Replica Analysis of General Multiuser Detection in MIMO DS-CDMA Channels with Imperfect CSI

Keigo Takeuchi*, Mikko Vehkaperä†, Toshiyuki Tanaka*, and Ralf R. Müller†
*Graduate School of Informatics, Kyoto University, Kyoto 606–8501, Japan
Email: takeuchi@sys.i.kyoto-u.ac.jp, tt@i.kyoto-u.ac.jp
†Dept Elec. & Telecommun., Norwegian University of Science and Technology, NO–7491 Trondheim, Norway
Email: mikko@iet.ntnu.no, ralf@iet.ntnu.no

***Abstract*—We consider impacts of channel estimation errors on performance of general multiuser detectors in MIMO DS-CDMA channels. We evaluate their performance in terms of asymptotic spectral efficiency, which is obtained via *decoupling structure*, by using the replica method. Numerical results imply that the performance of LMMSE detection is very close to that of MMSE detection for small system loads. Furthermore, we find that the spectral efficiency of MMSE detection grows discontinuously with the length of pilot sequences for large system loads, and that the critical length is close to the optimal length. While it is indistinguishable from that of LMMSE detection for short pilot sequences, the gap between the two is significantly large if the length of pilot sequences is longer than the critical length.**

I. Introduction

Multiple-input multiple-output direct-sequence code-division multiple-access (MIMO DS-CDMA) systems have been envisioned for future wireless communications in [1], [2]. Performance evaluation of multiuser detection schemes in MIMO DS-CDMA systems is an important fundamental issue. Perfect channel state information (CSI) at the receiver was assumed in previous studies [1]–[3]. In this paper, we evaluate performance of general multiuser detectors in MIMO DS-CDMA channels with imperfect CSI at the receiver, in order to elucidate influences of channel estimation errors on the performance of multiuser detection.

Most studies on performance evaluation of multiuser detection in conventional DS-CDMA channels with imperfect CSI consider large-system limit, in which the number of users and the spreading factor go to infinity with their ratio kept constant. Evans and Tse [4] focused on *decoupling structure* and elucidated impacts of channel estimation errors on linear multiuser detectors (MUDs) quantitatively and qualitatively. Li and Poor [5] approximately evaluated the performance of nonlinear MUDs with imperfect CSI at the receiver via the so-called replica method. We extend the philosophy of [4] and the methodology of [5] to the case of MIMO DS-CDMA systems.

The replica method is a powerful approach to the large-system analysis of random vector channels [6]–[10]. Despite the fact that the replica method still lacks mathematical justification, the replica strategy provides two advantages to us; one is that it allows us to deal with nonlinear MUDs. The other is that it is suitable for analyzing the decoupling structure of random vector channels [3]. In this paper we assume validity of the replica method.

Throughout this paper, $\mathcal{CN}(\mathbf{0},\mathbf{\Sigma})$, $\otimes$, $\log a$, and $\ln a$ stand for a circularly symmetric zero-mean complex Gaussian distribution with covariance matrix $\mathbf{\Sigma}$, the Kronecker product operator between two matrices, $\log_2 a$, and $\log_{\mathrm{e}} a$, respectively. $\boldsymbol{e}_n \in \mathbb{R}^n$ is defined as $\boldsymbol{e}_n = (1,\ldots,1)^T$. $\mathrm{KL}(\boldsymbol{A}\|\boldsymbol{B})$ denotes the Kullback-Leibler divergence with the logarithm to base 2 between $\mathcal{CN}(\mathbf{0},\boldsymbol{A})$ and $\mathcal{CN}(\mathbf{0},\boldsymbol{B})$. For an $n\times m$ matrix $\boldsymbol{A}=(\boldsymbol{a}_1,\ldots,\boldsymbol{a}_m)$, $\mathrm{vec}(\boldsymbol{A})$ means $\mathrm{vec}(\boldsymbol{A}) = (\boldsymbol{a}_1^T,\ldots,\boldsymbol{a}_m^T)^T$. For a zero-mean random vector $\boldsymbol{a}$, $\mathrm{cov}[\boldsymbol{a}]$ stands for the covariance matrix $\mathrm{E}[\boldsymbol{a}\boldsymbol{a}^H]$.

II. Channel Model

A. MIMO DS-CDMA Channels

We consider a block-fading randomly spread MIMO DS-CDMA channel with K fully-synchronized users, in which the kth user and the receiver use M_k transmit antennas and N receive antennas, respectively,

$$\boldsymbol{y}_{l,t}^{(\mathrm{p})} = \frac{1}{\sqrt{L}}\sum_{k=1}^{K}\boldsymbol{H}_k\boldsymbol{S}_{k,l}^{(\mathrm{p})}\boldsymbol{x}_{k,t}^{(\mathrm{p})} + \boldsymbol{n}_{l,t}^{(\mathrm{p})},\quad t=1,\ldots,\tau, \tag{1}$$

$$\boldsymbol{y}_{l,t}^{(\mathrm{d})} = \frac{1}{\sqrt{L}}\sum_{k=1}^{K}\boldsymbol{H}_k\boldsymbol{S}_{k,l}^{(\mathrm{d})}\boldsymbol{x}_{k,t}^{(\mathrm{d})} + \boldsymbol{n}_{l,t}^{(\mathrm{d})},\quad t=\tau+1,\ldots,T_{\mathrm{c}}, \tag{2}$$

where $\{\boldsymbol{x}_{k,t}^{(\mathrm{p})} \in \mathbb{C}^{M_k}\}_{t=1}^{\tau}$ and $\{\boldsymbol{x}_{k,t}^{(\mathrm{d})} \in \mathbb{C}^{M_k}\}_{t=\tau+1}^{T_{\mathrm{c}}}$ are the sequence of pilot symbol vectors of the kth user and the sequence of interleaved data symbol vectors of the kth user, respectively; $\boldsymbol{y}_{l,t}^{(\mathrm{p})} \in \mathbb{C}^N$ and $\boldsymbol{y}_{l,t}^{(\mathrm{d})} \in \mathbb{C}^N$ denote the received vectors at chip l ($l=1,\ldots,L$) in the respective times; $\boldsymbol{H}_k \in \mathbb{C}^{N\times M_k}$ is the channel matrix of the kth user; $\boldsymbol{S}_{k,l}^{(\mathrm{p})}, \boldsymbol{S}_{k,l}^{(\mathrm{d})} \in \mathbb{C}^{M_k\times M_k}$ are spatial spreading matrices at chip l of the kth user; and where $\boldsymbol{n}_{l,t}^{(\mathrm{p})}, \boldsymbol{n}_{l,t}^{(\mathrm{d})} \sim \mathcal{CN}(\mathbf{0}, N_0\boldsymbol{I}_N)$ represent additive white Gaussian noises (AWGNs). T_{c} corresponds to a coherence time.

$\boldsymbol{x}_{k,t}^{(\mathrm{p})}$ and $\boldsymbol{x}_{k,t}^{(\mathrm{d})}$ are assumed to satisfy $\mathrm{E}[\boldsymbol{x}_{k,t}^{(\mathrm{p})}] = \mathrm{E}[\boldsymbol{x}_{k,t}^{(\mathrm{d})}] = 0$ and $\mathrm{E}[\|\boldsymbol{x}_{k,t}^{(\mathrm{p})}\|^2] = \mathrm{E}[\|\boldsymbol{x}_{k,t}^{(\mathrm{d})}\|^2] = P_k$. Moreover, we assume that $\boldsymbol{S}_{k,l}^{(\mathrm{p})}$ and $\boldsymbol{S}_{k,l}^{(\mathrm{d})}$ are independent and identically distributed (i.i.d.) for all l. Let $\boldsymbol{C}_k \in \mathbb{C}^{M_k^2\times M_k^2}$ be the covariance matrix of $\mathrm{vec}(\boldsymbol{S}_{k,l}^{(i)})$ for $i=\mathrm{p},\mathrm{d}$, which satisfies $\|\boldsymbol{C}_k\|^2 = M_k$, and let $\mathrm{rank}(\boldsymbol{C}_k) = L_k$. Without loss of generality, we can assume $\boldsymbol{S}_{k,l}^{(i)} = \sum_{j=1}^{L_k} s_{k,j}^{(l,i)}\mathbf{\Gamma}_{k,j}$ [11], where $\mathrm{Re}(s_{k,j}^{(l,i)})$ and $\mathrm{Im}(s_{k,j}^{(l,i)})$ are i.i.d. zero-mean random variables with variance $1/2$ for all k, j, and l, and where $\mathbf{\Gamma}_{k,j} \in \mathbb{C}^{M_k \times M_k}$ is given by $\boldsymbol{C}_k = \mathbf{\Gamma}_k \mathbf{\Gamma}_k^H$ with $\mathbf{\Gamma}_k = (\mathrm{vec}(\mathbf{\Gamma}_{k,1}), \ldots, \mathrm{vec}(\mathbf{\Gamma}_{k,L_k}))$ and satisfies $\sum_{j=1}^{L_k} \|\mathbf{\Gamma}_{k,j}\|^2 = M_k$.

B. Channel and Data Estimations

We assume that the receiver knows the realizations of the spatial spreading matrices $\boldsymbol{\mathcal{S}}^{(i)} = \{\boldsymbol{S}_{k,l}^{(i)}; l = 1, \ldots, L,\ k = 1, \ldots, K\}$ for $i = \mathrm{p}, \mathrm{d}$ and of the pilot symbol vectors $\boldsymbol{\mathcal{X}}^{(\mathrm{p})} = \{\boldsymbol{x}_{k,t}^{(\mathrm{p})};\ \text{for all } k,\ t\}$ but does not know the realizations of the channel matrices $\boldsymbol{\mathcal{H}} = \{\boldsymbol{H}_k; k = 1, \ldots, K\}$.

In the first τ symbol periods, a channel estimator estimates the channel matrices from $\boldsymbol{\mathcal{I}}_{\mathrm{p}} = \{\boldsymbol{\mathcal{Y}}^{(\mathrm{p})}, \boldsymbol{\mathcal{X}}^{(\mathrm{p})}, \boldsymbol{\mathcal{S}}^{(\mathrm{p})}\}$, in which $\boldsymbol{\mathcal{Y}}^{(\mathrm{p})} = \{\boldsymbol{y}_{l,t}^{(\mathrm{p})}; \text{for all } l,\ t\}$, and feeds their estimates and uncertainties to a MUD in the form of a posterior distribution $p(\tilde{\boldsymbol{\mathcal{H}}} = \{\tilde{\boldsymbol{H}}_k; k = 1, \ldots, K\} | \boldsymbol{\mathcal{I}}_{\mathrm{p}})$ of the channel matrices conditioned on $\boldsymbol{\mathcal{I}}_{\mathrm{p}}$. We have added a tilde in order to distinguish $p(\tilde{\boldsymbol{\mathcal{H}}}|\boldsymbol{\mathcal{I}}_{\mathrm{p}})$ from the true posterior distribution $p(\boldsymbol{\mathcal{H}}|\boldsymbol{\mathcal{I}}_{\mathrm{p}})$. The estimate of $\boldsymbol{H}_k$ is given as $\hat{\boldsymbol{H}}_k = \int \tilde{\boldsymbol{H}}_k p(\tilde{\boldsymbol{\mathcal{H}}}|\boldsymbol{\mathcal{I}}_{\mathrm{p}}) d\tilde{\boldsymbol{\mathcal{H}}}$ and the (postulated) channel estimation error of $\boldsymbol{H}_k$ is represented by $\Delta\tilde{\boldsymbol{H}}_k = \tilde{\boldsymbol{H}}_k - \hat{\boldsymbol{H}}_k$.

We consider "one-shot" multiuser detection as in [4], in which $\boldsymbol{x}_{k,t}^{(\mathrm{d})}$ is estimated only from $\boldsymbol{\mathcal{Y}}_t^{(\mathrm{d})} = \{\boldsymbol{y}_{l,t}^{(\mathrm{d})}; l = 1, \ldots, L\}$, $\boldsymbol{\mathcal{S}}^{(\mathrm{d})}$, and $p(\tilde{\boldsymbol{\mathcal{H}}}|\boldsymbol{\mathcal{I}}_{\mathrm{p}})$, and in which $\{\boldsymbol{\mathcal{Y}}_{t'}^{(\mathrm{d})}; t' \neq t\}$ are not utilized for the estimation of $\boldsymbol{x}_{k,t}^{(\mathrm{d})}$. In order to derive a minimum mean-squared error (MMSE) MUD and a linear MMSE (LMMSE) MUD, we define a generalized posterior mean estimator (GPME)[1] as

$$\langle\langle \tilde{\boldsymbol{x}}_{k,t}^{(\mathrm{d})} \rangle\rangle = \frac{\mathrm{E}_{\tilde{\boldsymbol{\mathcal{X}}}_t^{(\mathrm{d})}}[\tilde{\boldsymbol{x}}_{k,t}^{(\mathrm{d})} p(\tilde{\boldsymbol{\mathcal{Y}}}_t^{(\mathrm{d})} = \boldsymbol{\mathcal{Y}}_t^{(\mathrm{d})} | \tilde{\boldsymbol{\mathcal{X}}}_t^{(\mathrm{d})}, \boldsymbol{\mathcal{I}}_{\mathrm{p}}, \boldsymbol{\mathcal{S}}^{(\mathrm{d})})]}{\mathrm{E}_{\tilde{\boldsymbol{\mathcal{X}}}_t^{(\mathrm{d})}}[p(\tilde{\boldsymbol{\mathcal{Y}}}_t^{(\mathrm{d})} = \boldsymbol{\mathcal{Y}}_t^{(\mathrm{d})} | \tilde{\boldsymbol{\mathcal{X}}}_t^{(\mathrm{d})}, \boldsymbol{\mathcal{I}}_{\mathrm{p}}, \boldsymbol{\mathcal{S}}^{(\mathrm{d})})]}, \tag{3}$$

$$p(\tilde{\boldsymbol{\mathcal{Y}}}_t^{(\mathrm{d})} | \tilde{\boldsymbol{\mathcal{X}}}_t^{(\mathrm{d})}, \boldsymbol{\mathcal{I}}_{\mathrm{p}}, \boldsymbol{\mathcal{S}}^{(\mathrm{d})}) = \int p(\tilde{\boldsymbol{\mathcal{Y}}}_t^{(\mathrm{d})} | \tilde{\boldsymbol{\mathcal{X}}}_t^{(\mathrm{d})}, \boldsymbol{\mathcal{S}}^{(\mathrm{d})}, \tilde{\boldsymbol{\mathcal{H}}}) p(\tilde{\boldsymbol{\mathcal{H}}}|\boldsymbol{\mathcal{I}}_{\mathrm{p}}) d\tilde{\boldsymbol{\mathcal{H}}}, \tag{4}$$

where $\tilde{\boldsymbol{\mathcal{Y}}}_t^{(\mathrm{d})} = \{\tilde{\boldsymbol{y}}_{l,t}^{(\mathrm{d})}; l = 1, \ldots, L\}$, $\tilde{\boldsymbol{\mathcal{X}}}_t^{(\mathrm{d})} = \{\tilde{\boldsymbol{x}}_{k,t}^{(\mathrm{d})}; k = 1, \ldots, K\}$, and where $p(\tilde{\boldsymbol{\mathcal{Y}}}_t^{(\mathrm{d})} | \tilde{\boldsymbol{\mathcal{X}}}_t^{(\mathrm{d})}, \boldsymbol{\mathcal{S}}^{(\mathrm{d})}, \tilde{\boldsymbol{\mathcal{H}}})$ characterizes the postulated channel by the MUD

$$\tilde{\boldsymbol{y}}_{l,t}^{(\mathrm{d})} = \frac{1}{\sqrt{L}} \sum_{k=1}^{K} \tilde{\boldsymbol{H}}_k \boldsymbol{S}_{k,l}^{(\mathrm{d})} \tilde{\boldsymbol{x}}_{k,t}^{(\mathrm{d})} + \tilde{\boldsymbol{n}}_{l,t}^{(\mathrm{d})}, \quad \tilde{\boldsymbol{n}}_{l,t}^{(\mathrm{d})} \sim \mathcal{CN}(\boldsymbol{0}, \tilde{N}_0 \boldsymbol{I}_N), \tag{5}$$

with $\tilde{\boldsymbol{x}}_{k,t}^{(\mathrm{d})} \in \mathbb{C}^{M_k}$ the postulated data symbol vector of the kth user at time t. The GPME (3) coincides with an MMSE estimator when there exists no model mismatch, i.e., $\tilde{\boldsymbol{x}}_{k,t}^{(\mathrm{d})} \sim \boldsymbol{x}_{k,t}^{(\mathrm{d})}$, $\tilde{N}_0 = N_0$, and $p(\tilde{\boldsymbol{\mathcal{H}}}|\boldsymbol{\mathcal{I}}_{\mathrm{p}}) = p(\boldsymbol{\mathcal{H}}|\boldsymbol{\mathcal{I}}_{\mathrm{p}})$.

An LMMSE estimator is obtained by regarding $\tilde{\boldsymbol{x}}_{k,t}^{(\mathrm{d})}$ as $\tilde{\boldsymbol{x}}_{k,t}^{(\mathrm{L,d})} \sim \mathcal{CN}(\boldsymbol{0}, \mathrm{E}[\boldsymbol{x}_{k,t}^{(\mathrm{d})} \boldsymbol{x}_{k,t}^{(\mathrm{d})H}])$ when perfect CSI is assumed [9]. However, the GPME (3) is nonlinear even if the replacement of $\tilde{\boldsymbol{x}}_{k,t}^{(\mathrm{d})}$ by $\tilde{\boldsymbol{x}}_{k,t}^{(\mathrm{L,d})}$ is performed because the channel estimation errors act as a multiplicative noise. The GPME (3) is reduced to an LMMSE estimator if the term which includes the multiplicative noise is approximated by an AWGN term. Our approach is to replace the postulated channel (5) by

$$\tilde{\boldsymbol{y}}_{l,t}^{(\mathrm{L,d})} = \frac{1}{\sqrt{L}} \sum_{k=1}^{K} \left(\hat{\boldsymbol{H}}_k \boldsymbol{S}_{k,l}^{(\mathrm{d})} \tilde{\boldsymbol{x}}_{k,t}^{(\mathrm{L,d})} + \sum_{j=1}^{L_k} s_{k,j}^{(l,d)} \boldsymbol{w}_{k,j,t} \right) + \tilde{\boldsymbol{n}}_{l,t}^{(\mathrm{d})}, \tag{6}$$

where $\{\boldsymbol{w}_{k,j,t};\ k = 1, \ldots, K,\ j = 1, \ldots, L_k\}$ are zero-mean complex Gaussian random vectors with $\mathrm{E}[\boldsymbol{w}_{k,j,t} \boldsymbol{w}_{k',j',t}^H] = \delta_{k,k'} \mathrm{E}[\Delta\tilde{\boldsymbol{H}}_k \mathbf{\Gamma}_{k,j} \tilde{\boldsymbol{x}}_{k,t}^{(\mathrm{L,d})} (\Delta\tilde{\boldsymbol{H}}_k \mathbf{\Gamma}_{k,j'} \tilde{\boldsymbol{x}}_{k,t}^{(\mathrm{L,d})})^H | \boldsymbol{\mathcal{I}}_{\mathrm{p}}]$. In this case, the GPME (3) is reduced to an LMMSE estimator $\langle\langle \tilde{\boldsymbol{x}}_{k,t}^{(\mathrm{L,d})} \rangle\rangle_{\mathrm{L}}$ if $\tilde{N}_0 = N_0$ and $p(\tilde{\boldsymbol{\mathcal{H}}}|\boldsymbol{\mathcal{I}}_{\mathrm{p}}) = p(\boldsymbol{\mathcal{H}}|\boldsymbol{\mathcal{I}}_{\mathrm{p}})$.

C. Assumptions

We assume that $\boldsymbol{x}_{k,t}^{(\mathrm{p})}$, $\boldsymbol{x}_{k,t}^{(\mathrm{d})}$, $\tilde{\boldsymbol{x}}_{k,t}^{(\mathrm{d})}$ (or $\tilde{\boldsymbol{x}}_{k,t}^{(\mathrm{L,d})}$), $\boldsymbol{S}_{k,l}^{(\mathrm{p})}$, $\boldsymbol{S}_{k,l}^{(\mathrm{d})}$, and $\boldsymbol{H}_k$ are independent for all k. Furthermore, $\tilde{\boldsymbol{H}}_k$ conditioned on $\boldsymbol{\mathcal{I}}_{\mathrm{p}}$ is also assumed to be independent for all k. Note that $p(\boldsymbol{\mathcal{H}}|\boldsymbol{\mathcal{I}}_{\mathrm{p}})$ are not generally decomposed into $\prod_{k=1}^{K} p(\boldsymbol{H}_k|\boldsymbol{\mathcal{I}}_{\mathrm{p}})$. Therefore, the last assumption may result in performance loss.

III. Theoretical Results

The spectral efficiency of the GPME MUD (3), which is defined as $(LT_{\mathrm{c}})^{-1} \sum_{k=1}^{K} \sum_{t=\tau+1}^{T_{\mathrm{c}}} I(\boldsymbol{x}_{k,t}^{(\mathrm{d})}; \langle\langle \tilde{\boldsymbol{x}}_{k,t}^{(\mathrm{d})} \rangle\rangle | \boldsymbol{\mathcal{I}}_{\mathrm{p}}, \boldsymbol{\mathcal{S}}^{(\mathrm{d})})$, is determined by $p(\langle\langle \tilde{\boldsymbol{x}}_{k,t}^{(\mathrm{d})} \rangle\rangle | \boldsymbol{x}_{k,t}^{(\mathrm{d})}, \boldsymbol{\mathcal{I}}_{\mathrm{p}}, \boldsymbol{\mathcal{S}}^{(\mathrm{d})})$. We consider the large-system limit, in which K and L go to infinity with their ratio $\beta = K/L$ fixed, in order to evaluate it analytically. Furthermore, we assume that it has self-averaging property with respect to $\boldsymbol{\mathcal{S}}^{(\mathrm{d})}$: The distribution of $\langle\langle \tilde{\boldsymbol{x}}_{k,t}^{(\mathrm{d})} \rangle\rangle$ conditioned on $\boldsymbol{x}_{k,t}^{(\mathrm{d})}$, $\boldsymbol{\mathcal{I}}_{\mathrm{p}}$, and $\boldsymbol{\mathcal{S}}^{(\mathrm{d})}$ converges weakly to a distribution, which is independent of $\boldsymbol{\mathcal{S}}^{(\mathrm{d})}$, in the large-system limit. The main result is that $p(\langle\langle \tilde{\boldsymbol{x}}_{k,t}^{(\mathrm{d})} \rangle\rangle | \boldsymbol{x}_{k,t}^{(\mathrm{d})}, \boldsymbol{\mathcal{I}}_{\mathrm{p}}, \boldsymbol{\mathcal{S}}^{(\mathrm{d})})$ is reduced to a conditional probability density function which describes a single-user channel. Therefore, the spectral efficiency of the GPME MUD (3) is given as that of the single-user channel. We first introduce the single-user channel.

Definition 1: We define a space-time coded single-user MIMO channel as

$$\boldsymbol{y}_{k,j,t} = \boldsymbol{H}_k \mathbf{\Gamma}_{k,j} \boldsymbol{x}_{k,t}^{(\mathrm{d})} + \boldsymbol{n}_{k,j,t}, \quad \boldsymbol{n}_{k,j,t} \sim \mathcal{CN}(\boldsymbol{0}, \mathbf{\Sigma}). \tag{7}$$

The receiver estimates $\boldsymbol{x}_{k,t}^{(\mathrm{d})}$ on the basis of the GPME $\langle \tilde{\boldsymbol{x}}_{k,t}^{(\mathrm{d})} \rangle$, defined as

$$\langle\ \cdot\ \rangle = \frac{\mathrm{E}_{\tilde{\boldsymbol{x}}_{k,t}^{(\mathrm{d})}, \tilde{\boldsymbol{H}}_k} \left[\ \cdot\ \prod_{j=1}^{L_k} p(\tilde{\boldsymbol{y}}_{k,j,t} = \boldsymbol{y}_{k,j,t} | \tilde{\boldsymbol{H}}_k, \tilde{\boldsymbol{x}}_{k,t}^{(\mathrm{d})}) \middle| \boldsymbol{\mathcal{I}}_{\mathrm{p}} \right]}{\mathrm{E}_{\tilde{\boldsymbol{x}}_{k,t}^{(\mathrm{d})}, \tilde{\boldsymbol{H}}_k} \left[\prod_{j=1}^{L_k} p(\tilde{\boldsymbol{y}}_{k,j,t} = \boldsymbol{y}_{k,j,t} | \tilde{\boldsymbol{H}}_k, \tilde{\boldsymbol{x}}_{k,t}^{(\mathrm{d})}) \middle| \boldsymbol{\mathcal{I}}_{\mathrm{p}} \right]}, \tag{8}$$

where the postulated channel $p(\tilde{\boldsymbol{y}}_{k,j} | \tilde{\boldsymbol{H}}_k, \tilde{\boldsymbol{x}}_{k,t}^{(\mathrm{d})})$ is given by

$$\tilde{\boldsymbol{y}}_{k,j} = \tilde{\boldsymbol{H}}_k \mathbf{\Gamma}_{k,j} \tilde{\boldsymbol{x}}_{k,t}^{(\mathrm{d})} + \tilde{\boldsymbol{n}}_{k,j,t}, \quad \tilde{\boldsymbol{n}}_{k,j,t} \sim \mathcal{CN}(\boldsymbol{0}, \tilde{\mathbf{\Sigma}}). \tag{9}$$

The actual covariance matrix $\mathbf{\Sigma}$ and the postulated covariance matrix $\tilde{\mathbf{\Sigma}}$ are solutions to the pair of the coupled fixed-point equations:

$$\mathbf{\Sigma} = N_0 \boldsymbol{I}_N + \beta \lim_{K \to \infty} \frac{1}{K} \sum_{k=1}^{K} \sum_{j=1}^{L_k} \boldsymbol{\mathcal{E}}_{k,j,t}(\mathbf{\Sigma}, \tilde{\mathbf{\Sigma}}), \tag{10}$$

[1] The GPME (3) in this paper is different from Li and Poor's [5] because we treat the channel estimation errors explicitly. They considered effective sequence errors approximately.

$$\tilde{\mathbf{\Sigma}} = \tilde{N}_0 \boldsymbol{I}_N + \beta \lim_{K\to\infty} \frac{1}{K} \sum_{k=1}^{K} \sum_{j=1}^{L_k} \boldsymbol{\mathcal{V}}_{k,j,t}(\mathbf{\Sigma}, \tilde{\mathbf{\Sigma}}), \tag{11}$$

where the error covariance matrix $\boldsymbol{\mathcal{E}}_{k,j,t}(\mathbf{\Sigma}, \tilde{\mathbf{\Sigma}})$ and the posterior covariance matrix $\boldsymbol{\mathcal{V}}_{k,j,t}(\mathbf{\Sigma}, \tilde{\mathbf{\Sigma}})$ of the GPME $\langle \tilde{\boldsymbol{H}}_k \mathbf{\Gamma}_{k,j} \tilde{\boldsymbol{x}}_{k,t}^{(\mathrm{d})} \rangle$ are defined as

$$\boldsymbol{\mathcal{E}}_{k,j,t}(\mathbf{\Sigma}, \tilde{\mathbf{\Sigma}}) = \mathrm{cov}\left[\boldsymbol{H}_k \mathbf{\Gamma}_{k,j} \boldsymbol{x}_{k,t}^{(\mathrm{d})} - \left\langle \tilde{\boldsymbol{H}}_k \mathbf{\Gamma}_{k,j} \tilde{\boldsymbol{x}}_{k,t}^{(\mathrm{d})} \right\rangle \middle| \boldsymbol{\mathcal{I}}_{\mathrm{p}} \right], \tag{12}$$

$$\boldsymbol{\mathcal{V}}_{k,j,t}(\mathbf{\Sigma}, \tilde{\mathbf{\Sigma}}) = \mathrm{cov}\left[\tilde{\boldsymbol{H}}_k \mathbf{\Gamma}_{k,j} \tilde{\boldsymbol{x}}_{k,t}^{(\mathrm{d})} - \left\langle \tilde{\boldsymbol{H}}_k \mathbf{\Gamma}_{k,j} \tilde{\boldsymbol{x}}_{k,t}^{(\mathrm{d})} \right\rangle \middle| \boldsymbol{\mathcal{I}}_{\mathrm{p}} \right]. \tag{13}$$

If there exist multiple solutions of $(\mathbf{\Sigma}, \tilde{\mathbf{\Sigma}})$, one should choose the solution which minimizes the so-called free energy:

$$\frac{\beta}{\ln 2}\mathcal{F} = \lim_{K\to\infty} \frac{1}{K} \sum_{k=1}^{K} \beta \tilde{\mathcal{C}}_k + F(\mathbf{\Sigma}, \tilde{\mathbf{\Sigma}}) + N \log(\pi N_0 \mathrm{e}), \tag{14}$$

where the function F is defined as $F(\mathbf{\Sigma}, \tilde{\mathbf{\Sigma}}) = \mathrm{KL}(N_0 \boldsymbol{I}_N \| \tilde{\mathbf{\Sigma}}) + \mathrm{KL}(\mathbf{\Sigma} \| \tilde{\mathbf{\Sigma}}) + \mathrm{KL}(\tilde{N}_0 \boldsymbol{I}_N \| \tilde{\mathbf{\Sigma}}) - \mathrm{KL}(\tilde{N}_0 \boldsymbol{I}_N \| \tilde{\mathbf{\Sigma}} \mathbf{\Sigma}^{-1} \tilde{\mathbf{\Sigma}})$, and where $\tilde{\mathcal{C}}_k$ is given by

$$\tilde{\mathcal{C}}_k = \mathrm{E}\left[\log \frac{\prod_{j=1}^{L_k} \tilde{p}(\boldsymbol{y}_{k,j,t} | \boldsymbol{H}_k, \boldsymbol{x}_{k,t}^{(\mathrm{d})})}{\mathrm{E}_{\tilde{\boldsymbol{x}}_{k,t}^{(\mathrm{d})}, \tilde{\boldsymbol{H}}_k}\left[\prod_{j=1}^{L_k} p(\boldsymbol{y}_{k,j,t} | \tilde{\boldsymbol{H}}_k, \tilde{\boldsymbol{x}}_{k,t}^{(\mathrm{d})}) \middle| \boldsymbol{\mathcal{I}}_{\mathrm{p}} \right]} \middle| \boldsymbol{\mathcal{I}}_{\mathrm{p}} \right], \tag{15}$$

with $\tilde{p}(\boldsymbol{y}_{k,j,t} | \boldsymbol{H}_k, \boldsymbol{x}_{k,t}^{(\mathrm{d})}) = p(\tilde{\boldsymbol{y}}_{k,j,t} = \boldsymbol{y}_{k,j,t} | \tilde{\boldsymbol{H}}_k = \boldsymbol{H}_k, \tilde{\boldsymbol{x}}_{k,t}^{(\mathrm{d})} = \boldsymbol{x}_{k,t}^{(\mathrm{d})})$.

Proposition 1: The distribution of the GPME (3) conditioned on $\boldsymbol{x}_{k,t}^{(\mathrm{d})}, \boldsymbol{\mathcal{I}}_{\mathrm{p}}$, and $\boldsymbol{\mathcal{S}}^{(\mathrm{d})}$ converges weakly to the distribution of $\langle \tilde{\boldsymbol{x}}_{k,t}^{(\mathrm{d})} \rangle$ conditioned on $\boldsymbol{x}_{k,t}^{(\mathrm{d})}$ and $\boldsymbol{\mathcal{I}}_{\mathrm{p}}$ in the large-system limit.

Proposition 1 is derived in Section V. It reveals the decoupling structure of the MIMO DS-CDMA channel with imperfect CSI: It is decoupled into a bank of the single-user space-time coded MIMO channels with imperfect CSI.

We next analyze the corresponding conditional distribution for the case of the linear estimator $\langle \langle \tilde{\boldsymbol{x}}_{k,t}^{(\mathrm{d})} \rangle \rangle_{\mathrm{L}}$.

Definition 2: We define the postulated channel by the receiver as

$$\tilde{\boldsymbol{y}}_{k,j,t}^{(\mathrm{L})} = \boldsymbol{a}_{k,j,t} + \tilde{\boldsymbol{n}}_{k,j,t}^{(\mathrm{L})}, \quad \tilde{\boldsymbol{n}}_{k,j,t}^{(\mathrm{L})} \sim \mathcal{CN}(\mathbf{0}, \tilde{\mathbf{\Sigma}}), \tag{16}$$

where $\boldsymbol{a}_{k,j,t} \in \mathbb{C}^N$ is a random vector defined as $\boldsymbol{a}_{k,j,t} = \hat{\boldsymbol{H}}_k \mathbf{\Gamma}_{k,j} \tilde{\boldsymbol{x}}_{k,t}^{(\mathrm{L,d})} + \boldsymbol{w}_{k,j,t}$. The receiver estimates $\boldsymbol{x}_{k,t}^{(\mathrm{d})}$ on the basis of the linear estimator

$$\langle \cdot \rangle_{\mathrm{L}} = \frac{\mathrm{E}_{\tilde{\boldsymbol{x}}_{k,t}^{(\mathrm{L,d})}, \{\boldsymbol{w}_{k,j,t}\}}\left[\cdot \prod_{j=1}^{L_k} p(\tilde{\boldsymbol{y}}_{k,j,t}^{(\mathrm{L})} = \boldsymbol{y}_{k,j,t} | \boldsymbol{a}_{k,j,t}) \middle| \boldsymbol{\mathcal{I}}_{\mathrm{p}} \right]}{\mathrm{E}_{\tilde{\boldsymbol{x}}_{k,t}^{(\mathrm{L,d})}, \{\boldsymbol{w}_{k,j,t}\}}\left[\prod_{j=1}^{L_k} p(\tilde{\boldsymbol{y}}_{k,j,t}^{(\mathrm{L})} = \boldsymbol{y}_{k,j,t} | \boldsymbol{a}_{k,j,t}) \middle| \boldsymbol{\mathcal{I}}_{\mathrm{p}} \right]}, \tag{17}$$

where $\boldsymbol{y}_{k,j,t}$ is given by (7). $\mathbf{\Sigma}$ and $\tilde{\mathbf{\Sigma}}$ satisfy the fixed-point equations (10) and (11) with $\boldsymbol{\mathcal{E}}_{k,j,t}$ and $\boldsymbol{\mathcal{V}}_{k,j,t}$ replaced by

$$\boldsymbol{\mathcal{E}}_{k,j,t}^{(\mathrm{L})}(\mathbf{\Sigma}, \tilde{\mathbf{\Sigma}}) = \mathrm{cov}\left[\boldsymbol{H}_k \mathbf{\Gamma}_{k,j} \boldsymbol{x}_{k,t}^{(\mathrm{L,d})} - \langle \boldsymbol{a}_{k,j,t} \rangle_{\mathrm{L}} \middle| \boldsymbol{\mathcal{I}}_{\mathrm{p}} \right], \tag{18}$$

$$\boldsymbol{\mathcal{V}}_{k,j,t}^{(\mathrm{L})}(\mathbf{\Sigma}, \tilde{\mathbf{\Sigma}}) = \mathrm{cov}\left[\boldsymbol{a}_{k,j,t} - \langle \boldsymbol{a}_{k,j,t} \rangle_{\mathrm{L}} \middle| \boldsymbol{\mathcal{I}}_{\mathrm{p}} \right]. \tag{19}$$

Proposition 2: The distribution of $\langle \langle \tilde{\boldsymbol{x}}_{k,t}^{(\mathrm{L,d})} \rangle \rangle_{\mathrm{L}}$ conditioned on $\boldsymbol{x}_{k,t}^{(\mathrm{d})}$, $\boldsymbol{\mathcal{I}}_{\mathrm{p}}$, and $\boldsymbol{\mathcal{S}}^{(\mathrm{d})}$ converges weakly to that of $\langle \tilde{\boldsymbol{x}}_{k,t}^{(\mathrm{L,d})} \rangle_{\mathrm{L}}$ conditioned on $\boldsymbol{x}_{k,t}^{(\mathrm{d})}$ and $\boldsymbol{\mathcal{I}}_{\mathrm{p}}$ in the large-system limit.

The derivation of Proposition 2 is omitted because it is almost the same as that of Proposition 1.

IV. Numerical Results

In numerical examples shown in this section, we assume that $\boldsymbol{x}_{k,t}^{(\mathrm{p})}$ and $\boldsymbol{x}_{k,t}^{(\mathrm{d})}$ have i.i.d elements, and that $\boldsymbol{H}_k$ are i.i.d. entries $(\boldsymbol{H}_k)_{nm} \sim \mathcal{CN}(0,1)$ (i.i.d Rayleigh fading). We also assume $M_k = M$ and $P_k = P$, and consider, for simplicity, only the (L)MMSE channel estimator, quadrature phase-shift-keying (QPSK) modulation, and the spreading scheme using different spreading sequences for different antennas [1]–[3], which can be dealt with by letting $\mathbf{\Gamma}_{k,j} = \mathrm{diag}(\mathbf{0}_{j-1},\ 1,\ \mathbf{0}_{M-j})$ for $j = 1, \ldots, M$.

Let us assume that the second terms of the right-hand sides of (10) and (11) have self-averaging property, i.e., they converge in probability to their expectations with respect to $\boldsymbol{\mathcal{I}}_{\mathrm{p}}$. Then, we have to evaluate $p(\hat{\boldsymbol{H}}_k)$ and $p(\Delta\tilde{\boldsymbol{H}}_k)$ in order to solve (10) and (11) because $\tilde{\boldsymbol{H}}_k$ depends on $\boldsymbol{\mathcal{Y}}^{(\mathrm{p})}$ only through $\hat{\boldsymbol{H}}_k$ and because $\Delta\tilde{\boldsymbol{H}}_k$ is uncorrelated with $\hat{\boldsymbol{H}}_k$ for the LMMSE channel estimator.

Lemma 1 (LMMSE Channel Estimation): In the large-system limit,

$$p(\hat{\boldsymbol{H}}_k) = \left(\frac{1 + \xi^2(\sigma_{\mathrm{p}}^2)}{\pi} \right)^{MN} \mathrm{e}^{-[1+\xi^2(\sigma_{\mathrm{p}}^2)]\|\hat{\boldsymbol{H}}_k\|^2}, \tag{20}$$

$$p(\Delta\tilde{\boldsymbol{H}}_k) = \frac{1}{(\pi \mathrm{MSE})^{MN}} \mathrm{e}^{-\frac{\|\Delta \boldsymbol{H}_k\|^2}{\mathrm{MSE}}}, \tag{21}$$

with $\xi^2(\sigma_{\mathrm{p}}^2) = M\sigma_{\mathrm{p}}^2/(\tau P)$ and $\mathrm{MSE} = \xi^2(\sigma_{\mathrm{p}}^2)/[1 + \xi^2(\sigma_{\mathrm{p}}^2)]$. σ_{p}^2 is the unique solution of the fixed-point equation

$$\sigma_{\mathrm{p}}^2 = N_0 + \beta \frac{P\xi^2(\sigma_{\mathrm{p}}^2)}{1 + \xi^2(\sigma_{\mathrm{p}}^2)}. \tag{22}$$

The proof of Lemma 1 is omitted since it is derived in the same manner as Proposition 1.

Figure 1 displays the multiuser efficiencies η, which are defined as $\mathbf{\Sigma} = (N_0/\eta)\boldsymbol{I}_N$, of the MMSE and LMMSE MUDs. There exist multiple solutions of (10) and (11) for the case of the MMSE MUD ($\beta = 2.1$). The solid line shows the solution which minimizes the free energy (14). The numerical result implies that the multiuser efficiency of the MMSE MUD exhibits a *waterfall* behavior for high β, in which its critical point β_{c} depends on M, N, τ, T_{c}, and P/N_0. The multiuser efficiency of the MMSE MUD is much larger than that of the LMMSE MUD if the ratio τ/T_{c} of the pilot sequence length to the coherence time is sufficiently large. However, it falls down toward that of the LMMSE MUD at the critical point where the correct solution changes.

The use of short pilot sequences leads to poor performance of one-shot multiuser detection. On the other hand, long pilot sequences cause reduced payload, leading to a low data rate. Therefore, there exists an optimal length of the pilot

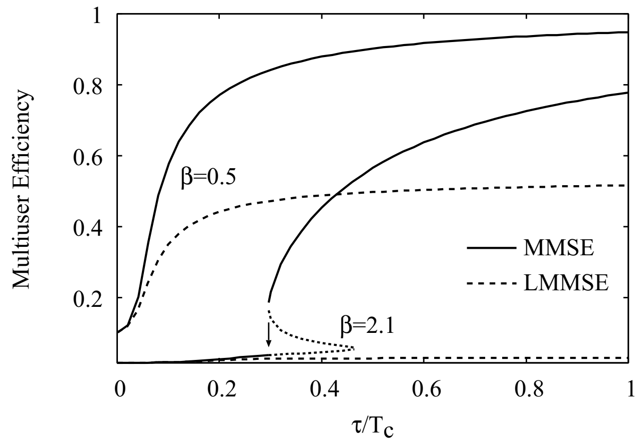

Fig. 1. Multiuser efficiency versus τ/T_{c} for $M = N = 4$, $T_{\mathrm{c}} = 40$, and $P/N_0 = 15$ dB. The solid and dashed lines represent the multiuser efficiencies of the MMSE and LMMSE MUDs, respectively. The dotted line shows solutions for the MMSE MUD which are not chosen by the criterion described in the main text.

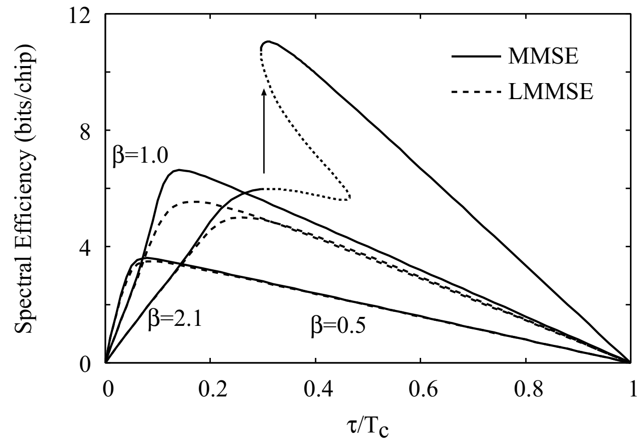

Fig. 2. Spectral efficiency versus τ/T_{c} for $M = N = 4$, $T_{\mathrm{c}} = 40$, and $P/N_0 = 15$ dB. The solid and dashed lines represent the spectral efficiencies of the MMSE and LMMSE MUDs, respectively. The dotted line shows solutions for the MMSE MUD which are not chosen by the criterion described in the main text.

sequences. Figure 2 shows the spectral efficiencies of the MMSE and LMMSE MUDs. The optimal length of the pilot sequences is short for small β, and the performance gap between the MMSE and LMMSE MUDs is marginal even though the computational complexity of the LMMSE MUD is much lower than that of the MMSE MUD. On the other hand, the optimal length increases as β grows, and the gap becomes significantly large. Furthermore, the optimal length for the MMSE MUD with high β is given by a slightly larger value than the critical point.

V. Derivation of Proposition 1

We define the free energy $\mathcal{F}$ as

$$\mathcal{F} = -\lim_{K=\beta L\to\infty}\lim_{u\to 0}\frac{\partial}{\partial u}\frac{1}{K}\ln \Xi^{(u)}, \tag{23}$$

where the moment generating function $\Xi^{(u)}$ is given by[2]

$$\Xi^{(u)} = \mathrm{E}\left[\left\{\mathrm{E}_{\tilde{\boldsymbol{\mathcal{X}}}^{(\mathrm{d})}}\left[p(\tilde{\boldsymbol{\mathcal{Y}}}^{(\mathrm{d})} = \boldsymbol{\mathcal{Y}}^{(\mathrm{d})}|\tilde{\boldsymbol{\mathcal{X}}}^{(\mathrm{d})}, \boldsymbol{\mathcal{I}}_{\mathrm{p}}, \boldsymbol{\mathcal{S}}^{(\mathrm{d})})\right]\right\}^u \middle| \boldsymbol{\mathcal{I}}_{\mathrm{p}}\right], \tag{24}$$

with $\tilde{\boldsymbol{\mathcal{Y}}}^{(\mathrm{d})} = \{\boldsymbol{\mathcal{Y}}_t^{(\mathrm{d})}\}_{t=\tau+1}^{T_{\mathrm{c}}}$, $\boldsymbol{\mathcal{Y}}^{(\mathrm{d})} = \{\boldsymbol{\mathcal{Y}}_t^{(\mathrm{d})}\}_{t=\tau+1}^{T_{\mathrm{c}}}$, and $\tilde{\boldsymbol{\mathcal{X}}}^{(\mathrm{d})} = \{\tilde{\boldsymbol{\mathcal{X}}}_t^{(\mathrm{d})}\}_{t=\tau+1}^{T_{\mathrm{c}}}$. For notational convenience, we drop the superscript $^{(\mathrm{d})}$ hereafter. We evaluate the free energy (23) using the replica method, in order to obtain Proposition 1.

Assume for a moment that u is a nonnegative integer. Let us define an $N \times \tau'$ matrix $\boldsymbol{V}_\alpha$ as

$$\boldsymbol{V}_\alpha = \frac{1}{\sqrt{K}}\sum_{k=1}^{K}\boldsymbol{H}_{k,\alpha}\boldsymbol{S}_k\boldsymbol{X}_{k,\alpha}, \quad \alpha = 0, 1, \ldots, u, \tag{25}$$

where $\boldsymbol{H}_{k,0} = \boldsymbol{H}_k$, $\boldsymbol{X}_{k,0} = (\boldsymbol{x}_{k,\tau+1}, \ldots, \boldsymbol{x}_{k,T_{\mathrm{c}}})$ and $\boldsymbol{X}_{k,\alpha} \sim (\tilde{\boldsymbol{x}}_{k,\tau+1}, \ldots, \tilde{\boldsymbol{x}}_{k,T_{\mathrm{c}}})$, $\boldsymbol{H}_{k,\alpha} \sim \tilde{\boldsymbol{H}}_k$ conditioned on $\boldsymbol{\mathcal{I}}_{\mathrm{p}}$ for $\alpha = 1, \ldots, u$, and where $\boldsymbol{S}_k$ follows $p(\boldsymbol{S}_{k,l}^{(\mathrm{d})})$. $\boldsymbol{H}_{k,\alpha}$, $\boldsymbol{S}_k$, and $\boldsymbol{X}_{k,\alpha}$ are independent for $\alpha = 1, \ldots, u$ and all k. Equation (24) depends on the channel matrices, the data symbol matrices, and the spatial spreading matrices only through $\boldsymbol{V}_\alpha$. Let us assume that the large-system limit and the operations with respect to u commute in (23). From the asymptotic normality of $\boldsymbol{V} = (\boldsymbol{V}_0, \ldots, \boldsymbol{V}_u)$ conditioned on $\{\boldsymbol{X}_{k,\alpha}\}$ and $\{\boldsymbol{H}_{k,\alpha}\}$ [11], (24) depends on them only via the covariance matrix

$$\begin{aligned}\boldsymbol{\mathcal{Q}} &= \mathrm{E}[\mathrm{vec}(\boldsymbol{V})\mathrm{vec}(\boldsymbol{V})^H|\{\boldsymbol{X}_{k,\alpha}\}, \{\boldsymbol{H}_{k,\alpha}\}] \\ &= \frac{1}{K}\sum_{k=1}^{K}\sum_{j=1}^{L_k}\mathrm{vec}(\boldsymbol{\Omega}_{k,j})\mathrm{vec}(\boldsymbol{\Omega}_{k,j})^H, \end{aligned}\tag{26}$$

with $\boldsymbol{\Omega}_{k,j} = (\boldsymbol{H}_{k,0}\boldsymbol{\Gamma}_{k,j}\boldsymbol{X}_{k,0}, \ldots, \boldsymbol{H}_{k,u}\boldsymbol{\Gamma}_{k,j}\boldsymbol{X}_{k,u})$. After some calculations [3], we obtain

$$\frac{1}{K}\ln \Xi^{(u)} = \frac{1}{K}\ln \mathrm{E}\left\{\mathrm{e}^{K\beta^{-1}G^{(u)}(\boldsymbol{\mathcal{Q}})}\middle|\boldsymbol{\mathcal{I}}_{\mathrm{p}}\right\} + o(K), \tag{27}$$

where $G^{(u)}(\boldsymbol{\mathcal{Q}})$ is given by

$$\begin{aligned}G^{(u)}(\boldsymbol{\mathcal{Q}}) = &-\ln\det(\boldsymbol{I} + \boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{Q}}) - \tau' N u \ln(\pi\tilde{N}_0) \\ &-\tau' N \ln\left(1 + \frac{N_0}{\tilde{N}_0}u\right),\end{aligned}\tag{28}$$

with

$$\boldsymbol{\mathcal{A}} = \frac{\beta}{\tilde{N}_0 + uN_0}\begin{pmatrix} u & -\boldsymbol{e}_u^T \\ -\boldsymbol{e}_u & \left(1 + \frac{uN_0}{\tilde{N}_0}\right)\boldsymbol{I}_u - \frac{N_0}{\tilde{N}_0}\boldsymbol{e}_u\boldsymbol{e}_u^T\end{pmatrix} \otimes \boldsymbol{I}_{\tau' N}. \tag{29}$$

Evaluating the expectation in (27) with the saddle-point method, we have

$$\lim_{K=\beta L\to\infty}\frac{1}{K}\ln \Xi^{(u)} = \sup_{\boldsymbol{\mathcal{Q}}}\left[\beta^{-1}G^{(u)}(\boldsymbol{\mathcal{Q}}) - I^{(u)}(\boldsymbol{\mathcal{Q}})\right], \tag{30}$$

[2]This definition is needed in the derivation of Lemma 1. Proposition 1 is obtained formally by letting $\tau' = T_{\mathrm{c}} - \tau$ be one. Let $\boldsymbol{H}_{k,a} \sim \boldsymbol{H}_k$ and $\boldsymbol{X}_{k,a} \sim \delta(\boldsymbol{X}_{k,a} - (\boldsymbol{x}_{k,1}^{(\mathrm{p})}, \ldots, \boldsymbol{x}_{k,\tau}^{(\mathrm{p})}))$ in order to show Lemma 1.

$$\mathcal{M}_k^{(u)}(\tilde{\mathcal{Q}}^{\mathrm{s}}) = \mathrm{E}\left[\int \prod_{j=1}^{L_k} q(\boldsymbol{Y}_{k,j}|\boldsymbol{H}_k\boldsymbol{\Gamma}_{k,j}\boldsymbol{X}_{k,0};\boldsymbol{\Sigma}) \left\{ \frac{\mathrm{E}_{\tilde{\boldsymbol{X}}_{k,1},\tilde{\boldsymbol{H}}_k}\left[\prod_{j=1}^{L_k} q(\boldsymbol{Y}_{k,j}|(\tilde{\boldsymbol{H}}_k\boldsymbol{\Gamma}_{k,j}\tilde{\boldsymbol{X}}_{k,1};\tilde{\boldsymbol{\Sigma}}) \middle| \boldsymbol{\mathcal{I}}_{\mathrm{p}} \right]}{\prod_{j=1}^{L_k} q(\boldsymbol{Y}_{k,j}|\boldsymbol{0};\tilde{\boldsymbol{\Sigma}})} \right\}^u \prod_{j=1}^{L_k} d\boldsymbol{Y}_{k,j} \middle| \boldsymbol{\mathcal{I}}_{\mathrm{p}} \right], \tag{39}$$

where the rate function $I^{(u)}(\mathcal{Q})$ is given by

$$I^{(u)}(\mathcal{Q}) = \sup_{\tilde{\mathcal{Q}}} \left[\mathrm{Tr}(\tilde{\mathcal{Q}}\mathcal{Q}) - \lim_{K\to\infty} \frac{1}{K} \sum_{k=1}^{K} \ln \mathcal{M}_k^{(u)}(\tilde{\mathcal{Q}}) \right], \tag{31}$$

with a $(u+1)N\tau' \times (u+1)N\tau'$ Hermitian matrix $\tilde{\mathcal{Q}}$ and with the moment generating function

$$\mathcal{M}_k^{(u)}(\tilde{\mathcal{Q}}) = \mathrm{E}\left\{ \exp\left[\Lambda_k(\tilde{\mathcal{Q}}) \right] \middle| \boldsymbol{\mathcal{I}}_{\mathrm{p}} \right\},$$

$$\Lambda_k(\tilde{\mathcal{Q}}) = \sum_{j=1}^{L_k} \mathrm{Tr}\left[\tilde{\mathcal{Q}} \mathrm{vec}(\boldsymbol{\Omega}_{k,j}) \mathrm{vec}(\boldsymbol{\Omega}_{k,j})^H \right]. \tag{32}$$

From the stationarity conditions in (30) and (31), we obtain the fixed-point equations:

$$\tilde{\mathcal{Q}}^{\mathrm{s}} = -\beta^{-1}(\boldsymbol{I} + \boldsymbol{\mathcal{A}}\mathcal{Q}^{\mathrm{s}})^{-1}\boldsymbol{\mathcal{A}}, \tag{33}$$

$$\mathcal{Q}^{\mathrm{s}} = \lim_{K\to\infty} \frac{1}{K} \sum_{k=1}^{K} \left\langle \sum_{j=1}^{L_k} \mathrm{vec}(\boldsymbol{\Omega}_{k,j}) \mathrm{vec}(\boldsymbol{\Omega}_{k,j})^H \right\rangle, \tag{34}$$

with $\langle \,\cdot\, \rangle = \mathcal{M}_k^{(u)}(\tilde{\mathcal{Q}}^{\mathrm{s}})^{-1} \mathrm{E}[\,\cdot\, \mathrm{e}^{\Lambda_k(\tilde{\mathcal{Q}}^{\mathrm{s}})} | \boldsymbol{\mathcal{I}}_{\mathrm{p}}]$.

At this point, we assume replica symmetry, i.e., $\boldsymbol{Q}_{\alpha\alpha'} = \mathrm{E}[\mathrm{vec}(\boldsymbol{V}_\alpha)\mathrm{vec}(\boldsymbol{V}_{\alpha'})^H | \{\boldsymbol{X}_{k,\alpha}\}, \{\boldsymbol{H}_{k,\alpha}\}]$ and $\tilde{\boldsymbol{Q}}_{\alpha\alpha'}$, the latter of which is a block of the block matrix $\tilde{\mathcal{Q}}^{\mathrm{s}}$ and satisfies $\mathrm{Tr}(\tilde{\mathcal{Q}}^{\mathrm{s}}\mathcal{Q}^{\mathrm{s}}) = \sum_{\alpha,\alpha'} \mathrm{Tr}(\tilde{\boldsymbol{Q}}_{\alpha\alpha'}\boldsymbol{Q}_{\alpha\alpha'})$, are invariant under exchange of non-zero replica indexes [3]. From (33), we obtain, in the limit $u \to 0$,

$$\tilde{\boldsymbol{Q}}_{00} = \boldsymbol{0},\ \tilde{\boldsymbol{M}} = \tilde{\boldsymbol{\Sigma}}^{-1},\ \tilde{\boldsymbol{Q}}_{11} = \tilde{\boldsymbol{Q}} - \tilde{\boldsymbol{M}},\ \tilde{\boldsymbol{Q}} = \tilde{\boldsymbol{\Sigma}}^{-1}\boldsymbol{\Sigma}\tilde{\boldsymbol{\Sigma}}^{-1}, \tag{35}$$

with $\tilde{\boldsymbol{M}} = \tilde{\boldsymbol{Q}}_{01}$ and a Hermitian matrix $\tilde{\boldsymbol{Q}} = \tilde{\boldsymbol{Q}}_{12}$. $\boldsymbol{\Sigma}$ and $\tilde{\boldsymbol{\Sigma}}$ are respectively given by

$$\boldsymbol{\Sigma} = N_0\boldsymbol{I}_{\tau' N} + \beta(\boldsymbol{Q}_{00} - \boldsymbol{M} - \boldsymbol{M}^H + \boldsymbol{Q}), \tag{36}$$

$$\tilde{\boldsymbol{\Sigma}} = \tilde{N}_0\boldsymbol{I}_{\tau' N} + \beta(\boldsymbol{Q}_{11} - \boldsymbol{Q}), \tag{37}$$

with $\boldsymbol{M} = \boldsymbol{Q}_{01}$ and a Hermitian matrix $\boldsymbol{Q} = \boldsymbol{Q}_{12}$. On the other hand, (28) yields

$$G^{(u)}(\mathcal{Q}^{\mathrm{s}}) = -(u-1)\ln\det\tilde{\boldsymbol{\Sigma}} - \ln\det(\tilde{\boldsymbol{\Sigma}} + u\boldsymbol{\Sigma}) - u\tau' N \ln\pi. \tag{38}$$

We next evaluate the moment generating function (32). Applying a transform that is similar to the Hubbard-Stratonovich transform to (32) after some algebra [3], we have (39) at the top of this page. The function $q(\boldsymbol{Y}|\boldsymbol{X};\boldsymbol{\Sigma})$ is defined as

$$q(\boldsymbol{Y}|\boldsymbol{X};\boldsymbol{\Sigma}) = \frac{1}{\pi^{\tau' N}\det\boldsymbol{\Sigma}} \mathrm{e}^{-\mathrm{vec}(\boldsymbol{Y}-\boldsymbol{X})^H\boldsymbol{\Sigma}^{-1}\mathrm{vec}(\boldsymbol{Y}-\boldsymbol{X})}. \tag{40}$$

We find that $\langle \,\cdot\, \rangle$ is reduced to (8) in $u \to 0$ and $\tau' = 1$ in the same manner. Evaluating (36) and (37) with (34), we obtain the fixed-point equations (10) and (11) by letting $\tau' = 1$.

The cumulant generating function (30) is well-defined for real u because (38), (39) and $\mathrm{Tr}(\tilde{\mathcal{Q}}^{\mathrm{s}}\mathcal{Q}^{\mathrm{s}})$ are well-defined. We assume that (30) for real u coincides with the true one in the neighborhood of $u = 0$. After some algebra, we find that the free energy (23) is given as (14) by letting $\tau' = 1$. If there exist multiple solutions of the fixed-point equations (10) and (11), one should choose the solution that yields the supremum of equation (30), which minimizes the free energy.

The expression of the free energy (14) implies that the MIMO DS-CDMA channel (2) is decoupled into a bank of the single-user MIMO channels (7). We can confirm that this statement is correct, by evaluating every moments of $p(\langle\langle \tilde{\boldsymbol{x}}_{k,t}^{(\mathrm{d})} \rangle\rangle | \boldsymbol{x}_{k,t}^{(\mathrm{d})}, \boldsymbol{\mathcal{I}}_{\mathrm{p}}, \boldsymbol{\mathcal{S}}^{(\mathrm{d})})$ with the replica method [3], although its calculation is omitted due to space limitation.

Acknowledgment

Tanaka acknowledges support through Grant-in-Aid for Scientific Research on Priority Areas (No. 18079010) from MEXT, Japan.

References

[1] A. Mantravadi, V. V. Veeravalli, and H. Viswanathan, "Spectral efficiency of MIMO multiaccess systems with single-user decoding," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 3, pp. 382–394, Apr. 2003.

[2] H. Li and H. V. Poor, "Spectral efficiency of equal-rate DS-CDMA systems with multiple transmit antennas," *IEEE Trans. Wireless Commun.*, vol. 5, no. 12, pp. 3680–3688, Dec. 2006.

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

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

[5] H. Li and H. V. Poor, "Impact of channel estimation errors on multiuser detection via the replica method," *EURASIP J. Wirel. Commun. Netw.*, vol. 2, pp. 175–186, 2005.

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

[7] A. L. Moustakas, S. H. Simon, and A. M. Sengupta, "MIMO capacity through correlated channels in the presence of correlated interferers and noise: A (not so) large N analysis," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2545–2561, Oct. 2003.

[8] R. R. Müller, "Channel capacity and minimum probability of error in large dual antenna array systems with binnary modulation," *IEEE Trans. Signal Process.*, vol. 51, no. 11, pp. 2821–2828, Nov. 2003.

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

[10] K. Takeda, S. Uda, and Y. Kabashima, "Analysis of CDMA systems that are characterized by eigenvalue spectrum," *Europhys. Lett.*, vol. 76, no. 6, pp. 1193–1199, 2006.

[11] K. Takeuchi and T. Tanaka, "Statistical-mechanics-based analysis of multiuser MIMO channels with linear dispersion codes," *J. Phys.: Conf. Ser.*, vol. 95, pp. 012 008–1–11, 2008.