

RIS-Assisted Massive MIMO Systems: Performance Limits and Realizations

Chun-Hung Liu

Dept. of Electrical & Computer Engineering
Mississippi State University, MS, USA
E-mail: chliu@ece.msstate.edu

Ngwe Thawdar

Air Force Research Laboratory
Information Directorate, Rome NY
E-mail: ngwe.thawdar@us.af.mil

Jwo-Yuh Wu

Institute of Communications Engineering
NYCU, Hsinchu, Taiwan
E-mail: jywu@nycu.edu.tw

Abstract—This paper considers a multi-antenna communication system that employs reconfigurable intelligent surfaces (RISs) to assist its downlink transmissions from a base station (BS) to a user. It aims to thoroughly investigate the fundamental performance limits of a massive multiple-input and multiple-output (MIMO) communication system and how to realize them. We first devise an equivalent channel model for RIS-assisted multi-antenna communication and analyze its controllable statistical properties that characterize how the RIS-assisted channel gain can be manipulated by controlling an RIS in order to achieve non-coherent and coherent signal combining at the user. The outage probabilities of the RIS-assisted channel model for the non-coherent and coherent combining situations are derived, and their asymptotic upper and lower limits for RIS-assisted massive MIMO systems are also found. Our analyses crucially show why coherent combining may only sometimes outperform non-coherent combining and how much the performance limits of channel outage can be achieved for a large number of antennas at a BS and reflecting elements on an RIS.

Index Terms—Reconfigurable Intelligent Surface, Massive MIMO, Random Matrix Theory, Outage Probability.

I. INTRODUCTION

Efficiently utilizing the radio spectrum is an imperative goal for future wireless networks that widely and seamlessly cover various mobile devices distributed over a large territory. Achieving this goal is a big challenge due to the uncontrollable random nature of wireless channels, making transmitted signals hard to detect anytime and anywhere. Such a challenge becomes even more severe when different kinds of mobile devices using multiple radio access technologies coexist in a wireless network. The recent meta-material technology of developing reconfigurable intelligent surfaces (RISs) provides a feasible solution to alleviating the uncontrollability of wireless propagation environments such that the random characteristics of wireless channels are no longer completely uncontrollable, enabling wireless networks to achieve broader and more uniform coverage [1], [2]. Nonetheless, how to effectively leverage signals reflected by RIS to enhance/assist multi-antenna transmission in a wireless network is a challenge in that how the fundamental performances of an RIS-assisted multi-antenna system relate to the system's settings, parameters, and controls is still unclear. As a result, there is a pressing need to exploit the performance limits of a communication system wherein multiple antennas and RISs can be jointly manipulated.

Prior works on the performance analysis and optimization of an RIS-assisted communication system mainly focused on

how to formulate an optimization problem that can jointly optimize the beamforming vector and the phases of an RIS (typically see [3]–[11]). For instance, references [3]–[6] studied the outage and coverage performances of an RIS-assisted communication system by mainly formulating an optimization problem of minimizing the outage probability or maximizing the coverage probability. Since they do not consider the massive MIMO model in their system, their analytical findings may not be able to generalize the massive MIMO scenario. The more recent works, e.g., [7]–[11], indeed studied the performances of a massive MIMO system equipped with RISs, yet their system models are more specific for some communication environments or scenarios, such as particular antenna setting and spatially-correlated channel models, wherefore their fundamental performance limits would be hard to be characterized due to complex modeling and intractability in analysis. Although much prior work has been on the performance evaluation of an RIS-assisted multi-antenna system, most current analytical outcomes in the literature were found in a communication system under some specific assumptions. Accordingly, one may be unable to employ them to exploit the performance limits of using massive MIMO and RIS in a communication system. In other words, they cannot fundamentally characterize the performance limits of an RIS-assisted communication system equipped with a large antenna array and a large RIS consisting of many reflecting elements.

This paper provides a feasible approach to evaluating the performance limits of an RIS-assisted massive MIMO system in terms of outage probability. Its first contribution is to devise an equivalent channel model in a downlink communication system equipped with multiple antennas and an RIS, which can integrate the signals from a direct link and an RIS-assisted reflecting link at a user so as to model and analyze the controllable statistical properties of an integrated RIS-assisted channel gain. Afterward, these statistical properties were employed to tractably analyze the distribution of the signal-to-noise-power ratio (SNR) for the equivalent channel model, thereby obtaining the low-complexity expressions of outage probability for the equivalent channel model in different non-coherent and coherent signal combining situations at the user, which is the second contribution. Our third contribution is to generalize the analytical outcomes of the outage probability to the massive MIMO scenario and find how the upper and lower limits of

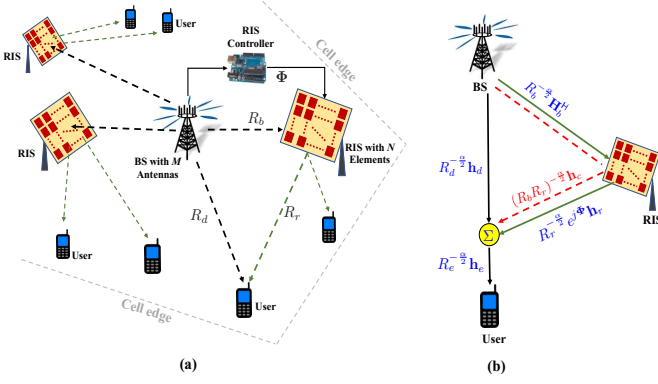


Fig. 1. (a) A RIS-assisted multi-antenna communication system: it consists of a base station (BS) equipped with M antennas, users equipped with a single antenna, and RISs with N reflecting elements. (b) The proposed equivalent (downlink) channel model of RIS-assisted MISO communication: the equivalent channel from the BS to the user is $\mathbf{h}_e(\Phi) \triangleq R_e^{-\frac{\alpha}{2}} \mathbf{h}_c(\Phi) + \mathbf{h}_d$.

the outage probability for the worst and best signal-combining scenarios can realize at a user when controlling an RIS. The correctness of the limits is validated with numerical simulation.

II. SYSTEM MODEL

In this paper, we consider a single-cell communication system in which a base station (BS) is equipped with M antennas, and all users in the communication system are equipped with a single antenna. Some RISs are deployed in the communication system to assist the communications between the BS and the users. Each of the RISs consists of N reflecting elements and is connected to the BS through a controller that can change the phases of the incident signals of each RIS connected to it. Now consider a downlink scenario where a particular RIS assists the signal transmission from the BS to a particular user. Let $\mathbf{H}_b \in \mathbb{C}^{N \times M}$ denote the $N \times M$ complex channel matrix from the BS to the RIS and $\mathbf{h}_r \in \mathbb{C}^N$ denote the $N \times 1$ complex multiple-input-single-output (MISO) channel vector from the RIS to the user. Thus, we can express the *cascaded* channel vector of the links from the BS to the user through the RIS as [12]

$$\mathbf{h}_c(\Phi) \triangleq \mathbf{H}_b^H e^{j\Phi} \mathbf{h}_r, \quad (1)$$

where superscript H stands for the Hermitian transpose, $\Phi \triangleq \text{diag}[\phi_1 \cdots \phi_N]$ is a diagonal phase matrix with elements $\{\phi_i\}$, and ϕ_i denotes the phase (angle) controlled by the i th element of the RIS¹. Note that $\mathbf{h}_c(\Phi) \in \mathbb{C}^M$ is an $M \times 1$ complex vector whose magnitude and phase vary with Φ . In addition, the $M \times 1$ MISO channel vector from the BS to the user is represented by $\mathbf{h}_d \in \mathbb{C}^M$. Suppose $\mathbf{H}_b$, $\mathbf{h}_r$, and $\mathbf{h}_d$ are used to model the fading effects in their corresponding channel. All their entries are *independent complex Gaussian* random variables (RVs) with zero mean and unit variance. An illustration of the RIS-assisted multi-antenna communications between the BS and the user is depicted in Fig. 1(a) where the three constants R_b , R_r , and R_d are the distance between the BS and the RIS,

¹Hence, Φ serves a “rotation matrix” that rotates vector $\mathbf{h}_r$ to vector $e^{j\Phi} \mathbf{h}_r$ and note that Φ does not change the magnitude of the vector it rotates since $(e^{j\Phi} \mathbf{h}_r)^H (e^{j\Phi} \mathbf{h}_r) = \mathbf{h}_r^H (e^{-j\Phi} e^{j\Phi}) \mathbf{h}_r = \mathbf{h}_r^H \mathbf{h}_r$.

the distance between the RIS and the user, and the distance between the BS and the user, respectively.

Consider the downlink scenario of the RIS-assisted multi-antenna communication system in Fig. 1(b) and suppose now the BS transmits signal x to the user. The received signal at the user can be expressed as

$$\begin{aligned} y &\triangleq \sqrt{P} R_d^{-\frac{\alpha}{2}} \left[\left(\frac{R_d}{R_b R_r} \right)^{\frac{\alpha}{2}} \mathbf{h}_c(\Phi) + \mathbf{h}_d \right]^H \mathbf{b} x + n \\ &= \sqrt{P} R_d^{-\frac{\alpha}{2}} \mathbf{h}_e^H(\Phi) \mathbf{b} x + n, \end{aligned} \quad (2)$$

where P is the transmit power of the BS, $\alpha > 0$ denotes the path-loss exponent, $\mathbf{b} \in \mathbb{C}^M$ denotes the beamforming unit vector (i.e., $\|\mathbf{b}\| = 1$ where $\|\mathbf{v}\|$ denotes the L_2 -norm of vector $\mathbf{v}$), n is the thermal noise with power σ_n^2 , and $\mathbf{h}_e(\Phi) \triangleq \left(\frac{R_d}{R_b R_r} \right)^{\frac{\alpha}{2}} \mathbf{h}_c(\Phi) + \mathbf{h}_d$. The received signal model in (2) suggests that the *direct* channel from the BS to the user and the *indirect* channel from the BS to the user through the RIS can be equivalently viewed as an RIS-assisted MISO channel with the channel vector $\mathbf{h}_e(\Phi)$. As a result, the SNR of the RIS-assisted MISO channel can be defined as

$$\text{SNR}(\Phi) \triangleq \frac{P |\mathbf{h}_e^H(\Phi) \mathbf{b}|^2}{\sigma_n^2 R_d^\alpha} \quad (3)$$

by assuming the power of the transmitted signal x is unity. The statistical properties of the RIS-assisted MISO channel gain $|\mathbf{h}_e^H(\Phi) \mathbf{b}|^2$ are partially controllable through the RIS and the BS, and thereby thoroughly investigating them can get some insights into how to ameliorate $\text{SNR}(\Phi)$ by properly manipulating the reflecting elements on the RIS and the beam direction of the antennas at the BS. In the following section, we will use the random matrix theory to first study some important preliminary results about the cascaded channel vector in (1) and then employ them to analyze the statistical properties of $|\mathbf{h}_e^H(\Phi) \mathbf{b}|^2$ given in (3).

III. CONTROLLABLE PROPERTIES OF RIS-ASSISTED MISO CHANNELS

To thoroughly understand how to manipulate the RIS to enhance the SNR in (3), we first need to analyze the controllable statistical properties of the channel gain $|\mathbf{h}_e^H(\Phi) \mathbf{b}|^2$. According to the definition of $\mathbf{h}_e(\Phi)$ in (2), there are two combining situations for the signals from the BS and the RIS happening at the user, i.e., *non-coherent combining* and *coherent combining*. When the RIS can manipulate Φ such that $\mathbf{h}_c(\Phi)$ is aligned with $\mathbf{h}_d$, the coherent combining situation happens at the user and otherwise the non-coherent combining situation occurs. The following theorem shows that the fundamental upper limits of $|\mathbf{h}_e^H(\Phi) \mathbf{b}|^2$ in the two combining situations almost surely exist and relate to $\lambda_b^{\min}$ and $\lambda_b^{\max}$, which are the smallest and the largest eigenvalues of $\mathbf{H}_b \mathbf{H}_b^H$, respectively.

²Throughout this paper, a complex vector $\mathbf{v} = [v_1 \cdots v_K \cdots v_K]^T \in \mathbb{C}^K$ is said to align with another complex vector $\mathbf{u} = [u_1 \cdots u_K \cdots u_K]^T \in \mathbb{C}^K$ if and only if the phases (angles) of the corresponding entries in $\mathbf{v}$ and $\mathbf{u}$ are equal, i.e., $\angle v_k = \angle u_k$ for all $k \in \{1, \dots, K\}$.

Theorem 1. *If the non-coherent combining situation happens, then the smallest upper limit of $|\mathbf{h}_e^H(\Phi)\mathbf{b}|^2$ in (2) can be almost surely derived as*

$$|\mathbf{h}_e^H(\Phi)\mathbf{b}|^2 \leq \left(\sqrt{\lambda_b^{nc} R_e^{-\frac{\alpha}{2}}} \|\mathbf{h}_r\| - \|\mathbf{h}_d\| \right)^2, \quad (4)$$

where $R_e \triangleq \frac{R_b R_r}{R_d}$ and $\lambda_b^{nc} \in [\lambda_b^{\min}, \lambda_b^{\max}]$. On the other hand, if the coherent combining situation occurs, the upper limit of $|\mathbf{h}_e^H(\Phi)\mathbf{b}|^2$ can be almost surely found as

$$|\mathbf{h}_e^H(\Phi)\mathbf{b}|^2 \leq \left(\sqrt{\lambda_b^{cc} R_e^{-\frac{\alpha}{2}}} \|\mathbf{h}_r\| + \|\mathbf{h}_d\| \right)^2, \quad (5)$$

where $\lambda_b^{cc} \in [\lambda_b^{\min}, \lambda_b^{\max}]$.

Proof: See Appendix A. ■

This theorem demonstrates some important phenomena in non-coherent and coherent combining situations when an RIS is employed to assist downlink MISO transmission. First of all, note that the two upper limits in (4) and (5) can be achieved when the beamforming vector $\mathbf{b}$ can be assigned as

$$\mathbf{b} = \frac{\mathbf{h}_e(\Phi)}{\|\mathbf{h}_e(\Phi)\|} = \frac{\mathbf{h}_c(\Phi) + R_e^{\frac{\alpha}{2}} \mathbf{h}_d}{\|\mathbf{h}_c(\Phi) + R_e^{\frac{\alpha}{2}} \mathbf{h}_d\|}. \quad (6)$$

In other words, the upper limits in (4) and (5) are achieved whenever $\mathbf{h}_c(\Phi)$ and $\mathbf{h}_d$ both are available at the BS. In addition, the RIS needs to find a Φ that simultaneously makes $\mathbf{h}_c(\Phi)$ align with $\mathbf{h}_d$ and lets $e^{j\Phi} \mathbf{h}_r$ align with an eigenvector of $\mathbf{H}_b \mathbf{H}_b^H$ corresponding to $\lambda_b^{\max}$ in order to make λ_b^{cc} increase to $\lambda_b^{\max}$. However, finding such a Φ for the RIS may not be possible because it must simultaneously satisfy these two conditions: $\mathbf{H}_b \mathbf{H}_b^H e^{j\Phi} \mathbf{h}_r = \lambda_b^{\max} e^{j\Phi} \mathbf{h}_r$ and $\mathbf{H}_b^H e^{j\Phi} \mathbf{h}_r$ aligned with $\mathbf{h}_d$, yet these two conditions essentially imply that $\mathbf{h}_d$ must be the eigenvector of $\mathbf{H}_b^H \mathbf{H}_b$ corresponding to $\lambda_b^{\max}$, which in general rarely occurs³. Thus, there does not exist a Φ such that λ_b^{cc} increases to $\lambda_b^{\max}$ if $\mathbf{h}_d$ is not the eigenvector of $\mathbf{H}_b \mathbf{H}_b^H$ corresponding to $\lambda_b^{\max}$. On the contrary, λ_b^{nc} can increase to $\lambda_b^{\max}$ since it is definitely possible for the RIS to find a Φ that merely makes $e^{j\Phi} \mathbf{h}_r$ align with an eigenvector corresponding to $\lambda_b^{\max}$. As a result, λ_b^{nc} is not necessarily smaller than λ_b^{cc} once it can go up to $\lambda_b^{\max}$.

In summary, we can conclude the following:

- Controlling Φ is able to manipulate $e^{j\Phi} \mathbf{h}_r$ such that it can be aligned with an eigenvector of $\mathbf{H}_b \mathbf{H}_b^H$ or/and $\mathbf{h}_d$.
- The upper limit in (4) can achieve its maximum (minimum) if Φ is manipulated so as to make $e^{j\Phi} \mathbf{h}_r$ align with the eigenvector of $\mathbf{H}_b \mathbf{H}_b^H$ corresponding to $\lambda_b^{\max}$ ($\lambda_b^{\min}$), i.e., making λ_b^{nc} equal to $\lambda_b^{\max}$ ($\lambda_b^{\min}$).
- The upper limit in (5) cannot achieve its maximum by controlling Φ to make λ_b^{cc} equal to $\lambda_b^{\max}$, unless $\mathbf{h}_d$ is the eigenvector of $\mathbf{H}_b^H \mathbf{H}_b$ corresponding to $\lambda_b^{\max}$.

The above conclusions further indicate that the upper limit of non-coherent combining in (4) can be greater than the upper limit

³This is because $\mathbf{H}_b^H e^{j\Phi} \mathbf{h}_r$ aligned with $\mathbf{h}_d$ leads to $\mathbf{H}_b^H e^{j\Phi} \mathbf{h}_r$ aligned with two vectors $\mathbf{h}_d$ and an eigenvector of $\mathbf{H}_b \mathbf{H}_b^H$, which means $\mathbf{h}_d$ must be an eigenvector of $\mathbf{H}_b^H \mathbf{H}_b$.

of coherent combining in (5) if the following inequality holds

$$\frac{(\lambda_b^{nc} - \lambda_b^{cc})}{2} \|\mathbf{h}_r\| > R_e^{\frac{\alpha}{2}} \|\mathbf{h}_d\|. \quad (7)$$

This inequality essentially reveals a crucial finding, that is, *coherent combining does not necessarily outperform non-coherence combining in the RIS-assisted multi-antenna communications*. More importantly, it is very likely to occur as N is sufficiently large since $\lambda_b^{\max}$ and $\|\mathbf{h}_r\|$ both increase with N . In the following section, more analytical outcomes will be provided to demonstrate this finding in terms of outage probability.

IV. ANALYSIS OF THE OUTAGE PROBABILITY

According to $\text{SNR}(\Phi)$ in (3), the (downlink) outage probability of the user with distance R_d away from the BS in Fig. 1(a) is defined as

$$Q(\theta, \Phi) \triangleq \mathbb{P}[\text{SNR}(\Phi) \leq \theta] = \mathbb{P}\left[\frac{P|\mathbf{h}_e^H(\Phi)\mathbf{b}|^2}{\sigma_n^2 R_d^\alpha} \leq \theta\right],$$

where θ is the threshold of successfully decoding received signals at the user. Since the optimal transmit beamforming vector $\mathbf{b}$ in (6) helps $|\mathbf{h}_e^H \mathbf{b}|$ achieve its upper limits as shown in Theorem 1, the following analyses are conducted by using $\mathbf{b}$ in (6) so that we can study the achievable limits of the outage probability in different situations. Namely, in this section, we will analyze the outage probability given by

$$Q(\theta, \Phi) = \mathbb{P}\left[\|\mathbf{h}_e(\Phi)\|^2 \leq \frac{\theta \sigma_n^2 R_d^\alpha}{P}\right]. \quad (8)$$

in the non-coherent and coherent combining situations, as specified in the following two subsections.

A. The outage probability of non-coherent combining

The explicit results of $Q(\theta, \Phi)$ in the non-coherent combining situation are shown in the following theorem.

Theorem 2. *Suppose the signals from the BS and the RIS are not coherently combined at the user. If $e^{j\Phi} \mathbf{h}_r$ is not an eigenvector of $\mathbf{H}_b \mathbf{H}_b^H$, $Q(\theta, \Phi)$ in (8) for the non-coherent combining situation can be explicitly derived as*

$$Q_{nc}(\theta) = \mathbb{E}\left[\bar{\gamma}\left(N, R_e^\alpha \left(\frac{\sigma_n^2 R_d^\alpha}{P \chi_M^2} \theta - \frac{\mathbf{G}_M}{\chi_M^2}\right)^+\right)\right], \quad (9)$$

where $(z)^+ \triangleq \max\{z, 0\}$ for $z \in \mathbb{R}$, $\bar{\gamma}(K, x) \triangleq \frac{1}{\Gamma(K)} \int_0^x t^{K-1} e^{-t} dt$ is the (regularized) lower incomplete gamma function, $\Gamma(K) \triangleq \int_0^\infty t^{K-1} e^{-t} dt$ is the gamma function, $\mathbf{G}_M \sim \text{Gamma}(M, 1)$ denotes a Gamma RV with shape M and scale 1, and χ_M^2 is a Chi-square RV with M degrees of freedom.

Proof: See Appendix B. ■

There are some important implications in Theorem 2 that are worth further addressing as follows:

- (i) $Q_{nc}(\theta)$ in (9) does not depend on Φ . This is because Φ completely makes $e^{j\Phi} \mathbf{h}_r$ independent of all the eigenvec-

tors of $\mathbf{H}_b \mathbf{H}_b^H$ and using such a Φ does not affect the outage probability.

- (ii) Equation (9) clearly shows how the outage probability of the user is affected by the system parameters (such as N , M , $\frac{P}{\sigma_n^2}$, etc.) such that we can understand which parameters may dominate $Q_{nc}(\theta)$ in different scenarios. For instance, $Q_{nc}(\theta)$ in (9) is very close to zero when P and M are so large that $\mathbb{P}[PG_M \geq \sigma_n^2 \theta R_d^\alpha]$ is very close to unity. Thus, we can know how much power P and how many antennas M the BS needs to make $Q_{nc}(\theta)$ reduce to a desired value based on R_d , θ , and σ_n^2 .
- (iii) Equation (9) is significantly influenced by R_e^α , especially for a transmitting environment with a high path-loss exponent. Accordingly, deploying RISs with an appropriate density so as to largely reduce R_e is essential to maintain a low outage probability.

In addition to the above implications, Equation (9) has a crucial property; we can show that it asymptotically reduces to a simpler form as M increases. This is due to $\frac{1}{\chi_M^2} \approx \frac{1}{M}$ and $\frac{G_M}{\chi_M^2} \approx 1$ for a large M , and thereby $Q_{nc}(\theta)$ in (9) reduces to the following asymptotic outage probability for a large M :

$$Q_{nc}^{\max}(\theta) = \bar{\gamma} \left(N, R_e^\alpha \left(\frac{\sigma_n^2 R_d^\alpha \theta}{MP} - 1 \right)^+ \right), \quad (10)$$

which is interpreted as the “maximum” outage probability of non-coherent combining achieved by an RIS-assisted massive MIMO system in the worst scenario that the RIS neither makes $\mathbf{h}_c(\Phi)$ align with $\mathbf{h}_d$ nor makes $e^{j\Phi} \mathbf{h}_r$ align with an eigenvector corresponding to $\lambda_b^{\max}$ (probably due to the lack of $\mathbf{h}_r$ and $\mathbf{H}_b$ at the BS). Moreover, when $e^{j\Phi} \mathbf{h}_r$ can be aligned with the eigenvector corresponding to $\lambda_b^{\max}$, $M, N \rightarrow \infty$ and $\frac{N}{M} \approx 1$, we can further show that $Q_{nc}(\theta)$ asymptotically reduces to the following lower limit:

$$Q_{nc}^{\min}(\theta) = \bar{\gamma} \left(N, \frac{R_e^\alpha}{4} \left(\frac{\sigma_n^2 R_d^\alpha \theta}{MP} - 1 \right)^+ \right). \quad (11)$$

This lower limit is apparently (much) smaller than $Q_{nc}^{\max}(\theta)$ in (10) owing to a large N . It can be interpreted as the “minimum” outage probability achieved by the RIS-assisted massive MIMO system in the “best” non-coherent combining situation as the RIS can make $e^{j\Phi} \mathbf{h}_r$ align with the eigenvector corresponding to $\lambda_b^{\max}$ and cannot make $\mathbf{h}_c(\Phi)$ align with $\mathbf{h}_d$. According to (10) and (11), we conclude the asymptotic upper and lower limits of $Q_{nc}(\theta)$ as follows

$$Q_{nc}^{\min}(\theta) \lesssim Q_{nc}(\theta) \lesssim Q_{nc}^{\max}(\theta). \quad (12)$$

Such an inequality provides the BS with useful strategies to significantly reduce the outage probability of a particular user when the non-coherent combining situation occurs at the user. For the situation of coherent combining, its outage probability will be analyzed in the following subsection.

B. The outage probability of coherent combining

The outage probability in the coherent combining situation is derived in the following theorem.

Theorem 3. Suppose the signals from the BS and the RIS are coherently combined at the user. If $e^{j\Phi} \mathbf{h}_r$ is not an eigenvector of $\mathbf{H}_b \mathbf{H}_b^H$, then $Q(\theta, \Phi)$ in (8) can be explicitly derived as

$$Q_{cc}(\theta) = \mathbb{E} \left\{ \bar{\gamma} \left(N, \frac{R_e^\alpha G_M}{\chi_M^2} \left[\left(\sqrt{\frac{\sigma_n^2 R_d^\alpha \theta}{G_M P}} - 1 \right)^+ \right]^2 \right) \right\}. \quad (13)$$

Proof: The proof is omitted due to limited space. ■

Some profound observations can be learned by comparing the result in Theorem 3 with that in Theorem 2. Let us first compare (13) with (9), and we can see that the second argument of $\bar{\gamma}(N, \cdot)$ in (9) is greater than or equal to the second argument of $\bar{\gamma}(N, \cdot)$ in (13), i.e.,

$$\left(\frac{\sigma_n^2 R_d^\alpha \theta}{\chi_M^2 P} - \frac{G_M}{\chi_M^2} \right)^+ \geq \left[\left(\sqrt{\frac{\sigma_n^2 R_d^\alpha \theta}{\chi_M^2 P}} - \sqrt{\frac{G_M}{\chi_M^2}} \right)^+ \right]^2$$

because the above-left term is greater than or equal to the second argument of $\bar{\gamma}(N, \cdot)$ in (13). As a result, we have $Q_{cc}(\theta) \leq Q_{nc}(\theta)$ since $\bar{\gamma}(N, x)$ is a monotonically increasing function of x for a given N . When $M \rightarrow \infty$, $\frac{1}{\chi_M^2} \rightarrow \frac{1}{M}$, and $\frac{G_M}{\chi_M^2} \rightarrow 1$, the asymptotic expression of $Q_{cc}(\theta)$ in (13) can be derived as

$$Q_{cc}^{\max}(\theta) \approx \bar{\gamma} \left(N, R_e^\alpha \left[\left(\sqrt{\frac{\sigma_n^2 R_d^\alpha \theta}{MP}} - 1 \right)^+ \right]^2 \right). \quad (14)$$

Since $Q_{cc}^{\max}(\theta)$ in (14) is smaller than $Q_{nc}(\theta)$ in (10), coherent combining always outperforms non-coherent combining in terms of the outage probability if $e^{j\Phi} \mathbf{h}_r$ cannot be aligned with an eigenvector corresponding to $\lambda_b^{\max}$. Similarly, the asymptotic outcome in (14) can be interpreted as the “maximum” outage probability achieved by the massive MIMO communication system in the “worst” coherent combining situation that the RIS cannot make $e^{j\Phi} \mathbf{h}_r$ align with an eigenvector corresponding to $\lambda_b^{\max}$ when making $\mathbf{h}_c(\Phi)$ align with $\mathbf{h}_d$. Moreover, (13) asymptotically reduces to

$$Q_{cc}^{\min}(\theta) = \bar{\gamma} \left(N, \frac{R_e^\alpha}{4} \left[\left(\sqrt{\frac{\sigma_n^2 R_d^\alpha \theta}{MP}} - 1 \right)^+ \right]^2 \right) \quad (15)$$

as M is fairly large. Note that the lower limit in (15) is realized only when the RIS can simultaneously make $e^{j\Phi} \mathbf{h}_r$ align with an eigenvector corresponding to $\lambda_b^{\max}$ as well as make $\mathbf{h}_c(\Phi)$ align with $\mathbf{h}_d$ so that it can be viewed as the minimum outage probability $Q_{cc}^{\min}(\theta)$ achieved in the “best” coherent combining situation. Accordingly, we can have the asymptotic upper and lower limits of $Q_{cc}(\theta, \Phi)$ given by⁴

$$Q_{cc}^{\min}(\theta) \lesssim Q_{cc}(\theta) \lesssim Q_{cc}^{\max}(\theta). \quad (16)$$

⁴Note that $Q_{cc}^{\min}(\theta)$ can hardly be achieved because it is almost impossible for the RIS to find a Φ that not only makes $e^{j\Phi} \mathbf{h}_r$ an eigenvector corresponding to $\lambda_b^{\max}$ but also lets $\mathbf{h}_c(\Phi)$ align with $\mathbf{h}_d$.

TABLE I
MAIN SYSTEM PARAMETERS FOR NUMERICAL SIMULATION

Parameter	Value
Path-loss Exponent α	3.5
SNR Threshold θ	-3 dB
Distance between BS and User R_d	25 (m)
Distance between BS and RIS R_b	20 (m)
Distance between RIS and User R_r	10 (m)
Number of Antennas M	128, 256, 512

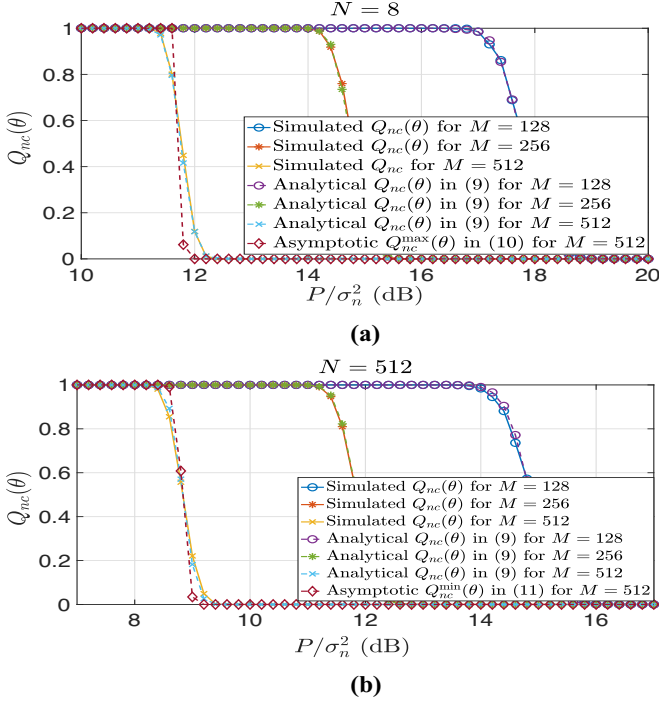


Fig. 2. Simulation results of $Q_{nc}(\theta)$ for $M = 128, 256, 512$: (a) $N = 8$ (b) $N = 512$.

In light of (12) and (16), the asymptotic lower and upper limits of the outage probability in the RIS-assisted massive MIMO system can be concluded as follows:

$$\min\{Q_{nc}^{\min}, Q_{cc}^{\min}\} \lesssim Q(\theta, \Phi) \lesssim Q_{nc}^{\max} \quad (17)$$

because $Q_{nc}^{\max} \geq Q_{cc}^{\max}$ and $Q_{cc}^{\max}$ may not be smaller than $Q_{nc}^{\max}$. The numerical results in the following section will illustrate these limits on $Q(\theta, \Phi)$.

V. NUMERICAL RESULTS

This section provides some numerical results to validate the analytical findings on the upper and lower limits of the outage probability in the previous section. The main system parameters adopted for numerical simulation are listed in Table I, and some system parameters that are altered explicitly for different simulation scenarios are annotated in their corresponding figures. Specifically, we consider the RIS-assisted communication system equipped with a large antenna array that performs non-coherent combining at the user, and the simulation results for the worst and best non-coherent combining situations are shown

in Fig. 2 (a) and (b), respectively. As shown in the figure, the analytical results of $Q_{nc}(\theta)$ in (9) perfectly coincide with their corresponding simulated results in the two situations, and the asymptotic outcomes, i.e., $Q_{nc}^{\max}(\theta)$ in (10) for large M and $Q_{nc}^{\min}(\theta)$ in (11) for large M and N , are also very close to their corresponding simulated and analytical outcomes. Therefore, the correctness of the expressions in (9), (10), and (11) is validated, and thereby the limits found in (16) are derived correctly. It is worth pointing out that significantly increasing the number of the reflecting elements on an RIS can remarkably improve the outage performance. For example, $Q_{nc}^{\min}$ in (b) is improved by about 3 dB if compared with $Q_{nc}^{\max}$ in (a), whereas we may need to double the number of antennas up to 1024 so as to achieve the same amount of improvement when using $N = 8$.

VI. CONCLUSION

In this paper, we studied the fundamental limits of the outage probability in an RIS-assisted massive MIMO communication system by devising an equivalent model of an RIS-assisted MISO channel from a BS to a user to characterize the SNR of the channel in the downlink. The controllable statistical properties of the SNR were investigated first, and the upper limits for the channel gain of the equivalent channel model were found for the two situations of non-coherent and coherent signal combining at the user. These upper limits indicate how to control the RIS to maximize the equivalent channel gain and also assist in analyzing the outage probability of the equivalent channel model. The outage probabilities for non-coherent and coherent signal combining were derived, and their asymptotic expressions for massive MIMO were found. Thus, we understand how the upper and lower limits of the outage probability can be realized by controlling an RIS, which was validated by numerical simulation.

ACKNOWLEDGMENT

The authors thank the Air Force Research Laboratory for supporting the work under Award #FA8650-20-2-5853 and clearing for its public release under Case Number AFRL-2023-5512.

APPENDIX

A. Proof of Theorem 1

Let $h_{e,i}$ and b_i denote the i th entry of vector $\mathbf{h}_e$ and the i th entry of vector $\mathbf{b}$, respectively. From the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned} |\mathbf{h}_e^H(\Phi)\mathbf{b}|^2 &= \left| \sum_{i=1}^M h_{e,i}^\dagger b_i \right|^2 \leq \left(\sum_{i=1}^M |h_{e,i}^\dagger|^2 \right) \cdot \left(\sum_{i=1}^M |b_i|^2 \right) \\ &= [\mathbf{h}_e^H(\Phi)\mathbf{h}_e(\Phi)] \cdot (\mathbf{a}^H\mathbf{b}) = \|\mathbf{h}_e(\Phi)\|^2 \|\mathbf{b}\|^2 \stackrel{(a)}{=} \|\mathbf{h}_e(\Phi)\|^2, \end{aligned}$$

where $\dagger$ denotes the complex conjugate operator and (a) is obtained due to $\|\mathbf{b}\| = 1$ by definition. Since $\mathbf{h}_c(\Phi) = \mathbf{H}_b^H e^{j\Phi} \mathbf{h}_r$, $R_e = \frac{R_b R_r}{R_d}$, and $\mathbf{h}_e(\Phi) = R_e^{-\frac{\alpha}{2}} \mathbf{h}_c(\Phi) + \mathbf{h}_d$, $\|\mathbf{h}_e(\Phi)\|^2$ can be explicitly written as

$$\|\mathbf{h}_e(\Phi)\|^2 = \|R_e^{-\frac{\alpha}{2}} \mathbf{H}_b^H e^{j\Phi} \mathbf{h}_r + \mathbf{h}_d\|^2.$$

Thus, we get the following upper bound on $\|\mathbf{h}_e(\Phi)\|^2$:

$$\begin{aligned} |\mathbf{h}_e^H(\Phi)\mathbf{b}|^2 &\leq \left\| R_e^{-\frac{\alpha}{2}} \mathbf{H}_b^H e^{j\Phi} \mathbf{h}_r + \mathbf{h}_d \right\|^2 \\ &\stackrel{(b)}{=} R_e^{-\alpha} \|\mathbf{h}_c(\Phi)\|^2 + 2R_e^{-\frac{\alpha}{2}} \|\mathbf{h}_c(\Phi)\| \|\mathbf{h}_d\| \cos(\phi_{cd}) + \|\mathbf{h}_d\|^2, \end{aligned}$$

where $\Re\{\mathbb{C}\}$ denotes the real part of a complex number $\mathbb{C}$, (b) follows from $\Re\{\langle \mathbf{h}_c(\Phi), \mathbf{h}_d \rangle\} = \|\mathbf{h}_c(\Phi)\| \|\mathbf{h}_d\| \cos(\phi_{cd})$, and $\phi_{cd} \in [-\pi, \pi]$ is the (random) angle between the two complex vectors $\mathbf{h}_c(\Phi)$ and $\mathbf{h}_d$. Note that ϕ_{cd} depends on Φ . Such an upper bound can be further simplified according to two situations of signal combining, i.e., non-coherent combining and coherent combining. For the situation of non-coherent combining at the user, the minimum of (b) happens when $\cos(\phi_{cd})$ is -1 , so we have the following upper bound:

$$|\mathbf{h}_e^H(\Phi)\mathbf{b}|^2 \leq \left(R_e^{-\frac{\alpha}{2}} \|\mathbf{h}_c(\Phi)\| - \|\mathbf{h}_d\| \right)^2.$$

In addition, we know that $\|\mathbf{h}_c(\Phi)\|^2 = \mathbf{h}_r^H (\mathbf{H}_b \mathbf{H}_b^H) \mathbf{h}_r$ can be almost surely bounded as

$$\lambda_b^{\min} \|\mathbf{h}_r\|^2 \leq \mathbf{h}_r^H (\mathbf{H}_b \mathbf{H}_b^H) \mathbf{h}_r \leq \lambda_b^{\max} \|\mathbf{h}_r\|^2.$$

We also know $\|\mathbf{h}_c(\Phi)\|^2 = \|\mathbf{h}_r\| \cdot \|\mathbf{H}_b \mathbf{H}_b^H \mathbf{h}_r\| \cdot |\cos(\phi_{br})| = \lambda_b^{nc} \cdot \|\mathbf{h}_r\|^2$ if letting ϕ_{br} be the angle between vectors $\mathbf{h}_r$ and $\mathbf{H}_b \mathbf{H}_b^H \mathbf{h}_r$ (i.e., $\phi_{br} \triangleq \angle \mathbf{h}_r - \angle \mathbf{H}_b \mathbf{H}_b^H \mathbf{h}_r$) and defining $\lambda_b^{nc} \triangleq |\cos(\phi_{br})| \cdot \|\mathbf{H}_b \mathbf{H}_b^H \mathbf{h}_r\| / \|\mathbf{h}_r\|$. Hence, we know $\lambda_b^{\min} \leq \lambda_b^{nc} \leq \lambda_b^{\max}$ because $-1 \leq \cos(\phi_{br}) \leq 1$. As a result, $\|\mathbf{h}_c(\Phi)\| \leq \sqrt{\lambda_b^{nc}} \|\mathbf{h}_r\|$ so that there must exist a λ_b^{nc} for non-coherent combining such that the upper limit in (4) exists. On the other hand, if the signals from the BS and the RIS experience the coherent combining situation at the user, then ϕ_{cd} is zero, and thereby we have

$$\begin{aligned} |\mathbf{h}_e^H(\Phi)\mathbf{b}|^2 &\leq R_e^{-\alpha} \|\mathbf{h}_c(\Phi)\|^2 + 2R_e^{-\frac{\alpha}{2}} \|\mathbf{h}_c(\Phi)\| \|\mathbf{h}_d\| \\ &\quad + \|\mathbf{h}_d\|^2 \stackrel{(d)}{=} \left[R_e^{-\frac{\alpha}{2}} \|\mathbf{h}_c(\Phi)\| + \|\mathbf{h}_d\| \right]^2. \end{aligned}$$

Since we know $\lambda_b(\phi_{br}) = |\cos(\phi_{br})| \cdot \|\mathbf{H}_b \mathbf{H}_b^H \mathbf{h}_r\| / \|\mathbf{h}_r\|$, defining $\lambda_b^{cc} \triangleq \lambda_b(\phi_{br})$ for the coherent combining situation gives rise to $\|\mathbf{h}_c(\Phi)\|^2 = \lambda_b^{cc} \|\mathbf{h}_r\|^2$. Note that λ_b^{cc} and λ_b^{nc} may not have the same distribution because the values of ϕ_{br} in the non-coherent and coherent combining situations could be different. Next, replacing $\|\mathbf{h}_c(\Phi)\|$ with $\sqrt{\lambda_b^{cc}} \|\mathbf{h}_r\|$ in the result of (d) leads to the result in (5) and the proof is completed.

B. Proof of Theorem 2

According to (4) and (8), the upper bound in (4) can be achieved so that $Q(\theta, \Phi)$ in the non-coherent combining situation can be expressed as

$$Q_{nc}(\theta) = \mathbb{P} \left[\lambda_b^{nc} R_e^{-\alpha} \|\mathbf{h}_r\|^2 + \|\mathbf{h}_d\|^2 \leq \frac{\sigma_n^2 R_d^\alpha}{P} \theta \right],$$

where $\|\mathbf{h}_r\|^2 \stackrel{d}{=} G_N$ and $\|\mathbf{h}_d\|^2 \stackrel{d}{=} G_M$ because all the entries in $\mathbf{h}_r$ and $\mathbf{h}_d$ are i.i.d. complex Gaussian RVs with zero mean and unit variance⁵. The cumulative density function (CDF) of

$ab\|\mathbf{h}_r\|^2 + \|\mathbf{h}_d\|^2$ for two non-negative constants a and b can be written as

$$\mathbb{P} [ab\|\mathbf{h}_r\|^2 + \|\mathbf{h}_d\|^2 \leq z] = \mathbb{E} \left[F_{a\|\mathbf{h}_r\|^2} \left(\frac{(z - G_M)^+}{b} \right) \right], \quad (\text{A.1})$$

where $F_Z(z)$ denotes the CDF of RV Z . Substituting $a = |\lambda_b^{nc}|$, $b = R_e^{-\alpha}$, and $z = \frac{\sigma_n^2 R_d^\alpha}{P} \theta$ into (A.1) gives rise to

$$Q_{nc}(\theta) = \mathbb{E} [F_{\lambda_b^{nc} G_N} (A(\theta))], \quad (\text{A.2})$$

where $A(\theta) \triangleq R_e^\alpha (\frac{\sigma_n^2}{P} \theta - \|\mathbf{h}_d\|^2)^+ \stackrel{d}{=} R_e^\alpha (\frac{\sigma_n^2 R_d^\alpha}{P} \theta - G_M)^+$.

Now consider $e^{j\Phi} \mathbf{h}_r$ is uncorrelated with the eigenvectors of $\mathbf{H}_b \mathbf{H}_b^H$. We can thus have $\|\mathbf{H}_b^H e^{j\Phi} \mathbf{h}_r\|^2 = \|\mathbf{H}_b^H \mathbf{h}_r\|^2 \stackrel{d}{=} G_N \chi_M^2$. Also, we know $\lambda_b^{nc} \|\mathbf{h}_r\|^2 = \|\mathbf{H}_b^H e^{j\Phi} \mathbf{h}_r\|^2$ and $\|\mathbf{h}_r\|^2 \stackrel{d}{=} G_N$ so that we can readily infer $\lambda_b^{nc} \stackrel{d}{=} \chi_M^2$ as long as $e^{j\Phi} \mathbf{h}_r$ is uncorrelated with any of the eigenvectors of $\mathbf{H}_b \mathbf{H}_b^H$. Thus, (A.2) can be rewritten as $Q_{nc}(\theta) = \mathbb{E}[F_{G_N}(A(\theta)/\chi_M^2)]$, which can be explicitly derived as shown in (9). The proof is complete.

REFERENCES

- [1] G. C. Alexandropoulos, N. Shlezinger, and P. del Hougne, "Reconfigurable intelligent surfaces for rich scattering wireless communications: Recent experiments, challenges, and opportunities," *IEEE Commun. Mag.*, vol. 59, no. 6, pp. 28–34, Jun. 2021.
- [2] Z. Chen, G. Chen, J. Tang, S. Zhang *et al.*, "Reconfigurable-intelligent-surface-assisted B5G/6G wireless communications: Challenges, solution, and future opportunities," *IEEE Commun. Mag.*, vol. 61, no. 1, pp. 16–22, Jan. 2023.
- [3] A. Hemanth, K. Umamaheswari, A. C. Pogaku, D.-T. Do, and B. M. Lee, "Outage performance analysis of reconfigurable intelligent surfaces-aided NOMA under presence of hardware impairment," *IEEE Access*, vol. 8, pp. 212 156–212 165, Nov. 2020.
- [4] L. Yang, Y. Yang, D. B. d. Costa, and I. Trigui, "Outage probability and capacity scaling law of multiple RIS-aided networks," *IEEE Wireless Commun. Lett.*, vol. 10, no. 2, pp. 256–260, Sep. 2020.
- [5] Z. Cui, K. Guan, J. Zhang, and Z. Zhong, "SNR coverage probability analysis of RIS-aided communication systems," *IEEE Trans. Veh. Technol.*, vol. 70, no. 4, pp. 3914–3919, Mar. 2021.
- [6] T. Van Chien, A. K. Papazafeiropoulos, L. T. Tu, R. Chopra, S. Chatzinotas, and B. Ottersten, "Outage probability analysis of IRS-assisted systems under spatially correlated channels," *IEEE Wireless Commun. Lett.*, vol. 10, no. 8, pp. 1815–1819, May 2021.
- [7] T. Van Chien, H. Q. Ngo, S. Chatzinotas, M. Di Renzo, and B. Ottersten, "Reconfigurable intelligent surface-assisted cell-free massive MIMO systems over spatially-correlated channels," *IEEE Trans. Wireless Commun.*, vol. 21, no. 7, pp. 5106–5128, Jul. 2022.
- [8] A. Shojaefard, K.-K. Wong, K.-F. Tong, Z. Chu, A. Mourad *et al.*, "MIMO evolution beyond 5G through reconfigurable intelligent surfaces and fluid antenna systems," *Proc. IEEE*, vol. 110, no. 9, pp. 1244–1265, Sep. 2022.
- [9] A. Papazafeiropoulos, Z. Abdullah, P. Kourtessis, S. Kisseleff, and I. Krikidis, "Coverage probability of STAR-RIS-assisted massive MIMO systems with correlation and phase errors," *IEEE Wireless Commun. Lett.*, vol. 11, no. 8, pp. 1738–1742, Aug. 2022.
- [10] E. Shi, J. Zhang, D. W. K. Ng, and B. Ai, "Uplink performance of RIS-aided cell-free massive MIMO system with electromagnetic interference," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 8, pp. 2431–2445, 2023.
- [11] Y. Zhang, W. Xia, H. Zhao, G. Zheng, S. Lambotharan, and L. Yang, "Performance analysis of RIS-assisted cell-free massive MIMO systems with transceiver hardware impairments," *IEEE Trans. Commun.*, Early Access 2023.
- [12] M. ElMossallamy, H. Zhang, L. Song, K. G. Seddik, Z. Han, and G. Y. Li, "Reconfigurable intelligent surfaces for wireless communications: Principles, challenges, and opportunities," *IEEE Trans. Cognitive Commun. and Netw.*, vol. 6, no. 3, pp. 990–1002, Sep. 2020.

⁵The notation $\stackrel{d}{=}$ denotes the equivalence in probability distribution.