Abstract—In this paper, a novel peak power reduction scheme based on trellis shaping is proposed for single-carrier pulse-shaped phase shift keying (PSK) systems. The use of PSK generally results in relatively low signal dynamic range, but its peak-to-average power ratio tends to increase as the bandwidth of pulse shaping filter becomes narrower. The simulation results demonstrate that the proposed approach can generate signals with almost constant envelope, even with the existence of pulse-shaping filters operated with roll-off factors as low as 0.1.

I. INTRODUCTION

Peak power reduction is one of the major issues in practical communication system design. Recently, a number of peak power reduction techniques have been proposed in the literature, but most are confined to orthogonal frequency division multiplexing (OFDM) or other multi-carrier modulation systems, since multi-carrier signals are known to behave Gaussian by nature. However, the single-carrier phase shift keying (PSK) systems also suffer from large dynamic range if the pulse shaping filters with low roll-off factors are employed for enhancing spectral efficiency.

One of the well-known approaches for dynamic range reduction of PSK signal is the use of $\pi/4$-shift QPSK constellation. It reduces signal envelope fluctuation with almost negligible increment of complexity, but its reduction capability is limited. It works for a moderate value of roll-off factor $\alpha$, but suffers from large dynamic range as well, if $\alpha$ is low. Also, it is effective for QPSK constellations, but not for eight or higher order PSK constellations.

Another interesting approach is the use of trellis shaping (TS), originally introduced by Forney [3], and first applied by Morrison [1] to peak power reduction of a single carrier PSK system. This technique makes use of a convolutional coding structure for constraining transition of successive signal constellations. The results of [1], however, do not show significant improvement over $\pi/4$-shift QPSK in terms of dynamic range reduction, even with the reduction of transmission rate due to shaping process. The reason for this may stem from the fact that the metric design proposed in [1] is determined empirically. Specifically, the criterion of code word selection is limited to the phase difference between the two consecutive symbols. Hence, the performance does not gain much compared to the analogous $\pi/4$-QPSK system.

In this paper, we propose a new TS design that can substantially reduce the dynamic range of PSK systems. Surprisingly, signals with almost constant envelope can be achieved. Furthermore, it works even if the pulse-shape filter with small roll off factor is used. It also turns out that the generated signal is also robust against inter-symbol interference associated with the sampling timing jitter.

A major drawback of the proposed TS scheme is its computational overhead. Since the proposed TS actually calculates the resulting oversampled signal waveform, it requires considerable complexity. Therefore, several approaches for computational complexity mitigation are also considered.

The TS requires additional bit redundancy. In our scenario, TS with one bit redundancy is considered. In this case, shaped 8PSK has the same information rate (per symbol) as QPSK without any increase of minimum Euclidean distance. Hence, SNR required for a given error rate increases when systems are compared under the condition of the same information rate. This loss in SNR may be more than compensated for by the benefit of peak power reduction gain in practical systems. Thus, the main focus of this paper is on its achievable dynamic range reduction capabilities. However, we note that combination of channel coding and TS as done in [2] is necessary for practical use in low SNR.

This paper is organized as follows: In the next section, a brief review of TS is given. A figure of merit for a quantitative representation of peak power is given in Section III. Section IV proposes a new TS design, and Section V compares proposed TS with conventional techniques through computer simulations. Finally, concluding remarks are given in Section VI. Note that we focus on shaping system with 8PSK format for comparison with $\pi/4$-QPSK, but the proposed scheme can be applied to higher order PSK as well in a straightforward manner.

II. THE PRINCIPLE OF TRELLIS SHAPING

In this section, we briefly review the principle of TS and notations used in the subsequent sections. The details can be found in [1–4].

Fig. 1 shows a general construction of TS. In this figure, $G_s$ is a $1 \times n_s$ generator matrix of convolutional code and $H^T_s$ is a corresponding syndrome matrix which satisfies $G_sH^T_s = 0$. 

A New Trellis Shaping Approach for Pulse-Shaped PSK Signals with Almost Constant Envelope

Makoto Tanahashi
Department of Electrical and Computer Engineering
Yokohama National University
Yokohama, Kanagawa 240-8501, Japan
Email: makoto@ochialab.dnj.ynu.ac.jp

Hideki Ochiai
Department of Electrical and Computer Engineering
Yokohama National University
Yokohama, Kanagawa 240-8501, Japan
Email: hideki@ynu.ac.jp
Our task here is to find a code word (or sequence) \( v \) given some criterion (which can be specified by applications) for \( v \) word syndrome) is denoted by \( (s) \). The left inverse of this syndrome matrix (simply called inverse syndrome) is \( s = v(H_s^{-1})^T \). Since the inverse syndrome is an \((n_s-1) \times n_s\) matrix, one bit redundancy is imposed here. Next, the code word \( v \), which is to be found by the Viterbi algorithm [5] based on some criterion (which can be specified by applications) for given \( s \) and \( u \), is modulo-2 added to \( s \) to form \( z = s + v \). Finally, \( z \) and \( u \) are mapped into complex symbol sequences \( S \) to transmit. To decode the information sequence at the receiver, \( z \) is simply fed into syndrome to obtain \( s \) as

\[
zH_s^T = (s + v)H_s^T = (s(H_s^{-1})^T + v)H_s^T = s + 0 = s,
\]

noticing that \( vH_s^T = 0 \) as long as \( v \) is a valid code word. Our task here is to find a code word (or sequence) \( v \) that may yield a complex baseband signal with low dynamic range. The details of code word selection metrics are given in Section IV.

III. FIGURE OF MERIT

In many practical applications, signal distortion or associated adjacent channel interference is permitted to some amount. In such a scenario, the relationship between given distortion level and required output back-off (OBO), rather than peak-to-average power ratio (PAR) itself, is of major concern. In this paper, as a simple measure of nonlinear distortion level, we use the average distortion power ratio (ADPR) proposed in [6]. This is used not only to evaluate systems but also to derive a design criterion of the proposed TS metric.

Let \( f(p) \) denote the probability density function (PDF) of instantaneous power \( p \) of complex baseband signal and be normalized such that the average input power is equal to unity, i.e.,

\[
p_{in} = E[p] = \int_0^\infty p f(p) dp = 1.
\]

Suppose that this signal is nonlinearly amplified with AM/AM characteristic of \( a(p) \). We then consider a typical power amplifier (PA) such that \( a(p) \) is assumed to satisfy

\[
a(p) \leq p, \quad 0 \leq p < \infty
\]

and

\[
\max_{0 \leq p \leq \infty} a(p) = p_{max},
\]

where \( p_{max} \) denotes a saturation point of the PA. The average output power is calculated by

\[
p_{out} = E[a(p)] = \int_0^\infty a(p) f(p) dp.
\]

Note that this value is always lower than \( p_{in} \) due to the assumption of (3). Accordingly, the OBO can be defined as

\[
OBO \triangleq \frac{p_{max}}{p_{out}}.
\]

In [6], the ADPR is defined as the average power of the distorted component caused by the nonlinearity of the PA, i.e.,

\[
ADPR = p_{in} - p_{out} = 1 - p_{out}.
\]

It should be noted that this ADPR, though easy to calculate, does not fully characterize the nonlinear distortion component as the correlation between original signal and distortion is not taken into account. In fact, the amount of ADPR considerably overestimates the actual nonlinear distortion components that affect the system performance. The above ADPR should be thus considered rather as a simple indicator of relative amount of distortion.

For a given AM/AM characteristic \( a(p) \), the relationship between nonlinear distortion (i.e., ADPR) and efficiency (i.e., OBO) can be obtained from (5), (6), and (7) with \( p_{max} \) being an intermediate parameter. It is often that baseband signal is fed to a soft envelope limiter before amplification to eliminate unacceptable peak beforehand. In such systems the nonlinearity \( a(p) \) is characterized almost by this limiter and therefore

\[
a(p) = \begin{cases} p_{max}, & p > p_{max} \\ p, & p \leq p_{max}. \end{cases}
\]

where \( p_{max} \) denotes a threshold level over which the instantaneous signal power is uniformly truncated. The ADPR with this limiter model is used for evaluation of the system as well as the metric design of TS.
IV. PROPOSED TRELlis Sscaping

Based on the observation in the above two sections, our task is to shape the PDF of the signal instantaneous power \( f(p) \) to minimize ADPR. This requires controlling transitions of successive symbol constellations. However, the conventional trellis shaping is operated symbol-wise similar to trellis coded modulation (TCM). That is, each transmission symbol is assigned each additional bit, and thus each trellis branch is associated with one PSK symbol. Therefore, the conventional symbol-wise trellis shaping configuration is not sufficient to support signal transition control.

Since the signal dynamic range is a quantity associated with the continuous (i.e. pulse-shaped) signal waveform, we associate each trellis branch with the pulse-shaped continuous waveform, which is referred to as a partial signal.

A. Partial Signal Representation

For a given complex symbol sequence \( \{ S_i \mid -\infty < i < \infty \} \) with \( i \) being a symbol index, we consider the following pulse train

\[
A(n) = \begin{cases} 
S_{n/N_s}, & n \text{ is multiple of } N_s \\
0, & \text{otherwise}, 
\end{cases}
\]

(9)

where \( N_s \) is an oversampling factor (i.e., the Nyquist interval divided by \( N_s \) sampling instants). The output of the filter with impulse response \( g(n) \) can be expressed as

\[
s(n) = \sum_{k=-\infty}^{\infty} A(k)g(n-k).
\]

(10)

Let \( K_s \) denote effective filter length i.e., \( g(n) \) is assumed to sustain non-zero amplitude over \( K_s \) symbol intervals and have a negligible effect elsewhere:

\[
g(n) \approx 0, \quad n < -\frac{K_s}{2}N_s, \quad \frac{K_s}{2}N_s \leq n.
\]

(11)

Let \( s_j(n) \) denote the \( j \)th partial signal that represents the symbol transition between \( S_i \) and \( S_{i+1} \) with \( i = j - (K_s/2 - 1) \), and is valid for one Nyquist symbol interval \( 0 \leq n < N_s \), defined by

\[
s_j(n) \triangleq s \left( n + \left( j - \frac{K_s}{2} + 1 \right) N_s \right).
\]

(12)

Substituting (10) into (12) and with some simplification, we obtain

\[
s_j(n) = \sum_{k=-K_s+1}^{0} S_{k+j}g \left( n - \left( k - \frac{K_s}{2} \right) N_s \right).
\]

(13)

Hence, \( s_j(n) \) is calculated from the \( j \)th symbol \( S_j \) and its \( (K_s-1) \) previous symbols.

B. Trellis Branch Associated with Multiple Previous Symbols

In order to associate the multiple previous symbols with one trellis branch, we extend Morrison’s approach [1] that attaches external memories to shaping encoder \( G_s \). Specifically, for a given shaping bit sequence \( x = \{ x_1, \ldots, x_j, \ldots \} \), the \( j \)th output of \( G_s \) is

\[
v_j = x_j G_s.
\]

(14)

The previous output is expressed using a delay operator \( D \) as

\[
v_{j-1} = x_{j-1} G_s = D x_j G_s = x_j \cdot DG_s.
\]

(15)

The multiple previous outputs for a given time instant can be obtained in a straightforward manner. The current and multiple, say \( m_{ex} \), additional previous outputs can be simultaneously captured to one trellis state transition by replacing \( G_s \) with \([G_s\ DG_s\ \cdots\ D^{m_{ex}}G_s]\).

In order to describe the replica of exact filter output signals, the number of required external memories \( m_{ex} \) should be equal to \( K_s - 1 \). Therefore, as the length of effective impulse response \( K_s \) increases, it causes a substantial increase of computational complexity. If \( G_s \) itself has \( m_G \) memories, the overall memories required for capturing the previous symbols are then given by \( m_G + m_{ex} \) and the number of states of the resulting trellis diagram is \( 2^{m_G+m_{ex}} \).

C. Branch Metric Calculation

As mentioned, the proposed TS is intended to minimize the ADPR. For this purpose, we transform the partial signals to the normalized instantaneous power defined as

\[
p_j(n) = \frac{1}{P_{av}} |s_j(n)|^2,
\]

(16)

where \( P_{av} \) is the average power calculated after trellis shaping. It should be noted that since we consider PSK for signal constellation and well-behaved pulse-shaping filter, unlike usual high-order QAM trellis shaping, the introduction of trellis shaping does not change the average power.

Although we have not yet presented our shaping metric, it is convenient for the following discussion to specify a prototype of branch metric \( B = B(S_j - m_{ex}, \ldots, S_j) \) as

\[
B(S_j - m_{ex}, \ldots, S_j) \triangleq \sum_{n=0}^{N_{j}-1} \mu(p_j(n)),
\]

(17)

where \( \mu(\cdot) \) denotes a metric function with respect to each sample of the normalized instantaneous power.

The overall structure of the branch metric calculation process is illustrated in Fig. 2.

D. Two Metric Functions

We consider the two metric functions referred to as limiter method and moment method.
1) Limiter Method: First metric function is intended to minimize ADPR directly. From (8), it is apparent that minimizing the ADPR based on limiter model is equivalent to minimizing normalized instantaneous power exceeding the given threshold level $p_{max}$. Accordingly, a metric function is formulated as

$$
\mu_l(p) = \begin{cases} 
    p - p_{max}, & p > p_{max} \\
    0, & p \leq p_{max}
\end{cases}.
$$

(18)

The ADPR-OBO characteristics can be controlled by adjusting $p_{max}$ in this equation.

2) Moment Method: It may be also effective to minimize the variance of the output signal around its average. Specifically, the second metric aims at minimizing the $\beta$th order central moment:

$$
\mu_m(p) = (p - 1)^\beta.
$$

(19)

As one can see, the value of statistical moment reflects various aspects of distribution, e.g., peakedness or smoothness. Thus, the choice of $\beta$ is capable of controlling the shape of resulting PDF and associated ADPR-OBO characteristics.

E. Complexity Mitigation

Our TS involves considerable computational effort, which stems mainly from the use of external memories and calculation of partial signals.

1) External Memory: $m_{ex} = K_s - 1$ is required for accurate realization of the partial signals. Since $K_s$ is the non-negligible duration of the impulse response, its typical length is considerable. As a result, a large amount of computational effort is unavoidable. In order to mitigate this, one apparent approach is to curtail $m_{ex}$ at a cost of reduced accuracy of partial signal representation.

2) Oversampling Factor: Unlike Viterbi decoder for channel coding, it is possible to compute metric and tabulate them beforehand in our TS scenario, so that branch metrics are obtained without any complex number operation. However, necessary memory to store the pre-calculated metrics increases exponentially and it even becomes prohibitively large when $m_{ex}$ grows. A possible and appropriate alternative in such a situation is to store a filtered symbol and calculate the partial signal at every branch. In this case, for a given oversampling factor $N_s$, $N_s(K_s - 1)$ complex additions, $N_s$ complex multiplications, and $N_s$ times use of $\mu(\cdot)$ function are necessary for every branch, which dominates the overall complexity. Thus, decreasing $N_s$ may considerably alleviate computational effort with reduced accuracy of partial signal representation.

V. Simulation Result

A. Parameters

We have performed exhaustive computer search for shaping encoder and it has turned out that the optimal bit allocation to shaped bits $s$ is two (hence, $u$ is absent) and the optimal generator in eight-state shaping convolutional encoder is

$$
G_s = \left[ 1 + D^3 \quad 1 + D + D^2 + D^3 \quad D \right].
$$

(20)

Throughout the following simulations, we use a square-root raised-cosine filter as a pulse shaper, with eight times oversampling for continuous output signal generation. The duration of impulse response $K_s$ varies depending on the roll-off factor $\alpha$: we choose $K_s = 12$ for $\alpha < 0.4$, and $K_s = 10$ for $\alpha \geq 0.4$.

B. Limiter Method and Moment Method

In the following simulation, we use $N_s = 8$ (which is equal to the oversampling factor of the pulse shaping filter) and $m_{ex} = K_s - 1$ for partial signal generation. This parameter
setting may be an ideal situation and expected to yield optimal performance. Example baseband waveforms with and without shaping are plotted in Fig. 3, where shaping is based on the limiter method with \( p_{\text{max}} = 1.33 \). Almost constant envelope is observed in both 8PSK and 32PSK.

The ADPR-OBO characteristics for the limiter and moment methods with \( \alpha = 0.1 \) and 0.4 are plotted in Fig. 4 in conjunction with those of \( \pi/4 \)-QPSK and the conventional TS with metrix design according to [1]. Fig. 5 shows the relationship between the roll-off factor \( \alpha \) and the resulting OBO required for achieving \( \text{ADPR} = -50 \) [dB]. We observe the following remarks from these results.

- The curves depend on the metric parameters (\( p_{\text{max}} \) or \( \beta \)). A similar tendency in terms of ADPR-OBO characteristics is observed for the two methods.
- The proposed TS can achieve very low peak power, e.g., OBO = 1.2 [dB] at \( \text{ADPR} = -50 \) [dB] when \( \alpha = 0.4 \). Another remarkable aspect is that the performance does not decay even when \( \alpha \) is decreased, whereas those without shaping show rapid increase of the required OBO. The OBO reduction of 4.6 [dB] is achievable when \( \alpha = 0.1 \).
- The limiter method is always superior to the moment method approximately by 0.2 [dB]. This stems from the fact that the former metric is designed to directly minimize ADPR presented in this paper.

C. Effects of External Memory and Oversampling Factor

The relationship between OBO and \( m_{\text{ex}} \) for a fixed ADPR is plotted in Fig. 6. From this figure, we can observe that compared to the moment method, the limiter method is more sensitive to the length of memory, or, in other words, inaccuracy of the partial waveforms representation.

This behavior is also observed when \( N_s \) is reduced. Fig. 7 shows the relationship between \( N_s \) and required OBO for a given ADPR. It is remarkable that the moment method has no degradation of performance even with \( N_s = 2 \), whereas the limiter method yields noticeable increase of required OBO.
One of major drawbacks of the use of small roll-off factor is its sensitivity to detection timing error caused by the sampling timing jitter at the receiver side. However, since the proposed TS reduces the fluctuation of signal envelope around sampling instants, it exhibits strong robustness against timing jitter. As an example, Fig. 8 shows the 8PSK signal constellation sampled with a fixed timing jitter $\Delta T = T_s/8$, where $T_s$ is the Nyquist interval, with and without proposed shaping. As observed, without shaping, detected symbols at the receiver scatter both in amplitude and phase, whereas only the phase noise dominates for the case with proposed shaping. We have confirmed by simulation that shaped signal with $\alpha = 0.1$ has almost the same error rate as unshaped one with $\alpha = 0.35$ in this amount of timing jitter.

VI. CONCLUSION

In this paper, we have proposed a powerful trellis shaping strategy for peak power reduction of PSK signals. Computer simulations have demonstrated its almost constant envelope behavior. Our method consists of partial signal generation and the associated two metric functions. Both metric functions enjoy adjustability of ADPR-OBO characteristics with their parameters (threshold level $p_{\max}$ or order of moment $\beta$). Simulation results show that the limiter method has better control capability, but is also sensitive to the number of external memories as well as the oversampling factors. Therefore, required complexity of the limiter method is higher than that of the moment method. Finally, we note that the proposed method is similarly applicable to peak power reduction of high-order QAM with some modifications.

REFERENCES