# Linear Gaussian Channels and Delayed Decision-Feedback Sequence Estimation

Shane M. Haas

November 29, 2000

## 1 Short Attention Span Summary

The output of a linear Gaussian channel is its input convolved with the channel's impulse response, plus additive white Gaussian noise. The detection or estimation of the channel's input based on its output is the focus of this presentation.

We will examine the following three methods to perform this detection:

**Maximum-Likelihood Sequence Estimation (MLSE)** is the optimal minimum probability of error detector [3]. If the channel has a finite impulse response, then the Viterbi algorithm (VA) provides a recursive implementation. Its complexity, however, increases exponentially with the length of channel response and the input alphabet size, making it infeasible in practice.

**Zero-Forcing Decision-Feedback Estimation (ZF-DFE)** is a suboptimal detector that subtracts from the received signal an estimate of the intersymbol interference based on previous decisions [4]. In general, its error probability is worse than that of a MLSE, but its implementation is far less complex.

**Delayed Decision-Feedback Sequence Estimation (DDFSE)** is a hybrid of MLSE and ZF-DFE, with each as special cases [1]. It truncates the channel impulse response and decodes with the VA, but uses a ZF-DFE on each branch of the trellis to remove intersymbol interference. Its complexity and performance is between that of a ZF-DFE and a MLSE.

## 2 Introduction

Maximum likelihood sequence estimation (MLSE) is the optimal minimum probability of error detector on intersymbol interference channels. Unfortunately, its complexity grows exponentially with the input alphabet size and channel memory, making it infeasible in practice.

When the channel has finite memory, the recursive Viterbi algorithm provides an efficient implementation. For infinite impulse response channels, however, the MLSE does not exist. For these channels, truncating the impulse response provides a suboptimal solution, and might not be sufficient in severe intersymbol interference. The purpose of this presentation is to examine other suboptimal detection methods that trade performance for reduced complexity.

The zero-forcing decision-feedback estimator (ZF-DFE) uses previous symbol decisions to estimate the intersymbol interference and subtracts it from the received sequence before detection. Because it requires symbol-by-symbol decisions, however, the ZF-DFE is not compatible with channel coding, which makes decisions on a block basis.

To overcome this limitation, transmitter pre-coding can effectively implement the feedback portion of a DFE at the transmitter [4]. Tomlinson and Harashima (TH) in 1970 proposed a pre-coding technique for use with $M$-PAM or $M \times M$-QAM signal constellations. TH pre-coding performs approximately the same as a ZF-DFE, but allows for channel coding. Unfortunately, it cannot be used with more general signal constellations, eliminating any possible shaping gains. Another drawback of TH pre-coding is that the transmitter must also have knowledge of the channel impulse response, which is usually estimated at the receiver.

Reduced-complexity MLSE algorithms such as delayed decision-feedback sequence estimation (DDFSE) [1] or reduced-state sequence estimation (RSSE) [2] are hybrids of MLSE and ZF-DFE, with each as special cases. They trade performance for reduced complexity by using a Viterbi algorithm on a smaller state space. We will examine the DDFSE in more detail in Section 6.

# 3 The Equivalent Discrete-Time Channel Model

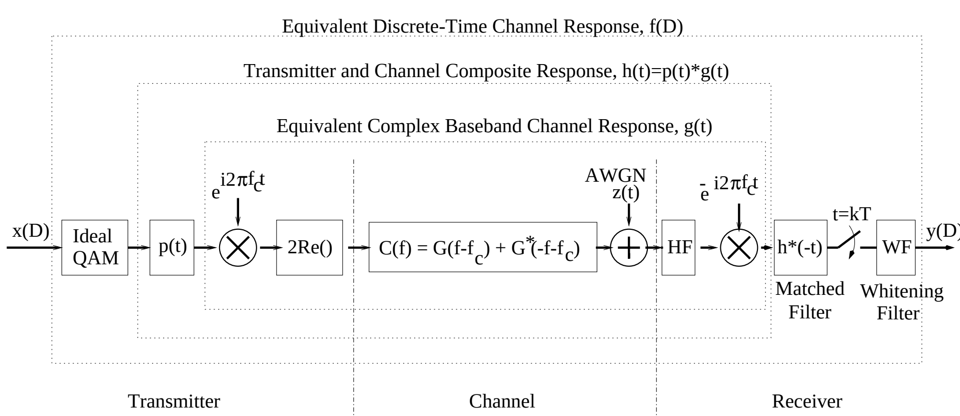


Figure 1: The linear Gaussian channel model

Figure 1 shows a model of the linear Gaussian channel. The sequence $x(D) = x_0 + x_1 D + \cdots + x_{N-1} D^{N-1}$, $x_i \in \mathcal{X}$, of symbols from a quadrature amplitude modulation (QAM) constellation $\mathcal{X}$ modulates a train of transmit filter pulses creating the baseband signal $\sum_{n=0}^{N-1} x_n p(t - nT)$. This signal is quadrature modulated to the carrier frequency $f_c$, and its real component transmitted over the channel.

The channel filters it and adds Gaussian noise that is spectrally white over the passband. Conceptually, the resulting signal passes through a Hilbert filter (HF) that only passes its positive frequency components. The Hilbert filter allows for an equivalent complex baseband representation $g(t)$ of the channel. In practice, however, this filter is unnecessary because the baseband matched filter will also remove the negative double frequency components after quadrature demodulation.

After demodulation, the receiver match filters the baseband signal by the composite response of the equivalent complex baseband channel $g(t)$ and the transmit filter $p(t)$. Sampling at integer multiples of $T$ produces the discrete-time baseband signal

$$y'(D) = R_{hh}(D)x(D) + z'(D)$$

where $R_{hh}(D)$ is the D-transform of the sampled autocorrelation function of $h(t)$, $z'(D)$ is a *non-white* stationary Gaussian sequence with autocorrelation sequence $R_{z'z'}(D) = S_z R_{hh}(D)$, and $S_z$ is the power spectral density of the AWGN process $z(t)$.

If $R_{hh}(D) = 1$, then the composite response is ideal Nyquist, and life is beautiful. The detection problem then becomes classic discrete-time detection in AWGN.

In practice, however, the channel is beyond our control, and we are not so lucky. If the aliased power spectrum of $h(t)$, i.e.

$$S_{hh}(f) = \sum_{k=-\infty}^{\infty} |H(f - k/T)|^2$$

is log-integrable over one period, then there exists a unique causal, monic, minimum-phase polynomial $f(D)$, such that

$$R_{hh}(D) = A^2 f(D) f^*(D^{-1})$$

where

$$\log A^2 = \int_0^{1/T} \log S_{hh}(f) df$$

Filtering $y'(D)$ by the anti-causal whitening filter $1/Af^*(D^{-1})$ yields the equivalent whitened matched-filtered discrete-time model

$$y(D) = Af(D)x(D) + z(D) \tag{1}$$

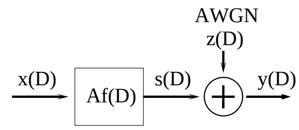


Figure 2: The equivalent whitened matched-filtered discrete-time linear Gaussian channel model, where $x(D)$ is the transmitted sequence; $f(D)$ is the monic, causal, minimum phase equivalent channel response; $z(D)$ is an iid Gaussian noise sequence with variance $S_z$; and $y(D)$ is the output sequence.

where $z(D)$ is a sequence of zero-mean independent identically distributed Gaussian random variables with variance $S_z$. This equivalent discrete-time channel model is shown in Figure 2.

For a more thorough discussion of the linear Gaussian channel see [4].

# 4 Maximum Likelihood Sequence Estimation (MLSE)

The minimum probability of error receiver uses the maximum likelihood (ML) or minimum distance rule:

$$\hat{x}(D) = \operatorname*{argmin}_{x(D)\in\mathcal{X}[D]} \|y(D) - Af(D)x(D)\|^2 \tag{2}$$

$$= \operatorname*{argmin}_{x(D)\in\mathcal{X}[D]} \|\frac{1}{A}y(D) - \underbrace{[f(D)-1]x(D)}_{\text{ISI}} - x(D)\|^2 \tag{3}$$

where $\|a(D)\|^2 = a_0^2 + a_1^2 + a_2^2 + \cdots$ and $\mathcal{X}[D] = \{x_0 + x_1 D + \cdots + x_{N-1}D^{N-1} | x_k \in \mathcal{X}\}$. In other words, the ML receiver chooses the sequence $x(D)$ whose response $Af(D)x(D)$ is closest to the received sequence $y(D)$. Notice that $[f(D)-1]x(D)$ is the intersymbol interference (ISI) that corrupts $x(D)$.

**Example 1:** ***FIR Channel*** Consider the finite impulse response (FIR) channel

$$f(D) = 1 - 1.5D + 0.5D^2 \tag{4}$$

with $A = 1$. Suppose that the input alphabet is $\mathcal{X} = \{-1, +1\}$. Find the most likely input sequence $x(D) = x_0 + x_1 D$ given the observed output sequence $y(D) = 2.1 - 2.9D$. The brute-force detection solution is shown in Table 1. For this received sequence, the most likely transmitted sequence is $x(D) = 1 - D$.

This brute-force solution is only feasible when the channel has a finite length, $\eta$, and requires approximately $|\mathcal{X}|^{N+\eta}$ comparisons. A more efficient implementation of this search is the recursive Viterbi algorithm (VA).

The VA operates on a trellis representation of the channel. Let $\mathbf{X_k} = (x_{k-1}, \ldots, x_{k-\eta})$ denote the state of the channel at time $k$. The number of states is exponential with alphabet size and channel length, i.e. $|\mathcal{X}|^\eta$. The VA

| Transmitted Sequence | | | | | State Sequence | | Filter Output | | | Distance Metrics | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $x_{-2}$ | $x_{-1}$ | $x_0$ | $x_1$ | | $(X_0, X_1, X_2)$ | | $s_0$ | $s_1$ | | $\|s_0 - y_0\|^2$ | $\|s_1 - y_1\|^2$ | $\|\mathbf{s} - \mathbf{y}\|^2$ |
| -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 |

Table 1: Maximum likelihood detection of $x(D)$ based on the observed output $y(D) = 2.1 - 2.9D$ of the FIR channel $f(D) = 1 - 1.5D + 0.5D^2$.

finds the state sequence through the trellis that minimizes the distance between its response and the received sequence.

**Example 1 (cont.):** ***FIR Channel*** Figure 3 shows the trellis representation of the FIR channel in (4). The branches of the trellis are labeled with their corresponding inputs and outputs.

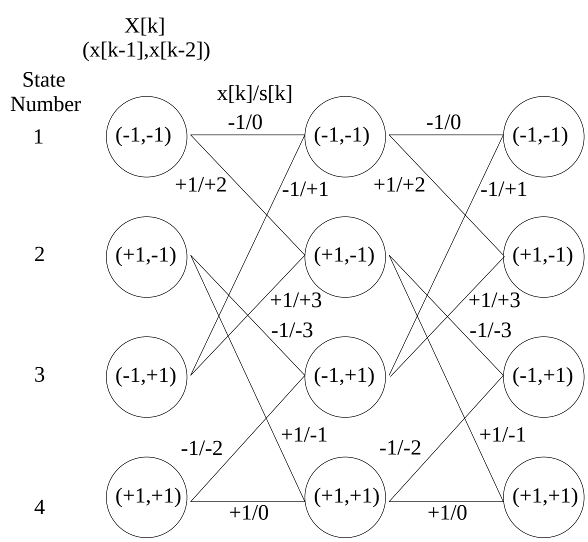


Figure 3: Trellis representation of the FIR channel $f(D) = 1 - 1.5D + 0.5D^2$.

The VA recursively finds the shortest path through the trellis in the following manner. For each state and next state pair ($\mathbf{X_k}$ and $\mathbf{X_{k+1}}$) it calculates the branch metric

$$|y_k - \sum_{i=0}^{\eta} A f_i x_{k-i}|^2 = |y_k - \sum_{i=1}^{\eta} A f_i x_{k-i} - A x_k|^2 \quad (5)$$

$$\propto |\frac{1}{A} y_k - \underbrace{\sum_{i=1}^{\eta} f_i x_{k-i}}_{\text{ISI}} - x_k|^2 \quad (6)$$

and the cumulative path metric, the sum of all branch metrics associated with the path through the trellis to state $\mathbf{X_k}$. Then, for each next state $\mathbf{X_{k+1}}$ it chooses the branch leading to it with the smallest path metric. These "surviving" branches form a minimum distance path to the state $\mathbf{X_{k+1}}$.

**Example 1 (cont.):** ***FIR Channel*** Figure 4 shows the branch and path metrics calculated by the Viterbi algorithm for the received sequence $y(D) = 2.1 - 2.9D$. Labeling each branch are its individual metric and its minimum distance path metric in brackets. The surviving branches are in bold, and minimum distance path at time $k = 1$ is dotted. This minimum distance path shows that the most likely transmitted sequence is $x(D) = 1 - D$.

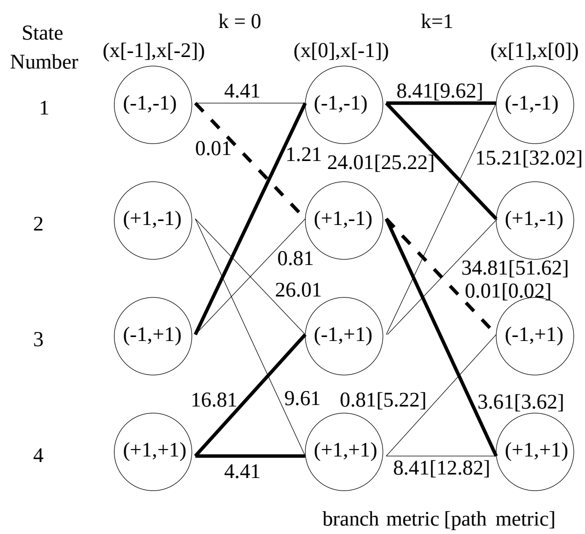


Figure 4: Branch and path metrics for the VA MLSE of x(D) on the FIR channel $f(D) = 1 - 1.5D + 0.5D^2$ based on the received sequence $y(D) = 2.1 - 2.9D$.

# 5 Zero-Forcing Decision Feedback Estimation (ZF-DFE)

The zero-forcing decision-feedback estimator (ZF-DFE) uses previous symbol decisions to subtract an estimate of the intersymbol interference from the received sequence. A block diagram of the ZF-DFE is shown in Figure 5. The decision rule is

$$\hat{x}_k = \underset{x \in \mathcal{X}}{\operatorname{argmin}} \left| \frac{1}{A} y_k - \underbrace{\sum_{i=1}^{\eta} f_i \hat{x}_{k-i}}_{\text{DFE of ISI}} - x \right|^2 \tag{7}$$

Notice the similarity between the branch metric calculation in the VA and this decision rule. In the ZF-DFE, previous decisions take the place of the state variables when calculating the intersymbol interference.

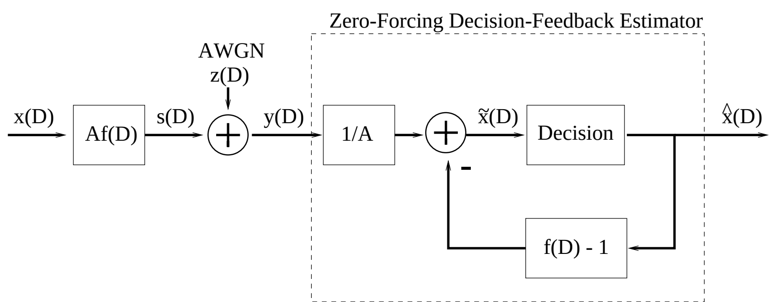


Figure 5: Block diagram of the linear Gaussian channel and zero-forcing decision-feedback estimator.

**Example 1 (cont.):** ***FIR Channel*** Figure 6 shows the ZF-DFE of $x(D)$ for the received sequence $y(D) = 2.1 - 2.9D$.

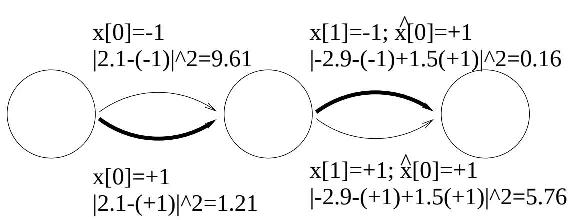


Figure 6: ZF-DFE of x(D) on the FIR channel $f(D) = 1 - 1.5D + 0.5D^2$ based on the received sequence $y(D) = 2.1 - 2.9D$.

## 6 Delayed Decision Feedback Sequence Estimation (DDFSE)

Delayed decision-feedback sequence estimation proposed in [1] is a hybrid of MLSE and ZF-DFE. It truncates the channel's impulse response and uses the VA, but with a ZF-DFE on each branch. The decision feedback estimate of the intersymbol interference comes from previous (delayed) decisions. For each state and next-state pair, the branch metric in the VA is

$$\begin{aligned} & |y_k - \sum_{i=0}^{\mu} A f_i x_{k-i} - \sum_{i=\mu+1}^{\eta} A f_i \hat{x}_{k-i}|^2 && (8) \\ \propto \; & |\frac{1}{A} y_k - \underbrace{\underbrace{\sum_{i=1}^{\mu} f_i x_{k-i}}_{\text{State Contribution}} - \underbrace{\sum_{i=\mu+1}^{\eta} f_i \hat{x}_{k-i}}_{\text{Delayed Decision Contribution}}}_{\text{ISI}} - x_k|^2 && (9) \end{aligned}$$

where $0 \leq \mu \leq \eta = deg\{f(D)\}$ and the state is defined as $\mathbf{X_k} = (x_{k-1}, \ldots, x_{k-\mu})$. The previous decisions $\hat{x}_{k-\mu-1}, \ldots, \hat{x}_{k-\eta}$ come from the states on the minimum distance path prior to state $\mathbf{X_k}$. Notice that when $\mu = 0$, the DDFSE reduces to the ZF-DFE. When $\mu = \eta$, the DDFSE becomes the MLSE.

When $\eta < \infty$, the computation of the branch metric in (8) requires the storage of only a finite number of past decisions. For $\eta = \infty$, we also want a way to compute the delayed decision contribution to ISI requiring the storage of only a finite number of past decisions.

Suppose that the channel impulse response is rational, i.e. $f(D) = \beta(D)/\gamma(D)$ for some polynomials $\beta(D)$ and $\gamma(D)$. Group $f(D)$ into two parts,

$$f(D) = f_\mu(D) + D^{\mu+1} f^+(D) \tag{10}$$

where

$$\begin{aligned} f_\mu(D) &= 1 + f_1 D + \cdots f_\mu D^\mu && (11) \\ f^+(D) &= f_{\mu+1} + f_{\mu+2} D + \cdots f_\eta D^\eta && (12) \\ &= \frac{(\beta(D) - f_\mu(D)\gamma(D))D^{-(\mu+1)}}{\gamma(D)} && (13) \\ &= \frac{\beta^+(D)}{\gamma(D)} && (14) \end{aligned}$$

and $\beta^+(D) = (\beta(D) - f_\mu(D)\gamma(D))D^{-(\mu+1)}$. Define $n = deg\{\beta^+(D)\}$ and $m = deg\{\gamma(D)\}$. Let

$$\begin{aligned} w(D) &= f^+(D) x(D) && (15) \\ &= \frac{\beta^+(D)}{\gamma(D)} x(D) && (16) \end{aligned}$$

So,

$$w_k = \begin{cases} \sum_{i=0}^{n} \beta_i^+ x_{k-i} - \sum_{i=1}^{m} \gamma_i w_{k-i} & (m > 0) \\ \sum_{i=0}^{n} f_{i+\mu+1} x_{k-i} & (m = 0) \end{cases} \tag{17}$$

We can then write the branch metric as

$$|y_k - \sum_{i=0}^{\mu} A f_i x_{k-i} - A\hat{w}_{k-\mu-1}|^2 \tag{18}$$

where

$$\hat{w}_{k-\mu-1} = \sum_{i=\mu+1}^{\eta} f_i \hat{x}_{k-i} \tag{19}$$

$$= \begin{cases} \sum_{i=0}^{n} \beta_i^+ \hat{x}_{k-\mu-1-i} - \sum_{i=1}^{m} \gamma_i \hat{w}_{k-\mu-1-i} & (m > 0) \\ \sum_{i=0}^{\eta-\mu-1} f_i \hat{x}_{k-\mu-1-i} & (m = 0) \end{cases} \tag{20}$$

Hence, for $m > 0$, each branch metric calculation only requires the current state $\mathbf{X_k} = (x_{k-1}, x_{k-2}, \cdots, x_{k-\mu})$, the previous $n$ decisions $\{\hat{x}_{k-\mu-1}, \ldots, \hat{x}_{k-\mu-n}\}$ from the minimum distance path, and the previous $m$ estimates $\{\hat{w}_{k-\mu-2}, \cdots, \hat{w}_{k-\mu-m-1}$ of the delayed decision contribution to the ISI.

**Example 1 (cont.):** ***FIR Channel [1]*** Again consider estimating the transmitted sequence sent into the FIR channel $f(D) = 1 - 1.5D + 0.5D^2$ based on the received sequence $y(D) = 2.1 - 2.9D$. With $\mu = 1$, the state of the DDFSE trellis is $\mathbf{X_k} = x_{k-1}$, and the decision feedback contribution to intersymbol interference is $\hat{w}_{k-2} = 0.5\hat{x}_{k-2}$. Figure 7 shows the DDFSE of the transmitted sequence $x(D) = 1 - D$.

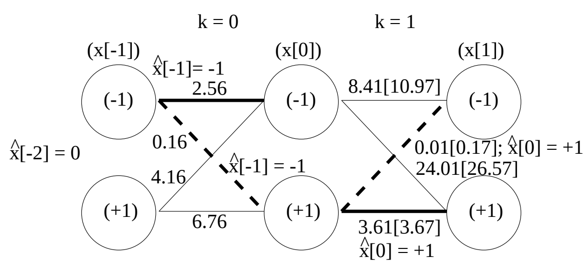


Figure 7: DDFSE of the transmitted sequence $x(D) = x_0 + x_1 D$ through the FIR channel $f(D) = 1 - 1.5D + 0.5D^2$ based on receive sequence $y(D) = 2.1 - 2.9D$.

**Example 2:** ***A One-Pole Channel Model [1]*** Consider the following IIR channel that often arises in optical recording:

$$f(D) = \sum_{i=0}^{\infty} \alpha^k D^k = \frac{A}{1 - \alpha D} \tag{21}$$

where $0 < \alpha < 1$. For the DDFSE with complexity parameter $\mu$,

$$f_\mu(D) = 1 + \alpha D + \alpha^2 D + \cdots + \alpha^\mu D^\mu$$

$$f^+(D) = \sum_{k=\mu+1}^{\infty} \alpha^k D^k = \frac{\alpha^{\mu+1}}{1 - \alpha D}$$

$$\hat{w}_k = \alpha^{\mu+1} \hat{x}_{k-\mu} + \alpha \hat{w}_{k-1}$$

# 7 Conclusions

We have summarized the implementation of the delayed decision-feedback sequence estimation (DDFSE) algorithm proposed in [1]. Its performance and complexity vary between that of a ZF-DFE and a MLSE, with each as special cases.


# 8 Acknowledgments

I would like to thank Prof. Forney for his direction in researching this topic. I would also like to thank Prof. Tarokh and Mike Lopez for suggesting papers for reduced-complexity MLSE algorithms.


# References


[1] A. Duel-Hallen. Delayed decision-feedback sequence estimation. *IEEE Transactions on Communications*, 37(5):428–436, May 1989.

[2] M.V. Eyuboglu and S. Qureshi. Reduced-state sequence estimation with set partitioning and decision feedback. *IEEE Transactions on Communications*, 36(1):13–20, January 1988.

[3] G.D. Forney. Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference. *IEEE Trans. Inform. Theory*, IT-18:363–378, May 1972.

[4] G.D. Forney. Chapter 10: Linear gaussian channels. In *Digital Communication Course Notes*, pages 125–145. Massachusetts Institute of Technology, 2000.