Trellis Coded CPM Type Space-Time Modulation Design Rules for Rayleigh Flat Fading Channel

Jan Sykora
Technicka 2
166 27 Praha 6
Czech Republic
Phone: +420 2 2435 3921
Fax: +420 2 2435 5829
Email: Jan.Sykora@feld.cvut.cz
http://radio.feld.cvut.cz
Abstract—A CPM type constant envelope space-time modulation is an attractive option especially for very high frequency and mobile applications due to its resistance to a nonlinear distortion. The paper analyzes the mean squared distance of the trellis coded CPM type constant envelope space-time modulated signal in a Rayleigh slowly flat fading spatial diversity channel with independent coefficients. It is shown that the distance evaluation depends on both—the modulator trellis and on the distance evaluation trellis. The latter has a special properties regarding to its free paths. These properties—a critical path and the distance increments behavior are the basis for the space-time trellis code design. We identify the trellis code (data to channel phase symbols mapping) design rules minimizing the mean squared distance. The critical path is shown to be the determining factor.

I. INTRODUCTION

Spatial diversity communication is currently one of the most targeted research topics. The most attractive is its potential for increasing the channel capacity in the bandwidth restricted environment. The topic is addressed from number of different viewpoints starting with the channel capacity evaluation (see e.g. [1], [2]) and going into the synthesis of the space-time modulation (e.g. [3]). However almost all resources have focused so far only on linear modulation schemes. Reference [4] belongs among rare exceptions. The nonlinear constant envelope modulations provide an attractive option from the implementation point of view especially on very high frequencies for indoor communication or for mobile devices with restricted power resources. Especially the resistance of such modulations to nonlinear distortion is very appealing. This paper is a follow up of the on going work started in [5]. The paper contains extended version of the derivations from [5] supplemented by additional results and correcting unfortunate numerical mistakes in the example application code of the [5].

We will focus on the problem by analyzing the distance properties of constant envelope space-time modulation and then by formulating the design rules for a trellis code maximizing the distance in Rayleigh spatial diversity channel with independent paths. Here, we will formulate the rules in terms of channel phase symbols properties unlike as it is developed in [4] where the rules are related to the transmitted signal itself. Our approach is more suitable for the design of the trellis code, i.e. data to channel phase symbols mapping.

The paper considers a general nonlinear constant envelope modulator with a finite correlation length. The nonlinearity of the modulation and the inherent memory of its phase expansion part inevitably makes the treatment more complicated then in the case of linear modulation. It is known that a class of binary constant envelope phase modulations can be equivalently expressed as a linear-like modulation using Laurent expansion (see [6]). However we will not proceed in this way in the paper. First because of the binary data constrain and also because of a lack (except for unity correlation length) of straightforward relation of the channel phase symbols to the data. This disqualifies the approach for the trellis code synthesis.

II. SYSTEM MODEL

A space-time (ST) modulated signal with constant envelope components is a vector signal

\[ s(t) = [s_1(t), \ldots, s_{N_t}(t)]^T \]

where \( N_t \) is the number of transmitter antennas. Each component is CPM-class modulated signal

\[ s_i(t) = e^{j2\pi \sum \varphi_{n,i} \beta(t-nT_S)} \]  

(1)

where \( \kappa \) is a positive real valued constant—modulation index, \( d_n \in \{0, 1, \ldots, (M_d - 1)\} \) is the data symbol, \( \sigma_n \in \{0, 1, \ldots, (M_{\varphi} - 1)\} \) is the state of the modulator and \( T_S \) is the symbol period. The whole data message is \( d = [d_0, \ldots, d_{N_t-1}]^T \). Variables \( q_{n,i} = q_{n,i}(d_n, \sigma_n) \) are discrete channel phase symbols. Function \( \beta(t) \) is a phase function (common to all \( i \)) obeying \( \beta(t) = 0, t \leq 0 \) and \( \beta(t) = 1/2, t \geq LT_S \). We also define \( q_n = [q_{n,1}, \ldots, q_{n,N_t}, \sigma_n]^T \) and \( q_n = q_{n,i}(\beta(t-nT_S) \) Modulator state \( \sigma_n \) obeys a state transition equation \( \sigma_{n+1} = \sigma(d_n, \sigma_n) \). We assume the

1 All vectors are column vectors. All derivations are carried out for complex envelope signals.
same modulation index $\kappa$ in all components of the signal. We also assume the unity and constant amplitude of the modulated signal over all components. In the next, we restrict our attention to the phase modulation functions with a finite correlation length $LT_s$, $L \in \mathbb{N}$.

The mapping between data and channel phase symbols $[d] \mapsto [q]$ represents the trellis code for the constant envelope ST modulation. The trellis code is a mapping between data and channel phase signal space. This needs to be distinguished from the case of linear ST modulation where the channel symbols are direct expansion of the modulated symbol in the signal space and therefore the trellis code is a mapping between data and signal space expansion of the modulated symbol.

The channel is considered to be a slowly flat fading one.

$$V_{ii'} = \int_{-\infty}^{\infty} (s_i^{(1)} - s_i^{(2)})(s_{i'}^{(1)} - s_{i'}^{(2)})^* dt.$$  

(6)

Defining $a_k = [a_{ik}, \ldots, a_{Mk}]^T$, $[A]_k = a_k$ and $V$ as the matrix with elements $V_{ii'}$ in $i$-th row and $i'$-th column, we can express this in a matrix form by

$$\rho_a^2 = \sum_{k=1}^{N_t} a_k^T Va_k = \sum_{k=1}^{N_t} tr(a_k a_k^H V^T) = tr(A^T VA^*).$$

(7)

Probability of the pairwise message error between transmitted message $\mathbf{d} = \mathbf{d}^{(1)}$ and the message $\mathbf{d}^{(2)}$ examined by the detector is for AWGN channel and for given particular $a_{ik}$

$$P_{2x,a} = \Pr\{\rho_{xa}(x, \mathbf{d}^{(2)}) < \rho_{xa}(x, \mathbf{d}^{(1)}) | \mathbf{d} = \mathbf{d}^{(1)}\} = Q\left(\rho_a \frac{1}{2\sqrt{N_0}}\right).$$

(8)

In order to get an upper bound on the average probability of the pairwise error we would have to perform the averaging over all values of $a_{ik}$, $P_{2x}^{(ub)} = \mathbb{E}[P_{2x,a}^{(ub)}]$. However, unlike for the case of a linear modulation treated for example in [3], this expression is mathematically difficult to further manipulate for our case of the nonlinear modulation when a simple relation to $g_{n,i}$ is required for the code synthesis. Instead, we average directly the squared bound

$$P_{2x}^{(ub)} = \frac{1}{2\sqrt{\pi}} \exp\left(-\frac{\rho_a^2}{8N_0}\right).$$

(9)

III. MEAN SQUARED DISTANCE EVALUATION TRELLIS

A. Mean squared distance

In this section, we derive an expression for the Euclidean distance for useful received signals corresponding to two different data messages $\mathbf{d}^{(1)}$ and $\mathbf{d}^{(2)}$. We denote the corresponding modulated signal components

$$s_i^{(1)} = e^{j2\pi x} \sum_{n=1}^{N_t} \Phi_{n,i}^{(1)}, s_i^{(2)} = e^{j2\pi x} \sum_{n=1}^{N_t} \Phi_{n,i}^{(2)}$$

(4)

and similarly $a_{ik}^{(1)}(t), a_{ik}^{(2)}(t)$. For the distance evaluation, the received signals $u_k(t)$ can be conveniently and equivalently organized in a sequence instead as a vector in parallel. This allows a simple expression for the squared distance of the signals conditioned by particular values $a_{ik}$ observed at the receiver

$$\rho_a^2 = \sum_{k=1}^{M} \int_{-\infty}^{\infty} |u_k^{(1)} - u_k^{(2)}|^2 dt = \sum_{k=1}^{N_t} \sum_{i=1}^{N_r} \sum_{i'=1}^{N_r} a_{ik} a_{ik}^* V_{ii'}$$

(5)

\footnote{Independent Identically Distributed.}

Of course there is no simple analytical relation between $P_{2x}^{(ub)}$ and $P_{2x}^{(ub)}$. However, we can physically interpret what is achieved by maximizing averaged $\rho^2$. This is in fact equivalent to minimizing $\mathbb{E} [\ln P_{2x,a}^{(ub)}]$ which can be easily seen from (9). A choice of the pairwise error probability logarithm minimization seems to be quite acceptable especially realizing that the mutual information and the channel capacity are closely related to the expectation of the error probability logarithm. Later in the paper, we...
will support this step by simulation results relating mutually \( \rho_{2}^{(ab)} \) and \( \rho_{2}^{(ab)} \) values.

Now we evaluate the average useful received signal squared distance over all values of \( a_{k} \). The averaging is performed over large number of data transfer sessions.

Considering the assumption of slow fading together with some kind of splitting the long message into smaller pieces, this also corresponds to the averaging the distance over the long data message split into independent smaller pieces. The average squared distance is

\[
\rho^{2} = E[\rho_{a}^{2}] = \sum_{k=1}^{N_{t}} \text{tr}(R_{ak}V^{T})
\]  

(11)

where \( R_{ak} = E[a_{i}a_{k}^{H}] \). Using the assumption of IID random variables \( a_{ik} \), we get \( R_{ak} = \sigma_{a}^{2}I \) and subsequently

\[
\rho^{2} = \sum_{k=1}^{N_{t}} \text{tr}(\sigma_{a}^{2}I^{T}) = \sigma_{a}^{2} \sum_{k=1}^{N_{t}} \text{tr}(V^{T}) = \sigma_{a}^{2} N_{r} \text{tr}(V).
\]

The trace of the matrix \( V \) is \( \text{tr}(V) = \sum_{i=1}^{N_{t}} V_{ii} \), where

\[
V_{ii} = \int_{-\infty}^{\infty} |x_{i}^{(1)}|^{2} + |x_{i}^{(2)}|^{2} - 2\Re[x_{i}^{(1)} \ast s_{i}^{(2)}] \ dt.
\]

(12)

A substitution of the modulated signal (4) into this expression and a normalization of the squared distance (12) by a suitable constant \( \rho^{2} = \rho^{2}/(\sigma_{a}^{2} N_{r} T_{S}) \) gives

\[
\rho^{2} = \sum_{i=1}^{N_{t}} \frac{1}{T_{S}} \int_{-\infty}^{\infty} 1 - \Re\left[ e^{j2\pi x \sum_{m=0}^{m-L} (\Phi_{n}^{(1)} - \Phi_{n}^{(2)})} \right] \ dt
\]

\[
= \sum_{i=1}^{N_{t}} \sum_{m=-\infty}^{\infty} \Delta \rho_{i}^{2}(m)
\]

(14)

The squared distance increments for the \( i \)-th branch are

\[
\Delta \rho_{i}^{2}(m) = 1 - \frac{1}{T_{S}} \int_{-\infty}^{T_{S}} 1 - \Re\left[ e^{j2\pi x \sum_{m=0}^{m-L} (\Phi_{n}^{(1)} - \Phi_{n}^{(2)})} \right] \ dt.
\]

(15)

Notice that \( \rho^{2} \) is the normalized value per one receiver antenna. A substitution for phase modulation functions and utilization of the finite correlation length assumption gives

\[
\Delta \rho_{i}^{2}(m) = 1 - \Re[\eta_{i}(m) \lambda_{i}(m)]
\]

(16)

where

\[
\eta_{i}(m) = \exp \left( j2\pi \kappa \sum_{n=-\infty}^{m-L} \Delta q_{m,n} \frac{1}{2} \right)
\]

(17)

and

\[
\lambda_{i}(m) = \frac{1}{T_{S}} \int_{0}^{T_{S}} e^{j2\pi x \sum_{n=m-L+1}^{m} \Delta q_{m,n, \beta}(t-nT_{S})} \ dt.
\]

(18)

The quantity \( \Delta q_{m,i} = q_{m,i}^{(1)} - q_{m,i}^{(2)} \) is the channel phase symbol difference. An overall squared distance increment reflecting all transmitted signals is

\[
\Delta \rho^{2}(m) = \sum_{i=1}^{N_{t}} \Delta \rho_{i}^{2}(m) = N_{r} - \sum_{i=1}^{N_{t}} \Re[\eta_{i}(m) \lambda_{i}(m)].
\]

(19)

If we consider a special case of high practical importance when the modulation index is a rational number \( \kappa = \kappa_{1}/\kappa_{2}, \kappa_{1}, \kappa_{2} \in \mathbb{N} \) and when channel phase symbols are drawn from the set \( q_{m,i} \in \{ \pm 1, \pm 3, \ldots, \pm (M_{q} - 1) \} \) then the differences are \( \Delta q_{m,i} \in \{ 0, \pm 2, \pm 4, \ldots, \pm 2(M_{q} - 1) \} \) and \( \sum_{m=-\infty}^{\infty} \Delta q_{m,i}/2 \) becomes an integer. The set of all possible \( \eta_{i}(m) \) becomes finite \( \eta_{i}(m) \in \{ \eta^{(i)} \}_{i=1}^{M_{q}} \). We can define

\[
\psi_{i}(m) = M_{q} \kappa \sum_{n=-\infty}^{m-L} \Delta q_{m,n} \frac{1}{2} \mod M_{q}
\]

(20)

\[
(\psi_{i}(m) \in \{ 0, 1, \ldots, (M_{q} - 1) \}) \text{ which is related to } \eta_{i}(m)
\]

by

\[
\eta_{i}(m) = \exp \left( j2\pi \frac{\psi_{i}(m)}{M_{q}} \right).
\]

(21)

We will assume this in the whole subsequent treatment.

\textbf{B. Distance evaluation trellis}

We can now define \( \Theta(m) = [\theta_{1}(m), \ldots, \theta_{N_{t}}(m)] \)

\[
\theta_{i}(m) = [\psi_{i}(m), \Delta q_{m-L+1,i}/2, \ldots, \Delta q_{m-1,i}/2].
\]

(22)

It can be easily seen that \( \Theta(m) \) and \( \Delta q_{m} = q_{m}^{(1)} - q_{m}^{(2)} \) can be interpreted as a \textit{state} and an \textit{input} of the distance evaluation procedure which itself possesses the memory. Compare this result with a typical trellis code design for linear modulations where the only memory needed to be considered is the one of the modulator trellis. The distance evaluation in that case is memoryless. The memory of the distance evaluation process is just the one that makes the modulator trellis code synthesis difficult for the nonlinear modulation.

The squared distance increment \( \Delta \rho^{2}(m) \) can be calculated having knowledge of the input \( \Delta q_{m} \) and the state \( \Theta(m) \). The quantity \( \Theta(m) \) will be referred as a \textit{distance evaluation state}. A state transition equation \( \theta_{i}(m+1) = \theta_{i}(m), \Delta q_{m,i} \) for \( i \)-th branch is clearly\(^{4}\)

\[
[\theta_{i}(m+1)]_{1} = ([\theta_{i}(m)]_{1} + M_{q} \kappa \theta_{i}(m)]_{2} \mod M_{q}
\]

\[
[\theta_{i}(m+1)]_{2} = [\theta_{i}(m)]_{3}
\]

\[
\vdots
\]

\[
[\theta_{i}(m+1)]_{L-1} = [\theta_{i}(m)]_{L-1}
\]

\[
[\theta_{i}(m+1)]_{L} = \Delta q_{m,i}/2.
\]

(23)

\(^{4}\text{x}_{i} \text{ denotes } i \text{-th component of the vector.}\)
The number of all possible distance evaluation states is

$$M_{\Theta} = ((2M_q - 1)^{j-1}M_q)^{N_t}. \quad (24)$$

The distance evaluation can be conveniently graphically represented by a path in a distance evaluation trellis. It is important to stress that the distance evaluation states and trellis are completely independent with the states and trellis of the modulator. The distance evaluation trellis serves as the distance evaluation tool. The particular form of the distance evaluation trellis depends only on the channel phase symbols difference and on $\kappa$, $N_t$ and $\beta(t)$. It is not directly related to the actual message data (only through $\Delta q_{m,i}$).

The distance evaluation trellis will be used later for an evaluation of the free distance and for a code design maximizing this free distance. From that point of view, the initial distance evaluation state is always $\Theta_{i}(0) = [0, 0, \ldots, 0]^T$, $i = 1, \ldots, N_t$ at the point$^5$ where the path in the modulator trellis splits for the first time for two different data messages $d^{(1)}, d^{(2)}$. This directly follows from the fact that the data and consequently the channel phase symbols are the same up to the point of the split ($m = 0$) for both messages and thus $\Delta q_{m,i} = 0$ for $m < 0$.

C. Example case

Now we present an example case as an illustration of the distance evaluation trellis concept. The phase function is

$$\beta(t) = \int_{t}^{1} \mu(\tau) \, d\tau \quad (25)$$

where $\mu(\tau)$ is the RC2 frequency impulse$^6$ (i.e. $L = 2$) and $q_{n,i} \in \{\pm 1\}$, $M_q = 2$ and $\kappa = 1/2$. We will consider the case of 2-space code (i.e. two transmitter antennas $N_t = 2$). These particular values imply also that $\Delta q_{n,i} \in [0, \pm 2]$ and $\psi (m) \in [0, 1]$, $M_\Theta = 2$. The distance evaluation state vector is

$$\Theta_i(m) = [\psi_i(m), \Delta q_{m-1,i}/2]^T \quad (26)$$

with initial state $\Theta_i(0) = [0, 0]^T$. The total number of distance evaluation states is $M_\Theta = 6^2 = 36$. In this example case, the first equation in (23) becomes particularly simple

$$[\Theta_i(m + 1)]_1 = ([\Theta_i(m)]_1 + [\Theta_i(m)]_2) \mod 2. \quad (27)$$

There is one special feature of the case with $M_\Theta = 2$ which is however not generally applicable for other values. This is the symmetry of the previous equation with respect to the sign of the $[\Theta_i(m)]_2$. We can therefore define a new state with reduced number of possible values

$$\Theta_i'(m) = [\psi_i(m), \Delta q_{m-1,i}/2]^T, \quad (28)$$

$^5$We can assign the index $m = 0$ to this point without any loss of generality.

$^6$Raised Cosine impulse with the length $2T_S$.

$\theta_i'(m) \in \{[0, 0]^T, [0, 1]^T, [1, 0]^T, [1, 1]^T\}$. \quad (29)

The number of all possible states is then reduced to $M_{\theta i}' = 4^2 = 16$.

For this particular example, we can evaluate mutual relation between $P_{2e}^{(ub)}$ and $P_{2e}^{(ub)}$ values by means of simulation. We plot a 2-dimensional graph $P_{2e}^{(ub)}$ versus $P_{2e}^{(ub)}$ where each point corresponds to one particular pairwise error probability event corresponding to particular pair of channel phase symbols sequences $\hat{q}_{i}^{(1)}$ and $\hat{q}_{i}^{(2)}$. We plot the point for all possible combinations of $\hat{q}_{i}^{(1)}$ and $\hat{q}_{i}^{(2)}$. The results for various parameters are on Figure 1. Signal to noise ratio is defined as

$$\gamma_s = \frac{\mathbf{E} \left[ \sum_{k=1}^{N_t} |u_k|^2 \right]}{N_0} = \frac{N_t N_r \sigma_e^2}{N_0}. \quad (30)$$

The most important is the behavior of relation in the north-east part of the graph. This is the region which dominates the overall error properties of the detection. We can see that the relation of $P_{2e}^{(ub)}$ and $P_{2e}^{(ub)}$ is close to a monotone function in that region. This justifies usage of $P_{2e}^{(ub)}$ as a code design objective function.

IV. MODULATOR TRELLOIS CODE DESIGN

A. Design objectives

In this section, we attempt to synthesize a modulator coder—i.e. to design a mapping$^7$ of data symbols $d_n$ and modulator states $\sigma_n$ on the channel phase symbols $q_n$ and modulator state transitions $\sigma_{n+1} = \sigma (d_n, \sigma_n)$ in such a way that the minimum mean squared distance over all different data messages

$$P_{\text{free}}^2 = \min_{d^{(1)}, d^{(2)}} \rho^2 \quad (31)$$

is maximized. The squared distance $\rho^2$ evaluation is however (as derived in the previous section) also ruled by the state-space model (state evaluation trellis). This makes the synthesis particularly difficult since we need to design a system with memory (discrete part of the modulator) in such a way that its output difference $\Delta q_n$ for two data messages fed into another system with memory (distance evaluation trellis) maximizes the output of the second one ($\rho^2$) for all $d^{(1)} \neq d^{(2)}$.

B. Initial and final state of distance evaluation trellis

We will assume that the first data message is identical to the first modulator trellis path split. We also realize that

$^7$The discrete function $(d_n, \sigma_n) \mapsto q_n$ forms a discrete part of the modulator.
the initial state of the distance evaluation trellis is always (for any particular data) given by (see section III-B)
\[ \mathbf{\Theta}_i(0) = [0, 0, \ldots 0]^T \] (32)
at the point of the path split.

A situation is much more interesting at the point in the modulator trellis where the two paths merge again. When
the paths merge again at \( n = K \), where \( K \) is the length of
the free path, then \( \Delta \mathbf{q}_s = \mathbf{0} \) for all \( n \geq K \). However
the behavior of the distance evaluation trellis is very different.
The distance evaluation trellis reaches its final state (after
some delay given by the memory of the evaluation
procedure) after the merge in the modulator trellis at \( n = K \).
A particular value of the final state depends on the distance
evaluation trellis and on the particular history of channel
phase symbol differences \( \Delta \mathbf{q}_f \). The number of possible
final states is \( M_0^N \), which can be easily deduced from the
distance evaluation state transition equation. All possible
final states are
\[ \mathbf{\Theta}_i(m) = [\psi_i(m), 0, 0, \ldots 0]^T, \] (33)
i = 1, \ldots, N_f.

One of the possible final states is equivalent to the initial
state \( \mathbf{\Theta}(m) = \mathbf{0} \). We realize that the squared distance
increment is zero \( \Delta \rho^2(m) = 0 \) if and only if the distance
evaluation state is \( \mathbf{\Theta}(m) = \mathbf{0} \) and \( \Delta \mathbf{q}_m = \mathbf{0} \). This means that
there are some paths (with the final state \( \mathbf{\Theta}(m) = \mathbf{0} \))
in the distance evaluation trellis that stop to increase the
accumulated squared distance when they reach the final
state. There are also paths (with the final state \( \mathbf{\Theta}(m) \neq \mathbf{0} \))
that continue to increase the squared distance even when
they reach the final state. The paths with the final state
\( \mathbf{\Theta}(m) = \mathbf{0} \) will be referred as critical distance evaluation
trellis paths. These paths affect the minimum free
distance of the modulated signal. All other paths continue
to increase the distance even after the merge in the mod-
ulator trellis and therefore cannot influence the minimum
free distance.

It is also useful to observe that in order to get into the
final state \( \mathbf{\Theta}(m) \) the last \( L - 1 \) channel symbol differences
must be zero \( \Delta \mathbf{q}_s = \mathbf{0} \) for \( n = (m - L + 1), \ldots, (m - 1) \).
Thus only first \( m - L + 1 \) differences \( \Delta \mathbf{q}_s \) needs to
be explored for reaching the final state at \( m \).

\section*{C. Modulator trellis design rules}

Here, we formulate modulator trellis design rules based
on the previous conclusions on the behavior of the distance
evaluation trellis (especially its final states). The
main concern is the critical distance evaluation trellis path.
The design rules are:

1. We analyze the distance evaluation trellis and find the
shortest free critical paths. The length of the shortest one
is denoted by \( \xi_0 \). The longer ones are denoted by \( \xi_n =
\xi_0 + n, n \in \mathbb{N} \).

2. We identify the set of all sequences of the channel
phase symbols differences
\[ \Lambda_\ell = \{ \Delta \mathbf{\hat{q}} = [\Delta \mathbf{\hat{q}}_0, \Delta \mathbf{\hat{q}}_1, \ldots, \Delta \mathbf{\hat{q}}_{\ell - L}(p)]_{p=1}^{M_\ell} \} \] (34)
that induce the critical path in the distance evaluation
trellis. Especially important is the set corresponding to the
shortest free path \( \Lambda_0 \). We also find corresponding
squared distances for all critical paths \( \rho_{\ell, c}^2(\Delta \mathbf{\hat{q}}) \) \( \rho^2 \)
for all \( \Delta \mathbf{\hat{q}} \in \Lambda_\ell \) and define
\[ \rho_{\ell, \text{min}, c}^2 = \min_{\Delta \mathbf{\hat{q}} \in \Lambda_\ell} \rho_{\ell, c}^2(\Delta \mathbf{\hat{q}}). \] (35)

3. We design the modulator trellis in such a way that we
avoid any sequence \( \Delta \mathbf{\hat{q}} \in \Lambda_0 \) for any pair of possible
paths in the modulator trellis that splits at the common
point and merges after \( \xi_0 - L + 1 \) transitions.

4. The pairs of free paths having the length \( \xi_1 - L + 1 \)
in the modulator trellis will be assigned to such channel
symbols that \( \Delta \mathbf{\hat{q}} \in \Lambda_1 \) and correspond to the
highest possible \( \rho_{\ell, 1}^2(\Delta \mathbf{\hat{q}}) \). This guarantees maximization of
Fig. 2. The distance evaluation trellis with reduced set of the states for the example III-C.

\[
\rho_{c_{\text{min},1}}^2(\mathcal{U}_1) \quad \text{where the minimum free squared distance over all used } \Delta \mathbf{q} \in \mathcal{U}_\ell \text{ is }
\rho_{c_{\text{min},1}}^2(\mathcal{U}_\ell) = \min_{\Delta \mathbf{q} \in \mathcal{U}_\ell} \rho_{c_{\text{min},1}}^2(\Delta \mathbf{q})
\]

(36)

and the set of all used difference sequences is defined as

\[ \mathcal{U}_\ell = \{ \Delta \mathbf{q}, \text{ used at critical path with } \xi_\ell \text{ length} \} \subset \mathcal{A}_\ell. \]

(37)

The expression \( \rho_{c_{\text{min},1}}^2(\mathcal{U}_1) \) can be greater then \( \rho_{c_{\text{min},1}}^2 \) since there can be some \( \Delta \mathbf{q} \in \mathcal{A}_1 \) not actually used in the modulator trellis. On the other side, the particular situation for given modulation parameters generally does not guarantee that \( \rho_{c_{\text{min},1}}^2(\mathcal{U}_1) \) would be greater than \( \rho_{c_{\text{min},1}}^2(\mathcal{U}_\ell) \) for \( \ell_2 > \ell_1 \). A careful observation of this behavior is needed. However the situation for high difference values \( \ell_2 - \ell_1 \) usually satisfies the condition.

D. Example case

Here, we will continue with the example case III-C. The distance evaluation trellis free paths are shown on Fig. 2 and 3. The reduced states are indexed by a new variable \( \Omega^g(\mathbf{m}) \) which is get from the reduced states \( \mathbf{y}_i(\mathbf{m}) \) as

\[
\Omega^g(\mathbf{m}) = [8, 4] \mathbf{y}_1(\mathbf{m}) + [2, 1] \mathbf{y}_2(\mathbf{m}).
\]

(38)

We use similar indexation by \( \Omega^g(\mathbf{m}) \) for ordinary states \( \mathbf{y}_i(\mathbf{m}) \). It can be clearly seen that the shortest free critical path has the length \( \xi_0 = 3 \).

The optimum code search in this simple case could be done by the following way.

1. We determine the sequences \( \mathcal{A}_0 \). An analysis of the distance evaluation trellis reveals that \( M_{\xi_0} = 24 \) and the set \( \mathcal{A}_0 \) is

\[
\mathcal{A}_0 = \bigcup_i \mathcal{A}_{0,i}
\]

where

\[
\begin{align*}
\mathcal{A}_{0,1} &= \{[\pm2, \pm2]^T, [\pm2, \pm2]^T\}, \\
\mathcal{A}_{0,2} &= \{[\pm2, 0]^T, [\pm2, 0]^T\}, \\
&\quad \{[0, \pm2]^T, [0, \pm2]^T\}. \\
\end{align*}
\]

(39)

(40)

Mean squared distances for these sets are

\[
\begin{align*}
\rho^2(\mathcal{A}_{0,1}) &\geq 4.0, \\
\rho^2(\mathcal{A}_{0,2}) &\geq 2.0.
\end{align*}
\]

(41)

(42)

2. Then we find sequences \( \mathcal{U}_1 = \bigcup_i \mathcal{U}_{1,i} \) with the largest accumulated squared distances. In this example, the subsets with largest distances are

\[
\begin{align*}
\mathcal{U}_{1,1} &= \{[\pm2, \pm2]^T, [0, 0]^T, [\pm2, \pm2]^T\}, \\
\mathcal{U}_{1,2} &= \{[\pm2, 0]^T, [\pm2, 0]^T, [\pm2, \pm2]^T\}, \\
&\quad \{[0, \pm2]^T, [\pm2, 0]^T, [0, \pm2]^T\}, \\
&\quad \{[\pm2, 0]^T, [\pm2, \pm2]^T, [0, \pm2]^T\}, \\
&\quad \{[0, \pm2]^T, [\pm2, \pm2]^T, [\pm2, 0]^T\}, \\
&\quad \{[\pm2, 0]^T, [\pm2, \pm2]^T, [\pm2, 0]^T\}, \\
&\quad \{[0, \pm2]^T, [\pm2, \pm2]^T, [\pm2, 0]^T\}. \\
\end{align*}
\]

(43)

(44)

\[
\mathcal{U}_{1,i} = \cdots, \; \forall i > 2.
\]

(45)

Mean squared distances for these sets are

\[
\begin{align*}
\rho^2(\mathcal{U}_{1,1}) &\geq 8.0, \\
\rho^2(\mathcal{U}_{1,2}) &\geq 6.0, \\
\end{align*}
\]

(46)

(47)

\[
\rho^2(\mathcal{U}_{1,i}) < \rho^2(\mathcal{U}_{1,2}), \; \forall i > 2.
\]

(48)

3. Then the trellis is set up in such a way that for \( m = 0 \) the diverting paths, for \( m = 1 \) the disjunct paths and
for $m = 2$ the merging paths are assigned such channel symbols that their difference falls into subset $U_{i,j}$ with largest possible distance.

4. Next level of the optimization procedure is the choice of the used paths set avoiding the critical paths even on the lengths $\xi_\ell$, $\ell \geq 1$. However the methodology for this type of design needs more research effort and currently relies on the ad-hoc approach.

Fig. 4 shows a simple 2-space 2-state ($M_d = 2$) modulator trellis for binary data ($M_d = 2$). In the case of the 2-state code we are however very restricted in possible mappings of channel symbols onto the code trellis. Therefore we, in fact, cannot use the best $U_{1,1}$. In order to have the disjunct paths of the $\Delta q_0 = [0, 1]^T$ type, both diverting and merging paths would be from the same set and therefore from the set $A_0$. In order to utilize $U_{1,1}$ the coder trellis would have to have more than 2 states. We must therefore use the set $U_{1,2}$. Diverting paths will be from $\Delta q_0 = [\pm 2, 0]^T$, disjunct paths from $\Delta q_0 = [0, \pm 2]^T$ and merging paths from $\Delta q_0 = [\pm 2, \pm 2]^T$.

There are shown several different codes among which the code A is the one with the best possible (for 2-state code) set $U_{i,j}$. For comparison purposes, Fig. 4 shows also other non-optimal codes. The code B has the shortest used critical path length equal to the minimal one—i.e. $\xi_0 = 3$. Code C has the shortest critical path length $\xi_1 = 4$ however the set $U_1$ is not the optimal one. There is also shown $\rho_{c_{\text{min}, \ell}}^2(U_{\ell j})$ where $\ell$ corresponds to the shortest used critical path having length $\xi_\ell$. The last shown code D is a “hand crafted” design trying to avoid the critical paths even on the length $\xi_1$. We are rewarded by the largest distance.

E. Open problems

There is a number of open problems that will need to be addressed in the future work. The most important ones are:

- general relation of $\rho_{c_{\text{min}}, \ell_2}^2$ and $\rho_{c_{\text{min}}, \ell_1}^2(U_{\ell_1})$ for $\ell_2 > \ell_1$,
- general relation between the number of the modulator states $M_d$ and the availability of sufficiently large set $U_{\ell}$ with high $\rho_{c_{\text{min}}, \ell}^2(U_{\ell})$,
- design avoiding critical paths on $\xi_\ell$, $\ell \geq 1$.

V. Conclusions

In this paper, we have presented a general procedure for the space-time constant envelope trellis coded modulation design. The work was concentrated around several key points. The first one was the derivation of the averaged squared distance in the slowly flat fading Rayleigh channel which was shown to correspond to the mean of the error probability logarithm. In the case of the nonlinear modulation this is the only mathematically tractable way. Second major point shown in the paper is the fact that the evaluation of the distance is controlled by two procedures with memory—the discrete part of the modulator and the distance evaluation algorithm. Both can be described by a trellis. The distance evaluation trellis has some special properties related to its initial and final state and also to the distance increments. All these properties were closely investigated. The most important is the concept of the critical distance evaluation trellis path which dominantly determines the overall modulation free distance. The principle of critical distance evaluation path can be generally utilized in any other situation where we design a system with memory while optimizing the objective function itself possessing the memory. In the second part of the paper a set of rules for the modulator trellis code design was suggested. The main attention was focused to the critical path treatment. As an example, a simple 2-space 2-state binary trellis code was designed.

References