Iterative Channel and Data Estimation: Framework and Analysis via Replica Method

Mikko Vehkaperä\textsuperscript{1}, Keigo Takeuchi\textsuperscript{2}, Ralf Müller\textsuperscript{1}, Toshiyuki Tanaka\textsuperscript{3}

\textsuperscript{1}Department of Electronics and Telecommunications, Norwegian University of Science and Technology
\textsuperscript{2}Department of Information and Communication Engineering, University of Electro-Communications, Tokyo, Japan.
\textsuperscript{3}Graduate School of Informatics, Kyoto University, Japan

July 3, 2009
Introduction
- Considered System at Glance
- Goals and Contributions

Mathematical Methods and Analysis
- Bayesian Inference and System Model
- Deriving a Class of Linear Data Estimators
- The Replica Method and Decoupling Principle

Results
- Theoretical Results
- Numerical Examples and Conclusions
Considered Uplink CDMA System

\[ P_1, X_1, H_1, S_1 \]

\[ P_K, X_K, H_K, S_K \]

Channel estimation

\[ Y, \{\langle H_k \rangle, \text{MSE}_k \} \]

Multiuser detection & decoding (MUD)

\[ \{\langle X_k \rangle_{\text{APP}}, \text{MSE}_k \} \]
Considered Uplink CDMA System

\[ P_1, X_1 \rightarrow H_1, S_1 \]

\[ \vdots \]

\[ P_K, X_K \rightarrow H_K, S_K \]

Channel estimation

\[ \gamma, \{\langle H_k \rangle, \text{MSE}_k \} \rightarrow \{\langle X_k \rangle_{\text{APP}}, \text{MSE}_k \} \rightarrow \text{Multiuser detection \& decoding (MUDD)} \]

Data estimation

\[ \gamma, \{\langle H_k \rangle, \text{MSE}_k \} \rightarrow \langle X_1 \rangle_{\text{EXT}} \rightarrow \{Z_1, \text{DEC \# 1} \}
\]

\[ \vdots \]

\[ Z_K \rightarrow \{Z_K, \text{DEC \# K} \} \rightarrow \langle X_K \rangle_{\text{EXT}} \]
Main Results

Large system analysis of iterative receivers with information feedback:

Decoupling result for an iterative receiver consisting of generic linear channel and data estimators utilizing soft / hard feedback.
Main Results

**Large system analysis** of iterative receivers with information feedback:

Decoupling result for an iterative receiver consisting of generic linear channel and data estimators utilizing soft / hard feedback. As corollaries we obtain, for example:

1. The **multiuser efficiency** of linear minimum mean square error (LMMSE) and single-user matched filter (SUMF) based iterative multiuser decoders (MUDD) under imperfect channel information.
Large system analysis of iterative receivers with information feedback:

Decoupling result for an iterative receiver consisting of generic linear channel and data estimators utilizing soft / hard feedback. As corollaries we obtain, for example:

1. The multiuser efficiency of linear minimum mean square error (LMMSE) and single-user matched filter (SUMF) based iterative multiuser decoders (MUDD) under imperfect channel information.

2. The mean square error of the iterative LMMSE channel estimator using training symbols and soft feedback from the MUDD.
Main Results

**Large system analysis** of iterative receivers with information feedback:

Decoupling result for an iterative receiver consisting of generic linear channel and data estimators utilizing soft / hard feedback. As corollaries we obtain, for example:

1. **The multiuser efficiency** of linear minimum mean square error (LMMSE) and single-user matched filter (SUMF) based iterative multiuser decoders (MUDD) under imperfect channel information.

2. **The mean square error** of the iterative LMMSE channel estimator using training symbols and soft feedback from the MUDD.

3. **MSE of the hard feedback** based linear channel estimator studied approximately by Li, Betz & Poor (Trans. SP., 2007).
Main Results

**Large system analysis** of iterative receivers with information feedback:

Decoupling result for an iterative receiver consisting of generic linear channel and data estimators utilizing soft / hard feedback. As corollaries we obtain, for example:

1. **The multiuser efficiency** of linear minimum mean square error (LMMSE) and single-user matched filter (SUMF) based iterative multiuser decoders (MUDD) under imperfect channel information.

2. **The mean square error** of the iterative LMMSE channel estimator using training symbols and soft feedback from the MUDD.

3. **MSE of the hard feedback** based linear channel estimator studied approximately by Li, Betz & Poor (Trans. SP., 2007).

Outline

1 Introduction
   - Considered System at Glance
   - Goals and Contributions

2 Mathematical Methods and Analysis
   - Bayesian Inference and System Model
   - Deriving a Class of Linear Data Estimators
   - The Replica Method and Decoupling Principle

3 Results
   - Theoretical Results
   - Numerical Examples and Conclusions
On Bayesian Inference

Let $x \sim P(x)$ be the assumed probability of RV $x \sim P(x)$. Given observation $y$ and the postulated distribution $Q(\tilde{x}, Q(y|\tilde{x}, \tilde{V}))$, the Generalized Posterior Mean Estimator (GPME) of $x$ reads

$$\langle x \rangle = \int \tilde{x} Q(d\tilde{x}|y; \tilde{V});$$

where $P$ denotes for the true probabilities and $Q$ for the assumed ones. In general, non-linear estimator unless $Q(\tilde{x}, y|\tilde{V})$ is Gaussian. Selecting appropriately the measures $Q$, GPME gives as special cases, e.g., linear MMSE and SUMF estimators.
On Bayesian Inference

Generalized Posterior Mean Estimator (GPME)

Let $Q(\tilde{x})$ be the assumed probability of RV $x \sim P(x)$. Given observation $y$ and the postulated distribution $Q(y|\tilde{x}, \tilde{\nu})$, the GPME of $x$ reads

$$\langle x \rangle = \int \tilde{x} Q(d\tilde{x}|y, \tilde{\nu}),$$

where $P$ denotes for the true probabilities and $Q$ for the assumed ones.
On Bayesian Inference

Generalized Posterior Mean Estimator (GPME)

Let $Q(\tilde{x})$ be the assumed probability of RV $x \sim P(x)$. Given observation $y$ and the postulated distribution $Q(y|\tilde{x}, \tilde{V})$, the GPME of $x$ reads

$$\langle x \rangle = \int \tilde{x} Q(d\tilde{x}|y, \tilde{V}),$$

where $P$ denotes for the true probabilities and $Q$ for the assumed ones.

- In general, non-linear estimator unless $Q(\tilde{x}, y|\tilde{V})$ is Gaussian.
- Selecting appropriately the measures $Q$, GPME gives as special cases, e.g., linear MMSE and SUMF estimators.
Synchronous Random Uplink CDMA

\[ y_n[c] = S \times H[c] \times u_n[c] + w_n[c] \]

\[
\begin{align*}
S & = L \times KM \\
H[c] & = KM \times K \\
u_n[c] & = K \times 1 \\
w_n[c] & = L \times 1
\end{align*}
\]
Synchronous Random Uplink CDMA

\[ y_n[c] = SH[c]u_n[c] + w_n[c] \]

IID

No ISI

\[ h_k[c] \text{ not known} \]

QPSK

\[ ... \]
Synchronous Random Uplink CDMA

\[
y_n[c] = S\tilde{H}[c]\tilde{u}_n[c] + \tilde{w}_n[c]
\]

\[
y_n[c] = SH[c]u_n[c] + w_n[c]
\]
Linear Data Estimation Under CSI Mismatch — (1)

\[ y_n = S(\langle H \rangle^{(i)} + \Delta H^{(i)}) x_n + w_n \]

\[ \Delta H^{(i)} = H - \langle H \rangle^{(i)} \]
Linear Data Estimation Under CSI Mismatch — (1)

LMMSE estimate

\[ y_n = S(\langle H \rangle^{(i)} + \Delta H^{(i)})x_n + w_n \]

\[ \Delta H^{(i)} = H - \langle H \rangle^{(i)} \]

uncorrelated

\[ y_n = S\langle H \rangle^{(i)} \tilde{x}_n + S\Delta \tilde{H}^{(i)} \tilde{x}_n + \tilde{w}_n \]

\[ Q(\tilde{h}^{(i)}_k) = \text{CN}(\langle h \rangle^{(i)}, \tilde{\Sigma}_{\Delta h_k}^{(i)}) \]
Linear Data Estimation Under CSI Mismatch — (1)

\[ y_n = S \langle H \rangle^{(i)} x_n + w_n \]

LMMSE estimate

\[ \Delta H^{(i)} = H - \langle H \rangle^{(i)} \]

uncorrelated

\[ y_n = S \langle H \rangle^{(i)} \tilde{x}_n + S \Delta \tilde{H}^{(i)} \tilde{x}_n + w_n \]

\[ Q(\tilde{h}_k) = CN(\langle h \rangle^{(i)}, \Sigma^{(i)}_{\Delta h_k}) \]

\[ Q(\tilde{x}_n) = CN(0, I_K) \]

\[ \langle x_n \rangle = \int \tilde{x}_n dQ(\tilde{x}_n|y_n, S, \langle H \rangle^{(i)}) \]
Linear Data Estimation Under CSI Mismatch — (1)

\[
\Delta H^{(i)} = H - \langle H \rangle^{(i)}
\]

\[
y_n = S\langle H \rangle^{(i)}\tilde{x}_n + S\Delta H^{(i)}\tilde{x}_n + w_n
\]

\[
\langle x_n \rangle = \int \tilde{x}_n dQ(\tilde{x}_n | y_n, S, \langle H \rangle^{(i)})
\]

\[
Q(h_k) = CN(\langle h \rangle^{(i)}, \Sigma_{\Delta h_k}^{(i)})
\]

\[
Q(\tilde{x}_n) = CN(0, I_K)
\]

non-linear estimator.
Data Estimation Under CSI Mismatch — (2)

\[ y_n = S\langle H\rangle^{(i)} x_n + S\Delta v_n^{(i)} + w_n \]

\[ \Delta v_n^{(i)} = \Delta H^{(i)} x_n \]
Data Estimation Under CSI Mismatch — (2)

\[ y_n = S\langle H\rangle^{(i)} x_n + S\Delta \tilde{v}_n^{(i)} + w_n \]

LMMSE estimate

\[ \Delta v_n^{(i)} = \Delta H^{(i)} x_n \]

uncorrelated

\[ \mathbb{Q}(\Delta \tilde{v}_n^{(i)}) = \mathcal{CN}(0, \Sigma_{\Delta h_k}^{(i)}) \]

\[ y_n = S\langle H\rangle^{(i)} \tilde{x}_n + S\Delta \tilde{v}_n^{(i)} + \tilde{w}_n \]
Data Estimation Under CSI Mismatch — (2)

\[ y_n = S\langle H\rangle^{(i)} x_n + S\Delta v_n^{(i)} + w_n \]

\[ \Delta v_n^{(i)} = \Delta H^{(i)} x_n \]

\[ y_n = S\langle H\rangle^{(i)} \tilde{x}_n + S\Delta \tilde{v}_n^{(i)} + \tilde{w}_n \]

\[ Q(\Delta \tilde{v}_n^{(i)}) = \text{CN}(0, \tilde{\Sigma}_{\Delta h_k}^{(i)}) \]

\[ Q(\tilde{x}_n) = \text{CN}(0, I_K) \]

\[ \langle x_n \rangle = \int \tilde{x}_n dQ(\tilde{x}_n|y_n, S, \langle H\rangle^{(i)}) \]
Data Estimation Under CSI Mismatch — (2)

\[
y_n = S\langle H \rangle^{(i)} x_n + S\Delta v_n^{(i)} + w_n
\]

LMMSE estimate

\[
\Delta v_n^{(i)} = \Delta H^{(i)} x_n
\]

uncorrelated

\[
y_n = S\langle H \rangle^{(i)} \tilde{x}_n + S\Delta \tilde{v}_n^{(i)} + \tilde{w}_n
\]

\[
Q(\Delta \tilde{v}_n^{(i)}) = \text{CN}(0, \Sigma_{\Delta h_k}^{(i)})
\]

\[
Q(\tilde{x}_n) = \text{CN}(0, I_K)
\]

\[
\langle x_n \rangle = \int \tilde{x}_n dQ(\tilde{x}_n | y_n, S, \langle H \rangle^{(i)})
\]

... linear estimator.
The following is all you need to know (yes, this is crude):

- The **replica method** is a set of mathematical tricks and assumptions to evaluate quantities of the form \( \lim_{K \to \infty} E \{ K^{-1} \log Z \} \).
The Replica Method...

The following is all you need to know (yes, this is crude):

- **The replica method** is a set of mathematical tricks and assumptions to evaluate quantities of the form \( \lim_{K \to \infty} \mathbb{E} \{ K^{-1} \log Z \} \).
- **To derive the decoupling principle**, shown in simplified form on the next few slides, requires exactly this.
  - The same tools apply on a very general class of CDMA systems with GPME based receivers (see also our Physcomnet’09 paper).
The Replica Method...

The following is all you need to know (yes, this is crude):

- The **replica method** is a set of mathematical tricks and assumptions to evaluate quantities of the form $\lim_{K \to \infty} \mathbb{E} \left\{ K^{-1} \log Z \right\}$.
- To derive the **decoupling principle**, shown in simplified form on the next few slides, requires exactly this.
  - The same tools apply on a very general class of CDMA systems with GPME based receivers (see also our Physcomnet’09 paper).

**What you really need to know about the replica method:**

Replica method is a complicated tool to transform certain types of intractable problems into solvable ones (under certain conditions).
Decoupling Principle — (1)

\[ y = \sum_k s_k h_k x_k + w \]

MULTIUSER CHANNEL

\[ x_1 \sim P_x \]

\[ \vdots \]

\[ x_K \sim P_x \]

Channel load: \( \alpha = K/L \).

RX

Generalized Posterior Mean Estimator

\[ \{ EQ(x_k|y, I) \} \]

\[ I = \{ \{ h_k \}, \{ s_k \}, \sigma^2 \} \]
Decoupling Principle — (1)

MULTIUSER CHANNEL

\[
\begin{align*}
\mathbf{x}_1 & \sim P \mathbf{x} \\
\vdots & \\
\mathbf{x}_K & \sim P \mathbf{x} \\
\end{align*}
\]

Channel load: \( \alpha = K/L \).

\[
y = \sum_k s_k h_k \mathbf{x}_k + \mathbf{w}
\]

\[
h_k \sim \text{CN}(0, \text{SNR}) \\
\mathbf{w} \sim \text{CN}(0, I_L)
\]

Spreading factor.

RX

Generalized Posterior Mean Estimator

\[
\left\{ \mathbb{E}_Q(\mathbf{x}_k | \mathbf{y}, \mathcal{I}) \right\}
\]

\[
\mathcal{I} = \left\{ \{ h_k \}, \{ s_k \}, \tilde{\sigma}^2 \right\}
\]

Channel load: \( \alpha = K/L \).
Decoupling Principle — (1)

Let $K = \alpha L \to \infty$.

MULTIUSER CHANNEL

$\mathbf{x}_1 \sim P_x$

$\vdots$

$\mathbf{x}_K \sim P_x$

Channel load: $\alpha = K/L$.

$\mathbf{y} = \sum_k s_k h_k \mathbf{x}_k + \mathbf{w}$

RX

Generalized Posterior Mean Estimator

$\{E_Q(x_k | \mathbf{y}, \mathcal{I})\}$

$\mathcal{I} = \{\{h_k\}, \{s_k\}, \sigma^2\}$

Iterative Channel and Data Estimation
Decoupling Principle — (2)

\[ z_1 = h_1 x_1 + \frac{1}{\sqrt{\eta}} w_1 \]

\[ z_K = h_K x_K + \frac{1}{\sqrt{\eta}} w_K \]

\[ w_k \sim \text{CN}(0, 1) \]
Outline

1. Introduction
   - Considered System at Glance
   - Goals and Contributions

2. Mathematical Methods and Analysis
   - Bayesian Inference and System Model
   - Deriving a Class of Linear Data Estimators
   - The Replica Method and Decoupling Principle

3. Results
   - Theoretical Results
   - Numerical Examples and Conclusions
Revisit the System Model

\[ P_1, X_1 \rightarrow H_1, S_1 \]

\[ \vdots \]

\[ P_K, X_K \rightarrow H_K, S_K \]

\[ W \rightarrow \mathcal{Y} \rightarrow \text{Channel estimation} \]

\[ \mathcal{Y}, \{\langle H_k \rangle, \text{MSE}_k \} \rightarrow \text{Multiuser detection (MUD)} \]

\[ \{\langle X_k \rangle_{\text{APP}}, \text{MSE}_k \} \]

\[ \mathcal{Y}, \{\langle H_k \rangle, \text{MSE}_k \} \rightarrow \text{Data estimation} \]

\[ \langle X_1 \rangle_{\text{EXT}} \rightarrow \text{DEC} \# 1 \rightarrow Z_1 \rightarrow \text{DEC} \# K \rightarrow Z_K \rightarrow \langle X_K \rangle_{\text{EXT}} \]
Main Results: Decoupling — (1)

Multiuser efficiency of iterative MUD with CSI mismatch

The problem of finding the SINR of the iterative MUD with feedback decouples into single-user, flat fading single-input multiple-output (SIMO) data estimation problem in the large system limit $K = \alpha L \rightarrow \infty$. 
Main Results: Decoupling — (1)

Multiuser efficiency of iterative MUD with CSI mismatch

The problem of finding the SINR of the iterative MUD with feedback decouples into single-user, flat fading single-input multiple-output (SIMO) data estimation problem in the large system limit $K = \alpha L \to \infty$.

MSE of linear channel estimator with feedback

The problem of finding the per-path MSE of the iterative channel estimator with feedback decouples into single-user, flat fading channel estimation problem in the large system limit $K = \alpha L \to \infty$. 
Consequence of the decoupling results

From the performance point of view, instead of studying an iterative multiuser CDMA system over a multipath fading channel, **we can study an iterative single-user system over a flat fading channel.**
Consequence of the decoupling results

From the performance point of view, instead of studying an iterative multiuser CDMA system over a multipath fading channel, we can study an iterative single-user system over a flat fading channel.

- The replica method gives us all the parameters that describe the statistics of the decoupled single-user system.
Main Results: What Does This Mean? — (2)

Consequence of the decoupling results

From the performance point of view, instead of studying an iterative multiuser CDMA system over a multipath fading channel, we can study an iterative single-user system over a flat fading channel.

- The replica method gives us all the parameters that describe the statistics of the decoupled single-user system.

E.g., the noise variances for LMMSE channel estimation with soft feedback

\[
\begin{align*}
\frac{1}{C_p^{(i)}} &= \sigma^2 + \frac{\alpha M \bar{p} \bar{M}}{1 + \frac{\bar{p}}{M} \left( \tau_p C_p^{(i)} + \tau_d C_d^{(i)} \right) E \left\{ |\langle x_{1n} \rangle_{\text{app}}|^2 (1 + C_d^{(i)} \frac{\bar{p}}{M} \Sigma_{\Delta x}^{\text{app}})^{-1} \right\}} \\
\frac{1}{C_d^{(i)}} &= \sigma^2 + \alpha ME \left\{ (\bar{p} / M) \Sigma_{\Delta x}^{\text{app}} / (1 + C_d^{(i)} (\bar{p} / M) \Sigma_{\Delta x}^{\text{app}}) \right\} \\
&\quad + \frac{\alpha M \bar{p} \bar{M} E \left\{ |\langle x_{1n} \rangle_{\text{app}}|^2 (1 + C_d^{(i)} \frac{\bar{p}}{M} \Sigma_{\Delta x}^{\text{app}})^{-2} \right\}}{1 + \frac{\bar{p}}{M} \left( \tau_p C_p^{(i)} + \tau_d C_d^{(i)} \right) E \left\{ |\langle x_{1n} \rangle_{\text{app}}|^2 (1 + C_d^{(i)} \frac{\bar{p}}{M} \Sigma_{\Delta x}^{\text{app}})^{-1} \right\}}
\end{align*}
\]
Three equal power paths ($M=3$), coherence time of $T = 100$ symbols, $\tau_p = 10$ training symbols, user load $\alpha = 1.2$, and average SNR of 4 dBs.
Three equal power paths (M=3), coherence time of $T = 100$ symbols, $\tau_p = 10$ training symbols, user load $\alpha = 1.2$, and average SNR of 4 dBs.
Three equal power paths (M=3), coherence time of $T = 100$ symbols, $\tau_p = 10$ training symbols, user load $\alpha = 1.2$, and average SNR of 4 dBs.
Bit Error Rate Performance

Three equal power paths (M=3), coherence time of $T = 100$ symbols, rate-1/2 convolutional code with generators $(753, 561)_8$ and Gray encoded QPSK.
Bit Error Rate Performance

Three equal power paths ($M=3$), coherence time of $T = 100$ symbols, rate-1/2 convolutional code with generators $(753, 561)_8$ and Gray encoded QPSK.
Bit Error Rate Performance

Three equal power paths (M=3), coherence time of $T = 100$ symbols, rate-1/2 convolutional code with generators $(753, 561)_8$ and Gray encoded QPSK.
Three equal power paths (M=3), coherence time of $T = 100$ symbols, rate-$1/2$ convolutional code with generators $(753, 561)_8$ and Gray encoded QPSK.
Analysis of uplink CDMA with iterative channel and data estimation

By utilizing the replica trick, we were able to perform large system analysis of synchronous uplink CDMA system utilizing iterative channel and data estimation.
Analysis of uplink CDMA with iterative channel and data estimation

By utilizing the replica trick, we were able to perform large system analysis of synchronous uplink CDMA system utilizing iterative channel and data estimation.

Hard vs. soft feedback

As with iterative data estimation, hard feedback can degrade the performance of linear channel estimator whereas this never happens with soft feedback.
Analysis of uplink CDMA with iterative channel and data estimation

By utilizing the replica trick, we were able to perform large system analysis of synchronous uplink CDMA system utilizing iterative channel and data estimation.

Hard vs. soft feedback

As with iterative data estimation, hard feedback can degrade the performance of linear channel estimator whereas this never happens with soft feedback.

Iterative channel estimation

Iterative channel estimation can provide near single user performance with (almost) negligible training overhead even for overloaded systems.