# Probabilistically Robust SWIPT for Secrecy MISOME Systems

Muhammad R. A. Khandaker, *Member, IEEE*, Kai-Kit Wong, *Fellow, IEEE*,
Yangyang Zhang, and Zhongbin Zheng

*Abstract*— This paper considers simultaneous wireless information and power transfer in a multiple-input single-output downlink system consisting of one multi-antenna transmitter, one single-antenna information receiver, multiple multi-antenna eavesdroppers (Eves), and multiple single-antenna energy-harvesting receivers (ERs). The main objective is to keep the probability of the legitimate user's achievable secrecy rate outage as well as the ERs' harvested energy outage caused by channel state information uncertainties below some prescribed thresholds. As is well known, the secrecy rate outage constraints present a significant analytical and computational challenge. Incorporating the energy harvesting outage constraints only intensifies that challenge. In this paper, we address this challenging issue using convex restriction approaches which are then proved to yield rank-one optimal beamforming solutions. Numerical results reveal the effectiveness of the proposed schemes.

*Index Terms*— Secrecy, outage, energy harvesting, SWIPT, probabilistic, robust secrecy.

## I. Introduction

RESEARCHERS have long been investigating conventional ambient energy resources, such as solar and wind, for energising low-power electronic devices. However, the sporadic and unpredictable nature of these ambient sources makes energy harvesting critical for applications where quality-of-service (QoS) is of priority, and most of these conventional harvesting technologies are only applicable in certain environments and/or weather conditions [1]. On the other hand, wireless power transfer via magnetic induction is not yet a viable means for widespread applications due to its extremely short distance of power transfer.

Recently, simultaneous wireless information and power transfer (SWIPT) has received enormous interest triggered

by the prospects of powering energy-limited wireless devices via RF energy harvesting (EH) due to reductions in power requirements of electronics [2]–[5]. As RF signals transport information and energy sumultaneously, mobile users are actually blessed with access to both energy and data at the same time through SWIPT. Surprisingly, this fact did not attract much attention until as late as the last decade.

A practical challenge for SWIPT is that information and energy receiver (ER) circuits operate at very different power sensitivity level (e.g., $-10$dBm for ERs versus $-60$dBm for information receivers (IRs)). To store a useful amount of energy from harvesting, the ERs need access to a higher receive power. Due to this incompatibility between the two forms of receivers, two practical schemes, namely, time switching (TS) and power splitting (PS) have been proposed in the literature in order to enable SWIPT [2], [3], [6]. In particular, the authors of [2] investigated both schemes in a multiple-input multiple-output (MIMO) broadcasting scenario from a base station (BS) to two mobile receivers decoding information and harvesting energy at the same time. Although the TS scheme simplifies the receiver design, it compromises the efficiencies of the SWIPT technology since the receivers decode information and harvest energy in alternating time slots. Thus, in [3], only the PS-based receiver architecture has been rigorously studied for SWIPT in a point-to-point system. The PS-based schemes have also been considered for multiple-input single-output (MISO) SWIPT multicasting systems with perfect as well as imperfect channel state information (CSI) in [6] and for MIMO multicasting systems in [7].

Nevertheless, due to channel fading, wireless power transfer efficiency decays drastically as the transmitter-receiver distance increases. Since multiple-antenna techniques provide additional degrees of freedom (DoF) exploiting spatial diversity, installing multiple antennas can help combat channel fading in order to improve wireless power transfer efficiency [2].

In order to further improve EH efficiency, a receiver-location based *near-far* scheduling scheme has also been proposed in the SWIPT literature for information and energy transmissions [2], [8], in which the receivers only in closer vicinity to the transmitter are scheduled for harvesting energy. Although the scheme apparently seems to be beneficial for EH, actually it gives rise to an undesired security vulnerability for transmitting secret information in scenarios where ERs are supposed to be kept in the dark about the secret message. Specifically, ERs in the scheme have better fading channels than IRs and thus have higher probability to successfully decode the information sent to the IRs [9], [10]. On the other

hand, the transmitter often needs to apply higher power in order to satisfy ERs' EH requirements which they normally do not need for information-only transmission. This makes secret messages susceptible to external eavesdropping attack as well. Information secrecy can be further degraded if the ERs start cooperating in order to perform joint decoding in an attempt to improve their interception. Thus, the SWIPT systems need to be carefully designed in order to be able to successfully transfer secret messages keeping the potential eavesdroppers ignorant of the secret message to the IR.

To make sure that the message is delivered secretly to the IR in a SWIPT system taking possible eavesdropping by the ERs into consideration, MISO secrecy beamforming schemes were proposed in [9]–[11]. It was assumed in [9] that the ERs do not collude to perform joint decoding and that the CSI of all the nodes, including the eavesdroppers, was perfectly known at the transmitter. Nonetheless, in order to guarantee maximum information secrecy, it would be more meaningful to consider the worst-case scenario, where the ERs collude together to attempt to decode the data jointly such that the eavesdropping rate is maximized. Also, obtaining the eavesdroppers' CSI perfectly is practically impossible. Hence the authors in [10] and [11] considered robust design based on deterministic channel uncertainty models for SWIPT in scenarios where the ERs may collude together to perform joint decoding in an attempt to improve their interception. In [12] and [13], physical layer security was studied for SWIPT in OFDMA networks.

Other works with secrecy in SWIPT either considered worst-case robust approaches in which the CSI errors are assumed to be within a bounded set, or correlation-based approaches in which the channel statistics is available. A delay-limited secrecy SWIPT system was considered in [14] with single-antenna nodes, while a randomization-guided rank-one suboptimal solution was proposed in [15] for worst-case MISO secrecy SWIPT systems, and in [16] for MIMO SWIPT systems. Worst-case based MISO secrecy SWIPT optimization has also been considered in [17] for norm-bounded channel uncertainty model based on successive convex approximation approach.

Unfortunately, due to inaccurate channel estimation methods, it may not always be possible that the legitimate transmitter obtains these deterministic models perfectly [18]. In such cases, secrecy as well as EH outage may occur. Hence, this paper considers robust secrecy optimization problems with probabilistic secrecy rate and EH constraints for MISO systems with multiple multi-antenna eavesdroppers (MISOME). Two probabilistic secrecy beamforming design problems have been considered namely (1) minimizing the transmit power subject to probabilistic QoS constraints, and (2) maximizing the secrecy rate subject to the total transmit power and secrecy and EH outage constraints. In contrast to the deterministic (worst- or average-case) models [9]–[11], [14]–[17], this approach offers a safe performance, guaranteeing a certain chance of successful QoS deliveries.

Our main objective is to maintain the probability of the legitimate user's achievable secrecy rate outage as well as the ERs' harvested energy outage caused by CSI

uncertainties below given thresholds. As is already well known, the secrecy rate outage constraints present a significant analytical and computational challenge. Inclusion of the EH outage constraints only proliferates that challenge. Unfortunately, quadratic chance constraints generally do not have closed-form expressions, and are unlikely to be tractable. Therefore, it is quite common in the robust optimization literature to develop safe tractable approximations of the outage-based QoS constraints that are computationally efficient and are good in accuracies. Motivated by this, in this paper, we present three conservative approximation approaches namely, Bernstein-type inequality (BTI), $\mathcal{S}$-procedure, and large deviation inequality (LDI) [19]–[22] based approaches in order to transform the probabilistic constraints into safe and tractable ones.

Applying semidefinite relaxation (SDR) technique, we show that the safe approximation results always yield rank-one optimal transmit covariance solution for the IR. Our simulation results also reveal that the BTI-based restriction approach is the best and the conventional $\mathcal{S}$-procedure based approach is comparatively the worst strategy in terms of secrecy rates, whereas the LDI-based approach appears to be a good compromise. Note that [23] considered secure robust transmit beamforming design for minimizing the total transmit power under QoS constraints including SINR and SINR outage constraints at the eavesdroppers. The nonconvex probabilistic constraint was replaced with a tractable convex deterministic constraint by bounding the radius of the uncertainty region (comparable to our $\mathcal{S}$-procedure based approach). However, [23] did not explicitly consider the secrecy rate or EH outage constraints, which are much more challenging to deal with but of great importance to SWIPT systems.

The rest of this paper is organized as follows. In Section II, the system model of a secret MISO SWIPT network is introduced. The secrecy rate constrained (SRC) power minimization problem is discussed in Section III for the imperfect CSI case whereas in Section IV, solutions to the secrecy rate maximization (SRM) problem are derived. Section V presents the simulation results that justify the significance of the proposed algorithms under various scenarios. Concluding remarks are provided in Section VI.

*Notations:* Throughout this paper, we use the following notations. Boldface lowercase and uppercase letters are used to represent vectors and matrices, respectively. The symbol $\mathbf{I}_n$ denotes an $n \times n$ identity matrix, $\mathbf{0}$ is a zero vector or matrix. Also, $\mathbf{A}^H$, $\mathrm{tr}(\mathbf{A})$, $\mathrm{rank}(\mathbf{A})$, and $|\mathbf{A}|$ represent the Hermitian (conjugate) transpose, trace, rank and determinant of a matrix $\mathbf{A}$; $\mathrm{Pr}[\cdot]$ represents the probability of an event; $\|\cdot\|$ and $\|\cdot\|_F$ represent the Euclidean norm and Frobenius norm, respectively; $\mathbf{A} \succeq \mathbf{0}\,(\mathbf{A} \succ \mathbf{0})$ means that $\mathbf{A}$ is a Hermitian positive semidefinite (definite) matrix. The notation $\mathbf{x} \sim \mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ means that $\mathbf{x}$ is a random vector following a complex circularly symmetric Gaussian distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$.

## II. SYSTEM MODEL

A MISO downlink system is considered for SWIPT with $K + L + 1$ receivers as shown in Fig. 1. The BS transmits
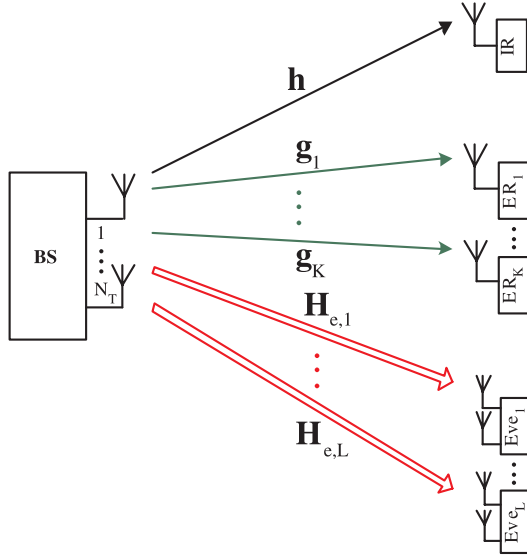
Fig. 1. A MISO SWIPT system with multiple-antenna eavesdroppers.

information to a legitimate user and energy to $K$ EH receivers keeping the information as secure as possible from $L$ eavesdroppers (Eves). It is assumed that the BS performs necessary user selection operation before transmission begins for decoding information. Note that similar assumption has also been made in numerous existing works in the secrecy literature for SWIPT, e.g., [9]–[11], [14]–[17]. The transmitter or BS has $N_T > 1$ transmitting antennas and each legitimate receiver (IR or ER) has single receiving antenna. Eves are equipped with $N_{e,i}$ antennas for $i = 1, \ldots, L$. The BS performs linear transmit beamforming to send secret information to the IR. We assume that the ERs are also legitimate users of the network authorized for EH only and do not intend to overhear the message destined to the IR.[1] By letting $\mathbf{x}$ be the transmit signal vector, the received signals at the IR, the $k$th ER, and the $i$th Eve can be modeled, respectively, as

$$y_d = \mathbf{h}^H \mathbf{x} + n_d, \tag{1}$$

$$y_{h,k} = \mathbf{g}_k^H \mathbf{x} + n_{h,k}, \text{ for } k = 1, \ldots, K, \tag{2}$$

$$\mathbf{y}_{e,i} = \mathbf{H}_{e,i}^H \mathbf{x} + \mathbf{n}_{e,i}, \text{ for } i = 1, \ldots, L, \tag{3}$$

where $\mathbf{h}$, $\mathbf{g}_k$, and $\mathbf{H}_{e,i}$ are the conjugated complex channel vector (matrix) between the BS and the IR, the $k$th ER, and the $i$th Eve, respectively, $n_d \sim \mathcal{CN}(0, \sigma_d^2)$, $n_{h,k} \sim \mathcal{CN}(0, \sigma_h^2)$, and $\mathbf{n}_{e,l} \sim \mathcal{CN}(\mathbf{0}, \sigma_e^2 \mathbf{I}_{N_l})$ are the additive Gaussian noises at the IR, the $k$th ER, and the $i$th Eve, respectively. For notational simplicity, the path loss factors $P_{f,d} = \sqrt{L_c d_I^{-\bar{\kappa}}}$, $P_{f,hk} = \sqrt{L_c d_{h,k}^{-\bar{\kappa}}}$, and $P_{f,ei} = \sqrt{L_c d_{e,i}^{-\bar{\kappa}}}$, with the path loss constant $L_c = G_T G_R \left(\frac{c}{4\pi f}\right)^2$ (in which $G_T$ and $G_R$ are the antenna gains of the transmitter and the receivers, respectively, $c$ is the speed of light, and $f$ is the carrier frequency), $d_I$, $d_{h,k}$, and $d_{e,i}$ indicating the distance of the IR, $k$th ER, and $i$th Eve, respectively, from the BS, and $\bar{\kappa}$ (typically between 2.7 and 3.5) being the path loss exponent, are all assumed

[1]The results of this paper are readily extendible to study the impacts of untrusted ERs on the SWIPT system performance.

to be absorbed in the corresponding channel gains. The BS chooses $\mathbf{x}$ as $\mathbf{x} = \mathbf{b}_I s_I$ where $s_I \sim \mathcal{CN}(0, 1)$ is the confidential information-bearing signal for the IR and $\mathbf{b}_I$ is the transmit beamforming vector. Hence, $\mathbf{x}$ is a beamformed version of the message.

Note that multi-antenna eavesdroppers may interchangeably mean that single-antenna eavesdroppers located in a favourable location to cooperate may collude together to improve their interception. For ease of exposition, we further assume that all the Eves are colluding into multiple groups. In particular, Eves are assumed to perform joint maximum signal-to-noise ratio (SNR) receive beamforming. By denoting $\mathbf{Q}_I \triangleq \mathbf{b}_I \mathbf{b}_I^H$ as the transmit covariance matrix, the mutual information (MI) between the BS and the IR is given by

$$C_I(\mathbf{Q}_I) = \log\left(1 + \frac{1}{\sigma_d^2} \mathbf{h}^H \mathbf{Q}_I \mathbf{h}\right), \tag{4}$$

and that between the BS and the colluded Eves is given by

$$C_{e,i}(\mathbf{Q}_I) = \log\left|\mathbf{I}_{N_{e,i}} + \frac{1}{\sigma_e^2} \mathbf{H}_{e,i}^H \mathbf{Q}_I \mathbf{H}_{e,i}\right|, \text{ for } i = 1, \ldots, L. \tag{5}$$

Given $\mathbf{Q}_I$, the achievable secrecy rate is given by [24]

$$C_s = \min_i \left\{C_I(\mathbf{Q}_I) - C_{e,i}(\mathbf{Q}_I)\right\}^+, \tag{6}$$

where $\{a\}^+ = \max(0, a)$. The harvested power at the $k$th ER is given by

$$E_k = \xi_k \mathbf{g}_k^H \mathbf{Q}_I \mathbf{g}_k, \tag{7}$$

where $\xi_k \in (0, 1]$ is the energy conversion efficiency of the energy transducers at the $k$th ER. For simplicity, it is assumed that the harvested energy due to the background noise at the EH receivers in (7) is negligible and as a consequence can be ignored [2].

In most of the existing works with secrecy for SWIPT, it is assumed that the instantaneous CSI of all the receivers is available at the transmitter. However, in practical wireless communication systems, perfect CSI is likely not available and an important issue is how to robustify a secure transmit design in the presence of imperfect CSI. As a consequence, our next exertion is to develop convex optimization algorithms that satisfy given chance constraints exploiting only imperfect CSI knowledge.

### III. ROBUST POWER MINIMIZATION WITH PROBABILISTIC CONSTRAINTS

In this section, we develop probabilistically robust algorithm for the SRC problem with secrecy rate and EH outage constraints. The interest here is in active Eves cases, where the Eves themselves are also users of the network and the transmitter aims to provide different services to different types of users. That is, the Eves are legitimate users for utilities other than the particular information destined to the IR. For such active eavesdroppers, the CSI can be estimated from the eavesdroppers' transmission. Thus we assume that the BS has incomplete knowledge of the ERs' as well as

Eves' channels while the IR's channel is perfectly known. The perfect IR's CSI assumption is quite widely exercised in the existing literature since the legitimate IR's CSI may be obtained at very high precision through secure control channels [9], [10], [16].

We consider the commonly used Gaussian channel error model for the imperfect CSI. To model the imperfect CSI, we assume that the actual channels $\mathbf{g}_k$, for $k = 1, \ldots, K$, lie in the neighbourhood of the estimated channels $\hat{\mathbf{g}}_k$, for $k = 1, \ldots, K$, available at the transmitter. The channel error vectors are assumed to have circularly symmetric complex Gaussian (CSCG) distribution. Thus, the actual channels are modeled as

$$\mathbf{g}_k = \hat{\mathbf{g}}_k + \boldsymbol{\delta}_{\mathrm{g},k}, \ \text{for } k = 1, \ldots, K, \tag{8}$$

where $\hat{\mathbf{g}}_k \in \mathbb{C}^{N_{\mathrm{T}} \times 1}$ is the estimated CSI of the $k$th ER and $\boldsymbol{\delta}_{\mathrm{g},k} \in \mathbb{C}^{N_{\mathrm{T}} \times 1}$, for $k = 1, \ldots, K$, represent the channel uncertainties such that $\boldsymbol{\delta}_{\mathrm{g},k} \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{R}_{\mathrm{g},k}\right)$, in which $\mathbf{R}_{\mathrm{g},k}$ is a positive semidefinite matrix. Similarly, the Eves' channel uncertainty model can be represented by

$$\mathbf{H}_{\mathrm{e},i} = \hat{\mathbf{H}}_{\mathrm{e},i} + \boldsymbol{\Delta}_{\mathrm{H},i}, \ \text{for } i = 1, \ldots, L, \tag{9}$$

where $\hat{\mathbf{H}}_{\mathrm{e},i} \in \mathbb{C}^{N_{\mathrm{T}} \times N_{\mathrm{e},i}}$ is the estimated CSI of the $i$th Eve and $\boldsymbol{\Delta}_{\mathrm{H},i} \in \mathbb{C}^{N_{\mathrm{T}} \times N_{\mathrm{e},i}}$, for $i = 1, \ldots, L$, represent the channel uncertainties such that $\boldsymbol{\delta}_{\mathrm{H},i} \triangleq \mathrm{vec}\left(\boldsymbol{\Delta}_{\mathrm{H},i}\right) \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{R}_{\mathrm{H},i}\right)$, where $\mathbf{R}_{\mathrm{H},i}$ is a positive semidefinite matrix. Thus, the probabilistically robust SRC power minimization problem can be formulated as

$$\min_{\mathbf{Q}_{\mathrm{I}}} \ \mathrm{tr}\left(\mathbf{Q}_{\mathrm{I}}\right) \tag{10a}$$

$$\text{s.t.} \ \Pr\left[\min_i \left\{C_{\mathrm{I}}\left(\mathbf{Q}_{\mathrm{I}}\right) - \hat{C}_{\mathrm{e},i}\left(\mathbf{Q}_{\mathrm{I}}\right)\right\}^+ \geq R\right] \geq 1 - p, \ \forall i, \tag{10b}$$

$$\Pr\left[\min_k \hat{E}_k \geq \eta_k\right] \geq 1 - q, \quad \forall k, \tag{10c}$$

$$\mathbf{Q}_{\mathrm{I}} \succeq \mathbf{0}, \tag{10d}$$

where $\hat{C}_{\mathrm{e},i}\left(\mathbf{Q}_{\mathrm{I}}\right) = \log\left|\mathbf{I}_{N_{\mathrm{e},i}} + \frac{1}{\sigma_{\mathrm{e}}^2}\mathbf{H}_{\mathrm{e},i}^H\mathbf{Q}_{\mathrm{I}}\mathbf{H}_{\mathrm{e},i}\right|$ is the $i$th Eves' MI with $\mathbf{H}_{\mathrm{e},i} = \hat{\mathbf{H}}_{\mathrm{e},i} + \boldsymbol{\Delta}_{\mathrm{H},i}$, for $i = 1, \ldots, L, \boldsymbol{\delta}_{\mathrm{H},i} \triangleq \mathrm{vec}\left(\boldsymbol{\Delta}_{\mathrm{H},i}\right) \sim \mathcal{CN}\left(0, \mathbf{R}_{\mathrm{H},i}\right)$ and $\hat{E}_k = \xi_k \mathbf{g}_k^H \mathbf{Q}_{\mathrm{I}} \mathbf{g}_k$ is the average energy harvested with $\mathbf{g}_k = \hat{\mathbf{g}}_k + \boldsymbol{\delta}_{\mathrm{g},k}$, for $k = 1, \ldots, K, \boldsymbol{\delta}_{\mathrm{g},k} \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{R}_{\mathrm{g},k}\right)$. The problem formulation in (10) guarantees that the IR can successfully decode its message at least $(1 - p) \times 100\%$ of the time. Similarly, the ERs can harvest the minimum required amount of power at least $(1 - q) \times 100\%$ of the time. As the desired outage probability decreases, the size of the feasible sets described by (10b) and (10c) decreases. Hence, one might expect an increase in the required transmit power with decreasing outage probabilities.

Notice that the rank constraint on $\mathbf{Q}_{\mathrm{I}}$ has been relaxed in problem (10). An important issue that arises from the relaxation is the rank of the resulting solution. The removal of the rank constraint $\mathrm{rank}(\mathbf{Q}_{\mathrm{I}}) = 1$ means that the solution obtained through solving problem (10) may have rank higher than one. A common practice of overcoming this is to apply some rank approximation procedure (e.g., randomization) to the optimal $\mathbf{Q}_{\mathrm{I}}^*$ to find a feasible beamforming solution $\mathbf{b}_{\mathrm{I}}$.

However, in this paper, we aim to prove the tightness of the rank relaxation.

The problem is still nonconvex due to the probabilistic constraints involving log det functions. To make those constraints more tractable, we introduce the following lemma.

*Lemma 1 [25]: For any positive semidefinite matrix $\mathbf{A}$, the following inequality holds*

$$|\mathbf{I} + \mathbf{A}| \geq 1 + \mathrm{tr}(\mathbf{A}) \tag{11}$$

*and the equality in (11) holds if and only if $\mathrm{rank}(\mathbf{A}) \leq 1$.*

*proof:* Let $r_A = \mathrm{rank}(\mathbf{A})$. While the case of $r_A = 0$ is trivial, for $r_A \geq 1$, let $\lambda_1 \geq \lambda_2 \geq \cdots \lambda_{r_A} > 0$ denote the nonzero eigenvalues of $\mathbf{A}$. Accordingly, we have that

$$|\mathbf{I} + \mathbf{A}| = \prod_{i=1}^{r_A}(1 + \lambda_i) = 1 + \sum_{i=1}^{r_A} \lambda_i + \sum_{i \neq k} \lambda_i \lambda_k + \ldots$$

$$\geq 1 + \sum_{i=1}^{r_A} \lambda_i = 1 + \mathrm{Tr}(\mathbf{A}).$$

Clearly, the above equality holds if and only if $r_A = 1$. □

Now, by applying *Lemma 1*, the secrecy rate outage constraint (10b) can be relaxed as

$$\Pr\left[\log\left(1 + \frac{1}{\sigma_{\mathrm{d}}^2}\mathbf{h}^H\mathbf{Q}_{\mathrm{I}}\mathbf{h}\right) - \log\left(1 + \frac{1}{\sigma_{\mathrm{e}}^2}\mathrm{tr}\left(\mathbf{H}_{\mathrm{e},i}^H\mathbf{Q}_{\mathrm{I}} \times \mathbf{H}_{\mathrm{e},i}\right)\right) \geq R\right] \geq 1 - p, \quad \forall i,$$

which is equivalent to

$$\Pr\left[\left(1 + \frac{1}{\sigma_{\mathrm{d}}^2}\mathbf{h}^H\mathbf{Q}_{\mathrm{I}}\mathbf{h}\right) \geq 2^R\left(1 + \frac{1}{\sigma_{\mathrm{e}}^2}\mathrm{tr}\left(\mathbf{H}_{\mathrm{e},i}^H\mathbf{Q}_{\mathrm{I}} \times \mathbf{H}_{\mathrm{e},i}\right)\right)\right] \geq 1 - p, \quad \forall i.$$

Rearranging the terms in the above equation yields

$$\Pr\left[\mathrm{tr}\left(\mathbf{H}_{\mathrm{e},i}^H\mathbf{Q}_{\mathrm{I}}\mathbf{H}_{\mathrm{e},i}\right) \leq \frac{\sigma_{\mathrm{e}}^2}{2^R}\left(1 + \frac{1}{\sigma_{\mathrm{d}}^2}\mathbf{h}^H\mathbf{Q}_{\mathrm{I}}\mathbf{h}\right) - \sigma_{\mathrm{e}}^2\right] \geq 1 - p, \quad \forall i. \tag{12}$$

Replacing $\mathbf{H}_{\mathrm{e},i} = \hat{\mathbf{H}}_{\mathrm{e},i} + \boldsymbol{\Delta}_{\mathrm{H},i}$, and then performing some mathematical manipulations, we finally obtain from (12)

$$\Pr\left[\mathrm{tr}\left(\boldsymbol{\Delta}_{\mathrm{H},i}^H\mathbf{Q}_{\mathrm{I}}\boldsymbol{\Delta}_{\mathrm{H},i} + \boldsymbol{\Delta}_{\mathrm{H},i}^H\mathbf{Q}_{\mathrm{I}}\hat{\mathbf{H}}_{\mathrm{e},i} + \hat{\mathbf{H}}_{\mathrm{e},i}^H\mathbf{Q}_{\mathrm{I}}\boldsymbol{\Delta}_{\mathrm{H},i} + \hat{\mathbf{H}}_{\mathrm{e},i}^H\mathbf{Q}_{\mathrm{I}}\hat{\mathbf{H}}_{\mathrm{e},i}\right) \leq \frac{\sigma_{\mathrm{e}}^2}{2^R}\left(1 + \frac{1}{\sigma_{\mathrm{d}}^2}\mathbf{h}^H\mathbf{Q}_{\mathrm{I}}\mathbf{h}\right) - \sigma_{\mathrm{e}}^2\right] \geq 1 - p, \quad \forall i. \tag{13}$$

Recall that the relaxation (12) is in fact tight according to *Lemma 1* if $\mathrm{rank}\left(\mathbf{Q}_{\mathrm{I}}\right) \leq 1$. Our goal is to reformulate problem (10) as a tractable convex problem and then prove that the relaxation in (12) is indeed tight for the chance-constrained secrecy problem by proving the rank-one structure of $\mathbf{Q}_{\mathrm{I}}$. To make the robust problem (10) more tractable to analyze and solve, we first transform the robust constraints in (10b) and (10c) into convex inequalities using advanced matrix inequality results in the optimization literature.

Note that the probability term in (13) does not have a closed-form expression. Now we apply the following matrix identities to reformulate the secrecy outage constraint

$$\text{vec}\left(\mathbf{AXB}\right) = \left(\mathbf{B}^H \otimes \mathbf{A}\right)\text{vec}\left(\mathbf{X}\right), \tag{14a}$$

$$\text{tr}\left(\mathbf{A}^H \mathbf{B}\right) = \text{vec}\left(\mathbf{A}\right)^H \text{vec}\left(\mathbf{B}\right), \tag{14b}$$

$$\left(\mathbf{A} \otimes \mathbf{B}\right)^H = \mathbf{A}^H \otimes \mathbf{B}^H. \tag{14c}$$

Applying the identities in (14), we can express (13) as

$$\Pr\left[\boldsymbol{\delta}_{\text{H},i}^H\left(\mathbf{I}_{N_{\text{e},i}}\otimes\mathbf{Q}_\text{I}\right)\boldsymbol{\delta}_{\text{H},i} + 2\Re\left\{\boldsymbol{\delta}_{\text{H},i}^H\left(\mathbf{I}_{N_{\text{e},i}}\otimes\mathbf{Q}_\text{I}\right)\hat{\mathbf{h}}_{\text{e},i}\right\}\right.$$
$$\left. +\hat{\mathbf{h}}_{\text{e},i}^H\left(\mathbf{I}_{N_{\text{e},i}}\otimes\mathbf{Q}_\text{I}\right)\hat{\mathbf{h}}_{\text{e},i} \leq \frac{\sigma_\text{e}^2}{2^R}\left(1+\frac{1}{\sigma_\text{d}^2}\mathbf{h}^H\mathbf{Q}_\text{I}\mathbf{h}\right)-\sigma_\text{e}^2\right]$$
$$\geq 1-p, \quad \forall i, \tag{15}$$

where $\hat{\mathbf{h}}_{\text{e},i} \triangleq \text{vec}(\hat{\mathbf{H}}_{\text{e},i})$. Since $\boldsymbol{\delta}_{\text{H},i} \sim \mathcal{CN}\left(\mathbf{0},\mathbf{R}_{\text{H},i}\right)$, $\boldsymbol{\delta}_{\text{H},i}$ can be reexpressed as $\boldsymbol{\delta}_{\text{H},i} = \mathbf{R}_{\text{H},i}^{\frac{1}{2}}\mathbf{v}_{\text{H},i}$ such that $\mathbf{v}_{\text{H},i} \sim \mathcal{CN}\left(\mathbf{0},\mathbf{I}_{N_\text{T}N_{\text{e},i}}\right)$. Thus, (15) can be reexpressed as

$$\Pr\left[\mathbf{v}_{\text{H},i}^H\left[-\mathbf{R}_{\text{H},i}^{\frac{1}{2}}\left(\mathbf{I}_{N_{\text{e},i}}\otimes\mathbf{Q}_\text{I}\right)\mathbf{R}_{\text{H},i}^{\frac{1}{2}}\right]\mathbf{v}_{\text{H},i} + 2\Re\left\{\mathbf{v}_{\text{H},i}^H\right.\right.$$
$$\left.\times\left[-\mathbf{R}_{\text{H},i}^{\frac{1}{2}}\left(\mathbf{I}_{N_{\text{e},i}}\otimes\mathbf{Q}_\text{I}\right)\hat{\mathbf{h}}_{\text{e},i}\right]\right\} + \frac{\sigma_\text{e}^2}{2^R}\left(1+\frac{1}{\sigma_\text{d}^2}\mathbf{h}^H\mathbf{Q}_\text{I}\mathbf{h}\right)$$
$$\left.-\sigma_\text{e}^2 - \hat{\mathbf{h}}_{\text{e},i}^H\left(\mathbf{I}_{N_{\text{e},i}}\otimes\mathbf{Q}_\text{I}\right)\hat{\mathbf{h}}_{\text{e},i} \geq 0\right] \geq 1-p, \quad \forall i. \tag{16}$$

Similarly, the EH outage constraint can be expressed as

$$\Pr\left[\mathbf{u}_{\text{g},k}^H\mathbf{R}_{\text{g},k}^{\frac{1}{2}}\mathbf{Q}_\text{I}\mathbf{R}_{\text{g},k}^{\frac{1}{2}}\mathbf{u}_{\text{g},k} + 2\Re\left\{\mathbf{u}_{\text{g},k}^H\mathbf{R}_{\text{g},k}^{\frac{1}{2}}\mathbf{Q}_\text{I}\hat{\mathbf{g}}_k\right\}\right.$$
$$\left. +\hat{\mathbf{g}}_k^H\mathbf{Q}_\text{I}\hat{\mathbf{g}}_k - \frac{\eta_k}{\zeta_k} \geq 0\right] \geq 1-q, \forall i, \tag{17}$$

where $\boldsymbol{\delta}_{\text{g},k} = \mathbf{R}_{\text{g},k}^{\frac{1}{2}}\mathbf{u}_{\text{g},k}$ with $\mathbf{u}_{\text{g},k} \sim \mathcal{CN}\left(\mathbf{0},\mathbf{I}_{N_\text{T}}\right)$. In the following, we tackle these probabilistic constraints pursuing some convex restriction approaches. Clearly, the constraints (16) and (17) are of the form:

$$\Pr\{\mathbf{e}^H\mathbf{Qe} + 2\Re\{\mathbf{e}^H\mathbf{r}\} + s \geq 0\} \geq 1-\rho.$$

If we can find a convex function $f(\mathbf{Q},\mathbf{r},s)$, such that

$$\Pr\{\mathbf{e}^H\mathbf{Qe} + 2\Re\{\mathbf{e}^H\mathbf{r}\} + s \geq 0\} \leq f(\mathbf{Q},\mathbf{r},\mathbf{s}),$$

then we will readily have the implication [21]

$$f(\mathbf{Q},\mathbf{r},s) \leq \rho$$
$$\implies \Pr\{\mathbf{e}^H\mathbf{Qe} + 2\Re\{\mathbf{e}^H\mathbf{r}\} + s \geq 0\} \geq 1-\rho. \tag{18}$$

Hence, the L.H.S. of the implication in (18) gives a safe approximation, which is convex, of the generally intractable probabilistic constraint in the R.H.S. In the following, we attempt to derive the convex restrictions $f(\mathbf{Q},\mathbf{r},\mathbf{s})$ for tackling the probabilistic constraints. The derived restriction methods differ in terms of both computational complexity and tightness.

## A. Robust Optimization Based on BTI

The relaxation step alone does not provide a convex approximation of the original problem. The semidefinite probabilistic constraints remain intractable. In order to make the secrecy and EH outage constraints more tractable, we consider BTI in this subsection. The Bernstein-type concentration inequalities play a central role to transform the probabilistic constraints into more tractable form based on large deviation inequality for complex Gaussian quadratic vector functions. For completeness, the inequality theorem is presented in *Lemma 2* below.

*Lemma 2 (BTI): Consider the chance constraint*

$$\Pr\left[\mathbf{x}^H\mathbf{Ax} + 2\Re\left\{\mathbf{x}^H\mathbf{r}\right\} + \theta \geq 0\right] \geq 1-\rho, \tag{19}$$

*where $\mathbf{x}$ is a standard complex Gaussian random vector, i.e., $\mathbf{x} \sim \mathcal{CN}\left(\mathbf{0},\mathbf{I}_n\right)$, the 3-tuple $(\mathbf{A},\mathbf{r},\theta)$ $(\mathbf{A} \in \mathbb{H}^{n\times n}$ is a complex hermitian matrix, $\mathbf{r} \in \mathbb{C}^n$, $b \in \mathbb{R})$ forms a set of deterministic optimization variables, and $\rho \in (0,1]$ is fixed. The following implication holds [19]:*

$$\Pr\left[\mathbf{x}^H\mathbf{Ax} + 2\Re\left\{\mathbf{x}^H\mathbf{r}\right\} + \theta \geq 0\right] \geq 1-\rho$$
$$\Longleftarrow \begin{cases} \text{tr}\left(\mathbf{A}\right) - \sqrt{-2\ln(\rho)}\psi + \ln(\rho)\omega + \theta \geq 0, & (20a) \\[2mm] \left\|\begin{bmatrix}\text{vec}\left(\mathbf{A}\right) \\ \sqrt{2}\mathbf{r}\end{bmatrix}\right\| \leq \psi, & (20b) \\[2mm] \omega\mathbf{I}_n + \mathbf{A} \succeq \mathbf{0}, \quad \psi,\omega \geq 0, & (20c) \end{cases}$$

*where $\psi,\omega \in \mathbb{R}$ are slack variables.*

Note that (20) is jointly convex in $\mathbf{A}$, $\psi$, and $\omega$ and represents an efficiently computable convex restriction of the chance constraint (19). Thus using BTI, we can derive closed-form upper bounds on the violation probability to construct an efficiently computable convex function.

Indeed, for each constraint in (16), the following correspondence can be shown for all $i$:

$$\mathbf{A}_i = -\mathbf{R}_{\text{H},i}^{\frac{1}{2}}\left(\mathbf{I}_{N_{\text{e},i}}\otimes\mathbf{Q}_\text{I}\right)\mathbf{R}_{\text{H},i}^{\frac{1}{2}}, \tag{21a}$$

$$\mathbf{r}_i = -\mathbf{R}_{\text{H},i}^{\frac{1}{2}}\left(\mathbf{I}_{N_{\text{e},i}}\otimes\mathbf{Q}_\text{I}\right)\hat{\mathbf{h}}_{\text{e},i}, \tag{21b}$$

$$\theta_i = \frac{\sigma_\text{e}^2}{2^R}\left(1+\frac{1}{\sigma_\text{d}^2}\mathbf{h}^H\mathbf{Q}_\text{I}\mathbf{h}\right)-\sigma_\text{e}^2-\hat{\mathbf{h}}_{\text{e},i}^H\left(\mathbf{I}_{N_{\text{e},i}}\otimes\mathbf{Q}_\text{I}\right)\hat{\mathbf{h}}_{\text{e},i}. \tag{21c}$$

Based on *Lemma 2*, a convex safe approximation of the secrecy rate constraint in (16) can be equivalently represented by

$$\text{tr}\left(\mathbf{R}_{\text{H},i}^{\frac{1}{2}}\left(\mathbf{I}_{N_{\text{e},i}}\otimes\mathbf{Q}_\text{I}\right)\mathbf{R}_{\text{H},i}^{\frac{1}{2}}\right) + \hat{\mathbf{h}}_{\text{e},i}^H\left(\mathbf{I}_{N_{\text{e},i}}\otimes\mathbf{Q}_\text{I}\right)\hat{\mathbf{h}}_{\text{e},i} - \frac{\sigma_\text{e}^2}{2^R}$$
$$+\sqrt{-2\ln(p)}\psi_i - \ln(p)\omega_i - \frac{\sigma_\text{e}^2}{2^R\sigma_\text{d}^2}\mathbf{h}^H\mathbf{Q}_\text{I}\mathbf{h} + \sigma_\text{e}^2 \leq 0, \tag{22a}$$

$$\left\|\begin{bmatrix}\text{vec}\left(\mathbf{R}_{\text{H},i}^{\frac{1}{2}}\left(\mathbf{I}_{N_{\text{e},i}}\otimes\mathbf{Q}_\text{I}\right)\mathbf{R}_{\text{H},i}^{\frac{1}{2}}\right) \\ \sqrt{2}\mathbf{R}_{\text{H},i}^{\frac{1}{2}}\left(\mathbf{I}_{N_{\text{e},i}}\otimes\mathbf{Q}_\text{I}\right)\hat{\mathbf{h}}_{\text{e},i}\end{bmatrix}\right\| \leq \psi_i, \tag{22b}$$

$$\omega_i\mathbf{I}_{N_\text{T}N_{\text{e},i}} - \mathbf{R}_{\text{H},i}^{\frac{1}{2}}\left(\mathbf{I}_{N_{\text{e},i}}\otimes\mathbf{Q}_\text{I}\right)\mathbf{R}_{\text{H},i}^{\frac{1}{2}} \succeq \mathbf{0}, \psi_i,\omega_i \geq 0, \forall i, \tag{22c}$$

and that of the EH outage constraint (17) can be equivalently recast into

$$\text{tr}\left(\mathbf{R}_{\text{g},k}^{\frac{1}{2}}\mathbf{Q}_\text{I}\mathbf{R}_{\text{g},k}^{\frac{1}{2}}\right)+\hat{\mathbf{g}}_k^H\mathbf{Q}_\text{I}\hat{\mathbf{g}}_k+\sqrt{-2\ln(q)}v_k-\ln(q)\varphi_k-\frac{\eta_k}{\xi_k}\geq 0,$$
(23a)

$$\left\|\begin{bmatrix}\text{vec}\left(\mathbf{R}_{\text{g},k}^{\frac{1}{2}}\mathbf{Q}_\text{I}\mathbf{R}_{\text{g},k}^{\frac{1}{2}}\right)\\\sqrt{2}\mathbf{R}_{\text{g},k}^{\frac{1}{2}}\mathbf{Q}_\text{I}\hat{\mathbf{g}}_k\end{bmatrix}\right\|\leq v_k,$$
(23b)

$$\varphi_k\mathbf{I}_{N_\text{T}}+\mathbf{R}_{\text{g},k}^{\frac{1}{2}}\mathbf{Q}_\text{I}\mathbf{R}_{\text{g},k}^{\frac{1}{2}}\succeq\mathbf{0},\quad v_k,\varphi_k\geq 0,\quad\forall k,$$
(23c)

where $q\in(0,1]$ is fixed, and $v,\varphi\in\mathbb{R}$ are slack variables. Incorporating (22) and (23), the power minimization problem (10) can be equivalently reformulated as

$$\min_{\mathbf{Q}_\text{I},\{\psi_i\},\{\omega_i\},\{v_k\},\{\varphi_k\}}\quad\text{tr}(\mathbf{Q}_\text{I})$$

$$\text{s.t.}\quad\text{tr}\left(\mathbf{R}_{\text{H},i}^{\frac{1}{2}}\left(\mathbf{I}_{N_{\text{e},i}}\otimes\mathbf{Q}_\text{I}\right)\mathbf{R}_{\text{H},i}^{\frac{1}{2}}\right)+\hat{\mathbf{h}}_{\text{e},i}^H\left(\mathbf{I}_{N_{\text{e},i}}\otimes\mathbf{Q}_\text{I}\right)\hat{\mathbf{h}}_{\text{e},i}$$
(24a)

$$+\sqrt{-2\ln(p)}\psi_i-\ln(p)\omega_i-\frac{\sigma_\text{e}^2}{2^R\sigma_\text{d}^2}\mathbf{h}^H\mathbf{Q}_\text{I}\mathbf{h}-\frac{\sigma_\text{e}^2}{2^R}+\sigma_\text{e}^2\leq 0,$$
(24b)

$$\left\|\begin{bmatrix}\text{vec}\left(\mathbf{R}_{\text{H},i}^{\frac{1}{2}}\left(\mathbf{I}_{N_{\text{e},i}}\otimes\mathbf{Q}_\text{I}\right)\mathbf{R}_{\text{H},i}^{\frac{1}{2}}\right)\\\sqrt{2}\mathbf{R}_{\text{H},i}^{\frac{1}{2}}\left(\mathbf{I}_{N_{\text{e},i}}\otimes\mathbf{Q}_\text{I}\right)\hat{\mathbf{h}}_{\text{e},i}\end{bmatrix}\right\|\leq\psi_i,$$
(24c)

$$\omega_i\mathbf{I}_{N_\text{T}N_{\text{e},i}}-\mathbf{R}_{\text{H},i}^{\frac{1}{2}}\left(\mathbf{I}_{N_{\text{e},i}}\otimes\mathbf{Q}_\text{I}\right)\mathbf{R}_{\text{H},i}^{\frac{1}{2}}\succeq\mathbf{0},$$
(24d)

$$\text{tr}\left(\mathbf{R}_{\text{g},k}^{\frac{1}{2}}\mathbf{Q}_\text{I}\mathbf{R}_{\text{g},k}^{\frac{1}{2}}\right)+\hat{\mathbf{g}}_k^H\mathbf{Q}_\text{I}\hat{\mathbf{g}}_k-\sqrt{-2\ln(q)}v_k$$
$$+\ln(q)\varphi_k-\frac{\eta_k}{\xi_k}\geq 0,$$
(24e)

$$\left\|\begin{bmatrix}\text{vec}\left(\mathbf{R}_{\text{g},k}^{\frac{1}{2}}\mathbf{Q}_\text{I}\mathbf{R}_{\text{g},k}^{\frac{1}{2}}\right)\\\sqrt{2}\mathbf{R}_{\text{g},k}^{\frac{1}{2}}\mathbf{Q}_\text{I}\hat{\mathbf{g}}_k\end{bmatrix}\right\|\leq v_k,$$
(24f)

$$\varphi_k\mathbf{I}_{N_\text{T}}+\mathbf{R}_{\text{g},k}^{\frac{1}{2}}\mathbf{Q}_\text{I}\mathbf{R}_{\text{g},k}^{\frac{1}{2}}\succeq\mathbf{0},$$
(24g)

$$\mathbf{Q}_\text{I}\succeq\mathbf{0},\quad\psi_i,\omega_i\geq 0,\ \forall i,v_k,\varphi_k\geq 0,\ \forall k.$$
(24h)

The problem (24) is convex and can be efficiently solved using interior-point based solvers [26].

*Theorem 1:* Suppose that the relaxed problem (24) *is feasible for $R > 0$. The optimal solution must satisfy* rank$(\mathbf{Q}_\text{I}) = 1$.

*proof:* See Appendix A. □

### B. $\mathcal{S}$-Procedure Based Approach

In this subsection, we develop a convex restriction approach in a conservative fashion for robust optimization. The main idea is to choose a set for the channel uncertainty region satisfying the probabilistic restriction. In contrast to norm-bounded CSI errors, we have the freedom to choose the set arbitrarily in this method according to the maximum tolerable outage probability. Towards this end, the following lemma is useful.

*Lemma 3: Consider an arbitrary set $\mathcal{A}\subset\mathbb{C}^{N_\text{T}\times 1}$ satisfying* $\Pr\{\mathbf{x}\in\mathcal{A}\}\geq 1-\rho$. *The following implication holds* [21]:

$$\mathbf{x}^H\mathbf{A}\mathbf{x}+2\Re\{\mathbf{x}^H\mathbf{r}\}+\theta\geq 0,\quad\forall\mathbf{x}\in\mathcal{A},$$
$$\Longrightarrow\Pr\{\mathbf{x}^H\mathbf{Q}\mathbf{x}+2\Re\{\mathbf{x}^H\mathbf{r}\}+\theta\geq 0\}\geq 1-\rho.$$
(25)

That is, the worst-case robust constraint on the L.H.S. of (25) is a safe approximation of the probabilistic constraint on the R.H.S. Based on *Lemma 3*, given the following deterministic quadratic constraint (from (16))

$$\mathbf{v}_{\text{H},i}^H\left[-\mathbf{R}_{\text{H},i}^{\frac{1}{2}}\left(\mathbf{I}_{N_{\text{e},i}}\otimes\mathbf{Q}_\text{I}\right)\mathbf{R}_{\text{H},i}^{\frac{1}{2}}\right]\mathbf{v}_{\text{H},i}+2\Re\left\{\mathbf{v}_{\text{H},i}^H\left[-\mathbf{R}_{\text{H},i}^{\frac{1}{2}}\right.\right.$$
$$\left.\left.\times\left(\mathbf{I}_{N_{\text{e},i}}\otimes\mathbf{Q}_\text{I}\right)\right]\hat{\mathbf{h}}_{\text{e},i}\right\}+\frac{\sigma_\text{e}^2}{2^R}\left(1+\frac{1}{\sigma_\text{d}^2}\mathbf{h}^H\mathbf{Q}_\text{I}\mathbf{h}\right)-\sigma_\text{e}^2$$
$$-\hat{\mathbf{h}}_{\text{e},i}^H\left(\mathbf{I}_{N_{\text{e},i}}\otimes\mathbf{Q}_\text{I}\right)\hat{\mathbf{h}}_{\text{e},i}\geq 0,\forall i,$$
(26)

choosing the following set for the channel uncertainty region

$$\mathcal{A}=\{\mathbf{v}_{\text{H},i}\in\mathbb{C}^{N_\text{T}\times 1}|\Pr\left(\mathbf{v}_{\text{H},i}^H\mathbf{v}_{\text{H},i}\leq\gamma_\text{e}^2\right)\geq 1-p\},\forall i,$$
(27)

is sufficient to guarantee the probabilistic constraint in (16). Interestingly, with $\mathbf{v}_{\text{H},i}$ defined as $\mathbf{v}_{\text{H},i}\sim\mathcal{CN}\left(0,\mathbf{I}_{N_\text{T}N_{\text{e},i}}\right)$, it can be easily verified that $\|\mathbf{v}_{\text{H},i}\|^2$ is a Chi-square ($\chi^2$) random variable with $2N_\text{T}N_{\text{e},i}$ degrees of freedom. The channel uncertainty region in (27) always holds for $\gamma_\text{e}=\sqrt{\frac{\mathcal{F}_{\chi_m^2}^{-1}(1-p)}{2}}$, where $\mathcal{F}_{\chi_m^2}^{-1}(a)$ is the inverse cumulative distribution function (CDF) of the Chi-square random variable $a$ with $m=2N_\text{T}N_{\text{e},i}$ degrees of freedom. In fact, $\gamma_{\text{e},i}$ can be interpreted as the radius of the ball $\mathcal{A}$ defining the channel uncertainty region. Thus the probabilistic constraint (16) can be equivalently expressed by the following set of inequalities:

$$\mathbf{v}_{\text{H},i}^H\left[-\mathbf{R}_{\text{H},i}^{\frac{1}{2}}\left(\mathbf{I}_{N_{\text{e},i}}\otimes\mathbf{Q}_\text{I}\right)\mathbf{R}_{\text{H},i}^{\frac{1}{2}}\right]\mathbf{v}_{\text{H},i}+2\Re\left\{\mathbf{v}_{\text{H},i}^H\left[-\mathbf{R}_{\text{H},i}^{\frac{1}{2}}\right.\right.$$
$$\left.\left.\times\left(\mathbf{I}_{N_{\text{e},i}}\otimes\mathbf{Q}_\text{I}\right)\right]\hat{\mathbf{h}}_{\text{e},i}\right\}+\frac{\sigma_\text{e}^2}{2^R}\left(1+\frac{1}{\sigma_\text{d}^2}\mathbf{h}^H\mathbf{Q}_\text{I}\mathbf{h}\right)-\sigma_\text{e}^2$$
$$-\hat{\mathbf{h}}_{\text{e},i}^H\left(\mathbf{I}_{N_{\text{e},i}}\otimes\mathbf{Q}_\text{I}\right)\hat{\mathbf{h}}_{\text{e},i}\geq 0,\quad\forall i,$$
(28a)
$$-\mathbf{v}_{\text{H},i}^H\mathbf{v}_{\text{H},i}+\gamma_\text{e}^2\geq 0,\forall i.$$
(28b)

At this point, we apply the so-called $\mathcal{S}$-procedure [27] to transform the constraint (28) into a more tractable linear matrix inequality (LMI). The $\mathcal{S}$-procedure is presented in *Lemma 4* below.

*Lemma 4 ($\mathcal{S}$-Procedure): Let $f_i(\mathbf{x}),i=1,2$, be defined as*

$$f_i(\mathbf{x})=\mathbf{x}^H\mathbf{A}_i\mathbf{x}+2\Re\left\{\mathbf{b}_i^H\mathbf{x}\right\}+c_i,$$
(29)

*where $\mathbf{A}_i\in\mathbb{C}^{n\times n},\mathbf{b}_i\in\mathbb{C}^n,c_i\in\mathbb{R}$. The implication $f_1(\mathbf{x})\leq 0\Rightarrow f_2(\mathbf{x})\leq 0$ holds if and only if there exists $\mu\geq 0$ such that*

$$\mu\begin{bmatrix}\mathbf{A}_1&\mathbf{b}_1\\\mathbf{b}_1^H&c_1\end{bmatrix}-\begin{bmatrix}\mathbf{A}_2&\mathbf{b}_2\\\mathbf{b}_2^H&c_2\end{bmatrix}\succeq\mathbf{0}$$
(30)

*provided that there exists a point $\hat{\mathbf{x}}$ such that $f_1(\hat{\mathbf{x}})<0$.*

According to *Lemma 4*, (28) holds if and only if there exists $\mu_{\text{H},i}\geq 0,\forall i$, such that $\boldsymbol{\Theta}_i\left(\mathbf{Q}_\text{I},\mu_{\text{H},i}\right)\succeq\mathbf{0}$

$$\boldsymbol{\Theta}_i\left(\mathbf{Q}_\mathrm{I}, \mu_{\mathrm{H},i}\right) \triangleq \begin{bmatrix} \mu_{\mathrm{H},i}\mathbf{I}_{N_\mathrm{T}N_{\mathrm{e},i}} - \mathbf{R}_{\mathrm{H},i}^{\frac{1}{2}}\left(\mathbf{I}_{N_{\mathrm{e},i}}\otimes\mathbf{Q}_\mathrm{I}\right)\mathbf{R}_{\mathrm{H},i}^{\frac{1}{2}} & -\mathbf{R}_{\mathrm{H},i}^{\frac{1}{2}}\left(\mathbf{I}_{N_{\mathrm{e},i}}\otimes\mathbf{Q}_\mathrm{I}\right)\hat{\mathbf{h}}_{\mathrm{e},i} \\ -\hat{\mathbf{h}}_{\mathrm{e},i}^H\left(\mathbf{I}_{N_{\mathrm{e},i}}\otimes\mathbf{Q}_\mathrm{I}\right)\mathbf{R}_{\mathrm{H},i}^{\frac{1}{2}} & \tau_{\mathrm{H},i} - \mu_{\mathrm{H},i}\gamma_\mathrm{e}^2 \end{bmatrix} \succeq \mathbf{0}, \tag{31}$$

$$\boldsymbol{\Upsilon}_k\left(\mathbf{Q}_\mathrm{I}, \mu_{\mathrm{g},k}\right) \triangleq \begin{bmatrix} \mu_{\mathrm{g},k}\mathbf{I}_{N_\mathrm{T}} + \mathbf{R}_{\mathrm{g},k}^{\frac{1}{2}}\mathbf{Q}_\mathrm{I}\mathbf{R}_{\mathrm{g},k}^{\frac{1}{2}} & \mathbf{R}_{\mathrm{g},k}^{\frac{1}{2}}\mathbf{Q}_\mathrm{I}\hat{\mathbf{g}}_k \\ \hat{\mathbf{g}}_k^H\mathbf{Q}_\mathrm{I}\mathbf{R}_{\mathrm{g},k}^{\frac{1}{2}} & \hat{\mathbf{g}}_k^H\mathbf{Q}_\mathrm{I}\hat{\mathbf{g}}_k - \frac{\eta_k}{\zeta_k} - \mu_{\mathrm{g},k}\gamma_\mathrm{e}^2 \end{bmatrix} \succeq \mathbf{0}, \tag{32}$$

(defined in (31), as shown at the top of this page), where $\tau_{\mathrm{H},i} \triangleq \frac{\sigma_\mathrm{e}^2}{2^R}\left(1 + \frac{1}{\sigma_\mathrm{d}^2}\mathbf{h}^H\mathbf{Q}_\mathrm{I}\mathbf{h}\right) - \sigma_\mathrm{e}^2 - \hat{\mathbf{h}}_{\mathrm{e},i}^H\left(\mathbf{I}_{N_{\mathrm{e},i}}\otimes\mathbf{Q}_\mathrm{I}\right)\hat{\mathbf{h}}_{\mathrm{e},i}, \forall i$.

Similarly, the EH outage constraint (17) can be transformed to the LMI in (32), as shown at the top of this page, where $\mu_{\mathrm{g},k} \geq 0, \forall k$. By exploiting $\mathcal{S}$-Procedure, the power minimization problem (10) can be equivalently reformulated as

$$\min_{\mathbf{Q}_\mathrm{I},\{\mu_{\mathrm{H},i}\},\{\mu_{\mathrm{g},k}\}} \quad \mathrm{tr}\left(\mathbf{Q}_\mathrm{I}\right) \tag{33a}$$

$$\mathrm{s.t.} \quad \boldsymbol{\Theta}_i\left(\mathbf{Q}_\mathrm{I}, \mu_{\mathrm{H},i}\right) \succeq \mathbf{0}, \forall i, \tag{33b}$$

$$\boldsymbol{\Upsilon}_k\left(\mathbf{Q}_\mathrm{I}, \mu_{\mathrm{g},k}\right) \succeq \mathbf{0}, \forall k, \tag{33c}$$

$$\mathbf{Q}_\mathrm{I} \succeq \mathbf{0}, \ \mu_{\mathrm{H},i} \geq 0, \quad \forall i, \ \mu_{\mathrm{g},k} \geq 0, \ \forall k. \tag{33d}$$

The SDP problem (33) is convex and can be efficiently solved using interior-point based solvers [26]. Interestingly, the following theorem states that the relaxed problem (33) always yields a rank-one transmit beamforming solution.

*Theorem 2: Suppose that the relaxed problem* (33) *is feasible for $R > 0$. The optimal solution must satisfy* $\mathrm{rank}\left(\mathbf{Q}_\mathrm{I}\right) = 1$.
    *proof:* See Appendix B. □

*Remark 1: At this point, we would like to comment that the BTI-based approach has higher computational complexity compared to the $\mathcal{S}$-procedure based approach as the former involves a more compound mixture of different types of constraints [21]. However, we will perform a full complexity analysis later in this section.*

### C. LDI Based Approach

Since the BTI and $\mathcal{S}$-procedure based approaches transform the chance-constrained optimization problem (10) into SDPs, the resulting safe designs are polynomial-time solvable [27]. The SDPs can, however, be very expensive to solve if the size of the LMI constraints in (24) and (33) is sufficiently large. Hence in this subsection, our endeavour is to develop convex restrictions involving simpler conic constraints. The method follows from the decomposition-based LDI [28] for complex Gaussian quadratic functions as defined in the following lemma.

*Lemma 5 [21, Lemma 2]: Let $\mathbf{x} \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{I}_n\right)$ be a standard complex Gaussian random vector, and let $\mathbf{A} \in \mathbb{H}^{n\times n}$ and $\mathbf{r} \in \mathbb{C}^n$ be given. Then, for any $v > \frac{1}{\sqrt{2}}$ and $\zeta > 0$, we have*

$$\Pr\left\{\mathbf{x}^H\mathbf{A}\mathbf{x} + 2\mathrm{Re}\{\mathbf{x}^H\mathbf{r}\} \leq \mathrm{tr}(\mathbf{A}) - \zeta\right\}$$
$$\leq \begin{cases} \exp\left(-\frac{\zeta^2}{4T^2}\right) & for \ 0 < \zeta \leq 2\bar{v}vT, \\ \exp\left(-\frac{\bar{v}v\zeta}{T} + (\bar{v}v)^2\right) & for \ \zeta > 2\bar{v}vT, \end{cases} \tag{34}$$

*where $\bar{v} = 1 - \frac{1}{2v^2}$ and $T = v\|\mathbf{A}\|_F + \frac{1}{\sqrt{2}}\|\mathbf{r}\|$.*

The merit of *Lemma 5* is that it helps decomposing a sum of dependent random variables into sums of independent random variables. This idea has been used extensively in the literature of probability theory; see, e.g., [28], [29].

Next, we concentrate on deriving convex restrictions of (16) and (17) based on the LDI approach using *Lemma 5*. For equation (16), we set

$$\zeta_i = \mathrm{tr}\left(-\mathbf{R}_{\mathrm{H},i}^{\frac{1}{2}}\left(\mathbf{I}_{N_{\mathrm{e},i}}\otimes\mathbf{Q}_\mathrm{I}\right)\mathbf{R}_{\mathrm{H},i}^{\frac{1}{2}}\right) + \tau_{\mathrm{H},i}, \tag{35}$$

$$T_{\mathrm{H},i} = v\| - \mathbf{R}_{\mathrm{H},i}^{\frac{1}{2}}\left(\mathbf{I}_{N_{\mathrm{e},i}}\otimes\mathbf{Q}_\mathrm{I}\right)\mathbf{R}_{\mathrm{H},i}^{\frac{1}{2}}\|_F$$
$$+ \frac{1}{\sqrt{2}}\| - \mathbf{R}_{\mathrm{H},i}^{\frac{1}{2}}\left(\mathbf{I}_{N_{\mathrm{e},i}}\otimes\mathbf{Q}_\mathrm{I}\right)\hat{\mathbf{h}}_{\mathrm{e},i}\|. \tag{36}$$

Here $v$ is obtained from the solution to the following quadratic equation

$$\bar{v}v = (1 - 1/(2v^2))v = \sqrt{-\ln(p)} \tag{37}$$

such that $v > \frac{1}{\sqrt{2}}$. It has been shown in [21] that such a $v$ must always exist, since $(1 - 1/(2v^2))v = 0$ when $v = \frac{1}{\sqrt{2}}$ and $(1 - 1/(2v^2))v$ is a monotonically increasing function of $v$ within the interval $[\frac{1}{\sqrt{2}}, \infty)$. Now from (37), we conclude that $p = \exp\left(-(\bar{v}v)^2\right)$. Furthermore, according to *Lemma 5*, the chance constraint (16) will be satisfied if we choose $\zeta_i = 2\sqrt{-\ln(p)}T_{\mathrm{H},i}$ for the interval $2\sqrt{-\ln(p)}T_{\mathrm{H},i} \leq \zeta_i \leq 2\bar{v}vT_{\mathrm{H},i}$. On the other hand, if $\zeta_i > 2\bar{v}vT_{\mathrm{H},i} = 2\sqrt{-\ln(p)}T_{\mathrm{H},i}$, then *Lemma 5* yields

$$\Pr\left\{\mathbf{v}_{\mathrm{H},i}^H\left[-\mathbf{R}_{\mathrm{H},i}^{\frac{1}{2}}\left(\mathbf{I}_{N_{\mathrm{e},i}}\otimes\mathbf{Q}_\mathrm{I}\right)\mathbf{R}_{\mathrm{H},i}^{\frac{1}{2}}\right]\mathbf{v}_{\mathrm{H},i}\right.$$
$$\left. + 2\Re\left\{\mathbf{v}_{\mathrm{H},i}^H\left[-\mathbf{R}_{\mathrm{H},i}^{\frac{1}{2}}\left(\mathbf{I}_{N_{\mathrm{e},i}}\otimes\mathbf{Q}_\mathrm{I}\right)\right]\hat{\mathbf{h}}_{\mathrm{e},i}\right\} + \tau_{\mathrm{H},i} \leq 0\right\}$$
$$\leq \exp\left(-\frac{\bar{v}v\zeta_i}{T_{\mathrm{H},i}} + (\bar{v}v)^2\right) < \exp\left(-(\bar{v}v)^2\right) = p, \tag{38}$$

which essentially indicates that the chance constraint (16) will still be satisfied. The resulting convex restriction can thus be expressed as

$$\mathrm{tr}(-\mathbf{R}_{\mathrm{H},i}^{\frac{1}{2}}\left(\mathbf{I}_{N_{\mathrm{e},i}}\otimes\mathbf{Q}_\mathrm{I}\right)\mathbf{R}_{\mathrm{H},i}^{\frac{1}{2}}) + \tau_{\mathrm{H},i} \geq 2\sqrt{-\ln(p)}T_{\mathrm{H},i}, \quad \forall i. \tag{39}$$

Thus using the definition of $T_{\mathrm{H},i}$, we obtain the following system of second-order cone (SOC) constraints from (39) in

order to tackle the probabilistic constraint (16):

$$
\begin{cases}
\mathrm{tr}\left(-\mathbf{R}_{\mathrm{H},i}^{\frac{1}{2}}\left(\mathbf{I}_{N_{\mathrm{e},i}}\otimes\mathbf{Q}_{\mathrm{I}}\right)\mathbf{R}_{\mathrm{H},i}^{\frac{1}{2}}\right)+\tau_{\mathrm{H}}\gtreqless 2\sqrt{-\ln(p)}\\
\qquad\qquad\qquad\qquad\qquad\times\left(\bar{\psi}_i+\bar{\omega}_i\right),\\
\frac{1}{\sqrt{2}}\left\|-\mathbf{R}_{\mathrm{H},i}^{\frac{1}{2}}\left(\mathbf{I}_{N_{\mathrm{e},i}}\otimes\mathbf{Q}_{\mathrm{I}}\right)\hat{\mathbf{h}}_{\mathrm{e},i}\right\|\leq\bar{\psi}_i,\\
v\left\|-\mathbf{R}_{\mathrm{H},i}^{\frac{1}{2}}\left(\mathbf{I}_{N_{\mathrm{e},i}}\otimes\mathbf{Q}_{\mathrm{I}}\right)\mathbf{R}_{\mathrm{H},i}^{\frac{1}{2}}\right\|_F\leq\bar{\omega}_i,
\end{cases}
\tag{40}
$$

where $\bar{\psi}_i,\bar{\omega}_i\in\mathbb{R},\forall i$, are slack variables. Similarly, defining the slack variables $\bar{v}_k$ and $\bar{\varphi}_k$, the EH outage constraint (17) can be expressed as

$$
\begin{cases}
\mathrm{tr}\left(\mathbf{R}_{\mathrm{g},k}^{\frac{1}{2}}\mathbf{Q}_{\mathrm{I}}\mathbf{R}_{\mathrm{g},k}^{\frac{1}{2}}\right)+\hat{\mathbf{g}}_k^H\mathbf{Q}_{\mathrm{I}}\hat{\mathbf{g}}_k-\frac{\eta_k}{\xi_k}\geq 2\sqrt{-\ln(q)}\\
\qquad\qquad\qquad\qquad\qquad\times\left(\bar{v}_k+\bar{\varphi}_k\right),\\
\frac{1}{\sqrt{2}}\left\|\mathbf{R}_{\mathrm{g},k}^{\frac{1}{2}}\mathbf{Q}_{\mathrm{I}}\hat{\mathbf{g}}_k\right\|\leq\bar{v}_k,\\
v\left\|\mathbf{R}_{\mathrm{g},k}^{\frac{1}{2}}\mathbf{Q}_{\mathrm{I}}\mathbf{R}_{\mathrm{g},k}^{\frac{1}{2}}\right\|_F\leq\bar{\varphi}_k.
\end{cases}
\tag{41}
$$

By applying the LDI method to the outage constrained problem (10), we obtain the convex restriction formulation of the power minimization problem as

$$
\min_{\mathbf{Q}_{\mathrm{I}},\{\bar{\psi}_i\},\{\bar{\omega}_i\},\{\bar{v}_k\},\{\bar{\varphi}_k\}}\quad\mathrm{tr}\left(\mathbf{Q}_{\mathrm{I}}\right)
\tag{42a}
$$

$$
\text{s.t.}\quad\mathrm{tr}\left(-\mathbf{R}_{\mathrm{H},i}^{\frac{1}{2}}\left(\mathbf{I}_{N_{\mathrm{e},i}}\otimes\mathbf{Q}_{\mathrm{I}}\right)\mathbf{R}_{\mathrm{H},i}^{\frac{1}{2}}\right)+\tau_{\mathrm{H},i}
$$
$$
\geq 2\sqrt{-\ln(p)}\left(\bar{\psi}_i+\bar{\omega}_i\right),
\tag{42b}
$$

$$
\frac{1}{\sqrt{2}}\left\|-\mathbf{R}_{\mathrm{H},i}^{\frac{1}{2}}\left(\mathbf{I}_{N_{\mathrm{e},i}}\otimes\mathbf{Q}_{\mathrm{I}}\right)\hat{\mathbf{h}}_{\mathrm{e},i}\right\|\leq\bar{\psi}_i,
\tag{42c}
$$

$$
v\left\|\mathrm{vec}\left(-\mathbf{R}_{\mathrm{H},i}^{\frac{1}{2}}\left(\mathbf{I}_{N_{\mathrm{e},i}}\otimes\mathbf{Q}_{\mathrm{I}}\right)\mathbf{R}_{\mathrm{H},i}^{\frac{1}{2}}\right)\right\|\leq\bar{\omega}_i,
\tag{42d}
$$

$$
\mathrm{tr}\left(\mathbf{R}_{\mathrm{g},k}^{\frac{1}{2}}\mathbf{Q}_{\mathrm{I}}\mathbf{R}_{\mathrm{g},k}^{\frac{1}{2}}\right)+\hat{\mathbf{g}}_k^H\mathbf{Q}_{\mathrm{I}}\hat{\mathbf{g}}_k-\frac{\eta_k}{\xi_k}
$$
$$
\geq 2\sqrt{-\ln(q)}\left(\bar{v}_k+\bar{\varphi}_k\right),
\tag{42e}
$$

$$
\frac{1}{\sqrt{2}}\left\|\mathbf{R}_{\mathrm{g},k}^{\frac{1}{2}}\mathbf{Q}_{\mathrm{I}}\hat{\mathbf{g}}_k\right\|\leq\bar{v}_k,
\tag{42f}
$$

$$
v\left\|\mathrm{vec}\left(\mathbf{R}_{\mathrm{g},k}^{\frac{1}{2}}\mathbf{Q}_{\mathrm{I}}\mathbf{R}_{\mathrm{g},k}^{\frac{1}{2}}\right)\right\|\leq\bar{\varphi}_k,
\tag{42g}
$$

$$
\mathbf{Q}_{\mathrm{I}}\succeq\mathbf{0},\quad\bar{\psi}_i,\bar{\omega}_i\geq 0,\quad\forall i,\bar{v}_k,\bar{\varphi}_k\geq 0,\forall k.
\tag{42h}
$$

Since the above convex problem contains only SOC constraints, it can be solved more efficiently than the convex restrictions obtained using BTI based and $\mathcal{S}$-procedure based approaches. Finally, the following theorem studies the tightness of the rank relaxation in problem (10).

*Theorem 3:* Suppose that the relaxed problem (42) *is feasible for* $R > 0$. *The optimal solution must satisfy* $\mathrm{rank}\left(\mathbf{Q}_{\mathrm{I}}\right)=1$.

*proof:* See Appendix C. $\qquad\square$

*Remark 2: It can be verified that the optimal solutions to problems (24), (33), and (42) obtained in Theorems 1, 2, and 3, respectively, are unique. Let us first consider that there are two distinct optimal solutions to problem (24), say* $\mathbf{Q}_1$ *and* $\mathbf{Q}_2$ *,such that* $\mathrm{rank}(\mathbf{Q}_1)=\mathrm{rank}(\mathbf{Q}_2)=1$. *Therefore, the range spaces of* $\mathbf{Q}_1$ *and* $\mathbf{Q}_2$ *must be different.*

*Following a basic concept in convex optimization, any* $\mathbf{Q}_3=\lambda\mathbf{Q}_1+(1-\lambda)\mathbf{Q}_2$, *for* $\lambda\in(0,1)$, *is also an optimal solution of (24) [27]. Since* $\mathbf{Q}_1$ *and* $\mathbf{Q}_2$ *are both rank-one and distinct, it is rigid that* $\mathbf{Q}_3$ *is of rank two, which contradicts with the result proved in* Theorem 1. *Hence problem (24) must have only one optimal solution* $\mathbf{Q}_{\mathrm{I}}$. *Similarly, the uniqueness of the optimal solutions of problems (33), and (42) can also be verified.*

### D. Complexity Analysis

In this subsection, we mathematically characterize the computational complexity of the proposed schemes. Note that the convex restriction formulations (24), (33), and (42) involve only LMI and SOC constraints, and hence can be solved using standard interior-point methods (IPM) [30, Lecture 6]. Therefore, we can use the worst-case computation time of IPM to compare the complexities of the three formulations. Now, using the transformation

$$
\mathbb{H}^n\ni\mathbf{S}\mapsto\begin{bmatrix}\mathrm{Re}(\mathbf{S}) & -\mathrm{Im}(\mathbf{S})\\\mathrm{Im}(\mathbf{S}) & \mathrm{Re}(\mathbf{S})\end{bmatrix}\in\mathbb{S}^{2n},
$$

where $\mathbb{S}^n$ and $\mathbb{H}^n$ represent the sets of $n\times n$ real symmetric matrices and complex Hermitian matrices, respectively, we can convert the complex-valued conic programs (24), (33), and (42) into equivalent real-valued conic programs of the form [21], [31]

$$
\min_{\mathbf{z}\in\mathbb{R}^n}\quad\mathbf{c}^T\mathbf{z}
$$
$$
\text{s.t.}\quad\sum_{i=1}^n z_i\mathbf{A}_i^j-\mathbf{B}^j\in\mathbb{S}_+^{k_j}\quad\text{for } j=1,\ldots,p,
$$
$$
\mathbf{T}^j\mathbf{z}-\mathbf{b}^j\in\mathbb{L}^{k_j}\quad\text{for } j=p+1,\ldots,m.
\tag{43}
$$

Here, $\mathbf{T}^j\in\mathbb{R}^{k_j\times n}$, $\mathbf{b}^j\in\mathbb{R}^{k_j}$ for $j=p+1,\ldots,m$, $\mathbf{c}\in\mathbb{R}^n$, $\mathbb{S}_+^k$ is the set of $k\times k$ real PSD matrices, and $\mathbb{L}^k$ is the second-order cone of dimention $k\geq 1$. Now the overall complexity of the IPM for solving the above problem consists of two components:

a) *Iteration Complexity:* The number of iterations required to reach an $\epsilon$-accurate ($\epsilon>0$) optimal solution of problem (43) is in the order of $\ln(1/\epsilon)\sqrt{\beta(\mathcal{K})}$, where $\beta(\mathcal{K})=\sum_{j=1}^p k_j+2(m-p)$ is known to be the barrier parameter.

b) *Per-Iteration Computation Cost:* A system of $n$ linear equations is required to be solved in each iteration. The computation tasks include the formation of the coefficient matrix $\mathbf{H}$ of the system of linear equations and the factorization of $\mathbf{H}$. The cost of forming $\mathbf{H}$ sums on the order of $\kappa_{\mathrm{for}}=n\sum_{j=1}^p k_j^3+n^2\sum_{j=1}^p k_j^2+n\sum_{j=p+1}^m k_j^2$ while the cost of factorization is on the order of $\kappa_{\mathrm{fac}}=n^3$ [21].

Thus the overall computation cost for solving (43) using IPM is on the order of $\ln(1/\epsilon)\sqrt{\beta(\mathcal{K})}\times(\kappa_{\mathrm{for}}+\kappa_{\mathrm{fac}})$. Using these concepts, we can now analyze the computational complexity of problems (24), (33), and (42). Note that in all three formulations, the number of decision variables ($n$ in (43)) is on the order of $N_{\mathrm{T}}^2$ (ignoring the slack variables). Let us first examine problem (24), which has $L$ LMI (trace) constraints of size 1, $L$ SOC constraints (of size 2), $L$ LMI constraints of

TABLE I
COMPLEXITY ANALYSIS OF THE PROPOSED APPROACHES

| Method | Complexity Order ($n = \mathcal{O}(N_{\mathrm{T}}^2)$) |
|---|---|
| BTI | $\ln(1/\epsilon)\sqrt{N_{\mathrm{T}}(K+L+1+)}n[(K+L)(N_{\mathrm{T}}^3 + nN_{\mathrm{T}}^2 + 3n + 3) + (K+L)(N_{\mathrm{T}}^2 + N_{\mathrm{T}} + 1)^2 + N_{\mathrm{T}}^3 + nN_{\mathrm{T}}^2 + n^2]$ |
| $\mathcal{S}$-procedure | $\ln(1/\epsilon)\sqrt{N_{\mathrm{T}}(K+L+1+)}n[(K+L)(N_{\mathrm{T}}+1)^3 + n(K+L)(N_{\mathrm{T}}+1)^2 + n(N_{\mathrm{T}}^2 + K + L) + N_{\mathrm{T}}^3 + K + L]$ |
| LDI | $\ln(1/\epsilon)\sqrt{N_{\mathrm{T}}(K+L+1+)}n[(N_{\mathrm{T}}^3 + 3K + 3L) + n(N_{\mathrm{T}}^2 + 3K + 3L) + (K+L)((N_{\mathrm{T}}^2+1)^2 + (N_{\mathrm{T}}+1)^2) + n^2]$ |

size $N_{\mathrm{T}}$, $K$ LMI (trace) constraints of size 1, $K$ SOC constraints, $K$ LMI constraints of size $N_{\mathrm{T}}$, and in (24h), 1 LMI constraints of size $N_{\mathrm{T}}$, $2L$ LMI constraints of size 1, $2K$ LMI constraints of size 1. Thus the complexity of the BTI-based algorithm is on the order shown in the first row of Table I. Similarly, the complexity of the $\mathcal{S}$-procedure based approach and the LDI-based approach can be quantified as shown in the second and the third row of Table I, respectively.

From Table I, it is straightforward to show that LDI-based method has the lowest computational complexity since it involves only SOC constraints, while the BTI-based approach has the highest computational cost since it involves a more complicated set of constraints. However, in terms of tightness, the $\mathcal{S}$-procedure based approach performs the worst, as evidenced by our numerical results.

## IV. ROBUST SRM

In the robust power minimization problem considered in Section III, attempt has been made to keep the transmit power as low as possible yet maintaining the predefined secrecy rate $R$, as well as the harvested energy, within the secrecy outage probability. However, in many practical wireless communication systems (e.g., secondary users' transmission in cognitive radio systems, small cell users' transmission in heterogeneous networks (HetNets)), the maximum allowable transmission power is limited to a certain level so as to keep the interference to other users below a given threshold. In those scenarios, SRC power minimization problem (10) may turn out to be infeasible [22], [32] and the designer may need to re-adjust the secrecy rate requirement in order to find a feasible solution. With insufficient channel knowledge, finding out the appropriate secrecy rate requirement can be a tedious job. Instead, a more attractive problem formulation can be to find the maximum secrecy rate $R$ that can be achieved subject to the same outage constraints and the additional transmit power constraint. Thus, the SRM problem with outage constraints for a given maximum transmission power can be represented as follows:

$$\max_{\mathbf{Q}_{\mathrm{I}}, R} R \tag{44a}$$

$$\text{s.t. } \Pr\left[\min_i \left\{C_{\mathrm{I}}(\mathbf{Q}_{\mathrm{I}}) - \hat{C}_{\mathrm{e},i}(\mathbf{Q}_{\mathrm{I}})\right\}^+ \geq R\right]$$
$$\geq 1 - p, \forall i, \tag{44b}$$

$$\Pr\left[\min_k \hat{E}_k \geq \eta_k\right] \geq 1 - q, \forall k, \tag{44c}$$

$$\text{tr}(\mathbf{Q}_{\mathrm{I}}) \leq P_{\mathrm{T}}, \quad \mathbf{Q}_{\mathrm{I}} \succeq \mathbf{0}, \tag{44d}$$

where $P_{\mathrm{T}}$ is the maximum available transmission power budget at the transmitter. This problem is not convex in terms of the outage constraints. To make this problem more tractable, we propose a two-stage optimization procedure as shown below:

$$= \max_R \begin{cases} \max_{\mathbf{Q}_{\mathrm{I}}} R & (45a) \\[2mm] \text{s.t. } \Pr\left[\min_i \left\{C_{\mathrm{I}}(\mathbf{Q}_{\mathrm{I}}) - \hat{C}_{\mathrm{e},i}(\mathbf{Q}_{\mathrm{I}})\right\}^+ \geq R\right] \\ \qquad\qquad \geq 1 - p, \forall i, & (45b) \\[2mm] \Pr\left[\min_k \hat{E}_k \geq \eta_k\right] \geq 1 - q, \quad \forall k, & (45c) \\[2mm] \text{tr}(\mathbf{Q}_{\mathrm{I}}) \leq P_{\mathrm{T}}, \quad \mathbf{Q}_{\mathrm{I}} \succeq \mathbf{0}. & (45d) \end{cases}$$

In the first stage, we solve the inner maximization problem of (45) for any given feasible $R$. In the second stage, we perform a one-dimensional line search over $R$ that leads to the optimal solution of the problem (44). Note that even with given $R$, the problem is not tractable due to the probabilistic constraints. Hence we apply the safe approximation approaches derived in the previous section for the probabilistic constraints.

### A. SRM Based on BTI

According to *Lemma 2*, the BTI based convex restrictions for constraints (45a) and (45a) are given by (22) and (23), respectively, for given $R$. Thus, the safe approximation for the inner maximization problem in (45) is given by

$$\max_{\mathbf{Q}_{\mathrm{I}}, \{\psi_i\}, \{\omega_i\}, \{\nu_k\}, \{\varphi_k\}} R \tag{46a}$$

$$\text{s.t. } \text{tr}(\mathbf{Q}_{\mathrm{I}}) \leq P_{\mathrm{T}}, \tag{46b}$$

$$\text{(24b)–(24h) satisfied.} \tag{46c}$$

Our next endeavour is to establish a link between the optimal solutions of (24) and (46). If we can prove that the the optimal solution of problem (24) is also optimal for problem (46), then we can readily obtain the optimal solution to the inner maximization problem in (45).

*Proposition 1: Any optimal solution to the power minimization problem* (24) *is also optimal to the problem* (46) *for identical specifications.*

*proof:* The proof is identical to that of [10, Th. 2] and is thus omitted for brevity. □

By *Theorem 1*, we know that the solution to the power minimization problem (24) is rank-one optimal, so is the solution of the problem (46) – we can immediately infer from *Proposition 1*. The remaining task is to find the optimal $R$ from the second stage of problem (45). A simple one-dimensional linear search (e.g., bisection or golden-section search) over $R$ to find the maximal $R$ that solves the feasibility problem (45) is sufficient. The lower boundary of the search is obviously 0 due to the assumption $R > 0$. The upper limit can be defined by assuming that the system vintages the highest secrecy rate at zero eavesdropping capacity i.e., at $C_{\mathrm{e},i} = 0, \forall i$. Thus, we obtain the upper search limit from (44b) as

$$R \leq \log\left(1 + \frac{1}{\sigma_{\mathrm{d}}^2}\mathbf{h}^H\mathbf{Q}_{\mathrm{I}}\mathbf{h}\right) \leq \log\left(1 + \frac{1}{\sigma_{\mathrm{d}}^2}\text{tr}(\mathbf{Q}_{\mathrm{I}})\|\mathbf{h}\|^2\right)$$

$$\leq \log\left(1 + \frac{P_{\mathrm{T}}}{\sigma_{\mathrm{d}}^2}\|\mathbf{h}\|^2\right). \tag{47}$$

Note that the last inequality is derived using the sum power constraint $\mathrm{tr}(\mathbf{Q}_I) \leq P_T$ in (44d).

### B. S-Procedure Based SRM

According to *Lemma 4*, the $\mathcal{S}$-procedure based convex restrictions for constraints (45a) and (45a) are given by (31) and (32), respectively. As a result, a safe approximation for the inner maximization problem in (45) is given by

$$\max_{\mathbf{Q}_I, \{\mu_{H,i}\}, \{\mu_{g,k}\}} R \tag{48a}$$

$$\text{s.t.} \, \mathrm{tr}(\mathbf{Q}_I) \leq P_T, \tag{48b}$$

$$(33b)\text{--}(33d) \text{ satisfied}. \tag{48c}$$

The following proposition establishes the solution equivalence between (48) and (33).

*Proposition 2: Any optimal solution to the power minimization problem* (33) *is also optimal to the problem* (48) *for identical specifications.*

*proof:* The proof is identical to that of [10, Th. 2] and is thus omitted for brevity. □

By *Theorem 2*, it is already known that the optimal solution to the power minimization problem (33) yields a rank-one transmit covariance $\mathbf{Q}_I$. Hence, the solution to the problem (48) should also be of unit-rank. The remaining task is to find the optimal $R$ for problem (45) through a one-dimensional line search as described in Subsection IV-A.

### C. LDI Based SRM

Similar to the BTI and the $\mathcal{S}$-procedure based approaches, a convex safe approximation for the inner maximization problem in (45) is given by *Lemma 5* as

$$\max_{\mathbf{Q}_I, \{\bar{\psi}_i\}, \{\bar{\omega}_i\}, \{\bar{v}_k\}, \{\bar{\varphi}_k\}} R \tag{49a}$$

$$\text{s.t.} \, \mathrm{tr}(\mathbf{Q}_I) \leq P_T, \tag{49b}$$

$$(42b)\text{--}(42h) \text{ satisfied}, \tag{49c}$$

which can be efficiently solved using existing solvers [26]. The solution obtained is identical to that of the power minimization problem (42) as described by the following proposition.

*Proposition 3: Any optimal solution to the power minimization problem* (42) *is also optimal to the problem* (49) *for identical specifications.*

*proof:* The proof is identical to that of [10, Th. 2] and is thus omitted for brevity. □

The solution has been proved to be rank-one optimal in *Theorem 3*. Finally, a line search is performed to find the optimal $R$ as described in Subsection IV-A.

## V. SIMULATION RESULTS

Here, we study the performance of the proposed algorithms in MISO secrecy SWIPT systems with probabilistic constraints through numerical simulations. For simplicity, it was assumed that $\eta_k = \eta, \xi_k = 1, \forall k, N_{e,i} = N_e, \forall i, p = q = \rho, \sigma_d^2 = \sigma_e^2 = 1, d_I = d_{h,k} = d_{e,i} = d$. We set $L_c = 35.97 \times 10^{-4}$, (i.e., when $G_T = 18$ dBi, $G_R = -2$ dBi, $c = 3 \times 10^8$, $f = 1000$ MHz), path loss exponent of $\bar{\kappa} = 2.7$, and distance
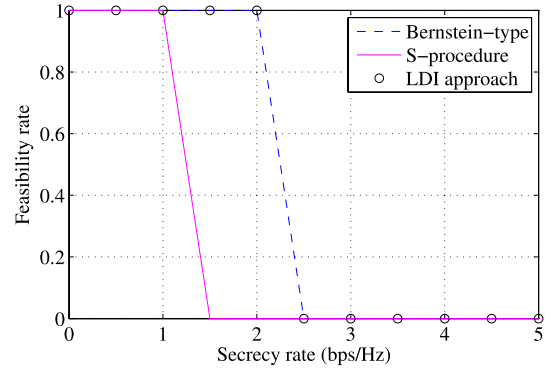


Fig. 2. Feasibility rates of the proposed methods with $N_T = 6$, $K = L = N_e = 3$, and $\eta = 0$ (dB).

$d = 10$ meters. In particular, we examine the case in which the QoS requirements are such that each user is provided with an outage probability of at most 10%; i.e., $\rho = 0.1$, unless otherwise specified. We simulated a flat Rayleigh fading environment where the channel vectors have entries with zero mean and variance $1/N_T$.

Since Monte-Carlo simulations become prohibitively expensive under very low outage requirements for the convex restriction approaches, the algorithms developed in this paper are generally not recommended [21]. However, to illustrate the performance of the proposed approaches in a larger error domain, we average the results over 500 realizations of the estimated channels.

We start the performance analysis of the proposed convex restriction formulations by comparing their feasibility rates, i.e., the chance of getting a feasible solution to the problem (10) in 500 realizations of the estimated channels. Fig. 2 shows the feasibility rates of the three approaches for $N_T = 6$, $K = L = N_e = 3$, and $\eta = 0$ (dB). Interestingly, the $\mathcal{S}$-procedure based approach has the lowest feasibility rate compared to the other two methods. The BTI-based and the LDI-based approaches yield almost the same feasibility rate. Therefore, one would expect the worst performance from the $\mathcal{S}$-procedure based approach in terms of secrecy rate as well as transmit power which we will observe in the remaining examples.

In the next example, we compare the transmit power consumptions of the proposed robust convex restriction solutions with that of the conventional perfect eavesdroppers' CSI based non-robust approach [25] for problem (10). In the non-robust scheme, the estimated CSIs have been used for designing the transmit covariance matrix. Fig. 3 shows the transmit power required by various methods against the rate outage threshold $R$ with $N_T = 8$, $K = 3$, $L = 2$, $N_e = 2$. We plot the results for $\eta = 10$ (dBm) and $\eta = 15$ (dBm). As can be seen from Fig. 3, the BTI-based approach yields slightly better transmit power performance compared to the LDI-based restriction approach. However, the $\mathcal{S}$-procedure based approach performs the worst amongst the proposed approaches. Interestingly, the robust algorithms guarantee the probabilistic constraints costing very little additional transmit power compared to the non-robust scheme. Obviously, with the increase in the required secrecy rate, all the algorithms demand higher transmit power.
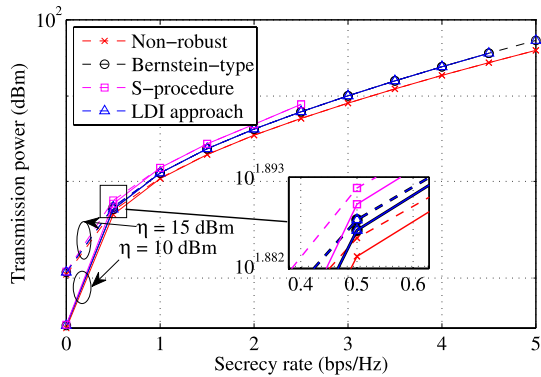
Fig. 3.   Transmit power versus secrecy outage requirement $R$ with $N_T = 8$, $K = 3$, $L = N_e = 2$, and $\eta = 10, 15$ (dBm).
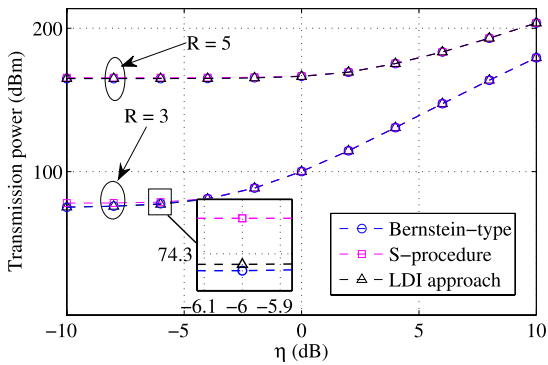


Fig. 5.   Achievable secrecy rate by the proposed algorithms versus $P_T$ with $N_T = 8$, $K = L = N_e = 3$, and $\eta = -5$.



Fig. 4.   Transmit power versus EH outage requirement $\eta$ with $N_T = 8$, $K = 3$, $L = N_e = 2$, and $R = 3, 5$ (bps/Hz).



Fig. 6.   The secrecy rate for different number of the eavesdroppers with $N_T = 6$, $K = 3$, $N_e = 2$, $P_T = 15, 20$ (dB), and $\eta = -5$ (dB).

The results in Fig. 3 indicate that when the EH requirement is increased from $\eta = -10$ (dB) to $\eta = 0$ (dB), all the approaches require more transmit power to satisfy the more demanding constraints. Hence, in the next example, we illustrate the transmit power consumptions of the solutions against the EH requirements. Fig. 4 shows the required transmit power versus the EH outage threshold $\eta$ with $N_T = 8$, $K = 3$, $L = 2$, $N_e = 2$. We plot the results for $R = 3$ and $R = 5$ (bps/Hz). The results in Fig. 4 show identical characteristics of the probabilistic restriction solutions as those in Fig. 3. Again, with increased EH outage requirement, all the algorithms demand higher transmit power.

Next, we examine the secrecy rate performance of the proposed safe approximation approaches for the SRM problem (44). Fig. 5 plots the worst-user secrecy rates of the various methods against the transmit power constraint $P_T$ for outage tolerance of 5% and 10% for both the secrecy rate and EH constraints. As a baseline scheme, we also plot the secrecy rate of the classic maximal ratio transmission (MRT) scheme. In the MRT scheme, we pick the IR's channel direction for transmit beamforming assuming no eavesdropper is present. Specifically, the transmit covariance matrix is defined as $\mathbf{Q}_I = \left(\frac{P_T}{\|\mathbf{h}\|^2}\right)\mathbf{h}\mathbf{h}^H$. Hence, it is a very suboptimal method and undergoes severe performance degradation compared with the proposed robust schemes. We set other parameters as $N_T = 8$, $K = L = 3$, $N_e = 2$, and $\eta = -5$ (dB). As we can see, the BTI- and LDI-based approaches yield almost
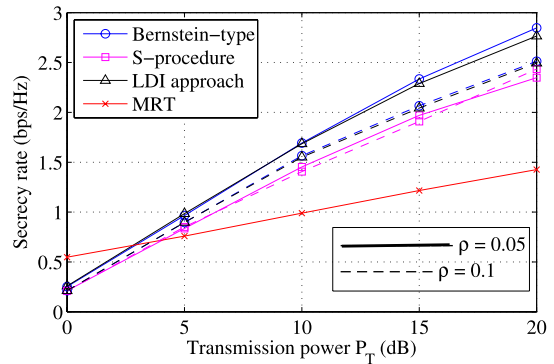
identical worst-user secrecy rate whereas the $\mathcal{S}$-procedure based approach has noticeable degradation. Since the secrecy rate $C_s$ in (6) is an increasing function of the transmit power $\mathrm{tr}(\mathbf{Q}_I)$, one can notice identical reflections in the results of Fig. 5 with increasing $P_T$. Also, it is no surprise that stricter outage constraint ($\rho = 0.05$) guarantees higher secrecy rate compared to relaxed outage constraint ($\rho = 0.1$). However, this observation does not apply to the conventional MRT scheme since it does not satisfy the outage constraints.

Finally, we analyze the achievable worst-user secrecy rate behaviors of the various methods when the number of eavesdroppers is increased. The achievable secrecy rate versus the number of eavesdroppers (i.e., $L$) is shown in Fig. 6 with $N_T = 6$, $K = 3$, $N_e = 2$, $P_T = 15$ (dB), and $\eta = -5$ (dB). It can be observed from this result that the worst-user secrecy rate decreases as more eavesdroppers are present which is reasonable since it is more likely to find an Eve with stronger channel (resulting in worse secrecy) among an increased number of Eves. As we observed in the previous examples, the BTI- and LDI-based schemes outperform the $\mathcal{S}$-Procedure based one in terms of the achievable secrecy rate.

## VI. CONCLUSIONS

This paper investigated the power minimization as well as the SRM problems with probabilistic QoS constraints in MISOME systems for SWIPT and proposed convex safe approximation based transmit beamforming algorithms with

imperfect CSI. Applying SDR techniques, we showed that rank-one optimal transmit covariance solutions are always obtainable for the safe approximation approaches. In particular, we found the maximum achievable secrecy rate under transmit power and outage constraints through optimally solving the power minimization problem. Simulation results have been provided to demonstrate the performance of the proposed approaches.

## APPENDIX

### A. Proof of Theorem 1

We start the proof by eliminating $\mathbf{Q}_\mathrm{I}$ from the constraints (24c) and (24f) such that these two constraints become irrelevant to proof of rank of $\mathbf{Q}_\mathrm{I}$. Note that the SOC program (SOCP) constraint (24c) can be equivalently expressed as

$$\sqrt{\left\| \mathbf{R}_{\mathrm{H},i}^{\frac{1}{2}} \left(\mathbf{I}_{N_{\mathrm{e},i}} \otimes \mathbf{Q}_\mathrm{I}\right) \mathbf{R}_{\mathrm{H},i}^{\frac{1}{2}} \right\|_F^2 + 2 \left\| \mathbf{R}_{\mathrm{H},i}^{\frac{1}{2}} \left(\mathbf{I}_{N_{\mathrm{e},i}} \otimes \mathbf{Q}_\mathrm{I}\right) \hat{\mathbf{h}}_{\mathrm{e},i} \right\|^2}$$

$$\leq \sqrt{\left\| \mathbf{R}_{\mathrm{H},i}^{\frac{1}{2}} \left(\mathbf{I}_{N_{\mathrm{e},i}} \otimes \mathbf{Q}_\mathrm{I}\right) \right\|_F^2 \left( \left\| \mathbf{R}_{\mathrm{H},i}^{\frac{1}{2}} \right\|_F^2 + 2 \left\| \hat{\mathbf{h}}_{\mathrm{e},i} \right\|^2 \right)}$$

$$\Longrightarrow \mathrm{tr}\left( \left(\mathbf{I}_{N_{\mathrm{e},i}} \otimes \mathbf{Q}_\mathrm{I}\right) \left(\mathbf{I}_{N_{\mathrm{e},i}} \otimes \mathbf{Q}_\mathrm{I}\right)^H \right) \leq \frac{\psi_i^2}{\alpha_i}, \quad (50)$$

where $\alpha_i \triangleq \mathrm{tr}^2\left(\mathbf{R}_{\mathrm{H},i}\right) + 2\mathrm{tr}\left(\mathbf{R}_{\mathrm{H},i}\right) \|\hat{\mathbf{h}}_{\mathrm{e},i}\|^2$.

Applying the matrix trace identities $\mathrm{tr}\left[(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D})\right] = \mathrm{tr}(\mathbf{A}\mathbf{C} \otimes \mathbf{B}\mathbf{D})$, $\mathrm{tr}(\mathbf{A} \otimes \mathbf{B}) = \mathrm{tr}(\mathbf{A})\mathrm{tr}(\mathbf{B})$ [33] and (14c), we obtain $\mathrm{tr}\left( \left(\mathbf{I}_{N_{\mathrm{e},i}} \otimes \mathbf{Q}_\mathrm{I}\right) \left(\mathbf{I}_{N_{\mathrm{e},i}} \otimes \mathbf{Q}_\mathrm{I}\right)^H \right) = \mathrm{tr}\left(\mathbf{I}_{N_{\mathrm{e},i}} \otimes \mathbf{Q}_\mathrm{I}\mathbf{Q}_\mathrm{I}^H\right) = N_{\mathrm{e},i}\mathrm{tr}\left(\mathbf{Q}_\mathrm{I}\mathbf{Q}_\mathrm{I}^H\right)$. Thus, (50) can be rewritten as

$$\mathrm{tr}\left(\mathbf{Q}_\mathrm{I}\mathbf{Q}_\mathrm{I}^H\right) \leq \frac{\psi_i^2}{\alpha_i N_{\mathrm{e},i}} \Rightarrow \mathbf{Q}_\mathrm{I}\mathbf{Q}_\mathrm{I}^H \preceq \kappa_i^2 \mathbf{I}_{N_\mathrm{T}}$$

$$\Rightarrow \begin{bmatrix} \kappa_i \mathbf{I}_{N_\mathrm{T}} & \mathbf{Q}_\mathrm{I} \\ \mathbf{Q}_\mathrm{I}^H & \kappa_i \mathbf{I}_{N_\mathrm{T}} \end{bmatrix} \succeq \mathbf{0}, \quad (51)$$

where $\kappa_i^2 \triangleq \frac{\psi_i^2}{\alpha_i N_{\mathrm{e},i}}$. Now we rewrite the constraint (51) as the following LMI:

$$\begin{bmatrix} \kappa_i \mathbf{I}_{N_\mathrm{T}} & \mathbf{0} \\ \mathbf{0} & \kappa_i \mathbf{I}_{N_\mathrm{T}} \end{bmatrix} + \begin{bmatrix} \mathbf{0} & \mathbf{Q}_\mathrm{I} \\ \mathbf{Q}_\mathrm{I}^H & \mathbf{0} \end{bmatrix} \succeq \mathbf{0}, \quad (52)$$

which can be eventually rewritten as

$$\begin{bmatrix} \kappa_i \mathbf{I}_{N_\mathrm{T}} & \mathbf{0} \\ \mathbf{0} & \kappa_i \mathbf{I}_{N_\mathrm{T}} \end{bmatrix} \succeq \begin{bmatrix} \mathbf{I}_{N_\mathrm{T}} \\ \mathbf{0} \end{bmatrix} \mathbf{Q}_\mathrm{I} \begin{bmatrix} \mathbf{0} & -\mathbf{I}_{N_\mathrm{T}} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ -\mathbf{I}_{N_\mathrm{T}} \end{bmatrix} \mathbf{Q}_\mathrm{I}^H$$
$$\times \begin{bmatrix} \mathbf{I}_{N_\mathrm{T}} & \mathbf{0} \end{bmatrix}, \ \|\mathbf{Q}_\mathrm{I}\| \leq \kappa_i. \quad (53)$$

The following lemma transforms the LMI in (53) into a more convenient form.

*Lemma 6 (Nemirovski Lemma):* [34, Lemma 2] *Given matrices* $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ *with* $\mathbf{A} = \mathbf{A}^H$,

$$\mathbf{A} \succeq \mathbf{B}^H \mathbf{X} \mathbf{C} + \mathbf{C}^H \mathbf{X}^H \mathbf{B}, \quad \forall \, \mathbf{X} : \|\mathbf{X}\| \leq \kappa,$$

*if and only if there exists a* $\beta \geq 0$ *such that*

$$\begin{bmatrix} \mathbf{A} - \beta \mathbf{C}^H \mathbf{C} & -\kappa \mathbf{B}^H \\ -\kappa \mathbf{B} & \beta \mathbf{I} \end{bmatrix} \succeq \mathbf{0}.$$

The merit of *Lemma 6* is that it transforms matrix inequalities into LMIs which do not involve the matrix variable. Based on *Lemma 6*, we obtain (54) (at the bottom of the page) from (53). Similarly, the SOCP constraint (24f) can be rewritten as (55), as shown at the bottom of this page, where $\varphi_k^2 \triangleq \frac{\nu_k^2}{\mathrm{tr}^2(\mathbf{R}_{\mathrm{g},k})+2\mathrm{tr}(\mathbf{R}_{\mathrm{g},k})\|\hat{\mathbf{g}}_k\|^2}$. Indeed, the SOCP constraints (24c) and (24f) have been rewritten without $\mathbf{Q}_\mathrm{I}$ in (54) and (55), respectively.

The Lagrangian dual function of problem (24) is given by

$$\mathcal{L}\left(\mathbf{Q}_\mathrm{I}, \mathbf{Z}, \lambda_i, \mathbf{C}_i, \mu_k, \mathbf{D}_k\right) \triangleq \mathrm{tr}\left(\mathbf{Q}_\mathrm{I}\right) - \mathrm{tr}\left(\mathbf{Z}\mathbf{Q}_\mathrm{I}\right) + \sum_{i=1}^{L} \sum_{n=1}^{N_{\mathrm{e},i}} \lambda_i$$

$$\times \mathrm{tr}\left( \mathbf{\Psi}_i^{(n,n)} \mathbf{Q}_\mathrm{I} \right) + \sum_{i=1}^{L} \sum_{n=1}^{N_{\mathrm{e},i}} \mathrm{tr}\left( \mathbf{\Phi}_i^{(n,n)} \mathbf{Q}_\mathrm{I} \right) + \sum_{i=1}^{L} \lambda_i \left( -\frac{\sigma_\mathrm{e}^2}{2^R \sigma_\mathrm{d}^2} \right)$$

$$\times \mathrm{tr}(\mathbf{h}\mathbf{h}^H \mathbf{Q}_\mathrm{I}) + \sqrt{-2\ln(p)}\psi_i - \ln(p)\omega_i - \frac{\sigma_\mathrm{e}^2}{2^R} + \sigma_\mathrm{e}^2 \Big)$$

$$- \sum_{i=1}^{L} \mathrm{tr}\left(\omega_i \mathbf{C}_i\right) - \sum_{k=1}^{K} \mu_k \mathrm{tr}\left( \mathbf{R}_{\mathrm{g},k} \mathbf{Q}_\mathrm{I} + \hat{\mathbf{g}}_k \hat{\mathbf{g}}_k^H \mathbf{Q}_\mathrm{I} \right) - \sum_{k=1}^{K} \mathrm{tr}\left(\varphi_k\right.$$

$$\times \mathbf{D}_k + \mathbf{R}_{\mathrm{g},k}^{\frac{1}{2}} \mathbf{D}_k \mathbf{R}_{\mathrm{g},k}^{\frac{1}{2}} \mathbf{Q}_\mathrm{I} \Big) + \sum_{k=1}^{K} \mu_k \left( \sqrt{-2\ln(q)}\nu_k - \ln(q)\varphi_k \right.$$

$$\left. + \frac{\eta_k}{\zeta_k} \right), \quad (56)$$

where $\mathbf{Z}, \lambda_i, \mathbf{C}_i, \mu_k, \mathbf{D}_k$ are the Lagrangian dual variables associated with $\mathbf{Q}_\mathrm{I}$, (24b), (24d), (24e), and (24g), respectively. For notational convenience, $\mathbf{\Psi}_i^{(n,n)}$ and $\mathbf{\Phi}_i^{(n,n)}$ are also defined as the diagonal block sub-matrices of $\mathbf{R}_{\mathrm{H},i} + \hat{\mathbf{h}}_{\mathrm{e},i}\hat{\mathbf{h}}_{\mathrm{e},i}^H$ and $\mathbf{R}_{\mathrm{H},i}^{\frac{1}{2}} \mathbf{C}_i \mathbf{R}_{\mathrm{H},i}^{\frac{1}{2}}$, respectively. In particular, $\mathbf{R}_{\mathrm{H},i} + \hat{\mathbf{h}}_{\mathrm{e},i}\hat{\mathbf{h}}_{\mathrm{e},i}^H = \begin{bmatrix} \mathbf{\Psi}_i^{(1,1)} & \cdots & \mathbf{\Psi}_i^{(1,N_{\mathrm{e},i})} \\ \vdots & \ddots & \vdots \\ \mathbf{\Psi}_i^{(N_{\mathrm{e},i},1)} & \cdots & \mathbf{\Psi}_i^{(N_{\mathrm{e},i},N_{\mathrm{e},i})} \end{bmatrix}$ and

$\mathbf{R}_{\mathrm{H},i}^{\frac{1}{2}} \mathbf{C}_i \mathbf{R}_{\mathrm{H},i}^{\frac{1}{2}} = \begin{bmatrix} \mathbf{\Phi}_i^{(1,1)} & \cdots & \mathbf{\Phi}_i^{(1,N_{\mathrm{e},i})} \\ \vdots & \ddots & \vdots \\ \mathbf{\Phi}_i^{(N_{\mathrm{e},i},1)} & \cdots & \mathbf{\Phi}_i^{(N_{\mathrm{e},i},N_{\mathrm{e},i})} \end{bmatrix}$. The relevant

$$\begin{bmatrix} \begin{bmatrix} \kappa_i \mathbf{I}_{N_\mathrm{T}} & \mathbf{0} \\ \mathbf{0} & \kappa_i \mathbf{I}_{N_\mathrm{T}} \end{bmatrix} - \beta_1 \begin{bmatrix} \mathbf{0} \\ -\mathbf{I}_{N_\mathrm{T}} \end{bmatrix} \begin{bmatrix} \mathbf{0} & -\mathbf{I}_{N_\mathrm{T}} \end{bmatrix} & -\kappa_i \begin{bmatrix} \mathbf{I}_{N_\mathrm{T}} \\ \mathbf{0} \end{bmatrix} \\ -\kappa_i \begin{bmatrix} \mathbf{I}_{N_\mathrm{T}} & \mathbf{0} \end{bmatrix} & \beta_1 \mathbf{I}_{N_\mathrm{T}} \end{bmatrix} \succeq \mathbf{0}. \quad (54)$$

$$\begin{bmatrix} \begin{bmatrix} \varphi_k \mathbf{I}_{N_\mathrm{T}} & \mathbf{0} \\ \mathbf{0} & \varphi_k \mathbf{I}_{N_\mathrm{T}} \end{bmatrix} - \beta_2 \begin{bmatrix} \mathbf{0} \\ -\mathbf{I}_{N_\mathrm{T}} \end{bmatrix} \begin{bmatrix} \mathbf{0} & -\mathbf{I}_{N_\mathrm{T}} \end{bmatrix} & -\varphi_k \begin{bmatrix} \mathbf{I}_{N_\mathrm{T}} \\ \mathbf{0} \end{bmatrix} \\ -\varphi_k \begin{bmatrix} \mathbf{I}_{N_\mathrm{T}} & \mathbf{0} \end{bmatrix} & \beta_2 \mathbf{I}_{N_\mathrm{T}} \end{bmatrix} \succeq \mathbf{0}. \quad (55)$$

Karush-Kuhn-Tucker (KKT) conditions can be defined as

$$\frac{\partial \mathcal{L}\left(\mathbf{Q}_{\mathrm{I}}, \mathbf{Z}, \lambda_i, \mathbf{C}_i, \mu_k, \mathbf{D}_k\right)}{\partial \mathbf{Q}_{\mathrm{I}}} = \mathbf{0}, \tag{57a}$$

$$\mathbf{Z}\mathbf{Q}_{\mathrm{I}} = \mathbf{0}, \quad \mathbf{Q}_{\mathrm{I}} \succeq \mathbf{0}, \tag{57b}$$

$$\mathbf{Z} \succeq \mathbf{0} \quad \lambda_i \geq 0, \quad \mathbf{C}_i \succeq \mathbf{0}, \forall i, \mu_k \geq 0, \mathbf{D}_k \succeq \mathbf{0}, \quad \forall k. \tag{57c}$$

Next, we prove that there exists at least one $\lambda_i > 0$ such that $t \triangleq \frac{\sigma_{\mathrm{e}}^2 \sum_{i=1}^{L} \lambda_i}{2^R \sigma_{\mathrm{d}}^2}$ is always positive definite. Denoting $\tau \triangleq \sum_{i=1}^{L} \lambda_i \left(-\frac{\sigma_{\mathrm{e}}^2}{2^R} + \sigma_{\mathrm{e}}^2\right) + \sum_{k=1}^{K} \frac{\mu_k \eta_k}{\zeta_k}$ that includes the terms in the Lagrangian dual function (56) not involving the primal variables, the dual problem of (24) is given by

$$\min_{\mathbf{Z}, \{\lambda_i\}, \{\mathbf{C}_i\}, \{\mu_k\}, \{\mathbf{D}_k\}} \sum_{i=1}^{L} \lambda_i \left(\frac{\sigma_{\mathrm{e}}^2}{2^R} - \sigma_{\mathrm{e}}^2\right) - \sum_{k=1}^{K} \frac{\mu_k \eta_k}{\zeta_k} \tag{58a}$$

$$\text{s.t.} \quad \mathbf{Z} \succeq \mathbf{0} \quad \lambda_i \geq 0, \quad \mathbf{C}_i \succeq \mathbf{0}, \forall i, \tag{58b}$$

$$\mu_k \geq 0, \mathbf{D}_k \succeq \mathbf{0}, \quad \forall k. \tag{58c}$$

Note that the primal problem (24) is convex and it can be easily verified that the problem satisfies Slater's condition [27]. Therefore, the duality gap is zero. Now, in order to successfully transfer information to the legitimate destination, the transmit power $\operatorname{tr}(\mathbf{Q}_{\mathrm{I}})$, which is the objective function of the primal problem (24), must be greater than zero. Therefore, the strict positivity $\sum_{i=1}^{L} \lambda_i \left(\frac{\sigma_{\mathrm{e}}^2}{2^R} - \sigma_{\mathrm{e}}^2\right) - \sum_{k=1}^{K} \frac{\mu_k \eta_k}{\zeta_k} > 0$ must also hold.

Let us now assume that $\lambda_i = 0, \forall i$. Then it can be easily observed that $\sum_{i=1}^{L} \lambda_i \left(\frac{\sigma_{\mathrm{e}}^2}{2^R} - \sigma_{\mathrm{e}}^2\right) - \sum_{k=1}^{K} \frac{\mu_k \eta_k}{\zeta_k} \leq 0$ for $\mu_k \geq 0$, $\forall k$, which contradicts with the already established fact. Thus we claim that there exists at least one $i$ for which $\lambda_i > 0$ and hence $t > 0$ must also hold.

Now, according to KKT condition (57a), we obtain

$$\mathbf{I}_{N_{\mathrm{T}}} + \sum_{i=1}^{L} \sum_{n=1}^{N_{\mathrm{e},i}} \lambda_i \mathbf{\Psi}_i^{(n,n)} + \sum_{i=1}^{L} \sum_{n=1}^{N_{\mathrm{e},i}} \mathbf{\Phi}_i^{(n,n)} - \sum_{i=1}^{L} \lambda_i \frac{\sigma_{\mathrm{e}}^2}{2^R \sigma_{\mathrm{d}}^2} \mathbf{h}\mathbf{h}^H$$

$$- \sum_{k=1}^{K} \mu_k \left(\mathbf{R}_{\mathrm{g},k} + \hat{\mathbf{g}}_k \hat{\mathbf{g}}_k^H\right) - \sum_{k=1}^{K} \mathbf{R}_{\mathrm{g},k}^{\frac{1}{2}} \mathbf{D}_k \mathbf{R}_{\mathrm{g},k}^{\frac{1}{2}} = \mathbf{Z}. \tag{59}$$

Let $\mathbf{X} \triangleq \mathbf{I}_{N_{\mathrm{T}}} + \sum_{i=1}^{L} \sum_{n=1}^{N_{\mathrm{e},i}} \lambda_i \mathbf{\Psi}_i^{(n,n)} + \sum_{i=1}^{L} \sum_{n=1}^{N_{\mathrm{e},i}} \mathbf{\Phi}_i^{(n,n)} - \sum_{k=1}^{K} \mu_k \left(\mathbf{R}_{\mathrm{g},k} + \hat{\mathbf{g}}_k \hat{\mathbf{g}}_k^H\right) - \sum_{k=1}^{K} \mathbf{R}_{\mathrm{g},k}^{\frac{1}{2}} \mathbf{D}_k \mathbf{R}_{\mathrm{g},k}^{\frac{1}{2}}$. Thus $\mathbf{Z} = \mathbf{X} - t\mathbf{h}\mathbf{h}^H$. From KKT condition (57b), we have $\mathbf{Z}\mathbf{Q}_{\mathrm{I}} = \mathbf{0}$. Furthermore, it can be verified that in order to meet the secrecy rate constraints, it must hold that $\mathbf{Q}_{\mathrm{I}} \neq \mathbf{0}$, or equivalently, $\operatorname{rank}(\mathbf{Q}_{\mathrm{I}}) \geq 1$. Then from (57b), it follows that $\operatorname{rank}(\mathbf{Z}) \leq N_{\mathrm{T}} - 1$.

A key step in the proof of rank-one solution with secrecy constraints is proving that the matrix $\mathbf{X}$ is positive definite. Note that for secrecy problems without the EH constraints, the last two negative terms in the expression of $\mathbf{X}$ do not appear, and hence is readily established, see e.g., in [22]. In contrast, for problems with EH constraints, this is a key challenging step which we overcome as follows.

Let $r_{\mathbf{X}} \triangleq \operatorname{rank}(\mathbf{X})$ denote the rank of $\mathbf{X}$. According to [10, Lemma 5], it holds true that $\operatorname{rank}(\mathbf{A} - \mathbf{B}) \geq \operatorname{rank}(\mathbf{A}) - \operatorname{rank}(\mathbf{B})$ for two matrices $\mathbf{A}$ and $\mathbf{B}$ of the same dimension. Thus $\operatorname{rank}(\mathbf{Z}) \geq \operatorname{rank}(\mathbf{X}) - \operatorname{rank}(t\mathbf{h}\mathbf{h}^H) = r_{\mathbf{X}} - 1$.

If $\mathbf{X}$ is positive-definite, $r_{\mathbf{X}} = N_{\mathrm{T}}$ and $\operatorname{rank}(\mathbf{Z}) \geq N_{\mathrm{T}} - 1$. However, if $\operatorname{rank}(\mathbf{Z}) = N_{\mathrm{T}}$, i.e., $\mathbf{Z}$ is of full-rank, then it follows from (57b) that $\mathbf{Q}_{\mathrm{I}} = \mathbf{0}$, which cannot be an optimal solution to (24). Therefore, we have $\operatorname{rank}(\mathbf{Z}) = N_{\mathrm{T}} - 1$. According to (57b), we have $\operatorname{rank}(\mathbf{Q}_{\mathrm{I}}) = 1$.

For the case when $r_{\mathbf{X}} < N_{\mathrm{T}}$, let $\mathbf{\Pi} = \left[\boldsymbol{\pi}_1, \boldsymbol{\pi}_2, \ldots, \boldsymbol{\pi}_{N_{\mathrm{T}} - r_{\mathbf{X}}}\right]$ with $\mathbf{\Pi}^H \mathbf{\Pi} = \mathbf{I}_{N_{\mathrm{T}} - r_{\mathbf{X}}}$ denote the orthogonal basis for the null space of $\mathbf{X}$, i.e., $\mathbf{X}\mathbf{\Pi} = \mathbf{0}$. Then we have

$$\boldsymbol{\pi}_i^H \mathbf{Z} \boldsymbol{\pi}_i = \boldsymbol{\pi}_i^H \left(\mathbf{X} - t\mathbf{h}\mathbf{h}^H\right) \boldsymbol{\pi}_i = -t|\mathbf{h}^H \boldsymbol{\pi}_i|^2 \leq 0,$$
$$\text{for } i = 1, \ldots, N_{\mathrm{T}} - r_{\mathbf{X}}. \tag{60}$$

To guarantee that the Lagrangian in (56) is bounded from below such that the dual function exists, it follows that $\mathbf{Z} \succeq \mathbf{0}$. Since $\mathbf{Z} \succeq \mathbf{0}$, it follows from (60) that $\mathbf{h}^H \boldsymbol{\pi}_i = 0, \forall i$, given $t > 0$. That is,

$$\mathbf{h}\mathbf{h}^H \mathbf{\Pi} = \mathbf{0}. \tag{61}$$

As a result, we have

$$\mathbf{Z}\mathbf{\Pi} = \mathbf{0}. \tag{62}$$

However, no information will be transferred to the IR in this case since all $\boldsymbol{\pi}_i$'s lie in the null space of $\mathbf{h}\mathbf{h}^H$ according to (61) [9, Proof of Proposition 4.1]. Hence positive secrecy rate constraint in (10b) cannot be satisfied in this case, which contradicts with the assumption $R > 0$. Thus $r_{\mathbf{X}} = N_{\mathrm{T}}$ must hold, which implies that $\mathbf{X}$ is positive definite and $\operatorname{rank}(\mathbf{Z}) \geq N_{\mathrm{T}} - 1$.

Combining this result with $\operatorname{rank}(\mathbf{Z}) \leq N_{\mathrm{T}} - 1$, it follows that $\operatorname{rank}(\mathbf{Z}) = N_{\mathrm{T}} - 1$. Accordingly, from (57b), we have

$$\operatorname{rank}(\mathbf{Q}_{\mathrm{I}}) = \operatorname{nullity}(\mathbf{Z}) = N_{\mathrm{T}} - \operatorname{rank}(\mathbf{Z}) = 1. \tag{63}$$

Theorem 1 is thus proved. ∎

### B. Proof of Theorem 2

By defining

$$\mathbf{\Sigma}_{\mathrm{H},i} \triangleq \begin{bmatrix} \mathbf{R}_{\mathrm{H},i}^{\frac{1}{2}} \left(\mathbf{I}_{N_{\mathrm{e},i}} \otimes \mathbf{Q}_{\mathrm{I}}\right) \mathbf{R}_{\mathrm{H},i}^{\frac{1}{2}} & \mathbf{R}_{\mathrm{H},i}^{\frac{1}{2}} \left(\mathbf{I}_{N_{\mathrm{e},i}} \otimes \mathbf{Q}_{\mathrm{I}}\right) \hat{\mathbf{h}}_{\mathrm{e},i} \\ \hat{\mathbf{h}}_{\mathrm{e},i}^H \left(\mathbf{I}_{N_{\mathrm{e},i}} \otimes \mathbf{Q}_{\mathrm{I}}\right) \mathbf{R}_{\mathrm{H},i}^{\frac{1}{2}} & \hat{\mathbf{h}}_{\mathrm{e},i}^H \left(\mathbf{I}_{N_{\mathrm{e},i}} \otimes \mathbf{Q}_{\mathrm{I}}\right) \hat{\mathbf{h}}_{\mathrm{e},i} \end{bmatrix},$$

$$\mathbf{\Lambda}_{\mathrm{H},i} \triangleq \begin{bmatrix} \mu_{\mathrm{H},i} \mathbf{I}_{N_{\mathrm{T}} N_{\mathrm{e},i}} & \mathbf{0} \\ \mathbf{0} & \frac{\sigma_{\mathrm{e}}^2}{2^R} - \sigma_{\mathrm{e}}^2 - \mu_{\mathrm{H},i} \gamma_{\mathrm{e}}^2 \end{bmatrix},$$

$\mathbf{\Theta}_i\left(\mathbf{Q}_{\mathrm{I}}, \mu_{\mathrm{H},i}\right)$ in (31) can be decomposed as

$$\mathbf{\Theta}_i\left(\mathbf{Q}_{\mathrm{I}}, \mu_{\mathrm{H},i}\right) = -\begin{bmatrix} \mathbf{R}_{\mathrm{H},i}^{\frac{1}{2}} & \hat{\mathbf{h}}_{\mathrm{e},i} \end{bmatrix}^H \left(\mathbf{I}_{N_{\mathrm{e},i}} \otimes \mathbf{Q}_{\mathrm{I}}\right) \begin{bmatrix} \mathbf{R}_{\mathrm{H},i}^{\frac{1}{2}} & \hat{\mathbf{h}}_{\mathrm{e},i} \end{bmatrix}$$
$$+ \mathbf{\Lambda}_{\mathrm{H},i} + \frac{\sigma_{\mathrm{e}}^2}{2^R \sigma_{\mathrm{d}}^2} [\mathbf{0} \quad \mathbf{h}]^H \mathbf{Q}_{\mathrm{I}} [\mathbf{0} \quad \mathbf{h}]. \tag{64}$$

Similarly, denoting $\mathbf{\Sigma}_{\mathrm{g},k} \triangleq \begin{bmatrix} \mathbf{R}_{\mathrm{g},k}^{\frac{1}{2}} \mathbf{Q}_{\mathrm{I}} \mathbf{R}_{\mathrm{g},k}^{\frac{1}{2}} & \mathbf{R}_{\mathrm{g},k}^{\frac{1}{2}} \mathbf{Q}_{\mathrm{I}} \hat{\mathbf{g}}_k \\ \hat{\mathbf{g}}_k^H \mathbf{Q}_{\mathrm{I}} \mathbf{R}_{\mathrm{g},k}^{\frac{1}{2}} & \hat{\mathbf{g}}_k^H \mathbf{Q}_{\mathrm{I}} \hat{\mathbf{g}}_k \end{bmatrix}$ and $\mathbf{\Lambda}_{\mathrm{g},k} \triangleq \begin{bmatrix} \mu_{\mathrm{g},k} \mathbf{I}_{N_{\mathrm{T}}} & \mathbf{0} \\ \mathbf{0} & \frac{\eta_k}{\zeta_k} - \mu_{\mathrm{g},k} \gamma_{\mathrm{e}}^2 \end{bmatrix}$, we obtain $\mathbf{\Upsilon}_k\left(\mathbf{Q}_{\mathrm{I}}, \mu_{\mathrm{g},k}\right) = \mathbf{\Sigma}_{\mathrm{g},k} + \mathbf{\Lambda}_{\mathrm{g},k} = \begin{bmatrix} \mathbf{R}_{\mathrm{g},k}^{\frac{1}{2}} & \hat{\mathbf{g}}_k \end{bmatrix}^H \mathbf{Q}_{\mathrm{I}} \begin{bmatrix} \mathbf{R}_{\mathrm{g},k}^{\frac{1}{2}} & \hat{\mathbf{g}}_k \end{bmatrix} + \mathbf{\Lambda}_{\mathrm{g},k}$ from (32).

Let us now define the Lagrangian dual function of problem (33):

$$\mathcal{L}\left(\mathbf{Q}_{\mathrm{I}}, \mathbf{Z}, \mathbf{V}_{\mathrm{H},i}, \mathbf{V}_{\mathrm{g},k}\right) \triangleq \mathrm{tr}\left(\mathbf{Q}_{\mathrm{I}}\right) - \mathrm{tr}\left(\mathbf{Z}\mathbf{Q}_{\mathrm{I}}\right)$$

$$+ \sum_{i=1}^{L} \sum_{n=1}^{N_{\mathrm{e},i}} \mathrm{tr}\left(\boldsymbol{\Omega}_i^{(n,n)} \mathbf{Q}_{\mathrm{I}}\right) - \sum_{i=1}^{L} \mathrm{tr}\left(\mathbf{V}_{\mathrm{H},i}\boldsymbol{\Lambda}_{\mathrm{H},i}\right) - \sum_{k=1}^{K} \mathrm{tr}\left(\mathbf{V}_{\mathrm{g},k}\right.$$

$$\times \boldsymbol{\Lambda}_{\mathrm{g},k}\right) - \sum_{k=1}^{K} \mathrm{tr}\left(\mathbf{V}_{\mathrm{g},k}\left[\mathbf{R}_{\mathrm{g},k}^{\frac{1}{2}} \quad \hat{\mathbf{g}}_k\right]^H \mathbf{Q}_{\mathrm{I}}\left[\mathbf{R}_{\mathrm{g},k}^{\frac{1}{2}} \quad \hat{\mathbf{g}}_k\right]\right)$$

$$- \frac{\sigma_{\mathrm{e}}^2}{2^R \sigma_{\mathrm{d}}^2} \sum_{i=1}^{L} \mathrm{tr}\left(\mathbf{V}_{\mathrm{H},i}\left[\mathbf{0} \quad \mathbf{h}\right]^H \mathbf{Q}_{\mathrm{I}}\left[\mathbf{0}\mathbf{h}\right]\right), \quad (65)$$

where $\mathbf{Z}, \mathbf{V}_{\mathrm{H},i}, \mathbf{V}_{\mathrm{g},k}$ are the Lagrangian dual variables associated with $\mathbf{Q}_{\mathrm{I}}$, (33b), and (33c), respectively. For notational convenience, $\boldsymbol{\Omega}_i^{(n,n)} \in \mathbb{H}_+^{N_{\mathrm{T}}}$ is defined as the diagonal block sub-matrices of $\left[\mathbf{R}_{\mathrm{H},i}^{\frac{1}{2}} \quad \hat{\mathbf{h}}_{\mathrm{e},i}\right] \mathbf{V}_{\mathrm{H},i}\left[\mathbf{R}_{\mathrm{H},i}^{\frac{1}{2}} \quad \hat{\mathbf{h}}_{\mathrm{e},i}\right]^H$. In particular,

$$\left[\mathbf{R}_{\mathrm{H},i}^{\frac{1}{2}} \quad \hat{\mathbf{h}}_{\mathrm{e},i}\right] \mathbf{V}_{\mathrm{H},i}\left[\mathbf{R}_{\mathrm{H},i}^{\frac{1}{2}} \quad \hat{\mathbf{h}}_{\mathrm{e},i}\right]^H$$

$$= \begin{bmatrix} \boldsymbol{\Omega}_i^{(1,1)} & \cdots & \boldsymbol{\Omega}_i^{(1,N_{\mathrm{e},i})} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\Omega}_i^{(N_{\mathrm{e},i},1)} & \cdots & \boldsymbol{\Omega}_i^{(N_{\mathrm{e},i},N_{\mathrm{e},i})} \end{bmatrix}. \quad (66)$$

The KKT conditions relevant to the proof can be defined as

$$\frac{\partial \mathcal{L}\left(\mathbf{Q}_{\mathrm{I}}, \mathbf{Z}, \mathbf{V}_{\mathrm{H},i}, \mathbf{V}_{\mathrm{g},k}\right)}{\partial \mathbf{Q}_{\mathrm{I}}} = \mathbf{0}, \quad (67a)$$

$$\mathbf{Z}\mathbf{Q}_{\mathrm{I}} = \mathbf{0}, \mathbf{Q}_{\mathrm{I}} \succeq \mathbf{0}, \quad (67b)$$

$$\mathbf{Z} \succeq \mathbf{0} \ \mathbf{V}_{\mathrm{H},i} \succeq \mathbf{0}, \quad \forall i,$$

$$\mathbf{V}_{\mathrm{g},k} \succeq \mathbf{0}, \quad \forall k. \quad (67c)$$

Next, we prove that $\mathbf{T} \triangleq \frac{\sigma_{\mathrm{e}}^2}{2^R \sigma_{\mathrm{d}}^2} \sum_{i=1}^{L} \mathbf{V}_{\mathrm{H},i}$ is always positive definite by contradiction. Denoting $\tau_{\mathrm{S}} \triangleq \sum_{i=1}^{L} \mathrm{tr}\left(\mathbf{V}_{\mathrm{H},i}\begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \frac{\sigma_{\mathrm{e}}^2}{2^R} - \sigma_{\mathrm{e}}^2 \end{bmatrix}\right) - \sum_{k=1}^{K} \mathrm{tr}\left(\mathbf{V}_{\mathrm{g},k}\begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\frac{\eta_k}{\xi_k} \end{bmatrix}\right)$ that includes the terms in the Lagrangian dual function not involving the primal variables, the dual problem of (33) is given by

$$\min_{\mathbf{Z}, \{\mathbf{V}_{\mathrm{H},i}\}, \{\mathbf{V}_{\mathrm{g},k}\}} -\tau_{\mathrm{S}} \quad (68a)$$

$$\text{s.t. } \mathbf{Z} \succeq \mathbf{0}, \ \mathbf{V}_{\mathrm{H},i} \succeq \mathbf{0}, \mathbf{V}_{\mathrm{g},k} \succeq \mathbf{0}, \quad \forall i, k. \quad (68b)$$

Note that the primal problem (33) is convex and it can be easily verified that the problem satisfies Slater's condition [27]. Thus the duality gap is zero. Now, in order to successfully transfer information to the legitimate destination, the transmit power $\mathrm{tr}\left(\mathbf{Q}_{\mathrm{I}}\right)$, which is the objective function of the primal problem (33), must be greater than zero. Therefore, the strict positivity on the objective of the dual problem (68) must also hold.

Now, if $\mathbf{V}_{\mathrm{H},i} = \mathbf{0}, \forall i$, resulting in $\mathbf{T} = \mathbf{0}$, then it can be easily observed that $\tau_{\mathrm{S}} \leq 0$ for $\mathbf{V}_{\mathrm{g},k} \succeq \mathbf{0}, \forall k$, which contradicts with the already established fact. Thus we claim that $\mathbf{T} \succ \mathbf{0}$ must also hold.

According to KKT condition (67a), we obtain

$$\mathbf{X}_{\mathrm{S}} - \left[\mathbf{0} \quad \mathbf{h}\right] \mathbf{T}\left[\mathbf{0} \quad \mathbf{h}\right]^H = \mathbf{Z}, \quad (69)$$

where $\mathbf{X}_{\mathrm{S}} \triangleq \mathbf{I}_{N_{\mathrm{T}}} + \sum_{i=1}^{L} \sum_{n=1}^{N_{\mathrm{e},i}} \boldsymbol{\Omega}_i^{(n,n)} - \sum_{k=1}^{K}\left[\mathbf{R}_{\mathrm{g},k}^{\frac{1}{2}} \quad \hat{\mathbf{g}}_k\right] \mathbf{V}_{\mathrm{g},k}\left[\mathbf{R}_{\mathrm{g},k}^{\frac{1}{2}} \quad \hat{\mathbf{g}}_k\right]^H$. From KKT condition (67b), we have $\mathbf{Z}\mathbf{Q}_{\mathrm{I}} = \mathbf{0}$. Furthermore, it can be verified that in order to meet the secrecy rate constraints, it must hold that $\mathbf{Q}_{\mathrm{I}} \neq \mathbf{0}$, or equivalently, $\mathrm{rank}\left(\mathbf{Q}_{\mathrm{I}}\right) \geq 1$. Then from (57b), it follows that $\mathrm{rank}\left(\mathbf{Z}\right) \leq N_{\mathrm{T}} - 1$.

Following similar reasoning as in Appendix A, it can be proved by contradiction that $\mathbf{X}_{\mathrm{S}} \succ \mathbf{0}$. Applying [10, Lemma 5], $\mathrm{rank}\left(\mathbf{Z}\right) \geq \mathrm{rank}\left(\mathbf{X}_{\mathrm{S}}\right) - \mathrm{rank}\left(\left[\mathbf{0} \quad \mathbf{h}\right] \mathbf{T}\left[\mathbf{0} \quad \mathbf{h}\right]^H\right) = N_{\mathrm{T}} - 1$. Combining this argument with $\mathrm{rank}\left(\mathbf{Z}\right) \leq N_{\mathrm{T}} - 1$, it follows that $\mathrm{rank}\left(\mathbf{Z}\right) = N_{\mathrm{T}} - 1$. Accordingly, from (57b), $\mathrm{rank}\left(\mathbf{Q}_{\mathrm{I}}\right) = 1$. ∎

### C. Proof of Theorem 3

Applying *Lemma 6* (Nemirovski lemma), constraints (42c), (42d), (42f), and (42g) can be equivalently transformed into LMIs not involving $\mathbf{Q}_{\mathrm{I}}$. Therefore, we ignore these constraints in the proof of rank-one $\mathbf{Q}_{\mathrm{I}}$ as in Appendix A. Accordingly, the Lagrangian of (42) is given by

$$\mathcal{L}\left(\mathbf{Q}_{\mathrm{I}}, \mathbf{Z}, \boldsymbol{\lambda}\right) \triangleq \mathrm{tr}\left(\mathbf{Q}_{\mathrm{I}}\right) - \mathrm{tr}\left(\mathbf{Z}\mathbf{Q}_{\mathrm{I}}\right) + \sum_{i=1}^{L} \sum_{n=1}^{N_{\mathrm{e},i}} \lambda_{b,i} \mathrm{tr}\left(\boldsymbol{\Psi}_i^{(n,n)} \mathbf{Q}_{\mathrm{I}}\right)$$

$$+ \sum_{i=1}^{L} \lambda_{b,i}\left(-\frac{\sigma_{\mathrm{e}}^2}{2^R \sigma_{\mathrm{d}}^2} \mathbf{h}^H \mathbf{Q}_{\mathrm{I}} \mathbf{h} - \frac{\sigma_{\mathrm{e}}^2}{2^R} + \sigma_{\mathrm{e}}^2 + 2\sqrt{-\ln(p)}\right.$$

$$\times \left(\bar{\psi}_i + \bar{\omega}_i\right)\right) - \sum_{k=1}^{K} \lambda_{e,k} \mathrm{tr}\left(\mathbf{R}_{\mathrm{g},k} \mathbf{Q}_{\mathrm{I}} + \hat{\mathbf{g}}_k \hat{\mathbf{g}}_k^H \mathbf{Q}_{\mathrm{I}}\right)$$

$$- \sum_{k=1}^{K} \lambda_{e,k}\left(-\frac{\eta_k}{\xi_k} - 2\sqrt{-\ln(q)}\left(\bar{v}_k + \bar{\varphi}_k\right)\right), \quad (70)$$

where $\boldsymbol{\lambda} \triangleq \left[\lambda_{b,1}, \ldots, \lambda_{b,L}, \lambda_{e,1}, \ldots, \lambda_{e,K}\right]^T$ is the vector of Lagrange multipliers associated with the constraints (42b) and (42e), and $\mathbf{Z}$ is the Lagrangian dual variable associated with $\mathbf{Q}_{\mathrm{I}}$. The KKT conditions can be defined as

$$\frac{\partial \mathcal{L}\left(\mathbf{Q}_{\mathrm{I}}, \mathbf{Z}, \boldsymbol{\lambda}\right)}{\partial \mathbf{Q}_{\mathrm{I}}} = \mathbf{0}, \quad (71a)$$

$$\mathbf{Z}\mathbf{Q}_{\mathrm{I}} = \mathbf{0}, \quad \mathbf{Q}_{\mathrm{I}} \succeq \mathbf{0}, \ \mathbf{Z} \succeq \mathbf{0}\boldsymbol{\lambda} \geq \mathbf{0}. \quad (71b)$$

Denoting $\tau_D \triangleq \sum_{i=1}^{L} \lambda_{b,i}\left(-\frac{\sigma_{\mathrm{e}}^2}{2^R} + \sigma_{\mathrm{e}}^2\right) + \sum_{k=1}^{K} \lambda_{e,k} \frac{\eta_k}{\xi_k}$ in (58), the dual problem can be defined as

$$\max_{\mathbf{Z}, \boldsymbol{\lambda}} -\tau_D \quad (72a)$$

$$\text{s.t. } \mathbf{Q}_{\mathrm{I}}, \ \mathbf{Z} \succeq \mathbf{0} \ \boldsymbol{\lambda} \geq \mathbf{0}. \quad (72b)$$

Following similar arguments as in Appendix A, it can be shown that there exists at least one $\lambda_{b,i} > 0$. The rest of the proof is identical to the corresponding part in Appendix A, and thus is omitted for brevity. ∎

## References

[1] I. Krikidis, S. Timotheou, S. Nikolaou, G. Zheng, D. W. K. Ng, and R. Schober, "Simultaneous wireless information and power transfer in modern communication systems," *IEEE Commun. Mag.*, vol. 52, no. 11, pp. 104–110, Nov. 2014.

[2] R. Zhang and C. K. Ho, "MIMO broadcasting for simultaneous wireless information and power transfer," *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 1989–2001, May 2013.

[3] X. Zhou, R. Zhang, and C. K. Ho, "Wireless information and power transfer: Architecture design and rate-energy tradeoff," *IEEE Trans. Commun.*, vol. 61, no. 11, pp. 4757–4767, Nov. 2013.

[4] L. Liu, R. Zhang, and K.-C. Chua, "Wireless information transfer with opportunistic energy harvesting," *IEEE Trans. Wireless Commun.*, vol. 12, no. 1, pp. 288–300, Jan. 2013.

[5] Z. Xiang and M. Tao, "Robust beamforming for wireless information and power transmission," *IEEE Wireless Commun. Lett.*, vol. 1, no. 4, pp. 372–375, Aug. 2012.

[6] M. R. A. Khandaker and K. K. Wong, "SWIPT in MISO multicasting systems," *IEEE Wireless Commun. Lett.*, vol. 3, no. 3, pp. 277–280, Jun. 2014.

[7] M. R. A. Khandaker and K. K. Wong, "QoS-based multicast beamforming for SWIPT," in *Proc. IEEE SECON Workshop Energy Harvesting Commun.*, Singapore, Jun./Jul. 2014, pp. 62–67.

[8] J. Xu, L. Liu, and R. Zhang, "Multiuser MISO beamforming for simultaneous wireless information and power transfer," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Mar. 2013, pp. 4754–4758.

[9] L. Liu, R. Zhang, and K.-C. Chua, "Secrecy wireless information and power transfer with MISO beamforming," *IEEE Trans. Signal Process.*, vol. 62, no. 7, pp. 1850–1863, Apr. 2014.

[10] M. R. A. Khandaker and K.-K. Wong, "Masked beamforming in the presence of energy-harvesting eavesdroppers," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 1, pp. 40–54, Jan. 2015.

[11] M. R. A. Khandaker and K. K. Wong, "Robust secrecy beamforming with energy-harvesting eavesdroppers," *IEEE Wireless Commun. Lett.*, vol. 4, no. 1, pp. 10–13, Feb. 2015.

[12] M. Zhang and Y. Liu, "Energy harvesting for physical-layer security in OFDMA networks," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 1, pp. 154–162, Jan. 2016.

[13] M. Zhang, Y. Liu, and R. Zhang, "Artificial noise aided secrecy information and power transfer in OFDMA systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 4, pp. 3085–3096, Apr. 2016.

[14] H. Xing, L. Liu, and R. Zhang, "Secrecy wireless information and power transfer in fading wiretap channel," *IEEE Trans. Veh. Technol.*, vol. 65, no. 1, pp. 180–190, Jan. 2016.

[15] R. Feng, Q. Li, Q. Zhang, and J. Qin, "Robust secure transmission in MISO simultaneous wireless information and power transfer system," *IEEE Trans. Veh. Technol.*, vol. 64, no. 1, pp. 400–405, Jan. 2015.

[16] S. Wang and B. Wang, "Robust secure transmit design in MIMO channels with simultaneous wireless information and power transfer," *IEEE Signal Process. Lett.*, vol. 22, no. 11, pp. 2147–2151, Nov. 2015.

[17] Z. Chu, Z. Zhu, M. Johnston, and S. L. Goff, "Simultaneous wireless information power transfer for MISO secrecy channel," *IEEE Trans. Veh. Technol.*, vol. 65, no. 9, pp. 6913–6925, Sep. 2016.

[18] A. Shojaeifard, K. A. Hamdi, E. Alsusa, D. K. C. So, and J. Tang, "Exact SINR statistics in the presence of heterogeneous interferers," *IEEE Trans. Inf. Theory*, vol. 61, no. 12, pp. 6759–6773, Dec. 2015.

[19] I. Bechar. (Sep. 2009). "A Bernstein-type inequality for stochastic processes of quadratic forms of Gaussian variables." [Online]. Available: http://arxiv.org/abs/0909.3595

[20] A. B. Tal and A. Nemirovski, "Robust solutions of linear programming problems contaminated with uncertain data," *Math. Program.*, vol. 88, pp. 411–424, Sep. 2000.

[21] K. Y. Wang, A. M. C. So, T. H. Chang, W. K. Ma, and C. Y. Chi, "Outage constrained robust transmit optimization for multiuser MISO downlinks: Tractable approximations by conic optimization," *IEEE Trans. Signal Process.*, vol. 62, no. 21, pp. 5690–5705, Nov. 2014.

[22] Z. Chu, H. Xing, M. Johnston, and S. L. Goff, "Secrecy rate optimizations for a MISO secrecy channel with multiple multiantenna eavesdroppers," *IEEE Trans. Wireless Commun.*, vol. 15, no. 1, pp. 283–297, Jan. 2016.

[23] 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.

[24] S. Goel and R. Negi, "Guaranteeing secrecy using artificial noise," *IEEE Trans. Wireless Commun.*, vol. 7, no. 6, pp. 2180–2189, Jun. 2008.

[25] Q. Li and W. K. Ma, "Optimal and robust transmit designs for miso channel secrecy by semidefinite programming," *IEEE Trans. Signal Process.*, vol. 59, no. 8, pp. 3799–3812, Aug. 2011.

[26] M. Grant and S. Boyd. (Apr. 2010). *CVX: MATLAB Software for Disciplined Convex Programming (Web Page and Software).* [Online]. Available: http://cvxr.com/cvx

[27] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[28] S. Janson, "Large deviations for sums of partly dependent random variables," *Random Struct. Algorithms*, vol. 24, no. 3, pp. 234–248, 2004.

[29] S. S. Cheung, A. M. C. So, and K. Wang, "Linear matrix inequalities with stochastically dependent perturbations and applications to chance-constrained semidefinite optimization," *SIAM J. Optim.*, vol. 22, no. 4, pp. 1394–1430, Oct. 2012.

[30] A. B. Tal and A. Nemirovski, *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*. Philadelphia, PA, USA: SIAM, 2001.

[31] M. X. Goemans and D. P. Williamson, "Approximation algorithms for MAX-3-CUT and other problems via complex semidefinite programming," *J. Comput. Syst. Sci.*, vol. 68, no. 2, pp. 442–470, Mar. 2004.

[32] K. Cumanan, Z. Ding, B. Sharif, G. Y. Tian, and K. K. Leung, "Secrecy rate optimizations for a MIMO secrecy channel with a multiple-antenna eavesdropper," *IEEE Trans. Veh. Technol.*, vol. 63, no. 4, pp. 1678–1690, May 2014.

[33] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U. K.: Cambridge Univ. Press, 1985.

[34] Y. Eldar, A. Ben-Tal, and A. Nemirovski, "Robust mean-squared error estimation in the presence of model uncertainties," *IEEE Trans. Signal Process.*, vol. 53, no. 1, pp. 168–181, Jan. 2005.

**Muhammad R. A. Khandaker** (S'10—M'13) received the B.Sc. degree (Hons.) in computer science and engineering from Jahangirnagar University, Dhaka, Bangladesh, in 2006, the M.Sc. degree in telecommunications engineering from East West University, Dhaka, in 2007, and the Ph.D. degree in electrical and computer engineering from Curtin University, Australia, in 2013. He was a Junior Hardware Design Engineer with Visual Magic Corporation Ltd., in 2005. He joined the Department of Computer Science and Engineering, IBAIS University, Dhaka, in 2006, as a Lecturer. In 2007, he joined the Department of Information and Communication Technology, Mawlana Bhasani Science and Technology University, as a Lecturer. He joined the Institute of Information Technology, Jahangirnagar University, Dhaka, in 2008, as a Lecturer. Since 2013, he has been a Postdoctoral Research Associate with the Department of Electronic and Electrical Engineering, University College London, U.K. He received the Curtin International Postgraduate Research Scholarship for his Ph.D. paper, in 2009. He also received the Best Paper Award at the 16th Asia-Pacific Conference on Communications, Auckland, New Zealand, 2010.
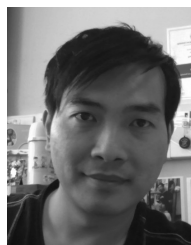
**Kai-Kit Wong** (SM'08–F'16) received the B.Eng, M.Phil., and Ph.D. degrees from the Hong Kong University of Science and Technology, Hong Kong, in 1996, 1998, and 2001, respectively, all in electrical and electronic engineering. He is currently a Professor of Wireless Communications with the Department of Electronic and Electrical Engineering, University College London, U.K. Prior to this, he took up faculty and visiting positions at the University of Hong Kong, Lucent Technologies, Bell-Labs, Holmdel, NJ, U.S., the Smart Antennas Research Group of Stanford University, and the Department of Engineering, the University of Hull, U.K. He is a Fellow of IET. He serves on the Editorial Board of the IEEE WIRELESS COMMUNICATIONS LETTERS, the IEEE COMMUNICATIONS LETTERS, and the *IEEE ComSoc/KICS Journal of Communications and Networks*. He also served as an Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS from 2005 to 2011 and the IEEE SIGNAL PROCESSING LETTERS from 2009 to 2012.

**Yangyang Zhang** received the B.S. and M.S. degrees in electronics and information engineering from Northeastern University, Shenyang, China, in 2002 and 2004 respectively, and the Ph.D. degree in electrical engineering from the University of Oxford, in 2008. He is currently with the Shenzhen Key Laboratory of Artificial Microstructure Design, Guangdong Key Laboratory of Meta-RF Microwave Radio Frequency and Kuang-Chi Institute of Advanced Technology. From 2008 to 2010, he was a Postdoctoral Research Fellow with the University College London. He is currently the Executive Vice President with the Kuang-Chi Institute of Advanced Technology, Shenzhen, China. His research interests are mainly focused on metamaterial-based future wireless communication system, such as MIMO communication system, metamaterial-based RF devices, and metamaterial-based spatial modulation technology. He received over 20 honors from various national and international competitions and published around 40 papers in various journals and conferences.

**Zhongbin Zheng** received the bachelor's and master's degrees in information and communications engineering from the Beijing University of Posts and Telecommunications in 2002 and 2005, respectively. He is currently the Vice Director of the China Academy of Information and Communications Technology, and the East China Institute of Telecommunications. He was also the Former Head of the Technology Department for the East China Institute of the Ministry of Industry and Information Technology. He is very active in research, resulting in not only a number of international paper publications, but also patents and draft standards.