Statistical mechanical analysis of the Kronecker channel model for multiple-input multiple-output wireless communication

Atsushi Hatabu

System IP Core Research Laboratories, NEC Corporation, Kawasaki 211-8666, Japan and Department of Computer Intelligence and Systems Science, Tokyo Institute of Technology, Yokohama 226-8502, Japan

Koujin Takeda and Yoshiyuki Kabashima

Department of Computer Intelligence and Systems Science, Tokyo Institute of Technology, Yokohama 226-8502, Japan

(Received 18 September 2009; published 17 December 2009)

The Kronecker channel model of wireless communication is analyzed using statistical mechanics methods. In the model, spatial proximities among transmission/reception antennas are taken into account as certain correlation matrices, which generally yield nontrivial dependence among symbols to be estimated. This prevents accurate assessment of the communication performance by naively using a previously developed analytical scheme based on a matrix integration formula. In order to resolve this difficulty, we develop a formalism that can formally handle the correlations in Kronecker models based on the known scheme. Unfortunately, direct application of the developed scheme is, in general, practically difficult. However, the formalism is still useful, indicating that the effect of the correlations generally increase after the fourth order with respect to correlation strength. Therefore, the known analytical scheme offers a good approximation in performance evaluation when the correlation strength is sufficiently small. For a class of specific correlation, we show that the performance analysis can be mapped to the problem of one-dimensional spin systems in random fields, which can be investigated without approximation by the belief propagation algorithm.

DOI: 10.1103/PhysRevE.80.061124 PACS number(s): 84.40.Ua, 75.10.Nr, 89.70.−a

I. INTRODUCTION

Recently, in the field of information science, techniques for efficiently handling systems with large amounts of data are strongly required, and statistical mechanics have attracted a great deal of attention. The number of applications in information science to which analytical schemes in statistical mechanics can be applied is increasing, and such applications offer a variety of consequences [1], some of which are not possible by standard techniques in information science. Information processing is a notable example.

In the present paper, we investigate wireless communication systems. Multiple-input multiple-output (MIMO) systems and code division multiple access (CDMA) systems in wireless communication have mathematical structures that are similar to those of disordered spin systems in physics, and analytical tools in statistical mechanics, such as the replica method and mean-field approximations, have enabled performance analysis and improved processing algorithms for actual communication systems [2–11]. In these studies, the communication process is described by a linear equation with transmitted signals $\boldsymbol{b} \in \mathbb{C}^K$ and received signals $\boldsymbol{r} \in \mathbb{C}^L$ using an $L \times K$ channel matrix $\boldsymbol{H} = (H_{lk}) \in \mathbb{C}^{LK}$ and noise $\boldsymbol{\eta} \in \mathbb{C}^L$ as

$$\boldsymbol{r} = \boldsymbol{H}\boldsymbol{b} + \sigma\boldsymbol{\eta}, \tag{1}$$

where σ^2 describes the noise power [see Fig. 1(A)]. Throughout the present paper, matrices and vectors are denoted in bold. In the above equation, the dimension K represents the number of multiple transmission antennas in the MIMO system, whereas L corresponds to the number of reception antennas. Clarifying the feature of the above-mentioned communication channel by the standard method of information theory is technically difficult because of the randomness in $\boldsymbol{H}$ and the discreteness of variable $\boldsymbol{b}$. However, statistical mechanical analysis enables us to avoid such difficulties in the limit of infinite system size.

In previous studies based on statistical mechanical methods, channel matrix $\boldsymbol{H}$ was characterized by a property that the cross correlation $\boldsymbol{H}^\dagger\boldsymbol{H}$ can be handled as a typical sample from a rotationally invariant matrix ensemble as follows:

$$\boldsymbol{H}^\dagger\boldsymbol{H} = \boldsymbol{U}\boldsymbol{D}\boldsymbol{U}^\dagger, \tag{2}$$

where $\boldsymbol{U}$ is a sample randomly chosen from the uniform distribution of K-dimensional unitary matrices, and $\boldsymbol{D}$ is a K-dimensional diagonal matrix. If $\boldsymbol{D}$ has an asymptotic and

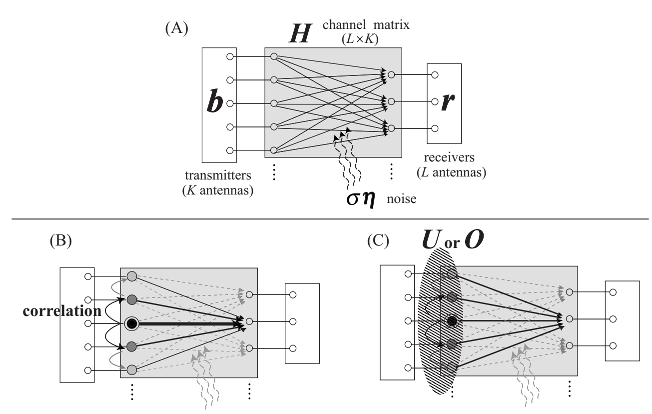

FIG. 1. (a) Schematic picture of the MIMO system. (b) The MIMO system with correlation. Correlation is propagated through adjacent transmission antennas. (c) The correlated system with unitary or orthogonal matrix multiplication. Correlations are smeared by the multiplication of unitary or orthogonal matrix.

deterministic eigenvalue distribution $\rho_D(\lambda)$ with a large matrix size limit $L,K\to\infty$ while keeping $\beta\equiv K/L$ finite, the features of this channel can be characterized by $\rho(\lambda)$, in conjunction with the replica method, and the performance of the channel can be assessed [12–14] by a matrix integration formula [15–19], which is defined for $\rho(\lambda)$.

However, one problem remains. For the simplest case in which each element of random matrix $\boldsymbol{H}$ is drawn from an independent and identically distributed (iid) Gaussian distribution, the property of rotational invariance concerning the cross correlation is satisfied. However, such property does not necessarily hold for general matrix ensembles of MIMO systems. For instance, in the Kronecker model [20], which is one of the standard models in the theory of wireless communication, the elements of the channel matrix are not drawn from an iid Gaussian distribution but are instead drawn from an $L\times K$-dimensional joint Gaussian distribution. More precisely, the channel matrix $\boldsymbol{H}$ is described as

$$\boldsymbol{H} = \sqrt{\boldsymbol{R}_{\mathrm{r}}}\,\Xi\,\sqrt{\boldsymbol{R}_{\mathrm{t}}}, \tag{3}$$

where each component of an $L\times K$ rectangular matrix $\Xi=(\Xi_{lk})$ is drawn from a complex iid Gaussian distribution: $P(\Xi_{lk})=L\pi^{-1}e^{-L|\Xi_{lk}|^2}(1\le l\le L, 1\le k\le K)$. $\boldsymbol{R}_{\mathrm{r}}\in\mathbb{C}^{L^2}$ and $\boldsymbol{R}_{\mathrm{t}}\in\mathbb{C}^{K^2}$ are L- and K-dimensional deterministic matrices, which are Hermitian and indicate correlations among reception antennas and transmission antennas, respectively. (Square root of a square matrix $\boldsymbol{A}$ is defined by the Cholesky decomposition $\boldsymbol{A}=\sqrt{\boldsymbol{A}}^{\dagger}\sqrt{\boldsymbol{A}}$ here.) In a previous paper [13], we analyzed this system by means of the matrix integration formula. However, for this system, the matrix ensemble is not rotationally invariant and, accordingly, the result of performance analysis via the matrix integration formula may not hold exactly.

One of the goals of the present paper is to develop a scheme that can handle the dependence on $\sqrt{\boldsymbol{R}_{\mathrm{r}}}$ and $\sqrt{\boldsymbol{R}_{\mathrm{t}}}$ in Eq. (3) explicitly. In other words, the method developed herein relies on the direct integration of each matrix element in Gaussian random matrix Ξ. The results of analysis for mutual information indicate that the roles of the deterministic matrices $\sqrt{\boldsymbol{R}_{\mathrm{r}}}$ and $\sqrt{\boldsymbol{R}_{\mathrm{t}}}$ are different. As will be shown later herein, the dependence on $\sqrt{\boldsymbol{R}_{\mathrm{r}}}$ can be treated using the matrix integration technique, whereas the dependence on $\sqrt{\boldsymbol{R}_{\mathrm{t}}}$ must be handled more carefully. The developed scheme can also be used to construct a practical demodulation algorithm. Another goal is to compare the performance of the Kronecker channel [Eq. (3)] via a analysis with the performance of the matrix integration formula applied to the entire cross-correlation matrix $\boldsymbol{H}^{\dagger}\boldsymbol{H}$, as we demonstrated in Ref. [13]. The two formulations are found to yield different result, which means that the application of the matrix integration formula to the cross-correlation matrix $\boldsymbol{H}^{\dagger}\boldsymbol{H}$, in general, does not yield correct results. However, when correlation among transmission antennas or when the off-diagonal element of the deterministic matrix $\sqrt{\boldsymbol{R}_{\mathrm{t}}}$ is sufficiently small, discrepancy between results of the scheme developed herein and that based on the matrix integration formula increases only after the fourth order with respect to correlation strength, implying that the formulation based on entire matrix integration yields good approximate results. This suggests that although the matrix-integration technique is generally an approximation, this technique is practically useful when the correlation is small because the matrix integration method enables the system to be characterized using only a few macroscopic variables, which significantly reduces the computational cost for analysis.

The remainder of the present paper is organized as follows. In Sec. II, we provide basic tools for performance analysis and propose a approach to analyze the Kronecker channel model. The analytical results differ from those obtained by matrix integration. In addition, we compare two results by a method of perturbative expansion with respect to the correlation parameter, and a discrepancy appears in the fourth-order coefficient of the correlation parameter, which indicates that the discrepancy is small when the correlation is small. In Sec. III, we show that the demodulation algorithm can be constructed from the minimization scheme of Gibbs free energy without the combined scheme of matrix integration and the Thouless-Anderson-Palmer approach as discussed in [12,13,21–23]. We present the experimental results of the demodulation algorithm for the Kronecker channel in Sec. IV. As a special case, we consider a system with a tridiagonal form of $\boldsymbol{R}_{\mathrm{t}}$, where the analytical scheme can be used for the random-field Ising chain. The results of a numerical experiment confirm the validity and usefulness of the proposed scheme. The final section presents a summary of the present paper.

II. ANALYSIS

Let us start with the communication channels described by Eq. (1). For the noise, we assume that $\boldsymbol{\eta}$ is drawn from a white normal complex Gaussian distribution $P(\boldsymbol{\eta})=\pi^{-L}e^{-|\boldsymbol{\eta}|^2}$. Each component of the transmit vector $\boldsymbol{b}$ is generated from an iid information source and modulated. For simplicity, the modulated components or symbols b_k are quantized to one of the elements in a set $\mathcal{B}$. For instance, for S-phase shift keying modulation $\mathcal{B}\equiv\{e^{2\pi is/S}\}(s=0,1,\ldots,S-1)$. As special and well-known cases, $\mathcal{B}\equiv\{\pm1\}$ for binary-phase shift keying (BPSK) modulation and $\mathcal{B}\equiv\{\pm1/\sqrt{2}\pm i/\sqrt{2}\}$ for quadrature-phase shift keying (QPSK) modulation. The prior of the transmit vector is denoted by $P(\boldsymbol{b})=\prod_{k=1}^{K}P(b_k)$. Here $P(b_k)=1/|\mathcal{B}|$ and $|\mathcal{B}|$ is the number of elements in $\mathcal{B}$.

As mentioned in the introduction, we investigate the Kronecker model described by the matrix of Eq. (3). In order to apply statistical mechanical schemes to the analysis of communication systems, we allow the number of antennas L and K to be sufficiently large while keeping $\beta=K/L$ finite. Next, let us assume that for the matrices $\boldsymbol{R}_{\mathrm{r}}$ and $\boldsymbol{R}_{\mathrm{t}}$ that there exist deterministic distributions $\rho_{\boldsymbol{R}_{\mathrm{r}}}(\lambda)$ and $\rho_{\boldsymbol{R}_{\mathrm{t}}}(\lambda)$, respectively, in the limit of infinite number of antennas. In addition, we assume that both distributions have compact supports and finite moments, which affects the applicability of the matrix integration formula.

In the following, we consider only the case in which the receivers know the channel matrix $\boldsymbol{H}$ and the noise power σ^2 in advance. The performance of the communication channels can be analyzed by estimating the mutual information between transmitted signals $\boldsymbol{b}$ and the received signals $\boldsymbol{r}$, denoted by $\mathcal{I}_{\boldsymbol{H}}$. For MIMO systems, we have

$$\mathcal{I}_{\boldsymbol{H}} = -\frac{1}{K}\int_{\mathbb{C}^L} d\boldsymbol{r} Z(\boldsymbol{r}) \ln Z(\boldsymbol{r}) - \frac{1}{\beta}\ln(\pi\sigma^2) - \frac{1}{\beta},$$

$$\text{where } Z(\boldsymbol{r}) \equiv \mathop{\mathrm{Tr}}_{\boldsymbol{b}} P(\boldsymbol{b}) \frac{1}{(\pi\sigma^2)^L}\exp\left[-\frac{|\boldsymbol{r}-\boldsymbol{H}\boldsymbol{b}|^2}{\sigma^2}\right]. \quad (4)$$

In this paper we use nat unit for mutual information and entropy, which means that information quantity is measured in a unit of natural logarithm. In statistical mechanics, $Z(\boldsymbol{r})$ serves as a partition function, which depends on quenched randomness $\boldsymbol{H}$, and $\mathcal{I}_{\boldsymbol{H}}$ is considered to represent the free energy.

Following the standard technique, we use the replica method to take the average over the channel matrix $\boldsymbol{H}$ in the mutual information:

$$\overline{\mathcal{I}_{\boldsymbol{H}}} = -\lim_{n\to 0}\frac{\partial}{\partial n}\frac{1}{K}\ln\overline{\int_{\mathbb{C}^L} d\boldsymbol{r} Z^{n+1}(\boldsymbol{r})} - \frac{1}{\beta}\ln(\pi\sigma^2) - \frac{1}{\beta}, \quad (5)$$

where $\overline{\cdots}$ denotes averaging over the distribution of channel matrix $\boldsymbol{H}$. As seen in Appendix A, inserting $\boldsymbol{H}$ in Eq. (3) and after some calculation we have as the final expression

$$\overline{\mathcal{I}_{\boldsymbol{H}}} = \mathop{\mathrm{Extr}}_{\lambda}\left\{\hat{G}_{\boldsymbol{\Xi}^\dagger \boldsymbol{R}_\mathrm{r}\boldsymbol{\Xi}}(\lambda) + I_{\boldsymbol{R}_\mathrm{t}}\left(\frac{\lambda}{\sigma^2}\right)\right\}, \quad (6)$$

where $\hat{G}_{\boldsymbol{\Xi}^\dagger \boldsymbol{R}_\mathrm{r}\boldsymbol{\Xi}}(\lambda)$ is the Legendre transform of $G_{\boldsymbol{\Xi}^\dagger \boldsymbol{R}_\mathrm{r}\boldsymbol{\Xi}}(\lambda)$, $\hat{G}_{\boldsymbol{\Xi}^\dagger \boldsymbol{R}_\mathrm{r}\boldsymbol{\Xi}}(\lambda) \equiv \mathrm{Extr}_\chi\{\lambda\chi - G_{\boldsymbol{\Xi}^\dagger \boldsymbol{R}_\mathrm{r}\boldsymbol{\Xi}}(\chi)\}$, $G_{\boldsymbol{\Xi}^\dagger \boldsymbol{R}_\mathrm{r}\boldsymbol{\Xi}}(\lambda)$ is given by

$$G_{\boldsymbol{\Xi}^\dagger \boldsymbol{R}_\mathrm{r}\boldsymbol{\Xi}}(\boldsymbol{A}) \equiv -\frac{1}{\beta}\int d\lambda \rho_{\boldsymbol{R}_\mathrm{r}}(\lambda)\ln(\boldsymbol{I} - \beta\lambda\boldsymbol{A}), \quad (7)$$

and $I_{\boldsymbol{R}_\mathrm{t}}(\chi)$ is defined as

$$\begin{aligned} I_{\boldsymbol{R}_\mathrm{t}}(\chi) \equiv &-\frac{1}{K}\int_{\mathbb{C}^K} d\boldsymbol{r}'\left\{\mathop{\mathrm{Tr}}_{\boldsymbol{b}} P(\boldsymbol{b})\left(\frac{\chi}{\pi}\right)^K \exp[-\chi|\boldsymbol{r}' - \sqrt{\boldsymbol{R}_\mathrm{t}}\boldsymbol{b}|^2]\right\} \\ &\times \ln\left\{\mathop{\mathrm{Tr}}_{\boldsymbol{b}} P(\boldsymbol{b})\left(\frac{\chi}{\pi}\right)^K \exp[-\chi|\boldsymbol{r}' - \sqrt{\boldsymbol{R}_\mathrm{t}}\boldsymbol{b}|^2]\right\} - \ln\left(\frac{\pi}{\chi}\right) \\ &- 1. \end{aligned} \quad (8)$$

Equation (6) is the primary result of the present paper. As we demonstrate in Sec. IV, it is convenient to use $I_{\boldsymbol{R}_\mathrm{t}}(\chi)$ for the discussion of the performance of the channel.

In the following, we consider three items. First, Eq. (6) provides a physical meaning for the performance analysis of the Kronecker channel. The term $I_{\boldsymbol{R}_\mathrm{t}}$ of the right-hand side corresponds to the mutual information of a channel

$$\boldsymbol{r}' = \sqrt{\boldsymbol{R}_\mathrm{t}}\boldsymbol{b}^0 + \frac{\sigma}{\sqrt{\lambda}}\boldsymbol{\eta}' \quad (9)$$

[see Eq. (4)], where $\boldsymbol{\eta}'$ is a K-dimensional normal complex Gaussian noise. Here, we let λ be a random variable that obeys probability distribution $P(\lambda) \simeq \exp[K\hat{G}_{\boldsymbol{\Xi}^\dagger \boldsymbol{R}_\mathrm{r}\boldsymbol{\Xi}}(\lambda)]$. Then, Eq. (6) means that, in the $K\to\infty$ limit $\overline{\mathcal{I}_{\boldsymbol{H}}}$ corresponds to the average of the exponential of the mutual information, $\exp[KI_{\boldsymbol{R}_\mathrm{t}}(\lambda/\sigma^2)]$ over λ. The extremization of Eq. (6) implies that the balance of the two λ-dependent functions $\hat{G}_{\boldsymbol{\Xi}^\dagger \boldsymbol{R}_\mathrm{r}\boldsymbol{\Xi}}(\lambda)$ and $I_{\boldsymbol{R}_\mathrm{t}}(\lambda/\sigma^2)$, which are dependent on correlations among reception antennas and among transmit antennas, respectively, is significant in the determination of $\overline{\mathcal{I}_{\boldsymbol{H}}}$.

Second, $I_{\boldsymbol{R}_\mathrm{t}}(\chi)$ can be evaluated using the following approximation method. After performing unitary transformation of the matrix $\boldsymbol{R}_\mathrm{t}$ to $\boldsymbol{U}^\dagger\boldsymbol{R}_\mathrm{t}\boldsymbol{U}$, we take the average of $I_{\boldsymbol{U}^\dagger\boldsymbol{R}_\mathrm{t}\boldsymbol{U}}(\chi)$ over unitary matrix $\boldsymbol{U}$ [denoted by $\overline{I_{\boldsymbol{U}^\dagger\boldsymbol{R}_\mathrm{t}\boldsymbol{U}}(\chi)}$ in the following] as in the case for the matrix $\boldsymbol{\Xi}\boldsymbol{R}_\mathrm{r}\boldsymbol{\Xi}^\dagger$. In a manner similar to the evaluation of $\mathcal{I}_{\boldsymbol{H}}$, we have

$$\overline{I_{\boldsymbol{U}^\dagger\boldsymbol{R}_\mathrm{t}\boldsymbol{U}}(\chi)} = \mathop{\mathrm{Extr}}_{\lambda}\{\hat{G}_{\boldsymbol{R}_\mathrm{t}}(\lambda) + I_{\boldsymbol{I}}(\lambda\chi)\}, \quad (10)$$

where $I_{\boldsymbol{I}}(\chi)$ is the mutual information of Eq. (8) after the substitution of $\boldsymbol{R}_\mathrm{t}=\boldsymbol{I}$, which can be decomposed to the mutual information of multiple single-output single-input channels.

Third, if the correlation among transmission antennas is sufficiently small, we can perform a perturbative expansion of $I_{\boldsymbol{R}_\mathrm{t}}$. Let us consider the case in which the matrix $\boldsymbol{R}_\mathrm{t}$ is expressed as $\boldsymbol{R}_\mathrm{t}=\boldsymbol{I}+\rho\boldsymbol{R}$ with a real small parameter ρ and K-dimensional matrix $\boldsymbol{R}$, the diagonal elements of which are all zero. After expansion, we have

$$\begin{aligned} I_{\mathbf{I}+\rho\boldsymbol{R}}(\chi) = I_{\boldsymbol{I}}(\chi) &- \frac{(\rho\chi)^2}{2}\frac{\mathrm{Tr}(\boldsymbol{R}^2)}{K}\{I_{\boldsymbol{I}}'(\chi)\}^2 \\ &+ \frac{(\rho\chi)^3}{3}\frac{\mathrm{Tr}(\boldsymbol{R}^3)}{K}\{I_{\boldsymbol{I}}'(\chi)\}^3 + \cdots, \end{aligned} \quad (11)$$

where $I_{\boldsymbol{I}}'(\chi)$ is the derivative of $I_{\boldsymbol{I}}(\chi)$ with respect to χ. Similarly, expanding the approximate mutual information $\overline{I_{\boldsymbol{I}+\rho\boldsymbol{U}^\dagger\boldsymbol{R}\boldsymbol{U}}(\chi)}$, we obtain the same result up to the third order of ρ. However, a discrepancy appears starting from the fourth-order coefficient. In the case of QPSK modulation, this discrepancy is expressed as (see Appendix B)

$$\begin{aligned} I_{\boldsymbol{I}+\rho\boldsymbol{R}}(\chi) - \overline{I_{\boldsymbol{I}+\rho\boldsymbol{U}^\dagger\boldsymbol{R}\boldsymbol{U}}(\chi)} = &-\frac{(\rho\chi)^4}{2}\frac{1}{K}\sum_i\left[(\boldsymbol{R}^2)_{ii} - \frac{\mathrm{Tr}(\boldsymbol{R}^2)}{K}\right]^2 \\ &\times[-I_{\boldsymbol{I}}''(\chi) - I_{\boldsymbol{I}}'(\chi)^2][I_{\boldsymbol{I}}'(\chi)]^2 \\ &- \frac{(\rho\chi)^4}{4}\frac{1}{K}\left\{\sum_{ij}[\mathrm{Re}(R_{ij})^4 + \mathrm{Im}(R_{ij})^4]\right\} \\ &\times\left\{[-I_{\boldsymbol{I}}''(\chi) - I_{\boldsymbol{I}}'(\chi)^2]^2 + \frac{C(\chi)^2}{6}\right\} \\ &+ O(\rho^5), \end{aligned} \quad (12)$$

where $C(\chi)$ is a function that depends on $P(\boldsymbol{b})$ as well as χ. This indicates that the approximate evaluation of mutual information by matrix integration yields a good result if the perturbation parameter ρ is sufficiently small. Under this condition, the evaluation using matrix integration described in [13] has an advantage in that it provides a good approximate solution that is more convenient than the exact evaluation of mutual information for the channel $\boldsymbol{r}=\boldsymbol{H}\boldsymbol{b}+\boldsymbol{\eta}$.

As described in Appendix B, we can prove that $\{-I_I''(\chi) - I_I'(\chi)^2\} \geq 0$, and accordingly, the right-hand side of Eq. (12) becomes nonpositive for a wide class of $P(\boldsymbol{b})$, including QPSK modulation, which means that approximate evaluation by $\overline{I_{I+\rho U^\dagger R U}(\chi)}$ gives an upper bound of $I_{I+\rho R}(\chi)$ up to the fourth order of the correlation parameter ρ.

III. DEMODULATION ALGORITHM

For practical communication, it is also significant to construct a computationally feasible demodulation algorithm. For inference of original signal $\mathbf{b}$ from received signal $\mathbf{r}$ and channel matrix $\boldsymbol{H}$, it is necessary to evaluate the following quantity:

$$\boldsymbol{m} = \sum_{\boldsymbol{b} \in \mathcal{B}^K} \boldsymbol{b} P(\boldsymbol{b}|\boldsymbol{r},\boldsymbol{H}). \tag{13}$$

However, it is computationally difficult to numerically evaluate $\boldsymbol{m}$ from this expression. Key to the practical solution of this problem is the use of the Gibbs free energy for the communication channel:

$$\Phi(\boldsymbol{m}) = \underset{\boldsymbol{h}}{\mathrm{Extr}}(-\ln \underset{\boldsymbol{b}}{\mathrm{Tr}} P(\boldsymbol{b}|\boldsymbol{r},\boldsymbol{H}) \exp\{\mathrm{Re}[\boldsymbol{h}^\dagger(\boldsymbol{b}-\boldsymbol{m})]\}), \tag{14}$$

and the quantity $\boldsymbol{m}$ can be estimated as the argument of the extremized Gibbs free energy. Substituting $P(\boldsymbol{b}|\boldsymbol{r},\boldsymbol{H}) = P(\boldsymbol{b})\exp[-|\boldsymbol{r}-\boldsymbol{H}\boldsymbol{b}|^2/\sigma^2]/Z$, with Z being the normalization and $\boldsymbol{H} = \sqrt{\boldsymbol{R}_\mathrm{r}}\boldsymbol{\Xi}\sqrt{\boldsymbol{R}_\mathrm{t}}$, we have

$$\Phi(\boldsymbol{m}) = \underset{\boldsymbol{h}}{\mathrm{Extr}}\left\{\frac{|\boldsymbol{r}-\boldsymbol{H}\boldsymbol{m}|^2}{\sigma^2} - \ln \underset{\boldsymbol{b}}{\mathrm{Tr}}\left(P(\boldsymbol{b}) \times \exp\left\{-\frac{|\sqrt{\boldsymbol{R}_\mathrm{r}}\boldsymbol{\Xi}\sqrt{\boldsymbol{R}_\mathrm{t}}(\boldsymbol{b}-\boldsymbol{m})|^2}{\sigma^2} + \mathrm{Re}[\boldsymbol{h}^\dagger(\boldsymbol{b}-\boldsymbol{m})]\right\}\right)\right\} + \ln Z. \tag{15}$$

Note that the extremization argument $\boldsymbol{h}$ is shifted as $\boldsymbol{h} + 2(\boldsymbol{H}^\dagger\boldsymbol{r} - \boldsymbol{H}^\dagger\boldsymbol{H}\boldsymbol{m})/\sigma^2 \to \boldsymbol{h}$ for making the calculation simpler. Although this distribution of the vector $\sqrt{\boldsymbol{R}_\mathrm{t}}(\boldsymbol{b}-\boldsymbol{m})$ is not isotropic but rather is biased by the matrix $\sqrt{\boldsymbol{R}_\mathrm{t}}$, the multiplication by the rectangular random matrix $\boldsymbol{\Xi}$ ensures the following approximation under the constraints $\chi = |\sqrt{\boldsymbol{R}_\mathrm{t}}(\boldsymbol{b}-\boldsymbol{m})|^2/K$ and $\kappa = \mathrm{Re}[\boldsymbol{h}^\dagger(\boldsymbol{b}-\boldsymbol{m})]/K$, where we introduce the auxiliary variables χ and κ, as follows:

$$\underset{\boldsymbol{b}}{\mathrm{Tr}} P(\boldsymbol{b})\exp\{-|\sqrt{\boldsymbol{R}_\mathrm{r}}\boldsymbol{\Xi}\sqrt{\boldsymbol{R}_\mathrm{t}}(\boldsymbol{b}-\boldsymbol{m})|^2/\sigma^2 + \mathrm{Re}[\boldsymbol{h}^\dagger(\boldsymbol{b}-\boldsymbol{m})]\}$$
$$\simeq \int d\chi \int d\kappa \exp[K\{G_{\boldsymbol{\Xi}^\dagger\boldsymbol{R}_\mathrm{r}\boldsymbol{\Xi}}(-\chi/\sigma^2) + \kappa\}] \times \underset{\boldsymbol{b}}{\mathrm{Tr}}\{P(\boldsymbol{b})\delta[|\sqrt{\mathbf{R}_\mathrm{t}}(\boldsymbol{b}-\boldsymbol{m})|^2 - K\chi]\delta[\mathrm{Re}[\boldsymbol{h}^\dagger(\boldsymbol{b}-\boldsymbol{m})] - K\kappa\}\}, \tag{16}$$

where $G_{\boldsymbol{\Xi}^\dagger\boldsymbol{R}_\mathrm{r}\boldsymbol{\Xi}}(x)$ is given by Eq. (7). Saddle point evaluation with respect to χ and κ yields the following approximate expression of the Gibbs free energy:

$$\Phi(\boldsymbol{m}) \simeq \underset{\chi,\hat{\chi}}{\mathrm{Extr}}\left\{\frac{|\boldsymbol{r}-\boldsymbol{H}\boldsymbol{m}|^2}{\sigma^2} - K G_{\boldsymbol{\Xi}^\dagger\boldsymbol{R}_\mathrm{r}\boldsymbol{\Xi}}\left(-\frac{\chi}{\sigma^2}\right) - K\hat{\chi}\chi + \Phi_\mathrm{t}(\boldsymbol{m};\hat{\chi})\right\} + \ln Z,$$

where $\Phi_\mathrm{t}(\boldsymbol{m};\hat{\chi}) = \underset{\boldsymbol{h}}{\mathrm{Extr}}\{-\ln \underset{\boldsymbol{b}}{\mathrm{Tr}}(\exp\{-\hat{\chi}|\sqrt{\boldsymbol{R}_\mathrm{t}}(\boldsymbol{b}-\boldsymbol{m})|^2 + \mathrm{Re}[\boldsymbol{h}^\dagger(\boldsymbol{b}-\boldsymbol{m})]\})\}$. (17)

From the Gibbs free energy we obtain a set of equations for estimating $\boldsymbol{m}$, and using these equations, we construct the demodulation algorithm or the method for finding the minimization argument $\boldsymbol{m}$ for the Gibbs free energy. In the following, we summarize the procedure for the minimization of $\Phi(\boldsymbol{m})$.

(i) (Step 0) Initialize variables as $\chi^{(0)}=1, \boldsymbol{m}^{(0)}=\boldsymbol{0}, \boldsymbol{h}^{(0)}=\boldsymbol{0}$ for step $t=0$ and set the number of steps as $t=1$.

(ii) (Step 1) For the tth step, update $\hat{\chi}$ and $\boldsymbol{h}$ as

$$\hat{\chi}^{(t)} = \frac{1}{\sigma^2}G'\left(-\frac{\chi^{(t-1)}}{\sigma^2}\right),$$

$$\boldsymbol{h}^{(t)} = \boldsymbol{h}^{(t-1)} + \sigma^2\hat{\chi}^{(t)}\left(2\frac{\boldsymbol{H}^\dagger\boldsymbol{r} - \boldsymbol{H}^\dagger\boldsymbol{H}\boldsymbol{m}^{(t-1)}}{\sigma^2} - \boldsymbol{h}^{(t-1)}\right).$$

(iii) (Step 2) Update $\boldsymbol{m}$ as

$$\boldsymbol{m}^{(t)} = \langle \boldsymbol{b} \rangle_t,$$

$$\chi^{(t)} = \frac{1}{K}\{\langle \boldsymbol{b}^\dagger\boldsymbol{R}_\mathrm{t}\boldsymbol{b}\rangle_t - \boldsymbol{m}^{(t)\dagger}\boldsymbol{R}_\mathrm{t}\boldsymbol{m}^{(t)}\},$$

where $\langle\cdot\rangle_t$ denotes the expectation

$$\langle f(\boldsymbol{b})\rangle_t \equiv \frac{\mathrm{Tr}_{\boldsymbol{b}} P(\boldsymbol{b})\exp[-\hat{\chi}^{(t)}|\sqrt{\boldsymbol{R}_\mathrm{t}}(\boldsymbol{b}-\boldsymbol{m})|^2 + \mathrm{Re}(\boldsymbol{h}^\dagger\boldsymbol{b})]f(\boldsymbol{b})}{\mathrm{Tr}_{\boldsymbol{b}} P(\boldsymbol{b})\exp[-\hat{\chi}^{(t)}|\sqrt{\boldsymbol{R}_\mathrm{t}}(\boldsymbol{b}-\boldsymbol{m})|^2 + \mathrm{Re}(\boldsymbol{h}^\dagger\boldsymbol{b})]}.$$

(iv) (Step 3) Update the number of recursion steps $t \to t+1$. Return to step 1 unless these variables converge, otherwise stop.

After termination of the above procedure, the transmit signal is estimated as $\hat{\boldsymbol{b}} = \mathrm{argmin}_{\boldsymbol{b}}\{|\boldsymbol{b}-\boldsymbol{m}^{(t)}|\}$.

The computational cost of step 1 is $O(KL)$ and is sufficiently small. In general, the cost of step 2 is not so small. However, we can reduce the cost of step 2 for the special forms of matrix $\boldsymbol{R}_\mathrm{t}$. For instance, when $\boldsymbol{R}_\mathrm{t}$ is a matrix of tridiagonal form, as considered in the next section, step 2 can be executed using the transfer matrix method. In such a case, the cost is $O(K)$, which is smaller than the cost of step 1.

IV. SIMPLE EXAMPLE

A. One-dimensional chain model

In Sec. II we have already discussed the discrepancy between two mutual information of the Kronecker model, the one by the rigorous evaluation and the one by the approximate evaluation averaging multiplied unitary matrix. For numerically validating the results obtained analytically, we performed numerical experiments for a simple but nontrivial example of Kronecker channel model $\boldsymbol{H}=\sqrt{\boldsymbol{R}_{\mathrm{r}}}\boldsymbol{\Xi}\sqrt{\boldsymbol{R}_{\mathrm{t}}}$, for which $\boldsymbol{R}_{\mathrm{r}}$ and $\boldsymbol{R}_{\mathrm{t}}$ are given as identity and tridiagonal matrices, respectively. More precisely, the correlation matrix of the transmission antennas is provided as

$$\boldsymbol{R}_{\mathrm{t}} \equiv \boldsymbol{I}+\rho\boldsymbol{R} = \begin{pmatrix} 1 & \rho & 0 & 0 & \cdots \\ \rho & 1 & \rho & 0 & \cdots \\ 0 & \rho & 1 & \rho & \cdots \\ 0 & 0 & \rho & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \qquad (18)$$

$$\boldsymbol{R} = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots \\ 1 & 0 & 1 & 0 & \cdots \\ 0 & 1 & 0 & 1 & \cdots \\ 0 & 0 & 1 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

which implies that correlations are taken into account only for pairs of adjacent transmission antennas. This model is referred to as the one-dimensional chain model because it can be mapped to a one-dimensional random-field Ising chain model as we see later. For simplicity, we analyze the real channel, and accordingly, all variables are set to be real. Here, $\boldsymbol{\Xi}$ represents the $L\times K$-dimensional iid Gaussian random matrix $\mathcal{N}(0,1/L)^{KL}$, $\boldsymbol{b}=(b_k)\in\mathbb{R}^K$ is a BPSK-modulated transmit signal ($b_k\in\{\pm 1\}$), and $\boldsymbol{\eta}\in\mathbb{R}^L$ is a normalized real Gaussian-distributed random vector $\mathcal{N}(0,1)$. The formulation so far for the complex channel can be reconstructed without difficultly for the real channel just by the replacement of unitary matrix $\boldsymbol{U}$ with orthogonal matrix $\boldsymbol{O}$.

In the following we analyze how different the conditional entropy and the bit error rate (BER) are between the system with tridiagonal transmission correlation matrix $\boldsymbol{R}_{\mathrm{t}}=\boldsymbol{I}+\rho\boldsymbol{R}$, which describes the correlation propagating through adjacent transmission antennas [Fig. 1(B)], and the one with $\boldsymbol{R}_{\mathrm{t}}=\boldsymbol{O}^T(\boldsymbol{I}+\rho\boldsymbol{R})\boldsymbol{O}=\boldsymbol{I}+\rho\boldsymbol{O}^T\boldsymbol{R}\boldsymbol{O}$ representing the system whose correlations are smeared by the multiplication of orthogonal matrix [Fig. 1(C)], which corresponds to the system evaluated by averaging over multiplied orthogonal matrix. As we see in Sec. II, there is no difference for $\boldsymbol{R}_{\mathrm{r}}$-dependent part in the mutual information between two evaluations, namely, the rigorous one and the one by averaging over multiplied unitary matrix. This is why we are allowed to set $\boldsymbol{R}_{\mathrm{r}}=\boldsymbol{I}$ for convenience of the analysis.

B. Analysis

Before the analysis of the entire Kronecker model, let us evaluate three pieces of mutual information, namely, $I_1(\chi)$, $I_{\boldsymbol{I}+\rho\boldsymbol{O}^T\boldsymbol{R}\boldsymbol{O}}(\chi)$, and $I_{\boldsymbol{I}+\rho\boldsymbol{R}}(\chi)$ [24], that appear as the partial mutual information in the expression of the mutual information of the entire system $\mathcal{I}_{\boldsymbol{H}}$, as described in Sec. II. For the BPSK modulation, the mutual information of the single-input single-output channel $I_1(\chi)$ is given by

$$I_1(\chi) = \chi - \int Dz\, \ln[\cosh(\chi+\sqrt{\chi}z)], \qquad (19)$$

where $Dz\equiv\exp(-z^2/2)/\sqrt{2\pi}$. For mutual information $I_{\boldsymbol{I}+\rho\boldsymbol{O}^T\boldsymbol{R}\boldsymbol{O}}(\chi)$, substitution of the tridiagonal form of $\boldsymbol{R}$ yields

$$\overline{I_{\boldsymbol{I}+\rho\boldsymbol{O}^T\boldsymbol{R}\boldsymbol{O}}(\chi)} = \underset{\lambda}{\mathrm{Extr}}\{\hat{G}_{\boldsymbol{I}+\rho\boldsymbol{O}^T\boldsymbol{R}\boldsymbol{O}}(\lambda) + I_1(\lambda\chi)\}. \qquad (20)$$

Here, $\hat{G}_{\boldsymbol{I}+\rho\boldsymbol{O}^T\boldsymbol{R}\boldsymbol{O}}(\lambda)=-(1/2)\ln[1-(\lambda-1)^2/4\rho^2]$, which is evaluated using the Stieltjes inversion formula for the function $G(\chi)$ (see Ref. [12]) and the relation $\rho_{\boldsymbol{R}_{\mathrm{t}}}(\lambda)=\lim_{\epsilon\to 0}(\pi N)^{-1}\partial_\lambda\,\mathrm{Im}\ln\det(\boldsymbol{R}_{\mathrm{t}}-(\lambda-i\epsilon)\boldsymbol{I})$. As described earlier, the discrepancy between $I_{\boldsymbol{I}+\rho\boldsymbol{R}}(\chi)$ and $I_{\boldsymbol{I}+\rho\boldsymbol{O}^T\boldsymbol{R}\boldsymbol{O}}(\chi)$ appears starting from the fourth-order term, which is expressed as the tridiagonal form of $\boldsymbol{R}$, as follows:

$$I_{\boldsymbol{I}+\rho\boldsymbol{R}}(\chi) = \overline{I_{\boldsymbol{I}+\rho\boldsymbol{O}^T\boldsymbol{R}\boldsymbol{O}}(\chi)} - \frac{(\rho\chi)^4}{4} \times\left(\{-2I_1''(\chi)-[2I_1'(\chi)]^2\}^2 + \frac{\hat{C}(\chi)^2}{6}\right) + O(\rho^5), \qquad (21)$$

where

$$\hat{C}(\chi) \equiv -2\int Dz\{1-\tanh^2(\chi+\sqrt{\chi}z)\}\{1-3\tanh^2(\chi+\sqrt{\chi}z)\}. \qquad (22)$$

For the tridiagonal form of $\boldsymbol{R}$, we can evaluate $I_{\boldsymbol{I}+\rho\boldsymbol{R}}(\chi)$ exactly using the transfer matrix method for the Ising chain in random fields. In order to demonstrate how this is accomplished, we transform the mutual information as follows:

$$\begin{aligned} I_{\boldsymbol{I}+\rho\boldsymbol{R}}(\chi) &= -\frac{1}{K}\int_{\mathbb{R}^K} d\boldsymbol{r}'\left\{\mathop{\mathrm{Tr}}_{\boldsymbol{b}}\frac{1}{2^K}\left(\frac{\chi}{2\pi}\right)^{K/2} \times\exp\left[-\frac{\chi}{2}|\boldsymbol{r}'-\sqrt{\boldsymbol{I}+\rho\boldsymbol{R}}\boldsymbol{b}|^2\right]\right\} \\ &\quad\times\ln\left\{\mathop{\mathrm{Tr}}_{\boldsymbol{b}}\frac{1}{2^K}\left(\frac{\chi}{2\pi}\right)^{K/2}\exp\left[-\frac{\chi}{2}|\boldsymbol{r}'-\sqrt{\boldsymbol{I}+\rho\boldsymbol{R}}\boldsymbol{b}|^2\right]\right\} \\ &\quad-\frac{1}{2}\ln\left(\frac{2\pi}{\chi}\right)-\frac{1}{2} \\ &\simeq -\frac{1}{2^K K}\int_{\mathbb{R}^K} D\boldsymbol{\eta}\mathop{\mathrm{Tr}}_{\bar{\boldsymbol{b}}}\ln \times\left\{\mathop{\mathrm{Tr}}_{\boldsymbol{b}}\prod_{k=1}^{K-1}\phi(b_k,b_{k+1}|\bar{b}_k,\bar{b}_{k+1},\eta_k)\right\} \\ &= -\frac{1}{2^K K}\int_{\mathbb{R}^K} D\boldsymbol{\eta}\mathop{\mathrm{Tr}}_{\bar{\boldsymbol{\tau}}}\ln\left\{\mathop{\mathrm{Tr}}_{\boldsymbol{\tau}}\prod_{k=1}^{K-1}\phi(\tau_k,\tau_{k+1}|\bar{\tau}_k,\eta_k)\right\}, \end{aligned} \qquad (23)$$

where a trivial constant is neglected in the second and the third lines. Note that gauge transformation as $\boldsymbol{b}\to\boldsymbol{\tau}$ and $\bar{\boldsymbol{b}}\to\bar{\boldsymbol{\tau}}$, where $\tau_k=b_k\bar{b}_k$ and $\bar{\tau}_k=\bar{b}_{k+1}\bar{b}_k$, and redefinition of $\boldsymbol{\eta}$ are used for simplification of the expression. Matrix elements $\phi(b_k,b_{k+1}|\bar{b}_k,\bar{b}_{k+1},\eta_k)$ and $\phi(\tau_k,\tau_{k+1}|\bar{\tau}_k,\eta_k)$ are defined as follows:

$$\phi(b_k,b_{k+1}|\bar{b}_k,\bar{b}_{k+1},\eta_k)=\frac{1}{2}\exp\left[-\frac{\chi}{2}|l_0(b_k-\bar{b}_k)+l_1(b_{k+1}-\bar{b}_{k+1})|^2+\sqrt{\chi}\eta_k(l_0b_k+l_1b_{k+1})\right],$$

$$\phi(\tau_k,\tau_{k+1}|\bar{\tau}_k,\eta_k)=\frac{1}{2}\exp\left[-\frac{\chi}{2}|l_0(\tau_k-1)+l_1\bar{\tau}_k(\tau_{k+1}-1)|^2+\sqrt{\chi}\eta_k(l_0\tau_k+l_1\bar{\tau}_k\tau_{k+1})\right], \quad (24)$$

where l_0 and l_1 are real constants that are obtained by Cholesky decomposition, i.e., $\boldsymbol{I}+\rho\boldsymbol{R}=\boldsymbol{\Lambda}^T\boldsymbol{\Lambda}$, where $\Lambda_{kk}=l_0,\Lambda_{(k+1)k}=l_1$, and zero otherwise;

$$\boldsymbol{R}_{\mathrm{t}}\equiv\boldsymbol{I}+\rho\boldsymbol{R}=\begin{pmatrix}\ddots & \vdots & \vdots & \vdots & \vdots & \cdots\\ \cdots & 1 & \rho & 0 & 0 & \cdots\\ \cdots & \rho & 1 & \rho & 0 & \cdots\\ \cdots & 0 & \rho & 1 & \rho & \cdots\\ \cdots & 0 & 0 & \rho & 1 & \cdots\\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots\end{pmatrix}=\boldsymbol{\Lambda}^T\boldsymbol{\Lambda}$$
$$\simeq\begin{pmatrix}\ddots & \vdots & \vdots & \vdots & \vdots & \cdots\\ \cdots & l_0 & 0 & 0 & 0 & \cdots\\ \cdots & l_1 & l_0 & 0 & 0 & \cdots\\ \cdots & 0 & l_1 & l_0 & 0 & \cdots\\ \cdots & 0 & 0 & l_1 & l_0 & \cdots\\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots\end{pmatrix}\begin{pmatrix}\ddots & \vdots & \vdots & \vdots & \vdots & \cdots\\ \cdots & l_0 & l_1 & 0 & 0 & \cdots\\ \cdots & 0 & l_0 & l_1 & 0 & \cdots\\ \cdots & 0 & 0 & l_0 & l_1 & \cdots\\ \cdots & 0 & 0 & 0 & l_0 & \cdots\\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots\end{pmatrix}. \quad (25)$$

These constants satisfy $l_0^2+l_1^2=1$, $l_0l_1=\rho$, and $l_0\geq l_1$ [25]. The matrix element of Eq. (24) corresponds to the Boltzmann weight of the Ising chain coupled with bimodal ($\bar{\boldsymbol{\tau}}$) and Gaussian ($\boldsymbol{\eta}$) random fields and consequently the BER for the demodulation $\hat{b}_k\equiv\arg\max_{b_k}\{\mathrm{Tr}_{\boldsymbol{b}\backslash b_k}P(\boldsymbol{b}|\boldsymbol{r})\}$, whose definition is given by $\mathrm{BER}\equiv\Sigma_k(1-b_k\hat{b}_k)/2K$, is evaluated analytically for the random-field Ising chain.

Several methods of analysis have been developed for handling the random-field Ising chain [26,27], and in the present study, we use the technique of belief propagation, which is equivalent to the transfer matrix method. After parametrization of the belief from site k to site $k+1$ by cavity field $h_{\to k+1}$ as $\mu(\tau_{k+1})=e^{h_{\to k+1}\tau_{k+1}}/2\cosh(h_{\to k+1})$, the propagation process from $h_{\to k}$ to $h_{\to k+1}$ is written as

$$\begin{aligned}h_{\to k+1}&=\mathrm{arctanh}\left[\sum_{\tau_{k+1}}\mu(\tau_{k+1})\tau_{k+1}\right]\\&=\mathrm{arctanh}\left[\frac{\Sigma_{\tau_k\tau_{k+1}}\phi(\tau_k,\tau_{k+1}|\bar{\tau}_k\eta_k)\exp(h_{\to k}\tau_k)\tau_{k+1}}{\Sigma_{\tau_k\tau_{k+1}}\phi(\tau_k,\tau_{k+1}|\bar{\tau}_k\eta_k)\exp(h_{\to k}\tau_k)}\right]\\&=\chi l_1^2+\chi\rho\bar{\tau}_k+\sqrt{\chi}l_1\bar{\tau}_k\eta_k\\&\quad-\bar{\tau}_k\,\mathrm{arctanh}[\tanh(\rho\chi)\tanh(h_{\to k}+\chi l_0^2+\chi\rho\bar{\tau}_k\\&\quad+\sqrt{\chi}l_0\eta_k)].\end{aligned} \quad (26)$$

Similarly, for the opposite direction, denoted by $h_{k\leftarrow}$,

$$\begin{aligned}h_{k\leftarrow}&=\mathrm{arctanh}\left[\frac{\Sigma_{\tau_k\tau_{k+1}}\phi(\tau_k,\tau_{k+1}|\bar{\tau}_k\eta_k)\exp(h_{k+1\leftarrow}\tau_{k+1})\tau_k}{\Sigma_{\tau_k\tau_{k+1}}\phi(\tau_k,\tau_{k+1}|\bar{\tau}_k\eta_k)\exp(h_{k+1\leftarrow}\tau_{k+1})}\right]\\&=\chi l_0^2+\chi\rho\bar{\tau}_k+\sqrt{\chi}l_0\eta_k-\bar{\tau}_k\,\mathrm{arctanh}[\tanh(\rho\chi)\\&\quad\times\tanh(h_{k+1\leftarrow}+\chi l_1^2+\chi\rho\bar{\tau}_k+\sqrt{\chi}l_1\bar{\tau}_k\eta_k)].\end{aligned} \quad (27)$$

The stationary distributions of beliefs for the numerically increasing and decreasing directions, denoted by π_+ and π_-, respectively, satisfy the following conditions:

$$\begin{aligned}\pi_+(h_{\to k+1})&=\int\pi_+(h_{\to k})dh_{\to k}\int D\eta\sum_{\bar{\tau}=\pm1}\frac{1}{2}\delta\{h_{\to k+1}\\&\quad-(\chi l_1^2+\chi\rho\bar{\tau}+\sqrt{\chi}l_1\bar{\tau}\eta)\\&\quad+\bar{\tau}\,\mathrm{arctanh}[\tanh(\rho\chi)\tanh(h_{\to k}+\chi l_0^2+\chi\rho\bar{\tau}\\&\quad+\sqrt{\chi}l_0\eta)]\},\end{aligned} \quad (28)$$

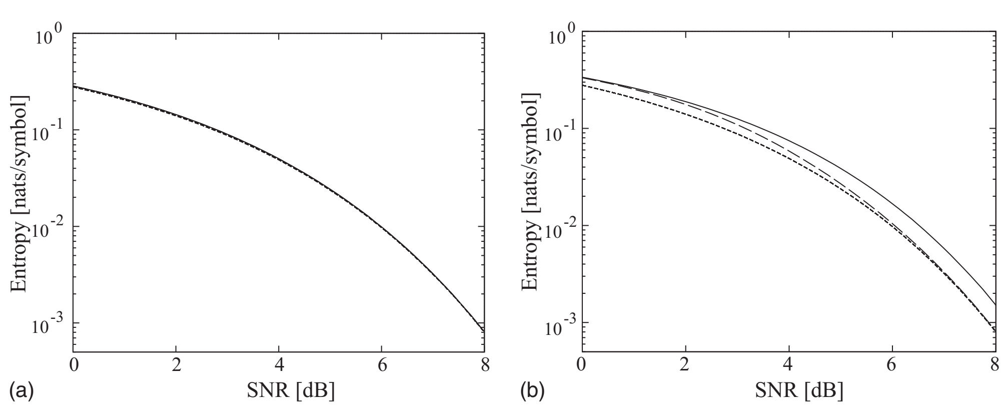

FIG. 2. Conditional entropy $h_{\boldsymbol{I}+\rho\boldsymbol{R}}(\boldsymbol{b}|\boldsymbol{r})$ for chainlike system [Eq. (1)]. Entropy vs SNR [(a) $\rho=0.2$ and (b) $\rho=0.5$]. The solid lines show the results obtained by exact analysis, the broken lines show the results obtained by matrix integration, and the dotted lines show the results for iid channel.

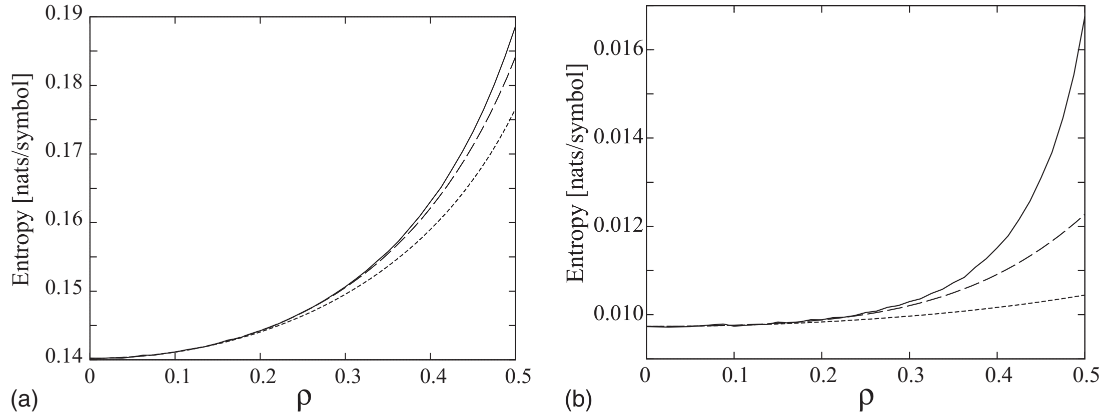

FIG. 3. Conditional entropy $h_{\boldsymbol{I}+\rho\boldsymbol{R}}(\boldsymbol{b}|\boldsymbol{r})$ for chainlike system [Eq. (2)]. Entropy vs correlation parameter ρ [(a) SNR=2 dB and (b) SNR=6 dB]. The solid lines show the results obtained by exact analysis, the broken lines show the results obtained by matrix integration with correction from the fourth order, and the dotted lines show the results obtained by matrix integration without correction.

$$\pi_-(h_{k\leftarrow}) = \int \pi_-(h_{k+1\leftarrow})dh_{k+1\leftarrow}\int D\eta\sum_{\bar{\tau}=\pm1}\frac{1}{2}\delta\{h_{k\leftarrow} - (\chi l_0^2+\chi\rho\bar{\tau}+\sqrt{\chi}l_0\eta) + \bar{\tau}\,\mathrm{arctanh}[\tanh(\rho\chi)\tanh(h_{k+1\leftarrow}+\chi l_1^2+\chi\rho\bar{\tau} + \sqrt{\chi}l_1\bar{\tau}\eta)]\}, \quad (29)$$

where $\int\pi_+(h_\rightarrow)dh_\rightarrow=\int\pi_-(h_\leftarrow)dh_\leftarrow=1$. (Absence of the site index means arbitrary site.) The functions $\pi_+(h_\rightarrow)$ and $\pi_-(h_\leftarrow)$ can be obtained numerically by the Monte Carlo method for a one-dimensional system. The bit error rate P_b is represented by $P_b=\int\pi_+(h_\rightarrow)dh_\rightarrow\int\pi_-(h_\leftarrow)dh_\leftarrow[1-\mathrm{sgn}(h_\rightarrow+h_\leftarrow)]/2$.

C. Result

In order to investigate the performance of communication channels, conditional entropy $h(\boldsymbol{b}|\boldsymbol{r})=h(\boldsymbol{b})-\mathcal{I}_{\boldsymbol{H}}$, where $h(\boldsymbol{b})$ denotes entropy, is a favorable measure because $h(\boldsymbol{b}|\boldsymbol{r})$ decreases to zero under smaller noise power. Figures 2 and 3 show the exact conditional entropy $h_{\boldsymbol{I}+\rho\boldsymbol{R}}(\boldsymbol{b}|\boldsymbol{r})$ estimated by the Monte Carlo method, the approximate entropy $h_{\boldsymbol{I}+\rho\boldsymbol{O}^T\boldsymbol{RO}}(\boldsymbol{b}|\boldsymbol{r})$ obtained by matrix integration, and the entropy of the iid channel $h_{\boldsymbol{I}}(\boldsymbol{b}|\boldsymbol{r})$. In both graphs, the entropy obtained by matrix integration, i.e., $h_{\boldsymbol{I}+\rho\boldsymbol{O}^T\boldsymbol{RO}}(\boldsymbol{b}|\boldsymbol{r})$, does not exceed the entropy obtained by exact evaluation $h_{\boldsymbol{I}+\rho\boldsymbol{R}}(\boldsymbol{b}|\boldsymbol{r})$, which implies the inequality $I_{\boldsymbol{I}+\rho\boldsymbol{R}}(\boldsymbol{b},\boldsymbol{r})\le I_{\boldsymbol{I}+\rho\boldsymbol{O}^T\boldsymbol{RO}}(\boldsymbol{b},\boldsymbol{r})$, which is given up to the fourth order of perturbation in Sec. II. The analysis indicates that the deviation of the approximate result from the exact result depends on the signal-to-noise ratio (SNR), defined in the present case by SNR $\equiv 10\log_{10}(1/2\sigma^2)$ (decibel=dB unit), and the correlation parameter ρ. For a small SNR and small correlation (=small ρ) the deviation is small, which means that the entropy, $h_{\boldsymbol{I}+\rho\boldsymbol{O}^T\boldsymbol{RO}}(\boldsymbol{b}|\boldsymbol{r})$ obtained by matrix integration gives a good approximation of the exact entropy, $h_{\boldsymbol{I}+\rho\boldsymbol{R}}(\boldsymbol{b}|\boldsymbol{r})$, while the deviation becomes greater in the case of a large SNR or large correlation.

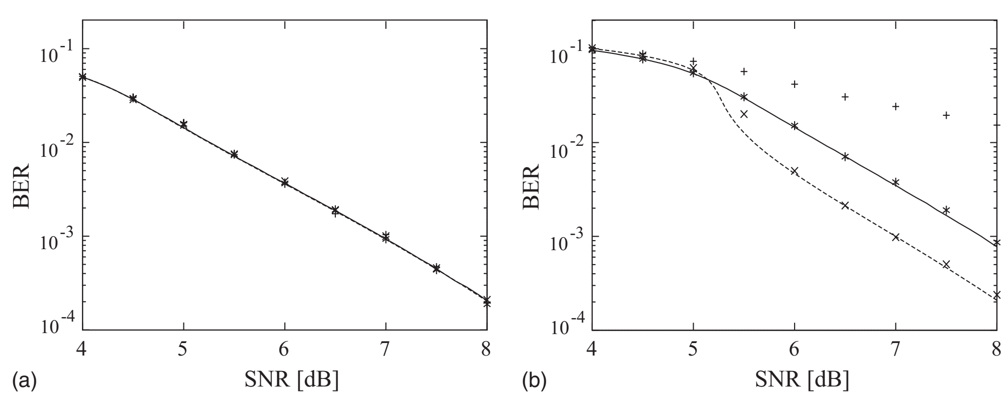

FIG. 4. Performance by the replica analysis and the result of demodulation experiment [Eq. (1)]. BER vs SNR [(a) ρ=0.2 and (b) ρ=0.5]. We set the parameters K=4400 and L=4000 and take the average of the results over 128 samples with various input signals $\boldsymbol{b}$, matrices $\boldsymbol{\Xi}$, and noises $\boldsymbol{\eta}$. We also varied the orthogonal matrix $\boldsymbol{O}$ for the model with orthogonal-matrix multiplication. The lines depict the results of the replica analysis for two models, the chainlike model $\boldsymbol{R}_t=\boldsymbol{I}+\rho\boldsymbol{R}$ with tridiagonal $\boldsymbol{R}$ (solid) and the model with orthogonal-matrix multiplication $\boldsymbol{R}_t=\boldsymbol{I}+\rho\boldsymbol{O}^T\boldsymbol{RO}$ (dotted). The symbols in the figure denote the results of demodulations, namely, demodulation for the chainlike model (*), demodulation for the model with orthogonal-matrix multiplication (×), inappropriate choice of the demodulation algorithm, i.e., the demodulation algorithm for the orthogonal-matrix multiplication model applied to the chainlike model (+).

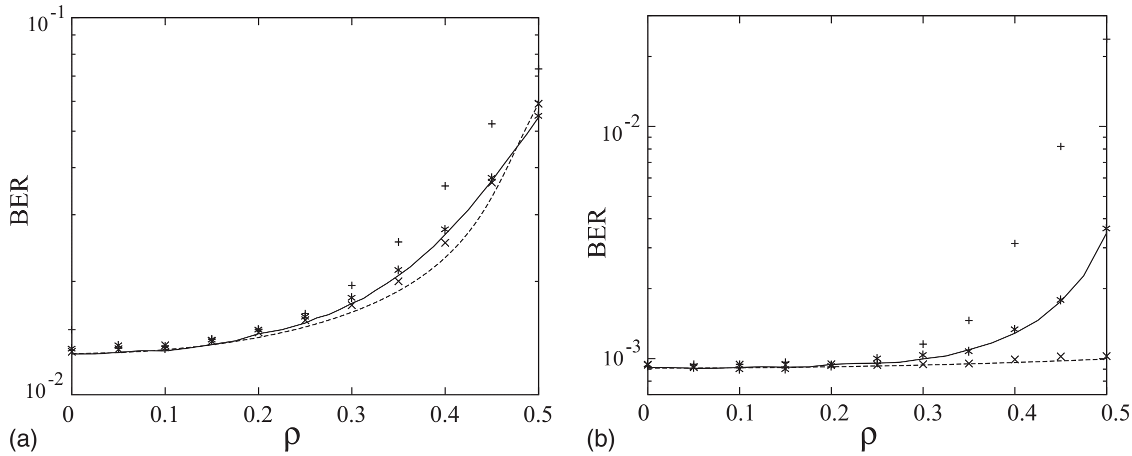

FIG. 5. Performance by the replica analysis and the results of demodulation experiment [Eq. (2)]. BER vs correlation parameter ρ [(a) SNR=5 dB and (b) SNR=7 dB']. We set the parameter $K=4400$ and $L=4000$ and take the average over 1024 samples for various values of $\boldsymbol{b}$, $\boldsymbol{\Xi}$, $\boldsymbol{\eta}$, and $\boldsymbol{O}$. The lines depict the results of the replica analysis for two models, the chainlike model $\boldsymbol{R}_t=\boldsymbol{I}+\rho\boldsymbol{R}$ with tridiagonal $\boldsymbol{R}$ (solid) and the model with orthogonal-matrix multiplication $\boldsymbol{R}_t=\boldsymbol{I}+\rho\boldsymbol{O}^T\boldsymbol{R}\boldsymbol{O}$ (dotted). The symbols in the figure denote the results of demodulations, namely, demodulation for the chainlike model $(*)$, demodulation for the model with orthogonal-matrix multiplication $(\times)$, inappropriate choice of the demodulation algorithm, i.e., the demodulation algorithm for the orthogonal-matrix multiplication model applied to the chain-like model $(+)$.

As we have discussed earlier, the approximate evaluation with matrix integration is useful because this simplifies the analysis. However, as shown by the numerical results for conditional entropy, this method is only valid when the correlation is small.

Next, we examine whether the proposed demodulation algorithm is practical. In Figs. 4 and 5, the results of demodulation for the Kronecker channels are depicted. These figures show two BER curves, namely, the BER obtained by exact analysis of the model $\boldsymbol{R}_t=\boldsymbol{I}+\rho\boldsymbol{R}$ with tridiagonal $\boldsymbol{R}$, as described in the previous subsection, and the BER obtained by the correlation matrix multiplied by an arbitrary orthogonal matrix, $\boldsymbol{R}_t=\boldsymbol{I}+\rho\boldsymbol{O}^T\boldsymbol{R}\boldsymbol{O}$ and averaged over the orthogonal matrix $\boldsymbol{O}$. For the latter model, matrix integration analysis can be applied due to multiplication of the orthogonal matrix. We have proposed an appropriate demodulation algorithm for each evaluation. For the former model, without the orthogonal matrix multiplication, we can use the belief propagation algorithm proposed in the previous subsection. For the latter model, with orthogonal matrix multiplication, the demodulation algorithm we proposed in [12,13] based on matrix integration and the Thouless-Anderson-Palmer method [21–23] is applicable.

The results of demodulation for each case show good agreement with the results obtained by replica analysis. For large noise power and large ρ [Figs. 4(B) and 5(B)], the BER of the model without orthogonal-matrix multiplication becomes larger than that with orthogonal-matrix multiplication, which reflects the discrepancy of mutual information from higher-order values of ρ, as mentioned in Sec. II. Therefore, in designing the demodulation algorithm for such correlated channels, appropriate treatment of the correlation matrix should be taken into consideration. We also examined the convergence speed of the algorithm. The proposed algorithm in the previous subsection requires the $O(K^2)$ matrix operation and the $O(K)$ belief propagation process in each step, and convergence of this algorithm requires dozens of iterations. Therefore, we conclude that this algorithm is computationally feasible.

V. SUMMARY

In the present paper, we proposed a performance analysis scheme and a demodulation algorithm for the Kronecker channel model in MIMO wireless communication systems. For a more exact evaluation than that of our previous paper using the matrix integration formula, we demonstrated that two separated manipulation steps for the product form of the channel matrix, i.e., Gaussian integration for the channel matrix and an appropriate scheme for dealing with transmitter correlation, are important for the correlated MIMO system. The numerical result for the tridiagonal correlation matrix model shows that the proposed scheme and algorithm are useful for performance analysis and for the construction of a practical demodulation algorithm.

ACKNOWLEDGMENTS

This work was supported by a Grant-in-Aid Scientific Research on Priority Areas "Deepening and Expansion of Statistical Mechanical Informatics (DEX-SMI)" from MEXT, Japan Grant No. 18079006. Y.K. was also supported by the JSPS Global COE program, "Computationism as a Foundation for the Sciences."

APPENDIX A: MUTUAL INFORMATION

We start from the expression in Eq. (5). For $n=0,1,2,\ldots$, we have

$$\int_{\mathbb{C}^L} d\boldsymbol{r} Z^{n+1}(\boldsymbol{r}) = \int_{\mathbb{C}^L} d\boldsymbol{r} \prod_{a=0}^{n} \left(\operatorname*{Tr}_{\boldsymbol{b}^a} P(\boldsymbol{b}^a) \frac{1}{(\pi\sigma^2)^L} \times \exp\left[-\frac{|\boldsymbol{r}-\boldsymbol{H}\boldsymbol{b}^a|^2}{\sigma^2} \right] \right)$$
$$= \left(\frac{\pi\sigma^2}{n+1} \right)^L \left(\prod_{a=0}^{n} \operatorname*{Tr}_{\boldsymbol{b}^a} P(\boldsymbol{b}^a) \frac{1}{(\pi\sigma^2)^L} \right) \times \exp\left[-\frac{1}{\sigma^2} \mathrm{Tr}(\boldsymbol{H}^\dagger \boldsymbol{H}\boldsymbol{L}) \right], \tag{A1}$$

where the K-dimensional square matrix $\boldsymbol{L}$ is defined by $L_{kk'} \equiv \Sigma_{a=0}^{n} b_k^a b_{k'}^{a*} - \Sigma_{ab} b_k^a b_{k'}^{b*}/(n+1)$. This can be rewritten as $L_{kk'} = \Sigma_{ab} b_k^a \mathcal{P}^{ab} b_{k'}^{b*}$ by introducing the $(n+1)$-dimensional projection matrix $\mathcal{P}^{ab} \equiv \delta^{ab} - 1/(n+1)$. Substituting channel matrix $\boldsymbol{H}$, given in Eq. (3), for the Kronecker model and integrating with respect to Ξ, we obtain

$$\overline{\int_{\mathbb{C}^L} d\boldsymbol{r} Z^{n+1}(\boldsymbol{r})} \propto \left[\prod_{a=0}^{n} \operatorname*{Tr}_{\boldsymbol{b}^a} P(\boldsymbol{b}^a) \right] \times \left[\det\left(\boldsymbol{I}_{LK} + \frac{1}{\sigma^2 L} \boldsymbol{R}_\mathrm{r} \otimes \sqrt{\boldsymbol{R}_\mathrm{t}} \boldsymbol{L} \sqrt{\boldsymbol{R}_\mathrm{t}}^\dagger \right) \right]^{-1}$$
$$= \left[\prod_{a=0}^{n} \operatorname*{Tr}_{\boldsymbol{b}^a} P(\boldsymbol{b}^a) \right] \exp\left[-L \int d\lambda \rho_{\boldsymbol{R}_\mathrm{r}}(\lambda) \times \mathrm{Tr}\ln\left(\boldsymbol{I}_K + \frac{\lambda}{\sigma^2 L} \sqrt{\boldsymbol{R}_\mathrm{t}} \boldsymbol{L} \sqrt{\boldsymbol{R}_\mathrm{t}}^\dagger \right) \right]$$
$$= \left[\prod_{a=0}^{n} \operatorname*{Tr}_{\boldsymbol{b}^a} P(\boldsymbol{b}^a) \right] \exp\left[K \,\mathrm{Tr}\, G_{\Xi^\dagger \boldsymbol{R}_\mathrm{r} \Xi} \times \left(-\frac{1}{\sigma^2 K} \sqrt{\boldsymbol{R}_\mathrm{t}} \boldsymbol{L} \sqrt{\boldsymbol{R}_\mathrm{t}}^\dagger \right) \right]$$
$$= \int d\boldsymbol{Q} \exp\left[K \,\mathrm{Tr}\, G_{\Xi^\dagger \boldsymbol{R}_\mathrm{r} \Xi}\left(-\frac{1}{\sigma^2} \mathcal{P}\boldsymbol{Q} \right) + \ln \Pi^{(n)} \times (\boldsymbol{Q}) \right], \tag{A2}$$

where $\boldsymbol{I}_D$ is the D-dimensional identity matrix, $\boldsymbol{Q}$ is an $(n+1)$-dimensional matrix, and $\otimes$ represents the Kronecker or direct product. Note that the trace in Eq. (A2) is K dimensional in the second and third lines and is $(n+1)$ dimensional in the last line. In the equation above, the following functions are defined:

$$\Pi^{(n)}(\boldsymbol{Q}) \equiv \left\{ \prod_{a=0}^{n} \operatorname*{Tr}_{\boldsymbol{b}^a} P(\boldsymbol{b}^a) \right\} \left\{ \prod_{a=0}^{n} \delta(\boldsymbol{b}^{a\dagger} \boldsymbol{R}_\mathrm{t} \boldsymbol{b}^a - K Q_{aa}) \right\} \times \left\{ \prod_{a<b}^{n} \delta(\boldsymbol{b}^{a\dagger} \boldsymbol{R}_\mathrm{t} \boldsymbol{b}^b - K Q_{ab}) \right\}. \tag{A3}$$

Note that the function $G_{\Xi^\dagger \boldsymbol{R}_\mathrm{r} \Xi}$ is the same function obtained from the matrix integration formula over unitary matrix Haar measure $d\boldsymbol{U}$ [15–17],

$$\int d\boldsymbol{U} \exp(\mathrm{Tr}\, \Xi^\dagger \boldsymbol{R}_\mathrm{r} \Xi \boldsymbol{A}) \simeq \exp[K \,\mathrm{Tr}\, G_{\Xi^\dagger \boldsymbol{R}_\mathrm{r} \Xi}(\boldsymbol{A}/K)]. \tag{A4}$$

Here, unitary matrix $\boldsymbol{U}$ to be integrated is defined as a K-dimensional unitary matrix that diagonalizes the random matrix product as $\Xi^\dagger \boldsymbol{R}_\mathrm{r} \Xi = \boldsymbol{U}^\dagger \boldsymbol{D} \boldsymbol{U}$, where $\mathbf{D}$ is a diagonal matrix and $\boldsymbol{A}$ is an arbitrary K-dimensional matrix. The result, given by Eq. (A2), indicates that the entire set of unitary matrices that appear in the diagonalization of all possible random matrix products $\Xi^\dagger \boldsymbol{R}_\mathrm{r} \Xi$ coincides with the entire set of unitary matrices, which guarantees that the matrix integration formula over the unitary matrix is applicable only to the $\Xi^\dagger \boldsymbol{R}_\mathrm{r} \Xi$ part of the entire cross-correlation matrix $\boldsymbol{H}^\dagger \boldsymbol{H}$. This is because the multiplication of the matrix Ξ serves as a unitary transformation. Note that we cannot apply the same argument to the entire cross-correlation matrix $\boldsymbol{H}^\dagger \boldsymbol{H} = \sqrt{\boldsymbol{R}_\mathrm{t}}^\dagger \Xi^\dagger \boldsymbol{R}_\mathrm{r} \Xi \sqrt{\boldsymbol{R}_\mathrm{t}}$ because the transmitter correlation matrix $\boldsymbol{R}_\mathrm{t}$ breaks rotational invariance, and more careful treatment is required as described below.

By using the saddle-point method and assuming replica symmetry as $q = Q_{ab}$ for $a \neq b$ and $q + \chi = Q_{aa}$, we can evaluate the replicated partition function in Eq. (A2) after introducing auxiliary variables $\hat{q} + \hat{\chi}$ and $-2\hat{q}$ for the delta functions of the diagonal and off-diagonal matrix elements, respectively,

$$\frac{1}{K} \ln \Pi^{(n)}(\boldsymbol{Q}) = \operatorname*{Extr}_{\hat{q},\hat{\chi}} \left\{ n\chi\hat{\chi} + [\hat{\chi} - (n+1)\hat{q}][\chi + (n+1)q] + \frac{1}{K} \ln\left\{ \left[\prod_{a=0}^{n} \operatorname*{Tr}_{\boldsymbol{b}^a} P(\boldsymbol{b}^a) \right] \exp\left[-\hat{\chi} \sum_{a=0}^{n} \boldsymbol{b}^{a\dagger} \boldsymbol{R}_\mathrm{t} \boldsymbol{b}^a + \hat{q} \sum_{ab}^{n} \boldsymbol{b}^{a\dagger} \boldsymbol{R}_\mathrm{t} \boldsymbol{b}^b \right] \right\} \right\}. \tag{A5}$$

The saddle-point condition for q yields $\hat{q} = \hat{\chi}/(n+1)$. After performing Hubbard-Stratonovich transformation,

$$\exp\left[-\hat{\chi} \sum_{a=0}^{n} \boldsymbol{b}^{a\dagger} \boldsymbol{R}_\mathrm{t} \boldsymbol{b}^a + \frac{\hat{\chi}}{n+1} \sum_{ab}^{n} \boldsymbol{b}^{a\dagger} \boldsymbol{R}_\mathrm{t} \boldsymbol{b}^b \right] = \left(\frac{(n+1)\hat{\chi}}{\pi} \right)^K \int_{\mathbb{C}^K} d\boldsymbol{r}' \exp\left[-\hat{\chi} \sum_{a=0}^{n} |\boldsymbol{r}' - \sqrt{\boldsymbol{R}_\mathrm{t}} \boldsymbol{b}^a|^2 \right], \tag{A6}$$

we have

$$\frac{1}{K} \ln \Pi^{(n)}(\boldsymbol{Q}) = \operatorname*{Extr}_{\hat{\chi}} \left\{ n\chi\hat{\chi} + \frac{1}{K} \ln \int_{\mathbb{C}^K} d\boldsymbol{r}' \left\{ \operatorname*{Tr}_{\boldsymbol{b}} P(\boldsymbol{b}) \left(\frac{\hat{\chi}}{\pi} \right)^K \times \exp[-\hat{\chi}|\boldsymbol{r}' - \sqrt{\boldsymbol{R}_\mathrm{t}} \boldsymbol{b}|^2] \right\}^{n+1} + n \ln\left(\frac{\pi}{\hat{\chi}} \right) + \ln(n+1) \right\}. \tag{A7}$$

Combining this equation with the remainder of the replicated partition function and noting that the matrix $\mathcal{PQ}$ has a single zero eigenvalue and n-degenerate χ under the replica-symmetric condition, we arrive at the final expression of the mutual information,

$$\overline{\mathcal{I}_{\boldsymbol{H}}} = \underset{\chi,\hat{\chi}}{\mathrm{Extr}}\left\{ - G_{\boldsymbol{\Xi}^\dagger \boldsymbol{R}_r \boldsymbol{\Xi}}\left(-\frac{\chi}{\sigma^2}\right) - \frac{\partial}{\partial n}\frac{1}{K}\ln \Pi^{(n)}(\boldsymbol{Q})|_{n=0}\right\}$$
$$= \underset{\lambda}{\mathrm{Extr}}\left\{ \hat{G}_{\boldsymbol{\Xi}^\dagger \boldsymbol{R}_r \boldsymbol{\Xi}}(\lambda) + I_{\boldsymbol{R}_t}\left(\frac{\lambda}{\sigma^2}\right)\right\}. \tag{A8}$$

APPENDIX B: PERTURBATIVE EXPANSION OF MUTUAL INFORMATION

In this appendix, we derive the perturbative expressions of free energies from two evaluations, namely, matrix integration and exact analysis. For convenience, we introduce the constant χ into the channel definition [see also Eqs. (8) and (9)] as follows:

$$\boldsymbol{r} = \sqrt{\chi}\boldsymbol{H}\boldsymbol{b} + \boldsymbol{\eta}, \tag{B1}$$

where we set $\sigma=1$ in Eq. (1). Taking QPSK modulation into account, let us assume the probability distribution $P(b_k)$, which satisfies the following conditions:

(1) $P(b_k)$ can be factorized into the same distributions for the real and the imaginary parts: $P(b_k) = \tilde{P}[\mathrm{Re}(b_k)]\tilde{P}[\mathrm{Im}(b_k)]$.

(2) Reflection symmetry: $\tilde{P}(x)=\tilde{P}(-x)$.

(3) Signal power condition: $\Sigma_{b_k}\tilde{P}[\mathrm{Re}(b_k)]\mathrm{Re}(b_k)^2 = \Sigma_{b_k}\tilde{P}[\mathrm{Im}(b_k)]\mathrm{Im}(b_k)^2 = 1/2$.

From these conditions, we have $\Sigma_{b_k}P(b_k)b_k^{2l-1} = 0, \Sigma_{b_k}P(b_k)|b_k|^{2l}b_k = 0, \Sigma_{b_k}P(b_k)|b_k|^2 = 1\,(l\in\mathbb{N})$ and so on. Quadrature-phase shift keying modulation is included in this case. As mentioned in the text, we also assume that the cross-correlation matrix $\boldsymbol{H}^\dagger\boldsymbol{H}$ can be written as $\boldsymbol{H}^\dagger\boldsymbol{H}=\boldsymbol{I}+\rho\boldsymbol{R}$ with zero diagonal elements of $\boldsymbol{R}, R_{kk}=0$ and convergence of the eigenvalue distribution of the cross-correlation matrix for $K\to\infty$.

1. Expansion of mutual information via matrix integration

As shown in Sec. II, mutual information is obtained via matrix integration as follows:

$$\overline{I_{\boldsymbol{I}+\rho \boldsymbol{U}^\dagger \boldsymbol{R}\boldsymbol{U}}(\chi)} = \underset{\xi,\hat{\xi}}{\mathrm{Extr}}\{I_{\boldsymbol{I}}(\hat{\xi}\chi) + \xi\hat{\xi} - G_{\boldsymbol{I}+\rho\boldsymbol{R}}(\xi)\}. \tag{B2}$$

The function $G_{\boldsymbol{I}+\rho\boldsymbol{R}}(z)$ can be decomposed to obtain

$$G_{\boldsymbol{I}+\rho\boldsymbol{R}}(z) = z + G_{\boldsymbol{R}}(\rho z). \tag{B3}$$

Let us define $\overline{\lambda^n} \equiv \mathrm{Tr}(\boldsymbol{R}^n)/K$. The function $G_{\boldsymbol{R}}(z)$ can be expressed in terms of $\overline{\lambda^n}$ as follows:

$$G_{\boldsymbol{R}}(z) = \frac{1}{2}\overline{\lambda^2}z^2 + \frac{1}{3}\overline{\lambda^3}z^3 + \frac{1}{4}(\overline{\lambda^4} - 2\overline{\lambda^2}^2)z^4 + \cdots, \tag{B4}$$

from the formula $G(z)=\int_0^z dx[f(x)-x^{-1}]$ so that $x = \int d\lambda \rho(\lambda)[f(x)-\lambda]^{-1}$. Note that $\bar{\lambda}=0$ for $\boldsymbol{R}$ and $G(0)=0$. Substitution into mutual information after redefinition of $\hat{\xi}$ yields the following:

$$\overline{I_{\boldsymbol{I}+\rho \boldsymbol{U}^\dagger \boldsymbol{R}\boldsymbol{U}}(\chi)} = \underset{\xi,\hat{\xi}}{\mathrm{Extr}}\{I_{\boldsymbol{I}}(\chi+\hat{\xi}\chi) + \xi\hat{\xi} - G_{\boldsymbol{R}}(\rho\xi)\} = \underset{\xi,\hat{\xi}}{\mathrm{Extr}}\left\{ I_{\boldsymbol{I}}(\chi) + \hat{\xi}\chi I_{\boldsymbol{I}}'(\chi) + \frac{(\hat{\xi}\chi)^2}{2}I_{\boldsymbol{I}}''(\chi) + \cdots + \xi\hat{\xi} - \frac{\overline{\lambda^2}}{2}(\rho\xi)^2 - \frac{\overline{\lambda^3}}{3}(\rho\xi)^3 - \frac{\overline{\lambda^4} - 2\overline{\lambda^2}^2}{4}(\rho\xi)^4 + \cdots\right\}. \tag{B5}$$

From the saddle-point conditions with respect to $\xi,\hat{\xi}$, we have

$$\xi = -\frac{\partial}{\partial\hat{\xi}}I_{\boldsymbol{I}}(\chi+\hat{\xi}\chi) = -\chi I_{\boldsymbol{I}}'(\chi) + c\rho^2 + O(\rho^3),$$

$$\hat{\xi} = \rho G_{\boldsymbol{R}}'(\rho\xi) = \rho G_{\boldsymbol{R}}'[-\rho\chi I_{\boldsymbol{I}}'(\chi) + c\rho^3] = \rho G_{\boldsymbol{R}}'[-\rho\chi I_{\boldsymbol{I}}'(\chi)] + c\rho^4 G_{\boldsymbol{R}}''[-\rho\chi I_{\boldsymbol{I}}'(\chi)] + O(\rho^7) = -\rho^2\overline{\lambda^2}\chi I_{\boldsymbol{I}}'(\chi) + \rho^3\overline{\lambda^3}\chi^2[I_{\boldsymbol{I}}'(\chi)]^2 - \rho^4(\overline{\lambda^4} - 2\overline{\lambda^2}^2)\chi^3[I_{\boldsymbol{I}}'(\chi)]^3 + c\rho^4\overline{\lambda^2} + O(\rho^5) \tag{B6}$$

for ρ up to the fourth order [c is an $O(1)$ constant]. Substituting these equations into the original expression for the mutual information, we obtain

$$\overline{I_{\boldsymbol{I}+\rho \boldsymbol{U}^\dagger \boldsymbol{R}\boldsymbol{U}}(\chi)} = I_{\boldsymbol{I}}(\chi) - \frac{(\rho\chi)^2}{2}\overline{\lambda^2}\{I_{\boldsymbol{I}}'(\chi)\}^2 + \frac{(\rho\chi)^3}{3}\overline{\lambda^3}\{I_{\boldsymbol{I}}'(\chi)\}^3 - \frac{(\rho\chi)^4}{4}(\overline{\lambda^4} - 2\overline{\lambda^2}^2)\{I_{\boldsymbol{I}}'(\chi)\}^4 + \frac{(\rho\chi)^4}{2}\overline{\lambda^2}^2\{I_{\boldsymbol{I}}''(\chi)\} \times\{I_{\boldsymbol{I}}'(\chi)\}^2 + O(\rho^5). \tag{B7}$$

2. Expansion of rigorous mutual information

Before studying perturbative expansion of the complex channel, let us start with the real channel for which mutual communication is given by

$$I_{\tilde{\boldsymbol{R}}_t}^{\mathrm{real}}(\chi) = -\frac{1}{K}\int_{\mathbb{R}^K} d\tilde{\boldsymbol{r}}\left\{\underset{\tilde{\boldsymbol{b}}}{\mathrm{Tr}}\tilde{P}(\tilde{\boldsymbol{b}})\left(\frac{\chi}{2\pi}\right)^{K/2}\exp\left(-\frac{\chi}{2}|\tilde{\boldsymbol{r}} - \sqrt{\tilde{\boldsymbol{R}}_t}\tilde{\boldsymbol{b}}|^2\right)\right\}\left\{\ln\underset{\tilde{\boldsymbol{b}}}{\mathrm{Tr}}\tilde{P}(\tilde{\boldsymbol{b}})\left(\frac{\chi}{2\pi}\right)^{K/2}\exp\left(-\frac{\chi}{2}|\tilde{\boldsymbol{r}} - \sqrt{\tilde{\boldsymbol{R}}_t}\tilde{\boldsymbol{b}}|^2\right)\right\} - \frac{1}{2}\ln\left(\frac{2\pi}{\chi}\right) - \frac{1}{2}, \tag{B8}$$

where all variables and matrices are real and denoted with tilde for discrimination between the real and the complex channels in this subsection. By substituting $\tilde{\boldsymbol{R}}_{\mathrm{t}}=\boldsymbol{I}+\rho\tilde{\boldsymbol{R}}$ into the above equation, where $\tilde{\boldsymbol{R}}$ is a symmetric matrix with zero diagonal elements, expanding with respect to ρ, and then performing some algebraic manipulation, we obtain the following:

$$
\begin{aligned}
I^{\mathrm{real}}_{\boldsymbol{I}+\rho\tilde{\boldsymbol{R}}}(\chi) &= I^{\mathrm{real}}_{\boldsymbol{I}}(\chi) + \sum_{k=1}^{4}\frac{\rho^k}{k!}\partial^k_\rho I^{\mathrm{real}}_{\boldsymbol{I}+\rho\tilde{\boldsymbol{R}}}(\chi)\Big|_{\rho=0} + O(\rho^5) = I^{\mathrm{real}}_{\boldsymbol{I}}(\chi) - \frac{(\rho\chi)^2}{2!}\frac{1}{2K}\sum_{i_1 i_2 j_1 j_2}\tilde{R}_{i_1 j_1}\tilde{R}_{i_2 j_2}\\
&\quad\times[\langle\tilde{b}_{i_1}\tilde{b}_{i_2}\rangle_{\mathrm{c}}\langle\tilde{b}_{j_1}\tilde{b}_{j_2}\rangle_{\mathrm{c}}] - \frac{(\rho\chi)^3}{3!}\frac{1}{2K}\sum_{i_1 i_2 i_3 j_1 j_2 j_3}\tilde{R}_{i_1 j_1}\tilde{R}_{i_2 j_2}\tilde{R}_{i_3 j_3}([\langle\tilde{b}_{i_1}\tilde{b}_{i_2}\tilde{b}_{i_3}\rangle_{\mathrm{c}}\langle\tilde{b}_{j_1}\tilde{b}_{j_2}\tilde{b}_{j_3}\rangle_{\mathrm{c}}] - 2[\langle\tilde{b}_{i_1}\tilde{b}_{j_2}\rangle_{\mathrm{c}}\langle\tilde{b}_{i_2}\tilde{b}_{j_3}\rangle_{\mathrm{c}}\langle\tilde{b}_{i_3}\tilde{b}_{j_1}\rangle_{\mathrm{c}}])\\
&\quad - \frac{(\rho\chi)^4}{4!}\frac{1}{2K}\sum_{i_1 i_2 i_3 i_4 j_1 j_2 j_3 j_4}\tilde{R}_{i_1 j_1}\tilde{R}_{i_2 j_2}\tilde{R}_{i_3 j_3}\tilde{R}_{i_4 j_4}\times([\langle\tilde{b}_{i_1}\tilde{b}_{i_2}\tilde{b}_{i_3}\tilde{b}_{i_4}\rangle_{\mathrm{c}}\langle\tilde{b}_{j_1}\tilde{b}_{j_2}\tilde{b}_{j_3}\tilde{b}_{j_4}\rangle_{\mathrm{c}}] - 12[\langle\tilde{b}_{i_1}\tilde{b}_{i_2}\tilde{b}_{i_3}\rangle_{\mathrm{c}}\langle\tilde{b}_{j_1}\tilde{b}_{j_2}\tilde{b}_{j_4}\rangle_{\mathrm{c}}\langle\tilde{b}_{i_3}\tilde{b}_{j_4}\rangle_{\mathrm{c}}]\\
&\quad + 6[\langle\tilde{b}_{i_1}\tilde{b}_{j_2}\rangle_{\mathrm{c}}\langle\tilde{b}_{i_2}\tilde{b}_{j_3}\rangle_{\mathrm{c}}\langle\tilde{b}_{i_3}\tilde{b}_{j_4}\rangle_{\mathrm{c}}\langle\tilde{b}_{i_4}\tilde{b}_{j_1}\rangle_{\mathrm{c}}]) + O(\rho^5) = I^{\mathrm{real}}_{\boldsymbol{I}}(\chi) - \frac{(\rho\chi)^2}{4}\frac{\mathrm{Tr}\tilde{\boldsymbol{R}}^2}{K}[\langle\tilde{b}^2\rangle_{\mathrm{c}}]^2 + \frac{(\rho\chi)^3}{6}\frac{\mathrm{Tr}\tilde{\boldsymbol{R}}^3}{K}[\langle\tilde{b}^2\rangle_{\mathrm{c}}]^3\\
&\quad - \frac{(\rho\chi)^4}{48}\frac{\sum_{ij}\tilde{R}^4_{ij}}{K}[\langle\tilde{b}^4\rangle_{\mathrm{c}}]^2 - \frac{(\rho\chi)^4}{8}\Bigg\{\left(\frac{\mathrm{Tr}\tilde{\boldsymbol{R}}^4}{K} - \frac{2\sum_i(\tilde{\boldsymbol{R}}^2)^2_{ii}}{K} + \frac{\sum_{ij}\tilde{R}^4_{ij}}{K}\right)[\langle\tilde{b}^2\rangle_{\mathrm{c}}]^4 - 2\left(\frac{\sum_i(\tilde{\boldsymbol{R}}^2)^2_{ii}}{K} - \frac{\sum_{ij}\tilde{R}^4_{ij}}{K}\right)[\langle\tilde{b}^2\rangle^2_{\mathrm{c}}][\langle\tilde{b}^2\rangle_{\mathrm{c}}]^2\\
&\quad + \frac{\sum_{ij}\tilde{R}^4_{ij}}{K}[\langle\tilde{b}^2\rangle^2_{\mathrm{c}}]^2\Bigg\} + O(\rho^5) = I^{\mathrm{real}}_{\boldsymbol{I}}(\chi) - \frac{(\rho\chi)^2}{4}\overline{\lambda^2}[\langle\tilde{b}^2\rangle_{\mathrm{c}}]^2 + \frac{(\rho\chi)^3}{6}\overline{\lambda^3}[\langle\tilde{b}^2\rangle_{\mathrm{c}}]^3 - \frac{(\rho\chi)^4}{8}\overline{\lambda^4}[\langle\tilde{b}^2\rangle_{\mathrm{c}}]^4\\
&\quad - \frac{(\rho\chi)^4}{8}\Bigg\{\frac{2\sum_i(\tilde{\boldsymbol{R}}^2)^2_{ii}}{K}([\langle\tilde{b}^2\rangle^2_{\mathrm{c}}] - [\langle\tilde{b}^2\rangle_{\mathrm{c}}]^2)[\langle\tilde{b}^2\rangle_{\mathrm{c}}]^2 + \frac{\sum_{ij}\tilde{R}^4_{ij}}{K}\left(([\langle\tilde{b}^2\rangle^2_{\mathrm{c}}] - [\langle\tilde{b}^2\rangle_{\mathrm{c}}]^2)^2 + \frac{[\langle\tilde{b}^4\rangle_{\mathrm{c}}]^2}{6}\right)\Bigg\} + O(\rho^5).
\end{aligned}
\tag{B9}
$$

Here, $\langle f(\tilde{\boldsymbol{b}})\rangle \equiv \mathrm{Tr}_{\tilde{\boldsymbol{b}}}\tilde{P}(\tilde{\boldsymbol{b}})f(\tilde{\boldsymbol{b}})\exp(-\chi|\tilde{\boldsymbol{r}}-\tilde{\boldsymbol{b}}|^2/2)/\mathrm{Tr}_{\tilde{\boldsymbol{b}}}\tilde{P}(\tilde{\boldsymbol{b}})\times\exp(-\chi|\tilde{\boldsymbol{r}}-\tilde{\boldsymbol{b}}|^2/2)$ and $\langle f(\tilde{\boldsymbol{b}})\rangle_{\mathrm{c}}$ is its cumulant. In addition, $\langle f(\tilde{\boldsymbol{b}})\rangle$ is a function of $\tilde{\boldsymbol{r}}$ and is always accompanied by the average $[F(\tilde{\boldsymbol{r}})] \equiv (\chi/2\pi)^{K/2}\int d\tilde{\boldsymbol{r}}\mathrm{Tr}_{\tilde{\boldsymbol{b}}}\tilde{P}(\tilde{\boldsymbol{b}})F(\tilde{\boldsymbol{r}})\exp(-\chi|\tilde{\boldsymbol{r}}-\tilde{\boldsymbol{b}}|^2/2)$, where b without a subscript denoted the signal at an arbitrary antenna. The expression in the third line of Eq. (B9) is obtained by repeatedly performing integration by parts with respect to $\tilde{\boldsymbol{r}}$. We can also show that $[\langle\tilde{b}^2\rangle_{\mathrm{c}}]=2I'^{\mathrm{real}}_{\boldsymbol{I}}(\chi)$ and $[\langle\tilde{b}^2\rangle^2_{\mathrm{c}}]=-2I''^{\mathrm{real}}_{\boldsymbol{I}}(\chi)$, from which we have

$$
\begin{aligned}
2I^{\mathrm{real}}_{\boldsymbol{I}+\rho\tilde{\boldsymbol{R}}}(\chi) &= 2I^{\mathrm{real}}_{\boldsymbol{I}}(\chi) - \frac{(\rho\chi)^2}{2}\overline{\lambda^2}[2I'^{\mathrm{real}}_{\boldsymbol{I}}(\chi)]^2\\
&\quad + \frac{(\rho\chi)^3}{3}\overline{\lambda^3}[2I'^{\mathrm{real}}_{\boldsymbol{I}}(\chi)]^3 - \frac{(\rho\chi)^4}{4}\overline{\lambda^4}[2I'^{\mathrm{real}}_{\boldsymbol{I}}(\chi)]^4\\
&\quad - \frac{(\rho\chi)^4}{4}\Bigg\{\frac{2\sum_i(\tilde{\boldsymbol{R}}^2)^2_{ii}}{K}\{-2I''^{\mathrm{real}}_{\boldsymbol{I}}(\chi) - [2I'^{\mathrm{real}}_{\boldsymbol{I}}(\chi)]^2\}\\
&\quad \times[2I'^{\mathrm{real}}_{\boldsymbol{I}}(\chi)]^2 + \frac{\sum_{ij}\tilde{R}^4_{ij}}{K}\Bigg(\{-2I''^{\mathrm{real}}_{\boldsymbol{I}}(\chi)\\
&\quad - [2I'^{\mathrm{real}}_{\boldsymbol{I}}(\chi)]^2\}^2 + \frac{[\langle\tilde{b}^4\rangle_{\mathrm{c}}]^2}{6}\Bigg)\Bigg\} + O(\rho^5).
\end{aligned}
\tag{B10}
$$

Next, we convert this result into the result for a complex channel, the mutual information of which is given by Eq. (6). We can easily show that the complex channel described by $\boldsymbol{r}=\sqrt{\chi}\sqrt{\boldsymbol{R}_{\mathrm{t}}}\boldsymbol{b}+\boldsymbol{z}$ is equivalent to the real channel of double size $\tilde{\boldsymbol{r}}=\sqrt{\chi}\sqrt{\tilde{\boldsymbol{R}}_{\mathrm{t}}}\tilde{\boldsymbol{b}}+\tilde{\boldsymbol{z}}$, where

$$
\tilde{\boldsymbol{z}} = \sqrt{2}\begin{pmatrix}\mathrm{Re}(\boldsymbol{z})\\ \mathrm{Im}(\boldsymbol{z})\end{pmatrix},\quad \tilde{\boldsymbol{r}} = \sqrt{2}\begin{pmatrix}\mathrm{Re}(\boldsymbol{r})\\ \mathrm{Im}(\boldsymbol{r})\end{pmatrix},\quad \tilde{\boldsymbol{b}} = \sqrt{2}\begin{pmatrix}\mathrm{Re}(\boldsymbol{b})\\ \mathrm{Im}(\boldsymbol{b})\end{pmatrix},
$$

$$
\begin{aligned}
\sqrt{\tilde{\boldsymbol{R}}_{\mathrm{t}}} &= \begin{bmatrix}\mathrm{Re}(\sqrt{\boldsymbol{R}_{\mathrm{t}}}) & -\mathrm{Im}(\sqrt{\boldsymbol{R}_{\mathrm{t}}})\\ \mathrm{Im}(\sqrt{\boldsymbol{R}_{\mathrm{t}}}) & \mathrm{Re}(\sqrt{\boldsymbol{R}_{\mathrm{t}}})\end{bmatrix},\quad P(\boldsymbol{b})\\
&= 2\tilde{P}[\sqrt{2}\,\mathrm{Re}(\boldsymbol{b})]\tilde{P}[\sqrt{2}\,\mathrm{Im}(\boldsymbol{b})].
\end{aligned}
\tag{B11}
$$

For such a system, we can show that $I_{\boldsymbol{I}}(\chi)=2I^{\mathrm{real}}_{\boldsymbol{I}}(\chi)$ and $I_{\boldsymbol{R}_{\mathrm{t}}(\chi)}=2I^{\mathrm{real}}_{\tilde{\boldsymbol{R}}_{\mathrm{t}}}(\chi)$. Since the eigenvalue distributions of $\boldsymbol{R}_{\mathrm{t}}=\boldsymbol{I}+\rho\boldsymbol{R}$ and the corresponding $\tilde{\boldsymbol{R}}_{\mathrm{t}}$ are the same, from the relationship between the real and complex channels, we have

$$
\begin{aligned}
I_{I+\rho R}(\chi) &= 2I^{\mathrm{real}}_{I+\rho\tilde{R}}(\chi) = I_I(\chi) - \frac{(\rho\chi)^2}{2}\overline{\lambda^2}[I_I'(\chi)]^2 \\
&+ \frac{(\rho\chi)^3}{3}\overline{\lambda^3}[I_I'(\chi)]^3 - \frac{(\rho\chi)^4}{4}\overline{\lambda^4}[I_I'(\chi)]^4 \\
&- \frac{(\rho\chi)^4}{4}\left\{ \frac{2\sum_i (R^2)_{ii}^2}{K}[-I_I''(\chi) - I_I'(\chi)^2][I_I'(\chi)]^2 \right. \\
&+ \frac{\sum_{ij}[\mathrm{Re}(R_{ij})^4 + \mathrm{Im}(R_{ij})^4]}{K}\left\{[-I_I''(\chi) - I_I'(\chi)^2]^2 \right. \\
&\left.\left. + \frac{C(\chi)^2}{6}\right\}\right\} + O(\rho^5). \qquad \text{(B12)}
\end{aligned}
$$

The function $C(\chi)$ is given by

$$
\begin{aligned}
C(\chi) \equiv 2(\chi/\pi)^K \int d\boldsymbol{r}\,\underset{\boldsymbol{b}}{\mathrm{Tr}}P(\boldsymbol{b})\langle \mathrm{Re}(b)^4 + \mathrm{Im}(b)^4\rangle_{\mathrm{c}}^{\mathrm{cmp}} \\
\times \exp(-\chi|\boldsymbol{r}-\boldsymbol{b}|^2),
\end{aligned}
$$

where $\langle f(\boldsymbol{b})\rangle^{\mathrm{cmp}} \equiv \underset{\boldsymbol{b}}{\mathrm{Tr}}P(\boldsymbol{b})f(\boldsymbol{b})\exp(-\chi|\boldsymbol{r}-\boldsymbol{b}|^2)/\underset{\boldsymbol{b}}{\mathrm{Tr}}P(\boldsymbol{b})$

$$
\times \exp(-\chi|\boldsymbol{r}-\boldsymbol{b}|^2), \qquad \text{(B13)}
$$

and the subscript of the angular bracket c denotes the cumulant. Substituting the definitions of $\overline{\lambda^2}$ and $\overline{\lambda^4}$, the discrepancy between the two results is obtained as

$$
\begin{aligned}
I_{I+\rho R}(\chi) - \overline{I_{I+\rho U^\dagger R U}(\chi)} &= -\frac{(\rho\chi)^4}{2}\frac{1}{K}\sum_i \left[(R^2)_{ii} - \frac{\mathrm{Tr}(R^2)}{K}\right]^2 \\
&\times [-I_I''(\chi) - I_I'(\chi)^2][I_I'(\chi)]^2 \\
&- \frac{(\rho\chi)^4}{4}\frac{1}{K}\left\{\sum_{ij}[\mathrm{Re}(R_{ij})^4 + \mathrm{Im}(R_{ij})^4]\right\} \\
&\times \left\{[-I_I''(\chi) - I_I'(\chi)^2]^2 + \frac{C(\chi)^2}{6}\right\} \\
&+ O(\rho^5), \qquad \text{(B14)}
\end{aligned}
$$

that is, the dominant term of the discrepancy is of the order ρ^4. The factor $-I_I''(\chi) - \{I_I'(\chi)\}^2 = -2I_I''^{\mathrm{real}}(\chi) - \{2I_I'^{\mathrm{real}}(\chi)\}^2 = [\langle\tilde{b}^2\rangle_{\mathrm{c}}^2] - [\langle\tilde{b}^2\rangle_{\mathrm{c}}]^2$ is non-negative, and the inequality $I_{I+\rho U^\dagger R U}(\chi) \geq I_{I+\rho R}(\chi)$ holds up to the fourth order.

[1] H. Nishimori, *Statistical Physics of Spin Glasses and Information Processing: An Introduction* (Oxford University Press, Oxford, 2001).
[2] T. Tanaka, Europhys. Lett. **54**, 540 (2001).
[3] T. Tanaka, IEEE Trans. Inf. Theory **48**, 2888 (2002).
[4] Y. Kabashima, J. Phys. A **36**, 11111 (2003).
[5] A. L. Moustakas, S. H. Simon, and A. M. Sengupta, IEEE Trans. Inf. Theory **49**, 2545 (2003).
[6] R. R. Müller, IEEE Trans. Signal Process. **51**, 2821 (2003).
[7] D. Guo and S. Verdú, IEEE Trans. Inf. Theory **51**, 1983 (2005).
[8] C.-K. Wen, Y.-N. Lee, J.-T. Chen, and P. Ting, IEEE Trans. Signal Process. **53**, 2059 (2005).
[9] J. P. Neirotti and D. Saad, Europhys. Lett. **71**, 866 (2005).
[10] C.-K. Wen, P. Ting, and J.-T. Chen, IEEE Trans. Commun. **54**, 349 (2006).
[11] H. Efraim, N. Yacov, O. Shental, and I. Kanter, J. Stat. Mech.: Theory Exp. (2009) P07039.
[12] K. Takeda, S. Uda, and Y. Kabashima, Europhys. Lett. **76**, 1193 (2006).
[13] K. Takeda, A. Hatabu, and Y. Kabashima, J. Phys. A: Math. Theor. **40**, 14085 (2007).
[14] R. R. Müller, D. Guo, and A. L. Moustakas, IEEE J. Sel. Areas Commun. **26**, 530 (2008).
[15] C. Itzykson and J. B. Zuber, J. Math. Phys. **21**, 411 (1980).
[16] E. Marinari, G. Parisi, and F. Ritort, J. Phys. A **27**, 7647 (1994).
[17] G. Parisi and M. Potters, J. Phys. A **28**, 5267 (1995).
[18] D. Voiculescu, K. J. Dykema, and A. Nica, *Free Random Variables*, CRM Monograph Series, Vol. 1 (American Mathematical Society, Providence, 1992).
[19] A. M. Tulino and S. Verdú, *Random Matrix Theory and Wireless Communications* (Now Publishers, Hanover, 2004).
[20] D.-S. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, IEEE Trans. Commun. **48**, 502 (2000).
[21] D. J. Thouless, P. W. Anderson, and R. G. Palmer, Philos. Mag. **35**, 593 (1977).
[22] M. Opper and O. Winther, Phys. Rev. Lett. **86**, 3695 (2001).
[23] M. Opper and O. Winther, Phys. Rev. E **64**, 056131 (2001).
[24] See Appendix B for the difference between the mutual informations of the real and the complex channels (factor 2 in particular). Note also that the functions G in matrix integration formulas for orthogonal matrix and unitary matrix differ by overall factor 2 and factor 1/2 in the argument [16].
[25] Strictly speaking this decomposition is not true due to the boundary. For simplicity we do not consider the boundary effect here. If we take the boundary into account, the constants l_0 and l_1 in the matrix must have the index of column (row). For numerical analysis we actually performed Cholesky decomposition numerically for finite size system without solving l_0 and l_1 using these relations.
[26] B. Derrida and H. J. Hilhorst, J. Phys. A **16**, 2641 (1983).
[27] M. Weigt and R. Monasson, Europhys. Lett. **36**, 209 (1996).