electronics

Article

Beamforming Optimization with the Assistance of Deep Learning in a Rate-Splitting Multiple-Access Simultaneous Wireless Information and Power Transfer System with a Power Beacon

Mario R. Camana [1,2], Carla E. Garcia [2] and Insoo Koo [1,*]

[1] Department of Electrical, Electronic and Computer Engineering, University of Ulsan, Ulsan 680-749, Republic of Korea; mario_camana@hotmail.com
[2] Interdisciplinary Centre for Security, Reliability and Trust, University of Luxembourg, 4365 Luxembourg City, Luxembourg; carli.garcia27@hotmail.com
* Correspondence: iskoo@ulsan.ac.kr

Abstract: This study examined the implementation of rate-splitting multiple access (RSMA) in a multiple-input single-output system using simultaneous wireless information and power transfer (SWIPT) technology. The coexistence of a base station and a power beacon was considered, aiming to transmit information and energy to two sets of users. One set comprises users who solely harvest energy, whereas the other can decode information and energy using a power-splitting (PS) structure. The main objective of this optimization was to minimize the total transmit power of the system while satisfying the rate requirements for PS users and ensuring minimum energy harvesting (EH) for both PS and EH users. The non-convex problem was addressed by dividing it into two subproblems. The first subproblem was solved using a deep learning-based scheme, combining principal component analysis and a deep neural network. The semidefinite relaxation method was used to solve the second subproblem. The proposed method offers lower computational complexity compared to traditional iterative-based approaches. The simulation results demonstrate the superior performance of the proposed scheme compared to traditional methods such as non-orthogonal multiple access and space-division multiple access. Furthermore, the ability of the proposed method to generalize was validated by assessing its effectiveness across several challenging scenarios.

Keywords: simultaneous wireless information and power transfer; SWIPT; power beacon; rate-splitting multiple access; RSMA; deep neural network; DNN; semidefinite relaxation; SDR

Citation: Camana, M.R.; Garcia, C.E.; Koo, I. Beamforming Optimization with the Assistance of Deep Learning in a Rate-Splitting Multiple-Access Simultaneous Wireless Information and Power Transfer System with a Power Beacon. *Electronics* **2024**, *13*, 872. https://doi.org/10.3390/electronics13050872

Academic Editors: Juan-Carlos Cano, Yichuang Sun, Haeyoung Lee and Oluyomi Simpson

Received: 18 December 2023
Revised: 10 February 2024
Accepted: 21 February 2024
Published: 23 February 2024

1. Introduction

Rate-splitting multiple access (RSMA) has become a promising multiple-access framework that is being considered for implementation in future 6G networks and beyond [1]. Recent studies have shown that RSMA outperforms traditional multiple-access techniques, such as non-orthogonal multiple access (NOMA) and space-division multiple access (SDMA), in terms of energy and spectral efficiency for multi-antenna systems [1–4]. The concept behind RSMA involves dividing user messages into multiple parts, which are transmitted using superposition coding at the transmitter and decoded using successive interference cancellation (SIC) at the receivers. The most common approach is based on single-layer rate splitting (RS), involving two parts: a common part and a private part. The common parts from all users are transmitted simultaneously using a shared codebook and must be decoded by all users, whereas the private parts are transmitted using private codebooks. This approach allows RSMA to transmit all the messages simultaneously and in the same frequency band using the power and spatial domains. Consequently, RSMA offers flexibility, where interference can be treated as noise or fully decoded, making it a versatile framework encompassing SDMA and power-division NOMA as special cases [3,5].

Modern wireless systems consider communication trends, such as the Internet of Things (IoT) and machine-type communications (MTC). The key priorities in wireless communications involve advancing green technology and reducing device power consumption [6]. Various studies have been undertaken to establish self-sustainable communication systems using energy-harvesting (EH) techniques. One of the EH technologies contributing to these objectives is simultaneous wireless information and power transfer (SWIPT) [7]. SWIPT is an effective technology that enables the base station (BS) to transmit energy and information simultaneously to wireless users. Within SWIPT, the most well-known architectures are the power-splitting (PS) architecture and time-switching (TS) architecture, where both schemes consider an EH module and an information decoding (ID) module at the receiver. In the TS architecture, the receiver periodically alternates between the ID and EH modules based on a TS sequence. In the PS architecture, the incoming radio frequency (RF) signal is divided into two streams based on a PS ratio, which are sent to the ID and EH modules. The authors of [7] focused on the PS architecture because it has been established in the literature as achieving the best balance between energy harvesting and information decoding.

A cost-effective solution was proposed in [8] using low-cost stations called power beacons (PBs) for wirelessly recharging devices using RF energy. PBs primarily serve as a source of wireless energy, extending the operational lifespan of battery-powered devices through wireless recharging. PBs do not necessitate complex computations and have low backhaul link requirements, enabling cost-effective and adaptable placement, making them a valuable addition to the system.

SWIPT has been widely studied, ranging from single-antenna systems [9] to multiple-antenna systems [10,11]. Shi et al. [10] optimized the precoding vectors and PS ratios in a multiuser multiple-input single-output (MU MISO) system to minimize the transmission power while considering both the minimum signal-to-interference-plus-noise ratio (SINR) and minimum EH at the user side. An extension to a multiuser multi-input multi-output (MIMO) system was explored in [11], considering the EH constraints and a maximum tolerable mean square error (MSE) for received information. SWIPT has also been implemented along with the NOMA method in multi-antenna systems to maximize data rates [12], minimize transmit power [13], and enhance energy efficiency [14]. On the other hand, there is limited research on the performance of networks implementing SWIPT in conjunction with the RSMA method.

In the context of RSMA, a pioneering study introduced the concept of RS [2]. The authors elucidated the principal limitations of conventional methods while emphasizing the potential advantages of RS in terms of spectral and energy efficiencies compared to traditional techniques. The initial investigations of RS in a multiuser MISO system [15,16] aimed to maximize the minimum rate and maximize the sum rate. The solutions to these optimization problems were based on the weighted minimum mean square error (WMMSE) method and the alternate optimization (AO) algorithm for optimizing the precoding vectors and common rate variables. These studies reported that the RS framework consistently outperformed conventional methods, as evidenced by a comprehensive analysis of data rates and complexity in scenarios involving imperfect channel state information at the transmitter (CSIT). In [17], an extension analysis was conducted on massive MIMO systems to maximize the minimum achievable rate of the common message. A hierarchical rate-splitting approach was proposed to address the challenges posed by the extensive array of antennas at the transmitter. The simulation results revealed the superior performance of RS over conventional broadcasting methods, considering perfect and imperfect CSIT. The RSMA method was initially introduced for downlink multiuser MISO systems [3]. The authors highlighted the generality of RSMA compared to SDMA and NOMA. They also addressed the maximization of the weighted sum rate while considering the minimum rate constraints and power limitations. The validity of the authors' assertions was confirmed by the simulation results, emphasizing the superior performance of RSMA across various scenarios with different network loads and numbers of users. An exhaustive analytical

analysis was conducted for a two-user case in [5], demonstrating the generality of RSMA by showcasing how it encompasses SDMA, NOMA, and orthogonal multiple access (OMA) as particular cases. Furthermore, the energy efficiency maximization and sum-rate maximization in multiuser MISO systems examined using RS emphasized the outstanding performance of RSMA, demonstrating its superior spectral and energy efficiency compared to SDMA and NOMA [18].

An initial study of RSMA and SWIPT was conducted to maximize the sum rate of users [19], considering two types of users: one for decoding information only and the other for harvesting energy only. A previous study [20] investigated the performance of RSMA with SWIPT in multiuser MISO systems, considering that users can decode information and harvest energy based on a PS factor. Furthermore, the scenario of RSMA with SWIPT when an IRS is deployed in the system was investigated in [21]. These studies reported significant improvements in performance provided by RSMA compared to traditional SDMA and NOMA methods. On the other hand, the aforementioned studies primarily focused on users close to the BS, driven by the EH requirements imposed by the optimization problem. This limitation becomes evident in real-world deployments, where users may be located far from the BS, posing challenges in meeting the EH requirements. Therefore, this paper investigated the deployment of a PB with RSMA as an efficient approach to address the previously mentioned challenge.

Huang et al. [22] introduced a system involving PBs for wireless power transfer (WPT). These PBs can wirelessly charge receivers and be strategically placed alongside femtocell base stations (BSs) to provide short-range SWIPT to wireless devices. A PB-assisted wireless-powered communication network (WPCN) was proposed in [23], consisting of a single-antenna PB and several single-antenna access points (APs). The PB provides RF energy to the APs, which transmit their information using the energy harvested from the PB. An extension of the PB-assisted WPCN, considering multi-antenna PBs, was reported in [24] to optimize the energy beamformer vector and maximize the spectrum efficiency. A previous study [25] considered the coexistence of a multi-antenna PB with a multi-antenna BS, assuming one ID user and several EH users, where the authors addressed the maximization of the total harvested energy. On the other hand, the aforementioned works did not consider SWIPT users equipped with a PS architecture and did not incorporate the RSMA framework.

Regarding state-of-the-art SWIPT systems assisted by a PB, the minimization of transmit power in a single-antenna system was investigated in [10]. Extensions to a multiuser MISO system incorporating SDMA and NOMA were introduced in [26,27], respectively. Vu et al. [28] considered a scenario in which a multi-antenna PB transmits RF energy to a single-antenna transmitter, serving two users with several relays. The transmitter employed the NOMA method and harvested energy from the PB to transmit the two messages. The relay applied the PS architecture to harvest energy and decode the message for the distant user, which was then forwarded. A part of this investigation was presented in a conference article [29], where the SWIPT system with RSMA was introduced with the aid of a PB, where the optimal scenario of decoding the whole interference from the PB at the user side was assumed. On the other hand, this paper considered the general case of treating the interference from the PB as noise. In addition, a low-complexity and efficient scheme was proposed based on deep learning (DL).

DL is gaining popularity as a technique for resource optimization in wireless networks because of its ability to significantly reduce computational complexity compared to traditional optimization methods. A review of the most common schemes based on machine learning and reinforcement learning methods for network optimization problems was published in [30,31]. A deep neural network (DNN) was introduced as an approximation method for solving the sum-rate maximization problem using the WMMSE algorithm in a single-antenna system [32]. The results of the simulations demonstrated the effectiveness of the DNN, which closely approximated the solution of the WMMSE algorithm while reducing the computational time. Xia et al. [33] addressed three popular optimization prob-

lems in a multiuser MISO system with SDMA. Their proposed solution leveraged a neural network module to predict key features based on the channel vectors and a beamforming module to construct the beamforming vector from the predicted vital features. Building upon the previous system model, the sum-rate maximization problem was investigated in [34], and two types of schemes based on a DNN were proposed: one where the beamformers are directly generated by the DNN output and the other utilizing a beamformer recovery module after the DNN module. The simulation results demonstrated performance close to traditional optimization methods, with the model incorporating the beamformer recovery module yielding the best results. On the other hand, the aforementioned work did not consider SWIPT technology and the RSMA method. The solutions are problem-dependent and rely on expert knowledge. Furthermore, state-of-the-art methods consider the channel vectors as direct inputs to the deep learning module. This approach involves using hundreds of features as the number of antennas and users increases, increasing the complexity of the deep learning model. A previous work [35] considered a MISO SWIPT system with RSMA, where a deep learning-based solution was proposed. The scheme comprised three modules: an autoencoder for dimension reduction, a DNN to predict a set of target variables, and a precoding module to obtain the precoding vectors. On the other hand, the authors did not account for EH users or the deployment of a PB, which significantly increased the number of input features. Furthermore, the autoencoder module required problem-dependent adjustments of its hyperparameters, such as the number of hidden layers, hidden nodes, activation functions, and learning rate. These factors made it difficult to generalize the scheme to other scenarios.

The RSMA framework has showcased significant enhancements in spectral and energy efficiencies compared to conventional SDMA and NOMA techniques. This establishes RSMA as a promising candidate for future 6G networks, prompting a thorough examination of its performance across diverse multi-antenna systems. Moreover, the deployment of PBs holds vital significance in extending the lifespan of wireless devices, especially in IoT scenarios where battery replacement poses challenges in hard-to-access areas. However, the interference generated by PBs during the information decoding process at the user side necessitates a comprehensive investigation into an efficient beamforming design at both the BS and PB. Consequently, this study aims to develop a high-performance, low-complexity solution to jointly optimize the beamforming vectors and PS ratios in a multi-antenna system. The proposed system model integrates a multi-antenna transmitter implementing RSMA with SWIPT, along with a multi-antenna PB. The primary objective of this study is to minimize the total transmission power while fulfilling data rates and EH requirements for EH and PS users. The main contributions of this paper are as follows:

- In the considered system model, the objective is to minimize the total transmission power of the BS and PB while meeting the minimum EH requirements for EH and PS users and ensuring a minimum data rate for PS users.
- A two-step approach is adopted to address the non-convex problem presented in this study. In the first step, we optimize the common rate variables through a DL-based scheme that combines the principal component analysis (PCA) technique for dimensionality reduction with a DNN. The second step focuses on optimizing the beamforming vectors of the BS and PB, along with the PS factors. The SDR technique is used to accomplish this.
- As a comparative scheme, the proposed minimization problem is addressed using a PSO-SDR approach, which is an iterative-based method that provides near-optimal solutions to the proposed problem. This scheme serves as a reference to analyze the performance of the proposed DL-based solution.
- Simulation results show that the proposed DL-based method can perform similarly to traditional iterative-based schemes while significantly reducing computational complexity. Furthermore, the proposed RSMA solution is compared with NOMA and SDMA, showing that the RSMA approach achieves the lowest transmit power. Moreover, the generalization performance of the proposed DL-based method is validated

by testing its performance across several challenging scenarios not included in the DL model training.

The remainder of this paper is structured as follows. Section 2 introduces the system model and formulates the problem. Section 3 outlines the proposed solution, and presents the comparative schemes. Section 4 provides the simulation results, and Section 5 reports the conclusions.

2. System Model and Problem Formulation

A MU MISO RSMA system assisted by a PB with SWIPT is considered, as shown in Figure 1. The number of antennas at the BS is $M \geq 2$, the number of antennas at the PB is $N \geq 2$, and there are K single-antenna PS users and G single-antenna EH users. EH users are exclusively dedicated to harvesting RF energy, whereas PS users are equipped with a PS mechanism, allowing them to partition the incoming RF signal into two components, serving both information decoding and EH based on a designated PS ratio, θ_k. The BS transmits information signals to both harvest energy at the EH module and decode information at the ID module for PS users. The PB, on the other hand, transmits energy-carrying signals used for energy harvesting by EH users and the EH module of PS users. However, these energy signals are considered interference at the ID module of PS users.

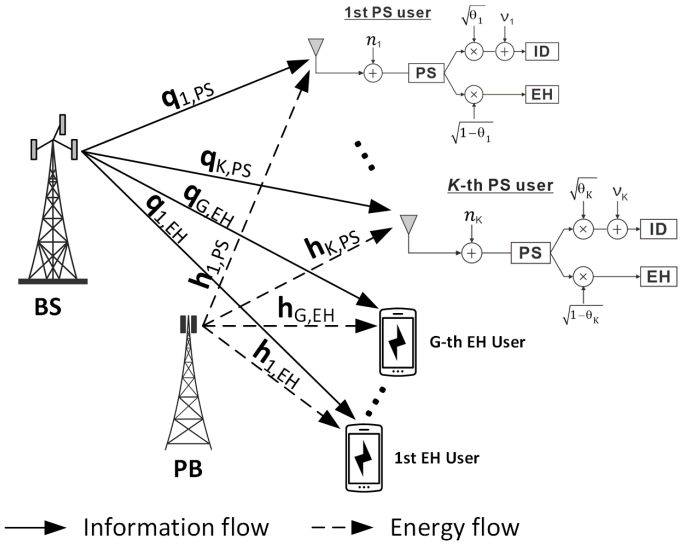

Figure 1. MISO SWIPT system aided by a PB.

In one-layer RSMA, the original message intended for the kth PS user, W_k, is separated into a common message, $W_{k,c}$, and a private message, $W_{k,p}$. A super-common message is produced by combining all the K common messages and encoding them into the common stream z_c^{PS}, which needs to be decoded by all PS users. The K private messages are independently encoded into K private streams $\{z_k^{PS}\}$ to be decoded by their respective user. At the PB, the energy-carrying signals are represented by $\left\{z_g^{EH}\right\}$, $g = 1, \ldots, G$. Therefore, the signals transmitted at the BS and PB after precoding are $\mathbf{x}_{BS} = \mathbf{p}_0 z_c^{PS} + \sum_{k=1}^{K} \mathbf{p}_k z_k^{PS}$ and $\mathbf{x}_{PB} = \sum_{g=1}^{G} \mathbf{e}_g z_g^{EH}$, respectively, where $\mathbf{p}_0 \in \mathbb{C}^{M\times 1}$ and $\mathbf{p}_k \in \mathbb{C}^{M\times 1}$ represent the information beamforming vectors, and $\mathbf{e}_g \in \mathbb{C}^{N\times 1}$ corresponds to the energy beamforming vector for z_c^{PS}, z_k^{PS}, and z_g^{EH}. Figure 2 shows the RSMA transmission scheme in the considered system model.

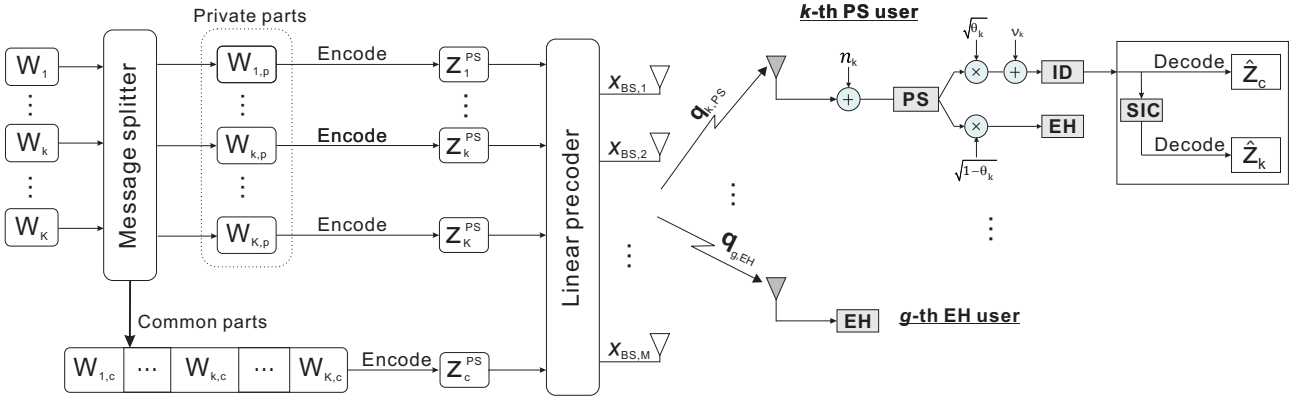

Figure 2. Proposed RSMA transmission scheme.

At the user side, the kth PS user first decodes the common stream z_c^{PS} with an achievable rate given by

$$R_{k,0} = \log_2\left(1 + \frac{\left|\mathbf{q}_{k,PS}^H \mathbf{p}_0\right|^2}{\sum_{i=1}^{K}\left|\mathbf{q}_{k,PS}^H \mathbf{p}_i\right|^2 + \sum_{g=1}^{G}\left|\mathbf{h}_{k,PS}^H \mathbf{e}_g\right|^2 + \sigma_k^2 + \frac{\gamma_k^2}{\theta_k}}\right), \forall k, \tag{1}$$

where $\mathbf{q}_{k,PS} \in \mathbb{C}^{M\times 1}$ represents the channel vector from the BS to the kth PS user; $\mathbf{h}_{k,PS} \in \mathbb{C}^{N\times 1}$ is the channel vector from the PB to the kth PS user; $\theta_k \in (0,1)$ is the PS factor; $n_k \sim \mathcal{CN}(0,\sigma_k^2)$ represents the Gaussian noise at the antenna of the kth PS user; and $v_k \sim \mathcal{CN}(0,\delta_k^2)$ is the data processing noise at the ID module of the kth PS user.

After decoding z_c^{PS}, RSMA employs the SIC procedure to eliminate interference due to the common stream z_c^{PS}. The kth PS can decode its private stream z_k^{PS} with an achievable rate given by

$$\mathrm{R}_k = \log_2\left(1 + \frac{\left|\mathbf{q}_{k,PS}^H \mathbf{p}_k\right|^2}{\sum_{i=1,i\neq k}^{K}\left|\mathbf{q}_{k,PS}^H \mathbf{p}_i\right|^2 + \sum_{g=1}^{G}\left|\mathbf{h}_{k,PS}^H \mathbf{e}_g\right|^2 + \sigma_k^2 + \frac{\gamma_k^2}{\theta_k}}\right), \forall k. \tag{2}$$

The rate at which z_c^{PS} is transmitted, expressed as R_0, must satisfy $R_0 \leq \min\{R_{1,0},\ldots,R_{K,0}\}$ because the common stream z_c^{PS} needs to be decoded by all users. Moreover, the rate R_0 is composed of the rates to transmit each $W_{k,c}$ and can be expressed as $R_0 = \sum_{k=1}^{K} \alpha_k$, where α_k is the rate to transmit the common part of the kth message, $W_{k,c}$.

The energy harvested at the EH module of the kth PS user can be expressed as follows:

$$\mathrm{EH}_k^{PS} = \varsigma_k^{PS}(1-\theta_k)\left(\sum_{i=0}^{K}\left|\mathbf{q}_{k,PS}^H \mathbf{p}_i\right|^2 + \sum_{g=1}^{G}\left|\mathbf{h}_{k,PS}^H \mathbf{e}_g\right|^2 + \sigma_k^2\right), \forall k, \tag{3}$$

where ς_k^{PS} is the EH efficiency at the kth PS user. Moreover, at the gth EH user, the harvested energy can be expressed as

$$\mathrm{EH}_g^{EH} = \varsigma_g^{EH}\left(\sum_{i=0}^{K}\left|\mathbf{q}_{g,EH}^H \mathbf{p}_i\right|^2 + \sum_{j=1}^{G}\left|\mathbf{h}_{g,EH}^H \mathbf{e}_j\right|^2\right), \forall g, \tag{4}$$

where ς_g^{EH} represents the EH efficiency at the gth EH user, and $\mathbf{h}_{g,EH} \in \mathbb{C}^{N\times 1}$ and $\mathbf{q}_{g,EH} \in \mathbb{C}^{M\times 1}$ are the channel vectors from the PB and BS to the gth EH user, respectively.

We seek to optimize the beamforming vectors, $\mathbf{p}_0, \{\mathbf{p}_k, \mathbf{e}_g\}$; the common rate variables, α_k; and the PS ratios, θ_k, which together achieve a minimum rate requirement, and at the same time, they can harvest a required minimum EH for future use. The minimization of the sum transmission power of the PB and BS can be formulated as follows:

$$\min_{\mathbf{p}_0, \{\mathbf{p}_k, \alpha_k, \theta_k, \mathbf{e}_g\}} \quad \sum_{i=0}^{K} \|\mathbf{p}_i\|^2 + \sum_{g=1}^{G} \|\mathbf{e}_g\|^2 \tag{5a}$$

$$\text{s.t.} \quad \alpha_k + R_k \geq \chi_k, \forall k \tag{5b}$$

$$\sum_{i=1}^{K} \alpha_i \leq R_{k,0}, \forall k \tag{5c}$$

$$\mathrm{EH}_k^{PS} \geq \varepsilon_k^{PS}, \forall k \tag{5d}$$

$$\mathrm{EH}_g^{EH} \geq \varepsilon_g^{EH}, \forall g \tag{5e}$$

$$0 < \theta_k < 1, \forall k \tag{5f}$$

$$\alpha_k \geq 0, \forall k, \tag{5g}$$

Constraint (5b) ensures that the kth PS user achieves a minimum rate requirement, denoted as χ_k. Constraint (5c) is set to guarantee that the common stream can be decoded by all PS users. Constraints (5d) and (5e) ensure that each kth PS user and gth EH user can harvest a minimum EH requirement, denoted as ε_k^{PS} and ε_g^{EH}, respectively. Solving the power minimization problem (5) is challenging because of its non-convex constraints (5b)–(5e). In the following, a near-optimal solution is proposed based on DL techniques and the SDR method.

3. Proposed Approach for Addressing Problem (5)

A DL-based scheme and SDR method are developed to address the non-convex problem (5). First, problem (5) is reformulated into two subproblems, denoted as follows:

$$\min_{\{\alpha_k\}} (\mathrm{Y}(\alpha_k)) \tag{6a}$$

$$\mathrm{Y}(\alpha_k) = \min_{\mathbf{p}_0, \{\mathbf{p}_k, \theta_k, \mathbf{e}_g\}} \quad \sum_{i=0}^{K} \|\mathbf{p}_i\|^2 + \sum_{g=1}^{G} \|\mathbf{e}_g\|^2$$

$$\text{s.t. } (5\text{b})\text{–}(5\text{f}), \tag{6b}$$

where the first subproblem, represented in (6a), optimizes the common rate variables, α_k, for the given beamforming vectors, $\mathbf{p}_0, \{\mathbf{p}_k, \mathbf{e}_g\}$, and PS ratios, θ_k. Section 3.1 presents a DNN-based scheme with PCA to solve this first subproblem. The second subproblem, represented by $\mathrm{Y}(\alpha_k)$ in (6b), optimizes the beamforming vectors, $\mathbf{p}_0, \{\mathbf{p}_k, \mathbf{e}_g\}$, and the PS ratios, θ_k, for the given common rate variables, α_k. The proposed solution for the second subproblem is based on the SDR technique with the penalty function method and is described in Section 3.2.

Figure 3 illustrates the overall procedure of the proposed scheme during both the training and online stages. In the training phase, the initial step involves generating the dataset. The dataset is generated by solving a minimization problem using conventional optimization methods, such as combining the PSO algorithm and SDR method. The features of the dataset consist of the rate and EH requirements, along with a reduced representation of the channel vectors obtained through PCA. The target values in this dataset correspond to the common rate, α_k. Subsequently, the K-fold cross-validation method is used to partition the dataset into training and validation subsets. These subsets are used to determine the optimal hyperparameters for the DNN module, including the number of hidden layers, hidden nodes, learning rates, batch size, and others. Once the DNN is trained, it is deployed at the BS for the online stage. During the online stage, the current channel vectors undergo dimensionality reduction via PCA. The reduced representations, rate, and EH requirements

serve as inputs for the trained DNN. The result of the DNN is the prediction of the common rate variables, α_k, which are used in the SDR module to optimize the PS factors, θ_k, and the beamforming vectors, $\mathbf{p}_0, \{\mathbf{p}_k, \mathbf{e}_g\}$.

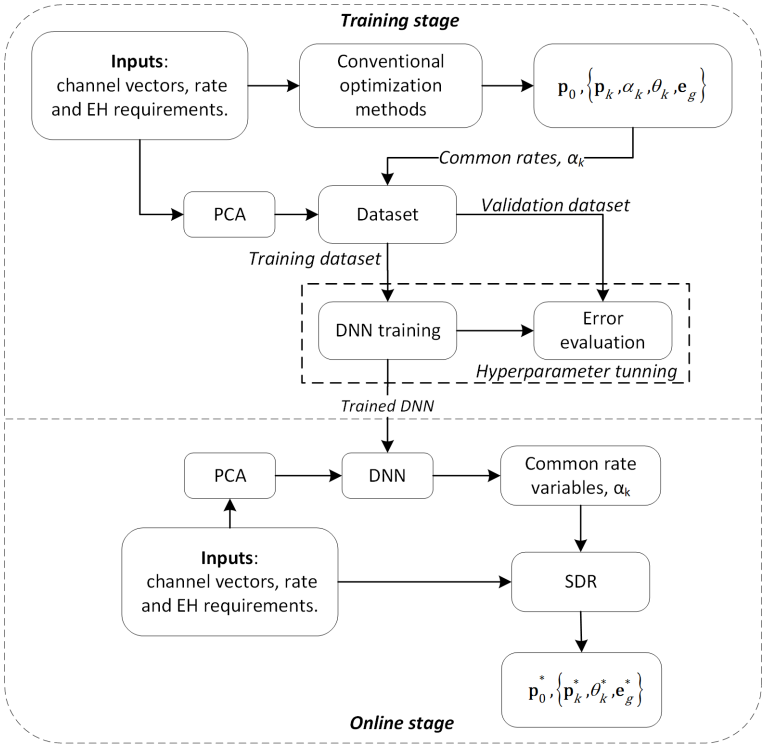

Figure 3. Overall procedure of the proposed approach.

Moreover, by utilizing the predicted common rate variables, α_k, the minimization problem (5) can be reformulated as follows:

$$\min_{\mathbf{p}_0, \{\mathbf{p}_k, \theta_k, \mathbf{e}_g\}} \sum_{i=0}^{K} \|\mathbf{p}_i\|^2 + \sum_{g=1}^{G} \|\mathbf{e}_g\|^2 \tag{7a}$$

subject to

$$\frac{\left|\mathbf{q}_{k,PS}^H \mathbf{p}_k\right|^2}{\sum_{i=1, i\neq k}^{K} \left|\mathbf{q}_{k,PS}^H \mathbf{p}_i\right|^2 + \sum_{g=1}^{G} \left|\mathbf{h}_{k,PS}^H \mathbf{e}_g\right|^2 + \sigma_k^2 + \frac{\gamma_k^2}{\theta_k}} \geq \omega_k, \forall k \tag{7b}$$

$$\frac{\left|\mathbf{q}_{k,PS}^H \mathbf{p}_0\right|^2}{\sum_{i=1}^{K} \left|\mathbf{q}_{k,PS}^H \mathbf{p}_i\right|^2 + \sum_{g=1}^{G} \left|\mathbf{h}_{k,PS}^H \mathbf{e}_g\right|^2 + \sigma_k^2 + \frac{\gamma_k^2}{\theta_k}} \geq \varphi, \forall k \tag{7c}$$

along with (5d), (5e) and (5g),

where $\varphi = 2^{\sum_{i=1}^{K} \alpha_i} - 1$, and $\omega_k = \max\{0, 2^{\chi_k - \alpha_k} - 1\}$.

3.1. Deep Learning-Based Scheme for Optimizing α_k

This subsection describes the PCA module for dimensionality reduction and the DNN-based method to predict the common rate variables, α_k.

Figure 4 presents the PCA module used to derive a reduced representation of each channel vector. The PCA technique [36] projects the input data onto a lower-dimensional subspace to maximize the variance of the projected data. The training dataset is defined

as $\mathbf{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_n, \ldots, \mathbf{x}_{N_D}\}$, where N_D is the total number of channel vectors used as the training dataset, and $\mathbf{x}_n$ is a W-dimensional vector composed of the real and imaginary components of the nth channel. The PCA module maps a channel vector onto an L-dimensional subspace that satisfies $L < W$. The projection of $\mathbf{x}_n$ can be represented by $\mathbf{y}_n = \mathbf{U}_L^T \mathbf{x}_n$, where $\mathbf{U}_L = [\mathbf{u}_1^T, \ldots, \mathbf{u}_L^T]$ with $\mathbf{u}_l^T \mathbf{u}_l = 1$ for $l = 1, \ldots, L$. A $\mathbf{U}_L$ that maximizes the trace of the covariance matrix of $\{\mathbf{y}_n\}$ is needed to maximize the variance of $\{\mathbf{y}_n\}$, which can be expressed as:

$$\mathbf{U}_L^* = \arg\max_{\mathbf{U}_L} \left(\mathrm{tr}\left(\mathbf{C}_\mathbf{y}\right)\right), \tag{8}$$

where $\mathbf{C}_\mathbf{y} = \frac{1}{N_D} \sum_{n=1}^{N_D} (\mathbf{y}_n - \bar{\mathbf{y}})(\mathbf{y}_n - \bar{\mathbf{y}})^T$, and $\bar{\mathbf{y}} = \frac{1}{N_D} \sum_{n=1}^{N_D} \mathbf{x}_n$. Let $\mathbf{C}_\mathbf{x} = \frac{1}{N_D} \sum_{n=1}^{N_D} (\mathbf{x}_n - \bar{\mathbf{x}})(\mathbf{x}_n - \bar{\mathbf{x}})^T$ represent the covariance matrix of $\{\mathbf{x}_n\}$, and $\bar{\mathbf{x}} = \frac{1}{N_D} \sum_{n=1}^{N_D} \mathbf{x}_n$. The solution of (8), described in [36,37], is obtained when

$$\mathbf{C}_\mathbf{x} \mathbf{u}_l = \lambda_l \mathbf{u}_l, \tag{9}$$

which means that the variance of the projected data is maximized when $\mathbf{u}_l$ is an eigenvector of $\mathbf{C}_\mathbf{x}$. Therefore, the optimal $\mathbf{U}_L^*$ is a matrix composed of the first L eigenvectors of $\mathbf{C}_\mathbf{x}$ as columns. Algorithm 1 summarizes the steps for performing the PCA technique.

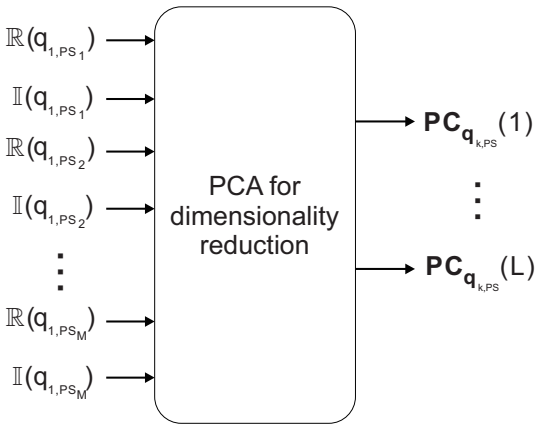

Figure 4. PCA module representation for the channel $\mathbf{q}_{1,PS}$.

Algorithm 1 PCA for dimensionality reduction

1: **inputs:** Training dataset $\mathbf{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_n, \ldots, \mathbf{x}_{N_D}\}$ containing the wireless channels and the number of components to keep, L.
2: Evaluate the mean $\bar{\mathbf{x}}$.
3: Calculate the covariance matrix $\mathbf{C}_\mathbf{x}$.
4: Decompose $\mathbf{C}_\mathbf{x}$ to obtain the eigenvectors and eigenvalues.
5: Select L eigenvectors corresponding to the L largest eigenvalues and create $\mathbf{U}_L$.
6: Map the W-dimensional channel vector into an L-dimensional representation by $\mathbf{y}_n = \mathbf{U}_L^T \mathbf{x}_n$.
7: **output:** Reduced representation of the channel vectors $\mathbf{q}_{k,PS}, \mathbf{q}_{g,EH}, \mathbf{h}_{k,PS}, \mathbf{h}_{g,EH}$ by performing step 6 for each channel vector.

Figure 5 shows the DNN model for the proposed approach, composed of an input layer, H_D hidden layers, and an output layer. The number of nodes in the input layer is determined by the user's rate requirements, EH requirements, and the reduced representation of the channel vectors after the PCA module, $\left\{\mathbf{PC}_{\mathbf{q}_{k,PS}}, \mathbf{PC}_{\mathbf{q}_{g,EH}}, \mathbf{PC}_{\mathbf{h}_{k,PS}}, \mathbf{PC}_{\mathbf{h}_{g,EH}}\right\}$. Each hidden layer consists of N_H hidden nodes, where the number of hidden layers and the number of hidden nodes are determined by fine-tuning the hyperparameters, as discussed in Section 4. The common rate variables specify the number of nodes in the output

layer, $\{\alpha_k\}$. During the training stage, the weights of the DNN are optimized using the backpropagation algorithm [38], given a training dataset with $\{\alpha_k\}$ as the real-valued targets. In conjunction with the grid search method, the K-fold cross-validation method is employed to select the best hyperparameters, including the number of hidden layers, hidden nodes, learning rates, and activation functions. Subsequently, during the online stage, after successfully training the DNN, the predicted common rate variables for the current channel vectors are constrained to feasible ranges, defined by the constraints (5b) and (5g), and are expressed as follows:

$$\alpha_k = \begin{cases} 0, & \text{if } \alpha_k < 0, \ \forall k, \quad (10a) \\ \alpha_k, & \text{if } 0 \le \alpha_k \le \chi_k, \ \forall k, \quad (10b) \\ \chi_k, & otherwise, \ \forall k \quad (10c) \end{cases}$$

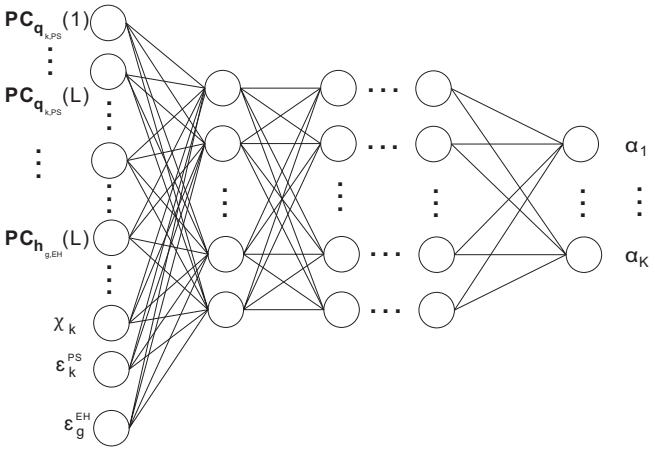

Figure 5. DNN module scheme for the proposed approach.

The modified common rate variables are passed to the SDR-based scheme to optimize the PS ratios, θ_k, and the beamforming vectors, $\mathbf{p}_0, \{\mathbf{p}_k, \mathbf{e}_g\}$. Within the DNN, we analyze the computational complexity in the online stage, as the training process takes place offline. This complexity is closely tied to the number of nodes and layers within the DNN. In particular, there are a total of $L(2K + 2G) + 2K + G$ nodes in the input layer, K nodes in the output layer, and each ith hidden layer has $N_{H,i}$ nodes, where $i = 1, \ldots, H_D$. Consequently, the computational complexity of the DNN module can be expressed as $\mathcal{O}((L(2K + 2G) + 2K + G)N_{H,1} + N_{H,1}N_{H,2} + \ldots + N_{H,H_D}K)$, even though this can be approximated to $\mathcal{O}(L(2K + 2G))$ when the DNN parameters are considered fixed.

3.2. SDR-Based Method for Optimizing $\mathbf{p}_0, \{\mathbf{p}_k, \mathbf{e}_g, \theta_k\}$

We introduce an approach based on the SDR technique to jointly optimize the PS factors and beamforming vectors, with a fixed value for $\{\alpha_k\}$. The matrix variables are denoted as $\mathbf{P}_i = \mathbf{p}_i\mathbf{p}_i^H$, $\mathbf{E}_g = \mathbf{e}_g\mathbf{e}_g^H$, $\mathbf{Q}_{k,PS} = \mathbf{q}_{k,PS}\mathbf{q}_{k,PS}^H$, $\mathbf{H}_{k,PS} = \mathbf{h}_{k,PS}\mathbf{h}_{k,PS}^H$, $\mathbf{Q}_{g,EH} = \mathbf{q}_{g,EH}\mathbf{q}_{g,EH}^H$, and $\mathbf{H}_{g,EH} = \mathbf{h}_{g,EH}\mathbf{h}_{g,EH}^H$. In addition, the matrix variable $\mathbf{P}_i$ enforces the conditions $\mathbf{P}_i \succeq 0$, $\text{rank}(\mathbf{P}_i) = 1$, while matrix $\mathbf{E}_g$ enforces $\mathbf{E}_g \succeq 0$ and $\text{rank}(\mathbf{E}_g) = 1$. Problem (7) can be transformed into a convex problem by removing the rank-one constraints and given $\{\alpha_k\}$, as follows:

$$\min_{\mathbf{P}_0, \{\mathbf{P}_k, \theta_k, \mathbf{E}_g\}} \ \sum_{i=0}^{K} \text{Tr}(\mathbf{P}_i) + \sum_{g=1}^{G} \text{Tr}(\mathbf{E}_g) \quad (11a)$$

subject to

$$\left(\sum_{i=1, i\neq k}^{K} \text{Tr}(\mathbf{Q}_{k,PS}\mathbf{P}_i) + \sum_{g=1}^{G} \text{Tr}(\mathbf{H}_{k,PS}\mathbf{E}_g) + \sigma_k^2 + \frac{\gamma_k^2}{\theta_k} \right) \omega_k - \text{Tr}(\mathbf{Q}_{k,PS}\mathbf{P}_k) \le 0, \ \forall k \quad (11b)$$

$$\left(\sum_{i=1}^{K}\mathrm{Tr}(\mathbf{Q}_{k,PS}\mathbf{P}_i)+\sum_{g=1}^{G}\mathrm{Tr}\left(\mathbf{H}_{k,PS}\mathbf{E}_g\right)+\sigma_k^2+\frac{\gamma_k^2}{\theta_k}\right)\varphi-\mathrm{Tr}(\mathbf{Q}_{k,PS}\mathbf{P}_0)\le 0,\ \forall k \tag{11c}$$

$$\frac{\varepsilon_k^{PS}}{\varsigma_k^{PS}(1-\theta_k)}-\sum_{i=0}^{K}\mathrm{Tr}(\mathbf{Q}_{k,PS}\mathbf{P}_i)-\sum_{g=1}^{G}\mathrm{Tr}\left(\mathbf{H}_{k,PS}\mathbf{E}_g\right)-\sigma_k^2\le 0,\ \forall k \tag{11d}$$

$$-\varsigma_g^{EH}\sum_{i=0}^{K}\mathrm{Tr}\left(\mathbf{Q}_{g,EH}\mathbf{P}_i\right)-\varsigma_g^{EH}\sum_{j=1}^{G}\mathrm{Tr}\left(\mathbf{H}_{g,EH}\mathbf{E}_j\right)+\varepsilon_g^{EH}\le 0,\ \forall g \tag{11e}$$

$$\mathbf{P}_0,\mathbf{P}_k,\mathbf{E}_g\succeq 0,\ \forall k,\ \forall g \tag{11f}$$

$$0<\theta_k<1,\ \forall k. \tag{11g}$$

Problem (11) is convex and can be solved efficiently using the CVX toolbox in MATLAB [39]. Problem (11) involves $K+1$ matrices of size $M\times M$ and G matrices of size $N\times N$, along with the $L_C=3K+G$ linear constraint variables. Therefore, the computational complexity of addressing problem (11) amounts to $\mathcal{O}\Big(\sqrt{(K+1)M+GN}\Big((K+1)^3M^6+G^3N^6+L_c\big((K+1)M^2+GN^2\big)\Big)\log(1/\zeta)\Big)$ while maintaining a solution accuracy of $\zeta>0$ [40]. The optimal solutions are represented in problem (11) as $\left\{\mathbf{P}_k^*,\mathbf{E}_g^*\right\}$. When the matrix solutions have a rank of one, the optimal beamforming vectors, $\left\{\mathbf{p}_i^*,\mathbf{e}_g^*\right\}$, are given by

$$\mathbf{p}_i=\sqrt{\lambda_H(\mathbf{P}_i)}\mathbf{v}_{H,\mathbf{P}_i},\ i=0,\ldots,K \tag{12a}$$

$$\mathbf{e}_g=\sqrt{\lambda_H\left(\mathbf{E}_g\right)}\mathbf{v}_{H,\mathbf{E}_g},\ g=1,\ldots,G, \tag{12b}$$

where $\lambda_H(\mathbf{A})$ represents the largest eigenvalue of the matrix $\mathbf{A}$, and $\mathbf{v}_{H,\mathbf{A}}$ denotes its corresponding eigenvector. If $\left\{\mathbf{P}_i^*,\mathbf{E}_g^*\right\}$ are rank-one, alternative approaches, such as the penalty function method [41] or the Gaussian randomization technique [20], can be employed to approximate the beamforming vectors.

The following details the penalty function method for the scenario where the matrix variables $\left\{\mathbf{P}_i^*,\mathbf{E}_g^*\right\}$ are not rank-one. The proposed penalty-based method is based on the definition of $\left\{\mathbf{P}_i^*,\mathbf{E}_g^*\right\}$ being positive semidefinite matrices, which satisfy the conditions $\mathrm{Tr}(\mathbf{P}_i)\ge\lambda_H(\mathbf{P}_i)$ and $\mathrm{Tr}\left(\mathbf{E}_g\right)\ge\lambda_H\left(\mathbf{E}_g\right)$, where the matrices $\left\{\mathbf{P}_i^*,\mathbf{E}_g^*\right\}$ are rank-one if $\mathrm{Tr}(\mathbf{P}_i)=\lambda_H(\mathbf{P}_i)$ and $\mathrm{Tr}\left(\mathbf{E}_g\right)=\lambda_H\left(\mathbf{E}_g\right)$. Hence, the proposed penalty-based approach aims to minimize $\mathrm{Tr}(\mathbf{P}_i)-\lambda_H(\mathbf{P}_i)$ and $\mathrm{Tr}\left(\mathbf{E}_g\right)-\lambda_H\left(\mathbf{E}_g\right)$ by introducing a penalty factor, κ, and incorporating these terms into the objective function as follows:

$$\begin{aligned}\min_{\mathbf{P}_0,\left\{\mathbf{P}_k,\theta_k,\mathbf{E}_g\right\}}\ &\sum_{i=0}^{K}\mathrm{Tr}(\mathbf{P}_i)+\sum_{g=1}^{G}\mathrm{Tr}\left(\mathbf{E}_g\right)+\kappa\sum_{i=0}^{K}\left(\mathrm{Tr}(\mathbf{P}_i)-\lambda_H(\mathbf{P}_i)\right)+\kappa\sum_{g=1}^{G}\left(\mathrm{Tr}\left(\mathbf{E}_g\right)-\lambda_H\left(\mathbf{E}_g\right)\right)\\ &\text{subject to (11b), \ldots, (11g).}\end{aligned} \tag{13}$$

Problem (13) is non-convex because of the terms $-\lambda_H(\mathbf{P}_i)$ and $-\lambda_H\left(\mathbf{E}_g\right)$. Therefore, the following inequality that holds for any $\mathbf{Z}_k\ge 0$ is used to handle the aforementioned non-convex terms:

$$\lambda_H(\mathbf{P}_i)\ge\lambda_H(\mathbf{Z}_i)+\mathbf{v}_{H,\mathbf{Z}_i}^H(\mathbf{P}_i-\mathbf{Z}_i)\mathbf{v}_{H,\mathbf{Z}_i}, \tag{14}$$

where $\mathbf{v}_{H,\mathbf{Z}_i}$ represents the unit-norm eigenvector associated with the largest eigenvalue of $\mathbf{Z}_i$. Hence, (14) is used to approximate $\lambda_H(\mathbf{P}_i)$ based on a feasible matrix $\mathbf{P}_i^{(j)}$ at the jth iteration as follows:

$$\lambda_H(\mathbf{P}_i)\ge\lambda_H\left(\mathbf{P}_i^{(j)}\right)+\mathbf{v}_{H,\mathbf{P}_i^{(j)}}^H\left(\mathbf{P}_i-\mathbf{P}_i^{(j)}\right)\mathbf{v}_{H,\mathbf{P}_i^{(j)}},\ \forall_k. \tag{15}$$

By applying a similar procedure as in (15), the term $\lambda_H(\mathbf{E}_g)$ can be equivalently transformed. As a result, problem (13) can be reformulated as follows:

$$\begin{aligned}&\min_{\mathbf{P}_0,\{\mathbf{P}_k,\theta_k,\mathbf{E}_g\}} \sum_{i=0}^{K}\mathrm{Tr}(\mathbf{P}_i)+\sum_{g=1}^{G}\mathrm{Tr}(\mathbf{E}_g)\\&+\kappa\sum_{i=0}^{K}\left(\mathrm{Tr}(\mathbf{P}_i)-\lambda_H\left(\mathbf{P}_i^{(j)}\right)+\mathbf{v}^H_{H,\mathbf{P}_i^{(j)}}\left(\mathbf{P}_i-\mathbf{P}_i^{(j)}\right)\mathbf{v}_{H,\mathbf{P}_i^{(j)}}\right)\\&+\kappa\sum_{g=1}^{G}\left(\mathrm{Tr}(\mathbf{E}_g)-\lambda_H\left(\mathbf{E}_g^{(j)}\right)+\mathbf{v}^H_{H,\mathbf{E}_g^{(j)}}\left(\mathbf{E}_g-\mathbf{E}_g^{(j)}\right)\mathbf{v}_{H,\mathbf{E}_g^{(j)}}\right)\\&\text{subject to (11b), ..., (11g).}\end{aligned} \tag{16}$$

Note that problem (16) is convex, and the solution can be obtained using the CVX toolbox in MATLAB. Algorithm 2 lists the proposed iterative scheme based on the penalty method, which is applied when the solutions to problem (11), $\left\{\mathbf{P}_i^*,\mathbf{E}_g^*\right\}$, are not rank-one.

Algorithm 2 Penalty-based method for solving problem (11)

1: **inputs:** Matrix solutions of problem (11), $\left\{\mathbf{P}_i^*,\mathbf{E}_g^*\right\}$, tolerance value, ϕ, penalty factor, κ, channel vectors, rate, and EH requirements.
2: Set iteration counter $j=0$.
3: Set initial feasible matrices, $\mathbf{P}_i^{(j)}=\mathbf{P}_i^*, \mathbf{E}_g^{(j)}=\mathbf{E}_g^*$.
4: **repeat**
5: Solve problem (16) given the feasible matrices $\left\{\mathbf{P}_i^{(j)},\mathbf{E}_g^{(j)}\right\}$ and denote the solutions as $\left\{\mathbf{P}_i^*,\mathbf{E}_g^*\right\}$.
6: Increase counter $j=j+1$.
7: Update the feasible matrices for the next iteration: $\mathbf{P}_i^{(j)}=\mathbf{P}_i^*, \mathbf{E}_g^{(j)}=\mathbf{E}_g^*$.
8: **until** $\sum_{i=0}^{K}\left(\mathrm{Tr}(\mathbf{P}_i)-\lambda_H(\mathbf{P}_i)\right)+\sum_{g=1}^{G}\left(\mathrm{Tr}(\mathbf{E}_g)-\lambda_H(\mathbf{E}_g)\right)\le\phi$
9: Obtain the beamforming vectors with (12a) and (12b).
10: **Output:** $\left\{\mathbf{p}_i^*,\mathbf{e}_g^*\right\}$.

3.3. PSO-Based Approach for Optimizing $\{\alpha_k\}$ with a Given $\mathbf{p}_0, \{\mathbf{p}_k, \mathbf{e}_g, \theta_k\}$

This subsection presents a comparative scheme for optimizing the common rate variables using a PSO-based approach [20] with given PS factors and beamforming vectors. PSO is a potent metaheuristic algorithm inspired by the social behavior of flocking birds, where the collective knowledge of the swarm guides each particle through the search space to discover the optimal solution. In the proposed PSO-based scheme, there is a population of S particles whose position represents the common ratio variables to be optimized, i.e., the position of the sth particle is given by $\mathbf{x}_s=\left[\alpha_1^s,\ldots,\alpha_K^s\right], s=1,\ldots,S$. The position of each particle is initialized randomly within the range $[0,\chi_k]$. The local best position for the sth particle, denoted as $\mathbf{x}_s^l$, represents the best location of the sth particle. Moreover, the global best position, denoted as $\mathbf{x}_g$, represents the best location of all the swarm particles. The location of the sth particle is based on its velocity and is expressed as

$$\mathbf{v}_s\leftarrow w\mathbf{v}_s+a_cu_1^n\left(\mathbf{x}_s^l-\mathbf{x}_s\right)+b_cu_2^n\left(\mathbf{x}_g-\mathbf{x}_s\right), \tag{17}$$

where w is the inertia weight; a_c, b_c are the acceleration parameters; and $u_1^n, u_2^n\sim U(0,1)$. Subsequently, the position of the sth particle is modified according to the following:

$$\mathbf{x}_s\leftarrow\mathbf{x}_s+\mathbf{v}_s. \tag{18}$$

The objective function, $f(\mathbf{x}_s)$, is defined by the sum transmission power (7a), obtained by solving problem (11), with $\{\alpha_k\}$ being determined by the position of the sth particle, $\mathbf{x}_s$. The computational complexity of the PSO-based method is given by $\mathcal{O}(S \cdot T_{\max} \cdot \mathcal{O}_{SDR})$, where $\mathcal{O}_{SDR}$ represents the complexity of solving problem (7) using the SDR method detailed in Section 3.2. Algorithm 3 presents the algorithm based on PSO designed to optimize $\{\alpha_k\}$.

Algorithm 3 Comparative scheme based on PSO for optimizing $\{\alpha_k\}$

inputs: Number of particles, S, maximum number of iterations, $T_{\max}$, rate, and EH requirements.
Set iteration counter $t = 0$ and initialize the position of each particle and its velocity, $\mathbf{v}_{s,t} = 0$.
Set the initial $\mathbf{x}^l_{s,t} = \mathbf{x}_{s,t}$ and $\mathbf{x}_{g,t} = \arg\min_{1 \le s \le S} f\left(\mathbf{x}^l_{s,t}\right)$.
While $t \le T_{\max}$ **do**
  for $s = 1, \ldots, S$, **do**
    Update the velocity and the position based on (17) and (18) to obtain $\mathbf{v}_{s,t+1}$ and $\mathbf{x}_{s,t+1}$.
    Restrict the value of the position based on (10).
    Solve problem (11) with the common rates given by $\mathbf{x}_{s,t+1}$ to get $f(\mathbf{x}_{s,t+1})$.
    if $f(\mathbf{x}_{s,t+1}) < f\left(\mathbf{x}^l_{s,t}\right)$ **then**
      $\mathbf{x}^l_{s,t} = \mathbf{x}_{s,t+1}$
    end if
  end for
  $\mathbf{x}_{g,t} = \arg\min_{1 \le s \le S} f\left(\mathbf{x}^l_{s,t}\right)$
  Increase counter $t = t + 1$.
end while
Output: Best common rates, $\left\{\alpha^*_k\right\}$, defined by $\mathbf{x}_g$.

4. Simulation Results

Numerical simulations were conducted to assess the performance of the proposed DL-based solution in minimizing the sum transmission power of the BS and PB in the considered MISO SWIPT system with RSMA. The simulation parameters were set to $K = 3$ PS users, $G = 2$ EH users, $M = 8$ antennas at the BS, $N = 8$ antennas at the PB, $\sigma_k^2 = \gamma_k^2 = -60$ dBm, $\varsigma_k^{PS} = \varsigma_g^{EH} = 1$, $\chi_k = \chi$, and $\varepsilon_k^{PS} = \varepsilon_g^{EH} = \varepsilon$. The channel vector between the kth PS user and the BS is given by

$$\mathbf{q}_{k,PS} = \sqrt{\beta d^{-\nu}_{BS-PS,k}}\,\tilde{\mathbf{q}}_{k,PS}, \ \forall k, \tag{19}$$

where $\nu = 2.2$ defines the path-loss exponent; $\beta = 10^{-3}$; $d^{-\nu}_{BS-PS,k}$ denotes the distance between the BS and the kth PS user; and $\tilde{\mathbf{q}}_{k,PS}$ is subject to independent Rician fading. $\mathbf{q}_{g,EH}, \mathbf{h}_{k,PS}, \mathbf{h}_{g,EH}$ are channels established according to (13). The BS is positioned at coordinates (8 m, 20 m), whereas the PB is located at (13 m, 7 m). PS users are randomly distributed within a designated region defined by $x_{PS} \in [13\text{ m}, 18\text{ m}]$ and $y_{PS} \in [13\text{ m}, 25\text{ m}]$. Similarly, EH users are randomly distributed within an area bounded by $x_{EH} \in [17\text{ m}, 22\text{ m}]$ and $y_{EH} \in [2\text{ m}, 12\text{ m}]$.

The effectiveness of the proposed DL-based method was evaluated by comparing it with a PSO-SDR method and conventional techniques, such as SDMA and NOMA. The proposed scheme is denoted as DNN-RSMA, and the comparative scheme based on PSO and SDR is denoted as PSO-RSMA. In SDMA [26], the kth PS user's message, W_k, is encoded directly into the data stream, z_k^{PS}, without common rate variables, and interference originating from other users is considered noise. In addition, a comparative

scenario was incorporated by assuming the absence of the power beacon (PB) deployment in the network.

In NOMA [27], interference originating from other users is mitigated by employing multiple layers in the SIC process. The message intended for the kth PS user, W_k, is encoded into the private stream, $z_{k,N}^{PS}$. The transmitted signal at the BS is given by $\mathbf{x}_{BS,NOMA} = \sum_{k=1}^{K} \mathbf{p}_{k,N} z_{k,N}^{PS}$, where $\mathbf{p}_{k,N} \in \mathbb{C}^{M\times 1}$ is the information beamforming vector for $z_{k,N}^{PS}$. The signal transmitted at the PB is the same as in the proposed RSMA-based framework. On the user side, the decoding order for PS users is determined by their respective channel strengths, denoted as $\|\mathbf{q}_{1,PS}\| \geq \ldots \geq \|\mathbf{q}_{K,PS}\|$. In particular, the kth PS user begins by decoding messages intended for the Kth, $(K-1)$th, …, $(k+1)$th PS users. Subsequently, the kth PS user proceeds to decode its intended message while treating the interference from the remaining messages as noise. The achievable rate to decode the lth message at the kth PS user is given by

$$\mathrm{R}_{k,NOMA}^{l} = \log_2\left(1 + \frac{\left|\mathbf{q}_{k,PS}^{H}\mathbf{p}_{l,N}\right|^2}{\sum_{k'=1}^{l-1}\left|\mathbf{q}_{k,PS}^{H}\mathbf{p}_{k',N}\right|^2 + \sum_{g=1}^{G}\left|\mathbf{h}_{k,PS}^{H}\mathbf{e}_{g}\right|^2 + \sigma_k^2 + \frac{\gamma_k^2}{\theta_k}}\right), \forall k, \forall l \geq k. \quad (20)$$

Two datasets were considered in the simulation analysis: a training/validation dataset and a testing dataset. As described in Section 3, the PSO-SDR method was used to process all the datasets considered. A fivefold cross-validation method was used to divide the training/validation dataset into separate training and validation datasets. These datasets were used to determine the optimal hyperparameters for the DNN module, including the number of hidden layers, hidden nodes, learning rates, activation functions, and batch sizes. Subsequently, the testing dataset was used to evaluate the ability of the model to generalize. The testing dataset encompassed scenarios not encountered during the training and validation phases to assess the generalization performance of the model. These scenarios included different user distances and varying numbers of antennas at the BS. The training/validation dataset consisted of 14,000 samples, covering data rates ranging from $\chi = 1$ bits/s/Hz to $\chi = 10$ bits/s/Hz, EH requirements from $\varepsilon = -16$ dBm to $\varepsilon = -30$ dBm, eight antennas at the BS, and the previously mentioned range of distances. The specifics of the testing dataset for each scenario are outlined later in this paper. In addition, the exponential linear unit (ELU) was selected as the activation function for the hidden layers, and the Adam algorithm was used as the optimizer. The training data for the PCA module were composed of 6000 channel vectors. Regarding the PSO-SDR method, the simulation parameters considered for the PSO algorithm after fine-tuning the parameters were $S = 15$, $w = 0.7$, $a_c = b_c = 1.494$, and $T_{max} = 25$.

Firstly, this paper presents an analysis of the convergence behavior of the DNN module while considering different hyperparameters. Figure 6 shows the convergence performance of the DNN module under varying learning rates (lr), numbers of hidden layers, and numbers of principal components (PCs) in the PCA module. In the legend in Figure 6, the last term represents the number of units × the number of hidden layers, where all hidden layers have the same number of hidden units. Regarding the learning rate, a high learning rate leads to rapid convergence, but it can also result in overfitting, as observed in the case of lr = 0.005. On the other hand, low learning rate values require a significantly higher number of epochs to achieve convergence, and they may not guarantee the lowest error value, as observed in the case of lr = 0.0001. Regarding the number of PCs, it is important to note that the PCA module was applied to each channel vector, resulting in each channel vector being represented by a certain number of PCs. The lowest error was achieved when the number of PCs was 2 or 3. In particular, when PC = 2, it implies a lower number of required input nodes in the DNN module. For the remainder of the simulations, a learning rate of lr = 0.001 and PC = 2 were selected based on these observations.

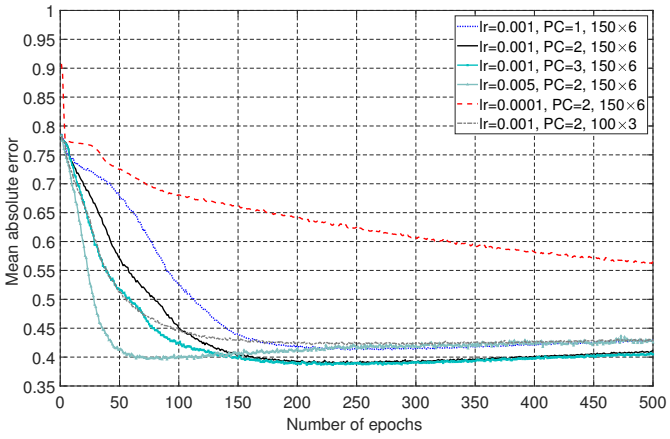

Figure 6. Convergence behavior of the proposed DNN module.

We utilized the recurrent neural network (RNN) model [38] in our simulation results for comparative analysis. RNN, a type of neural network, is specifically designed to analyze sequences of vectors, where values in successive vectors exhibit interrelationships. However, it is noteworthy that each sample in our dataset comprised a singular vector containing the user's rate requirements, EH requirements, and the reduced representation of the channel vectors post-PCA module. To incorporate the RNN into our simulations, we configured the number of vectors in the sequence to one and set the size of the hidden layer to 150.

Figure 7 shows the variation of the sum transmission power of the PB and BS with respect to the data rate requirements, χ, considering a minimum EH requirement of $\varepsilon = -24$ dBm. The data rate requirement serves as a constraint in the proposed optimization problem (5) and must be satisfied by the BS for all users in the system. As the rate requirement increases, the BS must allocate more transmission power to meet the specified rate, thereby increasing the total transmission power. Although the PB does not transmit information signals, their energy-carrying signals aid in meeting the EH requirements but are considered interference at the ID module of PS users. In particular, we observed that deploying the PB enabled a reduction of up to 3 dBm in the total transmission power for all the multiple-access methods compared. Furthermore, the RSMA scheme achieved significantly lower transmission power with and without the PB compared to their respective SDMA- and NOMA-based counterparts. Moreover, the proposed DNN-RSMA-based solution achieved a similar result to the near-optimal scheme of the PSO-SDR RSMA while significantly reducing the computational complexity. In particular, the complexity of the PSO-SDR method is denoted as $\mathcal{O}(S \cdot T_{\max} \cdot \mathcal{O}_{SDR})$, and the complexity of the DNN-based method is represented as $\mathcal{O}(\mathcal{O}_{DNN} + \mathcal{O}_{SDR})$, where $\mathcal{O}_{SDR}$ is the computational complexity of solving problem (7) using the SDR method, and $\mathcal{O}_{DNN}$ represents the computational complexity of predicting the common rate variables. Consequently, the proposed DNN-based method was approximately $S \times T_{\max}$ times faster than the PSO-SDR near-optimal scheme. In addition, we observed that the RNN model achieved performance comparable to that of the proposed DNN. This similarity arises because when the input is a single vector, the RNN exhibits a layer-wise architecture similar to that of a basic neural network. Furthermore, the proposed DNN demonstrated lower computational time compared to the RNN. Specifically, the RNN required a computational time of 191.232 s for training and 0.002 s for testing, whereas the proposed DNN required a computational time of 40.295 s for training and 0.001 s for testing. These simulations were conducted on a computer equipped with an Intel Core i7-6700 CPU and 16 GB of RAM. Therefore, the proposed DNN emerges as the most suitable neural network architecture for solving the proposed optimization problem.

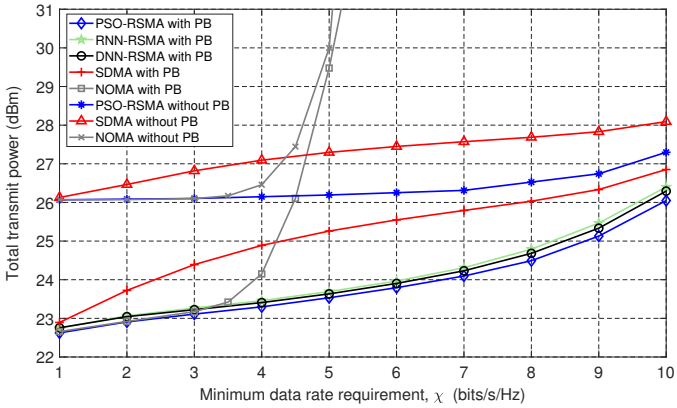

Figure 7. Total transmission power of the PB and BS versus the required data rate targets of PS users.

RSMA outperformed SDMA because of its capability to decode a portion of the interference through the SIC process on the common message (Figure 7), whereas SDMA treats all interference as noise. In the NOMA scheme, the difference arises from using the SIC process. In RSMA, the SIC process aims to cancel the interference from the common message. In contrast, in the NOMA scheme, the SIC process is employed to eliminate interference from the messages of users with weaker channel strengths. On the other hand, achieving excellent performance under the NOMA scheme in multi-antenna systems necessitates that users whose messages will be decoded by SIC have weaker channel strengths and sufficiently aligned channels. This alignment requirement is uncommon in real-world deployments and the channel model considered in Equation (13). Furthermore, as the rate requirement of the PS users increases, the transmission power needed for the beamforming vectors also increases, leading to performance degradation in the NOMA scheme after reaching $\chi = 4$ bits/s/Hz.

The variation of the sum transmission power of the PB and BS with respect to the EH requirements, ε, is presented in Figure 8, considering a minimum rate requirement of $\chi = 4$ bits/s/Hz. Similar to Figure 7, a significant reduction of approximately 3 dBm was observed in the total transmission power due to the deployment of the PB. Furthermore, the benefit of applying RSMA compared to conventional methods, such as NOMA and SDMA, in considerably reducing the transmission power was demonstrated. This is because RSMA employed the SIC procedure, which improved the data rate at the ID module of PS users while simultaneously reducing the PS factor to enhance harvested energy, thus reducing the transmission power. Furthermore, the proposed DNN-based approach performed similarly to the PSO-based method while significantly reducing computational complexity.

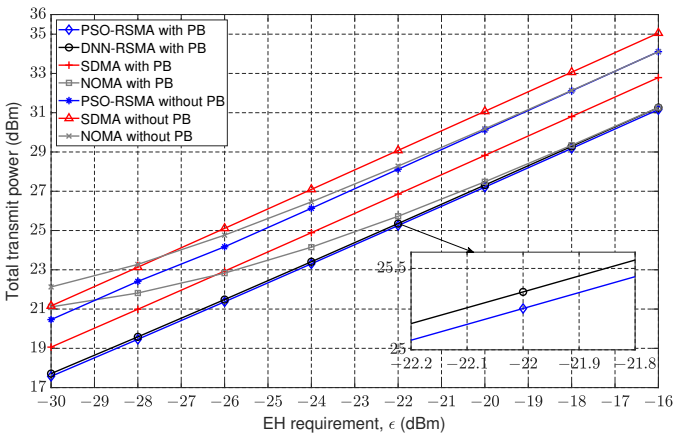

Figure 8. Total transmission power of the PB and BS versus the required EH of PS and EH users.

Next, we analyzed the generalization performance of the proposed DNN-based method across considerably different scenarios from those used in the training phase. In particular, we kept the trained DNN module fixed and tested its generalization performance by varying the number of antennas, resulting in a modification of the components of the channel vectors, and by varying the distance from the BS, resulting in a significant alteration of the channel strengths. The channel vectors, along with the requirements for data rates and energy harvesting, served as inputs for the PCA and DNN modules in the proposed scheme to generate the common rate variables, as detailed in Section 3.1.

Figure 9 presents the variation of the sum transmission power of the PB and BS with respect to the number of antennas equipped at the BS, considering a required minimum rate of $\chi = 4$ bits/s/Hz and a required minimum EH of $\varepsilon = -25$ dBm. As the number of antennas at the BS increased, there was a reduction in the transmission power due to the increased degrees of freedom. In the case of the DNN-based method, the training dataset solely consisted of samples representing a scenario with a BS equipped with eight antennas, as detailed at the beginning of Section 4. In contrast, the testing data in Figure 9 contained samples encompassing scenarios with varying numbers of antennas at the BS. The utilization of the same trained DNN module with different numbers of antennas can be attributed to the PCA module. In particular, independent PCA modules for each scenario were trained based on the number of antennas while maintaining the same number of principal components, L. Consequently, despite variations in the dimensionality of the channel vectors, the output of the PCA module remained consistent across all different numbers of antennas. Training a PCA module is a straightforward task because it only requires channel vector samples and does not necessitate labels or target values.

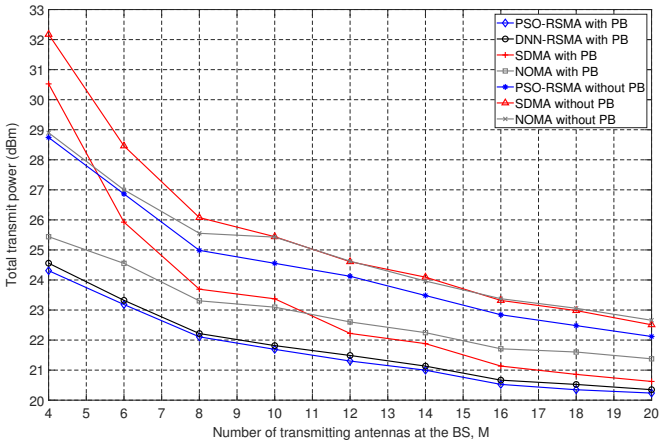

Figure 9. Total transmission power of the PB and BS versus the number of antennas at the BS.

The proposed DNN-based scheme demonstrated remarkable performance comparable to that of the PSO-based method, even when the number of antennas differed from the training scenarios (Figure 9). This indicates strong generalization capabilities and robustness to environmental changes. Furthermore, the RSMA-based methods consistently achieved lower total transmission power compared to the traditional NOMA and SDMA methods. Moreover, the deployment of the PB resulted in a significant decrease in the transmission power, even with an increasing number of antennas at the BS. This can be attributed to the ability of the PB to mitigate signal attenuation because of the distance from the transmitter, a critical factor in scenarios with EH requirements for users.

Figure 10 shows the variation of the sum transmission power of the PB and BS with respect to the position of the PS users, considering a minimum rate requirement of $\chi = 4$ bits/s/Hz and a required EH of $\varepsilon = -25$ dBm. In particular, the range of the position of the PS users on the x-axis, x_{PS}, was varied. This position was randomly selected within the region of $x_{PS} \in [d_x \text{ m}, (d_x + 5) \text{ m}]$, where d_x varied to analyze performance as the distance between the PS users and the BS increased. The transmission power increased

as d_x increased because the average distance from the user to the BS also increased with higher values of d_x. Additionally, the difference in the transmission power between scenarios with and without the PB diminished as the value of d_x increased, owing to the fixed location of the PB across all d_x values. As d_x increased, users were positioned farther from the PB, resulting in increased attenuation of the energy signal. Although the energy signal at the PS users was utilized for energy harvesting at the EH module, it was considered interference for the ID module. Consequently, as the received energy signal power from the PB diminished due to increased distance, it contributed less to the EH at the PS users. Meanwhile, reliance on the received power from the information signal transmitted by the BS became more dominant, leading to a reduction in the impact of the PB deployment. Moreover, as shown in the previous figures, the RSMA method consistently achieved a considerable decrease in the total transmission power in comparison to the NOMA and SDMA methods.

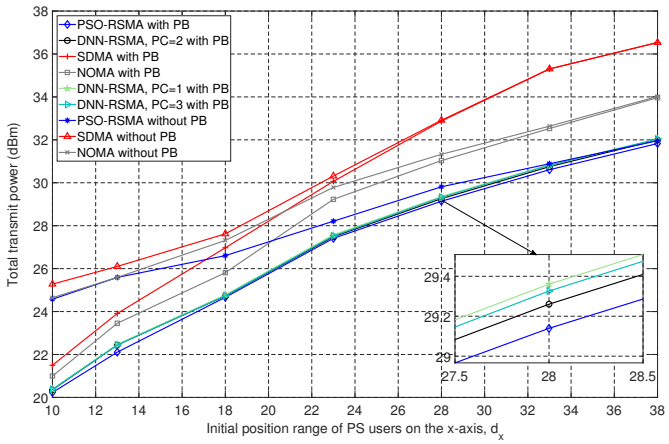

Figure 10. Total transmission power of the PB and BS versus the initial position range of PS users on the x-axis.

In the case of the DNN-based method, the training dataset consisted solely of samples representing a scenario with $d_x = 13$ m, as detailed at the beginning of Section 4. In contrast, the testing data in Figure 10 contained samples encompassing scenarios with varying values of d_x. The proposed DNN-based scheme exhibited high generalization performance, achieving comparable transmission power to the PSO-based method. Furthermore, this study analyzed the effect of slightly changing the number of PCs in the PCA module. Similar to the results shown in Figure 6, the cases of PC = 2 and PC = 3 achieved similar results, with PC = 2 having lower transmission power and demonstrating the best generalization performance.

5. Conclusions

A multiuser MISO SWIPT system using RSMA was evaluated with the assistance of a PB. The objective was to minimize the combined transmission power from the BS and PB while optimizing the beamforming vectors, common rate variables, and PS ratios. The proposed optimization problem was carried out under constraints that included EH requirements for both EH and PS users and data rate requirements for PS users. The proposed non-convex problem was divided into two parts. The first part was solved using the PCA method to reduce dimensionality and a DNN to predict the common rate variables. The second part used the SDR technique to optimize the PS factors and beamforming vectors. Comparative schemes were developed based on the PSO algorithm and SDR method for RSMA, along with a baseline scheme using NOMA.

Numerical simulations showed that RSMA significantly reduced the transmission power compared to conventional methods such as NOMA and SDMA. Moreover, the proposed DNN-based method achieved high performance, closely matching the results of the near-optimal PSO-based scheme while considerably reducing computational complexity.

Furthermore, this study tested the proposed DNN-based scheme across challenging scenarios that were significantly different from those used for training during data collection. Across these scenarios, the DNN-based scheme exhibited high generalization performance. Furthermore, strategically placing the PB close to EH users substantially decreased the overall transmit power, effectively meeting their EH requirements.

Author Contributions: Conceptualization, M.R.C. and C.E.G.; methodology, M.R.C.; software, M.R.C. and C.E.G.; validation, C.E.G. and I.K.; formal analysis, M.R.C.; investigation, M.R.C. and C.E.G.; resources, I.K.; data curation, M.R.C.; writing—original draft preparation, M.R.C.; writing—review and editing, C.E.G. and I.K.; visualization, M.R.C.; supervision, I.K.; project administration, I.K.; funding acquisition, I.K. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the National Research Foundation of Korea (NRF) through the Korean Government's Ministry of Science and ICT (MSIT) under Grant NRF-2021R1A2B5 B01001721, and in part by the Regional Innovation Strategy (RIS) through the NRF funded by the Ministry of Education (MOE) under Grant 2021RIS-003.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Mishra, A.; Mao, Y.; Dizdar, O.; Clerckx, B. Rate-Splitting Multiple Access for 6G—Part I: Principles, Applications and Future Works. *IEEE Commun. Lett.* **2022**, *26*, 2232–2236. [CrossRef]
2. Clerckx, B.; Joudeh, H.; Hao, C.; Dai, M.; Rassouli, B. Rate splitting for MIMO wireless networks: A promising PHY-layer strategy for LTE evolution. *IEEE Commun. Mag.* **2016**, *54*, 98–105. [CrossRef]
3. Mao, Y.; Clerckx, B.; Li, V.O. Rate-splitting multiple access for downlink communication systems: Bridging, generalizing, and outperforming SDMA and NOMA. *EURASIP J. Wirel. Commun. Netw.* **2018**, *2018* , 133. [CrossRef]
4. Mao, Y.; Dizdar, O.; Clerckx, B.; Schober, R.; Popovski, P.; Poor, H.V. Rate-Splitting Multiple Access: Fundamentals, Survey, and Future Research Trends. *IEEE Commun. Surv. Tutor.* **2022**, *24*, 2073–2126. [CrossRef]
5. Clerckx, B.; Mao, Y.; Schober, R.; Poor, H.V. Rate-splitting unifying SDMA, OMA, NOMA, and multicasting in MISO broadcast channel: A simple two-user rate analysis. *IEEE Wirel. Commun. Lett.* **2020**, *9*, 349–353. [CrossRef]
6. Ponnimbaduge Perera, T.D.; Jayakody, D.N.K.; Sharma, S.K.; Chatzinotas, S.; Li, J. Simultaneous Wireless Information and Power Transfer (SWIPT): Recent Advances and Future Challenges. *IEEE Commun. Surv. Tutor.* **2018**, *20*, 264–302. [CrossRef]
7. Bi, S.; Ho, C.K.; Zhang, R. Wireless powered communication: Opportunities and challenges. *IEEE Commun. Mag.* **2015**, *53*, 117–125. [CrossRef]
8. Huang, K.; Lau, V.K.N. Enabling Wireless Power Transfer in Cellular Networks: Architecture, Modeling and Deployment. *IEEE Trans. Wirel. Commun.* **2014**, *13*, 902–912. [CrossRef]
9. Pan, G.; Lei, H.; Yuan, Y.; Ding, Z. Performance Analysis and Optimization for SWIPT Wireless Sensor Networks. *IEEE Trans. Commun.* **2017**, *65*, 2291–2302. [CrossRef]
10. Shi, Q.; Liu, L.; Xu, W.; Zhang, R. Joint Transmit Beamforming and Receive Power Splitting for MISO SWIPT Systems. *IEEE Trans. Wirel. Commun.* **2014**, *13*, 3269–3280. [CrossRef]
11. Zhang, H.; Dong, A.; Jin, S.; Yuan, D. Joint Transceiver and Power Splitting Optimization for Multiuser MIMO SWIPT Under MSE QoS Constraints. *IEEE Trans. Veh. Technol.* **2017**, *66*, 7123–7135. [CrossRef]
12. Xu, Y.; Shen, C.; Ding, Z.; Sun, X.; Yan, S.; Zhu, G.; Zhong, Z. Joint Beamforming and Power-Splitting Control in Downlink Cooperative SWIPT NOMA Systems. *IEEE Trans. Signal Process.* **2017**, *65*, 4874–4886. [CrossRef]
13. Li, Z.; Chen, W.; Wu, Q.; Wang, K.; Li, J. Joint Beamforming Design and Power Splitting Optimization in IRS-Assisted SWIPT NOMA Networks. *IEEE Trans. Wirel. Commun.* **2022**, *21*, 2019–2033. [CrossRef]
14. Al-Obiedollah, H.M.; Cumanan, K.; Thiyagalingam, J.; Burr, A.G.; Ding, Z.; Dobre, O.A. Energy Efficient Beamforming Design for MISO Non-Orthogonal Multiple Access Systems. *IEEE Trans. Commun.* **2019**, *67*, 4117–4131. [CrossRef]
15. Joudeh, H.; Clerckx, B. Robust transmission in downlink multiuser MISO systems: A rate-splitting approach. *IEEE Trans. Signal Process.* **2016**, *64*, 6227–6242. [CrossRef]
16. Joudeh, H.; Clerckx, B. Sum-rate maximization for linearly precoded downlink multiuser MISO systems with partial CSIT: A rate-splitting approach. *IEEE Trans. Commun.* **2016**, *64*, 4847–4861. [CrossRef]
17. Dai, M.; Clerckx, B.; Gesbert, D.; Caire, G. A rate splitting strategy for massive MIMO with imperfect CSIT. *IEEE Trans. Wirel. Commun.* **2016**, *15*, 4611–4624. [CrossRef]
18. Mao, Y.; Clerckx, B.; Li, V.O.K. Rate-Splitting for Multi-Antenna Non-Orthogonal Unicast and Multicast Transmission: Spectral and Energy Efficiency Analysis. *IEEE Trans. Commun.* **2019**, *67*, 8754–8770. [CrossRef]

19. Mao, Y.; Clerckx, B.; Li, V.O.K. Rate-Splitting for Multiuser Multi-Antenna Wireless Information and Power Transfer. In Proceedings of the IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Cannes, France, 2–5 July 2019.
20. Camana Acosta, M.R.; Moreta, C.E.G.; Koo, I. Joint Power Allocation and Power Splitting for MISO-RSMA Cognitive Radio Systems With SWIPT and Information Decoder Users. *IEEE Syst. J.* **2021**, *15*, 289–5300. [CrossRef]
21. Camana, M.R.; Garcia, C.E.; Koo, I. Rate-Splitting Multiple Access in a MISO SWIPT System Assisted by an Intelligent Reflecting Surface. *IEEE Trans. Green Commun. Netw.* **2022**, *6*, 2084–2099. [CrossRef]
22. Huang, K.; Zhong, C.; Zhu, G. Some new research trends in wirelessly powered communications. *IEEE Wirel. Commun.* **2016**, *23*, 19–27. [CrossRef]
23. Ma, Y.; Chen, H.; Lin, Z.; Li, Y.; Vucetic, B. Distributed and Optimal Resource Allocation for Power Beacon-Assisted Wireless-Powered Communications. *IEEE Trans. Commun.* **2015**, *63*, 3569–3583. [CrossRef]
24. Xu, W.; Chen, W.; Fan, Y.; Zhang, Z.; Shi, X. Spectrum efficiency maximization for cooperative power beacon-enabled wireless powered communication networks. *China Commun.* **2021**, *18*, 230–251. [CrossRef]
25. Park, J.-H.; Jeon, Y.-S.; Han, S. Energy Beamforming for Wireless Power Transfer in MISO Heterogeneous Network With Power Beacon. *IEEE Commun. Lett.* **2017**, *21*, 1163–1166. [CrossRef]
26. Camana, M.R.; Garcia, C.E.; Koo, I. Transmit Beamforming for a MISO SWIPT System with a Power Beacon. In Proceedings of the 2020 International Conference on Information and Communication Technology Convergence (ICTC), Jeju, Republic of Korea, 21–23 October 2020.
27. Camana, M.R.; Garcia, C.E.; Koo, I. Power Minimization in a MISO NOMA SWIPT System Assisted by a Power Beacon. In Proceedings of the 2022 Winter Comprehensive Conference of the Korea Information and Communications Society (KICS), Pyeongchang, Republic of Korea, 9–11 February 2022.
28. Vu, T.-H.; Nguyen, T.-V.; Kim, S. Cooperative NOMA-Enabled SWIPT IoT Networks with Imperfect SIC: Performance Analysis and Deep Learning Evaluation. *IEEE Internet Things J.* **2022**, *9*, 2253–2266. [CrossRef]
29. Camana, M.R.; Garcia, C.E.; Koo, I. Rate Splitting Multiple Access for a MISO SWIPT System Aided by a Power Beacon. In Proceedings of the 2022 Thirteenth International Conference on Ubiquitous and Future Networks (ICUFN), Barcelona, Spain, 5–8 July 2022 .
30. Zhang, W.; Zhang, Z.; Chao, H.-C.; Guizani, M. Toward intelligent network optimization in wireless networking: An auto-learning framework. *IEEE Wirel. Commun.* **2019**, *26*, 76–82. [CrossRef]
31. Liang, L.; Ye, H.; Yu, G.; Li, G.Y. Deep-learning-based wireless resource allocation with application to vehicular networks. *Proc. IEEE* **2020**, *108*, 341–356. [CrossRef]
32. Sun, H.; Chen, X.; Shi, Q.; Hong, M.; Fu, X.; Sidiropoulos, N.D. Learning to optimize: Training deep neural networks for interference management. *IEEE Trans. Signal Process.* **2018**, *66*, 5438–5453. [CrossRef]
33. Xia, W.; Zheng, G.; Zhu, Y.; Zhang, J.; Wang, J.; Petropulu, A.P. A deep learning framework for optimization of MISO downlink beamforming. *IEEE Trans. Commun.* **2020**, *68*, 1866–1880. [CrossRef]
34. Kim, J.; Lee, H.; Hong, S.-E.; Park, S.-H. Deep Learning Methods for Universal MISO Beamforming. *IEEE Wirel. Commun. Lett.* **2020**, *9*, 1894–1898. [CrossRef]
35. Camana, M.R.; Garcia, C.E.; Koo, I. Deep Learning-Assisted Power Minimization in Underlay MISO-SWIPT Systems Based on Rate-Splitting Multiple Access. *IEEE Access* **2022**, *10*, 62137–62156. [CrossRef]
36. Bishop, C.M. *Pattern Recognition and Machine Learning*, 1st ed.; Springer: Cambridge, UK, 2006.
37. Wang, Q. Kernel principal component analysis and its applications in face recognition and active shape models. *arXiv* **2014**, arXiv:1207.3538v3.
38. Aggarwal, C.C. *Neural Networks and Deep Learning: A Textbook*, 1st ed.; Springer: Cham, Switzerland, 2018.
39. Grant, M.C.; Boyd, S.P. CVX: Matlab Software for Disciplined Convex Programming, Version 2.2. 2020. Available online: http://cvxr.com/cvx/ (accessed on 1 November 2023).
40. Karipidis, E.; Sidiropoulos, N.D.; Luo, Z.-Q. Quality of service and max–min fair transmit beamforming to multiple cochannel multicast groups. *IEEE Trans. Signal Process.* **2008**, *56*, 1268–1279. [CrossRef]
41. Phan, A.H.; Tuan, H.D.; Kha, H.H.; Ngo, D.T. Nonsmooth Optimization for Efficient Beamforming in Cognitive Radio Multicast Transmission. *IEEE Trans. Signal Process.* **2012**, *60*, 2941–2951. [CrossRef]