Energy Efficiency Optimization for Mutual-Coupling-Aware Wireless Communication System based on RIS-enhanced SWIPT

Ruoyan Ma, Jie Tang, *Senior Member, IEEE,* Xiuyin Zhang, *Fellow, IEEE,* Kai-Kit Wong, *Fellow, IEEE,* Jonathon A. Chambers, *Fellow, IEEE*

***Abstract*—The widespread deployment of the Internet of Things (IoT) is promoting interest in simultaneous wireless information and power transfer (SWIPT), the performance of which can be further improved by employing a reconfigurable intelligent surface (RIS). In this paper, we propose a novel RIS-enhanced SWIPT system built on an electromagnetic-compliant framework. The mutual-coupling effects in the whole system are presented explicitly. Moreover, the reconfigurability of RIS is no longer expressed by the reflection-coefficient matrix but by the impedances of the tunable circuit. For comparison, both the no-coupling and the coupling-awareness cases are discussed. In particular, the energy efficiency (EE) is maximized by cooperatively optimizing the impedance parameters of the RIS elements as well as the active beamforming vectors at the base station (BS). For the coupling-awareness case, the considered problem is split into several sub-problems and solved alternatively due to its nonconvexity. Firstly, it is transformed into a more solvable form by applying the Neuman series approximation, which can be resolved iteratively. Then an alternative optimization (AO) framework and semi-definite relaxation (SDR), successive convex approximation (SCA), and Dinkelbach's algorithm are applied to solve each sub-problem decomposed from it. Owning to the similarity between the two cases, the no-coupling one can be viewed as a reduced form of the coupling case and thus solved through a similar approach. Numerical results reveal the influence of mutual-coupling effects on the EE, especially in the RIS with closely spaced elements. In addition, physical beam designs in different conditions are presented to demonstrate how the RIS assists SWIPT through various reflecting states.**

***Index Terms*—Reconfigurable intelligent surface, simultaneous wireless information and power transfer, fractional programming, mutual coupling, energy efficiency.**

I. Introduction

THE Internet of Things (IoT) has ushered in a new era of development with the further application of fifth generation (5G) wireless communication technology. However, the performance limit of the 5G system cannot support the unprecedented expansion and more diversified information transmission demands required, such as by high-definition images or via multi-dimensional sensing information transfer. These demands are expected to be realized through the sixth generation (6G) technology, due to its outstanding performance, e.g., extremely high throughput, ultra-low latency and intelligent autonomous networks [1]. All these performance breakthroughs rely on several prospective enabling technologies, such as using a reconfigurable intelligent surface (RIS), terahertz communication and integrated sensing and communication (ISAC) [1]–[3]. Among them, since RIS has the controllability of the channel characteristics between the transmitting and the receiving ends, it is regarded as an immense promising technology expected to be employed in 6G systems.

The introduction of a RIS implies that the communication system will no longer passively adapt to the wireless channel, but actively adjust itself according to the needs. In particular, a RIS is capable of constructing the virtual line-of-sight (LoS) link where the LoS link is obstructed, which is more likely to occur in the high frequency band used by 6G [4]. In addition, the application of the RIS also includes the security reconstruction of the physical layer, assisting device-to-device (D2D) communications, suppression of the inter-cell interference of edge users and compensation for signal attenuation [4]–[6]. RIS has not only prominent capabilities and enriched application scenarios but also relatively compact structure and low energy consumption; as such it is anticipated to play a prominent role in the construction of 6G networks and future IoT scenarios [1].

Due to the smaller cells and the more accessible devices, energy consumption becomes an unavoidable problem in 6G. Energy efficiency (EE) therefore has a high priority in performance metrics of the 6G networks [2]. Indeed, for the IoT application scenario of 6G, since the terminal nodes will be uncountable and widely distributed, the difficulties of power supply and information exchange are more prominent. Simultaneous wireless information and power transfer (SWIPT), the key green-communication technology, makes it possible for node devices to receive power and exchange information at the same time [7]. Specifically, there are two categories of SWIPT technology: the near-field SWIPT and the far-field SWIPT [8]. The far-field has attracted much attention from academia and industry because of its flexibility and compatibility at the system level. There is a representative deployment scenario and a novel joint time allocation and power splitting (JTAPS)

Manuscript received October 14, 2022; revised December 21, 2022; This work was supported by the National Natural Science Foundation of China under Grant No. 62222105. *(Corresponding author: Jie Tang.)*

Ruoyan Ma, Jie Tang, Xiuyin Zhang are with the School of Electronic and Information Engineering, South China University of Technology, Guangzhou 510641, China (e-mail: eeruoyan_ma@mail.scut.edu.cn; eejtang@scut.edu.cn; eexyz@scut.edu.cn).

Kai-Kit Wong is with the Department of Electronic and Electrical Engineering, University College London, WClE 6BT London, U.K (e-mail: kai-kit.wong@ucl.ac.uk).

Jonathon A. Chambers is with the School of Engineering, University of Leicester, Leicester LE1 7RH, U.K (e-mail:jonathon.chambers@leicester.ac.uk).

scheme proposed in [9]. With this idea, SWIPT can help the relay node to achieve not only the information but also the power from the source and then transfer information through harvested energy flexibly. Nevertheless, compared to the near-field SWIPT, far-field SWIPT may be frequently impacted by ubiquitous obstructions and undergo tremendous path loss. As a technology that has the ability to control channels and further compensate for these losses, RIS can bring outstanding performance gain to the SWIPT system [10]. Next, we will introduce some works on this topic.

A. Related Works

1) RIS modeling: A RIS works mainly in reflection mode rather than propagation mode, and its reflection performance is influenced by various internal and external factors such as element types, array configurations, element spacings, and incident wave features [6], [11], [12]. Therefore, it is crucial to present these characteristics as accurately and comprehensively as possible during the modeling. The most commonly used model in network optimizations is the independent-reflection matrix model, where the reflection performance of the RIS is expressed by the reflection coefficient. Generally, this matrix is a diagonal matrix, and each reflection coefficient is independent. Although this model is simple enough to analyze, it doesn't describe the key physical characteristics of the RIS [6]. In addition, the physics-based model is widely studied in the field of electromagnetics (EMs), in which the radar anomalous reflection theory is introduced to construct the model. Indeed, this type of model can reveal more information about the physical characteristics during the RIS reflection than the previous one [13]. Moreover, the impedance model is based on the equivalent circuit analysis that describes the physical properties of RIS elements and tunable circuits. Thanks to it, the relationship between the amplitude and the phase shift of the reflection coefficient is presented clearly [14]. An end-to-end EM system model is proposed to analyze the influences of physical features on the system by combining the ideas of the physical model and the impedance model. In particular, it describes the entire RIS system in the form of the circuit, while the transmissions among various ends are based on mutual impedances [15], [16].

2) Reconfigurability of RIS: A RIS can adjust its reflection characteristics to fit the various channel circumstances, which sets it apart from conventional scatterers. As for engineering realizations, there are various ways to adjust the reflection parameters of the RIS elements, which mainly fall into discrete and continuous methods. The number of quantization bits is a worthwhile trade-off with regard to the discrete RIS. Concerning the cost of hardware design, although a higher number of quantization bits can improve performance gain, it is harder to deploy.1-bit discrete RIS is realized by PIN diodes working on the on-off states in [17] and the varactor diodes operating in two states with 180° phase difference in [12]. Moreover, the design based on the varactors is also able to achieve continuous tunability through controlling input voltages [18]. Regarding the theoretical performance analyses, the ideas of tunability design are dependent on the choices of models. As for the independent-reflection matrix model, there are several commonly adopted assumptions that the reflection phases of the elements are continuously adjustable from 0° to 360°, the reflection magnitudes are set to one and these two parameters about the reflection coefficient are independent. All these assumptions make the optimization problems resolved more easily but detached from physical reality [6]. A more practical optimization problem based on the RIS impedance model, which perspicaciously describes the dependence between phases and magnitudes of reflection coefficients, is shown to be solved by the penalty-based algorithm [14]. Furthermore, the imaginary parts of the tunable circuits connected to the RIS elements are taken as the optimization variables in the end-to-end EM system model [16], [19]. Instead of being idealistic, the hardware limits should be concerned integrated into these physical RIS models.

3) Optimization for the RIS-assisted system: A RIS has been shown to further enhance the performance of SWIPT-based systems [6] through its controllability of the channel. In the literature on RIS-assisted SWIPT, the hardware models, problem formulations, and scenario settings are distinct, which makes the optimization strategies different from each other. In addition to the RIS models having been analyzed, the models of the energy harvesting devices, namely the rectifying circuits, are also worth discussing. For expressing the actual features of the rectifying circuit, the non-linear model is more suitable than the linear model. Specifically, the non-linear model based on the sigmoidal function is used to describe the nonlinearity of the rectifiers [20]. Moreover, a rectenna model is introduced to analyze the effects brought by the waveform and the input power of the signal on the performances of the RIS-aided-SWIPT. Furthermore, the waveform is jointly optimized with active and passive beamforming variables [21]. The various objective functions, e.g., maximizing the weighted sum rate (WSR) [22], minimizing the total transmit power [23], maximizing the minimum power received [24] and maximizing the EE [25], are set according to the diverse performance priorities. Unlike the scenarios where information users and energy users are separated [22]–[24], [26], co-located information and power receivers based on power splitting are presented in [20], [21], [27]. Differing from the system with the single RIS, the performance analysis, adopting multi RISs to support hot-spot areas with diversified service requirements, is shown in [23].

B. Motivation and Contribution

Given the above inspirations and the design purposes for analyzing the importance of the hardware characteristics, a RIS-aided-SWIPT network based on the EM is constructed. Then we propose a system problem orienting to the key hardware parameters and aim to maximize the EE constrained to the QoS requirements of the various users. To address this problem, an effective optimization scheme is proposed. The impacts of the actual hardware characteristics on the system performance are presented, which is expected to guide the design and the deployment of the RIS in diversified scenarios. The main contributions of this paper are presented as follows:

- We build the entire RIS-aided-SWIPT MISO transmission model based on the end-to-end EM transfer model and the non-linear rectifier model. By using the EM model, the hardware impacts on EE, which includes the radiation features of the transmitting antennas, the reflection characteristics of the RIS, and especially the mutual-coupling effects influenced by the array configurations are presented.
- With the adopted model, the circuit impedances are utilized as the optimization variables to configure the RIS, which is more practical than the ideal assumptions. We propose two optimization schemes for the no-coupling as well as the coupling-awareness cases. Considering simplifying the original problems, the transformations and approximations are adopted for both cases. Concretely, we use the Neuman series approximation to introduce an iterative form of the transfer model for the coupling-awareness case. In each iteration, the non-convex problems can be approximately solved by the proposed optimization framework. Moreover, the no-coupling case can also be resolved by using a similar scheme.
- The optimization framework is established on the basis of alternative optimization (AO) strategy. With its help, we can decompose the problem into two tractable sub-problems. Specifically, we propose an approach mainly based on semi-definite relaxation (SDR), Dinkelbackh's algorithm, and successive convex approximation (SCA) for solving each sub-problems. Although it is effective for resolving the no-coupling case completely, the coupling-awareness case needs to be invoked multiple times until convergence.
- We employ both the numerical and the full-wave simulations for the performance analysis. As for the numerical simulation, we analyze the influences of QoS requirements and resource budgets on the EE. Moreover, the effects of RIS topologies and the mutual coupling are also considered. Additionally, we reveal the effects of the proposed algorithms on controlling the physical beams in distinct scenarios through full-wave simulations. The results further validate the accuracy of the adopted model and the effectiveness of the algorithms.

C. Organization

The rest of this paper is organized as follows. In Section II, we present the system model based on the EM analysis and the problem formulation. In Section III and Section IV, effective algorithms are proposed to tackle two non-convex problems, the no-coupling and the coupling-awareness problems, respectively. Numerical analyses and full-wave simulation results are shown in Section IV. Section V provides the conclusions. The notations adopted in this paper are listed in Table I.

II. System Model and Problem Formulation

A. Signal Model

In this paper, we consider a MISO downlink RIS-assisted SWIPT system with separate information decoder user (IDU) and energy harvester user (EHU). We assume the base station (BS) is equipped with $N_t > 1$ antennas, while each IDU and EHU adopts single antenna. Moreover, these two kinds of users, IDUs and EHUs, belong to the set $\mathcal{M}_I = \{1, \ldots, M_I\}$ and $\mathcal{M}_E = \{1, \ldots, M_E\}$, respectively. Generally, since the power management devices require more energy to operate, the service scopes of the EHUs are smaller than those of IDUs as presented in Fig. 1. This results in that IDUs, which are located further away from the BS, are likely to be blocked and the LoS links of them may not exist. To further improve the harvested power and signal-to-interference plus noise ratio (SINR) at the user ends, a nearly-passive RIS, which consists of N_i reflecting elements, is introduced into the system. In our scenario, it is assumed that the RIS can be controlled by the continuously adjustable varactors. Particularly, both the information and the energy beamforming are considered at the BS and therefore the transmitter signal is given by

$$\mathbf{x} = \sum_{j=1}^{M_I} \mathbf{w}_j s_j^I + \sum_{l=1}^{M_E} \mathbf{v}_l s_l^E \in \mathbb{C}^{N_T \times 1}, \tag{1}$$

where $\mathbf{w}_j \in \mathbb{C}^{N_T \times 1}$ and $\mathbf{v}_l \in \mathbb{C}^{N_T \times 1}$ denote the transmitting beamforming vectors for the j^{th} IDU and the l^{th} EHU, while s_j^I and s_l^E denote the information and energy signals. Particularly, the information signals $s_j^I, \forall j$ are assumed to

TABLE I: NOTATIONS

Symbol	Definition
$\lvert x\rvert$	The norm of x
$\mathrm{Re}(x)$ and $\mathrm{Im}(x)$	The real and imaginary parts of x
$\mathrm{zeros}(N, N)$	The $N \times N$ all zero matrix
$\mathbf{I}_{N\times N}$	The $N \times N$ identity matrix
$\lVert\mathbf{x}\rVert_2$	The Euclidean norm of $\mathbf{x}$
$\lVert\mathbf{X}\rVert$	The spectral norm of $\mathbf{X}$
$\mathbf{X}^H$	The Hermitian conjugate transpose of $\mathbf{X}$
$\mathbf{X}^T$	The transpose of $\mathbf{X}$
$\mathrm{Tr}(\mathbf{X})$	The trace of $\mathbf{X}$
$\mathrm{Rank}(\mathbf{X})$	The rank of $\mathbf{X}$
$\mathbf{A} \succeq 0$	$\mathbf{X}$ is a positive semidefinite matrix
$\mathbb{R}^{n\times m}, \mathbb{C}^{n\times m}$	The set of $n \times m$ real matrices and complex matrices
$\mathbb{E}(x)$	The statistical expectation of x

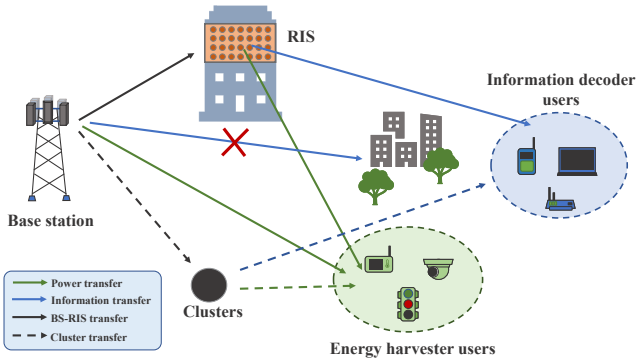


Fig. 1: The RIS-enhanced-SWIPT system.

be independent and identically distributed (i.i.d.) Gaussian random variables, in detail, $s_j^I \sim \mathcal{CN}(0,1), \forall j$. As for the energy signals, they do not transfer any information, so the arbitrary distributions with $\mathbb{E}\left\{\left|s_l^E\right|^2\right\} = 1, \forall l$ are assumed to describe their statistical properties without loss of generality. It is noteworthy that although paper [10] shows that energy signals are not required from the view of optimizing algorithms, we still hope to see their effects on the generations of the actual physical beams in distinct deployment scenarios from the EM perspective.

B. End-to-End EM Transfer Model

We assume that the scenario has two types of multipath propagation scatterers, the reconfigurable scatterer (i.e., the RIS) and the arbitrary scatterer, which coincides with the configuration of [16]. Thus the total transfer links can be divided into the direct link (BS-user), the RIS-assisted link (BS-RIS-user) and the scatterer link (BS-scatterer-user). Among them, the BS-user link and the BS-RIS-user link are based on deterministic physical entities (e.g., antennas, reflecting elements, and semiconductor elements). To further analyze the impacts of their key properties on the system performance, we adopt the end-to-end EM physical model described in the paper [15] to construct these two links and utilize the clustered channel model in [28] to set random scatterers in the system, hence the total transfer model can be depicted by

$$\mathbf{h}(\boldsymbol{\Omega}) = \mathbf{h}_{\mathrm{EM}}(\boldsymbol{\Omega}) + \mathbf{h}_{\mathrm{Cluster}} \in \mathbb{C}^{1\times N_t}, \tag{2}$$

where $\mathbf{h}_{\mathrm{EM}}$ is the EM physical transfer model and $\mathbf{h}_{\mathrm{Cluster}}$ is the clustered model.

As for $\mathbf{h}_{\mathrm{EM}}$, it is controlled by the tunable component matrix $\boldsymbol{\Omega} \in \mathbb{C}^{N_i\times N_i}$. Further, the form of the matrix is $\boldsymbol{\Omega} = \mathrm{diag}\left(\Omega_1, \cdots, \Omega_{N_i}\right)$ and its entries are the impedances of the back-end circuits of the RIS elements. The diodes or the varactors can be used to configure $\boldsymbol{\Omega}$. In this paper, the RIS adopts the varactors to achieve continuously adjustable reactances. In detail, $\mathbf{h}_{\mathrm{EM}}$ in our scenario can be given by (3), where $\mathbf{Z}_{\mathrm{G}} \in \mathbb{C}^{N_t\times N_t}$ denotes the source impedance matrix, a diagonal matrix, and $\mathrm{Z_L}$ is the load impedance. Generally, these two are determined on a case-by-case basis. Moreover, the transfer matrices are described below:

$$\mathbf{S}_{\mathrm{TR}}(\boldsymbol{\Omega}) = \boldsymbol{\mathcal{Z}}_{\mathrm{TR}} - \boldsymbol{\mathcal{Z}}_{\mathrm{TI}}\left(\boldsymbol{\mathcal{Z}}_{\mathrm{II}} + \boldsymbol{\Omega}\right)^{-1}\boldsymbol{\mathcal{Z}}_{\mathrm{IR}} \in \mathbb{C}^{N_t\times 1}, \tag{4}$$

$$\mathbf{S}_{\mathrm{TT}}(\boldsymbol{\Omega}) = \boldsymbol{\mathcal{Z}}_{\mathrm{TT}} - \boldsymbol{\mathcal{Z}}_{\mathrm{TI}}\left(\boldsymbol{\mathcal{Z}}_{\mathrm{II}} + \boldsymbol{\Omega}\right)^{-1}\boldsymbol{\mathcal{Z}}_{\mathrm{IT}} \in \mathbb{C}^{N_t\times N_t}, \tag{5}$$

$$\mathbf{S}_{\mathrm{RT}}(\boldsymbol{\Omega}) = \boldsymbol{\mathcal{Z}}_{\mathrm{RT}} - \boldsymbol{\mathcal{Z}}_{\mathrm{RI}}\left(\boldsymbol{\mathcal{Z}}_{\mathrm{II}} + \boldsymbol{\Omega}\right)^{-1}\boldsymbol{\mathcal{Z}}_{\mathrm{IT}} \in \mathbb{C}^{1\times N_t}, \tag{6}$$

$$\mathrm{S}_{\mathrm{RR}}(\boldsymbol{\Omega}) = \mathcal{Z}_{\mathrm{RR}} - \boldsymbol{\mathcal{Z}}_{\mathrm{RI}}\left(\boldsymbol{\mathcal{Z}}_{\mathrm{II}} + \boldsymbol{\Omega}\right)^{-1}\boldsymbol{\mathcal{Z}}_{\mathrm{IR}}, \tag{7}$$

where T, R and I signify the transmitter, the receiver and the RIS, respectively. $\boldsymbol{\mathcal{Z}}_{\mathrm{AB}} \in \mathbb{C}^{N_a\times N_b}$, in which $\mathrm{A}, \mathrm{B} \in \{\mathrm{T}, \mathrm{R}\}$ and the element numbers of them are N_a and N_b, represents the mutual-impedance matrix between the transmitter and the receiver, when A and B are diverse. On the contrary, $\boldsymbol{\mathcal{Z}}_{\mathrm{AB}}$ denotes the self-impedance matrix in the case that they are the same. Moreover, the diagonal and the off-diagonal elements of the self-impedance matrix are the self impedances and the mutual impedances of the elements in the array. Similarly, $\boldsymbol{\mathcal{Z}}_{\mathrm{AI}} \in \mathbb{C}^{N_A\times N_i}$ and $\boldsymbol{\mathcal{Z}}_{\mathrm{IB}} \in \mathbb{C}^{N_i\times N_B}$ refer to the mutual-impedance matrices between these ends and the RIS. $\boldsymbol{\mathcal{Z}}_{\mathrm{II}} \in \mathbb{C}^{N_i\times N_i}$ is the RIS self-impedance matrix. According to [15], when it is presupposed that: all radiators, including antennas and reflectors, are extremely thin dipoles and placed parallel to the z-axis, the entries of these impedance matrices are the mutual or the self impedances of the dipoles in various ends (i.e. the transmitter, the receiver and the RIS). In the above, all impedances can be calculated by (8), where $\eta_0 = \sqrt{\frac{\mu_0}{\epsilon_0}}$ and $k_0 = \frac{2\pi}{\lambda}$ are the characteristic impedance and the wavenumber, and μ_0, ϵ_0 and λ are the magnetic permeability, the electric permittivity and the wavelength, respectively. According to various impedances, a and b may be the same or the different physical entries. When considering the mutual impedance, $d_{ab}\left(z_a', z_b'\right) = \sqrt{\left(x_a - x_b\right)^2 + \left(y_a - y_b\right)^2 + \left(z_a' - z_b'\right)^2}$. In contrast, as for the self impedance, $d_{ab}\left(z_a', z_b'\right) = \sqrt{r^2 + \left(z_a' - z_b'\right)^2}$. In addition, (x_a, y_a, z_a), (x_b, y_b, z_b), l_a, l_b, r are the center positions, the lengths and the radii of the dipoles.

We make an assumption that all ends are located in the far-field regions of each other, which is a practical deployment scenario and consistent with the setup in [16], hence $\boldsymbol{\mathcal{Z}}_{\mathrm{TT}}$ in (5) and $\mathcal{Z}_{\mathrm{RR}}$ in (7) are more dominant compared to other items. Based on this, $\mathbf{S}_{\mathrm{TT}}$ and S_{RR} can be approximated as

$$\mathbf{S}_{\mathrm{TT}} \approx \boldsymbol{\mathcal{Z}}_{\mathrm{TT}}, \tag{9}$$

$$\mathrm{S}_{\mathrm{RR}} \approx \mathcal{Z}_{\mathrm{RR}}. \tag{10}$$

Similarly, we can let $\mathrm{S}_{\mathrm{RR}} - \mathbf{S}_{\mathrm{RT}}\left(\mathbf{Z}_{\mathrm{G}} + \mathbf{S}_{\mathrm{TT}}\right)^{-1}\mathbf{S}_{\mathrm{TR}} \approx \mathcal{Z}_{\mathrm{RR}}$ in (3) under the far-field condition. Then plugging

$$\mathbf{h}_{\mathrm{EM}}(\boldsymbol{\Omega}) = \frac{\mathrm{Z_L}\mathbf{S}_{\mathrm{RT}}(\boldsymbol{\Omega})\left(\mathbf{Z}_{\mathrm{G}} + \mathbf{S}_{\mathrm{TT}}(\boldsymbol{\Omega})\right)^{-1}}{\mathrm{Z_L} + \mathrm{S}_{\mathrm{RR}}(\boldsymbol{\Omega}) - \mathbf{S}_{\mathrm{RT}}(\boldsymbol{\Omega})\left(\mathbf{Z}_{\mathrm{G}} + \mathbf{S}_{\mathrm{TT}}(\boldsymbol{\Omega})\right)^{-1}\mathbf{S}_{\mathrm{TR}}(\boldsymbol{\Omega})} \in \mathbb{C}^{1\times N_t}, \tag{3}$$

$$\begin{aligned} Z_{ab} = \int_{z_a-l_a/2}^{z_a+l_a/2}\int_{z_b-l_b/2}^{z_b+l_b/2} \frac{j\eta_0}{(4\pi k_0)}\left(\frac{\left(z_a'-z_b'\right)^2}{d_{ab}^2\left(z_a', z_b'\right)}\left(\frac{3}{d_{ab}^2\left(z_a', z_b'\right)} + \frac{3jk_0}{d_{ab}\left(z_a', z_b'\right)} - k_0^2\right) - \frac{\left(jk_0 + \frac{1}{d_{ab}\left(z_a', z_b'\right)}\right)}{d_{ab}\left(z_a', z_b'\right)} + k_0^2\right) \\ - \frac{\exp\left(-jk_0 d_{ab}\left(z_a', z_b'\right)\right)\sin\left(k_0\left(\frac{l_b}{2} - \left|z_b' - z_b\right|\right)\right)\sin\left(k_0\left(\frac{l_a}{2} - \left|z_a' - z_a\right|\right)\right)}{d_{ab}\left(z_a', z_b'\right)\sin\left(k_0\frac{l_b}{2}\right)\sin\left(k_0\frac{l_a}{2}\right)} dz_b' dz_a', \end{aligned} \tag{8}$$

approximated (6), (9), (10) into (3), and $\mathbf{h}_{\mathrm{EM}}$ in (3) can be substituted as

$$\mathbf{h}_{\mathrm{EM}}(\boldsymbol{\Omega}) \approx \frac{\mathrm{Z_L}}{\mathrm{Z_L}+Z_{\mathrm{RR}}}\left(\boldsymbol{\mathcal{Z}}_{\mathrm{RT}}-\boldsymbol{\mathcal{Z}}_{\mathrm{RI}}\left(\boldsymbol{\mathcal{Z}}_{\mathrm{II}}+\boldsymbol{\Omega}\right)^{-1}\boldsymbol{\mathcal{Z}}_{\mathrm{IT}}\right) \times\left(\mathbf{Z}_{\mathrm{G}}+\boldsymbol{\mathcal{Z}}_{\mathrm{TT}}\right)^{-1}. \tag{11}$$

To further simplify the form of (11), abridged notations can be introduced as follows.

$$\frac{\mathrm{Z_L}}{\mathrm{Z_L}+Z_{\mathrm{RR}}}\boldsymbol{\mathcal{Z}}_{\mathrm{RT}}\left(\mathbf{Z}_{\mathrm{G}}+\boldsymbol{\mathcal{Z}}_{\mathrm{TT}}\right)^{-1}=\mathbf{e}\in\mathbb{C}^{1\times N_t}, \tag{12}$$

$$\frac{\mathrm{Z_L}}{\mathrm{Z_L}+Z_{\mathrm{RR}}}\boldsymbol{\mathcal{Z}}_{\mathrm{RI}}=\mathbf{r}\in\mathbb{C}^{1\times N_i}, \tag{13}$$

$$\boldsymbol{\mathcal{Z}}_{\mathrm{IT}}\left(\mathbf{Z}_{\mathrm{G}}+\boldsymbol{\mathcal{Z}}_{\mathrm{TT}}\right)^{-1}=\mathbf{L}\in\mathbb{C}^{N_i\times N_t}, \tag{14}$$

Based on (12)-(14), $\mathbf{h}_{\mathrm{EM}}$ in (11) can be formulated as

$$\mathbf{h}_{\mathrm{EM}}(\boldsymbol{\Omega})\approx\mathbf{e}-\mathbf{r}\left(\boldsymbol{\mathcal{Z}}_{\mathrm{II}}+\boldsymbol{\Omega}\right)^{-1}\mathbf{L}. \tag{15}$$

For clustered channels in (2), in the light of the description in [28], the summation of BS-scatterer-user multipath links can be given by

$$\mathbf{h}_{\mathrm{Cluster}}=\gamma\sum_{k=1}^{N_{\mathrm{cl}}}\sum_{p=1}^{N_{\mathrm{ray},k}}\alpha_{k,p}\sqrt{PL\left(r_{k,p}\right)}\mathbf{a}_t\left(\phi_{k,p}^t\right), \tag{16}$$

where N_{cl} and N_{ray} are the numbers of scatterers and their rays. $\gamma=\sqrt{N_T/\sum_{k=1}^{N_{\mathrm{cl}}}N_{\mathrm{ray},k}}$ denotes the normalization factor and $\alpha_{k,p}$ represents the complex gain of the p^{th} ray in the k^{th} scatterer. $PL\left(r_{k,p}\right)$ is the path loss such that $PL(r_{k,p})=-\left(\eta+\rho\log_{10}(r_{k,p})\right)/10$, where $r_{k,p}=r_k+\sqrt{\left(h_T-h_R+r_k\sin\theta_{k,p}^t\right)^2+\left(d-r_k\cos\theta_{k,p}^t\cos\phi_{k,p}^t\right)^2}$. $\theta_{k,p}$, and $\phi_{k,p}$ are the azimuth and elevation angles of departure. The heights of the transmitter and the receiver are represented by h_T and h_R, whereas r_k and d are the distances from the transmitter to the clusters and the receiver, respectively. Finally, the array response vector $\mathbf{a}_t\in\mathbb{C}^{1\times N_t}$ of a uniform linear array (ULA) on the y-axis can be expressed as $\mathbf{a}_t(\phi)=\frac{1}{\sqrt{N_t}}\left[1,e^{jk_0d_0\sin(\phi)},\ldots,e^{j(N_t-1)k_0d_0\sin(\phi)}\right]$, in which d_0 is the inter-element spacing of the array.

C. Probelm Formulation

Based on the above transfer model, the received signal of the j^{th} IDU is

$$y_j=\mathbf{h}_j(\boldsymbol{\Omega})\mathbf{x}=\underbrace{\mathbf{h}_j(\boldsymbol{\Omega})\mathbf{w}_js_j^I}_{\text{Desired signal}}+\underbrace{\sum_{k\neq j}^{M_I}\mathbf{h}_j(\boldsymbol{\Omega})\mathbf{w}_ks_k^I}_{\text{Inter-IDU interference}} +\underbrace{\sum_{l=1}^{M_E}\mathbf{h}_j(\boldsymbol{\Omega})\mathbf{v}_ls_l^E}_{\text{EHU interference}}+\underbrace{n_I}_{\text{AWGN}}, \tag{17}$$

where $n_I\sim\mathcal{CN}\left(0,\sigma^2\right)$ is the additive white Gaussian noise (AWGN) at the j^{th} IDU with the noise power σ^2. We assume that the IDUs have the ability to cancel the interference from the energy signals and therefore the j^{th} IDU only suffers the interference from other IDUs (e.g., $s_1,\cdots,s_{j-1},s_{j+1},\cdots,s_{M_I}$). Then the SINR of the IDU can be presented by

$$R_j\left(\mathbf{w}_j,\boldsymbol{\Omega}\right)=\log_2\left(1+\frac{\left|\mathbf{h}_j(\boldsymbol{\Omega})\mathbf{w}_j\right|^2}{\sum_{k\neq j}^{M_I}\left|\mathbf{h}_j(\Omega)\mathbf{w}_k\right|^2+\sigma^2}\right). \tag{18}$$

Similarly, the received model of the n^{th} EHU is given by

$$y_n=\mathbf{h}_n(\boldsymbol{\Omega})\mathbf{x}=\sum_{j=1}^{M_I}\mathbf{h}_n(\boldsymbol{\Omega})\mathbf{w}_js_j^I+\sum_{l=1}^{M_E}\mathbf{h}_n(\boldsymbol{\Omega})\mathbf{v}_ls_l^E+n_E, \tag{19}$$

where $n_E\sim\mathcal{CN}\left(0,\sigma^2\right)$ is the AWGN of the n^{th} EHU and can be ignored. Moreover, its received radio frequency (RF) power is expressed as

$$P_{\mathrm{RF},n}\left(\mathbf{w}_j,\mathbf{v}_l,\boldsymbol{\Omega}\right)=\mathbf{h}_n(\boldsymbol{\Omega})\left(\sum_{j=1}^{M_I}\mathbf{w}_j\mathbf{w}_j^H+\sum_{l=1}^{M_E}\mathbf{v}_l\mathbf{v}_l^H\right) \times\mathbf{h}_n(\boldsymbol{\Omega})^H. \tag{20}$$

The RF power in (20) needs to be rectified as direct current (DC) power for serving EHUs. Furthermore, when the input power is excessively high, due to the reverse bias of the diode, the non-linear characteristic of the rectifying circuit should be considered. To characterize this feature, the DC output power of the n^{th} EHU calculated by a non-linear EH model can be presented as (21), where ν and ϖ are determined by the circuit parameters. B is the maximum DC power represented by $B=\frac{V_{br}^2}{4R_L}$, where V_{br} and R_L denote the reverse breakdown voltage and the load resistance. Considering the convenience of the optimization, we can define an inverse function $\mathcal{G}(x)$ of (21) as

$$\mathcal{G}(x)=\varpi-\frac{\ln(\frac{B(1+\exp(\nu\varpi))}{x\exp(\nu\varpi)+B}-1)}{\nu}. \tag{22}$$

Moreover, the total required power of the system is given as follows.

$$P\left(\mathbf{w}_j,\mathbf{v}_l\right)=P_t\left(\mathbf{w}_j,\mathbf{v}_l\right)+P_c+N_iP_I, \tag{23}$$

Here, denoting the transmitting power as $P_t\left(\mathbf{w}_j,\mathbf{v}_l\right)=\sum_{j=1}^{M_I}\left\|\mathbf{w}_j\right\|_2^2+\sum_{l=1}^{M_E}\left\|\mathbf{v}_l\right\|_2^2$. P_c and P_I are the power consumption of the front-end modules (e.g., mixers and filters) at the BS and the control circuits of the RIS elements.

In this paper, we concentrate on both the sum rate of the system and the corresponding dissipated energy, thus the EE is introduced to trade off them. The fractional form of EE can be presented as

$$\mathrm{EE}\left(\mathbf{w}_j,\mathbf{v}_l,\boldsymbol{\Omega}\right)=\frac{\sum_{j=1}^{M_I}R_j\left(\mathbf{w}_j,\boldsymbol{\Omega}\right)}{P\left(\mathbf{w}_j,\mathbf{v}_l\right)}. \tag{24}$$

$$P_{\mathrm{DC},n}\left(\mathbf{w}_j,\mathbf{v}_l,\boldsymbol{\Omega}\right)=\frac{B(1+\exp\left(\nu\varpi\right))}{\exp\left(\nu\varpi\right)\left(1+\exp\left(-\nu\left(P_{\mathrm{RF},n}\left(\mathbf{w}_j,\mathbf{v}_l,\boldsymbol{\Omega}\right)-\varpi\right)\right)\right)}-\frac{B}{\exp\left(\nu\varpi\right)}, \tag{21}$$

For the RIS-enhanced SWIPT network, we set the problem to maximize the EE subject to the maximal transmitting power, the minimal rate of the IDU, the minimal DC output power of the EHU, and the hardware restrictions of the control circuit at the RIS. The above optimization problem has the form of the following $\mathbf{P}_0$.

$$\mathbf{P}_0: \max_{\{\mathbf{w}_j,\mathbf{v}_l,\boldsymbol{\Omega}\}} \mathrm{EE}\left(\mathbf{w}_j,\mathbf{v}_l,\boldsymbol{\Omega}\right) \tag{25a}$$

$$\text{s.t. } R_j\left(\mathbf{w}_j,\boldsymbol{\Omega}\right) \geq R_j^{(\mathrm{D})}, \forall j, \tag{25b}$$

$$P_t\left(\mathbf{w}_j,\mathbf{v}_l\right) \leq P_{\mathrm{Max}}, \tag{25c}$$

$$P_{\mathrm{RF},n}\left(\mathbf{w}_j,\mathbf{v}_l,\boldsymbol{\Omega}\right) \geq \mathcal{G}_n(E_n^{(\mathrm{D})}), \forall n, \tag{25d}$$

$$\mathrm{Re}\left(\boldsymbol{\Omega}_q\right) = R_0 \geq 0, \forall q, \tag{25e}$$

$$\mathrm{Im}\left(\boldsymbol{\Omega}_q\right) \in \mathbb{R}, \forall q. \tag{25f}$$

In (25b), $R_j^{(\mathrm{D})}$ denotes the rate requirement of the j^{th} IDU. As for (25c), P_{Max} is the maximal transmitting power budget. The $E_n^{(\mathrm{D})}$ in (25d) is the minimal DC output power demand of the n^{th} EHU. (25e) and (25f) represent the limitations of the control circuit. R_0 in (25e) denotes the loss of the circuit and $R_0 \geq 0$ is the guarantee that the RIS operates in the passive mode instead of the active mode [15]. In this paper, the reactances of the circuits in (25f) are reconfigurable and we assume that their scopes of tunability are the set of real numbers.

The proposed system EE optimization problem P_0 is quite challenging to resolve owing to the fractional objective function and the non-convex constraints. Particularly, to further show the influences of the hardware features on the system performances, the problem is separated into the no-coupling and the coupling-awareness situations for comparison. In addition, the transformations of the problems and the optimization schemes are introduced to solve them.

III. Solution Approach for No Mutual Coupling

Before starting to solve the problem, the form of (15) can be simplified under the no-coupling assumption for the RIS, since $\boldsymbol{\mathcal{Z}}_{\mathrm{II}}$ only keeps the self impedances of the RIS elements and turns into a diagonal matrix. It should be noted that the assumption is an ideal condition. Then the total transfer model is reformulated as

$$\mathbf{h}(\Lambda) \approx \mathbf{e} - \mathbf{r}\boldsymbol{\Lambda}\mathbf{L} + \mathbf{h}_{\mathrm{Cluster}}, \tag{26}$$

where $\boldsymbol{\Lambda} = \mathrm{diag}\left(\frac{1}{r_1+jx_1},\cdots,\frac{1}{r_{N_i}+jx_{N_i}}\right)$ and $r_q + jx_q$ for $q = 1, 2, \cdots, N_i$ is presented by

$$r_q = \mathrm{Re}\left(\boldsymbol{\mathcal{Z}}_{\mathrm{II}}(q,q)\right) + R_0, \tag{27}$$

$$x_q = \mathrm{Im}\left(\boldsymbol{\mathcal{Z}}_{\mathrm{II}}(q,q) + \boldsymbol{\Omega}(q,q)\right). \tag{28}$$

In our work, $\boldsymbol{\mathcal{Z}}_{\mathrm{II}}$ is calculated based on the dipole antenna and its real part is positive, thus $r_q > 0$. Further, the entries of $\boldsymbol{\Lambda}$ can be replaced as follows

$$\frac{1}{r_q + jx_q} = \frac{r_q - jx_q}{r_q^2 + x_q^2} = \frac{1 + \exp\left(j2\theta_q\right)}{2r_q}. \tag{29}$$

In which, $\tan\theta = -\frac{x_q}{r_q}$. Then plugging (29) into (26), we can obtain the transfer model for the users

$$\mathbf{h}(\Psi) \approx \mathbf{g} - \mathbf{r}\boldsymbol{\Psi}\mathbf{L}, \tag{30}$$

where shorthand notations are $\mathbf{g} = \mathbf{e} - \frac{\mathbf{rL}}{2r_q} + \mathbf{h}_{\mathrm{cluster}}$ and $\boldsymbol{\Psi} = \mathrm{diag}\left(\frac{\exp(j\phi_1)}{2r_1},\cdots,\frac{\exp\left(j\phi_{N_i}\right)}{2r_{N_i}}\right)$. According to $\cos\theta = \frac{r_q}{\sqrt{r_q^2+x_q^2}} > 0$, the angle constraints can be $\theta_q \in (-\pi/2, \pi/2)$ and $\phi_q = 2\theta_q \in (-\pi, \pi)$.

Relying on the above assumption without mutual-coupling effects of the RIS, the system problem $\mathbf{P}_0$ turns into $\mathbf{P}_{\mathrm{NC}-0}$ as follows

$$\mathbf{P}_{\mathrm{NC}-0}: \max_{\{\mathbf{w}_j,\mathbf{v}_l,\Psi\}} \mathrm{EE}\left(\mathbf{w}_j,\mathbf{v}_l,\Psi\right) \tag{31a}$$

$$\text{s.t. } (25c),$$

$$R_j\left(\mathbf{w}_j,\Psi\right) \geq R_j^{(\mathrm{D})}, \forall j, \tag{31b}$$

$$P_{\mathrm{RF},n}\left(\mathbf{w}_j,\mathbf{v}_l,\Psi\right) \geq \mathcal{G}_n(E_n^{(\mathrm{D})}), \forall n, \tag{31c}$$

$$\phi_q \in (-\pi,\pi), \forall q. \tag{31d}$$

Even though the structure of the channel model (30) has been further reduced and is similar to the cascaded channel model based on the independent diffusive scatterer (IDS) assumption in [6], the problem $\mathbf{P}_{\mathrm{NC}-0}$ remains a non-convex problem that is difficult to solve. To effectively address it, we adopt the SDR approach and the alternative optimization strategy. Specifically, for the beamforming and the circuit parameter subproblems, Dinkelbach's algorithm and SCA are introduced to tackle them.

A. Semi-Definite Relaxation

The introduction of the SDR method will further enhance the solvability of the problem and pave the way for the solution of the subproblems. For beamforming vectors $\mathbf{w}_j$ and $\mathbf{v}_l$ at the BS, denoting $\mathbf{W}_j = \mathbf{w}_j\mathbf{w}_j^H$ and $\mathbf{V}_l = \mathbf{v}_l\mathbf{v}_l^H$ while satisfying $\mathbf{W}_j \succeq 0$, $\mathrm{rank}(\mathbf{W}_j) = 1$, $\mathbf{V}_l \succeq 0$ and $\mathrm{rank}(\mathbf{V}_l) = 1$. Concerning $\mathbf{r}\boldsymbol{\Psi}\mathbf{L}$ in (30), it can be replaced by $\mathbf{f}\mathbf{L}_u$, where $\mathbf{f} \triangleq [\frac{\exp(j\phi_1)}{2r_1},\cdots,\frac{\exp\left(j\phi_{N_i}\right)}{2r_{N\,i}}]$ and $\mathbf{L}_u \triangleq -\mathrm{diag}(\mathbf{r})\mathbf{L}$. Moreover, we let $\hat{\mathbf{f}} = [\mathbf{f}, 1]$ and $\hat{\mathbf{L}}_u = [\mathbf{L}_u; \mathbf{g}]$. SDR can also applied in $\hat{\mathbf{f}}$ and we denote $\hat{\mathbf{F}} = \hat{\mathbf{f}}^H\hat{\mathbf{f}}$ subject to $\hat{\mathbf{F}} \succeq 0$ and $\mathrm{rank}(\hat{\mathbf{F}}) = 1$. Then the rate $R_j\left(\mathbf{W}_j,\hat{\mathbf{F}}\right)$ of the j^{th} IDU can be transformed as

$$R_j\left(\mathbf{W}_j,\hat{\mathbf{F}}\right) = \log_2\left(1 + \frac{\mathrm{Tr}\left(\hat{\mathbf{F}}\hat{\mathbf{L}}_{u,j}\mathbf{W}_j\hat{\mathbf{L}}_{u,j}^H\right)}{\mathrm{Tr}\left(\hat{\mathbf{F}}\hat{\mathbf{L}}_{u,j}\sum_{k\neq j}^{M_I}\mathbf{W}_k\hat{\mathbf{L}}_{u,j}^H\right) + \sigma^2}\right). \tag{32}$$

The received RF power $P_{\mathrm{RF},n}\left(\mathbf{W}_j,\mathbf{V}_l,\hat{\mathbf{F}}\right)$ of the n^{th} EHU can be substituted by

$$P_{\mathrm{RF},n}\left(\mathbf{W}_j,\mathbf{V}_l,\hat{\mathbf{F}}\right) = \mathrm{Tr}\left(\hat{\mathbf{F}}\hat{\mathbf{L}}_{u,n}(\sum_{j=1}^{M_I}\mathbf{W}_j + \sum_{l=1}^{M_E}\mathbf{V}_l)\hat{\mathbf{L}}_{u,n}^H\right). \tag{33}$$

The transmitting power $P_t\left(\mathbf{W}_j,\mathbf{V}_l\right)$ is relpaced by

$$P_t\left(\mathbf{W}_j,\mathbf{V}_l\right) = \sum_{j=1}^{M_I}\mathrm{Tr}\left(\mathbf{W}_j\right) + \sum_{l=1}^{M_E}\mathrm{Tr}(\mathbf{V}_l). \tag{34}$$

After SDR with dropping rank-one constraints, the problem $\mathbf{P}_{\mathrm{NC}-0}$ is transformed into $\mathbf{P}_{\mathrm{NC}-1}$

$$\mathbf{P}_{\mathrm{NC}-1}: \max_{\{\mathbf{W}_j,\mathbf{V}_l,\hat{\mathbf{F}}\}} \frac{\sum_{j=1}^{M_I} R_j\left(\mathbf{W}_j,\hat{\mathbf{F}}\right)}{P\left(\mathbf{W}_j,\mathbf{V}_l\right)} \tag{35a}$$

$$\text{s.t.} \ \frac{\mathrm{Tr}\left(\hat{\mathbf{F}}\hat{\mathbf{L}}_{u,j}\mathbf{W}_j\hat{\mathbf{L}}_{u,j}^H\right)}{\mathrm{Tr}\left(\hat{\mathbf{F}}\hat{\mathbf{L}}_{u,j}\sum_{k\neq j}^{M_I}\mathbf{W}_k\hat{\mathbf{L}}_{u,j}^H\right)+\sigma^2} \geq 2^{R_j^{(\mathrm{D})}}-1, \forall j, \tag{35b}$$

$$P_t\left(\mathbf{W}_j,\mathbf{V}_l\right) \leq P_{\mathrm{Max}}, \tag{35c}$$

$$\mathrm{Tr}\left(\hat{\mathbf{F}}\hat{\mathbf{L}}_{u,n}(\sum_{j=1}^{M_I}\mathbf{W}_j+\sum_{l=1}^{M_E}\mathbf{V}_l)\hat{\mathbf{L}}_{u,n}^H\right) \geq \mathcal{G}_n(E_n^{(\mathrm{D})}), \forall n, \tag{35d}$$

$$\hat{\mathbf{F}}_{q,q}=\frac{1}{4r_q^2}, \forall q, \hat{\mathbf{F}}_{q+1,q+1}=1, \tag{35e}$$

$$\mathbf{W}_j \succeq \mathbf{0}, \mathbf{V}_l \succeq \mathbf{0}, \hat{\mathbf{F}} \succeq \mathbf{0}, \forall j,l. \tag{35f}$$

However, the variables remain coupled, hence even if the SDR approach was adopted, the problem $\mathbf{P}_{\mathrm{NC}-1}$ is still non-convex. Nevertheless, the problem can be further transformed into a more solvable form using the AO strategy. More specifically, fix $\hat{\mathbf{F}}$ to optimize $\mathbf{W}_j$ and $\mathbf{V}_l$ firstly, then use the obtained $\mathbf{W}_j^*$ and $\mathbf{V}_l^*$ to obtain $\hat{\mathbf{F}}^*$, iteratively.

B. AO strategy

1) Optimization of $\{\mathbf{W}_j,\mathbf{V}_l\}$*:* As for the beamforming subproblem with fixed $\hat{\mathbf{F}}^*$, problem $\mathbf{P}_{\mathrm{NC}-1}$ can be simplified as follows.

$$\begin{aligned}\mathbf{P}_{\mathrm{NC}-2}: &\max_{\{\mathbf{W}_j,\mathbf{V}_l\}} \mathrm{EE}\left(\mathbf{W}_j,\mathbf{V}_l\right)\\ &\text{s.t. (35b)-(35d)}\\ &\mathbf{W}_j \succeq \mathbf{0}, \mathbf{V}_l \succeq \mathbf{0}, \forall j,l.\end{aligned} \tag{36}$$

First, the objective function of problem $\mathbf{P}_{\mathrm{NC}-2}$ has the form of the fraction, which makes the problem tough to optimize. The idea is converting the fraction into a more concise one. Luckily, the effective fractional programming approach, Dinkelbach's algorithm proposed in [29], can be adopted to transform it into a solvable form.

Lemma 1: Under the condition that ϵ^* is the unique zero, problem $\mathbf{P}_{\mathrm{NC}-2}$ can be equivalently transformed as the auxiliary problem $\mathbf{P}_{\mathrm{NC}-3}$ with the subtraction form of the objective function.

$$\begin{aligned}\mathbf{P}_{\mathrm{NC}-3}: &\max_{\{\mathbf{W}_j,\mathbf{V}_l\}} \sum_{j=1}^{M_I} R_j\left(\mathbf{W}_j\right)-\epsilon P\left(\mathbf{W}_j,\mathbf{V}_l\right)\\ &\text{s.t. (35b)-(35d), (36).}\end{aligned} \tag{37}$$

Proof: The proof was presented in [29], we will not go into detail here.

For tackling the problem $\mathbf{P}_{\mathrm{NC}-3}$, we utilize the steps in Algorithm 1 to iteratively solve the auxiliary problem until the objective function of the problem converges to a threshold value ε. Indeed, according to the algorithm, the crux of the algorithm is addressing the following problem $\mathbf{P}_{\mathrm{NC}-4}$ in each iteration i.

Algorithm 1 Beamforming vectors optimization framework based on Dinkelbach's algorithm

Initialize: $\epsilon(0)=0$, $\mathrm{EE}(0)=0$, ε;

1: Set $i=0$;
2: **Repeat**
3: Obtain $\mathbf{W}_j^*(i),\mathbf{V}_l^*(i)$ through tackling the problem $\mathbf{P}_{\mathrm{NC}-4}$;
4: Set

$$\Xi^*(i)=\sum_{j=1}^{M_I} R_j\left(\mathbf{W}_j^*(i)\right)-\epsilon(i)P\left(\mathbf{W}_j^*(i),\mathbf{V}_l^*(i)\right);$$

5: Update $\epsilon(i+1)=\frac{\sum_{j=1}^{M_I} R_j\left(\mathbf{W}_j^*(i)\right)}{P\left(\mathbf{W}_j^*(i),\mathbf{V}_l^*(i)\right)}$;
6: $i=i+1$;
7: **Until** $\Xi^*(\epsilon(i)) \leq \varepsilon$ or $\mathrm{EE}(i)-\mathrm{EE}(i-1) \leq \varepsilon$
8: Achieve $\mathbf{w}_j^*$ and $\mathbf{v}_l^*$ by the eigenvector of $\mathbf{W}_j^*(i)$ and $\mathbf{v}_l^*(i)$, if the rank-one condition is met. Otherwise, Gaussian randomization is adopted [25];

Output: the optimal $\mathbf{w}_j^*$ and $\mathbf{v}_l^*$.

$$\begin{aligned}\mathbf{P}_{\mathrm{NC}-4}: &\max_{\{\mathbf{W}_j(i),\mathbf{V}_l(i)\}} \sum_{j=1}^{M_I} R_j\left(\mathbf{W}_j(i)\right)\\ &-\epsilon(i)P\left(\mathbf{W}_j(i),\mathbf{V}_l(i)\right)\\ &\text{s.t. (35b)-(35d), (36).}\end{aligned} \tag{38}$$

Then, to further solve $\mathbf{P}_{\mathrm{NC}-4}$, the auxiliary variable τ is introduced to replace the fractional equation of the rate and thus the problem can be reformulated in the equivalent form as follows.

$$\begin{aligned}\mathbf{P}_{\mathrm{NC}-4}: &\max_{\left\{\begin{smallmatrix}\mathbf{W}_j(i),\mathbf{V}_l(i),\\ \tau_j(i)\end{smallmatrix}\right\}} \Xi(\tau_j(i),\mathbf{W}_j(i),\mathbf{V}_l(i)),\\ &\text{s.t.} R_j\left(\mathbf{W}_j(i)\right) \geq \tau_j(i), \forall j\\ &\text{(35b)-(35d), (36).}\end{aligned} \tag{39}$$

where the (39) is an essential constraint, which is utilized to ensure that the rate of IDU is not less than τ. Moreover, the objective equation $\Xi(\tau_j(i),\mathbf{W}_j(i),\mathbf{V}_l(i))$ of $\mathbf{P}_{\mathrm{NC}-4}$ is presented as

$$\Xi(\tau_j(i),\mathbf{W}_j(i),\mathbf{V}_l(i))=\sum_{j=1}^{M_I}\tau_j(i)-\epsilon(i)P\left(\mathbf{W}_j(i),\mathbf{V}_l(i)\right). \tag{40}$$

So far, the fractional parts of the objective are eliminated. Then for handling the introduced constraint (39), we define the other two variables $a_j(i)$ and $b_j(i)$ which satisfy the following equalities.

$$e^{a_j(i)}=\mathrm{Tr}\left(\hat{\mathbf{F}}\hat{\mathbf{L}}_{u,j}\sum_{k\neq j}^{M_I}\mathbf{W}_k(i)\hat{\mathbf{L}}_{u,j}^H\right), \tag{41}$$

$$e^{b_j(i)}=2^{\tau_j(i)}-1. \tag{42}$$

Algorithm 2 SCA algorithm

Initialize: $\bar{a}_j^{(0)}(i)$, $\bar{b}_j^{(0)}(i)$, ε;
1: Set $t=0$;
2: **Repeat**
3: Obtain $\mathbf{W}_j^{*(t)}(i)$ through tackling the convex problem $\mathbf{P}_{\mathrm{NC}-6}$;
4: Update $\bar{a}_j^{(t)}(i)$ and $\bar{b}_j^{(t)}(i)$
5: $t=t+1$;
6: **Until**
7: $\left|\bar{a}_j^{(t)}(i)-\bar{a}_j^{(t-1)}(i)\right| \le \varepsilon$ and $\left|\bar{b}_j^{(t)}(i)-\bar{b}_j^{(t-1)}(i)\right| \le \varepsilon$
Output: the optimal $\mathbf{W}_j^{*(t)}(i)$ and $\mathbf{V}_l^{*(t)}(i)$

With the help of (41) and (42), problem $\mathbf{P}_{\mathrm{NC}-4}$ can be rewriten as problem $\mathbf{P}_{\mathrm{NC}-5}$.

$$\mathbf{P}_{\mathrm{NC}-5}: \max_{\left\{\substack{\mathbf{w}_j(i),\mathbf{v}_l(i),\tau_j(i),\\ a_j(i),b_j(i)}\right\}} \Xi(\tau_j(i),\mathbf{W}_j(i),\mathbf{V}_l(i))$$
$$\text{s.t. } \mathrm{Tr}\left(\hat{\mathbf{F}}\hat{\mathbf{L}}_{u,j}\mathbf{W}_j\hat{\mathbf{L}}_{u,j}^H\right) \ge e^{a_j(i)+b_j(i)}+\sigma^2 e^{b_j(i)}, \forall j, \quad (43a)$$
$$e^{a_j(i)} \ge \mathrm{Tr}\left(\hat{\mathbf{F}}\hat{\mathbf{L}}_{u,j}\sum\nolimits_{k\neq j}^{M_I}\mathbf{W}_k(i)\hat{\mathbf{L}}_{u,j}^H\right), \forall j, \quad (43b)$$
$$e^{b_j(i)} \ge 2^{\tau_j(i)}-1, \forall j, \quad (43c)$$
(35b)-(35d), (36).

The first-order approximation can be adopted to linearize the left hand of (43b) and (43c), due to the following inequalities.

$$e^{a_j(i)} \ge e^{\bar{a}_j(i)}+e^{\bar{a}_j(i)}\left(a_j(i)-\bar{a}_j(i)\right), \quad (44)$$
$$e^{b_j(i)} \ge e^{\bar{b}_j(i)}+e^{\bar{b}_j(i)}\left(b_j(i)-\bar{b}_j(i)\right), \quad (45)$$

where $\bar{a}_j(i)$ and $\bar{b}_j(i)$ are the feasible values of $a_j(i)$ and $b_j(i)$. Moreover, (43b) and (43c) can be replaced as

$$e^{\bar{a}_j(i)}+e^{\bar{a}_j(i)}\left(a_j(i)-\bar{a}_j(i)\right) \ge \mathrm{Tr}\left(\hat{\mathbf{F}}\hat{\mathbf{L}}_{u,j}\sum\nolimits_{k\neq j}^{M_I}\mathbf{W}_k(i)\hat{\mathbf{L}}_{u,j}^H\right), \quad (46)$$

$$e^{\bar{b}_j(i)}+e^{\bar{b}_j(i)}\left(b_j(i)-\bar{b}_j(i)\right) \ge 2^{\tau_j(i)}-1. \quad (47)$$

Then, problem $\mathbf{P}_{\mathrm{NC}-5}$ has the following approximate form based on (46) and (47).

$$\mathbf{P}_{\mathrm{NC}-6}: \max_{\left\{\substack{\mathbf{w}_j(i),\mathbf{v}_l(i),\tau_j(i),\\ a_j(i),b_j(i)}\right\}} \Xi(\tau_j(i),\mathbf{W}_j(i),\mathbf{V}_l(i))$$
s.t. (35b)-(35d), (36), (43a),(46),(47).

Up to this point, problem $\mathbf{P}_{\mathrm{NC}-4}$ is transformed into a convex problem $\mathbf{P}_{\mathrm{NC}-6}$ after a series of substitutions and approximations. The convex programming solver, i.e., CVX, can be utilized to solve it. Nevertheless, the approximate solution can be more accurate through the SCA algorithm as presented in Algorithm 2. In particular, problem $\mathbf{P}_{\mathrm{NC}-6}$ will be solved continuously until the final condition is satisfied and the results of $\bar{a}_j^{(t)}(i)$ and $\bar{b}_j^{(t)}(i)$ in the t^{th} iteration are passed into the next solution process. Specifically, $\bar{a}_j^{(t)}(i)$ and $\bar{b}_j^{(t)}(i)$ can be updated by

$$\bar{a}_j^{(t)}(i)=\ln\left(\mathrm{Tr}\left(\hat{\mathbf{F}}\hat{\mathbf{L}}_{u,j}\sum\nolimits_{k\neq j}^{M_I}\mathbf{W}_k^{*(t)}(i)\hat{\mathbf{L}}_{u,j}^H\right)\right), \quad (48)$$
$$\bar{b}_j^{(t)}(i)=\ln\left(2^{\tau_j^{*(t)}(i)}-1\right), \quad (49)$$

where $\mathbf{W}_k^{*(t)}(i)$ is obtained from the t^{th} solution of the SCA iteration and $\tau_j^{*(t)}(i)$ can be replaced continuously by

$$\tau_j^{*(t)}(i)= \log_2\left(1+\frac{\mathrm{Tr}\left(\hat{\mathbf{F}}\hat{\mathbf{L}}_{u,j}\mathbf{W}_j^{*(t)}(i)\hat{\mathbf{L}}_{u,j}^{II}\right)}{\mathrm{Tr}\left(\sum_{k\neq j}^{M_I}\hat{\mathbf{F}}\hat{\mathbf{L}}_{u,j}\mathbf{W}_k^{*(t)}(i)\hat{\mathbf{L}}_{u,j}^{I}\right)+\sigma^2}\right). \quad (50)$$

2) Optimization of $\hat{\mathbf{F}}$*:* After achieving the optimized $\mathbf{W}_j^*=\mathbf{w}_j^*\mathbf{w}_j^{*H}$ and $\mathbf{V}_l^*=\mathbf{v}_l^*\mathbf{v}_l^{*H}$, problem $\mathbf{P}_{\mathrm{NC}-1}$ can be introduced with some auxiliary variables and reformulated as following

$$\mathbf{P}_{\mathrm{NC}-7}: \max_{\left\{\hat{\mathbf{F}},\varrho_j,p_j,q_j\right\}} \sum_{j=1}^{M_I}\varrho_j$$
$$\text{s.t. } \mathrm{Tr}\left(\hat{\mathbf{F}}\hat{\mathbf{L}}_{u,j}\mathbf{W}_j\hat{\mathbf{L}}_{u,j}^H\right) \ge e^{p_j+q_j}+\sigma^2 e^{q_j}, \forall j, \quad (51a)$$
$$e^{p_j} \ge \mathrm{Tr}\left(\hat{\mathbf{F}}\hat{\mathbf{L}}_{u,j}\sum\nolimits_{k\neq j}^{M_I}\mathbf{W}_k\hat{\mathbf{L}}_{u,j}^H\right), \forall j, \quad (51b)$$
$$e^{q_j} \ge 2^{\varrho_j}-1, \forall j, \quad (51c)$$
$$\hat{\mathbf{F}} \succeq \mathbf{0}, \quad (51d)$$
(35b), (35d), (35e).

Similar to problem $\mathbf{P}_{\mathrm{NC}-5}$, the above problem is a convex problem and the SCA strategy adopted in algorithm 2 can be used to solve it iteratively. For achieved $\hat{\mathbf{F}}^*$, eigenvalue decomposition or Gaussian randomization is introduced to get $\hat{\mathbf{f}}^*$. Through the AO strategy, the beamforming and the circuit parameters subproblems can be resolved iteratively until the termination condition of the total algorithm is reached.

IV. SOLUTION APPROACH FOR MUTUAL-COUPLING AWARENESS

When considering the mutual coupling among the RIS elements, $\mathbf{T}=\boldsymbol{\mathcal{Z}}_{\mathrm{II}}+\boldsymbol{\Omega}$ in (15) cannot be treated as a diagonal matrix and the matrix inverse must be retained, which makes the system problem difficult to be tackled. However, thanks to the Neuman series approximation [30], the inverse of the matrices can be simplified. Before adopting it, we assume the problem can be solved iteratively and the problem in the ℓ^{th} iteration is

$$\mathbf{P}_{\mathrm{CA}-0}^{(\ell)}: \max_{\left\{\mathbf{w}_j^{(\ell)},\mathbf{v}_l^{(\ell)},\boldsymbol{\Omega}^{(\ell)}\right\}} \mathrm{EE}\left(\mathbf{w}_j^{(\ell)},\mathbf{v}_l^{(\ell)},\boldsymbol{\Omega}^{(\ell)}\right)$$
$$\text{s.t. } P_t\left(\mathbf{w}_j^{(\ell)},\mathbf{v}_l^{(\ell)}\right) \le P_{\mathrm{Max}}, \quad (52a)$$
$$R_j\left(\mathbf{w}_j^{(\ell)},\boldsymbol{\Omega}^{(\ell)}\right) \ge R_j^{(\mathrm{D})}, \forall j, \quad (52b)$$
$$P_{\mathrm{RF},n}\left(\mathbf{w}_j^{(\ell)},\mathbf{v}_l^{(\ell)},\boldsymbol{\Omega}^{(\ell)}\right) \ge \mathcal{G}_n(E_n^{(\mathrm{D})}), \forall n, \quad (52c)$$
$$\mathrm{Re}\left(\boldsymbol{\Omega}_q^{(\ell)}\right)=R_0 \ge 0, \forall q, \quad (52d)$$
$$\mathrm{Im}\left(\boldsymbol{\Omega}_q^{(\ell)}\right) \in \mathbb{R}, \forall q. \quad (52e)$$

The optimized tunable component matrix $\mathbf{\Omega}^{*(\ell)}$ is introduced to update $\mathbf{T}$, thus $\mathbf{T}^{(\ell+1)}$ satisfies

$$\mathbf{T}^{(\ell+1)} = \mathbf{T}^{(\ell)} + \mathbf{G}^{(\ell)}, \tag{53}$$

where $\mathbf{G}^{(\ell)} = j\,\mathrm{Im}\left(\mathbf{\Omega}^{*(\ell)}\right)$ and $\mathbf{T}^{(0)} = \boldsymbol{\mathcal{Z}}_{\mathrm{II}} + R_0$ for fixing the resistances and adjusting the reactances of the circuits. Based on the Neuman series approximation, the inverse of $\mathbf{T}^{(\ell+1)}$ in (15) can be approximated by

$$\begin{aligned}
&\left(\mathbf{T}^{(\ell)} + \mathbf{G}^{(\ell)}\right)^{-1} \\
&= \left[\mathbf{I}_{N_i \times N_i} - \left(-{\mathbf{T}^{(\ell)}}^{-1}\mathbf{G}^{(\ell)}\right)\right]^{-1} {\mathbf{T}^{(\ell)}}^{-1} \qquad (54)\\
&= \sum_{n=0}^{+\infty}\left(-{\mathbf{T}^{(\ell)}}^{-1}\mathbf{G}^{(\ell)}\right)^{n} {\mathbf{T}^{(\ell)}}^{-1} \qquad (55)\\
&\approx {\mathbf{T}^{(\ell)}}^{-1} - {\mathbf{T}^{(\ell)}}^{-1}\mathbf{G}^{(\ell)}{\mathbf{T}^{(\ell)}}^{-1}. \qquad (56)
\end{aligned}$$

The transformation from (54) to (55) can be traced to [30]. Furthermore, the above approximation (56) only maintains the first and the second terms of the Neuman expansion. However, its accuracy needs to be analyzed. The difference $\mathbf{D}$ between (54) and (56) satisfies the below inequality.

$$\|\mathbf{D}\| \leq \frac{\left\|{\mathbf{T}^{(\ell)}}^{-1}\mathbf{G}^{(\ell)}\right\|^2}{1 - \left\|{\mathbf{T}^{(\ell)}}^{-1}\mathbf{G}^{(\ell)}\right\|}. \tag{57}$$

It is obvious that $\|{\mathbf{T}^{(k)}}^{-1}\mathbf{G}^{(\ell)}\|$ should be sufficiently small to ensure accuracy. For achieving this goal, we can let $\|\mathbf{G}^{(\ell)}\| = \frac{\zeta}{\|{\mathbf{T}^{(k)}}^{-1}\|}$ with $\zeta \ll 1$. Then the ℓ^{th} updated total transfer model for the users is simplified as

$$\begin{aligned}
\mathbf{h}^{(\ell)}(\mathbf{G}^{(\ell)}) \approx& \mathbf{e} + \mathbf{h}_{\text{Cluster}} - \mathbf{r}{\mathbf{T}^{(\ell)}}^{-1}\mathbf{L} \\
&+ \mathbf{r}{\mathbf{T}^{(\ell)}}^{-1}\mathbf{G}^{(\ell)}{\mathbf{T}^{(\ell)}}^{-1}\mathbf{L}. \qquad (58)
\end{aligned}$$

In terms of simplicity, we introduce some shorthand variables as following

$$\mathbf{e} + \mathbf{h}_{\text{Cluster}} - \mathbf{r}{\mathbf{T}^{(\ell)}}^{-1}\mathbf{L} = \mathbf{c}^{(\ell)}, \tag{59}$$

$$\mathbf{r}{\mathbf{T}^{(\ell)}}^{-1} = \mathbf{z}^{(\ell)}, \tag{60}$$

$${\mathbf{T}^{(\ell)}}^{-1}\mathbf{L} = \mathbf{E}^{(\ell)}. \tag{61}$$

According to (59)-(61), (58) can be further substituted by

$$\mathbf{h}^{(\ell)}(\mathbf{G}^{(\ell)}) \approx \mathbf{c}^{(\ell)} + \mathbf{z}^{(\ell)}\mathbf{G}^{(\ell)}\mathbf{E}^{(\ell)}. \tag{62}$$

The above equation also has a similar cascade form, which is relatively easy to fix. With its help, we can reformulate the system problem in each iteration.

$$\begin{aligned}
\mathbf{P}_{\text{CA}-1}^{(\ell)}: \max_{\left\{\mathbf{w}_j^{(\ell)}, \mathbf{v}_l^{(\ell)}, \mathbf{G}^{(\ell)}\right\}} & \mathrm{EE}\left(\mathbf{w}_j^{(\ell)}, \mathbf{v}_l^{(\ell)}, \mathbf{G}^{(\ell)}\right) \\
\text{s.t. } & (52a),\\
& R_j\left(\mathbf{w}_j^{(\ell)}, \mathbf{G}^{(\ell)}\right) \geq R_j^{(\mathrm{D})}, \forall j, \qquad (63a)\\
& P_{\mathrm{RF},n}\left(\mathbf{w}_j^{(\ell)}, \mathbf{v}_l^{(\ell)}, \mathbf{G}^{(\ell)}\right) \geq \mathcal{G}_n(E_n^{(\mathrm{D})}), \forall n, \qquad (63b)\\
& \mathbf{G}_{q,q}^{(\ell)} = j\Delta_q^{(\ell)}, \forall q, \qquad (63c)\\
& \Delta_q^{(\ell)} \leqslant \zeta, \forall q. \qquad (63d)
\end{aligned}$$

Algorithm 3 Alternative optimization based on Dinkelbach's algorithm and Neuman series approximation

Initialize: ε_t, ε_i;

Set $\ell = 0$, $\jmath = 1$, $\mathbf{T}^{(0)} = \boldsymbol{\mathcal{Z}}_{\mathrm{II}} + R_0$, $\mathbf{G}^{(0)} = \mathrm{zeros}(N_i, N_i)$, $D_t > \varepsilon_t$, $\mathrm{EE}^{(0)} = 0$;
while $D_t \leq \varepsilon_t$ **do**
  Set $D_i > \varepsilon_i$, $\mathrm{EE}\,(0) = 0$;
  Achieve $\mathbf{h}^{(\ell)}$ through (58) for every user;
  while $D_i \leq \varepsilon_i$ **do**
    Adopt AO strategy, SDR approach, Dinkelbach's algorithm and SCA method proposed in Algorithm 1 and Algorithm 2 to solve problem $\mathbf{P}_{\text{CA}-1}^{(\ell)}$;
    Output solved $\left\{\mathbf{w}_j^{*(\ell)}(\jmath), \mathbf{v}_l^{*(\ell)}(\jmath), \mathbf{G}^{*(\ell)}(\jmath)\right\}$ to calculate $\mathrm{EE}\,(\jmath)$;
    Set $D_i = \mathrm{EE}\,(\jmath) - \mathrm{EE}\,(\jmath - 1)$;
    Update $\jmath = \jmath + 1$;
  end while
  Return $\mathbf{G}^{*(\ell)}(\jmath)$ and $\mathrm{EE}^{(\ell)} = \mathrm{EE}\,(\jmath)$;
  Set $D_t = \mathrm{EE}^{(\ell)} - \mathrm{EE}^{(\ell-1)}$;
  Achieve $\mathbf{T}^{(\ell+1)} = \mathbf{T}^{(\ell)} + \mathbf{G}^{*(\ell)}$;
  Update $\jmath = 1$, $\ell = \ell + 1$;
end while

Output: the optimal $\left\{\mathbf{w}_j^{*(\ell)}(\jmath), \mathbf{v}_l^{*(\ell)}(\jmath), \mathbf{\Omega}^{*(\ell)}(\jmath)\right\}$

(63c) is adopted to ensure only the imaginary part of the tunable circuit impedance is adjustable as the same setting in problem $\mathbf{P}_0$. In addition, (63d) is the guarantee of the Neuman approximation's accuracy. It is worth noting that the choices of ζ need to be careful. If it is too small, it makes the approximation imprecise, while the convergence rate may be slow with large ζ.

As to resolve problem $\mathbf{P}_{\text{CA}-1}^{(\ell)}$, the SDR method can be utilized again. Then the beamforming and the circuit parameter subproblems are tackled iteratively. The optimized variable $\mathbf{G}^{(\ell)}$ is used to update the transfer model until convergence. The overall structure of the optimization frameworks is shown in Algorithm 3 detailedly. We can further find that the optimization scheme for solving the inner layer of the coupling-awareness problem can fully cope with the no-coupling case. In a sense, this also proves that the no-coupling case is indeed less complex and just like the choice of the IDS-based RIS model. Nevertheless, the cost of this simplification is to ignore the key hardware feature, the mutual coupling, of the RIS, and hence may cause optimization bias. The inaccuracy of this ideal assumption will be discussed in section V.

A. Semi-Definite Relaxation

For beamforming vectors, the SDR forms of them coincide with the previous section. Whereas, $u^{(\ell)}\mathbf{B}_u^{(\ell)}$ can be adopted to substitute $\mathbf{z}^{(\ell)}\mathbf{G}^{(\ell)}\mathbf{E}^{(\ell)}$ in (62), where $\mathbf{u}^{(\ell)} \triangleq [j\Delta_1^{(\ell)}, \ldots, j\Delta_{N_i}^{(\ell)}]$ and $\mathbf{B}_u^{(\ell)} \triangleq \mathrm{diag}(\mathbf{z}^{(\ell)})\mathbf{E}^{(\ell)}$. Further, we can make $\hat{\mathbf{u}}^{(\ell)} = [\mathbf{u}^{(\ell)}, j]$ and $\hat{\mathbf{B}}_u^{(\ell)} = [\mathbf{B}_u^{(\ell)}; (-j)\mathbf{c}^{(\ell)}]$, where unit imaginary j is introduced for recovering the original vector easily after achieving the optimized vectors. Moreover, according to SDR, $\hat{\mathbf{U}}^{(\ell)} = \hat{\mathbf{u}}^{(\ell)^H}\hat{\mathbf{u}}^{(\ell)}$ is introduced with

$\hat{\mathbf{U}}^{(\ell)^H} \succeq 0$ and $\text{rank}(\hat{\mathbf{U}}^{(\ell)}) = 1$. With the help of the relaxed variables, the $\mathbf{P}_{\text{CA}-1}^{(\ell)}$ can be transformed into the following form.

$$\mathbf{P}_{\text{CA}-2}^{(\ell)}: \max_{\left\{\mathbf{W}_j^{(\ell)}, \mathbf{V}_l^{(\ell)}, \hat{\mathbf{U}}^{(\ell)}\right\}} \frac{\sum_{j=1}^{M_I} R_j\left(\mathbf{W}_j^{(\ell)}, \hat{\mathbf{U}}^{(\ell)}\right)}{P\left(\mathbf{W}_j^{(\ell)}, \mathbf{V}_l^{(\ell)}\right)} \tag{64a}$$

$$\text{s.t.} \frac{\text{Tr}\left(\hat{\mathbf{U}}^{(\ell)}\hat{\mathbf{B}}_{u,j}^{(\ell)}\mathbf{W}_j^{(\ell)}\hat{\mathbf{B}}_{u,j}^{(\ell)^H}\right)}{\text{Tr}\left(\hat{\mathbf{U}}^{(\ell)}\hat{\mathbf{B}}_{u,j}^{(\ell)}\sum_{k\neq j}^{M_I}\mathbf{W}_k^{(\ell)}\hat{\mathbf{B}}_{u,j}^{(\ell)^H}\right)+\sigma^2} \geq 2^{R_j^{(\text{D})}}-1, \forall j, \tag{64b}$$

$$\sum_{j=1}^{M_I}\text{Tr}\left(\mathbf{W}_j^{(\ell)}\right)+\sum_{l=1}^{M_E}\text{Tr}(\mathbf{V}_l^{(\ell)}) \leq P_{\text{Max}}, \tag{64c}$$

$$\text{Tr}\left(\hat{\mathbf{U}}^{(\ell)}\hat{\mathbf{B}}_{u,n}^{(\ell)}(\sum_{j=1}^{M_I}\mathbf{W}_j^{(\ell)}+\sum_{l=1}^{M_E}\mathbf{V}_l^{(\ell)})\hat{\mathbf{B}}_{u,n}^{(\ell)^H}\right) \geq \mathcal{G}_n(E_n^{(\text{D})}), \forall n, \tag{64d}$$

$$\mathbf{U}_{q,q}^{(\ell)} \leqslant \zeta^2, \forall q, \mathbf{U}_{N_i+1,N_i+1}^{(\ell)} = 1, \tag{64e}$$

$$\mathbf{W}_j^{(\ell)} \succeq \mathbf{0}, \mathbf{V}_l^{(\ell)} \succeq \mathbf{0}, \hat{\mathbf{U}}^{(\ell)} \succeq \mathbf{0}, \forall j, l. \tag{64f}$$

where (64e) is the SDR form of accuracy constraint. Particularly, the rate of j^{th} IDU $R_j\left(\mathbf{W}_j^{(\ell)}, \hat{\mathbf{U}}^{(\ell)}\right)$ and total power consumption $P\left(\mathbf{W}_j^{(\ell)}, \mathbf{V}_l^{(\ell)}\right)$ of objective formulation (64a) can be given by

$$R_j\left(\mathbf{W}_j^{(\ell)}, \hat{\mathbf{U}}^{(\ell)}\right) = \log_2\left(1+\frac{\text{Tr}\left(\hat{\mathbf{U}}^{(\ell)}\hat{\mathbf{B}}_{u,j}^{(\ell)}\mathbf{W}_j^{(\ell)}\hat{\mathbf{B}}_{u,j}^{(\ell)^H}\right)}{\text{Tr}\left(\hat{\mathbf{U}}^{(\ell)}\hat{\mathbf{B}}_{u,j}^{(\ell)}\sum_{k\neq j}^{M_I}\mathbf{W}_k^{(\ell)}\hat{\mathbf{B}}_{u,j}^{(\ell)^H}\right)+\sigma^2}\right), \tag{65}$$

$$P\left(\mathbf{W}_j^{(\ell)}, \mathbf{V}_l^{(\ell)}\right) = \sum_{j=1}^{M_I}\text{Tr}\left(\mathbf{W}_j^{(\ell)}\right)+\sum_{l=1}^{M_E}\text{Tr}(\mathbf{V}_l^{(\ell)})+P_c+N_iP_I. \tag{66}$$

Compared to the no-coupling case, although the core of the problem is varied, problem $\mathbf{P}_{\text{CA}-2}^{(\ell)}$ is non-convex and has a similar form as $\mathbf{P}_{\text{NC}-1}$, thus the same approach used in the previous section can be adopted to resolve it. For decomposing the above problem into several subproblems, the AO strategy is presented in the next subsection.

B. AO strategy

1) Optimization of $\left\{\mathbf{W}_j^{(\ell)}, \mathbf{V}_l^{(\ell)}\right\}$*:* In each iteration, beamforming vectors and the reactances of the control elements are optimized iteratively until the objective convergences. When solving the beamforming subproblem, the Dinkelbach algorithm based on SCA proposed in Algorithm 1 and Algorithm 2 still works under the condition that $\hat{\mathbf{U}}^{(\ell)}$ is known. Since the solution approach is identical, the detailed process may not be repeated here.

2) Optimization of $\hat{\mathbf{U}}^{(\ell)}$*:* When obtaining resolved $\mathbf{W}_j^{*(\ell)}$ and $\mathbf{V}_l^{*(\ell)}$, the equivalent subproblem is presented as

$$\mathbf{P}_{\text{CA}-2}: \max_{\left\{\hat{\mathbf{U}}^{(\ell)}, \dot{\varrho}_j, \dot{p}_j, \dot{q}_j\right\}} \sum_{j=1}^{M_I}\dot{\varrho}_j,$$

$$\text{s.t.}\text{Tr}\left(\hat{\mathbf{U}}^{(\ell)}\hat{\mathbf{B}}_{u,j}^{(\ell)}\mathbf{W}_j^{(\ell)}\hat{\mathbf{B}}_{u,j}^{(\ell)^H}\right) \geq e^{\dot{p}_j+\dot{q}_j}+\sigma^2e^{\dot{q}_j}, \forall j, \tag{67a}$$

$$e^{\dot{p}_j} \geq \text{Tr}\left(\hat{\mathbf{U}}^{(\ell)}\hat{\mathbf{B}}_{u,j}^{(\ell)}\sum_{k\neq j}^{M_I}\mathbf{W}_k^{(\ell)}\hat{\mathbf{B}}_{u,j}^{(\ell)^H}\right), \forall j, \tag{67b}$$

$$e^{\dot{q}_j} \geq 2^{\dot{\varrho}_j}-1, \forall j \tag{67c}$$

$$\hat{\mathbf{U}}^{(\ell)} \succeq \mathbf{0}, \tag{67d}$$

(64b), (64d), (64e)

The above problem is solved by the SCA method iteratively until the stop condition is satisfied. Moreover, when using eigenvalue decomposition or Gaussian randomization to recover $\hat{\mathbf{u}}^{(\ell)}$ from $\hat{\mathbf{U}}^{(\ell)}$, the output vectors are real vectors and then every element of the vectors must be transformed into the pure imaginary value for coinciding with the assumption in (63c).

V. Numerical Results and Full-Wave Simulation

To analyze the effectiveness of our proposed algorithm and the impacts of relevant hardware attributes on system performance in more detail, we first conduct a series of numerical simulation experiments for different QoS requirements of the users (i.e., energy harvesting for EHUs and minimum transmission rate for IDUs), various transmitting power levels and diverse RIS topologies (i.e., various RIS element spacings with the same element number and different numbers of RIS elements under the uniform surface area). Specifically, simulations are all conducted under different situations including the ideal case, the mutual-coupling awareness, the mutual-coupling unawareness, the no RIS case, and the non-operating RIS case. Among them, the mutual-coupling unawareness means that the optimization results from the no-coupling case are deployed in practical condition. In addition, the non-operating case denotes that the reactances of control circuits are not optimized. Besides, the full-wave simulation is performed to analyze the physical active beams at the BS and the reflecting beams of the RIS in several typical scenarios.

A. Numerical Results

The two-dimensional experimental scene is shown in Fig. 2. The center location of the 8-element-array BS is located at (0,0) and the RIS with 2×8 array configuration is located at (1,1). Two EHUs and two IDUs are generated at the rectangular regions of (x_{IDU},y_{IDU}) and (x_{EHU},y_{EHU}) with $x_{\text{EHU}} \in [0.5, 2]$ $y_{\text{EHU}} \in [-1, -2]$, $x_{\text{IDU}} \in [4, 15]$ and $y_{\text{IDU}} \in [-5, -24]$. Moreover, all dipoles are perpendicular to the plane.

As for the transfer model in (2), it has two submodels, the EM transfer model and the clustered model. Concerning the EM transfer model, the system operates at 28GHz. All

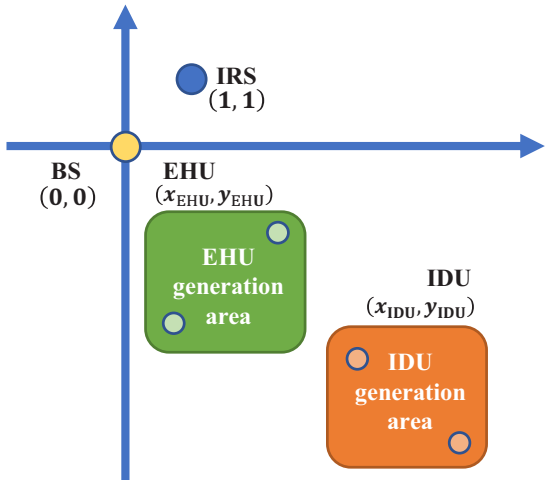


Fig. 2: The scenario setting.

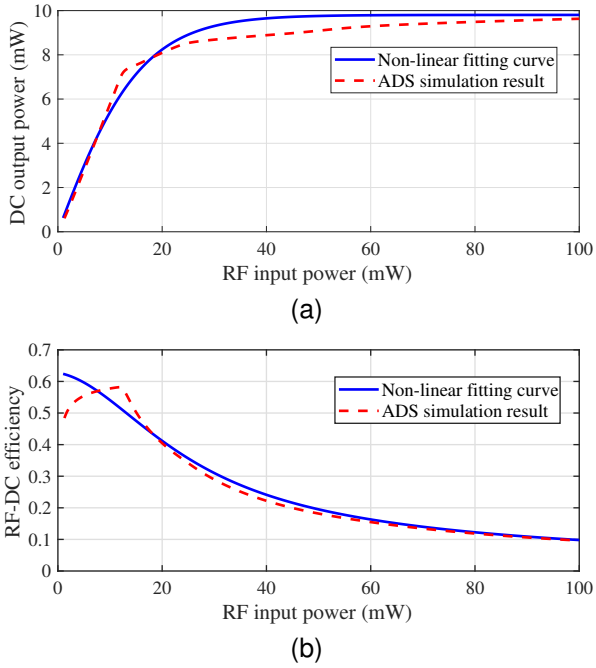
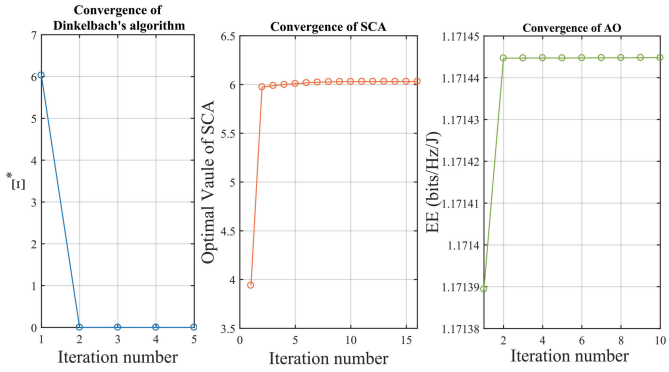
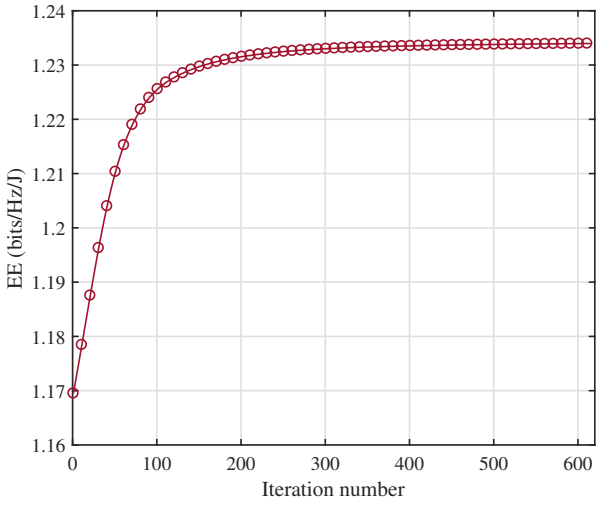


Fig. 3: (a) The DC output power versus RF input power. (b) The RF-DC efficiency versus RF input power.

antennas and RIS elements are dipoles with $\lambda/500$ radii and $\lambda/2$ lengths [15], [16]. The element spacings of the BS and the RIS are $\lambda/2$ and $\lambda/4$, respectively. The reason for the choice is that the RIS generally has more compact spacing for deployment. The characteristic impedance η_0 is set by 377 Ohm. The real part of the tunable circuit, the source impedance and the load impedance are 73 Ohm, which coincides with the resistance part of the antenna impedance as in [15], [16]. Regarding the clustered model in (16), η and ρ in the path loss equation are set by $\eta = 72$ and $\rho = 2.92$, respectively. The number of clusters follows $N_{\text{cl}} \sim \max\{\text{Poisson}(1.8), 1\}$ and the ray number in them follows uniform distribution $N_{\text{ray}} \sim U(0, 30)$. Complex gain $\alpha_{k,j}$ is a complex standard normal random variable, which follows: $\alpha_{k,p} \sim \mathcal{CN}(0, 1)$.



(a)



(b)

Fig. 4: (a) Convergence behaviors of AO strategy, Dinkelbach's algorithm and SCA method. (b) Convergence behavior of Algorithm 3.

Azimuth and elevation angles $(\theta_{k,p}^t, \phi_{k,p}^t)$ of departure are generated from the Laplacian distribution. The mean values of them uniformly distributed in $[-\pi/2, \pi/2]$ and standard deviations are set as 15°. The above parameter settings for the clustered model are set at 28GHz and can be traced back to [28], [31], [32].

The energy harvesting requirement for EHUs is $E_n^{(\text{D})} = 1.3\mu\text{W}, \forall n$ and the rate threshold for IDUs is $R_j^{(\text{D})} = 1\text{bit/s/Hz}, \forall j$. P_c at the BS and P_I at the control circuit are set by $P_c = 1$ W and $P_I = 10$ mW. Noise power is $\sigma^2 = 1 \times 10^{(-12)}$W. Transmitting power budget is configured to $P_{\text{Max}} = 4$ W. Moreover, we select the HSMS2850 rectifier diode as the core device to design a rectifier circuit through the Advanced Design System (ADS) 2017 software. The actual ADS simulation result is fit by the equation (21) as in Fig. 3. The fit circuit parameters of the rectifier are $B = 9.804$, $\varpi = -1.293$ and $\nu = 0.119$. All the above simulation configurations are the default unless specified in the comparison experiments.

1) Convergence behaviors of the proposed algorithms in coupling awareness: Both for the cases of no coupling and

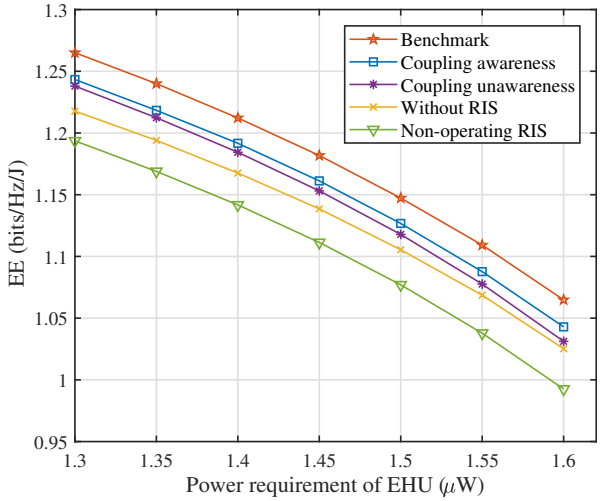


Fig. 5: EE versus power requirements of EHU.

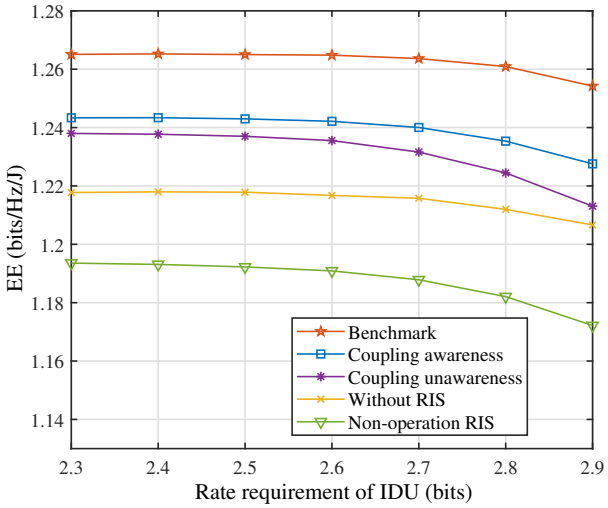


Fig. 6: EE versus rate requirements of IDU.

coupling awareness, the cores of the optimization framework are similar. The difference is that an additional iteration algorithm based on the Neuman series approximation should be introduced for tackling the coupling-awareness case. Interestingly, the algorithm used to resolve the no-coupling case can be seen as the inner solution of the algorithm adopted to solve the coupling-awareness case, hence the convergence behaviors of the algorithm in the case are exhaustive and the effectiveness of the proposed algorithm can be seen obviously from Fig. 4.

2) QoS requirements and resource allocation analysis: In the study of the EE performance under the different energy harvesting requirements in Fig. 5, the benchmark is set to the circumstance that the mutual-coupling effect doesn't exist, which is the case of section III. Furthermore, compared to the other situations, our proposed coupling-awareness algorithm is closest to it. Especially, the performance is better than the coupling-unawareness case, which indicates that the coupling attributes among the RIS elements cannot be ignored. In addition, the performances of the above conditions with the operating RIS are better than those in the no-RIS case and the non-operating RIS case, which demonstrates that RIS can improve EE performance effectively. Besides, the EE indicators of the above situations have the same downward trends with the increasing power demand of the EHU, as the transmitting power needs to be increased further to cope with the energy demands without enhancing the overall rate level.

As for the research that various rate thresholds impact EE performances in Fig. 6, the differences among all comparative situations are identical to those in the previous analysis. Specifically, the trends are the same and gradually declining. The reason for this phenomenon is that the power may be assigned to the IDUs with better channel conditions, while the rate constraints are accessible for all IDUs. Nevertheless, more power resources need to be allocated to the IDUs with poor channel conditions due to the continuously increasing rate requirements. Indeed, the cost of the enhancement is greater, especially in its later stage. In detail, the improvement of the

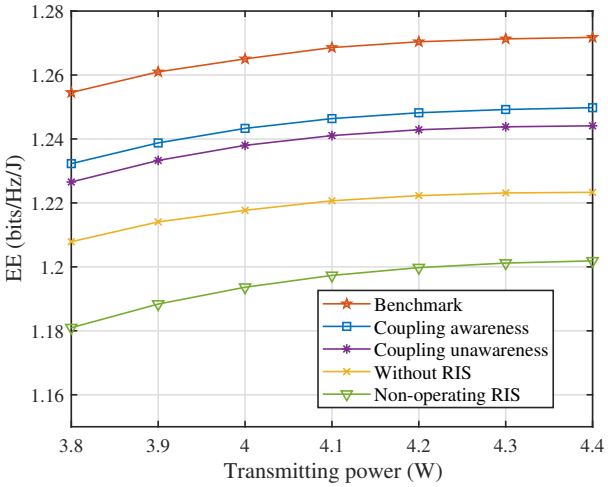


Fig. 7: EE versus transmitting power.

sum rate cannot keep up with the rate of resource consumption, which leads to decreasing EE levels.

In our work, the transmitting power is a guarantee of the QoS requirements and needs to be analyzed. From Fig. 7, we can find that the EE level rises as the transmitting power increases in every condition, however, the rate of growth decreases and levels off eventually. This is because at high transmit power values, the interference term grows noticeably and negatively affects the system's EE. Moreover, when in the later period of the power increment, the rate growth and power consumption are locked in a stalemate and thus the EE values maintain at the same levels. The above phenomenon shows that the transmitting power should be weighed carefully because of its dual character on EE.

3) The RIS topology analysis: For further improving the system performance, the RIS often contains more elements. Nevertheless, the array may be too cumbersome to deploy with a large number of elements. Therefore reducing the element spacing turns out to be a direct and effective method. From

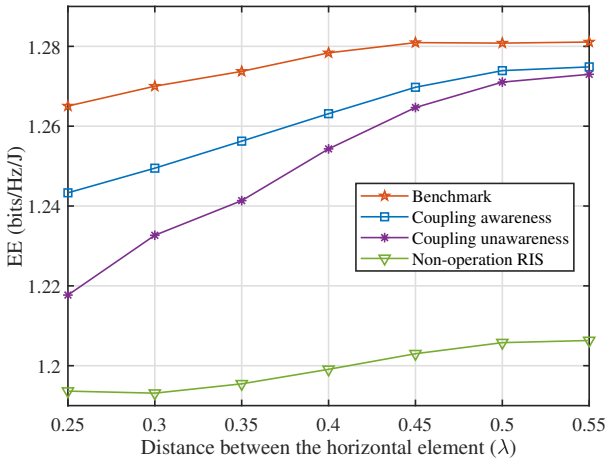


Fig. 8: EE versus element spacing of the IRS.

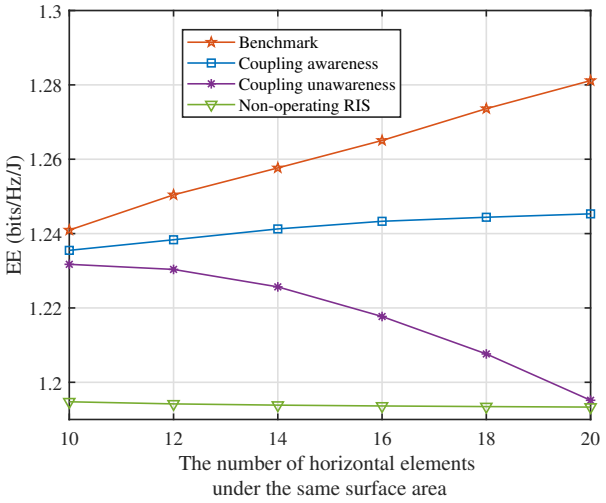


Fig. 9: EE versus number of RIS elements under the same surface area.

the perspective of engineering practice, the element spacing is an important design factor for the RIS, which is generally less than $\lambda/2$ [12]. Whereas, the smaller the spacing, the greater the mutual coupling. We research this influence on the EE in Fig. 8. After comparing the coupling-unawareness and the coupling-awareness case, it can be found that the tighter element spacing results in a larger performance gap between them. Additionally, the EE performances of all cases, including the ideal benchmark and the non-operating RIS case, improve as the spacing enlarges. We believe this is mainly owing to the increased effective area of the RIS.

Fixing the number of vertical elements and the total effective area of the RIS, we provide more horizontal elements to the RIS. From Fig. 9, as the increased mutual coupling caused by the high number of elements is not considered, the EE performance of the ideal benchmark has approximately linear enhancement with more elements. However, considering the condition of the actual coupling awareness, the EE improvement is slowing down, since more elements induce more serious mutual coupling. In detail, performance enhancements and coupling effects will work concurrently, resulting in later stages of sluggish growth. The mutual-coupling effect may be more obvious in the coupling-unawareness case, since this case ignores its effect and thus leads to optimization bias. Furthermore, its performance degrades as the aggravated coupling brought by more elements. Finally, since the effective area of the RIS remains at the same level, the non-operating RIS case has nearly unchanged EE performance. In short, the results show that the RIS configuration should be carefully taken into account as it affects EE performance.

B. Full-wave Simulation

For demonstrating the properties of the physical beams more clearly and to further prove the effectiveness of the algorithm, we perform a full-wave simulation by Ansys Electronics 2019 software. In detail, the radiation patterns and radar cross section (RCS) patterns are simulated for generating the active beams and reflecting beams according to the optimized parameters. In this section, we analyze the impacts of QoS performances, end locations and the role of the energy beamforming on the characteristics of the physical beams (i.e., active beams of the BS and reflecting beams of the RIS). To further demonstrate the beams' differences in various cases, we set the default power requirement satisfying $E^{(\mathrm{D})} = 1.0\mu\mathrm{W}$, and keep only one pair of the EHU and the IDU with positions, (0.5,-2) and (10,-11), while leaving other simulation parameters unchanged. Without loss of generality, we remove the clustered model and assume that there are no barriers blocking transmissions for simulation convenience. Moreover, we define the angles between two of these ends as $\vartheta_{\text{BS-RIS}}$, $\vartheta_{\text{BS-IDU}}$, $\vartheta_{\text{BS-EHU}}$, $\vartheta_{\text{RIS-IDU}}$ and $\vartheta_{\text{RIS-IDU}}$. The default settings will be used unless stated otherwise.

In Fig. 10, every subfig includes two parts, the active beams of the BS and the reflecting beams of the RIS. All active beams have the primary and the secondary sub-beams for serving the IDU and the EHU. However, the disparities between the two are various in different situations because of the varied QoS requirements of the users. In Fig. 10(b), as $E^{(\mathrm{D})}$ increases, the gain level of the IDU beam shrinks and it performs oppositely for the EHU beam compared to the default case in Fig. 10(a). Nevertheless, due to the intense demands of the IDU, this situation is reversed in Fig. 10(c). Furthermore, the differentiations of reflecting beams among them are insignificant, which point at 36°. This can be explained by the presence of the BS-IDU transfer link and the ease of covering users by the BS under the default position setting. This also means that the active beams play more essential roles and the reflecting beams only act as an aid to enhance the system performance.

When altering the positions of the ends as in Fig. 11, almost all the beams change significantly. As it can be seen in Fig. 11(a), adjusting the IDU to locate at (10,-4) causes the difference between $\vartheta_{\text{BS-IDU}}$ and $\vartheta_{\text{BS-EHU}}$ to increase, thus the BS is difficult to generate two sub-beams with a large gap for serving the users simultaneously. As a consequence, the RIS will share most of the responsibility for covering the

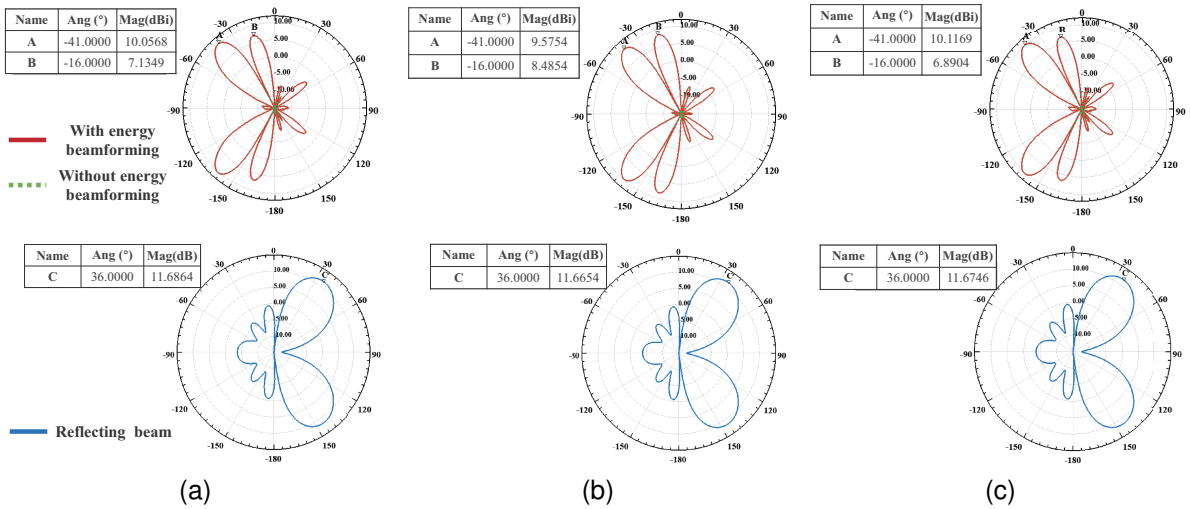


Fig. 10: (a) Beams for the default setting. (b) Beams for $E^{(\mathrm{D})} = 1.6\mu\mathrm{W}$. (c) Beams for $(\mathrm{D}) = 2.1\mathrm{bit/s/Hz}$.

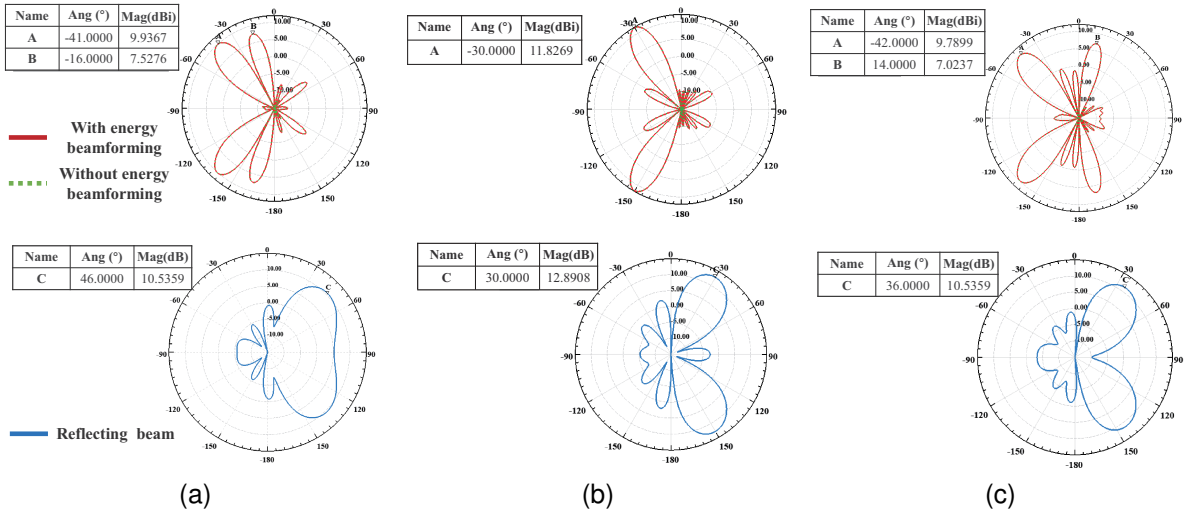


Fig. 11: (a) Beams for the IDU with position (10,-4). (b) Beams for the BS with position (0,-0.5). (c) Beams for the EHU with position (-0.5,-2)

IDU. Indeed, the form of the active beam is almost unaltered compared with the default case, but the angle of the main lobe of the reflecting beam changes to $46°$ in order to support the IDU with a large $\vartheta_{\text{RIS-IDU}}$. In 11(b), while placing the BS on (0,-0.5), the two sub-beams of the active beam turn into one. The reason for that is the closer difference between $\vartheta_{\text{BS-IDU}}$ and $\vartheta_{\text{BS-EHU}}$. Moreover, the reflecting beam points at $30°$ for assisting it to serve them. In 11(c), we put the EHU on the other side of the BS, while keeping the IDU unchanged. In line with expectations, the BS generates two opposite sub-beams at $-42°$ and $14°$, respectively. In this case, the reflecting beam is only produced for the IDU, which remains at $36°$.

In the above cases, with or without energy beamforming doesn't show any difference from the results of EM simulation. it also indicates that the energy beamforming is not essential and the absence of it will bring convenience to system design and algorithm optimization, which coincides with statements of [10] and [33].

VI. Conclusion

The hardware features and the performance analyses of an IRS-assisted-SWIPT MISO downlink system from the EM perspective are presented in this paper. Particularly, the transfer model based on the end-to-end EM transmission and multipath propagation was introduced to analyze the physical characteristics of the system. The problem was divided into two categories, the no coupling and the coupling awareness, and both of which were formulated as EE maximization problems subject to the QoS requirements and the hardware constraints. The effective algorithm frameworks were proposed to tackle the above non-convex problems. For the coupling-awareness case, the Neuman series approximation was adopted to simplify the transfer model. In each iteration of the approximation, the SDR transformation and the AO framework based

on Dinkelbach's algorithm and SCA approach were utilized to resolve the subproblems. Moreover, the no-coupling case can also adopt a similar scheme. The effectiveness of the proposed frameworks and the importance of mutual coupling were presented clearly in the numerical simulation. Specifically, the full-wave simulation results showed that the reflecting beams of the RIS may play different roles under various conditions and the energy beamforming is unnecessary. As for future research points, a more general end-to-end EM transfer system will be considered to reveal the deeper hardware features of the system. In detail, it will be compatible with more antenna configurations. Furthermore, the research about the more complex system settings with the co-located receivers will be conducted further.

REFERENCES

[1] D. C. Nguyen, M. Ding, P. N. Pathirana, A. Seneviratne, J. Li, D. Niyato, O. Dobre, and H. V. Poor, "6G internet of things: A comprehensive survey," *IEEE Internet Things J.*, vol. 9, no. 1, pp. 359–383, 2022.

[2] M. Z. Chowdhury, M. Shahjalal, S. Ahmed, and Y. M. Jang, "6G wireless communication systems: Applications, requirements, technologies, challenges, and research directions," *IEEE Open J. Commun. Soc.*, vol. 1, pp. 957–975, 2020.

[3] M. Giordani, M. Polese, M. Mezzavilla, S. Rangan, and M. Zorzi, "Toward 6G networks: Use cases and technologies," *IEEE Commun. Mag.*, vol. 58, no. 3, pp. 55–61, 2020.

[4] Q. Wu and R. Zhang, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network," *IEEE Commun. Mag.*, vol. 58, no. 1, pp. 106–112, 2019.

[5] R. Liu, Q. Wu, M. Di Renzo, and Y. Yuan, "A path to smart radio environments: An industrial viewpoint on reconfigurable intelligent surfaces," *IEEE Wireless Commun.*, vol. 29, no. 1, pp. 202–208, 2022.

[6] X. Yu, V. Jamali, D. Xu, D. W. K. Ng, and R. Schober, "Smart and reconfigurable wireless communications: From IRS modeling to algorithm design," *IEEE Wireless Commun.*, vol. 28, no. 6, pp. 118–125, 2021.

[7] T. D. P. Perera, D. N. K. Jayakody, S. K. Sharma, S. Chatzinotas, and J. Li, "Simultaneous wireless information and power transfer (SWIPT): Recent advances and future challenges," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 1, pp. 264–302, 2017.

[8] A. Costanzo, D. Masotti, G. Paolini, and D. Schreurs, "Evolution of SWIPT for the IoT world: Near- and far-field solutions for simultaneous wireless information and power transfer," *IEEE Microw. Mag.*, vol. 22, no. 12, pp. 48–59, 2021.

[9] Y. Ye, Y. Li, D. Wang, F. Zhou, R. Q. Hu, and H. Zhang, "Optimal transmission schemes for DF relaying networks using SWIPT," *IEEE Trans. Veh. Technol.*, vol. 67, no. 8, pp. 7062–7072, 2018.

[10] Q. Wu and R. Zhang, "Weighted sum power maximization for intelligent reflecting surface aided SWIPT," *IEEE Wireless Commun. Lett.*, vol. 9, no. 5, pp. 586–590, 2020.

[11] M. Di Renzo, A. Zappone, M. Debbah, M.-S. Alouini, C. Yuen, J. de Rosny, and S. Tretyakov, "Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2450–2525, 2020.

[12] X. Pei, H. Yin, L. Tan, L. Cao, Z. Li, K. Wang, K. Zhang, and E. Björnson, "RIS-aided wireless communications: Prototyping, adaptive beamforming, and indoor/outdoor field trials," *IEEE Trans. Commun.*, vol. 69, no. 12, pp. 8627–8640, 2021.

[13] M. Najafi, V. Jamali, R. Schober, and H. V. Poor, "Physics-based modeling and scalable optimization of large intelligent reflecting surfaces," *IEEE Trans. Commun.*, vol. 69, no. 4, pp. 2673–2691, 2021.

[14] S. Abeywickrama, R. Zhang, Q. Wu, and C. Yuen, "Intelligent reflecting surface: Practical phase shift model and beamforming optimization," *IEEE Trans. Commun.*, vol. 68, no. 9, pp. 5849–5863, 2020.

[15] G. Gradoni and M. Di Renzo, "End-to-end mutual coupling aware communication model for reconfigurable intelligent surfaces: An electromagnetic-compliant approach based on mutual impedances," *IEEE Wireless Commun. Lett.*, vol. 10, no. 5, pp. 938–942, 2021.

[16] A. Abrardo, D. Dardari, M. Di Renzo, and X. Qian, "MIMO interference channels assisted by reconfigurable intelligent surfaces: Mutual coupling aware sum-rate optimization based on a mutual impedance channel model," *IEEE Wireless Commun. Lett.*, vol. 10, no. 12, pp. 2624–2628, 2021.

[17] G. C. Trichopoulos, P. Theofanopoulos, B. Kashyap, A. Shekhawat, A. Modi, T. Osman, S. Kumar, A. Sengar, A. Chang, and A. Alkhateeb, "Design and evaluation of reconfigurable intelligent surfaces in real-world environment," *IEEE Open J. Commun. Soc.*, vol. 3, pp. 462–474, 2022.

[18] R. Fara, P. Ratajczak, D.-T. Phan-Huy, A. Ourir, M. Di Renzo, and J. de Rosny, "A prototype of reconfigurable intelligent surface with continuous control of the reflection phase," *IEEE Wireless Commun.*, vol. 29, no. 1, pp. 70–77, 2022.

[19] X. Qian and M. D. Renzo, "Mutual coupling and unit cell aware optimization for reconfigurable intelligent surfaces," *IEEE Wireless Commun. Lett.*, vol. 10, no. 6, pp. 1183–1187, 2021.

[20] S. Zargari, A. Khalili, Q. Wu, M. Robat Mili, and D. W. K. Ng, "Max-min fair energy-efficient beamforming design for intelligent reflecting surface-aided swipt systems with non-linear energy harvesting model," *IEEE Trans. Veh. Technol.*, vol. 70, no. 6, pp. 5848–5864, 2021.

[21] Y. Zhao, B. Clerckx, and Z. Feng, "IRS-aided SWIPT: Joint waveform, active and passive beamforming design under nonlinear harvester model," *IEEE Trans. Commun.*, vol. 70, no. 2, pp. 1345–1359, 2022.

[22] C. Pan, H. Ren, K. Wang, M. Elkashlan, A. Nallanathan, J. Wang, and L. Hanzo, "Intelligent reflecting surface aided MIMO broadcasting for simultaneous wireless information and power transfer," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1719–1734, 2020.

[23] Q. Wu and R. Zhang, "Joint active and passive beamforming optimization for intelligent reflecting surface assisted SWIPT under QoS constraints," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1735–1748, 2020.

[24] Y. Tang, G. Ma, H. Xie, J. Xu, and X. Han, "Joint transmit and reflective beamforming design for IRS-assisted multiuser MISO SWIPT systems," in *Proc. IEEE Int. Conf. Commun. (ICC), Dublin, Ireland, Jun. 2020*, pp. 1–6.

[25] J. Liu, K. Xiong, Y. Lu, D. W. K. Ng, Z. Zhong, and Z. Han, "Energy efficiency in secure IRS-aided SWIPT," *IEEE Wireless Commun. Lett.*, vol. 9, no. 11, pp. 1884–1888, 2020.

[26] Z. Chen, J. Tang, N. Zhao, M. Liu, and D. K. C. So, "Hybrid beamforming with discrete phase shifts for RIS-assisted multiuser SWIPT system," *IEEE Wireless Commun. Lett.*, pp. 1–1, 2022.

[27] S. Zargari, A. Khalili, and R. Zhang, "Energy efficiency maximization via joint active and passive beamforming design for multiuser MISO IRS-aided SWIPT," *IEEE Wireless Commun. Lett.*, vol. 10, no. 3, pp. 557–561, 2021.

[28] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, 2014.

[29] W. Dinkelbach, "On nonlinear fractional programming," *Manage. Sci.*, vol. 13, no. 7, pp. 492–498, 1967.

[30] G. W. Stewart, *Matrix algorithms: volume 1: basic decompositions*. Society for Industrial and Applied Mathematics, 1998.

[31] S. Buzzi and C. D'Andrea, "On clustered statistical MIMO millimeter wave channel simulation," *arXiv preprint arXiv:1604.00648*, 2016.

[32] M. R. Akdeniz, Y. Liu, M. K. Samimi, S. Sun, S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE J. Sel.Areas Commun.*, vol. 32, no. 6, pp. 1164–1179, 2014.

[33] J. Xu, L. Liu, and R. Zhang, "Multiuser MISO beamforming for simultaneous wireless information and power transfer," *IEEE Trans. Signal Process.*, vol. 62, no. 18, pp. 4798–4810, 2014.

Guest Editors
IEEE Internet of Things Journal
SI: Symbiotic Active/Passive Communications for the Internet of Things (IoT)

Subject: Responses to the Editor and Reviewers
Manuscript ID : IoT-26203-2022
Author: Ma, Ruoyan; Tang, Jie; Zhang, Xiu Yin; Wong, Kai-Kit; Chambers, Jonathon

December 19, 2022

Dear Editors,

We are very grateful for the constructive remarks of the editor and the anonymous reviewers, which have all been carefully incorporated into the revised manuscript. We have significantly modified the paper, where the analysis of the proposed solutions have been improved by providing some necessary explanation and modification. Furthermore, practical simulation results are also modified. As requested, point-by-point responses to received comments are provided in the following pages.

After addressing the major issues raised, we trust that the technical quality and clarity of the paper is significantly improved.

Yours sincerely,
Ma, Ruoyan; Tang, Jie; Zhang, Xiu Yin; Wong, Kai-Kit; Chambers, Jonathon

I. RESPONSE TO REVIEWER 1's COMMENTS

In this paper, the authors adopted a novel communication model based on electromagnetism to expose some crucial hardware features of the various ends. Remarkably, the mutual-coupling effect of the RIS was analyzed exhaustively. For comparison, the authors considered both the no-mutual-coupling case and the mutual-coupling-awareness case. Further, two optimization frameworks were proposed to solve the problems under these cases. Finally, the numerical results demonstrated the effectiveness of the algorithms. The article is innovative and logically rigorous. Nevertheless, there are some issues that need to be resolved further.

Response: Thanks for your positive comments. We greatly appreciate your useful comments that helped improve the original paper. All your comments are adopted in the revised manuscript. We have highlighted the changes in blue color in the revised manuscript.

Comment 1: *Why do the authors only consider the mutual-coupling effect of the RIS? It is evident that the effect also occurs at the transmitting array.*

Response: Thank you very much for your insightful comments. We completely agree with the reviewer that the mutual-coupling effect also occurs at the transmitting array. However, as for the adopted end-to-end EM transfer model, the coupling effect of the transmitting array has been considered in the model. In page 4, line 8, we define the self-impedance matrix $\boldsymbol{\mathcal{Z}}_{\mathrm{AB}}$, which includes the self impedances and mutual impedances of the elements in an array. In detail, as for the transmitting array, the form of the matrix can be presented as follows.

$$\boldsymbol{\mathcal{Z}}_{\mathrm{TT}} = \begin{bmatrix} \mathrm{z}_{11} & \cdots & \mathrm{z}_{1N_t} \\ \vdots & \ddots & \vdots \\ \mathrm{z}_{N_t1} & \cdots & \mathrm{z}_{N_tN_t} \end{bmatrix}$$

In the above matrix, the off-diagonal parts can express the mutual-coupling effects. Moreover, the parts are not ignored during the entire simulation, which means that the coupling effect of the transmitting array has been considered already.

Comment 2: *In this work, the authors fix the resistances of the control circuits and only optimize the reactance. Are these settings practical?*

Response: The authors are grateful for your positive comments. From the perspective of engineering deployment, we assume that the RIS elements are able to be adjusted by the varactors, which are semiconductors designed for controlling the capacitances of the circuits without changing their resistances. In detail, the varactors can be configured by the input voltages continuously. It is the reason why we only optimize the reactances of the control circuit and the rationality of the continuous reconfigurability of the RIS. Indeed, this kind of research setting coincides with the configurations of the references [15] and [16].

Comment 3: *How does equation (54) turn into equation (55)? I highly recommend that the authors provide the details of it.*

Response: The authors are thankful for exposing the problem. The transformation from equation (54) to equation (55) is crucial for Neuman series approximation. The detailed process can be given as follow. First, when supposing $(\mathbf{I}-\mathbf{P})$ is nonsingular, we have the equation as

$$(\mathbf{I}-\mathbf{P})\left(\mathbf{I}+\mathbf{P}+\mathbf{P}^2+\cdots+\mathbf{P}^k\right)=\mathbf{I}-\mathbf{P}^{k+1}$$

Where $\mathbf{P} = -\mathbf{T}^{(\ell)^{-1}}\mathbf{G}^{(\ell)}$. Multiplying this identity by $(\mathbf{I}-\mathbf{P})^{-1}$ and then removing the $(\mathbf{I}-\mathbf{P})^{-1}$ terms from both sides. We can get

$$\left(\mathbf{I}+\mathbf{P}+\mathbf{P}^2+\cdots+\mathbf{P}^k\right)-(\mathbf{I}-\mathbf{P})^{-1}=-(\mathbf{I}-\mathbf{P})\mathbf{P}^{k+1}$$

When $\lim_{k\to\infty}\mathbf{P}^k=0$, the $\mathbf{I}-\mathbf{P}$ will satisfy as

$$\mathbf{I}+\mathbf{P}+\mathbf{P}^2+\cdots+\mathbf{P}^k=(\mathbf{I}-\mathbf{P})^{-1}$$

The simplified proof of the transformation is done. It is worth mentioning that $\lim_{k\to\infty}\mathbf{P}^k=0$ is very crucial and also the origin of the condition $\zeta \ll 1$ in the manuscript. The related details about the above proof can be traced back to the reference [28]. As they are exhaustive, we omit the specific steps. However, just like the reviewer suggested, for clarity, we have included the explanation in the revised manuscript as follows:

Section IV. (page 9):

The transformation from (54) to (55) can be traced to [30]. Furthermore, the above approximation (56) only maintains the first and the second terms of the Neuman expansion. However, its accuracy needs to be analyzed. The difference $\mathbf{D}$ between (54) and (55) satisfies the below inequality.

$$\|\mathbf{D}\| \leq \frac{\left\|\mathbf{T}^{(\ell)^{-1}}\mathbf{G}^{(\ell)}\right\|^2}{1-\left\|\mathbf{T}^{(\ell)^{-1}}\mathbf{G}^{(\ell)}\right\|}. \tag{57}$$

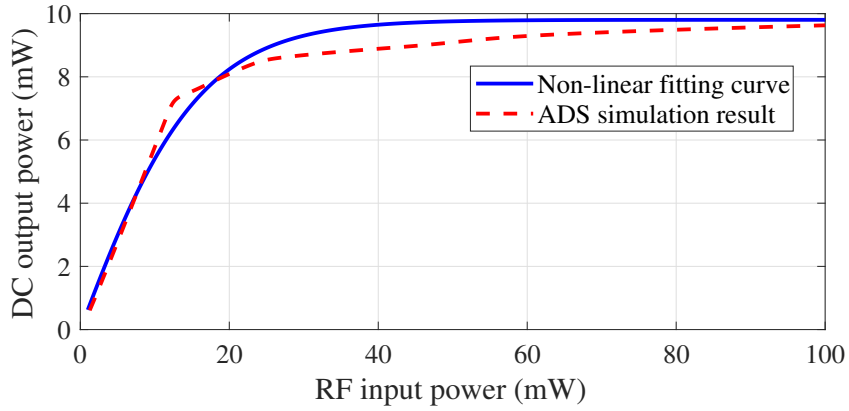
Comment 4: *Is it necessary to design a rectifying circuit for extracting the parameters of the non-linear fitting curve? The authors need to point out the significance of it.*

Response: The authors are grateful for your insightful remarks. Concerning the EHU, its essential hardware is the rectenna, which includes the receiving antenna and the rectifier. Moreover, the receiving antenna is considered a dipole antenna in this manuscript, thus the performance of EHU mainly relies on the design of the rectifier. The influence factors of rectifying efficiency are types of diodes, input power and load devices. Among them, diodes play an indispensable role and need to be treated carefully. From the point of view of the engineering design, the choices of diodes should follow the features of various scenarios due to their nonlinearity, which is caused by diode reverse bias. In detail, their maximum DC power is also related to this feature and can be presented by $\frac{V_{br}^2}{4R_L}$, where V_{br} and R_L denote the reverse breakdown voltage and the load resistance. In the following table, we present some types of diodes.

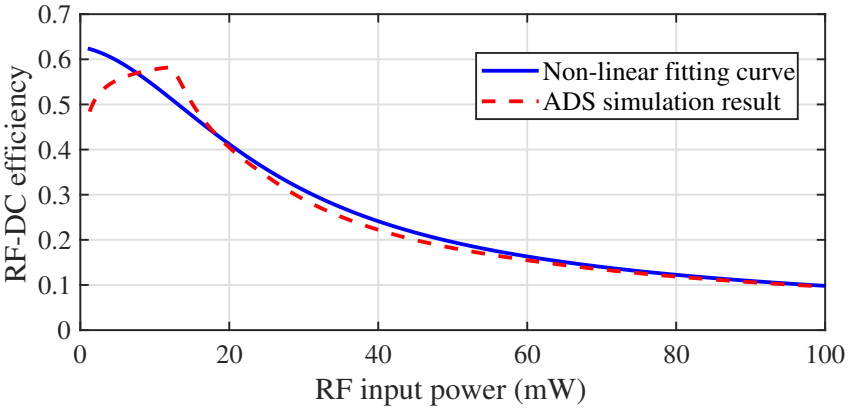
Type of diode	Reverse breakdown voltage (V)
HSMS − 2810	25
HSMS − 2820	15
HSMS − 2850	3.8
HSMS − 2860	7
SMS − 7630	2

The parameters of the various diodes show that the different diodes operate at distinct power ranges. For this reason, regarding our scenario, we approximate the input power level of the rectifier and then design a circuit based on the proper diode, which is identical to the practice during the engineering. Furthermore, considering the reviewer's suggestion, we have also modified the simulation results of the rectifying circuit.

Section V. (page 11):



(a)



(b)

(a) The DC output power versus RF input power. (b) The RF-DC efficiency versus RF input power.

Comment 5: *What is the meaning of the non-operating RIS? In order not to cause misconception, the authors should describe it clearly.*

Response: Thank you for bringing this to our attention. We completely agree with the reviewer that the definition of the non-operating RIS is unclear, and thus we have provided more explicit expressions for comparative situations as follows.

Section V. (page 10):

Specifically, simulations are all presented under different situations including the ideal case, the mutual-coupling awareness, the mutual-coupling unawareness, the no RIS case, and the non-operating RIS case. Among them, the mutual-coupling unawareness means that the optimization results from the no-coupling case are deployed in practical condition. In addition, the non-operating case denotes that the reactances of control circuits are not optimized.

Comment 6: *In Fig. 8, the distance between the horizontal element is spaced from 0.25 λ to 0.55λ.*

Why do the authors choose this range? Please clarify. .

Response: Thank you very much for your valuable comments. From the perspective of engineering practice, the element spacing is an important design factor for the RIS, which is generally less than $\lambda/2$. Because the RIS may have the relatively larger surface than the normal transmitting array, the element spacing need to be as small as possible for deployment. There are some representative examples of choosing small-element-spacing RIS for practical application. From [16] cited in the manuscript, we can see that the authors also choose the range of element spacing less than $\lambda/2$. Even in [12], the authors make reflective elements densely packed without spacing. To further illustrate the significance of choosing this range, we have modified the content in the manuscript.

Section V. (page 12):

We research this influence on the EE in Fig. 8. After comparing the coupling-unawareness and the coupling-awareness case, it can be found that the tighter element spacing results in a larger performance gap between them.

Section V. (page 13):

Finally, since the effective area of the RIS remains at the same level, the non-operating RIS case has nearly unchanged EE performance. In short, the results show that the RIS configuration should be carefully taken into account as it affects EE performance.

Comment 7: *How do the authors plot the reflecting beams of the RIS in the full-wave simulation? It is obvious that the reflecting beams are different from the norm active beams. The authors should explain the process of generating these plots.*

Response: Thank you very much for your kind suggestion. Indeed, all the active beams and reflecting beams are generated from the full-wave simulation by Ansys Electronics 2019 software. In addition, we import the beamforming vectors into it for generating radiation patterns of the transmitting array, namely, active beams. Then choosing the direction of incident waves and inputting the optimized impedances of control circuits for generating the radar cross section (RCS) patterns of the RIS, namely, the passive beams. For elaborating on this process as the reviewer recommend, we have revised the manuscript as follows.

Section V. (page 13):

For demonstrating the properties of the physical beam more clearly and to further prove the effectiveness of the algorithm, we perform a full-wave simulation by Ansys Electronics 2019 software. In detail, the radiation patterns and radar cross section (RCS) patterns are simulated for generating the active beams and reflecting beams according to the optimized parameters. In this section, we analyze the impacts of QoS performances, end locations and the role of the energy beamforming on the characteristics of the physical beam (i.e., active beams of the BS and reflecting beams of the RIS).

II. RESPONSE TO REVIEWER 2's COMMENTS

The authors view the RIS-aided SWIPT system from a novel perspective. The mutual-coupling influence, which is one of the most critical indicators in antenna designs, is introduced into the communication system. For solving intricate problems based on the adopted hardware system model, the authors use SDR, DinkelBach's strategy, and SCA to tackle them alternatively. In my opinion, this paper is valuable for analyzing the influences of the RIS hardware features on system performances. However, some unclear expressions and obvious errors must be mentioned.

Response: Thank you very much for your feedback. We are grateful for your expert comments that helped improve the original paper. All your remarks are adopted in the revised manuscript. We have highlighted the major changes in blue color in the revised manuscript.

Comment 1: *From the system model, it can be found that the authors use a kind of mixed transfer model, which consists of the determinate part and the random part. However, this kind of setting is rare in most articles. The authors need to explain it.*

Response: Thank you very much for your insightful comments. In our opinion, a RIS-assisted communication system is dominated mainly by the digitally controllable scatterers and the aggregated multipath scatterers. Moreover, for the transfer links based on the physical entities, the direct link (BS-user) and the RIS-assisted link (BS-RIS-user), EM transfer model can describe its characteristics exhaustively. However, as the scatterer links are constructed on arbitrary and irregular matters, we cannot utilize the EM approaches to analyze it and thus introduce the multipath links presented in (16). The idea of the hybrid transfer model coincides with the reference [16]. From the reference, the total transfer model also consists of two kinds of scatterers, which is identical to our setting. In order to show the rationality of this model, we have modified the related content of the manuscript.

Section II. (page 4):

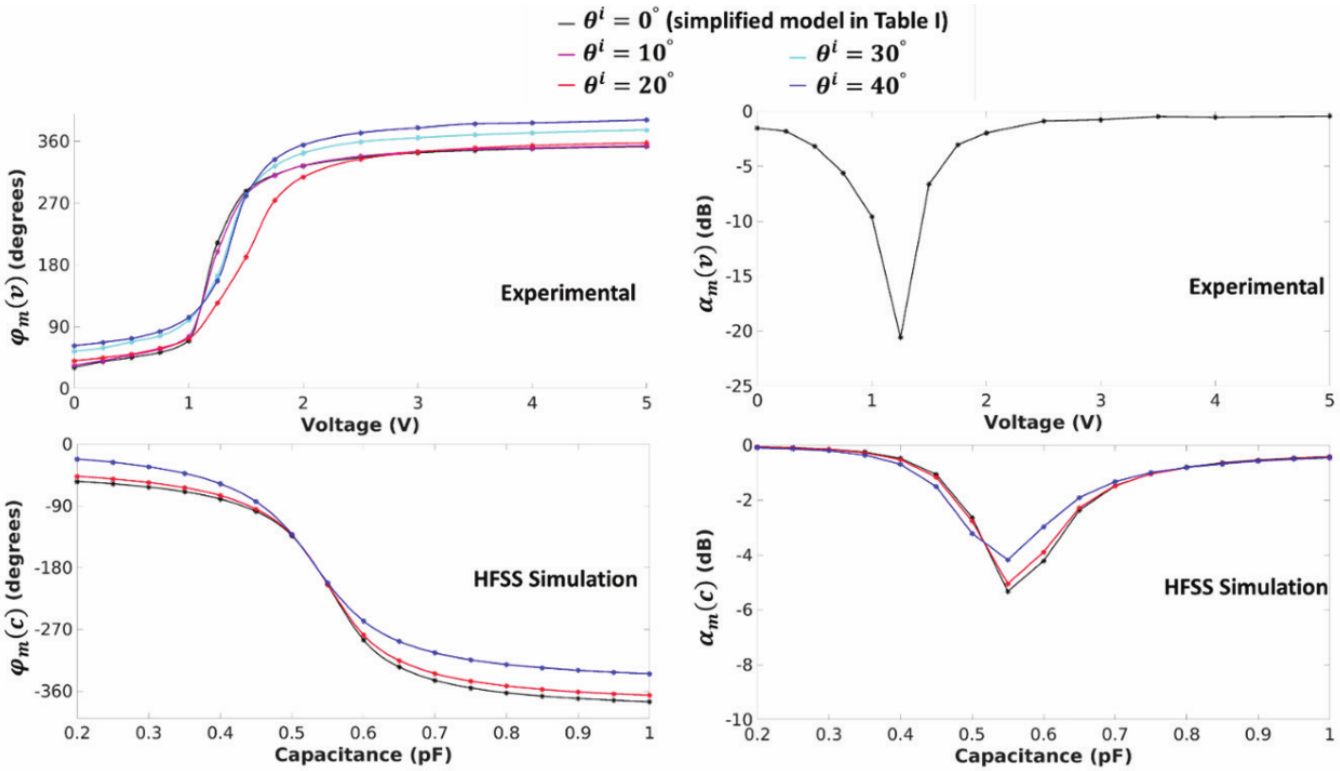
We assume that the scenario has two types of multipath propagation scatterers, the reconfigurable scatterer (i.e., the RIS) and the arbitrary scatterer, which coincides with the configuration of [16].

Comment 2: As for the EM transfer model, the authors assume that all the entities are dipole antennas. To the best of my knowledge, the elements of the RIS are the patch antennas and the RIS has a reflecting backplane, generally. The authors should explain why you make this assumption.

Response: The authors are grateful for your insightful remarks. All the entities are indeed dipole antennas. It is obvious that the current analysis of dipoles is simpler than other radiation and reflecting entities. On the other hand, their mutual impedances and self impedances can be calculated by the explicit expression. In short, simplifying the entities can save us a lot of computational trouble and make us only concentrate on the key physical features of the systems, which is the result of a trade-off. According to [15], a similar setting about the end-to-end EM transfer model is also limited to the dipole antenna. Nevertheless, we also agree with the reviewer that the dipole-antenna assumption is too limiting. Therefore, we will propose more general EM models based on it in the future as it is mentioned in the conclusion part.

Comment 3: In problem $\mathbf{P}_0$, constraint (25f) indicates that the control circuits have continuously adjustable capacity, whereas the statuses of the elements are discrete in most of the practical deployment. The authors need to explain the reason for this setting.

Response: Thank you very much for your valuable comments. In page 2, it mentions the control approaches of the RIS, which can be separated into two categories, the discrete one and the continuous one. The discrete RIS is the more common type during engineering practice, since it is easier to deploy. Furthermore, there also is a way, utilizing the varactors, to realize the continuous RIS. Particularly, the capacitances of the varactors can be configured by the continuously changing input power. This kind of setting coincides with the configurations of the references [15] and [16]. Moreover, the following figure from the reference [18] shows the properties of the varactors.



Different reflection coefficients under the continuously changing input voltages [18].

The above figure presents that the reflection coefficients can be continuously adjusted by the varactors, which demonstrates our setting is proper. For expressing more clearly, we have made the following modification.

Section I. (page 2):

1-bit discrete RIS is realized by PIN diodes working on the on-off states in [12] and the varactor diodes operating in two states with 180° phase difference in [17]. Moreover, the design based on the varactors is also able to achieve continuous tunability through controlling input voltages [18].

Comment 4: *In equation (49), "ln" is missing. Moreover, the update approach of the parameter $\tau_j^{*(t)}(i)$ is not given further. The authors should double-check these details.*

Section III. (page 8):

Response: Thank you for bringing this to our attention. They really are prominent problems. To comply with your comments, we have added the missing "ln" and the calculation equation of $\tau_j^{*(t)}(i)$ as follows.

$$\bar{a}_j^{(t)}(i) = \ln\left(\mathrm{Tr}\left(\hat{\mathbf{F}}\hat{\mathbf{L}}_{u,j}\sum\nolimits_{k\neq j}^{M_I}\mathbf{W}_k^{*(t)}(i)\hat{\mathbf{L}}_{u,j}^H\right)\right), \tag{48}$$

$$\bar{b}_j^{(t)}(i) = \ln\left(2^{\tau_j^{*(t)}(i)} - 1\right), \tag{49}$$

where $\mathbf{W}_k^{*(t)}(i)$ is obtained from the t^{th} solution of the SCA iteration and $\tau_j^{*(t)}(i)$ can be replaced continuously by

$$\tau_j^{*(t)}(i) = \log_2\left(1 + \frac{\mathrm{Tr}\left(\hat{\mathbf{F}}\hat{\mathbf{L}}_{u,j}\mathbf{W}_j^{*(t)}(i)\hat{\mathbf{L}}_{u,j}^{II}\right)}{\mathrm{Tr}\left(\sum_{k\neq j}^{M_I}\hat{\mathbf{F}}\hat{\mathbf{L}}_{u,j}\mathbf{W}_k^{*(t)}(i)\hat{\mathbf{L}}_{u,j}^{I}\right) + \sigma^2}\right). \tag{50}$$

Comment 5: *The authors should include the references that support the simulation parameters. Otherwise, the reliability of the numerical conclusions is in doubt.*

Response: The authors are thankful for the kind suggestion of the reviewer. We also agree that the support references are not enough. First, for the position setting of the various ends, the distance of BS-EHU is further than the distance of BS-IDU, the reason is mentioned in section II and we omit it in section V. For clarifying other simulation settings, the expressions have been modified. Moreover, additional references have been added to support the selected channel parameters for improving reliability.

Section V. (page 10):

The two-dimensional experimental scene is shown in Fig. 2. The location of the 8-element-array BS is located at (0,0) and the RIS with 2×8 array configuration is located at (1,1). Two EHUs and two IDUs are generated at the rectangular regions of (x_{IDU},y_{IDU}) and (x_{EHU},y_{EHU}) with $x_{\text{EHU}} \in [0.5, 2]$ $y_{\text{EHU}} \in [-1, -2]$, $x_{\text{IDU}} \in [4, 15]$ and $y_{\text{IDU}} \in [-5, -24]$. Moreover, all dipoles are perpendicular to the plane.

As for the transfer model in (2), it has two submodels, the EM transfer model and the clustered model. Concerning the EM transfer model, the system operates at 28GHz. All antennas and RIS elements are dipoles with $\lambda/500$ radii and $\lambda/2$ lengths [15], [16]. The element spacings of the BS and the RIS are $\lambda/2$ and $\lambda/4$, respectively. The reason is that the RIS generally has more compact spacing for deployment. The characteristic impedance η_0 is set by 377 Ohm. The real part of the tunable circuit, the source impedance and the load impedance R_0 are 73 Ohm, which coincides with the resistance part of the dipole as in [15], [16]. Regarding the clustered model in (17), η and ρ in the path loss equation are set by $\eta = 72$ and $\rho = 2.92$, respectively. The number of clusters follows $N_{\text{cl}} \sim \max\{\text{Poisson}(1.8), 1\}$ and the ray number in them follows uniform distribution $N_{\text{ray}} \sim U(0, 30)$. Complex gain $\alpha_{k,j}$ is a complex standard normal random variable, which follows: $\alpha_{k,p} \sim \mathcal{CN}(0, 1)$. Azimuth and elevation angles $(\theta_{k,p}^t, \phi_{k,p}^t)$ of departure are generated from the Laplacian distribution. The means of them uniformly distributed in $[-\pi/2, \pi/2]$ and standard deviations are set as 15°. The above parameter settings for the clustered model are set at 28GHz and can be traced back to [27],[30],[31].

Comment 6: *There are differences between the scenario settings in the numerical simulation and the full-wave simulation. Particularly, the authors assume that the BS-IDU links are blocked in the numerical simulation, but the links are existent in the full-wave simulation, and hence the settings are inconsistent.*

Response: Thank you very much for your valuable comments. We indeed make the simulation parameters of transfer models unchanged and alter the other settings in the full-simulation part. Firstly, in the comparative experiment of the full-wave simulation, the default energy requirement is changed to leave

more space for enlarging. Secondly, for distinguishing the beams more easily, we only keep a pair of EHU and IDU with fixed positions. Finally, because the Ansys software cannot simulate the conditions of BS-IDU links with barriers, we remove them from the scenario. For clarity, we have modified our expressions about this issue.

Section V. (page 13):

To further demonstrate the beams' differences in various cases, we set the default power requirement satisfying $E^{(\mathrm{D})} = 1.0\mu\mathrm{W}$, and keep only one pair of the EHU and the IDU with positions, (0.5,-2) and (10,-11), while leaving other simulation parameters unchanged. Without loss of generality, we remove the clustered model and assume that there are no barriers blocking transmissions for simulation convenience.

Comment 7: *What does "more general system" mean in the conclusion part? If the authors want to express that the radiation entities are no longer restricted to dipoles, the authors should state more clearly.*

Response: We are grateful for the reviewer's constructive opinion. We all agree the words are unclear in the conclusion part. Therefore the more explicit statements of future research have been revised.

Section VI. (page 16):

As for future research points, a more general end-to-end EM transfer system will be considered to reveal the deeper hardware features of the system. In detail, it will be compatible with more antenna configurations.

Comment 8: *The authors should pay attention to the references. For instance, "IEEE Transactions on Communications"−> "IEEE Trans. Commun".*

Response: Thank you very much for bringing the reference issues to our attention. As the reviewer's suggestion, we alter all the journal names of the references in the manuscript, e.g., "IEEE Transactions on Communications" −> "IEEE Trans. Commun", "IEEE Internet of things journal" −> "IEEE Internet Things J", "IEEE Open Journal of the Communications Society," −> " IEEE Open J. Commun. Soc" etc. The other details of the modification have been presented in the revised manuscript. Due to the limited space, we omit them here.

III. RESPONSE TO REVIEWER 3's COMMENTS

Unlike optimizations for reflection coefficients of the RIS, the authors focused on the practical circuit parameters. Overall, this paper is well written and easy to follow. However, several statements and analyses are unclear, which should be corrected before being considered for publication.

Response: We are thankful for the reviewer's constructive opinions. Several prominent problems in the manuscript are presented explicitly. Moreover, some insightful suggestion is proposed to improve its quality. All your remarks are adopted in the revised manuscript. We have highlighted the major changes in blue color in the revised manuscript.

Comment 1: *The authors construct the end-end transmission model, which consists of two submodels. The LoS links are built on the EM transfer submodel. According to the authors' description, this kind of submodel is determinate. However, this is contrary to the random channel environment. The authors should point out its validity and rationality.*

Response: The authors are grateful for the kind comments of the reviewer. In a RIS-assisted communication system, we hold the opinion that there are two kinds of scatterers, the reconfigurable scatterer (RIS) and the random scatterer. According to these scatterers, the total links can be separated into the direct link (BS-user), the RIS-assisted link (BS-RIS-user) and the scatterer link (BS-scatterer-user). Among them, the BS-user link and the BS-RIS-user link are based on physical entities, which means specific antenna configurations. Moreover, It is possible to analyze their transmission characteristics by using EM methods. However, the EM analysis mainly relies on the assumption of fixed environments. Therefore, the related EM transfer models are generally determinate. Indeed, the idea of the hybrid transfer model coincides with the reference [16]. Furthermore, it also can be found that the other physical models and circuit models are determinate in [6].

Comment 2: *In section II, the authors state that the information decoder users are more likely to be interrupted than the energy harvesting users, which should be explained in more details.*

Response: Thank you very much for your useful comments. In the manuscript, the service scopes of the EHUs are smaller than those of IDUs. The reason for choosing this setting is that the path loss is tremendous due to the 28GHz operating frequency. Moreover, the receiving power is too small to harvest for the EHUs but still enough for the IDUs at the relatively far distance from the BS. Particularly, the EHUs are set within two meters of the BS, so we assume the BS-EHU links cannot be interrupted. On the contrary, the IDUs with longer service distances have more possibility to be obstructed.

Comment 3: *In the problem formulation part, the parameter of constraint (25e) in problem* P_0 *is set as a fixed positive number. However, the authors should give more information about why this setting is appropriate.*

Response: Thank you for bringing this issue to our attention. R_0 in (25e) denotes the loss of the circuit and $R_0 \geq 0$ is the guarantee that the RIS operates in the passive mode instead of the active mode. Moreover, as the real part of dipole antennas is 73 Ohm, we assume the $R_0 = 73$ Ohm for matching. In addition, the varactors are utilized to configure the states of the RIS elements, so only the reactance can be optimized and the resistance R_0 is unchanged. Those are the reasons for choosing the fixed positive resistances of the control circuits. It is worth to be mentioned that this setting coincides with the configurations of the references [15] and [16]. For further expressing explicitly, we also have modified the content as the reviewer's comments.

Section V. (page 10):

The real part of the tunable circuit, the source impedance and the load impedance R_0 are 73 Ohm, which coincides with the resistance part of the dipole as in [15], [16].

Comment 4: *The upper limits of the sums in equation (34) are incorrect. This problem occurs several times. The authors should check them carefully.*

Response: We are thankful for your kind comments. We also find that the upper limits of the EHUs and IDUs are confused. For resolving it, we have checked the entire manuscript and corrected them as follows.

Section II. (page 6):

$$P_t\left(\mathbf{W}_j, \mathbf{V}_l\right) = \sum_{j=1}^{M_I} \mathrm{Tr}\left(\mathbf{W}_j\right) + \sum_{l=1}^{M_E} \mathrm{Tr}(\mathbf{V}_l). \tag{34}$$

Section IV. (page 10):

$$\sum_{j=1}^{M_I} \mathrm{Tr}\left(\mathbf{W}_j^{(\ell)}\right) + \sum_{l=1}^{M_E} \mathrm{Tr}(\mathbf{V}_l^{(\ell)}) \leq P_{\mathrm{Max}}, \tag{64c}$$

$$P\left(\mathbf{W}_j^{(\ell)}, \mathbf{V}_l^{(\ell)}\right) = \sum_{j=1}^{M_I} \mathrm{Tr}\left(\mathbf{W}_j^{(\ell)}\right) + \sum_{l=1}^{M_E} \mathrm{Tr}(\mathbf{V}_l^{(\ell)}) + P_c + N_i P_I. \tag{66}$$

Comment 5: *There is a missing end-term constraint of equation (35e) of problem* $\mathrm{P}_{\mathrm{NC}-1}$. *Please add it, otherwise, the problem formulation is incomplete.*

Response: Thank you for your kind suggestion. Compared to the coupling-awareness case, the $\mathrm{P}_{\mathrm{NC}-1}$ indeed misses an important constraint. It needs to be mentioned that it is a mistake in writing but not in coding, so the simulation results are still unrevised. The modified constraint of $\mathrm{P}_{\mathrm{NC}-1}$ is presented below.

Section II. (page 6):

$$\hat{\mathbf{F}}_{q,q} = \frac{1}{4r_q^2}, \forall q, \hat{\mathbf{F}}_{q+1,q+1} = 1 \tag{35e}$$

Comment 6: *For problem* $\mathrm{P}_{\mathrm{NC}-4}$, *the authors make the* $\epsilon(i)$ *as an optimization variable. However, the* $\epsilon(i)$ *is calculated by the given beamforming vectors in Algorithm 1. The authors should explain the contradiction.*

Response: Thank you for bringing this problem to our attention. Indeed, as the thoughts of the reviewer, the $\epsilon(i)$ is given by Algorithm 1. We wrote it in the optimization variable by mistake. We have corrected these mistakes in the manuscript.

Section III. (page 7):

$$\mathbf{P}_{\mathrm{NC}-4}: \max_{\{\mathbf{W}_j(i),\mathbf{V}_l(i)\}} \sum_{j=1}^{M_I} R_j\left(\mathbf{W}_j(i)\right) - \epsilon(i) P\left(\mathbf{W}_j(i), \mathbf{V}_l(i)\right) \tag{38}$$

$$\mathbf{P}_{\mathrm{NC}-4}: \max_{\left\{\begin{smallmatrix}\mathbf{W}_j(i),\mathbf{V}_l(i),\\ \tau_j(i)\end{smallmatrix}\right\}} \Xi(\tau_j(i), \mathbf{W}_j(i), \mathbf{V}_l(i))$$

$$\Xi(\tau_j(i), \mathbf{W}_j(i), \mathbf{V}_l(i)) = \sum_{j=1}^{M_I} \tau_j(i) - \epsilon(i) P\left(\mathbf{W}_j(i), \mathbf{V}_l(i)\right). \tag{40}$$

$$\mathbf{P}_{\mathrm{NC}-5}: \max_{\left\{\begin{smallmatrix}\mathbf{W}_j(i),\mathbf{V}_l(i),\tau_j(i),\\ a_j(i),b_j(i)\end{smallmatrix}\right\}} \Xi(\tau_j(i), \mathbf{W}_j(i), \mathbf{V}_l(i))$$

Comment 7: *In addition to the above issues, several typos should be addressed properly. This is not a full list, but presents some representative places that could be improved.*

(a) *Page 6, line 17, "*P_0*" −> "*$\mathrm{P}_{\mathrm{NC}-0}$*";*

(b) *Page 9, line 7, "adjust"−> "adjusting";*

(c) *Page 9, line 29, "(63d)" −> "(63e)";*

(d) *Page 10, line 43, "location"−> "central location";*

(e) *Page 10, line 46, "region"−> "regions"*

Response: All authors are grateful for your constructive suggestion. Indeed, the proposed problems by the reviewer are prominent, so we have modified them in the manuscript as follows. It should be mentioned that we have inspected the whole manuscript and also found other mistakes. The mistakes have been corrected carefully and presented in the revised version. Due to the limited space here, please check the resubmitted manuscript.

Section III. (page 6):

Even though the structure of the channel model (30) has been further reduced and is similar to the cascaded channel model based on the independent diffusive scatterer (IDS) assumption in [6], the problem $\mathbf{P}_{\mathrm{NC}-0}$ remains a non-convex problem that is difficult to solve.

Section IV. (page 9):

where $\mathbf{G}^{(\ell)} = j\,\mathrm{Im}\left(\mathbf{\Omega}^{*(\ell)}\right)$ and $\mathbf{T}^{(0)} = \boldsymbol{\mathcal{Z}}_{\mathrm{II}} + R_0$ for fixing the resistances and adjusting the reactances of the circuits.

Section IV. (page 9):

$$\mathbf{P}_{\mathrm{CA}-1}^{(\ell)}: \max_{\left\{\mathbf{w}_j^{(\ell)}, \mathbf{v}_l^{(\ell)}, \mathbf{G}^{(\ell)}\right\}} \mathrm{EE}\left(\mathbf{w}_j^{(\ell)}, \mathbf{v}_l^{(\ell)}, \mathbf{G}^{(\ell)}\right)$$

s.t. (52a),

$$R_j\left(\mathbf{w}_j^{(\ell)}, \mathbf{G}^{(\ell)}\right) \geq R_j^{(\mathrm{D})}, \forall j, \tag{63a}$$

$$P_{\mathrm{RF},n}\left(\mathbf{w}_j^{(\ell)}, \mathbf{v}_l^{(\ell)}, \mathbf{G}^{(\ell)}\right) \geq \mathcal{G}_n(E_n^{(\mathrm{D})}), \forall n, \tag{63b}$$

$$\mathbf{G}_{q,q}^{(\ell)} = j\Delta_q^{(\ell)}, \forall q, \tag{63c}$$

$$\Delta_q^{(\ell)} \leqslant \zeta, \forall q. \tag{63d}$$

(63c) is adopted to ensure only the imaginary part of the tunable circuit impedance is adjustable as the same setting in problem $\mathbf{P}_0$. In addition, (63d) is the guarantee of the Neuman approximation's accuracy. It is worth noting that the choices of ζ need to be careful.

Section V. (page 10):

The center location of the 8-element-array BS is located at (0,0) and the RIS with 2×8 array configuration is located at (1,1).

Section V. (page 10):

Two EHUs and two IDUs are generated at the rectangular regions of (x_{IDU},y_{IDU}) and (x_{EHU},y_{EHU}) with $x_{\mathrm{EHU}} \in [0.5, 2]$ $y_{\mathrm{EHU}} \in [-1, -2]$, $x_{\mathrm{IDU}} \in [4, 15]$ and $y_{\mathrm{IDU}} \in [-5, -24]$.