Miso: Misalignment Allowed Optimization for Multi-antenna Over-the-air Computation

Suhua Tang, *Senior Member, IEEE,* Chao Zhang, Jie Li, *Senior Member, IEEE,* Sadao Obana

***Abstract*—Over-the-air computation (AirComp), as an effective method to wireless data aggregation, has attracted much attention recently. It helps to improve network efficiency and scalability by integrating communication and computation in the air. In AirComp, both signal magnitude misalignment and noise lead to computation error. In the single antenna case, by allowing misalignment in signal magnitude, a good tradeoff can be achieved between signal distortion and noise power, which leads to a minimal error. In the multi-antenna case, usually the zero-forcing policy is used to enforce signal magnitude alignment (no distortion), which, however, increases noise and affects the overall computation error. To better exploit multiple antennas at the sink, in this paper, we propose a misalignment allowed optimization (Miso) method for AirComp. Specifically, a group of nodes whose signals may be misaligned are dynamically selected, and other signals are aligned to a higher level with higher quality. On this basis, the optimization of multi-antenna AirComp is converted to a difference of convex problem and is solved iteratively. Simulations confirm that the proposed method greatly reduces computation error and scales better with the number of nodes, compared with previous methods.**

***Index Terms*—Wireless data aggregation, Over-the-air computation, Multi-antenna, Difference of convex**

I. Introduction

Billions of sensors will be deployed and connected to the Internet by low power wide area (LPWA) technologies such as NB-IoT and LoRa [1] in the forthcoming smart society, where wireless data aggregation (WDA) will play an important role. In IoT networks, a sink node (access point or base station) collects data from its associated sensor nodes, and either processes the data locally or sends the data to the cloud. This, however, faces the scalability issue because separately collecting data from each sensor node is inefficient in wireless resource.

Over-the-air computation (AirComp), as an effective method to WDA, has attracted much attention recently. It helps to improve network efficiency and scalability by integrating communication and computation in the air. It was first put forward in [2], which integrates the data collection and fusion in the digital domain using structured codes. Later, AirComp in the analog domain using uncoded transmissions were studied in [3], [4]. All nodes, synchronized with the sink, transmit their signals simultaneously, which are added together in the air by exploiting the superposition property of wireless communication. Although analog communication seems to be error prone, actually it can achieve much smaller computation error than its digital counterpart when using the same amount of resources [5].

In its simplest form, AirComp only supports the sum operation, which corresponds to the signal superposition. It can be extended to support any kind of nomographic functions, such as geometric mean, weighted average, variance, *et al.* [6], [7], [8] by proper pre-processing and post-processing operations. Besides well-defined functions, exploiting deep neural network in the pre-processing and post-processing helps to approximate unknown functions via deep models learned from real data [9]. Except for sensor networks, AirComp has other applications, including distributed consensus [10], distributed spectrum sensing [11], and federated learning [12], [13], [14], [15].

To ensure unbiased estimations via AirComp, it is expected that all signals arrive at the sink, aligned in signal magnitude. This is usually achieved by pre-equalizing the channel via transmission power control. However, it cannot well deal with nodes in deep fading. In the single antenna case, by allowing misalignment in signal magnitude, a good tradeoff is achieved between signal distortion and noise, which leads to a minimal computation error [16], [17]. In the multi-antenna case, zero-forcing is still used to ensure signal magnitude alignment, with a sub-optimal performance.

In this paper, we propose a misalignment allowed optimization (Miso) method for the general AirComp in the multi-antenna case, to reduce the overall computation error of signals from all nodes. To the best of our knowledge, this is the first work in this field. Instead of making all signals aligned in magnitude, a group of nodes whose signal magnitude may be misaligned are dynamically selected, which enables to align signals from other nodes to a larger signal magnitude and improves signal quality at the sink. Thus, the computation mean squared error (MSE), taking into account both signal

Manuscript received xxxxxxxx, 2022; revised xxxxxxxx, 2022; accepted xxxxxxxx, 2022. Date of publication xxxxxxxx, 2022; date of current version xxxxxxxx, 2022. This work was supported in part by the SCAT Foundation, Japan. This work of Chao Zhang was supported by the Key Research and Development Program of Shaanxi under Grant 2023-YBGY-251. This work of Jie Li was supported by the National Key R&D Program of China No. 2020YFB1710900 and No. 2020YFB1806700, NSFC Grants 61932014 and 62232011.

The associate editor coordinating the review of this paper and approving it for publication was xxxxxxxx. (Corresponding author: Suhua Tang.)

Digital Object Identifier xxxxxxxxxxxxxxxxxxxx

S. Tang is with Graduate School of Informatics and Engineering, The University of Electro-Communications, Japan. (e-mail: shtang@uec.ac.jp).

C. Zhang is with School of Information and Communications Engineering, Xi'an Jiaotong University, China.

J. Li is with Department of Computer Science and Engineering, MoE Key Lab of Artificial Intelligence, AI Institute, Shanghai Jiao Tong University, Yancheng Blockchain Research Institute, China.

S. Obana is with The University of Electro-Communications, Japan.

misalignment and noise, is minimized. The main contribution of this paper is three-fold, as follows

- By explicitly selecting a set of nodes whose signal magnitude may be misaligned, the problem of minimizing computation MSE is reformulated, and approximated as a difference of convex problem.
- A low complexity, iterative algorithm is proposed for solving the problem, by using the matrix lifting technique.
- The set of nodes whose signal magnitude may be misaligned is dynamically updated based on channel state information and is jointly optimized with the antenna control.

Numerical analysis and Monte Carlo simulations confirm that the proposed method greatly reduces the computation MSE compared with state-of-the-art methods.

The rest of this paper is organized as follows: Sec. II reviews related work on improving the reliability of AirComp. Sec. III explains the system model of AirComp with multiple antennas at the sink. Then, Sec. IV proposes the misalignment allowed optimization method and its iterative solution. Sec. V presents the results achieved by Monte Carlo simulation, illustrating the impact of the number of antennas at the sink and the number of nodes in the network. Finally, Sec. VI concludes this paper.

II. Related Work

Noise in the wireless channels introduces computation error in AirComp and channel fading affects the alignment of signal magnitude. Here, we review previous efforts on dealing with the channel fading and noise in the wireless channels for AirComp and improving the reliability of AirComp.

A. Power Control

Power control is a straightforward method to reduce the computation MSE, using a large transmission power when the channel condition is poor, so that after pre-equalization, the equivalent channel gain is the same for each signal. Because of signal magnitude alignment, the transmission power on each link also depends on the status of other links. In the single-antenna scenario, this typically adopts the channel inversion policy [16], [17] unless the maximal transmission power is reached.

Reducing computation error usually is achieved at the cost of increased transmission power. Because it is difficult to replace the battery of sensor nodes deployed in wild environments, it is necessary to suppress the transmission power. The sum power constraint is considered in [18]. Wireless power transfer is also studied for AirComp [19], [20]. In the downlink, the sink delivers energy to nodes, and later in the uplink, nodes with enough power transmit their signals to the sink.

B. Time/Frequency/Path Diversity

An effective method to channel fading is diversity. Time diversity is exploited in [21], by extending AirComp to multiple slots. Each node can delay its transmission until its channel gain gets above a pre-defined threshold. The selection of transmission slot effectively improves channel gain, and the extra noise due to multiple slots is considered in the optimization.

A counterpart in frequency domain is suggested in [22]. With M channels available, each node selects a subset of the channels and broadcasts its signal over these channels under a certain power constraint, which helps to mitigate the impact of channel fading.

Path diversity, via amplify-and-forward based relay, is exploited in [23], using one relay to help multiple nodes. As a result, relay transmission power increases with the number of nodes using the relay. By controlling the number of nodes exploiting the relay, and coherently combining direct signals and forwarded signals, relay transmission power is suppressed while computation error is reduced. In addition, the tradeoff between computation error and transmission power is studied. Considering the constraint of relay transmission power, how to schedule nodes to use the relay is further studied in [24].

IRS (intelligent reflecting surface)-aided AirComp introduces controllable reflection paths via which signal phase is purposely adjusted [20], [25], [26]. In this way, reflected signals via multiple reflecting elements, after phase adjustment, can be coherently combined at the sink. Because signals are not amplified at the reflecting surface, there is no extra power consumption. On the other hand, the reflected signal usually is very weak, and it requires many reflecting elements to make the overall signal strength large enough.

C. Multiple Antennas

Using multiple antennas at the sink is another effective method to improve the performance of AirComp. But it also increases the technical difficulty. With multiple antennas at the sink, the computation MSE is not tractable any more, and it is difficult to consider signal misalignment. Therefore, usually zero-forcing is adopted to enforce magnitude alignment of all signals [19], [27], [28], [29], as a counterpart of channel inversion in the single antenna case. This removes the signal distortion, but increases noise power. As a result, its reduction of computation MSE is limited, especially when a link in deep fading affects the performance of the whole system.

The coverage of a single sink is limited. For a large-scale deployment, AirComp is extended to multiple cells in [30], [31]. This is somewhat similar to use a distributed antenna array, but potential interference needs to be taken into account.

D. CSI Detection

Accurate CSI (channel state information) usually is a necessity of AirComp. In a fast fading environment, however, the CSI detection itself becomes a challenging problem. Using the statistics of CSI instead of their instantaneous values and letting each node decide its transmission based on its own CSI, the distributed scheduling in [21] avoids the feedback of instantaneous CSI from each node to the sink.

In the application of AirComp to mobile environment, e.g., UAV (unmanned aerial vehicle) based data collection, high mobility makes it impractical to detect CSI per node and feed back to the sink. Alternatively, AirComp is exploited in

the CSI estimation. The whole process involves two stages. In the first stage, AirComp is used to estimate CSI, and in the second stage, it is used to perform the actual data collection/computation [32], [33]. Recently, researchers also tried to exploit blind AirComp without requiring CSI access [34].

E. Spatial Correlation

Usually, it is assumed that signals from all sensors are independent without correlation, which facilitates the analysis. But actually signals about the same event captured by adjacent sensors may be correlated. Researchers start to study signal correlation in [35], [36], but the analysis becomes more complex and it requires to know the channel correlation information. By exploiting the spatial correlation, it is possible to use signals from only part of nodes to get the correct sensing result, via spatial sampling [37]. This helps to reduce energy consumption in densely deployed networks.

F. Motivation of This Work

Computation error greatly affects the performance of AirComp. It is known that aligning the magnitude of all signals is not optimal, and misalignment in signal magnitude needs to be considered to minimize the overall computation MSE, allowing a small error in signal magnitude to largely reduce noise power. This is already achieved in the single antenna case [16], [17]. In some applications of AirComp, e.g., federated learning, signal misalignment is considered explicitly in the application itself [38].

Taking signal misalignment into account in the optimization remains a challenge for multi-antenna AirComp. Previous work [19], [27], [28], [29] still requires that all signals be aligned in magnitude. As a result, the weakest link is the bottleneck and the reduction of computation error is limited [16]. In general IoT tasks, usually it is required to collect data from all nodes, and it is expected to exploit multiple antennas at the sink instead of selecting one antenna [16]. To this end, this paper reduces the computation MSE, using multiple antennas at the sink simultaneously and considering signal magnitude misalignment in the optimization. On the one hand, the proposed method can be regarded as an extension of previous work [27], to better reduce the overall MSE by allowing signal distortion in the multi-antenna setting. On the other hand, it can be regarded as an extension of the work in [16], from single antenna to multiple antennas, to better exploit antennas at the sink.

This problem is not directly solvable. Therefore, we approximate the problem to a tractable one, and solve it by iteration. In the approximation, we need to know which signals are allowed to be distorted. Then, we minimize the overall MSE, on the one hand, by adjusting the Rx-scaling parameter, and on the other hand, by adjusting the set of signals that can be distorted. Because matrix lifting is used, we cannot directly get the optimal solution, and need to take Gaussian randomization to probe. In the optimization, it is ensured that the signals that cannot be distorted will be aligned, but for signals that can be distorted, in the probing stage, there is a chance that some of these signals will be aligned, which helps to reduce the overall MSE.

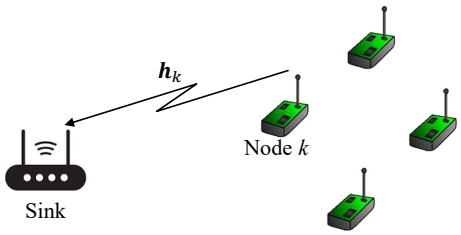


Fig. 1. Over-the-air computation for data collection/processing from K nodes with a single antenna to a sink with N antennas.

III. System Model

This paper considers a sensor network with K nodes and a sink, as shown in Fig. 1. Each node has a single antenna while the sink has N antennas. The sink tries to compute the sum of signals received from all nodes. The sink broadcasts a beacon signal to all nodes to initiate an over-the-air computation. Under this trigger, all nodes transmit their signals in the analog wave simultaneously. These signals are superimposed at the sink, generating the sum. This can be easily extended to support other nomographic functions by proper pre-processing and post-processing [7], [8], [9].

In the following analysis, $\mathbb{C}$ denotes the set of complex numbers, and $\mathbb{C}^N$ represents the set of column vectors with N complex numbers. Bold font is used for a vector or a matrix. $||\boldsymbol{x}||$ is the norm of a vector $\boldsymbol{x}$. $\boldsymbol{x}^H$ denotes the conjugate transpose of a vector or a matrix $\boldsymbol{x}$. For a matrix $\boldsymbol{x}$, $\mathrm{Tr}(\boldsymbol{x})$ and $\mathrm{rank}(\boldsymbol{x})$ represent its trace and rank, respectively. $\mathbb{E}\{\cdot\}$ is an operator of expectation. $|\mathbb{S}|$ is the number of elements in the set $\mathbb{S}$.

A. Basic Model

To deal with the difference in channel gains caused by different propagation distances and channel fading, the pre-processed signal at the k_{th} node, $x_k \in \mathbb{C}$, with zero mean and unit variance ($\mathbb{E}(|x_k|^2) = 1, |x_k| \leq v$), is amplified by its Tx-scaling factor $b_k \in \mathbb{C}$ and sent to the sink. All the transmissions are synchronized so that all signals arrive at the sink at the same time. It is assumed that all signals are non-correlated. Then, the received signal at the sink is

$$\boldsymbol{y} = \sum_{k=1}^{K} \boldsymbol{h}_k b_k x_k + \boldsymbol{n}, \tag{1}$$

where $\boldsymbol{h}_k \in \mathbb{C}^N$ is the channel coefficient between node k and the sink, and $\boldsymbol{n} \in \mathbb{C}^N$, $\boldsymbol{n} \sim \mathcal{CN}(0, \sigma^2 \boldsymbol{I})$, is the additive white Gaussian noise (AWGN). The channel in the sensor network is relatively stationary, so the sink can measure the channel coefficient h_k to each node k at a long interval by using a pilot signal, and notify node k of the Tx-scaling factor b_k.

The sink applies a Rx-scaling vector $\boldsymbol{a} \in \mathbb{C}^N$ to the received signal to get the computation result as

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3292882

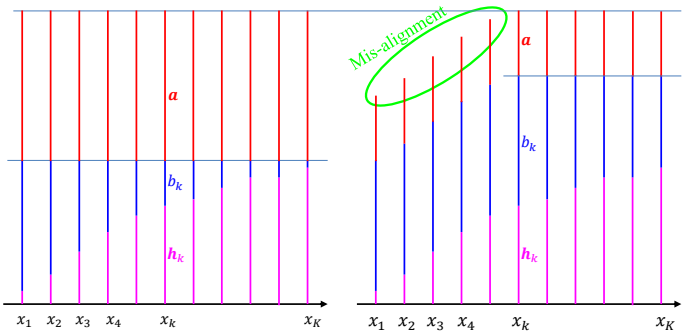


Fig. 2. Two policies of signal magnitude alignment in over-the-air computation. (a) Signal magnitude aligned to the weakest signal. (b) Signal magnitude aligned to a signal with a proper magnitude. Nodes are sorted in the ascending order of channel gains ($||\boldsymbol{h}_k||$, pink color), and the insufficient part is complemented by transmission power control (b_k, blue color). $\boldsymbol{a}$ (red color) helps to make the result reach the target value, but does not improve the signal to noise ratio.

$$r = \boldsymbol{a}^H \boldsymbol{y} = \sum_{k=1}^{K} \boldsymbol{a}^H \boldsymbol{h}_k b_k x_k + \boldsymbol{a}^H \boldsymbol{n}. \tag{2}$$

The computation MSE is measured as

$$\begin{aligned} \text{MSE} &= \mathbb{E}\left\{ \left| r - \sum_{k=1}^{K} x_k \right|^2 \right\} \\ &= \sum_{k=1}^{K} |\boldsymbol{a}^H \boldsymbol{h}_k b_k - 1|^2 + \sigma^2 ||\boldsymbol{a}||^2. \end{aligned} \tag{3}$$

$|b_k x_k|^2$ should be no more than P', the maximal transmission power. Let $P_{\max}$ denote P'/v^2. Then, $|b_k|^2 \le P'/v^2 \triangleq P_{\max}$, and parameters b_k and $\boldsymbol{a}$ are obtained by minimizing MSE as follows:

$$\mathcal{P}_1 \min_{\boldsymbol{a}, b_k} \text{MSE}, \tag{4a}$$

$$\text{s.t. } |b_k|^2 \le P_{\max}, k = 1, \cdots, K. \tag{4b}$$

This sum-of-squares of MSE cannot be directly optimized.

B. Different Optimization Policies

Different policies are used in the literature to minimize the computation MSE of AirComp.

In the 1st policy (Fig. 2(a)), it is required that all signals arrive at the sink aligned in signal magnitude [14], [25], [27]. In other words, $\boldsymbol{a}^H \boldsymbol{h}_k b_k - 1 = 0$ for $k = 1, \cdots, K$. Therefore, it is also called zero-forcing. When there is a node with very small $||\boldsymbol{h}_k||$, this requires to use a large $||\boldsymbol{a}||$. Although this removes the misalignment in signal magnitude, it leads to a large increase in noise power $\sigma^2||\boldsymbol{a}||^2$, which affects the overall computation MSE.

In the 2nd policy (Fig. 2(b)), misalignment in signal magnitude is allowed for some weak signals [16], [17]. In this way, other good signals are aligned to a larger magnitude, and a smaller $||\boldsymbol{a}||$ is used, which leads to a smaller noise power. Finding the optimal $\boldsymbol{a}$ requires a tradeoff between signal misalignment and noise power. The optimal solution is given for the single antenna case ($N = 1$) [16], [17]. But for the multi-antenna case, there is no complete solution, and the selection of a single antenna is suggested instead [16]. Although it outperforms the 1st policy in some cases, the performance is still limited.

This paper extends the 2nd policy to the multi-antenna case, using all antennas simultaneously instead of antenna selection, and considers the misalignment of signal magnitude in the optimization, so as to fully exploit multiple antennas at the sink.

IV. Misalignment-allowed Optimization for Multi-antenna AirComp

Here, we present the proposed method, misalignment-allowed optimization (Miso) for AirComp in the multi-antenna scenario, and analyze how to approximate the problem to a tractable one. Then, on this basis, we present an iterative solution.

A. Specifying a Set of Misaligned Signals

In the proposed model, misalignment in signal magnitude is allowed for some nodes. To facilitate the analysis, nodes are explicitly divided into two groups. $\mathbb{S}$ is used to denote the group of nodes whose signal magnitudes are allowed to be misaligned.

In the single antenna case, nodes are sorted in the ascending order of channel gain, and a critical number, $i^\star$, is defined in [16]. Then, the signal from a node whose index is below $i^\star$ is misaligned in signal magnitude even when the maximal transmission power is used.

In the proposed model, channel coefficient $\boldsymbol{h}_k$ is a complex vector, and signal magnitude further depends on $\boldsymbol{a}$ and b_k. To avoid the impact of transmission power, $\boldsymbol{a}^H \boldsymbol{h}_k$ is used as a criterion to divide nodes into two groups ($k \in \mathbb{S}$ and $k \notin \mathbb{S}$). Here, nodes are sorted in the ascending order of $|\boldsymbol{a}^H \boldsymbol{h}_k|$, and top i nodes are selected to form the set $\mathbb{S}_i$. By further changing the value of i, the minimal MSE can be found accordingly.

$$\min_{\mathbb{S}_i} \text{MSE}(\mathbb{S}_i). \tag{5}$$

B. Considering Signal Misalignment

A transmission power control policy similar to channel inversion is used, and transmission power at node k is computed as

$$b_k = \begin{cases} \frac{1}{\boldsymbol{a}^H \boldsymbol{h}_k} & \frac{1}{|\boldsymbol{a}^H \boldsymbol{h}_k|^2} \le P_{\max}, \\ \frac{\sqrt{P_{\max}} (\boldsymbol{a}^H \boldsymbol{h}_k)^H}{|\boldsymbol{a}^H \boldsymbol{h}_k|} & \frac{1}{|\boldsymbol{a}^H \boldsymbol{h}_k|^2} > P_{\max}. \end{cases} \tag{6}$$

For a signal from a node $k \notin \mathbb{S}_i$ to arrive at the sink aligned in signal magnitude, it is expected that $b_k = 1/(\boldsymbol{a}^H \boldsymbol{h}_k)$ will be used and satisfies $\boldsymbol{a}^H \boldsymbol{h}_k b_k - 1 = 0$ under the constraint $|b_k|^2 = 1/|\boldsymbol{a}^H \boldsymbol{h}_k|^2 \le P_{\max}$. On the other hand, for a signal from a node $k \in \mathbb{S}_i$, even with the maximal transmission power and proper phase control ($b_k = \sqrt{P_{\max}} (\boldsymbol{a}^H \boldsymbol{h}_k)^H / |\boldsymbol{a}^H \boldsymbol{h}_k|$),

$\boldsymbol{a}^H\boldsymbol{h}_k b_k$ will be less than 1, which leads to misalignment in signal magnitude.

Because b_k computed by (6) ensures $|1-\boldsymbol{a}^H\boldsymbol{h}_k b_k|^2=0$ for $k\notin\mathbb{S}_i$, minimizing MSE in (3) becomes

$$\mathcal{P}_2 \min_{\boldsymbol{a},b_k,\mathbb{S}_i} \sum_{k\in\mathbb{S}_i} \left|1-\boldsymbol{a}^H\boldsymbol{h}_k b_k\right|^2+\sigma^2||\boldsymbol{a}||^2, \tag{7a}$$

$$\text{s.t. } |\boldsymbol{a}^H\boldsymbol{h}_k|^2<\frac{1}{P_{\max}},\ k\in\mathbb{S}_i, \tag{7b}$$

$$|\boldsymbol{a}^H\boldsymbol{h}_k|^2\geq\frac{1}{P_{\max}},\ k\notin\mathbb{S}_i, \tag{7c}$$

$$|b_k|^2\leq P_{\max}. \tag{7d}$$

C. Problem Approximation

When $k\in\mathbb{S}_i$, $\boldsymbol{a}^H\boldsymbol{h}_k b_k$ equals to $|\boldsymbol{a}^H\boldsymbol{h}_k|\sqrt{P_{\max}}$ by using the maximal transmission power with a proper phase in (6) to reduce signal misalignment, and is less than 1. Then, MSE in (7a) becomes

$$\sum_{k\in\mathbb{S}_i}\left(1-|\boldsymbol{a}^H\boldsymbol{h}_k\sqrt{P_{\max}}|\right)^2+\sigma^2||\boldsymbol{a}||^2. \tag{8}$$

Under the constraint in (7b,7c), the number of variables is reduced, because now b_k is removed. But this is still not directly solvable.

Note that $|\boldsymbol{a}^H\boldsymbol{h}_k\sqrt{P_{\max}}|$ is less than 1, but approaches 1 to reduce the MSE error. Then, $|\boldsymbol{a}^H\boldsymbol{h}_k\sqrt{P_{\max}}|$ can be approximated by $|\boldsymbol{a}^H\boldsymbol{h}_k\sqrt{P_{\max}}|^2$, and

$$\begin{aligned}\left(1-|\boldsymbol{a}^H\boldsymbol{h}_k\sqrt{P_{\max}}|\right)^2 &\approx 1-2|\boldsymbol{a}^H\boldsymbol{h}_k\sqrt{P_{\max}}|^2\\ &\quad+|\boldsymbol{a}^H\boldsymbol{h}_k\sqrt{P_{\max}}|^2\\ &= 1-P_{\max}|\boldsymbol{a}^H\boldsymbol{h}_k|^2.\end{aligned} \tag{9}$$

Therefore, the problem $\mathcal{P}_2$ can be approximated as

$$\mathcal{P}_3 \min_{\boldsymbol{a},\mathbb{S}_i} \sigma^2||\boldsymbol{a}||^2+|\mathbb{S}_i|-P_{\max}\sum_{k\in\mathbb{S}_i}|\boldsymbol{a}^H\boldsymbol{h}_k|^2, \tag{10a}$$

$$\text{s.t. } (7b),(7c).$$

With more than 2 antennas, there is no simple optimal solution. At a fixed $\mathbb{S}_i$, $\boldsymbol{a}$ can be computed by minimizing the approximate metric. When jointly optimizing both $\boldsymbol{a}$ and $\mathbb{S}_i$ in (10a), it tends to select a small $|\mathbb{S}_i|$. Therefore, in the late computation, the approximation metric is used for computing $\boldsymbol{a}$, while the actual MSE is used to find $\mathbb{S}_i$.

D. Solving the problem by the DC approach

(10a) is a nonconvex QCQP (quadratically constrained quadratic program) problem [39], or more strictly, a difference of convex problem, which can be solved by using the semidefinite relaxation (SDR) technique [40].

Define $\boldsymbol{A}\triangleq\boldsymbol{a}\boldsymbol{a}^{\boldsymbol{H}}$ as a rank one semidefinite Hemitian matrix and $\boldsymbol{H}_k\triangleq\boldsymbol{h}_k\boldsymbol{h}_k^H$. Then, $||\boldsymbol{a}||^2=\boldsymbol{a}^H\boldsymbol{a}=\mathrm{Tr}(\boldsymbol{a}\boldsymbol{a}^{\boldsymbol{H}})=\mathrm{Tr}(\boldsymbol{A})$, and $|\boldsymbol{a}^H\boldsymbol{h}_k|^2=\boldsymbol{h}_k^H\boldsymbol{a}\boldsymbol{a}^H\boldsymbol{h}_k=\mathrm{Tr}(\boldsymbol{a}\boldsymbol{a}^H\boldsymbol{h}_k\boldsymbol{h}_k^H)=\mathrm{Tr}(\boldsymbol{A}\boldsymbol{H}_k)$.

Then, the difference of convex problem in (10a) is reformulated as a rank-one matrix optimization problem.

$$\mathcal{P}_4 \min_{\boldsymbol{A},\mathbb{S}_i} \sigma^2\mathrm{Tr}(\boldsymbol{A})+|\mathbb{S}_i|-P_{\max}\sum_{k\in\mathbb{S}_i}\mathrm{Tr}(\boldsymbol{A}\boldsymbol{H}_k), \tag{11a}$$

$$\text{s.t. } \mathrm{Tr}(\boldsymbol{A}\boldsymbol{H}_k)<\frac{1}{P_{\max}},\ k\in\mathbb{S}_i, \tag{11b}$$

$$\mathrm{Tr}(\boldsymbol{A}\boldsymbol{H}_k)\geq\frac{1}{P_{\max}},\ k\notin\mathbb{S}_i, \tag{11c}$$

$$\boldsymbol{A}\succeq 0,\mathrm{rank}(\boldsymbol{A})=1. \tag{11d}$$

Generally it is difficult to satisfy the constraint $\mathrm{rank}(\boldsymbol{A})=1$. It is proven in [14] that for a semidefinite matrix $\boldsymbol{A}$ with $\mathrm{Tr}(\boldsymbol{A})\geq 1$, $\mathrm{rank}(\boldsymbol{A})=1$ is equivalent to $\mathrm{Tr}(\boldsymbol{A})=||\boldsymbol{A}||_2$. Here, $\mathrm{Tr}(\boldsymbol{A})=\sum_i\sigma_i(\boldsymbol{A})$ and the spectral norm $||\boldsymbol{A}||_2=\sigma_1(\boldsymbol{A})$, where $\sigma_i(\boldsymbol{A})$ denotes the i^{th} largest singular value of matrix $\boldsymbol{A}$. $\mathrm{Tr}(\boldsymbol{A})-\sigma_1(\boldsymbol{A})\geq 0$ and the equality holds only if $\mathrm{Tr}(\boldsymbol{A})-||\boldsymbol{A}||_2=0$. Then, the problem becomes

$$\begin{aligned}\mathcal{P}_5 \min_{\boldsymbol{A},\mathbb{S}_i} &\sigma^2\mathrm{Tr}(\boldsymbol{A})+\rho\cdot(\mathrm{Tr}(\boldsymbol{A})-||\boldsymbol{A}||_2)\\ &+|\mathbb{S}_i|-P_{\max}\sum_{k\in\mathbb{S}_i}\mathrm{Tr}(\boldsymbol{A}\boldsymbol{H}_k),\end{aligned} \tag{12a}$$

$$\text{s.t. } \boldsymbol{A}\succeq 0, \tag{12b}$$

$$(11b),(11c),$$

where ρ is a positive parameter. This problem can be solved in an iterative way by exploiting CVX, a Matlab package for specifying and solving convex programs [41], [42]. At the t^{th} step, $\mathrm{Tr}(\boldsymbol{A})-||\boldsymbol{A}||_2$ is equivalent to $<\boldsymbol{A},I-\partial||\boldsymbol{A}^{t-1}||_2>$, an inner product between $\boldsymbol{A}$ and $I-\partial||\boldsymbol{A}^{t-1}||_2$, and $\partial||\boldsymbol{A}^{t-1}||_2$ is computed as $\boldsymbol{v}_1\boldsymbol{v}_1^H$ where $\boldsymbol{v}_1\in\mathbb{C}^N$ is the eigenvector of $\boldsymbol{A}^{t-1}$ corresponding to the singular value $\sigma_1(\boldsymbol{A}^{t-1})$.

The iterative solution of $\boldsymbol{A}$ cannot guarantee $\mathrm{Tr}(\boldsymbol{A})-||\boldsymbol{A}||_2=0$. If $\boldsymbol{A}$ is not rank one, usually the Gaussian randomization [40] will be used to find the optimal $\boldsymbol{a}$. With $\boldsymbol{a}$ solved, b_k will be computed by (6).

In the proposed method, actually Gaussian randomization is always adopted, for the following two reasons. (i) Two constraints (7b-7c) are used for the optimization via CVX, and this is an approximate solution. By Gaussian randomization, we expect to relax the constraint (remove the constraint (7b)) and directly minimize the real MSE. (ii) The iterative optimization depends on the initialization, and Gaussian randomization helps to alleviate this impact.

E. Alternate optimization

The optimization of $\boldsymbol{a}$ depends on the setting of $\mathbb{S}_i$, which changes with $\boldsymbol{a}$. Therefore, this is solved by alternate optimization of $\boldsymbol{a}$ and $\mathbb{S}_i$.

The whole procedure is shown in Algorithm 1, which consists of two parts. The procedure CompOptParam iterates over all possible $|\mathbb{S}_i|$. The i^{th} step is realized by the procedure FindOptA, which finds the optimal $\boldsymbol{a}^i$ given a specific $|\mathbb{S}_i|$. In FindOptA, $\hat{\boldsymbol{a}}^t$ and $\boldsymbol{a}$ represent Rx-scaling factors before and after Gaussian randomization, respectively. $\hat{\boldsymbol{a}}^0$ is initialized by random values (line 10). In the iteration, $\mathbb{S}_i$ is updated based

Algorithm 1 Find parameters for multi-antenna AirComp.

1: **procedure** CompOptParam($\{\boldsymbol{h}_k\}$)
2: **for** $i = 0, 1, 2, \cdots, K-1$ **do**
3: $|\mathbb{S}_i| \leftarrow i$ ▷ Specify the size of $\mathbb{S}_i$
4: $(\boldsymbol{a}^i, \text{MSE}_i)$ = FindOptA($\{\boldsymbol{h}_k\}$, $|\mathbb{S}_i|$)
5: **end for**
6: Find optimal solution, $\boldsymbol{a} = \underset{\boldsymbol{a}^i}{\text{argmin}}\ \text{MSE}_i$
7: return $\boldsymbol{a}$
8: **end procedure**
9: **procedure** FindOptA($\{\boldsymbol{h}_k\}$, $|\mathbb{S}_i|$)
10: Initialize $\hat{\boldsymbol{a}}^0$, $\boldsymbol{A}^0 = \hat{\boldsymbol{a}}^0(\hat{\boldsymbol{a}}^0)^H$, $\text{MSE}^0 \leftarrow \infty$
11: Compute $\boldsymbol{H}_k = \boldsymbol{h}_k\boldsymbol{h}_k^H$ for each node k
12: **for** $t = 1, 2, \cdots$ **do**
13: Sort nodes in the ascending order of $|(\hat{\boldsymbol{a}}^{t-1})^H\boldsymbol{h}_k|$
14: Select top $|\mathbb{S}_i|$ nodes as $\mathbb{S}_i$
15: Compute subgradient $\partial||\boldsymbol{A}^{t-1}||_2$
16: Compute $\boldsymbol{A}^t$ and MSE^t by (12a) using CVX
17: **if** $|\text{MSE}^t - \text{MSE}^{t-1}|/|\text{MSE}^t| \le \varepsilon_{\text{th}}$ **then**
18: break ▷ Stop iteration
19: **end if**
20: Decompose $\boldsymbol{A}^t$ as $\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{U}^H$
21: Compute $\hat{\boldsymbol{a}}^t = \boldsymbol{U}_1\sigma_1(\boldsymbol{\Sigma}^{1/2})$ ▷ Update $\hat{\boldsymbol{a}}^t$
22: **end for**
▷ Gaussian randomization
23: $\boldsymbol{a}_j = \boldsymbol{U}\boldsymbol{\Sigma}^{\mathbf{1/2}}\boldsymbol{n}_j$, $\boldsymbol{n}_j \in \mathcal{CN}(0, \boldsymbol{I})$, $j = 1, ..., L$
24: Normalize $\boldsymbol{a}_j = \frac{\boldsymbol{a}_j}{\min\limits_{k\notin\mathbb{S}_i}|\boldsymbol{a}_j^H\boldsymbol{h}_k|}\frac{1}{\sqrt{P_{\max}}}$ ▷ s.t. (7c)
25: **for** $j = 1, 2, \cdots, L$ **do**
26: Compute b_k by (6), $k = 1, \cdots, K$
27: Compute $\text{MSE}_j(\boldsymbol{a}_j)$ in (3) by using $\boldsymbol{a}_j$ and b_k
28: **end for**
29: Find optimal parameter, $\boldsymbol{a} = \underset{\boldsymbol{a}_j}{\text{argmin}}\ \text{MSE}_j(\boldsymbol{a}_j)$
30: return $\boldsymbol{a}$ and $\text{MSE}(\boldsymbol{a})$
31: **end procedure**

on $|(\hat{\boldsymbol{a}}^{t-1})^H\boldsymbol{h}_k|$ (line 13-14). Then, with $\mathbb{S}_i$, the optimal $\hat{\boldsymbol{a}}^t$ is found by the solution of $\boldsymbol{A}^t$ (line 15-16) via CVX and the decomposition of $\boldsymbol{A}^t = \hat{\boldsymbol{a}}^t(\hat{\boldsymbol{a}}^t)^H$ (line 20-21). After the convergence, Gaussian randomization (line 23-29) is performed. Specifically, candidate $\boldsymbol{a}_j$ is normalized to meet the constraint in (7c) (line 24), and the actual MSE (line 27) is computed, based on which the optimal solution $\boldsymbol{a}$ is found (line 29).

Finding the optimal parameter $\boldsymbol{a}$ depends on properly setting $\mathbb{S}$, a set of misaligned signals. It is non-trivial to decide $\mathbb{S}$. Therefore, two loops are used in Algorithm 1, finding an optimal parameter $\boldsymbol{a}^i$ for each $|\mathbb{S}_i|$. In this way, the computation time increases with the number of nodes.

1) Numerical results: With the same setting as in the first paragraph of Sec. V, we investigate the impact of the number of misaligned signals, using the network shown in Fig. 9.

Fig. 3 shows the computation MSE with respect to the value of $|\mathbb{S}_i|$, under different numbers of antennas at the sink. Generally, more antennas lead to smaller MSE. With the number of antennas fixed, MSE does not monotonically increase or decrease with $|\mathbb{S}_i|$. Too large or too small $|\mathbb{S}_i|$ leads to large MSE, and the exhaustive search is necessary to find the optimal value of $|\mathbb{S}_i|$.

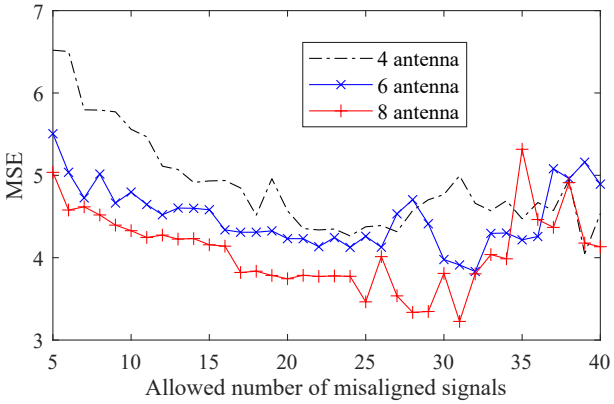


Fig. 3. Variation of MSE under different $|\mathbb{S}_i|$ values (#node = 50, $\sigma^2 = 1$).

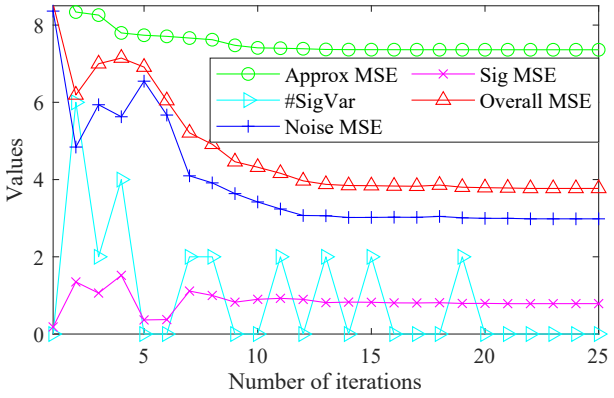


Fig. 4. Convergence of approximate MSE in (12a), actual overall MSE in (3), and $\mathbb{S}_i$ in the iteration (#node = 50, #antenna = 8, $\sigma^2 = 1$).

Fig. 4 shows an iterative process corresponding to $|\mathbb{S}_i| = 5$ (the number of antennas is 8) in Fig. 3. The variations in Sig MSE (corresponding to the distortion due to misalignment in signal magnitude) and Noise MSE (noise power) reflect the tradeoff between them. As the iteration goes, the approximate MSE in (12a) converges, so does the overall MSE in (3), but the latter shows larger variation before the convergence. The set of misaligned signals changes, and #SigVar shows the number of different signals in $\mathbb{S}_i$ between adjacent iterations. When MSE converges, the set of misaligned signals also gets stable, and the difference becomes 0.

As mentioned before, $\mathbb{S}_i$ represents a set of nodes whose signal may be misaligned, but in the Gaussian randomization, this constraint is relaxed. Therefore, the actual number of misaligned signals is less than $|\mathbb{S}_i|$, especially when $|\mathbb{S}_i|$ is large, as shown in Fig. 5.

F. Simplified iteration

To reduce the computation cost, here we exploit a simplified iteration, starting with an empty $\mathbb{S}$. This corresponds to the initial solution of the zero-forcing based method, where all signals are aligned to the same magnitude. In Algorithm 1, with a fixed $|\mathbb{S}|$, $\boldsymbol{a}$ and $\mathbb{S}$ are alternatively computed until the process converges, which is time consuming. In comparison, here, for each $|\mathbb{S}|$, $\mathbb{S}$ is decided by $\boldsymbol{a}$, and then $\boldsymbol{a}$ is updated.

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3292882

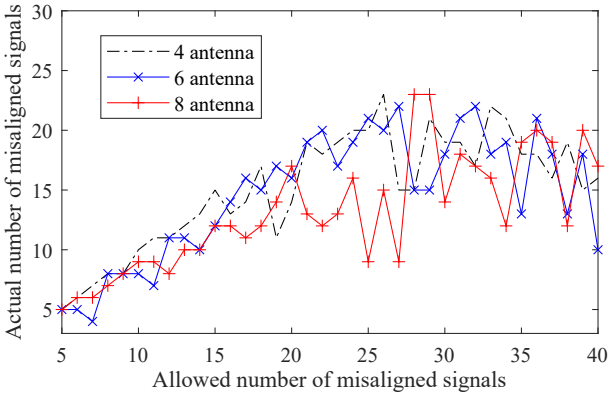


Fig. 5. Actual number of misaligned signals vs. allowed number of misaligned signals (#node = 50, $\sigma^2 = 1$).

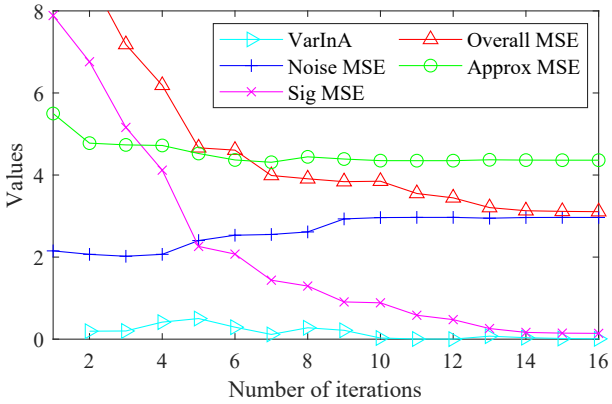


Fig. 6. Convergence of approximate MSE in (12a), actual overall MSE in (3) and the difference of optimal parameters in the iteration (#node = 50, #antenna = 8, $\sigma^2 = 1$).

At the beginning of next iteration, $|\mathbb{S}|$ is increased by 1 and $\mathbb{S}$ is adjusted, based on which $\boldsymbol{a}$ is updated again.

Algorithm 2 Find optimal parameters in single iteration.

1: **procedure** FINDOPTPARAMSIMPLEITERATION($\{\boldsymbol{h}_k\}$)
2: $|\mathbb{S}| \leftarrow 0$
3: Initialize $\hat{\boldsymbol{a}}^0$ with randoms, $\boldsymbol{a}^0 = \hat{\boldsymbol{a}}^0$, $\boldsymbol{A}^0 = \hat{\boldsymbol{a}}^0(\hat{\boldsymbol{a}}^0)^H$
4: Compute $\boldsymbol{H}_k = \boldsymbol{h}_k\boldsymbol{h}_k^H$ for each node k
5: **for** $t = 1, 2, \cdots$ **do**
6: Sort nodes in the ascending order of $|(\boldsymbol{a}^{t-1})^H\boldsymbol{h}_k|$
7: Select top $|\mathbb{S}|$ nodes as $\mathbb{S}$
8: Compute subgradient $\partial||\boldsymbol{A}^{t-1}||_2$
9: Compute $\boldsymbol{A}^t$ by (12a) using CVX
10: Decompose $\boldsymbol{A}^t$ as $\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{U}^H$
11: Compute $\hat{\boldsymbol{a}}^t = \boldsymbol{U}_1\sigma_1(\boldsymbol{\Sigma}^{1/2})$
12: **if** $||\hat{\boldsymbol{a}}^t - \hat{\boldsymbol{a}}^{t-1}||/||\hat{\boldsymbol{a}}^t|| \leq \varepsilon_{\text{th}}$ **then**
13: break ▷ Stop iteration
14: **end if**
▷ Gaussian randomization
15: $\boldsymbol{a}_j = \boldsymbol{U}\boldsymbol{\Sigma}^{\mathbf{1/2}}\boldsymbol{n}_j$, $\boldsymbol{n}_j \in \mathcal{CN}(0, \boldsymbol{I})$, $j = 1, ..., L$
16: Normalize $\boldsymbol{a}_j = \frac{\boldsymbol{a}_j}{\min_{k\notin\mathbb{S}}|\boldsymbol{a}_j^H\boldsymbol{h}_k|}\frac{1}{\sqrt{P_{\max}}}$ ▷ s.t. (7c)
17: **for** $j = 1, 2, \cdots, L$ **do**
18: Compute b_k by (6), $k = 1, \cdots, K$
19: Compute MSE_j in (3) by using $\boldsymbol{a}_j$ and b_k
20: **end for**
21: Find optimal parameter, $\boldsymbol{a}^t = \underset{\boldsymbol{a}_j}{\text{argmin}}\ \text{MSE}_j(\boldsymbol{a}_j)$
22: Increase $|\mathbb{S}|$ by 1
23: **end for**
24: return $\boldsymbol{a}^t$ and $\text{MSE}(\boldsymbol{a}^t)$
25: **end procedure**

The whole procedure is shown in Algorithm 2. $|\mathbb{S}|$ is initialized to 0 (line 2). At the end of each iteration, $|\mathbb{S}|$ is increased by 1 (line 22), and the misaligned signals will be selected at the beginning of next iteration (line 6-7). Solving the problem by CVX (line 8-11) and Gaussian randomization (line 15-21) are the same as in Algorithm 1.

It should be noted that the approximate MSE in (12a) varies with the size of $\mathbb{S}$, and $\boldsymbol{a}^t$ after Gaussian randomization also changes per iteration. In comparison, $\hat{\boldsymbol{a}}^t$ before Gaussian randomization is relatively stable, decided by $\boldsymbol{h}_k$, and gradually converges in a certain range. Therefore, the end of the iteration depends on the convergence of $\hat{\boldsymbol{a}}^t$. At each size of $\mathbb{S}$, the CVX operation is called only once. To improve system performance, Gaussian randomization is invoked in each iteration, and $\boldsymbol{a}^t$ after Gaussian randomization is used for selecting nodes to form $\mathbb{S}$ for the next iteration.

1) Numerical results: Using the same setting as in Sec. IV-E1, we investigate how the iteration converges. Fig. 6 shows the iterative process. Both Sig MSE and Noise MSE gradually converge, so does their sum, the overall MSE in (3). Here, we also show $||\hat{\boldsymbol{a}}^t - \hat{\boldsymbol{a}}^{t-1}||$, the difference of parameters between adjacent iterations (VarInA in the figure). It is clear that this difference approaches 0 in a certain range.

G. Analysis of complexity and optimality

1) Complexity: The computation mainly occurs in three places. The first one is the SDR by CVX. Its computation depends on two parameters, the number of nodes K and the number of antennas N. The computation cost of each CVX is $O(\max\{K, N\}^4 N^{1/2}\log(1/\epsilon))$ [40], where ϵ represents the solution accuracy. It is $O(K^4N^{1/2}\log(1/\epsilon))$ when $K \geq N$. The second one is singular value decomposition (SVD) on the $N \times N$ matrix $\boldsymbol{A}$, and the computation cost is $O(N^3)$. The third one is the Gaussian randomization, and the computation cost is $O(LK)$, where L is the number of candidates of $\boldsymbol{a}$.

In Algorithm 1, with a given $|\mathbb{S}|$, the CVX operation is repeated for a number $f(\varepsilon_{\text{th}})$, depending on ε_{th}. Because FindOptA is invoked for K times, the overall computation cost is $O(K^5N^{1/2}\log(1/\epsilon)f(\varepsilon_{\text{th}}))$, when $K \geq N$.

In comparison, Algorithm 2 has only one loop, and its computation complexity is $O(K^4N^{1/2}\log(1/\epsilon)\min\{K, f(\varepsilon_{\text{th}})\}$ when $K \geq N$.

The zero-forcing (ZF) method corresponds to the special case where $\mathbb{S}$ is an empty set in (12a), as follows,

$$\min_{\boldsymbol{A}} \sigma^2\text{Tr}(\boldsymbol{A}) + \rho \cdot (\text{Tr}(\boldsymbol{A}) - ||\boldsymbol{A}||_2), \tag{13a}$$

$$\text{s.t. } \text{Tr}(\boldsymbol{A}\boldsymbol{H}_k) \geq \frac{1}{P_{\max}},\ \forall k, \tag{13b}$$

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3292882

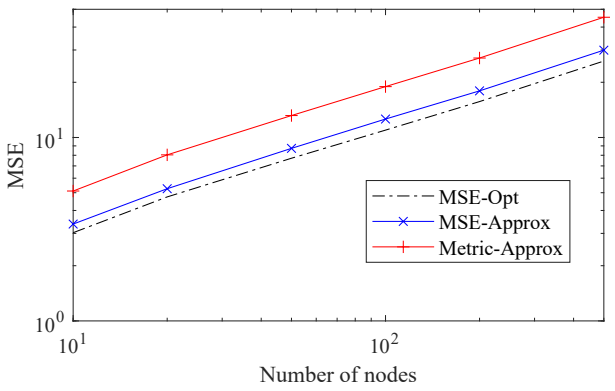


Fig. 7. Optimal MSE, approximate metric, and the actual MSE under the approximation with respect to the number of nodes (#antenna = 1, $\sigma^2 = 1$).

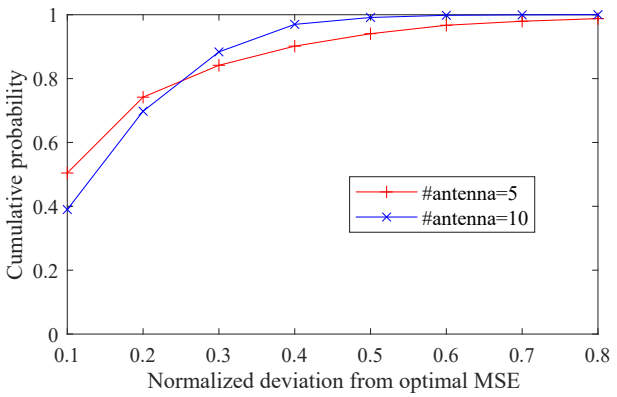


Fig. 8. Probability of MSE deviation from the optimum (#node = 50, $\sigma^2 = 1$).

$$\boldsymbol{A} \succeq 0, \tag{13c}$$

which is solved by FindOptA in Algorithm 1. Its complexity is $O(K^4 N^{1/2} \log(1/\epsilon) f(\varepsilon_{\text{th}}))$ when $K \geq N$. Therefore, Algorithm 2 almost has the same complexity as the ZF method.

2) Optimality : With a given $\mathbb{S}$, the solution of $\boldsymbol{A}$ by CVX is optimal. But for $\boldsymbol{a}$, its result is optimal only if $\boldsymbol{A}$ has a rank 1, where $\text{Tr}(\boldsymbol{A}) - ||\boldsymbol{A}||_2 = 0$. In the optimization, $\text{Tr}(\boldsymbol{A}) - ||\boldsymbol{A}||_2$, as a part of the overall metric, does not always reach 0. So the result of $\boldsymbol{a}$ is not optimal, but approaches the optimum.

The selection of $\mathbb{S}$ depends on an initial $\boldsymbol{a}$ (which is also the case in ZF). This makes the result depend on the initialization. Gaussian randomization helps to alleviate this problem.

When $N = 1$, both $\boldsymbol{a}$ and $\boldsymbol{h}_k$ are scalar, and (10a) becomes

$$|\mathbb{S}_i| + |a|^2 \left(\sigma^2 - P_{\max} \sum_{k \in \mathbb{S}_i} |h_k|^2 \right). \tag{14}$$

Let $|h_k|$ be arranged in the ascending order. Because $|\mathbb{S}_i| = i$, the problem $\mathcal{P}_3$ is approximated as

$$\mathcal{P}_3' \ \min_{a,i} \ i + |a|^2 \left(\sigma^2 - P_{\max} \sum_{k=1}^{i} |h_k|^2 \right), \tag{15a}$$

$$\text{s.t. } |a h_k|^2 < \frac{1}{P_{\max}}, \ k = 1, 2, \cdots, i, \tag{15b}$$

$$|a h_k|^2 \geq \frac{1}{P_{\max}}, \ k = i+1, \cdots, K. \tag{15c}$$

Its solution can be easily computed.

In the case of a single antenna, the optimal MSE has a closed form [16]. The optimal MSE, approximate metric in (15a), and the actual MSE under the approximation, are shown in Fig. 7. Although the approximate metric tends to be greater than the actual MSE, the actual MSE is very close to the optimal value.

When there are more than 2 antennas, $|\boldsymbol{a}^H \boldsymbol{h}_k|$ cannot be separated as $|\boldsymbol{a}||\boldsymbol{h}_k|$ and there is no simple optimal solution. Here, we use random seeds to initialize $\boldsymbol{a}$ and see how MSE varies in Algorithm 2. Assuming the minimal MSE (MSE_o) under all seeds is optimal, we will examine how each MSE deviates from this optimum. The normalized deviation of MSE is computed as $\delta_{\text{MSE}} = |\text{MSE} - \text{MSE}_o|/\text{MSE}_o$, and the probability $p(\delta_{\text{MSE}} < x)$ is shown in Fig. 8, with x as the horizontal axis. Here, 50 network topologies, each with 100 seeds, are used in the evaluation. When the number of antennas is 5, 50% of MSE deviates from the optimum by less than 10%, and 74% of MSE deviates by less than 20%. When the number of antennas is 10, 70% of MSE deviates by less than 20%. This confirms that Algorithm 2 converges towards the optimum with a high probability, with the help of Gaussian randomization.

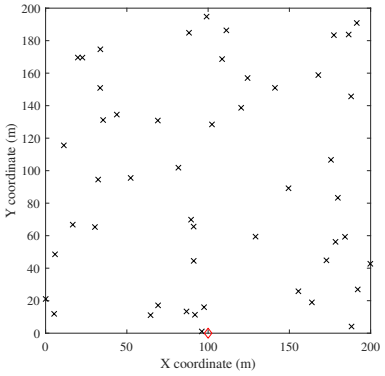


Fig. 9. Simulation scenario with 50 nodes (×) and 1 sink (♢) at the edge.

V. Simulation Evaluation

The proposed Miso method is evaluated by Monte Carlo simulation. Fig. 9 shows the evaluation scenario. Unless specified otherwise, 50 nodes are randomly distributed in a square area of 200m x 200m, and the sink is located at (100, 0), the middle of the bottom edge of the square. Antenna height is 1.5m, the same for all nodes and the sink. The 2.4GHz frequency band is used. A hybrid free-space/two-ray path loss model is assumed, the channel gain without fading is -80dB at a distance of 90m and the power amplification at the sink is 70dB. Block Rayleigh fading is assumed for each link. Parameters in the optimization are set as follows: $P_{\max} = 10$, $\sigma^2 = 1$, $\varepsilon_{\text{th}} = 0.001$.

The proposed Miso method (Algorithm 2) is compared with Algorithm 1, the antenna selection (AntSel) method [16] and the zero-forcing (ZF) method [14]. The MSE difference between Algorithm 1 and Algorithm 2 shows the cost of

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3292882

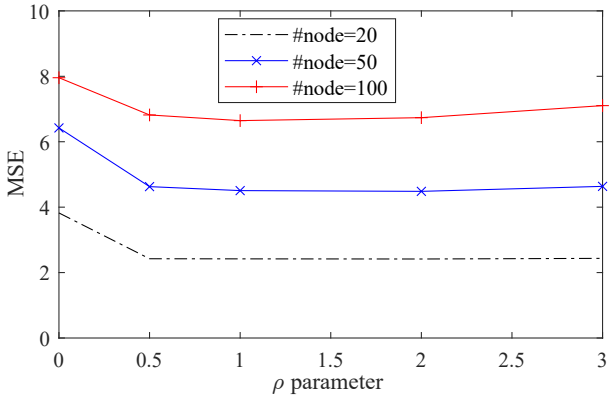


Fig. 10. Computation MSE under different values of ρ (#antenna = 5, $\sigma^2 = 1$, sink at the edge).

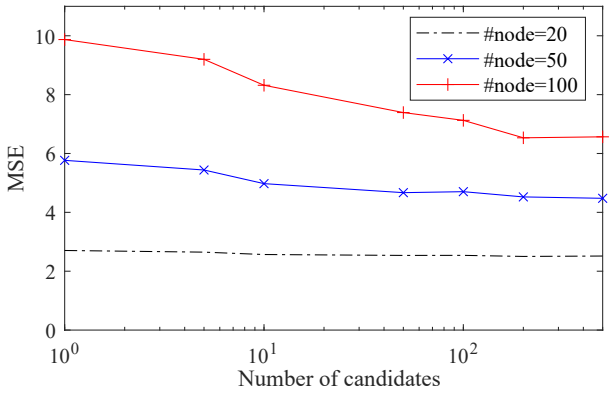


Fig. 11. Computation MSE under different numbers of candidates in Gaussian randomization (#antenna = 5, $\sigma^2 = 1$, sink at the edge).

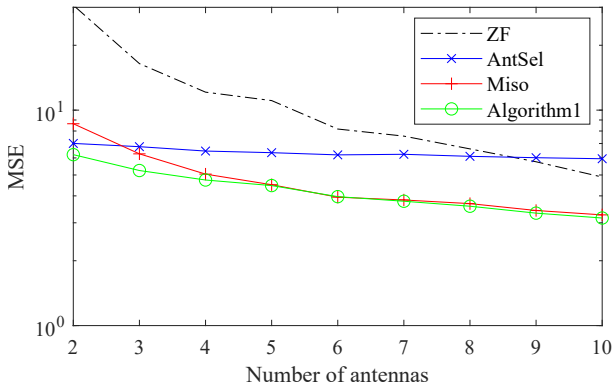


Fig. 12. Computation MSE under different numbers of antennas at the sink (#node = 50, $\sigma^2 = 1$, sink at the edge).

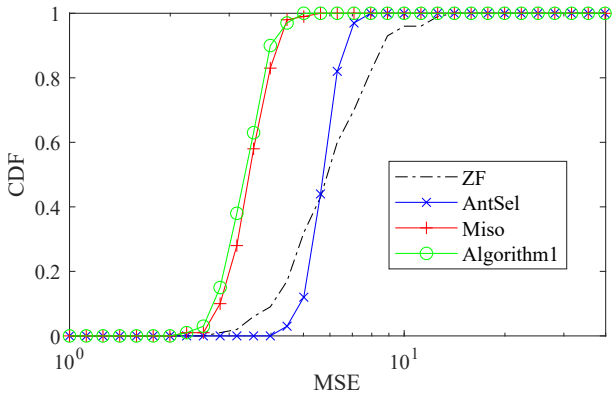


Fig. 13. Cumulative distribution of computation MSE (#node = 50, #antenna = 8, $\sigma^2 = 1$, sink at the edge).

improving the computation speed by the simplification in Algorithm 2. As for AntSel, single antenna AirComp is performed for each antenna at the sink, and the antenna that leads to the minimal computation MSE is selected.

A. Deciding parameters

First, we investigate the impact of parameter ρ. Fig. 10 shows how MSE varies with ρ. When $\rho = 0$, the constraint $\mathrm{rank}(\boldsymbol{A}) = 1$ is neglected in the optimization, and MSE is relatively large. When ρ is greater than 0.5, there is no significant change in MSE. Hereafter, ρ is set to 1.0.

Gaussian randomization plays an important role in finding the optimal $\boldsymbol{a}$ and avoiding the local optimum. Fig. 11 shows how MSE varies with the number of Gaussian randomization candidates. Obviously, when the number of nodes increases, more candidates will help to reduce MSE. Hereafter, the number of candidates is set to 500.

B. Impact of the number of antennas

We first evaluate the impact of the number of antennas at the sink. Fig. 12 shows the computation MSE with respect to the number of antennas. Generally, as the number of antennas increases, the computation MSE decreases in all methods. When the number of antennas is small, even AntSel, with simple antenna selection, outperforms ZF. This is because ZF tries to align magnitude of all signals and its performance is limited by the weakest link. When the number of antennas gets greater than 9, ZF starts to outperform AntSel, benefiting more from the increase of antennas than AntSel. Miso almost always outperforms AntSel by exploiting multiple antennas, and has a large gap compared with ZF by allowing signal misalignment in the optimization. Miso (Algorithm 2) is a little inferior to Algorithm 1, as a result of simplifying the computation. But the performance difference is small when the number of antennas is large.

Fig. 13 shows the cumulative distribution function (CDF) of computation MSE when the number of antennas is fixed to 8. Under this setting, the difference between Miso and Algorithm 1 is very small. On average, Miso reduces the computation MSE by 47.9% compared with ZF, by 43.6% compared with AntSel.

Fig. 14 shows average transmission power with respect to the number of antennas. Each node consumes a little more power in Miso than in AntSel, and consumes least power in ZF. When the number of antennas is fixed to 8, average transmission power in Miso is increased by 12.2% compared with AntSel, by 70.5% compared with ZF. This is affordable, considering that Miso greatly reduces the computation MSE. Algorithm 1, with better MSE performance than Miso, also requires larger transmission power.

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3292882

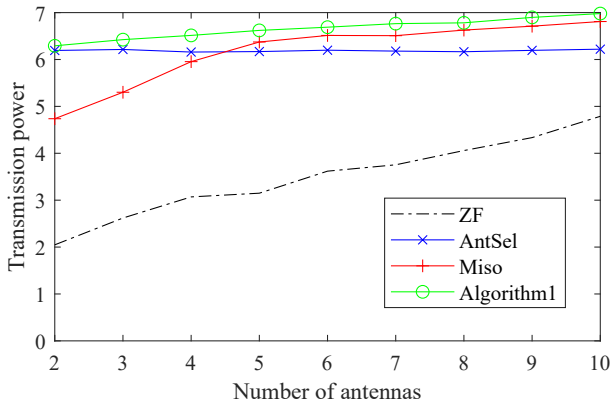


Fig. 14. Average transmission power per node under different numbers of antennas at the sink (#node = 50, $\sigma^2 = 1$, sink at the edge).

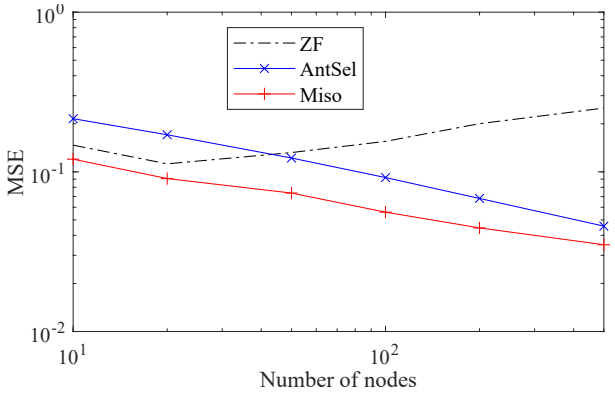


Fig. 15. Average computation MSE per node under different numbers of nodes (#antenna = 8, $\sigma^2 = 1$, sink at the edge).

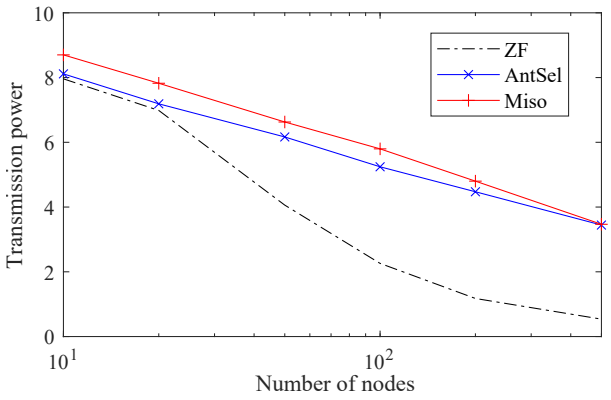


Fig. 16. Average transmission power under different numbers of nodes (#antenna = 8, $\sigma^2 = 1$, sink at the edge).

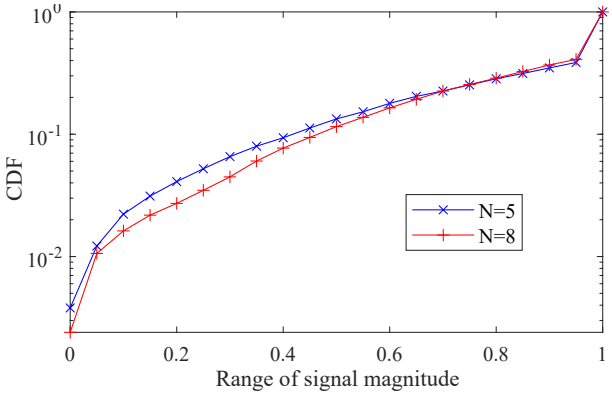


Fig. 17. Cumulative distribution of signal magnitude (#node = 50, $\sigma^2 = 1$, sink at the edge).

C. Impact of the number of nodes

Next we investigate the impact of the number of nodes, using the same setting as in Fig. 9 except that the number of nodes is changed.

With the increase of nodes, more weak links will appear, and unsurprisingly, the overall MSE increases with the number of nodes. But the trend changes when it comes to the average MSE per node (defined as overall MSE divided by the number of nodes), as shown in Fig. 15. Here, average MSE per node first decreases and then increases in ZF, because with more nodes, the probability of very weak links increases. In comparison, in AntSel and Miso, the average MSE per node always decreases with the increase of nodes, because on average the percentage of nodes with weak links is low. As the number of nodes increases, signals are aligned to a smaller magnitude. Accordingly, the average transmission power decreases in all methods, and the difference between Miso and AntSel also decreases, as shown in Fig. 16. These results confirm that the proposed Miso method scales better with the number of nodes than other methods.

D. Accuracy of approximation model

With $K = 50$ nodes and $N = 5, 8$ antennas, we investigate the distribution of signal magnitude to see the accuracy of the approximation in (9). The CDFs are shown in Fig. 17. Most of the signal magnitude (60.1% at $N = 5$, and 57.4% at $N = 8$) are aligned to 1. When also considering distorted signals, signal magnitude is greater than 0.8 with a probability 0.727 at $N = 5$ and $N = 8$, and greater than 0.5 with a probability 0.876 and 0.892 at $N = 5$ and $N = 8$, respectively. At $N = 5$, among the distorted signals, 68.8% signals have a signal magnitude in the range (0.5, 1), and only 15.2% signals have a signal magnitude less than 0.3. This indicates that the approximation in (9) is relatively accurate, but there is still room for further improvement.

E. Impact of noise power

Noise power greatly affects the performance, and its impact on MSE is shown in Fig. 18. Generally, the difference between Miso and ZF increases with noise power. When the noise is very small, the difference almost disappears. This is because Miso tries to take a tradeoff between signal distortion and noise power by applying the Rx-scaling parameter, and the effect of this tradeoff decreases with noise power. At very large noise levels, the AntSel method starts to outperform Miso. In the normal range of noise power, Miso achieves the best performance.

F. Results when sink is at the center

In the above evaluation, the scenario in Fig. 9 is used, where the sink is located at the edge. Next, the sink position

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2023.3292882

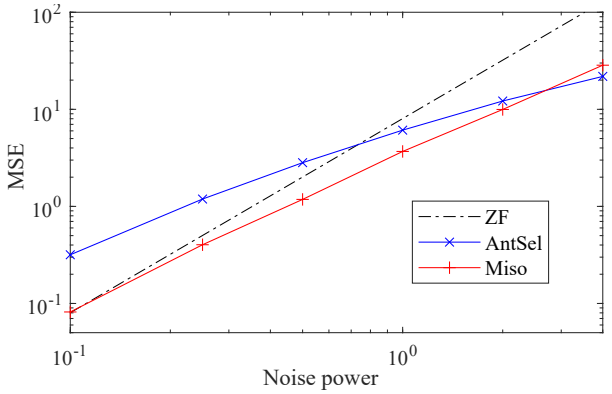


Fig. 18. Average computation MSE under different noise power levels (#node = 50, #antenna = 8, $\sigma^2 = 1$, sink at the edge).

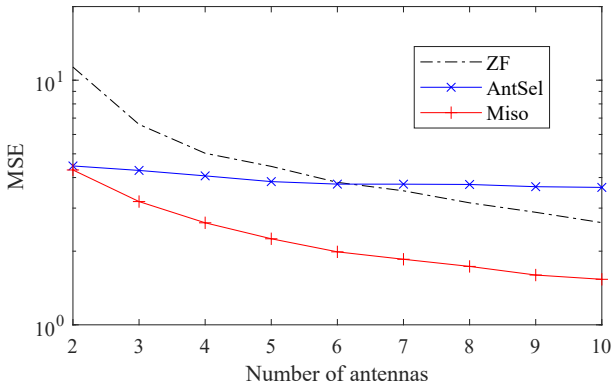


Fig. 19. Computation MSE under different numbers of antennas at the sink (#node = 50, $\sigma^2 = 1$, sink at the center).

is changed to the center of the experiment area, and the performance is evaluated again.

Fig. 19 shows how MSE varies with the number of antennas. Compared with Fig. 12, it is clear that the performance of MSE reduction gets more obvious. This is because when the sink is at the center, the variation in channel gains of different nodes tend to be smaller compared with the case where the sink is located at the edge. The impact of the number of nodes in Fig. 20 shows a similar trend, when compared with that in Fig. 15.

VI. Conclusion

To better exploit multiple antennas at the sink for AirComp, this paper studies allowing misalignment in signal magnitude to minimize the overall computation MSE involving both signal misalignment and noise. This is achieved by explicitly selecting a set of nodes whose signals may be misaligned, to improve the quality of other signals. The optimal parameters for transmission power control at each node and receiving at the sink are found by an iterative method, and the set of nodes whose signals may be misaligned is updated per iteration, which allows a joint optimization with other parameters. In addition, the iterative process is refined to reduce the computation time.

This paper approximates the problem of minimizing computation MSE in the multi-antenna scenario to a tractable one,

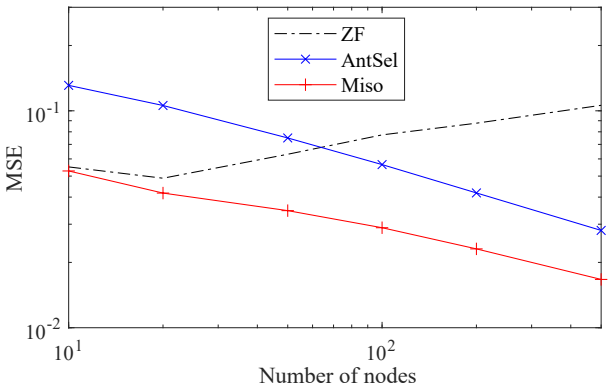


Fig. 20. Average computation MSE per node under different numbers of nodes (#antenna = 8, $\sigma^2 = 1$, sink at the center).

which introduces model error as well. In the future, we will further improve the accuracy of the approximation model.

References

[1] R. S. Sinha, Y. Wei, and S.-H. Hwang, "A survey on LPWA technology: LoRa and NB-IoT," *ICT Express*, vol. 3, no. 1, pp. 14 – 21, 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S2405959517300061

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

[3] J. Xiao, S. Cui, Z. Luo, and A. J. Goldsmith, "Linear coherent decentralized estimation," *IEEE Transactions on Signal Processing*, vol. 56, no. 2, pp. 757–770, 2008.

[4] M. Gastpar, "Uncoded transmission is exactly optimal for a simple gaussian sensor network," *IEEE Transactions on Information Theory*, vol. 54, no. 11, pp. 5247–5251, 2008.

[5] C.-Z. Lee, L. P. Barnes, and A. Ozgur, "Over-the-air statistical estimation," in *IEEE Globecom Proceedings*, 2020.

[6] R. C. Buck, "Approximate complexity and functional representation," *J. Math. Anal. Appl.*, vol. 70, pp. 280–298, 1979.

[7] M. Goldenbaum, H. Boche, and S. Staczak, "Nomographic functions: Efficient computation in clustered gaussian sensor networks," *IEEE Transactions on Wireless Communications*, vol. 14, no. 4, pp. 2093–2105, 2015.

[8] O. Abari, H. Rahul, and D. Katabi, "Over-the-air function computation in sensor networks," *CoRR*, vol. abs/1612.02307, 2016. [Online]. Available: http://arxiv.org/abs/1612.02307

[9] H. Ye, G. Y. Li, and B.-H. F. Juang, "Deep over-the-air computation," in *IEEE Globecom Proceedings*, 2020.

[10] F. Molinari, S. Stanczak, and J. Raisch, "Exploiting the superposition property of wireless communication for average consensus problems in multi-agent systems," in *2018 European Control Conference (ECC)*, 2018, pp. 1766–1772.

[11] L. Chen, N. Zhao, Y. Chen, F. R. Yu, and G. Wei, "Over-the-air computation for cooperative wideband spectrum sensing and performance analysis," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 11, pp. 10 603–10 614, 2018.

[12] G. Zhu, Y. Wang, and K. Huang, "Broadband analog aggregation for low-latency federated edge learning," *IEEE Transactions on Wireless Communications*, vol. 19, no. 1, pp. 491–506, 2020.

[13] M. M. Amiri and D. Gunduz, "Federated learning over wireless fading channels," *IEEE Transactions on Wireless Communications*, vol. 19, no. 5, pp. 3546–3557, 2020.

[14] K. Yang, T. Jiang, Y. Shi, and Z. Ding, "Federated learning via over-the-air computation," *IEEE Transactions on Wireless Communications*, vol. 19, no. 3, pp. 2022–2035, 2020.

[15] C. Sun, S. Wu, and T. Cui, "User selection for federated learning in a wireless environment: A process to minimize the negative effect of training data correlation and improve performance," *IEEE Vehicular Technology Magazine*, vol. 17, 2022.

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

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

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

[19] X. Li, G. Zhu, Y. Gong, and K. Huang, "Wirelessly powered data aggregation for IoT via over-the-air function computation: Beamforming and power control," *IEEE Transactions on Wireless Communications*, vol. 18, no. 7, pp. 3437–3452, 2019.

[20] 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 of Things Journal*, vol. 8, no. 3, pp. 1585–1598, 2021.

[21] S. Tang, P. Popovski, C. Zhang, and S. Obana, "Multi-slot over-the-air computation in fading channels," *IEEE Transactions on Wireless Communications*, pp. 1–11, 2023.

[22] T. Qin, W. Liu, B. Vucetic, and Y. Li, "Over-the-air computation via broadband channels," *IEEE Wireless Communications Letters*, vol. 10, no. 10, pp. 2150–2154, 2021.

[23] S. Tang, H. Yin, C. Zhang, and S. Obana, "Reliable over-the-air computation by amplify-and-forward based relay," *IEEE Access*, vol. 9, pp. 53 333–53 342, 2021.

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

[25] T. Jiang and Y. Shi, "Over-the-air computation via intelligent reflecting surfaces," in *2019 IEEE Global Communications Conference (GLOBECOM)*, 2019, pp. 1–6.

[26] W. Fang, Y. Jiang, Y. Shi, Y. Zhou, W. Chen, and K. B. Letaief, "Over-the-air computation via reconfigurable intelligent surface," pp. 8612–8626, 2021.

[27] L. Chen, X. Qin, and G. Wei, "A uniform-forcing transceiver design for over-the-air function computation," *IEEE Wireless Communications Letters*, vol. 7, no. 6, pp. 942–945, 2018.

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

[29] X. Zhai, X. Chen, J. Xu, and D. W. Kwan Ng, "Hybrid beamforming for massive MIMO over-the-air computation," *IEEE Transactions on Communications*, vol. 69, no. 4, pp. 2737–2751, 2021.

[30] Q. Lan, H. S. Kang, and K. Huang, "Simultaneous signal-and-interference alignment for two-cell over-the-air computation," *IEEE Wireless Communications Letters*, vol. 9, no. 9, pp. 1342–1345, 2020.

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

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

[33] A. Farajzadeh, O. Ercetin, and H. Yanikomeroglu, "Mobility-assisted over-the-air computation for backscatter sensor networks," *IEEE Wireless Communications Letters*, vol. 9, no. 5, pp. 675–678, 2020.

[34] J. Dong, Y. Shi, and Z. Ding, "Blind over-the-air computation and data fusion via provable wirtinger flow," *IEEE Transactions on Signal Processing*, vol. 68, pp. 1136–1151, 2020.

[35] W. Liu, X. Zang, B. Vucetic, and Y. Li, "Over-the-air computation with spatial-and-temporal correlated signals," *IEEE Wireless Communications Letters*, pp. 1–1, 2021.

[36] M. Frey, I. Bjelakovic, and S. Stanczak, "Over-the-air computation in correlated channels," *CoRR*, vol. abs/2101.04690, 2021. [Online]. Available: https://arxiv.org/abs/2101.04690

[37] S. T. Basaran, G. K. Kurt, and P. Chatzimisios, "Energy-efficient over-the-air computation scheme for densely deployed IoT networks," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 5, pp. 3558–3565, 2020.

[38] Y. Shao, D. Gündüz, and S. C. Liew, "Federated edge learning with misaligned over-the-air computation," *CoRR*, vol. abs/2102.13604, 2021. [Online]. Available: https://arxiv.org/abs/2102.13604

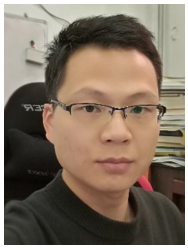
[39] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network: Joint active and passive beamforming design," in *2018 IEEE Global Communications Conference (GLOBECOM)*, 2018, pp. 1–6.

[40] Z.-Q. Luo, W.-K. Ma, A. M.-C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Processing Magazine*, vol. 27, no. 3, pp. 20–34, 2010.

[41] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," http://cvxr.com/cvx, Mar. 2014.

[42] ——, "Graph implementations for nonsmooth convex programs," in *Recent Advances in Learning and Control*, ser. Lecture Notes in Control and Information Sciences, V. Blondel, S. Boyd, and H. Kimura, Eds. Springer-Verlag Limited, 2008, pp. 95–110, http://stanford.edu/~boyd/graph_dcp.html.

Suhua Tang received the B.S. degree in Electronic Engineering in 1998 and the Ph.D. degree in Information and Communication Engineering in 2003, both from University of Science and Technology of China. From Oct. 2003 to Mar. 2014, he was with Adaptive Communications Research Laboratories, ATR, Japan. Since Apr. 2014, he is with Graduate School of Informatics and Engineering, The University of Electro-Communications, Japan. His research interests include green communications, ad hoc and sensor networks, inter-vehicle communications and high precision positioning.



Chao Zhang received the B.Sc. and Ph.D. degrees in electronics and information engineering from the University of Science and Technology of China (USTC), Hefei, Anhui, China, in 2005 and 2010, respectively. Since July 2010, he has been with Xi'an Jiaotong University, Xi'an, China, where he is currently an Associate Professor with the School of Information and Communications Engineering, Faculty of Electronic and Information Engineering. From Dec. 2018 to Dec. 2019, he was a Visiting Professor with the Iowa State University, Ames, IA, USA. His scientific interests include networking and wireless communications, green communications, satellite communications, and 6G related topics.

Jie Li (M'95, SM'01) received the B.E. degree in computer science from Zhejiang University, Hangzhou, China, the M.E. degree in electronic engineering and communication systems from China Academy of Posts and Telecommunications, Beijing, China. He received the Dr. Eng. degree from the University of Electro-Communications, Tokyo, Japan. He is a Chair Professor in Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai, China. His current research interests are in big data and AI, blockchain, edge computing, networking and security, OS, information system architecture and performance evaluation. He was a full professor in Department of Computer Science, University of Tsukuba, Japan. He was a visiting Professor in Yale University, USA, Inria, France. He is the founding Chair of IEEE ComSoc Technical Committee on Big Data and the Co-Chair of IEEE Big Data Community. He serves as an associated editor for many IEEE journals and transactions. He has also served on the program committees for several international conferences.

Sadao Obana received the B.E., M.E. and Ph.D. Degrees from Keio Univ. Tokyo, Japan, in 1976, 1978 and 1993 respectively. After joining KDDI (former KDD) in 1978, he was engaged in R&D in the field of packet exchange systems, network architecture, OSI (Open Systems Interconnection) protocols, database, distributed processing, network management and ITS (Intelligent Transport Systems). In 2004, he joined Advanced Telecommunication Research Institute International (ATR) and was a director of Adaptive Communications Research Laboratories in ATR. From 2011 to 2018, he was a professor with Graduate School of Informatics and Engineering, The University of Electro-Communications, Japan. Currently, he is an Executive Director of The University of Electro-Communications, Japan. He received an Award of Minister of Education, Culture, Sports, Science and Technology in 2001. Dr. Obana is a member of The Institute of Electronics, Information and Communication Engineers (IEICE) and a Fellow of Information Processing Society of Japan (IPSJ).