Joint Robust Beamforming and Power-Splitting Ratio Design in SWIPT-based Cooperative NOMA Systems with CSI Uncertainty

Yi Yuan, Peng Xu, *Member, IEEE,* Zheng Yang, *Member, IEEE,* Zhiguo Ding, *Senior Member, IEEE,* and Qianbin Chen, *Senior Member, IEEE*

Abstract—This paper investigates the effect of imperfect channel state information on the performance of the cooperative non-orthogonal multiple access (NOMA) transmission scheme, in which the cell-center user acts as a decode-and-forward relay to assist the information transmission. The cell-center user adopts simultaneous wireless information and power transfer (SWIPT) technique to harvest energy for information forwarding. Based on two channel error models, a maximization problem is formulated to jointly design the robust beamforming and the power splitting ratio while satisfying the minimum data rate for the far user and the successfully decoding requirement for the information of the far user at the near user. For the worst-case design under the deterministic error model, successive convex approximation and the semidefinite relaxation technique are utilized to approximate the non-convex problem to an iterative convex problem. For the outage-constrained design under the stochastic error model, Bernstein-type inequality-based and Large Deviation Inequality-based approaches are used to safely approximate the probabilistic constraints of the inner-level problem to the deterministic constraints and golden section search algorithm is employed to find out the optimal single variable of the outer-level problem. Furthermore, the rank proof is provided to prove that the relaxation is tight.

Index Terms—Non-orthogonal multiple access (NOMA), simultaneous wireless information and power transfer (SWIPT), robust beamforming, successive convex approximation (SCA).

I. INTRODUCTION

RECENTLY, non-orthogonal multiple access (NOMA) has been recognized as a promising technology for the upcoming fifth generation (5G) network due to its ability to improve the spectral efficiency (SE) and to provide massive

connectivity [1]–[7]. In contrast to the conventional orthogonal multiple access (OMA) scheme, NOMA is designed to simultaneously serve two or more users at the same base station in a single orthogonal resource block (e.g., time slot, subcarrier, spreading code) based on the different power level. By utilizing successive interference cancellation (SIC) at receivers, co-channel interference can be efficiently eliminated in NOMA systems. In order to maintain user fairness, more power is allocated to the user with the poor channel condition. Hence, compared to the conventional OMA, NOMA not only can provide more efficient utilization for scarce resources [2]–[4] but also can improve user fairness [5], [6]. To further enhance the performance of NOMA technique, the multi-antenna technique [8]–[10] and deep learning technique [11] have been integrated into NOMA system.

A. Related Works and Motivation

The spectral efficiency can be significantly enhanced by using NOMA technology since all users are allowed to share same resources. However, the performance of the cell-edge users may be seriously affected by the co-exist cell-center users. In order to improve reception reliability and to guarantee the quality-of-service (QoS) of the cell-edge users, an efficient cooperative NOMA transmission scheme was proposed, by which the near users with better channel conditions are served as relays to help the far users with poor channel conditions [12]. The simulation results also indicated that cooperative NOMA transmission outperforms non-cooperative NOMA transmission in terms of outage probability and diversity gain [12] . Based on the motivation of the cooperative NOMA transmission in [12], the impact of the cooperative NOMA with full-duplex relaying was investigated in [13] and the impact of relay selection strategy for cooperative NOMA was investigated in [14], [15].

Cell-center users may face a issue of energy scarcity since they consume the energy not only to decode their information but also to help information transmission of cell-edge users. The increased energy consumption will cause harmful effects on the energy-constrained devices. Building on the advantages of SWIPT technique [16]–[19], a SWIPT-based cooperative NOMA transmission strategy was proposed to deal with the energy scarcity issue [20]. The simulation and analytical results demonstrated that the application of SWIPT in cooperative NOMA systems can significantly enhance the system performance without decreasing the diversity gain

for the cell edge users compared to the cooperative NOMA without SWIPT [20]. In order to rich the results in [20], Do *et al.* [21] proposed a best-near best-far (BNBF) user selection scheme to investigate the effects of SWIPT based on three cooperative protocols. The analytical results indicated that the diversity gain of the BNBF scheme only depends on the number of cell edge users. A novel SWIPT-based cooperative NOMA transmission was proposed to investigate the effects of beamforming on the system performance in a multi-antenna system [22].

The investigation of the cooperative NOMA scheme in [12], [14], [20]–[22] is based on the assumption that channel state information (CSI) between the transmitter and all users is perfectly known by the transmitter. However, the perfect CSI assumption might not be valid due to the significant system overhead and the existence of channel estimation errors and quantization errors. Therefore, imperfect CSI scheme has been widely investigated in the conventional NOMA systems [23]–[26]. In [23], the authors studied the outage performance of NOMA under the long- and short-term transmission cases considering one-bit feedback. Based on the deterministic channel error model, the system outage performance of NOMA was investigated in [24] and the robust beamforming design for downlink NOMA multiple-input single-output (MISO) systems was studied in [25], [26]. The simulation results in [23]–[26] demonstrated that imperfect CSI at the transmitter can significantly deteriorate the system performance.

According to the open literature, we know that the practical imperfect CSI scenario causes the deterioration of system performance. Hence, the proposed power allocation strategy and beamforming algorithm in cooperative NOMA schemes in [12], [14], [20]–[22] cannot be used in the imperfect CSI scenarios. Motivated by this practical issue, we consider to investigate the impact of imperfect CSI on the system performance of the cooperative NOMA transmission scheme. The main contributions are presented in the following subsection.

B. Contribution

In this paper, we consider a two-user cooperative NOMA transmission scheme, in which the cell-center user acts as a decode-and-forward (DF) relay to forward information of the cell-edge user with the harvested energy. The power splitting (PS) receiver is adopted at the cell-center user to implement SWIPT. The system model considered in our work pertains to that in [20]–[22], but we design the algorithms under a practical channel model. Specifically, the designed protocols and algorithms in [20]–[22] are based on the perfect CSI assumption, which may cause the performance degradation in practical scenarios. In this paper, we consider to design the robust beamforming to improve the system performance under two imperfect CSI models, namely, a deterministic error model and a stochastic error model. The designed optimization problems under the two channel error models are difficult to solve compared to the perfect CSI model. In the following, the novel contributions of this paper are summarized:

- The robust beamforming and PS ratio are jointly designed through an optimization problem under the deterministic error model. The aim of the proposed problem is to maximize the data rate of the cell-center user while satisfying the minimum data rate of the cell-edge user and the successful decoding requirement of the cell-edge user information at the cell-center user. This optimization problem is non-convex and challenging to solve due to the coupled variables, quadratic terms, and the uncertainty channel errors in the objective function and the constraints. Based on the semidefinite relaxation (SDR), successive convex approximation (SCA), and two linear matrix inequality (LMI) lemmas, the proposed problem can be approximated to a semidefinite problem (SDP) at each iteration. A novel iterative algorithm is proposed to iteratively solve the approximated problem. Furthermore, the tightness of SDR is proved by investigating the rank-one property.

- An outage-constrained optimization problem is proposed to maximize the data rate of the cell-center user while satisfying the outage constraints of the cell-edge user and the system. The main difficulty to solve the proposed problem under the stochastic error model lies on the probability constraints, which have no closed-form expressions. We decompose the proposed problem to a two-level optimization problem. First, the outer-level problem is an one-variable problem, in which the optimal variable can be obtained by using golden section search (GSS)-based algorithm. Second, the inner-level problem can be safely approximated to the deterministic problem by using two conservative methods, namely that of Bernstein-type inequality (BTI)-based and Large Deviation Inequality (LDI)-based. Moreover, the tightness of SDR used in the inner-level problem is also proved by investigating the rank-one property.

- The computational complexity of the proposed algorithms under the deterministic error model and the stochastic error model is mathematically characterized. The BTI-based approach is more complex than LDI-based approach since BTI-based approach involves more LMI constraints.

- The simulation results show that the uncertainty channel error causes the serious performance deterioration and the proposed robust beamforming algorithm can provide significant performance gains compared with the non-robust design. Moreover, the proposed outage design based on the stochastic error model can provide higher robust compared to the worst-case design based on the deterministic error model.

C. Organization and Notations

The rest of this paper is organized as follows. The system model is presented in Section II. The worst-case design based on the deterministic error model is examined in Section III. Section IV presents the outage-constrained problem design and solutions under the stochastic error model. The computational complexity analysis for the proposed algorithms is given in Section V. Simulation results and conclusions are presented in Section VI and VII, respectively.

Notation: Column vectors and matrices are denoted by boldface lower case letters and capital letters, respectively. The Hermitian conjugate transpose, transpose, vectorization, trace, and rank of a matrix $\mathbf{A}$ are denoted by $\mathbf{A}^H$, $\mathbf{A}^T$, $\text{vec}(\mathbf{A})$, $\text{Tr}(\mathbf{A})$, and $\text{rank}(\mathbf{A})$, respectively. $\mathbf{A} \succeq \mathbf{0}$ denotes a positive semidefinite matrix. $\mathbf{C}^{M \times N}$ and $\mathbb{H}_+^N$ represent a M-by-N dimensional complex matrix set and a N-by-N Hermitian positive semidefinite matrix set, respectively. The Euclidean norm of a vector and the absolute value of a complex scalar are denoted by $\|\cdot\|$ and $|\cdot|$, respectively. $\mathbb{E}[\cdot]$ is the expectation operator.
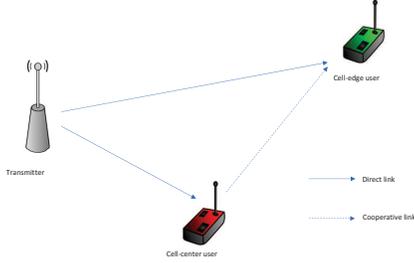
II. SYSTEM MODEL

Fig. 1: System model

A two-user cooperative NOMA transmission scheme is considered, by which the transmitter is equipped with N antennas and the two users are equipped with a single antenna, as shown in Fig. 1. In order to make the proposed designs and algorithms more suitable for practical scenarios, e.g., Internet of Things (IoT) or wireless sensor networks (WSN), we assume that both users are energy-constrained devices, which are fitted with the limited capacity battery. The battery fitted on the devices is used to cover the energy consumption of circuit operation and signal processing and needs to be replaced regularly. We define that the near user is the cell-center user and the far user is the cell-edge user. In order to improve the reliability of the far user, the near user will act as a friendly relay to forward information of the far user by using the DF protocol. The aim of the fitted battery on the device is to support circuit operation and signal processing, if the near user uses the battery energy for forwarding the information, the remained energy in the battery cannot support this device for its normal operation until the next replacement. Hence, we assume that the near user needs to use the harvested energy to forward the information at the cooperative phase in order to remain its lifetime. Each transmission block in the proposed system has two phases. Without loss of generality, it is assumed that the two phases have the same transmission periods. Each transmission period is listed below.

A. Phase 1: Direct Transmission

In the direct transmission phase, the transmitter sends the Superimposed signal $\mathbf{x} = \mathbf{w}_n s_n + \mathbf{w}_f s_f$ to the two users by utilizing NOMA technique [2], where $s_n, s_f \in \mathbf{C}^{1 \times 1}$ and $\mathbf{w}_n, \mathbf{w}_f \in \mathbf{C}^{N \times 1}$ represent the information-bearing messages delivered for the near and far users as well as the corresponding transmit beamforming vector, respectively. Without loss of generality, it is assumed that $\mathbb{E}[|s_n^2|] = \mathbb{E}[|s_f^2|] = 1$. The observation at the far user is given by

$$y_f^{(1)} = \sqrt{P_s}\mathbf{h}_f^H(\mathbf{w}_f s_f + \mathbf{w}_n s_n) + n_f^{(1)}, \qquad (1)$$

where P_s denotes the transmit power at the transmitter, $\mathbf{h}_f \in \mathbf{C}^{N \times 1}$ is the channel coefficient between the transmitter and the far user, and $n_f^{(1)} \sim \mathcal{CN}(0, \sigma_{f_1}^2)$ is the additive white Gaussian noise (AWGN) at the far user. The received signal-to-interference-plus-noise-ratio (SINR) at the far user detecting its own information is given by

$$\gamma_{f_1}^{(s_f)} = \frac{P_s|\mathbf{h}_f^H\mathbf{w}_f|^2}{P_s|\mathbf{h}_f^H\mathbf{w}_n|^2 + \sigma_{f_1}^2}. \qquad (2)$$

It is assumed that the near user is capable to recharge and store energy. Therefore, SWIPT can be implemented at the near user by using power splitting [17], [27]. Supercapacitors or short-term high-efficiency batteries can be fitted at the devices to implement this assumption [28]. Based on the power splitter at the near user, the received signal is divided into two streams. One is used for information decoding and the other is used for energy harvest. Then, the received signal at the near user for information decoding can be expressed as

$$y_n^{ID} = \sqrt{\rho}\sqrt{P_s}\mathbf{h}_n^H(\mathbf{w}_n s_n + \mathbf{w}_f s_f) + n_n, \qquad (3)$$

where $\rho \in (0,1]$ is the power splitting coefficient, $\mathbf{h}_n \in \mathbf{C}^{N \times 1}$ is the channel coefficient between the transmitter and the near user, and $n_n \sim \mathcal{CN}(0, \sigma_n^2)$ is AWGN. Based on the linear EH model assumption, the energy harvested at the near user is given by

$$E = \tau\xi(1-\rho)P_s(|\mathbf{h}_n\mathbf{w}_n|^2 + |\mathbf{h}_n\mathbf{w}_f|^2) \qquad (4)$$

where $\xi \in (0,1]$ is the energy conversion efficiency, and τ denotes the time period for the direct transmission phase. We assume the cooperative phase and the direct phase have the same transmission duration, i.e., $\tau = \frac{1}{2}$. Therefore, the maximum transmit power at the near user can be expressed as $P_t = \xi(1-\rho)P_s(|\mathbf{h}_n\mathbf{w}_n|^2 + |\mathbf{h}_n\mathbf{w}_f|^2)$. The SIC technique is applied at the near user to eliminate co-channel interference based on the concept of NOMA [2]. Particularly, the near user first decodes the information of the far user, and then decodes its own information by subtracting the information of the far user from its observation. Thus, the received SINR at the near user for decoding the information of the far user can be expressed as

$$\gamma_n^{(s_f)} = \frac{P_s\rho|\mathbf{h}_n^H\mathbf{w}_f|^2}{P_s\rho|\mathbf{h}_n^H\mathbf{w}_n|^2 + \sigma_n^2}. \qquad (5)$$

Note that the near user cannot decode its own information if it false to decode the information of the far user. After successfully decoding the information of the far user, the

corresponding signal-to-noise-ratio (SNR) at the near user is given by

$$\gamma_n^{(s_n)} = \frac{P_s \rho |\mathbf{h}_n^H \mathbf{w}_n|^2}{\sigma_n^2}. \tag{6}$$

B. Phase 2:Cooperative Transmission

In this phase, the near user forwards the information to the far user by using the harvested energy. The observation at the far user is given by

$$y_f^{(2)} = \sqrt{P_t} g s_f + n_f^{(2)}, \tag{7}$$

where g is the channel coefficient from the near user to the far user, and $n_f^{(2)} \sim \mathcal{CN}(0, \sigma_{f_2}^2)$ is the AWGN. Then, the received SNR at the far user in the cooperative phase can be expressed as

$$\gamma_2^{(s_f)} = \frac{\xi(1-\rho)P_s |g|^2 (|\mathbf{h}_n \mathbf{w}_n|^2 + |\mathbf{h}_n \mathbf{w}_f|^2)}{\sigma_{f_2}^2}. \tag{8}$$

By using maximal-ratio combining (MRC), the received final SINR at the far user is given by

$$\begin{aligned}
\gamma_{s_f} &= \gamma_1^{(s_f)} + \gamma_2^{(s_f)} \\
&= \frac{P_s |\mathbf{h}_f^H \mathbf{w}_f|^2}{P_s |\mathbf{h}_f^H \mathbf{w}_n|^2 + \sigma_{f_1}^2} + \frac{\xi(1-\rho)P_s |g|^2 (|\mathbf{h}_n \mathbf{w}_n|^2 + |\mathbf{h}_n \mathbf{w}_f|^2)}{\sigma_{f_2}^2}
\end{aligned} \tag{9}$$

The system designs in [12], [20]–[22] were based on the assumption that CSI between the BS and the users is perfectly known by the BS. However, it is not always possible for the BS to obtain perfect CSI of the users in practical scenarios due to channel estimations and quantization errors. Therefore, in this paper, we consider a more practical scenario, in which the transmitter only knows imperfect CSI of the two users. Note that this paper mainly focuses on the channel estimation errors at the multi-antenna transmitter. It is assumed that CSI between the two single-antenna users is perfectly known by the near user. Two different types of error model are adopted to describe the channel estimation errors. One is the deterministic error model, and the other is the stochastic error model. In the following parts, the optimization problems based on the two channel error models are designed.

III. WORST-CASE DESIGN BASED DETERMINISTIC ERROR MODEL

In this section, a problem of robust beamforming design is formulated based on the deterministic channel error model. The aim of the system is to maximize the data rate of the near user while satisfying the data rate requirement of the far user.

A. The Deterministic Error Model

In the deterministic uncertainty model, the estimated channel error is bounded by a constant value which is the edge of the ellipsoid. This model has been widely applied in the OMA system [29], [30] and NOMA system [25], [26]. The actual channel between the transmitter and the two users can be modelled as follows, respectively

$$\mathbf{h}_n = \hat{\mathbf{h}}_n + \mathbf{e}_n, \quad \mathbf{e}_n^H \mathbf{e}_n \leq \varepsilon_n^2, \tag{10a}$$

$$\mathbf{h}_f = \hat{\mathbf{h}}_f + \mathbf{e}_f, \quad \mathbf{e}_f^H \mathbf{e}_f \leq \varepsilon_f^2, \tag{10b}$$

where $\hat{\mathbf{h}}_n$ and $\hat{\mathbf{h}}_f$ denote the estimations of the corresponding channels, $\mathbf{e}_n$ and $\mathbf{e}_f$ represent the channel errors, ε_n and ε_f represent the size of the bounded error region.

B. Robust Beamforming Design based SDR Approach

Note that maximizing the data rate of the near user is equivalent to maximize its SINR. Hence, the proposed optimization problem can be defined as the following worst-case maximization problem based on the bounded CSI error model and on the definition of the positive semidefinite (PSD) matrices $\mathbf{Q}_n = \mathbf{w}_n \mathbf{w}_n^H$ and $\mathbf{Q}_f = \mathbf{w}_f \mathbf{w}_f^H$,

$$\text{P1}: \max_{\mathbf{Q}_n, \mathbf{Q}_f, \rho} \frac{P_s \rho (\hat{\mathbf{h}}_n + \mathbf{e}_n)^H \mathbf{Q}_n (\hat{\mathbf{h}}_n + \mathbf{e}_n)}{\sigma_n^2} \tag{11a}$$

$$\text{s.t.} \min_{\mathbf{e}_n} \frac{(\hat{\mathbf{h}}_n + \mathbf{e}_n)^H \mathbf{Q}_f (\hat{\mathbf{h}}_n + \mathbf{e}_n)}{(\hat{\mathbf{h}}_n + \mathbf{e}_n)^H \mathbf{Q}_n (\hat{\mathbf{h}}_n + \mathbf{e}_n) + \frac{\sigma_n^2}{P_s \rho}} \geq \gamma_1, \tag{11b}$$

$$\min_{\mathbf{e}_n, \mathbf{e}_f} \frac{(\hat{\mathbf{h}}_f + \mathbf{e}_f)^H \mathbf{Q}_f (\hat{\mathbf{h}}_f + \mathbf{e}_f)}{(\hat{\mathbf{h}}_f + \mathbf{e}_f)^H \mathbf{Q}_n (\hat{\mathbf{h}}_f + \mathbf{e}_f) + \frac{\sigma_{f_1}^2}{P_s}}$$

$$+ \frac{(\hat{\mathbf{h}}_n + \mathbf{e}_n)^H (\mathbf{Q}_n + \mathbf{Q}_f)(\hat{\mathbf{h}}_n + \mathbf{e}_n)}{\frac{\sigma_{f_2}^2}{\xi(1-\rho)P_s |g|^2}} \geq \gamma_1, \tag{11c}$$

$$0 \leq \text{Tr}(\mathbf{Q}_n + \mathbf{Q}_f) \leq 1, \tag{11d}$$

$$\mathbf{Q}_n \succeq \mathbf{0}, \ \mathbf{Q}_f \succeq \mathbf{0}, \tag{11e}$$

$$0 < \rho \leq 1, \tag{11f}$$

$$\text{rank}(\mathbf{Q}_n) \leq 1, \text{rank}(\mathbf{Q}_f) \leq 1, \tag{11g}$$

where $\gamma_1 = e^{2R_f} - 1$ is the minimum target SINR of the far user. Among the constraints (11b)-(11g), the constraint (11b) is imposed that the minimum SINR for successfully decoding the far user information at the near user should be greater than the threshold; (11c) is used to guarantee the minimum data rate of the far user; the power of the beamformers and the PS ratio is restricted by the constraint (11d) and (11f), respectively; the rank constraints (11g) is used to guarantee the optimality of solutions.

We remark that problem P1 is non-convex and is challenging to solve mainly because of the coupling variables and the semi-infinite programming in the objective function and constraints [31]. To circumvent these issues, an equivalent formulation is presented in the following proposition.

Proposition 1: The optimization problem P1 can be equiv-

alently expressed as

$$\text{P1-Equ}: \max_{\mathbf{Q}_n, \mathbf{Q}_f, \rho, u, x, t, v, a, b} u \tag{12a}$$

$$\text{s.t. } v^2 \geq u, \tag{12b}$$

$$t^2 \geq \gamma_1 - x, \tag{12c}$$

$$4a + (b-x)^2 \geq (b+x)^2, \tag{12d}$$

$$\text{Tr}((\hat{\mathbf{h}}_f + \mathbf{e}_f)\mathbf{Q}_f(\hat{\mathbf{h}}_f + \mathbf{e}_f)^H) \geq a, \tag{12e}$$

$$\text{Tr}((\hat{\mathbf{h}}_f + \mathbf{e}_f)\mathbf{Q}_n(\hat{\mathbf{h}}_f + \mathbf{e}_f)^H) + \frac{\sigma_{f_1}^2}{P_s} \leq b, \tag{12f}$$

$$\begin{bmatrix} \frac{P_s \rho}{\sigma_n^2} & v \\ v & (\hat{\mathbf{h}}_n + \mathbf{e}_n)^H \mathbf{Q}_n (\hat{\mathbf{h}}_n + \mathbf{e}_n) \end{bmatrix} \succeq \mathbf{0}, \tag{12g}$$

$$\begin{bmatrix} (1-\rho) & t \\ t & \frac{(\hat{\mathbf{h}}_n + \mathbf{e}_n)^H (\mathbf{Q}_n + \mathbf{Q}_f)(\hat{\mathbf{h}}_n + \mathbf{e}_n)}{\xi P_s |g|^2} \end{bmatrix} \succeq \mathbf{0}, \tag{12h}$$

$$(11b), (11d) - (11g) \tag{12i}$$

Proof: Please, refer to Appendix A. ∎

Although problem P1-Equ is still non-convex, it can be solved by applying *S-Procedure* lemma [31] to convert the non-convex constraints into linear matrix inequality (LMI), which includes finite number inequalities. They are detailed below.

Lemma 1: Define a function $f_m(\mathbf{x}) = \mathbf{z}^H \mathbf{A}_m \mathbf{z} + 2\text{Re}\{\mathbf{u}_m \mathbf{x}\} + c_m$, $m \in \{1, 2\}$, where $\mathbf{x} \in \mathbf{C}^{N \times 1}$, $\mathbf{A}_m \in \mathbb{H}^N$, $\mathbf{u}_m \in \mathbf{C}^{N \times 1}$ and $c_m \in \mathbb{R}$. Then, the implication $f_1(\mathbf{x}) \leq 0 \Rightarrow f_2(\mathbf{x}) \leq 0$ holds if and only if there exists a $\lambda \geq 0$ such that

$$\lambda \begin{bmatrix} \mathbf{A}_1 & \mathbf{u}_1 \\ \mathbf{u}_1^H & c_1 \end{bmatrix} - \begin{bmatrix} \mathbf{A}_2 & \mathbf{u}_2 \\ \mathbf{u}_2^H & c_2 \end{bmatrix} \succeq \mathbf{0} \tag{13}$$

provided that there exists a point $\tilde{\mathbf{x}}$ such that $f_m(\tilde{\mathbf{x}}) < 0$.

Based on Lemma 1, the non-convex constraints in (11b), (12e) and (12f) can be converted into the following convex LMI constraints, respectively

$$\mathbf{T}_n^f = \begin{bmatrix} \frac{\mathbf{Q}_f}{\gamma_1} - \mathbf{Q}_n + \lambda \mathbf{I} & (\frac{\mathbf{Q}_f}{\gamma_1} - \mathbf{Q}_n)\hat{\mathbf{h}}_n \\ \hat{\mathbf{h}}_n^H (\frac{\mathbf{Q}_f}{\gamma_1} - \mathbf{Q}_n) & \hat{\mathbf{h}}_n^H (\frac{\mathbf{Q}_f}{\gamma_1} - \mathbf{Q}_n)\hat{\mathbf{h}}_n - \frac{\sigma_n^2}{P_s \rho} - \lambda \varepsilon_n^2 \end{bmatrix} \succeq \mathbf{0}, \tag{14a}$$

$$\mathbf{T}_f = \begin{bmatrix} \mathbf{Q}_f + \mu \mathbf{I} & \mathbf{Q}_f \hat{\mathbf{h}}_f \\ \hat{\mathbf{h}}_f^H \mathbf{Q}_f & \hat{\mathbf{h}}_f^H \mathbf{Q}_f \hat{\mathbf{h}}_f - a - \mu \varepsilon_f^2 \end{bmatrix} \succeq \mathbf{0}, \tag{14b}$$

$$\mathbf{T}_n = \begin{bmatrix} \theta \mathbf{I} - \mathbf{Q}_n & -\mathbf{Q}_n \hat{\mathbf{h}}_f \\ -\hat{\mathbf{h}}_f^H \mathbf{Q}_n & -\theta \varepsilon_f^2 - \hat{\mathbf{h}}_f^H \mathbf{Q}_n \hat{\mathbf{h}}_f - \frac{\sigma_{f_1}^2}{P_s} + b \end{bmatrix} \succeq \mathbf{0}. \tag{14c}$$

Then, we consider to solve the more challenging non-convex constraints (12g) and (12h), in which the uncertainty error is included in the PSD matrices. Since Lemma 1 only can be used to tackle the uncertainty issue in the linear function, the following Lemma is introduced to solve the uncertainty issue in a quadratic matrix inequality (QMI).

Lemma 2: [32, Theorem 3.5] If there exists a PSD matrix $\mathbf{D} \succeq \mathbf{0}$, the following robust block QMI

$$\begin{bmatrix} \mathbf{H} & \mathbf{F} + \mathbf{GX} \\ (\mathbf{F} + \mathbf{GX})^H & \mathbf{C} + \mathbf{X}^H \mathbf{B} + \mathbf{B}^H \mathbf{X} + \mathbf{X}^H \mathbf{AX} \end{bmatrix} \succeq \mathbf{0},$$

$$\text{for all } \mathbf{I} - \mathbf{X}^H \mathbf{DX} \succeq \mathbf{0} \tag{15}$$

is equivalent to the following LMI with $t \geq 0$

$$\begin{bmatrix} \mathbf{H} & \mathbf{F} & \mathbf{G} \\ \mathbf{F}^H & \mathbf{C} & \mathbf{B}^H \\ \mathbf{G}^H & \mathbf{B} & \mathbf{A} \end{bmatrix} - t \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & -\mathbf{D} \end{bmatrix} \succeq \mathbf{0}. \tag{16}$$

To proceed, we set $\mathbf{X} = \mathbf{e}_n$, $\mathbf{D} = 1/\varepsilon_n^2 \mathbf{I}$, $\mathbf{H} = \frac{P_s \rho}{\sigma_n^2}$, $\mathbf{F} = v$, $\mathbf{G} = \mathbf{0}_{1 \times N}$, $\mathbf{C} = \hat{\mathbf{h}}_n^H \mathbf{Q}_n \hat{\mathbf{h}}_n$, $\mathbf{B} = \mathbf{Q}_n^H \hat{\mathbf{h}}_n$ and $\mathbf{A} = \mathbf{Q}_n$. By applying Lemma 2, the constraint in (12g) is converted into the following convex LMI

$$\mathbf{D}_n = \begin{bmatrix} \frac{P_s \rho}{\sigma_n^2} & v & \mathbf{0}_{1 \times N} \\ v & \hat{\mathbf{h}}_n^H \mathbf{Q}_n \hat{\mathbf{h}}_n - \alpha & \hat{\mathbf{h}}_n^H \mathbf{Q}_n \\ \mathbf{0}_{N \times 1} & \mathbf{Q}_n \hat{\mathbf{h}}_n & \mathbf{Q}_n + \frac{\alpha}{\varepsilon_n^2} \mathbf{I} \end{bmatrix} \succeq \mathbf{0}. \tag{17}$$

Similar to the constraint in (12g), the constraint in (12h) can be modified to the following LMI by using Lemma 2

$$\mathbf{D}_n^f = \begin{bmatrix} \xi(1-\rho)P_s |g|^2 & t & \mathbf{0}_{1 \times N} \\ t & \hat{\mathbf{h}}_n^H (\mathbf{Q}_n + \mathbf{Q}_f)\hat{\mathbf{h}}_n - \beta & \hat{\mathbf{h}}_n^H (\mathbf{Q}_n + \mathbf{Q}_f) \\ \mathbf{0}_{N \times 1} & (\mathbf{Q}_n + \mathbf{Q}_f)\hat{\mathbf{h}}_n & (\mathbf{Q}_n + \mathbf{Q}_f) + \frac{\beta}{\varepsilon_n^2} \mathbf{I} \end{bmatrix} \succeq \mathbf{0}. \tag{18}$$

Next, we will address the remaining non-convex constraints in (12b)-(12d). It obvious that the non-convexity of these three constraints is caused by the quadratic term on the left side of the inequality. Based on the SCA method [33], [34], we consider to propose an iterative approach to deal with the quadratic terms in (12b)-(12d). By performing the first-order Taylor approximation around the Taylor points, the constraints in (12e)-(12h) can be approximated as below

$$(v^n)^2 + 2v^n(v - v^n) \geq u, \tag{19a}$$

$$(t^n)^2 + 2t^n(t - t^n) \geq \gamma_1 - x, \tag{19b}$$

$$4a + (b^n - x^n)^2 + 2(b^n - x^n)(b - b^n - x + x^n) \geq (b+x)^2, \tag{19c}$$

where v^n, t^n, b^n and x^n are the value of variable v, t, b, x at the n-th iteration. Finally, the remaining obstacle for solving the problem P1 lies the rank constraints in (11g). By using the SDR method [35], problem P1 can be relaxed to the following problem

$$\max_{\mathbf{Q}_n, \mathbf{Q}_f, \rho, u, x, t, v, a, b, \lambda, \mu, \theta, \alpha, \beta} u \tag{20}$$

$$\text{s.t. } (11d) - (11f), (14), (17), (18), (19).$$

The objective function of the problem in (20) is a linear function. The constraints in (11d)-(11f) are affine functions. The constraints in (14),(17),(18) are LMIs, which can be expressed by several linear inequality constraints. With the given initial values v^n, t^n, b^n, x^n, the constraints in (19) are also affine functions. Hence, based on the convex theory in [31], we inform that the problem in (20) is convex at the n-th iteration and can be solved efficiently by numerical solvers such as CVX [36]. The algorithm to solve the problem in (20) is summarized in Algorithm 1. The main component of the proposed algorithm is the iterative procedure from step a) to step c). With the feasible initial values $v^{(1)}, t^{(1)}, b^{(1)}, x^{(1)}, u^{(1)}$ at the first iteration, the feasible solutions v', t', b', x', u' can be obtained by solving (20). Then, the initial values

---
Algorithm 1: The SCA-based algorithm for P1.

---
1) **Initialize:** Set $n = 1$, the tolerance $\varepsilon = 10^{-2}$, v^n, t^n, b^n, x^n, and u^n;
2) **Repeat:**
   a) Solving (20) by using CVX for the given initial values; ;
   b) Update v^n, t^n, b^n, x^n, and u^n;
   c) Update the iterative number $n = n + 1$;
3) **Until** $|u^n - u^{n-1}| \leq \varepsilon$
4) Output $\mathbf{Q}^n$ and $\mathbf{Q}^f$.

$v^{(1)}, t^{(1)}, b^{(1)}, x^{(1)}, u^{(1)}$ can be updated by the the feasible solutions v', t', b', x', u' at the end of the first iteration. This procedure will repeat until the gap of the objective value between two iterations belows the tolerance value. Since the problem in (20) is a maximization problem, Algorithm 1 can generate a non-increasing sequence of objective value. It means that the objective value of the problem in (20) converges to at least a stationary point.

Note that the solutions $\mathbf{Q}_n$ and $\mathbf{Q}_f$ to the problem in (20) are obtained by dropping the rank-one condition. If $\mathbf{Q}_n$ and $\mathbf{Q}_f$ are the rank-one constraints, the optimal beamforming $\mathbf{w}_n$ and $\mathbf{w}_f$ can be obtained by using eigenvalue decomposition (EVD) for $\mathbf{Q}_n$ and $\mathbf{Q}_f$, otherwise, Gaussian randomization will be used to generate the suboptimal solutions [35]. In order to obtain the optimal beamforming, the following theorem is presented to guarantee the tightness of the relaxation.

Theorem 1: Suppose the problem (20) is feasible with the given initial values, $\text{rank}(\mathbf{Q}_f) \leq 1$ and $\text{rank}(\mathbf{Q}_n) \leq 1$ always hold.

Proof: Please refer to Appendix B. ∎

IV. ROBUST BEAMFORMING DESIGN BASED STOCHASTIC ERROR MODEL

In this section, the robust beamforming is designed based on the stochastic error model to maximize the near user rate while satisfying the outage requirement of the far user.

A. The Stochastic Error Model

In the stochastic uncertainty model, the estimated channel error satisfies the Gaussian distribution. This model has been widely applied in the OMA system [19], [37], [38]. In this case, the actual channel between the transmitter and the two users can be respectively expressed as

$$\mathbf{h}_n = \hat{\mathbf{h}}_n + \mathbf{e}_n, \quad \mathbf{e}_n \sim \mathcal{CN}(\mathbf{0}, \mathbf{C}_n), \quad (21a)$$

$$\mathbf{h}_f = \hat{\mathbf{h}}_f + \mathbf{e}_f, \quad \mathbf{e}_f \sim \mathcal{CN}(\mathbf{0}, \mathbf{C}_f), \quad (21b)$$

where $\hat{\mathbf{h}}_n$ and $\hat{\mathbf{h}}_f$ are the estimated CSI; $\mathbf{e}_n$ and $\mathbf{e}_f$ denote the corresponding error; $\mathbf{C}_n \succeq \mathbf{0}$ and $\mathbf{C}_f \succeq \mathbf{0}$ are the associated error covariance.

B. Problem Formulation

Based on the aforementioned stochastic error model, the outage-constrained maximization problem can be reformulated as follows:

P2 :

$$\max_{\mathbf{Q}_n, \mathbf{Q}_f, \rho} \gamma_n \quad (22a)$$

$$\text{s.t. } \Pr\left\{\frac{P_s\rho(\hat{\mathbf{h}}_n+\mathbf{e}_n)^H\mathbf{Q}_n(\hat{\mathbf{h}}_n+\mathbf{e}_n)}{\sigma_n^2} \geq \gamma_n\right\} \geq 1-p, \quad (22b)$$

$$\Pr\left\{\frac{(\hat{\mathbf{h}}_n+\mathbf{e}_n)^H\mathbf{Q}_f(\hat{\mathbf{h}}_n+\mathbf{e}_n)}{(\hat{\mathbf{h}}_n+\mathbf{e}_n)^H\mathbf{Q}_n(\hat{\mathbf{h}}_n+\mathbf{e}_n)+\frac{\sigma_n^2}{P_s\rho}} \geq \gamma_1\right\} \geq 1-p_n^f, \quad (22c)$$

$$\Pr\left\{\frac{(\hat{\mathbf{h}}_f+\mathbf{e}_f)^H\mathbf{Q}_f(\hat{\mathbf{h}}_f+\mathbf{e}_f)}{(\hat{\mathbf{h}}_f+\mathbf{e}_f)^H\mathbf{Q}_n(\hat{\mathbf{h}}_f+\mathbf{e}_f)+\frac{\sigma_{f_1}^2}{P_s}} \right.$$
$$\left. + \frac{(\hat{\mathbf{h}}_n+\mathbf{e}_n)^H(\mathbf{Q}_n+\mathbf{Q}_f)(\hat{\mathbf{h}}_n+\mathbf{e}_n)}{\frac{\sigma_{f_2}^2}{\xi(1-\rho)P_s|g|^2}} \geq \gamma_1\right\} \geq 1-p, \quad (22d)$$

$$(11d) - (11g), \quad (22e)$$

where p and p_n^f denote the maximum outage probabilities. The constraints in (22b) and (22d) guarantee that both users have at least a $(1 - p)$ percent probability to reliably decode their own messages under CSI errors. (22c) ensures that there are $(1 - p_n^f)\%$ message sent to the far user can be reliably decoded by the near user under the channel error model. P2 is non-convex due to the coupled optimization variables and the channel errors [31]. In the sequel, we will introduce an approach to efficiently overcome the challenges.

C. Outer-Level Line Search Method

In order to circumvent the non-convexity of P2, we first decompose P2 to a two-level optimization problem by introducing a slack variable z. For a given z, the inner-level optimization problem can be formulated as

$$\Phi(z) = \max_{\mathbf{Q}_n, \mathbf{Q}_f, \rho} \gamma_n \quad (23a)$$

$$\text{s.t. } \Pr\left\{z + \frac{(\hat{\mathbf{h}}_n+\mathbf{e}_n)^H(\mathbf{Q}_n+\mathbf{Q}_f)(\hat{\mathbf{h}}_n+\mathbf{e}_n)}{\frac{\sigma_{f_2}^2}{\xi(1-\rho)P_s|g|^2}} \geq \gamma_1\right\} \geq 1-p, \quad (23b)$$

$$\frac{(\hat{\mathbf{h}}_f+\mathbf{e}_f)^H\mathbf{Q}_f(\hat{\mathbf{h}}_f+\mathbf{e}_f)}{(\hat{\mathbf{h}}_f+\mathbf{e}_f)^H\mathbf{Q}_n(\hat{\mathbf{h}}_f+\mathbf{e}_f)+\frac{\sigma_{f_1}^2}{P_s}} \geq z, \quad (23c)$$

$$(22b), (22c), (11d) - (11g), \quad (23d)$$

where $\Phi(z)$ is defined as the optimal value of the problem (23). The outer-level single variable optimization problem can be formulated as

$$\max_z \Phi(z)$$
$$\text{s.t. } z_{min} \leq z \leq z_{max}, \quad (24)$$

where z_{min} and z_{max} are the lower and upper bound of z, respectively. It is clearly observed that the lower bound of z can be set to zero, while the upper bound is calculated as

$$z \leq \frac{(\hat{\mathbf{h}}_f+\mathbf{e}_f)^H\mathbf{Q}_f(\hat{\mathbf{h}}_f+\mathbf{e}_f)}{(\hat{\mathbf{h}}_f+\mathbf{e}_f)^H\mathbf{Q}_n(\hat{\mathbf{h}}_f+\mathbf{e}_f)+\frac{\sigma_{f_1}^2}{P_s}}$$
$$\overset{(a)}{\leq} \frac{P_s\text{Tr}(\mathbf{Q}_f)\|\hat{\mathbf{h}}_f+\mathbf{e}_f\|^2}{\sigma_{f_1}^2} \leq \frac{P_s\text{Tr}(\hat{\mathbf{h}}_f\hat{\mathbf{h}}_f^H + \mathbf{C}_f)}{\sigma_{f_1}^2}, \quad (25)$$

where (a) follows from the fact that $\mathbf{h}_f^H \mathbf{Q}_f \mathbf{h}_f \leq \mathbf{Tr}(\mathbf{Q}_n)\|\mathbf{h}_f\|^2$ for any $\mathbf{Q}_f \succeq \mathbf{0}$ [39]. Note that z plays an important role for the power allocation. If we can obtain the value of $\Phi(z)$ by solving (23) with a given z, the optimal solutions to P_2 can be obtained by using a search method for (24). Hence, we will focus on solving the inner-level optimization problem (23) in the sequel.

D. Inner-Level Approximation Method

The main difficulty for solving the inner-level optimization problem (23) is the chance constraints in (22b), (22c), and (23b), in which the closed-form expressions are difficult to obtain. According to Challenge 1 provided in [37], an efficient way to solve the chance constraints is developing their computable upper bounds by using convex restriction approaches. Hence, we first transform the chance constraints to the following constraints by defining $\mathbf{e}_n = \mathbf{C}_n^{1/2}\mathbf{v}_n$, $\mathbf{v}_n \sim \mathcal{CN}(\mathbf{0},\mathbf{I})$

$$(22b) \Leftrightarrow \Pr\{\mathbf{v}_n^H \mathbf{C}_{\bar{n}}^{\frac{1}{2}} \mathbf{Q}_n \mathbf{C}_{\bar{n}}^{\frac{1}{2}} \mathbf{v}_n + 2\Re\{\mathbf{v}_n^H \mathbf{C}_{\bar{n}}^{\frac{1}{2}} \mathbf{Q}_n \hat{\mathbf{h}}_n\} + c_n \geq 0\}$$
$$\geq 1-p, \quad (26)$$

$$(22c) \Leftrightarrow \Pr\{\mathbf{v}_n^H \mathbf{C}_{\bar{n}}^{\frac{1}{2}}(\frac{\mathbf{Q}_f}{\gamma_1} - \mathbf{Q}_n)\mathbf{C}_{\bar{n}}^{\frac{1}{2}}\mathbf{v}_n + 2\Re\{\mathbf{v}_n^H \mathbf{C}_{\bar{n}}^{\frac{1}{2}}(\frac{\mathbf{Q}_f}{\gamma_1} - \mathbf{Q}_n)\hat{\mathbf{h}}_n\}$$
$$+ d_n \geq 0\} \geq 1 - p_n^f, \quad (27)$$

$$(23b) \Leftrightarrow \Pr\{\mathbf{v}_n^H \mathbf{C}_{\bar{n}}^{\frac{1}{2}}(\mathbf{Q}_n + \mathbf{Q}_f)\mathbf{C}_{\bar{n}}^{\frac{1}{2}}\mathbf{v}_n + 2\Re\{\mathbf{v}_n^H \mathbf{C}_{\bar{n}}^{\frac{1}{2}}(\mathbf{Q}_n + \mathbf{Q}_f)\hat{\mathbf{h}}_n\}$$
$$+ b_n \geq 0\} \geq 1 - p_f, \quad (28)$$

where $c_n = \hat{\mathbf{h}}_n^H \mathbf{Q}_n \hat{\mathbf{h}}_n - \frac{\gamma_n \sigma_n^2}{P_s \rho}$, $d_n = \hat{\mathbf{h}}_n^H(\frac{\mathbf{Q}_f}{\gamma_n} - \mathbf{Q}_n)\hat{\mathbf{h}}_n - \frac{\sigma_n^2}{P_s \rho}$, and $b_n = \hat{\mathbf{h}}_n^H(\mathbf{Q}_f + \mathbf{Q}_n)\hat{\mathbf{h}}_n - \frac{(\gamma_1 - z)\sigma_{f_2}^2}{\xi(1-\rho)P_s|g|^2}$. Based on the above transformation, we find that constraints in (26)-(28) have the same probability form as the following constraint

$$\Pr\{\mathbf{e}^H \mathbf{Q}\mathbf{e} + 2\Re\{\mathbf{e}^H \mathbf{r}\} + s \geq 0\} \geq 1 - \rho, \quad (29)$$

where $\mathbf{e} \sim \mathcal{CN}(\mathbf{0},\mathbf{I})$. The constraint (29) can be solved by developing its safe approximation upper bound $f(\mathbf{Q},\mathbf{r},s)$, which satisfies $f(\mathbf{Q},\mathbf{r},s) \leq \rho$. The safe approximation means that the obtained solution based on the constraint $f(\mathbf{Q},\mathbf{r},s) \leq \rho$ always satisfies the original constraint in (29). In the following, we attempt to develop the safe upper bound for the chance constraints in (26)-(28) by utilizing two convex restriction approaches: the BTI-based approach and the LDI-based approach.

1) BTI-Based Approach: Bernstein-type inequality is an efficient convex restriction approach to transform a chance constraint into a tractable constraint based on the following large deviation inequality for complex Gaussian quadratic functions [37].

Lemma 3: Consider the following outage constraint

$$\Pr\{\mathbf{x}^H \mathbf{A}\mathbf{x} + 2\Re\{\mathbf{x}^H \mathbf{u}\} + c \geq 0\} \geq 1 - \rho, \quad (30)$$

where $\mathbf{A} \in \mathbb{C}^{N \times N}$ is a complex Hermitian matrix, $\mathbf{u} \in \mathbb{C}^{N \times 1}$, $\mathbf{x} \sim \mathcal{CN}(\mathbf{0},\mathbf{I})$ and fixed $\rho \in (0,1]$. With any slack variables λ and μ, the outage constraint can be equivalently transformed

into the following three inequalities

$$\begin{cases} \mathbf{Tr}(\mathbf{A}) - \sqrt{-2\ln(\rho)}\lambda + \ln(\rho)\mu + c \geq 0, \\ \left\| \begin{bmatrix} \mathrm{vec}(\mathbf{A}) \\ \sqrt{2}\mathbf{u} \end{bmatrix} \right\| \leq \lambda, \\ \mu\mathbf{I}_N + \mathbf{A} \succeq \mathbf{0}, \quad \mu \geq 0. \end{cases} \quad (31)$$

Note that the three constraints in (31) are tractable deterministic constraints. If the solution satisfies the constraints in (31), it also satisfies the chance constraint in (30). Therefore, the safe upper bound of the chance constraint in (26) can be expressed as follows by using Lemma 3

$$\mathbf{Tr}(\mathbf{C}_{\bar{n}}^{\frac{1}{2}}\mathbf{Q}_n\mathbf{C}_{\bar{n}}^{\frac{1}{2}}) - \sqrt{-2\ln(p)}\lambda + \ln(p)\mu + c_n \geq 0, \quad (32a)$$

$$\left\| \begin{bmatrix} \mathrm{vec}(\mathbf{C}_{\bar{n}}^{\frac{1}{2}}\mathbf{Q}_n\mathbf{C}_{\bar{n}}^{\frac{1}{2}}) \\ \sqrt{2}\mathbf{C}_{\bar{n}}^{\frac{1}{2}}\mathbf{Q}_n\hat{\mathbf{h}}_n \end{bmatrix} \right\| \leq \lambda, \quad (32b)$$

$$\mu\mathbf{I}_N + \mathbf{C}_{\bar{n}}^{\frac{1}{2}}\mathbf{Q}_n\mathbf{C}_{\bar{n}}^{\frac{1}{2}} \succeq \mathbf{0}, \quad (32c)$$

where λ, μ are the slack variables. Since $\frac{\gamma_n}{\rho}$ included in c_n is non-convex term, we consider to approximate it to a linear function. By performing the first-order Taylor approximation around the Taylor points (γ_n^n, ρ^n), $\frac{\gamma_n}{\rho}$ can be approximated as

$$\frac{\gamma_n}{\rho} \leq \frac{\gamma_n^n}{\rho^n} - \frac{\gamma_n^n(\rho - \rho^n)}{(\rho^n)^2} = g(\gamma_n, \gamma_n^n, \rho, \rho^n). \quad (33)$$

After replacing $\frac{\gamma_n}{\rho}$ by $g(\gamma_n, \gamma_n^n, \rho, \rho^n)$, the constraint in (32a) is a linear function with the given γ_n^n and ρ^n. Next, based on Lemma 3, the chance constraints in (27) and (28) can be transformed into the following tractable deterministic constraints, respectively

$$\mathbf{Tr}(\mathbf{C}_{\bar{n}}^{\frac{1}{2}}(\frac{\mathbf{Q}_f}{\gamma_1} - \mathbf{Q}_n)\mathbf{C}_{\bar{n}}^{\frac{1}{2}}) - \sqrt{-2\ln(p_n^f)}\alpha + \ln(p_n^f)\beta + d_n \geq 0, (34a)$$

$$\left\| \begin{bmatrix} \mathrm{vec}(\mathbf{C}_{\bar{n}}^{\frac{1}{2}}(\frac{\mathbf{Q}_f}{\gamma_1} - \mathbf{Q}_n)\mathbf{C}_{\bar{n}}^{\frac{1}{2}}) \\ \sqrt{2}\mathbf{C}_{\bar{n}}^{\frac{1}{2}}(\frac{\mathbf{Q}_f}{\gamma_1} - \mathbf{Q}_n)\hat{\mathbf{h}}_n \end{bmatrix} \right\| \leq \alpha, \quad (34b)$$

$$\beta\mathbf{I}_N + \mathbf{C}_{\bar{n}}^{\frac{1}{2}}(\frac{\mathbf{Q}_f}{\gamma_n} - \mathbf{Q}_n)\mathbf{C}_{\bar{n}}^{\frac{1}{2}} \succeq \mathbf{0}, \quad (34c)$$

$$\mathbf{Tr}(\mathbf{C}_{\bar{n}}^{\frac{1}{2}}(\mathbf{Q}_f + \mathbf{Q}_n)\mathbf{C}_{\bar{n}}^{\frac{1}{2}}) - \sqrt{-2\ln(p_f)}\omega + \ln(p_f)\upsilon + b_n \geq 0, (35a)$$

$$\left\| \begin{bmatrix} \mathrm{vec}(\mathbf{C}_{\bar{n}}^{\frac{1}{2}}(\mathbf{Q}_f + \mathbf{Q}_n)\mathbf{C}_{\bar{n}}^{\frac{1}{2}}) \\ \sqrt{2}\mathbf{C}_{\bar{n}}^{\frac{1}{2}}(\mathbf{Q}_f + \mathbf{Q}_n)\hat{\mathbf{h}}_n \end{bmatrix} \right\| \leq \omega, \quad (35b)$$

$$\upsilon\mathbf{I}_N + \mathbf{C}_{\bar{n}}^{\frac{1}{2}}(\mathbf{Q}_f + \mathbf{Q}_n)\mathbf{C}_{\bar{n}}^{\frac{1}{2}} \succeq \mathbf{0}, \quad (35c)$$

where $\alpha, \beta, \omega, \upsilon$ are the slack variables. After solving (26)-(28), we will focus on the constraint in (23c), in which the channel error term cannot be solved by Lemma 3, because there is no probability form in the constraint in (23c) and the channel error is a Gaussian variable. In order to address the uncertainty error term in (23c), we consider to employ the *Sphere Bounding* method proposed in [37] to specify the deterministic region for the channel error. The spherical boundary ς_f is chosen from the following outage probability constraint $\Pr\{\|\mathbf{v}_f\| \leq \varsigma_f\} \geq 1 - p$. This constraint means that there are more than $1 - p$ portion of $\mathbf{v}_f$'s realizations lies in the sphere. Specifically, boundary ς_f can be obtained by

$\varsigma_f = \sqrt{\frac{F_{2N}^{-1}(1-p)}{2}}$, where $F_{2N}^{-1}(\cdot)$ denotes the inverse cumulative distribution function of a Chi-square random variable with $2N$ degrees of freedom. Note that all realizations of $\mathbf{v}_f$ lied in sphere satisfy the constraint in (23c). Based on Lemma 1, the constraint in (23c) can be converted into the following convex LMI with $\vartheta \geq 0$

$$\begin{bmatrix} \mathbf{C}_f^{\frac{1}{2}}(\frac{\mathbf{Q}_f}{z} - \mathbf{Q}_n)\mathbf{C}_f^{\frac{1}{2}} + \vartheta\mathbf{I} & \mathbf{C}_f^{\frac{1}{2}}(\frac{\mathbf{Q}_f}{z} - \mathbf{Q}_n)\hat{\mathbf{h}}_f \\ \hat{\mathbf{h}}_f^H(\frac{\mathbf{Q}_f}{z} - \mathbf{Q}_n)\mathbf{C}_f^{\frac{1}{2}} & \hat{\mathbf{h}}_f^H(\frac{\mathbf{Q}_f}{z} - \mathbf{Q}_n)\hat{\mathbf{h}}_f - \frac{\sigma_{f_1}^2}{P_s} - \vartheta\varsigma_f^2 \end{bmatrix} \succeq \mathbf{0}, \quad (36)$$

where $\vartheta \geq 0$.

By using the SDR method and by replacing the constraints, the inner level optimization problem can be approximated as

$$\max_{\mathbf{Q}_n, \mathbf{Q}_f, \rho, \gamma_n, \beta, \mu, v, \vartheta} \gamma_n$$
$$\text{s.t. } (32) - (36), (11d) - (11f). \quad (37)$$

With a given slack variable z, the objective function of the problem in (37) is a linear function. The constraints in (32)-(36) consist of linear constraints, SOC constraints, and LMI constraints. The constraints in (11d)-(11f) are all linear constraints. Therefore, the inner level problem in (37) is convex at the n-th iteration with the given z and can be solved efficiently by numerical solvers such as CVX. Since the solutions to the problem in (37) are obtained by dropping the rank-one constraints, the following theorem is presented to guarantee the tightness of the relaxation.

Theorem 2: Suppose that the relaxed problem (37) is feasible with a given z. The relaxed problem (37) yields an optimal solution satisfying $\text{rank}(\mathbf{Q}_n) \leq 1$ and $\text{rank}(\mathbf{Q}_f) \leq 1$.

Proof: Please refer to Appendix C. ∎

2) LDI-Based Approach: It is clearly shown that the problem (37) includes linear, second-order cone (SOC) and LMI constraints, which are polynomial-time solvable [31]. However, if the size of the LMI constraint is sufficiently large, it is time consuming to solve the problem in (37). Hence, we will use LDI-based approach, which is also developed based on the large deviation inequality for complex Gaussian quadratic functions [37], to solve the inner problem (23). The specification of LDI-based approach is illustrated in the following lemma.

Lemma 4: Fixing the Gaussian random vector $\mathbf{x} \sim \mathcal{CN}(0, \mathbf{I})$, a matrix $\mathbf{A} \in \mathbb{H}^{n \times n}$ and a vector $\mathbf{r} \in \mathbb{C}^n$, if exist any $v > \frac{1}{\sqrt{2}}$ and $\zeta > 0$, it has

$$\Pr\{\mathbf{x}^H\mathbf{A}\mathbf{x} + 2\Re\{\mathbf{x}^H\mathbf{r}\} \leq \text{tr}(\mathbf{A}) - \zeta\}$$
$$\leq \begin{cases} \exp\left(-\frac{\zeta^2}{4T^2}\right) & 0 < \zeta \leq 2\bar{v}vT \\ \exp\left(-\frac{\bar{v}v\zeta}{T} + (\bar{v}v)^2\right) & \zeta > 2\bar{v}vT \end{cases}, \quad (38)$$

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

The technical derivation of Lemma 4 refers to Appendix B in [37]. The merit of Lemma 4 is decomposing a sum of dependent random variables into two parts, each of which is a sum of certain independent random variables.

Next, we concentrate on deriving the safe upper bound of (26) based on Lemma 4 in the following. According to Lemma

4, we first set

$$a_n = \text{Tr}(\mathbf{C}_n^{\frac{1}{2}}\mathbf{Q}_n\mathbf{C}_n^{\frac{1}{2}}) + c_n, \quad (39)$$

$$T_n = v_n\|\mathbf{C}_n^{\frac{1}{2}}\mathbf{Q}_n\mathbf{C}_n^{\frac{1}{2}}\|_F + \frac{1}{\sqrt{2}}\|\mathbf{C}_n^{\frac{1}{2}}\mathbf{Q}_n\hat{\mathbf{h}}_n\|. \quad (40)$$

The value of v_a can be obtained by solving the quadratic equation $\bar{v}_n v_n = (1 - 1/(2v_n^2))v_n = \sqrt{\ln(1/p)}$. Note that $v_n > 1/\sqrt{2}$ must always exist, since equation $(1 - 1/(2v_n^2))v_n$ is a monotonically increasing function of v_a within the interval $[1/\sqrt{2}, \infty)$ and $(1 - 1/(2v_n^2))v_n = 0$ is satisfied when $v_n = 1/\sqrt{2}$. Furthermore, if $a_n = 2\sqrt{\ln(1/p)}T_n$ is chosen from the interval $[2\sqrt{\ln(1/p)}T_n, 2\bar{v}_n v_n T_n]$, $\exp(-\frac{a_n^2}{4T_n^2}) = p$, which means that the constraint (26) will be satisfied with $a_n = 2\sqrt{\ln(1/p)}T_n$. Similarly, if we choose $a_n > 2\bar{v}_n v_n T_n = 2\sqrt{(1/p)}T_n$, the constraint can be still satisfied by instituting a into (38). Based on the above analyses, we consider function $\exp(-(\text{Tr}(\mathbf{C}_n^{1/2}\mathbf{Q}_n\mathbf{C}_n^{1/2}) + c_n)/4T_n^2)$ as the safe upper bound of the chance constraint (26). Thus, the chance constraint (26) can be satisfied if the following constraint is satisfied

$$\text{Tr}(\mathbf{C}_n^{\frac{1}{2}}\mathbf{Q}_n\mathbf{C}_n^{\frac{1}{2}}) + c_n \geq 2\sqrt{\ln(1/p)}T_n. \quad (41)$$

By introducing the slack variables α and β, constraint (41) can be equivalently expressed as

$$\text{Tr}(\mathbf{C}_n^{\frac{1}{2}}\mathbf{Q}_n\mathbf{C}_n^{\frac{1}{2}}) + c_n \geq 2\sqrt{\ln(1/p)}(\alpha + \beta), \quad (42a)$$

$$v_a\|\text{vec}(\mathbf{C}_n^{\frac{1}{2}}\mathbf{Q}_n\mathbf{C}_n^{\frac{1}{2}})\| \leq \alpha, \quad (42b)$$

$$\frac{1}{\sqrt{2}}\|\mathbf{C}_n^{\frac{1}{2}}\mathbf{Q}_n\hat{\mathbf{h}}_n\| \leq \beta. \quad (42c)$$

Since the chance constraints in (26)-(28) have the same probability form, constraints (27) and (28) can be converted into the following constraints based on the same analysis, respectively

$$\text{Tr}(\mathbf{C}_n^{\frac{1}{2}}(\frac{\mathbf{Q}_f}{\gamma_1} - \mathbf{Q}_n)\mathbf{C}_n^{\frac{1}{2}}) + d_n \geq 2\sqrt{\ln(1/p_n^f)}(\lambda + \mu), \quad (43a)$$

$$v_n^f\|\text{vec}(\mathbf{C}_n^{\frac{1}{2}}(\frac{\mathbf{Q}_f}{\gamma_1} - \mathbf{Q}_n)\mathbf{C}_n^{\frac{1}{2}})\| \leq \lambda, \quad (43b)$$

$$\frac{1}{\sqrt{2}}\|\mathbf{C}_n^{\frac{1}{2}}(\frac{\mathbf{Q}_f}{\gamma_1} - \mathbf{Q}_n)\hat{\mathbf{h}}_n\| \leq \mu, \quad (43c)$$

$$\text{Tr}(\mathbf{C}_n^{\frac{1}{2}}(\mathbf{Q}_n + \mathbf{Q}_f)\mathbf{C}_n^{\frac{1}{2}}) + b_n \geq 2\sqrt{\ln(1/p_f)}(\omega + \nu), \quad (44a)$$

$$v_f\|\text{vec}(\mathbf{C}_n^{\frac{1}{2}}(\mathbf{Q}_n + \mathbf{Q}_f)\mathbf{C}_n^{\frac{1}{2}})\| \leq \omega, \quad (44b)$$

$$\frac{1}{\sqrt{2}}\|\mathbf{C}_n^{\frac{1}{2}}(\mathbf{Q}_n + \mathbf{Q}_f)\hat{\mathbf{h}}_n\| \leq \nu. \quad (44c)$$

Based on the aforementioned analyses, the inner-level problem in (23) can be equivalently reformulated to the following problem by using LDI-based approach

$$\max_{\mathbf{Q}_n, \mathbf{Q}_f, \gamma_n, \rho, \alpha, \beta, \lambda, \mu, \omega, \nu, \vartheta} \gamma_n$$
$$\text{s.t. } (36), (42) - (44), (11d) - (11f). \quad (45)$$

Similar to the problem in (37), the problem in (45) is convex at the n-th iteration with a given slack variable z. Since the optimal solution to the problem in (45) is obtained by dropping the rank-one constraints, we can use the method in Appendix

---
Algorithm 2: GSS-based algorithm for solving P2.

---
1) **Initialize:** Set $a = z_{min}$, $b = z_{max}$ and the tolerance $\epsilon = 10^{-2}$.
2) **while** $b - a > \epsilon$ **do**
3)   Calculate $c = a + 0.382(b - a)$, $d = a + 0.618(b - a)$;
4)   Obtain $\Phi(c)$ and $\Phi(d)$ by solving problem (37) (or (45) using SCA
5)   to update variables);
6)   **if** $\Phi(c) \leq \Phi(d)$ **then**
7)   Update a = c.
8)   **else**
9)   Update b = d.
10)   **endif**
11) **end while**
12) **Output:** The optimal solutions $\mathbf{Q}_n$ and $\mathbf{Q}_f$ can be obtained by solving (37) (or (45)) with the optimal $z = (a + b)/2$.

A, B to prove the optimality of the obtained solutions, and is omitted for brevity.

Compared to the problem in (37), the problem in (45) contains SOC constraints, which can be solved more efficiently than LMI constraints. This means that the LDI-based approach can achieve lower complexity compared with the BTI-based approach. More analyses for the computational complexity will be provided in the next section.

Based on the aforementioned approximation and transformation, the inner-level optimization can be converted into the convex form. Hence, the optimal solution to P_2 can be obtained by using GSS-based algorithm. The detailed process of solving P_2 is summarized in Algorithm 2.

V. COMPUTATIONAL COMPLEXITY

In this section, the computational complexities of the proposed schemes are mathematically characterized. We apply standard interior-point methods (IPM) to solve the problem in (20), (37), and (45) at each iteration step since these three problems involves LMI and SOC constraints [40, Lecture 6]. According to [37], the complexity consists of two components, namely, iteration complexity and per-iteration computation cost. Specifically, the complexity is mainly determined by the number of variables, the number and the size of LMI constraints as well as the number and the size of the SOC constraints and the size. We set k as the number of decision variables, set L^{worst}, L^{BTI} and L^{LDI} as the SCA iteration number for the worst-case method, the BTI-based approach and the LDI-based approach and set Q^{BTI}, Q^{LDI} as the line search number for the BTI-based approach and the LDI-based approach.

1) *Worst-case approach*: problem in (20) has $2N^2 + 1$ design variables and 11 slack variables, 3 LMI constraints of size $N + 1$, 2 LMI constraints of size $N + 2$, 2 LMI constraints of size N and 7 linear constraints.
2) *BTI-based approach*: problem in (37) has $2N^2 + 1$ design variables and 5 slack variables, 5 LMI constraints of size N, 1 LMI constraints of size $N + 1$, 3 SOC constraints of size $N^2 + N + 1$ and 11 linear constraints.
3) *LDI-based approach*: problem in (45) has $2N^2 + 1$ design variables and 8 slack variables, 1 LMI constraints of size $N + 1$, 2 LMI constraints of size N, 3 SOC constraints of size $N^2 + 1$, 3 SOC constraints of size $N + 1$ and 8 linear constraints.

The computational complexities of the proposed approaches are presented in Table I. As shown in the table, the computational complexity of the BTI-based approach is greater than that of the LDI-based approach, since the BTI-based approach contains too many LMI constraints.

VI. NUMERICAL RESULTS

In this section, via numerical results, we evaluate the performance of the proposed algorithms for a SWIPT-based cooperative NOMA system. Throughout the simulation, we assume that the two users are randomly deployed in a square area (10-meter $\times$ 20-meter) and the transmitter is allocated at the edge with a coordinate (0, 10-meter). The transmitter is equipped with 3 antennas. We use the Euclidean distance d and path loss exponent α to present the distance-dependent pass loss model $d^{-\alpha}$. In order to distinguish the channel conditions of the two user, we assume that $\alpha = 2$ is for the near user and $\alpha = 4$ is for the far user. For simplicity, it is assumed that the noise power spectral density is $\sigma_{f_1}^2 = \sigma_{f_2}^2 = \sigma_n^2 = -90$ dBm/Hz, the channel error covariance matrix is $\mathbf{C}_f = \mathbf{C}_n = \zeta^2 \mathbf{I}$, the outage probability $p = p_n^f = 0.01$ and the available bandwidth is 1 MHz. Without loss of generality, we use Rician fading model to generate all channels, in which the Rician factor is $K = 3$. All the simulation results are obtained by taking an average over 1000 independent channel realizations.

In order to provide meaningful insights into the design of the robust algorithm for the SWIPT-based cooperative NOMA system, we introduce the perfect CSI and non-robust schemes as two baselines. The performance of these two baselines is obtained based on the presumed CSI $\hat{\mathbf{h}}_n$ and $\hat{\mathbf{h}}_f$. Moreover, in order to provide a fair performance comparison between the deterministic error model and the stochastic error model, the radiuses of the uncertainty regions in the deterministic error model should be set as follows [37]:

$$\varepsilon_n = \varepsilon_f = \sqrt{\frac{\zeta^2 F_{2N}^{-1}(1-p)}{2}}. \qquad (46)$$

Fig. 2 displays the convergence behavior of the Algorithm 1 for three approximative approaches. As shown in figure, the proposed SCA iterative algorithms used in different cases converge after several iterations. Notice that the convergence speed of the proposed SCA algorithm depends on the initial values. If the initial values are relatively closer to the optimal solutions, the algorithm converges faster.

Fig. 3 presents the sum rate of the various designs versus the transmission power P. As shown in the figure, the performance gap between the perfect CSI case and imperfect CSI cases increases as the transmit power increases, because the transmitter needs more energy to overcome the negative effects caused by the uncertainty channel errors. An interesting phenomenon shows that the BTI- and LDI-based approaches yield almost identical the sum rate whereas the worst-case-based approach has noticeable degradation, because the worst-case approach provides more conservative design for the imperfect CSI system. In addition, the proposed NOMA designs achieve the better performance than the OMA design, which indicates that the NOMA strategy is more effective in improving the

TABLE I: Computational Complexity Analysis

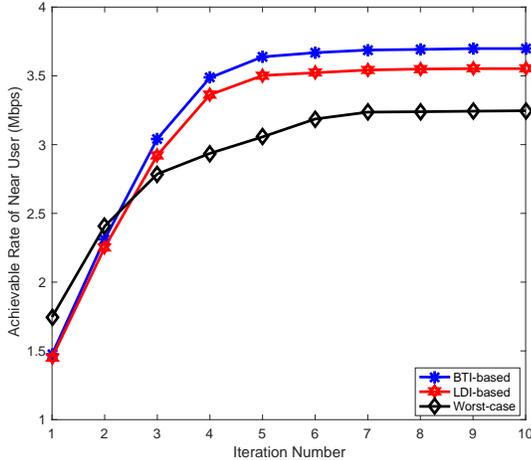| Methods	Complexity Order (Ignoring $\ln(1/\epsilon)$ in $\mathcal{O}(\cdot)$, where ϵ is accuracy).
Worst-Case	$\mathcal{O}(L^{worst} \cdot \sqrt{7(N+2)} \cdot k \cdot [3(N+1)^3 + 2(N+2)^3 + 2N^3 + k(3(N+1)^2 + 2(N+2)^2 + 2N^2) + 12(k+1) + k^2])$, where $k = \mathcal{O}(2N^2 + 12)$
BTI-based	$\mathcal{O}(Q^{BTI} \cdot L^{BTI} \cdot \sqrt{6(N+3)} \cdot k \cdot [5N^3(N+k) + (N+1)^2(N+k+1) + 3(N^2+1)^2 + 11(k+1) + k^2])$, where $k = \mathcal{O}(2N^2 + 6)$
LDI-based	$\mathcal{O}(Q^{LDI} \cdot L^{LDI} \cdot \sqrt{3(N+7)} \cdot k \cdot [(N+1)^2(N+k+1) + 2N^2(N+k) + 3((N^2+1)^2 + (N+1)^2) + 8(k+1) + k^2])$, where $k = \mathcal{O}(2N^2 + 9)$

Fig. 2: Convergence performance of the proposed SCA iterative algorithm with $P = 25$ dBm, $R_f = 1$ Mbps and $\zeta^2 = 0.01$.
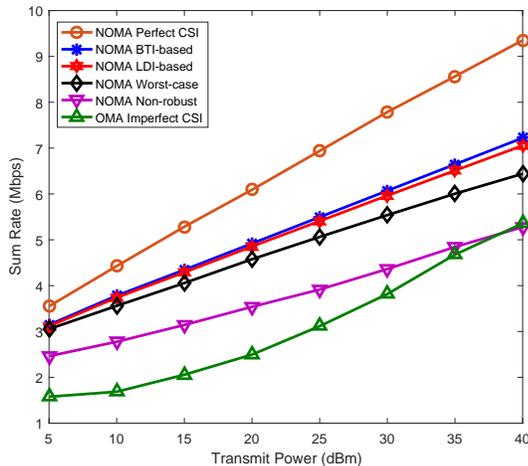

Fig. 4: The empirical CDF of sum rate with $P = 25$ dBm, $R_f = 1.5$ Mbps and $\zeta^2 = 0.01$.

Fig. 3: Sum rate performance comparison of the various methods with $R_f = 1$ Mbps and $\zeta^2 = 0.01$.

is because the non-robust method is designed based on the presumed CSI not the actual CSI. Also, there is no additional power can be used to overcome the channel uncertainty. The results in Fig. 3 and Fig. 4 indicates that the proposed robust methods can efficiently guarantee the sum rate compared to the non-robust method in the presence of imperfect CSIs. Moreover, the sum rate generated under the stochastic error model (BTI- and LDI-based approach) is higher than that under the deterministic error model (worst-case approach).

system SE under the imperfect CSI case. Besides, we observe a peculiar behavior that the non-robust scheme undergoes severe performance degradation compared with the proposed robust schemes, which means that the non-robust deign is sensitive to CSI uncertainties. To better understand the impact of imperfect CSI on the non-robust design, we show the empirical cumulative density function (CDF) of the achieved sum rate and set the target rate of the far user R_f to be 1.5 Mbps in Fig. 4. As shown in the figure approximately 7% of the sum rate is below 2 Mbps for the proposed robust methods whereas approximately 30% is for the non-robust method. This
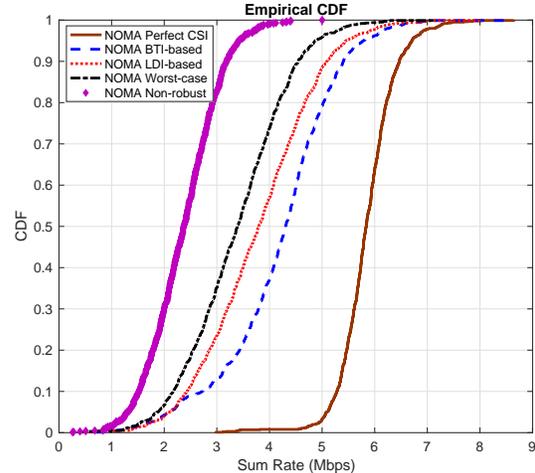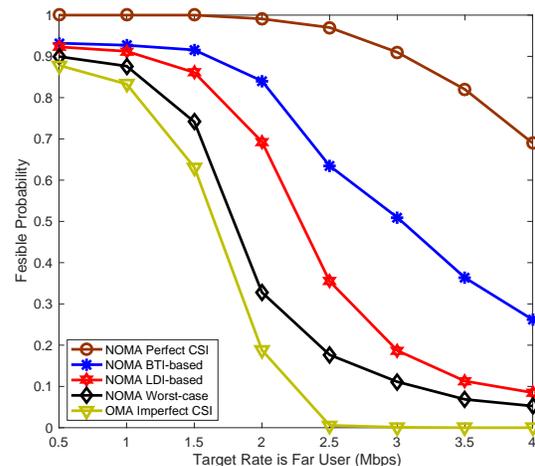

Fig. 5: Feasibility performance of the various methods with $P = 25$ dBm and $\zeta^2 = 0.01$.

Fig. 5 plots the feasible probability versus the target rate of the far user with various methods. One can observe that the

feasible probability of all schemes decreases as the target rate of the far user increases. Remarkably, the feasible probability of the perfect CSI scheme can achieve about 70% when the target rate of the far user is 4 Mbps, whereas the best feasible probability of the imperfect CSI schemes only can achieve about 28%. In particular, the BTI-based method has the best feasible probability performance followed by the LDI-based method. The worst-case method under the deterministic error model has the worst feasible probability performance. From another perspective, with the same feasible probability, the proposed NOMA strategy can support the higher data rate for the far user under the imperfect CSI case compared to the OMA strategy, which indicates that the proposed NOMA strategy can significantly mitigate the negative effects caused by the uncertainty channel.
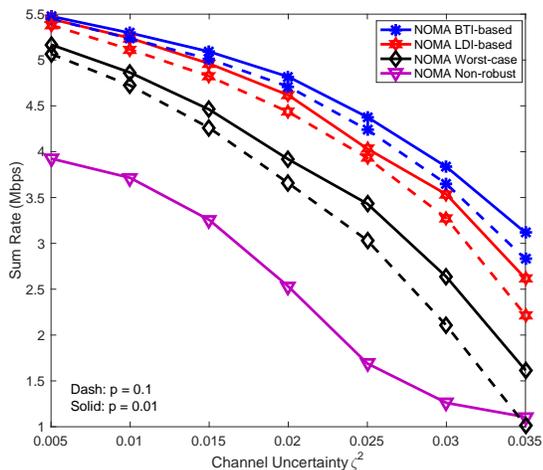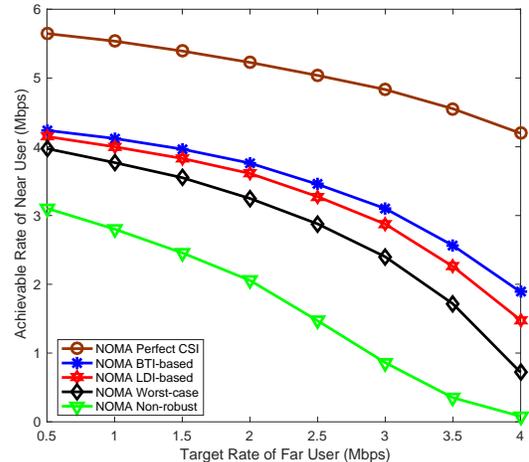

Fig. 7: Achievable rates tradeoff between the near user and far user with $P = 25$ dBm and $\zeta^2 = 0.01$.

low data requirements of the far user.

Fig. 6: Achievable rate performance of the near user versus channel uncertainty ζ^2 with $R_f = 1$ Mbps and $P = 25$ dBm.

Fig. 6 investigates the impact of channel uncertainties on the system performance with different outage tolerance p. As the figure shows, the sum rate of the system becomes worse when the channel uncertainty range increases, because the portion power used for information transmission is allocated to overcome the impacts of the channel uncertainties. Remarkably, the BTI- and LDI-based approaches can efficiently improve the system performance compared to the worst-case method in the presence of imperfect CSI. Furthermore, we find that the stricter outage tolerance ($p = 0.01$) generates the higher sum rate compared to the relaxed outage tolerance ($p = 0.1$).

Fig. 7 displays the rate tradeoff between the near user and the far user with the different approaches. It is observed that the achievable data rate of the near user decreases as the data rate requirements for the far user for all schemes increase, because the transmitter needs to allocate more power to satisfy the data rate requirement of the far user. This figure also proves that the non-robust method is more sensitive to uncertainty channels. Moreover, the performance gap between the three approaches under the imperfect CSI case becomes slightly bigger when the target rate of the far user increase. The BTI-based approach can provide more reliable transmission compared to other two approaches no matter for the high or
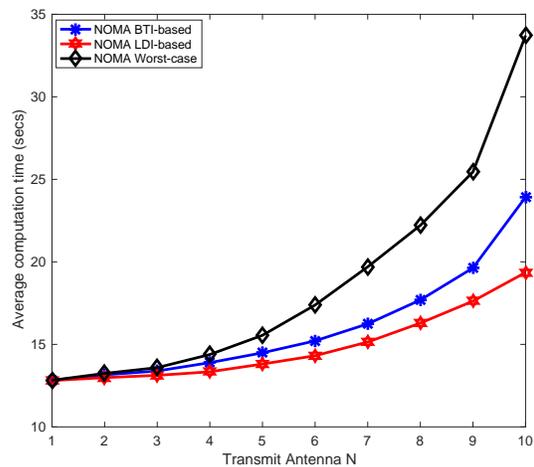

Fig. 8: Average runtimes of the various methods.

The average computation time of the various methods under the imperfect CSI case is shown in Fig. 8. The analysis of the computational complexity for the proposed approaches has been provided in section V. The results are obtained by using a desktop PC with 64-bit Intel(R) Core (TM) i7-7700HQ and 8 GB RAM. From the figure, we see that the computation time significantly increases with increasing the number of the transmit antenna. When the transmit antenna is greater than 4, the computation time of the worst-case method has a faster increase compared to the other approaches, and the LDI-based approach has the lowest computation time.

VII. CONCLUSION

In this paper, we have proposed a cooperative SWIPT-aided NOMA transmission scheme to design an efficient robust beamforming under two types of channel error models, namely, the deterministic error model and the stochastic

error model. The worst-case data rate maximization problem was proposed under the deterministic error model. By using SDR and SCA, the non-convex maximization problem can be iteratively solved. An outage-constrained maximization problem was proposed to design the robust beamforming under the stochastic error model. A two-level optimization algorithm was introduced to solve the problem. For the inner-level problem, BTI- and LDI-based approaches have been used to convert the chance constraints into the trackable deterministic constraints. The SCA method has been used to update the variables. The GSS-based algorithm was proposed to find the optimal single variable in the outer-level problem. Furthermore, a tightness analysis for the relaxation used in the two error models has been provided to prove the rank-one optimality. The computational complexity analyses of proposed approaches have also been provided. Finally, the simulation results show that the proposed robust algorithm can efficiently reduce the impacts of imperfect CSIs compared to the non-robust algorithm. Moreover, the stochastic error model has been shown more suitable for the imperfect CSI design compared to the deterministic error model.

APPENDIX A
PROOF OF PROPOSITION 1

By introducing several auxiliary variables (u, x, t, v, a, b), the problem P1 can be rewritten as

$$\max_{\mathbf{Q}_n, \mathbf{Q}_f, \rho, u, x, t, v, a, b} u \tag{47a}$$

$$\text{s.t. } v^2 \geq u, \tag{47b}$$

$$t^2 \geq \gamma_1 - x, \tag{47c}$$

$$4a + (b - x)^2 \geq (b + x)^2, \tag{47d}$$

$$\text{Tr}((\hat{\mathbf{h}}_f + \mathbf{e}_f)\mathbf{Q}_f(\hat{\mathbf{h}}_f + \mathbf{e}_f)^H) \geq a, \tag{47e}$$

$$\text{Tr}((\hat{\mathbf{h}}_f + \mathbf{e}_f)\mathbf{Q}_n(\hat{\mathbf{h}}_f + \mathbf{e}_f)^H) + \frac{\sigma_{f_1}^2}{P_s} \leq b, \tag{47f}$$

$$\frac{P_s\rho(\hat{\mathbf{h}}_n + \mathbf{e}_n)^H \mathbf{Q}_n(\hat{\mathbf{h}}_n + \mathbf{e}_n)}{\sigma_n^2} \geq v^2, \tag{47g}$$

$$(1-\rho)\xi P_s |g|^2 (\hat{\mathbf{h}}_n + \mathbf{e}_n)^H (\mathbf{Q}_n + \mathbf{Q}_f)(\hat{\mathbf{h}}_n + \mathbf{e}_n) \geq t^2, \tag{47h}$$

$$C3 - C6. \tag{47i}$$

We introduce the following lemma to convert the constraints in (47g) and (47h) into a quadratic matrix inequality (QMI).

Lemma 5: [31, Schur complement] Let $\mathbf{N}$ be a complex Hermitian matrix as

$$\mathbf{N} = \mathbf{N}^H = \begin{bmatrix} \mathbf{U}_1 & \mathbf{U}_2 \\ \mathbf{U}_2^H & \mathbf{U}_3 \end{bmatrix}. \tag{48}$$

$\mathbf{N} \succ \mathbf{0}$ is satisfied if and only if $\mathbf{U}_1 - \mathbf{U}_2^H \mathbf{U}_3^{-1} \mathbf{U}_2 \succeq \mathbf{0}$ with $\mathbf{U}_3 \succ \mathbf{0}$, or $\mathbf{U}_3 - \mathbf{U}_2^H \mathbf{U}_1^{-1} \mathbf{U}_2 \succeq \mathbf{0}$ with $\mathbf{U}_1 \succ \mathbf{0}$.
Based on Lemma (5), (47g) and (47h) can be expressed as

$$\begin{bmatrix} \frac{P_s\rho}{\sigma_n^2} & v \\ v & (\hat{\mathbf{h}}_n + \mathbf{e}_n)^H \mathbf{Q}_n(\hat{\mathbf{h}}_n + \mathbf{e}_n) \end{bmatrix} \succeq \mathbf{0}, \tag{49}$$

$$\begin{bmatrix} (1-\rho) & t \\ t & \frac{(\hat{\mathbf{h}}_n + \mathbf{e}_n)^H (\mathbf{Q}_n + \mathbf{Q}_f)(\hat{\mathbf{h}}_n + \mathbf{e}_n)}{\xi P_s |g|^2} \end{bmatrix} \succeq \mathbf{0}. \tag{50}$$

APPENDIX B
PROOF OF THEOREM 1

For ease of expression, we first rewrite $\mathbf{T}_n^f$, $\mathbf{T}_f$, $\mathbf{T}_n$, $\mathbf{D}_n$ and $\mathbf{D}_n^f$ in (14),(17),(18) as

$$\mathbf{T}_n^f = \mathbf{\Gamma}_n^f + \mathbf{H}_n^H \frac{\mathbf{Q}_f}{\gamma_1} \mathbf{H}_n - \mathbf{H}_n^H \mathbf{Q}_n \mathbf{H}_n, \tag{51}$$

$$\mathbf{T}_f = \mathbf{\Gamma}_f + \mathbf{H}_f^H \mathbf{Q}_f \mathbf{H}_f, \tag{52}$$

$$\mathbf{T}_n = \mathbf{\Gamma}_n - \mathbf{H}_f^H \mathbf{Q}_n \mathbf{H}_f, \tag{53}$$

$$\mathbf{D}_n = \mathbf{\Theta}_n + \overline{\mathbf{H}}_n^H \mathbf{Q}_n \overline{\mathbf{H}}_n, \tag{54}$$

$$\mathbf{D}_n^f = \mathbf{\Theta}_n^f + \overline{\mathbf{H}}_n^H \mathbf{Q}_n \overline{\mathbf{H}}_n + \overline{\mathbf{H}}_n^H \mathbf{Q}_f \overline{\mathbf{H}}_n, \tag{55}$$

$$\tag{56}$$

where $\mathbf{\Gamma}_n^f = \begin{bmatrix} \lambda\mathbf{I} & \mathbf{0} \\ \mathbf{0} & -\frac{\sigma_n^2}{P_s\rho} - \lambda\varepsilon_n^2 \end{bmatrix}$, $\mathbf{H}_n = \begin{bmatrix} \mathbf{I}_N & \hat{\mathbf{h}}_n \end{bmatrix}$,

$\mathbf{\Gamma}_f = \begin{bmatrix} \mu\lambda\mathbf{I} & \mathbf{0} \\ \mathbf{0} & -a - \mu\varepsilon_f^2 \end{bmatrix}$, $\mathbf{\Gamma}_n = \begin{bmatrix} \theta\lambda\mathbf{I} & \mathbf{0} \\ \mathbf{0} & b - \theta\varepsilon_f^2 - \frac{\sigma_{f_1}^2}{P_s} \end{bmatrix}$.

$\mathbf{\Theta}_n = \begin{bmatrix} \frac{P_s\rho}{\sigma_n^2} & v & \mathbf{0}_{1 \times N} \\ v & -\alpha & \mathbf{0}_{1 \times N} \\ \mathbf{0}_{N \times 1} & \mathbf{0}_{N \times 1} & \frac{\alpha}{\varepsilon_n^2}\mathbf{I} \end{bmatrix}$, $\mathbf{H}_f = \begin{bmatrix} \mathbf{I}_N & \hat{\mathbf{h}}_f \end{bmatrix}$, $\mathbf{\Theta}_n^f =$

$\begin{bmatrix} \xi(1-\rho)P_s|g|^2 & t & \mathbf{0}_{1 \times N} \\ t & -\beta & \mathbf{0}_{1 \times N} \\ \mathbf{0}_{N \times 1} & \mathbf{0}_{N \times 1} & \frac{\beta}{\varepsilon_n^2}\mathbf{I} \end{bmatrix}$, $\overline{\mathbf{H}}_n = \begin{bmatrix} \mathbf{0}_{N \times 1} & \hat{\mathbf{h}}_n & \mathbf{I}_N \end{bmatrix}$.

We denote $\mathbf{\Xi}$ as a collection of all the dual and primal variables related problem (20). Then, the Lagrangian function of the primal problem (20) is given by

$$\mathcal{L}(\mathbf{\Xi}) = \text{Tr}(\mathbf{A}_n^f \mathbf{\Gamma}_n^f) + \frac{\text{Tr}(\mathbf{A}_n^f \mathbf{H}_n^H \mathbf{Q}_f \mathbf{H}_n)}{\gamma_1} - \text{Tr}(\mathbf{A}_n^f \mathbf{H}_n^H \mathbf{Q}_n \mathbf{H}_n)$$
$$+ \text{Tr}(\mathbf{A}_f \mathbf{\Gamma}_f) + \text{Tr}(\mathbf{A}_f \mathbf{H}_f^H \mathbf{Q}_f \mathbf{H}_f) + \text{Tr}(\mathbf{A}_n \mathbf{\Gamma}_n)$$
$$- \text{Tr}(\mathbf{A}_n \mathbf{H}_f^H \mathbf{Q}_n \mathbf{H}_f) + \text{Tr}(\mathbf{B}_n \mathbf{\Theta}_n) + \text{Tr}(\mathbf{B}_n \overline{\mathbf{H}}_n^H \mathbf{Q}_n \overline{\mathbf{H}}_n)$$
$$+ \text{Tr}(\mathbf{B}_n^f \mathbf{\Theta}_n^f) + \text{Tr}(\mathbf{B}_n^f \overline{\mathbf{H}}_n^H \mathbf{Q}_n \overline{\mathbf{H}}_n) + \text{Tr}(\mathbf{B}_n^f \overline{\mathbf{H}}_n^H \mathbf{Q}_f \overline{\mathbf{H}}_n)$$
$$+ \text{Tr}(\mathbf{C}_n \mathbf{Q}_n) + \text{Tr}(\mathbf{C}_f \mathbf{Q}_f) - \delta(\text{Tr}(\mathbf{Q}_n + \mathbf{Q}_f) - 1) + \Lambda, \tag{57}$$

where Λ denotes the collections of variables not related to $\mathbf{Q}_n$ and $\mathbf{Q}_f$. $\mathbf{A}_n^f \in \mathbb{H}_+^{N+1}$, $\mathbf{A}_f \in \mathbb{H}_+^{N+1}$, $\mathbf{A}_n \in \mathbb{H}_+^{N+1}$, $\mathbf{B}_n \in \mathbb{H}_+^{N+2}$, $\mathbf{B}_n^f \in \mathbb{H}_+^{N+2}$, $\mathbf{C}_n \in \mathbb{H}_+^N$, $\mathbf{C}_f \in \mathbb{H}_+^N$, and $\delta \in \mathbb{R}_+$ are dual variables associated with $\mathbf{\Gamma}_n^f \succeq \mathbf{0}$, $\mathbf{\Gamma}_f \succeq \mathbf{0}$, $\mathbf{\Gamma}_n \succeq \mathbf{0}$, $\mathbf{D}_n \succeq \mathbf{0}$, $\mathbf{D}_n^f \succeq \mathbf{0}$, $\mathbf{Q}_n \succeq \mathbf{0}$, $\mathbf{Q}_f \succeq \mathbf{0}$, and $\text{Tr}(\mathbf{Q}_n + \mathbf{Q}_f) \leq 1$, respectively.

The KKT conditions relevant to the proof can be defined as

$$\mathbf{C}_n = \mathbf{C}_n' - \overline{\mathbf{H}}_n(\mathbf{B}_n + \mathbf{B}_n^f)\overline{\mathbf{H}}_n^H, \tag{58a}$$

$$\mathbf{C}_f = \mathbf{C}_f' - \overline{\mathbf{H}}_n \mathbf{B}_n^f \overline{\mathbf{H}}_n^H, \tag{58b}$$

$$\mathbf{A}_n^f \mathbf{T}_n^f = \mathbf{0}, \mathbf{A}_n \mathbf{T}_n = \mathbf{0}, \mathbf{A}_f \mathbf{T}_f = \mathbf{0} \ \mathbf{B}_n \mathbf{D}_n = \mathbf{0}, \tag{58c}$$

$$\mathbf{B}_n^f \mathbf{D}_n^f = \mathbf{0}, \mathbf{C}_n \mathbf{Q}_n = \mathbf{0}, \mathbf{C}_f \mathbf{Q}_f = \mathbf{0}, \tag{58d}$$

where $\mathbf{C}_n' = \delta\mathbf{I} + \mathbf{H}_n \mathbf{A}_n^f \mathbf{H}_n^H + \mathbf{H}_f \mathbf{A}_n \mathbf{H}_f^H$ and $\mathbf{C}_f' = \delta\mathbf{I} - \frac{\mathbf{H}_n \mathbf{A}_n^f \mathbf{H}_n^H}{\gamma_1} - \mathbf{H}_f \mathbf{A}_f \mathbf{H}_f^H$. Since $\delta > 0$, $\mathbf{A}_n^f \succeq \mathbf{0}$, and $\mathbf{A}_n \succeq \mathbf{0}$, we obtain $\mathbf{C}_n' \succ \mathbf{0}$. If $\mathbf{B}_n + \mathbf{B}_n^f = \mathbf{0}$, we obtain $\mathbf{C}_n \succ \mathbf{0}$, which indicates $\mathbf{Q}_n = \mathbf{0}$. It shows that $\mathbf{Q}_n = \mathbf{0}$ cannot be the solution to the problem. Hence, $\mathbf{B}_n + \mathbf{B}_n^f \succ \mathbf{0}$ always holds. Besides, since $\mathbf{B}_n \succeq \mathbf{0}$ and $\mathbf{B}_n^f \succeq \mathbf{0}$, $\mathbf{B}_n \succ \mathbf{0}$ and $\mathbf{B}_n^f \succ \mathbf{0}$

always hold. Then, left-multiplying the two sides of (58a) by $\mathbf{Q}_n$, we obtain

$$\mathbf{Q}_n(\delta\mathbf{I}+\mathbf{H}_n\mathbf{A}_n^f\mathbf{H}_n^H+\mathbf{H}_f\mathbf{A}_n\mathbf{H}_f^H)=\mathbf{Q}_n\overline{\mathbf{H}}_n(\mathbf{B}_n+\mathbf{B}_n^f)\overline{\mathbf{H}}_n^H. \quad (59)$$

We hold the following relation

$$\begin{aligned}
\mathrm{rank}(\mathbf{Q}_n) &\overset{(a)}{=} \mathrm{rank}(\mathbf{Q}_n(\delta\mathbf{I}+\mathbf{H}_n\mathbf{A}_n^f\mathbf{H}_n^H+\mathbf{H}_f\mathbf{A}_n\mathbf{H}_f^H) \\
&\overset{(b)}{=} \mathrm{rank}(\mathbf{Q}_n\overline{\mathbf{H}}_n(\mathbf{B}_n+\mathbf{B}_n^f)\overline{\mathbf{H}}_n^H) \\
&\overset{(c)}{\leq} \min\{\mathrm{rank}(\overline{\mathbf{H}}_n(\mathbf{B}_n+\mathbf{B}_n^f)\overline{\mathbf{H}}_n^H),\mathrm{rank}(\mathbf{Q}_n)\}, \quad (60)
\end{aligned}$$

where (a) is due to $\mathbf{C}'_n \succ \mathbf{0}$, (b) and (c) are based on the rank inequality property [39]. If we can prove that $\mathrm{rank}(\overline{\mathbf{H}}_n(\mathbf{B}_n+\mathbf{B}_n^f)\overline{\mathbf{H}}_n^H) \leq 1$, we will obtain $\mathrm{rank}(\mathbf{Q}_n) \leq 1$. Next, we consider the following two facts:

$$[\mathbf{0}_{N\times 1}\ \mathbf{0}_{N\times 1}\ \mathbf{I}_N]\overline{\mathbf{H}}_n^H = \mathbf{I}_N, \quad (61)$$

$$[\mathbf{0}_{N\times 1}\ \mathbf{0}_{N\times 1}\ \mathbf{I}_N]\mathbf{\Theta}_n = \frac{\alpha}{\varepsilon_n^2}(\overline{\mathbf{H}}_n - [\mathbf{0}_{N\times 1}\ \hat{\mathbf{h}}_n\ \mathbf{0}_{N\times 1}]). \quad (62)$$

Premultiplying $[\mathbf{0}_{N\times 1}\ \mathbf{0}_{N\times 1}\ \mathbf{I}_N]$ and postmultiplying $\overline{\mathbf{H}}_n^H$ to $\mathbf{D}_n\mathbf{B}_n$, we have

$$\begin{aligned}
&\frac{\alpha}{\varepsilon_n^2}(\overline{\mathbf{H}}_n - [\mathbf{0}_{N\times 1}\ \hat{\mathbf{h}}_n\ \mathbf{0}_{N\times 1}])\mathbf{B}_n\overline{\mathbf{H}}_n^H+\mathbf{Q}_n\overline{\mathbf{H}}_n\mathbf{B}_n\overline{\mathbf{H}}_n^H=\mathbf{0} \\
&\Rightarrow (\frac{\alpha}{\varepsilon_n^2}\mathbf{I}+\mathbf{Q}_n)\overline{\mathbf{H}}_n\mathbf{B}_n\overline{\mathbf{H}}_n^H= \frac{\alpha}{\varepsilon_n^2}[\mathbf{0}_{N\times 1}\ \hat{\mathbf{h}}_n\ \mathbf{0}_{N\times 1}]\mathbf{B}_n\overline{\mathbf{H}}_n^H. (63)
\end{aligned}$$

Similarly, we have

$$(\frac{\beta}{\varepsilon_n^2}\mathbf{I}+\mathbf{Q}_n+\mathbf{Q}_f)\overline{\mathbf{H}}_n\mathbf{B}_n^f\overline{\mathbf{H}}_n^H= \frac{\beta}{\varepsilon_n^2}[\mathbf{0}_{N\times 1}\ \hat{\mathbf{h}}_n\ \mathbf{0}_{N\times 1}]\mathbf{B}_n^f\overline{\mathbf{H}}_n^H. (64)$$

Since $\alpha > 0$ and $\beta > 0$ always hold, $\mathrm{rank}(\overline{\mathbf{H}}_n\mathbf{B}_n\overline{\mathbf{H}}_n^H) = \mathrm{rank}([\mathbf{0}_{N\times 1}\ \hat{\mathbf{h}}_n\ \mathbf{0}_{N\times 1}]\mathbf{B}_n\overline{\mathbf{H}}_n^H)$ and $\mathrm{rank}(\overline{\mathbf{H}}_n\mathbf{B}_n^f\overline{\mathbf{H}}_n^H) = \mathrm{rank}([\mathbf{0}_{N\times 1}\ \hat{\mathbf{h}}_n\ \mathbf{0}_{N\times 1}]\mathbf{B}_n^f\overline{\mathbf{H}}_n^H)$. Furthermore, we have $\mathrm{rank}(\overline{\mathbf{H}}_n(\mathbf{B}_n + \mathbf{B}_n^f)\overline{\mathbf{H}}_n^H) = \mathrm{rank}([\mathbf{0}_{N\times 1}\ \hat{\mathbf{h}}_n\ \mathbf{0}_{N\times 1}](\mathbf{B}_n + \mathbf{B}_n^f)\overline{\mathbf{H}}_n^H) \leq 1$. Therefore, $\mathrm{rank}(\mathbf{Q}_n) \leq 1$.

Next, we introduce the following lemma to prove the rank of $\mathbf{Q}_f$.

Lemma 6: Define $\mathbf{X}$ and $\mathbf{Y}$ as two matrices with the same dimension. Then, $\mathrm{rank}(\mathbf{X}+\mathbf{Y}) \geq \mathrm{rank}(\mathbf{X})-\mathrm{rank}(\mathbf{Y})$ always hold.

Proof: Based on the rank inequality property [39], we know that $\mathrm{rank}(\mathbf{X} + \mathbf{Y}) \leq \mathrm{rank}(\mathbf{X}) + \mathrm{rank}(\mathbf{Y})$ and $\mathrm{rank}(-\mathbf{X}) = \mathrm{rank}(\mathbf{X})$. Replacing $\mathbf{X}$ by $\mathbf{X} + \mathbf{Y}$, we have $\mathrm{rank}(\mathbf{X} + -\mathbf{Y}) + \mathrm{rank}(\mathbf{Y}) \geq \mathrm{rank}(\mathbf{X})$. Then, Lemma 6 is proved. ∎

Since $\mathbf{B}_n^f \succ \mathbf{0}$ always holds, $\mathrm{rank}(\mathbf{C}_f) \geq \mathrm{rank}(\mathbf{C}'_f) - \mathrm{rank}(\overline{\mathbf{H}}_n\mathbf{B}_n^f\overline{\mathbf{H}}_n^H) = \mathrm{rank}(\mathbf{C}'_f) - 1$. In addition, we obtain $\mathbf{C}_f$ is full rank matrix, $\mathbf{C}'_f \succ \mathbf{0}$. We have $\mathrm{rank}(\mathbf{C}_f) \geq N - 1$. If $\mathbf{C}_f$ is full rank, $\mathbf{Q}_f = \mathbf{0}$, which cannot be the solution to the problem. Hence, we have $\mathrm{rank}(\mathbf{C}_f) = N - 1$. According to (58d), we can prove that $\mathrm{rank}(\mathbf{Q}_f) \leq 1$ always holds. Combining the rank analysis of $\mathbf{Q}_n$ and $\mathbf{Q}_f$, the proof of theorem 1 is completed.

APPENDIX C
PROOF OF THEOREM 2

In this part, we will prove the relaxation tightness of the problem (37). First, we need to deal with the SOC constraints in the problem (37). The SOC constraint (32b) can be approximated as

$$\begin{aligned}
(37) \Leftrightarrow & \sqrt{\|\mathbf{C}_n^{\frac{1}{2}}\mathbf{Q}_n\mathbf{C}_n^{\frac{1}{2}}\|_F^2 + 2\|\mathbf{C}_n^{\frac{1}{2}}\mathbf{Q}_n\hat{\mathbf{h}}_n\|_2^2} \\
\overset{(a)}{\leq} & \sqrt{\|\mathbf{C}_n^{\frac{1}{2}}\mathbf{Q}_n\|_F^2(\|\mathbf{C}_n^{\frac{1}{2}}\|_F^2 + 2\|\hat{\mathbf{h}}_n\|_2^2)} \\
\overset{(b)}{\leq} & \sqrt{\mathrm{Tr}(\mathbf{Q}_n\mathbf{Q}_n^H)}\sqrt{\mathrm{Tr}^2(\mathbf{C}_n) + 2\mathrm{Tr}(\mathbf{C}_n)\|\hat{\mathbf{h}}_n\|_2^2} \\
= & \mathrm{Tr}(\mathbf{Q}_n\mathbf{Q}_n^H) \leq l^2 \quad (65)
\end{aligned}$$

where $l^2 = \frac{\lambda^2}{\mathrm{Tr}^2(\mathbf{C}_n)+2\mathrm{Tr}(\mathbf{C}_n)\|\hat{\mathbf{h}}_n\|_2^2}$. Step (a) and (b) is due to the property of the matrix norm [39]. Furthermore, (65) can be rewritten as the following LMI

$$\mathrm{Tr}(\mathbf{Q}_n\mathbf{Q}_n^H) \leq \frac{\lambda^2}{l^2} \Rightarrow \begin{bmatrix} l\mathbf{I} & \mathbf{Q}_n \\ \mathbf{Q}_n^H & l\mathbf{I} \end{bmatrix} \succeq \mathbf{0}, \quad (66)$$

which can eventually rewritten as

$$\begin{bmatrix} l\mathbf{I} & \mathbf{0} \\ \mathbf{0} & l\mathbf{I} \end{bmatrix} \succeq \begin{bmatrix} \mathbf{I} \\ \mathbf{0} \end{bmatrix} \mathbf{Q}_n \begin{bmatrix} \mathbf{0} & -\mathbf{I} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ -\mathbf{I} \end{bmatrix} \mathbf{Q}_n \begin{bmatrix} \mathbf{I} & \mathbf{0} \end{bmatrix}. \quad (67)$$

Also, we know $\|\mathbf{Q}_n\|_F \leq l$. Then, we introduce *Nemirovski lemma* to further deal with (67).

Lemma 7: [41, Lemma 2] For a given matrices $\mathbf{A} = \mathbf{A}^H$, $\mathbf{B}$, and $\mathbf{C}$, $\mathbf{A} \succeq \mathbf{B}^H\mathbf{X}\mathbf{C}+\mathbf{C}^H\mathbf{X}^H\mathbf{B}$ with $\|\mathbf{X}\| \leq t$ is satisfied if and only if the following LMI is satisfied with a real number $a \geq 0$

$$\begin{bmatrix} \mathbf{A} - a\mathbf{C}^H\mathbf{C} & -t\mathbf{B}^H \\ -t\mathbf{B} & a\mathbf{I} \end{bmatrix} \succeq \mathbf{0}. \quad (68)$$

By applying Lemma 7 to (67), the SOC constraint (32b) can be finally converted into the following LMI constraint, which does not involve $\mathbf{Q}_n$

$$\begin{bmatrix} \begin{bmatrix} l\mathbf{I} & \mathbf{0} \\ \mathbf{0} & l\mathbf{I} \end{bmatrix} - a\begin{bmatrix} \mathbf{0} \\ -\mathbf{I} \end{bmatrix}\begin{bmatrix} \mathbf{0} & -\mathbf{I} \end{bmatrix} & -l\begin{bmatrix} \mathbf{I} \\ \mathbf{0} \end{bmatrix} \\ -l\begin{bmatrix} \mathbf{I} & \mathbf{0} \end{bmatrix} & a\mathbf{I} \end{bmatrix} \succeq \mathbf{0}. \quad (69)$$

Similarly, the SOC constraints (34b) and (35b) also can be converted into the LMI constraints without the variables $\mathbf{Q}_n$ and $\mathbf{Q}_f$ based on Lemma 7.

The Lagrangian function of the problem (37) is given by

$$\begin{aligned}
\mathcal{L}=&-a\mathrm{Tr}(\mathbf{R}\mathbf{Q}_n)-\mathrm{Tr}(\mathbf{\Pi}\mathbf{Q}_n)-b\mathrm{Tr}(\mathbf{R}(\frac{\mathbf{Q}_f}{\gamma_1}-\mathbf{Q}_n)) \\
&-\mathrm{Tr}(\mathbf{\Xi}(\frac{\mathbf{Q}_f}{\gamma_1}-\mathbf{Q}_n))-c\mathrm{Tr}(\mathbf{R}(\mathbf{Q}_f+\mathbf{Q}_n))-\mathrm{Tr}(\mathbf{\Phi}_n\mathbf{Q}_n) \\
&-\mathrm{Tr}(\mathbf{\Psi}(\mathbf{Q}_f+\mathbf{Q}_n))-\frac{\mathrm{Tr}(\mathbf{G}\mathbf{H}_f^H\mathbf{Q}_f\mathbf{H}_f)}{z}+\mathrm{Tr}(\mathbf{G}\mathbf{H}_f^H\mathbf{Q}_n\mathbf{H}_f) \\
&-\mathrm{Tr}(\mathbf{\Phi}_f\mathbf{Q}_f)+\delta\mathrm{Tr}(\mathbf{Q}_f+\mathbf{Q}_n)+\Lambda \quad (70)
\end{aligned}$$

where $\mathbf{R} = \mathbf{C}_n + \hat{\mathbf{h}}_n\hat{\mathbf{h}}_n^H$, $\mathbf{\Pi} = \mathbf{C}_n^{\frac{1}{2}}\mathbf{A}\mathbf{C}_n^{\frac{1}{2}}$, $\mathbf{\Xi} = \mathbf{C}_n^{\frac{1}{2}}\mathbf{B}\mathbf{C}_n^{\frac{1}{2}}$, $\mathbf{\Psi} = \mathbf{C}_n^{\frac{1}{2}}\mathbf{D}\mathbf{C}_n^{\frac{1}{2}}$. And $\mathbf{A} \in \mathbb{H}_+^N$, $\mathbf{B} \in \mathbb{H}_+^N$, $\mathbf{D} \in \mathbb{H}_+^N$, $\mathbf{\Phi}_n \in \mathbb{H}_+^N$, $\mathbf{\Phi}_f \in \mathbb{H}_+^N$, $\mathbf{G} \in \mathbb{H}_+^{N+1}$, a, b, c, δ are dual variables associated

with the constraints in (37). Note that Λ denotes the collections of variables not related to $\mathbf{Q}_n$ and $\mathbf{Q}_f$. Then, the relevant KKT conditions can be defined as

$$\delta\mathbf{I} + (\frac{b}{\gamma_1} + c)\mathbf{R} + \frac{\mathbf{\Xi}}{\gamma_1} + \mathbf{\Psi} + \mathbf{C}_f = \frac{\mathbf{H}_f\mathbf{G}\mathbf{H}_f^H}{z}, \quad (71a)$$

$$\mathbf{C}_n\mathbf{Q}_n = \mathbf{0}, \ \mathbf{C}_f\mathbf{Q}_f = \mathbf{0}, \quad (71b)$$

$$\delta(\text{Tr}(\mathbf{Q}_n + \mathbf{Q}_f) - 1) = 0, \quad (71c)$$

$$\mathbf{Q}_n \succeq \mathbf{0}, \ \mathbf{Q}_f \succeq \mathbf{0}, \ \mathbf{C}_n \succeq \mathbf{0}, \ \mathbf{C}_f \succeq \mathbf{0}. \quad (71d)$$

Left-multiplying the two sides of (71d) by $\mathbf{Q}_f$, we obtain

$$\mathbf{Q}_f(\delta\mathbf{I} + (\frac{b}{\gamma_1} + c)\mathbf{R} + \frac{\mathbf{\Xi}}{\gamma_1} + \mathbf{\Psi} + \mathbf{C}_f) = \frac{\mathbf{H}_f\mathbf{G}\mathbf{H}_f^H\mathbf{Q}_f}{z} \quad (72)$$

From the prove in Appendix B, we know that $\text{rank}(\mathbf{Q}_f) = \min\{\text{rank}(\mathbf{H}_f\mathbf{G}\mathbf{H}_f^H), \text{rank}(\mathbf{Q}_f)\}$ with $\delta\mathbf{I} + (\frac{b}{\gamma_1} + c)\mathbf{R} + \frac{\mathbf{\Xi}}{\gamma_1} + \mathbf{\Psi} + \mathbf{C}_f \succ \mathbf{0}$. The proof of $\text{rank}(\mathbf{H}_f\mathbf{G}\mathbf{H}_f^H) \leq 1$ is similar to the proof of $\text{rank}(\mathbf{B}_n^f\mathbf{D}_n^f\overline{\mathbf{H}}_n^H)$ in Appendix B. Hence, we can prove that $\text{rank}(\mathbf{Q}_f) \leq 1$ and $\text{rank}(\mathbf{Q}_n) \leq 1$. The proof is completed.

REFERENCES

[1] Y. Saito, A. Benjebbour, Y. Kishiyama, and T. Nakamura, "Systemlevel performance evaluation of downlink non-orthogonal multiple access (NO-MA)," in *Proc. IEEE Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC)*, London, U.K., Sep. 2013, pp. 611-615.

[2] Z. Ding, Z. Yang, P. Fan, and H. V. Poor, "On the performance of non-orthogonal multiple access in 5G systems with randomly deployed users," *IEEE Signal Process. Lett.*, vol. 21, no. 12, pp. 1501-1505, Dec. 2014.

[3] Z. Ding *et al.*, "Application of non-orthogonal multiple access in LTE and 5G networks," *IEEE Commun. Mag.*, vol. 55, no. 2, pp. 185-191, Feb. 2017.

[4] Z. Ding, X. Lei, G. K. Karagiannidis, R. Schober, J. Yuan, and V. Bhargava, "A survey on non-orthogonal multiple access for 5G networks: Research challenges and future trends," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 10, pp. 2181-2195, Oct. 2017.

[5] S. Timotheou and I. Krikidis, "Fairness for non-orthogonal multiple access in 5G systems," *IEEE Signal Process. Lett.*, vol. 22, no. 10, pp. 1647-1651, Oct. 2015.

[6] P. Xu, K.Cumanan, "Optimal power allocation scheme for non-orthogonal multiple access with α-fairness," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 10, pp. 2357-2369, Oct. 2017.

[7] Y. Yuan and Z. Ding, "The application of non-orthogonal multiple access in wireless powered communication networks," in *Proc. IEEE 17th Int. Workshop on Signal Processing Advances in Wireless Commun.*, Edinburgh, UK, Jul. 2016.

[8] Q. Sun, S.Han, C.-L. I, and Z. Pan, "On the egrodic capacity of MIMO NOMA system," *IEEE Wireless Commun. Lett.*, vol. 4, no. 4, pp. 405-408, Aug. 2015.

[9] Z. Ding, F. Adachi, and H. V. Poor, "The application of MIMO to nonorthogonal multiple access," *IEEE Trans. Wireless Commun.*, vol. 15, no. 1, pp. 537-552, Jan. 2016.

[10] Z. Ding, R. Schober, and H. V. Poor, "A general MIMO framework for NOMA downlink and uplink transmission based on signal alignment," *IEEE Trans. Wireless Commun.*, vol. 15, no. 6, pp. 4438-4454, Jun. 2016.

[11] G. Gui, H. Huang, Y. Song, and H. Sari, "Deep Learning for An Effective Non-Orthogonal Multiple Access Scheme," *IEEE Trans. Veh. Technol.*, vol. 67, no. 9, pp. 8440-8450, Jun. 2018.

[12] Z. Ding, M. Peng, and H. V. Poor, "Cooperative non-orthogonal multiple access in 5G systems," *IEEE Commun. Lett.*, vol. 19, no. 8, pp. 1462-1465, Aug. 2015.

[13] H. Huang, J. Xiong, J. Yang, G. Gui, and H. Sari, "Rate region analysis in a full-duplex-aided cooperative nonorthogonal multiple-access system," *IEEE Access*, vol. 5, pp. 17869-17880, Aug. 2017.

[14] P. Xu, Z. Yang, Z.Ding, and Z. Zhang, "Optimal relay selection scheme for cooperative NOMA," *IEEE Trans. Veh. Technol.*, doi: 10.1109/TVT.2018.2821900.

[15] Z. Yang, Z. Ding, Y. Wu, and P. Fan, "Novel relay selection strategies for cooperative NOMA," *IEEE Trans. Veh. Technol.* vol. 66, no. 11, pp. 10114-10123, Nov. 2017.

[16] L. R. Varshney, "Transporting information and energy simultaneously," in *Proc. IEEE Int. Symp. Inf. Theory*, July 2008, pp. 1612-1616.

[17] 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.

[18] G. Pan, H. Lei, Y. Yuan and Z. Ding, "Performance analysis and optimization for SWIPT wireless sensor networks," *IEEE Trans. Commun.*, vol. 65, no.5, pp. 2291-2302, May 2017.

[19] Y. Yuan and Z. Ding, "Outage constrained secrecy rate maximization design with SWIPT in MIMO-CR systems," *IEEE Trans. Veh. Technol.*, vol. 67, no.6, pp. 5475-5480, June 2018.

[20] Y. Liu, Z. Ding, M. Elkashlan, and H. V. Poor, "Cooperative nonorthogonal multiple access with simultaneous wireless information and power transfer," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 938-953, Apr. 2016.

[21] N. T. Do, D. B. da Costa, T. Q. Duong, and B. An, "A BNBF user selection scheme for NOMA-based cooperative relaying systems with SWIPT," *IEEE Commun. Lett.*, vol. 21, no. 3, pp. 664-667, Mar. 2017.

[22] Y. Xu *et al.*, "Joint beamforming and power-splitting control in downlink cooperative SWIPT NOMA systems," *IEEE Trans. Signal Process.*, vol. 65, no. 18, pp. 4874-4886, Sep. 2017.

[23] P. Xu, Y. Yuan, Z. Ding, X. Dai, and R. Schober, "On the outage performance of non-orthogonal multiple access with 1-bit feedback," *IEEE Trans. Wireless Commun.*, vol. 15, no. 10, pp. 6716-6730, Oct. 2016.

[24] Z. Yang, Z. Ding, P. Fan, and G. Karagiannidis, "On the performance of 5G non-orthogonal multiple access systems with partial channel information," *IEEE Trans. Commun.*, vol. 64, no. 2, pp. 654-667, Feb. 2016

[25] Q. Zhang, Q. Li, and J. Qin, "Robust beamforming for nonorthongonal muliple-access systems in MISO channels," *IEEE Trans. Veh. Technol.*, vol. 65, no. 12, pp. 10231-10236, Dec. 2016.

[26] F. Alavi, K. Cumanan, Z. Ding, and A. G. Burr, "Robust beaforming techniques for non-orthogonal multiple access systems with bounded channel uncertainties," *IEEE Commun. Lett.*, vol. 21, no. 9, pp. 2033-2036, May. 2017.

[27] A. A. Nasir, X. Zhou, S. Durrani, and R. A. Kennedy, "Relaying protocols for wireless energy harvesting and information processing," *IEEE Trans. Wireless Commun.*, vol. 12, no. 7, pp. 3622-3636, Jul. 2013.

[28] I. Krikidis, S. Sasaki, S. Timotheou, and Z. Ding, "A low complexity antenna switching for joint wireless information and energy transfer in MIMO relay channels," *IEEE Trans. Commun.*, vol. 62, no. 5, pp. 1577-1587, May 2014.

[29] G. Zheng, K. K. Wong, and B. Ottersten, "Robust cognitive beamforming with bounded channel uncertainties," *IEEE Trans. Signal Process.*, vol. 57, no. 12, pp. 4871-4881, Dec. 2009.

[30] 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.

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

[32] Z.-Q. Luo, J. F. Sturm, and S. Zhang, "Multivariate nonnegative quadratic mappings," *SIAM J. Optim.*, vol. 14, no.4, pp. 1140-1162, May 2004.

[33] A. Beck, A. Ben-Tal, and L. Tetruashvili, "A sequential parametric convex approximation method with applications to nonconvex truss topology design problems," *J. Global Optim.*, vol. 47, no. 1, pp. 29-51, 2010.

[34] W.-C. Li, T.-H. Chang, C. Lin, and C.-Y. Chi, "Coordinated beamforming for multiuser MISO interference channel under rate outage constraints," *IEEE Trans. Signal Process.*, vol. 61, no. 5, pp. 1087-1103, Mar. 2013.

[35] Z.-Q. Luo, W.-K. Ma, A. M.-C.So, Y. YE, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Process. Mag*, vol. 27, no. 3, pp. 20-34, May 2010.

[36] M. Grant and S. Boyd, CVX: Matlab software for disciplined convex programming, version 1.21. Apr. 2011 [Online]. Available: http://cvxr.com/cvx/

[37] 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.

[38] Z. Chu, H. Xing, M. Johnson, and S. Le 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.

[39] R. A. Horn and C. R. Johnson, *Matrix Analysis,* Cambridge, U.K.: Cambridge Univ. Press, 2012.

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

[41] 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.