Performance Analysis of Multiple-Antenna Ambient Backscatter Systems at Finite Blocklengths

Likun Sui\textsuperscript{,}, Zhuai Lin\textsuperscript{,}, Senior Member, IEEE, Pei Xiao\textsuperscript{,}, Senior Member, IEEE, H. Vincent Poor\textsuperscript{,}, Life Fellow, IEEE, and Branka Vucetic\textsuperscript{,}, Life Fellow, IEEE

Abstract—This article analyzes the maximal achievable rate for a given blocklength and maximal error probability over a multiple-antenna ambient backscatter channel. The result consists of a finite blocklength channel coding achievable rate and a converse bound for the legacy system with finite alphabet constraints and multiple-input-multiple-output based on the Neyman–Pearson test, the Berry–Esseen theorem, and the Mellin transform. Then, we derive the closed-form expression of the mutual information and the information variance to reduce the complexity of the computation. By applying the low-complexity maximum-likelihood detection, the relation between the maximal error probability of the RF source signal and the average error probability of the tag symbol with respect to the blocklength is proposed. Finally, numerical evaluation of these bounds shows fast convergence to the maximal achievable rate as the blocklength increases and also proves that the information variance is an accurate measure of the backoff from the maximal achievable rate due to finite blocklength.

Index Terms—Achievability bound, ambient backscatter communications (AmBCs), converse bound, finite blocklength regime, multiple-input multiple-output (MIMO) channel.

I. INTRODUCTION

T he Internet of Things (IoT) has drawn considerable attention from academia and industry. However, as the demand for IoT increases dramatically, the provisioning of power to massive numbers of devices becomes a significant challenge. The novel concept of using passive communication techniques to enable communications for low-power devices is known as ambient backscatter [1], [2], [3], [4], [5]. This approach offers a promising communication solution between batteryless devices, enabling the future growth of IoT systems.

Ambient Backscatter Communications (AmBCs) use RF signals to transmit information symbols and harvest energy, resulting in battery-free operations. The basic operating principles of AmBCs are as follows.

1) A tag transmits a symbol, either 0 or +1, by backscattering and modulating an RF signal from an existing ambient source.
2) The receiver receives the signal from the ambient source and the backscattered signal from the tag.

Compared with active radio protocols, such as Wi-Fi, Bluetooth, and ZigBee, an AmBC system has a relatively limited data rate. Multiple antenna techniques can be used to mitigate this problem by increasing the data rate. For example, Yang et al. [6] used multiple antenna-based orthogonal frequency division multiplexing (OFDM) to design an AmBC system and cope with different channel conditions. Many current research works have focused on analyzing the achievable rate and capacity for AmBC systems in an infinite blocklength regime [7], [8]. However, in practice, evaluating how to maintain the desired maximal error probability at a given finite blocklength is critical. According to [9], the impact of channel coding on primary networks is often ignored in most of the open literature for AmBC systems. To address this, the authors analyzed the bit-error-rate (BER) performance of an AmBC system that employs LDPC-coded RF source signals. They derived the BER upper and lower bounds for both the RF source and tag signals, demonstrating a significant decrease in BERs compared to conventional AmBC systems that utilize uncoded RF source signals. Using multiple antennas in AmBC systems is an efficient way to enhance BER performance and operating range [10], [11]. Lee et al. [10] examined the BER performance of an AmBC system where the transmitter of the primary system has multiple antennas for transmitting beamforming. Long et al. [11] proposed an AmBC receiver with multiple antennas for receive beamforming to achieve higher transmission rates and better coverage. Recent studies have suggested that AmBC systems could use nonorthogonal multiple access (NOMA) signals in 5G networks [12], [13], [14], [15]. Zhang et al. [12], [13] investigated the outage probability and ergodic rate for AmBC systems that use NOMA signals. The optimization problem to maximize the ergodic capacity of such a system was studied in [14], while the reliability and security of an AmBC system using NOMA signals were explored in [15].

The fundamental theorem of reliable data transmission limits over a noisy channel in terms of the mutual information $I(X; Y)$ between input $X$ and output $Y$ has been established in [16]. The relationship between the data transmission rate and the maximal error probability has been demonstrated by various bounds in the finite blocklength regime [17]. In these bounds, the so-called information density plays an essential...
role, which is defined as follows:

\[ i(X; Y) = \log \frac{dP_{XY}}{d(P_X \times P_Y)}(X, Y) = \log \frac{dP_{Y|X=x}}{dP_Y}(y). \]

In this article, we consider a legacy system with multiple antennas on both the RF source and receiver and a single antenna tag. We analyze achievable and converse bound for the system together with normal approximations. We first study a multiple-input multiple-output (MIMO) channel of the legacy system from the information-theoretic point of view. We consider the channel between the receiver, the RF source, and the channel through the tag as a whole [7]. We derive a corresponding achievability bound, defined as a lower bound on the size of a code that can be guaranteed to exist with a given arbitrary blocklength and maximal error probability. Moreover, we further establish a corresponding converse bound, which is an upper bound on the size of any code with a given arbitrary blocklength and maximal error probability. We demonstrate that the number of transmitter and receiver antennas and the information variance in [17] and [18] affect the convergence speed to the maximal achievable rate as the blocklength \( n \) increases. We also show that the SNR gap between the tag symbol and RF source signal at a fixed BER decreases as the blocklength \( n \) increases.

A. Contributions

1) We use the Berry–Esseen theorem as a fundamental basis to provide achievability and converse bounds on the achievable rate \( R \) for a legacy system with multiple transmit and receive antennas. For our achievability bound, we utilize the Berry–Esseen theorem, the mutual information, and the information variance under the condition of the probability of the tag symbol to get the bound. Furthermore, we exploit the Mellin transform and Meijer G-function to obtain a maximum on the auxiliary channel’s probability density function (PDF), a product of \( m \) copies of PDF of Gamma distributed variables. Then, we apply the Lebesgue measure to get the maximum of its output space. To complete our achievability and converse bounds, we utilize the different modulation schemes in our legacy system.

2) To reduce the complexity of multiple integrals for deriving the mutual information and the information variance, we use the saddle point approximation and the Taylor expansion to obtain closed-form expressions of the mutual information and the information variance.

3) We apply a low-complexity maximum-likelihood (ML) detection to compute the average error probability of the tag symbol based on the received signal and estimated RF source signal for a variety of transmitter and receiver antennas. We determine the relationship between the average error probability of the tag symbol and the maximal error probability of the RF source signal as a function of the blocklength \( n \) and the number of transmitter and receiver antennas \( t \) and \( r \). We utilize different modulation schemes, i.e., BPSK and QPSK, and different coding methods, i.e., the EBCH code and the polar code, in our legacy system and AmBC system to validate the derived bounds and the error probability of the tag symbol.

B. Notation

We use lowercase letters to represent scalars. For a scalar complex value \( x \), we denote by \( |x|, \Re{x}, \text{and} \Im{x} \) its modulus, real part, and imaginary part, respectively. A bold uppercase letter such as \( \mathbf{X} \) denotes a random vector, and its realization is represented by a bold lowercase symbol such as \( x \). We use \( I_{di} \) to denote the identity matrix of size \( a \times a \). The superscript \( \mathbf{A}^H \) denotes the trace of the matrix \( \mathbf{A} \). \( \text{tr}(\mathbf{A}) \) denotes the trace of the matrix \( \mathbf{A} \). \( \mathcal{C}(\mu, \sigma^2) \) represents a complex Gaussian distribution with a mean of \( \mu \) and a variance of \( \sigma^2 \); in particular, \( \text{a complex Gaussian random variable} \ X \sim \mathcal{C}(0, \sigma^2) \) has independent and identically distributed (i.i.d.) zero mean Gaussian real and imaginary components which are circularly symmetric. \( \mathbb{E}[\cdot] \) stands for the Frobenius norm of a matrix \( \mathbf{A} \), which is \( \|\mathbf{A}\|_F = \sqrt{\text{tr}(\mathbf{A}^H \mathbf{A})} \). \( \mathbb{R}_+ \) stands for the non-negative real line; in particular, \( \mathbb{R}_+^m \) is the nonnegative m-dimensional real Euclidean spaces. \( \log(\cdot) \) denotes the binary logarithm. \( \mathbb{E}[x] \) denotes the statistical expectation of \( x \), and \( \mathbb{P}[A] \) denotes the probability of an event \( A \).

The remainder of this article is organized as follows. Section II describes the system model. Section III is devoted to performance analysis for our system model. Numerical results are presented in Section IV. Finally, Section V concludes this article.

II. System Model

Let us consider input and output sets \( \mathcal{A}^n, \mathcal{D}, \mathcal{B}^m, \hat{\mathcal{D}} \) and the conditional probability measures \( P_{Y|X} : \mathcal{A}^n \mapsto \mathcal{B}^m \), and \( P_{\hat{d}|d} : \mathcal{D} \mapsto \mathcal{D} \). We denote an RF source’s codebook with \( M \) codewords by \( (c_1, \ldots, c_M) \). An RF source signal decoder, which is defined as a random transformation \( P_{Z|Y} : \mathcal{B}^m \mapsto \{1, \ldots, M\} \), satisfies a maximal error probability of the RF source signal as follows:

\[ P_{Z|X}(i|C_i) \geq 1 - \epsilon_{\text{source}}, \quad i = 1, \ldots, M \]  \hspace{1cm} (1)

where \( \epsilon_{\text{source}} \) is the maximal error probability of the RF source signal. Additionally, a tag symbol decoder, which is a random transformation \( P_{\hat{d}|d} \), satisfies an average error probability of the tag symbol as follows:

\[ \epsilon_{\text{tag}} = 1 - \frac{1}{|D|} \sum_{i=0}^{|D|} P_{\hat{d}|d}(\hat{d}_i|d_i) \]  \hspace{1cm} (2)

where \( \hat{d}_i \in \mathcal{D} \) and \( d_i \in \mathcal{D} \), and \( \epsilon_{\text{tag}} \) is the average error probability of the tag symbol. The RF source’s codebook and decoder whose maximal error probability is smaller than \( \epsilon_{\text{source}} \) are called an \( (n, M, \epsilon_{\text{source}}) \) code and its corresponding coding rate is defined as \( R = (\log M/n) \).

We consider an AmBC system with one RF source, one receiver, and one backscatter tag with no battery, as depicted in Fig. 1. In the legacy system, the RF source and the receiver have \( t \) and \( r \) antennas, respectively, and the tag has a single antenna. We let \( m = \min(t, r) \) and denote by \( H_{xy}, H_{yr} \),
work in [17], our achievability and converse bounds examine the legacy MIMO system with finite alphabet constraints, i.e., BPSK and QPSK modulated signal. Due to the high complexity of multiple integrals in the calculation, we obtain the closed-form expression by applying the saddle point approximation to let our bounds have practical implementation. Furthermore, in the second section, we examine the relation between the maximal error probability of the RF source signal, $\epsilon_{\text{source}}$, and the average error probability of the tag symbol, $\epsilon_{\text{tag}}$, with respect to the blocklength $n$. Combining our derived bounds and the relation between $\epsilon_{\text{source}}$ and $\epsilon_{\text{tag}}$, we would find an $(n, M, \epsilon_{\text{source}})$ code to achieve a specific level of $\epsilon_{\text{tag}}$.

A. Achievability and Converse Bounds

In this section, our achievability and converse bounds for the legacy MIMO system are presented below.

**Theorem 1:** We consider a communication system having the finite input alphabet $\mathcal{A}$ and $\mathcal{D}$, and the continuous output alphabet $\mathcal{B}$. Let $p(Y; H(X))$ be the corresponding conditional PDF on $B$ for all $X \in \mathcal{A}$, where $H$ is a channel matrix. The input distribution $P(X) \triangleq [q_0, \ldots, q_L]^T$, where $q_i \in [0, \ldots, L_{\text{max}}]$ is equiprobable.

Thus for the legacy MIMO channel and arbitrary $0 < \epsilon_{\text{source}} < 1$, we have the achievability and converse bounds

$$I(X; Y|D) - \frac{\sqrt{U(X; Y|D)}}{n} Q^{-1}(\epsilon_{\text{source}}) + \frac{1}{n} + O\left(n^{-3/2}\right) \leq R \leq I(X; Y|D) - \frac{\sqrt{U(X; Y|D)}}{n} Q^{-1}(\epsilon_{\text{source}}) + \frac{m + 1}{2} \log n + O\left(n^{-3/2}\right)$$

where $I(X; Y|D)$ and $U(X; Y|D)$ denote the mutual information and the information variance with respect to $D$, respectively, and $Q$ is the complementary Gaussian cumulative distribution function $Q(x) = \frac{1}{\sqrt{2\pi}} \int_{x}^{\infty} e^{-u^2/2} du$.

The proof of Theorem 1 can be found in Appendix A.

In order to apply Theorem 1 in our system model, we make the assumption that the tag symbol $d$ is predetermined. We then apply a two-step decoding process: the first step is to decode the received signal $Y$ with the knowledge of $d$, and the second step is to decode the tag symbol $d$ with the use of the decoded received signal. According to Section II, we know that the set of $d$ consists of two different values, which are “0” and “+1,” respectively. One way to obtain the mutual information $I(X; Y|D)$ and the information variance $U(X; Y|D)$ is by taking a weighted sum on the probability of occurrence of that particular value of $d$.

1) **Case 1:** When $d = 0$, we have

$$I(X; Y|d = 0) \triangleq \int_{0}^{\infty} \int_{-\infty}^{\infty} \sum_{X \in \mathcal{A}} P(X)p(Y; H_0|X)$$

**III. Performance Analysis**

In this section, we provide the definitions of achievability and converse bounds. The achievability and converse bounds are essential to the proof of the channel coding theorem. The achievability bound is a lower bound on the size of a code that can be guaranteed to exist with a given arbitrary blocklength and error probability. The converse bound is an upper bound on the size of any code with a given arbitrary blocklength and error probability. Moreover, compared with the original

---

Fig. 1. System model for AmBCs.
\[
\log \left\{ \frac{p(Y, H_0|X)}{\sum_{X' \in \mathcal{A}'} p(X')p(Y, H_0|X')} \right\} dYdH_0 \\
= \sum_{X \in \mathcal{A}} \frac{q_X}{\det(\pi I_2)} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left( -\log e \sum_{j=0}^{m-1} \frac{1}{2} \left| y_j - h_{0,i}x_j \right|^2 - \log \left[ \sum_{k=0}^{m-1} \sum_{x'_k \in \mathcal{A}} q_k \exp \left\{ -\frac{1}{2} \left| y_k - h_{0,i}x'_k \right|^2 \right\} \right] \right) jydh_0 \\
(7)
\]

and

\[
U(X; Y|d = 0) \\
= \sum_{X \in \mathcal{A}} \frac{q_X}{\det(\pi I_2)} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left( -\log e \sum_{j=0}^{m-1} \frac{1}{2} \left| y_j - h_{0,i}x_j \right|^2 \\
- \log \left[ \sum_{k=0}^{m-1} \sum_{x'_k \in \mathcal{A}} q_k \exp \left\{ -\frac{1}{2} \left| y_k - h_{0,i}x'_k \right|^2 \right\} \right] \right) jydh_0 \\
- \left[ I(X; Y) \right]^2 \\
(9)
\]

where \( P(X) = [q_0, \ldots, q_t]^T \) denotes the input distribution, \( q_t = [(1/2), (1/2)] \) and \( q_4 = [(1/4), (1/4), (1/4), (1/4)] \), for BPSK and QPSK, respectively, and \( p(h_{0,i}) \) denotes the PDF of the channel between one transmit antenna and one receive antenna. In order to reduce the complexity of multiple integrals in the mutual information and the information variance, we give the closed-form approximation of (7) and (9) as follows:

\[
I(X; Y|d = 0) \\
\approx \log |A| - H_{h_0} \left[ \frac{1}{|A'|} \sum_{i=1}^{|A'|} \log \left( \frac{1}{|A'|} \sum_{j=1}^{d} \exp \left\{ -\frac{\| h_0(x_i - x_j) \|^2}{3 - \exp \left\{ -\| h_0(x_i - x_j) \|^2 / 4 \right\} \right\} \right) \right] \\
+ \mathbb{E}_{h_0} \left[ \frac{1}{|A'|} \sum_{i=1}^{d} \log \left( \frac{1}{|A'|} \sum_{j=1}^{d} \exp \left\{ -\frac{\| h_0(x_i - x_j) \|^2}{6 - \exp \left\{ -\| h_0(x_i - x_j) \|^2 / 16 \right\} \right\} \right) \right] \\
(10)
\]

and

\[
U(X; Y|d = 0) \approx - \left[ (\log |A|) - I(X; Y) \right]^2 \\
+ \mathbb{E}_{h_0} \left[ \frac{1}{|A'|} \sum_{i=1}^{d} \log \left( \frac{1}{|A'|} \sum_{j=1}^{d} \exp \left\{ -\frac{\| h_0(x_i - x_j) \|^2}{6 - \exp \left\{ -\| h_0(x_i - x_j) \|^2 / 16 \right\} \right\} \right) \right] \\
(11)
\]

where the proof of (10) and (11) can be found in Appendix B.

2) Case 2: When \( d = +1 \), in order to obtain the mutual informations and the information variance \( I(X; Y|d = +1) \) and \( U(X; Y|d = +1) \), we just change \( h_0 \) in (10) and (11) to \( h_1 \).

After getting the mutual information and the information variance of each case, then applying Theorem 1, we have

\[
I(X; Y|D) = \sqrt{\frac{U(X; Y|D)}{n}} Q^{-1}(\epsilon_{\text{source}}) + O\left(n^{-3/2}\right) \\
\leq R \leq I(X; Y|D) = \sqrt{\frac{U(X; Y|D)}{n}} Q^{-1}(\epsilon_{\text{source}}) \\
+ \frac{m + 1}{2} \log n \cdot n + O\left(n^{-3/2}\right) \\
(12)
\]

where

\[
I(X; Y|D) = \sum_{i=0}^{d} \mathbb{P}[d = d_i] I(X; Y|d = d_i) \\
(13)
\]

\[
U(X; Y|D) = \sum_{i=0}^{d} \mathbb{P}[d = d_i] U(X; Y|d = d_i). \\
(14)
\]

To compare with our result, we calculate the capacity of the channel whose input is a circularly symmetric complex Gaussian with zero mean and covariance \( P/\pi I_2 \). The theorem is shown below.

**Theorem 2 [22]:** Under the power constraint \( P \), we assume the same channel with the same number of transmitting and receiving antennas as our system model. Its capacity, as determined by the complex Gaussian input, is equal to

\[
C = \log \left( 1 + \frac{P}{T} g_0 \right) - \log \left( 1 + \frac{P}{T} g_1 \right) \\
\]

where \( g_0 \) and \( g_1 \) denote the eigenvalues of the matrix \( \mathbf{H}_0 \mathbf{H}_0^H \) and \( \mathbf{H}_1 \mathbf{H}_1^H \), respectively, where \( \mathbf{H}_0 \) and \( \mathbf{H}_1 \) are from (4), and their PDFs are given by

\[
p_{g_0}(g) = \frac{1}{2m} \sum_{i=1}^{m} \left( \frac{i-1}{i+1} \right) \left( \frac{\max[r,i-m]}{2} \right)^2 \cdot \left( \frac{\max[r,i-m]}{2} \right) \exp \left\{ -\frac{x}{4} \right\} \\
(16)
\]
and
\[ P_{R_1}(x) = \frac{1}{m} \sum_{i=1}^{m} (i-1)! \left( L_{i-1}^{\max(r,t)-m}(x) \right)^2 \exp \left\{-\frac{x}{2} \right\} \]
where \( L_k^m(x) \) is the associated Laguerre polynomial of order \( k \).

Thus, we have
\[
C = \frac{\mathbb{P}[d = 0]}{2} \int_{0}^{\infty} \log \left( 1 + \frac{P}{t g_0} \right) \sum_{i=1}^{m} \frac{(i-1)!}{(i-1 + \max(r,t)-m)!} \left( L_{i-1}^{\max(r,t)-m} \left( \frac{g_0}{2} \right) \right)^2 \exp \left\{-\frac{g_0}{2} \right\} \] 
+ \mathbb{P}[d = +1] \int_{0}^{\infty} \log \left( 1 + \frac{P}{t g_1} \right) \sum_{i=1}^{m} \frac{(i-1)!}{(i-1 + \max(r,t)-m)!} \left( L_{i-1}^{\max(r,t)-m} \left( \frac{g_1}{2} \right) \right)^2 \exp \left\{-\frac{g_1}{2} \right\}.
\]

B. Capacity Analysis of the AmBC System

In this section, we analyze the capacity of the AmBC system. We define \( C_{\text{AmBC}}(D; Y) \) as the capacity of the AmBC system. We assume that the input distribution \( P(X) \) is equiprobable and that the RF source signal is known in advance. Then, we can obtain the conditional capacity of the AmBC system. By taking a weighted sum on the probability of occurrence of that particular value of \( X, P(X) \), we have
\[
C_{\text{AmBC}}(D; Y|X) = \int_{0}^{\infty} \int_{0}^{\infty} \sum_{i=0}^{1} \mathbb{P}[d = i] p(Y|X) \log \frac{p(Y|X)}{\sum_{i=0}^{1} \mathbb{P}[d = i'] p(Y|X)} \] 
\[ = \mathbb{E}_h \left[ \sum_{i=0}^{1} \mathbb{P}[d = i] \log \left( \frac{1}{\mathbb{P}[d = i']} \right) \exp \left\{- \frac{||x_i(h_i - h_y)||^2}{3} \right\} \right] \] 
(20)
where the proof of (20) can be easily obtained by the similar method in Appendix B. Then, we have
\[
C_{\text{AmBC}}(D; Y) = \sum_{i=0}^{1} \mathbb{P}[d = i] C_{\text{AmBC}}(D; Y|X).
\]

C. Relation Between \( \varepsilon_{\text{source}} \) and \( \varepsilon_{\text{tag}} \)

In this section, we provide the relation between the maximal error probability of the RF source signal, \( \varepsilon_{\text{source}} \), and the average error probability of the tag symbol, \( \varepsilon_{\text{tag}} \), with respect to the blocklength \( n \). Due to the fact that the tag symbol affects the channel instead of being directly sent to the receiver end, we do not directly use the ML detection to estimate the tag symbol \( \hat{d} \). According to [23], the computational complexity of the ML detector is \( |A|^\log_{M}|D| \), which grows exponentially as the alphabet size of the modulation scheme increases. To reduce the complexity, we propose a low-complexity ML detection to decode the tag symbol from the received signal \( Y \). The computational complexity of low-complexity ML detection is \( \log_{M}|A||D| \), which is lower than that of the original ML detection.

The estimated tag symbol \( \hat{d} \) is given as follows:
\[
\hat{d} \equiv \arg \min_{d} \left\| Y - \hat{X} h_y - d X G \right\|^2
\]
where \( \hat{X} \) is the estimated RF source signal at the receiver side and \( G = A h_{r} h_y \). Then, we let
\[
Z = \frac{\mathbb{E}}{n} \mathbb{E} \left[ \frac{G^H \hat{X} H}{\|G\|^2} \right] d
\] 
\[ = \frac{\mathbb{E}}{n} \mathbb{E} \left[ \frac{G^H \hat{X} H}{\|G\|^2} \right] d
\] 
\[ + \frac{\mathbb{E}}{n} \mathbb{E} \left[ \frac{G^H \hat{X} X - \hat{X} H G}{\|G\|^2} \right] d + \omega \] 
(24)
where
\[
\omega = \frac{\mathbb{E}}{n} \mathbb{E} \left[ \frac{G^H \hat{X} W}{\|G\|^2} \right]
\]
and \( \omega \sim \mathcal{N}(0, 1/(n\|G\|^2)) \).

Due to the low-complexity ML detection, if \( Z > Z_{th} \), \( \hat{d} = +1 \), otherwise, \( \hat{d} = 0 \), where \( Z_{th} \) is the decision threshold.

1) Case 1: When \( \hat{X} = X \), we have \( 1 - \delta < (1/n)\mathbb{E} \left[ \frac{G^H \hat{X} H}{\|G\|^2} \right] < 1 + \delta \), where \( \delta \leq 1 \), and \( (1/n)\mathbb{E} \left[ \frac{G^H \hat{X} X - \hat{X} H G}{\|G\|^2} \right] \) is a discrete real number, then we can treat this value as the coefficient of \( d \), thus
\[
Z = \frac{\mathbb{E}}{n} \mathbb{E} \left[ \frac{G^H \hat{X} H}{\|G\|^2} \right] d + \omega.
\]
The coefficient of \( d \) only has value in the real part which means \( \mathbb{E} \left[ \frac{G^H \hat{X} H}{\|G\|^2} \right] = 0 \). Thus it shows that \( Z \) obeys the Gaussian distribution. Therefore, under \( \hat{X} = X \), the error probability of the event \( \hat{d} \neq d \) can be expressed as below
\[
\mathbb{P}[\hat{d} \neq d | \hat{X} = X] = Q \left( \sqrt{n\|G\|^2} \right)
\]
\[ + \mathbb{E} \left[ \frac{G^H \hat{X} X - \hat{X} H G}{\|G\|^2} \right] \right),
\]
(27)

Remark 1: For \( 2 \times 2 \) MIMO, we select \( Z_{th,1} = 0.75 \), while for \( 3 \times 3 \) MIMO, we select \( Z_{th,1} = 0.95 \).

2) Case 2: when \( \hat{X} \neq X \), we have \( -\delta < (1/n)\mathbb{E} \left[ \frac{G^H \hat{X} H}{\|G\|^2} \right] < +\delta \), where \( \delta \leq 1 \),
and (1/n)Trace\{\mathbf{G}^H \hat{\mathbf{X}} \hat{\mathbf{X}} \mathbf{G}/\|\mathbf{G}\|^2\} is a discrete complex number, thus we have
\[
Z = \frac{1}{n} \text{Trace}\left\{\mathbf{G}^H \hat{\mathbf{X}} \hat{\mathbf{X}} \mathbf{G}/\|\mathbf{G}\|^2\right\}d
+ \frac{1}{n} \text{Trace}\left\{\mathbf{G}^H (\hat{\mathbf{X}} \mathbf{X} - \hat{\mathbf{X}} \hat{\mathbf{X}}) \mathbf{H}_{sr}\right\}
+ \omega. \tag{28}
\]
Thus for \(d = +1\), we have
\[
P[d = 0, d = +1 | \hat{\mathbf{X}} \neq \mathbf{X}]
= Q\left(\sqrt{n} \|\mathbf{G}\| \left(- Z_{th,2}
+ \frac{1}{n} \mathbb{E}\left[\text{Trace}\left\{\frac{\mathbf{G}^H \hat{\mathbf{X}} \hat{\mathbf{X}} \mathbf{G}/\|\mathbf{G}\|^2}{\mathbf{G}^H (\hat{\mathbf{X}} \mathbf{X} - \hat{\mathbf{X}} \hat{\mathbf{X}}) \mathbf{H}_{sr}}\right\}\right]
+ \frac{G^H \hat{X} \hat{X} G}{\|G\|^2}\right)\right) \tag{29}
\]
and for \(d = 0\), we have
\[
P[d = +1, d = 0 | \hat{\mathbf{X}} \neq \mathbf{X}]
= Q\left(\sqrt{n} \|\mathbf{G}\| \left(- Z_{th,2}
+ \frac{1}{n} \mathbb{E}\left[\text{Trace}\left\{\frac{\mathbf{G}^H \hat{\mathbf{X}} \hat{\mathbf{X}} \mathbf{G}/\|\mathbf{G}\|^2}{\mathbf{G}^H (\hat{\mathbf{X}} \mathbf{X} - \hat{\mathbf{X}} \hat{\mathbf{X}}) \mathbf{H}_{sr}}\right\}\right]
+ \frac{G^H \hat{X} \hat{X} G}{\|G\|^2}\right)\right) \tag{30}
\]

\textit{Remark 2:} For \(2 \times 2\) MIMO, we select \(Z_{th,2} = 0.75\), while for \(3 \times 3\) MIMO, we select \(Z_{th,2} = 0.95\).

Once getting the average error probability of the tag symbol for both cases of \(\hat{\mathbf{X}} = \mathbf{X}\) and \(\hat{\mathbf{X}} \neq \mathbf{X}\), we have the equation which expresses the relationship between \(\epsilon\) and the average error probability of the tag symbol, \(\epsilon_{\text{tag}}\) by combining (27), (29), and (30)
\[
\epsilon_{\text{tag}} = (1 - \epsilon_{\text{source}})Q\left(\sqrt{n} \|\mathbf{G}\| \left(- Z_{th,1} + \frac{1}{n} \mathbb{E}\left[\text{Trace}\left\{\mathbf{G}^H \hat{\mathbf{X}} \hat{\mathbf{X}} \mathbf{G}/\|\mathbf{G}\|^2\right\}\right]\right)\right)
\]

\textbf{IV. Numerical Results}

In this section, we resort to numerical simulation to evaluate the proposed studies. We consider a legacy system consisting of multiple transmitter antennas, the tag with a single antenna and multiple receiver antennas. We assume all the channels, i.e., the channels between the transmitter and the tag, the tag and the receiver, and the transmitter and the receiver, are independent. Assuming that all the channels are Rayleigh fading channels, and the SNR is \(-5\) dB. Fig. 2 shows the mutual information of the legacy system in (13) for the BPSK and QPSK modulated signal for \(P[d = +1]\) from 0 to 1 with \(2 \times 2\) MIMO and \(3 \times 3\) MIMO, respectively. We observe that as the probability of the tag symbol \(d = +1\) increases, \(I(X; Y|D)\) increases accordingly, regardless of the modulation scheme of the RF source signal and the number of transmitter
and receiver antennas. Additionally, the gap between different modulated signals decreases as the number of transmitter and receiver antennas increases.

Fig. 3 demonstrates the value of $C_{\text{AmBC}}(D; Y)$ in (21) for BPSK and QPSK modulated RF source signal with $P[d = +1]$ from 0 to 1 over the $2 \times 2$ MIMO and $3 \times 3$ MIMO systems, respectively. By comparing the results in Fig. 2, we easily observe that under the same level of SNR, i.e., SNR = $-5$ dB, the mutual information of the legacy system $I(X; Y|D)$ is much larger than the capacity of the AmBC system $C_{\text{AmBC}}(D; Y)$. The mutual information of the legacy system as shown in Fig. 2 increases as $P[d = +1]$ increases. While the mutual information of the AmBC system $C_{\text{AmBC}}(D; Y)$ reaches the peak when $P[d = +1] = 0.5$, see Fig. 3, therefore, in the following simulations, we choose $P[d = +1] = 0.5$.

We consider the same legacy MIMO system as above, and we set the maximal error probability $\epsilon = 10^{-3}$. Assuming that all the channels are Rayleigh fading channels, $P[d = 0] = P[d = +1] = 0.5$ and the SNR is $-5$ dB. Fig. 4(a) and (b) show the numerical results of the derived bounds with BPSK and QPSK modulated signals and the capacity for $t = r = 2$ and $t = r = 3$, respectively. From Fig. 4(a), we can see that the capacity is 0.9611 bit/(channel use), which is calculated from (18), and the maximal achievable rate for the BPSK modulated signal is 0.7151 bit/(channel use), which is obtained based on (13). The blocklength $n$ required to achieve above 70% and 80% of its maximal achievable rate start at $n = 180$ and $n = 420$, respectively. The gap between the capacity and its maximal achievable rate is 0.2460 bit/(channel use). With the QPSK modulation, the maximal achievable rate is 0.8772 bit/(channel use), and the blocklength $n$ required to achieve above 70% and 80% of its maximal achievable rate start at $n = 380$ and $n = 860$, respectively. The gap in the QPSK case is 0.0839 bit/(channel use). In Fig. 4(b), we only change the number of transmitter and receiver antennas from $t = r = 2$ to $t = r = 3$, and the rest parameters remain the same. The capacity, in this case, is 1.4425 bit/(channel use). The BPSK modulated signal’s maximal achievable rate is 1.3006 bit/(channel use). The blocklength $n$, which can surpass 70% and 80% of its maximal achievable rates, decreases dramatically to 100 and 240, respectively, compared with the case of $2 \times 2$ MIMO. Moreover, the gap between the capacity and the maximal achievable rate decreases to 0.1419 bit/(channel use). For QPSK modulation, its maximal achievable rate is 1.415 bit/(channel use), and the blocklength $n = 260$ and $n = 580$ are required to achieve above 70% and 80% of its maximal achievable rate, respectively. The gap also falls to 0.0269 bit/(channel use) compared with the case for $2 \times 2$ MIMO. The findings are summarized in Table I. From Fig. 4, we can conclude that:

1) as the number of transmitter and receiver antennas increases, the maximal achievable rates of the BPSK and QPSK modulated signal accelerate, which indicates that the gap between the maximal achievable rate for different modulation schemes and the capacity decreases at the same SNR level;

2) the required blocklength $n$ falls significantly to achieve a given fraction of the maximal achievable rate as the number of transmitter and receiver antennas increases.

The information variance $U(X; Y|D)$ in (14) shows how quickly the performance converges to the maximal achievable
rate as blocklength $n$ grows. In the case of the BPSK and QPSK modulations shown in Fig. 4(a), the information variances for the BPSK and QPSK modulated signal are 0.9281 and 2.8267, respectively. From Fig. 4(b), we can see that the information variances are 1.6990 and 4.9729, respectively. Additionally, if the target is to transmit at a fraction of the maximal achievable rate 0 < $\eta$ < 1 with a predetermined $\epsilon_{\text{source}}$, the relationship between the required blocklength $n$ and the information variance is given as follows:

$$n \approx \frac{U(X; Y|D)}{I(X; Y|D)} \left( \frac{Q^{-1}(\epsilon_{\text{source}})}{1 - \eta} \right)^2.$$

To validate our results, we transform the achievability and converse bounds in Theorem 1 to the lower and upper bounds on the average error probability of the RF source signal. From Theorem 1, we have the achievability and converse bounds on maximal error probability. Since there always exists an $(n, M, \epsilon_{\text{source}})$-code in the maximal error probability $\epsilon_{\text{source}}$ that guarantees the existence of an $(n, M', \epsilon'_{\text{source}})$-code in the average error probability $\epsilon_{\text{source}}$, for any $\epsilon'_{\text{source}} < \epsilon_{\text{source}} < 1$ and 0 < $\xi$ < 1, where $M' = 2nR/(1-\xi)$ and $\epsilon'_{\text{source}} = \xi\epsilon_{\text{source}}$.

From (12), we have $I(X; Y|D) - \sqrt{U(X; Y|D)/nQ^{-1}(\epsilon_{\text{source}})} + O(n^{-3/2}) \leq R \leq I(X; Y|D) - \sqrt{U(X; Y|D)/nQ^{-1}(\epsilon_{\text{source}})} + (m+1)\log n/(2n) + O(n^{-3/2}).$

We transform the achievable and converse bounds into the lower and upper bounds on the maximal error probability of the RF source signal as follows:

$$Q\left( \frac{I(X; Y|D) + \frac{m+1}{2} \log n}{\sqrt{U(X; Y|D)/n}} - R \right) \leq \epsilon_{\text{source}}$$

$$\leq Q\left( \frac{I(X; Y|D) - R}{\sqrt{U(X; Y|D)/n}} \right).$$

Therefore, we obtain the lower and upper bounds on the average error probability of the RF source signal, i.e., $\epsilon_{\text{source}}$, which is shown below

$$\xi Q\left( \frac{I(X; Y|D) + \frac{m+1}{2} \log n}{\sqrt{U(X; Y|D)/n}} - R(1 - \xi) \right) \leq \epsilon_{\text{source}}$$

$$\leq \xi Q\left( \frac{I(X; Y|D) - R(1 - \xi)}{\sqrt{U(X; Y|D)/n}} \right).$$

Table I: Required blocklength to achieve a given fraction of the maximal achievable rate for a legacy MIMO system over a Rayleigh fading channel, SNR = −5 dB and $\epsilon = 10^{-3}$, and $P(d) = [0.5, 0.5]$.

|               | 2 × 2 MIMO | 3 × 3 MIMO |
|---------------|------------|------------|
|               | BPSK       | QPSK       | BPSK       | QPSK       |
| Required $n$ to Achieve 70% of the Maximal Achievable Rate | 180        | 380        | 100        | 260        |
| Required $n$ to Achieve 80% of the Maximal Achievable Rate | 420        | 860        | 240        | 580        |

By utilizing the polar code with a successive cancelation list (SCL) decoder and the extended BCH code with an ordered statistic decoder (OSD), we validate our derived results.

In Fig. 5(a), we set the number of the transmitter and receiver antennas to 2, the modulation scheme to BPSK, the coding rate $R = 0.5$, and the blocklength $n = 128$. All the simulations are averaged over $10^6$ Monte Carlo realizations. We choose two coding methods: one is the (128, 264)-polar code with SCL decoder (the list size is $L = 32$) and the other is the (128, 264)-EBCCH code with OSD decoder (the order is chosen to 4). We observe that for the EBCCH code, at the average error probability level of $10^{-2}$ and $10^{-4}$, the gap between the simulation result and the lower bound increases from 2.5 to 4 dB, respectively. As EbNo increases, the gap increases accordingly. The simulation result of the EBCCH code is slightly better than the one of the polar code. However, it still shows that the EBCCH code is still better than the polar code at the blocklength $n = 128$. In Fig. 5(b), we set the modulation scheme to QPSK and keep the rest parameters the same as in Fig. 5(a). We observe that the overall performance between the BPSK modulation scheme and the QPSK modulation scheme in the
2 × 2 MIMO system is similar. The simulation results and our derived bounds validate the observation.

In Fig. 6(a), we change the number of the transmitter and receiver antennas to 3, the modulation scheme to BPSK, and remain the rest parameters the same as in Fig. 5. At first, we compare the performance of two codes in different MIMO systems, i.e., 2 × 2 MIMO and 3 × 3 MIMO systems. At the same EbNo level of −4 dB, the average error probability drops from 0.4375 to 0.0907 for the EBCH code and from 0.5509 to 0.1089 for the polar code. At the same average error probability level of 10^{−2}, the gaps between different MIMO systems are 3 and 3.5 dB for the EBCH code and polar code, respectively. We observe that the bounds are closer in the 3 × 3 MIMO system than in the 2 × 2 MIMO system. In Fig. 6(b), we change the modulation scheme from BPSK to QPSK, and remain the rest parameters the same as in Fig. 6(a). The comparison between these two figures shows that the gap between the performance of the two codes in different modulation schemes becomes larger, i.e., 1 and 2 dB (at the same average error probability level of 10^{−2}) for the EBCH code and polar code, respectively.

Fig. 7. Comparison between ML detection and our proposed method with different coding methods with BPSK modulation over a Rayleigh fading channel and transmit antennas \( t = 2 \) and receive antennas \( r = 2 \), and transmit antennas \( t = 3 \) and receive antennas \( r = 3 \), respectively.

In Fig. 7, we set the number of the transmitter and receiver antennas to 2 and 3, respectively, the RF source signal modulation scheme to BPSK, and keep the rest parameters the same as in Figs. 5 and 6. We compare the performance of our proposed method and ML detection with different coding methods of the RF source signal, i.e., the EBCH code and polar code. When EbNo is small, the gaps between these two methods are 0.5 and 0.8 dB for the 2 × 2 MIMO and 3 × 3 MIMO system, respectively. As EbNo increases, the gap slightly increases to 1 and 1.2 dB for 2 × 2 MIMO and 3 × 3 MIMO systems, respectively. Additionally, we observe that as EbNo increases, the gap between different codes vanishes for both 2 × 2 MIMO and 3 × 3 MIMO systems.

Fig. 8 demonstrates the relation between the blocklength \( n \) and the probability of \( \mathbb{P}[\hat{d} \neq d | \hat{\mathbf{X}} = \mathbf{X}] \), \( \mathbb{P}[\hat{d} = 0, d = +1 | \hat{\mathbf{X}} \neq \mathbf{X}] \), and \( \mathbb{P}[\hat{d} = +1, d = 0 | \hat{\mathbf{X}} \neq \mathbf{X}] \), which are obtained from (27), (29), and (30), respectively. From Fig. 8(a), we observe that under the condition of \( \hat{\mathbf{X}} = \mathbf{X} \), the conditional probability of \( \hat{d} \neq d \) decreases as the blocklength \( n \) increases regardless of the number of transmitter and receiver antennas. When \( n = 200 \), the probability for \( \mathbb{P}[\hat{d} \neq d | \hat{\mathbf{X}} = \mathbf{X}] \) for 2 × 2 MIMO is 6.56 × 10^{−4}, while the one for 3 × 3 MIMO is 9.53 × 10^{−5}. When \( n \) moves to 1000, the probability for 2 × 2 MIMO decreases to 4.29 × 10^{−5} while the one for 3 × 3 MIMO drops to 1.23 × 10^{−7}. Furthermore, when \( n \) increases to 2000, the conditional probability, \( \mathbb{P}[\hat{d} \neq d] \) under the condition of \( \hat{\mathbf{X}} = \mathbf{X} \) for 2 × 2 MIMO case decreases to 1.30 × 10^{−5} in the meanwhile, that probability for 3 × 3 MIMO case falls to 2.41 × 10^{−8}. The decreasing trend of 2 × 2 MIMO is much slower than that for 3 × 3 MIMO from \( n = 0 \) to 1000. After \( n = 1000 \), the tendencies for both cases are flattened. Compared with ML detection, when \( n \) is less than 800, there is a small gap between the low-complexity ML detection that we mainly use in this article and the ML detection for 2 × 2 and 3 × 3 MIMO. As \( n \) grows larger, the performances of the two methods, i.e., the ML and the low-complexity ML detection methods, are basically the same. Fig. 8(b) and (c) show that under the condition of \( \hat{\mathbf{X}} \neq \mathbf{X} \) and \( d = +1 \), the conditional probability of \( \hat{d} \neq d \) increases as the blocklength \( n \) increases regardless of the number of the
transmitter and receiver antennas. Moreover, the figures also demonstrate that under the condition of \( \hat{X} \neq X \) and \( d = 0 \), the conditional probability of \( \hat{d} \neq d \) decreases as the blocklength \( n \) increases regardless of the number of transmitter and receiver antennas. Compared with the ML detection, for \( 2 \times 2 \) MIMO case, when \( n \) increases from 0 to 2000, the gap between the ML and the low-complexity ML detection shrinks to a constant, while the gap for \( 3 \times 3 \) MIMO falls to a very small margin.

B. Proof of Theorem 1

To complete this section, we need to introduce an important tool which is the Berry–Esseen theorem.

\textbf{Theorem 3:} Let \( X_k, k = 1, \ldots, n \) be independent with

\[
\mu_k = \mathbb{E}[X_k], \quad \sigma^2 = \text{Var}[X_k], \quad t_k = \mathbb{E}\left[|X_k - \mu_k|^3\right] \\
\sigma^2 = \sum_{k=1}^{n} \sigma_k^2 \quad \text{and} \quad T = \sum_{k=1}^{n} t_k.
\]

Fig. 8. Blocklength \( n \) versus the probability of \( \mathbb{P}[\hat{d} \neq d | \hat{X} = X] \), \( \mathbb{P}[\hat{d} = +1, d = 0 | \hat{X} = X] \), and \( \mathbb{P}[\hat{d} = 0, d = +1 | \hat{X} = X] \) over \( 2 \times 2 \) MIMO and \( 3 \times 3 \) MIMO, respectively. (a) \( \mathbb{P}[\hat{d} \neq d | \hat{X} = X] \) in (27). (b) \( \mathbb{P}[\hat{d} = 0, d = +1 | \hat{X} = X] \) in (29). (c) \( \mathbb{P}[\hat{d} = +1, d = 0 | \hat{X} = X] \) in (30).

Fig. 9. Comparison between the blocklength \( n \), the maximal error probability of the RF source signal \( \epsilon_{\text{source}} \), and the average error probability of the tag symbol \( \epsilon_{\text{tag}} \).

Basically, when \( n \) goes beyond 1000, the performances of the two detections are the same for \( 3 \times 3 \) MIMO case.

Fig. 9 demonstrates the average error probability of the tag symbol \( \epsilon_{\text{tag}} \) in (31) with different blocklength \( n \) and the maximal error probability of the RF source signal \( \epsilon_{\text{source}} \) for \( 2 \times 2 \) MIMO and \( 3 \times 3 \) MIMO, respectively. The relationship between \( \epsilon_{\text{source}} \) and the blocklength \( n \) is given by Theorem 1 and (12). Moreover, (31) illustrates the relation between \( \epsilon_{\text{source}} \) and \( \epsilon_{\text{tag}} \) and the blocklength \( n \). Therefore, we plot (31) in a 2-D plane with the \( x \)-axis representing the blocklength \( n \), the \( y \)-axis representing \( \epsilon_{\text{source}} \), and the \( z \)-axis representing \( \epsilon_{\text{tag}} \). From Fig. 9, we observe that for the same \( \epsilon_{\text{source}} \), \( 3 \times 3 \) MIMO significantly outperforms \( 2 \times 2 \) MIMO from the blocklength \( n = 0 \) to 1000. When \( n \) increases, the gap between these two cases shrinks. Furthermore, for the same blocklength \( n \), when \( \epsilon_{\text{source}} \) is less than \( 10^{-5} \), the performance of \( 3 \times 3 \) MIMO is substantially better than that of \( 2 \times 2 \) MIMO. As \( \epsilon_{\text{source}} \) becomes small, i.e., \( 10^{-10} \), the gap between the performances of two cases decreases.

\textbf{V. CONCLUSION}

In this article, we established achievability and converse bounds on the maximal achievable rate \( R \) at a given blocklength \( n \) and a maximal error probability \( \epsilon_{\text{source}} \) for a legacy MIMO system. We derived the relationship between \( \epsilon_{\text{source}} \) and \( \epsilon_{\text{tag}} \) with respect to the blocklength \( n \). The analytical results demonstrated that the number of transmit and receive antennas and the information variance \( U(X; Y|D) \) would affect the convergence speed to the maximal achievable rate as the blocklength \( n \) increases.

\textbf{APPENDIX A}

\textbf{Proof of Theorem 1}

\textbf{A. Proof of the Achievability Part}

To complete this section, we need to introduce an important tool which is the Berry–Esseen theorem.
Then, for any $-\infty < \lambda < \infty$

$$\left| \mathbb{P}\left[ \sum_{k=1}^{n} (X_k - \mu_k) \geq \lambda \sigma \right] - Q(\lambda) \right| \leq \frac{6T}{\sigma^3}. \quad (32)$$

For the proof of Theorem 1, we first need to prove that the second moment of $i(X; Y|D)$ is nonzero and its third moment is always less than infinite

$$U(X; Y|D) = \mathbb{E}\left[ |i(X; Y|D) - I(X; Y|D)|^2 \right]$$

$$= \int_{0}^{\infty} \int_{0}^{\infty} \sum_{X \in A_i} \left( P(X) p(Y, H|X) \left( 1 - p(Y, H|X) \right) \log^2 \left\{ \frac{p(Y, H|X)}{\sum_{X \in A_i} P(X)p(Y, H|X)} \right\} dYdH \right.$$

$$> 0.$$

Then, we need to show the third moment is less than infinite

$$T(X; Y|D) = \mathbb{E}\left[ |i(X; Y|D) - I(X; Y|D)|^3 \right]$$

$$\leq |B| \left( 3e^{-1} \log e \right)^3 + 2I(X; Y|D)^3.$$

According to the DT bound in [17], $\epsilon \leq \mathbb{E}\left[ \exp\left\{ -\left| i(X^n; Y^n|D) - \log \lambda \right| \right\} \right]$, where $\left| \cdot \right|$ denotes $\max\{\cdot, 0\}$. In the sequel, we prove that there exist some $\lambda$ values, so that

$$\epsilon \geq \mathbb{E}\left[ \exp\left\{ 0 \right\} \right] \left| i(X^n; Y^n|D) - \log \lambda \geq 0 \right\}$$

$$+ \mathbb{E}\left[ \exp\left\{ -i(X^n; Y^n|D) + \log \lambda \right\} \right] \left| i(X^n; Y^n|D) - \log \lambda < 0 \right\}.$$

Thus, we have

$$\mathbb{P}\left[ i(X^n; Y^n|D) \leq \log \lambda \right]$$

$$+ \lambda \mathbb{E}\left[ \exp\left\{ -i(X^n; Y^n|D) \right\} \right] \left| i(X^n; Y^n|D) - \log \lambda < 0 \right\}.$$

The first step is to obtain the upper bound of the first part of the right-hand side of (34). After applying Theorem 3, we have

$$\mathbb{P}\left[ i(X^n; Y^n|D) \leq nI(X; Y|D) - \tau \sqrt{nU(X; Y|D)} \right]$$

$$\leq \frac{6T(X; Y|D)}{\sqrt{nU(X; Y|D)^2}} + Q(\tau). \quad (35)$$

We assume

$$\log \lambda = nI(X; Y|D) - \tau \sqrt{nU(X; Y|D)}$$

and

$$\mathbb{P}\left[ i(X^n; Y^n|D) \leq \log \lambda \right] \leq \frac{6T(X; Y|D)}{\sqrt{nU(X; Y|D)^2}} + Q(\tau). \quad (37)$$

The upper bound of the second part of the right-hand side of (34) is given below. For $0 \leq i < \infty$ and any $\Delta > 0$

$$\mathbb{P}\left[ -\sqrt{nU(X; Y|D)} \left( \tau - \frac{i\Delta}{\sqrt{nU(X; Y|D)}} \right) \leq i(X^n; Y^n|D) \right]$$

$$- nI(X; Y|D) \leq -\sqrt{nU(X; Y|D)} \left( \tau - \frac{(i+1)\Delta}{\sqrt{nU(X; Y|D)}} \right). \quad (38)$$

$$\leq \frac{12T(X; Y|D)}{\sqrt{nU(X; Y|D)^2}} + Q\left( \tau + \frac{i\Delta}{\sqrt{nU(X; Y|D)}} \right) - Q\left( \tau + \frac{(i+1)\Delta}{\sqrt{nU(X; Y|D)}} \right). \quad (39)$$

where (38) is obtained by applying Theorem 3 twice. Then

$$\mathbb{E}\left[ \exp\left\{ -i(X^n; Y^n|D) \right\} \right] \left| i(X^n; Y^n|D) - \log \lambda \right\}$$

$$\leq \frac{\Delta}{\sqrt{2\pi} \sqrt{nU(X; Y|D)}} + \frac{12T(X; Y|D)}{\sqrt{nU(X; Y|D)^2}} \sum_{i=0}^{\infty} \exp\left\{ -(\log \lambda + i\Delta) \right\}. \quad (40)$$

Thus, we have

$$\lambda \mathbb{E}\left[ \exp\left\{ -i(X^n; Y^n|D) \right\} \right] \left| i(X^n; Y^n|D) - \log \lambda \right\}$$

$$\leq \left( \frac{\Delta}{\sqrt{2\pi} \sqrt{nU(X; Y|D)}} + \frac{12T(X; Y|D)}{\sqrt{nU(X; Y|D)^2}} \right) \exp(\Delta) \exp(\Delta - 1). \quad (41)$$

Substituting (37) and (41) into (34), we have

$$\mathbb{P}\left[ i(X^n; Y^n|D) \leq \log \lambda \right]$$

$$+ \lambda \mathbb{E}\left[ \exp\left\{ -i(X^n; Y^n|D) \right\} \right] \left| i(X^n; Y^n|D) - \log \lambda \right\}$$

$$\leq Q(\tau) + \frac{1}{\sqrt{nU(X; Y|D)^2}} \left( 1 + \frac{\exp(\Delta)}{\exp(\Delta - 1)} \right). \quad (42)$$

Based on (34), we can assume that the right-hand side of (42) equals to $\epsilon_{source}$, then we obtain the value of $\tau$

$$\tau = Q^{-1} \left( \epsilon_{source} - \frac{1}{\sqrt{nU(X; Y|D)^2}} \left( 1 + \frac{\exp(\Delta)}{\exp(\Delta - 1)} \right) \right), \quad (43)$$

For large $n$, the second item inside the $Q$ function of (43) vanishes. Therefore, we can obtain $\tau = Q^{-1}(\epsilon_{source}) + O(1/\sqrt{n})$. Then, we have $\log \lambda = nI(X; Y|D) - Q^{-1}(\epsilon_{source})\sqrt{nU(X; Y|D)} + O(1/\sqrt{n})$. Thus

$$R \geq I(X; Y|D) - Q^{-1}(\epsilon_{source}) \sqrt{\frac{U(X; Y|D)}{n}} + 1 + O(n^{-\frac{3}{2}}). \quad (44)$$
B. Proof of the Converse Part

We assume the transmitter is not aware of the realizations of the channel matrix $\mathbf{H}$. We denote the average power constraint

$$p(\mathbf{X}) \triangleq \frac{1}{n} \mathbf{XX}^H. \quad (45)$$

Based on [26], [27], and [28], to evaluate the converse bound for an auxiliary channel, we need to obtain the lower bound on $\epsilon'$, where $\epsilon'$ is the maximal error probability over the corresponding auxiliary channel. We thus denote the auxiliary channel $Q$ as follows:

$$Q_{Y|X,H} = \prod_{j=1}^{n} Q_{Y_{j}|X,H} \quad (46)$$

where

$$Q_{Y_{j}|X,H} = \mathcal{CN}\left(0, I_n + H p(\mathbf{X}) H^H \right). \quad (47)$$

We denote $\mathbf{B} \triangleq \mathbf{I}_n + H p(\mathbf{X}) H^H$ and let its eigenvector $\omega = [\omega_0, \ldots, \omega_m] = \lambda_{\text{max}}(\mathbf{B})$. Note that $\mathbf{P} = p(\mathbf{X})$ is the only factor that affects the output of the $Q_{Y|X,H}$ channel. Let the space $\mathbf{S} \triangleq p(\mathbf{Y}) = (1/n) \mathbf{YY}^H$ and its entry is defined as the square of the norm of $\mathbf{Y}$ and is then normalized by the blocklength $n$, which is shown below

$$S_j = \frac{\omega_j}{n} \sum_{i=1}^{n} |Z_{j,i}|^2, \quad j = 1, \ldots, m \quad (48)$$

where $Z_{j,i} \sim \mathcal{CN}(0,1)$. $\mathbf{S}$ can be seen as the statistical expression of the receiver’s detection of $\mathbf{X}$ from $\mathbf{Y}$ and $\mathbf{H}$. Thus the auxiliary channel $Q_{Y|X,H}$ can be seen as $Q_{S_{i|B}}$. From (48), we note that the $S_j$ follows the Gamma distribution, and its corresponding PDF is given by

$$q_{S_{j|B}}(s_j|\omega_j) = \frac{n^n}{(\omega_j)^n \Gamma(n)} s_j^{n-1} \exp\left\{ - \frac{n s_j}{\omega_j} \right\}. \quad (49)$$

Moreover, as $Q_{S_{i|P-P}}$ is a product of $m$ copies of the PDF of $S_j$ which is the Gamma distribution. We can obtain the PDF of $Q_{S_{i|P-P}}$ by the theorem shown below [24].

Theorem 4: Given $N$ independent gamma variables $x_i$ with the same shape parameter $k$ and the same scale parameter $\theta$ having density functions

$$f_i(x_i) = \frac{1}{\Gamma(k)\theta^k} x_i^{k-1} e^{-x_i/\theta}. \quad (50)$$

The PDF $g(z)$ of the product $z = x_1 x_2 \ldots x_N$ of $N$ independent gamma variables is a Meijer G-function multiplied by a normalizing constant $\mathcal{K}$

$$g(z) = K G_{0,N}^{N,0} \left( k-1 \left| \begin{array}{c} z \\ \theta \end{array} \right. \right) \quad (51)$$

where

$$\mathcal{K} = \left( \frac{1}{\theta} \right)^N \prod_{i=1}^{N} \frac{1}{\Gamma(k)} \quad (52)$$

and

$$G_{m,n}^{p,q} \left( j_1, j_2, \ldots, j_p \left| k_1, k_2, \ldots, k_q \right| \right) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} z^{-s} \prod_{j=1}^{m} \Gamma(s + j_1) \cdot \prod_{j=1}^{n} \Gamma(1 - j_2 - s) \cdot ds \quad (53)$$

where $c$ is a real constant defining a Bromwich path separating the poles of $\Gamma(s + j_1)$ from those of $\Gamma(1 - j_2 - s)$.

In our case, due to the shape parameter $k = n$ and the scale parameter $\theta = (1 + g_j p_j(\mathbf{X}))/2n$ and the number of copies $N = m$, applying Theorem 4, we can derive the PDF of $q_{S_{i|P}}$

$$q_{S_{i|P}}(z) = K G_{0,m}^{m,0} \left( n-1 \left| z \left( \frac{2n}{1 + g_j p_j(\mathbf{X})} \right) \right. \right) \quad (54)$$

where

$$\mathcal{K} = \left( \frac{2n}{1 + g_j p_j(\mathbf{X})} \right)^m \prod_{i=1}^{m} \frac{1}{\Gamma(n)} \quad (55)$$

and

$$G_{0,m}^{m,0} \left( n-1 \left| z \left( \frac{2n}{1 + g_j p_j(\mathbf{X})} \right) \right. \right) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \left( \frac{2n}{1 + g_j p_j(\mathbf{X})} \right)^{-s} \prod_{j=1}^{m} \Gamma(s + n - 1) ds. \quad (56)$$

Consider an arbitrary code for the auxiliary channel $Q$. Further we define $D_i, i = 1, \ldots, M$ be the decoding sets corresponding to the $M$ codewords. Let $\epsilon'_{\text{source}}$ be the maximal error probability over the auxiliary channel $Q$. Thus, we have

$$1 - \epsilon'_{\text{source}} = \frac{1}{M} \mathbb{E}_{\mathbf{R}} \left[ \sum_{i=0}^{M} D_i q_{S_{i|P}}(z) \right] \quad (57)$$

$$\leq \mathbb{E}_{\mathbf{R}} \left[ \sum_{i=0}^{M} q_{S_{i|P}}(z) \right] \quad (58)$$

$$\leq \mathbb{E}_{\mathbf{R}} \left[ \max \{q_{S_{i|P}}(z)\} \times \text{Leb}(D_0) \right]. \quad (59)$$

Next, we need to calculate the upper bounded Lebesgue of $D_0$. Because $p(\mathbf{X})$ is the power allocation vector, $\mathbf{p}$ belongs to a certain ball in $\mathbb{R}^m$. Due to the definition of $\mathbf{S}$, it belongs to a slightly larger ball. By definition of the Lebesgue measure [25], we can obtain the upper bounded Lebesgue measure of $D_0$

$$\text{Leb}(D_0) \leq \text{Leb}(\mathbf{S}) \leq \frac{K}{M} \quad (60)$$

where Leb is Lebesgue measure and $K$ is a constant.

Then for each codeword, the decoding set must have a Lebesgue measure smaller than $(K/M)$. Therefore, we have

$$1 - \epsilon'_{\text{source}} \leq \mathbb{E}_{\mathbf{R}} \left[ \max \{q_{S_{i|P}}(z)\} \times \frac{K}{M} \right] \leq \frac{n^{m/2}}{M}. \quad (61)$$
According to the binary hypothesis testing in [17], we have
\[
\Lambda(\epsilon_{\text{source}}) \geq \frac{1}{\kappa} \left( \epsilon_{\text{source}} - \mathbb{P}[I(X^n; Y^n) \leq \log \lambda] \right) \quad (62)
\]
\[
\geq \frac{1}{\kappa} \left( \epsilon_{\text{source}} - \frac{6T(X; Y)D}{\sqrt{n}U(X; Y)D} - Q(\tau) \right) \quad (63)
\]
where \( \Lambda(\epsilon_{\text{source}}) \) denotes the maximal probability of error under \( P_{Y|X,H} \) if the probability of error under \( Q_{Y|X,H} \) is \( \epsilon_{\text{source}} \) and (63) follows from (37). Then
\[
\log \Lambda(\epsilon_{\text{source}}) \geq -nI(X; Y|D) + \log \left( \frac{6T(X; Y)D}{\sqrt{n}U(X; Y)D} \right) - Q(\tau) \quad (64)
\]
where (64) follows from (36). We assume \( \tau = Q^{-1}(\epsilon_{\text{source}}(1 + [1/\sqrt{n}]) - [6T(X; Y)D]/[\sqrt{n}U(X; Y)D]) \). Thus
\[
\log \Lambda(\epsilon_{\text{source}}) \geq -nI(X; Y|D) + \log \left( \frac{6T(X; Y)D}{\sqrt{n}U(X; Y)D} \right) - \frac{1}{2} \log n. \quad (65)
\]
Due to the fact that \( \log \Lambda(\epsilon_{\text{source}}) \leq \log(1 - \epsilon'_{\text{source}}) \), we have
\[
-nI(X; Y|D) + \log \left( \frac{6T(X; Y)D}{\sqrt{n}U(X; Y)D} \right) \leq \log(1 - \epsilon'_{\text{source}}). \quad (66)
\]
Thus substituting (66) into (61), we have
\[
R \leq I(X; Y|D) - \sqrt{\frac{U(X; Y|D)}{n}} Q^{-1}(\epsilon_{\text{source}}) + \frac{(m+1) \log n}{2n} + O\left(\frac{n^{-\frac{1}{2}}}{\sqrt{n}}\right). \quad (67)
\]
This completes the proof.

**APPENDIX B**

**SADDLE POINT APPROXIMATION**

In this section, we give the proof of saddle point approximation of (10) and (11)
\[
I(X; Y|d = 0) = r \log |\mathcal{A}| - \frac{1}{|\mathcal{A}|} \sum_{i=1}^{\lvert \mathcal{A} \rvert} \int_{0}^{\infty} \int_{0}^{\infty} \frac{1}{\det(\pi L)} p(h_0) \exp\left\{-\frac{1}{2} \|y - h_0x_i\|^2\right\} \exp\left\{-\frac{1}{2} \|y - h_0x_i\|^2\right\} dydh_0 \quad (68)
\]
where (69) comes from the Taylor series expansion of (68) and
\[
\Pi(y, h_0) = \sum_{i=1}^{\lvert \mathcal{A} \rvert} \exp \left\{-\frac{1}{2} \|y - h_0x_i\|^2 - \frac{(q + 1) \|h_0(x_i - x_i)\|^2}{2q}\right\}. \quad (70)
\]
Before utilizing the saddle point approximation, we need to guarantee the existence of the saddle point. For convenience of notation, we use vector \( c_{i,r} \) to represent \( h_0(x_i - x_r) \) for any \( q \) positive integers, it easy for us to validate that \( \Pi^{-q}(y) = 0 \) and \( \lim_{y \to \infty} \Pi^{-q}(y) = 0 \). Thus there exists a maximum value of \( \Pi^{-q}(y) \), which satisfies the condition of the saddle point approximation. Then we can assume that \( \Pi^{-q}(y) \) achieves its maximum at \( y = y_0 \), which \( y_0 \) satisfies \( (\partial / \partial y) \Pi^{-q}(y) = y_0 = 0 \).

After solving (70), we have \( y_0 = \sum_{i=1}^{\lvert \mathcal{A} \rvert} q \rho_{i,r} c_{i,r} \), where \( \rho_{i,r} = \Pi(y_0) / \sum_{i=1}^{\lvert \mathcal{A} \rvert} \Pi(y_0) \) is a positive number from (0, 1) and satisfies that \( \sum_{i=1}^{\lvert \mathcal{A} \rvert} \rho_{i,r} = 1 \). Therefore, we have for a nonzero number \( q \), the multiple integrals over the complex number vector \( y \) can be approximated by the saddle point approximation
\[
\int_{-\infty}^{\infty} \frac{1}{\det(\pi L)} (\Pi(y, h_0))^{-q} dy \approx \sum_{i=1}^{\lvert \mathcal{A} \rvert} \exp \left\{-\frac{1}{2} \|h_0(x_i - x_i)\|^2 - \frac{(q + 1) \|c_{i,r}\|^2}{2q}\right\}. \quad (71)
\]
Combining (69) and (71), we eliminate the multiple integrals over the complex vector \( y \) as follows:
\[
I(X; Y|d = 0) \approx r \log |\mathcal{A}| - \frac{1}{|\mathcal{A}|} \sum_{i=1}^{\lvert \mathcal{A} \rvert} \int_{0}^{\infty} \int_{0}^{\infty} p(h_0) \exp\left\{-\frac{1}{2} \|h_0(x_i - x_i)\|^2 - \frac{(q + 1) \|c_{i,r}\|^2}{2q}\right\} dydh_0 \quad (72)
\]
Then by observing (72), we take advantage of inverse Taylor series expansion and we obtain
\[
I(X; Y|d = 0) \approx r \log |\mathcal{A}| - \int_{0}^{\infty} p(h_0) \frac{1}{|\mathcal{A}|} \sum_{i=1}^{\lvert \mathcal{A} \rvert} \exp\left\{-\frac{1}{2} \|h_0(x_i - x_i)\|^2 - \frac{(q + 1) \|c_{i,r}\|^2}{2q}\right\} dh_0. \quad (73)
\]
Moreover, we need to approximate the unconditional variance \( \mathbb{E}(X; Y|d = 0) \). The steps are basically same with the process of the approximation of \( I(X; Y) \), then we have
\[
\mathbb{E}(X; Y|d = 0) \approx \frac{1}{|A|} \sum_{i=1}^{|A|} \exp \left[ -\frac{1}{2} \left( \begin{bmatrix} x_t \mid a_i \end{bmatrix} - \begin{bmatrix} x_h \mid a_i \end{bmatrix} \right)^2 \right] \right] dh_0. 
\]

\[ (74) \]

REFERENCES

[1] L. Yan, Y. Zhang, L. T. Yang, and H. Ning, *The Internet of Things: From RFID to the Next-Generation Pervasive Networked Systems*. Boca Raton, FL, USA: Auerbach Publ., 2008.

[2] H. Jayakumar, K. Lee, W. S. Lee, A. Raha, Y. Kim, and V. Raghunathan, “Powering the Internet of Things,” in *Proc. Int. Symp. Low Power Electron. Des.* La Jolla, CA, USA, Aug. 2014, pp. 375–380.

[11] Q. Zhang, L. Zhang, Y.-C. Liang, and P. Y. Kam, “Backscatter-NOMA: Wireless communication out of thin air,” in *Proc. ACM SIGCOMM*, Hong Kong, Aug. 2013, pp. 39–50.

[12] S. Verdú and T. S. Han, “A general formula for channel capacity,” *IEEE Trans. Inf. Theory*, vol. 40, no. 4, pp. 1147–1157, Jul. 1994.

[13] G. J. Foschini and M. J. Gans, “On limits of wireless communications in a fading environment when using multiple antennas,” *Wireless Pers. Commun.*, vol. 6, no. 3, pp. 311–335, 1998.

[14] J. E. Telatar, “Capacity of multi-antenna Gaussian channels,” *Eur. Trans. Telecommun.*, vol. 10, pp. 585–595, Nov./Dec. 1999.

[15] C. Boyer and S. Roy, “Backscatter communication and RFID: Coding, energy, and MIMO analysis,” *IEEE Trans. Commun.*, vol. 62, no. 3, pp. 770–785, Mar. 2014.

[16] G. Yang, Y.-C. Liang, R. Zhang, and Y. Pei, “Modulation in the air: Backscatter communication over ambient OFDM carrier,” *IEEE Trans. Commun.*, vol. 66, no. 3, pp. 1219–1233, Mar. 2018.

[17] J. Qian, Y. Zhu, C. He, F. Gao, and S. Jin, “Achievable rate and capacity analysis for ambient backscatter communications,” *IEEE Trans. Commun.*, vol. 67, no. 9, pp. 6299–6310, Sep. 2019.

[18] R. Long, G. Yang, Y. Pei, and R. Zhang, “Transmit beamforming for cooperative ambient backscatter communication systems,” in *Proc. IEEE Global Commun. Conf.*, 2017, pp. 1–6.

[19] E. Biglieri, J. Proakis, and S. Shamai, “Fading channels: Information-theoretic and communications aspects,” *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2619–2692, Oct. 1998.

[20] D. N. C. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[21] G. J. Foschini and M. J. Gans, “On limits of wireless communications in a fading environment when using multiple antennas,” *Wireless Pers. Commun.*, vol. 6, no. 3, pp. 311–335, 1998.

[22] I. E. Telatar, “Capacity of multi-antenna Gaussian channels,” *Eur. Trans. Telecommun.*, vol. 10, pp. 585–595, Nov./Dec. 1999.

[23] M. D. Springer and W. E. Thompson, “The distribution of products of beta, gamma and Gaussian random variables,” *SIAM J. Appl. Math.*, vol. 18, no. 4, pp. 721–737, 1970.

[24] R. Ash, *Information Theory*, New York, NY, USA: Intersci. Publ., 1965.

[25] J. Neyman and E. S. Pearson, “On the problem of the most efficient tests of statistical hypotheses,” *Philos. Trans. Roy. Soc. A, Math. Phys. Eng. Sci.*, vol. 231, pp. 289–337, Feb. 1933.

[26] H. V. Poor and S. Verdú, “A lower bound on the probability of error in multihypothesis testing,” *IEEE Trans. Inf. Theory*, vol. 41, no. 6, pp. 1992–1994, Nov. 1995.

[27] R. E. Blahut, “Hypothesis testing and information theory,” *IEEE Trans. Inf. Theory*, vol. 20, no. 4, pp. 405–417, Jul. 1974.
H. Vincent Poor (Life Fellow, IEEE) received the Ph.D. degree in EECS from Princeton University, Princeton, NJ, USA, in 1977.

From 1977 to 1990, he was on the faculty of the University of Illinois at Urbana-Champaign, Champaign, IL, USA. Since 1990, he has been on the faculty at Princeton, where he is currently the Michael Henry Strater University Professor. From 2006 to 2016, he served as the Dean of Princeton’s School of Engineering and Applied Science. He has also held visiting appointments at several other universities, including most recently at Berkeley and Cambridge. His research interests are in the areas of information theory, machine learning and network science and their applications in wireless networks, and energy systems and related fields. Among his publications in these areas is the recent book *Machine Learning and Wireless Communications* (Cambridge University Press, 2022).

Dr. Poor received the IEEE Alexander Graham Bell Medal in 2017. He is a member of the National Academy of Engineering and the National Academy of Sciences and is a Foreign Member of the Chinese Academy of Sciences, the Royal Society, and other national and international academies.

Branka Vucetic (Life Fellow, IEEE) received the Ph.D. degree from the University of Belgrade, Belgrade, Serbia, in 1982.

She is an Australian Laureate Fellow, a Professor of Telecommunications, and the Director of the Centre for IoT and Telecommunications with the University of Sydney, Sydney, NSW, Australia. Her current research work is in wireless networks, and Industry 5.0. In the area of wireless networks, she works on communication system design for 6G and wireless AI. In the area of Industry 5.0, Vucetic’s research is focused on the design of cyber–physical–human systems and wireless networks for applications in healthcare, energy grids, and advanced manufacturing.

Dr. Vucetic is a Fellow of the Australian Academy of Technological Sciences and Engineering and the Australian Academy of Science.