Secure and Green SWIPT in Distributed Antenna Networks with Limited Backhaul Capacity

Derrick Wing Kwan Ng and Robert Schober

Abstract

This paper studies the resource allocation algorithm design for secure information and renewable green energy transfer to mobile receivers in distributed antenna communication systems. In particular, distributed remote radio heads (RRHs/antennas) are connected to a central processor (CP) via capacity-limited backhaul links to facilitate joint transmission. The RRHs and the CP are equipped with renewable energy harvesters and share their energies via a lossy micro-power grid for improving the efficiency in conveying information and green energy to mobile receivers via radio frequency (RF) signals. The considered resource allocation algorithm design is formulated as a mixed non-convex and combinatorial optimization problem taking into account the limited backhaul capacity and the quality of service requirements on simultaneous wireless information and power transfer (SWIPT). We aim at minimizing the total network transmit power when only imperfect channel state information of the wireless energy harvesting receivers that have to be powered by the wireless network is available at the CP. In light of the intractability of the problem, we reformulate it as an optimization problem with binary selection, which facilitates the design of an iterative resource allocation algorithm to solve the problem optimally using the generalized Bender’s decomposition (GBD). Furthermore, a suboptimal algorithm is proposed to strike a balance between computational complexity and system performance. Simulation results illustrate that the proposed GBD based algorithm obtains the global optimal solution and the suboptimal algorithm achieves a close-to-optimal performance. Besides, the distributed antenna network for SWIPT with renewable energy sharing is shown to require a lower transmit power compared to a traditional system with multiple co-located antennas.

Index Terms

Limited backhaul, physical layer security, wireless information and power transfer, distributed antennas, green energy sharing, non-convex optimization.

This paper has been accepted and presented in part at IEEE Globecom 2014 [1] and ICC 2014 [2], respectively. Derrick Wing Kwan Ng and Robert Schober are with the Institute for Digital Communications (IDC), Friedrich-Alexander-University Erlangen-Nürnberg (FAU), Germany (email:{kwan, schober}@int.de). The authors are also with the University of British Columbia, Vancouver, Canada.
I. INTRODUCTION

Next generation wireless communication systems are required to provide high speed, high security, and ubiquitous communication with guaranteed quality of service (QoS). These requirements have led to a tremendous energy consumption in both transmitter(s) and receiver(s). Multiple-input multiple-output (MIMO) technology has emerged as a viable solution in reducing the system power consumption. In particular, multiuser MIMO, where a transmitter equipped with multiple antennas serves multiple single-antenna receivers, is considered as an effective solution for realizing the performance gain offered by multiple antennas. On the other hand, energy harvesting based mobile communication system design facilitates self-sustainability for energy limited communication networks. For instance, the integration of energy harvesting devices into base stations for scavenging energy from traditional renewable energy sources such as solar and wind has been proposed for providing communication services [3]–[5]. However, these natural energy sources are usually location and climate dependent and may not be suitable for portable mobile receivers.

Recently, wireless power transfer has been proposed as an emerging alternative energy source, where the receivers scavenge energy from the ambient radio frequency (RF) signals [6]–[10]. The broadcast nature of wireless channels facilitates one-to-many wireless charging which eliminates the need for power cords and manual recharging. Besides, it enables the possibility of simultaneous wireless information and power transfer (SWIPT). However, the introduction of an RF energy harvesting capability at the receivers leads to many interesting and challenging new research problems which have to be solved to bridge the gap between theory and practice. In [6] and [7], the fundamental trade-off between harvested energy and wireless channel capacity was studied for point-to-point and multi-antenna wireless broadcast systems, respectively. In [8], it was shown that RF energy harvesting can improve the energy efficiency of communication networks. The combination of physical (PHY) layer security and SWIPT was recently investigated in [9] and [10] for total transmit power minimization and secrecy rate maximization, respectively. Nevertheless, despite the promising results in the literature [6]–[10], the performance of wireless power/energy transfer systems is still severely limited by the distance between the transmitter(s) and the receiver(s) due to the high signal attenuation caused by path loss and shadowing. Thus, it is expected that the energy consumption at the transmitters for wireless power transfer systems will become a financial burden to service providers if the efficiency
for wireless power transfer cannot be improved and the energy cost at the transmitters cannot be reduced.

Distributed antennas are an important technique for mitigating interference, reducing the network power consumption, and extending service coverage [11]–[15]. A promising architecture for distributed antenna networks is to split the functionalities of the base station between a central processor (CP) and a set of low cost remote radio heads (RRHs). In particular, the CP performs the power hungry and computationally intensive baseband signal processing while the RRHs are responsible for all simple RF operations such as analog filtering and power amplification. The RRHs are distributed across the network and connected to the CP via backhaul links. This system architecture is known as cloud computing network. The distributed antenna system architecture reduces the distance between the transmitters and receivers. Furthermore, it inherently provides spatial diversity for combating path loss and shadowing. It has been shown in [11], [12] that distributed antenna systems with full cooperation between the transmitters achieve a superior performance compared to co-located antenna systems. Yet, transferring the information data of all users from the CP to all RRHs, as is required for full cooperation, may be infeasible when the capacity of the backhaul links is limited. Hence, resource allocation for distributed antenna networks with finite backhaul capacity has attracted considerable attention in the research community [13]–[15]. In [13], the authors studied the energy efficiency of distributed antenna multi-cell networks with capacity constrained backhaul links. In [14] and [15], iterative algorithms were proposed to reduce the total system backhaul capacity consumption while guaranteeing reliable communication to the mobile users. However, the problem formulations in [14] and [15] do not constrain the capacity consumption of individual backhaul links which may lead to an information overflow in some backhaul links. Moreover, [13]–[15] assume the availability of an ideal power supply for each RRH such that a large amount of energy can be continuously used for operation of the system whenever needed. This ideal power supply requirement for the RRHs may not be feasible in practice, especially in developing countries or remote areas. In these cases, the assumption of a perpetual energy supply made in [13]–[15] may be overly optimistic. In addition, the receivers in [11]–[15] were assumed to be powered by constant energy sources which may also not be a valid assumption for energy-limited handheld devices. Although the transmitters can be powered by renewable green energy and the signals transmitted in the RF by the RRHs could be exploited as energy sources to the receivers for extending their lifetimes, resource allocation algorithm design for utilizing green energy in distributed antenna SWIPT systems has not been considered in the literature,
Motivated by the aforementioned observations, we propose the use of distributed antenna communication networks for transferring information and green renewable energy to mobile receivers wirelessly. We formulate the resource allocation algorithm design as a non-convex optimization problem. By taking into account the limited backhaul capacity, the harvested renewable energy sharing between RRHs, and the imperfect CSI of the energy harvesting receivers, we minimize the total network transmit power while ensuring the quality of service (QoS) of the wireless receivers for both secure communication and efficient wireless power transfer. To this end, we propose an optimal iterative algorithm based on the generalized Bender’s decomposition. In addition, we propose a low complexity suboptimal resource allocation scheme based on the difference of convex functions (d.c.) programming which provides a local optimal solution for the considered optimization problem.

II. System Model

A. Notation

We use boldface capital and lower case letters to denote matrices and vectors, respectively. $A^H$, $\text{Tr}(A)$, and $\text{Rank}(A)$ represent the Hermitian transpose, the trace, and the rank of matrix $A$; $A \succ 0$ and $A \succeq 0$ indicate that $A$ is a positive definite and a positive semidefinite matrix, respectively; $\text{vec}(A)$ denotes the vectorization of matrix $A$ by stacking its columns from left to right to form a column vector; $I_N$ is the $N \times N$ identity matrix; $\mathbb{C}^{N \times M}$ and $\mathbb{R}^{N \times M}$ denote the set of all $N \times M$ matrices with complex and real entries, respectively; $\mathbb{H}^N$ denotes the set of all $N \times N$ Hermitian matrices; $\text{diag}(x_1, \ldots, x_K)$ denotes a diagonal matrix with the diagonal elements given by $\{x_1, \ldots, x_K\}$; $|\cdot|$ and $\|\cdot\|_p$ denote the absolute value of a complex scalar and the $l_p$-norm of a vector, respectively. In particular, $\|\cdot\|_0$ is known as the $l_0$-norm of a vector and denotes the number of non-zero entries in the vector; the circularly symmetric complex Gaussian (CSCG) distribution is denoted by $\mathcal{CN}(\mu, \sigma^2)$ with mean $\mu$ and variance $\sigma^2$; $\sim$ stands for “distributed as”; $1$ denotes a column vector with all elements equal to one. $[\cdot]_{a,b}$ returns the $(a,b)$-th element of the input matrix, $\theta_n$ is the $n$-th unit column vector, i.e., $[\theta_n]_{t,1} = 1, t = n$, and $[\theta_n]_{t,1} = 0, \forall t \neq n$; and for a real valued continuous function $f(\cdot)$, $\nabla_x f(x)$ represents the gradient of $f(\cdot)$ with respect to vector $x$.

B. Distributed Antenna System Model and Central Processor

We consider a distributed antenna multiuser downlink communication network. The system consists of a CP, $L$ RRHs, $K$ information receivers (IRs), and $M$ energy harvesting receivers (ERs), cf. Figure
Fig. 1. Distributed antenna multiuser downlink communication system model with a central processor (CP), $L = 4$ remote radio heads (RRHs), $K = 2$ information receivers (IRs), and $M = 2$ energy harvesting receivers (ERs). The blue solid ellipsoids represent the information signals included for the different IRs. The red dotted ellipsoids illustrate the dual functionality of artificial noise in providing security and facilitating efficient energy transfer to the ERs.

Each RRH is equipped with $N_T > 1$ transmit antennas. The IRs and ERs are single antenna devices which exploit the received signal powers in the RF for information decoding and energy harvesting, respectively. In practice, the ERs may be idle IRs which are scavenging energy from the RF for extending their lifetimes. On the other hand, the CP is the core unit of the network, which has the data intended for all IRs. Besides, we assume that all computations are performed in the CP. In particular, based on the available CSI, the CP computes the resource allocation policy and broadcasts it to all RRHs. Each RRH receives the control signals for resource allocation and the data of the $K$ IRs from the CP via a backhaul link. The backhaul links can be implemented with different last mile communication technologies such as digital subscriber line (DSL) or out-of-band microwave links. Thus, the backhaul capacity may be limited. Furthermore, we assume that the CP is integrated with a constant energy source (e.g., a diesel generator) for supporting its normal operation, and the distributed RRHs are equipped with traditional energy harvesters such as solar panels and wind turbines for generating renewable energy. The harvested energy can be exchanged between the CP and the RRHs over a micro-power grid and the CP manages the energy flow in the micro-power grid, cf. Section II-F.
C. Channel Model

We focus on a frequency flat fading channel and a time division duplexing (TDD) system. The wireless information and power transfer between the RRHs and the receivers is divided into time slots. The received signals at IR $k \in \{1, \ldots, K\}$ and ER $m \in \{1, \ldots, M\}$ are given by

$$y_{IR}^k = h_k^H x + n_{IR}^k \quad \text{and} \quad y_{ER}^m = g_m^H x + n_{ER}^m,$$

respectively, where $x \in \mathbb{C}^{NT_L \times 1}$ denotes the joint transmit vector of the $L$ RRHs to the $K$ IRs and the $M$ ERs. The channel between the $L$ RRHs and IR $k$ is denoted by $h_k \in \mathbb{C}^{NT_L \times 1}$, and we use $g_m \in \mathbb{C}^{NT_L \times 1}$ to denote the channel between the $L$ RRHs and ER $m$. We note that the channel vector captures the joint effects of multipath fading and path loss. $n_{IR}^k \sim \mathcal{CN}(0, \sigma^2_s)$ and $n_{ER}^m \sim \mathcal{CN}(0, \sigma^2_s)$ represent additive white Gaussian noise (AWGN). We assume that the noise variances, $\sigma^2_s$, are identical at all receivers.

D. Channel State Information

We assume that $h_k, \forall k \in \{1, \ldots, K\}$, and $g_m, \forall m \in \{1, \ldots, M\}$, can be reliably obtained at the beginning of each scheduling slot by exploiting the channel reciprocity and the pilot sequences in the handshaking signals between the RRHs and the receivers. Besides, the estimate of $h_k$ is refined at the CP via the pilot sequences in acknowledgement packets during the entire scheduling slot. As a result, we can assume that the CSI for the RRHs-to-desired IR links is perfect during the entire transmission period. On the contrary, the ERs do not interact with the RRHs during information transmission. Thus, the CSI of the ERs may be outdated during transmission. In the following, we use a deterministic model [9], [16] for characterizing the resulting CSI uncertainty. The CSI of the link between the RRHs and ER $m$ is given by

$$g_m = \hat{g}_m + \Delta g_m \quad \text{and} \quad \Omega_m \triangleq \left\{ \Delta g_m \in \mathbb{C}^{NT_L \times 1} : \Delta g_m^H \Xi_m \Delta g_m \leq \varepsilon^2_m \right\}, \quad m \in \{1, \ldots, M\},$$

where $\hat{g}_m \in \mathbb{C}^{NT_L \times 1}$ is the channel estimate of ER $m$ available at the CP at the beginning of a scheduling slot. $\Delta g_m$ represents the unknown channel uncertainty of ER $m$ due to the slowly time varying nature of the channel during transmission. In [2], we define set $\Omega_m$ which contains all possible CSI uncertainties of ER $m$. Specifically, $\Omega_m$ specifies an ellipsoidal uncertainty region.

\footnote{The proposed system can be viewed as a hybrid information and energy distributed network. Specially, the harvested green energy at the RRHs is distributed to each others via the mico-grid and distributed to ERs via RF.}
for the estimated CSI of ER $m$, where $\varepsilon_m > 0$ and $\Xi_m \in \mathbb{C}^{N_T \times N_T}$, $\Xi_m \succ 0$ represent the radius and the orientation of the region, respectively. For instance, (2) represents an Euclidean sphere when $\Xi_m = \mathbf{I}_{N_T}$. In practice, the value of $\varepsilon^2_m$ depends on the coherence time of the associated channel and $\Xi_m$ depends on the adopted channel estimation method.

E. Signal and Backhaul Models

In each scheduling time slot, $K$ independent signal streams are transmitted simultaneously to the $K$ IRs. Specifically, a dedicated beamforming vector, $w_k^l \in \mathbb{C}^{N_T \times 1}$, is allocated to IR $k$ at RRH $l \in \{1, \ldots, L\}$ to facilitate information transmission. For the sake of presentation, we define a supervector $w_k \in \mathbb{C}^{N_T L \times 1}$ for IR $k$ as

$$w_k = \text{vec} \left( \begin{bmatrix} w_k^1 & w_k^2 & \ldots & w_k^L \end{bmatrix} \right).$$

(3)

$w_k$ represents the joint beamformer used by the $L$ RRHs for serving IR $k$. Then, the information signal to IR $k$, $x_k$, can be expressed as

$$x_k = w_k d_k,$$

(4)

where $d_k \in \mathbb{C}$ is the data symbol for IR $k$ and $\mathcal{E}\{|d_k|^2\} = 1, \forall k \in \{1, \ldots, K\}$, is assumed without loss of generality. The signals intended for the desired IRs can be overheard by ERs that are in the range of service coverage. If the ERs are malicious, they may eavesdrop the information signal of the IRs. Hence to provide secure communication for the desired IRs, the ERs must be treated as potential eavesdroppers and artificial noise is generated at the RRHs. In particular, the artificial noise can be used to degrade the channels between the RRHs and the potential eavesdroppers and be an energy source for the ERs. The RRHs choose the transmit signal vector $x$ as

$$x = \sum_{k=1}^{K} x_k + \underbrace{v}_{\text{artificial noise}},$$

(5)

where $v \in \mathbb{C}^{N_T L \times 1}$ is the artificial noise vector generated by the RRHs to combat the potential eavesdroppers. In particular, $v$ is modeled as a complex Gaussian random vector with $v \sim \mathcal{CN}(0, V)$, where $V \in \mathbb{H}^{N_T \times N_T}$, $V \succeq 0$, denotes the covariance matrix of the artificial noise. The artificial noise $v$ interferes both the IRs and ERs since $v$ is unknown to them. Hence, artificial noise transmission has to be carefully optimized to provide communication security while having a minimal effect on the IRs.
On the other hand, the data of each IR is delivered from the CP to the RRHs via backhaul links. The backhaul capacity consumption for backhaul link \( l \in \{1, \ldots, L\} \) is given by

\[
C_{l}^{\text{Backhaul}} = \sum_{k=1}^{K} \left\| \| w_{k}^{l} \|_{2} \right\|_{0} R_{k},
\]

where \( R_{k} \) is the required backhaul data rate for conveying the data of IR \( k \) to a RRH and \( \sum_{k=1}^{K} \left\| \| w_{k}^{l} \|_{2} \right\|_{0} \) counts the number of IRs consuming the capacity of backhaul \( l \). We note that the backhaul links may be capacity-constrained and the CP may not be able to send the data of all IRs to all RRHs as required for full cooperation. Thus, to reduce the load on the backhaul links, the CP can enable partial cooperation by sending the data of IR \( k \) only to a subset of the RRHs. In particular, by setting \( w_{l}^{k} = 0 \), RRH \( l \) is not participating in the joint data transmission to IR \( k \). Thus, the CP is not required to send the data for IR \( k \) to RRH \( l \) via the backhaul link which leads to a lower information flow in the backhaul link.

\[ F. \text{RRH Power Supply Model} \]

The constant energy source of the CP transfers energy to all RRHs via a dedicated power grid (micro-power grid) for supporting the power consumption at the RRHs and facilitating a more efficient network operation, cf. Figure 2. In particular, a bus in Figure 2 refers to the internal power line connection between two elements with zero impedance. The CP is connected to a point of common coupling to convey energy to the micro-power grid and has full control over the micro-power grid. Since each RRH is equipped with energy harvesters for harvesting renewable energy, the energy harvested by the RRHs can also be shared in the communication system via the micro-power grid. By exploiting the spatial diversity inherent to the distributed antenna network also for energy harvesting, we can overcome potential energy harvesting imbalances in the network for improving the system performance. In other words, there are \( L + 1 \) energy sources for supporting the CP and the \( L \) RRHs. We denote the unit of energy transferred from energy source \( n \in \{1, \ldots, L + 1\} \) to the micro-grid as \( E_{n}^{S} \), where the power generator at the CP is the \((L + 1)\)-th energy source. The power loss in delivering the power from all the \( L + 1 \) energy sources to the \( L \) RRH is given by

\[
P_{\text{Loss}} = \sum_{n=1}^{L+1} \sum_{m=1}^{L+1} E_{n}^{S} B_{n,m} E_{m}^{S} = (e^{S})^{T} B e^{S} > 0,
\]

where \( e^{S} = [E_{1}^{S} \ldots E_{L}^{S} \ldots E_{L+1}^{S}]^{T}, e^{S} \in \mathbb{R}^{L+1}. B_{n,m} = [B]_{n,m} \) is known as the B-coefficient and \( B \in \mathbb{R}^{(L+1) \times (L+1)}, B \succ 0 \), is the B-coefficient matrix which takes into account the distance
dependent power line resistance, the phase angles of the electrical currents, and the voltages generated by the different energy sources. We note that the $B$-coefficient matrix is a constant for a fixed number of loads and a fixed grid connection topology. We assume that the $B$-coefficient matrix is known to the CP for energy allocation from long term measurements. Furthermore, the maximum energy supplied by energy source $n$ is given by

$$
\text{Maximum supplied energy: } \frac{\theta_n^T e^S}{E^\text{max}_n} \leq E^\text{max}_n, \forall n \in \{1, \ldots, L + 1\},
$$

where $E^\text{max}_n$ is the maximum energy available at energy source $n$ and represents the total amount of energy generated by energy source $n$. In this paper, each energy source is able to adjust the amount of energy injected into the micro-power grid. Furthermore, the coherence time of the communication channel is on the order of milliseconds for receivers moving with pedestrian speeds, while the energy harvesting process evolves slowly compared to the communication channels. For instance, solar and wind energy remains constant for seconds in general [5]. As a result, for the resource allocation algorithm design for SWIPT, we assume that $E^\text{max}_n$ is a known constant.

### III. Problem Formulation

In this section, we define the quality of service (QoS) metrics for the design of secure communication and power efficient wireless energy transfer. Then, the resource allocation algorithm design is formulated as a non-convex optimization problem.
A. Achievable Data Rate and RF Energy Harvesting

The achievable data rate (bit/s/Hz) between the $L$ RRHs and IR $k$ is given by

$$C_k = \log_2(1 + \Gamma_k), \quad \text{where} \quad \Gamma_k = \frac{|h_k^H w_k|^2}{\sum_{j \neq k} |h_j^H w_j|^2 + \text{Tr}(Vh_k h_k^H) + \sigma_s^2}$$

(9)

is the receive signal-to-interference-plus-noise ratio (SINR) at IR $k$.

On the other hand, we assume that successive interference cancellation is performed at the ERs to remove the multiuser interference before decoding the desired information which represents the worst case for guaranteeing PHY layer security. Therefore, the achievable data rate between the RRHs and ER (potential eavesdropper) $m$ is given by

$$C_{ER_m} = \log_2 \left(1 + \Gamma_{ER_m}\right)$$

and

$$\Gamma_{ER_m} = \frac{|g_m^H w_k|^2}{\sum_{j \neq k} |g_m^H w_j|^2 + \text{Tr}(Vg_m g_m^H) + \sigma_s^2} \leq \frac{|g_m^H w_k|^2}{\text{Tr}(Vg_m g_m^H) + \sigma_s^2},$$

(10)

(11)

where $\Gamma_{ER_m}$ is the received SINR at ER $m$. $(a)$ constitutes an upper bound on the received SINR at ER $m$ for decoding the information of IR $k$ due to the aforementioned worst case assumption.

In the considered system, the information signal, $w_k d_k, \forall k \in \{1, \ldots, K\}$, serves as a dual purpose carrier for both information and energy. Besides, the artificial noise signal also acts as an energy source to the ERs. The total amount of energy harvested by ER $m \in \{1, \ldots, M\}$ is given by

$$E_{m}^{ER} = \mu \left(\text{Tr}(Vg_m g_m^H) + \sum_{k=1}^{K} |g_m^H w_k|^2\right),$$

(12)

where $0 < \mu \leq 1$ denotes the efficiency of converting the received RF energy to electrical energy for storage. We assume that $\mu$ is a constant and is identical for all ERs. Besides, the contribution of the antenna thermal noise power to the harvested energy is negligibly small compared to the energy harvested from the received signal, $\text{Tr}(Vg_m g_m^H) + \sum_{k=1}^{K} |g_m^H w_k|^2$, and thus is neglected in (12).

\(^2\)We note that the proposed framework can be easily extended to the case when a single-user detector is employed at the potential eavesdroppers.

\(^3\)We adopt the normalized energy unit Joule-per-second in this paper. Therefore, the terms “power” and “energy” are used interchangeably.
B. Optimization Problem Formulation

The system objective is to minimize the total network transmit power while providing QoS for reliable communication and efficient power transfer in a given time slot for given maximum backhaul capacities. The resource allocation algorithm design is formulated as the following optimization problem:

\[
\begin{align*}
\text{minimize} & & \sum_{k=1}^{K} \sum_{l=1}^{L} \left\| w_{lk} \right\|_2^2 + \text{Tr}(V) \\
\text{s.t.} & & \frac{|h_k^H w_k|^2}{\sum_{n \neq k} |h_n^H w_k|^2 + \text{Tr}(Vh_k h_k^H) + \sigma^2} \geq \Gamma_{\text{req}}, \forall k, \\
& & \min_{\Delta g_m \in \Omega_m} \frac{|g_m^H w_k|^2}{\text{Tr}(Vg_m g_m^H) + \sigma^2} \leq \Gamma_{\text{tol}}, \forall m, k, \\
& & \sum_{k=1}^{K} \left\| [\| w_k^l \|_2]_0 R_k \right\|_0 \leq C_l^{B_{\text{max}}}, \forall l, \\
& & P_{\text{CP}} + \sum_{l=1}^{L} \left\{ P_{l} + \rho \left( \sum_{k=1}^{K} \| w_k^l \|_2^2 + \text{Tr}(VR_l) \right) \right\} \leq 1^T e^S - (e^S)^T B e^S, \\
& & \theta_m^T e^S \leq E_m^{\text{max}}, \forall n \in \{1, \ldots, L + 1\}, \quad \text{C6: } \text{Tr}(VR_l) + \sum_{k=1}^{K} \| w_k^l \|_2^2 \leq P_{l}^{T_{\text{max}}}, \forall l, \\
& & \Delta g_m \in \Omega_m E_m^{\text{ER}} \geq P_m^{\text{min}}, \forall m, \quad \text{C8: } e^S \geq 0, \quad \text{C9: } V \succeq 0, 
\end{align*}
\]

where \( R_l \triangleq \text{diag} \left( \underbrace{0, \ldots, 0}_{(l-1)N_T}, \underbrace{1, \ldots, 1}_{N_T}, \underbrace{0, \ldots, 0}_{(L-l)N_T} \right), \forall l \in \{1, \ldots, L\} \), is a block diagonal matrix. \( \Gamma_{\text{req}} > 0 \) in constraint C1 indicates the required minimum receive SINR at IR \( k \) for information decoding. The corresponding data rate per backhaul link use for IR \( k \) is a constant given by \( R_k = \log_2(1 + \Gamma_{\text{req}}) \). Constraint C2 is imposed such that for a given CSI uncertainty set \( \Omega_m \), the maximum received SINR at ER \( m \) is less than the maximum tolerable received SINR \( \Gamma_{\text{tol}} \). In practice, the CP sets \( \Gamma_{\text{req}} \gg \Gamma_{\text{tol}} > 0, \forall k \in \{1, \ldots, K\} \), to ensure secure communication. Specifically, the adopted problem formulation guarantees that the achievable secrecy rate for IR \( k \) is bounded below by \( R_{\text{sec}} \geq \log_2(1 + \Gamma_{\text{req}}) - \log_2(1 + \Gamma_{\text{tol}}) \geq 0 \). We note that although \( \Gamma_{\text{req}} \) and \( \Gamma_{\text{tol}} \) in C1 and C2, respectively, are not optimization variables in this paper, a balance between secrecy capacity and system capacity can be struck by varying their values. In fact, when constraint C2 is removed from the optimization problem or \( \Gamma_{\text{tol}} \rightarrow \infty \), PHY layer security is not considered in the system. In other words, the adopted problem formulation is a generalized framework which provides flexibility in controlling the level of
communication security. In C3, the backhaul capacity consumption for backhaul link $l$ is constrained to be less than the maximum available capacity of backhaul link $l$, i.e., $C^B_{l_{\text{max}}}$. We note that once the covariance matrix $V$ is fixed, artificial noise $\nu$ can be generated locally at each RRH and is not required to be sent via the backhaul links. The right hand side of C4, $1^T e^S - (e^S)^T B e^S$, denotes the maximum available power in the power grid taking into account the power loss in the power lines. We note that $1^T e^S - (e^S)^T B e^S \geq 0$ always holds by the law of conservation of energy. The left hand side of C4 accounts for the total power consumption in the network. In C4, $P_{CP}^l$ and $P_{C_l}$ represent the fixed circuit power consumption in the CP and RRH $l$, respectively; the term $\sum_{k=1}^{K} \|w^l_k\|_2^2 + \text{Tr}(V R_l)$ denotes the output power of the power amplifier of RRH $l$, and $\rho \geq 1$ is a constant accounting for the power inefficiency of the power amplifier; C5 is a constraint on the maximum power supply from energy source $n \in \{1, \ldots, L + 1\}$. Constant $P_{l_{\text{max}}}^T$ in C6 is the maximum transmit power allowance for RRH $l$, which can be used to limit out-of-cell interference. $P_{m_{\text{min}}}$ in C7 is the minimum required power transfer to ER $m$. We note that for given CSI uncertainty sets $\Omega_m$, $\forall m$, the CP can only guarantee the minimum required power transfer to the $M$ ERs if they use all their received power for energy harvesting. C8 is the non-negativity constraint on the energy supply optimization variables. C9 and $V \in \mathbb{H}^{N_t \times L}$ constrain matrix $V$ to be a positive semi-definite Hermitian matrix, i.e., they ensure that $V$ is a valid covariance matrix.

Remark 1: We emphasize that the problem formulation considered in this paper is different from that in [14] and [15]. In particular, we focus on the capacity consumption of individual backhaul links while [14] and [15] studied the total network backhaul capacity consumption. Besides, we constrain the capacity consumption of the individual backhaul links which is not possible with the problem formulation adopted in [14] and [15]. On the other hand, although the combination of PHY layer security and SWIPT has been recently considered in [9] and [10], the results in [9] and [10] cannot be directly applied to our problem formulation due to the combinatorial constraints on the limited backhaul capacity and the exchange of harvested power between RRHs.

Remark 2: In practice, the coherence time of the communication channel is much shorter than the energy harvesting process and thus we assume that the energy arrival at the energy harvesters is a constant for resource allocation. The study of resource allocation algorithm design with a rapidly changing energy harvesting process is beyond the scope of this paper. We assume that when the energy harvested by RRHs exceeds the total energy consumption of the communication system, the remaining harvested energy is transferred to the external power grid. A possible scenario of the
considered system model is a smart grid setup. Specifically, the excess amount of harvested energy can be sold to the utility company. In fact, the proposed framework can be extended to the case of dynamic energy harvesting with energy storage in the RRHs by following similar approaches as in [4] and [18].

IV. RESOURCE ALLOCATION ALGORITHM DESIGN

The optimization problem in (13) is a non-convex problem. In the following, we first develop an iterative resource allocation algorithm for obtaining the global optimal solution based on the generalized Bender’s decomposition. Then, we propose a low computational complexity suboptimal algorithm inspired by the difference of convex functions programming.

A. Problem Reformulation

In this section, we reformulate the considered optimization problem to facilitate the development of resource allocation algorithms. First, we define \( W_k = w_k w_k^H \), \( H_k = h_k h_k^H \), and \( G_m = g_m g_m^H \) for notational simplicity. Besides, we introduce an auxiliary optimization variable \( s_{l,k} \) for simplifying the problem. Then, we recast the optimization problem as follows:

\[
\begin{align*}
\text{minimize} & \quad \sum_{k=1}^{K} \text{Tr}(W_k) + \text{Tr}(V) \\
\text{s.t.} \quad & \frac{\text{Tr}(H_k W_k)}{\Gamma_{\text{req}k}} \geq \sum_{j \neq k}^{K} \text{Tr}(H_k W_j) + \text{Tr}(H_k V) + \sigma_s^2, \forall k, \\
& \min_{\Delta g_m \in \Omega_k} \frac{\text{Tr}(W_k G_m)}{\Gamma_{\text{tol}k}} \leq \text{Tr}(G_m V) + \sigma_s^2, \forall m, k, \\
& \sum_{k=1}^{K} s_{l,k} R_k \leq C_{l}^{\text{Bmax}}, \forall l, \\
& P_{\text{CP}} + \sum_{l=1}^{L} \left\{ P_{C_l} + \varepsilon \left( \sum_{k=1}^{K} \text{Tr}(W_k R_l) + \text{Tr}(VR_l) \right) \right\} \leq 1^T e^S - (e^S)^T B e^S, \\
& \text{Tr}(VR_l) + \sum_{k=1}^{K} \text{Tr}(R_l W_k) \leq P_{l}^{\text{Tmax}}, \forall l, \\
& \min_{\Delta g_m \in \Omega_k} \mu \left[ \text{Tr} \left( \sum_{k=1}^{K} (W_k + V) G_m \right) \right] \geq P_{m}^{\text{min}}, \forall m, \\
& s_{l,k} \in \{0, 1\}, \forall k, l, \quad \text{C10: } \text{Tr}(W_k R_l) \leq s_{l,k} P_{l}^{\text{Tmax}}, \forall k, l, \\
& W_k \succeq 0, \forall k, \quad \text{C12: } W_k \succeq 0, \forall k, \\
& \text{Rank}(W_k) \leq 1, \forall k, \quad \text{C13: } \text{Rank}(W_k) \leq 1, \forall k,
\end{align*}
\]

(14)
Constraints C12, C13, and $W_k \in \mathbb{H}^{N_T \times L}$, $\forall k$, are imposed to guarantee that $W_k = w_k w_k^H$ holds after optimization. On the other hand, C10 and C11 are auxiliary constraints. In particular, constraints C10 and C11 restrict the optimization problem such that $s_{l,k} = 1$ must hold when the data of IR $k$ is conveyed to RRH $l$ for information transmission, i.e., $\text{Tr}(W_k R_l) > 0$. In other words, when $\text{Tr}(W_k R_l) > 0$, the data of IR $k$ consumes $R_k$ bit/s/Hz of the capacity of backhaul link $l$, cf. C3 in (14). On the other hand, it can be verified that the optimization problems in (14) and (13) are equivalent in the sense that they share the same optimal solution $\{W_k, V, e^S, s_{l,k}\}$. As a result, we focus on the design of an algorithm for solving the non-convex optimization problem in (14).

B. Iterative Resource Allocation Algorithm

In the following, we adopt the generalized Bender’s decomposition (GBD) to handle the constraints involving binary optimization variables [19]–[21], i.e., C3, C10, and C11. In particular, we decompose the problem in (14) into two problems, a primal problem and a master problem. The primal problem is a non-convex optimization problem when optimization variable $s_{l,k}$ is fixed and solving this problem with respect to $\{W_k, V, e^S\}$ yields an upper bound for the optimal value of (14). The master problem is a mixed-integer linear programming (MILP) with binary optimization variables $s_{l,k}$ for a fixed value of $\{W_k, V, e^S\}$. The solution of the master problem provides a lower bound for the optimal value of (14). We solve the primal and master problems iteratively until the solutions converge. In the following, we first propose algorithms for solving the primal and master problems in the $i$-th iteration, respectively. Then, we describe the iterative procedure between the master problem and the primal problem.

1) **Solution of the primal problem in the $i$-th iteration:** For given and fixed input parameters $s_{l,k}(i)$ obtained from the master problem in the $i$-th iteration, we solve the following optimization problem in the primal problem:

$$\begin{align*}
\text{minimize} & \quad w_k V \in \mathbb{H}^{N_T \times L}, e^S \\
& \quad \sum_{k=1}^{K} \text{Tr}(W_k) + \text{Tr}(V) \\
\text{s.t.} & \quad C1, C2, C4 - C9, C11 - C13.
\end{align*}$$

(15)

We note that constraints C3 and C10 in (14) will be handled by the master problem since they involve only the binary optimization variable $s_{l,k}$. Besides, $s_{l,k}$ is treated as a given constant in (15) and we minimize the objective function with respect to variables $\{W_k, V, e^S\}$. The first step in solving the primal problem in (15) is to handle the infinitely many constraints in C2 and C7 due to the imperfect
CSI. To facilitate the resource allocation algorithm design, we transform constraints C2 and C7 into linear matrix inequalities (LMIs) via the S-Procedure [22]. Exploiting [22] it can be shown that the original constraint C2 holds if and only if there exist \( \delta_{m,k} \geq 0, m \in \{1, \ldots, M\}, k \in \{1, \ldots, K\} \) such that the following LMI constraints hold:

\[
C2: \mathbf{S}_{C2_{m,k}}(W_k, V, \delta_{m,k}) = \begin{bmatrix} \delta_{m,k} \Xi_m + V & V\hat{g}_m \\ \hat{g}_m^H V & -\delta_{m,k} \Xi_m + \sigma_s^2 + \hat{g}_m^H V\hat{g}_m \end{bmatrix} - \frac{U_{gm}^H W_k U_{gm}}{\Gamma_{tol}} \succeq 0, \forall k, \tag{16}
\]

where \( U_{gm} = [I_{NTL} \hat{g}_m] \). Similarly, constraint C7 can be equivalently written as

\[
C7: \mathbf{S}_{C7_{m}}(W_k, V, \nu_m) = \begin{bmatrix} \nu_m \Xi_m + V & V\hat{g}_m \\ \hat{g}_m^H V & -\nu_m \Xi_m + \frac{P_{\text{min}}}{\mu} + \hat{g}_m^H V\hat{g}_m \end{bmatrix} + \sum_{k=1}^{K} U_{gm}^H W_k U_{gm} \succeq 0, \forall m, \tag{17}
\]

for \( \nu_m \geq 0, m \in \{1, \ldots, M\} \). Now, constraints C2 and C7 involve only a finite number of constraints which facilitates the resource allocation algorithm design. As a result, we can rewrite the primal problem as:

\[
\begin{array}{ll}
\text{minimize} & \sum_{k=1}^{K} \text{Tr}(W_k) + \text{Tr}(V) \\
\text{s.t.} & C1, C4, C5, C6, C8, C9, C12, \\
& C2: \mathbf{S}_{C2_{m,k}}(W_k, V, \delta_{m,k}) \succeq 0, \forall m, k, \\
& C7: \mathbf{S}_{C7_{m}}(W_k, V, \nu_m) \succeq 0, \forall m, \\
& C11: \text{Tr}(W_k R_l) \leq s_{l,k}(i) P_{l\text{max}}, \forall k, l, \\
& C13: \text{Rank}(W_k) \leq 1, \forall k, \ C14: \delta_{m,k}, \nu_m \geq 0, \forall m, k,
\end{array}
\]

where \( \delta \) and \( \nu \) are auxiliary optimization variable vectors, whose elements \( \delta_{m,k} \geq 0, m \in \{1, \ldots, M\}, k \in \{1, \ldots, K\} \), and \( \nu_m \geq 0, m \in \{1, \ldots, M\} \), were introduced in (16) and (17), respectively. Then, we relax constraint C13: \( \text{Rank}(W_k) \leq 1 \) by removing it from the problem formulation, such that the considered problem becomes a convex SDP. We note that the relaxed problem of (18) can be solved efficiently by convex programming numerical solvers such as CVX [23]. If the matrices \( W_k \) obtained from the relaxed problem (18) are rank-one matrices for all IRs, \( k \in \{1, \ldots, K\} \), then the problem in (18) and its relaxed version share the same optimal solution and the same optimal objective value. Otherwise, the optimal objective value of the relaxed version of (18) serves as a lower bound for the objective value of (18) since a larger feasible solution set is considered.

Now, we study the tightness of the adopted of the SDP relaxation. As the SDP relaxed optimization problem in (18) satisfies Slater’s constraint qualification and is jointly convex with respect to the
optimization variables, strong duality holds and thus solving the dual problem is equivalent to solving (18). For formulating the dual problem, we first define the Lagrangian of the relaxed version of (18) which can be expressed as

\[
\mathcal{L}(W_k, V, e^S, \delta, \nu, s_{l,k}(i), \Phi) = f_0(W_k, V) + f_1(W_k, V, e^S, \delta, \nu, \Phi) + f_2(W_k, s_{l,k}(i), \Phi),
\]

\[
f_0(W_k, V) = \sum_{k=1}^K \text{Tr}(W_k) + \text{Tr}(V),
\]

\[
f_1(W_k, V, e^S, \delta, \nu, \Phi) = -\text{Tr}(YV) - \sum_{k=1}^K \text{Tr}(Z_k W_k) + \sum_{n=1}^{L+1} \tau_n (\theta_n^T e^S - E_n^{\max})
\]

\[-\sum_{m=1}^M \sum_{k=1}^K \text{Tr}\left( S_{C_{2m,k}} W_k, V, \delta_{m,k} \right) D_{C_{2m,k}} - \sum_{m=1}^M \text{Tr}\left( S_{C_{7m}} W_k, V, \nu_m \right) D_{C_{7m}}
\]

\[+ \sum_{k=1}^K \alpha_k \left[ -\frac{1}{\Gamma_{\text{req},k}} \text{Tr}(H_k W_k) + \sum_{j \neq k}^K \text{Tr}(H_k W_j) + \text{Tr}(H_k V) + \sigma_s^2 \right] - \sum_{n=1}^{L+1} (\theta_n^T e^S) \chi_n
\]

\[+ \frac{\beta_{\text{CP}}}{M} \sum_{l=1}^L \left\{ P_C + \varepsilon \left( \sum_{k=1}^K \text{Tr}(W_k R_l) + \text{Tr}(V R_l) \right) \right\} - 1^T e^S + (e^S)^T B e^S
\]

\[-\sum_{k=1}^K \sum_{m=1}^M \delta_{m,k} \lambda_{m,k} - \sum_{m=1}^M \nu_m \theta_m + \sum_{l=1}^L \lambda_l \left( \text{Tr}(V R_l) + \sum_{k=1}^K \text{Tr}(R_l W_k) - P_{l,T}^{\max} \right), \text{ and}
\]

\[
f_2(W_k, s_{l,k}(i), \Phi) = \sum_{k=1}^K \sum_{l=1}^L \beta_{k,l} \left( \text{Tr}(W_k R_l) - s_{l,k}(i) P_{l,T}^{\max} \right).
\]

Here, \( \Phi = \{D_{C_{2m,k}}, D_{C_{7m}}, Y, Z_k, \alpha_k, \beta_k, \lambda_{m,k}, \theta_m \} \) is a collection of dual variables; \( D_{C_{2m,k}}, D_{C_{7m}}, \) \( Y, \) and \( Z_k \) are the dual variable matrices for constraints C2, C7, C9, and C12, respectively; \( \alpha_k, \beta_k, \lambda_{m,k}, \) \( \theta_m \) are scalar dual variables for constraints C1, C4, C5, C6, C8, C11, and C14, respectively. Function \( f_0(W_k, V) \) in (20) is the objective function of the SDP relaxed version of (18); \( f_1(W_k, V, e^S, \delta, \nu, \Phi) \) in (21) is a function involving only continuous optimization variables and dual variables; \( f_2(W_k, s_{l,k}(i), \Phi) \) in (22) is a function involving continuous optimization variables, dual variables, and binary optimization variable \( s_{l,k}(i) \). These functions are defined here for notational simplicity and will be exploited for facilitating the presentation of the solutions for both the primal problem and the master problem.

The dual problem of the relaxed SDP optimization problem in (18) is given by

\[
\max_{\Phi \geq 0} \min_{W_k, V \in \mathbb{R}^{N_T L}} \mathcal{L}(W_k, V, e^S, \delta, \nu, s_{l,k}(i), \Phi).
\]
We define $\Theta(i) = \{W^*_k, V^*, e^{S_k}, \delta^*, \nu^*\}$ and $\Phi(i) = \{\Phi^*\}$ as the optimal primal solution and the optimal dual solution of the SDP relaxed problem in \text{(18)} in the $i$-th iteration.

In the following, we introduce a theorem inspired by [10] for revealing the tightness of the SDP relaxation adopted in \text{(18)}. Let $C_k = I_{N_T L} + \sum_{m=1}^{M} \mathcal{U}_{g_m} \left( \frac{D_{C^2 m,k}}{\Gamma_{req_k}} - D_{C^7 m} \right) \mathcal{U}^H_{g_m} + \sum_{j \neq k} \mathcal{H}_j \mathcal{A}_j + \sum_{l=1}^{I} R_l (\rho \varepsilon + \gamma_l + \beta_{l,k})$ and $\text{Rank}(C_k) = r_k$. Besides, we denote the orthonormal basis of the null space of $C_k$ as $Y_k \in \mathbb{C}^{N_T L \times (N_T L - r_k)}$, and $\phi_{\omega_k} \in \mathbb{C}^{N_T L \times 1}$, where $1 \leq \omega_k \leq N_T L - r_k$, denotes the $\omega_k$-th column of $Y_k$. Hence, $C_k Y_k = 0$ and $\text{Rank}(Y_k) = N_T L - r_k$.

**Theorem 1:** Suppose the optimal solution of the SDP relaxed version of \text{(18)} is denoted by $\Theta^* = \{W^*_k, V^*, e^{S_k}, \delta^*, \nu^*\}$ and the dual problem solution is denoted as $\Phi^* = \{D^*_{C^2 m,k}, D^*_{C^7 m}, Y^*, Z^*_k, \alpha^*_k, \theta^*_k, \tau^*_m, \chi^*_n, \gamma^*_l, \beta^*_k, \delta^*_m, \theta^*_m\}$. Then, at the optimal point, for $\Gamma_{req_k} > 0$, and $\Gamma_{tol} > 0$, we have the following equalities:

$$H_k Y_k^* = 0, \quad W_k^* = \sum_{\omega_k=1}^{N_T L - r_k} \psi_{\omega_k} \phi_{\omega_k} \phi_{\omega_k}^H + \frac{f_k u_k u_k^H}{\text{rank-one}},$$

where $Y_k^*$ is the null space of matrix $C_k$ computed with the optimal dual variables $\Theta^*$. Variables $\psi_{\omega_k} \geq 0, \forall \omega_k \in \{1, \ldots, N_T L - r_k\}$, and $f_k > 0$ are positive scalars and $u_k \in \mathbb{C}^{N_T L \times 1}$, $\|u_k\| = 1$, satisfies $u_k^H Y_k = 0$. If $\exists k : \text{Rank}(W_k^*) > 1$, i.e., $\psi_{\omega_k} > 0$, then we can construct another solution of \text{(18)}, denoted by $\{\overline{W}_k, \overline{V}, \overline{e}^S, \overline{\delta}, \overline{\nu}\}$, which not only achieves the same objective value as $\Theta^*$, but also admits a rank-one matrix, i.e., $\text{Rank}(\overline{W}_k) = 1, \forall k$. The new optimal solution for the primal problem in the $i$-th iteration is given as

$$\overline{W}_k = f_k u_k u_k^H = W_k^* - \sum_{\omega_k=1}^{N_T L - r_k} \psi_{\omega_k} \phi_{\omega_k} \phi_{\omega_k}^H, \quad \overline{V} = V^* + \sum_{\omega_k=1}^{N_T L - r_k} \psi_{\omega_k} \phi_{\omega_k} \phi_{\omega_k}^H, \quad \overline{e}^S = e^{S^*_k}, \quad \overline{\delta} = \delta^*, \quad \overline{\nu} = \nu^*,$$

with $\text{Rank}(\overline{W}_k) = 1, \forall k \in \{1, \ldots, K\}$, where $f_k$ and $\psi_{\omega_k}$ can be easily found by substituting the variables in \text{(25)} into the relaxed version of \text{(18)} and solving the resulting convex optimization problem for $f_k$ and $\psi_{\omega_k}$.

**Proof:** The proof of Theorem 1 closely follows the proof of [10] Proposition 4.1] and is omitted here due to page limitation.

In other words, by applying Theorem 1 the optimal solution of the primal problem is obtained in each iteration. Besides, from the numerical solver, the dual variables corresponding to the constraints
in (18), i.e., $\Phi$, are obtained together with the primal solution $\Theta$. This information is used as an input to the master problem.

If problem (18) is infeasible for a given binary variable $s_{l,k}(i)$, then we formulate an $l_1$-minimization problem and use the corresponding dual variables and the optimal primal variables as the input to the master problem for the next iteration [20]. The $l_1$-minimization problem is given as:

$$\begin{align*}
\min_{W_k, V \in \mathbb{R}^{N_T L}} & \sum_{k=1}^{K} \sum_{l=1}^{L} \alpha_{l,k} \\
\text{s.t.} & \quad C1, C2, C4 - C9, C11, C12, C14,
\end{align*}$$

C11: $\text{Tr}(W_k R_l) \leq s_{l,k}(i) P_t^{\text{max}} + \alpha_{l,k}, \forall k, l, C15: \alpha_{l,k} \geq 0, \forall l, k.$ \hspace{1cm} (26)

The $l_1$-minimization problem is a convex optimization problem and can be solved by standard convex programming solvers. We adopt a similar notation as in (18) to denote the dual variables with respect to constraints C1, C2, C4 – C9, C11, C12, and C14 in (26). In particular, these variables are defined by: $\tilde{\Phi}(i) = \{\tilde{D}_{C_{2m,k}}, \tilde{D}_{C_{7m}}, \tilde{Z}_k, \tilde{Y}, \tilde{\alpha}_k, \tilde{\beta}_l, \tilde{\lambda}_{m,k}, \tilde{\theta}_m\}$. Also, the solution for the $l_1$-minimization problem in (26) is denoted as $\tilde{\Theta}(i) = \{W_k, V, e^S, \delta, \nu\}$.

2) Solution of the master problem in the $i$-th iteration: For notational simplicity, we define $F$ and $I$ as the sets of all iteration indices at which the primal problem is feasible and infeasible, respectively. Then, we formulate the master problem which utilizes the solutions of (18) and (26). The master problem in the $i$-th iteration is given as follows:

$$\begin{align*}
\min_{\mu, s_{l,k}} & \mu \\
\text{s.t.} & \quad \mu \geq \xi(\Phi(t), s_{l,k}), t \in \{1, \ldots, i\} \cap F, \\
& \quad 0 \geq \bar{\xi}(\tilde{\Phi}(t), s_{l,k}), t \in \{1, \ldots, i\} \cap I, \\
& \quad C3: \sum_{k=1}^{K} s_{l,k} R_k \leq C_l^{\text{max}}, \forall l, \quad C10: s_{l,k} \in \{0, 1\},
\end{align*}$$

(27)

where $s_{l,k}$ and $\mu$ are optimization variables for the master problem and

$$\begin{align*}
\xi(\Phi(t), s_{l,k}) &= \min_{W_k, V \in \mathbb{R}^{N_T L}} f_0(W_k, V) + f_1(W_k, V, e^S, \delta, \nu, \Phi(t)) + f_2(W_k, s_{l,k}, \Phi(t)), \\
\bar{\xi}(\tilde{\Phi}(t), s_{l,k}) &= \min_{W_k, V \in \mathbb{R}^{N_T L}} f_1(W_k, V, e^S, \delta, \nu, \tilde{\Phi}(t)) + f_2(W_k, s_{l,k}, \tilde{\Phi}(t)).
\end{align*}$$

(28)

(29)

If the problem in (26) is infeasible, this means that the original problem in (18) is infeasible even if the backhaul capacity is unlimited.
Equations (28) and (29) represent two different inner minimization problems inside the master problem. In particular, \( \mu \geq \xi(\Phi(t), s_{l,k}), t \in \{1, \ldots, i\} \cap \mathcal{F} \) and \( 0 \geq \tilde{\xi}(\tilde{\Phi}(t), s_{l,k}), t \in \{1, \ldots, i\} \cap \mathcal{I} \), denote the sets of hyperplanes spanned by *optimality cut* and *feasibility cut* from the first to the \( i \)-th iteration, respectively. The two different types of cuts are exploited to reduce the search region for the global optimal solution. Besides, both \( \xi(\Phi(t), s_{l,k}) \) and \( \tilde{\xi}(\tilde{\Phi}(t), s_{l,k}) \) are also functions of \( s_{l,k} \) which is the optimization variable of the outer minimization in (27).

Now, we introduce the following proposition for the solutions for the inner minimization problems.

**Proposition 1:** The solutions of (28) and (29) for index \( t \in \{1, \ldots, i\} \) are the solutions of (18) and (26) in the \( t \)-th iteration, respectively.

**Proof:** Please refer to Appendix for a proof of Proposition 1.

By substituting \( \Theta(t) \) and \( \tilde{\Theta}(t) \) into (28) and (29), respectively, the master problem is a standard MILP which can be solved by using standard numerical solvers for MILPs such as Mosek [24] and Gurobi [25]. We note that the objective value of (27), i.e., (27a), is a monotonically non-decreasing function with respect to the number of iterations as an additional constraint is imposed to the master problem in each additional iteration.

3) **Overall algorithm:** The overall iterative resource allocation algorithm is summarized in Table I. The algorithm is implemented by a repeated loop. We first set the iteration index \( i \) to zero and initialize the binary variables \( s_{l,k}(i) \). In the \( i \)-th iteration, we solve the problem in (18) by Theorem 1. If the problem is feasible (lines 6 – 7), then we obtain an intermediate resource allocation policy \( \Theta(i) \), the corresponding Lagrange multiplier set \( \Phi(i) \), and an intermediate objective value \( J_0' \). Both \( \Theta(i) \) and \( \Phi(i) \) are used to generate an *optimality cut* in the master problem. Besides, we update the performance upper bound \( UB(i) \) and the current optimal resource allocation policy when the current objective value is the lowest compared to those in all previous iterations. If the problem is infeasible (lines 9 – 10), then we solve the \( l_1 \)-minimization problem in (26) and obtain an intermediate resource allocation policy \( \tilde{\Theta}(i) \) and the corresponding Lagrange multiplier set \( \tilde{\Phi}(i) \). This information will be used to generate an *infeasibility cut* in the master problem. Then, we solve the master problem based on \( \tilde{\Theta}(t) \) and \( \Theta(i) \), \( t \in \{1, \ldots, i\} \), using a standard MILP numerical solver. The objective value of the master problem in each iteration serves as a system performance lower bound to the original optimization problem in (18). In the \( i \)-th iteration, when the difference between the \( i \)-th lower bound and the \( i \)-th upper bound is less than a predefined threshold \( \kappa \) (lines 12 – 14), the algorithm stops. We note that the convergence of the proposed iterative algorithm to the global optimal solution of (18) in...
**Algorithm** Generalized Bender’s Decomposition

1: Initialize the maximum number of iterations $L_{\text{max}}$ and a small constant $\kappa \to 0$
2: Set iteration index $i = 0$ and start with random values $s_{l,k}(i), \forall k, l$
3: repeat {Loop}
4: Solve (18) according to Theorem 1 for a given set of $s_{l,k}(i)$
5: if (18) is feasible then
6: Obtain an intermediate resource allocation policy $\Theta(i) = \{W'_k, V', e^{S'}, \delta', \nu'\}$, the corresponding Lagrange multiplier set $\Phi(i)$, and an intermediate objective value $f'_0$
7: Update the upper bound $\text{UB}(i) = \min_{t=1,\ldots,i-1} \{\text{UB}(t), f'_0\}$. If $\text{UB}(i) = f'_0$, set the current optimal policy $\Theta_{\text{current}} = \Theta(i)$, $s_{\text{current}} = s_{l,k}(i)$
8: else
9: Solve the feasibility problem in (26) and obtain an intermediate resource allocation policy $\tilde{\Theta}(i) = \{W'_k, V', e^{S'}, \delta', \nu'\}$ and the corresponding Lagrange multiplier set $\tilde{\Phi}(i)$
10: end if
11: Solve the master problem in (27) for $s_{l,k}$, save $s_{l,k}(i+1) = s_{l,k},$ and obtain the $i$-th lower bound, i.e., $\text{LB}(i)$
12: if $|\text{LB}(i) - \text{UB}(i)| \leq \kappa$ then
13: Global optimal = true, return $\{W'_k, V', e^{S'}, \delta', \nu', s_{\text{current}}\} = \{\Theta_{\text{current}}, s_{\text{current}}\}$
14: else
15: $i = i + 1$
16: end if
17: until $i = L_{\text{max}}$

A finite number of iterations is ensured even if $\kappa = 0$, provided that the master and primal problems can be solved in each iteration [20, Theorem 6.3.4]. We note that the optimal resource allocation algorithm has a non-polynomial time computational complexity. Please refer to the simulation section for the illustration of the convergence of the proposed optimal algorithm.

C. Suboptimal Resource Allocation Algorithm Design

The iterative resource allocation algorithm proposed in the last section leads to the optimal system performance. However, the algorithm has a non-polynomial time computational complexity since it needs to solve an MILP master problem in each iteration. In this section, we propose a suboptimal resource allocation algorithm which has a polynomial time computational complexity. We start the suboptimal resource allocation algorithm design by focusing on the reformulated optimization problem in (14).

1) **Problem reformulation via difference of convex functions programming:** The major obstacle in solving (14) is to handle the binary constraint. In fact, constraint C10 is equivalent to

$$C10a: 0 \leq s_{l,k} \leq 1 \quad \text{and} \quad C10b: \sum_{l=1}^{L} \sum_{k=1}^{K} s_{l,k} - \sum_{l=1}^{L} \sum_{k=1}^{K} s_{l,k}^2 \leq 0,$$

(30)
where optimization variable $s_{l,k}$ in C10a is a continuous value between zero and one and C10b is the difference of two convex functions. By using the SDP relaxation approach as in the optimal resource allocation algorithm, we can rewrite the optimization problem as

$$\min_{W_k, V \in \mathbb{H}^{N \times L}} \sum_{k=1}^{K} \text{Tr}(W_k) + \text{Tr}(V)$$

s.t. $C1 - C9, C10a, C11, C12, C14$. (31)

On the other hand, for a large constant value of $\phi \gg 1$, we can follow a similar approach as in [26] to show that the optimization problem in (31) is equivalent to the following problem:

$$\min_{W_k, V \in \mathbb{H}^{N \times L}} \sum_{k=1}^{K} \text{Tr}(W_k) + \text{Tr}(V) + \phi \left( \sum_{l=1}^{L} \sum_{k=1}^{K} s_{l,k} - \sum_{l=1}^{L} \sum_{k=1}^{K} s_{l,k}^2 \right)$$

s.t. $C1 - C9, C10a, C11, C12, C14$, (32)

where $\phi$ acts as a large penalty factor for penalizing the objective function for any $s_{l,k}$ that is not equal to 0 or 1. We note that the constraints in (32) span a convex set which allows the development of an efficient resource allocation algorithm. The problem in (32) is known as difference of convex functions (d.c.) programming due to the convexity of $g(s_{l,k}) = \sum_{l=1}^{L} \sum_{k=1}^{K} s_{l,k}^2$. Here, we can apply the successive convex approximation to obtain a local optimal solution of (32) [27].

2) Iterative suboptimal algorithm: The first step is to linearize the convex function $g(s_{l,k})$. Since $g(s_{l,k})$ is a differentiable convex function, then the following inequality [22]

$$g(s_{l,k}) \geq g(s_{l,k}^{(i)}) + \nabla s_{l,k} g(s_{l,k}^{(i)})(s_{l,k} - s_{l,k}^{(i)})$$

always holds for any feasible point $s_{l,k}^{(i)}$. As a result, for a given value $s_{l,k}^{(i)}$, the optimal value of the optimization problem,

$$\min_{W_k, V \in \mathbb{H}^{N \times L}} \sum_{k=1}^{K} \text{Tr}(W_k) + \text{Tr}(V) + \phi \left( \sum_{l=1}^{L} \sum_{k=1}^{K} s_{l,k} - \sum_{l=1}^{L} \sum_{k=1}^{K} s_{l,k}^2 \right)$$

s.t. $C1 - C9, C10a, C11, C12, C14$, (34)

leads to an upper bound of (32). Then, an iterative algorithm is used to tighten the upper bound as summarized in Table III. We first initialize the values of $s_{l,k}^{(i)}$ and the iteration index $i = 0$.

5 This method is also known as majorization minimization. There are infinite many of d.c. representations for (31) leading to different successive convex programs. Please refer to [26] for a more detailed discussion for d.c. programming.
TABLE II
SUBOPTIMAL ITERATIVE RESOURCE ALLOCATION ALGORITHM

| Algorithm Successive Convex Approximation |
|-------------------------------------------|
| 1: Initialize the maximum number of iterations $L_{\text{max}}$, penalty factor $\phi \gg 1$, iteration index $i = 0$, and $s_{l,k}^{(i)}$ |
| 2: repeat {Loop} |
| 3: Solve (34) for a given $s_{l,k}^{(i)}$ and obtain the intermediate resource allocation policy {$W'_k, V', e_{S'}^S, s_{l,k}^{'(i)}$} |
| 4: Set $s_{l,k}^{(i+1)} = s_{l,k}^{'(i)}, i = i + 1$ |
| 5: until Convergence or $i = L_{\text{max}}$ |

Then, we solve (34) for a given value of $s_{l,k}^{(i)}$, cf. line 3. Subsequently, we update $s_{l,k}^{(i+1)}$ with the intermediate solution $s_{l,k}^{(')}$. The main idea of the proposed iterative method is to generate a sequence of feasible solutions $s_{l,k}^{(i)}$ by successively solving the convex upper bound problem (34). The procedure is repeated iteratively until convergence or the maximum number of iterations is reached. We note that the proposed suboptimal algorithm converges to a local optimal solution of (32) with polynomial time computational complexity as shown in [27]. Besides, by exploiting Theorem 1, $\text{Rank}(W_k) = 1$ is guaranteed despite the adopted SDP relaxation. On the contrary, although the optimal resource allocation algorithm achieves the optimal system performance, it has a non-polynomial time computational complexity.

Remark 3: The proposed algorithm requires $s_{l,k}^{(i)}$ to be a feasible point for the initialization, i.e., $i = 0$. This point can be obtained by e.g. solving (32) for $\phi = 0$.

V. RESULTS

In this section, we evaluate the network performance of the proposed resource allocation design via simulations. There are $L = 3$ RRHs, $K = 5$ IRs, and $M = 2$ ERs in the system. We focus on the network topology shown in Figure 3. The inter-site distance between any two RRHs is 150 meters which is a typical distance for a micro-cellular setup. The three RRHs construct an equilateral triangle while the IRs and ERs are uniformly distributed inside a disc with radius 150 meters centered at the centroid of the triangle. The performance of the proposed algorithms is compared with the performances of a fully cooperative transmission scheme (cooperative transmission and energy cooperation), a fully cooperative transmission scheme with perfect CSI but without (w/o) energy cooperation, and a traditional system with co-located transmit antennas. For the full cooperation scheme, the solution is obtained by setting $C_{l\text{Bmax}} \rightarrow \infty$, and solving (13) by SDP relaxation. For the fully cooperative scheme with perfect CSI but w/o energy cooperation, we set $C_{l\text{Bmax}} \rightarrow \infty$, and $P_{l\text{Tmax}} = \infty$, restrict the RRHs to not share the harvested energy, and solve (13) by SDP relaxation. For the co-located transmit antenna system, we assume that there is only one RRH located at the center.
Fig. 3. A distributed antenna network simulation topology with $L = 3$ RRHs, $K = 5$ IRs, and $M = 2$ ERs.

Fig. 4. Normalized renewable energy harvesting profile for the considered distributed antenna network.

Fig. 5. Normalized renewable energy harvesting profile for the three RRHs.

of the system, which is equipped with the same number of antennas as all RRHs combined in the distributed setting, i.e., $N_T L$. Besides, the CP is at the same location as the RRH for the co-located transmit antenna system and the backhaul is not needed. Furthermore, we set $P_{T_{l}}^{\text{max}} = \infty$ and assume an unlimited energy supply for the co-located transmit antenna system to study its power consumption. Unless specified otherwise, we assume that the maximum SINR tolerance of each ER is set to $\Gamma_{\text{tol}} = 0$ dB. We adopt an Euclidean sphere for the CSI uncertainty region, i.e., $\Xi_m = \mathbf{I}_{N_T L}$. Furthermore, we define the normalized maximum channel estimation error of ER $m$ as $\sigma^2_{\text{est}} = \frac{\epsilon_m^2}{\|g_m\|^2} = 0.05$ where $\forall m \in \{1, \ldots, M\}$. The parameters adopted in the simulations are summarized in Table III.

On the other hand, we adopt the normalized renewable energy harvesting profile specified in Figure 4, for which the data was obtained at August 01, 2014, in Belgium\footnote{Please refer to [http://www.elia.be/en/grid-data/power-generation/] for details regarding energy harvesting data.}. The data is averaged every 15 minutes and there are 96 sample points per 24 hours. We denote the normalized renewable energy harvesting profile data points for wind energy and solar energy as $\xi_w = [\xi_{w,1}, \ldots, \xi_{w,96}]$ and
TABLE III

| System Parameters                                      | Value                                                                 |
|--------------------------------------------------------|----------------------------------------------------------------------|
| Carrier center frequency and path loss exponent        | 915 MHz and 2.7                                                      |
| Multipath fading distribution                          | Rayleigh fading                                                      |
| Total noise variance, $\sigma_n^2$                    | $-23$ dBm                                                           |
| Circuit power consumption at the CP and the $l$-th RRH  | 40 dBm and 30 dBm                                                   |
| Power amplifier power efficiency                       | $1/\rho = 0.38$                                                     |
| Max. transmit power allowance, $P_{t}^{T\text{max}}$, and min. required power transfer | 48 dBm and $-10$ dBm                                               |
| RF to electrical energy conversion efficiency, $\mu$, and penalty term, $\phi$ | $0.5$ and $10P_{t}^{T\text{max}}$                                  |
| B-coefficient matrix                                   | Obtained from example 4D in [17]                                    |

$\xi_s = [\xi_s,1, \ldots, \xi_s,96]$, respectively. We follow a similar approach as in [5] to generate the amount of harvested energy at each RRH for simulation. We assume that the CP has only enough energy to support its circuit power consumption and does not contribute energy to the energy cooperation between the RRHs. The three RRHs are equipped with both solar panels and wind turbines with different energy harvesting capabilities. The harvested energy over time at the three RRHs is given by $\xi_1 = E(0.5\xi_w + 0.5\xi_s)$, $\xi_2 = E(0.9\xi_w + 0.1\xi_s)$, and $\xi_3 = E(0.1\xi_w + 0.9\xi_s)$, respectively, as shown in Figure [5] where $E = 500$ Joules is a given constant indicating the maximum available energy from the solar panels and wind turbines. Thus, the maximum harvested energy for RRH $n \in \{1, \ldots, L\}$ at sample time $r \in \{1, \ldots, 96\}$ is given by $E_{n}^{\text{max}} = [\xi_n]_{1,r}$. The minimum required received SINRs for the five IRs are set to $\Gamma_{\text{req}_k} = [6, 9, 12, 15, 18]$ dB, respectively. In case of full cooperation, these five IRs require a total capacity of $20.5818$ bit/s/Hz averaged over the energy harvesting profiles. In particular, for each sample of energy harvesting profiles, we generate 1000 independent channel realizations for both path loss and multipath fading.

### A. Convergence of the Proposed Iterative Algorithms

Figure [6] illustrates the convergence of the proposed optimal and suboptimal algorithms for different total numbers of transmit antennas in the networks, $N_T L$. The backhaul capacity per link is 10 bits/s/Hz. It can be seen from the upper half of Figure [6] that the proposed optimal algorithm converges to the optimal solution, i.e., the upper bound value meets the lower bound value after less than 80 iterations. On the other hand, the suboptimal algorithm converges to a local optimal value after less than 10 iterations. We note that if a brute force approach is adopted to obtain a global optimal solution.

$^7$The minimum required power transfer is targeted at sensor applications.
without exploiting the structure of the problem, for $K = 5$ IRs and $L = 3$ RRHs, $2^{15}$ of SDPs need to be solved which may not be computational feasible in practice.

### B. Average Total Transmit Power

In Figure 7 we study the average total transmit power versus the total number of transmit antennas in the network, $N_T L$, for different resource allocation schemes. The performances of the proposed optimal and suboptimal iterative algorithms are shown for 80 and 10 iterations, respectively. It can be seen that the transmit power for the proposed optimal and suboptimal schemes decreases when the backhaul capacity per backhaul link increases from 10 bits/s/Hz to 15 bits/s/Hz. This is because the increased backhaul capacity facilitates joint transmission and thus reduces the total transmit power. However, the transmit power of all considered schemes/systems decreases gradually with the total number of transmit antennas in the network. In fact, extra degrees of freedom can be exploited for resource allocation when more antennas are available for the cooperation between RRHs. Furthermore, the performance gap between the proposed optimal algorithm and fully cooperative transmission is expected to decrease with increasing $N_T L$. For sufficiently large numbers of antennas at the RRHs, conveying the data of each IR to a subset of RRHs may be sufficient for guaranteeing the QoS requirements for reliable communication and efficient power transfer. The lower average total transmit power of fully cooperative transmission with energy cooperation comes at the expense of an exceedingly high backhaul capacity consumption. On the other hand, the proposed suboptimal algorithm achieves an excellent system performance even for the case of only 10 iterations.
Fig. 8. Average total transmit power (dBm) versus the normalized channel estimation error for different resource allocation schemes.

Compared to the two proposed schemes, it is expected that the co-located antenna scheme requires a higher transmit power since the co-located antenna system does not offer network type spatial diversity to combat the path loss. The performance of fully cooperative transmission with perfect CSI and w/o energy cooperation is not satisfactory compared to all other schemes. Specifically, RRH 3 mainly relies on the solar panel for energy harvesting and thus the available energy for RRH 3 is very limited during the night time. Therefore, despite the availability of perfect CSI and large numbers of distributed antennas in the system, the RRHs having more harvested renewable energy available are required to transmit with comparatively large powers for assisting the RRHs with smaller harvested renewable energy. In fact, the RRHs have to cooperate wirelessly which is less power efficient than the cooperation via the mico-grid.

In Figure 8 we show the average total transmit power (dBm) versus the normalized channel estimation error for the proposed schemes with $N_T L = 18$ and 10 bits/s/Hz capacity per backhaul link. As can be observed, the average transmit power increases with the normalized channel estimation error except for the case of perfect CSI. The reason behind this is twofold. First, a higher transmit power for the artificial noise, $v$, is required to satisfy constraints C2 and C7 due to a larger uncertainty set for the CSI, i.e., $\Xi_m$. Second, a higher amount of power also has to be allocated to the information signal $w_k s_k, \forall k$, cf. $w_k s_k, \forall k$, for neutralizing the interference caused by the artificial noise at the desired IRs.
C. Average Total Harvested Power

In Figure 9, we study the average total harvested RF power versus the total number of transmit antennas for different resource allocation schemes. It can be observed that the total harvested power of the proposed schemes decreases monotonically with increasing number of transmit antennas. This is because the extra degrees of freedom offered by the increasing number of antennas improve the efficiency of resource allocation. In particular, the direction of beamforming matrix $W_k$ can be more accurately steered towards the IRs which reduces the power allocation to $W_k$ and the leakage of power to the ERs. This also explains the lower harvested power for the full cooperation scheme with energy cooperation since it can exploit all transmit antennas in the network for joint transmission. Besides, the ERs in the full cooperation scheme w/o energy cooperation harvest the highest amount of power on average at the expense of the highest average total transmit power. Furthermore, although the system with co-located antennas consumes a higher transmit power, it does not always lead to the largest harvested power at the ERs in all considered scenarios. Indeed, a large portion of radiated power in the co-located antenna system is used to combat the path loss which emphasizes the inherent spatial diversity in distributed antenna systems for power efficient transmission. We also show in Figure 9 the minimum required total harvested power which is computed by assuming that constraint C7 is satisfied with equality for all ERs. Despite the imperfection of the CSI, because of the adopted robust optimization framework, the proposed optimal and suboptimal resource allocation schemes are able to guarantee the minimum harvested energy required by constraint C7 in every time instant. On the other hand, Figure 10 depicts the average total harvested power versus the normalized channel estimation error for the proposed schemes with $N_T L = 18$ and 10 bits/s/Hz backhaul capacity per backhaul link. For imperfect CSI, the harvested power increases with the channel estimation error. In fact, to fulfill the QoS requirements on power transfer and communication secrecy, more transmit power is required for larger $\sigma_{\text{est}}^2$, which leads to a higher energy level in the RF for energy harvesting.

D. Average System Secrecy Rate

Figure 11 illustrates the average system secrecy rate versus the total number of transmit antennas for different resource allocation schemes. For imperfect CSI, it can be seen that the average system secrecy rate decreases with the number of transmit antennas in the system. In fact, more transmit antennas can focus the energy of the information signals more accurately on the desired IRs which minimizes the information leakage to the ERs and improves the information delivery efficiency.
Fig. 10. Average total harvested RF power (dBm) versus the normalized channel estimation error for different resource allocation schemes.

Besides, as expected, the proposed suboptimal and optimal schemes achieve on average a higher secrecy rate compared to fully cooperative transmission with energy cooperation since for limited backhaul links the RRHs have to transmit with slightly higher powers. On the other hand, Figure 11 also reveals that all considered schemes/systems relying on imperfect CSI are able to guarantee the secrecy QoS requirement because of the proposed robust optimization.

VI. CONCLUSIONS

In this paper, we studied the resource allocation algorithm design for the wireless delivery of both secure information and renewable green energy to mobile receivers in distributed antenna communication systems. The algorithm design was formulated as a non-convex optimization problem with the objective to minimize the total network transmit power. The proposed problem formulation took into account the limited backhaul capacity, the sharing of harvested renewable green energy between RRHs, the imperfect CSI of the ERs, and QoS requirements for secure communication and efficient power transfer. An optimal iterative resource allocation algorithm was proposed for obtaining a global optimal solution based on the generalized Bender’s decomposition. To strike a balance between computational complexity and optimality, we also proposed a low complexity suboptimal algorithm. Simulation results showed that the proposed suboptimal iterative resource allocation scheme performs close to the optimal scheme. Besides, our results unveiled the potential power savings in SWIPT systems employing distributed antenna networks and renewable green energy sharing compared to centralized systems with multiple co-located antennas.
APPENDIX-PROOF OF PROPOSITION [1]

We start the proof by studying the solution of the dual problem in (23). For a given optimal dual variable $\Theta(i)$, we have

$$\Theta(i) = \arg \min_{W_k, V \in \mathbb{R}^{N_T \times L}, \Theta, \delta, \nu, s_{l,k}(i), \Phi(i)} \mathcal{L}(W_k, V, e^S, \delta, \nu, s_{l,k}(i), \Phi(i))$$

$$= \arg \min_{W_k, V \in \mathbb{R}^{N_T \times L}, \Theta, \delta, \nu} f_0(W_k, V) + f_1(W_k, V, e^S, \delta, \nu, \Phi(i)) + \sum_{k=1}^K \sum_{l=1}^L \beta_{k,l} \text{Tr}(W_k R_l) - s_{l,k}(i) P_{t_{\text{max}}})$$

$$= \arg \min_{W_k, V \in \mathbb{R}^{N_T \times L}, \Theta, \delta, \nu} f_0(W_k, V) + f_1(W_k, V, e^S, \delta, \nu, \Phi(i)) + \sum_{k=1}^K \sum_{l=1}^L \beta_{k,l} \text{Tr}(W_k R_l), \quad \text{(35)}$$

where the first equality is due to the Karush-Kuhn-Tucker (KKT) conditions of the SDP relaxed problem in (18). On the other hand, we can rewrite function $\xi(\Phi(t), s_{l,k}, t \in \{1, \ldots, i\}$ as

$$\xi(\Phi(t), s_{l,k}) = \minimize_{W_k, V \in \mathbb{R}^{N_T \times L}, \Theta, \delta, \nu} f_0(W_k, V) + f_1(W_k, V, e^S, \delta, \nu, \Phi(t)) + f_2(W_k, s_{l,k}, \Phi(t))$$

$$= \left\{ \begin{array}{l}
\minimize_{W_k, V \in \mathbb{R}^{N_T \times L}, \Theta, \delta, \nu} f_0(W_k, V) + f_1(W_k, V, e^S, \delta, \nu, \Phi(t)) + \sum_{k=1}^K \sum_{l=1}^L \beta_{k,l} \text{Tr}(W_k R_l)
\end{array} \right\}$$

$$- \sum_{k=1}^K \sum_{l=1}^L \beta_{k,l} s_{l,k} P_{t_{\text{max}}} \quad \text{(36)}$$

As a result, the primal solution in the $t$-th iteration, $\Theta(t)$, is also the solution for the minimization in the master problem in (36) for the $t$-th constraint in (27b). Similarly, we can use the same approach to prove that the solution of (26) is also the solution of (29).

REFERENCES

[1] D. W. K. Ng and R. Schober, “Resource Allocation for Coordinated Multipoint Networks with Wireless Information and Power Transfer,” in Proc. IEEE Global Telecommun. Conf., Dec. 2014.
[2] S. Leng, D. W. K. Ng, and R. Schober, “Power Efficient and Secure Multiuser Communication Systems with Wireless Information and Power Transfer,” in Proc. IEEE Intern. Commun. Conf., Jun. 2014.
[3] “Green Energy Solution by Huawei.” [Online]. Available: http://www.huawei.com/en/solutions/go-greener/hw-001339-greencommunication-energyefficient/energy.
[4] D. Ng, E. Lo, and R. Schober, “Energy-Efficient Resource Allocation in OFDMA Systems with Hybrid Energy Harvesting Base Station,” IEEE Trans. Wireless Commun., vol. 12, pp. 3412–3427, Jul. 2013.
[5] J. Xu and R. Zhang, “CoMP Meets Energy Harvesting: A New Communication and Energy Cooperation Paradigm,” to appear in IEEE Trans. Veh. Technol., Aug. 2014.
[6] P. Grover and A. Sahai, “Shannon Meets Tesla: Wireless Information and Power Transfer,” in Proc. IEEE Intern. Sympos. on Inf. Theory, Jun. 2010, pp. 2363 –2367.
[7] R. Zhang and C. K. Ho, “MIMO Broadcasting for Simultaneous Wireless Information and Power Transfer,” IEEE Trans. Wireless Commun., vol. 12, pp. 1989–2001, May 2013.
[8] D. W. K. Ng, E. S. Lo, and R. Schober, “Wireless Information and Power Transfer: Energy Efficiency Optimization in OFDMA Systems,” *IEEE Trans. Wireless Commun.*, vol. 12, pp. 6352 – 6370, Dec. 2013.

[9] ——, “Robust Beamforming for Secure Communication in Systems with Wireless Information and Power Transfer,” *IEEE Trans. Wireless Commun.*, vol. 13, pp. 4599–4615, Aug. 2014.

[10] L. Liu, R. Zhang, and K.-C. Chua, “Secrecy Wireless Information and Power Transfer with MISO Beamforming,” *IEEE Trans. Signal Process.*, vol. 62, pp. 1850–1863, Apr. 2014.

[11] D. Lee, H. Seo, B. Clerckx, E. Hardouin, D. Mazzarese, S. Nagata, and K. Sayana, “Transmission and Reception in LTE-Advanced: Deployment Scenarios and Operational Challenges,” *IEEE Commun. Mag.*, vol. 50, pp. 148–155, Feb. 2012.

[12] R. Irmer, H. Droste, P. Marsch, M. Griefer, G. Fettweis, S. Brueck, H. P. Mayer, L. Thiele, and V. Jungnickel, “Coordinated Multipoint: Concepts, Performance, and Field Trial Results,” *IEEE Commun. Mag.*, vol. 49, pp. 102–111, Feb. 2011.

[13] D. W. K. Ng, E. S. Lo, and R. Schober, “Energy-Efficient Resource Allocation in Multi-Cell OFDMA Systems with Limited Backhaul Capacity,” *IEEE Trans. Wireless Commun.*, vol. 11, pp. 3618–3631, Oct. 2012.

[14] J. Zhao, T. Quek, and Z. Lei, “Coordinated Multipoint Transmission with Limited Backhaul Data Transfer,” *IEEE Trans. Wireless Commun.*, vol. 12, pp. 2762–2775, Jun. 2013.

[15] B. Dai and W. Yu, “Sparse Beamforming for Limited-Backhaul Network MIMO System via Reweighted Power Minimization,” in *Proc. IEEE Global Telecommun. Conf.*, Dec. 2013.

[16] G. Zheng, K. K. Wong, and T. S. Ng, “Robust Linear MIMO in the Downlink: A Worst-Case Optimization with Ellipsoidal Uncertainty Regions,” *EURASIP J. Adv. Signal Process.*, vol. 2008, 2008, Article ID 609028.

[17] A. J. Wood and B. F. Wollenberg, *Power Generation, Operation, and Control*. Wiley-Interscience.

[18] J. Yang and S. Ulukus, “Optimal Packet Scheduling in an Energy Harvesting Communication System,” *IEEE Trans. Commun.*, vol. 60, pp. 220–230, Jan. 2012.

[19] R. Ramamonjison, A. Haghnejahdar, and V. Bhargava, “Joint Optimization of Clustering and Cooperative Beamforming in Green Cognitive Wireless Networks,” *IEEE Trans. Wireless Commun.*, vol. 13, pp. 982–997, Feb. 2014.

[20] C. A. Floudas, *Nonlinear and Mixed-Integer Optimization: Fundamentals and Applications*, 1st ed. Oxford University Press, 1995.

[21] D. Li and X. Sun, *Nonlinear Integer Programming*. Springer, 2006.

[22] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[23] M. Grant and S. Boyd, “CVX: Matlab Software for Disciplined Convex Programming, version 2.0 Beta,” [Online] Available: https://cvxr.com/cvx]. Sep. 2013.

[24] “MOSEK ApS: Software for Large-Scale Mathematical Optimization Problems, Version 7.0.0.111,” Apr. 2014. [Online]. Available: http://www.mosek.com/

[25] “GUROBI Optimization, State-of-the-Art Mathematical Programming Solver, v5.6,” Apr. 2014. [Online]. Available: http://www.gurobi.com/

[26] E. Che, H. Tuan, and H. Nguyen, “Joint Optimization of Cooperative Beamforming and Relay Assignment in Multi-User Wireless Relay Networks,” *IEEE Trans. Wireless Commun.*, vol. PP, no. 99, pp. 1–1, 2014.

[27] Q. T. Dinh and M. Diehl, “Local Convergence of Sequential Convex Programming for Nonlinear Programming,” in *Chapter of Recent Advances in Optimization and Its Application in Engineering*. Oxford University Press, 2010.