Robust Chance-Constrained Optimization for Power-Efficient and Secure SWIPT Systems

Tuan Anh Le, Quoc-Tuan Vien, Huan X. Nguyen, Derrick Wing Kwan Ng, and Robert Schober

Abstract—In this paper, we propose beamforming schemes to simultaneously transmit data securely to multiple information receivers (IRs) while transferring power wirelessly to multiple energy-harvesting receivers (ERs). Taking into account the imperfection of the instantaneous channel state information (CSI), we introduce a chance-constrained optimization problem to minimize the total transmit power while guaranteeing data transmission reliability, data transmission security, and power transfer reliability. As the proposed optimization problem is non-convex due to the chance constraints, we propose two robust reformulations of the original problem based on safe-convex-approximation techniques. Subsequently, applying semidefinite programming relaxation (SDR), the derived robust reformulations can be effectively solved by standard convex optimization packages. We show that the adopted SDR is tight and thus the globally optimal solutions of the reformulated problems can be recovered. Simulation results confirm the superiority of the proposed methods in guaranteeing transmission security compared to a baseline scheme. Furthermore, the performance of proposed methods can closely follow that of a benchmark scheme where perfect CSI is available for resource allocation.

I. INTRODUCTION

In a simultaneous wireless information and power transfer (SWIPT) system, in order to harvest meaningful amounts of energy, the energy-harvesting receivers (ERs) must be located closer to the transmitter than the conventional information receivers (IRs) [1]–[7]. Being closer to the transmitter, the ERs will receive stronger radio frequency (RF) signals than the IRs. Since information conveyed via RF signals is always at risks of being overheard by eavesdroppers due to the broadcast nature of wireless channels, the information intended to the IRs has to be protected in order to prevent potential eavesdropping by the ERs. Due to its high computational complexity, conventional upper-layer cryptography may cause a high energy consumption at the receivers. Therefore, such techniques may not be suitable for protecting information in SWIPT systems [2], [8] as SWIPT devices are usually energy limited. Instead, physical layer security [9]–[11], where fading, noise, and interference are exploited, is considered to be an effective method for providing secure information transmission in SWIPT systems [2], [12].

Both information transmission and power transfer are equally important in SWIPT systems. Due to current hardware limitations [1], [13], the RF-to-direct-current-energy-conversion efficiencies at ERs are typically low. As a result, a relatively high transmit power is required to compensate for the path loss in the propagation environment and the energy loss in the power-conversion circuitry. On the other hand, high transmit powers increase the risk of information leakage to the ERs. Therefore, a power-efficient strategy satisfying the information transmission and power transfer requirements of the IRs and ERs, respectively, is required. Beamforming is known to improve the power efficiency of wireless communications [14]–[18], and is also a promising candidate for power-efficient and secure SWIPT [4], [5].

The beamforming design is usually formulated as an optimization problem taking into account the system’s quality of service (QoS) requirements[4] which are specified by the system operator/designer. In particular, the channel state information (CSI) of the channels between the transmitter and the receivers, which can be obtained by appropriate channel estimation techniques in practice, is exploited to optimally control the powers and phases of the beamformer [19]. Hence, the CSI plays an important role in optimizing the performance of SWIPT systems. In this context, most related works have assumed the availability of perfect CSI, i.e., there are no CSI estimation errors, for the design of the beamformers for SWIPT systems, e.g., [12], [20]–[24]. Unfortunately, CSI estimation errors are unavoidable due to the nature of wireless channels [5], [25]. Achieving near-perfect CSI estimation, especially for multiple users and multiple antennas, entails a high cost in terms of the required signalling overhead. Hence, assuming perfect CSI for beamforming design either imposes a high burden on communication systems or results in a resource allocation mismatch such that the QoSs of the users of the system cannot be guaranteed. Therefore, robust beamforming designs taking into account the imperfection of the CSI are desirable for relieving the signalling-overhead burden while maintaining the users’ QoSs in practical SWIPT systems.

In the literature, the imperfection of the estimated CSI for the channel between a transmitter and a receiver is usually modeled as an error vector with random elements. Due to the randomness and continuity of the error vector, an infinite number of constraints have to be met to guarantee the QoS. This leads to an intractable beamforming design problem. To overcome this obstacle, the norms of the error vectors are often assumed to be bounded by known values [4], [5], [12], [20]. Then, using the S-procedure [27], the QoS constraints
of the SWIPT system can be replaced by a finite number of constraints representing upper bounds on the CSI errors\cite{4},\cite{5},\cite{12},\cite{25},\cite{26},\cite{28}. Such a conservative design approach requires an exceedingly large amount of system resources to protect rarely occurring worst cases. Hence, less conservative approaches have recently been proposed which tolerate the violation of the QoS constraints with a certain chance or probability\cite{29}–\cite{32}.

Motivated by such probabilistic approaches, this paper focuses on the design of power-efficient transmission strategies for secure SWIPT systems employing imperfect CSI. The contributions of this paper can be summarized as follows.

• Taking into account the imperfection of the CSI, we propose an outage-based chance optimization problem with the objective to minimize the total transmit power subject to the following three sets of QoS constraints: i) the probability/chance that the received SINRs at the IRs are above required levels is higher than predefined targets; ii) the probability/chance that the leakage SINRs at the ERs exceed secure levels is below a threshold; iii) the probability/chance that the powers received by the ERs are above required levels is greater than a prescribed value. The aforementioned three types of constraints guarantee data transmission reliability, data transmission security, and power transfer reliability, respectively.

• Since the adopted probabilistic/chance constraints are non-convex, we employ two different mathematical tools, i.e., the S-procedure\cite{27} and a Bernstein-type inequality\cite{33}, to develop two safe approximations\cite{34} of the original optimization problem. Using semidefinite programming (SDP) relaxation, the derived safe approximations are transformed into tractable SDPs which can be optimally solved by a standard interior-point method (IPM), e.g., the SeDuMi solver in CVX\cite{35}.

• The adopted SDP relaxations of the derived safe approximations are proved to be tight by showing that the relaxed/transformed problems always yield rank-one optimal solutions, and hence also constitute optimal solutions to the safely approximated problems. Particularly, in our rank-one proof, we introduce a novel method to convert a second-order-cone-programming (SOCP) constraint into a linear-matrix-inequality (LMI) constraint.

• Based on the worst-case runtime of the IPM\cite{36}, the computational complexities of the reformulated SDP versions of the proposed optimization problems are characterized.

This paper differs from the related works in\cite{4},\cite{5},\cite{12},\cite{29}–\cite{32} in terms of the problem formulation and the mathematical solution used.

Problem formulation: While this paper studies a secure SWIPT wireless system, the authors of\cite{30} considered a secure cognitive SWIPT system. This paper takes into account the imperfectness of the CSI of all IRs and all ERs whereas perfect knowledge of the IRs’ CSI was assumed in\cite{4},\cite{5}. A common target in the related literature is the minimization of the transmit power while ensuring the probabilities that the secrecy rate of each IR\cite{12},\cite{31},\cite{32} and the harvested power at each ER\cite{12},\cite{32} are above certain required levels. Thereby, the required secrecy rate constitutes one of the design parameters which can be set by the system operator/designer.

The authors of\cite{12} further proposed a two-stage optimization approach to decompose the constraint on the secrecy rate into constraints on the leakage SINRs at the ERs and the SINRs at the IRs by introducing two new optimization variables. Since these variables are tuned by the algorithm in\cite{12}, the operator/designer cannot enforce/guarantee any target values for the leakage SINRs or the IR SINRs. In contrast, the problem formulation in this paper enables the operator/designer to set target values for the maximum leakage SINRs at the ERs and the minimum SINRs at the IRs. As a result, the proposed problem formulation can guarantee that the IRs’ leakage information at the ERs remains below a secure level should they try to eavesdrop, e.g., below the decoding sensitivity of the ERs. In contrast, the approaches in\cite{12},\cite{30}–\cite{32} cannot accomplish this. If proper secrecy codes are used, e.g.,\cite{38}, then having only one constraint on the secrecy rate should be sufficient for a secure transmission. However, if conventional error correcting codes are used, even if the secrecy rate of the IR is kept above a certain required level\cite{12},\cite{30}–\cite{32}, the ERs may still be able to eavesdrop the IR’s message if their decoding sensitivity levels are lower than the leakage SINR. Hence, in this case, the schemes in\cite{12},\cite{30}–\cite{32} are less secure than the proposed schemes. Also, the optimization problem considered in this paper is more challenging than its counterparts in\cite{29}–\cite{32} as secure information transmission to multiple IRs is considered whereas\cite{29} does not consider secrecy at all and\cite{30}–\cite{32} protect only a single IR.

Mathematical aspects: References\cite{12},\cite{29}, and\cite{30} do not prove the rank-one property of the optimal beamforming solution when SDR is applied. On the other hand, the authors of\cite{31},\cite{32} have used some inequalities to transform their SOCP constraints into LMI constraints. Although the resulting transformed optimization problems are shown to yield rank-one solutions, the employed transformations reduce the size of the feasible region of the original problem which may lead to infeasibility. In contrast, in this paper, we transform an SOCP constraint into an LMI constraint without imposing any restriction. Therefore, the SDP relaxations of the considered optimization problems, i.e., the derived safe approximations, are tight.

Notation: Lower and upper case letter $y$ and $Y$: a scalar; bold lower case letter $y$: a column vector; bold upper case letter $Y$: a matrix; $||\cdot||$: the Euclidean norm; $||\cdot||_F$: the Frobenius norm; $(\cdot)^T$: the transpose operator; $(\cdot)^H$: the complex conjugate transpose operator; $\text{Tr}(\cdot)$: the trace operator; $\text{Pr}(\cdot)$: the probability of an event; $O(\cdot)$: the big-O notation; $Y \succeq 0$: $Y$ is positive semidefinite; $y \succeq 0$: all elements of vector $y$ are non-negative; $I_n$: an $n \times n$ identity matrix; $0_{m \times 1}$: an $m \times 1$ vector of all zero elements; $0_{n \times m}$: an $n \times m$ matrix of all zero elements; $\text{Re}(\cdot)$: the real part of a complex number; $\text{Eig}_{\text{max}}(Y)$: the maximum eigenvalue of $Y$; $s^*(Y)$: $\max[\text{Eig}_{\text{max}}(Y), 0]$; vec($Y$): stacking all the entries of $Y$ into a column vector; $\mathbb{R}$: the set of all
real scalars $c_{M 	imes 1}$: the set of all $M \times 1$ vectors with complex elements; $\mathbb{H}_{M \times M}$: the set of all $M \times M$ Hermitian matrices; $y \sim \mathcal{CN}(0, \sigma^2)$: $y$ is a zero-mean circularly symmetric complex Gaussian random variable with variance $\sigma^2$; $y \sim \mathcal{CN}(0, \mathbf{Y})$: $y$ is a zero-mean circularly symmetric complex Gaussian random vector with covariance matrix $\mathbf{Y}$; $\mathbf{Y}^{1/2}$: the square root of $\mathbf{Y}$; rank($\mathbf{Y}$): rank of $\mathbf{Y}$; $\forall$: for all.

II. System Model

In this paper, we consider a downlink SWIPT system where a transmitter equipped with $M > 1$ antennas simultaneously transmits information and power to $U$ IRs and $N$ ERs, respectively, using RF signals, see Fig. 1. Each IR and each ER is equipped with a single antenna. In order to provide secure communication, the ERs are treated as potential eavesdroppers. Hence, the power of the information-carrying signals intended for the IRs but received at the ERs should be kept low. On the other hand, due to the low energy conversion efficiency at the ERs, a high received power level is required at each ER to compensate for the power-conversion loss. Therefore, in this paper, we advocate the use of artificial noise to achieve both of these goals [4], [5], [26].

Let $\mathbf{h}_i \in \mathbb{C}^{M \times 1}$, $i \in \{1, \cdots, U\}$, and $\mathbf{g}_t \in \mathbb{C}^{M \times 1}$, $t \in \{1, \cdots, N\}$, represent the actual channel coefficients of the $i$-th IR and the $t$-th ER, respectively. Let $\mathbf{w}_i \in \mathbb{C}^{M \times 1}$ and $\mathbf{s}_j^{(I)} \sim \mathcal{CN}(0, 1)$, respectively, denote the beamforming vector and the data for the $i$-th IR. Let $\mathbf{v}_t \in \mathbb{C}^{M \times 1}$ and $\mathbf{s}_t^{(E)} \sim \mathcal{CN}(0, 1)$, respectively, be the artificial-noise beamforming vector and the artificial noise for the $t$-th ER. The signals received by the $i$-th IR and the $t$-th ER are, respectively, given by

$$
\begin{align*}
\mathbf{y}_i^{(I)} &= \sum_{j=1}^{U} \mathbf{h}_j^H \mathbf{w}_j \mathbf{s}_j^{(I)} + \sum_{r=1}^{M} \mathbf{h}_r^H \mathbf{v}_r \mathbf{s}_r^{(E)} + \mathbf{n}_i^{(I)} ,
\end{align*}
$$

and

$$
\begin{align*}
\mathbf{y}_t^{(E)} &= \sum_{j=1}^{U} \mathbf{g}_j^H \mathbf{w}_j \mathbf{s}_j^{(I)} + \sum_{p=1}^{N} \mathbf{g}_p^H \mathbf{v}_p \mathbf{s}_p^{(E)} + \mathbf{n}_t^{(E)}. 
\end{align*}
$$

Here, $\mathbf{n}_i^{(I)} \sim \mathcal{CN}(0, \sigma_{I,i}^2)$ and $\mathbf{n}_t^{(E)} \sim \mathcal{CN}(0, \sigma_{E,t}^2)$ are the zero-mean circularly symmetric complex additive white Gaussian noises observed at the $i$-th IR and the $t$-th ER, respectively.

We assume that the CSI estimation at the BS is imperfect. Similar to [4], [5], [12], [23], [25], [28], [32], we model the channel as $\mathbf{h}_i = \overline{\mathbf{h}}_i + \Delta \mathbf{h}_i$ and $\mathbf{g}_t = \overline{\mathbf{g}}_t + \Delta \mathbf{g}_t$, where $\overline{\mathbf{h}}_i \in \mathbb{C}^{M \times 1}$ and $\Delta \mathbf{h}_i \in \mathbb{C}^{M \times 1}$ are the estimated value of $\mathbf{h}_i$ and the corresponding estimation error, respectively; $\overline{\mathbf{g}}_t \in \mathbb{C}^{M \times 1}$ and $\Delta \mathbf{g}_t \in \mathbb{C}^{M \times 1}$ are the estimated value of $\mathbf{g}_t$ and the corresponding estimation error, respectively. We further assume that $\Delta \mathbf{h}_i \sim \mathcal{CN}(0, \mathbf{H}_i)$ and $\Delta \mathbf{g}_t \sim \mathcal{CN}(0, \mathbf{G}_t)$, where $\mathbf{H}_i \succeq 0$ and $\mathbf{G}_t \succeq 0$ are the channel estimation error covariance matrices which are assumed to be known for beamformer design. Interested readers are referred to [39] for techniques to estimate covariance matrices. Let $\Delta \mathbf{h}_i \sim \mathcal{CN}(0, \mathbf{H}_i)$ and $\Delta \mathbf{g}_t \sim \mathcal{CN}(0, \mathbf{G}_t)$, where $\mathbf{H}_i \geq 0$ and $\mathbf{G}_t \geq 0$ are the channel estimation error covariance matrices which are assumed to be known for beamformer design. Interested readers are referred to [39] for techniques to estimate covariance matrices. Let $\Delta \mathbf{h}_i \sim \mathcal{CN}(0, \mathbf{H}_i)$ and $\Delta \mathbf{g}_t \sim \mathcal{CN}(0, \mathbf{G}_t)$, where $\mathbf{H}_i \geq 0$ and $\mathbf{G}_t \geq 0$ are the channel estimation error covariance matrices which are assumed to be known for beamformer design. Interested readers are referred to [39] for techniques to estimate covariance matrices. Let $\Delta \mathbf{h}_i \sim \mathcal{CN}(0, \mathbf{H}_i)$ and $\Delta \mathbf{g}_t \sim \mathcal{CN}(0, \mathbf{G}_t)$, where $\mathbf{H}_i \geq 0$ and $\mathbf{G}_t \geq 0$ are the channel estimation error covariance matrices which are assumed to be known for beamformer design. Interested readers are referred to [39] for techniques to estimate covariance matrices.

We consider the following optimization problem:

A. Problem Formulation

Aiming to design a power-efficient beamforming scheme, we optimize the data beamforming vector set $\{\mathbf{w}_i\}$ and the artificial-noise beamforming vector set $\{\mathbf{v}_t\}$ for minimization of the total transmit power subject to probabilistic/chance constraints on SINR outages, leakage-SINR outages, and power-transfer outages. The design is formulated as the following optimization problem:

\[
\begin{align*}
\min_{\{\mathbf{w}_i\}, \{\mathbf{v}_t\}} & \sum_{i=1}^{U} ||\mathbf{w}_i||^2 + \sum_{t=1}^{N} ||\mathbf{v}_t||^2 \\
\text{s. t.} & \quad \Pr(\Gamma_i(\{\mathbf{w}_i\}, \{\mathbf{v}_t\}) \geq \gamma) \geq 1 - \rho_i, \forall i, \\
& \quad \Pr(\Gamma_t^{(E)}(\{\mathbf{w}_i\}, \{\mathbf{v}_t\}) \leq \gamma_t^{(E)}) \geq 1 - \rho_t^{(E)}, \forall t, \\
& \quad \Pr(\Phi_t(\{\mathbf{w}_i\}, \{\mathbf{v}_t\}) \geq P_t) \geq 1 - \gamma_t, \forall t,
\end{align*}
\]
where $\rho_i \in (0, 1], \rho_i^{(t)} \in (0, 1]$, and $\varrho_i \in (0, 1]$ are the predefined maximum tolerable probabilities/chances of SINR outages, leakage-SINR outages, and power-transfer outages, respectively. The events $\Gamma_i(\{|w_i\}, \{|v_i\}|) \geq \gamma_i$ and $\Gamma_i^{(t)}(\{|w_i\}, \{|v_i\}|) \leq \gamma_i^{(t)}$ in the first and second sets of probabilistic/chance constraints in (6) are non-convex with respect to $\{|w_i\}$ and $\{|v_i\}|$. In the sequel, we transform these events into convex forms by introducing new variables.

To this end, we define data beamforming matrix $W_t = w_t w_t^H$ and artificial-noise beamforming matrix $V_t = v_t v_t^H$, where $W_t \succeq \mathbf{0}, V_t \succeq \mathbf{0}, W_t \in \mathbb{H}^{M \times M}, V_t \in \mathbb{H}^{M \times M}$, and $W_t$ and $V_t$ are rank-one matrices. Using $x^H y = y^H x$, we rewrite the SINR event of IR $i$, $\Gamma_i(\{|w_i\}, \{|v_i\}|) \geq \gamma_i$, as:

$$\begin{equation}
(\mathbf{h}_i + \mathbf{H}_i^{1/2} \mathbf{e}_i)^H \mathbf{A}_i (\mathbf{h}_i + \mathbf{H}_i^{1/2} \mathbf{e}_i) \geq \sigma_i^2,
\end{equation}
$$

where $\mathbf{A}_i = (1 + \frac{1}{\gamma_i}) W_t - C$ and $C = \sum_{j=1}^U W_j + \sum_{j=1}^N V_j.$

Further manipulations exploiting the property $\mathbf{h}_i^H \mathbf{H}_i = \mathbf{h}_i$ of covariance matrices lead to the following equivalent form of (7):

$$f_i(\mathbf{e}_i) \triangleq e^{H_i} \mathbf{H}_i^{1/2} \mathbf{A}_i \mathbf{H}_i^{1/2} e_i + 2 \text{Re}(e^{H_i} \mathbf{H}_i^{1/2} \mathbf{h}_i) + \mathbf{h}_i^H \mathbf{A}_i \mathbf{h}_i - \sigma_i^2 \geq 0.
$$

Similarly, the information leakage event of IR $i$, $\Gamma_i^{(t)}(\{|w_i\}, \{|v_i\}|) \leq \gamma_i^{(t)}$, can be recast as:

$$k_i^{(t)}(\mathbf{r}_i) \triangleq r_i^H G_i^{1/2} B_i G_i^{1/2} r_i + 2 \text{Re}(r_i^H G_i^{1/2} B_i \mathbf{g}_i) + \mathbf{g}_i^H B_i \mathbf{g}_i + \sigma_e^2 \geq 0,
$$

where $B_i = C - (1 + \frac{1}{\gamma_i}) W_t$.

Furthermore, the harvested power event of ER $t$, $\Phi_t(\{|w_i\}, \{|v_i\}|) \geq \rho_t$, is equivalent to:

$$d_i(\mathbf{e}_i) \triangleq r_i^H G_i^{1/2} C_i G_i^{1/2} r_i + 2 \text{Re}(r_i^H G_i^{1/2} C_i \mathbf{g}_i) + \mathbf{g}_i^H C_i \mathbf{g}_i - \rho_t \geq 0.
$$

Using (8), (9), and (10), (6) can be equivalently stated as:

$$\min_{\{|w_i\}; |v_i| \in \mathbb{H}^{M \times M}} \text{Tr} \left( \sum_{i=1}^U W_i + \sum_{i=1}^N V_i \right)

\text{s. t.} \quad \text{Pr}(f_i(\mathbf{e}_i) \geq 0) \geq 1 - \rho_i, \forall i, \quad \text{Pr}(k_i^{(t)}(\mathbf{r}_i) \geq 0) \geq 1 - \rho_i^{(t)}, \forall i, \forall t, \quad \text{Pr}(d_i(\mathbf{e}_i) \geq 0) \geq 1 - \varrho_i, \forall i, \forall t, \quad W_t \succeq \mathbf{0}, \forall i, \forall v_i \succeq \mathbf{0}, \forall t, \quad \text{rank}(W_t) = 1, \forall t, \text{rank}(V_i) = 1, \forall t.$$

2Note that the event $\Phi_t(\{|w_i\}, \{|v_i\}|) \geq \rho_t$ in the third constraint in (6) is convex as it is a quadratic form in $\{|w_i\}$ and $\{|v_i\}$ with positive coefficients for the second-degree terms, see (6).

3A matrix is rank-one if and only if it has only one linearly independent column/row.

where $\{|W_i\} = \{|W_1, \cdots, W_U\|$ and $\{|V_i\} = \{|V_1, \cdots, V_N\|$ are two sets of beamforming matrices. Solving problem (11) is challenging due to the fact that the probabilistic constraints neither have simple closed-forms nor admit convexity. In other words, (11) is an NP-hard problem which cannot be solved in polynomial time. To overcome this challenge, our goal is to derive convex upper bounds for the chance constraints in (11).

### B. Safe Approximations

First, we tackle the intractable probabilistic constraints in (11) by replacing them by computationally tractable, i.e., convex approximations. These approximations result in a convex optimization problem with respect to $\{|W_i\|$ and $\{|V_i\|$, which can be regarded as a safe approximation if every feasible solution to the approximated problem is also feasible for the original problem. In other words, the optimal solution to the safe approximation problem is a feasible suboptimal solution to the original problem. Therefore, the problem based on safe approximations serves as an upper bound for the original problem.

Recently, three safe approximation methods have been introduced in [40], namely Method I: sphere bounding; Method II: Bernstein-type inequality; and Method III: decomposition-based large deviation inequality. In Method I, the chance constraints are approximated by assuming a spherical bound on the norm of the error vectors, while in Methods II and III, large deviation inequalities for complex Gaussian quadratic forms are utilized for constructing efficiently computable convex approximations. As reported in [40], Method II generally results in the tightest approximation among the three methods. The results in [40] also indicate that in terms of power efficiency and feasibility rate, Method II yields best performance followed by Method I and Method III. The poor performance of Method III is due to the fact that its approximation tightness is sacrificed for improved computational efficiency. Since, in this paper, we are interested in developing power-efficient strategies, in the following, we adopt Methods I and II to derive two safe approximations for (11). It is noted that our optimization problem is fundamentally different from the one considered in [40] where the objective is to minimize the total transmit power subject to IR rate outage constraints for a multiple-input single-output (MISO) downlink system.

4Although events $f_i(\mathbf{e}_i) \geq 0, k_i^{(t)}(\mathbf{r}_i) \geq 0,$ and $d_i(\mathbf{e}_i) \geq 0$ are convex, the corresponding probabilistic constraints in (11) are not convex.

5We note that a feasible solution to the original problem (11) may be infeasible for the approximated problem.

6The feasibility rate is the probability that the optimization problem is feasible [40].
IV. Proposed Safe Approximations

In this section, we introduce two safe approximations of (11) and perform a complexity analysis.

A. S-procedure Based Method

Let us assume that $e_i$ and $r_i$ are confined to the complex spherical sets $\xi_i \triangleq \{e_i \in C^{M \times 1} \mid \|e_i\|^2 \leq R_i^2\}$ and $\psi_i \triangleq \{r_i \in C^{M \times 1} \mid \|r_i\|^2 \leq Q_i^2\}$ having $M$ dimensions and radii $R_i$ and $Q_i$, respectively. Since the error vector $e_i \sim CN(0, I_M)$ is confined to the spherical set $\xi_i$, the probabilistic/chance constraint $\Pr(f_i(e_i) \geq 0) \geq 1 - \rho_i$ holds if

$$f_i(e_i) \geq 0 \quad \text{and} \quad \Pr(e_i \in \xi_i) \geq 1 - \rho_i. \quad (12)$$

The second condition in (12) is always true if the radius of the spherical set $\xi_i$ is selected such that $R_i = \sqrt{\frac{\xi_i(1-\rho_i)}{2}}$ where $\xi_i$ is the inverse cumulative distribution function of a Chi-square random variable with $m = 2M$ degrees of freedom. Therefore, using $|e_i|^2 = e_i^H I_M e_i$, the probabilistic constraint $\Pr(f_i(e_i) \geq 0) \geq 1 - \rho_i$ can be safely approximated by the following two constraints:

$$f_i(e_i) \geq 0 \quad \text{and} \quad e_i^H I_M e_i - \frac{\xi_i(1-\rho_i)}{2} \leq 0. \quad (13)$$

Applying similar steps for $\Pr(k_i^{(o)}(r_i) \geq 0) \geq \rho_i^{(o)}$ and $\Pr(d_k(r_k) \geq 0) \geq \psi_k$, we introduce the safe approximation of problem (11) as

$$\min_{\{W_i, V_i\} \in \mathbb{H}^{M \times M}} \sum_{i=1}^N \sum_{k=1}^N W_i \quad \text{s. t.} \quad f_i(e_i) \geq 0, \quad e_i^H I_M e_i - \frac{\xi_i(1-\rho_i)}{2} \leq 0, \quad \forall i,$$

$$k_i^{(o)}(r_i) \geq 0, \quad r_i^H I_M r_i - \frac{\xi_i(1-\rho_i)}{2} \leq 0, \quad \forall i, \quad \forall t,$$

$$d_k(r_k) \geq 0, \quad r_k^H I_M r_k - \frac{\xi_i(1-\psi_k)}{2} \leq 0, \quad \forall k, \quad \forall t,$$

$$W_i \geq 0, \quad \forall i, \quad \forall t, \quad V_i \geq 0, \quad \forall k, \quad \forall t,$$

$$\text{rank}(W_i) = 1, \quad \forall k, \quad \text{rank}(V_i) = 1, \quad \forall t. \quad (14)$$

Remark 1: The values of the radii $R_i$ and $Q_i$ of the spherical sets $\xi_i$ and $\psi_i$ are not required in (14) as they have been implicitly incorporated into the outages, e.g., $R_i = \sqrt{\frac{\xi_i(1-\rho_i)}{2}}$. In other words, these radii are determined by the maximum tolerable outage values, e.g., $\rho_i$.

The number of constraints in (14) is infinite due to the randomness and continuousness of the error vectors $e_i$ and $r_i$. To proceed, we introduce the following lemma.

Lemma 1 (S-procedure [22]): Let $m_n(x) = x^H Y_n x + 2Re(x^H Y_n) + c_n, \quad n \in \{1, 2\}$, where $Y_n \in \mathbb{H}^{M \times M}$, $y_n \in \mathbb{C}^{M \times 1}$, and $c_n \in \mathbb{R}$. If there exists an $x$ such that $m_n(x) < 0$, then $\forall x \in \mathbb{C}^{M \times 1}$, the following statements are equivalent:

1) $m_1(x) \geq 0$ and $m_2(x) \leq 0$ are satisfied $\forall x \in \mathbb{C}^{M \times 1}$.

2) There exists a $\beta \geq 0$ such that

$$\begin{bmatrix} Y_1 & y_1 \\ y_1^H & c_1 \end{bmatrix} + \beta \begin{bmatrix} Y_2 & y_2 \\ y_2^H & c_2 \end{bmatrix} \succeq 0.$$

Exploiting Lemma 1 and relaxing the rank-one constraints on $W_i$ and $V_i$, one can transform optimization problem (12) into the standard convex SDP form given in (15) at the top of next page where $a_i, A_i^{(o)}$, and $\beta$ are auxiliary optimization variables.

Remark 2: The LMI constraints in (15) are similar to those of the norm-bounded approaches in e.g., [20, 28]. However, the main difference between these two approaches is that the radii, i.e., the norms of the error vectors, are predefined values in the norm-bounded approaches whereas they can be controlled via the maximum tolerable outage probabilities in the proposed scheme.

To arrive at (15), we have relaxed the rank-one constraints on the beamforming matrices $W_i$ and $V_i$. In the following theorem, we will show that relaxing the rank-one constraints does not affect the optimality of the solutions.

Theorem 1: If problem (15) is feasible, then its optimal solution yields rank-one matrices $W_i$ and $V_i$.

Proof: Please refer to Appendix A.

As a consequence of Theorem 1, the optimal beamforming vectors $w^* \in C^{M \times 1}$ and $v^* \in C^{M \times 1}$ are the non-zero eigenvalues and $\vec{z}^{(w)}$ and $\vec{x}^{(v)}$ are the corresponding eigenvectors of the optimal rank-one matrices $W^*$ and $V^*$, respectively. Since (15) is a safe approximation of (11), $W_i^*$ and $V_i^*$ are suboptimal solutions to (11). As (11) is equivalent to (6), $w^*$ and $v^*$ are suboptimal solutions to the original problem (6).

B. Bernstein-type-inequality Based Method

In the previous subsection, bounded-norm conditions have been implicitly imposed on the error vectors via outage probability constraints to develop the S-procedure based method. Imposing these constraints may degrade the system performance as the feasible region of the original optimization problem is reduced. To overcome this problem, here, we adopt a different approach to obtain another type of robust formulation, i.e., another safe convex approximation of the original problem (11). This convex approximation approach is based on a large deviation inequality, i.e., a Bernstein-type inequality, which bounds the probability that a sum of random variables deviates from its mean [40]. To begin, let us recall the following lemma.

Lemma 2 (Bernstein-type inequality [33]): Consider the following random variable $f(x) = x^H Y x + 2Re(x^H u)$, where $x \sim CN(0, I_M), Y \in \mathbb{H}^{M \times M},$ and $u \in \mathbb{C}^{M \times 1}$. For all $\delta > 0$, the following statement holds:

$$\Pr\left(f(x) \geq \operatorname{Tr}(Y) - \sqrt{2\delta} \sqrt{\|Y\|_F^2 + 2\|u\|^2 - \delta^*}(Y)\right) \geq 1 - e^{-\delta^*}.$$

If (15) is infeasible, then some mechanisms such as admission control and quality-of-service adjustment are required at the Medium Access Control layer to restore the feasibility of the optimization problem. However, algorithms ensuring the feasibility of the optimization problem require a cross-layer design which is beyond the scope of this paper.
\[
\min_{\{W_i, V_i\} \in \mathbb{R}_{+}^{MxM}} \text{Tr}\left(\sum_{i=1}^{U} W_i + \sum_{i=1}^{N} V_i\right)
\]
\[
s. t.
\begin{align*}
& \text{Tr}\left(H_i^{1/2} A_i H_i^{1/2} + \alpha_i I_M\right) - \delta_i \alpha_i I_i - \delta_i b_i \geq 0, \\
& \text{Tr}\left(G_i^{1/2} B_i G_i^{1/2} + \beta_i I_M\right) - \delta_i \beta_i I_i - \delta_i a_i \geq 0,
\end{align*}
\]
\[
\text{where } a_i \text{ and } b_i \text{ are auxiliary optimization variables and } \mu_i = -\ln \delta_i. \text{ Therefore, the relaxed rank-one-constraint problem of the safe approximation of problem (11) can be written as}
\]
\[
\min_{\{W_i, V_i\} \in \mathbb{R}_{+}^{MxM}} \text{Tr}\left(\sum_{i=1}^{U} W_i + \sum_{i=1}^{N} V_i\right)
\]
\[
s. t.
\begin{align*}
& \text{Tr}\left(H_i^{1/2} A_i H_i^{1/2} + \alpha_i I_M\right) - \delta_i \alpha_i I_i - \delta_i b_i \geq 0, \\
& \text{Tr}\left(G_i^{1/2} B_i G_i^{1/2} + \beta_i I_M\right) - \delta_i \beta_i I_i - \delta_i a_i \geq 0,
\end{align*}
\]
\[
\text{The SOC constraints in (21), (23), and (27) are sub-cases of an SDP constraint since any SOC constraint can be recast in an LMI form using the Schur complement. Hence, the optimization problem in (20) is convex. Although the rank-one constraints on the beamforming matrices } W_i \text{ and } V_i \text{ have been relaxed to arrive at (20), the following theorem reveals that the relaxation preserves the optimality of the solution to the non-rank-one-relaxed problem, i.e., (20) with rank-one constraints on } W_i \text{ and } V_i.
\]
\[
\text{Theorem 2: For the optimal solution of (20) matrices } W_i \text{ and } V_i \text{ are always rank-one if the problem is feasible.}
\]
\[
\text{Proof: Please refer to Appendix}\]
\]
\[
\text{Hence, we can obtain beamforming vectors } w_i^* \text{ and } v_i^* \text{ from } W_i^* \text{ and } V_i^*, \text{ respectively, using the same technique as in Section IV-A as robust suboptimal solutions to the original problem (6).}
\]
\[
\text{C. Complexity Analysis}
\]
\[
\text{Hereafter, we refer to the S-procedure based method and the Bernstein-type-inequality based method as Method I and Method II, respectively. Since Method I, i.e., problem (15), and Method II, i.e., problem (20), contain LMI and SOC constraints, a standard IPM can be used to find their optimal solutions. To that end, we consider the worst-case runtime of the IPM to analyze the computational complexities of the proposed methods as follows.}
\]
\[
\text{Definition 1: For a given } \epsilon > 0, \text{ the set of } \{W_i^\epsilon\} \text{ and } \{V_i^\epsilon\} \text{ is called an } \epsilon\text{-solution to problem (15) or (20) if}
\]
\[
\text{Tr}\left(\sum_{i=1}^{U} W_i^\epsilon + \sum_{i=1}^{N} V_i^\epsilon\right) \leq \min_{\{W_i, V_i\} \in \mathbb{R}_{+}^{MxM}} \text{Tr}\left(\sum_{i=1}^{U} W_i + \sum_{i=1}^{N} V_i\right) + \epsilon.
\]
\[
\text{(31)}
\]
It can be observed that the number of decision variables of problems \(15\) and \(30\) is on the order of \((U + N)M^2\). Let \(\zeta = U + N\), \(\eta = \zeta + UN\), and \(n = O(\zeta M^3)\). We introduce the following lemma.

**Lemma 3:** The computational complexities to obtain \(\epsilon\)-solutions to problems \((15)\) and \((30)\) are respectively

\[
\ln(e^{-1}) \sqrt{\eta(M + 1) + \zeta M} \left[\eta(M + 1)\zeta(M + 1 + n) + \zeta M^2(M + n) + n^2\right] n
\]

(32)

and

\[
\ln(e^{-1}) \sqrt{(\zeta + \eta)M + 4\eta} \left[\eta \left(M^2 + M + 1\right)^2 + 2n + 2\right] + (\zeta + \eta)M^2(m + n) + n^2\right] n.
\]

(33)

**Proof:** Due to space limitation, we only provide a sketch of the proof. First, Method I, i.e., problem \((15)\), has \(\eta\) LMI constraints of size \((M + 1)\), and \(\zeta\) LMI constraints of size \(M\). Second, Method II, i.e., problem \((30)\), has 2\(\eta\) LMI constraints of size 1, \((\zeta + \eta)\) LMI constraints of size \(M\), and \(\eta\) SOC constraints of dimension \((M^2 + M + 1)\). Based on these observations, one can follow the same steps as in Section V-A to arrive at (32) and (33). Note that the terms \(\ln(e^{-1}) \sqrt{\eta(M + 1) + \zeta M}\) and \(\ln(e^{-1}) \sqrt{(\zeta + \eta)M + 4\eta}\) in (32) and (33) are the iteration complexities required for obtaining \(\epsilon\)-solutions to problems \((15)\) and \((30)\), respectively, while the remaining terms represent the per-iteration computation costs.

It can be observed from Lemma 3 that the computational complexity of Method I is lower than that of Method II in the following, we analyze the performance of the proposed methods in terms of power consumption.

V. SIMULATION RESULTS

We evaluate the performance of the two proposed methods and compare them against the probabilistic-constraint-based scheme introduced in [29], which is considered as the baseline scheme, the norm-bounded approach proposed in [12], and a benchmark scheme in [22]. For the baseline scheme, the data transmission and power transfer reliabilities are guaranteed. However, secure data transmission is not considered. The SeDuMi provided by the CVX optimization package [33] is employed to obtain the sets of the optimal beamforming matrices \(W^*_t\) and \(V^*_t\).

A. Simulation Setup

Consider a transmitter supporting two IRs and two ERs, i.e., \(U = N = 2\). The estimated channel vectors \(\hat{h}_i\) and \(\hat{g}_i\) are respectively modelled as: \(\hat{h}_i = H_i(t_i)\hat{h}_{iw}\) and \(\hat{g}_i = G_i(t_i^E)\hat{g}_{gw}\), where \(\hat{h}_{iw} \sim CN(0, I_M)\) and \(\hat{g}_{gw} \sim CN(0, I_M)\); \(H_i(t_i) = \overline{H}_i(\sqrt{f_c})\left(\frac{1}{\sqrt{f_c}}\right)^2\); \(G_i(t_i^E) = \overline{G}_i(\sqrt{f_c})\left(\frac{1}{\sqrt{f_c}}\right)^2\); \(t_i^E = 100\) ms, \(\forall i\); and \(t_i^E = 9\) m, \(\forall i\), are the distances from the transmitter to the IRs and ERs, respectively; \(c = 3 \times 10^8\) m/s is the speed of light; \(f_c = 900\) MHz is the carrier frequency; and \(\kappa = 2.7\) is the path loss exponent. The noise power at each IR and ER is assumed to be \(-70\) dBm. The error covariance matrices are given as \(H_i = \varepsilon(H_i(t_i))^{\dagger}I_M\) and \(G_i = \varepsilon(G_i(t_i^E))^{\dagger}I_M\) where \(\varepsilon = 0.001\). Monte-Carlo simulations have been carried out based on 500 channel realizations. The ER decoding sensitivity level is set to \(\gamma_i^0 = -5\) dB \(\forall i\). The SAR outage, leakage-SAR outage, and power-transfer outage probabilities are set to 10\%, i.e., \(P_i = \rho_i = \rho_i^0 = 0.1, \forall i\). The required power level at the ERs is \(P_i = -10\) dBm, \(\forall i\).

B. Performance Evaluation

Fig. 2 shows the average total transmit power versus the required SAR at each IR for different numbers of antennas for the proposed methods, the baseline scheme in [29], and the non-robust scheme in [22]. This figure indicates that for low-to-medium required IR SINRs, e.g. the SAR range from 0 dB to 12 dB, the tightnesses of the two approximations do not have much impact on their performances as the effect of the imperfect CSI is still easy to handle and hence the performances of the two proposed approaches are almost identical. However, as the IR target SINRs increase, the impact of the imperfect CSI is more severe and harder to cope with.
In such situation, Method II outperforms Method I as the approximation employed in the former is tighter than that in the latter. This is due to the fact that a bounded-norm model has been implicitly imposed on the uncertainty set of the CSI for the derivation of Method I but not for Method II. The same performance trend for the two types of approximations has also been reported in [40] for conventional information transmission in MISO downlink scenarios.

Fig. 2 reveals that the baseline scheme consumes less power, e.g., around 0.8 dB and 0.5 dB less at an IR target SINR of 10 dB, than the proposed approach for 6 and 8 antennas, respectively. The price paid for this lower power consumption is that the leakage SINR cannot be controlled, i.e., secrecy cannot be guaranteed. This will be shown and discussed more in detail in Fig. 3. At IR target SINRs less than 6 dB, the performances of the proposed methods are close to that of the baseline. However, as the IR target SINR increases, the performance gap between the proposed methods and the baseline widens. The reason for this is that the leakage SINR of the IRs increases with the IRs’ SINR requirements. Therefore, a higher power consumption is required by the proposed approaches as more artificial noise has to be generated to impact the ERs, hence, guaranteeing the required level of security for all IRs.

Since (6) and its equivalent form are non-convex optimization problems, finding the globally optimal solution by exhaustive search is computational infeasible given the large numbers of optimization variables. Therefore, a benchmark scheme is considered in Fig. 2. In particular, the benchmark scheme is based on [22] where perfect CSI is available for resource allocation. Thus, the performance of the benchmark scheme serves as a performance upper bound for the proposed Methods I and II. It can be observed from Fig. 2 that in the IR target SINR range from 0 dB to 20 dB, the performance gap between Method II and the benchmark scheme is less than 1 dB while the gap between Method I and the benchmark scheme is around 1 dB. This underlines the accuracy of our approximations. We note that the baseline scheme consumes slightly less transmit power than the benchmark scheme does. This benefit comes at the expense of no guaranteed communication secrecy.

Fig. 3 shows the histograms of the average leakage SINR at the ERs. To obtain the results in these figures, for each feasible channel realization, we first generated beamforming vectors for all methods. We then generated 1000 random error vectors for each channel realization to test the performance of each method. The resulting leakage SINRs, SINR levels, and ER power levels were averaged over the ERs and the feasible channel realizations. The results in Fig. 3 indicate that the information delivered to the IRs with the baseline scheme is at a very high risk of being decoded by the ERs as the leakage SINR is more than 10 dB higher than the ER decoding sensitivity level for all cases. The conservatism of Method I can be observed as it pushes the leakage SINR around 0.7 dB below the required value, i.e., 0.2 dB lower than Method II.

In Fig. 4 we compare our proposed methods against the norm-bounded scheme in [12]. As mentioned earlier, the two-stage optimization scheme in [12] can only ensure the individual secrecy rates of all IRs. In this comparison, we have given more privilege to the scheme by considering only one of those stages, i.e., Problem (36), where we have directly set the ER decoding sensitivity level to $\gamma_i^{(f)} = -5$ dB and varied the required IR target SINR $\gamma_i$ from $-10$ dB to $10$ dB. From the figure, it is clear that the proposed methods outperform the scheme in [12] in terms of power efficiency. For example, the optimization problem in [12] yields infeasible solutions for all channel realizations if the required IR SINRs exceed $-6$ dB while the proposed schemes can still operate.

Fig. 5 shows the average total transmit power versus the required power level at each ER for different required IR SINR targets.

We refer to a channel condition as feasible if all considered methods return feasible solutions.
widens to around 0.8 dB as the ER’s power demand increases to 10 dBm. This again confirms the improved tightness of the approximation employed in Method II compared to that employed in Method I for relatively high ER power demands. For the considered scenario, the required total transmit power is mainly determined by the required powers at the ERs which have a more significant effect on the total transmit power than the required IR SINR, cf. Figs. 2 and 5.

As can be observed, a slight increase in the transmit power level significantly improves the QoS, i.e., reduces the outage probability. For instance, for $M = 6$, an increase of 0.57 dB in the transmit power of Method I and an increase of 0.47 dB in that of Method II significantly reduce the outage probability from 30 % to 1 %.

Finally, from Figs. 4 and 5, it can be concluded that increasing the number of antennas reduces the power consumption of the considered schemes. This is a result of the improved beamformer resolution due to the extra spatial degrees of freedom introduced by additional antennas.

**VI. Conclusions and Future Work**

We have proposed a chance-constrained optimization problem to tackle the imperfection of the instantaneous CSI for the design of a power-efficient and secure SWIPT system. To handle non-convex QoS constraints, we have derived two robust reformulations of the proposed problem, i.e., Method I and Method II, adopting safe approximation techniques. Our analysis has revealed that Method I has a lower computational complexity than Method II. On the other hand, simulation results indicate that when the impact of imperfect CSI is more difficult to mitigate, i.e., at relatively high IR target SINRs, Method II outperforms Method I as it employs a tighter approximation. Both methods closely follow the performance of a benchmark scheme, where perfect CSI is available for resource allocation, while they outperform a baseline scheme in guaranteeing secure data transmissions. The results also show that a small increase in the transmit power leads to a significant reduction of the outage probability which in turn improves the QoS of the communication system.

A possible extension of this work is to consider a MIMO scenario where IRs and ERs are equipped with multiple-antennas. In such scenario, the expressions of the SINR, leakage SINR, and total power received at each ER will be different from those in this paper. Besides, the estimation errors will be characterized by matrices. As for Method I, it is possible to extend our robust beamforming design to the case of MIMO SWIPT involving error matrices. In particular, we can follow [41] Lemma 1 and apply the S-procedure. However, as for the generalization of the proposed algorithms for Method II, the extension to the case of MIMO is a non-trivial task which requires a new problem-solving methodology and more investigation.

**Acknowledgment**

Tuan Anh Le, Quoc-Tuan Vien and Huan Xuan Nguyen are funded by the Australian Research Council’s Discovery Early Career Researcher Award funding scheme (DE170100137). Robert Schober is supported by the Alexander von Humboldt Professorship Program.

**Appendix A**

**Proof of Theorem 1**

Now, we investigate the structure of the optimal beamforming matrix $W_i$. To this end, we rewrite [15] as:

$$
\min_{\{W_i, V_i\} \in \mathbb{M}^{M \times d_i \times M \times d_i}} \sum_{i=1}^{L} \mathbf{V}_i + \sum_{i=1}^{d_i} \mathbf{W}_i
$$

subject to

$$
\mathbf{D}_i(\alpha_i) + \mathbf{E}_i^H \mathbf{A}_i \mathbf{E}_i \succeq 0, \forall i
$$

$$
\mathbf{F}_i^{(i)}(\lambda_i^{(i)}) + \mathbf{L}_i^{(i)} \mathbf{B}_i \mathbf{L}_i \succeq 0, \forall i, \forall t
$$

$$
\mathbf{K}_i(\beta_i) + \mathbf{L}_i^{(i)} \mathbf{C}_i \mathbf{L}_i \succeq 0, \forall t
$$

$$
\alpha_i \succeq 0, \forall i, \lambda_i^{(i)} \succeq 0, \forall i, \forall t, \beta_i \succeq 0, \forall t,
$$

$$
\mathbf{W}_i \succeq 0, \forall i, \mathbf{V}_i \succeq 0, \forall t,
$$

where

$$
\mathbf{D}_i(\alpha_i) = \begin{bmatrix}
\alpha_i \mathbf{I}_M & \mathbf{0}_{M \times 1} \\
-\sigma_i^2 \mathbf{1}_d & -\alpha_i (1-\rho)
\end{bmatrix}
$$

$$
\mathbf{E}_i = \begin{bmatrix}
\mathbf{H}_i^{1/2} & \mathbf{0}_d
\end{bmatrix}
$$

$$
\mathbf{F}_i^{(i)}(\lambda_i^{(i)}) = \begin{bmatrix}
\lambda_i^{(i)} \mathbf{I}_M & \mathbf{0}_{M \times 1} \\
-\sigma_i^2 \mathbf{1}_d & -\lambda_i^{(i)} \mathbf{1}_d
\end{bmatrix}
$$

$$
\mathbf{L}_i = \begin{bmatrix}
\mathbf{G}_i^{1/2} & \mathbf{g}_i
\end{bmatrix}
$$

$$
\mathbf{K}_i(\beta_i) = \begin{bmatrix}
\beta_i \mathbf{I}_M & \mathbf{0}_{M \times 1} \\
-\mathbf{P}_i & -\beta_i \mathbf{1}_d
\end{bmatrix}
$$
The Lagrangian of (34) can be expressed as

\[
\mathcal{L}(\{W_i\}, \{V_i\}, \Upsilon) = \sum_{i=1}^{U} \text{Tr}(W_i) + \sum_{i=1}^{N} \text{Tr}(V_i) - \sum_{i=1}^{U} \text{Tr}(Q_i [D_i(\alpha_i) + E_i^H A_i E_i]) - \sum_{i=1}^{U} \kappa_i \alpha_i \\
- \sum_{i=1}^{U} \sum_{i=1}^{N} \text{Tr}(R_i^{(i)} [F_i^{(i)}(\lambda_{i}^{(i)}) + L_i^H B_i L_i]) \\
- \sum_{i=1}^{U} \sum_{i=1}^{N} \mu_i \lambda_{i}^{(i)} - \sum_{i=1}^{U} \text{Tr}(S_i [K_i(\beta_i) + L_i^H C_i L_i]) \\
- \sum_{i=1}^{U} \psi_i \beta_i - \sum_{i=1}^{U} \text{Tr}(T_i W_i) - \sum_{i=1}^{N} \text{Tr}(Z_i V_i),
\]

(35)

where \(Q_i, \kappa_i, R_i^{(i)}, \mu_i, S_i, \psi_i, T_i, \) and \(Z_i\) are the Lagrange multipliers associated with the constraints in (34). Furthermore, we have \(\alpha = [\alpha_1, \ldots, \alpha_U]^T, \mu = [\mu_1, \ldots, \mu_U, \ldots, \mu_1, \ldots, \mu_U]^T, \kappa = [\kappa_1, \ldots, \kappa_U], \psi = [\psi_1, \ldots, \psi_N]^T, \) \(\{Q_i\}, \ldots, \{U_i\}, \{R_i^{(i)}\} = [R_1^{(1)}, \ldots, R_1^{(N)}, \ldots, R_U^{(N)}], \{S_i\} = [S_1, \ldots, S_N], \) \(\{T_i\} = [T_1, \ldots, T_U], \{Z_i\} = [Z_1, \ldots, Z_N], \) and finally \(\Upsilon = [\alpha, \kappa, \mu, \psi, \{Q_i\}, \{R_i^{(i)}\}, \{S_i\}, \{T_i\}, \{Z_i\}]\).

Let \(g(\Upsilon)\) be the dual function of (35) given by

\[
g(\Upsilon) = \min_{\{W_i\},\{V_i\}\in\mathbb{M}^{M\times M}} \mathcal{L}(\{W_i\}, \{V_i\}, \Upsilon).
\]

(36)

Therefore, the corresponding dual problem of (34) can be stated as:

\[
\max_{\Upsilon} g(\Upsilon),
\]

s. t.

\[
\alpha > 0, \kappa > 0, \mu > 0, \psi > 0,
\]

\[
Q_i \succeq 0, T_i \succeq 0, S_i \succeq 0, Z_i \succeq 0, R_i^{(i)} \succeq 0, \forall i, \forall t.
\]

(37)

Since problem (34) is convex and satisfies Slater’s constraint qualification (27), the duality gap is zero and the optimal solution of (34) can be obtained by solving (37).

Let \(\Upsilon^* = [\alpha^*, \kappa^*, \mu^*, \psi^*, \{Q_i^*\}, \{R_i^{(i)*}\}, \{S_i^*\}, \{T_i^*\}, \{Z_i^*\}]\) be the optimal solution to dual problem (37), then the corresponding optimal solution \((W_i^*, \{V_i^*\})\) to problem (34) can be obtained as

\[
g(\Upsilon^*) = \min_{\{W_i\},\{V_i\}\in\mathbb{M}^{M\times M}} \mathcal{L}(\{W_i\}, \{V_i\}, \Upsilon^*).
\]

(38)

Substituting for \(A_i, B_i, \) and \(C\) in (35), after some mathematical manipulations, one can arrive at

\[
\mathcal{L}(\{W_i\}, \{V_i\}, \Upsilon^*) = \sum_{i=1}^{U} \text{Tr}(\Theta_i W_i) + \sum_{i=1}^{N} \text{Tr}(\Xi_i V_i) + \chi
\]

where \(\Theta_i = I_M - \frac{1}{\gamma_i} E_i Q_i^* E_i^H + \sum_{i=1}^{N} \frac{1}{\gamma_i} L_i R_i^{(i)*} L_i^H - \sum_{i=1}^{N} L_i S_i^* L_i^H - T_i^*, \quad \Xi_i = I_M + \sum_{i=1}^{U} E_i^H E_i + \sum_{i=1}^{N} L_i S_i^* L_i^H - \sum_{i=1}^{U} L_i^H R_i^{(i)*} L_i^H - \sum_{i=1}^{N} L_i S_i^* L_i^H - T_i^*, \quad \chi = \sum_{i=1}^{U} \text{Tr}(Q_i^* [D_i(\alpha_i^*) + E_i^H A_i E_i]) - \sum_{i=1}^{U} \kappa_i \alpha_i^* \quad \text{and}
\]

\[
\sum_{i=1}^{U} \sum_{i=1}^{N} \text{Tr}(R_i^{(i)*} L_i^H) - \sum_{i=1}^{N} \text{Tr}(S_i [K_i(\beta_i) + L_i^H C_i L_i]) - \sum_{i=1}^{U} \sum_{i=1}^{N} \mu_i \lambda_{i}^{(i)*} - \sum_{i=1}^{U} \sum_{i=1}^{N} \psi_i \beta_i^*.
\]

(40)

Therefore, we can rewrite (38) as

\[
g(\Upsilon^*) = \min_{\{W_i\},\{V_i\}\in\mathbb{M}^{M\times M}} \sum_{i=1}^{U} \text{Tr}(\Theta_i W_i) + \min_{\{V_i\}\in\mathbb{M}^{N\times N}} \sum_{i=1}^{N} \text{Tr}(\Xi_i V_i).
\]

(41)

Since (15) is feasible, the optimal value of its equivalent form, i.e., (34), is non-negative. Furthermore, the optimal duality gap between primary problem (34) and its Lagrange dual problem (37) is zero. Therefore, \(\Theta_i\) and \(\Xi_i\) must be positive semi-definite, i.e., \(\Theta_i \succeq 0, \forall i, \Xi_i \succeq 0, \forall t, \) to ensure that the Lagrangian dual function is bounded from below, i.e., the Lagrangian dual function cannot become \(-\infty\). We continue by introducing the following proposition:

**Proposition 1:** If an \(M\times M\) Hermitian matrix \(W_i\) has a rank of \(K \leq M\), then it can be expressed as \(W_i = \sum_{k=1}^{K} v_i k a_i k a_i k^H\), where \(v_i k\) and \(a_i k\) are the \(k\)th non-zero eigenvalue and the corresponding eigenvector of \(W_i\), respectively.

In the following, we prove the rank-one property of the solution of (40) by contradiction. Assuming that the optimal solution of (40), \(W_i^*\), has rank \(K > 1, \forall i, \) Proposition 1 indicates that \(W_i^* = \sum_{k=1}^{K} v_i k a_i k a_i k^H\). Now we construct another feasible solution to (41) as

\[
\tilde{W}_i = v_{i_{\rho}} a_{i_{\rho}} a_{i_{\rho}}^H, \quad \forall i, \forall \rho
\]

where \(p = \arg \min_{k \in \{1, \ldots, K\}} \langle a_i k, \Theta_i a_i k \rangle.\) Combining \(\Theta_i \succeq 0\) and (42) reveals that

\[
\sum_{i=1}^{U} \text{Tr}(\Theta_i \tilde{W}_i) < \sum_{i=1}^{U} \text{Tr}(\Theta_i W_i^*).
\]

(43)

The inequality in (43) contradicts the optimality of \(W_i^*\). Therefore, \(W_i^*\) must be a rank-one matrix for all \(i.\) Following similar arguments, one can show that \(V_i^*\) must be a rank-one matrix for all \(i.\)

**Appendix B**

**Proof of Theorem 2**

This proof is based on a fact that the SOC constraints in (21), (23), and (27) can be cast as LMIIs. With some abuse of notation, in this proof we reuse \(B, C, D, E, T, P, K, \) and \(u_i\) from Section III and Appendix A. In the sequel, we first show the transformation of (20). Exploiting the Schur complement with some mathematical manipulations, (21) can be equivalently written as \(B_i \succeq 0\) where

\[
B_i = \begin{bmatrix}
\theta I_{M^2 + M} - \frac{1}{\gamma_i} E_i^H E_i + \frac{1}{\gamma_i} L_i R_i^{(i)*} L_i^H - \frac{1}{\gamma_i} L_i^H R_i^{(i)*} L_i^H - \sum_{i=1}^{N} L_i^H R_i^{(i)*} L_i^H - \sum_{i=1}^{N} L_i^H R_i^{(i)*} L_i^H - T_i^* & \frac{\sqrt{2} H_i^H A_i^H}{{vec}(H_i^H A_i^H)} \\
\frac{\sqrt{2} H_i^H A_i^H}{{vec}(H_i^H A_i^H)} & \theta_i
\end{bmatrix}.
\]

(44)

11Proposition 1 can be proved by using the facts that the Hermitian matrix \(W_i\) has \(K\) real non-zero eigenvalues and \(K\) orthogonal eigenvectors.
We further decompose $B_i$ as
\[
B_i = C(\theta_i) + D_i + D_i^H + E_i + E_i^H
\]
(45)
where
\[
C(\theta_i) = \begin{bmatrix}
\theta_i I_{M^2+M} & 0_{(M^2+M)\times 1} \\
0_{1\times (M^2+M)} & \theta_i
\end{bmatrix},
\]
\[
D_i = \begin{bmatrix}
0_{(M^2+M)\times (M^2+M)} & 0_{(M^2+M)\times 1} \\
0_{M\times 1} & 0_{M^2+M}
\end{bmatrix},
\]
\[
E_i = \begin{bmatrix}
0_{M^2\times 1} \\
\operatorname{vec}(H_i^{1/2}A_iH_i^{1/2})
\end{bmatrix},
\]
(45)
Moreover, matrix $D_i$ can be further decomposed as
\[
D_i = \sqrt{\theta_i} \begin{bmatrix}
I_{M^2+M} & 0_{M\times 1} \\
0_{1\times (M^2+M)} & 0_{M\times 1}
\end{bmatrix},
\]
(46)
where $T_i(\bar{h}_i) = \sqrt{\theta_i} I_{M^2+M}$, $P_i(H_i^{1/2}) = H_i^{1/2} I_{M^2+M}$, and $[T_i(\bar{h}_i) A_i P_i(H_i^{1/2})]$ is an $M \times (M^2+M+1)$ matrix. Let $U_i = [0_{M\times 1} \cdots 0_{M\times 1} \mathbf{u}_i]^T$ be an $(M^2+M+1) \times M$ matrix where $\mathbf{u}_i$ is an $M \times 1$ vector with 1 at the $i$-th entry and all zeros elsewhere. Furthermore, let $K_i = \begin{bmatrix}0_{M\times M} & 0_{M\times M} & \cdots & 0_{M\times M} \\
1_{1\times (M^2+M)} & 1_{1\times (M^2+M)} & \cdots & 1_{1\times (M^2+M)}
\end{bmatrix}$
be an $M \times (M^2+M+1)$ matrix with $\mathbf{0}_{M\times 1}$ as the $i$-th block and all zeros elsewhere. One can express $E_i$ as
\[
E_i = \sum_{p=1}^{M} U_i H_i^{1/2} A_i H_i^{1/2} K_p.
\]
(47)
From (44), (45), (46), and (47), we conclude that constraint (21) is equivalent to the following LMI constraint
\[
\begin{aligned}
C(\theta) + T_i(\bar{h}_i) A_i P_i(H_i^{1/2}) + & \left[ T_i(\bar{h}_i) A_i P_i(H_i^{1/2}) \right]^H \\
+ & \sum_{p=1}^{M} U_i H_i^{1/2} A_i H_i^{1/2} K_p + \left[ \sum_{p=1}^{M} U_i H_i^{1/2} A_i H_i^{1/2} K_p \right]^H \geq 0.
\end{aligned}
\]
(48)
Similarly, constraint (23) can be equivalently cast in LMI form as
\[
\begin{aligned}
C(\theta^{(t)}) + T_i(\bar{g}_i) B_i P_i(G_i^{1/2}) + & \left[ T_i(\bar{g}_i) B_i P_i(G_i^{1/2}) \right]^H \\
+ & \sum_{p=1}^{M} U_p G_i^{1/2} B_i G_i^{1/2} K_p + \left[ \sum_{p=1}^{M} U_p G_i^{1/2} B_i G_i^{1/2} K_p \right]^H \geq 0.
\end{aligned}
\]
(49)
Finally, constraint (27) can be equivalently written as the following LMI
\[
\begin{aligned}
C(a) + T_i(\bar{g}_i) C P_i(G_i^{1/2}) + & \left[ T_i(\bar{g}_i) C P_i(G_i^{1/2}) \right]^H \\
+ & \sum_{p=1}^{M} U_p G_i^{1/2} C G_i^{1/2} K_p + \left[ \sum_{p=1}^{M} U_p G_i^{1/2} C G_i^{1/2} K_p \right]^H \geq 0.
\end{aligned}
\]
(50)
Using (48), (50), one can rewrite (30) as
\[
\begin{aligned}
\min_{\{W_i(V_i)\}_i \in \mathbb{R}^{M \times N}} & \operatorname{Tr} \left( \sum_{i=1}^{U} W_i + \sum_{i=1}^{N} V_i \right) \\
s.t. & \begin{cases}
\{W_i(V_i)\}_i \in \mathbb{R}^{M \times N}, & \forall i, \forall t; \\
\sum_{i=1}^{U} W_i \geq 0, & \forall i, \forall t, \\
\sum_{i=1}^{N} V_i \geq 0, & \forall i, \forall t,
\end{cases}
\end{aligned}
\]
(51)
The same technique as presented in Appendix A can now be adopted to prove that the optimal beamforming matrices $\{W_i\}$ and $\{V_i\}$ in (51) are always rank-one if the problem is feasible. In the following, we sketch some of the main steps.

Let $\overline{\mathcal{U}}(\{W_i\}, \{V_i\}, \overline{T})$ be the Lagrangian of (51) where $\overline{T}$ represents the collection of all Lagrangian multipliers associated with the constraints in (51). Furthermore, let $\overline{\mathcal{U}}(\overline{T})$ be the dual function of (51), i.e., $\overline{\mathcal{U}}(\overline{T}) = \min_{\{W_i(V_i)\}_i \in \mathbb{R}^{M \times N}} \overline{\mathcal{U}}(\{W_i(V_i)\}_i, \overline{T})$.

As problem (51) is convex and satisfies Slater’s constraint qualification [27], the duality gap is zero and the optimal solution of (51) can be obtained by solving its dual problem. Let $\overline{T}^*$ be the optimal solution of the dual problem of (51). Then, the corresponding optimal solution $(\{W_i^*, V_i^*\})$ to problem (51) can be obtained as
\[
\overline{\mathcal{U}}(\overline{T}^*) = \min_{\{W_i(V_i)\}_i \in \mathbb{R}^{M \times N}} \overline{\mathcal{U}}(\{W_i(V_i)\}_i, T^*).
\]
(52)
After some mathematical manipulations, one can rewrite (52) in the following form
\[
\overline{\mathcal{U}}(\overline{T}^*) = \min_{\{W_i\}_i \in \mathbb{R}^{M \times U}} \sum_{i=1}^{U} \operatorname{Tr} \left( \overline{\Theta}_i W_i \right) + \min_{\{V_i\}_i \in \mathbb{R}^{M \times N}} \sum_{i=1}^{N} \operatorname{Tr} \left( \overline{\Xi}_i V_i \right),
\]
(53)
where $\overline{\Theta}_i$ and $\overline{\Xi}_i$ are functions of $(\theta_i, T_i(\bar{h}_i), P_i(H_i^{1/2}), U_p, K_p, C(\theta^{(t)}), T_i(\bar{g}_i), C(a_i), \text{ and } \overline{T})$. By using a similar contradiction approach as in the proof of Theorem 1 one can show that the optimal solution, $(\{W_i^*, V_i^*\})$, to problem (53), which is also the optimal solution to (51), can only involve rank-one matrices.

References

[1] A. Ghasanfar, H. Tabassum, and E. Hossain, “Ambient RF energy harvesting in ultra-dense small cell networks performance and trade-offs,” IEEE Wireless Commun. Mag., vol. 23, no. 2, pp. 38–45, Apr. 2016.
[2] X. Chen, D. W. K. Ng, and H.-H. Chen, “Secrecy wireless information and power transfer: Challenges and opportunities,” IEEE Wireless Commun. Mag., vol. 23, no. 2, pp. 54–61, Apr. 2016.
[3] Q. Liu, K. S. Yildirim, P. Paweleczak, and M. Warner, “Safe and secure wireless power transfer networks: Challenges and opportunities in RF-based systems,” IEEE Commun. Mag., vol. 54, no. 9, pp. 74–79, Sep. 2016.
[4] D. W. K. Ng, E. S. Lo, and R. Schober, “Robust beamforming for secure communication in systems with wireless information and power transfer,” IEEE Trans. Wireless Commun., vol. 13, no. 8, pp. 4599–4615, Aug. 2014.
[5] D. W. K. Ng and R. Schober, “Secure and green SWIPT in distributed antenna networks with limited backhaul capacity,” IEEE Trans. Wireless Commun., vol. 14, no. 9, pp. 5082–5097, Sep. 2015.
