Non-Diagonal RIS Empowered Channel Reciprocity Attacks on TDD-based Wireless Systems

Haoyu Wang*, Zhu Han†, A. Lee Swindlehurst*
*Center for Pervasive Communications and Computing, University of California Irvine, CA, USA.
† Department of Electrical and Computer Engineering at the University of Houston, TX, USA.
Email: haoyuw30@uci.edu, hanzhu22@gmail.com, swindle@uci.edu.

ICC 2024 - IEEE International Conference on Communications | 978-1-7281-9054-9/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICC51166.2024.10622674

***Abstract*—Reconfigurable intelligent surface (RIS) technology can enhance the performance of wireless systems, but an adversary can use such technology to deteriorate communication links. This paper explores an RIS-based attack on multi-user wireless systems that require channel reciprocity for time-division duplexing (TDD). We demonstrate that deploying an RIS with a non-diagonal phase shift matrix can compromise channel reciprocity and lead to poor TDD performance. The attack can be achieved without transmission of signal energy, without channel state information (CSI), and without synchronization with the legitimate system, and thus it is difficult to detect and counteract. We provide an extensive set of simulation studies on the impact of such an attack on the achievable sum rate of the legitimate system, and we design a heuristic algorithm for optimizing the attack in cases where some partial knowledge of the CSI is available. Our results demonstrate that this channel reciprocity attack can significantly degrade the performance of the legitimate system.**

***Index Terms*—Channel Reciprocity, Passive Attack, Channel Estimation, Reconfigurable Intelligent Surfaces, Physical Layer Security.**

I. INTRODUCTION

A reconfigurable intelligent surface (RIS) is an essentially passive array of reflecting elements whose behavior can be dynamically configured. For example, based on available channel state information (CSI), the RIS can be designed to enhance desired signals and reduce the impact of blockages [1], [2]. Recent research has also demonstrated that an RIS can be effectively used to mitigate the impact of physical layer attacks such as jamming and eavesdropping [3]–[8].

On the other hand, it has recently been illustrated how an RIS can serve to disrupt rather than enhance wireless communications [9]–[13]. For example, an adversary could attack the RIS controller of a legitimate system and cause the RIS to act in an undesired way [9], or an illegitimate RIS could be located to enhance reception of sensitive information or direct interference towards certain receivers [10], [11]. A malicious RIS could cause its reflected signal to provide destructive interference at a receiver to cancel a desired direct-path signal, resulting in a reduction in signal-to-noise ratio (SNR) or a complete cancellation of the desired signal [12]. The type of attacks mentioned above are based on knowledge of the CSI for the legitimate links, and thus are difficult to implement.

CSI is not required by other types of RIS-based attacks, such as those that target multiuser multi-input single-output (MU-MISO) systems that employ a time-division duplex (TDD) protocol. In these systems, the users transmit uplink pilot data to a base station (BS) so it can estimate the CSI and design an appropriate precoder for downlink transmission. The approach in [14] is an example of such an attack. In this approach, the RIS introduces pilot contamination on a two-cell TDD system by randomly changing the RIS phase shifts between the uplink pilot transmission stage when the BS estimates the channel, and the downlink stage when the estimated CSI is used. Since the RIS has changed the channel, the designed precoder is no longer optimal and the downlink performance is degraded. A similar strategy is used in [15] for a single-cell MU-MISO system, and again the downlink rate is significantly impacted since the designed precoder cannot eliminate the multiuser interference. The approach presented in [16] uses the same attack philosophy, but with a “silent” (non-reflecting) RIS for the uplink and a randomly time-varying reflection pattern during the downlink. The attacks described above are sometimes referred to as “passive jamming” attacks, since they can be implemented without transmission of an active jamming signal.

While the methods in [14]–[16] do not require CSI, the attacking RIS has to be synchronized with the legitimate system so that it knows when the uplink and downlink periods are occurring. To overcome this significant drawback, we describe an alternative approach that still retains the advantages of the previous methods that do not require CSI or jamming transmissions. Instead of a standard RIS, our proposed approach is based on a new type of RIS architecture that is referred to as a “non-diagonal” RIS [17] (ND-RIS). In an ND-RIS, the signal received by one element is phase-shifted and reflected via a different RIS element. Such an RIS is called “non-diagonal” since it leads to a non-diagonal reflection matrix in the cascaded channel model, unlike a standard RIS whose reflection matrix is diagonal and hence symmetric. The non-diagonal reflection matrix, on the other hand is non-symmetric, and this produces non-reciprocal propagation for the uplink and downlink, which in turn violates the assumption

The work of H. Wang and A. L. Swindlehurst were supported by the U.S. National Science Foundation (NSF) under grants ECCS-2030029, and CNS-2107182. The work of Z. Han was supported by the US Department of Transportation, Amazon, Toyota, and NSF grants CNS-2107216, CNS-2128368, CMMI-2222810, and ECCS-2302469.

on which TDD communications are based. We refer to this novel approach as a Channel Reciprocity AttaCK (CRACK).

The only requirement for CRACK is that the ND-RIS be placed in a location where it can receive and reflect sufficient energy to/from the BS and the user terminals. Without any particular adjustment required, the ND-RIS will inherently produce non-reciprocal uplink and downlink channels without the need for any synchronization with or CSI from the legitimate MU-MISO system. However, if full or partial CSI is available, the ND-RIS can be configured to maximize the multiuser interference and hence minimize the sum rate of the MU-MISO system.

The outline of the paper is as follows. In Section II, we develop a mathematical model for the TDD-based MU-MISO uplink and downlink channels under CRACK to show how an ND-RIS can be used to sabotage the reciprocity of the channels and produce a large amount of multiuser interference. We emphasize that CRACK is a passive attack that does not require power for jamming, synchronization with the legitimate system, nor CSI for the uplink and downlink channels. Furthermore, CRACK is suitable for any type of channel model, but is most effective when the ND-RIS can receive and reflect significant energy to/from the BS and user terminals. Section III introduces the sum ergodic rate performance metric for maximum ratio transmission (MRT) and zero-forcing (ZF) precoding that will be used in the paper. We further design a heuristic optimization method that determines the ND-RIS connections and phase shifts that reduce the sum ergodic rate in cases where knowledge of the line-of-sight (LoS) components of the RIS-cascaded channels is available. Section IV presents the results of several simulation studies that demonstrate the effectiveness of the CRACK approach assuming the BS employs maximum ratio transmit (MRT) and zero forcing (ZF) precoding. Our simulation results demonstrate the significant degradation that CRACK can provide without any CSI, and also how information about only the LoS components of the CSI can dramatically increase the performance loss for both the MRT and ZF cases. Finally, Section V concludes the paper.

Notation: We will denote vectors and matrices by boldface lower and upper case letters, respectively. The transpose and Hermitian transpose operations will be respectively represented by $(\cdot)^T$ and $(\cdot)^H$. The operator $|x|$ indicates the absolute value of a scalar variable x, and $\|\mathbf{x}\|$ is the 2-norm of a vector variable $\mathbf{x}$. The diagonal matrix whose diagonal elements are given by the elements in $\mathbf{x}$ will be represented by $diag(\mathbf{x})$.

II. System Model

We consider the MU-MISO communication system illustrated in Fig. 1, with an M-antenna BS, K single-antenna legitimate users (LUs) denoted as $\mathrm{LU}_1, \mathrm{LU}_2, ..., \mathrm{LU}_\mathrm{K}$ ($K < M$), together with an ND-RIS composed of N elements. Here, we utilize the terms $\mathbf{h}_{k,r}, \mathbf{H}_{r,b}$ and $\mathbf{h}_{k,b}$, respectively, to denote the flat fading channels between LU_k and the ND-RIS, between the BS and the ND-RIS, and between LU_k and the BS. Moreover, the channel $\mathbf{H}_{up}$ is the overall uplink channel between the BS and LUs including both the direct and ND-RIS paths, while $\mathbf{H}_{down} = \mathbf{H}_{up}^T$ is the downlink channel *assumed* by the BS under the assumption of channel reciprocity.

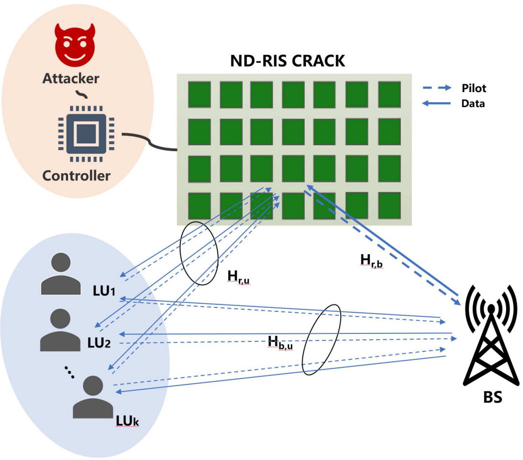

Fig. 1. Illustration of the MU-MISO system under malicious ND-RIS channel reciprocity attacks on the uplink pilot transmission (PT) channel (dashed lines) and the downlink data transmission (DT) channel (solid lines).

A. Time-Division Duplex Mode

For TDD MIMO systems, the LUs transmit pilot signals that the BS uses to estimate $\mathbf{H}_{up}$, and then the BS, assuming the channels are reciprocal, designs a precoder for the downlink, which it assumes to be $\mathbf{H}_{down} = \mathbf{H}_{up}^T$. Mathematically, the pilot signal from LU_k is denoted by $s_{p,k}$, and the received signal at the BS from all the LUs is thus

$$\mathbf{y}_p = \sum_{k=1}^{K} \sqrt{p_k}\mathbf{h}_{k,b}^* s_{p,k} + \mathbf{n}_p, \tag{1}$$

where p_k is the pilot signal power assuming that $E(|s_{p,k}|^2) = 1$, and $\mathbf{n}_p$ denotes the noise at receiver. With the ND-RIS present, the uplink channel for LU_k is $\mathbf{h}_{k,b}^* = \mathbf{H}_{r,b}\mathbf{\Phi}^*\mathbf{h}_{k,r} + \mathbf{h}_{k,b}$, where the direct path is given by $\mathbf{h}_{k,b}$, and the cascaded path through the ND-RIS is represented by $\mathbf{H}_{r,b}\mathbf{\Phi}^*\mathbf{h}_{k,r}$. The ND-RIS is responsible for the $\mathbf{\Phi}^*$ term in the cascaded channel. With knowledge of the pilot signals, the BS estimates the channels for all users, and we will assume that this estimation is error-free, and the overall uplink channel $\mathbf{H}_{up} = [\mathbf{h}_{1,b}^*, \cdots, \mathbf{h}_{K,b}^*] = \mathbf{H}_{r,b}\mathbf{\Phi}^*\mathbf{H}_{r,u} + \mathbf{H}_{b,u} \in \mathbb{C}^{M\times K}$ is exactly known to the BS, where $\mathbf{H}_{r,u} = [\mathbf{h}_{1,r}, \cdots, \mathbf{h}_{K,r}]$ and $\mathbf{H}_{b,u} = [\mathbf{h}_{1,b}, \cdots, \mathbf{h}_{K,b}]$ are the combined user-RIS channel and combined user-BS direct channel, respectively. Assuming channel reciprocity, the BS will then design an appropriate precoder for what it thinks the downlink channel is: $\mathbf{H}_{down} = \mathbf{H}_{up}^T = \mathbf{H}_{r,u}^T\mathbf{\Phi}^{*T}\mathbf{H}_{r,b}^T + \mathbf{H}_{b,u}^T$.

B. ND-RIS CRACK

Due to the existence of RIS, the *actual* downlink channel will be $\mathbf{H}_{down}^* = \mathbf{H}_{r,u}^T\mathbf{\Phi}^*\mathbf{H}_{r,b}^T + \mathbf{H}_{b,u}^T$. For a conventional RIS, $\mathbf{H}_{down} = \mathbf{H}_{down}^*$ are equal because of the diagonal and hence symmetric phase shift matrix. However, the phase shift matrix $\mathbf{\Phi}^*$ for an ND-RIS is generally neither diagonal or symmetric.

More specifically, the ND-RIS phase shift matrix is given by $\boldsymbol{\Phi}^* = \mathbf{J}_l\boldsymbol{\Phi}\mathbf{J}_r$, where $\boldsymbol{\Phi} = \mathrm{diag}\left[\theta_1, \cdots, \theta_N\right] \in \mathbb{C}^{N\times N}$ is a diagonal matrix with $\theta_i = e^{j\phi_i}$, and $\mathbf{J}_l$ and $\mathbf{J}_r$ are left and right permutation matrices, respectively. The permutation matrices $\mathbf{J}_l$ and $\mathbf{J}_r$ can be defined as mappings from one set of ordered indices into another. In particular, for arbitrary $N \times 1$ vectors $\mathbf{x}$ and $\mathbf{y}$, we have $\mathbf{x}^T\mathbf{J}_l = [x_{[1]}, x_{[2]}, \ldots, x_{[N]}]$, which rearranges the entries of $\mathbf{x}$, and $\mathbf{J}_r\mathbf{y} = [y_{(1)}, y_{(2)}, \ldots, y_{(N)}]^T$, which rearranges the entries of $\mathbf{y}$. The permutation mappings are thus denoted as $\mathbf{J}_l : n \rightarrow [n]$ and $\mathbf{J}_r : n \rightarrow (n)$. Since $\mathbf{J}_r$ and $\mathbf{J}_l$ are not necessarily related to each other, and in particular $\mathbf{J}_r \neq \mathbf{J}_l^T$, then $\boldsymbol{\Phi}^* \neq \boldsymbol{\Phi}^{*T}$ and the uplink and downlink channels will not be reciprocal.

III. MRT and ZF Precoding on CRACK

The downlink precoder $\mathbf{W} = [\mathbf{w}_1, \cdots, \mathbf{w}_K]$ designed by the BS is based on the assumed downlink channel $\mathbf{H}_{down}$, and will be suboptimal since the ND-RIS has created a non-reciprocal channel. We will see that this can lead to a severe degradation in system performance. Take maximum ratio transmit (MRT) precoding as an example, where $\mathbf{w}_k$ for $\mathrm{LU_k}$ is the conjugate transpose of the downlink channel for $\mathrm{LU_k}$, denoted as $\mathbf{w}_k = ((\mathbf{h}^*_{k,b})^T)^H \in \mathbb{C}^{M\times 1}$. The complete MRT precoder is expressed as $\mathbf{W} = \mathbf{H}^H_{down} = \left(\mathbf{H}^T_{up}\right)^H$, and the signals received by the K users are given by

$$\mathbf{y} = \mathbf{H}^*_{down}\mathbf{W}\mathbf{D}\mathbf{s} + \mathbf{n}_d, \tag{2}$$

where $\mathbf{D} = \mathrm{diag}(\sqrt{P_1}, \sqrt{P_2}, ..., \sqrt{P_K})$ is a $K \times K$ diagonal power allocation matrix and $\mathbf{n}_d$ is the received noise vector, whose elements are modeled as independent and identically distributed (i.i.d.) Gaussian random variables with zero mean and variance σ^2.

A. Achievable Rate for MRT Precoding

As explained before, the actual channel from the BS to $\mathrm{LU_k}$, denoted as $\mathbf{h}^*_{b,k} \in \mathbb{C}^{1\times M}$, is given by $\mathbf{h}^*_{b,k} = \mathbf{h}^T_{k,r}\boldsymbol{\Phi}^*\mathbf{H}^T_{r,b} + \mathbf{h}^T_{k,b}$. Therefore, the received downlink signal at $\mathrm{LU_k}$ can be written as

$$y_k = \sqrt{P_k}\mathbf{h}^*_{b,k}\mathbf{w}_k s_k + \mathbf{h}^*_{b,k}\sum_{i=1,i\neq k}^{K}\sqrt{P_i}\mathbf{w}_i s_i + n_k. \tag{3}$$

For MRT, the precoding vector for $\mathrm{LU_k}$ is

$$\mathbf{w}_k = ((\mathbf{h}^*_{k,b})^T)^H = (\mathbf{H}^T_{r,b})^H\widetilde{\boldsymbol{\Phi}}(\mathbf{h}^T_{k,r})^H + (\mathbf{h}^T_{k,b})^H \ ,$$

where $\widetilde{\boldsymbol{\Phi}} = ((\boldsymbol{\Phi}^*)^T)^H = \mathbf{J}_l\boldsymbol{\Phi}^H\mathbf{J}_r$ is the conjugate of $\boldsymbol{\Phi}^*$. For notational simplicity, we write $\mathbf{h}_{r,k} = \mathbf{h}^T_{k,r}$, $\mathbf{H}_{b,r} = \mathbf{H}^T_{r,b}$, and $\mathbf{h}_{b,k} = \mathbf{h}^T_{k,b}$. Thus, the actual received signal at $\mathrm{LU_k}$ can be re-written as

$$\begin{aligned} y_k = &\sqrt{P_k}\mathbf{h}^*_{b,k}(\mathbf{H}^H_{b,r}\widetilde{\boldsymbol{\Phi}}\mathbf{h}^H_{r,k} + \mathbf{h}^H_{b,k})s_k \\ &+ \mathbf{h}^*_{b,k}\sum_{i=1,i\neq k}^{K}\sqrt{P_i}(\mathbf{H}^H_{b,r}\widetilde{\boldsymbol{\Phi}}\mathbf{h}^H_{r,i} + \mathbf{h}^H_{b,i})s_i + n_k. \end{aligned} \tag{4}$$

Subsequently, the downlink signal-to interference-plus-noise ratio (SINR) for $\mathrm{LU_k}$ is determined by (5). Assuming ergodicity in all channels, the achievable ergodic rate for $\mathrm{LU_k}$ can be expressed as (6). The sum ergodic rate of the MU-MISO system is denoted as $R = \sum_{i=1}^{K} r^*_k$.

B. Achievable Rate for ZF Precoding

The conventional ZF precoder based on the estimated downlink channel is designed as

$$\mathbf{W} = \mathbf{H}^H_{down}\left(\mathbf{H}_{down}\mathbf{H}^H_{down}\right)^{-1} = [\mathbf{w}_1, \cdots, \mathbf{w}_K]. \tag{7}$$

Since the true channel is not equal to $\mathbf{H}_{down}$ the orthogonality between the precoding vector $\mathbf{w}_k$ for $\mathrm{LU_k}$ and the other users' actual downlink channels is broken, which leads to severe multiuser interference and thus dramatic degradation of the overall capacity of the MU-MISO system.

C. Assumed Channel Models

In the simulation analysis of the next section, we will adopt the following common channel models: Rayleigh fading for all direct BS-User channels, Rician fading for all User-RIS channels, and also for the BS-RIS channel. Mathematically, the channel between $\mathrm{LU_k}$ and BS is given by

$$\mathbf{h}_{k,b} = \sqrt{\alpha_{k,b}}\widetilde{\mathbf{h}}_{k,b}, \tag{8}$$

where $\alpha_{k,b} = \rho d_{k,b}^{-\iota_{k,b}}$ is the distance-dependent large-scale path-loss, ρ is the path loss at the reference distance $d_0 = 1$m, $d_{k,b}$ is the distance between $\mathrm{LU_k}$ and BS, and $\iota_{k,b}$ is the path loss exponent of the $\mathrm{LU_k}$-to-BS link. The elements of $\widetilde{\mathbf{h}}_{k,b} \in \mathbb{C}^{M\times 1}$ are drawn from an i.i.d. complex Gaussian distribution with zero mean and unit variance. The Rician channel $\mathbf{h}_{k,r}$ between $\mathrm{LU_k}$ and the ND-RIS is expressed as

$$\mathbf{h}_{k,r} = \sqrt{\alpha_{k,r}}\left(\sqrt{\frac{\kappa_{k,r}}{1+\kappa_{k,r}}}\overline{\mathbf{h}}_{k,r} + \sqrt{\frac{1}{1+\kappa_{k,r}}}\widetilde{\mathbf{h}}_{k,r}\right), \tag{9}$$

where $\kappa_{k,r}$ is the Rician factor, $\overline{\mathbf{h}}_{k,r}$ is the line-of-sight (LoS) channel component, and the elements of the non-LoS (NLoS) term $\widetilde{\mathbf{h}}_{k,r}$ follow an i.i.d. complex Gaussian distribution with zero mean and unit variance. Similarly, the Rician BS-RIS link is described as

$$\mathbf{H}_{r,b} = \sqrt{\alpha_{r,b}}\left(\sqrt{\frac{\kappa_{r,b}}{1+\kappa_{r,b}}}\overline{\mathbf{H}}_{r,b} + \sqrt{\frac{1}{1+\kappa_{r,b}}}\widetilde{\mathbf{H}}_{r,b}\right), \tag{10}$$

where the elements of the NLoS component $\widetilde{\mathbf{H}}_{r,b} \in \mathbb{C}^{M\times N}$ also follow an i.i.d. complex Gaussian distribution with zero mean and unit variance.

IV. Numerical Results

This section presents numerical results to assess the performance of the proposed ND-RIS CRACK approach introduced in Section II. We examine a TDD-based MU-MISO system where the BS adopts either MRT or ZF precoding, as detailed in Section III, for downlink transmission. Both the BS and ND-RIS are outfitted with uniform linear arrays (ULA), where each element is spaced one-half wavelength apart. The BS is located at the 3D coordinates (10m, 40m, 20m), and the LUs are randomly distributed in a circular region S centered at (10m, 0m, 1.5m) with a radius of 10m. The ND-RIS is deployed

$$\eta_k = \frac{P_k|(\mathbf{h}_{r,k}\mathbf{\Phi}^*\mathbf{H}_{b,r} + \mathbf{h}_{b,k})(\mathbf{H}_{b,r}^H\widetilde{\mathbf{\Phi}}\mathbf{h}_{r,k}^H + \mathbf{h}_{b,k}^H)|^2}{\sum_{i=1,i\neq k}^{K} P_i|(\mathbf{h}_{r,k}\mathbf{\Phi}^*\mathbf{H}_{b,r} + \mathbf{h}_{b,k})(\mathbf{H}_{b,r}^H\widetilde{\mathbf{\Phi}}\mathbf{h}_{r,i}^H + \mathbf{h}_{b,i}^H)|^2 + \sigma^2}. \tag{5}$$

$$r_k^* = \mathbb{E}[\log(1+\eta_k)] = \mathbb{E}\left[\log\left(1 + \frac{P_k|(\mathbf{h}_{r,k}\mathbf{\Phi}^*\mathbf{H}_{b,r} + \mathbf{h}_{b,k})(\mathbf{H}_{b,r}^H\widetilde{\mathbf{\Phi}}\mathbf{h}_{r,k}^H + \mathbf{h}_{b,k}^H)|^2}{\sum_{i=1,i\neq k}^{K} P_i|(\mathbf{h}_{r,k}\mathbf{\Phi}^*\mathbf{H}_{b,r} + \mathbf{h}_{b,k})(\mathbf{H}_{b,r}^H\widetilde{\mathbf{\Phi}}\mathbf{h}_{r,i}^H + \mathbf{h}_{b,i}^H)|^2 + \sigma^2}\right)\right], \tag{6}$$

at the location (0m, 35m, 15m). The path-loss exponents of the RIS-User link and BS-User link are set as $\iota_{k,r} = 2.5$ and $\iota_{k,b} = 3.5$, respectively. Unless otherwise specified, the path-loss exponent of the RIS-BS link is $\iota_{r,b} = 2$ and the BS transmit power is $P_k = 20$ dBm. Several different cases for the number of BS antennas and ND-RIS elements are investigated, as listed in Table I together with other simulation parameters.

TABLE I
SIMULATION PARAMETERS

Parameter	Value
Number of BS antennas M	[16,32,64,128,256,512]
Number of RIS elements N	[16,32,64,128,256,512]
Number of users K	4
Path loss ρ	-20 dB
Carrier frequency	28 GHz
Bandwidth BW	10 MHz
Noise power σ^2	$-170 + 10\log_{10}(\text{BW})$ dBm
Rician factor $\kappa_{k,r}$	3
Rician factor $\kappa_{r,b}$	16

A. Benchmark Schemes

Our proposed ND-RIS CRACK implementation will be compared with the following benchmark approaches:

* **Without ND-RIS**: No ND-RIS CRACK is present.
* **Random ND-RIS**: The values of the ND-RIS phase shifts are randomly chosen to implement the reciprocity attack. In particular, $\mathbf{J}_l$, $\mathbf{J}_r$, and $\mathbf{\Phi}$ follow uniform distributions where each possible state is equally likely.
* **Heuristic Algorithm (HA)**: The phase shifts $\mathbf{\Phi}$ are randomly generated in each transmission period, and based on knowledge of the MU-MISO system's LoS channels, $\mathbf{J}_l$ and $\mathbf{J}_r$ are designed by exhaustive search with 100 random generations to maximize the difference between the uplink and downlink of the LoS BS-RIS-User channels $\sum_{k=1}^{K}\beta_k\|\overline{\mathbf{H}}_{r,b}\mathbf{\Phi}^*\overline{\mathbf{h}}_{k,r} - \overline{\mathbf{H}}_{r,b}\mathbf{\Phi}^{*T}\overline{\mathbf{h}}_{k,r}\|^2$, where $\beta_k = \frac{\alpha_{k,r}\alpha_{r,b}\kappa_{k,r}\kappa_{r,b}}{(1+\kappa_{k,r})(1+\kappa_{r,b})}$.

The HA algorithm is an attempt to find an optimal choice of the ND-RIS phase shifts in order to inflict a maximum reduction in the MU-MISO system performance.

B. Non-Optimized CRACK

Here we focus on the performance of ND-RIS CRACK with random phase shifts. 300 random ND-RIS configurations are generated, and for each configuration, we calculate the average sum ergodic rate by averaging the achieved sum rate over 2000 random channel realizations.

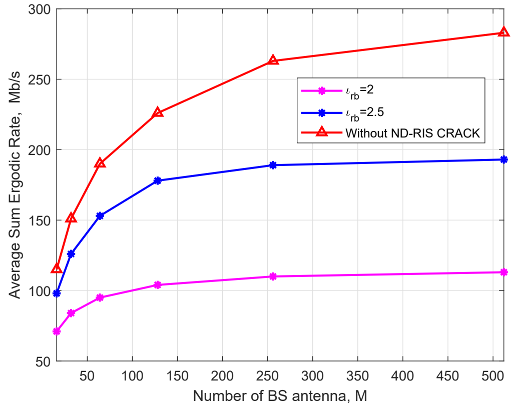

Fig. 2. Average sum ergodic rate versus different number of BS antennas (MRT, N=128).

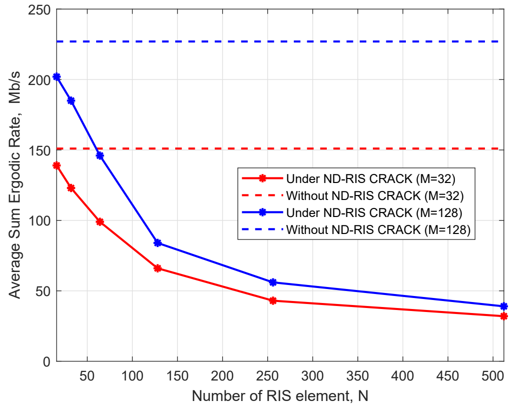

Fig. 3. Average sum ergodic rate versus ND-RIS size (MRT).

1) MRT beamforming on CRACK: Fig. 2 illustrates the simulated average sum ergodic rate in relation to the number of BS antennas and emphasizes the impact of the BS-RIS channel path loss on CRACK performance, assuming the BS utilizes MRT beamforming. With the ND-RIS in place, CRACK cannot be thwarted by increasing the number of BS antennas M, as the ergodic sum rate converges to a fixed value for large M. Conversely, as expected, the effectiveness of CRACK relies on the quality of the RIS-BS channel. A higher path loss leads to a reduced cascaded RIS channel gain, resulting in a weaker CRACK impact.

In Fig. 3, we show the average sum ergodic rate versus different numbers of ND-RIS elements N for $M = 32$ and $M = 128$ BS antennas, where the BS employs MRT

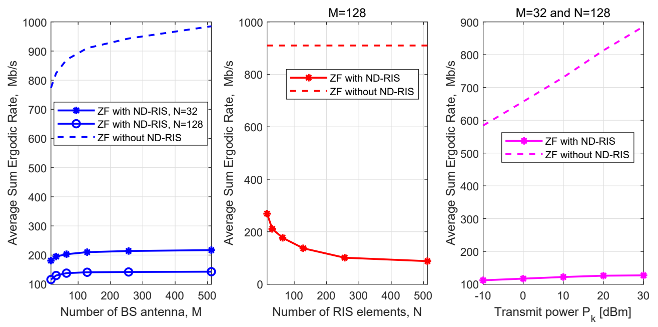

Fig. 4. Average ergodic sum rate of MU-MISO system with ZF precoding.

beamforming. As N increases, the sum ergodic rate initially experiences a sharp decrease before stabilizing at an extremely low level. This observation highlights the severe detrimental effect of CRACK on the MU-MISO system, where a substantial performance degradation can be realized without relying on CSI for the ND-RIS design. Notably, we observe a nearly 80% reduction in the sum ergodic rate when $N = 512$ and $M = 128$.

2) ZF beamforming on CRACK: The proposed CRACK approach can degrade the MU-MISO system throughput for other precoders besides MRT. In Fig. 4, we present results for the ZF precoding scenario, depicting the sum ergodic rate with respect to the number of BS antennas, RIS elements, and BS transmit power, assuming a randomly determined ND-RIS configuration. In comparison to the results with MRT from the previous section, we see that CRACK yields a significantly greater degradation in performance for ZF precoding, as ZF is designed to ideally eliminate all multiuser interference. Similarly, using a larger BS antenna array can provide little relief to the attack, while an increasing RIS size brings significantly increased performance loss. The rightmost figure further illustrates that increasing the BS transmit power is not an effective strategy to overcome CRACK, as would be the case for other types of physical layer attacks.

3) 1-bit ND-RIS CRACK: In this example, we demonstrate that the effectiveness of CRACK is not contingent on continuous phase control of the ND-RIS elements. To show this, we consider an ND-RIS whose phase shifts can be in one of two possible states, $\theta \in \{0, \pi\}$, requiring one bit of control. Fig. 5 illustrates the sum ergodic rate for both MRT and ZF precoding at the BS concerning the number of ND-RIS elements, comparing the performance degradation of a 1-bit ND-RIS, an ND-RIS with full phase control, and a system without an ND-RIS. We observe that the 1-bit results are similar to those obtained with full phase control.

4) ND-RIS Deployment Strategy: This example highlights that the ND-RIS can exert a more pronounced impact on the system performance when positioned in closer proximity to the BS than to the users. To demonstrate this, we position the ND-RIS near the line connecting the BS and the users and then systematically adjust the separation distance between the ND-RIS and the BS. Fig. 6 shows that as the distance increases and the ND-RIS approaches the users, the sum ergodic rate of the MU-MISO system experiences an increase for both ZF and MRT.

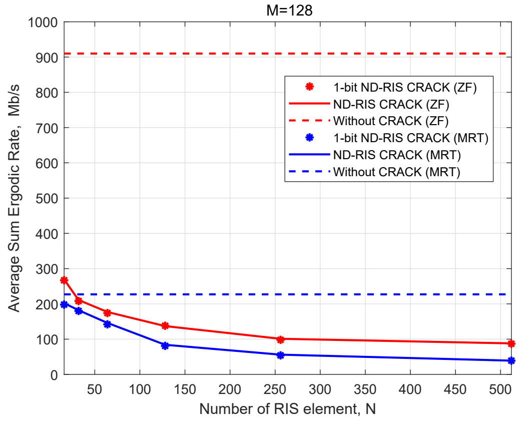

Fig. 5. Average ergodic sum rate under 1-bit ND-RIS CRACK with M=128.

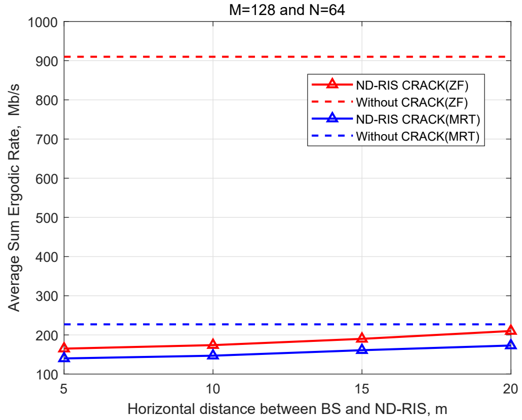

Fig. 6. Influence of RIS-BS distance on ergodic sum rate under CRACK.

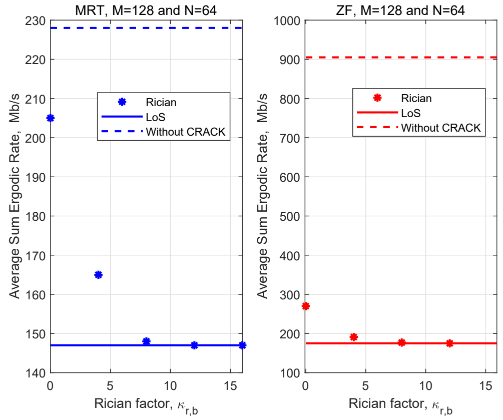

Fig. 7. Influence of RIS-BS Rician factor on ergodic sum rate under CRACK.

5) Influence of RIS-BS Rician Factor: Fig. 7 demonstrates that the presence of a LoS path between the RIS and the BS (RIS-BS LoS path) provides substantial potential for CRACK, in stark contrast to the scenario with $\kappa_{r,b} = 0$. As the strength of the LoS path intensifies, the average sum ergodic rate of the MU-MISO system under CRACK will decrease for both

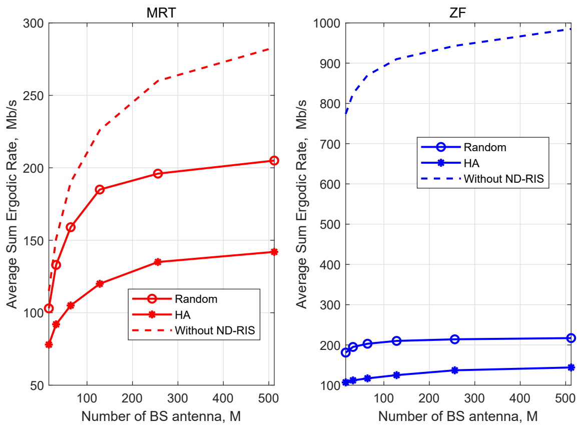

Fig. 8. HA performance for ND-RIS CRACK (N=32).

ZF and MRT, with a larger impact on MRT. When $\kappa_{r,b} \geq 12$, the Rician and LoS channel models produce similar results, and in such cases the influence of the NLoS component can be ignored. However, CRACK can provide significant performance degradation even for small $\kappa_{r,b}$.

C. Heuristic CRACK Optimization

Here, we demonstrate that when a malicious attacker can obtain the NLoS channel statistics CSI and the LoS channel components of the RIS-cascaded channel, the ND-RIS can be designed based on the proposed HA to implement a more severe attack on the system, even with a relatively small ND-RIS. Fig. 8 illustrates the average sum ergodic rate versus the number of BS antennas. As M increases, the average sum ergodic rate under the two ND-RIS schemes initially grows and then stabilizes at a much lower value than the ideal scenarios without CRACK. However, compared to randomly selecting the ND-RIS phases, HA can dramatically enhance the CRACK effect on the system whether the BS adopts MRT or ZF precoding. Specifically, HA realizes nearly a 47% performance degradation for MRT and nearly 86% for ZF, compared to the case without the ND-RIS.

V. Conclusions

In this study, we introduced a novel approach referred to as an ND-RIS-based channel reciprocity attack, or CRACK, to impair the performance of time-division duplex communication systems. CRACK operates by subtly disrupting the presumed reciprocity that underlies the legitimate channel, thereby injecting unexpected multi-user interference into the targeted communication system. Such interference cannot be counteracted by simply increasing the number of BS antennas or transmit power. Furthermore, this attack can be executed without the need for any CSI about the BS-user links or any links involving the ND-RIS. Additionally, CRACK operates independently of synchronization with the training or data phases of the communication system. However, if the NLoS channel statistics are available together with information about the LoS links, CRACK can realize a more severe attack using a simple heuristic approach. Simulations have also demonstrated that achieving significant performance degradation only requires selecting the phase of the ND-RIS elements with one-bit of precision.

References

[1] C. Pan, H. Ren, K. Wang, J. F. Kolb, M. Elkashlan, M. Chen, M. Di Renzo, Y. Hao, J. Wang, A. L. Swindlehurst, X. You, and L. Hanzo, “Reconfigurable intelligent surfaces for 6G systems: Principles, applications, and research directions,” *IEEE Commun. Mag.*, vol. 59, no. 6, pp. 14–20, June 2021.

[2] M. A. 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. Cogn. Commun. Netw.*, vol. 6, no. 3, pp. 990–1002, Sept. 2020.

[3] A. Almohamad, A. M. Tahir, A. Al-Kababji, H. M. Furqan, T. Khattab, M. O. Hasna, and H. Arslan, “Smart and secure wireless communications via reflecting intelligent surfaces: A short survey,” *IEEE Open and J. Commun. Soc.*, vol. 1, pp. 1442–1456, Sept. 2020.

[4] M. Hua, Q. Wu, W. Chen, O. A. Dobre, and A. Lee Swindlehurst, “Secure intelligent reflecting surface aided integrated sensing and communication,” *IEEE Trans. Wireless Commun., Early Access*, 2023.

[5] J. Zhang, H. Du, Q. Sun, B. Ai, and D. W. K. Ng, “Physical layer security enhancement with reconfigurable intelligent surface-aided networks,” *IEEE Trans. Info. Forensics & Security*, vol. 16, pp. 3480–3495, 2021.

[6] Z. Zhang, C. Zhang, C. Jiang, F. Jia, J. Ge, and F. Gong, “Improving physical layer security for reconfigurable intelligent surface aided NOMA 6G networks,” *IEEE Trans. Vehic. Tech.*, vol. 70, no. 5, pp. 4451–4463, May 2021.

[7] L. Dai, H. Huang, C. Zhang, and K. Qiu, “Silent flickering RIS aided covert attacks via intermittent cooperative jamming,” *IEEE Wireless Commun. Lett.*, vol. 12, no. 6, pp. 1027–1031, June 2023.

[8] S. Lin, Y. Xu, H. Wang, J. Gu, J. Liu, and G. Ding, “Secure multicast communications via RIS against eavesdropping and jamming with imperfect CSI,” *IEEE Trans. Vehic. Tech., Early Access*, 2023.

[9] H. Alakoca, M. Namdar, S. Aldirmaz-Colak, M. Basaran, A. Basgumus, L. Durak-Ata, and H. Yanikomeroglu, “Metasurface manipulation attacks: Potential security threats of RIS-aided 6G communications,” *IEEE Commun. Mag.*, vol. 61, no. 1, pp. 24–30, Jan. 2023.

[10] M. Wei, H. Zhao, V. Galdi, L. Li, and T. J. Cui, “Metasurface-enabled smart wireless attacks at the physical layer,” *Nat Electron 6*, p. 610–618 (2023), Aug. 2023.

[11] Y. Wang, H. Lu, D. Zhao, Y. Deng, and A. Nallanathan, “Wireless communication in the presence of illegal reconfigurable intelligent surface: Signal leakage and interference attack,” *IEEE Wireless. Commun.*, vol. 29, no. 3, pp. 131–138, June 2022.

[12] B. Lyu, D. T. Hoang, S. Gong, D. Niyato, and D. I. Kim, “IRS-based wireless jamming attacks: When jammers can attack without power,” *IEEE Wireless Commun. Lett.*, vol. 9, no. 10, pp. 1663–1667, Oct. 2020.

[13] R. Kaur, B. Bansal, S. Majhi, S. Jain, C. Huang, and C. Yuen, “A survey on reconfigurable intelligent surface for physical layer security of next-generation wireless communications,” *IEEE Open J. Veh. Technol.*, vol. 5, pp. 172–199, Jan. 2024.

[14] X. Luo and H. Zhu, “IRS-based TDD reciprocity breaking for pilot decontamination in massive MIMO,” *IEEE Wireless Commun. Lett.*, vol. 10, no. 1, pp. 102–106, Jan. 2021.

[15] H. Huang, Y. Zhang, H. Zhang, C. Zhang, and Z. Han, “Illegal intelligent reflecting surface based active channel aging: When jammer can attack without power and CSI,” *IEEE Trans. Vehic. Tech.*, vol. 72, no. 8, pp. 11 018–11 022, Aug. 2023.

[16] H. Huang, Y. Zhang, H. Zhang, Y. Cai, A. Lee Swindlehurst, and Z. Han, “Disco intelligent reflecting surfaces: Active channel aging for fully-passive jamming attacks,” *IEEE Trans. Wireless Commun., Early Access*, 2023.

[17] Q. Li, M. El-Hajjar, I. Hemadeh, A. Shojaeifard, A. A. M. Mourad, B. Clerckx, and L. Hanzo, “Reconfigurable intelligent surfaces relying on non-diagonal phase shift matrices,” *IEEE Trans. Vehic. Tech.*, vol. 71, no. 6, pp. 6367–6383, June 2022.