RIS-NOMA-Assisted Short-Packet Communication with Hardware Impairments
Chunli Xia, Zhongwu Xiang, Binbin Su, Hongbo Liu, Jin Meng and Gaofeng Pan, Senior Member, IEEE

Abstract—To address the stringent demands of next-generation networks for massive connectivity, ultra-low latency, and ultra-high spectral efficiency, the win-win integration of reconfigurable intelligent surface (RIS) and nonorthogonal multiple access (NOMA) is considered as a promising solution. In this paper, we investigate a downlink RIS-empowered NOMA system with short packet communications (SPC) in the presence of hardware impairments at the transceiver nodes. To characterize the performance of the proposed network, the approximate and asymptotic closed-form expressions of average block error rate (BLER) at far and near users are derived and analyzed. Based on asymptotic average BLER in high signal-to-noise ratio regime, the diversity order, minimum blocklength, and optimal power allocation are examined. RIS exhibits an improvement in diversity order and minimum blocklength. The achieved results show that users with poorer channel conditions have less sensitivity to hardware impairments whether for near users or RIS-assisted far users. Furthermore, our results manifest the benefits of NOMA and RIS in SPC over the benchmark orthogonal multiple access (OMA) scheme.

Index Terms—BLER, minimum blocklength, hardware impairments, non-orthogonal multiple access, short-packet communications, RIS.

I. INTRODUCTION

INTERNET of Things (IoT) has been regarded as a key technology for next generation wireless communication systems to establish ubiquitous connectivity for massive sensor nodes and IoT devices over the internet [1], [2]. Beyond the fifth generation (5G) and the sixth generation (6G) are expected to support a connection density 100 times higher than that of the fifth generation (5G) [3]. Furthermore, in various emerging key application scenarios in IoT, such as industrial automation [4] and autonomous driving [5], higher requirements are put forward for ultra-low delay and ultra-high reliability, requiring the delay to be kept within one millisecond and the reliability to reach 99.99% [6]. To this end, faced with more demanding and multi-user communication requirements, more intelligent technologies that can effectively improve transmission reliability and support large-scale connectivity are urgently needed.

To address the above challenges, reconfigurable intelligent surface (RIS) and non-orthogonal multiple access (NOMA) are considered as two appealing communication technologies to enhance the quality of service (QoS) communication for multiple users. On the one hand, RIS is a metasurface composed of multiple passive reflective elements and managed by a RIS controller [7]–[9], which can intelligently customize and program the wireless propagation environment by adjusting the phase of reflective elements, and thus effectively improving the achievable data rate and reliability for users [10]–[12]. In particular, RIS is a cost- and energy-efficient smart radio environment technology that does not include power-consuming components and provides performance improvement with extremely low power consumption [13]–[15]. On the other hand, NOMA employs power multiplexing and successive interference cancellation (SIC) to significantly improve spectral efficiency and user connectivity to meet different target requirements [16], [17]. Unlike conventional orthogonal multiple access (OMA) technologies, NOMA systems can serve multiple users in the same resource block, with the potential to improve user fairness and reduce transmission delays [18].

The intrinsic integration of NOMA and RIS can realize their potential benefits, as RIS can ensure a controllable and intelligently optimized wireless propagation environment for better implementation of NOMA, while NOMA can improve the spectrum efficiency and connectivity of RIS networks. Inspired by the advantages of both RIS and NOMA, the authors of [19] examined the effectiveness of the RIS in NOMA system with respect to the transmit power consumption by designing a joint optimization algorithm. To maximize the system average sum rate for delay-tolerate transmissions, joint phase shifters policy and resource assignment optimization problems in downlink RIS-aided NOMA and OMA systems were considered in [7], the results revealed that the RIS-aided NOMA system outperforms the conventional NOMA network over fading channels. In [20], RIS-enhanced multi-unmanned aerial vehicle (UAV) NOMA networks were investigated and they proved RIS-NOMA scheme achieves a higher sum rate in comparison with the benchmark schemes (i.e., RIS-OMA and NOMA without RIS) by block coordinate descent (BCD)-based iterative optimization algorithm. Besides, the performance of an RIS-assisted NOMA networks with imperfect
and perfect SIC was investigated in [21] and revealed that RIS-NOMA outperforms RIS-OMA scheme in terms of both outage behavior and ergodic rate.

The aforementioned research contributions have laid a solid foundation on the application of RIS-NOMA networks and its superiority has been widely proved [22]–[24]. However, most of literature focused on long packet services. On the one hand, relying on NOMA technology solely is not enough to meet the ultra-low latency of the IoT. On the other hand, multiple application scenarios in IoT require only a small amount of information to be exchanged among nodes. Therefore, short packet communication (SPC) is gradually becoming a feature of the IoT. In this case, the Shannon capacity theory no longer holds, and the reliability of transmission cannot be fully guaranteed at arbitrarily high signal-to-noise ratios (SNR). To fill the theoretical gap in SPC, authors in [25] formulated the relationship between achievable rate, decoding error probability and transmission delay, and the average block error rate (BLER) was defined as a key metric to evaluate the performance of SPC systems. Impelled by the discussion in [25], authors in [26] devoted to investigate the latency performance of RIS-aided downlink ultra-reliable and low-latency communication (URLLC) with user grouping in time division multiple access protocol. And the problem of ultra-high reliability in URLLC assisted by a mobile UAV and RIS in a SPC scenario was formulated and solved with Nelder-Mead simplex method in [27]. Their presented result revealed the UAV’s position is crucial to achieve ultra-high reliability for short packets. In [6], the performance of RIS-aided short-packet NOMA systems under perfect and imperfect SIC was investigated, and proved that the proposed system outperforms its OMA counterpart in terms of BLER and throughput. Out of the above-mentioned works, the research on RIS-NOMA networks in SPC is still scarce. Furthermore, numerous works on wireless communication systems are based on the assumption of perfect transceiver devices, in fact this assumption is unrealistic, and it is inevitable that radio-frequency transceiver hardware suffers losses due to phase noise, amplifier nonlinearity, etc., especially in IoT communication systems where low-cost hardware devices are commonly used [2], [28]–[30].

Impelled by the above facts, this paper investigates an IoT network consisting far and near users. Since the direct link channel conditions between far users and base station (BS) are not ideal, RIS is used to enhance the performance of far users. Far and near users can communicate simultaneously under the NOMA scheme. Besides, this paper considers a practical scenario where hardware impairments of the BS and the receiver nodes occur. The major contributions of this paper are as follows.

- In light of the demand for massive connectivity and strict communication performance for IoT, an RIS-empowered NOMA downlink communication system is investigated in this paper. In addition to serving users nearby the BS, the deployment of RIS improve access to services for far users at the edge. By assigning different power to users, far and near users can obtain reliable communication simultaneously via NOMA principle.
- To evaluate the system performance, the reliability performance for far and near users is analyzed by deriving closed-form expressions of average BLER in the presence of hardware impairments. To gain more insights, the asymptotic expressions for the average BLER are derived as the transmit power tends to infinity, the diversity order is also analyzed and the minimum block length and the corresponding optimal power allocation are investigated. In addition, the minimum blocklength of the OMA scheme is demonstrated as a benchmark. The analysis results show that increasing the element number of reflective surfaces enhances the diversity order and minimum blocklength.
- The numerical simulation results verify the correctness of the theoretical analysis, and simulation results validate the advantages of NOMA and RIS in SPC, and the proposed system outperforms the OMA benchmark scheme in terms of low latency transmission. It’s noteworthy to note that user with worse channel conditions is less susceptible to hardware impairments whether for near users or RIS-assisted far users.

Table I presents the comparisons between contributions of this work and the existing related literature.

The rest of the paper is organized as follows. The network model and transmission scheme are introduced in Section II. The closed-form expressions for the average BLER of far and near users are derived in Section III. Next, the asymptotic average BLER, diversity order, optimal power allocation, and minimum blocklength are expressed in Section IV. In Section V, numerical results are presented to verify the derived analysis models and some insights into system characteristics are presented. Finally, concluding remarks are provided in Section VI.

Notation: $\mathbb{C}^{M \times N}$ is an $M \times N$ space of complex matrices. $E[ \cdot ]$ denotes expectation operation. The superscript $\cdot^H$ stands for the conjugate-transpose operation. $\text{diag}(\cdot)$ represents a diagonal matrix. $\text{arg}(\cdot)$ denotes the phase of the complex-valued vector. $f_\mathbf{X}(\cdot)$ and $F_\mathbf{X}(\cdot)$ denote the probability density function (PDF) and cumulative distribution function (CDF) of a random variable $X$.

II. Network Model and Transmission Scheme

A. Network Model

As depicted in Fig. 1, we consider an RIS-empowered multi-user IoT NOMA network with short packets, which consists of a BS, an RIS, $Q$ near users $S_q \ (q \in \{1, \cdots, j, \cdots, k, \cdots, Q\})$ and $K$ far users $U_k \ (k \in \{1, \cdots, m, \cdots, n, \cdots, K\})$. Assuming that the direct communication links from BS to far users are not available due to severe path loss, deploying RIS can assist the signal transmission of far users, while the near users can leverage the direct links to communicate without taking up RIS resources. The presented network model can be applied to multiple IoT applications. To be specific, far users can be cell-edge user in cellular networks or user in dead zone, near users can be cell-central users that have favorable communication links with BS. The RIS consists of $N$ reconfigurable reflecting elements that can be divided into $H$ sub-surfaces, each of which has a different phase shift and serves a specific far
user [31], [32]. The satisfaction of $H \geq K$ is inevitable and obvious, it is assumed that $H = K$ for efficiency maximization of the RIS in this system. In addition, the RIS is managed by a smart controller that enables dynamic adjustment on the reflections of RIS elements. The BS and users are equipped with a single antenna and all nodes work under half-duplex mode, the additive white Gaussian noise (AWGN) powers at receivers are $\sigma^2$.

To characterize the optimal performance, it is assumed that the channel state information (CSI) of all channels is available at BS since it can be efficiently obtained by the recent research in channel estimation for RIS [33]–[35]. The channel between the BS and the $k$th sub-surface of the RIS is denoted as $h_{srk} = [h_{srk}^1, h_{srk}^2, \ldots, h_{srk}^H]^H \in \mathbb{C}^{E_k \times 1}$ and $E_k$ is the adjacent elements number of $k$th sub-surface. $h_{ru_k} = [h_{ru_k}^1, h_{ru_k}^2, \ldots, h_{ru_k}^H]^H \in \mathbb{C}^{E_k \times 1}$ represents the channel coefficient from $k$th sub-surface of RIS to $U_k$, and $h_{srk}$ denotes the channel coefficient from the BS to $S_k$. Furthermore, the channels are assumed to undergo Nakagami-$m$ fading, i.e., $|h_{srk}^i|^2 \sim \text{Nakagami}(m_{sr}, \Omega_{sr})$, $|h_{ru_k}^i|^2 \sim \text{Nakagami}(m_{ru}, \Omega_{ru})$, and $|h_{srk}^i|^2 \sim \text{Nakagami}(m_{sr}, \Omega_{sr})$, where $i \in \{1, 2, \ldots, E_k\}$, $m_{sr}$, $m_{ru}$ and $m_{sr}$ are the corresponding fading parameters, and corresponding fading powers are $\Omega_{sr}$, $\Omega_{ru}$ and $\Omega_{sr}$.

### B. NOMA Scheme

Under NOMA scheme, BS employs superposition coding to send signal $x_s$ to users. To facilitate analysis and maintain the advantages of NOMA as proposed in [36]–[38], the number of both far and near users served is assumed as two, and $x_s$ can be expressed as

$$x_s = \sqrt{a_u a_m P} x_{um} + \sqrt{a_u a_n P} x_{un} + \sqrt{b_j b_k P} x_{sk},$$

where $P$ is the transmit power at the BS, $x_{um}$ and $x_{un}$ are the normalized messages for $U_m$ and $U_n$ respectively, $x_{sk}$ stands for the normalized messages for $S_j$ and $S_k$ respectively. Besides, $a_u$ and $b_j$ represent power allocation factors of far and near users, respectively. It should be noted that under the NOMA principle, power is allocated based on channel gains, and users with lower channel gains are allocated more power. Hence, more power should be allocated to far users, i.e., $a_u > b_j$. Besides, the channel gain is primarily determined by the fading mean value, so this value can be utilized as a guide for efficient power allocation. Without loss of generality, the fading means from the BS to far users are ordered as $\Omega_{ru_m} < \Omega_{ru_n}$. Assume that the distance from the BS to the RIS is much larger than the size of the RIS, the channel difference from the BS to different sub-surfaces of the RIS is ignored, thus $\Omega_{sr} > \Omega_{sr_m}$, $\Omega_{sr} > \Omega_{sr_n}$, $\Omega_{sr} > \Omega_{sr_ij}$, $\Omega_{sr} > \Omega_{sr_ij}$, $\Omega_{sr} > \Omega_{sr_ij}$, $\Omega_{sr} > \Omega_{sr_ij}$, $\Omega_{sr} > \Omega_{sr_ij}$, respectively. Moreover, the fading means from the BS to near users are sorted as $\Omega_{sr} < \Omega_{sr_m}$, $\Omega_{sr} < \Omega_{sr_n}$, thus $b_j > b_k$, where $b_j$ and $b_k$ are the power allocation factors of $S_j$ and $S_k$, respectively.

Accordingly, the signal received at user $U_m$ and $U_n$ are respectively given by

$$y_{um} = h_{ru_m}^H \Theta h_{sr_m} (x_s + n_s) + n_{rm} + n_{um},$$

$$y_{un} = h_{ru_n}^H \Theta h_{sr_n} (x_s + n_s) + n_{rn} + n_{un},$$

where $\Theta$ is a diagonal phase shifting matrix, $n_s \sim \mathcal{CN}(0, \theta_s^2 P)$ represents the distortion noise for the BS, $n_{rm} \sim \mathcal{CN}(0, \theta_r m^2 P |h_{ru_m}^H \Theta h_{sr_m}|^2)$ and $n_{rn} \sim \mathcal{CN}(0, \theta_r n^2 P |h_{ru_n}^H \Theta h_{sr_n}|^2)$ respectively denote the distortion noise for node $U_m$ and $U_n$, with $\theta_s$, $\theta_r$, $m$ and $n$ being the level of impairments, $n_{um}$ and $n_{un}$ are the AWGN variables.
The signal received at user $S_j$ and $S_k$ are given by
\[ y_{ss_j} = h_{ss_j}(x_s+n_s) + n_j + n_{ss_j}, \]
\[ y_{ss_k} = h_{ss_k}(x_s+n_s) + n_k + n_{ss_k}, \]
respectively, where $n_j \sim \mathcal{C}\mathcal{N}(0,\theta_j^2 P|h_{ss_j}|^2)$ and $n_k \sim \mathcal{C}\mathcal{N}(0,\theta_k^2 P|h_{ss_k}|^2)$ respectively denote the distortion noise for user $S_j$ and $S_k$, with $\theta_j$ and $\theta_k$ being the level of impairments, $n_{ss_j}$ and $n_{ss_k}$ are the AWGN.

On the basis of NOMA principle, $U_m$ can directly decode its own message, the received signal-to-interference-plus-noise ratio (SINR) at $U_m$ can be given as
\[ \gamma_{u_m \rightarrow u_m} = \frac{\left( \sum_{e=1}^{E_m} |h^e_{r,u_m}|^2 |h^e_{s,r_m}|^2 \right) a_{u_m} \rho}{\left( \sum_{e=1}^{E_m} |h^e_{r,u_m}|^2 |h^e_{s,r_m}|^2 \right) (a_{u_m} \rho + b_{s} \rho + \rho \theta_m^2) + 1}, \]
where $\rho = \frac{P}{2}$, $\theta_m^2 = \theta_s^2 + \theta_r^2$. It is noteworthy that (6) holds only in the case of perfect CSI and the optimal phase shift of reflection elements in $n$th sub-surface $\phi_e = -\arg(h^e_{r,u_m}) - \arg(h^e_{s,r_m}) \in \{1,2,\cdots,E_m\}$ can be obtained [39].

As for $U_n$, the phase shift of reflection elements in $n$th sub-surface is $\phi'_e = -\arg(h^e_{r,u_n}) - \arg(h^e_{s,r_n}) \in \{1,2,\cdots,E_n\}$. Following NOMA principle, $U_n$ first decodes the signal of $U_m$ and removes it. After the successful implementation of SIC, $U_n$ then decodes its own signal, the SINR can be expressed as
\[ \gamma_{u_n \rightarrow u_n} = \frac{\left( \sum_{e=1}^{E_n} |h^e_{r,u_n}|^2 |h^e_{s,r_n}|^2 \right) a_{u_n} \rho}{\left( \sum_{e=1}^{E_n} |h^e_{r,u_n}|^2 |h^e_{s,r_n}|^2 \right) (b_{s} \rho + \rho \theta_n^2) + 1}, \]
where $\theta_n^2 = \theta_s^2 + \theta_r^2$.

Near users should first eliminate the interference from the signals intended to far users. The SINR of message $x_{sj}$ obtained at user $S_j$ is written as
\[ \gamma_{s_j \rightarrow s_j} = \frac{\left|h_{ss_j}\right|^2 b_s b_j \rho}{\left|h_{ss_j}\right|^2 (b_s b_j \rho + \rho \theta_j^2) + 1}, \]
where $\theta_j^2 = \theta_s^2 + \theta_r^2$.

Analogously, after eliminating the interference from the signals intended to user $S_k$, the SINR for $S_k$ to demodulate its own signal is given by
\[ \gamma_{s_k \rightarrow s_k} = \frac{\left|h_{ss_k}\right|^2 b_s b_k \rho}{\left|h_{ss_k}\right|^2 \rho \theta_k^2 + 1}, \]
where $\theta_k^2 = \theta_s^2 + \theta_r^2$. To facilitate the analysis, assume that all users have the same level of impairments, then $\theta_m^2 = \theta_n^2 = \theta_j^2 = \theta_k^2 = \theta^2$.

III. ANALYSIS OF BLOCK ERROR RATE IN SHORT PACKET COMMUNICATION

In this section, we first present some preliminaries on SPC. Then, we derive closed-form expressions of the average BLER for far and near users.

A. Preliminaries

SPC is becoming a favorable and inevitable trend in IoT. But Shannon theory built on the assumption of infinite blocklength is no longer applicable. In this case, Polyanskiy et al. in [25] pioneered the derivation of the maximal achievable rate for blocklength of $L$, the SINR $\gamma$, and the BLER $\epsilon$ as [40]
\[ R = \log_2(1 + \gamma) - \frac{\sqrt{V(\gamma)} Q^{-1}(\epsilon)}{L \ln 2}, \]
where $V(x) = 1 - (1 + x)^{-2}$, $Q^{-1}(x)$ is the inverse of the Gaussian Q-function $Q(x) = \int_{x}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-t^2} dt$. From (10), we can compute the instantaneous BLER of decoding the message of user $V$ at user $U (U,V \in \{U_m, U_n, S_j, S_k\})$ as follows:
\[ \epsilon_{U \rightarrow V} = Q(\ln 2 \log_2 (1 + \gamma_{U \rightarrow V}) - R_V), \]
where $R_V = \frac{\eta_V}{L_V}, \eta_V$ and $L_V$ denote the number of information bits and blocklength to user $V$, respectively. To facilitate subsequent analysis, (11) can be tightly approximated as
\[ \epsilon_{U \rightarrow V} = \begin{cases} 0 & \gamma_{U \rightarrow V} \leq A_V, \\ 1 & \gamma_{U \rightarrow V} > A_V. \end{cases} \]
where $g_V = \frac{1}{2g_V(L_V - 1)}$, $h_V = 2g_V - 1$, $A_V = h_V - \frac{1}{2g_V(L_V - 1)}$, and $B_V = h_V + \frac{1}{2g_V(L_V - 1)}$.

From (12), the average BLER $E[\epsilon_{U \rightarrow V}]$ is given by
\[ E[\epsilon_{U \rightarrow V}] = \int_0^\infty \epsilon_{U \rightarrow V} f_{\epsilon_{U \rightarrow V}}(x) dx = g_V \sqrt{L_V} \int_{A_V}^{B_V} F_{\epsilon_{U \rightarrow V}}(x) dx. \]

B. Average BLER Analysis of Far Users

**Lemma 1:** Suppose $X$ is a random variable (RV) with the first and second moments are $\mu_X = E[X]$ and $\mu_X(2)=E[X^2]$ respectively, then $X$ can be approximated as a gamma distribution with shape parameter $k'$ and scale parameter $\theta'$:
\[ k' = \frac{\mu_X}{\mu_X(2) - \mu_X^2}, \theta' = \frac{\mu_X(2) - \mu_X^2}{\mu_X}. \]

**Proof:** It can be referred to statistics books and relevant literature, such as [4], [41], [42].

Utilizing **Lemma 1**, the derivation of average BLER at users $U_m$ and $U_n$ is provided in the following theorems.

**Theorem 1:** The closed-form expression of the average
BLER for user $U_m$ in RIS-NOMA network is given by

$$E[\varepsilon_{u_m \rightarrow u_m}] = \frac{1}{\omega} \sum_{i=1}^{\omega} \gamma \left( k_1, \frac{\lambda_i u_m n - 1/n}{\theta_1} \right) \Gamma(k_1), \quad (15)$$

where $\omega$ is the complexity accuracy trade-off parameter, $k_1$ and $\theta_1$ are given in (35), $\lambda_1 = a_u u_m \rho + b_u \rho + \rho \theta^2$, and $x_i = (2(1-1)(B_{u_m} - A_{u_m})/2\omega$, $k_2$ and $\theta_2$ are given in (40).

Proof: See Appendix A.

**Corollary 1:** The diversity order of user $U_m$ and $U_n$ can be obtained as

$$D_{u_m} = - \lim_{\rho \to \infty} \log E[\varepsilon_{u_m \rightarrow u_m}] = \frac{k_1}{2}. \quad (17)$$

$$D_{u_n} = - \lim_{\rho \to \infty} \log E[\varepsilon_{u_n \rightarrow u_n}] = - \lim_{\rho \to \infty} \log \left( \frac{2\rho - 2\theta^2}{\rho} \right) = \frac{k_2}{2}. \quad (18)$$

Proof: When $\rho \to \infty$, $E[\varepsilon_{u_m \rightarrow u_m}]$ and $E[\varepsilon_{u_n \rightarrow u_n}]$ can be respectively expressed as

$$E[\varepsilon_{u_m \rightarrow u_m}] \rho \to \infty = \frac{1}{\omega} \sum_{i=1}^{\omega} \left( \frac{\gamma \left( k_1, \frac{\lambda_i u_m n - 1/n}{\theta_1} \right)}{\Gamma(k_1)} \right)^{k_1} \quad (19)$$

$$E[\varepsilon_{u_n \rightarrow u_n}] \rho \to \infty = \frac{1}{\omega} \sum_{i=1}^{\omega} \left( \frac{\gamma \left( k_1, \frac{\lambda_i u_n n - 1/n}{\theta_1} \right)}{\Gamma(k_1)} \right)^{k_1} \quad (20)$$

Invoking (19) and (20), we can find $E[\varepsilon_{u_m \rightarrow u_m}]$ and $E[\varepsilon_{u_n \rightarrow u_n}]$ are proportional to $\left( \frac{1}{\rho} \right)^{k_1/2}$ and $\left( \frac{1}{\rho} \right)^{k_2/2}$, respectively, thus the diversity order is obtained as (17) and (18), the proof is completed.

**Remark 1:** Observing (15) and (16), we can find that the average BLER of $U_m$ and $U_n$ drops with the increase of power. Besides, the rise in $k_1$ and $\theta_1$ improves the reliability of user $U_m$, and likewise increasing $k_2$ and $\theta_2$ enhances the reliability of user $U_n$. This indicates that optimizing the channel conditions of the cascade channel (the channel from BS to the RIS and then to the user) is beneficial to the user communication quality. Utilizing **Corollary 1**, it reveals increasing the number of RIS elements contributes to boosting the diversity order of user.

**C. Average BLER Analysis of Far Users**

The derivation of average BLER at users $S_j$ and $S_k$ is provided in the following.

**Theorem 3:** The closed-form expression of the average BLER for user $S_j$ in RIS-NOMA network is given by

$$E[\varepsilon_{s_j}] \approx E[\varepsilon_{s_j \rightarrow u_m}] + E[\varepsilon_{s_j \rightarrow u_n}] \approx \frac{1}{\omega} \sum_{i=1}^{\omega} \gamma \left( k_1, \frac{\lambda_i s_j n - 1/n}{\theta_1} \right) + \gamma \left( k_2, \frac{\lambda_i s_j n - 1/n}{\theta_2} \right), \quad (21)$$

where $\lambda_3 = b_j \rho + \rho \theta^2$, $\lambda_3 = b_j n_s \rho + \rho \theta^2$, and $\lambda_1 = A_{s_j} + \frac{(2(1-1)(B_{s_j} - A_{s_j})}{2\omega}, k_2$ and $\theta_2$ are given in (40).

Proof: See Appendix B.

**Remark 2:** Using **Theorem 3** and **Theorem 4**, we can observe that high power facilitates the execution of SIC and the demodulation of its own destination signal. In addition, the channel parameters and power allocation coefficient are the key factors that affect the reliability of near users, which should be set reasonably.

**IV. ANALYSIS OF OPTIMAL POWER ALLOCATION AND MINIMUM BLOCKLENGTH**

Based on the average BLER analysis presented in Section III, we provide the minimum blocklength at high SINR regime and the corresponding optimal power allocation in this section. To provide a comparison between NOMA and OMA, a benchmark OMA scheme is also analyzed.

**A. Asymptotic Average BLER Analysis**

Observing the expressions of the average BLER for far and near users presented in previous section, it is mathematically intractable to obtain the expression for the minimum blocklength. To facilitate this issue, we degenerate the expressions of average BLER to first-order Riemann integral approximation as employed by [43], [44]. Therefore, (13) can be further simplified as

$$E[\varepsilon_{u_{\rightarrow V}}] = g_{\mathbb{V}} \sqrt{B_{\mathbb{V}}} (B_{\mathbb{V}} - A_{\mathbb{V}}) F_{\gamma_{u_{\rightarrow V}}} \left( \frac{A_{\mathbb{V}} + B_{\mathbb{V}}}{2} \right). \quad (23)$$
\[
E[\varepsilon_{s_k}] \approx E[\varepsilon_{s_k} \rightarrow u_m] + E[\varepsilon_{s_k} \rightarrow u_n] + E[\varepsilon_{s_k} \rightarrow s_j] + E[\varepsilon_{s_k} \rightarrow s_k] \\
= \sum_{i=1}^{\omega} \gamma (k_i + \frac{u_m p_m^\rho}{\theta_1}) + \gamma (k_i + \frac{u_n p_n^\rho}{\theta_2}) + \gamma (k_i + \frac{p_j}{\theta_3}) + \gamma (k_i + \frac{p_j}{\theta_4})
\]

(22)

Leveraging (23), the approximate expression of the average BLER \( E[\varepsilon_{u_m} \rightarrow u_m] \) at user \( U_m \) is derived as

\[
E[\varepsilon_{u_m} \rightarrow u_m] = 1 - \sum_{r=0}^{k_1-1} \left( \frac{\lambda \Delta u_m \Sigma u_m}{\theta_1} \right)^r \frac{r}{r!} e^{-\frac{\lambda \Delta u_m \Sigma u_m}{\theta_1}} \approx \frac{k_2}{k_2!}
\]

(24)

where \( \Sigma u_m = \frac{A_{u_m} + B_{u_m}}{2} = 2R_{u_m} - 1 \). The asymptotic average BLER at user \( U_n \) is expressed as

\[
E[\varepsilon_{u_n} \rightarrow u_n] = 1 - \sum_{r=0}^{k_2-1} \left( \frac{\lambda \Delta u_n \Sigma u_n}{\theta_2} \right)^r \frac{r}{r!} e^{-\frac{\lambda \Delta u_n \Sigma u_n}{\theta_2}} \approx \frac{k_2}{k_2!}
\]

(25)

where \( \Sigma u_n = \frac{A_{u_n} + B_{u_n}}{2} = 2R_{u_n} - 1 \). Analogously, the asymptotic average BLER at user \( S_j \) is expressed as

\[
E[\varepsilon_{s_j} \rightarrow u_m] = 1 - \sum_{r=0}^{k_3-1} \left( \frac{\lambda \Delta s_j \Sigma s_j}{\theta_3} \right)^r \frac{r}{r!} e^{-\frac{\lambda \Delta s_j \Sigma s_j}{\theta_3}} \approx \frac{k_3}{k_3!}
\]

(26)

where \( \Sigma s_j = \frac{A_{s_j} + B_{s_j}}{2} = 2R_{s_j} - 1 \). The asymptotic average BLER at user \( S_k \) is given as

\[
E[\varepsilon_{s_k} \rightarrow u_m] = 1 - \sum_{r=0}^{k_4-1} \left( \frac{\lambda \Delta s_k \Sigma s_k}{\theta_4} \right)^r \frac{r}{r!} e^{-\frac{\lambda \Delta s_k \Sigma s_k}{\theta_4}} \approx \frac{k_4}{k_4!}
\]

(27)

where \( \Sigma s_k = \frac{A_{s_k} + B_{s_k}}{2} = 2R_{s_k} - 1 \).

**B. Optimal Power Allocation and Minimum Blocklength**

Based on the asymptotic average BLER, we will examine the optimal power allocation and minimum blocklength. We denote the required reliability of user \( U (\in \{U_m, U_n, S_j, S_k\}) \) in terms of BLER as \( \varepsilon_{U}^{\text{th}} \). By substituting \( \varepsilon_{U}^{\text{th}} \) into (24), (25), (26) and (27), the minimum required blocklength \( L_{U, \text{min}} \) of user \( U \) are respectively calculated as

\[
L_{u_m, \text{min}} = \log_2 \left( 1 + \frac{\eta_{u_m}}{1 + \Delta u_m \lambda_1 + 1} \right)
\]

(28)

\[
L_{u_n, \text{min}} = \log_2 \left( 1 + \frac{\eta_{u_n}}{1 + \Delta u_n \lambda_2} \right)
\]

(29)

\[
L_{s_j, \text{min}} = \log_2 \left( 1 + \frac{\eta_{s_j}}{1 + \Delta s_j} \right)
\]

(30)

and

\[
L_{s_k, \text{min}} = \log_2 \left( 1 + \frac{\eta_{s_k}}{1 + \Delta s_k} \right)
\]

(31)

where \( \Delta u_m = \Psi_{u_m} \frac{\epsilon_{U_m}^2}{\theta_2^2} \), \( \Delta u_n = \Psi_{u_n} \frac{\epsilon_{U_n}^2}{\theta_2^2} \), \( \Delta s_j = \Psi_{s_j} \frac{\epsilon_{U_j}^2}{\theta_3^2} \), \( \Delta s_k = \Psi_{s_k} \frac{\epsilon_{U_k}^2}{\theta_4^2} \), and \( \Psi_{U_m} = \varepsilon_{U_m}^\theta k_1 \), \( \Psi_{U_n} = \varepsilon_{U_n}^\theta k_2 \), \( \Psi_{S_j} = \varepsilon_{S_j}^\theta k_3 \), \( \Psi_{S_k} = \varepsilon_{S_k}^\theta k_4 \).
By calculating the derivatives of the minimum blocklength with respect to power allocation factor, we find that \(L_{um,\min}\) is a decreasing function of \(a_m\) and \(L_{us,\min}\) is an increasing function of \(a_m\). Then, \(L_{sj,\min}\) is a decreasing function of \(b_j\) and \(L_{sk,\min}\) is an increasing function of \(b_j\). Concurrently, \(L_{um,\min}\) and \(L_{us,\min}\) decrease monotonically with respect to \(a_u\), while \(L_{sj,\min}\) and \(L_{sk,\min}\) increase monotonically with respect to \(a_u\). Thus, the minimum common blocklength \(L_o^*\) is obtained by solving \(L_o^* = L_{um,\min} = L_{us,\min} = L_{sj,\min} = L_{sk,\min}\), and the corresponding optimal power allocation coefficient can be attained by Algorithm 1.

**Algorithm 1:** Algorithm for obtaining optimal power allocation coefficients

**Input:** \(\eta_{um}, \eta_{us}, \eta_{sj}, \eta_{sk}, \xi_{um}, \xi_{us}, \xi_{sj}, \xi_{sk}, \rho_s\), tolerance \(\delta\)

**Output:** Optimal power allocation coefficient \(a_{uo}\), \(a_{mo}\), \(b_{jo}\) and minimum blocklength \(L_o^*\)

1. Initialization: \(a_{uo}^+ \leftarrow 0.5, a_{uo}^- \leftarrow 1, a_{uo} = \frac{a_{uo}^++a_{uo}^-}{2}\), \(a_{mo}^+ \leftarrow 0.5, a_{mo}^- \leftarrow 1, a_{mo} = \frac{a_{mo}^++a_{mo}^-}{2}\), \(b_{jo}^+ \leftarrow 0.5, b_{jo}^- \leftarrow 1, b_{jo} = \frac{b_{jo}^++b_{jo}^-}{2}\)
2. While \(f_1(a_{uo}) = L_{um,\min} - L_{sj,\min} > \delta\)
   3. While \(f_2(a_{mo}) = L_{um,\min} - L_{us,\min} > \delta\)
      4. If \(f_2(a_{mo}) > 0\)
         5. Set \(a_{mo}^+ \leftarrow a_{mo}^-\)
      6. Else
         7. Set \(a_{mo}^- \leftarrow a_{mo}^+\)
   8. End
9. While \(f_3(b_{jo}) = L_{sj,\min} - L_{sk,\min} > \delta\)
   10. If \(f_3(b_{jo}) > 0\)
      11. Set \(b_{jo}^+ \leftarrow b_{jo}^-\)
   12. Else
      13. Set \(b_{jo}^- \leftarrow b_{jo}^+\)
   14. End
15. End
16. If \(f_1(a_{uo}) > 0\)
   17. Set \(a_{uo}^+ \leftarrow a_{uo}^-\)
   18. Else
      19. Set \(a_{uo}^- \leftarrow a_{uo}^+\)
   20. End
21. Compute \(f_1(a_{uo}), f_2(a_{mo})\) and \(f_3(b_{jo})\) based on (28), (29), (30) and (31).
22. End
23. Set \(a_{uo} \leftarrow a_{uo}^+, a_{mo} \leftarrow a_{mo}^+, b_{jo} \leftarrow b_{jo}^+, L_o \leftarrow L_{um,\min}^*\)
24. Return \(a_{uo}, a_{mo}, b_{jo}, L_o^*\)

**V. Numerical Results**

In this section, numerical and simulation results will be presented to analyze the performance of the considered communication network in terms of reliability performance and minimum blocklength and to provide some valuable insights for the system design. The effect of key parameters on the system performance is investigated. For clarity, unless otherwise stated, Table II summarizes the parameter settings adopted in this section [45], [46].

![Table II: Table of Parameters](image)

Fig. 2 plots the average BLERs of far and near users versus \(a_u\) for different power allocation coefficients, where the solid lines and discrete markers represent the analysis and simulation results, respectively. From Fig. 2(a), the average BLERs of two far users decline as \(a_u\) increases. In addition, when \(a_m\) rises, the reliability of \(U_m\) is enhanced, while that of \(U_n\) is weakened. This reflects the fact that performance trade-off between two far users can be achieved through power allocation. The average BLERs of the two near users first drop or stay unchanged in Fig. 2(b), then grows with the increase of \(a_u\). The part of reduction can be attributed to the increase in \(a_u\), which alleviates the difficulty of eliminating signals of far users and promotes the successful execution of SIC. However, as \(a_u\) continues to rise, less power is allocated to near users, which weakens the reliability. Furthermore, the reliability of both near users is weakened when the power allocation factors \(a_m\) and \(b_j\) increase simultaneously. However, with keeping \(a_m\) unchanged, the power allocation factor \(b_j\) can also realize the
Fig. 2. Average BLER at far and near users versus \( a \). As illustrated in Fig. 2(c), the BLERs of far users fall as the total power \( P \) goes up, confirming that high power is conducive to enhancing communication reliability. Additionally, the BLER is further reduced as the RIS assigns more reflective elements to far user, especially in high power region. This observation implies that there is a lot of space for the RIS to improve the reliability for far users. Moreover, when the RIS is employed in a multi-user scenario, the more stringent reliability requirements of users mean that the higher number of RIS elements is required, which also limits the total number of service users.

Fig. 4 depicts average BLERs at far users versus \( P \) for different \( \Omega_{sr} \). One can see from Fig. 4 that the reliability of far users is enhanced with the increase of the power, which is consistent with our expectations. In addition, when the distance between the BS and the RIS lengthens, BLER of users further climbs, which indicates that the shorter the distance between the BS and the RIS, the more beneficial for the reliability improvement of far users.

Fig. 5 illustrates average BLERs at far and near users versus \( \theta \) for different fading parameters. As shown in Fig. 5, the rise of transceiver hardware impairment degree leads to the reliability performance loss of users to varying degrees. Among them, users \( U_m \) and \( S_k \) are noticeably impacted, whereas the performance of users \( U_m \) and \( S_j \) is almost unaffected. This is due to the reliability of near users is directly affected by the power allocation factors of both far and near users. Specifically, the power allocation of far users affects the SIC implement, while the demodulation of their own signals is controlled by the power allocation of near users. Hence, near users need to comprehensively consider the power allocation to improve reliability provided that the total power is not too low.

In Fig. 3, we investigate the average BLERs at far users versus \( P \) for different number of elements of RIS sub-surface \( E \). As illustrated in Fig. 3, the BLERs of far users fall as the total power \( P \) goes up, confirming that high power is conducive to enhancing communication reliability. Additionally, the BLER is further reduced as the RIS assigns more reflective elements to far user, especially in high power region. This observation implies that there is a lot of space for the RIS to improve the reliability for far users. Moreover, when the RIS is employed in a multi-user scenario, the more stringent reliability requirements of users mean that the higher number of RIS elements is required, which also limits the total number of service users.

Fig. 3. Average BLER at far users versus \( P \) for different number of elements of RIS sub-surface \( E \).
In addition, it can be seen from Fig. 6(a) that the increase of blocklength will reduce the transmission rate. Nevertheless, the advanced fading parameter exhibits a downward trend as the fading parameters enlarge, which agrees with the outcome in Fig. 3, and that when the reliability requirements of the user are fixed, the increase in $E$ results in a shorter required packet length. According to Fig. 6(b), when two near users are closer to the BS, the reliability of the user is better.

In Fig. 7, we consider the minimum blocklength versus $P$ for different $E$ and $\eta$. It is shown that the minimum blocklength reduces with increasing $P$, which reveals high power can enhance the performance of short-packet networks. Regardless of reliability requirements of users are different (Fig. 7(a)) or same (Fig. 7(b)), the minimum block length can be shortened as $E$ increases, which indicates that the combination of RIS and short packet can yield better benefits since shorter blocklength contributes to higher spectrum utilization and lower latency. Additionally, the minimum block lengthens with the increasing data information bits, reflecting the inherent characteristics of SPC, which is well suited for the data transmission with less information bits, such as control.
empowered NOMA network with SPC over Nakagami-\textit{m} average BLER at far and near users. From the asymptotic view that the integration of NOMA and RIS can work well blocklength required at the same power. The results support the view that the integration of NOMA and RIS can work well for short packet transmission. Furthermore, the NOMA scheme surpasses the OMA scheme in terms of the minimum blocklength required at the same power. The results support the view that the integration of NOMA and RIS can work well for short packet transmission.

**VI. CONCLUSIONS**

In this paper, we analyzed the performance of an RIS-empowered NOMA network with SPC over Nakagami-\textit{m} fading channel in terms of the average BLER and minimum blocklength. Especially, transceiver hardware impairments have been taken into consideration to enhance the practicability and reasonability of the considered system model. We mathematically characterized the performance by deriving the approximate and asymptotic closed-form expressions for the average BLER at far and near users. From the asymptotic average BLER, we examined diversity order, minimum blocklength, and optimal power allocation. The results indicated that increasing the element number of reflective surfaces and lowering the distance between the BS and the RIS can improve the reliability for far users. Power allocation is a key factor for user reliability, and near users need to pay attention to the balance of SIC and self-signal demodulation provided that the total power is guaranteed. Furthermore, the user with worse channel conditions is less susceptible to hardware impairments whether for near users or RIS-assisted far users. Moreover, the system latency performance can be improved by increasing the number of RIS elements and decreasing the data information bits, and the advantages of NOMA and RIS in SPC are validated by the presented comparison.

**APPENDIX A**

**PROOF OF THEOREM 1**

Let $X_1 = \sum_{c=1}^{C_N} |h_{ru_m}|^2 |h_{sr_m}|$, we apply the causal form of the central limit theorem [47] and approximate the sum by gamma RV via Lemma 1. Observing (14), it is not difficult to find that the denominator of $k'$ and the numerator of $\theta'$ are variance. Thus, the mean and variance of $X_1$ can be calculated as

$$
E[X_1] = E\left[\sum_{c=1}^{C_N} |h_{ru_m}|^2 |h_{sr_m}| \right] = \sum_{c=1}^{C_N} E[|h_{ru_m}|^2] E[|h_{sr_m}|] = E_{ru_m} e_1,
$$

(33)

$$
Var[X_1] = Var\left[\sum_{c=1}^{C_N} |h_{ru_m}|^2 |h_{sr_m}| \right] = \sum_{c=1}^{C_N} Var[|h_{ru_m}|^2] E[|h_{sr_m}|] = E_{ru_m} (\Omega_{sr} - e_1^2),
$$

(34)

where $e_1 = \frac{\Gamma(m_{ru_m}+1/2)\Gamma(m_{sr}+1/2)}{\Gamma(m_{ru_m})\Gamma(m_{sr})\Omega_{sr}}$. As such, the distribution of $X_1$ can be approximated as $X_1 \sim \Gamma(k_1, k_2)$, $k_1$ and $\theta_1$ are given as

$$
k_1 = \frac{E_{ru_m} e_1^2}{\Omega_{sr} e_1^2 - e_1^2},
$$

(35)

After some mathematical manipulations, it yields the CDF of $X_1$:

$$
F_{X_1}(x) = \gamma \left( k_1, \frac{\sqrt{\theta_1}}{\theta_1} \right) \frac{1}{k_1}.
$$

(36)

Furthermore, the CDF of $\gamma_{ru_mru_m}$ is given as

$$
F_{\gamma_{ru_mru_m}}(x) = \gamma \left( k_1, \frac{\sqrt{\theta_1}}{\theta_1} \right),
$$

(37)

with (37) plugged into (13) and relying on the Riemann integral for the midpoint approximation, one can have (15).
APPENDIX B
PROOF OF THEOREM 2
Let us define $X_2 = \sum_{e=1}^{E_n} [h^e_{ru,n}] [h^e_{sr,n}]$, by using the similar methods in Appendix A, we have

$$
E[X_2] = E\sum_{e=1}^{E_n} [h^e_{ru,n}] [h^e_{sr,n}] = \sum_{e=1}^{E_n} E[|h^e_{ru,n}|] E[|h^e_{sr,n}|]
$$

(38)

$$
V a r[X_2] = V a r\sum_{e=1}^{E_n} [h^e_{ru,n}] [h^e_{sr,n}] = \sum_{e=1}^{E_n} V a r[|h^e_{ru,n}|] V a r[|h^e_{sr,n}|]
$$

(39)

where $e_2 = \frac{\Gamma(m_{sr:n}+1/2) \Gamma(m_{sr:n}+1/2)}{\Gamma(m_{sr:n}) \Gamma(m_{sr:n}+\delta_{sr:n})}$. The distribution of $X_2$ can be approximated as $X_2 \sim \Gamma(k_2, \theta_2)$, $k_2$ and $\theta_2$ are given as

$$
k_2 = \frac{E_n e_2}{\Omega_{ru} \Omega_{sr} - e_2^2}, \theta_2 = \frac{\Omega_{ru} \Omega_{sr} - e_2^2}{e_2}
$$

(40)

After some algebraic manipulations, the CDF of $\gamma_{u_n \rightarrow u_m}$ and $\gamma_{u_n \rightarrow u_n}$ can be respectively obtained as

$$
F_{\gamma_{u_n \rightarrow u_m}}(x) = \frac{\gamma(k_2, \frac{\lambda_{u_m u_n} \rho - 1}{\theta_2})}{\Gamma(k_2)}
$$

(41)

$$
F_{\gamma_{u_n \rightarrow u_n}}(x) = \frac{\gamma(k_2, \frac{\lambda_{u_n u_n} \rho}{\theta_2})}{\Gamma(k_2)}
$$

(42)

where $\gamma_{u_n \rightarrow u_m} = \frac{E_n}{\sum_{e=1}^{E_n} [h^e_{ru,n}] [h^e_{sr,n}] + \lambda_{u_n u_m}}$, and $\gamma_{u_n \rightarrow u_n}$ is given in (7), with (41) and (42) plugged into (13) respectively and relying on the Riemann integral for the midpoint approximation, we can calculate the average BLER of user $S_j$ as (21).

APPENDIX D
PROOF OF THEOREM 4
Utilizing the CDF of $|h_{s_{x,j}}|^2$, $F_{|h_{s_{x,j}}|^2}(x) = \frac{\gamma(k_4, \frac{x}{\theta_4})}{\Gamma(k_4)}$, similar to the procedure in Appendix C, we can compute the CDF of $\gamma_{s_k \rightarrow u_m}$, $\gamma_{s_k \rightarrow u_n}$, $\gamma_{s_k \rightarrow s_j}$, and $\gamma_{s_k \rightarrow s_k}$ as

$$
F_{\gamma_{s_k \rightarrow u_m}}(x) = \frac{\gamma(k_4, \frac{\lambda_{u_m u_n} \rho - 1}{\theta_4})}{\Gamma(k_4)}
$$

(46)

$$
F_{\gamma_{s_k \rightarrow u_n}}(x) = \frac{\gamma(k_4, \frac{\lambda_{u_n u_n} \rho}{\theta_4})}{\Gamma(k_4)}
$$

(47)

$$
F_{\gamma_{s_k \rightarrow s_j}}(x) = \frac{\gamma(k_4, \frac{\lambda_{s_j s_j} \rho - x}{\theta_4})}{\Gamma(k_4)}
$$

(48)

$$
F_{\gamma_{s_k \rightarrow s_k}}(x) = \frac{\gamma(k_4, \frac{\lambda_{s_k s_k} \rho - x}{\theta_4})}{\Gamma(k_4)}
$$

(49)

where $\gamma_{s_k \rightarrow u_m} = \frac{|h_{s_{x,j}}|^2 \lambda_{u_m u_n} \rho}{|h_{s_{x,j}}|^2 \lambda_{u_m u_n} + 1}$, $\gamma_{s_k \rightarrow u_n} = \frac{|h_{s_{x,j}}|^2 \lambda_{u_n u_n} \rho}{|h_{s_{x,j}}|^2 \lambda_{u_n u_n} + 1}$, $\gamma_{s_k \rightarrow s_j} = \frac{|h_{s_{x,j}}|^2 \lambda_{s_j s_j} \rho - x}{|h_{s_{x,j}}|^2 \lambda_{s_j s_j} + 1}$, and $\gamma_{s_k \rightarrow s_k}$ is given in (9). With (46), (47), (48) and (49) plugged into (13) respectively and relying on the Riemann integral for the midpoint approximation, the average BLER of user $S_k$ is obtained as (22). The proof is completed.

REFERENCES


[56] P. T. Tran, B. C. Nguyen, T. M. Hoang, X. H. Le, and V. D. Nguyen, “Exploiting multiple RISs and direct link for performance enhancement