

# Worst-case Design for RIS-aided Over-the-air Computation with Imperfect CSI

Wenhui Zhang, Jindan Xu, *Member, IEEE*, Wei Xu, *Senior Member, IEEE*,  
Xiaohu You, *Fellow, IEEE*, and Weijie Fu

**Abstract**—Over-the-air computation (AirComp) enables fast wireless data aggregation at the receiver through concurrent transmission by sensors in the application of Internet-of-Things (IoT). To further improve the performance of AirComp under unfavorable propagation channel conditions, we consider the problem of computation distortion minimization in a reconfigurable intelligent surface (RIS)-aided AirComp system. In particular, we take into account an additive bounded uncertainty of the channel state information (CSI) and the total power constraint, and jointly optimize the transceiver (Tx-Rx) and the RIS phase design from the perspective of worst-case robustness by minimizing the mean squared error (MSE) of the computation. To solve this intractable nonconvex problem, we develop an efficient alternating algorithm where both solutions to the robust sub-problem and to the joint design of Tx-Rx and RIS are obtained in closed forms. Simulation results demonstrate the effectiveness of the proposed method.

**Index Terms**—Reconfigurable intelligent surface, over-the-air computation, transceiver design, data aggregation.

## I. INTRODUCTION

AS the most prominent characteristic of the developing information age, massive data with astounding growth is of critical significance in facilitating the Internet-of-Things (IoT) applications. However, it has become harder for data collection from increasing smart devices of IoT. Meanwhile, it is even more challenging for information fusion of massive data, also known as data aggregation, from the numerous sensor devices with limited spectrum and low latency constraint.

To address this issue, a promising solution, namely over-the-air computation (AirComp), was proposed by leveraging the superposition property of wireless multiple access channels (MAC) [1]. Through concurrent transmissions by sensors and a weighted average function for the distributed local computing, fast data aggregation can be achieved by AirComp at the receiver, realizing a low transmission latency that is independent of the number of IoT devices. Recently, AirComp has attracted much attention in many areas like information theory [1], signal processing [2], and transceiver design [3]. It has been proved that AirComp can effectively improve communication efficiency and reduce the required bandwidth [4]. However, the advantages of the AirComp are promised upon that the propagation channel conditions are favorable.

To solve this problem, reconfigurable intelligent surface (RIS), emerging as a complementary technology [5], is believed to improve signal propagation conditions by reflecting incident signals at favorable angles utilizing the passive reflecting elements [6][7]. A RIS assisted AirComp system is proposed in [8] to boost the received signal power and mitigate

the performance bottleneck by reconfiguring the propagation channels. Thus with the assistance of using the RIS, energy focusing and nulling at desired locations can be realized in a AirComp system, shaping favorable propagation links for data collection. To reduce the mean squared error (MSE) of AirComp, the authors in [9] developed an alternating difference-of-convex (DC) programming algorithm by optimizing the transceiver and the RIS phase shifts. Besides, the authors in [10] tackled the non-convex optimization problem with the matrix lifting and concave convex procedure, solving the DC problem and designing the RISs phases and a linear detector for the RIS-aided cloud radio access network (C-RAN).

However, the assumption of perfect CSI in these studies for AirComp is hard to achieve in practice, resulting in noticeable performance degradation due to CSI inaccuracy in practical scenarios, not to mention the worst-case condition. To the best of our knowledge, one exception is [11], which proposed an efficient iterative algorithm to obtain a suboptimal solution for maximizing the system sum-rate considering the robustness against the impact of CSI imperfection. In addition, the wireless sensors of an AirComp system are usually powered by limited-power beacons [12], leading to a sum-power constraint of all the sensors.

Compared with these existing studies, in this paper we consider the optimization problem of transceiver design and RIS phase selection in a RIS-aided AirComp system under the sum-power constraint with imperfect CSI. Aiming to minimize the computation MSE under the total power constraint with CSI uncertainty, we cope with the joint design problem under the worst-case robustness for the AirComp. To solve this complicated nonconvex problem, we fortunately obtain closed-form solution to the subproblem of the robust design, and derive the closed-form solutions to the joint design of transceivers and RIS phases in an alternating optimization manner.

**Notations:** Throughout this paper,  $(\cdot)^H$ ,  $(\cdot)^T$ ,  $(\cdot)^*$ , and  $(\cdot)^\dagger$  denote the conjugate transpose, transpose, conjugate, and Moore-Penrose inverse, respectively.  $\mathbb{R}(\mathbb{C})$  and  $\mathbb{E}\{\cdot\}$  represent the space of real (complex) numbers and the expectation operator.  $|\cdot|$  denotes the modulus of a scalar.  $\mathcal{CN}(0, 1)$  represents the distribution of a circularly symmetric complex Gaussian variable with zero mean and unit variance.  $U[0, 2\pi)$  represents the uniform distribution drawn from 0 to  $2\pi$ . The  $k$ th entry of vector  $\mathbf{x}$  and the  $(i, j)$ th element of matrix  $\mathbf{X}$  are represented by  $[\mathbf{x}]_k$  and  $[\mathbf{X}]_{ij}$ , respectively.

## II. SYSTEM MODEL AND PROBLEM FORMULATION

We consider the uplink of a RIS-assisted AirComp system, which is composed of  $K$  single-antenna sensors and one single-antenna receiver. As shown in Fig. 1, assume that the direct links of sensor-receiver are blocked due to unfavorable

W. Zhang, J. Xu, W. Xu, and X. You are with the National Mobile Communications Research Laboratory (NCRL), Southeast University, Nanjing 210096, China ({whzhang, jdxu, wxu, xhyou}@seu.edu.cn).

W. Fu is with Guangdong Communications and Networks Institute, Guangdong 518131, China (fujw@gdnci.cn).

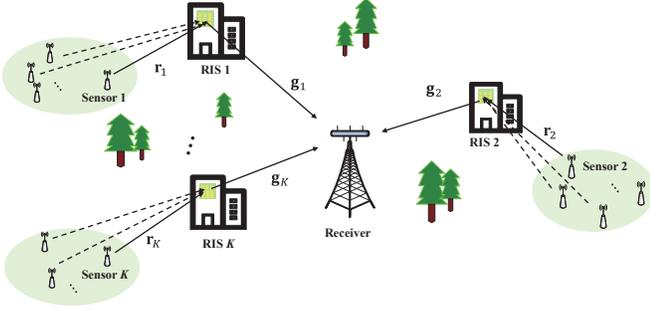


Fig. 1. Over-the-air computation with RIS.

propagation conditions, such as large obstacles [5][13], and are thus neglected. Each RIS with  $N$  passive reflecting antenna elements helps the communication from the sensors in a specific serving area to the receiver. Since the geometric locations of these RISs are relatively far from each other, the signal from one sensor is dominantly reflected by the RIS in the serving area, while propagations via the other RISs are ignorably weak due to large path loss. In a certain time slot, the RIS in each area only provides auxiliary reflection communication service for data uploading from a single scheduled sensor.

#### A. Processing of Air Computation

In the RIS-assisted AirComp system, each sensor pre-processes its own measurement signal and then sends it to the receiver for further calculation. Let  $x_k \in \mathbb{R}, \forall k \in \{1, \dots, K\}$  denote the pre-processed original signal of sensor  $k$ , which without loss of generality is assumed to have a normalized variance. Let  $\psi_k : \mathbb{R} \rightarrow \mathbb{R}, \forall k \in \{1, \dots, K\}$  denote a pre-processing function of sensor  $k$  and function  $\varphi : \mathbb{R} \rightarrow \mathbb{R}$  denotes the receiver's post-processing operation. Then, the received computation signal at the receiver is expressed as

$$y = \varphi \left( \sum_{k=1}^K \psi_k(x_k) \right). \quad (1)$$

From (1), by the pre-processing operation  $\psi_k$  by each sensor and the post-processing operation  $\varphi$  at the receiver, the receiver obtains the desired signal summation directly instead of requesting one-by-one transmissions of  $x_k$  and a specific computation [14]. By integrating computation and communication exploiting the signal superposition property of the wireless multiple-access channel, AirComp is accomplished via this concurrent transmission.

As to practical applications, for example, a central controller collects and computes the average real-time data (e.g., velocity, acceleration, temperature, humidity) sent by distributed sensors. By post-processing the operator like an averaging function of  $\varphi$ , the overall state of the environment is acquired.

#### B. Computation Model

In this system, each sensor data is linearly scaled by a factor  $t_k \in \mathbb{C}$  before transmission, realizing the pre-processing operation  $\psi_k$ . And the received signal is linearly scaled by a

Rx-scaling factor  $m$  realizing the post-processing operation  $\varphi$ . Applying this to (1), the received signal becomes

$$\begin{aligned} y &= m \left( \sum_{k=1}^K \mathbf{g}_k^H \Phi_k \mathbf{r}_k t_k x_k + n \right) \\ &= m \left( \sum_{k=1}^K \underbrace{\mathbf{g}_k^H \text{diag}\{\mathbf{r}_k\}}_{\mathbf{h}_k^H} \mathbf{v}_k t_k x_k + n \right), \end{aligned} \quad (2)$$

where  $\mathbf{g}_k \in \mathbb{C}^{N \times 1}$  and  $\mathbf{r}_k \in \mathbb{C}^{N \times 1}$  denote the channel between the  $k$ th RIS and the receiver and the channel between the  $k$ th sensor and its corresponding RIS, respectively. The diagonal matrix  $\Phi_k = \text{diag}(\beta_{k,1}e^{j\theta_{k,1}}, \dots, \beta_{k,N}e^{j\theta_{k,N}}) \in \mathbb{C}^{N \times N}$  is defined as the reflection matrix of RIS  $k$  and  $\mathbf{v}_k \triangleq [\beta_{k,1}e^{j\theta_{k,1}}, \dots, \beta_{k,N}e^{j\theta_{k,N}}]^T \in \mathbb{C}^{N \times 1}$  is a vector reshaped from the diagonal elements of  $\Phi_k$ , where  $\beta_{k,i} \in [0, 1]$  is the amplitude coefficient of the  $i$ th reflecting element of RIS  $k$  and  $\theta_{k,i} \in (0, 2\pi)$  is the phase coefficient for  $i \in \{1, 2, \dots, N\}$ . In this paper, we assume  $\beta_{k,i} = 1$  without loss of generality. Scalar  $n \in \mathbb{C}$  is the additive white Gaussian noise (AWGN) following  $\mathcal{CN}(0, \sigma^2)$ , and  $\mathbf{h}_k^H \triangleq \mathbf{g}_k^H \text{diag}\{\mathbf{r}_k\} \in \mathbb{C}^{1 \times N}$  is the cascaded equivalent channel. Especially in (2), the sum power of all the  $K$  sensors in the AirComp system is limited to  $P \triangleq \sum_{k=1}^K |t_k|^2$  for energy saving.

Then, the computation distortion in terms of the AirComp MSE, as defined in [12], is evaluated as

$$\begin{aligned} \text{MSE} &\triangleq \mathbb{E} \left\{ \left| y - \sum_{k=1}^K x_k \right|^2 \right\} \\ &= \sum_{k=1}^K |m \mathbf{h}_k^H \mathbf{v}_k t_k - 1|^2 + \sigma^2 |m|^2. \end{aligned} \quad (3)$$

Note that the MSE depends on the instantaneous CSI which can be obtained by channel estimation [13, 15] rather than the stochastic nature of  $\mathbf{g}_k$  and  $\mathbf{r}_k$ . However, the obtained CSI is always noisy, which follows the deterministic CSI error model [16] and is mathematically formulated as

$$\Delta_{\mathbf{h}_k} = \mathbf{h}_k^H - \hat{\mathbf{h}}_k^H, \quad (4)$$

where  $\mathbf{h}_k$ , for  $\forall k \in \{1, 2, \dots, K\}$ , is the nominal channel for the  $K$  links, and the CSI error  $\Delta_{\mathbf{h}_k}$  is bounded, i.e.,

$$\|\Delta_{\mathbf{h}_k}\|_2 \leq \varepsilon, \quad (5)$$

which means that the CSI uncertainty region is confined within a region of radius  $\varepsilon$ . By substituting (4), the MSE in (2) conditioned on the CSI estimate,  $\hat{\mathbf{h}}_k$ , is rewritten as

$$\text{MSE} = \sum_{k=1}^K |m(\hat{\mathbf{h}}_k^H + \Delta_{\mathbf{h}_k}) \mathbf{v}_k t_k - 1|^2 + \sigma^2 |m|^2. \quad (6)$$

#### C. Problem Formulation

The aim of our work is to minimize the AirComp MSE in (6) subject to the sum power constraint under the worst-case CSI error by jointly optimizing the design of the transceiver

and the RISs' reflection matrices. Now we are ready to formulate the optimization problem as follows

$$\min_{m, \{t_k\}, \{\mathbf{v}_k\}} \max_{\Delta_{\mathbf{h}_k}} \sum_{k=1}^K \left| m(\hat{\mathbf{h}}_k^H + \Delta_{\mathbf{h}_k}) \mathbf{v}_k t_k - 1 \right|^2 + \sigma^2 |m|^2 \quad (7a)$$

$$\text{s.t.} \quad \sum_{k=1}^K |t_k|^2 \leq P, \quad (7b)$$

$$|\mathbf{v}_k(n)|^2 = 1, \quad \forall n \in \{1, \dots, N\}, \quad (7c)$$

$$\|\Delta_{\mathbf{h}_k}\|_2 \leq \varepsilon. \quad (7d)$$

In particular, (7b) indicates the sum power constraint of all the  $K$  sensors, (7c) describes the unit-modulo constraints of the phase-shift element of RIS, and (7d) corresponds to the CSI error bound in (5). Observing this optimization problem, we find that it is hard to obtain the globally optimal solution, especially when coping with the nonconvex objective under the nonconvex unit-modulus constraint.

### III. PROPOSED WORST-CASE DESIGN FOR RIS-AIDED AIRCOMP SYSTEM

In this section, we tackle the optimization problem in (7) by jointly optimizing the design of the scaling factors  $m, \{t_k\}$  and the RISs' reflection matrices  $\{\mathbf{v}_k\}$  with the goal of minimizing the worst-case MSE in (7). We adopt the idea of alternative optimization, which is to alternately fix some variables while optimizing the others. We start with the robust design with  $m, \{t_k\}$  and  $\{\mathbf{v}_k\}$  fixed, which allows us to facilitate the problem solving and fortunately gives a closed-form solution for this worst-case solution.

#### A. Robust Design for the Worst-Case CSI Error

Assuming that  $m, \{t_k\}$  and  $\{\mathbf{v}_k\}$  are fixed, problem (7) is equal to

$$\max_{\Delta_{\mathbf{h}_k}, \forall k} \sum_{k=1}^K \left| m(\hat{\mathbf{h}}_k^H + \Delta_{\mathbf{h}_k}) \mathbf{v}_k t_k - 1 \right|^2 + \sigma^2 |m|^2 \quad (8a)$$

$$\text{s.t.} \quad \|\Delta_{\mathbf{h}_k}\|_2 \leq \varepsilon. \quad (8b)$$

Then, the above problem in (8a) can be decoupled into  $K$  subproblems as follows. For each  $k \in \{1, 2, \dots, K\}$ , we have

$$\min_{\Delta_{\mathbf{h}_k}} - \left| m t_k (\hat{\mathbf{h}}_k^H + \Delta_{\mathbf{h}_k}) \mathbf{v}_k - 1 \right|^2 \quad (9a)$$

$$\text{s.t.} \quad \|\Delta_{\mathbf{h}_k}\|_2 - \varepsilon \leq 0. \quad (9b)$$

**Theorem 1.** *The optimal solutions of  $\Delta_{\mathbf{h}_k}$ , and the optimized objective of problem (9) are given by the closed-form expressions as*

$$\Delta_{\mathbf{h}_k} = \frac{|\hat{t}_k|^2 \hat{\mathbf{h}}_k^H \mathbf{v}_k \mathbf{v}_k^H - \hat{t}_k \mathbf{v}_k^H}{\lambda_k - |\hat{t}_k|^2 N}, \quad (10)$$

$$\text{MSE}^* = \sum_{k=1}^K \left| \frac{\hat{t}_k \hat{\mathbf{h}}_k^H \mathbf{v}_k - 1}{1 - \lambda_k^{-1} |\hat{t}_k|^2 \mathbf{v}_k^H \mathbf{v}_k} \right|^2 + \sigma^2 |m|^2, \quad (11)$$

where  $\hat{t}_k \triangleq m t_k$ .

*Proof:* See Appendix A. ■

This theorem successfully copes with the worst-case CSI uncertainty in the system design.

#### B. Joint RIS Reflection and Transceiver Design

From the above subsection, we obtain the optimal solution in a closed form which achieves the worst-case MSE. Observing (11) and using  $\mathbf{v}_k^H \mathbf{v}_k = N$ , it is equivalent to rewrite the problem in (11) as

$$\min_{\{\hat{t}_k\}, \{\mathbf{v}_k\}} \sum_{k=1}^K \left| \frac{\hat{t}_k \hat{\mathbf{h}}_k^H \mathbf{v}_k - 1}{1 - \lambda_k^{-1} |\hat{t}_k|^2 N} \right|^2 + \sigma^2 |m|^2$$

$$\text{s.t.} \quad \sum_{k=1}^K |t_k|^2 \leq P,$$

$$|\mathbf{v}_k(n)|^2 = 1, \forall n \in \{1, \dots, N\}. \quad (12)$$

It is obvious that the sum power constraint is active since the larger  $t_k$  is, the smaller the MSE is. Then, it is safe to rewrite the constraint in (12) as the equality constraint, i.e.,  $\sum_{k=1}^K |t_k|^2 = P$ . The problem in (12) becomes

$$\min_{\{\hat{t}_k\}, \{\mathbf{v}_k\}} \sum_{k=1}^K \left| \frac{\hat{t}_k \hat{\mathbf{h}}_k^H \mathbf{v}_k - 1}{1 - \lambda_k^{-1} |\hat{t}_k|^2 N} \right|^2 + \frac{\sigma^2}{P} \sum_{k=1}^K |\hat{t}_k|^2$$

$$\text{s.t.} \quad |\mathbf{v}_k(n)|^2 = 1, n \in \{1, \dots, N\}. \quad (13)$$

Note that the problem in (13) is equivalent to the original problem in (7), but with respect to a reduced number of optimization variables including  $t_k$  and  $m$ . However, the problem is still less tractable due to the unit-modulus nonconvex constraints and the coupling variables. Since the problem in (13) can be decoupled into  $K$  subproblems for the independent variables with subscript  $k$ , without loss of generality, we focus on the  $k$ th subproblem and simplify the MSE objective in (11) as

$$\text{MSE}_k \triangleq \left| \frac{\hat{t}_k \hat{\mathbf{h}}_k^H \mathbf{v}_k - 1}{1 - \lambda_k^{-1} |\hat{t}_k|^2 N} \right|^2 + \frac{\sigma^2}{P} |\hat{t}_k|^2. \quad (14)$$

Then we apply the idea of alternating optimization and temporarily relax the unit-modulus constraint. Firstly, for any fixed  $m$  and  $t_k$ , we minimize the MSE by solving the following equality

$$\frac{\partial(\text{MSE}_k)}{\partial \mathbf{v}_k^*} = \frac{|\hat{t}_k|^2 \hat{\mathbf{h}}_k^H \mathbf{v}_k - \hat{t}_k^* \hat{\mathbf{h}}_k}{(1 - \lambda_k^{-1} N |\hat{t}_k|^2)^2} = 0, \quad (15)$$

which yields

$$\mathbf{v}_k = \frac{1}{\hat{t}_k} (\hat{\mathbf{h}}_k \hat{\mathbf{h}}_k^H)^\dagger \hat{\mathbf{h}}_k. \quad (16)$$

Considering the unit-modulus constraint which is temporarily relaxed, we can extract the phase parameters of (16) by normalizing the amplitude, and thus the unit-modulus solution for  $\mathbf{v}_k$  becomes irrelevant to  $\hat{t}_k$ . Then for the solution of  $\hat{t}_k$ , we have the following theorem.

**Theorem 2.** *The optimal solutions for the scaling factors  $m$  and  $t_k$ , equivalently  $\hat{t}_k$  in (12) are given as*

$$m = \sqrt{\frac{1}{P} \sum_{k=1}^K |\hat{t}_k|^2}, \quad (17)$$

$$t_k = \hat{t}_k \sqrt{\frac{P}{\sum_{k=1}^K |\hat{t}_k|^2}}, \quad (18)$$

$$\text{where } |\hat{t}_k|^2 = \frac{\sqrt[3]{\frac{2\lambda_k^{-1}NP|\hat{\mathbf{h}}_k^H(\hat{\mathbf{h}}_k\hat{\mathbf{h}}_k^H)^\dagger\hat{\mathbf{h}}_k-1|^2}{\sigma^2}} + 1}{\lambda_k^{-1}N}.$$

*Proof:* By substituting (16), the (14) becomes

$$\text{MSE}_k \triangleq \left| \frac{\hat{\mathbf{h}}_k^H(\hat{\mathbf{h}}_k\hat{\mathbf{h}}_k^H)^{-1}\hat{\mathbf{h}}_k - 1}{1 - \lambda_k^{-1}|\hat{t}_k|^2N} \right|^2 + \frac{\sigma^2}{P}|\hat{t}_k|^2. \quad (19)$$

Observing the  $\text{MSE}_k$  in (19), it is easy to check that

$$\frac{\partial^2 \text{MSE}_k}{\partial (|\hat{t}_k|^2)^2} = \frac{6(\lambda_k^{-1}N)^2 |\hat{\mathbf{h}}_k^H(\hat{\mathbf{h}}_k\hat{\mathbf{h}}_k^H)^{-1}\hat{\mathbf{h}}_k - 1|^2}{(1 - \lambda_k^{-1}N|\hat{t}_k|^2)^4} > 0, \quad (20)$$

which allows us to obtain the optimal solution by forcing the following derivative to zero. It follows

$$\frac{\partial \text{MSE}_k}{\partial (|\hat{t}_k|^2)} = \frac{2\lambda_k^{-1}N |\hat{\mathbf{h}}_k^H(\hat{\mathbf{h}}_k\hat{\mathbf{h}}_k^H)^\dagger\hat{\mathbf{h}}_k - 1|^2}{(1 - \lambda_k^{-1}N|\hat{t}_k|^2)^3} + \frac{\sigma^2}{P} = 0, \quad (21)$$

and thus

$$|\hat{t}_k|^2 = \frac{\sqrt[3]{\frac{2\lambda_k^{-1}NP|\hat{\mathbf{h}}_k^H(\hat{\mathbf{h}}_k\hat{\mathbf{h}}_k^H)^\dagger\hat{\mathbf{h}}_k-1|^2}{\sigma^2}} + 1}{\lambda_k^{-1}N}. \quad (22)$$

Using the equalities  $mt_k = \hat{t}_k$  and  $\sum_{k=1}^K |t_k|^2 = P$ , we are able to get the closed-form solutions in (17) and (18) for  $m$  and  $t_k$ , respectively. ■

Now we have obtained the optimal closed-form solutions for  $m$  in (17) and  $\{t_k\}$  in (18). By using (16), we also obtain the solution for the RISs' reflection matrices  $\{\mathbf{v}_k\}$ . In summary, the proposed solution for solving the original problem in (7) is described in Algorithm 1.

Note that the computational complexity of the proposed algorithm for solving problem (7) is dominated by the computations of matrix inversions and matrix multiplications. Concretely, the complexity of the calculations in  $\|\Delta_{\mathbf{h}_k}\|^2 = \varepsilon^2$  is  $\mathcal{O}(N)$ , and the complexity of the calculations in (16) is  $\mathcal{O}(N^3)$ . For calculating (17) and (18), the dominating complexity is the calculation of the intermediate variables  $\hat{t}_k$ , which is  $\mathcal{O}(N^3)$ . Then for each iteration in Algorithm 1, the overall computational complexity amounts to  $\mathcal{O}(KN^3)$ .

#### IV. SIMULATION RESULTS

In this section, we present the numerical results for the proposed algorithm. We define the CSI error level,  $\varepsilon = s \|\mathbf{h}_k\|_2$ , where  $s$  represents the CSI uncertainty coefficient which is a metric for evaluating the relative amount of the CSI uncertainty. For simulation,  $s$  is set as 0.4 and 0.6, and the channel coefficients  $\sigma_h^2$  is set as 0.5. We evaluate the performance in terms of the normalized mean squared error (NMSE) of the computation distortion at the receiver and the signal-to-noise ratio (SNR) is defined as  $10\log_{10}(P/\sigma_n^2)$ .

Fig. 2 evaluates the effectiveness of our proposed scheme compared with the non-robust scheme in [12], where  $N = 16$ ,  $K = 10$  and  $P = 10$ . It is obvious that the proposed method achieves better performance than the non-robust scheme. Meanwhile, CSI uncertainty coefficient  $s$  is correlated with the computation MSE, indicating that CSI error confined even within a smaller region results in lower MSE. In addition, we exploit the method by using many random initial values to acquire the globally optimum (approximate) in a numerical

---

#### Algorithm 1: Proposed algorithm for solving (7)

---

**Input:** Threshold  $\delta > 0$

- 1 initialization:  $n = 1$ ;
- 2 **for**  $k$  **to**  $K$  **do**
- 3     initialize  $\mathbf{v}_k^{(1)}$  satisfying (7c);
- 4      $m^{(1)} = 1, t_k^{(1)} = P/K$ ;
- 5 **end**
- 6 **do**
- 7     **for**  $k$  **to**  $K$  **do**
- 8         Calculate  $\lambda_k^{(n)}$  by solving the equation  $\|\Delta_{\mathbf{h}_k}^{(n)}\|^2 = \varepsilon^2$ ;
- 9         Update  $\mathbf{v}_k^{(n+1)}$  according to (16) and normalize the amplitude;
- 10         Update  $m^{(n+1)}$  according to (17);
- 11         Update  $t_k^{(n+1)}$  according to (18);
- 12     **end**
- 13      $n = n + 1$ ;
- 14 **while**  $\sum_{k=1}^K \|\mathbf{v}_k^{(n)} - \mathbf{v}_k^{(n-1)}\|_2^2 + |t_k^{(n)} - t_k^{(n-1)}|^2 + |m^{(n)} - m^{(n-1)}|^2 \leq \delta$ ;

**Output:**  $m = m^{(n-1)}, \{t_k\} = \{t_k^{(n-1)}\}, \{\mathbf{v}_k\} = \{\mathbf{v}_k^{(n-1)}\}$

---

way with tractable complexity. It is encouraging that the performance of the proposed method is close to the performance of a globally optimal solution.

Fig. 3 compares the computation NMSE at the receiver with different numbers of RIS reflecting elements, where  $K = 8$ ,  $s = 0.4$ , and  $P = 100$ . It is observed that with increasing RIS reflecting elements, the NMSE gradually improves. The reason is that with the growth of the RIS reflecting elements which can be regarded as more antennas, an extra channel gain is provided. However, with the increase of the RIS elements, the cost and power consumption problem will not be ignored, making it critical for the tradeoff design.

Fig. 4 evaluates the effect of the number of sensor devices on the performance of the system with  $N = 64$ ,  $P = 100$ , and  $s = 0.4$ . As we can see in Fig. 4, the performance deteriorates when more and more sensors are getting accessed. With an increasing number of sensors under the total power constraint, the performance gain of the proposed robust design becomes more pronounced.

#### V. CONCLUSION

In this paper, we consider the joint optimization of transceiver and the RIS phase design for a RIS-aided AirComp system with imperfect CSI. The complicated optimization problem is equivalently transformed into a subproblem of robust design and joint design of the system. We solve the joint design problem under the worst-case robustness for CSI error model and obtain the closed-form solutions by utilizing efficient alternating algorithms and solving the Karush-Kuhn-Tucker conditions.

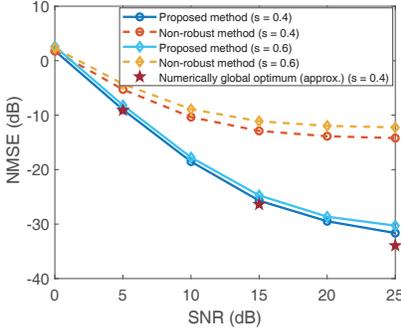
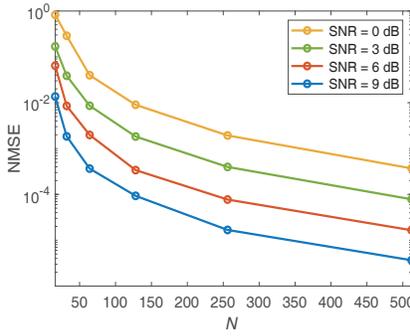
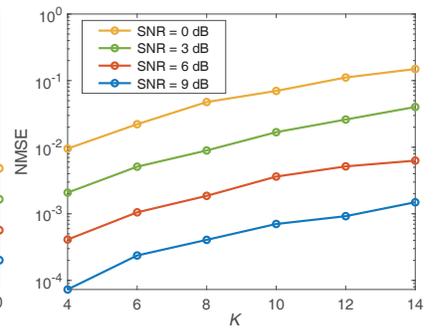


Fig. 2. The NMSE versus SNR.

Fig. 3. The NMSE versus  $N$ .Fig. 4. The NMSE versus  $K$ .

## APPENDIX A PROOF OF THEOREM 1

By applying the Karush-Kuhn-Tucker (KKT) conditions [17], it is equivalent to solve the following conditions as

$$\begin{aligned}
 \min_{\Delta_{\mathbf{h}_k}, \lambda_k} \quad & \mathcal{L}(\Delta_{\mathbf{h}_k}, \lambda_k) \\
 = & - \left| m(\hat{\mathbf{h}}_k^H + \Delta_{\mathbf{h}_k}) \mathbf{v}_k t_k - 1 \right|^2 + \lambda_k (\|\Delta_{\mathbf{h}_k}\|^2 - \varepsilon^2) \\
 \text{s.t.} \quad & \frac{\partial \mathcal{L}(\Delta_{\mathbf{h}_k}, \lambda_k)}{\partial \Delta_{\mathbf{h}_k}^*} = 0, \\
 & \lambda_k (\|\Delta_{\mathbf{h}_k}\|^2 - \varepsilon^2) = 0, \\
 & \|\Delta_{\mathbf{h}_k}\|^2 - \varepsilon^2 \leq 0, \\
 & \lambda_k \geq 0,
 \end{aligned} \tag{23}$$

where  $\lambda_k$ , for  $k = 1, 2, \dots, K$ , are the KKT multipliers.

Letting  $\hat{t}_k = m t_k$ , and from (23), we have

$$\begin{aligned}
 \frac{\partial \mathcal{L}(\Delta_{\mathbf{h}_k}, \lambda_k)}{\partial \Delta_{\mathbf{h}_k}^*} = & - |\hat{t}_k|^2 \hat{\mathbf{h}}_k^H \mathbf{v}_k \mathbf{v}_k^H + \hat{t}_k^* \mathbf{v}_k^H - |\hat{t}_k|^2 \Delta_{\mathbf{h}_k} \mathbf{v}_k \mathbf{v}_k^H \\
 & + \lambda_k \Delta_{\mathbf{h}_k}.
 \end{aligned} \tag{24}$$

Forcing (24) to zero, we obtain

$$\begin{aligned}
 \Delta_{\mathbf{h}_k} = & (|\hat{t}_k|^2 \hat{\mathbf{h}}_k^H \mathbf{v}_k \mathbf{v}_k^H - \hat{t}_k^* \mathbf{v}_k^H) (\lambda_k \mathbf{I}_N - |\hat{t}_k|^2 \mathbf{v}_k \mathbf{v}_k^H)^{-1} \\
 = & - \hat{\mathbf{h}}_k^H + (\lambda_k \hat{\mathbf{h}}_k^H - \hat{t}_k^* \mathbf{v}_k^H) (\lambda_k \mathbf{I}_N - |\hat{t}_k|^2 \mathbf{v}_k \mathbf{v}_k^H)^{-1} \\
 \stackrel{(a)}{=} & - \hat{\mathbf{h}}_k^H + \lambda_k \hat{\mathbf{h}}_k^H (\lambda_k^{-1} \mathbf{I}_N - \frac{-\lambda_k^{-2} |\hat{t}_k|^2 \mathbf{v}_k \mathbf{v}_k^H}{1 - \lambda_k^{-1} |\hat{t}_k|^2 \mathbf{v}_k^H \mathbf{v}_k})^{-1} \\
 & - \hat{t}_k^* \mathbf{v}_k^H (\lambda_k \mathbf{I}_N - |\hat{t}_k|^2 \mathbf{v}_k \mathbf{v}_k^H)^{-1} \\
 \stackrel{(b)}{=} & \frac{\lambda_k^{-1} |\hat{t}_k|^2 \hat{\mathbf{h}}_k^H \mathbf{v}_k \mathbf{v}_k^H}{1 - \lambda_k^{-1} |\hat{t}_k|^2 \mathbf{v}_k^H \mathbf{v}_k} - \frac{\lambda_k^{-1} \hat{t}_k^* \mathbf{v}_k^H}{1 - \lambda_k^{-1} |\hat{t}_k|^2 \mathbf{v}_k^H \mathbf{v}_k} \\
 = & \frac{\lambda_k^{-1} |\hat{t}_k|^2 \hat{\mathbf{h}}_k^H \mathbf{v}_k \mathbf{v}_k^H - \lambda_k^{-1} \hat{t}_k^* \mathbf{v}_k^H}{1 - \lambda_k^{-1} |\hat{t}_k|^2 \mathbf{v}_k^H \mathbf{v}_k},
 \end{aligned} \tag{25}$$

where (a) uses the matrix inversion lemma as

$$(\mathbf{A} + \mathbf{u}\mathbf{v}^T)^{-1} = \mathbf{A}^{-1} - \frac{1}{1 + \mathbf{v}^T \mathbf{A}^{-1} \mathbf{u}} \mathbf{A}^{-1} \mathbf{u}\mathbf{v}^T \mathbf{A}^{-1}, \tag{26}$$

and (b) uses the fact that

$$\mathbf{x}^H (\mathbf{A} + \tau \mathbf{x}\mathbf{x}^H)^{-1} = \frac{\mathbf{x}^H \mathbf{A}^{-1}}{1 + \tau \mathbf{x}^H \mathbf{A}^{-1} \mathbf{x}} \tag{27}$$

Substitute (25) into the KKT complementary condition in (23), i.e.,  $\lambda_k (\|\Delta_{\mathbf{h}_k}\|^2 - \varepsilon^2) = 0$  and solve this equality. By

for any invertible matrix  $\mathbf{A}$ .

removing unreasonable negative values, we get the value of  $\lambda_k$ .

Similarly, we have the desired worst-case MSE in (11).

## REFERENCES

- [1] B. Nazer and M. Gastpar, "Computation over multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3498–3516, Oct. 2007.
- [2] J. Dong, Y. Shi, and Z. Ding, "Blind over-the-air computation and data fusion via provable wirtinger flow," *IEEE Trans. Signal Process.*, vol. 68, pp. 1136–1151, Jan. 2020.
- [3] L. Chen, X. Qin, and G. Wei, "A uniform-forcing transceiver design for over-the-air function computation," *IEEE Wireless Commun. Lett.*, vol. 7, no. 6, pp. 942–945, Dec. 2018.
- [4] K. Yang, T. Jiang, Y. Shi, and Z. Ding, "Federated learning via over-the-air computation," *IEEE Trans. Wireless Commun.*, vol. 19, no. 3, pp. 2022–2035, Mar. 2020.
- [5] Q. Wu and R. Zhang, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network," *IEEE Commun. Mag.*, vol. 58, no. 1, pp. 106–112, Jan. 2020.
- [6] H. Shen, W. Xu, S. Gong, C. Zhao, and D. W. K. Ng, "Beamforming optimization for IRS-aided communications with transceiver hardware impairments," *IEEE Trans. Commun.*, vol. 69, no. 2, pp. 1214–1227, Feb. 2021.
- [7] W. Shi, W. Xu, X. You, C. Zhao, and K. Wei, "Intelligent reflection enabling technologies for integrated and green Internet-of-Everything beyond 5G: Communication, sensing, and security," *IEEE Wireless Commun.*, early access.
- [8] W. Fang *et al.*, "Over-the-air computation via reconfigurable intelligent surface," *IEEE Trans. Commun.*, vol. 69, no. 12, pp. 8612–8626, Dec. 2021.
- [9] T. Jiang and Y. Shi, "Over-the-air computation via intelligent reflecting surfaces," in *Proc. IEEE GLOBECOM*, Dec. 2019, pp. 1–6.
- [10] D. Yu, S.-H. Park, O. Simeone, and S. S. Shitz, "Optimizing over-the-air computation in IRS-aided C-RAN systems," in *Proc. IEEE SPAWC*, May 2020, pp. 1–5.
- [11] S. Hua, Z. Wei, Y. Cai, C. Liu, D. W. K. Ng, and J. Yuan, "Robust and secure sum-rate maximization for multiuser MISO downlink systems with self-sustainable IRS," *IEEE Trans. Commun.*, early access, Jul. 14, 2021.
- [12] 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, Sept. 2020.
- [13] W. Zhang *et al.*, "Cascaded channel estimation for IRS-assisted mmWave multi-antenna with quantized beamforming," *IEEE Commun. Lett.*, vol. 25, no. 2, pp. 593–597, Feb. 2021.
- [14] 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.
- [15] Z. Wang, L. Liu, and S. Cui, "Channel estimation for intelligent reflecting surface assisted multiuser communications: Framework, algorithms, and analysis," *IEEE Trans. Wireless Commun.*, early access, 2020.
- [16] W. Xu, Y. Cui, H. Zhang, G. Y. Li and X. You, "Robust beamforming with partial channel state information for energy efficient networks," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 12, pp. 2920–2935, Dec. 2015.
- [17] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.