Chapter 3
Quasi-Constant Envelope Modulations

In Chap. 2, we restricted our consideration to bandwidth-efficient modulations that were *strictly* constant envelope, thus rendering themselves maximally power efficient when transmitted over a nonlinear channel operating in saturation. As a compromise between the two extremes of constant and nonconstant envelope, we turn our attention to modulations that deviate slightly rom the former but make up for the attendent small loss in power efficiency by offering a more significant improvement in bandwidth efficiency. The most promising modulation in this category is Feher-patented quadriphase-shift-keying (FQPSK) [1], whose generic form finds its roots in cross-correlated PSK (XPSK), introduced in 1983 [2], and which has recently been given a more insightful interpretation [3], thereby allowing further enhancements [3,4]. Since the basic form of FQPSK and its predecessor, XPSK, are well documented in several of Feher's textbooks and papers [5-8], our focus here will be on the recent advancements [3,4] that allow additional improvements in power and bandwidth efficiencies. Nevertheless, we shall present a brief review of FQPSK in its originally conceived form, since it provides insight into the new interpretation and enhancements that followed. Before proceeding with the technical details, we present a brief historical perspective as well as the current state of the art regarding the practical application of FQPSK in government- and commercially developed hardware.

FQPSK was invented by Kamilo Feher (Digcom, Inc. and the University of California, Davis). It is covered by a number of U.S. and Canadian patents [1], and is exclusively licensed by Digcom, Inc. It has been adopted by the U.S. Department of Defense Joint Services Advanced Range Telemetry (ARTM) program as their Tier 1 modulation for missile, aircraft, and range applications to replace existing pulse code modulation/frequency modulation (PCM/FM) systems. FQPSK modems operating at a data rate of 20 Mb/s are currently

available as an off-the-shelf product from several commercial vendors. The suitability of FQPSK for high-speed application has been demonstrated under a joint program between Goddard Space Flight Center (GSFC) and the Jet Propulsion Laboratory (JPL), with the development of a 300-Mb/s modem based on the enhanced architecture suggested in Ref. 3. Actually, the receiver, which employs all-digital parallel processing, can operate at 600 Mb/s and although originally designed for BPSK/QPSK, can accommodate FQPSK simply by having one reprogram the detection filter coefficients. Furthermore, the GSFC development has demonstrated that the same synchronization (carrier, bit, etc.) and detection techniques used for QPSK can be used for FQPSK without hardware modification to achieve good (not necessarily optimal) performance. Finally, at the CCSDS/SFCG meeting held in October 1999, it was recommended "that either GMSK or FQPSK-B[1] be used for high data-rate transmissions whenever practicable, and, in any case, for operating at frequencies where the available bandwidth is limited."

As implied above, in its generic (unfiltered) form, FQPSK is conceptually the same as XPSK, introduced in 1983 by Kato and Feher [2].[2] This technique was in turn a modification of the previously introduced (by Feher et al. [10]) interference- and jitter-free QPSK (IJF-QPSK), with the express purpose of reducing the 3-dB envelope fluctuation characteristic of IJF-QPSK to 0 dB (or as close to zero as possible), thus making it appear constant envelope.[3] It is further noted that using a constant waveshape for the even pulse and a sinusoidal waveshape for the odd pulse, which was the case considered in [2], IJF-QPSK becomes identical to the staggered quadrature overlapped raised cosine (SQORC) scheme introduced by Austin and Chang [11]. (We shall demonstrate this shortly.) The means by which Kato and Feher achieved their 3-dB envelope reduction was the introduction of an intentional but controlled amount of cross-correlation between the I and Q channels. This cross-correlation operation was applied to the IJF-QPSK (SQORC) baseband signal prior to its modulation onto the I and Q carriers (see Fig. 3-1). Specifically, this operation constituted mapping in each *half*-symbol the 16 possible combinations of I and Q channel waveforms

---

[1] The acronym FQPSK-B refers to Butterworth-filtered FQPSK. The exact filter type and optimal value of bandwidth-symbol time product, BT_s, are proprietary.

[2] More recent versions of FQPSK, referred to as FQPSK-B [9], include proprietary designed filtering for additional spectrum containment. At the moment, such filtering is not germane to our discussions although it should be mentioned now that the enhancements to be discussed are also applicable to FQPSK-B and similarly provide improved performance.

[3] The reduction of the envelope fluctuation from 3 dB to 0 dB occurs only at the uniform sampling instants on the inphase (I) and quadrature (Q) channels. It is for this reason that XPSK is referred to as being pseudo- or quasi-constant envelope, i.e., its envelope has a small amount of fluctuation between the uniform sampling instants.

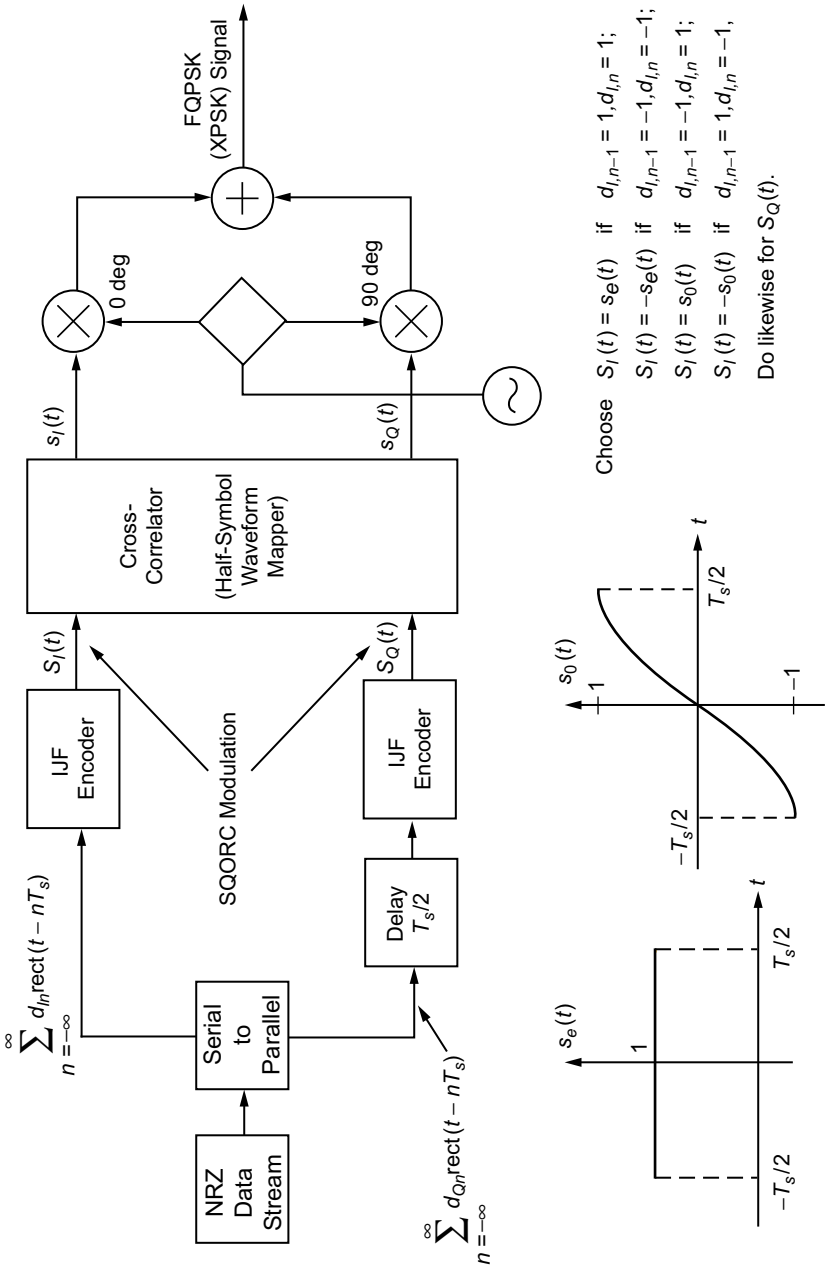

Fig. 3-1. Conceptual block diagram of FQPSK (XPSK). "IJF" (intersymbol interference/jitter-free encoder) is a waveform mapping function without any error-correcting capability.

present in the SQORC signal into a new[4] set of 16 waveform combinations chosen in such a way that the cross-correlator output is time continuous and has unit (normalized) envelope[5] at all I and Q uniform sampling instants. Because the cross-correlation mapping is based on a half-symbol characterization of the SQORC signal, there is no guarantee that the *slope* of the cross-correlator output waveform is continuous at the half-symbol transition points. In fact, it can be shown [3] that for a random-data input sequence, such a discontinuity in slope occurs one quarter of the time.

It is a well-known fact that the rate at which the sidelobes of a modulation's PSD roll off with frequency is related to the smoothness of the underlying waveforms that generate it. That is, the more derivatives of a waveform that are continuous, the faster its Fourier transform decays with frequency. Thus, since the first derivative of the FQPSK waveform is discontinuous (at half-symbol transition instants) on the average one quarter of the time, one can anticipate that an improvement in PSD rolloff could be had if the FQPSK cross-correlation mapping could be modified so that the first derivative is always continuous. By restructuring the cross-correlation mapping into a symbol-by-symbol representation, the slope discontinuity referred to above is placed in evidence and is particularly helpful in suggesting a means to eliminate it. This representation also has the advantage that it can be described directly in terms of the data transitions on the I and Q channels, and, thus, the combination of IJF encoder and cross-correlator can be replaced simply by a single modified cross-correlator. The replacement of the conventional FQPSK cross-correlator by this modified cross-correlator that eliminates the slope discontinuity leads to what is referred to as enhanced FQPSK [3]. Not only does enhanced FQPSK have a better PSD (in the sense of reduced out-of-band energy) than conventional FQPSK but from a modulation symmetry standpoint, it is a more logical choice.

A further and more important advantage of the reformulation as a symbol-by-symbol mapping is the ability to design a receiver for FQPSK or enhanced FQPSK that specifically exploits the correlation introduced into the modulation scheme to significantly improve power efficiency or, equivalently, error-probability performance. Such a receiver, which takes a form analogous to those used for trellis-coded modulations, will yield significant performance improvement over receivers that employ symbol-by-symbol detection, thus ignoring the inherent memory of the modulation.

---

[4] Of the 16 possible cross-correlator output combinations, only 12 of them are in fact new, i.e., for 4 of the input I and Q combinations, the cross-correlator outputs the identical combination.

[5] Actually, Kato and Feher allow (through the introduction of a transition parameter $k = 1 - A$ to be defined shortly) for a controlled amount of envelope fluctuation. For quasi-constant envelope, one should choose $A = 1/\sqrt{2}$.

3.1 Brief Review of IJF-QPSK and SQORC and Their Relation to FQPSK

The IJF-QPSK scheme (alternately called FQPSK-1) is based on defining waveforms, $s_o(t)$ and $s_e(t)$, which are respectively odd and even functions of time over the symbol interval $-T_s/2 \leq t \leq T_s/2$, and then using these and their negatives, $-s_o(t), -s_e(t)$, as a 4-ary signal set for transmission in accordance with the values of successive pairs of data symbols in each of the I and Q arms. Specifically, if d_{In} denotes the I channel data symbols in the interval $(n-(1/2))\,T_s \leq t \leq (n+(1/2))\,T_s$, then the transmitted waveform, $x_I\,(t)$, in this same interval would be determined as follows:

$$\left.\begin{array}{ll} x_I(t) = s_e\,(t-nT_s) \triangleq s_0\,(t-nT_s) & \text{if } d_{I,n-1}=1, d_{I,n}=1 \\ x_I\,(t) = -s_e\,(t-nT_s) \triangleq s_1\,(t-nT_s) & \text{if } d_{I,n-1}=-1, d_{I,n}=-1 \\ x_I\,(t) = s_o\,(t-nT_s) \triangleq s_2\,(t-nT_s) & \text{if } d_{I,n-1}=-1, d_{I,n}=1 \\ x_I\,(t) = -s_o\,(t-nT_s) \triangleq s_3\,(t-nT_s) & \text{if } d_{I,n-1}=1, d_{I,n}=-1 \end{array}\right\} \quad (3.1\text{-}1)$$

The Q channel waveform, $x_Q\,(t)$, would be generated by the same mapping as in (3.1-1), using instead the Q channel data symbols, $\{d_{Qn}\}$, and then delaying the resulting waveform by a half-symbol. If the odd and even waveforms, $s_o(t)$ and $s_e(t)$, are defined by

$$\left.\begin{array}{ll} s_e\,(t) = 1, & -\dfrac{T_s}{2} \leq t \leq \dfrac{T_s}{2} \\ s_o\,(t) = \sin\dfrac{\pi t}{T_s}, & -\dfrac{T_s}{2} \leq t \leq \dfrac{T_s}{2} \end{array}\right\} \quad (3.1\text{-}2)$$

then typical waveforms for the I and Q IJF encoder outputs are illustrated in Figs. 3-2(a) and 3-2(b).

An identical modulation to $x_I\,(t)$ (and likewise for $x_Q\,(t)$) generated from the combination of (3.1-1) and (3.1-2) can be obtained directly from the binary data sequence, $\{d_{In}\}$, itself, without the need for defining a 4-ary mapping based on the transition properties of the sequence. In particular, if we define the two-symbol-wide, raised-cosine pulse shape

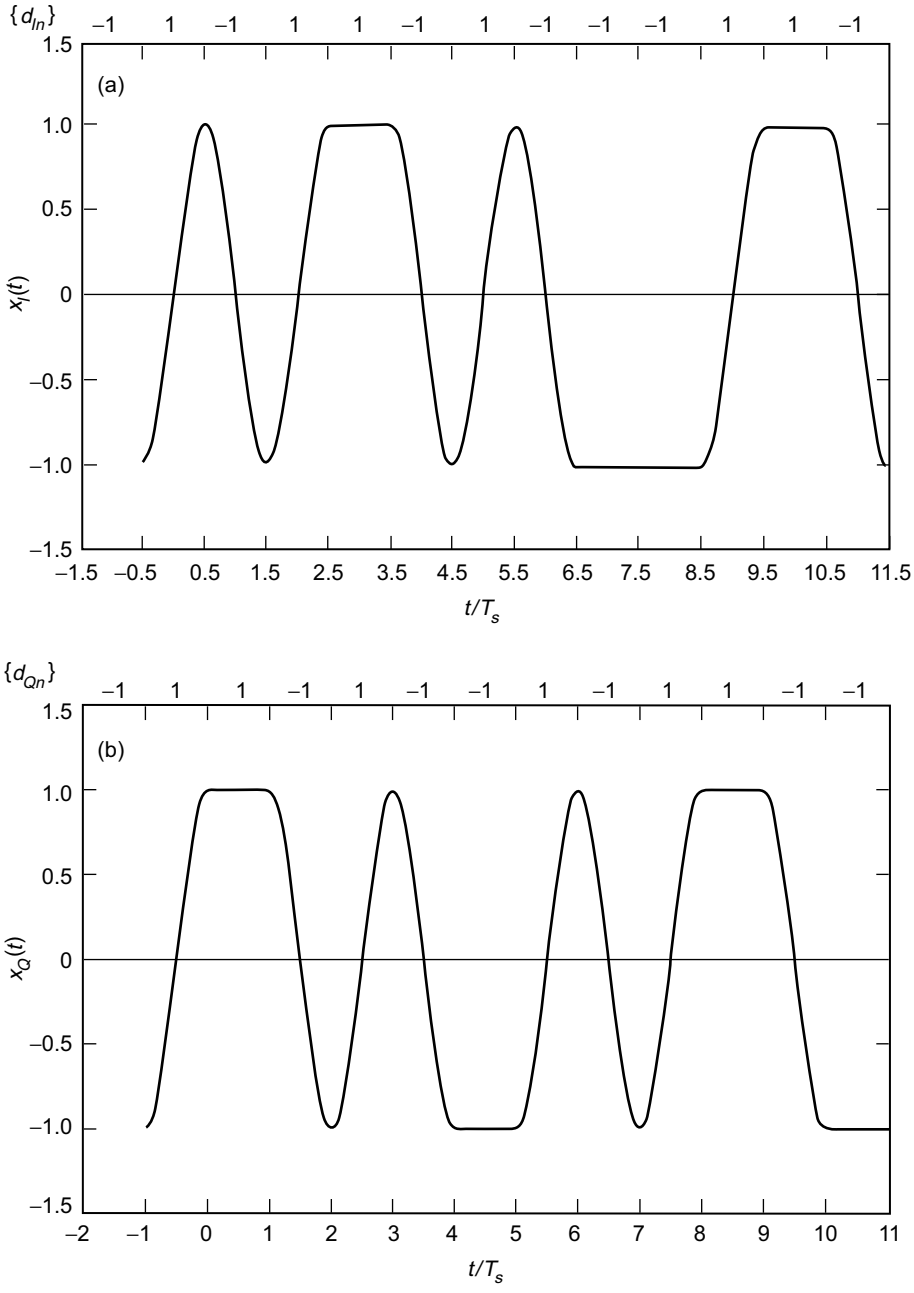

Fig. 3-2. IJF encoder output: (a) in-phase and (b) quadrature phase. Redrawn from [3].

$$p(t) = \sin^2\left(\frac{\pi\left(t + \dfrac{T_s}{2}\right)}{2T_s}\right), \quad -\frac{T_s}{2} \le t \le \frac{3T_s}{2} \tag{3.1-3}$$

then the I modulation

$$x_I(t) = \sum_{n=-\infty}^{\infty} d_{In}\, p(t - nT_s) \tag{3.1-4}$$

will be identical to that generated by the above IJF scheme, assuming the choice of odd and even waveforms as in (3.1-2). Similarly,

$$x_Q(t) = \sum_{n=-\infty}^{\infty} d_{Qn}\, p\left(t - \left(n + \frac{1}{2}\right)T_s\right) \tag{3.1-5}$$

would also be identical to that generated by the above IJF scheme. A quadrature modulation scheme formed from $x_I(t)$ of (3.1-4) and $x_Q(t)$ of (3.1-5) is precisely what Austin and Chang referred to as SQORC modulation [11], namely, independent I and Q staggered modulations with overlapping raised cosine pulses on each channel. The resulting carrier modulated waveform is described by

$$x(t) = x_I(t)\cos\omega_c t + x_Q(t)\sin\omega_c t \tag{3.1-6}$$

and is implementated as shown in Fig. 3-3. Figure 3-4 shows the equivalence of the transmitted SQORC baseband waveforms with the IJF-QPSK even and odd waveforms of (3.1-2) for a pair of consecutive data bits.

Although SQORC exhibits a 3-dB envelope fluctuation, it is nevertheless a highly bandwidth-efficient modulation. In fact, except for a normalization constant, its PSD is the *product* of the PSDs of OQPSK and MSK, i.e.,

$$S_{\text{SQORC}}(f) = \left(\frac{\sin\pi f T_s}{\pi f T_s}\right)^2 \frac{\cos^2 2\pi f T_s}{\left(1 - 16 f^2 T_s^2\right)^2} \tag{3.1-7}$$

which asymptotically decays as f^{-6}. It was with this in mind that Feher and Kato sought to tailor the transmitted waveform in such a fashion as to reduce the envelope fluctuation to near 0 dB, yet maintain the high bandwidth efficiency

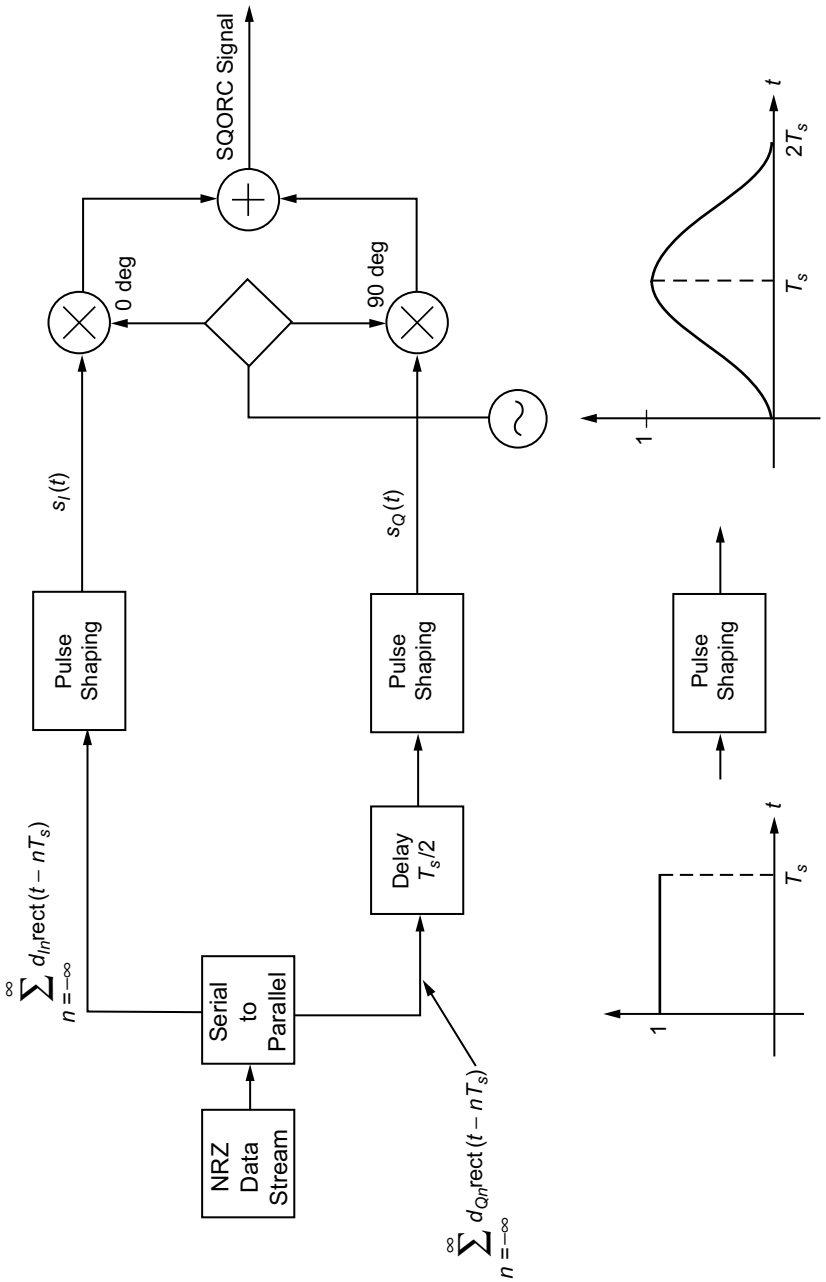

Fig. 3-3. Conceptual block diagram of an SQORC transmitter.

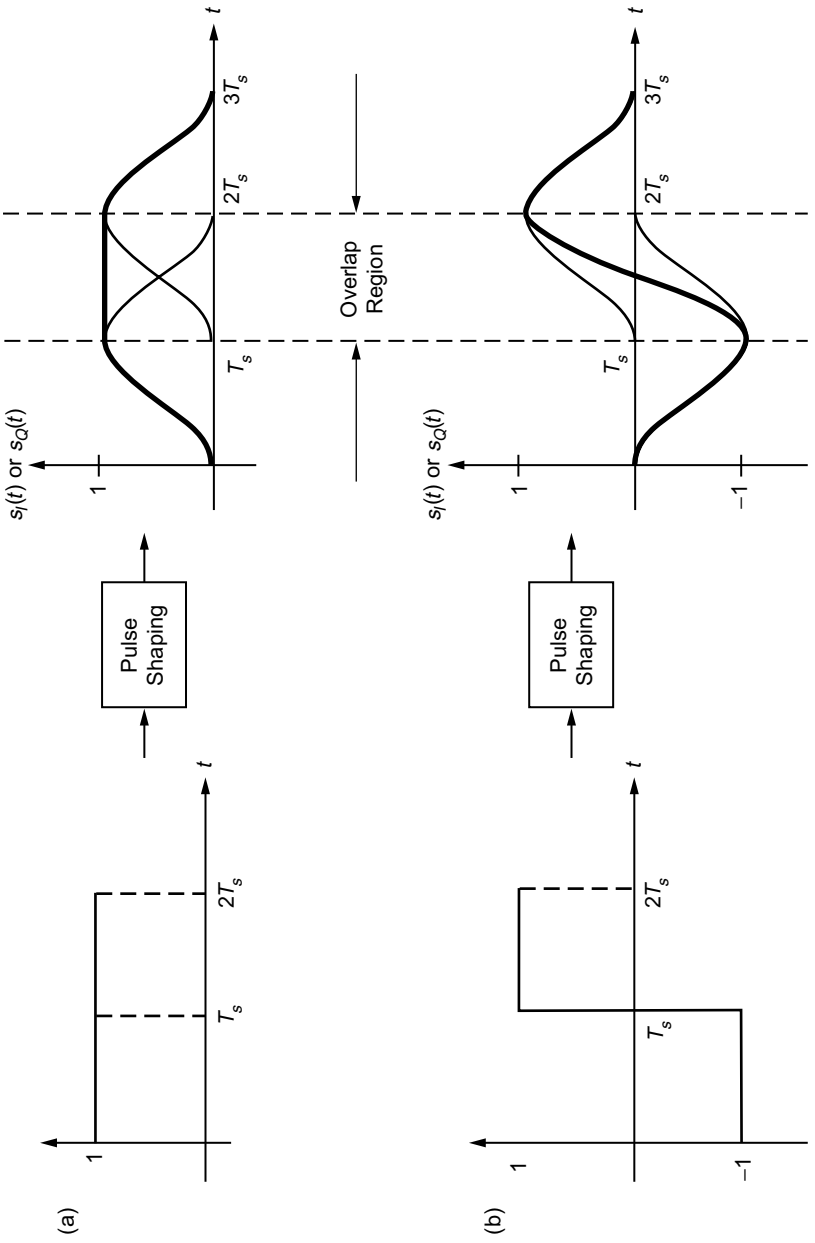

Fig. 3-4. Baseband SQORC waveforms: (a) consecutive symbols of like polarity and (b) consecutive symbols of alternating polarity.

inherent in SQORC. The specifics of how this is accomplished are wrapped up in the block labeled "cross-correlator" in Fig. 3-1 and are described below.

With reference to Fig. 3-1, in any given half-symbol, there are 16 possible combinations of the SQORC signal components, $S_I(t)$ and $S_Q(t)$, at the input of the cross-correlator. These combinations are illustrated in Fig. 3-5 and are composed of specific combinations of the signals $\pm 1, \pm \sin \pi t/T_s, \pm \cos \pi t/T_s$. For each of the I-Q component pairs, $S_I(t), S_Q(t)$, the cross-correlator generates a new I-Q component pair denoted by $s_I(t), s_Q(t)$, whose purpose is to reduce the envelope fluctuation of the resulting I and Q symbol steams. As such, the cross-correlator acts as a half-symbol waveform mapper. A mathematical description of the 16 possible cross-correlated signal combinations is given in Table 3-1.

Table 3-1. I and Q cross-correlated signal combinations.

$s_I(t)$ (or $s_Q(t)$)	$s_Q(t)$ (or $s_I(t)$)	Number of Combinations
$\pm \cos \pi t/T_s$	$\pm \sin \pi t/T_s$	4
$\pm A \cos \pi t/T_s$	$f_1(t)$ or $f_3(t)$	4
$\pm A \sin \pi t/T_s$	$f_2(t)$ or $f_4(t)$	4
$\pm A$	$\pm A$	4

The transition functions $f_i(t), i = 1, 2, 3, 4$ referred to in Table 3-1 are defined in the interval $0 \le t \le T_s/2$ by [2]

$$\left.\begin{aligned} f_1(t) &= 1 - (1-A)\cos^2 \frac{\pi t}{T_s} \\ f_2(t) &= 1 - (1-A)\sin^2 \frac{\pi t}{T_s} \\ f_3(t) &= -1 + (1-A)\cos^2 \frac{\pi t}{T_s} \\ f_4(t) &= -1 + (1-A)\sin^2 \frac{\pi t}{T_s} \end{aligned}\right\} \qquad (3.1\text{-}8)$$

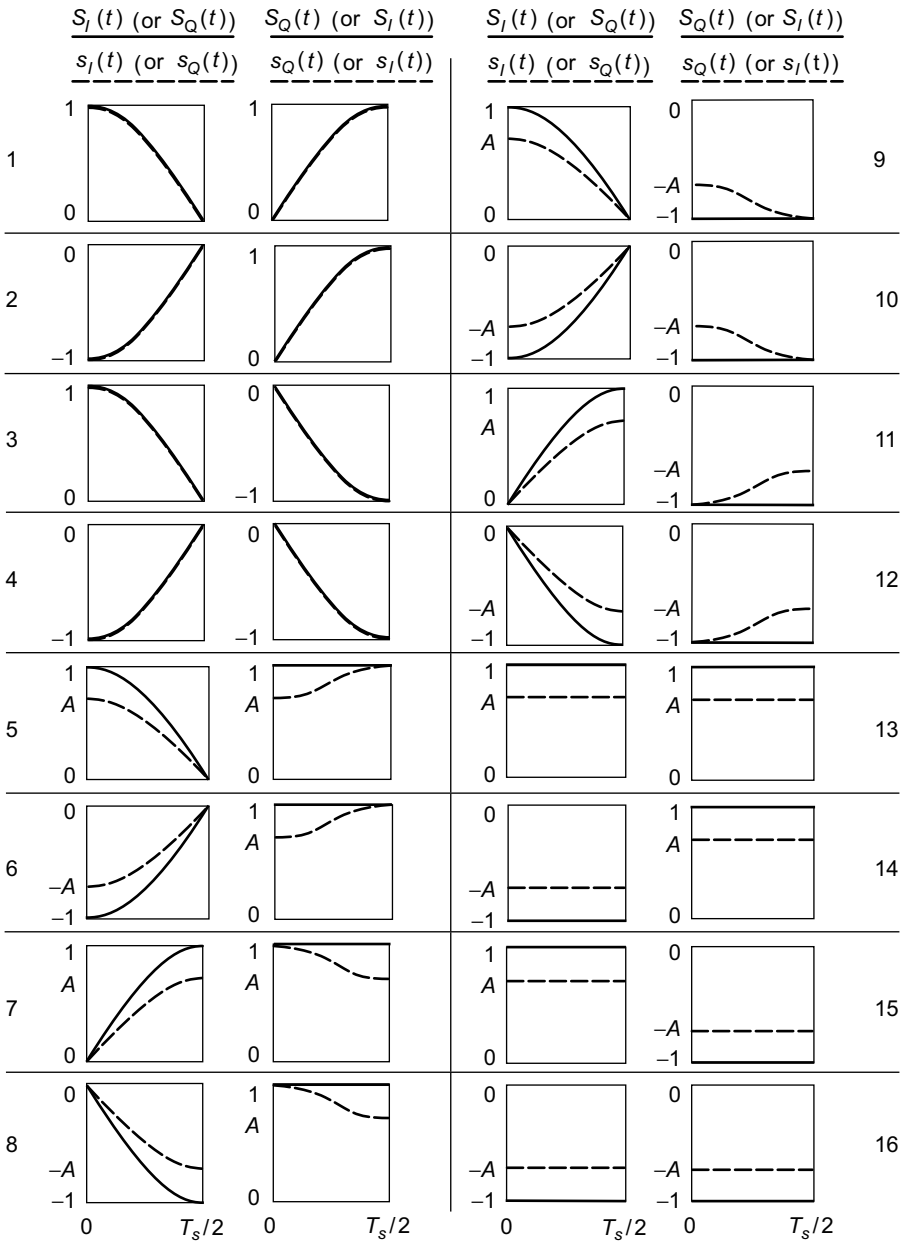

Fig. 3-5. FQPSK half-symbol waveform mapper ($A = 1 / \sqrt{2}$ for "constant" envelope).

where A is a transition parameter that can take on values in the interval $1/\sqrt{2}$ to 1 and is used to trade off between the amount of envelope fluctuation and bandwidth efficiency. The 16 mapped symbol pairs, $s_I(t), s_Q(t)$, that appear at the output of the cross-correlator, corresponding to the combinations in Table 3-1 are superimposed on the corresponding 16 possible input pairs, $S_I(t), S_Q(t)$, shown in Fig. 3-5. Note that the input pairs $S_I(t), S_Q(t)$ for combinations 1, 2, 3, and 4 are already constant envelope, and, thus, remapping of these signals is not necessary. For the same I and Q data sequences that generated the SQORC signal components in Figs. 3-2(a) and 3-2(b) and $A = 1/\sqrt{2}$, Figs. 3-6(a) and 3-6(b) illustrate the corresponding I and Q cross-correlator outputs. One can observe from these figures that at the uniform sample points on the I and Q channels, i.e., $t = nT_s$ and $t = (n + (1/2))T_s$ (n integer), the transmitted baseband signal is now *precisely* constant envelope. At other than the uniform sample points, the maximum fluctuation in the baseband signal envelope has been shown to be equal to 0.18 dB [2]—a small price for the significant bandwidth efficiency that has been afforded by this modulation, as will now be demonstrated.

Illustrated in Fig. 3-7 is the PSD of unfiltered FQPSK (as described above) along with that corresponding to other modulations previously discussed in this monograph. Figure 3-8 illustrates the corresponding plots of out-of-band energy versus normalized bandwidth, $BT_b = B/R_b$, as computed from

$$P_{ob} = 1 - \frac{\int_{-B}^{B} S_m(f)\,df}{\int_{-\infty}^{\infty} S_m(f)\,df} \tag{3.1-9}$$

When compared with OQPSK and MSK, which are both constant envelope, unfiltered FQPSK, which is virtually constant envelope, clearly provides an improvement in spectral efficiency. When compared with constant envelope GMSK, however, filtering must be applied to FQPSK in order to make it comparable in spectral efficiency. The PSD of FQPSK-B is superimposed on Figs. 3-7 and 3-8 and clearly shows a spectral advantage when compared, for example, with $BT_b = 0.5$ GMSK.

3.2 A Symbol-by-Symbol Cross-Correlator Mapping for FQPSK

In Ref. 3, the original characterization of FQPSK in terms of a cross-correlation operation performed on the pair of IJF encoder outputs every half-symbol interval was reformulated into a mapping performed directly on the input I and Q data sequences every full symbol interval. To do this, 16 waveforms, $s_i(t); i = 0, 1, 2, \cdots, 15$, were defined over the interval $-T_s/2 \le t \le T_s/2$, which

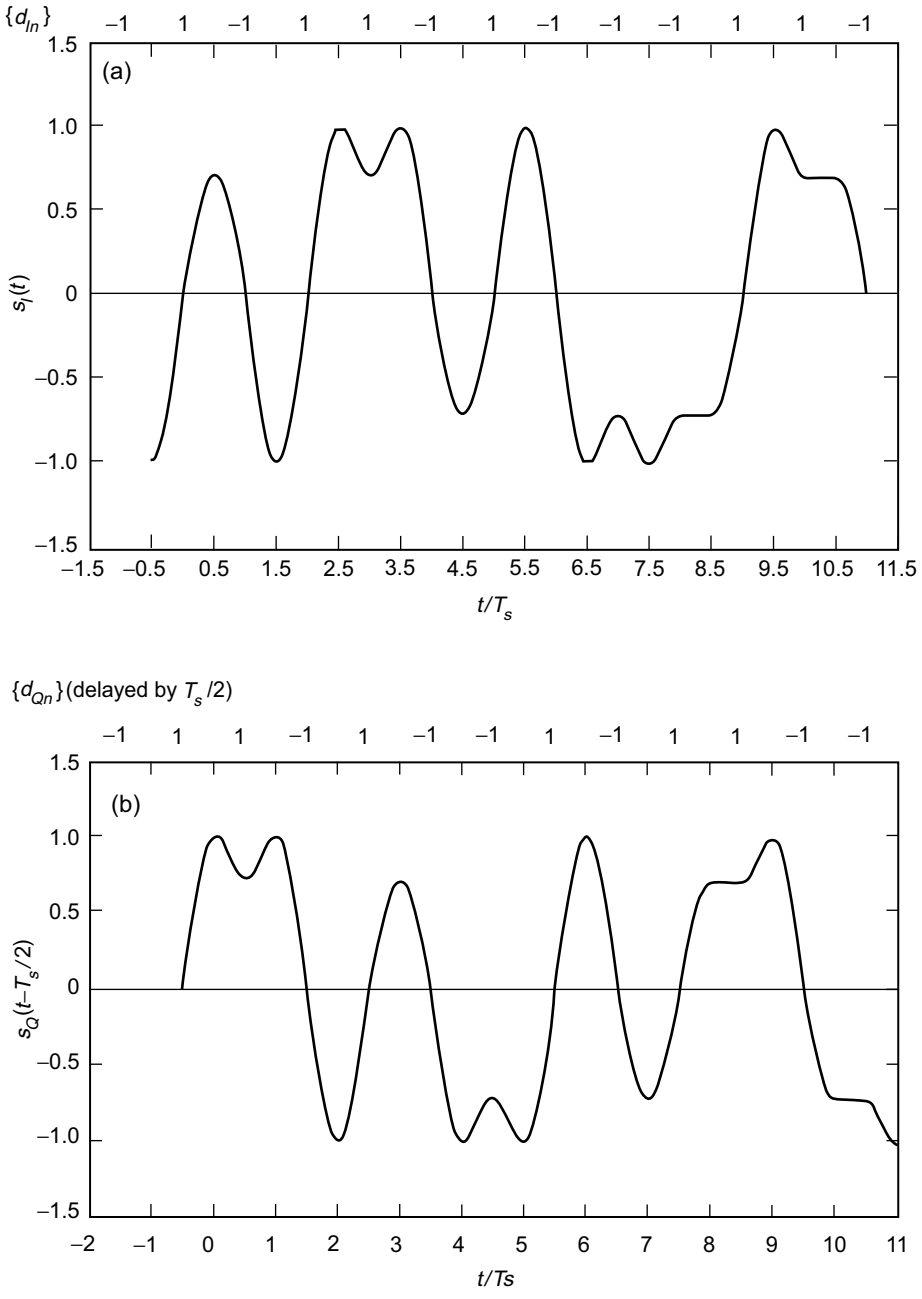

Fig. 3-6. FQPSK (XPSK) output: (a) in-phase and (b) quadrature-phase. Redrawn from [3].

collectively formed a transmitted signaling set for the I and Q channels. The particular I and Q waveforms chosen for any particular T_s-s signaling interval on each channel depended on the most recent data transition on that channel as well as the two most recent successive transitions on the other channel and such identified FQPSK as a modulation with inherent memory. The specific details are presented in Ref. 3 and are summarized as follows: Define (see Fig. 3-9)

$$
\left.
\begin{aligned}
s_0(t) &= A, && -\frac{T_s}{2} \le t \le \frac{T_s}{2}, && s_8(t) = -s_0(t) \\
s_1(t) &= \begin{cases} A, & -\frac{T_s}{2} \le t \le 0, \\ 1-(1-A)\cos^2\frac{\pi t}{T_s}, & 0 \le t \le \frac{T_s}{2}, \end{cases} && && s_9(t) = -s_1(t) \\
s_2(t) &= \begin{cases} 1-(1-A)\cos^2\frac{\pi t}{T_s}, & -\frac{T_s}{2} \le t \le 0, \\ A, & 0 \le t \le \frac{T_s}{2}, \end{cases} && && s_{10}(t) = -s_2(t) \\
s_3(t) &= 1-(1-A)\cos^2\frac{\pi t}{T_s}, && -\frac{T_s}{2} \le t \le \frac{T_s}{2}, && s_{11}(t) = -s_3(t) \\
s_4(t) &= A\sin\frac{\pi t}{T_s}, && -\frac{T_s}{2} \le t \le \frac{T_s}{2}, && s_{12}(t) = -s_4(t) \\
s_5(t) &= \begin{cases} A\sin\frac{\pi t}{T_s}, & -\frac{T_s}{2} \le t \le 0, \\ \sin\frac{\pi t}{T_s}, & 0 \le t \le \frac{T_s}{2}, \end{cases} && && s_{13}(t) = -s_5(t) \\
s_6(t) &= \begin{cases} \sin\frac{\pi t}{T_s}, & -\frac{T_s}{2} \le t \le 0, \\ A\sin\frac{\pi t}{T_s}, & 0 \le t \le \frac{T_s}{2}, \end{cases} && && s_{14}(t) = -s_6(t) \\
s_7(t) &= \sin\frac{\pi t}{T_s}, && -\frac{T_s}{2} \le t \le \frac{T_s}{2}, && s_{15}(t) = -s_7(t)
\end{aligned}
\right\} \qquad (3.2\text{-}1)
$$

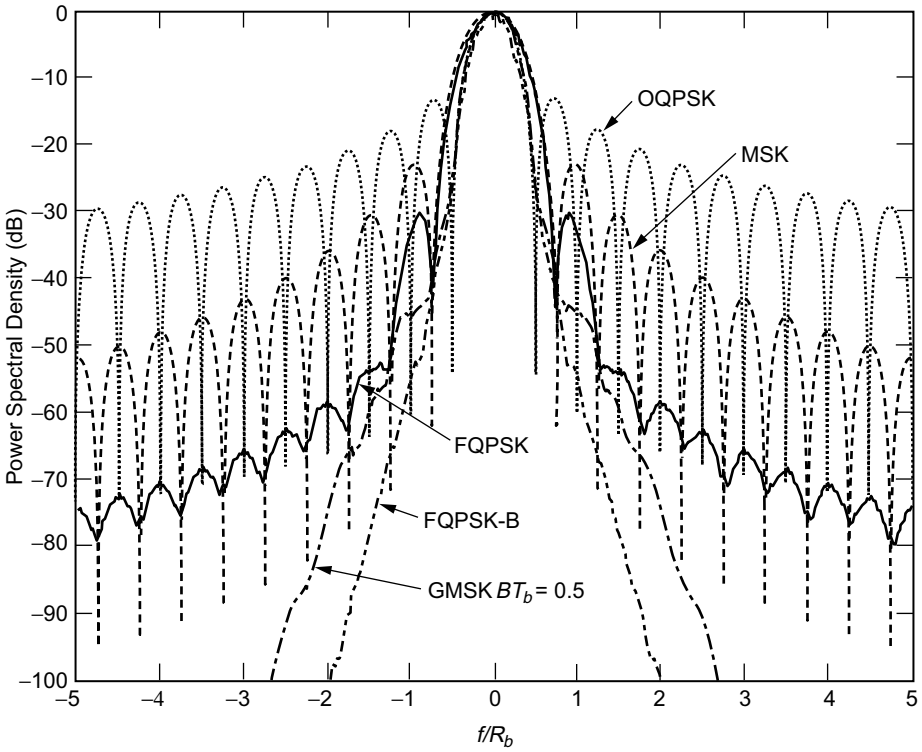

Fig. 3-7. Power spectral density of various modulations.

Note that for any value of A other then unity, e.g., $A = 1/\sqrt{2}$, $s_5(t)$ and $s_6(t)$ as well as their negatives, $s_{13}(t)$ and $s_{14}(t)$, will have a discontinuous slope at their midpoints (i.e., at $t = 0$) whereas the remaining 12 waveforms all have a continuous slope throughout their defining interval. Also, all 16 waveforms have zero slope at their endpoints and thus, concatenation of any pair of these will not result in a slope discontinuity. We will return to the issue of the discontinuous midpoint slope shortly.

The mapping function that assigns the particular baseband I and Q channel waveforms transmitted in the nth signaling interval chosen from the set in (3.2-1) is specified in terms of the transition properties of the I and Q data symbol sequences. For example, if $d_{I,n-1} = 1, d_{I,n} = 1$ (i.e., no transition on the I sequence and both data bits positive), then the transmitted I-channel signal, $y_I(t) = s_I(t)$, in the nth signaling interval $(n - (1/2))T_s \leq t \leq (n + (1/2))T_s$ is chosen as follows.

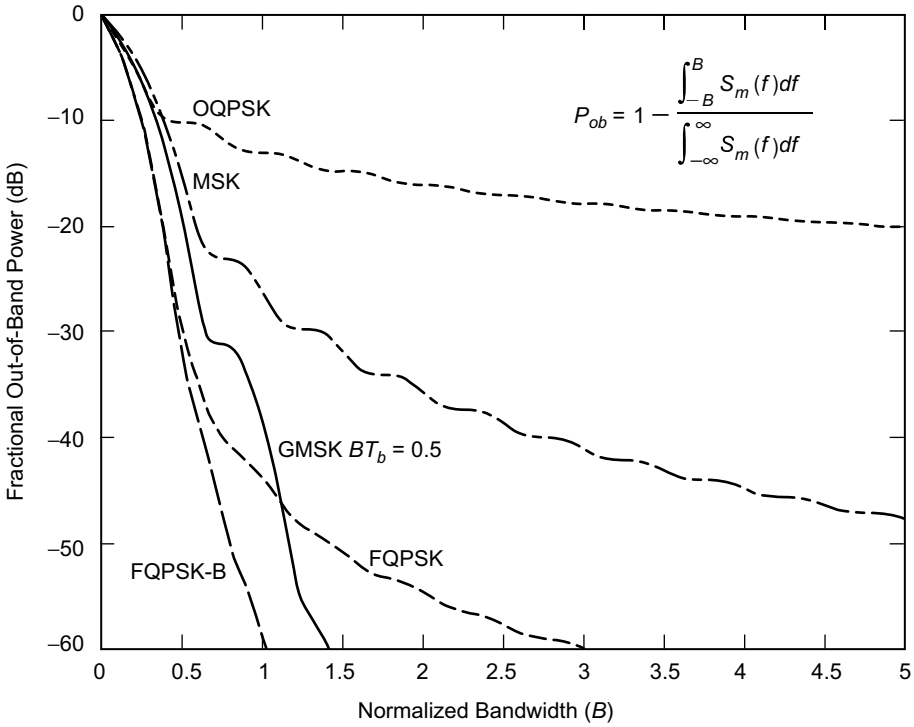

Fig. 3-8. Out-of-band power of various modulations.

A. $y_I(t) = s_0(t - nT_s)$ if $d_{Q,n-2}, d_{Q,n-1}$ results in no transition and $d_{Q,n-1}, d_{Q,n}$ results in no transition.

B. $y_I(t) = s_1(t - nT_s)$ if $d_{Q,n-2}, d_{Q,n-1}$ results in no transition and $d_{Q,n-1}, d_{Q,n}$ results in a transition (positive or negative).

C. $y_I(t) = s_2(t - nT_s)$ if $d_{Q,n-2}, d_{Q,n-1}$ results in a transition (positive or negative) and $d_{Q,n-1}, d_{Q,n}$ results in no transition.

D. $y_I(t) = s_3(t - nT_s)$ if $d_{Q,n-2}, d_{Q,n-1}$ results in a transition (positive or negative) and $d_{Q,n-1}, d_{Q,n}$ results in a transition (positive or negative).

Without belaboring the details, the assignments for the remaining three combinations of $d_{I,n-1}$ and $d_{I,n}$ follow similarly. Finally, making use of the signal properties in (3.2-1), the mapping conditions for the I-channel baseband output can be summarized in a concise form described by Table 3-2(a):

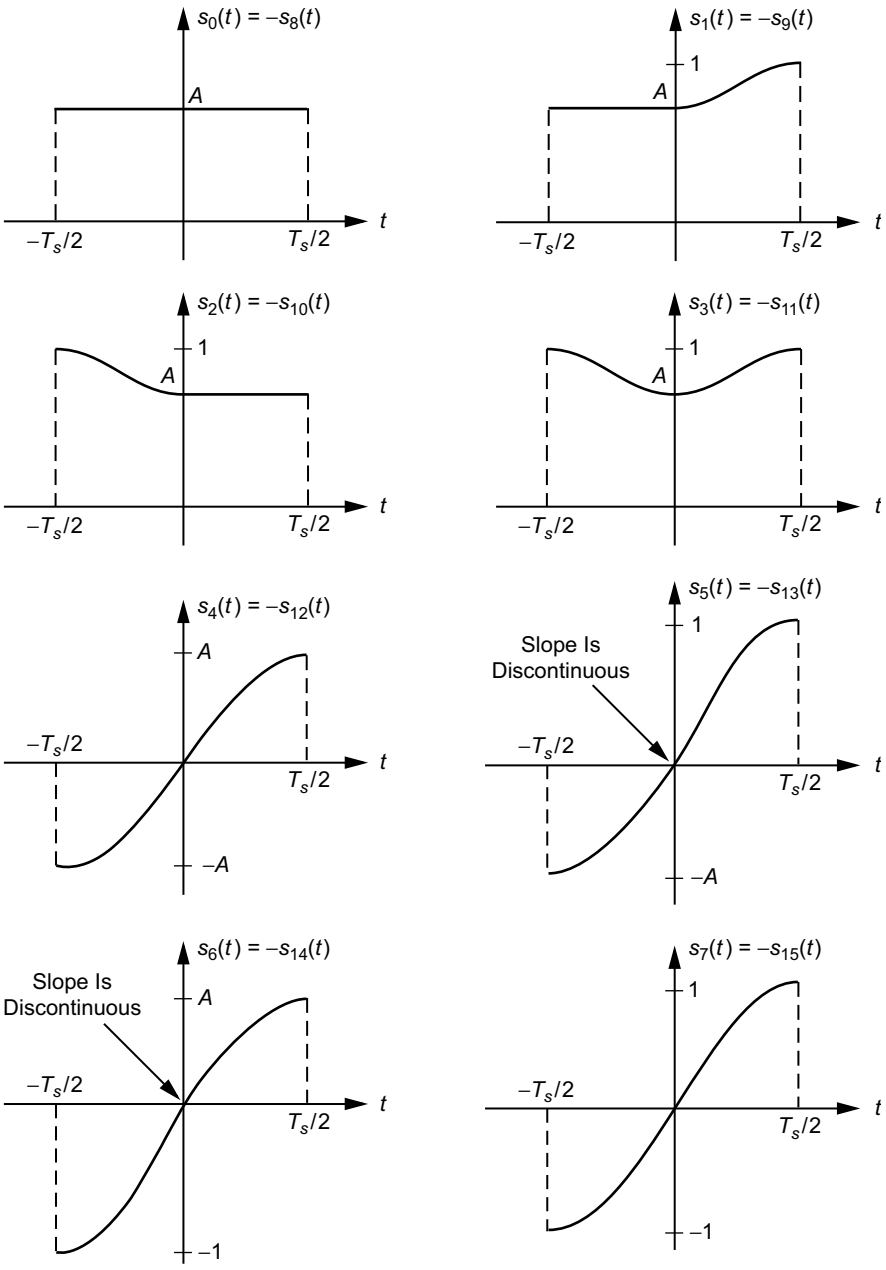

Fig. 3-9. FQPSK full-symbol waveforms
($A = 1 / \sqrt{2}$ for "constant" envelope). Redrawn from [3].

Table 3-2(a). Mapping for inphase (I)-channel baseband signal, $y_I(t)$, in the interval $(n - [1/2])\,T_s \le t \le (n + [1/2])\,T_s$.

$\left\lvert\frac{d_{In} - d_{I,n-1}}{2}\right\rvert$	$\left\lvert\frac{d_{Q,n-1} - d_{Q,n-2}}{2}\right\rvert$	$\left\lvert\frac{d_{Qn} - d_{Q,n-1}}{2}\right\rvert$	$s_I(t)$
0	0	0	$d_{In}s_0\,(t - nT_s)$
0	0	1	$d_{In}s_1\,(t - nT_s)$
0	1	0	$d_{In}s_2\,(t - nT_s)$
0	1	1	$d_{In}s_3\,(t - nT_s)$
1	0	0	$d_{In}s_4\,(t - nT_s)$
1	0	1	$d_{In}s_5\,(t - nT_s)$
1	1	0	$d_{In}s_6\,(t - nT_s)$
1	1	1	$d_{In}s_7\,(t - nT_s)$

A similar construction for the baseband Q-channel transmitted waveform, $y_Q\,(t) = s_Q\,(t - T_s/2)$, in the nth signaling interval, $nT_s \le t \le (n+1)\,T_s$, in terms of the transition properties of the I and Q data symbol sequences, $\{d_{In}\}$ and $\{d_{Qn}\}$, leads to Table 3-2(b):

Table 3-2(b). Mapping for quadrature (Q)-channel baseband signal, $y_Q(t)$, in the interval $nT_s \le t \le (n + 1)\,T_s$.

$\left\lvert\frac{d_{Qn} - d_{Q,n-1}}{2}\right\rvert$	$\left\lvert\frac{d_{I,n} - d_{I,n-1}}{2}\right\rvert$	$\left\lvert\frac{d_{I,n+1} - d_{I,n}}{2}\right\rvert$	$s_Q(t)$
0	0	0	$d_{Qn}s_0\,(t - nT_s)$
0	0	1	$d_{Qn}s_1\,(t - nT_s)$
0	1	0	$d_{Qn}s_2\,(t - nT_s)$
0	1	1	$d_{Qn}s_3\,(t - nT_s)$
1	0	0	$d_{Qn}s_4\,(t - nT_s)$
1	0	1	$d_{Qn}s_5\,(t - nT_s)$
1	1	0	$d_{Qn}s_6\,(t - nT_s)$
1	1	1	$d_{Qn}s_7\,(t - nT_s)$

Note from Tables 3-2(a) and 3-2(b) that subscript i of the transmitted signal $s_i\,(t - nT_s)$ or $s_i\,(t - (n + (1/2))\,T_s)$, as appropriate, is the binary-coded decimal (BCD) equivalent of the three transitions.

Applying the mappings in Tables 3-2(a) and 3-2(b) to the I and Q data sequences of Figs. 3-2(a) and 3-2(b) produces the identical I and Q baseband transmitted signals to those that would be produced by passing the I and Q IJF encoder outputs of this figure through the cross-correlator (half-symbol mapping) of the FQPSK (XPSK) scheme as described in Ref. 2 and illustrated in Figs. 3-6(a) and 3-6(b). Thus, we conclude that *for arbitrary I and Q input sequences, FQPSK can alternatively be generated by the symbol-by-symbol mappings of Tables 3-2(a) and 3-2(b) as applied to these sequences.*

3.3 Enhanced FQPSK

We now return to the issue of the midpoint slope discontinuity associated with four of the waveforms in Fig. 3-9. As discussed in Sec. 3.2, the symbol-by-symbol mapping representation of FQPSK identifies the fact that $s_5(t), s_6(t), s_{13}(t)$ and $s_{14}(t)$ have a slope discontinuity at their midpoint. Consequently, for random I and Q data symbol sequences, on the average, the transmitted FQPSK waveform will likewise have a slope discontinuity at one quarter of the uniform sampling time instants. To prevent this from occurring, we now redefine these four transmitted signals in a manner analogous to $s_1(t), s_2(t), s_9(t), s_{10}(t)$, namely,

$$\left.\begin{aligned} s_5(t) &= \begin{cases} \sin\dfrac{\pi t}{T_s} + (1-A)\sin^2\dfrac{\pi t}{T_s}, & -\dfrac{T_s}{2} \le t \le 0 \\ \sin\dfrac{\pi t}{T_s}, & 0 \le t \le \dfrac{T_s}{2} \end{cases} \qquad s_{13}(t) = -s_5(t) \\ s_6(t) &= \begin{cases} \sin\dfrac{\pi t}{T_s}, & -\dfrac{T_s}{2} \le t \le 0 \\ \sin\dfrac{\pi t}{T_s} - (1-A)\sin^2\dfrac{\pi t}{T_s}, & 0 \le t \le \dfrac{T_s}{2} \end{cases} \qquad s_{14}(t) = -s_6(t) \end{aligned}\right\} \quad (3.3\text{-}1)$$

Note that the signals $s_5(t), s_6(t), s_{13}(t), s_{14}(t)$ as defined in (3.3-1) do not have a slope discontinuity at their midpoint nor, for that matter, anywhere else in the defining interval. Also, the zero slope at their endpoints has been preserved. Therefore, using (3.3-1) in place of the corresponding signals of (3.2-1) will result in a modified FQPSK signal that has no slope discontinuity anywhere in time regardless of the value of A. Figure 3-10 illustrates a comparison of the

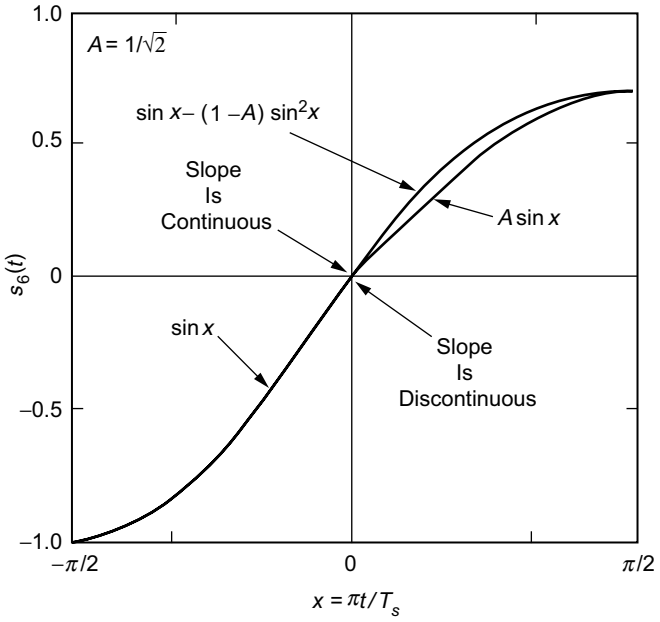

Fig. 3-10. Original and enhanced FQPSK pulse shapes. Redrawn from [3].

signal $s_6(t)$ of (3.3-1) with that of (3.2-1) for a value of $A = 1/\sqrt{2}$. Figure 3-11 illustrates the power spectral density of conventional FQPSK and its enhancement obtained by using the waveforms of (3.3-1) as replacements for those in (3.2-1). The significant improvement in spectral rolloff rate is clear from a comparison of the two.

As it currently stands, the signal set chosen for enhanced FQPSK has a symmetry property for $s_0(t), s_1(t), s_2(t), s_3(t)$ that is not present for $s_4(t), s_5(t), s_6(t), s_7(t)$. In particular, $s_1(t)$ and $s_2(t)$ are each composed of one-half of $s_0(t)$ and one-half of $s_3(t)$, i.e., the portion of $s_1(t)$ from $t = -T_s/2$ to $t = 0$ is the same as that of $s_0(t)$, whereas the portion of $s_1(t)$ from $t = 0$ to $t = T_s/2$ is the same as that of $s_3(t)$ and vice versa for $s_2(t)$. To achieve the same symmetry property for $s_4(t) - s_7(t)$, one would have to reassign $s_4(t)$ as

$$s_4(t) = \begin{cases} \sin\dfrac{\pi t}{T_s} + (1-A)\sin^2\dfrac{\pi t}{T_s}, & -\dfrac{T_s}{2} \le t \le 0 \\ & \\ \sin\dfrac{\pi t}{T_s} - (1-A)\sin^2\dfrac{\pi t}{T_s}, & 0 \le t \le \dfrac{T_s}{2} \end{cases} \qquad s_{12}(t) = -s_4(t) \qquad (3.3\text{-}2)$$

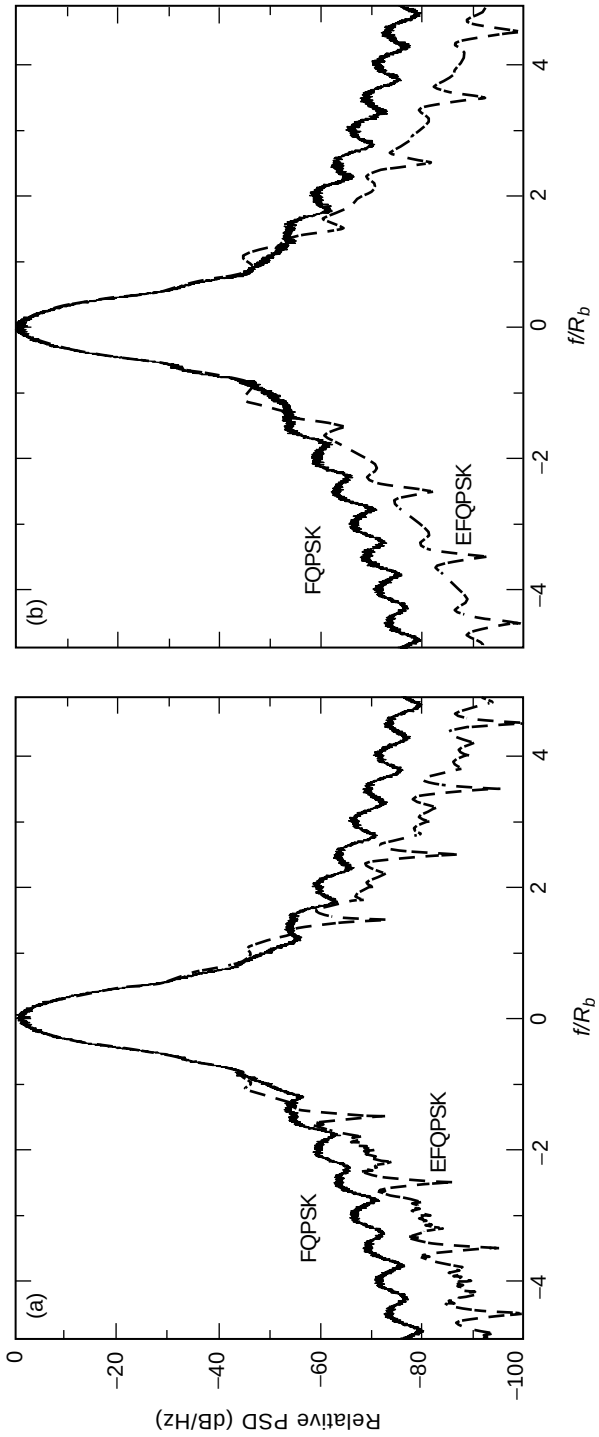

Fig. 3-11. Power spectrum of conventional and enhanced FQPSK: (a) without high-power amplifier and (b) with high-power amplifier. Redrawn from [3].

This minor change, which produces a complete symmetry in the waveform set, has an advantage from the standpoint of hardware implementation and produces a negligible change in spectral properties of the transmitted waveform. Nevertheless, for the remainder of the discussion, we shall ignore this minor change and assume the version of enhanced FQPSK first introduced in this section.

3.4 Interpretation of FQPSK as a Trellis-Coded Modulation

The I and Q mappings given in Tables 3-2a and 3-2b can alternatively be described in terms of the (0,1) representation of the I and Q data

$$\left.\begin{aligned} D_{In} &\triangleq \frac{1-d_{In}}{2} \\ D_{Qn} &\triangleq \frac{1-d_{Qn}}{2} \end{aligned}\right\} \tag{3.4-1}$$

which both range over the set (0,1). Then, defining the BCD representation of the indices i and j by

$$\left.\begin{aligned} i &= I_3 \times 2^3 + I_2 \times 2^2 + I_1 \times 2^1 + I_0 \times 2^0 \\ j &= Q_3 \times 2^3 + Q_2 \times 2^2 + Q_1 \times 2^1 + Q_0 \times 2^0 \end{aligned}\right\} \tag{3.4-2}$$

with

$$\left.\begin{aligned} I_0 &= D_{Qn} \oplus D_{Q,n-1}, & Q_0 &= D_{I,n+1} \oplus D_{In} \\ I_1 &= D_{Q,n-1} \oplus D_{Q,n-2}, & Q_1 &= D_{In} \oplus D_{I,n-1} = I_2 \\ I_2 &= D_{In} \oplus D_{I,n-1}, & Q_2 &= D_{Qn} \oplus D_{Q,n-1} = I_0 \\ I_3 &= D_{In}, & Q_3 &= D_{Qn} \end{aligned}\right\} \tag{3.4-3}$$

we have $y_I(t) = s_i(t - nT_s)$ and $y_Q(t) = s_j(t - (n+1/2))T_s)$. That is, in each symbol interval $((n-(1/2))T_s \le t \le (n+(1/2))T_s$ for $y_I(t)$ and $nT_s \le t \le (n+1)T_s$ for $y_Q(t)$), the I and Q channel baseband signals are each chosen from a set of 16 signals, $s_i(t), i = 0, 1, \cdots, 15$, in accordance with the 4-bit

BCD representations of their indices defined by (3.4-2) together with (3.4-3). A graphical illustration of the implementation of this mapping is given in Fig. 3-12.

Another interpretation of the mapping in Fig. 3-12 is as a 16-state trellis code with 2 binary (0,1) inputs $D_{I,n+1}, D_{Qn}$ and 2 waveform outputs, $s_i(t), s_j(t)$, where the state is defined by the 4-bit sequence, $D_{In}, D_{I,n-1}, D_{Q,n-1}, D_{Q,n-2}$. The trellis is illustrated in Fig. 3-13, and the transition mapping is given in Table 3-3. In this table, the entries in the column labeled "input" correspond to the values of the two input bits, $D_{I,n+1}, D_{Qn}$, that result in the transition, while the entries in the column labeled "output" correspond to the subscripts, i and j, of the pair of symbol waveforms, $s_i(t), s_j(t)$, that are outputted.

3.5 Optimum Detection

In designing receivers for FQPSK, the approach taken in the past has been to disregard the inherent memory of the transmitted modulation (actually the

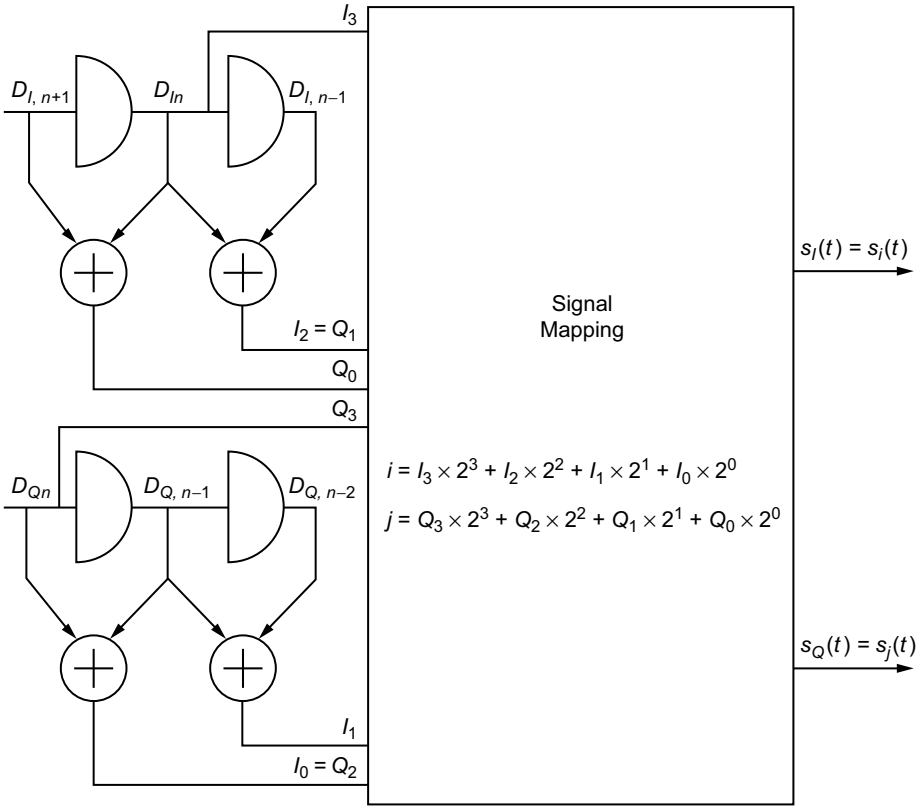

Fig. 3-12. Alternative implementation of FQPSK baseband signals. Redrawn from [3].

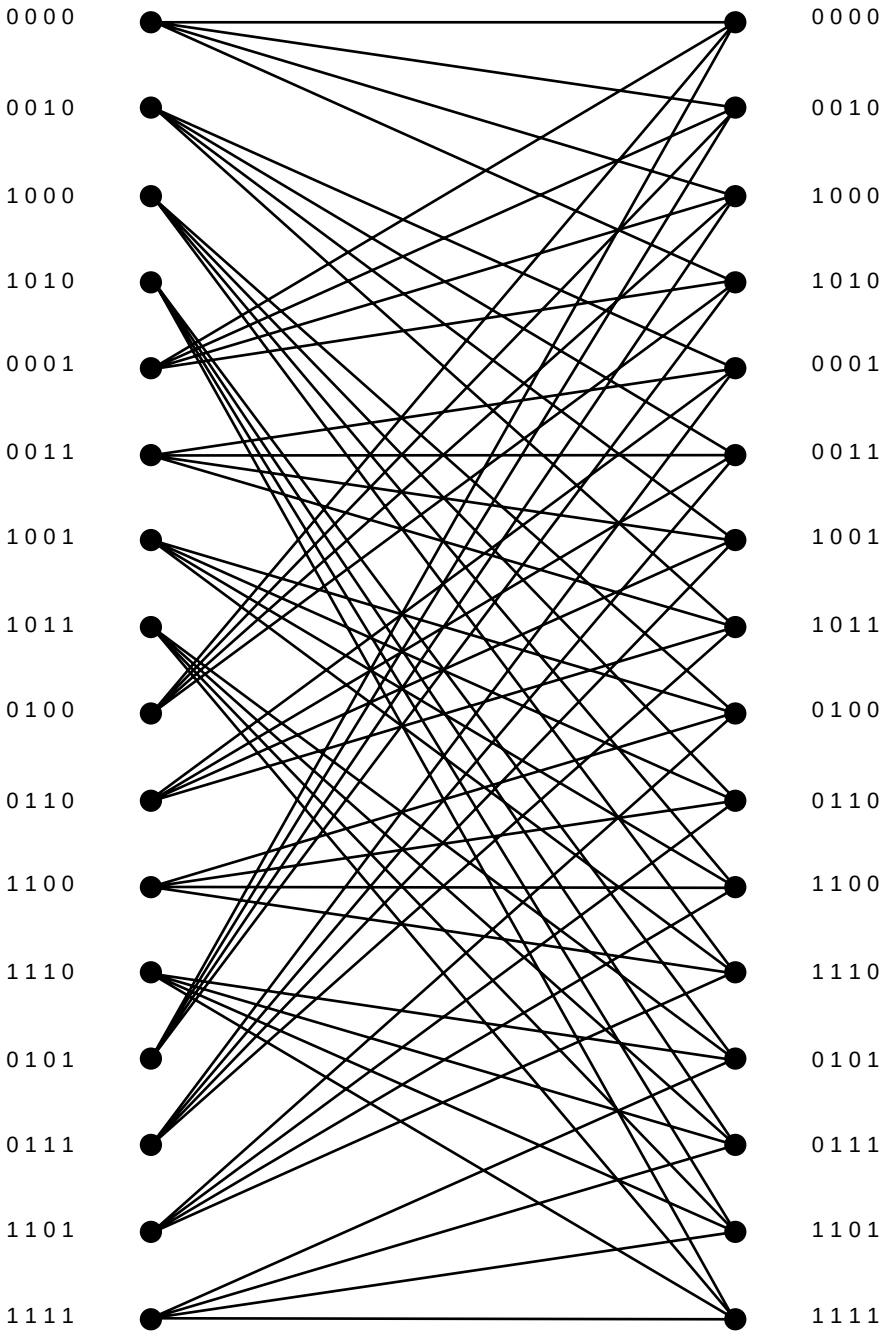

Fig. 3-13. The 16-state trellis diagram for FQPSK. Redrawn from [3].

Table 3-3. Trellis state transitions.

Current State	Input	Output	Next State
0 0 0 0	0 0	0 0	0 0 0 0
0 0 0 0	0 1	1 12	0 0 1 0
0 0 0 0	1 0	0 1	1 0 0 0
0 0 0 0	1 1	1 13	1 0 1 0
0 0 1 0	0 0	3 4	0 0 0 1
0 0 1 0	0 1	2 8	0 0 1 1
0 0 1 0	1 0	3 5	1 0 0 1
0 0 1 0	1 1	2 9	1 0 1 1
1 0 0 0	0 0	12 3	0 1 0 0
1 0 0 0	0 1	13 15	0 1 1 0
1 0 0 0	1 0	12 2	1 1 0 0
1 0 0 0	1 1	13 14	1 1 1 0
1 0 1 0	0 0	15 7	0 1 0 1
1 0 1 0	0 1	14 11	0 1 1 1
1 0 1 0	1 0	15 6	1 1 0 1
1 0 1 0	1 1	14 10	1 1 1 1
0 0 0 1	0 0	2 0	0 0 0 0
0 0 0 1	0 1	3 12	0 0 1 0
0 0 0 1	1 0	2 1	1 0 0 0
0 0 0 1	1 1	3 13	1 0 1 0
0 0 1 1	0 0	1 4	0 0 0 1
0 0 1 1	0 1	0 8	0 0 1 1
0 0 1 1	1 0	1 5	1 0 0 1
0 0 1 1	1 1	0 9	1 0 1 1
1 0 0 1	0 0	14 3	0 1 0 0
1 0 0 1	0 1	15 15	0 1 1 0
0 1 1 0	0 0	7 6	0 0 0 1
0 1 1 0	0 1	6 10	0 0 1 1
0 1 1 0	1 0	7 7	1 0 0 1
0 1 1 0	1 1	6 11	1 0 1 1
1 1 0 0	0 0	8 1	0 1 0 0
1 1 0 0	0 1	9 13	0 1 1 0
1 1 0 0	1 0	8 0	1 1 0 0
1 1 0 0	1 1	9 12	1 1 1 0
1 1 1 0	0 0	11 5	0 1 0 1
1 1 1 0	0 1	10 9	0 1 1 1

Table 3-3 (cont'd). Trellis state transitions.

Current State	Input	Output	Next State
1 1 1 0	1 0	11 4	1 1 0 1
1 1 1 0	1 1	10 8	1 1 1 1
0 1 0 1	0 0	6 2	0 0 0 0
0 1 0 1	0 1	7 14	0 0 1 0
0 1 0 1	1 0	6 3	1 0 0 0
0 1 0 1	1 1	7 15	1 0 1 0
0 1 1 1	0 0	5 6	0 0 0 1
0 1 1 1	0 1	4 10	0 0 1 1
0 1 1 1	1 0	5 7	1 0 0 1
0 1 1 1	1 1	4 11	1 0 1 1
1 1 0 1	0 0	10 1	0 1 0 0
1 1 0 1	0 1	11 13	0 1 1 0
1 1 0 1	1 0	10 0	1 1 0 0
1 1 0 1	1 1	11 12	1 1 1 0
1 1 1 1	0 0	9 5	0 1 0 1
1 1 1 1	0 1	8 9	0 1 1 1
1 1 1 1	1 0	9 4	1 1 0 1
1 1 1 1	1 1	8 8	1 1 1 1

original formulation of FQPSK as the combination of SQORC and a cross-correlator [half-symbol mapper] followed by I-Q carrier modulation did not recognize the existence of this inherent memory) and employ symbol-by-symbol detection. Based on the reformulation of FQPSK as a trellis-coded modulation (TCM) as discussed in Sec. 3.4, an optimum receiver would be one that exploited this characterization. In accordance with the foregoing representation of FQPSK as a TCM with 16 states, the optimum receiver (employing the Viterbi Algorithm) for FQPSK is illustrated in Fig. 3-14. Note that the 16 waveforms of Fig. 3-9 are not all equal in energy, and, thus, the matched filter outputs in this figure must be biased before applying them to the Viterbi decoder. Later on, we shall illustrate average BEP results obtained from a computer simulation of this receiver. For the moment, we shall just compare its asymptotic (limit of infinite energy-to-noise ratio) performance with that of the optimum receiver for conventional uncoded offset OQPSK based on normalized squared Euclidean distance, $d^2_{\min}/2\bar{E}_b$, where $\bar{E}_b$, denotes the average energy per bit. For the latter, $d^2_{\min}/2\bar{E}_b = 2$, which is the same as that for BPSK. For FQPSK, $d^2_{\min}/2\bar{E}_b$ is evaluated in Ref. 3 with the following results:

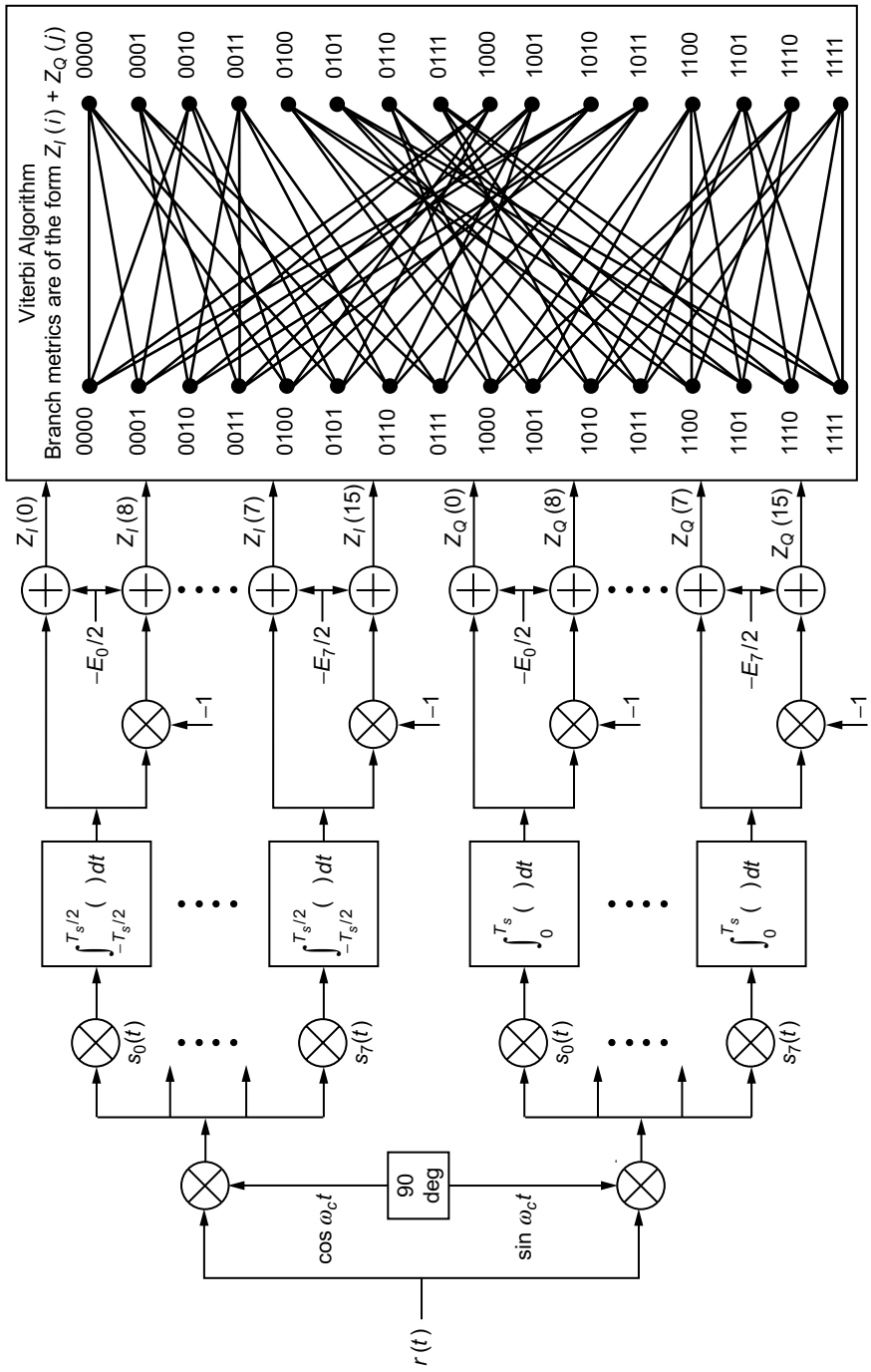

Fig. 3-14. Optimum trellis-coded receiver for FQPSK. Redrawn from [3].

$$\frac{d_{\min}^2}{2\bar{E}_b} = \frac{16\left[\frac{7}{4} - \frac{8}{3\pi} - A\left(\frac{3}{2} + \frac{4}{3\pi}\right) + A^2\left(\frac{11}{4} + \frac{4}{\pi}\right)\right]}{(7 + 2A + 15A^2)} \tag{3.5-1}$$

which for $A = 1/\sqrt{2}$ evaluates to

$$\frac{d_{\min}^2}{2\bar{E}_b} = 1.56 \tag{3.5-2}$$

Similarly, for enhanced FQPSK

$$\frac{d_{\min}^2}{2\bar{E}_b} = \frac{\left(3 - 6A + 15A^2\right)}{\frac{21}{8} - \frac{8}{3\pi} - A\left(\frac{1}{4} - \frac{8}{3\pi}\right) + \frac{29}{8}A^2} \tag{3.5-3}$$

which for $A = 1/\sqrt{2}$ coincidentally evaluates to (3.5-2), i.e., it is identical to that for FQPSK. Thus, the enhancement of FQPSK provided by using the waveforms of (3.3-1) as replacements for their equivalents in (3.2-1) is significantly beneficial from a spectral standpoint, with no penalty in receiver performance.

Finally, we conclude that as a trade against the significantly improved bandwidth efficiency afforded by FQPSK and its enhanced version relative to that of OQPSK, *an asymptotic loss of only 10 log(2/1.56) = 1.07 dB is experienced when optimum reception is employed.*[6]

3.6 Suboptimum Detection

As previously stated, FQPSK receivers traditionally use symbol-by-symbol detection, which is suboptimum and results in a significant loss relative to the performance of ideal OQPSK. In this section, we start by examining the extent of this degradation, following which we consider other suboptimum reception methods that also require less implementation complexity than the optimum receiver discussed in Sec. 3.5.

3.6.1 Symbol-by-Symbol Detection

Here we examine the performance of FQPSK when the detector makes decisions on a symbol-by-symbol basis, i.e., the inherent memory introduced by the trellis coding is ignored at the receiver. In order to understand how this can be accomplished, we will first establish the fact that in any typical transmission interval, there exists a fixed number (in particular, eight) of possible

[6] At smaller (finite) SNRs, the loss between uncoded OQPSK and trellis-decoded FQPSK will be even less.

waveforms (pulse shapes) that represent the FQPSK signal, and each of these occurs with equal probability. As such, from symbol-to-symbol, the FQPSK signal appears as an equiprobable M-ary signaling set (with $M = 8$) and thus can be detected accordingly. With this in mind, we shall investigate two possible simple structures, both of which are suboptimum relative to the trellis-coded receiver previously discussed that exploits the memory inherent in the modulation. The first structure is a standard offset QPSK receiver that employs simple I&Ds as detectors and, as such, ignores the pulse shaping associated with the above-mentioned M-ary symbol-by-symbol representation. The second structure, which shall be referred to as an average matched filter receiver, improves on the first one by replacing the I&Ds with matched filters, where the match is made to the average of the waveshapes in the M-ary signal set representation. Without loss in generality, the following description shall consider the case $n = 0$, corresponding to the I channel interval $-T_s/2 \leq t \leq T_s/2$ and the Q channel interval $0 \leq t \leq T_s$. We shall focus our attention only on the I channel, with the initial goal of defining the eight equally likely waveforms that typify an FQPSK waveform in the interval $0 \leq t \leq T_s$. To avoid confusion with the previously defined signals such as those defined in (3.2-1), we shall use upper-case notation, i.e., $S_i(t),\ i = 0, 1, \cdots, 7$ to describe these new waveforms. As we shall see momentarily, each of these new waveforms is composed of the latter half (i.e., that which occurs in the interval $0 \leq t \leq T_s/2$) of the I channel waveform transmitted in the interval $-T_s/2 \leq t \leq T_s/2$, followed by the first half (i.e., that which occurs in the interval $T_s/2 \leq t \leq T_s$) of the I channel waveform transmitted in the interval $T_s/2 \leq t \leq 3T_s/2$. As stated above, only eight such possible combinations can exist, and all are equiprobable.

3.6.1.1 Signal Representation. In Ref. 3, it is shown that for $d_{I0} = 1$ and $s_I(t) = s_0(t)$ in the interval $-T_s/2 \leq t \leq T_s/2$, the transmitted signal, $S_i(t)$, for the interval $0 \leq t \leq T_s$ is composed of the latter half of $s_0(t)$ followed by the first half of either $s_0(t), s_1(t), s_{12}(t)$ or $s_{13}(t)$. Looking at the definitions of $s_0(t), s_1(t), s_{12}(t), s_{13}(t)$ in (3.2-1), we see that this yields only two distinct possibilities for $S_i(t)$, namely,

$$\left.\begin{aligned} S_0(t) &= A, \quad 0 \leq t \leq T_s \\ S_1(t) &= \begin{cases} A, & 0 \leq t \leq \dfrac{T_s}{2} \\ A \sin \dfrac{\pi t}{T_s}, & \dfrac{T_s}{2} \leq t \leq T_s \end{cases} \end{aligned}\right\} \tag{3.6-1a}$$

both of which are equally likely.

Following a similar procedure (still for $d_{I0} = 1$), it can be shown that for each of the other possible waveforms in $-T_s/2 \leq t \leq T_s/2$, i.e., $s_1(t), s_2(t), s_3(t), s_4(t), s_5(t), s_6(t)$ and $s_7(t)$, there are only two possible distinct waveforms in $0 \leq t \leq T_s$, which are again equally likely. These possibilities are summarized in Table 3-4.

Table 3-4. Possible distinct signal pairs.

Signal in $-T_s/2 \leq t \leq T_s/2$	Signal in $0 \leq t \leq T_s$
$s_1(t)$	$S_2(t), S_3(t)$
$s_2(t)$	$S_0(t), S_1(t)$
$s_3(t)$	$S_2(t), S_3(t)$
$s_4(t)$	$S_4(t), S_5(t)$
$s_5(t)$	$S_6(t), S_7(t)$
$s_6(t)$	$S_4(t), S_5(t)$
$s_7(t)$	$S_6(t), S_7(t)$

where the signals $S_2(t), S_3(t), S_4(t), S_5(t), S_6(t), S_7(t)$ are defined as

$$\left.\begin{aligned}
S_2(t) &= 1-(1-A)\cos^2\frac{\pi t}{T_s}, \quad 0 \leq t \leq T_s \\
S_3(t) &= \begin{cases} 1-(1-A)\cos^2\dfrac{\pi t}{T_s}, & 0 \leq t \leq \dfrac{T_s}{2} \\ \sin\dfrac{\pi t}{T_s}, & \dfrac{T_s}{2} \leq t \leq T_s \end{cases} \\
S_4(t) &= \begin{cases} A\sin\dfrac{\pi t}{T_s}, & 0 \leq t \leq \dfrac{T_s}{2} \\ A, & \dfrac{T_s}{2} \leq t \leq T_s \end{cases} \\
S_5(t) &= A\sin\frac{\pi t}{T_s}, \quad 0 \leq t \leq T_s \\
S_6(t) &= \begin{cases} \sin\dfrac{\pi t}{T_s}, & 0 \leq t \leq \dfrac{T_s}{2} \\ 1-(1-A)\cos^2\dfrac{\pi t}{T_s}, & \dfrac{T_s}{2} \leq t \leq T_s \end{cases} \\
S_7(t) &= \sin\frac{\pi t}{T_s}, \quad 0 \leq t \leq T_s
\end{aligned}\right\} \tag{3.6-1b}$$

In comparing the performances of the suboptimum receivers of FQPSK to that of uncoded OQPSK, we shall reference them all to the same average transmitted power, $\bar{P}$, or, equivalently, the same average energy-per-bit to noise spectral density ratio, $\bar{E}_b/N_0 = \bar{P}T_b/N_0$. In order to do this, we must first compute the energy, $E_i = \int_0^{T_s} S_i^2(t)\,dt$, of each of the waveforms in (3.5-4a) and (3.5-4b) and take their average. The results are summarized below [3]:

$$\left.\begin{aligned}
E_0 &= A^2T_s \\
E_1 &= \frac{3}{4}A^2T_s \\
E_2 &= \left(\frac{3}{8} + \frac{1}{4}A + \frac{3}{8}A^2\right)T_s \\
E_3 &= \left(\frac{7}{16} + \frac{1}{8}A + \frac{3}{16}A^2\right)T_s \\
E_4 &= \frac{3}{4}A^2T_s \\
E_5 &= \frac{1}{2}A^2T_s \\
E_6 &= \left(\frac{7}{16} + \frac{1}{8}A + \frac{3}{16}A^2\right)T_s \\
E_7 &= \frac{1}{2}T_s
\end{aligned}\right\} \tag{3.6-2}$$

and

$$\bar{E} = \frac{1}{8}\sum_{i=0}^{7} E_i = \left(\frac{7 + 2A + 15A^2}{32}\right)T_s \tag{3.6-3}$$

Since the average power transmitted in the I channel is one-half the total (I+Q) average transmitted power, $\bar{P}$, then we have

$$\frac{\bar{P}}{2} = \frac{\bar{E}}{T_s} = \frac{7 + 2A + 15A^2}{32} \tag{3.6-4}$$

or, equivalently, the average energy-per-symbol is given by

$$\bar{P}T_s \triangleq \bar{E}_s = 2\bar{E}_b = \frac{7 + 2A + 15A^2}{16} T_s \tag{3.6-5}$$

Note that the evaluation of average energy per symbol based on the symbol-by-symbol M-ary representation of FQPSK is identical to that obtained from the representation as a trellis-coded modulation. Also note that for $A = 1$, which corresponds to SQORC modulation, we have $\bar{E}_s = (4/3)T_s$ which is consistent with the original discussions of this modulation in Ref. 11.

In accordance with our discussion at the beginning of Sec. 3.6.1, we shall consider two suboptimum receivers for symbol-by-symbol detection of FQPSK, the difference being the manner in which the detector is matched to the received signal. For the average matched filter case, the detector is implemented as a multiplication of the received signal by $\bar{S}(t) \triangleq \frac{1}{8}\sum_{i=0}^{7} S_i(t)$, followed by an I&D filter and binary hard decision device (see Fig. 3-15). For the OQPSK receiver, the detector is purely an I&D (i.e., matched to a rectangular pulse), which is tantamount to assuming $\bar{S}(t) = 1$. Thus, we can cover both cases at the same time, leaving $\bar{S}(t)$ as an arbitrary premultiplication pulse shape and later substitute the appropriate waveform.

Assuming the M-ary symbol-by-symbol representation of FQPSK just described, then the decision variable Z in Fig. 3-15 is given by

$$Z = \int_0^{T_s} S(t)\,\bar{S}(t)\,dt + \int_0^{T_s} n(t)\,\bar{S}(t)\,dt \triangleq \bar{Z} + N \tag{3.6-6}$$

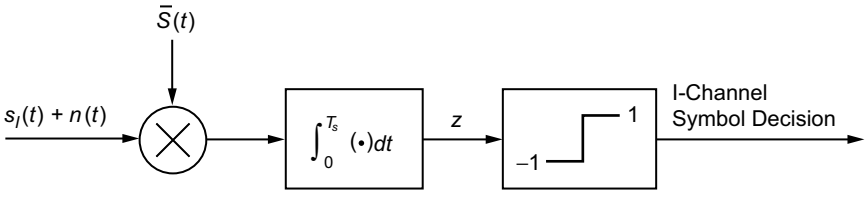

Fig. 3-15. Suboptimum receiver for FQPSK, based on symbol-by-symbol detection. Redrawn from [3].

where $S(t)$ is the transmitted waveform in $0 \leq t \leq T_s$ and ranges over the set of eight waveforms in (3.6-1a) and (3.6-2b) with equal probability. The random variable, N, is zero mean Gaussian with variance $\sigma_N^2 = N_0 E_{\bar{S}}/2$ where $E_{\bar{S}} \triangleq \int_0^{T_s} \bar{S}^2(t)\, dt$. Thus, the I channel symbol error probability (the same as the Q channel symbol error probability) conditioned on the particular $S(t) = S_i(t)$ corresponding to the transmitted symbol $d_{I0} = 1$ is shown to be

$$P_{si}(E) = \frac{1}{2}\ \text{erfc}\left(\sqrt{\frac{1}{N_0}\frac{\left(\int_0^{T_s} S_i(t)\,\bar{S}(t)\,dt\right)^2}{E_{\bar{S}}}}\right) \tag{3.6-7}$$

and hence the average symbol error probability is given by

$$P_s(E) \triangleq \frac{1}{8}\sum_{i=0}^{7} P_{si}(E) \tag{3.6-8}$$

3.6.1.2 Conventional OQPSK Receiver. For the conventional OQPSK receiver, we set $\bar{S}(t) = 1$, or equivalently, $E_{\bar{S}} = T_s$ in (3.6-7), resulting in

$$\begin{aligned} P_{si}(E) &= \frac{1}{2}\ \text{erfc}\left(\sqrt{\frac{T_s}{N_0}\left(\frac{1}{T_s}\int_0^{T_s} S_i(t)\,dt\right)^2}\right) \\ &= \frac{1}{2}\ \text{erfc}\left(\sqrt{\left(\frac{32}{7+2A+15A^2}\right)\frac{\bar{E}_b}{N_0}\left(\frac{E_i}{T_s}\right)^2}\right) \end{aligned} \tag{3.6-9}$$

Substituting the average energies from (3.6-2) in (3.6-9) for each signal and then performing the average as in (3.6-8) gives the final desired result for average symbol error probability, namely,

$$P_{si}(E) = \frac{1}{16}\,\text{erfc}\left(\sqrt{\left(\frac{32A^4}{7+2A+15A^2}\right)\frac{\bar{E}_b}{N_0}}\right) + \frac{1}{8}\,\text{erfc}\left(\sqrt{\left(\frac{18A^4}{7+2A+15A^2}\right)\frac{\bar{E}_b}{N_0}}\right) + \frac{1}{16}\,\text{erfc}\left(\sqrt{\left(\frac{(3+2A+3A^2)^2}{2(7+2A+15A^2)}\right)\frac{\bar{E}_b}{N_0}}\right) + \frac{1}{8}\,\text{erfc}\left(\sqrt{\left(\frac{(7+2A+3A^2)^2}{8(7+2A+15A^2)}\right)\frac{\bar{E}_b}{N_0}}\right) + \frac{1}{16}\,\text{erfc}\left(\sqrt{\left(\frac{8A^4}{7+2A+15A^2}\right)\frac{\bar{E}_b}{N_0}}\right) + \frac{1}{16}\,\text{erfc}\left(\sqrt{\left(\frac{8}{7+2A+15A^2}\right)\frac{\bar{E}_b}{N_0}}\right) \tag{3.6-10}$$

3.6.1.3 Average Matched Filter Receiver. For the average matched filter, we need to compute the correlations of each of the pulse shapes in (3.6-1a) and (3.6-1b) with the average pulse shape, $\bar{S}(t)$, and also the energy, $E_{\bar{S}}$, of the average pulse shape. Rewriting (3.6-7) in a form analogous to (3.6-9), namely,

$$P_{si}(E) = \frac{1}{2}\,\text{erfc}\left(\sqrt{\left(\frac{32}{7+2A+15A^2}\right)\frac{\bar{E}_b}{N_0}\,\frac{\left(\frac{1}{T_s}\int_0^{T_s} S_i(t)\,\bar{S}(t)\,dt\right)^2}{\frac{1}{T_s}E_{\bar{S}}}}\right) \tag{3.6-11}$$

then the results necessary to evaluate (3.6-11) are tabulated below:

$$\frac{1}{T_s}\int_0^{T_s} S_0(t)\,\bar{S}(t)\,dt = \frac{A}{4}\left[\frac{1}{2}+\frac{2}{\pi}+A\left(\frac{3}{2}+\frac{2}{\pi}\right)\right]$$

$$\frac{1}{T_s}\int_0^{T_s} S_1(t)\,\bar{S}(t)\,dt = \frac{1}{T_s}\int_0^{T_s} S_4(t)\,\bar{S}(t)\,dt$$

$$= \frac{A}{4}\left[\frac{1}{2}+\frac{5}{3\pi}+A\left(1+\frac{7}{3\pi}\right)\right]$$

$$\frac{1}{T_s}\int_0^{T_s} S_2(t)\,\bar{S}(t)\,dt = \frac{1}{4}\left[\frac{3}{8}+\frac{4}{3\pi}+A\left(\frac{3}{4}+\frac{2}{\pi}\right)+A^2\left(\frac{7}{8}+\frac{2}{3\pi}\right)\right] \tag{3.6-12}$$

$$\frac{1}{T_s}\int_0^{T_s} S_3(t)\,\bar{S}(t)\,dt = \frac{1}{T_s}\int_0^{T_s} S_6(t)\,\bar{S}(t)\,dt$$

$$= \frac{1}{4}\left[\frac{7}{16}+\frac{4}{3\pi}+A\left(\frac{5}{8}+\frac{7}{3\pi}\right)+A^2\left(\frac{7}{16}+\frac{1}{3\pi}\right)\right]$$

$$\frac{1}{T_s}\int_0^{T_s} S_5(t)\,\bar{S}(t)\,dt = \frac{A}{2}\left[\frac{1}{4}+\frac{2}{3\pi}+A\left(\frac{1}{4}+\frac{4}{3\pi}\right)\right]$$

$$\frac{1}{T_s}\int_0^{T_s} S_7(t)\,\bar{S}(t)\,dt = \frac{1}{2}\left[\frac{1}{4}+\frac{2}{3\pi}+A\left(\frac{1}{4}+\frac{4}{3\pi}\right)\right]$$

and

$$\frac{1}{T_s}E_{\bar{S}} = \frac{1}{16}\left[(1+A)^2\left(\frac{3}{2}+\frac{4}{\pi}\right)+\frac{3}{8}(1-A)^2-2\left(1-A^2\right)\left(\frac{1}{2}+\frac{2}{3\pi}\right)\right] \tag{3.6-13}$$

Finally, substituting (3.6-12) and (3.6-13) into (3.6-11) and averaging as in (3.6-8) gives the desired result, which we shall not explicitly write in closed form.

3.6.2 Average Bit-Error Probability Performance

The average BEP of the two suboptimum receivers discussed in Sec. 3.5.2.2 is illustrated in Fig. 3-16 for the case $A = 1/\sqrt{2}$. These results are obtained directly

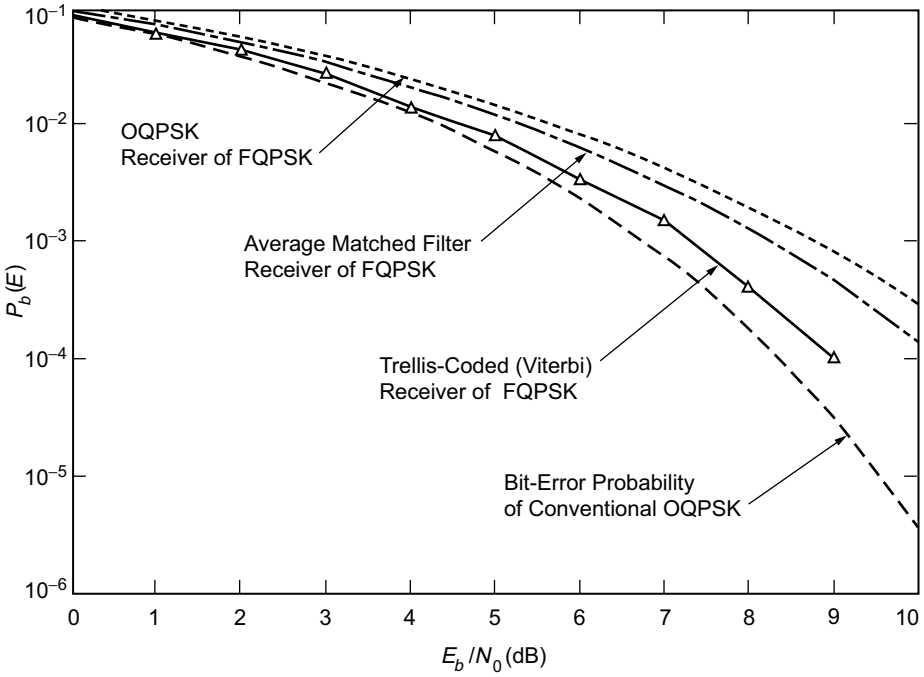

Fig. 3-16. Bit-error probability performance of various receivers of FQPSK modulation (reference curve is bit-error probability of OQPSK). Redrawn from [3].

from (3.6-10) for the OQPSK receiver and from (3.6-8) in combination with (3.6-11)–(3.6-13) for the average matched filter receiver. Also included in this figure is the performance corresponding to the optimum uncoded OQPSK receiver (same performance as for uncoded BPSK), i.e., $P_b(E) = (1/2)\text{erfc}\sqrt{E_b/N_0}$ as well as simulation results obtained for the optimum trellis-coded receiver of Fig. 3-14. We observe, as one might expect, that the average matched filter receiver outperforms the OQPSK receiver, since an attempt to match the transmitted pulse shape (even on an average basis) is better than no attempt at all. We also observe that the trellis-coded receiver at $P_b(E) = 10^{-4}$ is more than 1 dB better than the average matched filter receiver, granted that the latter is considerably simpler in implementation. Finally, for the same average BEP, the trellis-coded receiver of FQPSK is only about 0.6 dB inferior to uncoded OQPSK performance, which is a relatively small penalty paid for the vast improvement in PSD afforded by the former relative to the latter.

3.6.3 Further Receiver Simplifications and FQPSK-B Performance

As discussed above, symbol-by-symbol detection of FQPSK pays a significant penalty relative to optimum detection at the cost of a significant increase in implementation complexity. In Ref. 4, the authors introduce a simplified Viterbi-type receiver that still exploits the memory inherent in the modulation but has a reduced trellis and significantly less complexity (fewer states in the VA), with only a slight BEP degradation compared to that of the optimum (full Viterbi) receiver. The reduction in the number of states of the trellis comes about by grouping signal waveforms (see Fig. 3-9) with similar characteristics and using a single averaged matched filter for each group. In this sense, this simplified receiver acts as a compromise between the very simple averaged matched filter of Sec. 3.6.1.3, which uses a single matched filter equal to the average of all waveforms, and the optimum receiver, which uses a full bank of filters individually matched to each waveform. In discussing this simplified receiver, we shall consider its performance (obtained by simulation) in the context of FQPSK-B since, as shown in Figs. 3-7 and 3-8, FQPSK-B is much more spectrally efficient than unfiltered FQPSK but has ISI introduced by the filtering.

With reference to Fig. 3-14, the optimum Viterbi receiver for FQPSK implements a bank of matched filters to produce the correlations of the received signal with each of the 16 waveforms in Fig. 3-9 (actually only eight matched filters are required for each of the I and Q channels since $s_8(t), s_9(t), \cdots, s_{15}(t)$ are the negatives of $s_0(t), s_1(t), \cdots, s_7(t)\, s_0(t), s_1(t), \cdots, s_7(t)$). The Viterbi receiver then acts on these 32 correlation values (actually the branch metrics are the sum of each of the 16 energy-biased correlations from the corresponding I and Q channels) to produce a joint decision on the I and Q signals transmitted in a given symbol interval. A simplified FQPSK (or FQPSK-B) Viterbi receiver can be formed by observing certain similarities in the waveforms of Fig. 3-9 and thereby separating them into four different groups. For example, for $A = 1$, waveforms $s_0(t), s_1(t), \cdots, s_3(t)$ would become identical (similarly, for waveforms $s_8(t), s_9(t), \cdots, s_{11}(t)$). Thus, it is reasonable for arbitrary A to form group 1 as $s_0(t), s_1(t), \cdots, s_3(t)$ and also group 3 as $s_8(t), s_9(t), \cdots, s_{11}(t)$. Likewise, for $A = 1$, waveforms $s_4(t), s_5(t), \cdots, s_7(t)$ would become identical (similarly, for waveforms $s_{12}(t), s_{12}(t), \cdots, s_{15}(t)$). Consequently, for arbitrary A, it is again reasonable to form group 2 as $s_4(t), s_5(t), \cdots, s_7(t)$ and also group 4 as $s_{12}(t), s_{12}(t), \cdots, s_{15}(t)$. A close examination of the mapping that produced the trellis of Fig. 3-13 reveals that with this grouping, the trellis-coded structure of FQPSK splits into two independent I and Q two-state trellises. By independent, we mean that the I and Q decisions are no longer produced jointly but rather separately by individual VAs acting on the enery-biased correlations derived from the I and Q demodulated signals, respectively. A block diagram of

the simplified FQPSK-B receiver[7] is illustrated in Fig. 3-17. First, the received signal is demodulated and correlated against the arithmetic average of each of the above waveform groups as given by

$$\left.\begin{aligned} q_0(t) &= \frac{1}{4}\sum_{i=0}^{3} s_i(t) \\ q_1(t) &= \frac{1}{4}\sum_{i=4}^{7} s_i(t) \\ q_2(t) &= \frac{1}{4}\sum_{i=8}^{11} s_i(t) = -q_0(t) \\ q_3(t) &= \frac{1}{4}\sum_{i=12}^{15} s_i(t) = -q_1(t) \end{aligned}\right\} \tag{3.6-14}$$

and illustrated in Fig. 3-18. Since $q_2(t)$ and $q_3(t)$ are the negatives of $q_0(t)$ and $q_1(t)$, only two correlators (matched filters) are needed for the I and Q channels. Next, the VA metrics are formed by energy-biasing the matched filter outputs, where the appropriate energies are now those corresponding to the group averaged waveforms in (3.6-14). Figure 3-19 shows the two-state trellis associated with the grouped signals for each of the I and Q channels. The trellis is symmetric and has two transitions to each state. The two VAs could also be combined into a single four-state VA. When compared to the full Viterbi receiver of Fig. 3-14, the simplified Viterbi receiver has 12 fewer correlators and an 8-fold reduction in the number of VA computations performed per decoded bit.

Figure 3-20 illustrates the simulated BEP of the simplified and full FQPSK-B Viterbi receivers and compares them with that of the conventional symbol-by-symbol I&D (also referred to as sample-and-hold (S&H) in Ref. 4) receiver and ideal QPSK (or equivalently OQPSK). The simulated channel includes a nonlinear solid-state power amplifier (SSPA) operating in full saturation, which restores a constant envelope to the transmitted signal. For the full 16-state Viterbi receiver, a truncation path length (number of bits decoding delay before making decisions) of 50 bits was used. Due to the short constraint length nature of the reduced trellis in the simplified receiver, a truncation path length of only

[7] The primary difference between the simplified receiver for unfiltered FQPSK and FQPSK-B is the inclusion of an appropriate bandpass filter at the input to the receiver to match the filtering applied to the modulation at the transmitter.

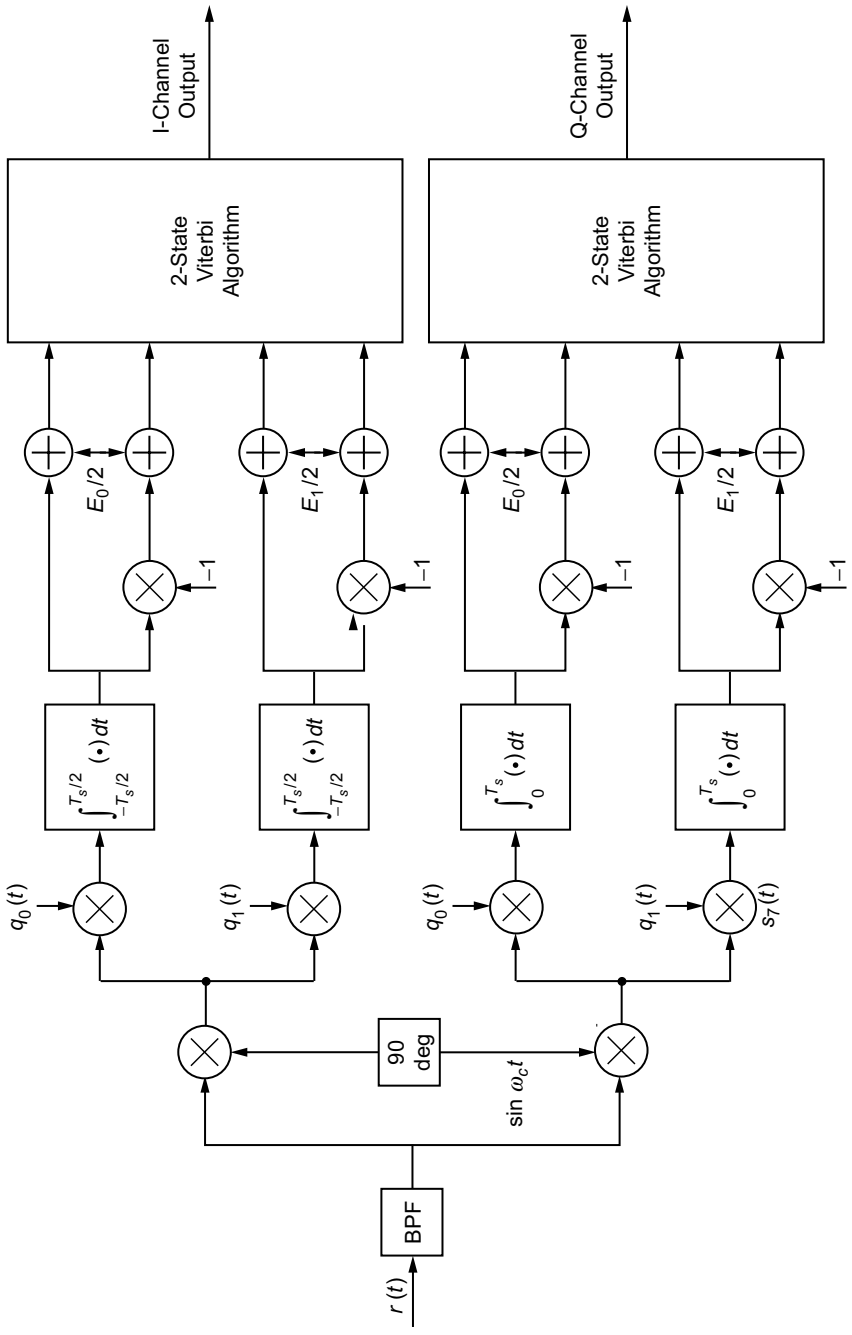

Fig. 3-17. Simplified FQPSK-B Viterbi receiver.

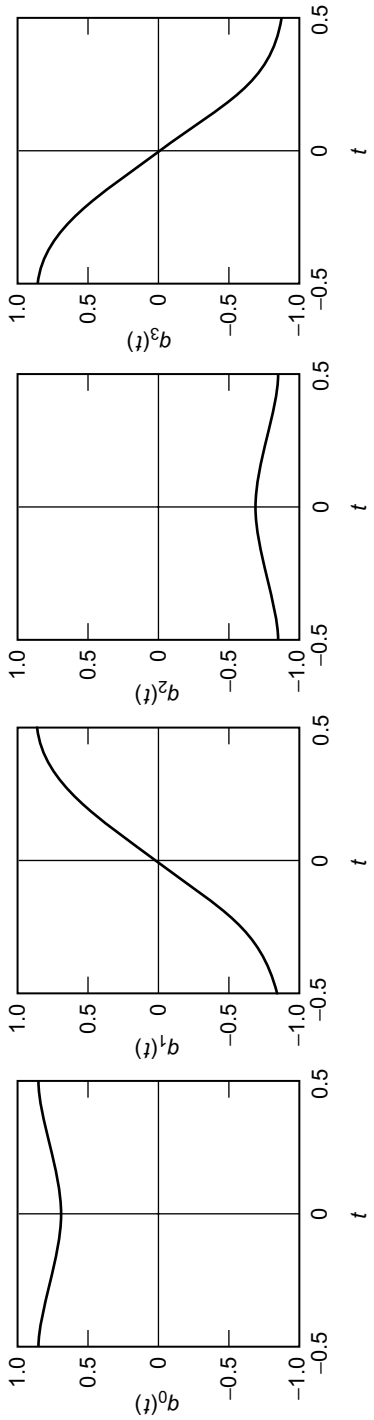

Fig. 3-18. Averaged waveforms for a simplified Viterbi receiver. Redrawn from [4].

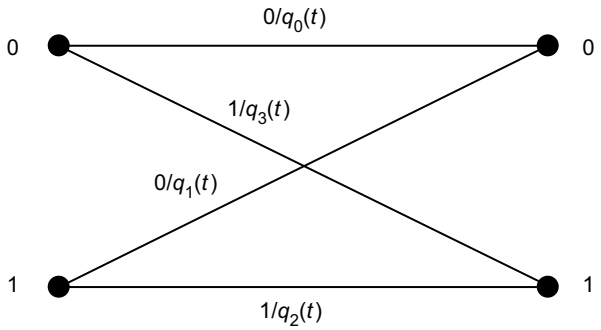

Fig. 3-19. Trellis diagram for a simplified FQPSK-B Viterbi receiver. Redrawn from [4].

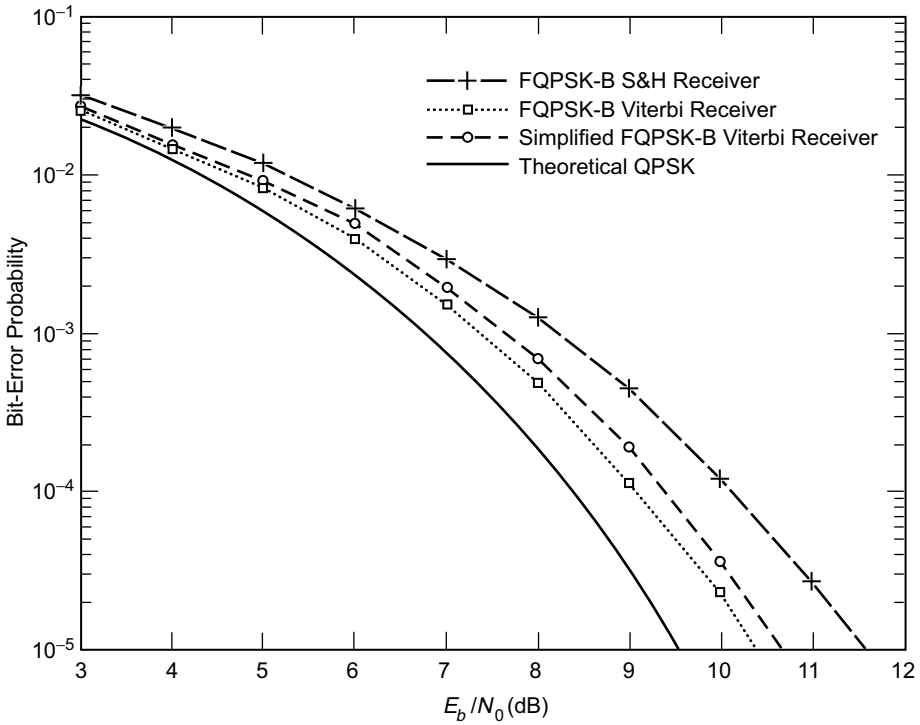

Fig. 3-20. Bit-error probability performance of FQPSK-B S&H and Viterbi receivers, with saturated SSPA. Redrawn from [4].

10 bits was shown to be necessary. Using the results in Fig. 3-20, Table 3-5 summarizes a comparison of the performances of the three FQPSK-B receivers at BEPs of 10^{-3} and 10^{-5}.

Table 3-5. Comparison of FQPSK-B performances.

Receiver	E_b/N_0 (dB) for 10^{-3} BEP	Loss Compared to Ideal QPSK at $P_b(E) = 10^{-3}$	E_b/N_0 (dB) for 10^{-5} BEP	Loss Compared to Ideal QPSK at $P_b(E) = 10^{-5}$
Full Viterbi receiver	7.4	0.6	10.4	0.8
Simplified receiver	7.65	0.85	10.7	1.1
S&H receiver	8.2	1.4	11.6	2.0

We observe from this table that the full Viterbi receiver performs 0.8 dB better than the symbol-by-symbol S&H receiver at a BEP of 10^{-3}, which is comparable to the analogous comparison made in Ref. 3 for unfiltered FQPSK. Thus, we can conclude that the Viterbi receiver works almost as well for the filtered version of FQPSK as it does for the unfiltered version. When compared with the full Viterbi receiver, the simplified FQPSK-B receiver suffers a slight degradation (0.25 dB) but is still better than the S&H receiver at a BEP of 10^{-3}. At a BEP of 10^{-5}, the full and simplified FQPSK-B Viterbi receivers are, respectively, 1.2 and 0.9 dB better than the S&H receiver.

Finally, to allow evaluation of BEP at values sufficiently small as to make simulation impractical, the BEP performance of the symbol-by-symbol S&H receiver for FQPSK-B was derived using superposition arguments in the appendix of Ref. 4 where the channel was modeled as being linear. Because of the ISI inherent in FQPSK-B, there are a large number of terms in the analytical expression for BEP. However, it was shown there that 32 terms is sufficient to give a very close match to simulation results (see Fig. 3-21). Comparing the results for the linear channel, given in Fig. 3-21, with the S&H simulation results for the nonlinear channel in Fig. 3-20 indicates a very small difference between them, the latter being the worse of the two.

3.7 Cross-Correlated Trellis-Coded Quadrature Modulation

Cross-correlated trellis-coded quadrature modulation (XTCQM) [12] is a technique that expands on the notion of combined bandwidth/power efficiency

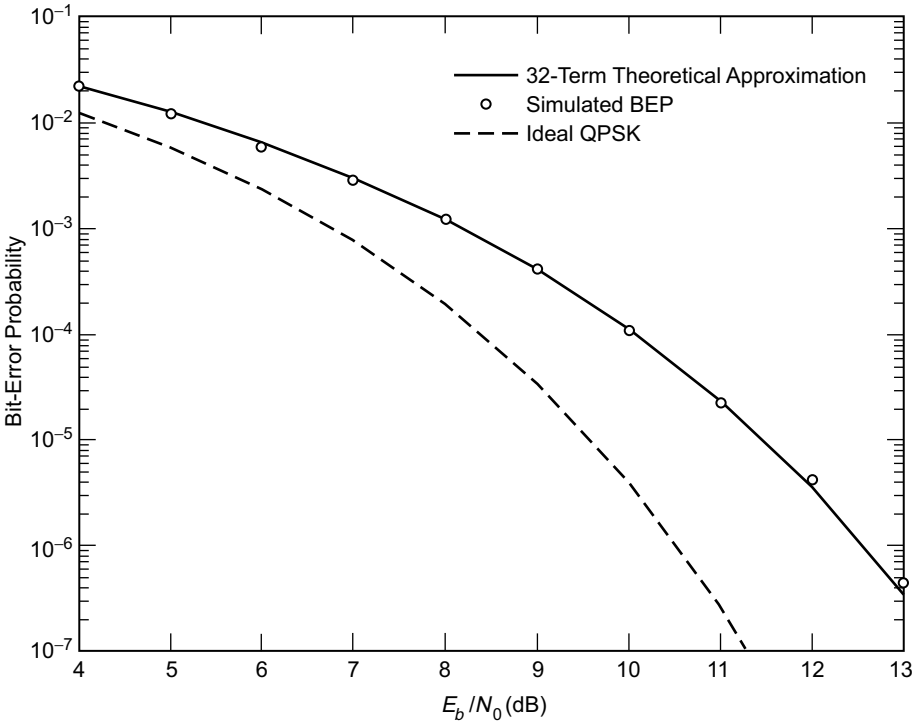

Fig. 3-21. Comparison of theoretical and simulated bit-error probability for an FQPSK-B S&H receiver.

indigenous to TCM, with particular emphasis on the spectral occupancy of the transmitted signal, while at the same time paying careful attention to the desirability of small envelope fluctuation. Whereas TCM combines conventional multilevel or multiphase modulations with error-correction coding through a suitable mapping that simultaneously exploits the desirable properties of these two functions, XTCQM focuses on a quadrature structure with coded I and Q channels that are cross-correlated and whose outputs are mapped into an M-ary waveform modulation. By virtue of its form, XTCQM produces a transmitted waveform with high spectral efficiency and allows for the design of a highly-power-efficient receiver. Its transmitter and receiver make use of standard, currently available devices, e.g., OQPSK modulator, convolutional encoders, matched filters, and Viterbi decoders, for its implementation. As we shall see, specific embodiments of XTCQM manifest themselves as FQPSK and trellis-coded versions of OQPSK and SQORC modulation. However, the generic structure provides considerably more flexibility for trading off between power and spectral efficiencies than these more restrictive embodiments.

3.7.1 Description of the Transmitter

With reference to Fig. 3-22, consider an input binary (± 1) i.i.d. data (information) sequence, $\{d_n\}$, at a bit rate, $R_b = 1/T_b$. This sequence is split into inphase (I) and quadrature (Q) sequences, $\{d_{In}\}$ and $\{d_{Qn}\}$, respectively, which consist of the even and odd bits of the information bit sequence, $\{d_n\}$, occurring at a rate, $R_s = 1/T_s = 1/2T_b$. We assume that the I and Q sequences, $\{d_{In}\}$ and $\{d_{Qn}\}$, are time synchronous and that the bit, d_{In} (or d_{Qn}), occurs during the interval $(n-(1/2))\,T_s \leq t \leq (n+(1/2))\,T_s$. As was the case for the TCM representation of FQPSK illustrated in Fig. 3-12, it is more convenient to work with the $(0,1)$ equivalents of the I and Q data sequences, namely, $\{D_{In}\}$ and $\{D_{Qn}\}$, as in (3.4-1). The sequences, $\{D_{In}\}$ and $\{D_{Qn}\}$, are applied to I and Q rate $r = 1/N$ convolutional encoders (the two encoders are in general different, i.e., they have different tap connections and different modulo 2 summers but are assumed to have the same code rate). Let $\left\{E_{Ik}\,\middle|_{k=1}^{N}\right\}$ and $\left\{E_{Qk}\,\middle|_{k=1}^{N}\right\}$ respectively denote the sets of N $(0,1)$ output symbols of the I and Q convolutional encoders corresponding to a single-bit input to each. These sets of output symbols will be used to determine a pair of baseband waveforms, $s_I(t), s_Q(t)$, which ultimately modulate I and Q carriers for transmission over the channel. In order to generate an offset form of modulation, the signal, $s_Q(t)$, will be delayed by $T_s/2 = T_b$ s prior to modulation on the quadrature carrier.[8] The mapping of the symbol sets $\left\{E_{Ik}\,\middle|_{k=1}^{N}\right\}$ and $\left\{E_{Qk}\,\middle|_{k=1}^{N}\right\}$ into $s_I(t)$ and $s_Q(t)$ and the size and content (waveshapes) of the waveform sets from which the latter are selected are the two most significant constituents of the XTCQM modulation scheme.

3.7.1.1 The Mapping. The mapping of the sets $\left\{E_{Ik}\,\middle|_{k=1}^{N}\right\}$ and $\left\{E_{Qk}\,\middle|_{k=1}^{N}\right\}$ into $s_I(t)$ and $s_Q(t)$ is illustrated in Fig. 3-23. Consider that each of these sets of N $(0,1)$ output symbols is partitioned into three groups as follows: For the first group, let $I_{l_1}, I_{l_2}, \cdots, I_{l_{N_1}}$ be a subset containing N_1 elements of $\left\{E_{Ik}\,\middle|_{k=1}^{N}\right\}$ that will be used only in the selection of $s_I(t)$. For the second group, let $Q_{l_1}, Q_{l_2}, \cdots, Q_{l_{N_2}}$ be a subset containing N_2 elements of $\left\{E_{Ik}\,\middle|_{k=1}^{N}\right\}$ that will be used only in the selection of $s_Q(t)$. Finally, for the third group, let $I_{l_{N_1+1}}, I_{l_{N_1+2}}, \cdots, I_{l_{N_1+N_3}} = Q_{l_{N_2+1}}, Q_{l_{N_2+2}}, \cdots, Q_{l_{N_2+N_3}}$ be a subset containing N_3 elements of $\left\{E_{Ik}\,\middle|_{k=1}^{N}\right\}$ that will be used both for the selection of $s_I(t)$ and $s_Q(t)$; hence, the term "cross-correlation" in the name of the modulation

[8] Note that delaying the waveform one-half of a symbol at the output of the mapping allows synchronous demodulation and computation of the path metric at the receiver. This is also true for the TCM implementation of FQPSK and, as such, is different than the conventional FQPSK approach, which applies the half-symbol delay to the Q data stream prior to any further processing (see Fig. 3-1).

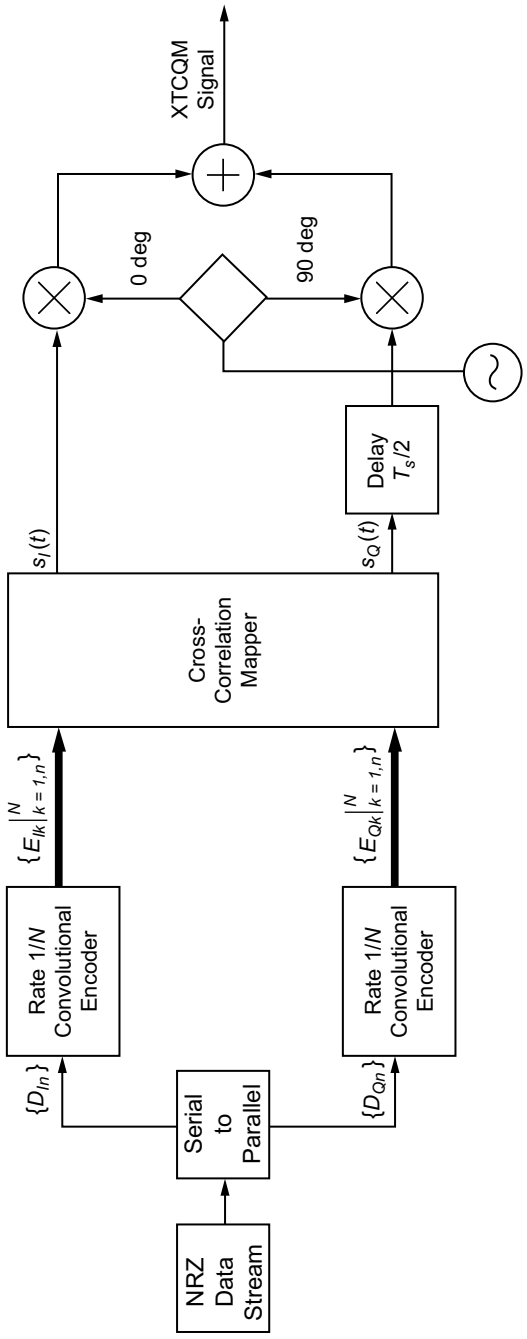

Fig. 3-22. Conceptual block diagram of an XTCQM transmitter.

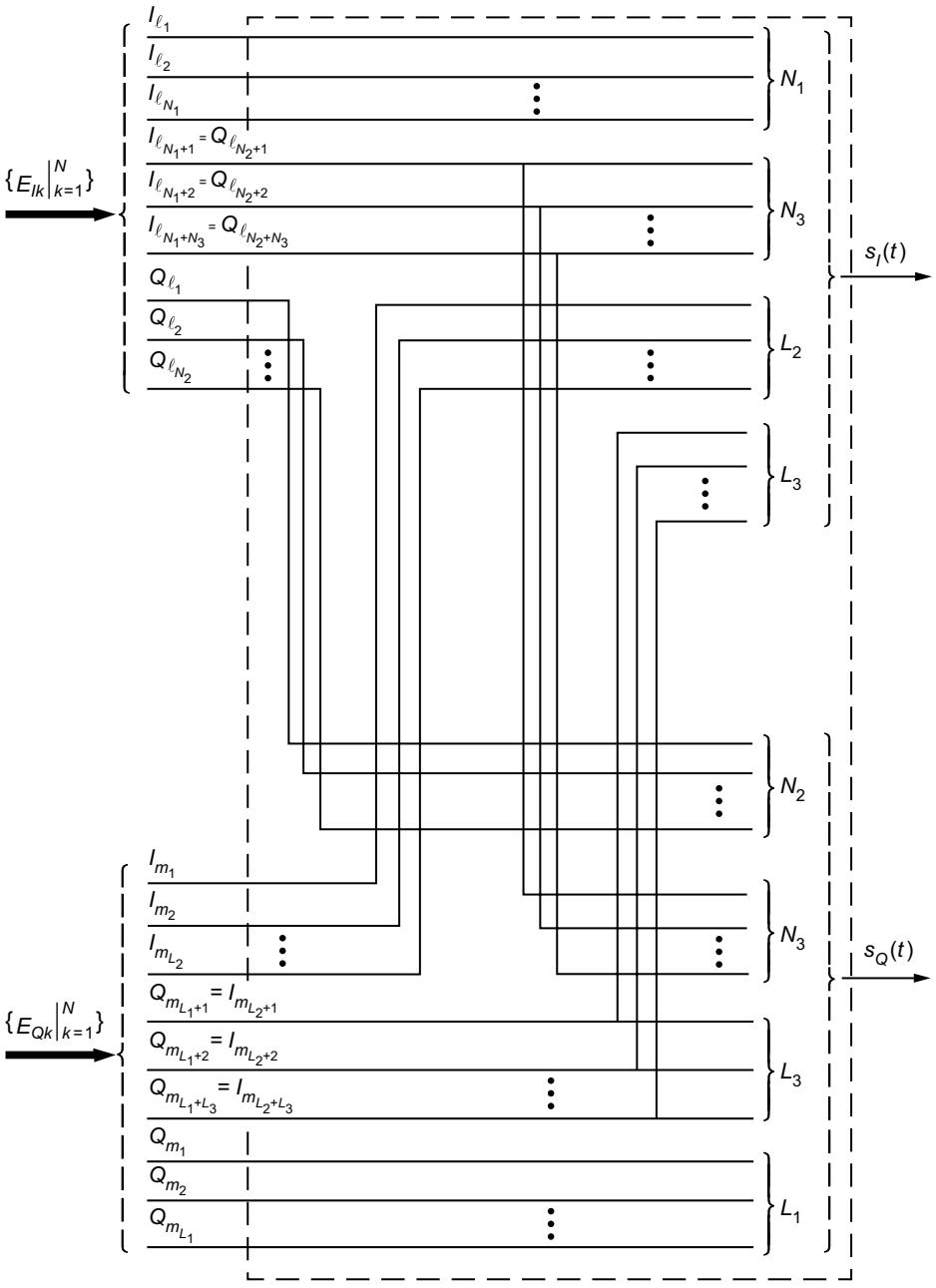

Fig. 3-23. Cross-correlation mapper.

scheme. Since all of the output symbols of the I encoder are used either to select $s_I(t)$ or $s_Q(t)$, or both, then, clearly, we must have $N_1+N_2+N_3=N$. A similar three-part grouping of the Q encoder ouput symbols $\left\{E_{Qk}\,\middle|_{k=1}^{N}\right\}$ is assumed to occur. That is, for the first group, let $Q_{m_1}, Q_{m_2}, \cdots, Q_{m_{L_1}}$ be a subset containing L_1 elements of $\left\{E_{Qk}\,\middle|_{k=1}^{N}\right\}$ that will be used only in the selection of $s_Q(t)$. For the second group, let $I_{m_1}, I_{m_2}, \cdots, I_{m_{L_2}}$ be a subset containing L_2 elements of $\left\{E_{Qk}\,\middle|_{k=1}^{N}\right\}$ that will be used only in the selection of $s_I(t)$. Finally, for the third group let $Q_{m_{L_1+1}}, Q_{m_{L_1+2}}, \cdots, Q_{m_{L_1+L_3}} = I_{m_{L_2+1}}, I_{m_{L_2+2}}, \cdots, I_{m_{L_2+L_3}}$ be a subset containing L_3 elements of $\left\{E_{Qk}\,\middle|_{k=1}^{N}\right\}$ that will be used both for the selection of $s_I(t)$ and $s_Q(t)$. Once again, since all of the output symbols of the Q encoder are used either to select $s_I(t)$ or $s_Q(t)$, or both, then, clearly, we must have $L_1+L_2+L_3=N$. More often than not, because of symmetry properties associated with the resulting modulation, we shall want to choose $L_1=N_1, L_2=N_2$ and $L_3=N_3$; however, the proposed XTCQM scheme is not restricted to this particular selection.

In summary, based on the above, the signal $s_I(t)$ will be determined from symbols $I_{l_1}, I_{l_2}, \cdots, I_{l_{N_1+N_3}}$ from the output of the I encoder and symbols $I_{l_1}, I_{l_2}, \cdots, I_{l_{L_2+L_3}}$ from the output of the Q encoder. Therefore, the size of the signaling alphabet used to select $s_I(t)$ will be $2^{N_1+N_3+L_2+L_3} \triangleq 2^{N_I}$. Similarly, the signal $s_Q(t)$ will be determined from symbols $Q_{l_1}, Q_{l_2}, \cdots, Q_{l_{L_1+L_3}}$ from the output of the Q encoder and symbols $Q_{l_1}, Q_{l_2}, \cdots, Q_{l_{N_2+N_3}}$ from the output of the I encoder. Thus, the size of the signaling alphabet used to select $s_Q(t)$ will be $2^{L_1+L_3+N_2+N_3} \triangleq 2^{N_Q}$. If it is desired that the size of the signaling alphabets for selecting $s_I(t)$ and $s_Q(t)$ be equal (a case of common interest), we need to have $N_I=N_Q$ or, equivalently, $L_1+N_2=N_1+L_2$. This condition is clearly satisfied if the condition $L_1=N_1, L_2=N_2$ is met; however, the former condition is less restrictive and does not require the latter to be true.

Having now assigned the encoder output symbols to either $s_I(t)$ or $s_Q(t)$ or both, the final step in the signal mapping is to form appropriate BCD numbers from these symbols and use these as indices i and j for choosing $s_I(t)=s_i(t)$ and $s_Q(t)=s_j(t)$, where $\left\{s_i(t)\,\middle|_{i=1}^{N_I}\right\}$ and $\left\{s_j(t)\,\middle|_{j=1}^{N_Q}\right\}$ are the signal waveform sets assigned for transmission of the I and Q channel signals. Specifically, let $I_0, I_1, \cdots, I_{N_I}$ be the particular set of symbols (taken from both I and Q encoder outputs) used to select $s_I(t)$, and let $Q_0, Q_1, \cdots, Q_{N_Q}$ be the particular set of symbols (taken from both I and Q encoder outputs) used to select $s_Q(t)$. Then, the BCD indices needed above are $i = I_{N_I-1}\times 2^{N_I-1}+\cdots+I_1\times 2^1+\cdots+I_0\times 2^0$ and $j = Q_{N_Q-1}\times 2^{N_Q-1}+\cdots+Q_1\times 2^1+\cdots+Q_0\times 2^0$.

3.7.1.2 The Signal Sets (Waveforms). While, in principle, any set of N_I waveforms of duration T_s s defined on the interval $-T_s/2 \le t \le T_s/2$, can be used for selecting the I channel transmitted signal, $s_I(t)$, and,

likewise, any set of N_Q waveforms of duration T_s s also defined on the interval $-T_s/2 \leq t \leq T_s/2$, can be used for selecting the Q channel transmitted signal, $s_Q(t)$, there are certain properties that should be invoked on these waveforms to make them desirable both from a power and spectral efficient standpoint. For the purpose of this discussion, we shall assume the special case $N_I = N_Q \triangleq N^*$ although, as pointed out previously, this is not a limitation on the invention. First, to achieve maximum distance in the waveform set (i.e., good power efficiency) one should divide the signal set, $\left\{ s_i(t) \big|_{i=1}^{N^*} \right\}$, into two equal parts-the signals in the second part being antipodal to (the negatives of) those in the first part. Mathematically, the signal set would have the composition $s_0(t), s_1(t), \cdots, s_{N^*/2-1}(t), -s_0(t), -s_1(t), \cdots, -s_{N^*/2-1}(t)$. Second, to achieve good spectral efficiency, one should choose the waveforms to be as smooth (i.e., as many continuous derivatives) as possible. Furthermore, to prevent discontinuities at the symbol transition-time instants, the waveforms should have a zero first derivative (slope) at their endpoints, $t = \pm T_s/2$.

In the next section, we explore the FQPSK embodiment as well as several other embodiments corresponding to well-known modulation schemes previously discussed in this monograph.

3.7.2 Specific Embodiments

3.7.2.1 FQPSK. Consider as an example the mapping scheme of Sec. 3.7.1.1, corresponding to $N_1 = N_2 = N_3 = 1$, i.e., rate $r = 1/N = 1/3$ encoders, and $L_1 = L_2 = L_3 = 1$, e.g., of the three output symbols from the I encoder, one is used to choose the I-channel signal, one is used to choose the Q-channel signal, and one is used to choose both the I- and Q-channel signals. Suppose now that the specific symbol assignments for the three partitions of the I encoder output are: I_3 (group 1), Q_0 (group 2), $I_2 = Q_1$ (group 3), and, similarly, the specific symbol assignments for the three partitions of the Q encoder output are: Q_3 (group 1), I_1 (group 2), $I_0 = Q_2$ (group 3). Furthermore, since $N_I = N_Q = 4$, the size of the signaling alphabet from which both $s_I(t)$ and $s_Q(t)$ are to be selected will be composed of $2^4 = 16$ signals. If, then, the rate 1/3 encoders are specifically chosen as in Fig. 3-12, and the 16 waveforms are selected as in Fig. 3-9, it immediately follows that FQPSK becomes a particular embodiment of the XTCQM scheme.

3.7.2.2 Trellis-Coded OQPSK. Consider an XTCQM scheme in which the mapping function is performed identical to that in the FQPSK embodiment (i.e., as in Fig. 3-12) but the waveform assignment is made as follows (see Fig. 3-24):

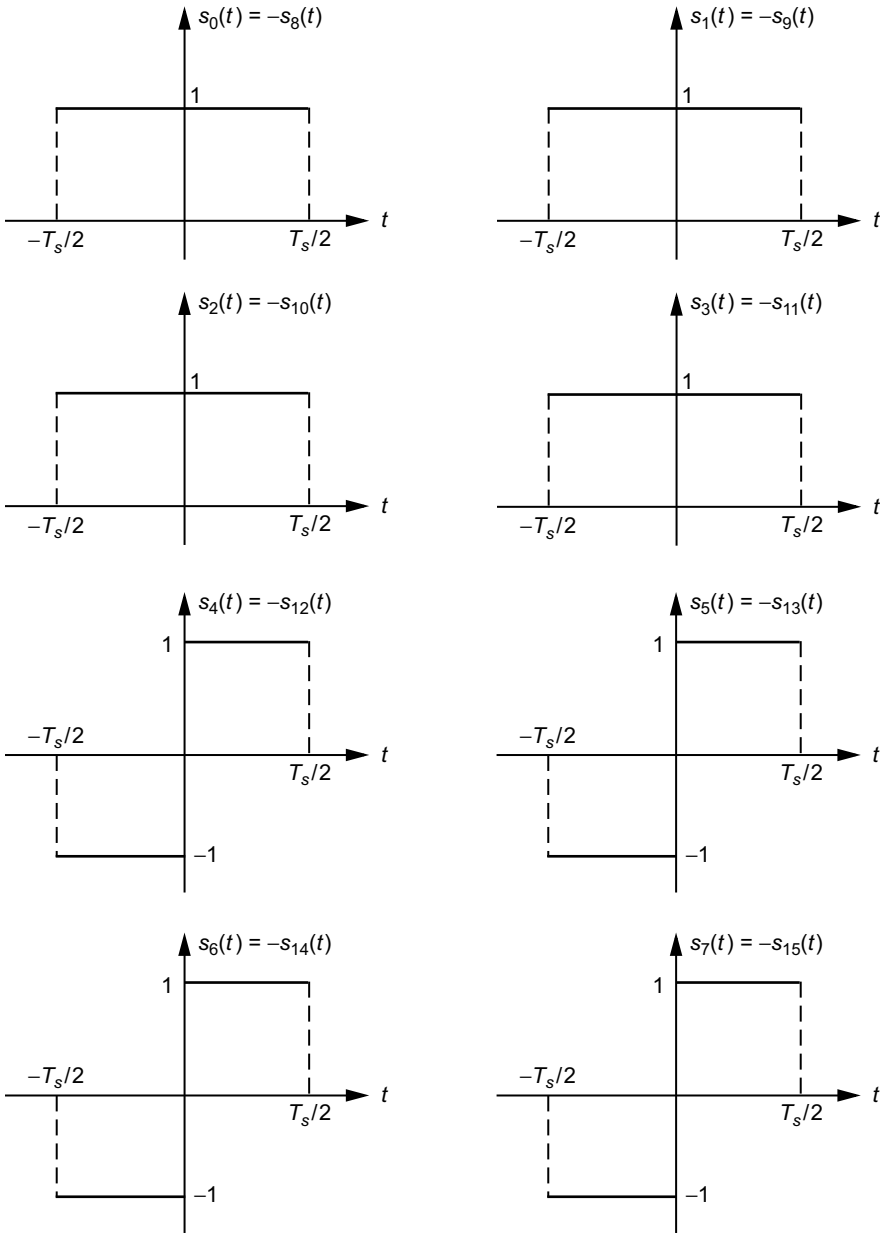

Fig. 3-24. Trellis-coded OQPSK full-symbol waveforms.

$$\left(\begin{array}{l} s_0(t) = s_1(t) = s_2(t) = s_3(t) = 1, \quad -\dfrac{T_s}{2} \le t \le \dfrac{T_s}{2} \\ \\ s_4(t) = s_5(t) = s_6(t) = s_7(t) = \begin{cases} -1, & -\dfrac{T_s}{2} \le t \le 0 \\ 1, & 0 \le t \le \dfrac{T_s}{2} \end{cases} \\ \\ s_i(t) = -s_{i-8}(t), \quad i = 8, 9, \cdots, 15 \end{array} \right) \tag{3.7-1}$$

that is, the first four waveforms are identical (a rectangular unit pulse) as are the second four (a split rectangular unit pulse), and the remaining eight waveforms are the negatives of the first eight. As such, there are only four unique waveforms, which we denote by $c_i(t)\big|_{i=0}^{3}$, where $c_0(t) = s_0(t), c_1(t) = s_4(t), c_2(t) = s_8(t), c_3(t) = s_{12}(t)$. Since in the BCD representations for each group of four identical waveforms, the two least significant bits are irrelevant, i.e., the two most significant bits are sufficient to define the common waveform for each group, we can simplify the mapping scheme by eliminating the need for I_0, I_1 and Q_0, Q_1. With reference to Fig. 3-12, elimination of I_0, I_1 and Q_0, Q_1 accomplishes two purposes. First, each encoder (both of which are now identical) needs only a single shift-register stage, and, second, the correlation between the two encoders insofar as the mapping of either one's output symbols to both $s_I(t)$ and $s_Q(t)$ has been eliminated, which, therefore, results in what might be termed a degenerate form of XTCQM, dubbed trellis-coded OQPSK [13]. The resulting embodiment is illustrated in Fig. 3-25. Since, insofar as the mapping is concerned, the I and Q channels are now decoupled (as indicated by the dashed line in the signal mapping block of Fig. 3-25), it is sufficient to examine the trellis structure and its distance properties for only one of the two channels (I or Q). The trellis diagram for either the I or Q channel of this modulation scheme would simply have two states and is illustrated in Fig. 3-26. The dashed line indicates a transition caused by an input "0," and the solid line indicates a transition caused by an input "1." Also, the branches are labeled with the output signal waveform that results from the transition. An identical trellis diagram would exist for the Q channel.

What is interesting about this embodiment of XTCQM is that, as far as the transmitted signal is concerned, it has a PSD identical to that of uncoded OQPSK (which is the same as for uncoded QPSK). In particular, because of the constraints imposed by the signal mapping, the waveforms $c_1(t) = s_4(t)$ and $c_3(t) = s_{12}(t)$ can never occur twice in succession. Thus, for any input information sequence, the sequence of signals $s_I(t)$ and $s_Q(t)$ cannot transition at a rate

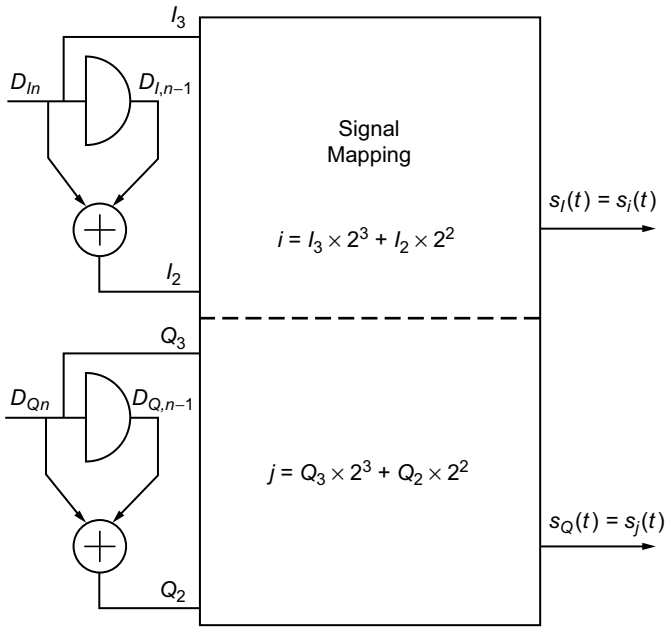

Fig. 3-25. The trellis-coded OQPSK embodiment of an XTCQM transmitter.

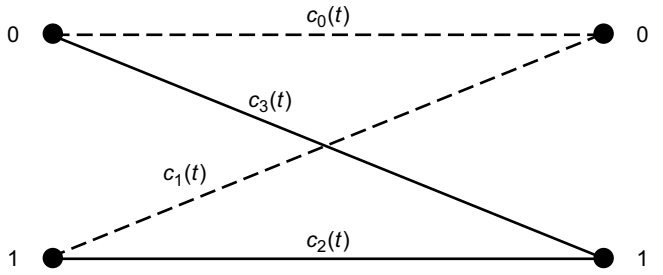

Fig. 3-26. Two-state trellis diagram for OQPSK.

faster than $1/T_s$. Stated another way, there cannot be any segment of sequences of $s_I\,(t)$ or $s_Q\,(t)$ that is constant for less than T_s s. This additional spectrum conservation constraint imposed by the signal mapping function of XTCQM will naturally result in a reduction of the coding (power) gain relative to that which could be achieved with another mapping, which does not prevent the successive repetition of $c_1\,(t)$ and $c_3\,(t)$. However, the latter occurrence would result in a bandwidth expansion by a factor of two.

3.7.2.3 Trellis-Coded SQORC. If, instead of a split rectangular pulse in (3.7-1), a sinusoidal pulse is used, namely,

$$\left.\begin{aligned} s_4(t) = s_5(t) = s_6(t) = s_7(t) = \sin\frac{\pi t}{T_s}, \quad -\frac{T_s}{2} \le t \le \frac{T_s}{2} \\ s_i(t) = -s_{i-8}(t), \quad i = 12, 13, 14, 15 \end{aligned}\right\} \tag{3.7-2}$$

the simplification of the mapping function shown in Fig. 3-25 again occurs (i.e., decoupling of the I and Q channels) and the trellis diagram of Fig. 3-26 is still appropriate for either the I or Q channel [13]. Once again, insofar as the transmitted signal is concerned, it has a PSD identical to that of uncoded SQORC (which is the same as for uncoded QORC).

3.7.2.4 Uncoded OQPSK. If we further simplify the signal assignment and mapping of Fig. 3-12 such that

$$\left.\begin{aligned} s_0(t) = s_1(t) = \cdots = s_7(t) = 1, \quad -\frac{T_s}{2} \le t \le \frac{T_s}{2} \\ s_i(t) = -s_{i-8}(t), \quad i = 8, 9, \cdots, 15 \end{aligned}\right\} \tag{3.7-3}$$

then in the BCD representations for each group of eight identical waveforms, the three least significant bits are irrelevant, i.e., only the first significant bit is needed to define the common waveform for each group. Hence, we can simplify the mapping scheme by eliminating the need for I_0, I_1, I_2 and Q_0, Q_1, Q_2. Defining the two unique waveforms, $c_0(t) = s_0(t), c_1(t) = s_8(t)$, we obtain the simplified degenerate mapping of Fig. 3-27, which corresponds to uncoded OQPSK with NRZ data formatting.

Likewise, if instead of the signal assignment in (3.7-3), we were to use

$$\left.\begin{aligned} s_0(t) = s_1(t) = \cdots = s_7(t) = \begin{cases} -1, & -\frac{T_s}{2} \le t \le 0 \\ 1, & 0 \le t \le \frac{T_s}{2} \end{cases} \\ s_i(t) = -s_{i-8}(t), \quad i = 8, 9, \cdots, 15 \end{aligned}\right\} \tag{3.7-4}$$

then the mapping of Fig. 3-27 produces uncoded OQPSK with Manchester (biphase) data formatting.

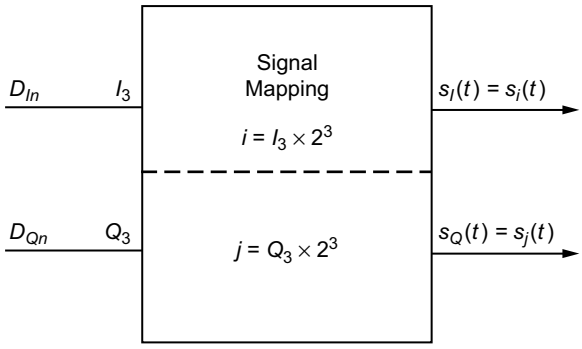

Fig. 3-27. Uncoded OQPSK embodiment of an XTCQM transmitter with NRZ data formatting.

3.8 Other Techniques

In the early 1980s, shaped BPSK (SBPSK) was introduced by Dapper and Hill [14] as a means of bandlimiting a BPSK signal while, at the same time, keeping its envelope constant. Further development of the SBPSK concept led to a variant of this scheme for offset quadrature modulation referred to as shaped offset QPSK (SOQPSK). In 2000, Hill [15] reported on a specific SOQPSK scheme with an enhanced waveform that offers spectral containment and detection efficiency comparable to or better than FQPSK-B, depending on the specifics of the comparison. Since SOPQSK is nonproprietary, whereas FQPSK-B is not, then, in view of the above-mentioned performance similarity, the former should be considered as a potential candidate in bandwidth-efficient modulation applications. In this section, we briefly review the results presented in Ref. 14 placing them in the context and notation of previous sections of this monograph. The material that follows should have been included in Chap. 2 since SOQPSK is truly constant envelope. However, we delayed discussing it there so that we might first present the material on FQPSK, thereby allowing the reader a better understanding of the performance comparison.

3.8.1 Shaped Offset QPSK

As a prelude to understanding the concept behind shaped offset QPSK (SOQPSK), it is instructive to first demonstrate that conventional OQPSK (rectangular pulse shaping implied) can be represented as a special case of CPM modulation. Specifically, OQPSK has the form in (2.8-1) together with (2.8-2), where $h = 1/2$; the frequency pulse, $g(t)$, of (2.8-3) is a delta function, i.e., $g(t) = (1/2)\delta(t)$ [equivalently, the phase pulse, $q(t)$, is a step function, i.e., $q(t) = (1/2)u(t)$]; and the ith element of the effective data se-

quence, α_i, can be shown to be related to the true input data bit sequence $\mathbf{a} = (\cdots, a_{-2}, a_{-1}, a_0, a_1, a_2, \cdots)$ by[9]

$$\alpha_i = (-1)^{i+1} \frac{a_{i-1}(a_i - a_{i-2})}{2} \tag{3.8-1}$$

Since the a_i's take on ± 1 values, then the α_i's come from a ternary (−1,0,+1) alphabet. However, in any given bit (half-symbol) interval, the α_i's can only assume one of two equiprobable values, namely, 0 and +1 or 0 and −1, with the further restriction that a +1 cannot be followed by a −1, or vice versa. Thus, in reality, the modulation scheme is a binary CPM but one whose data alphabet can vary (between two choices) from bit interval to bit interval. Another way of characterizing the variation rule for the data alphabet is as follows: If the previous bit is 0, then the data alphabet for the current bit is switched relative to that available for the previous bit, i.e., if it was (0,+1) for the previous transmission, it becomes (0,−1) for the current transmission, and vice versa. On the other hand, if the previous bit is a +1 or a −1, then the data alphabet for the current bit remains the same as that available for the previous bit, e.g., if it was (0,+1) for the previous transmission, it is again (0,+1) for the current transmission.

Since $h = 1/2$ together with the factor of 1/2 in $g(t)$ corresponds to a phase change of $\pi/2$ rad, then a value of $\alpha_i = 0$ suggests no change in carrier phase (no transition occurs in the I (or Q) data symbol sequence at the midsymbol time instant of the Q (or I) data symbol), whereas a value of $\alpha_i = \pm 1$ suggests a carrier phase change of $\pm\pi/2$ (a transition occurs in the I (or Q) data symbol sequence at the midsymbol time instant of the Q (or I) data symbol). Finally, note that since the duration of the frequency pulse does not exceed the baud (bit) interval, then, in accordance with the discussion in Sec. 2.7, the CPM representation of OQPSK is full response and can be implemented with the cascade of a precoder satisfying (3.8-1) and a conventional CPM modulator such as in Fig. 2-7.

In the early conception of SOQPSK, a rectangular pulse of duration equal to the bit period was used for $g(t)$. In this sense, one might think that SOQPSK resembled MSK; however, we remind the reader that for the latter, the data alphabet was fixed at $-1, +1$ whereas for the former it varies between $0, -1$ and $0, +1$. Thus, whereas in a given bit interval, the phase for MSK is always linearly varying with either a positive or negative slope, the phase for SOQPSK can either vary linearly or remain stationary. As such, the phase trellis for SOQPSK will have plateaus during the bit intervals where $\alpha_i = 0$. Since for OQPSK

[9] Note that the I and Q data symbols a_{In}, a_{Qn} of (2.2-4) are respectively obtained as the even and odd bits of the sequence $\mathbf{a}$. Also note that, whereas the I-Q representation of OQPSK contains I and Q data sequences at the symbol rate $1/T_s$, the effective data sequence for the CPM representation occurs at the half-symbol (bit) rate, $1/(T_s/2) = 1/T_b$.

itself, the phase trellis would only have plateaus (no linear variations), then in this sense, SOQPSK with a square frequency pulse can be viewed as a hybrid of OQPSK and MSK.

In Ref. 14, two variants of SOQPSK, referred to as SOQPSK-A and SOQPSK-B, were considered based upon a frequency pulse shape that is a minor modification of the impulse response corresponding to a spectral raised cosine filter. The modification corresponds to applying a raised cosine (in the time domain) window to the above impulse response, which alone would have doubly infinite extent. Specifically,

$$g(t) = g_1(t)\, g_2(t) \tag{3.8-2}$$

where

$$\left.\begin{aligned} g_1(t) &\triangleq \frac{A\cos\pi\alpha Bt/T_s}{1-4\left(\alpha Bt/T_s\right)^2}\,\frac{\sin\pi Bt/T_s}{\pi Bt/T_s} \\ g_2(t) &\triangleq \begin{cases} 1, & \left|\dfrac{t}{T_s}\right| \le \varepsilon_1 \\ \dfrac{1}{2}+\dfrac{1}{2}\cos\dfrac{\pi\left(|t/T_s|-\varepsilon_1\right)}{\varepsilon_2}, & \varepsilon_1 < \left|\dfrac{t}{T_s}\right| \le \varepsilon_1+\varepsilon_2 \\ 0, & \left|\dfrac{t}{T_s}\right| > \varepsilon_1+\varepsilon_2 \end{cases} \end{aligned}\right\} \tag{3.8-3}$$

In (3.8-3), $g_1(t)$ is the impulse response of the spectral raised cosine filter with amplitude, A, fractional rolloff factor, α, and additional time-scaling factor, B, and $g_2(t)$ is the above-mentioned windowing function that limits the duration of $g(t)$ to $2(\varepsilon_1+\varepsilon_2)T_s$. The values of the parameters in (3.8-3) that define SOQPSK-A and SOQPSK-B are tabulated in Table 3-6.

Table 3-6. Parameter values for SOQPSK-A and SOQPSK-B.

Parameter	SOQPSK-A	SOQPSK-B
α	1.0	0.5
B	1.35	1.45
ε_1	1.4	2.8
ε_2	0.6	1.2

From Table 3-6, we see that SOQPSK-A has a frequency pulse duration of $4T_s$ ($8T_b$) and SOQPSK-B has a frequency pulse duration of $8T_s$ ($16T_b$); thus, both of these schemes correspond to partial-response CPM. Hill [15] clearly indicates that the parameter values chosen to represent SOQPSK-A and SOQPSK-B "are not 'optimum' in any mathematical sense" but are simply representative examples of what can be achieved with the functional form in (3.8-2) together with (3.8-3).

Figure 3-28 illustrates the simulated PSDs of SOQPSK-A and SOQPSK-B along with that of the earlier full-response version of SOQPSK, which uses a rectangular T_b-s frequency pulse (Ref. 14 refers to the latter as military standard (MIL-STD) SOQPSK, since this version was in fact adopted as a military standard). We observe from this figure that the difference between SOQPSK-A and SOQPSK-B down to a level of −40 dB is virtually nil; however, below that level, SOQPSK-A offers a significant spectral improvement over SOQPSK-B. To compare the spectral behavior of SOQPSK with that of FQPSK-B, Hill [15] uses the results illustrated in Fig. 3-28 (which were verified by experimental measurement) along with measured hardware results obtained from an FQPSK-B modem built by RF Networks, Inc. and tested at the ARTM Project facility at Edwards

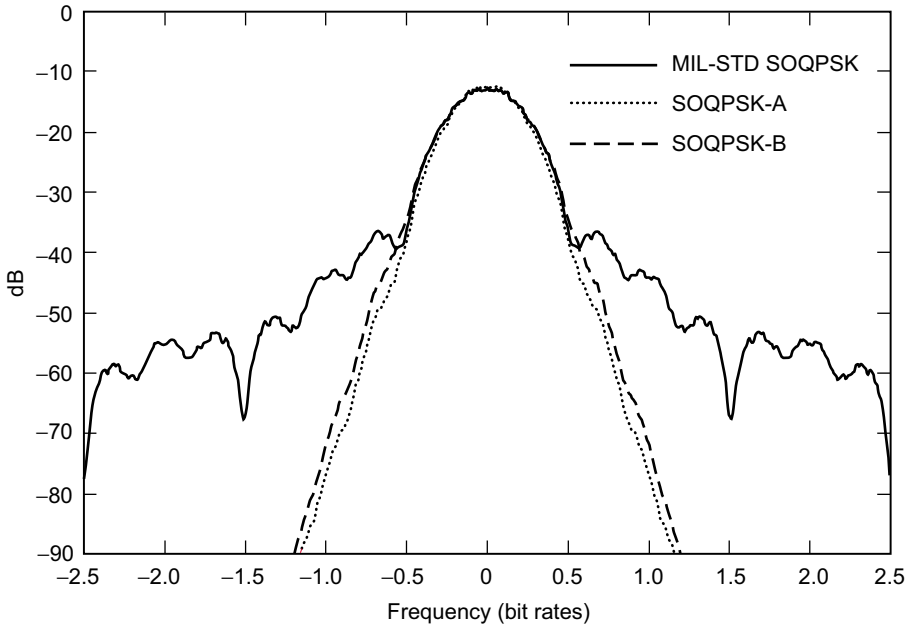

Fig. 3-28. Simulated SOQPSK power spectral densities. Resolution bandwidth = 20 kHz. Redrawn from [15].

Air Force Base. Two sets of comparisons were made. In one case, the PSDs of SOQPSK-A and SOQPSK-B were compared with that of FQPSK-B without nonlinear amplification (as such, FQPSK-B is therefore nonconstant envelope). Figure 3-29 illustrates this comparison, where it can be observed that: (a) down to −25 dB, the three PSDs are virtually indistinguishable from one another, and (b) below −25 dB, FQPSK-B is the most compact, SOQPSK-A is slightly wider, and SOQPSK-B is wider still. The second comparison pits SOQPSK-A against FQPSK-B with nonlinear amplification (to produce a constant envelope modulation). Figure 3-30 illustrates this comparison, where SOQPSK-A now has a narrower PSD than FQPSK-B. (Note that since SOQPSK-A is constant envelope, the nonlinear amplification theoretically has no effect on its PSD. This was also confirmed experimentally, as indicated in Ref. 14).

To complete this discussion, Figs. 3-31(a) and 3-31(b) illustrate the simulated and measured BEP of MIL-STD SOQPSK, SOQPSK-A, SOQPSK-B, and FQPSK-B. The simulated results were obtained using a conventional OQPSK receiver that for all three modulations is suboptimum, since no attempt is made to match the equivalent I and Q pulse shapes. We observe from Fig. 3-31(a) that

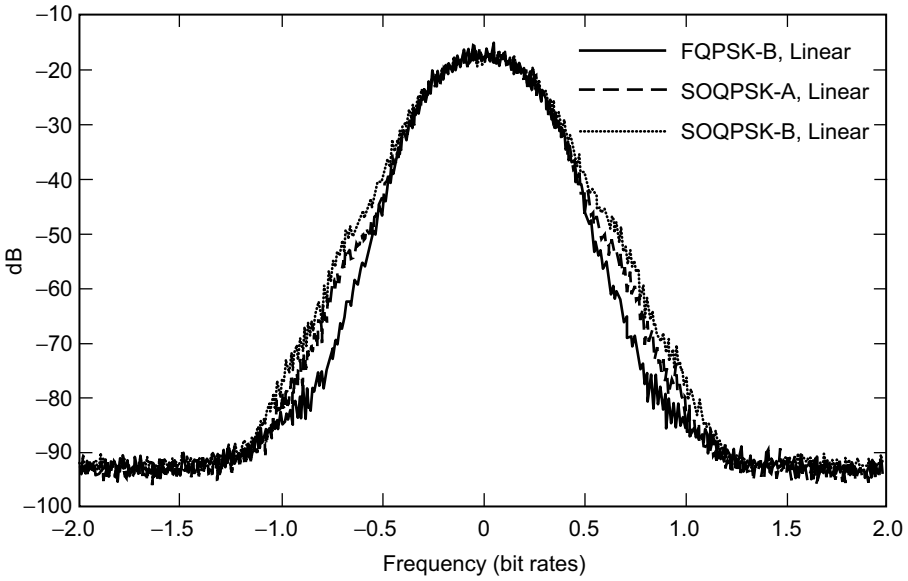

Fig. 3-29. A comparison of the power spectral densities of SOQPSK and FQPSK-B without nonlinear amplification. Resolution bandwidth = 3 kHz, video bandwidth = 10 kHz, and data rate = 1.0 Mb/s. Data from the Advanced Range Telemetry Lab at Edwards Air Force Base, California. Redrawn from [15].

at a BEP of 10^{-5}, MIL-STD SOQPSK and SOQPSK-B produce about the same E_b/N_0 performance penalty, i.e., 2.4 dB, relative to ideal unfiltered OQPSK, whereas SOQPSK-A is only about 0.25 dB worse. The conclusion to be drawn from this fact is that the MIL-STD variant of SOQPSK, which employs a rectangular frequency pulse, can be modified to generate SOQPSK-A or SOQPSK-B, with virtually no penalty in detection (power) efficiency but a considerable improvement in bandwidth efficiency. It seems clear that further optimizing the receiver by including appropriate matched filtering and trellis decoding (recall that SOQPSK-A or SOQPSK-B are memory modulations by virtue of the fact that they are partial-response CPMs) would yield additional improvement in power efficiency. The measured BEP performance curves in Fig. 3-31(b) reveal that at a BEP of 10^{-5}, SOQPSK-A is comparable but about 0.5 dB worse than nonlinearly amplified FQPSK-B, whereas SOQPSK-B is about 0.75 dB better.

As this book was going to press, the author became aware of results [16] describing a trellis detector for SOQPSK-A and SOQPSK-B that, by accounting for the pulse shaping and memory inherent in the modulations, provides superior detection performance as compared to the traditional OQPSK detector considered in Ref. 15.

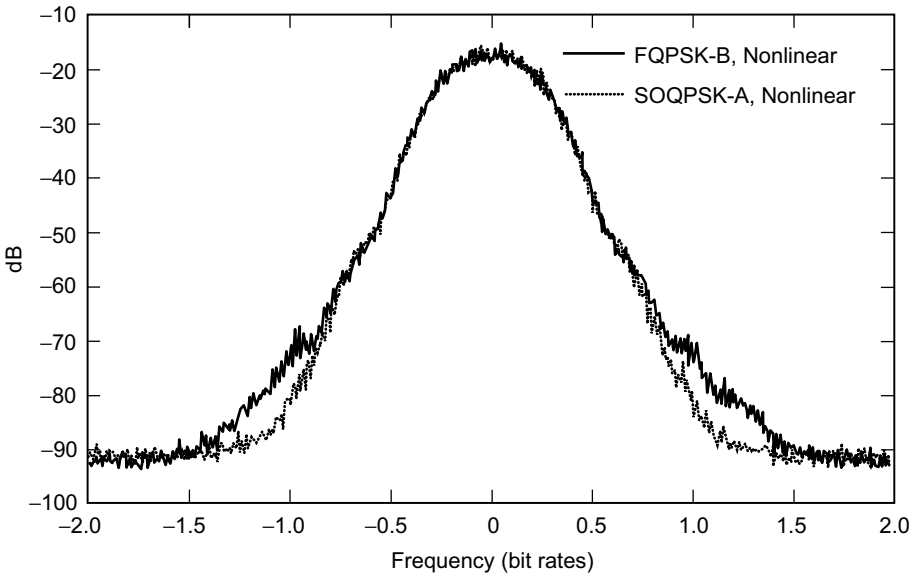

Fig. 3-30. A comparison of the power spectral densities of SOQPSK-A and FQPSK with nonlinear amplification. Resolution bandwidth = 3 kHz, video bandwidth = 10 kHz, and data rate = 1.0 Mb/s. Data from the Advanced Range Telemetry Lab at Edwards Air Force Base, California. Redrawn from [15].

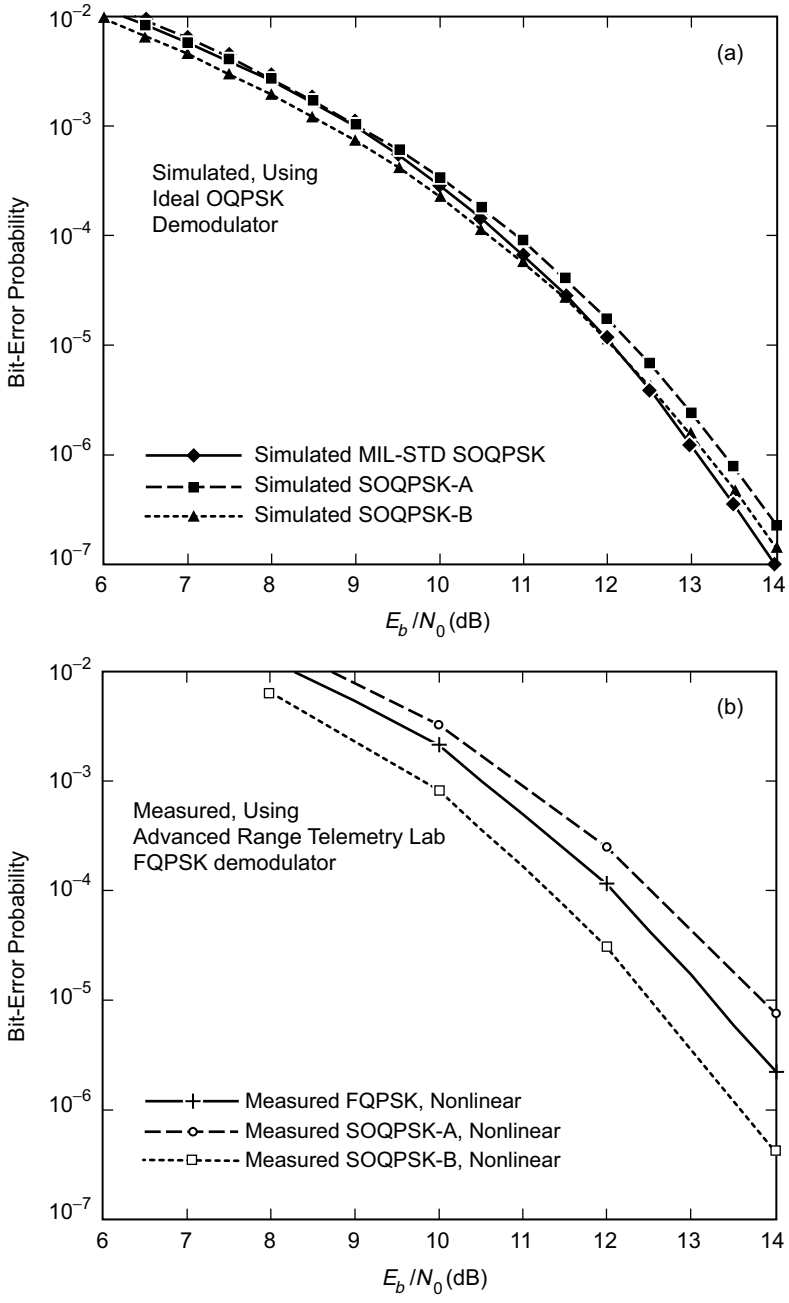

Fig. 3-31. Bit-error probability results for FQPSK-B, SOQPSK-A, and SOQPSK-B: (a) simulated and (b) measured. Redrawn from [15].

References

[1] K. Feher et al., U.S. patents: 4,567,602; 4,339,724; 4,644,565; 5,784,402; 5,491,457. Canadian patents: 1,211,517; 1,130,871; 1,265,851.

[2] S. Kato and K. Feher, "XPSK: A new cross-correlated phase-shift-keying modulation technique," *IEEE Transactions on Communications*, vol. 31, no. 5, pp. 701–707, May 1983.

[3] M. K. Simon and T.-Y. Yan, "Unfiltered Feher-patented quadrature phase-shift-keying (FQPSK): Another interpretation and further enhancements: Parts 1, 2," *Applied Microwave & Wireless Magazine*, pp. 76–96/pp. 100–105, February/March 2000.

[4] D. Lee, M. K. Simon, and T.-Y. Yan, "Enhanced Performance of FQPSK-B Receiver Based on Trellis-Coded Viterbi Demodulation," *International Telemetering Conference*, San Diego, California, October 23–26, 2000.

[5] K. Feher, *Wireless Digital Communications: Modulation and Spread Spectrum Applications*, Upper Saddle River, New Jersey: Prentice Hall, 1995.

[6] K. Feher, *Digital Communications: Satellite/Earth Station Engineering*, Littleton, Colorado: Crestone Engineering, 1996.

[7] K. Feher, "F-QPSK-A superior modulation technique for mobile and personal communications," *IEEE Transactions on Broadcasting*, vol. 39, no. 2, pp. 288–294, June 1993.

[8] K. Feher, "FQPSK transceivers double the spectral efficiency of wireless and telemetry systems," *Applied Microwave & Wireless Magazine*, June 1998. Also presented at *European Telemetry Conference*, Garmish-Patternk, Germany, May 5–8, 1998.

[9] W. L. Martin, T-Y. Yan, and L. V. Lam, "CCSDS-SFCG: Efficient modulation methods study at NASA/JPL, Phase 3: End-to end performance," *Proceedings of the SFGC Meeting*, Galveston, Texas, September 16–25, 1997.

[10] T. Le-Ngoc, K. Feher, and H. Pham Van, "New modulation techniques for low-cost power and bandwidth efficient satellite earth stations," *IEEE Transactions on Communications*, vol. 30, no. 1, pp. 275–283, January 1982.

[11] M. C. Austin and M. V. Chang, "Quadrature overlapped raised-cosine modulation," *IEEE Transactions on Communications*, vol. 29, no. 3, pp. 237–249, March 1981.

[12] M. K. Simon and T.-Y. Yan, "Cross-correlated trellis coded quadrature modulation," patent filed October 5, 1999.

[13] M. K. Simon, P. Arabshahi, and M. Srinivasan, "Trellis-coded quadrature phase shift keying (QPSK) with variable overlapped raised-cosine pulse shaping," *Telecommunications and Mission Operations Progress Report 42-136*, vol. October–December 1998, February 15, 1999.
http://tmo.jpl.nasa.gov/progress_report/issues.html
Accessed March 2, 2001.

[14] M. J. Dapper and T. J. Hill, "SBPSK: A robust bandwidth-efficient modulation for hard-limited channels," *MILCOM Conference Record*, Los Angeles, California, pp. 31.6.1–31.6.6, October 21–24, 1984.

[15] T. J. Hill, "An enhanced, constant envelope, interoperable shaped offset QPSK (SOQPSK) waveform for improved spectral efficiency," *International Telemetering Conference*, San Diego, California, October 23–26, 2000. Also see "A non-proprietary, constant envelope, variant of shaped offset QPSK (SOQPSK) for improved spectral containment and detection efficiency," *MILCOM Conference Record*, vol. 1, Los Angeles, California, pp. 347–352, October 23–26, 2000.

[16] M. Geoghegan, "Implementation and performance results for trellis detection of SOQPSK," to be presented at *International Telemetering Conference 2001*, Las Vegas, Nevada, October 22–25, 2001.