Analysis of Semiblind Channel Estimation for MIMO Systems

Tianbin Wo, Peter Adam Hoeher\textsuperscript{1}
Ansgar Scherb, Karl-Dirk Kammeyer\textsuperscript{2}

\textsuperscript{1} Information and Coding Theory Lab (ICT),
University of Kiel, Germany
\{wtb,ph\}@tf.uni-kiel.de

\textsuperscript{2} Department of Communications Engineering (ANT),
University of Bremen, Germany
\{scherb,kammeyer\}@ant.uni-bremen.de

Autumn School, Paris, France
October 24–28, 2005
1. Motivation

2. Channel Model

3. Semiblind Channel Estimation

4. Cramer-Rao Lower Bound

5. Conclusion
Training-based channel estimation (TBCE)

channel estimation quality \(\leftrightarrow\) trade-off \(\rightarrow\) system throughput
(bit error rate) \(\rightarrow\) (bandwidth efficiency)

Blind system identification / channel equalization

system throughput maximized but \(\rightarrow\) ambiguity (permutation, phase)

Semiblind channel estimation (SBCE)

training-based channel estimation + blind techniques \(\Rightarrow\) achieve high system throughput without sacrificing robustness
Motivation

▶ Training-based channel estimation (TBCE)

channel estimation quality \(\leftrightarrow\) system throughput (bit error rate) \(\rightarrow\) (bandwidth efficiency)

▶ Blind system identification / channel equalization

system throughput maximized \(\rightarrow\) ambiguity (permutation, phase) and local minima \(\rightarrow\) system throughput maximized but
Motivation

- Training-based channel estimation (TBCE)

  channel estimation quality \leftrightarrow \text{trade-off} \rightarrow \text{system throughput}
  \begin{align*}
  \text{(bit error rate)} \quad \text{(bandwidth efficiency)}
  \end{align*}

- Blind system identification / channel equalization

  system throughput maximized \rightarrow \text{ambiguity (permutation, phase)}
  \text{and local minima}

- Semiblind channel estimation (SBCE)

  training-based channel estimation + blind techniques

  \rightarrow \text{achieve high system throughput without sacrificing robustness}
Equivalent discrete-time channel model:

\[
y_i[k] = \sum_{j=1}^{NT} h_{ij}[k] \cdot x_j[k] + n_i[k],
\]

with \( 1 \leq i \leq NR \)

Within a channel coherence interval, channel model can be rewritten in matrix form:

\[
Y = H \cdot X + N
\]
Task of the Receiver

- Task: Given the observation $Y = HX + N$ and the knowledge of training, $X$ shall be detected.

- Solution 1: estimate $H$ first by training, then detect $X$.

Solution 1
Task of the Receiver

- Task: Given the observation $Y = HX + N$ and the knowledge of training, $X$ shall be detected.

- Solution 1: estimate $H$ first by training, then detect $X$.

- Solution 2: estimate $H$ and $X$ iteratively.

Solution 1

Solution 2
### One burst: $K_T$ training symbols, $K_I$ data symbols, $K = K_T + K_I$
Semiblind Channel Estimation

Algorithm

Iterative Channel Estimation and Data Detection

<table>
<thead>
<tr>
<th>Tx-1:</th>
<th>0 0 1 0 0 0 1 1</th>
<th>0 0 1 0 0 0 1 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial CE</td>
<td>Training ((K_T) symbols)</td>
<td></td>
</tr>
<tr>
<td>Tx-2:</td>
<td>1 1 1 0 0 1 0 0</td>
<td>1 1 1 0 0 1 0 0</td>
</tr>
</tbody>
</table>

1 Burst: \(K\) symbols

Algorithm:

1. Initial channel estimation by using the training only
2. * Given channel knowledge, perform data detection
   * Given data decisions, perform channel estimation by taking the whole burst as a virtual training
3. Repeat step 2 until a certain stopping criterion is reached
Semiblind Channel Estimation Algorithm

Iterative Channel Estimation and Data Detection

<table>
<thead>
<tr>
<th>Tx-1:</th>
<th>Initial CE</th>
</tr>
</thead>
<tbody>
<tr>
<td>...................</td>
<td>0 0 1 0 0 0 1 1</td>
</tr>
<tr>
<td>Training (K symbols)</td>
<td></td>
</tr>
<tr>
<td>Tx-2:</td>
<td>1 1 1 0 0 1 0 0</td>
</tr>
<tr>
<td>...................</td>
<td></td>
</tr>
</tbody>
</table>

1 Burst: K symbols

ML-SBCE:

1. $i = 0$: $\mathbf{H}_0 = \mathbf{Y}_T \cdot \mathbf{X}_T^\dagger$ ("$i$" denotes the iteration index)

2. $i = i + 1$
   
   $\mathbf{X}_i = \arg \min_{\mathbf{X} \in \mathcal{X}} \| \mathbf{Y} - \mathbf{H}_{i-1} \mathbf{X} \|_F^2$  
   
   "ML data detection"

   $\mathbf{H}_i = \mathbf{Y} \cdot \mathbf{X}_i^\dagger$

   "LS channel estimation"

3. Repeat 2 until $(\mathbf{H}_i, \mathbf{X}_i) = (\mathbf{H}_{i-1}, \mathbf{X}_{i-1})$
Semiblind Channel Estimation
Performance

- TBCE: training-based CE,
  SBCE: semiblind CE
- MSE curves are linear
  at high and low SNR
- Crossover SNR range
- longer burst,
  larger performance improvement

TBCE: training-based CE,
SBCE: semiblind CE

MSE curves are linear
at high and low SNR

Crossover SNR range

longer burst,
larger performance improvement

$N_R = N_T = 4$, $K_T = 4$
Semiblind Channel Estimation Performance

- TBCE: training-based CE, SBCE: semiblind CE
- MSE curves are linear at high and low SNR
- Crossover SNR range
- longer burst, larger performance improvement
- TBCE: training-based CE, SBCE: semiblind CE
- MSE curves are linear at high and low SNR
- Crossover SNR range
- longer burst, larger performance improvement
Semiblind Channel Estimation

Performance

- TBCE: training-based CE, SBCE: semiblind CE
- MSE curves are linear at high and low SNR
- Crossover SNR range
- longer burst, larger performance improvement

\[ N_R = N_T = 4, \ K_T = 4 \]

\[ E_s / N_0 \text{ in dB} \]

\[ \text{Mean Squared Error} \]

\[ \text{TBCE, SBCE: } K = 200 \]

17 dB

\[ N_R = N_T = 4, \ K_T = 4 \]
Cramer-Rao Lower Bound (CRLB)

- Assuming ML-SBCE is unbiased
  - its performance will be limited by the Cramer-Rao lower bound
Assuming ML-SBCE is unbiased
→ its performance will be limited by the Cramer-Rao lower bound

Cramer-Rao lower bound

\[ C_{\hat{H}} - F^{-1} \geq 0, \quad \text{where} \quad C_{\hat{H}} = \mathbb{E}\{\text{vec}(H - \hat{H})\text{vec}(H - \hat{H})^H\} \]
Assuming ML-SBCE is unbiased
→ its performance will be limited by the Cramer-Rao lower bound

Cramer-Rao lower bound

\[ C_{\hat{H}} - F^{-1} \geq 0, \quad \text{where} \quad C_{\hat{H}} = E\{\text{vec}(H - \hat{H})\text{vec}(H - \hat{H})^H\} \]

Fisher information matrix is given by

\[
F = \frac{1}{|A|\sigma_n^2} \int_{\mathbb{C}^{N_R \times K}} \sum_{\mathbf{X}_i, \mathbf{X}_j \in A} p(\mathbf{Y}|H, \mathbf{X}_i)p(\mathbf{Y}|H, \mathbf{X}_j) \sum_{\mathbf{X}_i \in A} p(\mathbf{Y}|H, \mathbf{X}_i) \left( (\mathbf{X}_i \otimes I_{N_R}) \mathbf{d}_i^* \mathbf{d}_j^T (\mathbf{X}_j^H \otimes I_{N_R}) \right) d\mathbf{Y}
\]

with

\[ \mathbf{d}_i = \text{vec} \left\{ \frac{\mathbf{Y} - H\mathbf{X}_i}{\sigma_n} \right\} \]
Minimum distance of noiseless channel outputs

\[ \delta_{\text{min}} = \min_{X_i, X_j \in X} \| H(X_i - X_j) \|_F^2, \quad i \neq j \]
Minimum distance of noiseless channel outputs

\[ \delta_{\text{min}} = \min_{X_i, X_j \in \mathcal{X}} \| H(X_i - X_j) \|_F^2, \quad i \neq j \]

For high SNR values where \( \delta_{\text{min}} \gg \sigma^2_n \), we will obtain

\[ F_{\text{high}} = \frac{K}{\sigma^2_n} I_{NRNT} \]
Cramer-Rao Lower Bound
Asymptotic Approximations

- Minimum distance of noiseless channel outputs

\[ \delta_{\text{min}} = \min_{X_i, X_j \in \mathcal{X}} \|H(X_i - X_j)\|_F^2, \quad i \neq j \]

- For high SNR values where \( \delta_{\text{min}} \gg \sigma_n^2 \), we will obtain

\[ F_{\text{high}} = \frac{K}{\sigma_n^2} \mathbf{I}_{N_R N_T} \]

- For low SNR values where \( \delta_{\text{min}} \ll \sigma_n^2 \), we will obtain

\[ F_{\text{low}} = \frac{K_T}{\sigma_n^2} \mathbf{I}_{N_R N_T} \]
Cramer-Rao Lower Bound
Asymptotic Approximations

- Minimum distance of noiseless channel outputs

\[ \delta_{\text{min}} = \min_{X_i, X_j \in \mathcal{X}} \| \mathbf{H}(X_i - X_j) \|_F^2, \quad i \neq j \]

- For high SNR values where \( \delta_{\text{min}} \gg \sigma_n^2 \), we will obtain

\[
\text{CRLB: MSE} \geq \text{tr} \left\{ \mathbf{F}^{-1}_{\text{high}} \right\} = \frac{N_R N_T}{K} \sigma_n^2
\]

- For low SNR values where \( \delta_{\text{min}} \ll \sigma_n^2 \), we will obtain

\[
\text{CRLB: MSE} \geq \text{tr} \left\{ \mathbf{F}^{-1}_{\text{low}} \right\} = \frac{N_R N_T}{K_T} \sigma_n^2
\]
Cramer-Rao Lower Bound

Simulation Results

- **CRLB at low SNR**
  \[ MSE \geq \frac{N_R N_T}{K_T} \sigma_n^2 \]
  → only training symbols are used for channel estimation

- **CRLB at high SNR**
  \[ MSE \geq \frac{N_R N_T}{K} \sigma_n^2 \]
  → all data symbols are known and used for channel estimation
Least squares channel estimator:

\[
\hat{H} = YX^H (XX^H)^{-1} = (HX + N) X^H (XX^H)^{-1}
\]

It is easy to find that \(E\{\hat{H}\} = H\), i.e. the estimator is unbiased.
Semiblind Channel Estimation

Semiblind channel estimator:

$$\hat{H} = Y \hat{X}^H (\hat{X} \hat{X}^H)^{-1} = (HX + N) \hat{X}^H (\hat{X} \hat{X}^H)^{-1}$$

In case of $\hat{X} \neq X$, do we still have $E\{\hat{H}\} = H$?
Semiblind Channel Estimation
Biasing

Semiblind channel estimator:

\[
\hat{H} = Y\hat{X}^H (\hat{X}\hat{X}^H)^{-1} = (HX + N)\hat{X}^H (\hat{X}\hat{X}^H)^{-1}
\]

In case of \( \hat{X} \neq X \), do we still have \( \mathbb{E}\{\hat{H}\} = H \) ?

Given large burst length, following approximations are valid:

\[
\hat{X}\hat{X}^H \approx K\mathbf{I}_{N_T}
\]

\[
\mathbb{E}\{X\hat{X}\} \approx (K_T + (1 - 2P_s)K_1)\mathbf{I}_{N_T}
\]
Semiblind Channel Estimation
Biasing

Semiblind channel estimator:

\[ \hat{H} = Y \hat{X}^H (\hat{X} \hat{X}^H)^{-1} = (HX + N) \hat{X}^H (\hat{X} \hat{X}^H)^{-1} \]

In case of \( \hat{X} \neq X \), do we still have \( E\{\hat{H}\} = H \)?

Given large burst length, following approximations are valid:

\[ \hat{X} \hat{X}^H \approx K I_{N_T} \]
\[ E\{X \hat{X}\} \approx (K_T + (1 - 2P_s)K_1)I_{N_T} \]

Let \( E = \hat{X} - X \) denote the detection error, we will obtain

\[ E\{\hat{H}\} \approx \frac{K_T + (1 - 2P_s)K_1}{K} H + \frac{1}{K} E\{NE^H\} \]
Degree of biasing:

\[ \text{BIAS} = \| H - \mathbb{E}\{ \hat{H} \} \|^2_F \]
Degree of biasing:

\[ \text{BIAS} = \| H - \mathbb{E}\{\hat{H}\} \|_F^2 \]

Noise-error cross correlation:

\[ \text{NEC} = \| \frac{1}{K} \mathbb{E}\{\text{NE}^H\} \|_F^2 \]
Semiblind Channel Estimation
Biasing – Simulation Results

Degree of biasing:
\[ \text{BIAS} \equiv \|H - \mathbb{E}\{\hat{H}\}\|_F^2 \]

Noise-error cross correlation:
\[ \text{NEC} \equiv \left\| \frac{1}{K} \mathbb{E}\{N\mathbf{E}^H\} \right\|_F^2 \]
Semiblind Channel Estimation
Mean Squared Error – Analytical Approximation

Definition:

\[ \text{MSE} = E \left\{ \| \hat{H} - H \|^2_F \right\} \]

Assumption:

\[ \hat{X} \hat{X}^H \approx K I_N \]

Analytical approximation:

\[ \text{MSE} \approx \left( 2 P_s K I \right)^2 \frac{1}{2} \left( \alpha - 4 P_s K \right) \]

\[ + \frac{1}{2} K^2 E \left\{ \| \hat{N} \hat{X}^H \|^2_F \right\} \]
Semiblind Channel Estimation
Mean Squared Error – Analytical Approximation

Definition:

\[
\text{MSE} = \mathbb{E}\left\{ \| \hat{H} - H \|^2_F \right\}
\]

Assumption:

\[
\begin{align*}
\hat{X}\hat{X}^H & \approx KI_{N_T} \\
X\hat{X}^H & \approx (K_T + (1 - 2P_s)K_1)I_{N_T}
\end{align*}
\]
Semiblind Channel Estimation
Mean Squared Error – Analytical Approximation

Definition:

\[ \text{MSE} = \mathbb{E}\{ \| \hat{H} - H \|^2_F \} \]

Assumption:

\[ \hat{X}\hat{X}^H \approx KI_{NT} \]
\[ X\hat{X}^H \approx (K_T + (1 - 2P_s)K_1)I_{NT} \]

Analytical approximation:

\[ \text{MSE} \approx \left( \frac{2P_sK_1}{K} \right)^2 \| H \|_F^2 - \frac{4P_sK_1}{K^2} \text{tr} \left\{ \Re \{ \mathbb{E}\{ N\hat{E}^H \} H^H \} \right\} + \frac{1}{K^2} \mathbb{E}\{ \| N\hat{X}^H \|_F^2 \} \]
- MSE-Anal. = $\alpha - \beta + \gamma$

- The analytical expression coincides with the true MSE at reasonable SNR values.
Semiblind Channel Estimation
Mean Squared Error – Simulation Results

- **MSE-Anal.** = $\alpha - \beta + \gamma$

- The analytical expression coincides with the true MSE at reasonable SNR values.

- **Observation:** $\alpha \equiv \beta$
  
  $\alpha = \left( \frac{2P_s K_1}{K} \right)^2 \text{tr}\{\Re\{HH^H\}\}$
  
  $\beta = \frac{4P_s K_1}{K^2} \text{tr}\{\Re\{E\{NE^H\}H^H\}\}$
  
  ⇒ pattern of noise-error correlation is determined by the channel matrix

- **Conclusion:** $\text{MSE} \approx \gamma = \frac{1}{K} \|X_H\|_2^2$
MSE-Anal. = \( \alpha - \beta + \gamma \)

The analytical expression coincides with the true MSE at reasonable SNR values.

Observation: \( \alpha \equiv \beta \)
\( \Rightarrow \) pattern of noise-error correlation is determined by the channel matrix

Conclusion:
\[
\text{MSE} \approx \gamma = \frac{1}{K^2} \mathbb{E}\left\{ \| \tilde{N} \hat{X}^H \|_F^2 \right\}
\]
Semiblind Channel Estimation
Mean Squared Error – Simulation Results

- MSE-Anal. = $\alpha - \beta + \gamma$

- The analytical expression coincides with the true MSE at reasonable SNR values.

- Observation: $\alpha \equiv \beta$
  $\Rightarrow$ pattern of noise-error correlation is determined by the channel matrix

- Conclusion:
  \[
  \text{MSE} \approx \gamma = \frac{1}{K^2} \mathbb{E} \{ \| \hat{X}^H \|^2_F \} \\
  = \frac{1}{K^2} \mathbb{E} \{ \| NX^H + NE^H \|^2_F \} = \frac{N_R N_T}{K} \sigma_n^2 + \frac{1}{K^2} \mathbb{E} \{ \| NE^H \|^2_F \}
  \]
  
  “CRLB at high SNR”

Faculty of Engineering
University of Kiel

Information and Coding
Theory Lab
Semiblind Channel Estimation
Mean Squared Error – Simulation Results

- MSE-Anal. = $\alpha - \beta + \gamma$

- The analytical expression coincides with the true MSE at reasonable SNR values.

- Observation: $\alpha \equiv \beta$
  $\Rightarrow$ pattern of noise-error correlation is determined by the channel matrix

- Conclusion:
  \[
  \text{MSE} \approx \gamma = \frac{1}{K^2} E \left\{ \|N\hat{X}^H\|_F^2 \right\} = \frac{1}{K^2} E \left\{ \|NX^H + NE^H\|_F^2 \right\} = \frac{N_R N_T}{K} \sigma_n^2 + \frac{1}{K^2} E \left\{ \|NE^H\|_F^2 \right\}
  \]
  "CRLB at high SNR"

Faculty of Engineering
University of Kiel

Information and Coding Theory Lab
Approximation of CRLB at high SNR fits for all cases.

As $N_R$ increases, the cross-over SNR range shifts to the left.

Reason:

$$\min_{i \neq j} \left\{ \| H\tilde{x}_i - H\tilde{x}_j \|_F^2 \right\}$$ increases,

→ higher capability of error correction

For $N_R < N_T$, the reduction of noise-error correlation may be of practical interests.
SBCE is effective in improving channel estimation quality

ML-SBCE is biased at low SNR and tends to be unbiased at high SNR
Reasons of biasing: Erroneous data detection
Noise-error correlation

ML-SBCE is well bounded by CRLB at high SNR, while may exceed CRLB at low SNR due to biasing

Semi-analytically: $\text{MSE} \approx \frac{1}{K^2} \mathbb{E}\{\|NE^H\|^2\}$

As $N_R$ increases, the cross-over range shifts to lower SNRs
$\rightarrow N_R$ plays an important role in the performance of SBCE

Further improvements: biasing compensation, MSE reduction
Appendix-1: Cramer-Rao Lower Bound

Cross-Over Range

\[ N_R = N_T = 4, \ K_T = 4, \ K = 200 \]

- CRLB at low SNR
- MSE of SBCE
- CRLB at high SNR

\[ \text{Mean Squared Error} \]

\[ \text{E}_s/N_0 \text{ in dB} \]
Appendix-1: Cramer-Rao Lower Bound
Cross-Over Range

\[ N_R = N_T = 4, K_T = 4, K = 200 \]

\[
F = \frac{1}{|\mathcal{A}|\sigma_n^2} \int_{\mathbb{C}^{NR \times K}} \frac{\sum_{X_i, X_j \in \mathcal{A}} p(Y|H, X_i)p(Y|H, X_j)}{\sum_{X_i \in \mathcal{A}} p(Y|H, X_i)} (X_i \otimes I_{NR}) d_i^* d_j^T (X_j^H \otimes I_{NR}) dY
\]
The SNR range where $F$ transits from $F_{\text{high}}$ to $F_{\text{low}}$

\[ \alpha \cdot \delta_{\min} \leq \sigma_n^2 \leq \beta \cdot \delta_{\min} \]
Appendix-2: Cramer-Rao Lower Bound
Saturation Point

- The SNR range where $F$ transits from $F_{\text{high}}$ to $F_{\text{low}}$

$$\alpha \cdot \delta_{\text{min}} \leq \sigma_n^2 \leq \beta \cdot \delta_{\text{min}}$$

- Open issue:

$$\alpha = ?$$
$$\beta = ?$$
The SNR range where $\mathbf{F}$ transits from $\mathbf{F}_{\text{high}}$ to $\mathbf{F}_{\text{low}}$

$$\alpha \cdot \delta_{\text{min}} \leq \sigma_n^2 \leq \beta \cdot \delta_{\text{min}}$$

Open issue:

$$\alpha = \ ?$$

$$\beta = \ ?$$

Minimum channel output distance

$$\delta_{\text{min}} = 4N_T \int_0^\infty \delta^{N_R} e^{-\delta} \left[ \Gamma(N_R, \delta) \right]^{N_T-1} [(N_R - 1)!]^{N_T} d\delta$$
Appendix-2: Cramer-Rao Lower Bound
Saturation Point

- The SNR range where $F$ transits from $F_{\text{high}}$ to $F_{\text{low}}$

$$\alpha \cdot \delta_{\text{min}} \leq \sigma_n^2 \leq \beta \cdot \delta_{\text{min}}$$

- Open issue:
  \begin{align*}
    \alpha & = ? \\
    \beta & = ?
  \end{align*}

- Minimum channel output distance

$$\delta_{\text{min}} = 4N_T \int_{0}^{\infty} \delta^{N_R} e^{-\delta} \left[ \Gamma(N_R, \delta) \right]^{N_T-1} \frac{d\delta}{[(N_R - 1)!]^{N_T}}$$

- Observation: as $N_R$ increases, $\delta_{\text{min}}$ increases

  $\rightarrow$ saturation point shifts to the leftside

<table>
<thead>
<tr>
<th>Number of receive antennas</th>
<th>Minimum Squared Distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Simul. $\delta_{\text{min}}$</td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
</tr>
</tbody>
</table>

$N_T = 2$

- Graph showing $\delta_{\text{min}}$ simulation and analysis for different numbers of receive antennas.