

A Power and Spectrum Efficient Uplink Transmission Scheme for QoS-Constrained IoT Networks

Hamad Yahya, *Graduate Student Member, IEEE*, Arafat Al-Dweik, *Senior Member, IEEE*, Youssef Iraqi, *Senior Member, IEEE*, Emad Alsusa, *Senior Member, IEEE* and Ashfaq Ahmed

Abstract—Non-orthogonal multiplexing (NOM) is a novel superposition coding inspired scheme that has been recently proposed for improving the power, spectrum efficiency and delay of wireless links with packet error rate (PER) constraints. Despite its efficiency, restricting the number of multiplexed packets to two limits the throughput improvement to 100%. Therefore, this work presents a novel NOM design with unlimited number of multiplexed packets by manipulating the repeated transmissions in automatic repeat request (ARQ) to enhance the power and spectrum efficiency by multiplexing new and repeated packets while taking into account the channel conditions and varying the power per packet in different transmissions. The proposed scheme employs an efficient heuristic algorithm to perform the power assignment and multiplexing decisions. Moreover, the complexity of the proposed NOM can be controlled by enforcing a limit on the maximum number of multiplexed packets per transmission, making it suitable for different types of Internet of Things (IoT) nodes with various computational capabilities. The obtained results demonstrate the effectiveness of proposed scheme, which offers up to 200% spectral efficiency improvement at moderate signal to noise ratios (SNRs), and up to 700% at high SNRs. Furthermore, the new scheme can reduce the transmission power consumption by up to 6 dB in the high SNR region.

Index Terms—Internet of Things (IoT), Internet of Vehicles (IoV), Internet of Drones (IoD), automatic repeat request (ARQ), vehicle to infrastructure (V2I), vehicle-to-vehicle (V2V), non-orthogonal multiple access (NOMA), non-orthogonal multiplexing (NOM), throughput, spectral efficiency.

I. INTRODUCTION

The integration of Internet of Things (IoT) in various applications has been growing vastly in the last decade, which triggered the introduction of more specific applications such as Internet of Drones (IoD) [1]–[3] and Internet of Vehicles (IoV) [4]–[7], or even more advanced configurations that include both IoD and IoV as shown in Fig. 1. In such applications, ultra reliable low latency communication (URLLC)

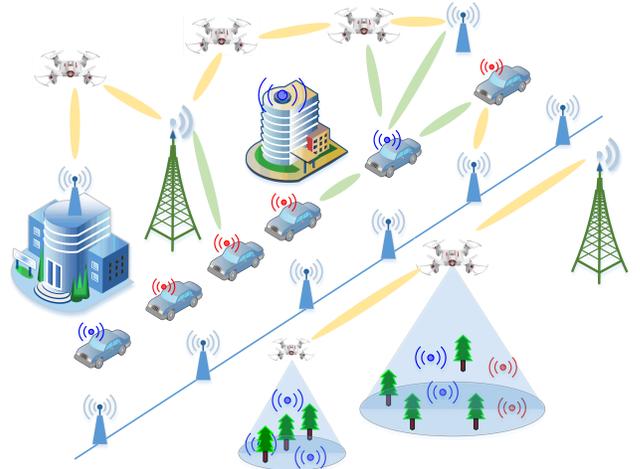


Fig. 1. A general IoT diagram showing examples of IoV and IoD with V2V and V2I connectivity.

transmission is crucial because erroneous or delayed information can cause fatal consequences. For example, vehicles in IoV environment should perform vehicle-to-vehicle (V2V) and vehicle to infrastructure (V2I) connectivity to enable vehicles make certain critical decisions [8]. IoD applications have similar requirements because the drones typically have to collect and relay location-specific critical information [1]–[3]. Therefore, the received data should be verified for correctness before being used to make decisions. In the context of the decision making processes, decision fusion has been recognized as one of the most efficient approaches to improve the reliability of the decisions made about a certain phenomenon [9], [10]. Nevertheless, having redundant information from multiple sources to perform fusion is infeasible in several applications.

In addition to IoT applications, having reliable data is required for most current applications that rely on wireless communications. The international telecommunication union (ITU) has specified that the packet error rate (PER) for any class of service should not be less than 10^{-4} [11, Table I]. In this context, automatic repeat request (ARQ) is considered as the prominent technique that can be used to achieve this goal. Consequently, ARQ is adopted for the fifth generation (5G) [12] and for several other wireless communications standards such as long term evolution (LTE), wireless fidelity

Hamad Yahya and Emad Alsusa are with the Department of Electrical and Electronic Engineering, The University of Manchester, Manchester M13 9PL, U.K. (email: hamad.mohamadaliyahya@manchester.ac.uk; e.alsusa@manchester.ac.uk).

A. Al-Dweik is with the Center for Cyber-Physical Systems (C2PS), Khalifa University, 127788 Abu Dhabi, UAE, and also with the Department of Electrical and Computer Engineering, Western University, London, ON N6A 3K7, Canada (e-mail: dweik@fulbrightmail.org; arafat.dweik@ku.ac.ae);

Y. Iraqi is with the Center for Cyber-Physical Systems (C2PS), Khalifa University, 127788 Abu Dhabi, UAE, (e-mail: youssef.iraqi@ku.ac.ae).

A. Ahmed is with the Department of Electrical Engineering and Computer Science, Khalifa University, 127788 Abu Dhabi, UAE (email: ashfaq.ahmed@ku.ac.ae).

(WiFi), wireless personal area networks (WPANs) and world-wide interoperability for microwave access (WiMAX) [11]. Moreover, ARQ has been considered for several technologies such as non-orthogonal multiple access (NOMA) [13]–[15], multiple-input multiple-output (MIMO), [16], massive MIMO (mMIMO) [17] and cognitive radio (CR) [18].

ARQ allows nearly error-free transmission by checking all received packets for errors and instructing the transmitter to retransmit the packets that fail the error check. Packets that never pass the error check will be eventually dropped [11]. Despite its advantages, ARQ may cause severe throughput and power degradation, particularly at low and moderate signal to noise ratios (SNRs) [19] due to the repeatedly transmitted data and the need for frequent feedback messages from the receiver to the transmitter. Consequently, extensive research in the literature is devoted to mitigate the adverse effects of ARQ. For example, integrating forward error correction (FEC) and ARQ, denoted as hybrid-ARQ (HARQ), can reduce the number of retransmissions by reducing the PER [20]. Moreover, HARQ provides more flexibility to control the type and amount of retransmitted data, which can be optimized to improve the system efficiency [11]. However, FEC may also reduce the system throughput and power due to the redundant parity bits, and the additional encoding and decoding complexity and delay. The packet length is a critical parameter that affects the PER, and thus, optimizing the packet length can significantly improve the system throughput [21].

Utilizing time diversity by combining the multiple retransmitted packets, known as chase combining (CC), can significantly improve ARQ system efficiency. In ARQ-CC systems, maximum ratio combining (MRC) can be used to combine all the transmissions that correspond to a particular packet to reduce the PER and improve the throughput. CC also enables a remarkable power saving because the retransmitted packets usually do not require the same power used in the initial packet transmission [19]. Moreover, the fact that full power is not required for retransmitted data enables multiplexing new and retransmitted bits in the power domain [22]–[24]. In [22], the focus is to reduce the delay using an early retransmission technique. The paper also considers combining new and retransmitted packets using superposition coding where new packets are allocated fixed high power as compared to retransmitted packets. In each transmission session, a maximum of one new and one retransmitted packets can be combined and transmitted simultaneously. However, the fixed power assignment and false negative acknowledgments (NACKs) caused by early retransmission decisions eliminate the throughput gain that could have been achieved using the superposition coding. Khreis *et al.* [23] proposed a multi-layer HARQ that utilizes the idle time that the transmitter usually experiences while waiting for a feedback message (F_M). The main concept is to retransmit certain packets multiple times without waiting for a F_M . To avoid throughput losses, the redundant data is combined with new transmission using superposition coding. The results presented in [23] show that the throughput improvement is about 10%. Non-orthogonal multiplexing (NOM) is used in [24] to utilize the power and spectrum efficiently where a low complexity heuristic

algorithm is proposed to combine new and retransmitted data. Although the proposed algorithm provides up to 100% throughput improvement at high SNRs, the work considers fixed power assignment and the combining is limited only to two packets.

As can be noted from the discussed literature and the reference listed therein, NOM has a great potential to overcome the limitations inherent in ARQ. Therefore, this article presents a novel NOM for ARQ-based wireless systems to improve their power and spectral efficiency, and as consequence, the throughput will be improved as well. In the proposed NOM-ARQ, the transmitter computes the power coefficients of the multiplexed NOM packets to stack the maximum possible number of packets in a single transmission while satisfying a predetermined PER constraint. Moreover, the total transmit power will be fixed and normalized to unity. The obtained numerical results show that a significant spectral and throughput gain can be achieved using the proposed scheme when compared to conventional a ARQ system. The power efficiency is also evaluated in terms of average power per packet. To suit typical IoT nodes hardware limitations, the proposed system design is tuned to provide efficient power consumption, small buffering requirements, and reduced decoding complexity.

The rest of the paper is organized as follows. The conventional ARQ system and channel models are described in Section II. The proposed ARQ system with adaptive power allocation is explained in Section III. The proposed ARQ transmission protocol is presented in Section IV. Simulation and numerical results are discussed in Section V. The conclusion is drawn in Section VI.

II. ARQ SYSTEM AND CHANNEL MODELS

In this work, multiplexing data packets that belong to a single user in the power domain is denoted as NOM, while if the data packets belong to multiple users the system is denoted as NOMA. The ARQ system considered in this work consists of an IoT node that needs to send M packets over a wireless channel. Therefore, the system configuration corresponds to the uplink transmission in IoT networks. The transmitter and receiver utilize truncated ARQ with stop-and-wait (SW) flow control protocol where the maximum number of allowed transmissions is \mathcal{M} , which includes the initial transmission. A packet that is transmitted \mathcal{M} times and still has errors will be dropped. Without loss of generality, it is assumed that binary phase shift keying (BPSK) modulation is adopted at the transmitter to ensure reliable transmission. BPSK modulation is adopted in various standards such as the NB-IoT standard [25]. Nevertheless, it is straightforward to extend the work to other modulation schemes. The channel between the transmitter and receiver is considered as a block fading where the channel remains fixed for the period of one packet but changes randomly over consecutive packets. Therefore, the received packet that corresponds to the k th transmitted packet can be written as

$$\mathbf{y}_{k,l}^{(u)} = h_{k,l}^{(u)} \sqrt{\alpha_{k,l}^{(u)}} \mathbf{x}_{k,l}^{(u)} + \mathbf{w}_{k,l}^{(u)} \quad (1)$$

where l is the transmission slot, u is the transmission/retransmission counter for each packet, α is the transmit

power which is normalized to unity, $h \sim \mathcal{CN}(0, \sigma_h^2)$ is the channel frequency response (CFR), \mathbf{w} is the additive white Gaussian noise (AWGN) vector, $\mathbf{w} = [w_1, w_2, \dots, w_L]$, $w_i \sim \mathcal{CN}(0, \sigma_w^2)$, and $\mathbf{x} = [x_1, x_2, \dots, x_L]$ is the transmitted data packet, $x_i \in \{-1, 1\} \forall i$, and L is the packet length in symbols.

At the receiver, the received symbols in each packet are demodulated, and then the packet undergoes an error detection process, which is typically considered to be perfect, i.e., probability of false alarm (PFA) and probability of miss (PM) are equal to zero. After error detection, an F_M is sent back to the transmitter to instruct it to either retransmit the previously transmitted packet when the F_M is a NACK, or send a new packet when F_M is an acknowledgment (ACK). In this work, we consider Type-I ARQ, and thus, the same packet can be transmitted up to \mathcal{M} times, $u \in \{1, 2, \dots, \mathcal{M}\}$. The receiver can exploit the channel temporal variations and combine the multiple received versions of the same packet using CC. Therefore, the PER for the k^{th} packet after the u^{th} transmission can be written as

$$\bar{P}_{k,l}^{(u)} = 1 - (1 - P_{k,l}^{(u)})^L \quad (2)$$

where

$$P_{k,l}^{(u)} = Q \left(\sqrt{\sum_{i=1}^u \alpha_{k,l+i}^{(i)} \gamma_{k,l+i}^{(i)}} \right) \quad (3)$$

and $\gamma_{k,l+i}^{(i)} = \frac{2|h_{k,l+i}^{(i)}|^2}{\sigma_w^2}$. The transmission slot index l corresponds to the slot in which packet k was transmitted for the first time.

III. PROPOSED NOM-ARQ SYSTEM

In conventional ARQ, the power per packet can be considered fixed, i.e., $\alpha_{k,l}^{(u)} = \alpha \forall \{k, u, l\}$. Alternatively, $\alpha_{k,l}^{(u)}$ can be optimized to save power [19], which is necessary in the case that a packet is transmitted more than once. In such scenarios, the transmitter may frequently experience cases with $\alpha_{k,l}^{(u)} \leq p_{\max}$, where p_{\max} is the transmitter maximum transmit power. Consequently, the power amplifier efficiency at the transmitter could be significantly harmed. Therefore, we propose in this work to opportunistically multiplex multiple packets at the transmitter in the power domain such that the total transmit power is reasonably large. In addition to resolving the power amplifier efficiency problem, a significant spectral efficiency improvement can be gained because multiple packets can be transmitted using the same transmission time/frequency resources that are used in conventional ARQ systems. As a consequence of the spectral efficiency improvement, the delay can be reduced since it is inversely proportional to the transmission rate. Moreover, the average power per packet will be reduced because several packets with small power values will be successfully received.

A. Proposed NOM

Similar to conventional ARQ, consider that M packets are buffered at the transmitting node, and will be transmitted over a wireless channel to a receiver. We assume that the CFR of

that transmission session is known at the transmitter, which is a widely adopted assumption in stationary, low, and moderate mobility systems. If the channel conditions allow multiplexing more than one packet into a particular transmission session, then the transmitted NOM packets during the first transmission slot can be written as

$$\mathbf{d}_1 = \sum_{i=1}^{N_1} \sqrt{\alpha_{i,1}^{(1)}} \mathbf{x}_{i,1}^{(1)} \quad (4)$$

where N_1 is the number of packets that are multiplexed in the current transmission slot $N_1 \leq N_{\text{cap}}$, N_{cap} is the maximum number of NOM packets that can be multiplexed, $\alpha_{i,1}^{(1)}$ is the power coefficient of the i^{th} NOM packet, the total power in the transmission slot is normalized to unity, and thus, $\sum_{i=1}^{N_1} \alpha_{i,1}^{(1)} = 1$.

Although the power per packet in (4) can be assigned in an arbitrary manner, the selected powers will affect the receiver design in terms of computational complexity, buffering requirements and packet delay. Therefore, to enable using the low complexity successive interference cancellation (SIC) receiver, minimize the buffering requirements, and avoid jitter, the power selection should be performed following the general rules used for NOMA. To achieve these goals, the system design should adhere to the following general design rules:

- 1) The packets' sequence numbers are used as a priority indicators, that is, a packet with the lowest index has the highest priority, and so forth. Packets with higher priority should be assigned more power than packets with lower priority, $\alpha_{1,l}^{(u)} > \alpha_{2,l}^{(u)} > \dots > \alpha_{N_l,l}^{(u)}$. More specific rules should be specified based on the adopted modulation scheme and the number of multiplexed packets [26].
- 2) Because the packet with the highest priority will be allocated the highest power, it will be highly likely that lower priority packets will not pass the error check if the primary packet fails the error check. Consequently, all multiplexed packets will be simultaneously dropped after \mathcal{M} unsuccessful transmissions. To prevent such scenarios, the packet with the minimum index, denoted as the primary packet, is the only packet that its transmission counter is incremented when it is received unsuccessfully. Consequently, only the primary packet will be dropped after \mathcal{M} unsuccessful transmissions. For example, given that the primary packet \mathbf{x}_1 has failed $m - 1$ times, then the composite packet in the m^{th} transmission time slot can be written as

$$\mathbf{d}_m = \sqrt{\alpha_{1,m}^{(m)}} \mathbf{x}_{1,m}^{(m)} + \sum_{i=2}^{N_m} \sqrt{\alpha_{i,m}^{(1)}} \mathbf{x}_{i,m}^{(1)}. \quad (5)$$

As can be noted from (5), the transmission counter u is incremented only for \mathbf{x}_1 .

At the receiver, the received composite packet that corresponds to \mathbf{d}_m can be expressed as

$$\mathbf{y}_m = h_m \left(\sqrt{\alpha_{1,m}^{(m)}} \mathbf{x}_{1,m}^{(m)} + \sum_{i=2}^{N_m} \sqrt{\alpha_{i,m}^{(1)}} \mathbf{x}_{i,m}^{(1)} \right) + \mathbf{w}_m. \quad (6)$$

For notational convenience, the CFR h for the composite packet \mathbf{d}_m will be written as $h_{k,l}^{(u)}$. Therefore,

$$\mathbf{y}_m = h_{1,m}^{(m)} \sqrt{\alpha_{1,m}^{(m)}} \mathbf{x}_{1,m}^{(m)} + \sum_{i=2}^{N_m} h_{i,m}^{(1)} \sqrt{\alpha_{i,m}^{(1)}} \mathbf{x}_{i,m}^{(1)} + \mathbf{w}_m \quad (7)$$

where $h_{k,l}^{(u)} = h_l \forall \{k, u\}$. Moreover, the transmission time slot index, l , will be dropped unless it is necessary to include it. Following the SIC detection principle, the primary packet is detected first, and then the secondary packets are detected according to their index values. However, if the primary packet is not detected successfully, then the interference cancellation will fail and most likely all secondary packets will not be detected successfully as well. Consequently, after the detection stage, the primary packet will be tested for errors, and if it fails, the detection process for all secondary packets will be aborted, and these packets will be considered erroneous. The same procedure is applied to the secondary packets, i.e., if \mathbf{x}_i fails the error check, then $\mathbf{x}_l \forall l > i$ will be automatically considered erroneous. Consequently, F_M will have the form $\mathcal{A} = [1, 1, \dots, 1, 0, 0, \dots, 0]$ where 1 stands for ACK and 0 stands for NACK.

B. Power Allocation

The number of multiplexed NOM packets per transmission slot mostly depends on the CFR and noise power. Before the commencement of the packet transmission process, a power allocation routine is performed to allocate the available power to maximize the possible number of multiplexed packets N_l during each transmission slot to maximize the spectral efficiency while satisfying the PER constraint. To reduce the complexity and the signaling overhead, the routine is called only when the counter of the primary packet $u = 1$, and all transmission parameters are kept fixed during any retransmission process, if there is any. The maximum number of packets that can be multiplexed in a particular transmission session is obtained using Algorithm 1, which can be described as follows:

- 1) Compute the PER for a single packet using (2) and (3), and compare it to the PER threshold τ . If $\bar{P}_{1,l}^{(1)} < \tau$, go to next step, otherwise $N_l = 1$ and thus, only one packet will be transmitted.
- 2) Compute the power coefficients to transmit a NOM packet. The process starts with two multiplexed packets and may continue up to N_{cap} . The optimization problem

for a given N_l can be formulated as follows

$$\arg \min_{\alpha \in (0,1)} P_{N_l,l}^{(1)} \quad (8a)$$

subject to:

$$\bar{P}_{j,l} \leq \tau, \quad \forall j \in \{1, \dots, N_l - 1\} \quad (8b)$$

$$\sqrt{\alpha_{j,l}^{(1)}} > \sum_{r=j+1}^{N_l} \sqrt{\alpha_{r,l}^{(1)}}, \quad \forall j \in \{1, \dots, N_l - 1\} \quad (8c)$$

$$\sqrt{\alpha_{N_l,l}^{(1)}} > 0 \quad (8d)$$

$$\sum_{r=j+1}^{N_l} \sqrt{\alpha_{r,l}^{(1)}} = 1 \quad (8e)$$

where the objective in (8a) is to find power coefficient α for which the bit error rate (BER) of the last secondary packet, $\mathbf{x}_{N_l,l}$, is minimized while satisfying the constraints (8b)–(8e). The constraint in (8b) is non-linear, and it is used to satisfy the PER threshold for the primary and all secondary packets except $\mathbf{x}_{N_l,l}$ which will be transmitted regardless of the PER constraint. Constraints (8c)–(8d) are required to satisfy power coefficient bounds [26, Eq. (21)], (8e) ensures a unity sum of power coefficients during a transmission slot. The BER for NOM can be written as [27, Eq. (23)]

$$P_{k,l}^{(u)} = \frac{1}{2^{N_l-1}} \sum_j c_j Q \left(\sqrt{|\Delta_j|^2 \sum_{i=1}^u \gamma_{k,l+i}^{(i)}} \right) \quad (9)$$

where c_j and Δ_j values depend on α and N_l . Similarly, the PER can be calculated according to (2).

- 3) Once α is calculated for $N_l = 2$, the PER constraint satisfaction is checked for the secondary packet. If it is strictly satisfied or not satisfied, then $N_l = 2$ and the computed α are returned. Otherwise, α for $N_l = 3$ will be computed. This loop will continue until $N_l = N_{\text{cap}}$ unless the PER constraint of the N_l^{th} packet is strictly satisfied or not satisfied.

It is worth mentioning that if a primary packet is not successfully decoded, then a retransmission is requested with the same parameters. Upon receiving the retransmitted realizations, the receiver applies MRC to combine the received packets up to the current transmission. Whenever the primary packet is not detected correctly, the received composite packets are buffered for combining with the upcoming retransmissions. If the primary packet exceeds \mathcal{M} without being successfully detected, then the composite packet and all its buffered versions are dropped and the receiving buffer is cleared. Otherwise, the primary packet can be subtracted from the combined packet to allow combining and detecting the secondary packets.

IV. NOM-ARQ TRANSMISSION PROTOCOL

As can be noted from the power assignment process, the transmitter may send one packet, or up to N_{cap} multiplexed

Algorithm 1: Optimal power allocation

Input: $\gamma_{1,l}^{(1)}, \tau, L, N_{\text{cap}}$
Output: N_l, α_l

```

1  $i = 1$ 
2 Compute PER using (2) and (3).
3  $\mathbf{p}_i = 1$ 
4 if  $\bar{P}_{1,l}^{(1)} \geq \tau$  then
5    $N_l = i$ 
6    $\alpha_l = [1]$ 
7   return
8 else
9   for  $i = 2 : N_{\text{cap}}$  do
10    Compute  $\alpha$  using (8a)–(8e) with  $N_l = i$ 
11     $\mathbf{p}_i = \alpha$ 
12    Compute PER using (2) and (9).
13    if  $\bar{P}_{i,l}^{(1)} \geq \tau$  then
14       $N_l = i$ 
15       $\alpha_l = \mathbf{p}_i$ 
16      break
17    else
18      continue

```

packets in each transmission slot. According to the SW protocol, the transmitter will be expecting a F_M for each transmitted packet. The F_M will contain the feedback for all multiplexed packets in a vector \mathcal{A} , which can be $\mathcal{A} = [\mathcal{A}_1]$, $\mathcal{A} = [\mathcal{A}_1, \mathcal{A}_2]$, or $\mathcal{A} = [\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_{N_{\text{cap}}}]$, $\mathcal{A}_i \in \{0, 1\} \forall i$, where 1 represents an ACK and 0 is for a NACK. However, in the case that a packet is not detected correctly, it will cause severe interference to the low-power packets, and thus, most likely they will not be detected correctly as well. Therefore, to save time and computational complexity, if packet \mathbf{x}_i fails, then the detection of \mathbf{x}_{i+1} or $\{\mathbf{x}_{i+1}, \mathbf{x}_{i+2}\}$ is aborted and all lower power packets will be considered incorrect. Consequently, the F_M for the $N_{\text{cap}} = 3$ case, can only be $\mathcal{A} = [0, 0, 0]$, $\mathcal{A} = [1, 0, 0]$, $\mathcal{A} = [1, 1, 0]$, or $\mathcal{A} = [1, 1, 1]$. In the case that $\mathcal{A} = [0, 0, 0]$, the transmitter considers that all three packets have failed. Therefore, the packet with index i remains the primary packet during the following transmission, but the transmission index l is increased by one. If the primary packet is decoded correctly during its initial transmission but the secondary packets are unable to pass the error check, the acknowledgment $\mathcal{A} = [1, 0, 0]$ is sent back to the transmitter. In the next transmission slot, the primary packet becomes the packet with index $i + 1$. In addition, the buffer containing the observations and channel realization of the previous transmission slot will be cleared. Moreover, the primary packet's transmission counter u is reset to 1 for the new primary packet. It is worth noting that the primary packet will be dropped if the primary packet transmission counter exceeds the maximum number of allowed transmissions \mathcal{M} .

A. Receiver Design

Assuming that the power assignment routine during the first transmission slot, $l = 1$, returns $N_1 = 3$ and $\alpha_1 = [\alpha_1, \alpha_2, \alpha_3]$. Hence, the system transmits three packets during this transmission slot. The modulated packets, simply termed as packets, are denoted as $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$, and packet \mathbf{x}_1 is the primary packet. The receiver employs SIC to remove inter-packet interference for the detection of individual packets from the received composite packet. For example, during the first transmission, the received composite packet is given as

$$\mathbf{y}_1 = \left(\sqrt{\alpha_{1,1}^{(1)}} \mathbf{x}_{1,1}^{(1)} + \sqrt{\alpha_{2,1}^{(1)}} \mathbf{x}_{2,1}^{(1)} + \sqrt{\alpha_{3,1}^{(1)}} \mathbf{x}_{3,1}^{(1)} \right) h_1 + \mathbf{w}_1. \quad (10)$$

The receiver directly decodes the primary packet \mathbf{x}_1 using the BPSK hard detector during its initial transmission. If the decoding is successful, the extraction and decoding of the second packet \mathbf{x}_2 is initiated using the decoded \mathbf{x}_1 . Similarly, the third packet \mathbf{x}_3 is extracted and decoded using the successfully decoded packets \mathbf{x}_1 and \mathbf{x}_2 .

If the decoded packet \mathbf{x}_1 does not pass the error check, the channel information h_1 and the currently received composite packet \mathbf{y}_1 are stored in the buffer, and F_M of $\mathcal{A} = [0, 0, 0]$ is returned. After the retransmission with the same transmission parameters, the MRC is applied to combine the realizations of the primary packet received till l^{th} transmission. The output of the combining process for the primary packet is given as

$$\hat{\mathbf{y}}_{1,l} = \frac{\sum_{k=1}^l \sqrt{\alpha_{1,k}^{(k)}} \mathbf{y}_k h_k^*}{\sum_{k=1}^l \alpha_{1,k}^{(k)} |h_k|^2}. \quad (11)$$

The combined primary packet $\hat{\mathbf{y}}_{1,l}$ is then decoded using BPSK detector to get $\bar{\mathbf{x}}_1$. If $\bar{\mathbf{x}}_1 \neq \mathbf{x}_1$ and $l \leq \mathcal{M}$, the received composite packet \mathbf{y}_l , the channel information h_l , and $\alpha_l, \forall l$ are stored in the receiver's buffer, and a retransmission is requested. If $l > \mathcal{M}$, the primary packet, \mathbf{x}_1 , is dropped, and the next packet, \mathbf{x}_2 , becomes the primary packet in the following transmission. In case the receiver successfully decodes the primary packet, $\bar{\mathbf{x}}_1$ is further used to extract all the realizations of the secondary packets up to l^{th} transmission round. The extraction of the second packet's realizations is given as

$$\tilde{\mathbf{y}}_{2,k} = \mathbf{y}_k - \sqrt{\alpha_{1,k}^{(k)}} \bar{\mathbf{x}}_1 h_k, \quad \forall k \in \{1, \dots, l\}. \quad (12)$$

The extracted secondary packets are then combined as,

$$\hat{\mathbf{y}}_{2,l} = \frac{\sum_{k=1}^l \sqrt{\alpha_{2,k}^{(k)}} \tilde{\mathbf{y}}_{2,k} h_k^*}{\sum_{k=1}^l \alpha_{2,k}^{(k)} |h_k|^2}. \quad (13)$$

The combined secondary packet is then decoded to get $\bar{\mathbf{x}}_2$. If $\bar{\mathbf{x}}_2 \neq \mathbf{x}_2$, the realizations of the third packet are extracted using $\bar{\mathbf{x}}_1$ and $\bar{\mathbf{x}}_2$. The extraction of the third packet is performed as

$$\tilde{\mathbf{y}}_{3,k} = \mathbf{y}_k - \left(\sqrt{\alpha_{1,k}^{(k)}} \bar{\mathbf{x}}_1 + \sqrt{\alpha_{2,k}^{(k)}} \bar{\mathbf{x}}_2 \right) h_k, \quad \forall k \in \{1, \dots, l\}. \quad (14)$$

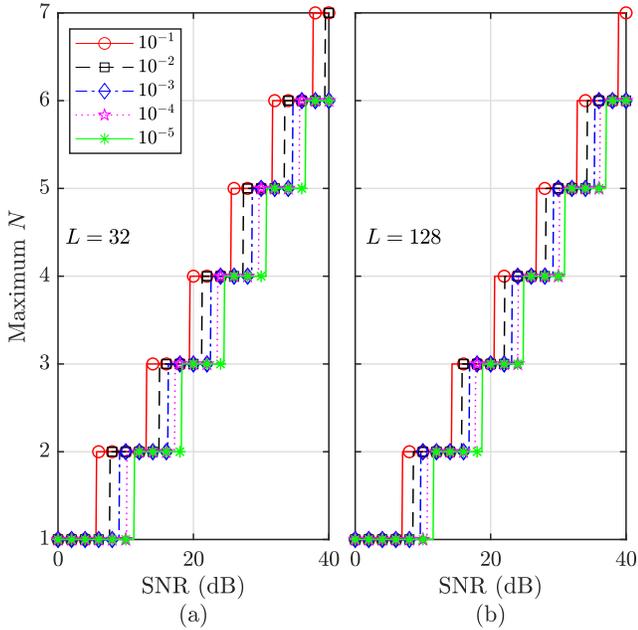


Fig. 2. Spectral efficiency, N , of the proposed NOM-ARQ scheme for various PER thresholds in AWGN channel. (a) $L = 32$. (b) $L = 128$.

The realizations of the third packet are then combined as

$$\hat{\mathbf{y}}_{3,l} = \frac{\sum_{k=1}^l \sqrt{\alpha_{3,k}^{(k)}} \tilde{\mathbf{y}}_{3,k} h_k^*}{\sum_{k=1}^l \alpha_{3,k}^{(k)} |h_k|^2}. \quad (15)$$

Finally, the combined tertiary packet is decoded to get $\bar{\mathbf{x}}_3$.

V. NUMERICAL RESULTS AND DISCUSSIONS

This section presents the performance evaluation of the proposed transmission model. Monte Carlo simulation is performed for $M = 3 \times 10^6$ packet realization with $N_{\text{cap}} = 7$ packets per transmission slot. Two packet lengths are considered, $L = 32$ and 128 symbols, and maximum allowable transmissions $\mathcal{M} = 3$. The performance is evaluated for PER threshold $\tau \in \{10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}\}$. The performance of the proposed system is measured through several metrics and compared to an ARQ system in which a single packet with normalized unity power is transmitted. The SNR shown in the figures is defined as $\text{SNR} \triangleq 1/\sigma_w^2$. Additionally, in the legend, NOM denotes the proposed NOM-ARQ while ARQ denotes the conventional ARQ system.

The maximum number of multiplexed packets versus SNR over AWGN channels is shown in Fig. 2. It should be noted that N , in this context, represents spectral efficiency with a unit of bit per channel use. As can be noted from the figure, when the PER threshold decreases, higher SNR is required to obtain the same value of N . Additionally, the same conclusion applies for packets with higher length as the PER is higher for longer packets. Furthermore, the optimal power allocation per packet for the AWGN channel is shown in Fig. 3. It is noted that power coefficients for the secondary packets decreases significantly when the total number of multiplexed packets increases. Moreover, the more stringent the PER requirements, the smaller the power coefficients for the secondary packets

get, whereas the primary packet power coefficient increases. It is worth mentioning that at relatively high SNRs, the power coefficients per packet converge regardless of PER threshold.

The average power allocations per packet for the proposed NOM-ARQ in a Rayleigh fading channel is demonstrated in Fig. 4. Before each transmission, the value of N and the power coefficients for the N packets are determined using Algorithm 1. The average power allocation per packet is computed as the ratio of the sum of powers allocated per packet to the total channel uses. As depicted in the figure, when the PER threshold reduces from 10^{-1} to 10^{-5} , more power is allocated to the primary packet to satisfy its PER threshold. Therefore, decreasing the PER threshold for the primary packet increases its average power allocation. Such performance is obtained because the primary packet has the highest priority in terms of PER satisfaction. Moreover, similar to the AWGN case, the power coefficients per packet at high SNRs converge to the same value regardless of the PER threshold.

Throughput is considered as one of the important metrics to examine the performance of a retransmission system. The throughput of the proposed system for various PER thresholds is shown in Fig. 5, where the throughput is defined as the total number of successfully received packets divided by a total number of transmission sessions. The spectral efficiency is considered as an upper bound for the throughput because the spectral efficiency does not consider dropped and retransmitted packets. From the results, it is evident that a significant throughput gain is achieved using the proposed NOM. For example, the throughput gain exceeds 350% at SNR of 30 dB for $L = 32$. **The impact of the PER threshold is generally mild for the considered range of PER thresholds, where the difference is about 0.3 bits for SNRs more than 25 dB.** As can be noted from the figure, the throughput degrades by decreasing the PER threshold because higher power is required to satisfy the PER constraint. Therefore, the number of multiplexed packets is reduced. Using the same justification, the throughput is improved when the PER is relaxed because more packets can be multiplexed at the transmitter. However, the throughput will not consistently improve because increasing the PER thresholds will increase the packet retransmission rates. The NOM-ARQ and the conventional ARQ are performing roughly the same at low SNRs because the outcome of the adaptation process most likely will result in a single packet transmission.

The packet drop rate (PDR) for NOM-ARQ and the conventional ARQ is demonstrated in Fig. 6. The PDR is defined as the ratio of the total number of dropped packets to M . The figure shows that PDR improves with a more stringent PER threshold, which is expected because the power allocated to each packet is inversely proportional to the PER threshold. Nevertheless, it can be noted that the PDR of the proposed and conventional ARQ converges to the same PDR for all the considered PER thresholds, except the 10^{-1} case. According to [11], the PDR required for various applications classified by ITU is 10^{-3} . Therefore, it can be noted from Fig. 6 that the proposed system can successfully achieve a PDR of 10^{-3} for all demonstrated thresholds at relatively low SNRs.

In ARQ systems, a packet may undergo multiple transmissions before it is successfully received. Consequently,

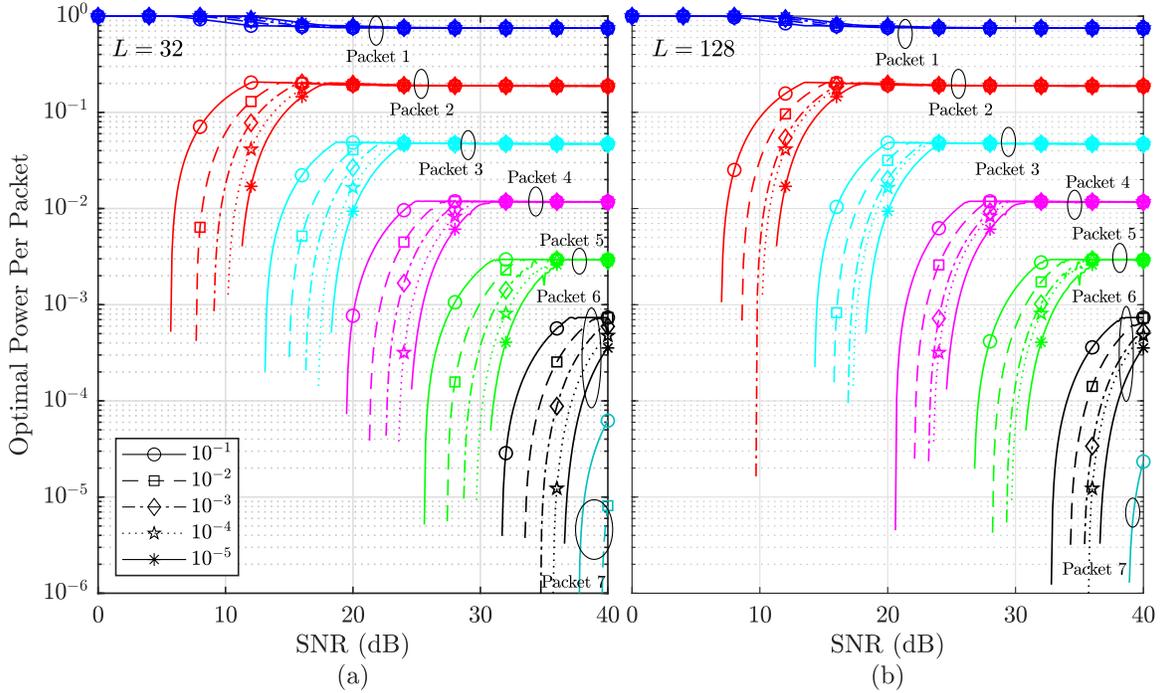


Fig. 3. Optimal power allocation per packet for NOM-ARQ packets for different PER thresholds in AWGN channel. (a) $L = 32$. (b) $L = 128$.

such packets will experience significant delay. Typically, more transmissions are performed in the low SNR region, whereas a single transmission is generally sufficient for a packet to be received successfully at high SNRs. The average number of transmission rounds for various PER thresholds is depicted in Fig. 7, where the average number of transmissions per packet is defined as the ratio of the total channel uses to M . As can be noted from the figure, the average number of transmission approaches \mathcal{M} at low SNRs. However, the results are significantly improved when SNR is increased. Another riveting result is that the ARQ-NOM delay may become less than one, which can never be achieved by conventional ARQ. Such a result is obtained because the total time required to transmit M packets can be much less than $M \times T_{PCK}$, where T_{PCK} is the time interval of one packet. Such performance can never be achieved with conventional ARQ because in the best-case scenario, one packet is transmitted in each T_{PCK} time interval.

The average consumed power per correctly received packet is shown in Fig. 8. The average consumed power is defined as the total power used to transmit M packets divided by the total number of correctly received packets. The power used to transmit each packet is normalized to unity. As can be noted from the figure, the average power per packet is about 6 dB. Such high power consumption is partially due to the retransmission process and since the dropped packets' power is added to the total consumed power. However, the power consumption decreases significantly by increasing SNR, where it becomes about -6 dBw at SNR of 30 dB. It is worth noting that Fig. 8 is generated with a fixed transmission power for the proposed and conventional ARQ systems. Although this might seem unfair because power adaptation can be also applied to

conventional ARQ, the efficiency of the power amplifier will be dropped significantly due to the small transmission power. Therefore, both systems consider a fixed transmission power equal to $p_{\max} = 1$.

Fig. 9 shows the throughput of the proposed NOM-ARQ for various N_{cap} , where $\tau = 10^{-1}$ and $L = 32$. Controlling the value of N_{cap} can be used to trade-off the spectral efficiency, or throughput, versus the system complexity and signalling overhead. More specifically, the value of N_{cap} affects the complexity of the power assignment and SIC processes. Therefore, lower N_{cap} values can be used for IoT nodes with limited computational capabilities, at the expense of spectral efficiency reduction. To provide a rough estimate of the complexity versus N_{cap} , Fig. 10 shows the average simulation time in seconds for Algorithm 1 while considering various SNRs. As the figure shows, the time complexity grows exponentially versus N_{cap} , and hence, it is necessary to optimize its value. Moreover, it can be noted that the time requirements are generally high due to power assignment through linear search methods, and the search step should be decreased significantly when N_{cap} is increased [26]. Moreover, the BER computational complexity also increases significantly versus N_{cap} [27]. Therefore, although the proposed heuristic algorithm is generally efficient, the complexity of the BER formulas remains a limiting issue. To resolve this matter, more efficient power searching methods can be used, and simpler BER formulas, possibly approximations, should be adopted.

VI. CONCLUSION

In this paper, a novel NOM-ARQ transmission scheme for high power and spectrum efficient communications is proposed and evaluated under various conditions. In this scheme, several

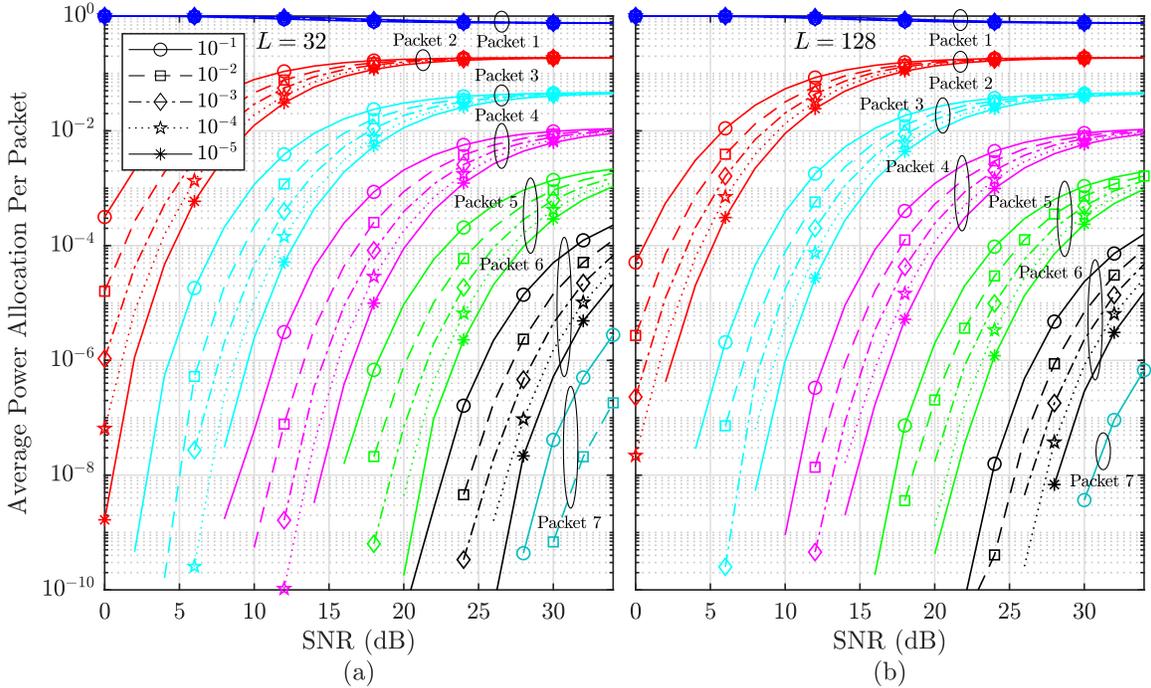


Fig. 4. Average power allocations per packet for NOM-ARQ using various PER thresholds in Rayleigh fading channels. (a) $L = 32$. (b) $L = 128$.

packets are adaptively multiplexed in the power domain where each multiplexed packet contains new and re-transmitted data symbols. To enable integration in IoT systems with limited processing capabilities the complexity of the proposed scheme can be manipulated by limiting the maximum number of multiplexed packets which results in a trade-off between system throughput and complexity. Nonetheless, it was demonstrated that at high SNRs, the proposed NOM-ARQ achieves significant spectral efficiency gain over the conventional ARQ. Similarly, the proposed NOM-ARQ was shown to outperform the conventional ARQ in terms of power efficiency and delay.

REFERENCES

- [1] A. Ahmed, M. Naem, and A. Al-Dweik, "Joint optimization of sensors association and UAVs placement in IoT applications with practical network constraints," *IEEE Access*, vol. 9, pp. 7674–7689, Jan. 2021.
- [2] N. Hossein Motlagh, T. Taleb, and O. Arouk, "Low-altitude unmanned aerial vehicles-based internet of things services: Comprehensive survey and future perspectives," *IEEE Internet Things J.*, vol. 3, no. 6, pp. 899–922, Dec. 2016.
- [3] S. Zhang, J. Liu, and W. Sun, "Stochastic geometric analysis of multiple unmanned aerial vehicle-assisted communications over internet of things," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 5446–5460, Jun. 2019.
- [4] J. Contreras-Castillo, S. Zeadally, and J. A. Guerrero-Ibañez, "Internet of vehicles: Architecture, protocols, and security," *IEEE Internet Things J.*, vol. 5, no. 5, pp. 3701–3709, Oct. 2018.
- [5] L.-L. Wang, J.-S. Gui, X.-H. Deng, F. Zeng, and Z.-F. Kuang, "Routing algorithm based on vehicle position analysis for internet of vehicles," *IEEE Internet Things J.*, vol. 7, no. 12, pp. 11701–11712, Dec. 2020.
- [6] Y. Ni, J. He, L. Cai, J. Pan, and Y. Bo, "Joint roadside unit deployment and service task assignment for internet of vehicles (IoV)," *IEEE Internet of Things J.*, vol. 6, no. 2, pp. 3271–3283, Apr. 2019.
- [7] K. Z. Ghafoor, L. Kong, D. B. Rawat, E. Hosseini, and A. S. Sadiq, "Quality of service aware routing protocol in software-defined internet of vehicles," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 2817–2828, Apr. 2019.

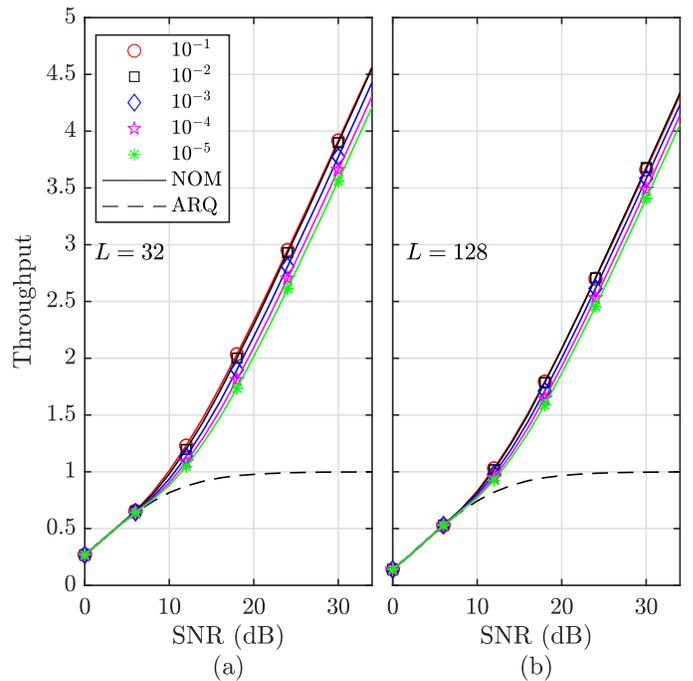


Fig. 5. Throughput for various PER thresholds. (a) $L = 32$. (b) $L = 128$.

- [8] A. J. Al-Dweik, M. Mayhew, R. Muresan, S. M. Ali, and A. Shami, "Using technology to make roads safer: Adaptive speed limits for an intelligent transportation system," *IEEE Veh. Technol. Mag.*, vol. 12, no. 1, pp. 39–47, Mar. 2017.
- [9] M. A. Al-Jarrah, A. Al-Dweik, M. Kalil, and S. S. Ikki, "Decision fusion in distributed cooperative wireless sensor networks," *IEEE Trans. Veh. Technol.*, vol. 68, no. 1, pp. 797–811, Jan. 2019.
- [10] M. Al-Jarrah *et al.*, "Decision fusion for IoT-based wireless sensor networks," *IEEE Internet Things J.*, vol. 7, no. 2, pp. 1313–1326, Feb. 2020.

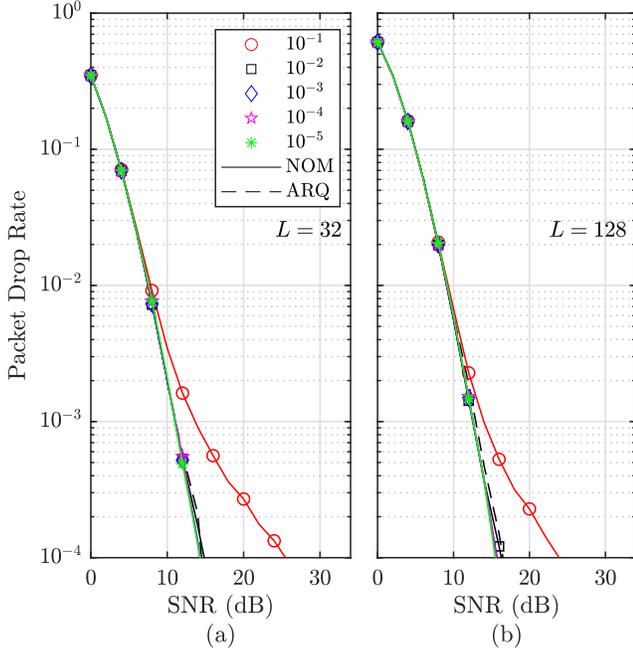


Fig. 6. Packet drop rate. (a) $L = 32$. (b) $L = 128$.

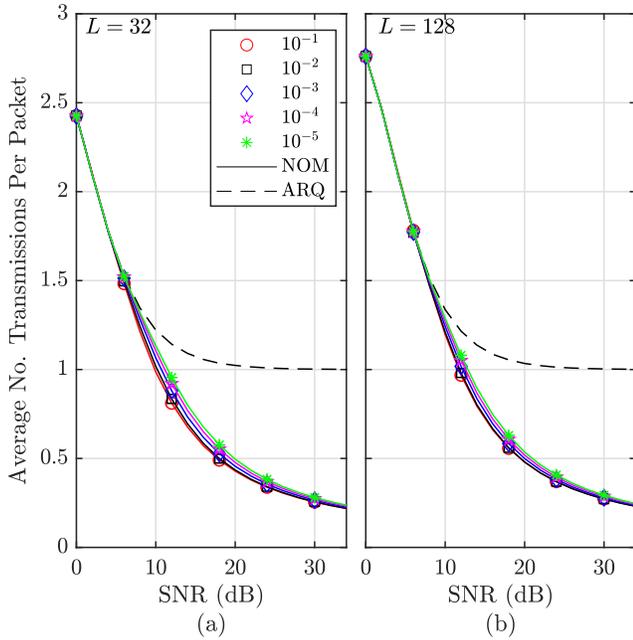


Fig. 7. Average number of transmission per packet. (a) $L = 32$. (b) $L = 128$.

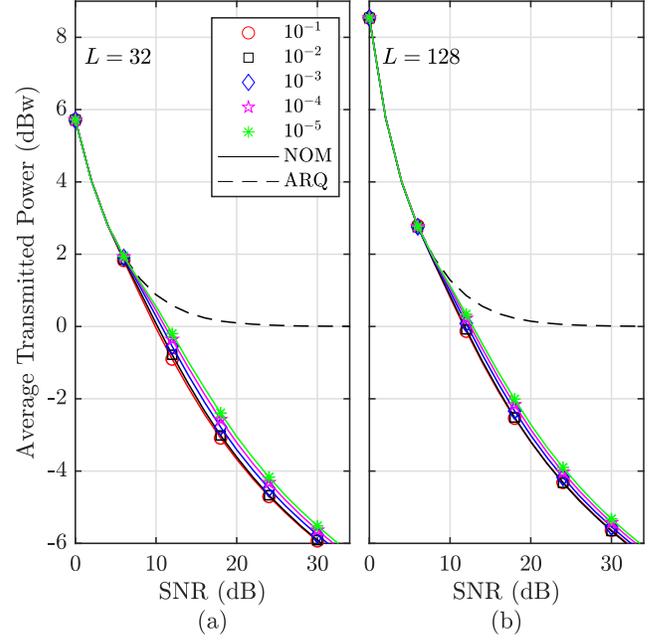


Fig. 8. Average consumed power per packet. (a) $L = 32$. (b) $L = 128$.

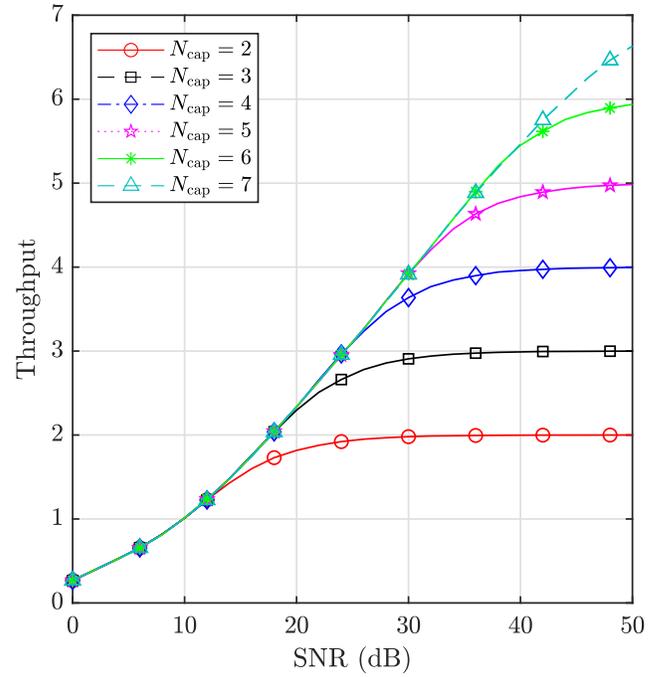


Fig. 9. Throughput of NOM-ARQ for different N_{cap} , where $\tau = 10^{-1}$ and $L = 32$.

[11] A. Ahmed *et al.*, "Hybrid automatic repeat request (HARQ) in wireless communications systems and standards: A contemporary survey," *IEEE Commun. Surveys Tuts.*, pp. 1–1, 2021.

[12] "3GPP 38 series of specifications: 5G New Radio," <https://www.3gpp.org/DynaReport/38-series.htm>, accessed: 2020-08-18.

[13] D. Cai, Z. Ding, P. Fan, and Z. Yang, "On the performance of NOMA with hybrid ARQ," *IEEE Trans. Veh. Technol.*, vol. 67, no. 10, pp. 10033–10038, Oct. 2018.

[14] Z. Yu, C. Zhai, W. Ni, and D. Wang, "Non-orthogonal multiple access with cooperative truncated arq and relay selection," *IEEE Access*, vol. 7, pp. 56 228–56 243, Apr. 2019.

[15] F. Cao, Y. Song, and Y. Yang, "ARQ assisted short-packet communications for NOMA networks over Nakagami- m fading channels," *IEEE Access*, vol. 8, pp. 158 263–158 272, Aug. 2020.

[16] S. Park and S. Choi, "Performance of symbol-level combining and bit-level combining in MIMO multiple ARQ systems," *IEEE Trans. Commun.*, vol. 64, no. 4, pp. 1517–1528, Apr. 2016.

[17] R. Mai, T. V. K. Chaitanya, and T. Le-Ngoc, "Progressive hybrid precoding and combining for massive MIMO ARQ systems," *IEEE Access*, vol. 6, pp. 34 503–34 515, Jun. 2018.

[18] A. U. Rehman, C. Dong, L. L. Yang, and L. Hanzo, "Performance of cognitive stop-and-wait hybrid automatic repeat request in the face of imperfect sensing," *IEEE Access*, vol. 4, pp. 5489–5508, Jul. 2016.

[19] H. Mukhtar, A. Al-Dweik, M. Al-Mualla, and A. Shami, "Low complexity power optimization algorithm for multimedia transmission over wireless networks," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 1,

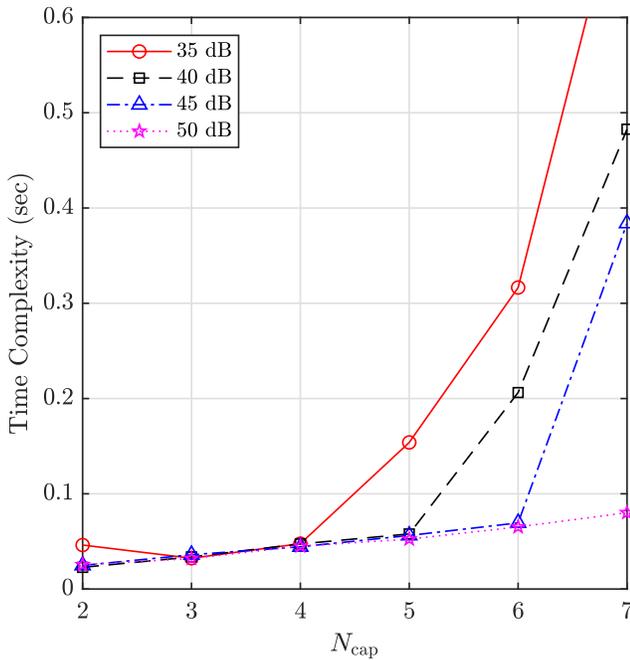


Fig. 10. Algorithm time complexity, in seconds, versus N_{cap} for various SNRs, where $L = 32$.

pp. 113–124, Feb. 2015.

- [20] H. Mukhtar, A. Al-Dweik, M. Al-Mualla, and A. Shami, “Adaptive hybrid ARQ system using turbo product codes with hard/soft decoding,” *IEEE Commun. Lett.*, vol. 17, no. 11, pp. 2132–2135, Nov. 2013.
- [21] K. Vakiliinia, S. V. S. Ranganathan, D. Divsalar, and R. D. Wesel, “Optimizing transmission lengths for limited feedback with nonbinary LDPC examples,” *IEEE Trans. Commun.*, vol. 64, no. 6, pp. 2245–2257, Jun. 2016.
- [22] K. Miura, Y. Kishiyama, and K. Higuchi, “Low latency HARQ method using early retransmission before channel decoding based on superposition coding,” in *Int. Symp. Intelligent Signal Process. Commun. Syst. (ISPACS)*, Taipei, Taiwan, Feb. 2019, pp. 1–2.
- [23] A. Khreis, F. Bassi, P. Ciblat, and P. Duhamel, “Multi-layer HARQ with delayed feedback,” *IEEE Trans. Wireless Commun.*, vol. 19, no. 9, pp. 6224–6237, Sep. 2020.
- [24] A. Al-Dweik and Y. Iraqi, “High throughput wireless links for time-sensitive WSNs with reliable data requirements,” *IEEE Sensors J.*, Sep. 2021. [Online]. Available: https://www.techrxiv.org/articles/preprint/ARQ_The_Gateway_from_NOMA_to_NOM/13157714.
- [25] A. Høglund, X. Lin, O. Liberg, A. Behravan, E. A. Yavuz, M. Van Der Zee, Y. Sui, T. Tirronen, A. Ratilainen, and D. Eriksson, “Overview of 3GPP Release 14 enhanced NB-IoT,” *IEEE Network*, vol. 31, no. 6, pp. 16–22, 2017.
- [26] Y. Iraqi and A. Al-Dweik, “Power allocation for reliable SIC detection of rectangular QAM-based NOMA systems,” *IEEE Trans. Veh. Technol.*, vol. 70, no. 8, pp. 8355–8360, Aug. 2021.
- [27] H. Yahya, E. Alsusa, and A. Al-Dweik, “Exact BER analysis of NOMA with arbitrary number of users and modulation orders,” *IEEE Trans. Commun.*, pp. 1–1, 2021.