Chapter 5 Convolutional Codes and Trellis Coded Modulation

- 5.1 Encoder Structure and Trellis Representation
- 5.2 Systematic Convolutional Codes
- 5.3 Viterbi Decoding
- 5.4 Soft-Decision Viterbi Decoding
- 5.5 BCJR Decoding
- 5.6 Trellis Coded Modulation

§ 5.1 Encoder Structure and Trellis Representation

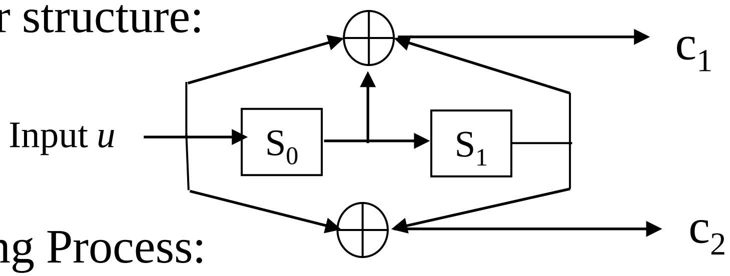
- Introduction
  - Encoder: contains memory (order m: m memory units);
  - Output: encoder output at time unit t depends on the input and the memory units status at time unit t;
  - By increasing the memory order m, one can increase the convolutional code's minimum distance ($d_{\min}$) and achieve low bit error rate performance (P_b);
  - Decoding Methods:
    - Viterbi algorithm [1]: Maximum Likelihood (ML) decoding algorithm;
    - Bahl, Cocke, Jelinek, and Raviv (BCJR) [2] algorithm: Maximum *A Posteriori* Probability (MAP) decoding algorithm, used for iterative decoding process, e.g. turbo decoding.

[1] A. J. Viterbi, "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," IEEE Trans. Inform. Theory, IT-13, 260-269, April, 1967.

[2] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," IEEE Trans, Inform. Theory, IT-20; 284-287, March, 1974.

§ 5.1 Encoder Structure and Trellis Representation

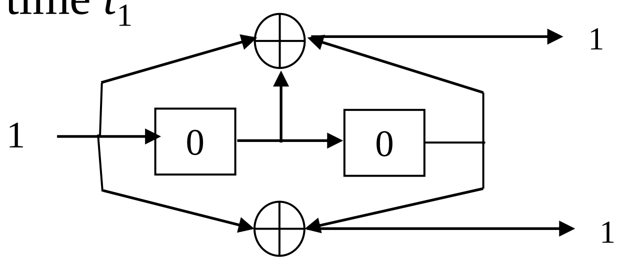
- The $(7, 5)_8$ conv. code
  - Encoder structure:



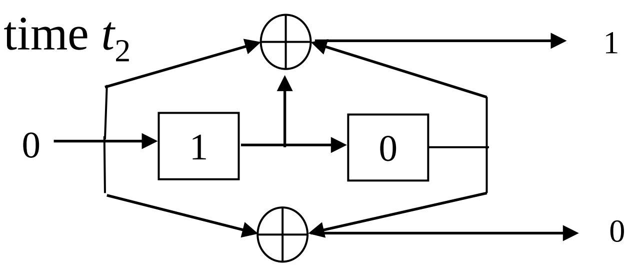
  - Encoding Process:

(Initialized state $S_0S_1 = 00$)

At time t_1



At time t_2



Code rate: ½;

Memory: $m = 2$;

Constraint length: $m + 1 = 3$;

Encoding:

$c_1 = u \oplus S_0 \oplus S_1$;

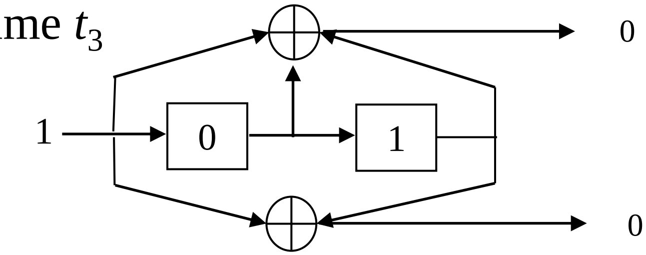
$c_2 = u \oplus S_1$;

Registers update:

$S_1' = S_0$.

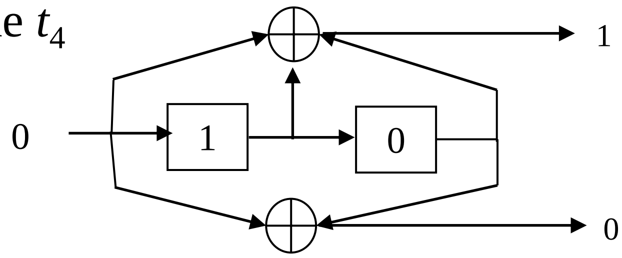
$S_0' = u$.

At time t_3

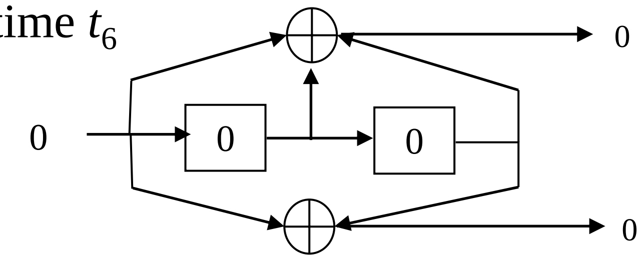


§ 5.1 Encoder Structure and Trellis Representation

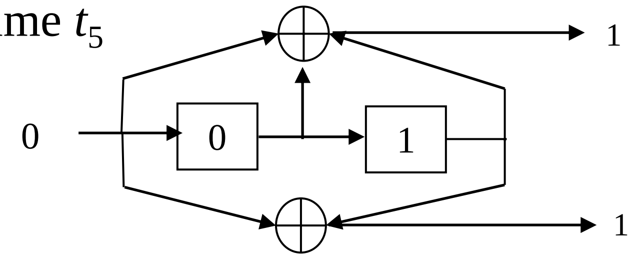
At time t_4



At time t_6



At time t_5

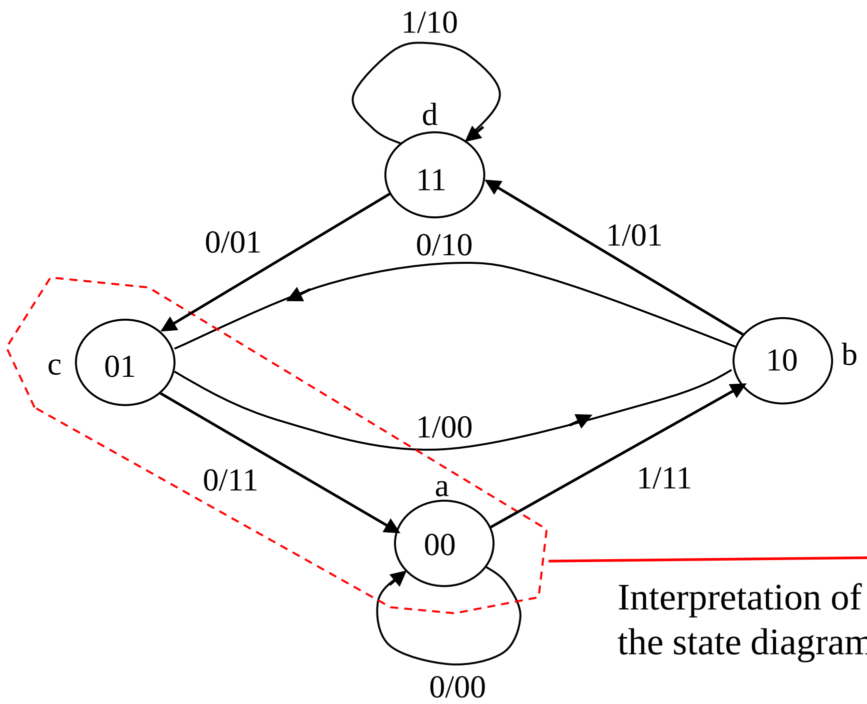


Input sequence $[u_1\ u_2\ u_3\ u_4\ u_5\ u_6] = [1\ 0\ 1\ 0\ 0\ 0]$

Output sequence $[c_1^1 c_1^2\ c_2^1 c_2^2\ c_3^1 c_3^2\ c_4^1 c_4^2\ c_5^1 c_5^2\ c_6^1 c_6^2] = [11\ 10\ 00\ 10\ 11\ 00]$

§ 5.1 Encoder Structure and Trellis Representation

A state transition diagram of the $(7, 5)_8$ conv. code



Interpretation of the state diagram

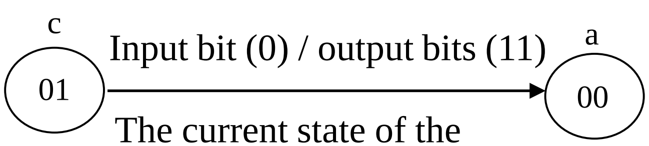
State definition (S_0 S_1)

a = 00

b = 10

c = 01

d = 11



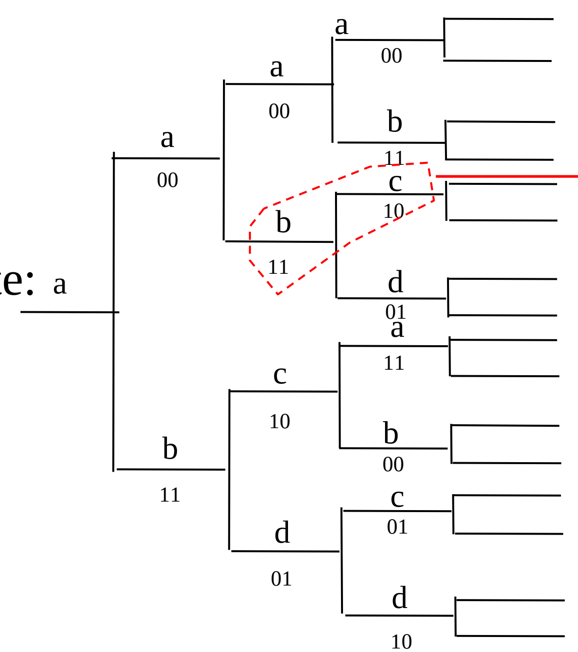
The current state of the encoder is **<u>c</u>**. If the input bit is **<u>0</u>**, it will output **<u>11</u>** and the next state of the encoder is **<u>a</u>**.

§ 5.1 Encoder Structure and Trellis Representation

Tree Representation of the $(7, 5)_8$ conv. code

Time unit: 1 2 3 4 →

Initialised state:



Tree diagram interpretation:

The current state of the encoder is **b**. If the input bit is **0**, the output will be **10**, and the next state of the encoder is **c**.

↑ Input bit as 0

↓ Input bit as 1

$\frac{\text{State after transition}}{\text{Output from transition}}$

Example 5.1 Determine the codeword that corresponds to message [0 1 1 0 1]

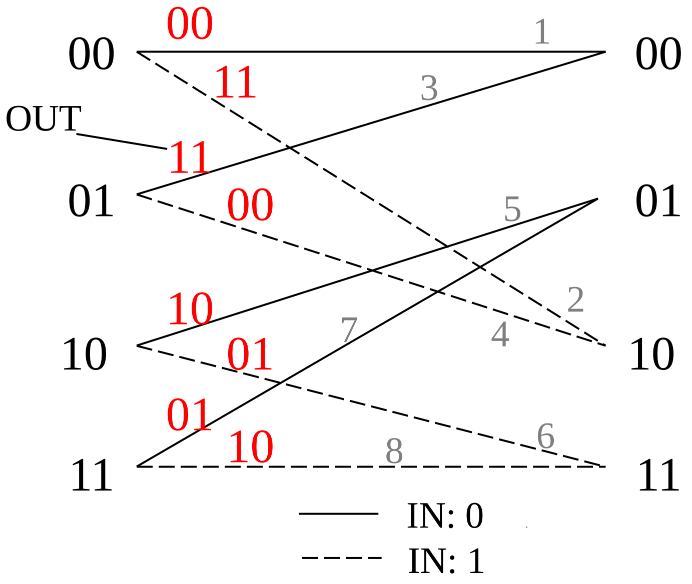
§ 5.1 Encoder Structure and Trellis Representation

Trellis of the $(7, 5)_8$ conv. code

State Table

IN	Current State	Next State	Out	ID
0	00	00	00	1
1	00	10	11	2
0	01	00	11	3
1	01	10	00	4
0	10	01	10	5
1	10	11	01	6
0	11	01	01	7
1	11	11	10	8

Trellis



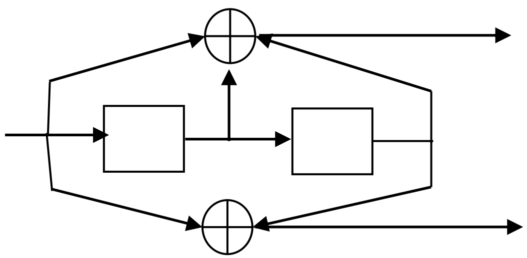
Remark: A trellis tells the state transition and IN/OUT relationship. It can be used to yield a convolutional codeword of a sequential input.

Example 5.2 Use the above trellis to determine the codeword that corresponds to message [0 1 1 0 1].

§ 5.1 Encoder Structure and Trellis Representation

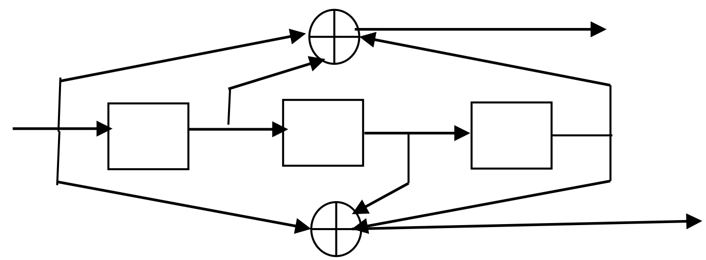
Some practical conv. codes

$(7, 5)_8$ conv. code



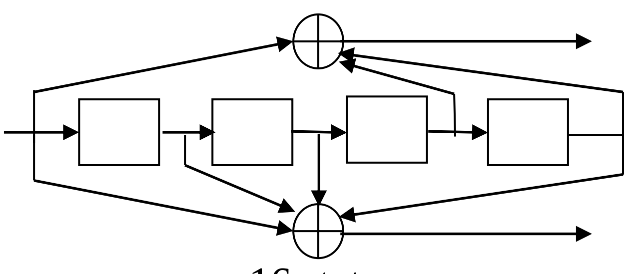
4 states

$(15, 13)_8$ conv. code



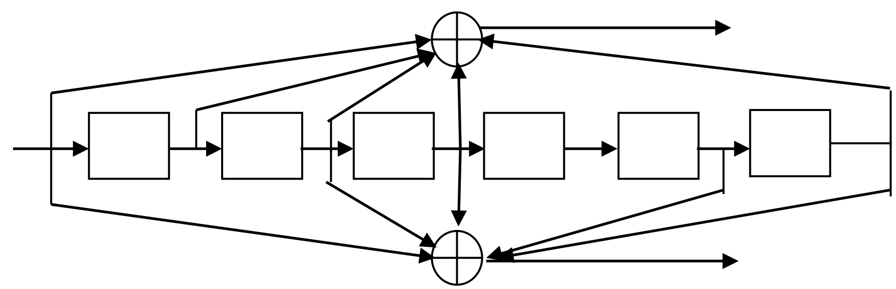
8 states

$(23, 35)_8$ conv. code



16 states

$(171, 133)_8$ conv. code



64 states

Remark: A convolutional code's error-correction capability improves by increasing the number of the encoder states.

Remark: m tailing bits are needed to force the encoder back to the all-zero state. For the above nonsystematic conv. codes, m 0s are needed.

§ 5.1 Encoder Structure and Trellis Representation

The encoder structure can also be represented by generator sequences or transfer functions.

Example 5.3

- The $(7,5)_8$ conv. code can also be written as:

  A rate ½ conv. code with generator sequences

  $$g^{(1)} = [1\ 1\ 1],\ \ g^{(2)} = [1\ 0\ 1]$$

  A rate ½ conv. code with transfer functions

  $$g^{(1)}(x) = 1 + x + x^2,\ \ g^{(2)}(x) = 1 + x^2$$

- The $(15,13)_8$ conv. code can also be written as:

  A rate ½ conv. code with generator sequences:

  $$g^{(1)} = [1\ 1\ 0\ 1],\ \ g^{(2)} = [1\ 0\ 1\ 1]$$

  A rate ½ conv. code with transfer functions:

  $$g^{(1)}(x) = 1 + x + x^3,\ \ g^{(2)}(x) = 1 + x^2 + x^3$$

§ 5.1 Encoder Structure and Trellis Representation

When the codeword (message) length is finite, conv. code is also a linear block code. That says its encoding can be defined by $\bar{c} = \bar{u} \cdot \mathbf{G}$

Example 5.4 For the (7,5) convolutional code with message $\bar{u} = (u_1, u_2, u_3, u_4, u_5)$, we have

$$u_1 \rightarrow \begin{cases} c_1^1 = 0 \oplus 0 \oplus u_1 \\ c_1^2 = 0 \oplus u_1 \end{cases} \quad u_2 \rightarrow \begin{cases} c_2^1 = 0 \oplus u_1 \oplus u_2 \\ c_2^2 = 0 \oplus u_2 \end{cases} \quad u_3 \rightarrow \begin{cases} c_3^1 = u_1 \oplus u_2 \oplus u_3 \\ c_3^2 = u_1 \oplus u_3 \end{cases}$$

$$u_4 \rightarrow \begin{cases} c_4^1 = u_2 \oplus u_3 \oplus u_4 \\ c_4^2 = u_2 \oplus u_4 \end{cases} \quad u_5 \rightarrow \begin{cases} c_5^1 = u_3 \oplus u_4 \oplus u_5 \\ c_5^2 = u_3 \oplus u_5 \end{cases}$$

$$\mathbf{G} = \begin{array}{c} \\ u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \end{array} \begin{array}{c} \begin{array}{cccccccccc} c_1^1 & c_1^2 & c_2^1 & c_2^2 & c_3^1 & c_3^2 & c_4^1 & c_4^2 & c_5^1 & c_5^2 \end{array} \\ \begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \end{array}$$

§ 5.1 Encoder Structure and Trellis Representation

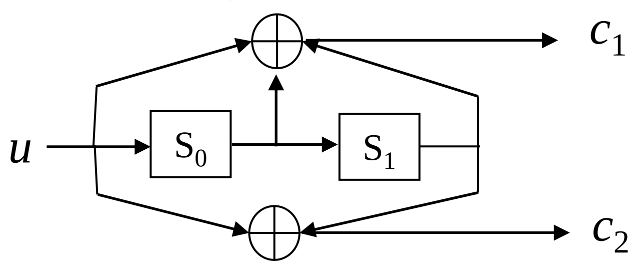
In general, given a rate ½ $m = 2$ conv. code can be defined by $g^{(1)}(x) = g_0^{(1)} + g_1^{(1)}x + g_2^{(1)}x^2$ and $g^{(2)}(x) = g_0^{(2)} + g_1^{(2)}x + g_2^{(2)}x^2$. Its generator matrix **G** is

$$\mathbf{G} = \begin{bmatrix} g_0^{(1)} & g_0^{(2)} & g_1^{(1)} & g_1^{(2)} & g_2^{(1)} & g_2^{(2)} & & & & & & & \\ & & g_0^{(1)} & g_0^{(2)} & g_1^{(1)} & g_1^{(2)} & g_2^{(1)} & g_2^{(2)} & & & & & \\ & & & & g_0^{(1)} & g_0^{(2)} & g_1^{(1)} & g_1^{(2)} & g_2^{(1)} & g_2^{(2)} & & & \\ & & & & & & g_0^{(1)} & g_0^{(2)} & g_1^{(1)} & g_1^{(2)} & \ddots & & \\ & & & & & & & & g_0^{(1)} & g_0^{(2)} & \ddots & & \\ & & & & & & & & & & \ddots & & \end{bmatrix}$$

§ 5.2 Systematic Convolutional Codes

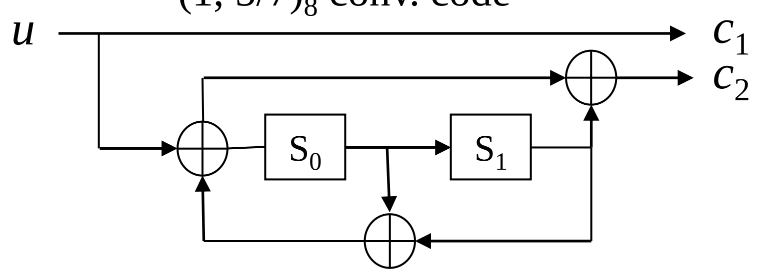
- The $(7, 5)_8$ conv. code's systematic counterpart is:

$(7, 5)_8$ conv. code



Nonsystematic code

$(1, 5/7)_8$ conv. code



Systematic code

Encoding and Registers' updating rules:

$[S_0\ S_1]$ are initialized as $[0\ 0]$;

$c_1 = u$; (systematic feature) feedback $= S_0 \oplus S_1$;

$c_2 = u \oplus$ feedback $\oplus S_1$; $S_1' = S_0$; $S_0' = u \oplus$ feedback;

Remark: Systematic encoding structure is often used to constitute Turbo codes (Chapter 6).

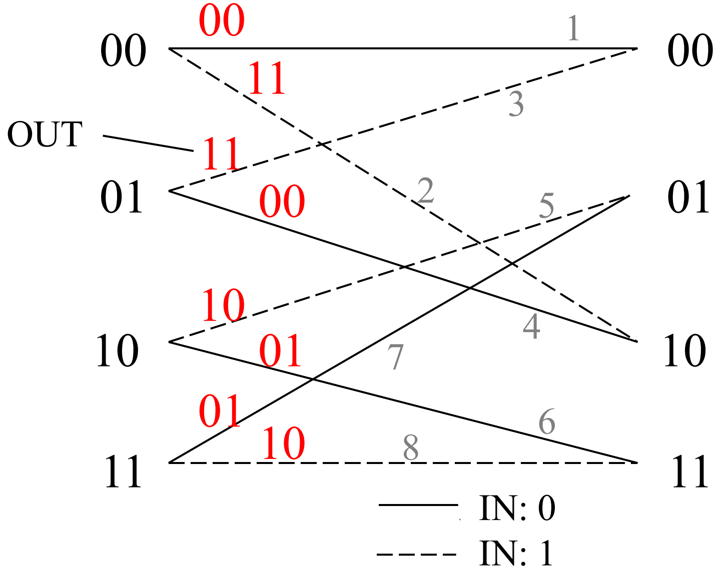
§ 5.2 Systematic Convolutional Codes

For the $(1, 5/7)_8$ conv. code

State Table

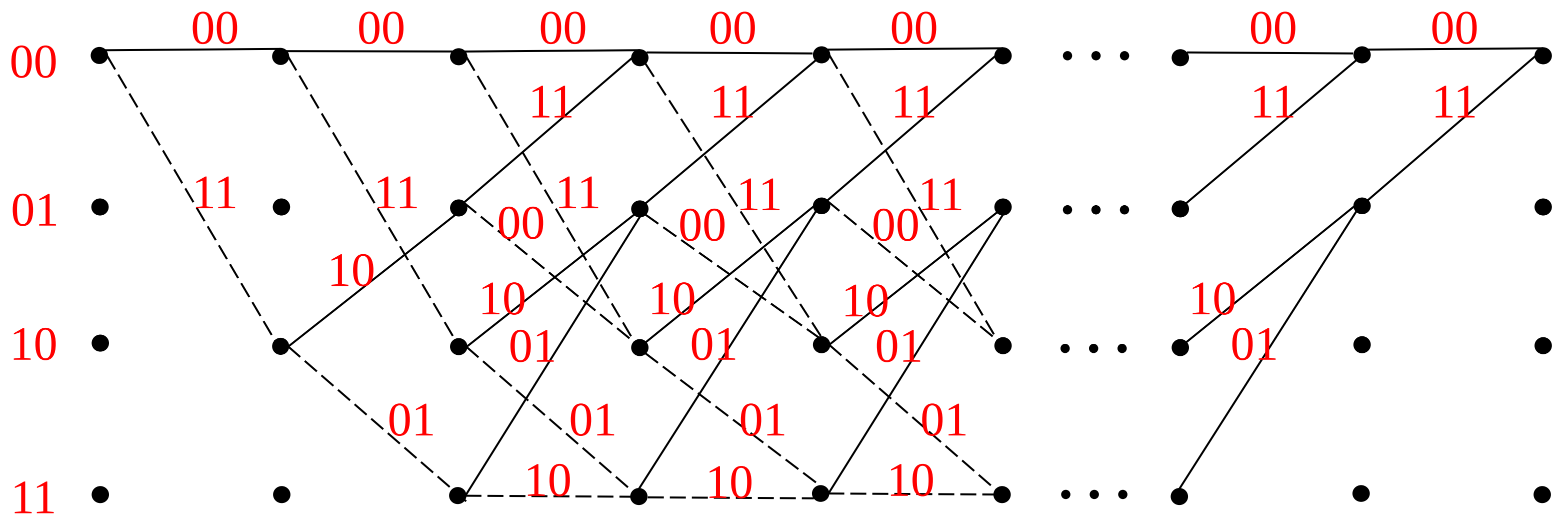
IN	Current State	Next State	Out	ID
0	00	00	00	1
1	00	10	11	2
0	01	10	00	3
1	01	00	11	4
0	10	11	01	5
1	10	01	10	6
0	11	01	01	7
1	11	11	10	8

Trellis



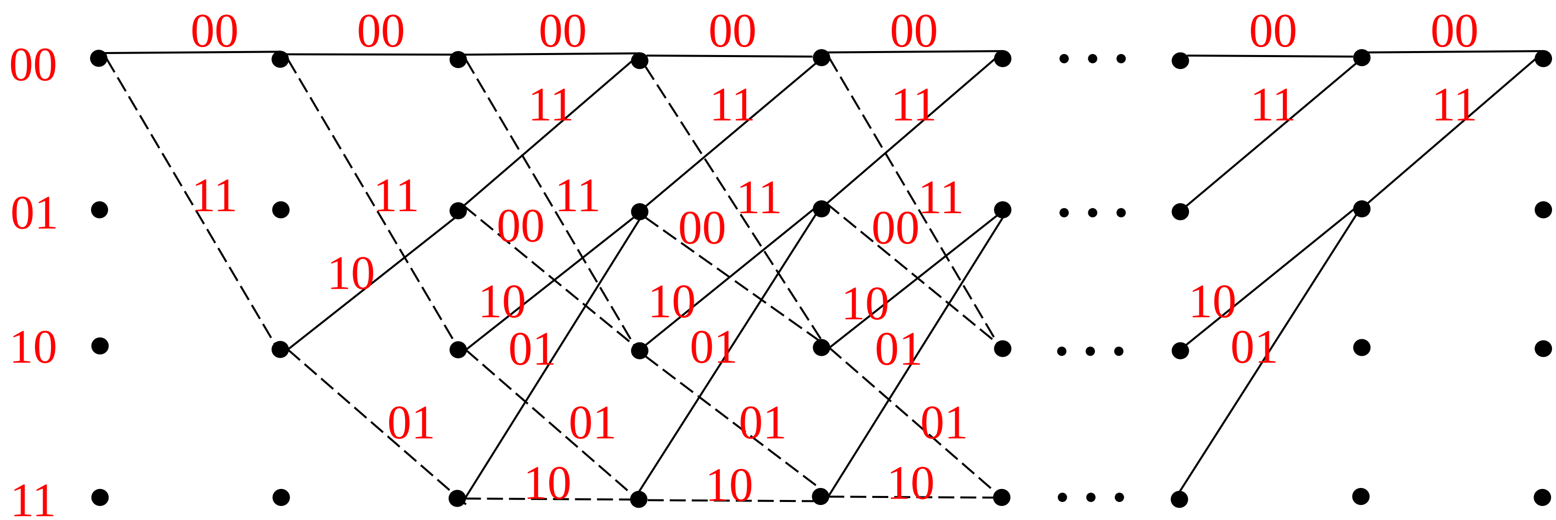
§ 5.3 Viterbi Decoding

Let us extend the trellis of the $(7, 5)_8$ conv. code as if there is a sequential input.



- Such an extension results in a **Viterbi trellis**
- A path in the Viterbi trellis represents a convolutional codeword that corresponds to a sequential input (message).

§ 5.3 Viterbi Decoding



- Decoding motivation: Given a received word $\bar{r}$, find the mostly likely codeword $\hat{\bar{c}}$ such that the Hamming distance $d(\bar{r}, \hat{c})$ is minimized.
- Since $\hat{\bar{c}}$ corresponds to a path in the Viterbi trellis, tracing back the path of $\hat{\bar{c}}$ enables us to find out the message.
- Branch metrics: Hamming distance between a transition branch's output and the corresponding received symbol (or bits).
- Path metrics: Accumulated Hamming distance of the previous branch metrics.

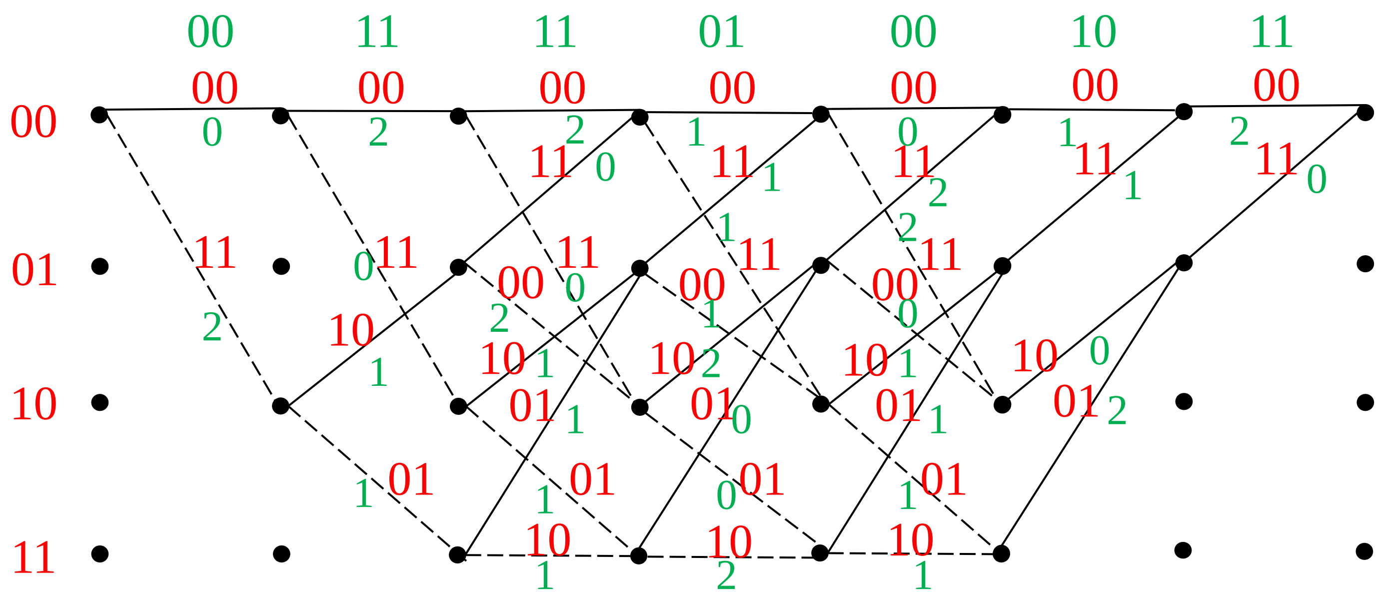
§ 5.3 Viterbi Decoding

Example 5.5 Given the $(7, 5)_8$ conv. code as in *Examples 5.1 - 5.3*. The transmitted codeword is $\bar{c} = [00\ 11\ 01\ 01\ 00\ 10\ 11]$(codeword produced by adding tailing bits). After channel, the received word is

$$\bar{r} = [00\ 11\ \boxed{1}1\ 01\ 00\ 10\ 11]$$

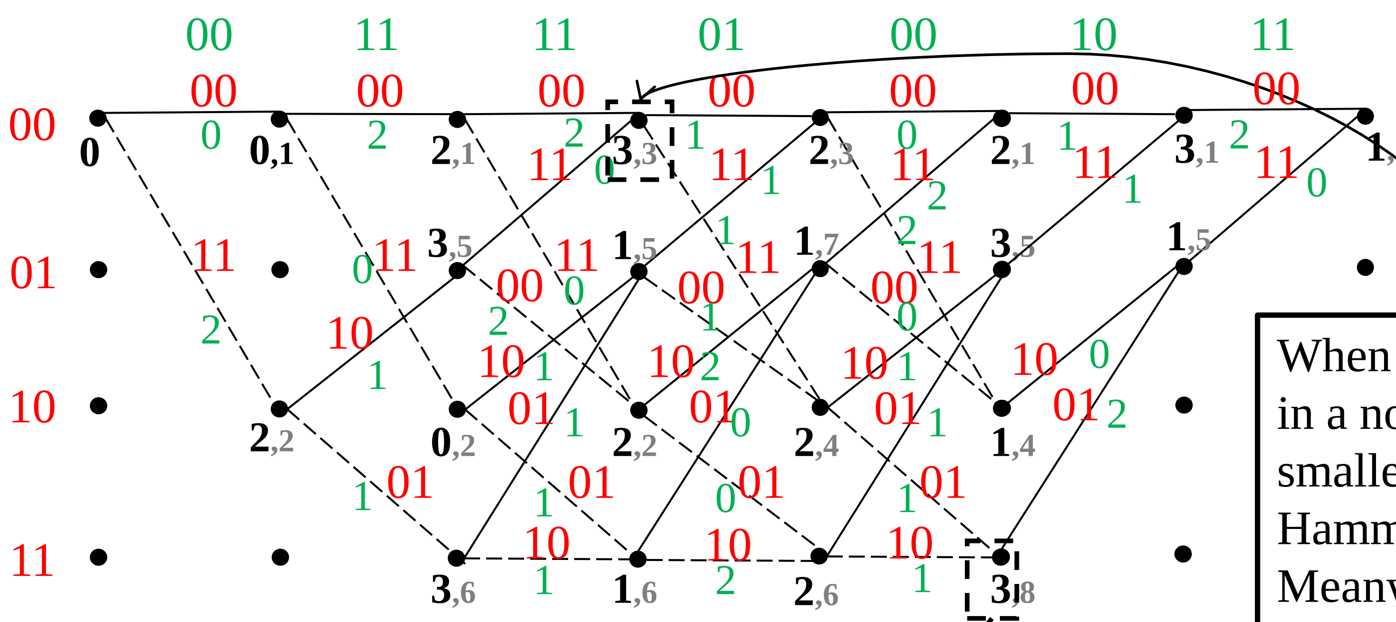
Try to use the following Viterbi trellis to decode it.

Step 1: Calculate all the branch metrics.



§ 5.3 Viterbi Decoding

Step 2: Calculate the path metrics and memorize the path IDs.



When the two joining paths give the same accumulated Hamming distance, pick up one randomly.

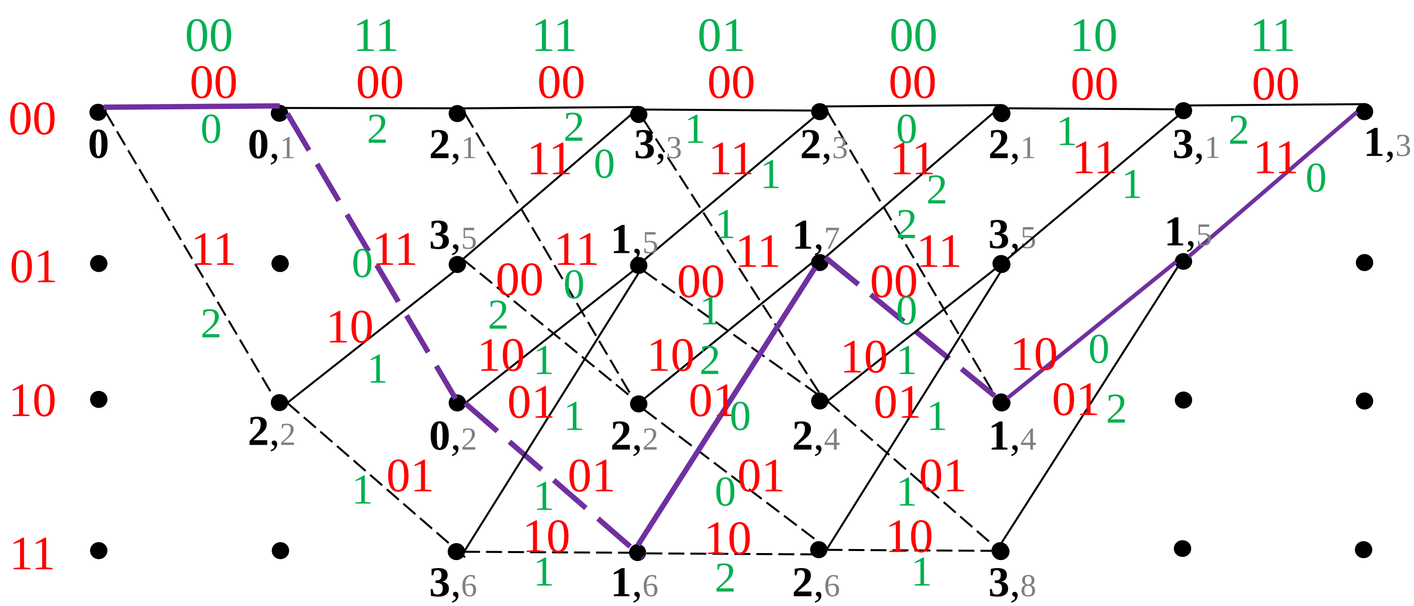
When two paths join in a node, keep the smaller accumulated Hamming distance. Meanwhile, memorize ID of the (selected) branch that leads to the node.

§ 5.3 Viterbi Decoding

Step 3: Pick up the minimal path metric and trace back to determine the message.

Tracing rules: (1) Trellis connection;

(2) The tracing route should match the trellis transition ID.



Decoding output: 0 1 1 0 1 0 0

§ 5.3 Viterbi Decoding

Metrics of the Viterbi Decoding Process

Branch Metrics Table

ID	Branch Metric						
1	0	2	2	1	0	1	2
2	2	0	0	1	2	∞	∞
3	∞	∞	0	1	2	1	0
4	∞	∞	2	1	0	∞	∞
5	∞	1	1	2	1	0	∞
6	∞	1	1	0	1	∞	∞
7	∞	∞	1	0	1	2	∞
8	∞	∞	1	2	1	∞	∞

Path Metrics Table

00	0	0	2	3	2	2	3	1
01	∞	∞	3	1	1	3	1	∞
10	∞	2	0	2	2	1	∞	∞
11	∞	∞	3	1	2	3	∞	∞

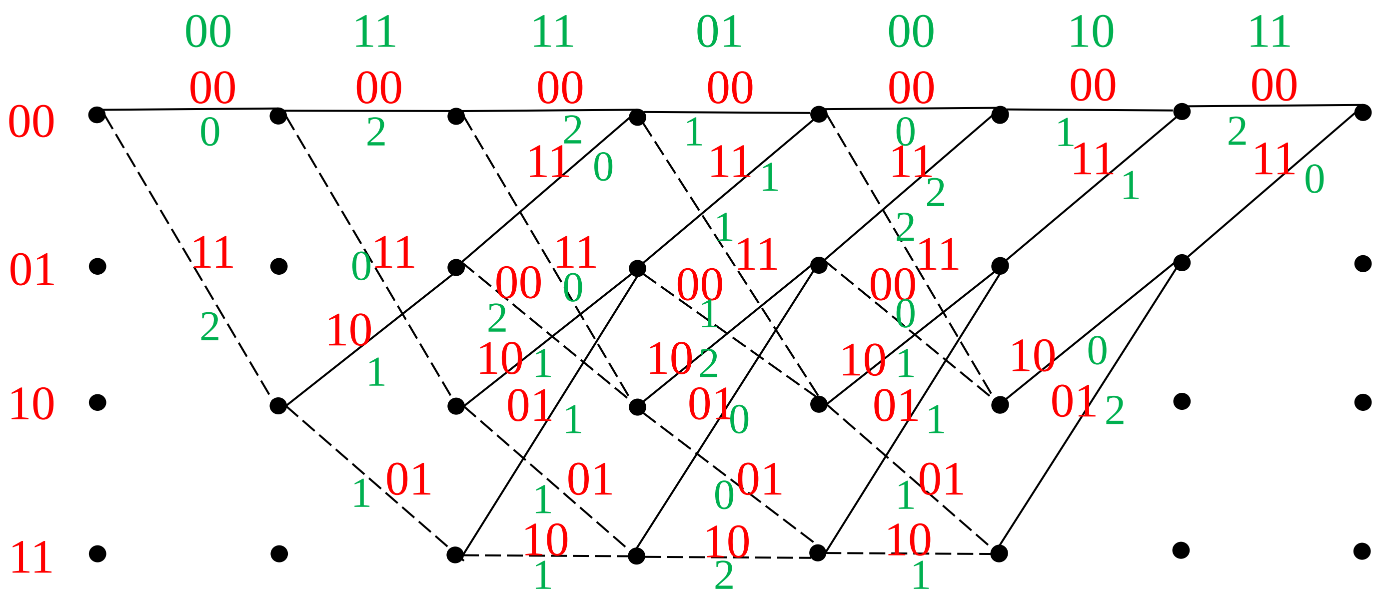
Trellis Transition ID Table

00	1	1	3	3	1	1	3
01	×	5	5	7	5	5	×
10	2	2	2	4	4	×	×
11	×	6	6	6	8	×	×

Remark: With tailing bits, the backward trace always starts from the all-zero state.

§ 5.3 Viterbi Decoding

Free distance of a convolutional code



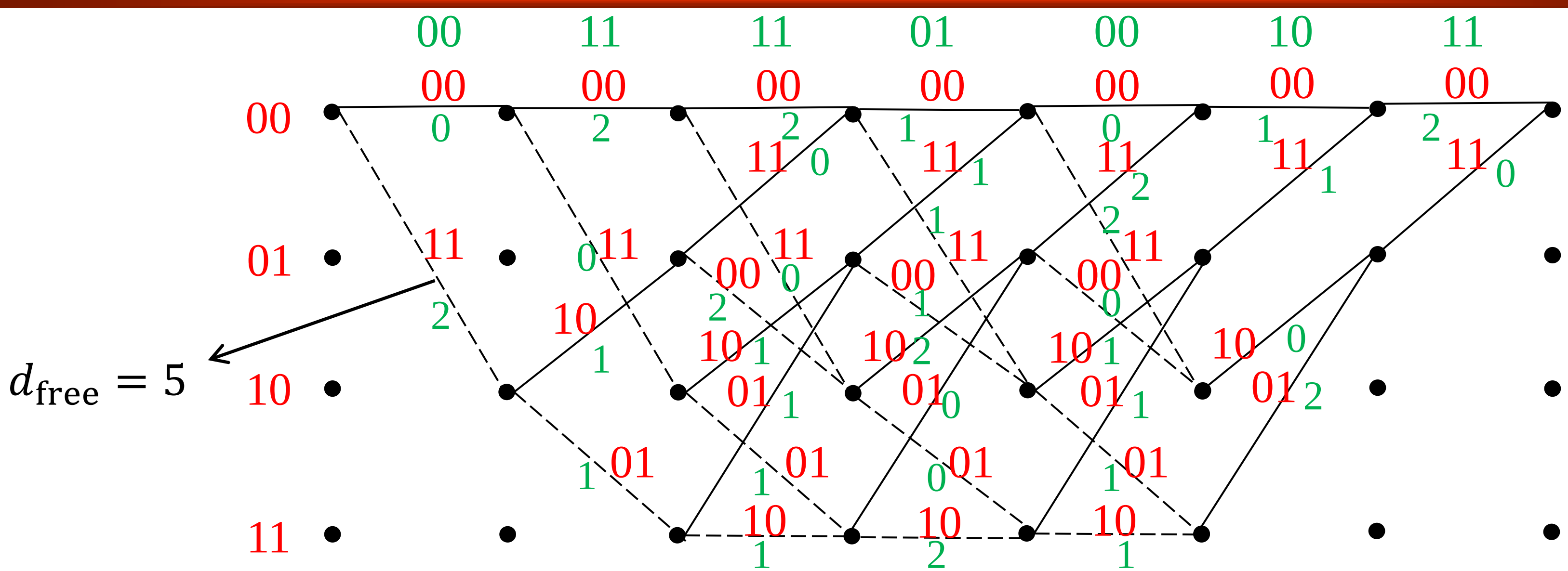
- A conv. code's performance is determined by its free distance.
- Free distance

$$d_{\text{free}} = \min\{d_{\text{Ham}}(\bar{c_1}, \bar{c_2}), \bar{c_1} \neq \bar{c_2}\}$$

- With knowing $\mathbf{0} = [0\ 0\ 0\ \ldots 0]$ is also a convolutional codeword

$$d_{\text{free}} = \min\{weight(\bar{c}), \bar{c} \neq \mathbf{0}\}.$$

§ 5.3 Viterbi Decoding



Hence, it is the minimum weight of all finite length paths in the Viterbi trellis that diverge from and emerge with the all zero state.

- The tailing bits are needed to ensure a greater free distance for the code.
- Convolutional code with a large number of states will have a great d_{free}, and hence stronger error-correction capability

§ 5.3 Viterbi Decoding

Remark: Convolutional code is more competent in correcting spread errors, but not burst errors.

error

E.g., with $\bar{r}_1 = [0\ \ 1\ \ e\ \ 1\ \ \boxed{e}\ \ 1\ \ 0\ \ 1\ \ 0\ \ 0\ \ e\ \ 1]$

and $\bar{r}_2 = [0\ \ 1\ \ 0\ \ e\ \ e\ \ e\ \ 0\ \ 1\ \ 0\ \ 0\ \ 1\ \ 1]$

Viterbi algorithm is more competent in correcting received vector $\bar{r}_1$.

§ 5.4 Soft-Decision Viterbi Decoding

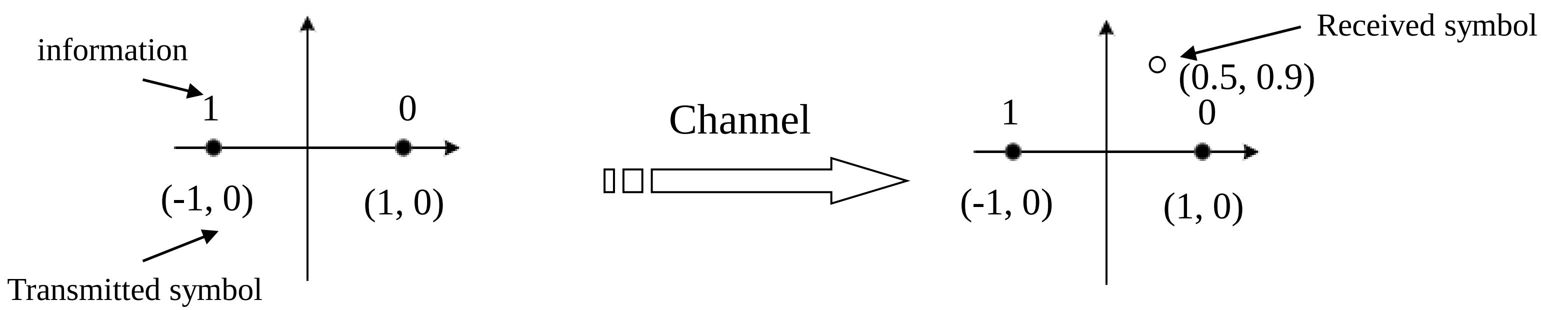
- Soft-decision Viterbi decoding

  • While we are performing the hard-decision Viterbi decoding, we have the scenario that two joining paths yield the same accumulated Hamming distance. This would cause decoding ‘ambiguity’ and performance penalty;
  • Such a performance loss can be compensated by utilizing soft-decision decoding, e.g., soft-decision Viterbi decoding

- Modulation and Demodulation

  • Modulation: mapping the coded symbol into a transmitted symbol;
  • Demodulation: determining the codeword symbol with a received symbol;

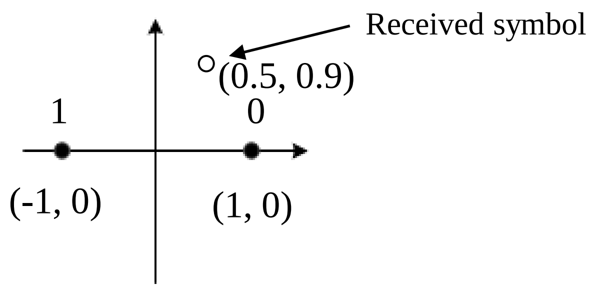
<u>BPSK</u>



<u>Modulation</u> <u>Demodulation</u>

§ 5.4 Soft-Decision Viterbi Decoding

- Modulation and Demodulation (e.g., BPSK)



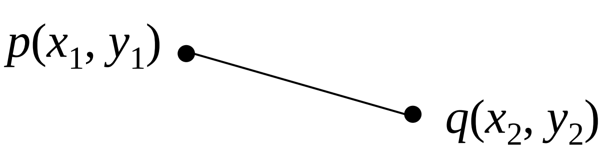
Demodulation

Hard-decision: the information bit is 0. The Hamming distance becomes the Viterbi decoding metrics;

Soft-decision: the information bit has Pr. of 0.7 being 0 and Pr. of 0.3 bing 1. The *Euclidean distance* (*or probability*) becomes the Viterbi decoding metrics;

- Euclidean Distance

Definition: The Euclidean distance between points *p* and *q* is the length of the line segment connecting them.



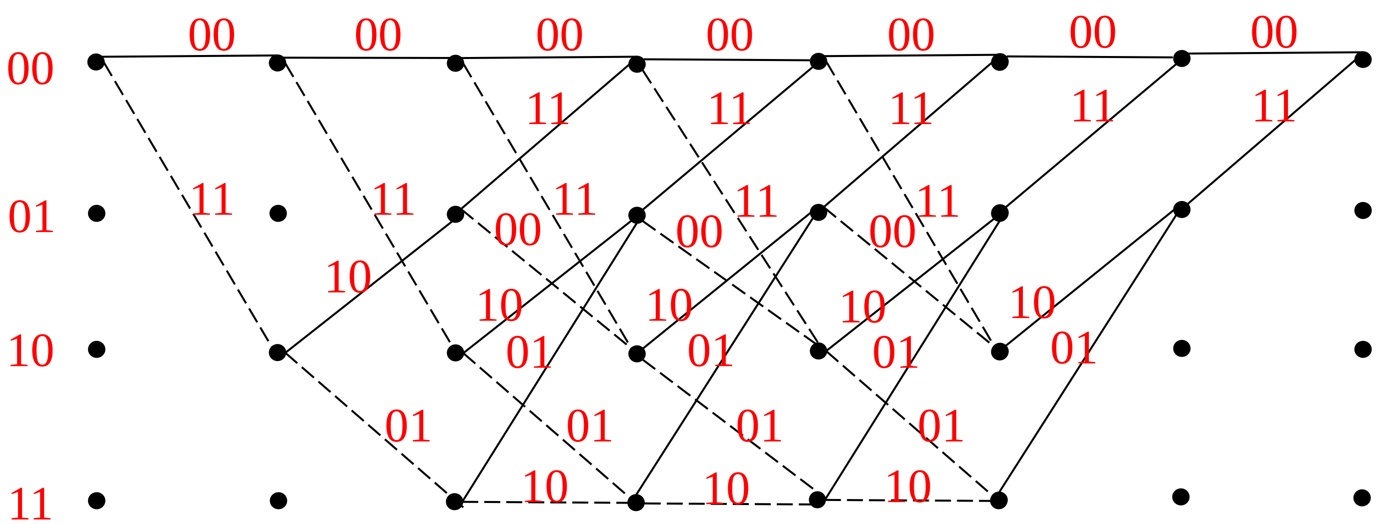
$$d_{Eud} = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

§ 5.4 Soft-Decision Viterbi Decoding

Example 5.6. Given the $(7, 5)_8$ conv. code as in ***Examples 5.5***.
The transmitted codeword is $\bar{c} = [00\ 11\ 01\ 01\ 00\ 10\ 11]$

After BPSK modulation, the transmitted symbols are:
(1, 0), (1, 0), (-1, 0), (-1, 0), (1, 0), (-1, 0), (1, 0), (-1, 0), (1, 0), (1, 0), (-1, 0), (1, 0), (-1, 0), (-1, 0).

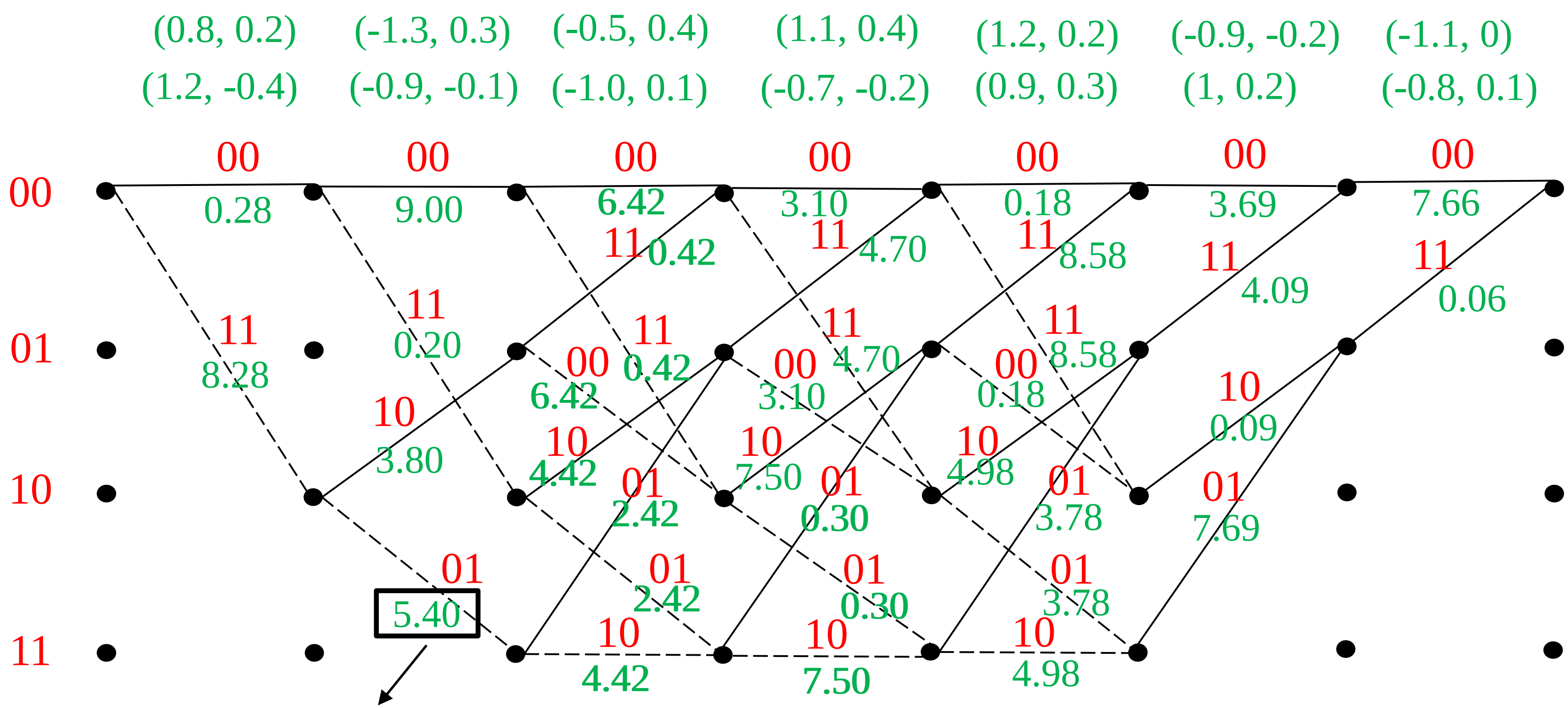
After the channel, the received symbols are:
(0.8, 0.2), (1.2, -0.4), (-1.3, 0.3), (-0.9, -0.1), (-0.5, 0.4), (-1.0, 0.1), (1.1, 0.4), (-0.7, -0.2),
(1.2, 0.2), (0.9, 0.3), (-0.9, -0.2), (1, 0.2), (-1.1, 0), (-0.8, 0.1).



§ 5.4 Soft-Decision Viterbi Decoding

$0 \rightarrow (1, 0)$
$1 \rightarrow (-1, 0)$

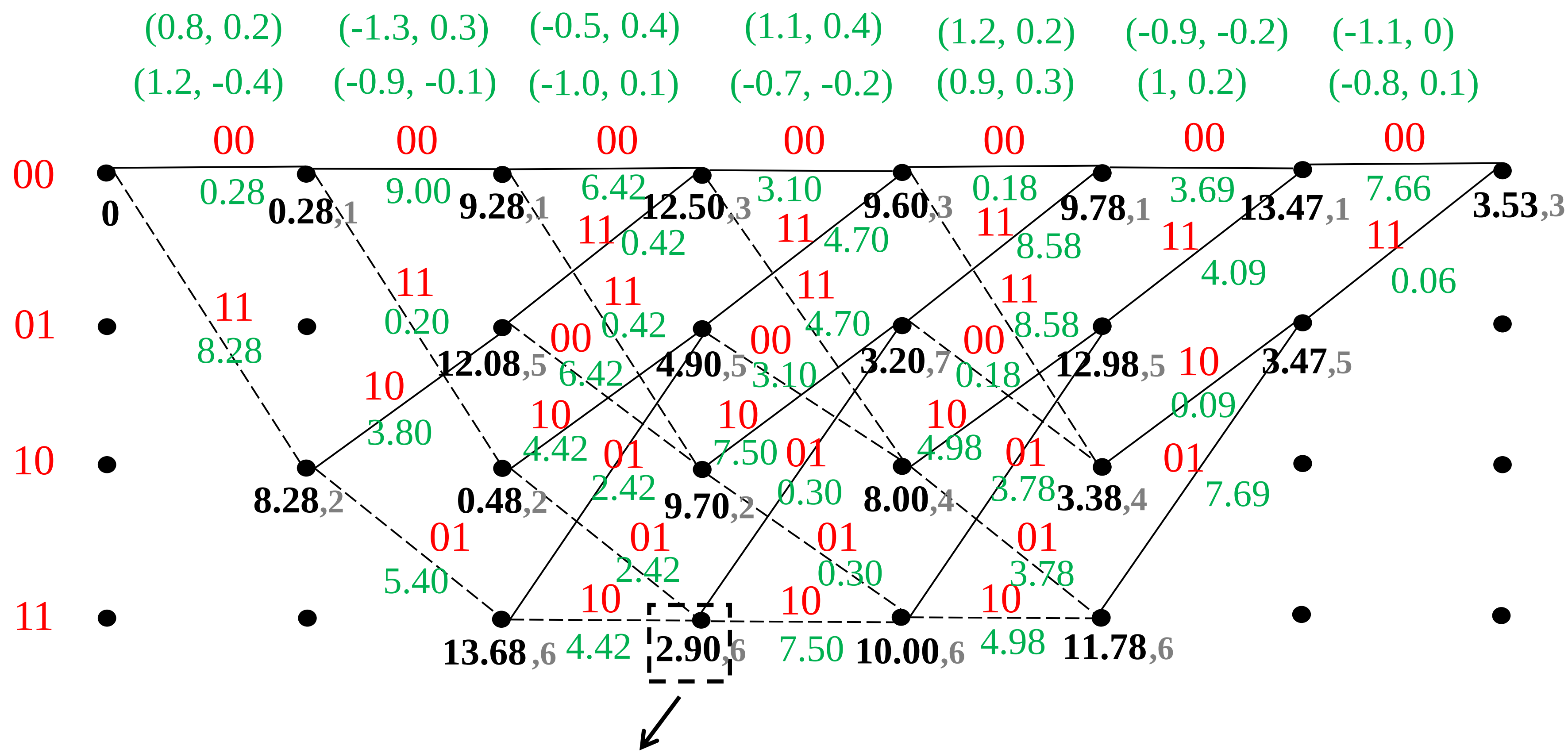
Step 1: Calculate all the branch metrics.



Branch metric: $(-1.3-1)^2 + (0.3-0)^2 + (-0.9+1)^2 + (-0.1-0)^2 = 5.40$

§ 5.4 Soft-Decision Viterbi Decoding

Step 2: Calculate the path metrics and memorize the path IDs.

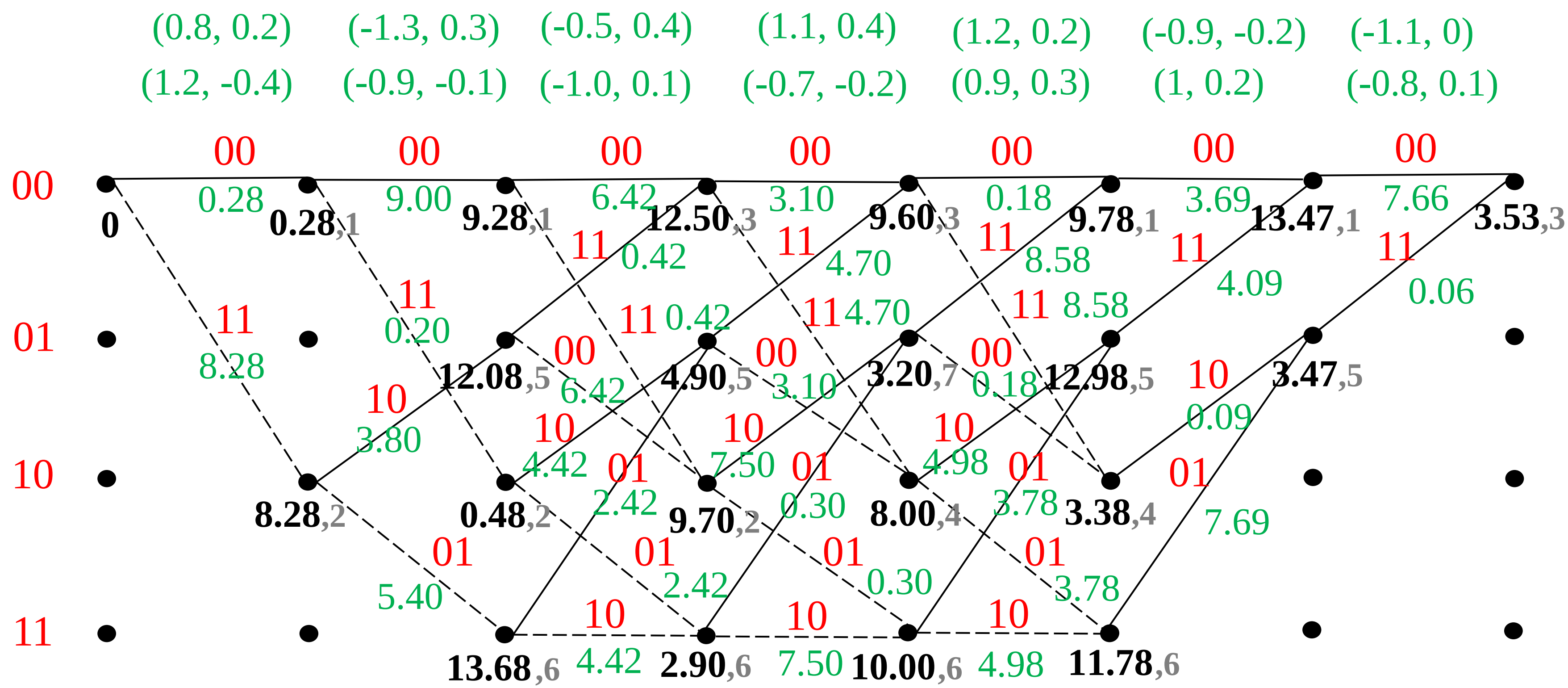


When two paths join in a node, keep the smaller accumulated squared Euclidean distance.

§ 5.4 Soft-Decision Viterbi Decoding

Step 3: Pick up the minimal path metric and trace back to determine the message.

Tracing rules: The same as hard-decision Viterbi decoding algorithm.



Decoding output: 0 1 1 0 1 0 0

§ 5.4 Soft-Decision Viterbi Decoding

Metrics of the Soft-Decision Viterbi Decoding Process

Path Metrics Table

00	0	0.28	0.98	12.50	9.60	9.78	13.47	3.53
01	∞	∞	12.08	4.90	3.20	12.98	3.47	∞
10	∞	8.28	0.48	9.70	8.00	3.38	∞	∞
11	∞	∞	13.68	2.90	10.00	11.78	∞	∞

Branch Metrics Table

ID	**Branch Metric**						
1	0.28	9.00	6.42	3.10	0.18	3.69	7.66
2	8.28	0.20	0.42	4.70	8.58	∞	∞
3	∞	∞	0.42	4.70	8.58	4.09	0.06
4	∞	∞	6.42	3.10	0.18	∞	∞
5	∞	3.80	4.42	7.50	4.98	0.09	∞
6	∞	5.40	2.42	0.30	3.78	∞	∞
7	∞	∞	2.42	0.30	3.78	7.69	∞
8	∞	∞	4.42	7.50	4.98	∞	∞

Trellis Transition ID Table

00	1	1	3	3	1	1	3
01	×	5	5	7	5	5	×
10	2	2	2	4	4	×	×
11	×	6	6	6	6	×	×

§ 5.5 BCJR Decoding

- Hard-decision Viterbi Decoding : a Hard-In-Hard-Out (HIHO) decoding.
  Soft-decision Viterbi Decoding : a Soft-In-Hard-Out (SIHO) decoding.
  BCJR Decoding: a Soft-In-Soft-Out (SISO) decoding.

- A Soft-In-Soft-Out (SISO) decoding algorithm that takes probabilities as the input and delivers probabilities as the output.

- With an attempt to deliver both the *a posteriori probabilities* of $P(c_t|y_t)$ and $P(u_{t'}|y_t)$, it is also called the maximum *a posteriori* (MAP) algorithm.

<u>In light of a rate ½ code</u>

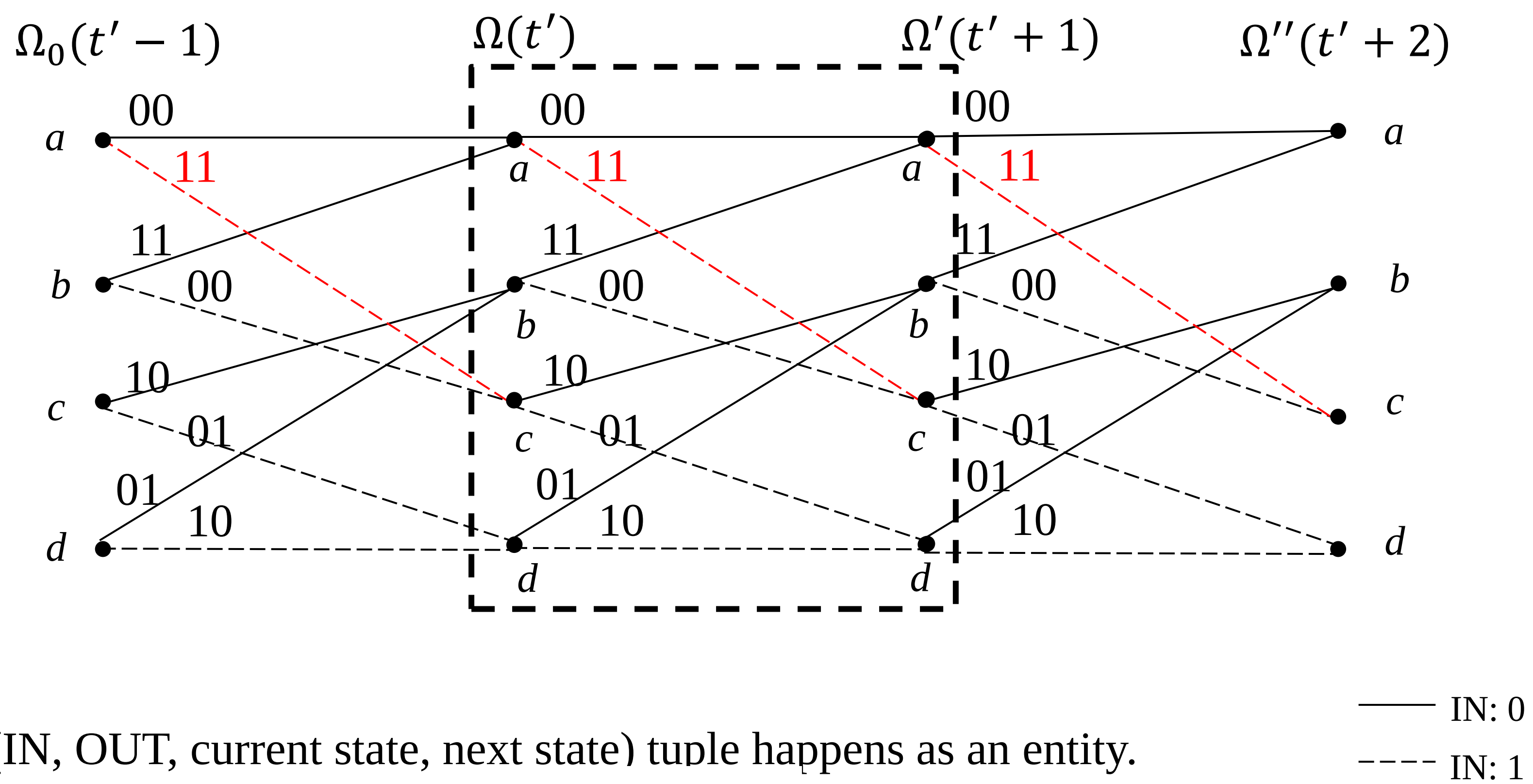
message symbols: $u_1, u_2, \ldots, u_{t'}, \ldots$

codeword symbols: $c_1, c_2, \ldots, c_t, \ldots$ $\longrightarrow$ $c_1^1, c_1^2, c_2^1, c_2^2, \ldots, c_{t'}^1, c_{t'}^2, \ldots$

received symbols: $y_1, y_2, \ldots, y_t, \ldots$ $\longrightarrow$ $y_1^1, y_1^2, y_2^1, y_2^2, \ldots, y_{t'}^1, y_{t'}^2, \ldots$

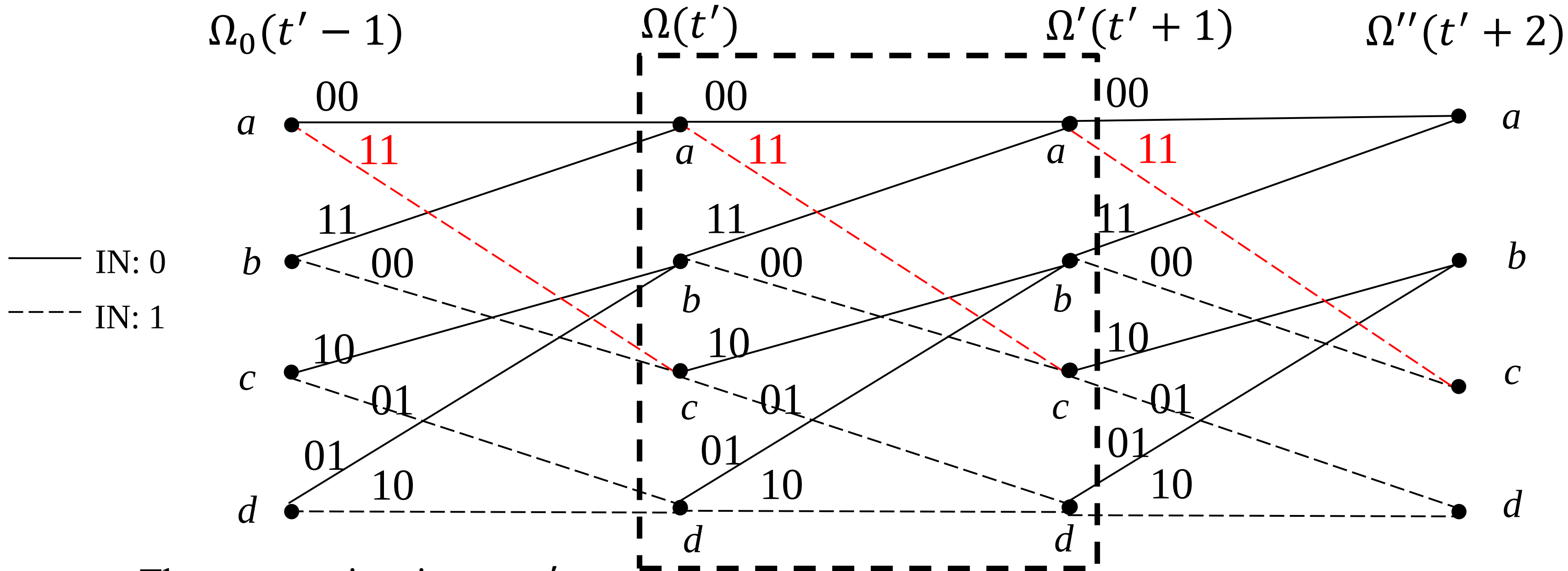
§ 5.5 BCJR Decoding

- In a trellis (e.g., trellis of the $(7, 5)_8$ conv. code).



The (IN, OUT, current state, next state) tuple happens as an entity.

§ 5.5 BCJR Decoding



That says at time instant t'

$\sum$ Prob [trellis transition w.r.t. an input θ] = Prob [$u_{t'} = \theta$], $\theta \in \{ 0,1 \}$.

$\sum$ Prob [trellis transition w.r.t. an output of θ] = Prob [$c_{t'}^{1(2)} = \theta$], $\theta \in \{ 0,1 \}$.

We seek to determine all the Prob[trellis transition w.r.t. an **input** θ] at time t' to know

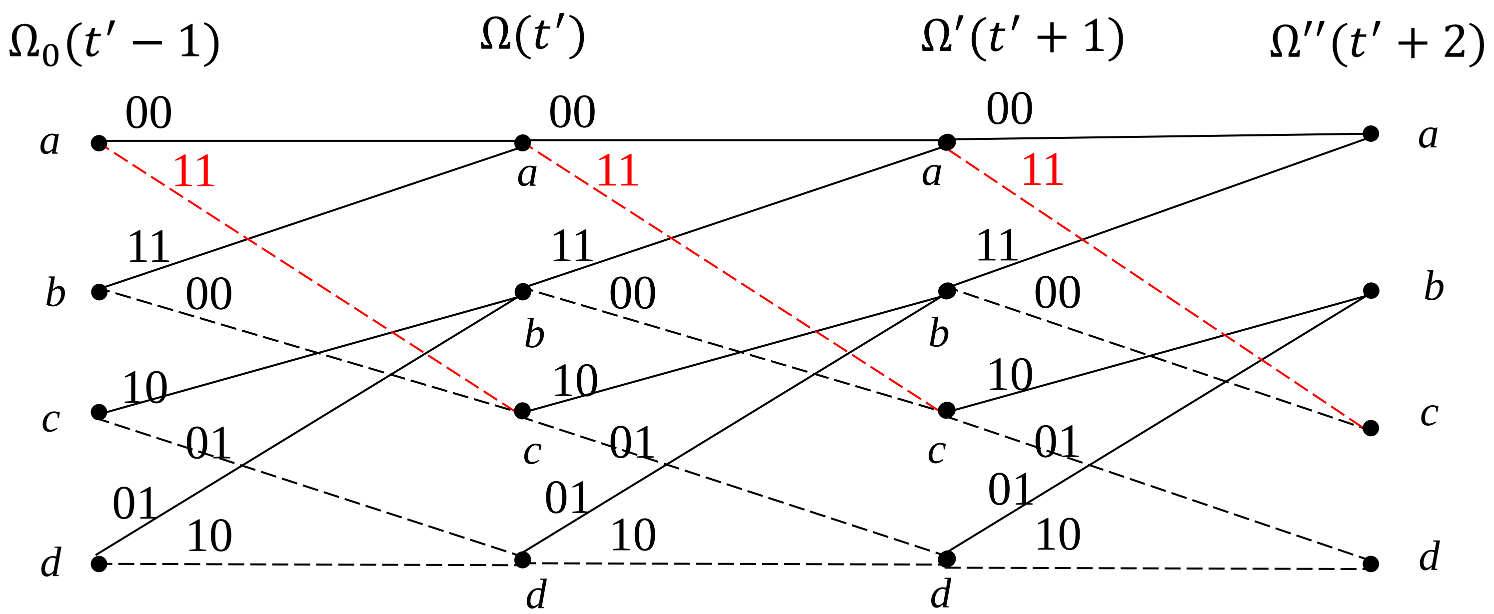
$$P(u_{t'} = \theta | y_t)$$

We seek to determine all the Prob[trellis transition w.r.t. an **output** θ] at time t' to know

$$P\left(c_{t'}^{1(2)} = \theta | y_t\right)$$

§ 5.5 BCJR Decoding

- Determine the state transition probabilities.

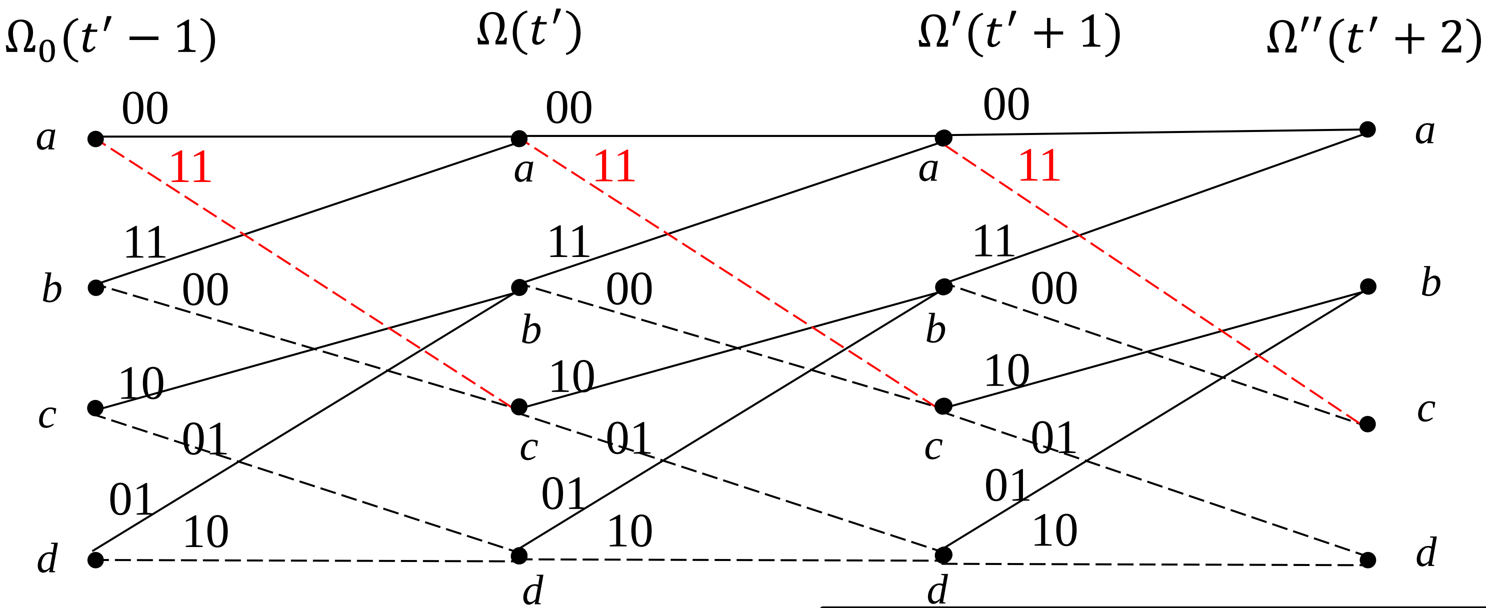


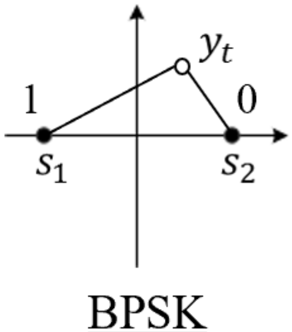
- For a rate half conv. code, $u_{t'} \rightarrow c_{t'}^1, c_{t'}^2$
- Trellis state transition probability: $(\Omega, \Omega') \in \{a, \ b, \ c, \ d\}$

$$\Gamma_{\Omega \rightarrow \Omega'} = P_a(u_{t'}) \ P_{\mathrm{ch}}(c_{t'}^1) P_{\mathrm{ch}}(c_{t'}^2)$$

§ 5.5 BCJR Decoding

- Determine the state transition probabilities.





$$\Gamma_{\Omega\to\Omega'} = P_a(u_{t'})\ P_{\mathrm{ch}}\left(c_{t'}^1\right)P_{\mathrm{ch}}\left(c_{t'}^2\right)$$

A priori prob. of information bit. E.g. w/o knowledge of $u_{t'}$,
$P_a(u_{t'}=0)$= $P_a(u_{t'}=1)=0.5$

Channel observations: E.g. BPSK is used

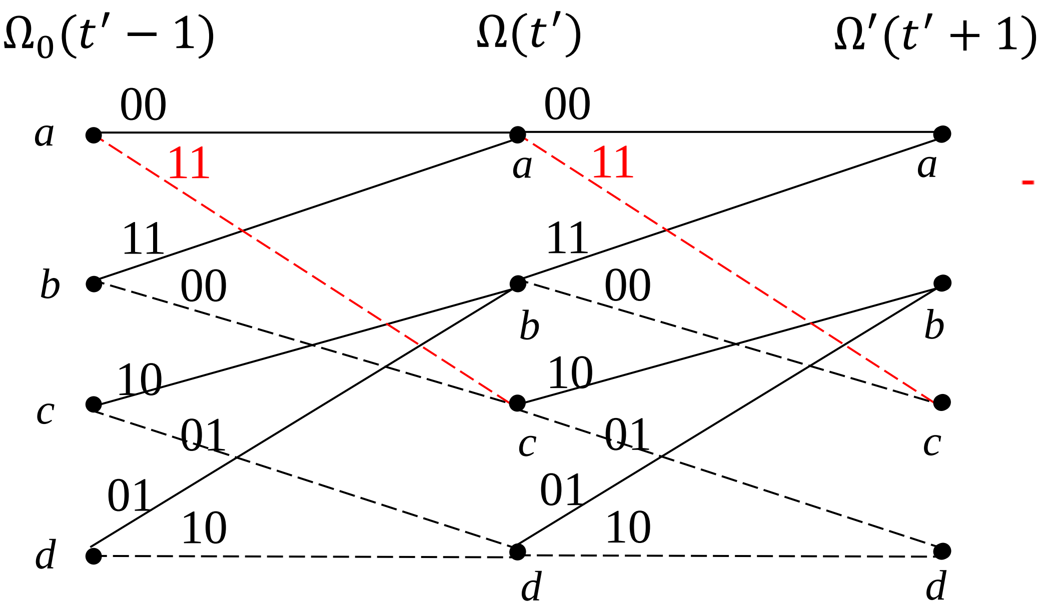
$$P_{\mathrm{ch}}\left(c_{t'}^1=0\right)=P\left(y_t\middle|c_{t'}^1=0\right)=\frac{1}{\sqrt{\pi N_0}}\exp\left(-\frac{\|y_t-s_2\|^2}{N_0}\right)$$

$$P_{\mathrm{ch}}\left(c_{t'}^1=1\right)=P\left(y_t\middle|c_{t'}^1=0\right)=\frac{1}{\sqrt{\pi N_0}}\exp\left(-\frac{\|y_t-s_1\|^2}{N_0}\right)$$

$P_{\mathrm{ch}}\left(c_{t'}^2\right)$ can be calculated similarly.

§ 5.5 BCJR Decoding

- Determine the probability of each beginning state.



- Probability of beginning a trellis transition ($\Omega \rightarrow \Omega'$) from state Ω (Determined by a forward trace).

$$A_{t'}(\Omega) = N_A \sum_{(\Omega_0, \Omega)} A_{t'-1}(\Omega_0) \cdot \Gamma_{\Omega_0 \to \Omega} ,$$
$$t' = 1, 2, \cdots, k.$$

N_A: Normalization factor that ensures $A_{t'}(a) + A_{t'}(b) + A_{t'}(c) + A_{t'}(d) = 1$.

- Knowing the Viterbi trellis starts from the all-zero state, we initialize:

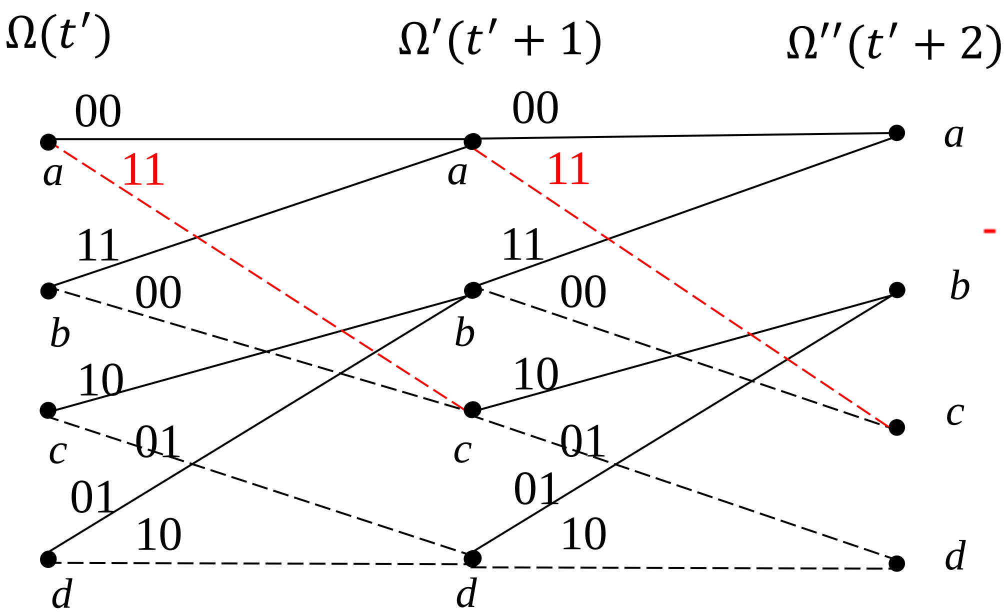
$$A_0(a) = 1, \text{ and } A_0(b) = A_0(c) = A_0(d) = 0.$$

- E.g., in the highlighted trellis transition

$$A_{t'}(a) = A_{t'-1}(a) \cdot \Gamma_{a \to a} + A_{t'-1}(b) \cdot \Gamma_{b \to a}.$$

§ 5.5 BCJR Decoding

- Determine the probability of each ending state.



- Probability of ending the trellis transition ($\Omega \rightarrow \Omega'$) at state Ω' (Determined by a backward trace).

  $B_{t'+1}(\Omega') = N_B \sum_{(\Omega', \Omega'')} B_{t'+2}(\Omega'') \cdot \Gamma_{\Omega' \rightarrow \Omega''}$.

  N_B: Normalization factor that ensures $B_{t'+1}(a) + B_{t'+1}(b) + B_{t'+1}(c) + B_{t'+1}(d) = 1$.

- By ensuring after encoding, the shift registers (encoder) are restored to the all zero state (achieved by bit tailing), we can initialize:

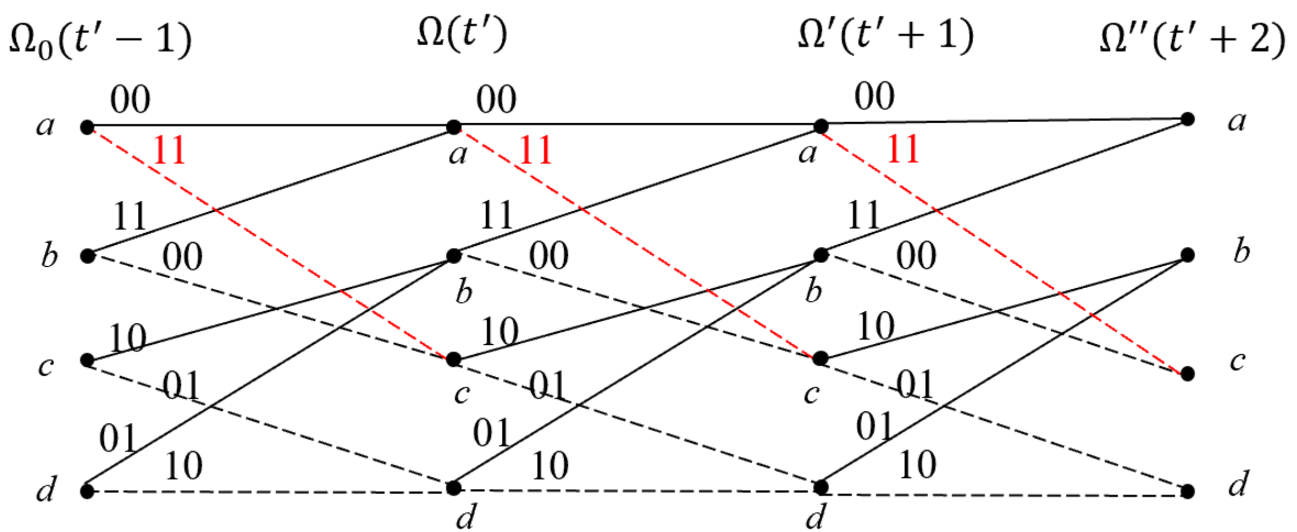
$$B_{k+2}(a) = 1, \text{ and } B_{k+2}(b) = B_{k+2}(c) = B_{k+2}(d) = 0.$$

- E.g., in the highlighted trellis transition

$$B_{t'+1}(c) = B_{t'+2}(b) \cdot \Gamma_{c \rightarrow b} + B_{t'+2}(d) \cdot \Gamma_{c \rightarrow d}.$$

§ 5.5 BCJR Decoding

- Determine the *a posteriori* probability of each information bit



After the **Forward Trace** and **Backward Trace**, we obtain all the $A_{t'}(\Omega), B_{t'+1}(\Omega')$ and $\Gamma_{\Omega\to\Omega'}$ of each time instant t'. We can now determine the *a posteriori probabilities* $P(u_{t'}|y_t)$ for each information bit as

$$P(u_{t'}=0|y_t) = N_P \sum_{\underline{(\Omega,\Omega')_0}} A_{t'}(\Omega)\cdot\Gamma_{\Omega\to\Omega'}\cdot B_{t'+1}(\Omega')$$

State transition indicated by ——

$$P(u_{t'}=1|y_t) = N_P \sum_{\underline{(\Omega,\Omega')_1}} A_{t'}(\Omega)\cdot\Gamma_{\Omega\to\Omega'}\cdot B_{t'+1}(\Omega')$$

State transition indicated by ----

N_p: Normalization factor that ensures $P(u_{t'}=0|y_t)+P(u_{t'}=1|y_t)=1$

§ 5.5 BCJR Decoding

- E.g.,

$$P(u_{t'}=0|y_t) \;= N_P\cdot(A_{t'}(a)\cdot\Gamma_{a\to a}\cdot B_{t'+1}(a)+A_{t'}(b)\cdot\Gamma_{b\to a}\cdot B_{t'+1}(a)$$
$$A_{t'}(c)\cdot\Gamma_{c\to b}\cdot B_{t'+1}(b)+A_{t'}(d)\cdot\Gamma_{d\to b}\cdot B_{t'+1}(b)).$$

$$N_p = P(u_{t'}=0|y_t)+P(u_{t'}=1|y_t)$$

- Decision based on the *a posteriori* probabilities.

$$\hat{u}_{t'}=0, \text{ if } P(u_{t'}=0|y_t)\geq P(u_{t'}=1|y_t)$$

$$\hat{u}_{t'}=1, \text{ if } P(u_{t'}=1|y_t)> P(u_{t'}=0|y_t)$$

Similarly, $P\left(c_{t'}^1=0|y_t\right)$, $P\left(c_{t'}^1=1|y_t\right)$, $P\left(c_{t'}^2=0|y_t\right)$, $P\left(c_{t'}^2=1|y_t\right)$ can be made.

§ 5.5 BCJR Decoding

Example 5.7. With the same transmitted codeword and received symbols of ***Example 5.6***, use the BCJR algorithm to decode it.

With the received symbols, we can determine

$$\begin{cases} P_{ch}(c_1^1 = 0) = 0.83 \\ P_{ch}(c_1^1 = 1) = 0.17 \end{cases} \quad \begin{cases} P_{ch}(c_1^2 = 0) = 0.92 \\ P_{ch}(c_1^2 = 1) = 0.08 \end{cases} \quad \begin{cases} P_{ch}(c_2^1 = 0) = 0.07 \\ P_{ch}(c_2^1 = 1) = 0.93 \end{cases} \quad \begin{cases} P_{ch}(c_2^2 = 0) = 0.14 \\ P_{ch}(c_2^2 = 1) = 0.86 \end{cases}$$

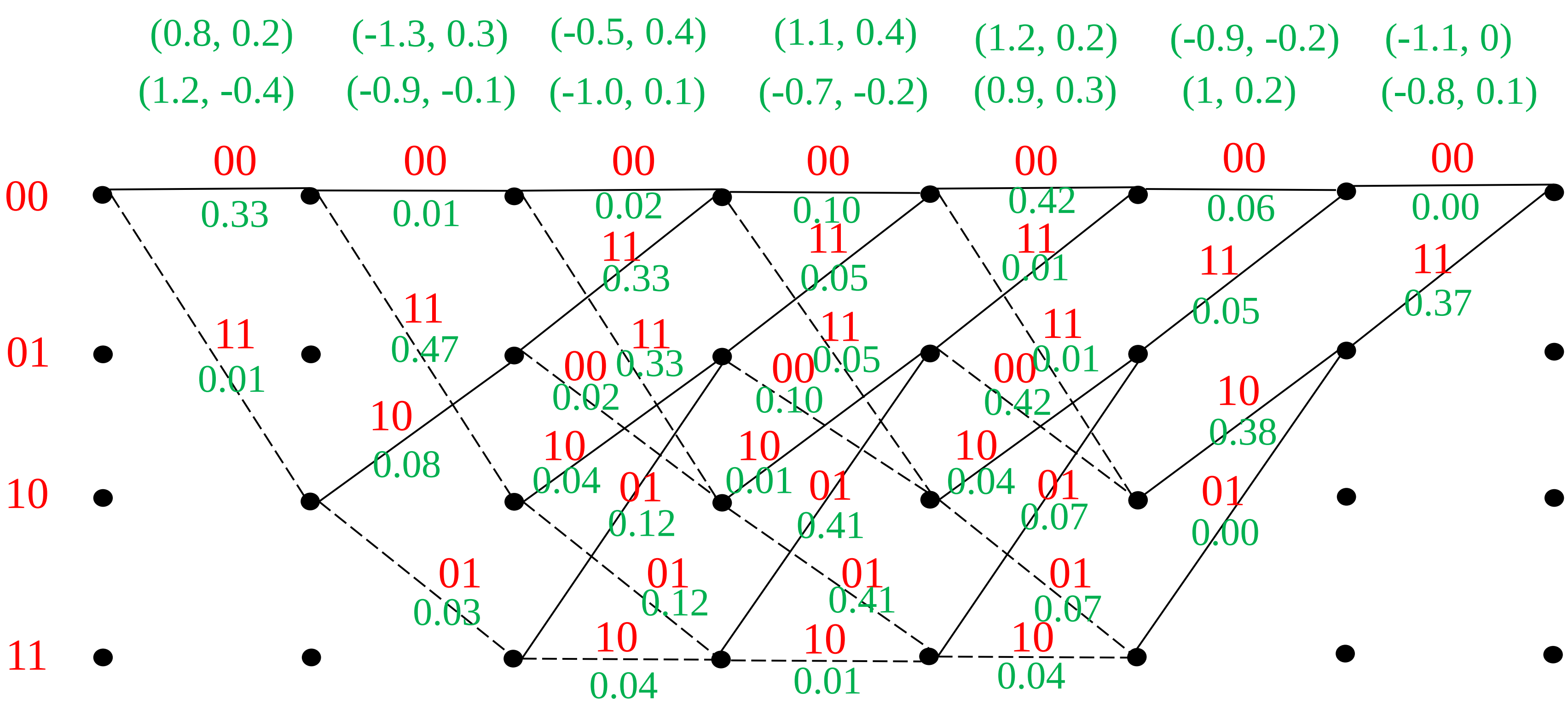
$$\begin{cases} P_{ch}(c_3^1 = 0) = 0.27 \\ P_{ch}(c_3^1 = 1) = 0.73 \end{cases} \quad \begin{cases} P_{ch}(c_3^2 = 0) = 0.12 \\ P_{ch}(c_3^2 = 1) = 0.88 \end{cases} \quad \begin{cases} P_{ch}(c_4^1 = 0) = 0.90 \\ P_{ch}(c_4^1 = 1) = 0.10 \end{cases} \quad \begin{cases} P_{ch}(c_4^2 = 0) = 0.20 \\ P_{ch}(c_4^2 = 1) = 0.80 \end{cases}$$

$$\begin{cases} P_{ch}\left(c_5^1 = 0\right) = 0.92 \\ P_{ch}\left(c_5^1 = 1\right) = 0.08 \end{cases} \quad \begin{cases} P_{ch}\left(c_5^2 = 0\right) = 0.86 \\ P_{ch}\left(c_5^2 = 1\right) = 0.14 \end{cases} \quad \begin{cases} P_{ch}(c_6^1 = 0) = 0.14 \\ P_{ch}(c_6^1 = 1) = 0.86 \end{cases} \quad \begin{cases} P_{ch}(c_6^2 = 0) = 0.88 \\ P_{ch}(c_6^2 = 1) = 0.12 \end{cases}$$

$$\begin{cases} P_{ch}(c_7^1 = 0) = 0.10 \\ P_{ch}(c_7^1 = 1) = 0.90 \end{cases} \quad \begin{cases} P_{ch}(c_7^2 = 0) = 0.17 \\ P_{ch}(c_7^2 = 1) = 0.83 \end{cases}$$

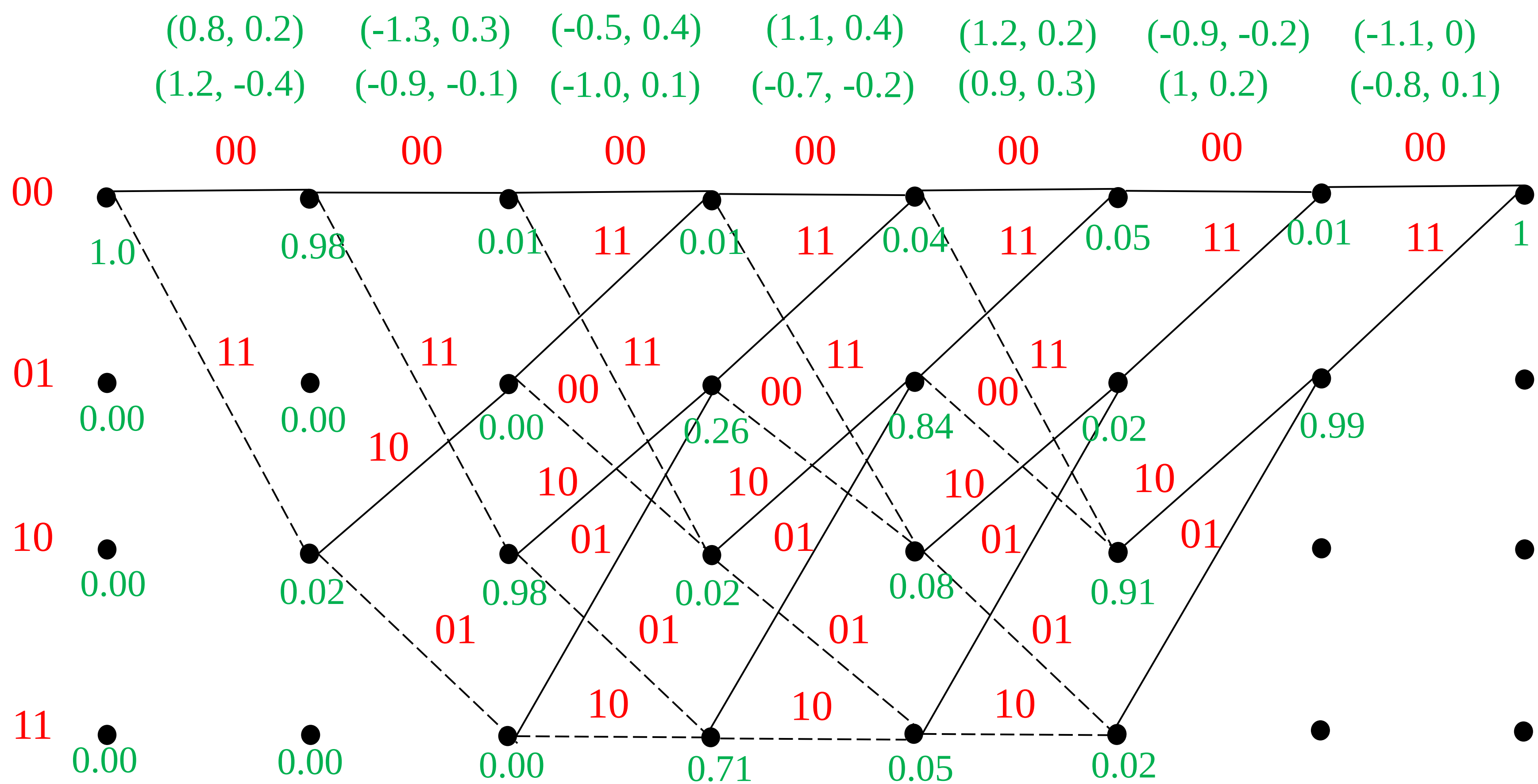
§ 5.5 BCJR Decoding

Step 1: Determine $\Gamma_{\Omega\rightarrow\Omega'}$ of all transitions.



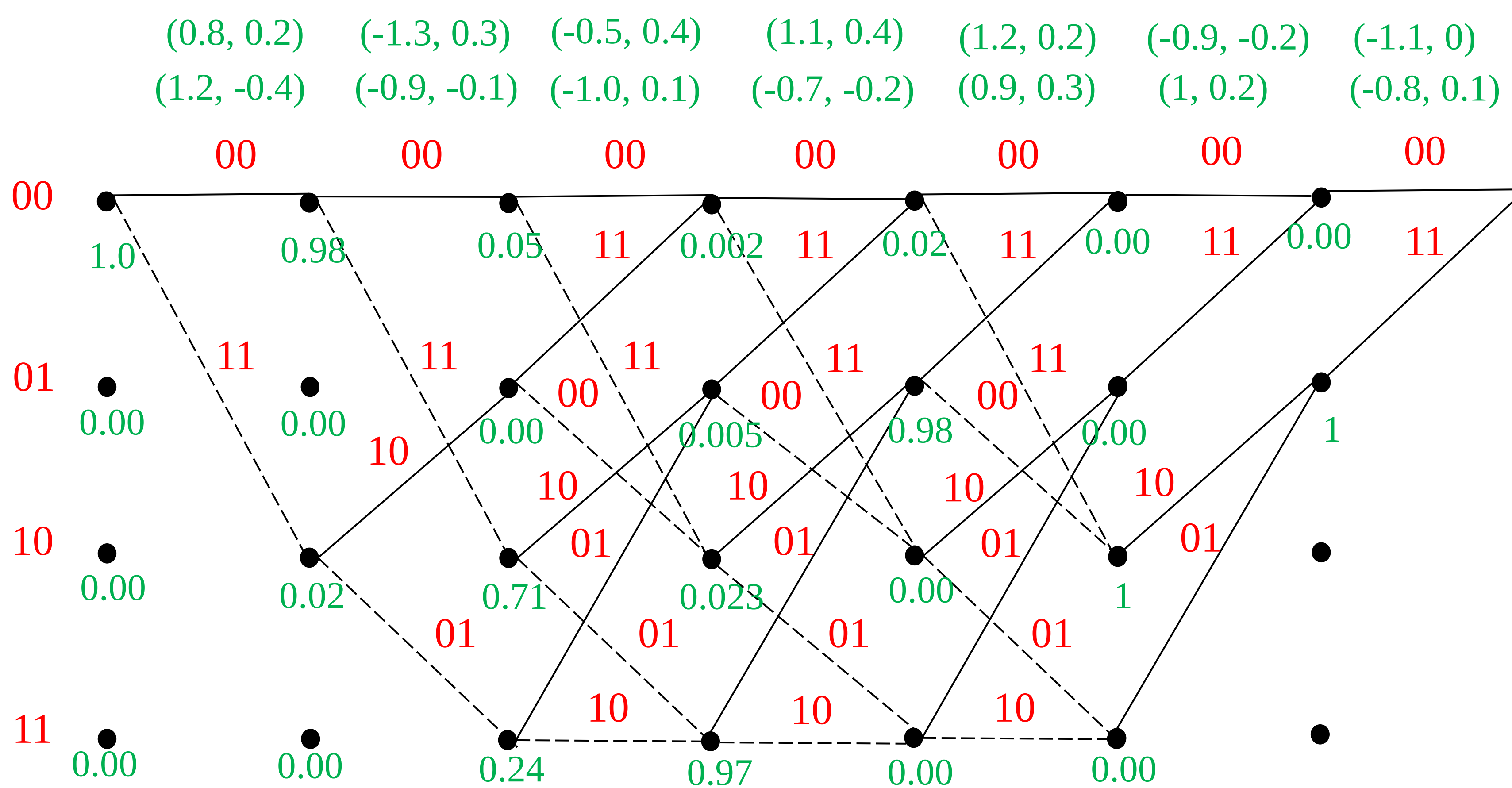
§ 5.5 BCJR Decoding

Step 2: Forward trace, determine $A_{t'}(\Omega)$ of values.



§ 5.5 BCJR Decoding

Step 3: Backward Trace, determine $B_{t'+1}(\Omega')$ of values.



§ 5.5 BCJR Decoding

Step 4: Determine the *a posteriori* probabilities of each information bit.

$$\begin{cases} P(u_1 = 0|y_1) = 1 \\ P(u_1 = 1|y_1) = 0 \end{cases} \Rightarrow \quad \hat{u}_1 = 0$$

$$\begin{cases} P(u_2 = 0|y_2) = 0 \\ P(u_2 = 1|y_2) = 1 \end{cases} \Rightarrow \quad \hat{u}_2 = 1$$

$$\begin{cases} P(u_3 = 0|y_3) = 0 \\ P(u_3 = 1|y_3) = 1 \end{cases} \Rightarrow \quad \hat{u}_3 = 1$$

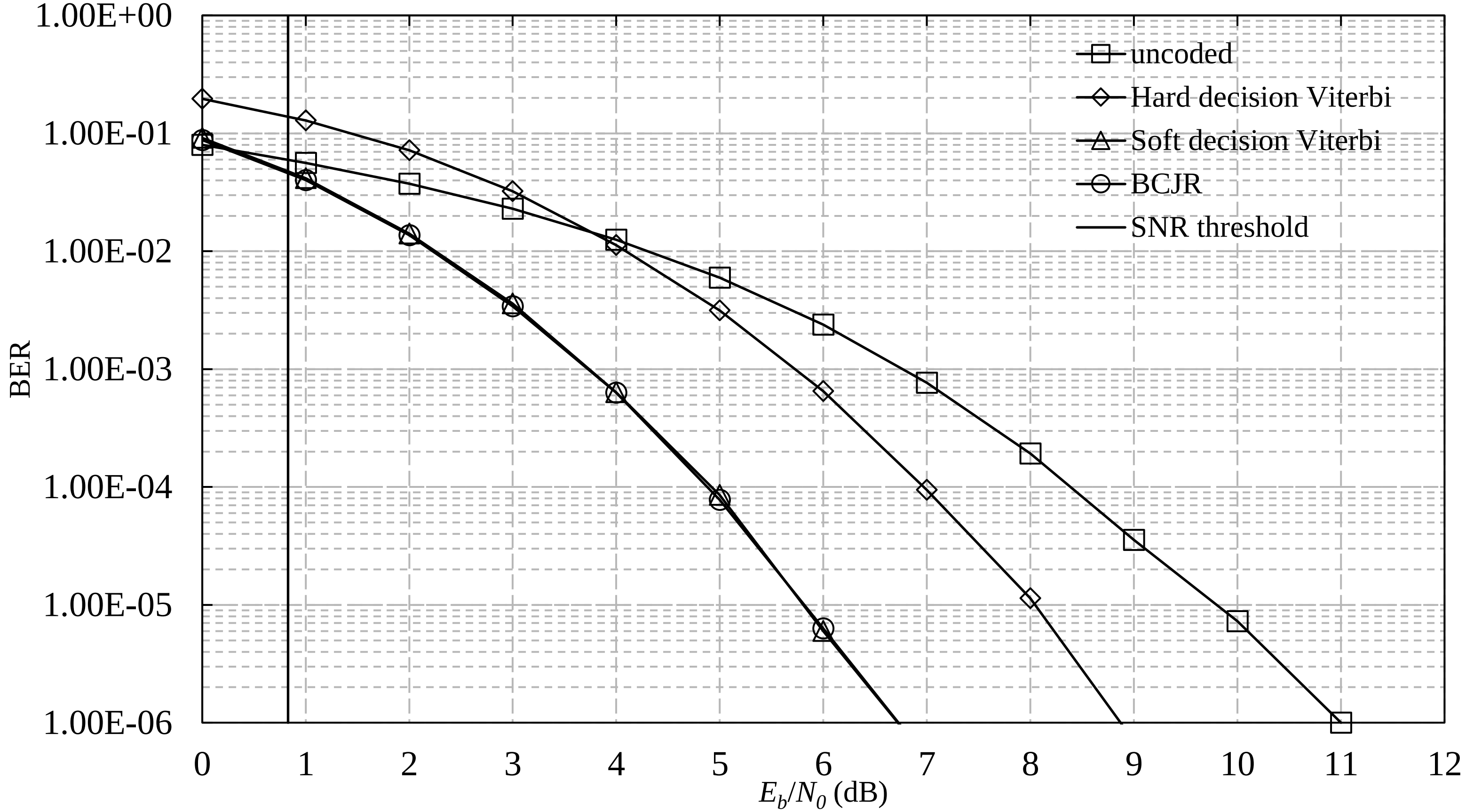
$$\begin{cases} P(u_4 = 0|y_4) = 1 \\ P(u_4 = 1|y_4) = 0 \end{cases} \Rightarrow \quad \hat{u}_4 = 0$$

$$\begin{cases} P(u_5 = 0|y_5) = 0 \\ P(u_5 = 1|y_5) = 1 \end{cases} \Rightarrow \quad \hat{u}_5 = 1$$

$$\hat{u}_6 = \hat{u}_7 = 0$$

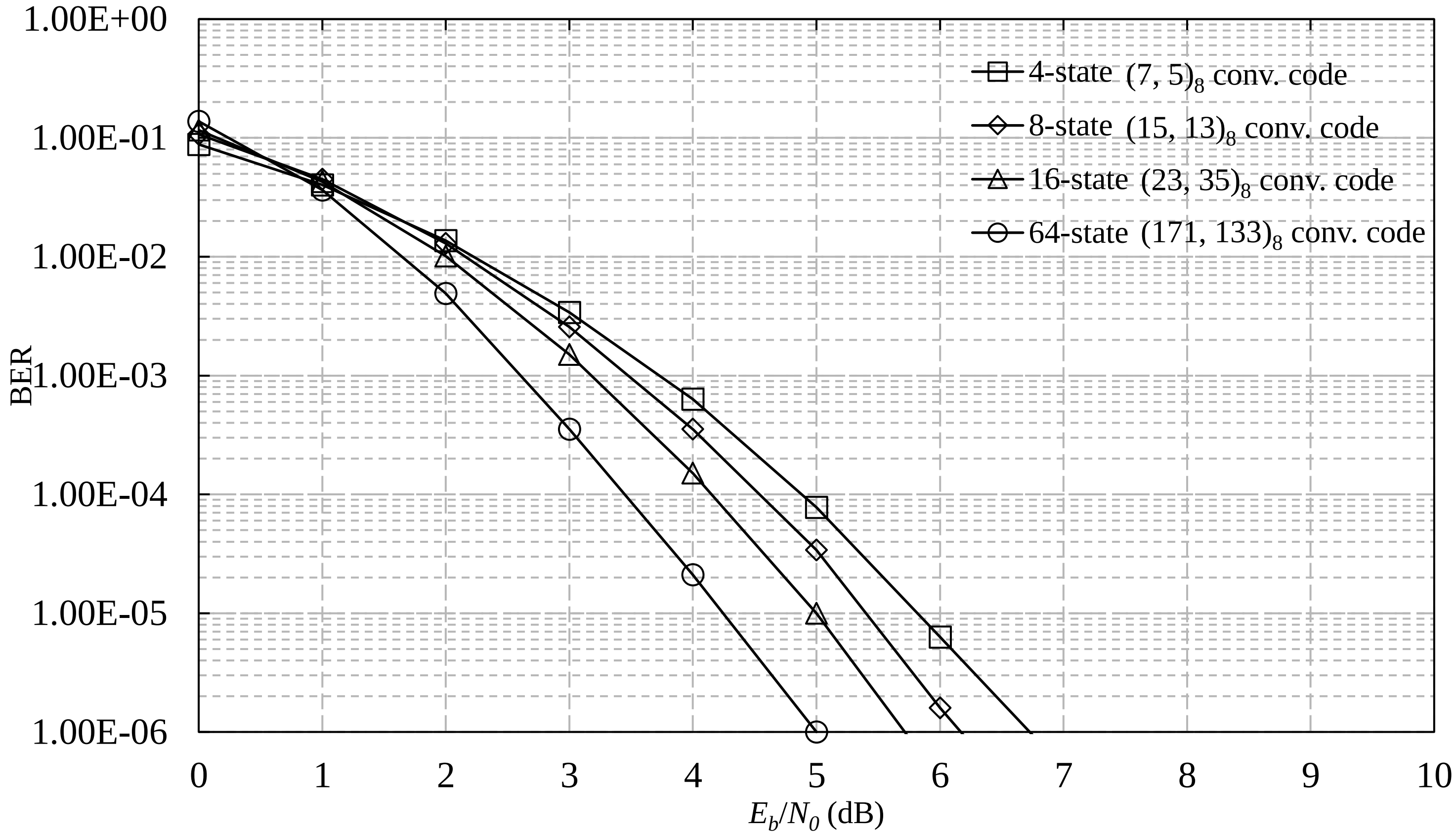
§ 5.5 BCJR Decoding

BER performance of $(7, 5)_8$ conv. code over AWGN channel using BPSK.



§ 5.5 BCJR Decoding

BER performance of different conv. code over AWGN channel using BPSK.



§ 5.6 Trellis Coded Modulation

- Convolutional code enables reliable communications. But as a channel code, its error-correction function is on the expense of spectral efficiency.
- Spectral efficiency $(\eta) = \dfrac{Nr.\,of\; information\; bits}{transmitted\; symbol}$
- E.g., an uncoded system using BPSK

  η = 1 info bits/symbol

  A rate 1/2 conv. coded system using BPSK

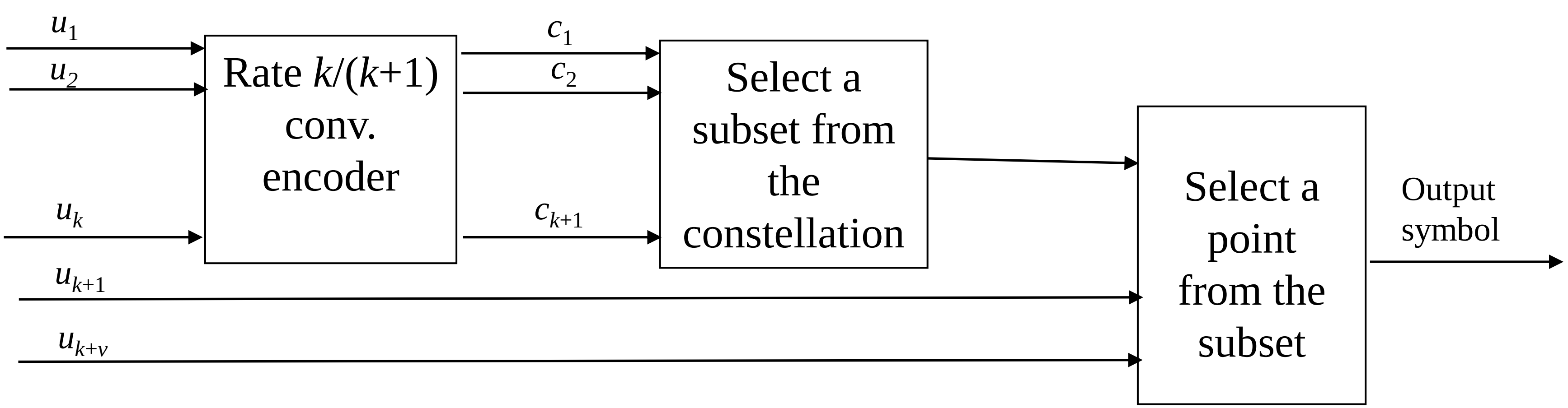
  η = 0.5 info bits/symbol
- Can we achieve reliable and yet spectrally efficient communication?

  Solution: Trellis Coded Modulation (TCM) that integrates a conv. code with a high order modulation [3].

[3] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 55-67, 1982.

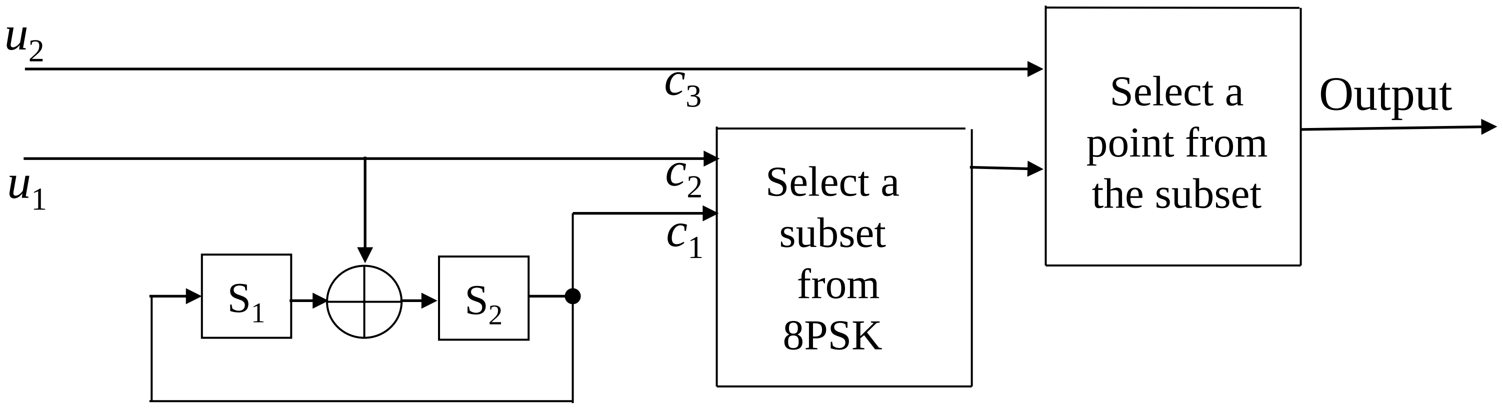
§ 5.6 Trellis Coded Modulation

- A general structure of the TCM scheme



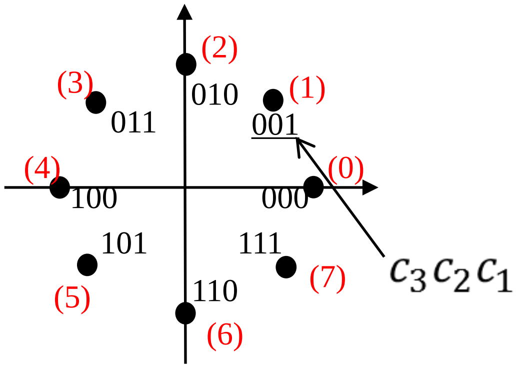
§ 5.6 Trellis Coded Modulation

- A rate 2/3 TCM code.



Rate ½ 4-state Convolutional Code

8PSK Constellation



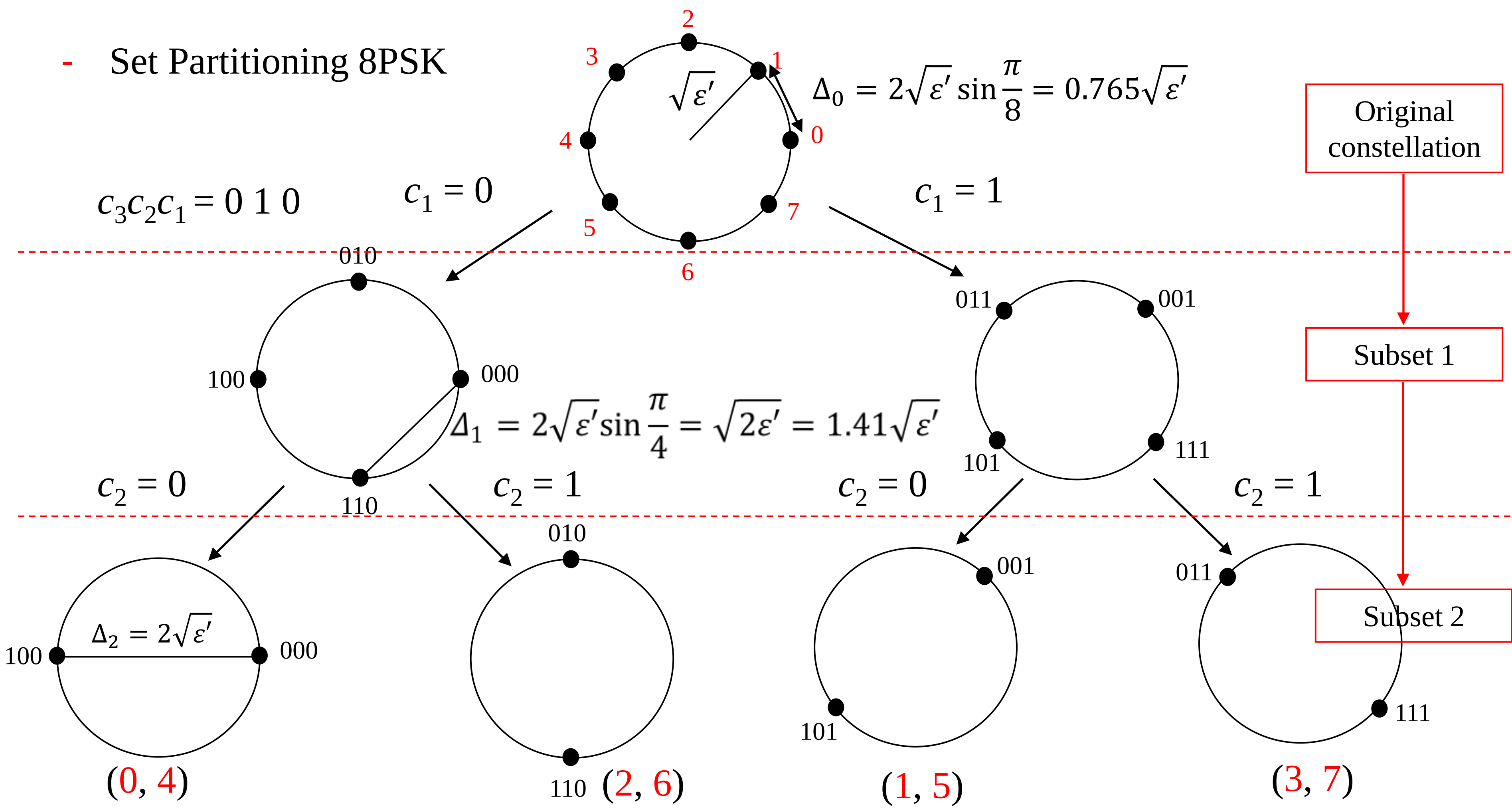
§ 5.6 Trellis Coded Modulation

- State table of the rate 2/3 TCM code

Input		Current State		Next State		Output			Symbol
u_1	u_2	S_1	S_2	S_1'	S_2'	c_1	c_2	c_3	**8PSK sym**
0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	1	0	1	0	2
0	1	0	0	0	0	0	0	1	4
1	1	0	0	0	1	0	1	1	6
0	0	0	1	1	0	1	0	0	1
1	0	0	1	1	1	1	1	0	3
0	1	0	1	1	0	1	0	1	5
1	1	0	1	1	1	1	1	1	7
0	0	0	0	0	1	0	0	0	0
1	0	1	0	0	0	0	1	0	2
0	1	1	0	0	1	0	0	1	4
1	1	1	0	0	0	0	1	1	6
0	0	1	1	1	1	1	0	0	1
1	0	1	1	1	0	1	1	0	3
0	1	1	1	1	1	1	0	1	5
1	1	1	1	1	0	1	1	1	7

§ 5.6 Trellis Coded Modulation

- Set Partitioning 8PSK



§ 5.6 Trellis Coded Modulation

- Set Partitioning 8PSK

By doing set partitioning, the minimum distance between point within a subset is increasing as: $\Delta_0 < \Delta_1 < \Delta_2$.

Set Partitioning

Original constellation

$$\Delta_0 = d(0,1) = 2\sqrt{\varepsilon'}\sin\frac{\pi}{8} = 0.765\sqrt{\varepsilon'}$$

Subset 1

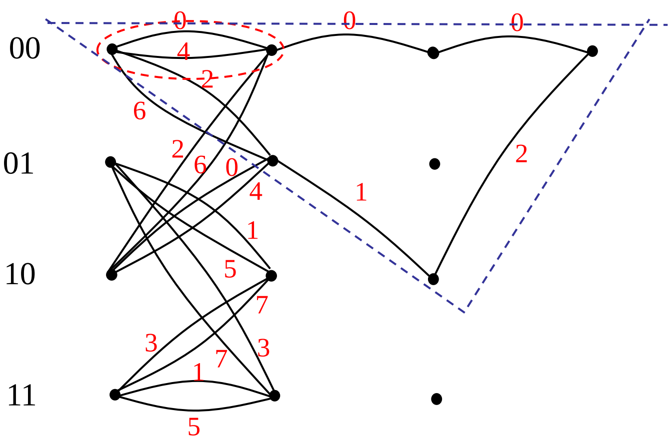
$$\Delta_1 = d(0,2) = \sqrt{2\varepsilon'} = 1.414\sqrt{\varepsilon'}$$

Subset 2

$$\Delta_2 = d(0,4) = 2\sqrt{\varepsilon'}$$

§ 5.6 Trellis Coded Modulation

- Viterbi trellis of the rate 2/3 TCM code



For diverse/remerge transition:

$$d_{\text{free}}^2 = [d^2(0,2) + d^2(0,1) + d^2(0,2)]$$
$$= 2\varepsilon' + \left(2 - \sqrt{2}\right)\varepsilon' + 2\varepsilon' = 4.568\varepsilon'$$

For parallel transition:

$$d_{\text{free}}^2 = d^2(0,4) = 4\varepsilon'$$

Choose the smaller one as the free distance of the code:

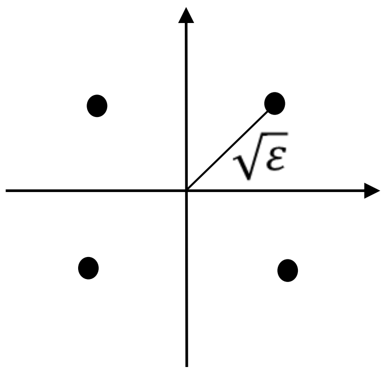
$$d_{\text{free}}^2 = 4\varepsilon'$$

- Bit $c_3 = 0$ and $c_3 = 1$ result in two parallel transition branches. By doing set partitioning, we are trying to maximize the Euclidean distance between the two parallel branches. So that the free distance of the TCM code can be maximized.

§ 5.6 Trellis Coded Modulation

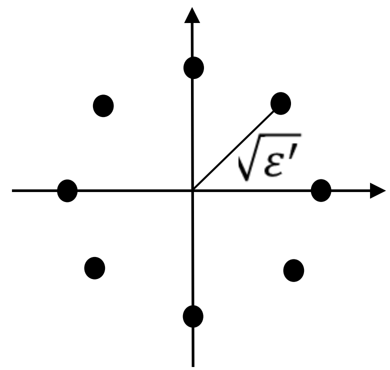
- Asymptotic coding gain over an uncoded system.
- Spectral efficiency (η) = 2 info bits/sym.

uncoded QPSK



$d_{min}^2 = 2\varepsilon$

rate 2/3 coded 8PSK



$d_{free}^2 = 4\varepsilon'$

Asymptotic coding gain $\gamma = \dfrac{\left(d_{free}^2/\varepsilon'\right)}{\left(d_{min}^2/\varepsilon\right)} = 2.$

Asymptotic coding gain in dB $= 10\log_{10}\gamma = 3$ dB.

- With the same transmission spectral efficiency of 2 info bits/sym, the TCM coded system achieves 3 dB coding gain over the uncoded system asymptotically.