Trellis-Based Equalization for Sparse ISI Channels Revisited

Jan Mietzner¹  Sabah Badri-Hoeher¹  Ingmar Land²
Peter A. Hoeher¹

¹Information and Coding Theory Lab, University of Kiel, Germany
{jm,sbh,ph}@tf.uni-kiel.de

²Digital Communications Division, Aalborg University, Denmark
il@kom.aau.dk

ISIT 2005, Adelaide, Australia
September 4-9, 2005
Sparse ISI Channels

**Sparse ISI channels** are encountered in many high-data-rate communication systems (wireless & wireline)
Sparse ISI channels are encountered in many high-data-rate communication systems (wireless & wireline)

Discrete-time channel impulse response (CIR): **Large** memory length $L$, only **few** non-zero channel coefficients ($G \ll L$)

\[
\mathbf{h} := \begin{bmatrix}
  h_0 & 0 & \ldots & 0 & h_1 & 0 & \ldots & 0 & h_2 & \ldots & h_{G-1} & 0 & \ldots & 0 & h_G
\end{bmatrix}^T
\]

- $f_0$ zeros
- $f_1$ zeros
- $f_{G-1}$ zeros
Sparse ISI channels are encountered in many high-data-rate communication systems (wireless & wireline)

Discrete-time channel impulse response (CIR): Large memory length $L$, only few non-zero channel coefficients ($G \ll L$)

$$h := [h_0 \underbrace{0 \ldots 0}_{f_0 	ext{ zeros}} h_1 \underbrace{0 \ldots 0}_{f_1 	ext{ zeros}} h_2 \ldots h_{G-1} \underbrace{0 \ldots 0}_{f_{G-1} 	ext{ zeros}} h_G]^T$$

Special case: Zero-pad channel

$$f_0 = f_1 = \ldots = f_{G-1} =: f \geq 1$$
Equalization for Sparse ISI Channels

Discrete-time channel model

\[ y[k] = h_0 x[k] + \sum_{g=1}^{G} h_g x[k-d_g] + n[k] \]

- \( y[k] \): \( k \)th received sample
- \( x[k] \): \( k \)th transmitted data symbol
- \( n[k] \): \( k \)th AWGN sample
- \( d_g \): Position of \( h_g \) within \( h \)
### Discrete-time channel model

\[
y[k] = h_0 x[k] + \sum_{g=1}^{G} h_g x[k-d_g] + n[k]
\]

- \( y[k] \): \( k \)th received sample
- \( x[k] \): \( k \)th transmitted data symbol
- \( n[k] \): \( k \)th AWGN sample
- \( d_g \): Position of \( h_g \) within \( h \)

Due to large channel memory length, **efficient** equalization with reasonable **complexity** is a demanding task.
Discrete-time channel model

\[ y[k] = h_0 x[k] + \sum_{g=1}^{G} h_g x[k-d_g] + n[k] \]

- \( y[k] \): \( k \)th received sample
- \( x[k] \): \( k \)th transmitted data symbol
- \( n[k] \): \( k \)th AWGN sample
- \( d_g \): Position of \( h_g \) within \( h \)

Due to large channel memory length, efficient equalization with reasonable complexity is a demanding task.

Here: Trellis-based equalization (based on VA or BCJRA)

MLSE prohibitive \( \Rightarrow M^L \) trellis states (\( M \)-ary data symbols)
Exploiting the sparse channel structure, reduced-complexity algorithms can be derived.
Existing Solutions

Exploiting the sparse channel structure, reduced-complexity algorithms can be derived

Zero-pad channel:
- McGinty/Kennedy/Hoeher’98: Parallel-trellis VA (P-VA)
- Lee/McLane’02: Parallel-trellis BCJRA (P-BCJRA)
Existing Solutions

**Exploiting** the sparse channel structure, **reduced-complexity** algorithms can be derived

Zero-pad channel:

- McGinty/Kennedy/Hoehler’98: Parallel-trellis VA (P-VA)
- Lee/McLane’02: Parallel-trellis BCJRA (P-BCJRA)

⇒ Still **optimal** in the sense of MLSE
⇒ Based on parallel **regular** trellises
Existing Solutions

Exploiting the sparse channel structure, reduced-complexity algorithms can be derived.

Zero-pad channel:
- McGinty/Kennedy/Hoeher’98: Parallel-trellis VA (P-VA)
- Lee/McLane’02: Parallel-trellis BCJRA (P-BCJRA)

⇒ Still optimal in the sense of MLSE
⇒ Based on parallel regular trellises

General sparse channel:
- Benvenuto/Marchesani’96: Multi-trellis VA (M-VA)

⇒ Based on parallel irregular (i.e., time-varying) trellises
Introduction

Complexity Reduction Without Loss of Optimality

- Unified Framework Based on Factor Graphs
- Recapitulation of the P-VA and the M-VA
- Drawbacks of the Existing Solutions

Simple Equalization Scheme for General Sparse ISI Channels

Conclusions
Decomposition into multiple parallel trellises

Key question:

Which symbol decisions $\hat{x}[k], 1 \leq k \leq K_B$ ($K_B$ block length) are influenced by a certain symbol hypothesis $\tilde{x}[k_0]$?
Decomposition into multiple parallel trellises

Key question:
Which symbol decisions $\hat{x}[k]$, $1 \leq k \leq K_B$ ($K_B$ block length) are influenced by a certain symbol hypothesis $\tilde{x}[k_0]$?

- Suppose, a certain decision $\hat{x}[k_1]$ is not influenced by $\tilde{x}[k_0]$
- Suppose, there are no symbol decisions $\hat{x}[k]$ that are influenced by both $\tilde{x}[k_0]$ and $\tilde{x}[k_1]$
Decomposition into multiple parallel trellises

Key question:

Which symbol decisions $\hat{x}[k], 1 \leq k \leq K_B$ ($K_B$ block length) are influenced by a certain symbol hypothesis $\tilde{x}[k_0]$?

- Suppose, a certain decision $\hat{x}[k_1]$ is not influenced by $\tilde{x}[k_0]$
- Suppose, there are no symbol decisions $\hat{x}[k]$ that are influenced by both $\tilde{x}[k_0]$ and $\tilde{x}[k_1]$

$\Rightarrow \tilde{x}[k_0]$ and $\tilde{x}[k_1]$ can be accommodated in separate trellises without loss of optimality
Example 1: \( h := [h_0 \ 0 \ 0 \ 0 \ 0 \ 0 \ h_1 \ 0 \ h_2]^T \) \( (L=8, \ G=2) \)
Example 1: \[ h := [ h_0 \ 0 \ 0 \ 0 \ 0 \ 0 \ h_1 \ 0 \ h_2 ]^T \quad (L=8, \ G=2) \]
Framework for Complexity Reduction

Example 1: \( h := \begin{bmatrix} h_0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^T \quad (L=8, \ G=2) \)

\[
\begin{align*}
y[k_0] &= h_0 x[k_0] + h_1 x[k_0-6] + h_2 x[k_0-8] \\
y[k_0+6] &= h_0 x[k_0+6] + h_1 x[k_0] + h_2 x[k_0-2] \\
y[k_0+8] &= h_0 x[k_0+8] + h_1 x[k_0+2] + h_2 x[k_0]
\end{align*}
\]
Example 1: \( h := [h_0 \ 0 \ 0 \ 0 \ 0 \ 0 \ h_1 \ 0 \ h_2]^T \) \( (L=8, \ G=2) \)
Example 1: \( h := [h_0 \ 0 \ 0 \ 0 \ 0 \ 0 \ h_1 \ 0 \ h_2]^T \) \( (L=8, \ G=2) \)

\[\begin{align*}
\bar{x}[k_0] \\
 x[k_0] & \ x[k_0+1] & \ x[k_0+2] & +3 & +4 & +5 & +6 & +7 & +8 & +9 & +10 & +11 & +12 & +13 & +14 & +15 & x[k_0+16] \\
y[k_0] & \ y[k_0+1] & \ y[k_0+2] & +3 & +4 & +5 & +6 & +7 & +8 & +9 & +10 & +11 & +12 & +13 & +14 & +15 & y[k_0+16]
\end{align*}\]

\( \Rightarrow \) Two parallel (regular) trellises are still optimal!

\( \Rightarrow \) Parallel-trellis VA/BCJRA
Example 2: \[ h := [h_0 \ 0 \ 0 \ 0 \ 0 \ 0 \ h_1 \ h_2]^T \quad (L=8, \ G=2) \]
Framework for Complexity Reduction

Example 2: \[ h := \begin{bmatrix} h_0 & 0 & 0 & 0 & 0 & 0 & h_1 & h_2 \end{bmatrix}^T \quad (L=8, \ G=2) \]

⇒ Decomposition into parallel regular trellises not possible
(without loss of optimality)!
Example 2: \( h := [ h_0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ h_1 \ h_2 ]^T \) \( (L=8, \ G=2) \)

⇒ Decomposition into parallel regular trellises not possible (without loss of optimality)!

Multi-trellis VA neglects most of the dependencies ⇒ suboptimal!
Alternative solution for **general sparse channels**: Suboptimal parallel-trellis VA/BCJRA (McGinty/Kennedy/Hoeher’98, Lee/McLane’02)
Alternative solution for **general sparse channels**: Suboptimal parallel-trellis VA/BCJRA (McGinty/Kennedy/Hoeher’98, Lee/McLane’02)

(a) Find an underlying zero-pad CIR similar to the given CIR
Alternative solution for **general sparse channels**: Suboptimal parallel-trellis VA/BCJRA (McGinty/Kennedy/Hoeher’98, Lee/McLane’02)

(a) Find an underlying zero-pad CIR similar to the given CIR
(b) Define the parallel trellis diagrams
Alternative solution for **general sparse channels:**

**Suboptimal** parallel-trellis VA/BCJRA

(McGinty/Kennedy/Hoeher’98, Lee/McLane’02)

(a) Find an underlying zero-pad CIR similar to the given CIR
(b) Define the parallel trellis diagrams
(c) Perform decision feedback between the parallel trellises
Drawbacks

(a) Fading channel
(a) Fading channel $\Rightarrow$ Start all over again!
(a) Fading channel $\Rightarrow$ Start all over again!
(b) In practice, no exact zero coefficients
Drawbacks

(a) Fading channel ⇒ Start all over again!
(b) In practice, no exact zero coefficients

Our approach

It does not seem useful to **explicitly** utilize the **sparse** channel structure

⇒ How good are **standard** suboptimal equalization techniques?
Drawbacks

(a) Fading channel ⇒ Start all over again!
(b) In practice, no exact zero coefficients

Our approach

Use **prefiltering** in conjunction with **standard** reduced-complexity **trellis-based equalizer**

⇒ Tackle **general** sparse fading CIRs & provide performance **close** to the matched filter bound (MFB)
Outline

▶ Introduction

▶ Complexity Reduction Without Loss of Optimality

▶ Simple Equalization Scheme for General Sparse ISI Channels
  - Considered Receiver Structure
  - Numerical Results

▶ Conclusions
Considered Receiver Structure

\[ x[k] \rightarrow \text{ISI channel + AWGN} \rightarrow y[k] \rightarrow \text{Linear prefilter} \rightarrow z[k] \rightarrow \text{Trellis-based equalizer} \rightarrow \hat{x}[k] \]

- **Linear prefilter** that can be computed **efficiently**
  (with standard techniques available in the literature)
Considered Receiver Structure

- **Linear prefilter** that can be computed **efficiently** (with standard techniques available in the literature)
- **Standard** reduced-complexity **trellis-based equalizer** (not specifically designed for sparse ISI channels, since sparse CIR structure is normally lost after prefiltering)

⇒ Solely the linear prefilter is adjusted to the current CIR
Considered Receiver Structure

\[ x[k] \xrightarrow{\text{ISI channel} + \text{AWGN}} y[k] \xrightarrow{\text{linear prefilter}} z[k] \xrightarrow{\text{Trellis-based equalizer}} \hat{x}[k] \]

- **Linear prefilter** that can be computed **efficiently** (with standard techniques available in the literature)
- **Standard** reduced-complexity **trellis-based equalizer** (not specifically designed for sparse ISI channels, since sparse CIR structure is normally lost after prefiltering)

⇒ Solely the linear prefilter is adjusted to the current CIR

**Example:** Minimum-phase filter (WMF) in conjunction with delayed decision-feedback sequence estimator (DDFSE)
Comparison with sub-P-BCJRA

Static CIR $h = [h_0 0 0 0 h_4 0 0 h_7 0 \ldots 0 h_{15}]^T$ (no zero-pad)

$h_0 = 0.87, h_4 = h_7 = h_{15} = 0.29$

Binary transmission; WMF with $L_F = 40$ filter taps;

DDFSE with memory length $K \ll L$
Comparison with sub-P-BCJRA

Static CIR \( \mathbf{h} = [h_0 \ 0 \ 0 \ 0 \ h_4 \ 0 \ 0 \ h_7 \ 0 \ldots 0 \ h_{15}]^T \) (no zero-pad)

\[ h_0 = 0.87, \ h_4 = h_7 = h_{15} = 0.29 \]

Binary transmission; WMF with \( L_F = 40 \) filter taps;
DDFSE with memory length \( K \ll L \)

DDFSE \( (K = 4) \) + WMF: Similar performance as sub-P-BCJRA
Comparison with sub-P-BCJRA

Static CIR $h = [h_0 \ 0 \ 0 \ 0 \ h_4 \ 0 \ 0 \ h_7 \ 0 \ldots \ 0 \ h_{15}]^T$ (no zero-pad)

$h_0 = 0.87$, $h_4 = h_7 = h_{15} = 0.29$

Binary transmission; WMF with $L_F = 40$ filter taps;

DDFSE with memory length $K \ll L$

**Diagram:**

- **DDFSE ($K=4$) + WMF:** Similar performance as sub-P-BCJRA
- **DDFSE ($K=3$) + WMF:** Reduced complexity at expense of small loss
Fading CIR with $h_g \sim \mathcal{CN}(0, \sigma_{h,g}^2)$ and power profile

$$p := [\sigma_{h,0}^2 \underbrace{0 \ldots 0}_{f \text{ zeros}} \sigma_{h,1}^2 0 0 0 \sigma_{h,2}^2 \sigma_{h,3}^2]^T, \quad \sigma_{h,g}^2 = 0.25$$
Fading CIR with $h_g \sim \mathcal{CN}(0, \sigma_{h,g}^2)$ and power profile

$$p := \left[ \sigma_{h,0}^2 \underbrace{0 \ldots 0}_{f \text{ zeros}} \sigma_{h,1}^2 \ 0 \ 0 \ 0 \ \sigma_{h,2}^2 \ \sigma_{h,3}^2 \right]^T, \ \sigma_{h,g}^2 = 0.25$$

Memory length $L = 6$ ($K = 5$)
Fading CIR with $h_g \sim \mathcal{CN}(0, \sigma^2_{h,g})$ and power profile

$$p := \begin{bmatrix} \sigma^2_{h,0} & 0 \cdots 0 & \sigma^2_{h,1} & 0 & 0 & 0 & \sigma^2_{h,2} & \sigma^2_{h,3} \end{bmatrix}^T, \quad \sigma^2_{h,g} = 0.25$$

Memory length $L = 12$ ($K = 5$)
Fading CIR, Different Memory Lengths

Fading CIR with $h_g \sim \mathcal{CN}(0, \sigma_{h,g}^2)$ and power profile

$$p := [\sigma_{h,0}^2, 0 \ldots 0, \sigma_{h,1}^2, 0, 0, 0, \sigma_{h,2}^2, \sigma_{h,3}^2]^T, \quad \sigma_{h,g}^2 = 0.25$$

Memory length $L = 20 \ (K = 5)$
Fading CIR, Different Memory Lengths

Fading CIR with $h_g \sim \mathcal{CN}(0, \sigma_{h,g}^2)$ and power profile

$$p := [\sigma_{h,0}^2 \underbrace{0 \ldots 0}_{f \text{ zeros}} \sigma_{h,1}^2 0 0 0 \sigma_{h,2}^2 \sigma_{h,3}^2]^T, \quad \sigma_{h,g}^2 = 0.25$$

Memory length $L = 20$ ($K = 5$)

DDFSE with WMF deviates only 1-2 dB from the MFB (at BER $10^{-3}$) even for a large memory length $L$

WMF makes huge difference
Conclusions

- Efficient equalization of **sparse** ISI channels at reasonable complexity is a **demanding task**
- **Optimal** trellis-based solutions are only applicable for zero-pad channels (factor graph)
- Existing **suboptimal** solutions explicitly exploit the sparse channel structure and seem impracticable for **fading** channels
- **Our approach:** Use linear prefilter in conjunction with standard reduced-complexity trellis-based equalizer
Conclusions

- Efficient equalization of sparse ISI channels at reasonable complexity is a demanding task
- Optimal trellis-based solutions are only applicable for zero-pad channels (factor graph)
- Existing suboptimal solutions explicitly exploit the sparse channel structure and seem impracticable for fading channels
- Our approach: Use linear prefilter in conjunction with standard reduced-complexity trellis-based equalizer

⇒ General sparse ISI channels can be tackled
⇒ Only the linear prefilter is adjusted to the current CIR
⇒ Performance close to the MFB