Antifragility of RIS-assisted Communication Systems under Jamming Attacks

Mounir Bensalem, Thomas Röthig and Admela Jukan

Technische Universität Braunschweig, Germany; {mounir.bensalem, t.roethig, a.jukan}@tu-bs.de

Abstract—Antifragility of communication systems is defined as measure of benefits gained from the adverse events and variability of its environment. In this paper, we introduce the notion of antifragility in Reconfigurable Intelligent Surface (RIS) assisted communication systems affected by a jamming attack. We analyzed the antifragility of the two hop systems, where the wireless path contains source node, RIS, destination node, and a eavesdropping/jamming node. We propose and analyze the antifragility performance for several jamming models, such as Digital Radio Frequency Memory (DRFM) and phase and amplitude shifting. Our paper shows that antifragility throughput can indeed be achieved under certain power thresholds and for various jamming models. In particular, high jamming power combined with low baseline data rates yields an antifragile gain factor of approximately five times. The results confirm that reconfigurable intelligent surfaces, when coupled with an antifragile design philosophy, can convert hostile interference from a liability into a throughput gain.

I. INTRODUCTION

Conventional resilient communication seeks to prevent or mitigate the effects of disruptions through hostile conditions, or attacks. Whereas resilience aims at preserving or restoring the existing functionality, a new paradigm, referred to as *antifragility*, goes a step further by leveraging uncertainty and adverse conditions for performance gains. According to its original principles first presented in [1], antifragile systems thrive under randomness, by improving their operational metrics when faced with challenges that would simply degrade or disrupt other systems. The first work on application of antifragility in communication systems by [2] showed that network throughput can greatly improve when subjected to adversarial jamming attacks. Despite the growing interest in designing and operating future mobile networks reliably, antifragility remains a relatively under-explored concept in communications and networking.

We are interested in the specific research question which is whether one new technology carries to potential for antifragile performance gains, i.e., Reconfigurable Intelligent Surfaces (RIS). Aside from being specific, our research question is also relevant, since RIS, by employing a two-dimensional array of reflecting elements, where the parameter, such as amplitude and phase, can dynamically be tuned, is known to effectively mitigate interference and extend coverage [3]. It is this capability especially that carries potential to achieving antifragile gains under jamming attacks. Analyzing RIS-assisted systems under the jamming scenarios is therefore fundamental to assessing their adaptability and to developing strategies that leverage deliberate interference for performance gains.

In this paper, we study analytically whether RIS-assisted communication systems can achieve antifragile performance gains under jamming attack. We focus on a two-hop communication system consisting of source node, RIS, and destination node, with an eavesdropping or jamming node acting as an adversarial entity. We subject this system to various jamming models, including Digital Radio Frequency Memory (DRFM) and phase and amplitude shifting. Our analysis reveals that by leveraging the reflections of the RIS, communication systems subjected to high jamming power can, in fact, experience improved throughput under certain conditions, demonstrating a remarkably antifragile behavior. Specifically, when the system baseline data rate is low and adversarial power is high, we observe a notable antifragile gain factor of up to $5\times$. Although the gain factor is smaller when baseline data rates are larger, it still represents a notable improvement under strong jamming. By addressing a range of jamming techniques and analyzing the role of RIS in achieving antifragility, this work provides a foundation for developing more generally adaptive, self-improving wireless networks.

The rest of the paper is organized as follows. Section II presents the related work. Section III provides the system model and Section IV the antifragility scheme design. Section V discusses numerical results. Section VI concludes the paper.

II. RELATED WORK

Paper [4] presented a method to carefully design frequency-shift keying (FSK) waveforms in order to exploit reactive jammers, effectively forcing an attacker, i.e., the jammer node, to act as an unintentional relay and thereby enhance the data rates of the legitimate users. Building on this fundamental idea, the early work in [2] studied the antifragile wireless communications showing the antifragile throughput gain under reactive jamming scenarios. More recently, [5], basing its antifragility scheme on paper [4] focused on reducing the outage probability in multi-relay cognitive networks, however without consideration of the jamming model type, signal separation and spoofing signal creation.

Our study is the first to explore antifragility achieved in wireless communications with RIS. We investigate how RIS-assisted systems can exhibit antifragile behavior under diverse and potentially severe jamming models, including Digital Radio Frequency Memory (DRFM) and phase/amplitude-shifting

attacks, as well as passive and active RIS configurations. We adopt the RIS channel model used in several related papers, such as in [6]–[8]. The attackers' channel models are modified and adapted from [4], [6], [9]. Recent RIS studies have addressed reactive jammers and emphasized the need for resilient systems. The paper [10] proposed a joint beamforming optimization scheme for active and passive transmission that maximizes both the secrecy rate and the system throughput, reducing the impact of the jamming attack. Contrary to this approach, we leverage the jammer as an information relaying node and therefore configure the RIS phases solely to boost the legitimate users SNR, with no mechanism to degrade the eavesdropping link. Another resilient RIS-assisted system is presented in [11]. This work proposes two partitioned RIS optimization schemes, which integrate beamforming with artificial noise injection to suppress the eavesdroppers SNR and improve the legitimate users secrecy capacity. In contrast to this approach, we harness the jammers signal as a constructive relay. Hence, our design refrains from artificial noise injection and instead leverages the jammers' SNR to boost the achievable throughput in the system.

In sum, this paper provides a novel antifragile communication framework, which can however effectively integrate known techniques, such as beamforming and jammer piggybacking, with jammer localization and mitigation. To the best of our knowledge, this is the first work to analyze antifragility in RIS-assisted wireless communications.

III. SYSTEM MODEL

A. Reference Scenarios

The system is illustrated in Figure 1. It includes a transmitter (source), receiver (destination), and one RIS node. Two types of jamming nodes are illustrated, one that eavesdrop either the source (Jammer 1) or the signal reflected by the RIS (Jammer 2). We refer to the Jammer 1 scenario as Source-Aware Eavesdropping and to the Jammer 2 scenario as RIS-Aware-Eavesdropping.

Figure 1: RIS-based jamming scenarios

B. Channel Model

We consider a Rayleigh fading channel model for legitimate communication. The received signal at the destination node for a jammer free system is here given as [6]–[8],

$$y(t) = \mathbf{h}_{SR}\mathbf{R}^{1/2}\Phi\mathbf{R}^{1/2}\mathbf{h}_{RD}x(t) + w(t)$$
$$= \left(\sum_{a=1}^{M}\sum_{k=1}^{M}\sum_{l=1}^{M}\rho_{a,k}^{1/2}\rho_{a,l}^{1/2}h_{SR,k}h_{RD,l}e^{j\phi_a}\right)x(t) + w(t) \quad (1)$$

Where M denotes the number of elements in the legitimate RIS, $x(t)$ is the transmitted signal from the source to the destination node, $\mathbf{h}_{SR} = [h_{SR,1},..,h_{SR,k},...,h_{SR,M}]$ is the fading channel from the source to the RIS, with $h_{SR,k} = g_{SR,k}e^{-j\theta_k}\sqrt{(d_{SR})^{-\delta}}$ is the channel coefficient related to the element k, and $\mathbf{h}_{RD} = [h_{RD,1},...,h_{RD,l},...,h_{RD,M}]^T$ is the fading channel from the RIS to the destination, with $h_{RD,l} = g_{RD,l}e^{-j\psi_{k,l}}\sqrt{(d_{RD})^{-\delta}}$ is the channel coefficient related to the element l, δ denotes the path loss exponent, $g_{SR,k}$ and $g_{RD,l}$ the channel gains for S-to-R through element k and R-to-D through element l, d_{SR} and d_{RD} the distance from the source to the RIS and RIS to the destination, respectively. Moreover, the correlation matrix $\mathbf{R}$ is defined as an MxM matrix, that defines the correlation coefficients $\rho_{i,j}$ between the ith and j^{th} elements of the RIS. Thus, $\forall i,j, 0 \leq \rho_{i,j} \leq 1$ and $\rho_{i,j} = 1$ for $i = j$. In addition, the reflection coefficients of the RIS are within the diagonal phase matrix $\Phi = \text{diag}[e^{-j\phi_1}, e^{-j\phi_2}, ..., e^{-j\phi_M}]$.

1) Source-Aware Eavesdropping: Jammer 1 eavesdrops directly on the source node and then attacks the destination node. To build an antifragile strategy, it is necessary to model the channel in the presence of a jammer. As a jammer could have several types of behavior, it needs to be modeled accordingly.

In order to consider the impact of Jammer 1 on the received signal, we denote by $\mathbf{h}_{E1}$ the eavesdropping CSI from the source to Jammer 1, and by $\mathbf{h}_{J1}$ the jamming CSI from Jammer 1 to destination, which is given by h_s [4], [9]:

$$h_s = \sqrt{\frac{\kappa}{\kappa+1}}\sigma e^{j\theta_{s,i}} + \sqrt{\frac{1}{\kappa+1}}\sum_{i=1}^{L}R_{s,i}e^{j\theta_{s,i}}, s = J1 \text{ or } E1$$
$$(2)$$

Where κ is the Rician factor, σ^2 the average power, L the number of paths, and $R_{s,i}$ the Rayleigh distributed amplitude and $\theta_{s,i}$ the uniformly distributed phase of the channel. When the Rician factor $\kappa = 0$, the Rician fading channel coefficients are simplified to the Rayleigh fading coefficients. The received signal, considering the presence of a jammer, is given by:

$$y(t) = \mathbf{h}_{SR}\,\mathbf{R}^{1/2}\,\mathbf{\Phi}\,\mathbf{R}^{1/2}\,\mathbf{h}_{RD}\,x(t) + y_{J1}(t) + w(t)$$
$$= \left(\sum_{a=1}^{M}\sum_{k=1}^{M}\sum_{l=1}^{M}h_{SR,k}\,h_{RD,l}\,\rho_{a,k}^{\frac{1}{2}}\rho_{a,l}^{\frac{1}{2}}\,e^{j\phi_a}\right)x(t)$$
$$+ A\left(\sum_{i=1}^{L}R_{E1,i}\,e^{j\theta_{E1,i}}\right)\left(\sum_{j=1}^{L}R_{J1,j}\,e^{j\theta_{J1,j}}\right)x(t-\tau_{J1}) + w(t)$$
$$(3)$$

With A being a multiplicative factor that can be defined based on the type of jamming attack (Table I), explained in Section III-C, and τ_{J1} denotes the delay of the jamming path through Jammer 1.

2) RIS-Aware Eavesdropping: We denote the eavesdropping CSI from the RIS to Jammer 2 by $\mathbf{h}_{RJ}$ and by $\mathbf{h}_{J2}$ the jamming CSI from Jammer 2 to destination, calculated using Eq. (2). The received signal $y(t)$ is a superposition of the legitimate signal defined by Eq. (1), and the jamming signal (Jammer 2), which is defined by $y_{J2}(t)$ and given as [6], [9],

$$y_{\mathrm{J2}}(t) = \left(\sum_{a=1}^{M} \sum_{k=1}^{M} \sum_{l=1}^{M} \rho_{a,k}^{1/2} \rho_{a,l}^{1/2} \, h_{SR,k} \, h_{RJ,l} \, e^{j\phi_a} \right)$$
$$A \left(\sum_{j=1}^{L} R_{\mathrm{J2},j} \, e^{j\theta_{\mathrm{J2},j}} \right) x(t - \tau_{\mathrm{J2}}) \quad (4)$$

where τ_{J2} denotes the delay of the jamming path through Jammer 2. Thus, the received signal $y(t)$ is given by:

$$y(t) = \mathbf{h}_{SR} \mathbf{R}^{1/2} \mathbf{\Phi} \mathbf{R}^{1/2} \mathbf{h}_{RD} x(t) + y_{\mathrm{J2}}(t) + w(t)$$
$$= \left(\sum_{a=1}^{M} \sum_{k=1}^{M} \sum_{l=1}^{M} \rho_{a,k}^{1/2} \rho_{a,l}^{1/2} \, h_{\mathrm{SR},k} \, h_{\mathrm{RD},l} \, e^{j\phi_a} \right) x(t)$$
$$+ \left(\sum_{a=1}^{M} \sum_{k=1}^{M} \sum_{l=1}^{M} \rho_{a,k}^{1/2} \rho_{a,l}^{1/2} \, h_{SR,k} \, h_{RJ,l} \, e^{j\phi_a} \right)$$
$$A \left(\sum_{j=1}^{L} R_{\mathrm{J2},j} \, e^{j\theta_{\mathrm{J2},j}} \right) x(t - \tau_{\mathrm{J2}})) + w(t) \quad (5)$$

It should be noted that Jammer 2 can also be a malicious RIS, able to make attacks using phase shifting or amplitude shifting, as an active malicious RIS could do. If the destination is equipped with m antennas, the resulting signals become:

$$\mathbf{Y} = \mathbf{s_v} \mathbf{y} + \mathbf{w(t)} \quad (6)$$

With $\mathbf{Y} \in \mathbb{C}^{m \times 1}$ for a single received symbol, $\mathbf{y} \in \mathbb{C}^{1 \times 1}$, $\mathbf{s_v} \in \mathbb{C}^{m \times 1}$ being the steering vector for the angle of arrival (AoA) and y being $y = h_1 x(t) + h_{\mathrm{J2}} x(t - \tau_{\mathrm{j}})$ for the single antenna and $y = h_1 x(t) + h_{\mathrm{J2}} x(t)$ for the multipath jammer.

C. Jammer Model

A reactive (repeater) jammer operates by capturing the waveform emitted by the legitimate transmitter, optionally applying a deterministic transformation, and retransmitting the modified signal with the objective of degrading the receiver's performance. We adopt three categories of jammer models: digital radio frequency memory (DRFM), phase shifting (PS) and Amplitude shifting (AS), which are summarized in Table I. The DRFM attack retransmits the target signal on a sample-by-sample basis with a constant amplification gain denoted as β_a. Assuming that h_{E1} and h_{J1} are defined using Eq. (2), the DRFM jamming signal can be expressed as follows:

$$y_J(t) = \beta_a h_{\mathrm{E1}} h_{\mathrm{J1}} x(t - \tau_{\mathrm{J1}}) + w(t) \quad (7)$$

The PS jammer transforms the signal, randomly inverting the phase of the intercepted symbols. We denote by $U(t)$ a random sequence drawn from the set $U(t) \in \{1, -1\}$ with probability mass function $f_U(u) = 0.5 : U = 1, -1$.

$$y_J(t) = U(t) h_{\mathrm{E1}} h_{\mathrm{J1}} x(t - \tau_{\mathrm{J1}}) + w(t) \quad (8)$$

Similarly, the AS jammer introduces random amplitude perturbations, defined as $V(t)$. Thus the signal is given by:

$$y_J(t) = V(t) h_{\mathrm{E1}} h_{\mathrm{J1}} x(t - \tau_{\mathrm{J1}}) + w(t) \quad (9)$$

Jammer Model	Factor A	Signal Manipulation
DRFM	β_a	Amplification using fixed β_a
PS	$U(t)$	Phase shift using $U(t) : u \in \{1, -1\}$
AS	$V(t)$	Amplitude amplification attenuation using $V(t) : 2 \geq v \geq 0$

Table I: Jammer models summarized [2].

IV. ANTIFRAGILITY SCHEME DESIGN

In order to achieve antifragile gain, several factors related to both the jammer and the legitimate signal must be considered. The system must first detect a jamming attack and estimate the jammers delay relative to the legitimate signal, τ_{J1} or τ_{J2}. Additionally, the signals must be received orthogonally, either in time or space, to avoid overlap that would hinder demodulation and decoding. Thus, this section is organized as follows: first, we present general system assumptions about jamming detection and initial delay estimation. Then, the orthogonality strategy is discussed, leading to jamming classification, and finally, signal, modulation, and coding adaptation.

A. Jamming Detection and Delay Estimation

The jamming signal can be detected by evaluating the Bit Error Rate (BER) of the received signal. Without loss of generality, we assume that Reed-Solomon (RS) code is used; other codes would need to be separately handled. The presence of a jammer is detected when the BER exceeds a certain threshold, rendering the RS code unable to decode the message. When this condition occurs, a Cross-Correlation (CC) analysis is performed on the signal. The CC metric is obtained by correlating the received sequence y with its time-aligned replica $\tilde{y}$; the presence of a peak is taken as evidence of a jammer. Because the delay estimate τ provides only coarse alignment-insufficient for symbol- or sample-level precision-we compute a complete cross-correlation (i.e., a sliding inner product) rather than a single dot product, given by [2]:

$$R_{y\tilde{y}}(\tau) = \sum_{n=0}^{F_{max}} y[n] \cdot \tilde{y}[n + \tau] \quad (10)$$

where F_{max} is the frame length. If a malicious signal is present, the CC analysis displays a secondary sharp peak, in dependence on the strength and similarity of the malicious signal to the original signal. This peak is then used as an initial estimate for the delay introduced by the jammer and is defined as the delay that maximizes the magnitude of the cross-correlation:

$$\hat{\tau} = \arg \max_{\tau \in [-\gamma, \gamma]} |R_{y\tilde{y}}(\tau)| \quad (11)$$

where γ is maximum anticipated delay value. Since most jammers operate in cycles, the receiver stores the timing of the first jamming attack and continues the transmission. If a second attack occurs, the system calculates the duration of the jamming cycle and can adapt the legitimate signal.

B. Orthogonality Approach

To ensure orthogonal reception of desired and jamming waveforms, we consider two operating modes: (i) spatial orthogonality via directional separation, and, (ii) temporal orthogonality via time-domain partitioning.

For spatial separation, we estimate the angle of arrival (AoA). Because first-order multipath reflections dominate mmWave links, the AoA distribution exhibits sharp peaks [12]. We therefore apply a forward-backward spatially smoothed MUSIC algorithm [13]. The resulting AoAs for both desired and jamming signals enable beamforming-based spatial filtering at the receiver, implemented with a linearly constrained minimum-variance (LCMV) beamformer [14].

The beamforming weights $\mathbf{w}$ are computed using the covariance matrix $\mathbf{R}$, the constraint matrix and vector $\mathbf{C}$ and $\mathbf{f}$, respectively:

$$\mathbf{w} = \mathbf{R}^{-1}\mathbf{C}\left(\mathbf{C}^H\mathbf{R}^{-1}\mathbf{C}\right)^{-1}\mathbf{f} \tag{12}$$

The legitimate and malicious signal can be reconstructed using the following equation, i.e.,

$$\mathbf{s} = \mathbf{w}^H\mathbf{X} \tag{13}$$

If the receiver cannot separate the signals spatially, temporal orthogonality must be established.

For temporal orthogonality, the duration of the legitimate signal is reduced based on the estimated jamming delay, to prevent an overlap. Let T_{active} denote the time it takes for a signal to trigger the jamming process and T_{OFF} the time for the jamming process to stop. The legitimate signal must fulfill the property:

$$T_{\text{active}} \leq D(s_{\text{legit}}) < \tau_{\text{j}} \tag{14}$$

, with $D(\cdot)$ is the duration of the signal, and τ_{j} the time it takes for the jamming signal s_{legit} to be received at the destination.

C. Jamming Classification

To differentiate among the attack types listed in Table I, the jammer's fading channel must first be estimated so that channel effects do not bias the classification. A maximum-likelihood estimator is employed for this purpose. After the initial estimate, the channel coefficient is refreshed and stored whenever a jamming waveform is observed.

For the classification of the DRFM jammer, we propose a stepwise classification process, which employs a novel similarity ratio normalizing cross-correlation by self-correlation (SC) to provide robust detection less sensitive to inherent signal structures. First, the self-correlation $R_{yy}(\tau)$ of the legitimate signal is calculated based on Eq. (10). Afterwards the SC

maximum value is normalized along the signal duration, and used as reference, i.e.,

$$SC_{max} = \frac{\max\left(|R_{yy}(\tau)|\right)}{F_{max}} \tag{15}$$

Next, the cross-correlation between the jamming signal and the legitimate signal is calculated, given by Eq. (10). The maximum cross-correlation CC_{max} value is also normalized.

Finally, the similarity ratio between the cross-correlation CC and the self-correlation SC is calculated as:

$$\text{Sim} = \frac{\max(|R_{y\hat{y}}(\tau)|}{\max(|R_{yy}(\tau)|)} \tag{16}$$

A similarity ratio that surpasses a predefined threshold leads to the classification of the interference as DRFM jamming.

For amplitude-shifting (AS) and phase-shifting (PS) jammers, the isolated waveform is first demodulated and correlated with pilot symbols. If the bit-inversion count surpasses a preset threshold, the jammer is classified as AS or PS according to the active modulation. A similarity ratio is likewise computed; since AS and PS randomly invert parts of the signal, their similarity to the legitimate waveform is significantly lower than that of a DRFM jammer.

If classification confidence is insufficient, the identification routine is executed again. Once a jammer type is conclusively determined, the receiver logs the category, the estimated delay and other relevant metrics and then forwards this information to the transmitter. The transmitter subsequently initiates the waveform and coding-adaptation procedure. Malicious RIS devices can be similarly recognized by the propagation delay they impose: both passive and active malicious RIS configurations introduce latencies typically shorter than those incurred by signal-processing reactive jammers.

D. Signal Adaptation

Antifragile gains can only be obtained when the legitimate waveform is modified such that the jammer affects none of its information dimensions. Once such orthogonality is secured, i.e., via temporal, spatial, or spectral separation, the jammer's emission can be exploited to convey additional information without corrupting the desired data. Signal adaptations therefore focus on preserving the legitimate signal's phase, amplitude, and other dimensions, while treating the jamming waveform as an independent, utilizable resource. For an AS jammer that perturbs the signal envelope, the system remaps its waveform to an M-PSK constellation, thereby removing amplitude dependence. With a PS jammer, the transmitter adopts ASK whose constellation points reside exclusively in the positive real half-plane, allowing any $180°$ phase inversions to be easily corrected. For a DRFM jammer no modulation change is necessary, since the strategy leverages the PSK signal's coherent addition with the jammer's delayed replica.

E. Modulation and Code Adaptation

Once the signal adaptation is completed, a new code rate is determined based on the increased SNR value. Therefore, we

denote the additional jamming SNR as SNR_J and legitimate SNR as SNR_L:

$$\text{SNR}_J = \frac{\gamma_E \gamma_J}{\gamma_E + \gamma_J + 1}, \qquad \text{SNR}_L = \gamma_L. \qquad (17)$$

Since (n, k) Reed-Solomon code can correct up to $t = (n - k)/2$ symbol errors, choosing the optimal code depends on the post-demodulation error profile. Thus, the measured SNR is used to compute the bit error rate (BER) for each adaptive modulation scheme, which is then converted into the residual symbol error rate $P_{\text{res}}^{\text{RS}}$ following the formulation in [15].

$$P_{\text{rs}}^{\text{RS}} = \frac{n \cdot \text{SER} - t}{n} \qquad (18)$$

with $\text{SER} = 1 - (1 - \text{BER})^{\log_2(M)}$. Using this factor, thresholds for each jamming category can be defined such that $P_{\text{rs}}^{\text{RS}} \leq \Delta < 0$, with Δ being a negative real number.

F. Antifragile gain

We denote the throughput in the baseline scenario by T_L and under jamming attack by T_J respectively. The optimal Reed-Solomon code is selected according to Eq. (18), which ensures that all symbol errors within each block are corrected. The throughput is then defined based on the code rate R_C, bandwidth B, modulation order M and SNR index i (Eq. 17):

$$T_i = B R_c^i \log_2(M^i), i = L \text{ or } J \qquad (19)$$

The system demonstrates antifragile gain only if its throughput during jamming surpasses its baseline, jammer-free throughput. To evaluate this gain across different SNRs, the jamming-to-signal ratio (JSR) is defined. It represents the ratio of jamming power P_J to legitimate signal power P_L:

$$\text{JSR} = 10 * \log_{10}(P_J) - 10 * \log_{10}(P_L) \qquad (20)$$

V. NUMERICAL EVALUATION

A. Assumptions

Fo simulations, we assume that is RIS positioned with respect to the receiver and transmitter at $d_{\text{SR}} = 18$ m and $d_{\text{RD}} = 7$ m, respectively. We set the path-loss exponent to $\delta = 2.7$ and the rate $\lambda_R = 0.05$. The transmitter sends the legitimate signal with a fixed power of 20 dBm on a 28 GHz carrier. Because the receiver's signal-to-noise ratio (SNR) can fluctuate significantly depending on the size of the RIS, the jamming-to-signal ratio (JSR) is defined based on Eq. 20. This choice confines the jammer output to the interval 0-40 dBm; an SNR-based definition would require substantially higher jammer powers to span the same JSR range at larger baseline SNRs. Adaptive modulation and coding (AMC) employs Reed-Solomon codes whose code rates vary from 0.70 ($\text{RS}(178, 255)$) to 0.94 ($\text{RS}(240, 255)$) based on Eq. 18. The modulation format is selected according to the attack detected, with orders of up to $M = 64$. Results are reported separately for source-aware and RIS-aware eavesdropping cases, and antifragile gains are emphasized with light blue shading. In all simulations, the channel coefficients were iterated 200 times, to obtain average values.

B. Source-Aware Eavesdropping

Figure 2 compares the three attacks with a fixed code rate. Without spatial orthogonality the jammer and desired signal overlap for half the frame; the transmitter shortens its burst (Sec. IV-B, Eq. 14), reducing payload and throughput (Eq. 19). Thus, antifragile gains appear only after raising the modulation order, i.e., at JSR = 3 dB for DRFM and 15 dB for AS. With full orthogonality (Eq. 12) and a higher baseline SNR, throughput exceeds the jammer, free reference from about -5 dB, DRFM allowing for the largest boost due to its high mutual information with the desired waveform. Figure 3 introduces adaptive RS coding based on the Eqs. 17-18 and a higher receiver SNR. Variable code rates enable both an earlier appearance and a larger magnitude of antifragile gain by pairing higher-order modulations (increasing M in Eq. 18) with moderately reduced code rates (R_C in Eq. 18), thereby enhancing error-correction capability and elevating the peak throughput. Figure 4 scales the RIS aperture. As the surface grows, the legitimate SNR rises faster than the fixed-power jammer can follow, so antifragile benefits fade; the PS attack offers no gain beyond 128 elements. Across the source-aware scenario DRFM delivers the highest data rates, followed by AS and PS, consistent with mutual-information limits under fixed jammer power and position.

(a) Without orthogonality

(b) With orthogonality

Figure 2: Throughput vs JSR for a fixed coding rate. Baseline SNR = 7 dB

Figure 3: Throughput vs. JSR for optimal selected coding rate. Baseline SNR = 10 dB

Figure 4: Throughput vs JSR for optimal coding rate and varying RIS sizes.

VI. CONCLUSION

This paper has presented the first study of *antifragility* in RIS-assisted wireless links under jamming attacks. We considered three attack models, i.e., DRFM, phase shifting (PS), amplitude shifting (AS). We proposed a receiver-side identification framework that was able to exploit spatial or temporal orthogonality to separate the desired and malicious waveforms, classifies the jammer type, and send the estimate back to the transmitter. A joint waveform-coding adaptation was devised: PSK/ASK remapping for AS and PS attacks, unchanged PSK for DRFM, and rate-adaptive Reed-Solomon coding that balances higher modulation orders against stronger error control. The results confirm that reconfigurable intelligent surfaces, when coupled with an antifragile design philosophy, can convert hostile interference into a throughput resource. This insight opens up a new design dimension for resilient wireless networks, where deliberate or accidental interference may be exploited rather than merely avoided.

C. RIS-Aware-Eavesdropping

In RIS-aware eavesdropping scenario, we keep the RIS location, code rate, and modulation order identical to source-aware case. Figure 5 presents the reactive-jammer results under the same baseline SNR and adaptive coding settings used in Fig. 4. Here, every jammer achieves a noticeably higher throughput. This outcome is anticipated because the adversary receives the RIS-reflected waveform and thereby benefits from its beam-forming gain; consequently, the effective jamming SNR scales more rapidly as the JSR grows.

REFERENCES

[1] N. N. Taleb and R. Douady, "Mathematical definition, mapping, and detection of (anti)fragility," *Quantitative Finance*, vol. 13, no. 11, 2013.

[2] M. Lichtman and et. al., "Antifragile communications," *IEEE Systems Journal*, vol. 12, no. 1, pp. 659–670, 2018.

[3] X. Tang, X. Lan, D. Zhai, R. Zhang, and Z. Han, "Securing wireless transmissions with ris-receiver coordination: Passive beamforming and active jamming," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 6, pp. 6260–6265, 2021.

[4] M. Lichtman and et. al., "Fsk-based reactive jammer piggybacking," *IEEE Communications Letters*, vol. 21, no. 1, pp. 68–71, 2017.

[5] B. Ji and et. al., "Multi-relay cognitive network with anti-fragile relay communication for intelligent transportation system under aggregated interference," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 7, pp. 7736–7745, 2023.

[6] N. Mensi and D. B. Rawat, "On the performance of partial ris selection vs. partial relay selection for vehicular communications," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 9, pp. 9475–9489, 2022.

[7] E. Björnson and L. Sanguinetti, "Rayleigh fading modeling and channel hardening for reconfigurable intelligent surfaces," *IEEE Wireless Communications Letters*, vol. 10, no. 4, pp. 830–834, 2020.

[8] M. Bensalem and A. Jukan, "Outage probability analysis of wireless paths with faulty reconfigurable intelligent surfaces," in *2024 IEEE Latin-American Conference on Communications (LATINCOM)*. IEEE, 11/6/2024 - 11/8/2024, pp. 1–6.

[9] C. Huang and et. al., "Multi-hop ris-empowered terahertz communications: A drl-based hybrid beamforming design," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 6, pp. 1663–1677, 2021.

[10] Y. Sun, K. An, J. Luo, Y. Zhu, G. Zheng, and S. Chatzinotas, "Intelligent reflecting surface enhanced secure transmission against both jamming and eavesdropping attacks," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 10, pp. 11017–11022, 2021.

[11] S. Arzykulov and et. al., "Artificial noise and ris-aided physical layer security: Optimal ris partitioning and power control," *IEEE Wireless Communications Letters*, vol. 12, no. 6, pp. 992–996, 2023.

[12] C. E. O'Lone, H. S. Dhillon, and R. M. Buehrer, "Characterizing the first-arriving multipath component in 5g millimeter wave networks: Toa, aoa, and non-line-of-sight bias," *IEEE Transactions on Wireless Communications*, vol. 21, no. 3, pp. 1602–1620, 2022.

[13] S. U. Pillai and B. H. Kwon, "Forward/backward spatial smoothing techniques for coherent signal identification," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 1, pp. 8–15, 1989.

[14] J. Bourgeois and W. Minker, *Time-Domain Beamforming and Blind Source Separation*. Boston, MA: Springer US, 2009, vol. 3.

[15] C. V. Phung and et. al., "Performance analysis of mdpc and rs codes in two-channel thz communication systems," in *2022 45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO)*. IEEE, 2022, pp. 482–487.

Figure 5: Throughput vs JSR for optimal selected coding rate. Baseline SNR = 10 dB