Another Application for Trellis Shaping: PAR Reduction for DMT (OFDM)

Werner Henkel, *Member, IEEE*, and Björn Wagner

Abstract—A bound for the possible reduction of the peak-to-average ratio (PAR) dependent on the rate as well as possible practical procedures are presented. The idea of trellis shaping, originally used to minimize average transmit power in single-carrier systems, is applied to the problem of PAR reduction in multi-carrier transmission. Its impact, as a function of code rate, as well as design practicability is considered using metrics in both time and frequency domain.

Index Terms—Crest factor, DMT, OFDM, PAR, peak-to-average ratio, shaping

I. INTRODUCTION

ACCORDING to the central limit theorem, the superposition of many carriers in multitone signaling leads to a Gaussian-like density with a high peak-to-average ratio. Here we investigate the impact of trellis shaping on the PAR. We first determine the symbol-error rate due to clipping (dependent on the clipping level and the size of the modulation signal set) and derive theoretical limits for the achievable PAR dependent on the rate. While we concentrate on discrete multitone (DMT), i.e., on real time-domain signals, the method can be applied to orthogonal frequency-division multiplexing (OFDM), as well.

Due to space limitations, we reference only a few alternative approaches. A pure time-domain method for PAR reduction has been proposed in [1] which seems to yield good performance at moderate complexity. It is based on iteratively subtracting impulse-like functions in time domain at the positions where high peak values occur. The functions are defined using reserved carriers. Maybe one direct algebraic coding approach in [2] should also be mentioned that uses Golay sequences and describes them as cosets of Reed-Muller codes. Unfortunately, the code rate there is quite low. During the reviewing process of this paper, a new procedure has been developed using so-called analog codes to correct the effects of clipping at the receiver [3].

To obtain a first estimate of the impact of clipping on the error rate, we simplify the problem by idealizing the density to be exactly Gaussian.

For the calculation of the symbol-error probability, let us assume an ideal Gaussian i.i.d. density of the time-domain signal samples. The average power P of a QAM signal constellation with $M = 2^m$ points is known to be

$$P = \frac{a^2}{6}(M - 1),$$ (1)

with a denoting the minimum Euclidean distance between points.

Choosing a definition of the discrete Fourier transform (DFT) such that the average power in time and transform domain is identical (factor $1/\sqrt{N}$ for both transforms), the variance of the Gaussian time-domain signal equals P. Let the clipping level be $\hat{s}$ and the ratio of the clipping level over the rms value $\sqrt{P}$ be $\mathcal{L}_c = \hat{s}/\sqrt{P}$. We obtain the clipping noise power[1]

$$N = 2 \int_{\mathcal{L}_c\sqrt{P}}^{\infty} (u - \mathcal{L}_c\sqrt{P})^2 \frac{1}{\sqrt{2\pi P}} e^{-\frac{u^2}{2P}} du.$$ (2)

Assuming a Gaussian (i.i.d.) noise density, the probability p_d of crossing the half distance between two neighboring QAM signal points is

$$p_d = \frac{1}{2} \operatorname{erfc}\left(\frac{a/2}{\sqrt{2N}}\right)$$ (3)

The symbol-error probability for an inner QAM-constellation point is

$$p_s = 4 \cdot p_d - 4 \cdot p_d^2.$$ (4)

Following the derivation in the equations (1) to (4), in Fig. 1, we obtain results for the symbol-error probability p_s in relation to the bit allocation m and the ratio $\mathcal{L}_c$ of the clipping level over the rms value. Although we used a quite simplified model for the amplitude density of the time-domain samples which may be considered to be not too exact for small FFT sizes and $L > 3$, we can observe that there is a demand for only a slightly higher A/D and D/A resolution. This is of less practical importance compared to the increased power consumption, the nonlinear echoes, and the out-of-band power that are not treated in here.
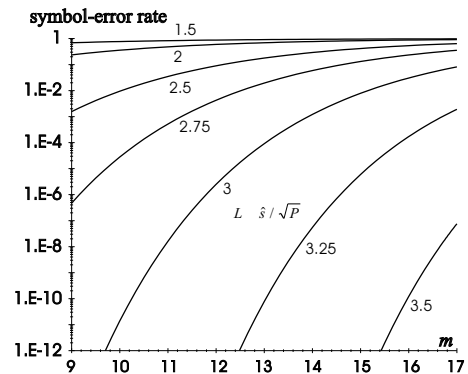

Fig. 1. Symbol-error probability P_s as a function of the number of bits per carrier m and the clipping ratio $\mathcal{L}_c = \hat{s}/\sqrt{P}$ under idealizing Gaussian i.i.d. assumptions

Now, we are presenting a theoretical limit for the achievable PAR dependent on the code rate R or, equivalently, the redundancy $1 - R$ offered for PAR reduction. We still do this based on the simplifying assumption of a Gaussian i.i.d. time-domain density. After this theoretical treatment we will present simulation results, too.

Let $\hat{s}$ again be the voltage limit. The average power belonging to a limited Gaussian density is given by

$$\bar{P} = \frac{2 \cdot \int_0^{\hat{s}} x^2 \frac{e^{-\frac{x^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma} dx}{2 \cdot \int_0^{\hat{s}} \frac{e^{-\frac{x^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma} dx} = \frac{2\sigma^2}{\frac{\sqrt{\pi}}{2}\operatorname{erf}(\frac{\hat{s}}{\sqrt{2}\sigma})} \int_0^{\frac{\hat{s}}{\sqrt{2}\sigma}} x^2 e^{-x^2} dx.$$ (5)

W. Henkel is with Telecommunications Research Center Vienna (FTW), A-1040 Vienna, Austria (e-mail: werner.henkel@ieee.org).

B. Wagner is with T-Nova, Bonn, Germany, (e-mail: Bjoern.Wagner@telekom.de).

[1] For a closed-form solution, see [4]

The probability of an amplitude lower than $\hat{s}$ is $\mathcal{P}_i = \text{erf}(\frac{\hat{s}}{\sqrt{2}\sigma})$. The probability that this is the case for all N time-domain samples is $\prod_{i=0}^{N-1} \mathcal{P}_i = \text{erf}^N(\frac{\hat{s}}{\sqrt{2}\sigma})$. The code rate R follows to be[2]

$$R = \frac{\log_2(2^{mN_{act}} \prod_{i=0}^{N-1} \mathcal{P}_i)}{mN_{act}} = 1 + \frac{N \cdot \log_2(\text{erf}(\frac{\hat{s}}{\sqrt{2}\sigma}))}{N_{act} \cdot m} , \quad (6)$$

$N/N_{act} = $ constant, being the ratio of the number of DFT components relative to the number of independently usable carriers ($N/N_{act} \approx 2$ for baseband transmission due to conjugacy constraints). Note that the rate R is dependent on m, i.e., the number of bits per carrier, not on N. A parametric plot using equations (5) and (6) of the principal achievable PAR $= \hat{s}^2/\bar{P}$ at a certain rate R is shown in Fig. 2. Although the derivation was based on some simplification, the results give the indication that a quite low percentage or redundancy (2 %) should be sufficient to reduce the PAR to the value of single-carrier QAM (or CAP, Carrierless AM/PM) covering the same frequency band. However, this does not tell, how the reduction can be achieved. Note that the lower limit of the PAR is 3. This is the PAR of a rectangular distribution, which is the approximation of the clipped Gaussian density for very low limit $\hat{s}$.[3] In Fig. 2, simulation results for $N_{act} = 15, 63, 255$ and $N = 32, 128, 512$, respectively, have been included; they are almost identical and hence indistinguishable. In the simulations, we also limit the voltage at a certain level $\hat{s}$ and computed the PAR and corresponding rate R when discarding sequences with peaks above $\hat{s}$. We see that the theoretical results due to the idealistic assumptions are somewhat far from the outcome of the simulations. Nevertheless, important properties like the independence of N could be confirmed. Although the simulations have only been carried out down to R around 0.8, the lower PAR limit of 3 seems to hold for the simulation outcomes as well. Especially, the statement still holds that only a very low redundancy should be required to achieve a PAR below the one of single-carrier QAM.

We did not consider oversampling or interpolation between the time-domain samples resulting from the N-point IDFT, which may lead to an increase in the PAR values, resulting in some degradation in the PAR(R) relation of Fig. 2. For a discussion of the effect of, e.g., a Lagrange interpolation, the reader is referred to [7].
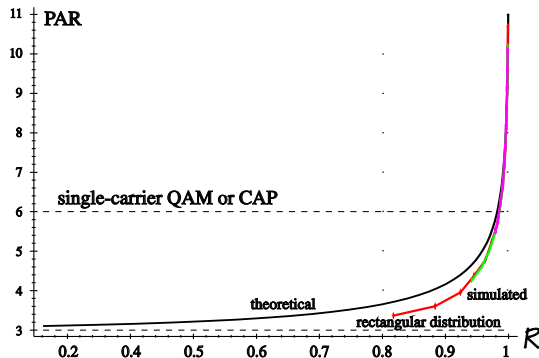

Fig. 2. Achievable peak-to-average ratio (PAR, linear scale) at a certain code rate R with 16-QAM ($m = 4$) on the carriers (theoretical and simulation results)

II. TRELLIS SHAPING FOR PEAK LIMITATION

We first give a short introduction to trellis shaping. For a detailed treatment, the reader is referred to Forney's paper [8].

In trellis shaping, a valid code sequence of a convolutional code is added (modulo 2) to a data sequence (see, Fig.3). The code sequence is chosen according to some criteria using a Viterbi algorithm. In the

[2]For a treatment of the complex OFDM case, see [5] and [6].

[3]Derivations of the PAR for single-carrier modulation and for a rectangular density can be obtained from the authors or by visiting http://www.ftw.at

original application described by Forney, e.g., the optimization criterion was the average power.

Let H be the parity-check matrix (syndrome former) of the convolutional code. Then, together with a valid code sequence y, we obtain $y \cdot H^T = 0$. This allows to eliminate the superimposed code sequence at the receiver side. However, hereto it is necessary to preprocess the information sequence i with the left inverse of the syndrome former $z = i \cdot (H^T)^{-1}$. With this, the information can be retrieved:

$$z' \cdot H^T = (z \oplus y) \cdot H^T = (z \cdot H^T) \oplus \underbrace{(y \cdot H^T)}_{=0} = z \cdot H^T = i . \quad (7)$$

Unlike Forney's proposal for trellis shaping, according to Fig. 3 we apply a multi-dimensional shaper that sequentially influences the last partition in a binary partition tree. For the multitone application, a DFT framing has to be introduced after the QAM mapping. The principal structures of the one- and multi-dimensional shapers are quite the same. However, the matrix dimensions of the syndrome former and its inverse are different. Instead of 2×1 and 1×2 we have $n \times (n-k)$ and $(n-k) \times n$ for H^T and $(H^T)^{-1}$, respectively. We apply the shaping in the last partition since this is the bit position with the biggest possible change in the signal-point location in a 2^m-QAM or 2^m-PSK alphabet.
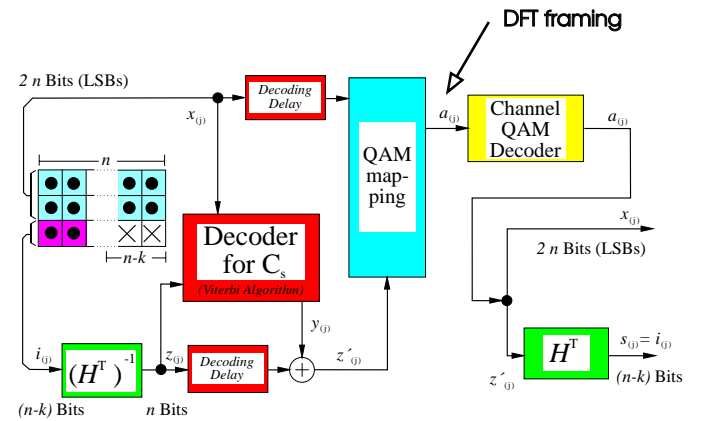

Fig. 3. Multidimensional Trellis Shaping

The well-known $(5, 7, 7)_8$ code may serve as an example for a possible shaping code. The corresponding matrices are

$$
\begin{aligned}
G &= \begin{pmatrix} 1+D^2 & 1+D+D^2 & 1+D+D^2 \end{pmatrix} \\[4pt]
H^T &= \begin{pmatrix} 1+D+D^2 & 1+D+D^2 \\ 1+D^2 & 0 \\ 0 & 1+D^2 \end{pmatrix} \\[4pt]
H^{T-1}_1 &= \begin{pmatrix} 0 & \frac{1}{1+D^2} & 0 \\ 0 & 0 & \frac{1}{1+D^2} \end{pmatrix} \\[4pt]
H^{T-1}_2 &= \begin{pmatrix} \frac{1}{1+D+D^2} & 0 & \frac{1}{1+D^2} \\ \frac{1}{1+D+D^2} & \frac{1}{1+D^2} & 0 \end{pmatrix}
\end{aligned}
\quad (8)
$$

Two possible inverse syndrome formers have been given. We also checked codes with construction principles other than the maximum free Hamming distance, but did not see better results.

For the desired application, we subdivide the DFT vector into blocks[4] of length n. The rate of the convolutional code is chosen to be k/n, which means an overall redundancy of $1 - R = k/(n \cdot m)$.

[4]A method called 'partial transmit sequences' described in [5], [9] and the iterative method described in [10] also use some block structure

As an optimization criteria inside the Viterbi algorithm, we could think of metrics in time or DFT domain. We describe possible approaches in the following two sections.

A. Metric in Time Domain

As a possible metric let us consider the peak power itself. This metric, however, does not fulfil the typical requirement for a metric inside the Viterbi algorithm, namely to be additive.

The time-domain peak power has to be determined for every path segment in the trellis and the time-domain vector needs to be updated for every additional block according to

$$f_{\kappa_\nu} = f_{\kappa_{\nu-1}} + \sum_{l=(\nu-1)n+1}^{\nu n} F_l e^{j\frac{2\pi}{N}l\kappa} \left(+ F_l^\star e^{j\frac{2\pi}{N}(N-l)\kappa} \right) . \quad (9)$$

The terms in brackets are for baseband transmission. Conjugacy constraints have to be fulfilled in order to obtain a real time-domain signal. For a four-state $1/n$-rate shaping code this means that in total eight DFTs ($4 \cdot 2$ trellis paths) instead of only one FFT have to be computed which means a relatively high complexity. There may, however, be a chance to re-express the DFT/FFT as has been mentioned in [5] to reduce the complexity. Despite the high complexity, the procedure may still be of interest for broadcast applications because then the complexity has to be installed only once at the transmitter side. The additional complexity due to trellis shaping at the receiver side can be neglected. Results of trellis shaping with the peak power in time domain as the Viterbi metric are given in Fig. 4. The strong limiting effect of the shaping becomes obvious. The PAR results are nearly as good as for QAM/CAP, having a PAR of 6 ($\hat{s} = 2.45 \cdot \sigma$). The dashed vertical lines at $\pm 2.65\sigma$ are valid for the shaping-code rate 1/3. Note that the total redundancy for the shaping is the shaping-code rate divided by the number of bits of the signal sets m, i.e., for 16-QAM, Fig. 4 shows results with redundancies of 1/8, 1/12, 1/16, and 1/32 (code rates 1/2, 1/3, 1/4, 1/8). The results do not take any interpolation filter into account. One should note, however, that overshooting caused by the filter will worsen the PAR.
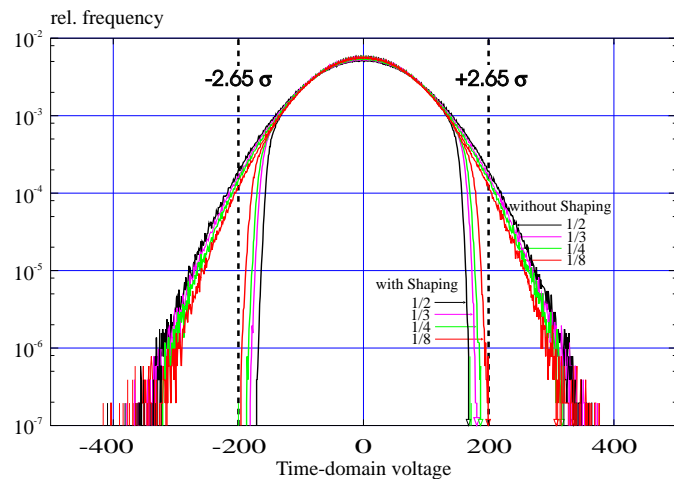

Fig. 4. Histogram for shaping with the time-domain metric (16-QAM, $N = 512$; shaping-code rates $1/n$)

B. Metric in DFT Domain

In DFT domain, we have the problem that unlike Parseval's formula for the average power, there is no simple expression for the peak power that would relate time and DFT domain.

As a simple indication for a high voltage peak in time domain, however, one may at first sight concentrate on cases where the phases

of the carrier symbols $F_l = |F_l|e^{j\phi_l}$ are nearly a linear function of the frequency and would eliminate the phase term inside the IDFT $f_{\kappa_\nu} = \sum_{l=0}^{N-1} F_l e^{j\frac{2\pi}{N}l\kappa}$. A large variance from an estimated linear function (and also from parallels at a distance of π) would then be some indication for low peak values. However, practical investigations show that estimating the peak position by, e.g., linear regression is too unreliable.

Another procedure, which works at least for small signal sets and small carrier numbers is based on the tabulation of block-transition metrics in DFT domain. There, we separately investigate the influence of two (equal-sized) segments of the DFT frame on the time-domain signal, setting remaining components of the DFT frame to zero. We speak of 'block-transitions' since the relation of the contents of the segments to one another is of importance for the time-domain amplitudes. For example, repeating the pattern of one segment in the other segment will often cause higher amplitudes in time domain than independent choices. When subdividing the DFT block into equal-sized segments, the time-domain samples may be computed by adding the per-segment DFTs. This means that the influence of the segments on the time-domain result is additive. For defining a shaping metric, as has already been mentioned, we consider pairs of segments, which may lead to counting segments multiply, but still the relation to the time-domain result is additive. With four segments, we may, e.g., consider the combinations 1:2, 2:3, 3:4, and 4:1, which would mean taking all the segments twice. When using the Viterbi algorithm and starting from the first segment of the DFT frame, we may have different counts of using the different segments. For example, when considering only neighboring segments for the metric computation (1:2, 2:3, and 3:4) the border segments are only counted once, while the others appear twice. This will not be taken into account.

As metric we actually choose the absolute value of the time-domain peak voltage[5]. Ignoring that the time-domain peaks that correspond to different segments may have different positions and signs, we still define this to be an additive metric. This additive metric can easily be used inside the Viterbi algorithm.

To conclude, the metrics are determined as the time-domain peak amplitude resulting from certain block transitions, setting the remaining components of the DFT frame to zero. All such possible block transitions together with the corresponding (normalized) peak amplitudes are tabulated.

The normalized peak amplitude can either be used directly as a metric inside the Viterbi algorithm, meaning that only the current transition (neighboring segments) is considered, or linear combinations of transition metrics may be used (see, Fig. 5). For example, for a decision upon block i, all transitions $0 \rightarrow i$, $1 \rightarrow i$, ..., $i - 1 \rightarrow i$ could be incorporated. A part of a normalized metric table for one transition is shown in Table I.

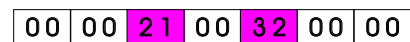Fig. 5. Block transitions

Fig. 6 shows all the $M^{2B} = 2^{2mB}$ (B: block size) normalized metrics of one transition for $m = 2$, $B = 2$. The size of the table is determined by the signal-set size and the segment size (times 2), which has to grow with N to obtain a significant PAR reduction. The use of the procedure is thus restricted to only small alphabets and small carrier numbers.

For illustration we show a few steps of the Viterbi algorithm in Fig. 7. The underlying modulation alphabet is 4-QAM and as a shaping code, we use the $(5, 7)_8$ code. As an inverse syndrome former, we have

---

[5]Using the time-domain peak power, instead, leads to only slightly worse results.

TABLE I

TRANSITION METRIC TABLE FOR ONE SINGLE BLOCK TRANSITION FOR BLOCKS OF
SIZE 2 AND 4-QAM (NUMBERED 0-3) AS SIGNAL SET

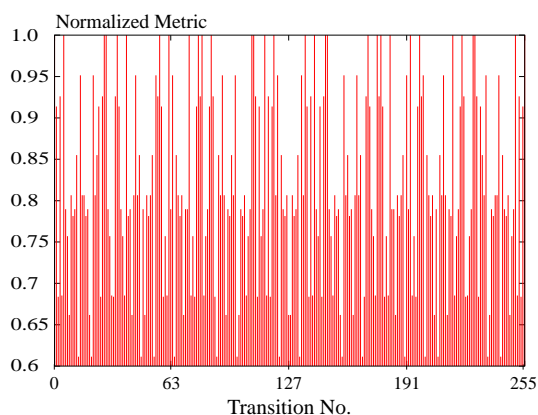| Transition	Normalized Metric
$00 \longrightarrow 00$	1.000
$10 \longrightarrow 00$	0.914
$20 \longrightarrow 00$	0.684
$30 \longrightarrow 00$	0.926
$01 \longrightarrow 00$	0.685
$11 \longrightarrow 00$	1.000
$\vdots$	$\vdots$
$33 \longrightarrow 00$	0.807
$00 \longrightarrow 10$	0.807
$10 \longrightarrow 10$	0.781
$\vdots$	$\vdots$
$33 \longrightarrow 33$	1.000

Fig. 6. Block-transition metrics for blocks of length 2 with 4-QAM signal sets

chosen one with leading zeros. This leads to zeros in the first position of the lower line segments, which is the line to be modified by the shaping sequence. The trellis paths are labeled by the coder input (irrelevant for shaping purposes), the coded binary sequence, and the QAM outcomes of the addition of the lower line output of the inverse syndrome former and the components of the binary shaping code sequence. After the first segment (DC component is zero, anyway) the metric is initialized with zero. Then for the next segment, we use the metric table with transitions between segments 1 and 2. Correspondingly, for the next segment, the table with transitions between segments 2 and 3 is chosen. In this example we have only used tables for neighboring segments, i.e., no linear combination of different transition metrics has been applied.

The advantage of the procedure is that the complexity for this block-transition shaping is extremely small because only a Viterbi algorithm for a four-state trellis needs to be computed ($\lfloor (N_{act} - 1)/n \rfloor \cdot 4$ add-compare-select operations) and no additional FFT or DFT is required.

The accumulated DFT-domain metrics are not in a direct simple algebraic relation to the peak power in time domain, but can offer some 'evidence' about the peak value. Thus, as expected, the results (see, Fig. 8) are a little worse than using direct transforms into time domain.

One referee pointed us to some further possibilities for a DFT-domain metric described in [11], [12], [13], which may be worth studying.
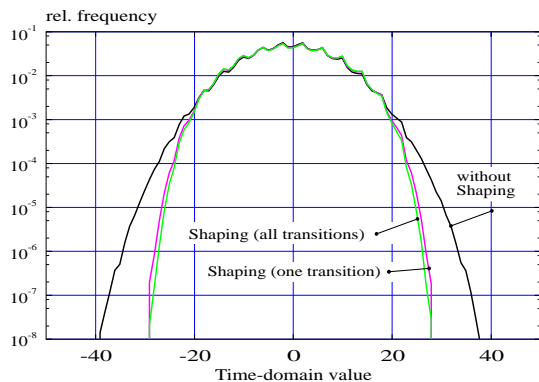

Fig. 8. Histogram for shaping with the DFT-domain metric (4-QAM, $N = 64$; shaping-code rate $1/3$)

III. CONCLUSIONS

It has been outlined that trellis shaping is in principal applicable to reduce the peak-to-average ratio of a multitone signal. A time-domain metric yielded quite good results with relatively high complexity. A DFT-domain metric was shown to be a worthwhile alternative with low complexity for small signal sets and carrier numbers. We showed results with shaping redundancies between 3 and 12.5%.

ACKNOWLEDGMENTS

REFERENCES

[1] Tellado, J., Cioffi , J.M., "PAR Reduction in Multicarrier Transmission Systems", *Delayed Contribution ITU-T 4/15*, D.150 (WP 1/15), Geneva, February 9-20, 1998.

[2] Davis, J.A., Jedwab, J., "Peak-to-Mean Power Control and Error Correction for OFDM Transmission using Golay Sequences and Reed-Muller Codes", *Electronics Letters*, Vol. 33, No. 4, Feb. 13, 1997, pp. 267-268.

[3] Henkel, W., "Analog Codes for Peak-to-Average Ratio Reduction", Proc. *3rd ITG Conference Source and Channel Coding*, Munich, Germany, January 17–19, 2000, pp. 151-155. (http://www.ftw.at)

[4] Mestdagh, D.J.G., Spruyt, P.M.P., "A Method to Reduce the Probability of Clipping in DMT-Based Transceivers", *IEEE Trans. on Communications*, Vol. 44, No. 10, Oct. 1996, pp. 1234-1238.

[5] Müller, St.H., Bäuml, R.W., Fischer, R.F.H., Huber, J.B., "OFDM with Reduced Peak-to-Average Power Ratio by Multiple Signal Representation", *Annals of Telecommunications*, Vol. 52, No. 1-2, Feb. 1997, pp. 58-67.

[6] Shepherd, S., Orriss, J., Barton, St., "Asymptotic Limits in Peak Envelope Power Reduction by Redundant Coding in Orthogonal Frequency-Division Multiplex Modulation" *IEEE Trans. on Communications*, Vol. 46, No. 1, Jan. 1998, pp. 5-10.

[7] Paterson, K.G., Tarokh, V.: "On the Existence and Construction of Good Codes with Low Peak-to-Average Power Ratios", *HP Technical Report*, HPL-1999-51.

[8] Forney, G.D., "Trellis Shaping", *IEEE Trans. on Information Theory*, Vol. IT-38, No. 2, pp. 281-300, March 1992.

[9] Müller, St.H., Huber, J.B., "OFDM with Reduced Peak-to-Average Power Ratio by Optimum Combination of Partial Transmit Sequences", *Electronics Letters*, Vol. 33, No. 5, Feb. 1997, pp. 368-369.

[10] Friese, M., "Multitone Signals with Low Crest Factor" *IEEE Trans. on Communications*, Vol. 45, No. 10, Oct. 1997, pp. 1338-1344.

[11] Eetvelt, P. van, Wade, G., Tomlinson, M., "Peak to Average Power Reduction for OFDM Schemes by Selective Scrambling", *Electronics Letters*, Vol. 32, No. 21, Oct. 10, 1996, pp. 1963-1964.

[12] Tellambura, C., "Upper Bound on Peak Factor of N-Multiple Carriers", *Electronics Letters*, Vol. 33, No. 19, Sept. 11, 1997, pp. 1608-1609.

[13] Tellambura, C., "Phase Optimisation Criterion for Reducing Peak-to-Average Power Ratio in OFDM", *Electronics Letters*, Vol. 34, No. 2, Jan. 22, 1998, pp. 169-170.

[14] Henkel, W., Wagner, B., "Verfahren zur Reduktion des Spitzen-Mittelwertverhältnisses der Leistung bei Mehrträgerübertragung", *DBP 196 25 054.4* (German patent application).

[15] Henkel, W., Wagner, B., "Trellis Shaping for Reducing the Peak-to-Average Ratio of Multitone Signals", Proc. *6th Benelux-Japan Workshop*, Essen, Germany, August 21-23, 1996, pp. 4.1-4.2.

[16] Henkel, W., Wagner, B., "Trellis-Shaping zur Reduzierung des Spitzen-Mittelwert-Verhältnisses bei DMT/OFDM", Proc. *OFDM-Fachgespräch*, Braunschweig, Germany, Sept. 10-11, 1996. (in German)

[17] Henkel, W., Wagner, B., "Trellis Shaping for Reducing the Peak-to-Average Ratio of Multitone Signals", Proc. *ISIT 1997*, Ulm, Germany, June 29 – July 4, 1997, p. 519.

4-QAM

DFT-domain vector
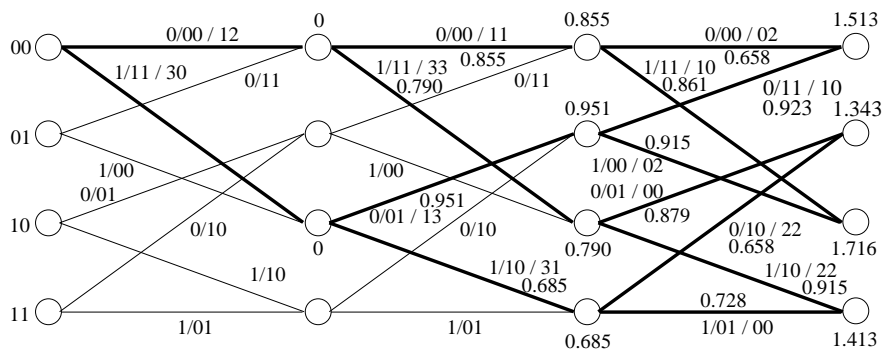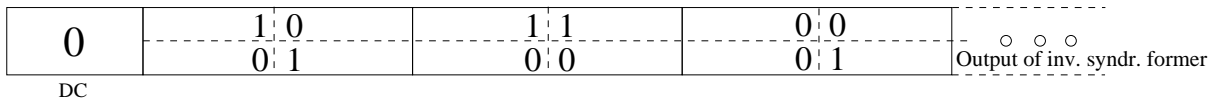
DC
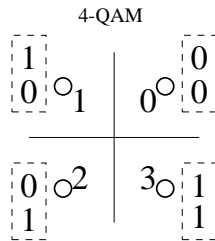
Output of inv. syndr. former

Fig. 7. Example of the metric computations carried out by the Viterbi algorithm of the trellis shaper