Linear Gaussian Channels and Delayed Decision Feedback Sequence Estimation

Shane M. Haas

Laboratory for Information and Decision Systems

Massachusetts Institute of Technology

I – Motivation

II – The Equivalent Discrete-Time Channel Model

III – Maximum Likelihood Sequence Estimation (MLSE)

IV – Zero-Forcing Decision-Feedback Estimation (ZF-DFE)

V – Delayed Decision-Feedback Sequence Estimation (DDFSE)

VI – Performance

VII – Conclusions

I – Motivation

- **THE PROBLEM:** Detection in the presence of intersymbol interference

- **SOME SOLUTIONS:**

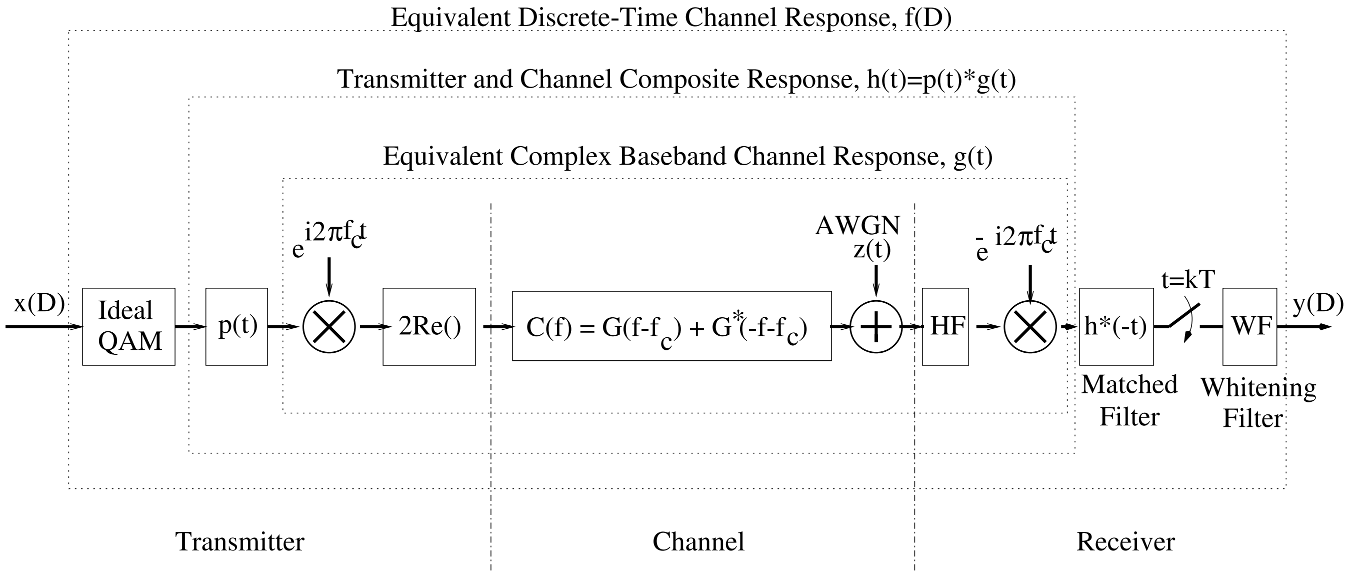
  MLSE – optimal detector

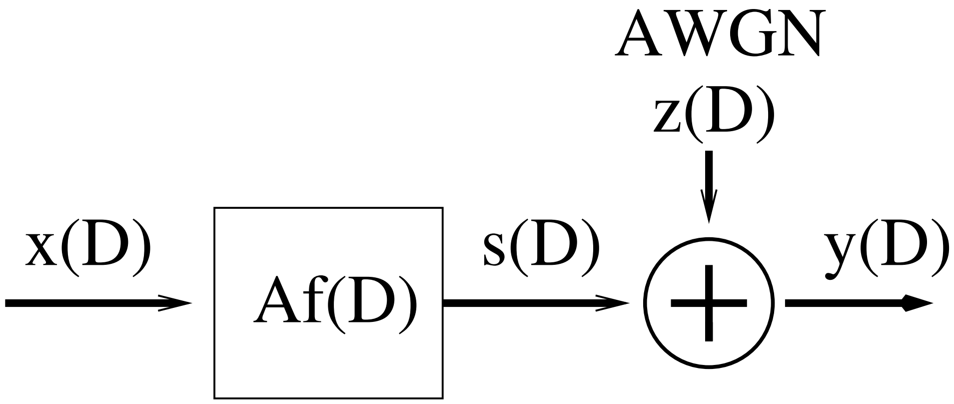
  ZF-DFE – suboptimal detector

  TH Precoding – ZF-DFE at transmitter

  DDFSE – hybrid of MLSE and ZF-DFE

II — The Equivalent Discrete-Time Channel Model





$$y(D) = Af(D)x(D) + z(D)$$

III – Maximum Likelihood Sequence Estimation (MLSE)

- The optimum detector is the minimum distance rule:

$$\begin{aligned}
\widehat{x}(D) &= \underset{x(D)\in\mathcal{X}[D]}{\text{argmin}} \|y(D) - Af(D)x(D)\|^2 \\
&= \underset{x(D)\in\mathcal{X}[D]}{\text{argmin}} \|\frac{1}{A}y(D) - \underbrace{[f(D)-1]x(D)}_{\text{ISI}} \\
&\quad -x(D)\|^2
\end{aligned}$$

- **EXAMPLE 1: FIR Channel**
  - Given:
    * Input Alphabet: $\mathcal{X} = \{-1, +1\}$
    * Channel: $f(D) = 1 - 1.5D + 0.5D^2$, $A = 1$
    * Output: $y(D) = 2.1 - 2.9D$
  - Find: Most likely transmitted sequence: $x(D) = x_0 + x_1 D$

Transmitted Sequence				State Seq.	Filter Output		Distance Metrics		
x_{-2}	x_{-1}	x_0	x_1	(X_0, X_1, X_2)	s_0	s_1	$\lvert s[0]-y[0]\rvert^2$	$\lvert s[1]-y[1]\rvert^2$	$\lvert s-y\rvert^2$
-1	-1	-1	-1	(1,1,1)	0	0	4.41	8.41	12.82
-1	-1	-1	1	(1,1,2)	0	2	4.41	24.01	28.42
-1	-1	1	-1	(1,2,3)	2	-3	0.01	0.01	0.02
-1	-1	1	1	(1,2,4)	2	-1	0.01	3.61	3.62
-1	1	-1	-1	(2,3,1)	-3	1	26.01	15.21	41.22
-1	1	-1	1	(2,3,2)	-3	3	26.01	34.81	60.82
-1	1	1	-1	(2,4,3)	-1	-2	9.61	0.81	10.42
-1	1	1	1	(2,4,4)	-1	0	9.61	8.41	18.02
1	-1	-1	-1	(3,1,1)	1	0	1.21	8.41	9.62
1	-1	-1	1	(3,1,2)	1	2	1.21	24.01	25.22
1	-1	1	-1	(3,2,3)	3	-3	0.81	0.01	0.82
1	-1	1	1	(3,2,4)	3	-1	0.81	3.61	4.42
1	1	-1	-1	(4,3,1)	-2	1	16.81	15.21	32.02
1	1	-1	1	(4,3,2)	-2	3	16.81	34.81	51.62
1	1	1	-1	(4,4,3)	0	-2	4.41	0.81	5.22
1	1	1	1	(4,4,4)	0	0	4.41	8.41	12.82

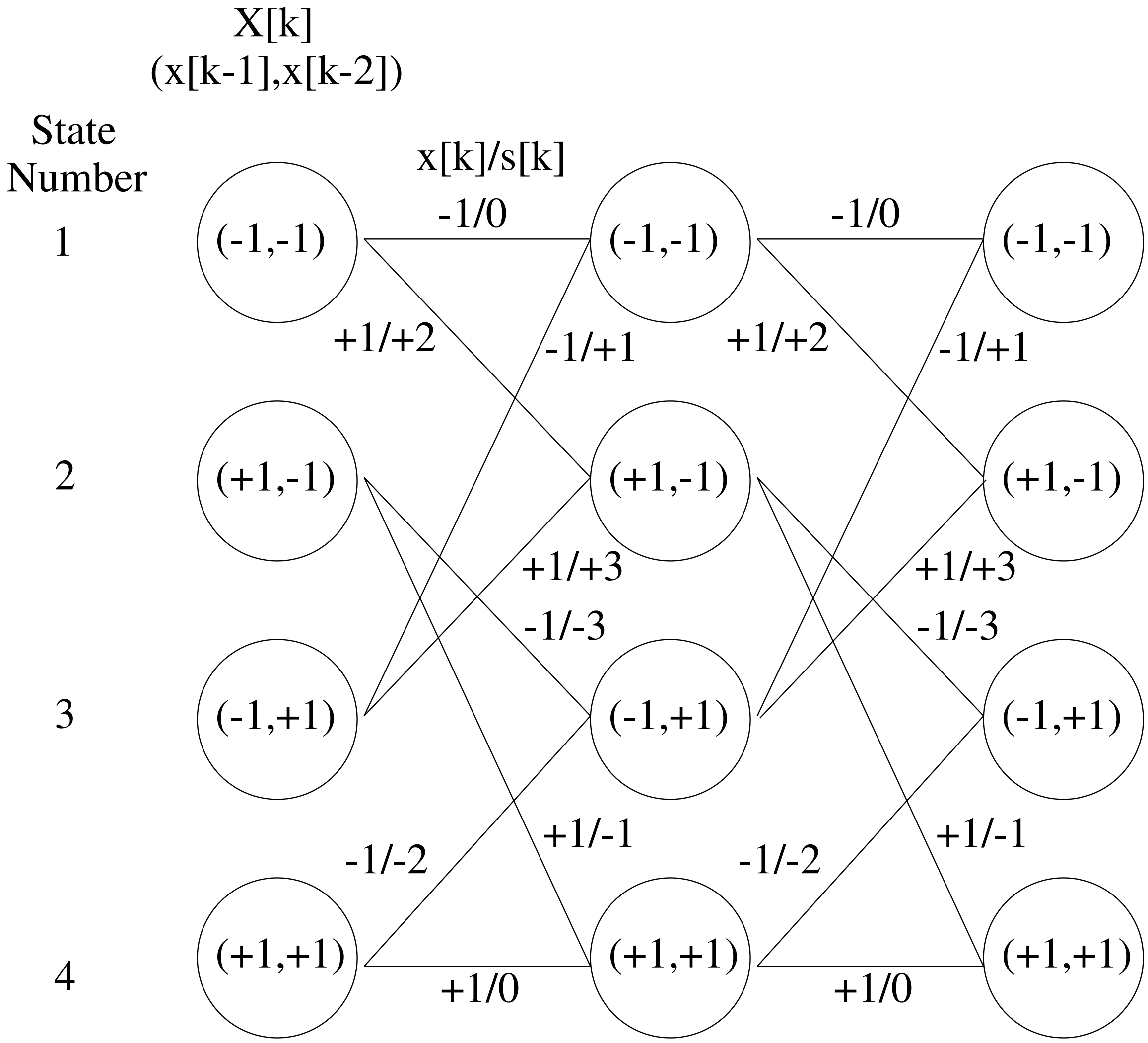
- The Viterbi Algorithm
  - State: $\mathbf{X_k} = (x_{k-1}, \ldots, x_{k-\eta})$; $\eta = deg\{f(D)\}$
  - Branch metric for $\mathbf{X_k}$ and $\mathbf{X_{k+1}}$

$$
\begin{aligned}
& |y_k - \sum_{i=0}^{\eta} A f_i x_{k-i}|^2 \\
& \quad = \; |y_k - \sum_{i=1}^{\eta} A f_i x_{k-i} - A x_k|^2 \\
& \quad \propto \; |\frac{1}{A} y_k - \underbrace{\sum_{i=1}^{\eta} f_i x_{k-i}}_{\text{ISI}} - x_k|^2
\end{aligned}
$$

  - For each next state, choose smallest path metric

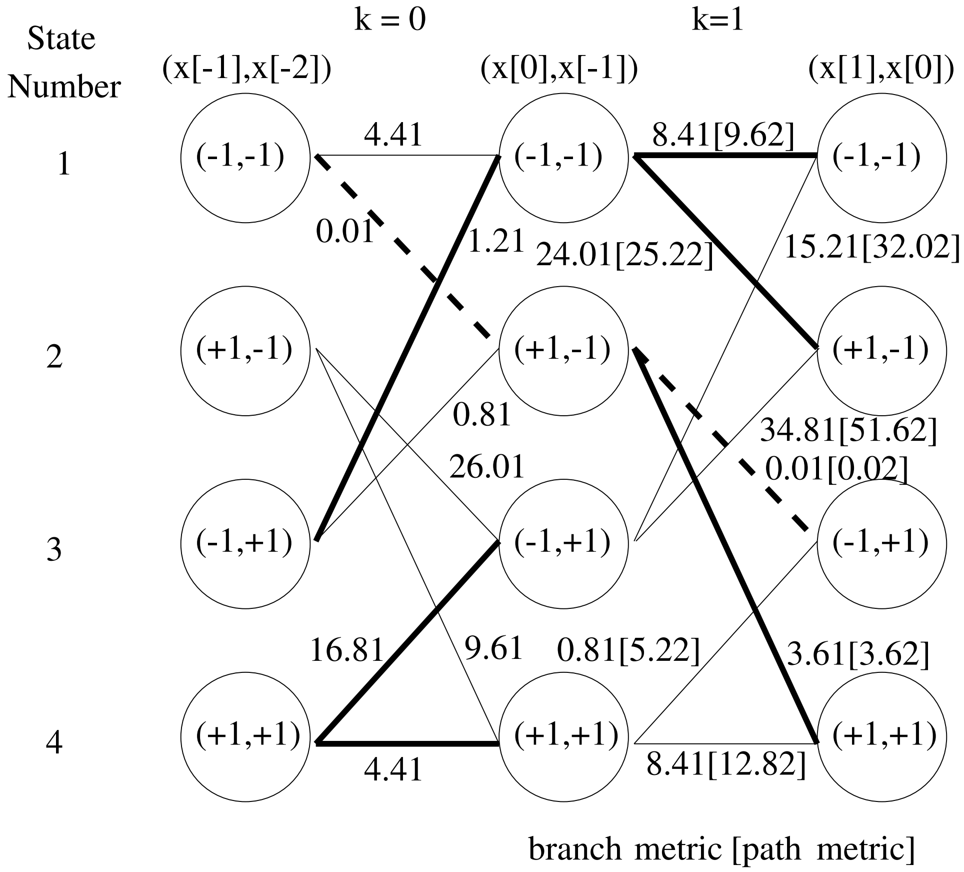
- **EXAMPLE (cont.): FIR Channel**

  * $f(D) = 1 - 1.5D + 0.5D^2$, $A = 1$,



- **EXAMPLE (cont.): FIR Channel**

  * $y(D) = 2.1 - 2.9D$

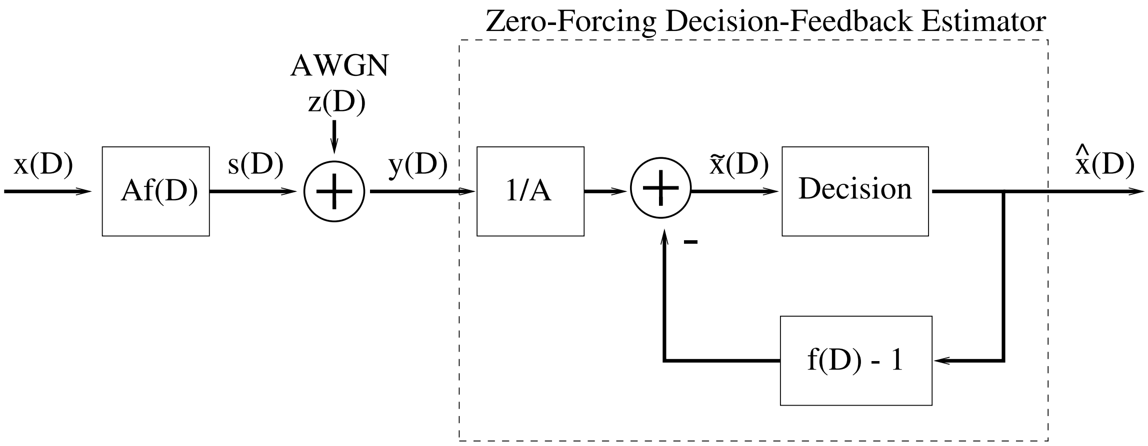


- Advantages of MLSE:
  - Optimal in terms of error probability
  - VA provides recursive implementation
- Disadvantages of MLSE:
  - Complexity is exponential, i.e. $|\mathcal{X}|^\eta$
  - Does not exist for IIR channels

IV – Zero-Forcing Decision-Feedback Estimation (ZF-DFE)

- Decision Rule:

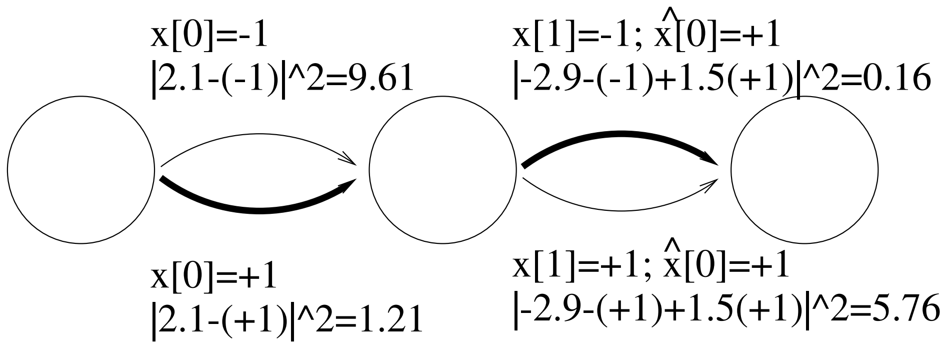
$$\hat{x}_k = \underset{x \in \mathcal{X}}{\text{argmin}} \left| \frac{1}{A} y_k - \underbrace{\sum_{i=1}^{\eta} f_i \hat{x}_{k-i}}_{\text{DFE of ISI}} - x \right|^2$$



- **EXAMPLE (cont.): FIR Channel**

  * $f(D) = 1 - 1.5D + 0.5D^2$, $A = 1$,

  * $y(D) = 2.1 - 2.9D$



- Advantages of ZF-DFE:
  - Very low complexity
- Disadvantages of ZF-DFE:
  - Non-optimal probability of error
  - Not compatible with channel coding

- **Tomlinson-Harashima (TH) precoding** implements DFE at transmitter

- Advantages of TH Precoding:
  - Very low complexity
  - Allows channel coding

- Disadvantages of TH Precoding:
  - Non-optimal probability of error (same as ZF-DFE)
  - Not compatible with general shaped constellations ($M \times M$ QAM only)

V – Delayed Decision-Feedback Sequence Estimation (DDFSE)

- Hybrid of MLSE and ZF-DFE, with each as special cases
- Truncates channel response
- Uses VA on reduced state trellis
- Uses ZF-DFE on each branch

- State: $\mathbf{X_k} = (x_{k-1}, \ldots, x_{k-\mu})$, $0 \leq \mu \leq \eta = deg\{f(D)\}$

- Branch Metric:

$$|y_k - \sum_{i=0}^{\mu} A f_i x_{k-i} - \sum_{i=\mu+1}^{\eta} A f_i \widehat{x}_{k-i}|^2$$

$$\propto \; |\frac{1}{A} y_k - \underbrace{\underbrace{\sum_{i=1}^{\mu} f_i x_{k-i}}_{\text{from State}} - \underbrace{\sum_{i=\mu+1}^{\eta} f_i \widehat{x}_{k-i}}_{\text{from DFE}}}_{\text{ISI}} - x_k|^2$$

- **For FIR Channels**: Requires storage of $\eta - \mu$ previous decisions from minimum distance path to $\mathbf{X_k}$, i.e. $\{\widehat{x}_{k-\mu-1}, \ldots, \widehat{x}_{k-\eta}\}$

- **For IIR Channels**: We also want the finite storage of past decisions...

- Assume $f(D) = \beta(D)/\gamma(D)$. Write

$$f(D) = f_\mu(D) + D^{\mu+1} f^+(D)$$

where

$$\begin{aligned}
f_\mu(D) &= 1 + f_1 D + \cdots f_\mu D^\mu \\
f^+(D) &= f_{\mu+1} + f_{\mu+2} D + \cdots f_\eta D^\eta \\
&= \frac{(\beta(D) - f_\mu(D)\gamma(D))D^{-(\mu+1)}}{\gamma(D)} \\
&= \frac{\beta^+(D)}{\gamma(D)} \\
\beta^+(D) &= (\beta(D) - f_\mu(D)\gamma(D))D^{-(\mu+1)}
\end{aligned}$$

- Define the IIR DFE contribution to ISI as

$$
\begin{aligned}
w(D) &= f^{+}(D)x(D) \\
&= \frac{\beta^{+}(D)}{\gamma(D)}x(D)
\end{aligned}
$$

So,

$$
w_k = \begin{cases} \sum_{i=0}^{n} \beta_i^{+} x_{k-i} - \sum_{i=1}^{m} \gamma_i w_{k-i} & (m > 0) \\ \sum_{i=0}^{n} f_{i+\mu+1} x_{k-i} & (m = 0) \end{cases}
$$

where $n = deg\{\beta^{+}(D)\}$ and $m = deg\{\gamma(D)\}$

- Rewrite the branch metric as:

$$|y_k - \sum_{i=0}^{\mu} A f_i x_{k-i} - A\widehat{w}_{k-\mu-1}|^2$$

where

$$\begin{aligned}
&\widehat{w}_{k-\mu-1} \\
&= \sum_{i=\mu+1}^{\eta} f_i \widehat{x}_{k-i} \\
&= \begin{cases} \sum_{i=0}^{n} \beta_i^{+} \widehat{x}_{k-\mu-1-i} \\ \quad - \sum_{i=1}^{m} \gamma_i \widehat{w}_{k-\mu-1-i} & (m > 0) \\ \sum_{i=0}^{\eta-\mu-1} f_i \widehat{x}_{k-\mu-1-i} & (m = 0) \end{cases}
\end{aligned}$$

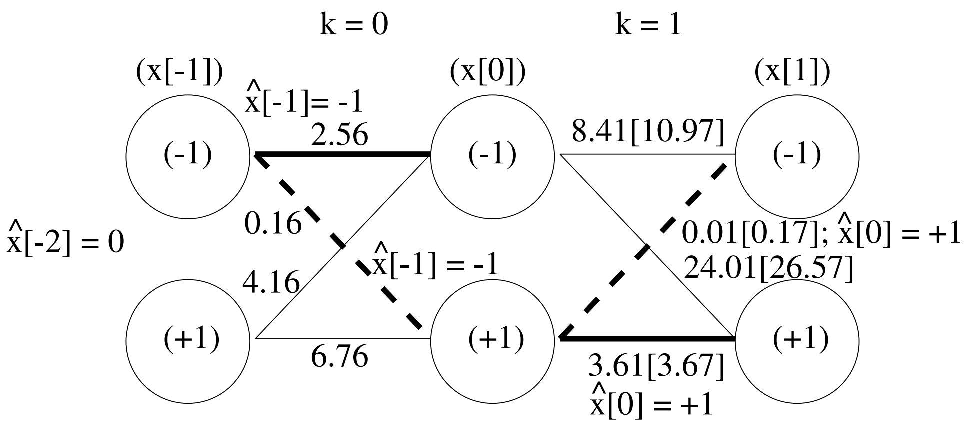
- **For IIR Channel:** Requires
  - Current state: $\mathbf{X_k} = (x_{k-1}, x_{k-2}, \cdots, x_{k-\mu})$,
  - Previous n decisions: $\{\hat{x}_{k-\mu-1}, \ldots, \hat{x}_{k-\mu-n}\}$
  - Previous m estimates of DFE IIR part of ISI: $\{\hat{w}_{k-\mu-2}, \ldots, \hat{w}_{k-\mu-m-1}\}$

- **EXAMPLE (cont.): FIR Channel**

  * $f(D) = 1 - 1.5D + 0.5D^2,\ A = 1$

  * $y(D) = 2.1 - 2.9D$

  * $\mu = 1$



- Advantages of DDFSE:
  - Trades performance for complexity
  - Exists for IIR channels
  - Compatible with trellis coding
- Disadvantages of DDFSE:
  - Suboptimal in terms of error probability

VI – Performance

- In general, difficult to evaluate for MLSE and DDFSE

- Based on the notion of error events

- Performance of DDFSE is between MLSE and ZF-DFE

- Term in upper bound (see paper):

$$Q\left(\frac{1}{2}\frac{|d^{\mu}_{min}|}{N_0}\right)$$

$$\sum_{\lambda \in \Lambda_{d^{\mu}_{min}}} w(\lambda) \prod_{i=0}^{n-1} \frac{m - |e_i|}{m}$$

where

  - $|d^{\mu}_{min}|$ is the minimum distance achieved by an error event
  - $\Lambda_{d^{\mu}_{min}}$ is the set of error events which achieve the minimum distance
  - $w(\lambda)$ number of symbol errors entailed by error event λ

– n is the duration of error event $w(\lambda)$

– m is the number of points in the PAM constellation $\mathcal{X}$

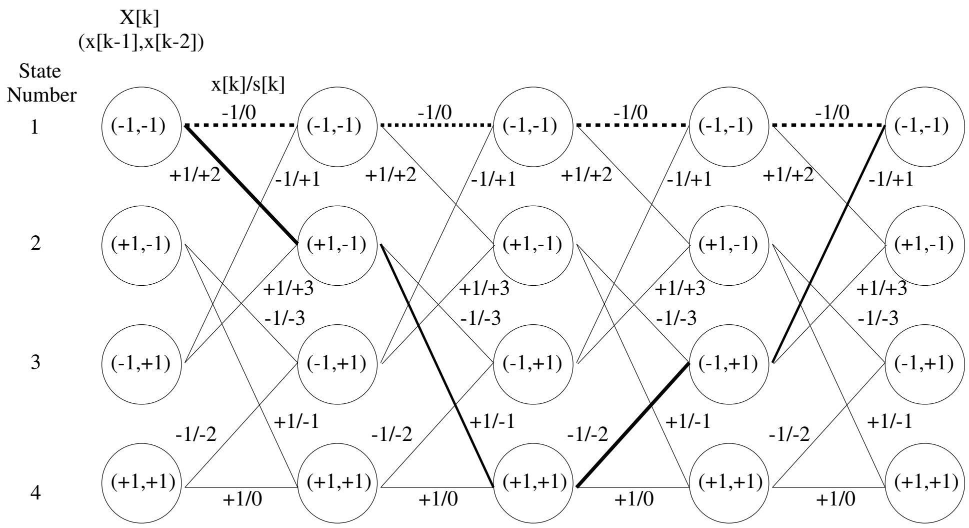
– $e_i = x_i - \widehat{x}_i$ is the error sequence

- Error probability in the absence of preceding decision errors

- **EXAMPLE (cont.): FIR Channel**

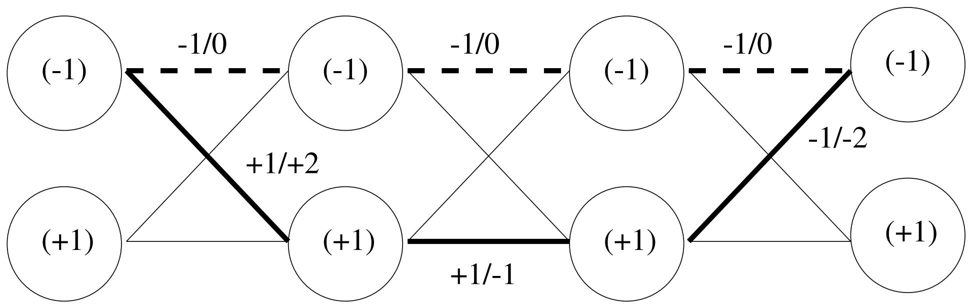
  * $f(D) = 1 - 1.5D + 0.5D^2$, $A = 1$

  * An error event in the VA with $|d|^2 = 10$



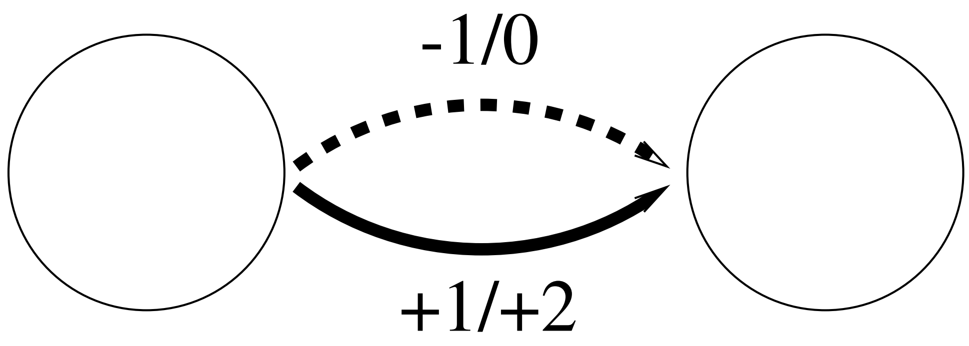
- **EXAMPLE (cont.): FIR Channel**

  * The same error event in the DDFSE with $|d|^2 = 9$



- **EXAMPLE (cont.): FIR Channel**

  * The same error event in the ZF-DFE with $|d|^2 = 4$



- **EXAMPLE (cont.): FIR Channel**
  - Plot of Error Performance (from Duel-Hallen, 1989):

VII – Conclusions

- Equivalent DT Channel Model
- MLSE – optimal detector (high complexity)
- ZF-DFE – suboptimal detector (low complexity)
- DDFSE – hybrid of MLSE and ZF-DFE