WARWICK
THE UNIVERSITY OF WARWICK

Manuscript version: Author's Accepted Manuscript
The version presented in WRAP is the author's accepted manuscript and may differ from the published version or Version of Record.

Persistent WRAP URL:
http://wrap.warwick.ac.uk/139784

How to cite:
Please refer to published version for the most recent bibliographic citation information. If a published version is known of, the repository item page linked to above, will contain details on accessing it.

Copyright and reuse:
The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions.

Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Publisher's statement:
Please refer to the repository item page, publisher's statement section, for further information.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk.

Computation over MAC: Achievable Function Rate Maximization in Wireless Networks

Li Chen, Nan Zhao, *Senior Member, IEEE,* Yunfei Chen, *Senior Member, IEEE,* Xiaowei Qin and F. Richard Yu, *Fellow, IEEE,*

***Abstract*—The next generation wireless network is expected to connect billions of nodes, which brings up the bottleneck on the communication speed for distributed data fusion. To overcome this challenge, computation over multiple access channel (CoMAC) was recently developed to compute the desired functions with a summation structure (e.g., mean, norm, etc.) by using the superposition property of wireless channels. This work aims to maximize the achievable function rate of reliable CoMAC in wireless networks. More specifically, considering channel fading and transceiver design, we derive the achievable function rate adopting the quantization and the nested lattice coding, which is determined by the number of nodes, the maximum value of messages and the quantization error threshold. Based on the derived result, the transceiver design is optimized to maximize the achievable function rate of the network. We first study a single cluster network without inter-cluster interference (ICI). Then, a multi-cluster network is further analyzed in which the clusters work in the same channel with ICI. In order to avoid the global channel state information (CSI) aggregation during the optimization, a low-complexity signaling procedure irrelevant with the number of nodes is proposed utilizing the channel reciprocity and the defined effective CSI.**

***Index Terms*—Computation, interference, multiple access channel, MIMO, signaling procedure, transceiver design**

I. Introduction

Both communication and computation are important for the next generation wireless networks. As the computation speed is growing rapidly, the communication speed of the air interface has become a bottleneck, especially for large-scale networks. For example, the 5G cellular network is predicted to provide an Internet of Things (IoT) that interconnects up to 1 trillion devices, and a million connections per square kilometer [1]. Traditional "communicate-then-compute" principle attempts to avoid the multi-node interference through orthogonal multiple access channel (MAC), which will result in an extensive air interface latency with limited radio resources [2].

In order to overcome this bottleneck, computation over MAC (CoMAC) provides a promising design to integrate communication and computation. It utilizes the superposition property of wireless channel to compute a class of functions with a summation structure (e.g., sum, norm, etc.) directly. CoMAC allows simultaneous transmission of multiple nodes instead of traditional orthogonal access, which dramatically reduces the air interface latency. The study on CoMAC can be tracked back to the pioneer work in information theory. B. Nazer and M. Gastpar first pointed out that it was beneficial to utilize the superposition property of wireless channels to compute some target functions with a similar structure [3].

From the implementation point of view, uncoded analog function computation is the most direct way to realize CoMAC. Uncoded transmission where the node's channel input is merely a scaled version of its noisy observation has been proved to be optimal for a standard Gaussian multiple-access channel in [4]. The analog CoMAC for a generalized network consisting of multiple fusion centers was studied in [5], where the network was divided into several clusters with independent target functions computed. Various experiment platforms have been built to verify the idea of the analog CoMAC in [6], [7].

Furthermore, several practical issues were discussed during the implementation of the analog CoMAC. Considering the imperfect synchronization between different nodes, a robust analog function computation scheme was proposed in [8], where the synchronization error has been transferred into an additive noise through random sequences. Modeling the channel uncertainty with the worst-case model, a robust design problem for parallel analog function computation was formulated and optimized in [9], [10]. In [11], a uniform forcing transceiver design for analog CoMAC was proposed to compensate the non-uniform fading of different nodes. More complicated multiple-input and multiple-output (MIMO) transceiver designs were provided in [12] for the analog CoMAC, which realized multiple functions computed over MAC with antenna arrays. The objective of all these designs is to minimize the computation error measured by mean squared error (MSE) between the target function and the computed one. However, the reliable computation cannot be guaranteed by the analog CoMAC.

In order to achieve reliable CoMAC, digital CoMAC was further developed. Using channel coding to compute the noisy modulo sum was investigated in [13] based on the linear property of nested lattice coding [14]. Inspired by the idea of recovering the linear combination first, various compute-and-forward schemes based on nested lattice coding have been proposed to improve the communication rate [15]–[17], where the integer coefficient of linear combination should be chosen to be close to the channel coefficient [18], [19]. M. Goldenbaum et al. first proposed a unified digital scheme to compute

L. Chen and X. Qin are with Department of Electronic Engineering and Information Science, University of Science and Technology of China. (e-mail: {chenli87, qinxw}@ustc.edu.cn). N. Zhao is with the School of Info. and Commun. Eng., Dalian University of Technology, Dalian, 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).

structured functions over MAC in [20], where the achievable function rate was first provided assuming the uniform fading of different nodes. With non-uniform fading considered, the work in [21] found that the achievable function decreased as the number of nodes increases, and eventually went to zero due to fading MAC. They also proposed an opportunistic CoMAC to achieve a non-vanishing computation rate in [22], where a subset of nodes opportunistically participated in the transmission at each time slot. A wide-band CoMAC based on orthogonal frequency division multiplexing (OFDM) was studied in [23], where functions were transmitted like bit sequences through division and reconstruction.

However, the afore-mentioned works of digital CoMAC do not consider the transceiver optimization to maximize the achievable function rate. This motivates our work to derive the achievable function rate considering the channel fading and the transceiver design, and then maximize the achievable function rate in wireless networks. Actually, MIMO transceiver optimization to maximize the communication rate is a classical topic that has been extensively studied for information-centric networks [24]–[26]. The objective of the existing MIMO transceiver optimization for information-centric networks is maximizing the communication rate [27]. It is often achieved by decoupling the messages of multiple nodes and suppressing the interference between different nodes. In contrast to traditional networks, the digital CoMAC network is function-centric and the objective of MIMO transceiver design is to maximize the achievable function rate through compensating the non-uniform fading and suppressing the receive noise or interference. Thus, the traditional MIMO transceiver design cannot be used for the CoMAC network.

In this work, we study the transceiver optimization for digital CoMAC in wireless networks. Each node quantizes its reading and then encodes the quantized vector. After transmit beamforming, all nodes transmit the message concurrently. The fusion center (FC) receives the estimate of the target function directly. After receive beamforming and channel decoding, the target function can be recovered reliably. The corresponding achievable function rate is derived, which is defined as functions per channel use. Then the MIMO transceiver is optimized to maximize the derived achievable function rate for single cluster network without inter-cluster interference (ICI) and multi-cluster network with ICI. The optimization depends on the global CSI of the network, which incurs a massive CSI aggregation. A low-complexity signaling procedure is further proposed using the channel reciprocity and the defined effective CSI. The main contributions of this work are summarized as follows.

- **The achievable function rate.** As mentioned, the existing works of digital CoMAC do not provide the achievable function rate considering both the channel fading and the transceiver design. This work derives the corresponding achievable function rate, which is determined by the number of nodes, the maximum value of messages and the quantization error threshold.
- **The CoMAC transceiver design.** Based on the derived results, we formulate the problem to maximize the achievable function rate of the CoMAC network under the power constraint at each node. Both single cluster networks and multi-cluster networks are studied based on the definition of the effective noise.
- **A low-complexity signaling procedure.** Because the traditional signaling procedure of the transceiver optimization requires the global CSI of the network, it incurs a high time-complexity. We propose a low-complexity signaling procedure irrelevant with the number of nodes to avoid individual CSI aggregation based on the channel reciprocity and the defined effective CSI.

Table I
SOME COMMON NOMOGRAPHIC FUNCTIONS

Name	φ_k	ψ	f
Arithmetic Mean	$\varphi_k = s_k$	$\psi = \frac{1}{K}(\cdot)$	$f = \frac{1}{K}\sum_{k=1}^{K} s_k$
Weighted Sum	$\varphi_k = \omega_k s_k$	$\psi = (\cdot)$	$f = \sum_{k=1}^{K} \omega_k s_k$
Geometric Mean	$\varphi_k = \log(s_k)$	$\psi = \exp(\cdot)$	$f = \left(\prod_{k=1}^{K} s_k\right)^{\frac{1}{K}}$
Polynomial	$\varphi_k = \omega_k {s_k}^{\beta_k}$	$\psi = (\cdot)$	$f = \sum_{k=1}^{K} \omega_k {s_k}^{\beta_k}$
Euclidean Norm	$\varphi_k = {s_k}^2$	$\psi = (\cdot)^{\frac{1}{2}}$	$f = \sqrt{\sum_{k=1}^{K} {s_k}^2}$

The remainder of the paper is organized as follows. Section II presents the system model. Section III derives the achievable function rate. The transceiver optimization and the corresponding signaling procedure design are studied in Section IV. The study is further extended to multi-cluster networks in Section V. Simulation results are provided in Section VI, followed by concluding remarks in Section VII.

Notation: $\mathcal{R}$ denotes the reals, and $\mathcal{F}_q$ denotes the finite field of size q with the corresponding subset of the integers $\{0, \cdots, q-1\}$. $\bigoplus$ stands for the modulo summation over the finite field. We use boldface lowercase letter to denote column vectors and boldface uppercase letters to denote matrices.

II. SYSTEM MODEL

We consider a wireless network composed of an FC and K nodes indexed by $k \in \{1, 2, \cdots K\}$. The t-th reading of the node k is $s_{k,t}$, where $t \in \{1, \cdots, T\}$. The network is assumed to be function-centric. That is the FC does not care about the individual reading of each node but the target function thereof $f_t\left(s_{k,1}, \cdots, s_{k,T}\right)$. Furthermore, the target function is assumed to be Nomografic function, which is defined as follows

Definition 1. (Nomografic function [20]) The Nomografic function is represented as

$$f_t = \psi\left[\sum_{k=1}^{K} \varphi_k\left(s_{k,t}\right)\right], \tag{1}$$

where $\psi(\cdot)$ is the post-processing function of the FC, and $\varphi_k(\cdot)$ is the pre-processing function of the node k. Some common Nomographic functions are listed in Table I.

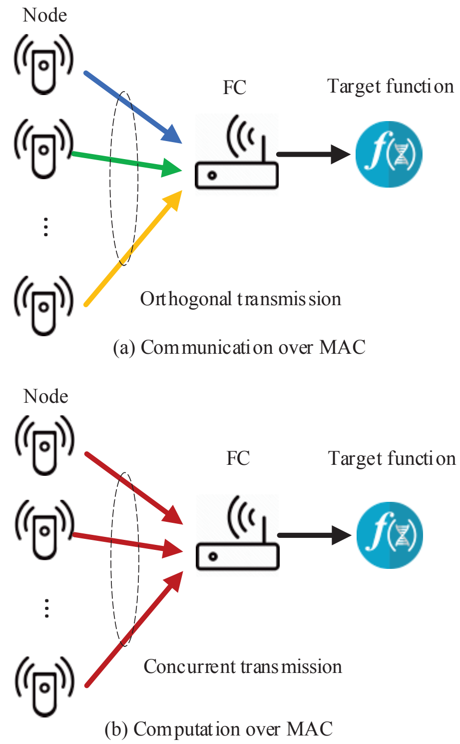


Figure 1. Communication over MAC versus Computation over MAC.

Due to the summation structure of the target function, there are two schemes to recover it at the FC.

- **Communication over MAC**: As illustrated in Fig. 1(a), communication over MAC is a communication and computation separated scheme, where the FC first recovers the individual reading $\{s_{k,t}\}$ of all nodes through orthogonal MAC, and then computes the target function f_t thereof. In order to avoid inter-node interference, orthogonal multiple access, e.g., TDMA, should be adopted, which incurs a high latency.
- **Computation over MAC**: As illustrated in Fig. 1(b), computation over MAC is a communication and computation integrated scheme, which utilizes the superposition property of the wireless channel to receive the summation part of the target function $\left\{\sum_{k=1}^{K}\varphi_k\left(s_{k,t}\right)\right\}$ with all nodes' concurrent transmission. It harnesses the inter-node interference rather than avoiding it, which is promising for a large network.

In order to achieve reliable computation, data processing flow including quantization and channel coding is required at both the node k and the FC.

For the node k, the data processing flow is given as

$$\{s_{k,t}\} \to \{\varphi_k\left(s_{k,t}\right)\} \to \{\mathbf{v}_{k,t}\} \to \mathbf{q}_k \to \mathbf{x}_k{}^{1}, \tag{2}$$

where the reading $s_{k,t}$ is first processed by the pre-processing function $\varphi_k\left(\cdot\right)$, the pre-processed reading $\varphi_k\left(s_{k,t}\right)$ is then quantized into a length v binary vector $\mathbf{v}_{k,t}$, $\{\mathbf{v}_{k,t}\}$ is then mapped into a length m q-ary vector $\mathbf{q}_k \in \mathcal{F}_q^m$, and the quantized vector $\mathbf{q}_k$ is encoded into a length M dithered nested lattice codeword $\mathbf{x}_k \in \mathcal{R}^M$ finally.

After transmit beamforming, all nodes concurrently transmit with symbol-level synchronization assumed. Then, the received signal after receive beamforming of the FC is

$$\begin{aligned}\mathbf{u} &= \mathbf{a}^T\sum_{k=1}^{K}\mathbf{H}_k\mathbf{b}_k\mathbf{x}_k + \mathbf{a}^T\mathbf{n}\\ &= \sum_{k=1}^{K}\mathbf{x}_k + \underbrace{\sum_{k=1}^{K}\left(\mathbf{a}^T\mathbf{H}_k\mathbf{b}_k - 1\right)\mathbf{x}_k + \mathbf{a}^T\mathbf{n}}_{\text{effective noise}},\end{aligned} \tag{3}$$

where $\mathbf{a} \in \mathcal{R}^{N_r}$ is the receive beamforming vector of the FC, $\mathbf{b}_k \in \mathcal{R}^{N_t}$ is the transmit beamforming vector of the node k, $\mathbf{H}_k \in R^{N_r \times N_t}$ is the channel matrix between the node k and the FC, which is assume to be block-fading and keeps constant during the transmission of $\mathbf{x}_k$, and $\mathbf{n} \in R^{N_r}$ is the noise vector with each element distributed as $\mathcal{N}\left(0, \sigma_n^2\right)$. It can be seen that the effective noise of CoMAC is composed of the distortion caused by the non-uniform fading and the received noise, which is different from the traditional transmission noise.

For the FC, the data processing flow is given as

$$\mathbf{u}^{2} \to \sum_{k=1}^{K}\mathbf{q}_k \to \left\{\sum_{k=1}^{K}\mathbf{v}_{k,t}\right\} \to \left\{\sum_{k=1}^{K}\varphi_k\left(s_{k,t}\right)\right\} \to \{f_t\}, \tag{4}$$

where the received signal $\mathbf{u}$ is first decoded into the sum of the quantized vector $\sum_{k=1}^{K}\mathbf{q}_k$ after removing the dithers, $\left\{\sum_{k=1}^{K}\mathbf{v}_{k,t}\right\}$ and $\left\{\sum_{k=1}^{K}\varphi_k\left(s_{k,t}\right)\right\}$ are then recovered, and the target functions $\{f_t\}$ are computed by the post-processing function $\psi(\cdot)$.

In the following discussion, we adopt the achievable function rate as the performance metric, which specifies how many functions can be computed per channel use with predefined quantization error and decoding error, i.e.,

Definition 2. (Achievable function rate) Let Δ be quantization error threshold of $\varphi_k\left(s_k\right) \to \mathbf{q}_k$, ε decoding accuracy of $\mathbf{u} \to \left\{\sum_{k=1}^{K}\mathbf{v}_{k,t}\right\}$. Then R_C is said to be an achievable function rate with fixed quantization error $\Delta > 0$, if for every rate $R' = T/M \leq R_C$ and every $\delta > 0$, the decoding error satisfies $\Pr\left(\cup_{t=1}^{T}\{|\hat{\mathbf{g}}_t - \mathbf{g}_t| > \varepsilon\}\right) < \delta$ with sufficiently large channel uses M, where $\mathbf{g}_t = \sum_{k=1}^{K}\mathbf{v}_{k,t}$ and $\hat{\mathbf{g}}_t$ is its corresponding estimate at the FC.

III. Achievable Function Rate

In this section, we first derive the achievable function rate as the performance metric, where both the channel fading and the transceiver design are considered.

We adopt nested lattice coding as channel coding, and the most attractive property of nested lattice coding is its linear property [28]. That is

[1]To makes the transmitted codewords independent from the underlying lattice points, the nodes have to dither their lattice points using common randomness.

[2]At the FC, the received signal after receive beamforming removes the dithers, and then is mapped to the estimate of the sum.

$$\sum\nolimits_{k=1}^{K} \mathbf{x}_k \mod \Lambda_c \in \mathcal{L}^M, \tag{5}$$

where $\mathcal{L}^M$ is an M-dimensional nested lattice codebook, and a summation of nested lattice codewords $\mathbf{x}_k \in \mathcal{L}^M$ modulo the coarse lattice Λ_c only takes values on the codebook $\mathcal{L}^M$. Thus, the nested lattice coding can be adopted to protect a modulo-q sum of the quantized vectors, i.e., $\mathbf{v} = \bigoplus_{k=1}^{K} \mathbf{q_k}$.

Lemma 1. (The achievable rate to decode modulo-q sum) Given the received signal of the FC in (3), the modulo-q sum is decoded as $\mathcal{D}(\mathbf{u})$. The corresponding decoding error is ε. With sufficiently large number of channel uses M, the following rate is achievable for arbitrary $\varepsilon > 0$ with nested lattice coding

$$R_{\mathcal{L}} \leq \frac{1}{2} \log_2{}^{+} \left(\frac{P_0}{\sum\limits_{k=1}^{K} \left\| \mathbf{a}^T \mathbf{H}_k \mathbf{b}_k - 1 \right\|^2 P_0 + \left\| \mathbf{a} \right\|^2 \sigma_n^2} \right), \tag{6}$$

where $\log_2{}^{+}(\cdot) = \max\{\log_2(\cdot), 0\}$, P_0 is the average transmit power constraint, and

$$\Gamma = \sum_{k=1}^{K} \left\| \mathbf{a}^T \mathbf{H}_k \mathbf{b}_k - 1 \right\|^2 P_0 + \left\| \mathbf{a} \right\|^2 \sigma_n^2 \tag{7}$$

is the MSE of the effective noise of CoMAC.

Proof. The proof follows from Theorem 1 in [13] and Theorem 2 in [29], which provide the achievable rate to decode modulo-q sum for single antenna MAC and multi-antenna MAC. □

For compute-and-forward scheme, the target function of the relay is the modulo sum of transmitted messages with arbitrary coefficients, i.e. $\mathbf{v} = \oplus_{k=1}^{K} a_k \mathbf{q_k}$, where $\{a_k\}$ are coefficients and $\{\mathbf{q_k}\}$ are transmitted messages. And the corresponding wrapping around is inevitable. However, the goal of CoMAC is to recover the function in (1). The desired function of the FC is the sum of transmitted messages with uniform coefficients, i.e., $\mathbf{v} = \sum_{k=1}^{K} \mathbf{q}_k$. Further, in order to compute $\mathbf{v} = \sum_{k=1}^{K} \mathbf{q}_k$, we adopt the nested lattice coding, which has been adopted to compute $\mathbf{v} = \oplus_{k=1}^{K} a_k \mathbf{q_k}$ in compute-and-forward scheme. Thus, the challenge of CoMAC is that the message vector $\mathbf{q}_k$ should be carefully designed to avoid the wrapping around of modulo sum. Thus, the length of quantization vector $\mathbf{q}_k$ should satisfy the following lemma.

Lemma 2. (The length of quantization vector) Without loss of generality, we assume that $\varphi_k(s_k) \in (0, \varphi_{\max})$. The quantization process is $\{\varphi_k(s_{k,t})\} \rightarrow \{\mathbf{v}_{k,t}\} \rightarrow \mathbf{q}_k$. The length of $\mathbf{q}_k$ should satisfy

$$m \geq \frac{T\left(\log_2 K + \log_2 \varphi_{\max} - \log_2 \Delta\right)}{\log_2 p}. \tag{8}$$

Proof. First, we consider the process of $\{\varphi_k(s_{k,t})\} \rightarrow \{\mathbf{v}_{k,t}\}$. The most significant bit of $\mathbf{v}_{k,t}$ v_m should satisfy $2^{v_m} \geq \varphi_{\max}$, and the least significant bit of $\mathbf{v}_{k,t}$ v_l should satisfy $2^{-v_l} \leq \Delta$. Thus, the length of $\mathbf{v}_{k,t}$ satisfies

$$v = v_m + v_l \geq \log_2 \varphi_{\max} - \log_2 \Delta. \tag{9}$$

Then, we consider the process of $\{\mathbf{v}_{k,t}\} \rightarrow \mathbf{q}_k$. The mapping process can be expressed as

$$\mathbf{q}_k = \left(\sum_{t=1}^{\tau} \mathbf{v}_{k,t} q^{t-1}, \cdots, \sum_{t=1}^{\tau} \mathbf{v}_{k,t+(m-1)\tau} q^{t-1} \right), \tag{10}$$

where $m\tau = T$. In order to avoid wrapping around during the modulo sum, one has

$$\sum_{k=1}^{K} \mathbf{v}_{k,t} \leq K \frac{\varphi_{\max}}{\Delta} \leq q, \tag{11}$$

and then

$$\sum_{k=1}^{K} \sum_{t=1}^{\tau} \mathbf{v}_{k,t} q^{t-1} = \sum_{t=1}^{\tau} \sum_{k=1}^{K} \mathbf{v}_{k,t} q^{t-1} \leq q^{\tau} - 1. \tag{12}$$

Thus, for fixed q, the alphabet size of nested lattice coding p should satisfy $q^{\tau} - 1 \leq p - 1$, and we have

$$\tau \leq \frac{\log_2 p}{\log_2 q} \leq \frac{\log_2 p}{\log_2 K + \log_2 \varphi_{\max} - \log_2 \Delta} \tag{13}$$

Because $m\tau = T$, the length of $\mathbf{q}_k$ should satisfy (8), which completes the proof. □

According to Lemma 1 and Lemma 2, the achievable function rate of CoMAC considering the channel fading and the transceiver design is given as follows.

Proposition 1. (The achievable function rate) The achievable function rate of CoMAC (functions per channel use) in Definition 2 can be given by

$$R_C \leq \frac{1}{2\left[\log_2(K) + \log_2(\varphi_{\max}) - \log_2(\Delta)\right]} \log_2{}^{+}\left(\frac{P_0}{\Gamma}\right), \tag{14}$$

where Γ is given in (7).

Proof. Given the alphabet size p, the message rate of nested lattice coding is

$$R_{\mathcal{L}} = \frac{m}{M} \log_2 p \overset{(a)}{\geq} \frac{T\left(\log_2 K + \log_2 \varphi_{\max} - \log_2 \Delta\right)}{M}, \tag{15}$$

where the procedure (a) is due to Lemma 2. Thus, one has

$$\frac{T}{M} < \frac{R_{\mathcal{L}}}{\log_2 K + \log_2 \varphi_{\max} - \log_2 \Delta}. \tag{16}$$

According to $R_{\mathcal{L}}$ given jn Lemma 1 and R_C defined in Definition 2, it completes the proof. □

IV. Transceiver Design and Signaling Procedure

In this section, we discuss the transceiver design to maximize the achievable function rate. Furthermore, a low-complexity signaling procedure is provided to avoid massive CSI aggregation.

A. Transceiver design

Since the power constraints of the transmit beamforming vector for each node satisfy $\|\mathbf{b}_k\| \leq 1, \forall k$, the transceiver optimization problem to maximize the achievable function rate can thus be formulated as

$$\text{(P1)} \quad \begin{array}{ll} \max\limits_{\mathbf{a},\{\mathbf{b}_k\}} & R_C \\ \text{s.t.} & \|\mathbf{b}_k\|^2 \leq 1, \forall k \end{array} \tag{17}$$

According to the expression of R_C given in Proposition 1, P1 can be further transformed into the equivalent problem to minimize the MSE of the effective noise of CoMAC. That is

$$\text{(P2)} \quad \begin{array}{ll} \min\limits_{\mathbf{a},\{\mathbf{b}_k\}} & \Gamma \\ \text{s.t.} & \|\mathbf{b}_k\|^2 \leq 1, \forall k, \end{array} . \tag{18}$$

where Γ is given in (7). MSE measures the distortion between the received signal and the target signal. Minimizing the MSE of analog CoMAC is essentially different from maximizing the achievable function rate from the physical point of view. Because the MSE measures the distortion between the received signal and the target signal in an analog way, but the achievable function rate measures how many reliable functions can be computed per channel use in a digital way. We can utilize the equivalence of minimizing the MSE and maximizing the achievable function rate from the mathematical point of view. Then the Lagrangian dual objective of the problem P2 is

$$\mathcal{L}_1 = \sum_{k=1}^{K} \left\|\mathbf{a}^T \mathbf{H}_k \mathbf{b}_k - 1\right\|^2 P_0 + \|\mathbf{a}\|^2 \sigma_n^2 + \sum_{k=1}^{K} \lambda_k \left(\|\mathbf{b}_k\|^2 - 1\right), \tag{19}$$

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

Taking the partial derivative of $\mathcal{L}_1$ with respect to $\mathbf{a}$, the optimal receive beamforming vector can be obtained as

$$\begin{aligned} \mathbf{a} &= \left(\sum_{k=1}^{K} \mathbf{H}_k \mathbf{b}_k {\mathbf{b}_k}^T {\mathbf{H}_k}^T + \gamma^{-1}\mathbf{I}\right)^{-1} \sum_{k=1}^{K} \mathbf{H}_k \mathbf{b}_k \\ &= \left(\mathbf{H}\mathbf{B}\mathbf{B}^T\mathbf{H}^T + \gamma^{-1}\mathbf{I}\right)^{-1} \mathbf{H}\mathbf{B}\mathbf{1}^T \end{aligned}, \tag{20}$$

where $\gamma = P_0/\sigma_n^2$ is the transmit SNR, $\mathbf{H} = [\mathbf{H}_1, \mathbf{H}_2, \cdots, \mathbf{H}_K]$, $\mathbf{B} = \text{diag}\{\mathbf{b}_1, \mathbf{b}_2, \cdots, \mathbf{b}_K\}$, and $\mathbf{1}$ is a length K vector with all elements being 1. The above obtained receiver is also the MMSE receiver of the effective noise, and the corresponding MSE is

$$\begin{aligned} \Gamma &= \left(\mathbf{a}^T\mathbf{H}\mathbf{B}\mathbf{B}^T\mathbf{H}^T\mathbf{a} - 2\mathbf{a}^T\mathbf{H}\mathbf{B}\mathbf{1} + K\right) P_0 + \mathbf{a}^T\mathbf{a}\sigma_n^2 \\ &= KP_0 - \mathbf{1}\mathbf{B}^T\mathbf{H}^T\left(\mathbf{H}\mathbf{B}\mathbf{B}^T\mathbf{H}^T + \gamma^{-1}\mathbf{I}\right)^{-1}\mathbf{H}\mathbf{B}\mathbf{1}^T P_0, \\ &\overset{(a)}{=} P_0\mathbf{1}\left(\mathbf{I} + \gamma\mathbf{B}^T\mathbf{H}^T\mathbf{H}\mathbf{B}\right)^{-1}\mathbf{1}^T \end{aligned} \tag{21}$$

where the procedure (a) is because the matrix inversion lemma $\left(\mathbf{M}_1 + \mathbf{M}_2\mathbf{M}_3\mathbf{M}_4\right)^{-1} = {\mathbf{M}_1}^{-1} - {\mathbf{M}_1}^{-1}\mathbf{M}_2\left({\mathbf{M}_3}^{-1} + \mathbf{M}_4{\mathbf{M}_1}^{-1}\mathbf{M}_2\right)^{-1}\mathbf{M}_4{\mathbf{M}_1}^{-1}$.

Algorithm 1 Traditional transceiver optimization of CoMAC

Initialize $\mathbf{b}_k^{(0)}$, the convergence threshold ε, and $i = 0$.
The FC estimates the CSI $\mathbf{H}_k$ of all nodes.
The FC computes the corresponding MSE $\Gamma^{(0)}$ according to (21).
repeat
  The FC optimizes the receive beamforming vector $\mathbf{a}^{(i+1)}$ according to (20).
  The FC optimizes the transmit beamforming vector $\mathbf{b}_k^{(i+1)}$ according to (22).
  The FC computes the corresponding MSE $\Gamma^{(i+1)}$ according to (21).
until $\left|\Gamma^{(i+1)} - \Gamma^{(i)}\right| \leq \varepsilon$
The FC transmits the optimized $\mathbf{b}_k^{(i+1)}$ to each node in turn.

Taking the partial derivative of $\mathcal{L}_1$ with respect to $\mathbf{b}_k$, the optimal transmit beamforming vector can be obtained as

$$\mathbf{b}_k = \left({\mathbf{H}_k}^T\mathbf{a}\mathbf{a}^T\mathbf{H}_k + \lambda_k\mathbf{I}\right)^{-1}{\mathbf{H}_k}^T\mathbf{a}, \tag{22}$$

where λ_k can be found with the one dimensional numerical search as

$$\begin{aligned} \|\mathbf{b}_k\|^2 &= \text{Tr}\left[\left({\mathbf{H}_k}^T\mathbf{a}\mathbf{a}^T\mathbf{H}_k + \lambda_k\mathbf{I}\right)^{-2}{\mathbf{H}_k}^T\mathbf{a}\mathbf{a}^T\mathbf{H}_k\right] \\ &= \text{Tr}\left[\left(\mathbf{\Sigma}_k + \lambda_k\mathbf{I}\right)^{-2}{\mathbf{U}_k}^H{\mathbf{H}_k}^T\mathbf{a}\mathbf{a}^T\mathbf{H}_k\mathbf{U}_k\right] \\ &= \sum_{i=1}^{N_t}\frac{\mathbf{\Lambda}_k(i,i)}{\left[\mathbf{\Sigma}_k(i,i) + \lambda_k\right]^2} \leq 1 \end{aligned}, \tag{23}$$

where $\text{SVD}\left({\mathbf{H}_k}^T\mathbf{a}\mathbf{a}^T\mathbf{H}_k\right) = \mathbf{U}_k\mathbf{\Sigma}_k{\mathbf{U}_k}^H$, and $\mathbf{\Lambda}_k = {\mathbf{U}_k}^H{\mathbf{H}_k}^T\mathbf{a}\mathbf{a}^T\mathbf{H}_k\mathbf{U}_k$.

Since the problem P2 is convex over $\mathbf{a}$ or $\mathbf{b}_k$ but not jointly, a standard technique to solve the problem is to use alternating optimization. Since iterations of $\mathbf{a}$ and $\mathbf{b}_k$ optimization can decrease the MSE relative to that of the last iteration and the MSE is lower bounded, the convergence of the iterations is guaranteed. Although the global convergence is not guaranteed in general, it ensures local convergence and often leads to a good suboptimal design with proper initial values.

B. Low-complexity signaling procedure

The traditional transceiver optimization algorithm is given in Algorithm 1, where the CSI of all nodes is aggregated at the FC and then the optimized $\mathbf{b}_k$ is transmitted back to each node after alternating optimization.

The detailed signaling procedure of Algorithm 1 is illustrated in Fig. 2(a). The advantage of CoMAC is to avoid the distributed data aggregation by computing the desired function directly. However, the transceiver optimization above requires the global CSI aggregated at the FC. This will incur a great signaling cost and outweigh the advantage of CoMAC. The corresponding time complexity of Algorithm 1 is provided as follows.

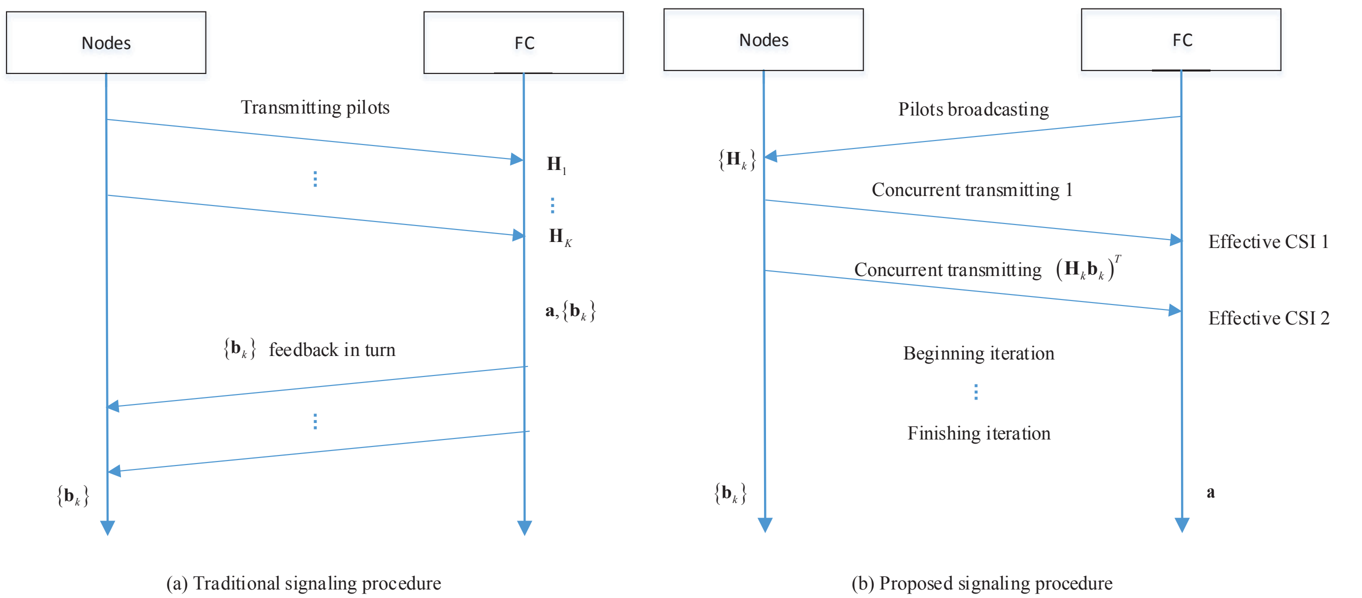


Figure 2. The signaling procedure of transceiver optimization of CoMAC.

Proposition 2. (The time complexity of the traditional signaling procedure) Without considering the time consumption of the optimization, the time complexity of the traditional signaling procedure of CoMAC requires $2N_tK$ symbol slots.

Proof. Each node transmits a pilot matrix to the FC in turn, and the pilot matrix of each node should be no smaller than an $N_t \times N_t$ matrix. Thus, it takes at least $N_t \times K$ symbol slots to complete the channel training process. After alternating optimization at the FC, the optimized $\mathbf{b}_k$ is transmitted back to each node in turn, which is an N_t-length vector. Thus, it takes at least $N_t \times K$ symbol slots to complete the feedback process, which completes the proof. □

One sees that the time complexity of the traditional signaling procedure linearly increases with the number of devices K. In order to avoid massive CSI aggregation, we propose a low-complexity signaling procedure based on channel reciprocity and the defined effective CSI.

The optimized receiver $\mathbf{a}$ in (20) can be regarded as a function of the CSI $\mathbf{H}_k$. Instead of collecting the global CSI of the network, it only requires the effective CSI defined as follows.

$$\text{effective CSI 1: } \sum_{k=1}^{K} \mathbf{H}_k \mathbf{b}_k, \tag{24}$$

which can be obtained by the FC when all nodes concurrently transmit 1 with beamforming vector $\mathbf{b}_k$.

$$\text{effective CSI 2: } \sum_{k=1}^{K} \mathbf{H}_k \mathbf{b}_k {\mathbf{b}_k}^T {\mathbf{H}_k}^T, \tag{25}$$

which can be obtained by the FC when all nodes concurrently transmit $(\mathbf{H}_k\mathbf{b}_k)^T$ with beamforming vector $\mathbf{b}_k$.

Also, the optimized transmitter of the node k in (22) can be regarded as a function of its own CSI $\mathbf{H}_k$. We can use channel reciprocity to allow each node to measure its channel to the FC with very low overhead.

Algorithm 2 Proposed transceiver optimization of CoMAC

Initialize $\mathbf{b}_k^{(0)}$, the convergence threshold ε, and $i = 0$.
Each device estimates its own CSI $\mathbf{H}_k$ based on the broadcasting pilots.
The FC computes the corresponding MSE $\Gamma^{(0)}$ according to (21).
repeat
  The FC estimates the effective CSI 1 in (24).
  The FC estimates the effective CSI 2 in (25).
  The FC optimizes $\mathbf{a}^{(i+1)}$ according to (20).
  The FC broadcasts $\mathbf{a}^{(i+1)}$ to all nodes.
  The FC computes the corresponding MSE $\Gamma^{(i+1)}$ according to (21).
  The node k optimizes the transmit beamforming vector $\mathbf{b}_k^{(i+1)}$ according to (22).
until $\left|\Gamma^{(i+1)} - \Gamma^{(i)}\right| \le \varepsilon$

Thus, we can propose Algorithm 2 that has a low-complexity signaling procedure, where each node requires its own CSI and the FC requires the effective CSI in (24) and (25). The transmitter $\mathbf{b}_k$ is optimized at the node k and receiver $\mathbf{a}$ is optimized at the FC in an iterative way. The detailed signaling procedure of Algorithm 2 is illustrated in Fig. 2(b), where the time complexity is given as follows.

Proposition 3. (The time complexity of the proposed signaling procedure) Without considering the time consumption of the optimization, the time complexity of the proposed signaling procedure requires $N_r + (1 + 2N_r) N_{iter}$ symbol slots, where N_{iter} is the number of optimization iterations.

Proof. The pilot matrix broadcasted to each node for estimating their own CSI should be no smaller than $N_r \times N_r$ matrix.

Thus, it takes at least N_r symbol slots. For each iteration, it takes 1 symbol slot for the FC to estimate the effective CSI 1 in (24) and N_r symbol slots to estimate the effective CSI 2 in (25). After $\mathbf{a}$ is optimized at the FC, it takes N_r symbol slots to broadcast the N_r-length $\mathbf{a}$ to all devices. Assuming the number of iterations is N_{iter}, it takes $N_r + (1+2N_r)\,N_{iter}$ symbol slots, which completes the proof. $\square$

Remark 1. (Comparison between the traditional signaling procedure and the proposed one) In order to illustrate the low complexity of the proposed signaling procedure, we consider a typical dense network composed of $K = 100$ nodes. The node and the FC has $N_t = N_r = 2$ antennas. The symbol slots required for the traditional signaling procedure is $2N_tK = 400$. According to the simulation results in Section VI, the average number of iteration is about $N_{iter} = 10$. The symbol slots required for the proposed signaling procedure is $N_r + (1+2N_r)\,N_{iter} = 52$, which is 13% of the time complexity of the traditional signaling procedure.

Intuitively, the complexity of the proposed signaling procedure decreases due to two reasons. The one is the use of the broadcasting property and the channel reciprocity property of the wireless channel to estimate each node's own CSI. The other is the use of the superposition property of the wireless channel to obtain the effective CSI with a summation structure.

C. Robust design with imperfect CSI

Further, we present the transceiver design for handling CSI uncertainty in a robust manner with the traditional signaling procedure and the proposed signaling procedure. We only consider the CSI error at the FC, and the signaling and the pilots broadcast from FC are noiseless.

In the traditional signaling procedure, CSI is represented by the individual estimated channel for each node, i.e.,

$$\mathbf{H}_k = \widehat{\mathbf{H}}_k + \Delta\mathbf{H}_k, \forall k \tag{26}$$

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.

Thus, the MSE of effective noise of CoMAC in the Problem P2 can be rewritten as

$$(\text{P2.1}) \quad \begin{array}{ll} \min\limits_{\mathbf{a},\{\mathbf{b}_k\}} & \Gamma|\widehat{\mathbf{H}}_k \\ \text{s.t.} & \|\mathbf{b}_k\|^2 \le 1, \forall k, \end{array} . \tag{27}$$

where

$$\begin{aligned} \Gamma|\widehat{\mathbf{H}}_k =& \sum_{k=1}^{K} \|\mathbf{a}^T\hat{\mathbf{H}}_k\mathbf{b}_k - 1\|^2 + \sigma_n^2\|\mathbf{a}^T\|^2 \\ &+ \sigma_h^2\|\mathbf{a}^T\|^2 \sum_{k=1}^{K}\|\mathbf{b}_k\|^2, \end{aligned} \tag{28}$$

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 $\mathbb{E}\left\{\Delta\mathbf{H}_k\cdot\Delta\mathbf{H}_j^T\right\} = \mathbf{0}_{N_r\times N_r}$, $k\neq j$.

Taking the partial derivation of the Lagrangian dual objective of the Problem P2.1, the optimal transceiver can be given by

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

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

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

In the proposed signaling procedure, the estimated uncertainty of the effective CSI can be expressed as

$$\mathbf{g} = \sum_{k=1}^{K}\mathbf{H}_k\mathbf{b}_k + \Delta\mathbf{g} = \widehat{\mathbf{g}} + \Delta\mathbf{g}, \tag{31}$$

and

$$\mathbf{F} = \sum_{k=1}^{K}\mathbf{H}_k\mathbf{b}_k\mathbf{b}_k^T\mathbf{H}_k^T + \Delta\mathbf{F} = \widehat{\mathbf{F}} + \Delta\mathbf{F}, \tag{32}$$

where $\hat{\mathbf{F}}$ and $\hat{\mathbf{g}}$ are the nominal effective CSI available at the FC, and $\Delta\mathbf{F}$, $\Delta\mathbf{g}$ represent the estimated uncertainty at the FC.

Thus, the MSE of effective noise of CoMAC in the Problem P2 can be rewritten as

$$(\text{P2.2}) \quad \begin{array}{ll} \min\limits_{\mathbf{a},\{\mathbf{b}_k\}} & \Gamma|\widehat{\mathbf{g}}, \widehat{\mathbf{F}} \\ \text{s.t.} & \|\mathbf{b}_k\|^2 \le 1, \forall k, \end{array} . \tag{33}$$

where

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

and $\Delta\mathbf{F}_i$ is the i-th column vector of matrix $\Delta\mathbf{F}$, a_i is the i-th element of vector $\mathbf{a}$, $\mathbb{E}\{\Delta\mathbf{F}_i\} = \mathbb{E}\{\Delta\mathbf{g}\} = \mathbf{0}_{N_r\times 1}$, $\mathbb{E}\left\{\Delta\mathbf{F}_i\cdot\Delta\mathbf{F}_i^T\right\} = \mathbb{E}\left\{\Delta\mathbf{g}\cdot\Delta\mathbf{g}^T\right\} = \sigma_h^2\mathbf{I}_{N_r}$, and $\mathbb{E}\left\{\Delta\mathbf{F}_i\cdot\Delta\mathbf{F}_j^T\right\} = \mathbf{0}_{N_r\times N_r}, i\neq j$. So that $\Gamma|\widehat{\mathbf{g}}, \widehat{\mathbf{F}}$ can be calculated as

$$\Gamma|\widehat{\mathbf{g}}, \widehat{\mathbf{F}} = \mathbf{a}^T\hat{\mathbf{F}}\mathbf{a} - \mathbf{a}^T\hat{\mathbf{g}} - \mathbf{a}\hat{\mathbf{g}}^T + K + \sigma_n^2\mathbf{a}^T\mathbf{a}, \tag{35}$$

which is the same with that of the Problem P2, and the transceiver design with perfect CSI in (20) and (22) is applicable to that with imperfect CSI for the proposed signaling procedure.

V. Discussion for Multi-Cluster Networks

In this section, we extend our study to multi-cluster networks, which are more complicated due to ICI. Both transceiver design to maximize the sum achievable function rate and a low-complexity signaling procedure are discussed.

As illustrated in Fig. 3, the network is composed of L clusters indexed by $l\in\{1,\cdots,L\}$. The cluster l is composed of K_l nodes and an FC l. The node set of the cluster l is

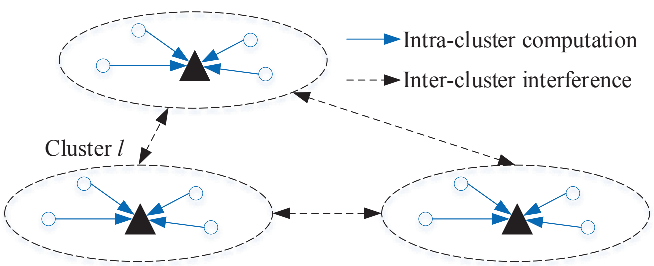


Figure 3. Multi-cluster networks.

given as $\mathcal{C}_l$ with the cardinality $|\mathcal{C}_l| = K_l$. Assuming that the clusters do not overlap, the node set of the network can be expressed as $\mathcal{C} = \bigcup_{l=1}^{L} \mathcal{C}_l$ with the cardinality $|\mathcal{C}| = K$, and the target function of the cluster l can be written as

$$f_l = \psi_l \left[\sum_{k \in \mathcal{C}_l} \varphi_k \left(s_k \right) \right]. \tag{36}$$

where $\psi_l(\cdot)$ is the post-processing function of the FC l.

The received signal of the FC l is

$$\begin{aligned} \mathbf{u}_l &= \underbrace{\mathbf{a}_l{}^T \sum_{k \in \mathcal{C}_l} \mathbf{H}_{k,l} \mathbf{b}_k \mathbf{x}_k}_{\text{intra-cluster CoMAC}} + \underbrace{\mathbf{a}_l{}^T \sum_{k \in \mathcal{C} - \mathcal{C}_l} \mathbf{H}_{k,l} \mathbf{b}_k \mathbf{x}_k}_{\text{ICI}} + \mathbf{a}_l{}^T \mathbf{n} \\ &= \sum_{k \in \mathcal{C}_l} \mathbf{x}_k + \\ &\quad \underbrace{\sum_{k \in \mathcal{C}_l} \left(\mathbf{a}_l{}^T \mathbf{H}_{k,l} \mathbf{b}_k - 1 \right) \mathbf{x}_k + \sum_{k \in \mathcal{C} - \mathcal{C}_l} \mathbf{a}_l{}^T \mathbf{H}_{k,l} \mathbf{b}_k \mathbf{x}_k + \mathbf{a}_l{}^T \mathbf{n}}_{\text{effective noise}}, \end{aligned} \tag{37}$$

where $\mathbf{a}_l \in \mathcal{R}^{N_r}$ denotes the receive beamforming vector of the FC l, $\mathbf{b}_k \in \mathcal{R}^{N_t}$ is the transmit beamforming vector of the node k, $\mathbf{x}_k \in \mathcal{R}^n$ is the transmit codewords of the node k, $\mathbf{H}_{k,l} \in R^{N_r \times N_t}$ is the channel matrix between the node k and the FC l, and $\mathbf{n} \in R^{N_r}$ is the noise vector with each element distributed as $\mathcal{N}\left(0, \sigma_n^2\right)$. The effective noise is composed of the distortion caused by non-uniform fading, the received noise, and the ICI.

Proposition 4. (The achievable function rate) For multi-cluster networks with ICI, the achievable function rate of CoMAC in the cluster l with q-ary quantization and nested lattice coding is given as

$$R_{C,l} \le \frac{1}{2\left[\log_2\left(K_l\right) + \log_2\left(\varphi_{\max}\right) - \log_2\left(\Delta\right)\right]} \log_2{}^+ \left(\frac{P_0}{\Gamma_l} \right), \tag{38}$$

where

$$\begin{aligned} \Gamma_l &= \sum_{k \in \mathcal{C}_l} \left\| \mathbf{a}_l{}^T \mathbf{H}_{k,l} \mathbf{b}_k - 1 \right\|^2 P_0 \\ &\quad + \sum_{k \in \mathcal{C} - \mathcal{C}_l} \left\| \mathbf{a}_l{}^T \mathbf{H}_{k,l} \mathbf{b}_k \right\|^2 P_0 + \left\| \mathbf{a}_l \right\|^2 \sigma_n^2 \end{aligned} \tag{39}$$

is the MSE of the effective noise of CoMAC in the cluster l.

Proof. The proof is similar to that of the Proposition 1, which is omitted for conciseness. □

A. Transceiver design

The achievable function rate of the network is the sum of the achievable function rate of each cluster $R_{C,l}$. And the transceiver optimization to maximize the achievable function rate of the network is formulated as

$$\begin{aligned} \text{(P3)} \quad & \max_{\{\mathbf{a}_l\},\{\mathbf{b}_k\}} \quad \sum_{l=1}^{L} R_{C,l} \\ & \text{s.t.} \qquad \left\| \mathbf{b}_k \right\|^2 \le 1, \forall k \end{aligned} \tag{40}$$

Since $\mathbf{a}_l$ of the FC l is not related to the other clusters, it is designed to maximize $R_{C,l}$, which can be further transformed into the MMSE receiver of the effective noise of CoMAC in the cluster l.

Taking the partial derivative of Γ_l in (39) to $\mathbf{a}_l$, the optimal receive beamforming vector of the FC l can be obtained as

$$\begin{aligned} \mathbf{a}_l &= \left(\sum_{k \in \mathcal{C}} \mathbf{H}_{k,l} \mathbf{b}_k \mathbf{b}_k{}^T \mathbf{H}_{k,l}{}^T + \gamma^{-1} \mathbf{I} \right)^{-1} \sum_{k \in \mathcal{C}_l} \mathbf{H}_{k,l} \mathbf{b}_k \\ &= \left(\mathbf{G}_l \mathbf{B} \mathbf{B}^T \mathbf{G}_l{}^T + \gamma^{-1} \mathbf{I} \right)^{-1} \mathbf{G}_l \mathbf{B} \mathbf{d}_l{}^T \end{aligned}, \tag{41}$$

where $\mathbf{G}_l = [\mathbf{H}_{1,l}, \mathbf{H}_{2,l}, \cdots, \mathbf{H}_{K,l}]$, $\mathbf{B} = \text{diag}\{\mathbf{b}_1, \mathbf{b}_2, \cdots, \mathbf{b}_K\}$, and $\mathbf{d}_l = [d_1, d_2, \cdots, d_K]$ is a length-K vector with K_l elements corresponding to the nodes belonging to the cluster l being 1 and the other elements being 0, that is $d_k = 1, k \in \mathcal{C}_l$ and $d_k = 0, k \notin \mathcal{C}_l$. The corresponding MSE of the effective noise of CoMAC in the cluster l can be calculated as

$$\begin{aligned} \Gamma_l &= \left(\mathbf{a}_l{}^T \mathbf{G}_l \mathbf{B} \mathbf{B}^T \mathbf{G}_l{}^T \mathbf{a}_l - 2\mathbf{a}_l{}^T \mathbf{G}_l \mathbf{B} \mathbf{d}_l{}^T + K_l \right) P_0 + \mathbf{a}_l{}^T \mathbf{a}_l \sigma_n^2 \\ &= \left(K_l - \mathbf{d}_l \mathbf{B}^T \mathbf{G}_l{}^T \left(\mathbf{G}_l \mathbf{B} \mathbf{B}^T \mathbf{G}_l{}^T + \gamma^{-1} \mathbf{I} \right)^{-1} \mathbf{G}_l \mathbf{B} \mathbf{d}_l{}^T \right) P_0, \\ &\overset{(a)}{=} \mathbf{d}_l \left(\mathbf{I} + \gamma \mathbf{B}^T \mathbf{G}_l{}^T \mathbf{G}_l \mathbf{B} \right)^{-1} \mathbf{d}_l{}^T P_0 \end{aligned} \tag{42}$$

where the procedure (a) is also due to the matrix inversion lemma as that in (21).

In order to optimize the transmit beamforming vector for each node, we utilize the relationship between the sum rate maximization and the weighted sum MSE minimization. We first establish the relationship between the achievable function rate and the weighted sum MSE of CoMAC, which is formulated as

$$\begin{aligned} \text{(P4)} \quad & \max_{\{\mathbf{b}_k\}} \quad \sum_{l=1}^{L} w_l \Gamma_l \\ & s.t. \qquad \left\| \mathbf{b}_k \right\| \le 1, \forall k \end{aligned}, \tag{43}$$

where w_l is an adaptive weight associated with the cluster l. This argument is similar to that given for MIMO interference channel [30] and MIMO full-duplex interference channel [31].

Proposition 5. (The relationship between sum rate and weighted sum MSE of CoMAC) The gradient of sum rate problem and the gradient of weighted sum MSE problem are equal if the adaptive weight w_l is chosen as

$$w_l = \frac{P_0}{2\ln 2\left[\log\left(K_l\right)+\log\left(\varphi_{\max}\right)-\log\left(\Delta\right)\right]\Gamma_l}, \tag{44}$$

where Γ_l is given in (39). Then we can solve the sum rate maximization of the problem P3 through solving the weighted sum MSE minimization of the problem P4.

Proof. The Lagrangian functions of the optimization problems P3 and P4 can be given as

$$\mathcal{L}_2 = -\sum_{l=1}^{L} R_{C,l} + \sum_{k=1}^{K}\lambda_k\left(\left\|\mathbf{b}_k\right\|^2-1\right) \tag{45}$$

and

$$\mathcal{L}_3 = \sum_{l=1}^{L} w_l\Gamma_l + \sum_{k=1}^{K}\lambda_k\left(\left\|\mathbf{b}_k\right\|^2-1\right) \tag{46}$$

respectively. The gradients of both Lagrangian functions with respect to $\mathbf{b}_k$ can be calculated as

$$\frac{\partial\mathcal{L}_2}{\partial\mathbf{b}_k} = -\sum_{l=1}^{L}\frac{R_{C,l}}{\partial\mathbf{b}_k}+\lambda_k\mathbf{b}_k, \tag{47}$$

and

$$\frac{\partial\mathcal{L}_3}{\partial\mathbf{b}_k} = \sum_{l=1}^{L} w_l\frac{\partial\Gamma_l}{\partial\mathbf{b}_k}+\lambda_k\mathbf{b}_k. \tag{48}$$

Then the KKT conditions of the problem P3 and the problem P4 can be satisfied simultaneously with the choice of the weight in (44), which completes the proof. □

The optimal $\mathbf{b}_k$ of the Problem P4 can be calculated as

$$\mathbf{b}_k = \left(\sum_{j=1}^{L} w_j{\mathbf{H}_{k,j}}^T\mathbf{a}_j{\mathbf{a}_j}^T\mathbf{H}_{k,j}+\lambda_k\mathbf{I}\right)^{-1} w_l{\mathbf{H}_{k,l}}^T\mathbf{a}_l, \tag{49}$$

where l is the index of the cluster that the node k belongs to, and λ_k is Lagrange multiplier associated with the power constraint of the node k, which can be found with the following one dimensional numerical search.

$$\begin{aligned}\left\|\mathbf{b}_k\right\|^2 &= \mathrm{Tr}\left[\left(\mathbf{\Sigma}_k+\lambda_k\mathbf{I}\right)^{-2}{w_l}^2{\mathbf{U}_k}^H{\mathbf{H}_{k,l}}^T\mathbf{a}_l{\mathbf{a}_l}^T\mathbf{H}_{k,l}\mathbf{U}_k\right] \\ &= \sum_{l=1}^{L}\frac{\mathbf{\Xi}_k\left(i,i\right)}{\left(\mathbf{\Sigma}_k\left(i,i\right)+\lambda_k\right)^2}\le 1,\end{aligned}, \tag{50}$$

where $\mathrm{SVD}\left(\sum_{j=1}^{L} w_j{\mathbf{H}_{k,j}}^T\mathbf{a}_j{\mathbf{a}_j}^T\mathbf{H}_{k,j}\right) = \mathbf{U}_k\mathbf{\Sigma}_k{\mathbf{U}_k}^H$, $\mathbf{\Xi}_k = {\mathbf{U}_k}^H{\mathbf{H}_{k,l}}^T\mathbf{a}_l{w_l}^2{\mathbf{a}_l}^T\mathbf{H}_{k,l}\mathbf{U}_k$.

Since the problem P4 is also convex over $\mathbf{a}_l$ and $\mathbf{b}_k$ with fixed w_l but not jointly, a standard technique to solve the problem is using alternating optimizing of $\mathbf{a}_l$ and $\mathbf{b}_k$ with adaptive weight w_l. The convergence of the iterations with adaptive w_l is proved as follows.

Proposition 6. (The convergence of the iterations with adaptive w_l) The iterations of $\mathbf{b}_k$, $\mathbf{a}_l$ optimization with adaptive w_l ensure local convergence.

Proof. Because the sum CoMAC rate is a convex function of the MSE of the effective noise, we have

$$\begin{aligned}\sum_{l=1}^{L} R_{C,l}\left[\Gamma_l\left(i+1\right)\right] \ge& \sum_{l=1}^{L} R_{C,l}\left[\Gamma_l\left(i\right)\right] \\ &+\sum_{l=1}^{L}\frac{\partial R_{C,l}\left[\Gamma_l(i)\right]}{\partial\Gamma_l(i)}\left[\Gamma_l\left(i+1\right)-\Gamma_l\left(i\right)\right]\end{aligned}, \tag{51}$$

where $\Gamma_l\left(i\right)$ is the MSE of the effective noise in the i-th iteration. According to $R_{C,l}\left(\Gamma_l\right)$ in (38), we have

$$\frac{\partial R_{C,l}\left(\Gamma_l\right)}{\partial\Gamma_l} = -w_l. \tag{52}$$

Thus,

$$\begin{aligned}\sum_{l=1}^{L} R_{C,l}\left[\Gamma_l\left(i+1\right)\right] \overset{(a)}{\ge}& \sum_{l=1}^{L} R_{C,l}\left[\Gamma_l\left(i\right)\right] \\ &-\sum_{l=1}^{L} w_l\left(i\right)\left[\Gamma_l\left(i+1\right)-\Gamma_l\left(i\right)\right], \\ \overset{(b)}{>}& \sum_{l=1}^{L} R_{C,l}\left[\Gamma_l\left(i\right)\right]\end{aligned} \tag{53}$$

where (a) is because the rate $R_{C,l}\left[\Gamma_l\left(i\right)\right]$ is a convex function of the MSE $\Gamma_l\left(i\right)$, and (b) is because the iteration optimization of $\mathbf{b}_k$, $\mathbf{a}_l$ with fixed w_l is to minimize the weighted sum MSE in P3.

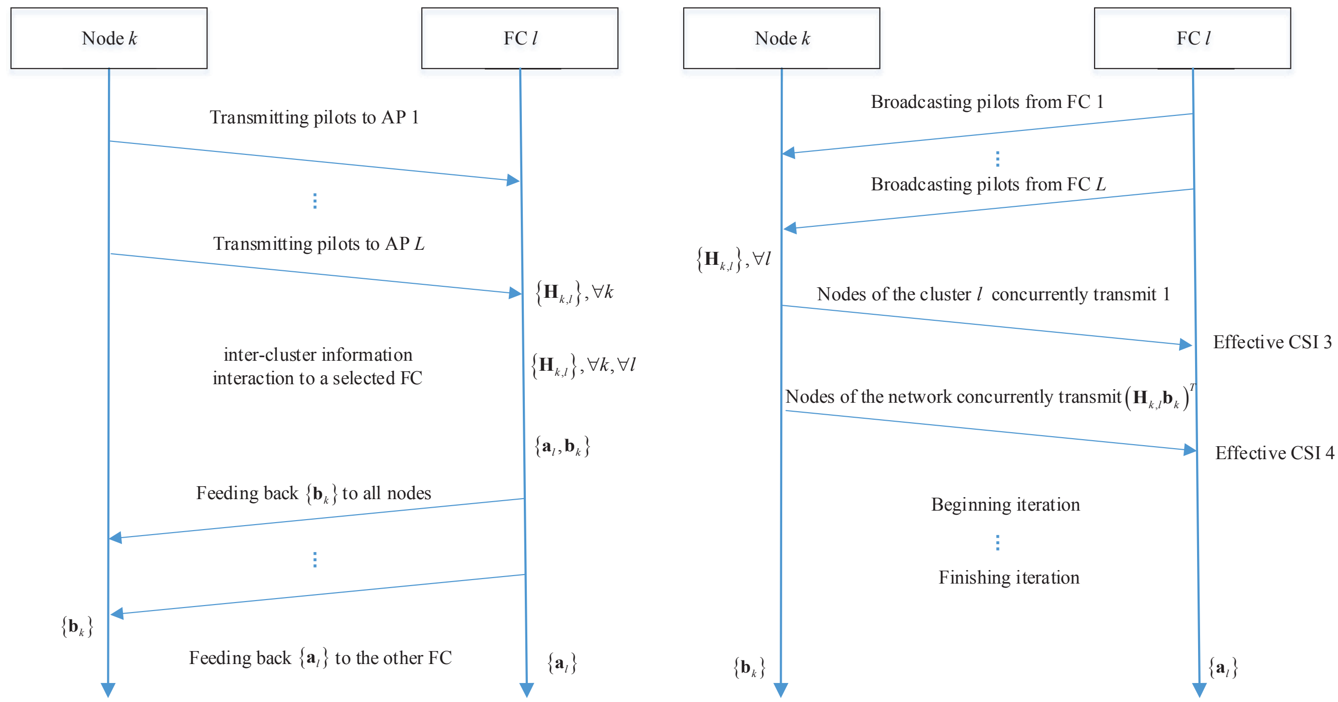
It can be seen that the iterations of $\mathbf{b}_k$, $\mathbf{a}_l$ and w_l optimization will increase the sum rate relative to that of the last iteration. Because the sum rate is upper bounded, the convergence can be guaranteed, which completes the proof. □

B. Low-complexity signaling procedure

The traditional transceiver optimization algorithm is given in Algorithm 3, where the CSI of all nodes is aggregated at a central FC and then the optimized $\mathbf{b}_k$ is transmitted back to each node after alternating optimization.

The detailed signaling procedure of Algorithm 1 is illustrated in Fig. 4(a). It will not only incur a high signaling cost for a large-scale network but also require the information interaction between different FCs. The corresponding time complexity of Algorithm 3 is provided as follows.

Proposition 7. (The time complexity of the traditional signaling procedure) Without considering the time complexity of the optimization and the information interaction between different FCs, the time complexity of the traditional signaling procedure in multi-cluster networks requires $(L+1)N_tK$ symbol slots.



(a) Traditional signaling procedure

(b) Proposed signaling procedure

Figure 4. The signaling procedure of transceiver optimization.

Algorithm 3 Traditional transceiver optimization of CoMAC with ICI.

Initialize $\mathbf{a}_l^{(0)}$, convergence threshold ε, $i = 0$, and the sum computation rate $\sum_{l=1}^{L} R_{C,l}{}^{(0)}$.
Each FC estimates the CSI between itself to all devices.
The central FC aggregates the global CSI via inter-cluster interaction.
repeat
  The central FC optimizes $\mathbf{b}_k^{(i+1)}$ according to (49).
  The central FC optimizes $\mathbf{a}_l^{(i+1)}$ according to (41).
  The central FC updates $w_l^{(i+1)}$ according to (44).
  The central FC computes the corresponding sum computation rate $\sum_{l=1}^{L} R_{C,l}{}^{(i+1)}$.
until $\left\| \sum_{l=1}^{L} R_{C,l}{}^{(i+1)} - \sum_{l=1}^{L} R_{C,l}{}^{(i)} \right\| \leq \varepsilon$
The central FC transmits the optimized $\mathbf{a}_l$ to the other FCs via inter-cluster information interaction.
The central FC transmits back the optimized $\mathbf{b}_k$ to each node in turn.

Proof. For the cluster l, it takes at least $N_t \times K$ symbol slots for estimating the CSI between all nodes and the FC l. Thus, L clusters require $L \times N_t \times K$ symbol slots to complete the channel training process. After optimization at the central FC, it takes at least $N_t \times K$ symbol slots to transmit back the optimized $\mathbf{b}_k$ to each node in turn. Without considering the time complexity of the optimization and the information interaction between different FCs, it completes the proof. □

The time complexity of the traditional signaling procedure linearly increases with the number of nodes K even without considering the time complexity of the information interaction between different FCs. Following a similar idea of Algorithm 2, we propose a signaling procedure based on the channel reciprocity and the defined effective CSI.

The optimized receiver $\mathbf{a}_l$ in (41) can be also regarded as a function of the CSI $\mathbf{H}_{k,l}, k \in \mathcal{C}$. Instead of collecting the global CSI of the network, it only requires the effective CSI defined as follows.

$$\text{effective CSI 3: } \sum_{k \in \mathcal{C}_l} \mathbf{H}_{k,l}\mathbf{b}_k, \forall l, \tag{54}$$

which can be obtained by the FC l when all nodes of the cluster l concurrently transmit 1 with beamforming vector $\mathbf{b}_k$.

$$\text{effective CSI 4: } \sum_{k \in \mathcal{C}} \mathbf{H}_{k,l}\mathbf{b}_k{\mathbf{b}_k}^T{\mathbf{H}_{k,l}}^T, \forall l, \tag{55}$$

which can be obtained by the FC l when all nodes of the network concurrently transmit $(\mathbf{H}_{k,l}\mathbf{b}_k)^T$ with beamforming vector $\mathbf{b}_k$.

Also, the optimized transmitter of the node k in (44) can be regarded as a function of the CSI $\mathbf{H}_{k,l}, \forall l$ between itself and all FCs. We can use channel reciprocity to allow each node to measure its channel to all FCs with very low overhead. Note that the forward and reverse channels are always the same since they operate on the same carrier frequency in TDD systems.

Thus, we can propose Algorithm 4 with a novel signaling procedure, where each node requires the CSI between itself and all FCs and the FC l requires the effective CSI in (54)

Algorithm 4 Proposed transceiver optimization of CoMAC with ICI.

Initialize $\mathbf{a}_l^{(0)}$, convergence threshold ε, $i=0$, and the sum computation rate $\sum_{l=1}^{L} R_{C,l}{}^{(0)}$.
Each node estimates the CSI $\mathbf{H}_{k,l}, \forall l$ between itself and the FC l based on the broadcasting pilots.
repeat
  Each node optimizes $\mathbf{b}_k^{(i+1)}$ based on (49).
  Each FC estimates the effective CSI 3 in (54) in turn.
  Each FC estimates the effective CSI 4 in (55) in turn.
  Each FC optimizes $\mathbf{a}_l^{(i+1)}$ based on (41).
  Each FC optimizes $w_l^{(i+1)}$ based on (44).
  Each FC transmits back $\mathbf{a}_l^{(i+1)}$ and $w_l^{(i+1)}$ to all nodes in turn.
  The central FC computes the corresponding sum computation rate $\sum_{l=1}^{L} R_{C,l}{}^{(i+1)}$.
until $\left\|\sum_{l=1}^{L} R_{C,l}{}^{(i+1)} - \sum_{l=1}^{L} R_{C,l}{}^{(i)}\right\| \le \varepsilon$

and (55). The transmitter $\mathbf{b}_k$ is optimized at the node k and the receiver $\mathbf{a}_l$ is optimized at the FC l in an iterative way. The signaling procedure of Algorithm 4 is illustrated in Fig. 4(b), where the time complexity is given as follows.

Proposition 8. (The complexity of the proposed signaling procedure) Assuming the number of iterations is N_{iter}, the total symbol slots of the proposed signaling procedure in multi-cluster networks is $L[N_r + 2(N_r+1)N_{iter}]$.

Proof. Each FC broadcasts $N_r \times N_r$ pilot matrix in turn for each node to estimate the CSI $\mathbf{H}_{k,l}, \forall l$ between itself and each FC, which takes $N_r \times L$ symbol slots. For each iteration, it takes 1 symbol slot for the FC l to estimate the effective CSI 3 in (54) and N_r symbol slots to estimate the effective CSI 4 in (55). After optimization at the FC l, it takes (N_r+1) symbol slots to broadcast $\mathbf{a}_l$ and w_l to all devices. In order to avoid interference, each cluster follows the signaling procedure in turn, it takes $2L(N_r+1)$ for all clusters. Assuming the number of iterations is N_{iter}, it takes totally $2L(N_r+1)N_{iter}$ for the signaling iteration, which completes the proof. □

Remark 2. (Comparison between the traditional signaling procedure and the proposed one) In order to illustrate the low complexity of the proposed signaling procedure, we consider a typical dense network composed of $L=4$ clusters. Each cluster has 50 devices and the network has $K=200$ devices. Each node and each FC has $N_t = N_r = 2$ antennas. The symbol slots required for the traditional signaling procedure is $(L+1)N_tK = 1000$ even without considering the time complexity of inter-cluster information interaction. According to the simulation results in Section VI, the number of iteration is about $N_{iter}=10$. The symbol slots required for the proposed signaling procedure is $L[N_r+2(N_r+1)N_{iter}] = 268$, which is 26.8% of the complexity of the traditional signaling procedure. Besides, the proposed signaling procedure does not require inter-cluster information interaction.

VI. Simulation Results And Discussion

In this section, we provide some simulation results to illustrate the performances of the new designs. Both single cluster networks without ICI and multi-cluster networks with ICI are discussed. The simulation parameters are set as follows unless specified otherwise. For multi-cluster networks, the network is composed of $L=2$ clusters and the number of nodes in each clusters K_l is the same. The elements of the channel matrices are generated as i.i.d. Gaussian random variables $\mathcal{N}\left(0, \sigma_h^2\right)$. For the channel between the nodes and their own FC, we assume that $\sigma_h^2 = 0dB$. For the channel between the nodes and the other FC, we assume that $\sigma_h^2 = -20dB$, where the ICI is limited due to the cluster distribution. All numerical results are obtained using 10^5 Monte-Carlo simulations.

A. Convergence and power consumption

The convergence properties of the proposed transceiver optimization in single cluster network and multi-cluster network are shown in Fig. 5. The transmit SNR is $\gamma = 20dB$. We adopt a random channel realization for both figures, where the convergence results of the sum computation rate with different numbers of nodes $K = 20$, 40 and 60 are provided. Although the convergence speed varies for the different channel realizations, the results indicate that the number of iterations achieving convergence is around 10 in all cases. The convergence speed slightly decreases with the increase of the number of nodes K. That is because the searching space of the optimization increases with the increase of the number of nodes K.

The power consumption comparison between the traditional and the proposed signaling procedure is illustrated in Fig. 6. The signaling power is $P_{sig} = \gamma_{sig} N_o W$, where γ_{sig} is the signaling SNR, signaling bandwidth $W = 10$ kHz, and Thermal noise power $N_o = -141$ dBW/MHz. The circuit power of each node is assumed to be 100 mW. Assuming the time duration of signaling slot is 10 ms. For the traditional signaling procedure, the total power consumption increases almost linearly with the number of nodes. This is because the time complexity of traditional signaling procedure increases linearly with the number of nodes. Both the transmit power consumption and the static circuit power consumption depend on the time complexity of the signaling procedure. For the proposed signaling procedure, the increasing rate of the total power consumption is much smaller. Although all nodes are required to transmit signaling for the proposed signaling procedure, the signaling power will be superposed coherently. Thus, the signaling power of each node is much smaller than that of the traditional signaling procedure. Also, the time complexity of the proposed signaling procedure is much smaller than that of the traditional signaling procedure. Thus, the transmit power and static circuit power consumption decrease correspondingly.

B. Achievable function rate

The achievable function rate of the network versus the transmit SNR ranging from 0dB to 20dB is illustrated in Fig. 7 for single cluster network and in Fig. 8 for multi-cluster network, respectively. The number of nodes is $K=20$.

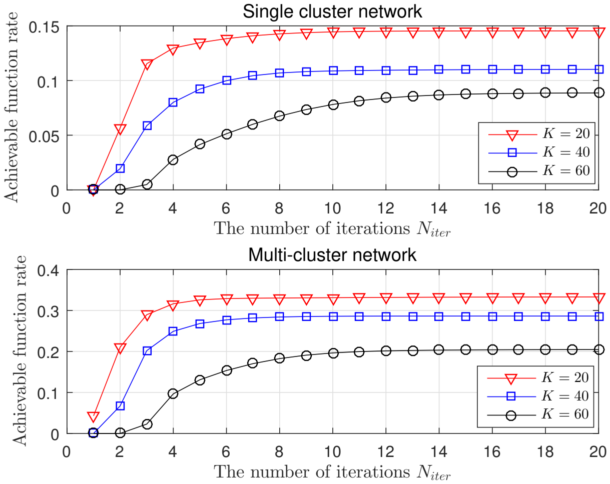


Figure 5. The converge properties of transceiver optimization.

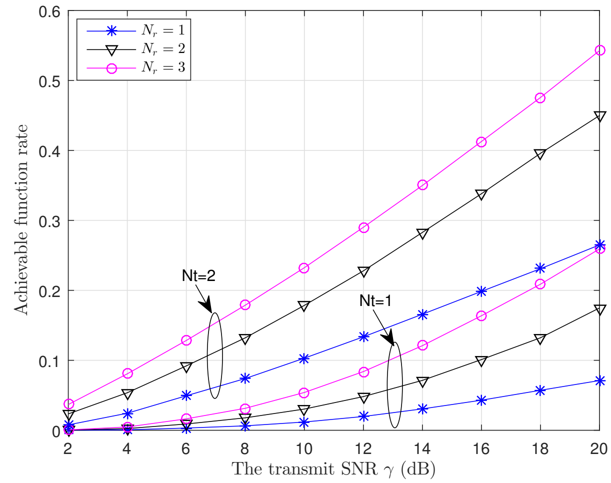


Figure 7. The achievable function rate versus SNR in single cluster network.

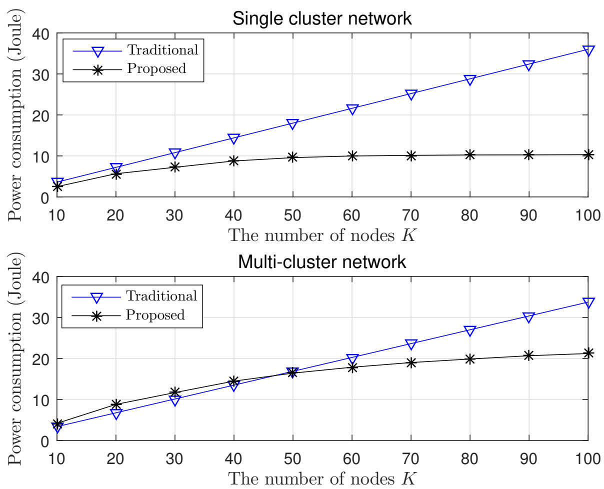


Figure 6. The power consumption of the signaling procedure.

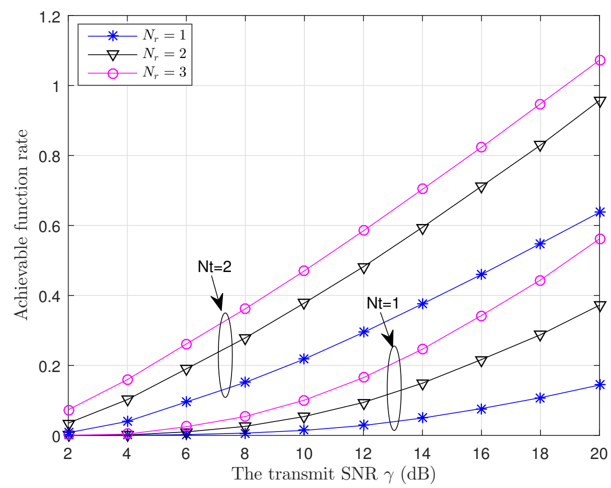


Figure 8. The sum function rate versus SNR in multi-cluster network.

Different numbers of transmit antennas $N_t = 1, 2$ and receive antennas $N_r = 1, 2, 3$ are considered. Firstly, the sum achievable function rate increases with the increase of the transmit SNR γ for both scenarios, where the increasing rate slightly increases with the number of the transmit antennas N_t and the number of the receive antennas N_r. That can be explained as the increase of the multi-antenna diversity [3]. Compared the achievable function rate in Fig. 7 with that in Fig. 8, it can be seen that the sum achievable function rate for multi-cluster network is higher than that for single cluster network. That is because the design with ICI will optimize the achievable function rate in a joint way.

The achievable function rate of the network versus different number of the nodes ranging from 10 to 100 is shown in Fig. 9 for single cluster network and in Fig. 10 for multi-cluster network, respectively. The transmit SNR is $\gamma = 20dB$. Different numbers of transmit antennas $N_t = 1, 2$ and receive antennas $N_r = 1, 2, 3$ are considered. Firstly, the achievable function rate decreases with the increase of the number of nodes K for both scenarios. That is because the increase of the number of nodes will incur a larger distortion error caused by the non-uniform fading between different nodes. And the corresponding effective noise of the CoMAC will also increases, which will decrease the achievable function rate. Also, the achievable function rate also increases with the number of antennas for both scenarios due to the increase of the multi-antenna diversity, and the achievable function rate in multi-cluster network is higher than that in single cluster network due to the benefit of joint optimization when the interference between different clusters is limited.

C. CSI error and synchronization phase offsets

We compare the achievable function rate between the proposed signaling procedure and the traditional signaling proce-

[3]Because the target function experiences independently fading channels, the MSE of the effective noise decreases due to the coherent combining of the replicas of the target function.In this work, we only consider single function computed over MAC through multi-antennas without multiplexing gain.

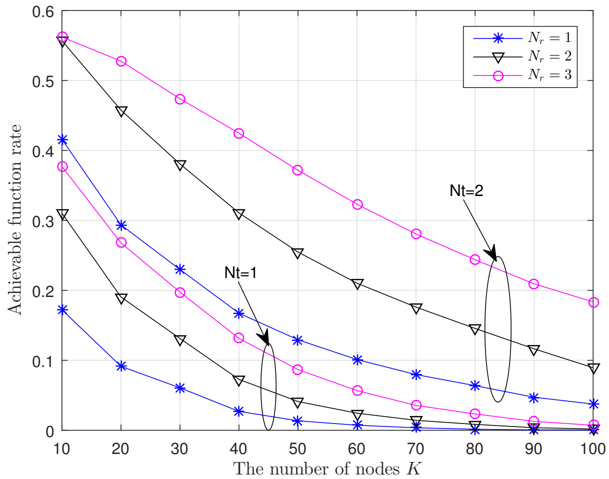


Figure 9. The sum computation rate versus different number of nodes in single cluster network.

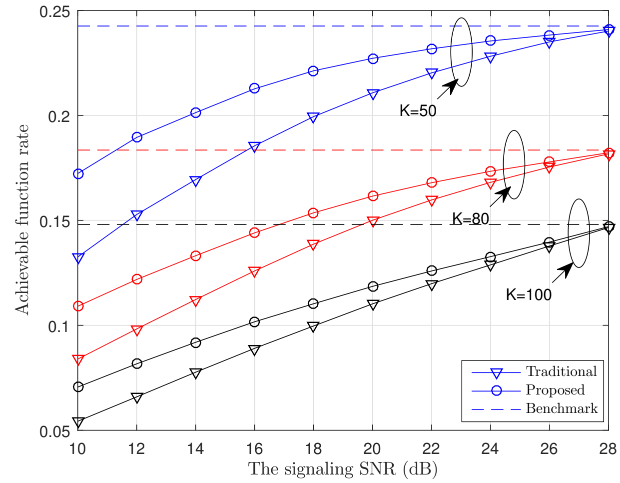


Figure 11. The achievalbe function rate versus different number of nodes in multi-cluster network.

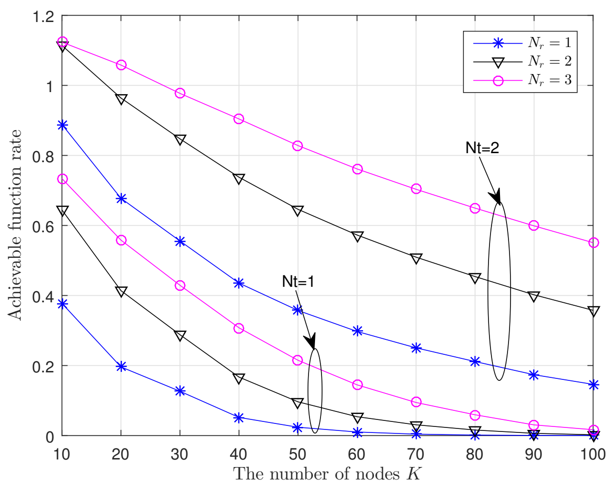


Figure 10. The sum computation rate versus different number of nodes in multi-cluster network.

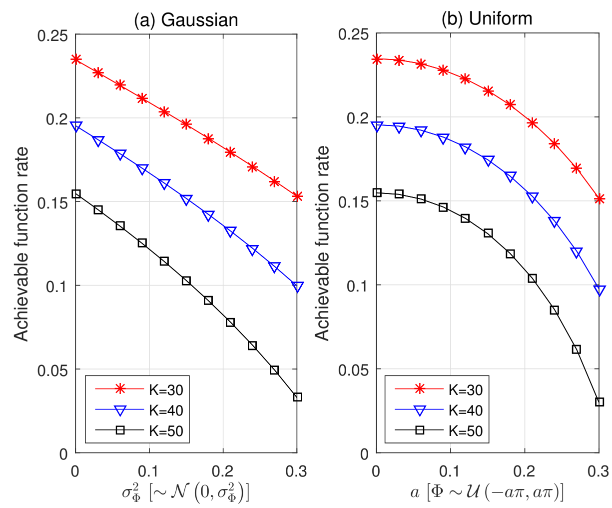


Figure 12. The sum computation rate versus different number of nodes in multi-cluster network.

dure in Fig. 11, where different signaling SNR is considered (10 dB$\sim$30 dB). The benchmark is obtained with perfect CSI. It can be seen that CSI error will cause performance degradation for both signaling procedures. However, the performance degradation of the proposed signaling procedure is smaller than that of the traditional signaling procedure. That is because the signaling power will be superposed coherently for the proposed signaling procedure. And the impacts of individual CSI error will not accumulate.

For the line-of-sight scenario, the synchronization phase offset of "AirShare" is not significant. But for non-line-of-sight scenario, the synchronization phase offset of "AirShare" can not be neglected due to the phase offset of broadcasting channel. Further, we provide the effect of synchronization phase offsets in Fig. 12. The transmit SNR is 20 dB, and different numbers of nodes $K = 30, 40, 50$ are simulated. In Fig. 12(a), we consider Gaussian synchronization phase offsets $\Phi \sim \mathcal{N}(0, \sigma_\Phi^2)$ with variance $\sigma_\Phi^2 = 0 \sim 0.3$. In Fig. 12(b), we consider uniform synchronization phase offsets $\Phi \sim \mathcal{U}(-a\pi, a\pi)$ and with $a = 0 \sim 0.3$. It can be seen that the increase of synchronization phase offsets will decrease the achievable function rate. Also, when the number of nodes increases, the achievable function rate decreases. This is because the impact of synchronization phase offsets will be intensified with more nodes.

VII. Conclusion

In this work, we have studied the transceiver design of CoMAC in wireless networks. The achievable function rate of CoMAC considering both the channel fading and the transceiver design has been derived using the quantization and the nested lattice coding. Based on the derived results, we have first maximized the achievable function rate through the transceiver optimization in single cluster network. Then a more complicated multi-cluster network has been further analyzed, where ICI has been considered. In order to avoid massive

CSI aggregation during the optimization and reduce the time-complexity, low-complexity signaling procedures have been proposed for both scenarios by using the channel reciprocity and the defined effective CSI. The work has provided a novel design for the MIMO transceiver design in function-centric networks that is different from the traditional one in information-centric networks.

REFERENCES

[1] L. Xu, R. Collier, and G. M. P. OHare, "A survey of clustering techniques in WSNs and consideration of the challenges of applying such to 5G IoT scenarios," *IEEE Internet Things J.*, vol. 4, no. 5, pp. 1229–1249, Oct. 2017.

[2] Y. Liu, Z. Qin, M. Elkashlan, Z. Ding, A. Nallanathan, and L. Hanzo, "Nonorthogonal multiple access for 5G and beyond," *Proc. IEEE*, vol. 105, no. 12, pp. 2347–2381, Dec. 2017.

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

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

[5] M. Goldenbaum, H. Boche, and S. Staczak, "Harnessing interference for analog function computation in wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 61, no. 20, pp. 4893–4906, Oct. 2013.

[6] S. Sigg, P. Jakimovski, and M. Beigl, "Calculation of functions on the RF-channel for IoT," in *2012 3rd IEEE International Conference on the Internet of Things*, Oct. 2012, pp. 107–113.

[7] A. Kortke, M. Goldenbaum, and S. Staczak, "Analog computation over the wireless channel: A proof of concept," in *IEEE SENSORS 2014 Proceedings*, Nov. 2014, pp. 1224–1227.

[8] M. Goldenbaum and S. Stanczak, "Robust analog function computation via wireless multiple-access channels," *IEEE Trans. Commun.*, vol. 61, no. 9, pp. 3863–3877, Sep. 2013.

[9] J. Huang, Q. Zhang, Q. Li, and J. Qin, "Robust parallel analog function computation via wireless multiple-access MIMO channels," *IEEE Signal Process. Lett.*, vol. 22, no. 9, pp. 1297–1301, Sep. 2015.

[10] F. Ang, L. Chen, N. Zhao, Y. Chen, and F. R. Yu, "Robust design for massive csi acquisition in analog function computation networks," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 2361–2373, March 2019.

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

[12] L. Chen, N. Zhao, Y. Chen, F. R. Yu, and G. Wei, "Over-the-air computation for IoT networks: Computing multiple functions with antenna arrays," *IEEE Internet Things J.*, vol. 5, no. 6, pp. 5296–5306, Dec. 2018.

[13] B. Nazer and M. Gastpar, "Compute-and-forward: Harnessing interference through structured codes," *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6463–6486, Oct. 2011.

[14] R. Zamir, B. Nazer, Y. Kochman, and I. Bistritz, *Lattice Coding for Signals and Networks: A Structured Coding Approach to Quantization, Modulation and Multiuser Information Theory*. Cambridge University Press, 2014.

[15] Z. Chen, P. Fan, and K. B. Letaief, "Compute-and-forward: Optimization over multisource multirelay networks," *IEEE Trans. Veh. Technol.*, vol. 64, no. 5, pp. 1806–1818, May 2015.

[16] M. Hejazi, S. M. Azimi-Abarghouyi, B. Makki, M. Nasiri-Kenari, and T. Svensson, "Robust successive compute-and-forward over multiuser multirelay networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 10, pp. 8112–8129, Oct. 2016.

[17] A. S. Guerreiro, G. Fraidenraich, and S. Kumar, "Approximate sum rate for integer-forcing receiver," *IEEE Trans. Commun.*, vol. 65, no. 11, pp. 4899–4910, Nov. 2017.

[18] Y. C. Huang, K. R. Narayanan, and P. C. Wang, "Lattices over algebraic integers with an application to compute-and-forward," *IEEE Trans. Inf. Theory*, pp. 1–1, 2018.

[19] W. Liu and C. Ling, "Efficient integer coefficient search for compute-and-forward," *IEEE Trans. Wireless Commun.*, vol. 15, no. 12, pp. 8039–8050, Dec. 2016.

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

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

[22] S. W. Jeon and B. C. Jung, "Opportunistic function computation for wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 6, pp. 4045–4059, Jun. 2016.

[23] F. Wu, L. Chen, N. Zhao, Y. Chen, F. R. Yu, and G. Wei, "Computation over wide-band multi-access channels: Achievable rates through sub-function allocation," *IEEE Trans. Wireless Commun.*, vol. 18, no. 7, pp. 3713–3725, July 2019.

[24] F. Zhu, F. Gao, S. Jin, H. Lin, and M. Yao, "Robust downlink beamforming for BDMA massive MIMO system," *IEEE Trans. Commun.*, vol. 66, no. 4, pp. 1496–1507, Apr. 2018.

[25] C. Song and Y. Jeon, "Weighted MMSE precoder designs for sum-utility maximization in multi-user SWIPT network-MIMO with per-BS power constraints," *IEEE Trans. Veh. Technol.*, vol. 67, no. 3, pp. 2809–2813, Mar. 2018.

[26] C. Xing, Y. Ma, Y. Zhou, and F. Gao, "Transceiver optimization for multi-hop communications with per-antenna power constraints," *IEEE Trans. Signal Process.*, vol. 64, no. 6, pp. 1519–1534, Mar. 2016.

[27] X. Sun, N. Yang, S. Yan, Z. Ding, D. W. K. Ng, C. Shen, and Z. Zhong, "Joint beamforming and power allocation in downlink NOMA multiuser mimo networks," *IEEE Trans. Wireless Commun.*, vol. 17, no. 8, pp. 5367–5381, Aug. 2018.

[28] R. Zamir, "Lattices are everywhere," in *2009 Information Theory and Applications Workshop*, Feb. 2009, pp. 392–421.

[29] J. Zhan, B. Nazer, M. Gastpar, and U. Erez, "MIMO compute-and-forward," in *2009 IEEE International Symposium on Information Theory*, June 2009, pp. 2848–2852.

[30] J. Shin and J. Moon, "Weighted-sum-rate-maximizing linear transceiver filters for the k-user MIMO interference channel," *IEEE Trans. Commun.*, vol. 60, no. 10, pp. 2776–2783, Oct. 2012.

[31] A. C. Cirik, R. Wang, Y. Hua, and M. Latva-aho, "Weighted sum-rate maximization for full-duplex MIMO interference channels," *IEEE Trans. Commun.*, vol. 63, no. 3, pp. 801–815, Mar. 2015.