EM-Based Soft Noncoherent Equalization of Doubly Selective Channels using Tree Search and Basis Expansion

Sung-Jun Hwang and Phil Schniter

SPAWC-2009, Perugia

June 22, 2009

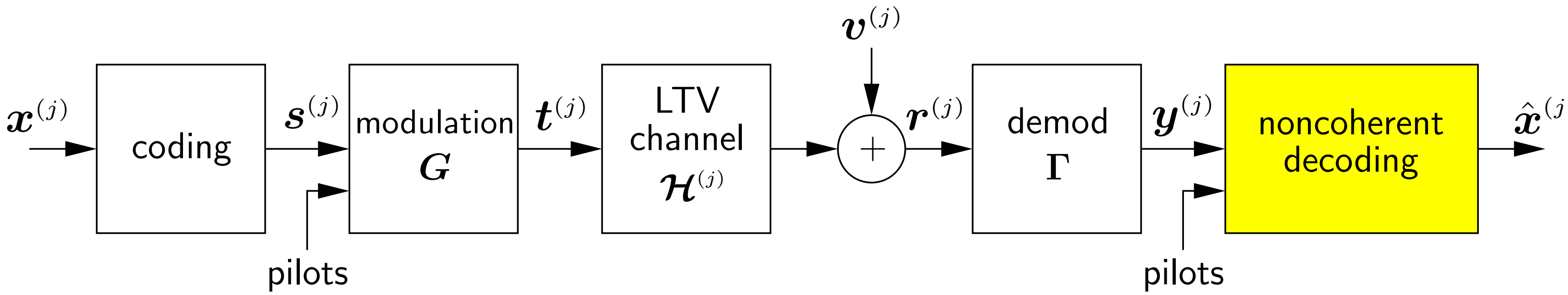
Problem Description:

- Coded block transmission over a doubly selective channel.
- Channel realizations *unknown*, but channel statistics known.
- Goal: near-optimal decoding with low complexity and few pilots.

Approach:

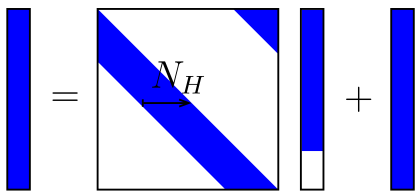
- Turbo reception (soft noncoherent equalization $\rightleftarrows$ soft decoding).
- Soft decoder: off-the-shelf LDPC.
- Soft noncoherent equalizer: a novel design leveraging. . .
  - the EM algorithm $\rightsquigarrow$ joint soft channel-estimation/equalization,
  - a basis expansion model (BEM) for channel variation,
  - a tree search for soft equalization.

Block Transmission/Reception Model:



The demodulated symbols for the j^{th} block take the form:

$$y^{(j)} = \underbrace{\Gamma \mathcal{H}^{(j)} G}_{H^{(j)}} s^{(j)} + z^{(j)}$$



where

- *single carrier* (ZP): $G = I_N$ and $\Gamma = \begin{pmatrix} I_{N_h-1} & 0 & I_{N_h-1} \\ 0 & I_{N-N_h+1} & 0 \end{pmatrix}$.
- *multi-carrier* (PS): $G = \mathcal{D}(g) F_t^H$ and $\Gamma = F_r \mathcal{D}(\gamma)$.

Basis Expansion Model:

We parameterize the d^{th} diagonal $\boldsymbol{h}_d^{(j)}$ of the matrix $\boldsymbol{H}^{(j)}$ using a BEM:

$$\boldsymbol{h}_d^{(j)} \approx \boldsymbol{B}\boldsymbol{\eta}_d^{(j)}, \qquad \boldsymbol{\theta}^{(j)} \triangleq \begin{bmatrix} \boldsymbol{\eta}_0^{(j)} \\ \vdots \\ \boldsymbol{\eta}_{N_H-1}^{(j)} \end{bmatrix} \in \mathbb{C}^{N_H N_b}.$$

yielding the system model

$$\boldsymbol{y}^{(j)} = \underbrace{\left[\mathcal{D}_0(\boldsymbol{s}^{(j)})\boldsymbol{B}, \ldots, \mathcal{D}_{N_H-1}(\boldsymbol{s}^{(j)})\boldsymbol{B}\right]}_{\boldsymbol{A}^{(j)}} \boldsymbol{\theta}^{(j)} + \boldsymbol{z}^{(j)},$$

Typical choices:

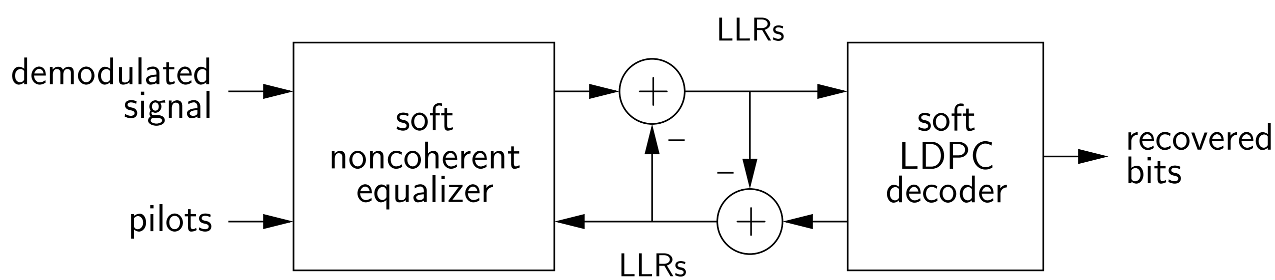
- *Single Carrier*: Karhunen-Loève, Polynomial, oversampled CE, DPS (models variation across time).
- *Multi-carrier*: complex exponential (models variation across frequency—a function of the delay profile).

Noncoherent Turbo Equalization

- Large performance gains are possible through the use of sophisticated coding schemes (e.g., LDPC).
- For complexity reasons, noncoherent decoding is split into
  1. *noncoherent equalization*, which leverages channel structure,
  2. *decoding*, which leverages the code structure.
- By *iterating* the two steps ("turbo equalization"), we hope to get

  near-optimal noncoherent decoding with practical complexity. ☺

  Note: Doing so requires *soft* equalization (and *soft* decoding).



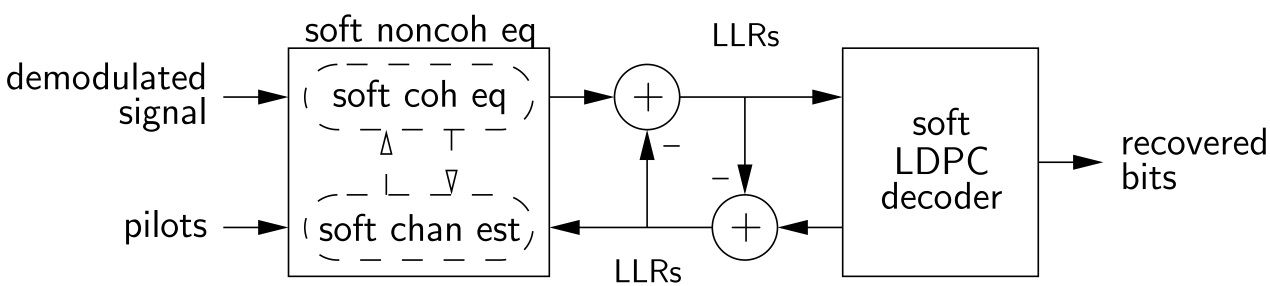
Soft Noncoherent Equalization

By "soft noncoherent equalization" we mean

computing coded-bit LLRs in the presence of an unknown channel.

Possible approaches:

1. Joint equalization/chan-est (MAP inspired)
2. Iterative equalization & chan-est (EM inspired)



3. Iterative equalization & chan-est (ad hoc)
4. Non-iterative equalization (with pilot-aided channel estimation)

Bayesian EM Algorithm:

Using symbols $\boldsymbol{s}$ as the "missing data," the i^{th} EM iteration becomes

$$\hat{\boldsymbol{\theta}}[i+1] \;=\; \arg\max_{\hat{\boldsymbol{\theta}}} \mathrm{E}\left\{ \ln p(\boldsymbol{y}, \boldsymbol{s} \,|\, \hat{\boldsymbol{\theta}}) \,\middle|\, \boldsymbol{y}, \hat{\boldsymbol{\theta}}[i] \right\} + \ln p(\hat{\boldsymbol{\theta}})$$

With the Ricean fading assumption $\boldsymbol{\theta} \sim \mathcal{CN}(\bar{\boldsymbol{\theta}}, \boldsymbol{R}_\theta)$, we get

$$\hat{\boldsymbol{\theta}}[i+1] \;=\; \bar{\boldsymbol{\theta}} + \left(\boldsymbol{C} + \sigma^2 \boldsymbol{R}_\theta^{-1}\right)^{-1} \left(\bar{\boldsymbol{A}}^H \boldsymbol{y} - \boldsymbol{C}\bar{\boldsymbol{\theta}}\right)$$

where

$$\bar{\boldsymbol{A}} \;=\; \left[\mathcal{D}_0(\bar{\boldsymbol{s}})\boldsymbol{B}, \ldots, \mathcal{D}_{N_H-1}(\bar{\boldsymbol{s}})\boldsymbol{B}\right]$$

$$\boldsymbol{C} \;=\; \bar{\boldsymbol{A}}^H \bar{\boldsymbol{A}} + \begin{bmatrix} \boldsymbol{B}^H \mathcal{D}_0(\boldsymbol{c})\boldsymbol{B} & & \\ & \ddots & \\ & & \boldsymbol{B}^H \mathcal{D}_{N_H-1}(\boldsymbol{c})\boldsymbol{B} \end{bmatrix}$$

use symbol means $\bar{\boldsymbol{s}} \triangleq \mathrm{E}\{\boldsymbol{s} \,|\, \boldsymbol{y}, \hat{\boldsymbol{\theta}}[i]\}$ & variances $\mathcal{D}(\boldsymbol{c}) \triangleq \mathrm{cov}\{\boldsymbol{s}, \boldsymbol{s} \,|\, \boldsymbol{y}, \hat{\boldsymbol{\theta}}[i]\}$ calculated via the previous channel estimate $\hat{\boldsymbol{\theta}}[i]$.

Soft Symbol Estimation:

We can use the (coherent) bit LLRs

$$L(x_k \,|\, \hat{\boldsymbol{\theta}}[i]) \triangleq \ln \frac{\Pr\{x_k = 1 \,|\, \boldsymbol{y}, \hat{\boldsymbol{\theta}}[i]\}}{\Pr\{x_k = 0 \,|\, \boldsymbol{y}, \hat{\boldsymbol{\theta}}[i]\}}$$

to calculate the symbol means/variances. For QPSK $s_n \in \{\pm 1 \pm j\}$, get

$$\begin{aligned} \bar{s}_n &= \tanh\{\tfrac{1}{2} L(x_{2n} \,|\, \hat{\boldsymbol{\theta}}[i])\} + j \tanh\{\tfrac{1}{2} L(x_{2n+1} \,|\, \hat{\boldsymbol{\theta}}[i])\} \\ c_n &= 2 - |\bar{s}_n|^2. \end{aligned}$$

The bit LLRs can be written using the metrics $\{\mu(\boldsymbol{x} \,|\, \hat{\boldsymbol{\theta}}[i])\}_{\boldsymbol{x} \in \{0,1\}^{QN}}$:

$$\begin{aligned} \mu(\boldsymbol{x} \,|\, \hat{\boldsymbol{\theta}}[i]) &= -\tfrac{1}{\sigma^2} \|\boldsymbol{y} - \boldsymbol{A} \hat{\boldsymbol{\theta}}[i]\|^2 + \boldsymbol{l}^T \boldsymbol{x}, \\ L(x_k \,|\, \hat{\boldsymbol{\theta}}[i]) &= \ln \frac{\sum_{\boldsymbol{x}:\, x_k = 1} \exp \mu(\boldsymbol{x} \,|\, \hat{\boldsymbol{\theta}}[i])}{\sum_{\boldsymbol{x}:\, x_k = 0} \exp \mu(\boldsymbol{x} \,|\, \hat{\boldsymbol{\theta}}[i])}, \end{aligned}$$

where $\boldsymbol{l} \triangleq [\ldots, L_a(x_k), \ldots]^T$ are prior LLRs (obtained from the decoder).

Simplified LLR Evaluation:

To avoid the 2^{QN}-term summations, we use the "max-log" approximation:

$$L(x_k \mid \hat{\boldsymbol{\theta}}[i]) \approx \max_{\boldsymbol{x} \in \mathcal{X}[i] \cap \{\boldsymbol{x}: x_k = 1\}} \mu(\boldsymbol{x} \mid \hat{\boldsymbol{\theta}}[i]) - \max_{\boldsymbol{x} \in \mathcal{X}[i] \cap \{\boldsymbol{x}: x_k = 0\}} \mu(\boldsymbol{x} \mid \hat{\boldsymbol{\theta}}[i])$$

$$\mathcal{X}[i] \ : \ \text{set containing the } M \text{ most probable } \boldsymbol{x},$$

which requires relatively few evaluations of $\mu(\boldsymbol{x} \mid \hat{\boldsymbol{\theta}}[i])$.

The set $\mathcal{X}[i]$ can be found efficiently using a (soft coherent) tree search, e.g., using the M-algorithm. The required complexity is $\mathcal{O}(M 2^Q N N_b N_H)$:

- *linear* in the block length N,
- *linear* in the number of channel coefficients $N_b N_H$.
- *linear* in the constellation size 2^Q.

Simplified Soft Channel Estimation — Multicarrier Case

We would like to avoid an $\mathcal{O}(N^3)$ matrix inversion in

$$\hat{\boldsymbol{\theta}}[i+1] \;=\; \bar{\boldsymbol{\theta}} + \big(\boldsymbol{C} + \sigma^2 \boldsymbol{R}_\theta^{-1}\big)^{-1}\big(\bar{\boldsymbol{A}}^H \boldsymbol{y} - \boldsymbol{C}\bar{\boldsymbol{\theta}}\big),$$

where

$$\begin{aligned}
\bar{\boldsymbol{A}} &= \big[\mathcal{D}_0(\bar{\boldsymbol{s}})\boldsymbol{B}, \dots, \mathcal{D}_{N_H-1}(\bar{\boldsymbol{s}})\boldsymbol{B}\big] \\
\boldsymbol{C} &= \bar{\boldsymbol{A}}^H \bar{\boldsymbol{A}} + \begin{bmatrix} \boldsymbol{B}^H \mathcal{D}_0(\boldsymbol{c})\boldsymbol{B} & & \\ & \ddots & \\ & & \boldsymbol{B}^H \mathcal{D}_{N_H-1}(\boldsymbol{c})\boldsymbol{B} \end{bmatrix}.
\end{aligned}$$

In the multicarrier case, we can exploit the facts that $\boldsymbol{R}_\theta$ is block diagonal and that multiplication-by-$\boldsymbol{B}$ can be calculated via an FFT.

Main idea: Use conjugate-gradient algorithm to solve for $\hat{\boldsymbol{\theta}}[i+1]$ iteratively. Complexity: $\mathcal{O}(N \log_2 N)$.

Simplified Soft Channel Estimation — General Case

Using the approximation $\boldsymbol{c} \approx \boldsymbol{0}$, we get

$$\hat{\boldsymbol{\theta}}[i{+}1] \;\approx\; \bar{\boldsymbol{\theta}} + \big(\bar{\boldsymbol{A}}^H \boldsymbol{A} + \sigma^2 \boldsymbol{R}_\theta^{-1}\big)^{-1} \bar{\boldsymbol{A}}^H \big(\boldsymbol{y} - \bar{\boldsymbol{A}}\bar{\boldsymbol{\theta}}\big),$$

which allows us to solve for $\hat{\boldsymbol{\theta}}[i{+}1]$ using a sequential-Bayes recursion:

set $\{\boldsymbol{\Sigma}_{-1}^{-1},\ \hat{\boldsymbol{\theta}}_{-1}[i{+}1]\} \triangleq \{\sigma^{-2}\boldsymbol{R}_\theta,\ \bar{\boldsymbol{\theta}}\}$;
for $n = 0, 1, 2, \dots, N-1$,
$\quad \boldsymbol{a}_n = [\bar{s}_n \boldsymbol{b}_n^H, \cdots, \bar{s}_{n-N_H+1}\boldsymbol{b}_n^H]^H$;
$\quad \boldsymbol{d}_n = \boldsymbol{\Sigma}_{n-1}^{-1}\boldsymbol{a}_n$;
$\quad \alpha_n = (1 + \boldsymbol{a}_n^H \boldsymbol{d}_n)^{-1}$;
$\quad \boldsymbol{\Sigma}_n^{-1} = \boldsymbol{\Sigma}_n^{-1} - \alpha_n \boldsymbol{d}_n \boldsymbol{d}_n^H$;
$\quad \hat{\boldsymbol{\theta}}_n[i{+}1] = \hat{\boldsymbol{\theta}}_{n-1}[i{+}1] + \alpha_n(y_n - \boldsymbol{a}_n^H \boldsymbol{\theta}_{n-1}[i{+}1])\boldsymbol{d}_n$;
end

Complexity: $\mathcal{O}(N(N_H N_b)^2)$.

Numerical Experiments — Single-carrier:

Channel:

- WSSUS Rayleigh (via Jakes), $N_h = 3$ taps, $f_{\mathsf{D}} T_s = 0.002$.
  (e.g., f_c =60GHz, BW=1MHz, 36km/hr, τ_h =3μs)

Transmitter:

- rate-$\frac{1}{2}$ LDPC, 4096-bit frame, QPSK ($Q = 2$)
- block length: $N = 64$
- $N_p = 6$ pilots at start of each block.

Receiver:

- BEM: Karhunen Loève with $N_b = 3$
- EM iterations: $K = 3$, tree-search parameter $M = 64$
- LDPC decoding iterations: ≤ 60
- turbo iterations: ≤ 8

Description of Curves:

The proposed EM algorithm with K iterations is denoted "**(cT+sBE)**^{K}" since it iterates *coherent tree-search* (**cT**) with *soft BEM estimation* (**sBE**).

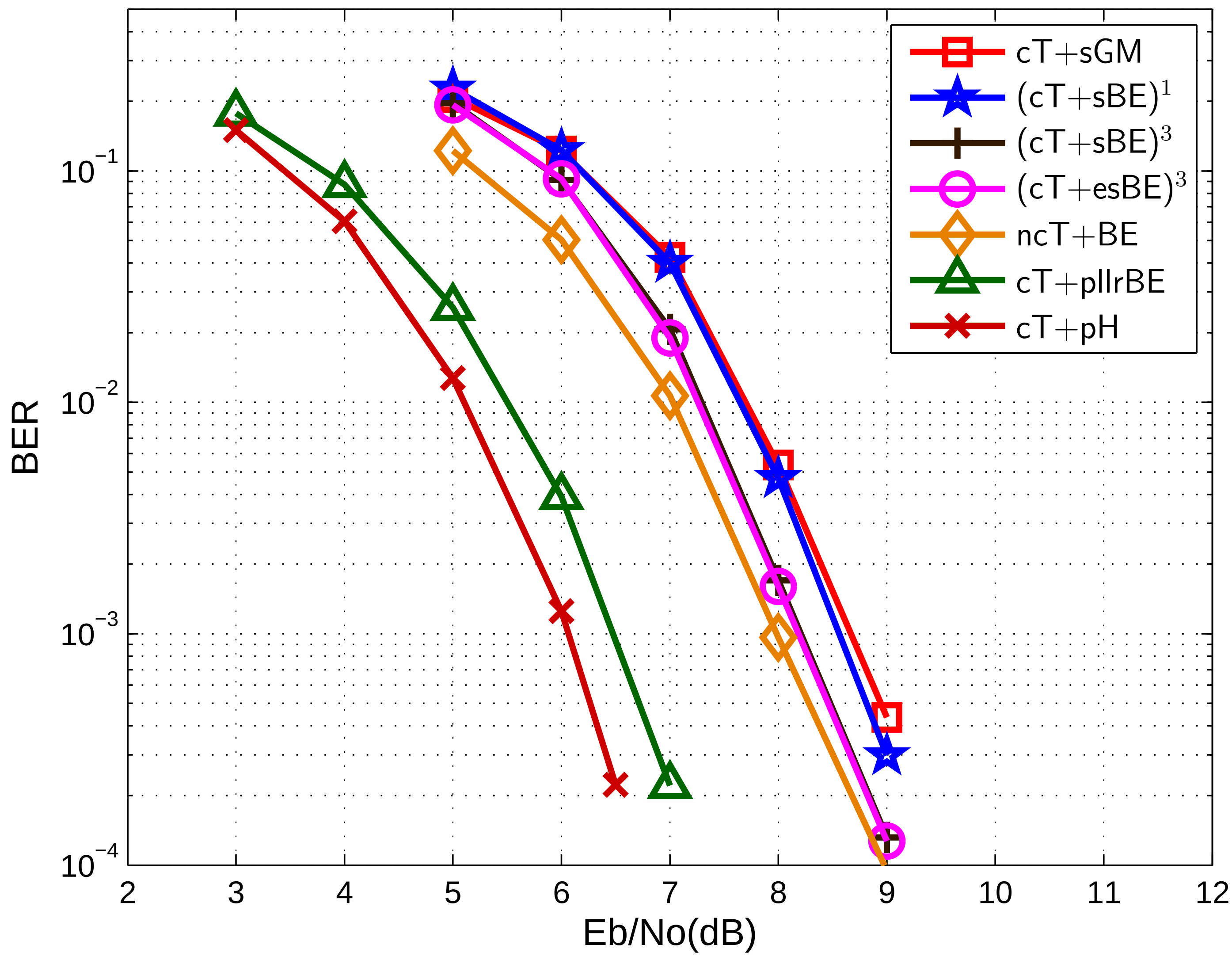
The two genie-aided bounds are: coherent tree search with *perfect knowledge of* $\boldsymbol{H}$ (**cT+pH**), and soft BEM estimation using *perfect LLR feedback from the decoder* (**cT+pLLRBE**). *Only about 2 dB better!*

The conventional technique uses *soft 2nd-order Gauss-Markov channel estimation* (**sGM**). Here we combine this with coherent tree search.

An approximate MAP-optimal approach is our *non-coherent tree search* (**ncT**) from Asilomar-07. Generally, it is more computationally complex.

We also tried the EM algorithm with "exact" soft BEM estimation **(cT+esBE)**^{K} to show that it performs only slightly better than **(cT+sBE)**^{K}.

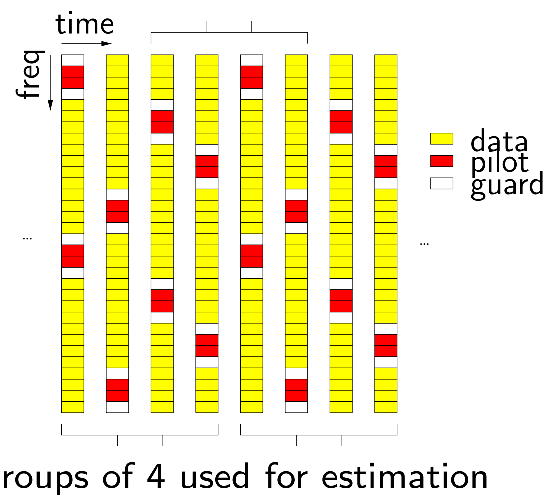
Single-carrier Performance



Numerical Experiments — Multi-carrier:

Channel:

- WSSUS Rayleigh (via Jakes), $N_h = 3$ taps, $f_{\mathsf{D}} T_s = 0.002$.
  (e.g., f_c =60GHz, BW=1MHz, 36 km/hr, τ_h =3μs)



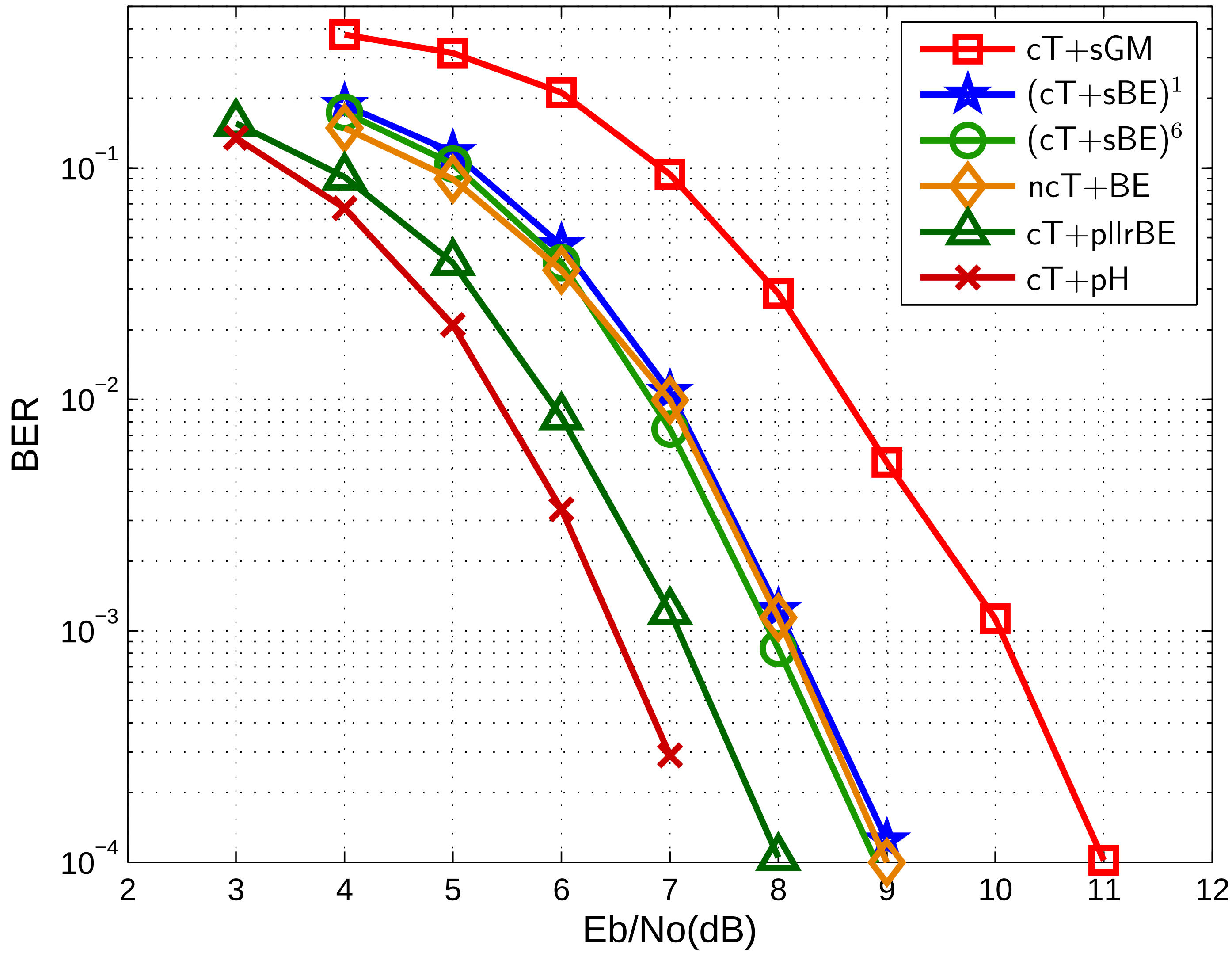
Transmitter:

- rate-$\frac{1}{2}$ LDPC, 4096-bit frame, QPSK
- $N = 64$ subcarriers
- $N_p = 9$ pilot subcarriers

Receiver:

- BEM: CE with $N_b = N_h$, ICI taps: $N_H = 3$
- EM iterations $K = 6$, tree-search parameter $M = 64$
- LDPC decoding iterations ≤ 60
- turbo iterations ≤ 8

Multi-carrier Performance



Conclusions:

- We proposed a novel soft noncoherent equalization algorithm based on the Bayesian EM algorithm.
- The algorithm alternates between two steps: soft MMSE estimation of BEM coefficients, and computation of (coherent) coded-bit LLRs.
- To calculate the LLRs, we proposed to use a (soft) tree search implemented via the M-algorithm.
- To calculate soft MMSE estimates, we presented two simplified algs:
  - an $\mathcal{O}\big(N(N_H N_b)^2\big)$ algorithm based on sequential Bayes,
  - an $\mathcal{O}\big(N \log_2 N\big)$ algorithm based on the conjugate gradient algorithm and FFT; applicable only in the multi-carrier case.
- The EM-based soft noncoherent equalizer performs only ≈ 2dB away from genie-aided bounds.