

Simultaneously Transmitting and Reflecting (STAR) RIS Assisted Over-the-Air Computation Systems

Xiongfei Zhai, *Member, IEEE*, Guojun Han, *Senior Member, IEEE*, Yunlong Cai, *Senior Member, IEEE*, Yuanwei Liu, *Senior Member, IEEE*, and Lajos Hanzo, *Life Fellow, IEEE*

Abstract—The performance of over-the-air computation (AirComp) systems degrades due to the hostile channel conditions of wireless devices (WDs), which can be significantly improved by the employment of reconfigurable intelligent surfaces (RISs). However, the conventional RISs require that the WDs have to be located in the half-plane of the reflection space, which restricts their potential benefits. To address this issue, the novel family of simultaneously transmitting and reflecting reconfigurable intelligent surfaces (STAR-RIS) is considered in AirComp systems to improve the computation accuracy across a wide coverage area. To minimize the computation mean-squared-error (MSE) in STAR-RIS assisted AirComp systems, we propose a joint beamforming design for optimizing both the transmit power at the WDs, as well as the passive reflect and transmit beamforming matrices at the STAR-RIS, and the receive beamforming vector at the fusion center (FC). Specifically, in the updates of the passive reflect and transmit beamforming matrices, closed-form solutions are derived by introducing an auxiliary variable and exploiting the coupled binary phase-shift conditions. Moreover, by assuming that the number of antennas at the FC and that of elements at the STAR-RIS/RIS are sufficiently high, we theoretically prove that the STAR-RIS assisted AirComp systems provide higher computation accuracy than the conventional RIS assisted systems. Our numerical results show that the proposed beamforming design outperforms the benchmark schemes relying on random phase-shift constraints and the deployment of conventional RIS. Moreover, its performance is close to the lower bound achieved

by the beamforming design based on the STAR-RIS dispensing with coupled phase-shift constraints.

I. INTRODUCTION

The ultra-fast aggregation of the deluge of data generated by distributed wireless devices (WDs) will become an urgent demand for the Internet-of-things (IoT) [1]–[3]. However, due to the fact that the fusion center (FC) of conventional communication systems has to recover the individual data stream of each WD and then perform data aggregation, there are three dominant challenges: 1) the inter-WD interference will degrade the performance of data stream recovery; 2) the successive data stream recovery will lead to extremely high latency; 3) the spectral- and energy-efficiency will be severely affected by the increasing number of WDs. Since the FC typically has to compute a specific function (e.g., the arithmetic mean, the weighted sum, or the geometric mean) in data aggregation, the compelling technique of over-the-air computation (AirComp) has been proposed as a potential remedy [4]–[11]. However, as the number of WDs increases rapidly, the computation accuracy of AirComp tends to erode, since it is challenging to design a single receiver for the FC to aggregate the data from an escalating number of WDs. To address this issue, advanced massive multiple-input multiple-output (MIMO) and reconfigurable intelligent surface (RIS) aided techniques may be harnessed in AirComp. Specifically, the RIS is more attractive due to its flexible deployment and reduced power consumption. Although the traditional RISs beneficially improve the computational performance of AirComp, they can only provide half-plane coverage, while the WDs are spread across the entire 360-degree angular range. To extend the coverage, the recent simultaneously transmitting and reflecting RISs (STAR-RISs) have excellent potential in AirComp due to the fact that they can reflect and transmit the incident signals simultaneously. Therefore, full-plane coverage is possible in the STAR-RIS assisted AirComp systems.

A. Prior Works

In the literature, the computation mean-square-error (MSE) metric is the most popular one in AirComp systems [4]–[11]. By assuming fading channels and single-antenna FC/WDs, the power allocation algorithms of AirComp systems based on popular convex optimization methods have been investigated in [4]–[6]. Instead of minimizing the computation MSE, a new computation metric, namely the outage probability, is defined as the probability that the computation MSE exceeds

The work of X. Zhai was supported in part by the Science and Technology Program of Guangzhou under Grant 202102020869, and the Guangdong Basic and Applied Basic Research Foundation under Grants 2022A1515010153 and 2021A1515011645. The work of G. Han was supported in part by the Joint Funds of the National Natural Science Foundation of China and Guangdong under Grant U2001203, Key-Area R&D Program of Guangdong Province under Grant 2021B1101270001, and Guangdong Provincial Key Laboratory of Photonics Information Technology under Grant 2020B121201011. The work of Y. Cai was supported in part by the National Natural Science Foundation of China under Grants 61971376 and U22A2004. This work of Y. Liu was supported in part by the CHIST-ERA grant under the project CHIST-ERA-20-SICT-005, by the Engineering and Physical Sciences Research Council (EPSRC) under Project EP/W035588/1, in part by the Royal Society under grant RGS/R1/221050 and grant IEC/NSFC/201112, and in part by the PHC Alliance Franco-British Joint Research Programme under Grant 822326028. L. Hanzo would like to acknowledge the financial support of the Engineering and Physical Sciences Research Council projects EP/W016605/1 and EP/P003990/1 (COALESCE) as well as of the European Research Council's Advanced Fellow Grant QuantCom (Grant No. 789028). (Corresponding author: Yunlong Cai)

X. Zhai and G. Han are with the School of Information Engineering, Guangdong University of Technology, Guangzhou 510006, China (e-mail: zhaixiongfei@gdut.edu.cn; gjhan@gdut.edu.cn).

Y. Cai is with the College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China (e-mail: ylcai@zju.edu.cn).

Y. Liu is with the School of Electronic Engineering and Computer Science, Queen Mary University of London, London E1 4NS, UK, (email: yuanwei.liu@qmul.ac.uk).

L. Hanzo is with the Department of Electronics and Computer Science, University of Southampton, Southampton, UK (Email: lh@ecs.soton.ac.uk).

a given threshold and is minimized in [7]. However, the aforementioned contributions focused their attention on single-input single-output (SISO) AirComp systems, which cannot support multi-functional computation. To deal with this issue, the authors of [9] exploited the MIMO techniques in their AirComp systems. Specifically, a zero-forcing (ZF) based beamforming design was proposed for MIMO-aided AirComp systems in [9] for facilitating multi-functional computations and for minimizing the computation MSE. To guarantee low computation MSE, accurate time synchronization is important for AirComp systems. To achieve time synchronization both among the WDs as well as between the FC and WDs, a synchronization scheme, which is referred to as AirShare, has been developed in [10]. In this scheme, the FC broadcasts a synchronization block to all WDs for time synchronization.

Nevertheless, the computation accuracy of AirComp will be severely degraded by the rapidly increasing number of WDs due to the ones having poor channel conditions. To overcome this challenge, one of the effective techniques is to provide higher spatial degrees. Thus the massive MIMO techniques were exploited for AirComp in [11]–[13]. In particular, in our previous work [12] we conceived a hybrid analog-digital (AD) beamformer for massive MIMO AirComp systems and proposed a successive-convex-approximation (SCA) based beamforming design to achieve low computation MSE. Moreover, the computation MSE has been proven to be inversely proportional to the number of receive antennas at the FC. Although the computation accuracy can be significantly improved by massive MIMO schemes, this is generally at the cost of huge energy consumption and complexity. Thus finding a more efficient scheme for enhancing the computation performance of AirComp is of pivotal importance.

Again, the passive RIS concept has attracted substantial attention worldwide due to its compelling capability of increasing the spatial degrees of freedom, whilst relying on passive reflecting elements [14]–[18]. By adjusting the phase shift and even potentially the amplitude of the incident signal with the aid of passive reflecting elements, the propagation environment is beneficially influenced. As compared to massive MIMO, the fabrication cost and energy consumption of RIS is modest, since no radio frequency (RF) chains are used at the RIS. Given the above advantages, RISs have been exploited in diverse applications, such as RIS-aided wireless power transfer [19], [20], RIS assisted physical layer security [21]–[24] and RIS-aided orthogonal frequency division multiplexing (OFDM) [25], [26]. Specifically, the authors in [20] investigated the RIS-assisted simultaneous wireless information and power transfer (SWIPT) systems with the quality-of-service (QoS) constraints at the information and energy users. A penalty-based algorithm was proposed to address the non-convexity of the corresponding optimization problem. To achieve the secure transmission, the RIS-assisted MIMO systems was studied in [24] and an alternating optimization algorithm with SCA was developed to jointly optimize the beamforming design at the transceiver and RIS. To apply the RIS in the frequency-selected channels, the combination of RIS and OFDM was exploited in [25]. To estimate the channel state information (CSI) with reduced overhead, a

RIS-element-grouping method was proposed. Then the system rate maximization problem was formulated and solved by alternately optimizing the power allocation and passive array coefficients.

RISs have also been exploited in AirComp systems [27]–[30]. In particular, a joint active and passive beamforming design based on difference-of-convex (DC) programming and matrix lifting has been proposed in [27] for minimizing the computation MSE. Furthermore, the authors of [29] combined AirComp and energy-harvesting, where the RIS-assisted beamforming design problem is separated into two phases: the energy beamforming phase and the AirComp phase. To reduce the heavy signaling overhead due to the estimation of CSI, a two-stage stochastic beamforming algorithm has been proposed in [30], where the passive beamforming matrix employed at the RIS is optimized based on the channel statistics, while the transmit power at the WDs and the receive beamforming vector at the FC are updated with the aid of the effective near-real-time CSI. Although the RISs can significantly improve the computation performance of AirComp, the aforementioned treatises only considered conventional RISs, which can only reflect the incident signals to a 180-degree half-plane. In other words, the FC and WDs have to be located on the same side of the RIS. Hence this geographical constraint restricts the efficiency, appeal and flexibility of RISs, especially when the FC and WDs are located on both sides of a RIS on a 360-degree disc.

To address this issue, a novel technique, which is referred to as STAR-RISs, has been developed in [31] and [32]. In particular, the incident signals impinging on the elements of STAR-RIS from either side will be divided into two parts, where one of them is reflected to the same side of the incident signals into the so-called *reflection space*, while the other is transmitted to the opposite side, namely to the *transmission space*. This can be achieved by adjusting the electric and magnetic currents of the elements at the STAR-RIS. Thus the incident signals can be transmitted and reflected with the aid of two types of coefficients, which are referred to as the transmission and reflection coefficients. Some prototypes of STAR-RISs have been proposed in the literature based on metasurfaces [33], [34]. The authors of [33] considered the passive elements with a parallel resonant LC tank and small metallic loops, which can provide the required electric and magnetic surface reactance. Due to the aforementioned advantages, some authors have also considered the exploitation of STAR-RIS in wireless communication systems [35]–[37]. In particular, three different operating protocols have been developed for STAR-RISs in [35], namely energy splitting, mode switching, and time switching. As a further advance, the authors of [35] proposed the joint active and passive beamforming concept for minimizing the power consumption of STAR-RIS aided unicast and multicast systems. By considering the coupled phase-shift constraints between the transmit and reflect matrices at the STAR-RIS, the authors of [36] investigated the power consumption minimization problems under specific user-rate constraints for both non-orthogonal multiple access (NOMA) and orthogonal multiple access (OMA) systems by relying on an element-wise alternating

optimization algorithm.

B. Motivations and Contributions

However, to the best of our knowledge, the application of STAR-RISs in AirComp systems and the corresponding beamforming design have not been investigated in the open literature. Hence closing this knowledge-gap is the main motivation of this work. Furthermore, as compared to traditional RISs, which can only provide half-plane coverage for the AirComp systems, STAR-RISs is capable of covering the full plane. Furthermore, the limited coverage of the traditional RISs introduces increased design/coordination complexity, especially for a dense network, while STAR-RISs can be regarded as being transparent to the communication environment. Our main contributions of this paper are summarized as follows:

- We propose a STAR-RIS assisted AirComp system, where a multi-antenna FC aims for aggregating the signals from a number of single-antenna WDs with the aid of a STAR-RIS. In contrast to the traditional RIS-assisted AirComp systems, the WDs located in both the reflection and transmission spaces may be supported. Thus the transmit power at the WDs, the receive beamforming vector at the FC, and the passive reflect/transmit beamforming matrices at the STAR-RIS have to be jointly optimized for minimizing the computation MSE. No research on such systems has been reported in the open literature as yet. Besides, the formulated problem is challenging due to the fact that the passive reflect and transmit beamforming matrices at the STAR-RIS are coupled in the phases and are both subjected to the non-convex unit modulus constraints.
- An efficient optimization algorithm is proposed for updating the variables in an alternating manner. First, the receive beamforming vector at the FC and the transmit power at the WDs are optimized by exploiting the first-order optimality condition and the Lagrange duality method, respectively. Then we reformulate the original problem by introducing an auxiliary binary vector and exploiting the binary phase-coupled constraints. The closed-form solutions of the transmit/reflect beamforming matrices at the STAR-RIS are derived based on the constant modulus constraints and the binary nature. The complexity and convergence of the proposed algorithm are also analyzed.
- Given a sufficiently high numbers of antennas at the FC and elements at the STAR-RIS, we theoretically derive approximate expression for the computation MSEs of both the STAR-RIS assisted and conventional RIS assisted systems. Then it is proven that the STAR-RIS assisted AirComp systems can achieves lower computation MSE than their conventional RIS assisted AirComp counterparts, which further validate the effectiveness of STAR-RIS.
- Our numerical results verify the convergence of the proposed algorithm. It is also shown that our proposed algorithm outperforms other benchmark schemes based on the random phase-shift constraints and conventional RIS.

Thus our algorithm achieves both reduced computation MSE and full-plane coverage compared to its conventional RIS-aided counterpart. Moreover, the computation MSE achieved by the proposed algorithm is close to the lower bound attained by the scheme based on STAR-RIS without coupled phase-shift constraints.

C. Organization and Notations

The remainder of this paper is organized as follows. Section II introduces the system model and the problem formulated. The proposed joint beamforming design and the analysis of its convergence and complexity are presented in Section III. The comparison between the STAR-RIS assisted AirComp systems and their conventional RIS-aided counterparts are shown in Section IV. Finally, Section V presents our simulation results, followed by our conclusions in Section VI.

Notations: Throughout this paper, we adopt bold upper-case letters for matrices and bold lower-case letters for vectors. $\mathbf{A}(i, j)$ denotes the entry on the i^{th} row and the j^{th} column for a matrix \mathbf{A} , while \mathbf{A}^T , \mathbf{A}^H , \mathbf{A}^* , and \mathbf{A}^{-1} denote its transpose, Hermitian transpose, conjugate, and inverse, respectively. We denote \mathbf{I} as the identity matrix whose dimension will be clarified from the context, and denote $\mathbb{C}^{m \times n}$ ($\mathbb{R}^{m \times n}$) as the m -by- n dimensional complex (real) space. The notations $\mathbb{E}(\cdot)$, $|\cdot|$, $\Re(\cdot)$, and $\text{sign}(\cdot)$ represent the expectation, absolute value, Real part, and sign of an input variable, respectively. We denote \circ as the Hadamard product between two matrices. $\mathcal{CN}(\Upsilon, \Phi)$ denotes the circularly symmetric complex Gaussian (CSCG) distribution with mean Υ and covariance matrix Φ .

II. SYSTEM DESCRIPTION

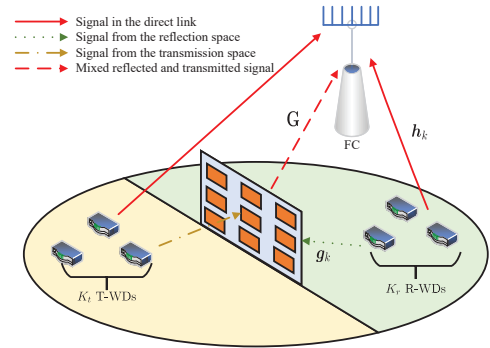


Fig. 1: A STAR-RIS assisted AirComp system.

In this section, we first introduce the investigated STAR-RIS assisted AirComp system. Then the optimization problem is formulated to minimize the computation MSE.

A. System Model

In this paper, we study a STAR-RIS assisted AirComp system as shown in Fig. 1, which consists of a multi-antenna FC, a STAR-RIS, and K single-antenna WDs. We assume that the FC is equipped with an uniform linear array (ULA) and the number of antennas is N . In this system, K WDs first collect one heterogeneous time-varying parameter

from the environment (e.g., temperature, noise, humidity), and then transmit the collected parameter to the FC via wireless channels for functional computation. We employ a STAR-RIS with M passive reflecting/transmitting elements to improve the computation MSE performance. In the conventional RIS assisted AirComp system, all WDs should be at the same side of the RIS due to the fact that the RIS can only perform signal reflection. However, the STAR-RIS is also able to transmit the incident signal, allowing that the WDs can be located at the different sides. Hence, the deployment of the STAR-RIS is much more flexible than the conventional one. Specifically, the space in Fig. 1 is separated into two parts by the STAR-RIS, i.e., the reflection space and the transmission space. Similarly, the WDs are also divided into two groups, called the R-WD and T-WD groups, which are located in the reflection and transmission space, respectively. We assume that the numbers of R-WDs and T-WDs are K_r and K_t with $K_r + K_t = K$. Therefore, the STAR-RIS can reflect the signal from the reflection space (the R-WDs) as well as transmit the signal from the transmission space (the T-WDs) after the processing of passive beamforming. For convenience, we respectively denote $\mathcal{K} \triangleq \{1, 2, \dots, K\}$, $\mathcal{K}_r \triangleq \{1, 2, \dots, K_r\}$, $\mathcal{K}_t \triangleq \{K_r + 1, \dots, K\}$, and $\mathcal{M} \triangleq \{1, 2, \dots, M\}$ as the sets of all WDs, R-WDs, T-WDs, and the passive reflect/transmit elements at the STAR-RIS. By applying the AirShare technique in [10], we assume that the synchronization of different WDs is established.

At a particular time slot, let $s_k \in \mathbb{C}$ denote the collected parameter by WD k . We would like to note that the parameter s_k is assumed to be normalized and independent of those from different WDs, i.e., $\mathbb{E}(s_k s_k^*) = 1$ and $\mathbb{E}(s_k s_j^*) = 0$, $k, j \in \mathcal{K}$. For swift data aggregation, the FC aims at computing the sum of collected parameters from all the WDs (i.e., $s = \sum_{k \in \mathcal{K}} s_k$). Besides, the design investigated in this paper can be readily extended to other nomographic functions [38]. By denoting v_k as the transmit coefficient at WD k , then the corresponding transmitted data is given by

$$x_k = v_k s_k. \quad (1)$$

Let P denote the transmit power budget and the transmit power constraints at the WDs is expressed as $|v_k|^2 \leq P$, $k \in \mathcal{K}$. By considering the deployment of the STAR-RIS, the incident signals at the STAR-RIS from two groups of WDs, i.e., the R-WDs and T-WDs, are written as

$$\mathbf{y}_r = \sum_{k \in \mathcal{K}_r} \mathbf{g}_k v_k s_k, \quad (2)$$

$$\mathbf{y}_t = \sum_{k \in \mathcal{K}_t} \mathbf{g}_k v_k s_k, \quad (3)$$

respectively, where $\mathbf{g}_k \in \mathbb{C}^{M \times 1}$ denotes the channel vector from WD k to the STAR-RIS. In this paper, we assume that the STAR-RIS operates based on the energy splitting protocol [35].¹ In this protocol, a single passive beamforming

matrix at the conventional RIS is divided into two parts at the STAR-RIS. One is designed for signal reflection, namely the passive reflect beamforming matrix, while the other is used for signal transmission, called the passive transmit beamforming matrix. Then the passive reflect and transmit beamforming matrices are denoted by $\mathbf{\Theta}_r \triangleq \sqrt{\beta_r} \text{diag}(\boldsymbol{\theta}_r^*) \in \mathbb{C}^{M \times M}$ and $\mathbf{\Theta}_t \triangleq \sqrt{\beta_t} \text{diag}(\boldsymbol{\theta}_t^*) \in \mathbb{C}^{M \times M}$, respectively, where $\beta_r, \beta_t \in [0, 1]$, $\boldsymbol{\theta}_r \triangleq [e^{j\theta_1^r}, e^{j\theta_2^r}, \dots, e^{j\theta_M^r}]^T \in \mathbb{C}^{M \times 1}$ and $\boldsymbol{\theta}_t \triangleq [e^{j\theta_1^t}, e^{j\theta_2^t}, \dots, e^{j\theta_M^t}]^T \in \mathbb{C}^{M \times 1}$ represent the diagonal reflect and transmit beamforming vectors, θ_m^r and θ_m^t denote the reflection and transmission phase-shift coefficients, respectively, with $\theta_m^r, \theta_m^t \in [0, 2\pi]$, $\forall m \in \mathcal{M}$.² When $\beta_t = 0$, the considered STAR-RIS reduces to a conventional RIS. Hence, the latter can be viewed as a special case. Since the correlation between β_r and β_t makes the system model and the corresponding problem much more complicated, to avoid high complexity and show more insights, we consider a simplified equal-energy splitting protocol in our initial attempt, i.e., $\beta_r = \beta_t = 0.5$. Furthermore, for ease of presentation, β_r and β_t are both scaled up to be 1, which means the STAR-RIS can perform full-power signal reflection and transmission simultaneously. To perform the signal reflection and transmission, the passive elements at the STAR-RIS need to support both electric and magnetic currents [32], [33]. Hence, we consider a coupled phase-shift model for STAR-RIS, which leads to the following condition for the phase-shift coefficients of the passive elements [36]:

$$\cos(\theta_m^t - \theta_m^r) = 0, \forall m \in \mathcal{M}. \quad (4)$$

This condition shows that the reflection and transmission phase-shift coefficients are coupled one by one and satisfy the constraint $|\theta_m^t - \theta_m^r| = \frac{\pi}{2}, \frac{3\pi}{2}, \forall m \in \mathcal{M}$. By jointly optimizing the passive reflect and transmit beamforming matrices at the STAR-RIS, we can obtain a high degree of flexibility for reducing the computation MSE in STAR-RIS assisted AirComp systems. After the processing of the passive reflect and transmit beamforming matrices, the mixed reflected and transmitted signal by the STAR-RIS is given by

$$\mathbf{x}_{\text{rt}} = \mathbf{\Theta}_r \sum_{k \in \mathcal{K}_r} \mathbf{g}_k v_k s_k + \mathbf{\Theta}_t \sum_{k \in \mathcal{K}_t} \mathbf{g}_k v_k s_k. \quad (5)$$

Hence, the received signal at the FC can be expressed as

$$\begin{aligned} \mathbf{y} = & \underbrace{\sum_{k \in \mathcal{K}_r} (\mathbf{h}_k + \mathbf{G} \mathbf{\Theta}_r \mathbf{g}_k) v_k s_k}_{\text{Signal from the reflection space}} + \underbrace{\sum_{k \in \mathcal{K}_t} (\mathbf{h}_k + \mathbf{G} \mathbf{\Theta}_t \mathbf{g}_k) v_k s_k}_{\text{Signal from the transmission space}} \\ & + \mathbf{n}, \end{aligned} \quad (6)$$

where $\mathbf{h}_k \in \mathbb{C}^{N \times 1}$, $\mathbf{G} \in \mathbb{C}^{N \times M}$, and $\mathbf{n} \in \mathbb{C}^{N \times 1}$ denote the channel vector from WD k to the FC, the channel matrix from the STAR-RIS to the FC, and the additive white Gaussian noise (AWGN) vector with $\mathbb{E}(\mathbf{n}) = \mathbf{0}$ and $\mathbb{E}(\|\mathbf{n}\|^2) = \sigma^2 \mathbf{I}$, respectively. As we can see, the received signal at the FC

¹In this work, for ease of analysis, the CSI related to the WDs, the STAR-RIS, and the FC is assumed to be known. This can be achieved by exploiting the uplink/downlink training methods in [26], [39]–[41]. A robust beamforming design can be developed by following the previous robust design techniques [42]–[44], which will be addressed in our future research.

²The continuous phase adjustment is assumed at the RIS's elements for ease of analysis. However, we would like to note that our proposed beamforming design can be also applied to the cases with the discrete phase adjustment. This can be achieved by projecting the continuous phase to the nearest element of the discrete-phase set.

can be separated into three parts, i.e., the signal from the reflection space, the signal from the transmission space, and the noise vector. Then the FC aims to compute the sum of the transmitted parameters from all the WDs (in the reflection and transmission spaces). Let $\mathbf{u} \in \mathbb{C}^{N \times 1}$ denote the receive beamforming vector at the FC, thus the obtained result after computing is expressed as

$$\hat{s} = \mathbf{u}^H \mathbf{y}. \quad (7)$$

B. Problem Formulation

With the above derivation, the computation performance can be measured by the MSE between s and \hat{s} , namely the computation MSE, which is given by

$$\begin{aligned} \zeta &= \mathbb{E}(|\hat{s} - s|^2) \\ &= \sum_{k \in \mathcal{K}_r} |\mathbf{u}^H (\mathbf{h}_k + \mathbf{G}\boldsymbol{\Theta}_r \mathbf{g}_k) v_k - 1|^2 \\ &\quad + \sum_{k \in \mathcal{K}_t} |\mathbf{u}^H (\mathbf{h}_k + \mathbf{G}\boldsymbol{\Theta}_t \mathbf{g}_k) v_k - 1|^2 + \sigma^2 \mathbf{u}^H \mathbf{u} \\ &= \sum_{k \in \mathcal{K}_r} |\mathbf{u}^H \mathbf{h}_k v_k + \boldsymbol{\Theta}_r^H \text{diag}(\mathbf{u}^H \mathbf{G}) \mathbf{g}_k v_k - 1|^2 \\ &\quad + \sum_{k \in \mathcal{K}_t} |\mathbf{u}^H \mathbf{h}_k v_k + \boldsymbol{\Theta}_t^H \text{diag}(\mathbf{u}^H \mathbf{G}) \mathbf{g}_k v_k - 1|^2 + \sigma^2 \mathbf{u}^H \mathbf{u}. \end{aligned} \quad (8)$$

In this work, we focus on minimizing the computation MSE by optimizing the transmit coefficients at the WDs, the receive beamforming vectors at the FC, and the diagonal reflect/transmit beamforming vectors at the STAR-RIS, i.e., solving the following problem:

$$\begin{aligned} \mathcal{P}1: \quad &\underset{\{v_k\}, \mathbf{u}, \boldsymbol{\Theta}_r, \boldsymbol{\Theta}_t}{\text{minimize}} \quad \zeta \\ &\text{subject to } |v_k|^2 \leq P, \forall k \in \mathcal{K}, \\ &\quad \boldsymbol{\Theta}_r(m) = e^{j\theta_m^r}, \forall m \in \mathcal{M}, \\ &\quad \boldsymbol{\Theta}_t(m) = e^{j\theta_m^t}, \forall m \in \mathcal{M}, \\ &\quad |\theta_m^t - \theta_m^r| = \frac{\pi}{2} \text{ or } \frac{3\pi}{2}, \forall m \in \mathcal{M}, \\ &\quad \theta_m^r, \theta_m^t \in [0, 2\pi], \forall m \in \mathcal{M}. \end{aligned} \quad (9)$$

One can see that, there are two kinds of constraints in problem $\mathcal{P}1$, i.e., the transmit power constraints at the WDs and the structure constraints at the STAR-RIS. Specifically, the second and third constraints are constant modulus constraints due to the structure of STAR-RIS, while the coupled phase-shift condition leads to the last constraint. Multiple variables are coupled in the non-convex and non-linear constraints, which makes problem $\mathcal{P}1$ more challenging. To the best of our knowledge, there lacks efficient algorithms to solve this non-convex beamforming design problem.

III. THE PROPOSED ALTERNATING OPTIMIZATION BEAMFORMING DESIGN

In this section, we propose an alternating-optimization (AO) algorithm based on the binary phase-shift constraints (BPCs), namely AO-BPC, to solve problem $\mathcal{P}1$. In particular, we first optimize the receive beamforming vector at the FC by fixing

the transmit coefficients at the WDs and the diagonal reflect/transmit beamforming vectors at the STAR-RIS based on the first-order optimality condition. Then the transmit coefficients are updated given the receive beamforming vector and the diagonal reflect/transmit beamforming vectors by exploiting the Lagrange duality method. Finally, the diagonal reflect and transmit beamforming vectors are decoupled by introducing an auxiliary binary vector and exploiting the binary phase-shift constraints. Then the closed-form solutions are derived based on the constant modulus constraints and the binary nature. The computational complexity and convergence of the proposed algorithm are both analyzed.

A. Optimization of Receive Beamforming Vector \mathbf{u}

By fixing $\{v_k\}$, $\boldsymbol{\Theta}_r$, and $\boldsymbol{\Theta}_t$, the subproblem with respect to the receive beamforming vector at the FC is given by

$$\mathcal{P}2: \underset{\mathbf{u}}{\text{minimize}} \quad \zeta. \quad (10)$$

Obviously, problem $\mathcal{P}2$ is unconstrained and convex, which can be solved by checking the first-order optimality condition. Therefore, the optimal solution for \mathbf{u} is expressed as

$$\begin{aligned} \mathbf{u} &= \left(\sum_{k \in \mathcal{K}_r} |v_k|^2 \mathbf{h}_{k,r} \mathbf{h}_{k,r}^H + \sum_{k \in \mathcal{K}_t} |v_k|^2 \mathbf{h}_{k,t} \mathbf{h}_{k,t}^H + \sigma^2 \mathbf{I} \right)^{-1} \\ &\quad \times \left(\sum_{k \in \mathcal{K}_r} \mathbf{h}_{k,r} v_k + \sum_{k \in \mathcal{K}_t} \mathbf{h}_{k,t} v_k \right), \end{aligned} \quad (11)$$

where $\mathbf{h}_{k,r} = \mathbf{h}_k + \mathbf{G}\boldsymbol{\Theta}_r \mathbf{g}_k$, $k \in \mathcal{K}_r$ and $\mathbf{h}_{k,t} = \mathbf{h}_k + \mathbf{G}\boldsymbol{\Theta}_t \mathbf{g}_k$, $k \in \mathcal{K}_t$ denote the equivalent channel vectors from the R-WDs and T-WDs to the FC, respectively. Compared to the sum-MMSE receive beamforming in conventional RIS assisted AirComp systems, which is in the form of $(\sum_{k \in \mathcal{K}} |v_k|^2 \mathbf{h}_k \mathbf{h}_k^H + \sigma^2 \mathbf{I})^{-1} \sum_{k \in \mathcal{K}} \mathbf{h}_k v_k$, the terms inside and outside the matrix inversion are quite different due to the fact that the R-WDs and T-WDs transmit signal to the FC via different paths.

B. Optimization of the Transmit Coefficients $\{v_k\}$

With given \mathbf{u} , $\boldsymbol{\Theta}_r$, and $\boldsymbol{\Theta}_t$, we can formulate the subproblem with respect to $\{v_k\}$ as follows:

$$\begin{aligned} \mathcal{P}3: \quad &\underset{\{v_k\}}{\text{minimize}} \quad \zeta \\ &\text{subject to } |v_k|^2 \leq P, \forall k \in \mathcal{K}. \end{aligned} \quad (12)$$

It can be observed that the transmit coefficients at the R-WDs and those at the T-WDs are separated in the objective function and constraint. Furthermore, the transmit coefficients among the R-WDs/T-WDs are also separated. Hence, we can solve problem $\mathcal{P}3$ by optimizing v_k in the order of k . Then the problem with respect to the transmit coefficient at R-WD k can be formulated as

$$\begin{aligned} \mathcal{P}4: \quad &\underset{v_k}{\text{minimize}} \quad |\mathbf{u}^H \mathbf{h}_{k,r} v_k - 1|^2 \\ &\text{subject to } |v_k|^2 \leq P. \end{aligned} \quad (13)$$

Similarly, the transmit coefficient optimization problem for T-WD k is given by

$$\begin{aligned} \mathcal{P5}: \text{minimize}_{v_k} \quad & |\mathbf{u}^H \mathbf{h}_{k,t} v_k - 1|^2 \\ \text{subject to} \quad & |v_k|^2 \leq P. \end{aligned} \quad (14)$$

Problem $\mathcal{P4}$ and $\mathcal{P5}$ are both convex problems with one quadratic constraint. Hence, the Lagrange duality method can be exploited to solve these two problems efficiently, leading to the following lemma.

Lemma 1: The optimal solutions for problem $\mathcal{P4}$ and $\mathcal{P5}$ are given by

$$v_k^* = \frac{\mathbf{h}_{k,r}^H \mathbf{u}}{\mathbf{u}^H \mathbf{h}_{k,r} \mathbf{h}_{k,r}^H \mathbf{u} + \mu_{k,r}^*}, \forall k \in \mathcal{K}_r, \quad (15)$$

$$v_k^* = \frac{\mathbf{h}_{k,t}^H \mathbf{u}}{\mathbf{u}^H \mathbf{h}_{k,t} \mathbf{h}_{k,t}^H \mathbf{u} + \mu_{k,t}^*}, \forall k \in \mathcal{K}_t, \quad (16)$$

respectively, where $\mu_{k,r}^*$ and $\mu_{k,t}^*$ denote the optimal Lagrange multipliers associated with the transmit power constraints in problem $\mathcal{P4}$ and $\mathcal{P5}$. If $\mathbf{u}^H \mathbf{h}_{k,r} \mathbf{h}_{k,r}^H \mathbf{u}$ and $\mathbf{u}^H \mathbf{h}_{k,t} \mathbf{h}_{k,t}^H \mathbf{u}$ are non-zero and

$$\frac{1}{\mathbf{u}^H \mathbf{h}_{k,r} \mathbf{h}_{k,r}^H \mathbf{u}} < P, \forall k \in \mathcal{K}_r, \quad (17)$$

$$\frac{1}{\mathbf{u}^H \mathbf{h}_{k,t} \mathbf{h}_{k,t}^H \mathbf{u}} < P, \forall k \in \mathcal{K}_t, \quad (18)$$

we choose $\mu_{k,r}^* = 0$ and $\mu_{k,t}^* = 0$; otherwise, we choose $\mu_{k,r}^*$ and $\mu_{k,t}^*$ such that the following conditions are satisfied

$$\left| \frac{\mathbf{u}^H \mathbf{h}_{k,r}}{\mathbf{u}^H \mathbf{h}_{k,r} \mathbf{h}_{k,r}^H \mathbf{u} + \mu_{k,r}^*} \right|^2 = P, \forall k \in \mathcal{K}_r, \quad (19)$$

$$\left| \frac{\mathbf{u}^H \mathbf{h}_{k,t}}{\mathbf{u}^H \mathbf{h}_{k,t} \mathbf{h}_{k,t}^H \mathbf{u} + \mu_{k,t}^*} \right|^2 = P, \forall k \in \mathcal{K}_t. \quad (20)$$

Proof: The proof of Lemma 1 is omitted due to the space limitation. ■

As we can see from Lemma 1, we obtain the solution for v_k by considering the following two cases: 1) When the transmit power budget is sufficiently large, we choose v_k to force the value of objective function to be zero; 2) when the transmit power budget is limited, v_k is optimized by fully exploiting the transmit power, namely the equality of transmit power constraint needs to be satisfied.

C. Optimization of the Diagonal Reflect and Transmit Beamforming Vectors $\boldsymbol{\theta}_r, \boldsymbol{\theta}_t$

By fixing $\{v_k\}$ and \mathbf{u} , the subproblem with respect to $\boldsymbol{\theta}_r$ and $\boldsymbol{\theta}_t$ is given by

$$\begin{aligned} \mathcal{P6}: \text{minimize}_{\boldsymbol{\theta}_r, \boldsymbol{\theta}_t} \quad & \zeta \\ \text{subject to} \quad & \boldsymbol{\theta}_r(m) = e^{j\theta_m^r}, \forall m \in \mathcal{M}, \\ & \boldsymbol{\theta}_t(m) = e^{j\theta_m^t}, \forall m \in \mathcal{M}, \\ & |\theta_m^t - \theta_m^r| = \frac{\pi}{2} \text{ or } \frac{3\pi}{2}, \forall m \in \mathcal{M}, \\ & \theta_m^r, \theta_m^t \in [0, 2\pi], \forall m \in \mathcal{M}. \end{aligned} \quad (21)$$

It is clear that problem $\mathcal{P6}$ is challenging due to the coupled phase-shift constraints and the constant modulus constraints. In the following, we first reformulate problem $\mathcal{P6}$ to a more tractable form by introducing an auxiliary binary vector and exploiting the coupled binary phase-shift constraints. Then the constant modulus constraints can be exploited to solve this relaxed problem by updating the elements in $\boldsymbol{\theta}_r$ and the introduced binary vector separately. Finally, $\boldsymbol{\theta}_t$ is optimized by considering the coupled phase-shift constraints.

From the coupled phase-shift constraints, it can be inferred that $\theta_m^t = \theta_m^r \pm \frac{\pi}{2}$ or $\frac{3\pi}{2}$. Then, we have

$$\boldsymbol{\theta}_t(m) = e^{j\theta_m^t} = \mathbf{q}(m) \sqrt{-1} e^{j\theta_m^r} = \mathbf{q}(m) \sqrt{-1} \boldsymbol{\theta}_r(m), \quad (22)$$

where $\mathbf{q} \in \mathbb{R}^{M \times 1}$ is a binary vector with $\mathbf{q}(m) \in \{-1, 1\}$. Thus the relationship between θ_m^t and θ_m^r depends on the value of \mathbf{q} , leading to a binary phase-shift constraint. Therefore, problem $\mathcal{P6}$ is reformulated to

$$\begin{aligned} \mathcal{P7}: \text{minimize}_{\boldsymbol{\theta}_r, \mathbf{q}} \quad & \hat{\zeta} \\ \text{subject to} \quad & \boldsymbol{\theta}_r(m) = e^{j\theta_m^r}, \forall m \in \mathcal{M}, \\ & \mathbf{q}(m) \in \{-1, 1\}, \forall m \in \mathcal{M}, \end{aligned} \quad (23)$$

where

$$\begin{aligned} \hat{\zeta} = & \sum_{k \in \mathcal{K}_r} |\mathbf{u}^H \mathbf{h}_k v_k + \boldsymbol{\theta}_r^H \text{diag}(\mathbf{u}^H \mathbf{G}) \mathbf{g}_k v_k - 1|^2 \\ & + \sum_{k \in \mathcal{K}_t} |\mathbf{u}^H \mathbf{h}_k v_k + (\boldsymbol{\theta}_r \circ \sqrt{-1} \mathbf{q})^H \text{diag}(\mathbf{u}^H \mathbf{G}) \mathbf{g}_k v_k - 1|^2 \\ & + \sigma^2 \mathbf{u}^H \mathbf{u}. \end{aligned} \quad (24)$$

One can see that, by denoting $\boldsymbol{\theta}_t$ as the Hadamard product of $\boldsymbol{\theta}_r, \mathbf{q}$, and the unit imaginary vector, the variables of problem $\mathcal{P7}$ are no longer coupled in the constraints and the challenge is only due to the constant modulus constraints. To deal with this issue, we apply the first-order optimality condition with constant modulus constraints. Since the variables $\boldsymbol{\theta}_r$ and \mathbf{q} are separated in the constraints, we can address problem $\mathcal{P7}$ in an alternating manner. Besides, due to the fact that the elements of $\boldsymbol{\theta}_r$ are also separated, $\boldsymbol{\theta}_r$ can be optimized by updating $\boldsymbol{\theta}_r(m)$ in the order of $m \in \mathcal{M}$, i.e., solving the following subproblem:

$$\begin{aligned} \mathcal{P8}: \text{minimize}_{\boldsymbol{\theta}_r(m)} \quad & \sum_{k \in \mathcal{K}_r} |\boldsymbol{\theta}_r^*(m) \mathbf{a}_k(m) + \mathbf{b}_k(m)|^2 \\ & + \sum_{k \in \mathcal{K}_t} |(\boldsymbol{\theta}_r(m) \sqrt{-1} \mathbf{q}(m))^* \mathbf{c}_k(m) + \mathbf{d}_k(m)|^2 \\ \text{subject to} \quad & \boldsymbol{\theta}_r(m) = e^{j\theta_m^r}, \end{aligned} \quad (25)$$

where

$$\mathbf{a}_k = \text{diag}(\mathbf{u}^H \mathbf{G}) \mathbf{g}_k v_k, \forall k \in \mathcal{K}_r, \quad (26)$$

$$\mathbf{b}_k(m) = \mathbf{u}^H \mathbf{h}_k v_k - 1 + \sum_{j \neq m} \boldsymbol{\theta}_r^*(j) \mathbf{a}_k(j), \quad (27)$$

$$\forall k \in \mathcal{K}_r, \forall m, j \in \mathcal{M}, \quad (28)$$

$$\mathbf{c}_k = \text{diag}(\mathbf{u}^H \mathbf{G}) \mathbf{g}_k v_k, \forall k \in \mathcal{K}_t, \quad (28)$$

$$\mathbf{d}_k(m) = \mathbf{u}^H \mathbf{h}_k v_k - 1 + \sum_{j \neq m} (\boldsymbol{\theta}_r(j) \sqrt{-1} \mathbf{q}(j))^* \mathbf{c}_k(j), \quad (29)$$

$$\forall k \in \mathcal{K}_t, \forall m, j \in \mathcal{M}. \quad (29)$$

Note that $|\boldsymbol{\theta}_r(m)|^2 = 1$. By ignoring some constant terms, problem $\mathcal{P}8$ can be equivalently rewritten as follows:

$$\begin{aligned} \mathcal{P}9: \text{minimize} \quad & \Re[\boldsymbol{\theta}_r^*(m)\eta] \\ & \text{subject to } \boldsymbol{\theta}_r(m) = e^{j\theta_r^m}, \end{aligned} \quad (30)$$

where $\eta = \sum_{k \in \mathcal{K}_r} \mathbf{a}_k(m) \mathbf{b}_k^*(m) + \sum_{k \in \mathcal{K}_t} \sqrt{-1} \mathbf{q}^*(m) \mathbf{c}_k(m) \mathbf{d}_k^*(m)$. Without the constant modulus constraint, the minimum objective function value of problem $\mathcal{P}9$ can be obtained by choosing $\boldsymbol{\theta}_r(m) = -\eta$. Therefore, the feasible solution for $\boldsymbol{\theta}_r(m)$ is given by

$$\boldsymbol{\theta}_r(m) = \frac{-\eta}{|\eta|}. \quad (31)$$

Similarly, $\mathbf{q}(m)$ can be updated element-wisely, which leads to the following problem:

$$\begin{aligned} \mathcal{P}10: \text{minimize} \quad & \sum_{k \in \mathcal{K}_t} |(\boldsymbol{\theta}_r(m) \sqrt{-1} \mathbf{q}(m))^* \mathbf{c}_k(m) + \mathbf{d}_k(m)|^2 \\ & \text{subject to } \mathbf{q}(m) \in \{-1, 1\}. \end{aligned} \quad (32)$$

Due to the fact that $\mathbf{q}(m)$ is a real variable and $|\mathbf{q}(m)|^2 = 1$, the objective function can be equivalently rewritten as $\mathbf{q}(m) \Re[(\boldsymbol{\theta}_r(m) \sqrt{-1})^* \mathbf{c}_k(m) \mathbf{d}_k^*(m)]$. It can be inferred that the objective function value is minimized by choosing $\mathbf{q}(m) = -1$ if $\Re[\sum_{k \in \mathcal{K}_t} (\boldsymbol{\theta}_r(m) \sqrt{-1})^* \mathbf{c}_k(m) \mathbf{d}_k^*(m)] \geq 0$; otherwise, $\mathbf{q}(m) = 1$. Hence, $\mathbf{q}(m)$ is updated by

$$\mathbf{q}(m) = -\text{sign}\{\Re[\sum_{k \in \mathcal{K}_t} (\boldsymbol{\theta}_r(m) \sqrt{-1})^* \mathbf{c}_k(m) \mathbf{d}_k^*(m)]\}. \quad (33)$$

Finally, $\boldsymbol{\theta}_t$ can be optimized by considering the binary phase-shift constraints, i.e.,

$$\boldsymbol{\theta}_t = \boldsymbol{\theta}_r \circ (\sqrt{-1} \mathbf{q}). \quad (34)$$

According to the above derivation, the AO-BPC algorithm for solving problem $\mathcal{P}1$ is summarized in Algorithm 1. One can see that, by exploiting the alternating optimization scheme and some conventional methods (e.g., the first-order optimality condition and the Lagrange multiplier method), we derive the closed-form solutions for \mathbf{u} and $\{v_k\}$ and then address problems $\mathcal{P}2$ and $\mathcal{P}3$ efficiently. Beside, since the coupled phase-shift constraints between $\boldsymbol{\theta}_r$ and $\boldsymbol{\theta}_t$ in the subproblem $\mathcal{P}6$ are non-convex and uncertain, we introduce the auxiliary binary vector \mathbf{q} and rewrite these constraints in the form of Hadamard product and reformulate $\mathcal{P}6$ into a more tractable problem with respect to $\boldsymbol{\theta}_r$ and \mathbf{q} , where the variables are

Algorithm 1 The AO-BPC Algorithm for Solving Problem $\mathcal{P}1$.

Initialize $\epsilon > 0$, $\{v_k\}$, \mathbf{u} , $\boldsymbol{\theta}_r$, and $\boldsymbol{\theta}_t$ such that they meet all the constraints;

Repeat

Update \mathbf{u} according to (11);

Update $\{v_k\}$ according to Lemma 1;

Update $\boldsymbol{\theta}_r$ and $\boldsymbol{\theta}_t$ according to (31), (33), and (34);

Until the decrease of the objective is below ϵ .

separated in the constraints. Finally, to tackle the binary constraints in the reformulated problem with respect to $\boldsymbol{\theta}_r$ and \mathbf{q} , inspiring by the binary nature of \mathbf{q} , we update these variables with their closed-form solutions in an alternating manner. Then the solution of $\boldsymbol{\theta}_t$ can be derived by considering the Hadamard-product constraint.

D. The Analysis of Complexity and Convergence

Considering (11), the complexity of updating \mathbf{u} is $\mathcal{O}(K(M^2 + MN + N) + N^3)$. Similarly, the complexity of optimizing $\{v_k\}$ is $\mathcal{O}(K(M^2 + MN + N))$. One can see that, the complexities of optimizing $\boldsymbol{\theta}_r$ and \mathbf{q} are both $\mathcal{O}(KM)$. The complexity of updating $\boldsymbol{\theta}_t$ based on (34) is $\mathcal{O}(M)$. Then the overall complexity of Algorithm 1 is $\mathcal{O}(K(M^2 + MN + N) + N^3 + KM + M)$. The convergence of Algorithm 1 is also analyzed, which is summarized in the following theorem.

Theorem 1: It is guaranteed that any limit point of the sequence generated by the proposed AO-BPC algorithm in Algorithm 1 converges to a stationary solution of problem $\mathcal{P}1$.

Proof: See Appendix A. ■

IV. COMPARISON WITH THE CONVENTIONAL RIS ASSISTED AIRCOMP SYSTEMS

In this section, we first derive the computation MSE for the conventional RIS assisted AirComp systems. Then under some reasonable assumptions, the computation MSEs for the STAR-RIS assisted and conventional RIS assisted systems are simplified. Finally, we theoretically prove that the STAR-RIS can provide more performance gain than the conventional RIS with respect to the computation MSE.

By replacing the STAR-RIS with a conventional RIS, the received signal at the FC is given by

$$\tilde{\mathbf{y}} = \sum_{k \in \mathcal{K}_r} (\mathbf{h}_k + \mathbf{G} \boldsymbol{\Theta} \mathbf{g}_k) v_k s_k + \sum_{k \in \mathcal{K}_t} \mathbf{h}_k v_k s_k + \mathbf{n}, \quad (35)$$

where $\boldsymbol{\Theta} \in \mathbb{C}^{M \times M}$ is the passive beamforming matrix at the RIS. Note that since the conventional RIS cannot perform signal transmission, there are only direct links between the FC and the WDs located in the transmission space. By exploiting the receive beamforming vector $\tilde{\mathbf{u}} \in \mathbb{C}^{N \times 1}$, the computed result is expressed as

$$\tilde{s} = \tilde{\mathbf{u}}^H \tilde{\mathbf{y}}. \quad (36)$$

Then the computation MSE for the conventional RIS assisted AirComp systems is formulate as follows:

$$\begin{aligned}\tilde{\zeta} &= \mathbb{E}(|\tilde{s} - s|^2) \\ &= \sum_{k \in \mathcal{K}_r} |\tilde{\mathbf{u}}^H (\mathbf{h}_k + \mathbf{G}\mathbf{\Theta} \mathbf{g}_k) v_k - 1|^2 + \sum_{k \in \mathcal{K}_t} |\tilde{\mathbf{u}}^H \mathbf{h}_k v_k - 1|^2 \\ &\quad + \sigma^2 \tilde{\mathbf{u}}^H \tilde{\mathbf{u}}.\end{aligned}\quad (37)$$

By assuming that $\tilde{\mathbf{u}}$ is also optimized by the first-order optimality condition, then $\tilde{\mathbf{u}}$ can be updated by

$$\begin{aligned}\tilde{\mathbf{u}} &= \left(\sum_{k \in \mathcal{K}_r} |v_k|^2 \tilde{\mathbf{h}}_{k,r} \tilde{\mathbf{h}}_{k,r}^H + \sum_{k \in \mathcal{K}_t} |v_k|^2 \mathbf{h}_k \mathbf{h}_k^H + \sigma^2 \mathbf{I} \right)^{-1} \\ &\quad \times \left(\sum_{k \in \mathcal{K}_r} \tilde{\mathbf{h}}_{k,r} v_k + \sum_{k \in \mathcal{K}_t} \mathbf{h}_k v_k \right).\end{aligned}\quad (38)$$

where $\tilde{\mathbf{h}}_{k,r} = \mathbf{h}_k + \mathbf{G}\mathbf{\Theta} \mathbf{g}_k$. By substituting (38) into (37), the computation MSE can be rewritten as

$$\begin{aligned}\tilde{\zeta}(v_k, \mathbf{\Theta}) &= K - \left(\sum_{k \in \mathcal{K}_r} \tilde{\mathbf{h}}_{k,r} v_k + \sum_{k \in \mathcal{K}_t} \mathbf{h}_k v_k \right)^H \\ &\quad \times \left(\sum_{k \in \mathcal{K}_r} |v_k|^2 \tilde{\mathbf{h}}_{k,r} \tilde{\mathbf{h}}_{k,r}^H + \sum_{k \in \mathcal{K}_t} |v_k|^2 \mathbf{h}_k \mathbf{h}_k^H + \sigma^2 \mathbf{I} \right)^{-1} \\ &\quad \times \left(\sum_{k \in \mathcal{K}_r} \tilde{\mathbf{h}}_{k,r} v_k + \sum_{k \in \mathcal{K}_t} \mathbf{h}_k v_k \right).\end{aligned}\quad (39)$$

Similarly, the computation MSE for the STAR-RIS assisted AirComp systems is given by

$$\begin{aligned}\zeta(v_k, \mathbf{\Theta}_r, \mathbf{\Theta}_t) &= K - \left(\sum_{k \in \mathcal{K}_r} \mathbf{h}_{k,r} v_k + \sum_{k \in \mathcal{K}_t} \mathbf{h}_{k,t} v_k \right)^H \\ &\quad \times \left(\sum_{k \in \mathcal{K}_r} |v_k|^2 \mathbf{h}_{k,r} \mathbf{h}_{k,r}^H + \sum_{k \in \mathcal{K}_t} |v_k|^2 \mathbf{h}_{k,t} \mathbf{h}_{k,t}^H + \sigma^2 \mathbf{I} \right)^{-1} \\ &\quad \times \left(\sum_{k \in \mathcal{K}_r} \mathbf{h}_{k,r} v_k + \sum_{k \in \mathcal{K}_t} \mathbf{h}_{k,t} v_k \right).\end{aligned}\quad (40)$$

For the ease of analysis, we consider the following assumption:

Assumption 1: We assume that the number of antennas at the FC and the number of elements at the STAR-RIS/RIS are sufficiently large such that the channel vectors of different WDs are orthogonal, i.e.,

$$\begin{aligned}\frac{1}{N} \mathbf{h}_k^H \mathbf{h}_{k'} &= 0, \quad \frac{1}{N} \mathbf{h}_{j,r}^H \mathbf{h}_{j',r} = 0, \\ \forall k, k' \in \mathcal{K}, k \neq k', \forall j, j' \in \mathcal{K}_r, j \neq j',\end{aligned}\quad (41)$$

$$\begin{aligned}\frac{1}{N} \mathbf{h}_{k,t}^H \mathbf{h}_{k',t} &= 0, \quad \frac{1}{N} \mathbf{h}_{j,r}^H \mathbf{h}_{j',t} = 0, \\ \forall k, k' \in \mathcal{K}_t, k \neq k', \forall j \in \mathcal{K}_r, j' \in \mathcal{K}_t,\end{aligned}\quad (42)$$

$$\begin{aligned}\frac{1}{N} \tilde{\mathbf{h}}_{k,r}^H \tilde{\mathbf{h}}_{k',r} &= 0, \quad \frac{1}{N} \tilde{\mathbf{h}}_{j,r}^H \mathbf{h}_{j'} = 0, \\ \forall k, k' \in \mathcal{K}_r, k \neq k', \forall j \in \mathcal{K}_r, j' \in \mathcal{K}_t.\end{aligned}\quad (43)$$

With the above assumption, we can simplify the computation MSEs in (39) and (40) to a more tractable form, leading to the following lemma.

Lemma 2: With Assumption 1, (39) and (40) can be approximately written as follows:

$$\begin{aligned}\tilde{\zeta}(v_k, \mathbf{\Theta}) &\approx K - \frac{1}{\sigma^2} \left(\sum_{k \in \mathcal{K}_r} \frac{\sigma^2 |v_k|^2 \tilde{\mathbf{h}}_{k,r}^H \tilde{\mathbf{h}}_{k,r}}{\sigma^2 + |v_k|^2 \tilde{\mathbf{h}}_{k,r}^H \tilde{\mathbf{h}}_{k,r}} \right. \\ &\quad \left. + \sum_{k \in \mathcal{K}_t} \frac{\sigma^2 |v_k|^2 \mathbf{h}_k^H \mathbf{h}_k}{\sigma^2 + |v_k|^2 \mathbf{h}_k^H \mathbf{h}_k} \right),\end{aligned}\quad (44)$$

$$\begin{aligned}\zeta(v_k, \mathbf{\Theta}_r, \mathbf{\Theta}_t) &\approx K - \frac{1}{\sigma^2} \left(\sum_{k \in \mathcal{K}_r} \frac{\sigma^2 |v_k|^2 \mathbf{h}_{k,r}^H \mathbf{h}_{k,r}}{\sigma^2 + |v_k|^2 \mathbf{h}_{k,r}^H \mathbf{h}_{k,r}} \right. \\ &\quad \left. + \sum_{k \in \mathcal{K}_t} \frac{\sigma^2 |v_k|^2 \mathbf{h}_{k,t}^H \mathbf{h}_{k,t}}{\sigma^2 + |v_k|^2 \mathbf{h}_{k,t}^H \mathbf{h}_{k,t}} \right),\end{aligned}\quad (45)$$

respectively.

Proof: See Appendix B. ■

According to Lemma 2, we theoretically prove that the computation MSE of STAR-RIS assisted AirComp system is less than that of conventional RIS assisted AirComp system, which is summarized in the following theorem.

Theorem 2: Let \tilde{v}_k^* (v_k^*) and $\tilde{\mathbf{\Theta}}^*$ ($\mathbf{\Theta}_r^*, \mathbf{\Theta}_t^*$) denote the optimal transmit coefficient at the WDs and passive beamforming matrix for conventional RIS assisted (STAR-RIS assisted) AirComp systems. The computation MSE of the STAR-RIS assisted AirComp system is less than that of the conventional RIS assisted system, i.e., $\zeta(v_k^*, \mathbf{\Theta}_r^*, \mathbf{\Theta}_t^*) < \tilde{\zeta}(\tilde{v}_k^*, \tilde{\mathbf{\Theta}}^*)$.

Proof: See Appendix C. ■

V. SIMULATION RESULTS

In this section, we evaluate the computation MSE performance of the proposed AO-BPC algorithm by comparing with the following benchmark schemes:

- **CRIS:** In this benchmark scheme, the STAR-RIS is replaced with the conventional RIS. Hence, the signal from T-WDs will not be processed by the RIS, which means that there are only direct links between the FC and the T-WDs. Hence, by excluding the update of the diagonal transmit beamforming vector, we can also optimize the receive beamforming vector at the FC, the transmit coefficients at the WDs, and the diagonal reflect vector at the RIS according to the AO-BPC algorithm.
- **AO-RPC:** In this benchmark scheme, the random phase-shift constraint between $\mathbf{\theta}_t$ and $\mathbf{\theta}_r$ is considered, i.e., $\mathbf{\theta}_t = \mathbf{\theta}_r \circ \hat{\mathbf{q}}$ with $\hat{\mathbf{q}} \in \mathbb{C}^{M \times 1}$ and $|\hat{q}(m)| = 1$. In particular, we assume that $\hat{\mathbf{q}}$ is fixed in the beamforming design. Then we can update $\mathbf{\theta}_r$ and $\mathbf{\theta}_t$ by solving a reformulated subproblem in the form of problem $\mathcal{P}8$ and considering the random phase-shift constraint.
- **AO-WPC:** In this benchmark scheme, the phase-shift constraint between $\mathbf{\theta}_t$ and $\mathbf{\theta}_r$ is removed. Therefore, $\mathbf{\theta}_t$ and $\mathbf{\theta}_r$ are optimized separately by exploiting the method applied to problem $\mathcal{P}9$.

In the following simulations, the tightly spaced antennas and the limited angular spread of scattering environment are

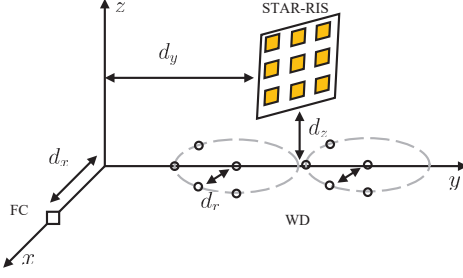


Fig. 2: The location setup in the simulations.

assumed. Thus we consider a general spatially correlated Rician fading channel model for the WD-RIS, WD-FC, and RIS-FC links, which consists of both line-of-sight (LoS) and non-LoS components and is widely used in literature [35]. In the LoS component, the distance-dependent path loss is defined as $L = C_0 \left(\frac{d}{D_0}\right)^{-\alpha}$ with C_0 , d , and α being the path loss at the reference distance $D_0 = 1$ meter (m), the individual distance, and the path loss exponent. To be clear, we denote α_{WR} ($\alpha_{tWF}/\alpha_{tWF}$ and α_{RF}) and β_{WR} (β_{tWF}/β_{tWF} and β_{SF}) as the path loss exponent and the Rician factor for the WD-RIS (WD-FC and RIS-FC) links. In the simulation setup, a three-dimensional system of Fig. 2 is considered, where the FC and the STAR-RIS are on the x -axis and y - z plane, respectively. The FC is assumed to equipped with a ULA, while an uniform rectangular array (URA) with $M = M_y M_z$ reflecting elements are deployed at the STAR-RIS with M_y and M_z denoting the number of elements along the y -axis and z -axis, respectively. The reference antenna/element at the FC/STAR-RIS are located at $(d_x = 2 \text{ m}, 0, 0)$ and $(0, d_y = 50 \text{ m}, d_z = 3 \text{ m})$. Besides, the R-WDs/T-WDs uniformly locate on a circle with radius $d_r = 5 \text{ m}$ whose center is at $(0, d_y - d_r = 45 \text{ m}, 0)/(0, d_y + d_r = 55 \text{ m}, 0)$. We define the signal-to-noise-ratio (SNR) as $\text{SNR} = \frac{P}{\sigma^2}$. Furthermore, the simulation parameter settings in Table II are considered unless otherwise specified.

TABLE I: The settings of the simulation parameters.

Parameter	Value	Parameter	Value
M	64	α_{tWF}	4
M_y	4	α_{RF}	3
M_z	M/M_y	d_x	2 m
N	64	d_y	50 m
K	64	d_z	3 m
K_r	$K/2$	d_r	3 m
K_t	$K/2$	β_{WS}	3 dB
σ^2	-80 dBm	β_{SF}	3 dB
C_0	-30 dB	β_{tWF}	-3 dB
D_0	1 m	β_{tWF}	-3 dB
α_{WR}	2.2	ϵ	10^{-4}
α_{tWF}	3.8		

A. Convergence

Fig. 3 shows the convergence behaviour of the AO-BPC algorithm with SNR = 5 dB, 15 dB, and 25 dB. In different SNR regions, the proposed algorithm converges to a stationary point after a number of iterations. The stationary point decreases when SNR increases and tends to be zero. Furthermore, in

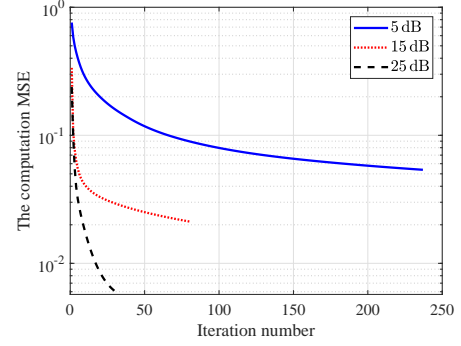


Fig. 3: Convergence behaviour of the AO-BPC algorithm with SNR = 5 dB, 15 dB, and 25 dB.

the moderate and high SNR region, the AO-BPC algorithm terminates within only a few iterations. However, one can see that the convergence speed slows down as the SNR becomes extremely low, which takes over 200 iterations to converge with SNR = 5 dB.

B. Impact of the Number of WDs

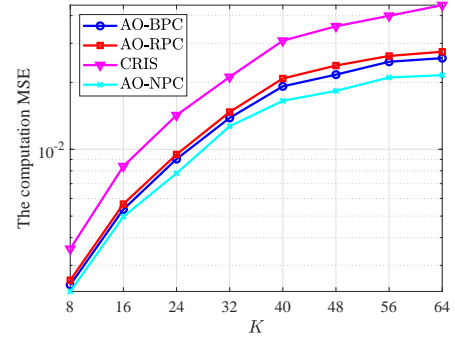


Fig. 4: The computation MSE performance versus K with SNR = 15 dB.

Fig. 4 compares the computation MSE performance of the considered schemes versus K with SNR = 15 dB. One can see that the computation MSEs of all the schemes increase with the number of WDs due to the fact that it is more challenging to design a single passive reflect/transmit matrix at the STAR-RIS/RIS and a single receive beamforming vector at the FC to aggregate the collected data from more WDs. Besides, the AO-BPC algorithm outperforms the CRIS algorithm with an increasing gap. This coincides with our discussion in Section IV that the STAR-RIS assisted AirComp system shows less computation MSE than the conventional RIS assisted system. The computation MSE achieved by the AO-BPC algorithm is less than that achieved by the AO-RPC algorithm with a minor gap, since these two algorithms both consider the coupled phase-shift constraints. The phase-shift constraints in the AO-BPC algorithm are binary, which can be decoupled by introducing an auxiliary binary vector. This vector can be optimized by exploiting its binary nature to improve the system performance. However, there lack efficient methods to address the random phase-shift constraints in the AO-RPC

algorithm, which leads to the minor gap in Fig. 4. Also, the random coupled phase-shift constraints may lead to high complexity and take more iterations to converge. The proposed algorithm approaches the AO-NPC algorithm, which can be viewed as a lower bound since the passive reflect and transmit beamforming matrices at the STAR-RIS are optimized by ignoring the coupled phase-shift constraints, which provides more spatial degrees. This further verifies the effectiveness of our proposed algorithm.

C. Impact of the Number of R-WDs

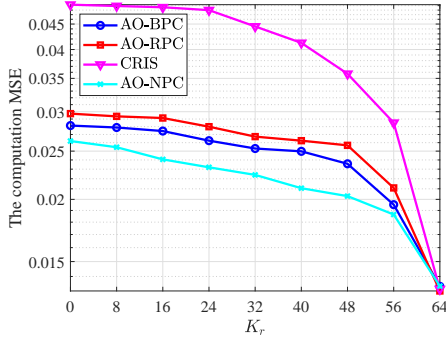


Fig. 5: The computation MSE performance versus K_r with $K = 64$ and $\text{SNR} = 15$ dB.

We compare the computation MSE performance versus K_r of the considered schemes with $K = 64$ and $\text{SNR} = 15$ dB. As we can see, as K_r increases, the computation MSEs of all the schemes decrease since more WDs become R-WDs with less pathloss according to Fig. 2. Also, these R-WDs can benefit from the passive reflection matrix at the STAR-RIS/RIS. In particular, the effect of the passive reflection matrix is much more obvious in the CRIS algorithm since the conventional RIS works only when the WDs is located at the reflection space. When $K_r = 0$, i.e., all the WDs are T-WDs, the STAR-RIS assisted AirComp system shows much less computation MSE than the conventional RIS assisted system. This is due to the fact that the signals from T-WDs are only transmitted via direct links between the T-WDs and the FC in the conventional RIS assisted systems since the RIS can only perform signal reflection, while in the STAR-RIS assisted system the signals can be transmitted by the STAR-RIS. Moreover, the computation MSEs achieved by the considered schemes coincide when $K_r = 64$. In this case, all the WDs become R-WDs and the STAR-RIS does not perform signal transmission, i.e., the STAR-RIS assisted AirComp systems reduce to the conventional RIS assisted ones.

D. Impact of the Number of Passive Reflecting/Transmitting Elements at the STAR-RIS/RIS

Fig. 6 shows the computation MSE performance of the considered schemes versus the number of passive reflecting/transmitting elements at the STAR-RIS/RIS with $\text{SNR} = 15$ dB. As M increases, the STAR-RIS/RIS can provide more channel gain and then the computation MSEs decrease. The

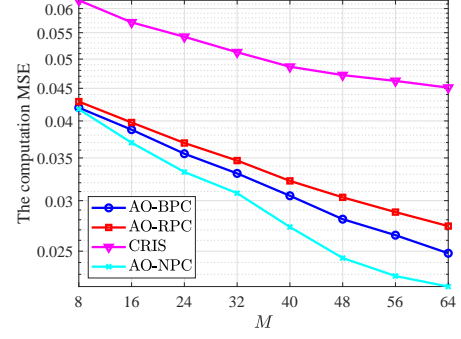


Fig. 6: The computation MSE performance versus M with $\text{SNR} = 15$ dB.

AO-BPC algorithm outperforms the AO-RPC and CRIS algorithms with increasing gaps due to the facts that the former can not only offer more spatial degrees for the T-WDs (i.e., signal transmission) but also better deal with the coupled phase-shift constraints between the passive reflect and transmit beamforming matrices at the STAR-RIS when M becomes larger. Also, the computation MSE of our proposed algorithm is lower bounded by that achieved by the AO-NPC algorithm. But the gap is quite small especially when M is less than 32.

E. Impact of the Number of Receive Antennas at the FC

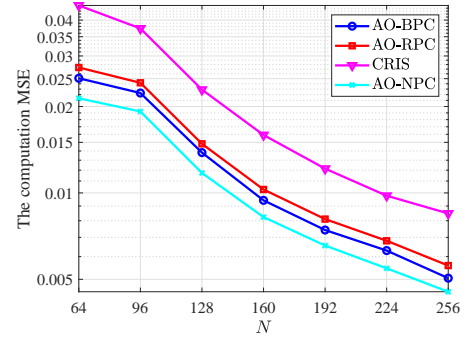


Fig. 7: The computation MSE performance versus N with $\text{SNR} = 15$ dB.

Fig. 7 compares the computation MSE performance versus the number of receive antennas at the FC with $\text{SNR} = 15$ dB. Since more receive antennas provide more spatial degrees, the computation MSEs of all the curves decrease as N increases. Specially, the computation MSE decreases slowly when N increases from 64 to 96, while the speed becomes faster when $N > 96$. It is due to the fact that the channel gain of direct links between the WDs and the FC tends to be dominant with sufficiently large number of receive antennas. The proposed AO-BPC algorithm still outperforms the AO-RPC and CRIS algorithms. Specifically, the gap between the AO-BPC and CRIS algorithms is much larger than that between the AO-BPC and AO-RPC algorithms, which verifies the outstanding advantages of STAR-RIS in AirComp systems.

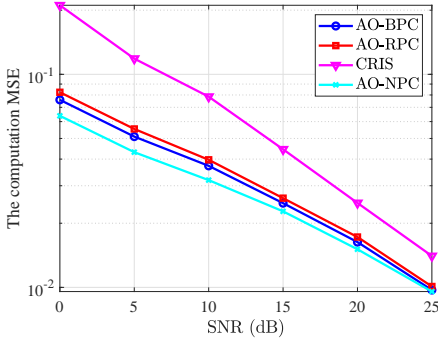


Fig. 8: The computation MSE performance versus SNR.

F. Impact of the SNR

Fig. 8 presents the computation MSE of the considered schemes versus SNR. As SNR increases, the computation MSE performance improves significantly due to the channel gain. Besides, it can be observed that the gap between the AO-BPC and AO-NPC algorithms turn to be smaller when SNR changes from 0 dB to 25 dB. It means that though the binary phase-shift constraints are considered in the AO-BPC algorithm, the computation MSE achieved by our proposed algorithm is quite close to that achieved by the AO-NPC algorithm in the high-SNR region, which is viewed as the performance lower bound. Such observation further verifies the efficiency of the AO-BPC algorithm.

VI. CONCLUSION

In this work, the joint beamforming design for the STAR-RIS assisted AirComp systems was investigated. An AO-BPC algorithm was proposed to jointly optimize the transmit coefficients at the R-WDs/T-WDs, the passive reflect/transmit beamforming matrices at the STAR-RIS, and the receive beamforming vector at the FC. In particular, we optimized the transmit coefficients and the receive beamforming vector by applying the Lagrange duality method and the first-order optimality condition. By introducing an auxiliary variable and exploiting the binary phase-shift constraints, we derived the closed-form solutions for the passive reflect and transmit beamforming matrices. The analysis of complexity and convergence for the proposed AO-BPC algorithm has been also presented. Moreover, under some reasonable assumptions, the expressions of the computation MSEs for the STAR-RIS assisted and conventional RIS assisted AirComp systems are simplified. Then it is proven that the STAR-RIS assisted AirComp system can achieve less computation MSE than the conventional RIS assisted system. Our simulation results showed that our proposed AO-BPC algorithm outperforms the AO-RPC and CRIS benchmark schemes. Besides, the computation MSE of the AO-BPC algorithm is close to the performance lower bound achieved by optimizing the passive reflect and transmit beamforming matrices without coupled phase-shift constraints. Our proposed AO-BPC algorithm can serve as an excellent candidate for STAR-RIS assisted AirComp system.

In the following, we discuss some issues that are not addressed yet in this work to motivate future research.

- As an initial attempt, we only considered the energy splitting protocol in this work. In general, STAR-RIS can operate in different protocols, e.g., mode switching and time switching. The performance and complexity of beamforming design may be significantly affected by the protocols. Therefore, it is important to investigate the STAR-RIS assisted AirComp systems under different protocols in the future.
- In this work, the CSI is assumed to be perfectly known at the WDs, the FC, and the STAR-RIS. In practise, since the number of elements at the STAR-RIS is large, it is challenging to estimate channel information and the error may be large. Hence, the robust beamforming design for STAR-RIS assisted AirComp systems is worthy of further investigation.
- Considering the hardware limitation, the STAR-RIS may only perform discrete phase shift in practice. Therefore, the beamforming design for STAR-RIS assisted AirComp systems with discrete phase shift is also an important topic.

APPENDIX A PROOF OF THEOREM 1

Since the alternating optimization method is exploited to solve problem $\mathcal{P}1$, we need to prove that the objective function value is non-increasing in each iteration. Obviously, considering the fact that the receive beamforming vector \mathbf{u} at the FC and the transmit beamforming coefficients $\{v_k\}$ are updated based on the first-order optimality condition and Lagrange duality method, respectively, we have

$$\zeta(\mathbf{u}_{n+1}, \{v_{k,n}\}, \boldsymbol{\theta}_{r,n}, \boldsymbol{\theta}_{t,n}) \leq \zeta(\mathbf{u}_n, \{v_{k,n}\}, \boldsymbol{\theta}_{r,n}, \boldsymbol{\theta}_{t,n}), \quad (46)$$

$$\zeta(\mathbf{u}_n, \{v_{k,n+1}\}, \boldsymbol{\theta}_{r,n}, \boldsymbol{\theta}_{t,n}) \leq \zeta(\mathbf{u}_n, \{v_{k,n}\}, \boldsymbol{\theta}_{r,n}, \boldsymbol{\theta}_{t,n}), \quad (47)$$

where n denotes the iteration number. Therefore, the optimization of \mathbf{u} and $\{v_k\}$ leads to the non-increasing objective function value.

Next, let us focus on problem $\mathcal{P}7$. By considering the constant modulus constraints, we update the diagonal reflect vector $\boldsymbol{\theta}_r$ at the STAR-RIS and the auxiliary variable \mathbf{q} element-wisely according to (31) and (33), which are the optimal and feasible solutions for problems $\mathcal{P}9$ and $\mathcal{P}10$, respectively. Furthermore, the objective functions of $\mathcal{P}7$, $\mathcal{P}9$, and $\mathcal{P}10$ are all equivalent. Then we have

$$\hat{\zeta}(\mathbf{u}_n, \{v_{k,n}\}, \boldsymbol{\theta}_{r,n+1}, \mathbf{q}_n) \leq \hat{\zeta}(\mathbf{u}_n, \{v_{k,n}\}, \boldsymbol{\theta}_{r,n}, \mathbf{q}_n), \quad (48)$$

$$\hat{\zeta}(\mathbf{u}_n, \{v_{k,n}\}, \boldsymbol{\theta}_{r,n}, \mathbf{q}_{n+1}) \leq \hat{\zeta}(\mathbf{u}_n, \{v_{k,n}\}, \boldsymbol{\theta}_{r,n}, \mathbf{q}_n). \quad (49)$$

Note that the diagonal transmit vector $\boldsymbol{\theta}_t$ is optimized by (34), which satisfies the coupled phase-shift constraints. Thus we have

$$\hat{\zeta}(\mathbf{u}_n, \{v_{k,n}\}, \boldsymbol{\theta}_{r,n}, \mathbf{q}_n) = \zeta(\mathbf{u}_n, \{v_{k,n}\}, \boldsymbol{\theta}_{r,n}, \boldsymbol{\theta}_{t,n}). \quad (50)$$

Combining (46)-(50), the objective function value is proven to be non-increasing in each iteration. Besides, since the objective

function is defined as the computation MSE, its value is lower-bounded by zero. Hence, the convergence of Algorithm 1 follows immediately.

APPENDIX B PROOF OF LEMMA 2

We first focus on the simplification of (39). As we can see, the matrix inverse term is the most challenging to deal with. By exploiting its special structure, we can apply the Sherman-Morrison formula [45]. The matrix inverse term can be rewritten as

$$\mathbf{D}_n = \begin{cases} \sigma^2 \mathbf{I}, & \text{if } n = 0, \\ \mathbf{D}_{n-1} + |v_n|^2 \tilde{\mathbf{h}}_{n,r} \tilde{\mathbf{h}}_{n,r}^H, & \text{else if } 1 < n \leq K_r, \\ \mathbf{D}_{n-1} + |v_n|^2 \mathbf{h}_n \mathbf{h}_n^H, & \text{otherwise.} \end{cases} \quad (51)$$

Then we have the following recursive formula for calculate the inverse of \mathbf{D}_n

$$\mathbf{D}_n^{-1} = \begin{cases} \mathbf{D}_{n-1}^{-1} - \frac{\mathbf{D}_{n-1}^{-1} |v_n|^2 \tilde{\mathbf{h}}_{n,r} \tilde{\mathbf{h}}_{n,r}^H \mathbf{D}_{n-1}^{-1}}{1 + |v_n|^2 \tilde{\mathbf{h}}_{n,r}^H \mathbf{D}_{n-1}^{-1} \tilde{\mathbf{h}}_{n,r}}, & \text{if } 1 \leq n \leq K_r \\ \mathbf{D}_{n-1}^{-1} - \frac{\mathbf{D}_{n-1}^{-1} |v_n|^2 \mathbf{h}_n \mathbf{h}_n^H \mathbf{D}_{n-1}^{-1}}{1 + |v_n|^2 \mathbf{h}_n^H \mathbf{D}_{n-1}^{-1} \mathbf{h}_n}, & \text{otherwise.} \end{cases} \quad (52)$$

By assuming that $K_r \geq 2$ and letting $n = 1$, we can obtain the inverse of \mathbf{D}_1

$$\mathbf{D}_1^{-1} = \frac{1}{\sigma^2} \left(\mathbf{I} - \frac{|v_1|^2 \tilde{\mathbf{h}}_{1,r} \tilde{\mathbf{h}}_{1,r}^H}{\sigma^2 + |v_1|^2 \tilde{\mathbf{h}}_{1,r}^H \tilde{\mathbf{h}}_{1,r}} \right). \quad (53)$$

Then we calculate the inverse of \mathbf{D}_2

$$\begin{aligned} \mathbf{D}_2^{-1} &= \mathbf{D}_1^{-1} - \frac{\mathbf{D}_1^{-1} |v_2|^2 \tilde{\mathbf{h}}_{2,r} \tilde{\mathbf{h}}_{2,r}^H \mathbf{D}_1^{-1}}{1 + |v_2|^2 \tilde{\mathbf{h}}_{2,r}^H \mathbf{D}_1^{-1} \tilde{\mathbf{h}}_{2,r} / \sigma^2} \\ &\approx \frac{1}{\sigma^2} \left(\mathbf{I} - \frac{|v_1|^2 \tilde{\mathbf{h}}_{1,r} \tilde{\mathbf{h}}_{1,r}^H}{\sigma^2 + |v_1|^2 \tilde{\mathbf{h}}_{1,r}^H \tilde{\mathbf{h}}_{1,r}} - \frac{|v_2|^2 \tilde{\mathbf{h}}_{2,r} \tilde{\mathbf{h}}_{2,r}^H}{\sigma^2 + |v_2|^2 \tilde{\mathbf{h}}_{2,r}^H \tilde{\mathbf{h}}_{2,r}} \right), \end{aligned} \quad (54)$$

where the approximation is due to (43). Therefore, we can derive the inverse of \mathbf{D}_K as follows

$$\begin{aligned} \mathbf{D}_K^{-1} &\approx \frac{1}{\sigma^2} \left(\mathbf{I} - \sum_{k \in \mathcal{K}_r} \frac{|v_k|^2 \tilde{\mathbf{h}}_k \tilde{\mathbf{h}}_k^H}{\sigma^2 + |v_k|^2 \tilde{\mathbf{h}}_{k,r}^H \tilde{\mathbf{h}}_{k,r}} \right. \\ &\quad \left. - \sum_{k \in \mathcal{K}_t} \frac{|v_k|^2 \mathbf{h}_k \mathbf{h}_k^H}{\sigma^2 + |v_k|^2 \mathbf{h}_k^H \mathbf{h}_k} \right). \end{aligned} \quad (55)$$

By substituting (55) into (39), we have

$$\begin{aligned} \tilde{\zeta}(v_k, \boldsymbol{\Theta}) &\approx K - \left(\sum_{k \in \mathcal{K}_r} \tilde{\mathbf{h}}_{k,r} v_k + \sum_{k \in \mathcal{K}_t} \mathbf{h}_k v_k \right)^H \\ &\times \mathbf{D}_K^{-1} \left(\sum_{k \in \mathcal{K}_r} \tilde{\mathbf{h}}_{k,r} v_k + \sum_{k \in \mathcal{K}_t} \mathbf{h}_k v_k \right) \\ &\approx K - \frac{1}{\sigma^2} \left(\sum_{k \in \mathcal{K}_r} |v_k|^2 \tilde{\mathbf{h}}_{k,r}^H \tilde{\mathbf{h}}_{k,r} + \sum_{k \in \mathcal{K}_t} |v_k|^2 \mathbf{h}_k^H \mathbf{h}_k \right. \\ &\quad \left. - \sum_{k \in \mathcal{K}_r} \frac{|v_k|^4 |\tilde{\mathbf{h}}_{k,r}^H \tilde{\mathbf{h}}_{k,r}|^2}{\sigma^2 + |v_k|^2 \tilde{\mathbf{h}}_{k,r}^H \tilde{\mathbf{h}}_{k,r}} - \sum_{k \in \mathcal{K}_t} \frac{|v_k|^4 |\mathbf{h}_k^H \mathbf{h}_k|^2}{\sigma^2 + |v_k|^2 \mathbf{h}_k^H \mathbf{h}_k} \right) \\ &= K - \frac{1}{\sigma^2} \left(\sum_{k \in \mathcal{K}_r} \frac{\sigma^2 |v_k|^2 \tilde{\mathbf{h}}_{k,r}^H \tilde{\mathbf{h}}_{k,r}}{\sigma^2 + |v_k|^2 \tilde{\mathbf{h}}_{k,r}^H \tilde{\mathbf{h}}_{k,r}} \right. \\ &\quad \left. + \sum_{k \in \mathcal{K}_t} \frac{\sigma^2 |v_k|^2 \mathbf{h}_k^H \mathbf{h}_k}{\sigma^2 + |v_k|^2 \mathbf{h}_k^H \mathbf{h}_k} \right), \end{aligned} \quad (56)$$

where the second approximation is due to (42) and (43). (45) can be obtained similarly and the results follows immediately.

APPENDIX C PROOF OF THEOREM 2

First, we assume that the STAR-RIS assisted AirComp systems also adopt the optimal transmit coefficients and passive reflect beamforming matrix of the conventional RIS assisted system. Then we have

$$\begin{aligned} &\zeta(\tilde{v}_k^*, \tilde{\boldsymbol{\Theta}}^*, \boldsymbol{\Theta}_t) - \tilde{\zeta}(\tilde{v}_k^*, \tilde{\boldsymbol{\Theta}}^*) \\ &= \frac{1}{\sigma^2} \sum_{k \in \mathcal{K}_t} \left(\frac{\sigma^2 |v_k|^2 \mathbf{h}_k^H \mathbf{h}_k}{\sigma^2 + |v_k|^2 \mathbf{h}_k^H \mathbf{h}_k} - \frac{\sigma^2 |v_k|^2 \mathbf{h}_{k,t}^H \mathbf{h}_{k,t}}{\sigma^2 + |v_k|^2 \mathbf{h}_{k,t}^H \mathbf{h}_{k,t}} \right) \\ &= \frac{1}{\sigma^2} \sum_{k \in \mathcal{K}_t} \left(\frac{\sigma^2 e_k}{\sigma^2 + e_k} - \frac{\sigma^2 e_k + f_k}{\sigma^2 + e_k + f_k} \right), \end{aligned} \quad (57)$$

where $e_k = |v_k|^2 \mathbf{h}_k^H \mathbf{h}_k$ and $f_k = \sigma^2 (2\Re(\mathbf{h}_k^H \mathbf{G} \boldsymbol{\Theta}_t \mathbf{g}_k) + \mathbf{g}_k^H \boldsymbol{\Theta}_t^H \mathbf{G}^H \mathbf{G} \boldsymbol{\Theta}_t \mathbf{g}_k)$. Clearly, we can choose $\boldsymbol{\Theta}_t = \tilde{\boldsymbol{\Theta}}_t$ such that the term $\Re(\mathbf{h}_k^H \mathbf{G} \tilde{\boldsymbol{\Theta}}_t \mathbf{g}_k)$ is positive. Then f_k is positive. Since we have $\frac{a}{b} < \frac{a+c}{b+c}$ for any positive numbers a, b, c , the following inequality holds

$$\frac{\sigma^2 e_k}{\sigma^2 + e_k} < \frac{\sigma^2 e_k + f_k}{\sigma^2 + e_k + f_k}, \quad (58)$$

which leads to $\zeta(\tilde{v}_k^*, \tilde{\boldsymbol{\Theta}}_t^*, \tilde{\boldsymbol{\Theta}}_t) < \tilde{\zeta}(\tilde{v}_k^*, \tilde{\boldsymbol{\Theta}}^*)$. Note that $\zeta(v_k^*, \boldsymbol{\Theta}_r^*, \boldsymbol{\Theta}_t^*) \leq \zeta(\tilde{v}_k^*, \tilde{\boldsymbol{\Theta}}^*, \tilde{\boldsymbol{\Theta}}_t)$ always holds, therefore the results follow immediately.

REFERENCES

- [1] M. Agiwal, A. Roy, and N. Saxena, "Next generation 5G wireless networks: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 3, pp. 1617-1655, 3rd quater, 2016.
- [2] J. Lin, W. Yu, N. Zhang, X. Yang, H. Zhang, and W. Zhao, "A survey on internet of things: Architecture, enabling technologies, security and privacy, and applications," *IEEE Internet Things J.*, vol. 4, no. 5, pp. 1125-1142, Oct. 2017.
- [3] X. Chen, D. W. K. Ng, W. Yu, E. G. Larsson, N. Al-Dhahir, and R. Schober, "Massive access for 5G and beyond," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 3, pp. 615-637, Mar. 2021.

- [4] X. Cao, G. Zhu, J. Xu, and K. Huang, "Optimal power control for over-the-air computation in fading channels," *IEEE Trans. Wireless Commun.*, vol. 19, no. 11, pp. 7498-7513, Nov. 2020.
- [5] 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.
- [6] J.-J. Xiao, S. Cui, Z.-Q. Luo, and A. J. Goldsmith, "Linear coherent decentralized estimation," *IEEE Trans. Signal Process.*, vol. 56, no. 2, pp. 757-770, Feb. 2008.
- [7] C.-H. Wang, A. S. Leong, and S. Dey, "Distortion outage minimization and diversity order analysis for coherent multiaccess," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 6144-6159, Dec. 2011.
- [8] 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.
- [9] G. Zhu and K. Huang, "MIMO over-the-air computation for high-mobility multi-modal sensing," *IEEE Internet Things J.*, vol. 6, no. 4, pp. 6089-6103, Aug. 2019.
- [10] O. Abari, H. Rahul, D. Katabi, and M. Pant, "Airshare: Distributed coherent transmission made seamless," in *Proc. IEEE INFOCOM*, pp. 1742-1750, Apr. 2015.
- [11] 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.
- [12] X. Zhai, X. Chen, J. Xu, and D. W. K. Ng, "Hybrid beamforming for massive MIMO over-the-air computation," *IEEE Trans. Commun.*, vol. 69, no. 4, pp. 2737-2751, Apr. 2021.
- [13] Y.-S. Jeon, M. M. Amiri, J. Li, and H. V. Poor, "Gradient estimation for federated learning over massive MIMO communication systems," [Online]. <https://arxiv.org/pdf/2003.08059.pdf>, 2020.
- [14] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5394-5409, Nov. 2019.
- [15] Q. Wu and R. Zhang, "Beamforming optimization for wireless network aided by intelligent reflecting surface with discrete phase shifts," *IEEE Trans. Commun.*, vol. 68, no. 3, pp. 1838-1851, Mar. 2020.
- [16] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, and C. Yuen, "Reconfigurable intelligent surfaces for energy efficiency in wireless communication," *IEEE Trans. Wireless Commun.*, vol. 18, no. 8, pp. 4157-4170, Aug. 2019.
- [17] M. Di Renzo *et al.*, "Smart radio environments empowered by AI reconfigurable meta-surfaces: An idea whose time has come," *EURASIP J. Wireless Commun. Netw.*, May 2019. [Online]. <https://doi.org/10.1186/s13638-019-1438-9>.
- [18] E. Basar, M. Di Renzo, J. De Rosny, M. Debbah, M. Alouini, and R. Zhang, "Wireless communications through reconfigurable intelligent surfaces," *IEEE Access*, vol. 7, pp. 116753-116773, Aug. 2019.
- [19] Q. Wu and R. Zhang, "Weighted sum power maximization for intelligent reflecting surface aided SWIPT," *IEEE Wireless Commun. Lett.*, vol. 9, no. 5, pp. 587-590, May 2020.
- [20] Q. Wu and R. Zhang, "Joint active and passive beamforming optimization for intelligent reflecting surface assisted SWIPT under QoS constraints," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1735-1748, Aug. 2020.
- [21] M. Cui, G. Zhang, and R. Zhang, "Secure wireless communication via intelligent reflecting surface," *IEEE Wireless Commun. Lett.*, vol. 8, no. 5, pp. 1410-1414, Oct. 2019.
- [22] D. Xu, X. Yu, Y. Sun, D. W. K. Ng, and R. Schober, "Resource allocation for secure IRS-assisted multiuser MISO systems," in *Proc. IEEE GLOBECOM Wkshps.*, pp. 1-6, Dec. 2019.
- [23] X. Guan, Q. Wu, and R. Zhang, "Intelligent reflecting surface assisted secrecy communication: Is artificial noise helpful or not?" *IEEE Wireless Commun. Lett.*, vol. 9, no. 6, pp. 778-782, Jun. 2020.
- [24] W. Jiang, Y. Zhang, J. Wu, W. Feng, and Y. Jin, "Intelligent reflecting surface assisted secure wireless communications with multiple-transmit and multiple-receive antennas," *IEEE Access*, vol. 8, pp. 86659-86673, May 2020.
- [25] Y. Yang, B. Zheng, S. Zhang, and R. Zhang, "Intelligent reflecting surface meets OFDM: Protocol design and rate maximization," *IEEE Trans. Commun.*, vol. 68, no. 7, pp. 4522-4535, Jul. 2020.
- [26] B. Zheng and R. Zhang, "Intelligent reflecting surface-enhanced OFDM: Channel estimation and reflection optimization," *IEEE Wireless Commun. Lett.*, vol. 9, no. 4, pp. 518-522, Apr. 2020.
- [27] T. Jiang and Y. Shi, "Over-the-air computation via intelligent reflecting surfaces," in *Proc. IEEE GLOBECOM*, pp. 1-6, Feb. 2020.
- [28] D. Yu, S.-H. Park, O. Simeone, S. S. Shitz, "Optimizing over-the-air computation in IRS-aided C-RAN systems," in *Proc. IEEE SPAWC*, pp. 1-5, Aug. 2020.
- [29] Z. Wang, Y. Shi, Y. Zhou, H. Zhou, and N. Zhang, "Wireless-powered over-the-air computation in intelligent reflecting surface aided IoT networks," *IEEE Internet Things J.*, vol. 8, no. 3, pp. 1585-1598, Feb. 2021.
- [30] X. Zhai, G. Han, Y. Cai, and L. Hanzo, "Beamforming design based on two-stage stochastic optimization for RIS-assisted over-the-air computation systems," *IEEE Internet Things J.*, Early Access, Aug. 2021.
- [31] J. Xu, Y. Liu, X. Mu, and O. A. Dobre, "STAR-RISs: Simultaneous transmitting and reflecting reconfigurable intelligent surfaces," *IEEE Commun. Lett.*, vol. 25, no. 9, pp. 3134-3138, Sep. 2021.
- [32] Y. Liu, X. Mu, J. Xu, R. Schober, Y. Hao, H. V. Poor, and L. Hanzo, "STAR: Simultaneous transmission and reflection for 360° coverage by intelligent surfaces," *IEEE Wireless Commun.*, vol. 28, no. 6, pp. 102-109, Dec. 2021.
- [33] B. Zhu, K. Chen, N. Jia, L. Sun, J. Zhao, T. Jiang, and Y. Feng, "Dynamic control of electromagnetic wave propagation with the equivalent principle inspired tunable metasurface," *Sci. Rep.*, vol. 4, no. 4971, May 2014.
- [34] NTT DOCOMO, "Docomo conducts world's first successful trial of transparent dynamic metasurface," [Online]. www.nttdocomo.co.jp/english/info/mediacenter/pr/2020/0117_00.html.
- [35] X. Mu, Y. Liu, L. Guo, J. Lin, and R. Schober, "Simultaneously transmitting and reflecting (STAR) RIS aided wireless communications," *IEEE Trans. Wireless Commun.*, Early Access, pp. 1-16, Oct. 2021.
- [36] Y. Liu, X. Mu, R. Schober, and H. V. Poor, "Simultaneously transmitting and reflecting (STAR)-RISs: A coupled phase-shift model," [Online]. <https://arxiv.org/pdf/2110.02374v1.pdf>, 2021.
- [37] H. Niu and X. Liang, "Weighted sum-rate maximization for STAR-RISs-aided networks with coupled phase-shifters," *IEEE Syst. J.*, Early Access, pp. 1-10, Apr. 2022.
- [38] M. Goldenbaum, H. Boche, and S. Stańczak, "Nomographic functions: Efficient computation in clustered Gaussian sensor networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 4, pp. 2093-2105, Apr. 2015.
- [39] X. Rao and V. K. N. Lau, "Distributed compressive CSIT estimation and feedback for FDD multi-user massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 62, no. 12, pp. 3261-3271, Jun. 2014.
- [40] M. Soltanalian, M. M. Naghsh, N. Shariati, P. Stoica, and B. Hassibi, "Training signal design for correlated massive MIMO channel estimation," *IEEE Trans. Wireless Commun.*, vol. 16, no. 2, pp. 1135-1143, Feb. 2017.
- [41] M. Goldenbaum and S. Stańczak, "On the channel estimation effort for analog computation over wireless multiple-access channels," *IEEE Wireless Commun. Lett.*, vol. 3, pp. 261-264, Jun. 2014.
- [42] M. Zhao, Y. Cai, Q. Shi, B. Champagne, and M. Zhao, "Robust transceiver design for MISO interference channel with energy harvesting," *IEEE Trans. Signal Process.*, vol. 64, no. 17, pp. 4618-4633, Sep. 2016.
- [43] Y. Cai, M. Zhao, Q. Shi, B. Champagne, and M. Zhao, "Joint transceiver design algorithms for multiuser MISO relay systems with energy harvesting," *IEEE Trans. Commun.*, vol. 64, no. 10, pp. 4147-4164, Oct. 2016.
- [44] Y. Cai, Q. Shi, B. Champagne, and G. Y. Li, "Joint transceiver design for secure downlink communications over an amplify-and-forward MIMO relay," *IEEE Trans. Commun.*, vol. 65, no. 9, pp. 3691-3704, Sep. 2017.
- [45] K. B. Petersen and M. S. Pedersen, "The matrix cookbook," Nov. 2012 [Online]. <https://www.math.uwaterloo.ca/hwolkowi/matrixcookbook.pdf>



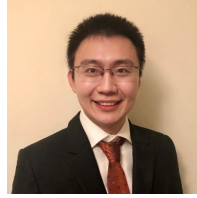
Xiongfei Zhai (M'19) received the B.S. and Ph.D degrees from Zhejiang University, Hangzhou, China, in 2009 and 2018, respectively. He was a visiting student with the Department of Electrical and Computer Engineering, Iowa State University, Ames, IA, USA. He is currently an assistant professor with the School of Information Engineering, Guangdong University of Technology, Guangzhou, China. His research interests include over-the-air computation, massive MIMO communications, millimeter-wave communications, and application of optimization

algorithm in communications.



Guojun Han (M'12-SM'14) received his Ph.D. from Sun Yatsen University, Guangzhou, China, and the M.E. degree from South China University of Technology, Guangzhou, China. From March 2011 to August 2013, he was a Research Fellow at the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. From October 2013 to April 2014, he was a Research Associate at the Department of Electrical and Electronic Engineering, Hong Kong University of Science and Technology. He is now a Full Professor

and Executive Dean at the School of Information Engineering, Guangdong University of Technology, Guangzhou, China. He has been a Senior Member of IEEE since 2014. His research interests are in the areas of wireless communications, signal processing, coding and information theory. He has more than 14 years' experience on research and development of advanced channel coding and signal processing algorithms and techniques for various data storage and communication systems.



Yuanwei Liu (S'13-M'16-SM'19, <http://www.eecs.qmul.ac.uk/~yuanwei>) received the PhD degree in electrical engineering from the Queen Mary University of London, U.K., in 2016. He was with the Department of Informatics, King's College London, from 2016 to 2017, where he was a Post-Doctoral Research Fellow. He has been a Senior Lecturer (Associate Professor) with the School of Electronic Engineering and Computer Science, Queen Mary University of London, since Aug. 2021, where he was a Lecturer (Assistant

Professor) from 2017 to 2021. His research interests include non-orthogonal multiple access, reconfigurable intelligent surface, integrated sensing and communications, and machine learning.

Yuanwei Liu is a Web of Science Highly Cited Researcher in 2021 and 2022, an IEEE Communications Society Distinguished Lecturer, an IEEE Vehicular Technology Society Distinguished Lecturer, the academic Chair for the Next Generation Multiple Access Emerging Technology Initiative, and the vice chair of the IEEE Technical Committee on Cognitive Networks (TCCN). He received IEEE ComSoc Outstanding Young Researcher Award for EMEA in 2020. He received the 2020 IEEE Signal Processing and Computing for Communications (SPCC) Technical Early Achievement Award, IEEE Communication Theory Technical Committee (CTTC) 2021 Early Achievement Award. He received IEEE ComSoc Outstanding Nominee for Best Young Professionals Award in 2021. He is the co-recipient of the Best Student Paper Award in IEEE VTC2022-Fall, the Best Paper Award in ISWCS 2022 and IEEE SPCC Best Paper Award. He serves as a Senior Editor of IEEE Communications Letters, an Editor of the IEEE Transactions on Wireless Communications and the IEEE Transactions on Communications. He serves as the Guest Editor for IEEE JSAC on Next Generation Multiple Access, IEEE JSTSP on Signal Processing Advances for Non-Orthogonal Multiple Access, IEEE Network on Next Generation Multiple Access for 6G. He serves as the Publicity Co-Chair for IEEE VTC 2019-Fall, Symposium Co-Chair for Cognitive Radio & AI-Enabled Networks for IEEE GLOBECOM 2022 and Communication Theory for IEEE GLOBECOM 2023. He serves as the chair of Special Interest Group (SIG) in SPCC Technical Committee on signal processing Techniques for next generation multiple access, the vice-chair of SIG WTC on Reconfigurable Intelligent Surfaces for Smart Radio Environments.



Yunlong Cai (S'07-M'10-SM'16) received the M.Sc. degree in electronic engineering from the University of Surrey, Guildford, U.K., in 2006, and the Ph.D. degree in electronic engineering from the University of York, York, U.K., in 2010. From 2010 to 2011, he was a Post-Doctoral Fellow with the Electronics and Communications Laboratory, Conservatoire National des Arts et Metiers, Paris, France. Since February 2011, he has been with the College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou, China,

where he is currently a Professor. From August 2016 to January 2017, he was a Visiting Scholar at the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, USA. His research interests include transceiver design for multiple-antenna systems, cooperative and relay communications, UAV communications, and machine learning for communications. He has published over 200 journal and conference papers in these areas. He regularly sits on the technical program committee (TPC) boards of prominent IEEE conferences such as ICC, GLOBECOM, and VTC. He is currently an Associate Editor of IEEE SIGNAL PROCESSING LETTERS.



Lajos Hanzo (<http://www-mobile.ecs.soton.ac.uk>, https://en.wikipedia.org/wiki/Lajos_Hanzo)

(FIEEE'04) received his Master degree and Doctorate in 1976 and 1983, respectively from the Technical University (TU) of Budapest. He was also awarded the Doctor of Sciences (DSc) degree by the University of Southampton (2004) and Honorary Doctorates by the TU of Budapest (2009) and by the University of Edinburgh (2015). He is a former Editor-in-Chief of the IEEE Press. He has served several terms as Governor of both IEEE

ComSoc and of VTS. He has published 1900+ contributions at IEEE Xplore, 19 Wiley-IEEE Press books and has helped the fast-track career of 123 PhD students. Over 40 of them are Professors at various stages of their careers in academia and many of them are leading scientists in the wireless industry. He is also a Fellow of the Royal Academy of Engineering, FIEEE, FIET, Fellow of EURASIP and a Foreign Member of the Hungarian Academy of Sciences. He was bestowed upon the IEEE Eric Sumner Technical Field Award.