Trellis Shaping with Flexible Control of Peak and Average Power for Single-Carrier High-Order QAM

Makoto Tanahashi and Hideki Ochiai

Department of Electrical and Computer Engineering - Yokohama National University

Yokohama, Kanagawa 240-8501, Japan

Email: makoto@ochiailab.dnj.ynu.ac.jp, hideki@ynu.ac.jp

Abstract—We extend our trellis shaping approach of reducing peak-to-average power ratio (PAR) of single-carrier PSK systems to high-order quadrature amplitude modulations (QAMs). Simulation results show that, with only one-bit redundancy per transmitted symbol, significant PAR reduction can be achieved. For example, band-limited transmission with PAR below 3 dB can be achieved with a square 64-QAM after pulse shaping with a roll off factor of as low as 0.1. Furthermore, the newly developed shaping metric offers a capability of simultaneously reducing the average power. The reduction of the PAR and average power can be flexibly controlled by adjusting a parameter associated with the shaping metric.

I. INTRODUCTION

In the future wireless communications systems, one of the emerging issues is to achieve higher data rate with limited spectrum resource. Bandwidth-efficient air interference such as high-order quadrature amplitude modulations (QAMs) and orthogonal frequency division multiplexing (OFDM) is therefore seen as an enabling technology. However, these modulation techniques are prone to have a large signal dynamic range or high peak-to-average power ratio (PAR). If the PAR is large, operation of a power amplifiers (PA) with a large back off is required, which significantly degrades the PA efficiency and thus reduces the operation time by battery. In order to achieve a high PA efficiency, a PAR reduction technique at the baseband stage therefore plays an important role. In the literature, the PAR issue has been extensively addressed in the context of OFDM signaling, but the single-carrier (SC) signals also suffer from high PAR in the high bandwidth efficiency regime: using a high-order QAM and a pulse-shaping filter with nearly rectangular spectrum shape considerably increases the resulting PAR, typically higher than 7 dB. In this paper, we therefore aim at PAR reduction of strictly band-limited high-order SC-QAM.

The trellis shaping (TS), originally proposed by Forney [1] for shaping gain (i.e., reduction of average power), has also found its application in the PAR reduction. In the literature, its applications to SC are found in [2, 4] and those for OFDM are also found in [5, 6] (see also references therein). However, an attempt to the PAR reduction of SC-QAM has so far been made very few [2], and the achievable PAR reduction performance was not significant. The precoding approach of [3] shares a concept similar to the TS, but it does not provide a significant PAR reduction.

Recently, in [4], the authors have proposed a new TS design for SC that takes into account pulse-shaped continuous waveform, with main focus on phase shift keying (PSK) signaling. This paper extends our work of [4] to more intricate high-order QAM scenarios and propose a high bandwidth efficient QAM system with a PAR adaptation capability. The TS of [4] consists of 1) replicating the pulse-shaped (i.e., interpolated) waveforms as opposed to discrete symbols [2]; 2) simple branch metric definitions for each sample of the replicated waveform. Since the process for generating the replica of the pulse-shaped waveforms is common to PSK and QAM, the required modification for QAM is reduced to the design of a suitable branch metric. Nevertheless, it should be noted that since QAM inherently has multiple amplitude levels, this modification may be nontrivial. Specifically, the average power of QAM signals can vary before and after shaping process, which makes the systematic design for PAR reduction difficult, considering that the PAR is a ratio of the peak power over the average power. Conversely, however, by exploiting this fact we demonstrate that it is possible to design a shaping metric that is capable of reducing not only the PAR but also the average power (i.e., achieving shaping gain [1]) simultaneously, similar to [3] but with much larger gain. The trade-off between the reduction of the PAR and average power can be flexibly controlled by a parameter associated with the shaping metric.

We mainly focus on widely used square and cross QAMs which have a large minimum Euclidean distance (MED) in conjunction with trellis coded modulation (TCM) scheme, but also consider circular (star) QAMs which have a smaller MED but with inherently lower PAR. Comparing the shaping results for these two types of constellations turns out that, when our TS is applied, the square QAM is preferable in terms of both the MED and the PAR. Notations used in this paper follow those in [4] and the reader is referred to [4] for detailed description.

II. CODED AND SHAPED QAM MODULATION

A. Square and Cross 64, 128, and 256-ary QAMs

Fig. 1 illustrates the baseline communication system model for SC-QAM with TS considered in this paper. As illustrated in the figure, the most significant bits (MSBs) s that have a relatively large MED within each subset (after the set partitioning) are used for shaping, and the least significant bits

Fig. 1. End-to-end communication model with trellis shaping and trellis coded modulation for square and cross QAMs.

(LSBs) u that have a small MED are protected by channel codes. We note that instead of this parallel concatenation of the TS and coding, a serial concatenation is investigated in [7], motivated by the idea of serially concatenated turbo codes. This approach enables us to exploit shaping for error correction, but its extension to high-order QAMs appears to be challenging considering the requirement for decoding complexity. We thus restrict our attention to the parallel concatenation shown in Fig. 1 for square and cross QAMs.

In the case of 64-QAM, we assign both s and u 2-bit information each. For higher order QAM cases, we consider additional information bits t , subject to neither shaping nor coding, as denoted by the dotted paths in Fig. 1. Let us denote the order of QAM constellation and the number of bits that constitute one constellation symbol by M and $m = \log_2 M$, respectively. Then the overall information rate [bit/symbol] of this shaped case is given by $m - 2$ [bit/symbol]. For comparison, we also consider an unshaped transmission case that imposes only one-bit coding redundancy and thus has an information rate of $m - 1$ [bit/symbol].

B. Star 64-ary QAM

As opposed to the square or cross QAMs where set-partition labeling is established with the TCM, the bit labeling of star QAMs is typically split into phase and amplitude [8]. As an example, we consider the 64-ary star QAM presented in [8] (called 64-STAR in this paper). The 64-STAR can be viewed as a combination of four concentric 16-ary PSK rings having different radii, where two bits determine the amplitude and the remaining four bits determine the phase of the 16-PSK ring. We adopt the optimum ring ratio given in [8]. The shape of this constellation will be later shown along with the simulation results. Note that we do not explicitly consider combination of the 64-STAR with error correcting codes as our main objective is to investigate the shaping capability. We assign the two amplitude bits and one of the phase bits to the shaped bits (corresponding to $\tilde{s}$ in Fig. 1). This bit mapping design has been found effective by our simulation.

C. SC Signaling with TS for Controlling Waveforms

The notation for the TS is identical to that given in [4], except that now there is a channel coding in parallel with

shaping process as illustrated in Fig. 1. In what follows, we briefly review the underlying principle of the TS for controlling waveforms.

In Fig. 1, the multiplexed bit sequence $[z, t, c]$ where

$$z = \tilde{s} + v = \tilde{s} + xG_s \quad (1)$$

is mapped onto 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

$$s(t) = \sum_{l=0}^{N-1} S_l g(t - lT_s), \quad (2)$$

where T_s denotes the symbol rate [symbol/sec] i.e., the Nyquist interval. We assume $g(t)$ to be a square-root raised-cosine with an effective (non-zero) duration spanning over K_s symbols. In (1), the x , called *waveform controlling sequence*, can be chosen arbitrary [4] and thus has capable of controlling the symbol sequence S or the baseband signal $s(t)$. Since the G_s is a generator matrix of any convolutional code, the Viterbi algorithm (VA) can be used for the selection of x , but with a suitable metric definition that can efficiently reduce the PAR.

D. Shaping Process

This subsection outlines how to select an optimum controlling sequence x . The discrete-time representation of the $s(t)$ of (2) is obtained by sampling $s(t)$ with a time interval $\Delta T = T_s/N_s$, 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 $g[n] \triangleq g(n\Delta T)$. Upon determining the l th symbol S_l , let us consider the signal samples $s[n]$ truncated over one Nyquist symbol interval from $n = (l - \frac{K_s}{2})N_s$, i.e.,

$$\begin{aligned} s_l[\tilde{n}] &\triangleq s\left[\tilde{n} + \left(l - \frac{K_s}{2}\right)N_s\right] \quad \text{for } 0 \leq \tilde{n} < N_s \\ &= \sum_{k=-K_s+1}^0 S_{l+k} g\left[\tilde{n} - \left(k + \frac{K_s}{2}\right)N_s\right]. \end{aligned} \quad (4)$$

Since our primary interest is to control the peak and average power characteristics, (4) is transformed into an instantaneous power form as $p_l[\tilde{n}] = |s_l[\tilde{n}]|^2$. We can observe that $p_l[\tilde{n}]$ is a function of the l th symbol S_l and its $(K_s - 1)$ previous symbols. The combination of these symbols is controlled by the corresponding part of $\mathbf{x}$, i.e., $\mathbf{x}_l \triangleq \{x_{l-(K_s-1)-(K-1)} \dots, x_{l-K_s+1}, \dots, x_{l-1}, x_l\}$, where K is the constraint length of $\mathbf{G}_s$. Since the size of $\mathbf{x}_l$ is $K_s + (K - 1)$, there are $2^{K_s+(K-1)}$ possible realizations of $p_l[\tilde{n}]$.

The VA proposed in [4] calculates the branch metric:

$$B_l \triangleq \sum_{\tilde{n}=0}^{N_s-1} \mu(p_l[\tilde{n}]) \quad (5)$$

for all the patterns at the l th interval, where $\mu(\cdot)$ is a *metric function* that quantifies the PAR from a given instantaneous power sample. With this branch metric, the VA finds a path with the minimum sum metric. As a result, the selected waveform controlling sequence $\mathbf{x}$ associated with the chosen path is an optimum controlling sequence that satisfies

$$\mathbf{x} = \arg \min \left[\sum_{l=0}^{N-1} B_l \right] = \arg \min \left[\sum_{n=0}^{N_s N-1} \mu(p[n]) \right], \quad (6)$$

where $p[n] = |s[n]|^2$. The most important observation here is that, with a suitable definition of $\mu(\cdot)$, we can arbitrary shape band-limited (i.e., continuous) waveforms. Also, it should be noted that since the equations from (3) to (6) are not restricted to a particular constellation shape, we can apply this shaping process to any QAM constellation. Our next task is to design $\mu(\cdot)$ for a given QAM constellation.

We here note the complexity of this shaping process briefly. In the VA, the number of branch metric computations for one transmitted symbol is given by $2^{K_s+(K-1)}$. Nevertheless, this can be considered to be even feasible, since it does not depend on the modulation order M , whereas a similar approach given in [9] has the computation order proportional to M^{K_s} . Although the mitigation of complexity given in [4], such as curtailment of the observed duration, is applicable to the QAM cases in a straightforward manner, we do not discuss complexity mitigation in this paper, as our main focus is on evaluating achievable maximum shaping capability.

III. SHAPING METRIC FOR QAM

In the case of QAM, the average power should vary before and after shaping, and this effect should be taken into account for designing a shaping metric. In this section, we systematically develop a metric function suitable for QAMs.

A. Step 1: Estimating Peak Power Control Capability

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

$$\Pr[p > p_{peak}] = \varepsilon. \quad (7)$$

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

$$\mu_{\text{test}}(p) = \mathbb{I}_{p > p_{peak}}, \quad (8)$$

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

$$\varepsilon = \Pr[p[t] > p_{peak}] \approx \frac{N_a}{NN_s} \quad (9)$$

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

B. Step 2: Adjustment of PAR and 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 to an enhancement of MED, 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 the PAR and average power requirement. 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} :

$$\mu(p) = \begin{cases} C, & p > p_{peak} \\ |p - p_{ref}|, & \text{otherwise,} \end{cases} \quad (10)$$

where C is an arbitrary large constant. If we set $p_{ref} = p_{peak}$, (6) can be approximately rewritten as

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

where $p_{av} = \frac{1}{N_s N} \sum_{n=0}^{N_s N-1} p[n]$ 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

$$\mathbf{x} = \arg \min \left[\sum_{n=0}^{N_s N-1} p[n] \right] = \arg \min [p_{av}], \quad (12)$$

which means that the average power is simply minimized, similar to the original operation of the TS as average power

Fig. 2. BER performance of the shaped and unshaped 64-QAMs.

reduction [1]. 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 has high linearity, one can enhance MED (and hence noise margin) by setting p_{ref} 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.

IV. NUMERICAL RESULTS

A. Figure of Merit

The proposed TS is capable of reducing both the PAR and average power, but it also reduces the information rate. In order to make a fair comparison of this trade-off, we evaluate the PAR reduction and the shaping gain by imposing a constraint such that the comparing shaped and unshaped systems achieve the same information rate [bit/sec] for a given bandwidth [Hz]. Note that the achievable error rate performance for given signal-to-noise power ratio is almost identical between the two schemes as they employ the same constellation size. In Fig. 2, we show the bit error rate (BER) performances, where a convolutional code G_c with the constraint length 7 is used. As observed, there is little degradation which stems from the non-catastrophic error propagation effect of the syndrome H_s^T (see [7]). In the following, $T_s^{(s)}$, $\mathcal{E}^{(s)}$, and $\alpha^{(s)}$ denote the symbol rate [symbol/sec], the average energy of one QAM symbol, and the roll-off factor of the pulse-shaping filter in the case of the shaped case. Likewise, $T_s^{(u)}$, $\mathcal{E}^{(u)}$, and $\alpha^{(u)}$ denote those of the unshaped case. Note that we assume the unshaped system has a unit power, i.e., $\frac{\mathcal{E}^{(u)}}{T_s^{(u)}} = 1$.

1) *Bandwidth Efficiency*: The shaped and unshaped systems should have the same information rate [bit/sec] occupying the same bandwidth [Hz], which is formulated as

$$\frac{m-2}{T_s^{(s)}} = \frac{m-1}{T_s^{(u)}}, \quad \frac{1+\alpha^{(s)}}{T_s^{(s)}} = \frac{1+\alpha^{(u)}}{T_s^{(u)}}. \quad (13)$$

Eliminating $T_s^{(s)}$ and $T_s^{(u)}$ leads us to the condition on the roll-off factor for the same bandwidth efficiency:

$$\alpha^{(u)} = \frac{m-1}{m-2}(1+\alpha^{(s)}) - 1. \quad (14)$$

2) *Peak-to-Average Power Ratio*: As described in the section III-A, we first fix the peak power p_{peak} by carrying out a simulation with the metric function μ_{test} . Next, with the metric function given in (10), we perform a computer simulation and obtain the reduced average power p_{av} . Since the peak power is fixed by p_{peak} , we can immediately obtain the PAR as

$$PAR^{(s)} = p_{peak}/p_{av}. \quad (15)$$

As a detailed figure of dynamic range characteristic, we also use the complementary cumulative distribution function (CCDF) of the normalized instantaneous power. The PAR reduction is evaluated as

$$\eta_p \triangleq \left[\frac{PAR^{(u)}}{PAR^{(s)}} \right]_{dB} = [PAR^{(u)}]_{dB} - [p_{peak}]_{dB} + [p_{av}]_{dB}, \quad (16)$$

where $PAR^{(u)}$ represents the PAR of the unshaped case and is defined as the normalized instantaneous power at CCDF = ε .

3) *Shaping Gain*: For the average power p_{av} obtained through simulation, we have the following equations:

$$\frac{\mathcal{E}^{(s)}}{T_s^{(s)}} = p_{av}, \quad \frac{\mathcal{E}^{(u)}}{T_s^{(u)}} = 1. \quad (17)$$

Substituting (13) to the above equations leads to the save of energy, i.e., the shaping gain η_s expressed as

$$\eta_s \triangleq \left[\frac{\mathcal{E}^{(u)}}{\mathcal{E}^{(s)}} \right]_{dB} = -[p_{av}]_{dB} - \left[\frac{m-1}{m-2} \right]_{dB}, \quad (18)$$

where the second term represents the loss in terms of information rate. The shaping gain can be treated as an increase of the MED (or equivalently noise margin), if the use of the same amount of energy is assumed. In this case, BER performance (Fig. 2) is shifted left by η_s . Note that, conventionally, the term “shaping gain” is defined in discrete symbol domain, whereas the shaping gain in our definition takes into account the band-limitation effect. Due to this difference, the theorem of ultimate shaping gain (1.53 dB) [1] does not hold in our results. From (16) and (18), it is apparent that the lower the PAR η_p , the lower the shaping gain η_s , and vice versa. Moreover, the total gain $\eta_p + \eta_s$ is independent of p_{av} (and hence the choice of the reference power p_{ref}).

B. Simulation Results

In the simulations, we always set $\alpha^{(s)} = 0.1$, while from (14) we have $\alpha^{(u)} = 0.38, 0.32$, and 0.28 for 64-QAM (64-STAR), 128-QAM, and 256-QAM yielding the bandwidth efficiencies of 3.64, 4.55, and 5.45 [bit/sec/Hz], respectively. The employed shaping encoder is $G_s = [1 + D^3 \quad 1 + D + D^2 + D^3 \quad 1]$. First, we determine p_{peak} . The reference parameter ε is set to $\varepsilon = 10^{-4}$ as an example. In this case, it has been found by simulation that the peak power

Fig. 3. CCDF of shaped/unshaped QAM signals. The values attached to the legends denote p_{peak} and p_{av} after shaping (in dB), from left to right.

Fig. 4. Shaping examples: baseband waveforms of 64-QAM and 64-STAR with $\alpha = 0.1$. Constellation points are also plotted.

is $p_{peak} = 1.6, 1.7$, and 1.7 for 64-QAM, 128-QAM, and 256-QAM, and $p_{peak} = 1.2$ for 64-STAR. Fig. 3 shows the CCDF characteristics of shaped and unshaped QAMs, where we set $p_{ref} = p_{peak}$, and therefore the maximal performance in terms of dynamic range reduction can be expected. From this figure, we can observe that the reduction of the signal dynamic range is significant. There are PAR reductions η_p of 3.2, 3.0, 3.4, and 2.4 dB for 64QAM, 128QAM, 256QAM, and 64STAR, respectively. See also shaped waveforms (trajectories) shown in Fig. 4 to confirm the effect of the proposed TS. Moreover, in Table I the relationship among p_{ref} , η_s , and η_p is summarized for the 64-QAM case. As observed, decreasing the reference power level p_{ref} results in larger η_s but smaller η_p and vice versa. It is noteworthy that even if we set p_{ref} in favor of the

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

reference power p_{ref}	1.6	1.2	1.0	0.8	0.6	0.4	0.0
PAR reduction η_p (dB)	3.2	3.1	3.0	2.6	1.8	1.2	0.9
shaping gain η_s (dB)	0.0	0.0	0.1	0.5	1.3	1.9	2.2

PAR reduction (i.e., high reference power level), there is no negative shaping gain observed, which stems from the fact that a small amount of reduction in the average power is preserved even in such a case. At low to moderate values of p_{ref} , we can achieve both PAR reduction and shaping gain at the same time. In the case of 64-STAR, the achievable maximal PAR reduction is not as significant as the 64-QAM as observed from Figs. 3. Although the significant reduction in the average power is in turn achieved, we have confirmed that the MED of the 64-STAR is still lower than that of the 64-QAM.

V. CONCLUSION

In this paper, we have applied the TS to the PAR reduction of single-carrier high-order QAM signals and demonstrated that the proposed shaping metric achieves significant PAR reduction together with noticeable shaping gain without sacrificing much of bandwidth efficiency. Conventionally, a QAM constellation with large MED tends to have a high PAR. However, as demonstrated by our results, with the application of the TS, there is no explicit trade-off between the MED of the employed QAM constellation and the resulting PAR. In future work, we will pursue the optimization of QAM constellation for various TS constraints.

ACKNOWLEDGMENT

This work was in part supported by the Strategic Information and Communications R&D Promotion Programme (SCOPE), the Ministry of Internal Affairs and Communications, Japan.

REFERENCES

- [1] G. D. Forney, Jr., "Trellis shaping," *IEEE Trans. Inform. Theory*, vol. 38, pp. 281–300, Mar. 1992.
- [2] 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.
- [3] 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.
- [4] M. Tanahashi and H. Ochiai, "A new trellis shaping approach for pulse-shaped PSK signals with almost constant envelope," in *Proc. IEEE ICC'07*, pp. 5598–5603, June 2007 (see also full paper version to be published in *IEEE Trans. Commun.*).
- [5] 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.
- [6] 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.
- [7] M. Tanahashi and H. Ochiai, "Iterative decoding of concatenated channel coding and trellis shaping based on markov model," in *Proc. IEEE Globecom'07*, pp. 3905–3909, Nov. 2007.
- [8] T. May, H. Rohling, and V. Engels, "Performance analysis of Viterbi decoding for 64-DAPSK and 64-QAMmodulated OFDM signals," *IEEE Trans. Commun.*, vol. 46, pp. 182–190, Feb. 1998.
- [9] 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.