Robust Design for Massive CSI Acquisition in Analog Function Computation Networks

Fan Ang, Li Chen, Nan Zhao, *Senior Member, IEEE,* Yunfei Chen, *Senior Member, IEEE,*
F. Richard Yu, *Fellow, IEEE*

***Abstract*—Analog function computation utilizes the superposition property of *multi-access channel* (MAC) to compute the target function in an efficient way. However, its corresponding transceiver requires global *channel state information* (CSI) of the network, which incurs large latency. To tackle this challenge, a novel scheme called over-the-air signaling procedure is proposed by exploiting a defined effective CSI in this paper. We first derive the training complexity of the proposed scheme and compare it with the conventional design. It is shown that the training complexity of the proposed scheme can be greatly reduced for massive CSI acquisition by avoiding collecting individual CSI. To account for the difference of the desired CSI, a corresponding robust model is further discussed. Through modeling the channel uncertainties under the expectation-based model and the worst-case model, we formulate the transceiver optimization for both the conventional scheme and the over-the-air signaling procedure. The computational time complexity is derived as a polynomial expression, and it can be significantly reduced for the over-the-air signaling procedure due to its independence of the number of nodes. Finally, the *mean-square error* (MSE) improvement and complexity reduction of the proposed design are demonstrated via simulation.**

***Index Terms*—Analog function computation, expectation-based model, transceiver design, worst-case model.**

I. Introduction

FUTURE wireless applications require higher rate, lower latency and reliable connections with numerous devices [1]. This makes the traditional orthogonal multi-access schemes less feasible as the excessive latency and low efficiency in spectrum utilization. A seminal scheme, analog function computation, was proposed to solve this challenge. It is intelligent to exploit the signal-superposition property of *multi-access channel* (MAC) to compute the desired functions of distributed sensing data from nodes in [2], [3].

This work has been supported by National Natural Science Foundation of China (Grant No. 61601432), and the Fundamental Research Funds for the Central Universities. This work has been supported in part by the National Natural Science Foundation of China under Grant 61871065, and in part by the open research fund of National Mobile Communications Research Laboratory, Southeast University (No. 2018D03). (*Corresponding author: Li Chen*)

F. Ang and L. Chen are with Department of Electronic Engineering and Information Science, University of Science and Technology of China.(e-mail: angfan@mail.ustc.edu.cn, chenli87@ustc.edu.cn).

N. Zhao is with the School of Info. and Commun. Eng., Dalian University of Technology, Dalian 116024 China, and also with National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China (email: zhaonan@dlut.edu.cn).

Y. Chen is with the School of Engineering, University of Warwick, Coventry CV4 7AL, U.K. (e-mail: Yunfei.Chen@warwick.ac.uk).

F.R. Yu is with the Department of Systems and Computer Engineering, Carleton University, Ottawa, ON, K1S 5B6, Canada (email: richard.yu@carleton.ca).

The idea of analog function computation can be tracked back to the information theory analysis in the pioneering work [3]. A novel structure known as compute-and-forward was provided to recover integer coefficient linear combinations, which could obtain higher computation rate by improving communication rate in [4], [5]. Afterwards, a computation scheme was studied to achieve higher computation rate by exploiting the interference property of the Gaussian MAC in [6], [7]. It was based on two categories of nomographic functions proposed in [8], including linear and non-linear functions. To reduce the time and energy complexity in practical scenarios, a novel subfunction allocation was proposed to handle the frequency selective fading and vanishing computation rate issue through the division, allocation and reconstruction of the functions in [9].

Compared with the digital function computation in the aforementioned works, analog function computation can achieve lower complexity and higher energy efficiency for future wireless networks. It is only interested in the desired function values rather than individual messages of all users. The seminal work proposed the analog function computation scheme aiming at compute a variety of functions that used a simple data pre-processing and post-processing strategy in [10]. The feasibility of the analog function computation was proved by the implementation on self-developed software defined radio devices in [11]. Moreover, various applications of analog function computation have been developed for future network scenarios [12], [13]. However, several practical issues of the implementation for analog function computation should be further discussed.

The synchronization required for all nodes is still an open problem in the analog function computation networks. To the best of our knowledge, some existed methods were proposed to solve this problem. The simple robust analog joint source-channel computation was developed in [10], which transformed synchronization error as random noise to solve the synchronization problem. The design transforms the function computation to power detection while synchronization error appears as random noise. Subsequently, another novel way called *AirShare* was studied to utilize broadcasted reference-clock signal to complete the transmission in [14], and the implementation in a network of software radios was provided.

One practical issue originates from the fading property of practical MAC, which motivates the adaptive transceiver design to compensate the non-uniform fading of active nodes. A uniform-forcing transceiver design was proposed to normalize the channel fading for single function computation in [15]. To compute multiple functions, related work discussed the beamforming and channel feedback design to minimize sum *mean-square error* (MSE) via spatial diversity in [16]. In parallel with the above research, the combination of transmitter design with zero-forcing beamforming was introduced to cancel the intra-node interference of multiple functions, and they studied the uniform-forcing power control to compensate the non-uniform fading in [17].

It is worth mentioning that the above research was based on an ideal model with perfect *channel state information* (CSI). Practical transceivers have to operate with uncertain CSI, which inspires researchers to adopt the robust design for conventional networks in [18]–[20]. An intuitive robust precoding technique was developed to eliminate the need for postprocessing at the fusion center for wireless sensor networks in [21]. In analog function computation networks, they discussed the robust transceiver optimization for parallel analog functions computation with MAC in [22].

To the best of our knowledge, the existing methods of the transceiver design for analog function computation are based on the global CSI, which incurs extremely high training complexity for large-scale nodes, and makes the analog function computation lose its prime superiority of avoiding individual data aggregation. Moreover, building on the individual CSI acquisition scheme, the known robust designs are found to have high computational time complexity in [20]. This makes the efficient training process of the transceiver design an open problem.

Motivated by this observation, we propose the over-the-air signaling procedure to reduce the training complexity for massive CSI acquisition. Furthermore, since the known conventional robust designs are inapplicable for the defined effective CSI acquisition, the robust designs of the proposed solution are introduced using the expectation-based model and the worst-case model. The former is based on the statistical properties of CSI uncertainty [23], and the latter represents the fixed CSI uncertainty sets [24]. The robust designs for both models are not jointly convex for the transmitter and the receiver. Thus, we adopt an iterative algorithm to find the efficient optimal solutions. The main contributions of this work are summarized as follows.

- **Over-the-air signaling procedure**: A novel signaling procedure is proposed to obtain the defined effective CSI exploiting the superposition property and the channel reciprocity of MAC. The receiver only requires the defined effective CSI instead of the global CSI to achieve transceiver optimization. The proposed solution is superior to the conventional scheme in terms of both training complexity and MSE performance.
- **Robust design with imperfect CSI**: The robust design is compared between the conventional scheme and the proposed solution. Since the conventional design is infeasible due to the difference of the desired CSI, an iterative algorithm is proposed to obtain the closed form of optimal transceiver under the expectation-based model. Furthermore, we convert the non-convex optimization into *semidefinite programming* (SDP) via S-procedure under the worst-case model [25], [26].
- **Computational time complexity of the design**: The computational time complexity is derived for both the conventional design and the over-the-air signaling procedure. It is found that the computational time complexity of proposed solution is only related to the number of receive antennas, which proves that the proposed solution outperforms the conventional design on the computational time complexity for massive CSI acquisition.

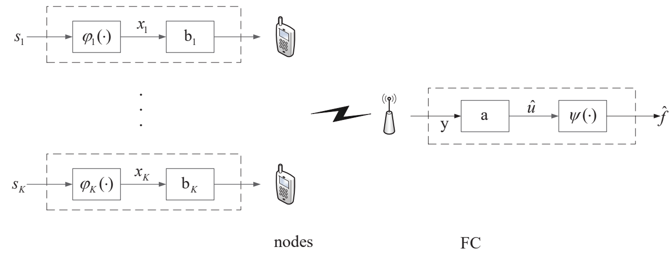


Fig. 1. The analog computation network.

The remainder of the paper is organized as follows. Section II introduces the system model of analog function computation. Section III presents the optimal objective of training complexity and robust design. The over-the-air signaling procedure and the comparison between the conventional scheme and the proposed one are presented in Section IV. Section V shows the transceiver design with imperfect CSI, and the analysis of the computational time complexity. Simulation results are provided in Section VI.

Throughout the paper, we use boldface lowercase to refer to vectors and boldface uppercase to refer to matrices respectively. The real numbers are denoted as $\mathbb{R}$. Let $\mathbf{A}^{-1}$ denote the inverse of a matrix $\mathbf{A}$. Let $\|\cdot\|$ denote the 2-norm of a vector or matrix, and let $(\cdot)^T$ denote the transpose of a vector or matrix. $\mathbf{0}_{m\times n}$ denotes zero matrix with m rows and n columns, $\mathbf{I}_m$ denotes unit matrix with m rows and m columns, and $\mathbf{1}$ denotes unit vector. $\mathcal{N}(0,1)$ is the distribution of real Gaussian with mean 0 and covariance 1. $\mathbb{E}\{\cdot\}$ is the expectation function.

II. System Model

As illustrated in Fig. 1, we consider an uplink system with a single *fusion center* (FC) and K nodes. Each node is equipped with N_t antennas and the FC is equipped with N_r antennas. The data observed by the node k is $s_k \in \mathbb{R}$. Instead of collecting individual data, the FC aims at computing the desired functions. The class of functions to compute by analog function computation is called Nomographic functions.

Definition 1 (Nomographic function [8]). If there exist K pre-processing functions $\varphi_k(\cdot) : \mathbb{R} \to \mathbb{R}$ along with a post-processing function $\Psi(\cdot) : \mathbb{R} \to \mathbb{R}$, the function $f(s_1, s_2, ..., s_K)$ is called Nomographic, which can be expressed as

$$f(s_1, s_2, ..., s_K) = \Psi\left[\sum_{k=1}^{K}\varphi_k(s_k)\right]. \tag{1}$$

Some common nomographic functions with given pre-processing function and post-processing function are listed in Table I.

Each node sends the symbol $x_k = \varphi_k(s_k)$ simultaneously. At the FC, the received symbol with perfect CSI can be expressed as

$$\mathbf{y} = \sum_{k=1}^{K}\mathbf{H}_k\mathbf{b}_k x_k + \mathbf{n}, \tag{2}$$

where $\mathbf{H}_k \in \mathbb{R}^{N_r\times N_t}$ denotes the channel matrix of node k, $\mathbf{b}_k \in \mathbb{R}^{N_t}$ is the transmit beamforming vector of node k, $\mathbf{n}$ represents the noise vector with each element distributed as $\mathcal{N}(0, \sigma_n^2)$, and x_i satisfies $\mathbb{E}\left\{x_i^2\right\} = 1$, $i = 1, 2, ..., K$, $\mathbb{E}\left\{x_i x_j\right\} = 0$, $i \neq j$ and $\mathbb{E}\left\{x_i n_l\right\} = 0$, $l = 1, 2, ..., N_r$, where n_l denotes the l-th element of $\mathbf{n}$.

The estimated value after receive beamforming can be denoted as

$$\hat{u} = \mathbf{a}^T\sum_{k=1}^{K}\mathbf{H}_k\mathbf{b}_k x_k + \mathbf{a}^T\mathbf{n}, \tag{3}$$

where $\mathbf{a}^T \in \mathbb{R}^{N_r}$ denotes the receive beamforming vector.

Therefore, the estimated function $\hat{f}$ can be formulated as

$$\hat{f} = \Psi\left(\hat{u}\right). \tag{4}$$

Compared with the desired function $f = \Psi\left(u\right)$ from all nodes, where

$$u = \sum_{k=1}^{K}x_k, \tag{5}$$

the distortion of $\hat{f}$ in (4) can be represented as

$$\text{MSE}(\hat{f}, f) = \mathbb{E}\left\{\|\hat{f} - f\|^2\right\}. \tag{6}$$

The criterion of the transceiver design is to minimize $\text{MSE}(\hat{f}, f)$. However, it is difficult to obtain the corresponding optimal transceiver for the general form of post-processing function $\psi(\cdot)$. Since the desired value u is close to the received value $\hat{u}$, we consider the Taylor expansion of the desired computed function $\hat{f} = \Psi\left(\hat{u}\right)$ at u so that the distortion of $\hat{f}$ can be formulated as

$$\text{MSE}(\hat{f}, f) \approx [\psi^{'}(u)]^2\text{MSE}(\hat{u}, u), \tag{7}$$

where

$$\text{MSE}(\hat{u}, u) = \mathbb{E}\left\{\|\hat{u} - u\|^2\right\}. \tag{8}$$

Based on the above statement, the objective of analog function computation is to minimize $\text{MSE}(\hat{u}, u)$ through transceiver design.

TABLE I
SOME EXAMPLES OF f

Name	φ_k	ψ	f
Arithmetic Mean	$\varphi_k = s_k$	$\psi = \frac{1}{K}$	$f = \frac{1}{K}\sum\limits_{k=1}^{K}s_k$
Weighted Sum	$\varphi_k = w_k s_k$	$\psi = 1$	$f = \sum\limits_{k=1}^{K}w_k s_k$
Geometric Mean	$\varphi_k = \log(s_k)$	$\psi = \exp(\cdot)$	$f = (\prod\limits_{k=1}^{K}s_k)^{\frac{1}{K}}$
Polynomial	$\varphi_k = w_k {s_k}^{\beta_k}$	$\psi = 1$	$f = \sum\limits_{k=1}^{K}w_k {s_k}^{\beta_k}$
Euclidean Norm	$\varphi_k = {s_k}^2$	$\psi = (\cdot)^{\frac{1}{2}}$	$f = \sqrt{\sum\limits_{k=1}^{K}{s_k}^2}$

Definition 2 (MSE of Analog Function Computation)**.** The MSE of equivalent desired function value u can be written as

$$\begin{aligned}\text{MSE}(\hat{u}, u) &= \mathbb{E}\left\{\text{tr}[(\hat{u} - u)(\hat{u} - u)^T]\right\}\\ &= \sum_{k=1}^{K}\|\mathbf{a}^T\mathbf{H}_k\mathbf{b}_k\|^2 + \sigma_n^2\|\mathbf{a}^T\|^2,\end{aligned} \tag{9}$$

where $\hat{u}$ denotes the estimated function value in (3).

III. PROBLEM STATEMENT

In this section, the training complexity of the conventional optimization is discussed, and the robust transceiver design is further considered.

A. Training Complexity of Optimal Transceiver Design

We consider the joint adaptive transceiver design subject to the transmission power constraint, i.e., the average transmission power of each symbol of node k cannot exceed a given positive threshold P_k. Since the signal x_i satisfies $\mathbb{E}\left\{x_i^2\right\} = 1$, the problem of $\text{MSE}(\hat{u}, u)$ optimization with perfect CSI can be expressed as

$$\begin{aligned}\text{P1}:\ &\min_{\mathbf{a},\mathbf{b}_k}\ \text{MSE}(\hat{u}, u)\\ &\text{s.t.}\ \|\mathbf{b}_k\|^2 \leq P_k, k = 1, 2, ..., K,\end{aligned} \tag{10}$$

where $\text{MSE}(\hat{u}, u)$ is given in (9).

The objective function of P1 is convex over each of the transmit vectors or receive vector, but not jointly convex. Thus, we adopt a classic efficient algorithm to find the optimal solution illustrated in [27], where the optimal results can be expressed as

$$\mathbf{a} = \left(\sigma_n^2\mathbf{I} + \sum_{k=1}^{K}\mathbf{H}_k\mathbf{b}_k\mathbf{b}_k^T\mathbf{H}_k^T\right)^{-1}\left(\sum_{k=1}^{K}\mathbf{H}_k\mathbf{b}_k\right), \tag{11}$$

$$\mathbf{b}_k = \left(\mathbf{H}_k^T\mathbf{a}\mathbf{a}^T\mathbf{H}_k + \mu_k\mathbf{I}\right)^{-1}\mathbf{H}_k^T\mathbf{a}, k = 1, 2, ..., K, \tag{12}$$

where $\mu_k \geq 0$ and satisfies

$$\mu_k \left[\mathrm{tr}(\mathbf{b}_k^T \mathbf{b}_k) - P_k \right] = 0, k = 1, 2, ..., K. \tag{13}$$

The main idea of the transceiver design is to find the optimal transmit vector $\mathbf{b}_k$ via (12) for a fixed receive vector $\mathbf{a}$, and find the optimal receive vector $\mathbf{a}$ via (11) for the newly obtained transmit vector $\mathbf{b}_k$. The iterative algorithm will stop under a specific condition.

It can be seen that the global CSI acquisition of all nodes is essential for the transceiver design in P1. In fact, the training complexity of the conventional approach would incur large latency and huge overhead with massive nodes, and it also makes the analog function computation lose its prime advantage of avoiding collecting individual data.

To tackle this challenge, an intelligent scheme called over-the-air signaling procedure is proposed in Section IV, whose training complexity is positively irrelevant with the number of nodes K but increases significantly slower than the conventional design.

B. Optimal Transceiver with Imperfect CSI

The aforementioned transceiver design generally requires perfect CSI at both sides. However, practical transceivers usually operate under uncertain CSI. Robust design is proposed to ensure a certain level of the performance under the CSI uncertainty model, which can be generally expressed as

$$\mathbf{CSI} = \hat{\mathbf{CSI}} + \Delta\mathbf{CSI}, \tag{14}$$

where $\hat{\mathbf{CSI}}$ denotes the nominally available value of the CSI, and $\Delta\mathbf{CSI}$ is the channel uncertainty set, which can be protrayed as the expectation-based robust and the worst-case robust manners.

Expectation-based robust model is adopted to handle the channel robust manner while only channel statistical properties are available.

Definition 3 (Expectation-Based Robust Model [23])**.** In the expectation-based robust model, the entries of the uncertainty matrix are assumed to be Gaussian distributed with $\mathbb{E}\left\{\Delta\mathbf{CSI}\right\} = \mathbf{0}$, and $\mathbb{E}\left\{\Delta\mathbf{CSI} \cdot \Delta\mathbf{CSI}^T\right\} = \sigma_h^2 \mathbf{I}$.

Robust MSE optimization problem in the expectation-based model can be formed as

$$\begin{aligned} \mathrm{P2}: \ & \min_{\mathbf{a}, \mathbf{b}_k} \mathrm{MSE}|\hat{\mathbf{CSI}} \\ & \text{s.t. } \|\mathbf{b}_k\|^2 \leq P_k, k = 1, 2, ..., K, \end{aligned} \tag{15}$$

where $\mathrm{MSE}|\hat{\mathbf{CSI}}$ is the conditional MSE with given $\hat{\mathbf{CSI}}$.

Since the P2 is non-convex, we utilize an iterative algorithm to convert it into a convex problem and illustrate the closed-form solution of the optimal transceiver in Section V.

However, the expectation-based robust model is not proper for all systems, especially for the system with strict MSE requirements. The alternative model is to have fixed uncertainty sets and to maximize the performance under the worst channel uncertainty, known as the worst-case model and defined as below.

Definition 4 (Worst-Case Robust Model [24])**.** In the worst-case robust model, the norm of the channel uncertainties matrix $\Delta\mathbf{CSI}$ is bounded by the spherical region, which can be expressed as

$$\|\Delta\mathbf{CSI}\|^2 \leq \sigma_h^2, \tag{16}$$

where $\sigma_h^2 \geq 0$ denotes the radius of the spherical channel uncertainty region.

Thus, the transceiver design becomes a min-max problem as

$$\begin{aligned} \mathrm{P3}: \ & \min_{\mathbf{a}, \mathbf{b}_k} \max_{\Delta\mathbf{CSI}} \mathrm{MSE}|\hat{\mathbf{CSI}} \\ & \text{s.t. } \|\mathbf{b}_k\|^2 \leq P_k, k = 1, 2, ..., K, \\ & \|\Delta\mathbf{CSI}\|^2 \leq \sigma_h^2. \end{aligned} \tag{17}$$

One challenge to resolve P3 is the unavailable worst channel uncertain condition during the optimization. We can settle this utilizing a novel scheme known as S-procedure, which transforms the channel uncertain constraints into SDP by adding extra variables which can be optimized. The other challenge is the non-convex objective function over transceiver, which is solved by an iterative algorithm in Section V.

IV. Over-the-Air Signaling Procedure

In this section, we introduce the conventional scheme and propose the over-the-air signaling procedure which requires the defined effective CSI.

A. Conventional Signaling Procedure

Based on the iterative process of the transceiver design, the conventional signaling procedure can be illustrated as Fig. 2. (a).

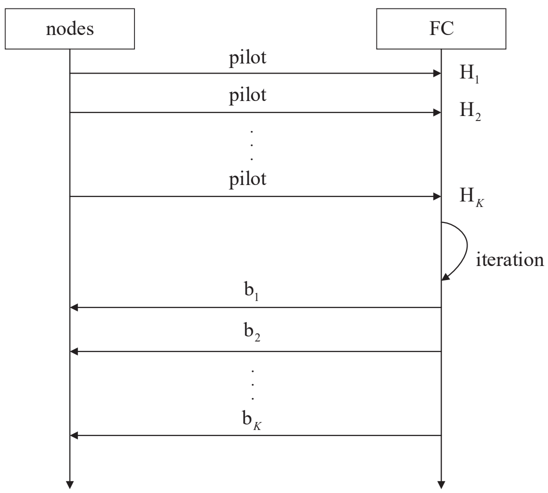
The training process of conventional signaling procedure can be mainly divided into the following steps.

- CSI acquisition: Each node sends a pilot to the FC in turn to estimate its global CSI.
- Algorithm operation: The iterative algorithm of the transceiver optimization is executed at the FC.
- Optimal results acquisition: The FC sends the corresponding optimal transmit vector to each node in turn.

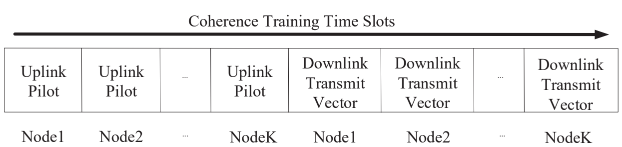
According to the signaling procedure, we derive the training complexity of the conventional signaling procedure in the following.

Proposition 1 (Training Complexity of Conventional Signaling Procedure)**.** The conventional signaling procedure takes at least $2N_tK$ time slots.

Proof: As shown in Fig. 2. (b), each node first transmits a pilot vector to the FC in turn, which takes at least N_tK time slots to acquire the global CSI. In order to send corresponding optimal transmit vector $\mathbf{b}_k$, another N_tK time slots should be spent. In conclusion, it would spend at least $2N_tK$ time slots for training process. ∎

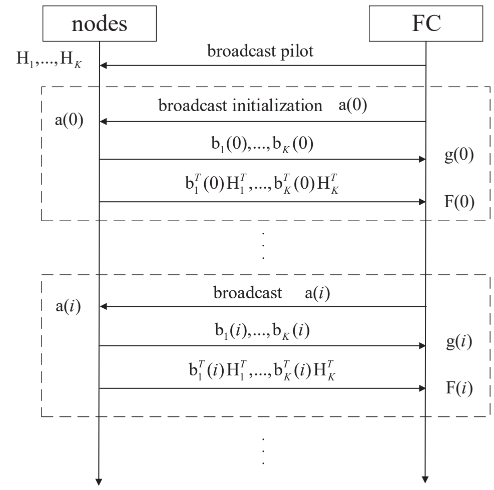


(a) Conventional Signaling Procedure

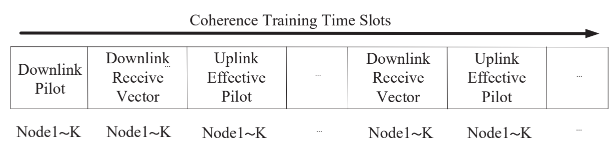


(b) Conventional Training Time Slots

Fig. 2. Schematic diagram of conventional signaling procedure and training time slots.



(a) Over-the-Air Signaling Procedure



(b) Over-the-Air Training Time Slots

Fig. 3. Schematic diagram of over-the-air signaling procedure and training time slots.

B. Proposed Signaling Procedure

The analysis above indicates that the training complexity of the conventional signaling procedure is linear increasing with the number of nodes K. It will cause serious latency with massive nodes. Note that the FC only needs the combination of global CSI instead of individual CSI in (11), and the optimal transmit vector $\mathbf{b}_k$ of node k is only related to its own global CSI in (12). As illustrated in Fig. 3. (a), we propose the over-the-air signaling procedure to reduce the training complexity.

The training process of the over-the-air signaling procedure can be roughly divided into two steps.

- Individual CSI acquisition of nodes: The FC broadcasts a pilot to estimate their own CSI.
- Transceiver optimization iteration: The iterative algorithm continues until meets the specific condition. The details are illustrated as Algorithm 1.

To introduce the proposed signaling procedure more specifically, we provide the iteration process in details as follows. In each iteration, the FC broadcasts the current receive vector $\mathbf{a}$ to all nodes so that the current optimal transmit vector $\mathbf{b}_k$ can be obtained through (12). Later, the FC would obtain the effective CSI vector $\mathbf{g}$ when all nodes send the pilot $\mathbf{1}$ simultaneously, which can be expressed as

$$\mathbf{g} = \sum_{k=1}^{K} \mathbf{H}_k \mathbf{b}_k. \tag{18}$$

Similarly, the effective CSI matrix $\mathbf{F}$ is obtained by sending the pilot $\mathbf{b}_k^T \mathbf{H}_k^T$ where

$$\mathbf{F} = \sum_{k=1}^{K} \mathbf{H}_k \mathbf{b}_k \mathbf{b}_k^T \mathbf{H}_k^T. \tag{19}$$

The FC can obtain its optimal receive vector $\mathbf{a}$ at the current iterative time through bringing effective CSI (18) and (19) into (11).

In conclusion, we illustrate the optimal transceiver as

$$\mathbf{a} = \left(\sigma_n^2 \mathbf{I} + \mathbf{F}\right)^{-1} \mathbf{g}, \tag{20}$$

$$\mathbf{b}_k = \left(\mathbf{H}_k^T \mathbf{a}\mathbf{a}^T \mathbf{H}_k + \mu_k \mathbf{I}\right)^{-1} \mathbf{H}_k^T \mathbf{a}, \tag{21}$$

where μ_k satisfies

$$\mu_k \left[\mathrm{tr}(\mathbf{b}_k^T \mathbf{b}_k) - P_k\right] = 0. \tag{22}$$

The values of μ_k are either positive such that the power constraint holds or zero. The details of the iterative algorithm are shown in Algorithm 1, which elaborate the updating process of the Lagrange multipliers μ_k, the receive vector $\mathbf{a}$ and the transmit vector $\mathbf{b}_k$ during each iterative time. The iteration continues until stopping criterion can be met. We analyze the training complexity of the over-the-air signaling procedure in the following.

Proposition 2 (Training Complexity of Over-the-Air Signaling

Procedure)**.** Assuming the number of iterative times is N_{iter}, the number of essential time slots of the over-the-air signaling procedure is $N_r + (1 + 2N_r)N_{iter}$.

Proof: As shown in Fig. 3. (b), the first process of broadcasting pilot vector would take at least N_r time slots. During each iterative time, we spend one time slot to estimate the effective CSI vector $\mathbf{g}$ in (18) and N_r time slots for the effective CSI matrix $\mathbf{F}$ in (19). The optimal receive vector $\mathbf{a}$ should be broadcasted to all nodes, which would consume another N_r time slots. Thus, it would take $N_r+(1+2N_r)N_{iter}$ time slots for training process. ■

Obviously, the training complexity is irrelevant with the number of nodes K. It is caused by utilizing the broadcasting property, the reciprocity property and the superposition property of MAC to obtain the effective CSI. Through the simulation results, the number of iterative times N_{iter} is positively increasing with the number of nodes K. However, it grows significantly slower than the conventional design.

We compare the two signaling procedures to elaborate the corresponding performance in the following.

Remark 1 (Training Complexity Comparison)**.** The conventional signaling procedure takes $2N_tK$ times slots, while the over-the-air signaling procedure requires $N_r+(1+2N_r)N_{iter}$. The training complexity of the proposed scheme is positively correlated with the number of nodes K but grows slower than the conventional scheme. There exists an intersection point with the corresponding number of nodes $M = [N_r + (1 + 2N_r)N_{iter}]/2N_t$. The over-the-air signaling procedure shows superior performance when $K > M$.

For more intuitive comparison of the training complexity, we use a system composed of $K = 100$ active nodes as an example. Each node and the FC are equipped with $N_t = N_r = 2$ antennas. According to the simulation results in Section VI, the conventional scheme demands $2N_tK = 400$ time slots. The number of iterative times satisfies $N_{iter} \leq 15$ so that the required time slots of the over-the-air signaling procedure is $N_r + (1+2N_r)N_{iter} = 77$, which is 19% of the conventional scheme. Moreover, the gap between the two training methods becomes larger with the growth of the number of nodes.

In each iterative time of the proposed algorithm, the MSE is reduced by the transceiver results updating. Since the MSE is lower bounded by zero, the convergence of Algorithm 1 can be guaranteed. Nevertheless, it is difficult to ensure the global convergence caused by its non-convex optimization.

Algorithm 1 Training Procedure for Over-the-Air Signaling Procedure

Input: $\mathbf{H}_k$,Initialization $\mathbf{a}(0)$
Output: $\mathbf{a}$, $\mathbf{b}_k$, μ_k

number of iteration time n = 0.
Initialize $\mathbf{a}$ = $\mathbf{a}(0)$.
Update $\mathbf{a}$, $\mathbf{b}_k$, μ_k:
repeat
  Update $\mathbf{b}_k(n)$ based on Equation (21), $k = 1, 2, ..., K$
  Update $\mathbf{g}(n + 1)$ based on Equation (18)
  Update $\mathbf{F}(n + 1)$ based on Equation (19)
  Update $\mathbf{a}(n + 1)$ based on Equation (20)
  Update $\mu_k(n + 1)$ based on Equation (22), $k = 1, 2, ..., K$
  $n \leftarrow n + 1$
until converge

V. Robust Design with Imperfect CSI

In this section, we present the transceiver design for handling CSI uncertainty in a robust manner with the conventional method and the proposed design, which are discussed under the expectation-based robust model and the worst-case robust model.

The above transceiver design is based on the global accurate CSI for all nodes and the FC. Since it is difficult to obtain accurate CSI in wireless communications, which is originated from a variety of sources, e.g., imperfect channel estimation, feedback quantization, and delay in CSI acquisition on fading channels, we propose the robust design.

In the conventional signaling procedure, CSI is represented by the individual estimated channel for each node as defined below.

Definition 5 (Robust CSI Model with Conventional Signaling Procedure)**.** In the conventional signaling procedure, the CSI uncertainty model can be expressed as

$$\begin{aligned} \hat{\mathbf{CSI}} &= \hat{\mathbf{H}}_k, k = 1, 2, ..., K, \\ \Delta\mathbf{CSI} &= \Delta\mathbf{H}_k, k = 1, 2, ..., K, \end{aligned} \tag{23}$$

where $\hat{\mathbf{H}}_k$ denotes the nominally global CSI available at both sides, and $\Delta\mathbf{H}_k$ is the estimated channel uncertainty at FC.

Similarly, we can model the CSI uncertainty with the over-the-air signaling procedure.

Definition 6 (Robust CSI Model with Over-the-Air Signaling Procedure)**.** In the over-the-air signaling procedure, the estimated uncertainty of the effective CSI contains two parts. The one is the effective CSI vector uncertainty via (19) as

$$\begin{aligned} \hat{\mathbf{CSI}}_1 &= \hat{\mathbf{F}} = \sum_{k=1}^{K} \mathbf{H}_k \mathbf{b}_k \mathbf{b}_k^T \mathbf{H}_k^T, \\ \Delta\mathbf{CSI}_1 &= \Delta\mathbf{F}, \end{aligned} \tag{24}$$

and the other one is the effective CSI matrix uncertainty via (18)

$$\begin{aligned} \hat{\mathbf{CSI}}_2 &= \hat{\mathbf{g}} = \sum_{k=1}^{K} \mathbf{H}_k \mathbf{b}_k, \\ \Delta\mathbf{CSI}_2 &= \Delta\mathbf{g}, \end{aligned} \tag{25}$$

where $\mathbf{H}_k$ denotes the available individual CSI at node k, $\hat{\mathbf{F}}$ and $\hat{\mathbf{g}}$ are the nominal effective CSI only available at FC, and $\Delta\mathbf{F}$, $\Delta\mathbf{g}$ represent the estimated uncertainty at FC.

In the following part, we propose the corresponding transceiver design with the conventional scheme and the over-the-air signaling procedure under the expectation-based and the worst-case robust models.

A. Expectation-Based Robust Design with Conventional Signaling Procedure

According to the robust model in (23), the optimization of MSE with the conventional signaling procedure under the expectation-based robust model in P2 can be expressed as

$$\begin{aligned} \text{P4}: \ & \min_{\mathbf{a},\mathbf{b}_k} \text{MSE}|\hat{\mathbf{H}}_k \\ & \text{s.t. } \|\mathbf{b}_k\|^2 \leq P_k, k=1,2,...,K, \end{aligned} \tag{26}$$

where

$$\begin{aligned} \text{MSE}|\hat{\mathbf{H}}_k = & \sum_{k=1}^{K} \|\mathbf{a}^T\hat{\mathbf{H}}_k\mathbf{b}_k - 1\|^2 + \sigma_n^2\|\mathbf{a}^T\|^2 \\ & + \sigma_h^2\|\mathbf{a}^T\|^2\sum_{k=1}^{K}\|\mathbf{b}_k\|^2, \end{aligned} \tag{27}$$

and $\Delta\mathbf{H}_k$ satisfies $\mathbb{E}\{\Delta\mathbf{H}_k\} = \mathbf{0}_{N_r\times N_t}$, the second-order statistics of $\Delta\mathbf{H}_k$ satisfies $\mathbb{E}\left\{\Delta\mathbf{H}_k\cdot\Delta\mathbf{H}_k^T\right\} = \sigma_h^2\mathbf{I}_{N_r}$, and $\mathbb{E}\left\{\Delta\mathbf{H}_k\cdot\Delta\mathbf{H}_j^T\right\} = \mathbf{0}_{N_r\times N_r}$, $k\neq j$.

Since $\text{MSE}|\hat{\mathbf{H}}_k$ is not jointly convex on transmit vector $\mathbf{a}$ and receive vector $\mathbf{b}_k$, the similar iterative algorithm of P1 can be adopted to find the optimal transceiver.

According to the problem formulation in P4, the Lagrange dual objective function can be constructed as

$$L(\mathbf{a},\mathbf{b}_k,\mu_k) = \text{MSE}|\hat{\mathbf{H}}_k + \sum_{k=1}^{K}\mu_k\left(\mathbf{b}_k^T\mathbf{b}_k - P_k\right), \tag{28}$$

where μ_k is the Lagrange multiplier associated with the power constraint of the node k.

Proposition 3 (Optimal Robust Transceiver with Conventional Signaling Procedure)**.** The expression of the optimal transmit vectors and receive vector can be formulated as

$$\mathbf{a} = \left[\sigma_n^2\mathbf{I} + \sum_{k=1}^{K}\left(\mathbf{H}_k\mathbf{b}_k\mathbf{b}_k^T\mathbf{H}_k^T + \sigma_h^2\mathbf{b}_k^T\mathbf{b}_k\mathbf{I}\right)\right]^{-1}\left(\sum_{k=1}^{K}\mathbf{H}_k\mathbf{b}_k\right), \tag{29}$$

$$\mathbf{b}_k = \left(\mathbf{H}_k^T\mathbf{a}\mathbf{a}^T\mathbf{H}_k + \mu_k\mathbf{I} + \sigma_h^2\mathbf{a}^T\mathbf{a}\mathbf{I}\right)^{-1}\mathbf{H}_k^T\mathbf{a}, k=1,2,...,K, \tag{30}$$

where $\mu_k \geq 0$ and satisfies

$$\mu_k\left[\text{tr}(\mathbf{b}_k^T\mathbf{b}_k) - P_k\right] = 0, k=1,2,...,K. \tag{31}$$

Proof: Taking the partial derivative of L with respect to the receive vector $\mathbf{a}$ and transmit vector $\mathbf{b}_k$ and using the KKT conditions, we can obtain the optimality conditions as shown in (29) and (30). ■

B. Expectation-Based Robust Design with Over-the-Air Signaling Procedure

In this part, we assume that the CSI uncertainty is originated from the same source for analytic convenience. It implies that the elements of uncertainty matrix $\mathbf{F}$ and vector $\mathbf{g}$ satisfy the same distribution. We can model the CSI error as

$$\begin{aligned} & \mathbf{F}_i = \hat{\mathbf{F}}_i + \boldsymbol{\Delta}_i, i=1,2,...,N_r, \\ & \mathbf{g} = \hat{\mathbf{g}} + \boldsymbol{\Delta}_{N_r+1}, \end{aligned} \tag{32}$$

where $\boldsymbol{\Delta}$ denotes the channel uncertainty sets matrix, $\mathbf{A}_i$ is the column vector of matrix $\mathbf{A}$, $\boldsymbol{\Delta}_i$ satisfies $\mathbb{E}\{\boldsymbol{\Delta}_i\} = \mathbf{0}_{N_r\times 1}$, $\mathbb{E}\left\{\boldsymbol{\Delta}_i\cdot\boldsymbol{\Delta}_i^T\right\} = \sigma_h^2\mathbf{I}_{N_r}$, and $\mathbb{E}\left\{\boldsymbol{\Delta}_i\cdot\boldsymbol{\Delta}_j^T\right\} = \mathbf{0}_{N_r\times N_r}$, $i\neq j$, $i=1,2,...,N_r+1$.

According to (32), the expectation-based robust design in P2 can be written as

$$\begin{aligned} \text{P5}: \ & \min_{\mathbf{a},\mathbf{b}_k} \text{MSE}|\left\langle\hat{\mathbf{F}},\hat{\mathbf{g}}\right\rangle \\ & \text{s.t. } \|\mathbf{b}_k\|^2 \leq P_k, k=1,2,...,K, \end{aligned} \tag{33}$$

where

$$\begin{aligned} \text{MSE}|\left\langle\hat{\mathbf{F}},\hat{\mathbf{g}}\right\rangle & = \mathbb{E}\left\{\mathbf{a}^T\mathbf{F}\mathbf{a} - \mathbf{a}^T\mathbf{g} - \mathbf{a}\mathbf{g}^T + K + \sigma_n^2\mathbf{a}^T\mathbf{a}\right\} \\ & = \mathbf{a}^T\hat{\mathbf{F}}\mathbf{a} - \mathbf{a}^T\hat{\mathbf{g}} - \mathbf{a}\hat{\mathbf{g}}^T + K + \sigma_n^2\mathbf{a}^T\mathbf{a} \\ & + \mathbb{E}\left\{\sum_{i=1}^{N_r}\mathbf{a}^T\boldsymbol{\Delta}_i a_i - \mathbf{a}^T\boldsymbol{\Delta}_{N_r+1} - \mathbf{a}\boldsymbol{\Delta}_{N_r+1}^T\right\}. \end{aligned} \tag{34}$$

The CSI uncertainty variable $\boldsymbol{\Delta}_i$ is a zero mean variable so that $\text{MSE}|\left\langle\hat{\mathbf{F}},\hat{\mathbf{g}}\right\rangle$ can be formed as

$$\text{MSE}|\left\langle\hat{\mathbf{F}},\hat{\mathbf{g}}\right\rangle = \mathbf{a}^T\hat{\mathbf{F}}\mathbf{a} - \mathbf{a}^T\hat{\mathbf{g}} - \mathbf{a}\hat{\mathbf{g}}^T + K + \sigma_n^2\mathbf{a}^T\mathbf{a}. \tag{35}$$

The same structure of (35) and (9) makes the method of transceiver design under perfect CSI being totally applicable for P5. Thus, we can obtain the expression of the transceiver utilizing the Lagrange dual objective function which is

$$L(\mathbf{a},\mathbf{b}_k,\mu_k) = \text{MSE}|\left\langle\hat{\mathbf{F}},\hat{\mathbf{g}}\right\rangle + \sum_{k=1}^{K}\mu_k\left(\mathbf{b}_k^T\mathbf{b}_k - P_k\right), \tag{36}$$

where μ_k is the Lagrange multiplier associated with the power constraint of the node k.

Proposition 4 (Optimal Robust Transceiver with Over-the-air Signaling Procedure)**.** The expression of the optimal transmit vectors and receive vector can be formulated as

$$\mathbf{a} = \left(\sigma_n^2\mathbf{I} + \hat{\mathbf{F}}\right)^{-1}\hat{\mathbf{g}}, \tag{37}$$

$$\mathbf{b}_k = \left(\mathbf{H}_k^T\mathbf{a}\mathbf{a}^T\mathbf{H}_k + \mu_k\mathbf{I}\right)^{-1}\mathbf{H}_k^T\mathbf{a}, k=1,2,...,K, \tag{38}$$

where $\mu_k \geq 0$ and satisfies

$$\mu_k\left[\mathrm{tr}(\mathbf{b}_k^T\mathbf{b}_k)-P_k\right]=0, k=1,2,...,K. \tag{39}$$

Proof: Similar to the optimization in P4, we take the partial derivative of $L(\mathbf{a},\mathbf{b}_k,\mu_k)$ with respect to the receive vector $\mathbf{a}$ and transmit vector $\mathbf{b}_k$ and using the KKT conditions, the optimality conditions can be obtained as (37) and (38). ■

Similarly, the results can be obtained via iterative algorithm under the perfect CSI. The iterative algorithm will stop under a specific condition. The transceiver design can improve the MSE performance than the conventional scheme due to the noise averaging effect, which is illustrated by simulation results in Section VI.

C. Worst-Case Robust Design with Conventional Signaling Procedure

In section III, we indicate that the statistical properties of the CSI uncertainties are no longer suitable to guarantee the MSE constraints exactly. An alternative robust model is developed to maximize the performance under the worst channel uncertainty and known as the worst-case model.

According (24) and (25), the optimization of MSE for P3 can be formulated as

$$\begin{aligned}\mathrm{P6}:\ \min_{\mathbf{a},\mathbf{b}_k}\max_{\Delta\mathbf{H}_k}\ & \mathrm{MSE}|\hat{\mathbf{H}}_k,\\ \text{s.t. } & \|\mathbf{b}_k\|^2\le P_k, k=1,2,...,K,\\ & \|\Delta\mathbf{H}_k\|^2\le\sigma_h^2, k=1,2,...,K,\end{aligned} \tag{40}$$

where $\|\Delta\mathbf{H}_k\|^2\le\sigma_h^2$ denotes the constraint of uncertainty set, and $\mathrm{MSE}|\hat{\mathbf{H}}_k$ is the MSE under estimated CSI which can be formed as

$$\mathrm{MSE}|\hat{\mathbf{H}}_k=\sum_{k=1}^{K}\|\mathbf{a}^T(\hat{\mathbf{H}}_k+\Delta\mathbf{H}_k)\mathbf{b}_k-1\|^2+\sigma_n^2\|\mathbf{a}^T\|^2. \tag{41}$$

To simplify the objective function, we rewrite P6 into

$$\begin{aligned}\mathrm{P7}:\min_{\mathbf{a},\mathbf{b}_k}\max_{\Delta\mathbf{H}_k}\ & \sum_{k=1}^{K}t_k+\sigma_n^2\|\mathbf{a}^T\|^2\\ \text{s.t. } & \|\mathbf{b}_k\|^2\le P_k, k=1,2,...,K,\\ & \|\Delta\mathbf{H}_k\|^2\le\sigma_h^2, k=1,2,...,K,\\ & \|\mathbf{a}^T(\hat{\mathbf{H}}_k+\Delta\mathbf{H}_k)\mathbf{b}_k-1\|^2\le t_k.\end{aligned} \tag{42}$$

To handle the channel uncertainty matrix into an available manner, we introduce new variables via an important tool in robust optimization, which is known as the S-procedure and primarily generalized as follows.

Lemma 1 (S-procedure)**.** Let $\theta_m(\boldsymbol{\epsilon})=\boldsymbol{\epsilon}^T\mathbf{Z}_m\boldsymbol{\epsilon}+\mathbf{z}_m^T\boldsymbol{\epsilon}+\boldsymbol{\epsilon}^T\mathbf{z}_m+\tilde{z}_m$, $m=1,2$ be two quadratic functions in $\boldsymbol{\epsilon}$ and let $\mathbf{Z}_m$ be Hermitian. Suppose there exists $\hat{\boldsymbol{\epsilon}}$ such that $\theta_1(\hat{\boldsymbol{\epsilon}})>0$, then the implication [25] [26]

$$\theta_1(\boldsymbol{\epsilon})\ge 0\Rightarrow\theta_2(\boldsymbol{\epsilon})\ge 0, \tag{43}$$

holds true if and only if there exists $\lambda\ge 0$ such that

$$\begin{bmatrix}\mathbf{Z}_2 & \mathbf{z}_2\\ \mathbf{z}_2{}^T & \tilde{z}_2\end{bmatrix}-\lambda\begin{bmatrix}\mathbf{Z}_1 & \mathbf{z}_1\\ \mathbf{z}_1{}^T & \tilde{z}_1\end{bmatrix}\succeq\mathbf{0}. \tag{44}$$

Let $\theta_1(\boldsymbol{\epsilon})\ge 0$ describe the $\Delta\mathbf{H}_k$ constraint and let $\theta_2(\boldsymbol{\epsilon})\ge 0$ describe MSE constraint specifically, and the channel uncertainty constraint and the MSE constraint can be combined into a second-order cone constraint which is convex. If such $\lambda\ge 0$ exists, the transceiver design can hold the MSE constraint for all points in the channel uncertainty set, even for the worst channel uncertainty condition. Based on it, we convert the channel uncertainty constraint holding into positive variable λ finding.

To solve the challenge of the non-convex objective function over $\mathbf{a}^T,\mathbf{b}_k$, we adopt the iterative algorithm to find the optimal solution.

In conclusion, the P7 is varied as follows.

Proposition 5 (Equivalent Convex Robust Design Problem with Conventional Signaling Procedure)**.** For a fixed receive vector $\mathbf{a}$, P7 is equivalent to P7.1

$$\begin{aligned}\mathrm{P7.1}:\min_{\mathbf{b}_k,\lambda_k}\ & \sum_{k=1}^{K}t_k+\sigma_n^2\|\mathbf{a}^T\|^2\\ \text{s.t. } & \begin{bmatrix}\lambda_k\mathbf{I}_{N_tN_r}-\boldsymbol{\phi}_k\boldsymbol{\phi}_k^T & -\boldsymbol{\phi}_k\boldsymbol{\upsilon}_k^T\\ -\boldsymbol{\upsilon}_k\boldsymbol{\phi}_k^T & t_k-\|\boldsymbol{\upsilon}_k\|^2-\lambda_k\sigma_h^2\end{bmatrix}\succeq\mathbf{0},\\ & \lambda_k\ge 0, k=1,2,...,K.\end{aligned} \tag{45}$$

where

$$\begin{aligned}\boldsymbol{\upsilon}_k&=\mathbf{a}^T\hat{\mathbf{H}}_k\mathbf{b}_k-1,\\ \boldsymbol{\phi}_k^T&=\mathrm{vec}(\mathbf{a}\mathbf{b}_k^T).\end{aligned} \tag{46}$$

For fixed transmit vectors $\mathbf{b}_k$, $k=1,2,...,K$, the optimization of P7 can be formulated as the similar form to P7.1 and defined as P7.2. The main idea of the transceiver design is to find the optimal receive vector $\mathbf{a}$ via fixed transmit vectors $\mathbf{b}_k$, and vise versa with the iteration between P7.1 and P7.2.

Proof: Denote the channel uncertainty constraint and MSE constraint as $\theta_1(\boldsymbol{\zeta}_k)\ge 0$ and $\theta_2(\boldsymbol{\zeta}_k)\ge 0$ respectively,

$$\begin{aligned}\theta_1(\boldsymbol{\zeta}_k)&=-\boldsymbol{\zeta}_k^T\mathbf{I}_{N_tN_r}\boldsymbol{\zeta}_k+\sigma_h^2,\\ \theta_2(\boldsymbol{\zeta}_k)&=-\boldsymbol{\zeta}_k^T\boldsymbol{\phi}_k\boldsymbol{\phi}_k^T\boldsymbol{\zeta}_k-\boldsymbol{\zeta}_k^T\boldsymbol{\phi}_k\boldsymbol{\upsilon}_k^T-\boldsymbol{\upsilon}_k\boldsymbol{\phi}_k^T\boldsymbol{\zeta}_k+t_k-\|\boldsymbol{\upsilon}_k\|^2,\end{aligned} \tag{47}$$

where

$$\boldsymbol{\zeta}_k=\mathrm{vec}(\Delta\mathbf{H}_k), \tag{48}$$

we can obtain the expression of the new optimization constraints via S-procedure in (43) and (44). ■

Thereby, we raise a robust MSE minimization algorithm with iteratively $\mathbf{a}$, $\mathbf{b}_k$ and slack variables λ_k under the worst-case model. The iterative process will stop under a specific condition. We give a brief derivation regarding the computational time complexity of employing S-procedure method.

Proposition 6 (Computational Time Complexity of Conventional Signaling Procedure)**.** The computational time complexity is $\mathcal{O}(K^{\frac{7}{2}})$.

Proof: The constraints in (45) satisfy the real-valued standard SDP form [28]

$$\min_{\mathbf{x}\in\mathbb{R}^n} \left\{ \mathbf{m}^T\mathbf{x} : \mathbf{A}_0 + \sum_{j=1}^{n} x_j \mathbf{A}_j \succeq \mathbf{0} \right\}, \tag{49}$$

where $\mathbf{A}_j$ are symmetric block-diagonal matrices with K diagonal blocks of sizes $a_i \times a_i$, where $a_i = N_t \times N_r + 1$. $n = K + N_t$ represents the number of unknown variables, where the first term of n represents the number of slack variables λ_k and the second corresponds the length of the receive vector. The number of arithmetic operations for the termination of interior point methods that solve this problem is known to be upper bounded by

$$\mathcal{O}\left[\left(1+\sum_{i=1}^{K} a_i\right)^{\frac{1}{2}} n \left(n^2 + n\sum_{i=1}^{K} {a_i}^2 + \sum_{i=1}^{K} {a_i}^3\right)\right]. \tag{50}$$

Based on it, the computational time complexity is about $\mathcal{O}(K^{\frac{7}{2}})$ which is the highest order in (50). ■

We introduce K extra variables to represent the worst channel condition utilizing S-procedure, and adopt the iterative algorithm to transform the optimization into a convex problem. The MSE performance is shown through simulations in Section VI.

D. Worst-Case Robust Design with Over-the-Air Signaling Procedure

In the over-the-air signaling procedure, the effective CSI uncertainty model in (32) is assumed to satisfy $\|\boldsymbol{\Delta}_i\|^2 \le \sigma_h^2$. The optimization of MSE at the FC can be written as

$$\begin{aligned} \text{P8}: \ & \min_{\mathbf{a},\mathbf{b}_k} \max_{\boldsymbol{\Delta}_i} \text{MSE}|\left\langle \hat{\mathbf{F}}, \hat{\mathbf{g}} \right\rangle \\ & \quad \text{s.t. } \operatorname{tr}(\mathbf{b}_k\mathbf{b}_k^T) \le P_k, k = 1,2,...,K, \\ & \quad \quad \|\boldsymbol{\Delta}_i\|^2 \le \sigma_h^2, i = 1,2,...,N_r+1, \end{aligned} \tag{51}$$

where

$$\begin{aligned} \text{MSE}|\left\langle \hat{\mathbf{F}}, \hat{\mathbf{g}} \right\rangle &= \mathbb{E}\left\{\mathbf{a}^T\mathbf{F}\mathbf{a} - \mathbf{a}^T\mathbf{g} - \mathbf{a}\mathbf{g}^T + K + \sigma_n^2\mathbf{a}^T\mathbf{a}\right\} \\ &= \mathbf{a}^T\hat{\mathbf{F}}\mathbf{a} - \mathbf{a}^T\hat{\mathbf{g}} - \mathbf{a}\hat{\mathbf{g}}^T + K + \sigma_n^2\mathbf{a}^T\mathbf{a} \\ &+ \left(\sum_{i=1}^{N_r} \mathbf{a}^T\boldsymbol{\Delta}_i a_i\right) - \mathbf{a}^T\boldsymbol{\Delta}_{N_r+1} - \boldsymbol{\Delta}_{N_r+1}^T\mathbf{a}. \end{aligned} \tag{52}$$

Similar to P7, the challenges to resolve P8 include the uncertain worst channel condition and the non-convex objective function. Based on the analysis above, we can rebuild P8 utilizing S-procedure as follows.

Proposition 7 (Equivalent Convex Robust Design Problem with Over-the-Air Signaling Procedure)**.** For fixed transmit vectors $\mathbf{b}_k$, $k = 1,2,...,K$, P8 is equivalent to

$$\text{P8.1}: \ \min_{\mathbf{a},\lambda_i} \sum_{i=1}^{N_r+1} t_i + \text{MSE}_0 \tag{53a}$$

$$\text{s.t.} \begin{bmatrix} \lambda_{N_r+1}\mathbf{I}_{N_r} & \mathbf{a} \\ \mathbf{a}^T & t_{N_r+1} - \lambda_{N_r+1}\sigma_h^2 \end{bmatrix} \succeq \mathbf{0}, \tag{53b}$$

$$\begin{bmatrix} \lambda_i\mathbf{I}_{N_r} & -\frac{1}{2}a_i\mathbf{a} \\ -\frac{1}{2}a_i\mathbf{a}^T & t_i - \lambda_i\sigma_h^2 \end{bmatrix} \succeq \mathbf{0}, i = 1,2,...,N_r, \tag{53c}$$

$$\lambda_i \ge 0, i = 1,2,...,N_r+1. \tag{53d}$$

where

$$\text{MSE}_0 = \mathbf{a}^T\hat{\mathbf{F}}\mathbf{a} - \mathbf{a}^T\hat{\mathbf{g}} - \mathbf{a}\hat{\mathbf{g}}^T + K + \sigma_n^2\mathbf{a}^T\mathbf{a}. \tag{54}$$

For a fixed receive vector $\mathbf{a}$, P8 can be formulated as

$$\begin{aligned} \text{P8.2}: \ & \min_{\mathbf{b}_k} \text{MSE}_0 \\ & \quad \text{s.t. } \|\mathbf{b}_k\|^2 \le P_k, k = 1,2,...,K, \end{aligned} \tag{55}$$

which is convex over $\mathbf{b}_k$ and the optimal transmit vectors can be defined as

$$\begin{aligned} \mathbf{b}_k &= \left(\mathbf{H}_k^T\mathbf{a}\mathbf{a}^T\mathbf{H}_k + \mu_k\mathbf{I}\right)^{-1}\mathbf{H}_k^T\mathbf{a}, k = 1,2,...,K, \\ & \mu_k\left[\operatorname{tr}(\mathbf{b}_k^T\mathbf{b}_k) - P_k\right] = 0, k = 1,2,...,K, \end{aligned} \tag{56}$$

where $\mu_k \ge 0$.

Proof: For fixed transmit vectors $\mathbf{b}_k$, $k = 1,2,...,K$, the uncertainty estimated channel constraint and the MSE constraint can be expressed as $\theta_1(\boldsymbol{\Delta}_{N_r+1}) \ge 0$ and $\theta_2(\boldsymbol{\Delta}_{N_r+1}) \ge 0$, where

$$\begin{aligned} & -\mathbf{a}^T\boldsymbol{\Delta}_{N_r+1} - \mathbf{a}\boldsymbol{\Delta}_{N_r+1}^T \le t_{N_r+1}, \\ & \theta_1(\boldsymbol{\Delta}_{N_r+1}) = -\boldsymbol{\Delta}_{N_r+1}^T\mathbf{I}_{N_r}\boldsymbol{\Delta}_{N_r+1} + \sigma_h^2, \\ & \theta_2(\boldsymbol{\Delta}_{N_r+1}) = \mathbf{a}^T\boldsymbol{\Delta}_{N_r+1} + \mathbf{a}\boldsymbol{\Delta}_{N_r+1}^T + t_{N_r+1}. \end{aligned} \tag{57}$$

If there exists $\lambda_{N_r+1} \ge 0$ such that $\theta_1(\boldsymbol{\Delta}_{N_r+1}) \ge 0 \Rightarrow \theta_2(\boldsymbol{\Delta}_{N_r+1}) \ge 0$ holds true, we can obtain the SDP constraint in (53b).

For the other uncertainty estimated channel constraints $\|\boldsymbol{\Delta}_i\|^2 \le \sigma_h^2$, $i = 1,2,...,N_r$, the uncertainty channel constraints and the MSE constraints can be expressed as $\theta_1(\boldsymbol{\Delta}_i) \ge 0$ and $\theta_2(\boldsymbol{\Delta}_i) \ge 0$, where

$$
\begin{aligned}
&\mathbf{a}^T\boldsymbol{\Delta}_i a_i \le t_i, i = 1,2,...,N_r,\\
&\theta_1(\boldsymbol{\Delta}_i) = -\boldsymbol{\Delta}_i^T \mathbf{I}_{N_r}\boldsymbol{\Delta}_i + \sigma_h^2,\\
&\theta_2(\boldsymbol{\Delta}_i) = -\frac{1}{2}a_i\mathbf{a}^T\boldsymbol{\Delta}_i - \frac{1}{2}a_i\boldsymbol{\Delta}_i^T\mathbf{a} + t_i.
\end{aligned} \tag{58}
$$

Similarly, the constraints of $\boldsymbol{\Delta}_i$ can be expressed as the SDP constraints in (53c).

Based on the above statement, we can achieve the convex optimization for receive vector under fixed transmit vectors.

If the receive vector $\mathbf{a}$ is fixed, the optimization of $\mathbf{b}_k$ is irrelevant with $\boldsymbol{\Delta}_i$ in P8, which results in that the objective function of the optimization in P8 is equals to MSE_0, and the P8 becomes the convex optimization for fixed receive vector $\mathbf{a}$ in (55). The Lagrange dual objective function can be constructed as

$$
L(\mathbf{b}_k,\mu_k) = \mathrm{MSE}_0 + \sum_{k=1}^{K}\mu_k(\mathbf{b}_k^T\mathbf{b}_k - P_k), \tag{59}
$$

where the Lagrange multiplier μ_k is associated with the power constraint of transmitter k.

According to the KKT conditions given as

$$
\begin{aligned}
\mathrm{tr}\left(\mathbf{b}_k^T\mathbf{b}_k\right) - P_k \le 0, k = 1,2,...,K,\\
\mu_k \ge 0, k = 1,2,...,K,\\
\mu_k\left[\mathrm{tr}(\mathbf{b}_k^T\mathbf{b}_k) - P_k\right] = 0, k = 1,2,...,K,\\
\frac{\partial L}{\partial \mathbf{b}_k} = 0, k = 1,2,...,K,
\end{aligned} \tag{60}
$$

we take the partial derivative of L with the respect to the vector $\mathbf{b}_k$ and obtain optimal transmit vectors in (56). ■

Proposition 8 (Computational Time Complexity of Over-the-Air Signaling Procedure)**.** The computational time complexity of the over-the-air signaling procedure is about $\mathcal{O}(N_r^6)$.

Proof: Due to the fact that the constraints of P8 in (53) are SDP which contain N_r+1 blocks of sizes $a_i = N_r+1$, the complexity is upper bounded by

$$
\mathcal{O}\left[\left(1+\sum_{i=1}^{N_r+1} a_i\right)^{\frac{1}{2}} n\left(n^2 + n\sum_{i=1}^{N_r+1} {a_i}^2 + \sum_{i=1}^{N_r+1} {a_i}^3\right)\right], \tag{61}
$$

The unknown vector to be determined is of size $n = N_r + (N_r+1)$, where the first corresponds the receive vector of length N_r and the second represents the number of slack variables λ_k equals to N_r+1. The highest order in the polynomial is N_r^6. Therefore, the computational time complexity is about $\mathcal{O}(N_r^6)$. The optimization in (56) is simple multiplication, whose computational time complexity can be ignored. ■

Based on (50) and (61), we compare the conventional design and the over-the-air signaling procedure in the following.

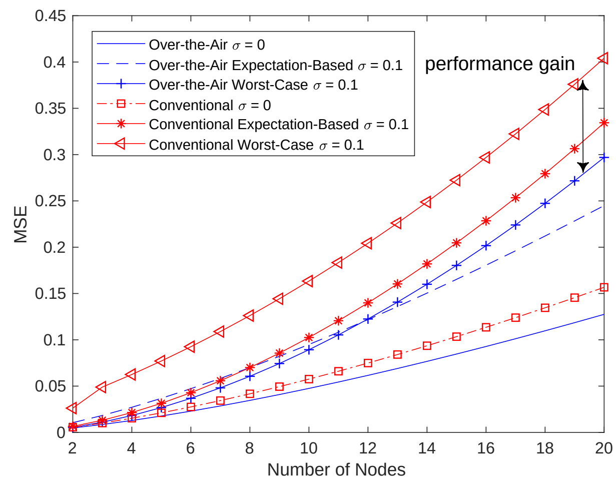


Fig. 4. MSE performance with different number of nodes of 2×2 channel matrix and SNR = 30dB.

Remark 2 (Computational Time Complexity Comparison)**.** The computational time complexity of the conventional scheme is about $\mathcal{O}(K^{\frac{7}{2}})$, which is extremely increasing with massive nodes. Compared with the conventional design, the computational time complexity of the over-the-air signaling procedure is about $\mathcal{O}(N_r^6)$. When the number of nodes K satisfies $K \gg N_r$, the robust design with the over-the-air signaling procedure significantly reduces the computational time complexity.

VI. Simulation Results and Discussion

In this section, we illustrate the performance of the proposed algorithms through simulations. Each transmitter is limited by the same transmit power constraint. The locations of the users satisfy the random distribution. With the normalization of the large-scale fading and the consideration of the small-scale fading, a quasi-static flat Rayleigh fading channel is used, which is modeled as independently and identically distributed (i.i.d) Gaussian random variables with zero mean and unit variance. We set the initialization receive vector as $\mathbf{b} = \mathbf{1}$.

Fig. 4 shows the MSE performance for different numbers of nodes. This agrees with our intuition that more connected nodes makes it harder to design one common receive beamformer to equalize all channels. The MSE performance of the over-the-air signaling procedure is obviously better than the conventional scheme with numerous nodes, which is caused by noise averaging effect. The performance gap between two designed schemes increases with the growth of the number of nodes.

We evaluate the MSE performance as a function of *signal-to-noise ratio* (SNR) in Fig. 5. It can be seen that the performance can be exceedingly enhanced with the increase of SNR. The guarantee of the worst uncertainty constraint leads that the MSE performance under the worst-case model is no more better than the expectation-based robust design. It it caused by the fact that the worst channel uncertainty could have very low probability in practice, and the transceiver result is suboptimal

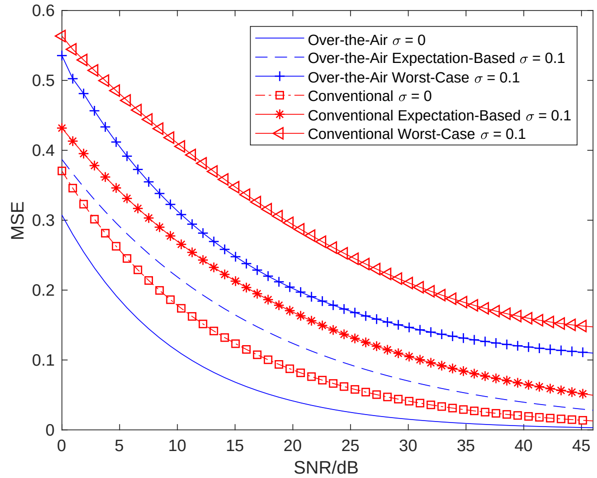


Fig. 5. MSE performance with different SNR of 2×2 channel matrix and the number of nodes = 2.

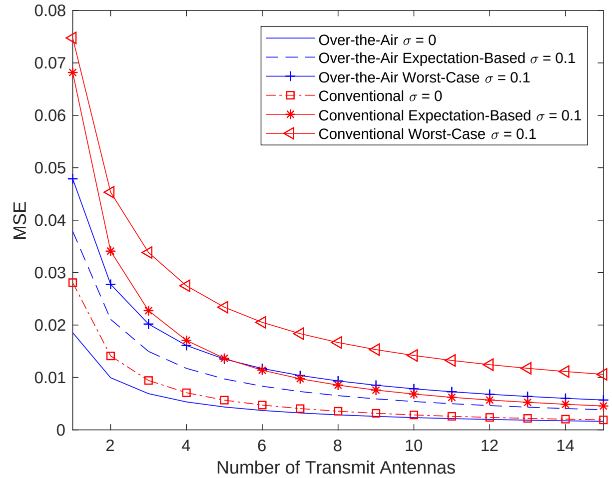


Fig. 7. MSE performance with different number of transmit antenna number of 2 nodes and SNR = 30dB.

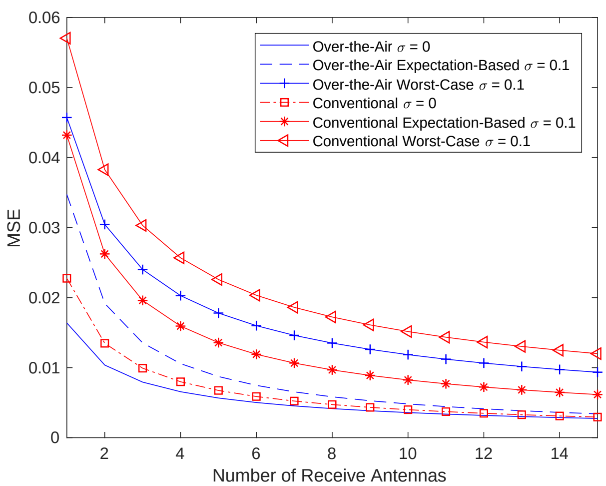


Fig. 6. MSE performance with different number of receive antenna number of 2 nodes and SNR = 30dB.

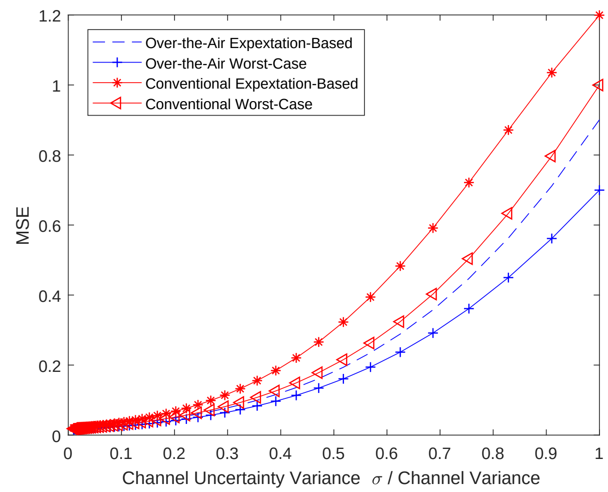


Fig. 8. MSE performance with different channel uncertainty variance of 2 nodes and SNR = 30dB.

for the most conditions. It also verifies the conclusion that the over-the-air signaling procedure outperforms the conventional scheme on the MSE performance. Although the proposed algorithm is not guaranteed to converge to global optimal solutions, the simulations show that it quickly converges and the proposed algorithm is not sensitive to initialization.

Fig. 6 and Fig. 7 illustrate the MSE performance of the different number of receive antennas N_r and number of transmit antennas N_t. The MSE performance increases with not only the number of receive antennas N_r but also the number of transmit antennas N_t. Deploying multi-antennas provide extra channel gain and compensate for the performance degradation via space diversity.

We show the MSE performance with different channel uncertainty variance σ in Fig. 8. With the increases of channel uncertainty variance, the MSE performance of both robust designs are decreased. However, the robust design of the over-the-air signaling procedure performs lower MSE than the conventional design. The performance gap between the conventional design and the proposed method also becomes deep with the increases of uncertain variance, which is caused by noise averaging effect.

The distribution of iterative times is counted for the over-the-air signaling procedure with the perfect CSI in Fig. 9. (a). We count the pseudo-random numbers which satisfy the uniform distribution for comparison. The result shows that the iterative times basically satisfies uniform distribution in [1,15], which is much less than the number of nodes $K = 100$. Fig. 9. (b) shows the comparison of the number of training time slots between the conventional design and the over-the-air signaling procedure. In the over-the-air signaling procedure, we first obtain the iterative times with different number of nodes via simulation results. According to the derived results in **Proposition 2**, we plot the essential training time slots of the over-the-air signaling procedure which is linear increasing with the iterative times. There exists a cross

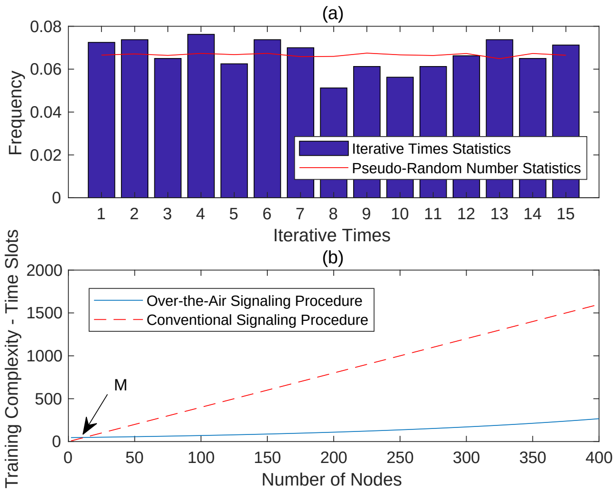


Fig. 9. (a) is the distribution of iterative times for 2×2 channel matrix. Number of nodes = 100 and SNR = 30dB. It is based on the over-the-air signaling procedure under perfect CSI. (b) is the training time slots of the conventional scheme and the over-the-air signaling procedure

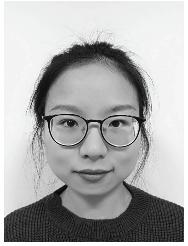
point M for the conventional scheme and the over-the-air signaling procedure, and the proposed signaling procedure extremely reduces training complexity after the cross point. The simulation results support our previous analysis in Section IV that the training complexity of the proposed solution is significantly reduced for the massive CSI acquisition. Thereby, the proposed solution features low complexity and is preferred in the networks with large scale nodes.

VII. Conclusions

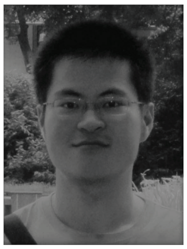
In this paper, we have proposed a robust design for massive CSI acquisition in analog function computation networks. The over-the-air signaling procedure has been proposed to solve the excessive latency problem of the conventional design. The training complexity of the conventional signaling procedure and the proposed one have been developed and further compared. The robust design under the expectation-based model and the worst-case model have also been discussed for both the conventional scheme and the proposed signaling procedure. Moreover, we have provided the computational time complexity analysis. We have derived the conclusion that the training complexity has been significantly reduced for the proposed signaling procedure via the defined effective CSI, and its corresponding transceiver optimization has behaved lower computational time complexity than the conventional design. Simulation results have shown that the transceiver design with the over-the-air signaling procedure has improved MSE performance owing to the noise averaging effect.

References

[1] G. T. V14.2.0, "Study on scenarios and requirements for next generation access technologies," *3rd Generation Partnership Project (3GPP)*, 2017.
[2] O. Abari, H. Rahul, and D. Katabi, "Over-the-air function computation in sensor networks," *arXiv preprint arXiv:1612.02307*, 2016.
[3] B. Nazer and M. Gastpar, "Computation over multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3498–3516, 2007.
[4] ——, "Compute-and-forward: Harnessing interference through structured codes," *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6463–6486, 2011.
[5] J. Zhan, B. Nazer, U. Erez, and M. Gastpar, "Integer-forcing linear receivers," *IEEE Trans. Inf. Theory*, vol. 60, no. 12, pp. 7661–7685, 2014.
[6] M. Goldenbaum, H. Boche, and S. Stańczak, "Harnessing interference for analog function computation in wireless sensor networks," *IEEE Trans. Signal Process*, vol. 61, no. 20, pp. 4893–4906, 2013.
[7] ——, "Nomographic functions: Efficient computation in clustered Gaussian sensor networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 4, pp. 2093–2105, 2015.
[8] R. C. Buck, "Approximate complexity and functional representation," *Journal of Mathematical Analysis & Applications*, vol. 70, no. 1, pp. 280–298, 1976.
[9] F. Wu, L. Chen, N. Zhao, Y. Chen, F. R. Yu, and G. Wei, "Computation over wide-band mac: Improved achievable rate through sub-function allocation," *arXiv:1806.08632*, 2018.
[10] M. Goldenbaum and S. Stańczak, "Robust analog function computation via wireless multiple-access channels," *IEEE Trans. Commun*, vol. 61, no. 9, pp. 3863–3877, 2013.
[11] A. Kortke, M. Goldenbaum, and S. Stańczak, "Analog computation over the wireless channel: A proof of concept," *SENSORS*, pp. 1224–1227, 2014.
[12] M. Goldenbaum, S. Stanczak, and H. Boche, "On achievable rates for analog computing real-valued functions over the wireless channel," *IEEE International Conference on Communications (ICC)*, pp. 4036–4041, 2015.
[13] S. Sigg, P. Jakimovski, and M. Beigl, "Calculation of functions on the rf-channel for iot," *International Conference on Internet of Things (IOT)*, pp. 107–113, 2012.
[14] O. Abari, H. Rahul, D. Katabi, and M. Pant, "Airshare: Distributed coherent transmission made seamless," *IEEE Conference on Computer Communications (INFOCOM)*, pp. 1742–1750, 2015.
[15] 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.
[16] G. Zhu, L. Chen, and K. Huang, "Over-the-air computation in MIMO multi-access channels: Beamforming and channel feedback," *arXiv preprint arXiv:1803.11129*, 2018.
[17] 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.
[18] A. Tajer, N. Prasad, and X. Wang, "Robust linear precoder design for multi-cell downlink transmission," *IEEE Trans. Signal Process*, vol. 59, no. 1, pp. 235–251, 2011.
[19] E. Björnson, G. Zheng, M. Bengtsson, and B. Ottersten, "Robust monotonic optimization framework for multicell miso systems," *IEEE Trans. Signal Process*, vol. 60, no. 5, pp. 2508–2523, 2012.
[20] N. Vucic, H. Boche, and S. Shi, "Robust transceiver optimization in downlink multiuser mimo systems," *IEEE Trans. Signal Process*, vol. 57, no. 9, pp. 3576–3587, 2009.
[21] N. K. Venkategowda, B. B. Narayana, and A. K. Jagannatham, "Precoding for robust decentralized estimation in coherent-mac-based wireless sensor networks," *IEEE Trans. Signal Process. Lett*, vol. 24, no. 2, pp. 240–244, 2017.
[22] J. Huang, Q. Zhang, Q. Li, and J. Qin, "Robust parallel analog function computation via wireless multiple-access mimo channels," *IEEE Trans. Signal Process .Lett*, vol. 22, no. 9, pp. 1297–1301, 2015.
[23] A. B. Gershman, N. D. Sidiropoulos, S. Shahbazpanahi, M. Bengtsson, and B. Ottersten, "Convex optimization-based beamforming," *IEEE Signal Process Mag.*, vol. 27, no. 3, pp. 62–75, 2010.
[24] S. A. Vorobyov, A. B. Gershman, and Z.-Q. Luo, "Robust adaptive beamforming using worst-case performance optimization: A solution to the signal mismatch problem," *IEEE Trans. Signal Process*, vol. 51, no. 2, pp. 313–324, 2003.
[25] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.
[26] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear matrix inequalities in system and control theory*. Siam, 1994, vol. 15.
[27] H. Shen, B. Li, M. Tao, and X. Wang, "Mse-based transceiver designs for the mimo interference channel," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3480–3489, 2010.
[28] A. Ben-Tal and A. Nemirovski, *Lectures on modern convex optimization: analysis, algorithms, and engineering applications*. Siam, 2001, vol. 2.



Fan Ang received the B.S. degree in information and communications engineering from Xi'an Jiaotong University, Xian, China, in 2017. She is currently pursuing the M.S. degree with the Department of Electronic Engineering and Information Science, University of Science and Technology of China. Her research interests include wireless IoT communications and federated learning.



Li Chen received the B.E. in electrical and information engineering from Harbin Institute of Technology, Harbin, China, in 2009 and the Ph.D.degree in electrical engineering from the University of Science and Technology of China, Hefei, China, in 2014. He is currently a faculty member with the Department of Electronic Engineering and Information Science, University of Science and Technology of China. His research interests include wireless IoT communications and wireless optical communications.

Nan Zhao (S'08-M'11-SM'16) is currently an Associate Professor at Dalian University of Technology, China. He received the B.S. degree in electronics and information engineering in 2005, the M.E. degree in signal and information processing in 2007, and the Ph.D. degree in information and communication engineering in 2011, from Harbin Institute of Technology, Harbin, China. His recent research interests include UAV Communications, Interference Alignment, and Physical Layer Security.

Dr. Zhao is serving or served on the editorial boards of 7 SCI-indexed journals. He received Top Reviewer Award from IEEE Transactions on Vehicular Technology in 2016, and was nominated as an Exemplary Reviewer by IEEE Communications Letters in 2016. He won the best paper awards in IEEE VTC'2017-Spring and MLICOM 2017.

Yunfei Chen (S'02-M'06-SM'10) received his B.E. and M.E. degrees in electronics engineering from Shanghai Jiaotong University, Shanghai, P.R.China, in 1998 and 2001, respectively. He received his Ph.D. degree from the University of Alberta in 2006. He is currently working as an Associate Professor at the University of Warwick, U.K. His research interests include wireless communications, cognitive radios, wireless relaying and energy harvesting.

F. Richard Yu (S'00-M'04-SM'08-F'18) received the PhD degree in electrical engineering from the University of British Columbia (UBC) in 2003. From 2002 to 2006, he was with Ericsson (in Lund, Sweden) and a start-up in California, USA. He joined Carleton University in 2007, where he is currently a Professor. He received the IEEE Outstanding Service Award in 2016, IEEE Outstanding Leadership Award in 2013, Carleton Research Achievement Award in 2012. His research interests include wireless cyber-physical systems, connected/autonomous vehicles, security, distributed ledger technology, and deep learning.