Energy Efficiency Optimization for MISO WIPT Systems With Zero-Forcing Beamforming

Qingjiang Shi, Cheng Peng, Weiqiang Xu, Senior Member, IEEE, Mingyi Hong, and Yunlong Cai, Member, IEEE

Abstract—This paper considers a power splitting based multi-user multiple-input-single-output (MISO) downlink system with simultaneous wireless information and power transfer, where each single antenna receiver splits the received signal into two streams of different power for decoding information and harvesting energy separately. Assuming that the most common zero-forcing (ZF) beamforming scheme is employed by the base station, we aim to maximize the system energy efficiency in bits per Joule by joint beamforming and power splitting under both the signal-to-interference-plus-noise ratio constraints and energy harvesting constraints. The energy efficiency optimization (EEO) problem is nonconvex and very hard to solve. In this paper, by exploiting the problem structure, we first simplify the EEO problem to a joint transmit power allocation and receive power splitting problem. Then, with a careful reformulation, we propose a Lagrangian relaxation (LR) method coupled with Dinkelbach method to address the simplified EEO problem, whilst devising a nearly closed-form solution for the subproblems involved in the Dinkelbach method. It is proven that the proposed LR method is optimum under some condition and can guarantee at least a feasible solution, which is a notable advantage over the existing methods. Besides, we develop a low complexity EEO algorithm by proportionally distributing the total power to users. Finally, numerical results validate the excellent efficiency of the proposed algorithms.

Index Terms—Dinkelbach method, Lagrangian relaxation method, multi-user MISO downlink, power splitting, wireless information and power transfer, zero-forcing beamforming.

I. INTRODUCTION

Currently, a variety of new applications of information and communication technologies have sprung up, which demand for high data rate communications with certain quality of service (QoS) guarantees and lead to a huge amount of energy consumption. On the other hand, traditional wireless devices are usually battery-powered with limited electricity capacity. Therefore, energy efficient design of communication systems incorporating energy harvesting technologies has recently drawn an upsurge of interest in both industrial and academia. Particularly, integrating radio frequency (RF) energy harvesting with traditional wireless information transmission has driven a new research area—(simultaneous) wireless information and power transfer ((S)WIPT). This necessitates a fundamental paradigm change in designing signal processing/communication/energy harvesting systems and requires for a unified study [1].

WIPT has been studied for various communication systems in different context [2]–[12]. While initial works in this field focused on point-to-point single antenna systems and studied tradeoffs between information rate and power transfer [2]–[4], most of the recent studies on WIPT focused on multi-antenna systems [5]–[12]. For example, as one of the pioneering research on multi-antenna WIPT systems, Zhang et al. [5] considered a MIMO information-energy broadcast system of three nodes (one transmitter, one information receiver and one energy receiver), and investigated the relevant rate-energy region and optimal transmission schemes with time switching and power splitting receivers. The works [6]–[12] investigated multi-user MISO WIPT systems. In [6], Xu et al. investigated the optimal information/energy beamforming strategy to achieve the maximum harvested energy for multi-user MISO WIPT systems with separated information/energy receivers. Based on the PS receiving scheme, Shi et al. [7] studied the optimal joint beamforming and power splitting (JBPS) to achieve the minimum transmission power of a multi-user MISO WIPT downlink system subject to both signal-to-interference-plus-noise (SINR) constraints and energy harvesting (EH) constraints. A similar JBPS problem for multi-user MISO interference channel (IFC) was studied in [8], [9] where efficient suboptimal solutions based on conic programming relaxation methods were obtained. Multi-antenna WIPT interference channels were also considered in the works [10]–[12] where the researchers investigated the transmission strategy with energy harvesting constraints. In addition, WIPT has been studied in other channel setups such as relay channels [13]–[16], orthogonal frequency division multiplexing (OFDM) channels [17], [18], and physical layer security channels [19]–[22].
Almost all of the above works focused on WIPT system design with the goal of throughput maximization, transmission power minimization, or harvested power/energy maximization. However, high energy efficiency (EE), defined as the number of information bits delivered per unit energy (bits per Joule), will be pursued in the design of future green wireless communication systems [23]. While there are a great number of works on energy efficiency optimization (EEO) for traditional communication systems [23], energy efficiency optimization was also recently introduced to energy-efficient WIPT system design [25]–[30]. In [25], Ng et al. considered energy-efficiency based resource allocation for a single-antenna point-to-point OFDM WIPT system with a PS-based receiver. The EEO problem in [25] is complicated by the coupling of the PS ratio and the subcarrier power allocation. Based on Dinkelbach method [24], the authors of [25] proposed an iterative resource allocation algorithm in aid of one dimensional exhaustive search for the PS ratio. In [26], Ng et al. studied energy efficient joint user selection and power allocation for single-antenna multi-user downlink OFDM WIPT systems where only one user is selected to be an information receiver while the rest are energy receivers. Due to user selection, the EEO problem in [26] is a mixed integer programming problem and thus it is much more complicated than that in [25]. Resorting to linear relaxation for the discrete user selection variables, [26] proposed an iterative resource allocation algorithm for multi-user downlink OFDM WIPT systems. By extending [25], [26], the work [27] investigated energy efficiency based joint subcarrier, PS ratio, and power allocation for single-antenna multiuser downlink OFDM WIPT systems with PS-based receivers, and proposed a similar iterative resource allocation algorithm as in [25], [26]. In [28], Chen et al. considered a time division duplex (TDD) large-scale MISO WIPT system, where, a terminal $S_1$ equipped with a large-scale antenna array first transfers wireless power through energy beamforming to a single-antenna terminal $S_2$ in slot 1 and then the terminal $S_2$ transmits information to $S_1$ in slot 2 using the harvested energy. Based on Dinkelbach method, an energy-efficient resource allocation scheme is proposed in [28] by jointly optimizing transmit power and transfer duration. In [29], He et al. investigated energy efficiency optimization for multicell multiuser MIMO downlink WIPT systems with energy harvesting constraints only. They assumed in [29] that the PS ratios are equal for all users and then proposed a two-tier suboptimal algorithm to tackle an equivalent problem of their EEO problem, i.e., iteratively adjusting the common PS ratio in the outer loop while optimizing the precoders with fixed PS ratios using block coordinate descent method in the inner loop. Different from [25]–[29], the work [30] studied the downlink and uplink energy efficiency tradeoff in TDD multi-user OFDMA systems where the users are equipped with PS-based receivers, respectively. The established multi-objective resource allocation problem is also a mixed integer programming. Through relaxation and transformation, Xiong et al. [30] developed a near-optimal resource allocation strategy that approaches the Pareto optimal tradeoff performance.

In this paper, we consider a multi-user MISO downlink SWIPT system as in [7], where a base station transmits radio signals to multiple single antenna receivers that can split the received signal into two streams of different power for decoding information and harvesting energy separately. Different from [7] where transmission power is minimized with QoS guarantee, this paper studies energy efficiency optimization by joint transmit beamforming and receiving power splitting. Specifically, we assume that the most common zero-forcing (ZF) beamforming scheme is employed by the system and aim to maximize the system energy efficiency under both SINR constraints and EH constraints. The established EEO problem is more complicated than the transmission power minimization problem in [7] due to the highly nonconvex fractional EE objective function. In this paper, by exploiting the problem structure, we first simplify the EEO problem to a joint transmit power allocation and receiving power splitting problem. With a tacit reformulation, we then propose a Lagrangian relaxation (LR) method coupled with the Dinkelbach method to address the simplified EEO problem, whilst devising a nearly closed-form solution for the subproblems involved in the Dinkelbach method. We prove that the proposed LR method is optimum under some sufficient optimality condition1 and can guarantee at least a feasible solution, which is a notable advantage over the methods used in [25]–[28] (See Prop. 3.1 and Remark 3.1 below). Besides, by proportionally distributing the total power to each user, we develop a low complexity EEO algorithm. Finally, numerical results verify that the proposed algorithms are very efficient in terms of both the convergence performance and the achieved energy efficiency performance. In particular, simulation results show that the LR method can achieve the optimum energy efficiency. In addition, it is emphasized that the proposed low complexity algorithm can globally solve the EEO problem of multi-user interference channels with PS-type WIPT, which is also a contribution of this paper.

The remainder of this paper is organized as follows. In the next section, we describe the system model and state the EE maximization problem formulation. Section III presents Lagrangian relaxation method for EE optimization while Section IV presents the Dinkelbach method for calculating the dual function of the EE maximization problem. A low complexity EE optimization method is proposed in Section V. Section VII provides numerical examples while Section VI concludes the paper.

Notations: scalars are denoted by lower-case letters; bold-face lower-case letters are used for vectors, while bold-face upper-case letters are for matrices. $A^H$ denote the conjugate transpose of a complex matrix $A$, $\|x\|$ denotes the Euclidean norm of a complex vector $x$, while $\alpha$ denotes the absolute value of a complex scalar $a$. For a real function $f(x)$, $\frac{\partial f}{\partial x}$ denotes its first-order derivative at $x$. The distribution of a circularly symmetric complex Gaussian (CSCG) random vector with mean $\mu$ and covariance matrix $C$ is denoted by $CN(\mu, C)$, and ‘~’ stands for ‘distributed as’. Finally, $C^{n \times n}$ denotes the space of $n \times n$ complex matrices.

II. SYSTEM MODEL AND PROBLEM FORMULATION

We first illustrate the system model and then give the problem formulation with some simplification.

1 Unfortunately, we are not able to show in theory how often the sufficient optimality condition is met. And strictly speaking, we have not yet theoretically proven the optimality of the LR method. However, it is found in simulations that the LR method can achieve the optimal energy efficiency performance.
A. System Description

As in [7], we consider a multi-user MISO downlink SWIPT system where one base station (BS) equipped with $N_t > 1$ antennas serves $K$ single-antenna users over a given frequency band of bandwidth $W$, as shown in Fig. 1. The system considered here differs from the traditional MISO downlink system in that each receiver splits the received signal into two independent streams: one stream is used for information decoding (ID) while the other stream is used for energy harvesting (EH). We assume that linear beamforming is employed by the BS to transmit signals to users. Thus the complex baseband transmitted signal at BS is expressed as

$$ x = \sum_{k=1}^{K} \mathbf{v}_k \mathbf{h}_k $$

(1)

where $s_k \sim \mathcal{C}\mathcal{N}(0, 1)$ denotes the transmitted data symbol and $\mathbf{v}_k$ denotes the transmit beamforming vector for user $k$.

Assume flat fading for channels between the BS and all users and let $\mathbf{h}_k$ denote the conjugate channel vector between the BS and user $k$. Then the received signal at user $k$ before power splitting is given by

$$ y_k = \mathbf{h}_k^H x + n_k, \quad k = 1, 2, \ldots, K $$

(2)

where $n_k \sim \mathcal{C}\mathcal{N}(0, \sigma_k^2)$ denotes the antenna noise at user $k$.

After power splitting with ratio $\rho_k$, the signal for information decoding at user $k$ is expressed as

$$ y_{k}^{\text{ID}} = \sqrt{\rho_k} \left( \mathbf{h}_k^H \sum_{j=1}^{K} \mathbf{v}_j s_j + n_k \right) + z_k, \quad \forall k $$

(3)

where $z_k \sim \mathcal{C}\mathcal{N}(0, \delta_k^2)$ is the additional noise introduced by the ID at user $k$, and the signal for energy harvesting at user $k$ is expressed as

$$ y_{k}^{\text{EH}} = \sqrt{1 - \rho_k} \left( \mathbf{h}_k^H \sum_{j=1}^{K} \mathbf{v}_j s_j + n_k \right), \quad \forall k $$

(4)

Accordingly, the SINR for ID at user $k$ is given by

$$ \text{SINR}_k = \frac{\rho_k |\mathbf{h}_k^H \mathbf{v}_k|^2}{\rho_k \sum_{j \neq k} |\mathbf{h}_k^H \mathbf{v}_j|^2 + \rho_k \sigma_k^2 + \delta_k^2}, \quad \forall k $$

(5)

and the harvested power by the EH of user $k$ is given by

$$ E_k = \zeta_k (1 - \rho_k) \left( \sum_{j=1}^{K} |\mathbf{h}_k^H \mathbf{v}_j|^2 + \sigma_k^2 \right), \quad \forall k $$

(6)

where it is assumed that all the noise components and symbols are independent from each other, $\zeta_k \in (0, 1]$ denotes the energy conversion efficiency at the EH of user $k$.

B. Problem Statement

We study the EEO problem for the considered system. EE is equivalently defined as the ratio between the system sum rate and the system power consumption [23]. In the following, we first describe the two components of energy efficiency—sum rate and power consumption, and then give the problem formulation with some simplification.

Assuming that perfect channel state information (CSI) is available, the data rate of user $k$ is given by

$$ R_k \triangleq W \log(1 + \text{SINR}_k), \quad \forall k. $$

(7)

Thus the system sum rate is $R \triangleq \sum_{k=1}^{K} R_k$.

Now we consider the system power consumption. In general, RF energy harvesting generates a very small amount of energy and thus may not contribute too much to the system energy efficiency. However, it is not difficult to see that, the energy efficiency of a wireless communication system can be indeed improved by energy harvesting, since the system power consumption is counteracted by the harvested power. Hence, as in [27], we take the harvested power into consideration in the formulation of the system power consumption (hence, the energy efficiency formulation for the multi-user MISO SWIPT system).

Specifically, the total system power consumption is expressed as follows

$$ P_s = \vartheta \sum_{k=1}^{K} \| \mathbf{v}_k \|^2 + P_c - \sum_{k=1}^{K} E_k. $$

(8)

where the first term $\vartheta \sum_{k=1}^{K} \| \mathbf{v}_k \|^2$ represents the power dissipation in the power amplifier of the transmitter ($\vartheta \geq 1$ is the power inefficiency of the amplifier), the second term $P_c$, which is independent of the first term, stands for the constant power consumption of the transceivers induced mainly by signal processing (it will be elaborated in Section VI), and the last term represents the harvested power of the $K$ receivers.

Given the expression of the system sum rate and power consumption, we are ready to state the problem formulation. The goal of this paper is to do joint transceiver design so that the system energy efficiency, defined by $R_k/P_s$ [27], is maximized while meeting two kinds of QoS constraints—SINR and EH constraints. Particularly, to simplify our transceiver design, we assume $N_t \geq K$ and consider the most common beamforming scheme—zero-forcing (ZF) processing, which nulls the multiuser interference signals. Mathematically, the problem is

$$ \min_{\mathbf{V}} \frac{R_k}{P_s} \text{subject to} \quad \text{SINR}_k \geq \gamma_k, \quad \forall k, \quad \text{EH} \geq \epsilon_k, \quad \forall k. $$

(9)

where the right hand side of (9) describes the system energy efficiency formulation with harvested power as shown in (9) is more general and the proposed algorithm for (9) can be straightforwardly extended to the EEO problem whose EE formulation contains no harvested power.

ZF beamforming is commonly used in transceiver design, since it is simple yet possibly offer asymptotically optimum performance, e.g., [7] has showed that ZF beamforming is asymptotically optimum in achieving the minimum transmission power of the multi-user MISO downlink SWIPT system. Furthermore, it is known that ZF beamforming works when $N_t \geq K$. 

2From the point of view of optimization, the energy efficiency formulation with harvested power as shown in (9) is more general and the proposed algorithm for (9) can be straightforwardly extended to the EEO problem whose EE formulation contains no harvested power.

3ZF beamforming is commonly used in transceiver design, since it is simple yet possibly offer asymptotically optimum performance, e.g., [7] has showed that ZF beamforming is asymptotically optimum in achieving the minimum transmission power of the multi-user MISO downlink SWIPT system. Furthermore, it is known that ZF beamforming works when $N_t \geq K$. 

---

Fig. 1. A multi-user MISO SWIPT system with PS-based receivers [7].
equivalently formulated as follows (by neglecting the constant $W$)

$$\max_{\{\mathbf{u}_k, \rho_k\}} \left\{ \sum_{k=1}^{K} R_k \right\}$$

$$\text{s.t.} \frac{\rho_k \mathbf{h}_k^H \mathbf{v}_k}{\mathbf{h}_k^H \mathbf{v}_k} \geq \gamma_k, \forall k$$

$$\sum_{k=1}^{K} E_k \leq P_{\text{total}},$$

$$\sum_{k=1}^{K} \mathbf{v}_k^H \mathbf{v}_k \leq P_{\text{total}},$$

where $\mathbf{H}_k \triangleq \{\mathbf{h}_1, \ldots, \mathbf{h}_{k-1}, \mathbf{h}_{k+1}, \ldots, \mathbf{h}_K\}$, $\forall k$, $E_k$ and $R_k$ are functions of $(\mathbf{u}_k, \rho_k)$ (cf. (6) and (7)), (9b) and (9c) are respectively the SINR and EH constraints with $\gamma_k$'s and $e_k$'s being the corresponding QoS targets, (9d) are the constraints induced by the ZF beamforming which nulls the interference signal, (9e) is the total power constraint with $P_{\text{total}}$ being the allowed maximum transmission power, while (9f) are the inherent constraints for PS ratios. Obviously, it is difficult to solve problem (9) since the problem has a nonconcave objective function and a nonconvex constraint set. However, we can simplify the problem by exploring the ZF constraints (9d) and exploiting the problem structure. Specifically, with the following definitions:

$$\tilde{\mathbf{u}}_k \triangleq \frac{\mathbf{U}_k \tilde{\mathbf{H}}_k^H \mathbf{h}_k}{\| \mathbf{U}_k \tilde{\mathbf{H}}_k^H \mathbf{h}_k \|} \quad \text{and} \quad \tilde{g}_k \triangleq \mathbf{h}_k^H \tilde{\mathbf{u}}_k, \forall k$$

where $\mathbf{u}_k$ denotes the orthogonal basis of the null space of $\mathbf{H}_k^H$, we have the following proposition.

**Proposition 2.1:** Problem (9) can be simplified as follows

$$\max_{\{\tilde{p}_k, \rho_k\}} \left\{ \sum_{k=1}^{K} \log \left( 1 + \frac{p_k \tilde{g}_k}{p_k \tilde{g}_k + \tilde{g}_k^2} \right) \right\}$$

$$\text{s.t.} \frac{\rho_k \tilde{p}_k \tilde{g}_k}{\rho_k \tilde{g}_k^2 + \tilde{g}_k^2} \geq \gamma_k, \forall k,$$

$$\tilde{g}_k (1 - \rho_k) (p_k \tilde{g}_k + \sigma_k^2) \geq e_k, \forall k,$$

$$\sum_{k=1}^{K} p_k \leq P_{\text{total}},$$

$$0 \leq \rho_k \leq 1, \forall k$$

with $\mathbf{v}_k = \sqrt{p_k} \tilde{\mathbf{u}}_k, k = 1, 2, \ldots, K$.

**Proof:** Please see Appendix A.

**Proposition 2.1** suggests that the optimal solution to problem (9) can be obtained by solving problem (11). Hence, the main task of the rest of this paper is to solve problem (11). Before demonstrating our solution, we study the feasibility of problem (11). Lemma 2.1 shows a sufficient and necessary condition for the feasibility of problem (11).

**Lemma 2.1:** Define $l_1(\rho_k) \triangleq \frac{\gamma_k (p_k \sigma_k^2 + \tilde{g}_k^2)}{p_k \tilde{g}_k}$ and $l_2(\rho_k) \triangleq \frac{1}{\rho_k} \left( \frac{\gamma_k \sigma_k^2}{\rho_k} - \tilde{g}_k^2 \right), k = 1, 2, \ldots, K$. Furthermore, define $p_{k,\text{min}} \triangleq \min_{1 \leq k \leq K} \max(l_1(\rho_k), l_2(\rho_k)), k = 1, 2, \ldots, K$. Problem (11) is feasible if and only if $\sum_{k=1}^{K} p_{k,\text{min}} \leq P_{\text{total}}$.

**Proof:** Please see Appendix B.

Note that $p_{k,\text{min}}$ is attained when $l_1(\rho_k) = l_2(\rho_k)$, which can be easily obtained by solving a quadratic equation. Hence, Lemma 2.1 indicates that the feasibility of problem (11) (equivalently (9)) can be easily verified. Without loss of generality, in the rest of this paper we assume that problem (11) is feasible, i.e., $\sum_{k=1}^{K} p_{k,\text{min}} \leq P_{\text{total}}$.

**III. LA GRANGIAN RELAXATION METHOD FOR ENERGY EFFICIENCY OPTIMIZATION**

For notational convenience, we define $E_k(\rho_k, \rho_k) \triangleq \tilde{g}_k (1 - \rho_k) (p_k \tilde{g}_k + \sigma_k^2)$ and $R_k(\rho_k, \rho_k) \triangleq \log \left( 1 + \frac{p_k \tilde{g}_k + \sigma_k^2}{p_k \tilde{g}_k^2 + \tilde{g}_k^2} \right)$. Thus problem (11) is compactly written as

$$\max_{\{p_k, \rho_k\}} \left\{ \sum_{k=1}^{K} R_k(\rho_k, \rho_k) \right\}$$

$$s.t. \sum_{k=1}^{K} \tilde{p}_k \tilde{g}_k + \tilde{g}_k \geq \gamma_k, \forall k,$$

$$\gamma_k (1 - \rho_k) (p_k \tilde{g}_k + \sigma_k^2) \geq e_k, \forall k,$$

$$\sum_{k=1}^{K} p_k \leq P_{\text{total}},$$

$$0 \leq \rho_k \leq 1, \forall k.$$
power consumption $\vartheta \sum_{k=1}^{K} p_k + P_c - \sum_{k=1}^{K} E_k(p_k, \rho_k)$ is positive. Thus, the total power constraint is equivalent to

$$\frac{\sum_{k=1}^{K} p_k - P_{total}}{\sum_{k=1}^{K} p_k + P_c - \sum_{k=1}^{K} E_k(p_k, \rho_k)} \leq 0.$$  

It follows that problem (12) is equivalent to

$$\max_{\{p_k, \rho_k\}} \frac{\sum_{k=1}^{K} R_k(p_k, \rho_k)}{\sum_{k=1}^{K} p_k + P_c - \sum_{k=1}^{K} E_k(p_k, \rho_k)}$$

s.t. $\frac{p_k \rho_k g_k}{\rho_k \sigma_k^2 + \delta_k^2} \geq \gamma_k, \ \forall k,$

$$\zeta_k (1 - \rho_k) \left( p_k g_k + \sigma_k^2 \right) \geq \epsilon_k, \ \forall k,$$

$$0 < \rho_k < 1, \ \forall k,$$

$$\sum_{k=1}^{K} p_k - P_{total}.$$

Problem (13) appears to be harder than problem (12). However, it will be clear later that the dual function of problem (13) can be obtained in nearly closed-form and consequently the corresponding dual problem can be easily solved.

Now, we are ready to use the LR method to address problem (13) (equivalently, (12)). To this end, we introduce Lagrange multiplier $\lambda$ to the last constraint of problem (13) and define the partial Lagrangian associated with problem (13) as

$$L(p_k, \rho_k, \lambda) = \sum_{k=1}^{K} R_k(p_k, \rho_k)$$

$$- \frac{\sum_{k=1}^{K} p_k + P_c - \sum_{k=1}^{K} E_k(p_k, \rho_k)}{\sum_{k=1}^{K} p_k + P_c - \sum_{k=1}^{K} E_k(p_k, \rho_k)} \lambda.$$

With the above partial Lagrangian, the dual function, denoted by $d(\lambda)$, is written as

$$d(\lambda) = \max_{\{p_k, \rho_k\}} L(p_k, \rho_k, \lambda)$$

s.t. $\frac{p_k \rho_k g_k}{\rho_k \sigma_k^2 + \delta_k^2} \geq \gamma_k, \ \forall k,$

$$\zeta_k (1 - \rho_k) \left( p_k g_k + \sigma_k^2 \right) \geq \epsilon_k, \ \forall k,$$

$$0 \leq \rho_k \leq 1, \ \forall k.$$  

It is noteworthy that the optimal dual solution $\lambda$ may not lead to a feasible primal solution $\{p_k(\lambda), r_k(\lambda)\}$. However, since $d(\lambda)$ is a convex function and $\lambda^{\ast} \left( \sum_{k=1}^{K} p_k(\lambda^{\ast}) - P_{total} \right)$ is a subgradient of $d(\lambda)$ [33, pp. 12], we have the following proposition.

**Proposition 3.1:** Let $\lambda^{\ast}$ be the optimal dual variable. Then

1) If $p_k(\lambda^{\ast})$ satisfy the total power constraint and $\lambda^{\ast} \left( \sum_{k=1}^{K} p_k(\lambda^{\ast}) - P_{total} \right) = 0$, $\{p_k(\lambda^{\ast}), r_k(\lambda^{\ast})\}$ is an optimal solution to problem (12);

2) For any $\lambda > \lambda^{\ast}$, $\{p_k(\lambda), r_k(\lambda)\}$ satisfies the total power constraint, i.e., $\sum_{k=1}^{K} p_k(\lambda) \leq P_{total}$.

**Proof:** See Appendix C.

According to Proposition 3.1, we infer that either an optimal solution or a feasible solution can be obtained by the LR method. The main task of the LR method is to solve the dual problem (16). This can be done by using Bisection method [32] with the aid of the subgradient of $d(\lambda)$. We summarize the proposed LR method in Algorithm 1 in Table I, where Steps 1–5 check whether or not $\{p_k(\lambda), r_k(\lambda)\}$ solves problem (12) globally, Steps 6–16 represent the Bisection method which solves the dual problem (16) globally, and Steps 17–20, if necessary, are carried out to provide a feasible solution to problem (12). Now, the remaining challenge is to solve problem (15), which will be addressed in the following section.

Before proceeding to the next section, we make two remarks on Algorithm 1. The first remark points out the advantage of our

Note that the power amplifier inefficiency factor $\vartheta$ is greater than 1 while all $g_k$ ’s are much smaller than 1 due to large-scale path loss.
Remark 3.1: Problem (15) will be solved by an iterative algorithm called Dinkelbach method [24]. Thus the main part of Algorithm 1 (cf. Steps 8–16) can be viewed as a two-tier algorithm, where the inner loop solves problem (15) using the Dinkelbach method while the outer loop performs the bisection algorithm [32]. Note that, the iterative algorithms in [25]–[27] can be also viewed as two-tier algorithms, where, however, the outer loop implements the Dinkelbach method while the inner loop solves the subproblem of subtractive form by using the dual-decomposition method [33]. Hence, Algorithm 1 is fundamentally different from the algorithms in [25]–[27]. Furthermore, since the subproblems of subtractive form in [25]–[27] are nonconvex and much complicated, the authors of [25]–[27] have not yet proved that zero duality gap must be guaranteed for their problems. Consequently, the dual decomposition algorithms in [25]–[27] may not arrive at the optimal (even not feasible) solutions to the EEO problems in theory. However, as shown in the second part of Proposition 3.1, our algorithm guarantees at least a feasible solution. Moreover, simulations later show that the proposed algorithm can achieve the optimal energy efficiency.

Remark 3.2: If problem (15) (or equivalently (17) below) has a unique solution, then the subgradient is just the gradient of . Hence, by the first order optimality condition of the dual problem, we have

\[ \nabla \frac{\sum_{k=1}^{K} p_k(\lambda)}{\theta \sum_{k=1}^{K} p_k(\lambda) + P_c - \sum_{k=1}^{K} E_k(p_k, \rho_k(\lambda))} = \frac{\lambda}{\theta}. \]

is just the gradient of . Hence, by the first order optimality condition of the dual problem, we have \[ P_{\text{total}} - \sum_{k=1}^{K} p_k(\lambda^*) = 0 \] when a unique solution is obtained for (15). In other words, if problem (15) has a unique solution for \( \lambda = \lambda^* \), the LR method comes up with an optimal solution to the primal problem (13) (equivalently (11)) according to Proposition 3.1.

IV. DINKELBACH METHOD FOR PROBLEM (15)

Problem (15) is a nonlinear fractional programming. It is clearly not a convex problem and hard to solve. However, it is observed that both the numerator and the denominator of the objective function, as well as the constraint set of problem (15), are separable across user \( k = 1, 2, \ldots, K \). Hence, if we transform the objective function (i.e., the fractional form) into a numerator-denominator subtractive form, then the resultant problem can be decomposed into \( K \) tractable subproblems.

This is our basic idea for solving problem (15), realized by using the Dinkelbach method [24].

A. Dinkelbach Method

We propose using the Dinkelbach method, a popular technique for solving nonlinear fractional programmings, to solve problem (15). The following lemma lays a theoretical base for the Dinkelbach method.

Lemma 4.1: Define

\[ \Omega_k \triangleq \left\{ (p_k, \rho_k) \bigg| \frac{p_k \rho_k g_k}{\rho_k \sigma_k^2 + \delta_k^2} \geq \gamma_k, \right. \]

\[ \left. \zeta_k (1 - \rho_k) \left( p_k g_k + \sigma_k^2 \right) > \epsilon_k, 0 \leq p_k < 1 \right\}. \]

Then \( \{p^*_k, \rho^*_k\} \) solves problem (15) if and only if it solves

\[ \max_{(p_k, \rho_k) \in \Omega_k, \forall k} \sum_{k=1}^{K} R_k(p_k, \rho_k) - \lambda \left( \sum_{k=1}^{K} p_k - P_{\text{total}} \right) - \eta \left( \frac{\theta}{\sum_{k=1}^{K} p_k + P_c - \sum_{k=1}^{K} E_k(p_k, \rho_k)} \right) \]

with \( \eta - \eta^* \) given by

\[ \eta^* \triangleq \sum_{k=1}^{K} R_k(p_k^*, \rho_k^*) - \lambda \left( \sum_{k=1}^{K} p_k^* - P_{\text{total}} \right) - \frac{\theta}{\sum_{k=1}^{K} p_k^* + P_c - \sum_{k=1}^{K} E_k(p_k^*, \rho_k^*)}. \]

Lemma 4.1 can be proven by following a similar approach as in [24]. Clearly, \( \eta^* \) is just the maximum energy efficiency. Let \( F(\eta) \) denote the optimal value of problem (17). The lemma implies that the maximum energy efficiency \( \eta^* \) satisfies \( F(\eta^*) = 0 \). Note that, the Dinkelbach method is an iterative algorithm that generates a sequence of values of \( \eta \) converging to the maximum energy efficiency \( \eta^* \) monotonically such that \( F(\eta^*) = 0 \). We summarize the algorithm in Table II. Once problem (17) is globally solved in Step 4, Algorithm 2 will finally achieve an optimal solution to problem (15). The following subsection shows how to solve problem (17).

B. Solution to Problem (17)

Different from problem (15), problem (17) is separable. That is, problem (17) can be decomposed into \( K \) subproblems with the \( k \)-th subproblem given by

\[ \max_{p_k, \rho_k} R_k(p_k, \rho_k) - (\lambda + \eta \theta) p_k + \eta E_k(p_k, \rho_k) \]

s.t. \( \frac{p_k \rho_k g_k}{\rho_k \sigma_k^2 + \delta_k^2} \geq \gamma_k, \)

\[ \zeta_k (1 - \rho_k) \left( p_k g_k + \sigma_k^2 \right) \geq \epsilon_k, 0 \leq p_k \leq P_{\text{total}}, 0 \leq \rho_k \leq 1 \]

The Dinkelbach method is in essence a Newton method for solving the equation \( F(\eta) = 0 \) [34]. Since \( F(\eta) \) is a convex function, the Dinkelbach method yields an optimal solution to the equation.

We have imposed a bound on \( p_k \) for two reasons. First, \( p_k \) must be no greater than \( P_{\text{total}} \) and having this bound can make the algorithm numerically stable. Second, the technique developed in this section will be applied to the development of the low complexity method in Section V where the optimization model (38) has upper bounds on \( p_k \)’s.
where we have neglected some constant terms in the objective function. Clearly, problem (19) is nonconvex. However, it is readily seen that both $R_k(p_k, \rho_k)$ and $E_k(p_k, \rho_k)$ are concave functions of $p_k$. Moreover, all the constraint functions related to $p_k$ are linear in $p_k$. Thus, once $\rho_k$ is fixed, problem (19) is a convex problem with respect to $p_k$. This implies that problem (19) can be solved by one dimensional exhaustive search. However, the exhaustive search method is not efficient. In what follows, we show how problem (19) can be efficiently solved.

First, we recast problem (19) as follows

$$\max_{\rho_k, p_k} R_k(p_k, \rho_k) - (\lambda + \eta \theta) p_k + \eta E_k(p_k, \rho_k)$$

s.t. $p_k \geq l_1(\rho_k)$,

$p_k \geq l_2(\rho_k)$,

$0 \leq p_k \leq P_{\text{total}}$,

$0 \leq \rho_k \leq 1$ \hspace{1cm} (20)

where $l_1(\rho_k)$ and $l_2(\rho_k)$ are defined in Lemma 2.1. Then, by denoting the constraint set of $p_k$ as $\mathcal{P}_k(\rho_k) \triangleq \{ p_k \mid \max(l_1(\rho_k), l_2(\rho_k)) \leq p_k \leq P_{\text{total}} \}$, we rewrite problem (20) as the following equivalent two-tier maximization problem

$$\max_{0 \leq \rho_k \leq 1} \max_{p_k \in \mathcal{P}_k(\rho_k)} R_k(p_k, \rho_k) - (\lambda + \eta \theta) p_k + \eta E_k(p_k, \rho_k) \tag{21}$$

which sheds light on our solution to problem (19). As discussed above, for fixed $\rho_k$, the inner maximization problem of (21) is a convex problem. Moreover, since $R_k(p_k, \rho_k)$ is a strictly concave function of $p_k$, the inner maximization problem has a unique solution. Let $p_k(\rho_k)$ denote the unique optimal solution to the inner maximization problem of (21). Then the optimal $\rho_k$ of problem (19) can be obtained by solving the following problem

$$\max_{0 \leq \rho_k \leq 1} \psi_k(\rho_k) \tag{22}$$

where $\psi_k(\rho_k) \triangleq R_k(p_k(\rho_k), \rho_k) - (\lambda + \eta \theta) p_k(\rho_k) + \eta E_k(p_k(\rho_k), \rho_k)$.

Next, we solve problem (22) by considering all possible cases for $p_k(\rho_k)$. The inner maximization problem of problem (21) for fixed $\rho_k$ can be explicitly written as

$$\max_{p_k} R_k(p_k, \rho_k) - (\lambda + \eta \theta) p_k + \eta E_k(p_k, \rho_k)$$

s.t. $\max(l_1(\rho_k), l_2(\rho_k)) \leq p_k \leq P_{\text{total}}$ \hspace{1cm} (23)

By checking the first order condition, we obtain the stationary point of the objective function of problem (23) as follows

$$p_k^*(\rho_k) = \frac{1}{\lambda + \eta \theta - \eta \theta \rho_k g_k \sigma_k^2 - \delta_k^2} \tag{24}$$

Since problem (23) is a univariate convex problem with a bound constraint, its optimal value must be attained either on the boundary of the constraint or at the stationary point $p_k^*(\rho_k)$ in terms of the relative magnitude of $p_k^*(\rho_k)$, $\max(l_1(\rho_k), l_2(\rho_k))$.

1) $p_k(\rho_k) = P_{\text{total}}$: In this case, we have

$$p_k(\rho_k) = P_{\text{total}}$$

Taking the derivative of $\psi_k(\rho_k)$ with respect to $\rho_k$, we obtain

$$\frac{d \psi_k(\rho_k)}{d \rho_k} = \frac{P_{\text{total}} g_k + \sigma_k^2}{\rho_k (\rho_k g_k + \sigma_k^2)^2} - \eta \frac{g_k}{\rho_k} \tag{25}$$

Thus, $\frac{d \psi_k(\rho_k)}{d \rho_k} = 0$ implies that $\rho_k$ satisfies the following quadratic equation

$$a_1 \rho_k^2 + b_1 \rho_k + c_1 = 0 \tag{26}$$

where $a_1$, $b_1$, and $c_1$ are constants.

According to (25), problem (22) can be solved through first separately checking the above four cases and then keeping the best $\rho_k$ as the solution of problem (22). It can be shown that each condition in (25) can be reduced to one or several intervals of $\rho_k$. Hence, problem (22) under the above four conditions can be recast as optimization problems in the form of

$$\max_{\rho_k \in I_k} \psi_k(\rho_k) \tag{26}$$

where $I_k$ denotes the union of some intervals of $\rho_k$ (including the interval $0 \leq \rho_k \leq 1$). It is known that, for an optimization problem with interval constraints, its optimal solution must be either some endpoint of the intervals or some feasible stationary point of the objective function. Hence, regarding the optimal solution to problem (26), we have the following lemma.

**Lemma 4.2**: Suppose that $p_k(\rho_k)$ is differentiable. Then the optimal value of problem (26) must be attained either at the point that satisfies

$$\frac{d \psi_k(\rho_k)}{d \rho_k} = 0$$

or at some endpoint of $I_k$.

**Lemma 4.2** tells us that, besides the endpoints of $I_k$, if we can find all possible $\rho_k$’s that satisfy condition (27), the optimal $\rho_k$ of problem (26) can be obtained by simply checking the objective function value $\psi_k(\rho_k)$. This is our basic idea for solving problem (26) (hence, (22) and (19)). In the following, we consider the above four cases one by one and particularly examine the corresponding first-order condition $\frac{d \psi_k(\rho_k)}{d \rho_k} = 0$.

1. $p_k(\rho_k) = P_{\text{total}}$: In this case, we have

$$\psi_k(\rho_k) = \log \left( 1 + \frac{p_k P_{\text{total}} g_k}{\rho_k (\rho_k g_k + \sigma_k^2)} \right) - (\lambda + \eta \theta) P_{\text{total}}$$

Taking the derivative of $\psi_k(\rho_k)$ with respect to $\rho_k$, we obtain

$$\frac{d \psi_k(\rho_k)}{d \rho_k} = \frac{1}{\rho_k (\rho_k g_k + \sigma_k^2)^2} - \frac{P_{\text{total}} g_k + \sigma_k^2}{\rho_k (\rho_k g_k + \sigma_k^2)^2} - \eta \frac{g_k}{\rho_k} \tag{27}$$

Thus, $\frac{d \psi_k(\rho_k)}{d \rho_k} = 0$ implies that $\rho_k$ satisfies the following quadratic equation

$$a_1 \rho_k^2 + b_1 \rho_k + c_1 = 0 \tag{28}$$

Quartic equations can be solved in nearly closed-form [35].

---

7The four conditions in (25) can be simplified as trivial or quadratic inequalities which correspond to some intervals for $\rho_k$.

8Quartic equations can be solved in nearly closed-form [35].
with \( a_1 = \sigma_k^2(P_{\text{total}}g_k + \sigma_k^2) \), \( b_1 = \delta_k^2(P_{\text{total}}g_k + 2\sigma_k^2) \), and
\[
c_1 = \left( \frac{\delta_k^2}{\eta \psi_k'((\rho_kg_k) + \sigma_k^2)} \right).
\]

2) \( p_k(\rho_k) = l_1(\rho_k) \). In this case, we have
\[
\psi_k(\rho_k) - \log \left( 1 + \frac{\rho_kl_1(\rho_k)g_k}{\rho_k\sigma_k^2 + \delta_k^2} \right) - (\lambda + \eta \theta)l_1(\rho_k) + \eta \psi_k((\rho_kg_k) + \sigma_k^2).
\]

Let us recall \( l_1(\rho_k) \), which implies that
\[
\log \left( 1 + \frac{\rho_kl_1(\rho_k)g_k}{\rho_k\sigma_k^2 + \delta_k^2} \right)
\]
is a constant and \( \frac{d\psi_k(\rho_k)}{\rho_k} = \frac{\eta \psi_k((\rho_kg_k) + \sigma_k^2)}{\rho_k} \). Thus, we have
\[
\frac{d\psi_k(\rho_k)}{d\rho_k} = -\eta \psi_k((\rho_kg_k) + \sigma_k^2),
\]
\[
- \eta \psi_k(l_1(\rho_k)g_k + \sigma_k^2),
\]
\[
- (\lambda + \eta \theta)l_2(\rho_k) - \eta \psi_k((\rho_kg_k) + \sigma_k^2).
\]

3) \( p_k(\rho_k) = l_2(\rho_k) \). In this case, we have
\[
\psi_k(\rho_k) = \log \left( 1 + \frac{\rho_kl_2(\rho_k)g_k}{\rho_k\sigma_k^2 + \delta_k^2} \right) - (\lambda + \eta \theta)l_2(\rho_k).
\]

Let us recall \( l_2(\rho_k) \), which implies that
\[
\eta \psi_k((\rho_kg_k) + \sigma_k^2) \]
\( \) is a constant. Thus, we have
\[
\frac{d\psi_k(\rho_k)}{d\rho_k} = \frac{1}{\rho_k} \left( \frac{\eta \psi_k((\rho_kg_k) + \sigma_k^2)}{\eta \psi_k'((\rho_kg_k) + \sigma_k^2)} \right).
\]

Equating the above equation to zero, i.e., \( \frac{d\psi_k(\rho_k)}{d\rho_k} = 0 \), leads to
\[
\left( g_kl_2(\rho_k) + \rho_kg_kd\frac{\psi_k(\rho_k)}{d\rho_k} \right) + \sigma_k^2 = 0.
\]

By plugging \( l_2(\rho_k) = \frac{1}{\delta_k^2} \left( \frac{\psi_k'}{\psi_k'} - \sigma_k^2 \right) \) into (34) and with some manipulations, we can finally reduce (34) to the following quartic equation
\[
a_3\rho_k^4 + b_3\rho_k^3 + c_3\rho_k^2 + d_3\rho_k + e_3 = 0
\]
with \( a_3 = -2\delta_k^2\sigma_k^2g_k \), \( b_3 = -6\delta_k^2\sigma_k^2g_k + 2\delta_k^2 \), \( c_3 = (6\delta_k^2\sigma_k^2g_k - 3\delta_k^2g_k + 3\delta_k^2\sigma_k^2) \), \( d_3 = -\delta_k^2\sigma_k^2g_k(2\delta_k^2 - \delta_k^2) \), and \( e_3 = (6\delta_k^2\sigma_k^2g_k - 4\delta_k^2g_k + 3\delta_k^2\sigma_k^2) \).

4) \( p_k(\rho_k) = \psi_k(\rho_k) \). In this case, we have
\[
\psi_k(\rho_k) = \log \left( 1 + \frac{\rho_k\psi_k(\rho_k)g_k}{\rho_k\sigma_k^2 + \delta_k^2} \right) - (\lambda + \eta \theta)p_k(\rho_k).
\]

Similarly, we can finally reduce \( \frac{d\psi_k(\rho_k)}{d\rho_k} = 0 \) to the following quartic equation
\[
a_4\rho_k^4 + b_4\rho_k^3 + c_4\rho_k^2 + d_4\rho_k + e_4 = 0
\]
with \( a_4 = -g_k^2\sigma_k^2\eta^2\sigma_k^2 \), \( b_4 = -\delta_k^2\sigma_k^2\eta^2\sigma_k^2 \), \( c_4 = (\delta_k^2\eta^2 - 2\delta_k^2\lambda^2 + \eta^2 + 1)g_k^2\sigma_k^2\eta^2 \), \( d_4 = -\delta_k^2\sigma_k^2\eta^2 - 2\delta_k^2\lambda^2 + 1)g_k^2\sigma_k^2\eta^2 \), and \( e_4 = g_k^2\sigma_k^2\eta^2 \).

To sum up, we can obtain the candidates of the optimal \( \rho_k \) by solving problem (26) in four cases, equivalently, first solving (29), (32), (35) and (37) and then checking the objective values of all the endpoints of \( T_k \) as well as the obtained feasible stationary points. The best candidate with the maximum value of \( \psi_k(\rho_k) \) is picked as the optimal \( \rho_k \). Accordingly, we obtain the optimal \( \rho_k \) in terms of (25). Therefore, all the subproblems of (17) can be solved with complexity of \( O(1) \) in parallel. As a result, the worst-case complexity of Algorithm 1 is \( O(T_DN) \), where \( T_D \) and \( N \) denote the number of iterations required by the Dinkelbach method and Bisection method, respectively. Generally, both the Dinkelbach method and Bisection method achieve convergence in about ten iterations (which
is validated with numerical examples in Section VI. Thus Algorithm 1 is very efficient.

V. LOW COMPLEXITY ALGORITHM FOR PROBLEM (12)

In Section III, we have proposed Algorithm 1 to address the energy efficiency optimization problem (12). As argued in Remark 3.1, Algorithm 1 is in essence a two-tier iterative algorithm. In this section, we propose a suboptimal but single-tier (hence, low complexity) iterative algorithm based on only the Dinkelbach method.

It is observed that, if the total power constraint of problem (12) is replaced with independent user power constraints, i.e., $p_k \leq P_k, \forall k$, the corresponding energy efficiency optimization problem, i.e.,

$$\max_{\{p_k, \rho_k\}} \frac{\sum_{k=1}^{K} R_k(p_k, \rho_k)}{\vartheta \sum_{k=1}^{K} p_k + P_c - \sum_{k=1}^{K} E_k(p_k, \rho_k)}$$

s.t.

$$p_k \rho_k g_k \geq \gamma_k, \forall k,$$

$$\zeta_k (1 - \varrho_k) \left( \frac{p_k g_k + \sigma_k^2}{\sigma_k^2} \right) \geq \epsilon_k, \forall k,$$

$$0 \leq p_k \leq P_k, \forall k,$$

$$0 \leq \rho_k \leq 1, \forall k,$$

(38)
can be globally solved by using the Dinkelbach method in Section IV (without need of the LR method). Motivated by this observation, we instead consider solving problem (38) in order to achieve a low complexity suboptimal transceiver design. A direct way for achieving $K$ individual power constraints is just equally distributing the total power to $K$ users, i.e., $p_k \leq P_{\text{total}}/K, \forall k$. However, problem (38) with $K$ equally distributed power constraints has smaller feasible region than problem (12). Moreover, it may become even infeasible. To address this challenge while obtaining $K$ feasible individual user power constraints, we proportionally distribute the total power to each user and have each $P_k$ proportional to $p_{k,\text{min}}$ (defined in (2.1)), i.e., let $P_k = \frac{P_{\text{total}} p_{k,\text{min}}}{\sum_{k=1}^{K} p_{k,\text{min}}}, \forall k$. It is not difficult to see that problem (38) with such $H_k$'s must be feasible if problem (11) is feasible. Hence, our low complexity method just solves problem (38) with $P_k = \frac{P_{\text{total}} p_{k,\text{min}}}{\sum_{k=1}^{K} p_{k,\text{min}}}, \forall k$. According to Section IV, it is readily known that problem (38) can be globally solved using the Dinkelbach method, leading to a single-tier iterative algorithm. Obviously, the proposed single-tier iterative algorithm has lower complexity (of $O(T_\text{ijk})$) than Algorithm 1. Moreover, it will be seen from simulations that this low complexity algorithm could achieve a performance very close to that of Algorithm 1. Two important remarks are made as follows on the proposed low complexity algorithm.

Remark 5.1: If the optimal solution $p_k$ to problem (12) satisfies the individual power constraints of problem (38), we can obtain the optimal solution to problem (12) through solving problem (38). This means that the proposed low complexity algorithm could achieve the optimal solution to problem (12) when $P_{\text{total}}$ is large, although it is suboptimal in general for problem (12).

Remark 5.2: It is worth noting that, when ZF beamforming is employed in MISO SWIPT interference channels [8], [9], the energy efficient transceiver design problem of MISO SWIPT interference channels is in almost the same form of (38) with $P_k$ being the maximum transmission power of the $k$-th transmitter. Hence, the proposed low complexity algorithm can provide an optimal energy efficient transceiver design method for MISO SWIPT interference channels with ZF beamforming. We emphasize that this can be viewed as a contribution of this paper.

VI. SIMULATION RESULTS

In this section, we numerically evaluate the performance of the proposed energy efficiency optimization algorithms in downlink MISO SWIPT systems of $K = 4$ users. Unless specified otherwise, it is assumed that the BS is equipped with $N_t = 4$ antennas and the total transmission power is $P_{\text{total}} = 30$ dBm. We assume that the system bandwidth is $W = 15$ kHz and all users have the same set of parameters, i.e., $\zeta_k = \zeta$, $\delta_k = \delta$, $\sigma_k^2 = \sigma^2$, $\epsilon_k = \epsilon$, and $\gamma_k = \gamma, \forall k$. Moreover, we set $\vartheta = \delta$, $\zeta = 0.65$, $\alpha^2 = -70$ dBm and $\delta^2 = -50$ dBm in all simulations. It is further assumed that the signal attenuation from BS to all users is 40 dB corresponding to an identical distance of 5 meters. With this transmission distance, the line-of-sight (LOS) signal is dominant, and thus the Rician fading is used to model the channel [7]. Specifically, $h_k$ is expressed as

$$h_k = \sqrt{\frac{K_R}{1 + K_R}} h_k^{\text{LOS}} + \sqrt{\frac{K_R}{1 + K_R}} h_k^{\text{NLOS}}. \quad (39)$$

where $h_k^{\text{LOS}} \in \mathbb{C}^{N_t \times 1}$ and $h_k^{\text{NLOS}} \in \mathbb{C}^{N_t \times 1}$ denotes the LOS deterministic component and the Rayleigh fading component with each element being a CSCG random variable with zero mean and covariance of $\sigma^2 = 40$ dB, respectively, and $K_R$ is the Rician factor set to 5 dB. Note that the far-field uniform linear antenna array model is adopted for the LOS component, i.e., $h_k^{\text{LOS}} = \sqrt{10^{-2/1} e^{j\phi_1} e^{j\phi_2} \ldots e^{j\phi_{N_t-1} d + j\phi_1}}$ with $\phi_k = -2\pi d \sin(\theta_k)$, where $d$ is the spacing between successive antenna elements at BS, $\mu$ is the carrier wavelength, and $\phi_k$ is the direction of user $k$ to the BS. We set $d = \frac{\lambda}{2}$ and $\{\phi_1, \phi_2, \phi_3, \phi_4\} = \{-30^\circ, -60^\circ, 60^\circ, 30^\circ\}$ as in [7]. In simulations, the constant circuitry power $P_c$ varies for different antenna setups. Specifically, $P_c$ is set according to [37]

$$P_c = N_t (P_{\text{DAC}} + P_{\text{mix}} + P_{\text{filt}}) + 2P_{\text{SN}} + K (P_{\text{LNA}} + P_{\text{mix}} + P_{\text{IFA}} + P_{\text{filt}} + P_{\text{ADC}}) \quad (40)$$

where $P_{\text{DAC}}, P_{\text{mix}}, P_{\text{filt}}, P_{\text{SN}}, P_{\text{LNA}}, P_{\text{IFA}}, P_{\text{filt}}$, and $P_{\text{ADC}}$ denote the power consumption of the digital to analog conversion (DAC), the mixer, the active filters at the transmitter side, the frequency synthesizer, the low-noise amplifier (LNA), the intermediate frequency amplifier (IFA), the active filters at the receiver side, and the analog to digital conversion (ADC), respectively. Table III lists all parameters used for computing $P_c$, where the values are the same as in [37] except $P_{\text{DAC}}$ and $P_{\text{SN}}$. Every component of the constant circuitry power $P_c$ is detailed in [37]. However, other type of power consumption, e.g., CSI acquisition, can be also taken into consideration by including them in $P_c$ and the proposed energy-efficient design approach still works.
TABLE III
SYSTEM PARAMETERS FOR COMPUTING $P_r$

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>The mixer $P_{mix}$</td>
<td>30.3mW</td>
</tr>
<tr>
<td>The frequency synthesizer $P_{syn}$</td>
<td>50mW</td>
</tr>
<tr>
<td>The active filter $P_{T_{tri}}$</td>
<td>2.5mW</td>
</tr>
<tr>
<td>The low-noise amplifier $P_{LNA}$</td>
<td>20mW</td>
</tr>
<tr>
<td>The intermediate frequency amplifier $P_{IFA}$</td>
<td>3mW</td>
</tr>
<tr>
<td>The ADC $P_{ADC}$</td>
<td>6.75mW</td>
</tr>
<tr>
<td>The DAC $P_{DAC}$</td>
<td>15.44mW</td>
</tr>
</tbody>
</table>

$P_{ADC}$ which are estimated according to the model introduced in [38] with channel bandwidth $W = 15$ kHz.

Firstly, we study the convergence performance of the proposed algorithms. Fig. 2 illustrates the convergence performance of the Dinkelbach method (i.e., Algorithm 2) for ten random problem instances of (15). It is observed that Algorithm 2 can always converge very quickly within only several iterations. We also examine the convergence performance of the LR method (i.e., Algorithm 1). It is seen from simulations that Algorithm 1 can also converge very quickly, as shown in Fig. 3, where both the primal objective values (i.e., (13)) and the dual objective values (i.e., (15)) at feasible iterations\(^{10}\) are presented. From Fig. 3, it can be also observed that the primal objective value coincides with the dual objective value as the iteration proceeds, implying that the LR method can achieve optimal solutions. This is due to the fact that unique solutions are always observed for problem (15) in our simulations (cf. Remark 3.2).

Secondly, we examine the performance of the LR method and the low complexity algorithm in terms of the achieved average energy efficiency over 100000 random channel realizations. Fig. 4 depicts that the average system energy efficiency improves with the total transmission power, $P_{t_{total}}$, under different SINR targets when $\gamma = 16$ dB, and $1.97 \times 10^8$ Bits/Joule when $\gamma = 20$ dB. Fig. 4 that, the LR method coincides with the upper bound, implying again that the LR method can achieve optimal solutions. Furthermore, it is seen from Fig. 4 that (also from Fig. 6), the performance of the low complexity method could be very close to that of the LR method and particularly coincides with the latter when $P_{t_{total}}$ exceeds 30 dBm (cf. Remark 5.1). In addition, one can see that better energy efficiency performance can be obtained in the scenario of less stringent QoS targets. Similar observations can be made from Fig. 5 which illustrates the average system energy efficiency versus the total transmission power, $P_{t_{total}}$, for different EH targets when $\gamma = 15$ dB. These observations confirm that there exists a tradeoff between the system energy efficiency and the QoS levels achieved by users.

For comparison, we also provide the average energy efficiency performance of the power minimization method [7] as a

\(^{10}\)A feasible iteration refers to $\{p_v(\lambda), p_u(\lambda)\}$ that is feasible to problem (13).
Fig. 5. Average energy efficiency Vs. the total transmission power for different EH targets. For all $P_{\text{total}}$, the average energy efficiency achieved by the power minimization method [7] is 2.91 when $\epsilon = -25$ dBm, and 0.78 x 10^5 Bts/Joule when $\epsilon = -15$ dBm.

Fig. 6. Average energy efficiency Vs. the number of transmit antennas. It can be seen (also from Fig. 6 below) that the proposed methods have better average energy efficiency than the power minimization method.

VII. CONCLUSION

In this paper, we have studied EE-based joint ZF beamforming and receive power splitting for multi-user MISO SWIPT systems. Based on a simple reformulation of the total power constraint, the Lagrangian relaxation method coupled with the Dinkelbach method is proposed to handle the non-convex fractional EE maximization problem. It is proven that the LR method can guarantee at least a feasible solution by slightly adjusting the optimal dual variable if necessary. A low complexity joint transceiver design method has been also proposed for EE maximization. Simulation results validate the superior energy efficiency performance of the proposed solutions. Lastly, we remark that the proposed LR method cannot be extended to the multi-user interference case (i.e., solving problem (9) without the ZF constraints) due to the inseparable SINR and EH constraints and thus another techniques should be investigated to address the energy efficiency optimization problem with multi-user interference. In addition, it has been shown in simulations that antenna selection could improve the system energy efficiency of multi-user MISO SWIPT systems with large-scale antenna arrays at BS and thus it is worthy of studying in the future.

APPENDIX A

THE PROOF OF PROPOSITION 2.1

First, letting $v_k = \sqrt{p_k |\mathbf{v}_k|}$ in (9) with $|\mathbf{v}_k| = 1$, we recast problem (9) as

$$\max_{\{p_k, \mathbf{v}_k, \rho_k\}} \sum_{k=1}^{K} \log \left( 1 + \frac{p_k |\mathbf{h}_k^H \mathbf{v}_k|^2}{\rho_k \sigma_k^2 + \delta_k} \right)$$

s.t. $\rho_k \mathbf{h}_k^H \mathbf{v}_k = 0, \forall k$, $|\mathbf{v}_k| = 1, \forall k$,

$$\delta_k (1 - \rho_k) \left( p_k |\mathbf{h}_k^H \mathbf{v}_k|^2 + \sigma_k^2 \right) \geq \epsilon_k, \forall k,$$

$$\mathbf{H}_k^H \mathbf{v}_k = 0, \forall k,$$

$$\sum_{k=1}^{K} p_k \leq P_{\text{total}},$$

$$0 \leq \rho_k \leq 1, \forall k.$$
where \( \tilde{E}_k(\tilde{v}_k, \rho_k, \rho_k) \triangleq \zeta_k(1 - \rho_k) \left( \rho_k |h_k^H \tilde{v}_k| ^2 + \sigma_k^2 \right) \); the interference terms in (41a)–(41c) have been canceled by applying the ZF conditions (41d). Observing that increasing the values of the terms \( h_k^H \tilde{v}_k \) will increase the objective function while maintaining the constraints (41b) and (41c), we infer that the optimal \( \tilde{v}_k \)'s of problem (41) are obtained when \( h_k^H \tilde{v}_k \)'s are maximized subject to the constraints (41d) and (41e), i.e., when the following problem

\[
\begin{align*}
\max_{\tilde{v}_k} & \quad h_k^H \tilde{v}_k^2 \\
\text{s.t.} & \quad |\tilde{v}_k|^2 - 1, \quad H_k^H \tilde{v}_k = 0,
\end{align*}
\]

is solved for all \( k \). Note that, according to [7, Prop. 5.1], \( \tilde{v}_k \) (cf. (10)) is an optimal solution to problem (42) and \( g_k \) (cf. (10)) is the corresponding optimal value. As a direct result, problem (41) can be simplified as (11). This completes the proof.

**APPENDIX B**

**THE PROOF OF LEMMA 2.1**

Clearly, problem (11) is feasible if and only if the optimal value of the following problem

\[
\begin{align*}
\min_{\{p_k, \rho_k\}} & \quad \sum_{k=1}^{K} p_k \\
\text{s.t.} & \quad \frac{p_k \rho_k g_k}{\rho_k \sigma_k^2 + \sigma_k^2} \geq \gamma_k, \forall k, \\
& \quad \zeta_k(1 - \rho_k) \left( \rho_k g_k + \sigma_k^2 \right) \geq \epsilon_k, \forall k, \\
& \quad 0 \leq \rho_k \leq 1, \forall k.
\end{align*}
\]

is not greater than \( P_{\text{total}} \). Note that problem (43) is separable and thus we consider its \( k \)-th subproblem only

\[
\begin{align*}
\min_{p_k, \rho_k} & \quad p_k \\
\text{s.t.} & \quad \frac{p_k \rho_k g_k}{\rho_k \sigma_k^2 + \sigma_k^2} \geq \gamma_k, \\
& \quad \zeta_k(1 - \rho_k) \left( \rho_k g_k + \sigma_k^2 \right) \geq \epsilon_k, \\
& \quad 0 \leq \rho_k \leq 1.
\end{align*}
\]

By the definition of \( l_1(\rho_k) \) and \( l_2(\rho_k) \), it can be shown that the first two constraints of problem (44) are respectively equivalent to \( p_k \geq l_1(\rho_k) \) and \( p_k \geq l_2(\rho_k) \). By combing these two constraints, we rewrite problem (44) equivalently as

\[
\begin{align*}
\min_{p_k, \rho_k} & \quad p_k \\
\text{s.t.} & \quad p_k \geq \max\{l_1(\rho_k), l_2(\rho_k)\}, \\
& \quad 0 \leq \rho_k \leq 1.
\end{align*}
\]

It is seen that the minimum \( p_k \) is attained when the feasible set of \( p_k \) becomes largest, that is, the optimal \( p_k \) must minimize \( \max\{l_1(\rho_k), l_2(\rho_k)\} \) in order to make the feasible set of \( p_k \) as large as possible. Hence, the optimal value of problem (45) is just \( p_{\text{min}} \), in terms of the definition of \( p_{\text{min}} \), and the optimal value of problem (43) is \( \sum_{k=1}^{K} p_k \). This completes the proof.

**APPENDIX C**

**THE PROOF OF PROPOSITION 3.1**

The first part follows directly from [36, Prop. 3.3.4]. Thus we prove the second part only. First, we show by contradiction that, \( d(\lambda) \geq d(\lambda), \forall \lambda > \lambda^* \). Assume the contrary, i.e., \( d(\lambda) < d(\lambda) \) for some \( \lambda > \lambda^* \). There must exist \( t \in (0, 1) \) such that \( \lambda = t\lambda^* + (1 - t)\lambda^* \). Hence, by the convexity of \( d(\lambda) \) and using the fact that \( d(\lambda) \geq d(\lambda^*) \), we have

\[
d(\lambda) < t d(\lambda^*) + (1 - t)d(\lambda) \leq d(\lambda)
\]

implying a contradiction. Hence, we have \( d(\lambda) > d(\lambda) \), \( \forall \lambda > \lambda^* \). This implies that any subgradient of \( d(\lambda) \) is nonnegative for \( \lambda > \lambda^* \). Thus, by noting that

\[
P_{\text{total}} - \sum_{k=1}^{K} p_k(\lambda) \leq \sum_{k=1}^{K} p_k(\lambda)
\]

is a subgradient of \( d(\lambda) \), we have \( \sum_{k=1}^{K} p_k(\lambda) \leq P_{\text{total}} \) for \( \lambda > \lambda^* \). This completes the proof.

**REFERENCES**


Weiqiang Xu (M’09–SM’13) received the Ph.D. degree in control science and engineering from Zhejiang University, China, in 2003 and 2006, respectively.

He also was a postdoctoral researcher at Zhejiang University, and a visiting scholar at Columbia University, New York, NY, USA. He is currently a professor with the School of Information Science and Technology, Zhejiang Sci-Tech University, Hangzhou, China. His research interests include multicol networks, ad hoc networks, wireless sensor networks, wireless optical networks, congestion control, and smart grid.


Mingyi Hong received the Ph.D. degree in systems engineering from the University of Virginia in 2011. He has been a Research Assistant Professor with the Department of Electrical and Computer Engineering, University of Minnesota, Twin Cities, USA. He is currently an Assistant Professor with the Department of Industrial and Manufacturing Systems Engineering (IMSE), Iowa State University, Ames, USA. His research interests include signal processing, wireless communications, large-scale optimization and its applications in compressive sensing, complex networks, and high-dimensional data analysis.

Yunlong Cai (S’07–M’10) received the Ph.D. degree in electronic engineering from the University of York, York, U.K., in 2010.

From February 2010 to January 2011, he was a Postdoctoral Fellow at the Electronics and Communications Laboratory, Conservatoire National des Arts etMetiers (CNAM), Paris, France. Since February 2011, he has been with the Department of Information Science and Electronic Engineering, Zhejiang University, Hangzhou, China, where he is currently an Associate Professor. His research interests include spread spectrum communications, adaptive signal processing, multiuser detection, and multiple antenna systems.

Dr. Cai has served on the Technical Program Committee (TPC) as member in various international conferences including for example IEEE GLOBECOM, IEEE ISWCS, and ITGWSA. He has also served as the Publicity Co-Chair for the 5th International Conference on Wireless Communications and Signal Processing (WCSP).
Energy Efficiency Optimization for MISO SWIPT Systems With Zero-Forcing Beamforming

Qingjiang Shi, Cheng Peng, Weiqiang Xu, Senior Member, IEEE, Mingyi Hong, and Yunlong Cai, Member, IEEE

Abstract—This paper considers a power splitting based multi-user multiple-input-single-output (MISO) downlink system with simultaneous wireless information and power transfer, where each single antenna receiver splits the received signal into two streams of different power for decoding information and harvesting energy separately. Assuming that the most common zero-forcing (ZF) beamforming scheme is employed by the base station, we aim to maximize the system energy efficiency in bits per Joule by joint beamforming and power splitting under both the signal-to-interference-plus-noise ratio constraints and energy harvesting constraints. The energy efficiency optimization (EEO) problem is nonconvex and very hard to solve. In this paper, by exploiting the problem structure, we first simplify the EEO problem to a joint transmit power allocation and receive power splitting problem. Then, with a careful reformulation, we propose a Lagrangian relaxation (LR) method coupled with Dinkelbach method to address the simplified EEO problem, whilst devising a nearly closed-form solution for the subproblems involved in the Dinkelbach method. It is proven that the proposed LR method is optimum under some condition and can guarantee at least a feasible solution, which is a notable advantage over the existing methods. Besides, we develop a low complexity EEO algorithm by proportionally distributing the total power to users. Finally, numerical results validate the excellent efficiency of the proposed algorithms.

Index Terms—Dinkelbach method, Lagrangian relaxation method, multiuser MISO downlink, power splitting, wireless information and power transfer, zero-forcing beamforming.

Manuscript received February 13, 2015; revised July 12, 2015; accepted October 02, 2015. Date of publication October 26, 2015. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Eduard Jorswieck. The work of Q. Shi and W. Xu is supported by the National Nature Science Foundation of China under grant 61302076, 61374024, Zhejiang Provincial Natural Science Foundation of China under grant LQ13F010008, LR15F010002, the Science Foundation of Zhejiang Sci-Tech University (ZSTU) under grant 1203805Y, and the open project of Provincial Key Laboratory of Information Networks, Zhejiang, China. The work of Y. Cai is supported by NSFC under grant 61471319. The work of M. Hong is supported by NSF under grant CCF-1526078.

Q. Shi is with the School of Information Science and Technology, Zhejiang Sci-Tech University, Hangzhou 310018, China. He is also with the Provincial Key Laboratory of Information Networks, Zhejiang, China. (e-mail: qingjiangshi@gmail.com).

C. Peng and W. Xu are with the School of Information Science and Technology, Zhejiang Sci-Tech University, Hangzhou 310018, China. (e-mail: peng18ccheng@gmail.com; wx.xu@126.com).

M. Hong is with the Department of Industrial and Manufacturing Systems Engineering, Iowa State University, IA 50011 USA (e-mail: mingyi@iastate.edu).

Y. Cai is with the Department of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China (e-mail: ylcai@zju.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2015.2489603

I. INTRODUCTION

Currently, a variety of new applications of information and communication technologies have sprung up, which demand for high data rate communications with certain quality of service (QoS) guarantees and lead to a huge amount of energy consumption. On the other hand, traditional wireless devices are usually battery-powered with limited electricity capacity. Therefore, energy efficient design of communication systems incorporating energy harvesting technologies has recently drawn an upsurge of interest in both industrial and academia. Particularly, integrating radio frequency (RF) energy harvesting with traditional wireless information transmission has driven a new research area—(simultaneous) wireless information and power transfer ((S)WIPT). This necessitates a fundamental paradigm change in designing signal processing/communication/energy harvesting systems and requires for a unified study [1].

WIPT has been studied for various communication systems in different context [2]–[12]. While initial works in this field focused on point-to-point single antenna systems and studied tradeoffs between information rate and power transfer [2]–[4], most of the recent studies on WIPT focused on multi-antenna systems [5]–[12]. For example, as one of the pioneering research on multi-antenna WIPT systems, Zhang et al. [5] considered a MIMO information-energy broadcast system of three nodes (one transmitter, one information receiver and one energy receiver), and investigated the relevant rate-energy region and optimal transmission schemes with time switching and power splitting receivers. The works [6]–[12] investigated multi-user MISO WIPT systems. In [6], Xu et al. investigated the optimal information/energy beamforming strategy to achieve the maximum harvested energy for multi-user MISO WIPT systems with separated information/energy receivers. Based on the PS receiving scheme, Shi et al. [7] studied the optimal joint beamforming and power splitting (JBPS) to achieve the minimum transmission power of a multi-user MISO SWIPT downlink system subject to both signal-to-interference-plus-noise (SINR) constraints and energy harvesting (EH) constraints. A similar JBPS problem for multi-user MISO interference channel (IFC) was studied in [8], [9] where efficient suboptimal solutions based on conic programming relaxation methods were obtained. Multi-antenna WIPT interference channels were also considered in the works [10]–[12] where the researchers investigated the transmission strategy with energy harvesting constraints. In addition, WIPT has been studied in other channel setups such as relay channels [13]–[16], orthogonal frequency division multiplexing (OFDM) channels [17], [18], and physical layer security channels [19]–[22].
Almost all of the above works focused on WIPT system design with the goal of throughput maximization, transmission power minimization, or harvested power/energy maximization. However, high energy efficiency (EE), defined as the number of information bits delivered per unit energy (bits per Joule), will be pursued in the design of future green wireless communication systems [23]. While there are a great number of works on energy efficiency optimization (EEO) for traditional communication systems [23], energy efficiency optimization was also recently introduced to energy-efficient WIPT system design [25]–[30]. In [25], Ng et al. considered energy-efficiency based resource allocation for a single-antenna point-to-point OFDM WIPT system with a PS-based receiver. The EEO problem in [25] is complicated by the coupling of the PS ratio and the subcarrier power allocation. Based on Dinkelbach method [24], the authors of [25] proposed an iterative resource allocation algorithm in aid of one dimensional exhaustive search for the PS ratio. In [26], Ng et al. studied energy efficient joint user selection and power allocation for single-antenna multi-user downlink OFDM WIPT systems where only one user is selected to be an information receiver while the rest are energy receivers. Due to user selection, the EEO problem in [26] is a mixed integer programming problem and thus it is much more complicated than that in [25]. Resorting to linear relaxation for the discrete user selection variables, [26] proposed an iterative resource allocation algorithm for multi-user downlink OFDM WIPT systems. By extending [25], [26], the work [27] investigated energy efficiency based joint subcarrier, PS ratio, and power allocation for single-antenna multiuser downlink OFDM WIPT systems with PS-based receivers, and proposed a similar iterative resource allocation algorithm as in [25], [26]. In [28], Chen et al. considered a time division duplex (TDD) large-scale MISO WIPT system, where, a terminal $S_1$ equipped with a large-scale antenna array first transfers wireless power through energy beamforming to a single-antenna terminal $S_2$ in slot 1 and then the terminal $S_2$ transmits information to $S_1$ in slot 2 using the harvested energy. Based on Dinkelbach method, an energy-efficient resource allocation scheme is proposed in [28] by jointly optimizing transmit power and transfer duration. In [29], He et al. investigated energy efficiency optimization for multicell multiuser MIMO downlink WIPT systems with energy harvesting constraints only. They assumed in [29] that the PS ratios are equal for all users and then proposed a two-tier suboptimal algorithm to tackle an equivalent problem of their EEO problem, i.e., iteratively adjusting the common PS ratio in the outer loop while optimizing the precoders with fixed PS ratios using block coordinate descent method in the inner loop. Different from [25]–[29], the work [30] studied the downlink and uplink energy efficiency tradeoff in TDD multi-user OFDMA systems where the users are equipped with PS-based receivers, respectively. The established multi-objective resource allocation problem is also a mixed integer programming. Through relaxation and transformation, Xiong et al. [30] developed a near-optimal resource allocation strategy that approaches the Pareto optimal tradeoff performance.

In this paper, we consider a multi-user MISO downlink SWIPT system as in [7], where a base station transmits radio signals to multiple single antenna receivers that can split the received signal into two streams of different power for decoding information and harvesting energy separately. Different from [7] where transmission power is minimized with QoS guarantee, this paper studies energy efficiency optimization by joint transmit beamforming and receive power splitting. Specifically, we assume that the most common zero-forcing (ZF) beamforming scheme is employed by the system and aim to maximize the system energy efficiency under both SINR constraints and EH constraints. The established EEO problem is more complicated than the transmission power minimization problem in [7] due to the highly nonconvex fractional EE objective function. In this paper, by exploiting the problem structure, we first simplify the EEO problem to a joint transmit power allocation and receive power splitting problem. With a tacitual reformulation, we then propose a Lagrangian relaxation (LR) method coupled with the Dinkelbach method to address the simplified EEO problem, whilst devising a nearly closed-form solution for the subproblems involved in the Dinkelbach method. We prove that the proposed LR method is optimum under some sufficient optimality condition\(^1\) and can guarantee at least a feasible solution, which is a notable advantage over the methods used in [25]–[28] (See Prop. 3.1 and Remark 3.1 below). Besides, by proportionally distributing the total power to each user, we develop a low complexity EEO algorithm. Finally, numerical results verify that the proposed algorithms are very efficient in terms of both the convergence performance and the achieved energy efficiency performance. In particular, simulation results show that the LR method can achieve the optimum energy efficiency. In addition, \textit{it is emphasized that} the proposed low complexity algorithm can globally solve the EEO problem of multi-user interference channels with PS-type WIPT, which is also a contribution of this paper.

The remainder of this paper is organized as follows. In the next section, we describe the system model and state the EE maximization problem formulation. Section III presents Lagrangian relaxation method for EE optimization while Section IV presents the Dinkelbach method for calculating the dual function of the EE maximization problem. A low complexity EE optimization method is proposed in Section V. Section VII provides numerical examples while Section VI concludes the paper.

\textbf{Notations:} scalars are denoted by lower-case letters; boldface lower-case letters are used for vectors, while boldface upper-case letters are for matrices. $\textbf{A}^H$ denote the conjugate transpose of a complex matrix $\textbf{A}$. $\|\textbf{z}\|$ denotes the Euclidean norm of a complex vector $\textbf{z}$, while $a$ denotes the absolute value of a complex scalar $a$. For a real function $f(x)$, $\frac{df(x)}{dx}$ denotes its first-order derivative at $x$. The distribution of a circularly symmetric complex Gaussian (CSCG) random vector with mean $\textbf{m}$ and covariance matrix $\textbf{C}$ is denoted by $\mathcal{CN}(\textbf{m}, \textbf{C})$, and ‘$\sim$’ stands for ‘distributed as’. Finally, $\textbf{C}^{m\times n}$ denotes the space of $m \times n$ complex matrices.

\section{System Model and Problem Formulation}

We first illustrate the system model and then give the problem formulation with some simplification.

\(^1\)Unfortunately, we are not able to show in theory how often the sufficient optimality condition is met. And strictly speaking, we have not yet theoretically proven the optimality of the LR method. However, it is found in simulations that the LR method can achieve the optimal energy efficiency performance.
A. System Description

As in [7], we consider a multi-user MISO downlink SWIPT system where one base station (BS) equipped with $N_t > 1$ antennas serves $K$ single-antenna users over a given frequency band with bandwidth $W$, as shown in Fig. 1. The system considered here differs from the traditional MIMO downlink system in that each receiver splits the received signal into two independent streams: one stream is used for information decoding (ID) while the other stream is used for energy harvesting (EH). We assume that linear beamforming is employed by the BS to transmit signals to users. Thus the complex baseband transmitted signal at BS is expressed as

$$x = \sum_{k=1}^{K} \mathbf{v}_k \beta_k \tag{1}$$

where $s_k \sim \mathcal{CN}(0,1)$ denotes the transmitted data symbol and $\mathbf{v}_k$ denotes the transmit beamforming vector for user $k$.

Assume flat fading for channels between the BS and all users and let $\mathbf{h}_k$ denote the conjugate channel vector between the BS and user $k$. Then the received signal at user $k$ before power splitting is given by

$$y_k = \mathbf{h}_k^H x + n_k, \quad k = 1, 2, \ldots, K \tag{2}$$

where $n_k \sim \mathcal{CN}(0,\sigma_n^2)$ denotes the antenna noise at user $k$.

After power splitting with ratio $\rho_k$, the signal for information decoding at user $k$ is expressed as

$$y_{k}^{\text{ID}} = \sqrt{\rho_k} \left( \mathbf{h}_k^H \sum_{j=1}^{K} \mathbf{v}_j s_j + n_k \right) + z_k, \quad \forall k \tag{3}$$

where $z_k \sim \mathcal{CN}(0,\delta_k^2)$ is the additional noise introduced by the ID at user $k$, and the signal for energy harvesting at user $k$ is expressed as

$$y_{k}^{\text{EH}} = \sqrt{1 - \rho_k} \left( \mathbf{h}_k^H \sum_{j=1}^{K} \mathbf{v}_j s_j + n_k \right), \quad \forall k. \tag{4}$$

Accordingly, the SINR for ID at user $k$ is given by

$$\text{SINR}_k = \frac{\rho_k |\mathbf{h}_k^H \mathbf{v}_k|^2}{\rho_k \sum_{j \neq k} |\mathbf{h}_k^H \mathbf{v}_j|^2 + \rho_k \sigma_n^2 + \delta_k^2}, \quad \forall k \tag{5}$$

and the harvested power by the EH of user $k$ is given by

$$E_k = \zeta_k (1 - \rho_k) \left( \sum_{j=1}^{K} |\mathbf{h}_k^H \mathbf{v}_j|^2 + \sigma_k^2 \right), \quad \forall k \tag{6}$$

where it is assumed that all the noise components and symbols are independent from each other, $\zeta_k \in (0,1]$ denotes the energy conversion efficiency at the EH of user $k$.

B. Problem Statement

We study the EEO problem for the considered system. EE is equivalently defined as the ratio between the system sum rate and the system power consumption [23]. In the following, we first describe the two components of energy efficiency—sum rate and power consumption, and then give the problem formulation with some simplification.

Assuming that perfect channel state information (CSI) is available, the data rate of user $k$ is given by

$$R_k = W \log(1 + \text{SINR}_k), \quad \forall k. \tag{7}$$

Thus the system sum rate is

$$R = \sum_{k=1}^{K} R_k. \tag{8}$$

Now we consider the system power consumption. In general, RF energy harvesting generates a very small amount of energy and thus may not contribute too much to the system energy efficiency. However, it is not difficult to see that the energy efficiency of a wireless communication system can be indeed improved by energy harvesting, since the system power consumption is counteracted by the harvested power. Hence, as in [27], we take the harvested power into consideration in the formulation of the system power consumption (hence, the energy efficiency formulation for the multi-user MISO SWIPT system).

Specifically, the total system power consumption is expressed as follows

$$P_s = \vartheta \sum_{k=1}^{K} ||\mathbf{v}_k||^2 + P_c - \sum_{k=1}^{K} E_k. \tag{9}$$

where the first term $\vartheta \sum_{k=1}^{K} ||\mathbf{v}_k||^2$ represents the power dissipation in the power amplifier of the transmitter ($\vartheta \geq 1$ is the power inefficiency of the amplifier), the second term $P_c$, which is independent of the first term, stands for the constant power consumption of the transceivers induced mainly by signal processing (it will be elaborated in Section VI), and the last term represents the harvested power of the $K$ receivers.

Given the expression of the system sum rate and power consumption, we are ready to state the problem formulation. The goal of this paper is to do joint transceiver design so that the system energy efficiency, defined by [27], is maximized while meeting two kinds of QoS constraints—SINR and EH constraints. Particularly, to simplify our transceiver design, we assume $N_t \geq K$ and consider the most common beamforming scheme—zero-forcing (ZF) processing, which nulls the multiuser interference signals. Mathematically, the problem is

\[ \text{maximize} \quad \frac{R}{P_s} \]
equivalently formulated as follows (by neglecting the constant W)

\[
\max_{\{w_k, p_k\}} \sum_{k=1}^{K} R_k
\]

\[
\text{s.t.} \quad \frac{\rho_k \| h_k^T w_k \|^2}{\rho_k \| h_k^T v_j \|^2 + \rho_k \sigma_k^2 + \delta_k^2} \geq \gamma_k, \forall k, \quad (9a)
\]

\[
\zeta_k (1 - \rho_k) \left( \sum_{j=1}^{K} h_k^H v_j \|^2 + \sigma_k^2 \right) \geq \epsilon_k, \forall k, \quad (9c)
\]

\[
\| h_k^H w_k \| = 0, \forall k, \quad (9d)
\]

\[
\sum_{k=1}^{K} \| v_k \|^2 \leq P_{\text{total}}, \quad (9e)
\]

\[
0 \leq \rho_k \leq 1, \forall k, \quad (9f)
\]

where \( h_k \triangleq [h_1, \ldots, h_{k-1}, h_{k+1}, \ldots, h_K] \), \( \forall k \), \( E_k \) and \( R_k \) are functions of \( (w_k, \rho_k) \) (cf. (6) and (7)). (9b) and (9c) are respectively the SINR and EH constraints with \( \gamma_k \)'s and \( \epsilon_k \)'s being the corresponding QoS targets, (9d) are the constraints induced by the ZF beamforming which nulls the interference signal, (9e) is the total power constraint with \( P_{\text{total}} \) being the allowed maximum transmission power, while (9f) are the inherent constraints for PS ratios. Obviously, it is difficult to solve problem (9) since the problem has a nonconcave objective function and a nonconvex constraint set. However, we can simplify the problem by exploring the ZF constraints (9d) and exploiting the problem structure. Specifically, with the following definitions:

\[
u_k \triangleq \frac{U_k U_k^H h_k}{\| U_k U_k^H h_k \|} \quad \text{and} \quad g_k \triangleq h_k^H v_k, \forall k, \quad (10)
\]

where \( U_k \) denotes the orthogonal basis of the null space of \( h_k^H \), we have the following proposition.

**Proposition 2.1:** Problem (9) can be simplified as follows

\[
\max_{\{p_k, \rho_k\}} \sum_{k=1}^{K} \log \left( 1 + \frac{p_k g_k + \sigma_k^2}{\rho_k \sigma_k^2 + \delta_k^2} \right)
\]

\[
\text{s.t.} \quad \frac{p_k \rho_k g_k}{\rho_k \sigma_k^2 + \delta_k^2} \geq \gamma_k, \forall k,
\]

\[
\zeta_k (1 - \rho_k) \left( p_k g_k + \sigma_k^2 \right) \geq \epsilon_k, \forall k,
\]

\[
\sum_{k=1}^{K} p_k \leq P_{\text{total}},
\]

\[
0 \leq \rho_k \leq 1, \forall k. \quad (11)
\]

where \( v_k = \sqrt{p_k} \nu_k \), \( k = 1, 2, \ldots, K \).

**Proof:** Please see Appendix A.

Proposition 2.1 suggests that the optimal solution to problem (9) can be obtained by solving problem (11). Hence, the main task of the rest of this paper is to solve problem (11). Before demonstrating our solution, we study the feasibility of problem (11). Lemma 2.1 shows a sufficient and necessary condition for the feasibility of problem (11).

**Lemma 2.1:** Define \( l_1(\rho_k) = \frac{\gamma_k (p_k g_k + \sigma_k^2)}{\rho_k \sigma_k^2 + \delta_k^2} \) and \( l_2(\rho_k) = \frac{\gamma_k (p_k g_k + \sigma_k^2)}{\rho_k \sigma_k^2 + \delta_k^2} \). Furthermore, define \( p_{k, \text{min}} = \min_{0 \leq p_k \leq 1} \max\{l_1(\rho_k), l_2(\rho_k)\}, k = 1, 2, \ldots, K \). Problem (11) is feasible if and only if \( \sum_{k=1}^{K} p_{k, \text{min}} \leq P_{\text{total}} \).

**Proof:** Please see Appendix B.

Note that \( p_{k, \text{min}} \) is attained when \( l_1(\rho_k) = l_2(\rho_k) \), which can be easily obtained by solving a quadratic equation. Hence, Lemma 2.1 indicates that the feasibility of problem (11) (equivalently (9)) can be easily verified. Without loss of generality, in the rest of this paper we assume that problem (11) is feasible, i.e., \( \sum_{k=1}^{K} p_{k, \text{min}} \leq P_{\text{total}} \).

**III. LAGRANGIAN RELAXATION METHOD FOR ENERGY EFFICIENCY OPTIMIZATION**

For notational convenience, we define \( E_k(p_k, \rho_k) \triangleq \zeta_k (1 - \rho_k) (p_k g_k + \sigma_k^2) \) and \( R_k(p_k, \rho_k) \triangleq \log \left( 1 + \frac{p_k g_k + \sigma_k^2}{\rho_k \sigma_k^2 + \delta_k^2} \right) \). Thus problem (11) is compactly written as

\[
\max_{\{p_k, \rho_k\}} \sum_{k=1}^{K} R_k(p_k, \rho_k)
\]

\[
\text{s.t.} \quad \phi \sum_{k=1}^{K} p_k + P_c - \sum_{k=1}^{K} E_k(p_k, \rho_k)
\]

\[
\zeta_k (1 - \rho_k) (p_k g_k + \sigma_k^2) \geq \epsilon_k, \forall k,
\]

\[
\sum_{k=1}^{K} p_k \leq P_{\text{total}},
\]

\[
0 \leq \rho_k \leq 1, \forall k. \quad (12)
\]

Since problem (12) is a nonconvex nonlinear fractional programming, it is difficult to directly solve it. To develop an efficient algorithm, we propose using Lagrangian relaxation method to address problem (12).

The LR method is a popular technique for solving difficult optimization problems, which can provide not only bounds (i.e., weak duality) but also good suboptimal solutions with some heuristic [31]. The basic idea of the LR method is to build some complicated constraints into objective functions and solve the dual problems of nonconvex problems instead of the original problems. In the following, we will resort to the LR method to address problem (12).

It is readily seen that, without the total power constraint, the constraint set of problem (12) is separable over user \( k = 1, 2, \ldots, K \). Hence, we try to tackle the total power constraint to make the problem more tractable. However, directly building the total power constraint into the objective function of problem (12) with a Lagrange multiplier will make the resultant dual problem hard to solve. This motivates us to tactfully reformulate problem (12) so that the corresponding dual problem can be easily solved. Note that, the equivalent system
power consumption \( \vartheta \sum_{k=1}^{K} p_k + P_c - \sum_{k=1}^{K} E_k(p_k, \rho_k) \) is positive\(^4\). Thus, the total power constraint is equivalent to

\[
\frac{\sum_{k=1}^{K} p_k - P_{\text{total}}}{\vartheta \sum_{k=1}^{K} p_k + P_c - \sum_{k=1}^{K} E_k(p_k, \rho_k)} \leq 0.
\]

It follows that problem (12) is equivalent to

\[
\max_{\{p_k, \rho_k\}} \sum_{k=1}^{K} R_k(p_k, \rho_k)
\]

\[
\text{s.t. } \frac{p_k \rho_k g_k}{\rho_k \sigma_k^2 + \delta_k^2} \geq \gamma_k, \ \forall k,
\]

\[
\frac{\zeta_k (1 - \rho_k)}{\rho_k \sigma_k^2 + \delta_k^2} \geq e_k, \ \forall k,
\]

\[
0 \leq \rho_k \leq 1, \ \forall k.
\]

\[
\frac{\sum_{k=1}^{K} p_k - P_{\text{total}}}{\vartheta \sum_{k=1}^{K} p_k + P_c - \sum_{k=1}^{K} E_k(p_k, \rho_k)} \leq 0. \tag{13}
\]

Problem (13) appears to be harder than problem (12). However, it will be clear later that the dual function of problem (13) can be obtained in nearly closed-form and consequently the corresponding dual problem can be easily solved.

Now, we are ready to use the LR method to address problem (13) (equivalently, (12)). To this end, we introduce Lagrange multiplier \( \lambda \) to the last constraint of problem (13) and define the partial Lagrangian associated with problem (13) as [32]

\[
L(\{p_k\}, \{\rho_k\}, \lambda) \triangleq \sum_{k=1}^{K} R_k(p_k, \rho_k)
\]

\[
\text{s.t. } \frac{p_k \rho_k g_k}{\rho_k \sigma_k^2 + \delta_k^2} \geq \gamma_k, \ \forall k,
\]

\[
\frac{\zeta_k (1 - \rho_k)}{\rho_k \sigma_k^2 + \delta_k^2} \geq e_k, \ \forall k,
\]

\[
0 \leq \rho_k \leq 1, \ \forall k.
\]

(14)

With the above partial Lagrangian, the dual function, denoted by \( d(\lambda) \), is written as [32]

\[
d(\lambda) \triangleq \max_{\{p_k, \rho_k\}} L(\{p_k\}, \{\rho_k\}, \lambda)
\]

\[
\text{s.t. } \frac{p_k \rho_k g_k}{\rho_k \sigma_k^2 + \delta_k^2} \geq \gamma_k, \ \forall k,
\]

\[
\frac{\zeta_k (1 - \rho_k)}{\rho_k \sigma_k^2 + \delta_k^2} \geq e_k, \ \forall k,
\]

\[
0 \leq \rho_k \leq 1, \ \forall k.
\]

(15)

Let \( p_k(\lambda) \) and \( \rho_k(\lambda) \) denote an optimal solution to problem (15). It is easily known that, when \( \lambda = 0 \), if \( p_k(0), k = 1, 2, \ldots, K \), satisfy the total power constraint, we obtain an optimal solution to problem (12), i.e., \( p_k(0) \) and \( \rho_k(0), k = 1, 2, \ldots, K \); otherwise, we need to augment \( \lambda \) to increase the dominance of the second term of the partial Lagrangian (14) and force \( p_k(\lambda)'s \) to satisfy the total power constraint. On the other hand, by weak duality [32], we have \( p^* < d(\lambda) \) for any \( \lambda > 0 \), where \( p^* \) is the optimal value of problem (12). Hence, to obtain a possibly tight bound, we need to find a nonnegative \( \lambda \) to minimize the dual function \( d(\lambda) \), i.e., solving the dual problem

\[
\min_{\lambda \geq 0} d(\lambda). \tag{16}
\]

It is noteworthy that the optimal dual solution \( \lambda \) may not lead to a feasible primal solution \( \{p_k(\lambda), \rho_k(\lambda)\} \). However, since \( d(\lambda) \) is a convex function and

\[
\frac{\sum_{k=1}^{K} p_k(\lambda) + P_c - \sum_{k=1}^{K} E_k(p_k(\lambda), \rho_k(\lambda))}{\vartheta \sum_{k=1}^{K} p_k + P_c - \sum_{k=1}^{K} E_k(p_k, \rho_k)}
\]

is a subgradient of \( d(\lambda) \) [33, pp. 12], we have the following proposition.

**Proposition 3.1:** Let \( \lambda^* \) be the optimal dual variable. Then

1) If \( p_k(\lambda^*) \) satisfies the total power constraint and \( \lambda^* \left( \sum_{k=1}^{K} p_k(\lambda^*) - P_{\text{total}} \right) = 0 \), \( \{p_k(\lambda^*), \rho_k(\lambda^*)\} \) is an optimal solution to problem (12);

2) For any \( \lambda > \lambda^* \), \( \{p_k(\lambda)\} \) satisfies the total power constraint, i.e., \( \sum_{k=1}^{K} p_k(\lambda) \leq P_{\text{total}} \).

**Proof:** See Appendix C.

According to Proposition 3.1, we infer that either an optimal solution or a feasible solution can be obtained by the LR method. The main task of the LR method is to solve the dual problem (16). This can be done by using Bisection method [32] with the aid of the subgradient of \( d(\lambda) \). We summarize the proposed LR method in Algorithm 1, where Steps 1–5 check whether or not \( \{p_k(0), \rho_k(0)\} \) solves problem (12) globally, Steps 6–16 represent the Bisection method which solves the dual problem (16) globally, and Steps 17–20, if necessary, are carried out to provide a feasible solution to problem (12).

Before proceeding to the next section, we make two remarks on Algorithm 1. The first remark points out the advantage of our

---

\(^4\)Note that the power amplifier inefficiency factor \( \vartheta \) is greater than 1 while all \( g_k's \) are much smaller than 1 due to large-scale path loss.
TABLE II

<table>
<thead>
<tr>
<th>Algorithm 2: The Dinkelbach Method for Problem (15).</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 initialize feasible $p_k$, $\rho_k$, $k = 1, 2, \ldots, K$</td>
</tr>
<tr>
<td>2 repeat</td>
</tr>
<tr>
<td>3 $\eta \leftarrow \frac{\sum_{k=1}^{K} R_k(p_k, \rho_k) - \lambda(\sum_{k=1}^{K} p_k - P_{\text{total}})}{p_k \sum_{k=1}^{K} p_k + P_{\text{total}} - \sum_{k=1}^{K} R_k(p_k, \rho_k)}$</td>
</tr>
<tr>
<td>4 update ${p_k, \rho_k}$ by solving problem (17)</td>
</tr>
<tr>
<td>5 until some termination condition is met</td>
</tr>
</tbody>
</table>

algorithm over the algorithms in [25]–[27], while the second remark indicates a sufficient condition under which the algorithm will arrive at optimal solutions.

Remark 3.1: Problem (15) will be solved by an iterative algorithm called Dinkelbach method [24]. Thus the main part of Algorithm 1 (cf. Steps 8–16) can be viewed as a two-tier algorithm, where the inner loop solves problem (15) using the Dinkelbach method while the outer loop performs the bisection algorithm [32]. Note that, the iterative algorithms in [25]–[27] can be also viewed as two-tier algorithms, where, however, the outer loop implements the Dinkelbach method while the inner loop solves the subproblem of subtractive form by using the dual-decomposition method [33]. Hence, Algorithm 1 is fundamentally different from the algorithms in [25]–[27]. Furthermore, since the subproblems of subtractive form in [25]–[27] are nonconvex and much complicated, the authors of [25]–[27] have not yet proven that zero duality gap must be guaranteed for their problems. Consequently, the dual decomposition algorithms in [25]–[27] may not arrive at the optimal (even not feasible) solutions to the EEO problems in theory. However, as shown in the second part of Proposition 3.1, our algorithm guarantees at least a feasible solution. Moreover, simulations later show that the proposed algorithm can achieve the optimal energy efficiency.

Remark 3.2: If problem (15) (or equivalently (17) below) has a unique solution, then the subgradient

$$P_{\text{total}} \sum_{k=1}^{K} p_k(\lambda)$$

$$\theta \sum_{k=1}^{K} p_k(\lambda) + P_c - \sum_{k=1}^{K} E_k(p_k, \rho_k)$$

is just the gradient of $d(\lambda)$. Hence, by the first order optimality condition of the dual problem, we have $P_{\text{total}} - \sum_{k=1}^{K} p_k(\lambda^*) = 0$ when a unique solution is obtained for (15). In other words, if problem (15) has a unique solution for $\lambda = \lambda^*$, the LR method comes up with an optimal solution to the primal problem (13) (equivalently (11)) according to Proposition 3.1.

IV. DINKELBACH METHOD FOR PROBLEM (15)

Problem (15) is a nonlinear fractional programming. It is clearly not a convex problem and hard to solve. However, it is observed that both the numerator and the denominator of the objective function, as well as the constraint set of problem (15), are separable across user $k = 1, 2, \ldots, K$. Hence, if we transform the objective function (i.e., the fractional form) into a numerator-denominator subtractive form, then the resultant problem can be decomposed into $K$ tractable subproblems. This is our basic idea for solving problem (15), realized by using the Dinkelbach method [24].

A. Dinkelbach Method

We propose using the Dinkelbach method, a popular technique for solving nonlinear fractional programmings, to solve problem (15). The following lemma lays a theoretical base for the Dinkelbach method.

Lemma 4.1: Define

$$\Omega_k \triangleq \left\{ (p_k, \rho_k) \right\} \frac{p_k \rho_k g_k}{\rho_k \sigma_k + \delta_k} \geq \gamma_k, \quad \zeta_k(1 - \rho_k) (p_k g_k + \sigma_k^2) > e_k, \quad 0 < p_k < 1 \right\}.$$ 

Then $\{p_k^*, \rho_k^*\}$ solves problem (15) if and only if it solves

$$\max_{(p_k, \rho_k) \in \Omega_k} \sum_{k=1}^{K} R_k(p_k, \rho_k) - \lambda \left( \sum_{k=1}^{K} p_k - P_{\text{total}} \right)$$

$$- \eta \left( \theta \sum_{k=1}^{K} p_k + P_c - \sum_{k=1}^{K} E_k(p_k, \rho_k) \right)$$

with $\eta - \eta^*$ given by

$$\eta^* \triangleq \sum_{k=1}^{K} R_k(p_k^*, \rho_k^*) - \lambda \left( \sum_{k=1}^{K} p_k^* - P_{\text{total}} \right)$$

$$\theta \sum_{k=1}^{K} p_k^* + P_c - \sum_{k=1}^{K} E_k(p_k^*, \rho_k^*)$$

Lemma 4.1 can be proven by following a similar approach as in [24]. Clearly, $\eta^*$ is just the maximum energy efficiency. Let $F(\eta)$ denote the optimal value of problem (17). The lemma implies that the maximum energy efficiency $\eta^*$ satisfies $F(\eta^*) = 0$. Note that, the Dinkelbach method is an iterative algorithm that generates a sequence of values of $\eta$ converging to the maximum energy efficiency $\eta^*$ monotonically such that $F(\eta^*) = 0$. We summarize the algorithm in Table II. Once problem (17) is globally solved in Step 4, Algorithm 2 will finally achieve an optimal solution to problem (15). The following subsection shows how to solve problem (17).

B. Solution to Problem (17)

Different from problem (15), problem (17) is separable. That is, problem (17) can be decomposed into $K$ subproblems with the $k$-th subproblem given by

$$\max_{p_k, \rho_k} R_k(p_k, \rho_k) - (\lambda + \eta \theta) p_k + \eta E_k(p_k, \rho_k)$$

s.t. $\frac{p_k \rho_k g_k}{\rho_k \sigma_k^2 + \delta_k^2} \geq \gamma_k, \quad \zeta_k(1 - \rho_k) (p_k g_k + \sigma_k^2) \geq e_k, \quad 0 \leq p_k \leq P_{\text{total}}, \quad 0 \leq \rho_k \leq 1$ (19)
where we have neglected some constant terms in the objective function. Clearly, problem (19) is nonconvex. However, it is readily seen that, both $R_k(p_k, \rho_k)$ and $E_k(p_k, \rho_k)$ are concave functions of $p_k$. Moreover, all the constraint functions related to $p_k$ are linear in $p_k$. Thus, once $\rho_k$ is fixed, problem (19) is a convex problem with respect to $p_k$. This implies that problem (19) can be solved by one dimensional exhaustive search. However, the exhaustive search method is not efficient. In what follows, we show how problem (19) can be efficiently solved.

First, we recast problem (19) as follows

$$\max_{\rho_k \in \mathcal{A}_k} R_k(p_k, \rho_k) - (\lambda + \eta \theta) p_k + \eta E_k(p_k, \rho_k)$$

s.t. $p_k \geq l_1(\rho_k);$ $p_k \geq l_2(\rho_k);$ $0 \leq p_k \leq P_{total};$ $0 \leq \rho_k \leq 1$ (20)

where $l_1(\rho_k)$ and $l_2(\rho_k)$ are defined in Lemma 2.1. Then, by denoting the constraint set of $p_k$ as $\mathcal{P}_k(\rho_k) \triangleq \{p_k | \max(l_1(\rho_k), l_2(\rho_k)) \leq p_k \leq P_{total}\}$, we rewrite problem (20) as the following equivalent two-tier maximization problem

$$\max_{0 \leq p_k \leq 1} \max_{p_k \in \mathcal{P}_k(\rho_k)} R_k(p_k, \rho_k) - (\lambda + \eta \theta) p_k + \eta E_k(p_k, \rho_k)$$ (21)

which sheds light on our solution to problem (19). As discussed above, for fixed $\rho_k$, the inner maximization problem of (21) is a convex problem. Moreover, since $R_k(p_k, \rho_k)$ is a strictly concave function of $p_k$, the inner maximization problem has a unique solution. Let $\hat{p}_k(\rho_k)$ denote the unique optimal solution to the inner maximization problem of (21). Then the optimal $\rho_k$ of problem (19) can be obtained by solving the following problem

$$\max_{0 \leq \rho_k \leq 1} \psi_k(\rho_k)$$ (22)

where $\psi_k(\rho_k) \triangleq R_k(p_k, \rho_k) - (\lambda + \eta \theta) p_k + \eta E_k(p_k, \rho_k) + \eta E_k(\hat{p}_k(\rho_k), \rho_k).$

Next, we solve problem (22) by considering all possible cases for $\hat{p}_k(\rho_k)$. The inner maximization problem of problem (21) for fixed $\rho_k$ can be explicitly written as

$$\max_{p_k} R_k(p_k, \rho_k) - (\lambda + \eta \theta) p_k + \eta E_k(p_k, \rho_k)$$

s.t. $\max(l_1(\rho_k), l_2(\rho_k)) \leq p_k \leq P_{total}.$ (23)

By checking the first order condition, we obtain the stationary point of the objective function of problem (23) as follows

$$\hat{p}_k^*(\rho_k) = \frac{1}{\lambda + \eta \theta - \eta \bar{E}(1 - \rho_k) g_k} \left( \frac{\sigma_k^2}{\rho_k g_k} - \frac{\delta_k^2}{g_k} \right).$$ (24)

Since problem (23) is a univariate convex problem with a bound constraint, its optimal value must be attained either on the boundary of the constraint or at the stationary point $\hat{p}_k^*(\rho_k)$ in terms of the relative magnitude of $\hat{p}_k^*(\rho_k)$, $\max(l_1(\rho_k), l_2(\rho_k))$, and $P_{total}$. Specifically, the optimal solution to problem (23), $p_k(\rho_k)$, can be expressed as follows

$$p_k(\rho_k) = \begin{cases} P_{total}, & \text{if } \hat{p}_k^*(\rho_k) \geq P_{total} \\
\max(1, \hat{p}_k^*(\rho_k)), & \text{if } P_{total} \geq \hat{p}_k^*(\rho_k) \geq (l_1(\rho_k), l_2(\rho_k)), \\
\hat{p}_k^*(\rho_k), & \text{if } (l_1(\rho_k), l_2(\rho_k)) \geq \hat{p}_k^*(\rho_k). \end{cases}$$ (25)

According to (25), problem (22) can be solved through first separately checking the above four cases and then keeping the best $\rho_k$ as the solution of problem (22). It can be shown that each condition in (25) can be reduced to one or several intervals of $\rho_k$. Hence, problem (22) under the above four conditions can be recast as optimization problems in the form of

$$\max_{\rho_k \in \mathcal{I}_k} \psi_k(\rho_k)$$ (26)

where $\mathcal{I}_k$ denotes the union of some intervals of $\rho_k$ (including the interval $0 \leq \rho_k \leq 1$). It is known that, for an optimization problem with interval constraints, its optimal solution must be either some endpoint of the intervals or some feasible stationary point of the objective function. Hence, regarding the optimal solution to problem (26), we have the following lemma.

**Lemma 4.2**: Suppose that $p_k(\rho_k)$ is differentiable. Then the optimal value of problem (26) must be attained either at the point that satisfies

$$\frac{d\psi_k(\rho_k)}{d\rho_k} = 0 \text{ and } \rho_k \in \mathcal{I}_k$$ (27)

or at some endpoint of $\mathcal{I}_k$.

Lemma 4.2 tells us that, besides the endpoints of $\mathcal{I}_k$, if we can find all possible $\rho_k$ s that satisfy condition (27), the optimal $\rho_k$ of problem (26) can be obtained by simply checking the objective function value $\psi_k(\rho_k)$. This is our basic idea for solving problem (26) (hence, (22) and (19)). In the following, we consider the above four cases one by one and particularly examine the corresponding first-order condition.

1) $p_k(\rho_k) = P_{total}$: In this case, we have

$$\psi_k(\rho_k) = \log \left( \frac{1 + \rho_k P_{total} g_k}{\rho_k \bar{g}^2_k + \delta_k^2} \right) - (\lambda + \eta \theta) P_{total} + \eta E_k(1 - \rho_k) g_k + \eta g_k.$$ (28)

Taking the derivative of $\psi_k(\rho_k)$ with respect to $\rho_k$, we obtain

$$\frac{d\psi_k(\rho_k)}{d\rho_k} = \frac{1}{1 + \frac{P_{total} g_k}{\rho_k \bar{g}^2_k + \delta_k^2}} \left( \frac{P_{total} g_k \delta_k^2}{\rho_k \bar{g}^2_k + \delta_k^2} - \eta \bar{E}_k (P_{total} g_k + \bar{g}_k) \right)$$

$- \frac{P_{total} g_k \delta_k^2}{\rho_k \bar{g}^2_k + \delta_k^2} \left( \frac{P_{total} g_k + \bar{g}_k}{\rho_k \bar{g}^2_k + \delta_k^2} \right) - \eta \bar{E}_k (P_{total} g_k + \bar{g}_k).$

Thus, $\frac{d\psi_k(\rho_k)}{d\rho_k} = 0$ implies that $\rho_k$ satisfies the following quadratic equation

$$a_1 \rho_k^2 + b_1 \rho_k + c_1 = 0$$ (29)

7The four conditions in (25) can be simplified as trivial or quadratic inequalities which correspond to some intervals for $\rho_k$.

8Quartic equations can be solved in nearly closed-form [35].
with $a_1 = \sigma_k^2 (P_{total} g_k + \sigma_k^2)$, $b_1 = \delta_k^2 (P_{total} g_k + 2\sigma_k^2)$, and $c_1 = \left( \frac{\delta_k^2}{\eta_k g_k} + \frac{\sigma_k^2}{\eta_k g_k} \right)$. In this case, we have

$$\psi_k(\rho_k) = \log \left(1 + \frac{\rho_k l_1(\rho_k) g_k}{\rho_k \sigma_k^2 + \delta_k^2} \right) - (\lambda + \eta\theta) l_1(\rho_k) + \eta_k (1 - \rho_k) [l_2(\rho_k) g_k + \sigma_k^2] \tag{30}$$

Let us recall $l_1(\rho_k) = \frac{\rho_k (\rho_k^2 + \delta_k^2)}{\rho_k \sigma_k^2 + \delta_k^2}$, which implies that

$$\log \left(1 + \frac{\rho_k l_1(\rho_k) g_k}{\rho_k \sigma_k^2 + \delta_k^2} \right)$$

is a constant and $\frac{d l_1(\rho_k)}{d \rho_k} = -\frac{\delta_k^2}{\rho_k \sigma_k^2 + \delta_k^2}$. Thus, we have

$$\frac{d \psi_k(\rho_k)}{d \rho_k} = -\left(\lambda + \eta\theta\right) \frac{d l_1(\rho_k)}{d \rho_k} + \eta_k (1 - \rho_k) \frac{d l_1(\rho_k)}{d \rho_k} + \eta_k \delta_k^2 \tag{31}$$

Furthermore, $\frac{d \psi_k(\rho_k)}{d \rho_k} = 0$ implies that $\rho_k$ satisfies the following quadratic equation

$$a_2 \rho_k^2 + c_2 = 0 \tag{32}$$

with $a_2 = -\eta_k g_k \sigma_k^2 (1 + \gamma_k)$ and $c_2 = (\lambda + \eta\theta - \eta_k g_k) \gamma_k \delta_k^2$. In this case, we have

$$\psi_k(\rho_k) = \log \left(1 + \frac{\rho_k l_2(\rho_k) g_k}{\rho_k \sigma_k^2 + \delta_k^2} \right) - (\lambda + \eta\theta) l_2(\rho_k) + \eta_k (1 - \rho_k) [l_2(\rho_k) g_k + \sigma_k^2] \tag{33}$$

Let us recall $l_2(\rho_k) = \frac{1}{s_k} \left( \frac{\rho_k \sigma_k^2 + \delta_k^2}{g_k (1 - \rho_k)} - \sigma_k^2 \right)$, which implies that

$$\eta_k (1 - \rho_k) [l_2(\rho_k) g_k + \sigma_k^2]$$

is a constant. Thus, we have

$$\frac{d \psi_k(\rho_k)}{d \rho_k} = -\left(\lambda + \eta\theta\right) \frac{d l_2(\rho_k)}{d \rho_k} + \eta_k (1 - \rho_k) \frac{d l_2(\rho_k)}{d \rho_k} + \eta_k \delta_k^2 \tag{34}$$

Equating the above equation to zero, i.e., $\frac{d \psi_k(\rho_k)}{d \rho_k} = 0$, leads to

$$\left(\frac{g_k l_2(\rho_k) + \rho_k g_k \frac{d l_2(\rho_k)}{d \rho_k}}{\rho_k \sigma_k^2 + \delta_k^2} \right) \frac{d l_2(\rho_k)}{d \rho_k} + \sigma_k^2 \rho_k l_2(\rho_k) g_k - (\lambda + \eta\theta) \left(\rho_k \sigma_k^2 + \delta_k^2 \right) \frac{d l_2(\rho_k)}{d \rho_k} = 0 \tag{35}$$

By plugging $l_2(\rho_k) = \frac{1}{s_k} \left( \frac{\rho_k \sigma_k^2 + \delta_k^2}{g_k (1 - \rho_k)} - \sigma_k^2 \right)$ into (34) and with some manipulations, we can finally reduce (34) to the following quartic equation

$$a_3 \rho_k^4 + b_3 \rho_k^3 + c_3 \rho_k^2 + d_3 \rho_k + e_3 = 0 \tag{36}$$

with $a_3 = -2 \delta_k^2 \sigma_k^2 g_k$, $b_3 = (6 \delta_k^2 \sigma_k^3 + 3 \delta_k^2 \sigma_k^2 + \delta_k^3 \sigma_k g_k)$, $c_3 = (6 g_k \delta_k^2 \sigma_k^2 - 3 g_k e_k \delta_k^3 - 3 g_k \delta_k^2 \delta_k^2 e_k^2 \lambda - e_k^2 \delta_k^2 \eta)$, $d_3 = -e_k \delta_k^2 \sigma_k^2 g_k (2 \delta_k^2 - \delta_k^2)$, and $e_3 = (g_k \delta_k^2 \sigma_k^2 - e_k \delta_k^2 \sigma_k^2 - e_k \delta_k^2 \delta_k^2 \eta - e_k \delta_k^2 \eta - \eta_k \delta_k^2 \eta)$. In this case, we have

$$\psi_k(\rho_k) = \log \left(1 + \frac{\rho_k l_3(\rho_k) g_k}{\rho_k \sigma_k^2 + \delta_k^2} \right) - (\lambda + \eta\theta) p_k(\rho_k) + \eta_k (1 - \rho_k) \left(\rho_k g_k + \sigma_k^2 g_k \right) \tag{37}$$

Similarly, we can finally reduce $\frac{d \psi_k(\rho_k)}{d \rho_k} = 0$ to the following quartic equation

$$a_4 \rho_k^4 + b_4 \rho_k^3 + c_4 \rho_k^2 + d_4 \rho_k + e_4 = 0 \tag{38}$$

with $a_4 = -g_k^2 \delta_k^2 \eta^2 \sigma_k^2$, $b_4 = -\delta_k^2 \delta_k^2 \eta^2 \sigma_k^2 \eta^2$, $c_4 = (\delta_k^2 \delta_k^2 \eta^2 + \delta_k^2 \delta_k \lambda \eta \sigma_k^2 + \eta \eta \lambda g_k^2 \sigma_k^2 \eta \sigma_k^2 \eta)$, $d_4 = (\delta_k^2 \delta_k^2 \eta^2 + \delta_k^2 \delta_k \lambda \eta \sigma_k^2 + \eta \eta \lambda g_k^2 \sigma_k^2 \eta \sigma_k^2 \eta)$, and $e_4 = (\delta_k^2 \delta_k^2 \eta^2 \sigma_k^2 + \delta_k^2 \delta_k \lambda \eta \sigma_k^2 + \eta \eta \lambda g_k^2 \sigma_k^2 \eta \sigma_k^2 \eta)$.

To sum up, we can obtain the candidates of the optimal $\rho_k$ by solving problem (26) in four cases, equivalently, first solving (29), (32), (35) and (37) and then checking the objective values of all the endpoints of $T_k$ as well as the obtained feasible stationary points. The best candidate with the maximum value of $\psi_k(\rho_k)$ is picked as the optimal $\rho_k$. Accordingly, we obtain the optimal $\rho_k$ in terms of (25). Therefore, all the subproblems of (17) can be solved with complexity of $O(1) \text{ in parallel}$. As a result, the worst-case complexity of Algorithm 1 is $O(T_D T_H)$, where $T_D$ and $T_H$ denote the number of iterations required by the Dinkelbach method and Bisection method, respectively. Generally, both the Dinkelbach method and Bisection method achieve convergence in about ten iterations (which...
is validated with numerical examples in Section VI). Thus Algorithm 1 is very efficient.

V. LOW COMPLEXITY ALGORITHM FOR PROBLEM (12)

In Section III, we have proposed Algorithm 1 to address the energy efficiency optimization problem (12). As argued in Remark 3.1, Algorithm 1 is in essence a two-tier iterative algorithm. In this section, we propose a suboptimal but single tier (hence, low complexity) iterative algorithm based on only the Dinkelbach method.

It is observed that, if the total power constraint of problem (12) is replaced with independent user power constraints, i.e., $p_k \leq P_k$, $\forall k$, the corresponding energy efficiency optimization problem, i.e.,

$$\max_{\{p_k, \rho_k\}} \sum_{k=1}^{K} \frac{R_k(p_k, \rho_k)}{p_k + P_c - \sum_{k=1}^{K} P_k(p_k, \rho_k)}$$

subject to

$$p_k\rho_k g_k \geq \gamma_k, \quad \forall k,$$

$$\zeta_k(1 - \rho_k) (p_k g_k + \sigma^2_k) \geq \epsilon_k, \quad \forall k,$$

$$0 \leq p_k \leq P_k, \quad \forall k,$$

$$0 \leq \rho_k \leq 1, \quad \forall k,$$

(38)
can be globally solved by using the Dinkelbach method in Section IV (without need of the LR method). Motivated by this observation, we instead consider solving problem (38) in order to achieve a low complexity suboptimal transceiver design. A direct way for achieving $K$ individual power constraints is just equally distributing the total power to $K$ users, i.e., $p_k \leq P_{total}/K$, $\forall k$. However, problem (38) with $K$ equally distributed power constraints has smaller feasible region than problem (12). Moreover, it may become even infeasible. To address this challenge while obtaining $K$ feasible individual power constraints, we proportionally distribute the total power to each user and have each $P_k$ proportional to $p_{k,m,n}$ (defined in (2.1)), i.e., let $P_k = P_{total}\sum_{k=1}^{K} p_{k,m,n}/\sum_{k=1}^{K} p_{k,m,n}$, $\forall k$. It is not difficult to see that problem (38) with such $H_k$'s must be feasible if problem (11) is feasible. Hence, our low complexity method just solves problem (38) with $P_k = P_{total}\sum_{k=1}^{K} p_{k,m,n}/\sum_{k=1}^{K} p_{k,m,n}$, $\forall k$. According to Section IV, it is readily known that problem (38) can be globally solved using the Dinkelbach method, leading to a single-tier iterative algorithm. Obviously, the proposed single-tier iterative algorithm has lower complexity (of $O(T_k)$) than Algorithm 1. Moreover, it will be seen from simulations that this low complexity algorithm could achieve a performance very close to that of Algorithm 1. Two important remarks are made as follows on the proposed low complexity algorithm.

Remark 5.1: If the optimal solution $p_k$ to problem (12) satisfies the individual power constraints of problem (38), we can obtain the optimal solution to problem (12) through solving problem (38). This means that the proposed low complexity algorithm could achieve the optimal solution to problem (12) when $P_{total}$ is large, although it is suboptimal in general for problem (12).

Remark 5.2: It is worth noting that, when ZF beamforming is employed in MISO SWIPT interference channels [8], [9], the energy efficient transceiver design problem of MISO SWIPT interference channels is in almost the same form of (38) with $P_k$ being the maximum transmission power of the $k$-th transmitter. Hence, the proposed low complexity algorithm can provide an optimal energy efficient transceiver design method for MISO SWIPT interference channels with ZF beamforming. We emphasize that this can be viewed as a contribution of this paper.

VI. SIMULATION RESULTS

In this section, we numerically evaluate the performance of the proposed energy efficiency optimization algorithms in downlink MISO SWIPT systems of $K = 4$ users. Unless specified otherwise, it is assumed that the BS is equipped with $N_t = 4$ antennas and the total transmission power is $P_{total} = 30$ dBm. We assume that the system bandwidth is $W = 15$ kHz and all users have the same set of parameters, i.e., $\zeta_k = \zeta$, $\delta_k^2 - \delta$, $\sigma_k^2 = \sigma^2$, $\epsilon_k = \epsilon$, and $\gamma_k = \gamma$, $\forall k$. Moreover, we set $\theta = \frac{\pi}{3}$, $\zeta = 0.05$, $\sigma^2 = -70$ dBm and $\delta^2 = -50$ dBm in all simulations. It is further assumed that the signal attenuation from BS to all users is 40 dB corresponding to an identical distance of 5 meters. With this transmission distance, the line-of-sight (LOS) signal is dominant, and thus the Rician fading is used to model the channel [7]. Specifically, $h_k$ is expressed as

$$h_k = \sqrt{\frac{K_R}{1 + K_R}} h_k^{LOS} + \sqrt{\frac{K_R}{1 + K_R}} h_k^{NLOS}.$$  

where $h_k^{LOS} \in \mathbb{C}^{N_t \times 1}$ and $h_k^{NLOS} \in \mathbb{C}^{N_t \times 1}$ denotes the LOS deterministic component and the Rayleigh fading component with each element being a CSCG random variable with zero mean and covariance of $40$ dB, respectively, and $K_R$ is the Rician factor set to $5$ dB. Note that the far-field uniform linear antenna array model is adopted for the LOS component, i.e., $h_k^{LOS} = 10^{-0.1} \left| 1 e^{j\varphi_1} e^{j\varphi_2} \ldots e^{j(N_t-1)\varphi_{N_t}} \right|^T$ with $\varphi_k = -\frac{2\pi d \sin(\theta_k)}{\lambda}$, where $d$ is the spacing between successive antenna elements at BS, $\mu$ is the carrier wavelength, and $\phi_k$ is the direction of user $k$ to the BS. We set $d = \frac{\lambda}{2}$ and $\{\phi_1, \phi_2, \phi_3, \phi_4\} = \{-30^\circ, -60^\circ, 60^\circ, 30^\circ\}$ as in [7].

In simulations, the constant circuitry power $P_c$ varies for different antenna setups. Specifically, $P_c$ is set according to [37]

$$P_c = N_t (P_{DAC} + P_{mix} + P_{f \bar{t}}) + 2 P_{sys} + K (P_{LNA} + P_{mix} + P_{IFA} + P_{f \bar{t}} + P_{ADC})$$

where $P_{DAC}$, $P_{mix}$, $P_{f \bar{t}}$, $P_{sys}$, $P_{LNA}$, $P_{IFA}$, $P_{f \bar{t}}$, and $P_{ADC}$ denote the power consumption of the digital to analog conversion (DAC), the mixer, the active filters at the transmitter side, the frequency synthesizer, the low-noise amplifier (LNA), the intermediate frequency amplifier (IFA), the active filters at the receiver side, and the analog to digital conversion (ADC), respectively. Table III lists all parameters used for computing $P_c$, where the values are the same as in [37] except $P_{DAC}$ and

*Every component of the constant circuitry power $P_c$ is detailed in [37]. However, other type of power consumption, e.g., CSI acquisition, can be also taken into consideration by including them in $P_c$ and the proposed energy-efficient design approach still works.*
which are estimated according to the model introduced in [38] with channel bandwidth $W = 15$ kHz.

Firstly, we study the convergence performance of the proposed algorithms. Fig. 2 illustrates the convergence performance of the Dinkelbach method (i.e., Algorithm 2) for ten random problem instances of (15). It is observed that Algorithm 2 can always converge very quickly within only several iterations. We also examine the convergence performance of the LR method (i.e., Algorithm 1). It is seen from simulations that Algorithm 1 can also converge very quickly, as shown in Fig. 3, where both the primal objective values (i.e., (13)) and the dual objective values (i.e., (15)) at feasible iterations are presented. From Fig. 3, it can be also observed that the primal objective value coincides with the dual objective value as the iteration proceeds, implying that the LR method can achieve optimal solutions. This is due to the fact that unique solutions are always observed for problem (15) in our simulations (cf. Remark 3.2).

Secondly, we examine the performance of the LR method and the low complexity algorithm in terms of the achieved average energy efficiency over 100000 random channel realizations. Fig. 4 depicts that the average system energy efficiency improves with the total transmission power, $P_{\text{total}}$, under different SINR targets when $\gamma = 15$ dB. For comparison, we also provide the dual optimal values in the plot as an upper bound for the system energy efficiency. It can be observed from Fig. 4 that, the LR method coincides with the upper bound, implying again that the LR method can achieve optimal solutions. Furthermore, it is seen from Fig. 4 that (also from Fig. 6), the performance of the low complexity method could be very close to that of the LR method and particularly coincides with the latter when $P_{\text{total}}$ exceeds 30 dBm (cf. Remark 5.1). In addition, one can see that better energy efficiency performance can be obtained in the scenario of less stringent QoS targets. Similar observations can be made from Fig. 5 which illustrates the average system energy efficiency versus the total transmission power, $P_{\text{total}}$, for different EH targets when $\gamma = 20$ dB.

For comparison, we also provide the average energy efficiency performance of the power minimization method [7] as a
Fig. 5. Average energy efficiency Vs. the total transmission power for different EH targets. For all $P_{\text{total}}$, the average energy efficiency achieved by the power minimization method [7] is $2.91 \times 10^5$ Bts/Joule when $c = -25$ dBm, and $0.78 \times 10^5$ Bts/Joule when $c = -15$ dBm.

Fig. 6. Average energy efficiency Vs. the number of transmit antennas. As shown in Fig. 6 with fixed $\gamma = 15$ dB, $e = -20$ dBm but varying $P_c$, is it seen that the system energy efficiency first improves when $N_t$ increases to 12 and then decreases with $N_t$. This implies that the BS equipped with a very large-scale antenna array may not achieve the highest energy efficiency. Thus antenna selection should be adopted in multi-antenna SWIPT system design to achieve a good energy efficiency. It is noted that a similar phenomenon was also observed in general massive MIMO systems [39], [40]. The reason behind such a phenomenon is that the constant circuitry power $P_c$, which is taken into consideration in the energy efficiency formulation, scales linearly with the number of active transmit antennas [39], [40].

VII. CONCLUSION

In this paper, we have studied EE-based joint ZF beamforming and receive power splitting for multi-user MISO SWIPT systems. Based on a simple reformulation of the total power constraint, the Lagrangian relaxation method coupled with the Dinkelbach method is proposed to handle the non-convex fractional EE maximization problem. It is proven that the LR method can guarantee at least a feasible solution by slightly adjusting the optimal dual variable if necessary. A low complexity joint transceiver design method has been also proposed for EE maximization. Simulation results validate the superior energy efficiency performance of the proposed solutions. Lastly, we remark that the proposed LR method cannot be extended to the multi-user interference case (i.e., solving problem (9) without the ZF constraints) due to the inseparable SINR and EH constraints and thus another techniques should be investigated to address the energy efficiency optimization problem with multi-user interference. In addition, it has been shown in simulations that antenna selection could improve the system energy efficiency of multi-user MISO SWIPT systems with large-scale antenna arrays at BS and thus it is worthy of studying in the future.

APPENDIX A

THE PROOF OF PROPOSITION 2.1

First, letting $\psi_k \triangleq \sqrt{p_k} \psi_k$ in (9) with $|\psi_k| = 1$, we recast problem (9) as

$$\max_{\{\psi_k,p_k,\rho_k\}} \quad \frac{\sum_{k=1}^{K} \log \left( 1 + \frac{p_k \rho_k \|h_k^H \psi_k\|^2}{p_k \sigma_k^2 + \delta_k} \right)}{p_k + P_c - \sum_{k=1}^{K} \tilde{E}_k (p_k, \rho_k)}$$

s.t.

$$\rho_k \frac{p_k \psi_k}{\sigma_k^2 + \delta_k} \geq \gamma_k, \forall k, \quad (41b)$$

$$\zeta_k (1 - \rho_k) \left( p_k \psi_k^H \psi_k^2 + \sigma_k^2 \right) \geq \epsilon_k, \forall k, \quad (41c)$$

$$H_k^H \psi_k = 0, \forall k, \quad (41d)$$

$$|\psi_k| = 1, \forall k, \quad (41e)$$

$$\sum_{k=1}^{K} p_k \leq P_{\text{total}}, \quad (41f)$$

$$0 \leq \rho_k \leq 1, \forall k. \quad (41g)$$

baseline. In the power minimization method, the transmit beamforming and receive PS ratios are designed to achieve the minimum transmission power while satisfying both the SINR constraints and EH constraints. Hence, for different allowed total transmission power $P_{\text{total}}$, the average energy efficiency of the power minimization method is invariant. To keep the figures neat, we present the average energy efficiency results of the power minimization method [7] in the captions of Figs. 4 and 5. It can be seen (also from Fig. 6 below) that the proposed methods have better average energy efficiency than the power minimization method.

At last, we investigate the impact of the number of transmit antennas, $N_t$, on the system energy efficiency for the proposed solutions. As shown in Fig. 6 with fixed $\gamma = 15$ dB, $e = -20$ dBm but varying $P_c$, is it seen that the system energy efficiency first improves when $N_t$ increases to 12 and then decreases with $N_t$. This implies that the BS equipped with a very large-scale antenna array may not achieve the highest energy efficiency. Thus antenna selection should be adopted in multi-antenna SWIPT system design to achieve a good energy efficiency. It is noted that a similar phenomenon was also observed in general massive MIMO systems [39], [40]. The reason behind such a phenomenon is that the constant circuitry power $P_c$, which is taken into consideration in the energy efficiency formulation, scales linearly with the number of active transmit antennas [39], [40].
where $E_k(\vec{v}_k, p_k, \rho_k) \triangleq \zeta_k(1 - \rho_k) \left( \rho_k |\vec{h}_k^H \vec{v}_k|^2 + \sigma_k^2 \right)$, the interference terms in (41a)–(41c) have been canceled by applying the ZF conditions (41d). Observing that increasing the values of the terms $\vec{h}_k^H \vec{v}_k^2$ will increase the objective function while maintaining the constraints (41b) and (41c), we infer that the optimal $\vec{v}_k$’s of problem (41) are obtained when $|\vec{h}_k^H \vec{v}_k|^2$ is maximized subject to the constraints (41d) and (41e), i.e., when the following problem

$$\begin{align*}
\max_{\vec{v}_k} & \quad \vec{h}_k^H \vec{v}_k^2 \\
\text{s.t.} & \quad |\vec{v}_k|^2 \leq 1, \\
& \quad \vec{h}_k^H \vec{v}_k = 0,
\end{align*}$$

(42)
is solved for all $k$. Note that, according to [7, Prop. 5.1], $\vec{v}_k$ (cf. (10)) is an optimal solution to problem (42) and $g_k$ (cf. (10)) is the corresponding optimal value. As a direct result, problem (41) can be simplified as (11). This completes the proof.

**APPENDIX B**

**THE PROOF OF LEMMA 2.1**

Clearly, problem (11) is feasible if and only if the optimal value of the following problem

$$\begin{align*}
\min_{\{p_k, \rho_k\}} & \quad \sum_{k=1}^{K} p_k \\
\text{s.t.} & \quad \frac{p_k \rho_k \sigma_k^2}{\rho_k \sigma_k^2 + \sigma_k^2} \geq \gamma_k, \quad \forall k, \\
& \quad \zeta_k(1 - \rho_k) \left( \rho_k \sigma_k^2 + \sigma_k^2 \right) \geq \epsilon_k, \quad \forall k, \\
& \quad 0 \leq \rho_k \leq 1, \quad \forall k.
\end{align*}$$

(43)
is not greater than $P_{\text{total}}$. Note that problem (43) is separable and thus we consider its $k$-th subproblem only

$$\begin{align*}
\min_{p_k, \rho_k} & \quad p_k \\
\text{s.t.} & \quad \frac{p_k \rho_k \sigma_k^2}{\rho_k \sigma_k^2 + \sigma_k^2} \geq \gamma_k, \\
& \quad \zeta_k(1 - \rho_k) \left( \rho_k \sigma_k^2 + \sigma_k^2 \right) \geq \epsilon_k, \\
& \quad 0 \leq \rho_k \leq 1.
\end{align*}$$

(44)

By the definition of $l_1(\rho_k)$ and $l_2(\rho_k)$, it can be shown that the first two constraints of problem (44) are respectively equivalent to $p_k \geq l_1(\rho_k)$ and $\rho_k \geq l_2(\rho_k)$. By combing these two constraints, we rewrite problem (44) equivalently as

$$\begin{align*}
\min_{p_k, \rho_k} & \quad p_k \\
\text{s.t.} & \quad p_k \geq \max\{l_1(\rho_k), l_2(\rho_k)\}, \\
& \quad 0 \leq \rho_k \leq 1.
\end{align*}$$

(45)

It is seen that the minimum $p_k$ is attained when the feasible set of $p_k$ becomes largest, that is, the optimal $p_k$ must minimize $\max\{l_1(\rho_k), l_2(\rho_k)\}$ in order to make the feasible set of $p_k$ as large as possible. Hence, the optimal value of problem (45) is just $p_{k,\text{min}}$, in terms of the definition of $p_{k,\text{min}}$, and the optimal value of problem (43) is $\sum_{k=1}^{K} p_{k,\text{min}}$. This completes the proof.

**APPENDIX C**

**THE PROOF OF PROPOSITION 3.1**

The first part follows directly from [36, Prop. 3.3.4]. Thus we prove the second part only. First, we show by contradiction that, $d(\lambda) \geq d(\lambda^*)$, $\forall \lambda > \lambda^*$. Assume the contrary, i.e., $d(\lambda) < d(\lambda^*)$ for some $\lambda > \lambda^*$. There must exist $t \in (0, 1)$ such that $\lambda = t \lambda^* + (1 - t) \lambda^*$. Hence, by the convexity of $d(\lambda)$ and using the fact that $d(\lambda) \geq d(\lambda^*)$, we have

$$d(\lambda) = (1 - t)d(\lambda^*) + t d(\lambda) < t d(\lambda^*) + (1 - t) d(\lambda) = d(\lambda)$$

implying a contradiction. Hence, we have $d(\lambda) > d(\lambda^*)$ for $\lambda > \lambda^*$. This implies that any subgradient of $d(\lambda)$ is nonnegative for $\lambda > \lambda^*$. Thus, by noting that

$$\frac{\delta \sum_{k=1}^{K} p_k(\lambda) + \rho_k(\lambda) - \sum_{k=1}^{K} E_k(\vec{v}_k, p_k(\lambda), \rho_k(\lambda))}{d(\lambda)}$$

is a subgradient of $d(\lambda)$, we have $\sum_{k=1}^{K} p_k(\lambda) \leq P_{\text{total}}$ for $\lambda > \lambda^*$. This completes the proof.

**REFERENCES**


