Trellis Shaping for Controlling Envelope of Single-Carrier High-Order QAM Signals

Makoto Tanahashi*, *Student Member, IEEE*, and Hideki Ochiai†, *Member, IEEE*
 Department of Electrical and Computer Engineering, Yokohama National University
 79-5 Tokiwadai, Hodogaya, Yokohama, Kanagawa 240-8501, Japan
 Email: *makoto@ochiailab.dnj.ynu.ac.jp, †hideki@ynu.ac.jp

This paper was in part presented at the *IEEE Global Communications Conference 2008 (GLOBECOM 2008)*, New Orleans, LA, December 2008.

Abstract—In this paper, trellis shaping (TS) is applied to dynamic range control of band-limited single-carrier high-order quadrature amplitude modulation (QAM) signals. With a newly designed shaping metric, we show that a signal with very low peak-to-average power ratio (PAR) can be achieved without significant loss of data rate. A specific example demonstrates that a band-limited transmission with spectral efficiency of 4.55 bit/sec/Hz (including the redundancy due to the shaping) and PAR below 3 dB is achievable using a square 64-ary QAM constellation and a root raised-cosine filter with a roll-off factor 0.1. Furthermore, the proposed TS for high-order QAM can simultaneously reduce the average power and thus offers a shaping gain. The reduction of the PAR and average power can be flexibly controlled by adjusting a parameter associated with the shaping process.

Index Terms—Single-carrier QAM, peak-to-average power ratio, trellis shaping, star QAM

I. INTRODUCTION

To meet an increasing demand in high speed communications with limited spectrum resources, the use of high-order quadrature amplitude modulation (QAM) is inevitable. Even though the peak-to-average power ratio (PAR) of a single-carrier (SC) modulation signal is in general lower than that of orthogonal frequency division multiplexing (OFDM), it can be considerably high in the high bandwidth efficient communication regime: using a high-order QAM constellation and a steep pulse-shaping filter considerably increases the resulting PAR, typically higher than 7 dB. This is because the highly fluctuating impulse response makes the instantaneous power distribution more Gaussian, similar to OFDM signals [1]. Therefore, in order to achieve high power amplifier (PA) efficiency, a PAR reduction technique at the baseband stage plays an important role even for SC-QAM signals.

There have been several PAR reduction approaches proposed in the literature that directly control pulse-shaped (continuous) waveforms of SC signals. In [2, 3], an amplitude scaling method has been proposed. This technique can tightly reduce the peak power to a given level. However, its average power is also reduced and effectiveness in terms of the PAR reduction is not clear. Another method based on amplitude scaling has been proposed in [4]. This method does not affect the spectra, but introduces intentional amplitude fluctuation on signal constellations that contribute to high PAR. Therefore,

the scaling may cause performance degradation for high-order QAM systems.

Another approach that can flexibly control envelope fluctuation is the use of *trellis shaping* (TS). The original TS introduced by Forney [5] was not intended to be used as a PAR reduction scheme, but TS-based PAR reduction approaches have been subsequently studied for SC signal application in [6–8] and for OFDM in [9–11]. The pioneering work by Morrison [6], however, does not offer significant reduction of PAR due to its empirical shaping metric design. The precoding approach of [12], sharing a concept similar to TS, does not provide a significant PAR reduction, either. Recently, in [8], the authors have proposed an elaborate design that takes into account continuous waveform and have shown a significant improvement in terms of the PAR reduction performance for phase shift keying (PSK) signals. This paper extends our work to more intricate QAM constellations and proposes a highly bandwidth-efficient QAM system with a PAR adaptation capability.

The key issue when attempting PAR reduction for high-order QAM constellations is how to control the peak as well as average power effectively. Among many existing PAR reduction techniques in the literature, some techniques also increase the average power. If one maintains the peak power level while increasing the average power, the resulting PAR will be reduced. However, even though a reduced PAR is always preferable from the viewpoint of the PA designers, increasing transmission average power exclusively for PAR reduction may not be well justified from the system designers' viewpoint; the reduced average power in turn increases the minimum squared Euclidean distance (MSED) among constellation symbols (and thus noise immunity) for fixed total energy use, which may improve the link budget. We thus propose a systematic design approach that can offer both peak and average power reduction for a given targeted PAR.

The paper first derives a new unified shaping metric that is applicable to *any* QAM format. With this metric, we also address the following issues.

- TS imposes redundant bits to control signals. With this in mind, we evaluate the PAR and average power reduction capability of the proposed TS, comparing to the unshaped system that has the same bandwidth efficiency.
- Seemingly, star QAM constellations have lower PAR than square or cross QAM, possibly with comparable MSED property. In [13], a trade-off between PAR and MSED

Fig. 1. End-to-end communication model with trellis shaping (TS) for high-order QAMs.

is studied for various 16-ary QAM constellations, and it is shown that star shapes have good performance. The studies on 16 and 32-ary star constellations in [14] also suggest the superiority of these constellations. In this paper, we make a comparative study of various constellation shapes in conjunction with our TS. It turns out that a square QAM is always preferable.

The rest of this paper is organized as follows: Section II describes a trellis shaping system that can control the SC-QAM signal envelope, and Section III proposes a shaping metric suitable for QAM. Section IV presents various QAM constellations considered in the paper. Section V compares these QAMs with the proposed TS. Section VI discusses performance of shaped QAM systems along with their efficiencies. Finally, concluding remarks are stated in Section VII. Notations used in this paper basically follow those of [8] and the reader is referred to [8] for the detailed description.

II. SYSTEM DESCRIPTION

A. Generation of Band-Limited SC-QAM Signals

Figure 1 illustrates a baseline communication system model for SC-QAM with TS considered in this paper. As illustrated in the figure, the most significant bits (MSBs) used for shaping are called shaping bits and denoted by s , while the least significant bits (LSBs), called non-shaping bits and denoted by u , are directly mapped onto a QAM constellation. The size of s is denoted by¹ n_s , while in this paper we restrict $n_s = 2$ for simplicity.

The TS process makes use of the convolutional coding structure: let G_s denote a $1 \times (n_s + 1)$ generator matrix of a binary convolutional code, the entries of which are polynomials in D similar to ordinary convolutional codes for error correction. We denote the constraint length of G_s by K . An example of this shaping code is presented in Section V-A along with some simulation parameters. An $(n_s + 1) \times n_s$ corresponding syndrome former matrix, denoted by H_s^T , satisfies $G_s H_s^T = \mathbf{0}$, where $\mathbf{0}$ is a zero row vector of length n_s . The $n_s \times (n_s + 1)$ left inverse of this syndrome former matrix (hereafter simply called inverse syndrome former) is denoted by $(H_s^{-1})^T$ and satisfies $(H_s^{-1})^T H_s^T = \mathbf{I}$, where $\mathbf{I}$ is an identity matrix of size n_s .

¹In our previous work [8], n_s is defined as the size of $\tilde{s}$. For convenience of explanation, we have defined n_s as the size of s in this paper.

After the convolutional encoding by the inverse syndrome former, s is transformed into $\tilde{s} = s(H_s^{-1})^T$ with one bit redundancy imposed². The size of $\tilde{s}$ is thus $n_s + 1$. Next, an arbitrary bit sequence x is encoded by G_s , added to $\tilde{s}$, and forms³ $z = \tilde{s} + xG_s = \tilde{s} + v$. Finally, the multiplexed bit sequence $[z, u]$ is mapped onto the QAM constellation to generate complex symbols $S = \{S_0, \dots, S_{N-1}\}$, where N is the number of transmitted QAM symbols for a given frame. This symbol sequence is then converted to a continuous-time baseband signal $s(t)$ by a pulse-shaping filter $g(t)$ as

$$s(t) = \sum_{l=0}^{N-1} S_l g(t - lT), \quad (1)$$

where T denotes the symbol period, i.e., the Nyquist interval. We assume $g(t)$ to be the impulse response of a root raised-cosine (RRC) filter with an effective (non-zero) duration spanning over K_s symbols.

At the receiver, the demapper estimates mapped bits from the detected samples R . Note that, in Fig. 1, the notation $\hat{\cdot}$ (hat) denotes the estimate of a given symbol. To retrieve s at the receiver, the estimate of z is fed to the syndrome former H_s^T . Then the following equation is satisfied as long as the estimation of z is perfect.

$$\begin{aligned} z H_s^T &= (\tilde{s} + v) H_s^T = (s(H_s^{-1})^T + v) H_s^T \\ &= s(H_s^{-1})^T H_s^T + x G_s H_s^T = s \mathbf{I} + x \mathbf{0} = s. \end{aligned} \quad (2)$$

An important property of TS is that even though the actually transmitted signal $s(t)$ depends on x , (2) holds for arbitrary x . Thus, without loss of information, the waveform of $s(t)$ can be controlled by choosing x . We thus refer to x as a waveform controlling sequence.

B. Shaping Process

This subsection outlines how to select an optimum waveform controlling sequence x .

The discrete-time representation of the continuous waveform $s(t)$ of (1) is obtained by sampling $s(t)$ with a time

²For simplicity, we do not consider multidimensional TS [8] that can mitigate this redundancy at the cost of the PAR reduction capability.

³The Viterbi Algorithm directly selects a valid codeword v from the convolutional code defined by G_s . Therefore, encoding of x by G_s need not be implemented in practice.

interval $\Delta T = T/N_s$, where N_s is an oversampling factor. The n th sample $s[n] \triangleq s(n\Delta T)$ is then given by

$$s[n] = \sum_{l=0}^{N-1} S_l g[n - lN_s], \quad (3)$$

where $g[n] \triangleq g(n\Delta T)$. Furthermore, the associated instantaneous power form is given by $p[n] = |s[n]|^2$. The optimum $\mathbf{x}$ is selected with respect to the previously determined input data such that it minimizes the total summation of a certain metric for each sample, i.e.,

$$\mathbf{x} = \arg \min \left[\sum_{n=0}^{N_s N - 1} \mu(p[n]) \right], \quad (4)$$

where $\mu(\cdot)$ is a *metric function* that provides a metric of a given instantaneous power sample. To search such $\mathbf{x}$, (4) is implemented with the Viterbi algorithm (VA), where the l th trellis section is associated with the truncated version of $p[n]$ for the $(l - K_s/2)$ -th⁴ Nyquist symbol interval, i.e., $p_l[n] \triangleq p[n + (l - K_s/2)N_s]$, $n = 0, 1, \dots, N_s - 1$.

The most important observation here is that, with a suitable definition of $\mu(\cdot)$, we can arbitrarily shape the band-limited waveforms. Also, it should be noted that since the equations from (3) to (4) are not restricted to a particular constellation shape, we can apply this shaping process to any QAM constellation. Thus, our single task is to design $\mu(\cdot)$ that can efficiently reduce the PAR of QAM signals.

We briefly note the complexity of this shaping process. In the VA, the number of the trellis states is given by $2^{(K-1+K_s-1)}$, while the original TS [5] requires only $2^{(K-1)}$ states. Nevertheless, it is worth noting that the complexity of the VA does not depend on employed constellation size M . (Another trellis-based approach proposed in [15] has the computation order proportional to M^{K_s}). Although a possible reduction in complexity given in [8] is applicable to the QAM cases in a straightforward manner, we do not discuss complexity issues further in this paper, as our main focus is on evaluating achievable maximum shaping capability.

III. SHAPING METRIC FOR QAM

In general, the average power should vary before and after shaping in the case of QAM, and this effect should be taken into account for designing a shaping metric. In this section, we systematically develop a metric function applicable to any QAM format.

A. Step 1: Estimating Peak Power Control Capability

Since the signal PAR depends on the data sequence, it can be characterized as a probabilistic variable. Thus, we associate the peak power with its probability of occurrence. For a given small probability ε , we define the threshold peak power p_{peak} as

$$\Pr[p > p_{peak}] = \varepsilon. \quad (5)$$

Our interest here is that for a given allowable excess probability ε , how much the peak power p_{peak} can be reduced by the TS, or alternatively, how much the resulting clipping probability ε is for a given peak power p_{peak} . Finding the relationship between ε and p_{peak} for specified settings of TS (e.g., shaping encoders and allocations of shaping bits) and QAM format is the objective of this first step. To this end, we define a metric function as

$$\mu_{\text{test}}(p) = \mathbb{I}_{p > p_{peak}}, \quad (6)$$

where $\mathbb{I}_{\mathcal{A}}$ is an indicator function that has a unit value if the event $\mathcal{A}$ occurs and 0 otherwise. We perform a simulation for a given TS system using this *test* metric which estimates the capability of the peak power control. Specifically, we use the fact that, with this shaping metric, the sum metric after completion of the VA provides the number of samples that exceed a given threshold p_{peak} . Hence, if the sum metric is N_a , then the threshold probability ε is calculated by

$$\varepsilon = \Pr[p > p_{peak}] \approx \frac{N_a}{NN_s} \quad (7)$$

for sufficiently large number of symbols N . Based on this procedure, we obtain the desired ε - p_{peak} relationship.

B. Step 2: Adjustment of Average Power

The previous step allows us to fix the threshold peak power p_{peak} for a given excess probability ε . Proper setting of p_{peak} may also result in the average power reduction since the entire amplitude is made smaller. Although the reduction of the average power can be reflected by an enhancement of MSED (hence, noise margin), it in turn may increase the resulting PAR, as this figure is a ratio of the peak and average power. Thus, we wish to control the average power to meet both PAR and average power requirements. From this perspective, we propose a metric function that minimizes the difference between a certain reference power level p_{ref} and instantaneous power of the resulting signal, while at the same time eliminating the occurrence of the instantaneous power above the previously determined threshold p_{peak} :

$$\mu(p) = \begin{cases} C, & p > p_{peak} \\ |p - p_{ref}|, & \text{otherwise,} \end{cases} \quad (8)$$

where C is an arbitrary large positive constant. The choice of C does not affect the performance as long as it is much larger than a maximum value of $|p - p_{ref}|$. In simulations, we set $C = 10^6$.

If we set $p_{ref} = p_{peak}$, (4) can be rewritten as

$$\begin{aligned} \mathbf{x} &= \arg \min \left[\sum_{n=0}^{N_s N - 1} |p[n] - p_{peak}| \right] \\ &= \arg \min [p_{peak} - p_{av}] = \arg \min \left[\frac{p_{peak}}{p_{av}} \right], \end{aligned} \quad (9)$$

where

$$p_{av} = \frac{1}{N_s N} \sum_{n=0}^{N_s N - 1} p[n] \quad (10)$$

⁴The shift by $K_s/2$ is to compensate for non-causality of $g(t)$.

denotes the resulting average power. Hence, p_{av} is maximized and then the PAR is minimized. On the contrary, setting $p_{ref} = 0$ is equivalent to

$$\mathbf{x} = \arg \min \left[\sum_{n=0}^{N_s N-1} p[n] \right] = \arg \min [p_{av}], \quad (11)$$

which means that the average power is simply minimized, similar to the original operation of the TS as average power reduction [5].

By changing p_{ref} within the range $[0, p_{peak}]$, an arbitrary trade-off between the PAR reduction and the average power reduction can be achieved. Therefore, if the PA under operation does not have severe nonlinearity, one can enhance the MSE by setting p_{ref} to a small value. On the other hand, if the PA efficiency is of primary concern, one may set $p_{ref} = p_{peak}$ for full PAR reduction capability.

It should be noted that once the desired p_{peak} and p_{ref} are determined by the above two steps, the actual metric to be calculated for every TS processing is (8).

IV. QAM CONSTELLATIONS

In this section, we describe various QAM constellations to be examined along with their bit labelings.

A. Square and Cross QAMs

For square (or cross) 16, 64, 128, and 256-ary QAMs, we simply refer to these standard QAMs as M -QAMs in the following.

Since these high-order QAMs are suitable for trellis-coded modulation (TCM), we assume that their bit labeling is Ungerboeck's set-partition labeling, where the 3 bits associated with the deepest partition levels (having the largest intra-subset MSEs) are used for the shaping (see z in Fig. 1). The remaining bits can be used for TCM. However, in this paper, we do not explicitly consider the use of TCM, since the TS and TCM, influencing different bits, do not affect the performance of each other.

B. Star QAMs

As opposed to the above mentioned standard QAMs for which the set-partition labeling is established in conjunction with the TCM, the bit labeling of star QAMs is typically split into phase and amplitude [16]. As example constellations, we consider the popular 16-ary star QAM (see, e.g., [13]) and the 64-ary star QAM presented in [16]. In this paper, we refer to them as 16-STAR and 64-STAR, respectively. The 16-STAR can be viewed as a combination of two concentric 8-PSK rings having different radii, where one bit determines the amplitude and the remaining three bits determine the phase of the 8-PSK ring. Similar to the structure of 16-STAR, 64-STAR is a combination of the four concentric 16-PSK rings, where two bits determine the amplitude and the remaining four bits determine the phase of the 16-PSK ring. The optimum ring ratios given in [13, 16] are assumed for the two star QAMs. The shape of the constellations will be later shown along with the simulation results.

Fig. 2. The ε - p_{peak} relationships for various QAM constellations at the roll-off factor $\alpha = 0.1$.

Fig. 3. CCDF of shaped/unshaped QAM signals with $\alpha = 0.1$. The reference power is set to $p_{ref} = p_{peak}$ for maximal PAR reduction performance. The values attached to the legends denote the peak power p_{peak} and the average power p_{av} after shaping (in dB), from left to right.

In the case of 16-STAR, we examine the following two assignments of the shaping bits: 1) all the 3 phase bits are to be shaped, and 2) the amplitude bit and 2 of the 3 phase bits are to be shaped. In the former assignment, referred to as phase-shaped 16-STAR (P-16-STAR), we only control phase patterns so that the average power never varies before and after shaping. Thus, the choice of the reference power p_{ref} in (8) does not affect the result. Conversely, the latter assignment, called amplitude-shaped 16-STAR (A-16-STAR), can control the average power by proper setting of p_{ref} .

Unlike 16-STAR, the constellation of 64-STAR itself has a large dynamic range. Thus, amplitude bits should be shaped so as to mitigate the dynamic range. We assign the two amplitude bits and one of the phase bits to the shaping part (i.e., z in Fig. 1).

Fig. 4. Shaping examples: baseband waveforms of 16 and 64-ary QAM constellations with $\alpha = 0.1$. The constellation points are also plotted.

V. NUMERICAL PERFORMANCE

A. Parameter Setting

In the simulations, we use the following example shaping encoder that offers good shaping capability (from [8]):

$$\mathbf{G}_s = \begin{bmatrix} 1 + D^3 & 1 + D + D^2 + D^3 & D \end{bmatrix}, \quad (K = 4).$$

As an example of strict band-limitation, a RRC filter is used with a roll-off factor of $\alpha = 0.1$. The effective duration of this filter is chosen to be $K_s = 12$, and the oversampling factor for the discrete representation of the filtered signal is $N_s = 8$. These parameters are chosen large enough to exhibit best achievable performance. As shown in [8], reducing K_s and N_s may result in the system with lower complexity but with degradation of envelope controlling capability.

Note that, in the simulation results, the values representing relative powers such as p_{peak} , p_{av} , and p_{ref} are normalized such that the average power without shaping, i.e., when all constellation points are equally likely, is 1 (0 dB).

B. Peak and Average Power Reduction

First, we investigate the ε - p_{peak} relationship by using the test shaping metric μ_{test} given in (6). The results are plotted in Fig. 2. From this figure, it is observed that 64-STAR can achieve the lowest peak power. However, it should be noted that this does not necessarily result in the lowest PAR as its average power is also reduced by this procedure. We have to fix the reference power p_{ref} to evaluate peak and average power reduction performance. In what follows, we will adopt the value of p_{peak} at $\varepsilon = 10^{-4}$ as an example.

Next, we set $p_{ref} = p_{peak}$ (full PAR reduction mode) to see the maximum achievable performance in terms of PAR reduction. Figure 3 shows the complementary cumulative distribution function (CCDF) characteristics of the normalized instantaneous power of shaped and unshaped QAM signals in this case. This quantity, defined as

$$\text{CCDF}(p_0) = \Pr \left[\frac{p}{p_{av}} > p_0 \right], \quad (12)$$

Fig. 5. Trade-off of the average power p_{av} and PAR p_{peak}/p_{av} with respect to a given reference power p_{ref} in the case of 64-QAM.

represents the probability that a normalized instantaneous power p/p_{av} exceeds a given level p_0 . The predetermined peak power, p_{peak} , and resulting average power after shaping, p_{av} , are also shown in the legend of the figure. From the CCDF curves, we can observe that the reduction of the signal dynamic range is significant. Moreover, it is noteworthy that even if we set p_{ref} in favor of the PAR reduction (i.e., high reference power level), a certain amount of reduction in the average power is preserved in all the constellations simulated here, except for P-16-STAR that does not have capability of controlling its amplitude.

Several further remarks can be made: from (5) and (12), it is apparent that the gap between p_{peak} (abscissa of Fig. 2) and threshold instantaneous power p_0 (abscissa of Fig. 3) observed at $CCDF = \varepsilon$ is precisely equal to p_{av} in dB. For instance, the normalized instantaneous power at $CCDF = 10^{-4}$ is $p_0 = 3.1$ dB for the 256-QAM case. Since the peak power p_{peak} at $\varepsilon = 10^{-4}$ is 2.3 dB, the amount of an average power reduction can be expressed as $[p_0]_{dB} - [p_{peak}]_{dB} = 0.8$ dB, which agrees with the value of p_{av} listed in Fig. 3. Note that some of the shaped curves exhibit a floor in the region $CCDF < \varepsilon$, which stems from the fact that peak power larger than p_{peak} can occur with probability less than ε . This floor can be arbitrarily lowered by using a larger p_{peak} in the metric function associated with a smaller value of ε .

The shaped waveforms (trajectories) of the various QAM constellations with and without shaping are shown in Fig. 4. From this figure, the effectiveness of the proposed TS can be confirmed, as the occurrence of the outermost symbols is completely eliminated (except for the P-16-STAR case). This observation verifies the fact that both peak power and average power are reduced. Moreover, it is worthwhile to point out that our results do not comply with the empirical lore that QAM constellations with star shape have PAR lower than square or cross QAM. It is observed that when a RRC filter with a roll-off factor as small as 0.1 is used, the shape of an underlying constellation becomes irrelevant to the PAR of the pulse-shaped signals, even without constraint by shaping.

Finally, in Fig. 5 the relationship among a given p_{ref} ,

the resulting p_{av} , and the PAR $\frac{p_{peak}}{p_{av}}$ is plotted for the 64-QAM case as a specific example. As observed, decreasing the reference power level p_{ref} results in smaller average power p_{av} and thus higher PAR $\frac{p_{peak}}{p_{av}}$, and vice versa.

VI. DISCUSSIONS

In this section, we make comparative studies on the performance of the proposed TS based on the numerical results in the previous section, focusing on the following two different viewpoints: 1) overall system efficiency in terms of the PAR reduction capability, bandwidth efficiency, and noise margin (i.e., minimum Euclidean distance), and 2) effect of QAM constellation shapes in terms of the noise margin and PAR.

A. Overall System Evaluation

First, we introduce the following notations. Let $T^{(s)}, \mathcal{E}^{(s)}$, and $\alpha^{(s)}$ denote the Nyquist interval, the average energy of one QAM symbol, and the roll-off factor of the RRC filter, respectively, all in the shaped case. Likewise, let $T^{(u)}, \mathcal{E}^{(u)}$, and $\alpha^{(u)}$ denote those of the unshaped case. For a particular constellation order M , the information conveyed by one symbol is given by $m - 1$ bit/symbol for the shaped case, where $m = \log_2 M$, due to the one bit redundancy added by the shaping process.

The proposed TS is capable of reducing both the PAR and average power even with a narrow pulse-shaping filter and thus it has a potential of increasing bandwidth efficiency by using a small roll-off factor. However, it in turn reduces an information rate due to the introduction of the shaping bits. In order to make a fair comparison of this trade-off, we evaluate these improvements and rate loss by imposing a constraint such that the comparing shaped and unshaped systems achieve the same information rate with occupying the same bandwidth. Hence, the following equations should hold:

$$\frac{m-1}{T^{(s)}} = \frac{m}{T^{(u)}} \quad \text{and} \quad \frac{1+\alpha^{(s)}}{T^{(s)}} = \frac{1+\alpha^{(u)}}{T^{(u)}}. \quad (13)$$

If this is satisfied by using different roll-off factors, in other words, the rate loss associated with shaping is compensated for by reducing the roll-off factor, eliminating $T^{(s)}$ and $T^{(u)}$ leads us to the following condition on the roll-off factor for the same bandwidth efficiency:

$$\alpha^{(u)} = \frac{m}{m-1}(1+\alpha^{(s)}) - 1. \quad (14)$$

Hence, to make a comparison with the shaped system with $\alpha^{(s)} = 0.1$ shown in Figs. 3 and 5, we consider reference unshaped QAM systems with $\alpha^{(u)} = 0.47, 0.32, 0.28$, and 0.26 for 16-QAM, 64-QAM, 128-QAM, and 256-QAM, yielding the bandwidth efficiencies of 2.73, 4.55, 5.45, and 6.37 bit/sec/Hz, respectively.

1) *PAR Reduction*: By evaluating the CCDF characteristics of the shaped and unshaped signals, we can immediately obtain their PARs as the normalized instantaneous powers at $CCDF = \varepsilon$. Let $PAR^{(s)}$ and $PAR^{(u)}$ denote these PARs. Note

TABLE I
TRADE-OFF OF η_p AND η_s FOR THE 64-QAM CASE.

reference power p_{ref}	1.6	1.0	0.8	0.6	0.4	0.0
PAR reduction η_p (dB)	3.4	3.3	2.9	2.1	1.4	1.2
shaping gain η_s (dB)	0.2	0.3	0.7	1.5	2.1	2.3

that we can evaluate the PAR of the shaped case alternatively as $\text{PAR}^{(s)} = \frac{p_{peak}}{p_{av}}$. The PAR reduction is then calculated as

$$\eta_p \triangleq \left[\frac{\text{PAR}^{(u)}}{\text{PAR}^{(s)}} \right]_{\text{dB}} = [\text{PAR}^{(u)}]_{\text{dB}} - [p_{peak}]_{\text{dB}} + [p_{av}]_{\text{dB}}. \quad (15)$$

2) *Shaping Gain*: Since the average power of the shaped signals is decreased for a fixed average energy of a QAM symbol, this in turn increases the utilizable average energy of a QAM symbol when the average power of band-limited signals is constrained. Recall that, in the simulation results (see Section V-A), the average power is decreased from 1 to p_{av} by shaping. Thus, for the shaped systems, the average energy of a QAM symbol is determined such that the average power is increased up to $1/p_{av}$. The average energies $\mathcal{E}^{(s)}$ and $\mathcal{E}^{(u)}$ are then expressed as

$$\mathcal{E}^{(s)} = (1/p_{av}) \cdot T^{(s)} \quad \text{and} \quad \mathcal{E}^{(u)} = 1 \cdot T^{(u)}. \quad (16)$$

Taking the ratio of the above and substituting (13) lead us to shaping gain η_s expressed as

$$\eta_s \triangleq \left[\frac{\mathcal{E}^{(s)}}{\mathcal{E}^{(u)}} \right]_{\text{dB}} = -[p_{av}]_{\text{dB}} - \left[\frac{m}{m-1} \right]_{\text{dB}}. \quad (17)$$

This quantity, representing how much more energy a QAM symbol can use, can be equivalently treated as an increase in the noise immunity. Thus, error rate curves versus signal-to-noise power ratio (not shown in this paper for space limitation) will be shifted left by η_s for a fixed noise variance. Note that the second term of (17) represents the loss in terms of the shaping redundancy. We also note that the conventional definition of the “shaping gain” refers to discrete QAM constellation prior to pulse-shaping filtering, whereas the shaping gain in our definition takes into account the band-limitation effect. Due to this difference, the theorem of the ultimate shaping gain (1.53 dB) [17] is not applicable to our results.

3) *Evaluation*: From our simulation, we have obtained PAR reductions η_p of 2.0, 3.4, 3.1, and 3.4 dB, and shaping gain η_s of 0.9, 0.2, -0.3, and 0.2 dB for 16-QAM, 64-QAM, 128-QAM, and 256-QAM, respectively, at the highest reference power setting (i.e., $p_{ref} = p_{peak}$). The trade-offs of η_p and η_s for the 64-QAM case are summarized in Table I. Note that, from (15) and (17), it is apparent that the total gain $\eta_p + \eta_s$ is independent of p_{av} , and hence the choice of the reference power p_{ref} . This fact can be confirmed from Table I. The balance between η_p and η_s is, however, practically important (and thus we adjust it by the reference power p_{ref}). The amplifier designers may request low PAR, but the communication system designers may request large noise margin. Optimizing total efficiency requires the precise modeling of the PA under operation, and is beyond the scope of this paper.

B. Comparison of QAM Constellations

As shown in Section V-B, there is no explicit relationship between constellation shapes and pulse-shaped waveforms if the proposed TS is applied. Thus, we quantitatively evaluate the constellation shapes in terms of PAR and MED. In what follows, we compare 16-QAM versus 16-STAR and 64-QAM versus 64-STAR shaped at $p_{ref} = p_{peak}$ as a specific example.

1) *MSED*: For given average energy of constellation symbols denoted by $\mathcal{E}$, the MSEDs are calculated as

$$\begin{aligned} &0.40\mathcal{E} \text{ (16-QAM)}, \quad 0.28\mathcal{E} \text{ (16-STAR)}, \\ &0.095\mathcal{E} \text{ (64-QAM)}, \quad 0.042\mathcal{E} \text{ (64-STAR)}. \end{aligned} \quad (18)$$

Substituting $\mathcal{E}^{(s)} = (1/p_{av}) \cdot T^{(s)}$ and $\mathcal{E}^{(u)} = 1 \cdot T^{(u)}$ of (16) into $\mathcal{E}$ in (18) yields the MSEDs for the shaped and unshaped cases taking into account the savings in the average power due to the proposed TS. Note that, since our objective here is to evaluate the MSEDs of the constellations having the same number of symbol points, both $T^{(s)}$ and $T^{(u)}$ are assumed to be a unit value.

2) *PAR*: We obtain the PARs of the pulse-shaped continuous signals in the same way as that presented in the previous subsection. For comparison, we also evaluate PARs without pulse-shaping taken into consideration, which can be defined as the ratios of the maximum and average energies in the given constellation shapes⁵. These quantities are deterministic and given by (not in dB)

$$\begin{aligned} &1.80 \text{ (16-QAM)}, \quad 1.51 \text{ (16-STAR)}, \\ &2.33 \text{ (64-QAM)}, \quad 2.10 \text{ (64-STAR)}. \end{aligned} \quad (19)$$

Having defined the MSED and PAR, we now define the total efficiency (TE) as $[\text{TE}]_{\text{dB}} = [\text{MSED}]_{\text{dB}} - [\text{PAR}]_{\text{dB}}$, which serves as a simple measure for evaluating the MSED and the PAR in the same framework.

3) *Comparison Results*: Based on the simulations, we have obtained MSED, PAR, and TE for given constellations and they are tabulated in Table II. It is observed that the highest TE can be achieved by the square 16-QAM with the proposed TS, which even outperforms the TE of unshaped 16-QAM evaluated prior to pulse-shaping by 1.6 dB. It is also interesting to observe that the square QAM in general performs better than the star QAM in terms of *both* MSED and PAR (except for P-16-STAR, which has much lower PAR but has a considerable MSED penalty).

VII. CONCLUSION

In this paper, we have applied trellis shaping to the PAR reduction of single-carrier QAM signals and demonstrated that the proposed shaping metric achieves significant PAR reduction together with noticeable shaping gain without sacrificing bandwidth efficiency. Conventionally, a QAM constellation with large MSED tends to have large fluctuation of the envelope and yields a high PAR. However, as demonstrated by our results on various QAM constellations, with the application of the TS, there is no explicit dependency between the MSED of

⁵Since this kind of PAR is calculated based on energy, it should be interpreted as a peak-to-average *energy* ratio.

TABLE II

COMPARISON OF SQUARE AND STAR QAM CONSTELLATIONS IN TERMS OF MSED AND PAR. ALL VALUES ARE WRITTEN IN dB. “SHAPED” AND “UNSHAPED” REFER TO WHETHER OR NOT THE TS IS APPLIED.

size M	shape	unshaped (no pulse-shaping)			unshaped			shaped		
		MSED	PAR	TE	MSED	PAR	TE	MSED	PAR	TE
16	16-QAM	-4.0	2.6	-6.6	-4.0	7.0	-11.0	-1.9	3.1	-5.0
	A-16-STAR	-5.5	1.8	-7.3	-5.5	6.6	-12.1	-2.6	4.6	-7.2
	P-16-STAR							-5.5	2.5	-8.0
64	64-QAM	-10.2	3.7	-13.9	-10.2	7.1	-17.3	-9.3	2.8	-12.1
	64-STAR	-13.8	3.2	-17.0	-13.8	7.4	-20.2	-11.1	3.4	-14.5

the employed QAM constellation and the resulting PAR. Our comparison has revealed that even square QAM constellations can achieve very low PAR.

Future work includes optimization of the constellation shape under the TS constraint and its application to single-carrier frequency domain equalization (SC-FDE) systems over frequency-selective fading channels.

REFERENCES

- [1] D. Wulich and L. Goldfeld, “Bound of the distribution of instantaneous power in single carrier modulation,” *IEEE Trans. Wireless Commun.*, vol. 4, pp. 1773–1778, July 2005.
- [2] A. Tomlinson, A. Ambroze, and G. Wade, “Power and bandwidth efficient modulation and coding for small satellite communication terminals,” in *Proc. IEEE ICC’02*, pp. 2943–2946, Apr. May 2002.
- [3] A. Ambroze, M. Tomlinson, and G. Wade, “Magnitude modulation for small satellite earth terminals using QPSK and OQPSK,” in *Proc. IEEE ICC’03*, pp. 2099–2103, May 2003.
- [4] S. L. Miller and R. J. O’Dea, “Peak power and bandwidth efficient linear modulation,” *IEEE Trans. Commun.*, vol. 46, pp. 1639–1648, Dec. 1998.
- [5] G. D. Forney, Jr., “Trellis shaping,” *IEEE Trans. Inform. Theory*, vol. 38, pp. 281–300, Mar. 1992.
- [6] I. S. Morrison, “Trellis shaping applied to reducing the envelope fluctuations of MQAM and band-limited MPSK,” in *Proc. Int. Conf. Digital Satellite Commun. (ICDSC’92)*, pp. 143–149, May 1992.
- [7] M. Litzemberger and W. Rupprecht, “Combined trellis shaping and coding to control the envelope of a bandlimited PSK-signal,” in *Proc. IEEE ICC’94*, pp. 630–634, June 1994.
- [8] M. Tanahashi and H. Ochiai, “Near constant envelope trellis shaping for PSK signaling,” *IEEE Trans. Commun.*, vol. 57, pp. 450–458, Feb. 2009.
- [9] W. Henkel and B. Wagner, “Another application for trellis shaping: PAR reduction for DMT (OFDM),” *IEEE Trans. Commun.*, vol. 48, pp. 1471–1476, Sept. 2000.
- [10] H. Ochiai, “A novel trellis-shaping design with both peak and average power reduction for OFDM systems,” *IEEE Trans. Commun.*, vol. 52, pp. 1916–1926, Nov. 2004.
- [11] T. T. Nguyen and L. Lampe, “On trellis shaping for PAR reduction in OFDM systems,” *IEEE Trans. Commun.*, vol. 55, pp. 1678–1682, Sept. 2007.
- [12] R. F. H. Fischer, R. Tzschoppe, and J. B. Huber, “Signal shaping for peak-power and dynamics reduction in transmission schemes employing precoding,” *IEEE Trans. Commun.*, vol. 50, pp. 735–741, May 2002.
- [13] Y. Zhou, P. J. McLane, and C. Loo, “Performance of predistorted APK modulation for one- and two-link nonlinear power amplifier satellite communication channels,” *IEEE Trans. Veh. Technol.*, vol. 54, pp. 629–638, Mar. 2005.
- [14] R. D. Gaudenzi, A. G. i Fabregas, and A. Martinez, “Performance analysis of turbo-coded APSK modulations over nonlinear satellite channels,” *IEEE Trans. Wireless Commun.*, vol. 5, pp. 2396–2407, Sept. 2006.
- [15] M. Chen and O. M. Collins, “Trellis pruning for peak-to-average power ratio reduction,” in *Proc. IEEE ISIT’05*, pp. 1261–1265, Sept. 2005.
- [16] T. May, H. Rohling, and V. Engels, “Performance analysis of Viterbi decoding for 64-DAPSK and 64-QAM modulated OFDM signals,” *IEEE Trans. Commun.*, vol. 46, pp. 182–190, Feb. 1998.
- [17] R. F. H. Fischer, *Precoding and Signal Shaping for Digital Transmission*. John Wiley & Sons Inc, 2002.

PLACE
PHOTO
HERE

Makoto Tanahashi (S’07) was born in Gifu, Japan, in December 1983. He received the B.E. and M.E. degrees in electrical engineering from Yokohama National University, Japan, in 2006 and 2008, respectively. He is currently working toward the Ph.D. degree in the same university. Since April 2009, he has been a Research Fellow (DC) of the Japan Society for the Promotion of Science.

He was a recipient of a Student Paper Award from the Telecommunications Advancement Foundation in 2009.

PLACE
PHOTO
HERE

Hideki Ochiai (S’97–M’01) received the B.E. degree in communication engineering from Osaka University, Osaka, Japan, in 1996, and the M.E. and Ph.D. degrees in information and communication engineering from the University of Tokyo, Tokyo, Japan, in 1998 and 2001, respectively.

From 1994 to 1995, he was with the Department of Electrical Engineering, University of California (UCLA), Los Angeles, CA, under the scholarship of the Ministry of Education, Science, and Culture.

From 2001 to 2003, he was a Research Associate at the Department of Information and Communication Engineering, the University of Electro-Communications, Tokyo, Japan. Since April 2003, he has been with the Department of Electrical and Computer Engineering, Yokohama National University, Yokohama, Japan, where he is currently an Associate Professor. From 2003 to 2004, he was a Visiting Scientist at the Division of Engineering and Applied Sciences, Harvard University, Cambridge, MA.

Dr. Ochiai currently serves as an Editor for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS.