



Article Saddle Point Approximation of Mutual Information for Finite-Alphabet Inputs over Doubly Correlated MIMO Rayleigh Fading Channels

Yuyu Liu¹, Jinbao Zhang^{1,2} and Dan Zhang^{1,3,*}

- ¹ Institute of Electromagnetic Compatibility, Beijing Jiaotong University, Beijing 100044, China; 19120084@bjtu.edu.cn (Y.L.); jbzhang@bjtu.edu.cn (J.Z.)
- ² Frontiers Science Center for Smart High-speed Railway System, Beijing 100044, China
- Beijing Engineering Research Center of EMC and GNSS Technology for Rail Transportation, Beijing 100044, China
- Correspondence: zhang.dan@bjtu.edu.cn

Abstract: Given the mutual information of finite-alphabet inputs cannot be calculated concisely and accurately over fading channels, this paper proposes a new method to calculate the mutual information. First, the applicability of the saddle point method is studied, and then the mutual information is estimated by the saddle point approximation method with known channel state information. Furthermore, we induce the expectation of mutual information over doubly correlated multiple-input multiple-output (MIMO) Rayleigh fading channels. The validity of the saddle point approximation method is verified by comparing the numerical results of the Monte Carlo method and the saddle point approximation method under different doubly correlated MIMO fading channel scenarios.

Keywords: mutual information; finite-alphabet inputs; doubly correlated MIMO Rayleigh fading channels; saddle point approximation

1. Introduction

Mutual information plays an irreplaceable role in the theoretical analysis of communication system performance, including analysis, evaluation and optimization of transceiver structure [1], encoding and decoding schemes [2], and communication system bit error rate (BER) performance [3], etc., so it attracts increasing research interest. Channel capacity, defined as the upper bound of mutual information, is realized under Gaussian inputs over additive white Gaussian noise (AWGN) channels [4]. A large number of theoretical analyses and research are also based on this concept [5–7]. By means of Minkowski's inequality [1], two lower bounds of system capacity are obtained and are used as selection indexes to discuss the selection of antenna subsets in spatial multiplexing systems. Under the condition of Gaussian inputs, the hybrid encoder and combiners are designed by maximizing the achievable SE [2].

However, Gaussian inputs are rarely realized in practice because the unbounded amplitude of Gaussian distribution may lead to infinite transmitting power, and the continuity of Gaussian distribution will make it difficult to detect and decode the signal at the receiver. In practical communication systems, inputs are usually taken from finite-alphabet constellation sets with average distribution, rather than Gaussian inputs [8]. Considerable gaps in terms of transmitting performance exist [9–11] due to the differences between Gaussian and finite-alphabet inputs and then lead to deviations from optimal strategies. For example, it is believed that the traditionally optimal strategy to achieve capacity for Gaussian inputs is to allocate higher power to the sub-channels with a larger signal-to-noise ratio (SNR). In [10], it is demonstrated that such strategy may be quite suboptimal for the reason that the mutual information with finite-alphabet inputs is upper bounded, and there is little incentive to allocate more power to the sub-channels already close to saturation.



Citation: Liu, Y.; Zhang, J.; Zhang, D. Saddle Point Approximation of Mutual Information for Finite-Alphabet Inputs over Doubly Correlated MIMO Rayleigh Fading Channels. *Appl. Sci.* **2021**, *11*, 4700. https://doi.org/10.3390/app11104700

Academic Editors: Ireneusz Kubiak, Tadeusz Wieckowski and Yevhen Yashchyshyn

Received: 28 April 2021 Accepted: 19 May 2021 Published: 20 May 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). At the same time, channel capacity reflects the upper bound of communication system performance, so the performance of digital communication systems can be accurately evaluated by mutual information under the condition of finite alphabet sets inputs and actual transmission environment [12,13].

Due to the great complexity of the direct calculation of mutual information, it is almost impossible to obtain a closed-form solution. Monte Carlo trials are usually used for direct and accurate calculation [14]. In order to reduce the computational complexity, a bit-level algorithm using PDF of a log-likelihood ratio (LLR) to calculate mutual information is proposed [15]. However, each modulation mode of the algorithm requires a lot of prior simulations, and it is only suitable for specific scenarios without universality. To optimize linear precoding with finite-alphabet inputs, the authors in [16,17] deduced the closedform lower and upper bounds of mutual information as alternatives, which reduce the computational effort by several orders of magnitude compared to calculating the average mutual information directly. Then the bounds of mutual information are applied to multiple antennas, secure cognitive radio networks [18], and relay networks [19]. The study in [20] utilized the cutoff rate (CR) as the alternative of mutual information (MI) to design the linear precoders. Mutual information was also used to develop a two-step algorithm to enhance the achievable secrecy rate of cooperative jamming for secure communication with finitealphabet inputs in [11]. However, the gaps between approximation and accurate mutual information are still ambiguous, which limits the range where mutual information can be applied. Recently, the authors of [21] approximated ergodic mutual information based on multi-exponential decay curve fitting under M-ary quadrature amplitude modulation (M-QAM) signaling, but other modulation modes were neglected.

This study takes a step toward evaluating accurate mutual information for finitealphabet-based transmissions over doubly correlated MIMO fading channels. After discussing the applicability of the saddle point method, we obtain the approximate solution of mutual information for any known CSI and modulation mode by using this method, which is universal. On this basis, the mutual information expectation of doubly-correlated MIMO Rayleigh fading channels is further derived. This proposition highlights the considerable accuracy with radically reduced complexity.

The outline of this paper is as follows. The second section introduces the MIMO transmission model and its preliminary research. In the third section, the saddle point approximation method is used to estimate the mutual information, and then we calculate the mutual information expectation over doubly correlated MIMO Rayleigh fading channels. Then the validity and accuracy of the proposed method are verified under different doubly correlated MIMO fading channels scenarios in the fourth section. Finally, the fifth section gives conclusions.

2. Problem Formulation

2.1. Model of MIMO Transmission

Consider MIMO system with N_T transmitting antennas and N_R receiving antennas. Let $\tilde{\tilde{x}} \in \mathbb{C}^{N_T \times 1}$ ($\mathbb{C}^{N \times m}$ denotes the $N \times m$ complex spaces) be a transmitting signal vector, satisfying $\mathbb{E}_{\tilde{x}} \{\tilde{\tilde{x}}\} = 0$ and $\mathbb{E}_{\tilde{x}} \{\tilde{\tilde{x}}\tilde{\tilde{w}}^H\} = I$, where $\mathbb{E}_{(\bullet)}\{*\}$ stands for the statistical expectation of random * with respect to its variable \cdot ; I and 0 denote an unit and zero matrix of appropriate dimensions, respectively. $(\bullet)^H$ is the conjugate transpose of matrix \cdot . MIMO transmission is generally modeled by

$$\widetilde{\widetilde{y}} = \widetilde{H}\widetilde{\widetilde{x}} + \widetilde{\widetilde{w}}$$
(1)

where preliminaries are made as below.

1. $\widetilde{\tilde{H}} \in \mathbb{C}^{N_{R} \times N_{T}}$ is a complex fading channel matrix between transmitting antenna and receiving antenna arrays. The doubly correlated MIMO Rayleigh fading channel is modeled by $\Psi_{R}^{1/2} \widetilde{\tilde{H}}_{WG} \Psi_{T}^{1/2} \in \mathbb{C}^{N_{R} \times N_{T}}$ [1], where $\widetilde{\tilde{H}}_{WG} \in \mathbb{C}^{N_{R} \times N_{T}}$ is a matrix

consisted of independent and identically distributed CN(0,1) complex Gaussian entries; Ψ_T and Ψ_R are transmitting and receiving correlation matrices, respectively. Ψ_T and Ψ_R can be expressed as

$$\Psi_{\rm T} = \boldsymbol{U}_{\rm T} \boldsymbol{\Sigma}_{\rm T} \boldsymbol{U}_{\rm T}^{\rm H} \text{ and } \Psi_{\rm R} = \boldsymbol{U}_{\rm R} \boldsymbol{\Sigma}_{\rm R} \boldsymbol{U}_{\rm R}^{\rm H}$$
⁽²⁾

where U_T and U_R are unitary matrices whose columns are eigenvectors of Ψ_T and Ψ_R ; Σ_T and Σ_R represent diagonal matrices whose diagonal entries are the eigenvalues of Ψ_T and Ψ_R , respectively.

2. $\widetilde{\widetilde{w}} \in \mathbb{C}^{N_R \times 1}$ stands for AWGN corresponding to N_R receiving antennas, where each element is independent and identically complex Gaussian distributed, satisfying $\mathbb{E}_{\widetilde{w}} \left\{ \widetilde{\widetilde{w}} \right\} = 0$ and $\mathbb{E}_{\widetilde{w}} \left\{ \widetilde{\widetilde{w}} \widetilde{\widetilde{w}}^H \right\} = \sigma^2 I$.

2.2. Mutual Information for Finite-Alphabet Inputs

When a linear unitary transform U_R^H is applied on the receiving signal $\tilde{\tilde{y}}$, the MIMO model in (1) is equivalent to a model with channel matrix $U_R^H \tilde{\tilde{H}}$ and noise $U_R^H \tilde{\tilde{w}}$ [16], which is written as

$$\mathbf{y} = \boldsymbol{U}_{\mathrm{R}}^{\mathrm{H}\widetilde{\mathbf{y}}} = \boldsymbol{\Sigma}_{\mathrm{R}}^{1/2} \boldsymbol{H}_{\mathrm{WG}} \boldsymbol{\Sigma}_{\mathrm{T}}^{1/2} \boldsymbol{U}_{\mathrm{T}}^{\mathrm{H}\widetilde{\mathbf{x}}} + \boldsymbol{w}$$
(3)

where $H_{WG} = U_R^H \tilde{H}_{WG} U_T$ and $w = U_R^H \tilde{\tilde{w}}$. $\tilde{\tilde{x}}$ is selected equiprobably from N_T -dimension constellation consisted of $N = \prod_{I=1}^{N_T} N_I$ vectors, where N_I denotes the number of symbols in the *i*-th discrete constellation Ω_I . The mutual Information for Finite-Alphabet inputs between x and y is given by [22]

$$\mathcal{I}(\tilde{\tilde{x}}; \tilde{\tilde{y}}) = \mathcal{I}(\boldsymbol{U}_{T}^{H} \tilde{\tilde{x}}; \mathbf{y})$$

$$= \log_{2} N + \frac{1}{N} \sum_{m=1}^{N} \mathbb{E}_{H_{WG}, w} \left\{ \log_{2} \left[\exp\left(-\frac{||w||^{2}}{\sigma^{2}}\right) / \sum_{k=1}^{N} \exp\left(-\frac{||c_{m,k} + w||^{2}}{\sigma^{2}}\right) \right] \right\}$$
(4)

where $c_{m,k} = \Sigma_R^{1/2} H_{WG} \Sigma_T^{1/2} U_T^H d_{m,k}$ and $d_{m,k} = \mathbf{q}_m - \mathbf{q}_k \cdot \mathbf{q}_m$ and \mathbf{q}_k are the *m*-th and *k*-th points in the constellation of $\widetilde{\widetilde{x}}$, and $||\bullet||$ stands for the Euclidean norm of the variable \bullet .

Since the statistical Channel State Information (CSI) is varying much slower than instantaneous $\tilde{\tilde{H}}$, and can be obtained by channel estimation, this work assumed the statistical CSI was a perfectly-known constant. Consequently, the problem was to calculate the average mutual information with given statistical CSI.

3. Saddle Point Approximation for Mutual Information

Mutual information by (4) needs multiple integrals to compute expectation over H_{WG} and w. As N increases, it leads to prohibitive complexity and becomes the most significant obstacle in achieving accurate mutual information. Therefore, we used the idea of the mean value theorem of integrals to simplify multiple integrals by finding an appropriate point. In this section, we explore the saddle point approximation method and highlight the convenient calculation with a weighted mean over constellation set of x, instead of expectation over all possible samples of random H_{WG} and w.

3.1. Saddle Point Approximation

We first considered the expectation over the AWGN vector w. The Taylor series of (4) is expanded to

$$\mathcal{I}(\mathbf{x}; \mathbf{y} | \mathbf{H}) = \log_2 N - \frac{1}{N \ln 2} \sum_{m=1}^N \sum_{\sigma=1}^{+\infty} \frac{1}{\sigma} \sum_{q=0}^\sigma \frac{\sigma! (-1)^q}{q! (\sigma-q)!} \frac{1}{\pi^{N_R} \sigma^{2N_R}} \int_{\mathbf{w}} \left(\sum_{k=1}^N \sigma_{m,k}(\mathbf{w}) \right)^{-q} \mathrm{d}\mathbf{w}$$
(5)

where

$$p_{m,k}(w) = \exp\left(\frac{||w - qc_{m,k}||^2}{q\sigma^2} - \frac{(q+1)||c_{m,k}||^2}{\sigma^2}\right) \text{ and } H = \Sigma_{\mathrm{R}}^{1/2} H_{\mathrm{WG}} \Sigma_{\mathrm{T}}^{1/2}$$
(6)

Before proceeding to the saddle point approximation, we needed to establish the following lemma, which guarantees the existence of the saddle point.

Lemma 1. For non-zero natural number q, the maximum of $\left(\sum_{k=1}^{N} \sigma_{m,k}(w)\right)^{-q}$ exists and is achieved at $w = w_0$, where w_0 is the weighted average vector of $c_{m,1}, c_{m,2}, \ldots$, and $c_{m,N}$.

$$\boldsymbol{w}_{0} = \sum_{k=1}^{N} q \rho_{m,k} \boldsymbol{c}_{m,k} \triangleq q \, \boldsymbol{c}_{m} \tag{7}$$

where $\rho_{m,k} = \sigma_{m,k}(w_0) / \sum_{k=1}^{N} \sigma_{m,k}(w_0) = \sigma_{m,k}(w_0) / \sigma_m(w_0)$ is a positive real number over an open interval (0, 1) and satisfies $\sum_{k=1}^{N} \rho_{m,k} = 1$.

The proof of Theorem 1 is shown in Appendix A.

We are now at $w = w_0$ to perform saddle point approximation.

Proposition 1. For non-zero natural number q, integral over complex AWGN vector w is approximated by

$$\int_{\boldsymbol{w}} \frac{1}{\pi^{N_{\mathrm{R}}} \sigma^{2N_{\mathrm{R}}}} \left(\sum_{k=1}^{N} \sigma_{m,k}(\boldsymbol{w}) \right)^{-q} \mathrm{d}\boldsymbol{w} \approx \left[\sum_{k=1}^{N} \exp\left(-\frac{\alpha_{m,k} || \boldsymbol{c}_{m,k} ||^{2}}{\sigma^{2}} \right) \right]^{-q}$$
(8)

$$\alpha_{m,k} = \left[1 - q \frac{||\vec{c}_{m}||^{2}}{||\vec{c}_{m,k}||^{2}} + q \frac{\vec{c}_{m,k} \cdot \vec{c}_{m} + \vec{c}_{m} \cdot \vec{c}_{m,k}}{||\vec{c}_{m,k}||^{2}} - \left(\frac{||\vec{c}_{m,k}||^{2}}{\sigma^{2}}\right)^{-1} \ln\left(\sum_{k=1}^{N} \frac{\rho_{m,k}||\vec{c}_{m,k}||^{2}}{\sigma^{2}} - \frac{||\vec{c}_{m}||^{2}}{\sigma^{2}}\right) \right]$$
(9)

The proof of Proposition 1 is shown in Appendix B. Mutual information is approximated by (5) and (8) as

$$\mathcal{I}(\mathbf{x};\mathbf{y}|\mathbf{H}) \approx \mathcal{I}(\alpha_{m,k},\mathbf{H}) = \log_2 N - \frac{1}{N} \sum_{m=1}^N \log_2 \left[\sum_{k=1}^N \exp\left(-\frac{\alpha_{m,k}||\mathbf{c}_{m,k}||^2}{\sigma^2}\right) \right]$$
(10)

Generally speaking, the close-form solution is hardly obtained. That is to say, we cannot write down the exact expression of $\alpha_{m,k}$. Optionally, we adopt a numerical method to obtain approximated $\alpha_{m,k}$,

$$\underset{\alpha_{m,k}:k=1,2}{\operatorname{argmin}} \left\{ |\mathcal{I}(\mathbf{x};\mathbf{y}|1) - \mathcal{I}(\alpha_{m,k},1) \right\} \approx \left\{ 3 - \exp[-||\mathbf{c}_{m,k}||^2 / (4\sigma^2)] \right\}^{-1}$$
(11)

where $\mathcal{I}(\mathbf{x}; \mathbf{y}|1)$ is computed by Monte Carlo method by taking BPSK over single-input and single-output (SISO) over AWGN channel (that is $\mathbf{H} = 1$) as an example, and $\alpha_{m,k}$ is fixed at each signal to noise ratio (SNR). Thus, (10) can be written as

$$\mathcal{I}(\mathbf{x}; \mathbf{y}|\mathbf{H}) \approx \log_2 N - \frac{1}{N} \sum_{m=1}^N \log_2 \left[\sum_{k=1}^N \exp\left(-\frac{||\mathbf{c}_{m,k}||^2}{\sigma^2} \frac{1}{3 - \exp[-||\mathbf{c}_{m,k}||^2/(4\sigma^2)]}\right) \right]$$

$$= \log_2 N - \frac{1}{N} \sum_{m=1}^N \log_2 \left[\sum_{k=1}^N \exp\left(-\frac{||\mathbf{H}\mathbf{d}_{m,k}||^2}{\sigma^2} \frac{1}{3 - \exp[-||\mathbf{H}\mathbf{d}_{m,k}||^2/(4\sigma^2)]}\right) \right]$$
(12)

3.2. Average Mutual Information over Doubly Correlated MIMO Rayleigh Fading Channels

The average mutual information over doubly correlated MIMO Rayleigh fading channels is computed as below,

$$\mathcal{I}(\mathbf{x};\mathbf{y}) \approx \log_2 N - \frac{1}{N} \sum_{m=1}^{N} \mathbb{E}_{H_{WG}} \left\{ \log_2 \left[\sum_{k=1}^{N} \exp\left(-\frac{||\boldsymbol{c}_{m,k}||^2}{\sigma^2} \alpha_{m,k}\right) \right] \right\}$$
(13)

Since $c_{m,k} = \Sigma_R^{1/2} H_{WG} \Sigma_T^{1/2} U_T^H d_{m,k}$, it is still quite hard to compute the expectation of H_{WG} . Consequently, (13) remains unsuitable for theoretical applications. Obviously, when SNR varies from $-\infty$ to $+\infty$, $\alpha_{m,k}$ satisfies $1/3 < \alpha_{m,k} < 1/2$ by (11), so average mutual information is approximated by

$$\mathcal{I}(\mathbf{x}; \mathbf{y}) \approx \log_2 N - \frac{1}{2N} \sum_{m=1}^{N} \mathbb{E}_{\mathbf{H}_{WG}} \left\{ \log_2 \left[\sum_{k=1}^{N} \exp\left(-\frac{||\boldsymbol{c}_{m,k}||^2}{2\sigma^2}\right) \right] \right\} - \frac{1}{2N} \sum_{m=1}^{N} \mathbb{E}_{\mathbf{H}_{WG}} \left\{ \log_2 \left[\sum_{k=1}^{N} \exp\left(-\frac{||\boldsymbol{c}_{m,k}||^2}{3\sigma^2}\right) \right] \right\}$$
(14)

For the simplified calculation, the following proposition provides approximate solution:

Proposition 2. Average mutual information integral over H_{WG} is lower bounded by

$$\mathcal{I}(\mathbf{x};\mathbf{y}) \geq \log_2 N - \frac{1}{2N} \sum_{m=1}^N \log_2 \left[\sum_{k=1}^N \prod_{l=1}^{N_R} \left(1 + \frac{[\mathbf{\Sigma}_R]_{l,l}}{2\sigma^2} d_{m,k}^{H} \mathbf{\sigma}^H \mathbf{\Sigma}_T \mathbf{\sigma} d_{m,k} \right)^{-1} \right] - \frac{1}{2N} \sum_{m=1}^N \log_2 \left[\sum_{k=1}^N \prod_{l=1}^{N_R} \left(1 + \frac{[\mathbf{\Sigma}_R]_{l,l}}{3\sigma^2} d_{m,k}^{H} \mathbf{\sigma}^H \mathbf{\Sigma}_T \mathbf{\sigma} d_{m,k} \right)^{-1} \right]$$
(15)

where $\boldsymbol{P} = \boldsymbol{U}_{\mathrm{T}}^{\mathrm{H}}$.

The proof of Proposition 2 is shown in Appendix C.

4. Simulation Verification and Result Analysis

This section presents examples to illustrate that the saddle point approximation method is very accurate. We considered an exponential correlation model. According to [23], the correlation matrix elements of transmitting and receiving antennas can be expressed as:

$$\begin{cases} [\Psi_{\rm T}(\rho_{\rm T})]I_{\rm T}, j_{\rm T} = \rho_{\rm T}^{|I_{\rm T}-j_{\rm T}|} \\ [\Psi_{\rm R}(\rho_{\rm R})]I_{\rm R}, j_{\rm R} = \rho_{\rm R}^{|I_{\rm R}-j_{\rm R}|} \end{cases} \text{ and } \begin{cases} I_{\rm T}, j_{\rm T} = 1, 2, \dots, N_{\rm T} \\ I_{\rm R}, j_{\rm R} = 1, 2, \dots, N_{\rm R} \end{cases}$$
(16)

where $\rho_T, \rho_R \in [0, 1)$.

4.1. Accuracy of Saddle Point Approximation

In Figures 1 and 2, doubly correlated Rayleigh fading and Rice fading channel models were considered, respectively. We compared the average mutual information by the Monte Carlo method and saddle point approximation method by (12). Different input types (BPSK, QPSK, QAM, 8PSK, and 16QAM) were assigned to transmitting antennas to ensure generality. Obviously, with the increase in SNR, mutual information presented an upward trend. When the SNR was greater than 15dB, mutual information tended to be stable. In these cases, (12) offered a very good approximation to the average mutual information for known channel state information.



Figure 1. Comparison on MI calculated by Monte Carlo and saddle point approximation under different input types and correction parameters ($\rho_T = \rho_R = \rho$) over doubly correlated Rayleigh fading channel model.



Figure 2. Comparison on MI calculated by Monte Carlo and saddle point approximation under different input types and correction parameters ($\rho_T = \rho_R = \rho$) over doubly correlated Rice fading channel model.

Figure 3 compares normalized MI calculated by Monte Carlo and saddle point approximation according to upper and lower bounds of $\alpha_{m,k}$ by (14) over doubly correlated Rayleigh fading channels. At low SNR, the normalized average mutual information of different modulation signals had little difference. With the increase in SNR, the normalized average mutual information of different input types tended to 1, and the growth rate of BPSK was the fastest. All simulation values were very close to the approximation values.



Figure 3. Comparison on normalized MI calculated by Monte Carlo and saddle point approximation according to upper and lower bounds of $\alpha_{m,k}$ by (14) under different input types, correction parameters ($\rho_T = \rho_R = \rho = 0.5$) over doubly correlated Rayleigh fading channel model ($N_T = N_R = 2$).

Comparison of MI calculated by Monte Carlo and the lower bound of saddle point approximation methods by (15) over doubly correlated Rayleigh fading channel are shown in Figure 4. With the increase in SNR, the mutual information increased, and the accuracy of saddle point approximation became higher. It was also clear that the lower bound of saddle point approximation achieved considerable accuracy for different doubly correlated MIMO fading channel scenarios.



Figure 4. Comparison on MI calculated by Monte Carlo and the lower bound of saddle point approximation method under different input types and correction parameters ($\rho_T = \rho_R = 0.5$) over doubly correlated Rayleigh fading channel model.

4.2. Conciseness of Saddle Point Approximation

The average mutual information has no closed-form expression, so it is usually calculated by the Monte Carlo method. The more sample points, the more accurate the calculation result is. We denoted the sample points as N_W . In order to obtain a relatively accurate value of mutual information, N_W was at least 10^4 . Tables 1 and 2 compare the computational complexity of the Monte Carlo method and the saddle point approximation method according to the number of operations and CPU time under the condition that N_W was 10^4 . The codes of mutual information calculation based on the Monte Carlo method and the saddle point approximation method were executed on an Intel Core i5-5200U 2.20 GHz processor. The results showed that the computational complexity of the proposed saddle point approximation method was much lower than that of the traditional Monte Carlo method. For example, as shown in Table 2, when N_T and N_R were equal to 2 and the input type of the two transmitting antennas was 16QAM, the CPU time of saddle point approximation methods by (15) was several orders of magnitude less than that of Monte Carlo method.

Table 1. Comparison of computational complexity between Monte Carlo method and saddle point approximation method according to the number of operations.

Number of Operations	Monte Carlo Method	Formula (15)
Exponential operation	$N_W * (N^2 + 1)$	0
Logarithm operation	N _W * N	2N + 1

Table 2. Comparison of computational complexity between Monte Carlo method and saddle point approximation method according to CPU time (seconds) under different input types and correction parameters ($\rho_{\rm T} = \rho_R = 0.4$). The symbol/indicates the CPU time is more than half an hour.

Cases	Input Type	Monte Carlo Method	Formula (15)	
$N_{\rm T} = N_{\rm R} = 2$	BPSK	3.043959	0.001049	
$N_{\rm T} = N_{\rm R} = 3$	BPSK	4.44435	0.019326	
$N_T = N_R = 4$	BPSK	8.427808	0.029586	
$N_{\rm T} = N_{\rm R} = 2$	QPSK	7.457080	0.027120	
$N_{\rm T} = N_{\rm R} = 3$	QPSK	69.558626	0.081062	
$N_{\rm T} = N_{\rm R} = 4$	QPSK	/	0.558082	
$N_{\rm T} = N_{\rm R} = 2$	8PSK	58.627341	0.070549	
$N_{\rm T} = N_{\rm R} = 3$	8PSK	/	1.573137	
$N_T = N_R = 4$	8PSK	/	/	
$N_{\rm T} = N_{\rm R} = 2$	16QAM	1281.853	0.612448	
$N_{\rm T} = N_{\rm R} = 3$	16QAM	/	255.736633	
$N_T = N_R = 4$	16QAM	/	/	

5. Conclusions

This paper studied the numerical calculation of mutual information for finite-alphabetbased transmissions over doubly correlated MIMO fading channels. The average mutual information was dominated by statistical CSI, and the obstacle of computation was complexity. We examined the appropriateness of the saddle point method first. Then mutual information over any known channel model was calculated by saddle point approximation. Furthermore, we induced the expectation of mutual information over doubly correlated MIMO Rayleigh-fading channels. Numerical results for various MIMO scenarios showed the efficacy of the proposed method. Compared to existing conclusions, the proposed approximation is of considerable accuracy in estimating the average mutual information with radically reduced complexity. It is promising that its accuracy and convenience will facilitate the practical application of mutual information.

Author Contributions: Conceptualization, J.Z., D.Z., and Y.L.; methodology, J.Z. and Y.L; validation, J.Z. and Y.L.; investigation, J.Z. and D.Z.; writing—original draft preparation, Y.L.; writing—review and editing, J.Z. and D.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Basic Research Business Expenses of Beijing Jiaotong University—Special Project of Frontier Science Center of Smart High Speed Railway System (number 2020JBZD004, 2020JBZD010); and Enterprise project (number AIPEC/BJJT/20BW1219).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors wish to thank the reviewers for their valuable comments and suggestions concerning this manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Define $\tilde{c}_{m,k}$ and \tilde{w} for computational simplicity,

$$\widetilde{c}_{m,k} = \begin{bmatrix} \operatorname{Re}\{c_{m,k}\} \\ \operatorname{Im}\{c_{m,k}\} \end{bmatrix} \text{ and } \widetilde{w} = \begin{bmatrix} \operatorname{Re}\{w\} \\ \operatorname{Im}\{w\} \end{bmatrix}$$
(A1)

where Re{•} and Im{•} stand for the real and imaginary components of a complex number •. $\tilde{c}_{m,k}$ and \tilde{w} are $2N_{\rm R} \times 1$ dimensional real vectors ($\tilde{c}_{m,k} \in \mathbb{R}^{2N_{\rm R} \times 1}$ and $\tilde{w} \in \mathbb{R}^{2N_{\rm R} \times 1}$) that satisfy $||\tilde{c}_{m,k}||^2 = ||c_{m,k}||^2$ and $||\tilde{w}||^2 = ||w||^2$.

By (6) and (A1), we have

$$\left(\sum_{k=1}^{N}\sigma_{m,k}(\widetilde{w})\right)^{-q} = \sigma_m^{-q}(\widetilde{w}) = \left[\sum_{k=1}^{N}\exp\left(\frac{||\widetilde{w} - q\widetilde{c}_{m,k}||^2}{q\sigma^2} - \frac{(q+1)||\widetilde{c}_{m,k}||^2}{\sigma^2}\right)\right]^{-q}$$
(A2)

Since *q* is positive integers, it is easy to verify that

$$\begin{aligned}
\sigma_m^{-q}(\widetilde{\boldsymbol{w}}) &= \left[\sum_{k=1}^N \exp\left(\frac{||\widetilde{\boldsymbol{w}} - q\widetilde{\boldsymbol{c}}_{m,k}||^2}{q\sigma^2} - \frac{(q+1)||\widetilde{\boldsymbol{c}}_{m,k}||^2}{\sigma^2}\right)\right]^{-q} > 0 \\
\lim_{[\widetilde{\boldsymbol{w}}]_l \to +\infty} \{\sigma_m^{-q}(\widetilde{\boldsymbol{w}})\} &= \lim_{[\widetilde{\boldsymbol{w}}]_l \to +\infty} \left[\sum_{k=1}^N \exp\left(\frac{||\widetilde{\boldsymbol{w}} - q\widetilde{\boldsymbol{c}}_{m,k}||^2}{q\sigma^2} - \frac{(q+1)||\widetilde{\boldsymbol{c}}_{m,k}||^2}{\sigma^2}\right)\right]^{-q} \to 0 \quad (A3) \\
\lim_{[\widetilde{\boldsymbol{w}}]_l \to -\infty} \{\sigma_m^{-q}(\widetilde{\boldsymbol{w}})\} &= \lim_{[\widetilde{\boldsymbol{w}}]_l \to -\infty} \left[\sum_{k=1}^N \exp\left(\frac{||\widetilde{\boldsymbol{w}} - q\widetilde{\boldsymbol{c}}_{m,k}||^2}{q\sigma^2} - \frac{(q+1)||\widetilde{\boldsymbol{c}}_{m,k}||^2}{\sigma^2}\right)\right]^{-q} \to 0
\end{aligned}$$

Therefore, there is a maximum value of $\sigma_m^{-q}(\tilde{w})$, which satisfies the conditions of saddle point approximation calculation. Assuming $\sigma_m^{-q}(\tilde{w})$ achieves the maximum at $\tilde{w} = \tilde{w}_0$, \tilde{w}_0 satisfies

$$\begin{cases} \operatorname{grad}\{\ln[\sigma_m^{-q}(\widetilde{\boldsymbol{w}})]\}|_{\widetilde{\boldsymbol{w}}_0} = 0\\ H\{\ln[\sigma_m^{-q}(\widetilde{\boldsymbol{w}})]\}|_{\widetilde{\boldsymbol{w}}_0} \prec 0 \end{cases}$$
(A4)

where $H\{\ln[\sigma_m^{-q}(\tilde{w})]\}|_{\tilde{w}_0}$ is the Hessian matrix of $\ln[\sigma_m^{-q}(\tilde{w})]$ at $\tilde{w} = \tilde{w}_0$. By the fact of $\sigma_m(\tilde{w}) > 0$ and q > 0, for $I = 1, 2, ..., 2N_R$, grad $\{\ln[\sigma_m^{-q}(\tilde{w})]\}|_{\tilde{w}_0} = 0$ is equivalent to

$$\sum_{k=1}^{N} \frac{2(\tilde{w} - q\tilde{c}_{m,k})}{q\sigma^2} \exp\left(\frac{||\tilde{w} - q\tilde{c}_{m,k}||^2}{q\sigma^2} - \frac{(q+1)||\tilde{c}_{m,k}||^2}{\sigma^2}\right)\Big|_{\tilde{w} = \tilde{w}_0} = 0$$
(A5)

and then the Hessian matrix $H\{\ln[\sigma_m^{-q}(\widetilde{w})]\}_{L_i}(\widetilde{w})|_{\widetilde{w}_0}$ is rewritten as

$$\begin{aligned} H\{\ln[\sigma_m^{-q}(\widetilde{w})]\}I, j(\widetilde{w})|_{\widetilde{w}_0} &= -q\mathbf{H}\{\ln[\sigma_m(\widetilde{w})]\}I, j(\widetilde{w})|_{\widetilde{w}_0} \\ &= -\frac{q}{\sigma_m(\widetilde{w})} \left[\left\{ \frac{\partial^2}{\partial[\widetilde{w}]_I \partial[\widetilde{w}]_j} \sigma_m(\widetilde{w}) \right\}_{I, j=1, 2, \dots, 2N_R} \right]|_{\widetilde{w}=\widetilde{w}_0} \\ &= -\sum_{k=1}^N \frac{4(\widetilde{w}_0 - q\widetilde{c}_{m,k})(\widetilde{w}_0 - q\widetilde{c}_{m,k})^T \widetilde{\sigma}_{m,k}(\widetilde{w}_0)}{q\sigma^4 \sigma_m(\widetilde{w}_0)} - \frac{2}{\sigma^2} \mathbf{I}_{2N_R} \prec 0 \end{aligned}$$
(A6)

Note that (A5) is equivalent to an implicit function of $\tilde{c}_{m,1}$, $\tilde{c}_{m,2}$, ..., $\tilde{c}_{m,N}$ and \tilde{w}_0 as below

$$\mathcal{F}(\widetilde{c}_{m,1},\ldots,\widetilde{c}_{m,N},\widetilde{w}_0) = \left[\left\{ f_I(\widetilde{c}_{m,1},\ldots,\widetilde{c}_{m,N},\widetilde{w}_0) \right\} |_{I=1,2,\ldots,2N_{\mathrm{R}}} \right] = 0$$
(A7)

where

$$f_{I}(\widetilde{\boldsymbol{c}}_{m,1},\ldots,\widetilde{\boldsymbol{c}}_{m,N},\widetilde{\boldsymbol{w}}) = \sum_{k=1}^{N} \left([\widetilde{\boldsymbol{w}}]_{j} - q[\widetilde{\boldsymbol{c}}_{m,k}]_{j} \right) \exp\left(\frac{1}{q\sigma^{2}} \sum_{l=1}^{2N_{\mathrm{R}}} \left([\widetilde{\boldsymbol{w}}]_{l} - q[\widetilde{\boldsymbol{c}}_{m,k}]_{l} \right)^{2} - \frac{q+1}{\sigma^{2}} \sum_{l=1}^{2N_{\mathrm{R}}} \left[\widetilde{\boldsymbol{c}}_{m,k} \right]_{l}^{2} \right)$$
(A8)

for $i = 1, 2, ..., 2N_R$. Then the Jacobi matrix of $\mathcal{F}(\tilde{c}_{m,1}, ..., \tilde{c}_{m,N}, \tilde{w})$ is computed as below

$$\mathbf{J}_{\mathcal{F}}|_{(\widetilde{c}_{m,1},\dots,\widetilde{c}_{m,N},\widetilde{w})} = \left[\left\{ \frac{\partial}{\partial [\widetilde{w}]_{j}} f_{I}(\widetilde{c}_{m,1},\dots,\widetilde{c}_{m,N},\widetilde{w}) \right\}_{I,j=1,2,\dots,2N_{R}} \right]$$

$$= \sigma_{m}(\widetilde{w}) \left(I_{2N_{R}} + \frac{2}{q\sigma^{2}\sigma_{m}(\widetilde{w})} \sum_{k=1}^{N} (\widetilde{w} - q\widetilde{c}_{m,k}) (\widetilde{w} - q\widetilde{c}_{m,k})^{\mathrm{T}} \sigma_{m,k}(\widetilde{w}) \right)$$
(A9)

Recalling (A6), $\mathbf{J}_{\mathcal{F}}|_{(\tilde{c}_{m,1},...,\tilde{c}_{m,N},\tilde{w})}$ is a positive definite matrix at \tilde{w}_0 , so it is invertible. Consequently, \tilde{w}_0 can, in principle, express in terms of $\tilde{c}_{m,1}$, $\tilde{c}_{m,2}$, ..., $\tilde{c}_{m,N}$ by implicit function theorem. Namely, the maximum of $\sigma_m^{-q}(\tilde{w})$ is achieved on the condition that \tilde{w}_0 satisfies (A5). Note that a complex number is zero when and only when both its real and imaginary parts are zero vectors, so we have,

$$w_0 = \sum_{k=1}^N q \rho_{m,k} c_{m,k} \triangleq q \bar{c}_m \tag{A10}$$

where $\rho_{m,k} = \sigma_{m,k}(w_0) / \sigma_m(w_0)$ is a positive real number over an open interval (0, 1) and satisfies $\sum_{k=1}^{N} \rho_{m,k} = 1$. So w_0 and c_m are both weighted average vectors of $c_{m,1}, c_{m,2}, \ldots$, and $c_{m,N}$.

Appendix **B**

Lemma 1 denotes that $\sigma_m^{-q}(\tilde{w})$ is maximized at \tilde{w}_0 . By (A5), we have,

$$\operatorname{grad}\left\{\ln[p_m^{-q}(\widetilde{w})]\right\}|_{\widetilde{w}_0} = -\frac{q}{p_m(\widetilde{w})} \left[\left\{ \frac{\partial \sigma_m(\widetilde{w})}{\partial[\widetilde{w}]_j} \right\} \Big|_{I=1,2,\dots,2N_{\rm R}} \right] \Big|_{\widetilde{w}=\widetilde{w}_0} = 0 \quad (A11)$$

By (A5), (A6), and (A10), the Taylor series of $\ln[\sigma_m^{-q}(\tilde{w})]$ is expanded to

$$\ln[\sigma_m^{-q}(\widetilde{w})] \approx \ln[\sigma_m^{-q}(\widetilde{w}_0)] + \frac{1}{2} (\widetilde{w} - \widetilde{w}_0)^{\mathrm{T}} H_{\widetilde{w}_0}(\widetilde{w} - \widetilde{w}_0)$$
(A12)

Note that a positive definite matrix *A* is invertible, and the determinant of *A* can be computed by $\exp[\operatorname{Tr} \ln(A)]$, where Tr*A* stands for the trace of *A*. Recalling (A1) and (A6), the saddle point approximation can be computed by Gaussian integral

$$\int_{\boldsymbol{w}} \frac{p_m^{-q}(\boldsymbol{w})}{\pi^{N_{\mathsf{R}}} \sigma^{2N_{\mathsf{R}}}} d\boldsymbol{w} \approx \frac{p_m^{-q}(\tilde{\boldsymbol{w}}_0)}{\pi^{N_{\mathsf{R}}} p^{2N_{\mathsf{R}}}} \int_{\tilde{\boldsymbol{w}}} \exp\left(\frac{1}{2}(\tilde{\boldsymbol{w}}-\tilde{\boldsymbol{w}}_0)^{\mathsf{T}} \boldsymbol{H}_{\tilde{\boldsymbol{w}}_0}(\tilde{\boldsymbol{w}}-\tilde{\boldsymbol{w}}_0)\right) d\tilde{\boldsymbol{w}} \\
= \left\{ p_m(\tilde{\boldsymbol{w}}_0) \det^{1/2q} \left(-\frac{\sigma^2}{2} \boldsymbol{H}_{\tilde{\boldsymbol{w}}_0}\right) \right\}^{-q} \ge \left\{ p_m(\tilde{\boldsymbol{w}}_0) * \frac{1}{2q} \operatorname{Tr} \left(-\frac{\sigma^2}{2} \boldsymbol{H}_{\tilde{\boldsymbol{w}}_0}\right) \right\}^{-q} \\
\ge \sum_{k=1}^{N} \exp\left\{ \left[\begin{array}{c} -\frac{||\boldsymbol{c}_m||^2}{\sigma^2} \\ 1 - q \frac{||\boldsymbol{c}_m||^2}{||\boldsymbol{c}_{m,k}||^2} + q \frac{\boldsymbol{c}_{m,k} + \boldsymbol{c}_m + \boldsymbol{c}_m + \boldsymbol{c}_m + \boldsymbol{c}_m}{||\boldsymbol{c}_{m,k}||^2}} - \left(\frac{||\boldsymbol{c}_m||^2}{\sigma^2}\right)^{-1} \\ \ln\left(\sum_{k=1}^{N} \frac{|\boldsymbol{\rho}_m]_k ||\boldsymbol{c}_{m,k}||^2}{\sigma^2} - \frac{||\boldsymbol{c}_m||^2}{\sigma^2}\right) \right] \right\}^{-q}$$
(A13)

According to (A5), we can induce

$$\sum_{k=1}^{N} \left(\vec{c}_m - c_{m,k} \right) \exp\left[-\frac{||c_{m,k}||^2}{\sigma^2} \left(1 + q \frac{c_{m,k}^{H} \vec{c}_m + \vec{c}_m^{H} c_{m,k}}{\sigma^2} - q \frac{||\vec{c}_m||^2}{\sigma^2} \right) \right] = 0 \quad (A14)$$

Therefore, (A13) and (A14) demonstrate that Gaussian integral is dominated by $||c_{m,k}||^2/\sigma^2$ in terms of exponential. Define a multiplier $\alpha_{m,k}$ dominated by $||c_{m,k}||^2/\sigma^2$,

$$\alpha_{m,k} = \begin{bmatrix} 1 - q \frac{||\vec{c}_{m}||^{2}}{||c_{m,k}||^{2}} + q \frac{c_{m,k} + \vec{c}_{m} + \vec{c}_{m} + c_{m,k}}{||c_{m,k}||^{2}} - \left(\frac{||c_{m,k}||^{2}}{\sigma^{2}}\right)^{-1} \\ \ln \left(\sum_{k=1}^{N} \frac{|\rho_{m}]_{k} ||c_{m,k}||^{2}}{\sigma^{2}} - \frac{||\vec{c}_{m}||^{2}}{\sigma^{2}}\right) \end{bmatrix}$$
(A15)

So

$$\int_{\boldsymbol{w}} \frac{1}{\pi^{N_{\mathrm{R}}} \sigma^{2N_{\mathrm{R}}}} \left(\sum_{k=1}^{N} p_{m,k}(\boldsymbol{w}) \right)^{-q} \mathrm{d}\boldsymbol{w} \approx \left[\sum_{k=1}^{N} \exp\left(-\frac{\alpha_{m,k} ||\boldsymbol{c}_{m,k}||^2}{\sigma^2}\right) \right]^{-q}$$
(A16)

Appendix C

Obviously, when SNR varies from $-\infty$ to $+\infty$, $\alpha_{m,k}$ satisfies $1/3 < \alpha_{m,k} < 1/2$ by (11), so the average mutual information over doubly correlated MIMO Rayleigh-fading channels is approximated by

$$\mathcal{I}(\mathbf{x};\mathbf{y}) \approx -\frac{1}{2N} \sum_{m=1}^{N} \mathbb{E}_{H_{WG}} \left\{ \log_2 \left[\frac{1}{N} \sum_{k=1}^{N} \exp\left(-\frac{||\boldsymbol{c}_{m,k}||^2}{2\sigma^2}\right) \right] \right\} -\frac{1}{2N} \sum_{m=1}^{N} \mathbb{E}_{H_{WG}} \left\{ \log_2 \left[\frac{1}{N} \sum_{k=1}^{N} \exp\left(-\frac{||\boldsymbol{c}_{m,k}||^2}{3\sigma^2}\right) \right] \right\}$$
(A17)

Because $log_2(x)$ is a concave function, by Jensen's inequality, we have

$$\mathbb{E}_{\boldsymbol{H}_{\mathrm{WG}}}\left\{\log_{2}\left[\sum_{k=1}^{N}\exp\left(-\frac{||\boldsymbol{c}_{m,k}||^{2}}{\alpha_{m,k}\sigma^{2}}\right)\right]\right\} \leq \log_{2}\left[\sum_{k=1}^{N}\mathbb{E}_{\boldsymbol{H}_{\mathrm{WG}}}\left\{\exp\left(-\frac{||\boldsymbol{c}_{m,k}||^{2}}{\alpha_{m,k}\sigma^{2}}\right)\right\}\right]$$
(A18)

where $c_{m,k} = \Sigma_R^{1/2} H_{WG} \Sigma_T^{1/2} U_T^H d_{m,k}$, so expectation in (A18) is rewritten as

$$\mathbb{E}_{\boldsymbol{H}_{\mathrm{WG}}}\left\{\exp\left(-\frac{||\boldsymbol{c}_{m,k}||^{2}}{\alpha_{m,k}\sigma^{2}}\right)\right\} = \mathbb{E}_{\boldsymbol{H}_{\mathrm{WG}}}\left\{\exp\left(-\sum_{l=1}^{N_{\mathrm{R}}}\frac{[\boldsymbol{H}_{\mathrm{WG}}]_{l}([\boldsymbol{\Sigma}_{\mathrm{R}}]_{l,l}\boldsymbol{q}_{m,k})[\boldsymbol{H}_{\mathrm{WG}}]_{l}^{\mathrm{H}}}{\alpha_{m,k}\sigma^{2}}\right)\right\}$$
(A19)

where

$$\mathbf{q}_{m,k} = \mathbf{\Sigma}_{\mathrm{T}}^{1/2} \boldsymbol{\sigma} \boldsymbol{d}_{m,k} \boldsymbol{d}_{m,k}^{\mathrm{H}} \boldsymbol{\sigma}^{\mathrm{H}} \mathbf{\Sigma}_{\mathrm{T}}^{1/2} \text{ and } \boldsymbol{\sigma} = \boldsymbol{U}_{\mathrm{T}}^{\mathrm{H}}$$
(A20)

 $[\mathbf{H}_{WG}]_l$ stands for the l^{th} row of \mathbf{H}_{WG} .

Since H_{WG} is an independent and identically distributed complex AWGN matrix, $[H_{WG}]_{l}$ is an independent and identically distributed complex AWGN vector. So

$$\mathbb{E}_{\boldsymbol{H}_{WG}}\left\{\exp\left(-\frac{||\boldsymbol{c}_{m,k}||^{2}}{\alpha_{m,k}\sigma^{2}}\right)\right\} = \mathbb{E}_{\boldsymbol{H}_{WG}}\left\{\exp\left(-\frac{1}{\alpha_{m,k}\sigma^{2}}\sum_{l=1}^{N_{R}}\left[\boldsymbol{H}_{WG}\right]_{l}([\boldsymbol{\Sigma}_{R}]_{l,l}\boldsymbol{Q}_{m,k})[\boldsymbol{H}_{WG}]_{l}^{H}\right)\right\} \\
= \prod_{l=1}^{N_{R}}\mathbb{E}_{[\boldsymbol{H}_{WG}]_{l}}\left\{\exp\left(-\frac{[\boldsymbol{H}_{WG}]_{l}([\boldsymbol{\Sigma}_{R}]_{l,l}\boldsymbol{q}_{m,k})[\boldsymbol{H}_{WG}]_{l}^{H}}{\alpha_{m,k}\sigma^{2}}\right)\right\} \\
= \prod_{l=1}^{N_{R}}\int_{[\boldsymbol{H}_{WG}]_{l}}p([\boldsymbol{H}_{WG}]_{l})\exp\left(\begin{array}{c}-\frac{1}{\alpha_{m,k}\sigma^{2}}\\[\boldsymbol{H}_{WG}]l([\boldsymbol{\Sigma}_{R}]_{l,l}\boldsymbol{1}\boldsymbol{q}_{m,k})\\[\boldsymbol{H}_{WG}]l^{H}\end{array}\right)d[\boldsymbol{H}_{WG}]_{l} \tag{A21} \\
= \prod_{l=1}^{N_{R}}\int_{[\boldsymbol{H}_{WG}]_{l}}\frac{1}{\pi^{N_{T}}}\exp\left(\begin{array}{c}-[\boldsymbol{H}_{WG}]l\\[\boldsymbol{H}_{WG}]l^{H}\end{array}\right)d[\boldsymbol{H}_{WG}]_{l}$$

According to [24],

$$\int_{h\in\mathbb{C}^{N\times 1}} \exp\left(-h^{h}(A+\mathbf{j}B)h\right) dh = \frac{\pi^{N}}{\det(A+\mathbf{j}B)}$$
(A22)

where $[h_1]$ is iid $\mathcal{CN}(0, \sigma)$. (A21) can be written as

$$\mathbb{E}_{H_{WG}}\left\{\exp\left(-\frac{||\boldsymbol{c}_{m,k}||^{2}}{\alpha_{m,k}\sigma^{2}}\right)\right\} = \prod_{l=1}^{N_{R}}\left[\det\left(\boldsymbol{I}_{N_{T}}+\frac{[\boldsymbol{\Sigma}_{R}]_{l,l}}{\alpha_{m,k}\sigma^{2}}\boldsymbol{Q}_{m,k}\right)\right]^{-1}$$
$$= \prod_{l=1}^{N_{R}}\left[\det\left(\begin{array}{c}\boldsymbol{I}_{N_{T}}+\\\frac{[\boldsymbol{\Sigma}_{R}]_{l,l}}{\alpha_{m,k}\sigma^{2}}\boldsymbol{\Sigma}_{T}^{1/2}\boldsymbol{P}\boldsymbol{d}_{m,k}\boldsymbol{d}_{m,k}^{H}\boldsymbol{P}^{H}\boldsymbol{\Sigma}_{T}^{1/2}\end{array}\right)\right]^{-1}$$
(A23)

Note that for column vector α ,

$$\det(\mathbf{I}_N + \boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{H}}) = 1 + [\boldsymbol{\Sigma}_{\boldsymbol{\alpha}}]_1 = 1 + \operatorname{tr}(\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{H}}) = 1 + \boldsymbol{\alpha}^{\mathrm{H}}\boldsymbol{\alpha}$$
(A24)

we have

$$\mathbb{E}_{\boldsymbol{H}_{\mathrm{WG}}}\left\{\exp\left(-\frac{||\boldsymbol{c}_{m,k}||^{2}}{\alpha_{m,k}\sigma^{2}}\right)\right\} = \prod_{l=1}^{N_{\mathrm{R}}} \left(1 + \frac{[\boldsymbol{\Sigma}_{\mathrm{R}}]_{l,l}}{\alpha_{m,k}\sigma^{2}} \boldsymbol{d}_{m,k}^{\mathrm{H}} \boldsymbol{P}^{\mathrm{H}} \boldsymbol{\Sigma}_{\mathrm{T}} \boldsymbol{P} \boldsymbol{d}_{m,k}\right)^{-1}$$
(A25)

recall (A17), we have

$$\mathcal{I}(\mathbf{x};\mathbf{y}) \geq \log_2 N - \frac{1}{2N} \sum_{m=1}^{N} \log_2 \left[\sum_{k=1}^{N} \prod_{l=1}^{N_{\mathrm{R}}} \left(1 + \frac{[\mathbf{\Sigma}_{\mathrm{R}}]_{l,l}}{2\sigma^2} d_{m,k}^{\mathrm{H}} \mathbf{P}^{\mathrm{H}} \mathbf{\Sigma}_{\mathrm{T}} \mathbf{P} d_{m,k} \right)^{-1} \right] - \frac{1}{2N} \sum_{m=1}^{N} \log_2 \left[\sum_{k=1}^{N} \prod_{l=1}^{N_{\mathrm{R}}} \left(1 + \frac{[\mathbf{\Sigma}_{\mathrm{R}}]_{l,l}}{3\sigma^2} d_{m,k}^{\mathrm{H}} \mathbf{P}^{\mathrm{H}} \mathbf{\Sigma}_{\mathrm{T}} \mathbf{P} d_{m,k} \right)^{-1} \right]$$
(A26)

References

- Jin, S.; Gao, X. Statistical antenna selection for MIMO systems in double-sided correlated rayleigh fading channels. In Proceedings
 of the IEEE Wireless Communications and Networking Conference, Las Vegas, NV, USA, 3–6 April 2006; pp. 729–733.
- Zhang, Y.; Huo, Y.; Wang, D.; Dong, X.; You, X. Channel Estimation and Hybrid Precoding for Distributed Phased Arrays Based MIMO Wireless Communications. *IEEE Trans. Veh. Technol.* 2020, *69*, 12921–12937. [CrossRef]
- Jin, X.L.; Yang, J.D.; Song, K.Y.; No, J.S.; Shin, D.J. On the relationship between mutual information and bit error probability for some linear dispersion codes. *IEEE Tran. Wirel. Commun.* 2009, *8*, 90–94. [CrossRef]
- 4. Shannon, C.E. A mathematical theory of communication. Bell Syst. Technol. J. 1948, 27, 379–423, 623–656. [CrossRef]
- Smith, P.J.; Shafi, M. On a Gaussian approximation to the capacity of wireless MIMO systems. In Proceedings of the 2002 IEEE International Conference on Communications. Conference Proceedings, ICC 2002, New York, NY, USA, 28 April–2 May 2002; pp. 406–410.
- Chuah, D.N.; Tse, D.N.C.; Kahn, J.M.; Valenzuela, R.A. Capacity scaling in MIMO wireless systems under correlated fading. *IEEE Trans. Inf. Theory* 2002, 48, 637–650. [CrossRef]
- Levin, G.; Loyka, S. On the Outage Capacity Distribution of Correlated Keyhole MIMO Channels. *IEEE Trans. Inf. Theory* 2008, 54, 3232–3245. [CrossRef]
- 8. Hsu, H. Digital communications over fading channels. *IEEE Circuits Devices Mag.* 2001, 17, 57.
- 9. Lozona, A.; Tulino, A.M.; Verdu, S. Optimum power allocation for parallel Gaussian channels with arbitrary input distributions. *IEEE Trans. Inf. Theory* **2006**, *52*, 3033–3051. [CrossRef]
- 10. Xiao, C.; Zheng, Y.R.; Ding, Z. Globally Optimal Linear Precoders for Finite Alphabet Signals Over Complex Vector Gaussian Channels. *IEEE Trans. Signal Process.* **2011**, *59*, 3301–3314. [CrossRef]
- 11. Cao, K.; Cai, Y.; Wu, Y.; Yang, W. Cooperative Jamming for Secure Communication with Finite Alphabet Inputs. *IEEE Commun. Lett.* **2017**, *21*, 2025–2028. [CrossRef]
- 12. Sadeghi, P.; Vontobel, P.O.; Shams, R. Optimization of information rate upper and lower bounds for channels with memory. *IEEE Trans. Inf. Theory* **2009**, *55*, 663–688. [CrossRef]
- 13. Güney, N.; Delic, H.; Alagöz, F. Achievable information rates of PPM impulse radio for UWB channels and rake reception. *IEEE Trans. Commun.* **2010**, *58*, 1524–1535. [CrossRef]

- 14. Owen, A.B. Monte Carlo extension of quasi-Monte Carlo. In Proceedings of the 1998 Winter Simulation Conference, Washington, DC, USA, 13–16 December 1998; pp. 571–577.
- Sayana, K.; Zhuang, J.; Stewart, K. Short Term Link Performance Modeling for ML Receivers with Mutual Information per Bit Metrics. In Proceedings of the IEEE GLOBECOM 2008–2008 IEEE Global Telecommunications Conference, New Orleans, LA, USA, 30 November–4 December 2008; pp. 1–6.
- Zeng, W.; Xiao, C.; Wang, M.; Lu, J. Linear precoding for finite alphabet inputs over MIMO fading channels with statistical CSI. *IEEE Trans. Signal Process.* 2012, 60, 3134–3148. [CrossRef]
- 17. Yang, P.; Yang, H. A Low-Complexity Linear Precoding for Secure Transmission over MIMOME Wiretap Channels with Finite-Alphabet Inputs. *IEEE Trans. Veh. Technol.* **2019**, *68*, 9896–9907. [CrossRef]
- 18. Zeng, W.; Zheng, Y.R.; Xiao, C. Multiantenna secure cognitive radio networks with finite-alphabet inputs: A global optimization approach for precoder design. *IEEE Trans. Wirel. Commun.* **2016**, *15*, 3044–3057. [CrossRef]
- Zeng, W.; Zheng, Y.R.; Wang, M.; Lu, J. Linear precoding for relay networks: A perspective on finite-alphabet inputs. *IEEE Trans.* Wireless Commun. 2012, 11, 1146–1157. [CrossRef]
- Yadav, A.; Juntti, M.; Lilleberg, J. Linear Precoder Design for Doubly Correlated Partially Coherent Fading MIMO Channels. *IEEE Trans. Wirel. Commun.* 2014, 13, 3621–3635. [CrossRef]
- Ouyang, C.; Wu, S.; Jiang, C.; Cheng, J.; Yang, H. Approximating Ergodic Mutual Information for Mixture Gamma Fading Channels with Discrete Inputs. *IEEE Commun. Lett.* 2020, 24, 734–738. [CrossRef]
- Xiao, C.; Zheng, Y.R. On the mutual information and power allocation for vector Gaussian channels with finite discrete inputs. In Proceedings of the IEEE GLOBECOM 2008–2008 IEEE Global Telecommunications Conference, New Orleans, LA, USA, 30 November–4 December 2008; pp. 1–5.
- 23. Wang, W.; Guyet, T.; Guiniou, R. Autonomic intrusion detection: Adaptively detecting anomalies over unlabeled audit data streams in computer networks. *Knowl. Based Syst.* **2014**, *70*, 103–117. [CrossRef]
- 24. Hassibi, B.; Marzetta, T. Multiple-antennas and isotropically random unitary inputs: The received signal density in closed form. *IEEE Trans. Inf. Theory* **2002**, *48*, 1473–1484. [CrossRef]