the trellis and following the branches which are saved, back to the beginning stage.

7. Calculate the Bit Error Rate (BER)

The estimated data is compared to the generated data one by one and the error is accumulated. Thus the bit error rate can be easily calculated.

The simulation results of trellis-coding for $\binom{6}{3}$ and $\binom{8}{4}$ permutation modulation in terms of BER vs. $E_b/N_0$ are shown in Figure 5-13 and Figure 5-14 respectively.

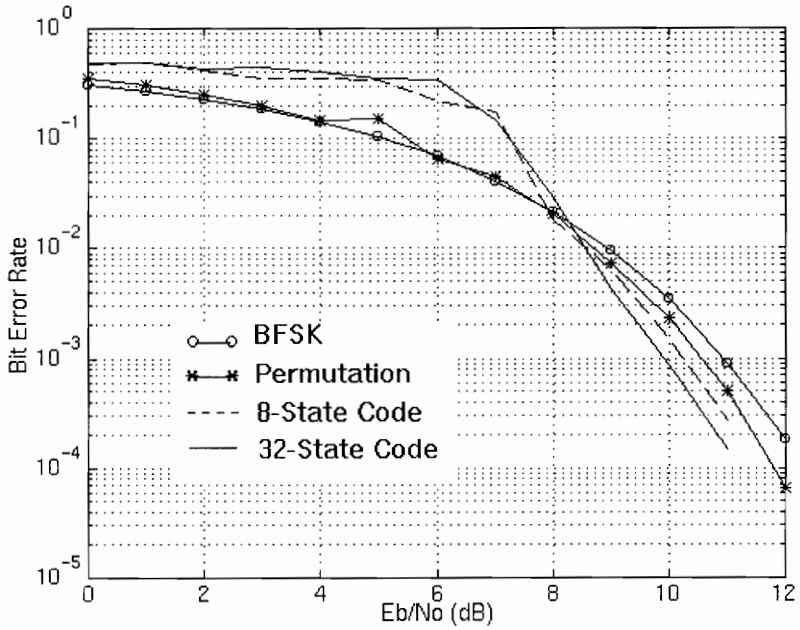



Figure 5-14. BER vs. $E_b/N_0$ for Trellis-Coded $\binom{6}{3}$ Permutation Modulation

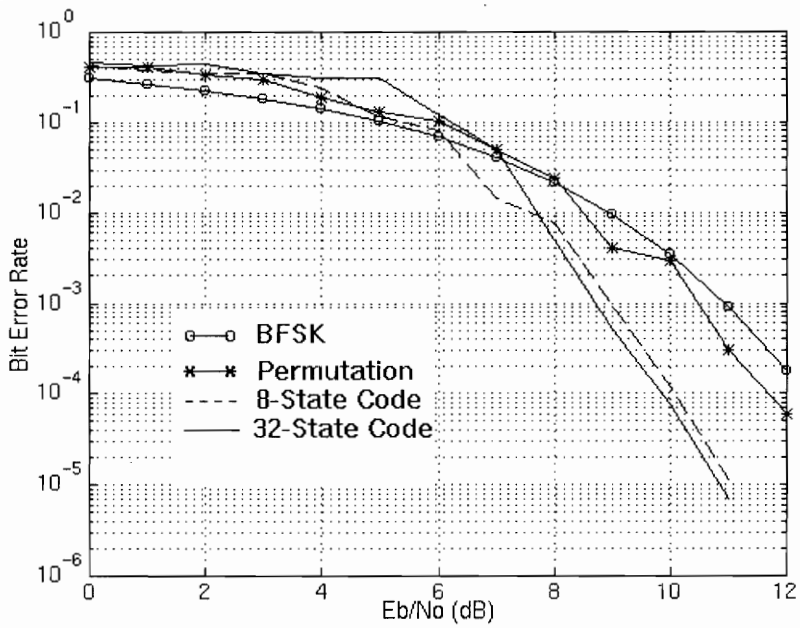



Figure 5-15. BER vs. $E_b/N_0$ for Trellis-Coded $\binom{8}{4}$ Permutation Modulation

For comparison, the simulation results for $\binom{6}{3}$ and $\binom{8}{4}$ permutation modulation are also shown on the related figures. In each figure, we illustrate the simulation results for using two different convolutional encoders, one is 8-state and the other is 32-state. Both figures show that the trellis coding provide modest performance improvement over the permutation modulation in terms of BER. The increase in the constraint length and state of the convolutional code have moderate effect on the asymptotic coding gain. This is also noted that a slightly increase in the constraint length results in a substantially increase of the complexity of the trellis and thus the decoder.

To evaluate the overall performance of the trellis-coded permutation modulation, we draw a new energy efficiency & spectral efficiency graph in Figure 5-15, comparing to the similar graph in Figure 3-4.

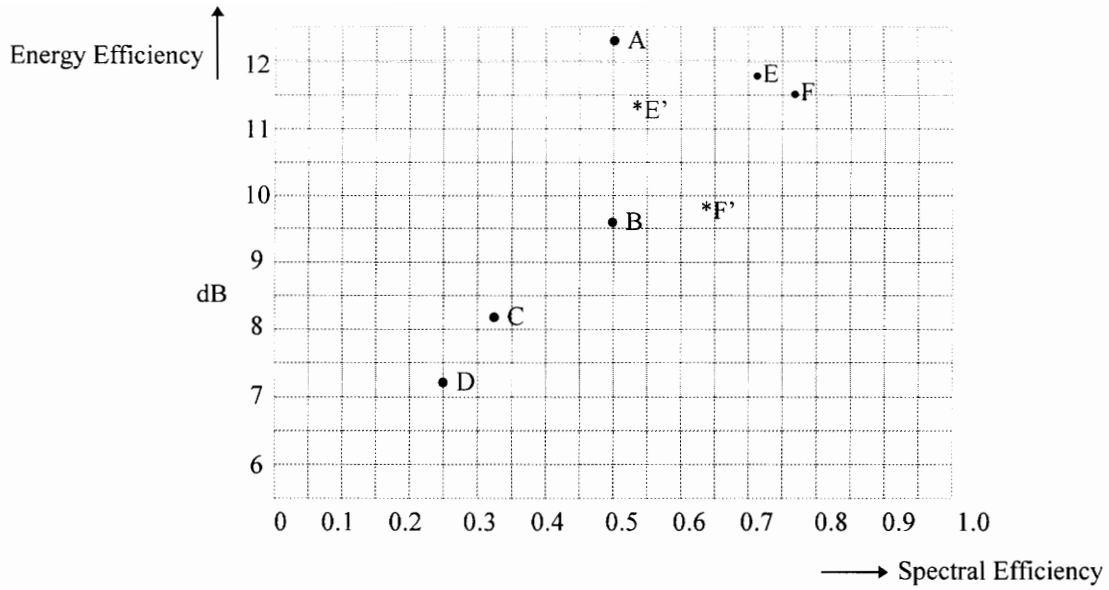



Figure 5-16. New Energy Efficiency & Spectral Efficiency

| Symbol | A | B | C | D | E | F | E′ | F′ |
|---|---|---|---|---|---|---|---|---|
| Alphabet | $\binom{2}{1}$ | $\binom{4}{1}$ | $\binom{8}{1}$ | $\binom{16}{1}$ | $\binom{6}{3}$ | $\binom{8}{4}$ | $\binom{6}{3}$ (with trellis coding) | $\binom{8}{4}$ |

## 5.7 Conclusion

In this chapter, we introduce a new technique, trellis-coding for permutation modulation (TCPM). It is the application of TCM to FSK permutation modulation, a noncoherent modulation scheme. We have shown that the same concepts behind TCM are also applicable to TCPM. The keys of TCPM are also set partitioning and selection of partitions by convolutional encoder. In the decoder, both TCM and TCPM use the

Viterbi algorithm to perform decoding and demodulation. The only difference is the definition of branch metric. For TCM, the branch metric is defined as the smallest squared of Euclidean distance between the received signal and the candidate signal. However, for TCPM, based on the characteristic of noncoherent modulation, we define a new metric which leads to superior performance.

Examples of TCPM scheme, such as trellis-coding for $\binom{6}{3}$ and $\binom{8}{4}$ permutation modulation, are studied and simulated. It is shown that using trellis coding to permutation modulation can modestly improve the performance. It is possible to achieve some combination of energy and spectral efficiency, which are not attainable with other techniques.

# Chapter 6 Conclusion and Future Work

## 6.1 Conclusion

Trellis-coded modulation (TCM) is a technique which combines convolutional coding and modulation. Ungerboeck, as well as other researchers have shown that using trellis-coded modulation can effectively provide coding gain without increasing the bandwidth and energy requirement.

TCM is usually applied to coherent modulation, i.e., QPSK and QAM. In this thesis, we explored the application of TCM to FSK permutation modulation, a noncoherent modulation scheme. FSK permutation modulation is a subclass of general permutation modulation. For $\binom{m}{n}$ permutation modulation, the information signal is transmitted simultaneously on n out of m carrier frequencies. Thus it can convey $\log_2\binom{m}{n}$ information bits per symbol. It is shown that using permutation modulation can increase the spectral efficiency, which is attractive to bandwidth limited applications. However, the improvement of spectral efficiency is obtained by trading of the energy efficiency.

In order to provide some coding gain to FSK permutation modulation without increasing the bandwidth utilization as in conventional coding, we apply trellis coding to FSK permutation modulation (TCPM). The same concepts as in TCM are employed in TCPM. Set partitioning and selection of partitions by convolutional encoder are the keys in the encoding procedure. In the receiver, the Viterbi algorithm is used as in standard TCM, to decode the sequence and estimate the transmitted signal. But the definition of metric in TCPM is different from that in TCM. It is known that the squared

Euclidean distance between the received signal and the candidate signals is used in TCM as metric. While in TCPM system, signal energy is used as the metric for the noncoherent reception.

Examples, such as trellis coding for $\binom{6}{3}$ and $\binom{8}{4}$ FSK permutation modulation, are studied and the performance in terms of error probability were simulated. The simulation results show that trellis coding can provide modest performance improvement. As a result, trellis-coded permutation modulation can possibly achieve some combinations of energy efficiency and spectral efficiency which are not otherwise achievable using noncoherent modulation.

## 6.2 Future Work

Several extensions of this work may be possible. It may be possible to construct other trellis structures which may be optimal for noncoherent modulation. The trellis structures we chose in our work are Ungerboeck encoders which are known to be good for coherent modulation. However they might not be the best codes for noncoherent modulation. Searching for optimal encoders for noncoherent modulation may be a useful and interesting research topic.

Another future extension to this work may be to develop analytical results which describe the performance for trellis-coded permutation modulation. This will help us better understand this new technique, and supplement the simulation performance analysis.

Exploration of other metrics in the decoding procedure of TCPM might also be a useful extension of the work. There might exits some metrics which provide better performance, although we believe that signal energy is a good metric for noncoherent modulation.

Finally, the application of trellis-coded permutation modulation to practical usage should be explored in future work, particularly, for paging system

# Reference


1. G.Ungerboeck, "Channel coding with multilevel/phase signals," IEEE Trans. Information Theory, vol. 1T-28, pp. 56-67, Jan. 1982.

2. G. Ungerboeck, "Trellis-coded modulation with redundant signal sets. Part I: Introduction; Part II: State of the art," IEEE Communication Mag., vol. 25, No. 2, Feb. 1987.

3. Rade Petrevic, "Multicarrier permutation for narrowband PCS," Proceeding on Virginia Tech Wireless Personal Communication Symposium, June 1993.

4. E. Brookner, "The permutation modulations in fading channels and their comparison based on a general method for the comparison of M-ary Modulations," IEEE Transaction on Communication Technology., vol. COM-17. Mp/ 6. Dec. 1969, pp. 616-639.

5. Stephen B. Wacher, Error Control Systems for Digital Communication and Storage", Prentice Hall, Englewood NJ 07632, 1995.

6. Leon W. Couch II, Digital and Analog Communication Systems", 4th edition, Macmillan Publishing Company, 1993.

7. John G. Proakis, "Digital communications", McGraw-Hill Books Company, 1983.

8. Michael C. Jeruchim, Philip Balaban, K. Sam Shanmugan, Simulation of Communication Systems, Plenum Press, NY, 1992.

9. Rodger E. Ziemer, Rodger L. Peterson, Introduction to Digital Communication,

Macmillan Publishing Company, NY, 1992.

10. H. L. Schneider, "Data transmission with permutation modulation," Bell Sys. Tech. J., vol. 47, pp. 1131-1138, July/August 1968.

11. Rieger, S., "Error rate in data transmission," Proc. IRE, 46 (May 1958) pp. 919-920.

12. Helstrom, C. W., "The resolution of signals in white Gaussian noise," Proc. IRE, 43 (September 1955), pp. 1111-1118.

13. Ezio Biglieri, Dariush Divsalar, Peter J. McLane, Marvin K. Simon, Introduction to Trellis-Coded Modulation with Applications, Macmillan Publishing Company, NY, 1991.

14. David Slepian, "Permutation Modulation", Proceedings of the IEEE, March 1965, pp. 228-237

15. S. S. Pietiobon, R. H. Deng, A. Lafannechere, G. Ungerboeck and D. J. Costello Jr, "Trellis-coded multidimensional phase modulation," IEEE Transactions on Information Theory, Vol. IT-36, No 1, pp 63-89, Jan. 1990.

16. Brian Woerner, "Coded modulation for direct-sequence spread spectrum communication," Ph.D Dissertation, University of Michigan, August 1991.

17. R. Calderbank and J. E. Mazo, "A new description of trellis codes," IEEE Transaction on Information Theory, vol, IT-30, no. 6, pp784-791, November 1984.

18. D.Divsalar and M. K. Simon, "Trellis coded modulation for 4800-9600 bits/s transmission over a fading mobile satellite channel," IEEE Journal on Selection Area in Communications, vol. SAC-5, no. 2. pp162-174, February 1987.

19. IBM Eurpoe, "Trellis-coded modulation scheme with 8-state symmetric encoder and 90° symmetric for use in data modems transmitting 3-7 bits per modulation interval,' CCITT SG XVII contribution COM XVII, no. D180, October 1983.

20. E. Biglieri, "High-level modulation and coding for nonlinear satellite channels," IEEE Transaction on Communications, vol. COM-32, no.5, pp616-676, May 1984.

21. A. J. Viterbi and J. K. Omura, Principles of Digital Communication and Coding, New York, McGraw-Hill, 1979.

22. M. K. Simon, et.al, Spread-Spectrum Communications Handbook, New York, McGraw-Hill, 1995.

23. G. D. Forney, Jr., "The Viterbi Algorithm," Processdings of the IEEE, vol. 01, pp. 268-278, March 1973.

24. S. Kato, M. Morikuia and S. Kuboto, "Inplementation of Coded Modems," IEEE Communications Magazine, December 1991, pp. 88-98.

# Vita

Xu Lin was born in Guangzhou, P.R. China. She obtained her B.S. in Electrical Engineering from South China University of Technology, Guangzhou, P.R. China. She had been working for Bruel & Kjaer as an application engineer for several years before she came to the United State in 1991.

She started her graduate study in Electrical Engineering at Virginia Tech in fall 1994. Her major is in Communications. She joined MPRG in June 1995, working with Dr. Woerner. She is expected to work at Hughes Network System, Germentown, Maryland, after she graduate in May, 1996.

She is a member of Eta Kappa Nu (HKN), Electrical Engineering Honor Society.