ISIT 2009, Seoul, Korea, June 28 - July 3, 2009

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

Mikko Vehkaperä*, Keigo Takeuchi†, Ralf R. Müller* and Toshiyuki Tanaka‡
*Depart. of Electr. and Telecomm., Norwegian University of Science and Technology, NO-7491 Trondheim, Norway.
E-mail: {mikko, ralf}@iet.ntnu.no
†Depart. of Inform. and Commun. Engineering, University of Electro-Communications, Tokyo 182-8585, Japan.
E-mail: takeuchi@ice.uec.ac.jp
‡Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan.
E-mail: tt@i.kyoto-u.ac.jp

***Abstract*—The large system analysis of a randomly spread direct-sequence code-division multiple-access system operating over a frequency-selective fading channel is considered. Iterative multiuser detection and decoding based on generalized posterior mean estimation and interference cancellation is assumed. The channel is mismatched and provided by a linear estimator whose initial pilot-based decisions are iteratively refined by using a feedback from the single-user decoders. By an application of the replica method, a tool from statistical physics, and density evolution with Gaussian approximation, we show that the performance metrics of the considered multiuser system converge in distribution at the large system limit to that of a simple single-user system operating over a flat fading channel. We also give the exact result of the hard decision feedback based channel estimator analyzed approximately by Li *et al.* (2007).**

I. Introduction

The large system analysis of a randomly spread direct-sequence code-division multiple-access system (DS-CDMA) operating over an uplink multipath fading channel is considered. Since straightforward joint decoding of the users is infeasible in practical systems and naive separation of detection and decoding reduces the system capacity significantly [1]–[5], we employ here the iterative multiuser detection and decoding (MUDD) framework proposed in [6]–[8]. The channel state information (CSI) is assumed to be mismatched and provided by a linear channel estimator. To reduce the pilot overhead and improve the reliability of the CSI, an information feedback from the single-user decoders is used to refine the initial pilot-based channel estimates. Although the use of iterative channel and data estimation has been proposed and numerically studied by several authors in the literature, the only effort to mathematically analyze the performance of such a receiver has been made to our knowledge by Li *et al.* [9].

Instead of pre-defining the estimators explicitly before the analysis (see, e.g., [1], [2], [7]–[9]), we consider here a more general Bayesian framework based on generalized posterior mean estimation (GPME) [5], [10], [11]. The class of estimators described by such a parametrized GPME encloses as special cases, e.g., the linear and non-linear minimum mean square error (MMSE) estimators. Our approach allows us also to consider both the hard and soft feedback schemes under the same unified Bayesian framework. The treatment of the CSI mismatch at the MUDD in this paper follows [2]. This differs slightly from [12], where some heuristic assumptions [12, Sec. 2.2] were made before the application of the replica method. Also the aforementioned analysis by Li *et al.* [9], where a linear hard feedback based "maximum likelihood"[1] channel estimator and single-user matched filter (SUMF) with hard interference cancellation were considered, makes assumptions similar to [12, Sec. 2.2]. The result obtained in [9] is also an approximation, even in the large system limit. As an example of our general result, we report the exact large system performance of the system studied in [9].

To assess the performance of the GPME-based iterative receiver described in Sections II and III, we combine the replica method, recently applied with great success to problems in telecommunications [5], [10]–[14], with the density evolution under Gaussian approximation (GA-DE) [7] to derive so-called "decoupling results" (see, e.g., [5], [10], [11]) in Section IV. We remark that, although the replica method is a standard tool in statistical physics, the assumptions made in the replica analysis are mostly heuristic and their formal justification is an open problem in mathematical physics [15] (see also [5], [13]). All proofs in this paper are omitted due to space constraints.

We write $\boldsymbol{x} \sim \mathbb{P}$ and $\tilde{\boldsymbol{x}} \sim \mathbb{Q}$ for a random vector (RV) with true $\mathbb{P}$ and postulated $\mathbb{Q}$ probability measures, respectively. For a proper complex Gaussian RV $\boldsymbol{x}$ with mean $\boldsymbol{\mu}_{\boldsymbol{x}} = \mathsf{E}\{\boldsymbol{x}\}$ and covariance $\boldsymbol{\Sigma}_{\boldsymbol{x}} = \mathrm{Cov}\{\boldsymbol{x}\} = \mathsf{E}\{(\boldsymbol{x}-\mathsf{E}\{\boldsymbol{x}\})(\boldsymbol{x}-\mathsf{E}\{\boldsymbol{x}\})^{\mathsf{H}}\}$, we write $\boldsymbol{x} \sim \mathsf{CN}(\boldsymbol{\mu}_{\boldsymbol{x}}, \boldsymbol{\Sigma}_{\boldsymbol{x}})$. For matrix $\boldsymbol{A} = [\boldsymbol{a}_1\ \boldsymbol{a}_2\ \cdots\ \boldsymbol{a}_N] \in \mathbb{C}^{M\times N}$, we define operator $\mathrm{vec}(\boldsymbol{A}) = [\boldsymbol{a}_1^{\mathsf{T}}\ \boldsymbol{a}_2^{\mathsf{T}}\ \cdots\ \boldsymbol{a}_N^{\mathsf{T}}]^{\mathsf{T}} \in \mathbb{C}^{MN}$, and let $\boldsymbol{D} = \mathrm{diag}(\boldsymbol{d})$ be the diagonal matrix defined by the vector $\boldsymbol{d} = [d_1 \cdots d_M]^{\mathsf{T}}$.

II. System Model and Assumptions

Consider a synchronous uplink DS-CDMA system operating over an M-path block fading channel. Let the delay spread of the channel be small compared to the symbol time so that we can neglect the effects of intersymbol interference. The transmission takes place over independent identically distributed (IID) fading blocks $c = 1, 2, \ldots, C$ having a coherence time of T symbols. For initial channel estimation, τ_{p} IID

[1] The estimator treats the hard estimates of the data symbols as pilots.

pilot symbols denoted by vector $\boldsymbol{p}_k[c] = \left[p_{k1}[c] \cdots p_{k\tau_{\mathsf{p}}}[c]\right]^{\mathsf{T}}$ are drawn uniformly for each user $k = 1, \ldots, K$ and block c from the quarternary phase shift keying (QPSK) signal set $\mathcal{M} = \{\pm\frac{1}{\sqrt{2}} \pm \frac{j}{\sqrt{2}}\}$. We assume that the same binary error correction codes are used by all users, but the random uniform bit-interleavers concatenated with them are independent. For the kth user we denote by $\mathcal{C}_k$ the composition of the code and the interleaver. The interleaved binary code word is modulated by using a symbol-by-symbol Gray mapping onto $\mathcal{M}$ and denoted by the vector $\boldsymbol{x}_k = \mathrm{vec}\big(\left[\boldsymbol{x}_k[1] \cdots \boldsymbol{x}_k[C]\right]\big) \in \mathcal{M}^N$, where $\boldsymbol{x}_k[c] = \left[x_{k1}[c] \cdots x_{k\tau_{\mathsf{d}}}[c]\right]^{\mathsf{T}}$ are the $\tau_{\mathsf{d}} = T - \tau_{\mathsf{p}}$ code symbols of the code word transmitted during the cth fading block. We also assume that due to random bit-interleaving $\mathsf{E}\{\boldsymbol{x}_k\} = \boldsymbol{0}$ and $\mathrm{Cov}\{\boldsymbol{x}_k\} = \boldsymbol{I}$ in the limit $N = \tau_{\mathsf{d}}C \to \infty$, and with τ_{d} fixed, the channel can be considered to be ergodic over the entire code word.

The nth received vector within the cth block reads [2]

$$\boldsymbol{y}_n[c] = \frac{1}{\sqrt{L}} \sum_{k=1}^{K} u_{kn}[c] \boldsymbol{S}_k \boldsymbol{h}_k[c] + \sigma \boldsymbol{w}_n[c], \qquad (1)$$

where $\boldsymbol{y}_n[c] = \left[y_{1n}[c] \cdots y_{Ln}[c]\right]^{\mathsf{T}}$, $\boldsymbol{w}_n[c] \sim \mathsf{CN}(\boldsymbol{0}, \boldsymbol{I})$, and

$$u_{kn}[c] = \begin{cases} p_{kn}[c], & n = 1, \ldots, \tau_{\mathsf{p}}, \\ x_{k(n-\tau_{\mathsf{p}})}[c], & n = \tau_{\mathsf{p}} + 1, \ldots, T. \end{cases}$$

For later use, we write $\mathcal{Y}_c = \{\boldsymbol{y}_n[c] \,|\, \forall n\}$, and let the set of all channel coefficients, pilot symbols and data symbols at cth fading block to be denoted by $\mathcal{H}_c = \{\boldsymbol{h}_k[c] \,|\, \forall k\}$, $\mathcal{P}_c = \{\boldsymbol{p}_k[c] \,|\, \forall k\}$ and $\mathcal{X}_c = \{\boldsymbol{x}_k[c] \,|\, \forall k\}$, respectively. The spreading matrix $\boldsymbol{S}_k = \left[\boldsymbol{s}_k^1 \cdots \boldsymbol{s}_k^M\right] = \left[\vec{\boldsymbol{s}}_{k1}^{\mathsf{T}} \cdots \vec{\boldsymbol{s}}_{kL}^{\mathsf{T}}\right]^{\mathsf{T}}$, where M is the number of multipaths, is modified without loss of generality [2, Thms. 3 and 4] to have IID entries with zero mean and unit variance for all $k = 1, \ldots, K$. The set of all spreading sequences is denoted by $\mathcal{S} = \{\boldsymbol{S}_k \,|\, \forall k\}$. The channel vectors $\boldsymbol{h}_k[c] = \left[h_k^1[c] \cdots h_k^M[c]\right]^{\mathsf{T}}$ are drawn according to $\mathbb{P}(\boldsymbol{h}_k[c]) = \mathsf{CN}(\boldsymbol{0}, \overline{\boldsymbol{P}}_k) \,\forall c$, where $\overline{\boldsymbol{P}}_k = \mathrm{diag}\big(\left[\overline{p}_k^1 \cdots \overline{p}_k^M\right]\big)$ contains the power delay profile of the kth users channel. The average received signal-to-noise ratio (SNR) is defined as $\overline{\gamma}_k = \overline{p}_k/\sigma^2$, where $\overline{p}_k = \mathrm{tr}(\overline{\boldsymbol{P}}_k)$.

Denote by $\tilde{\mathcal{H}}_c = \{\tilde{\boldsymbol{h}}_k[c] = \left[\tilde{h}_k^1[c] \;\cdots\; \tilde{h}_k^M[c]\right]^{\mathsf{T}} \,|\, \forall k\}$ and $\tilde{\mathcal{X}}_c = \{\tilde{\boldsymbol{x}}_k[c] = \left[\tilde{x}_{k1}[c] \;\cdots\; \tilde{x}_{k\tau_{\mathsf{d}}}[c]\right]^{\mathsf{T}} \,|\, \forall k\}$ the set of postulated channel coefficients and data symbols for cth fading block, and let $\mathcal{I}_c = \{\mathcal{P}_c, \mathcal{S}\} \,\forall c$ be information about the transmitted signal that is available to both the channel estimator and the MUDD. Assuming the first pass at the iterative receiver, the general procedure for iteratively updating the *a posteriori* probabilities (APPs) of the estimates at both the channel and the data estimator is described in Algorithm 1.

Assumption 1. For all iterations $i = 0, 1, \ldots$, we assume that:

1) For all $k = 1, \ldots, K$, $\mathbb{Q}(\tilde{\boldsymbol{h}}_k) = \mathbb{P}(\boldsymbol{h}_k)$ [2].
2) The distribution $\mathbb{Q}^{(i)}\big(\tilde{\mathcal{H}}_c | \mathcal{Y}_c, \mathcal{I}_c\big)$ is fully factorized in $\tilde{\mathcal{H}}_c$, with complex Gaussian marginals [2].
3) The mean and covariance of $\mathbb{P}_{\mathsf{app}}^{(i)}\big(\boldsymbol{x}_k\big)$ coincide with the ones obtained by using GA-DE [7], [8].

Algorithm 1 Iterative channel and data estimation.

1) Given $\{\mathbb{Q}(\tilde{\boldsymbol{h}}_k)\}$ and $\mathcal{I}_c = \{\mathcal{P}_c, \mathcal{S}\} \,\forall c$, the channel estimator calculates $\mathcal{Q}_{\mathcal{H}}^{(0)} = \{\mathbb{Q}^{(0)}\big(\tilde{\boldsymbol{h}}_k[c] | \mathcal{Y}_c, \mathcal{I}_c\big) \,|\, \forall k, c\}$ and sends it to MUDD.
2) The MUDD uses $\mathcal{Q}_{\mathcal{H}}^{(0)}$ and the postulated priors $\{\mathbb{Q}(\tilde{\boldsymbol{x}}_k)\}$ to obtain $\langle\langle \tilde{\mathcal{X}}_k \rangle\rangle_{(0)} = \{\langle\langle \tilde{x}_{kn} \rangle\rangle_{(0)} = \int \mathbb{Q}^{(0)}\big(\mathrm{d}\tilde{x}_{kn}[c] | \boldsymbol{y}_n[c], \mathcal{I}_c\big) \,|\, \forall n, c\}$.
3) Given $\langle\langle \tilde{\mathcal{X}}_k \rangle\rangle_{(0)}$ and the code constraints $\mathcal{C}_k$, each decoder calculates the approximate APPs $\mathcal{P}_{\mathcal{X}_k} = \{\mathbb{P}_{\mathsf{app}}^{(0)}(x_{kn}[c]) | \forall n, c\}$ of data symbols.
4) The MUDD applies an operator $F : \mathbb{P}(x | \cdot\,) \mapsto \mathbb{Q}(\tilde{x} | \cdot\,), x, \tilde{x} \in \mathbb{C}$, on the APPs $\{\mathcal{P}_{\mathcal{X}_k}\}$ obtained by the decoders and sends the transformed APPs $\mathcal{Q}_{\mathcal{X}_c}^{(0)} = \{\mathbb{Q}_{\mathsf{app}}^{(0)}(\tilde{x}_{kn}[c]) | \forall k, n\} \,\forall c$ to the channel estimator.
5) Channel estimator obtains $\mathcal{Q}_{\mathcal{H}_c}^{(1)} = \{\mathbb{Q}^{(1)}\big(\tilde{\boldsymbol{h}}_k[c] | \mathcal{Y}_c, \mathcal{I}_c\big) \,|\, \forall k, m\}$ given $\mathcal{Q}_{\mathcal{X}_c}^{(0)}$, and sends $\mathcal{Q}_{\mathcal{H}}^{(1)} = \{\mathcal{Q}_{\mathcal{H}_c}^{(1)} | \forall c\}$ to MUDD. Step 2) is repeated.

III. Channel And Data Estimation With Feedback

A. Linear Data Estimation With PIC

Let us consider the estimation of the data symbols $x_{1n}, n = 1, \ldots, \tau_{\mathsf{d}}$ of the first user at the ith iteration when parallel interference cancellation (PIC) and linear filtering is used. For notational convenience we drop the time index $n = 1, \ldots, \tau_{\mathsf{d}}$.

Let $\langle\langle \tilde{\boldsymbol{h}}_k \rangle\rangle_{(i)} = \left[\langle\langle \tilde{h}_k^1 \rangle\rangle_{(i)} \cdots \langle\langle \tilde{h}_k^M \rangle\rangle_{(i)}\right]^{\mathsf{T}}$ and $\boldsymbol{\Sigma}_{\tilde{\boldsymbol{h}}_k}^{(i)}$ be the mean and covariance of $\mathbb{Q}^{(i)}\big(\tilde{\boldsymbol{h}}_k[c] | \mathcal{Y}_c, \mathcal{I}_c\big)$ respectively. Denote by $\langle\langle \tilde{\boldsymbol{x}}_{\setminus 1} \rangle\rangle_{\mathsf{ext}}^{(i-1)} = \left[\langle\langle \tilde{x}_2 \rangle\rangle_{(i-1)} \cdots \langle\langle \tilde{x}_K \rangle\rangle_{(i-1)}\right]^{\mathsf{T}}$ and $\boldsymbol{\Sigma}_{\boldsymbol{x}_{\setminus 1}}^{(i-1)}$ the mean and covariance of $\mathbb{Q}_{\mathsf{ext}}^{(i-1)}(\tilde{\boldsymbol{x}}_{\setminus 1}) = \prod_{k=2}^{K} \mathbb{Q}_{\mathsf{ext}}^{(i-1)}(\tilde{x}_k)$, respectively. Analogous to Algorithm 1, we let $\mathbb{Q}_{\mathsf{ext}}^{(i-1)}\big(\tilde{x}_k\big)$ be the probabilities transformed by an operator F_{ext} from the *extrinsic* probabilities $\mathbb{P}_{\mathsf{ext}}^{(i-1)}\big(x_k\big)$ obtained by the sum-product algorithm [7]. Let $\mathcal{I}_{(i)} = \{\mathcal{I}, \langle\langle \tilde{\boldsymbol{h}}_k \rangle\rangle_{(i)}, \langle\langle \tilde{\boldsymbol{x}}_{\setminus 1} \rangle\rangle_{\mathsf{ext}}^{(i-1)}\}$ and re-write (1) as

$$\boldsymbol{y} = \frac{1}{\sqrt{L}} \sum_{k=1}^{K} \boldsymbol{S}_k \langle\langle \tilde{\boldsymbol{h}}_k \rangle\rangle_{(i)} x_k + \frac{1}{\sqrt{L}} \sum_{k=1}^{K} \boldsymbol{S}_k \Delta \boldsymbol{v}_k^{(i)} + \sigma \boldsymbol{w}, \qquad (2)$$

where $\Delta \boldsymbol{v}_k^{(i)} = \Delta \boldsymbol{h}_k^{(i)} x_k$ and $\Delta \boldsymbol{h}_k^{(i)} = \boldsymbol{h}_k - \langle\langle \tilde{\boldsymbol{h}}_k \rangle\rangle_{(i)}, \forall k$, and $x_k = \big(\langle\langle \tilde{x}_k \rangle\rangle_{\mathsf{ext}}^{(i-1)} + \Delta x_k^{(i-1)}\big)$, $\Delta x_k^{(i-1)} = x_k - \langle\langle \tilde{x}_k \rangle\rangle_{\mathsf{ext}}^{(i-1)}$ for $k = 2, \ldots, K$. We also define for later use a RV

$$\boldsymbol{\xi}_k^{(i)} = \begin{cases} \Delta \boldsymbol{h}_1^{(i)} x_1, & k = 1, \\ \Delta \boldsymbol{h}_k^{(i)} x_k + \langle\langle \tilde{\boldsymbol{h}}_k \rangle\rangle_{(i)} \Delta x_k^{(i-1)}, & k = 2, \ldots, K. \end{cases} \qquad (3)$$

Now, postulate in (2) a new noise variance $\tilde{\sigma}^2$, a Gaussian prior $\mathbb{Q}(\tilde{x}_1) = \mathsf{CN}(0, 1)$, $\mathbb{Q}\big(\tilde{\boldsymbol{x}}_{\setminus 1}\big) = \mathsf{CN}\big(\langle\langle \tilde{\boldsymbol{x}}_{\setminus 1} \rangle\rangle_{\mathsf{ext}}^{(i-1)}, \tilde{\boldsymbol{\Sigma}}_{\boldsymbol{x}_{\setminus 1}}^{(i-1)}\big)$, and assume that $\mathbb{Q}^{(i)}\big(\Delta \tilde{\boldsymbol{v}}_k | \mathcal{I}_{(i)}\big) = \mathsf{CN}(\boldsymbol{0}, \tilde{\boldsymbol{\Sigma}}_{\Delta \boldsymbol{v}_k}^{(i)})$. The resulting GPME gives arise to class of linear estimators

$$\langle\langle \tilde{x}_1 \rangle\rangle_{(i)} = \frac{\frac{1}{\sqrt{L}} \langle\langle \tilde{\boldsymbol{h}}_1 \rangle\rangle_{(i)}^{\mathsf{H}} \boldsymbol{S}_1^{\mathsf{H}} \tilde{\boldsymbol{\Sigma}}_{\mathsf{d}}^{-1}}{1 + \frac{1}{L} \langle\langle \tilde{\boldsymbol{h}}_1 \rangle\rangle_{(i)}^{\mathsf{H}} \boldsymbol{S}_1^{\mathsf{H}} \tilde{\boldsymbol{\Sigma}}_{\mathsf{d}}^{-1} \boldsymbol{S}_1 \langle\langle \tilde{\boldsymbol{h}}_1 \rangle\rangle_{(i)}} \tilde{\boldsymbol{y}}^{(i)}, \qquad (4)$$

where $\tilde{\boldsymbol{y}}^{(i)} = \boldsymbol{y} - \frac{1}{\sqrt{L}} \sum_{k=2}^{K} \boldsymbol{S}_k \langle\langle \tilde{\boldsymbol{h}}_k \rangle\rangle_{(i)} \langle\langle \tilde{x}_k \rangle\rangle_{\mathsf{ext}}^{(i-1)}$, $\tilde{\boldsymbol{\Sigma}}_{\mathsf{d}} = \frac{1}{L} \sum_{k=1}^{K} \boldsymbol{S}_k \tilde{\boldsymbol{\Sigma}}_{\boldsymbol{\xi}_k}^{(i)} \boldsymbol{S}_k^{\mathsf{H}} + \tilde{\sigma}^2 \boldsymbol{I}$, and (4) is parametrized by:

1) $F_{\mathsf{ext}} : \mathbb{P} \mapsto \mathbb{Q}$, defines the type of IC (e.g., hard / soft);
2) $\tilde{\boldsymbol{\Sigma}}_{\Delta \boldsymbol{v}_k}^{(i)}$, quantifies the estimator's knowledge about the error statistics of CSI;
3) $\tilde{\sigma}^2$, defines the type of linear filtering used.

Example 1. Assume $\tilde{\sigma}^2 = \sigma^2$, and let F_{ext} be the identity operator, i.e., $\mathbb{Q}^{(i-1)}_{\mathsf{ext}}(\tilde{x}_k) = \mathbb{P}^{(i-1)}_{\mathsf{ext}}(x_k)\ \forall k$. Furthermore, let $\tilde{\boldsymbol{\Sigma}}^{(i)}_{\boldsymbol{\xi}_k} = \boldsymbol{\Sigma}^{(i)}_{\boldsymbol{\xi}_k}$, where $\boldsymbol{\Sigma}^{(i)}_{\boldsymbol{\xi}_1} = \boldsymbol{\Sigma}^{(i)}_{\boldsymbol{h}_1}$, for $k = 1$ and $\boldsymbol{\Sigma}^{(i)}_{\boldsymbol{\xi}_k} = \langle\langle\tilde{\boldsymbol{h}}_k\rangle\rangle_{(i)}\langle\langle\tilde{\boldsymbol{h}}_k\rangle\rangle^{\mathsf{H}}_{(i)}[\boldsymbol{\Sigma}^{(i-1)}_{\boldsymbol{x}_{\setminus 1}}]_{kk} + \boldsymbol{\Sigma}^{(i)}_{\boldsymbol{h}_k}$, for $k = 2, \dots, K$. The GPME (4) reduces then to an extension of the LMMSE data estimator [2, Thm. 1] to include soft PIC.

Example 2. Let $\tilde{\sigma}^2 \to \infty$, and define F_{ext} to be the identity operator or $F_{\mathsf{ext}} : \mathbb{P}^{(i)}_{\mathsf{ext}}(x_k) \mapsto 1_{\tilde{x}_k}\big(\arg\min_{x_k \in \mathcal{M}} \mathbb{P}^{(i)}_{\mathsf{ext}}(x_k)\big)$. Then, the GPME (4) reduces respectively to the "SUMF-Based Soft IC" and the "Hard-IC" receivers, studied under the assumption of perfect CSI in [7, Prop. 2].

B. Linear Channel Estimation With Information Feedback

Let us drop the block index and concentrate on one fading block. Let $\langle\langle\tilde{\boldsymbol{x}}_k\rangle\rangle_{(i-1)}$ and $\boldsymbol{\Sigma}^{(i-1)}_{\boldsymbol{x}_k}$ be the true mean and covariance of $\mathbb{Q}^{(i-1)}_{\mathsf{app}}(\tilde{\boldsymbol{x}}_k)$. We re-write the channel model (1)

$$\begin{bmatrix}\boldsymbol{y}^{\mathsf{p}}_l\\ \boldsymbol{y}^{\mathsf{d}}_l\end{bmatrix} = \frac{1}{\sqrt{L}}\sum_{k=1}^{K}\begin{bmatrix}\boldsymbol{p}_k\\ \langle\langle\tilde{\boldsymbol{x}}_k\rangle\rangle_{(i-1)}\end{bmatrix}\vec{\boldsymbol{s}}_{kl}\boldsymbol{h}_k + \frac{1}{\sqrt{L}}\sum_{k=1}^{K}\sum_{m=1}^{M} s^m_{kl}h^m_k\underbrace{\begin{bmatrix}\boldsymbol{0}\\ \boldsymbol{x}_k - \langle\langle\tilde{\boldsymbol{x}}_k\rangle\rangle_{(i-1)}\end{bmatrix}}_{=\Delta\boldsymbol{x}^{(i-1)}_k} + \sigma\boldsymbol{w}_l, \tag{5}$$

and denote $\Delta\boldsymbol{w}^{(i-1)}_{km} = h^m_k\Delta\boldsymbol{x}^{(i-1)}_k \in \mathbb{C}^T$ for later use. Assuming the fading is an ergodic process over the code words, $\boldsymbol{h}_k$ and $\Delta\boldsymbol{x}^{(i-1)}_k$ are statistically independent. Now, postulate a noise variance $\tilde{\sigma}^2$, and $\mathbb{Q}^{(i-1)}\big(\Delta\tilde{\boldsymbol{w}}_{km}\,|\,\mathcal{I}, \langle\langle\tilde{\mathcal{X}}\rangle\rangle_{(i-1)}\big) = \mathsf{CN}(\boldsymbol{0}, \tilde{\boldsymbol{\Sigma}}^{(i-1)}_{\Delta\boldsymbol{w}_{km}})$, where[2] $\tilde{\boldsymbol{\Sigma}}^{(i-1)}_{\Delta\boldsymbol{w}_{km}} = \overline{p}^m_k\big[\boldsymbol{0}^{\mathsf{T}}\ \big(\tilde{\boldsymbol{\Sigma}}^{(i-1)}_{\boldsymbol{x}_k}\big)^{\mathsf{T}}\big]^{\mathsf{T}}$ and $\tilde{\boldsymbol{\Sigma}}^{(i-1)}_{\boldsymbol{x}_k}$ is the postulated covariance matrix of $\mathbb{Q}^{(i-1)}_{\mathsf{app}}(\tilde{\boldsymbol{x}}_k)$. The resulting GPME yields a class of linear channel estimators

$$\langle\langle\, h^m_k\,\rangle\rangle_{(i)} = \int \tilde{h}^m_k\,\mathbb{Q}\big(\mathrm{d}\tilde{\mathcal{H}}\big|\,\mathcal{Y}, \mathcal{I}, \langle\langle\tilde{\mathcal{X}}\rangle\rangle_{(i-1)}\big) = \frac{\mathsf{E}_{\tilde{\mathcal{H}},\Delta\tilde{\mathcal{W}}_{(i-1)}}\big\{\tilde{h}^m_k\,\mathbb{Q}\big(\mathcal{Y}\big|\,\mathcal{I}, \tilde{\mathcal{H}}, \langle\langle\tilde{\mathcal{X}}\rangle\rangle_{(i-1)}, \Delta\tilde{\mathcal{W}}_{(i-1)}\big)\big\}}{\mathsf{E}_{\tilde{\mathcal{H}},\Delta\tilde{\mathcal{W}}_{(i-1)}}\big\{\mathbb{Q}\big(\mathcal{Y}\big|\,\mathcal{I}, \tilde{\mathcal{H}}, \langle\langle\tilde{\mathcal{X}}\rangle\rangle_{(i-1)}, \Delta\tilde{\mathcal{W}}_{(i-1)}\big)\big\}}, \tag{6}$$

where $\Delta\tilde{\mathcal{W}}_{(i-1)} = \big\{\Delta\tilde{\boldsymbol{w}}^{(i-1)}_{km}\,|\,\forall k, m\big\}$, parametrized by:

1) the operator F (see Algorithm 1);
2) the postulated covariance matrix $\tilde{\boldsymbol{\Sigma}}^{(i-1)}_{\boldsymbol{x}_k}$;
3) the postulated noise variance $\tilde{\sigma}^2$.

By choosing these parameters appropriately, we can derive all the usual iterative channel estimators with linear filtering.

Example 3. Let F be the identity operator, so that $\mathbb{Q}^{(i-1)}_{\mathsf{app}}(\tilde{x}_{kn}) = \mathbb{P}^{(i-1)}_{\mathsf{app}}(x_{kn})$, for all k, n. Furthermore, let $\tilde{\boldsymbol{\Sigma}}^{(i-1)}_{\boldsymbol{x}_k} = \boldsymbol{\Sigma}^{(i-1)}_{\boldsymbol{x}_k}$ and $\tilde{\sigma}^2 = \sigma^2$. The GPME (6), is the LMMSE channel estimator for the system (1).

Example 4. Postulating $\tilde{\sigma}^2 = 0$, $\tilde{\boldsymbol{\Sigma}}^{(i-1)}_{\boldsymbol{x}_k} = \boldsymbol{0}$ and defining $F : \mathbb{P}^{(i)}_{\mathsf{app}}(x_{kn}) \mapsto 1_{\tilde{x}_{kn}}\big(\arg\min_{x_{kn}\in\mathcal{M}} \mathbb{P}^{(i)}_{\mathsf{app}}(x_{kn})\big)$ in (6) yields the hard feedback based linear channel estimator in [9].

[2] We interpret $\tilde{\boldsymbol{x}} \sim \mathsf{CN}(\boldsymbol{0}, \boldsymbol{0}) \iff \tilde{\boldsymbol{x}} = \boldsymbol{0}$.

IV. Decoupling Results

In the following we present the decoupling results, derived along the lines of [5], [10] by using the replica method, for the estimators described in Section III. The proofs are omitted due to space constraints.

A. Linear Data Estimation With PIC

Consider the system model (2), and let

$$\boldsymbol{z}_k = \langle\langle\boldsymbol{h}_k\rangle\rangle_{(i)}x_k + \Delta\boldsymbol{v}^{(i)}_k + \boldsymbol{w}_k, \quad \boldsymbol{w}_k \sim \mathsf{CN}(\boldsymbol{0}, \boldsymbol{D}_{(i)}), \tag{7}$$

be a related single-user single-input multiple output flat fading channel. Denote $\mathcal{I}^{(i)}_k = \big\{\boldsymbol{p}_k, \langle\langle\boldsymbol{h}_k\rangle\rangle_{(i)}\big\}$, postulate $\tilde{\boldsymbol{w}}_{km} \sim \mathsf{CN}\big(\boldsymbol{0}, \tilde{\boldsymbol{D}}_{(i)}\big)$, $\mathbb{Q}(\tilde{x}_k) = \mathsf{CN}\big(\langle\langle\tilde{x}_k\rangle\rangle^{(i-1)}_{\mathsf{ext}}, [\tilde{\boldsymbol{\Sigma}}^{(i-1)}_{\boldsymbol{x}_{\setminus 1}}]_{kk}\big)$, and let $\mathbb{Q}^{(i)}\big(\Delta\tilde{\boldsymbol{v}}_k|\mathcal{I}_{(i)}\big) = \mathsf{CN}(\boldsymbol{0}, \tilde{\boldsymbol{\Sigma}}^{(i)}_{\Delta\boldsymbol{v}_k})$. The resulting GPME reads

$$\langle\cdots\rangle_{(i)} = \frac{\mathsf{E}_{\tilde{x}_k,\Delta\boldsymbol{v}^{(i)}_k}\big\{\cdots\,\mathbb{Q}\big(\boldsymbol{z}_k\big|\,\tilde{x}_k, \Delta\boldsymbol{v}^{(i)}_k, \mathcal{I}^{(i)}_k\big)\big|\,\langle\langle\tilde{x}_k\rangle\rangle^{(i-1)}_{\mathsf{ext}}\big\}}{\mathsf{E}_{\tilde{x}_k,\Delta\boldsymbol{v}^{(i)}_k}\big\{\mathbb{Q}\big(\boldsymbol{z}_k\big|\,\tilde{x}_k, \Delta\boldsymbol{v}^{(i)}_k, \mathcal{I}^{(i)}_k\big)\big|\,\langle\langle\tilde{x}_k\rangle\rangle^{(i-1)}_{\mathsf{ext}}\big\}}. \tag{8}$$

Furthermore, let the noise covariance matrices be given by the solutions to the coupled fixed point equations

$$\boldsymbol{D}_{(i)} = \sigma^2\boldsymbol{I}_M + \alpha M\lim_{K\to\infty}\frac{1}{K}\sum_{k=1}^{K}\boldsymbol{V}_k\big(\boldsymbol{D}_{(i)}, \tilde{\boldsymbol{D}}_{(i)}\big), \tag{9}$$

$$\tilde{\boldsymbol{D}}_{(i)} = \tilde{\sigma}^2\boldsymbol{I}_M + \alpha M\lim_{K\to\infty}\frac{1}{K}\sum_{k=1}^{K}\tilde{\boldsymbol{V}}_k\big(\boldsymbol{D}_{(i)}, \tilde{\boldsymbol{D}}_{(i)}\big), \tag{10}$$

where, after denoting $\mathsf{E}_{\mathrm{d}}\{\,\cdot\,\} = \mathsf{E}\{\,\cdot\,|\boldsymbol{h}_k, \mathcal{I}^{(i)}_k\}$,

$$\boldsymbol{V}_k\big(\boldsymbol{D}_{(i)}, \tilde{\boldsymbol{D}}_{(i)}\big) = \mathsf{E}_{\mathrm{d}}\left\{\Big(\boldsymbol{h}_kx_k - \big\langle\tilde{\boldsymbol{u}}^{(i)}_k\big\rangle\Big)\Big(\boldsymbol{h}_kx_k - \big\langle\tilde{\boldsymbol{u}}^{(i)}_k\big\rangle\Big)^{\mathsf{H}}\right\}, \tag{11}$$

$$\tilde{\boldsymbol{V}}_k\big(\boldsymbol{D}_{(i)}, \tilde{\boldsymbol{D}}_{(i)}\big) = \mathsf{E}_{\mathrm{d}}\left\{\Big(\tilde{\boldsymbol{u}}^{(i)}_k - \big\langle\tilde{\boldsymbol{u}}^{(i)}_k\big\rangle\Big)\Big(\tilde{\boldsymbol{u}}^{(i)}_k - \big\langle\tilde{\boldsymbol{u}}^{(i)}_k\big\rangle\Big)^{\mathsf{H}}\right\}, \tag{12}$$

and $\tilde{\boldsymbol{u}}^{(i)}_k = \langle\langle\tilde{\boldsymbol{h}}_k\rangle\rangle_{(i)}\tilde{x}_k + \Delta\tilde{\boldsymbol{v}}^{(i)}_k$.

Claim 1. *Let $N = \tau_{\mathsf{d}}C \to \infty$ and $K = \alpha L \to \infty$ with α and τ_{d} fixed. Conditioned on $\{\mathcal{H}, \mathcal{Q}^{(i)}_{\mathcal{H}}\}$, the joint distribution of the true and postulated inputs and the estimate $\langle\langle x_1\rangle\rangle_{(i)}$ of the multiuser system* (2)*, converges in probability to the joint distribution of the true and postulated inputs and the estimate $\langle x_k\rangle_{(i)}$ of the single-user system* (7)*.*

B. Linear Channel Estimation With Information Feedback

Let us recall the system model (5) and define a related flat fading single-user channel for $m = 1, 2, \dots, M$ as

$$\boldsymbol{z}_{km} = \underbrace{\begin{bmatrix}\boldsymbol{p}_k\\ \langle\langle\tilde{\boldsymbol{x}}_k\rangle\rangle_{(i-1)}\end{bmatrix}}_{=\tilde{\boldsymbol{x}}^{(i-1)}_k} h^m_k + \underbrace{\begin{bmatrix}\boldsymbol{0}\\ \boldsymbol{x}_k - \langle\langle\tilde{\boldsymbol{x}}_k\rangle\rangle_{(i-1)}\end{bmatrix}}_{=\Delta\boldsymbol{x}^{(i-1)}_k} h^m_k + \boldsymbol{w}_{km}, \tag{13}$$

where $\boldsymbol{w}_{km} \sim \mathsf{CN}\big(\boldsymbol{0}, \boldsymbol{C}_{(i)}\big)$. Let $\mathcal{I}^{(i-1)}_k = \big\{\boldsymbol{p}_k, \langle\langle\tilde{\boldsymbol{x}}_k\rangle\rangle_{(i-1)}\big\}$ be known and postulate the noise is distributed as

$\tilde{\boldsymbol{w}}_{km} \sim \mathsf{CN}\big(\mathbf{0}, \tilde{\boldsymbol{C}}_{(i)}\big)$. If we let $\mathbb{Q}^{(i-1)}\big(\Delta\tilde{\boldsymbol{w}}_{km}\,|\,\mathcal{I}_k^{(i-1)}\big) = \mathsf{CN}(\mathbf{0}, \tilde{\boldsymbol{\Sigma}}^{(i-1)}_{\Delta \boldsymbol{w}_{km}})$, the resulting single-user GPME reads

$$\langle\cdots\rangle_{(i)} = \frac{\mathsf{E}_{\tilde{\boldsymbol{h}}_k, \Delta\tilde{\mathcal{W}}_k^{(i-1)}}\big\{\cdots\prod_{m=1}^M \mathbb{Q}\big(\boldsymbol{z}_{km}\,\big|\,\tilde{h}_k^m, \Delta\tilde{\boldsymbol{w}}_{km}^{(i-1)}, \mathcal{I}_k^{(i-1)}\big)\big\}}{\mathsf{E}_{\tilde{\boldsymbol{h}}_k, \Delta\tilde{\mathcal{W}}_k^{(i-1)}}\big\{\prod_{m=1}^M \mathbb{Q}\big(\boldsymbol{z}_{km}\,\big|\,\tilde{h}_k^m, \Delta\tilde{\boldsymbol{w}}_{km}^{(i-1)}, \mathcal{I}_k^{(i-1)}\big)\big\}}. \tag{14}$$

Furthermore, let the noise covariance matrices be given by the solutions to the coupled fixed point equations

$$\boldsymbol{C}_{(i)} = \sigma^2 \boldsymbol{I}_T + \alpha \lim_{K\to\infty} \frac{1}{K}\sum_{k=1}^K \sum_{m=1}^M \boldsymbol{V}_{km}\big(\boldsymbol{C}_{(i)}, \tilde{\boldsymbol{C}}_{(i)}\big), \tag{15}$$

$$\tilde{\boldsymbol{C}}_{(i)} = \tilde{\sigma}^2 \boldsymbol{I}_T + \alpha \lim_{K\to\infty} \frac{1}{K}\sum_{k=1}^K \sum_{m=1}^M \tilde{\boldsymbol{V}}_{km}\big(\boldsymbol{C}_{(i)}, \tilde{\boldsymbol{C}}_{(i)}\big). \tag{16}$$

where, after denoting $\mathsf{E}_{\mathsf{c}}\{\cdot\} = \mathsf{E}\{\cdot\,|\,\boldsymbol{x}_k, \mathcal{I}_k^{(i-1)}\}$,

$$\boldsymbol{V}_{km}\big(\boldsymbol{C}_{(i)}, \tilde{\boldsymbol{C}}_{(i)}\big) = \mathsf{E}_{\mathsf{c}}\left\{\Big(\boldsymbol{x}_k h_k^m - \big\langle\tilde{\boldsymbol{u}}_{km}^{(i-1)}\big\rangle\Big)\Big(\boldsymbol{x}_k h_k^m - \big\langle\tilde{\boldsymbol{u}}_{km}^{(i-1)}\big\rangle\Big)^{\mathsf{H}}\right\}, \tag{17}$$

$$\tilde{\boldsymbol{V}}_{km}\big(\boldsymbol{C}_{(i)}, \tilde{\boldsymbol{C}}_{(i)}\big) = \mathsf{E}_{\mathsf{c}}\left\{\Big(\tilde{\boldsymbol{u}}_{km}^{(i-1)} - \big\langle\tilde{\boldsymbol{u}}_{km}^{(i-1)}\big\rangle\Big)\Big(\tilde{\boldsymbol{u}}_{km}^{(i-1)} - \big\langle\tilde{\boldsymbol{u}}_{km}^{(i-1)}\big\rangle\Big)^{\mathsf{H}}\right\}, \tag{18}$$

and $\tilde{\boldsymbol{u}}_{km}^{(i-1)} = \tilde{\boldsymbol{x}}_k^{(i-1)}\tilde{h}_k^m + \Delta\tilde{\boldsymbol{w}}_{km}^{(i-1)}$.

Claim 2. *Let $N = \tau_{\mathsf{d}} C \to \infty$ and $K = \alpha L \to \infty$ with α and τ_{d} finite and fixed. Conditioned on $\{\mathcal{X}, \mathcal{P}, \mathcal{Q}_{\mathcal{X}}^{(i-1)}\}$, the joint distribution of the true and postulated channel coefficients and the estimates $\big\{\langle\langle h_k^m\rangle\rangle_{(i)}\big\}$ of the multiuser system (5), converges in probability to the joint distribution of the true and postulated channel coefficients and the estimates $\big\{\langle h_k^m\rangle_{(i)}\big\}$ of the single-user system (13).*

Remark 1. Due to Claims 1 and 2, the performance of the general iterative multiuser DS-CDMA system described in Sections II and III, can be analyzed by concentrating on the much simpler single-user system described by (7) – (18).

V. Performance Analysis

For simplicity we assume in the following that $\overline{\boldsymbol{P}}_k = \frac{\overline{p}}{M}\boldsymbol{I}$ for all $k = 1, \ldots, K$.

A. Linear Data Estimation With PIC

Corollary 1. *Consider the LMMSE-PIC or "SUMF-Based Soft IC" in Examples 1 and 2, respectively. Let the channel information be given by the LMMSE estimator in Example 3. The post-detection SINR at ith iteration is then given by*

$$\mathsf{SINR}_{(i)} = \frac{\|\tilde{\boldsymbol{h}}_{(i)}\|^2 D_{(i)}}{1 + \xi_{(i)} D_{(i)}}, \tag{19}$$

where $\xi^{(i)}$ is the per-path MSE of the channel estimate and $\tilde{\boldsymbol{h}}_{(i)} \sim \mathsf{CN}\big(\mathbf{0}, \big(\frac{\overline{p}}{M} - \xi_{(i)}\big)\boldsymbol{I}\big)$. The noise covariance $\boldsymbol{D}_{(i)} = \frac{1}{D_{(i)}}\boldsymbol{I}$ for the SUMF-receiver is

$$\frac{1}{D_{(i)}} = \sigma^2 + \alpha\mathsf{E}\left\{\Sigma_{\Delta x}^{\mathsf{ext}}\|\tilde{\boldsymbol{h}}_{(i)}\|^2 + \xi_{(i)}\right\}, \tag{20}$$

and for the LMMSE-based receiver the solution to

$$\frac{1}{D_{(i)}} = \sigma^2 + \alpha(M-1)\frac{\xi_{(i)}}{1 + \xi_{(i)} D_{(i)}} + \alpha\mathsf{E}\left\{\frac{\Sigma_{\Delta x}^{\mathsf{ext}}\|\tilde{\boldsymbol{h}}_{(i)}\|^2 + \xi_{(i)}}{1 + D\big(\Sigma_{\Delta x}^{\mathsf{ext}}\|\tilde{\boldsymbol{h}}_{(i)}\|^2 + \xi_{(i)}\big)}\right\}, \tag{21}$$

where $\Sigma_{\Delta x}^{\mathsf{ext}}$ denotes the MSE of the soft symbols, based on extrinsic information of a single-user system over an ergodic Rayleigh fading channel whose received SNR has the same distribution as $\mathsf{SINR}_{(i-1)}$.

Remark 2. It can be easily verified that Claim 1 reproduces also the results reported in [1], [2], [7].

B. Linear Channel Estimation With Information Feedback

Corollary 2. *For the LMMSE channel estimator of Example 3, $\boldsymbol{C}_{(i)} = \tilde{\boldsymbol{C}}_{(i)}$ and $\boldsymbol{C}_{(i)} = \sigma^2\boldsymbol{I} + \alpha M\mathsf{E}\big\{\boldsymbol{V}\big(\boldsymbol{C}_{(i)}, \boldsymbol{C}_{(i)}\big)\big\}$, where*

$$\mathsf{E}\big\{\boldsymbol{V}\big(\boldsymbol{C}_{(i)}, \boldsymbol{C}_{(i)}\big)\big\} = \mathsf{E}\left\{\boldsymbol{C}\big(\boldsymbol{\Sigma}_{\Delta\boldsymbol{w}} + \boldsymbol{C}_{(i)}\big)^{-1} \times\left(\tilde{\boldsymbol{\Sigma}}_{\Delta\boldsymbol{w}} + \frac{\frac{\overline{p}}{M}\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}^{\mathsf{H}}\big(\boldsymbol{\Sigma}_{\Delta\boldsymbol{w}} + \boldsymbol{C}_{(i)}\big)^{-1}\boldsymbol{C}}{1 + \frac{\overline{p}}{M}\tilde{\boldsymbol{x}}^{\mathsf{H}}\big(\boldsymbol{\Sigma}_{\Delta\boldsymbol{w}} + \boldsymbol{C}_{(i)}\big)^{-1}\tilde{\boldsymbol{x}}}\right)\right\}, \tag{22}$$

and $\tilde{\boldsymbol{x}} = \mathrm{vec}([\boldsymbol{p} \;\; \langle\tilde{\boldsymbol{x}}\rangle_{\mathsf{app}}])$, where $\langle\tilde{\boldsymbol{x}}\rangle_{\mathsf{app}}$ denotes the soft feedback and $\Sigma_{\Delta\boldsymbol{x}}^{\mathsf{app}}$ the corresponding MSE, based on APPs of a single-user system operating over an ergodic Rayleigh fading channel whose received SNR has the same distribution as $\mathsf{SINR}_{(i-1)}$. For large coherence time T, the solution to

$$\frac{1}{C_{\mathsf{p}}^{(i)}} = \sigma^2 + \frac{\alpha M\frac{\overline{p}}{M}}{1 + \frac{\overline{p}}{M}\big(\tau_{\mathsf{p}} C_{\mathsf{p}}^{(i)} + \tau_{\mathsf{d}} C_{\mathsf{d}}^{(i)}\mathsf{E}\big\{|\langle\tilde{x}\rangle_{\mathsf{app}}|^2(1 + C_{\mathsf{d}}^{(i)}\frac{\overline{p}}{M}\Sigma_{\Delta x}^{\mathsf{app}})^{-1}\big\}\big)} \tag{23}$$

$$\frac{1}{C_{\mathsf{d}}^{(i)}} = \sigma^2 + \alpha M\mathsf{E}\left\{\frac{\frac{\overline{p}}{M}\Sigma_{\Delta x}^{\mathsf{app}}}{1 + C_{\mathsf{d}}^{(i)}\frac{\overline{p}}{M}\Sigma_{\Delta x}^{\mathsf{app}}}\right\} + \frac{\alpha M\frac{\overline{p}}{M}\mathsf{E}\big\{|\langle\tilde{x}\rangle_{\mathsf{app}}|^2(1 + C_{\mathsf{d}}^{(i)}\frac{\overline{p}}{M}\Sigma_{\Delta x}^{\mathsf{app}})^{-2}\big\}}{1 + \frac{\overline{p}}{M}\big(\tau_{\mathsf{p}} C_{\mathsf{p}}^{(i)} + \tau_{\mathsf{d}} C_{\mathsf{d}}^{(i)}\mathsf{E}\big\{|\langle\tilde{x}\rangle_{\mathsf{app}}|^2(1 + C_{\mathsf{d}}^{(i)}\frac{\overline{p}}{M}\Sigma_{\Delta x}^{\mathsf{app}})^{-1}\big\}\big)} \tag{24}$$

satisfies $\frac{1}{C_{\mathsf{p}}^{(i)}} = \sigma^2 + \alpha M\xi_{(i)}$, where $\xi_{(i)}$ is the per-path MSE.

Remark 3. If we let $\tau_{\mathsf{d}} = 0$ or $\langle\tilde{x}\rangle_{\mathsf{app}} = x$ so that $\Sigma_{\Delta x}^{\mathsf{app}} = 0$, Corollary 2 reduces to [2, Thm. 2], as expected. We also note from (23) and (24) that the use of soft feedback can never increase the per-path MSE of this channel estimator.

Corollary 3. *If $\alpha M < T$, the per-path MSE of the estimator given in Example 4 converges in the large system limit to*

$$\xi_{(i)} = \frac{\sigma^2}{T - \alpha M} + \frac{4\frac{\overline{p}}{M}\varepsilon_{\mathsf{app}}\tau_{\mathsf{d}}\big[1 + \varepsilon_{\mathsf{app}}(\tau_{\mathsf{d}} - 1)\big]}{T^2} + \alpha\frac{4\overline{p}\varepsilon_{\mathsf{app}}\tau_{\mathsf{d}}\big[T - \big(1 + \varepsilon_{\mathsf{app}}(\tau_{\mathsf{d}} - 1)\big)\big]}{T^2(T - \alpha M)}, \tag{25}$$

where $\varepsilon_{\mathsf{app}}$ denotes the probability of error of the information bits based on APPs of a single-user system operating over an ergodic Rayleigh fading channel whose received SNR has the same distribution as $\mathsf{SINR}_{(i-1)}$.

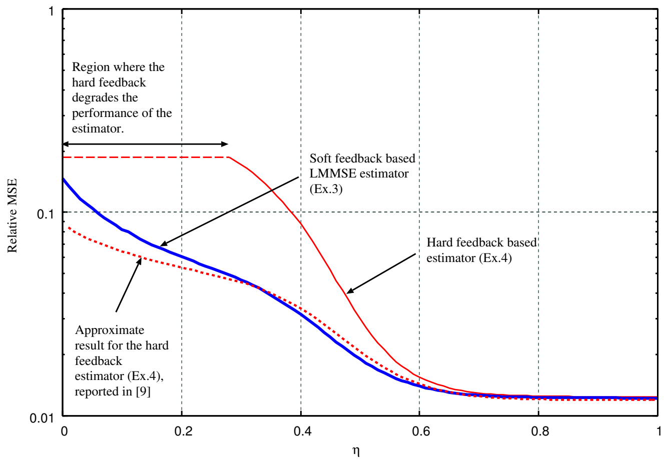

Fig. 1. Relative MSE $\frac{\xi}{\overline{p}/M}$ vs. multiuser efficiency η for the linear channel estimators with soft or hard feedback, as given in Examples 3 (blue line) and 4 (red lines), respectively. Solid lines correspond to the exact large system results reported here, dotted line is for the approximate result from [9]. Three equal power paths, coherence time of $T = 100$ symbols, $\tau_\mathsf{p} = 10$ pilots, load $\alpha = 1.2$ and average SNR $\overline{\gamma} = 4$ dB.

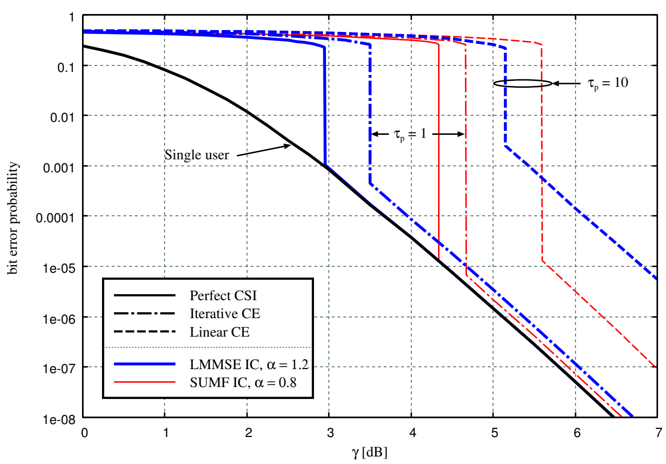

Fig. 2. Bit error probability vs. SNR $\overline{\gamma}$ of the LMMSE (Ex. 1) and SUMF (Ex. 2) based MUDD with soft IC. Perfect CSI, LMMSE channel estimation (CE) [2] and iterative CE (Ex. 3) are considered for given loads α and numbers of pilots τ_p. Three equal power paths, coherence time of $T = 100$ symbols, rate-$1/2$ convolutional code with generators $(753, 561)_8$ and Gray QPSK.

Remark 4. The main result of [9, Sec. III] is an approximation, whereas (25) is exact at the large system limit. The results coincide if $\varepsilon_\mathsf{app} \approx 0$, and $T \gg M$ as in [9].

VI. Numerical Examples

Let us consider a half-rate convolutional code with generator polynomials $(753, 561)_8$. With QPSK, the average SNR per information bit is $\overline{\gamma}$. We remark that the numerical examples given here are based on the asymptotic results obtained in Section V and for finite systems are approximations.

The relative MSE $\frac{\xi}{\overline{p}/M}$ vs. the average multiuser efficiency $\eta = \mathsf{E}\{\mathsf{SINR}/\overline{\gamma}\}$ for the channel estimators given in Examples 3 and 4 is plotted in Figure 1. The analysis in [9] gives far too optimistic results in this case, and misses the region where the unreliable hard feedback causes increase instead of decrease in the MSE (dashed part of the upper curve). Since the soft feedback cannot degrade the performance of the estimator (cf. Remark 3), it is more suitable for iterative receivers than the hard feedback. We therefore consider only soft feedback based receivers in the following.

Bit error probability vs. SNR for an iterative system consisting of the LMMSE or SUMF-based MUDD given in Examples 1 and 2, respectively, and the channel estimator in Example 3 is shown in Figure 2. Channel coherence time of $T = 100$ symbols and three $M = 3$ equal power paths are assumed. Remarkably, the system with load $\alpha = 1.2$ and LMMSE-based iterative receiver converges to near single-user performance at relatively low SNR with only one pilot symbol. The performance of the system with linear LMMSE channel estimation [2], on the other hand, suffers significantly from the channel estimation errors even with ten pilots. For the iterative system with SUMF-based MUDD, the user load was dropped to $\alpha = 0.8$ in order to converge within the given range. This is due to the relatively poor multiple access interference suppression capability of the SUMF-based MUDD.

References

[1] D. N. C. Tse and S. V. Hanly, "Linear multiuser receivers: Effective interference, effective bandwidth and user capacity," *IEEE Trans. Inform. Theory*, vol. 45, no. 2, pp. 641–657, Mar. 1999.

[2] J. Evans and D. N. C. Tse, "Large system performance of linear multiuser receivers in multipath fading channels," *IEEE Trans. Inform. Theory*, vol. 46, no. 6, pp. 2059–2078, Sep. 2000.

[3] S. Shamai and S. Verdú, "The impact of frequency-flat fading on the spectral efficiency of CDMA," *IEEE Trans. Inform. Theory*, vol. 47, no. 4, pp. 1302–1327, May 2001.

[4] R. R. Müller and W. H. Gerstacker, "On the capacity loss due to separation of detection and decoding," *IEEE Trans. Inform. Theory*, vol. 50, no. 8, pp. 1769–1778, Aug. 2004.

[5] D. Guo and S. Verdú, "Randomly spread CDMA: Asymptotics via statistical physics," *IEEE Trans. Inform. Theory*, vol. 51, no. 6, pp. 1983–2010, Jun. 2005.

[6] A. Worthen and W. Stark, "Unified design of iterative receivers using factor graphs," *IEEE Trans. Inform. Theory*, vol. 47, no. 2, pp. 843–849, Feb. 2001.

[7] J. Boutros and G. Caire, "Iterative multiuser joint decoding: Unified framework and asymptotic analysis," *IEEE Trans. Inform. Theory*, vol. 48, no. 7, pp. 1772–1793, Jul. 2002.

[8] G. Caire, R. R. Müller, and T. Tanaka, "Iterative multiuser joint decoding: Optimal power allocation and low-complexity implementation," *IEEE Trans. Inform. Theory*, vol. 50, no. 9, pp. 1950–1973, Sep. 2004.

[9] H. Li, S. M. Betz, and H. V. Poor, "Performance analysis of iterative channel estimation and multiuser detection in multipath DS-CDMA channels," *IEEE Trans. Signal Processing*, vol. 55, no. 5, pp. 1981–1993, May 2007.

[10] K. Takeuchi, T. Tanaka, and T. Yano, "Asymptotic analysis of general multiuser detectors in MIMO DS-CDMA channels," *IEEE J. Select. Areas Commun.*, vol. 26, no. 3, pp. 486–496, Apr. 2008.

[11] K. Takeuchi, M. Vehkaperä, T. Tanaka, and R. R. Müller, "Replica analysis of general multiuser detection in MIMO DS-CDMA channels with imperfect CSI," in *Proc. IEEE Int. Symp. Inform. Theory*, Toronto, ON, Canada, Jul. 6–11 2008, pp. 514–518.

[12] H. Li and H. V. Poor, "Impact of channel estimation errors on multiuser detection via the replica method," *EURASIP J. Wirel. Commun. Netw.*, no. 2, pp. 175–186, Apr. 2005.

[13] T. Tanaka, "A statistical-mechanics approach to large-system analysis of CDMA multiuser detectors," *IEEE Trans. Inform. Theory*, vol. 48, no. 11, pp. 2888–2910, Nov. 2002.

[14] R. R. Müller, D. Guo, and A. Moustakas, "Vector precoding for wireless MIMO systems and its replica analysis," *IEEE J. Select. Areas Commun.*, vol. 26, no. 3, pp. 530–540, Apr. 2008.

[15] M. Talagrand, *Spin Glasses: A Challenge for Mathematicians, Cavity and Mean Field Models*. Berlin Heidelberg: Springer-Verlag, 2003.