Outage and Finite-SNR DMT Analysis for IRS-aided MIMO Systems: How Large IRSs Need to Be?

Xin Zhang, Graduate Student Member, IEEE, Xianghao Yu, Member, IEEE, and S.H. Song, Member, IEEE

Abstract—Intelligent reflecting surfaces (IRSs) are promising enablers for high-capacity wireless communication systems by constructing favorable channels between the transmitter and receiver. However, general, accurate, and tractable outage analysis for IRS-aided multiple-input-multiple-output (MIMO) systems is not available in the literature. In this paper, we first characterize the mutual information (MI) of IRS-aided MIMO systems by capitalizing on large random matrix theory (RMT). Based on this result, a closed-form approximation for the outage probability is derived and a gradient-based algorithm is proposed to minimize the outage probability with statistical channel state information (CSI). We also investigate the diversity-multiplexing tradeoff (DMT) with the finite signal-to-noise ratio (SNR). Based on these theoretical results, we further study the impact of the IRS size on system performance. In the high SNR regime, we provide closed-form expressions for the ergodic mutual information (EMI) and outage probability as a function of the IRS size, which analytically reveal that the benefit of increasing the IRS size saturates quickly. Simulation results validate the accuracy of the theoretical analysis and confirm the increasing cost for deploying larger IRSs to improve system performance. For example, for an IRS-aided MIMO system with 20 antennas at both the transmitter and receiver, we need to double the size of the IRS to increase the throughput from 90% to 95% of its maximum value.

Index Terms—Intelligent reflecting surface (IRS), multiple-input-multiple-output (MIMO), outage probability, random matrix theory (RMT).

I. INTRODUCTION

Intelligent reflecting surfaces (IRSs) have attracted extensive interests from both academia and industry and are considered as one of the promising solutions for high-capacity future communication systems [1]. By designing the controllable phase shifts, IRSs can customize favorable channels between the transceivers [2], thus increasing the throughput and the reliability of wireless links [3]. In addition, IRSs are energy-efficient due to their passive nature. Inspired by these advantages, IRSs have been applied to various wireless systems such as massive multiple-input-multiple-output (MIMO) [4], simultaneous wireless information and power transfer (SWIPT) [5], and non-orthogonal multiple access (NOMA) [6] systems.

There have been works investigating the design and performance analysis of IRS-aided systems [7], [8], [9]. To this end, the capacity (throughput) and outage probability (reliability) are two important performance measures. The ergodic mutual information (EMI) (average throughput) has been well studied for single-input-single-output (SISO) [10] and multiple-input-single-output (MISO) systems [9]. The EMI of IRS-aided MIMO systems over Rician channels was investigated by random matrix theory (RMT) with an accurate and closed-form approximation [11].

The outage probability of IRS-aided SISO systems was evaluated in previous works. In [6], a closed-form expression for the outage probability of NOMA SISO networks was obtained under Nakagami fading using central limit theorem (CLT). The outage probability of a SISO system with multiple IRSs over Rician channels was analyzed in [8] using non-central chi-squared distribution. In [12], the outage probability of IRS-aided vehicular communication systems was evaluated using CLT, showing that the outage probability is strongly related to the vehicle density and the size of the IRS. Atapattu et al. derived a closed-form expression for the outage probability and provided the optimal phase shifts design [13]. It was also shown that the decreasing rate of the outage probability is related to the size of IRSs.

The outage probability of IRS-aided multiple-input-single-output (MISO) systems was also studied. In [14], considering reflection pattern modulation (RPM), a closed-form approximation for the asymptotic outage probability over Rician channels was obtained using Gamma approximation. In [15], with maximum-ratio transmission (MRT), the expression of the outage probability and its asymptotically-optimal form were given, while the phase shifts were optimized to minimize the outage probability. In [16] and [17], the robust design of IRS-aided MISO systems was investigated with imperfect channel state information (CSI). The Bernstein-type inequality was utilized to characterize the outage probability.

However, as far as the authors know, the outage probability of IRS-aided MIMO systems was only investigated in [18], where Mellin transform and finite-regime RMT were leveraged to derive the outage probability over Rayleigh fading channels with correlations only at one side of the transceivers. With statistical CSI, the outage probability was further minimized by optimizing the phase shifts using a genetic algorithm. Although the Mellin transform-based method gives an exact expression for the outage probability, there exist two main demerits. First, antennas at both transceivers and reflecting elements at IRSs are typically compactly packed together and correlated with each other. However, the result in [18] can only handle correlation at one side of the communication link due to the intractability of the Mellin transform. Second, the complexity of the exact expression and the implicit relation-
ship between the outage probability and phase shifts make it challenging to design the phase shifts and to reveal insights of the system. In other words, a generic and tractable outage characterization of IRS-aided MIMO systems has not been given. In this paper, we first characterize the distribution of the MI for IRS-aided MIMO systems and utilize it to evaluate the outage probability. We then propose a gradient descent algorithm to minimize the outage probability by optimizing the phase shifts. To better characterize the trade-off between diversity and multiplexing gain, we also investigate the finite-SNR diversity-multiplexing tradeoff (DMT) for IRS-aided systems. Finally, the impact of the IRS size on the system performance is studied to answer the question: How large IRSs need to be?

A. Contributions

The main contributions of this paper are listed as follows.

1) The distribution of the MI for IRS-aided MIMO systems is first derived. Based on the result, an approximation on the outage probability over general correlated channels is obtained. To the best of the authors' knowledge, this is the first analytical result regarding the outage probability in IRS-aided MIMO systems over general correlated channels. Numerical results validate the accuracy of the proposed method.

2) With the derived outage probability at hand, a gradient algorithm is proposed to minimize the outage probability by optimizing the phase shifts at the IRS, assuming only statistical CSI. Numerical results show that the algorithm can efficiently decrease the outage probability.

3) A closed-form expression is obtained for the finite-SNR DMT of IRS-aided MIMO systems, which is not available in the literature. An interesting observation is that the finite-SNR DMT is highly related to the ratio between the mean and the standard deviation of the MI. The accuracy of the expression is validated by numerical results.

4) The impact of the size of the IRS on system performance is investigated. To this end, we first propose the concept called IRS efficiency to measure the efficiency of increasing the IRS size in achieving the maximum throughput. The expression of the outage probability with respect to the IRS size is explicitly given in the high SNR regime. Based on the theoretical analysis and simulation results, we have two key observations. First, when the size of the IRS is infinitely large, the performance of the two-hop IRS system is the same as that of the single-hop link. Second, the benefit of increasing the size of the IRS saturates quickly. For example, for an IRS-aided system with 20 antennas at the transceivers over independent channels, we need to double the size of the IRS to increase the throughput from 90% to 95% of its maximum value.

B. Organizations

The rest of this paper is organized as follows. Section III presents the system model for the IRS-aided MIMO system. Section IV introduces the main results including the characterization of the distribution for the MI and its special version for the uncorrelated case. The analysis and optimization of the outage probability and the finite-SNR DMT are given in Section V. Section VI discusses the effect of the IRS size on system performance. The theoretical results are validated in Section VII by numerical simulations. Finally, Section VIII concludes the paper.

Notations: Bold upper case letters and bold lower case letters represent the matrix and vector, respectively. Re{·} denotes the real part of a complex number. P(·) is the probability measure. \( \mathbb{C}^N \) and \( \mathbb{C}^{M \times N} \) represent the space of N-dimensional vectors and the space of M-by-N matrices. \( A^H \) represents the conjugate transpose of \( A \). \( [A]_{i,j} \) represents the i, j-th entry of \( A \). \( \otimes \) denotes the element-wise product of matrices. \( \operatorname{Tr} A \) and \( ||A|| \) represent the trace and the spectral norm of \( A \). \( \Phi(x) \) is the cumulative distribution function (CDF) of standard Gaussian distribution and \( Q(\cdot) \) is the Q-function. \( \mathcal{CN} \) and \( \mathcal{N} \) represent the circularly complex Gaussian and real Gaussian distribution, respectively. \( D_{N \to \infty} \) and \( \mathcal{Q}_{N \to \infty} \) denote the convergence in distribution, the convergence in probability. \( O(\cdot) \) and \( o(\cdot) \) represent the big-O and the little-o notations, respectively.

II. System Model

A. System Model

We consider a point-to-point MIMO communication system, where an IRS is deployed to establish favorable communication links for the user that would otherwise be blocked. There are \( M \) antennas at the transmitter and \( N \) antennas at the receiver, and the number of elements at the IRS is \( L \). Given the line-of-sight (LoS) path is blocked, the received signal \( y \in \mathbb{C}^N \) is given by

\[
y = H_1 \Psi H_2 s + n, \tag{1}
\]

where \( s \in \mathbb{C}^M \) and \( n \in \mathbb{C}^N \sim \mathcal{CN}(0, \sigma^2 I) \) denote the transmitted signal with unit average transmit power, i.e., \( \mathbb{E}||s||^2 = 1, i = 1, 2, ..., M \), and the additive white Gaussian noise (AWGN) with variance \( \sigma^2 \), respectively. \( H_2 \in \mathbb{C}^{L \times M} \) and \( H_1 \in \mathbb{C}^{N \times L} \) represent the channel matrices from the transmitter to the IRS and from the IRS to the receiver, respectively. \( \Psi \in \mathbb{C}^{L \times L} = \text{diag}(\psi_1, \psi_2, ..., \psi_L) = \text{diag}(e^{j\theta_1}, e^{j\theta_2}, ..., e^{j\theta_L}) \) with \( \theta_i \in [0, 2\pi], i = 1, 2, ..., L \), are the phase shifts imposed by the IRS. In this paper, we consider the general Rayleigh model for \( H_1 \) and \( H_2 \) with

\[
H_1 = R_1^T X T_1^2, \quad H_2 = R_2^T Y T_2^2, \tag{2}
\]

where \( R_1 \) and \( T_1 \), \( i = 1, 2 \), are four positive semi-definite correlation matrices. \( R_1 \) and \( T_2 \) denote the correlation matrices of receive and transmit antennas, respectively. \( T_1 \) and \( R_2 \) denote the transmit and receive correlation matrices of the IRS. \( X \in \mathbb{C}^{N \times L} \) and \( Y \in \mathbb{C}^{L \times M} \) are two independent and identically distributed (i.i.d.) Gaussian random matrices, whose entries follow \( \mathcal{CN}(0, \frac{1}{L}) \) and \( \mathcal{CN}(0, \frac{1}{N}) \), respectively. We assume that statistical CSI, i.e. correlation matrices of the channel, is available.
B. Mutual Information and Outage Probability

The MI of the IRS-aided MIMO system is given by
\[ I(\rho) = \log \det(I + \rho H_1 \Psi H_2^H \Psi^H H_1^H), \tag{3} \]
where \( \rho = \frac{1}{N} \) denotes the SNR. Based on the distribution of the MI, we can evaluate the performance of IRS-aided MIMO systems with different metrics. For example, the average throughput can be determined by the EMI as \( \mathbb{E} I(\rho) \). On the other hand, the reliability of the system can be measured by the outage probability, which, for a preset transmission rate \( R \), can be written as
\[ P_{\text{out}}(R) = \mathbb{P}(I(\rho) < R). \tag{4} \]
Finally, the trade-off between the throughput and reliability can be measured by the diversity and multiplexing trade-off (DMT).

The EMI of IRS-aided MIMO systems has been obtained in the literature \cite{11}. In the following, we first investigate the distribution of \( I(\rho) \) and then utilize the result to analyze the outage probability and finite-SNR DMT of IRS-aided MIMO systems.

III. Characterization of the MI for IRS-aided MIMO Systems

Before introducing our main results, we first present some preliminary results including the approximation of the EMI. The analysis is based on RMT, which has been widely used and shown to be efficient in the performance analysis of MIMO systems \cite{19, 20}.

A. Assumptions and Existing Results on EMI

The results of this paper are developed based on the following assumptions.

Assumption 1. \( 0 < \lim \inf_{M \to \infty} \frac{M}{N} \leq \frac{M}{N} \leq \lim \sup_{M \to \infty} \frac{M}{N} < \infty, \) 
\[ 0 < \lim \inf_{M \to \infty} \frac{M}{N} \leq \frac{M}{N} \leq \lim \sup_{M \to \infty} \frac{M}{N} < \infty. \]

Assumption 2. \( \lim \sup_{M \to \infty} \left\| R_1 \right\| < \infty, \lim \sup_{M \to \infty} \left\| T_1 \right\| < \infty, i = 1, 2 \) \cite{19, 20}.

Assumption 3. \( \inf_{M \geq 1} \frac{1}{M} \text{Tr} R_1 > 0, \inf_{M \geq 1} \frac{1}{M} \text{Tr} T_2 > 0, \)
\[ \inf_{M \geq 1} \frac{1}{M} \text{Tr} T_1 \Psi R_2 \Psi^H > 0. \]

\( \Psi \) is the asymptotic regime considered for the large-scale system, indicating that the dimensions of the system \((M, N, L)\) grow to infinity with the same pace. Due to the unitary invariant attributes of Gaussian random matrices, \( X \) and \( X' \) are equivalent statistically, where \( U \) and \( V \) are both unitary matrices. The MI in (3) can be written as
\[ I(\rho) = \log \det(I + \rho R X' S X Y' S^H Y S^H R X'), \tag{5} \]
where \( R = U_R R_1 U_R^H, S = U_S T_S^H \Psi R_2 \Psi^H T_S^H U_R^H, \) and \( T = U_T T_2 U_T^H \) are the eigenvalue decompositions with \( R, S, \) and \( T \) being diagonal matrices. Given \( X \) (\( Y \)) and \( X' \) (\( Y' \)) are statistically equivalent, we will not differentiate them in the following. Thus, the equivalent channel matrix can be given by
\[ H = R X S Y T. \tag{6} \]

A.2 indicates that \( \lim \sup_{M \to \infty} \left\| S \right\| < \infty. \)

The following theorem gives an approximation for the EMI.

Theorem 1. \((\cite{22} \text{ Theorem 2}, \cite{24} \text{ Theorem 5} \) and \cite{27} \text{ Corollary 1}) With the channel matrix \( H \) given in (6), if A.1 and A.2 are satisfied, it holds true that, for general random matrices \( X \) and \( Y, \)
\[ \frac{1}{N} \mathbb{E} I(\rho) \overset{N \to \infty}{\to} \frac{1}{N} T(\rho), \tag{7} \]
where \( T(\rho) \) is given by
\[ T(\rho) = \log \det(I + \frac{\rho M g g^H}{\delta} R) + \log \det(I + \delta g S) + \log \det(I + g T) - 2M g g^H. \tag{9} \]

Here, \((\delta, g, \Psi)\) is the unique positive solution of the following system of equations
\[ \delta = \frac{1}{L} \text{Tr} R Q R, \quad g = \frac{1}{M} \text{Tr} S Q S, \quad \Psi = \frac{1}{M} T Q T, \tag{10} \]
where
\[ Q_R = \left(\frac{1}{\rho} I + \frac{M g g^H}{L \delta} \right)^{-1}, \quad Q_S = \left(\frac{1}{\delta} I + \Psi \Psi^H \right)^{-1}, \quad Q_T = \left(I + g T \right)^{-1}. \tag{11} \]

Remark 1. This theorem indicates that \( T(\rho) \) is a good approximation for \( \mathbb{E} I(\rho) \). For \( T \) to hold, the random matrices \( X \) and \( Y \) are not necessarily Gaussian where the EMI is normalized by \( N \). From \( \mathbb{E} \), we know that, when \( X \) and \( Y \) are Gaussian, the convergence holds even if we get rid of factor \( \frac{1}{\delta} \), which may not hold true for non-Gaussian matrices as discussed in \cite{25, 27}. For IRS-aided MIMO systems, the matrix \( S \) is related to the phase shifts matrix \( \Psi \), which indicates that the solution of (10) is also related to the phase shifts.

Next, we investigate the fluctuation of \( I(\rho) \) by RMT. The challenge arises from the fact that the effective channel is the product of two random matrices. There are some related results in the literature. The CLT of linear spectral statistics for \( F \)-matrices was given in \cite{20}. In \cite{27}, the CLT of linear spectral statistics with general non-Gaussian entries was given for the product of two i.i.d. random matrices, which was also investigated in \cite{28} by a free probability approach for Gaussian random matrices. The authors of \cite{28} gave the CLT for the MI of double-Rayleigh channels when \( M = N \). However, the CLT for the MI over correlated channels has not been considered and is one of the main contributions of this paper. Based on the CLT, we also investigate the outage probability for IRS-aided MIMO systems using large RMT.

B. Asymptotic Gaussianity of the MI

To characterize the distribution of the MI, we first prove the Gaussianity of the MI.
Theorem 2. (CLT for the MI) If A.1 - A.3 are satisfied, it holds true that

\[
\frac{I(\rho) - T(\rho)}{\sqrt{V(\rho)}} \xrightarrow{N \to \infty} N(0,1),
\]

(12)

where \( T(\rho) \) is given in (9). The asymptotic variance \( V(\rho) \) is given by

\[
V(\rho) = -\log(1 - \gamma_R \Gamma_L) - \log(1 - \gamma_S \gamma_T),
\]

(13)

where \( \gamma_R, \gamma_S, \gamma_T, \) and \( \Gamma_L \) are listed in Table I.

Proof. The proof of Theorem 2 is given in Appendix A.

Remark 2. Theorem 2 indicates the asymptotic Gaussianity of the MI. Note that the variance consists of two terms caused by \( X \) and \( Y \), respectively. Different from the entries-based CLT used in SISO systems [6], [12], this CLT is developed based on RMT. From the proof of Theorem 2, we can observe that the term \(-\gamma_S \eta_T \) in \( \Gamma_L \) can be omitted if \( L \) is large, where \( \psi_T \) is given in Table I. In this case, we can use \( \Gamma \), listed in Table I, to replace \( \Gamma_L \) and the variance is given by

\[
V(\rho) = -\log(1 - \gamma_R \Gamma) - \log(1 - \gamma_S \gamma_T) - \log(1 - \gamma_S \eta_T).
\]

(14)

The large \( L \) case will be used in the following analysis as we will investigate the asymptotic performance when \( L \) goes to infinity.

C. Uncorrelated Cases

To better illustrate the impact of the size of the IRS, we further consider the special case where the channels are i.i.d., i.e., \( R = S = T = I \), which is referred to as the double-Rayleigh model [28]. To derive the CLT, we first determine the EMI.

Proposition 1. If A.1 holds true, the EMI with \( N = M \) is given by

\[
T(\rho) = 2N \log(1+g) + \frac{N}{\tau} \log(1 + \frac{\tau \rho}{(1+g)^2}) - 2Ng, (15)
\]

where \( \tau = \frac{M}{L} \), and \( g \) is determined by the following cubic equation

\[
g^3 + 2g^2 + (1 + \rho \tau - \rho)g - \rho = 0,
\]

(16)

with \( g > 0 \) and \( 1 + (1 - \tau)g > 0 \).

Proof. Proposition 1 can be obtained by Theorem 1 by setting \( R = S = T = I \), so we omit the proof.

Remark 3. In this case, the parameters in Theorem 1 degenerate to one parameter \( g \), which is described by a cubic equation. Thus, the performance only depends on the ratio between the number of transmit antennas (\( M = N \)) and the size of the IRS (\( L \)), i.e., \( \tau \).

Proposition 2. If A.1 holds true, \( N = M \), and \( L \) is comparable with \( M \), the CLT for the MI is given by

\[
\frac{I(\rho) - T(\rho)}{\sqrt{V(\rho)}} \xrightarrow{N \to \infty} N(0,1),
\]

(17)

where

\[
V(\rho) = \log(1+g)^2 - \log(2g^3 + 2g^2),
\]

(18)

and \( g \) is identical to that in Proposition 1.

Proof. The proof of Proposition 2 is given in Appendix B.

Remark 4. Proposition 2 was also obtained in [28] Proposition 3 by a free probability approach while in this paper it is shown as a special case of Theorem 2. The different signs of \( g^2 \) and \( \rho \) originate from the fact that the signs of Cauchy transform and Stieljes transform are opposite.

Proposition 1 and Proposition 2 indicate that the mean and variance of the MI are determined by the root \( g \) of the cubic equation (16), whose coefficients are related to \( \tau \) and SNR \( \rho \).

IV. OUTAGE AND FINITE-SNR DMT ANALYSIS

A. Outage of IRS-aided MIMO Systems

As a direct result of Theorem 2, the approximation of the outage probability is given by the following theorem.

Theorem 3. (A closed-form approximation for the outage probability of IRS-aided MIMO systems) Given a preset transmission rate \( R \), the outage probability can be approximated by

\[
P_{out}(R) \approx \Phi \left( \frac{R - T(\rho)}{\sqrt{V(\rho)}} \right).
\]

(19)

On the other hand, given the outage probability \( P_{out} \), the outage rate can be approximated by \( R \approx T(\rho) + \sqrt{V(\rho)} \Phi^{-1}(p_{out}) \).

Our results are different from [18] in the sense that both \( R_1 \) and \( T_2 \) are not restricted to be an identity matrix, indicating that the proposed method can handle the general correlated channels. In fact, our results are also applicable for double-scattering channels [23], [29].
B. Outage Probability Optimization

Given statistical CSI, the optimization problem for the outage probability can be formulated as

\[
P_1: \min_{\Psi} P_{out}(R), \quad \text{s.t.} \quad \Psi = \text{diag}(\psi_1, \psi_2, \ldots, \psi_L),
\]

where \( \psi_l = \exp(\gamma_l|\theta_l|) \). The challenges arise from two aspects: (1) Evaluation of the objective function; (2) The non-convexity of the problem caused by the uni-modular constraints of \( \psi_l \). The first one can be resolved by approximating the outage probability using \([19]\) in Theorem 1. Given the SNR \( \rho \), the mean \( \mathbf{T}(\rho) \) and variance \( \mathbf{V}(\rho) \) are functions of phase shifts \( \psi_l, l = 1, 2, \ldots, L \), so we use the notations \( \mathbf{T}(\Psi) \) and \( \mathbf{V}(\Psi) \) instead. Thus, \( \mathcal{P}_1 \) can be rewritten as

\[
P_2: \min_{\Psi} G(\Psi) = \Phi \left( \frac{R - \mathbf{T}(\Psi)}{\sqrt{\mathbf{V}(\Psi)}} \right), \quad \text{s.t.} \quad \Psi = \text{diag}(\psi_1, \psi_2, \ldots, \psi_L), |\psi_l| = 1, l = 1, 2, \ldots, L.
\]

Although the parameters \( \delta, g, \) and \( \mathbf{T} \) are coupled with \( \theta_l \) as explained in Remark 1, the partial derivatives of the objective function with respect to \( \theta_l \) can be given in a closed form. Therefore, \( \mathcal{P}_2 \) can be solved using the gradient descent method. Specifically, in each iteration, the update of \( \theta_l \) is obtained by searching on the negative gradient direction, until the value of the objective function converges to a stationary point.

Next, we compute the partial derivatives with respect to \( \theta_l \), \( l = 1, 2, \ldots, L \), and we use the notation \( \gamma'_{\ell} = \frac{\partial \gamma_{\ell}}{\partial \theta_{\ell}} \) to represent the partial derivatives for simplicity. By the chain rule, the partial derivative of \( G(\Psi) \) with respect to \( \theta_l \) is given by

\[
G_l'(\Psi) = \frac{\exp\left(-\frac{\theta^T(\Psi)}{\delta}\right)}{2\pi} T_l'(\Psi),
\]

where

\[
T_l'(\Psi) = \frac{R - \mathbf{T}(\rho)}{\sqrt{\mathbf{V}(\rho)}},
\]

\[
T_l'(\Psi) = \frac{-\mathbf{T}_l'(\rho)\mathbf{V}(\rho) - \frac{1}{\delta}(R - \mathbf{T}(\rho))\mathbf{V}_l'(\rho)}{\mathbf{V}(\rho)^{\frac{3}{2}}},
\]

\[
T_l'(\rho) = \mathbf{G}_l(\mathbf{T}, \mathbf{T}, \mathbf{T}, \mathbf{T})^{-1} \mathbf{F}_l.
\]

\( \mathbf{F}_l \) is defined as

\[
\mathbf{F}_l = \mathbf{T}_l^\dagger (\mathbf{G}_l \otimes \mathbf{R}_2) \mathbf{T}_l^\dagger, \quad l = 1, 2, \ldots, L,
\]

where

\[
\mathbf{G}_l|_{p,q} = \begin{cases} e^{\gamma_l|\theta_l|}, & p = l, \\ -e^{\gamma_l|\theta_l|}, & q = l, \\ 0, & \text{otherwise}. \end{cases}
\]

In fact, \( \mathbf{F}_l = (\mathbf{T}_l^\dagger \mathbf{R}_2 \mathbf{G}_l^H \mathbf{T}_l^\dagger)^{-1} \) and the term \( V_l'(\rho) \) can be given by

\[
V_l'(\rho) = \frac{\gamma_{\ell} \gamma_{\ell} T + \gamma_{\ell} S \gamma_{\ell} T}{\Delta Y} + \frac{\gamma_{\ell} R T + \gamma_{\ell} S \gamma_{\ell} T}{\Delta X}.
\]

where

\[
\Gamma_l = \frac{M - 2\gamma_{\ell} R T}{\Delta X} + \frac{\gamma_{\ell} S \gamma_{\ell} T}{\Delta Y} + 2gT_1 + g^2 T_1,
\]

The derivatives of \( \gamma'_{\ell} \) in (27) are given as

\[
\gamma'_{R\ell} = -\frac{2M R (\delta g' - \delta g)}{\delta},
\]

\[
\gamma'_{T\ell} = -\frac{2gT_1}{\delta},
\]

\[
\gamma'_{S\ell} = -\frac{2\gamma S (\delta g' - \delta g)}{\delta^2} + 2\gamma S (\delta g' - \delta g).
\]

Lemma 1. Given that \( (\delta, g, \mathbf{g}) \) is the positive solution of (19) and \( \mathbf{p}_l = (\delta_{g_l}, g_{g_l}, \mathbf{g}_l)^T \), it holds true that

\[
\mathbf{A} \mathbf{p}_l = \mathbf{q}_l
\]

and \( |\mathbf{A}| > 0 \), where \( \mathbf{A} \) and \( \mathbf{q}_l \) are defined by

\[
\mathbf{A} = \begin{bmatrix} 
\frac{M R}{\delta} & M R \gamma S & M R S \\
-\frac{M R}{\delta} & 1 & \gamma S \\
0 & \gamma T & 1 
\end{bmatrix},
\]

\[
\mathbf{q}_l = \begin{bmatrix} 
g_{g_l} \gamma S (\mathbf{F}_l) 
\end{bmatrix}^T.
\]

Therefore, we have

\[
\mathbf{p}_l = \mathbf{A}^{-1} \mathbf{q}_l.
\]

Proof. The proof can be obtained by taking derivatives of \( (\delta, g, \mathbf{g}) \) with respect to \( \theta_l \) at both sides of (19), and the details are omitted here. We have \( |\mathbf{A}| > 0 \) because

\[
|\mathbf{A}| = z \gamma_R (1 - \gamma_S^2) \gamma_T + \frac{M R \gamma S \gamma T}{\Delta X}
\]

and the two terms on the right-hand side (RHS) are positive, so \( \mathbf{A} \) is invertible.

Next, we provide the gradient descent method. We use the Armijo-Goldstein (AG) line search method \([30]\) to find an expected decrease of the objective function based on the local gradients. The AG condition is given in the 4th line of Algorithm 1. The algorithm can be proved to converge to a stationary point by following the proof of Theorem 1 in \([31]\) or Theorem 1 in \([32]\). The performance of the algorithm will be evaluated in Section VI.

C. Finite SNR DMT for IRS-aided Systems

In this section, we investigate the finite-SNR DMT of IRS-aided systems. DMT was proposed in \([33]\) to characterize the trade-off between diversity and multiplexing gain, which is typically referred to as the asymptotic-SNR DMT. Here, we investigate the finite-SNR DMT \([34], [35]\), which provides more insightful information for the low and moderate SNR regimes.

By the definition of finite-SNR DMT \([35]\), the multiplexing gain is given by

\[
m = \frac{dR}{dI(\rho)} \approx \frac{dR}{I(\rho)}.
\]
**Algorithm 1** Gradient Descent Algorithm for the Phase Shift Matrix $\Psi$

**Input:** $\theta^{(0)}$, initial stepsize $\alpha_0$, scaling factor $0 < c < 1$ and control parameter $0 < \beta < 1$. Set $t = 0$.

1: repeat
2: Compute the gradient
$$\nabla_\theta G(\Psi) = \left( \frac{\partial G(\Psi)}{\partial \theta_1}, \frac{\partial G(\Psi)}{\partial \theta_2}, \ldots, \frac{\partial G(\Psi)}{\partial \theta_L} \right)^T,$$

according to (22) and its direction $d^{(t)} = \frac{\nabla_\theta G(\Psi)}{\|\nabla_\theta G(\Psi)\|}$.
3: $\alpha \leftarrow c\alpha$
4: while $G(\Psi^{(t)}) - G(\text{diag}[\exp(j\theta^{(t)} - \alpha_jd^{(t)})]) < 0$ do
5: $\alpha \leftarrow \alpha_0$
6: end while
7: $\theta^{(t+1)} \leftarrow \theta^{(t)} - \alpha_jd^{(t)}$
8: $\Psi^{(t+1)} \leftarrow \text{diag}[\exp(j\theta^{(t+1)})]$.
9: $t \leftarrow t + 1$
10: until Convergence.

**Output:** $\theta^{(t)}, \Psi^{(t)}$.

where $d = \min(L, M, N)$ and $R$ is the data rate. By approximating the outage probability using the CLT in Theorem 2, we obtain the following theorem regarding the finite-SNR DMT for IRS-aided systems in Rayleigh channels with equal power allocation among different antennas.

**Theorem 4.** The DMT in the finite-SNR regime can be approximated by
$$d(m, \rho) = \frac{z(m - d)}{\sqrt{2\pi}} \exp\left(-\frac{(m - d)^2}{2}\right)H'(z),$$

where $H(z) = \frac{T(z)}{\sqrt{V(z)}}$, $T(z)$ is obtained by replacing $\rho$ in $T(\rho)$ with $z = \rho^{-1}$, and the same holds for $V(z)$.

$H'(z)$ is given as
$$H'(z) = \frac{T'(z)V(z) - \frac{z}{2}T'(z)V'(z)}{V(z)^2},$$

where
$$T'(z) = -\frac{N}{z} + \text{Tr}Q_R,$$
$$V'(\rho) = \gamma_S\gamma_T' + \gamma_S\gamma_T + \gamma_R$$
$$\Gamma' = \frac{M - 2}{L\delta^2} + \gamma_L + \frac{\gamma_j}{\Delta Y},$$
$$\gamma_L' = \frac{2\gamma_T'}{\Delta Y} + \gamma_I' + \frac{1}{\Delta X},$$
$$\gamma_R' = \frac{2\gamma_T'}{\Delta Y} - \frac{M\delta'}{L\delta^2} - \frac{2}{\Delta Y},$$
$$\gamma_I' = \frac{2\gamma_T'}{\Delta Y} - \frac{M\delta'}{L\delta^2} - \frac{2}{\Delta Y},$$
$$\gamma_L = \frac{2\gamma_T'}{\Delta Y} - \frac{M\delta'}{L\delta^2} - \frac{2}{\Delta Y},$$
$$\gamma_I = \frac{2\gamma_T'}{\Delta Y} - \frac{M\delta'}{L\delta^2} - \frac{2}{\Delta Y},$$
$$[\delta', g', \gamma']^T = A^{-1}[\delta\gamma_{R,1}, 0, 0]^T.$$  

A is defined in (30) and other symbols are given in Table I.

**Proof.** The proof of Theorem 4 is given in Appendix C.

**Remark 5.** Similar to (22), the derivatives in (36) are obtained by the chain rule. It can be observed that $H(z)$ plays an important role in the finite-SNR DMT. The expression can be further simplified by an approximation, which is given in the following proposition.

**Proposition 3.** (An approximation for the finite-SNR DMT) The finite-SNR DMT in Theorem 2 can be approximated by
$$d(m, \rho) \approx -\frac{z(m - d)^2}{\sqrt{2\pi}} H'(z)$$

where $H'(z)$ is obtained from (37).

**Remark 6.** This approximation is obtained by the approximation $Q(x) \approx \frac{e^{-x^2}}{\sqrt{2\pi}}$, $x > 0$. From this approximation, we can observe that the finite-SNR DMT is highly related to the ratio between the mean and the standard deviation of the MI, i.e., $H(z)$. This indicates that the effect of the SNR on the finite-SNR DMT is represented through the term $-zH(z)H'(z)$.

V. HOW LARGE THE IRS NEEDS TO BE?

### A. Impact of the Size of IRSs on EMI

We first characterize the relation between the EMI and the size of the IRS. For simplicity, we assume that $M = N$ and focus on the parameter $\tau$, i.e., the ratio between the number of antennas $N$ and the size of the IRS $L$. To study the performance gain obtained by increasing the size of the IRS, we focus on the case when $\tau < 1$, which means that the size of the IRS is larger than the number of antennas at the transmitter and receiver. Results with $\tau > 1$ can be handled similarly. Inspired by the concept of massive MIMO efficiency proposed in [37], we define the concept called the IRS efficiency to characterize the efficiency of increasing the IRS size to achieve higher EMI.

**Definition 1.** Given an SNR $\rho$, the IRS efficiency $\eta \in (0, 1)$ is defined as
$$\eta = \frac{T(\rho)}{T_\infty(\rho)},$$

where $T(\rho)$ denotes the EMI achieved by a given size of the IRS and $T_\infty(\rho)$ represents the maximum throughput with infinitely large IRSs. Specifically, the IRS efficiency $\eta$ represents the percentage of the maximum throughput that is achieved by a given size of IRS.

Given $\eta$, we can determine the minimum size of the IRS needed to achieve at least the fraction $\eta$ of the asymptotic performance. To better illustrate the impact of the IRS size, we consider independent channels and ignore the optimization of phase shifts. According to Proposition 1 for given $\rho$ and $M$, $T(\rho)$ will be determined only by $\tau$. Next, we will investigate $T_\infty(\rho)$.

### B. Asymptotic Mean and Variance for the MI

We first determine $T_\infty(\rho)$ for a given finite SNR $\rho$. The roots of the equation in (16) can be given by Cardano’s formula [28] as
$$g_m = \frac{e^{\frac{\pi}{3}} w(\tau, \rho) + e^{-\frac{2\pi}{3}} w^*(\tau, \rho) - 2}{3}, m = 1, 2, 3,$$
where $\omega(\tau, \rho)$ and $\omega^*(\tau, \rho)$ are given by
\[
\omega(\tau, \rho) = (1 + 9\rho^2 + 9\rho^2)^{1/3},
\]
\[
\omega(\tau, \rho)^* = (1 - 3\rho^2 + 3\rho) / \omega(\tau, \rho).
\]
\[
(\cdot)^\frac{1}{3} \text{ and } (\cdot)^\frac{2}{3} \text{ take any cubic root and square root, respectively.}
\]
It is straightforward to verify that the term under the square root operator is negative when $\tau = 0$ and there are three real roots. In fact, by Vieta’s formulas, the only positive root is the one we desire. When the size of IRS goes to infinity, i.e., $\tau \to 0$, the asymptotic solution of the cubic equation is given by
\[
g_\infty = \frac{2(\text{Re}(\omega(0, \rho)) - 1)}{3},
\]
where $\omega(0, \rho)$ is chosen as the one whose real and imaginary parts are both positive. In this case, the asymptotic EMI with $\tau \to 0$ is given by
\[
T_\infty(\rho) = 2N \log(1 + g_\infty) + \frac{N\rho}{1 + g_\infty^2} - \frac{2Ng_\infty}{1 + g_\infty}.
\]
In fact, the asymptotic mean and variance can be obtained by an easier approach as they are equal to those of the MI for a single-hop MIMO system over independent Rayleigh channels. Details are presented by the following proposition.

**Proposition 4.** When $N = M$, the mean and variance of the MI of a single-hop and independent Rayleigh channel are given by:
\[
\begin{align*}
T_{\text{Rayleigh}}(\rho) &= N \log(\delta + 1 + \rho) - \frac{N\delta}{1 + \delta}, \\
V_{\text{Rayleigh}}(\rho) &= \log(\frac{(1 + \delta)^2}{2\delta + 1}),
\end{align*}
\]
where $\delta$ is the positive solution of $\delta^2 + \delta - \rho = 0$. It holds true that
\[
T_\infty(\rho) = T_{\text{Rayleigh}}(\rho), \quad V_\infty(\rho) = V_{\text{Rayleigh}}(\rho).
\]

**Proof.** The expression in (43) can be obtained as a special case of Proposition 1 in [22]. Then, we have $g_\infty = \delta$ since the cubic equation in (16) degenerates to a quadratic one, i.e., $g(1 + g) = \rho$. Therefore, it follows from (42)
\[
\begin{align*}
T_\infty(\rho) &= N \log[(1 + g_\infty)^2] + \frac{Ng_\infty}{1 + g_\infty^2} - \frac{2Ng_\infty}{1 + g_\infty} \\
&= N \log[1 + g_\infty + g_\infty(1 + g_\infty)] - \frac{Ng_\infty}{1 + g_\infty} \\
&= N \log(1 + \rho + g_\infty) - \frac{Ng_\infty}{1 + g_\infty} \\
&= T_{\text{Rayleigh}}(\rho).
\end{align*}
\]
With similar manipulations, the asymptotic variance can be written as
\[
V_\infty(\rho) = \log[1 + g_\infty^2] - \log(\rho + 2\rho g_\infty) \\
= \log\left(\frac{1 + g_\infty^2}{2g_\infty + 1}\right) = V_{\text{Rayleigh}}(\rho).
\]

**Remark 7.** (44) indicates that the maximum throughput and the asymptotic variance of the MI with the two-hop IRS channel are equal to those of a single-hop channel.

Based on (42), we can determine how large the IRS needs to be to achieve a certain performance of the asymptotic performance. The results for the general correlated cases can be derived similarly by Theorem 1 but with an implicit expression for EMI with respect to the size of the IRS. The variance of the MI, the outage probability, and the finite-SNR DMT can be handled similarly according to (41) and Proposition 2.

In the above derivation, we assume that the transmitter and receiver have the same number of antennas. The derivations for unequal cases can also be performed in similar ways. Furthermore, the correlated case can also be investigated numerically using Theorem 1 and Theorem 2.

### C. The Impact of the IRS Size in the High SNR Regime

At high SNRs, the EMI and outage probability can be approximated as a more explicit function of $\tau$ compared with the cubic root form in (39), which is presented by the following theorem.

**Theorem 5.** (High SNR case) When $\tau < 1$ and $\rho \gg 1$, the only positive solution of the cubic equation in (16) can be approximated as
\[
g = \sqrt{(1 - \tau)^3} + \left[\frac{1}{2(1 - \tau)} - 1\right] + o(1)
\]
\[
= a\rho^\frac{3}{4} + b + o(1).
\]
When $N$ is fixed and $L \ll \sqrt{\rho}$, the mean and variance of the MI can be approximated by
\[
T(\rho) = N \log(\rho) - 2 - (\frac{1}{\tau} - 1) \log(1 - \tau) + \frac{2a}{\rho^{1/4}}
\]
\[
+ o(\rho^{-1/4}),
\]
and
\[
V(\rho) = \frac{1}{2} \log(\frac{\rho}{3a^2}) + \frac{2a^2(1 - b) - 1}{2a^3} \rho^{-1/4} + o(\rho^{-1/4}),
\]
respectively. Furthermore, when $N$ is fixed and $L$ is comparable to $\sqrt{\rho}$ or far larger than $\sqrt{\rho}$, the mean and variance of the MI can be approximated by
\[
T(\rho) = N \log(\frac{2}{\tau}) - \frac{2N}{L} + 2\rho^{-1/4} + o(\rho^{-1/4}),
\]
\[
V(\rho) = \frac{1}{2} \log(\frac{\rho}{4}) + \frac{N}{2L} + \rho^{-1/4} + o(\rho^{-1/4}).
\]

The outage probability can be further approximated by $P_{\text{out}}(R) \approx \Phi\left(\frac{R - T(\rho)}{V(\rho)}\right)$.

Based on Theorem 5, we have the following remarks.

**Remark 8.** **Asymptotic performance:** At high SNRs, as $L$ goes to infinity, $T(\rho)$ and $V(\rho)$ converge to an asymptotic value as shown in (29), which is dominated by the term $\log(\rho)$.

**Remark 9.** **Fast convergence:** $T(\rho)$ and $V(\rho)$ converges fast to the asymptotic value. When $L$ is large, the rate of
convergence to the asymptotic performance is $O\left(\frac{1}{L}\right)$. This indicates that the gain of EMI achieved by increasing the size of the IRS decreases when the IRS size is getting larger.

**Remark 10. SNR vs. size of IRS:** We can observe from (49) that the impact of the IRS size is very limited compared with that of the SNR as the mean and variance are dominated by the SNR.

**Remark 11. Comparison with the single-hop Rayleigh channel:** At high SNRs, the mean and variance of the MI for a single-hop i.i.d. Rayleigh channel is given in [35]:

$$
T(\rho) = N[\log\left(\frac{\rho}{e}\right) + 2\rho^{-\frac{1}{2}}] + o(\rho^{-\frac{1}{2}}),
$$

$$
V(\rho) = \frac{1}{2} \log\left(\frac{\rho}{4}\right) + \rho^{-\frac{1}{2}} + o(\rho^{-\frac{1}{2}}),
$$

(50)

Comparing the results in (50) and (49), we can observe that the difference induced by the IRS is of order $O\left(\frac{1}{L}\right)$, which indicates that these two results converge as the IRS size goes to infinity.

The analysis in this section revealed the effect of increasing the size of the IRS, which provides guidance on the design of practical systems. Numerical results in the next section will further validate the observations.

**VI. Numerical Results**

In this section, simulations are performed to validate the theoretical results. Here we adopt the exponential correlation model. Specifically, the entries of $R_i$, $T_i$, $i = 1, 2$, can be represented by $[C(\mu)]_{i,j} = \mu^{[i-j]}$. The unit of the preset transmission rate $R$ is bits/s/Hz.

**A. Outage Probability Approximation**

First, we consider the case when there is no correlation at the transmitter side, for which both the proposed method and the method based on Mellin transform [18] can be applied. The parameters are set to be $M = L = N = 3$, correlation matrices $R_1 = R_2 = T_1 = C(\mu)$ with $\mu = 0.5$, and $T_2 = I$. The phase shift matrix is set to be $\Psi = I$.

The Monte-Carlo simulation values are computed over $10^8$ realizations, and we use the variance in (13) for the small $L$. It can be observed from Fig. 1 that the performance of the proposed method is almost the same as the one based on Mellin transform but with a simpler expression. The results also indicate that the derived analytical results are accurate for a small number of antennas although it was derived in the asymptotic regime.

In Fig. 2 we consider the correlated case with $T_2 = C(0.5)$ and plot the outage probability with respect to the rate threshold when $L = 3, 16, 32$, respectively. Markers represent the simulation results. It can be observed that the proposed method works well when there are correlations at both the transmitter and receiver sides. Note that the Mellin transform-based method cannot handle this case.

![Fig. 1. Performance comparison with the Mellin transform-based method in [18].](image)

![Fig. 2. General correlated and rank-deficient cases.](image)

![Fig. 3. Effectiveness of the optimization algorithm.](image)

**B. Effectiveness of the Proposed Optimization Algorithm**

Fig. 3 verifies the effectiveness of Algorithm 1 in minimizing the outage probability when $M = N = 4$ and $\mu = 0.8$ is used for all correlation matrices. The initial phase shifts are set to be $\psi_i = e^{\frac{(2\pi i}{L} \cdot \alpha}$, $i = 1, 2, ..., L$. The initial step size is set as $\alpha_0 = 0.0005$ and the scale factor $\beta = 0.5$. Simulation results are denoted by the markers. It can be observed that the outage probability is efficiently improved by Algorithm 1.

**C. Finite SNR DMT**

Fig. 4 validates the accuracy of the closed-form expression for the finite SNR DMT as shown in (34). The parameters are set as $N = M = 4$, $L = 2$, $\mu = 0.5$, and $\Psi = I$. It can be observed that the closed-form expression is very accurate as validated by simulation results (plotted by markers). Furthermore, the finite-SNR DMT is more accurate at low SNRs compared with the asymptotic SNR results in [38].
D. Impact of the Size of the IRS

Asymptotic Performance and Convergence: In Fig. 5 we show how the EMI changes when the IRS size increases. The parameters are set as \( M = N = 20 \) with independent channels. The range of the size of IRS is from 4 to 100. The asymptotic EMI in (42) and the simulation results are plotted with dash-dotted lines and circular markers, respectively. It can be observed that EMI increases and approaches the asymptotic performance as the size of the IRS becomes larger. However, the increasing rate of EMI decreases, which means that the performance gain by the increment of the IRS size varnishes for larger IRS. This agrees with Remarks 8 and 9.

Similar comparisons are performed for the variance of the MI with \( M = N = 4 \) and \( R = 4 \). It can be observed from Fig. 6 that the variance decreases and approaches the asymptotic one when the size of the IRS becomes larger and the approximations for the variance are accurate.

IRS Efficiency: In Fig. 7 the size of the IRS needed to achieve \( \eta \) of the asymptotic performance is plotted when \( M = N = 20 \). It is worth noting that the required size for \( \eta = 0.95 \) is almost twice of that for \( \eta = 0.9 \), which means that the 5% performance increment requires a huge increase in the size of the IRS. This agrees with Remark 9. In practice, we need to decide whether it is worth the cost of deploying larger IRSs to obtain the additional performance.

High SNR Regime: In Figs. 8 and 9, the approximation performance of Theorem 5 for the outage probability and EMI in the high SNR regime is validated with \( M = N = 4 \), respectively. The range of the IRS size is from 8 to 256 and the SNR is set to be 20 dB in Fig. 8. It can be observed that the approximation in (48) is accurate in the high SNR regime.

SNR vs. IRS Size: We can observe from Figs. 5, 8 and 9 that, when the size of the IRS increases, the performance (outage probability and EMI) improves but the rate of change decreases, leading to the convergence to the asymptotic value. Furthermore, the EMI gain caused by the increment of SNR is far larger than that caused by the increment of the size of IRS, which agrees with Remark 10.

VII. Conclusion

In this paper, we investigated the outage probability and finite-SNR DMT of IRS-aided MIMO communication systems assuming statistical CSI. By leveraging RMT, the CLT of the MI was derived and the outage probability was approximated in a closed-form. A gradient descent algorithm was proposed to minimize the outage probability by optimizing the phase shifts of the IRS, which was shown to be efficient by numerical results. To better characterize the trade-off between the bit error rate and the throughput in an operational SNR regime, a closed-form finite SNR DMT was presented. Finally, the impact of the IRS size on the EMI and outage probability was studied. Both theoretical and simulation results showed that
larger IRSs provide better performance but the gain saturates quickly as the size of IRS increases. For example, to improve the throughput from 90% to 95% of its maximum value, a huge hardware cost (IRS size) is required. The results in this work not only provided accurate characterization for the MI and outage probability of IRS-aided systems, which is not available in the literature, but also revealed the impact of the IRS size on system performance, offering valuable guidance for practical system deployment.

APPENDIX A
Proof of Theorem 2

Proof. As the channel matrix $\mathbf{H}$ in (6) is the product of two random matrices, i.e., $\mathbf{X}$ and $\mathbf{Y}$, we use the martingale method to show the CLT [26]. First, we rewrite the process $I(\rho) - E(I(\rho))$ as a summation of two processes:

$$\{I(\rho) - E(I(\rho))\} = \sum_{i=1}^{M} \frac{1}{M} \left( \sum_{j \neq i} t_{ij} \rho_{ij} + \frac{1}{M} \rho \right) \tilde{y}_{ij}.$$

Then, we show that the two processes are asymptotically Gaussian so that $I(\rho) - E(I(\rho))$ will be Gaussian. In the following, we first provide the outline of the proof in steps and then give the details for each step.

Step 1: Consider the first process $I(\rho) - E(I(\rho)|Y)$ given $Y = \{all \ Y\}$ and show that it converges to a Gaussian process. Furthermore, it is proved that the asymptotic mean and variance are independent of the condition $Y$, which indicates that the asymptotic distribution of the first process is independent of that for the second process $E(I(\rho)|Y) - E(I(\rho))$ because its mean and variance are deterministic.

Step 2: Show the asymptotic distribution of $E(I(\rho)|Y) - E(I(\rho))$ is Gaussian.

Case 3: The sum of two independent Gaussian random processes will result in a Gaussian random process and we will compute the expression of the asymptotic variance in this step.

Step 1: Denote $\mathbf{W} = \tilde{S}^T \tilde{Y} \tilde{T}^T$ and it follows from (6) that the channel matrix can be given by $\mathbf{H} = \mathbf{R} \tilde{X} \mathbf{W}$.

$$\|\mathbf{W} \tilde{\mathbf{W}}\| \leq \|\tilde{S}\|\|\mathbf{T}\|(1 + \sqrt{\frac{1}{M}}) < \infty \quad \lim \inf \frac{1}{M} \text{Tr} \mathbf{Y} \mathbf{Y}^H > 0$$

and the condition of Theorem 1 in [22] is satisfied with probability one. Therefore, the condition of Theorem 1 in [22] is satisfied with probability one, from which we can conclude that $I(\rho) - E(I(\rho)|Y)$ converges to a zero-mean Gaussian process and the asymptotic variance $V_1$ is given by

$$V_1 = \frac{\text{Var}(I(\rho)|Y)}{M} \xrightarrow{\text{a.s.}} \frac{1}{M} \text{Tr} \mathbf{Q}_f.$$
\[
\tilde{\gamma}_\rho = \frac{1}{L} \text{Tr} \mathbf{W}^H \mathbf{C} \mathbf{WW}^H \mathbf{C} = \frac{1}{L} \sum_{i=1}^{L} (1 + t_i y_i^H \mathbf{S}^+ \tilde{\mathbf{C}_i} \mathbf{WW}^H \mathbf{C} \mathbf{S}^+ y_i) = \frac{1}{L} \sum_{i=1}^{L} \frac{t_i y_i^H \mathbf{S}^+ \tilde{\mathbf{C}_i} \mathbf{WW}^H \mathbf{C} \mathbf{S}^+ y_i + o(1)}{1 + t_i L \text{Tr} \mathbf{SC}} + \frac{t_i^2 y_i^H \mathbf{S}^+ \tilde{\mathbf{C}_i} \mathbf{WW}^H \mathbf{C} \mathbf{S}^+ y_i + o(1)}{(1 + t_i L \text{Tr} \mathbf{SC})^2} + \frac{t_i^3 (\frac{1}{L} \text{Tr} \mathbf{SC})^2}{(1 + \frac{t_i}{L \text{Tr} \mathbf{SC}})^2} + o(1) = Y_1 + Y_2 + o(1).
\]

Then by letting \(\alpha = 1\) and \(l = 0\), we have \(h = g\) and \(\tilde{h} = \tilde{g}\).

Taking the derivative of \(h\) with respect to \(l\), we have
\[
\frac{d}{dl} S(\rho L + \mathbf{D} + \mathbf{f} \mathbf{W}^H)^{-1} = \frac{1}{M} \text{Tr} S \left( \delta^{-1} (\mathbf{I} + \tilde{\mathbf{h}}) S^{-1} \delta^{-1} (\mathbf{I} + \tilde{\mathbf{h}}) S \right)^{-1} \frac{d\tilde{h}}{dl} + \frac{1}{M} \text{Tr} S \left( \delta^{-1} (\mathbf{I} + \tilde{\mathbf{h}}) S \right)^{-1} \mathbf{D} \left( \delta^{-1} (\mathbf{I} + \tilde{\mathbf{h}}) S \right)^{-1},
\]

Therefore, we have the following system of equations
\[
\begin{bmatrix}
1 \\
\gamma_T \\
\gamma_S
\end{bmatrix}
\begin{bmatrix}
\frac{d\gamma_T}{dl} \\
\frac{d\gamma_S}{dl}
\end{bmatrix}
= \begin{bmatrix}
\frac{2 \gamma(S)}{\delta^2} \\
\delta
\end{bmatrix} + o(1).
\]

According to (58) and letting \(\mathbf{D} = \mathbf{S}\), we have
\[
\frac{1}{M} \text{Tr} \mathbf{SC} \mathbf{SC} = \frac{1}{M} \frac{d\gamma_S}{dl} \bigg|_{l=0} - \frac{1}{\delta^2} \frac{d\gamma_S}{dl}.
\]

Hence, we have
\[
\gamma_T = \frac{M \gamma_S (\gamma_T^2 - \frac{1}{M \gamma^\rho})}{L \delta^2 \Delta_Y}, \quad \gamma_S = \frac{M \gamma_T g^2}{L \delta^2}.
\]

where \(\gamma^\rho\) is given in Table II. By applying the continuous mapping theorem [40] to logarithm function, the asymptotic variance is given by
\[
Y_1 \xrightarrow{p_{M \rightarrow \infty}} - \log\left[1 - \frac{M \gamma^R (\gamma_T^2 - \frac{1}{M \gamma^\rho})}{L \delta^2} + \frac{\gamma^T g^2}{\Delta_Y}\right] = - \log(1 - \gamma^R \Gamma L),
\]

which does not depend on \(Y\). Given that the term \(-\frac{2 \gamma^R (\gamma_T^2 - \frac{1}{M \gamma^\rho})}{L \delta^2} \Delta_Y\) will vanish as \(L\) goes to infinity, we have \(\Gamma L \xrightarrow{M \rightarrow \infty} \Gamma\). For a better approximation of the variance, we will use \(\Gamma L\) for the cases with small \(L\) and \(\Gamma\) for the cases with large \(L\), respectively.

Similar analysis can be performed on the asymptotic mean to show that the asymptotic mean also does not depend on \(Y\), which indicates that the asymptotic distribution of the first process is independent of that for the second process as the asymptotic mean and variance are deterministic.

Step 2: In this step, we will show the asymptotic Gaussianity of \(\mathbb{E}(I(\rho) Y) - \mathbb{E}(I(\rho))\) using the martingale method (CLT for martingales) in [40]. Let \(Z = \mathbf{R}^2 \mathbf{X} \mathbf{S}^+\), then \(H = ZY^T \mathbf{S}^+\).

By the bookkeeping approach in [25] and [41], we have
\[
\mathbb{E}(I(\rho) Y) - \mathbb{E}(I(\rho))Y = \sum_{m=1}^{M} (E_m - E_{m-1}) \log(1 + \Lambda_m),
\]

where
\[
\Lambda_m = \frac{\mathbf{y}_m^H \mathbf{Z}^H \mathbf{Q}_m \mathbf{Z} \mathbf{y}_m - \frac{L}{M} \text{Tr} \mathbf{Z}^H \mathbf{Z} \mathbf{Q}_m}{1 + \frac{t_i}{M} \text{Tr} \mathbf{Z} \mathbf{Z}^H \mathbf{Q}_m} + \frac{t^2_m}{M^2} \text{Tr} \mathbf{Z} \mathbf{Z}^H \mathbf{Q}_m.
\]

Here \(E_m(\cdot) = \mathbb{E}(\cdot | \mathbf{y}_m, \mathbf{y}_{m+1}, ..., \mathbf{y}_M)\) and \(E_{m+1}(\cdot) = \mathbb{E}(\cdot)\).

We skip the verification of the Lyapunov’s condition and turn to the computation of the asymptotic variance for the second process, \(V_2\). Following a similar method as in [25], [41], we have
\[
V_2 \xrightarrow{p_{M \rightarrow \infty}} \sum_{m=1}^{M} (\mathbb{E}(e_m))^2
\]

where step \((a)\) follows from the rank-one lemma [25] Lemma 3.1). By using Lemma 2 in Appendix D twice and similar computations in [25], we have
\[
\mathbb{E}(1 - \frac{L}{M} \text{Tr} \mathbf{Z} \mathbf{Z}^H \mathbf{Q}_m) \rightarrow \frac{1}{M} \text{Tr} \mathbb{E}(\mathbf{Z}^H \mathbf{Q}_m) \rightarrow \frac{1}{L} \text{Tr} \mathbb{E}(\mathbf{X} \mathbf{Q}_m) \rightarrow \frac{1}{L} \text{Tr} \mathbb{E}(\mathbf{X} \mathbf{R}_m) = \frac{1}{L} \text{Tr} \mathbb{E}(\mathbf{X} \mathbf{R}_m) = \left(\frac{1}{L} \text{Tr} \mathbb{E}(\mathbf{X} \mathbf{R}_m)\right)^{-1} \mathbf{I} + \mathbf{SY}^T \mathbf{Y}^{-1} -1
\]

\[
\text{and convergence follows}
\]

\[
\mathbb{E}(\text{LHS} - \text{RHS}) \leq \frac{K}{M^2} \mathbb{E}(\mathbf{Y}^H \mathbf{Y}) = O(M^{-2})
\]

according to Lemma 2 in [42] and Markov’s inequality, where LHS denotes the left-hand side of step \((a)\). Similarly,
Step (b) follows by taking $B = E_m\left(\frac{1}{L} Tr EXRQ\right)^{-1}\mathbf{I} + SYTY^H)^{-1}S$. Therefore, we have

$$\frac{1}{M} E_m Tr E_m\left((\frac{1}{L} Tr EXRQ)^{-1}\mathbf{I} + SYTY^H)^{-1}\right)S((\frac{1}{L} Tr EXRQ)^{-1}\mathbf{I} + SYTY^H)^{-1}S \xrightarrow{P(a)} \frac{1}{M} E_m Tr E_m((\delta^{-1}\mathbf{I} + SYTY^H)^{-1})S \times (\delta^{-1}\mathbf{I} + SYTY^H)^{-1}S \xrightarrow{P(b)} \frac{1}{M} E E_m[\mathbf{C}(\delta^{-1})]\mathbf{S}(\delta^{-1}) \xrightarrow{\Delta} \Omega_m,$$

(70)

Steps (a), (b) follow from $\text{Var}(\text{Tr} RQ) = O(1)$ and $\text{Var}(\text{Tr} SC(\delta^{-1})) = O(1)$, respectively, which can be obtained by the variance control using Poincaré-Nash inequality [22, 29]. According to [68] and [25] Section 5, it holds true that

$$\Omega_m = \frac{\gamma}{1 - \gamma \gamma'} + o(1),$$

(71)

where $\gamma'(m) = \frac{1}{M} \sum_{l=1}^{m} |Q_l|^2 |r_l|^2$. Also, we have $\gamma'(m) - \gamma'(m-1) = \frac{1}{M} \sum |Q|^2 |r|^2$. Therefore,

$$V_2 \xrightarrow{P} \sum_{m=1}^{M} E_{m-1}(E_m A_m)^2 \xrightarrow{P} \sum_{m=1}^{M} \frac{\gamma}{1 - \gamma \gamma'} M \rightarrow \infty,$$

(72)

$$-\log(1 - \gamma \gamma') = -\log \Delta_Y.$$

Step 3: Since $I(\rho) - E(I(\rho)Y)$ and $E(I(\rho)Y) - EI(\rho)$ are asymptotically independent Gaussian random processes, the asymptotic distribution of their sum is Gaussian and the variance is given by

$$V = V_1 + V_2.$$

(73)

This completes the proof.

**APPENDIX B**

**PROOF OF PROPOSITION [2]**

Proof. As this proposition is the special case of Theorem 2 in which the Gaussianity has been proved, we only need to compute the variance. By (10), we have

$$\mathbf{g} = \frac{1}{g + 1}, \frac{g}{\gamma}, \gamma = \gamma^2;$$

(74)

and

$$\gamma_R = \frac{\gamma^2}{\gamma}, \gamma_S = \gamma g^2, \gamma_T = \gamma^2.$$  

(75)

Then, the variance can be written as

$$V(\rho) = -\log[(1 + g)^2 - \gamma g^2 - \gamma g^2 g^2 - g^2(1 - \gamma g^2 g^2)] - \log(\gamma^2) = -\log(T_1) - \log(T_2),$$

(76)

where

$$T_1 = 1 + 2g - \gamma g^2 + \gamma^2 (g - 1)^2$$

(77)

and (a) follows from $g^3 + 2g^2 + (1 + \tau/z - 1/z)g - 1/z = 0$. The conclusion follows by noticing that $T_2 = (1 + g)^2 - 2$. $\square$

**APPENDIX C**

**PROOF OF THEOREM 4**

Proof. By the definition of multiplexing gain in [35], we can obtain the rate $R$ as,

$$R = \frac{m}{d} \tilde{T},$$

(78)

For the ease of manipulations, we use $\tilde{T}(z)$ and $V(z)$, which are obtained by letting $z = \frac{1}{\tau}$ in $T(\rho)$ and $V(\rho)$, respectively. Then, by the definition of finite-SNR DMT in [35], we have

$$d(m, \rho) = -\frac{\partial \log(P_{out}(m-d)(z))}{\partial \log(\rho)} = \frac{\partial \log(P_{out}(m-d)(z))}{\partial \rho} = \frac{z(m - d)}{2\sqrt{\pi}} \Phi^{-1}(\frac{m-d}{2d}) H'(z).$$

(79)

$H'(z)$ can be obtained by the chain rule. $\square$

**APPENDIX D**

**COMPUTATIONAL RESULTS**

The following lemma is used in the derivation of [69].

**Lemma 2.** Let $B$ and $D$ be deterministic matrices with bounded norm. For $Q = (MDZ^H + I)^{-1}$, where $Z = R^H X S^H$ is defined in Appendix A, it holds true that

$$\frac{1}{M} E Tr Z^H Q Z B = \frac{1}{M} E Tr S B$$

(80)

$$\times (\frac{1}{L} E E R Q^{-1} I + S D)^{-1} + O(\frac{\|D\|^2}{M}).$$

Proof. Here we omit the condition on $Y$. By the integration-by-parts formula [22], we have

$$E[Z^H Q Z B]_{i,j} = \frac{1}{L} E [S^H R S B]_{i,j} = \frac{1}{L} E [E_{n,m}[R^H X S Z B]_{i,j} + \frac{1}{L} E [E_{n,m}[S D Z^H R S Z B]_{i,j}]$$

(81)

Therefore, by moving the first term of the RHS to the LHS, we have

$$\frac{1}{M} E Tr Z Z^H Q Z B = \frac{1}{M} E Tr ((\frac{1}{L} E E R Q)^{-1} I + S D)^{-1} S B + \varepsilon_z,$$

(82)

where

$$|\varepsilon_z| = \frac{1}{M} E E Tr (\frac{1}{L} E Tr R Q)^{-1} I + S D)^{-1} S B + \varepsilon_z.$$

(83)

$$\leq \frac{K}{M} \text{Var}^2(Tr(RQ)) = O(\frac{\|D\|^2}{M}),$$

(84)
with $K$ being a constant. The conclusion follows from the variance control in Appendix B of [42].

REFERENCES

[1] Q. Wu and R. Zhang, “Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network,” *IEEE Wireless Commun. Mag.*, vol. 58, no. 1, pp. 106–112, Jan. 2019.

[2] X. Yu, V. Jamali, D. Xu, D. W. K. Ng, and R. Schober, “Smart and reconfigurable wireless communications: From IRS modeling to algorithm design,” *IEEE Wireless Commun. Mag.*, to appear.

[3] B. Matthiesen, E. Björnson, E. De Carvalho, and P. Popovski, “Intelligent reflecting surface operation under predictable receiver mobility: A continuous time propagation model,” *IEEE Wireless Commun. Lett.*, vol. 10, no. 2, pp. 216–220, Feb. 2020.

[4] Z. Wang, L. Liu, S. Zhang, and S. Cui, “Massive MIMO communication with intelligent reflecting surface,” *arXiv preprint arXiv:2107.04255*, 2021.

[5] D. Xu, V. Jamali, X. Yu, D. W. K. Ng, and R. Schober, “Optimal resource allocation design for large IRS-assisted SWIPT systems: A scalable optimization framework,” *arXiv preprint arXiv:2104.03546*, 2021.

[6] Y. Cheng, K. H. Li, Y. Liu, K. C. Teh, and H. V. Poor, “Downlink and uplink intelligent reflecting surface aided networks: NOMA and OMA,” *IEEE Trans. Wireless Commun.*, vol. 20, no. 6, pp. 3988–4000, June 2021.

[7] W. Mei and R. Zhang, “Performance analysis and user association optimization for wireless network aided by multiple intelligent reflecting surfaces,” *IEEE Trans. Commun.*, vol. 9, no. 9, pp. 6296–6312, Sep. 2021.

[8] Z. Zhang, Y. Cui, F. Yang, and L. Ding, “Analysis and optimization of outage probability in multi-intelligent reflecting surface-assisted systems,” *arXiv preprint arXiv:1909.02193*, 2019.

[9] Y. Han, W. Tang, S. Jin, C.-K. Wen, and X. Ma, “Large intelligent surface-assisted wireless communication exploiting statistical CSI,” *IEEE Trans. Veh. Technol.*, vol. 68, no. 8, pp. 8238–8242, Sep. 2019.

[10] T. Van Chien, L. T. Tu, S. Chatzinotas, and B. Ottersten, “Coverage probability and ergodic capacity of intelligent reflecting surface-enhanced communication systems,” *IEEE Commun. Lett.*, vol. 25, no. 1, pp. 69–73, Jan. 2020.

[11] J. Zhang, J. Liu, S. Ma, C.-K. Wen, and S. Jin, “Large system achievable rate analysis of RIS-assisted MIMO wireless communication with statistical CSI,” *IEEE Trans. Wireless Commun.*, vol. 20, no. 9, Sep. 2021.

[12] J. Wang, W. Zhang, X. Bao, T. Song, and C. Pan, “Outage analysis for intelligent reflecting surface assisted vehicular communication networks,” in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Taipei, Taiwan, Dec. 2020, pp. 1–6.

[13] S. Atapattu, R. Fan, P. Dharmawansa, G. Wang, J. Evans, and T. A. Tsiftsis, “Reconfigurable intelligent surface assisted two-way communications: Performance analysis and optimization,” *IEEE Trans. Commun.*, vol. 68, no. 10, pp. 6552–6567, 2020.

[14] S. Lin, B. Zheng, G. C. Alexandropoulos, M. Wen, M. Di Renzo, and F. Chen, “Reconfigurable intelligent surfaces with reflection pattern modulation: Beamforming design and performance analysis,” *IEEE Trans. Wireless Commun.*, vol. 20, no. 2, pp. 741–754, Feb. 2020.

[15] C. Guo, Y. Cui, F. Yang, and L. Ding, “Outage probability analysis and minimization in intelligent reflecting surface-assisted MISO systems,” *IEEE Commun. Lett.*, vol. 24, no. 7, pp. 1563–1567, Jul. 2020.

[16] G. Zhou, C. Pan, H. Ren, K. Wang, and A. Nallanathan, “A framework of robust transmission design for IRS-aided MISO communications with imperfect cascaded channels,” *IEEE Trans. Signal Process.*, vol. 68, pp. 5092–5106, Aug. 2020.

[17] S. Hong, C. Pan, H. Ren, K. Wang, K. K. Chai, and A. Nallanathan, “Robust transmission design for intelligent reflecting surface-aided secure communication systems with imperfect cascaded CSI,” *IEEE Trans. Commun.*, vol. 20, no. 4, pp. 2487–2501, Apr. 2020.

[18] Z. Shi, H. Wang, Y. Fu, G. Yang, S. Ma, and F. Gao, “Outage analysis of reconfigurable intelligent surface aided MIMO communications with statistical CSI,” *IEEE Trans. Wireless Commun.*, to appear.

[19] R. Couillet, M. Debah, and J. W. Silverstein, “A deterministic equivalent for the analysis of correlated MIMO multiple access channels,” *IEEE Trans. Inf. Theory*, vol. 57, no. 6, pp. 3493–3514, Jan. 2011.

[20] C.-K. Wen, G. Pan, K.-K. Wong, M. Guo, and J.-C. Chen, “A deterministic equivalent for the analysis of non-Gaussian correlated MIMO multiple access channels,” *IEEE Trans. Inf. Theory*, vol. 59, no. 1, pp. 329–352, Sep. 2012.