Joint Optimization of Secure Over-the-Air Computation and Reliable Multicasting Assisted by a MIMO Untrusted Two-Way Relay

Quanzhong Li, Hualiang Luo, and Liang Yang

***Abstract*—In this article, we investigate the joint optimization of secure Over-the-Air Computation (AirComp) and reliable multicasting assisted by a multiple-input–multiple-output-untrusted two-way relay, where artificial noise is employed at the access point (AP) to interfere the relay for ensuring secure AirComp. We aim to minimize the computation distortion at the AP by jointly designing the transmit variables of all the nodes and the aggregation variables at the AP and relay, under the secure AirComp constraint, the reliable multicasting constraint, and the transmit power constraints of all the nodes. We consider two scenarios that perfect and imperfect channel state information (CSI) are available. For the former, the formulated optimization problem is highly nonconvex, and we propose an effective block coordinate descent (BCD)-penalty successive convex approximation (penaltySCA) method to solve the nonconvex problem. For the latter, we model the imperfect CSI by using the worst case criterion and the formulated robust optimization problem is much more challenging than its counterpart with perfect CSI. To solve the robust problem effectively, we first transform it into a deterministic optimization problem by employing some powerful mathematical lemmas and then apply the proposed BCD-penaltySCA method to solve the reformulated deterministic problem. The proposed methods are shown by simulations to significantly reduce the computation distortion compared with other benchmarks under considering secure AirComp and reliable multicasting.**

***Index Terms*—Beamforming, channel state information (CSI), multiple-input multiple-output (MIMO), over-the-air computation (AirComp), security optimization, untrusted two-way relay.**

I. Introduction

OVER-THE-AIR computation (AirComp) is a promising technique to aggregate the sensing data from the distributed sensors in the Internet of Things (IoT) network by exploiting the supposition of wireless signals over the multiple access channel, bringing the advantage of very high spectral efficiency and extremely low communication delay [1], [2], [3]. Although some works consider computational rate as the objective function [4], [5], the objective function of most prior works is the computation error, namely, mean squared error (MSE) of the aggregation result of the sensors' data [6], [7], [8], [9], [10], [11]. Many early works focus on the point-to-point (i.e., single-hop) AirComp networks where the sensors directly communicate with the access point (AP), and optimize the power control in single-antenna AirComp networks [6], [7], [8] or the beamforming in multiple-input–multiple-output (MIMO) AirComp networks [9], [10], [11]. Besides wireless communication, the AirComp is also used in other fields. In federated learning, Yan et al. [12] transmitted each distributed device's gradient through AirComp, while Zhang et al. [13] proposed a coded AirComp scheme for fast model aggregation. In [14], the AirComp is introduced to enable efficient computation of the control signal for control systems.

However, the AP in the AirComp network may suffer from severe fading when sensors are far from it. In order to address the issue caused by severe fading, relays have been widely applied in the traditional wireless communication networks [15], [16], [17], [18], [19], [20], [21], [22], [23], [24], [25]. In order to obtain better quality of service, single-antenna relay [15], [16] and multiantenna relay [17], [18] are investigated in wireless communication networks. Some existing research works also consider the cooperation with relay [19], [20] or selection of multiple relays [21], [22] schemes in wireless communication networks. Furthermore, due to the high maneuverability and flexible deployment, the works in [23], [24], and [25] consider utilizing the unmanned aerial vehicle as a relay.

Motivated by the great success of relaying in wireless communications, relays have been introduced in the evolving AirComp networks [26], [27], [28], [29], [30], [31], [32], [33], [34]. However, to the best of our knowledge, an untrusted two-way relay-assisted AirComp is not in the literature and its security issue is worth studying. In this article, we investigate the joint optimization of secure AirComp and reliable multicasting in a MIMO untrusted two-way relay-assisted communication and computation network, where the AP and the relay have multiple antennas and the sensors have a single antenna. The signal transmission between the AP and the sensors is finished in two time slots. In the first time slot, the sensors send their preprocessed sensing data to the relay and

Manuscript received 8 March 2023; revised 6 May 2023; accepted 9 May 2023. Date of publication 11 May 2023; date of current version 25 September 2023. This work was supported in part by the National Natural Science Foundation of China under Grant 62272493 and Grant 61802447; and in part by the Guangdong Basic and Applied Basic Research Foundation under Grant 2023A1515011201. *(Quanzhong Li and Hualiang Luo contributed equally to this work.) (Corresponding author: Quanzhong Li.)*

Quanzhong Li is with the School of Computer Science and Engineering, Sun Yat-sen University, Guangzhou 510006, China (e-mail: liquanzh@mail.sysu.edu.cn).

Hualiang Luo is with the School of Electronics and Information Technology, Sun Yat-sen University, Guangzhou 510006, China (e-mail: luohliang@mail2.sysu.edu.cn).

Liang Yang is with the College of Computer Science and Electronic Engineering, Hunan University, Changsha 410082, China (e-mail: liangy@hnu.edu.cn).

Digital Object Identifier 10.1109/JIOT.2023.3275564

meanwhile the AP transmits the multicasting signal plus the artificial noise (AN) with beamformers to the relay, and in the second time slot, the relay broadcasts the received signal multiplied by a beamforming matrix to the AP and sensors. To ensure secure AirComp and reliable multicasting, we jointly optimize the transmit variables of all the nodes and the aggregation variables at the AP and relay, with the objective to minimize the computation distortion at the AP under the constraints that the minimum computation distortion at the relay is larger than some given threshold, the signal-to-interference-plus-noise ratio (SINR) at each sensor is higher than some predetermined value, and the transmit power at all the nodes cannot exceed their budget.

A. Novelties and Contributions

The novelty of our work comes from the following three main aspects. First, compared with the existing relay-assisted AirComp [26], [27], [28], [29], [30], [31], [32], [33], [34], the proposed untrusted relay-assisted AirComp wants to prevent the relay from wiretapping the aggregation of the sensors' data and thus has an additional security constraint based on the computation distortion, which is highly nonconvex and make the optimization problem much more challenging to deal with. Second, conventional untrusted relay-assisted communication needs to decode each data of the sources (sensors), aiming to maximize the secrecy rate or improve the communication reliability [36], [37], [38], [39], [40], [41], while our proposed untrusted relay-assisted AirComp focuses on the secure aggregation of the sensors' data instead of decoding individual data and its performance metric is the computation distortion in the received value of the desired function, which is a totally different objective from convectional untrusted relay-assisted communication. Third, the existing relay-assisted AirComp [26], [27], [28], [29], [30], [31], [32], [33], [34] only consider the case of perfect channel state information (CSI) which is only applicable to a specific scenario, while our work considers both cases of perfect and imperfect CSI and has a greater flexibility to adapt to different application scenarios.

The main contributions of our work are summarized as follows.

1) We propose a MIMO untrusted two-way relay-assisted secure AirComp and reliable multicasting network, where we employ the multicasting signal and the AN to interfere the relay to ensure secure AirComp. We formulate the MSE minimization problem with the nonconvex secure AirComp constraint and reliable multicasting constraints, considering both perfect and imperfect CSI.
2) With perfect CSI available, we propose an effective block coordinate descent (BCD)-penalty successive convex approximation (penaltySCA) method to solve the nonconvex MSE minimization problem, where the aggregation variables are found in closed form and the transmit variables are optimized by the SCA or penaltySCA method in which a second-order cone programming (SOCP) or a semidefinite programming (SDP) is solved.
3) When the CSI is imperfect, we employ the worst case criterion to model the CSI, and the corresponding robust MSE minimization problem is much more challenging than its counterpart with perfect CSI. To make the robust problem tractable, we transform it to a deterministic optimization problem by employing some powerful mathematical lemmas. Since he reformulated deterministic problem is still nonconvex, we apply the proposed BCD-penaltySCA method to deal with it.
4) We provide numerical results to show that compared with other benchmarks, the proposed methods have a significant improvement in terms of the computation distortion under the constraints of secure AirComp and reliable multicasting.

B. Organization and Notations

The remainder of this article is organized as follows. In Section III, the secure AirComp and reliable multicasting network model is described. In Section IV, we formulate the optimization problem for the case of perfect CSI and propose an effective BCD-penaltySCA algorithm. In Section V, we formulate the robust optimization problem for the case of imperfect CSI and solve the robust problem by the proposed BCD-penaltySCA algorithm. Simulation results are presented in Section VI, and we conclude this article in Section VII.

Notations: $\mathbf{U}^\dagger$, $\mathbf{U}^*$, $\mathbf{U}^T$, $\mathrm{tr}(\mathbf{U})$, and $\|\mathbf{U}\|$ denote the conjugate transpose, conjugate, transpose, trace, and Frobenius norm of the matrix $\mathbf{U}$, respectively. $\mathrm{vec}(\mathbf{V})$ means stacking the columns of the matrix $\mathbf{V}$ into a single vector. $\odot$ denotes the Hadamard product. $\mathbf{W} \succeq \mathbf{0}$ mean that $\mathbf{W}$ is positive semidefinite. $\Re\{x\}$ denotes the real part of the complex number x.

II. Related Works

Since the relay can substantially overcome the disadvantage of short coverage of single-hop AirComp networks, *relay-assisted* (i.e., two-hop) AirComp networks have been studied recently [26], [27], [28], [29], [30], [31], [32], [33]. Wang et al. [26] studied a hierarchical Aircomp network assisted by multiple single-antenna relays and proposed a centralized algorithm and a decentralized algorithm with global and partial CSI, respectively. In [27], a part of sensors transmit their sensing signals to the AP with the assistance of a single-antenna cooperative relay, while other sensors directly send their sensing signals to the AP, and the impact caused by the relay position has been studied. A cooperative multiantenna amplify-and-forward (AF) relay-aided AirComp network has been considered in [30], while the corresponding case of cooperating with multiple multiantenna AF relays has been investigated in [31]. The works [32], [33] mainly focused on the performance analysis of relay-assisted AirComp networks. Jiang et al. [32] analyzed the MSE outage probability and diversity order in the context of AirComp and developed a relay selection scheme to achieve the full diversity order, and the work [33] studied wireless edge federated learning system based on relay-assisted AirComp and characterized the decoding failure probability at the AP. Different from the above two-hop AirComp networks, the work [34] proposed a multihop digital AirComp network where the authors derived the computation rate and proposed a time allocation and power control scheme to improve the network performance.

In the aforementioned scenarios, the relay is considered to be trusted, in the sense that, the relay will not infer the aggregation result of the sensing data from the sensors. However, when the relay is interested in the aggregation result and tends to decipher the result for the unallowed operations, despite operating with the desired relaying protocols, the relay becomes *untrusted* [35]. To maximize the secrecy rate, the security issue of an untrusted relay-assisted communication has been widely investigated in [36], [37], [38], [39], [40], and [41]. In [36], the secrecy outage probability and ergodic secrecy rate in a cooperative network in presence of multiple untrusted single-antenna relays and multiple passive eavesdroppers has been investigated. A nonorthogonal multiple access scheme has been considered in [37], where the single-antenna relay that assists the communications between the source and the far user is assumed to be untrusted. In [38], both scenarios of single-antenna and multiple-antenna relay are investigated, and the analytical expression for a lower bound on the ergodic secrecy sum rate has been derived. In [39], both cooperative and noncooperative secure beamforming schemes with an untrusted MIMO AF relay have been investigated, and the result in [39] has been extended to MIMO two-way untrusted relay systems [40]. Later, the MIMO two-way untrusted relay in [40] has been considered to perform full-duplex operation in [41].

With regard to the security design for AirComp, in [42], a friendly jammer is used to against passive eavesdropping, where the jammer's signal is reconstructed and fully canceled by the legitimate receiver but deteriorates the eavesdropper's signal-noise ratio (SNR), and thus inhibits the illegitimate receiver's ability to estimate the value of the objective function. Different from the secrecy rate of the physical layer security defined in traditional wireless communications, a δ-semantically secure based on the total variation norm on signed measures and a V-MSE-secure are defined in [42], which means that the estimation MSE at the eavesdropper is at least V under a uniformly distributed objective, regardless of which estimator the eavesdropper uses. In [43], a multiantenna full-duplex AP utilizes the AN to degrade the eavesdropper's links while receiving sensors' preprocessed sensing signals.

III. System Model

Consider a MIMO untrusted two-way relay-assisted AirComp and multicasting networks as shown in Fig. 1, where K sensors want to send their sensing data (e.g., temperature and humidity) to the AP for aggregation and meanwhile the AP transmits a common signal to all the sensors for some multicasting applications (e.g., updating the status of sensors), both with the help of an untrusted two-way relay. We assume that the sensors are equipped with a single antenna, and the relay and AP has N_r and N_p antennas, respectively [30].

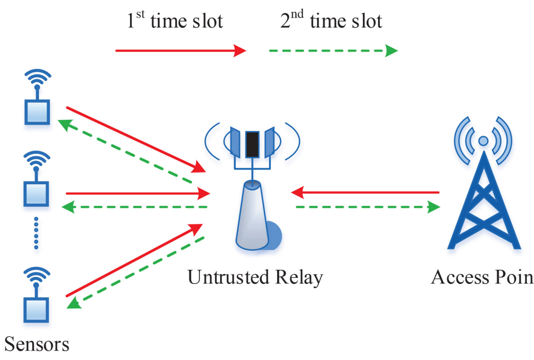

Fig. 1. MIMO untrusted two-way relay-assisted AirComp and multicasting networks.

The transmission of data is finished in two consecutive time slots. In the first time slot, all the sensors send their data to the relay simultaneously, in which the transmitted data by the kth sensor is denoted as

$$x_k = b_k s_k \tag{1}$$

where s_k is the sensing data with zero mean and unit variable, and b_k is the corresponding transmit coefficient.

Meanwhile, the AP transmits a common signal s_c to the relay, and in order to protect the aggregation of sensors' data from being wiretapped by the untrusted relay, the AP also transmits an AN signal $\mathbf{z}$ to interfere the relay. Thus, the transmitted signal by the AP is expressed as

$$\mathbf{x}_c = \mathbf{w}s_c + \mathbf{V}\mathbf{z} \tag{2}$$

where $\mathbf{w}$ and $\mathbf{V}$ is the beamforming vector/matrix for s_c and $\mathbf{z}$, respectively, and $s_c \in \mathcal{CN}(0, 1)$, $\mathbf{z} \in \mathcal{CN}(\mathbf{0}, \mathbf{I})$.

From (1) and (2), the received signal at the relay in the first time slot is given by

$$\mathbf{y}_a = \sum_{k=1}^{K} \mathbf{h}_k b_k s_k + \mathbf{H}_p(\mathbf{w}s_c + \mathbf{V}\mathbf{z}) + \mathbf{n}_r \tag{3}$$

where $\mathbf{h}_k$ and $\mathbf{H}_p$ denote the channel from the kth sensor and AP to the relay, respectively, and $\mathbf{n}_r \in \mathcal{CN}(\mathbf{0}, \sigma_r^2\mathbf{I})$ is the additive noise at the relay.

In the second time slot, the signal received at the relay is multiplied by a beamforming matrix $\mathbf{F}$ and then forwarded to the sensors and AP. From (3), the signal received at the AP is given by

$$\mathbf{y}_p = \mathbf{G}_p\mathbf{F}\left(\sum_{k=1}^{K} \mathbf{h}_k b_k s_k + \mathbf{H}_p(\mathbf{w}s_c + \mathbf{V}\mathbf{z}) + \mathbf{n}_r\right) + \mathbf{n}_p \tag{4}$$

where $\mathbf{g}_k$ and $\mathbf{G}_p$ denotes the channel from the relay to the kth sensor and AP, respectively, and $n_k \in \mathcal{CN}(0, \sigma_k^2)$ and $\mathbf{n}_p \in \mathcal{CN}(\mathbf{0}, \sigma_p^2\mathbf{I})$ are the additive noise at the kth sensor and AP, respectively.

To achieve secure AirComp, the minimum computation distortion measured by MSE at the untrusted relay should be larger than some given threshold, such that the relay cannot compute the aggregation of the sensors' data accurately enough. Considering the relay is interested to compute the arithmetic sum of the sensors' data, i.e., $\bar{s} = \sum_{k=1}^{K} s_k$, and applies the aggregation beamforming vector $\mathbf{a}$ to estimate the value of $\bar{s}$ as

$$\hat{s} = \mathbf{a}^{\dagger}\mathbf{y}_a. \tag{5}$$

Based on (3) and (5), the computation distortion at the relay is given by [26], [27], [28]

$$\begin{aligned}\mathrm{MSE}_r &= \mathbb{E}\left(|\bar{s}-\hat{s}|^2\right)\\ &= \sum_{k=1}^{K}\left|\mathbf{a}^\dagger\mathbf{h}_k b_k - 1\right|^2 + \underbrace{\left|\mathbf{a}^\dagger\mathbf{H}_p\mathbf{w}\right|^2 + \left\|\mathbf{a}^\dagger\mathbf{H}_p\mathbf{V}\right\|^2}_{\text{by interference from AP}}\\ &\quad + \sigma_r^2\|\mathbf{a}\|^2 \end{aligned} \tag{6}$$

from which we can see that the interference from the AP can effectively increase the computation distortion at the relay and thus provide computation security against the untrusted relay.

At the AP, the sensors' data are aggregated by a received beamforming vector $\mathbf{d}$. We assume that the CSI about $\mathbf{H}_p$ and $\mathbf{G}_p$ are accurate, which is reasonable when the AP and relay are fixed and the channels are estimated through pilot signals with high transmit power [26], [27], [28]. Then, the AP can completely remove the self-interference signal $\mathbf{w}s_c+\mathbf{V}\mathbf{z}$ in (4), and the remaining signal at the AP is given by

$$\tilde{\mathbf{y}}_p = \sum_{k=1}^{K}\mathbf{G}_p\mathbf{F}\mathbf{h}_k b_k s_k + \mathbf{G}_p\mathbf{F}\mathbf{n}_r + \mathbf{n}_p. \tag{7}$$

From (7), the MSE for the AP to estimate the arithmetic sum of the sensors' data is expressed as

$$\mathrm{MSE}_p = \sum_{k=1}^{K}\left|\mathbf{d}^\dagger\mathbf{G}_p\mathbf{F}\mathbf{h}_k b_k - 1\right|^2 + \sigma_r^2\left\|\mathbf{d}^\dagger\mathbf{G}_p\mathbf{F}\right\|^2 + \sigma_p^2\|\mathbf{d}\|^2. \tag{8}$$

From (6) and (8), we can see that the signal transmitted by the AP will only affect the computation distortion at the relay not that at the AP, which possibly provides an effective method for computation security against the relay while not interfering the data aggregation at the AP.

At the kth sensor, the received signal given by

$$y_k = \mathbf{g}_k^\dagger\mathbf{F}\left(\sum_{i=1}^{K}\mathbf{h}_i b_i s_i + \mathbf{H}_p(\mathbf{w}s_c+\mathbf{V}\mathbf{z}) + \mathbf{n}_r\right) + n_k. \tag{9}$$

Note that y_k contains the signal s_k sent by the kth sensor in the first time slot, and s_k is known at the kth sensor. Thus, the kth sensor can eliminate the self-interference caused by s_k before decoding the common signal s_c. Let r_k denote the achievable information rate of the kth sensor, and we use the achievable information rate of each sensor to measure the quality of the multicasting. Obviously, the effect of self-interference elimination depends on the corresponding CSI, i.e., $\mathbf{h}_k$ and $\mathbf{g}_k$, and the cases of perfect and imperfect CSI will be discussed in the rest of this article.

Based on the above, we formulate the optimization problem as

$$\begin{aligned}&\min_{\mathbf{F},\mathbf{V},\mathbf{w},\mathbf{d},\{b_k\}} \quad \mathrm{MSE}_p && (10a)\\ &\text{s.t.} \quad \min_{\mathbf{a}} \ \mathrm{MSE}_r \geq \xi_r && (10b)\\ &\qquad r_k \geq R_k,\ k=1,\ldots,K && (10c)\\ &\qquad |b_k|^2 \leq P_k,\ k=1,\ldots,K && (10d)\\ &\qquad \|\mathbf{w}\|^2 + \|\mathbf{V}\|^2 \leq P_{ap} && (10e)\\ &\qquad \sum_{k=1}^{K}\|\mathbf{F}\mathbf{h}_k b_k\|^2 + \left\|\mathbf{F}\mathbf{H}_p\mathbf{w}\right\|^2 + \left\|\mathbf{F}\mathbf{H}_p\mathbf{V}\right\|^2\\ &\qquad + \sigma_r^2\|\mathbf{F}\|^2 \leq P_r. && (10f)\end{aligned}$$

In problem (10), we aim to minimize the computation distortion at the AP [i.e., (10a)], based on the conditions that the minimum computation distortion at the untrusted relay should be larger than a given threshold ξ_r [i.e., (10b)] to protect the computation security, the achievable information rate of kth sensor is greater than a target transmission rate R_k [i.e., (10c)] to ensure the reliable multicasting, and the transmit power constraints at the sensors, AP and relay [i.e., (10d)–(10f)] with the corresponding power budget $\{P_k\}$, P_{ap} and P_r.

Remark 1 (Encryption Technique): Besides the PLS, another security technology, called the encryption technique, is also widely used [44], [45], [46], [47]. The homomorphic encryption, which has the property that the result of decryption after adding ciphertext is equivalent to the addition of plaintext, could be applied in this AirComp network. Combining with the PLS technology, the ciphertext received by the relay may be scrambled. Due to the avalanche effect [44], it is more difficult for the untrusted relay to recover the plaintext from the received ciphertext, and the protection of the aggregation result has been strengthened. As for some classic homomorphic encryption methods, such as the elliptic curve cryptography [45] and Paillier [46], we need to change the computing function of AirComp into a polynomial to adapt to their operation on the ciphertext. In addition, since the encryption methods, especially some nonlightweight methods, may bring a large computational burden to the sensors [47], and whether the encryption method can be applied to the network depends on the sensors' computing resources and power budget.

IV. Joint Optimization With Perfect CSI

In this section, we assume that the CSI about $\{\mathbf{h}_k\}$ and $\{\mathbf{g}_k\}$ is perfect, which is reasonable when the sensors are static and the channels vary slow [29], [30], [31]. With perfect CSI, the kth sensor can eliminate its own data s_k in (9), and the remaining signal at the kth sensor is given by

$$\tilde{y}_k = \mathbf{g}_k^\dagger\mathbf{F}\left(\sum_{i=1,i\neq k}^{K}\mathbf{h}_i b_i s_i + \mathbf{H}_p(\mathbf{w}s_c+\mathbf{V}\mathbf{z}) + \mathbf{n}_r\right) + n_k. \tag{11}$$

From (11), the SINR to decode the multicasting signal s_c at the kth sensor can be expressed as

$$\gamma_k = \frac{\left|\mathbf{g}_k^\dagger\mathbf{F}\mathbf{H}_p\mathbf{w}\right|^2}{\sum_{i=1,i\neq k}^{K}\left|\mathbf{g}_k^\dagger\mathbf{F}\mathbf{h}_i b_i\right|^2 + \left\|\mathbf{g}_k^\dagger\mathbf{F}\mathbf{H}_p\mathbf{V}\right\|^2 + \sigma_r^2\left\|\mathbf{g}_k^\dagger\mathbf{F}\right\|^2 + \sigma_k^2} \tag{12}$$

which should be designed to ensure the reliable multicasting.

Combining (12) and let $\rho_k = 2^{R_k}-1$ denote the target SINR corresponding to the given target transmission rate of the kth sensor, the optimization problem (10) can be expressed as

$$\min_{\mathbf{F},\mathbf{V},\mathbf{w},\mathbf{d},\{b_k\}} \quad \mathrm{MSE}_p \tag{13a}$$

$$\text{s.t.} \ \min_{\mathbf{a}} \ \text{MSE}_r \geq \xi_r \tag{13b}$$

$$\gamma_k \geq \rho_k, \ \ k = 1, \ldots, K \tag{13c}$$

$$|b_k|^2 \leq P_k, \ \ k = 1, \ldots, K \tag{13d}$$

$$\|\mathbf{w}\|^2 + \|\mathbf{V}\|^2 \leq P_{ap} \tag{13e}$$

$$\sum_{k=1}^{K} \|\mathbf{F}\mathbf{h}_k b_k\|^2 + \left\|\mathbf{F}\mathbf{H}_p\mathbf{w}\right\|^2 + \left\|\mathbf{F}\mathbf{H}_p\mathbf{V}\right\|^2 + \sigma_r^2 \|\mathbf{F}\|^2 \leq P_r. \tag{13f}$$

The challenges of solving problem (13) lie in the following two main aspects. First, the optimization variables in problem (13) are highly coupled in the objective and constraints. Second, the computation security constraint (13b) and the reliable multicasting constraint (13c) are highly nonconvex even with decoupled optimization variables. In general, there is no standard method for solving such nonconvex optimization problems optimally. Thus, in the following, we propose an effective BCD-penaltySCA algorithm to solve problem (13), which decouples the coupling optimization variables and makes problem (13) more traceable by splitting the problem into several subproblems, and then solves them alternatively.

A. Optimizing the Aggregation Vectors $\{\boldsymbol{d}, \boldsymbol{a}\}$

With given the variables $\{\mathbf{F}, \mathbf{V}, \mathbf{w}, b_k\}$, the optimization problem with respective to the aggregation vectors $\{\mathbf{d}, \mathbf{a}\}$ are two unconstrained convex quadratic problems

$$\min_{\mathbf{d}} \ \sum_{k=1}^{K} \left|\mathbf{d}^\dagger \mathbf{G}_p \mathbf{F} \mathbf{h}_k b_k - 1\right|^2 + \sigma_r^2 \left\|\mathbf{d}^\dagger \mathbf{G}_p \mathbf{F}\right\|^2 + \sigma_p^2 \|\mathbf{d}\|^2 \tag{14}$$

$$\min_{\mathbf{a}} \ \sum_{k=1}^{K} \left|\mathbf{a}^\dagger \mathbf{h}_k b_k - 1\right|^2 + \left|\mathbf{a}^\dagger \mathbf{H}_p \mathbf{w}\right|^2 + \left\|\mathbf{a}^\dagger \mathbf{H}_p \mathbf{V}\right\|^2 + \sigma_r^2 \|\mathbf{a}\|^2 \tag{15}$$

whose optimal solution can be obtained just by setting the first-order derivations of the objectives to being zero as

$$\mathbf{d} = \left(\mathbf{G}_p \mathbf{F} \mathbf{S} \mathbf{F}^\dagger \mathbf{G}_p^\dagger + \sigma_p^2 \mathbf{I}\right)^{-1} \mathbf{G}_p \mathbf{F} \mathbf{t} \tag{16}$$

$$\mathbf{a} = \left(\mathbf{S} + \mathbf{H}_p \mathbf{w} \mathbf{w}^\dagger \mathbf{H}_p^\dagger + \mathbf{H}_p \mathbf{V} \mathbf{V}^\dagger \mathbf{H}_p^\dagger\right)^{-1} \mathbf{t} \tag{17}$$

where

$$\mathbf{S} \triangleq \sum_{k=1}^{K} |b_k|^2 \mathbf{h}_k \mathbf{h}_k^\dagger + \sigma_r^2 \mathbf{I} \tag{18}$$

$$\mathbf{t} \triangleq \sum_{k=1}^{K} b_k \mathbf{h}_k. \tag{19}$$

B. Optimizing the Relay Beamforming Matrix $\boldsymbol{F}$

With given the variables $\{\mathbf{d}, \mathbf{a}, \mathbf{V}, \mathbf{w}, b_k\}$, the optimization problem with respective to the relay beamforming matrix $\mathbf{F}$ can be expressed as

$$\min_{\mathbf{F}} \ \left\|\hat{\mathbf{d}}^\dagger \mathbf{F} \mathbf{S}^{\frac{1}{2}}\right\|^2 - 2\Re\left\{\hat{\mathbf{d}}^\dagger \mathbf{F} \mathbf{t}\right\} \tag{20a}$$

$$\text{s.t.} \ \ \rho_k \left\|\mathbf{g}_k^\dagger \mathbf{F} \mathbf{T}_k^{\frac{1}{2}}\right\|^2 - \left\|\mathbf{g}_k^\dagger \mathbf{F} \mathbf{R}^{\frac{1}{2}}\right\|^2 + \rho_k \sigma_k^2 \leq 0, \ \ k = 1, \ldots, K \tag{20b}$$

$$\left\|\mathbf{F}\left(\mathbf{S} + \mathbf{R} + \mathbf{H}_p \mathbf{V} \mathbf{V}^\dagger \mathbf{H}_p^\dagger\right)^{\frac{1}{2}}\right\|^2 \leq P_r \tag{20c}$$

where

$$\mathbf{T}_k \triangleq \mathbf{S} - |b_k|^2 \mathbf{h}_k \mathbf{h}_k^\dagger + \mathbf{H}_p \mathbf{V} \mathbf{V}^\dagger \mathbf{H}_p^\dagger \tag{21}$$

$$\mathbf{R} \triangleq \mathbf{H}_p \mathbf{w} \mathbf{w}^\dagger \mathbf{H}_p^\dagger \tag{22}$$

$$\hat{\mathbf{d}} \triangleq \mathbf{G}_p^\dagger \mathbf{d}. \tag{23}$$

Problem (20) is a nonconvex optimization problem since the left hand side of each constraint in (20b) is the difference of two convex quadratic functions. To solve problem (20) effectively, we employ the SCA method to (20).

According to the property of the convex function, we have the following inequality:

$$\left\|\mathbf{g}_k^\dagger \mathbf{F} \mathbf{R}^{\frac{1}{2}}\right\|^2 \geq 2\Re\left\{\mathbf{g}_k^\dagger \mathbf{F} \mathbf{R} \mathbf{F}^{(n)\dagger} \mathbf{g}_k\right\} - \left\|\mathbf{g}_k^\dagger \mathbf{F}^{(n)} \mathbf{R}^{\frac{1}{2}}\right\|^2 \tag{24}$$

where $\mathbf{F}^{(n)}$ is the optimal solution at the nth iteration.

By replacing $\|\mathbf{g}_k^\dagger \mathbf{F} \mathbf{R}^{(1/2)}\|^2$ in (20b) by the lower bound [i.e., the right hand side of (24)], we solve the following convex problem at the $(n+1)$th iteration:

$$\min_{\mathbf{F}} \ \left\|\hat{\mathbf{d}}^\dagger \mathbf{F} \mathbf{S}^{\frac{1}{2}}\right\|^2 - 2\Re\left\{\hat{\mathbf{d}}^\dagger \mathbf{F} \mathbf{t}\right\} \tag{25a}$$

$$\text{s.t.} \ \ \rho_k \left\|\mathbf{g}_k^\dagger \mathbf{F} \mathbf{T}_k^{\frac{1}{2}}\right\|^2 - 2\Re\left\{\mathbf{g}_k^\dagger \mathbf{F} \mathbf{R} \mathbf{F}^{(n)\dagger} \mathbf{g}_k\right\} + \tau_k \leq 0, \ \ k = 1, \ldots, K \tag{25b}$$

$$\left\|\mathbf{F}\left(\mathbf{S} + \mathbf{R} + \mathbf{H}_p \mathbf{V} \mathbf{V}^\dagger \mathbf{H}_p^\dagger\right)^{\frac{1}{2}}\right\|^2 \leq P_r \tag{25c}$$

where $\tau_k = \|\mathbf{g}_k^\dagger \mathbf{F}^{(n)} \mathbf{R}^{(1/2)}\|^2 + \rho_k \sigma_k^2$ is a constant. Note that constraint (25b) holds is a sufficient condition for constraint (20b) to hold, and the two constraints are equivalent when the optimization variable $\mathbf{F}$ converges.

The optimization problem (25) is a convex quadratically constrained quadratic problem (QCQP), and we can recast it as an SOCP as

$$\min_{\mathbf{F}, \theta} \ \theta \tag{26a}$$

$$\text{s.t.} \ \left\|\begin{bmatrix} \left(\hat{\mathbf{d}}^\dagger \mathbf{F} \mathbf{S}^{\frac{1}{2}}\right)^{\mathrm{T}} \\ \frac{2\Re\left\{\hat{\mathbf{d}}^\dagger \mathbf{F} \mathbf{t}\right\} + \theta - 1}{2} \end{bmatrix}\right\| \leq \frac{2\Re\left\{\hat{\mathbf{d}}^\dagger \mathbf{F} \mathbf{t}\right\} + \theta + 1}{2} \tag{26b}$$

$$\left\|\begin{bmatrix} \sqrt{\rho_k}\left(\mathbf{g}_k^\dagger \mathbf{F} \mathbf{T}_k^{\frac{1}{2}}\right)^{\mathrm{T}} \\ \frac{2\Re\left\{\mathbf{g}_k^\dagger \mathbf{F} \mathbf{R} \mathbf{F}^{(n)\dagger} \mathbf{g}_k\right\} - \tau_k - 1}{2} \end{bmatrix}\right\| \leq \frac{2\Re\left\{\mathbf{g}_k^\dagger \mathbf{F} \mathbf{R} \mathbf{F}^{(n)\dagger} \mathbf{g}_k\right\} - \tau_k + 1}{2}, \ \ k = 1, \ldots, K \tag{26c}$$

$$\left\|\text{vec}\left(\mathbf{F}\left(\mathbf{S} + \mathbf{R} + \mathbf{H}_p \mathbf{V} \mathbf{V}^\dagger \mathbf{H}_p^\dagger\right)^{\frac{1}{2}}\right)\right\| \leq \sqrt{P_r} \tag{26d}$$

which is solved effectively by using the interior-point algorithm [48] or the CVX software [49].

C. Optimizing the Transmit Variables $\{\boldsymbol{V}, \boldsymbol{w}, b_k\}$

Inserting (17) into (15) and given the other variables, the optimization problem with respective to the transmit variables

$\{\mathbf{V}, \mathbf{w}, b_k\}$ can be expressed as

$$\min_{\mathbf{V},\mathbf{w},\{b_k\}} \sum_{k=1}^{K}\left|\hat{\mathbf{d}}^{\dagger}\mathbf{F}\mathbf{h}_k\right|^2|b_k|^2-2\sum_{k=1}^{K}\Re\left\{\hat{\mathbf{d}}^{\dagger}\mathbf{F}\mathbf{h}_k b_k\right\} \tag{27a}$$

$$\text{s.t. } \mathbf{t}^{\dagger}\left(\mathbf{S}+\mathbf{H}_p\mathbf{w}\mathbf{w}^{\dagger}\mathbf{H}_p^{\dagger}+\mathbf{H}_p\mathbf{V}\mathbf{V}^{\dagger}\mathbf{H}_p^{\dagger}\right)^{-1}\mathbf{t}\leq\hat{\xi}_r \tag{27b}$$

$$\rho_k\sum_{i=1,i\neq k}^{K}\left|\mathbf{g}_k^{\dagger}\mathbf{F}\mathbf{h}_i\right|^2|b_i|^2+\rho_k\left\|\mathbf{g}_k^{\dagger}\mathbf{F}\mathbf{H}_p\mathbf{V}\right\|^2 -\left|\mathbf{g}_k^{\dagger}\mathbf{F}\mathbf{H}_p\mathbf{w}\right|^2+\hat{\sigma}_k^2\leq 0,\ k=1,\ldots,K \tag{27c}$$

$$\text{(13d)–(13f)} \tag{27d}$$

where $\hat{\xi}_r = K-\xi_r$ and $\hat{\sigma}_k^2=\rho_k\sigma_r^2\|\mathbf{g}_k^{\dagger}\mathbf{F}\|^2+\rho_k\sigma_k^2$ are constants.

Although the objective (27a) and the transmit power constraints (13d)–(13f) are convex, problem (27) is still nonconvex since the secure AirComp constraint (27b) and reliable multicasting constraint (27c) are nonconvex. To solve problem (27) effectively, in this section, we will propose an efficient penaltySCA method that converts the nonconvex problem into a differential convex (DC) form and then linearizes the nonconvex terms.

Because the nonconvex constraint (27b) contains the inverse operation of the matrix, it is very difficult to solve it directly. In order to facilitate the expression and handling, we introduce some slack variables $\{\mathbf{X}, \mathbf{Y}, \mathbf{T}, \mathbf{t}\}$ and combine (18) and (19), and then the constraint (27b) can be represented as

$$\mathbf{X}=[b_1\mathbf{h}_1,\ldots,b_K\mathbf{h}_K] \tag{28a}$$

$$\mathbf{T}=\left[\mathbf{X}\ \mathbf{H}_p\mathbf{V}\ \mathbf{H}_p\mathbf{w}\right] \tag{28b}$$

$$\mathbf{t}=\mathbf{H}\mathbf{b} \tag{28c}$$

$$\mathbf{Y}=\mathbf{T}\mathbf{T}^{\dagger} \tag{28d}$$

$$\mathbf{t}^{\dagger}\left(\sigma_r^2\mathbf{I}+\mathbf{Y}\right)^{-1}\mathbf{t}\leq\hat{\xi}_r \tag{28e}$$

where

$$\mathbf{H}=[\mathbf{F}\mathbf{h}_1,\ldots,\mathbf{F}\mathbf{h}_K] \tag{29}$$

$$\mathbf{b}=[b_1,\ldots,b_K]^{\mathrm{T}}. \tag{30}$$

Note that the equality constraints (28a)–(28c) are linear and thus are convex, and the left-hand side of the inequality constraint (28e) is a matrix fractional function [48], which means that (28e) is also convex and can be represented as a linear matrix inequality (LMI)

$$\begin{bmatrix}\hat{\xi}_r & \mathbf{t}^{\dagger}\\ \mathbf{t} & \sigma_r^2\mathbf{I}+\mathbf{Y}\end{bmatrix}\succeq\mathbf{0}. \tag{31}$$

However, the equality constraint (28d) is quadratic over $\mathbf{T}$ and thus is a nonconvex constraint. To tackle this equality constraint, we need the following result.

Lemma 1: The equality constraint $\mathbf{Y}=\mathbf{T}\mathbf{T}^{\dagger}$ is equivalent to

$$\begin{bmatrix}\mathbf{Y} & \mathbf{T}\\ \mathbf{T}^{\dagger} & \mathbf{I}\end{bmatrix}\succeq\mathbf{0} \tag{32}$$

$$\mathrm{tr}\left(\mathbf{Y}-\mathbf{T}\mathbf{T}^{\dagger}\right)\leq 0 \tag{33}$$

where the first constraint is an LMI and the second is a difference of convex constraint.

Proof: See the Appendix. ■

Substituting (27b) by (28), the optimization problem (27) can be expressed as

$$\min_{\mathbf{V},\mathbf{w},\{b_k\}} \sum_{k=1}^{K}\left|\hat{\mathbf{d}}^{\dagger}\mathbf{F}\mathbf{h}_k\right|^2|b_k|^2-2\sum_{k=1}^{K}\Re\left\{\hat{\mathbf{d}}^{\dagger}\mathbf{F}\mathbf{h}_k b_k\right\} \tag{34a}$$

$$\text{s.t. } \mathrm{tr}\left(\mathbf{Y}-\mathbf{T}\mathbf{T}^{\dagger}\right)\leq 0 \tag{34b}$$

$$-\left|\mathbf{g}_k^{\dagger}\mathbf{F}\mathbf{H}_p\mathbf{w}\right|^2+\hat{\sigma}_k^2\leq 0,\ k=1,\ldots,K \tag{34c}$$

$$\text{(27d), (28a)–(28c), (31), (32).} \tag{34d}$$

Now, we can employ the penalty method [48] to solve problem (34), where a penalty term is incorporated into the objective function in order to handle the nonconvex constraint

$$\min_{\mathbf{V},\mathbf{w},\{b_k\},\mathbf{X},\mathbf{Y},\mathbf{T},\mathbf{t}} \sum_{k=1}^{K}\left|\hat{\mathbf{d}}^{\dagger}\mathbf{F}\mathbf{h}_k\right|^2|b_k|^2 -2\sum_{k=1}^{K}\Re\left\{\hat{\mathbf{d}}^{\dagger}\mathbf{F}\mathbf{h}_k b_k\right\}+\varsigma\,\mathrm{tr}\left(\mathbf{Y}-\mathbf{T}\mathbf{T}^{\dagger}\right) \tag{35a}$$

$$\text{s.t. } -\left|\mathbf{g}_k^{\dagger}\mathbf{F}\mathbf{H}_p\mathbf{w}\right|^2+\hat{\sigma}_k^2\leq 0,\ k=1,\ldots,K \tag{35b}$$

$$\text{(27d), (28a)–(28c), (31), (32)} \tag{35c}$$

where $\varsigma>0$ is a sufficiently large penalty factor.

Heretofore, we have transformed the nonconvex constraint (27b) into a series of convex constraints and a DC penalty term. Nevertheless, the optimization problem (35) is nonconvex due to the last penalty term in (35a) and the nonconvex constraint (35b) [i.e., (27c)]. Since both (35a) and (35b) are in the DC form, we employ the SCA method to solve it. By the convexity, we first have

$$\mathrm{tr}\left(\mathbf{T}\mathbf{T}^{\dagger}\right)\geq 2\Re\left\{\mathrm{tr}\left(\mathbf{T}^{(n)\dagger}\mathbf{T}\right)\right\}-\mathrm{tr}\left(\mathbf{T}^{(n)}\mathbf{T}^{(n)\dagger}\right)\triangleq f^{(n)}(\mathbf{T}) \tag{36}$$

$$\left|\mathbf{g}_k^{\dagger}\mathbf{F}\mathbf{H}_p\mathbf{w}\right|^2\geq 2\Re\left\{\mathbf{w}^{(n)\dagger}\mathbf{H}_p^{\dagger}\mathbf{F}^{\dagger}\mathbf{g}_k\mathbf{g}_k^{\dagger}\mathbf{F}\mathbf{H}_p\mathbf{w}\right\} -\left|\mathbf{g}_k^{\dagger}\mathbf{F}\mathbf{H}_p\mathbf{w}^{(n)}\right|^2\triangleq g_k^{(n)}(\mathbf{w}) \tag{37}$$

where $(\mathbf{V}^{(n)}, \mathbf{w}^{(n)}, \{b_k^{(n)}\}, \mathbf{T}^{(n)})$ is the optimal solution at the nth iteration.

Note that $f^{(n)}(\mathbf{T})$ and $g_k^{(n)}(\mathbf{w})$ in (36) and (37) are two linear functions. Applying them to (35), at the $(n+1)$th iteration, we solve the following convex problem:

$$\min_{\mathbf{V},\mathbf{w},\{b_k\},\mathbf{X},\mathbf{Y},\mathbf{T},\mathbf{t}} \sum_{k=1}^{K}\left|\hat{\mathbf{d}}^{\dagger}\mathbf{F}\mathbf{h}_k\right|^2|b_k|^2 -2\sum_{k=1}^{K}\Re\left\{\hat{\mathbf{d}}^{\dagger}\mathbf{F}\mathbf{h}_k b_k\right\}+\varsigma\left(\mathrm{tr}(\mathbf{Y})-f^{(n)}(\mathbf{T})\right) \tag{38a}$$

$$\text{s.t. } \rho_k\sum_{i=1,i\neq k}^{K}\left|\mathbf{g}_k^{\dagger}\mathbf{F}\mathbf{h}_i\right|^2|b_i|^2+\rho_k\left\|\mathbf{g}_k^{\dagger}\mathbf{F}\mathbf{H}_p\mathbf{V}\right\|^2 -g_k^{(n)}(\mathbf{w})+\hat{\sigma}_k^2\leq 0,\ k=1,\ldots,K \tag{38b}$$

$$\text{(27d), (28a)–(28c), (31), (32).} \tag{38c}$$

Algorithm 1 Proposed BCD-PenaltySCA Algorithm to Solve Problem (13)

```
Initializing: $n = 0$, $\{\mathbf{F}^{(0)}, \mathbf{V}^{(0)}, \mathbf{w}^{(0)}, b_k^{(0)}\}$;
Repeat:
    Compute $\mathbf{d}^{(n+1)}$ and $\mathbf{a}^{(n+1)}$ according to (16) and (17);
    Solve the SOCP (26) to obtain $\mathbf{F}^{(n+1)}$;
    Solve the SDP (40) to obtain $\{\mathbf{V}^{(n+1)}, \mathbf{w}^{(n+1)}, b_k^{(n+1)}\}$;
    $n = n + 1$;
Until: Convergence.
```

Denote

$$\hat{\mathbf{g}}_k = \left[\mathbf{g}_k^\dagger \mathbf{F}\mathbf{h}_1, \ldots, \mathbf{g}_k^\dagger \mathbf{F}\mathbf{h}_{k-1}, 0, \mathbf{g}_k^\dagger \mathbf{F}\mathbf{h}_{k+1}, \ldots, \mathbf{g}_k^\dagger \mathbf{F}\mathbf{h}_K\right]^{\mathrm{T}} \tag{39}$$

we can rewrite the convex problem (38) as the following SDP:

$$\min_{\mathbf{V},\mathbf{w},\{b_k\},\mathbf{X},\mathbf{Y},\mathbf{T},\mathbf{t},\lambda} \quad \lambda \tag{40a}$$

$$\text{s.t.} \quad \left\| \begin{bmatrix} \left(\mathbf{d}^\dagger \mathbf{H}\right)^{\mathrm{T}} \odot \mathbf{b} \\ \frac{2\Re\left\{\hat{\mathbf{d}}^\dagger \mathbf{t}\right\} - \varsigma\left(\mathrm{tr}(\mathbf{Y}) - f^{(n)}(\mathbf{T})\right) + \lambda - 1}{2} \end{bmatrix} \right\| \leq \frac{2\Re\left\{\hat{\mathbf{d}}^\dagger \mathbf{t}\right\} - \varsigma\left(\mathrm{tr}(\mathbf{Y}) - f^{(n)}(\mathbf{T})\right) + \lambda + 1}{2} \tag{40b}$$

$$\left\| \begin{bmatrix} \sqrt{\rho_k}\hat{\mathbf{g}}_k \odot \mathbf{b} \\ \sqrt{\rho_k}\left(\mathbf{g}_k^\dagger \mathbf{F}\mathbf{H}_p \mathbf{V}\right)^{\mathrm{T}} \\ \frac{g_k^{(n)}(\mathbf{w}) - \hat{\sigma}_k^2 - 1}{2} \end{bmatrix} \right\| \leq \frac{g_k^{(n)}(\mathbf{w}) - \hat{\sigma}_k^2 + 1}{2}, \quad k = 1, \ldots, K \tag{40c}$$

$$|b_k| \leq \sqrt{P_k}, \quad k = 1, \ldots, K \tag{40d}$$

$$\left\| \begin{bmatrix} \mathbf{w} \\ \mathrm{vec}(\mathbf{V}) \end{bmatrix} \right\| \leq \sqrt{P_{ap}} \tag{40e}$$

$$\left\| \begin{bmatrix} \mathrm{vec}(\mathbf{H} \odot \mathbf{b}) \\ \mathbf{F}\mathbf{H}_p \mathbf{w} \\ \mathrm{vec}\left(\mathbf{F}\mathbf{H}_p \mathbf{V}\right) \\ \sigma_r \mathrm{vec}(\mathbf{F}) \end{bmatrix} \right\| \leq \sqrt{P_r} \tag{40f}$$

$$\text{(28a)} - \text{(28c)}, \quad \text{(31)}, \text{(32)} \tag{40g}$$

which is solved effectively by using the interior-point algorithm [48] or the CVX software [49].

D. Proposed Algorithm

In the proposed BCD-penaltySCA algorithm, we solve problem (13) by alternately updating the blocks $\{\mathbf{d}, \mathbf{a}\}$, $\{\mathbf{F}\}$ and $\{\mathbf{V}, \mathbf{w}, b_k\}$ until convergence. The solution of updating $\{\mathbf{d}, \mathbf{a}\}$ is given in closed form. The subproblems of updating $\{\mathbf{F}\}$ and $\{\mathbf{V}, \mathbf{w}, b_k\}$ are solved by applying the SCA and the proposed penaltySCA, respectively. The details of the proposed BCD-penaltySCA algorithm are summarized in Algorithm 1.

Remark 2 (Convergence Analysis of Algorithm 1): By (16), (17), (26), and (40), we update the blocks $\{\mathbf{d}, \mathbf{a}\}$, $\{\mathbf{F}\}$ and $\{\mathbf{V}, \mathbf{w}, b_k\}$ iteratively. Because (16) and (17) have the closed form and problems (26) and (40) are convex, updating $\{\mathbf{d}, \mathbf{a}\}$, $\{\mathbf{F}\}$ and $\{\mathbf{V}, \mathbf{w}, b_k\}$ iteratively will only decease or maintain the objective value of problem (13). By updating $\{\mathbf{d}, \mathbf{a}\}$, $\{\mathbf{F}\}$ and $\{\mathbf{V}, \mathbf{w}, b_k\}$ iteratively, we obtain a monotonically decreasing sequence of the objective values of problem (13) which is lower bounded by zero since MSE is positive. Therefore, the proposed algorithm (Algorithm 1) converges.

Remark 3 (Complexity of Algorithm 1): The main computational burden of Algorithm 1 is from solving the SOCP (26) and SDP (40), whose complexity is about $\mathcal{O}(K^{0.5}N_m^6 \log(1/\varepsilon))$ and $\mathcal{O}(K^{0.5}N_m^7 \log(1/\varepsilon))$ where $N_m = \max(N_r, N_p)$ [50]. Thus, the complexity of Algorithm 1 is about $\mathcal{O}(L_1 K^{0.5}N_m^7 \log(1/\varepsilon))$ where L_1 is the iterative number for the convergence of Algorithm 1.

Remark 4 (Limitation of the Proposed Algorithm): In general, the proposed BCD-penaltySCA algorithm can only obtain the local optimal solution to problem (13). Although the outer polyblock approximation algorithm in [51] is employed to obtain a global optimal solution for a DC program, it cannot be applied to obtain a globally optimal solution to the DC program (35) due to the nonconvex quadratic and LMI constraints in problem (35). Since problem (35) is a subproblem of the original problem (13), how to obtain a global optimal solution to problem (13) is still unknown. In addition, the penalty factor ς needs to be selected appropriately according to a specific problem to obtain a faster convergence speed.

V. Robust Optimization With Imperfect CSI

In the above section, we consider the available CSI about $\{\mathbf{h}_k\}$ and $\{\mathbf{g}_k\}$ is perfect. Here, a more practical scenario is considered where the available CSI about $\{\mathbf{h}_k\}$ and $\{\mathbf{g}_k\}$ is imperfect, which may occur when the sensors are located in the mobile devices. As in [52], [53], and [54], we employ the worst case criterion to model the imperfect CSI, i.e., $\mathbf{h}_k$ and $\mathbf{g}_k$ are given by

$$\mathcal{H}_k = \left\{\mathbf{h}_k | \mathbf{h}_k = \bar{\mathbf{h}}_k + \Delta_{h,k}, ||\Delta_{h,k}|| \leq \epsilon_{h,k}\right\} \tag{41}$$

$$\mathcal{G}_k = \left\{\mathbf{g}_k | \mathbf{g}_k = \bar{\mathbf{g}}_k + \Delta_{g,k}, ||\Delta_{g,k}|| \leq \epsilon_{g,k}\right\} \tag{42}$$

where $\bar{\mathbf{h}}_k$ and $\bar{\mathbf{g}}_k$ is the estimated channel, and $\Delta_{h,k}$ and $\Delta_{g,k}$ represents the channel error, which is norm bounded with a corresponding given radius $\epsilon_{h,k}$ and $\epsilon_{g,k}$.

In the second time slot, because of the imperfect CSI, the self-interference term at the received signal y_k given in (9) cannot be removed completely. Only part of the self-interference term is removed, thus the remaining received signal at the kth sensor is

$$\tilde{\tilde{y}}_k = \mathbf{g}_k^\dagger \mathbf{F}\left(\sum_{i=1, i\neq k}^{K} \mathbf{h}_i b_i s_i + \mathbf{H}_p(\mathbf{w}s_c + \mathbf{V}\mathbf{z}) + \mathbf{n}_r\right) + n_k + \left(\Delta_{g,k}^\dagger \mathbf{F}\bar{\mathbf{h}}_k + \bar{\mathbf{g}}_k^\dagger \mathbf{F}\Delta_{h,k} + \Delta_{g,k}^\dagger \mathbf{F}\Delta_{h,k}\right) b_k s_k. \tag{43}$$

From (43), the SINR to decode the multicasting signal s_c at the kth sensor is

$$\tilde{\gamma}_k = \frac{\left|\mathbf{g}_k^\dagger \mathbf{F}\mathbf{H}_p \mathbf{w}\right|^2}{\sum_{i=1, i\neq k}^{K} \left|\mathbf{g}_k^\dagger \mathbf{F}\mathbf{h}_i b_i\right|^2 + \left\|\mathbf{g}_k^\dagger \mathbf{F}\mathbf{H}_p \mathbf{V}\right\|^2 + \sigma_r^2 \left\|\mathbf{g}_k^\dagger \mathbf{F}\right\|^2 + \tilde{\sigma}_k^2} \tag{44}$$

where

$$\tilde{\sigma}_k^2 = \left|\Delta_{g,k}^\dagger \mathbf{F}\bar{\mathbf{h}}_k + \bar{\mathbf{g}}_k^\dagger \mathbf{F}\Delta_{h,k} + \Delta_{g,k}^\dagger \mathbf{F}\Delta_{h,k}\right|^2 |b_k|^2 + \sigma_k^2. \tag{45}$$

Based on (13), (41), (42), and (44), the optimization problem (10) can be further expressed as

$$\begin{aligned}
&\min_{\mathbf{F},\mathbf{V},\mathbf{w},\mathbf{d},\{b_k\}} \ \max_{\mathbf{h}_k\in\mathcal{H}_k} \ \mathrm{MSE}_p && (46a)\\
&\text{s.t.} \min_{\mathbf{a}} \ \min_{\mathbf{h}_k\in\mathcal{H}_k} \ \mathrm{MSE}_r \ge \xi_r && (46b)\\
&\qquad \tilde{\gamma}_k \ge \rho_k,\ \ k = 1,\ldots,K \ \forall \mathbf{h}_k \in \mathcal{H}_k &&\\
&\qquad \mathbf{g}_k \in \mathcal{G}_k && (46c)\\
&\qquad |b_k|^2 \le P_k,\ \ k = 1,\ldots,K && (46d)\\
&\qquad \|\mathbf{w}\|^2 + \|\mathbf{V}\|^2 \le P_{ap} && (46e)\\
&\qquad \sum_{k=1}^{K} \|\mathbf{F}\mathbf{h}_k b_k\|^2 + \left\|\mathbf{F}\mathbf{H}_p\mathbf{w}\right\|^2 + \left\|\mathbf{F}\mathbf{H}_p\mathbf{V}\right\|^2 &&\\
&\qquad + \sigma_r^2\|\mathbf{F}\|^2 \le P_r \ \forall \ \mathbf{h}_k \in \mathcal{H}_k. && (46f)
\end{aligned}$$

Problem (46) is a robust optimization problem and is more challenging than its counterpart with perfect CSI, i.e., problem (13), mainly due to the following two reasons. One is that the existing channel errors $\{\Delta_{h,k}\}$ and $\{\Delta_{g,k}\}$ make the problem highly nonconvex even when the BCD algorithm is applied. The other is that the constraints become semi-infinite and we cannot deal with them like the deterministic constraints.

A. Deterministic Problem Transformation

To make the robust problem (46) tractable, a key step is to eliminate the channel errors $\{\Delta_{h,k}\}$ and $\{\Delta_{g,k}\}$. In what follows, we successively handle the terms with channel uncertainty and convert the robust problem into a deterministic problem.

1) Transformation of the Objective Function (46): By using the epigraph reformulation [48], the objective (46a) can be rewritten as

$$\begin{aligned}
&\min_{\mathbf{F},\mathbf{V},\mathbf{w},\mathbf{d},\{b_k\},t} \ t && (47a)\\
&\quad \text{s.t.}\ \ \mathrm{MSE}_p \le t \ \forall \ \mathbf{h}_k \in \mathcal{H}_k. && (47b)
\end{aligned}$$

Obviously, the constraint (47b) is still hard to deal with because it is a semi-infinite constraint. In the following, we make our efforts to convert this semi-infinite constraint to the finite one by eliminating the channel errors $\{\Delta_{h,k}\}$.

For proceeding, we reexpress the MSE_p in the constraint (47b) as

$$\mathrm{MSE}_p = \|\boldsymbol{\psi}_p\|^2 \tag{48}$$

where

$$\boldsymbol{\psi}_p \triangleq \bar{\boldsymbol{\psi}}_p + \sum_{k=1}^{K} \Psi_{p,k}\Delta_{h,k} \tag{49}$$

$$\bar{\boldsymbol{\psi}}_p = \begin{bmatrix} \mathbf{d}^\dagger\mathbf{G}_p\mathbf{F}\bar{\mathbf{h}}_1 b_1 - 1\\ \vdots \\ \mathbf{d}^\dagger\mathbf{G}_p\mathbf{F}\bar{\mathbf{h}}_K b_K - 1\\ \mathrm{vec}\left(\sigma_r\mathbf{d}^\dagger\mathbf{G}_p\mathbf{F}\right)\\ \sigma_p\mathbf{d} \end{bmatrix} \tag{50}$$

and

$$\Psi_{p,k} = \begin{bmatrix} \mathbf{0}\\ \mathbf{d}^\dagger\mathbf{G}_p\mathbf{F}b_k\\ \mathbf{0} \end{bmatrix}. \tag{51}$$

By employing the Schur complement lemma [55], the constraint (47b) is represented as

$$\begin{bmatrix} t & \bar{\boldsymbol{\psi}}_p^\dagger\\ \bar{\boldsymbol{\psi}}_p & \mathbf{I} \end{bmatrix} \succeq -\sum_{k=1}^{K}\begin{bmatrix} 0 & \left(\Psi_{p,k}\Delta_{h,k}\right)^\dagger\\ \Psi_{p,k}\Delta_{h,k} & \mathbf{0} \end{bmatrix}$$
$$\forall \ \ ||\ \Delta_{h,k}|| \le \epsilon_{h,k},\ \ k = 1,\ldots,K. \tag{52}$$

Constraint (52) still contains the channel errors. In order to eliminate these errors, we need the following sign-definiteness lemma.

Lemma 2 [54]: Assume $\mathbf{C}$ is a Hermitian matrix and given arbitrary matrices $\{\mathbf{U}_i, \mathbf{V}_i\}_{i=1}^N$, the following semi-infinite LMI:

$$\mathbf{C} \succeq \sum_{i=1}^{N}\left(\mathbf{U}_i^\dagger\mathbf{X}_i\mathbf{V}_i + \mathbf{V}_i^\dagger\mathbf{X}_i^\dagger\mathbf{U}_i\right) \ \forall \ \mathbf{X}_i : \|\mathbf{X}_i\| \le \epsilon_i$$

holds if and only if there are $\{\lambda_i \ge 0\}_{i=1}^N$ such that

$$\begin{bmatrix} \mathbf{C} - \sum\limits_{i=1}^{N}\lambda_i\mathbf{V}_i^\dagger\mathbf{V}_i & -\epsilon_1\mathbf{U}_1^\dagger & \cdots & -\epsilon_N\mathbf{U}_N^\dagger\\ -\epsilon_1\mathbf{U}_1 & \lambda_1\mathbf{I} & \cdots & \mathbf{0}\\ \vdots & \vdots & \ddots & \vdots\\ -\epsilon_N\mathbf{U}_N & \mathbf{0} & \cdots & \lambda_N\mathbf{I} \end{bmatrix} \succeq \mathbf{0}.$$

Choosing the parameters appropriately in Lemma 2 as

$$\mathbf{C} = \begin{bmatrix} t & \bar{\boldsymbol{\psi}}_p^\dagger\\ \bar{\boldsymbol{\psi}}_p & \mathbf{I} \end{bmatrix},\ \ \mathbf{V}_k = [-1\ \ \mathbf{0}],\ \ \mathbf{U}_k = \begin{bmatrix}\mathbf{0} \ \ \Psi_{p,k}^\dagger\end{bmatrix} \tag{53}$$

we can rewrite the constraint (52), i.e., (47b), as an LMI

$$\begin{bmatrix} \begin{bmatrix} t - \sum\limits_{k=1}^{K}\mu_{p,k} & \bar{\boldsymbol{\psi}}_p^\dagger\\ \bar{\boldsymbol{\psi}}_p & \mathbf{I} \end{bmatrix} & \Omega_p^\dagger\\ \Omega_p & \mathrm{diag}\left(\left\{\mu_{p,k}\mathbf{I}\right\}_{k=1}^K\right) \end{bmatrix} \succeq \mathbf{0} \tag{54}$$

where $\Omega_p = [-\epsilon_{h,1}\mathbf{U}_1^{\mathrm{T}},\ldots,-\epsilon_{h,K}\mathbf{U}_K^{\mathrm{T}}]^{\mathrm{T}}$.

2) Transformation of the Constraint (46b): To make the constraint (46b) tractable, we introduce slack variables and transform the constraints involving channel uncertainty into a more straightforward form

$$\min_{\mathbf{a}} \ \max_{\{\tau_k\}} \ \sum_{k=1}^{K}\tau_k + \left|\mathbf{a}^\dagger\mathbf{H}_p\mathbf{w}\right|^2 + \left\|\mathbf{a}^\dagger\mathbf{H}_p\mathbf{V}\right\|^2 + \sigma_r^2\|\mathbf{a}\|^2 \ge \xi_r \tag{55}$$

$$\left|\mathbf{a}^\dagger\mathbf{h}_k b_k - 1\right|^2 \ge \tau_k \ \forall \ \mathbf{h}_k \in \mathcal{H}_k. \tag{56}$$

Due to the infinite number of constraints parameterized by $\mathbf{h}_k$, we need to borrow the following S-Lemma.

Lemma 3 [56]: Define the functions

$$h_l(\mathbf{y}) = \mathbf{y}^\dagger\mathbf{C}_l\mathbf{y} + 2\Re\left\{\mathbf{b}_l^\dagger\mathbf{y}\right\} + a_l,\ \ l = 1, 2$$

where $\mathbf{C}_l = \mathbf{C}_l^\dagger$. The implication $h_1(\mathbf{y}) \leq 0 \Rightarrow h_2(\mathbf{y}) \leq 0$ holds if and only if there exists $\zeta \geq 0$ such that

$$\zeta \begin{bmatrix} \mathbf{C}_1 & \mathbf{b}_1 \\ \mathbf{b}_1^\dagger & a_1 \end{bmatrix} - \begin{bmatrix} \mathbf{C}_2 & \mathbf{b}_2 \\ \mathbf{b}_2^\dagger & a_2 \end{bmatrix} \succeq \mathbf{0}$$

provided that there exists a point $\mathbf{y}_0$ such that $h_1(\mathbf{y}_0) < 0$.

Applying Lemma 3 again to the constraint (56), we have

$$\begin{bmatrix} \lambda_k \mathbf{I} + \theta_k \Upsilon & \theta_k \Upsilon^\dagger \bar{\mathbf{h}}_k + b_k^* \mathbf{a} \\ \theta_k \bar{\mathbf{h}}_k^\dagger \Upsilon + b_k \mathbf{a}^\dagger & -\tilde{\tau}_k - \epsilon_{h,k}^2 \lambda_k + \theta_k \bar{\mathbf{h}}_k^\dagger \Upsilon \bar{\mathbf{h}}_k \end{bmatrix} \succeq \mathbf{0} \quad (57)$$

where $\tilde{\tau}_k = \tau_k + 2\Re\{b_k \mathbf{a}^\dagger \bar{\mathbf{h}}_k\} - 1$, and

$$\Upsilon = \mathbf{a}\mathbf{a}^\dagger, \;\; \theta_k = b_k b_k^*. \quad (58)$$

3) Transformation of (46c): Since there are fractions in constraint (46c) and both the numerator and denominator contain channel uncertainty, we cannot rewrite the constraints as the form that the norm of a vector is less than or equal to some value and apply the Lemma 2 to convert them into LMIs. By introducing the slack variables $\{\omega_k\}$, we rewrite (46c) as

$$\sum_{i=1, i\neq k}^{K} \left|\mathbf{g}_k^\dagger \mathbf{F}\mathbf{h}_i b_i\right|^2 + \left\|\mathbf{g}_k^\dagger \mathbf{F}\mathbf{H}_p \mathbf{V}\right\|^2 + \sigma_r^2 \left\|\mathbf{g}_k^\dagger \mathbf{F}\right\|^2 + \tilde{\sigma}_k^2 \leq \omega_k \quad (59a)$$

$$\left|\mathbf{g}_k^\dagger \mathbf{F}\mathbf{H}_p \mathbf{w}\right|^2 \geq \rho_k \omega_k \quad (59b)$$

for all $\mathbf{h}_k \in \mathcal{H}_k, \;\; \mathbf{g}_k \in \mathcal{G}_k, K = 1, \ldots, K$.

For (59a), neglecting higher order error terms, we can rewrite it as

$$\|\boldsymbol{\psi}_k\|^2 \leq \omega_k \quad (60)$$

where

$$\boldsymbol{\psi}_k \triangleq \bar{\boldsymbol{\psi}}_k + \sum_{i=1}^{K} \Psi_{k,i} \Delta_{h,i} + \Psi_g \Delta_{g,k} \quad (61)$$

$$\bar{\boldsymbol{\psi}}_k = \begin{bmatrix} \bar{\mathbf{g}}_k^\dagger \mathbf{F}\bar{\mathbf{h}}_1 b_1 \\ \vdots \\ \bar{\mathbf{g}}_k^\dagger \mathbf{F}\bar{\mathbf{h}}_{k-1} b_{k-1} \\ 0 \\ \bar{\mathbf{g}}_k^\dagger \mathbf{F}\bar{\mathbf{h}}_{k+1} b_{k+1} \\ \vdots \\ \bar{\mathbf{g}}_k^\dagger \mathbf{F}\bar{\mathbf{h}}_K b_K \\ \mathbf{V}^\dagger \mathbf{H}_p^\dagger \mathbf{F}^\dagger \bar{\mathbf{g}}_k \\ \sigma_r \mathbf{F}^\dagger \bar{\mathbf{g}}_k \end{bmatrix} \quad (62)$$

and

$$\Psi_{k,i} = \begin{bmatrix} \mathbf{0} \\ \bar{\mathbf{g}}_k^\dagger \mathbf{F} b_i \\ \mathbf{0} \end{bmatrix}, \quad \Psi_g = \begin{bmatrix} b_1^* \bar{\mathbf{h}}_1^\dagger \mathbf{F}^\dagger \\ \vdots \\ b_K^* \bar{\mathbf{h}}_K^\dagger \mathbf{F}^\dagger \\ \mathbf{V}^\dagger \mathbf{H}_p^\dagger \mathbf{F}^\dagger \\ \sigma_r \mathbf{F}^\dagger \end{bmatrix}. \quad (63)$$

Then, (60), i.e., (59a), is rewritten as an LMI

$$\begin{bmatrix} \begin{bmatrix} \omega_k - \sum_{i=1}^{K+1} \mu_{k,i} & \bar{\boldsymbol{\psi}}_k^\dagger \\ \bar{\boldsymbol{\psi}}_k & \mathbf{I} \end{bmatrix} & \Omega_k^\dagger \\ \Omega_k & \operatorname{diag}\left(\{\mu_{k,i}\mathbf{I}\}_{i=1}^{K+1}\right) \end{bmatrix} \succeq \mathbf{0} \quad (64)$$

where

$$\mathbf{U}_{k,i} = \left[\mathbf{0} \;\; \Psi_{k,i}^\dagger\right], i = 1, \ldots, K, \;\; \mathbf{U}_{k,K+1} = \left[\mathbf{0} \;\; \Psi_g^\dagger\right] \quad (65)$$

$$\Omega_k = \left[-\epsilon_{h,1}\mathbf{U}_{k,1}^{\mathrm{T}}, \ldots, -\epsilon_{h,K}\mathbf{U}_{k,K}^{\mathrm{T}}, -\epsilon_{g,k}\mathbf{U}_{k,K+1}^{\mathrm{T}}\right]^{\mathrm{T}}. \quad (66)$$

For (59b), its form is similar to the constraint (56), and we apply the Lemma 3 again. For ease of exposition, we insert $\mathbf{g}_k = \bar{\mathbf{g}}_k + \Delta_{g,k}$ into (59b) and recast it as

$$\begin{cases} \forall \; \Delta_{g,k}{:}\Delta_{g,k}^\dagger \Delta_{g,k} - \epsilon_{g,k}^2 \leq 0, \\ -\Delta_{g,k}^\dagger \Gamma \Delta_{g,k} - 2\Re\left\{\bar{\mathbf{g}}_k^\dagger \Gamma \Delta_{g,k}\right\} - \bar{\mathbf{g}}_k^\dagger \Gamma \bar{\mathbf{g}}_k + \rho_k \omega_k \leq 0 \end{cases} \quad (67)$$

where

$$\Gamma = \mathbf{F}\mathbf{H}_p \mathbf{w}\mathbf{w}^\dagger \mathbf{H}_p^\dagger \mathbf{F}^\dagger. \quad (68)$$

Applying Lemma 3 to (67), we covert the constraint (59b) to an LMI

$$\begin{bmatrix} \zeta_k \mathbf{I} + \Gamma & \Gamma^\dagger \bar{\mathbf{g}}_k \\ \bar{\mathbf{g}}_k^\dagger \Gamma & -\epsilon_{g,k}^2 \zeta_k - \rho_k \omega_k + \bar{\mathbf{g}}_k^\dagger \Gamma \bar{\mathbf{g}}_k \end{bmatrix} \succeq \mathbf{0}. \quad (69)$$

4) Transformation of (46f): Similar to (47b), we can use Lemma 2 again and rewrite the robust transmit power constraint (46f) as an LMI

$$\begin{bmatrix} \begin{bmatrix} P_r - \sum_{k=1}^{K} \mu_{r,k} & \bar{\boldsymbol{\psi}}_r^\dagger \\ \bar{\boldsymbol{\psi}}_r & \mathbf{I} \end{bmatrix} & \Omega_r^\dagger \\ \Omega_r & \operatorname{diag}\left(\{\mu_{r,k}\mathbf{I}\}_{k=1}^{K}\right) \end{bmatrix} \succeq \mathbf{0} \quad (70)$$

where

$$\bar{\boldsymbol{\psi}}_r = \begin{bmatrix} \mathbf{F}\bar{\mathbf{h}}_1 b_1 \\ \vdots \\ \mathbf{F}\bar{\mathbf{h}}_K b_K \\ \mathbf{F}\mathbf{H}_p \mathbf{w} \\ \operatorname{vec}\left(\mathbf{F}\mathbf{H}_p \mathbf{V}\right) \\ \sigma_r \operatorname{vec}(\mathbf{F}) \end{bmatrix} \quad (71)$$

and

$$\Psi_{r,k} = \left[\mathbf{0} \;\; \mathbf{F}^{\mathrm{T}} b_k \;\; \mathbf{0}\right]^{\mathrm{T}}, \;\; \mathbf{U}_{r,k} = \left[\mathbf{0} \;\; \Psi_{r,k}^\dagger\right] \quad (72)$$

$$\Omega_r = \left[-\epsilon_{h,1}\mathbf{U}_{r,1}^{\mathrm{T}}, \ldots, -\epsilon_{h,K}\mathbf{U}_{r,K}^{\mathrm{T}}\right]^{\mathrm{T}}. \quad (73)$$

5) Deterministic Problem: Until now, we have eliminated all the channel errors, and we can transform the robust problem (46) into a deterministic one

$$\min_{\mathbf{F}, \mathbf{V}, \Gamma, \Upsilon, \mathbf{w}, \mathbf{d}, \mathbf{a}, \{b_k, \mu_{p,k}, \mu_{r,k}, \omega_k, \mu_{k,i}, \lambda_k, \tau_k, \zeta_k, \theta_k\}, t} \; t \quad (74a)$$

$$\text{s.t. } (46d), (46e), (54), (55), (57), (58), \; (64), (68), (69), (70). \quad (74b)$$

Although problem (74) have finite constraints, it is still hard to be solved because the variables are highly coupled. In the following, we propose an effective BCD-penaltySCA method to solve problem (74).

B. Proposed BCD-penaltySCA Method

Similar to the case of perfect CSI, we decouple the deterministic problem (74) into several subproblems and alternately solve these subproblems by convex technique or penaltySCA method.

1) Optimizing $\mathbf{d}$ and $\{\tau_k\}$: With given other variables, the optimization problem with respective to $\mathbf{d}$ and $\{\tau_k\}$ are two separated subproblems, shown as follows:

$$\min_{\mathbf{d},t,\{\mu_{p,k}\}} \; t \quad \text{s.t. (54)} \tag{75}$$

and

$$\max_{\{\tau_k,\lambda_k\}} \sum_{k=1}^{K} \tau_k \quad \text{s.t. (57).} \tag{76}$$

The two optimization problems are convex SDPs and can be effectively solved by using the interior-point algorithm [48] or the CVX software [49].

2) Optimizing $\mathbf{a}$ and Υ: With given other variables, the optimization problem with respective to $\mathbf{a}$ and Υ is

$$\min_{\mathbf{a},\Upsilon,\{\lambda_k\}} \left|\mathbf{a}^\dagger \mathbf{H}_p \mathbf{w}\right|^2 + \left\|\mathbf{a}^\dagger \mathbf{H}_p \mathbf{V}\right\|^2 + \sigma_r^2 \|\mathbf{a}\|^2 \tag{77a}$$

$$\text{s.t.} \;\; \Upsilon = \mathbf{a}\mathbf{a}^\dagger \tag{77b}$$

$$\text{(57).} \tag{77c}$$

Although the objective and (57) are convex, problem (77) is still nonconvex due to the equality constraint $\Upsilon = \mathbf{a}\mathbf{a}^\dagger$.

Using Lemma 1, the equality constraint $\Upsilon = \mathbf{a}\mathbf{a}^\dagger$ is equivalent to

$$\begin{bmatrix} \Upsilon & \mathbf{a} \\ \mathbf{a}^\dagger & 1 \end{bmatrix} \succeq \mathbf{0} \tag{78}$$

$$\operatorname{tr}(\Upsilon) - \|\mathbf{a}\|^2 \le 0. \tag{79}$$

Replacing the nonconvex constraint (77b) with (78) and (79), we can recast problem (77) as

$$\min_{\mathbf{a},\Upsilon,\{\lambda_k\}} \left|\mathbf{a}^\dagger \mathbf{H}_p \mathbf{w}\right|^2 + \left\|\mathbf{a}^\dagger \mathbf{H}_p \mathbf{V}\right\|^2 + \sigma_r^2 \|\mathbf{a}\|^2 \tag{80a}$$

$$\text{s.t.} \;\; \operatorname{tr}(\Upsilon) - \|\mathbf{a}\|^2 \le 0 \tag{80b}$$

$$\begin{bmatrix} \Upsilon & \mathbf{a} \\ \mathbf{a}^\dagger & 1 \end{bmatrix} \succeq \mathbf{0} \tag{80c}$$

$$\text{(57).} \tag{80d}$$

Now, we can apply the proposed penaltySCA method to solve (80), and in the $(n+1)$th iteration, we solve the following convex SDP:

$$\min_{\mathbf{a},\Upsilon,\{\lambda_k\}} \left|\mathbf{a}^\dagger \mathbf{H}_p \mathbf{w}\right|^2 + \left\|\mathbf{a}^\dagger \mathbf{H}_p \mathbf{V}\right\|^2 + \sigma_r^2 \|\mathbf{a}\|^2 + \varsigma\left(\operatorname{tr}(\Upsilon) - 2\Re\left\{\mathbf{a}^{(n)\dagger}\mathbf{a}\right\}\right) \tag{81a}$$

$$\text{s.t. (57), (80c).} \tag{81b}$$

3) Optimizing $\mathbf{F}$ and Γ: With given other variables, the optimization problem with respective to $\mathbf{F}$ and Γ is

$$\min_{\mathbf{F},\Gamma,\{\mu_{p,k},\mu_{r,k},\mu_{k,i},\omega_k,\zeta_k\},t} \; t \tag{82a}$$

$$\text{s.t.} \;\; \Gamma = \mathbf{F}\mathbf{H}_p\mathbf{w}\mathbf{w}^\dagger\mathbf{H}_p^\dagger\mathbf{F}^\dagger, \tag{82b}$$

$$\text{(54), (64), (69), (70).} \tag{82c}$$

Due to the equality constraint (82b), problem (82) is still nonconvex. Applying Lemma 1, we transform problem (82) into

$$\min_{\mathbf{F},\Gamma,\{\mu_{p,k},\mu_{r,k},\mu_{k,i},\omega_k,\zeta_k\},t} \; t \tag{83a}$$

$$\text{s.t.} \;\; \begin{bmatrix} \Gamma & \mathbf{F}\mathbf{H}_p\mathbf{w} \\ (\mathbf{F}\mathbf{H}_p\mathbf{w})^\dagger & 1 \end{bmatrix} \succeq \mathbf{0} \tag{83b}$$

$$\operatorname{tr}(\Gamma) - \left\|\mathbf{F}\mathbf{H}_p\mathbf{w}\right\|^2 \le 0 \tag{83c}$$

$$\text{(54), (64), (69) (70)} \tag{83d}$$

and then we can apply the proposed penaltySCA method to solve (83). Defining $\mathbf{a}_f = \mathbf{F}\mathbf{H}_p\mathbf{w}$, in the $(n+1)$th iteration, we solve the following convex SDP:

$$\min_{\mathbf{F},\Gamma,\{\mu_{p,k},\mu_{r,k},\mu_{k,i},\omega_k,\zeta_k\},t} \; t + \varsigma\left(\operatorname{tr}(\Gamma) - 2\Re\left\{\mathbf{a}_f^{(n)\dagger}\mathbf{a}_f\right\}\right) \tag{84a}$$

$$\text{s.t.} \;\; \text{(83b), (83d)} \tag{84b}$$

where $\mathbf{a}_f^{(n)} = \mathbf{F}^{(n)}\mathbf{H}_p\mathbf{w}$.

4) Optimizing $\mathbf{w}$, $\mathbf{V}$, and $\{b_k\}$: With given other variables, the optimization problem with respective to $\mathbf{w}$, $\mathbf{V}$ and $\{b_k\}$ is

$$\min_{\mathbf{w},\mathbf{V},\{b_k,\mu_{p,k},\mu_{r,k},\mu_{k,i},\omega_k,\zeta_k,\lambda_k,\theta_k\},t} \; t \tag{85a}$$

$$\text{s.t.} \;\; \left|\mathbf{a}^\dagger\mathbf{H}_p\mathbf{w}\right|^2 + \left\|\mathbf{a}^\dagger\mathbf{H}_p\mathbf{V}\right\|^2 \ge \hat{\hat{\xi}}_r \tag{85b}$$

$$\Gamma = \mathbf{F}\mathbf{H}_p\mathbf{w}\mathbf{w}^\dagger\mathbf{H}_p^\dagger\mathbf{F}^\dagger \tag{85c}$$

$$\theta_k = b_k b_k^* \tag{85d}$$

$$\text{(46d), (46e), (54), (57), (64), (69), (70)}$$

where $\hat{\hat{\xi}}_r = \xi_r - \sigma_r^2\|\mathbf{a}\|^2 - \sum_{k=1}^{K}\tau_k$.

There are two nonconvex constraints (85c) and (85d) in the optimization problem (85). Following the same technique above, we apply the proposed penaltySCA method to solve problem (85). In the $(n+1)$th iteration, the convex SDP that needs to be solved is given by

$$\min_{\mathbf{w},\mathbf{V},\{b_k,\mu_{p,k},\mu_{r,k},\mu_{k,i},\omega_k,\zeta_k,\lambda_k,\theta_k\},t} \; t + \varsigma\left(\theta_k - 2\Re\left\{b_k^{(n)*}b_k\right\}\right) + \varsigma\left(\operatorname{tr}(\Gamma) - 2\Re\left\{\mathbf{a}_w^{(n)\dagger}\mathbf{a}_f\right\}\right) \tag{86a}$$

$$\text{s.t.} \;\; 2\Re\left\{\mathbf{w}^{(n)\dagger}\mathbf{H}_p^\dagger\mathbf{a}\mathbf{a}^\dagger\mathbf{H}_p\mathbf{w} + \mathbf{a}^\dagger\mathbf{H}_p\mathbf{V}\mathbf{V}^{(n)\dagger}\mathbf{H}_p^\dagger\mathbf{a}\right\} \ge \hat{\hat{\xi}}_r + \left|\mathbf{a}^\dagger\mathbf{H}_p\mathbf{w}^{(n)}\right|^2 + \left\|\mathbf{a}^\dagger\mathbf{H}_p\mathbf{V}^{(n)}\right\|^2 \tag{86b}$$

$$\begin{bmatrix} \theta_k & b_k \\ b_k^* & 1 \end{bmatrix} \succeq \mathbf{0} \tag{86c}$$

$$\text{(83b), (85e)} \tag{86d}$$

where $\mathbf{a}_w^{(n)} = \mathbf{F}\mathbf{H}_p\mathbf{w}^{(n)}$.

5) Proposed Algorithm: The proposed BCD-penaltySCA algorithm for solving problem (46) is summarized in Algorithm 2.

Remark 5 (Convergence Analysis of Algorithm 2): Similar to the proposed algorithm (Algorithm 1), updating the variables in steps 3–6 of the proposed algorithm (Algorithm 2), we will obtain a monotonically decreasing sequence of the objective

Algorithm 2 Proposed BCD-PenaltySCA Algorithm to Solve Robust Problem (46)

```
Initializing: $n = 0$, $\{\mathbf{F}^{(0)}, \mathbf{V}^{(0)}, \mathbf{w}^{(0)}, \mathbf{a}^{(0)}, b_k^{(0)}\}$;
repeat
    Update $\mathbf{d}$ and $\{\tau_k\}$ by solving SDPs (75) and (76);
    Update $\mathbf{a}$ and $\Upsilon$ by solving SDP (81);
    Update $\mathbf{F}$ and $\Gamma$ by solving SDP (84);
    Update $\mathbf{w}$, $\mathbf{V}$ and $\{b_k\}$ by solving SDP (86);
    $n = n + 1$;
until Convergence.
```

values of problem (46) which is lower bounded by zero. Thus, the proposed algorithm (Algorithm 2) also converges.

Remark 6 (Complexity of Algorithm 2): The computational complexity of Algorithm 2 is mainly from solving the SDPs, whose complexity is about $\mathcal{O}(K^4 N_m^7 \log(1/\varepsilon))$ [50]. Thus, the computational complexity of Algorithm 2 is about $\mathcal{O}(L_2 K^4 N_m^7 \log(1/\varepsilon))$, where L_2 is the iterative number for the convergence of Algorithm 2.

Remark 7 (The Choice Between Robust and Approximate CSI): Another way to handle the channel uncertainty is to use the approximated CSI. As in [57], [58], and [59], the AP can act as a central processor to perform the optimization process. Thus, the AP has the ability to estimate the error caused by using the approximate CSI and decides whether to use Algorithm 1 to obtain a approximate solution with a lower computational complexity or Algorithm 2 to obtain a robust solution.

VI. Simulation Results

In this section, we provide the numerical results to verify the effectiveness of the proposed schemes for an untrusted relay-assisted AirComp network. As in [26], we assume that all the entries of the channel matrices/vectors $(\mathbf{H}_p, \mathbf{G}_p)$ and $(\mathbf{h}_k, \mathbf{g}_k)$ are independent and identically distributed complex Gaussian random variables with zero mean and variances d_0^{-n} and d_k^{-n}, respectively, where d_0 and d_k denote the distance from AP to relay and from relay to the kth sensor, respectively, where $n = 3.5$ is the path-loss exponent. For simplicity, we also assume that $d_0 = 1$, d_k is randomly generated from $[0.5, 1.5]$, and the powers of sensors and noises, the SINR thresholds, and the channel error radius are the same, i.e., $P_k = P_s$, $\sigma_r^2 = \sigma_p^2 = \sigma_k^2 = \sigma^2$, $\rho_k = \rho$, and $\epsilon_{h,k} = \epsilon_{g,k} = \epsilon$. We set the antenna number of the AP and relay is $N_p = N_r = 20$, and the number of sensors is $K = 10$ [30], [31].

A. Performances of Proposed Scheme for Perfect CSI

With perfect CSI available, we compare our proposed scheme with three benchmarks, including the zero-forcing reception and zero-forcing transmission scheme proposed in [60] (denoted as "ZFR-ZFT") and the maximal-ratio reception and maximal-ratio transmission scheme proposed in [61] (denoted as "MRR-MRT") and our proposed scheme with randomly generated AN (denoted as "RandAN"). If not specified, we set the transmit power budget of the sensors, relay and AP is $P_s/\sigma^2 = 10$ dB, $P_r/\sigma^2 = 30$ dB, and $P_{ap}/\sigma^2 = 40$ dB, the minimum MSE constraint at the relay is $\xi_r = 3$, and the SINR threshold at the sensors is $\rho = 2$ dB.

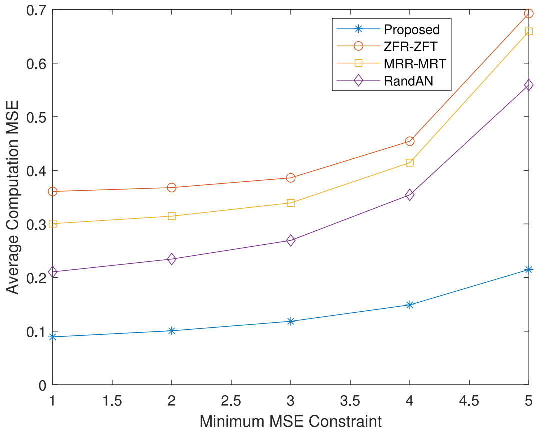

Fig. 2. Average computation MSE at the AP under different minimum MSE constraints at the relay.

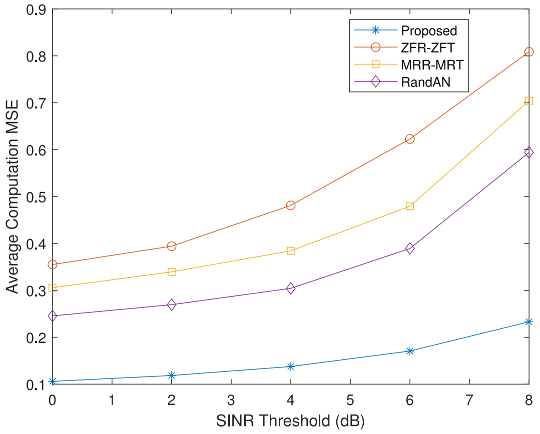

Fig. 3. Average computation MSE at the AP under different SINR thresholds at the sensors.

In Fig. 2 , we present the average computation MSE at the AP under different minimum MSE constraint at the relay. From Fig. 2 , we see that the proposed scheme has a significant reduction of the computation MSE at the AP over the "RandAN" scheme for different minimum MSE constraint at the relay, which indicates that the AN plays an important role to satisfy the secure AirComp constraint at the relay. Furthermore, the RandAN scheme outperforms the "ZFR-ZFT" and "MRR-MRT" schemes, which means that using specified relay beamforming schemes cannot reduce the computation MSE at the AP effectively. In Fig. 2 , we also see that when the minimum MSE constraint at the relay is not too large, the average computation MSE at the AP by all the schemes grows not too fast, however, when the minimum MSE constraint at the relay is larger, the average computation MSE at the AP has a relatively rapid increase, which is maybe because that when the minimum MSE constraint at the relay is too large, most of the power is allocated to satisfy the secure AirComp constraint and thus significantly increases the average computation MSE at the AP.

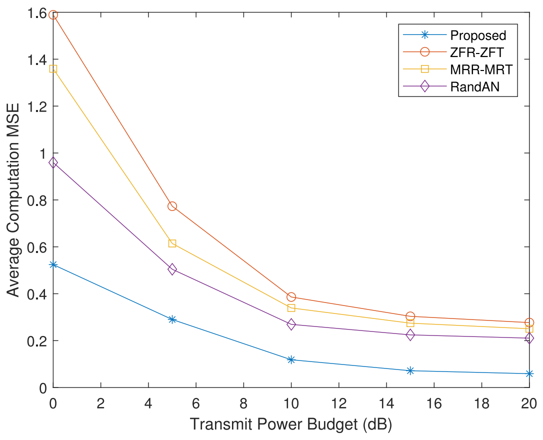

Fig. 4. Average computation MSE at the AP under different transmit power budgets of the sensors.

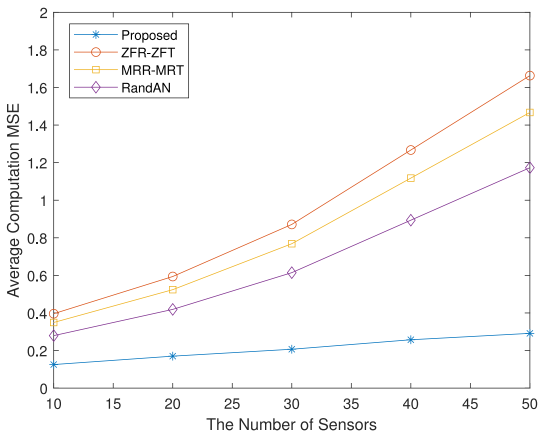

Fig. 5. Average computation MSE at the AP under different network configurations where $N_p = N_r = 1.1K$.

In Fig. 3, we present the average computation MSE at the AP under different SINR threshold at the sensors. From Fig. 3, we can see that the proposed scheme is better than all the other benchmarks for different reliable multicasting requirements. We can also see from Fig. 3 that as the SINR threshold at the sensors increases, the average computation MSE at the AP by all the schemes grows, which is because that as the SINR threshold at the sensors increases, more power has to be allocated to satisfy the reliable multicasting requirement and then less power is allocated to reduce the computation MSE at the AP.

In Fig. 4, we present the average computation MSE at the AP under different transmit power budget of the sensors. From Fig. 4, it is seen that for different transmit power budget of the sensors, the proposed scheme outperforms all the other benchmarks in terms of the average computation MSE at the AP. We can also see from Fig. 4 that as the transmit power budget of the sensors increases, the average computation MSE at the AP by all the schemes decreases, which is because as the transmit power of the sensors increases, the AP is easier to align the signals of all the sensors to compress the calculation error.

In Fig. 5, we investigate the average computation MSE at the AP under different network configuration. If we only increase the antenna number of AP and relay, N_p and N_r, the network obtains more antenna resources and has greater space freedom to optimize the beamforming design at the AP and relay, thus achieving better network performance [62]. But, if we only increase the number of sensors, K, the network needs to support more sensors to achieve reliable multicast, and at this time, fewer transmit antennas at the AP and relay are difficult to support more sensors [63]. Therefore, as the number of sensors increases, the antenna number of AP and relay should also increase accordingly, and in Fig. 5, we set $N_p = N_r = 1.1K$. From Fig. 5, we can see that as the network size increases, the average computation MSE at the AP by our proposed algorithm slowly increases. Thus, in order to meet the accuracy requirements of AirComp in practical applications, it is necessary to choose an appropriate network configuration.

TABLE I
COMPUTING TIME OF ALGORITHM 1 UNDER DIFFERENT NETWORK CONFIGURATIONS

(K, N_p, N_r)	Computing Time(s)
$(10, 11, 11)$	2.1
$(20, 22, 22)$	22.3
$(30, 33, 33)$	270.4
$(40, 44, 44)$	1317.5
$(50, 55, 55)$	4520.1

In Table I, we show the time to compute the solution by our proposed algorithm (Algorithm 1) under different network configuration. The simulations were performed in MATLAB on a Windows desktop with four Intel i3 cores and 16 GB of RAM. From Table I, we can see that as the network size increases, the computing time of our proposed algorithm (Algorithm 1) increases, which is because Algorithm 1 needs to solve a larger scale SOCP or SDP. In order to accelerate the computing of the optimization solutions, a feasible approach is to design a parallel algorithm for Algorithm 1, such as using the consensus alternating direction method of multipliers (ADMMs) [64] to solve the SOCP and using parallel primal-dual interior-point method [65] to solve the SDP. This is reserved for an interesting future work.

B. Performances of Proposed Scheme for Imperfect CSI

With imperfect CSI available, we compare our proposed robust scheme (dented as "Proposed Robust") with the proposed robust scheme with randomly generated AN (denoted as "Robust RandAN"), and the proposed scheme with perfect CSI is given which serves as a performance upper (or MSE lower) bound (dented as "Perfect CSI"). If not specified, the simulation parameter setting is the same as that in the case of perfect CSI, and a normalized channel error radius is used and set as 0.02.

In Fig. 6, we show the average worst case computation MSE at the AP under different minimum MSE constraint at the relay. From Fig. 6, we can see that the proposed robust scheme is close to the scheme with perfect CSI, which indicates that the proposed robust scheme can effectively reduce

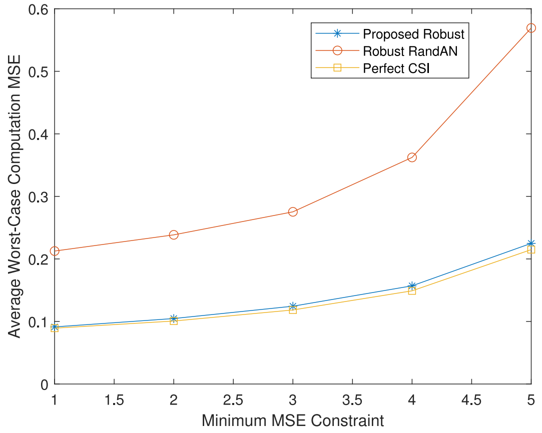

Fig. 6. Average worst case computation MSE at the AP under different minimum MSE constraints at the relay.

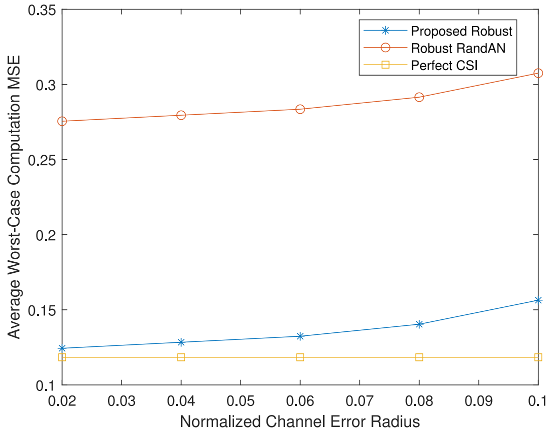

Fig. 7. Average worst case computation MSE at the AP under different normalized channel error radii.

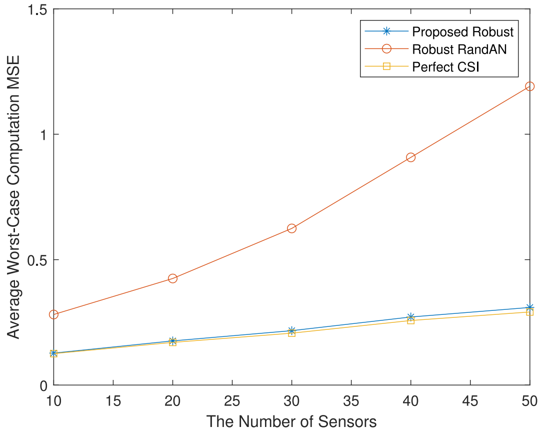

Fig. 8. Average worst case computation MSE at the AP under different network configurations where $N_p = N_r = 1.1K$.

TABLE II
COMPUTING TIME OF ALGORITHM 2 UNDER DIFFERENT NETWORK CONFIGURATIONS

(K, N_p, N_r)	Computing Time(s)
$(10, 11, 11)$	4.5
$(20, 22, 22)$	56.8
$(30, 33, 33)$	621.9
$(40, 44, 44)$	3162.1
$(50, 55, 55)$	9492.2

the performance loss raised by the channel errors. We can also see from Fig. 6 that the proposed robust scheme is better than the Robust RandAN scheme, which again shows the importance of the AN to satisfy the secure AirComp constraint even with the channel errors.

In Fig. 7, we show the average worst case computation MSE at the AP under different normalized channel error radius. From Fig. 7, it is seen that the average worst case computation MSE at the AP by the proposed robust scheme and the Robust RandAN scheme increases with the channel error radius increasing. We can also see from Fig. 7 that the performance gap between the proposed robust scheme, the Robust RandAN scheme and the Perfect CSI scheme tends to be greater when the channel error radius goes to be larger. This is because as the channel error radius increases, the robust secure AirComp constraint and reliable multicasting requirement become more strict and thus increase the worst case computation MSE.

In Fig. 8, we investigate the average worst case computation MSE at the AP under different network configuration, where we set $N_p = N_r = 1.1K$. From Fig. 8, we can see that as the network size increases, the average worst case computation MSE at the AP by our proposed robust scheme increases slowly. But the average worst case computation MSE by the Robust RandAN scheme grows more quickly, which is due to the fact that randomly generated AN becomes harder to satisfy the secure AirComp constraint at the relay and easier to raise the computation MSE at the AP. From Fig. 8, we can also see that the proposed robust scheme is close to the scheme with perfect CSI, which indicates again that the proposed robust scheme can effectively reduce the performance loss due to the channel errors.

In Table II, we show the time to compute the solution by our proposed robust algorithm (Algorithm 2) under different network configuration. The simulations were performed on the same desktop as Algorithm 1. From Table II, we can see that as the network size increases, the proposed robust algorithm (Algorithm 2) needs to solve larger scale SDPs, and thus the computing time increases. Similar to Algorithm 1, in order to accelerate the computing of the solutions by robust algorithm (Algorithm 2), a feasible approach is to design a parallel algorithm for Algorithm 2, e.g., using parallel primal-dual interior-point methods [65] to solve the SDPs.

VII. CONCLUSION

In this article, we have studied the joint optimization of secure AirComp and reliable multicasting in a MIMO untrusted two-way relay-assisted computation and communication networks, where AN is employed at the AP to interfere the relay for ensuring secure AirComp. We have formulated the optimization problems for two cases of perfect and imperfect CSI. We propose an effective BCD-penaltySCA algorithm for the case of perfect CSI, while for the case of imperfect CSI, we first transform the robust optimization problem to a deterministic problem and then employ the proposed BCD-penaltySCA algorithm to solve the reformulated deterministic problem. Numerical results have demonstrated that the proposed schemes outperform other benchmarks in terms of the computation distortion under the constraints of secure AirComp and reliable multicasting.

Appendix
Proof of Lemma 1

Applying the Schur complement, $\begin{bmatrix} \mathbf{Y} & \mathbf{T} \\ \mathbf{T}^{\dagger} & \mathbf{I} \end{bmatrix} \succeq \mathbf{0}$ equals

$$\mathbf{Y} - \mathbf{T}\mathbf{T}^{\dagger} \succeq \mathbf{0}. \tag{87}$$

Combining (87) with $\mathrm{tr}(\mathbf{Y} - \mathbf{T}\mathbf{T}^{\dagger}) \leq 0$, we can obtain

$$\mathrm{tr}\Big(\mathbf{Y} - \mathbf{T}\mathbf{T}^{\dagger}\Big) = 0. \tag{88}$$

From (87) and (88), we last have

$$\mathbf{Y} = \mathbf{T}\mathbf{T}^{\dagger}. \tag{89}$$

References

[1] O. Abari, H. Rahul, and D. Katabi, "Over-the-air function computation in sensor networks," 2016, *arXiv:1612.02307*.

[2] B. Nazer and M. Gastpar, "Computation over multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3498–3516, Oct. 2007.

[3] M. Gastpar, "Uncoded transmission is exactly optimal for a simple Gaussian 'sensor' network," *IEEE Trans. Inf. Theory*, vol. 54, no. 11, pp. 5247–5251, Nov. 2008.

[4] L. Chen, N. Zhao, Y. Chen, X. Qin, and F. R. Yu, "Computation over MAC: Achievable function rate maximization in wireless networks," *IEEE Trans. Commun.*, vol. 68, no. 9, pp. 5446–5459, Sep. 2020.

[5] S.-W. Jeon, C.-Y. Wang, and M. Gastpar, "Computation over Gaussian networks with orthogonal components," *IEEE Trans. Inf. Theory*, vol. 60, no. 12, pp. 7841–7861, Dec. 2014.

[6] W. Liu, X. Zang, Y. Li, and B. Vucetic, "Over-the-air computation systems: Optimization, analysis and scaling laws," *IEEE Trans. Wireless Commun.*, vol. 19, no. 8, pp. 5488–5502, Aug. 2020.

[7] X. Zang, W. Liu, Y. Li, and B. Vucetic, "Over-the-air computation systems: Optimal design with sum-power constraint," *IEEE Wireless Commun. Lett.*, vol. 9, no. 9, pp. 1524–1528, Sep. 2020.

[8] X. Cao, G. Zhu, J. Xu, and K. Huang, "Optimized power control for over-the-air computation in fading channels," *IEEE Trans. Wireless Commun.*, vol. 19, no. 11, pp. 7498–7513, Nov. 2020.

[9] G. Zhu and K. Huang, "MIMO over-the-air computation for high-mobility multimodal sensing," *IEEE Internet Things J.*, vol. 6, no. 4, pp. 6089–6103, Aug. 2019.

[10] D. Wen, G. Zhu, and K. Huang, "Reduced-dimension design of MIMO over-the-air computing for data aggregation in clustered IoT networks," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5255–5268, Nov. 2019.

[11] X. Chen, A. Liu, and M.-J. Zhao, "High-mobility multi-modal sensing for IoT network via MIMO AirComp: A mixed-timescale optimization approach," *IEEE Commun. Lett.*, vol. 24, no. 10, pp. 2295–2299, Oct. 2020.

[12] N. Yan, K. Wang, C. Pan, and K. K. Chai, "Private federated learning with misaligned power allocation via over-the-air computation," *IEEE Commun. Lett.*, vol. 26, no. 9, pp. 1994–1998, Sep. 2022.

[13] N. Zhang, M. Tao, J. Wang, and S. Shao, "Coded over-the-air computation for model aggregation in federated learning," *IEEE Commun. Lett.*, vol. 27, no. 1, pp. 160–164, Jan. 2023.

[14] P. Park, P. D. Marco, and C. Fischione, "Optimized over-the-air computation for wireless control systems," *IEEE Commun. Lett.*, vol. 26, no. 2, pp. 424–428, Feb. 2022.

[15] X. Jia, C. Zhang, J.-M. Kang, and I.-M. Kim, "Joint beamforming design and time allocation for wireless powered asymmetric two-way multirelay network," *IEEE Trans. Veh. Technol.*, vol. 67, no. 10, pp. 9641–9655, Oct. 2018.

[16] B. Mahboobi, S. Mehrizi, and M. Ardebilipour, "Multicast relay beamforming in CDMA networks: Nonregenerative approach," *IEEE Commun. Lett.*, vol. 19, no. 8, pp. 1418–1421, Aug. 2015.

[17] A. Almradi and K. A. Hamdi, "MIMO full-duplex relaying in the presence of co-channel interference," *IEEE Trans. Veh. Technol.*, vol. 66, no. 6, pp. 4874–4885, Jun. 2017.

[18] U. Rashid, H. D. Tuan, H. H. Kha, and H. H. Nguyen, "Joint optimization of source precoding and relay beamforming in wireless MIMO relay networks," *IEEE Trans. Commun.*, vol. 62, no. 2, pp. 488–499, Feb. 2014.

[19] V. S. Krishna and M. R. Bhatnagar, "A joint antenna and path selection technique in single-relay-based DF cooperative MIMO networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 3, pp. 1340–1353, Mar. 2016.

[20] M. M. Amiri, A. Olfat, and N. C. Beaulieu, "Novel beamforming scheme for multicasting in cooperative wireless networks with a multiple antenna relay," *IEEE Trans. Wireless Commun.*, vol. 14, no. 8, pp. 4482–4493, Aug. 2015.

[21] J. Lee and N. Al-Dhahir, "Exploiting sparsity for multiple relay selection with relay gain control in large AF relay networks," *IEEE Commun. Lett.*, vol. 2, no. 3, pp. 347–350, Jun. 2013.

[22] L. Yang, K. Qaraqe, E. Serpedin, and X. Gao, "Performance analysis of two-way relaying networks with the *N*th worst relay selection over various fading channels," *IEEE Trans. Veh. Technol.*, vol. 64, no. 7, pp. 3321–3327, Jul. 2015.

[23] X. Jiang, Z. Wu, Z. Yin, W. Yang, and Z. Yang, "Trajectory and communication design for UAV-relayed wireless networks," *IEEE Commun. Lett.*, vol. 8, no. 6, pp. 1600–1603, Dec. 2019.

[24] X. Jiang, Z. Wu, Z. Yin, and Z. Yang, "Joint power and trajectory design for UAV-relayed wireless systems," *IEEE Commun. Lett.*, vol. 8, no. 3, pp. 697–700, Jun. 2019.

[25] J.-Y. Wang, Y. Ma, R.-R. Lu, J.-B. Wang, M. Lin, and J. Cheng, "Hovering UAV-based FSO communications: Channel modelling, performance analysis, and parameter optimization," *IEEE J. Sel. Areas Commun.* , vol. 39, no. 10, pp. 2946–2959, Oct. 2021.

[26] F. Wang, J. Xu, V. K. N. Lau, and S. Cui, "Amplify-and-forward relaying for hierarchical over-the-air computation," *IEEE Trans. Wireless Commun.*, vol. 21, no. 12, pp. 10529–10543, Dec. 2022.

[27] S. Tang, H. Yomo, C. Zhang, and S. Obana, "Node scheduling for AF-based over-the-air computation," *IEEE Wireless Commun. Lett.*, vol. 11, no. 9, pp. 1945–1949, Sep. 2022.

[28] Z. Lin, H. Liu, and Y.-J. A. Zhang, "Relay-assisted cooperative federated learning," *IEEE Trans. Wireless Commun.*, vol. 21, no. 9, pp. 7148–7164, Sep. 2022.

[29] S. Tang, H. Yin, C. Zhang, and S. Obana, "Reliable over-the-air computation by amplify-and-forward based relay," *IEEE Access*, vol. 9, pp. 53333–53342, 2021.

[30] Y. Li, M. Jiang, G. Zhang, and M. Cui, "Joint optimization for multi-antenna AF-relay aided over-the-air computation," *IEEE Trans. Veh. Technol.*, vol. 71, no. 6, pp. 6744–6749, Jun. 2022.

[31] M. Jiang, Y. Li, G. Zhang, and M. Cui, "Joint beamforming optimization in multi-relay assisted MIMO over-the-air computation for multimodal sensing data aggregation," *IEEE Commun. Lett.*, vol. 25, no. 12, pp. 3937–3941, Dec. 2021.

[32] R. Jiang, S. Zhou, and K. Huang, "Achieving cooperative diversity in over-the-air computation via relay selection," in *Proc. IEEE VTC-Fall*, 2020, pp. 1–6.

[33] R. Jiang and S. Zhou, "Cluster-based cooperative digital over-the-air aggregation for wireless federated edge learning," in *Proc. IEEE/CIC ICCC*, 2020, pp. 887–892.

[34] F. Wu, L. Chen, N. Zhao, Y. Chen, F. R. Yu, and G. Wei, "Computation over multi-access channels: Multi-hop implementation and resource allocation," *IEEE Trans. Commun.*, vol. 69, no. 2, pp. 1038–1052, Feb. 2021.

[35] X. He and A. Yener, "Cooperation with an untrusted relay: A secrecy perspective," *IEEE Trans. Inf. Theory*, vol. 56, no. 8, pp. 3807–3827, Aug. 2010.

[36] A. Kuhestani, A. Mohammadi, and M. Mohammadi, "Joint relay selection and power allocation in large-scale MIMO systems with untrusted relays and passive eavesdroppers," *IEEE Trans. Inf. Forensics Security*, vol. 13, pp. 341–355, 2018.

[37] L. Lv, H. Jiang, Z. Ding, L. Yang, and J. Chen, "Secrecy-enhancing design for cooperative downlink and uplink NOMA with an untrusted relay," *IEEE Trans. Commun.*, vol. 68, no. 3, pp. 1698–1715, Mar. 2020.

[38] L. Lv, F. Zhou, J. Chen, and N. Al-Dhahir, "Secure cooperative communications with an untrusted relay: A NOMA-inspired jamming and relaying approach," *IEEE Trans. Inf. Forensics Security*, vol. 14, pp. 3191–3205, 2019.

[39] C. Jeong, I.-M. Kim, and D. I. Kim, "Joint secure beamforming design at the source and the relay for an amplify-and-forward MIMO untrusted relay system," *IEEE Trans. Signal Process.*, vol. 60, no. 1, pp. 310–325, Jan. 2012.

[40] J. Mo, M. Tao, Y. Liu, and R. Wang, "Secure beamforming for MIMO two-way communications with an untrusted relay," *IEEE Trans. Signal Process.*, vol. 62, no. 9, pp. 2185–2199, May 2014.

[41] S. Zhao, J. Liu, Y. Shen, X. Jiang, and N. Shiratori, "Secure beamforming for full-duplex MIMO two-way untrusted relay systems," *IEEE Trans. Inf. Forensics Security*, vol. 15, pp. 3775–3790, 2020.

[42] M. Frey, I. Bjelaković, and S. Stańczak, "Towards secure over-the-air computation," in *Proc. IEEE ISIT*, Jul. 2021, pp. 700–705.

[43] C. Hu, Q. Li, Q. Zhang, and J. Qin, "Secure transceiver design and power control for over-the-air computation networks," *IEEE Commun. Lett.*, vol. 26, no. 7, pp. 1509–1513, Jul. 2022.

[44] J. Wang, J. Mu, S. Wei, C. Jiang, and N. C. Beaulieu, "Statistical characterization of decryption errors in block-ciphered systems," *IEEE Trans. Commun.*, vol. 63, no. 11, pp. 4363–4376, Nov. 2015.

[45] N. Koblitz, "Elliptic curve cryptosystems," *Math. Comput.*, vol. 48, no. 177, pp. 203–209, Jan. 1987.

[46] M. Nassar, A. Erradi, and Q. M. Malluhi, "Paillier's encryption: Implementation and cloud applications," in *Proc. Int. Conf. Appl. Res. Comput. Sci. Eng.*, 2015, pp. 1–5.

[47] M. Shakiba-Herfeh, A. Chorti, and H. V. Poor, "Physical layer security: Authentication, integrity, and confidentiality," in *Physical Layer Security*. Cham, Switzerland: Springer, 2021, pp. 129–150.

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

[49] M. Grant and S. Boyd. "CVX: MATLAB software for disciplined convex programming." Mar. 2014. [Online]. Available: http://cvxr.com/cvx

[50] Y. Nesterov and A. Nemirovskii, *Interior Point Polynomial Algorithms in Convex Programming*. Philadelphia, PA, USA: SIAM, 1994.

[51] Y. Sun, D. W. K. Ng, Z. Ding, and R. Schober, "Optimal joint power and subcarrier allocation for full-duplex multicarrier non-orthogonal multiple access systems," *IEEE Trans. Commun.*, vol. 65, no. 3, pp. 1077–1091, Mar. 2017.

[52] F. Zhou, Z. Chu, H. Sun, R. Q. Hu, and L. Hanzo, "Artificial noise aided secure cognitive beamforming for cooperative MISO-NOMA using SWIPT," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 4, pp. 918–931, Apr. 2018.

[53] H. Sun, F. Zhou, R. Q. Hu, and L. Hanzo, "Robust beamforming design in a NOMA cognitive radio network relying on SWIPT," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 1, pp. 142–155, Jan. 2019.

[54] E. A. Gharavol and E. G. Larsson, "The sign-definiteness lemma and its applications to robust transceiver optimization for multiuser MIMO systems," *IEEE Trans. Signal Process.*, vol. 61, no. 2, pp. 238–252, Jan. 2013.

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

[56] A. Beck and Y. C. Eldar, "Strong duality in nonconvex quadratic optimization with two quadratic constraints," *SIAM J. Opt.*, vol. 17, no. 3, pp. 844–860, Jul. 2006.

[57] X. Cao, G. Zhu, J. Xu, and K. Huang, "Cooperative interference management for over-the-air computation networks," *IEEE Trans. Wireless Commun.*, vol. 20, no. 4, pp. 2634–2651, Apr. 2021.

[58] Z. Wang, Y. Zhou, Y. Shi, and W. Zhuang, "Interference management for over-the-air federated learning in multi-cell wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 8, pp. 2361–2377, Aug. 2022.

[59] W. Ni, Y. Liu, Z. Yang, H. Tian, and X. Shen, "Federated learning in multi-RIS-aided systems," *IEEE Internet Things J.*, vol. 9, no. 12, pp. 9608–9624, Jun. 2022.

[60] C. Hu, Q. Li, Q. Zhang, and J. Qin, "Security optimization for an AF MIMO two-way relay-assisted cognitive radio nonorthogonal multiple access networks with SWIPT," *IEEE Trans. Inf. Forensics Security*, vol. 17, pp. 1481–1496, 2022.

[61] X. Wang, J. Liu, and C. Zhai, "Wireless power transfer-based multipair two-way relaying with massive antennas," *IEEE Trans. Wireless Commun.*, vol. 16, no. 11, pp. 7672–7684, Nov. 2017.

[62] M. Tao and R. Wang, "Linear precoding for multi-pair two-way MIMO relay systems with max-min fairness," *IEEE Trans. Signal Process.*, vol. 60, no. 10, pp. 5361–5370, Oct. 2012.

[63] Z. Zhang, Z. Chen, M. Shen, and B. Xia, "Spectral and energy efficiency of multipair two-way full-duplex relay systems with massive MIMO," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 848–863, Apr. 2016.

[64] K. Huang and N. D. Sidiropoulos, "Consensus-ADMM for general quadratically constrained quadratic programming," *IEEE Trans. Signal Process.*, vol. 64, no. 20, pp. 5297–5310, Oct. 2016.

[65] K. Nakata, M. Yamashita, K. Fujisawa, and M. Kojima, "A parallel primal–dual interior-point method for semidefinite programs using positive definite matrix completion," *Parallel Comput.*, vol. 32, pp. 24–43, Jan. 2006.

Quanzhong Li received the B.S. and Ph.D. degrees from Sun Yat-sen University (SYSU), Guangzhou, China, in 2009 and 2014, respectively.

He is currently an Associate Professor with the School of Computer Science and Engineering, SYSU. His research interests include the area of wireless communications and signal processing, with main focus on optimization techniques for resource allocation and physical-layer security.

Hualiang Luo received the B.Eng. and M.S. degrees from Wuhan University of Technology, Wuhan, China, in 2017 and 2020, respectively. He is currently pursuing the Ph.D. degree with the School of Electronics and Information Technology, Sun Yat-sen University, Guangzhou, China.

His research interests focus on the optimization design of wireless communications systems.

Liang Yang was born in Hunan, China. He received the Ph.D. degree in electrical engineering from Sun Yat-sen University, Guangzhou, China, in 2006.

From 2006 to 2013, he was a Teacher with Jinan University, Guangzhou. He joined Guangdong University of Technology, Guangzhou, in 2013. He is currently a Professor with Hunan University, Changsha, China. His current research interest includes the performance analysis of wireless communication systems.