Joint Channel Estimation and Sequence Detection over Doubly Dispersive Channels

Sung-Jun Hwang and Phil Schniter

Nov. 2006

Problem Description:

- Uncoded block transmission over an ISI channel that varies significantly over the block.
- Data symbols and channel are both unknown. At least one known pilot symbol.
- Interested in near-optimal sequence detection with reasonable complexity.

Related Work:

- Per-survivor processing (PSP): trellis-based equalization using the surviving partial-paths as training for adaptive channel estimation.
- Joint estimation/MLSD for singly-dispersive channels.

System Model:

Received samples are $\{r_n\}_{n=0}^{N-1}$, where

$$r_n = \sum_{l=0}^{N_h-1} h_{n,l} s_{n-l} + v_n,$$

$h_{n,l}$: time-n response to an impulse at time $(n-l)$.

N_h : discrete channel spread.

$\{s_n\}_{n=0}^{N-1}$: symbols from finite alphabet $\mathcal{Q}$

$\{v_n\}_{n=0}^{N-1}$: CWGN with variance σ^2.

We assume WSSUS fading:

$$E\{h_{n,l} h^*_{n-m,l-p}\} = \rho_m \sigma_l^2 \delta_p$$

Note: holds for single-carrier transmission over a time-varying ISI channel, or multicarrier transmission over a frequency-varying ICI channel.

BEM Approximation (Used by Receiver):

The receiver employs a basis expansion model (BEM)

$$h_{n,l} \approx \sum_{p=0}^{N_b-1} b_{n,p}\theta_{p,l} \quad \text{for } n \in \{0,\dots,N-1\}.$$

$$\begin{aligned} \{b_{n,p}\}_{n=0}^{N-1} &: \ p^{th} \text{ basis waveform} \\ N_b &: \ \text{number of basis waveforms} \\ \theta_{p,l} &: \ \text{coefficient for } p^{th} \text{ basis waveform and } l^{th} \text{ lag} \end{aligned}$$

BEM options include:

- oversampled complex exponential: $b_{n,p} = e^{j\frac{2\pi}{NK}pn}, K \geq 1$
- polynomial: $b_{n,p} = n^p$
- Karhunen-Loeve: $\{b_{n,p}\}_{n=0}^{N-1}$ is the p^{th} largest eigenvector of Toeplitz correlation matrix defined from $\{\rho_m\}_{m=0}^{N-1}$

BEM-Approximated System Model:

$$\boxed{\boldsymbol{r}_n = \boldsymbol{B}_n \boldsymbol{S}_0^n \boldsymbol{\theta} + \boldsymbol{v}_n}$$

where, by example,

$$\underbrace{\begin{bmatrix} r_2 \\ r_1 \\ r_0 \end{bmatrix}}_{\boldsymbol{r}_2} = \underbrace{\begin{bmatrix} \boldsymbol{b}_2^H & & \\ & \boldsymbol{b}_1^H & \\ & & \boldsymbol{b}_0^H \end{bmatrix}}_{\boldsymbol{B}_2} \underbrace{\begin{bmatrix} s_2 \boldsymbol{I}_{N_b} & s_1 \boldsymbol{I}_{N_b} \\ s_1 \boldsymbol{I}_{N_b} & s_0 \boldsymbol{I}_{N_b} \\ s_0 \boldsymbol{I}_{N_b} & s_{-1} \boldsymbol{I}_{N_b} \end{bmatrix}}_{\boldsymbol{S}_0^2} \underbrace{\begin{bmatrix} \boldsymbol{\theta}_0 \\ \boldsymbol{\theta}_1 \end{bmatrix}}_{\boldsymbol{\theta}} + \underbrace{\begin{bmatrix} v_2 \\ v_1 \\ v_0 \end{bmatrix}}_{\boldsymbol{v}_2}$$

$$\boldsymbol{b}_n = [b_{n,0}, \ldots, b_{n,N_b-1}]^H \quad : \text{time-}n \text{ basis values}$$

$$\boldsymbol{\theta}_l = [\theta_{0,l}, \ldots, \theta_{N_b,l}]^T \quad : \text{lag-}l \text{ BEM coefficients}$$

Note:

- $\boldsymbol{\theta} \in \mathbb{C}^{N_b N_h}$ contains all time-varying channel parameters
- $\boldsymbol{S}_0^n$ contains data symbols $\boldsymbol{s}_n = [s_n, \ldots, s_0]^T$

Noncoherent Data Detection:

MLSD criterion:

$$\hat{\boldsymbol{s}}_n = \arg\max_{\boldsymbol{s}_n} p(\boldsymbol{r}_n|\boldsymbol{s}_n)$$

With prior channel pdf $p(\boldsymbol{\theta})$,

$$p(\boldsymbol{r}_n|\boldsymbol{s}_n) = \int_{\boldsymbol{\theta}} \overbrace{p(\boldsymbol{r}_n|\boldsymbol{s}_n,\boldsymbol{\theta})}^{\mathcal{CN}(\boldsymbol{B}_n\boldsymbol{S}_0^n\boldsymbol{\theta},\sigma^2\boldsymbol{I})} \underbrace{p(\boldsymbol{\theta})}_{\mathcal{CN}(\boldsymbol{0},\boldsymbol{R}_\theta)} d\boldsymbol{\theta}$$

After some algebra, we obtain a quadratic noncoherent metric

$$\hat{\boldsymbol{s}}_n = \arg\min_{\boldsymbol{s}_n} \Big\{ \underbrace{\boldsymbol{r}_n^H \boldsymbol{\Phi}_{\boldsymbol{s}_n} \boldsymbol{r}_n}_{\mu(\boldsymbol{s}_n)} + \log\det(\boldsymbol{\Sigma}_{\boldsymbol{s}_n}) \Big\} \approx \arg\min_{\boldsymbol{s}_n} \mu(\boldsymbol{s}_n)$$

for $\boldsymbol{\Phi}_{\boldsymbol{s}_n} = \big(\boldsymbol{B}_n\boldsymbol{S}_0^n\boldsymbol{R}_\theta(\boldsymbol{B}_n\boldsymbol{S}_0^n)^H + \sigma^2\boldsymbol{I}_{n+1}\big)^{-1}$.

Estimation/Detector Interpretation:

Can write noncoherent metric as

$$\mu(\boldsymbol{s}_n) \;=\; \underbrace{\sigma^{-2}\|\boldsymbol{r}_n - \boldsymbol{B}_n \boldsymbol{S}_0^n \hat{\boldsymbol{\theta}}_{\boldsymbol{s}_n}\|^2}_{\text{``coherent'' ML metric}} + \underbrace{\sigma^{-2}\hat{\boldsymbol{\theta}}_{\boldsymbol{s}_n}^H \boldsymbol{R}_\theta^{-1}\hat{\boldsymbol{\theta}}_{\boldsymbol{s}_n}}_{\text{prior reconciliation}}$$

where $\hat{\boldsymbol{\theta}}_{\boldsymbol{s}_n}$ is the MMSE estimate of $\boldsymbol{\theta}$ from $\boldsymbol{r}_n$ given $\boldsymbol{s}_n$:

$$\hat{\boldsymbol{\theta}}_{\boldsymbol{s}_n} \;=\; \mathrm{E}\{\boldsymbol{\theta}\boldsymbol{r}_n^H|\boldsymbol{s}_n\}\,\mathrm{E}\{\boldsymbol{r}_n\boldsymbol{r}_n^H|\boldsymbol{s}_n\}^{-1}\boldsymbol{r}_n$$

In other words, the noncoherent metric $\mu(\boldsymbol{s}_n)$ *adapts* to the channel that is implicitly estimated with $\boldsymbol{s}_n$ as training.

Note: Brute-force search evaluates $|\mathcal{Q}|^N$ metrics!!

Fast Adaptive Sequential Decoding:

1. Suboptimal breadth-first tree search via the M-algorithm:
   - Say $\boldsymbol{\mathcal{S}}_n$ contains the M "best" estimates of $\boldsymbol{s}_n$. For each extension $\boldsymbol{s}_{n+1} = \left[\begin{smallmatrix} s \\ \boldsymbol{s}_n \end{smallmatrix}\right]$, where $\boldsymbol{s}_n \in \boldsymbol{\mathcal{S}}_n$ and $s \in \mathcal{Q}$, calculate the noncoherent metric $\mu(\boldsymbol{s}_{n+1})$. Then keep M best in $\boldsymbol{\mathcal{S}}_{n+1}$.
   - In total, evaluates $M|\mathcal{Q}|N$ noncoherent metrics.
   - Performance almost indistinguishable from brute-force.
2. Fast metric computation:
   - Updating $\hat{\boldsymbol{\theta}}_{\boldsymbol{s}_n}$ requires only about $nN_bN_h + 4N_b^2N_h^2$ operations.

Assuming $N > N_bN_h$ (i.e., an underspread channel), the total complexity of calculating $\hat{\boldsymbol{s}}_{N-1}$ is $\boxed{\mathcal{O}(M|\mathcal{Q}|N^2N_bN_h)}$.

Fast Metric Computation:

Can write MMSE estimate as

$$\begin{aligned}\hat{\boldsymbol{\theta}}_{\boldsymbol{s}_n} &= \boldsymbol{\Sigma}_{\boldsymbol{s}_n}^{-1}\boldsymbol{A}_n^H\boldsymbol{r}_n\\ \boldsymbol{\Sigma}_{\boldsymbol{s}_n} &= \boldsymbol{A}_n^H\boldsymbol{A}_n + \sigma^2\boldsymbol{R}_\theta^{-1}\\ \boldsymbol{A}_n &= \boldsymbol{B}_n\boldsymbol{S}_0^n \ \in \mathbb{C}^{(n+1)\times N_b N_h}\end{aligned}$$

Noticing a rank-one update:

$$\begin{aligned}\boldsymbol{\Sigma}_{\boldsymbol{s}_{n+1}} &= \boldsymbol{\Sigma}_{\boldsymbol{s}_n} + \boldsymbol{a}_{n+1}\boldsymbol{a}_{n+1}^H\\ \boldsymbol{a}_{n+1}^H &= \boldsymbol{b}_{n+1}^H\boldsymbol{S}_{n+1}^{n+1} \ \in \mathbb{C}^{N_b N_h}\\ \boldsymbol{\Sigma}_{\boldsymbol{s}_{n+1}}^{-1} &= \boldsymbol{\Sigma}_{\boldsymbol{s}_n}^{-1} - \frac{(\boldsymbol{\Sigma}_{\boldsymbol{s}_n}^{-1}\boldsymbol{a}_{n+1})(\boldsymbol{\Sigma}_{\boldsymbol{s}_n}^{-1}\boldsymbol{a}_{n+1})^H}{1+\boldsymbol{a}_{n+1}^H\boldsymbol{\Sigma}_{\boldsymbol{s}_n}^{-1}\boldsymbol{a}_{n+1}},\end{aligned}$$

so complexity of calculating $\hat{\boldsymbol{\theta}}_{\boldsymbol{s}_{n+1}}$ is $\mathcal{O}(nN_bN_h)$ when $n > N_bN_h$.

Given $\hat{\boldsymbol{\theta}}_{\boldsymbol{s}_{n+1}}$, the complexity of calculating $\mu(\boldsymbol{s}_{n+1})$ is also $\mathcal{O}(nN_bN_h)$ when $n > N_bN_h$.

Construction of the Transmission Frame:

Pilots:

- One pilot symbol needed to resolve channel/data phase ambiguity.
- N_h leading pilots useful for "initializing" the metric, allowing for good M-alg performance with small M.
- $N_h N_b$ pilots required for pilot-aided estimation-then-detection.

Diversity:

- $N_h - 1$ trailing zeros needed to make full delay-diversity accessible.
- Doppler diversity not accessible without coding/precoding. (This issue will be treated in future work.)

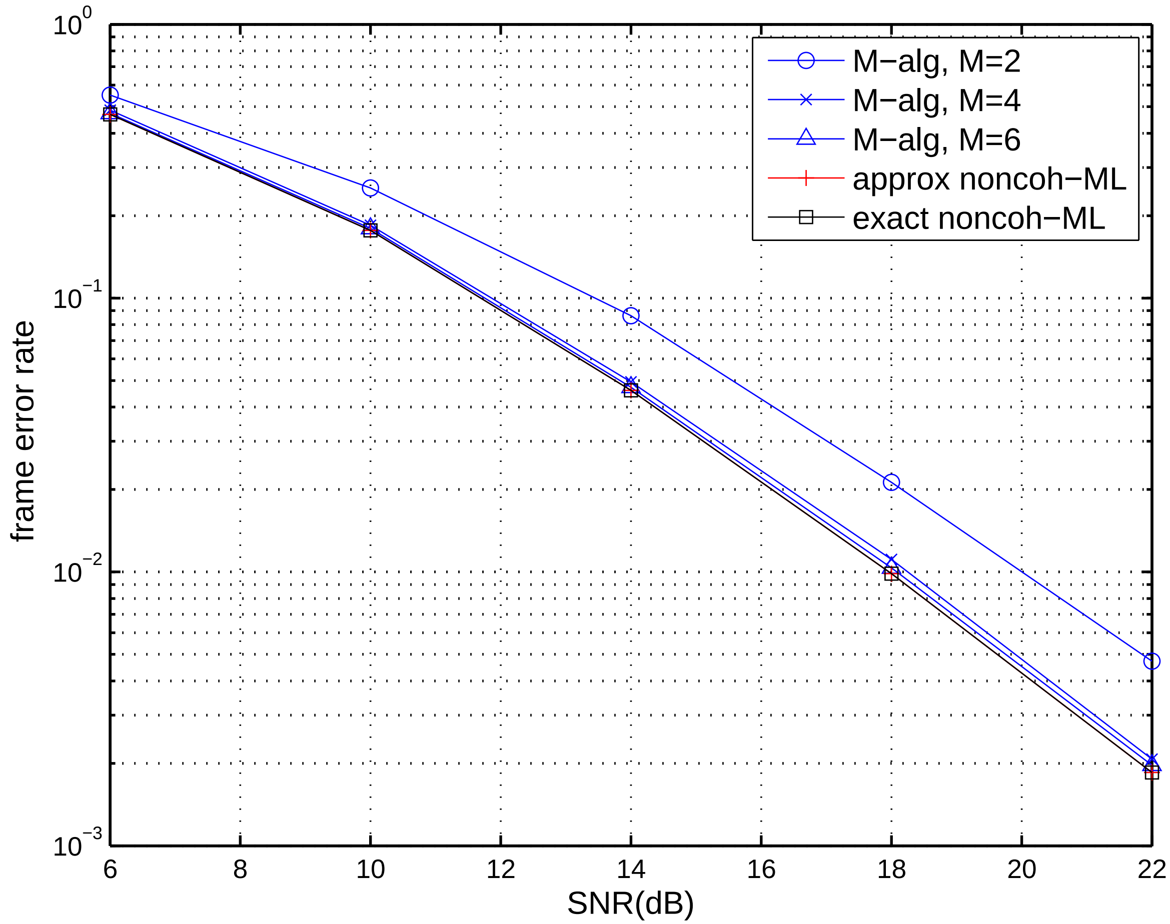
$\rightsquigarrow$*We insert $N_h N_b$ leading pilots to facilitate a fair comparison with estimation-then-detection schemes, and we insert $N_h - 1$ trailing zeros to make delay-diversity accessible.*

Numerical Experiments:

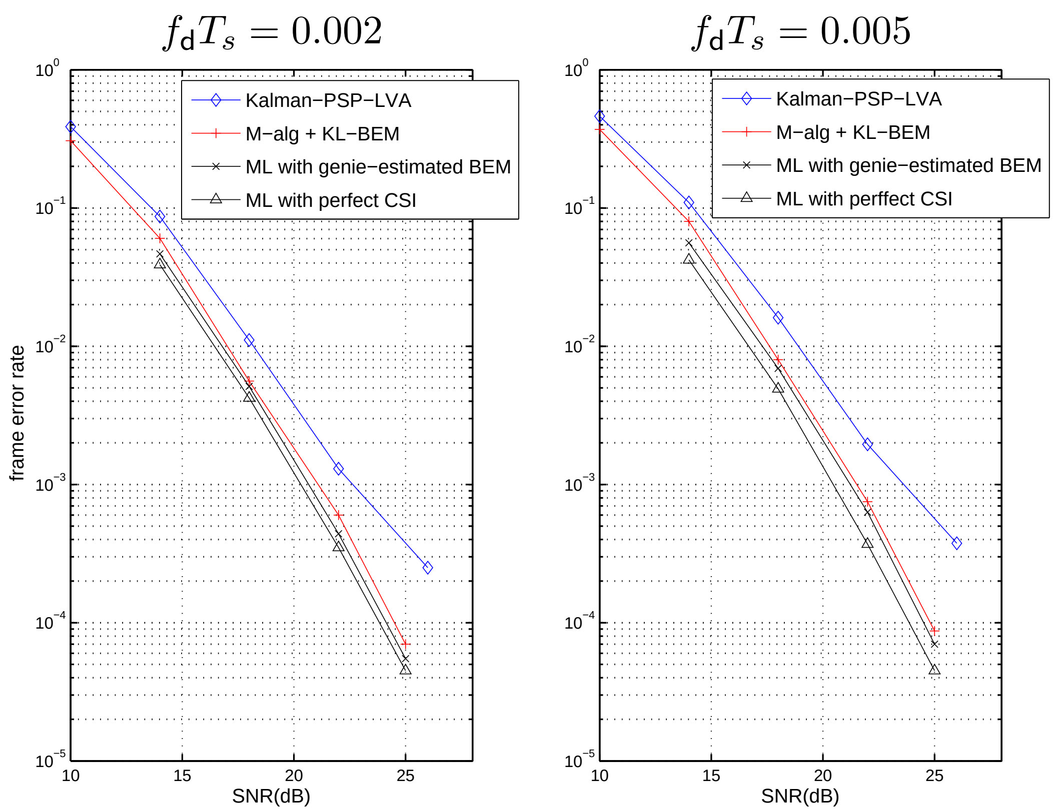
- BPSK symbols, $N = 25$
- WSSUS Jakes channel with delay spread $N_h = 2$ and single-sided Doppler spread $f_{\mathsf{d}} T_s \in \{0.002, 0.005\}$.
- Receiver BEMs ($N_b = 2$):
  1. Karhunen-Loeve (KL)
  2. Oversampled Complex Exponential (OCE)
- Reference Algorithms
  1. ML with perfect $\{h_{n,l}\}$ (genie-aided)
  2. ML with MMSE-$\hat{\boldsymbol{\theta}}$ from pilots+data (genie-aided)
  3. ML with MMSE-$\hat{\boldsymbol{\theta}}$ from pilots
  4. PSP with RLS-$\{\hat{h}_{n,l}\}$

Effect of Metric Approximation and Choice of M:

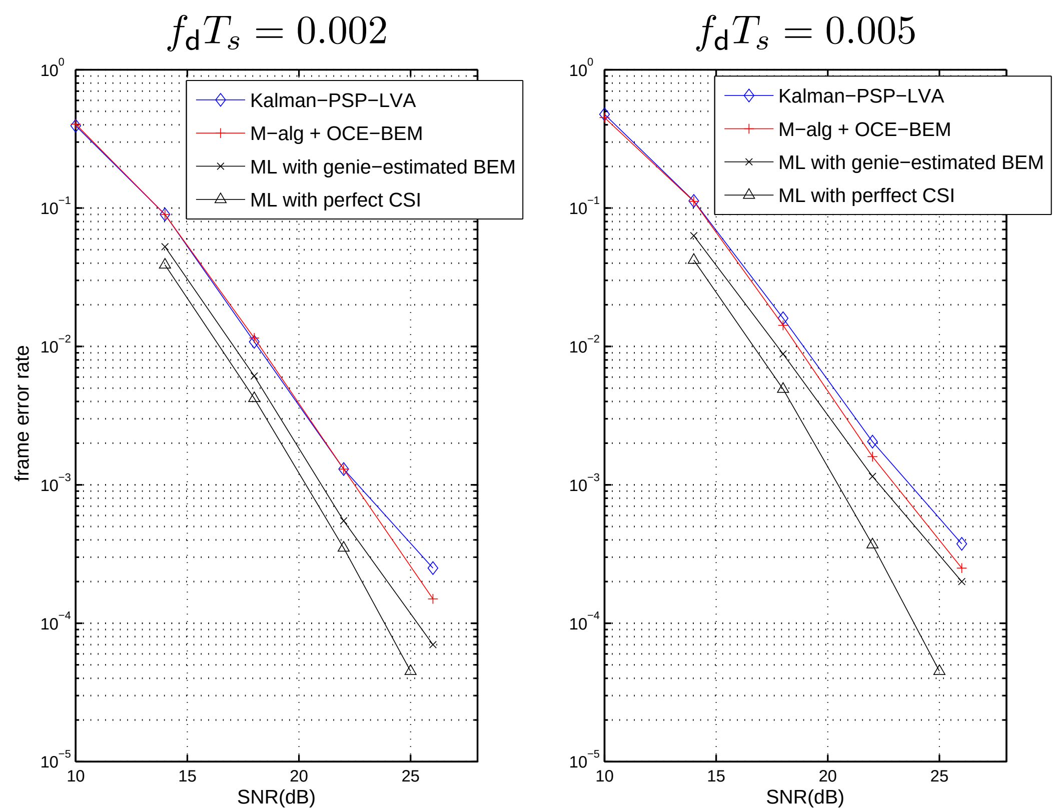
$$f_{\mathsf{d}}T_s = 0.005$$



Performance with KL-BEM:



Performance with OCE-BEM:



Conclusions:

- Joint channel/symbol estimation for quickly varying ISI channels.
- Leveraged BEM channel approximation, M-algorithm, fast MMSE-channel estimation.
- Less than 1 dB from optimal performance at complexity $\mathcal{O}(M|\mathcal{Q}|N^2 N_b N_h)$.
- Significantly outperforms decoupled channel/symbol estimation.
- Outperforms PSP-RLS, especially at high Doppler spreads.