Harvesting Devices' Heterogeneous Energy Profiles and QoS Requirements in IoT: WPT-NOMA vs BAC-NOMA

Zhiguo Ding, *Fellow, IEEE*

***Abstract*— The next generation Internet of Things (IoT) exhibits a unique feature that IoT devices have different energy profiles and quality of service (QoS) requirements. In this paper, two energy and spectrally efficient transmission strategies, namely wireless power transfer assisted non-orthogonal multiple access (WPT-NOMA) and backscatter communication (Back-Com) assisted NOMA (BAC-NOMA), are proposed by utilizing this feature of IoT and employing spectrum and energy cooperation among the devices. In particular, the use of NOMA ensures that the devices with different QoS requirements can share the same spectrum, and WPT and Back-Com are employed to utilize the cooperation among the devices with different energy profiles, which avoids the use of a dedicated power beacon. Furthermore, for the proposed WPT-NOMA scheme, the application of hybrid successive interference cancelation (SIC) decoding order is also considered, and analytical results are developed to demonstrate that WPT-NOMA can avoid outage probability error floors and realize the full diversity gain. Unlike WPT-NOMA, BAC-NOMA suffers from an outage probability error floor, and the asymptotic behavior of this error floor is analyzed in the paper by applying the extreme value theory. In addition, the effect of a unique feature of BAC-NOMA, i.e., employing one device's signal as the carrier signal for another device, is studied, and its impact on the diversity gain is revealed. Simulation results are also provided to compare the performance of the proposed strategies and verify the developed analytical results.**

***Index Terms*— Non-orthogonal multiple access (NOMA), simultaneous wireless information and power transfer (SWIPT), backscatter communications (BackCom).**

I. Introduction

A. Motivations and the Related Existing Works

The next generation Internet of Things (IoT) is envisioned to support various important applications, including smart home, intelligent transportation, wireless health-care, environment monitoring, etc [1]. The key step to implement the IoT is to ensure that a massive number of IoT devices with heterogeneous energy profiles and quality of service (QoS) requirements can be connected in a spectrally efficient manner, which results in the two following challenges. From the spectral efficiency perspective, it is challenging to support massive connectivity, given the scarce bandwidth resources available for wireless communications. Non-orthogonal multiple access (NOMA) has been recognized as a spectrally efficient solution to support massive connectivity by encouraging spectrum sharing among wireless devices with different QoS requirements [2]–[4]. For example, in conventional orthogonal multiple access (OMA), a delay-sensitive IoT device is allowed to solely occupy a bandwidth resource block, which is not helpful to support massive connectivity and can also result in low spectral efficiency, particularly if this device has a small amount of data to send. By using NOMA, additional users, such as delay-tolerate devices, can be admitted to the channel. As a result, the overall spectral efficiency is improved, and the use of advanced forms of NOMA can ensure that massive connectivity is supported while strictly guaranteeing all devices' QoS requirements [5]–[7]. It is worth to point out that successive interference cancellation (SIC) is a key component of NOMA systems, and conventionally SIC decoding orders are mainly determined by the devices' channel state information (CSI). Recent studies show that the spectral efficiency and reception reliability of uplink NOMA can be significantly improved by using hybrid SIC decoding order which is to dynamically decide the decoding order by simultaneously using the devices' CSI as well as their QoS requirements [8]–[10].

From the energy perspective, the challenge is due to the fact that some IoT devices might be equipped with continuous power supplies, but there are many other devices which are battery powered and hence severely energy constrained. This challenge motivates the use of two techniques, wireless power transfer (WPT) and backscatter communication (BackCom). The key idea of WPT is to use radio frequency (RF) signals for energy transfer. In particular, an energy-constrained IoT device can first carry out energy harvesting by using the RF signals sent by a power station or another non-energy-constrained node in the wireless network, where the harvested energy can be used to power the transmission of the energy-constrained device [11]–[14]. Similar to WPT, BackCom is another low-power and low-complexity technique to connect energy-constrained devices [15]–[17]. The key idea of BackCom is to ask an energy-constrained IoT device, termed a tag, to carry out passive reflection and modulation of a single-tone sinusoidal continuous wave sent by a BackCom reader. Instead of relying on the continuous wave sent by a reader, a variation

Manuscript received July 27, 2020; revised October 16, 2020 and December 11, 2020; accepted January 14, 2021. Date of publication January 19, 2021; date of current version May 18, 2021. This work was supported by the Engineering and Physical Sciences Research Council (EPSRC) under Grant EP/P009719/2. The associate editor coordinating the review of this article and approving it for publication was N. Tran.

The author is with the School of Electrical and Electronic Engineering, The University of Manchester, Manchester M14 9PL, U.K. (e-mail: zhiguo.ding@manchester.ac.uk).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TCOMM.2021.3052948.

Digital Object Identifier 10.1109/TCOMM.2021.3052948

of BackCom, termed symbiotic radio, was recently proposed to use the information-bearing signal sent by a non-energy-constrained device to power a batteryless device [18], [19].

In order to simultaneously address the aforementioned spectral and energy challenges, it is natural to consider the combination of NOMA with the two energy-cooperation transmission techniques in the next generation IoT, which will be the focus of this paper. Early examples of WPT assisted NOMA (WPT-NOMA) have considered the cooperative communication scenario, where relay transmission is powered by the energy harvested from the signals sent by a source [20]–[22]. In downlink scenarios, the use of WPT-NOMA can yield a significant improvement in the spectral and energy efficiency as demonstrated by [23]–[25]. The application of WPT to uplink NOMA has been previously studied in [26], where users use the energy harvested from the signal sent by the base station to power their uplink NOMA transmission. Compared to WPT-NOMA, the application of BackCom to NOMA received less attention. In [27] and [28], NOMA was used to ensure that multiple backscatter devices can communicate with the same access point (a reader) simultaneously by modulating the continuous wave sent by the access point. More recently, the application of NOMA to a special case of BackCom, symbiotic radio, has been considered in [29], [30].

B. Aim and Contributions of This Work

The aim of this paper is to consider a NOMA uplink scenario, where a delay-sensitive non-energy-constrained IoT device and multiple delay-tolerant energy-constrained devices communicate with the same access point, and the use of NOMA ensures that the devices with different QoS requirements can share the same spectrum. In particular, following the semi-grant-free protocol proposed in [5] and [8], one of the delay-tolerant devices is granted access to the channel which would be solely occupied by the delay-sensitive device in OMA. Because some IoT devices are energy constrained, WPT and Back-Com are employed to utilize the cooperation among the devices with different energy profiles, which avoids the use of a dedicated power beacon. The contributions of the paper are listed as follows:

- A new WPT-NOMA scheme is proposed by applying hybrid SIC decoding order, where the transmission of an energy-constrained device is powered by the energy harvested from the signals sent by the non-energy-constrained device. An intermediate benefit for using hybrid SIC decoding order for NOMA uplink is to avoid an outage probability error floor, which is not possible if a fixed SIC decoding order is used. In this paper, the outage performance of WPT-NOMA with hybrid SIC decoding order is analyzed, and the obtained analytical results demonstrate that outage probability error floors can be avoided and the full diversity gain is still achievable.
- A general multi-user BAC-NOMA scheme is proposed, where an energy-constrained device reflects and modulates the signals sent by the non-energy-constrained device. Note that the BAC-NOMA scheme considered in [30] can be viewed as a special case of this general framework. In addition, the two key features of BAC-NOMA are analyzed in detail. Firstly, we focus on the outage probability error floor suffered by BAC-NOMA, by applying the extreme value theory (EVT) [31], [32].[1] Secondly, we focus on another feature of BAC-NOMA, i.e., modulating the energy-constrained device's signal on the non-energy-constrained device's signal. This feature means that the relationship between the two devices' signals is multiplicative, instead of additive. Or in other words, the non-energy-constrained device's signal can be viewed as a type of fast fading for the energy-constrained device. The analytical results developed in the paper show that this virtual fading is damaging to the reception reliability, and the diversity gain achieved by BAC-NOMA is capped by one.

The remainder of the paper is organized as follows. In Section II, the two proposed NOMA schemes are described and their achievable data rates are illustrated. In Sections III and IV, the performance of the proposed WPT-NOMA and BAC-NOMA schemes is analyzed, respectively. In Section V, computer simulations are presented, and Section VI concludes the paper. The details of all proofs are collected in the appendix.

II. System Model

Consider a NOMA uplink scenario with one access point and $(M+1)$ IoT devices, denoted by $\mathrm{U_m}$, $0 \leq m \leq M$. For illustration purposes, assume that $\mathrm{U_0}$ is a non-energy-constrained delay-sensitive device, whereas $\mathrm{U_m}$, $1 \leq m \leq M$, are energy constrained and delay tolerant. The channel from the access point to $\mathrm{U_i}$ is denoted by h_i, $0 \leq i \leq M$. The channel from $\mathrm{U_0}$ to $\mathrm{U_m}$ is denoted by g_m, $1 \leq m \leq M$.

Because $\mathrm{U_0}$ is delay sensitive, it is allowed to solely occupy a bandwidth resource block in OMA, which is spectrally inefficient for supporting massive connectivity. Following the designs shown in [5] and [8], we consider that one of the delay-tolerant IoT devices is to be granted access to the resource block which would be solely occupied by $\mathrm{U_0}$ in OMA.

Assumption: To facilitate performance analysis, we assume that the energy-constrained devices are located in a small-size cluster, such that that the distances between $\mathrm{U_0}$ and U_m, $m \geq 1$, are same. A similar assumption is also made to the distances between the access point and the devices. For example, the devices can be sensors in a self-driving vehicle or on an autonomous robot. For smart home applications, the devices can be sensors for different functionalities fixed in the same room. Therefore, we assume that g_m, $1 \leq m \leq M$, are modeled as independent and identically distributed (i.i.d.) Rayleigh fading, i.e., complex Gaussian distributed with zero mean and variance λ_g, $g_m \sim CN(0, \lambda_g)$, where $\lambda_g \triangleq d_g^{\phi}$, d_g denotes the distance between $\mathrm{U_0}$ and U_m, $m \geq 1$, and ϕ denotes the path loss exponent. Similarly, we also assume that $h_m \sim CN(0, \lambda_h)$ and $h_0 \sim CN(0, \lambda_0)$, where $\lambda_h \triangleq d_h^{\phi}$, $\lambda_0 \triangleq d_0^{\phi}$, d_h denotes the distance between the access point

[1] EVT can be viewed as a special case of order statistics and is to characterize the distribution of ordered random variables, when the number of random variables becomes infinity [33], [34].

and U_m, $m \geq 1$, and d_0 denotes the distance between U_0 and the access point.

A. WPT Assisted NOMA

Without loss of generality, assume that U_m is granted access, where the details for the scheduling strategy will be provided at the end of this subsection. Suppose that the energy-constrained devices can support WPT, and time-switching WPT is used for its simplicity, which consists of two phases [35]. During the first αT seconds, U_m performs energy harvesting by using U_0's signal, denoted by s_0, and then uses the harvested energy for its transmit power to send its signal s_m to the access point, where α denotes the time-switching parameter, $0 \leq \alpha \leq 1$ and T denotes the block period. Therefore, the amount of energy harvested at U_m is $\eta P|g_m|^2\alpha T$, where P denotes U_0's transmit power, η denotes the energy harvesting efficiency coefficient. This means that the observation at the access point is given by

$$y_{\mathrm{AP}} = \sqrt{P}h_0 s_0 + \sqrt{\frac{\eta P|g_m|^2\alpha}{1-\alpha}}h_m s_m + n_{\mathrm{AP}}, \tag{1}$$

where n_{AP} denotes the noise.

For the proposed WPT-NOMA scheme, hybrid SIC decoding order is applied [9], [10]. In particular, if s_m is decoded first, U_m's maximal data rate without causing the failure of SIC (or degrading U_0's performance) is given by

$$\mathrm{R}_m^{WP,1} = (1-\alpha)\log\left(1 + \frac{\eta P\bar{\alpha}|g_m|^2|h_m|^2}{P|h_0|^2+1}\right), \tag{2}$$

where $\bar{\alpha} = \frac{\alpha}{1-\alpha}$ and the noise power is assumed to be normalized. If U_0's signal is decoded first, U_0's achievable data rate is given by

$$\mathrm{R}_{0,m}^{WP,2} = (1-\alpha)\log\left(1 + \frac{P|h_0|^2}{\eta P\bar{\alpha}|g_m|^2|h_m|^2+1}\right). \tag{3}$$

Denote U_0's target data rate by R_0. If $\mathrm{R}_{0,m}^{WP,2} \geq R_0$, s_0 can be successfully decoded and removed, which means that s_m can be decoded correctly with the following data rate:

$$\mathrm{R}_m^{WP,2} = (1-\alpha)\log\left(1 + \eta P\bar{\alpha}|g_m|^2|h_m|^2\right). \tag{4}$$

Device Scheduling for WPT-NOMA: The aim of device scheduling is to ensure that the delay-tolerant device which yields the largest data rate can be selected, under the condition that U_0's QoS requirements are strictly guaranteed. Note that $\mathrm{R}_{0,m}^{WP,2} \geq R_0$ is equivalent to the following inequality:

$$\gamma_m \leq \frac{|h_0|^2}{\bar{\epsilon}_0\eta\bar{\alpha}} - \frac{1}{\eta P\bar{\alpha}}, \tag{5}$$

where $\bar{\epsilon}_0 = 2^{\frac{R_0}{1-\alpha}} - 1$. Furthermore, define $\bar{\epsilon}_s = 2^{\frac{R_s}{1-\alpha}} - 1$ and $\tau(h_0) = \max\left\{0, \frac{|h_0|^2}{\bar{\epsilon}_0\eta\bar{\alpha}} - \frac{1}{\eta P\bar{\alpha}}\right\}$, where it is assumed that U_m, $1 \leq m \leq M$, have the same target data rate, denoted by R_s. The delay-tolerant IoT devices can be divided into the two groups, denoted by $\mathcal{S}_1$ and $\mathcal{S}_2$, respectively, as defined in the following:

- $\mathcal{S}_1$ contains the devices whose channel gains satisfy $\gamma_m > \tau(h_0)$. If one device from $\mathcal{S}_1$ is scheduled, its signal has to be decoded at the first stage of SIC, which yields the data rate $\mathrm{R}_m^{WP,1}$.
- $\mathcal{S}_2$ contains the devices whose channel gains satisfy $\gamma_m \leq \tau(h_0)$. If one device from $\mathcal{S}_2$ is scheduled, its signal can be decoded either at the first stage of SIC (which yields the data rate $\mathrm{R}_m^{WP,1}$) or at the second stage of SIC (which yields the data rate $\mathrm{R}_m^{WP,2}$). Since $\mathrm{R}_m^{WP,1} \leq \mathrm{R}_m^{WP,2}$ always holds, U_m always prefers its signal to be decoded at the second stage of SIC.

The access point selects the delay-tolerant device which yields the largest data rate, i.e.,

$$m^* = \arg\max\left\{\max\left\{\mathrm{R}_m^{WP,1}, m\in\mathcal{S}_1\right\}, \max\left\{\mathrm{R}_m^{WP,2}, m\in\mathcal{S}_2\right\}\right\}. \tag{6}$$

Remark 1: As can be observed from (2), (3), and (4), the use of time-switching reduces the time duration for data transmission, since the first αT seconds are used for energy harvesting. This feature of WPT-NOMA can lead to a potential performance loss compared BAC-NOMA which can support continuous data transmission. It is worth to point out that massive connectivity can be still supported, even though a single IoT device is scheduled at each time. For example, assume that U_0 is scheduled for nT time slots, where n denotes the number of time slots. During each time slot, one IoT device is served and then can be removed from the competing device set ($\mathcal{S}_1 \bigcup \mathcal{S}_2$) due to the short packet transmission feature of IoT, which means that n devices can be served without consuming any extra bandwidth resources.

B. BackCom-Assisted NOMA

Again assume that U_m is granted access, where the details for the BAC-NOMA scheduling strategy will be provided later. Suppose that the energy-constrained devices are capable to carry out backscatter communications. Different from WPT-NOMA, energy harvesting is not performed by BAC-NOMA. Instead, $\mathrm{U_m}$ modulates and reflects the signals sent by U_0. Therefore, the access point receives the following signal[2]:

$$y_{\mathrm{AP}} = \sqrt{P}h_0 s_0 + \sqrt{P}\beta g_m h_m s_0 s_m + n_{\mathrm{AP}}, \tag{7}$$

where β denotes the BackCom power reflection coefficient. Because $\mathrm{U_m}$'s signals are modulated on U_0's signal, the detection strategy of BAC-NOMA is different from that of WPT-NOMA. In particular, for BAC-NOMA, there is only one choice for the SIC decoding order, which is to decode U_0's signal first. The reason for this is that U_0's signal can be viewed as a fading channel for U_m's signal. In order to implement coherent detection, U_0's signal, i.e., the virtual fading channel, needs to be decoded first.

We note that the addressed system can be interpreted as a conventional communication scenario with imperfect channel state information, where $\beta g_m h_m s_m$ can be treated as a channel

[2]We assume that the symbol periods of different devices are same, where the design of BAC-NOMA for the case with devices using different symbol periods is beyond the scope of this paper.

estimation error. As discussed in [36], [37], treating the channel estimation error as a noise yields a worst-case scenario, and the resulting data rate can be used as an insightful lower bound on the channel capacity in practice. Therefore, in BAC-NOMA, U_0's achievable data rate is given by

$$\mathrm{R}_{0,m}^{BAC} = \log\left(1 + \frac{P|h_0|^2}{P\beta^2|g_m|^2|h_m|^2 + 1}\right). \tag{8}$$

Assuming that U_0's signal can be correctly decoded, i.e., $\mathrm{R}_{0,m}^{BAC} \geq R_0$, U_0's signal can be removed, which leads to the following system model:

$$y_{\mathrm{AP}} - \sqrt{P}h_0 s_0 = \sqrt{P}\beta g_m h_m s_0 s_m + n_{\mathrm{AP}}. \tag{9}$$

Therefore, an achievable data rate for decoding s_m is given by

$$\mathrm{R}_{m}^{BAC} = \log\left(1 + P\beta^2|g_m|^2|h_m|^2|s_0|^2\right), \tag{10}$$

where U_0's signal, s_0, is viewed as a fast fading channel gain for s_m. Similar to [18], [30], it is assumed that $s_m \sim CN(0,1)$, i.e., the probability density function (pdf) of this virtual fading channel, $|s_0|^2$, is $f_{|s_0|^2}(x) = e^{-x}$.

Device Scheduling for BAC-NOMA: In OMA, U_0 is allowed to solely occupy the channel, whereas the use of NOMA ensures that the backscatter devices can also be granted access. In order to guarantee U_0's QoS requirements, device U_m can be granted access only if $\mathrm{R}_{0,m}^{BAC} \geq R_0$ which can be rewritten as follows:

$$|g_m|^2|h_m|^2 \leq \beta^{-2}\epsilon_0^{-1}|h_0|^2 - \beta^{-2}P^{-1} \tag{11}$$

where $\epsilon_0 = 2^{R_0} - 1$.

On the other hand, it is ideal to admit the device which can maximize the data rate R_m^{BAC}. Therefore, the device scheduling criterion is given by

$$m^* = \arg\max\left\{\mathrm{R}_m^{BAC}, m \in \mathcal{S}_0\right\}, \tag{12}$$

where $\mathcal{S}_0 = \left\{m : \mathrm{R}_{0,m}^{BAC} \geq R_0, 1 \leq m \leq M\right\}$.

Remark 2: Unlike WPT-NOMA, BAC-NOMA can support one SIC decoding order only, which is the reason why it suffers an outage probability error floor, as shown in the next section. Another feature of BAC-NOMA is that s_0 is treated as a virtual fading channel, which means s_m suffers additional fading attenuation. The impact of this virtual fading channel on the reception reliability of s_m will be investigated in the following section.

Remark 3: We note that the two proposed device scheduling strategies can be carried out in a distributed manner. Take BAC-NOMA as an example. Each backscatter device decides to participate in contention, if $\mathrm{R}_{0,m}^{BAC} > R_0, m \in \mathcal{S}$, otherwise it switches to the match state. Each device calculates its backoff time inversely proportionally to its achievable data rate R_m^{BAC}, which ensures that U_{m^*} can be granted access in a distributed manner.

III. Performance Analysis for WPT-NOMA

Since the implementation of WPT-NOMA is transparent to U_0, we only focus on the performance of the admitted delay-tolerant energy-constrained device. Denote the effective channel gains of the devices by $\gamma_m = |g_m|^2|h_m|^2$. In order to simplify notations, without of loss of generality, assume that the delay-tolerant devices are ordered according to their effective channel gain as follows:

$$\gamma_1 \leq \cdots \leq \gamma_M. \tag{13}$$

With this channel ordering, the impact of device scheduling on the NOMA transmission can be shown explicitly. Particularly, denote $\bar{E}_m$ by the event that the size of $\mathcal{S}_2$ is m, i.e., $\bar{E}_m$ can be expressed as follows:

$$\bar{E}_m = \{\gamma_m < \tau(h_0), \gamma_{m+1} > \tau(h_0)\}, \tag{14}$$

for $1 \leq m \leq M-1$, where $\bar{E}_0 = \{\gamma_1 > \tau(h_0)\}$ and $\bar{E}_M = \{\gamma_M < \tau(h_0)\}$.

The outage probability achieved by WPT-NOMA can be expressed as follows:

$$\mathrm{P}^{WP} = \sum_{m=1}^{M} \underbrace{\mathrm{P}\left(\max\left\{\mathrm{R}_m^{WP,2}, \mathrm{R}_M^{WP,1}\right\} < R_s, |\mathcal{S}_2| = m\right)}_{T_m} + \underbrace{\mathrm{P}\left(\mathrm{R}_M^{WP,1} < R_s, |\mathcal{S}_2| = 0\right)}_{T_0}. \tag{15}$$

We note that the performance analysis requires the pdf and cumulative distribution function (CDF) of the ordered channel gain γ_m, which can be found by using the density functions of the unordered channel gain. In particular, the pdf of the unordered effective channel gain is given by [30]

$$f_\gamma(x) = 2\lambda_h\lambda_g K_0\left(2\sqrt{\lambda_h\lambda_g x}\right), \tag{16}$$

where $K_i(\cdot)$ denotes the i^{th}-order modified Bessel function of the second kind. The CDF of the unordered channel gain, denoted by $F_\gamma(x)$, can be obtained as follows:

$$\begin{aligned} F_\gamma(x) &= \int_0^x 2\lambda_h\lambda_g K_0\left(2\sqrt{\lambda_h\lambda_g y}\right)dy \\ &= \frac{4}{\lambda_h\lambda_g}x\int_0^1 K_0\left(2t\sqrt{x}\sqrt{\frac{1}{\lambda_h\lambda_g}}\right)tdt \\ &= 1 - 2\sqrt{\lambda_h\lambda_g x}K_1\left(2\sqrt{\lambda_h\lambda_g x}\right), \end{aligned} \tag{17}$$

where [38, (6.561.8)] is used. As can be observed from (16) and (17), the density functions of the unordered channel gains contain Bessel functions, which makes it difficult to obtain an exact expression for the outage probability achieved by WPT-NOMA. However, the diversity gain achieved by WPT-NOMA can be obtained, as shown in the following theorem.

Theorem 1: For the considered NOMA uplink scenario, WPT-NOMA can realize a diversity gain of M, if $\bar{\epsilon}_0\bar{\epsilon}_s < 1$.

Proof: See Appendix A. □

Remark 3: Theorem 1 shows that the diversity gain achieved by WPT-NOMA is not zero, which implies that WPT-NOMA does not suffer any outage probability error floors, a feature not achievable to BAC-NOMA, as shown in the next section. Therefore, WPT-NOMA is a more robust transmission solution, compared to BAC-NOMA, particularly at high SNR.

Remark 4: Note that M is the maximal multi-user diversity gain achievable to the considered NOMA uplink scenario, since there are M delay-tolerant devices competing for the access. Theorem 1 shows that the maximal diversity gain can be realized by WPT-NOMA, even though battery-less transmission is used. Therefore, WPT-NOMA is particularly attractive for energy-constrained IoT devices which have strict requirements for reception reliability.

Remark 5: We note that the conclusion that there is no outage probability error floor also holds for the special case $M = 1$, i.e., there is a single delay-tolerant device and device scheduling is not carried out. This implies that the outage probability error floor is avoided due to the use of hybrid SIC decoding order, instead of device scheduling

IV. Performance Analysis for BAC-NOMA

Again because the implementation of NOMA is transparent to U_0, we only focus on the performance of the admitted delay-tolerant device. The outage probability of interest is expressed as follows:

$$\mathrm{P}^{BAC} = \mathrm{P}\left(\mathrm{R}_{m^*}^{BAC} < R_s, |\mathcal{S}_0| \neq 0\right) + \mathrm{P}\left(|\mathcal{S}_0| = 0\right), \quad (18)$$

where $|\mathcal{S}|$ denotes the size of set $\mathcal{S}$.

Assume that the devices' channel gains are ordered as in (13). Denote E_m by the event that the size of $\mathcal{S}_0$ is m, i.e., E_m can be expressed as follows:

$$E_m = \{\gamma_m < \theta(h_0), \gamma_{m+1} > \theta(h_0)\}, \quad (19)$$

for $1 \leq m \leq M-1$, where $\theta(h_0) = \beta^{-2}\epsilon_0^{-1}|h_0|^2 - \beta^{-2}P^{-1}$. We note that $E_0 = \{\gamma_1 > \theta(h_0)\}$ and $E_M = \{\gamma_M < \theta(h_0)\}$.

The use of (12) and (13) means that U_m will be granted access, for the event E_m. Therefore, the outage probability can be further written as follows:

$$\mathrm{P}^{BAC} = \sum_{m=1}^{M} \underbrace{\mathrm{P}\left(\mathrm{R}_m^{BAC} < R_s, E_m\right)}_{Q_m} + \mathrm{P}\left(E_0\right). \quad (20)$$

We note that P^{BAC} is more challenging to analyze, compared to P^{WP}, because there are more random variables involved. In the following, we focused on two key features of WPT-NOMA.

A. Outage Probability Error Floor

In this subsection, we will show that BAC-NOMA suffers from an outage probability error floor. The existence of the error floor can be sufficiently proved by focusing on a lower bound on the outage probability as shown in the following:

$$\mathrm{P}^{BAC} \geq \mathrm{P}\left(E_0\right). \quad (21)$$

The simulation results provided in Section V show that E_0 is indeed the most damaging event at high SNR, compared to the terms Q_m, $1 \leq m \leq M$. $\mathrm{P}(E_0)$ can be expressed as follows:

$$\begin{aligned}\mathrm{P}\left(E_0\right) &= \mathrm{P}\left(\gamma_1 > \beta^{-2}\epsilon_0^{-1}|h_0|^2 - \beta^{-2}P^{-1}\right)\\ &= \mathrm{P}\left(\beta^2\epsilon_0\gamma_1 + \epsilon_0 P^{-1} > |h_0|^2 > \epsilon_0 P^{-1}\right)\\ &\quad + \mathrm{P}\left(|h_0|^2 < \epsilon_0 P^{-1}\right). \quad (22)\end{aligned}$$

Denote $f_{\gamma_1}(x) \triangleq M f_\gamma(x)\left(1 - F_\gamma(x)\right)^{M-1}$ by the marginal pdf of the smallest order statistics, and hence P_{E_0} can be expressed as follows:

$$\begin{aligned}\mathrm{P}\left(E_0\right) &= \int_0^\infty \left(e^{-\lambda_0\epsilon_0 P^{-1}} - e^{-\lambda_0(\beta^2\epsilon_0 x + \epsilon_0 P^{-1})}\right) f_{\gamma_1}(x)dx\\ &\quad + 1 - e^{-\lambda_0\epsilon_0 P^{-1}}\\ &= 1 - M e^{-\lambda_0\epsilon_0 P^{-1}} \int_0^\infty e^{-\lambda_0\beta^2\epsilon_0 x} f_\gamma(x)\\ &\quad \times \left(1 - F_\gamma(x)\right)^{M-1} dx. \quad (23)\end{aligned}$$

At high SNR, i.e., $P \to \infty$, $\mathrm{P}(E_0)$ can be approximated as follows:

$$\begin{aligned}\mathrm{P}\left(E_0\right) &\approx 1 - M \int_0^\infty e^{-\lambda_0\beta^2\epsilon_0 x} f_\gamma(x)\left(1 - F_\gamma(x)\right)^{M-1} dx\\ &\approx \lambda_0\beta^2\epsilon_0 \int_0^\infty e^{-\lambda_0\beta^2\epsilon_0 x}\left(1 - F_\gamma(x)\right)^M dx, \quad (24)\end{aligned}$$

which is constant and not a function of P. Combining (21) with (24), it is sufficient to conclude that BAC-NOMA transmission suffers an outage probability error floor.

Remark 6: This finding is consistent to the conclusions made in [30]. The reason for the existence of this error floor is due to the fact that only one SIC decoding order can be used by BAC-NOMA. Compared to BAC-NOMA, WPT-NOMA can avoid this error floor and hence outperform BAC-NOMA at high SNR.

Remark 7: Theorem 1 indicates that WPT-NOMA can utilize the multi-user diversity, and hence a nature question is whether BAC-NOMA can also use the multi-user diversity, i.e., whether it is beneficial to invite more delay-tolerant devices to participate in transmission in BAC-NOMA. By applying the EVT, the following lemma can be obtained for this purpose.

Lemma 1: The error floor caused by $\mathrm{P}(E_0)$ *can be reduced to zero by increasing the number of participating delay-tolerant devices* M *and the transmit power* P.

Proof: See Appendix B. □

B. Impact of s_0 on Reception Reliability

Recall that s_0 is treated as a type of fast fading when the signal from the delay-tolerant device is decoded. In this section, we will show that this fast fading has a harmful impact on the outage probability. To obtain an insightful conclusion, we consider an ideal situation, in which E_0 does not happen. We will show that even in such an ideal situation, the full multi-user diversity gain cannot be realized. Recall the term Q_m, $1 \leq m \leq M-1$, shown in (20) can be evaluated as follows:

$$\begin{aligned}Q_m &= \mathrm{P}\left(\mathrm{R}_m^{BAC} < R_s, \gamma_m < \theta(h_0), \gamma_{m+1} > \theta(h_0)\right)\\ &= \mathrm{P}\left(\gamma_m < \min\{\epsilon_s P^{-1}\beta^{-2}|s_0|^{-2}, \theta(h_0)\}, \gamma_{m+1} > \theta(h_0)\right).\end{aligned} \quad (25)$$

Define $a_{s_0,h_0} = \min\{\epsilon_s P^{-1}\beta^{-2}|s_0|^{-2}, \theta(h_0)\}$. By applying order statistics, the joint pdf of γ_m and γ_{m+1} is given by [31]

$$\begin{aligned}f_{\gamma_m,\gamma_{m+1}}(x,y) &= \mu_0 f_\gamma(x) f_\gamma(y)\left(F_\gamma(x)\right)^{m-1}\\ &\quad \times \left(1 - F_\gamma(y)\right)^{M-m-1}, \quad (26)\end{aligned}$$

for $x < y$, where $\mu_0 = \frac{M!}{(m-1)!(M-m-1)!}$. Therefore, Q_m can be expressed as follows:

$$Q_m = \bar{\mu}_0 \mathcal{E}_{h_0,s_0} \left\{ \int_{\theta(h_0)}^{\infty} f_\gamma(y) \left(1 - F_\gamma(y)\right)^{M-m-1} dy \times \int_0^{a_{s_0,h_0}} f_\gamma(x) \left(F_\gamma(x)\right)^{m-1} dx \right\} \tag{27}$$

$$= \bar{\mu}_0 \mathcal{E}_{h_0,s_0} \left\{ \left(1 - F_\gamma(\theta(h_0))\right)^{M-m} \left(F_\gamma(a_{s_0,h_0})\right)^m \right\}, \tag{28}$$

where $\bar{\mu}_0 = \mu_0 \mathcal{E}_{h_0,s_0}$. Because the density functions of γ_m contain Bessel functions, a closed-form expression for Q_m is difficult to obtain, and hence we consider an ideal scenario, in which the connection from U_0 to U_m, $1 \le m \le M$, is lossless. This assumption yields a lower bound on Q_m as follows:

$$Q_m \ge \bar{\mu}_0 \mathcal{E}_{h_0,s_0} \left\{ \left(1 - \bar{F}_\gamma(\theta(h_0))\right)^{M-m} \left(\bar{F}_\gamma(a_{s_0,h_0})\right)^m \right\}, \tag{29}$$

where $\bar{F}_\gamma(x) = 1 - e^{-\lambda_h x}$. For the case E_m, $m \ge 1$, we have $\theta(h_0) \ge 0$, which means that $|h_0|^2 \ge \epsilon_0 P^{-1}$. In addition, $a_{s_0,h_0} = \min\{\epsilon_s P^{-1}\beta^{-2}|s_0|^{-2}, \theta(h_0)\} = \theta(h_0)$ implies the following

$$|h_0|^2 \le \epsilon_0 \epsilon_s P^{-1} |s_0|^{-2} + \epsilon_0 P^{-1}. \tag{30}$$

By applying the simplified CDF, $\bar{F}_\gamma(x)$, the lower bound on Q_m can be expressed as follows:

$$\begin{aligned} Q_m \ge\ & \bar{\mu}_0 \int_0^\infty e^{-y} \int_{\frac{\epsilon_0}{P}}^{\frac{\epsilon_0\epsilon_s}{Py} + \frac{\epsilon_0}{P}} \left(1 - e^{-\lambda_h \theta(x)}\right)^m \\ & \times e^{-(M-m)\lambda_h \theta(x)} \lambda_0 e^{-\lambda_0 x} dx dy \\ & + \bar{\mu}_0 \int_0^\infty e^{-y} \left(1 - e^{-\lambda_h \epsilon_s P^{-1} \beta^{-2} y^{-1}}\right)^m \\ & \times \int_{\frac{\epsilon_0\epsilon_s}{Py} + \frac{\epsilon_0}{P}}^\infty e^{-(M-m)\lambda_h \theta(x)} \lambda_0 e^{-\lambda_0 x} dx dy. \end{aligned} \tag{31}$$

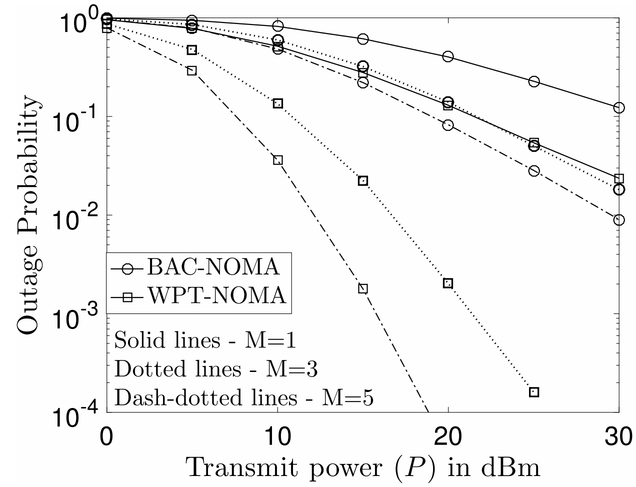
With some algebraic manipulations, the lower bound on Q_m can be approximated at high SNR as follows:

$$\begin{aligned} Q_m \ \ge\ & \bar{\mu}_0 \lambda_0 \sum_{p=0}^m \binom{m}{p} (-1)^p \breve{\mu}_p^{-1} \left[-\frac{\breve{\mu}_p \epsilon_0 \epsilon_s}{P} \ln \frac{\breve{\mu}_p \epsilon_0 \epsilon_s}{P} \right] \\ & + \bar{\mu}_0 \lambda_0 \breve{\mu}_0^{-1} \sum_{p=0}^m \binom{m}{p} (-1)^p \left(\frac{4\tilde{\mu}_p}{P} \ln \frac{4\tilde{\mu}_p}{P} \right) \end{aligned} \tag{32}$$

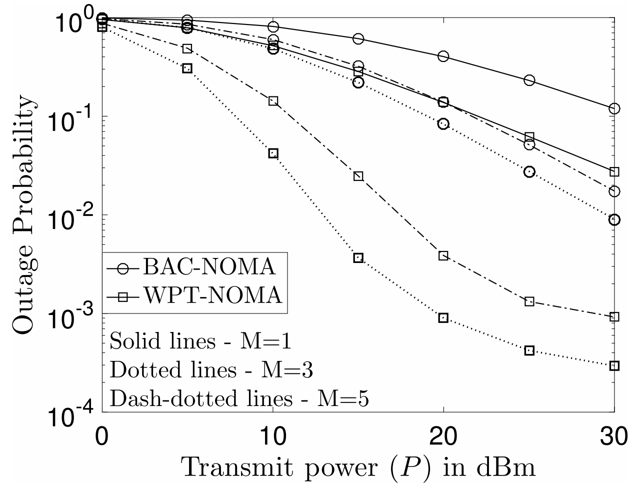
$$\rightarrow \frac{1}{P \ln^{-1} P}, \tag{33}$$

where the last approximation follows from the fact that each term in (32) can be approximated as $\frac{1}{P \ln^{-1} P}$.

Remark 8: Following the steps in the proof for Theorem 1 and also using (33), it is straightforward to show that the achievable diversity gain is one. In other words, the approximation obtained in (33) shows that the existence of virtual fast fading $|s_0|^2$ caps the diversity gain achieved by BAC-NOMA by one, even if the outage probability error floor can be discarded.



(a) $R_0 = 0.1$ BPCU



(b) $R_0 = 2$ BPCU

Fig. 1. Outage performance of BAC-NOMA and WPT-NOMA. $R_s = 1.2$ bit per channel use (BPCU). $d_h = d_0 = 50$ m and $d_g = 5$ m. $\alpha = 0.5$, $\beta = 0.1$, and $\eta = 0.1$.

V. Simulation Results

In this section, the performance of the two considered transmission schemes, BAC-NOMA and WPT-NOMA, is investigated by using computer simulation results. For all the carried out simulations, we choose $\phi = 3.5$ and the noise power is -94 dBm. In Fig. 1, the outage performance achieved by WPT-NOMT and BAC-NOMA is studied with different choices of R_0. In Fig. 1(a), the choice $R_0 = 0.1$ bits per channel use (BPCU) is used. With $R_0 = 0.1$ BPCU and $R_s = 1.2$ BPCU, it is straightforward to verify that the condition $\bar{\epsilon}_0 \bar{\epsilon}_s < 1$ holds. As indicated in Theorem 1, if $\bar{\epsilon}_0 \bar{\epsilon}_s < 1$ holds, WPT-NOMA can avoid outage probability error floors, which is consistent to the observations made from Fig. 1(a). In addition, Fig. 1(a) shows that the slope of the outage probability curve for WPT-NOMA is increased when increasing M, which indicates that the diversity gain achieved by WPT-NOMA is increased by increasing M, an observation also consistent to the conclusion made in Theorem 1. In Fig. 1(b), the choice $R_0 = 2$ BPCU is used, which leads

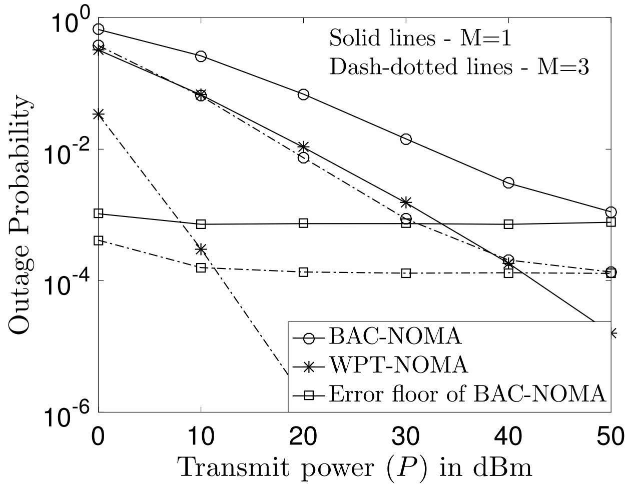
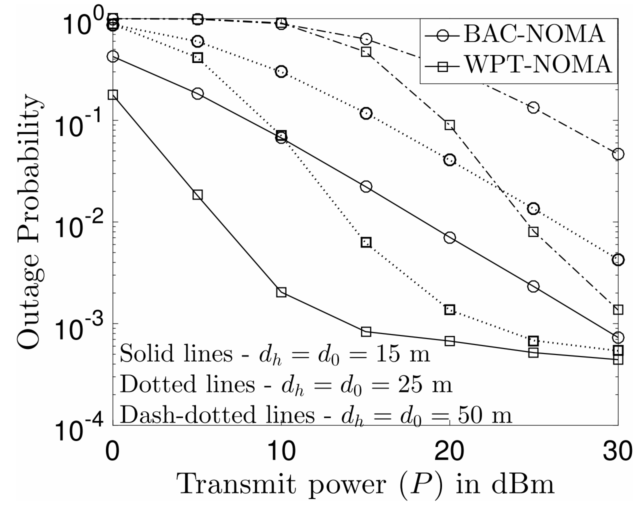
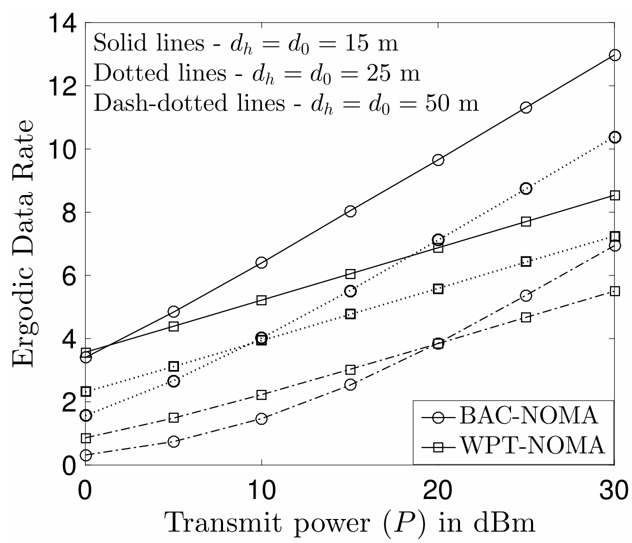


Fig. 2. Illustration of the outage probability error floor of BAC-NOMA. $R_0 = 0.1$ BPCU and $R_s = 1.2$ BPCU. $d_h = d_0 = 100$ m and $d_g = 1$ m. $\alpha = 0.5$, $\beta = 0.1$, and $\eta = 0.1$.



(a) Outage Probability



(b) Ergodic Data Rate

Fig. 3. Impact of path loss on the performance of BAC-NOMA and WPT-NOMA. $M = 5$, $R_0 = 2$ BPCU, $R_s = 3$ BPCU. $d_g = 5$ m. $\alpha = 0.5$, $\beta = 0.1$, and $\eta = 0.1$.

to the violation of the condition $\bar{\epsilon}_0\bar{\epsilon}_s < 1$. As a result, there are error floors for the outage probabilities achieved by WPT-NOMA, as shown in Fig. 1(b).

On the other hand, the two figures in Fig. 1 show that BAC-NOMA always suffers outage probability error floors, which is due to the fact that hybrid SIC decoding order cannot be implemented in BAC-NOMA systems. In addition, the figures also demonstrate that the performance of BAC-NOMA can be improved by increasing M, i.e., inviting more delay-tolerant devices to participate in NOMA transmission is beneficial to improve reception reliability. But unlike WPT-NOMA, increasing M does not change the slope of the outage probability curve for BAC-NOMA. It is worth to point out that for the two considered choices of R_0, WPT-NOMA can always realize a smaller outage probability than BAC-NOMA, as shown in Fig. 1.

In Fig. 2, the outage probability error floor experienced by BAC-NOMA is studied, where the term in the legend, 'Error Floor of BAC-NOMA', refers to $\mathrm{P}(E_0)$. In order to clearly show the asymptotic behaviour of the outage probability, a larger transmit power range than those in Fig. 1 is used. As can be observed from the figure, $\mathrm{P}(E_0)$ is a tight lower bound on the outage probability, and it is constant at high SNR, which implies that E_0 is the most damaging event and is the cause for the error floor of the outage probability. Another important observation is that increasing M is useful to reduce the error floor, which confirms Lemma 1. On the other hand, WPT-NOMA does not suffer any outage probability error floor because the used target rate choices satisfy $\bar{\epsilon}_0\bar{\epsilon}_s < 1$.

In Fig. 3, the impact of path loss on the performance of WPT-NOMA and BAC-NOMA is studied. In Fig. 3(a), the outage probability is used as the metric for the performance evaluation, whereas the ergodic data rate is used as the metric in Fig. 3(b). The two figures in Fig. 3 show that the performance of the two NOMA schemes is degraded when path loss becomes more severe. This deteriorating effect of path loss can be explained by using WPT-NOMA as an example. Increasing path loss does not only increase the attenuation of the signal strength, but also reduces the energy harvested at the delay-tolerant devices. For a similar reason, the performance of BAC-NOMA is also significantly affected by path loss. Therefore, the ideal applications of BAC-NOMA and WPT-NOMA are indoor communication scenarios, e.g., the distances between the nodes are not large. We note that WPT-NOMA also exhibits outage probability error floors in Fig. 3(a), since the condition $\bar{\epsilon}_0\bar{\epsilon}_s < 1$ does not hold, an observation consistent to the previous figures. In addition, Fig. 3(a) shows that WPT-NOMA outperforms BAC-NOMA, if the outage probability is used as the metric for performance evaluation, which is also consistent to the previous numerical studies. However, Fig. 3(b) shows an interesting result that BAC-NOMA can outperform WPT-NOMA if the ergodic rate is used as the performance metric, particularly at high SNR and with small path loss. One possible reason is that WPT-NOMA relies on the time-switching WPT strategy, i.e., the first αT seconds are used for energy harvesting, and the

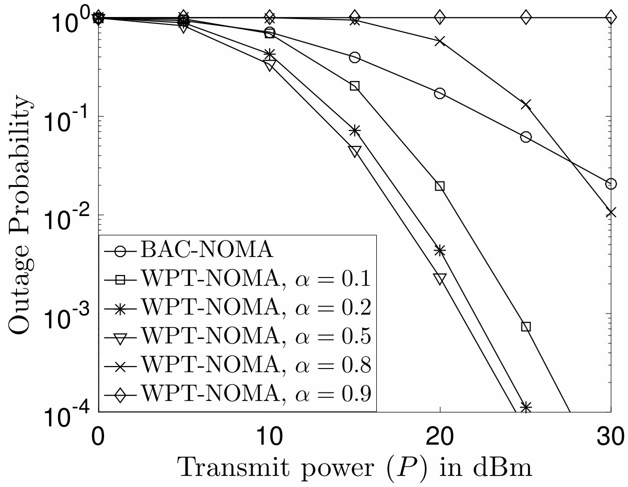
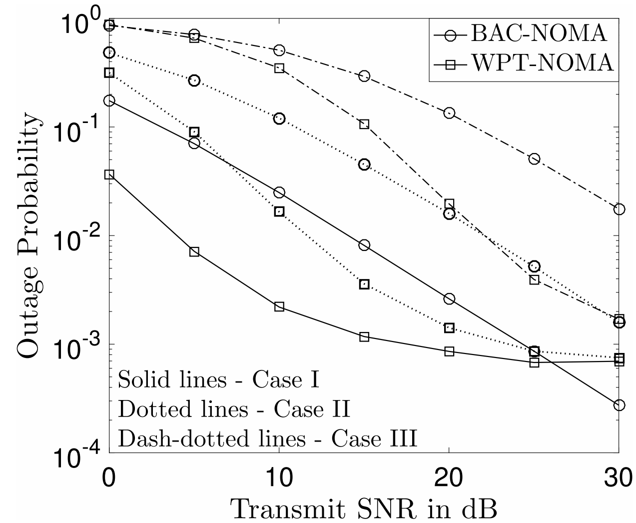


Fig. 4. Impact of the choices of α on the performance of WPT-NOMA. $R_0 = 0.1$ BPCU and $R_s = 2$ BPCU. $d_h = d_0 = 50$ m and $d_g = 5$ m. $M = 5$, $\beta = 0.1$, and $\eta = 0.1$.

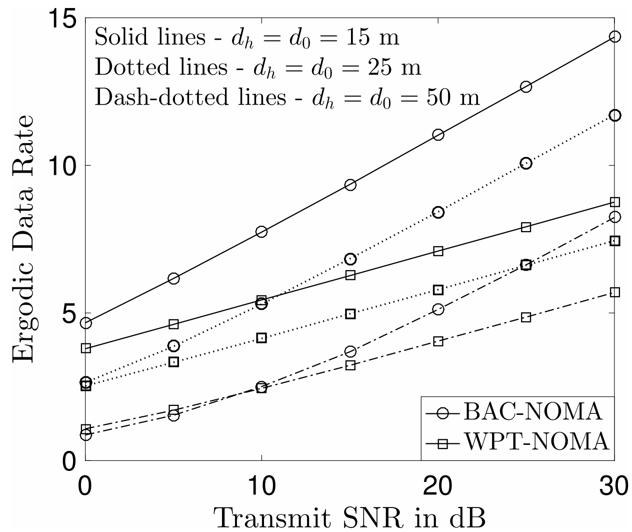
remaining $(1-\alpha)T$ seconds are used for data transmission. On in other words, there is less time available for WPT-NOMA to transmit, whereas BAC-NOMA can carry out transmission continuously.

In order to clearly demonstrate the impact of α on the performance of WPT-NOMA, in Fig. 4, different choices of α are used. In particular, $\alpha = 0.1$ and $\alpha = 0.9$ are a pair of choices of interest, as explained in the following. The use of $\alpha = 0.1$ means that the delay-tolerant devices use a small amount of time for energy harvesting and the majority time for data transmission, whereas $\alpha = 0.9$ means that the majority time is used for energy harvesting. Fig. 4 demonstrates that the choice of $\alpha = 0.9$ results in the poorest performance among all the choices shown in the figure. This is due to the fact that there is not sufficient time for data transmission, even though a good amount of energy has been harvested and the delay-tolerant devices can use larger transmit powers than that in the case with $\alpha = 0.1$. It is worth pointing out that the choice of $\alpha = 0.5$ yields the best performance among the choices shown in the figure.

Recall that, in order to facilitate the performance analysis, $\mathrm{U_m}$, $1 \leq m \leq M$, are assumed to be located in a cluster whose size is so small that the distances from $\mathrm{U_m}$ to the access point are similar. In Fig. 5, the performance of BACK-NOMA and WPT-NOMA is evaluated without this assumption. In particular, the locations of the access point and U_0 are fixed, whereas $\mathrm{U_m}$, $1 \leq m \leq M$, are uniformly located in a square with sides of length 10 m and centred at U_0. For the case in which the distance between $\mathrm{U_m}$ and U_0 is smaller than one meter, the corresponding path loss is kept as 1. It is important to point out that all the observations made to Fig. 3 are still applicable to Fig. 5. For example, WPT-NOMA can offer a large outage performance gain over BAC-NOMA, but BAC-NOMA outperforms WPT-NOMA if the ergodic data rate is used as the performance metric. It is worth to note that the performance of the two NOMA schemes in Fig. 5 is improved, compared to Fig. 3. The reason for this observation is that for Fig. 3, the distance between $\mathrm{U_m}$ and U_0 is fixed



(a) Outage Probability



(b) Ergodic Data Rate

Fig. 5. Impact of path loss on the performance of BAC-NOMA and WPT-NOMA. $M = 5$, $R_0 = 2$ BPCU, $R_s = 3$ BPCU. $\alpha = 0.5$, $\beta = 0.1$, and $\eta = 0.1$. The access point is located at the origin. $\mathrm{U_m}$, $1 \leq m \leq M$, are uniformly located in a square with sides of length 10 m and centred at U_0. U_0 is located at $(15m, 15m)$ for Case I, $(25m, 25m)$ for Case II, and $(50\ m, 50\ m)$ for Case III, respectively.

at 5 m, but for Fig. 5, it is possible that the distance between $\mathrm{U_m}$ and U_0 is smaller than 5 m.

VI. Conclusion

In this paper, two energy and spectrally efficient transmission strategies, namely WPT-NOMA and BAC-NOMA, were proposed by employing the energy and spectrum cooperation among the IoT devices. For the proposed WPT-NOMA scheme, hybrid SIC decoding order was used to improve reception reliability, and the developed analytical results demonstrate that WPT-NOMA can avoid outage probability error floors and realize the full diversity gain. Unlike WPT-NOMA, BAC-NOMA suffers from an outage probability error floor, and the asymptotic behaviour of this error floor was analyzed in the paper by applying EVT. In addition, the effect of using one device's signal as the carrier signal was studied, and its harmful impact on the diversity gain was revealed. In this paper, we have studied the two techniques from the communication perspective, and it is worth to point out that different hardware circuits are required by the two techniques. Therefore, comparing the two schemes using more sophisticated metrics from both the communication and circuit theories will be an important direction for future research.

We note that the provided simulation results show that the choice of α has a significant impact on the performance of WPT-NOMA, and therefore an important direction for future research is to develop low-complexity algorithms for optimizing α. In addition, we note that the reason for BAC-NOMA to suffer the outage probability error floor is due to the fact that hybrid SIC decoding order cannot be implemented. However, provided that U_n, $1 \leq n \leq M$, can carry out non-coherent detection, it is possible to apply hybrid SIC decoding order to BAC-NOMA, which is another important direction for future research. Furthermore, a linear energy harvesting model was used to obtain insightful understandings about the performance of WPT-NOMA, and it is an important direction of future research to consider the use of non-linear energy harvesting models.

APPENDIX A
PROOF FOR THEOREM 1

The proof for the theorem can be divided to four steps, where the first three steps are to analyze the asymptotic behavior of T_0, T_m, $1 \leq m \leq M-1$, and T_M, respectively., and the last step is to study the overall diversity gain.

A. Asymptotic Study of T_0

This section focuses on the high-SNR approximation of T_0 which can be rewritten as follows:

$$\begin{aligned} T_0 &= \mathrm{P}\left(\mathrm{R}_M^{WP,1} < R_s, |\mathcal{S}_2| = 0\right) \\ &= \mathrm{P}\left(\gamma_M < \frac{\bar{\epsilon}_s(P|h_0|^2+1)}{\eta P\bar{\alpha}}, \gamma_1 > \tau(h_0)\right). \end{aligned} \tag{34}$$

As can be observed from (34), T_0 is a function of two order statistics, γ_1 and γ_M, whose joint pdf is given by [31]

$$f_{\gamma_1,\gamma_M}(x,y) = \frac{M!}{(M-2)!} f_\gamma(x) f_\gamma(y) \left[F_\gamma(y) - F_\gamma(x)\right]^{M-2}. \tag{35}$$

Denote $T_{0|h_0}$ by the value of T_0 when h_0 is treated as a constant. Therefore, $T_{0|h_0}$ can be expressed as follows:

$$\begin{aligned} T_{0|h_0} &= \frac{M!}{(M-2)!}\int_{\tau(h_0)}^{\frac{\bar{\epsilon}_s(P|h_0|^2+1)}{\eta P\bar{\alpha}}} f_\gamma(x) \int_x^{\frac{\bar{\epsilon}_s(P|h_0|^2+1)}{\eta P\bar{\alpha}}} f_\gamma(y) \\ &\quad \times \left[F_\gamma(y) - F_\gamma(x)\right]^{M-2} dydx \\ &= \frac{M!}{(M-1)!}\int_{\tau(h_0)}^{\frac{\bar{\epsilon}_s(P|h_0|^2+1)}{\eta P\bar{\alpha}}} f_\gamma(x) \\ &\quad \times \left[F_\gamma\left(\frac{\bar{\epsilon}_s(P|h_0|^2+1)}{\eta P\bar{\alpha}}\right) - F_\gamma(x)\right]^{M-1} dx. \end{aligned}$$

$T_{0|h_0}$ can be further simplified as follows:

$$T_{0|h_0} = \left[F_\gamma\left(\frac{\bar{\epsilon}_s(P|h_0|^2+1)}{\eta P\bar{\alpha}}\right) - F_\gamma(\tau(h_0))\right]^M. \tag{36}$$

Therefore, T_0 can be obtained by finding the expectation of $T_{0|h_0}$ with respect to h_0:

$$T_0 = \mathcal{E}_{h_0}\left\{T_{0|h_0}\right\}.$$

We note that $\tau(h_0)$ can have different forms depending on the choice of $|h_0|^2$. In particular, $\tau(h_0) = 0$ means

$$\frac{|h_0|^2}{\bar{\epsilon}_0\eta\bar{\alpha}} - \frac{1}{\eta P\bar{\alpha}} \leq 0, \tag{37}$$

which requires

$$|h_0|^2 \leq \frac{\bar{\epsilon}_0}{P}. \tag{38}$$

For the case $\tau(h_0) \neq 0$, the probability shown in (34) requires $\tau(h_0) < \frac{\bar{\epsilon}_s(P|h_0|^2+1)}{\eta P\bar{\alpha}}$. This hidden constraint imposes another constraint on $|h_0|^2$ as follows:

$$\frac{|h_0|^2}{\bar{\epsilon}_0} - \frac{1}{P} < \frac{\bar{\epsilon}_s(P|h_0|^2+1)}{P}, \tag{39}$$

which can be explicitly expressed as follows:

$$|h_0|^2 < \frac{\bar{\epsilon}_0(1+\bar{\epsilon}_s)}{P(1-\bar{\epsilon}_0\bar{\epsilon}_s)}. \tag{40}$$

By using the constraints shown in (38) and (40), T_1 can be expressed as follows:

$$\begin{aligned} T_0 &= \lambda_0\int_0^{\frac{\bar{\epsilon}_0}{P}} \left[F_\gamma\left(\frac{\bar{\epsilon}_s(Px+1)}{\eta P\bar{\alpha}}\right) - F_\gamma(0)\right]^M e^{-\lambda_0 x}dx \\ &\quad + \lambda_0\int_{\frac{\bar{\epsilon}_0}{P}}^{\frac{\bar{\epsilon}_0(1+\bar{\epsilon}_s)}{P(1-\bar{\epsilon}_0\bar{\epsilon}_s)}} e^{-\lambda_0 x}\left[F_\gamma\left(\frac{\bar{\epsilon}_s(Px+1)}{\eta P\bar{\alpha}}\right)\right. \\ &\quad \left. - F_\gamma\left(\frac{x}{\bar{\epsilon}_0\eta\bar{\alpha}} - \frac{1}{\eta P\bar{\alpha}}\right)\right]^M dx. \end{aligned} \tag{41}$$

We note that the upper bound on $|h_0|^2$, $\frac{\bar{\epsilon}_0(1+\bar{\epsilon}_s)}{P(1-\bar{\epsilon}_0\bar{\epsilon}_s)}$, is crucial to remove outage probability error floors and realize the full diversity gain, as shown in the following.

In particular, one can observe that both $\frac{\bar{\epsilon}_s(Px+1)}{\eta P\bar{\alpha}}$ and $\frac{x}{\bar{\epsilon}_0\eta\bar{\alpha}} - \frac{1}{\eta P\bar{\alpha}}$ go to zero for $P\to\infty$ in the two integrals considered in (41). Therefore, the parameters of the Bessel functions in T_0 go to zero for $P\to\infty$. Recall that $xK_1(x) \approx 1 + \frac{x^2}{2}\ln\frac{x}{2}$, for $x \to 0$ [39]. Therefore, the CDF of the unordered channel gain can be approximated as follows:

$$\begin{aligned} F_\gamma(x) &= 1 - 2\sqrt{\lambda_h\lambda_g x}K_1\left(2\sqrt{\lambda_h\lambda_g x}\right) \\ &\approx 1 - (1 + \lambda_h\lambda_g x\ln(\lambda_h\lambda_g x)) = -\lambda_h\lambda_g x\ln(\lambda_h\lambda_g x), \end{aligned} \tag{42}$$

for $x \to 0$. We note that for $x \to 0$, $\ln(\lambda_h\lambda_g x) < 0$ and hence the approximation for $F_\gamma(x)$ in (42) is still positive.

Therefore, T_0 can be approximated at high SNR as follows:

$$\begin{aligned}T_0 \approx& \int_0^{\frac{\bar{\epsilon}_0}{P}} \left[-\lambda_h\lambda_g \frac{\bar{\epsilon}_s(Px+1)}{\eta P\bar{\alpha}} \ln\left(\lambda_h\lambda_g \frac{\bar{\epsilon}_s(Px+1)}{\eta P\bar{\alpha}}\right)\right]^M \\ &\times dx\lambda_0 + \lambda_0 \int_{\frac{\bar{\epsilon}_0}{P}}^{\frac{\bar{\epsilon}_0(1+\bar{\epsilon}_s)}{P(1-\bar{\epsilon}_0\bar{\epsilon}_s)}} \\ &\times \left[\lambda_h\lambda_g\left(\frac{x}{\bar{\epsilon}_0\eta\bar{\alpha}} - \frac{1}{\eta P\bar{\alpha}}\right)\ln\left(\lambda_h\lambda_g\left(\frac{x}{\bar{\epsilon}_0\eta\bar{\alpha}} - \frac{1}{\eta P\bar{\alpha}}\right)\right)\right. \\ &\left. - \lambda_h\lambda_g \frac{\bar{\epsilon}_s(Px+1)}{\eta P\bar{\alpha}} \ln\left(\lambda_h\lambda_g \frac{\bar{\epsilon}_s(Px+1)}{\eta P\bar{\alpha}}\right)\right]^M dx. \end{aligned} \tag{43}$$

In order to obtain a more insightful asymptotic expression of T_0, the expression in (43) can be rewritten as follows:

$$\begin{aligned}T_0 \approx& \frac{\lambda_0}{P}\int_0^{\bar{\epsilon}_0} \left[-\lambda_h\lambda_g \frac{\bar{\epsilon}_s(y+1)}{\eta P\bar{\alpha}} \ln\left(\lambda_h\lambda_g \frac{\bar{\epsilon}_s(y+1)}{\eta P\bar{\alpha}}\right)\right]^M dy \\ &+ \frac{\lambda_0}{P}\int_{\bar{\epsilon}_0}^{\frac{\bar{\epsilon}_0(1+\bar{\epsilon}_s)}{(1-\bar{\epsilon}_0\bar{\epsilon}_s)}} \left[\frac{\lambda_h\lambda_g}{P}\left(\frac{y}{\bar{\epsilon}_0\eta\bar{\alpha}} - \frac{1}{\eta\bar{\alpha}}\right)\right. \\ &\times \ln\left(\frac{\lambda_h\lambda_g}{P}\left(\frac{y}{\bar{\epsilon}_0\eta\bar{\alpha}} - \frac{1}{\eta\bar{\alpha}}\right)\right) - \lambda_h\lambda_g \\ &\left.\times \frac{\bar{\epsilon}_s(y+1)}{\eta P\bar{\alpha}} \ln\left(\lambda_h\lambda_g \frac{\bar{\epsilon}_s(y+1)}{\eta P\bar{\alpha}}\right)\right]^M dy \\ =& \frac{\lambda_0}{P}\int_0^{\bar{\epsilon}_0}\left[-\frac{b_1(y)}{P}\ln\left(\frac{b_1(y)}{P}\right)\right]^M dy + \frac{\lambda_0}{P}\int_{\bar{\epsilon}_0}^{\frac{\bar{\epsilon}_0(1+\bar{\epsilon}_s)}{(1-\bar{\epsilon}_0\bar{\epsilon}_s)}} \\ &\times \left[\frac{b_2(y)}{P}\ln\left(\frac{b_2(y)}{P}\right) - \frac{b_1(y)}{P}\ln\left(\frac{b_1(y)}{P}\right)\right]^M dy, \end{aligned} \tag{44}$$

where $y = Px$, $b_1(y) = \lambda_h\lambda_g\frac{\bar{\epsilon}_s(y+1)}{\eta\bar{\alpha}}$ and $b_2(y) = \lambda_h\lambda_g\left(\frac{y}{\bar{\epsilon}_0\eta\bar{\alpha}} - \frac{1}{\eta\bar{\alpha}}\right)$. It is important to point out that both $b_1(y)$ and $b_2(y)$ are constant and not functions of P.

Denote the two integrals in (44) by $\tilde{Q}_1$ and $\tilde{Q}_2$, respectively. For $\tilde{Q}_1$, the following approximation can be used:

$$\frac{b_1(y)}{P}\ln\left(\frac{b_1(y)}{P}\right) = \frac{b_1(y)}{P}\left[\ln b_1(y) - \ln P\right] \tag{45}$$

$$\underset{P\to\infty}{\approx} -\frac{b_1(y)}{P}\ln P = -\frac{b_1(y)}{P\ln^{-1}P}, \tag{46}$$

since $b_1(y)$ is finite and strictly larger than zero for the integral considered in $\tilde{Q}_1$. Therefore, Q_1 can be approximated as follows:

$$\begin{aligned}\tilde{Q}_1 \approx& \int_0^{\bar{\epsilon}_0}\left[\frac{b_1(y)}{P\ln^{-1}P}\right]^M dy = \frac{e_1}{P^M\ln^{-M}P} \\ =& \mathcal{O}\left(\frac{1}{P^M\ln^{-M}P}\right),\end{aligned} \tag{47}$$

where $\mathcal{O}$ denotes the approximation operation by omitting the constant multiplicative coefficient, and the last approximation follows from the fact that $e_1 = \int_0^{\bar{\epsilon}_0}\left[b_1(y)\right]^M dy$ is constant and not a function of P.

The approximation for $\tilde{Q}_2$ is more complicated since $b_2(y)$ can be zero for the considered integral and hence $\ln b_2(y)$ can be unbounded. Unlike $\tilde{Q}_1$, $\tilde{Q}_2$ can be approximated as follows:

$$\begin{aligned}\tilde{Q}_2 =& \sum_{p=0}^{M}\frac{(-1)^p}{P^M}\binom{M}{p}\int_{\bar{\epsilon}_0}^{\frac{\bar{\epsilon}_0(1+\bar{\epsilon}_s)}{(1-\bar{\epsilon}_0\bar{\epsilon}_s)}} b_2(y)^{M-p}b_1(y)^p \\ &\times\left[\ln b_2(y) - \ln P\right]^{M-p}\left[\ln b_1(y) - \ln P\right]^p dy \\ =& \sum_{p=0}^{M}\frac{(-1)^p}{P^M}\int_{\bar{\epsilon}_0}^{\frac{\bar{\epsilon}_0(1+\bar{\epsilon}_s)}{(1-\bar{\epsilon}_0\bar{\epsilon}_s)}} b_2(y)^{M-p}b_1(y)^p \\ &\times\left(\sum_{i=0}^{M-p}(-1)^i\binom{M-p}{i}(\ln b_2(y))^{M-p-i}(\ln P)^i\right) \\ &\times\left(\sum_{j=0}^{p}\binom{p}{j}(-1)^j(\ln b_1(y))^{p-j}(\ln P)^j\right)dy. \end{aligned} \tag{48}$$

At high SNR, the term with $(\ln P)^M$ is dominant, compared to the terms with $(\ln P)^m$, $m < M$, which means that (48) can be further approximated as follows:

$$\begin{aligned}\tilde{Q}_2 \approx& \frac{(\ln P)^M}{P^M}\sum_{p=0}^{M}\sum_{i+j=M}(-1)^{p+i+j}\binom{M-p}{i}\binom{p}{j} \\ &\times\int_{\bar{\epsilon}_0}^{\frac{\bar{\epsilon}_0(1+\bar{\epsilon}_s)}{(1-\bar{\epsilon}_0\bar{\epsilon}_s)}} b_2(y)^{M-p}b_1(y)^p(\ln b_1(y))^{p-j} \\ &\times(\ln b_2(y))^{M-p-i}dy = \mathcal{O}\left(\frac{1}{P^M\ln^{-M}P}\right). \end{aligned} \tag{49}$$

Therefore, with $P \to \infty$, T_0 can be approximated as follows:

$$T_0 = \frac{\lambda_0}{P}\tilde{Q}_1 + \frac{\lambda_0}{P}\tilde{Q}_2 = \mathcal{O}\left(\frac{1}{P^{M+1}\ln^{-M}P}\right). \tag{50}$$

B. Asymptotic Study of T_m, $1 \leq m \leq M$

This section is to focus on T_m, $1 \leq m \leq M-1$, which can be expressed as follows:

$$\begin{aligned}T_m =& \mathrm{P}\left(\mathrm{R}_m^{WP,2} < R_s, \mathrm{R}_M^{WP,1} < R_s, |\mathcal{S}_2| = m\right) \\ =& \mathrm{P}\left(\gamma_m < \frac{\bar{\epsilon}_s}{\eta P\bar{\alpha}}, \gamma_m < \tau(h_0), \gamma_{m+1} > \tau(h_0),\right. \\ &\left.\gamma_M < \frac{\bar{\epsilon}_s(P|h_0|^2+1)}{\eta P\bar{\alpha}}\right).\end{aligned}$$

For the case of $1 \leq m \leq M$, $\tau(h_0) \neq 0$, which means

$$\frac{|h_0|^2}{\bar{\epsilon}_0\eta\bar{\alpha}} - \frac{1}{\eta P\bar{\alpha}} > 0, \tag{51}$$

or equivalently $|h_0|^2 > \frac{\bar{\epsilon}_0}{P}$. Furthermore, the requirement $\tau(h_0) < \frac{\bar{\epsilon}_s(P|h_0|^2+1)}{\eta P\bar{\alpha}}$ leads to the constraint $|h_0|^2 < \frac{\bar{\epsilon}_0(1+\bar{\epsilon}_s)}{P(1-\bar{\epsilon}_0\bar{\epsilon}_s)}$, as discussed in (40).
c

Therefore, T_m can be rewritten as follows:

$$\begin{aligned}T_m =& \mathrm{P}\left(\gamma_m < \frac{\bar{\epsilon}_s}{\eta P\bar{\alpha}}, \gamma_M < \frac{\bar{\epsilon}_s(P|h_0|^2+1)}{\eta P\bar{\alpha}},\right. \\ &\left.\gamma_m < \frac{|h_0|^2}{\bar{\epsilon}_0\eta\bar{\alpha}} - \frac{1}{\eta P\bar{\alpha}}, \gamma_{m+1} > \frac{|h_0|^2}{\bar{\epsilon}_0\eta\bar{\alpha}} - \frac{1}{\eta P\bar{\alpha}}\right) \\ =& \mathrm{P}\left(\gamma_m < b_{h_0}, \gamma_{m+1} > \tau(h_0), \gamma_M < a(h_0)\right), \end{aligned} \tag{52}$$

where $a(h_0) = \frac{\bar{\epsilon}_s(P|h_0|^2+1)}{\eta P\bar{\alpha}}$ and $b_{h_0} = \min\left\{\frac{\bar{\epsilon}_s}{\eta P\bar{\alpha}}, \tau(h_0)\right\}$.

As can be observed from (52), T_m, $1 \leq m \leq M-1$, is a function of three order statistics, γ_m, γ_{m+1}, and γ_M. Recall that the joint pdf of three order statistics is given by [31]

$$\begin{aligned} &f_{\gamma_m,\gamma_{m+1},\gamma_M}(x,y,z) \\ &\quad = c_m F_\gamma(x)^{m-1} \\ &\quad\quad \times \left(F_\gamma(z) - F_\gamma(y)\right)^{M-m-2} f_\gamma(x) f_\gamma(y) f_\gamma(z), \end{aligned} \tag{53}$$

where $c_m = \frac{M!}{(m-1)!(M-m-2)!}$.

Denote $T_{m|h_0}$ by the value of T_m by assuming that h_0 is fixed. By using the joint pdf in (53), $T_{m|h_0}$ can be expressed as follows:

$$\begin{aligned} T_{m|h_0} &= \mathrm{P}\left(\gamma_m < b_{h_0}, \gamma_{m+1} > \tau(h_0), \gamma_M < a(h_0)\right) \\ &= c_m \int_0^{b_{h_0}} F_\gamma(x)^{m-1} f_\gamma(x) dx \int_{\tau(h_0)}^{a(h_0)} f_\gamma(y) \\ &\quad \times \int_y^{a(h_0)} \left(F_\gamma(z) - F_\gamma(y)\right)^{M-m-2} f_\gamma(z) dz. \end{aligned} \tag{54}$$

By using the property of CDFs, $T_{m|h_0}$ can be more explicitly expressed as follows:

$$\begin{aligned} T_{m|h_0} &= \bar{c}_m F_\gamma(b_{h_0})^m \int_{\tau(h_0)}^{a(h_0)} \Big[\left(F_\gamma(a(h_0)) - F_\gamma(y)\right)^{M-m-1} \\ &\quad - \left(F_\gamma(y) - F_\gamma(y)\right)^{M-m-1}\Big] f_\gamma(y) dy \\ &= \bar{c}_m F_\gamma(b_{h_0})^m \\ &\quad \times \int_{\tau(h_0)}^{a(h_0)} \left[F_\gamma(a(h_0)) - F_\gamma(y)\right]^{M-m-1} f_\gamma(y) dy, \end{aligned} \tag{55}$$

where $\bar{c}_m = \frac{M!}{m!(M-m-1)!}$. The expression of $T_{m|h_0}$ can be further simplified as follows:

$$\begin{aligned} T_{m|h_0} &= \tilde{c}_m F_\gamma(b_{h_0})^m \Big(\left[F_\gamma(a(h_0)) - F_\gamma(\tau(h_0))\right]^{M-m} \\ &\quad - \left[F_\gamma(a(h_0)) - F_\gamma(a(h_0))\right]^{M-m}\Big) \\ &= \tilde{c}_m F_\gamma(b_{h_0})^m \left[F_\gamma(a(h_0)) - F_\gamma(\tau(h_0))\right]^{M-m}, \end{aligned} \tag{56}$$

where $\tilde{c}_m = \frac{M!}{m!(M-m)!}$.

T_m can be obtained by calculating the expectation of $T_{m|h_0}$ with respect of $|h_0|^2$ as follows:

$$\begin{aligned} T_m &= \mathcal{E}_{h_0}\left\{T_{m|h_0}\right\} \\ &= \tilde{c}_m \mathcal{E}_{h_0}\Big\{F_\gamma(b_{h_0})^m \left[F_\gamma(a(h_0)) - F_\gamma(\tau(h_0))\right]^{M-m}\Big\}. \end{aligned} \tag{57}$$

Recall that $b_{h_0} = \tau(h_0)$ if the constraint $\frac{\bar{\epsilon}_s}{\eta P \bar{\alpha}} > \frac{|h_0|^2}{\bar{\epsilon}_0 \eta \bar{\alpha}} - \frac{1}{\eta P \bar{\alpha}}$ is satisfied, which imposes the following constraint on $|h_0|^2$:

$$|h_0|^2 < \frac{\bar{\epsilon}_0(1+\bar{\epsilon}_s)}{P}. \tag{58}$$

Therefore, T_m can be more explicitly expressed as follows:

$$\begin{aligned} T_m &= \tilde{c}_m \lambda_0 \int_{\frac{\bar{\epsilon}_0}{P}}^{\frac{\bar{\epsilon}_0(1+\bar{\epsilon}_s)}{P}} \left(\left[F_\gamma(a(x)) - F_\gamma(\tau(x))\right]^{M-m}\right) \\ &\quad \times F_\gamma(\tau(x))^m e^{-\lambda_0 x} dx + \tilde{c}_m \lambda_0 F_\gamma\left(\frac{\bar{\epsilon}_s}{\eta P \bar{\alpha}}\right)^m \\ &\quad \times \int_{\frac{\bar{\epsilon}_0(1+\bar{\epsilon}_s)}{P}}^{\frac{\bar{\epsilon}_0(1+\bar{\epsilon}_s)}{P(1-\bar{\epsilon}_0\bar{\epsilon}_s)}} \left(\left[F_\gamma(a(x)) - F_\gamma(\tau(x))\right]^{M-m}\right) e^{-\lambda_0 x} dx, \end{aligned} \tag{59}$$

where the constraints on $|h_0|^2$ shown in (40), (51) and (58) have been used.

We note that for the integrals considered in (59), $\tau(x) \to 0$ for $P \to \infty$, which can be explained in the following. Recall that

$$\tau(x) = \frac{x}{\bar{\epsilon}_0 \eta \bar{\alpha}} - \frac{1}{\eta P \bar{\alpha}}. \tag{60}$$

For the integrals considered in (59), $\frac{\bar{\epsilon}_0}{P} \leq x \leq \frac{\bar{\epsilon}_0(1+\bar{\epsilon}_s)}{P}$ and $\frac{\bar{\epsilon}_0(1+\bar{\epsilon}_s)}{P} \leq x \leq \frac{\bar{\epsilon}_0(1+\bar{\epsilon}_s)}{P(1-\bar{\epsilon}_0\bar{\epsilon}_s)}$. Therefore, indeed $x \to 0$ for $P \to \infty$, which means that $\tau(x) \to 0$. Similarly, for the integrals considered in (59), the following approximation also holds

$$a(x) = \frac{\bar{\epsilon}_s x}{\eta \bar{\alpha}} + \frac{\bar{\epsilon}_s}{\eta P \bar{\alpha}} \underset{P\to\infty}{\longrightarrow} 0. \tag{61}$$

By using these asymptotic behaviours of $\tau(x)$ and $a(x)$, the probability T_m can be approximated as follows:

$$\begin{aligned} T_m &\approx \tilde{c}_m \lambda_0 \int_{\frac{\bar{\epsilon}_0}{P}}^{\frac{\bar{\epsilon}_0(1+\bar{\epsilon}_s)}{P}} \left[-\lambda_h \lambda_g \tau(x) \ln(\lambda_h \lambda_g \tau(x))\right]^m \left[\lambda_h \lambda_g \tau(x)\right. \\ &\quad \times \left.\ln(\lambda_h \lambda_g \tau(x)) - \lambda_h \lambda_g a(x) \ln(\lambda_h \lambda_g a(x))\right]^{M-m} dx \\ &\quad + \tilde{c}_m \lambda_0 \left[-\lambda_h \lambda_g \frac{\bar{\epsilon}_s}{\eta P \bar{\alpha}} \ln\left(\lambda_h \lambda_g \frac{\bar{\epsilon}_s}{\eta P \bar{\alpha}}\right)\right]^m \\ &\quad \times \int_{\frac{\bar{\epsilon}_0(1+\bar{\epsilon}_s)}{P}}^{\frac{\bar{\epsilon}_0(1+\bar{\epsilon}_s)}{P(1-\bar{\epsilon}_0\bar{\epsilon}_s)}} \left[\lambda_h \lambda_g \tau(x) \ln(\lambda_h \lambda_g \tau(x))\right. \\ &\quad \left. - \lambda_h \lambda_g a(x) \ln(\lambda_h \lambda_g a(x))\right]^{M-m} dx, \end{aligned}$$

for $P \to \infty$.

Define $\bar{\tau}(h_0) = P\lambda_h\lambda_g\tau(h_0)$ and $\bar{a}(h_0) = P\lambda_h\lambda_g a(h_0)$. Therefore, T_m can be expressed as follows:

$$\begin{aligned} T_m &\approx \tilde{c}_m \lambda_0 \int_{\frac{\bar{\epsilon}_0}{P}}^{\frac{\bar{\epsilon}_0(1+\bar{\epsilon}_s)}{P}} \left[-\frac{\bar{\tau}(x)}{P} \ln\left(\frac{\bar{\tau}(x)}{P}\right)\right]^m \\ &\quad \times \left[\frac{\bar{\tau}(x)}{P} \ln\left(\frac{\bar{\tau}(x)}{P}\right) - \frac{\bar{a}(x)}{P} \ln\left(\frac{\bar{a}(x)}{P}\right)\right]^{M-m} dx \\ &\quad + \tilde{c}_m \lambda_0 \left[-\frac{\lambda_h \lambda_g \bar{\epsilon}_s}{\eta P \bar{\alpha}} \ln\left(\frac{\lambda_h \lambda_g \bar{\epsilon}_s}{\eta P \bar{\alpha}}\right)\right]^m \int_{\frac{\bar{\epsilon}_0(1+\bar{\epsilon}_s)}{P}}^{\frac{\bar{\epsilon}_0(1+\bar{\epsilon}_s)}{P(1-\bar{\epsilon}_0\bar{\epsilon}_s)}} \\ &\quad \times \left[\frac{\bar{\tau}(x)}{P} \ln\left(\frac{\bar{\tau}(x)}{P}\right) - \frac{\bar{a}(x)}{P} \ln\left(\frac{\bar{a}(x)}{P}\right)\right]^{M-m} dx. \end{aligned} \tag{62}$$

In order to obtain a more insightful asymptotic expression, we substitute the following three parameters, $y = Px$,

$$\tilde{\tau}(y) = \lambda_h \lambda_g \left(\frac{y}{\bar{\epsilon}_0 \eta \bar{\alpha}} - \frac{1}{\eta \bar{\alpha}}\right), \tag{63}$$

and

$$\tilde{a}(y) = \lambda_h \lambda_g \left(\frac{\bar{\epsilon}_s y}{\eta \bar{\alpha}} + \frac{\bar{\epsilon}_s}{\eta \bar{\alpha}} \right), \tag{64}$$

into the expression of T_m, which yields the following expression:

$$\begin{aligned} T_m \approx & \frac{\tilde{c}_m \lambda_0}{P} \int_{\bar{\epsilon}_0}^{\bar{\epsilon}_0(1+\bar{\epsilon}_s)} \left[-\frac{\tilde{\tau}(y)}{P} \ln \left(\frac{\tilde{\tau}(y)}{P} \right) \right]^m \\ & \times \left[\frac{\tilde{\tau}(y)}{P} \ln \left(\frac{\tilde{\tau}(y)}{P} \right) - \frac{\tilde{a}(y)}{P} \ln \left(\frac{\tilde{a}(y)}{P} \right) \right]^{M-m} dy \\ & + \frac{\tilde{c}_m \lambda_0}{P} \left[-\frac{\lambda_h \lambda_g \bar{\epsilon}_s}{\eta P \bar{\alpha}} \ln \left(\frac{\lambda_h \lambda_g \bar{\epsilon}_s}{\eta P \bar{\alpha}} \right) \right]^m \int_{\bar{\epsilon}_0(1+\bar{\epsilon}_s)}^{\frac{\bar{\epsilon}_0(1+\bar{\epsilon}_s)}{(1-\bar{\epsilon}_0\bar{\epsilon}_s)}} \\ & \times \left[\frac{\tilde{\tau}(y)}{P} \ln \left(\frac{\tilde{\tau}(y)}{P} \right) - \frac{\tilde{a}(y)}{P} \ln \left(\frac{\tilde{a}(y)}{P} \right) \right]^{M-m} dy. \end{aligned} \tag{65}$$

It is important to point out that both $\tilde{\tau}(y)$ and $\tilde{a}(y)$ are constant and not functions of P. By using the steps similar to those to obtain the approximation of T_0, T_m can be approximated as follows:

$$T_m = \mathcal{O}\left(\frac{1}{P^{M+1} \ln^{-M} P} \right). \tag{66}$$

C. Asymptotic Study of T_M

For the special case T_M, we first recall that T_M can be expressed as follows:

$$\begin{aligned} T_M &= \mathrm{P}\left(\mathrm{R}_M^{WP,2} < R_s, \mathrm{R}_M^{WP,1} < R_s, |\mathcal{S}_2| = M \right) \\ &= \mathrm{P}\left(\gamma_M < \frac{\bar{\epsilon}_s}{\eta P \bar{\alpha}}, \gamma_M < \tau(h_0) \right). \end{aligned} \tag{67}$$

By using the marginal pdf of the largest order statistics, T_M can be be straightforwardly expressed as follows:

$$T_M = \mathcal{E}_{h_0} \left\{ F_\gamma \left(\min \left\{ \frac{\bar{\epsilon}_s}{\eta P \bar{\alpha}}, \tau(h_0) \right\} \right)^M \right\}. \tag{68}$$

As can be observed from (67), T_M is a function of γ_M only, which is different from T_m, $1 \leq m \leq M-1$. It is important to point out that the constraint of $|h_0|^2$ shown in (40) does not exist for T_M. This causes the reduction of the diversity gain from $M+1$ to M, as shown in the following. T_M can be more explicitly expressed as follows:

$$T_M = \underbrace{\tilde{c}_m \lambda_0 \int_{\frac{\bar{\epsilon}_0}{P}}^{\frac{\bar{\epsilon}_0(1+\bar{\epsilon}_s)}{P}} F_\gamma(\tau(x))^M e^{-\lambda_0 x} dx}_{T_{M,1}} \tag{69}$$

$$+ \underbrace{\tilde{c}_m \lambda_0 F_\gamma \left(\frac{\bar{\epsilon}_s}{\eta P \bar{\alpha}} \right)^M}_{T_{M,2}} \underbrace{\int_{\frac{\bar{\epsilon}_0(1+\bar{\epsilon}_s)}{P}}^{\infty} e^{-\lambda_0 x} dx}_{T_{M,3}}. \tag{70}$$

By following steps similar to those to analyze T_m, $1 \leq m \leq M-1$, it is straightforward to show that $T_{M,1} = \mathcal{O}\left(\frac{1}{P^{M+1} \ln^{-M} P} \right)$ and $T_{M,2} = \mathcal{O}\left(\frac{1}{P^M \ln^{-(M-1)} P} \right)$.

What makes the high SNR behaviour of T_M different from those of T_m, $0 \leq m \leq M-1$, is $T_{M,3}$. It is important to point out that the upper end of the integral range of $T_{M,3}$ is ∞, instead of a value which goes to zero for $P \to \infty$. As a result, $\lambda_0 T_{M,3} = e^{-\lambda_0 \frac{\bar{\epsilon}_0(1+\bar{\epsilon}_s)}{P}} \underset{P \to \infty}{\longrightarrow} 1$, instead of $\frac{1}{P}$. Therefore, T_M can be approximated at high SNR as follows:

$$T_M = \mathcal{O}\left(\frac{1}{P^M \ln^{-(M-1)} P} \right). \tag{71}$$

D. Overall High-SNR Approximation

By substituting (50), (66) and (71) in (15), we can conclude that the overall outage probability can be approximated as follows:

$$\mathrm{P}^{WP} = \mathcal{O}\left(\frac{1}{P^M \ln^{-(M-1)} P} \right), \tag{72}$$

for $P \to \infty$. (72) indicates that T_M is the most dominant term in (15) at high SNR.

The diversity gain achieved by WPT-NOMA can be obtained as follows:

$$\begin{aligned} d &= \lim_{P \to \infty} -\frac{\log \mathrm{P}^{WP}}{\log P} = \lim_{P \to \infty} \frac{\log \left(P^M \ln^{-(M-1)} P \right)}{\log P} \\ &= \lim_{P \to \infty} \left[\frac{\log P^M}{\log P} - \frac{\log \ln^{M-1} P}{\log P} \right]. \end{aligned} \tag{73}$$

The following limit holds at high SNR

$$\begin{aligned} \lim_{P \to \infty} \frac{\log \ln^{M-1} P}{\log P} &= \lim_{P \to \infty} \frac{\log e \ln \left(\ln^{M-1} P \right)}{\log e \ln P} \\ &= \lim_{P \to \infty} \frac{M-1}{\ln P} = 0, \end{aligned} \tag{74}$$

where L'Hospital's rule is used. Therefore, the diversity gain achieved by WPT-NOMA can be obtained as follows:

$$d = \lim_{P \to \infty} \frac{\log P^M}{\log P} = M, \tag{75}$$

and the theorem is proved.

Appendix B
Proof for Lemma 1

In order to study the asymptotic behaviour of $\mathrm{P}(E_0)$, EVT is applied in the following. Recall that the limiting CDF of the smallest order statistics should follow one of the three distributions, namely the Frechét type, the modified Weibull type and the extreme value CDF [31, Theorem 8.3.5]. For the considered order statistics, γ_1, the modified Weibull type is applicable as explained in the following.

Denote $F_\gamma^{-1}(a)$ by the inverse function of the CDF of the unordered channel gain, i.e., $F_\gamma \left(F_\gamma^{-1}(a) \right) = a$. The first condition to show that the considered CDF is the modified Weibull type of EVT is that $F_\gamma^{-1}(0)$ should be finite [31, Theorem 8.3.6]. For the considered CDF, we have

$F_\gamma^{-1}(0)=0$ which is indeed finite. The second condition is to show whether the following limitation exists

$$\lim_{\epsilon\to 0^+}\frac{F_\gamma\left(F_\gamma^{-1}(0)+\epsilon x\right)}{F_\gamma\left(F_\gamma^{-1}(0)+\epsilon\right)}=x^{\breve{\alpha}}, \tag{76}$$

for all $x>0$, where $\breve{\alpha}$ denotes a constant parameter.

For the considered CDF, the limitation can be expressed as follows:

$$\lim_{\epsilon\to 0^+}\frac{F_\gamma\left(F_\gamma^{-1}(0)+\epsilon x\right)}{F_\gamma\left(F_\gamma^{-1}(0)+\epsilon\right)} = \lim_{\epsilon\to 0^+}\frac{F_\gamma(\epsilon x)}{F_\gamma(\epsilon)} = \lim_{\epsilon\to 0^+}\frac{1-2\sqrt{\lambda_h\lambda_g\epsilon x}K_1\left(2\sqrt{\lambda_h\lambda_g\epsilon x}\right)}{1-2\sqrt{\lambda_h\lambda_g\epsilon}K_1\left(2\sqrt{\lambda_h\lambda_g\epsilon}\right)}. \tag{77}$$

Note that in (77), x is constant, and the limitation is with respect to ϵ. When $\epsilon\to 0$, the approximation in (42) can be applied and the limitation can be obtained as follows:

$$\lim_{\epsilon\to 0^+}\frac{F_\gamma\left(F_\gamma^{-1}(0)+\epsilon x\right)}{F_\gamma\left(F_\gamma^{-1}(0)+\epsilon\right)} = \lim_{\epsilon\to 0^+}\frac{-\lambda_h\lambda_g\epsilon x\ln(\lambda_h\lambda_g\epsilon x)}{-\lambda_h\lambda_g\epsilon\ln(\lambda_h\lambda_g\epsilon)} = \lim_{\epsilon\to 0^+}\frac{x\ln(\lambda_h\lambda_g\epsilon x)}{\ln(\lambda_h\lambda_g\epsilon)}. \tag{78}$$

By applying L'Hospital's rule, the limitation can be obtained as follows:

$$\lim_{\epsilon\to 0^+}\frac{F_\gamma\left(F_\gamma^{-1}(0)+\epsilon x\right)}{F_\gamma\left(F_\gamma^{-1}(0)+\epsilon\right)} = \lim_{\epsilon\to 0^+}\frac{x\frac{\lambda_h\lambda_g x}{\lambda_h\lambda_g\epsilon x}}{\frac{\lambda_h\lambda_g}{\lambda_h\lambda_g\epsilon}}=x, \tag{79}$$

which means that $\breve{\alpha}=1$ for the considered order statistics.

As a result, the smallest channel gain will follow the modified Weibull type with $\breve{\alpha}=1$, i.e.,

$$\frac{\gamma_1-a_m}{b_m}\sim G_2^*(x;\breve{\alpha}), \tag{80}$$

where $G_2^*(x;\breve{\alpha})$ denotes the modified Weibull distribution:

$$G_2^*(x;\breve{\alpha})\triangleq 1-G_2(-x;\breve{\alpha})=1-e^{-x}, \tag{81}$$

and $G_2(x;\breve{\alpha})$ denotes the Weibull distribution defined as follows:

$$G_2(x;\breve{\alpha})\triangleq\begin{cases}e^{-(-x)^{\breve{\alpha}}}, & x<0\\ 1, & x\geq 0.\end{cases} \tag{82}$$

The two parameters in (80), a_m and b_m are given by

$$a_m\triangleq F_\gamma^{-1}(0)=0, \tag{83}$$

and

$$b_m\triangleq F_\gamma^{-1}\left(\frac{1}{M}\right)-F_\gamma^{-1}(0)=F_\gamma^{-1}\left(\frac{1}{M}\right). \tag{84}$$

The challenging step is to find an explicit expression of b_m, which can be obtained by solving the following equation:

$$1-2\sqrt{\lambda_h\lambda_g b_m}K_1\left(2\sqrt{\lambda_h\lambda_g b_m}\right)=\frac{1}{M}. \tag{85}$$

For $M\to\infty$, we have $\frac{1}{M}\to 0$ and hence $b_m\to 0$. Because $b_m\to 0$, the use of the approximation in (42) can be used to simplify the equation (85) as follows:

$$-\lambda_h\lambda_g b_m\ln(\lambda_h\lambda_g b_m)=\frac{1}{M}. \tag{86}$$

In order to apply the Lambert W function, (86) needs to be written as follows:

$$-\frac{1}{M}=-\frac{1}{M\lambda_h\lambda_g b_m}e^{-\frac{1}{M\lambda_h\lambda_g b_m}}, \tag{87}$$

which means that the solution of (87) can be expressed as follows:

$$-\frac{1}{M\lambda_h\lambda_g b_m}=W\left(-\frac{1}{M}\right), \tag{88}$$

or equivalently

$$b_m=-\frac{1}{M\lambda_h\lambda_g W\left(-\frac{1}{M}\right)}, \tag{89}$$

where $W(\cdot)$ denotes the Lambert W function.

Because $-\frac{1}{M}$ is negative, there are two solutions for $W\left(-\frac{1}{M}\right)$, namely $W_0\left(-\frac{1}{M}\right)$ and $W_{-1}\left(-\frac{1}{M}\right)$ [40]. Recall that $W_0(x)\to 0$ for $x\to 0$, which means that $b_m=-\frac{1}{M\lambda_h\lambda_g W_0\left(-\frac{1}{M}\right)}\to\infty$ for $M\to\infty$. This is contradicted to (84) which indicates that $b_m\to 0$ for $M\to\infty$. Therefore, $W_0\left(-\frac{1}{M}\right)$ is not the solution of the considered case, and we are interested the other branch, $W_{-1}\left(-\frac{1}{M}\right)$. Recall that $W_{-1}(x)$ can be bounded as follows: [40]

$$-1-\sqrt{2u}-u<W_{-1}\left(-e^{-u-1}\right)<-1-\sqrt{2u}-\frac{2}{3}u, \tag{90}$$

for $u>0$.

By applying the bounds, $W_{-1}\left(-\frac{1}{M}\right)$ can be bounded as follows:

$$\ln\frac{1}{M}<W_{-1}\left(-\frac{1}{M}\right)<\frac{2}{3}\ln\frac{1}{M}, \tag{91}$$

which yields the following approximation:

$$W_{-1}\left(-\frac{1}{M}\right)=-\mathcal{O}(\ln M). \tag{92}$$

Therefore, b_m can be approximated as follows:

$$b_m=\frac{1}{\lambda_h\lambda_g M\mathcal{O}(\ln M)}. \tag{93}$$

By applying (83) and (93) to (80), we have $\frac{\gamma_1}{b_m}\sim e^{-x}$ and the limiting CDF of the smallest channel gain is given by

$$F_{\gamma_1}(y)=1-e^{yM\lambda_h\lambda_g W\left(-\frac{1}{M}\right)}, \tag{94}$$

and the corresponding pdf is given by $f_{\gamma_1}(y)=M\lambda_h\lambda_g W\left(-\frac{1}{M}\right)e^{yM\lambda_h\lambda_g W\left(-\frac{1}{M}\right)}$.

By using this pdf, $\mathrm{P}(E_0)$ can be expressed as follows:

$$\begin{aligned}\mathrm{P}(E_0)&=\int_0^\infty\left(e^{-\lambda_0\epsilon_0P^{-1}}-e^{-\lambda_0(\beta^2\epsilon_0x+\epsilon_0P^{-1})}\right)f_{\gamma_1}(x)dx\\&\quad+1-e^{-\lambda_0\epsilon_0P^{-1}}\\&\approx 1+\frac{M\lambda_h\lambda_g W\left(-\frac{1}{M}\right)}{\lambda_0\beta^2\epsilon_0-M\lambda_h\lambda_g W\left(-\frac{1}{M}\right)},\end{aligned}$$

which can be approximated as follows:

$$\begin{aligned}\mathrm{P}(E_0)&\approx 1-\frac{M\lambda_h\lambda_g\mathcal{O}(\ln M)}{\lambda_0\beta^2\epsilon_0+M\lambda_h\lambda_g\mathcal{O}(\ln M)}\\&\approx 1-\frac{1}{1+\frac{\lambda_0\beta^2\epsilon_0}{M\lambda_h\lambda_g\mathcal{O}(\ln M)}}\to\frac{\lambda_0\beta^2\epsilon_0}{\lambda_h\lambda_g M\mathcal{O}(\ln M)},\end{aligned} \tag{95}$$

where the last approximation follows from the fact that $\frac{1}{1+x} \approx 1 - x$ for $x \to 0$. By increasing M, (95) clearly shows that $\mathrm{P}(E_0)$ approaches zero, and the proof for the lemma is complete.

References

[1] *Roadmap for IoT Research, Innovation and Development in Europe*, EU NGIoT, Jan. 2020.

[2] Z. Ding, X. Lei, G. K. Karagiannidis, R. Schober, J. Yuan, and V. K. Bhargava, "A survey on non-orthogonal multiple access for 5G networks: Research challenges and future trends," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 10, pp. 2181–2195, Oct. 2017.

[3] H. Nikopour and H. Baligh, "Sparse code multiple access," in *Proc. IEEE Int. Symp. Pers. Indoor Mobile Radio Commun.*, London, U.K., Sep. 2013, pp. 332–336.

[4] Y. Saito, A. Benjebbour, Y. Kishiyama, and T. Nakamura, "System-level performance evaluation of downlink non-orthogonal multiple access (NOMA)," in *Proc. IEEE 24th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*, London, U.K., Sep. 2013, pp. 611–615.

[5] Z. Ding, R. Schober, P. Fan, and H. V. Poor, "Simple semi-grant-free transmission strategies assisted by non-orthogonal multiple access," *IEEE Trans. Commun.*, vol. 67, no. 6, pp. 4464–4478, Jun. 2019.

[6] J. Zhang, X. Tao, H. Wu, N. Zhang, and X. Zhang, "Deep reinforcement learning for throughput improvement of the uplink grant-free NOMA system," *IEEE Internet Things J.*, vol. 7, no. 7, pp. 6369–6379, Jul. 2020.

[7] J. Choi, "NOMA-based compressive random access using Gaussian spreading," *IEEE Trans. Commun.*, vol. 67, no. 7, pp. 5167–5177, Jul. 2019.

[8] Z. Ding, R. Schober, and H. V. Poor, "A new QoS-guarantee strategy for NOMA assisted semi-grant-free transmission," submitted to, *IEEE Trans. Wireless Commun.* [Online]. Available: http://arxiv.org/abs/2004.12997

[9] Z. Ding, R. Schober, and H. V. Poor, "Unveiling the importance of SIC in NOMA systems—Part 1: State of the art and recent findings," *IEEE Commun. Lett.*, vol. 24, no. 11, pp. 2373–2377, Nov. 2020.

[10] Z. Ding, R. Schober, and H. V. Poor, "Unveiling the importance of SIC in NOMA systems—Part II: New results and future directions," *IEEE Commun. Lett.*, vol. 24, no. 11, pp. 2378–2382, Nov. 2020.

[11] R. Zhang and C. K. Ho, "MIMO broadcasting for simultaneous wireless information and power transfer," *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 1989–2001, May 2013.

[12] X. Lu, P. Wang, D. Niyato, D. I. Kim, and Z. Han, "Wireless networks with RF energy harvesting: A contemporary survey," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 2, pp. 757–789, 2nd Quart., 2015.

[13] Z. Ding *et al.*, "Application of smart antenna technologies in simultaneous wireless information and power transfer," *IEEE Commun. Mag.*, vol. 53, no. 4, pp. 86–93, Apr. 2015.

[14] K. W. Choi *et al.*, "Toward realization of long-range wireless-powered sensor networks," *IEEE Wireless Commun.*, vol. 26, no. 4, pp. 184–192, Aug. 2019.

[15] K. Han and K. Huang, "Wirelessly powered backscatter communication networks: Modeling, coverage, and capacity," *IEEE Trans. Wireless Commun.*, vol. 16, no. 4, pp. 2548–2561, Apr. 2017.

[16] G. Wang, F. Gao, R. Fan, and C. Tellambura, "Ambient backscatter communication systems: Detection and performance analysis," *IEEE Trans. Commun.*, vol. 64, no. 11, pp. 4836–4846, Nov. 2016.

[17] W. Liu, K. Huang, X. Zhou, and S. Durrani, "Next generation backscatter communication: Systems, techniques, and applications," *J. Wireless Commun. Netw.*, vol. 69, pp. 1–10, Dec. 2019.

[18] R. Long, Y.-C. Liang, H. Guo, G. Yang, and R. Zhang, "Symbiotic radio: A new communication paradigm for passive Internet of Things," *IEEE Internet Things J.*, vol. 7, no. 2, pp. 1350–1363, Feb. 2020.

[19] W. Liu, Y.-C. Liang, Y. Li, and B. Vucetic, "Backscatter multiplicative multiple-access systems: Fundamental limits and practical design," *IEEE Trans. Wireless Commun.*, vol. 17, no. 9, pp. 5713–5728, Sep. 2018.

[20] Y. Liu, Z. Ding, M. Elkashlan, and H. V. Poor, "Cooperative non-orthogonal multiple access with simultaneous wireless information and power transfer," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 938–953, Apr. 2016.

[21] A. Agarwal, A. K. Jagannatham, and L. Hanzo, "Finite blocklength non-orthogonal cooperative communication relying on SWIPT-enabled energy harvesting relays," *IEEE Trans. Commun.*, vol. 68, no. 6, pp. 3326–3341, Jun. 2020.

[22] Z. Yang, Z. Ding, P. Fan, and N. Al-Dhahir, "The impact of power allocation on cooperative non-orthogonal multiple access networks with SWIPT," *IEEE Trans. Wireless Commun.*, vol. 16, no. 7, pp. 4332–4343, Jul. 2017.

[23] J. Tang *et al.*, "Decoupling or learning: Joint power splitting and allocation in MC-NOMA with SWIPT," *IEEE Trans. Commun.*, vol. 68, no. 9, pp. 5834–5848, Sep. 2020.

[24] T.-V. Nguyen, V.-D. Nguyen, D. B. da Costa, and B. An, "Hybrid user pairing for spectral and energy efficiencies in multiuser MISO-NOMA networks with SWIPT," *IEEE Trans. Commun.*, vol. 68, no. 8, pp. 4874–4890, Aug. 2020.

[25] H. Zhang, M. Feng, K. Long, G. K. Karagiannidis, V. C. M. Leung, and H. V. Poor, "Energy efficient resource management in SWIPT enabled heterogeneous networks with NOMA," *IEEE Trans. Wireless Commun.*, vol. 19, no. 2, pp. 835–845, Feb. 2020.

[26] P. D. Diamantoulakis, K. N. Pappi, Z. Ding, and G. K. Karagiannidis, "Wireless-powered communications with non-orthogonal multiple access," *IEEE Trans. Wireless Commun.*, vol. 15, no. 12, pp. 8422–8436, Dec. 2016.

[27] J. Guo, X. Zhou, S. Durrani, and H. Yanikomeroglu, "Design of non-orthogonal multiple access enhanced backscatter communication," *IEEE Trans. Wireless Commun.*, vol. 17, no. 10, pp. 6837–6852, Oct. 2018.

[28] F. D. Ardakani and V. W. S. Wong, "Joint reflection coefficient selection and subcarrier allocation for backscatter systems with NOMA," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, May 2020, pp. 1–6.

[29] Y. Liao, G. Yang, and Y.-C. Liang, "Resource allocation in NOMA-enhanced full-duplex symbiotic radio networks," *IEEE Access*, vol. 8, pp. 22709–22720, 2020.

[30] Q. Zhang, L. Zhang, Y.-C. Liang, and P.-Y. Kam, "Backscatter-NOMA: A symbiotic system of cellular and Internet-of-Things networks," *IEEE Access*, vol. 7, pp. 20000–20013, 2019.

[31] B. C. Arnold, N. Balakrishnan, and H. N. Nagaraja, *A First Course in Order Statistics*. Hoboken, NJ, USA: Wiley, 1992.

[32] G. Song and Y. Li, "Asymptotic throughput analysis for channel-aware scheduling," *IEEE Trans. Commun.*, vol. 54, no. 10, pp. 1827–1834, Oct. 2006.

[33] M. Sharif and B. Hassibi, "On the capacity of MIMO broadcast channels with partial side information," *IEEE Trans. Inf. Theory*, vol. 51, no. 2, pp. 506–522, Feb. 2005.

[34] T. Yoo and A. Goldsmith, "On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, pp. 528–540, Mar. 2006.

[35] X. Zhou, R. Zhang, and C. K. Ho, "Wireless information and power transfer: Architecture design and rate-energy tradeoff," *IEEE Trans. Commun.*, vol. 61, no. 11, pp. 4754–4767, Nov. 2013.

[36] M. Medard, "The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel," *IEEE Trans. Inf. Theory*, vol. 46, no. 3, pp. 933–946, May 2000.

[37] M. Ding and S. D. Blostein, "Maximum mutual information design for MIMO systems with imperfect channel knowledge," *IEEE Trans. Inf. Theory*, vol. 56, no. 10, pp. 4793–4801, Oct. 2010.

[38] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products*, 6th ed. New York, NY, USA: Academic, 2000.

[39] Z. Ding, I. Krikidis, B. Sharif, and H. V. Poor, "Wireless information and power transfer in cooperative networks with spatially random relays," *IEEE Trans. Wireless Commun.*, vol. 13, no. 8, pp. 4440–4453, Aug. 2014.

[40] I. Chatzigeorgiou, "Bounds on the lambert function and their application to the outage analysis of user cooperation," *IEEE Commun. Lett.*, vol. 17, no. 8, pp. 1505–1508, Aug. 2013.