

Microscopic Analysis for Decoupling Principle of Linear Vector Channel

Kazutaka Nakamura and Toshiyuki Tanaka

Department of Systems Science, Graduate School of Informatics, Kyoto University

36-1, Yoshida-Honmachi, Sakyo-ku, Kyoto, 606-8501, Japan

Email: {knakamur, tt}@i.kyoto-u.ac.jp

Abstract—This paper studies the decoupling principle of a linear vector channel, which is an extension of CDMA and MIMO channels. We show that the scalar-channel characterization obtained via the decoupling principle is valid not only for collections of a large number of elements of input vector, as discussed in previous studies, but also for individual elements of input vector, i.e. the linear vector channel for individual elements of channel input vector is decomposed into a bank of independent scalar Gaussian channels in the large-system limit, where dimensions of channel input and output are both sent to infinity while their ratio fixed.

I. INTRODUCTION

Recently, the replica method, developed in statistical mechanics, has been applied to problems of performance evaluation of various digital wireless communication systems, especially code-division multiple-access (CDMA) and multi-input multi-output (MIMO) systems [1]–[4]. The replica method provides us with a description of these channels, called, the *decoupling principle*; that is, a CDMA channel, or equivalently a MIMO channel, is decoupled, under a certain randomness assumption of the channel, into a bank of independent scalar Gaussian channels in the large-system limit, where dimensions of channel input and output are both sent to infinity while their ratio fixed.

Existing results of replica analysis, however, rely on saddle-point evaluation of integrals, which is only valid for evaluating *macroscopic* quantities, such as an empirical mean of many *microscopic* quantities, such as individual elements of input, which are many in the sense that their number goes to infinity as the dimensions of the system in the large-system limit. It is therefore not clear as to whether the scalar-channel characterization of CDMA or MIMO channels obtained via the replica analysis is still valid if we are interested in microscopic quantities in the large-system limit.

In this paper we show that the scalar-channel characterization is still valid for microscopic quantities, by performing replica analysis on a linear vector channel, which is an extension of CDMA or MIMO channels.

II. LINEAR VECTOR CHANNEL

We consider a K -input N -output linear vector channel, defined as follows. Let $\mathbf{x}_0 = (x_{01}, \dots, x_{0K})^T$ denote the input vector of the channel, and $\mathbf{y} = (y_1, \dots, y_N)^T$ denote the output vector given a linear transform $H\mathbf{x}_0$ of the inputs, where H is an $N \times K$ channel matrix. Assuming the channel

to be memoryless, the input-output characteristic of the linear vector channel is represented as

$$P_0(\mathbf{y}|H\mathbf{x}_0) = \prod_{\mu=1}^N \rho_0 \left(y_\mu \left| \frac{\mathbf{h}_\mu^T \mathbf{x}_0}{\sqrt{N}} \right| \right), \quad (1)$$

where $\mathbf{h}_\mu^T/\sqrt{N}$ denotes μ th row of H . We define a true prior as $P_0(\mathbf{x})$. Inference of the input vector \mathbf{x}_0 , given the output vector \mathbf{y} and the channel matrix H , can be solved by a detection scheme based on Bayesian inference. The detector assumes a channel model to be $P(\mathbf{y}|H\mathbf{x}) = \prod_{\mu=1}^N \rho(y_\mu|\mathbf{h}_\mu^T \mathbf{x}/\sqrt{N})$, and a prior distribution to be $P(\mathbf{x})$. We also assume perfect channel state information at the detector. These assumptions yield the posterior distribution

$$P(\mathbf{x}|\mathbf{y}, H) = \frac{P(\mathbf{y}|H\mathbf{x})P(\mathbf{x})}{\int P(\mathbf{y}|H\mathbf{x})P(\mathbf{x})d\mathbf{x}}. \quad (2)$$

The posterior mean estimator (PME) $\bar{\mathbf{x}} = \int \mathbf{x} P(\mathbf{x}|\mathbf{y}, H) d\mathbf{x}$ is the optimal inference scheme to minimize the mean squared error, if the assumed model is matched to the true model.

In this paper, we study joint distributions of L ($\ll K$) elements of input vector and their estimates based on the posterior distribution (2), given a channel matrix H . Without loss of generality we consider the first L elements of input vector, $\mathbf{x}_0^L = (x_{01}, \dots, x_{0L})^T$, and their estimates \mathbf{x}^L . The joint distribution to be studied is thus

$$\mathcal{P}(\mathbf{x}_0^L, \mathbf{x}^L|H) = \int P(\mathbf{x}^L|\mathbf{y}, H) P_0(\mathbf{y}|H\mathbf{x}_0) P_0(\mathbf{x}_0) d\mathbf{x}_0^{\setminus L}. \quad (3)$$

We assume the channel matrix H to be random and evaluate expectation of $\mathcal{P}(\mathbf{x}_0^L, \mathbf{x}^L|H)$ over H in the large-system limit where $K, N \rightarrow \infty$ while $\beta = K/N$ is kept finite:

$$\mathcal{P}(\mathbf{x}_0^L, \mathbf{x}^L) = \lim_{K, N \rightarrow \infty} E_H [\mathcal{P}(\mathbf{x}_0^L, \mathbf{x}^L|H)]. \quad (4)$$

$E_u[\dots]$ denotes expectation over the random variable u . Note that if the scalar-channel characterization is derived for the joint distribution (4) using the replica method, it is easy to show the scalar-channel characterization is still valid for arbitrary microscopic quantities depend on \mathbf{x}_0^L and \mathbf{x}^L .

To simplify the analysis, we assume the following:

- Random channel matrix: The elements $\{h_{\mu k}\}$ are independent and identically distributed (i.i.d.) with mean

zero, unit variance, odd-order moments being zero and $(2m)$ th-order moments being finite.

- The first L elements of input vector \mathbf{x}_0^L and the remaining elements $\mathbf{x}_0^{\setminus L} = (x_{0(L+1)}, \dots, x_{0K})^T$ are independent, so that the prior distribution of \mathbf{x} is factorized as

$$P_0(\mathbf{x}_0) = P_0^L(\mathbf{x}_0^L) P_0^{\setminus L}(\mathbf{x}_0^{\setminus L}). \quad (5)$$

The factorized form $P(\mathbf{x}) = P^L(\mathbf{x}^L) P^{\setminus L}(\mathbf{x}^{\setminus L})$ is also used as the postulated prior distribution.

- The conditional distributions $\rho_0(y|u)$ and $\rho(y|u)$ are one and two times differentiable with respect to u , respectively.

III. MAIN RESULT

Our main result is the following claim.

Claim 1: In the large-system limit and under the assumption of replica symmetry (see Sect. IV), the joint distribution $\mathcal{P}(\mathbf{x}_0^L, \mathbf{x}^L)$ defined in (4) is asymptotically equivalent to the joint distribution

$$\begin{aligned} \mathcal{P}(\mathbf{x}_0^L, \mathbf{x}^L) &= \int \frac{\prod_{k=1}^L \rho_G(z_k|x_k) \tilde{P}^L(\mathbf{x}^L)}{\int \prod_{k=1}^L \rho_G(z_k|x_k) \tilde{P}^L(\mathbf{x}^L) d\mathbf{x}^L} \\ &\times \prod_{k=1}^L \rho_{G0}(z_k|x_{0k}) P_0(\mathbf{x}_0^L) d\mathbf{z}^L, \end{aligned} \quad (6)$$

where $\rho_{G0}(z|x)$ and $\rho_G(z|x)$ represent input-output characteristics of the scalar Gaussian channels

$$\rho_{G0}(z|x) = \sqrt{\frac{E^2}{2\pi F}} \exp\left[-\frac{E^2(z-x)^2}{2F}\right], \quad (7)$$

$$\rho_G(z|x) = \sqrt{\frac{E}{2\pi}} \exp\left[-\frac{E(z-x)^2}{2}\right], \quad (8)$$

respectively, and where $\mathbf{z}^L = (z_1, \dots, z_L)^T$. $\tilde{P}^L(\mathbf{x}^L)$ is a “modulated” version of the assumed prior, defined as

$$\tilde{P}^L(\mathbf{x}^L) = \frac{\exp\left[\frac{G-F+E}{2} \|\mathbf{x}^L\|^2\right] P^L(\mathbf{x}^L)}{\int \exp\left[\frac{G-F+E}{2} \|\mathbf{x}^L\|^2\right] P^L(\mathbf{x}^L) d\mathbf{x}^L}, \quad (9)$$

where $\|\mathbf{x}\|^2 = \mathbf{x}^T \mathbf{x}$.

The parameters $\{G, E, F\}$ are determined by solving the following equations for $\{G, E, F, r, m, q\}$,

$$G = \int \bar{\rho}_0\left(y \sqrt{\frac{\beta m^2}{q}} t\right) \frac{\bar{\rho}''(y|\sqrt{\beta q} t)}{\bar{\rho}(y|\sqrt{\beta q} t)} Dt dy, \quad (10)$$

$$E = \int \bar{\rho}'_0\left(y \sqrt{\frac{\beta m^2}{q}} t\right) \frac{\bar{\rho}'(y|\sqrt{\beta q} t)}{\bar{\rho}(y|\sqrt{\beta q} t)} Dt dy, \quad (11)$$

$$F = \int \bar{\rho}_0\left(y \sqrt{\frac{\beta m^2}{q}} t\right) \left[\frac{\bar{\rho}'(y|\sqrt{\beta q} t)}{\bar{\rho}(y|\sqrt{\beta q} t)}\right]^2 Dt dy, \quad (12)$$

$$r = \lim_{K \rightarrow \infty} \frac{1}{K} \langle \|\langle \mathbf{x} \rangle\|^2 \rangle, \quad (13)$$

$$m = \lim_{K \rightarrow \infty} \frac{1}{K} \langle \mathbf{x}_0^T \langle \mathbf{x} \rangle \rangle, \quad (14)$$

$$q = \lim_{K \rightarrow \infty} \frac{1}{K} \langle \|\langle \mathbf{x} \rangle\|^2 \rangle, \quad (15)$$

where $\int(\dots) Du = \int_{-\infty}^{\infty}(\dots) \exp(-u^2/2) du / \sqrt{2\pi}$. The distributions $\bar{\rho}_0$ and $\bar{\rho}$ are defined as

$$\begin{aligned} \bar{\rho}_0\left(y \sqrt{\frac{\beta m^2}{q}} t\right) &= \int \rho_0\left(y \sqrt{\frac{\beta m^2}{q}} t + \sqrt{\beta\left(r_0 - \frac{m^2}{q}\right)} u\right) Du, \\ \bar{\rho}\left(y \sqrt{\beta q} t\right) &= \int \rho\left(y \sqrt{\beta q} t + \sqrt{\beta(r-q)} u\right) Du, \end{aligned} \quad (16)$$

respectively, where $f'(y|u) = \frac{\partial}{\partial u} f(y|u)$, and where

$$r_0 = \lim_{K \rightarrow \infty} \frac{1}{K} \int \|\mathbf{x}_0\|^2 P_0(\mathbf{x}_0) d\mathbf{x}_0. \quad (18)$$

The brackets $\langle \dots \rangle$ and $\langle \dots \rangle$ denote the averages with respect to the joint distribution of \mathbf{x}_0 and $\mathbf{z} = (z_1, \dots, z_K)^T$,

$$\langle \dots \rangle = \iint (\dots) \prod_{k=1}^K \rho_{G0}(z_k|x_{0k}) P_0(\mathbf{x}_0) d\mathbf{z} d\mathbf{x}_0, \quad (19)$$

and the posterior distribution of \mathbf{x} given \mathbf{z} ,

$$\langle \dots \rangle = \frac{\int (\dots) \prod_{k=1}^K \rho_G(z_k|x_k) \tilde{P}(\mathbf{x})}{\int \prod_{k=1}^K \rho_G(z_k|x_k) \tilde{P}(\mathbf{x}) d\mathbf{x}}, \quad (20)$$

respectively, where

$$\tilde{P}(\mathbf{x}) = \frac{\exp\left[\frac{G-F+E}{2} \|\mathbf{x}\|^2\right] P(\mathbf{x})}{\int \exp\left[\frac{G-F+E}{2} \|\mathbf{x}\|^2\right] P(\mathbf{x}) d\mathbf{x}}. \quad (21)$$

If more than one solution exists for (10)–(15), the correct solution is the one that minimizes the function \mathcal{F} defined as

$$\begin{aligned} \mathcal{F} &= \frac{1}{\beta} \iint \bar{\rho}_0\left(y \sqrt{\frac{\beta m^2}{q}} t\right) \log \bar{\rho}\left(y \sqrt{\beta q} t\right) Dt dy \\ &+ \frac{1}{2} Gr - Em + \frac{1}{2} Fq + \frac{F}{2E} + \frac{1}{2} Er_0 - \frac{1}{2} \log \frac{E}{2\pi} \\ &+ \lim_{K \rightarrow \infty} \frac{1}{K} \iint \prod_{k=1}^K \rho_{G0}(z_k|x_{0k}) P_0(\mathbf{x}_0) \\ &\times \left\{ \log \int \prod_{k=1}^K \rho_G(z_k|x_k) \tilde{P}(\mathbf{x}) d\mathbf{x} \right\} d\mathbf{x}_0 d\mathbf{z}. \end{aligned} \quad (22)$$

Detailed derivation of the claim is given in Section IV. The claim implies that the scalar-channel characterization is valid for the joint distribution $\mathcal{P}(\mathbf{x}_0^L, \mathbf{x}^L)$, this is, the joint distribution $\mathcal{P}(\mathbf{x}_0^L, \mathbf{x}^L)$ defined in (4) can be asymptotically identified as the joint distribution of \mathbf{x}_0^L and \mathbf{x}^L where the elements of \mathbf{x}_0^L are independently transmitted over the scalar Gaussian channel $\rho_{G0}(z|x)$ and where the detector postulates the channel model $\rho_G(z|x)$ and the modulated version of the assumed prior $\tilde{P}(\mathbf{x}^L)$ (Fig. 1). This result is a finer version of the decoupling principle, which is first stated by Tse and Hanly [5], and named by Guo and Verdú [2].

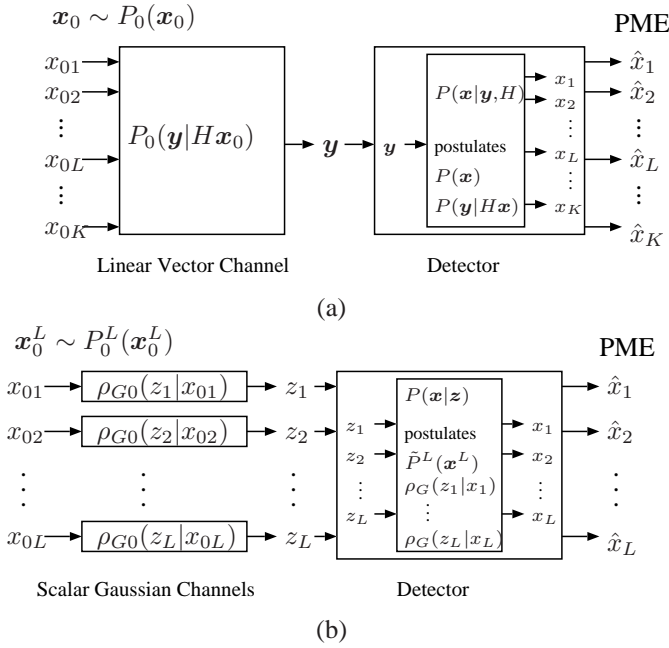


Fig. 1. The linear vector channel and the corresponding detector (a). The bank of scalar Gaussian channels and their corresponding detector (b).

IV. DERIVATION OF THE CLAIM

A. Replica method

We evaluate $\mathcal{P}(\mathbf{x}_0^L, \mathbf{x}^L)$ defined in (4) via replica method. Introducing a real number n , (4) can be rewritten as

$$\begin{aligned} \mathcal{P}(\mathbf{x}_0^L, \mathbf{x}^L) &= \lim_{K, N \rightarrow \infty} \lim_{n \rightarrow 0} \mathbb{E}_H \left[\int \left\{ \int P(\mathbf{y}|\mathbf{H}\mathbf{x}) P(\mathbf{x}) d\mathbf{x}^{\setminus L} \right\} \right. \\ &\quad \times \left. \left\{ \int P(\mathbf{y}|\mathbf{H}\mathbf{x}) P(\mathbf{x}) d\mathbf{x} \right\}^{n-1} P_0(\mathbf{y}|\mathbf{H}\mathbf{x}_0) P_0(\mathbf{x}_0) d\mathbf{x}_0^{\setminus L} \right]. \end{aligned} \quad (23)$$

According to the standard prescription of replica method, we first evaluate

$$\begin{aligned} \mathcal{Z}_n(\mathbf{x}_0^L, \mathbf{x}_1^L) &= \lim_{K, N \rightarrow \infty} \mathbb{E}_H \left[\prod_{a=0}^n \left\{ \int P_a(\mathbf{y}|\mathbf{H}\mathbf{x}_a) P_a(\mathbf{x}_a) \right\} \right. \\ &\quad \times \left. d\mathbf{x}_0^{\setminus L} d\mathbf{x}_1^{\setminus L} \prod_{a=2}^n d\mathbf{x}_a \right] \end{aligned} \quad (24)$$

for a positive integer n , where $P_a(\mathbf{y}|\mathbf{H}\mathbf{x}_a) = P(\mathbf{y}|\mathbf{H}\mathbf{x}_a)$ and $P_a(\mathbf{x}) = P(\mathbf{x})$ for $a = 1, \dots, n$, and then the result is continued to real n in order to take the limit $n \rightarrow 0$ to obtain

$$\lim_{n \rightarrow 0} \mathcal{Z}_n(\mathbf{x}_0^L, \mathbf{x}_1^L) \Big|_{\mathbf{x}_1^L = \mathbf{x}^L} = \mathcal{P}(\mathbf{x}_0^L, \mathbf{x}^L). \quad (25)$$

Although there is no rigorous justification for the replica method, we assume validity of the replica method and related techniques throughout this paper.

B. Average over channel matrix

To evaluate (24), we first take the average over the channel matrix \mathbf{H} . Using the assumptions of random channel matrix and memoryless channels, one has

$$\begin{aligned} \mathcal{Z}_n(\mathbf{x}_0^L, \mathbf{x}_1^L) &= \lim_{K, N \rightarrow \infty} \int \cdots \int \left\{ \mathbb{E}_{\mathbf{h}} \left[\int \prod_{a=0}^n \rho_a \left(y \left| \frac{\mathbf{h}^T \mathbf{x}_a}{\sqrt{N}} \right) dy \right] \right\}^N \\ &\quad \times \prod_{a=0}^n P_a(\mathbf{x}_a) d\mathbf{x}_0^{\setminus L} d\mathbf{x}_1^{\setminus L} \prod_{a=2}^n d\mathbf{x}_a, \end{aligned} \quad (26)$$

where $\rho_a(y|u) = \rho(y|u)$ for $a = 1, \dots, n$. We let

$$A = \left\{ \mathbb{E}_{\mathbf{h}} \left[\int \prod_{a=0}^n \rho_a \left(y \left| \frac{\mathbf{h}^T \mathbf{x}_a}{\sqrt{N}} \right) dy \right] \right\}^N \quad (27)$$

and introduce auxiliary random variables $\mathbf{v} = (v_0, \dots, v_n)^T$, $v_a = \mathbf{h}^T \mathbf{x}_a / \sqrt{K}$. The average over \mathbf{h} in (27) can be rewritten in terms of an integral over the conditional distribution of \mathbf{v} given $\{\mathbf{x}_a; a = 0 \dots, n\}$, denoted by $V(\mathbf{v}|\{\mathbf{x}_a\})$, as

$$A = \left\{ \int V(\mathbf{v}|\{\mathbf{x}_a\}) \int \prod_{a=0}^n \rho_a \left(y \left| \sqrt{\beta} v_a \right) dy d\mathbf{v} \right\}^N. \quad (28)$$

To obtain an explicit expression for $V(\mathbf{v}|\{\mathbf{x}_a\})$, we evaluate the characteristic function of $V(\mathbf{v}|\{\mathbf{x}_a\})$, as

$$\begin{aligned} \hat{V}(\hat{\mathbf{v}}|\{\mathbf{x}_a\}) &= \int e^{i\hat{\mathbf{v}}^T \mathbf{v}} V(\mathbf{v}|\{\mathbf{x}_a\}) d\mathbf{v} \\ &= \exp \left[-\frac{1}{2} \hat{\mathbf{v}}^T Q \hat{\mathbf{v}} \right] \\ &\quad \times \left\{ 1 - \frac{3-\kappa}{24K} \sum_{a,b,c,d=0}^n W_{abcd} \hat{v}_a \hat{v}_b \hat{v}_c \hat{v}_d + O(K^{-2}) \right\}, \end{aligned} \quad (29)$$

where $\hat{\mathbf{v}} = (\hat{v}_0, \dots, \hat{v}_n)^T$, where κ is fourth-order moment of $h_{\mu k}$, and where $(n+1) \times (n+1)$ symmetric matrix Q and fourth-order symmetric tensor W are defined as

$$Q_{ab} = \frac{1}{K} \sum_{k=1}^K x_{ak} x_{bk} \quad (0 \leq a \leq b \leq n), \quad (30)$$

$$W_{abcd} = \frac{1}{K} \sum_{k=1}^K x_{ak} x_{bk} x_{ck} x_{dk} \quad (0 \leq a \leq b \leq c \leq d \leq n). \quad (31)$$

Note that in the above we have to evaluate $\hat{V}(\hat{\mathbf{v}}|\{\mathbf{x}_a\})$ up to $O(K^{-1})$ terms. The inverse Fourier transform yields

$$V(\mathbf{v}|\{\mathbf{x}_a\}) = V_G(\mathbf{v}) - \frac{1}{K} V_{\Delta}(\mathbf{v}) + O(K^{-2}), \quad (32)$$

where

$$V_G(\mathbf{v}) = \left[(2\pi)^{n+1} \det(Q) \right]^{-\frac{1}{2}} \exp \left[-\frac{1}{2} \mathbf{v}^T Q^{-1} \mathbf{v} \right], \quad (33)$$

$$V_{\Delta}(\mathbf{v}) = \frac{3-\kappa}{24} \sum_{a,b,c,d=0}^n W_{abcd} \frac{\partial^4}{\partial v_a \partial v_b \partial v_c \partial v_d} V_G(\mathbf{v}). \quad (34)$$

Collecting these expressions, we have

$$A = \exp \left[N\mathcal{G}_0(Q) - \mathcal{G}_1(Q, W) + O(K^{-1}) \right], \quad (35)$$

where

$$\mathcal{G}_0(Q) = \log \int V_G(\mathbf{v}) \int \prod_{a=0}^n \rho_a(y|\sqrt{\beta}v_a) dy d\mathbf{v}, \quad (36)$$

$$\mathcal{G}_1(Q, W) = \frac{\int V_\Delta(\mathbf{v}) \int \prod_{a=0}^n \rho_a(y|\sqrt{\beta}v_a) dy d\mathbf{v}}{\beta \int V_G(\mathbf{v}) \int \prod_{a=0}^n \rho_a(y|\sqrt{\beta}v_a) dy d\mathbf{v}}. \quad (37)$$

C. Integral over Q and W

Since the quantity A depends on $\{\mathbf{x}_a\}$ only through Q and W , one can rewrite (26) in terms of an integral over Q and W , as

$$\begin{aligned} \mathcal{Z}_n(\mathbf{x}_0^L, \mathbf{x}_1^L) &= \lim_{K, N \rightarrow \infty} \iint \exp \left[N\mathcal{G}_0(Q) - \mathcal{G}_1(Q, W) + O(K^{-1}) \right] \\ &\times \mu_K(Q, W; \mathbf{x}_0^L, \mathbf{x}_1^L) dQ dW, \end{aligned} \quad (38)$$

where

$$\begin{aligned} \mu_K(Q, W; \mathbf{x}_0^L, \mathbf{x}_1^L) &= \int \cdots \int \prod_{0 \leq a \leq b \leq n} \delta \left(Q_{ab} - \frac{1}{K} \sum_{k=1}^K x_{ak} x_{bk} \right) \\ &\times \prod_{0 \leq a \leq b \leq c \leq d \leq n} \delta \left(W_{abcd} - \frac{1}{K} \sum_{k=1}^K x_{ak} x_{bk} x_{ck} x_{dk} \right) \\ &\times \prod_{a=0}^n P_a(\mathbf{x}_a) d\mathbf{x}_0^L d\mathbf{x}_1^L \prod_{a=2}^n d\mathbf{x}_a, \end{aligned} \quad (39)$$

and $dQ = \prod_{0 \leq a \leq b \leq n} dQ_{ab}$, $dW = \prod_{0 \leq a \leq b \leq c \leq d \leq n} dW_{abcd}$.

We evaluate (39) in the large-system limit by following the derivation in [6], [7]. We introduce parameters $\hat{Q} = \{\hat{Q}_{ab}; 0 \leq a \leq b \leq n\}$ and $\hat{W} = \{\hat{W}_{abcd}; 0 \leq a \leq b \leq c \leq d \leq n\}$, which are conjugates to Q and W , respectively, and define some functions of them for later use:

$$\begin{aligned} \Lambda(\hat{Q}, \hat{W}) &= \frac{1}{K} \log \int \cdots \int \prod_{k=1}^K \exp \left[\sum_{0 \leq a \leq b \leq n} \hat{Q}_{ab} x_{ak} x_{bk} \right. \\ &\left. + \sum_{0 \leq a \leq b \leq c \leq d \leq n} \hat{W}_{abcd} x_{ak} x_{bk} x_{ck} x_{dk} \right] \prod_{a=0}^n \left\{ P_a(\mathbf{x}_a) d\mathbf{x}_a \right\}, \end{aligned} \quad (40)$$

$$\begin{aligned} \lambda_x(\hat{Q}, \hat{W}; \mathbf{x}_0^L, \mathbf{x}_1^L) &= \log \int \cdots \int \prod_{k=1}^L \exp \left[\sum_{0 \leq a \leq b \leq n} \hat{Q}_{ab} x_{ak} x_{bk} \right. \\ &\left. + \sum_{0 \leq a \leq b \leq c \leq d \leq n} \hat{W}_{abcd} x_{ak} x_{bk} x_{ck} x_{dk} \right] \prod_{a=0}^n P_a^L(\mathbf{x}_a^L) \prod_{a=2}^n d\mathbf{x}_a^L, \end{aligned} \quad (41)$$

$$\begin{aligned} \lambda(\hat{Q}, \hat{W}) &= \log \int \cdots \int \prod_{k=1}^L \exp \left[\sum_{0 \leq a \leq b \leq n} \hat{Q}_{ab} x_{ak} x_{bk} \right. \\ &\left. + \sum_{0 \leq a \leq b \leq c \leq d \leq n} \hat{W}_{abcd} x_{ak} x_{bk} x_{ck} x_{dk} \right] \prod_{a=0}^n \left\{ P_a^L(\mathbf{x}_a^L) d\mathbf{x}_a^L \right\}. \end{aligned} \quad (42)$$

We further assume that $\Lambda(\hat{Q}, \hat{W})$ has a limit as $K \rightarrow \infty$. Using the functions (40)–(42), the Fourier transform of (39) is given by

$$\begin{aligned} \hat{\mu}_K(\hat{Q}, \hat{W}; \mathbf{x}_0^L, \mathbf{x}_1^L) &= \exp \left[K\Lambda \left(i\frac{\hat{Q}}{K}, i\frac{\hat{W}}{K} \right) \right. \\ &\left. + \lambda_x \left(i\frac{\hat{Q}}{K}, i\frac{\hat{W}}{K}; \mathbf{x}_0^L, \mathbf{x}_1^L \right) - \lambda \left(i\frac{\hat{Q}}{K}, i\frac{\hat{W}}{K} \right) \right], \end{aligned} \quad (43)$$

and its inverse Fourier transform yields

$$\begin{aligned} \mu_K(Q, W; \mathbf{x}_0^L, \mathbf{x}_1^L) &= \left(\frac{K}{2\pi} \right)^{\{(n+2) + (n+4)\}} \\ &\times \iint \exp \left[K \left\{ -iQ \cdot \hat{Q} - iW \cdot \hat{W} + \Lambda(i\hat{Q}, i\hat{W}) \right\} \right] \\ &\times \exp \left[\lambda_x(i\hat{Q}, i\hat{W}; \mathbf{x}_0^L, \mathbf{x}_1^L) - \lambda(i\hat{Q}, i\hat{W}) \right] d\hat{Q} d\hat{W}, \end{aligned} \quad (44)$$

where $Q \cdot \hat{Q}$ and $W \cdot \hat{W}$ are abbreviations of $\sum_{0 \leq a \leq b \leq n} Q_{ab} \hat{Q}_{ab}$ and $\sum_{0 \leq a \leq b \leq c \leq d \leq n} W_{abcd} \hat{W}_{abcd}$, respectively.

To evaluate the integral over \hat{Q} and \hat{W} in (44), let $\hat{Q}^* = \{\hat{Q}_{ab}^*; 0 \leq a \leq b \leq n\}$ and $\hat{W}^* = \{\hat{W}_{abcd}^*; 0 \leq a \leq b \leq c \leq d \leq n\}$ denote the solution of the equations

$$Q_{ab} = \frac{\partial \Lambda(\hat{Q}, \hat{W})}{\partial \hat{Q}_{ab}}, \quad W_{abcd} = \frac{\partial \Lambda(\hat{Q}, \hat{W})}{\partial \hat{W}_{abcd}}. \quad (45)$$

Applying three operations to (44); a change of variables

$$i\hat{Q}_{ab} \rightarrow i\frac{\hat{Q}_{ab}}{\sqrt{K}} + \hat{Q}_{ab}^*, \quad i\hat{W}_{abcd} \rightarrow i\frac{\hat{W}_{abcd}}{\sqrt{K}} + \hat{W}_{abcd}^*, \quad (46)$$

Taylor expansion of Λ , λ_x and λ , and a change of integration paths to real axes, one can find that the integral in (44) leads to a Gaussian integration. Then, one obtains

$$\begin{aligned} \mu_K(Q, W; \mathbf{x}_0^L, \mathbf{x}_1^L) &= \left(\frac{K}{2\pi} \right)^{\frac{1}{2}\{(n+2) + (n+4)\}} \det \left(\mathcal{H}(\Lambda|\hat{Q}^*, \hat{W}^*) \right)^{-\frac{1}{2}} \\ &\times \exp \left[K \left\{ -Q \cdot \hat{Q}^* - W \cdot \hat{W}^* + \Lambda(\hat{Q}^*, \hat{W}^*) \right\} \right. \\ &\left. + \lambda_x(\hat{Q}^*, \hat{W}^*; \mathbf{x}_0^L, \mathbf{x}_1^L) - \lambda(\hat{Q}^*, \hat{W}^*) + O(K^{-\frac{1}{2}}) \right], \end{aligned} \quad (47)$$

where $\mathcal{H}(f|\mathbf{u}^*)$ represents a Hessian matrix of the function $f(\mathbf{u})$ at $\mathbf{u} = \mathbf{u}^*$. Use of Gaussian integration requires the Hessian matrix $\mathcal{H}(\Lambda|\hat{Q}^*, \hat{W}^*)$ being positive definite. Note that a similar evaluation is still possible when $\mathcal{H}(\Lambda|\hat{Q}^*, \hat{W}^*)$ is non-negative definite [8].

D. Saddle-point evaluation

We evaluate the integral over Q and W in (38) via the saddle-point method [9]. We obtain

$$\begin{aligned} \mathcal{Z}_n(\mathbf{x}_0^L, \mathbf{x}_1^L) &= \lim_{K, N \rightarrow \infty} D \exp \left[Kn\mathcal{F}_n(Q^*, W^*) - \mathcal{G}_1(Q^*, W^*) \right. \\ &\quad \left. + \lambda_x(\hat{Q}^*, \hat{W}^*; \mathbf{x}_0^L, \mathbf{x}_1^L) - \lambda(\hat{Q}^*, \hat{W}^*) + O(K^{-1}) \right], \end{aligned} \quad (48)$$

where the function $\mathcal{F}_n(Q, W)$ is defined as

$$\begin{aligned} \mathcal{F}_n(Q, W) &= \frac{1}{n} \left[\frac{1}{\beta} \mathcal{G}_0(Q) - Q \cdot \hat{Q}^* - W \cdot \hat{W}^* + \Lambda(\hat{Q}^*, \hat{W}^*) \right] \end{aligned} \quad (49)$$

Note that \hat{Q}^* and \hat{W}^* depend on Q and W via (45). The saddle points $Q^* = \{Q_{ab}^*; 0 \leq a \leq b \leq n\}$ and $W^* = \{W_{abcd}^*; 0 \leq a \leq b \leq c \leq d \leq n\}$ are determined as the solution of

$$\frac{\partial \mathcal{F}_n(Q, W)}{\partial Q_{ab}} = 0, \quad \frac{\partial \mathcal{F}_n(Q, W)}{\partial W_{abcd}} = 0. \quad (50)$$

If more than one solution exists for (50), the correct solution is the one that maximizes (49). The normalization factor D is given by

$$D = \left[\det(\mathcal{H}(\Lambda|\hat{Q}^*, \hat{W}^*)) \det(\mathcal{H}(-n\mathcal{F}_n|Q^*, W^*)) \right]^{-\frac{1}{2}}. \quad (51)$$

Application of the saddle-point method here requires that the Hessian matrix $\mathcal{H}(-n\mathcal{F}_n|Q^*, W^*)$ is positive definite.

Since our final result will be a function of \mathbf{x}_0^L and \mathbf{x}_1^L , we can ignore terms in (48) which are independent of these variables, obtaining

$$\mathcal{Z}_n(\mathbf{x}_0^L, \mathbf{x}_1^L) \propto \exp \left[\lambda_x(\hat{Q}^*, 0; \mathbf{x}_0^L, \mathbf{x}_1^L) \right]. \quad (52)$$

Note that one obtains $\hat{W}_{abcd} = 0$ by solving (50), and that the overall factor, which we have just ignored, can be determined via normalization. It turns out, from $\hat{W}_{abcd} = 0$, (45), and (50), that Q^* and \hat{Q}^* do not depend on W^* .

E. Replica symmetric ansatz

To proceed further, we assume replica symmetry (RS) [10], under which we let

$$\begin{aligned} Q_{00}^* &= r_0, & Q_{aa}^* &= r, & Q_{0a}^* &= m, & Q_{ab}^* &= q, \\ \hat{Q}_{00}^* &= \frac{1}{2}G_0, & \hat{Q}_{aa}^* &= \frac{1}{2}G, & \hat{Q}_{0a}^* &= E, & \hat{Q}_{ab}^* &= F, \end{aligned} \quad (53)$$

for positive integers $a < b$. Then, $\mathcal{F} \equiv \lim_{n \rightarrow 0} \mathcal{F}_n(Q, W)$ is reduced to (22), and the saddle-point equations (45) and (50) become (10)–(15), (18) and $G_0 = 0$ (For detailed derivation, see [10]). Notice that the condition for the Hessian matrix $\mathcal{H}(-n\mathcal{F}_n|Q^*, W^*)$ being positive definite yields the de Almeida-Thouless (AT) condition for local stability of RS solutions [11].

Inserting the RS assumption (53) into (52), one obtains

$$\begin{aligned} \mathcal{Z}_n(\mathbf{x}_0^L, \mathbf{x}_1^L) &\propto \int \left[\prod_{k=1}^L \rho_G(z_k | x_{1k}) e^{\frac{G-F+E}{2} \|\mathbf{x}_1^L\|^2} P_1^L(\mathbf{x}_1^L) \right] \\ &\quad \times \left[\int \prod_{k=1}^L \rho_G(z_k | x_k) e^{\frac{G-F+E}{2} \|\mathbf{x}^L\|^2} P^L(\mathbf{x}^L) d\mathbf{x}^L \right]^{n-1} \\ &\quad \times \prod_{k=1}^L \left\{ \rho_{G_0}(z_k | x_{0k}) e^{\frac{1}{2}(nEz_k^2 + G_0x_{0k}^2)} \right\} P_0^L(\mathbf{x}_0^L) d\mathbf{z}^L. \end{aligned} \quad (54)$$

Taking the limit $n \rightarrow 0$, one finally arrives at (6).

V. CONCLUSION

In this paper, we have considered the decoupling principle of the linear vector channel. We have shown that the scalar-channel characterization obtained via decoupling principle is valid for the joint distributions of $L (\ll K)$ elements of input vector and their estimates based on the posterior probability, in the large-system limit. This implies that the scalar-channel characterization is valid not only for macroscopic quantities, but also for microscopic quantities on the linear vector channel.

ACKNOWLEDGMENT

The authors would like to acknowledge support from the Grant-in-Aid for Scientific Research on Priority Areas (No. 18079010), the Ministry of Education, Culture, Sports, Science and Technology, Japan.

REFERENCES

- [1] 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.
- [2] 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.
- [3] R. R. Müller, "Channel capacity and minimum probability of error in large dual antenna array systems with binary modulation," *IEEE Trans. Signal Process.*, vol. 51, no. 11, pp. 2821–2828, Nov. 2003.
- [4] K. Takeuchi and T. Tanaka, "Hierarchical decoupling principle of a MIMO-CDMA channel in asymptotic limits," in *2007 IEEE Int. Symp. Inform. Theory, Nice, France, 2007*, pp. 1271–1275.
- [5] D. Tse and S. V. Hanly, "Linear multiuser receivers: Effective interference, effective bandwidth and user capacity," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 641–657, Mar. 1999.
- [6] A. Martin-Löf, "A Laplace approximation for sums of independent random variables," *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, vol. 59, pp. 101–115, 1982.
- [7] E. Bolthausen, "Laplace approximations for sums of independent random vectors," *Prob. Th. Rel. Fields*, vol. 72, pp. 305–318, 1986.
- [8] —, "Laplace approximations for sums of independent random vectors. Part II. Degenerate maxima and manifolds of maxima," *Prob. Th. Rel. Fields*, vol. 76, pp. 167–206, 1987.
- [9] E. T. Copson, *Asymptotic Expansions*. Cambridge, U.K.: Cambridge University Press, 1965.
- [10] T. Tanaka, "Statistical learning in digital wireless communications," in *Lecture Notes in Computer Science 3244: Algorithmic Learning Theory*, S. Ben-David, J. Case, and A. Maruoka, Eds. Springer-Verlag, Berlin, Heidelberg, 2004, pp. 464–478.
- [11] J. R. L. de Almeida and D. J. Thouless, "Stability of the sherrington-kirkpatrick solution of a spin glass model," *J. Phys. A: Math. Gen.*, vol. 11, pp. 983–990, 1978.