Analysis of Semiblind Channel Estimation for MIMO Systems

Tianbin Wo, Peter Adam Hoeher¹ Ansgar Scherb, Karl-Dirk Kammeyer²

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

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 $\xleftarrow{}$ trade-off (bit error rate) system throughput (bandwidth efficiency)





Training-based channel estimation (TBCE)

channel estimation quality $\xleftarrow{}$ trade-off (bit error rate) system throughput (bandwidth efficiency)

Blind system identification / channel equalization

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





Training-based channel estimation (TBCE)

Blind system identification / channel equalization

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

Semiblind channel estimation (SBCE)
 training-based channel estimation + blind techniques

 ~ achieve high system throughput without sacrificing robustness





Channel Model

 Equivalent discrete-time channel model:

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

with 1 $\leq i \leq N_R$

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

$$\mathbf{Y} = \mathbf{H} \cdot \mathbf{X} + \mathbf{N}$$







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









• One burst: K_T training symbols, K_I data symbols, $K = K_T + K_I$







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



ML-SBCE:

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

2.
$$i = i + 1$$

* $\widehat{\mathbf{X}}_i = \arg \min_{\widetilde{\mathbf{X}} \in \mathcal{X}} \{ \|\mathbf{Y} - \widehat{\mathbf{H}}_{i-1} \widetilde{\mathbf{X}}\|_{\mathsf{F}}^2 \}$ "ML data detection"
* $\widehat{\mathbf{H}}_i = \mathbf{Y} \cdot \widehat{\mathbf{X}}_i^{\dagger}$ "LS channel estimation"

3. Repeat 2 until
$$\left(\widehat{\mathbf{H}}_{i}, \widehat{\mathbf{X}}_{i}\right) = \left(\widehat{\mathbf{H}}_{i-1}, \widehat{\mathbf{X}}_{i-1}\right)$$





- 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







- 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





Cramer-Rao Lower Bound (CRLB)

Assuming ML-SBCE is unbiased

 \rightarrow its performance will be limited by the Cramer-Rao lower bound





Cramer-Rao Lower Bound (CRLB)

Assuming ML-SBCE is unbiased

 \rightarrow its performance will be limited by the Cramer-Rao lower bound

Cramer-Rao lower bound

$$\mathbf{C}_{\widehat{\mathbf{H}}} - \mathbf{F}^{-1} \geqslant \mathbf{0}, \quad \text{where} \quad \mathbf{C}_{\widehat{\mathbf{H}}} = \mathsf{E} \big\{ \mathsf{vec} \{ \mathbf{H} - \widehat{\mathbf{H}} \} \mathsf{vec} \{ \mathbf{H} - \widehat{\mathbf{H}} \}^H \big\}$$





Cramer-Rao Lower Bound (CRLB)

Assuming ML-SBCE is unbiased

 \rightarrow its performance will be limited by the Cramer-Rao lower bound

Cramer-Rao lower bound

$$\mathbf{C}_{\widehat{\mathbf{H}}} - \mathbf{F}^{-1} \geqslant \mathbf{0}, \quad \text{where} \quad \mathbf{C}_{\widehat{\mathbf{H}}} = \mathsf{E} \big\{ \mathsf{vec} \{ \mathbf{H} - \widehat{\mathbf{H}} \} \mathsf{vec} \{ \mathbf{H} - \widehat{\mathbf{H}} \}^H \big\}$$

Fisher information matrix is given by

$$\mathbf{F} = \frac{1}{|\mathcal{A}|\sigma_n^2} \int_{\mathbb{C}^{N_R \times K}} \frac{\sum_{\mathbf{X}_i, \mathbf{X}_j \in \mathcal{A}} p(\mathbf{Y}|\mathbf{H}, \mathbf{X}_i) p(\mathbf{Y}|\mathbf{H}, \mathbf{X}_j)}{\sum_{\mathbf{X}_i \in \mathcal{A}} p(\mathbf{Y}|\mathbf{H}, \mathbf{X}_i)} (\mathbf{X}_i \otimes \mathbf{I}_{N_R}) \mathbf{d}_i^* \mathbf{d}_j^T (\mathbf{X}_j^H \otimes \mathbf{I}_{N_R}) \ d\mathbf{Y}$$
with

$$\underline{\mathbf{d}}_i = \mathsf{vec} \left\{ \frac{\mathbf{Y} - \mathbf{H} \mathbf{X}_i}{\sigma_n} \right\}$$



Minimum distance of noiseless channel outputs

$$\delta_{\min} = \min_{\mathbf{X}_i, \mathbf{X}_j \in \mathcal{X}} \left\| \mathbf{H} (\mathbf{X}_i - \mathbf{X}_j) \right\|_F^2, \quad i \neq j$$





Minimum distance of noiseless channel outputs

$$\delta_{\min} = \min_{\mathbf{X}_i, \mathbf{X}_j \in \mathcal{X}} \left\| \mathbf{H} (\mathbf{X}_i - \mathbf{X}_j) \right\|_F^2, \quad i \neq j$$

▶ For high SNR values where $\delta_{\min} \gg \sigma_n^2$, we will obtain

$$\mathbf{F}_{\mathsf{high}} = \frac{K}{\sigma_n^2} \mathbf{I}_{N_{\mathsf{R}}N_{\mathsf{T}}}$$





Cramer-Rao Lower Bound Asymptotic Approximations

Minimum distance of noiseless channel outputs

$$\delta_{\min} = \min_{\mathbf{X}_i, \mathbf{X}_j \in \mathcal{X}} \left\| \mathbf{H} (\mathbf{X}_i - \mathbf{X}_j) \right\|_F^2, \quad i \neq j$$

• For high SNR values where $\delta_{\min} \gg \sigma_n^2$, we will obtain

$$\mathbf{F}_{\mathsf{high}} = \frac{K}{\sigma_n^2} \mathbf{I}_{N_{\mathsf{R}}N_{\mathsf{T}}}$$

▶ For low SNR values where $\delta_{\min} \ll \sigma_n^2$, we will obtain

$$\mathbf{F}_{\mathsf{low}} = \frac{K_{\mathsf{T}}}{\sigma_n^2} \mathbf{I}_{N_{\mathsf{R}}N_{\mathsf{T}}}$$





Cramer-Rao Lower Bound Asymptotic Approximations

Minimum distance of noiseless channel outputs

$$\delta_{\min} = \min_{\mathbf{X}_i, \mathbf{X}_j \in \mathcal{X}} \left\| \mathbf{H} (\mathbf{X}_i - \mathbf{X}_j) \right\|_F^2, \quad i \neq j$$

• For high SNR values where $\delta_{\min} \gg \sigma_n^2$, we will obtain

CRLB: MSE
$$\geq$$
 tr $\left\{ \mathbf{F}_{\mathsf{high}}^{-1} \right\} = \frac{N_{\mathsf{R}}N_{\mathsf{T}}}{K}\sigma_n^2$

▶ For low SNR values where $\delta_{\min} \ll \sigma_n^2$, we will obtain

CRLB: MSE
$$\geq$$
 tr $\{\mathbf{F}_{low}^{-1}\}$ = $\frac{N_{\mathsf{R}}N_{\mathsf{T}}}{K_{\mathsf{T}}}\sigma_n^2$





CRLB at low SNR

$$MSE \geq \frac{N_{\mathsf{R}}N_{\mathsf{T}}}{K_{\mathsf{T}}}\sigma_n^2$$

- \rightarrow only training symbols are used for channel estimation
- CRLB at high SNR

$$MSE \geq \frac{N_{\mathsf{R}}N_{\mathsf{T}}}{K}\sigma_n^2$$

 \rightarrow all data symbols are known and used for channel estimation





Least squares channel estimator:

$$\widehat{\mathbf{H}} = \mathbf{Y} \mathbf{X}^{H} ig(\mathbf{X} \mathbf{X}^{H} ig)^{-1} = ig(\mathbf{H} \mathbf{X} + \mathbf{N} ig) \mathbf{X}^{H} ig(\mathbf{X} \mathbf{X}^{H} ig)^{-1}$$

It is easy to find that $\mathsf{E}\big\{\widehat{\mathbf{H}}\big\}=\mathbf{H},$ i.e. the estimator is unbiased





Semiblind channel estimator:

$$\widehat{\mathbf{H}} = \mathbf{Y} \widehat{\mathbf{X}}^{H} \big(\widehat{\mathbf{X}} \widehat{\mathbf{X}}^{H} \big)^{-1} = \big(\mathbf{H} \mathbf{X} + \mathbf{N} \big) \widehat{\mathbf{X}}^{H} \big(\widehat{\mathbf{X}} \widehat{\mathbf{X}}^{H} \big)^{-1}$$

In case of $\widehat{\mathbf{X}}\neq\mathbf{X},$ do we still have $\mathsf{E}\big\{\widehat{\mathbf{H}}\big\}=\mathbf{H}$?





Semiblind channel estimator:

$$\widehat{\mathbf{H}} = \mathbf{Y} \widehat{\mathbf{X}}^{H} \big(\widehat{\mathbf{X}} \widehat{\mathbf{X}}^{H} \big)^{-1} = \big(\mathbf{H} \mathbf{X} + \mathbf{N} \big) \widehat{\mathbf{X}}^{H} \big(\widehat{\mathbf{X}} \widehat{\mathbf{X}}^{H} \big)^{-1}$$

In case of $\widehat{\mathbf{X}}\neq\mathbf{X},$ do we still have $\mathsf{E}\big\{\widehat{\mathbf{H}}\big\}=\mathbf{H}$?

Given large burst length, following approximations are valid:

$$\begin{aligned} \widehat{\mathbf{X}} \widehat{\mathbf{X}}^{H} &\approx K \mathbf{I}_{N_{\mathsf{T}}} \\ \mathsf{E} \big\{ \mathbf{X} \widehat{\mathbf{X}} \big\} &\approx (K_{\mathsf{T}} + (1 - 2P_{s})K_{\mathsf{I}}) \mathbf{I}_{N_{\mathsf{T}}} \end{aligned}$$





Semiblind channel estimator:

$$\widehat{\mathbf{H}} = \mathbf{Y} \widehat{\mathbf{X}}^{H} \big(\widehat{\mathbf{X}} \widehat{\mathbf{X}}^{H} \big)^{-1} = \big(\mathbf{H} \mathbf{X} + \mathbf{N} \big) \widehat{\mathbf{X}}^{H} \big(\widehat{\mathbf{X}} \widehat{\mathbf{X}}^{H} \big)^{-1}$$

In case of $\widehat{\mathbf{X}}\neq\mathbf{X},$ do we still have $\mathsf{E}\big\{\widehat{\mathbf{H}}\big\}=\mathbf{H}$?

Given large burst length, following approximations are valid:

$$\begin{aligned} & \widehat{\mathbf{X}} \widehat{\mathbf{X}}^{H} \approx K \mathbf{I}_{N_{\mathsf{T}}} \\ & \mathsf{E} \big\{ \mathbf{X} \widehat{\mathbf{X}} \big\} \approx (K_{\mathsf{T}} + (1 - 2P_{s})K_{\mathsf{I}}) \mathbf{I}_{N_{\mathsf{T}}} \end{aligned}$$

 \blacktriangleright Let $\mathbf{E}=\widehat{\mathbf{X}}-\mathbf{X}$ denote the detection error, we will obtain

$$\mathsf{E}\left\{\widehat{\mathbf{H}}\right\} \approx \frac{K_{\mathsf{T}} + (1 - 2P_s)K_{\mathsf{I}}}{K}\mathbf{H} + \frac{1}{K}\mathsf{E}\left\{\mathbf{N}\mathbf{E}^H\right\}$$





Biasing – Simulation Results

► Degree of biasing: $\mathbf{BIAS} \doteq \|\mathbf{H} - \mathsf{E}\{\widehat{\mathbf{H}}\}\|_{F}^{2}$



I C T

Biasing – Simulation Results

- Degree of biasing: $\mathbf{BIAS} \doteq \|\mathbf{H} - \mathsf{E}\{\widehat{\mathbf{H}}\}\|_{F}^{2}$
- Noise-error cross correlation: $\mathbf{NEC} \doteq \left\| \frac{1}{K} \mathsf{E} \{ \mathbf{NE}^H \} \right\|_F^2$





Biasing – Simulation Results

- ► Degree of biasing: $\mathbf{BIAS} \doteq \|\mathbf{H} - \mathsf{E}\{\widehat{\mathbf{H}}\}\|_{F}^{2}$
- Noise-error cross correlation: $\mathbf{NEC} \doteq \left\| \frac{1}{K} \mathsf{E} \{ \mathbf{NE}^H \} \right\|_F^2$



г с т

Mean Squared Error – Analytical Approximation

Definition:

$$\mathsf{MSE} = \mathsf{E} \left\{ \| \widehat{\mathbf{H}} - \mathbf{H} \|_{\mathsf{F}}^2 \right\}$$





Mean Squared Error – Analytical Approximation

Definition:

$$\mathsf{MSE} = \mathsf{E} \big\{ \| \widehat{\mathbf{H}} - \mathbf{H} \|_{\mathsf{F}}^2 \big\}$$

Assumption:

$$\begin{split} \widehat{\mathbf{X}} \widehat{\mathbf{X}}^{H} &\approx K \mathbf{I}_{N_{\mathsf{T}}} \ \mathbf{X} \widehat{\mathbf{X}}^{H} &\approx (K_{\mathsf{T}} + (1 - 2P_{s})K_{\mathsf{I}}) \mathbf{I}_{N_{\mathsf{T}}} \end{split}$$





Mean Squared Error – Analytical Approximation

Definition:

$$\mathsf{MSE} = \mathsf{E} \big\{ \| \widehat{\mathbf{H}} - \mathbf{H} \|_{\mathsf{F}}^2 \big\}$$

Assumption:

$$\begin{split} \widehat{\mathbf{X}} \widehat{\mathbf{X}}^H &pprox K \mathbf{I}_{N_{\mathsf{T}}} \ \mathbf{X} \widehat{\mathbf{X}}^H &pprox (K_{\mathsf{T}} + (1 - 2P_s)K_{\mathsf{I}}) \mathbf{I}_{N_{\mathsf{T}}} \end{split}$$

Analytical approximation:

$$\mathsf{MSE} \approx \underbrace{\left(\frac{2P_sK_{\mathsf{I}}}{K}\right)^2 \|\mathbf{H}\|_{\mathsf{F}}^2}_{\alpha} - \underbrace{\frac{4P_sK_{\mathsf{I}}}{K^2}\mathsf{tr}\Big\{\Re\big\{\mathsf{E}\big\{\mathbf{NE}^H\big\}\mathbf{H}^H\big\}\Big\}}_{\beta} + \underbrace{\frac{1}{K^2}\mathsf{E}\big\{\|\mathbf{N}\widehat{\mathbf{X}}^H\|_{\mathsf{F}}^2\big\}}_{\gamma}$$





Mean Squared Error - Simulation Results

- MSE-Anal. = $\alpha \beta + \gamma$
- The analytical expression coincides with the true MSE at reasonable SNR values.





Mean Squared Error - Simulation Results

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

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

• Observation:
$$\alpha \equiv \beta$$

 $\alpha = \left(\frac{2P_sK_1}{K}\right)^2 \operatorname{tr}\left\{\Re\left\{\mathbf{H}\mathbf{H}^H\right\}\right\}$
 $\beta = \frac{4P_sK_1}{K^2}\operatorname{tr}\left\{\Re\left\{\mathbf{E}\left\{\mathbf{N}\mathbf{E}^H\right\}\mathbf{H}^H\right\}\right\}$
 \Rightarrow pattern of noise-error correlation

is determined by the channel matrix





IC T



14

Mean Squared Error - Simulation Results

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

- The analytical expression coincides with the true MSE at reasonable SNR values.
- Observation: α ≡ β
 ⇒ pattern of noise-error correlation is determined by the channel matrix
- Conclusion: $MSE \approx \gamma = \frac{1}{K^2} E\{\|\mathbf{N}\widehat{\mathbf{X}}^H\|_{\mathsf{F}}^2\}$





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:



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



"CRLB at high SNR"



Conclusion:



Mean Squared Error w.r.t. $N_{\rm R}$

- Approximation of CRLB at high SNR fits for all cases
- As N_R increases, the cross-over SNR range shifts to the left
 - Reason: $\min_{i \neq j} \left\{ \left\| \mathbf{H} \widetilde{\mathbf{X}}_i - \mathbf{H} \widetilde{\mathbf{X}}_j \right\|_F^2 \right\} \text{ increases,}$
 - \rightarrow higher capability of error correction
- For $N_{\mathsf{R}} < N_{\mathsf{T}}$,

the reduction of noise-error correlation may be of practical interests







Conclusion

- 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:
$$MSE \approx \frac{1}{K^2} E\{ \| \mathbf{NE}^H \|_F^2 \}$$

- ▶ As N_R increases, the cross-over range shifts to lower SNRs → 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







Appendix-1: Cramer-Rao Lower Bound Cross-Over Range



$$\mathbf{F} = \frac{1}{|\mathcal{A}|\sigma_n^2} \int_{\mathbb{C}^{N_R \times K}} \frac{\mathbf{x}_i, \mathbf{x}_j \in \mathcal{A}}{p(\mathbf{Y}|\mathbf{H}, \mathbf{X}_i) p(\mathbf{Y}|\mathbf{H}, \mathbf{X}_j)}}{\sum\limits_{\mathbf{X}_i \in \mathcal{A}} p(\mathbf{Y}|\mathbf{H}, \mathbf{X}_i)} (\mathbf{X}_i \otimes \mathbf{I}_{N_R}) \underline{\mathbf{d}}_i^* \underline{\mathbf{d}}_j^T (\mathbf{X}_j^H \otimes \mathbf{I}_{N_R}) \ d\mathbf{Y}$$





The SNR range where F transits from F_{high} to F_{low}

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



ПСТ

The SNR range where F transits from F_{high} to F_{low}

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

Open issue:

 $\begin{array}{rcl} \alpha & = & ? \\ \beta & = & ? \end{array}$



ПСТ

The SNR range where F transits from F_{high} to F_{low}

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

Open issue:

 $\begin{array}{rcl} \alpha & = & ? \\ \beta & = & ? \end{array}$

Minimum channel output distance

$$\delta_{\min} = 4N_{\mathrm{T}} \int_{0}^{\infty} \frac{\delta^{N_{\mathrm{R}}} e^{-\delta} \left[\Gamma(N_{\mathrm{R}}, \delta) \right]^{N_{\mathrm{T}}-1}}{\left[(N_{\mathrm{R}} - 1)! \right]^{N_{\mathrm{T}}}} d\delta$$







The SNR range where F transits from F_{high} to F_{low}

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

Open issue:

- $\begin{array}{rcl} \alpha & = & ? \\ \beta & = & ? \end{array}$
- Minimum channel output distance

$$\delta_{\min} = 4N_{\mathrm{T}} \int_{0}^{\infty} \frac{\delta^{N_{\mathrm{R}}} e^{-\delta} \left[\Gamma(N_{\mathrm{R}}, \delta) \right]^{N_{\mathrm{T}}-1}}{\left[(N_{\mathrm{R}} - 1)! \right]^{N_{\mathrm{T}}}} d\delta$$

• Observation: as N_R increases,

 δ_{min} increases

 \rightarrow saturation point shifts to the leftside





