Defeating Super-Reactive Jammers With Deception Strategy: Modeling, Signal Detection, and Performance Analysis

Nguyen Van Huynh, Diep N. Nguyen, Dinh Thai Hoang, Thang X. Vu, Eryk Dutkiewicz, and Symeon Chatzinotas

Abstract

This paper aims to develop a novel framework to defeat a super-reactive jammer, one of the most difficult jamming attacks to deal with in practice. Specifically, the jammer has an unlimited power budget and is equipped with the self-interference suppression capability to simultaneously attack and listen to the transmitter's activities. Consequently, dealing with super-reactive jammers is very challenging. Thus, we introduce a smart deception mechanism to attract the jammer to continuously attack the channel and then leverage jamming signals to transmit data based on the ambient backscatter communication which is resilient to radio interference/jamming. To decode the backscattered signals, the maximum likelihood (ML) detector can be adopted. However, the method is notorious for its high computational complexity and require a specific mathematical model for the communication system. Hence, we propose a deep learning-based detector that can dynamically adapt to any channel and noise distributions. With the Long Short-Term Memory network, our detector can learn the received signals' dependencies to achieve the performance close to that of the optimal ML detector. Through simulation and theoretical results, we demonstrate that with proposed approaches, the more power the jammer uses to attack the channel, the better bit error rate performance we can achieve.

Index Terms

Anti-jamming, ambient backscatter communications, signal detection, reactive jammer, deep learning, LSTM, and physical layer security.

Nguyen Van Huynh, Diep N. Nguyen, Dinh Thai Hoang, and Eryk Dutkiewicz are with University of Technology Sydney, Australia. E-mails: huynh.nguyenvan@student.uts.edu.au, {Diep.Nguyen, Hoang.Dinh, and Eryk.Dutkiewicz}@uts.edu.au.

Thang X. Vu and Symeon Chatzinotas are with the Interdisciplinary Centre for Security, Reliability and Trust (SnT), University of Luxembourg, L-1855 Luxembourg. Emails: {thang.vu, symeon.chatzinotas}@uni.lu.

Preliminary results in this paper were presented at the IEEE WCNC Conference, 2020 [1].

I. Introduction

Wireless communications play an essential role in many areas by facilitating tetherless and ubiquitous transmissions through the broadcast medium. However, due to the exposed nature of wireless communications, current wireless networks are extremely vulnerable to jamming attacks. In particular, the jammer intentionally injects high-power interference signals to the target wireless channels to significantly reduce the signal-to-interference-plus-noise ratio (SINR) at the legitimate receiver, and thus it can disrupt or even bring down the whole system. Among all types of jamming attacks, the super-reactive jamming attack is one of the most difficult jamming attacks to deal with in practice. In particular, the super-reactive jammer has unlimited power budget and is equipped with the self-interference suppression (SiS) capability [2], [3] allowing it to simultaneously attack and listen to the activities of the transmitter on the target channel. Consequently, all current anti-jamming approaches such as frequency hopping [4]-[8], rate adaption [9], and deception [10], [11] cannot effectively defeat such attack since with virtually unlimited power, the super-reactive jammer can theoretically jam all channels over which it discerns activity of legitimate users (thanks to the jammer's SiS capability).

In this paper, we introduce a novel anti-jamming framework to effectively defeat super-reactive jamming attacks. This framework is composed of three main components: (i) intelligent deception strategy, (ii) smart multi-source backscatter communications, and (iii) a smart detection mechanism. Specifically, when a legitimate transmitter detects a reactive jamming attack on its communications channel, it will "pretend to be oblivious" and continue the data transmission to keep the jammer from continuing its attack. Then, the transmitter changes the communication method by activating the backscatter communications mode to transmit data instead of using the current active transmission mode. The ambient backscatter communication technology, unlike conventional wireless communications, is resilient to radio jamming interference/jamming as it can reflect or absorb the jamming signals to backscatter information to the receiver [12], [13]. To be more specific, the transmitter will use an auxiliary low-power tag device to backscatter its information simultaneously on its own transmitted signals *and* the jamming signals, referred to as multi-source backscattering communications. As such, this solution allows the system to communicate under super-reactive jamming attacks. Finally, to enhance the detection efficiency at the receiver, we develop a novel deep learning (DL) based detection mechanism. It is widely known that the maximum likelihood (ML) detection method is optimal in terms of bit error rate (BER). However, the ML detector requires sophisticated mathematical operations to model the system in advance [19]. As such, it cannot be directly applied to other scenarios with different channel and noise distributions (e.g., when the jammer adjusts its attack strategy or other anti-jamming systems in different wireless environments). Instead, our proposed DL-based detector does not require a specific model with sophisticated mathematical operations for each type of wireless channel, and thus it can be directly applied to other scenarios without any modification, e.g., when the reactive jammer adjusts its attack strategy or when the system is placed in different wireless environments. As a result, the proposed DL detector can be adopted in various anti-jamming systems with different channel and noise distributions as well as different jamming attack strategies.

Through simulation results and theoretical analysis, we demonstrate that the proposed framework can not only successfully defeat the super-reactive jammer, but also be able to “exploit” the jamming signals to improve the system performance. In particular, we first show that with our proposed framework, the backscatter rate and BER performance increases with the jamming power of the super-reactive jammer. As such, the proposed framework is resilient to the super-reactive jamming attacks. Furthermore, the proposed DL-based detector can achieve the BER performance close to that of the optimal ML detector without requiring information of the jammer in advance (e.g., jamming signals’ characteristics). Finally, through simulation results, it it revealed that the system performance can be significantly improved if we use multiple antennas at the receiver.

The rest of this paper is organized as follows. In Section II, we give an overview of related work in anti-jamming and signal detection. Section III and Section IV present the system model and the channel model, respectively. Section V discusses the decoding process at the receiver. Section VI introduces our proposed DL-based signal detector. After that, the evaluation results are discussed in Section VII. Finally, conclusions are drawn in Section VIII.

II. Related Work and Main Contributions

A. *Anti-Jamming Approaches*

Several anti-jamming solutions have been proposed in the literature [14]. Although these solutions are feasible in dealing with conventional jamming attacks, e.g., proactive jamming, they are not effective or even impractical in defeating super-reactive jamming attacks. In this section, we first highlight existing anti-jamming techniques and their limitations in coping with the super-reactive jamming.

1) Rate Adaptation: One common anti-jamming solution is the rate adaptation (RA) technique [9]. This technique allows the transmitter to adapt its transmission rate under jamming attacks. For example, when the jammer attacks the channel at a high power level, the transmitter can reduce its transmission rate to make it easier for the receiver to decode the information. In contrast, the transmitter can increase its transmission rate when the jammer attacks at low power levels. However, in [9], the authors demonstrated that the RA technology is not effective in dealing with smart jamming like reactive jamming, especially when the jamming power is very high. Additionally, the authors pointed out that in some cases, the transmitter fails to recover from the lowest data rate even when the jammer ceases its attacks. For the super-reactive jammer, the RA technique does not work well as the jammer can always attack the channel with a very high power level. More importantly, by simultaneously listening to activities of the transmitter while jamming by using recent advances in signal detection and sensing, e.g., self-interference suppression technology [2], [3], the jammer can quickly adapt its attack power as soon as the transmitter changes its transmission rate. Consequently, the RA technology is not possible in dealing with super-reactive jammers.

2) Frequency Hopping: Another anti-jamming solution widely adopting in existing studies is the frequency hopping (FH) technology [4]-[8]. The key idea of this solution is that the transmitter can switch to another channel when the jammer attacks the current channel. However, these solutions and other FH approaches in the literature inherit several drawbacks in dealing with jamming attacks, especially under reactive jamming. First, the FH technology requires a pre-shared key at both the transmitter and the receiver, resulting in the problem of scalability due to the need of distributing the shared key [4]. Second, under the reactive jamming, when the transmitter switches to a new communication channel, the reactive jammer can also quickly detect and attack the new channel. Finally, if the jammer has a very high power budget, e.g., the super-reactive jammer considered in this paper, to attack all the channels at the same time, the FH technology is no longer effective.

3) State-of-the-art Solutions: Recently, some state-of-the-art solutions have been proposed to overcome the limitations of conventional anti-jamming mechanisms. Specifically, in [22], the authors equipped the transmitter with the radio frequency (RF) energy harvesting capability to harvest energy from the jamming signals when the jammer attacks the channel. The harvested energy is then stored in energy storage and used to support the internal operations and active transmissions when the jammer is idle. Differently, the authors in [13] proposed to use the ambient backscatter communications [12], [15] to allow the transmitter to backscatter the jamming signals to transmit information. The authors demonstrated that by leveraging the jamming signals, the average throughput of the transmitter increases with the jamming power. However, all these solutions are only feasible in dealing with proactive jammers, i.e., the jammers attack the channel regardless of the transmitter's activities. In contrast, the reactive jammer can quickly cease its attacks when there is no active transmission from the transmitter. As such, the performance of these solutions in dealing with the super-reactive jammer is very limited.

Several solutions have been proposed recently to deal with reactive jammers. In [10], the authors proposed a deception tactic to deal with jamming attacks in a cognitive radio communication system. The key idea of this solution is that the secondary user can transmit "fake" signals to attract the jammer to attack the channel, and thus it can undermine the attack abilities by draining the power of the jammer. As the authors considered scenarios in which the jammer has a limited power budget, this solution works well because the jammer does not have enough power to attack when the secondary user transmits the actual information. However, if the jammer has a sufficient power budget (e.g., the super-reactive jammer considered in this paper), this solution is no longer efficient. Similarly, in [11], the authors proposed a deception-based anti-jamming framework for IoT networks. In particular, the transmitter can choose either to actively transmit or perform deception at the beginning of the time slot. If the transmitter performs deception, it will first generate "fake" signals to lure the jammer. After that, once the jammer attacks the channel, the transmitter can harvest energy from the jamming signals or backscatter information to the receiver through the jamming signals. Nevertheless, for reactive jammers equipped with the self-interference suppression technology [3], this solution is not effective. The reason is that the reactive jammer can simultaneously attack the channel and listen to the activities of the transmitter. As soon as the transmitter ceases its active transmissions, the jammer will stop its jamming attacks, and thus there is no jamming signals for the transmitter to leverage, i.e., harvest energy and backscatter information. To the best of our knowledge, all current anti-jamming solutions cannot effectively deal with reactive jammers, especially the super-reactive jammer which has self interference suppression ability and is equipped with an unlimited power budget.

B. Deep Learning for Signal Detection

There are a few research works using deep learning for signal detection [16]-[19]. In [16], the authors proposed a deep neural network to estimate the wireless channel and detect signals in orthogonal frequency-division multiplexing systems. In [17], the authors focused on signal detection for radio communication signals by using a one-dimension convolutional neural network architecture. Nevertheless, to the best of our knowledge, there is no work considering deep learning for signal detection in ambient backscatter communication systems. The main challenge when detecting backscattered bits is that the backscattered signals are very weak with low data rates. In addition, the received signals at the receiver are a combination of the backscattered signals from the backscatter tag and the direct signals from the RF source. Thus, classifying the backscattered signals is very challenging. To address this problem, in this work we propose to adopt the LSTM network, which is designed to learn the sequence and time-series data, to learn the long-term dependencies between the received signals at different antennas over several RF source symbols for each backscattered bit. By doing this, the receiver can better estimate the weak backscattered signals because the LSTM network can store information learned from previous input data (i.e., received signals) and combine them to achieve better estimation performance. Moreover, the proposed DL-based signal detector can be straightforwardly applied to other scenarios, e.g., when the reactive jammer changes its attack strategy or other anti-jamming systems in different wireless environments. To the best of our knowledge, the proposed detector is the first solution using deep learning in signal detection for ambient backscatter communications in which the backscattered signals are usually weak.

C. The Novelty and Main Contributions

The major contributions of this paper can be summarized as follows.

- We propose a novel anti-jamming framework to defeat super-reactive jammers who has unlimited power budget and is equipped with the self-interference suppression capability to simultaneously attack and listen to the activities of the transmitter on the target channel. Specifically, when the transmitter detects jamming attacks on the channel, it will “pretend to be oblivious” and continue to transmit data to lure/trap the jammer. Then, the transmitter switches to the backscatter communication mode by using the ambient backscatter tag to send the actual information by simultaneously backscattering the jammer’s signals and the transmitter’s signals. As such, the transmitter can communicate with the receiver even under very strong jamming attacks. To the best of our knowledge, this is the first anti-jamming framework that can efficiently deal with super-reactive jamming attacks.
- We introduce the receiver design with multiple antennas to leverage the diversity gain at the receiver, and thus enhancing the backscattered signals (which are usually very weak). We then present the signal models of direct links (from the jammer and the transmitter) and the backscattered link (from the tag). Finally, we theoretically quantify the maximum achievable backscatter rate of the ambient backscatter tag.

- We propose a novel DL-based signal detector by using the LSTM network to predict the backscatter bit at the receiver. With this DL-based solution, the proposed signal detector does not require the specific channel model to formulate the detection hypotheses with sophisticated mathematical models. As such, our proposed DL-based signal detector can be straightforwardly applied to other scenarios, e.g., when the reactive jammer changes its attack strategy or other anti-jamming systems in different wireless environments. To the best of our knowledge, the proposed detector is the first solution using deep learning in signal detection for ambient backscatter communications.
- Finally, we perform extensive simulations with the aims of not only demonstrating the efficiency of proposed solutions but also providing insightful analytical results for the implementation of our framework. In particular, the proposed framework can allow the transmitter to transmit data even under very strong and super-reactive jamming attacks. Importantly, by using our proposed solution, the BER performance of the system can be enhanced as the attack power of the jammer increases.

Notation: The notations used in this paper are listed in the following. Lowercase letter g denotes a scalar variable. Boldface lowercase letter $\mathbf{g}$ denotes a vector. Boldface uppercase letter $\mathbf{G}$ denotes a matrix. $\mathcal{CN}(\mu, \xi^2)$ is the circularly symmetric complex Gaussian (CSCG) distribution with variance ξ^2 and mean μ. $\mathbf{I}_M$ is the $M \times M$ identical matrix. $\mathbf{G}^{\mathrm{T}}$ denotes the transpose of matrix $\mathbf{G}$. $\mathbf{G}^{\mathrm{H}}$ is the conjugate transpose of matrix $\mathbf{G}$. $\mathbb{E}[.]$ is the statistical expectation. $\odot$ denotes the Hadamard product, i.e., element-wise multiplication of vectors. $I(X;Y)$ denotes the mutual information of random variables X and Y. $H(X|Y)$ is the conditional entropy. $\oplus$ is the addition modulo 2 operator.

III. System Model

In this paper, we consider a legitimate wireless communication system which includes a transmitter and a receiver. The transmitter transmits data to the receiver on a dedicated channel. A super-reactive jammer, which is located near the legitimate system, aims to disrupt/bring down the communications between the transmitter and the receiver by injecting strong interference signal to the dedicated channel.

A. Jammer Model

We consider that the super-reactive jammer only attacks the channel once it detects active transmissions from the transmitter. In addition, with state-of-the-art technologies such as self-interference suppression [2], [3], the reactive jammer can simultaneously attack the channel and

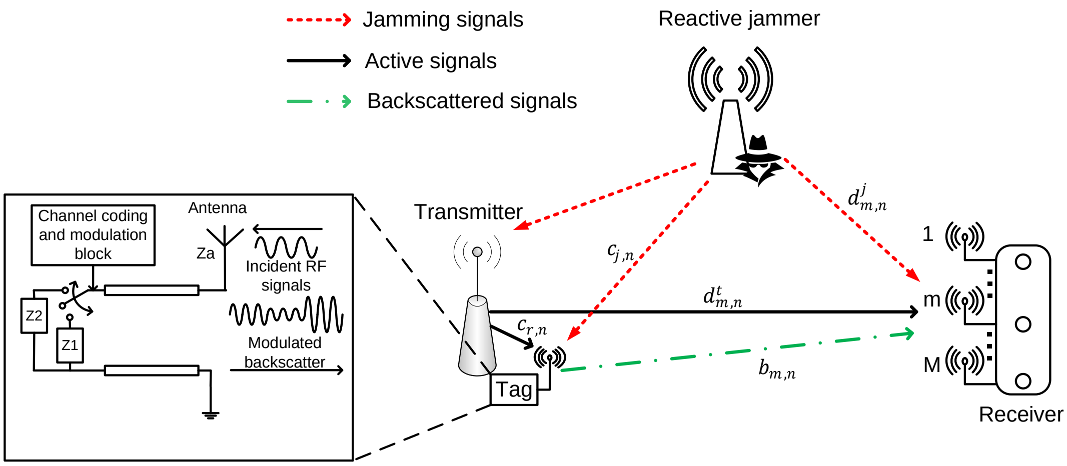


Fig. 1: System model.

listen to the activities of the transmitter. In particular, this technique allows a wireless device to suppress its self-interference up to the noise floor, and thus wireless signals can be transmitted and received on the same frequency channel [3]. As such, the reactive jammer can detect the status of the transmitter to cease the jamming attacks when the transmitter stops transmitting on the channel or quickly launch jamming attacks when the transmitter actively transmits the very first bits on the channel. Given this type of jamming attacks, the deception strategies proposed in existing works (e.g., [10], [11]) are no longer effective because the reactive jammer can quickly stop its attacks when the transmitter finishes the deception process, and thus there are no jamming signals for the transmitter to leverage. Moreover, with the unlimited power budget, the super-reactive jammer can simultaneously inject high-power interference to multiple channels. Hence, conventional approaches like RA and FH technologies also cannot defend from the considered jamming attacks, as discussed in the previous section.

B. Transmitter Design

To defeat the reactive jammer, at the transmitter's side, we implement an auxiliary low-power ambient backscatter tag equipped with ambient backscatter communication technology, as illustrated in Fig. 1. In particular, if the reactive jammer attacks the dedicated channel, the transmitter can detect the jamming attacks by using common mechanisms such as measuring the signal strength on the channel, based on the feedback message from the receiver, or the packet delivery ratio [34]. Upon detecting the jamming attacks, the transmitter keeps its transmissions to attract the reactive jammer. Concurrently, the transmitter switches the transmission mode by using the ambient backscatter tag. Then, the tag uses the ambient backscatter communication technology to send the actual data by backscattering the transmitter's signals and the jammer's signals. In other words, the jammer and the transmitter are considered as two ambient RF sources that are leveraged to backscatter information to the receiver [33].

To be more specific, the ambient backscatter tag modulates the information sent from the transmitter and then backscatters the transmitter's signals and the jammer's signals to send the information to the receiver. To that end, the ambient backscatter tag is equipped with an RF switch, e.g., ADG902, that is directly connected to the antenna of the tag, as illustrated in Fig. 1. The bit stream sent from the transmitter is first modulated by using the modulo-2 modulation technique. When the tag wants to transmit bits "0", it will switch to load Z_1 to change to the non-reflecting state (also known as the absorbing state). In contrast, the tag switches to load Z_2 to change to the reflecting state to transmit bits "1". In this way, the modulated information can be backscattered to the receiver by just switching between the non-reflecting and reflecting states. Thus, unlike conventional wireless communications, the ambient backscatter communication technology is resilient to radio interference/jamming as it can absorb or reflect the jamming signals to transmit information to the receiver. This is a fundamental principle of backscatter communication systems, and interested readers can find more information from the following references [15], [12]. It is worth noting that the ambient backscatter tag conveys information to the receiver by reflecting/observing the ambient RF signals instead of generating active RF signals like the conventional active transmission. As a result, even under strong jamming attacks, our proposed method still allows the system to communicate by backscattering the jammer's signals and the transmitter's signals. The process of backscattering and decoding information under super-reactive jamming attacks will be presented in details in Section IV and Section V.

C. *Receiver Design*

As mentioned, the received signals contain the backscattered signals and the direct link signals from the transmitter and the jammer, resulting in a low signal detection accuracy. To address this problem, we employ M antennas[1] at the receiver. As such, the receiver can exploit the diversity gain, and thus improving the BER performance. More importantly, we propose a deep learning based signal detector that can apply to any channel and noise distribution and achieve the BER performance close to that of the optimal ML detector. The process of decoding the backscattered

[1] Our proposed approach still works well even the receiver equipped with a single antenna. Results for this case can be found in the simulation section.

signals will be presented in Section V and Section VI in details. It is worth noting that the receiver can also detect the jamming attacks from the jammer and switch its reception mode to receive and decode the backscattered signals. Moreover, the proposed model and analysis in this paper can be straightforwardly adopted in other scenarios with multiple backscatter tags. To that end, the receiver can deploy popular scheduling mechanisms to eliminate the collisions between tags. Another potential solution is backscattering data at different backscatter rates [12], [15]. In this way, the receiver can easily distinguish the backscattered signals from different tags. More importantly, our proposed approach can also directly apply to scenarios with multiple jammers. In particular, these jammers are also considered as ambient RF sources to support the backscatter communications of the tag with the same signal model and analysis of the jammer considered in this paper.

IV. Channel Model

To make favorable conditions for the receiver to successfully decode the backscattered signals from the tag, the backscatter rate should be lower than the sampling rate of ambient RF source signals [12], [15]. As such, in this work, we assume that the ambient backscatter tag backscatters data at the rate N times lower than the sampling rate of the ambient RF signals [20], [27]. In other words, each backscatter symbol will be backscattered over N RF source symbols. In the system under consideration, the received signals at the receiver include the direct link signals from the transmitter and the jammer, the backscattered signals from the tag, and noise. Denote y_{mn} as the n-th received sample at the m-th antenna of the receiver. We have

$$y_{mn} = \underbrace{d^{(\mathrm{t})}_{mn} + d^{(\mathrm{j})}_{mn}}_{\text{direct links}} + \underbrace{b_{mn}}_{\text{backscatter links}} + \sigma_{mn}, \tag{1}$$

where $d^{(\mathrm{t})}_{mn}$ and $d^{(\mathrm{j})}_{mn}$ are the direct link signal from the transmitter and the jammer, respectively. b_{mn} is the backscattered signal from the backscatter tag and σ_{mn} is the CSCG noise with zero mean and unit variance, i.e., $\sigma_{mn} \sim \mathcal{CN}(0,1)$.

A. Direct Links

Denote s_{jn} and s_{tn} as the signals generated from the jammer and the transmitter at time instant n, respectively. Similar to other research works in the literature, s_{jn} and s_{tn} are assumed to be independent and identically distributed (i.i.d) at every time instant n. Note that in this paper, the tag considers the transmitter as an ambient RF source to support its backscatter communications. We assume that the RF signals generated from the transmitter during the jamming attacks are unknown at the receiver and are random (e.g., the transmitter actively transmits random/blank signals to attract the jammer and save energy). Thus, we assume that s_{tn} follows the standard CSCG distribution, i.e., $s_{tn} \sim \mathcal{CN}(0,1)$ [23]. Moreover, similar to other studies [20], [24], [25], [26], we assume that s_{jn} follows the standard CSCG distribution as the jamming signals are also usually random, i.e., $s_{jn} \sim \mathcal{CN}(0,1)$[2].

Then, the direct link signals from the jammer received at the m-th antenna of the receiver are expressed as follows [20], [27]:

$$d_{mn}^{(\mathrm{j})} = f_{jm}\sqrt{P_{tj}}s_{jn}, \tag{2}$$

where f_{jm} denotes the Rayleigh fading [32] from the jammer to the receiver. Without loss of generality, let $\mathbb{E}[|f_{jm}|^2] = 1$. P_{tj} presents the average received power of the direct link from the jammer to the receiver. P_{tj} can be obtained as follows [21]:

$$P_{tj} = \frac{\kappa P_j G_j G_r}{{L_j}^{\upsilon}}, \tag{3}$$

where P_j is the transmit power of the jammer, G_j is the antenna gain of the jammer, G_r is the antenna gain of the receiver, and L_j is the distance between the receiver and the jammer. υ is the path loss exponent and $\kappa = (\frac{\lambda}{4\pi})^2$ with λ is the wavelength.

In the same way, the direct link signals from the transmitter received at the m-th antenna of the receiver are also expressed as follows:

$$d_{mn}^{(\mathrm{t})} = f_{rm}\sqrt{P_{tr}}s_{tn}, \tag{4}$$

where f_{rm} denotes the Rayleigh fading [32] from the transmitter to the receiver. Without loss of generality, let $\mathbb{E}[|f_{rm}|^2] = 1$. P_{tr} is the average received power of the direct link from the transmitter to the receiver. P_{tr} can be calculated as follows:

$$P_{tr} = \frac{\kappa P_t G_t G_r}{{L_r}^{\upsilon}}, \tag{5}$$

where P_t presents the transmit power of the transmitter, L_r denotes the distance between the receiver and the transmitter, and G_t is the antenna gain of the transmitter.

[2] Note that our proposed deep learning-based detector does not require this information in advance. The proposed deep learning-based detector can be directly applied for any other channel distributions

B. Backscatter Link

The ambient backscatter tag can backscatter the RF signals from both the reactive jammer and the transmitter to send data to the receiver by using the ambient backscatter communication technology [12], [15]. In the following, we formulate the backscattered signals received at the receiver in details. Specifically, the jamming signals received at the backscatter tag are expressed as follows:

$$c_{jn} = l_j \sqrt{P_{bj}} s_{jn}, \tag{6}$$

where l_j denotes the Rayleigh fading [32] from the jammer to the backscatter tag. Without loss of generality, let $\mathbb{E}[|l_j|^2] = 1$. P_{bj} is the average jamming power received at the backscatter tag. P_{bj} is calculated as follows:

$$P_{bj} = \frac{\kappa P_j G_j G_b}{{L_s}^v}, \tag{7}$$

where L_s denotes the distance between the backscatter tag and the jammer and G_b is the tag's antenna gain. Similarly, the transmitter's signals received at the backscatter tag are derived as follows:

$$c_{rn} = l_r \sqrt{P_{br}} s_{tn}, \tag{8}$$

where l_r is the Rayleigh fading [32] from the transmitter to the tag. Without loss of generality, let $\mathbb{E}[|l_r|^2] = 1$. P_{br} denotes the average power from the transmitter received at the backscatter tag and can be calculated as follows:

$$P_{br} = \frac{\kappa P_t G_t G_b}{{L_b}^v}, \tag{9}$$

where L_b is the distance between the tag and the transmitter.

Denote c_n as the total RF signals received at the backscatter tag. From (6) and (8), we have

$$c_n = c_{rn} + c_{jn} = l_r \sqrt{P_{br}} s_{tn} + l_j \sqrt{P_{bj}} s_{jn}. \tag{10}$$

From the fundamental of the ambient backscatter communication technology [15], [12], the backscatter tag has two states (i.e., *on-off keying (OOK)*) including the reflecting state denoted by $e = 1$ and the non-reflecting state (i.e., absorbing state) denoted by $e = 0$. As mentioned, the backscatter rate must be N times lower than the sampling rate of the RF sources. Thus, state e remains unchanged during N source symbols. Denote $s_{b,n}$ as the backscattered signals from the backscatter tag, we have

$$s_{b,n} = \gamma c_n e, \tag{11}$$

where γ denotes the reflection coefficient which combines all the effects of the load impedance, antenna impedance, and other components at the ambient backscatter tag [27].

The backscatter then backscatters signals to the receiver to send the actual information. We then formulate the backscattered signals received at the m-th antenna of the receiver as follows:

$$b_{mn} = f_{bm}\sqrt{\frac{G_b G_r \kappa}{{L_e}^{\delta}}} s_{b,n}, \tag{12}$$

where L_e denotes the distance between the receiver and the backscatter tag, δ is the path loss exponent, and f_{bm} is the Rayleigh fading [32] from the tag to the receiver. Without loss of generality, let $\mathbb{E}[|f_{bm}|^2] = 1$. Substituting (10) and (11) into (12), we have

$$\begin{aligned} b_{mn} &= f_{bm}\sqrt{\frac{G_b G_r \kappa}{{L_e}^{\delta}}}\gamma e\Big(l_r\sqrt{P_{br}} s_{tn} + l_j\sqrt{P_{bj}} s_{jn}\Big) \\ &= f_{bm} e\Big(l_r\sqrt{\frac{\kappa|\gamma|^2 P_{tr} {G_b}^2 {L_r}^{v}}{{L_b}^{v}{L_e}^{\delta}}} s_{tn} + l_j\sqrt{\frac{\kappa|\gamma|^2 P_{tj} {G_b}^2 {L_j}^{v}}{{L_s}^{v}{L_e}^{\delta}}} s_{jn}\Big). \end{aligned} \tag{13}$$

Denote $\tilde{\alpha}_r = \frac{\kappa|\gamma|^2 {G_b}^2 {L_r}^{v}}{{L_b}^{v}{L_e}^{\delta}}$ and $\tilde{\alpha}_j = \frac{\kappa|\gamma|^2 {G_b}^2 {L_j}^{v}}{{L_s}^{v}{L_e}^{\delta}}$, we can rewrite (13) as follows:

$$b_{mn} = f_{bm} e\Big(l_r\sqrt{\tilde{\alpha}_r P_{tr}} s_{tn} + l_j\sqrt{\tilde{\alpha}_j P_{tj}} s_{jn}\Big). \tag{14}$$

C. Received Signals at the Receiver

Substituting (4), (2), and (14) to (1), we have

$$\begin{aligned} y_{mn} &= d^{(\mathrm{t})}_{mn} + d^{(\mathrm{j})}_{mn} + b_{mn} + \sigma_{mn} \\ &= f_{rm}\sqrt{P_{tr}} s_{tn} + f_{jm}\sqrt{P_{tj}} s_{jn} + f_{bm} e\Big(l_r\sqrt{\tilde{\alpha}_r P_{tr}} s_{tn} + l_j\sqrt{\tilde{\alpha}_j P_{tj}} s_{jn}\Big) + \sigma_{mn}. \end{aligned} \tag{15}$$

Note that as we consider $\sigma_{mn} \sim \mathcal{CN}(0,1)$, the noise power equals to 1. Thus, we denote $\alpha_{dj} \triangleq P_{tj}$ and $\alpha_{dt} \triangleq P_{tr}$ as the average signal-to-noise ratio (SNR) of the direct link from the jammer to the receiver and the average SNR of the direct link from the transmitter to the receiver, respectively. Similarly, we denote $\alpha_{bj} \triangleq \tilde{\alpha}_j P_{tj}$ and $\alpha_{bt} \triangleq \tilde{\alpha}_r P_{tr}$ as the average SNR of the backscatter links (jammer-tag-receiver and transmitter-tag-receiver, respectively). Thus, (15) can be rewritten as follows:

$$y_{mn} = \underbrace{f_{rm}\sqrt{\alpha_{dt}} s_{tn} + f_{jm}\sqrt{\alpha_{dj}} s_{jn}}_{\text{direct links}} + \underbrace{f_{bm} e\Big(l_r\sqrt{\alpha_{bt}} s_{tn} + l_j\sqrt{\alpha_{bj}} s_{jn}\Big)}_{\text{backscatter links}} + \sigma_{mn}. \tag{16}$$

As the receiver is equipped with M antennas, the channel response vector can be defined as follows:

$$\mathbf{f}_r = [f_{r1}, \ldots, f_{rm}, \ldots, f_{rM}]^{\mathsf{T}}, \tag{17}$$

$$\mathbf{f}_j = [f_{j1}, \ldots, f_{jm}, \ldots, f_{jM}]^{\mathsf{T}}, \tag{18}$$

$$\mathbf{f}_b = [f_{b1}, \dots, f_{bm}, \dots, f_{bM}]^\mathsf{T}, \tag{19}$$

As such, the total signals received at M antennas can be expressed as follows:

$$\mathbf{y}_n = \underbrace{\mathbf{f}_r \sqrt{\alpha_{dt}} s_{tn} + \mathbf{f}_j \sqrt{\alpha_{dj}} s_{jn}}_{\text{direct links}} + \underbrace{\mathbf{f}_b e \Big(l_r \sqrt{\alpha_{bt}} s_{tn} + l_j \sqrt{\alpha_{bj}} s_{jn} \Big)}_{\text{backscatter links}} + \boldsymbol{\sigma}_n, \tag{20}$$

where

$$\mathbf{y}_n = [y_{1n}, \dots, y_{mn}, \dots, y_{Mn}]^\mathsf{T}. \tag{21}$$

$$\boldsymbol{\sigma}_n = [\sigma_{1n}, \dots, \sigma_{mn}, \dots, \sigma_{Mn}]^\mathsf{T}. \tag{22}$$

In this paper, each backscatter frame $\mathbf{b}$ is assumed to contain I information bits as follows:

$$\mathbf{b} = [b^{(1)}, \dots, b^{(i)}, \dots, b^{(I)}], \tag{23}$$

where $b^{(i)} \in \{0, 1\}, \forall i = 1, 2, \dots, I$. For simplicity, we assume that the channel remains invariant during one backscatter frame. Before being backscattered, each original bit is decoded as follows [20]:

$$e^{(i)} = e^{(i-1)} \oplus b^{(i)}, \tag{24}$$

where $\oplus$ is the modulo-2 operator and $\mathbf{e} = [e^{(1)}, \dots, e^{(i)}, \dots, e^{(I)}]$ is the modulated symbol vector in which $e^{(0)} = 1$ is the reference symbol. By using the modulo-2 operator, the probabilities of backscattering bit 0 and 1 are equal. Thus, we can achieve the maximum achievable backscatter rate as demonstrated in Section VII. Note that, each backscatter symbol $e^{(i)}$ is backscattered over N RF source symbols as illustrated in Fig. 2. As such, the received signal during the i-th backscatter symbol period can be obtained as follows:

$$\mathbf{y}_n^{(i)} = \mathbf{f}_r \sqrt{\alpha_{dt}} s_{tn}^{(i)} + \mathbf{f}_j \sqrt{\alpha_{dj}} s_{jn}^{(i)} + \mathbf{f}_b e^{(i)} \Big(l_r \sqrt{\alpha_{bt}} s_{tn}^{(i)} + l_j \sqrt{\alpha_{bj}} s_{jn}^{(i)} \Big) + \boldsymbol{\sigma}_n^{(i)}, \tag{25}$$

where $n = 1, 2, \dots, N$ and $i = 1, 2, \dots, I$.

Next, we present the theoretical analysis to obtain the maximum achievable backscatter rate. First, the backscatter rate of the backscatter tag when using the OOK modulation technique (i.e., non-reflecting and reflecting mechanism) is denoted as R_b. However, obtaining the close-form of the backscatter rate is impossible as the ambient RF signals are unknown and usually random [12], [20]. Hence, in Theorem 1, we obtain the maximum achievable backscatter rate $R_\mathrm{b}^\dagger$ of the backscatter tag to evaluate the performance of the system.

Theorem 1. *The maximum achievable backscatter rate $R_b^\dagger$ of the backscatter tag can be numerically obtained as follows:*

$$R_b^\dagger = C(\theta_0) - \mathbb{E}_{\mathbf{y}_0}[C(\omega)] = C(\theta_0) - \int_{\mathbf{y}_0} (\theta_0 p(\mathbf{y}_0 | e = 0) + \theta_1 p(\mathbf{y}_0 | e = 1)) C(\omega_0) d\mathbf{y}_0, \tag{26}$$

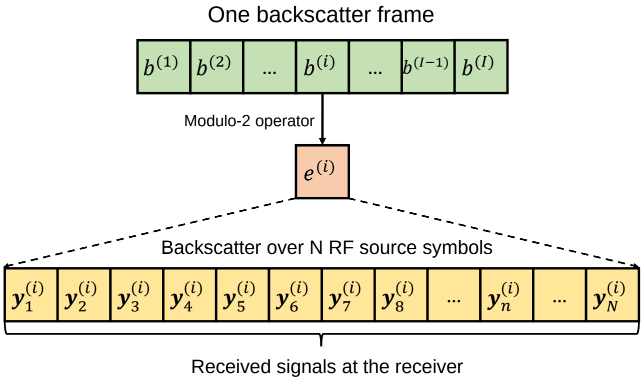


Fig. 2: Backscatter frame.

where θ_0 *is the prior probability when backscattering bits 0 (the prior probability when backscattering bits 1 is* $\theta_1 = 1 - \theta_0$*),* $\mathbf{y}_0$ *is a realization of* $\mathbf{y}$*,* C *is the binary entropy function, and* $p(e = j|\mathbf{y}_0), j \in \{0, 1\}$ *is the posterior probability of backscattered bit* e *given the received signal* $\mathbf{y}_0$.

Proof. The proof of Theorem 1 is provided in Appendix A. □

Denote $\mathbf{Y}^{(i)} = [\mathbf{y}_1^{(i)}, \dots, \mathbf{y}_n^{(i)}, \dots, \mathbf{y}_N^{(i)}]^\mathsf{T}$ as the received signal sequence during the i-th backscatter symbol period. In the next sections, we present the process of recovering all original bits from $\mathbf{Y}^{(i)}$ by using the ML detector and our proposed DL-based detector.

V. Decoding Backscattered Information at the Receiver with Maximum Likelihood Detector

As mentioned, to the best of our knowledge, there is no solution in the literature that can deal with the considered super-reactive jammer. Our proposed deception solution can enable the transmitter to lure/trap the jammer to allow the backscatter tag to leverage the RF signals from both the jammer and the transmitter to send the actual information to the receiver. However, decoding the weak backscattered signals is very challenging as the received signals are combinations of the backscattered signals and the direct link signals from the transmitter and the jammer. In the following, we present how the receiver can decode the backscattered signals by using the optimal ML detector. It is worth noting that the ML detector is well known as an optimal signal detector that can achieve high and stable BER performance, and thus it can be used as an upper-bound in this work.

A. Likelihood Functions of Received Signals

The fundamental task of the ML detector is deriving the likelihood functions of the received signals at the receiver. Specifically, when the backscatter tag backscatters bits "0", i.e., $e^{(i)} = 0$, there is no backscattered signals in the received signals. Otherwise, the received signals contain both the direct link signals and the backscattered signals when the tag backscatters bits "1", i.e., $e^{(i)} = 1$. Thus, we first derive the channel statistical covariance matrices corresponding to these cases as follows [20], [27]:

$$\begin{aligned} \mathbf{K}_1 &= (\mathbf{h}_1 + \mathbf{h}_2)(\mathbf{h}_1 + \mathbf{h}_2)^{\mathrm{H}} + (\mathbf{h}_3 + \mathbf{h}_4)(\mathbf{h}_3 + \mathbf{h}_4)^{\mathrm{H}} + \mathbf{I}_M, \\ \mathbf{K}_0 &= \mathbf{h}_1\mathbf{h}_1^{\mathrm{H}} + \mathbf{h}_3\mathbf{h}_3^{\mathrm{H}} + \mathbf{I}_M, \end{aligned} \tag{27}$$

where $\mathbf{h}_1 = \mathbf{f}_r\sqrt{\alpha_{dt}}$, $\mathbf{h}_2 = l_r\mathbf{f}_b\sqrt{\alpha_{bt}}$, $\mathbf{h}_3 = \mathbf{f}_a\sqrt{\alpha_{dj}}$, $\mathbf{h}_4 = l_j\mathbf{f}_b\sqrt{\alpha_{bj}}$, $\mathbf{I}_M$ is the $M \times M$ identity matrix, and $\mathbf{A}^{\mathrm{H}}$ is the conjugate transpose of matrix $\mathbf{A}$. $\mathbf{K}_1$ is the channel statistical covariance matrix when backscattering bits "1", and $\mathbf{K}_0$ is the channel statistical covariance matrix when backscattering bits "0". Note that the noise and the RF signals generated from the transmitter and the jammer follow the CSCG distribution[3]. As such, $\mathbf{y}_n^{(i)}$ is a CSCG distributed vector. Given backscatter symbol $e^{(i)}$ and received signals $\mathbf{y}_n^{(i)}$, we can derive the conditional probability density functions (PDFs) corresponding to $e^{(i)} = 0$ and $e^{(i)} = 1$ as follows:

$$\begin{aligned} p(\mathbf{y}_n^{(i)}|e^{(i)} = 0) &= \frac{1}{\pi^M|K_0|}e^{-\mathbf{y}_n^{(i)\mathrm{H}}K_0^{-1}\mathbf{y}_n^{(i)}}, \\ p(\mathbf{y}_n^{(i)}|e^{(i)} = 1) &= \frac{1}{\pi^M|K_1|}e^{-\mathbf{y}_n^{(i)\mathrm{H}}K_1^{-1}\mathbf{y}_n^{(i)}}. \end{aligned} \tag{28}$$

Based on (28), the likelihood functions of the received signals sequence $\mathbf{Y}^{(i)} = [\mathbf{y}_1^{(i)}, \ldots, \mathbf{y}_n^{(i)}, \ldots, \mathbf{y}_N^{(i)}]^{\mathsf{T}}$ given $e^{(i)} = 0$ and $e^{(i)} = 1$ can be expressed as follows [20]:

$$\begin{aligned} \mathcal{L}(\mathbf{Y}^{(i)}|e^{(i)} = 0) &= \prod_{n=1}^{N}\frac{1}{\pi^M|K_0|}e^{-\mathbf{y}_n^{(i)\mathrm{H}}K_0^{-1}\mathbf{y}_n^{(i)}}, \\ \mathcal{L}(\mathbf{Y}^{(i)}|e^{(i)} = 1) &= \prod_{n=1}^{N}\frac{1}{\pi^M|K_1|}e^{-\mathbf{y}_n^{(i)\mathrm{H}}K_1^{-1}\mathbf{y}_n^{(i)}}. \end{aligned} \tag{29}$$

B. Maximum Likelihood Detector

Based on (29), the original symbol $b^{(i)}$ can be obtained through the following ML criterion (i.e., hypothesis) [20], [27]:

$$\hat{e}^{(i)} = \begin{cases} 0, & \mathcal{L}(\mathbf{Y}^{(i)}|e^{(i)} = 0) > \mathcal{L}(\mathbf{Y}^{(i)}|e^{(i)} = 1), \\ 1, & \mathcal{L}(\mathbf{Y}^{(i)}|e^{(i)} = 0) < \mathcal{L}(\mathbf{Y}^{(i)}|e^{(i)} = 1), \end{cases} \tag{30}$$

[3] Note that our proposed deep learning-based detector does not require this information in advance.

where $\hat{e}^{(i)}$ is the decision result of $e^{(i)}$. By substituting the likelihood functions in (29), the ML detector for the backscattered symbol $e^{(i)}$ is expressed as follows:

$$\hat{e}^{(i)} = \begin{cases} 0, & \prod_{n=1}^{N} \frac{1}{\pi^M |K_0|} e^{-\mathbf{y}_n^{(i)\mathrm{H}} K_0^{-1} \mathbf{y}_n^{(i)}} > \prod_{n=1}^{N} \frac{1}{\pi^M |K_1|} e^{-\mathbf{y}_n^{(i)\mathrm{H}} K_1^{-1} \mathbf{y}_n^{(i)}}, \\ 1, & \prod_{n=1}^{N} \frac{1}{\pi^M |K_0|} e^{-\mathbf{y}_n^{(i)\mathrm{H}} K_0^{-1} \mathbf{y}_n^{(i)}} < \prod_{n=1}^{N} \frac{1}{\pi^M |K_1|} e^{-\mathbf{y}_n^{(i)\mathrm{H}} K_1^{-1} \mathbf{y}_n^{(i)}}. \end{cases} \tag{31}$$

Using the logarithm operations, we have

$$\hat{e}^{(i)} = \begin{cases} 0, & \sum_{n=1}^{N} \mathbf{y}_n^{(i)\mathrm{H}} (K_0^{-1} - K_1^{-1}) \mathbf{y}_n^{(i)} < N \ln \frac{|K_1|}{|K_0|}, \\ 1, & \sum_{n=1}^{N} \mathbf{y}_n^{(i)\mathrm{H}} (K_0^{-1} - K_1^{-1}) \mathbf{y}_n^{(i)} > N \ln \frac{|K_1|}{|K_0|}. \end{cases} \tag{32}$$

Based on $\hat{e}^{(i)}$ we can derive the original bit $b^{(i)}$.

VI. Deep Learning-Based Signal Detector

As mentioned, the performance of conventional signal detection techniques greatly depends on the channel and noise distributions. Specifically, these techniques require a specific sophisticated mathematical model for each type of wireless channel and noise. For example, the optimal ML detector in this paper is formulated for Gaussian-like signals and noise (i.e., CSCG). Nevertheless, if the jammer changes its attack strategy or the system is placed at different wireless environments, these conventional detectors may need to be reformulated based on the environment's parameters to achieve a good performance. To address this problem, in this section, we propose a deep learning (DL) based signal detector to detect backscattered signals at the receiver that can automatically adapt to any channel and noise distributions, and thus broadening the applications of the proposed detector, especially in robust anti-jamming strategies as the reactive jammer may perform different attack strategies. Note that a few deep learning-based detectors have been proposed in the literature [16]-[19]. However, these detectors may not be feasible for backscattered signals which are weak and contain the mixture versions of the direct link signals from the jammer and the transmitter. Our proposed DL-based signal detector can deal with this problem by using the LSTM network to learn the long-term dependencies of the received signals at the multiple antennas and over N RF source symbols for each backscattered bit. To the best of our knowledge, our proposed detector is the first DL-based detector for ambient backscatter communication systems.

A. Data Preprocessing

Data processing, as known as data mining, plays a vital role in deep learning. In particular, the quality of the training data and the useful information that can be obtained from it greatly affect the learning performance of the deep learning model. In practice, the raw data will not be "clean" enough and will not be in a proper format that the deep learning model can easily learn to achieve good results. As such, the data must be preprocessed before feeding to the deep learning model. In the following, we describe in details how the data is processed in this work.

Recall that the received signal during the i-th backscatter symbol period can be obtained as follows:

$$\mathbf{y}_n^{(i)} = \mathbf{f}_r \sqrt{\alpha_{dt}} s_{tn}^{(i)} + \mathbf{f}_j \sqrt{\alpha_{dj}} s_{jn}^{(i)} + \mathbf{f}_b e^{(i)} \left(l_r \sqrt{\alpha_{bt}} s_{tn}^{(i)} + l_j \sqrt{\alpha_{bj}} s_{jn}^{(i)} \right) + \boldsymbol{\sigma}_n^{(i)}, \tag{33}$$

where $n = 1, 2, \dots, N$ and $i = 1, 2, \dots, I$. Denote $\mathbf{Y}^{(i)} = [\mathbf{y}_1^{(i)}, \dots, \mathbf{y}_n^{(i)}, \dots, \mathbf{y}_N^{(i)}]^\mathsf{T}$ as the received signal sequence in the i-th symbol period. It is worth noting that the received signal sequence $\mathbf{Y}^{(i)}$ can be shortened to the sample covariance matrix to reduce the size of the training data while maintaining the effectiveness of the deep learning model. Specifically, the sample covariance matrix can be expressed as follows:

$$\mathbf{C}^{(i)} = \frac{1}{N} \sum_{n=1}^{N} \mathbf{y}_n^{(i)} \mathbf{y}_n^{(i)^\mathrm{H}}, \tag{34}$$

where $\mathbf{y}_n^{(i)^\mathrm{H}}$ is the conjugate transpose of received signal matrix $\mathbf{y}_n^{(i)}$. In this way, the sample covariance matrix $\mathbf{C}^{(i)}$ represents all the useful information of the received signal sequence $\mathbf{Y}^{(i)}$.

Moreover, the estimation accuracy of the detector can be further improved by exploiting the channel statistical covariance matrices $\mathbf{K}_0$ and $\mathbf{K}_1$ to obtain the better estimations of the sample covariance matrix as follows:

$$\begin{aligned} \widetilde{\mathbf{C}}_0^{(i)} &= \mathbf{C}^{(i)} \mathbf{K}_0^{-1}, \\ \widetilde{\mathbf{C}}_1^{(i)} &= \mathbf{C}^{(i)} \mathbf{K}_1^{-1}. \end{aligned} \tag{35}$$

It is worth noting that the channel statistical covariance matrices can be estimated at the receiver through several common estimation techniques in the literature. However, the receiver does not know whether the tag backscatters bit "0" or bit "1". Thus, we need to feed both $\widetilde{\mathbf{C}}_0^{(i)}$ and $\widetilde{\mathbf{C}}_1^{(i)}$ to the deep learning model, and then the deep learning model can learn and estimate the backscattered bit. In Fig. 3, we show how to reshape $\widetilde{\mathbf{C}}_0^{(i)}$ and $\widetilde{\mathbf{C}}_1^{(i)}$ and feed them to the deep learning model. For simplicity, we remove the indicator (i). $c_{i,j}^0$ with $i, j \in \{1, 2, \dots, M\}$ represents an element of $\widetilde{\mathbf{C}}_0^{(i)}$ and $c_{i,j}^1$ with $i, j \in \{1, 2, \dots, M\}$ represents an element of $\widetilde{\mathbf{C}}_1^{(i)}$. Note that with M antennas at the receive, $\widetilde{\mathbf{C}}_0^{(i)}$ and $\widetilde{\mathbf{C}}_1^{(i)}$ are $M \times M$ matrices. We then reshape $\widetilde{\mathbf{C}}_0^{(i)}$ and $\widetilde{\mathbf{C}}_1^{(i)}$ into two arrays, and each array contains $M \times M$ elements. Thus, we have total $2 \times M \times M$ elements. We then construct the input sequence of the deep learning model by taking the real part $\bar{c}_{i,j}^k$, the imaginary part $\hat{c}_{i,j}^k$, and the absolute value $|c_{i,j}^k|$ of each element $c_{i,j}^k$

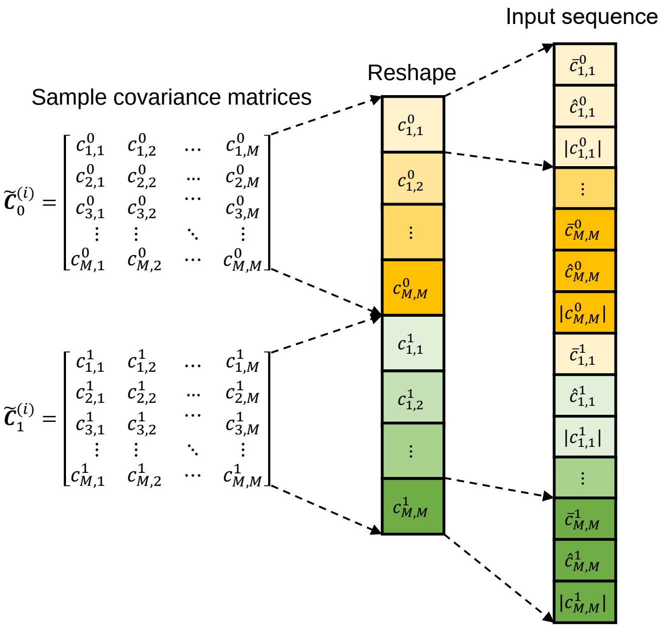

Sample covariance matrices
$\widetilde{\mathbf{C}}_0^{(i)} = \begin{bmatrix} c_{1,1}^0 & c_{1,2}^0 & \cdots & c_{1,M}^0 \\ c_{2,1}^0 & c_{2,2}^0 & \cdots & c_{2,M}^0 \\ c_{3,1}^0 & c_{3,2}^0 & \cdots & c_{3,M}^0 \\ \vdots & \vdots & \ddots & \vdots \\ c_{M,1}^0 & c_{M,2}^0 & \cdots & c_{M,M}^0 \end{bmatrix}$
$\widetilde{\mathbf{C}}_1^{(i)} = \begin{bmatrix} c_{1,1}^1 & c_{1,2}^1 & \cdots & c_{1,M}^1 \\ c_{2,1}^1 & c_{2,2}^1 & \cdots & c_{2,M}^1 \\ c_{3,1}^1 & c_{3,2}^1 & \cdots & c_{3,M}^1 \\ \vdots & \vdots & \ddots & \vdots \\ c_{M,1}^1 & c_{M,2}^1 & \cdots & c_{M,M}^1 \end{bmatrix}$
Reshape
$c_{1,1}^0$
$c_{1,2}^0$
$\vdots$
$c_{M,M}^0$
$c_{1,1}^1$
$c_{1,2}^1$
$\vdots$
$c_{M,M}^1$
Input sequence
$\bar{c}_{1,1}^0$
$\hat{c}_{1,1}^0$
$|c_{1,1}^0|$
$\vdots$
$\bar{c}_{M,M}^0$
$\hat{c}_{M,M}^0$
$|c_{M,M}^0|$
$\bar{c}_{1,1}^1$
$\hat{c}_{1,1}^1$
$|c_{1,1}^1|$
$\vdots$
$\bar{c}_{M,M}^1$
$\hat{c}_{M,M}^1$
$|c_{M,M}^1|$

Fig. 3: Data preprocessing.

with $k \in \{0, 1\}$ and $i, j \in \{1, 2, \ldots, M\}$. Then, the input sequence is fed into the deep neural network for learning. In this way, the deep learning model can learn all aspects of the received signals and thus improve the learning efficiency.

B. Deep Neural Network Architecture

In this work, we adopt Long Short-Term Memory (LSTM) [28], [29] network that can quickly learn the received signals to detect backscattered bits. An LSTM network is a type of recurrent neural network that can capture and learn the long-term dependencies between input data [28]. As such, this network architecture is widely adopted in sequence prediction problems. The reasons for using the LSTM network in this work are twofold.

- First, the input sequence consists of the estimated channel statistical covariances at M antennas (with the same backscattered bit) of the receiver. Thus, with the LSTM network, the dependencies between these covariances are kept track by the memory part of the LSTM unit during the training process. As a result, the LSTM network can detect the backscattered bits more effectively by learning from all received signals at M antennas and the relations between them.
- Second, as mentioned, the channel remains invariant during one backscatter frame which consists of I bits being backscattered. Each backscattered bit corresponds to an input

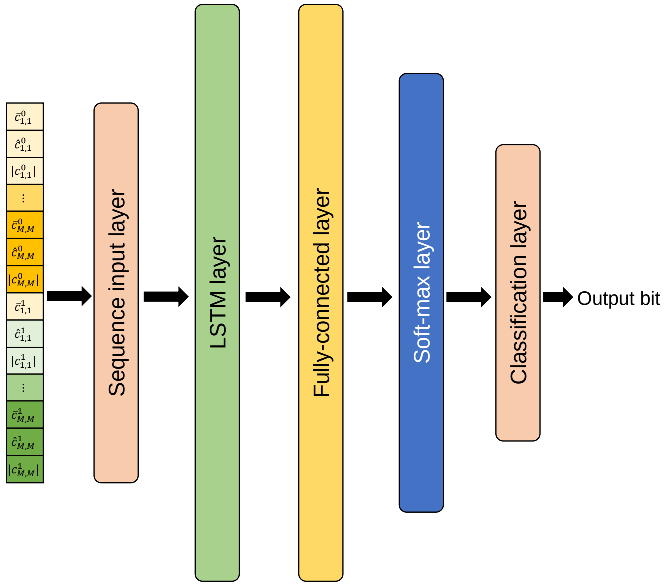


Fig. 4: Deep neural network architecture.

sequence for the deep learning model to predict. Thus, by using the LSTM network, we can capture and learn the dependencies between these input sequences, and thus keeping track of the channel state during one backscatter frame. As such, the deep learning model can learn the input data more efficiently resulting in a high predict accuracy.

In Fig. 4, we illustrate the deep neural network architecture in this work. In particular, the deep neural network consists of one sequence input layer, one LSTM layer, one fully connected layer, one softmax layer, and one classification layer. In particular, the sequence input layer is used to import the input data to the deep neural network. After that, the input data is forwarded to the LSTM layer. Specifically, the LSTM layer has an output (also known as the hidden state) denoted by $\mathbf{w}$ and a cell state denoted by s as illustrated in Fig. 5 [29]. At each time step t, the LSTM layer uses the current output $\mathbf{w}_{t-1}$ and cell state s_{t-1} as well as the input sequence data to calculate output $\mathbf{w}_t$ and cell state s_t.

Cell state s_t represents the information learned from the previous time steps and output state $\mathbf{w}_t$ contains the output for the current time step. In the training process, the LSTM layer updates its weights by using gates. There are three types of weights in the LSTM layer including the input weights $\mathbf{W}$, the recurrent weights $\mathbf{R}$, and the bias b. These matrices are concatenated from the weights of each gates in the LSTM layers as follows:

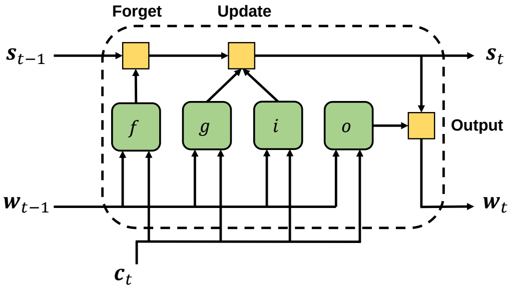


Fig. 5: Flow of data at the LSTM layer [29].

$$\mathbf{W} = \begin{bmatrix} W_i \\ W_f \\ W_g \\ W_o \end{bmatrix}, \mathbf{R} = \begin{bmatrix} R_i \\ R_f \\ R_g \\ R_o \end{bmatrix}, \mathbf{b} = \begin{bmatrix} b_i \\ b_f \\ b_g \\ b_o \end{bmatrix}, \tag{36}$$

where i, f, g, o represent the input gate, the forget gate, the cell candidate, and the output gate, respectively. In particular, the input state is used to decide what information we will update. At time step t, the input gate can be formulated as follows:

$$i_t = \sigma_g(W_i\mathbf{c}_t + R_i\mathbf{w}_{t-1} + b_i), \tag{37}$$

where σ_g is the gate activation function. In this work, we use the sigmoid function which is a common activation function in deep learning [31]. σ_g is given by $\sigma_g(x) = (1 + e^{-x})^{-1}$. $\mathbf{c}_t$ is the input sequence at time step t. The forget state is used to decide what information we will remove from the cell state, i.e., remove the long-term dependencies. The forget state can be formulated as follows:

$$f_t = \sigma_g(W_f\mathbf{c}_t + R_f\mathbf{w}_{t-1} + b_f). \tag{38}$$

The cell candidate is used to add information to the cell state and is formulated as follows:

$$g_t = \sigma_c(W_g\mathbf{c}_t + R_g\mathbf{w}_{t-1} + b_g), \tag{39}$$

where σ_c is the state activation function. In this work, we use the hyperbolic tangent function (tanh) [31], which is commonly used to express the cell state of LSTM networks [29], to derive the state activation function. Finally, the output gate is used control the output of the LSTM layer. This gate can be formulated as follows:

$$o_t = \sigma_g(W_o\mathbf{c}_t + R_o\mathbf{w}_{t-1} + b_o). \tag{40}$$

Based on these gates, the LSTM can update the cell state and the hidden state (i.e., output). In particular, the cell state can be derived as follows:

$$\mathbf{s}_t = f_t \odot \mathbf{s}_{t-1} + i_t \odot g_t, \tag{41}$$

where $\odot$ denotes the element-wise multiplication of vectors. Similarly, the hidden state can be expressed as follows:

$$\mathbf{w}_t = o_t \odot \sigma_c(\mathbf{s}_t). \tag{42}$$

By recurrently updating the cell state and the hidden state, the LSTM layer can learn useful information from the input data. We then use the fully connected hidden layer and the softmax layer between the LSTM and the output layer to transform the output of the LSTM layer to the form that the output layer can use to predict backscattered bits, i.e., bits “0” or bits “1”. In particular, the fully connected hidden layer multiplies its input by a weight matrix and then adds a bias vector. In this work, the weights are initialized with the Glorot initializer [30], and the bias is initialize with zeros. Then, the softmax layer applies the softmax function [31] to the output of the fully connected layer. Finally, the classification layer calculates the cross entropy loss for the output of the softmax layer. In this way, the receiver can predict the backscattered bits sent from the transmitter.

VII. Analytical and Simulation Results

A. Simulation Setup

Unless otherwise stated, the average signal-to-noise ratios of the direct link from the jammer and from the transmitter to the receiver are set at 7 dB and 5 dB, respectively. The relative signal-to-noise ratios of the backscatter links (i.e., jammer-tag-receiver and transmitter-tag-receiver) are set at $\tilde{\alpha}_j = \tilde{\alpha}_r = -15$ dB. These parameters will be varied to evaluate the performance of the system in different scenarios. Moreover, to observe the performance of the proposed approach in different settings, the number of antennas at the receiver is varied from 1 to 10. For the DL-based signal detector, the training set is generated by running 10^5 Monte Carlo runs. After finishing the training process, the trained model is used to obtain the BER performance through 10^6 Monte Carlo runs. For the deep neural network, we employ one sequence input layer, one LSTM layer, one fully connected layer, one softmax layer, and one classification layer as shown in Fig. 4. The number of hidden units of the LSTM layer is 1000. The fully connected layer has the output size of 2 corresponding to two classes of output bits, i.e., bits “0” and bits “1”. We use Adam optimizer to train the network with the learning rate of 0.001. It is worth noting that

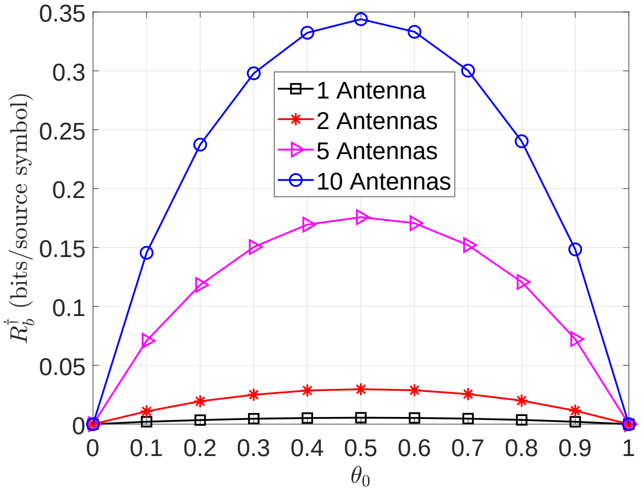


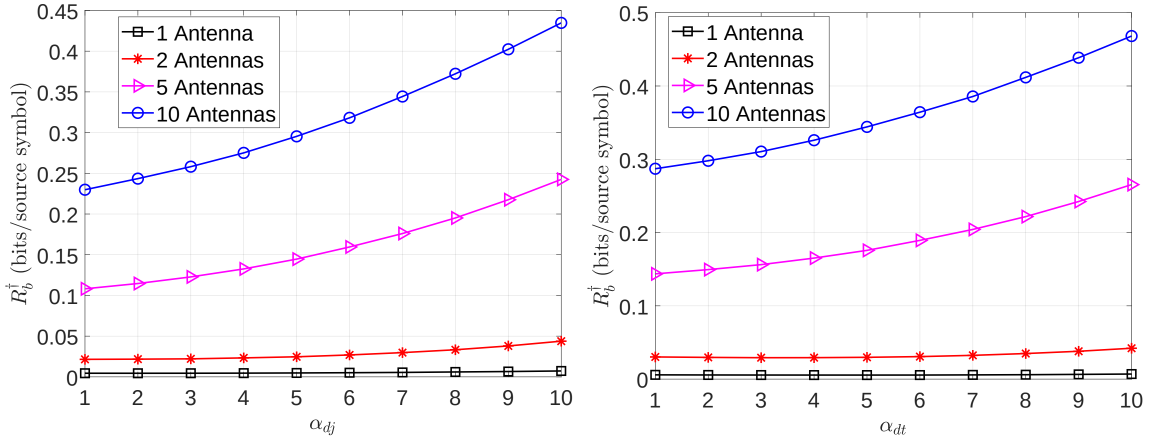
Fig. 6: Maximum achievable backscatter rate vs. prior probability when backscattering bits "0".

conventional solutions are not feasible in dealing with the super-reactive jammer as discussed in Section I. Thus, in this section, we focus on analyzing and demonstrating the effectiveness of our proposed solutions in different scenarios.

B. Maximum Achievable Backscatter Rate

We first vary the prior probability of backscattering bits "0" and observe the maximum achievable backscatter rate of the tag with different number of antennas at the receiver as illustrated in Fig. 6. In particular, the maximum achievable backscatter rate is obtained based on Theorem 1 through 10^6 Monte Carlo runs. It can be observed that the backscatter rate increases with the number of antennas at the receiver. The reason is that with multiple antennas, the receiver can leverage the antenna gain to eliminate the effects of the fading and the direct link interference. As a result, the backscattered signals receiver at the receiver can be enhanced under reactive jamming attacks. It is worth noting that when the probability of backscattering bits "0" equals 0.5, the backscatter rate is maximized. Thus, in this paper we employ the modulo-2 operator to encode the original bits as discussed in Section IV-C to backscatter equiprobable symbols, i.e., $\theta_0 = \theta_1 = 0.5$.

Next, we vary the average SNRs of the direct link from the jammer and from the transmitter to the receiver and evaluate the maximum achievable backscatter rate in various settings as shown in Fig. 7(a) and Fig. 7(b), respectively. Specifically, in Fig. 7(a), the SNR of the direct link from the jammer is varied from 1 dB to 10 dB while fixing the SNR of the direct from the transmitter at 5 dB. It can be observed that the backscatter rate increases with the SNR of the



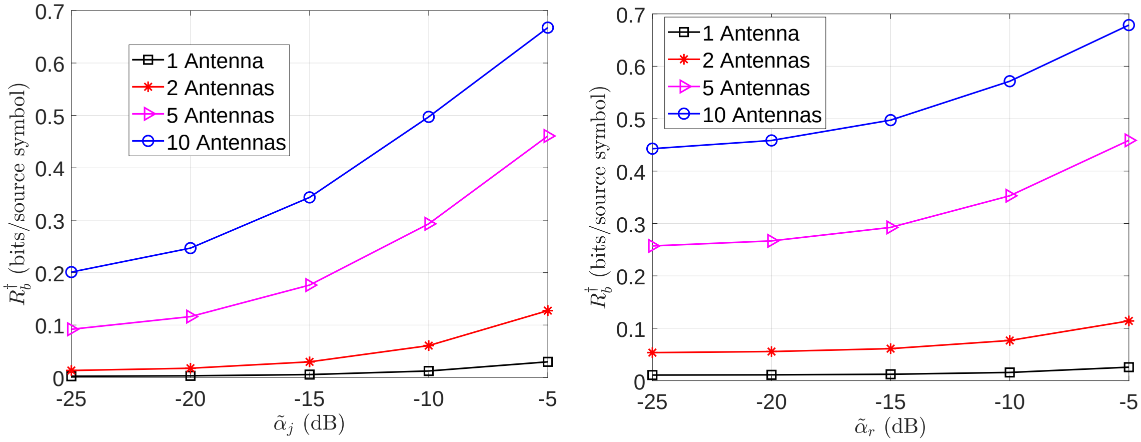
(a) $\alpha_{dt} = 5$ dB, $\tilde{\alpha}_j = \tilde{\alpha}_r = -15$ dB.

(b) $\alpha_{dj} = 7$ dB, $\tilde{\alpha}_j = \tilde{\alpha}_r = -15$ dB.

Fig. 7: Maximum average achievable backscatter rate vs. the average SNRs of the direct link from (a) the jammer and (b) the transmitter.

direct link from the jammer. In other words, with our proposed approach, the system can improve the average throughput when the jammer attacks the channel at higher power levels. This is an important finding as our proposed solution can effectively deal with jamming attacks. As we analyzed in Sections II and III, conventional methods cannot deal with the considered jamming attacks as the super-reactive jammer can simultaneously attack and listen to the activities of the transmitter. It is worth noting that the higher number of antennas at the receiver, the higher backscatter rate we can achieve. This is due to the fact that the receiver can effectively deal with the direct link interference from the transmitter and the jammer by using multiple antennas. Similarly, in Fig. 7(b), we fix the SNR of the direct link from the jammer at 7 dB and vary the SNR of the direct link from the transmitter from 1 dB to 10 dB. It can be observed that, when the transmitter transmits at higher power levels, the tag can backscatter more information to the receiver. Thus, to achieve better performance under jamming attacks, one feasible solution is increasing the transmit power of the transmitter. Note that with two antennas implemented at the receiver, the backscatter rate is very low due to the fading and direct link interference. Again, these negative effects can be mitigated by using more antennas at the receiver.

Next, we vary the relative SNRs (corresponding to the reflection coefficient, antenna gain, path loss, and tag-to-receiver distance as expressed in (15)) of the backscatter links from the jammer and the transmitter and observe the maximum achievable backscatter rate of the tag as shown in Fig. 8(a) and Fig. 8(b), respectively. It can be observed that, when the relative



(a) $\alpha_{dt} = 5$ dB, $\alpha_{dj} = 7$ dB, $\tilde{\alpha}_r = -15$ dB.

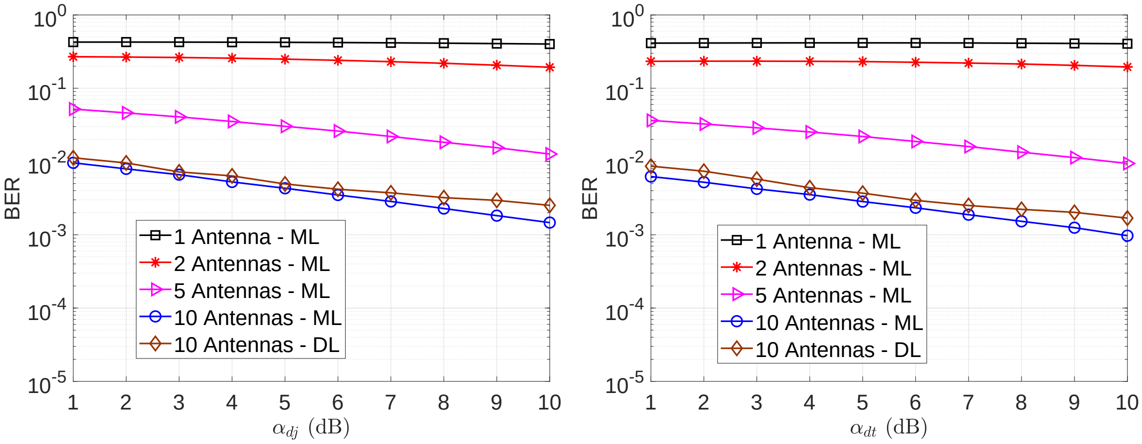
(b) $\alpha_{dt} = 5$ dB, $\alpha_{dj} = 7$ dB, $\tilde{\alpha}_j = -10$ dB.

Fig. 8: Maximum average achievable backscatter rate vs. the SNRs of the backscatter links from (a) the jammer and (b) the transmitter.

SNRs increases, the maximum achievable backscatter rate of the tag also increases. The reason is that with higher SNRs, the backscattered signals received at the receiver are improved, and thus the receiver can efficiently extract the original information. Again, the receiver can deal with the fading and direct link interference from the transmitter and the jammer to increase the backscatter rate by employing more antennas. Note that in above results, we theoretically obtain the maximum achievable backscatter rate of the backscatter tag which does not require to detect backscattered bits at the receiver. Thus, deep learning is not applied. In the following, we will demonstrate the effectiveness of the DL-based signal detector in terms of BER.

C. *BER Performance*

In this subsection, we evaluate the BER of the system in different scenarios. We first vary the average SNR of the direct link from the jammer and from the transmitter to the receiver as shown in Fig. 9(a) and Fig. 9(b), respectively. In particular, it can be observed from Fig. 9(a) that the BER reduces when the SNR of the direct link from the jammer to the receiver α_{dj} increases. The reason is that with stronger jamming signals, the backscatter tag can backscatter information more effectively, resulting in lower BERs. The similar trend can be observed when increasing the SNR of the direct link from the transmitter to the receiver α_{dt} as shown in Fig. 9(b). As a result, with our proposed deception mechanism, the higher power the jammer uses to attack the channel, the better BER performance we can achieve. In addition, increasing the transmit power of the transmitter can also improve the BER performance. More importantly, our proposed

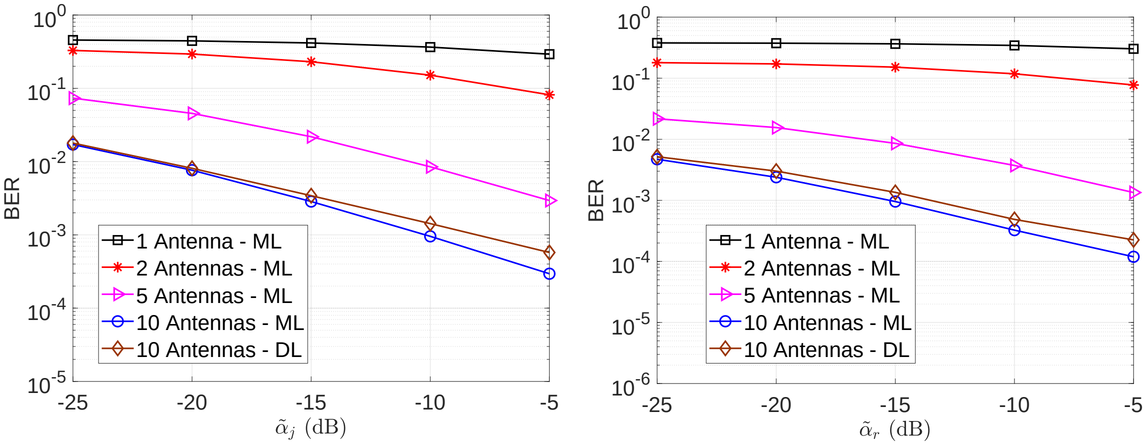


(a) $\alpha_{dt} = 5$ dB, $\tilde{\alpha}_j = \tilde{\alpha}_r = -15$ dB.

(b) $\alpha_{dj} = 7$ dB, $\tilde{\alpha}_j = \tilde{\alpha}_r = -15$ dB.

Fig. 9: BER vs. the average SNRs of the direct link from (a) the jammer and (b) the transmitter.

DL-based detector can achieve the BER performance close to that of the ML detector without using any sophisticated mathematical model for the wireless communication system. Again, the BER reduces when we employ more antennas at the receiver. Note that when the SNR increases, the gap between the BERs obtained by our proposed DL-based detector and the optimal ML detector increases. The reason is that with higher SNRs, the optimal ML detector can detect the backscattered bits based on the likelihood functions of the received signals more efficiently. However, our proposed detector can always achieve the BER close to that of the optimal ML detector without requiring complicated mathematical models and distributions.

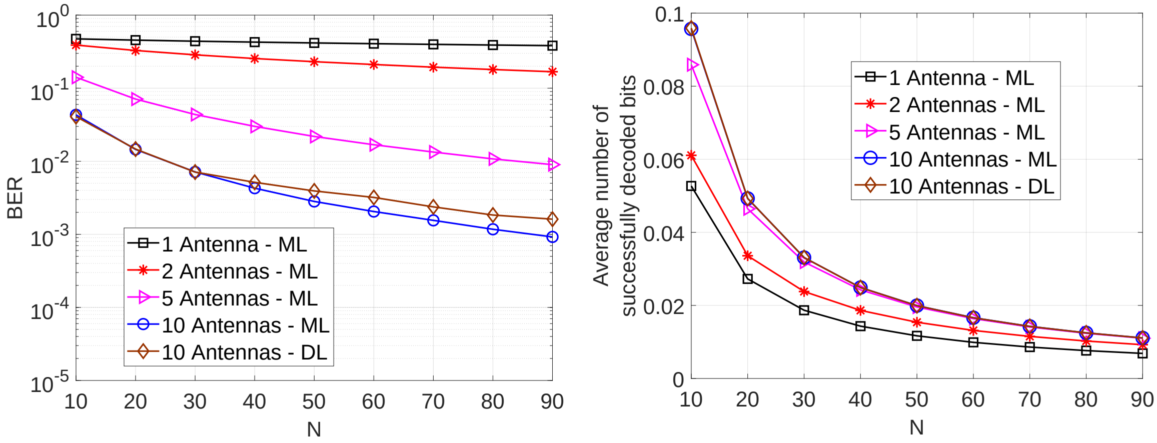


(a) $\alpha_{dt} = 5$ dB, $\alpha_{dj} = 7$ dB, $\tilde{\alpha}_r = -15$ dB.

(b) $\alpha_{dt} = 5$ dB, $\alpha_{dj} = 7$ dB, $\tilde{\alpha}_j = -10$ dB.

Fig. 10: BER vs. the SNRs of the backscatter links from (a) the jammer and (b) the transmitter.

Next, the SNRs of the backscatter links (i.e., jammer-tag-receiver and transmitter-tag-receiver) are varied to evaluate the BER performance as shown in Fig. 10(a) and Fig. 10(b), respectively. In particular, when $\tilde{\alpha}_j$ and $\tilde{\alpha}_r$ increase, the BER is reduced. This is stemmed from the fact that when the SNRs of the backscatter links increase, the received backscattered signals are stronger, and thus the backscatter communications are more reliable. Again, the more antennas we have at the receiver, the better BER performance we can achieve. It is worth noting that our proposed DL-based detector achieves the BER performance close to that of the ML detector.



(a)

(b)

Fig. 11: (a) BER and (b) Average number of successfully decoded bits vs. N.

Finally, we vary the spreading factor N to evaluate the BER performance of the system. Recall that the backscatter rate must be lower than the data rate of the RF sources to guarantee the successful decoding process at the receiver. As such, each symbol is backscattered over N RF source symbols. It can be observed from Fig. 11(a) that the BER reduces when N is increased. However, as shown in Fig. 11(b), a higher value of N leads to a lower throughput of the system as the backscatter rate is reduced. Hence, the trade-off between N and the backscatter rate should be considered when designing the system.

VIII. Conclusion

In this paper, we have proposed the state-of-the-art framework to deal with the super-reactive jammer that has virtually unlimited power and can simultaneously attack the channel and detect activity of the transmitter on the channel. To the best of our knowledge, there is no effective solution that can defeat such kinds of jamming attacks. To address this problem, in this paper, we have first introduced the novel deception mechanism to allows the transmitter to attract the jammer and leverage the jamming signals to transmit information to the receiver by using the ambient backscatter tag. To improve the BER performance, we have proposed to use multiple antennas at the receiver. Then, the maximum likelihood detector is developed to detect the backscattered bits at the receiver. However, this detector requires sophisticated mathematical operations to model the wireless channel and is not feasible to directly apply when the system changes. Thus, we have proposed the novel deep learning-based detector with the latest advances in Long Short Team Memory networks. To the best of our knowledge, this is the first deep learning-based signal detector for ambient backscatter communications in which the backscattered signals are usually weak. Through the simulation results and theoretical analysis, we have shown that with our proposed solution, the more power the jammer uses to attack the channel, the better performance in terms of BER and throughput we can achieve. In addition, the proposed deep learning-based signal detector can achieve the BER performance close to that of the optimal maximum likelihood detector.

APPENDIX A
THE PROOF OF THEOREM 1

Similar to [20], in the following, we will show how to obtain the maximum achievable backscatter rate of the backscatter tag. It can be observed that, $R_{\mathrm{b}} = I(e;\mathbf{y})$ is the mutual information between the received signals $\mathbf{y}$ at the receiver and the modulated information e. Thus, the maximum achievable backscatter rate $R_{\mathrm{b}}^{\dagger}$ can be calculated as in (43) [20].

$$R_{\mathrm{b}}^{\dagger} = \mathbb{E}[I(e,\mathbf{y})]. \tag{43}$$

The mutual information $I(e,\mathbf{y})$ is formulated as follows:

$$I(e,\mathbf{y}) = C(\theta_0) - \mathbb{E}_{\mathbf{y}_0}[H(e|\mathbf{y}_0)], \tag{44}$$

where $H(e|\mathbf{y}_0)$ is the conditional entropy of e given $\mathbf{y}_0$, and $C(\theta_0)$ denotes the binary entropy function which is expressed in (45).

$$C(\theta_0) \triangleq -\theta_0 \log_2 \theta_0 - \theta_1 \log_2 \theta_1. \tag{45}$$

It is worth noting that $C(\theta_0)$ is independent at all the channel coefficients. Thus, $R_{\mathrm{b}}^{\dagger}$ can be rewritten as follows:

$$R_{\mathrm{b}}^{\dagger} = \mathbb{E}[I(e,\mathbf{y})] = C(\theta_0) - \mathbb{E}_{\mathbf{y}_0}[H(e|\mathbf{y}_0)]. \tag{46}$$

The posterior probability of e when receiving $\mathbf{y}_0$ is formulated as follows:

$$p(e=j|\mathbf{y}_0)=\frac{\theta_j p(\mathbf{y}|e=j)}{\theta_0 p(\mathbf{y}_0|e=0)+\theta_1 p(\mathbf{y}_0|e=1)}, \tag{47}$$

with $j\in\{0,1\}$. We then define $\omega_j=p(e=j|\mathbf{y}_0)$ with $j\in\{0,1\}$. Given the above, we can derive $H(e|\mathbf{y}_0)$ as follows:

$$H(e|\mathbf{y}_0)=-\sum_{j=0}^{1}\omega_j\log_2\omega_j=C(\omega_0). \tag{48}$$

As a result, the maximum achievable backscatter rate is obtained as follows:

$$R_{\mathsf{b}}^{\dagger}=C(\theta_0)-\mathbb{E}_{\mathbf{y}_0}[C(\omega)]=C(\theta_0)-\int_{\mathbf{y}_0}(\theta_0 p(\mathbf{y}_0|e=0)+\theta_1 p(\mathbf{y}_0|e=1))C(\omega_0)d\mathbf{y}_0. \tag{49}$$

REFERENCES

[1] N. V. Huynh, D. T. Hoang, D. N. Nguyen, E. Dutkiewicz, and M. Mueck, "Defeating Smart and Reactive Jammers with Unlimited Power", *IEEE WCNC*, Virtual Conference, 25-28 May 2020.

[2] W. Afifi and M. Krunz, "Exploiting Self-Interference Suppression for Improved Spectrum Awareness/efficiency in Cognitive Radio Systems," *IEEE INFOCOM*, Turin, Italy, 14-19 Apr. 2013.

[3] M. K. Hanawal, D. N. Nguyen, and M. Krunz, "Cognitive Networks With In-Band Full-Duplex Radios: Jamming Attacks and Countermeasures," *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, no. 1, pp. 296-309, Mar. 2020.

[4] L. Xiao, H. Dai, and P. Ning, "Jamming-Resistant Collaborative Broadcast Using Uncoordinated Frequency Hopping," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 1, pp. 297-309, Feb. 2012.

[5] K. Pelechrinis, C. Koufogiannakis, and S. V. Krishnamurthy, "On the Efficacy of Frequency Hopping in Coping With Jamming Attacks in 802.11 Networks," *IEEE Transactions on Wireless Communications*, vol. 9, no. 10, pp. 3258-3271, Oct. 2010.

[6] Y. Wu, B. Wang, K. R. Liu, and T. C. Clancy, "Anti-Jamming Games in Multi-Channel Cognitive Radio Networks," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 1, pp. 4-15, Jan. 2012.

[7] C. Popper, M. Strasser, and S. Capkun, "Anti-Jamming Broadcast Communication Using Uncoordinated Spread Spectrum Techniques," *IEEE Journal on Selected Areas in Communications*, vol. 28, no. 5, pp. 703-715, Jun. 2010.

[8] Y. Gao, Y. Xiao, M. Wu, M. Xiao, and J. Shao, "Game Theory-Based Anti-Jamming Strategies for Frequency Hopping Wireless Communications," *IEEE Transactions on Wireless Communications*, vol. 17, no. 8, pp. 5314-5326, Aug. 2018.

[9] G. Noubir, R. Rajaraman, B. Sheng, and B. Thapa, "On the Robustness of IEEE 802.11 Rate Adaptation Algorithms Against Smart Jamming," *ACM Conference on Wireless Network Security*, Hamburg, Germany, Jun. 2011.

[10] D. T. Hoang, D. Niyato, P. Wang, and D. I. Kim, "Performance Analysis of Wireless Energy Harvesting Cognitive Radio Networks Under Smart Jamming Attacks," *IEEE Transactions on Cognitive Communications and Networking*, vol. 1, no. 2, pp. 200-216, Jun. 2015.

[11] D. T. Hoang, D. N. Nguyen, M. A. Alsheikh, S. Gong, E. Dutkiewicz, D. Niyato, and Z. Han, ""Borrowing Arrows with Thatched Boats": The Art of Defeating Reactive Jammers in IoT Networks," *IEEE Wireless Communications*, vol. 27, no. 3, pp. 79-87, Jun. 2020.

[12] N. V. Huynh, D. T. Hoang, X. Lu, D. Niyato, P. Wang, and D. I. Kim, "Ambient Backscatter Communications: A Contemporary Survey," *IEEE Communications Surveys & Tutorials*, vol. 20 , no. 4, pp. 2889-2922, Fourth Quarter 2018.

[13] N. V. Huynh, D. N. Nguyen, D. T. Hoang and E. Dutkiewicz, “Jam Me If You Can: Defeating Jammer with Deep Dueling Neural Network Architecture and Ambient Backscattering Augmented Communications,” *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 11, pp. 2603-2620, Nov. 2019.

[14] A. Mpitziopoulos, D. Gavalas, C. Konstantopoulos, and G. Pantziou, “A Survey on Jamming Attacks and Countermeasures in WSNs,” *IEEE Communications Surveys & Tutorials*, vol. 11, no. 4, pp. 42-56, Fourth Quarter 2009.

[15] V. Liu, A. Parks, V. Talla, S. Gollakota, D. Wetherall, and J. R. Smith, “Ambient Backscatter: Wireless Communication Out of Thin Air,” *ACM SIGCOMM*, Hong Kong, China, Aug. 2013.

[16] H. Ye, G. Y. Li, and B.-H. Juang, “Power of Deep Learning for Channel Estimation and Signal Detection in OFDM Systems,” *IEEE Wireless Communications Letters*, vol. 7, no. 1, pp. 114-117, Feb. 2018.

[17] T. J. O’Shea, T. Roy, and T. C. Clancy, “Over-the-Air Deep Learning Based Radio Signal Classification,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 1, pp. 168-179, Feb. 2018.

[18] S. Rajendran, W. Meert, D. Giustiniano, V. Lenders, and S. Pollin, “Deep Learning Models for Wireless Signal Classification With Distributed Low-Cost Spectrum Sensors,” *IEEE Transactions on Cognitive Communications and Networking*, vol. 4, no. 3, pp. 433-445, Sept. 2018.

[19] C. Fan, X. Yuan, and Y.-J. Zhang, “CNN-Based Signal Detection for Banded Linear Systems,” *IEEE Transactions on Wireless Communications*, vol. 18, no. 9, pp. 4394-4407, Sept. 2019.

[20] H. Guo, Q. Zhang, D. Li, and Y.-C. Liang, “Noncoherent Multiantenna Receivers for Cognitive Backscatter System with Multiple RF Sources,” [Online]. Available: arXiv:1808.04316.

[21] T. S. Rappaport, *Wireless Communications Principles and Practice*, 2nd ed. Prentice Hall, 2002.

[22] J. Guo, N. Zhao, F. R. Yu, X. Liu, and V. CM. Leung, “Exploiting Adversarial Jamming Signals for Energy Harvesting in Interference Networks,” *IEEE Transactions on Wireless Communications*, vol. 16, no. 2, pp. 1267-1280, Feb. 2017.

[23] J. Qian, F. Gao, G. Wang, S. Jin, and H. Zhu, “Noncoherent Detections for Ambient Backscatter System,” *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1412-1422, Mar. 2017.

[24] H. Xing, K.-K. Wong, Z. Chu, and A. Nallanathan, “To Harvest and Jam: A Paradigm of Self-Sustaining Friendly Jammers for Secure AF Relaying,” *IEEE Transactions on Signal Processing*, vol. 63, no. 24, pp. 6616-6631, Sept. 2015.

[25] S. Goel and R. Negi, “Guaranteeing Secrecy Using Artificial Noise,” *IEEE Transactions on Wireless Communications*, vol. 7, no. 6, pp. 2180-2189, Jun. 2008.

[26] H. Pirzadeh, S. M. Razavizadeh, and E. Bjornson, “Subverting Massive MIMO by Smart Jamming,” *IEEE Wireless Communications Letters*, vol. 5, no. 1, pp. 20-23, Feb. 2016.

[27] Q. Zhang, H. Guo, Y.-C. Liang, and X. Yuan, “Constellation Learning-Based Signal Detection for Ambient Backscatter Communication Systems,” *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 2, pp. 452-463, Feb. 2019.

[28] S. Hochreiter, and J. Schmidhuber, “Long Short-Term Memory,” *Neural computation*, vol. 9, no. 8, pp. 1735-1780, 1997.

[29] Long Short-Term Memory Networks, [Online]. Available: https://www.mathworks.com/help/deeplearning/ug/long-short-term-memory-networks.html

[30] X. Glorot and Y. Bengio, “Understanding the Difficulty of Training Deep Feedforward Neural Networks,” in *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, Sardinia, Italy, 2010.

[31] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, 2016.

[32] H. Cai, Q. Zhang, Q. Li, and J. Qin, “Proactive Monitoring via Jamming for Rate Maximization Over MIMO Rayleigh Fading Channels,” *IEEE Communications Letters*, vol. 21, no. 9, pp. 2021-2024, Jun. 2017.

[33] C. Yang, J. Gummeson, and A. Sample, “Riding the Airways: Ultra-Wideband Ambient Backscatter via Commercial Broadcast Systems,” *IEEE INFOCOM*, Atlanta, GA, USA, 1-4 May 2017.

[34] M. K. Hanawal, D. N. Nguyen, and M. Krunz, “Jamming Attack on In-Band Full-Duplex Communications: Detection and Countermeasures,” *IEEE INFOCOM*, San Francisco, CA, USA, 10-14 Apr. 2016.