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/131647

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.

NOMA-Enhanced Computation Over Multi-Access Channels

Fangzhou Wu, Li Chen, Nan Zhao, *Senior Member, IEEE,* Yunfei Chen, *Senior Member, IEEE,*
F. Richard Yu, *Fellow, IEEE,* and Guo Wei

Abstract—**Massive numbers of nodes will be connected in future wireless networks. This brings great difficulty to collect a large amount of data. Instead of collecting the data individually, computation over multi-access channels (CoMAC) provides an intelligent solution by computing a desired function over the air based on the signal-superposition property of wireless channels. To improve the spectrum efficiency in conventional CoMAC, we propose the use of non-orthogonal multiple access (NOMA) for functions in CoMAC. The desired functions are decomposed into several sub-functions, and multiple sub-functions are selected to be superposed over each resource block (RB). The corresponding achievable rate is derived based on sub-function superposition, which prevents a vanishing computation rate for large numbers of nodes. We further study the limiting case when the number of nodes goes to infinity. An exact expression of the rate is derived that provides a lower bound on the computation rate. Compared with existing CoMAC, the NOMA-based CoMAC not only achieves a higher computation rate but also provides an improved non-vanishing rate. Furthermore, the diversity order of the computation rate is derived, which shows that the system performance is dominated by the node with the worst channel gain among these sub-functions in each RB.**

Index Terms—**Achievable computation rate, limiting rate, NOMA, sub-function superposition, wireless networks.**

I. INTRODUCTION

In 5G and the Internet of Things, tens of billions of devices are expected to be deployed to serve our societies [1]–[4]. With such an enormous number of nodes, the collection of data using the conventional multi-access schemes is impractical because this would result in excessive network latency with limited radio resources.

To solve the problem, a promising solution is computation over multi-access channels (CoMAC). It exploits the signal-superposition property of wireless channels by computing

the desired function through concurrent node transmissions [5]–[19]. Many networks computing a class of nomographic functions of the distributed data can employ CoMAC [5]. For example, wireless sensor networks can use the framework of CoMAC since it only aims to obtain a functional value of the sensor readings (e.g., arithmetic mean, polynomial or the number of active nodes) instead of requiring all readings from all sensors.

Analog CoMAC was first studied in [6]–[10], where pre-processing at each node and post-processing at the fusion center were used to resist fading and compute functions. The designs of pre-processing and post-processing used to compute linear and non-linear functions have been proposed in [7], and the effect of channel estimation error was characterized in [8]. In order to verify the feasibility of analog CoMAC in practice, software-defined radio was built in [9]. Also, the authors in [10] studied how to compute multiple functions over-the-air with antennas arrays at devices and the access point, where different linear combinations with arbitrary coefficients for the Gaussian sources were computed. In summary, the simple analog CoMAC has led to an active area focusing on the design and implementation techniques for receiving the desired function.

Since analog CoMAC is not robust to noise, digital CoMAC was proposed to use joint source-channel coding in [11]–[20] to improve the equivalent signal-to-noise ratio (SNR). The potential of linear source coding was discussed in [11], and its application was presented in [12] for CoMAC. Compared with linear source coding, nested lattice coding can approach the performance of a standard random coding [13]. The lattice-based CoMAC was extended to a general framework in [15] for relay networks with linear channels and additive white Gaussian noise (AWGN). In order to combat non-uniform fading, a uniform-forcing transceiver design was given in [14]. Achievable computation rates were given in [15]–[17] for digital CoMAC through theoretical analysis.

The implementation of CoMAC faces several practical issues. One is the synchronization of all active sensors required for coherent combining at nodes. The frequency synchronization has been solved by an attractive solution, called "AirShare", which was developed in [21] for synchronizing nodes by broadcasting a reference-clock signal. To cope with the phase offset, a design has been proposed to estimate synchronization phase offset and to equalize the corresponding error in [22].

Especially in wide-band CoMAC (WB-CoMAC), the system has to face more issues caused by orthogonal frequency-

division multiplexing access (OFDMA) such as peak-to-average power ratio (PAPR) and inter-carrier interference. Considering the influence of PAPR, [23] suggests that block coding provides a way to tackle the PAPR problem. [24] provides a method to replace the conventional constellations by lattice codes that satisfy some geometrical properties, e.g., reducing PAPR or reducing average energy. Since the implement of CoMAC is based on lattice codes [15], the CoMAC system with low PAPR is desired by choosing a well-designed lattice code. The inter-carrier interference for OFDMA uplink has been extensively researched. Many works provided promising solutions to suppressing the influence of the inter-carrier interference [25], [26]. Since the inter-carrier interference is caused by the carrier frequency offset that exists in OFDM instead of CoMAC, the conventional schemes can be used in WB-CoMAC to remove the inter-carrier interference.

Besides, the computation rate of digital CoMAC eventually goes to zero in the fading multi-access channels (MAC) as the number of nodes increases. This is a serious issue that digital CoMAC endures. In order to avoid the vanishing rate, narrow-band CoMAC (NB-CoMAC) with opportunistic computation has been studied in [17]. WB-CoMAC with sub-function allocation has been extended in [20] against both frequency selective fading and the vanishing computation rate.

To the best of our knowledge, all the aforementioned CoMAC works only considered the use of orthogonal multiple access (OMA) for functions by transmitting the function in different resource blocks (RBs), i.e, time slots or sub-carriers. Non-orthogonal multiple access (NOMA) is well-known for improving spectrum efficiency but has never been considered in CoMAC [27]–[29]. Different from NOMA for information transmission, NOMA-based CoMAC (NOMA-CoMAC) superposes multiple functions instead of different bit sequences from different nodes in each RB. Also, nodes with disparate channel conditions are allowed to be served simultaneously in conventional NOMA to improve the performance, whereas the node with poor channel condition only makes the computation rate vanishing in NOMA-CoMAC system since the function computed by nodes requires the uniform fading. It suggests that the direct use of NOMA in CoMAC system is not suitable.

Motivated by the above observations, in this work, we propose a NOMA-CoMAC system through the division, superposition, SIC and reconstruction of the desired functions. The analytical expression for the achievable computation rate with sub-function superposition is derived based on nested lattice coding [16]–[19]. Furthermore, several limiting cases are considered to characterize the lower bound of the computation rate and the diversity order. Our contributions can be summarized as follows:

- *Novel NOMA-CoMAC*. We propose a NOMA-CoMAC system with sub-function superposition. Unlike NOMA systems for information transmission, NOMA-CoMAC decomposes the desired functions into several sub-functions, superposes these sub-functions with large equivalent channel gains in each RB, executes the process of SIC and reconstructs the desired functions at the fusion center.
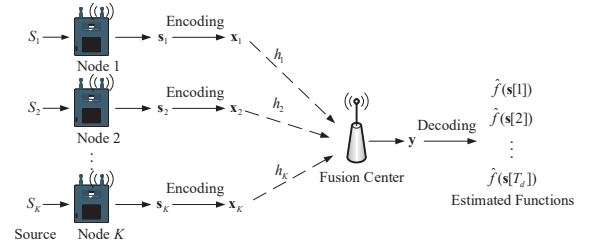

Fig. 1. Framework of narrow-band CoMAC

- *Improved computation rate*. The analytical expression of the computation rate of NOMA-CoMAC is derived. Using the closed-form expression, the power allocated to each node can be calculated directly with low computational complexity. Compared with the conventional CoMAC schemes, both achievable computation rate and non-vanishing rate with massive nodes are improved.
- *Limiting cases*. We characterize the lower bound of the computation rate with an exact expression as the number of nodes goes to infinity. It provides a straightforward way to evaluate the system performance. As the power of each node goes to infinity, we obtain the diversity order of the computation rate of NOMA-CoMAC. It shows that the node with the worst channel gain among these sub-functions in each RB plays a dominant role.

The rest of the paper is organized as follows. Section II introduces the definitions and the existing results of NB-CoMAC and WB-CoMAC. In Section III-B, we first provide the system model of NOMA-CoMAC and then summarize the main results of this paper. Section IV presents the proposed NOMA-CoMAC with sub-function superposition in detail and analyzes the computation rate. Section V focuses on the performance of the proposed NOMA-CoMAC, which includes power control and outage analysis. Simulation results and the corresponding discussion are presented in Section VI, and conclusions are given in Section VII.

II. PRELIMINARIES

We introduce two typical CoMAC frameworks in this section, and review the main results of several previous works. We define $\mathsf{C}^+(x) = \max\left\{\frac{1}{2}\log(x), 0\right\}$ and $\lceil x \rceil = \min\{n \in \mathbb{Z} | x \leq n\}$ as the ceiling function. Let $[1:n]$ denote a set $\{1, 2, \cdots, n\}$ and $(\cdot)^{\mathrm{T}}$ represent the transpose of a vector or matrix. For a set $\mathcal{A}$, $|\mathcal{A}|$ denotes the cardinality of $\mathcal{A}$. Let the entropy of a random variable A be $H(A)$ and $\mathrm{diag}\{a_1, a_2, \cdots, a_n\}$ denote the diagonal matrix of which the diagonal elements are from a_1 to a_n. A set $\{x_1, x_2, \cdots, x_N\}$ is written as $\{x_i\}_{i \in [1:N]}$ or $\{x_i\}_{i=1}^N$ for short.

A. Narrow-Band CoMAC

As shown in Fig. 1, the node i draws data from the corresponding random source S_i and obtains a data vector $\mathbf{s}_i$ with length T_d. The node i encodes the data vector $\mathbf{s}_i$ with length T_d into the transmitted vector $\mathbf{x}_i$ with length n. All the

nodes transmit the vector over the air simultaneously. At the fusion center, the received vector $\mathbf{y}$ is the superposition of all the transmitted vector from all these nodes. After decoding, the desired functions are obtained. Since the length of the data vector of each node is T_d, the number of the desired functions computed at the fusion center is also T_d. $s_i[j]$ is the j-th data of the i-th node from the random source S_i, $\mathbf{s}[j] = [s_1[j], \cdots, s_K[j]]$ is the j-th data of all K nodes and $\mathbf{s}_i = [s_i[1], \cdots, s_i[T_d]]$ is the data vector of node i. Each data $s_i[j]$ belongs to $[0 : p - 1]$, which means that it is mapped to a number between 0 and $p - 1$ through quantization. Let $\mathbf{s_r} = [S_1, S_2, \cdots, S_K]$ be a random vector associated with a joint probability mass function $p_{\mathbf{s_r}}(\cdot)$ as $\mathbf{s}[j]$ is independently drawn from $p_{\mathbf{s_r}}(\cdot)$.

A function $f(\mathbf{s_r})$ with respect to the random source vector $\mathbf{s_r}$ is called the desired function. Its definition is given as follows.

Definition 1 (Desired Function). For all $j \in [1 : T_d]$, a function with independent variables $\{s_1[j], s_2[j], \cdots, s_K[j]\}$ is called the desired function with the form as

$$f(s_1[j], s_2[j], \cdots, s_K[j]) = f(\mathbf{s}[j]), \quad (1)$$

where $\mathbf{s}[j]$ is independently drawn from $p_{\mathbf{s_r}}(\cdot)$. Every function $f(\mathbf{s}[j])$ is seen as a realization of $f(\mathbf{s_r})$. Thus, the fusion center computes T_d desired functions when each node gets data from each random source for T_d times.

Remark 1 (Typical Desired Functions). As studied in [19], CoMAC is designed to compute different types of desired functions. There are two typical functions that we focus on. A function $f(\mathbf{s}[j])$ whose values are in the set $\{\sum_{i=1}^{K} a_{1,i} s_i[j], \cdots, \sum_{i=1}^{K} a_{L_s,i} s_i[j]\}$ is called the arithmetic sum function, where $a_{l,i} \in \mathbb{R}$ is the weighting factor for node i, and L_s belongs to $\mathbb{N}$. The arithmetic sum function is a weighted sum function, which includes the mean function for all K nodes $f(\mathbf{s}[j]) = \frac{1}{K} \sum_{i=1}^{K} s_i[j]$ and the function for the active node only $f(\mathbf{s}[j]) = \{s_1[j], s_2[j], \cdots, s_K[j]\}$ as special cases. Apart from this, a function $f(\mathbf{s}[j])$ with values in the set of $\{\sum_{i=1}^{K} \mathbf{1}_{s_i[j]=0}, \cdots, \sum_{i=1}^{K} \mathbf{1}_{s_i[j]=p}\}$ is regarded as the type function where $\mathbf{1}_{(\cdot)}$ denotes the indicator function. As pointed out in [30], any symmetric function such as mean, variance, maximum, minimum and median can be attained from the type function.

In order to be robust against noise, we use sequences of nested lattice codes [15] throughout this paper. Let $\mathbf{s}_i$ denote the data vector for the i-th node whose length is T_d. After encoding $\mathbf{s}_i$, it is mapped to $\mathbf{x}_i = [x_i[1], x_i[2], \cdots, x_i[n]]$ as the length-n transmitted vector for node i. At the fusion center, the received vector with length n is given by $\mathbf{y} = [y[1], y[2], \cdots, y[n]]$ and decoded into T_d desired functions.

Considering the block code with length n, the definition of computation rate [5], [16], [17], [19] can be given as follows.

Definition 2 (Computation rate). The computation rate specifies how many function values can be computed per channel use within a predefined accuracy. It can be written as $R = \lim_{n \to \infty} \frac{T_d}{n} H(f(\mathbf{s_r}))$ where T_d is the number of function
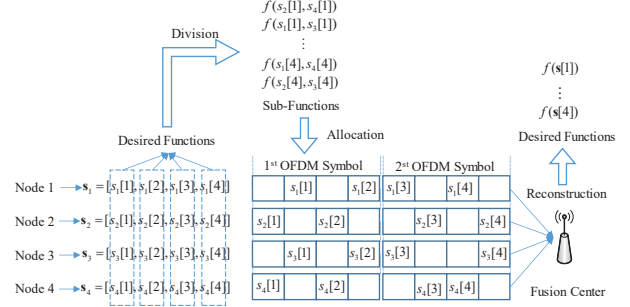

Fig. 2. Framework of wide-band CoMAC

values, n is the length of the block code and $H(f(\mathbf{s_r}))$ is the entropy of $f(\mathbf{s_r})$. Apart from this, R is achievable only if there is a length-n block code so that the probability $\Pr\left(\bigcup_{j=1}^{T_d} \left\{\hat{f}(\mathbf{s}[j] \neq f(\mathbf{s}[j]))\right\}\right) \to 0$ as n increases.

B. Wide-Band CoMAC

As shown in Fig. 2, we provide a simplified explanation on the implementation of WB-CoMAC. Each node draws data from the corresponding random source and obtains a length-four data vector. At the fusion center, four desired functions needs to be computed. Unlike NB-CoMAC, only part of nodes participate in the computation in each sub-carrier. Thus, the four desired functions are divided into eight sub-functions, and allocated to eight sub-carriers. The fusion center receives a superposition of the OFDM symbols and obtains eight sub-functions. At last, the desired functions are reconstructed by these sub-functions.

The sub-function is only part of the desired function which is computed by a subset of nodes. Assuming that a sub-function is computed by M chosen nodes and the number of all nodes is K, the desired function is split into $B = \frac{K}{M}$ parts.

Definition 3 (Sub-Function). Let

$$\tau_u = \{x \in [1 : K] : |\tau_u| = M\} \quad (2)$$

denote a set where each element x is the index from the M chosen nodes. Suppose that $\bigcup_{u=1}^{B} \tau_u = [1 : K]$ and $\tau_u \bigcap \tau_v = \emptyset$ for all $u, v \in [1 : B]$, a function $f\left(\{s_i[j]\}_{i \in \tau_u}\right)$ is said to be a sub-function if and only if there exists a function $f_c(\cdot)$ satisfying $f(\mathbf{s}[j]) = f_c(\{f(\{s_i[j]\}_{i \in \tau_u})\}_{u \in [1:B]})$.

Remark 2 (Detachable functions). From the definition of the sub-function, both arithmetic sum functions and type functions are detachable functions. Assuming K nodes are divided into $B = 4$ parts, we have $f(\mathbf{s}[j]) = \sum_{i=1}^{K} a_i s_i[j] = \sum_{u=1}^{4} \sum_{i \in \tau_u} a_i s_i[j]$ where $\sum_{i \in \tau_u} a_i s_i[j]$ represents a sub-function. Also, for the type function, we have $f(\mathbf{s}[j]) = \sum_{i=1}^{K} \mathbf{1}_{s_i[j]=b} = \sum_{u=1}^{4} \sum_{i \in \tau_u} \mathbf{1}_{s_i[j]=b}$ where $\sum_{i \in \tau_u} \mathbf{1}_{s_i[j]=b}$ represents a sub-function. To restructuring the desired function $f(\mathbf{s}[j])$, the fusion center has to obtain the four sub-functions first.

C. Existing Results

The computation rates of NB-CoMAC and WB-CoMAC are given as follows.

Theorem 1 (Rate of NB-CoMAC). As shown in [17, Theorem 1], for any $M, B \in \mathbb{N}$ satisfying $MB = K$, the ergodic computation rate of NB-CoMAC is given by

$$R = \frac{1}{B} \mathsf{E} \left[\mathsf{C}^+ \left(\frac{1}{M} + \frac{|h_{\mathcal{I}_M}|^2 KP}{M \mathsf{E} \left[\frac{|h_{\mathcal{I}_M}|^2}{|h|^2} \right]} \right) \right], \qquad (3)$$

where K is the number of nodes, M is the number of the chosen nodes to compute a sub-function, $|h_{\mathcal{I}_M}|^2$ is the channel gain of the $\mathcal{I}_M$-th node, $\mathcal{I}_M$ is the M-th element of the set of the ordered indexes of channel gains $\{\mathcal{I}_i\}_{i \in [1:K]}$ such that $|h_{\mathcal{I}_1}|^2 \geq |h_{\mathcal{I}_2}|^2 \geq \cdots \geq |h_{\mathcal{I}_K}|^2$ and h is used to represent a representative coefficient without loss of generality..

Theorem 1 considered the NB-CoMAC with flat fading, [20] expanded it to a WB-CoMAC with frequency-selective fading to focus on high-speed transmission.

Theorem 2 (Rate of WB-CoMAC). As mentioned in [20, Corollary 2 and Eq. (27)], for any $M, N \in \mathbb{N}$ satisfying $K = BM$, the ergodic computation rate of WB-CoMAC over frequency selective fading MAC is given by

$$R = \frac{1}{BN} \mathsf{E} \left[\sum_{g=1}^{N} \mathsf{C}^+ \left(\frac{N}{M} + \frac{KP|h_{\mathcal{I}_M^g}|^2}{M \mathsf{E} \left[\frac{|h_{\mathcal{I}_M^g}|^2}{|h|^2} \right]} \right) \right], \quad (4)$$

where N is the number of sub-carriers, $|h_{\mathcal{I}_M^g}|^2$ is the channel gain of the $\mathcal{I}_M^g$-th node at the g-th sub-carrier and $\mathcal{I}_M^g$ is the M-th element of the set of ordered indexes of channel gains $\{\mathcal{I}_i^g\}_{i \in [1:K]}$ at the g-th sub-carrier such that $|h_{\mathcal{I}_1^g}|^2 \geq |h_{\mathcal{I}_2^g}|^2 \geq \cdots \geq |h_{\mathcal{I}_K^g}|^2$.

One sees that both CoMAC schemes in the above only consider the use of OMA to transmit a function in each RB. This results in low spectrum efficiency. Since NOMA can offer extra improvement in spectrum efficiency, we apply NOMA to CoMAC to improve the computation rate. Thus, we propose a NOMA-CoMAC system with sub-function superposition, where each sub-carrier can serve these sub-functions with large equivalent channel gains simultaneously. It can not only achieve a higher computation rate but also can provide an improved non-vanishing rate with massive nodes.

III. SYSTEM MODEL & MAIN RESULTS

In this section, we first provide the system model of NOMA-CoMAC, and them summarize the main results of our paper. The comparison between NOMA-CoMAC and existing results is further presented.

A. Novel NOMA for Wide-Band MAC

The framework of WB-CoMAC discussed in Section II-B will be used to transmit multiple functions simultaneously in each sub-carrier using NOMA. We consider an OFDM-based system with N sub-carriers during T_s OFDM symbols while the length of the block code is n. In each sub-carrier, L functions are chosen to be transmitted. Then, the m-th received OFDM symbol at the fusion center can be expressed as

$$\mathbf{Y}[m] = \sum_{l=1}^{L} \sum_{i=1}^{K} \mathbf{V}_i^l[m] \mathbf{X}_i^l[m] \mathbf{H}_i[m] + \mathbf{W}[m], \qquad (5)$$

where $m \in [1 : T_s]$, $T_s = \lceil \frac{n}{N} \rceil$, K is the number of nodes, the power allocation matrix of node i is $\mathbf{V}_i^l[m] = \mathrm{diag}\{v_{i,1}^l[m], \cdots, v_{i,N}^l[m]\}$ whose diagonal element is the power allocated to compute the l-th function at each sub-carrier, $\mathbf{X}_i^l[m] = \mathrm{diag}\{x_{i,1}^l[m], x_{i,2}^l[m], \cdots, x_{i,N}^l[m]\}$ is the transmitted diagonal matrix of node i to compute the l-th function, a diagonal matrix $\mathbf{H}_i[m] = \mathrm{diag}\{h_{i,1}[m], \cdots h_{i,N}[m]\}$ is the channel response matrix of which the diagonal element is the channel response of each sub-carrier for node i and the diagonal element of $\mathbf{W}[m]$ is identically and independently distributed (i.i.d.) complex Gaussian random noise following $\mathcal{CN}(0, 1)$.

Assume the frequency synchronization and the phase synchronization are carried out based on the schemes proposed in [21], [22] respectively and the carrier frequency offset is solved by the scheme proposed in [26]. Based on Eq. (5), the received signal in the g-th sub-carrier at the m-th OFDM symbol can be given as

$$y_g[m] = \sum_{l=1}^{L} \sum_{i=1}^{K} v_{i,g}^l[m] x_{i,g}^l[m] h_{i,g}[m] + w_{i,g}[m], \quad (6)$$

where $v_{i,g}^l[m]$ is the power of node i allocated in the sub-carrier g for computing the l-th function, $x_{i,g}^l[m]$ is the transmitted symbol of node i for the l-th function in the g-th sub-carrier from the transmitted vector $\mathbf{x}_i$, $h_{i,g}[m]$ is the channel response of the sub-carrier g for node i and $w_{i,g}[m]$ is i.i.d. complex Gaussian random noise following $\mathcal{CN}(0, 1)$.

$$R = \frac{ML}{KNT_s} \sum_{m=1}^{T_s} \left[\sum_{g=1}^{N} \min_{i \in [1:L]} \left[\mathsf{C}^+ \left(\frac{N}{M} + \frac{N \min_{u \in \mathcal{M}_l} \left[|h_{\mathcal{I}_u^g}[m]|^2 P_{\mathcal{I}_u^g}[m] \right]}{1 + N \sum_{j=i+1}^{L} \min_{u \in \mathcal{M}_j} \left[|h_{\mathcal{I}_u^g}[m]|^2 P_{\mathcal{I}_u^g}[m] \right]} \right) \right] \right] \qquad (7)$$

B. Main Results

In the OFDM-based system, all nodes are sorted by their channel gains in each sub-carrier, and the ordered nodes are divided into $B = \frac{K}{M} \in \mathbb{N}$ parts to compute B sub-functions. Only the first $L \leq B$ sub-functions with large equivalent channel gains are chosen to be superposed in a sub-carrier. Then, the computation rate of NOMA-CoMAC is achievable with the limit of large n.

Theorem 3 (Rate of NOMA-CoMAC). For any $M, L, N, K \in \mathbb{N}$ satisfying $L \leq B$ and $K = BM$, the ergodic computation rate of NOMA-CoMAC over wide-band MAC is given as Eq. (7), where K is the number of nodes, M is the number of chosen nodes to compute a sub-function, L is the number of chosen sub-functions in each sub-carrier, N is the number of sub-carriers, $T_s = \lceil \frac{n}{N} \rceil$ is the number of OFDM symbols, $\mathcal{M}_x = [M(x-1)+1 : Mx]$ is a set including the indexes of the chosen nodes to compute the corresponding sub-function, $\mathcal{I}_u^g[m]$ is the u-th index of the ordered indexes of K nodes in the g-th sub-carrier at the-m OFDM symbol, $|h_{\mathcal{I}_u^g}[m]|^2$ is the channel gain of the $\mathcal{I}_u^g[m]$-th node in the g-th sub-carrier at the m-th OFDM symbol and $P_{\mathcal{I}_u^g}[m]$ is the power allocated to node $\mathcal{I}_u^g[m]$.

Proof: Please refer to Section IV for proof. ∎

Remark 3 (Property of NOMA-CoMAC). Theorem 3 presents a general rate that can be used with power control. It shows that the rate of NOMA-CoMAC is determined by the sub-function with the slowest rate among the L sub-functions in every sub-carrier since the desired function cannot be reconstructed unless all sub-functions are received at the fusion center.

Remark 4 (Generalization of rates for NB-CoMAC and WB-CoMAC). The rate of NOMA-CoMAC in Theorem 3 generalizes the rates of NB-CoMAC and WB-CoMAC. By setting $L = 1$ in Theorem 3, we can obtain the rate of WB-CoMAC in Theorem 2. The rate of NB-CoMAC in Theorem 1 can be also obtained by setting $L = 1, N = 1$ in Theorem 3.

In the proposed scheme, we choose the first L sub-functions with large channel gains in each sub-carrier. Since the superposition transmission of too many sub-functions makes SIC at the fusion center difficult, only two sub-functions ($L = 2$) as a pair are chosen to be transmitted in a single sub-carrier.

Corollary 1 (Rate of NOMA-CoMAC with average power control). Considering an OFDM-based system where each sub-carrier serves a sub-function pair ($L = 2$), the ergodic computation rate of NOMA-CoMAC considering the average power control, i.e, $\mathsf{E}\left[P_{i,g}[m]\right] \leq \frac{P}{N}$ for node i, can be obtained as

$$R = \frac{2}{BN}\mathsf{E}\left[\sum_{g=1}^N \mathsf{C}^+\left(\frac{N}{M} + \frac{2P\frac{K}{M}|h_{\mathcal{I}_M^g}|^2}{\Gamma + \sqrt{\Gamma^2 + 4P\frac{K}{M}|h_{\mathcal{I}_M^g}|^2\varpi_{1,g}}}\right)\right],$$
$$(8)$$

where $\Gamma = \frac{|h_{\mathcal{I}_M^g}|^2}{|h_{\mathcal{I}_{2M}^g}|^2}\varpi_{2,g} + \varpi_{1,g}$ and $\varpi_{l,g} = \mathsf{E}\left[\frac{|h_{\mathcal{I}_{Ml}^g}|^2}{|h|^2}\right]$.

Proof: Please refer to Section V and Appendix A for proof. ∎

Corollary 1 provides an easy way to allocate the power into each sub-function when average power control is considered since the power allocated to each node can be calculated directly using the closed-form expression with low computational complexity.

Similar to the previous works, the rate of NOMA-CoMAC in Corollary 1 can also prevent the rate from vanishing as the number of nodes K increases. Nevertheless, the previous works only verified the non-vanishing rate through simulation and did not obtain its exact value through mathematical analysis. We characterize the lower bound of the computation rate as the limiting rate. It can be used to calculate the accurate value of the non-vanishing computation rate with given parameters.

Corollary 2 (Limiting Rate of NOMA-CoMAC). As K increases, the computation rate of NOMA-CoMAC approaches an exact value which is only determined by $r = \frac{M}{K}$ and can be given as

$$R(r) = 2r\mathsf{C}^+\left(\frac{N}{rK} + \frac{2P\xi_{1-r}\xi_{1-2r}}{r\Delta_{\mathrm{M}} + \sqrt{(r\Delta_{\mathrm{M}})^2 + 4r\varpi_1 P\xi_{1-r}\xi_{1-2r}^2}}\right),$$
$$(9)$$

where $\Delta_{\mathrm{M}} = \varpi_1\xi_{1-2r} + \varpi_2\xi_{1-r}$, $\varpi_l = \mathsf{E}\left[\frac{|h_{\mathcal{I}_{Ml}}|^2}{|h|^2}\right]$, $F_{|h|^2}(\xi_x) = x$ and $F_{|h|^2}$ is the cumulative distribution function (CDF) of $|h|^2$. For i.i.d. Rayleigh fading, $F_{|h|^2}$ is the CDF of the exponential distribution with parameter one, i.e., $F_{|h|^2} = 1 - \exp(-x)$ and $\xi_x = -\ln(1-x)$.

Proof: Please refer to Appendix B for proof. ∎

Remark 5 (Determination of the Number of Nodes in Each Sub-Carrier). Making M_g, the number of nodes in the g-th sub-carrier, adaptive that maximizes the computation rate is

TABLE I
SUMMARY OF RATES OF NB-CoMAC, WB-CoMAC AND NOMA-CoMAC

CoMAC Scheme	Achievable Rate	Limiting Rate
NB-CoMAC	$R = \frac{1}{B}\mathsf{E}\left[\mathsf{C}^+\left(\frac{1}{M} + \frac{|h_{\mathcal{I}_M}|^2 KP}{M\varpi_1}\right)\right]$ [17]	$R(r) = r\mathsf{C}^+\left(\frac{1}{rK} + \frac{\xi_{1-r}P}{r\varpi_1}\right)$
WB-CoMAC	$R = \frac{1}{BN}\mathsf{E}\left[\sum_{g=1}^N \mathsf{C}^+\left(\frac{N}{M} + \frac{KP|h_{\mathcal{I}_M^g}|^2}{M\varpi_{1,g}}\right)\right]$ [20]	$R(r) = r\mathsf{C}^+\left(\frac{N}{rK} + \frac{\xi_{1-r}P}{r\varpi_1}\right)$
NOMA-CoMAC	$R = \frac{2}{BN}\mathsf{E}\left[\sum_{g=1}^N \mathsf{C}^+\left(\frac{N}{M} + \frac{2P\frac{K}{M}|h_{\mathcal{I}_M^g}|^2}{\Gamma + \sqrt{\Gamma^2 + 4P\frac{K}{M}|h_{\mathcal{I}_M^g}|^2\varpi_{1,g}}}\right)\right]$	$R(r) = 2r\mathsf{C}^+\left(\frac{N}{rK} + \frac{2P\xi_{1-r}\xi_{1-2r}}{r\Delta_{\mathrm{M}} + \sqrt{(r\Delta_{\mathrm{M}})^2 + 4r\varpi_1 P\xi_{1-r}\xi_{1-2r}^2}}\right)$

quite challenging because it depends on K and P as well as the ordered channel gains. Such an integer optimization problems is combinatorial in nature and therefore hard to solve for large-scale problems. To make the problem more tractable, we only consider the case where each M_g is equal to the same M and derive the optimal M as K is extremely large by setting the first derivative of $R(r)$ in Corollary 2 to zero with respect to r. It is shown that the proposed NOMA-CoMAC attains a non-vanishing computation rate even if K tends to infinity.

Note that previous works only proved that the computation rate was non-vanishing through simulation. Corollary 2 provides the lower bound of the computation rate of NOMA-CoMAC, which is easier to evaluate the performance. Using a similar proof of Corollary 2, we can obtain the limiting rates of WB-CoMAC and NB-CoMAC.

Remark 6 (Limiting Rates for NB-CoMAC and WB-Co-MAC)**.** No exact lower bound of the computation rates and their limiting rates are available in previous works. Hence, we derive the exact expression of these limiting rates, which can calculate the exact values of these non-vanishing rates with given parameters. Following a similar proof, the limiting rate of WB-CoMAC in Theorem 2 can be obtained easily as

$$R(r) = r\mathsf{C}^+ \left(\frac{N}{rK} + \frac{\xi_{1-r}P}{r\varpi_1} \right). \qquad (10)$$

It also generalizes the limiting rate of NB-CoMAC in Theorem 1 as $N = 1$. Unlike conventional works with respect to a series of random variables and M, these limiting rates are only determined by M(or r).

In conclusion, we summarize these achievable computation rates and limiting rates in Table I.

IV. PROPOSED NOMA-CoMAC SCHEME

In this section, we first introduce our framework of NOMA-CoMAC through sub-function superposition. Based on the proposed scheme, we further derive the computation rate step by step.

A. Proposed Scheme

As shown in Fig. 3, we provide a simplified description of the proposed scheme in a hybrid OFDM-NOMA system.

- *Sub-Function Process.* In each sub-carrier, we sort all the nodes depending on the corresponding channel gains. Then, every M nodes in such an order compute a function which is regarded as a sub-function in Fig. 3(a). Referring to Definition 2, τ denotes the set whose elements belong to these indexes of M nodes to compute a sub-function $f\left(\{s_i[j]\}_{i \in \tau}\right)$. Then, let the set

$$\mathcal{S} = \{\tau \subseteq [1:K] : |\tau| = M\} \qquad (11)$$

include all the possible sub-functions[1], and the cardinality of $\mathcal{S}$ is $|\mathcal{S}| = \binom{K}{M}$.

---

[1]For easy presentation, we use the element $\tau \in \mathcal{S}$ stands for the sub-function $f\left(\{s_i[j]\}_{i \in \tau}\right)$ which is computed by these nodes in τ.

- *Superposition Process.* As shown in Fig. 3(a), let the worst channel gain in the sub-function stand for the equivalent channel gain of the sub-function. Then, we sort all the sub-functions in each sub-carrier according to these equivalent channel gains. Only the first L sub-functions are chosen to be simultaneously transmitted, which is seen as a superposition. Then, one possible superposition can be defined as

$$\delta = \{\tau_1, \cdots, \tau_L : \tau_u \cap \tau_v = \emptyset, |\delta| = L\}, \qquad (12)$$

where $u \neq v$, $u, v \in [1:L]$ and sub-functions $\tau_u, \tau_v \in \mathcal{S}$. All the possible superpositions are in a set

$$\mathcal{H} = \left\{ \delta \in \mathcal{S} : |\mathcal{H}| = \prod_{l=0}^{L-1} \binom{K - Ml}{M} \right\}. \qquad (13)$$

- *SIC Process.* As shown in Fig. 3(b), all the OFDM symbols are received at the fusion center. Each sub-carrier contains a superposition with L sub-functions. Through SIC given in [15], we can obtain all the sub-functions.

- *Reconstruction Process.* As mentioned in Definition 2, all the sub-functions need to be reconstructed at the fusion center. The set

$$\mathcal{X} = \left\{ \varphi = \{\delta_1, \delta_2, \cdots, \delta_D\} : \delta_u \cap \delta_v = \emptyset, \right.$$
$$\left. \bigcup_{u=1}^{D} \delta_u = [1:K], |\varphi| = D = \frac{B}{L} \right\} \qquad (14)$$

contains all the possible combinations whose element φ can reconstruct a whole desired function, and the cardinality of $\mathcal{X}$ is $|\mathcal{X}| = \prod_{l=0}^{L-1} \binom{ML-Ml}{M} \prod_{d=0}^{D-1} \binom{K-Md}{ML}$. After all the sub-functions are collected in Fig. 3(b), we can recover the desired functions by using the relationship between the sub-functions and the desired functions.

B. Computation Rate of NOMA-CoMAC

As shown in Fig. 3, the desired function is divided into $B \in \mathbb{N}$ parts, and each part can be regarded as a sub-function which is computed at the fusion center individually. In each sub-carrier, L sub-functions are chosen to be transmitted. Based on those definitions in the previous sub-section, we use the following parts to derive the computation rate step by step.

Rate of Sub-Function τ. As mentioned in Fig. 3(a), L sub-functions are chosen for the g-th sub-carrier at the m-th OFDM symbol to be transmitted. The l-th sub-function is computed by M nodes whose indexes are in the set $\{\mathcal{I}_u^g[m]\}_{u \in \mathcal{M}_l}$ at the fusion center. For the sake of fairness, we assume the bandwidth of the hybrid OFDM-NOMA system with N sub-carriers is the same as the mentioned conventional CoMAC system. The bandwidth of each sub-carrier is $\frac{1}{N}$ of the total bandwidth, and the noise power in a sub-carrier scales down with the number of sub-carriers N. Then, the computation rate of the l-th sub-function in a sub-carrier at the m-th OFDM symbol can be given as follows.

Lemma 1 (Computation Rate of a Sub-Function)**.** With the limit of large n and L sub-functions in the g-th sub-carrier,

(a) Sub-Function Process and Superposition Process.
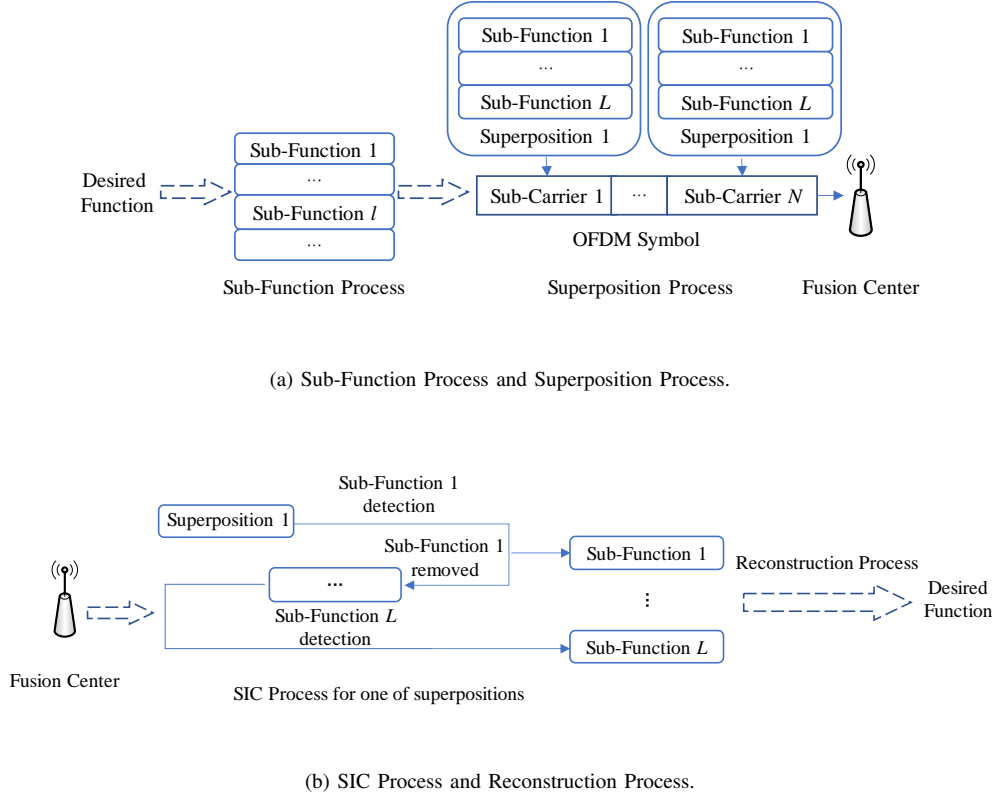

(b) SIC Process and Reconstruction Process.

Fig. 3. Design of NOMA-CoMAC

the instantaneous computation rate of the l-th sub-function at the m-th OFDM symbol with AWGN whose variance is $\frac{1}{N}$ can be express as

$$R_{l,g}[m] = \frac{1}{N}\mathsf{C}^+\left(\frac{N}{M} + \frac{N \min\limits_{u\in\mathcal{M}_l}\left[|h_{\mathcal{I}_u^g}[m]|^2 P_{\mathcal{I}_u^g}[m]\right]}{1 + N\sum\limits_{j=l+1}^{L}\min\limits_{u\in\mathcal{M}_j}\left[|h_{\mathcal{I}_u^g}[m]|^2 P_{\mathcal{I}_u^g}[m]\right]}\right), \tag{15}$$

where $|h_{\mathcal{I}_u^g}[m]|^2$ is the channel gain of the g-th sub-carrier for the node $\mathcal{I}_u^g[m]$, $P_{\mathcal{I}_u^g}[m]$ is the power allocated to the $\mathcal{I}_u^g[m]$-th node in the g-th sub-carrier and $\mathcal{M}_l$ is the set including the index of the chosen nodes to compute the l-th sub-function (See Eq. (7)).

Proof: As demonstrated in [15], CoMAC can subtract part of the contribution from the channel observation to compute several functions at the fusion center based on successive cancellation. Let $\mathbf{h}$ denote the channel vector and $\mathbf{a}_l$ denote the coefficient vector to compute the l-th function. From [15, Theorem 12], the computation rate of the l-th function from the channel observation with a noise variance of σ_Z^2 can be express as

$$R_l = \mathsf{C}^+\left(\frac{P}{|\alpha_l|^2 + P\|\alpha_l\mathbf{h} - \sum_{j=1}^{l}\mathbf{a}_j\|^2}\right), \tag{16}$$

where α_l is the scalar parameter to move the channel coefficients closer to the l-th desired function. By giving the optimal

α_l following [15, Remark 11], $\frac{1}{N}$ of the noise variance and the i-th element of the coefficient vector

$$\mathbf{a}_l[i] = \begin{cases} \mathbf{h}[i] & i \in \{\mathcal{I}_u^g[m]\}_{u\in\mathcal{M}_l}, \\ 0 & \text{otherwise} \end{cases}, \tag{17}$$

the computation rate of the l-th sub-function in single sub-carrier can be given as

$$R_l = \mathsf{C}^+\left(\frac{N}{M} + \frac{NP\|\mathbf{a}_l\|^2}{1 + NP\sum_{j=i+1}^{L}\|\mathbf{a}_j\|^2}\right). \tag{18}$$

Then combining Eq. (18) with [16, Theorem 3], the computation rate considering fading channel and power control at the t-th time slot is further expressed as

$$R_l[t] = \mathsf{C}^+\left(\frac{N}{M} + \frac{N \min\limits_{u\in\mathcal{M}_l}\left[|h_{\mathcal{I}_u}[t]|^2 P_{\mathcal{I}_u}[t]\right]}{1 + N\sum_{j=l+1}^{L}\min\limits_{u\in\mathcal{M}_j}\left[|h_{\mathcal{I}_u}[t]|^2 P_{\mathcal{I}_u}[t]\right]}\right). \tag{19}$$

Since the propagation time of a sub-carrier symbol in OFDM needs N time slots as mentioned in [20, Lemma 1], $R_{l,g}[m]$ for the l-th sub-function in the g-th sub-carrier at the m-th OFDM symbol is $\frac{1}{N}$ of $R_l[t]$. In conclusion, Lemma 1 has been proved. ∎

Rate of Superposition δ. Compared with conventional CoMAC schemes, each sub-function in the same sub-carrier has to face inter-function interference in NOMA-CoMAC, which causes the different computation rates of different sub-functions in the same sub-carrier. Lemma 1 demonstrates the

rate of the l-th function in the g-th sub-carrier is a part of the sum rate of the superposition of the L sub-functions in single sub-carrier. From Fig. 3(a), it shows that we need to determine the sum rate of all the L sub-functions in a superposition since those sub-functions are transmitted as a whole.

Lemma 2 (Computation Rate of a Superposition). As the limit of large n, the instantaneous computation rate of the superposition of L sub-functions in the g-th sub-carrier at the m-th OFDM symbol with AWGN whose variance is $\frac{1}{N}$ can be express as

$$R_{\varphi,g}^{\delta}[m] = \min_{l \in [1:L]} R_{l,g}[m] \qquad (20)$$

Proof: The rate of the superposition of L sub-functions is determined by the minimum $R_{l,g}[m]$ for all $l \in [1:L]$, since each sub-function is a part of the original desired function and the desired function can be reconstructed if and only if all parts have been received at the fusion center. ∎

Rate of Combination φ. In the hybrid OFDM-based NOMA system with N sub-carriers, the number of OFDM symbols T_s is at least $\frac{n}{N}$[2]. It also implies that the number of all the sub-carriers during T_s OFDM symbols is n, and each sub-carrier serves one superposition $\delta \in \mathcal{H}$. We define a set $\mathcal{M}_{\varphi}$ including those sub-carriers that serve the combination φ and a set $\mathcal{M}_{\varphi}^{\delta}$ containing the sub-carriers that serve the specific superposition δ in the combination φ. Since the superpositions and the combinations in practice are random depending on channel realizations, it causes that $|\mathcal{M}_{\varphi}|$ and $|\mathcal{M}_{\varphi}^{\delta}|$ are stochastic. As the limit of large n, the set $\mathcal{M}_{\varphi}^{\delta}$ and $\mathcal{M}_{\varphi}$ contain, respectively, $|\mathcal{M}_{\varphi}^{\delta}| = \frac{n}{D|\mathcal{X}|}$ sub-carriers and $|\mathcal{M}_{\varphi}| = D|\mathcal{M}_{\varphi}^{\delta}|$ sub-carriers referring to [20, Lemma 2]. Then, the transmission of the specific superposition δ in the combination φ totally occupies $T_{\varphi}^{\delta} = \lceil \frac{|\mathcal{M}_{\varphi}^{\delta}|}{N} \rceil$ OFDM symbols[3].

Lemma 3 (Computation Rate of a Combination). The average rate for computing those sub-functions in the combination φ during $T_{\varphi} = DT_{\varphi}^{\delta}$ OFDM symbols can be given as

$$R_{\varphi} = \frac{1}{T_{\varphi}} \sum_{m=1}^{T_{\varphi}^{\delta}} \frac{1}{N} \sum_{g=1}^{N} R_{\varphi,g}^{\delta}[m]. \qquad (21)$$

Proof: According to Eqs. (5) and (6), the received signal in the g-th sub-carrier for the superposition δ based on the combination φ can be given as

$$y_{\varphi,g}^{\delta}[m] = \sum_{\tau_l \in \delta} \sum_{i \in \tau_l} v_{i,\varphi,g}^{\tau_l}[m] x_{i,\varphi,g}^{\tau_l}[m] h_{i,\varphi,g}^{\tau_l}[m] + w_{i,\varphi,g}^{\tau_l}[m] \qquad (22)$$

$$= \sum_{\tau_l \in \delta} \sum_{i \in \tau_l} x_{i,\varphi,g}^{\tau_l}[m] h_{i,\varphi,g}^{'\tau_l}[m] + w_{i,\varphi,g}^{\tau_l}[m], \qquad (23)$$

where $m = [1 : T_{\varphi}^{\delta}]$, τ_l contains the indexes of the l-th sub-function in the superposition δ and $h_{i,\varphi,g}^{'\tau_l}[m] = |h_{i,\varphi,g}^{\tau_l}[m]|\sqrt{P_{i,\varphi,g}^{\tau_l}[m]}$ is equivalent channel.

---

[2]In order to simplify the derivation, the number of OFDM symbols in Eq. (5) is given as $\frac{n}{N}$ instead.

[3]Although these superpositions are sent separately in different sub-carrier in different OFMD symbol, we can consider that they are transmitted centrally when obtaining the achievable rate.

Depending on the conclusions of Lemmas 1 and 2, and Eq. (23), we can obtain the average rate for computing the superposition δ in the combination φ during T_{φ}^{δ} OFDM symbols

$$R_{\varphi}^{\delta} = \frac{1}{T_{\varphi}^{\delta}} \sum_{m=1}^{T_{\varphi}^{\delta}} \frac{1}{N} \sum_{g=1}^{N} R_{\varphi,g}^{\delta}[m]. \qquad (24)$$

Then, the average rate for computing those sub-functions in the combination φ is expressed as

$$R_{\varphi} \overset{(a)}{=} \frac{1}{D} \min_{\delta \in \varphi} R_{\varphi}^{\delta} \overset{(b)}{=} \frac{1}{D} R_{\varphi}^{\delta}, \qquad (25)$$

where condition (a) follows because of the similar result to Lemma 2 and condition (b) follows as each average rate R_{φ}^{δ} approaches the same value with the limit of large n. ∎

With the help of Eq. (21) and the length of the transmitted vector $|\mathcal{M}_{\varphi}|$, the length of the data vector is $U_{\varphi} = \frac{R_{\varphi}|\mathcal{M}_{\varphi}|}{H(f(\mathbf{s_r}))}$ as the same as the number of the desired function values reconstructed by combination φ (see Definition 2). U_{φ} is only the part of all the values of desired function T_d, and the exact number of desired function values for all $\varphi \in \mathcal{X}$ during T_s OFDM symbols is

$$T_d = \sum_{\varphi \in \mathcal{X}} U_{\varphi}. \qquad (26)$$

Finally, with the help of Lemmas 1, 2, and 3, the computation rate of NOMA-CoMAC based on Definition 2 can be given as

$$R = \lim_{n \to \infty} \frac{T_d}{n} H(f(\mathbf{s_r}))$$
$$= \lim_{n \to \infty} \frac{\sum_{\varphi \in \mathcal{X}} \frac{R_{\varphi}|\mathcal{M}_{\varphi}|}{H(f(\mathbf{s_r}))}}{n} H(f(\mathbf{s_r}))$$
$$= \frac{ML}{KN} \frac{1}{T_s} \sum_{m=1}^{T_s} \sum_{g=1}^{N} R_{\varphi,g}^{\delta}[m]. \qquad (27)$$

In conclusion, the rate in Theorem 3 is achievable as n increases.

As the number of sub-carriers and users are increasing, the estimation error of channel state information (CSI) could be more severe. In the following remark, we provide some discussion on the impact of the channel estimation error.

Remark 7 (Computation Rate With Imperfect CSI). The channel estimation error could be more severe as the number of sub-carriers and users are increasing. As pointed out in [31], the channel estimation error in the g-th sub-carrier is often measured by the mean square error (MSE), which can be defined as

$$\text{MSE}_g = \mathsf{E}\left[\frac{1}{K} \sum_{i=1}^{K} \left| h_{i,g}' - h_{i,g} \right|^2 \right]. \qquad (28)$$

We choose to characterize the estimation error as additive Gaussian noise. The estimated channel response is $h_{i,g}' = h_{i,g} + e_{i,g}$, where the estimation error $e_{i,g}$, independent from $h_{i,g}$, is a complex Gaussian random variable with zero mean

and a variance, σ_e^2, equal to MSE_g.[4] Following the similar steps of Lemma 1 and the derivation of the computation rate with imperfect CSI in [33], the rate $R_{l,g}[m]$ considering imperfect CSI can be rewritten as

$$
R'_{l,g}[m] = \frac{1}{N}\mathsf{C}^+\left(\left(\frac{N}{M} + \frac{N\min\limits_{u\in\mathcal{M}_l}\left[|h_{\mathcal{I}_u}[m]|^2 P_{\mathcal{I}_u}[m]\right]}{1+N\sum_{j=l+1}^{L}\min\limits_{u\in\mathcal{M}_j}\left[|h_{\mathcal{I}_u}[m]|^2 P_{\mathcal{I}_u}[m]\right]}\right)\right.
$$
$$
\times\left[1+\sum_{j=1}^{L}\Delta_j^2[m]\left(\frac{1}{\min\limits_{u\in\mathcal{M}_j}\left[|h_{\mathcal{I}_u}[m]|^2 P_{\mathcal{I}_u}[m]\right]M}\right.\right.
$$
$$
\left.\left.\left.+2+\min\limits_{u\in\mathcal{M}_j}\left[|h_{\mathcal{I}_u}[m]|^2 P_{\mathcal{I}_u}[m]\,M\right]\right)\right]^{-1}\right),
$$
(29)

where $\Delta_j[m] = \alpha'_j[m] - \alpha_j[m]$ is the instantaneous MMSE coefficient error and

$$
\alpha'_l = \frac{P\mathbf{h}'^{\mathrm{T}}\mathbf{a}_l}{1+P\sum_{i\in\{\mathcal{I}_u^g[m]\}_{u\in\mathcal{M}_l}}|\mathbf{h}'[i]|^2}.
\tag{30}
$$

Finally, with the help of Lemmas 1, 2, and 3, the computation rate of NOMA-CoMAC with imperfect CSI is given as

$$
R' = \frac{ML}{KNT_s}\sum_{m=1}^{T_s}\sum_{g=1}^{N}\min_{i\in[1:L]}R'_{l,g}[m].
\tag{31}
$$

One can observe that the computation rate will reduce when the variance of the estimation error, σ_e^2, increases. When the number of nodes to compute a sub-function increases, the computation rate will also reduce. For a special case with perfect CSI, i.e., $\sigma_e^2 = 0$, the MMSE coefficient error $\Delta_j[m]$ is equal to zero, leading to the rate with perfect CSI as Theorem 3. Also, by setting $L = 1$ and $N = 1$ and considering non-fading case, it is reduced to the rate in [33].

V. PERFORMANCE OF PROPOSED NOMA-CoMAC SCHEME

In order to discuss the performance of the proposed scheme, in this section, we first derive the achievable computation rate of the proposed NOMA-CoMAC with average power control based on the general rate in Theorem 3. We further analyze the outage performance and obtain the diversity order for the corresponding rate.

A. Average Power Control

We consider an average power control method where the average power of each node in each sub-carrier is no more than $\frac{P}{N}$ for all $g \in [1:N]$, i.e., $\mathsf{E}\left[P_{i,g}[m]\right] \leq \frac{P}{N}$. Let $P_{i,g}[m]$

represent the transmitted power in the g-th sub-carrier at the m-th OFDM symbol for the node i, and it can be express as

$$
P_{\mathcal{I}_i^g}[m] = \begin{cases} \beta_{l,g}[m]\dfrac{KP\dfrac{|h_{\mathcal{I}_{Ml}^g}[m]|^2}{|h_{\mathcal{I}_i^g}[m]|^2}}{N\sum_{l=1}^{L}\beta_{l,g}[m]\varpi_{l,g}} & i\in\mathcal{M}_l, \forall l\in[1:L] \\[4pt] 0 & \text{otherwise} \end{cases},
\tag{32}
$$

where $\varpi_{l,g} = \mathsf{E}\left[\frac{|h_{\mathcal{I}_{Ml}^g}[m]|^2}{|h|^2}\right] = \mathsf{E}\left[\frac{|h_{\mathcal{I}_{Ml}^g}|^2}{|h|^2}\right]$ as a constant, $\beta_{l,g}[m]$ can be regarded as the power factor to compute the l-th sub-function in the g-th sub-carrier and the detailed derivation is given in Appendix A. By putting Eq. (32) into Eq. (15), the rate of the l-th sub-function $R_{l,g}[m]$ in Lemma 1 can be rewritten as

$$
R_{l,g}[m] = \frac{1}{N}\mathsf{C}^+\left(\frac{N}{M} + \frac{PK|h_{\mathcal{I}_{Ml}^g}[m]|^2\beta_{l,g}[m]}{M\sum_{l=1}^{L}\beta_{l,g}[m]\varpi_{l,g}+\sum_{j=l+1}^{L}PK|h_{\mathcal{I}_{Mj}^g}[m]|^2\beta_{j,g}[m]}\right).
\tag{33}
$$

We work on maximizing the instantaneous rate of each OFDM symbol to improve the ergodic rate since the rate in Theorem 3 can be regarded as the mean of the instantaneous rate. Then we formulate the following optimization.

Problem 1.

$$
\underset{\beta_{l,g}[m]}{\text{maximize}}\quad \sum_{g=1}^{N}\min_{l\in[1:L]}R_{l,g}[m]
$$
$$
s.t.\quad \beta_{l,g}[m] \geq 0 \quad \forall l\in[1:L]
$$

Because the superposition transmission of too many sub-functions brings the difficulty of SIC at the fusion center and makes the mathematical analysis hard, we choose two sub-functions as a pair to be transmitted in single sub-carrier. By setting $L = 2$ in Problem 2, the relationship between $\beta_{1,g}[m]$ and $\beta_{2,g}[m]$ can be obtained as

$$
\beta_{1,g}[m] = \left(\sqrt{(\Upsilon[m]\varpi_{2,g}+\varpi_{1,g})^2 + 4P\frac{K}{M}|h_{\mathcal{I}_M^g}[m]|^2\varpi_{1,g}}\right.
$$
$$
\left.+\varpi_{1,g} - \Upsilon[m]\varpi_{2,g}\right)\frac{\beta_{2,g}[m]}{2\Upsilon[m]\varpi_{1,g}},
$$
(34)

where $\Upsilon[m] = \frac{|h_{\mathcal{I}_M^g}[m]|^2}{|h_{\mathcal{I}_{2M}^g}[m]|^2}$ and the detailed derivation is given in Appendix A.

Therefore, based on Theorem 3 and the increasing n, the ergodic computation rate of NOMA-CoMAC with a sub-function pair

$$
R \overset{(a)}{=} \frac{2M}{KNT_s}\sum_{m=1}^{T_s}\left[\sum_{g=1}^{N}R_{1,g}[m]\right] \overset{(b)}{=} \frac{2M}{KNT_s}\sum_{m=1}^{T_s}\left[\sum_{g=1}^{N}R_{2,g}[m]\right]
\tag{35}
$$

is achievable, where the conditions (a) and (b) follow because $R_{1,g}[m] = R_{2,g}[m]$ with the given optimal power allocation factors $\beta_{1,g}[m]$ and $\beta_{2,g}[m]$ in Eq. (34). As a result, Corollary 1 has been proved.

Note that the power allocated to each node for computing the corresponding sub-function in Eq. (32) is a closed-form

---

[4]To simplify the analysis, we assume MSE_g in each sub-carrier is the same one. This simplified model has also been used in the literature (for example, [32]).

expression with the given relationship between $\beta_{1,g}[m]$ and $\beta_{2,g}[m]$ in Eq. (34). It implies that the allocated power can be calculated directly with low computational complexity. Hence, the power control method is suitable to be deployed in a system that requires low latency.

We provide a signaling procedure to use the power control method as follows.

Step 1. The fusion center broadcasts a pilot signal to all nodes, and each node calculates its own CSI simultaneously.

Step 2. The fusion center collects the N channel gains from each node.

Step 3. The fusion center calculates the Ml-th channel gain $|h_{\mathcal{I}_{Ml}^g}[m]|^2$ in the ordered channel gains and obtains the set $\{\mathcal{I}_i^g[m]\}_{i \in \mathcal{M}_l}$ including M indexes with large channel gains at the g-th sub-carrier.

Step 4. The fusion center broadcasts each $|h_{\mathcal{I}_{Ml}^g}[m]|^2$ and the ID[5] of $\{\mathcal{I}_i^g[m]\}_{i \in \mathcal{M}_l}$ to all nodes for $l \in [1 : L]$.

Step 5. At the same time, each node finds each set with the corresponding ID in its memory, and calculates the transmitted power through Eq. (32) as $\beta_{l,g}[m]$ is obtained by the nodes itself.

B. Optimal Power Control

Recalling Theorem 3, we formula an optimization problem considering $L = 2$ as follows.

Problem 2.

$$\underset{P_{\mathcal{I}_i^g}[m]}{\text{maximize}} \quad \frac{2M}{KN} \sum_{g=1}^{N} \min\left[R_{1,g}[m], R_{2,g}[m]\right]$$

$$\text{s.t.} \quad \sum_{g=1}^{N} P_{i,g}[m] \leq P \quad \forall i \in [1 : K] \qquad (36)$$

.

Eq. (36) is difficult to solve, because the $\min$ function is nested inside the $\log$ function, and the ordered indexes also bring difficulty. We introduce two $K \times N$ sub-function allocation matrix, $\boldsymbol{\omega}^1$ and $\boldsymbol{\omega}^2$ to make it tractable. Each element in the sub-function allocation matrix is given as

$$\omega_{\mathcal{I}_i^g,g}^l[m] = \begin{cases} 1 & i \in \mathcal{M}_l \\ 0 & \text{otherwise} \end{cases} \quad \forall g \in [1 : N], l \in \{1, 2\}. \tag{37}$$

Furthermore, $\min_{i \in \mathcal{M}_l}\left[|h_{\mathcal{I}_i^g}[m]|^2 P_{\mathcal{I}_i^g}[m]\right]$ in Theorem 3 implies that the computation rate of the l-th sub-function is only determined by the node with the minimum product among the chosen nodes in $\mathcal{M}_l$. This means that other nodes do not need to put much power to this sub-carrier except for the node with the minimum product, and the unused power can be allocated to other sub-carriers to improve the rate. Thus, we use the level $\eta_g^l[m]$ to replace the $\min$ function by making $|h_{\mathcal{I}_i^g}[m]|^2 P_{\mathcal{I}_i^g}[m] = \eta_g^l[m]$ for all $i \in \mathcal{M}_l$ in the g-th sub-carrier.

[5]Because all the possible sets $\{\mathcal{I}_i^g[m]\}_{i \in \mathcal{M}_l}$ are known and can be stored in the memories of the nodes and the fusion center, we only need to broadcast an ID of it instead of all the elements in the set.

Then, by converting the $\max - \min$ optimization to a $\max$ optimization though relaxation, Problem 2 is rewritten as follows.

Problem 3.

$$\underset{\eta_g^1[m],\eta_g^2[m]}{\text{maximize}} \quad \frac{2M}{KN} \sum_{g=1}^{N} \mathsf{C}^+\left(\frac{N}{M} + N\eta_g^2[m]\right)$$

$$\text{s.t.} \quad \sum_{g=1}^{N} G_{i,g}\left(\eta_g^1[m]\omega_{i,g}^1[m] + \eta_g^2[m]\omega_{i,g}^2[m]\right) \leq P \quad \forall i \in [1 : K]$$

$$\mathsf{C}^+\left(\frac{N}{M} + N\frac{\eta_g^1[m]}{1 + \eta_g^2[m]}\right) \geq \mathsf{C}^+\left(\frac{N}{M} + N\eta_g^2[m]\right),$$

where $\eta_g^l[m] \geq 0$ and $P_{i,g}[m] \geq 0$ are necessary, $G_{i,g}[m] = 1/|h_{i,g}[m]|^2$, $P_{i,g}[m] = G_{i,g}[m]\eta_g^l[m]\omega_{i,g}^l[m]$ and $\omega_{i,g}^l[m]$ is the element of the sub-function allocation matrix in the m-th OFDM symbol.

Since Problem 3 is concave and has a unique maximum, we introduce the Lagrangian multipliers $\{\lambda_i\}$ to formulate the Lagrangian function. By setting the first derivative of it with respect to $\eta_g^1[m]$ and $\eta_g^2[m]$ to zero for each $g, g = 1, \cdots, N$, we have the optimal solutions as $\eta_g^{*1}[m] = \eta_g^{*2}[m] + \eta_g^{*2}[m]^2$ and

$$\eta_g^{*2}[m] = \frac{1}{4}\left(\sqrt{(1 + \varphi[m])^2 + \frac{16M}{K\varepsilon_1[m]N} - \frac{4 + 4\varphi[m]}{M} + \frac{4}{M^2}} - \frac{M + 2 + M\varphi[m]}{M}\right)^+, \tag{38}$$

where $\varepsilon_1[m] = \sum_{i=1}^{K} \lambda_i G_{i,g}[m]\omega_{i,g}^1[m]$, $\varepsilon_2[m] = \sum_{i=1}^{K} \lambda_i G_{i,g}[m]\omega_{i,g}^2[m]$ and $\varphi[m] = \varepsilon_2[m]/\varepsilon_1[m]$.

While $\eta_g^{*1}[m]$ and $\eta_g^{*2}[m]$ satisfy

$$\max_{i=1}^{K}\left[\sum_{g=1}^{N} G_{i,g}\left(\eta_g^1[m]\omega_{i,g}^1[m] + \eta_g^2[m]\omega_{i,g}^2[m]\right)\right] = P, \tag{39}$$

the optimal power in each sub-carrier of each node is

$$P_{i,g}^*[m] = G_{i,g}[m]\left(\eta_g^{*1}[m]\omega_{i,g}^1[m] + \eta_g^{*2}[m]\omega_{i,g}^2[m]\right). \tag{40}$$

C. Outage Probability and Diversity Order

It is worth obtaining the outage probability and the diversity order for NOMA-CoMAC system, which can help us to analyze the performance theoretically. An exact expression for the outage probability and the diversity order are provided in the following corollary.

Corollary 3 (Diversity Order of NOMA-CoMAC). The outage probability for NOMA-CoMAC with a sub-function pair transmission in the OFDM-based system can be given as

$$P_{out} = \left\{1 - \Theta \sum_{u_1=0}^{M-1} \sum_{u_2=0}^{K-2M} \binom{M-1}{u_1}\binom{K-2M}{u_2}(-1)^{K-2M+u_1-u_2} \right.$$

$$\times \left[\int_{\frac{\xi}{P}}^{\infty} \int_{\phi_1}^{\phi_2} \mathrm{Ex}_1(x)\mathrm{d}x\mathrm{Ex}_2(y)\mathrm{d}y \right.$$

$$\left.\left. + \int_{\frac{\varepsilon_R T}{NP\beta_{2M}}}^{\infty} \int_{\phi_3}^{\infty} \mathrm{Ex}_1(x)\mathrm{d}x\mathrm{Ex}_2(y)\mathrm{d}y\right]\right\}^N, \tag{41}$$

where $\Theta = \frac{K!}{(M-1)!(M-1)!(L-2M)!}$, $\mathrm{Ex}_1(x) = e^{-(M+u_1)x}$, $\mathrm{Ex}_2(y) = e^{-(K-M-u_1-u_2)y}$, $\phi_1 = \frac{\varepsilon_R T}{N\beta_{1,g}P} + \frac{\varepsilon_R T\beta_{2,g}}{\beta_{1,g}}|h_{\mathcal{I}_{2M}^g}|^2$,

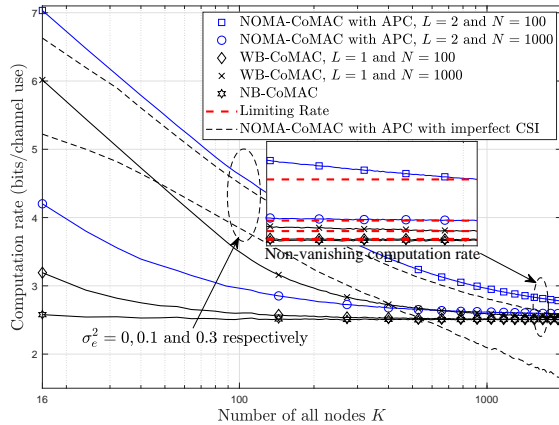Fig. 4. Comparison between NOMA-CoMAC and conventional CoMAC schemes versus the number of all nodes K, $\mathrm{SNR} = 10$ dB

Fig. 5. Computation rates of NOMA-CoMAC versus SNR and the number of sub-carriers N

$\phi_2 = \frac{\beta_{2,g}}{\beta_{1,g}}|h_{\mathcal{I}_{2M}^g}|^2 + NP\frac{\beta_{2,g}}{\beta_{1,g}}|h_{\mathcal{I}_{2M}^g}|^4$, $\phi_3 = \frac{\beta_{2,g}}{\beta_{1,g}}|h_{\mathcal{I}_{2M}^g}|^2 + NP\frac{\beta_{2,g}}{\beta_{1,g}}|h_{\mathcal{I}_{2M}^g}|^4$, $\xi = \frac{\varepsilon_{R^T}-1+\sqrt{\frac{4\varepsilon_{R^T}}{\beta_{2,g}}+(\varepsilon_{R^T}-1)^2}}{2N}$ and $\varepsilon_{R^T} = 2^{\frac{R^T KN}{2M}} - \frac{N}{M}$.

The diversity order achieved by NOMA-CoMAC with a sub-function pair transmission is given by

$$-\lim_{P\to\infty}\frac{\log P_{out}}{\log P} = \underbrace{N}_{o_1}\cdot\underbrace{(K-2M)}_{o_2} \tag{42}$$

Proof: See Appendix C. ∎

Corollary 3 illustrates that the diversity order in our proposed NOMA-CoMAC with a sub-function pair transmission consists of two parts. The first part o_1 is achievable benefiting from the OFDM design, and the superposition of the first $L = 2$ sub-functions brings the gain as the second part o_2. Furthermore, the part o_2 shows that the node with the worst channel gain in these two sub-functions plays a dominant role.

VI. NUMERICAL RESULTS AND DISCUSSION

In this section, numerical results of the computation rate based on NOMA-CoMAC are provided and compared with the conventional CoMAC schemes. Besides, the outage performance of NOMA-CoMAC is demonstrated, and the accuracy of the derived analytical results are verified through Monte Carlo simulation. For easy presentation, the abbreviation for average power control and optimal power control are APC and OPC, respectively.

A. Computation Rate

We compare the conventional CoMAC schemes, i.e., NB-CoMAC and WB-CoMAC, with NOMA-CoMAC in Fig. 4. We see that NOMA-CoMAC can improve the computation rate over the conventional CoMAC schemes as the number of nodes K is small. When K increases, all the rates shown in Fig. 4 are decreasing but keep a non-vanishing rate. Compared with the conventional CoMAC schemes, the proposed NOMA-CoMAC can provide an improved non-vanishing rate. It also
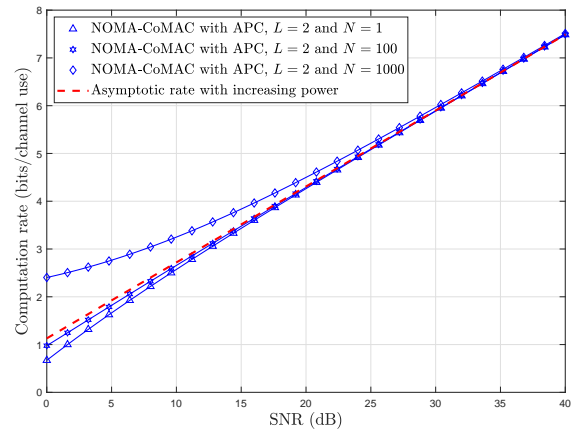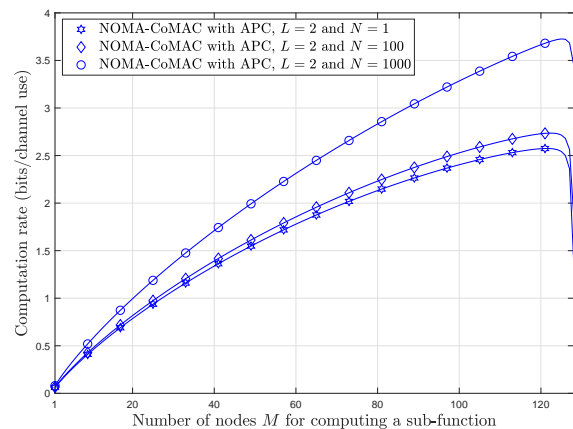

Fig. 6. Computation rates of NOMA-CoMAC versus the number of chosen nodes M, $\mathrm{SNR} = 10$ dB

verifies the limiting rates in Corollary 2 and implies that the exact non-vanishing rates as the lower bounds can be obtained without any simulation using Corollary 2. As for the impact of the estimation error, the computation rate decreases as σ_e^2 increases. When K becomes larger, the gap between the rate with perfect CSI and the one with imperfect CSI also becomes larger. This is because, for a large K, M will become larger and then enlarge the impact of estimation error (referring to Eq. (31)).

The relationship between the computation rate of NOMA-CoMAC and SNR is shown in Fig. 5. When SNR increases, all three rates with different N increase. When SNR is small, the number of sub-carriers plays the main role to improve the rate since the increasing N can reduce the equivalent noise power. However, when SNR increases, the contribution of N becomes smaller and can be ignored compared to SNR. Hence, all three rates are asymptotically equal. From Corollary 2, the asymptotic rate with increasing SNR can be given as $2r\mathsf{C}^+(P^{\frac{1}{2}}\xi_{1-r}^{\frac{1}{2}})$.

Fig. 6 illustrates that the computation rate of NOMA-CoMAC referred to Corollary 1 versus the increase of the
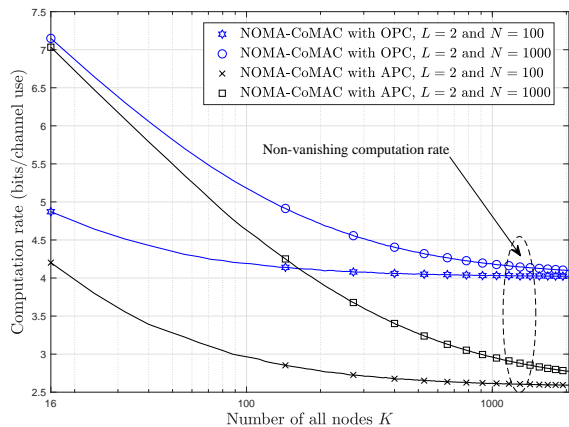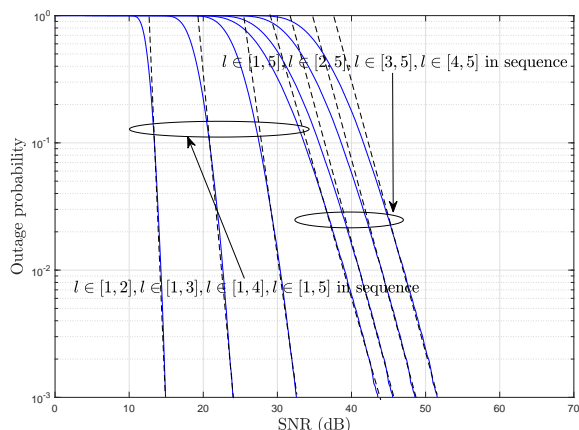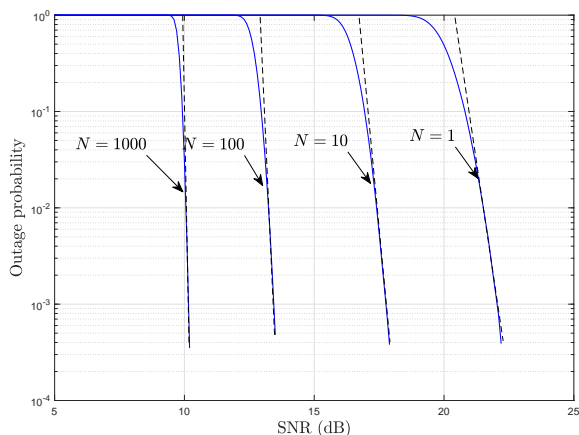
Fig. 7. Comparison of computation rates between NOMA-CoMAC with OPC and NOMA-CoMAC with APC.

(a) the outage probabilities with different sub-function pairs.

(b) the outage probabilities with different number of sub-carriers.

Fig. 8. Outage probabilities with different sub-function pairs and different number of sub-carriers.

number of chosen nodes M. It shows that the rate in Eq. (8) does not increase monotonically as M increases. The reason is that the rate gain $\frac{2M}{KN}$ increases with M, whereas the worse channel gain $|h_{\mathcal{I}_{2M}^g[m],g}[m]|^2$ becomes vanishing as M increases. Hence, there is a trade-off between the rate gain and the worse channel gain. It also implies that there is an optimal M that achieves the maximum rate. In addition, with the increase of the number of sub-carriers N, the rate can be improved because the equivalent noise power of lattice code decreases.

As shown in Fig. 7, the comparison of the computation rate between NOMA-CoMAC with OPC and NOMA-CoMAC with APC is given. When the number of nodes K increases, both rates will decrease and eventually keep stable. However, NOMA-CoMAC with OPC provides a higher rate than NOMA-CoMAC with APC. It shows that the computation rate is further improved by optimal power control and a non-vanishing computation rate is also provided as K increases.

B. Outage Performance

As mentioned in Corollary 1, NOMA-CoMAC only chooses a sub-function pair to be computed in a single sub-carrier. Fig. 8(a) shows the outage performance of different sub-function pairs with the same first sub-function including $l \in [1,2], l \in [1,3], l \in [1,4], l \in [1,5]$. One can observe that the outage performance of the sub-function pair $l \in [1,2]$ is always superior to that of other sub-function pairs. Furthermore, the outage performance of different sub-function pairs with the same second sub-function including $l \in [1,5], l \in [2,5], l \in [3,5], l \in [4,5]$ is also demonstrated in Fig. 8(a). We can see that the outage probabilities for all five sub-function pairs have the same diversity order while the sub-function pair $l \in [1,5]$ offers a constant performance gain over the others. As shown in Fig. 8(b), both the diversity order and performance gain increase as the number of sub-carriers increases.

Based on Fig. 8, we can see that the first function in the sub-function pair determines the performance gain. The diversity order depends on the second function of the sub-function pairs. Hence, the sub-function pair chosen in Corollary 1 outperforms any other sub-function pair. Besides, the figures above verify Corollary 3 where the diversity order is only associated with the sub-function with poorer equivalent channel gain in the sub-function pair, and the diversity order increases as N increases.

VII. CONCLUSION

In order to provide extra improvement in spectrum efficiency, we have proposed a NOMA-CoMAC system through the division, superposition, SIC and reconstruction of the desired functions. Unlike those NOMA designs for information transmission, the proposed NOMA-CoMAC aims at computing functions over the MAC through superposing sub-functions. The expression of achievable computation rate has been derived based on nested lattice coding. It not only improves the spectrum efficiency but also enhances the non-vanishing rate via sub-function superposition. Furthermore, we

have considered some limiting cases to find more insights. As the number of nodes goes to infinity, we have obtained the exact expression of the limiting computation rate which characterizes the lower bound of the computation rate of NOMA-CoMAC. It can be used to evaluate the performance of the CoMAC system without time-consuming simulation. We have discussed the outage performance and analyzed the diversity order as the power goes to infinity. It shows the node with the worst channel gain among these sub-functions in each RB plays a dominant role.

APPENDIX A
PROOF OF COROLLARY 1

We consider the average power control method where $\mathsf{E}\left[P_{i,g}[m]\right] \le \frac{P}{N}$. Let $P_{i,g}[m]$ represent the transmitted power in the g-th sub-carrier at the m-th OFDM symbol for the node i, which can be express as

$$P_{\mathcal{I}_i^g}[m] = \begin{cases} c\beta_{l,g}[m]\dfrac{|h_{\mathcal{I}_{Ml}^g}[m]|^2}{|h_{\mathcal{I}_i^g}[m]|^2} & i \in \mathcal{M}_l, \forall l \in [1:L] \\ 0 & \text{otherwise} \end{cases} \tag{43}$$

In Eq. (43), c is a constant which is needed to be solved, and $\beta_{l,g}[m]$ can be regarded as the power factor to compute the l-th sub-function in the g-th sub-carrier. With the average power control, we have

$$\mathsf{E}[P_{i,g}[m]] = \sum_{j=1}^{K}\Pr(i = \mathcal{I}_j^g[m])\mathsf{E}[P_{i,g}[m]|i = \mathcal{I}_j^g[m]]$$
$$\overset{(a)}{=} c\frac{M}{K}\sum_{l=1}^{L}\beta_{l,g}[m]\underbrace{\mathsf{E}\left[\frac{|h_{\mathcal{I}_{Ml}^g}[m]|^2}{|h|^2}\right]}_{\varpi_{l,g}} = \frac{P}{N}, \quad (44)$$

where condition (a) follows because the channel gains in different sub-carriers are i.i.d. random variables and h is used as a representative random variable without loss of generality. Then we can calculate the constant c and obtain Eq. (32).

Assuming that the channel gain in each sub-carrier is i.i.d., and we further convert the $\max-\min$ problem shown as Problem 2 into a $\max$ problem. A new form of the optimization is given in the following.

Problem 4.

$$\begin{aligned} \text{maximize} \quad & \beta_{L,g}[m] \\ \text{s.t.} \quad & R_{1,g}[m] = R_{2,g}[m] = \cdots = R_{L,g}[m] \\ & \beta_{l,g}[m] \ge 0 \quad \forall l \in [1:L] \end{aligned} \tag{45}$$

By setting $L = 2$, the relationship between $\beta_{1,g}[m]$ and $\beta_{2,g}[m]$ can be calculated through quadratic formula by solving $R_{1,g}[m] = R_{2,g}[m]$. Then, the relationship between $\beta_{1,g}[m]$ and $\beta_{2,g}[m]$ in Eq. (34) can be further obtained.

APPENDIX B
PROOF OF COROLLARY 2

Let $\frac{M}{K} = r$ and $\varpi_l = \mathsf{E}\left[\frac{|h_{\mathcal{I}_{Ml}}|^2}{|h|^2}\right]$. The limiting rate of NB-CoMAC in Theorem 1 is given as

$$\begin{aligned} R &= \frac{1}{B}\mathsf{E}\left[\mathsf{C}^+\left(\frac{1}{M} + \frac{|h_{\mathcal{I}_M}|^2 KP}{M\mathsf{E}\left[\frac{|h_{\mathcal{I}_M}|^2}{|h|^2}\right]}\right)\right] \\ &= r\mathsf{E}\left[\mathsf{C}^+\left(\frac{1}{rK} + \frac{\mathcal{N}_r P}{r\varpi_1}\right)\right] \\ &\overset{(a)}{=} r\mathsf{C}^+\left(\frac{1}{rK} + \frac{\xi_{1-r}P}{r\varpi_1}\right), \end{aligned} \tag{46}$$

where $\mathcal{N}_r$ asymptotically follows the Gaussian distribution with mean ξ_{1-r} and variance $\frac{r(1-r)}{Kf_{|h|^2}^2(\xi_{1-r})}$, $F_{|h|^2}(\xi_{1-r}) = 1-r$ [34] and condition (a) holds while the variance of $\mathcal{N}_r$ approaches 0.

With similar steps, the limiting rate of WB-CoMAC in Theorem 2 is easy to obtain. And the limiting rate of NOMA-CoMAC with average power control in Corollary 1 can be express as

$$\begin{aligned} R &= \frac{2}{BN}\mathsf{E}\left[\sum_{g=1}^{N}\mathsf{C}^+\left(\frac{N}{M} + \frac{2P\frac{K}{M}|h_{\mathcal{I}_M^g}|^2}{\Gamma + \sqrt{\Gamma^2 + 4P\frac{K}{M}|h_{\mathcal{I}_M^g}|^2\varpi_{1,g}}}\right)\right] \\ &= 2r\mathsf{E}\left[\mathsf{C}^+\left(\frac{N}{rK} + \frac{2P\mathcal{N}_r\mathcal{N}_{2r}}{r\Delta + \sqrt{(r\Delta)^2 + 4r\varpi_1 P\mathcal{N}_r\mathcal{N}_{2r}^2}}\right)\right] \\ &= 2r\mathsf{C}^+\left(\frac{N}{rK} + \frac{2P\xi_{1-r}\xi_{1-2r}}{r\Delta_{\mathrm{M}} + \sqrt{(r\Delta_{\mathrm{M}})^2 + 4r\varpi_1 P\xi_{1-r}\xi_{1-2r}^2}}\right), \end{aligned} \tag{47}$$

where $\Gamma = \frac{|h_{\mathcal{I}_M^g}|^2}{|h_{\mathcal{I}_{2M}^g}|^2}\varpi_{2,g} + \varpi_{1,g}$, $\Delta = \varpi_1\mathcal{N}_{2r} + \varpi_2\mathcal{N}_r$ and $\Delta_{\mathrm{M}} = \mathsf{E}[\Delta] = \varpi_1\xi_{1-2r} + \varpi_2\xi_{1-r}$.

APPENDIX C
PROOF OF COROLLARY 3

Recalling Problem 2 and setting $L = 2$, the instantaneous computation rate is given as

$$R[m] = \frac{2M}{KN}\sum_{g=1}^{N}\min\left[R_{1,g}[m], R_{2,g}[m]\right] \tag{48}$$

with fixed power allocation factors $\beta_{1,g}[m]$ and $\beta_{2,g}[m]$.

Since the analysis of outage probability can be seen as a single-input and multiple-output system approximately referring to [35] in the OFDM-based system, we only aim at the outage probability of the rate in each sub-carrier P_{out}^{sub}, and the outage probability P_{out} can be evaluated as

$$\begin{aligned} P_{out} &= (P_{out}^{sub})^N \\ &= (1 - \Pr\left\{R_{1,g}[m] > R^T, R_{2,g}[m] > R_{1,g}[m]\right\} - \\ &\quad \Pr\left\{R_{2,g}[m] > R^T, R_{1,g}[m] > R_{2,g}[m]\right\})^N, \end{aligned} \tag{49}$$

where R^T is the target rate.

In the following derivation, we will ignore $[m]$ which stands for the m-th OFDM symbol, since it has no influence on the

result and makes derivation obscure. Hence, $\bar{P}_{out}^{sub} = 1 - P_{out}^{sub}$ can be further expressed as

$$\bar{P}_{out}^{sub} = \underbrace{\Pr\left\{\phi_1 < |h_{\mathcal{I}_M^g}|^2 < \phi_2, |h_{\mathcal{I}_{2M}^g}|^2 > \frac{\xi}{P}\right\}}_{\alpha_1}$$
$$+ \underbrace{\Pr\left\{|h_{\mathcal{I}_M^g}|^2 > \phi_3, |h_{\mathcal{I}_{2M}^g}|^2 > \frac{\varepsilon_{RT}}{NP\beta_{2,g}}\right\}}_{\alpha_2} \quad (50)$$

where $\phi_1 = \frac{\varepsilon_{RT}}{N\beta_{1,g}P} + \frac{\varepsilon_{RT}\beta_{2,g}}{\beta_{1,g}}|h_{\mathcal{I}_{2M}^g}|^2$, $\phi_2 = \frac{\beta_{2,g}}{\beta_{1,g}}|h_{\mathcal{I}_{2M}^g}|^2 + NP\frac{\beta_{2,g}}{\beta_{1,g}}|h_{\mathcal{I}_{2M}^g}|^4$, $\phi_3 = \frac{\beta_{2,g}}{\beta_{1,g}}|h_{\mathcal{I}_{2M}^g}|^2 + NP\frac{\beta_{2,g}}{\beta_{1,g}}|h_{\mathcal{I}_{2M}^g}|^4$, $\xi = \frac{\varepsilon_{RT}-1+\sqrt{\frac{4\varepsilon_{RT}}{\beta_{2,g}}+(\varepsilon_{RT}-1)^2}}{2N}$ and $\varepsilon_{RT} = 2^{\frac{R_T KN}{2M}} - \frac{N}{M}$.

As shown in [36] and binomial theorem, the joint probability density function of the order statistics $|h_{\mathcal{I}_M^g}|^2$ and $|h_{\mathcal{I}_{2M}^g}|^2$ is

$$f_{|h_{\mathcal{I}_M^g}|^2, |h_{\mathcal{I}_{2M}^g}|^2}(x,y) = \Theta \sum_{u_1=0}^{M-1} \sum_{u_2=0}^{K-2M} \binom{2M-1}{u_1}\binom{K-2M}{u_2}$$
$$\times (-1)^{K-2M+u_1-u_2} \mathrm{Ex}_1(x)\mathrm{Ex}_2(y), \quad (51)$$

where $\Theta = \frac{K!}{(M-1)!(M-1)!(L-2M)!}$, $\mathrm{Ex}_1(x) = e^{-(M+u_1)x}$ and $\mathrm{Ex}_2(y) = e^{-(K-M-u_1-u_2)y}$.

Since $\bar{P}_{out}^{sub}$ can be divided into two parts, the first part α_1 is evaluated as follows:

$$\alpha_1 = \Theta \sum_{u_1=0}^{M-1} \sum_{u_2=0}^{K-2M} \binom{M-1}{u_1}\binom{K-2M}{u_2}$$
$$\times (-1)^{K-2M+u_1-u_2} \int_{\frac{\xi}{P}}^{\infty} \int_{\phi_1}^{\phi_2} \mathrm{Ex}_1(x)\mathrm{d}x\,\mathrm{Ex}_2(y)\mathrm{d}y. \quad (52)$$

And the second part can be evaluated as follows:

$$\alpha_2 = \Theta \sum_{u_1=0}^{M-1} \sum_{u_2=0}^{K-2M} \binom{M-1}{u_1}\binom{K-2M}{u_2}$$
$$\times (-1)^{K-2M+u_1-u_2} \int_{\frac{\varepsilon_{RT}}{NP\beta_{2M}}}^{\infty} \int_{\phi_3}^{\infty} \mathrm{Ex}_1(x)\mathrm{d}x\,\mathrm{Ex}_2(y)\mathrm{d}y. \quad (53)$$

Substituting the above equations into Eq. (49), the outage probability P_{out} in Eq. (41) can be obtained.

In order to find the diversity gain, we focus on the case with high SNR, i.e., $P \to \infty$. First of all, a more exact expression of α_1 without the integral can be expressed as

$$\alpha_1 = \Theta \sum_{u_1=0}^{M-1} \sum_{u_2=0}^{K-2M} \binom{M-1}{u_1}\binom{K-2M}{u_2}(-1)^{K-2M+u_1-u_2}$$
$$\times \frac{1}{2\sqrt{(M+u_1)D}(M+u_1)((-B+1)u_1+(-B+1)M-K+u_2)}$$
$$\times \left(\left(\mathrm{erf}\left(\frac{(M+u_1)C+2DE(M+u_1)+K-M-u_1-u_2}{2\sqrt{(M+u_1)D}}\right)-1\right)\right.$$
$$\times ((-B+1)u_1+(-B+1)M-K+u_2)e^{\frac{((C-1)u_1+(C-1)M+K-u_2)^2}{4(M+u_1)D}}\sqrt{\pi}$$
$$\left.\times -2e^{((-B+1)u_1+(-B+1)M-K+u_2)E-(M+u_1)A}\sqrt{(M+u_1)D}\right), \quad (54)$$

where $A = \frac{\varepsilon_{RT}}{N\beta_{1,g}P}$, $B = \frac{\varepsilon_{RT}\beta_{2,g}}{\beta_{1,g}}$, $C = \frac{\beta_{2,g}}{\beta_{1,g}}$, $D = NP\frac{\beta_{2,g}}{\beta_{1,g}}$ and $E = \frac{\xi}{P}$. With the case as $P \to \infty$, $A \to 0$, $E \to 0$ and $D \to \infty$. Applying Taylor expansion of the exponential functions, α_1 can be further given as

$$\alpha_1 = \Theta \sum_{u_1=0}^{M-1} \sum_{u_2=0}^{K-2M} \binom{M-1}{u_1}\binom{K-2M}{u_2}(-1)^{K-2M+u_1-u_2+1}$$
$$\times \frac{\left(\frac{\sqrt{\pi}((B-1)u_1+(B-1)M-u_2+K)}{\sqrt{(M+u_1)D}}\sum_{n=0}^{\infty}\frac{((C-1)u_1+(C-1)M+K-u_2)^{2n}}{n!(4(M+u_1)D)^n}\right)}{2((B-1)u_1+(B-1)M-u_2+K)(M+u_1)} \quad (55)$$

By using the property of CDF of $|h_{\mathcal{I}_{2M}^g}|^2$ as $F_{|h_{\mathcal{I}_M^g}|^2,|h_{\mathcal{I}_{2M}^g}|^2}(\infty,\infty) = 1$ and the binomial theorem, (55) can be expressed as follows:

$$\alpha_1 = \frac{1}{2}\left(1 + \Theta \sum_{u_1=0}^{M-1}\binom{M-1}{u_1}(-1)^{K-2M+u_1+1}\frac{\sqrt{\pi}((B-1)u_1+(B-1)M-u_2+K)}{((B-1)u_1+(B-1)M-u_2+K)(M+u_1)}\right.$$
$$\times \sum_{n=0}^{\infty}\frac{1}{n!(4(M+u_1)D)^n}\sum_{u_2=0}^{K-2M}(-1)^{u_2}\binom{K-2M}{u_2}$$
$$\left.\times \sum_{k=0}^{2n-1}\binom{2n-1}{k}((C-1)u_1+(C-1)M+K-u_2)^{2n-1-k}(-1)^k u_2^k\right). \quad (56)$$

Recalling two sums of the binomial coefficients from [37, Eq. (0.153)] and [37, Eq. (0.154)], we can have the following expression

$$\alpha_1 \approx \frac{1}{2} - \Theta\sum_{u_1=0}^{M-1}\binom{M-1}{u_1}\frac{((C-1)u_1+(C-1)M+K-u_2)^{2(K-2M)}}{(K-2M)!(4(M+u_1)D)^{(K-2M)}}, \quad (57)$$

since the factor with u_2^k, $k < (K-2M)$ is equal to zero by using the above equations. Furthermore, the terms with u_2^k, $k > (K-2M)$ can be also removed as the dominant factor is $k = K-2M$. Following steps similar to the ones for obtaining Eq. (57), we can prove that the dominant factor is also $k = K-2M$ in α_2. In conclusion, the diversity gain is obtained.

REFERENCES

[1] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of things: A survey on enabling technologies, protocols, and applications," *IEEE Commun. Surv. Tutorials*, vol. 17, no. 4, pp. 2347–2376, 2015.

[2] M. Vaezi, Z. Ding, and H. V. Poor, *Multiple access techniques for 5G wireless networks and beyond*. Springer, 2019.

[3] G. Zhu, S. Ko, and K. Huang, "Inference from randomized transmissions by many backscatter sensors," *IEEE Trans. Wireless Commun.*, vol. 17, no. 5, pp. 3111–3127, May 2018.

[4] S. Ko, K. Han, and K. Huang, "Wireless networks for mobile edge computing: Spatial modeling and latency analysis," *IEEE Trans. Wireless Commun.*, vol. 17, no. 8, pp. 5225–5240, Aug 2018.

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

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

[7] O. Abari, H. Rahul, and D. Katabi, "Over-the-air function computation in sensor networks," *arXiv preprint arXiv:1612.02307*, 2016.

[8] M. Goldenbaum and S. Stanczak, "On the channel estimation effort for analog computation over wireless multiple-access channels," *IEEE Wireless Commun. Lett.*, vol. 3, no. 3, pp. 261–264, 2014.

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

[10] 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, 2018.

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

[12] R. Appuswamy and M. Franceschetti, "Computing linear functions by linear coding over networks," *IEEE Trans. Inf. Theory*, vol. 60, no. 1, pp. 422–431, 2014.

[13] U. Erez, S. Litsyn, and R. Zamir, "Lattices which are good for (almost) everything," *IEEE Trans. Inf. Theory*, vol. 51, no. 10, pp. 3401–3416, 2005.

[14] 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, 2018.

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

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

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

[18] C. Y. Wang, S. W. Jeon, and M. Gastpar, "Interactive computation of type-threshold functions in collocated Gaussian networks," *IEEE Trans. Inf. Theory*, vol. 61, no. 9, pp. 4765–4775, 2015.

[19] M. Goldenbaum, "Computation of real-valued functions over the channel in wireless sensor networks," Ph.D. dissertation, Technische Universität München, 2014.

[20] 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.

[21] O. Abari, H. Rahul, D. Katabi, and M. Pant, "Airshare: Distributed coherent transmission made seamless," in *2015 IEEE Conference on Computer Communications (INFOCOM)*. IEEE, 2015, pp. 1742–1750.

[22] 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 Trans. Veh. Technol.*, vol. 67, no. 11, pp. 10 603–10 614, 2018.

[23] G. Wunder, R. F. Fischer, H. Boche, S. Litsyn, and J.-S. No, "The PAPR problem in OFDM transmission: New directions for a long-lasting problem," *IEEE Signal Process Mag.*, vol. 30, no. 6, pp. 130–144, 2013.

[24] A. Mobasher and A. K. Khandani, "Integer-based constellation-shaping method for PAPR reduction in OFDM systems," *IEEE Trans. Commun.*, vol. 54, no. 1, pp. 119–127, 2006.

[25] S.-W. Hou and C. C. Ko, "Intercarrier interference suppression for OFDMA uplink in time-and frequency-selective fading channels," *IEEE Trans. Veh. Technol.*, vol. 58, no. 6, pp. 2741–2754, 2008.

[26] A. Farhang, N. Marchetti, L. E. Doyle, and B. Farhang-Boroujeny, "Low complexity CFO compensation in uplink OFDMA systems with receiver windowing," *IEEE Trans. Signal Process.*, vol. 63, no. 10, pp. 2546–2558, 2015.

[27] Z. Ding, X. Lei, G. K. Karagiannidis, R. Schober, J. Yuan, and V. K. Bhargava, "A survey on non-orthogonal multiple access for 5G networks: Research challenges and future trends," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 10, pp. 2181–2195, 2017.

[28] B. Li, X. Qi, K. Huang, Z. Fei, F. Zhou, and R. Q. Hu, "Security-reliability tradeoff analysis for cooperative NOMA in cognitive radio networks," *IEEE Trans. Commun.*, vol. 67, no. 1, pp. 83–96, 2018.

[29] H. Sun, F. Zhou, R. Q. Hu, and L. Hanzo, "Robust beamforming design in a NOMA cognitive radio network relying on SWIPT," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 1, pp. 142–155, 2018.

[30] A. Giridhar and P. R. Kumar, "Computing and communicating functions over sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 4, pp. 755–764, 2005.

[31] S. Ye, R. S. Blum, and L. J. Cimini, "Adaptive ofdm systems with imperfect channel state information," *IEEE Trans. Wireless Commun.*, vol. 5, no. 11, pp. 3255–3265, 2006.

[32] H. Cheon and D. Hong, "Effect of channel estimation error in ofdm-based wlan," *IEEE Commun. Lett.*, vol. 6, no. 5, pp. 190–192, 2002.

[33] K. N. Pappi, G. K. Karagiannidis, and R. Schober, "How sensitive is compute-and-forward to channel estimation errors?" in *2013 IEEE International Symposium on Information Theory*. IEEE, 2013, pp. 3110–3114.

[34] S. S. Wilks, "Order statistics," *Bulletin of the American Mathematical Society*, vol. 54, no. 1, pp. 6–50, 1948.

[35] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[36] H.-C. Yang and M.-S. Alouini, *Order Statistics in Wireless Communications: diversity, adaptation, and scheduling in MIMO and OFDM systems*. Cambridge University Press, 2011.

[37] A. Jeffrey and D. Zwillinger, *Table of Integrals, Series, and Products*. Elsevier Science, 2000.
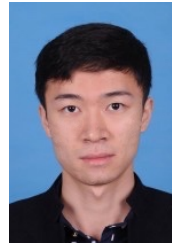
Fangzhou Wu received the B.S. degree in electronic information engineering from North China Electric Power University, Beijing, China, in 2016. He is currently pursuing the Ph.D. degree with the Department of Electronic Engineering and Information Science, CAS Key Laboratory of Wireless-Optical Communications, University of Science and Technology of China. His research interests include visible light communications, and wireless communication and computation.
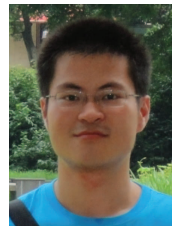
Li Chen received the B.E. in electrical and information engineering from Harbin Institute of Technology, Harbin, China, in 2009 and the Ph.D. degree in electrical engineering from the University of Science and Technology of China, Hefei, China, in 2014. He is currently an Associate Professor with the 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 Ph.D. degree in information and communication engineering in 2011, from Harbin Institute of Technology, Harbin, China.

Dr. Zhao is serving or served on the editorial boards of 7 SCI-indexed journals, including IEEE Transactions on Green Communications and Networking. He won the best paper awards in IEEE VTC 2017 Spring, MLICOM 2017, ICNC 2018, WCSP 2018 and CSPS 2018. He also received the IEEE Communications Society Asia Pacific Board Outstanding Young Researcher Award in 2018.
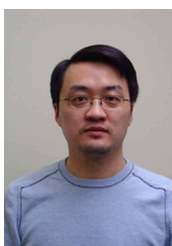
Yunfei Chen (S'02-M'06-SM'10) received his B.E. and M.E. degrees in electronics engineering from Shanghai Jiaotong University, Shanghai, P.R.China, in 1998 and 2001, respectively. He received his Ph.D. degree from the University of Alberta in 2006. He is currently working as an Associate Professor at 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, the Ontario Early Researcher Award (formerly Premiers Research Excellence Award) in 2011, the Excellent Contribution Award at IEEE/IFIP TrustCom 2010, the Leadership Opportunity Fund Award from Canada Foundation of Innovation in 2009 and the Best Paper Awards at IEEE ICNC 2018, VTC 2017 Spring, ICC 2014, Globecom 2012, IEEE/IFIP TrustCom 2009 and Int'l Conference on Networking 2005. His research interests include wireless cyber-physical systems, connected/autonomous vehicles, security, distributed ledger technology, and deep learning.

He serves on the editorial boards of several journals, including Co-Editor-in-Chief for Ad Hoc & Sensor Wireless Networks, Lead Series Editor for IEEE Transactions on Vehicular Technology, IEEE Transactions on Green Communications and Networking, and IEEE Communications Surveys & Tutorials. He has served as the Technical Program Committee (TPC) Co-Chair of numerous conferences. Dr. Yu is a registered Professional Engineer in the province of Ontario, Canada, a Fellow of the Institution of Engineering and Technology (IET), and a Fellow of the IEEE. He is a Distinguished Lecturer, the Vice President (Membership), and an elected member of the Board of Governors (BoG) of the IEEE Vehicular Technology Society.

Guo Wei received the B.S. degree in electronic engineering from the University of Science and Technology of China (USTC), Hefei, China, in 1983 and the M.S. and Ph.D. degrees in electronic engineering from the Chinese Academy of Sciences, Beijing, China, in 1986 and 1991, respectively. He is currently a Professor with the School of Information Science and Technology, USTC. His current research interests include wireless and mobile communications, wireless multimedia communications, and wireless information networks.