A BAC-NOMA Design for 6 G umMTC With Hybrid SIC: Convex Optimization or Learning-Based?

DOI:
10.1109/TVT.2024.3380056

Document Version
Accepted author manuscript

A BAC-NOMA Design for 6G umMTC with Hybrid SIC: Convex Optimization or Learning-based?

Shiyu Jiao, Ximing Xie, Kaidi Wang, Member, IEEE, and Zhiguo Ding, Fellow, IEEE

Abstract—This paper presents a new backscattering communication (BackCom)-assisted non-orthogonal multiple access (BAC-NOMA) transmission scheme for device-to-device (D2D) communications. This scheme facilitates energy and spectrum cooperation between BackCom devices and cellular downlink users in 6G ultra-massive machine-type communications (umMTC) scenarios. Given its quasi-uplink nature, the hybrid successive interference cancellation (SIC) is applied to further improve performance. The data rate of BackCom devices with high quality of service (QoS) requirements is maximized by jointly optimizing backscatter coefficients and the beamforming vector. The use of hybrid SIC and BackCom yields two non-concave sub-problems involving transcendental functions. To address this problem, this paper designs and compares convex optimization-based and unsupervised deep learning-based algorithms. In the convex optimization, the closed-form backscatter coefficients of the first sub-problem are obtained, and then semi-definite relaxation (SDR) is utilized to design the beamforming vector. On the other hand, the second sub-problem is approximated by using a combination of sequential convex approximation (SCA) and SDR. For unsupervised deep learning-based optimization, a loss function is properly designed to satisfy constraints. Computer simulations show the following instructive results: i) the superiority of the hybrid SIC strategy; ii) the distinct sensitivities and efficacies of these two algorithms in response to varying parameters; iii) the superior robustness of the unsupervised deep learning-based optimization.

Index Terms—Non-orthogonal multiple access (NOMA), backscatter communications(BackCom), hybrid SIC, convex optimization, unsupervised deep learning

I. Introduction

Ultra-massive machine type communications (umMTC) is proposed based on mMTC of the fifth-generation (5G) era and has been viewed as one of the key communication scenarios to be realized by the envisioned sixth-generation (6G) mobile networks [1]. The umMTC scenario is expected to serve various devices with different quality of services (QoS), such as wearable devices, high-definition (HD) cameras, smart homes, environment sensors and so on [2]–[4]. However, limited by spectrum and energy, deploying an Internet of Everything (IoE) consisting of an extremely large number of devices faces enormous challenges, which motivate researchers to work on various novel schemes. For enhancing spectrum utilization, non-orthogonal multiple access (NOMA) has been widely used in various wireless communication networks due to its higher spectrum efficiency compared to traditional orthogonal multiple access (OMA). Different from OMA, NOMA allows users non-orthogonally utilize time, frequency and coding but with different power levels [5]. For example, if devices are appended in conventional OMA networks, each of the devices is allocated with a solely occupied resource block regardless of its QoS requirement, which is a spectral inefficient scheme and not helpful to construct 6G umMTC. On the other hand, if different QoS-required devices are added to legacy networks with NOMA protocol, these devices can be admitted to the same channel with legacy users [6], [7], which can construct a high spectral efficiency and umMTC-helpful scheme. In NOMA, users are split in the power domain, where superposition coding is used in transmitters and successive interference cancellation (SIC) is used in receivers to remove interferences. Therefore, optimal power allocation for NOMA systems with individual QoS constraints[1] is very important [8]. For uplink NOMA, there are generally two types of SIC, including channel state information (CSI)-based SIC and QoS-based SIC [9]. CSI-based SIC is an intuitive decoding strategy which decides the decoding order on the basis of users' channel conditions but it can not provide performance gain when NOMA users' channel quality is similar. QoS-based SIC, as the name implies, implements SIC according to users' QoS requirements, for example, in cognitive-radio NOMA networks, the primary user's signal can be decoded first by guaranteeing the QoS threshold [10]. However, it has the risk of decoding failure when the first decoded signal can not meet its target rate. The proposal of hybrid SIC not only conquers those shortcomings [9] but also guides the selection of decoding order. The key idea of hybrid SIC is to dynamically select the optimal decoding order that achieves the maximum data rate. For instance, consider a two-user uplink NOMA scenario where a secondary user is admitted to share the same resource block with a primary user. If the CSI-based SIC is used, the signal of the user having a higher channel gain will be decoded first, which yields two decoding orders. If the QoS-based SIC is used, the primary user's QoS is preferentially guaranteed. In particular, if the primary user's signal can be successfully decoded first after

Shiyu Jiao, Ximing Xie and Kaidi Wang are with School of Electrical and Electronic Engineering, The University of Manchester, M1 9BB, UK, (e-mail: jiao.shiyu@outlook.com, ximing.xie@manchester.ac.uk and kaidi.wang@ieee.org).

Zhiguo Ding is with 6G center, Khalifa University, Abu Dhabi, UAE, and the University of Manchester, Manchester, M1 9BB, UK, (zhiguo.ding@ieee.org).

[1]The reflection coefficients optimization problem in this study is indeed an optimal power allocation for NOMA systems.

the secondary user is added (i.e., decoding the primary user's signal under the secondary user's interference), the decoding order will be decoding the primary user's signal first and then the secondary user's signal. In this case, hybrid SIC can improve system performance by dynamically selecting the decoding order in different scenarios.

Energy limitation is another issue besides spectrum utilization. For the purpose of saving energy or prolonging battery life, a variety of schemes that enable energy cooperation have been proposed, such as, wireless power transfer (WPT) and backscattering communication (BackCom) [11]. The key idea of WPT is to harvest the energy of the radio frequency (RF) sent by a power station or other non-energy-constrained devices, where the harvested energy is utilized to transmit signals. In [12], the energy consumption problem in Machine-to-Machine communications with multiple access and non-linear energy harvesting for IoT, in which an energy effective resource allocation algorithm has been proposed. BackCom was first proposed in [13]. It has also been regarded as a mature and efficient technology to realize energy cooperation in 6G [14]. The key idea of BackCom is to utilize the signal sent by non-energy-constrained devices to activate the circuit of BackCom devices (BDs), and then BDs will re-modulate their signals and reflect them [15]–[17]. In the future 6G networks, BackCom has been regarded as a promising technique for various applications, especially for green communications and the Internet of things (IoT), due to its advantages of high energy efficiency and spectrum friendly [18]–[20]. Therefore, BAC-NOMA is a wise combination that can simultaneously enhance spectrum and energy efficiency.

A. Related Works

BAC-NOMA is a wise combination that can simultaneously realize spectrum and energy cooperation and thus attracts a large number of researchers. In [21], the minimum throughput among all BDs is maximized by jointly optimizing backscatter time durations and power reflection coefficients, where multiple BDs and a backscatter receiver are considered. The energy efficiency of a green network is maximized for a Backcom-enabled intelligent reflection surface (IRS)-NOMA network [22] by jointly optimizing the phase shifts of IRS and power allocation of the base station (BS). In this system, a BD, two downlink users and an IRS are considered, where the BackCom is employed as the secondary transmission of a downlink user, and the IRS is deployed to assist another downlink user. The authors in [23] maximized the energy efficiency of a BackCom-enabled two-user NOMA downlink network, subject to users' QoS requirements. In this work, a Dinkelbach-based algorithm is applied to iteratively optimize the power allocation of the BS and the reflection coefficient of the BD. [24] analyzed the outage probability of a simple downlink BAC-NOMA network over the Nakagami fading channel, where a downlink user is served simultaneously by the BS and the BD. [25] highlighted the benefits of utilizing BAC-NOMA in legacy orthogonal frequency division multiple access (OFDMA)-based and spatial division multiple access (SDMA)-based networks, where the reflection coefficients of the two schemes are both optimized for the fair comparison.

B. Motivations and Contributions

Although BAC-NOMA has been extensively investigated, research in this area that can support 6G umMTC is still rare, which motivates us to propose a novel and easy-to-implement [2] BackCom-based scheme for legacy cellular networks that can realize spectrum and energy cooperation. Therefore, this paper considers the deployment of BAC-NOMA-enabled device-to-device (D2D) communications to legacy OMA networks. Specifically, two BDs and a data fusion center are admitted to share the same resource block with a legacy OMA downlink user. With the help of optimizations, the QoS of the legacy OMA user can be guaranteed while the spectrum and energy cooperation can also be realized. The quasi-uplink characteristic of BackCom-based D2D communications inspires us to use hybrid SIC to further improve its performance. However, several challenges are also introduced, for example, BackCom brings the transcendental function and hybrid SIC discretizes the optimization problem, which makes the original problem non-concave and complicated. To decompose the original problem, two subproblems are derived. For the first subproblem, the optimal solution can be achieved by using convex optimization. However, the second one can only be solved by applying approximation. This further motivates us to use the learning-based algorithm and compare it with conventional convex optimization.

This paper aims to maximize the high QoS required BD's data rate and explore the differences between the unsupervised deep learning-based optimization and convex optimization in solving problems corresponding to different decoding orders. The main contributions of this paper are summarized as follows:

- A novel BAC-NOMA D2D transmission scheme is proposed, where hybrid SIC is applied for D2D communication. This scheme can realize energy and spectrum cooperation with legacy OMA users and is easy to implement in legacy OMA networks. The data rate of a high data rate required BD is maximized by guaranteeing the low data rate required BD's and the legacy OMA user's QoS.

[2]The term "easy-to-implementation" implies that comparing integrating more devices by using the conventional OMA scheme, employing the proposed BAC-NOMA-based scheme is less complex and incurs lower overhead. If more devices are integrated into the legacy OMA system with the OMA scheme, new devices will carve up the OMA user's resource block, consequently causing a substantial reduction in the QoS for the legacy OMA users. Moreover, the original communication system needs to be re-designed, for example, the signal processing flow, the protocol and resource scheduling. In contrast, the deployment of extra devices via our proposed scheme only requires the BS to execute the proposed algorithm based on the global CSI.

- To efficiently address the non-concave problem, it is divided into two sub-problems according to different decoding orders. For the first sub-problem, after proving the monotonicity of the transcendental objective function, the coupling between variables is avoided by deriving the optimal closed-form backscattering coefficients. Subsequently, semi-definite relaxation (SDR) is utilized to optimize its beamforming vector. For the second sub-problem, the method of solving the first subproblem is no longer available, and hence the successive convex approximation (SCA) and SDR are employed for the joint beamforming and reflection coefficient design. Meanwhile, the optimality of the two sub-problem is analyzed.
- Unsupervised deep learning is applied to solve the problem, where the loss function is dexterously designed. The constraints can be guaranteed by the penalty term in the loss function. This method is broadly applicable to many optimization problems.
- Simulations demonstrate the following fascinating findings: 1) convex optimization and unsupervised deep learning-based optimization have different sensitivities to different parameters; 2) unsupervised deep learning-based optimization is more robust than convex optimization when channels are outdated. This provides guidance for the selection of optimization methods in real-world design.

C. Organisation

The rest of this paper is organized as follows. Section II describes the system model and formulates it as an optimization problem with hybrid SIC. In section III, the problem is solved by using convex optimization. In section IV, unsupervised deep learning-based optimization is discussed. Section V demonstrates simulation results and compares the corresponding performances of the two optimization paths. Finally, the conclusion is provided in section VI.

II. System Model and Problem Formulation

This paper proposes a communication scenario with one BS, two BackCom transmitters denoted by B_1 and B_2, one data fusion centre denoted by B_r and a legacy OMA downlink user U_0, where B_1 is assumed to be a high data rate required device while B_2 only requires low data rate. This is a common scenario in 6G networks [26]–[28]. For example, B_1 is a high-definition (HD) that is uploading videos and B_2 is a healthcare wearable device that needs a stable low data rate. Assume that the BDs have been admitted to share the same communication resources with U_0 [3]. In this network, the BS is equipped with M antennas while two BDs and the downlink user are equipped with a signal antenna. Because BDs B_1 and B_2 transmit their own data simultaneously on the same spectrum to B_r, it can be viewed as an uplink BAC-NOMA scenario. For BAC-NOMA, BDs' circuit will be excited by the signal sent from the BS, i.e., $x_0 = \boldsymbol{\omega}_0 s_0$, where s_0 is the desired signal of U_0, satisfying $\mathbb{E}\{|s_0|^2\} = 1$, where $\{\cdot\}$ denotes the expectation operation, and $\boldsymbol{\omega}_0 \in \mathbb{C}^{M\times 1}$ is the corresponding beamforming vector. According to the ambient BackCom's principle, BDs will modulate and reflect their own information over the incident signal x_0. Therefore, the BackCom signal reflected by $B_n, \forall n \in \{1,2\}$ can be represented as $x_n = \boldsymbol{h}_n^H \boldsymbol{\omega}_0 s_0 s_n$, where $\boldsymbol{h}_n \in \mathbb{C}^{M\times 1}$ denotes the channel from the BS to B_n and s_n is the transmit signal of B_n, also satisfying $\mathbb{E}\{|s_n|^2\} = 1, \forall n \in \{1,2\}$. Due to the spectrum sharing mechanism, the admitted BAC-NOMA transmission and the legacy downlink transmission will mutually interfere with each other. Therefore, the received signal at B_r is given by

$$y_{B_r} = \underbrace{\sum_{n=1}^{2} \sqrt{\eta_n} g_n \boldsymbol{h}_n^H \boldsymbol{\omega}_0 s_0 s_n}_{\text{desired signal}} + \underbrace{\boldsymbol{h}_B^H \boldsymbol{\omega}_0 s_0}_{\text{interference}} + n_B, \quad (1)$$

where η_n is B_n's BackCom reflection coefficient, g_n denotes the channel gain between B_r and B_n, $\boldsymbol{h}_B \in \mathbb{C}^{M\times 1}$ denotes the channel vector from the BS to B_r and n_B denotes the noise which follows $\mathcal{CN}(0, \sigma^2)$ distribution. For U_0, its received signal is given as follows

$$y_0 = \underbrace{\boldsymbol{h}_0^H \boldsymbol{\omega}_0 s_0}_{\text{desired signal}} + \underbrace{\sum_{n=1}^{2} \sqrt{\eta_n} g_{n,0} \boldsymbol{h}_n^H \boldsymbol{\omega}_0 s_0 s_n}_{\text{interference}} + n_0, \quad (2)$$

where $\boldsymbol{h}_0 \in \mathbb{C}^{M\times 1}$ is the channel vector between the BS and U_0 and $g_{n,0}$ is the channel between B_n and U_0. Since s_n is unknown at U_0, the second term of (2) is treated as interference. On the other hand, due to the double fading effect, the direct link signal is much stronger than the backscatter link signal, a proper decoding order at B_r and U_0 is employed to decode s_0 first and then the BackCom signals [29]. By this principle, U_0's data rate is given by

$$R_0 = \log\left(1 + \frac{|\boldsymbol{h}_0^H \boldsymbol{\omega}_0|^2}{\sum_{n=1}^{2} \eta_n |g_{n,0}|^2 |\boldsymbol{h}_n^H \boldsymbol{\omega}_0|^2 + \sigma^2}\right). \quad (3)$$

[3]The proposed system and its corresponding algorithms can be readily expanded to the multi-user multi-BD scenario in which BDs and legacy OMA users are pre-grouped. However, it's important to note that grouping an excessive number of BDs with a single OMA user is not practical due to two key reasons: 1) The BDs and OMA user sharing the same resource block in a NOMA setup will cause significant interference for the OMA user. This necessitates the BS performing an extensive number of SIC to ensure the QoS of the OMA user. 2) As the number of BDs increases, the number of potential SIC decoding sequences and corresponding optimization problems also increases in a factorial manner, requiring the BS to solve these optimization problems simultaneously, which yields extremely high complexity. Both factors contribute to considerable latency.

Unlike U_0, B_r needs to perform SIC to remove interferences and decode the desired signal. To remove the interference, B_r needs to decode s_0 first and remove it. The data rate of s_0 observed at B_r is given by

$$R_{0\to B} = \log\left(1 + \frac{|\boldsymbol{h}_B^H \boldsymbol{\omega}_0|^2}{\sum_{n=1}^{2} \eta_n |g_n|^2 |\boldsymbol{h}_n^H \boldsymbol{\omega}_0|^2 + \sigma^2}\right). \quad (4)$$

In the next stage, B_r will still perform SIC to decode its desired signal (i.e.,s_1 and s_2). In this case, the transmission between B_r and BDs can be viewed as an uplink NOMA scenario, and hence the hybrid SIC strategy [9] can be utilized to further improve the performance. According to the hybrid SIC strategy, the following two decoding orders need to be considered:

1) SIC decoding order 1: If s_1 is decoded first, B_1's achievable rate is given by

$$R_1^{(1)} = \log_2\left(1 + \frac{\eta_1 |g_1|^2 |\boldsymbol{h}_1^H \boldsymbol{\omega}_0|^2 |s_0|^2}{\eta_2 |g_2|^2 |\boldsymbol{h}_2^H \boldsymbol{\omega}_0|^2 |s_0|^2 + \sigma^2}\right). \quad (5)$$

After implementing SIC, B_2's achievable rate is given by

$$R_2^{(1)} = \log_2\left(1 + \frac{\eta_2 |g_2|^2 |\boldsymbol{h}_2^H \boldsymbol{\omega}_0|^2 |s_0|^2}{\sigma^2}\right). \quad (6)$$

2) SIC decoding order 2: If s_2 is decoded first, the achievable rate of B_1 and B_2 can be respectively represented as

$$R_1^{(2)} = \log_2\left(1 + \frac{\eta_1 |g_1|^2 |\boldsymbol{h}_1^H \boldsymbol{\omega}_0|^2 |s_0|^2}{\sigma^2}\right), \quad (7)$$

and

$$R_2^{(2)} = \log_2\left(1 + \frac{\eta_2 |g_2|^2 |\boldsymbol{h}_2^H \boldsymbol{\omega}_0|^2 |s_0|^2}{\eta_1 |g_1|^2 |\boldsymbol{h}_1^H \boldsymbol{\omega}_0|^2 |s_0|^2 + \sigma^2}\right), \quad (8)$$

where the superscript denotes the decoding order case while the subscript indicates the BDs' index. In BackCom, the signal s_0 is perfectly known at B_r according to (4). Subsequently, s_0 can be viewed as the fast-fading channel component when B_r is decoding s_n [30]. Therefore, optimizing the average data rate (i.e., the expectation of the uplink data rate with respect to $|s_0|^2$) is a more appropriate choice, as discussed in [16]. The average data rate of B_1 of the SIC decoding order 1 can be represented by

$$\bar{R}_1^{(1)} = \mathbb{E}_{s_0}\{R_1^{(1)}\} = \int_0^{\infty} R_1^{(1)}(|s_0|^2) f_{|s_0|^2}(x) dx, \quad (9)$$

where $f_{|s_0|^2}(x)$ is the probability density function (pdf) of $|s_0|^2$. The average data rate of B_1 and B_2 for other cases can be obtained in the same way. Assume that s_0 and s_n both follow complex Gaussian distribution with zero mean and unit variance, their envelope will follow the exponential distribution, and hence the pdf of $|s_0|^2$ is $f_{|s_0|^2}(x) = e^{-x}$. With some algebraic manipulation, $\bar{R}_1^{(1)}$,$\bar{R}_2^{(1)}$,$\bar{R}_1^{(2)}$ and $\bar{R}_2^{(2)}$ can be derived as follows:

$$\begin{aligned}\bar{R}_1^{(1)} &= \log_2(e) \\ &\left[-e^{\frac{\sigma^2}{\sum_{n=1}^{2} \eta_n |g_n|^2 |h_n^H \boldsymbol{\omega}_0|^2}} E_i\left(-\frac{\sigma^2}{\sum_{n=1}^{2} \eta_n |g_n|^2 |h_n^H \boldsymbol{\omega}_0|^2}\right)\right. \\ &\left. + e^{\frac{\sigma^2}{\eta_2 |g_2|^2 |h_2^H \boldsymbol{\omega}_0|^2}} E_i\left(-\frac{\sigma^2}{\eta_2 |g_2|^2 |h_2^H \boldsymbol{\omega}_0|^2}\right)\right], \end{aligned} \quad (10)$$

$$\bar{R}_2^{(1)} = -\log_2(e) e^{\frac{\sigma^2}{\eta_2 |g_2|^2 |\boldsymbol{h}_2^H \boldsymbol{\omega}_0|^2}} E_i\left(-\frac{\sigma^2}{\eta_2 |g_2|^2 |\boldsymbol{h}_2^H \boldsymbol{\omega}_0|^2}\right), \quad (11)$$

$$\bar{R}_1^{(2)} = -\log_2(e) e^{\frac{\sigma^2}{\eta_1 |g_1|^2 |\boldsymbol{h}_1^H \boldsymbol{\omega}_0|^2}} E_i\left(-\frac{\sigma^2}{\eta_1 |g_1|^2 |\boldsymbol{h}_1^H \boldsymbol{\omega}_0|^2}\right), \quad (12)$$

$$\begin{aligned}\bar{R}_2^{(2)} &= \log_2(e) \\ &\left[-e^{\frac{\sigma^2}{\sum_{n=1}^{2} \eta_n |g_n|^2 |h_n^H \boldsymbol{\omega}_0|^2}} E_i\left(-\frac{\sigma^2}{\sum_{n=1}^{2} \eta_n |g_n|^2 |h_n^H \boldsymbol{\omega}_0|^2}\right)\right. \\ &\left. + e^{\frac{\sigma^2}{\eta_1 |g_1|^2 |h_1^H \boldsymbol{\omega}_0|^2}} E_i\left(-\frac{\sigma^2}{\eta_1 |g_1|^2 |h_1^H \boldsymbol{\omega}_0|^2}\right)\right], \end{aligned} \quad (13)$$

where $E_i(x) \triangleq \int_{-\infty}^{x} (\frac{1}{u}) e^u \quad du, x \leq 0$ denotes an exponential integral function [31]. Since B_1 demands a high data rate, this paper aims to maximize B_1's uplink data rate while guaranteeing the QoS of the legacy cellular user U_0 and the low data rate required BD B_2.[4] By using hybrid SIC, the optimization problem can be formulated as follows:

$$\begin{aligned} \text{P1:} \max_{\{\eta_n, \boldsymbol{\omega}_0, u_n\}} \quad & \sum_{n=1}^{2} u_n \bar{R}_1^{(n)} && (14a) \\ \text{s.t.} \quad & \sum_{n=1}^{2} u_n \bar{R}_2^{(n)} \geq R_{2,t} && (14b) \\ & \min\{R_0, R_{0\to B}\} \geq R_{0,t}, && (14c) \\ & 0 \leq \eta_n \leq 1, \quad n = 1, 2 && (14d) \\ & |\boldsymbol{\omega}_0|^2 \leq P_t, && (14e) \\ & \sum_{n=1}^{2} u_n = 1. && (14f) \end{aligned}$$

where $R_{2,t}$ and $R_{0,t}$ denote the target rate of B_2 and U_0, P_t represents maximum transmit power of the BS, respectively. $u_n, n = 1, 2$ is a binary variable which controls the decoding order selection.

$$\begin{cases} u_1 = 1, u_2 = 0, & \text{if } s_1 \text{ is decoded first} \\ u_1 = 0, u_2 = 1, & \text{if } s_2 \text{ is decoded first} \end{cases} \quad (15)$$

[4]While the algorithmic description may initially appear to prioritize the rate maximization of User B_1, it is important to note that the optimization framework is inherently designed to consider the performance requirements of both users. This is achieved through carefully structured constraints to ensure that user B_2's QoS is not compromised. The balance of data rates and service quality between B_1 and B_2 is a fundamental characteristic of our proposed approach. Note that weight sum rate maximization can also provide a flexible trade-off between B_1 and B_2, but this framework does not meet the needs of the considered system.

(14b) is to guarantee B_2's QoS and (14c) is to guarantee U_0's QoS and the implementation of SIC. (14e) is the power constraint of the BS. (14f) indicates that only one decoding order can be selected.

The problem P1 is non-convex and can be viewed as a discrete problem. In order to solve the problem effectively, it is divided into two sub-problems according to two different decoding orders and solved in a parallel manner at the BS. Based on the solutions, the BS can decide which decoding order should be selected and inform it to Br through the control channel. Note that the BS has sufficient energy and computing resources in practical communication systems, thus the proposed scheme is possible to be realized. To efficiently solve the problem P1, it is decomposed as follows:

1) For the first case, SIC decoding order 1 is selected (i.e., B_1's signal is decoded first) which derives the problem P2 as follows:

$$\text{P2:} \max_{\{\eta_1,\eta_2,\boldsymbol{\omega}_0\}} \quad \bar{R}_1^{(1)} \tag{16a}$$
$$\text{s.t.} \quad \bar{R}_2^{(1)} \geq R_{2,t}, \tag{16b}$$
$$\text{(14c), (14d) and (14e).} \tag{16c}$$

2) For the second case, SIC decoding order 2 is selected (i.e., B_2's signal is decoded first) which derives the problem P3 as follows:

$$\text{P3:} \max_{\{\eta_1,\eta_2,\boldsymbol{\omega}_0\}} \quad \bar{R}_1^{(2)} \tag{17a}$$
$$\text{s.t.} \quad \bar{R}_2^{(2)} \geq R_{2,t}, \tag{17b}$$
$$\text{(14c), (14d) and (14e).} \tag{17c}$$

III. Convex Optimization-based Algorithm Design

A. Optimal Solution for SIC Decoding Order 1

At first glance, these two questions might seem remarkably similar, but in reality, they are substantially different and cannot be addressed by the same approach, especially for applying convex optimization. The reasons are briefly described as follows. The objective function of P2 (i.e. (18a)) is monotonically decreasing with respect to η_2, as proved in Lemma 1, while the objective function of P3 (i.e., (22a)) is monotonically increasing with respect to η_1 and even does not include η_2. On the other hand, P2 and P3 have different convexity. $\bar{R}_2^{(1)}$ is a concave function of η_2 [25] while $\bar{R}_2^{(2)}$ is a convex function of η_2, which means (18b) is a convex set but (22b) is not. Different convex properties and monotonicity prevent them from being solved by the same method. Assisted by the monotonicity analysis, the closed-form expression of η_2 and η_2 in P2 can be derived, and then the beamforming vector $\boldsymbol{\omega}_0$ can be solved by using SDR. However, P3 can only be solved by a SCA and SDR-based alternating algorithm. A more detailed discussion is provided in the following two subsections. Attention is initially drawn to the decoding order 1, designated as P2.

$$\text{P2:} \max_{\{\eta_1,\eta_2,\boldsymbol{\omega}_0\}} \quad \bar{R}_1^{(1)} \tag{18a}$$
$$\text{s.t.} \quad \bar{R}_2^{(1)} \geq R_{2,t}, \tag{18b}$$
$$\min\{R_0, R_{0\to B}\} \geq R_{0,t}, \tag{18c}$$
$$0 \leq \eta_n \leq 1, \quad n = 1, 2, \tag{18d}$$
$$|\boldsymbol{\omega}_0|^2 \leq P_t, \tag{18e}$$

Lemma 1. $f(x) \triangleq -e^{\frac{1}{a+x}} E_i(-\frac{1}{a+x}) + e^{\frac{1}{x}} E_i(-\frac{1}{x}), a \geq 0, x \geq 0$ is a monotonically decreasing function of x.

Proof. Please refer to Appendix A. □

In this quasi-uplink BAC-NOMA, with the decoding order 1, the QoS constraint of B_2 is still indispensable even if it can enjoy no-interference decoding. The reasoning behind this is articulated as follows: According to Lemma 1, it can be deduced that (10) is a monotonically decreasing function of η_2, which implies that $\bar{R}_1^{(1)}$ can reach its maximum value when η_2 approaches its lower limit. In other words, if the QoS of B_2 is not assured, B_1 could reach its maximum data rate by preventing B_2 from backscattering its signal (i.e., $\eta_2 = 0$), which is impractical in real-world scenarios. This assertion can be further manifested in the proof of the following proposition.

Proposition 1. For given $\boldsymbol{\omega}_0$, the optimal reflection coefficient of B_1 and B_2 of problem P2 can be expressed as

$$\begin{cases} \eta_1^* = \min\{b_1, b_2, 1\}, \\ \eta_2^* = \frac{\beta}{|g_2|^2 |\boldsymbol{h}_2^H \boldsymbol{\omega}_0|^2}, \end{cases} \tag{19}$$

where $b_1 = \frac{|\boldsymbol{h}_0^H \boldsymbol{\omega}_0|^2 - \gamma_{0,t}(\frac{|g_{2,k}|^2}{|g_2|^2}\beta + \sigma^2)}{\gamma_{0,t}|g_{1,k}|^2 |\boldsymbol{h}_1^H \boldsymbol{\omega}_0|^2}$, $b_2 = \frac{|\boldsymbol{h}_B^H \boldsymbol{\omega}_0|^2 - \gamma_{0,t}(\beta+\sigma^2)}{\gamma_{0,t}|g_1|^2 |\boldsymbol{h}_1^H \boldsymbol{\omega}_0|^2}$ and β is obtained by solving $-e^{\frac{\sigma^2}{\beta}} E_i(-\frac{\sigma^2}{\beta}) = \ln(2) R_{2,t}$.

Proof. Please refer to Appendix B. □

By substituting η_2^* and η_1^*, P2 can be rewritten as follows:

$$\text{P2-1:} \max_{\{\boldsymbol{\omega}_0\}} \quad \eta_1^* |g_1|^2 |\boldsymbol{h}_1^H \boldsymbol{\omega}_0|^2 \tag{20a}$$
$$\text{s.t.} \quad 0 \leq \min\{b_1, b_2, 1\}, \tag{20b}$$
$$\frac{\beta}{|g_2|^2 |\boldsymbol{h}_2^H \boldsymbol{\omega}_0|^2} \leq 1, \tag{20c}$$
$$|\boldsymbol{\omega}_0|^2 \leq P_t. \tag{20d}$$

Considering the constraint (20b), the following three cases need to be discussed:

1) $\eta_1^* = b_1$, $b_1 \leq b_2$ and $0 \leq b_1 \leq 1$.
2) $\eta_1^* = b_2$, $b_2 \leq b_1$ and $0 \leq b_2 \leq 1$.
3) $\eta_1^* = 1$, $1 \leq b_1$ and $1 \leq b_2$.

Based on the above cases, three optimization problems can be derived. Note that all the three cases can be solved in a similar way, and then case 1 is described as an example. By

using the SDR technique to ignore the rank one constraint [32], P2-1 can be recast as

$$\text{P2-2:} \max_{\{\boldsymbol{W}\}} \quad \text{Tr}(\boldsymbol{H}_0\boldsymbol{W}) \tag{21a}$$
$$\text{s.t.} \quad \text{Tr}(\boldsymbol{H}_0\boldsymbol{W})|g_1|^2 - \text{Tr}(\boldsymbol{H}_B\boldsymbol{W})|g_{1,k}|^2 + a \le 0, \tag{21b}$$
$$\text{Tr}(\boldsymbol{H}_0\boldsymbol{W}) - \gamma_{0,t}|g_{1,k}|^2\text{Tr}(\boldsymbol{H}_1\boldsymbol{W}) + b \le 0, \tag{21c}$$
$$\text{Tr}(\boldsymbol{H}_0\boldsymbol{W}) - b \ge 0, \tag{21d}$$
$$\beta - |g_2|^2\text{Tr}(\boldsymbol{H}_2\boldsymbol{W}) \le 0, \tag{21e}$$
$$\text{Tr}(\boldsymbol{W}) \le P_t, \tag{21f}$$
$$\boldsymbol{W} \succeq 0, \tag{21g}$$

where $\boldsymbol{H}_0 = \boldsymbol{h}_0\boldsymbol{h}_0^H$ and $\boldsymbol{W} = \boldsymbol{\omega}_0\boldsymbol{\omega}_0^H$ are two rank one symmetric positive semidefinite (PSD) matrix, $a = \gamma_{0,t}(\beta(|g_{1,k}|^2 - \frac{|g_1|^2|g_{2,k}|^2}{|g_2|^2}) + \sigma^2(|g_{1,k}|^2 - |g_1|^2))$ and $b = \gamma_{0,t}(\frac{|g_{2,k}|^2}{|g_2|^2}\beta + \sigma^2)$. P2-2 is a semidefinite programming (SDP) and can be solved by convex optimization tools. Afterward, $\boldsymbol{\omega}_0$ can be reconstructed by applying Gaussian randomization procedure if $\text{rank}(\boldsymbol{W}) \neq 1$ [32] or by eigenvalue decomposition (EVD) if $\text{rank}(\boldsymbol{W}) = 1$. Finally, the solution corresponding to the maximum $\bar{R}_1^{(1)}$ is chosen as the optimal solution of P2.

B. Suboptimal Solution for SIC Decoding Order 2

Recall the problem related to the second decoding order:

$$\text{P3:} \max_{\{\eta_1,\eta_2,\boldsymbol{\omega}_0\}} \quad \bar{R}_1^{(2)} \tag{22a}$$
$$\text{s.t.} \quad \bar{R}_2^{(2)} \ge R_{2,t}, \tag{22b}$$
$$\text{(14c), (14d) and (14e).} \tag{22c}$$

P3 is a non-concave problem even when SDR is applied, because a convex function $\bar{R}_2^{(2)}(\eta_n, \boldsymbol{W})$ greater than or equal to zero is not a convex set, i.e., (22b) is not convex. Additionally, due to the more complicated transcendental function in (22b), the approach employed to derive closed-form solutions for η_1 and η_2 with respect to $\boldsymbol{\omega}_0$ in P2 is no longer applicable to P3. Therefore, SDR is combined with SCA to iteratively approximate (22b). A detailed discussion for solving P3 is provided in the following.

First, backscattering coefficient optimization is studied. Denote $[\eta_1, \eta_2]^T$ by $\boldsymbol{\eta}$, (11) and (13) can be rewritten as:

$$\bar{R}_1^{(2)} = -\log_2(e)\left[-e^{\frac{\sigma^2}{\boldsymbol{\eta}^T\boldsymbol{B}}}E_i\left(-\frac{\sigma^2}{\boldsymbol{\eta}^T\boldsymbol{B}}\right)\right], \tag{23}$$

and

$$\bar{R}_2^{(2)} = -\log_2(e)\left[e^{\frac{\sigma^2}{\boldsymbol{\eta}^T\boldsymbol{A}}}E_i\left(-\frac{\sigma^2}{\boldsymbol{\eta}^T\boldsymbol{A}}\right) + e^{\frac{\sigma^2}{\boldsymbol{\eta}^T\boldsymbol{B}}}E_i\left(-\frac{\sigma^2}{\boldsymbol{\eta}^T\boldsymbol{B}}\right)\right], \tag{24}$$

where $\boldsymbol{A} = [|g_1|^2|\boldsymbol{h}_1^H\boldsymbol{\omega}_0|^2, |g_2|^2|\boldsymbol{h}_2^H\boldsymbol{\omega}_0|^2]^T = [|g_1|^2\text{Tr}(\boldsymbol{H}_1\boldsymbol{W}), |g_2|^2\text{Tr}(\boldsymbol{H}_2\boldsymbol{W})]^T$ and $\boldsymbol{B} = [|g_1|^2|h_1^H\boldsymbol{\omega}_0|^2, 0]^T = [|g_1|^2\text{Tr}(\boldsymbol{H}_1\boldsymbol{W}), 0]^T$. By defining $\bar{R}_2^{(2)}(\boldsymbol{\eta}) \triangleq f(\boldsymbol{\eta})$ and applying the fact that $f(\frac{1}{x}) = -e^{\frac{1}{x}}Ei(-\frac{1}{x})$ is a monotonically increasing function of x, for a given $\boldsymbol{W}$, P3 can be recast as follows:

$$\text{P3-1:} \max_{\{\boldsymbol{\eta}\}} \quad \boldsymbol{\eta}^T\boldsymbol{B} \tag{25a}$$
$$\text{s.t.} \quad f(\boldsymbol{\eta}_0) + (\boldsymbol{\eta} - \boldsymbol{\eta}_0)f'(\boldsymbol{\eta}_0) \ge R_{2,t}\ln(2), \tag{25b}$$
$$\text{Tr}(\boldsymbol{H}_0\boldsymbol{W}) - \gamma_{0,t}(\boldsymbol{\eta}^T\boldsymbol{C} + \sigma^2) \ge 0, \tag{25c}$$
$$\text{Tr}(\boldsymbol{H}_B\boldsymbol{W}) - \gamma_{0,t}(\boldsymbol{\eta}^T\boldsymbol{A} + \sigma^2) \ge 0, \tag{25d}$$
$$0 \preceq \boldsymbol{\eta} \preceq 1, \tag{25e}$$

where $\boldsymbol{C} = [|g_{1,k}|^2\text{Tr}(\boldsymbol{H}_1\boldsymbol{W}), |g_{2,k}|^2\text{Tr}(\boldsymbol{H}_2\boldsymbol{W})]^T$ and $f'(\boldsymbol{\eta}_0)$ represents the first-order derivative of $f(\boldsymbol{\eta})$ at $\boldsymbol{\eta}_0$. After the first-order Taylor expansion, (25b) becomes a linear function of $\boldsymbol{\eta}$, greater than or equal to zero, which is a convex set. P3-1 is a linear programming (LP) and can be solved by convex optimization tools.

Subsequently, the SDR and SCA are used to design the beamforming of the BS. With the given $\boldsymbol{\eta}$, P3 can be reduced to

$$\text{P3-2:} \max_{\{\boldsymbol{W}\}} \quad \eta_1|g_1|^2\text{Tr}(\boldsymbol{H}_1\boldsymbol{W}) \tag{26a}$$
$$\text{s.t.} \quad f(\boldsymbol{W}_0) + \text{Tr}((\boldsymbol{W} - \boldsymbol{W}_0)f'(\boldsymbol{W}_0)) \ge R_{2,t}\ln(2), \tag{26b}$$
$$\text{(25c) and (25d),} \tag{26c}$$
$$\text{Tr}(\boldsymbol{W}) \le P_t, \tag{26d}$$
$$\boldsymbol{W} \succeq 0, \tag{26e}$$

where $f'(\boldsymbol{W}_0)$ is the first-order derivative of $f(\boldsymbol{W})$ at $\boldsymbol{W}_0$. P3-2 is a concave optimization problem with respect to $\boldsymbol{W}$. As a result, P3 can be resolved by alternatively solving P3-1 and P3-2. In the end, the same as P2-2, Gaussian randomization or EVD can be applied to retrieve the beamforming vector $\boldsymbol{\omega}_0$. Thus far, the variables $\boldsymbol{\eta}$ and $\boldsymbol{\omega}_0$ of problem P3 have been solved independently. For the purpose of iteritive optimization, the alternating operation of these two algorithms is needed to be conducted.

The maximum data rate of B_1 with its corresponding optimal decoding order and solution can be obtained by simultaneously solving P2 and P3 and selecting the optimum one. The proposed convex optimization-based algorithm with the hybrid SIC strategy is presented in Algorithm 1.

In simulations, almost all experiments can meet the rank-one constraint even if SDR is used, which drives the exploration of when the optimality of the obtained SDR solution can be established.

Proposition 2. With the randomly generated complex-valued channels, the obtained solution of P2 is optimal when (20b) is strictly satisfied and the obtained $\boldsymbol{W}$ of P3-2 via SDR can satisfy the rank-one constrain when (26c) is strictly satisfied.

Proof. Please refer to Appendix C. □

Remark 1. The solution of P1 is optimal if P2's solution is selected as well as Proposition 2 can be met, otherwise, it is suboptimal. The reason is as follows: the solution of P3 is still suboptimal even if the obtained beamforming

Algorithm 1 Convex optimization with hybird SIC

Calculate: β in Proposition 1.
Solve P2 and select the optimal solution: $\arg\max_{\eta_1^{(1)},\eta_2^{(1)},\boldsymbol{\omega}_0^{(2)}}\{R_1^{(1)}\}$
if $\text{Rank}(\boldsymbol{W}_0^{(1)})=1$ then
  Perform EVD to recover $\boldsymbol{\omega}_0^{(1)}$.
else
  Perform Gaussian Randomization to recover $\boldsymbol{\omega}_0^{(1)}$.
end if
Initialize: $\epsilon=0.0001$, $j=0$, $\boldsymbol{\eta}_0^{(2)}$ and $\boldsymbol{W}_0^{(2)}$
while $\bar{R}_1^{(2)(j+1)}-\bar{R}_1^{(2)(j)}>\epsilon$ do
  Update $\boldsymbol{W}_0^{(2)(j)}$ by solving P3-2.
  Update $\boldsymbol{\eta}_0^{(2)(j)}$ by solving P3-1.
  $j=j+1$
end while
if $\text{Rank}(\boldsymbol{W}_0^{(2)})=1$ then
  Perform EVD to recover $\boldsymbol{\omega}_0^{(2)}$.
else
  Perform Gaussian Randomization to recover $\boldsymbol{\omega}_0^{(2)}$.
end if
Output: $\eta_1^*,\eta_2^*,\boldsymbol{\omega}_0^*=\arg\max\{\max\{R_1^{(1)},R_1^{(2)}\}\}$

vector of P3-2 in each iteration via SDR can meet the rank-one constraint because the non-convex constraint (22b) is approximated by the first-order Taylor expansion. Although the solution of P3 is suboptimal, the probability of choosing the solution of P3 is still not low (see the simulation part). This further demonstrates the superiority of the hybrid SIC strategy.

C. Complexity and Convergence Analysis

The complexity analysis of Algorithm 1 should start with solving P2 and P3 respectively. In fact, solving P2 requires solving three SDPs with the same form as P2-2. The complexity of solving the SDP using the interior point method is given by $\mathcal{O}(\max\{m,n\}^{4.5}\sqrt{n}\log(\frac{1}{\epsilon}))$ [32], where n represents the problem size and m denotes the number of constraints, and hence the total complexity of solving P2 is $\mathcal{O}(3\max\{m,n\}^{4.5}\sqrt{n}\log(\frac{1}{\epsilon}))$. On the other hand, solving P3 requires alternatively solving an LP (i.e., P3-1) and an SDP (i.e. P3-2). The complexity of solving the LP using the interior point method is given by $\mathcal{O}((n+m)^{3.5})$ [33]. Moreover, the complexity of implementing EVD is $\mathcal{O}(n^3)$. Therefore, the total complexity of Algorithm 1 is $\mathcal{O}(3\max\{m,n\}^{4.5}\sqrt{n}\log(\frac{1}{\epsilon})+I_{iter}(\max\{m,n\}^{4.5}\sqrt{n}\log(\frac{1}{\epsilon})+(n+m)^{3.5})+2n^3)$, where I_{iter} is the number of iteration for solving P3.

Denote $Y(\{\boldsymbol{\eta}^l,\boldsymbol{\omega}_0^l\})$ the value of the objective function of P3 in the l-th iteration. For P3-2, with a given beamforming vector, the following equation can be derived

$$\begin{aligned}Y(\{\boldsymbol{\eta}^l,\boldsymbol{\omega}_0^l\}) &= Y_{\boldsymbol{\eta}}(\{\boldsymbol{\eta}^l,\boldsymbol{\omega}_0^l\})\\ &\le Y_{\boldsymbol{\eta}}(\{\boldsymbol{\eta}^{l+1},\boldsymbol{\omega}_0^l\})\\ &\le Y(\{\boldsymbol{\eta}^{l+1},\boldsymbol{\omega}_0^l\}).\end{aligned}\tag{27}$$

where $Y_{\boldsymbol{\eta}}$ represents the value of (25a). The first equality can hold due to the fact that the first-order Taylor expansions are tight at the given local point. The second inequality can hold because P3-1 is optimized. The third inequality can hold since the objective function's value of Problem P3-1 serves as a lower bound on that of P3. Consequently, it can be inferred that the objective function of P3-1 is non-decreasing with each iteration. The similar analysis can be adapted to P3-2

$$\begin{aligned}Y(\{\boldsymbol{\eta}^{l+1},\boldsymbol{\omega}_0^l\}) &= Y_{\boldsymbol{\omega}_0}(\{\boldsymbol{\eta}^{l+1},\boldsymbol{\omega}_0^l\})\\ &\le Y_{\boldsymbol{\omega}_0}(\{\boldsymbol{\eta}^{l+1},\boldsymbol{\omega}_0^{l+1}\})\\ &\le Y(\{\boldsymbol{\eta}^{l+1},\boldsymbol{\omega}_0^{l+1}\}).\end{aligned}\tag{28}$$

According to (27) and (28), the following inequality can be obtained:

$$Y(\{\boldsymbol{\eta}^l,\boldsymbol{\omega}_0^l\})\le Y(\{\boldsymbol{\eta}^l,\boldsymbol{\omega}_0^l\}).\tag{29}$$

Given the fact that the data rate of B_1 must be a finite positive value, the convergence of the proposed algorithm can be guaranteed.

IV. Unsupervised Deep Learning-based Algorithms Design

This section investigates the application of unsupervised deep learning to jointly optimize the reflection coefficients and the beamforming vector. The primary motivations for using learning-based method are as follows: i) as discussed in the last section, convex optimization does not offer a universal algorithm to address both Problems P2 and P3, and there is a significant difference in the complexity of the proposed algorithms; ii) convex optimization only finds as suboptimal solution to P3; iii) learning-based methods are widely used and proven to be effective in dealing with complex, high-dimensional and non-linear optimization problems in the field of wireless communications [34]–[36]. Deep learning-based methods are generally divided into three broad categories: supervised deep learning, unsupervised deep learning and deep reinforcement learning. Supervised learning needs labelled data to train. For optimization problems, labels are optimal solutions and hence supervised deep learning can only be implemented by using the solutions of algorithm 1 to train, which means that the output solutions of supervised learning will not be better than algorithm 1. Therefore, it is difficult to compare the proposed convex optimization-based algorithm with the supervised deep learning-based method. In contrast, unsupervised deep learning with proper deep neural network (DNN) and loss function design can be directly used to solve optimization problems. On the other hand, implementing deep reinforcement learning can be more complex and resource-intensive compared to unsupervised learning, especially in terms of algorithm design, tuning, and computational resources. In summary, for this system, unsupervised deep learning can offer a more straightforward and resource-efficient approach. In this section, both P2 and P3 are addressed using unsupervised deep learning, employing identical DNN architectures, loss function design, and training datasets. Moreover, it is worth highlighting the difference in the dimensions of their output layers. P2 can focus on

the optimization of only the beamforming vector, as its reflection coefficients have been derived in closed-form, whereas P3 necessitates the simultaneous optimization of two reflection coefficients along with the beamforming vector. In the following, this section will delve into three key areas: training/validation data set generation, the construction of deep neural networks, and the design of loss functions.

A. Training/Validation Data Set Generation

Unlike supervised learning that requires feature-label pairs, unsupervised learning uses features to directly train the model. To the specific problem of this paper, the training/validation data sets are constructed by using randomly generated channels. For each time step, the training/validation sample can be expressed as follows:

$$s^{(t)} = \{\boldsymbol{h}_0, \boldsymbol{h}_n, \boldsymbol{h}_B, g_n, g_{n,0}\}, \forall n = 1, 2. \tag{30}$$

To improve the DNN's ability to learn the characteristics of the training set, it is important to include a diverse range of training examples. In a straightforward word, the training set should be large enough and includes as many as possible channel samples. In this paper, the Rician channel model is selected and 50000 channel samples are generated to train the DNN. Note that the input data must be real numbers and hence the real part and imaginary part of the complex channels should be separated and aligned before being stored in the memory.

B. Deep Neural Network Construction

DNN usually includes an input layer, multiple hidden layers and an output layer. The dimension of the input layer depends on the size of the input data while the dimension of the output layer is decided by the output actions. In this paper, the dimension of the input layer for the two SIC decoding orders should be set to $2(3M + 2)$. However, the dimensions of the output layer in the two SIC decoding orders (i.e., P2-1 and P3) are different, because only the beamforming vector is needed to be optimized in the first SIC decoding order case. Therefore, the dimension of the output layer of the DNN corresponding to the SIC decoding order 1 is set to $2M$ while the dimension of the output layer of the DNN corresponding to the SIC decoding order 2 should be set to $2M + 2$. In addition, the Relu function is selected to be the active function and is inserted between every two adjacent layers to introduce non-linear transformation. This can enable the DNN to learn and represent complex patterns, alleviates the gradient vanishing problems, activate the sparsity of the DNN, improve the DNN's robustness and enhance computational efficiency. In deep learning, normalization is a classical skill to improve the training stage as it has the advantages of mitigating internal covariate shift, accelerating training speed, facilitating higher learning rate and improving model generalization. In general, there are two types of normalizations:

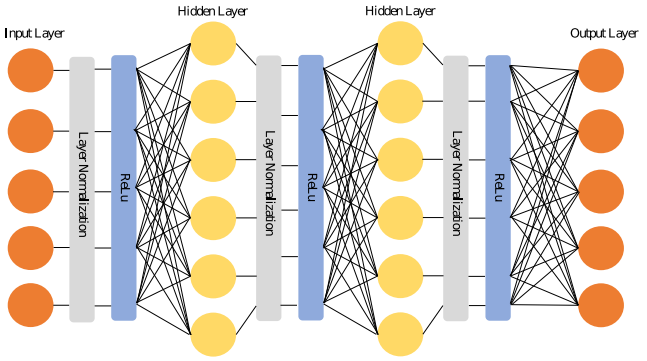


Fig. 1: Structure of the constructed DNN.

- Batch normalization (BN) [37]: This method refers to the normalization of data for each mini-batch during training. Specifically, BN computes the mean and standard deviation of all data within a batch for a particular feature. This method is beneficial in enhancing stability when training large networks, allowing for higher learning rates, and preventing gradient vanishing. However, one substantial limitation of BN is its dependency on operating with mini-batch data. This could pose a problem when dealing with sequence models that require extensive contexts, such as text, time-series data and signal.
- Layer normalization (LN) [38]: LN addresses the issue faced by BN when handling sequence data. LN involves normalizing a single data sample across all features. Consequently, the operations of LN are conducted internally within a specific layer, independent of other samples. This implies that it calculates the mean and standard deviation for each sample individually. This design makes LN particularly advantageous when handling sequence data of variable lengths. Furthermore, LN is not dependent on batch size, and hence the large batch size is allowed to further enhance the training speed by using CUDA.

Considering that the training model is not large and the input data samples (channels) are sequences and are mutually independent, LN is utilized in this paper to improve the training performance. According to [38], LN should be deployed after the normalization but before the non-linearity to give each neuron its own adaptive bias and gain, which helps to stabilize the hidden state dynamics in recurrent networks. The constructed deep neural network is shown in Fig. 1.

C. Loss Function Design for Constrained Problems

For the applications of unsupervised learning to solve constrained optimization problems, it is crucial that design a loss function properly and wisely, because the output of the DNN must satisfy all the constraints. The loss function design of this paper is inspired by the Lagrange dual function in Karush-Kuhn-Tucker (KKT) conditions of an optimization problem. The problem P2 is taken as an

example to describe our loss function design philosophy. The Lagrange dual function of P2 can be expressed as

$$\begin{aligned} L(\eta_1,\eta_2,\boldsymbol{\omega}_0,\boldsymbol{\lambda}) &= -\bar{R}_1^{(1)} - \lambda_1(\bar{R}_2^{(1)} - R_{2,t}) - \lambda_2(R_0 - R_{0,t}) \\ &- \lambda_3(R_{0\to B} - R_{0,t}) - \lambda_4\eta_1 - \lambda_5\eta_2 \\ &+ \lambda_6(\eta_1 - 1) + \lambda_7(\eta_2 - 1) + \lambda_8(|\boldsymbol{\omega}_0|^2 - P_t), \end{aligned} \tag{31}$$

where $\boldsymbol{\lambda}$ is Lagrange multipliers. According to the complementary slackness of KKT conditions, the terms with Lagrange multipliers must be zero if (31) achieves its minimum value. Consequently, the loss function is formulated as an integrated version of the objective functions and constraints by employing a penalty strategy. Specifically, the penalty function is designed by utilizing the degree of deviation of the output from the feasible region, which can be expressed as follows:

$$\mathcal{P}_i(\boldsymbol{x}^{(t)}) = \begin{cases} \mu_i \mathcal{C}_i(\boldsymbol{x}^{(t)})^{\nu_i}, & \text{if } \mathcal{C}_i(\boldsymbol{x}^{(t)}) > 0\ , \\ 0, & \text{otherwise}, \end{cases} \tag{32}$$

where $\boldsymbol{x}^{(t)}$ represent the output of DNN variables in t-th step, $\mu_i > 0$ and $\nu_i > 0$ are called penalty factors and $\mathcal{C}_i(\boldsymbol{x}^{(t)})$ represents the function that corresponding to the i-th constraint. Note that the constraint should be organized into a uniform form (i.e., $\mathcal{C}_i(\boldsymbol{x}^{(t)}) \le 0$) to guarantee the validity of penalty functions. In fact, the total power constraint can be always satisfied by normalizing the output beamforming vector in each time step by using (33), and hence it does not need to be included in the loss function as punishment.

$$\boldsymbol{\omega}_0^{*(t)} = \sqrt{P_t}\frac{\boldsymbol{\omega}_0^{(t)}}{|\boldsymbol{\omega}_0^{(t)}|^2}. \tag{33}$$

By following the above discussion, the loss functions corresponding to the two decoding orders are given by

$$\mathcal{L}_1(\boldsymbol{\omega}_0^{(t)}) = -\mathcal{F}_1(\boldsymbol{\omega}_0^{(t)}) + \sum_i \mathcal{P}_{1,i}(\boldsymbol{\omega}_0^{(t)}) \tag{34}$$

and

$$\begin{aligned} \mathcal{L}_2(\eta_1^{(t)},\eta_2^{(t)},\boldsymbol{\omega}_0^{(t)}) = &-\mathcal{F}_2(\eta_1^{(t)},\eta_2^{(t)},\boldsymbol{\omega}_0^{(t)}) \\ &+ \sum_i \mathcal{P}_{2,i}(\eta_1^{(t)},\eta_2^{(t)},\boldsymbol{\omega}_0^{(t)}), \end{aligned} \tag{35}$$

where

$$\begin{cases} \mathcal{F}_1(\boldsymbol{\omega}_0^{(t)}) = \eta_1^*|g_1|^2|\boldsymbol{h}_1^H\boldsymbol{\omega}_0^{(t)}|^2, \\ \mathcal{F}_2(\eta_1^{(t)},\eta_2^{(t)},\boldsymbol{\omega}_0^{(t)}) = \bar{R}_2^{(2)}(\eta_1^{(t)},\eta_2^{(t)},\boldsymbol{\omega}_0^{(t)}). \end{cases} \tag{36}$$

A mini-batch typically acts as the training dataset for each epoch to improve the training speed and accuracy. Under this mechanism, the network should minimize the mean of the loss functions over the mini-batch. The average loss function is given as follows:

$$\bar{\mathcal{L}} = \frac{1}{\mathcal{B}}\sum_{t=1}^{\mathcal{B}} \mathcal{L}(\boldsymbol{x}^{(t)}), \tag{37}$$

where $\mathcal{B}$ represents the size of a mini-batch. The unsupervised deep learning-based optimization is summarised Algorithm 2.

Algorithm 2 Unsupervised deep learning-based optimization

1: Initialize: Two deep networks according to Fig. 1.
2: Initialize: Generate training set and by (30) with size 50000 and 10000.
3: Initialize: Size if mini-batch $\mathcal{B} = 64$ and Learning rate $\alpha = 0.0001$.
4: Loss Function Design
5: Normalize the output beamforming vector by (33).
6: Map the output $\eta_1^{(2)}$ and $\eta_2^{(2)}$ to $(0, 1]$ by using Sigmoid function.
7: Calculate the loss by (34) and (35).
8: Training Stage
9: for Epoch = 1:10000 do
10: Randomly select training sets with size $\mathcal{B}$.
11: Update the weights and bias of the two DNNs by respectively minimizing (34) and (35) with the form of (37), where Adam optimizer is applied.
12: end for
13: Validation Stage
14: Input: $\{\boldsymbol{h}_0, \boldsymbol{h}_n, \boldsymbol{h}_b, g_n, g_{n,0}\}, \forall n = 1, 2$ in the validation set.
15: if $\mathcal{P}_i^{(1)}(\boldsymbol{x}) \neq 0, \forall i$ then
16: $\eta_1^{(1)*}, \eta_2^{(1)*}, \boldsymbol{\omega}_0^{(1)*} = \arg\max\{R_1^{(1)}\}$
17: else
18: The output for decoding order 1 is infeasible, $R_1^{(1)} = 0$.
19: end if
20: if $\mathcal{P}_i^{(2)}(\boldsymbol{x}) \neq 0, \forall i$ then
21: $\eta_1^{(2)*}, \eta_2^{(2)*}, \boldsymbol{\omega}_0^{(2)*} = \arg\max\{R_1^{(2)}\}$
22: else
23: The output for decoding order 1 is infeasible, $R_1^{(2)} = 0$.
24: end if
25: Output: $\eta_1^*, \eta_2^*, \boldsymbol{\omega}_0^* = \arg\max\{\max\{R_1^{(1)}, R_1^{(2)}\}\}$.

Remark 2. Note that the outputs (i.e., optimization variables) of the two DNNs are different. For SIC decoding order 1, we can apply P2-1 to design its loss function and optimize $\boldsymbol{\omega}_0$ only, as two reflection coefficients can be obtained by their closed-form expressions. Furthermore, the partial closed-form-based loss function design not only avoids the discussion of P2-1 being divided, but also improves the optimality of the output solution of unsupervised deep learning, which will be illustrated in the simulation part. For SIC decoding order 2, we directly use the original form of the problem to design its loss function, as any ensuing analysis for approximation is unlikely to yield practical guidance for the implementation of unsupervised deep learning.

Remark 3. The penalty factors hold significant influence over the pace and precision of the training process, given its impact on the gradient of the loss function. An overly large penalty factor allows the training results to quickly meet the constraint conditions, but it is challenging to provide an optimal solution. Conversely, an excessively

TABLE I: Hyperparameters.

Hyperparameters	Value
Noise power	-94dBm/Hz
Location of BS	(0 m, 0 m)
Locations of B_1, B_2 and U_0	$x_i, y_i \in [3\text{m}, 15\text{m}]$
Path loss coefficient	2.5
Number of neurons of the input layer,	256
Number of neurons of the hidden layer 1	512
Number of neurons of the hidden layer 2	256
Size of training set	50000
Size of validation set	10000
Number of epochs	10000
Size of mini-batch	64
Learning rate	0.0001

small penalty factor slows down the speed at which the output reaches the feasible region. Thus, finding proper penalty factors is very important but challenging.

Remark 4. The feasibility of the output solution from DNN needs to be checked. There are two ways to check the feasibility of the solution. One is to judge whether the output of the DNN can make all penalty (i.e., $\mathcal{P}_i(\boldsymbol{x})$) terms of the loss function zero. The other one is to substitute the output of the DNN into the constraints of the original problem and check whether they are all satisfied. Please note that the former is used in the pseudocode of algorithm 2, which corresponds to our simulations.

V. Simulation Results

In this section, computer simulations are provided to demonstrate the superiority of the proposed BAC-NOMA scheme with hybrid SIC. The time division multiple access (TDMA) scheme is included as a benchmark, where backscatter coefficients and beamforming vectors are also optimized for fair comparison. All hyperparameters related to convex optimization and unsupervised learning-based optimization are listed in table I. Rician fading channel model is utilized and the path loss is also taken into account. For simplicity notation, c_1 denotes the case of SIC decoding order 1, c_2 denotes the case of SIC decoding order 2 and H-SIC denotes the case of hybrid SIC decoding.

Fig. 2 illustrates the trend of B_1's average data rate variation in relation to the BS's transmission power on various SIC decoding schemes, algorithms and MA schemes. In this simulation, the number of antennas is $M = 8$, and the minimum target rate of U_0 and B_2 are 2 bps/Hz and 1 bps/Hz, respectively. As observed, the hybrid SIC decoding strategy consistently delivers superior performance compared to fixed SIC decoding, regardless of the algorithm chosen. In the case of hybrid SIC, the convex optimization-based algorithm slightly outperforms the unsupervised learning-based algorithm, the gap can also be observed in the case of c_1. This improvement originates from the optimal solution with closed-form reflection coefficients presented in c_1. Conversely, the use of unsupervised deep learning for the case of c_2 can provide superior performance compared to the use of convex optimization (i.e., the proposed SCA and SDR-based alternating method), which reveals that the unsupervised deep learning can better fit the optimal solution for this non-concave problem. On the other hand, BAC-NOMA significantly outperforms BAC-TDMA.

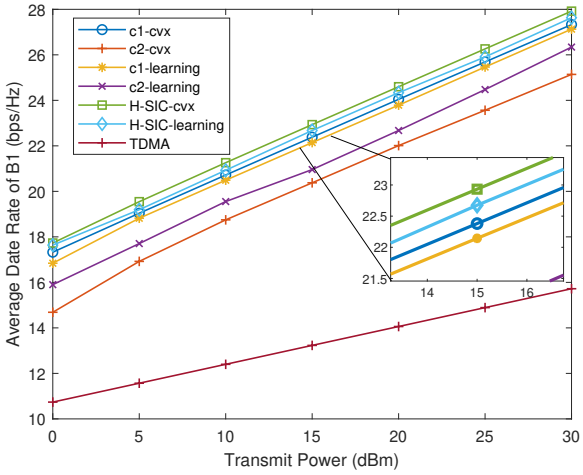


Fig. 2: B_1's average data rate versus the transmit power of the BS.

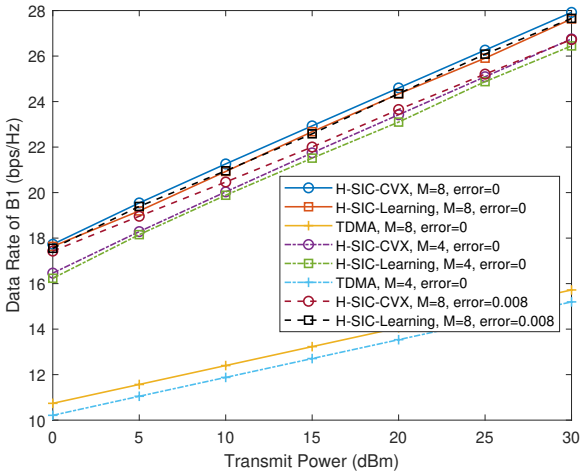
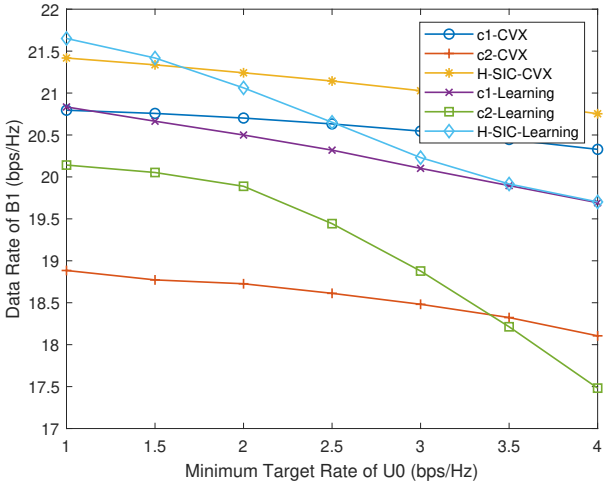
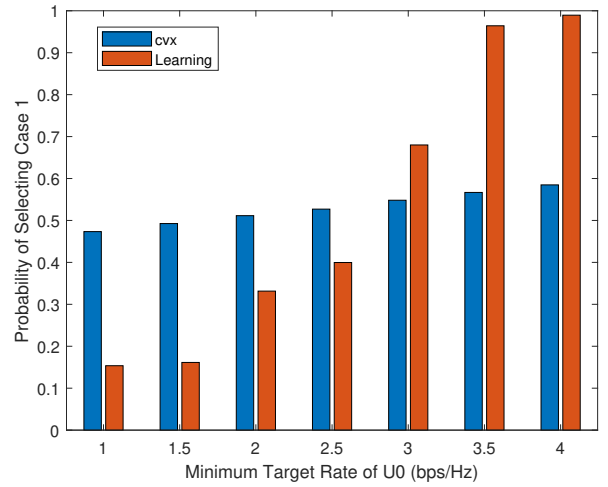


Fig. 3: B_1's average data rate versus the transmit power of the BS with different antennas and channel error.

Fig. 3 shows the impact of the number of antennas and the channel errors on B_1's data rate, where the target rates of B_1 and U_0 are the same as the last simulation. The channel errors indicated that the obtained CSI are outdated. Specifically, the channel employed for optimization and training lags behind the real channel at the time of transmission. As can be seen, the outdated channels degrade B_1's data rate. However, different from that when CSI is perfect, unsupervised deep learning shows superior robustness compared to convex optimization when channels are outdated. Further exploration of the impact of outdated channels will be shown in a later simulation. It can also be observed in this figure that more antennas can provide higher performance gain.

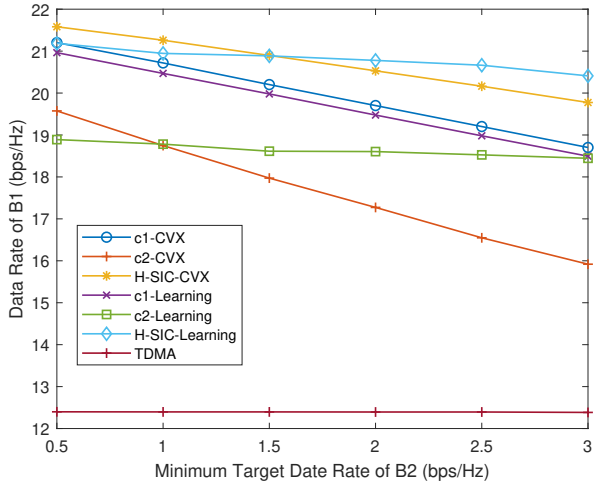


(a) B_1's data rate versus U_0's target rate.

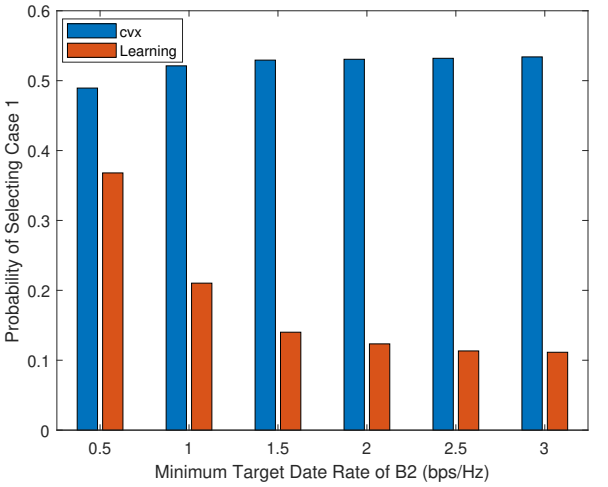


(b) Probability of selecting c_1.

Fig. 4: The impact of the minimum target rate of U_0. M = 8, $P_t = 10$ dBm and $R_{2,t} = 1$ bps/Hz.



(a) B_1's data rate versus B_2's target rate.

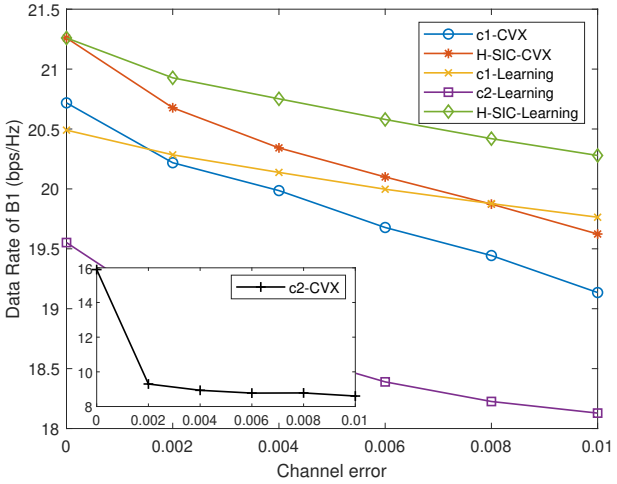


(b) Probability of selecting c_1.

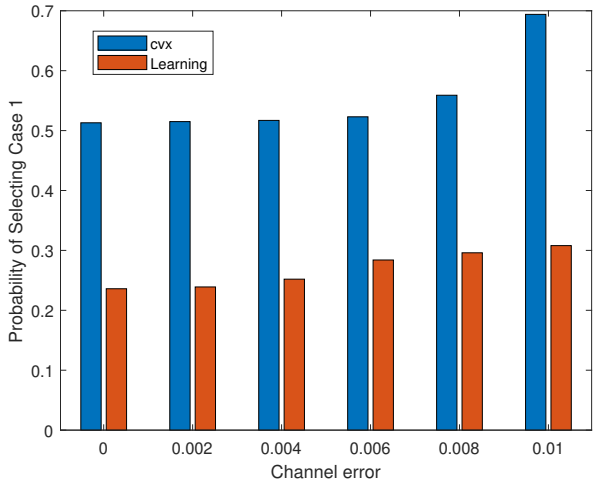
Fig. 5: The impact of the minimum target rate of B_2. M = 8, $P_t = 10$ dBm and $R_{0,t} = 2$ bps/Hz.

The impact of the minimum target rate of U_0 on B_1's data rate performance and SIC decoding order selecting are shown in Fig. 4(a) and Fig. 4(b), where the obtained CSI are perfect. B_1's data rate is decreasing with the increase of U_0's target rate. When the target rate of U_0 is small, hybrid SIC with unsupervised deep learning outperforms that with convex optimization, whereas convex optimization shows better performance when U_0's target rate is large. For the fixed SIC decoding order, in most cases, convex optimization proves to be more suitable for c_1, while unsupervised deep learning is more appropriate for c_2. This trend shown in Fig. 4(a) can correspond to the decoding order selection probabilities shown in Fig. 4(b). As the target rate of U_0 increases, the data rate of B_1 undergoes a sharp decline when unsupervised deep learning is applied, especially for c_2. However, this drop can be mitigated by employing convex optimization. Corresponding to Fig. 4(b), the probability of choosing c_1 rises sharply if unsupervised deep learning is used, while the use of convex optimization only leads to a slight increase in the probability of choosing c_1.

Fig. 5(a) and Fig. 5(b) respectively demonstrate the impact of the minimum target rate of B_2 on B_1's data rate performance and SIC decoding order selecting. Comparing the impact of the minimum target rate of U_0 on B_1's data rate as shown in Fig. 4(a), Fig. 5(a) shows a completely opposite trend. For hybrid SIC, convex optimization outperforms unsupervised deep learning when B_2's minimum target rate is low (i.e., from 0.5 bps/Hz to 1.5 bps/Hz), which is reversed after B2's minimum target rate is greater than 1.5 bps/Hz. Consistent with previous simulations, c1 with convex optimization is always better than that with unsupervised deep learning, which shows again the superiority of the closed-form optimal solution. With the increased minimum target rate of B_2, B_1's data rate degradation of c_1 is more severe than that of c_2 in



(a) B_1's data rate versus channel error.



(b) Probability of selecting c_1.

Fig. 6: The impact of channel errors. M = 8, P_t = 10 dBm, $R_{0,t} = 2$ bps/Hz and $R_{2,t} = 1$ bps/Hz.

the case that unsupervised deep learning is used. This corresponds to the results of Fig. 5(b) that with the increasing minimum target rate of B_2 the probability of choosing c_1 is decreased. However, in the case that applying convex optimation, the speed of B_1's data rate reduction is similar in c1 and c2. Therefore, it is important to choose different algorithms for different $R_{0,t}$ and $R_{2,t}$ setups.

In Fig. 6(a), the scenario that channels are outdated is investigated. Unsupervised deep learning exhibits significant superiority compared to convex optimization when overcoming the adverse effects brought by outdated channels. For c_1, the performance degradation induced by channel errors is tolerable. However, for c_2, such degradation is catastrophic, especially for the use of convex optimization. Moreover, Fig. 6(a) further underscores the superior performance of the hybrid SIC strategy compared to the fixed SIC decoding order. For the decoding order selection probability, owing to c_1's superior robustness compared to c_2, the likelihood of choosing c_1 predictably rises with the increase of channel errors, regardless of the algorithm employed.

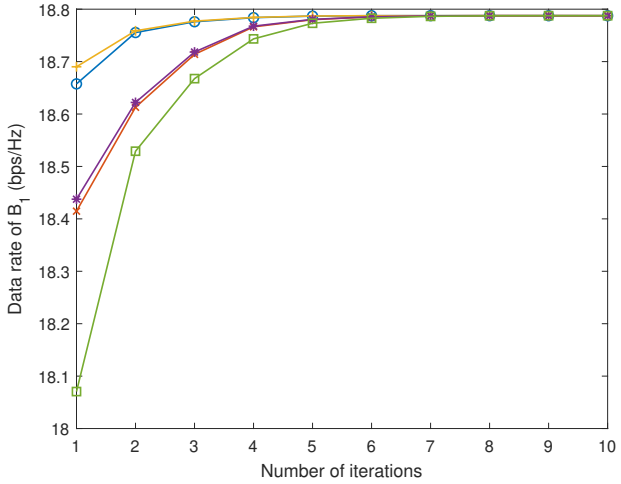
The reasons that unsupervised deep learning outperforms convex optimization under outdated CSI can be analyzed from two perspectives as follows. From the qualitative point of view, in environments with perfect CSI, convex optimization excels due to its ability to utilize accurate and complete information for optimal solutions. However, its performance decreases with outdated CSI, as it relies heavily on data accuracy. Conversely, deep learning models adapt well to uncertainties, maintaining effectiveness even with imperfect CSI. From an optimality point of view, convex optimization provides an optimal solution (the solution of P2) or a near-optimal solution (the solution of P3) based on the current input. When the input CSI is outdated, convex optimization will offer the optimal solution corresponding to the outdated channel, which is impossible to be the optimal solution under the real channel. However, the unsupervised deep learning-based approach usually provides a suboptimal solution. When the input CSI is outdated, the output solution from DNN is still suboptimal to the outdated CSI but more importantly, is possibly optimal to the real CSI. Therefore, the long-term average data rate that the unsupervised learning-based algorithm provided outperforms that provided by the convex optimization.

In all simulations, it can be observed that hybrid SIC always provide the best performance compared with the other two fixed decoding orders (i.e., c_1 and c_1), regardless of which algorithm is used. This benefits from the fact that hybrid SIC always chooses the optimal decoding order.

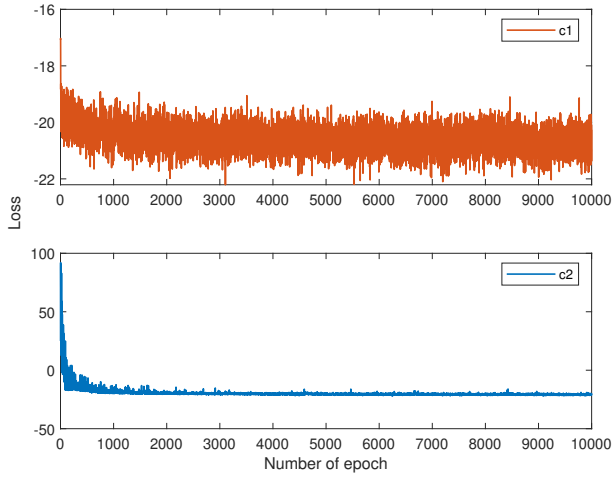
Fig. 7(a) shows the convergence of random 5 experiments under the same hyperparameter setup ($M = 8$, P_t = 0dBm, $R_{0,t} = 2$ bps/Hz and $R_{2,t} = 1$ bps/Hz). As can be observed, the proposed algorithm can converge to a stable state in about 7 iterations. Please note that these 5 experiments converged on different but similar values. Due to scale reasons, they seem to overlap. Fig. 7(b) examines the convergence of the loss function for the two DNNs. As can be observed, c_1 is converged with 1000 epochs while c_2 needs about 2000 epochs. Additionally, the difference between the initial loss value and converged loss values for c_1 is less pronounced than that for c_2. The above two advantages of c1 benefit from the simplification of the problem by introducing the closed-form solution of η_1 and η_2. Therefore, it can be concluded that convex optimization can significantly help the training efficiency of unsupervised deep learning.

VI. Conclusion

In this paper, a novel BAC-NOMA D2D transmission scheme with hybrid SIC has been proposed to enable 6G umMTC networks. The data rate of the high QoS demanding device has been maximized by jointly optimizing their reflection coefficients and designing the



(a) Convergence of the proposed algorithm for solving P3 with random channels.



(b) Convergence of the loss function in the learning-based algorithm.

Fig. 7: Convergence illustration.

beamforming vector. Two subproblems have been formulated because of the use of hybrid SIC. With the convex optimization path, for the first decoding order, the closed-form reflection coefficients have been derived and the beamforming vector is optimized by using SDR, while the second decoding order is optimized alternatively by the algorithm combining SCA and SDR. For the unsupervised deep learning path, the loss function is wisely designed under the inspiration of the Lagrange function, where constraints can be guaranteed by penalty terms. In simulations, curve charts have shown the perspective advantages of convex optimization and unsupervised deep learning-based optimization with different QoS requirements and channel errors, which provides guidance for the selection of optimization methods. Moreover, those bar charts can provide guidance for establishing the optimal decoding order in practical works. The design of user pairing algorithms for multi-user scenarios is also a promising research direction.

Appendix A
Proof of Lemma 1

Proving $f(x)$ is a monotonically decreasing function of x is equivalent to proving $\frac{df(x)}{dx}$ is non-positive when $x \geq 0$. With the knowledge of $\frac{dE_i(-\frac{1}{x})}{dx} = -\frac{e^{-\frac{1}{x}}}{x}$ [16], [31], the first order derivative of $f(x)$ is given by

$$\frac{df(x)}{dx} = \frac{e^{\frac{1}{a+x}}}{(a+x)^2}E_i(-\frac{1}{a+x}) + \frac{1}{a+x} - \left(\frac{e^{\frac{1}{x}}}{x^2}E_i(-\frac{1}{x}) + \frac{1}{x}\right). \tag{38}$$

Defining a new function that $g(x) \triangleq \frac{e^{\frac{1}{x}}}{x^2}E_i(-\frac{1}{x}) + \frac{1}{x}$. Note that (38) can be recast to $\frac{df(x)}{dx} = g(a+x) - g(x), \forall a, x \geq 0$. Hence, to prove (38) is non-positive with non-negative x, it is sufficient to prove $g(x)$ is a monotonically decreasing function of x. Because the monotonically decreasing of $g(x)$ implies that $g(a+x) - g(x) \leq 0$ can always hold, which guarantees the non-positivity of $\frac{df(x)}{dx}$. The first order derivative of $g(x)$ can be derived as follows

$$\frac{dg(x)}{x} = -\frac{1}{x^4}\left(e^{\frac{1}{x}}(1+2x)E_i(-\frac{1}{x}) + x + x^2\right). \tag{39}$$

Similarly, proving $h(x) \triangleq e^{\frac{1}{x}}(1+2x)E_i(-\frac{1}{x}) + x + x^2 \geq 0$ is equivalent to proving (39) is non-positive when $x \geq 0$. To show $h(x)$ is non-negative, the following function of x is defined

$$H(x) = E_i(-\frac{1}{x}) + \frac{x}{1+2x}e^{-\frac{1}{x}} + \frac{x^2}{1+2x}e^{-\frac{1}{x}}. \tag{40}$$

It is noted that proving $H(x) \geq 0$ is equivalent to showing the non-negativity of $h(x)$. The first derivative of $H(x)$ is given by

$$\begin{aligned}\frac{dH(x)}{dx} &= e^{-\frac{1}{x}}\left(-\frac{1}{x} + \frac{2x^2+2x+1}{(1+2x)^2} + \frac{1+x}{(1+2x)x}\right) \\ &= e^{-\frac{1}{x}}\frac{2x^2}{(1+2x)^2} \geq 0.\end{aligned} \tag{41}$$

(41) means $H(x)$ is a monotonically increasing function of x. The lower bound of $H(x)$ can be obtained by following the approximation of $E_i(x) \approx \frac{e^x}{x}$ when $x \to -\infty$ [31]. Hence, $H(x) \geq H(0) = 0$ can be obtained, which implies $\frac{df(x)}{dx}$ is a monotonically decreasing function with respect to x, and the lemma is proved.

Appendix B
Proof of Proposition 1

By defining a constant β such that $-e^{\frac{\sigma^2}{\beta}}E_i(-\frac{\sigma^2}{\beta}) = \ln(2)R_{2,t}$, (18b) can be recast as follows:

$$-e^{\frac{\sigma^2}{\eta_2|g_2|^2|\boldsymbol{h}_2^H\boldsymbol{\omega}_0|^2}}E_i\left(-\frac{\sigma^2}{\eta_2|g_2|^2|\boldsymbol{h}_2^H\boldsymbol{\omega}_0|^2}\right) \geq -e^{\frac{\sigma^2}{\beta}}E_i(-\frac{\sigma^2}{\beta}), \tag{42}$$

where β can be determined through the resolution of the transcendental function, achievable with software utilities like Matlab's vpasolve function. Given that $f(x) = -e^{\frac{1}{x}}E_i(-\frac{1}{x}), x \geq 0$ is a monotonically increasing function

of x as demonstrated in [16], one of the lower bounds of x can be derived as:

$$\frac{\beta}{|g_2|^2|\boldsymbol{h}_2^H\boldsymbol{\omega}_0|^2} \leq \eta_2. \tag{43}$$

According to (14c), two of the upper bounds of x can also be obtained as:

$$\begin{cases} \eta_2 \leq \frac{|\boldsymbol{h}_0^H\boldsymbol{\omega}_0|^2-\gamma_{0,t}\eta_1|g_{1,k}|^2|\boldsymbol{h}_1^H\boldsymbol{\omega}_0|^2}{\gamma_{0,t}|g_{2,k}|^2|\boldsymbol{h}_2^H\boldsymbol{\omega}_0|^2}, \\ \eta_2 \leq \frac{|\boldsymbol{h}_B^H\boldsymbol{\omega}_0|^2-\gamma_{0,t}\eta_1|g_1|^2|\boldsymbol{h}_1^H\boldsymbol{\omega}_0|^2}{\gamma_{0,t}|g_2|^2|\boldsymbol{h}_2^H\boldsymbol{\omega}_0|^2}, \end{cases} \tag{44}$$

where $\gamma_{0,t} = 2^{R_{0,t}} - 1$. Given that (10) is a monotonically decreasing function of η_2 (i.e., Lemma 1), (18a) can achieve its maximum value if η_2 is located at its lower bound. Therefore, the optimal reflection coefficient of B_2 can be given by the following closed-form expression:

$$\eta_2^* = \frac{\beta}{|g_2|^2|\boldsymbol{h}_2^H\boldsymbol{\omega}_0|^2}. \tag{45}$$

For the case that (45) does not satisfy (44) the lower bound of η_2 will be 0 (i.e., B_2 does not transmit signal) and hence the problem becomes infeasible.

By substituting η_2^* into P2, (18a) can be rewritten as follows:

$$\begin{aligned} \bar{R}_1^{(1)} = &-\log_2(e)e^{\frac{\sigma^2}{\eta_1|g_1|^2|\boldsymbol{h}_1^H\boldsymbol{\omega}_0|^2+\beta}} \\ &\times E_i\left(-\frac{\sigma^2}{\eta_1|g_1|^2|\boldsymbol{h}_1^H\boldsymbol{\omega}_0|^2+\beta}\right) - R_{2,t}, \end{aligned} \tag{46}$$

and (14c) can be recast as follows:

$$\eta_1 \leq \frac{|\boldsymbol{h}_0^H\boldsymbol{\omega}_0|^2 - \gamma_{0,t}(\frac{|g_{2,k}|^2}{|g_2|^2}\beta + \sigma^2)}{\gamma_{0,t}|g_{1,k}|^2|\boldsymbol{h}_1^H\boldsymbol{\omega}_0|^2}, \tag{47}$$

$$\eta_1 \leq \frac{|\boldsymbol{h}_B^H\boldsymbol{\omega}_0|^2 - \gamma_{0,t}(\beta + \sigma^2)}{\gamma_{0,t}|g_1|^2|\boldsymbol{h}_1^H\boldsymbol{\omega}_0|^2}. \tag{48}$$

As for η_1, (46) is an increasing function of it, and hence the maximum value of (46) can be reached when η_1 is located at its upper bound. By defining $b_1 = \frac{|\boldsymbol{h}_0^H\boldsymbol{\omega}_0|^2-\gamma_{0,t}(\frac{|g_{2,k}|^2}{|g_2|^2}\beta+\sigma^2)}{\gamma_{0,t}|g_{1,k}|^2|\boldsymbol{h}_1^H\boldsymbol{\omega}_0|^2}$ and $b_2 = \frac{|\boldsymbol{h}_B^H\boldsymbol{\omega}_0|^2-\gamma_{0,t}(\beta+\sigma^2)}{\gamma_{0,t}|g_1|^2|\boldsymbol{h}_1^H\boldsymbol{\omega}_0|^2}$, the closed-form optimal η_1 is given by

$$\eta_1^* = \min\{b_1, b_2, 1\}. \tag{49}$$

The proposition is proved.

C Proof of Proposition 2

Without loss of generality, P2-2 is considered as an example. The proof for other problems can be derived in the similar way. After neglecting the rank-one constraint, the Lagrange function of P2-2 is given by

$$\begin{aligned} &L(\boldsymbol{W}, \lambda_i, \boldsymbol{\lambda}) = -\mathrm{Tr}(\boldsymbol{H}_0\boldsymbol{W}) + \lambda_1(\mathrm{Tr}(\boldsymbol{H}_0\boldsymbol{W})|g_1|^2 \\ &- \mathrm{Tr}(\boldsymbol{H}_B\boldsymbol{W})|g_{1,k}|^2 + a) + \lambda_2(\mathrm{Tr}(\boldsymbol{H}_0\boldsymbol{W}) - \gamma_{0,t}|g_1|^2 \\ &\times \mathrm{Tr}(\boldsymbol{H}_1\boldsymbol{W}) - \gamma_{0,t}(\frac{|g_{2,k}|^2}{|g_2|^2}\beta + \sigma^2)) + \lambda_3(b - \mathrm{Tr}(\boldsymbol{H}_0\boldsymbol{W})) \\ &+ \lambda_4(\beta - |g_2|^2\mathrm{Tr}(\boldsymbol{H}_2\boldsymbol{W})) + \lambda_5(\mathrm{Tr}(\boldsymbol{W}) - P_t) - \mathrm{Tr}(\boldsymbol{\lambda}\boldsymbol{W}), \end{aligned} \tag{50}$$

where λ_i and $\boldsymbol{\lambda}$ are Lagrange multipliers. Without the rank-one constraint, P2-2 is a concave problem and the Karush-Kuhn-Tucker (KKT) can be satisfied. According to the stationarity and complementary slackness, the following two equations are given by

$$\boldsymbol{\lambda} = \lambda_5\boldsymbol{I} + (\lambda_1|g_1|^2 + \lambda_2)\boldsymbol{H}_0 - \Delta, \tag{51}$$

and

$$\boldsymbol{\lambda}\boldsymbol{W} = 0, \tag{52}$$

where $\Delta = \lambda_1\boldsymbol{H}_B|g_{1,k}|^2 + \lambda_2\boldsymbol{H}_1 + (\lambda_3+1)\boldsymbol{H}_0 + \lambda_4\boldsymbol{H}_2$. Given the inherent randomness of channels, the probability that the channel-determined matrix Δ possesses two or more identical eigenvalues approaches zero. Denote the first two maximum eigenvalues of Δ by δ_1 and δ_2, $\delta_1 > \delta_2$. According to slackness complementary, if (20b) is strictly satisfied in case 1), λ_1 and λ_2 are both zero and hence the second term of (51) is 0, which leads to the following discussion to establish the rank-one conclusion.

- If $\lambda_5 = \delta_1$ and $\lambda_5 > \delta_2$, the dimension of the null space of $\boldsymbol{\lambda}$ is one. Further according to (52), $\mathrm{rank}(\boldsymbol{W}) = 1$ can be concluded.
- If $\lambda_5 > \delta_1$, $\boldsymbol{\lambda}$ is a full rank matrix and hence $\boldsymbol{W} = 0$, which is not reasonable in practice.
- If $\lambda_5 \leq \delta_2$, $\boldsymbol{\lambda}$ has at least one negative value and hence $\boldsymbol{\lambda}$ is not a PSD matrix which contradicts with KKT.

The proposition is proved.

References

[1] X. You, C.-X. Wang, J. Huang, X. Gao, Z. Zhang, M. Wang, Y. Huang, C. Zhang, Y. Jiang, J. Wang et al., "Towards 6G wireless communication networks: Vision, enabling technologies, and new paradigm shifts," Science China Information Sciences, vol. 64, no. 1, pp. 1–74, 2021.

[2] S. Nayak and R. Patgiri, "6G communication: A vision on the potential applications," in Edge Analytics: Select Proceedings of 26th International Conference—ADCOM 2020. Springer, 2022, pp. 203–218.

[3] P. K. Padhi and F. Charrua-Santos, "6G enabled industrial internet of everything: Towards a theoretical framework," Applied System Innovation, vol. 4, no. 1, p. 11, 2021.

[4] F. Guo, F. R. Yu, H. Zhang, X. Li, H. Ji, and V. C. Leung, "Enabling massive IoT toward 6G: A comprehensive survey," IEEE Internet of Things Journal, vol. 8, no. 15, pp. 11 891–11 915, 2021.

[5] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, "Non-orthogonal multiple access (NOMA) for cellular future radio access," in 2013 IEEE 77th vehicular technology conference (VTC Spring). IEEE, 2013, pp. 1–5.

[6] J. Choi, "NOMA-based compressive random access using Gaussian spreading," IEEE Transactions on Communications, vol. 67, no. 7, pp. 5167–5177, 2019.

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

[8] Z. Yang, W. Xu, C. Pan, Y. Pan, and M. Chen, "On the optimality of power allocation for noma downlinks with individual qos constraints," IEEE Communications Letters, vol. 21, no. 7, pp. 1649–1652, 2017.

[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 Communications Letters, vol. 24, no. 11, pp. 2373–2377, 2020.

[10] Z. Ding, R. Schober, and H. V. Poor, "No-pain no-gain: DRL assisted optimization in energy-constrained CR-NOMA networks," IEEE Transactions on Communications, vol. 69, no. 9, pp. 5917–5932, 2021.

[11] Z. Ding, "Harvesting devices' heterogeneous energy profiles and QoS requirements in IoT: WPT-NOMA vs BAC-NOMA," IEEE Transactions on Communications, vol. 69, no. 5, pp. 2837–2850, 2021.

[12] Z. Yang, W. Xu, Y. Pan, C. Pan, and M. Chen, "Energy efficient resource allocation in machine-to-machine communications with multiple access and energy harvesting for iot," IEEE Internet of Things Journal, vol. 5, no. 1, pp. 229–245, 2017.

[13] A. R. Koelle, S. W. Depp, and R. W. Freyman, "Short-range radio-telemetry for electronic identification, using modulated RF backscatter," Proceedings of the IEEE, vol. 63, no. 8, pp. 1260–1261, 1975.

[14] S. J. Nawaz, S. K. Sharma, B. Mansoor, M. N. Patwary, and N. M. Khan, "Non-coherent and backscatter communications: Enabling ultra-massive connectivity in 6G wireless networks," IEEE Access, vol. 9, pp. 38 144–38 186, 2021.

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

[16] Z. Ding and H. V. Poor, "On the application of BAC-NOMA to 6G umMTC," IEEE Communications Letters, vol. 25, no. 8, pp. 2678–2682, 2021.

[17] 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 of Things Journal, vol. 7, no. 2, pp. 1350–1363, 2019.

[18] J.-P. Niu and G. Y. Li, "An overview on backscatter communications," Journal of Communications and Information Networks, vol. 4, no. 2, pp. 1–14, 2019.

[19] N. Van Huynh, D. T. Hoang, X. Lu, D. Niyato, P. Wang, and D. I. Kim, "Ambient backscatter communications: A contemporary survey," IEEE Communications surveys & tutorials, vol. 20, no. 4, pp. 2889–2922, 2018.

[20] Y.-C. Liang, Q. Zhang, E. G. Larsson, and G. Y. Li, "Symbiotic radio: Cognitive backscattering communications for future wireless networks," IEEE Transactions on Cognitive Communications and Networking, vol. 6, no. 4, pp. 1242–1255, 2020.

[21] G. Yang, X. Xu, and Y.-C. Liang, "Resource allocation in NOMA-enhanced backscatter communication networks for wireless powered IoT," IEEE Wireless Communications Letters, vol. 9, no. 1, pp. 117–120, 2019.

[22] Y. Zhuang, X. Li, H. Ji, and H. Zhang, "Exploiting intelligent reflecting surface for energy efficiency in ambient backscatter communication-enabled NOMA networks," IEEE Transactions on Green Communications and Networking, vol. 6, no. 1, pp. 163–174, 2022.

[23] Y. Xu, Z. Qin, G. Gui, H. Gacanin, H. Sari, and F. Adachi, "Energy efficiency maximization in NOMA enabled backscatter communications with QoS guarantee," IEEE Wireless Communications Letters, vol. 10, no. 2, pp. 353–357, 2020.

[24] M. Elsayed, A. Samir, A. A. A. El-Banna, X. Li, and B. M. ElHalawany, "When NOMA multiplexing meets symbiotic ambient backscatter communication: Outage analysis," IEEE Transactions on Vehicular Technology, vol. 71, no. 1, pp. 1026–1031, 2021.

[25] Z. Ding and H. V. Poor, "Advantages of NOMA for Multi-User BackCom Networks," IEEE Communications Letters, vol. 25, no. 10, pp. 3408–3412, 2021.

[26] S. Louvros, M. Paraskevas, and T. Chrysikos, "Qos-aware resource management in 5g and 6g cloud-based architectures with priorities," Information, vol. 14, no. 3, p. 175, 2023.

[27] S. Wan, J. Hu, C. Chen, A. Jolfaei, S. Mumtaz, and Q. Pei, "Fair-hierarchical scheduling for diversified services in space, air and ground for 6g-dense internet of things," IEEE Transactions on Network Science and Engineering, vol. 8, no. 4, pp. 2837–2848, 2020.

[28] O. A. Nwogu, G. Diaz, and M. Abdennebi, "Differential traffic qos scheduling for 5g/6g fronthaul networks," in 2021 31st International Telecommunication Networks and Applications Conference (ITNAC). IEEE, 2021, pp. 113–120.

[29] 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. 20 000–20 013, 2019.

[30] 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 of Things Journal, vol. 7, no. 2, pp. 1350–1363, 2020.

[31] I. S. Gradshteyn and I. M. Ryzhik, Table of integrals, series, and products. Academic press, 2014.

[32] Z.-Q. Luo, W.-K. Ma, A. M.-C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," IEEE Signal Processing Magazine, vol. 27, no. 3, pp. 20–34, 2010.

[33] S. P. Boyd and L. Vandenberghe, Convex optimization. Cambridge university press, 2004.

[34] V. P. Rekkas, S. Sotiroudis, P. Sarigiannidis, G. K. Karagiannidis, and S. K. Goudos, "Unsupervised machine learning in 6g networks-state-of-the-art and future trends," in 2021 10th international conference on modern circuits and systems technologies (MOCAST). IEEE, 2021, pp. 1–4.

[35] J. Gao, C. Zhong, X. Chen, H. Lin, and Z. Zhang, "Unsupervised learning for passive beamforming," ieee communications letters, vol. 24, no. 5, pp. 1052–1056, 2020.

[36] H. Hojatian, J. Nadal, J.-F. Frigon, and F. Leduc-Primeau, "Unsupervised deep learning for massive mimo hybrid beamforming," IEEE Transactions on Wireless Communications, vol. 20, no. 11, pp. 7086–7099, 2021.

[37] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in International conference on machine learning. pmlr, 2015, pp. 448–456.

[38] J. L. Ba, J. R. Kiros, and G. E. Hinton, "Layer normalization," arXiv preprint arXiv:1607.06450, 2016.