EFQPSK versus CERN: A Comparative Study

by

Deva K. Borah and Stephen Horan

NMSU-ECE-01-014

UNCLASSIFIED

NMSU Technical Report No. NMSU-ECE-01-014

EFQPSK versus CERN: A Comparative Study

Deva K. Borah and Stephen Horan

New Mexico State University
Klipsch School of Electrical & Computer Engineering
Las Cruces, New Mexico 88003

Prepared for NASA
under Grant # NAG 5-9323

September 2001

UNCLASSIFIED

Abstract

This report presents a comparative study on Enhanced Feher's Quadrature Phase Shift Keying (EFQPSK) and Constrained Envelope Root Nyquist (CERN) techniques. These two techniques have been developed in recent times to provide high spectral and power efficiencies under nonlinear amplifier environment. The purpose of this study is to gain insights into these techniques and to help system planners and designers with an appropriate set of guidelines for using these techniques.

The comparative study presented in this report relies on effective simulation models and procedures. Therefore, a significant part of this report is devoted to understanding the mathematical and simulation models of the techniques and their set-up procedures. In particular, mathematical models of EFQPSK and CERN, effects of the sampling rate in discrete time signal representation, and modeling of nonlinear amplifiers and predistorters have been considered in detail.

The results of this study show that both EFQPSK and CERN signals provide spectrally efficient communications compared to filtered conventional linear modulation techniques when a nonlinear power amplifier is used. However, there are important differences. The spectral efficiency of CERN signals, with a small amount of input backoff, is significantly better than that of EFQPSK signals if the nonlinear amplifier is an ideal clipper. However, to achieve such spectral efficiencies with a practical nonlinear amplifier, CERN processing requires a predistorter which effectively translates the amplifier's characteristics close to those of an ideal clipper. Thus, the spectral performance of CERN signals strongly depends on the predistorter. EFQPSK signals, on the other hand, do not need such predistorters since their spectra are almost unaffected by the nonlinear amplifier.

This report discusses several receiver structures for EFQPSK signals. It is observed that optimal receiver structures can be realized for both coded and uncoded EFQPSK signals with not too much increase in computational complexity. When a nonlinear amplifier is used, the bit error rate (BER) performance of the CERN signals with a matched filter receiver is found to be more than one decibel (dB) worse compared to the bit error performance of EFQPSK signals. Although channel coding is found to provide BER performance improvement for both EFQPSK and CERN signals, the performance of EFQPSK signals remains better than that of CERN. Optimal receiver structures for CERN signals with nonlinear equalization is left as a possible future work.

Based on the numerical results, it is concluded that, in nonlinear channels, CERN processing leads towards better bandwidth efficiency with a compromise

in power efficiency. Hence for bandwidth efficient communications needs, CERN is a good solution provided effective adaptive predistorters can be realized. On the other hand, EFQPSK signals provide a good power efficient solution with a compromise in bandwidth efficiency.

Contents

List of Figures

Acknowledgements

The authors would like to thank Richard Reinhart, Gene Fujikawa, and Monty Andro of NASA Glenn Research Center, Cleveland, Ohio, for their constructive comments and valuable suggestions during the course of the work.

1 Summary

This report presents a comparative study on EFQPSK and CERN modulation techniques. Both these techniques have been developed in recent times to provide high spectral and power efficiencies under nonlinear amplifier environment. The EFQPSK achieves this by producing a nearly constant envelope signal by deliberately introducing correlation between the inphase and the quadrature channel components. On the other hand, the CERN technique employes conventional root Nyquist pulses and reduces the amplitude fluctuation by repetitively superimposing suitably scaled and delayed versions of the root Nyquist pulse.

It is observed in this report that in an additive white Gaussian (AWGN) channel, without a nonlinear amplifier, the spectral efficiency of CERN is significantly better than that of EFQPSK. The spectral efficiency of CERN in this case is exactly the same as that of a conventional root Nyquist filtered modulation technique. In addition, when unbalanced data is employed, CERN processing does not generate significant spectral lines (except a zero frequency line) while EFQPSK signals are observed to generate strong spectral lines. However, the spectral quality of CERN degrades significantly in the presence of a nonlinear amplifier that generates AM/AM and AM/PM effects, while the spectra of EFQPSK remain unaffected. The spectra of CERN, however, can be significantly improved by using a predistorter, also called a linearizer. Thus, the spectral performance of CERN greatly depends on the predistorter that has to track and compensate for the nonideal characteristics of a practical power amplifier.

The bit error rate (BER) performance has been studied for both EFQPSK and CERN signals. It is observed that optimal receiver structures can be easily realized for both coded and uncoded EFQPSK signals. When a nonlinear amplifier is used, the bit error rate performance of the CERN signals with a matched filter receiver is found to be more than one decibel (dB) worse compared to the bit error performance of EFQPSK signals. Although channel coding is found to provide BER performance improvement for both EFQPSK and CERN signals, EFQPSK's performance still remains better than that of CERN signals. The deterioration in CERN's BER performance is mainly due to two reasons: 1) CERN introduces a signal dependent noise, and 2) the matched filter receiver used in the study is suboptimal. Improved receiver structures with nonlinear equalization are complex structures, and they are beyond the scope of this report.

2 Introduction

Power and bandwidth efficiencies are two very important performance characteristics of any modulation technique. It is well known that conventional QPSK and OQPSK signals have good spectral effficiencies. However, their spectral efficiencies are seriously affected when such signals pass through a nonlinear amplifier that generates Amplitude Modulation to Amplitude Modulation (AM/AM) and Amplitude modulation to Phase Modulation (AM/PM) effects [1, 2]. A nonlinearly modulated signal, such as the Gaussian Minimum Shift Keying (GMSK), can safely pass through nonlinear amplifiers without spectrum regeneration and performance degradation. However, its spectrum efficiency is low. Therefore, a number of techniques have been developed in recent times to achieve better spectrum utilization, yet with little performance degradation under nonlinear amplification environment. Feher-Patented Quadrature Phase-Shift-Keying (FQPSK) and Constrained Envelope Root Nyquist (CERN) are two such modulation techniques.

The FQPSK technique is conceptually the same as the cross-correlated phase shift keying (XPSK) modulation presented in [3] where an intentional but controlled amount of cross correlation between the inphase (I) and quadrature (Q) channels is introduced. This gives rise to a nearly constant envelope signal making it suitable for nonlinear amplification. In [4], an enhanced version of the FQPSK (EFQPSK) signal is presented. This technique emphasizes on symbol by symbol representation of the cross correlation operation. As a result, instead of the crosscorrelator of the conventional FQPSK, the EFQPSK can be described directly in terms of data transitions on the I and Q channels. The EFQPSK also improves the smoothness of these waveforms, thus improving the power spectral density roll off. Further, the error probability performance has also been significantly improved by exploiting the correlation inherent in the modulation.

The CERN method is a properiety technique developed by SiCOM which is later acquired by Intersil. This technique tries to utilize the high spectral efficiency of the conventional Nyquist (or root Nyquist) filtered linear modulation schemes. The main idea is to reduce the envelope fluctuation of the signal by adding appropriately delayed and scaled versions of the same pulse waveform. This reduction in envelope fluctuations enables the system to decrease the backoff level of the nonlinear amplifier and, therefore, efficient amplification can be obtained without sacrificing the spectral quality. However, to be effective in a practical nonlinear amplifier environment, the CERN technique must employ a good linearizer (or predistorter) which should be able to track the amplifier characteristics.

In this report, we present a comparative study on EFQPSK and CERN signals in order to understand their relative advantages and disadvantages. The comparisons essentially focus on their spectral properties and the bit error rate performance. Specifically the following questions are addressed through simulation studies: (1) How do EFQPSK and CERN compare in terms of spectrum occupation, with and without a nonlinear amplifier? (2) Are their spectra significantly affected when unbalanced data are transmitted? (3)How does the bit error rate (BER) performance of EFQPSK compare with that of CERN? (4) Does channel coding affect the BER performance comparison? (5) What are the different receiver structures and how do they perform? (6) What are the main drawbacks of the two techniques?

This report is organized as follows. In the next section, a mathematical description of the EFQPSK and CERN signal generations is given, simulation setup procedures are described and receiver structures are discussed. This section also presents models for the nonlinear amplifier and the predistorter. Section 4 describes the numerical results and their implications. Section 5 summarizes the study and, finally, Section 6 makes recommendations based on the findings of this report.

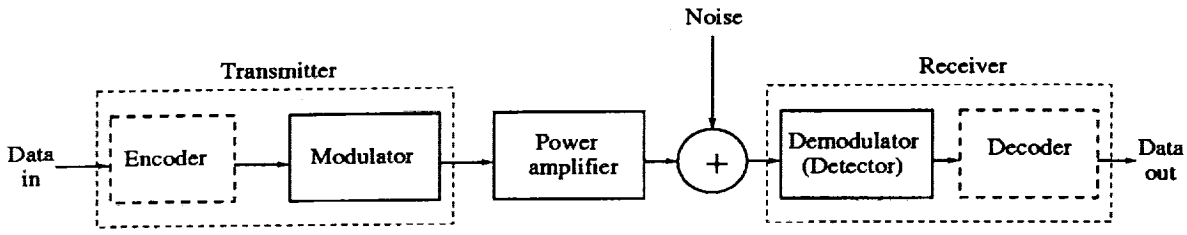


Figure 1: Communication System

3 Methods, Assumptions, and Procedures

This section describes the mathematical models of EFQPSK and CERN signals, their simulation set-up procedures, and receiver structures. Figure 1 shows the block diagram of the system under consideration. Binary data, encoded by a channel encoder (optional), are modulated using one of the following three modulation techniques: (1) conventional root-raised cosine filtered QPSK, (2) EFQPSK and (3) CERN with QPSK. The modulated signal then passes through a nonlinear amplifier (NLA) that generates AM/AM and AM/PM effects. The NLA, in general, produces considerable spectral spreading of the modulated signal, thus seriously degrading the bandwidth efficiency of the system. In order to reduce the effects of nonlinear distortion, the operating point of the NLA is 'backed off' from saturation. However, a large amount of backoff reduces the available output power, and thus degrades the bit error rate performance of the system. Hence a meaningful comparison of modulation schemes must include studies on both spectral qualities and bit error rate (BER) performance. Since the BER performance strongly depends on the receiver structure employed, several demodulator/decoder structures are also described in this section in order to prepare the groundwork for their simulation in the next section.

3.1 Enhanced Feher's QPSK (EFQPSK)

3.1.1 Signal Model

The enhanced Feher's Quadrature Phase Shift Keying (EFQPSK)[4] employs sixteen waveforms. Each waveform is one symbol interval long. In every symbol interval, a particular waveform is chosen for the I channel and another waveform is chosen for the Q channel. The selection of an I or a Q waveform depends on the most recent data transition on that channel as well as two most recent successive transitions on the other channel. The sixteen waveforms for EFQPSK are

1. $s_0(t) = A,\ -T/2 \le t \le T/2,\ \ s_8(t) = -s_0(t)$

2. $s_1(T) = \begin{cases} A, & -T/2 \le t \le 0 \\ 1-(1-A)\cos^2(\pi t/T), & 0 \le t \le T/2 \end{cases}$
   $s_9(t) = -s_1(t)$

3. $s_2(t) = \begin{cases} 1-(1-A)\cos^2(\pi t/T), & -T/2 \le t \le 0 \\ A, & 0 \le t \le T/2 \end{cases}$
   $s_{10}(t) = -s_2(t)$

4. $s_3(t) = 1-(1-A)\cos^2(\pi t/T),\ -T/2 \le t \le T/2,\ \ s_{11}(t) = -s_3(t)$

5. $s_4(t) = A\sin(\pi t/T),\ -T/2 \le t \le T/2,\ \ s_{12}(t) = -s_4(t)$

6. $s_5(t) = \begin{cases} \sin(\pi t/T)+(1-A)\sin^2(\pi t/T), & -T/2 \le t \le 0 \\ \sin(\pi t/T), & 0 \le t \le T/2 \end{cases}$
   $s_{13}(t) = -s_5(t)$

7. $s_6(t) = \begin{cases} \sin(\pi t/T), & -T/2 \le t \le 0 \\ \sin(\pi t/T)-(1-A)\sin^2(\pi t/T), & 0 \le t \le T/2 \end{cases}$
   $s_{14}(t) = -s_6(t)$

8. $s_7(t) = \sin(\pi t/T)\ \ -T/2 \le t \le T/2,\ \ s_{15}(t) = -s_7(t)$

The constant A is $1/\sqrt{2}$. The selection of the waveforms and the generation of the EFQPSK signal are well summarized in [4], as shown here via Tables 1 and 2, where $a_{I,n}$ represents I-channel bit and $a_{Q,n}$ represents Q-channel bit. Depending on the input bits as shown in the table, the I-channel waveform is $x_I(t) = a_{I,n}s_\mu(t-nT)$ during the n-th signaling interval $(n-1/2)T \le t \le (n+1/2)T$. Similarly, the Q-channel waveform is $x_Q(t-T/2) = a_{Q,n}s_\nu(t-nT)$ in the signaling interval $nT \le t \le (n+1)T$. Observe that I-channel and Q-channel waveforms are kept offset by an interval T/2. The waveform selection procedures can be mathematically described as

$$\mu = 4 \times |(a_{I,k} - a_{I,k-1})/2| + 2 \times |(a_{Q,k-1} - a_{Q,k-2})/2| + |(a_{Q,k} - a_{Q,k-1})/2| \tag{3.1}$$

$$\nu = 4 \times |(a_{Q,k} - a_{Q,k-1})/2| + 2 \times |(a_{I,k} - a_{I,k-1})/2| + |(a_{I,k+1} - a_{I,k})/2| \tag{3.2}$$

$\lvert(a_{I,n} - a_{I,n-1})/2)\rvert$	$\lvert(a_{Q,n-1} - a_{Q,n-2})/2\rvert$	$\lvert(a_{Q,n} - a_{Q,n-1})/2)\rvert$	$x_I(t)$
0	0	0	$a_{I,n}s_0(t-nT)$
0	0	1	$a_{I,n}s_1(t-nT)$
0	1	0	$a_{I,n}s_2(t-nT)$
0	1	1	$a_{I,n}s_3(t-nT)$
1	0	0	$a_{I,n}s_4(t-nT)$
1	0	1	$a_{I,n}s_5(t-nT)$
1	1	0	$a_{I,n}s_6(t-nT)$
1	1	1	$a_{I,n}s_7(t-nT)$

Table 1: Generation of inphase signal in the interval $(n-1/2)T \leq t \leq (n+1/2)T$.

$\lvert(a_{Q,n} - a_{Q,n-1})/2\rvert$	$\lvert(a_{I,n} - a_{I,n-1})/2\rvert$	$\lvert(a_{I,n+1} - a_{I,n})/2\rvert$	$x_Q(t)$
0	0	0	$a_{Q,n}s_0(t-nT)$
0	0	1	$a_{Q,n}s_1(t-nT)$
0	1	0	$a_{Q,n}s_2(t-nT)$
0	1	1	$a_{Q,n}s_3(t-nT)$
1	0	0	$a_{Q,n}s_4(t-nT)$
1	0	1	$a_{Q,n}s_5(t-nT)$
1	1	0	$a_{Q,n}s_6(t-nT)$
1	1	1	$a_{Q,n}s_7(t-nT)$

Table 2: Generation of quadrature phase signal in the interval $nT \leq t \leq (n+1)T$.

3.1.2 Simulation Set-up Procedures

Each waveform $s_i(t)$ is represented by $r = 20$ or $r = 22$ samples. Although each waveform is individually time-limited and hence not bandlimited, the overall signal becomes approximately bandlimited due to the continuity of the waveforms, and samples at the rate $r \geq 20$ per symbol duration provide a good quality representation of the signal. Let $\mathbf{s}_i$ be the vector containing the r samples of the waveform $s_i(t)$. Considering both the I and Q channels, the average symbol energy over all the sixteen waveforms is

$$E_s = \frac{1}{8}\sum_{i=0}^{15} \mathbf{s}_i^T \mathbf{s}_i = \frac{1}{4}\sum_{i=0}^{7} \mathbf{s}_i^T \mathbf{s}_i \tag{3.3}$$

where $(\cdot)^T$ denotes the transpose of a matrix or a vector. We normalize each waveform by a constant k so that the normalized waveform satisfies the following relation.

$$\frac{1}{4}\sum_{i=0}^{7}(k\mathbf{s}_i)^T(k\mathbf{s}_i) = \frac{r}{2} \Rightarrow k = \sqrt{\frac{r}{E_s}} \tag{3.4}$$

This normalization is called energy scaling. As a result of energy scaling, the average energy over a symbol period (considering both I and Q channels) becomes r. The variance of the real (or imaginary) white Gaussian noise σ_n is calculated as

$$\sigma_n = \sqrt{\frac{r}{4 \times 10^{\frac{\text{SNR}}{10}}}} \tag{3.5}$$

where SNR is the signal-to-noise ratio (SNR) in dB. The total noise variance (considering both real and imaginary components) is $2\sigma_n^2$. However, it is to be noted that noise does not have any role in our spectrum study, and it is used in the bit error rate (BER) performance study only.

Before the EFQPSK signal is passed through the TWT, the signal has to be appropriately scaled to meet the backoff specifications for the TWT. This is called backoff scaling. This is done as follows. The input backoff (IBO) in dB is defined as

$$\text{IBO} = 10\log(\frac{P_{\text{in}}^{\text{sat}}}{\overline{P}_{\text{in}}}) \tag{3.6}$$

where $P_{\text{in}}^{\text{sat}}$ is the input saturation power when the output power begins to saturate, and $\overline{P}_{\text{in}}$ represents average input power. The complex signal (that is, including

both I and Q channels), after energy scaling, is subjected to backoff scaling by multiplying with a constant γ given by

$$\gamma = \frac{1}{\sqrt{10^{\frac{IBO}{10}}}} \tag{3.7}$$

As an example, for 0 dB backoff, the scaling factor γ becomes 1. Note that the constant γ is not considered in the noise variance calculation expression (3.5).

3.1.3 Receiver Structures

The optimal receiver structure and an averaged MF receiver structure have been studied in [5] for uncoded EFQPSK. In this section, we categorize the receiver structures into optimal and suboptimal structures and discuss them for both coded and uncoded EFQPSK. Moreover, a linear optimal filter receiver and a per-waveform matched filter (PWMF) structure are proposed and their complexity requirements, bit error rate (BER) performance and adaptive versions are discussed.

3.2 Optimal structure

The optimal receiver structure uses the Viterbi algorithm (VA) with 16 states. Each state $(a_{I,k-1},\ a_{I,k},\ a_{Q,k-2},\ a_{Q,k-1})$ represents 4 bits consisting of two I channel bits, $a_{I,k-1}$, $a_{I,k}$, and two Q channel bits, $a_{Q,k-2}$, $a_{Q,k-1}$. The transition from state $\sigma_l = (a_{I,k-2}, a_{I,k-1}, a_{Q,k-3}, a_{Q,k-2})$ to state $\sigma_p = (a_{I,k-1}, a_{I,k}, a_{Q,k-2}, a_{Q,k-1})$ is associated with the branch metric

$$\lambda(\sigma_l, \sigma_p, k) = ||\mathbf{y}_{I,k} - a_{I,k}\mathbf{s}_\mu||^2 + ||\mathbf{y}_{Q,k} - a_{Q,k}\mathbf{s}_\nu||^2$$

where the waveform identification numbers μ and ν are obtained from (3.1) and (3.2) respectively. The vector $\mathbf{y}_{I,k} = a_{I,k}\mathbf{s}_i + \mathbf{n}_I$, where the vector $\mathbf{s}_i$ consists of samples of $s_i(t)$, and $\mathbf{n}_I$ is the inphase noise sample vector. The branch metric can also be alternatively obtained as

$$\lambda'(\sigma_l, \sigma_p, k) = -2a_{I,k}\mathrm{Re}\{\mathbf{s}_\mu^H \mathbf{y}_{I,k}\} + \| \mathbf{s}_\mu \|^2 - 2a_{Q,k}\mathrm{Re}\{\mathbf{s}_\nu^H \mathbf{y}_{Q,k}\} + \| \mathbf{s}_\nu \|^2 \tag{3.8}$$

The complexity of this receiver in terms of the number of multiplications is of the order of $64r$ per detected bit.

When a convolutional code of constraint length K is used, a joint detector and decoder is implemented with 2^{K+2} states. In this case, each state is associated

with 6 encoder output channel bits. A state transition provides 2 additional bits. Therefore, each state transition provides 8 channel bits. These bits determine the waveform numbers (3.1) and (3.2), which are used in the calculation of the branch metric given by (3.8).

3.3 Suboptimal structures

3.3.1 Single Filter Receiver

Let $\mathbf{w}$ be a linear filter of r taps. Consider the detection of the I channel bit using this filter. Let the input signal vector to the filter while detecting the k-th bit be denoted by $\mathbf{y}_{I,k}$. This vector is of the form

$$\mathbf{y}_{I,k} = a_{I,k}\mathbf{s}_i + \mathbf{n}_I$$

where the vector $\mathbf{s}_i$ consists of samples of $s_i(t)$. The mean-squared error (MSE) at the output of the filter is $\mathcal{E} = E\{|\mathbf{w}^T(a_{I,k}\mathbf{s}_i + \mathbf{n}_I) - a_{I,k}|^2\}$. Taking the derivative of $\mathcal{E}$ with respect to $\mathbf{w}$ and setting it to zero, the optimal MMSE filter is obtained as

$$\mathbf{w}_{opt} = \mathbf{R}^{-1}\bar{\mathbf{s}} \tag{3.9}$$

where $\mathbf{R} = \frac{1}{8}\sum_{i=0}^{7}\mathbf{s}_i\mathbf{s}_i^T + \sigma_n^2\mathbf{I}$, $E\{\mathbf{n}_I\mathbf{n}_I^T\} = \sigma_n^2\mathbf{I}$, and $\bar{\mathbf{s}} = \frac{1}{8}\sum_{i=0}^{7}\mathbf{s}_i$. The same optimal filter is used for detecting both the I channel and the Q channel symbols $a_{I,k}$ and $a_{Q,k}$ using the input vector $\mathbf{y}_{I,k}$ and $\mathbf{y}_{Q,k}$ respectively.

At very low SNR, $\mathbf{R} \approx \sigma_n^2\mathbf{I}$, and then, $\mathbf{w}_{opt} \approx \sigma_n^{-2}\bar{\mathbf{s}}$ which is the averaged matched filter receiver discussed in [5]. It is to be noted that a direct realization of the optimal linear receiver requires knowledge of the noise variance. However, an adaptive implementation, discussed later in this section, avoids this problem. The implementation of the linear receiver requires rN multiplications per bit, where rN is the number of filter taps ($N = 1$ for a filter spanning exactly over one bit duration).

The BER performance analysis for the single filter receiver is similar to that of the averaged matched filter receiver presented in [5]. The bit error probability is

$$P_b = \frac{1}{8}\sum_{i=0}^{7}\frac{1}{2}\mathrm{erfc}\left(\frac{\alpha_i}{\sigma_n\sqrt{2\gamma}}\right)$$

where $\alpha_i = \mathbf{w}_{opt}^T\mathbf{s}_i$, σ_n is real (or imaginary) noise variance, and $\gamma = \mathbf{w}_{opt}^T\mathbf{w}_{opt}$.

3.3.2 Per-Waveform Matched Filter (PWMF) Receiver

The PWMF receiver consists of 8 parallel matched filters, each corresponding to one of the waveforms $\mathbf{s}_i, 0 \leq i \leq 7$. During each symbol interval, a pair of matched filters is chosen. We observe that only the following pairs of I and Q channel waveforms are allowed in EFQPSK: $(\mathbf{s}_0, \mathbf{s}_0)$, $(\mathbf{s}_0, \mathbf{s}_1)$, $(\mathbf{s}_1, \mathbf{s}_4)$, $(\mathbf{s}_1, \mathbf{s}_5)$, $(\mathbf{s}_2, \mathbf{s}_0)$, $(\mathbf{s}_2, \mathbf{s}_1)$, $(\mathbf{s}_3, \mathbf{s}_4)$, $(\mathbf{s}_3, \mathbf{s}_5)$, $(\mathbf{s}_4, \mathbf{s}_2)$, $(\mathbf{s}_4, \mathbf{s}_3)$, $(\mathbf{s}_5, \mathbf{s}_6)$, $(\mathbf{s}_5, \mathbf{s}_7)$, $(\mathbf{s}_6, \mathbf{s}_2)$, $(\mathbf{s}_6, \mathbf{s}_3)$, $(\mathbf{s}_7, \mathbf{s}_6)$, $(\mathbf{s}_7, \mathbf{s}_7)$. Let us denote the set of all permissible waveform index pairs as F. Then the k-th I and Q channel bits are detected as

$$\begin{aligned}(\hat{a}_{I,k}, \hat{a}_{Q,k}) &= \arg \min_{\tilde{a}_{I,k}, \tilde{a}_{Q,k}} \min_{(m,n)\in F} (\| \mathbf{y}_{I,k} - \tilde{a}_{I,k}\mathbf{s}_m \|^2 + \| \mathbf{y}_{Q,k} - \tilde{a}_{Q,k}\mathbf{s}_n \|^2) \\ &= \arg \min_{\tilde{a}_{I,k}, \tilde{a}_{Q,k}} \min_{(m,n)\in F} [-2\tilde{a}_{I,k}\mathrm{Re}\{\mathbf{s}_m^H \mathbf{y}_{I,k}\} + \| \mathbf{s}_m \|^2 \\ &\quad -2\tilde{a}_{Q,k}\mathrm{Re}\{\mathbf{s}_n^H \mathbf{y}_{Q,k}\} + \| \mathbf{s}_n \|^2] \qquad (3.10)\end{aligned}$$

This receiver is called PWMF, because the matched filter is implemented for all permissible waveform pairs, and the one with the lowest metric is taken for detection. The complexity in terms of the number of multiplications is eight times the complexity of the single filter receiver.

The coded EFQPSK is detected with the PWMF receiver using the VA. The decoder VA has 2^{K-1} states. The branch metric is calculated as

$$\lambda(\sigma_l^{(c)}, \sigma_p^{(c)}, k) = \min_{m,n} (\| \mathbf{y}_{I,k} - \tilde{a}_{I,k}\mathbf{s}_m \|^2 + \| \mathbf{y}_{Q,k} - \tilde{a}_{Q,k}\mathbf{s}_n \|^2)$$

where $\sigma_l^{(c)}$ and $\sigma_p^{(c)}$ denote decoder states in the trellis diagram and the transition from $\sigma_l^{(c)}$ to $\sigma_p^{(c)}$ produces the symbol pair $(\tilde{a}_{I,k}, \tilde{a}_{Q,k})$.

An approximate BER analysis for the PWMF receiver can be given as follows. Let $\mathbf{s}_m$ be the waveform for the k-th I channel bit. An upper bound for the probability of error is $\sum_{i=0}^{7}(1/2)\mathrm{erfc}((d_i/2)/\sqrt{2}\sigma_n)$, where $d_i = \sqrt{(\mathbf{s}_m + \mathbf{s}_i)^T(\mathbf{s}_m + \mathbf{s}_i)}$. Averaging over all waveforms $\mathbf{s}_m$, the upper bound is $(1/8)\sum_{m=0}^{7}\sum_{i=0}^{7}(1/2)\mathrm{erfc}(\ (d_i/2)/\sqrt{2}\sigma_n)$. An approximate BER is obtained by using $(1/8)\sum_{m=0}^{7}(1/2)$ $\mathrm{erfc}((\ \min_i d_i/2)/\sqrt{2}\sigma_n)$. This approximate BER is plotted in our results and is shown to agree well with simulation results.

3.3.3 Adaptive Receiver Structures

In order to avoid the direct estimation of the noise variance or exact symbol synchronization, adaptive implementation of the single filter receiver and the PWMF

receiver is realized using the LMS algorithm [6]. The filter (or the set of filters in the case of the PWMF) is updated during a training period. After the filter (or the set of filters) has converged, it is used for data detection. In the case of the single filter receiver, the tap update equation is

$$\mathbf{w}(n+1) = \mathbf{w}(n) - \mu e(n)\mathbf{y}_{I,n}$$

where $\mathbf{w}(n)$ is the filter tap vector at time n, μ is the step size of the algorithm, and $e(n) = a_{I,n} - \mathbf{w}^T(n)\mathbf{y}_{I,n}$. For the Q-channel symbol, $\mathbf{y}_{I,n}$ is replaced by $\mathbf{y}_{Q,n}$, and $e(n)$ is replaced by $e(n) = a_{Q,n} - \mathbf{w}^T(n)\mathbf{y}_{Q,n}$. In other words, both the I and Q channel bits are used for adaptation.

The PWMF receiver is adaptively realized using 8 parallel filters. The training symbols provide the waveform numbers m and n. Therefore, each training bit (I and Q) is used to update the particular filter depending on the values of m and n. The other filters remain unaffected for that bit period. In studying the convergence of the MSE, we use the same squared error of the unaffected filters in the next symbol interval and calculate the average over all the filters.

3.4 Constrained Envelope Root Nyquist (CERN)

3.4.1 Signal Generation

The CERN signal generation, as presented in the patent [7], requires three main components: (1) symbol filtering, (2) on-time constrained envelope generation, and (3) off-time constrained envelope generation. The output signal samples from these three components are combined to generate the CERN signal as demonstrated in Fig. 2.

Symbol filtering

The symbol filtering component generates a baseband equivalent linearly modulated signal $s_o(t)$ as

$$s_o(t) = \sum_{i=-N}^{N} a_i p(t - iT) \tag{3.11}$$

where $\{a_i\}$ is the sequence of M-ary symbols, $p(t)$ is a pulse-spreading filter, T is the symbol period, and $(2N+1)$ is the total number of symbols transmitted. Each M-ary symbol is obtained by mapping a k-bit block, where k is a positive integer, into one of $M = 2^k$ signal points in the appropriate signal space diagram. Figure 3 shows signal space diagrams for 8-PSK and 16-P-APSK signals. The transmit

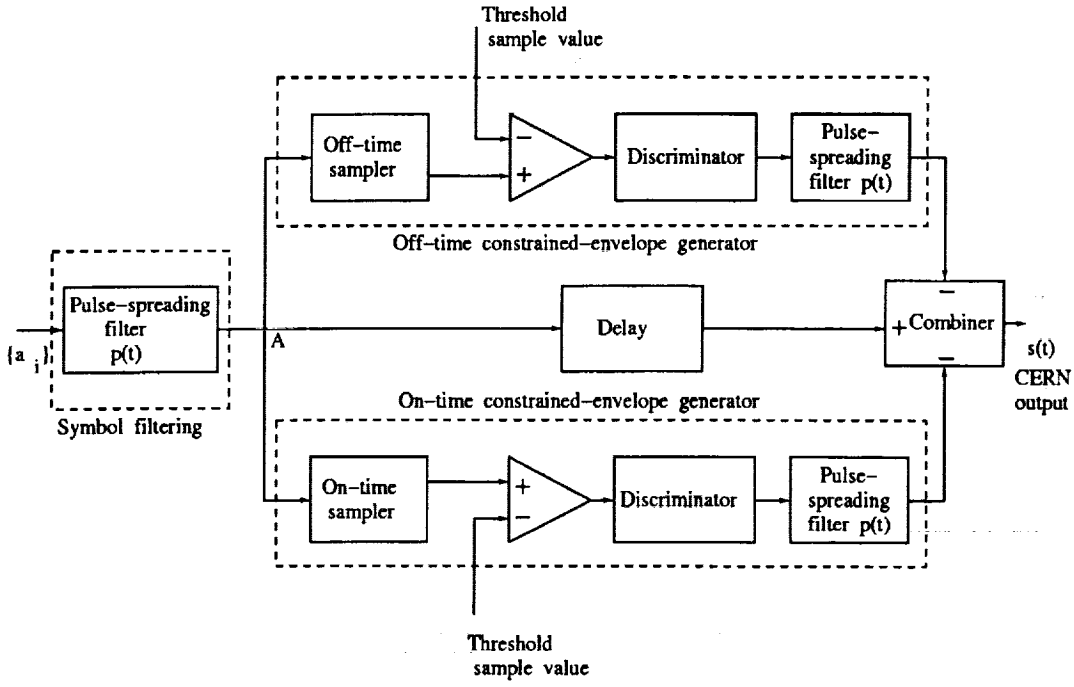
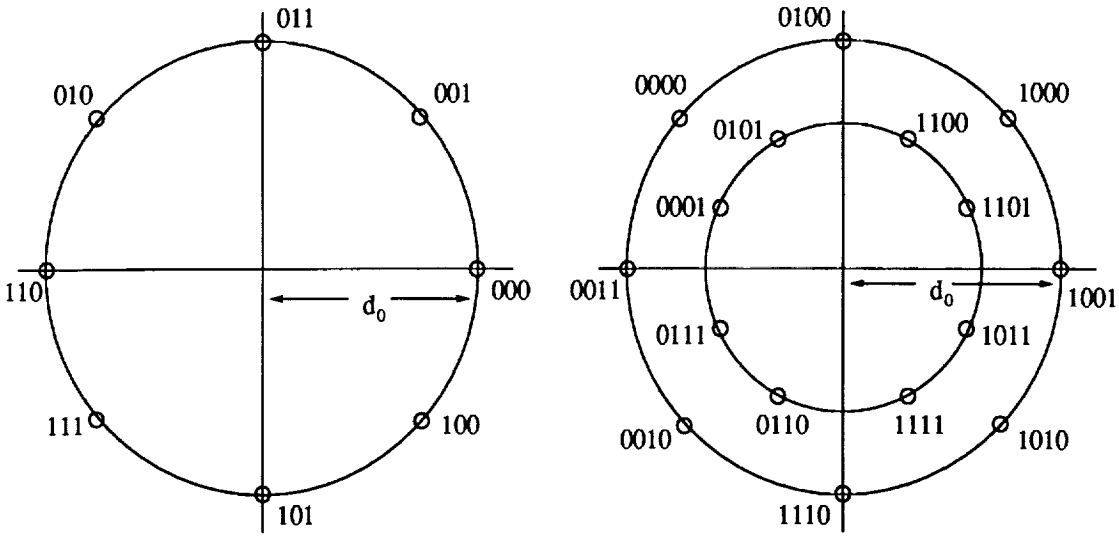


Figure 2: Block diagram showing CERN generation

filter, in our study, is a root raised cosine (RRC) filter with an excess bandwidth α, $0 \leq \alpha \leq 1$. Therefore, the bandwidth of the pulse-spreading transmit filter is limited to $|f| \leq (1+\alpha)/2T$. Thus, discrete time samples at the Nyquist rate and above can adequately describe the signal $s_o(t)$. In this case, we consider a sampling rate of $r = 2$ samples per symbol interval. These fractional rate (higher than symbol rate) samples of (3.11) at time instants $t = 0T, 0.5T, T, 1.5T, 2T, \cdots$ are

$$\{s_o(0T), s_o(0.5T), s_o(T), s_o(1.5T), s_o(2T), s_o(2.5T), s_o(3T), \cdots\} \qquad (3.12)$$

The signal samples in (3.12) can be divided into two groups: (1) on-time samples and (2) off-time samples. The on-time samples are the samples at time instants $t = iT$, where i is an integer, $-N \leq i \leq N$. The remaining samples, i.e., samples at time instants $t = (i+0.5)T$ are the off-time samples. As shown in Fig. 2, all signal samples at the output of the symbol filtering unit pass through a delay unit on way to the combiner. At the same time, all the on-time samples at the output of the symbol filtering unit also pass into the on-time constrained envelope generator. Similarly, all the off-time samples pass into the off-time constrained envelope



(a) 8–PSK

(b) 16–P–APSK

Figure 3: Signal space diagrams

generator.

On-time constrained envelope generator

The i-th on-time sample in (3.12) is

$$s_o(iT) = \cdots + a_{i-1}p(T) + a_i p(0) + a_{i+1}p(-T) + \cdots \tag{3.13}$$

Since $p(0)$ is the largest magnitude of the filter response samples (point E in Fig. 4), the most significant contribution to $s_o(iT)$ comes from the term $a_i p(0)$. The remaining terms in (3.13) are individually small since the corresponding filter response samples, e.g., points A, C, G, I etc. in Fig.4, have small magnitudes. However, depending on the transmitted symbols, these samples may all add up in correct phase to produce a large magnitude sample $s_o(iT)$. For example, in the case of an RRC filter with $\alpha = 0.2$, the maximum magnitude of this sample can be as high as 1.4 times the magnitude of $a_i p(0)$. Thus, to faithfully transmit the symbol a_i, the transmitter circuit would require an output power of 1.96 $(= 1.4^2)$ times the power required to transmit the symbol a_i in the absence of intersymbol interference (ISI) due to the pulse-spreading filter. To reduce this increased power

requirement, an on-time constrained generator is used in the CERN technique. This is shown in Fig. 2. Each on-time complex sample, $s_o(iT)$, from the output of the pulse-spreading filter at point A is fed to the on-time constrained-envelope generator where the magnitude of the sample, $|s_o(iT)|$, is compared to the threshold value $d_0p(0)$. If $|s_o(iT)| \leq d_0p(0)$ then the discriminator output is made zero. However, if the on-time signal magnitude $|s_o(iT)|$ exceeds the threshold $d_0p(0)$, then the discriminator output generates an error signal sample e_i as

$$e_i = s_o(iT)\left(\frac{|s_o(iT)| - d_0p(0)}{|s_o(iT)|p(0)}\right) \tag{3.14}$$

This complex error sequence is passed through the pulse-spreading filter $p(t)$ to generate the on-time constrained bandwidth error signal

$$e_{on}(t) = \sum_i e_i p(t - iT) \tag{3.15}$$

A combiner combines this error signal with the signal $s_o(t)$ along with an off-time error signal as shown in Fig.2. Let us first ignore the off-time error signal. The combination of $s_o(t)$ and $e_{on}(t)$ results in $s'(t)$, where $s'(t) = s_o(t) - e_{on}(t)$. We observe that the i-th on-time sample of this combined signal $s'(t)$ now becomes

$$\begin{aligned} s'(iT) &= s_o(iT) - s_o(iT)\left(\frac{|s_o(iT)| - d_0p(0)}{|s_o(iT)|}\right) \\ &= s_o(iT)\left(\frac{d_0p(0)}{|s_o(iT)|}\right) \end{aligned} \tag{3.16}$$

Thus the magnitude of the on-time sample of the combined signal $s'(iT)$ is reduced to $d_0p(0)$. Finally, we note that for certain types of pulse-spreading filter $p(t)$, the on-time constrained envelope generator is not needed. For example, if $p(t)$ is a raised cosine filter, then the i-th on-time sample is $a_ip(0)$ which is free from intersymbol interference and CERN generation in that case does not require an on-time constrained envelope generator.

Off-time constrained envelope generator

The i-th off-time sample at the output of the symbol filter is

$$s_o((i + 0.5)T) = \cdots + a_{i-1}p(1.5T) + a_ip(0.5T) + a_{i+1}p(-0.5T) + \cdots \tag{3.17}$$

Figure 4 shows that the pulse samples $p(0.5T)$ and $p(-0.5T)$ are equally large. Thus, the off-time samples cannot be ISI free. Depending on the transmitted symbols, the off-time sample $s_o((i + 0.5)T)$ given by (3.17) can be as large as 1.9

times the magnitude of the peak on-time information carrying phase point, i.e, $d_0 p(0)$ for an RRC filter with $\alpha = 0.2$. Hence the output power requirement for this sample is 3.6 times the power required to transmit the symbol a_i as an on-time sample in the absence of the ISI. Therefore, in order to reduce this inefficient use of power, an off-time constrained envelope generator is used. The principle of the off-time constrained envelope generator is exactly same as the on-time generator. The off-time error samples are obtained as

$$e_{i+0.5} = s_o((i+0.5)T)\left(\frac{|s_o((i+0.5)T)| - d_0 p(0)}{|s_o((i+0.5)T)|p(0)}\right) \tag{3.18}$$

This error sequence generates the constrained bandwidth error signal

$$e_{off}(t) = \sum_i e_{i+0.5} p(t-(i+0.5)T) \tag{3.19}$$

The combiner output due to symbol filtering, on-time generator and off-time generator gives the final CERN signal as

$$\begin{aligned} s(t) &= s_o(t) - e_{on}(t) - e_{off}(t) \\ &= \sum_i a_i p(t-iT) - \sum_i e_i p(t-iT) - \sum_i e_{i+0.5} p(t-(i+0.5)T) \end{aligned} \tag{3.20}$$

Note that the delay unit in the figure ensures that the symbol filtered signal $s_o(t)$ combines with the off-time/on-time generator in correct time.

3.4.2 Simulation Set-up Procedures

The generation of CERN signal is already described in Section 3.4. One problem with CERN generation is that the error signals (off time and on time) are not known beforehand. Therefore, the generated CERN signal cannot be scaled by a predetermined analytic constant. Generating all possible error values and averaging over them is time consuming when the length of the filter is long. In our simulation, we calculate the average value of the squared magnitudes of CERN samples by generating all the samples once and then rescaling the samples so that the average magnitude becomes unity. In practice, it is possible to group data into data frames and to scale each data frame separately for power normalization, albeit with additional frame delay.

The simulation of CERN signal processing is done by representing a truncated RRC filter with r equally spaced samples per symbol interval. This discrete time

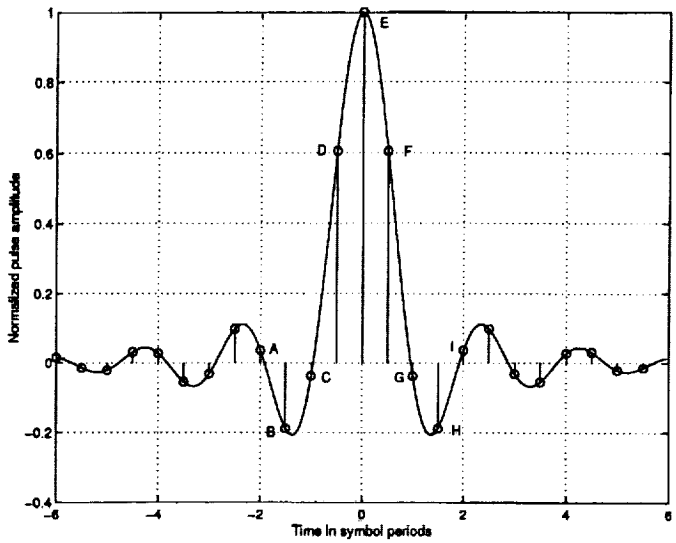


Figure 4: The root raised cosine filter response for $\alpha = 0.15$

filter is used for symbol filtering. The off-time/on-time samples of the filtered symbols are processed for constraining the envelope of the signal as shown in Fig. 2. The combined signal samples are passed through a predistorter (linearizer), to be described in Section 3.6, and then through a nonlinear amplifier (NLA).

3.4.3 Receiver Structures

The receiver used for CERN signals is a matched filter (MF) receiver. Although improved detection using nonlinear equalization can be obtained, the resulting receiver structures become more complicated and are not considered in this report. The MF consists of rL taps, where rL is also the number of transmit filter taps. In our simulation, after the i-th symbol has been generated and filtered through the RRC filter, the MF $\mathbf{p}$ considers the received sample vector $\mathbf{y} = [y((i-2L-2)r+1), \cdots, y((i-L-2)r)]$, and produces the soft output $\tilde{a}_\xi = \mathbf{p}^T\mathbf{y}$, where $\xi = i - L - L/2 - 1$, and L is an even positive integer. For uncoded transmission, these soft estimates are converted to hard decisions $\{\hat{a}_\xi\}$. When coding is used, the soft decisions are directly used in the decoder algorithm.

3.5 Nonlinear Amplifier Modeling

The TWT nonlinear amplifier model of [8] is used in the study. To understand the effects of the nonlinear amplifier, let us consider an input signal

$$x(t) = m(t)\cos(\omega_c t + \psi(t)) \tag{3.21}$$

where ω_c is the carrier frequency, and $m(t)$ and $\psi(t)$ are the modulated envelope and phase, respectively. The corresponding TWT output is

$$y(t) = A[m(t)]\cos(\omega_c t + \psi(t) + \Phi[m(t)]) \tag{3.22}$$

where $A[m(t)]$ is an odd function of $m(t)$ representing AM/AM conversion, and $\Phi[m(t)]$ is an even function of $m(t)$, representing AM/PM conversion. The following equations are used to represent $A(m)$ and $\Phi(m)$ in our study.

$$A(m) = \frac{\alpha_a m}{1 + \beta_a m^2} \tag{3.23}$$

$$\Phi(m) = \frac{\alpha_\phi m^2}{1 + \beta_\phi m^2} \tag{3.24}$$

In our simulation study, we use $\alpha_a = 1.9638$, $\alpha_\phi = 2.5293$, $\beta_a = 0.9945$, and $\beta_\phi = 2.8168$. These values are directly taken from [8]. Accordingly, the AM/AM conversion and the AM/PM conversion curves are obtained as shown in Figs. 5 and 6 respectively.

3.6 Predistorter for CERN

It will be observed in our results section that CERN processing alone cannot deliver its full potential spectral benefits without a linearizer (predistorter). The linearizer must be adaptive in order to effectively compensate for the NLA characteristics, which will change with temperature, aging etc. However, derivation of adaptive linearizer structures is beyond the scope of this work. Only a characterization of CERN's performance with a linearizer is required. Therefore, the fixed linearizer structure of [9] is used in our study and it is described in Appendix A. The overall effects on AM/AM and AM/PM characteristics due to the predistorter are shown in Figs. 7 and 8 respectively.

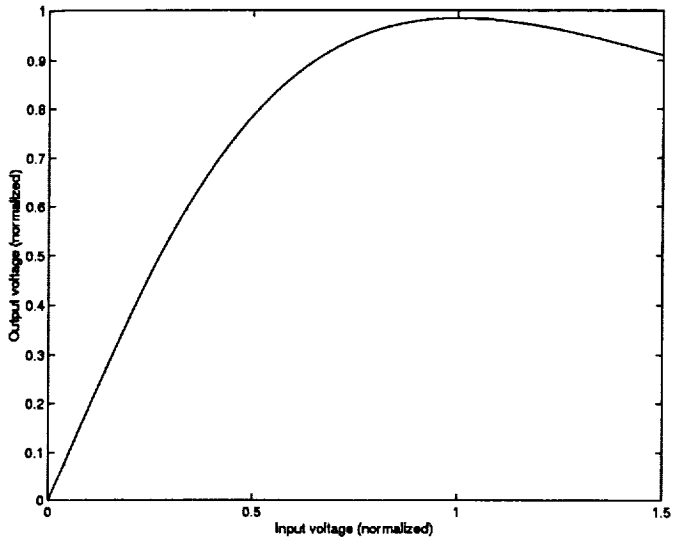


Figure 5: AM/AM characteristics of a commercial TWT

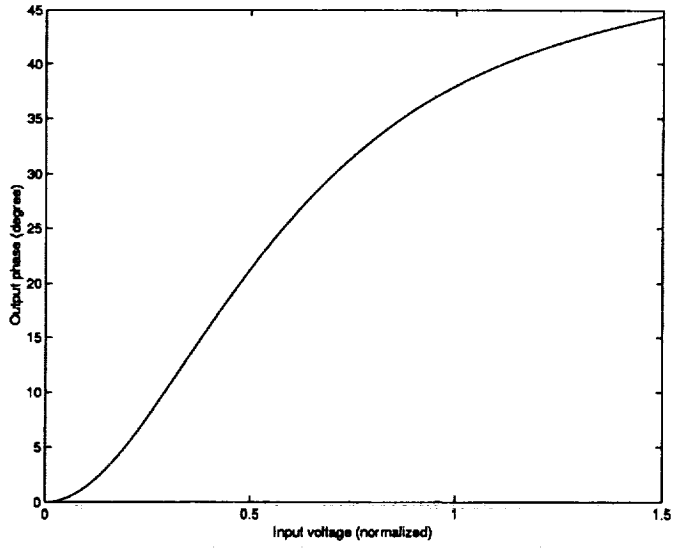


Figure 6: AM/PM characteristics of a commercial TWT

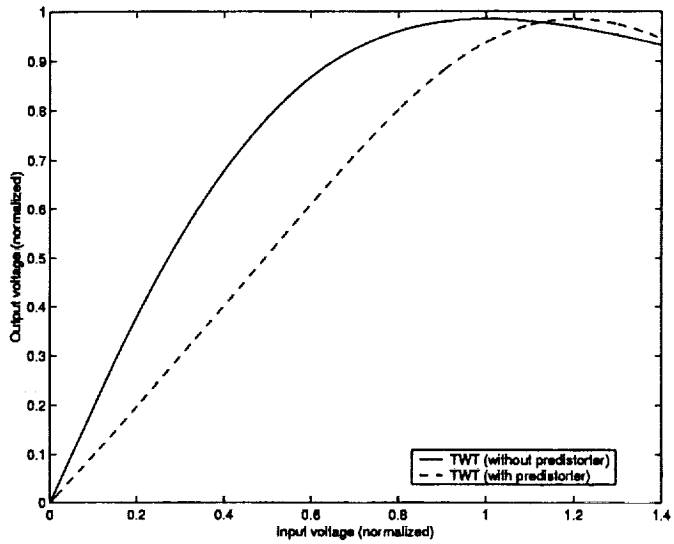


Figure 7: Effect of the predistorter on AM/AM Characteristics

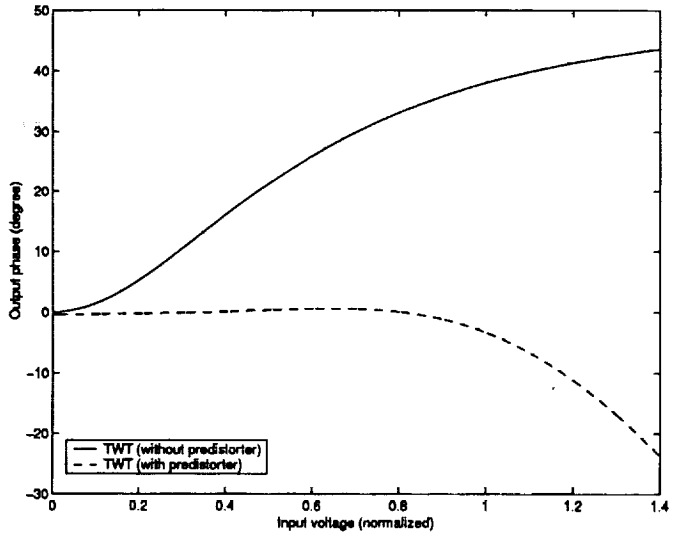


Figure 8: Effect of the predistorter on AM/PM Characteristics

4 Results and Discussion

4.1 EFQPSK and CERN Signal Shapes

The complex EFQPSK signal plotted in Fig. 9 shows that the signal is nearly constant envelope. Further the TWT has negligible effects on the signal. In contrast, the CERN signal envelope fluctuates much more. This is shown via a plot of the magnitude probability density functions (pdfs) in Fig. 10 for conventional 8 PSK and 8 PSK with CERN. It is observed that the envelope magnitudes vary over a wide range with high probabilities. However, the conventional 8 PSK signal has signal magnitudes that approach close to 2.0, while the CERN signal magnitudes abruptly stop at about 1.4. The probability distribution functions (PDFs) are shown in Fig. 11, while the portion of the curves near the PDF value of 100% is shown in Fig. 12. It is observed that for less than 0.01% clipping, the clipping levels for conventional 8 PSK and CERN are 1.91 and 1.36 respectively. So the backoff improvement for CERN processing is about $20\log_{10}(1.91/1.36) = 2.95$ dB. This approximately agrees with the improvement of 3.3 dB found in [10].

The magnitude pdf curve for conventional 8 PSK presented in Fig. 9 of [10] shows two peaks, while we have found only one peak in our study. Our one peak result also agrees with the histogram plot on p. 84 in [11].

4.2 EFQPSK and CERN Spectra

The spectrum of the EFQPSK signal is shown in Fig. 13. The spectrum agrees very well with the results reported in [4]. We observe that even when the TWT is operated at 0 dB backoff, the effects on the spectrum are negligible. Further the spectra remain well within the Space Frequency Coordination Group (SFCG) mask even for data rates higher than 2 Mbps.

The behavior of the EFQPSK spectra for unbalanced data is, however, different from the above results shown for balanced data (data bits with equal probabilities). Figure 14 shows the EFQPSK spectrum for unbalanced data when a +1 occurs with probability 0.4 and a -1 occurs with probability 0.6. Strong spectral lines at frequencies 0 and $1/2T_b$ are observed. Assuming that the DC line at zero frequency is eliminated, the line at frequency $1/2T_b$ still exceeds the SFCG mask. The strengths of the spectral lines with respect to the peak spectrum level (at zero frequency) for balanced data are given in Table 3.

The spectra of conventional QPSK signals are shown in Figs. 15 and 16. These figures show that for conventional QPSK there is severe spectral regrowth when

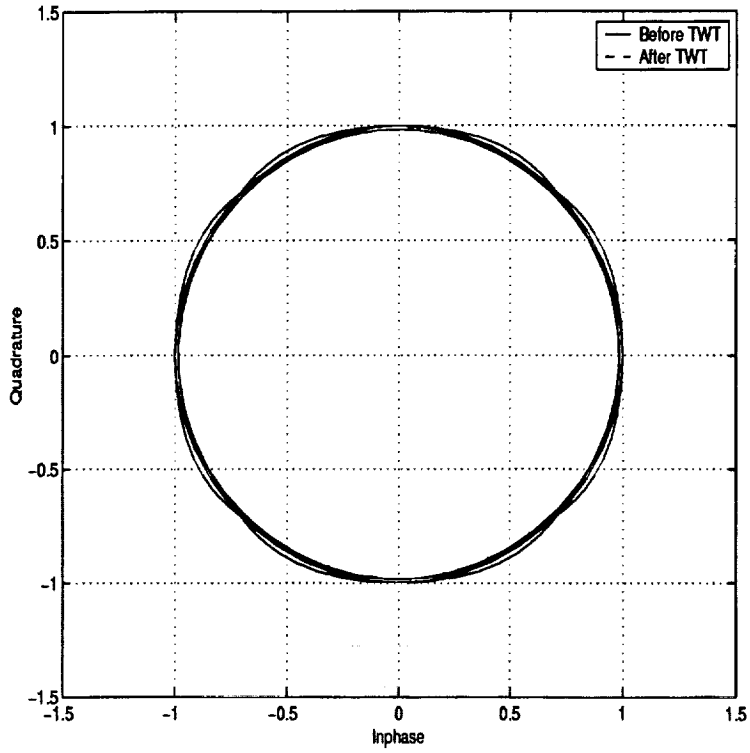


Figure 9: Vector plot of EFQPSK. The TWT operates at 0 dB backoff

Probability of a bit being a +1	Relative strength at zero frequency (dB)	Relative strength at $f = 1/2T_b$ (dB)
0.45	+16.0	-10.0
0.40	+21.0	-4.0
0.35	+25.0	-0.5
0.30	+30.0	0.0
0.25	+30.0	+0.5
0.15	+38.0	+1.5
0.10	+39.0	+2.5
0.05	+41.0	+1.0

Table 3: Strengths of the lines when unbalanced data are used. The relative strength measures strengths with respect to the spectrum level at zero frequency for balanced data.

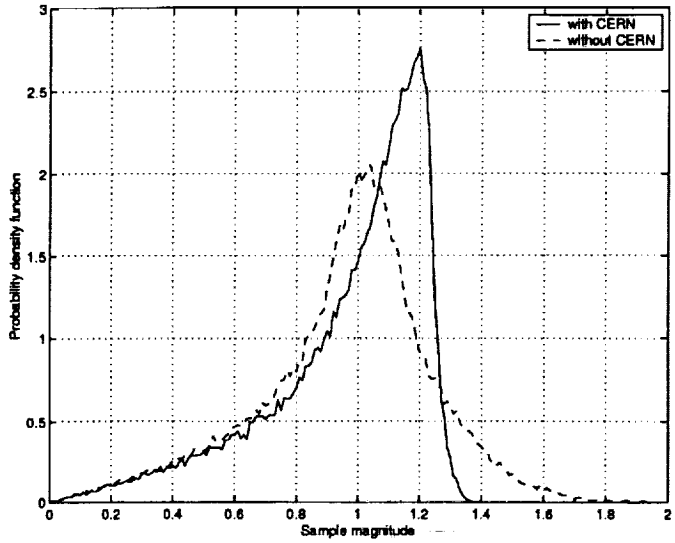


Figure 10: Magnitude probability density functions for 8-PSK, sampling rate, $r = 22$, length of the RRC filter, 60 symbol periods, roll-off factor=0.12.

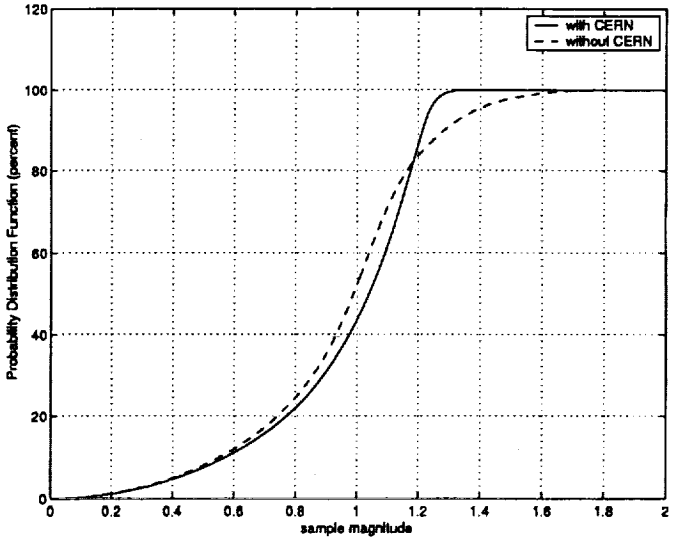


Figure 11: Magnitude probability distribution functions for 8-PSK, sampling rate, $r = 22$, length of the RRC filter, 60 symbol periods, roll-off factor=0.12.

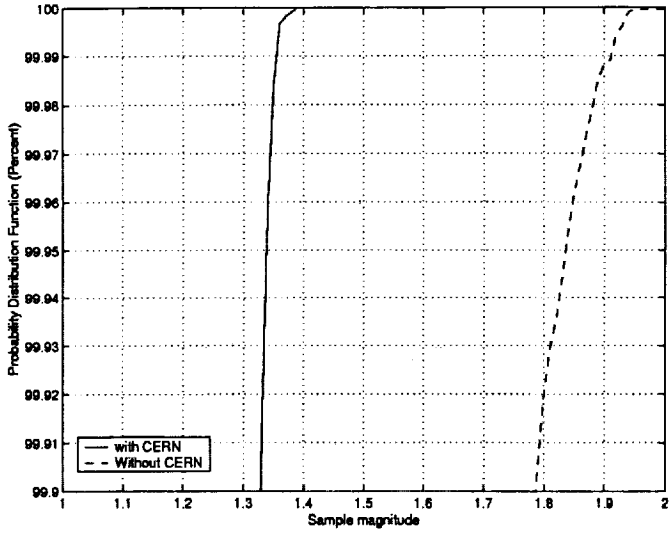


Figure 12: Magnitude probability distribution functions for 8-PSK, sampling rate, $r = 22$, length of the RRC filter, 60 symbol periods, roll-off factor=0.12.

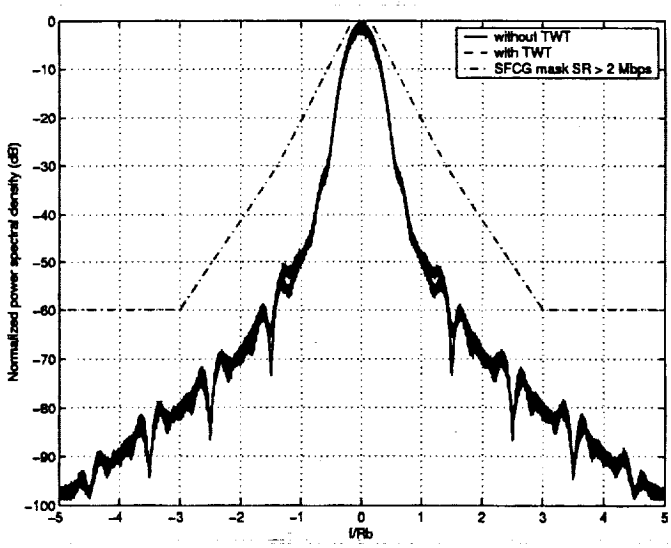


Figure 13: Spectra of EFQPSK. The TWT operates at 0 dB backoff. The figure shows that the effects of the TWT on the spectrum are negligible.

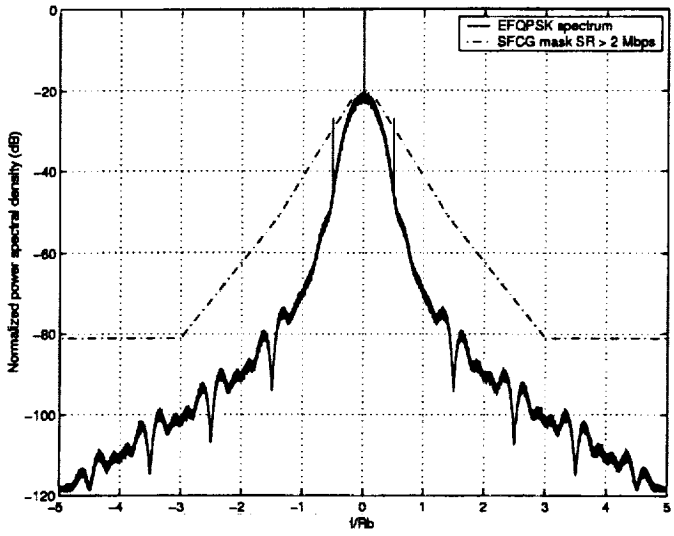


Figure 14: Spectra of EFQPSK for unbalanced data (with +1 occuring 40% and -1 occuring 60%). The DC line is ignored in plotting the SFCG mask.

the TWT is used. This is true irrespective of the TWT model being used, for example, both an ideal clipper model as in Fig. 15 and a more practical model (described in Section 3.5) as in Fig. 16 produce severe spectral regrowth. Figures 17 shows the performance for CERN processed QPSK with an ideal TWT clipper model. It is observed that CERN improves the spectral performance as compared to conventional QPSK of Fig. 15. This agrees with the results presented in [10]. However, when a more practical TWT model (described in Section 3.5) is used, Fig. 18 shows that CERN signal generates considerable spectral regrowth in this case.

Since CERN signal shows good spectral properties with an ideal clipper, a natural choice for improving its spectral properties in a practical TWT environment is to use a predistorter. The combination of the predistorter and the TWT produces an approximate clipper characteristics. The spectrum as a result becomes much more compact as shown in Fig. 19. Even for 0 dB input backoff, a comparison with Fig. 18 shows that the predistorter has improved upon the spectrum shown in Fig. 20. Therefore, an effective predistorter must be realized in order to derive the benefits of CERN.

The predistorter's beneficial role in CERN signals raises an important ques-

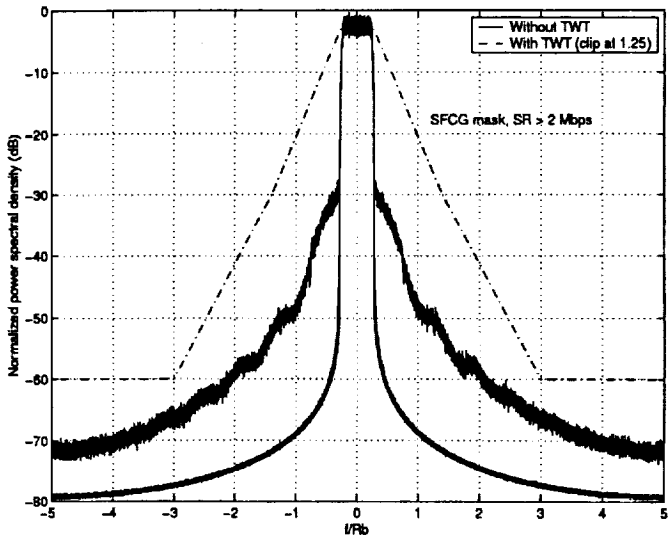


Figure 15: Spectra of conventional QPSK with square root raised cosine filtering (roll-off factor=0.12). The TWT is an ideal clipper.

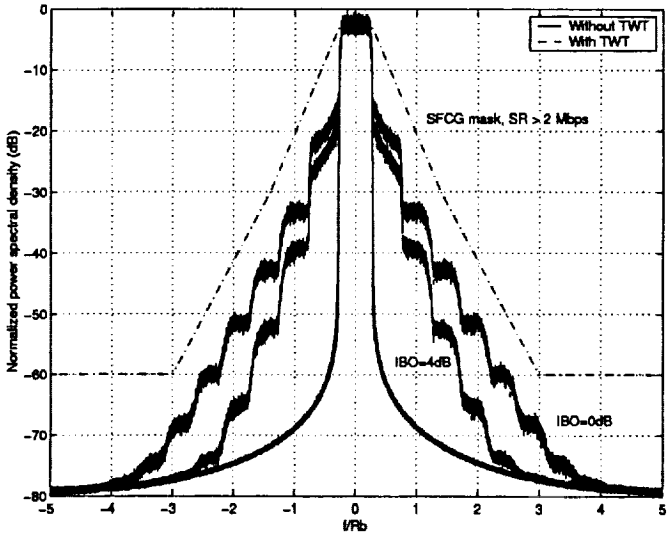


Figure 16: Spectra of conventional QPSK with square root raised cosine filtering (roll-off factor=0.12). The TWT model is more realistic with both AM/AM and AM/PM effects.

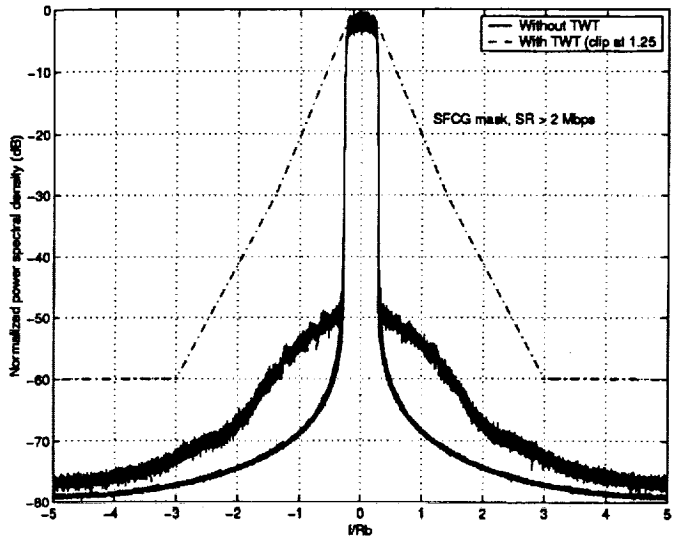


Figure 17: Spectra of QPSK CERN with square root raised cosine filtering (roll-off factor=0.12). The TWT is an ideal clipper.

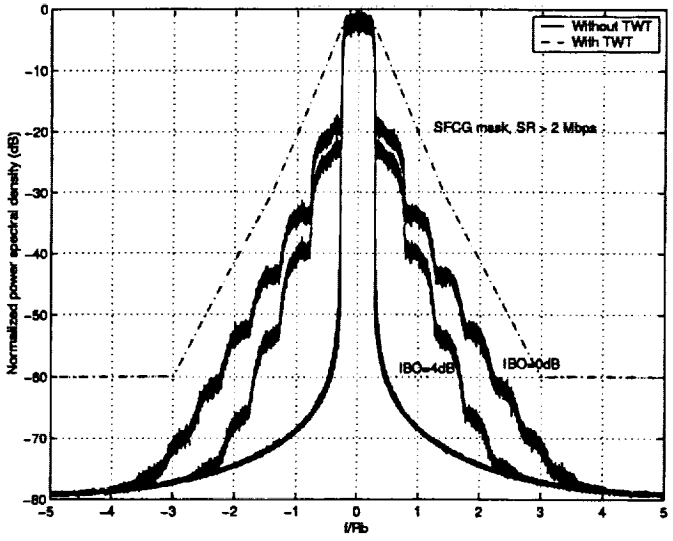


Figure 18: Spectra of QPSK CERN with square root raised cosine filtering (roll-off factor=0.12). The TWT model is more realistic with both AM/AM and AM/PM effects.

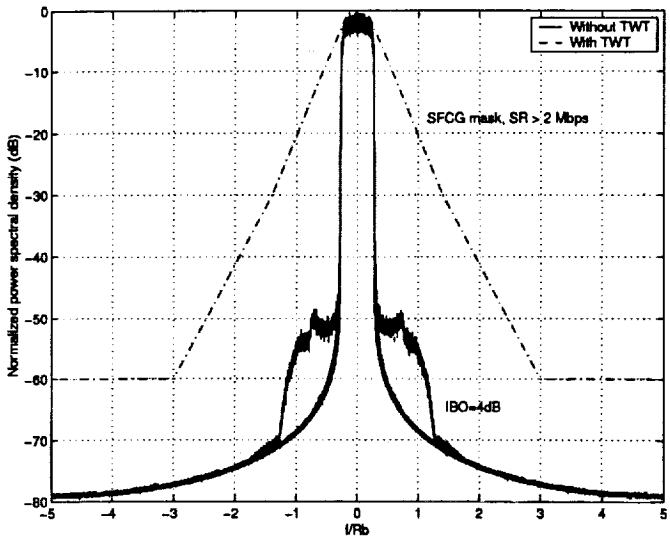


Figure 19: Spectra of QPSK CERN with square root raised cosine filtering (roll-off factor=0.12). The TWT model is more realistic with both AM/AM and AM/PM effects and a predistorter is used.

tion: How much benefit do we get from a predistorter if we directly use conventional QPSK instead of CERN processing? Figure 21 shows that the spectral quality has improved considerably even for conventional QPSK. However, CERN's spectrum is better than the spectrum of conventional QPSK. On the other hand, CERN also introduces signal dependent noise [10], thus compromising the power efficiency (of about 1 dB as reported in [10]) for the bandwidth efficiency.

4.3 BER performance of EFQPSK and CERN

The bit error rate (BER) performance of the receivers for uncoded EFQPSK is shown in Fig. 22. It is observed that the performance degradation due to the TWT is negligible in the SNR range shown. The TWT is operated with 0 dB input back-off. The trellis coded Viterbi receiver's performance agrees with similar results reported in [5]. However, the average matched filter (MF) receiver's performance does not agree with the corresponding results in [5]. The figure also demonstrates our proposed per waveform matched filter (PWMF) receiver's performance. The PWMF receiver provides a performance versus complexity tradeoff since its per-

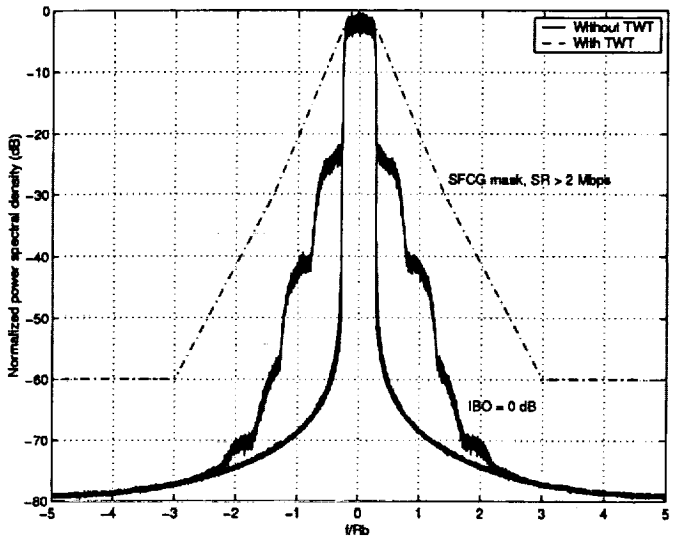


Figure 20: Spectra of QPSK CERN with square root raised cosine filtering (roll-off factor=0.12). The TWT model is more realistic with both AM/AM and AM/PM effects and a predistorter is used.

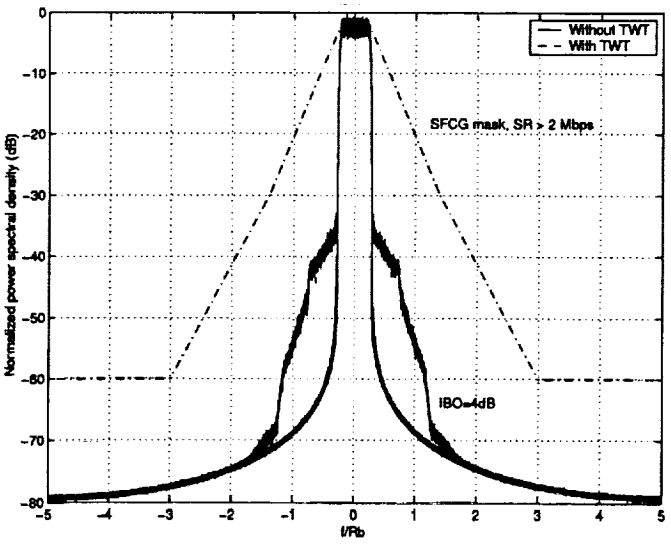


Figure 21: Spectra of conventional QPSK with square root raised cosine filtering (roll-off factor=0.12). The practical TWT model and a predistorter are used.

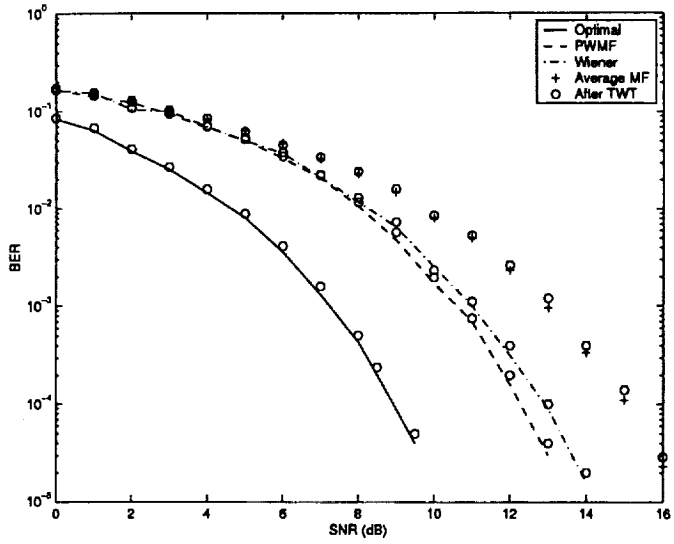


Figure 22: BER performance for uncoded EFQPSK

formance is better than that of the average MF receiver but its complexity is much less than the trellis coded Viterbi receiver. The Wiener filter receiver's performance is also better than the average MF receiver's performance. The BER performance of the receivers is compared with analytical results in Fig. 23 and good agreement is observed. In general, the performance gap among the receivers is found to increase with the increase in SNR.

Figure 24 shows the performance of coded EFQPSK. Convolutional codes of constraint length $K = 7$, rate $1/2$, with generators $g_1 = 1111001$ and $g_2 = 1011011$ are used in the study. This code is used in NASA's Tracking Data Relay Satellite (TDRS). It is observed that coding provides a gain of more than 5 dB near a BER of 10^{-4} for the optimal structure.

Since EFQPSK is not a strictly bandlimited signal, a BER performance study with different receiver structures for different sampling rates r is presented in Fig. 25. This is important for the simulation study and also for the digital implementation of the receivers. It is observed that the optimal receiver's performance is little affected by the sampling rate. However, the suboptimal receivers require higher sampling rates for good performance. To be on a safer side, we use a sampling rate of $r = 22$ samples per symbol interval in simulating our receiver structures.

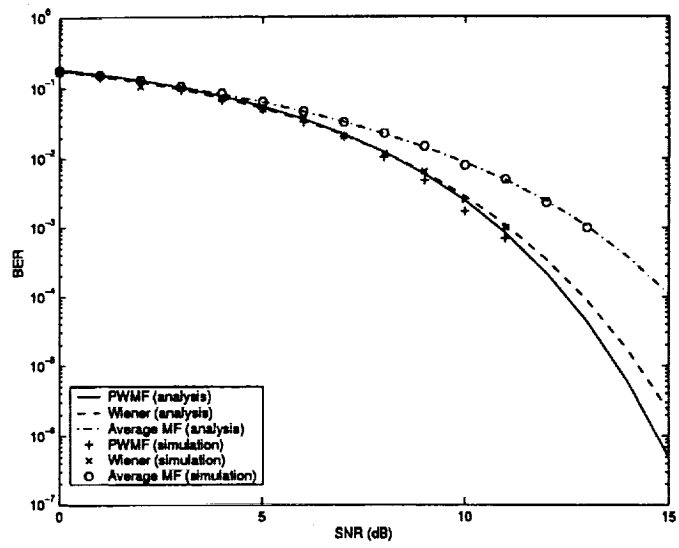


Figure 23: Simulation and analytical BER performance for uncoded EFQPSK

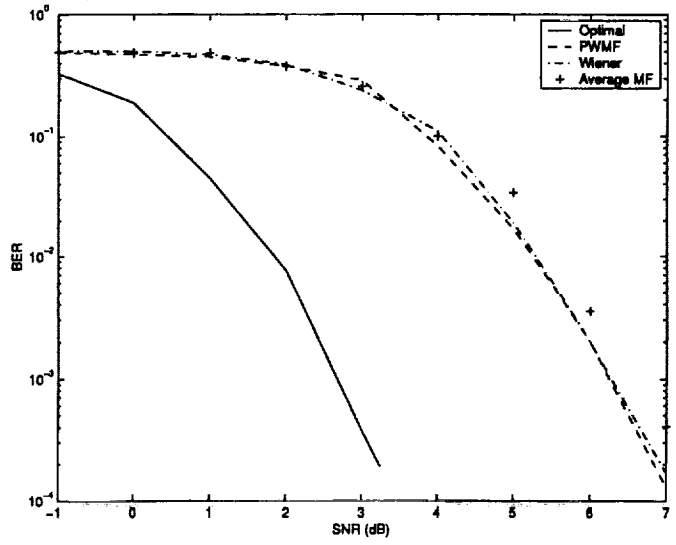


Figure 24: BER performance for coded EFQPSK

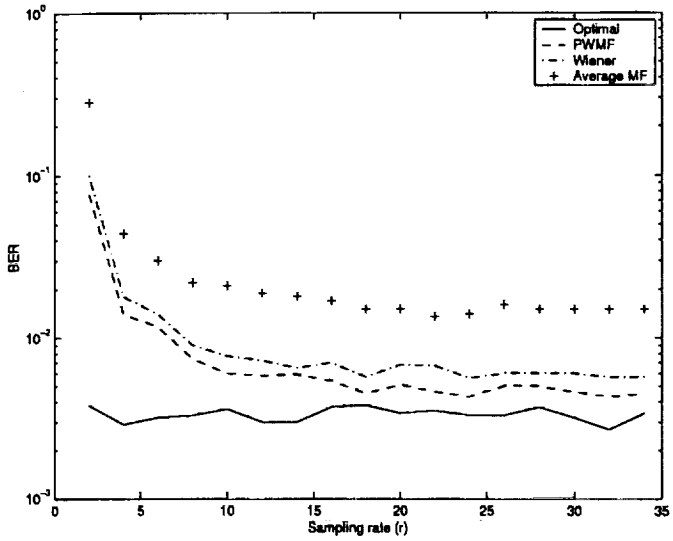


Figure 25: BER performance of uncoded EFQPSK against the sampling rate (r). The optimal receiver operates at 6 dB SNR while the SNR for the other receivers is 9 dB

The nearly fixed envelope of EFQPSK implies that an almost constant AM/PM effect takes place. Therefore, this nearly fixed phase distortion can be compensated by a predistorter which can use the knowledge of the average signal power to calculate and compensate for the phase distortion. Since this process can introduce phase error, we present a study on the effects of this error on various receiver structures. Figure 26 shows that even a large phase error (up to several degrees) in the phase predistortion does not affect performance by much.

The BER performance for uncoded and coded CERN signals is shown in Fig. 27. The receiver structure is a matched filter receiver. The TWT operates at 0 dB input backoff and a predistorter is also used. The figure shows that the use of the TWT results in more than 1 dB performance loss. It is to be noted that a TWT actually amplifies an input signal and so the performance degradation in using the TWT should be seen as a loss in comparison with an ideal perfectly linear amplifier. Coding is found to provide more than 6 dB improvement in performance at a BER of 10^{-4}.

The BER performance comparison of EFQPSK and CERN QPSK is shown in Fig. 28. The data rate is same for both the techniques and the TWT operates at 0

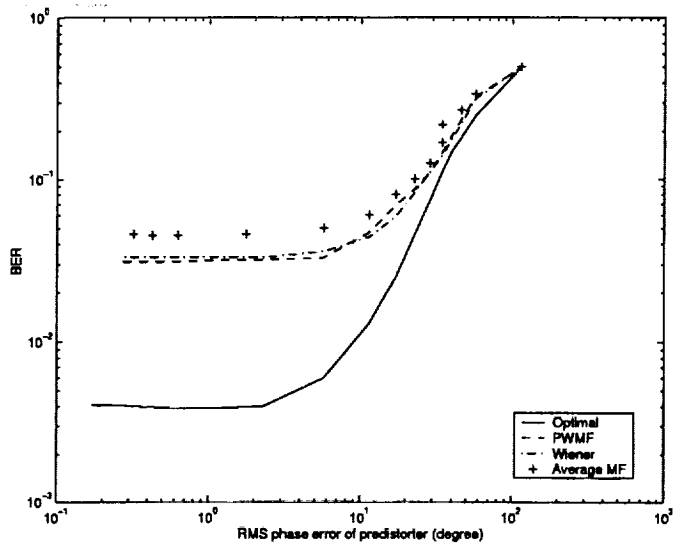


Figure 26: Effects on BER due to phase error in the predistorter

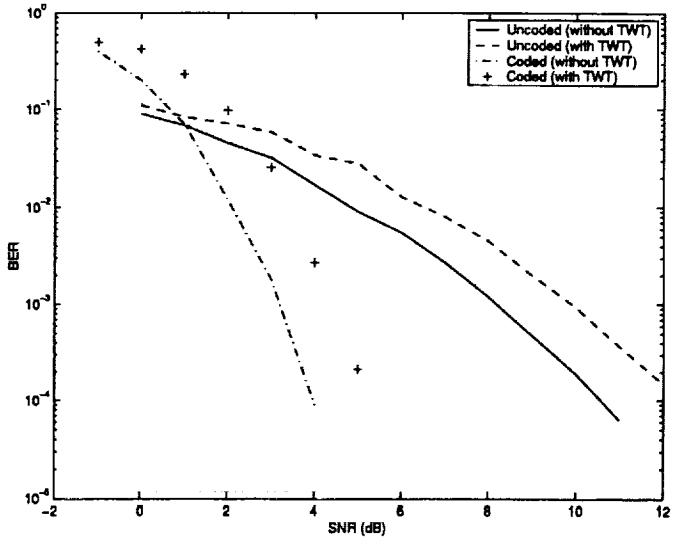


Figure 27: BER performance of CERN QPSK

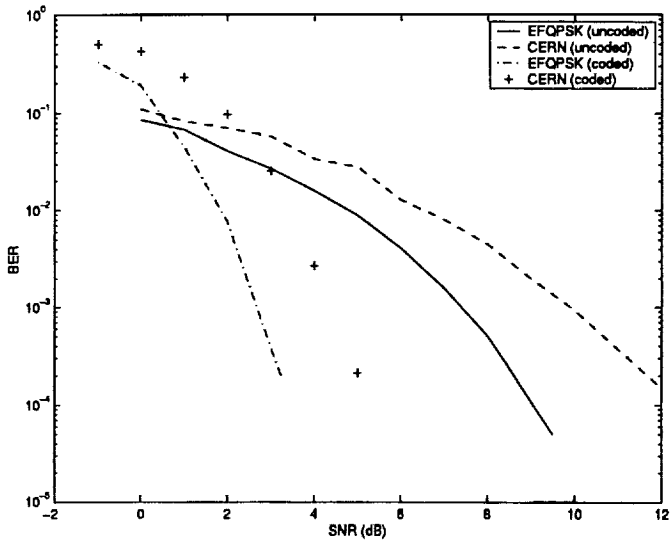


Figure 28: BER performance comparison of EFQPSK and CERN QPSK

dB input backoff for both. Thus the spectral efficiency of CERN is worse than that of EFQPSK in this case (Figs. 13 and 20). The EFQPSK uses the optimal receiver structure while CERN uses an MF receiver. Although EFQPSK uses an optimal receiver structure, it can be implemented easily, and its complexity in the coded case is still not too high compared to the CERN MF receiver. This comparison shows that coded CERN loses about 2 dB in comparison with coded EFQPSK near a BER of 10^{-4}.

4.4 Discussion

Both EFQPSK and CERN signals are found to have advantages and disadvantages. The advantages of EFQPSK include 1) Negligible effects by the NLA. Spectra well within the SFCG mask. 2) Optimal receiver structure can be easily implemented, and high power efficiency can be achieved. The disadvantages of EFQPSK include 1) Its sensitivity to unbalanced data, in which case spectral lines are generated. 2) Spectral efficiency in linear amplifier environment is poorer than conventional QPSK.

The important advantages of CERN signals include 1) Its insensitivity to unbalanced data which do not produce spectral lines except a zero frequency DC

line. 2) High spectral efficiency in conjunction with a predistorter. Moreover, higher order modulation techniques can be used to improve spectral efficiency significantly, albeit at the cost of power efficiency. There are also a few disadvantages of CERN and they include 1) Unknown power normalization factor. 2) Considerable spectral regrowth when a TWT amplifier is used. A predistorter needs to be used to improve the spectral performance.

5 Conclusion

A comparative study on Enhanced Feher's QPSK (EFQPSK) and Constrained Envelope Root Nyquist (CERN) signals has been presented. It is observed that both these techniques provide bandwidth and power efficient communications in nonlinear channels. The EFQPSK achieves this by producing a nearly constant envelope signal by introducing correlation between the inphase and the quadrature channel components. On the other hand, the CERN technique employes conventional root Nyquist pulses and reduces the amplitude fluctuation by repetitively superimposing suitably scaled and delayed versions of the root Nyquist pulse.

It is observed in this study that in a linear channel, the spectral efficiency of CERN is significantly better than that of EFQPSK. The spectral efficiency of CERN in this case is exactly the same as that of a conventional root Nyquist filtered modulation technique. Further, for unbalanced data, CERN processing does not generate spectral lines (except a zero frequency line) while EFQPSK signals are observed to generate strong spectral lines. However, the spectral quality of CERN degrades significantly in the presence of a nonlinear amplifier that generates AM/AM and AM/PM effects, while the spectra of EFQPSK remain unaffected. The spectra of CERN, however, can be significantly improved by using a predistorter. Thus, the spectral performance of CERN greatly depends on the predistorter that has to track and compensate for the nonideal characteristics of a practical power amplifier. It is also observed that such a predistorter can also improve the spectral quality of conventional QPSK signals.

The bit error rate (BER) performance has been studied for both EFQPSK and CERN signals. It is observed that optimal receiver structures can be easily realized for both coded and uncoded EFQPSK signals with not too much increase in computational complexity. When a nonlinear amplifier is used, the bit error rate performance of the uncoded CERN signals with a matched filter receiver is found to be a few decibels (dBs) worse compared to the bit error performance of uncoded EFQPSK signals. Although channel coding is found to provide performance improvement for both EFQPSK and CERN signals, EFQPSK's performance still remains better than that of CERN signals. The deterioration in CERN's BER performance is mainly due to two reasons: 1) CERN introduces a signal dependent noise, and 2) the matched filter receiver used in the study is suboptimal. Improved receiver structures with nonlinear equalization are complex structures, and they are beyond the scope of this report.

Based on the numerical results, it is concluded that, in nonlinear channels, CERN processing leads towards better bandwidth efficiency with a compromise

in power efficiency. Hence for bandwidth efficient communications needs, CERN is a good solution provided effective adaptive predistorters can be realized. In our study, we have used a fixed predistorter structure available in the open literature, as the actual predistorter model used by Intersil could not be obtained. The EFQPSK signals provide a good power efficient solution with a compromise in bandwidth efficiency when compared to CERN signals with good predistorters.

6 Recommendations

The study shows that both EFQPSK and CERN techniques meet the spectral mask requirements of the Space Frequency Co-ordination Group's (SFCG's) recommendation 17-2R1. The only concern for EFQPSK is the case of unbalanced data in which case care must be taken to limit the strong spectral lines. CERN's spectra highly depend on the predistorter's performance. In our study, we have used a fixed predistorter model, as the actual predistorter used by Intersil could not be obtained. Therefore, a study with an adaptive realistic predistorter model is recommended to ensure that such a system meets the spectral requirements in practice. The study should also consider filtered conventional linear modulation, such as the conventional QPSK without CERN processing, since such techniques do not rely on commercial patents. A successful predistorter may also make such conventional techniques superior to EFQPSK in terms of bandwidth efficiency.

This report also shows that, for nonlinear channels, EFQPSK is a more power efficient solution compared to CERN. However, a nonlinear equalizer, although more complex, is expected to reduce this BER performance gap between CERN and EFQPSK. Hence a study of nonlinear equalization techniques for CERN and conventional QPSK is also recommended to find out the maximum power efficiencies, in terms of BER performance, that can be achieved for these techniques.

7 References

References

[1] S. Benedetto and E. Biglieri, *Principles of Digital Transmission with Wireless Applications*, Kluwer Academic/Plenum Publishers, New York, 1999.

[2] H. Yan and K. Feher, "Improved modulation techniques for wireless communications: Raised cosine filters FQPSK- FQPSK(RC)," *IEEE Trans. Broadcasting*, vol. 43, pp. 221–226, June 1997.

[3] S. Kato and K. Feher, "XPSK: A new cross-correlated phase shift keying modulation technique," *IEEE Trans. Commun.*, vol. 31, May 1983.

[4] M. K. Simon and T.-Y. Yan, "Unfiltered FQPSK: Another interpretation and further enhancements, part 1: Transmitter implementation and optimum reception," *Applied Microwave & Wireless*, pp. 76–96, Feb. 2000.

[5] M. K. Simon and T.-Y. Yan, "Unfiltered FQPSK: Another interpretation and further enhancements, part 2: Suboptimum reception and performance comparisons," *Applied Microwave & Wireless*, pp. 100–105, March 2000.

[6] S. Haykin, *Adaptive Filter Theory*, Prentice-Hall, Inc., New Jersey, 1996.

[7] R. D. McCallister, B. A. Cochran, and B. P. Badke, "Constrained-envelope digital-communications transmission system and method therefor," *US Patent 6,104,761*, Aug.15 2000.

[8] A. A. M. Saleh, "Frequency-independent and frequency-dependent nonlinear models of TWT amplifiers," *IEEE Trans. Commun.*, vol. COM-29, pp. 1715–1720, Nov. 1981.

[9] V. Lottici A. N. D'Andrea and R. Reggiannini, "Rf power amplifier linearization through amplitude and phase predistortion," *IEEE Trans. Commun.*, vol. 44, pp. 1477–1484, Nov. 1996.

[10] R. McCallister, R. Putnam, M. Andro, and G. Fujikawa, "Radiation-hardened high-performance modulator module," in *Proc. AIAA'99*, 1999.

[11] J. B. Anderson, *Digital Transmission Theory*, the Institute of Electrical and Electronics Engineers, Inc., New York, 1999.

A Predistorter Model

In this appendix, the fixed linearizer (predistorter) presented in [9] is described. There are two reasons for using this particular predistorter in this study. First, the scope of this report is limited to a study on the effects of the predistorter without the implementation issues, such as the convergence rate etc., of an adaptive predistorter. Hence this fixed predistorter structure is chosen for simplicity. Second, this predistorter assumes perfect knowledge of the NLA characteristics, and since it is not available in practice, an actual predistorter's performance should be worse. However, although perfect knowledge of the NLA characteristics is used, the compensator uses only a suboptimal polynomial model. Thus the predistorter's performance should be slightly worse than that of an optimal predistorter, and therefore, it should be closer to a practical realization.

The complex envelopes at the input and the output of the TWT are rewritten from (3.21) and (3.22) respectively as

$$x(t) = m_x(t)\exp(j\psi_x(t)) \tag{A.1}$$

$$\begin{aligned} y(t) &= A[m_x(t)]\exp(j(\psi_x(t) + \Phi[m_x(t)])) \\ &= x(t)\frac{A[m_x(t)]}{m_x(t)}\exp(j\Phi[m_x(t)]) \end{aligned} \tag{A.2}$$

Using (3.23), (3.24) and (A.2),

$$y = x\left(\frac{\alpha_a}{1+\beta_a m_x^2}\right)\exp\left(j\frac{\alpha_\phi m_x^2}{1+\beta_\phi m_x^2}\right) \tag{A.3}$$

where the dependence on time t has been dropped for notational convenience. The modulus and argument of the complex envelope of the signal at the output of the NLA can be written as

$$m_y = A[m_x] \tag{A.4}$$

$$\psi_y = \psi_x + \Phi[m_x] \tag{A.5}$$

Observe that the right-hand-side in (A.3) is a function of m_x^2. Therefore, it is convenient to reformulate the above equations as

$$\begin{aligned} m_y^2 &= A^2[m_x] \triangleq A_2[m_x] \\ \psi_y &= \psi_x + \Phi[m_x] \end{aligned} \tag{A.6}$$

To obtain an exact linearization of the NLA, the input-output response of the predistorter should be equal to the inverse response of the NLA. That is, the input to the NLA should be

$$\begin{aligned} m_x &= A_2^{-1}[m_y^2] \\ \psi_x &= \psi_y - \Phi[m_x] \end{aligned} \tag{A.7}$$

Therefore, the ideal predistorter output (input to the NLA) is

$$\begin{aligned} x &= A_2^{-1}[m_y^2]\exp\left(j(\psi_y - \Phi[A_2^{-1}[m_y^2]])\right) \\ &= A_2^{-1}[m_y^2]\exp(j\psi_y)\exp(-j\Phi[A_2^{-1}[m_y^2]]) \\ &= A_2^{-1}[m_y^2]\left(\frac{y}{m_y}\right)\exp(-j\Phi[A_2^{-1}[m_y^2]]) \\ &= yG[m_y^2]\exp(-j\Theta[m_y^2]) \end{aligned} \tag{A.8}$$

where

$$G[m_y^2] \triangleq \frac{A_2^{-1}[m_y^2]}{\sqrt{m_y^2}} \tag{A.9}$$

$$\Theta[m_y^2] \triangleq \Phi[A_2^{-1}[m_y^2]] \tag{A.10}$$

The functions $G[m_y^2]$ and $\Theta[m_y^2]$ are approximated by polynomials in the interval $m_y \in (m_{y,1}, m_{y,2})$ as

$$G_\gamma[m_y^2] \approx \sum_{i=0}^{\gamma} \alpha_i m_y^{2i} \tag{A.11}$$

$$\Theta_\eta[m_y^2] \approx \sum_{i=0}^{\eta} \beta_i m_y^{2i} \tag{A.12}$$

where γ and η are the orders of the polynomials in m_y^2. Next, the following two-step procedure is used. First, the polynomial coefficients(A.11) are determined by considering a finite number N_s of uniformly-spaced samples $m_{y,i} = A[m_{x,i}]$, $i = 1, 2, \cdots, N_s$ of the function $A[m_{x,i}]$ in the interval $m_{y,i} \in (m_{y,1}, m_{y,2})$. These values of $m_{y,i} = A[m_{x,i}]$ are used in (A.11) to obtain the corresponding values of $G[m_y^2]$, and then to obtain the coefficients $\{\alpha_i\}$ that minimize the quadratic cost function

$$J \triangleq \sum_{i=1}^{N_s} w(m_y^2)[G[m_y^2] - G_\gamma[m_y^2]]^2 \tag{A.13}$$

where the function $w(m_y^2)$ depends in general on modulation parameters. Once (A.9) has been replaced by polynomial model (A.11), the phase response function $\Theta[m_y^2]$ is replaced by $\Theta'[m_y^2]$ which equals

$$\Theta'[m_y^2] \triangleq \Phi[\sqrt{m_y^2} G_\gamma[m_y^2]] \tag{A.14}$$

The coefficients $\{\beta_i\}$ are then derived so that $\Theta_\eta[m_y^2]$ matches the $\Theta'[m_y^2]$ according to the same fitting rule used in the previous case. In our study, we have used $\gamma = \eta = 1$.

List of Symbols, Abbreviations, and Acronyms

AM/AM	Amplitude Modulation to Amplitude Modulation
AM/PM	Amplitude Modulation to Phase Modulation
AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
CERN	Constrained Envelope Root Nyquist
dB	decibel
EFQPSK	Enhanced Feher's QPSK
FQPSK	Feher-patented QPSK
GMSK	Gaussian Minimum Shift Keying
IBO	Input backoff
ISI	Intersymbol Interference
LMS	Least Mean Squares
MF	Matched Filter
MMSE	Minimum Mean Squared Error
MSE	Mean Squared Error
NLA	Nonlinear Amplifier
pdf	probability density function
psd	power spectral density
PSK	Phase Shift Keying
PWMF	Per-Waveform Matched Filter
QPSK	Quadrature Phase Shift Keying
RF	Radio Frequency
RRC	Root Raised Cosine
SFCG	Space Frequency Coordination Group
SNR	Signal-to-Noise Ratio
TWT	Travelling Wave Tube
XPSK	Cross-correlated Phase Shift Keying

Functions and Operators

$\lfloor \cdot \rfloor$	Floor function
$\lceil \cdot \rceil$	Ceiling function
$*$	Convolution
$\lvert \cdot \rvert$	Absolute value

$\|\cdot\|$	Euclidean norm
$\exp(\cdot)$	Exponential function
$m!$	Factorial
$E\{\cdot\}$	Expectation
$a \bmod b$	Modulus (signed remainder after division a/b)
$\mathrm{erfc}(\cdot)$	Complementary error function
$\mathrm{diag}(\cdot)$	Form a diagonal matrix with the elements shown in brackets as diagonal elements
$(\cdot)^T$	Transpose of a matrix
$(\cdot)^H$	Hermitian transpose of a matrix
$\mathbf{A}^{-1}$	Inverse of a matrix $\mathbf{A}$
$\mathrm{Re}\{\cdot\}$	Real part
$\mathrm{Im}\{\cdot\}$	Imaginary part