Beamforming Design with Bilevel Optimization for RIS-Assisted SWIPT Systems

Yinghui He, Yunlong Cai, and Guanding Yu
College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China
E-mail: {2014hyh, ylcai, yuguanding}@zju.edu.cn

Abstract—**Simultaneous wireless information and power transfer (SWIPT) has been regarded as one of the most promising technologies for the Internet of Things (IoT). However, in some scenarios where one function takes priority over the other one, existing single-level optimization models do not apply to formulate this class of problems. To address it, we employ bilevel optimization in this paper and consider a reconfigurable intelligent surface (RIS)-assisted SWIPT system. Specifically, the harvested power is maximized in the lower-level problem and the sum rate is maximized in the upper-level problem. We first derive the optimal set of the lower-level problem and then transform the bilevel problem into an equivalent but more tractable form. Subsequently, a block-wise bilevel-based beamforming design (B^3D) algorithm is developed, which converges to the stationary point of the bilevel problem. Finally, simulation results are presented to verify the performance of the proposed algorithm. Compared to the existing optimization schemes, the proposed algorithm achieves a higher sum rate with the same harvested power.**

I. Introduction

Leveraging the advantages of low cost and the abilities of wireless transmission, sensing, and computation, massive Internet of Things (IoT) devices have been extensively deployed for a wide range of applications across various fields [1]. However, the sustained growth of IoT devices is hindered by the energy supply bottleneck. Traditional energy supply methods relying on batteries or wired connections limit the application scenarios. To address this issue, energy harvesting (EH) technology has emerged as a solution, enabling the wireless supply of power to IoT devices through radio frequency (RF) signals. Meanwhile, RF signals inherently possess the function of information delivery. Consequently, simultaneous wireless information and power transfer (SWIPT) has been proposed and attracted a great deal of attention as it can use a single signal to accomplish two objectives: power transfer and information delivery [2], [3]. While existing research focused on exploring the tradeoff between these two objectives, three primary schemes have emerged. The first one aims to maximize the performance of the information delivery under the requirement of the power transfer [4], while the second one pursues the maximization of the harvested power [5], [6]. The third one directly maximizes the tradeoff between the two objectives through appropriate weighting [7].

However, in some scenarios, the priorities of the two objectives might be in order, with one objective taking precedence over the other. For instance, in situations where a majority of IoT devices are grappling with power shortages, the base station (BS) must prioritize maximizing harvested power before optimizing the information delivery rate. Achieving such a goal poses significant challenges with existing optimization methods. For example, when employing the weighting method, if the weight assigned to one objective significantly outweighs the other, the performance of the former objective can be maximized while that of the latter one is hard to be guaranteed. To address this challenge, another burgeoning optimization method, namely bilevel optimization [8], [9], has gained prominence. In the bilevel problem, there is an upper-level problem (i.e., the leader) and a lower-level problem (i.e., the follower), where the optimization variables are coupled over both levels. The optimal set of the lower-level problem determines the feasible set of the upper-level problem. Motivated by this, our paper aims to utilize bilevel optimization for cases where one function holds higher priority. Additionally, to enhance data rates, a reconfigurable intelligent surface (RIS) is deployed near information decoding (ID) users, allowing for channel reconfiguration [11].

In the formulated bilevel problem, the lower-level problem strives to maximize the harvested power of EH users and the upper-level problem seeks to maximize the sum rate of ID users. Due to the non-convexity of the lower-level problem, the existing bilevel optimization methods (such as first-order penalty methods [10]) cannot be applied. Instead, we first solve the lower-level problem, deriving necessary and sufficient conditions for the optimal solution. After that, we develop a block-wise bilevel-based beamforming design (B^3D) algorithm by solving the upper-level problem and show that the proposed algorithm converges to a Karush-Kuhn-Tucker (KKT) (stationary) point of the original bilevel problem. Notably, our solving process reveals that the lower-level problem determines the direction of the beamforming and the upper-level problem decides the power allocation among beamforming vectors for ID users. Finally, through comparisons with mainstream schemes, we demonstrate that our proposed algorithm achieves a higher sum rate under the same harvested power.

The paper is structured as follows. Section II briefly describes the system model and formulates the bilevel problem of interest. In Section III, after deriving the optimal set to the lower-level problem, we develop a beamforming design algorithm for the bilevel problem and analyze its convergence properties and complexity. Simulation results are presented in

Section IV, and Section V concludes the whole paper.

Notations: In this paper, scalars are denoted in lowercase, vectors are denoted in boldface lowercase, and matrices are denoted in boldface uppercase. $\boldsymbol{I}$ represents an identity matrix. $(\cdot)^*$, $(\cdot)^T$, and $(\cdot)^H$ denote complex conjugate, transpose, and Hermitian transpose, respectively.

II. System Model and Problem Formulation

In this section, we first introduce the model of the RIS-assisted SWIPT system, and then formulate the beamforming design as a bilevel optimization problem.

A. System Model

In this paper, we consider an RIS-assisted SWIPT system, comprising a BS, J single-antenna ID users, and K single-antenna EH users, as shown in Fig. 1. The BS is equipped with M transmit antennas, and it aims to convey information to ID users and transfer power to EH users, simultaneously. Meanwhile, one RIS is deployed near ID users, enhancing the communication signals from the BS to ID users. The RIS is equipped with N reflecting elements with phase being ϕ_n. Since the RIS is far from the BS and EH users are near the BS, the effect of the RIS to EH users can be neglected.

For the j-th ID user, let s_j denote the transmit symbol with zero mean and unit variance, i.e., $\mathbb{E}\left\{|s_j|^2\right\} = 1$, and $\boldsymbol{w}_j \in \mathbb{C}^{M\times 1}$ be the corresponding beamforming vector at the BS. To transfer enough power towards K EH users, the BS randomly generates an artificial redundant signal vector, denoted by $\boldsymbol{v} \in \mathbb{C}^{M\times 1}$ with mean being zero and variance being $\boldsymbol{V} \in \mathbb{C}^{M\times M}$, i.e., $\mathbb{E}\left\{\boldsymbol{v}\boldsymbol{v}^H\right\} = \boldsymbol{V}$. Then, the total transmit signal at the BS can be expressed as

$$\boldsymbol{x} = \sum\nolimits_{j=1}^{J} \boldsymbol{w}_j s_j + \boldsymbol{v}. \tag{1}$$

Then, the transmit power of the BS is given by

$$P^{\text{BS}}(\boldsymbol{w}_j, \boldsymbol{V}) = \mathbb{E}\left\{\boldsymbol{x}^H\boldsymbol{x}\right\} = \sum\nolimits_{j=1}^{J} \boldsymbol{w}_j^H\boldsymbol{w}_j + \text{tr}\left\{\boldsymbol{V}\right\}. \tag{2}$$

Let $\boldsymbol{g}_k \in \mathbb{C}^{M\times 1}$ denote the channel between the BS and the k-th EH user. Then, the received signal at the k-th EH is

$$y_k^{\text{EH}} = \boldsymbol{g}_k^H\Big(\sum\nolimits_{j=1}^{J} \boldsymbol{w}_j s_j + \boldsymbol{v}\Big) + n_k^{\text{EH}}, \tag{3}$$

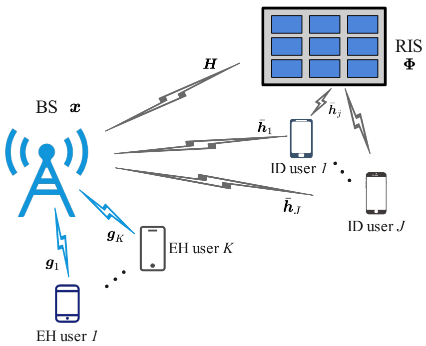

Fig. 1. RIS-assisted SWIPT system.

where n_k^{EH} is the complex circular Gaussian noise with mean being zero and variance being σ^2. Then, the harvested power at the k-th EH user is

$$\begin{aligned} P_k^{\text{EH}}(\boldsymbol{w}_j, \boldsymbol{V}) &= \eta\mathbb{E}\left\{\left(y_k^{\text{EH}}\right)^H y_k^{\text{EH}}\right\} &(4)\\ &= \eta\left(\boldsymbol{g}_k^H\left(\sum_{j=1}^{J}\boldsymbol{w}_j\boldsymbol{w}_j^H + \boldsymbol{V}\right)\boldsymbol{g}_k + \sigma^2\right), &(5)\end{aligned}$$

where η is the energy conversion efficiency of the EH user.

There are two links between the BS and each ID user, that is, the "BS–ID user" link and the "BS–RIS–ID user" link. For the former link, the channel vector is denoted by $\boldsymbol{h}_j \in \mathbb{C}^{M\times 1}$. For the latter link, $\boldsymbol{H} \in \mathbb{C}^{M\times N}$ denotes the channel matrix between the BS and the RIS, and $\bar{\boldsymbol{h}}_j \in \mathbb{C}^{N\times 1}$ denotes the channel vector between the RIS and the j-th ID user. Then, the equivalent channel of the "BS–RIS–ID user" link is $\bar{\boldsymbol{h}}_j^H\boldsymbol{\Phi}\boldsymbol{H}$, where $\boldsymbol{\Phi} = \text{Diag}\left(e^{j\phi_1}, \cdots, e^{j\phi_N}\right) \in \mathbb{C}^{N\times N}$ is the diagonal passive beamforming matrix at RIS. By defining $\bar{\boldsymbol{H}}_j = \text{Diag}\{\bar{\boldsymbol{h}}_j\}^H\boldsymbol{H} \in \mathbb{C}^{N\times M}$, it can be rewritten as $\boldsymbol{\phi}^T\bar{\boldsymbol{H}}_j$ with $\boldsymbol{\phi} = \left[e^{j\phi_1}, \cdots, e^{j\phi_N}\right]^T \in \mathbb{C}^{N\times 1}$. Based on the aforementioned analysis, the received signal at the j-th ID user is given by

$$y_j^{\text{ID}} = \left(\boldsymbol{h}_j^H + \boldsymbol{\phi}^T\bar{\boldsymbol{H}}_j\right)\left(\boldsymbol{v} + \sum\nolimits_{j=1}^{J}\boldsymbol{w}_j s_j\right) + n_j^{\text{ID}}, \tag{6}$$

where n_j^{ID} is the complex circular Gaussian noise with mean being zero and variance being σ^2. Then, the signal-to-interference-plus-noise ratio (SINR) of j-th ID user is

$$\begin{aligned}&\text{SINR}_j^{\text{ID}}(\boldsymbol{w}_j, \boldsymbol{V})\\ &= \frac{\left(\boldsymbol{h}_j^H + \boldsymbol{\phi}^T\bar{\boldsymbol{H}}_j\right)\boldsymbol{w}_j\boldsymbol{w}_j^H\left(\boldsymbol{h}_j^H + \boldsymbol{\phi}^T\bar{\boldsymbol{H}}_j\right)^H}{\left(\boldsymbol{h}_j^H + \boldsymbol{\phi}^T\bar{\boldsymbol{H}}_j\right)\left(\boldsymbol{V} + \sum\limits_{i\neq j}\boldsymbol{w}_i\boldsymbol{w}_i^H\right)\left(\boldsymbol{h}_j^H + \boldsymbol{\phi}^T\bar{\boldsymbol{H}}_j\right)^H + \sigma^2}. \end{aligned}\tag{7}$$

B. Problem Formulation

In this paper, the BS has two objectives, i.e., power transfer for EH users and information delivery for ID users. The former objective holds precedence and the latter one is secondary. Thus, our objective is to maximize the sum rate of ID users following the maximization of the total harvested power of EH users. Meanwhile, to further improve the data rate of ID users, we also optimize the phase of each reflecting element at the RIS. Therefore, this problem can be formulated as the following constrained bilevel programming:

$$\max_{\{\boldsymbol{w}_j, \boldsymbol{V}, \boldsymbol{\phi}\}} \sum_{j=1}^{J}\log_2\left(1 + \text{SINR}_j^{\text{ID}}(\boldsymbol{w}_j, \boldsymbol{V})\right), \tag{8a}$$

$$\text{s.t.} \quad \{\boldsymbol{w}_j, \boldsymbol{V}\} \in \arg\max_{\{\boldsymbol{w}_j', \boldsymbol{V}'\}}\Bigg\{\sum_{k=1}^{K} P_k^{\text{EH}}(\boldsymbol{w}_j', \boldsymbol{V}') \tag{8b}$$

$$\Big| P^{\text{BS}}(\boldsymbol{w}_j', \boldsymbol{V}') \le P^{\max}\Bigg\}, \tag{8c}$$

$$|\boldsymbol{\phi}(n)| = 1, \ \forall n. \tag{8d}$$

Here, the lower-level problem aims to maximize the total harvested power (8b) under the power limitation (8c) of the BS with $P^{\max}$ being the budget. Within the set of the optimal solution to the lower-level problem, the upper-level problem aims to maximize the sum rate of ID users under uni-modulus constraint (8d) on all elements of the RIS passive beamforming vector. Different from the conventional single-level problem, we need to solve the lower-level problem and obtain the corresponding set of optimal solutions, rather than only one optimal solution. Meanwhile, existing methods for bilevel optimization cannot be directly applied since they are designed for the case where the lower-level problem is convex [10] while the lower-level problem of this paper is non-convex.

III. Beamforming Design Algorithm

In this section, we first solve the lower-level problem and then solve the upper-level one. Finally, a block-wise bilevel-based beamforming design ($\mathrm{B}^3\mathrm{D}$) algorithm is developed.

A. Optimal Set to Lower-Level Problem

It is easy to show that the objective function of the lower-level problem is convex, but the problem is not convex since it aims to maximize a convex function. Hence, we cannot directly solve it with existing convex optimization methods. To address it, we will convert the problem into a more easily tractable yet equivalent form. Specifically, we can define $\boldsymbol{G} = \sum_{k=1}^{K} \boldsymbol{g}_k \boldsymbol{g}_k^H$ and rewrite $\boldsymbol{V}$ as the linear combination of M vectors, i.e., $\boldsymbol{V} = \sum_{m=1}^{M} \boldsymbol{f}_m \boldsymbol{f}_m^H$. As a result, the lower-level problem can be rewritten as

$$\max_{\{\boldsymbol{w}_j, \boldsymbol{f}_m\}} \quad \eta \left(\sum_{j=1}^{J} \boldsymbol{w}_j^H \boldsymbol{G} \boldsymbol{w}_j + \sum_{m=1}^{M} \boldsymbol{f}_m^H \boldsymbol{G} \boldsymbol{w}_m + K\sigma^2 \right), \tag{9a}$$

$$\text{s.t.} \quad \sum_{j=1}^{J} \boldsymbol{w}_j^H \boldsymbol{w}_j + \sum_{m=1}^{M} \boldsymbol{f}_m^H \boldsymbol{f}_m \leq P^{\max}. \tag{9b}$$

It is similar to a Rayleigh quotient maximization problem, and thus the optimal set can be given in the following Theorem.

Theorem 1: The necessary and sufficient condition of the optimal solution to the lower-level problem is

$$\boldsymbol{w}_j = \alpha_j \bar{\boldsymbol{g}}_1, \ \boldsymbol{V} = |\beta|^2 \bar{\boldsymbol{g}}_1 \bar{\boldsymbol{g}}_1^H, \tag{10}$$

$$\sum_{j=1}^{J} |\alpha_j|^2 + |\beta|^2 = P^{\max}, \ \alpha_j \in \mathbb{C}, \ \beta \in \mathbb{C}, \tag{11}$$

where $\bar{\boldsymbol{g}}_1$ is the eigenvector of the largest eigenvalue for $\boldsymbol{G}$.

From Theorem 1, we can observe that, to maximize the sum of harvested power, all beamforming vectors for ID users and artificial redundant signal matrix for EH users are directly determined by the eigenvector of the maximum eigenvalue of $\boldsymbol{G}$, i.e., $\bar{\boldsymbol{g}}_1$. This result meets our intuitive perception since $\bar{\boldsymbol{g}}_1$ is the most efficient beamforming direction and all transmit power should be focused on this direction for maximizing the summation of harvested power in the lower-level problem. Meanwhile, the power allocation among beamforming vectors and the artificial redundant signal matrix is not given and only needs to satisfy the condition (11).

Until now, we can derive the closed-form of the optimal solution set of the lower-level problem, as shown in Theorem 1. In the next part, we will apply the obtained optimal set to solve the upper-level problem.

B. Proposed Algorithm to Bilevel Problem

Substituting the derived optimal set into the lower-level problem, we can get the following reformulated constrained optimization problem

$$\max_{\{\alpha_j, \beta, \boldsymbol{\phi}\}} \sum_{j=1}^{J} \log_2 \left(1 + \mathrm{SINR}_j^{\mathrm{ID}}(\alpha_j \bar{\boldsymbol{g}}_1, |\beta|^2 \bar{\boldsymbol{g}}_1 \bar{\boldsymbol{g}}_1^H)\right), \tag{12a}$$

$$\text{s.t.} \quad \sum_{j=1}^{J} |\alpha_j|^2 + |\beta|^2 = P^{\max}, \tag{12b}$$

$$|\boldsymbol{\phi}(n)| = 1, \ \forall n. \tag{12c}$$

Before solving it, we have the following Lemma for β.

Lemma 1: To maximize the sum rate of ID uses, β should be zero.

Proof: From the expression of SINR, we can find that SINR decreases with $|\beta|^2$. Therefore, the sum rate also decreases with $|\beta|^2$ and β should be zero for maximizing the sum rate, which ends the proof. ■

Next, we can adopt the weighted sum mean square error (WMMSE) algorithm [12] to transfer the objective function into an equivalent form. Specifically, we introduce auxiliary variables μ_j, $\omega_j > 0$, $\forall j$, which stand for the receiver gain and weighting coefficient of ID user j. Let e_j denote the mean square error (MSE) of ID user j, which is given by

$$\begin{aligned} e_j &= \mathbb{E}\left\{(\mu_j^H y_j^{\mathrm{ID}} - s_j)(\mu_j^H y_j^{\mathrm{ID}} - s_j)^H\right\} \\ &= -2\mathcal{R}\left\{\mu_j^H \left(\boldsymbol{h}_j^H + \boldsymbol{\phi}^T \bar{\boldsymbol{H}}_j\right) \alpha_j \bar{\boldsymbol{g}}_1\right\} + |\mu_j|^2\sigma^2 \\ &\quad + |\mu_j \alpha_j|^2 \sum_{i=1}^{J} \left| \left(\boldsymbol{h}_j^H + \boldsymbol{\phi}^T \bar{\boldsymbol{H}}_j\right) \bar{\boldsymbol{g}}_1 \right|^2 + 1. \end{aligned} \tag{13}$$

Then, the constrained optimization problem (12) is equivalent to the following WMMSE problem:

$$\min_{\{\alpha_j, \boldsymbol{\phi}, \mu_j, \omega_j\}} \sum_{j=1}^{J} \omega_j e_j - \log \omega_j, \tag{14a}$$

$$\text{s.t.} \quad \sum_{j=1}^{J} |\alpha_j|^2 = P^{\max}, \tag{14b}$$

$$|\boldsymbol{\phi}(n)| = 1, \ \forall n. \tag{14c}$$

Based on the block structure of this problem, we can partition the variables into four blocks, α_j, μ_j, ω_j, and $\boldsymbol{\phi}$ and apply the block coordinate descent (BCD) method to solve it. To be more specific, the blocks of the variables are updated successively by solving each corresponding subproblems. In the following, we will show the closed-form solutions of each block.

Block-ω_j. We optimize ω_j, $\forall j$ in parallel by fixing the other variables. In this case, the corresponding subproblem is given by

$$\min_{\omega_j} \quad \omega_j e_j - \log \omega_j. \tag{15}$$

Based on the first order optimality condition, the optimal solution to problem (15) is given by

$$\omega_j^\star = 1/e_j, \ \forall j. \tag{16}$$

Block-μ_j. Similarly, we optimize μ_j, $\forall j$ in parallel by fixing the other variables, which ends up with the following subproblem

$$\begin{aligned}\min_{\mu_j} \quad & -2\mathcal{R}\left\{\mu_j^H\left(\boldsymbol{h}_j^H + \boldsymbol{\phi}^T\bar{\boldsymbol{H}}_j\right)\alpha_j\bar{\boldsymbol{g}}_1\right\} + |\mu_j|^2\sigma^2 \\ & + |\mu_j\alpha_j|^2\sum_{i=1}^{J}\left|\left(\boldsymbol{h}_j^H + \boldsymbol{\phi}^T\bar{\boldsymbol{H}}_j\right)\bar{\boldsymbol{g}}_1\right|^2. \end{aligned} \tag{17}$$

Therefore, the optimal solution to this problem is given by

$$\mu_j^\star = \frac{\alpha_j\left(\boldsymbol{h}_j^H + \boldsymbol{\phi}^T\bar{\boldsymbol{H}}_j\right)\boldsymbol{w}_j}{|\alpha_j|^2\sum_{i=1}^{J}\left|\left(\boldsymbol{h}_j^H + \boldsymbol{\phi}^T\bar{\boldsymbol{H}}_j\right)\bar{\boldsymbol{g}}_1\right|^2 + \sigma^2}. \tag{18}$$

Block-α_j. We optimize α_j by fixing the other variables. The subproblem can be given as,

$$\min_{\{\alpha_j\}} \ -2\omega_j\mathcal{R}\left\{\mu_j^H\left(\boldsymbol{h}_j^H + \boldsymbol{\phi}^T\bar{\boldsymbol{H}}_j\right)\alpha_j\bar{\boldsymbol{g}}_1\right\}$$

$$+\omega_j|\mu_j\alpha_j|^2\sum_{i=1}^{J}\left|\left(\boldsymbol{h}_j^H + \boldsymbol{\phi}^T\bar{\boldsymbol{H}}_j\right)\bar{\boldsymbol{g}}_1\right|^2, \tag{19a}$$

$$\text{s.t.} \quad \sum_{j=1}^{J}|\alpha_j|^2 = P^{\max}, \tag{19b}$$

which is not convex due to the constraint. To address it, constraint (19b) can be relaxed into: $\sum_{j=1}^{J}|\alpha_j|^2 \le P^{\max}$. Since the objective function decreases with $|\alpha|$, it can be shown that the optimal solution after the relaxation satisfies the constraint (19b). Thus, the optimal solution is

$$\alpha_j^\star = \frac{\left(\omega_j\mu_j^*\left(\boldsymbol{h}_j^H + \boldsymbol{\phi}^T\bar{\boldsymbol{H}}_j\right)\bar{\boldsymbol{g}}_1\right)^*}{\sum_{j=1}^{J}\omega_j\left|\mu_j^*\left(\boldsymbol{h}_j^H + \boldsymbol{\phi}^T\bar{\boldsymbol{H}}_j\right)\bar{\boldsymbol{g}}_1\right|^2 + \lambda}, \tag{20}$$

where $\lambda \ge 0$ is the optimal Lagrange multiplier for constraint (19b). Here, scalar λ can be obtained via the search algorithm until constraint (19b) is satisfied.

Block-$\boldsymbol{\phi}(n)$. By defining $\boldsymbol{x} = \sum_{j=1}^{J}\omega_j\mu_j\bar{\boldsymbol{H}}_j\alpha_j\bar{\boldsymbol{g}}_1$, $\boldsymbol{y} = \sum_{j=1}\omega_j|\mu_j|^2\bar{\boldsymbol{H}}_j\boldsymbol{B}\boldsymbol{h}_j$, $\boldsymbol{A} = \sum_{j=1}\omega_j|\mu_j|^2\bar{\boldsymbol{H}}_j\boldsymbol{B}\bar{\boldsymbol{H}}_j^H$, $\boldsymbol{B} = \sum_{j=1}^{J}\boldsymbol{w}_j\boldsymbol{w}_j^H$, the subproblem with respect to $\boldsymbol{\phi}(n)$ becomes

$$\min_{\boldsymbol{\phi}(n)} \quad 2\mathcal{R}\left\{\boldsymbol{\phi}(n)\left(\boldsymbol{y}(n) - \boldsymbol{x}(n)\right)\right\}$$

$$+ 2\mathcal{R}\left\{\boldsymbol{\phi}(n)\sum_{m\neq n}\boldsymbol{A}(n,m)\boldsymbol{\Phi}(m)^*\right\}. \tag{21a}$$

$$\text{s.t.} \quad |\boldsymbol{\phi}(n)| = 1. \tag{21b}$$

Algorithm 1: B^{3}D Algorithm for Bilevel Problem (8).

Initialize variables α_j, μ_j, $\omega_j, \forall j$, and $\boldsymbol{\phi}$. Set the tolerance of accuracy $\epsilon > 0$ and maximum number of iterations $I^{\max}$ and the iteration number $i = 0$;
repeat
  Update $\omega_j, \forall j$ by (16);
  Update $\mu_j, \forall j$ by (18);
  Update $\alpha_j, \forall j$ by (20);
  Update $\boldsymbol{\phi}$ by (22);
until *the gap between consecutive values of the objective function is under ϵ or $i > I^{\max}$.*

Thus, the optimal solution is

$$\boldsymbol{\phi}(n)^\star = \exp\left(i\pi - i\angle\left(\boldsymbol{y}(n) - \boldsymbol{x}(n) + \sum_{m\neq n}\boldsymbol{A}(n,m)\boldsymbol{\phi}(m)^*\right)\right). \tag{22}$$

So far, we have obtained closed-form solutions for these four subproblems. The proposed algorithm is summarized as Algorithm 1. In each iteration, the above four steps are implemented sequentially.

C. Analysis and Discussion

Until now, we have proposed the B^{3}D algorithm for bilevel problem (8) as shown in Algorithm 1. To develop this algorithm, we first obtain the necessary and sufficient conditions of the optimal solution to the lower-level problem as shown in Theorem 1 and then equivalently transfer the bilevel problem to problem (12). Furthermore, problem (12) is transferred to problem (14) using the WMMSE method. Note that they share the same set of global optimal solutions and KKT conditions. After that, the BCD method is adopted to solve problem (14), ensuring convergence of the iterates generated by our algorithm to a stationary point of problem (14). Based on the above discussion, it can be concluded that the solution obtained by Algorithm 1 is a stationary point of bilevel problem (8). Furthermore, from the above process, it can be seen that the direction of beamforming vectors is determined by the lower-level problem, and the power allocation among beamforming vectors is decided by the upper-level problem.

Next, we can analyze the computational complexity of the B^{3}D algorithm. Specifically, in each iteration, the computational complexities of four steps are $\mathcal{O}(1)$, $\mathcal{O}(JMN)$, $\mathcal{O}(JMN + \log\frac{1}{\epsilon})$, and $\mathcal{O}(JMN)$, respectively. Here, ϵ is the tolerance of accuracy. Meanwhile, the maximum number of iterations is $I^{\max}$. Therefore, the overall computational complexity of the B^{3}D algorithm is

$$\mathcal{O}\left(I^{\max}\left(3JMN + \log\frac{1}{\epsilon}\right)\right). \tag{23}$$

Our proposed algorithm offers two advantages in comparison to existing algorithms for bilevel problems, e.g., [9], [10]:

• The existing algorithms mainly address the specific bilevel problem with the lower problem being convex or strongly convex. Unfortunately, the lower problem in the bilevel problem of interest is non-convex and there is a lack of algorithms tailored for such a bilevel problem. Our proposed B^3D algorithm effectively solves this bilevel problem, obtaining a first-order stationary point.

• To deal with the bilevel problem, the existing algorithms tend to either use double-loop penalty methods [10], where the outer loop updates parameters for penalty and the inner loop solves the problem under given the penalty parameter or compute the inverse of the Hessian matrix of the lower-level problem to get the hyper-gradient of the upper-level objective function [9]. As a result, those algorithms have a high computational complexity in general. In contrast, the B^3D algorithm has a much lower computational complexity as it requires only a single-loop structure.

IV. Simulation Results

In this section, we aim to evaluate the performance of the proposed beamforming design algorithm for the RIS-assisted SWIPT system.

A. Simulation Setup

The central points of the BS and the RIS are located at $(0, 0)$ and $(200, 100)$, in meter (m), respectively. The number of transmit antennas at the BS is 12, i.e., $M = 12$, and the number of reflecting elements at the RIS is 50, i.e., $N = 50$. The number of EH users is set as 8, and they are randomly located within the radius of 5m from the BS. The energy conversion efficiency is set as 0.5. Meanwhile, the number of ID users also is 8, and they are randomly located within the radius of 200m from the coordinate $(150, 150)$. For the large-scale fading of the wireless channel, it is modeled as

$$L(d) = 30 + 36\log(d), \tag{24}$$

where d is the individual link distance in the meter. For the small-scale fading, it follows Rician fading and is modeled as

$$\boldsymbol{H}^{\mathrm{s}} = \sqrt{\frac{\varepsilon}{\varepsilon+1}}\boldsymbol{a}_r(\theta^r)\boldsymbol{a}_t(\theta^t)^H + \sqrt{\frac{1}{\varepsilon+1}}\boldsymbol{H}_0, \tag{25}$$

where ε is the Rician factor, $\boldsymbol{H}_0$ is the non-line-of-sight component whose entries follow the distribution $\mathcal{CN}(0, 1)$, $\boldsymbol{a}_t(\theta^t)$ and $\boldsymbol{a}_r(\theta^r)$ are the transmit and receive array response vectors with θ^t and θ^r being the azimuth angles of departure and arrival, respectively. The universal expression for array response vector is $\boldsymbol{a}(\theta) = M^{-1/2}[1, e^{i\frac{2\pi\Delta}{\lambda}\sin(\theta)}, \cdots, e^{i\frac{2\pi\Delta}{\lambda}(M-1)\sin(\theta)}]$, where λ is the wavelength and Δ is the antenna spacing. Δ is set as $\lambda/2$. For the RIS-related channel, i.e., $\boldsymbol{H}$ and $\bar{\boldsymbol{h}}$, ε is set as 3dB. For the remained channel, i.e., $\boldsymbol{h}_j$ and $\boldsymbol{g}_k$, ε is set as 9dB. Moreover, the noise power is 10^{-9}W and the transmit power limitation of the BS is set as 40dBm.

B. Performance Verification

To better show the performance of the proposed algorithm, we adopt two benchmark schemes as

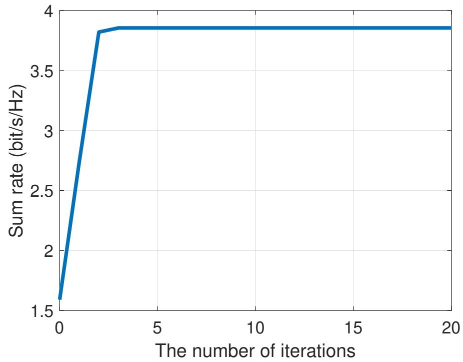

Fig. 2. Convergence behavior of the proposed algorithm.

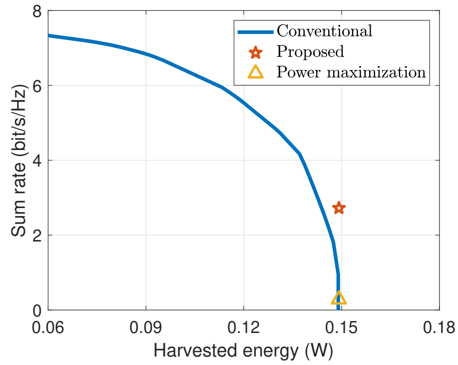

Fig. 3. Performance comparison among three schemes.

- **Conventional scheme.** It aims to optimize a single-level problem. Specifically, the sum rate is considered as the objective function with the requirement of harvested power being P^{req}. For such a problem, due to the constraint of the power requirement being non-convex, we apply the penalty dual decomposition (PDD)-based method as the baseline for finding the first-order stationary point of non-convex problems [13].
- **Power maximization scheme.** It only considers maximizing the sum of harvested power for EH users without considering the sum rate.

Fig. 2 shows the convergence behavior of the proposed algorithm, that is, the evolution of the sum rate versus the iteration number. Notably, the proposed algorithm demonstrates rapid convergence, achieving its optimal result within a few iterations. Additionally, we also compare the running time of the proposed algorithm with that of two benchmark schemes, as shown in Tab. I. Note that all schemes are tested on a desktop Intel (i5-13600K) CPU with 16 GB RAM and the tolerance of accuracy is set as 5×10^{-3}. From the table, we can see that both the proposed algorithm and energy maximization scheme exhibit minimal running times, while the running time of the conventional scheme is considerably higher. This discrepancy arises because the PDD-based method for the conventional scheme involves double loops, whereas our proposed algorithm operates with a single loop, offering closed-form solutions for each subproblem. Meanwhile, when the required harvested power in the conventional scheme is the same as the maximized harvested power of the proposed

TABLE I
PERFORMANCE COMPARISON AMONG THREE SCHEMES.

Scheme	Proposed	Conventional	Power Maximization
Running time (s)	0.126	488	2.87×10^{-4}
Sum rate (bit/s/Hz)	5.1015	1.6032	0.6586
Maximized/Required harvested power (W)	0.6948	0.6948	0.6948

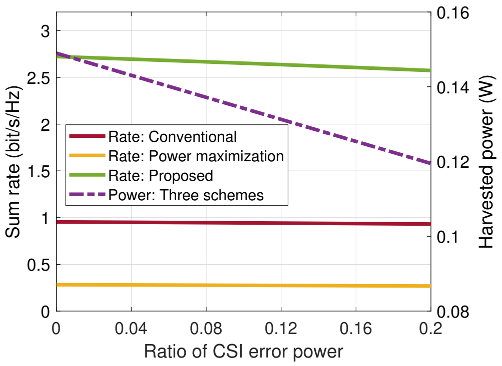

Fig. 4. Effect of CSI estimation error.

algorithm, the sum rate of the latter is much higher, further proving the effectiveness of the proposed algorithm.

Next, to comprehensively compare the performance (the sum rate and the harvested power) among three schemes, we plot the relationship between the sum rate and the harvested power for each scheme, as shown in Fig. 3. In the conventional scheme, the sum rate varies with the requirement of the harvested power, resulting in a curve depicting the tradeoff between the sum rate and harvested power. This curve represents the achievable region of the conventional scheme. For the proposed algorithm and energy maximization scheme, the harvested power is maximized first, resulting in a single point for each scheme. From the figure, it is evident that the obtained point of the proposed algorithm is located at the top right of the curve of the conventional scheme. It means that under the same sum rate/harvested power, the harvested power/sum rate of the proposed algorithm is higher than that of the conventional scheme. This result is reasonable due to the following reasons. We derive the optimal set of the lower-level problem and then maximize the sum rate within the optimal set, while the conventional scheme simultaneously considers the sum rate and harvested power, leading to convergence to a stationary point with low performance.

Moreover, we evaluate the effect of CSI estimation error on the harvested power and the sum rate, as shown in Fig 4. Note that we set P^{req} as the maximum value obtained by the power maximization scheme for the conventional scheme. Meanwhile, the channel with the estimation error is modeled as: $\hat{\boldsymbol{h}} = \sqrt{1-\rho}\boldsymbol{h} + \sqrt{\rho}\boldsymbol{n}_h$, where ρ denotes the ratio of the CSI error power to $||\hat{\boldsymbol{h}}||^2$ and $\boldsymbol{n}_h$ is the error vector whose entries follow the distribution $\mathcal{CN}(0,1)$. For the harvested power, three schemes share the same curve and the harvested power decreases with the ratio of the CSI error power. The sum rates of the proposed algorithm and the conventional scheme decrease with the ratio of the CSI error power, while that of the power maximization scheme remains invariant. This is because the former two opt to maximize the sum rate, but the latter one does not consider the sum rate.

V. CONCLUSION

In this paper, we considered the case where the power transfer function takes precedence over the information delivery function in the RIS-assisted system. Consequently, we formulated a bilevel optimization problem, where the lower-level problem maximizes the harvested power and the upper-level problem maximizes the sum rate. We first derived the optimal set to the lower-level problem, which determines the direction of beamforming vectors. After that, we developed a B^{3}D algorithm for the original problem with the help of the BCD method. Finally, by comparing it with benchmark schemes, numerical results showed that the proposed algorithm achieves a higher sum rate with the same harvested power, which underscores the effectiveness and superiority.

REFERENCES

[1] S. H. Shah and I. Yaqoob, "A survey: Internet of Things (IoT) technologies, applications and challenges," in *Proc. IEEE Smart Energy Grid Eng. (SEGE)*, 2016, pp. 381–385.
[2] L. R. Varshney, "Transporting information and energy simultaneously," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2008, pp. 1612–1616.
[3] R. Zhang and C. K. Ho, "MIMO broadcasting for simultaneous wireless information and power transfer," *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 1989–2001, Mar. 2013.
[4] X. Peng, P. Wu, H. Tan, and M. Xia, "Optimization for IRS-assisted MIMO-OFDM SWIPT system with nonlinear EH model," *Internet Things J.*, vol. 9, no. 24, pp. 25 253–25268, 2022.
[5] 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, Sep. 2014.
[6] Q. Shi, L. Liu, W. Xu, and R. Zhang, "Joint transmit beamforming and receive power splitting for MISO SWIPT systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 6, pp. 3269–3280, Jun. 2014.
[7] W. Khalid, H. Yu, J. Cho, Z. Kaleem, and S. Ahmad, "Rate-energy tradeoff analysis in RIS-SWIPT systems with hardware impairments and phase-based amplitude response," *IEEE Access*, vol. 10, pp. 31821–31835, 2022.
[8] J. F. Bard. *Practical bilevel optimization: algorithms and applications*. Springer Science & Business Media, 2013.
[9] S. Lu, "Bilevel optimization with coupled decision-dependent distributions," in *Proc. Int. Conf. Mach. Learn.*, Apr. 2023, pp. 1–32.
[10] Z. Lu and S. Mei, "First-order penalty methods for bilevel optimization," 2023, *arXiv:2301.01716v1*.
[11] Q. Wu, S. Zhang, B. Zheng, C. You, and R. Zhang, "Intelligent reflecting surface-aided wireless communications: A tutorial," *IEEE Trans. Commun.*, vol. 69, no. 5, pp. 3313–3351, May 2021.
[12] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and C. He, "An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel," *IEEE Trans. Singal Process.*, vol. 59, no. 9, pp. 4331–4340, 2011.
[13] Q. Shi and M. Hong, "Penalty dual decomposition method for nonsmooth nonconvex optimization–Part I: Algorithms and convergence analysis," *IEEE Trans. Signal Process.*, vol. 68, pp. 4108–4122, 2020.