MULTI-LEVEL TRELLIS CODED MODULATION AND MULTI-STAGE DECODING*

Daniel J. Costello, Jr.
Jiantian Wu
Department of Electrical and Computer Engineering
University of Notre Dame
Notre Dame, IN 46556

Shu Lin
Department of Electrical Engineering
University of Hawaii
Honolulu, HI 96822, USA

presented at
1990 IEEE Information Theory Workshop
Eindhoven, The Netherlands
June 1990

*This work was supported by NSF Grant NCR89-03429 and NASA Grant NAG5-557.
Outline of Paper

• Introduction

• Basic multi-level trellis coding

• Multi-level trellis coding based on multi-way partition chains

• Combined Ungerboeck-type and multi-level trellis coding

• Generalized multi-level trellis coding

• A level-spanning, multi-level trellis coding scheme

• 90° rotationally invariant multi-level trellis codes

• Conclusions
**Introduction**

- The goal in designing a coded modulation system is to achieve a good trade-off between coding gain, decoding complexity, and decoding delay.

- Multi-level coding is a powerful technique for constructing bandwidth efficient coded modulation codes. Good multi-level coding schemes can be designed by using previously known codes as component codes.

- Multi-stage decoding provides a simple decoder implementation for multi-level codes with a small loss in coding gain.

- For coded QAM, the total power gain over uncoded QAM is composed of two parts: the coding gain \( \gamma(C) \) and the shaping gain \( \gamma(S) \).
• At a bit error rate of \(10^{-5} \sim 10^{-6}\), the maximum coding gain is about 7.5 dB, and the maximum shaping gain is about 1.5 dB.

• Because these gains can be achieved independently, for coded QAM we focus on coding gain only and choose the signal set to be \(Z^N\) (\(N\) dimensional integer lattice).

• For coded MPSK, we also focus only on coding gain, since no shaping gain is possible.

• The phase invariant property (or phase symmetry) is useful in resolving carrier-phase ambiguity and ensuring rapid carrier-phase resynchronization after a temporary loss of synchronization. It is desirable for a coded modulation system to have as much phase symmetry as possible.

• We present necessary and sufficient conditions for a QAM code to be \(90^\circ\) rotationally invariant, and some \(90^\circ\) rotationally invariant multi-level codes are constructed.
Multi-level Trellis Coding Based on Set Partitioning

- Figure 1 shows a multi-level trellis coding scheme based on set partitioning. $\Lambda_0$ is a signal set, $\Lambda_i$ is a subset of $\Lambda_{i-1}$, and $\Lambda_m$ is the all-zero vector. $C_1, C_2, \ldots, C_m$ represent the different component codes, and the overall multi-level code is denoted by $C$.

Fig. 1 Multi-level trellis coding based on set partitioning
Related previous work

- Leech (1964) and Leech and Sloane (1971) used a multi-level structure to construct lattices.

- Multi-level codes using "proper indexing", which is the same as Ungerboeck's "set partitioning", of two dimensional signal sets was proposed by Imai and Hirakawa (1977). They also presented a multi-stage decoding method using a posteriori probabilities based on channel statistics.


- Sayegh (1986) showed how Imai and Hirakawa's method can be combined with set partitioning to create multi-level block coded modulation systems.

- Pottie and Taylor (1989) proposed a hierarchy of codes to match the partitioning of signal sets by generalizing Imai and Hirakawa's and Ginzburg's coding schemes.
• Calderbank (1989) investigated the path multiplicity for a variety of multi-level codes.

• Tanner (1990?) studied linking subspaces of vector spaces to guarantee a large minimum separation between signals in the resulting signal set so that good multi-level codes can be designed.

Basic multi-level trellis codes

• This construction is based on two-way partition chains, where all component codes are binary codes (block or convolutional).

• Let \( \Delta_i \) be the minimum squared Euclidean distance (MSED) of \( \Lambda_i \) for \( i = 0, 1, \ldots, m \).

• Let \( d_i \) be the minimum Hamming distance of binary code \( C_i \) for \( i = 1, 2, \ldots, m \).

• Then the MSED of the multi-level code is (Leech & Sloane, Ginzburg, Sayegh, etc.)

\[
D(C) = \min\{d_i\Delta_{i-1}, 1 \leq i \leq m\}
\]
• The normalized redundancy $\rho(C)$ is defined as (Forney) the number of redundant bits per two dimensional signal (symbol).

• The spectral efficiency $\eta(C)$ is defined as (Ungerboeck) the number of information bits per two dimensional signal (symbol).

• Basic multi-level codes with normalized redundancy $\rho(C) = 1 \text{ bit/symbol}$ were presented by Yamaguchi and Imai (1987).

• Basic multi-level codes with smaller normalized redundancies can be constructed by using two-way partition chains with multi-dimensional signal sets and binary convolutional or block codes. Some four and eight dimensional basic multi-level codes were constructed by Wu and Zhu (1990?).

• We present some new basic multi-level codes based on set partitioning of one and two dimensional signal sets. Some of these new codes have non-integer normalized redundancies $\rho(C)$. 
• **Example 1.** A three-level trellis code using an 8-PSK signal set with mapping by set partitioning is shown in Figure 2.
• Let $C_1$ be a 16-state rate-$1/4$ convolutional code with minimum free Hamming distance 16,
$C_2$ an 8-state rate-$3/4$ convolutional code with free distance 4,

$C_3 = P_n$, the $(n, n-1)$ single parity check code.
• The spectral efficiency of this multi-level code is

$$\eta(C) = 1 + \frac{(n-1)}{n} \text{ bits/symbol}$$

• The minimum free squared Euclidean distance is

$$D(C) = \min\{0.586 \times 16, 2 \times 4, 4 \times 2\} = 8$$

• The nominal coding gain (Ungerboeck) over uncoded QPSK is

$$\gamma(C) = 10 \log_{10} \left( \frac{D(C)}{D(QPSK)} \right) = 6.02 dB.$$  

• The 256-state, rate-$2/3$, $\eta(C) = 2$ bits/symbol Ungerboeck code has $D(C) = 7.515$ and $\gamma(C) = 5.75$ dB.
Multi-Stage Decoding of Example 1

- A three-level multi-stage decoder for Example 1 is shown in Figure 3.

![Diagram of multi-stage decoding](image-url)

**Fig. 3** Multi-stage decoding for Example 1
• The normalized complexity $N_D$ of multi-stage decoding is the number of required binary operations (additions and comparisons) per 2 dimensional symbol.

• For a $2^\nu$-state, $k$ input bit, $n$ output bit convolutional (trellis) code, the Add-Compare-Select (ACS) operation of the Viterbi algorithm requires $2^k$ additions and a comparison of $2^k$ numbers, or $2^k - 1$ binary comparisons, for each of the $2^\nu$ states, so its complexity is $2^{k+\nu+1} - 2^{\nu}$. (This number should be normalized to the complexity per 2 dimensional symbol.)
First-stage of Decoding

- For each state transition period, the symbol metrics of both QPSK subsets (see Figure 4) must be computed. Complexity = 2 binary operations/symbol
• Then the branch metrics within each state transition period must be computed by adding the four symbol metrics on each branch.
  Complexity = 6 binary operations/symbol

• The ACS operation of the Viterbi algorithm is then used to determine the surviving path at each state.
  Complexity = 12 binary operations/symbol
Second-stage of Decoding

- The decoded information from the first stage is passed on to the second-stage.

- For each state transition period, the symbol metrics of both BPSK subsets (see Figure 5) must be computed.

Complexity = 2 binary operations/symbol

Fig. 5
• Then the branch metrics within each state transition period must be computed by adding the four symbol metrics on each branch.
Complexity = 6 binary operations/symbol

• The ACS operation of the Viterbi algorithm is then used to determine the surviving path at each state.
Complexity = 30 binary operations/symbol

• If the parallel transitions in the trellis are resolved by table-look-up, the complexity reduces to 14 binary operations/symbol.
Third-stage of Decoding (assume $n = 32$)

- The decoded information from the first and second stages is made available to the third stage.

- For each state transition period, the metrics of both the 0 and 1 symbols must be computed.
  Complexity = 2 binary operations/symbol

- In this case, the branch metrics are the symbol metrics computed above (one symbol per trellis branch).

- After 8 branches (32 symbols) in the first and second trellis are decoded, the Viterbi algorithm is used to make a decoding decision for the block code $C_3 = P_{32}$.
  Complexity = 6 binary operations/symbol

- The total decoding complexity is $N_D = 66$ binary operations per 2 dimensional symbol, or $N_D = 50$ not counting the parallel transitions.
• If we take $C_3$ to be $P_n$, where $n \to \infty$ (i.e., a 2-state, non-redundant, catastrophic trellis code), the multi-level code has $1 + 3 + 4 = 8$ input bits and $4 + 4 + 4 = 12$ output bits for every four 8-PSK transmitted symbols. Overall, this can be viewed as a $16 \times 8 \times 2 = 256$-state 8-dimensional trellis code.

• Without considering the computation of the symbol and branch metrics, the ACS complexity of maximum likelihood decoding of the overall trellis code is 

$$\frac{2^{8+8+1} - 2^8}{4} = 2^{15} - 2^6 > 3 \times 10^4$$

binary operations/2 dimensional symbol.

• Note that the complexity of the multi-stage decoder in this example is only about 0.2% of the complexity of the overall maximum likelihood decoder.

• However, the performance of the multi-stage decoder is close to that of the maximum likelihood decoder.

• For the 256-state Ungerboeck code, the ACS complexity alone is 1792 binary operations/symbol.
Example 2. The one dimensional partition chain $\mathbb{Z}/2\mathbb{Z}/4\mathbb{Z}/\ldots$ has MSED $1/4/16/\ldots$ (see Figure 6).

Fig. 6 Set partitioning of $\mathbb{Z}$
Let $C_1$ be a 16-state rate-$1/4$ convolutional code with minimum free Hamming distance 16,

$C_2$ be an 8-state rate-$3/4$ convolutional code with free distance 4,

and $C_3, C_4, \ldots$ be rate-1 codes (no coding).

Since there are two levels of coding, this is a two-level code (see Figure 7).

The normalized redundancy $\rho(C)$ is 2 bits per symbol.

The MSED is

$$D(C) = \min \{1 \times 16, \ 4 \times 4, \ 16\} = 16$$
• The nominal coding gain (Forney) is

\[ \gamma(C) = 10 \log_{10} \frac{D(C)}{2^{\rho(C)}} = 6.02 \text{ (dB)} \]

• This code has the same nominal coding gain and normalized redundancy as the 24-dimensional Leech lattice \( \Lambda_{24} \) but much less decoding complexity.

• Due to a large path multiplicity, the effective coding gain of this two-level code is less than the nominal coding gain. To reduce the path multiplicity, we can choose longer convolutional codes (with larger constraint lengths and free distances).

• For example, if \( C_1 \) is a 32-state rate-1/4 convolutional code with free distance 18 and \( C_2 \) is a 32-state rate-3/4 convolutional code with free distance 5, the path multiplicity is reduced and the effective coding gain is closer to 6.02 dB.
Using multi-stage decoding, an additional loss of coding gain occurs, but the decoding complexity is less than a 64-state Gerberboeck code and much less than the Leech lattice \( \Lambda_{24} \).

Table 1. Comparison of multi-level trellis codes and other codes (spectral efficiency \( \eta(C) = 4 \) \text{ c/symbol} \) using 8-PAM modulation

<table>
<thead>
<tr>
<th>Codes</th>
<th>#S</th>
<th>( R_i )</th>
<th>( \gamma(C) )</th>
<th>( N_D )</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-level</td>
<td>16 &amp; 8</td>
<td>1/4 &amp; 3/4</td>
<td>5.81</td>
<td>116</td>
<td>14</td>
</tr>
<tr>
<td>5-level</td>
<td>32 &amp; 8</td>
<td>1/4 &amp; 3/4</td>
<td>5.81</td>
<td>130</td>
<td>16</td>
</tr>
<tr>
<td>7-level</td>
<td>32 &amp; 32</td>
<td>1/4 &amp; 3/4</td>
<td>5.81</td>
<td>350</td>
<td>20</td>
</tr>
<tr>
<td>8-level</td>
<td>256</td>
<td></td>
<td>5.81</td>
<td>( \sim 1264 )</td>
<td></td>
</tr>
<tr>
<td>Gerberboeck</td>
<td>32</td>
<td>2/3</td>
<td>4.77</td>
<td>232</td>
<td>5</td>
</tr>
<tr>
<td>Gerberboeck</td>
<td>64</td>
<td>2/3</td>
<td>5.44</td>
<td>456</td>
<td>6</td>
</tr>
</tbody>
</table>

The decoding delay of a multi-level trellis code is proportional to

\[
D \triangleq \frac{1}{2} \sum N_i K_i,
\]

where \( N_i \) is the dimensionality of the signals (cosets) associated with a branch transition of the \( i \)th component code and \( K_i \) is the constraint length of the \( i \)th component code.
Multi-level trellis codes based on a set partition chain with strictly increasing distances

- Multi-level trellis codes using multi-dimensional signal sets can achieve higher spectral efficiencies (lower normalized redundancies) than multi-level codes based on two dimensional signal sets.

- For two-way partitioning of multi-dimensional signal sets, the MSED at successive partition levels may be equal. For example, the partition chain $Z^4/D_4/RZ^4/RD_4/2Z^4/2D_4/...$ of the four dimensional integer lattice $Z^4$ has distances $1/2/2/4/4/8/...$, where $R$ represents the rotation operation, $R^2 = 2$, and $D_4$ is the densest known four dimensional lattice.

- Reducing the number of component codes can reduce the decoding delay and the path multiplicity.
• Some partition levels can be joined to form a new multi-way partition chain with strictly increasing distances. For example, the partition chain $Z^4/D_4/RD_4/2D_4/\ldots$ has distances $1/2/4/8/\ldots$. Since $|Z^4/D_4| = 2$ and $|R^i D_4/R^{i+1} D_4| = 4$ for $i = 0, 1, 2, \ldots$, the first component code can be a binary code, and other component codes can be binary input, 4-ary output codes or codes over $GF(4)$.

• The lower bound on the MSED of these multi-level codes is given by

$$D(C') \geq \min\{d_i \Delta_{i-1}, \ 1 \leq i \leq m\}$$

where $d_i$ is now the minimum free Hamming distance of code $C_i$ (binary or 4-ary).
Example 3. This code is based on the partition chain $Z^4/D_4/RD_4/\ldots$ and includes two component codes (see Figure 8):

$C_1$ is an 8 state rate-3/4 convolutional code with free distance 4 and $C_2$ is an $(N, N-1)$ block code over $GF(4)$ with minimum distance 2 (4 states).

• The normalized redundancy is $\rho(C) = \frac{1}{8} + \frac{1}{N}$, the MSED is 4, and the nominal coding gain is

$$\gamma(C) = 5.64 - \frac{3.01}{N} \text{ (dB)} = 5.48 \text{ dB } (N = 19)$$

• The decoding complexity is $N_D = 37$, and the decoding delay is $D = 24$ (excluding the decoding delay of the block code).

![Diagram](Fig.8 Multi-level code of Example 3.)

• The 64-state, rate-4/5 Ungerboeck code for $Z^4$ has $\gamma(C) = 5.48 \text{ dB}$, $N_D \approx 496$, and $D = 12$. 
Combined Ungerboeck-type and multi-level trellis codes

- For the above two classes of multi-level codes, each output symbol of a component encoder corresponds to a single coset of a subset of a signal constellation (two dimensional or multi-dimensional). For Ungerboeck-type codes, all the encoder output symbols associated with a single trellis branch correspond to a single coset of a subset of a signal constellation. Ungerboeck-type codes can be used as component codes at some levels in conjunction with a multi-way partition chain.

- Instead of using several high rate codes at higher levels of partitioning, we use an Ungerboeck-type code to reduce the decoding delay and path multiplicity.

- The usual lower bound on the MSED cannot be applied to this construction. A more general lower bound on the MSED of these multi-level codes (Kasami & Lin) is given by

$$D(C) \geq \min\{D(C_i), \ 1 \leq i \leq m\}$$

where $D(C_i)$ is the MSED of code $C_i$. 
Example 4. The encoding structure is shown in Figure 9.

- Let $C_1$ be a 16-state rate-1/4 convolutional code with minimum free Hamming distance 16,

$C_2$ a 16-state rate-7/8 trellis code with MSED 8 (Pietrobon, Deng, et.al., 1990).

![Diagram of multi-level code of example 4.](image)

- The 64-state, rate-4/5, $\eta(C) = 2$, $4 \times 8$-PSK code constructed by Pietrobon, Deng, et.al. (1990) has $D(C) = 7.029$, $\gamma(C) = 5.46$ dB, $N_D \approx 496$, and $D = 12$. 

27
• Encoding procedure:
  The information sequence is divided into blocks of 8 bits each:
  the first bit in each block enters encoder $C_1$, and the 4 output bits specify 4 consecutive cosets of $4 \times (8\text{-PSK/QPSK})$,
  the other 7 bits of the block enter encoder $C_2$ and the 8 output bits specify a $4 \times$ QPSK signal.

• The spectral efficiency of this multi-level code is
  \[ \eta(C) = 8/4 = 2 \text{ bits/symbol} \]

• The MSED is
  \[ D(C) = \min\{16 \times 0.586, 8\} = 8 \]
  where $D(C_1) = 16 \times 0.586$.

• The nominal coding gain over uncoded QPSK is $\gamma(C) = 6.02$ dB.

• The decoding complexity for multi-stage decoding is $N_D \approx 100$ binary operations per symbol.

• The decoding delay is only $D = 16$, which is less than a multi-level code with more stages.
Generalized multi-level trellis codes

- The previous examples were all based on Ungerboeck's set partitioning. A modified set partitioning method can be used to construct generalized multi-level trellis codes.

- **Example 5.** The four dimensional 8-state code $C(Z^4/RD_4)$ constructed by Wei (1987) has MSED 4. Mapping the same binary code to $RZ^4/2D_4$ rather than $Z^4/RD_4$, we obtain a trellis code, denoted by $C_2(RZ^4/2D_4)$, with MSED 8. Using an 8-state rate-1/3 convolutional code as the first component code $C_1(Z^2/RZ^2)$ and $C_2(RZ^4/2D_4)$ as the second component code gives the two-level trellis code shown in Figure 10.

![Diagram](image)

Fig. 10 Multi-level code of Example 5.
• Encoding procedure:
The information sequence is divided into blocks of a specified number of bits according to the desired spectral efficiency: the first 2 bits in each block enter encoder $C_1$, and the 6 output bits specify 6 consecutive cosets of $Z^2/RZ^2$, i.e., 3 consecutive cosets of $Z^4/RZ^4$, the next 6 bits of the block enter encoder $C_2$, and the 9 output bits specify 3 consecutive cosets of $RZ^4/2D_4$.
Together with uncoded bits, each coded block determines 3 consecutive four dimensional signals, i.e., 6 two dimensional signals.
• Note that the first coding level partitions a six dimensional signal set whereas the second coding level partitions a four dimensional signal set.
• The nominal coding gain is $\gamma(C) = 5.52$ dB, which is 1.00 dB greater than Wei's code.
• The decoding delay of this two-level code is $D = 15$, whereas the delay of Wei's code is $D = 6$.
• The decoding complexity of this two-level code is $N_D = 56$, whereas the complexity of Wei's code is $N_D = 44$. 
Example 6. Consider the generalized multi-level trellis code shown in Figure 11.

The first component code, associated with the partition $\mathbb{Z}/2\mathbb{Z}$, is a 32 state rate-$1/2$ convolutional code with free distance 8.

The second component code, associated with the partition $2\mathbb{Z}^8/2D_8$, is a single parity check block code of length $N$. 

Fig. 11 Multi-level code of Example 6.
• Encoding procedure (spectral efficiency = m bits/symbol):

• The information sequence is divided into blocks of $4Nm$ bits each, with three subsequences of length $4N, N - 1$, and $4Nm - 4N - (N - 1)$ corresponding to $C_1, C_2$, and uncoded bits, respectively.

• Each output bit of encoder $C_1$ specifies a coset of partition $Z/2Z$, i.e., $8N$ output bits specify $N$ cosets of $Z^8/2Z^8$.

• Each output bit of encoder $C_2$ specifies a coset of $2Z^8/2D_8$, i.e., $N$ output bits specify $N$ cosets of $2Z^8/2D_8$.

• Together with the $4Nm - 4N - (N - 1)$ uncoded bits, each coded block of $N$ eight dimensional signals, i.e., $4N$ two dimensional signals contains $4Nm$ bits of information and the spectral efficiency in $m$ bits/symbol.
• The normalized redundancy is \( \rho(C) = 1 + \frac{1}{4N} \) and the MSED is 8. Therefore the nominal coding gain is

\[
\gamma(C) = 6.02 - \frac{3.01}{4N} \text{ (dB)}.
\]

• The decoding complexity is \( N_D = 140 \) and the decoding delay (excluding the block code) is \( D = 5 \).

• The 128-state, rate-4/5 Ungerboeck code for \( \mathbb{Z}^8 \) has \( \gamma(C) = 5.27 \text{ dB} \), number of nearest neighbors \( N_{\text{free}} = 112 \), \( N_D \approx 992 \), and \( D = 28 \).
In general, let $\Lambda_0$ be a signal set and $\Lambda_0^{(1)}$ be a set such that $\Lambda_0 \times \cdots \times \Lambda_0 = \Lambda_0^{(1)} \times \cdots \times \Lambda_0^{(1)}$, for some integers $j_0, k_0 \geq 1$.

If there are $2m$ sets $\Lambda_{i-1}^{(i)}$ and $\Lambda_i^{(i)}$, for $i = 1, 2, \ldots, m$, where $\Lambda_m^{(m)}$ is the empty set, satisfying the following conditions:

1. $\Lambda_{i-1}^{(i-1)} \times \cdots \times \Lambda_{i-1}^{(i-1)} = \Lambda_{i-1}^{(i)} \times \cdots \times \Lambda_{i-1}^{(i)}$, for $i = 2, 3, \ldots, m$, and for some $j_i, k_i \geq 1$;

2. $\Lambda_{i-1}^{(i)} \supseteq \Lambda_i^{(i)}$, for $i = 1, 2, \ldots, m$;

then we can construct a multi-level code having the form shown in Figure 12:

$$C = C_1 \left( \Lambda_0^{(1)}/\Lambda_1^{(1)} \right) + C_2 \left( \Lambda_1^{(2)}/\Lambda_2^{(2)} \right) + \ldots + C_m \left( \Lambda_{m-1}^{(m)}/\Lambda_m^{(m)} \right)$$
The MSED of this multi-level code is lower bounded by

\[ D(C) \geq \min\{D[C_i(\Lambda_{i-1}^{(i)}/\Lambda_i^{(i)})], 1 \leq i \leq m\} \]

where \( D[C_i(\Lambda_{i-1}^{(i)}/\Lambda_i^{(i)})] \) is MSED of the \( i \)th component code.
Level spanning multi-level trellis codes

- Level spanning provides an approach to constructing rotationally invariant multi-level trellis codes.
- However, the lower bound on MSED of generalized multi-level codes may not hold for this class of codes.

**Example 7.** Consider the multi-level coding scheme shown in Figure 13.

![Diagram of multi-level trellis code](image)
• $C_1$ is a 16-state rate-2/3 Ungerboeck code, which has MSED 6 when used with the partition $Z^2/RZ^2/2Z^2$.

• $C_2$ is an 8-state rate-7/8 binary convolutional code with free distance 3.

• $C_3$ is a 2-state $(8, 7)$ block code with minimum distance 2.

• Encoding procedure (spectral efficiency = $m$ bits/symbol):

• The information sequence is divided into blocks of $16m$ bits each, with four subsequences of length 16, 7, 7, and $16m - 30$ corresponding to $C_1, C_2, C_3$, and uncoded bits, respectively.

• Each state transition period of encoder $C_1$ outputs three bits which specify cosets of the partitions $Z^4/D_4, D_4/RZ^4$, and $RD_4/2Z^4$, respectively, i.e., 24 output bits specify 8 cosets of each partition.
Each output bit of encoder $C_2$ specifies a coset of $RZ^4/RD_4$, i.e., 8 output bits specify 8 cosets of $RZ^4/RD_4$.

Each output bit of encoder $C_3$ specifies a coset of $2Z^4/2D_4$, i.e., 8 output bits specify 8 cosets of $2Z^4/2D_4$.

Together with $16m - 30$ uncoded bits, each coded block of 8 four dimensional signals, i.e., 16 two dimensional signals, contains $16m$ bits of information and the spectral efficiency is $m$ bits/symbol.

Since the MSED’s of $Z^4, D_4,$ and $RD_4$ are the same as $Z^2, RZ^2,$ and $2Z^2,$ respectively, the MSED of $C_1$ is the same as the corresponding Ungerboeck code, i.e., $D(C_1) = 6$. Therefore, assuming the lower bound on MSED holds in this case, $D(C) = \min\{6, 3 \times 2, 2 \times 4, 8\} = 6$.

The normalized redundancy is $\rho(C) = 5/8$, the nominal coding gain is $\gamma(C) = 5.90$ dB, the decoding complexity of multi-stage decoding is $N_D = 108$, and the decoding delay is $D = 56$.

However, due to the uncertainty regarding the bound, the actual values of $D(C)$ and $\gamma(C)$ may be less than stated above.
It can be shown that the two bits corresponding to the partition levels $D_4/RZ^4$ and $RD_4/2Z^4$ are the only ones affected by a $90^\circ$ phase rotation. So this code can be combined with a differential encoder to achieve $90^\circ$ rotational invariance as shown in Figure 14.

Fig. 14 Diagram of four dimensional 90 rotationally invariant encoder with a differential encoder
Conclusions

- Several constructions for multi-level trellis codes are presented and many codes with better performance than previously known codes are found. These codes provide a flexible trade-off between coding gain, decoding complexity, and decoding delay.

- New multi-level trellis coded modulation schemes using generalized set partitioning methods are developed for QAM and PSK signal sets.

- New rotationally invariant multi-level trellis codes which can be combined with differential encoding to resolve phase ambiguity are presented.
Appendix B

New Multi-Level Codes over $GF(q)$