Received October 23, 2019, accepted November 4, 2019, date of publication November 6, 2019, date of current version November 15, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2951999

Resource Allocation Based on Interference Alignment With Clustering for Data Stream Maximization in Dense Small Cell Networks

HAO ZHANG[1,2], KUNDE YANG[1,2], (Member, IEEE), AND SHUN ZHANG[3], (Member, IEEE)

[1]School of Marine Science and Technology, Northwestern Polytechnical University, Xi'an 710072, China
[2]Key Laboratory of Ocean Acoustics and Sensing, Northwestern Polytechnical University, Ministry of Industry and Information Technology, Xi'an 710072, China
[3]State Key Laboratory of Integrated Services Networks, Xidian University, Xi'an 710071, China

Corresponding author: Kunde Yang (ykdzym@nwpu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61801397, in part by the China Postdoctoral Science Foundation under Grant 2018M643735, in part by the Natural Science Foundation of Shaanxi Province under Grant 2019JQ-624, in part by the Fundamental Research Funds for the Central Universities under Grant 3102018jcc033 and Grant 3102018bzc001, and in part by the National Natural Science Foundation of China under Grant 61871455.

ABSTRACT In this paper, by allocating the limited subchannel resources based on interference alignment (IA) with clustering to mitigate the severe co-tier interference in dense small cell networks (SCNs), we aim at maximizing the number of data streams achieved by all small cell user equipments (SUEs) while guaranteeing the data stream requirement for each SUE. As the corresponding optimization problem is NP-hard, we propose a two-phases algorithm based on graph theory to obtain a suboptimal but efficient solution, which requires much lower complexity and notably reduced feedback overhead. In the first phase, all SBS-SUE pairs are grouped into disjoint clusters by the recursive partition of the constructed graph, where the edge weight is only determined by the path losses from small cell base stations (SBSs) to SUEs, so the estimation of perfect global channel state information (CSI) is avoided. Furthermore, when guaranteeing that IA is feasible within each cluster, the number of formed clusters and the size of each cluster are determined by the distributions of dense SCNs instead of being specified. In the second phase, by further treating each cluster as a new vertex, graph coloring algorithm is proposed for subchannel allocation, which only requires to estimate the perfect CSI within each cluster. The analysis of computational complexity demonstrates that the proposed two-phases algorithm has a much lower complexity compared with the optimal solution obtained by the exhaustive search. Numerical results show that the proposed solution outperforms other related schemes and exhibits a performance close to the optimal solution.

INDEX TERMS Clustering, data streams, dense small cell networks (SCNs), graph coloring, graph partitioning, interference alignment (IA), subchannel allocation.

I. INTRODUCTION

According to the latest report from Cisco, the global smart devices such as smart phones and tablets grew to 4.5 billion, while this number will grow to 7 billion in the fifth generation (5G) wireless communication networks by 2022 [1]. Such explosive growth will create a massive number of data stream requirements (e.g., online video and game) for simultaneously accessing the network, which will become demanding especially in public areas such as airports, shopping malls and offices [2]. Network densification, which can provide enhanced coverage and improved spectral efficiency for macrocell networks by densely deploying small cell base stations (SBSs) in the existing macrocells, is one of the key technologies to meet the aforementioned requirements in 5G networks [3]–[5]. Additionally, the orthogonal deployment of dense small cell networks (SCNs) in existing macrocell networks, i.e., small cells and macrocells are allocated with orthogonal spectrum resources, is attractive to operators because they will not necessitate managing the cross-tier interference, which results in low interaction with macrocell networks [6].

Nevertheless, the severe co-tier interference between SBSs and small cell user equipments (SUEs) in orthogonal

The associate editor coordinating the review of this manuscript and approving it for publication was Qiang Ni.

frequency division multiple access (OFDMA) based dense SCNs will become a dominant factor that degrades the network performance [7], [8]. So how to effectively management the co-tier interference to further support the massive number of data stream requirements for SUEs in dense SCNs is one of the most challenging issues in 5G networks [4], [6], [9]. Additionally, due to the limited spectrum resources and explosive growth in the number of users [10], it is almost impossible to mitigate the co-tier interference only through the traditional interference management techniques such as subchannel allocation, power control or a combination of them in dense SCNs [9]. Consequently, it will be of great necessity to utilize much more advanced interference management techniques for co-tier interference mitigation in dense SCNs.

Recently, Interference alignment (IA) proposed in [11] is considered as an advanced technique of interference management. IA enables multiple transmitter-receiver (Tx-Rx) pairs to simultaneously transmit data over the same frequencies free from interference, which remarkably improves the spectral efficiency of dense SCNs where the limited spectrum resources are far from enough to meet the massive number of data stream requirements for SUEs. IA and its applications were comprehensively surveyed in [12].

However, exploiting IA in dense SCNs is limited by its feasibility condition given in [13]; besides, it requires to estimate the perfect global channel state information (CSI), which will incur heavy feedback overhead [12]. By dividing all users in the network into disjoint clusters with each only containing a certain number of users that meets the feasibility condition of IA, IA with clustering [14], [15] only requires to estimate the perfect CSI within each cluster, so the feedback overhead is also notably reduced. Through IA with clustering, the intra-cluster interference is completely eliminated. However, dense SCNs will also experience significant performance degradation because some of the inter-cluster interference will still be very strong. Accordingly, to meet the massive number of data stream requirements for SUEs, resource allocation based on IA with clustering is expected to be exploited for mitigating both intra-cluster and inter-cluster interference in dense SCNs.

In this paper, by allocating the limited subchannel resources based on IA with clustering to mitigate the co-tier interference in dense SCNs, we aim at maximizing the total number of data streams achieved by all SUEs while guaranteeing each SUE's data stream requirement. The corresponding optimization problem is formulated as a combinatorial optimization problem which is NP-hard. Obtaining its optimal solution by exhaustive search requires huge computational complexity, so it is unable to be implemented by a central entity called Home eNB Gateway (HeNB GW) [16] in practical system; furthermore, it necessitates estimating the perfect global CSI that requires heavy feedback overhead. Therefore, to find a suboptimal but efficient solution, we propose to solve it through a two-phases centralized algorithm which requires much lower complexity and notably reduced feedback overhead.

In the first phase, all SBS-SUE pairs are grouped into disjoint clusters through recursively partitioning the constructed graph, which only needs the information of path losses from SBSs to SUEs; then, to ensure the feasibility of IA within each cluster, we further remove some vertices from the clusters with sizes not meeting the feasibility condition of IA. Note that the number of clusters and the size of each cluster formed in the first phase are determined by the distributions of dense SCNs instead of being specified before clustering. In the second phase, by further treating each cluster as a new vertex, a proposed graph coloring algorithm is used for solving the subproblem of subchannel allocation based on IA with clustering, which requires a lower computational complexity. Besides, this phase only needs to estimate the perfect CSI with each cluster, so the feedback overhead is also notably reduced. In practical systems, HeNB GW will execute each of the aforementioned phases.

The main contributions of our paper are summarized as follows:

- We propose to efficiently utilize the limited subchannel resources based on IA with clustering to maximize the number of achievable data streams while guaranteeing each SUE's data stream requirement in dense SCNs.
- We propose a two-phases centralized algorithm based on graph theory, which requires much lower computational complexity and notably reduced feedback overhead, to obtain an efficient suboptimal solution to the corresponding NP-hard optimization problem.
- In the first phase, we group all vertices into disjoint clusters by recursively partitioning the constructed graph, where the proposed edge weight is only determined by the path losses from SBSs to SUEs; then, we remove some vertices from the clusters where IA is infeasible to ensure that IA can be performed within each cluster.
- In the second phase, by further treating each cluster as a new vertex, we propose a graph coloring algorithm for subchannel allocation, which only requires to estimate the perfect CSI within each cluster.
- We analyze the computational complexities of the optimal solution obtained by exhaustive search, the proposed solution and three other related solutions, and we also analyze the feedback overhead required by estimating the perfect global CSI and the proposed solution.

The remainder of this paper is structured as follows. Section II reviews the related work. Section III presents the system model and then formulates the corresponding optimization problem. In Section IV, we propose a two-phases suboptimal but efficient algorithm with low complexity and reduced feedback overhead. Section V analyzes the computational complexities of the optimal solution, proposed solution and three other related solutions. The feedback overhead required by estimating the perfect global CSI and the proposed solution are analyzed in Section VI. Simulation results are exhibited in Section VII. Finally, conclusions are drawn in Section VIII. A summary of key notations is given in Table 1.

TABLE 1. Summary of key notations.

Notation	Physical meaning
$\mathcal{K}$, K	The set and the number of SBSs in dense SCNs
$\mathcal{N}$, N	The set and the number of subchannels available in dense SCNs
M_T, M_R	The number of transmit and receive antennas at each SBS and SUE
$\mathcal{C}_i$, $\lvert\mathcal{C}_i\rvert$	An arbitrary cluster performing IA and its size
d	The number of data stream requirements for each SUE
M	The maximum size of each cluster performing IA
d_k	The number of data streams achieved by SUE k in cluster $\mathcal{C}_i$ after performing IA over each subchannel
$\mathsf{d}_{\mathcal{C}_i}$	The total number of data streams achieved by SUEs in cluster $\mathcal{C}_i$ after performing IA over each subchannel
$\mathbf{V}_k^n$	The transmit precoding matrix of SBS k for SUE k that is exclusively served by SBS k over subchannel n
$\mathbb{C}^{M_T\times \mathsf{d}_k}$	The set of complex $M_T \times \mathsf{d}_k$ matrices
$\mathbf{U}_k^n$, $(\mathbf{U}_k^n)^H$	The interference suppression matrix at SUE k over subchannel n and its Hermitian transpose
$\mathbf{H}_{kl}^n$	The channel gain matrix from SBS l to SUE k over subchannel n
$\mathbf{0}_{\mathsf{d}_k\times\mathsf{d}_k}$, $\mathbf{I}_{\mathsf{d}_k}$	The $\mathsf{d}_k \times \mathsf{d}_k$ matrix with all elements being 0 and $\mathsf{d}_k \times \mathsf{d}_k$ identity matrix
$\mathrm{rank}(\mathbf{A})$	The rank of matrix $\mathbf{A}$
PL_{km}	The path loss from SBS m to SUE k in dense SCNs
$\mathbf{s}_k^n$	The vector of transmitted symbols from SBS k to SUE k over subchannel n
$\mathcal{CN}(\mathbf{a},\mathbf{A})$	The complex Gaussian distributions with mean $\mathbf{a}$ and covariance matrix $\mathbf{A}$
p	The transmit power of each SBS over each subchannel that always keeps constant
$\mathbf{n}_k^n$	The circularly symmetric additive white Gaussian noise at SUE k over subchannel n
$(\sigma_k^n)^2$	The noise power at SUE k over subchannel n
$\mathsf{PL}_{\mathcal{C}_i\mathcal{C}_j}$	The path loss from cluster $\mathcal{C}_j$ to cluster $\mathcal{C}_i$ in dense SCNs
PL_0	A prescribed threshold of path loss used to distinguish the strong and weak interference
$\varphi_{\mathcal{C}_i}^n$	A decision variable that takes the value of 1 if subchannel n is allocated to cluster $\mathcal{C}_i$ and 0 otherwise
$\bigcap$, $\bigcup$, $\setminus$	The union, intersection and subtraction operations of sets
$\Theta_{\mathrm{clus}}^{\min}$	The minimum number of possible cases of clustering for SBS-SUE pairs that need to be compared
$\Phi_{\min}$	The minimum number of finally formed clusters performing IA
$\left\lceil \frac{K}{\mathsf{M}} \right\rceil$	The smallest positive integer bigger than or equal to $\frac{K}{\mathsf{M}}$
$\Theta_{\mathrm{allo}}^{\min}$	The total number of possible cases of subchannel allocation that need to be compared after forming $\Phi_{\min}$ clusters
$\mathcal{N}_{\mathcal{C}_i}$	The set of subchannels allocated to cluster $\mathcal{C}_i$
Ψ	The complexity of designing the precoding and interference suppression matrices for IA in each cluster over each subchannel
$\mathcal{G}$	The constructed graph
$\mathcal{V}$, $\mathcal{E}$	The set of vertices and edges in the constructed graph $\mathcal{G}$
$\mathbf{W}$	The weighted adjacency matrix of graph $\mathcal{G}$
w_{km}	The weight of edge connecting vertices k and m, i.e., the element in the kth row and the mth column of $\mathbf{W}$
PL'_{km}, $\mathsf{PL}'_{\mathcal{C}_i\mathcal{C}_j}$	The path loss between vertices k and m and that between clusters $\mathcal{C}_i$ and $\mathcal{C}_j$ in graph $\mathcal{G}$
$\mathsf{PL}'_{\mathrm{Min}}$	The minimum value of path losses between any two vertices in graph $\mathcal{G}$
C_0	A sufficiently big positive constant in the definition of edge weight
$\mathsf{W}_{\mathcal{C}_i}$	The weigh of cluster $\mathcal{C}_i$
$\mathrm{cut}(\mathcal{C}_i, \mathcal{V}\setminus\mathcal{C}_i)$	The cut between two disjoint vertex sets $\mathcal{C}_i$ and $\mathcal{V}\setminus\mathcal{C}_i$
$\mathcal{G}_1$, $\mathcal{G}_2$	The subgraphs obtained after partitioning graph $\mathcal{G}$ for the first time
$\mathcal{V}_{\mathrm{IA}}$, $\mathcal{V}_{\mathrm{lar}}$	The set of clusters performing IA and clusters with sizes larger than M
Φ	The number of finally formed clusters performing IA
$\mathcal{G}_s$	The simplified graph obtained by further treating each cluster performing IA as a vertex
$\mathcal{V}_{\mathcal{G}_s}$, $\mathcal{E}_{\mathcal{G}_s}$	The set of vertices and edges in the simplified graph $\mathcal{G}_s$
$\widetilde{v}_i$	The ith vertex in vertex set $\mathcal{V}_{\mathcal{G}_s}$, where $i = 1,\ldots,\Phi$
$\mathbf{A}_{\mathcal{G}_s}$	The adjacency matrix of simplified graph $\mathcal{G}_s$
$a_{\widetilde{v}_i\widetilde{v}_j}$	The element in $\widetilde{v}_i$th row and $\widetilde{v}_j$ column of adjacency matrix $\mathbf{A}_{\mathcal{G}_s}$
$\deg(\widetilde{v}_i)$	The degree of vertex $\widetilde{v}_i$ in the simplified graph $\mathcal{G}_s$
$S_{\widetilde{v}_i}$	The serial number of vertex $\widetilde{v}_i$ in $\mathcal{V}_{\mathcal{G}_s}$
Γ	The number of finally formed groups
Ω_{i_1}	The i_1th group that contains vertices without any edges among them, where $i_1 \in \{1,\ldots,\Gamma\}$
α, r_0	The quotient and the remainder of N divided by Γ

II. RELATED WORK

In this section, we respectively review IA with clustering and resource allocation based on IA in existing works, and then introduce the recent research of resource allocation based on IA with clustering.

A. IA WITH CLUSTERING AND RESOURCE ALLOCATION BASED ON IA

IA with clustering was firstly exploited in large cellular networks [14] and then in wireless ad hoc networks [15]. In [17], all Tx-Rx pairs were divided into disjoint clusters according

to the proposed criterion only determined by path losses. Then, IA was performed within each cluster to completely eliminate the strong intra-cluster interference, treating the relatively weak inter-cluster interference as noise. In [18], according to the distances between different Tx-Rx pairs, the whole network was divided into only one IA subnetwork (i.e., a cluster) and several spatial multiplexing subnetworks with each consisting of only one user and far away from the IA subnetwork. In [19], through poisson cluster process in a generic two-tier heterogeneous wireless network, all Tx-Rx pairs were grouped into disjoint clusters with the same size according to the distance between transmitters and receivers, and IA was performed within each cluster to maximize the second-tier throughput by using the tradeoff between signal-to-interference ratio and multiplexing gain. In dense SCNs, however, some of the inter-cluster interference is still strong and cannot be treated as noise. Consequently, without further resource allocation, it is unable to meet the massive number of data stream requirements for all SUEs only through IA with clustering.

The following works investigate resource allocation based on IA. In [20], an efficient approach was proposed to judge the feasibility condition of IA in full-duplex-enabled SCNs by relaxing the corresponding optimization problem into a linear programming one, and then an adaptive power allocation scheme based IA was proposed for interference elimination. In [21], by aligning the interference generated in different time slots to the same subspace, a scheme of time slot allocation based on partial IA was proposed to simultaneously mitigate both the inter-cell and intra-cell interference in a two-cell relay heterogeneous networks. However, in dense SCNs, without clustering, it is infeasible to mitigate all the interference to further meet the massive number of data stream requirements for all SUEs only by exploiting resource allocation based on IA.

B. RESOURCE ALLOCATION BASED ON IA WITH CLUSTERING

There are also several works that investigate resource allocation based on IA with clustering for interference mitigation. In [22], subchannel and power allocation based on IA with clustering were performed to maximum the sum rate of all secondary users (SUs) after performing IA while guaranteeing the quality of service (QoS) requirement for each secondary, and the total interference caused by SUs performing IA at each PU was ensured to be below the given threshold. In [23], to maximize the number of users with satisfactory QoS requirements in femtocell networks, a subchannel allocation scheme based on IA with clustering was proposed, and a suboptimal solution to the corresponding optimization problem was obtained through the transformed conflict graph. In [24], aiming at maximizing the sum rate of all femtocell users, a scheme of orthogonal subchannel allocation based on IA with clustering was proposed for fully connected ultra-dense femtocell networks, then an efficient algorithm with low complexity consisting of two phases was proposed for clustering and subchannel allocation. In [25], an IA and soft-space-reuse (IA-SSR) based cooperative transmission scheme was proposed for multi-cell massive MIMO networks, and the optimal power allocation and low overhead channel training framework based on IA-SSR were developed to maximize the sum capacity of the network. However, how to form disjoint clusters with specific clustering criterion is not investigated.

It is noteworthy that all clusters formed in [22]–[25] have equal sizes, i.e., clusters formed in [22]–[24] have a size of 3, and the sizes of those formed in [25] are assumed to be the same as the number of base stations, which is not the general case of performing IA. In dense SCNs, on the premise of guaranteeing IA is feasible and the intra-cluster interference is strong enough within each cluster, the cluster size should be determined by the distributions of dense SCNs.

In [26], base stations were grouped into disjoint clusters by the coalitional game theory according to large-scale fading. When the inter-cluster interference was strong, clusters were allocated with different time slots; otherwise, they would share the same time slots. In [27], a distributed multi-domain interference management scheme jointly utilized OFDMA scheduling, TDMA scheduling, IA and power control was proposed to mitigate interference while maximizing the achievable rate of SUEs in ultra-dense SCNs; furthermore, clusters were formed through an overlapping coalition formation game, and the inter-cluster interference was mitigated by TDMA scheduling. Nevertheless, the schemes above are not suitable for supporting the data stream requirements for the massive number of SUEs in dense SCNs because they need to simultaneously access the network.

In [28], an efficient subchannel allocation scheme based on IA with similarity clustering, i.e., both QoS requirements and path losses are similar within each cluster, was proposed to maximize the number of QoS guaranteed SUEs in dense SCNs, and a low-complexity algorithm with notably reduced feedback overhead was proposed to solve the optimization problem. However, instead of being determined by the distributions of dense SCNs, the number of clusters is a priori. In [29], IA with clustering was exploited in full-duplex-based SCNs. Furthermore, two clustering methods called minimized spectrum consumption clustering (MSCC) and minimized interference leakage clustering (MILC) were proposed, where the inter-cluster interference was mitigated through allocating orthogonal resource blocks in MSCC and treated as noise in MILC. However, in dense SCNs, all small cells are randomly distributed, so some of the inter-cluster interference is strong enough while some is relatively weak after clustering, and only utilizing MILC or MSCC will be not enough for inter-cluster interference mitigation.

To conclude, schemes of resource allocation based on IA with clustering which require low complexity and notably reduced feedback overhead should be investigated to effectively mitigate the interference and hence further meet the massive number of data stream requirements for SUEs; additionally, under the premise that IA is feasible and the

intra-cluster interference is strong enough within each cluster, the number of clusters and the size of each cluster should be determined by the distributions of dense SCNs instead of being specified before clustering.

III. SYSTEM MODEL AND PROBLEM FORMULATION

A. SYSTEM MODEL

As shown in Fig. 1, we consider the downlink transmission of OFDMA-based dense SCNs deployed in an existing macrocell, where small cells and the macrocell operate on orthogonal subchannels [2], [30]. Therefore, the cross-tier interference, including both the interference incurred by the macrocell base station (MBS) at SUEs and that incurred by SBSs at macrocell user equipments (MUEs), is mitigated. Assume that the subchannel allocation for MUEs has already been finished, so we only focus on mitigating the co-tier interference incurred by SBSs at SUEs through allocating the limited subchannels based on IA with clustering to maximize the number of data streams achieved by all SUEs in dense SCNs.

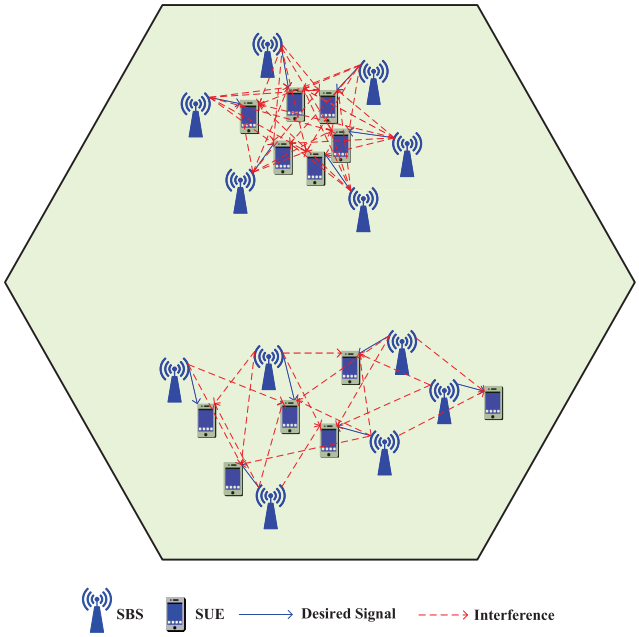

FIGURE 1. An illustrative example of dense SCNs composed of 12 small cells.

The set of SBSs in dense SCNs is denoted as $\mathcal{K} = \{1, 2, \ldots, K\}$. It is assumed that all SUEs have already been associated with their serving SBSs before clustering and subchannel allocation [23], [28], [31], and combining SUE association with resource allocation based on IA with clustering in dense SCNs will be our future work. Therefore, in dense SCNs, each SBS serves multiple SUEs associated with it, which is quite different from device-to-device (D2D) networks. Furthermore, in this paper, each SBS provides data stream transmission for its multiple associated SUEs through different time slots, so we assume that each SBS only serves one SUE in a given time slot of subchannel allocation [23], [28]. How to select an exclusive SUE for each SBS in a given time slot is beyond the scope of this paper, accordingly, it is also assumed to be finished before clustering and subchannel allocation. In this paper, SUE k is referred to as the exclusive SUE served by SBS k in a given time slot. Furthermore, an SBS and the SUE it exclusively serves in a given time slot are called an SBS-SUE pair, so clustering in this paper is referred to as grouping all SBS-SUE pairs into disjoint clusters, each of which contains a certain number of SBS-SUE pairs performing IA. In a new time slot of subchannel allocation, each SBS exclusively serves another one SUE associated with it, and the process of clustering and subchannel allocation for clusters will be performed again.

Let $\mathcal{N} = \{1, 2, \ldots, N\}$ denote the set of subchannels available in each time slot of subchannel allocation in dense SCNs, where N is much smaller than K. Each SBS and SUE is equipped with M_T and M_R antennas, respectively. Besides, each SUE wants to achieve at least d simultaneously transmitted interference-free data streams from its serving SBS, where $\mathsf{d} \leq \min\{M_T, M_R\}$ [13]. Furthermore, we only consider the case of all SUEs having the same and less data stream requirements, e.g., $\mathsf{d} = 1$, and the case of SUEs with differentiated data stream requirements will be investigated in our future work.

In an arbitrary cluster $\mathcal{C}_i$ that performs IA, the following inequality must be met [13]:

$$\mathsf{d} \leq \frac{M_T + M_R}{|\mathcal{C}_i| + 1}, \quad \forall \mathcal{C}_i \subseteq \mathcal{K}, \tag{1}$$

where $|\mathcal{C}_i|$ denotes the size of cluster $\mathcal{C}_i$, i.e., the number of SBS-SUE pairs contained by cluster $\mathcal{C}_i$. From inequality (1) we know that $|\mathcal{C}_i|$ is limited by

$$|\mathcal{C}_i| \leq \left\lfloor \frac{M_T + M_R}{\mathsf{d}} - 1 \right\rfloor \triangleq \mathsf{M}, \quad \forall \mathcal{C}_i \subseteq \mathcal{K}, \tag{2}$$

where $\left\lfloor \frac{M_T+M_R}{\mathsf{d}} - 1 \right\rfloor$ is the floor function that returns to the biggest integer smaller than or equal to $\frac{M_T+M_R}{\mathsf{d}} - 1$. In this paper, for simplicity, we use M to denote $\left\lfloor \frac{M_T+M_R}{\mathsf{d}} - 1 \right\rfloor$, i.e., the maximum size of each cluster performing IA.

After cluster $\mathcal{C}_i$ has formed, from (1), we know that the number of data streams achieved by SUE k in cluster $\mathcal{C}_i$ after performing IA over each subchannel is

$$\mathsf{d}_k = \left\lfloor \frac{M_T + M_R}{|\mathcal{C}_i| + 1} \right\rfloor \geq \mathsf{d}, \quad \forall k \in \mathcal{C}_i, \ \forall \mathcal{C}_i \subseteq \mathcal{K}, \tag{3}$$

which implies that each SUE in a cluster performing IA will achieve guaranteed data stream requirement as long as this cluster is allocated with one subchannel. Then, the total number of data streams achieved by all SUEs in cluster $\mathcal{C}_i$ after performing IA over each subchannel is

$$\mathsf{d}_{\mathcal{C}_i} = |\mathcal{C}_i| \cdot \mathsf{d}_k. \tag{4}$$

Besides, to perform IA within cluster $\mathcal{C}_i$ over subchannel n, the transmit precoding matrix of SBS k for SUE k that is exclusively served by SBS k and the interference suppression

matrix at SUE k over subchannel n, i.e., $\mathbf{V}_k^n \in \mathbb{C}^{M_T \times \mathsf{d}_k}$ and $\mathbf{U}_k^n \in \mathbb{C}^{M_R \times \mathsf{d}_k}$, must be designed to meet the following equations [32]:

$$\begin{cases} \left(\mathbf{U}_k^n\right)^H \mathbf{H}_{kl}^n \mathbf{V}_l^n = \mathbf{0}_{\mathsf{d}_k \times \mathsf{d}_k}, \\ \text{rank}\left(\left(\mathbf{U}_k^n\right)^H \mathbf{H}_{kk}^n \mathbf{V}_k^n\right) = \mathsf{d}_k, \quad \forall k,\ l \in \mathcal{C}_i,\ k \neq l, \end{cases} \tag{5}$$

where $\mathbf{H}_{kl}^n \in \mathbb{C}^{M_R \times M_T}$ denotes the channel gain matrix from SBS l to SUE k over subchannel n.

In our model, before performing IA, an SUE k in cluster $\mathcal{C}_i$ only suffers intra-cluster interference from other SBSs in cluster $\mathcal{C}_i$. This is because the strong inter-cluster interference will be mitigated by orthogonal subchannel allocation, while the weak interference will be treated as noise. Then, after completely eliminating the intra-cluster interference by IA in cluster $\mathcal{C}_i$, the received signal at SUE k in $\mathcal{C}_i$ over subchannel n is expressed as

$$\left(\mathbf{U}_k^n\right)^H \mathbf{y}_k^n = \sqrt{\frac{1}{\mathsf{PL}_{kk}}} \left(\mathbf{U}_k^n\right)^H \mathbf{H}_{kk}^n \mathbf{V}_k^n \mathbf{s}_k^n + \left(\mathbf{U}_k^n\right)^H \mathbf{n}_k^n, \tag{6}$$

where PL_{kk} denotes the path loss from SBS k to SUE k it exclusively serves in the given time slot. $\mathbf{s}_k^n \in \mathbb{C}^{\mathsf{d}_k \times 1}$ is the vector of transmitted symbols from SBS k to SUE k over subchannel n, where $\mathbf{s}_k^n \sim \mathcal{CN}\left(\mathbf{0}_{M_T \times 1}, \frac{p_k^n}{\mathsf{d}_k}\mathbf{I}_{\mathsf{d}_k}\right)$, and p_k^n denotes the transmit power of SBS k over subchannel n. In this paper, we only investigate subchannel allocation, hence we assume that $p_k^n = p$, $\forall k \in \mathcal{K}$, $\forall n \in \mathcal{N}$, where p keeps constant when performing IA with clustering and subchannel allocation. Such assumption is also used in [23] and [28]. Finally, $\mathbf{n}_k^n \in \mathbb{C}^{M_R \times 1}$ represents the circularly symmetric additive white Gaussian noise (AWGN) at SUE k over subchannel n, and $\mathbf{n}_k^n \sim \mathcal{CN}\left(\mathbf{0}_{M_R \times 1}, \left(\sigma_k^n\right)^2 \mathbf{I}_{M_R}\right)$, where $\left(\sigma_k^n\right)^2$ denotes the noise power at SUE k over subchannel n.

B. PROBLEM FORMULATION

To formulate the corresponding optimization problem, we first give the following definition.

Definition 1: In dense SCNs, the path loss from cluster $\mathcal{C}_j$ to cluster $\mathcal{C}_i$, i.e., $\mathsf{PL}_{\mathcal{C}_i\mathcal{C}_j}$, is defined as the minimum value of path losses from SBSs in $\mathcal{C}_j$ to SUEs in $\mathcal{C}_i$:

$$\mathsf{PL}_{\mathcal{C}_i\mathcal{C}_j} = \min\left\{\mathsf{PL}_{kf} \,|\, \forall f \in \mathcal{C}_j, \forall k \in \mathcal{C}_i\right\}, \tag{7}$$

where PL_{kf} is the path loss from SBS f in cluster $\mathcal{C}_j$ to SUE k in cluster $\mathcal{C}_i$.

Therefore, a threshold of path loss used to distinguish the strong and weak interference, i.e., PL_0, can be reasonably prescribed according to the transmit power of each SBS over each subchannel and the noise power at each SUE, so that when $\mathsf{PL}_{\mathcal{C}_i\mathcal{C}_j} > \mathsf{PL}_0$, the inter-cluster interference between $\mathcal{C}_i$ and $\mathcal{C}_j$ will be weak enough and can be treated as noise, hence they are allowed to use the same subchannels; otherwise, they must be allocated with orthogonal subchannels to mitigate the strong inter-cluster interference.

Aiming at maximizing the total number of data streams achieved by all SUEs while guaranteeing the data stream requirement for each SUE in dense SCNs, the corresponding optimization problem is formulated as follow:

$$\begin{aligned} \max_{\mathcal{C}_i,\, \varphi_{\mathcal{C}_i}^n} \ & \sum_{n \in \mathcal{N}} \sum_{\mathcal{C}_i \subseteq \mathcal{K}} \varphi_{\mathcal{C}_i}^n \cdot \mathsf{d}_{\mathcal{C}_i} \\ \text{s.t. C1}: \ & \mathcal{C}_i \bigcap \mathcal{C}_j = \emptyset, \quad \forall \mathcal{C}_i, \mathcal{C}_j \subseteq \mathcal{K} \\ \text{C2}: \ & \bigcup \mathcal{C}_i = \mathcal{K} \\ \text{C3}: \ & 2 \leq |\mathcal{C}_i| \leq \mathsf{M}, \quad \forall \mathcal{C}_i \\ \text{C4}: \ & \mathsf{PL}_{kl} \leq \mathsf{PL}_0, \quad \forall k, l \in \mathcal{C}_i,\ \forall \mathcal{C}_i \\ \text{C5}: \ & \varphi_{\mathcal{C}_i}^n \cdot \varphi_{\mathcal{C}_j}^n = \begin{cases} 1, \text{ if } \mathsf{PL}_{\mathcal{C}_i\mathcal{C}_j} > \mathsf{PL}_0 \\ 0, \text{ else}, \end{cases} \quad \forall \mathcal{C}_i,\ \mathcal{C}_j,\ \forall n \\ \text{C6}: \ & \sum_{n \in \mathcal{N}} \varphi_{\mathcal{C}_i}^n \geq 1, \quad \forall \mathcal{C}_i \\ \text{C7}: \ & \varphi_{\mathcal{C}_i}^n \in \{0, 1\}, \quad \forall \mathcal{C}_i,\ \forall n. \end{aligned} \tag{8}$$

In problem (8), $\varphi_{\mathcal{C}_i}^n$ is a decision variable that takes the value of 1 if subchannel n is allocated to cluster $\mathcal{C}_i$, and 0 otherwise. Constraints C1 and C2 ensure that all clusters are disjoint and their union constitutes the set of all SUEs $\mathcal{K}$, respectively. The size of each cluster is limited in constraint C3, where the case of two users performing IA is investigated in [33]. Constraint C4 requires the path loss from SBS l to SUE k in cluster $\mathcal{C}_i$ not to exceed the given threshold PL_0, so that the intra-cluster interference is strong enough. Constraint C5 indicates that if $\mathsf{PL}_{\mathcal{C}_i\mathcal{C}_j} > \mathsf{PL}_0$, the inter-cluster interference will be treated as noise, so that clusters $\mathcal{C}_i$ and $\mathcal{C}_j$ can reuse the same subchannels; otherwise, they must be allocated with orthogonal subchannels. Constraint C6 ensures that each cluster is allocated with at least one subchannel, so that each SUE can achieve at least d data streams. This is because IA enables all SUEs in cluster $\mathcal{C}_i$ to share subchannel n free from intra-cluster interference, allocating subchannel n to cluster $\mathcal{C}_i$ is equivalent to allocating it to all SUEs in cluster $\mathcal{C}_i$. Optimization problem (8) is a combinatorial optimization problem, and we have the following lemma.

Lemma 1: Optimization problem (8) is NP-hard.

Proof: Obtaining the optimal solution to problem (8) requires to compare all the possible cases of clustering for SBS-SUE pairs and those of subchannel allocation for each clustering result.

Firstly, the total number of all possible cases of clustering for SBS-SUE pairs that need to be compared is

$$\Theta_{\text{clus}} = \sum_{\substack{2 \leq |\mathcal{C}_i| \leq \mathsf{M} \\ i=1,\ldots,\Phi}} \frac{\binom{K}{|\mathcal{C}_1|}\left[\prod_{i_1=2}^{\Phi-1}\binom{K-\sum_{i_2=1}^{i_1-1}|\mathcal{C}_{i_2}|}{|\mathcal{C}_{i_1}|}\right]}{\Phi!} \tag{9}$$

$$\approx \sum_{\substack{2 \leq |\mathcal{C}_i| \leq \mathsf{M} \\ i=1,\ldots,\Phi}} K^K \tag{10}$$

$$= (\mathsf{M}-1)^{\Phi} K^K. \tag{11}$$

In (9), Φ is the number of clusters performing IA, where $\Phi \geq \Phi_{\min}$. Here $\Phi_{\min} = \lceil \frac{K}{\mathsf{M}} \rceil$ denotes the minimum number of clusters performing IA, and $\lceil \frac{K}{\mathsf{M}} \rceil$ is the smallest positive

integer bigger than or equal to $\frac{K}{\mathsf{M}}$. Moreover, $\binom{K}{|\mathcal{C}_1|} = \frac{K!}{|\mathcal{C}_1|!(K-|\mathcal{C}_1|!)}$ is the number of cases choosing $|\mathcal{C}_1|$ SBS-SUE pairs from all the K SBS-SUE pairs, where $K! = \prod_{k=1}^{K} k$ denotes the factorial of K. Besides, the approximation in (10) follows the result in [31].

Then, the total number of possible cases of subchannel allocation that need to be compared after forming Φ clusters is

$$\Theta_{\text{allo}} = \prod_{i=1}^{\Phi}\left[\sum_{1\le|\mathcal{N}_{\mathcal{C}_i}|\le N}\binom{N}{|\mathcal{N}_{\mathcal{C}_i}|}\left(\Psi\left|\mathcal{N}_{\mathcal{C}_i}\right|\right)\right] > \left(2^N\Psi\right)^{\Phi}, \tag{12}$$

where $\mathcal{N}_{\mathcal{C}_i}$ is the set of subchannels allocated to cluster $\mathcal{C}_i$, and Ψ denotes the complexity of designing the precoding and interference suppression matrices for IA in each cluster over each subchannel [22].

From (11) and (12) we know that obtaining the optimal solution to problem (8) requires to compare the number of $\Theta_{\text{clus}}\Theta_{\text{allo}}$ cases, which is bigger than $\left[(\mathsf{M}-1)2^N\Psi\right]^{\Phi}K^K$ and exponentially increases with both K and N. Therefore, optimization problem (8) is NP-hard. ■

Lemma 1 indicates that obtaining the optimal solution to problem (8) through comparing all the possible cases of clustering and subchannel allocation for clusters requires prohibitive computational complexity; besides, it necessitates estimating the perfect global CSI, which will result in huge feedback overhead. Consequently, we propose a two-phases algorithm with lower complexity and reduced feedback overhead based on graph theory to obtain a suboptimal but efficient solution to problem (8). The first phase groups all SBS-SUE pairs into disjoint clusters through graph partition based on path loss, where each SBS and the SUE it exclusively serves in a given time slot are represented as a vertex in the constructed graph, and then removes vertices from the clusters with sizes larger than M. The second phase allocates subchannels to the formed clusters by the proposed graph coloring algorithm. Such method (i.e., clustering and then resource allocation for clusters) is also adopted in [31], [23] and our previous work [28].

IV. PROPOSED SOLUTION

A. PHASE 1: CLUSTERING BASED ON PATH LOSSES

The goal of clustering is to group all SBS-SUE pairs into disjoint clusters, each of which has a size that meets the feasibility condition of IA (i.e., constraint C3 in problem (8)) so that IA can be performed within it to completely eliminate the strong intra-cluster interference. However, such combinatorial optimization problem is still NP-hard. As a suboptimal but efficient solution, we utilize graph partitioning to finish clustering for SBS-SUE pairs in this phase, which requires a much lower computational complexity.

Firstly, we need to construct the weighted graph $\mathcal{G} \triangleq (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is the set of all vertices in the constructed graph $\mathcal{G}$, and $\mathcal{E}$ denotes the set of all edges in $\mathcal{G}$. In this paper, each SBS-SUE pair, i.e., each vertex in $\mathcal{V}$ represents an SBS and the SUE it exclusively serves in a given time slot (see Fig. 2). The reasons are as follows. Firstly, the path loss from each SBS to the SUE it exclusively serves is smaller than the path losses between this SBS and the SUEs it does not serve. Secondly, in the first phase, we only consider the interference incurred by SBSs at SUEs they do not serve. So the property of an SBS and the SUE it exclusively serves is the same as that of a vertex. The aforementioned method is also used in [23] and [28]. As shown in Fig. 2, the vertex that represents SBS k and SUE k that is exclusively served by SBS k in dense SCNs is called vertex k, where $k = 1, \ldots, K$. Consequently, the vertex set $\mathcal{V}$ in graph $\mathcal{G}$ consists of K vertices because there are altogether K SUEs in dense SCNs. Note that we consider the downlink transmission of OFDMA-based dense SCNs in this paper, so each SBS plays a role of transmitter, and the SUE it exclusively serves plays a role of intended receiver. Therefore, treating each SBS-SUE pair as one vertex is equivalent to merging each transmitter and its intended receiver into one vertex, which implies that there is only one type of vertices in graph $\mathcal{G}$.

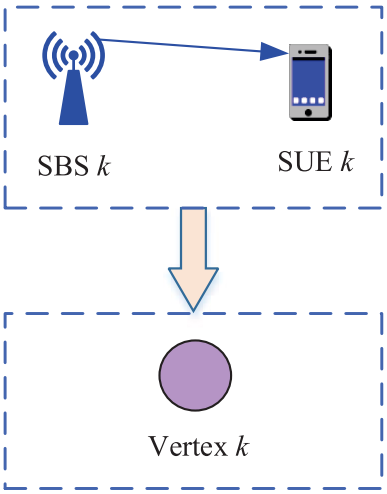

FIGURE 2. Treating each SBS-SUE pair as a vertex.

Moreover, we use $\mathbf{W} = [w_{km}]$ to denote the weighted adjacency matrix of graph $\mathcal{G}$, where w_{km} is the weight of edge connecting vertices k and m. As the intra-cluster interference must be strong, the weight of edges in each cluster should be as big as possible. Note that each SBS has the same transmit power, but the path losses from SBSs to the SUEs they do not serve are differentiated due to the random deployment of small cells. Consequently, the weight of edge connecting any two vertices should be determined by the path loss between them. Besides, $\mathcal{G}$ is an undirected graph, so the edge weight must be symmetric, i.e., $w_{km} = w_{mk}, \forall k, m \in \mathcal{V}$, and the path loss between vertices k and m must also be symmetric, i.e., $\mathsf{PL}'_{km} = \mathsf{PL}'_{mk}$, which can be guaranteed by defining PL'_{km} as the minimum value between the path loss from SBS

m to SUE k and that from SBS k to SUE m in dense SCNs:

$$\mathsf{PL}'_{km} = \min\left\{\mathsf{PL}_{km}, \mathsf{PL}_{mk} \middle| \forall k, m \in \mathcal{V}\right\}. \tag{13}$$

Therefore, if either PL_{km} or PL_{mk} does not exceed PL_0, we have $\mathsf{PL}'_{km} \leq \mathsf{PL}_0$, i.e., there will be strong interference between vertices k and m, and vice versa. Then, similar to Definition 1, we have the following definition.

Definition 2: In the constructed graph $\mathcal{G}$, the path loss between clusters $\mathcal{C}_i$ and $\mathcal{C}_j$, i.e., $\mathsf{PL}'_{\mathcal{C}_i\mathcal{C}_j}$, is defined as the minimum value of path losses between all vertices in $\mathcal{C}_i$ and those in $\mathcal{C}_j$:

$$\mathsf{PL}'_{\mathcal{C}_i\mathcal{C}_j} = \min\left\{\mathsf{PL}'_{fk} \middle| \forall f \in \mathcal{C}_j, \forall k \in \mathcal{C}_i\right\}. \tag{14}$$

Furthermore, we use the ratio of the transmit power of each SBS over each subchannel to the path loss between vertices k and m, i.e., p/PL'_{km}, to quantify the strength of interference induced by vertex m at vertex k. Accordingly, the bigger the ratio, the stronger interference vertex m will induce at vertex k, and vice versa. In this paper, w_{km} is defined as

$$w_{km} = \begin{cases} 1, & \text{if } k = m \\ \dfrac{p/\mathsf{PL}'_{km}}{p/\mathsf{PL}'_{\mathrm{Min}}}, & \text{else if } \mathsf{PL}'_{km} \leq \mathsf{PL}_0, \quad \forall k \neq m \\ -\mathrm{C}_0, & \text{else,} \end{cases} \tag{15}$$

where $\mathsf{PL}'_{\mathrm{Min}}$ is the minimum value of path losses between any two vertices in graph $\mathcal{G}$, i.e.,

$$\mathsf{PL}'_{\mathrm{Min}} = \min\left\{\mathsf{PL}'_{km} \middle| \forall k, m \in \mathcal{V}\right\}. \tag{16}$$

Here note that in order to make w_{km} dimensionless, we normalize the term p/PL'_{km} when $\mathsf{PL}'_{km} \leq \mathsf{PL}_0, \forall k \neq m$. Moreover, the constant C_0 in (15) is a sufficiently big positive number. In graph $\mathcal{G}$, when $\mathsf{PL}'_{km} \leq \mathsf{PL}_0$, the interference between vertices k and l is regarded as strong interference, and they are connected by a solid line; otherwise, the interference between them is weak, and they are connected by a dashed line. Fig. 3 shows the constructed graph $\mathcal{G}$ corresponding to Fig. 1.

Then, the subproblem of clustering for SBS-SUE pairs is converted into the problem of clustering for vertices which is formulated as

$$\begin{aligned} &\max_{\mathcal{C}_i} \sum_{\mathcal{C}_i \subseteq \mathcal{V}} \mathsf{W}_{\mathcal{C}_i} \\ &\text{s.t. C1} - \text{C3} \\ &\quad \text{C4}' : \mathsf{PL}'_{kl} \leq \mathsf{PL}_0, \quad \forall k, l \in \mathcal{C}_i, \ \forall \mathcal{C}_i \subseteq \mathcal{V}. \end{aligned} \tag{17}$$

In (17), $\mathsf{W}_{\mathcal{C}_i}$ denotes the weigh of cluster $\mathcal{C}_i$, which is defined as the sum weight of edges in $\mathcal{C}_i$, i.e.,

$$\mathsf{W}_{\mathcal{C}_i} = \sum_{k,l \in \mathcal{C}_i} w_{kl}. \tag{18}$$

Nevertheless, subproblem (17) is still NP-hard. This is because obtaining its optimal solution requires to compare all the possible cases of clustering for K vertices. From (11) it can be known that the minimum total number of possible cases that need to be compared, i.e., $(\mathsf{M}-1)^{\Phi_{\min}} K^K$, still exponentially increases with K, which demonstrates the NP-hardness of subproblem (17). To solve (17), we first convert it into the following equivalent subproblem through the following lemma.

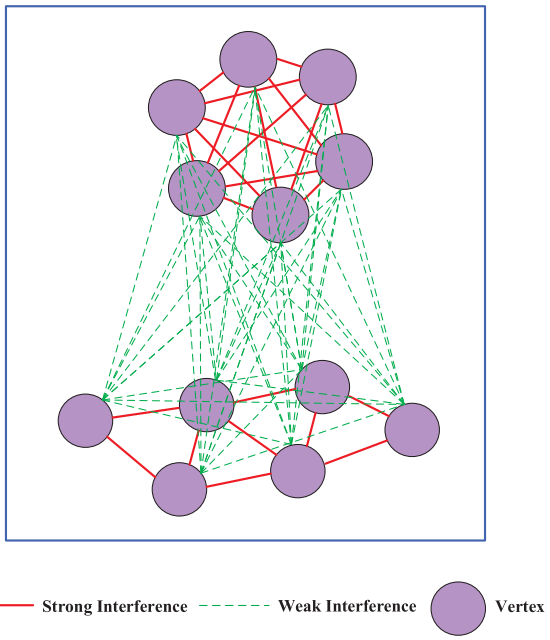

FIGURE 3. The constructed graph $\mathcal{G}$.

Lemma 2: Subproblem (17) is equivalent to the following subproblem:

$$\begin{aligned} &\min_{\mathcal{C}_i} \sum_{\mathcal{C}_i \subseteq \mathcal{V}} \mathrm{cut}\left(\mathcal{C}_i, \mathcal{V} \setminus \mathcal{C}_i\right) \\ &\text{s.t. C1} - \text{C3, C4}' \end{aligned} \tag{19}$$

where

$$\mathrm{cut}\left(\mathcal{C}_i, \mathcal{V} \setminus \mathcal{C}_i\right) = \sum_{k \in \mathcal{C}_i, q \in \mathcal{V} \setminus \mathcal{C}_i} w_{kq} \tag{20}$$

denotes the cut between two disjoint vertex sets $\mathcal{C}_i$ and $\mathcal{V} \setminus \mathcal{C}_i$.

Proof: See Appendix A.

Subproblem (19) is equivalent to the NP-hard subproblem (17), so it is also NP-hard. However, as an efficient solution, we can recursively partition graph $\mathcal{G}$ through the minimum cut criterion, and the well-known Stoer-Wagner (S-W) algorithm proposed in [34] can be used to compute the minimum cut of an arbitrary subgraph of $\mathcal{G}$. Firstly, graph $\mathcal{G}$ will be partitioned into two subgraphs $\mathcal{G}_1$ and $\mathcal{G}_2$ by S-W algorithm. Then, $\mathcal{G}_1$ and $\mathcal{G}_2$ will be recursively partitioned until the weighted adjacency matrix of each subgraph partitioned from them has no elements with value $-\mathrm{C}_0$. More detail is shown in Algorithm 1. Note that in step 1 of Algorithm 1, $\mathcal{V}_{\mathrm{IA}}$ denotes the set of clusters performing IA, i.e., the size of each cluster contained by $\mathcal{V}_{\mathrm{IA}}$ does not exceed M, and $\mathcal{V}_{\mathrm{lar}}$ represents the

Algorithm 1 Clustering for Vertices by the Recursive Partition of Graph $\mathcal{G}$

Initialize $\mathcal{V}_{\mathrm{IA}} = \emptyset$ and $\mathcal{V}_{\mathrm{lar}} = \emptyset$.
Compute the weighted adjacency matrix $\mathbf{W}$ according to (15), and then construct graph $\mathcal{G}$.
if there are no elements with value $-\mathrm{C}_0$ in $\mathbf{W}$ **then**
  Stop and treat $\mathcal{V}$ as a cluster.
else
  Exploit S-W algorithm to partition $\mathcal{G}$ into two subgraphs, i.e., $\mathcal{G}_1$ and $\mathcal{G}_2$.
end if
for each subgraph $\mathcal{G}_i$ $(i = 1, 2)$ partitioned from $\mathcal{G}$ **do**
  if there are no elements with value $-\mathrm{C}_0$ in the weighted adjacency matrix of $\mathcal{G}_i$ **then**
    Stop and treat the vertex set of $\mathcal{G}_i$ as a cluster.
  else
    Recursively execute steps 2-6 to partition subgraph $\mathcal{G}_i$ until there are no elements with value $-\mathrm{C}_0$ in the weighted adjacency matrix of each subgraph partitioned from $\mathcal{G}_i$, and treat the vertex set of each subgraph obtained after the recursive partition as one cluster.
  end if
end for
Delete all edges with weight $-\mathrm{C}_0$.
for each obtained cluster $\mathcal{C}_i$ **do**
  if $|\mathcal{C}_i| \leq \mathsf{M}$ **then**
    Update $\mathcal{V}_{\mathrm{IA}} = \mathcal{V}_{\mathrm{IA}} \bigcup \mathcal{C}_i$.
  else
    Update $\mathcal{V}_{\mathrm{lar}} = \mathcal{V}_{\mathrm{lar}} \bigcup \mathcal{C}_i$.
  end if
end for
Output: The set of clusters performing IA, i.e., $\mathcal{V}_{\mathrm{IA}}$, and the set of clusters with sizes larger than M, i.e., $\mathcal{V}_{\mathrm{lar}}$.

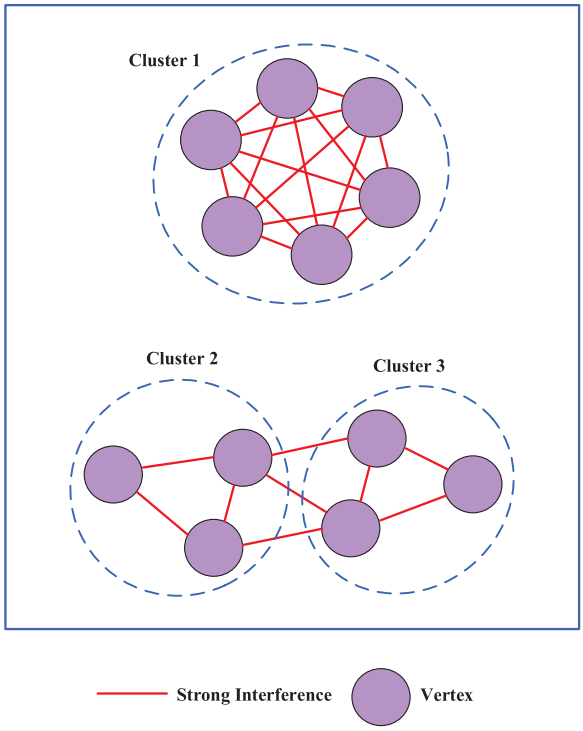

FIGURE 4. Clusters formed after Algorithm 1.

set of clusters with sizes larger than M. After Algorithm 1, we have the following theorem.

Theorem 1: The path loss between any two vertices in each cluster formed after Algorithm 1 is smaller than or equal to PL_0.

Proof: See Appendix B.

Note that Algorithm 1 finally outputs $\mathcal{V}_{\mathrm{IA}}$ and $\mathcal{V}_{\mathrm{lar}}$, where $\mathcal{V}_{\mathrm{IA}} \bigcup \mathcal{V}_{\mathrm{lar}} = \mathcal{V}$. When $\mathcal{V}_{\mathrm{lar}} \neq \emptyset$, there exist some clusters with sizes larger than M, where IA is still infeasible. So we further need to address the aforementioned feasibility issue of IA by removing some vertices from the clusters whose sizes exceed M. Specifically, if $|\mathcal{C}_i| > \mathsf{M}$, we first select one cluster (e.g., $\mathcal{C}_{i_1}$) which contains M vertices and has the maximum weight from $\mathcal{C}_i$. Then, if $|\mathcal{C}_i \setminus \mathcal{C}_{i_1}| > \mathsf{M}$, the process above will be repeatedly performed until each of the newly formed cluster contains at most M vertices; otherwise, $\mathcal{C}_i \setminus \mathcal{C}_{i_1}$ will be treated as another new cluster. For example, when $M_T = M_R = 2$, and $\mathsf{d} = 1$, we have $\mathsf{M} = 3$. Therefore, in Fig. 4, IA is infeasible in Cluster 1, and we first need to select 3 vertices that has the maximum weight from Cluster 1 and treat them as a new cluster, i.e., Cluster 1′; and then the remaining 3 vertices will form another new cluster, i.e., Cluster 4. Note that when there are no elements with value $-\mathrm{C}_0$ in $\mathbf{W}$, i.e., the interference between any two vertices in $\mathcal{V}$ are strong enough. In this special case, $|\mathcal{V}|$ (i.e., K) is much bigger than M, so IA is definitely infeasible in $\mathcal{V}$ if it is treated as a big cluster. Then, M (i.e., the maximum cluster size of performing IA) vertices will be removed from $\mathcal{V}$ to form a new cluster performing IA, e.g., cluster $\mathcal{C}_i$, and then $\mathcal{V}$ is updated to $\mathcal{V} \setminus \mathcal{C}_i$. The aforementioned procedure will be repeatedly performed until $|\mathcal{V}| \leq \mathsf{M}$. More detail about removing vertices is given in Algorithm 2.

It is also worth mentioning that Φ in steps 6 and 14 of Algorithm 2 denotes the number of finally formed clusters performing IA, and it is not specified before clustering but determined by the distributions of dense SCNs. Moreover, $\Phi_{\min} = \lceil \frac{K}{\mathsf{M}} \rceil$ in step 6 represents the minimum number of finally formed clusters performing IA. We know that only when there are no elements with value $-\mathrm{C}_0$ in $\mathbf{W}$, $\Phi_{\min}$ can be achieved. Finally, Algorithm 2 outputs the set of finally formed Φ clusters performing IA, i.e., $\mathcal{V}_{\mathrm{IA}} = \{\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_\Phi\}$.

Executing Algorithm 1 and Algorithm 2 require to know the weight of edge between any two vertices that is only determined by the path losses between them. In practical communication systems, the path losses can be estimated by each SUE through computing the ratios of transmit power of pilot symbols from different SBSs to the receive power

Algorithm 2 Removing Vertices From Clusters With Sizes Larger Than M

if there are no elements with value $-C_0$ in **W then**
  Set $i = 1$.
  while $|\mathcal{V}| > \mathsf{M}$ **do**
    Determine $\mathcal{C}_i$ which has the maximum weight and meets $|\mathcal{C}_i| = \mathsf{M}$ in $\mathcal{V}$, and update $\mathcal{V}_{\mathrm{IA}} = \mathcal{V}_{\mathrm{IA}} \bigcup \mathcal{C}_i$, $\mathcal{V} = \mathcal{V} \setminus \mathcal{C}_i$, and $i = i + 1$.
  end while
  Set $\Phi = \Phi_{\min}$ and $\mathcal{C}_\Phi = \mathcal{V}$, where $\Phi_{\min} = \lceil \frac{K}{\mathsf{M}} \rceil$, and update $\mathcal{V}_{\mathrm{IA}} = \mathcal{V}_{\mathrm{IA}} \bigcup \mathcal{C}_\Phi$.
else
  while $|\mathcal{V}_{\mathrm{lar}}| > 0$ **do**
    for each $\mathcal{C}_j \subseteq \mathcal{V}_{\mathrm{lar}}$ **do**
      Execute steps 2-6 to remove vertices from cluster $\mathcal{C}_j$, then update $\mathcal{V}_{\mathrm{lar}}$ and $\mathcal{V}_{\mathrm{IA}}$.
    end for
  end while
end if
Output: The set of finally formed Φ clusters performing IA, i.e., $\mathcal{V}_{\mathrm{IA}} = \{\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_\Phi\}$.

of corresponding pilot symbols received at this SUE. In this way, the estimation of perfect global CSI is avoided when grouping all SBS-SUE pairs into disjoint clusters, which notably reduces the feedback overhead in dense SCNs.

B. PHASE 2: SUBCHANNEL ALLOCATION BY THE PROPOSED GRAPH COLORING ALGORITHM

After the phase of clustering, the subproblem of subchannel allocation for clusters becomes

$$\max_{\varphi^n_{\mathcal{C}_i}} \sum_{n\in\mathcal{N}} \sum_{\mathcal{C}_i\subseteq\mathcal{V}_{\mathrm{IA}}} \varphi^n_{\mathcal{C}_i} \cdot \mathsf{d}_{\mathcal{C}_i}$$
$$\text{s.t. C5}' : \varphi^n_{\mathcal{C}_i} \cdot \varphi^n_{\mathcal{C}_j} = \begin{cases} 1, & \text{if } \mathsf{PL}'_{\mathcal{C}_i\mathcal{C}_j} > \mathsf{PL}_0 \\ 0, & \text{else,} \end{cases} \quad \forall \mathcal{C}_i,\ \mathcal{C}_j,\ \forall n$$
$$\text{C6, C7.} \tag{21}$$

Subproblem (21) is also NP-hard owing to the combinatorial nature of subchannel allocation. Because each cluster performing IA still contains multiple vertices without intra-cluster interference among them, so subproblem (21) is quite different from the traditional subchannel allocation problems. Fortunately, each cluster can be further treated as a vertex, which is also adopted in [23]. The reason is that there is no intra-cluster interference within each cluster performing IA, and each cluster only suffers the inter-cluster interference, which indicates that from the aspect of subchannel allocation, a cluster has the same property as that of a vertex. In this way, graph $\mathcal{G}$ can be further converted into the simplified graph $\mathcal{G}_s = (\mathcal{V}_{\mathcal{G}_s}, \mathcal{E}_{\mathcal{G}_s})$. Here $\mathcal{V}_{\mathcal{G}_s} = \{\widetilde{v}_1, \ldots, \widetilde{v}_\Phi\}$ is the set of vertices in $\mathcal{G}_s$. Note that after Algorithm 2, there are altogether Φ clusters performing IA, therefore, in the simplified graph $\mathcal{G}_s$, we use vertex $\widetilde{v}_i$ to represent cluster $\mathcal{C}_i$, where $i = 1, \ldots, \Phi$. Besides, $\mathcal{E}_{\mathcal{G}_s}$ denotes the set of edges in $\mathcal{G}_s$. To obtain $\mathcal{G}_s$,

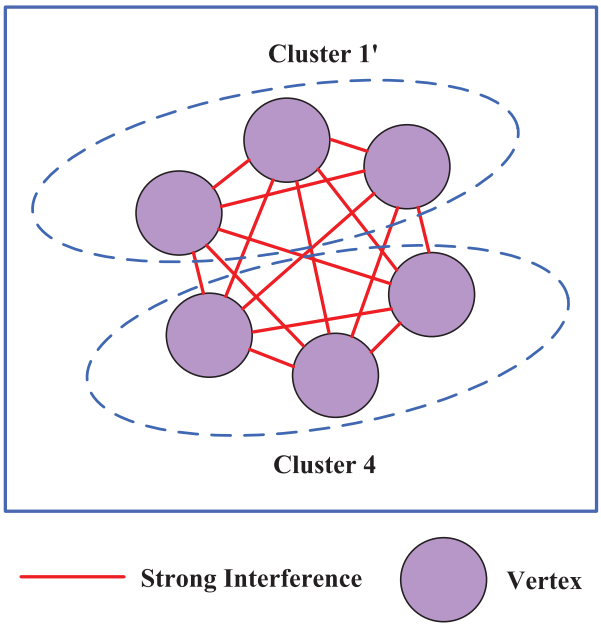

FIGURE 5. Removing vertices from Cluster 1.

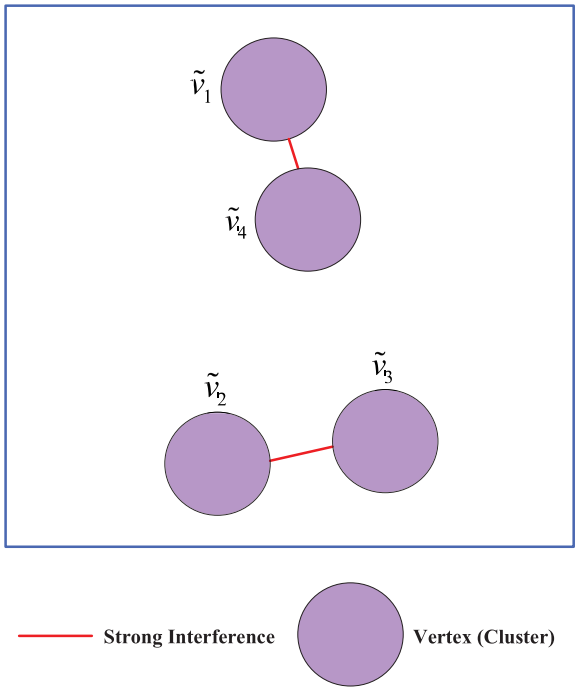

FIGURE 6. The simplified graph $\mathcal{G}_s$.

we first need to determine its adjacency matrix through the following definition.

Definition 3: The adjacency matrix of simplified graph $\mathcal{G}_s$ is defined as a $\Phi \times \Phi$ symmetric matrix $\mathbf{A}_{\mathcal{G}_s} = [a_{\widetilde{v}_i\widetilde{v}_j}]$, where $a_{\widetilde{v}_i\widetilde{v}_j} = 1$ if $\mathsf{PL}'_{\mathcal{C}_i\mathcal{C}_j} \le \mathsf{PL}_0$, $\forall \widetilde{v}_i \ne \widetilde{v}_j$; otherwise, $a_{\widetilde{v}_i\widetilde{v}_j} = 0$.

In $\mathcal{G}_s$, if $a_{\widetilde{v}_i\widetilde{v}_j} = 1$, there is an edge connecting vertices $\widetilde{v}_i$ and $\widetilde{v}_j$; otherwise, there is no edge between them. Fig. 6 shows the simplified graph $\mathcal{G}_s$. Then, based on Definition 3,

the degree of vertex $\widetilde{v}_i$ in $\mathcal{G}_s$ is defined as the number of edges connecting with it, i.e., the sum of elements in the row (column) corresponding to $\widetilde{v}_i$ in $\mathbf{A}_{\mathcal{G}_s}$:

$$\deg(\widetilde{v}_i) = \sum_{\widetilde{v}_j \in \mathcal{V}_{\mathcal{G}_s}} a_{\widetilde{v}_i \widetilde{v}_j}. \tag{22}$$

Next we utilize the proposed graph coloring algorithm which has a lower complexity to obtain a suboptimal but efficient solution to (21). Briefly speaking, each subchannel is represented by one color, and if two vertices in $\mathcal{V}_{\mathcal{G}_s}$ is connected with one edge, they must be allocated with different colors; otherwise, they can be allocated with the same colors. Moreover, vertices with bigger degrees will be given higher priority to subchannel allocation. The process of subchannel allocation for vertices by the proposed coloring algorithm is described as follows.

We first determine $\mathbf{A}_{\mathcal{G}_s}$ and compute the degree of each vertex in $\mathcal{V}_{\mathcal{G}_s}$; then, we sort all vertices by the descending order of their degrees, which follows step 1 of the Dsatur algorithm firstly proposed in [35] to efficiently color the vertices of a graph. However, the number of SUEs is much bigger than that of subchannels available in dense SCNs, and our goal is to maximize the total number of data streams achieved by all SUEs while guaranteeing the data stream requirement for each SUE, which are not taken into consideration in the Dsatur algorithm. So all the steps after step 3 are proposed by us to utilize the limited subchannel resources more efficiently.

Specifically speaking, firstly, after step 3, there may exist some vertices with the same degrees. Therefore, secondly, these vertices are further sorted by the descending order of the number of their achievable data streams, so that the vertices with bigger number of data streams are given higher priority to subchannel allocation, which is in accordance with maximizing the total number of data streams achieved by all SUEs. More detail is shown in steps 4-11 of Algorithm 3. Secondly, in order to utilize the N subchannels more efficiently, we divide all the sorted vertices into disjoint groups, each of which contains vertices without any edges among them, so that all vertices in each group are allocated with the same colors, and vertices in different groups will be allocated with different colors. Note that a group is quite different from a cluster performing IA. In this paper, a group in the simplified graph $\mathcal{G}_s$ is a set of vertices without interference among them, while a cluster performing IA in the constructed graph $\mathcal{G}$ is a set of vertices with strong interference. Detail about group formation is shown in steps 12-19. Finally, the disjoint groups formed through steps 12-19 will be allocated with different colors. Note that the number of groups is smaller than that of subchannels available, so each group can be allocated with at least one subchannel, i.e., all vertices contained by each group can also be allocated with at least one subchannel. Therefore, constraint C6 of subproblem (21) can be definitely met. Then, the remaining procedures of subchannel allocation are used to further increase the total number of achievable data streams. Steps 20-24 show the detail about subchannel allocation for the formed groups.

Next, we give the definitions of new notations used in Algorithm 3 and make some necessary explanations for the algorithm. $\widetilde{v}_1$, $\widetilde{v}_{\Phi-1}$, and $\widetilde{v}_{i+1}$ denote the first vertex, the $(\Phi-1)$th vertex, and the $(i+1)$th vertex in $\mathcal{V}_{\mathcal{G}_s}$, respectively. Note that $\mathcal{V}_{\mathcal{G}_s}$ in step 4 has already been updated to the vertex set where all vertices are sorted by the descending order of their degrees. Besides, $\widetilde{v}_i = \widetilde{v}_1 : \widetilde{v}_{\Phi-1}$ in step 4 means that vertex $\widetilde{v}_i$ will traverse the first $\Phi-1$ vertices (i.e., from the first vertex $\widetilde{v}_1$ to the $(\Phi-1)$th vertex $\widetilde{v}_{\Phi-1}$) of the updated vertex set $\mathcal{V}_{\mathcal{G}_s}$ with a step size of 1. $S_{\widetilde{v}_i}$ is the serial number of vertex $\widetilde{v}_i$ in $\mathcal{V}_{\mathcal{G}_s}$, i.e., the smaller the serial number of a vertex, the higher its ranking in $\mathcal{V}_{\mathcal{G}_s}$. Therefore, vertex $\widetilde{v}_j$ in step 5 denotes the vertex after $\widetilde{v}_i$ in $\mathcal{V}_{\mathcal{G}_s}$. Γ is the number of finally formed groups. Similar to the number of finally formed clusters Φ after Algorithm 2, Γ is not specified but determined by the distributions of dense SCNs. Ω_{i_1} denotes group i_1, where $i_1 \in \{1, \ldots, \Gamma\}$. α and r_0 denote the quotient and the remainder of the number of subchannels N divided by the number of groups Γ, respectively. Consequently, if $r_0 = 0$, subchannel allocation will be performed in α procedures, i.e., in procedure i_2, subchannel $i_1 + (i_2 - 1)\Gamma$ is allocated to all vertices in group Ω_{i_1}, where $i_1 \in \{1, \ldots, \Gamma\}$, and $i_2 \in \{1, \ldots, \alpha\}$; otherwise, subchannel allocation will be performed in $\alpha + 1$ procedures, i.e., the first α procedures are the same as those performed above, and in procedure $\alpha + 1$, subchannel $\alpha\Gamma + i_3$ is allocated to all vertices in group Ω_{i_3}, where $i_3 \in \{1, \ldots, r_0\}$.

Note that Algorithm 3 outputs the Γ formed groups and the corresponding subchannels allocated to each group. Finally, according to (3), (4), and the output of Algorithm 3, the total number of achievable data streams in dense SCNs can be obtained. Here it is noteworthy that executing Algorithm 3 only requires to estimate the perfect CSI within each cluster, which also avoids estimating the perfect global CSI and hence greatly reduces the feedback overhead. More detail about the estimation of perfect CSI within each cluster in practical systems can be found in [28].

V. COMPUTATIONAL COMPLEXITY ANALYSIS

From the proof of Lemma 1 it can be known that obtaining the optimal solution to optimization problem (8) by exhaustive search requires to compare more than $\left[(\mathsf{M}-1)2^N\Psi\right]^\Phi K^K$ cases, which is exponentially increases with both K and N. So obtaining the optimal solution to (8) needs a computational complexity higher than $\mathcal{O}\left(\left[(\mathsf{M}-1)2^N\Psi\right]^\Phi K^K\right)$, which is prohibitive in practical communication systems.

Next, we analyze the computational complexity of proposed solution. In Algorithm 1, partitioning graph $\mathcal{G}$ into subgraphs $\mathcal{G}_1$ and $\mathcal{G}_2$ by S-W algorithm in step 6 requires a complexity not exceeding $\mathcal{O}\left(\frac{K^2(K-1)}{2} + K^2\log_2 K\right)$ [34], which is lower than $\mathcal{O}\left(K^3\right)$. Moreover, both the number of subgraphs that need to be partitioned and the number of vertices in each subgraph are much smaller than K, so steps 8-15 have a complexity lower than $\mathcal{O}\left(K^3\right)$. Besides, the complexity of step 15 is lower than $\mathcal{O}\left(K^2\right)$, and steps 16-22 requires

Algorithm 3 Subchannel Allocation by the Proposed Graph Coloring Algorithm

1: Determine $\mathbf{A}_{\mathcal{G}_s}$ according to Definition 3.
2: Compute the degree of each vertex by $\mathcal{V}_{\mathcal{G}_s}$ according to (22).
3: Sort all vertices in $\mathcal{V}_{\mathcal{G}_s}$ by the descending order of their degrees, and set the obtained vertex set as $\mathcal{V}_{\mathcal{G}_s}$.
4: **for** $\widetilde{v}_i = \widetilde{v}_1 : \widetilde{v}_{\Phi-1}$ **do**
5: Let $\widetilde{v}_j = \widetilde{v}_{i+1}$.
6: **if** $\deg(\widetilde{v}_i) = \deg(\widetilde{v}_j)$ **then**
7: **if** $\mathsf{d}_{\widetilde{v}_i} > \mathsf{d}_{\widetilde{v}_j}$ and $S_{\widetilde{v}_i} > S_{\widetilde{v}_j}$, or $\mathsf{d}_{\widetilde{v}_i} < \mathsf{d}_{\widetilde{v}_j}$ and $S_{\widetilde{v}_i} < S_{\widetilde{v}_j}$ **then**
8: Exchange the order of vertices $\widetilde{v}_i$ and $\widetilde{v}_j$.
9: **end if**
10: **end if**
11: **end for**
12: Set $\Gamma = 1$.
13: **while** $|\mathcal{V}_{\mathcal{G}_s}| > 0$ **do**
14: Let $\Omega_\Gamma = \Omega_\Gamma \bigcup \{\widetilde{v}_1\}$ and $\mathcal{V}_{\mathcal{G}_s} = \mathcal{V}_{\mathcal{G}_s} \setminus \{\widetilde{v}_1\}$.
15: **while** $\exists \widetilde{v}_j \in \mathcal{V}_{\mathcal{G}_s}$, s.t. $\forall \widetilde{v}_i \in \Omega_\Gamma$, $a_{\widetilde{v}_i\widetilde{v}_j} = 0$ **do**
16: Update $\Omega_\Gamma = \Omega_\Gamma \bigcup \{\widetilde{v}_j\}$, and $\mathcal{V}_{\mathcal{G}_s} = \mathcal{V}_{\mathcal{G}_s} \setminus \{\widetilde{v}_j\}$.
17: **end while**
18: Update $\Gamma = \Gamma + 1$.
19: **end while**
20: Let $\alpha = \lfloor \frac{N}{\Gamma} \rfloor$ and $r_0 = N - \alpha \cdot \Gamma$.
21: Allocate subchannel $i_1 + (i_2 - 1)\Gamma$ to all vertices in group Ω_{i_1}, where $i_1 \in \{1, \ldots, \Gamma\}$, and $i_2 \in \{1, \ldots, \alpha\}$.
22: **if** $r_0 \neq 0$ **then**
23: Allocate subchannel $\alpha\Gamma + i_3$ to all vertices in group Ω_{i_3}, where $i_3 \in \{1, \ldots, r_0\}$.
24: **end if**
25: **Output:** The Γ formed groups $\Omega_1, \Omega_2, \ldots, \Omega_\Gamma$ and the corresponding subchannels allocated to each group.

a complexity of $\mathcal{O}(K)$. Consequently, Algorithm 1 has a complexity lower than $\mathcal{O}\left(K^3\right)$. In Algorithm 2, the complexity of steps 1-6 is $\mathcal{O}\left(\sum_{i_0=0}^{\Phi_{\min}-1}\binom{K-i_0\cdot\mathsf{M}}{\mathsf{M}}\right)$. Obviously, $\mathcal{O}\left(\sum_{i_0=0}^{\Phi_{\min}-1}\binom{K-i_0\cdot\mathsf{M}}{\mathsf{M}}\right) < \mathcal{O}\left(\Phi_{\min}\cdot\binom{K}{\mathsf{M}}\right)$, where $\Phi_{\min}$ determined by step 6 of Algorithm 2 is the minimum number of clusters performing IA, and $\binom{K}{\mathsf{M}} = \frac{K!}{\mathsf{M}!(K-\mathsf{M}!)}$ denotes the number of cases that choose M vertices from K vertices. Moreover, step 8-12 also has a complexity lower than $\mathcal{O}\left(\Phi_{\min}\cdot\binom{K}{\mathsf{M}}\right)$. So Algorithm 2 requires a complexity lower than $\mathcal{O}\left(\Phi_{\min}\cdot\binom{K}{\mathsf{M}}\right)$. In Algorithm 3, as $|\mathcal{V}_{\mathcal{G}_s}|$ is much smaller than K, the complexity of steps 1-11 is lower than $\mathcal{O}(K^2)$. Furthermore, steps 13-19 requires a complexity much lower than $\mathcal{O}(K^2)$, and steps 20-24 has a complexity of $\mathcal{O}(N\Psi)$. So the complexity of Algorithm 3 is lower than $\mathcal{O}(K^2 + N\Psi)$. Finally, the computational complexity of proposed solution is lower than $\mathcal{O}\left(K^3 + \Phi_{\min}\cdot\binom{K}{\mathsf{M}} + N\Psi\right)$, and it is also much lower than that of the exhaustive search.

From Section II we know that there is no related work about resource allocation based on IA with clustering to maximize the number of data streams achieved by all SUEs while guaranteeing the data stream requirement for each SUE. Therefore, as effective benchmarks, solutions obtained by clustering algorithms proposed in other literature and Algorithm 3 in this paper or Algorithms 1 and 2 in this paper and graph coloring algorithms proposed in other literature are used for the comparisons of computational complexity. Specifically, the computational complexities of following three related solutions are analyzed.

Firstly, similarity clustering algorithms proposed in [28] are used for grouping SBS-SUE pairs into disjoint clusters, then each cluster performing IA is further treated as a vertex, and Algorithm 3 is used for allocating subchannels to the formed clusters performing IA ([28]+Alg. 3). Similarity clustering algorithms in [28] are comprised of similarity clustering for SBS-SUE pairs and further adjusting the clusters with sizes not meeting the feasibility condition of IA, which require complexities lower than $\mathcal{O}(K\Phi_{\min}T_1T_2)$ and $\mathcal{O}\left(K^2\Phi_{\min}\right)$, respectively. Here T_1 and T_2 are the number of iterations taken by Lanczos and K-means algorithms to converge [28], respectively. Besides, recall that the complexity of Algorithm 3 is lower than $\mathcal{O}(K^2 + N\Psi)$. Consequently, [28]+Alg. 3 requires a complexity lower than $\mathcal{O}(K\Phi_{\min}T_1T_2 + K^2\Phi_{\min} + N\Psi)$.

Secondly, Algorithms 1 and 2 are used for clustering, then each cluster performing IA is further treated as a vertex, and coloring algorithm proposed in [36] is used for subchannel allocation (Algs. 1+2+ [36]). The overall complexity of Algorithms 1 and 2 is $\mathcal{O}\left(K^3 + \Phi_{\min}\cdot\binom{K}{\mathsf{M}}\right)$. Moreover, from [36] we know that its proposed coloring algorithm requires a complexity of $\mathcal{O}(K^2)$. However, when subchannel allocation based on IA with clustering is performed, steps 8-14 of the coloring algorithm in [36] requires an extra complexity of $\mathcal{O}(\Phi\Psi)$, where $\Phi_{\min} \leq \Phi \leq \lceil\frac{K}{2}\rceil$. Accordingly, the complexity of Algs. 1+2+ [36] is $\mathcal{O}\left(K^3 + \Phi_{\min}\cdot\binom{K}{\mathsf{M}} + \Phi\Psi\right)$.

Thirdly, similar to Algs. 1+2+[36], Algorithms 1 and 2 are used for clustering, then each cluster performing IA is further treated as a vertex, and coloring algorithm proposed in [37] is used for subchannel allocation (Algs. 1+2+[37]). From [37] it can be known that the complexity of its proposed coloring algorithm is $\mathcal{O}(K^2)$. Similarly, steps 6-12 of the coloring algorithm in [37] requires an extra complexity of $\mathcal{O}(\Phi\Psi)$ when performing subchannel allocation based on IA with clustering. Therefore, Algs. 1+2+[37] also has a complexity of $\mathcal{O}\left(K^3 + \Phi_{\min}\cdot\binom{K}{\mathsf{M}} + \Phi\Psi\right)$.

In conclusion, the proposed solution, Algs. 1+2+[36] and Algs. 1+2+ [37] have the same computational complexities. Furthermore, when the number of SUEs K is bigger, [28]+Alg. 3 will require a much higher complexity than the proposed solution because it will take lots of time to make K-means algorithm converge in [28]+Alg. 3; otherwise, their complexities are considered to be the same.

VI. FEEDBACK OVERHEAD ANALYSIS

In this section, we analyze the feedback overhead required by estimating perfect global CSI as well as the proposed solution, respectively.

Estimating the perfect global CSI requires to obtain the channel matrix from each SBS to each SUE over each subchannel. In our scenario, there are N subchannels available and K SBSs, each only serving one SUE in a given time slot of subchannel allocation. Besides, over each subchannel, there are K^2 channel matrices that need to be estimated, and each channel matrix consists of $M_T \times M_R$ channel coefficients. We assume that the feedback overhead of estimating each channel coefficient is quantified to B bit, so the total feedback overhead of estimating the perfect global CSI is $NK^2M_TM_RB$ bit.

Executing Algorithms 1 and 2 only require to estimate the path loss from each SBS to the $K-1$ SUEs it does not serve, which requires a feedback overhead of $(K-1)^2B$ bit. Moreover, from steps 20-24 of Algorithm 3 we know that the subchannel allocation for the groups will be performed in $\alpha+1$ procedures at most, which implies that each cluster performing IA will be allocated with $\alpha+1$ subchannels at most. Therefore, executing Algorithm 3 only requires to estimate the perfect CSI within each cluster over $\alpha+1$ subchannels at most, which needs a feedback overhead not exceeding $\Phi_{\min}(\alpha+1)\mathsf{M}^2M_TM_RB$ bit, where $\Phi_{\min} = \lceil \frac{K}{\mathsf{M}} \rceil$ denotes the minimum number of finally formed clusters performing IA, and $\alpha+1 = \lceil \frac{N}{\Gamma} \rceil$. So the total feedback overhead required by the proposed solution does not exceed $(K-1)^2B+\lceil \frac{K}{\mathsf{M}} \rceil \lceil \frac{N}{\Gamma} \rceil \mathsf{M}^2M_TM_RB \approx (K-1)^2B+K\mathsf{M}^2M_TM_RB$ bit, which is notably reduced compared with that required by estimating the perfect global CSI especially when K is bigger.

VII. NUMERICAL RESULTS

In our simulation, all SBSs are deployed in a single macrocell, and their locations are modeled as an independent homogeneous Poisson point process (PPP) with density $\lambda = \frac{K}{\pi r^2}$, where r denotes the macrocell radius. Here it is noteworthy that each SBS only serves one SUE in a given time slot of resource allocation, so λ also represents the density of SUEs performing IA. Table 2 lists the system parameters that keep unchanged in the simulation. In the simulation, the algorithm proposed in [38] and the formula given in [17] are respectively utilized to determine the precoding and interference suppression matrices of IA, which is also adopted in our previous work [28]. This is because when the number of users performing IA exceeds 3, there will not exist the closed forms of precoding and interference suppression matrices of IA [38]; besides, instead of designing the precoding and interference suppression matrices of IA, we focus on clustering and subchannel allocation for the formed clusters performing IA with limited subchannel resources to maximize the total number of data streams in dense SCNs.

In the simulation, all SBSs are deployed in indoor environment, and the percentages of SUEs distributed in indoor and outdoor environments are 80% and 20% [39], respectively. Besides, in a given time slot of subchannel allocation, each SBS only serves 1 SUE that randomly locates in a circular disc around this SBS with an inner and outer radii of 3 m and 10 m, respectively [40]. In this paper, we use the dual strip model with urban deployment in [41] to model the propagation environment. Firstly, when SBS k and SUE k that is exclusively served by SBS k in a given time slot are in the same apartment, PL_{kk} is given by

$$\mathsf{PL}_{kk}(\text{dB}) = 38.46 + 20\log_{10}d_{kk} + 0.7d_{\text{2D, indoor}} + qL_{\text{iw}}, \tag{23}$$

where d_{kk} represents the distance between SBS k and SUE k, and $d_{\text{2D, indoor}}$ denotes the distance inside the apartment, and q is the number of inner walls separating SBS k and SUE k, and L_{iw} with a value of 5 dB denotes the penetration loss of an inner wall. Besides, when SUE k is in outdoor environment, PL_{kk} becomes

$$\begin{aligned}\mathsf{PL}_{kk}(\text{dB}) &= \max\left(15.3 + 37.6\log_{10}d_{kk}, 38.46 + 20\log_{10}d_{kk}\right) \\ &\quad + 0.7d_{\text{2D, indoor}} + qL_{\text{iw}} + L_{\text{ow}},\end{aligned} \tag{24}$$

where L_{ow} with a value of 20 dB is the penetration losses of an outdoor wall. When SBS m and SUE k that is not served by SBS m are in the same apartment or SUE k is in outdoor environment, the expression of path loss from SBS m to SUE k, i.e., PL_{km}, is similar to that of PL_{kk} in (23) or (24), respectively. In addition, when SBS m and SUE k are respectively in two different apartments, PL_{km} becomes

$$\begin{aligned}\mathsf{PL}_{km}(\text{dB}) &= \max\left(15.3 + 37.6\log_{10}d_{km}, 38.46 + 20\log_{10}d_{km}\right) \\ &\quad + 0.7d_{\text{2D, indoor}} + qL_{\text{iw}} + L_{\text{ow},1} + L_{\text{ow},2}\end{aligned} \tag{25}$$

where $L_{\text{ow},1}$ and $L_{\text{ow},2}$ are the penetration losses due to outdoor walls for the two apartments.

From Table 2 we know that the noise power at each SUE over each subchannel is $N_0 \cdot \Delta f = -121.4$ dBm. Therefore, the given threshold PL_0 should meet that $\frac{p}{\mathsf{PL}_0} \geq -121.4$ dBm. As $p = 30$ dBm, we have $\mathsf{PL}_0 \leq 151.4$ dB. However, due to the postprocessing of interference suppression matrices for performing IA at each SUE, the strength of inter-cluster interference at each SUE also experiences a certain degree of attenuation. Consequently, here PL_0 is set

TABLE 2. System parameters.

Parameter	Value
Carrier frequency	2 GHz
Noise power density, N_0	-174 dBm/Hz
Subchannel bandwidth, Δf	0.18 MHz
Macrocell radius	150 m
Number of receive antennas at each SUE, M_R	2
Transmit power of each SBS, p	30 dBm (1 W)
Number of data stream requirement for each SUE, d	1

to 135 dB, which leaves some margin for the aforementioned attenuation.

In the simulation, we compare the performances of following six schemes. Here it is worth mentioning that in dense SCNs, using exhaustive search to find the optimal solution to problem (8) will lead to unaffordable computational complexity. So we utilize Algorithms 1 and 2 to group all SBS-SUE pairs into disjoint clusters, and then utilize exhaustive search to obtain the optimal solution to the subproblem of subchannel allocation for the formed clusters (OSAC), i.e., subproblem (21), which can be regarded as an efficient approximation of the optimal solution to problem (8).

1) **OSAC**.
2) **Proposed Solution**.
3) **[28]+Alg. 3**.
4) **Algs. 1+2+[36]**.
5) **Algs. 1+2+[37]**.
6) **Random Clustering**. All SBS-SUE pairs are randomly grouped into disjoint clusters, each of which has a size that meets constraint C3 of problem (8), and then Algorithm 3 is utilized for subchannel allocation.

It is also noteworthy that the performance curve of OSAC and those of five other solutions are obtained by taking the average of 5 and 200 realizations, respectively. This is because evaluating the performance of OSAC needs huge time. In each realization, the distributions of all SBSs and SUEs as well as the channel matrix gain between each SBS and each SUE over each subchannel are varied. Accordingly, the number of data streams achieved by each solution under various parameters is not necessarily integers.

Fig. 7 shows the number of achievable data streams against the number of SUEs, where $N = 8$, $\mathsf{PL}_0 = 135$ dB and $M_T = 3$. Here the number of achievable data streams in the Y-axis is referred to as the total number of data streams achieved by the SUEs that have been allocated with subchannels in dense SCNs. As each SUE only has 1 data stream requirement, from (2) we know that $\mathsf{M} = 4$. When the number of SUEs K increases from 30 to 60 with a step size

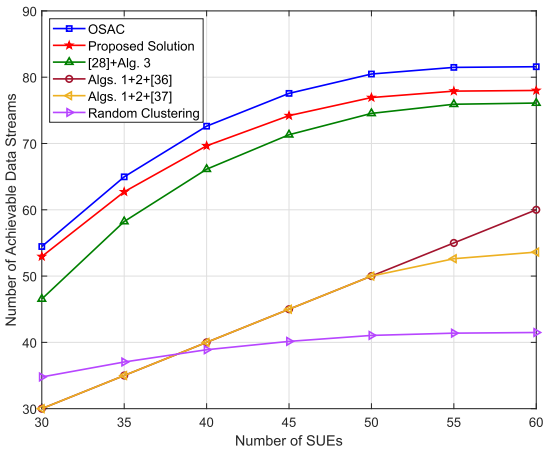

FIGURE 7. Number of achievable data streams vs. number of SUEs K.

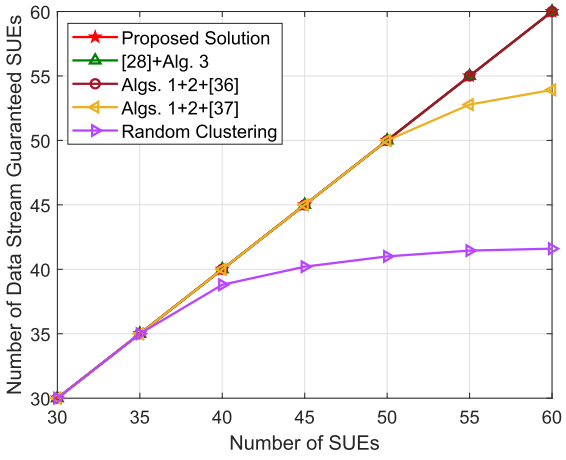

FIGURE 8. Number of data stream guaranteed SUEs vs. number of SUEs K.

of 5, λ takes the values of 0.4246×10^{-3}, 0.4954×10^{-3}, 0.5662×10^{-3}, 0.6369×10^{-3}, 0.7077×10^{-3}, 0.7785×10^{-3} and 0.8493×10^{-3}, respectively. Firstly, we can observe that the number of data streams achieved by OSAC, the proposed solution and [28]+Alg. 3 increases with K but at a much slower pace when $K \geq 45$, and hardly increases when $K \geq 55$. The reason is that when K is relatively small, e.g., $K \leq 45$, the inter-cluster interference is relatively weak, so each subchannel can be reused by more clusters; otherwise, orthogonal subchannel allocation for disjoint clusters becomes dominant to mitigate the stronger inter-cluster interference, so only 8 subchannels available is not enough to support the rapid increase in the number of achievable data streams. Furthermore, the proposed solution outperforms [28]+Alg. 3, but their performance gap becomes smaller as K increases. We can also observe that when $K \leq 50$, the performances of both Algs. 1+2+[36] and Algs. 1+2+[37] exhibit a linear increase but are still much worse than that of proposed solution because the coloring algorithms in both [36] and [37] allocate only one subchannel to each cluster; when $K > 50$, Algs. 1+2+[37] has a worse performance than Algs. 1+2+[36] as all vertices are not sorted by the descending order of their degrees in the coloring algorithms proposed by [37]. Finally, OSAC outperforms the proposed solution especially when K is bigger; however, it will require a much higher computational complexity.

Fig. 8 plots the number of data stream guaranteed SUEs against the number of SUEs, where $N = 8$, $\mathsf{PL}_0 = 135$ dB and $M_T = 3$. Here the number of data stream guaranteed SUEs is referred to as the number of SUEs that can achieve the minimum data stream requirements in dense SCNs, i.e., d simultaneously transmitted interference-free data streams from their serving SBSs. Generally, the number of data stream guaranteed SUEs achieved by all the five solutions increases with K. Furthermore, even though K increases, the proposed solution, [28]+Alg. 3 and Algs. 1+2+[36] enable all SUEs to achieve satisfactory connectivity requirements; however, when $K > 50$, this is unable to be guaranteed by

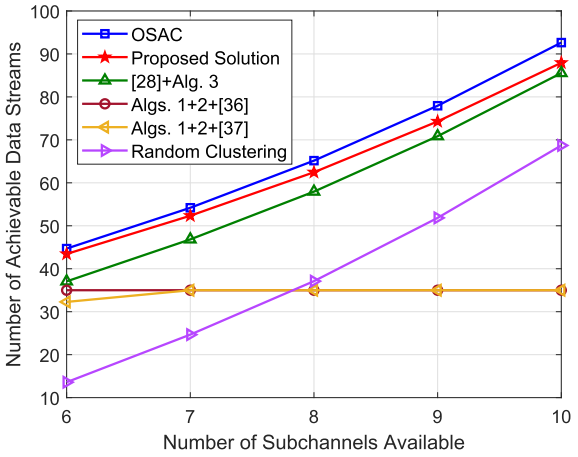

FIGURE 9. Number of achievable data streams vs. number of subchannels available N.

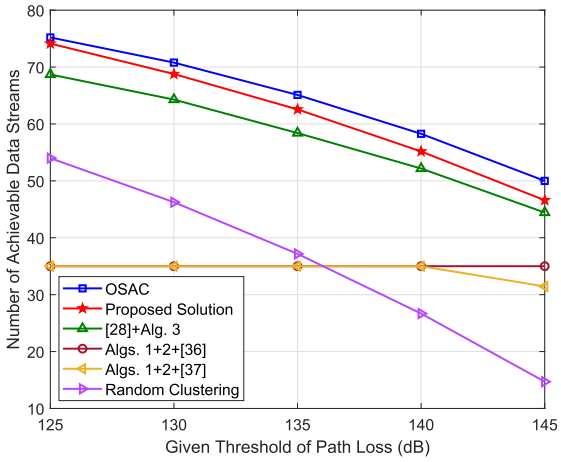

FIGURE 10. Number of achievable data streams vs. the given threshold of path loss PL_0.

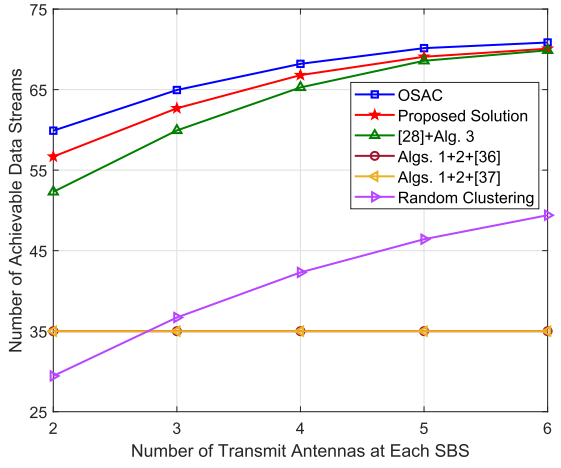

FIGURE 11. Number of achievable data streams vs. the number of transmit antennas at each SBS M_T.

Algs. 1+2+[37] with only 8 subchannels available. Finally, when $K > 35$, the solution of random clustering exhibits the worst performance, and it only makes some of the SUEs meet their connectivity requirements. This is because the intra-cluster interference within each cluster formed by the solution of random clustering may be weak, while the inter-cluster interference tends to be strong enough, and more clusters will be allocated with orthogonal subchannels for inter-cluster interference mitigation, in this way, some clusters cannot be allocated with even one subchannel.

It is noteworthy that when K is bigger, OSAC still has a much higher computational complexity, which implies that evaluating its performance will need huge time. Consequently, K is set to a relatively small value (i.e., $K = 35$) in the following simulation.

Fig. 9 investigates the number of achievable data streams with respect to the number of subchannels available, where $K = 35$, $\mathsf{PL}_0 = 135$ dB, $M_T = 3$ and $\lambda = 0.4954 \times 10^{-3}$. We can observe that the number of data streams achieved by OSAC, the proposed solution, [28]+Alg. 3 and random clustering almost exponentially increases with N. However, when $N \geq 7$, the number of data streams achieved by Algs. 1+2+[36] and Algs. 1+2+[37] always keeps constant. The reason is that the coloring algorithms proposed by [36] and [37] require each cluster to be allocated with only one subchannel. In addition, although the solution of random clustering mitigates some of the relatively weak inter-cluster interference by orthogonal subchannel allocation, it enables each cluster to be allocated with more subchannels as N increases, so it outperforms both Algs. 1+2+[36] and Algs. 1+2+[37] when $N \geq 7$. Finally, the performance gap between OSAC and proposed solution becomes a little larger as N increase, but the computational complexity of proposed solution is much lower than that of OSAC.

Fig. 10 plots the number of achievable data streams against the given threshold of path loss, where $K = 35$, $N = 8$, $M_T = 3$ and $\lambda = 0.4954 \times 10^{-3}$. As PL_0 increases, more inter-cluster interference which will be considered to be strong enough is mitigated through allocating orthogonal subchannels; therefore, the 8 subchannels are utilized with a lower efficiency, and OSAC, the proposed solution, [28]+Alg. 3 and random clustering experience significant performance degradation. Besides, when 125 dB $\leq \mathsf{PL}_0 \leq$ 140 dB, the number of data streams achieved by Algs. 1+2+[36] and Algs. 1+2+[37] is always 35; however, when $\mathsf{PL}_0 >$ 140 dB, Algs. 1+2+ [37] becomes unable to support the connectivity requirements for all SUEs only with 8 subchannels. We can also notice that the proposed solution exhibits a performance better than [28]+Alg. 3. This is mainly because our clustering criterion in (15) and Algorithms 1 and 2 can guarantee that the intra-cluster interference is strong enough, which necessitates performing IA within each cluster; moreover, instead of being a priori, the number of clusters formed after Algorithms 1 and 2 is determined by the distributions of dense SCNs, which can utilize the limited subchannels more efficiently.

Fig. 11 presents the number of achievable data streams with respect to the number of transmit antennas at each SBS, where

TABLE 3. Elapsed time of obtaining all the six solutions under various parameters (s).

Parameter \ Solution		OSAC	Proposed Solution	[28]+Alg. 3	Algs. 1+2+[36]	Algs. 1+2+[37]	Random Clustering
$N = 8$, $\mathsf{PL}_0 = 135$ dB, $M_T = 3$	$K = 30$	216754	5.3917	5.1052	5.4720	5.3251	3.8725
	$K = 35$	235736	8.4152	8.4093	8.3763	8.5167	5.3207
	$K = 40$	256437	12.1834	11.5960	12.5680	12.4078	7.7086
	$K = 45$	278653	17.6351	18.6721	17.2913	17.9279	9.8513
	$K = 50$	302189	23.8993	35.8829	23.5201	24.1894	12.9758
	$K = 55$	328145	31.3019	61.0273	31.0019	30.9218	16.7852
	$K = 60$	357632	40.2095	116.6370	39.6720	40.4053	22.6718
$K = 35$, $\mathsf{PL}_0 = 135$ dB, $M_T = 3$	$N = 6$	57632	7.0287	6.8329	6.9204	6.9553	4.2357
	$N = 7$	118243	7.8649	7.9853	8.1896	7.9253	4.7498
	$N = 8$	235109	8.6582	8.5372	8.4306	8.6791	5.2319
	$N = 9$	472617	9.0473	9.2507	9.2637	8.9024	5.6641
	$N = 10$	937158	9.7215	9.6324	9.5327	9.5908	6.1235
$K = 35$, $N = 8$, $M_T = 3$	$\mathsf{PL}_0 = 125$ dB	215378	7.4968	7.3026	7.3896	7.4913	5.5027
	$\mathsf{PL}_0 = 130$ dB	222105	7.7628	7.9217	7.9938	7.8339	5.5874
	$\mathsf{PL}_0 = 135$ dB	235984	8.1582	8.3705	8.0137	8.3694	5.4685
	$\mathsf{PL}_0 = 140$ dB	259162	9.5387	11.0616	9.7108	9.2971	5.6092
	$\mathsf{PL}_0 = 145$ dB	298757	12.8024	14.9807	12.2360	12.3486	5.4391
$K = 35$, $N = 8$, $\mathsf{PL}_0 = 135$ dB	$M_T = 2$	156289	4.9132	4.8758	4.7089	4.8256	3.1783
	$M_T = 3$	235617	8.4005	8.6824	8.5007	8.3561	5.1938
	$M_T = 4$	351267	12.9276	12.4029	13.2660	13.1593	8.2597
	$M_T = 5$	529706	18.2739	18.6270	18.0031	18.6928	12.8230
	$M_T = 6$	790493	25.3481	25.7945	25.7830	24.7404	17.9368

$K = 35$, $N = 8$, $\mathsf{PL}_0 = 135$ dB and $\lambda = 0.4954 \times 10^{-3}$. From (2) we know that when M_T increases from 2 to 6, M, i.e., the maximum cluster size, will increase from 3 to 7, which indicates that each cluster can contain more SBS-SUE pairs to perform IA. Note that in OSAC and proposed solution, it is necessary to exploit IA for intra-cluster interference mitigation only when it is strong enough within each cluster; however, the increase in M_T has no influence on the strength of intra-cluster interference. Accordingly, as M_T increases, few clusters will contain more SBS-SUE pairs to perform IA in OSAC, proposed solution and [28]+Alg. 3 when $K = 35$, and their performances are improved at a relatively slow pace but still much better compared with that of random clustering. In addition, [28]+Alg. 3 provides a performance that gradually approaches to that of proposed solution as M_T becomes bigger, and they exhibit almost the same performances when $M_T \geq 5$. It is also noteworthy that the number of data streams achieved by Algs. 1+2+ [36] and Algs. 1+2+ [37] remains unchanged regardless of the increase in M_T. Finally, the proposed solution exhibits a performance much closer to OSAC as M_T increases, and it also requires a computational complexity much lower than OSAC.

In Table 3, we list the elapsed time of obtaining all the six solutions under various parameters. Note that same as obtaining the performance curves Figs. 7 to 11, each elapsed time of OSAC and those of five other solutions in the table is obtained by taking the average of 5 and 200 realizations, respectively. It can be observed that the elapsed time of obtaining OSAC is much longer than five other solutions, so OSAC is infeasible in practical communication systems. Besides, the elapsed time of obtaining the proposed solution, Algs. 1+2+ [36] and Algs. 1+2+ [37] is almost the same under various parameters, and the elapsed time of obtaining [28]+Alg. 3 is longer than that of proposed solution when $K \geq 50$. Finally, obtaining the solution of random clustering need the shortest time, however, it offers almost the worst performance under some parameters.

VIII. CONCLUSION

In this paper, we have allocated limited subchannel resources based on IA with clustering to maximize the number of data streams achieved by SUEs in dense SCNs while guaranteeing the data stream requirement for each SUE. Due to the NP-hardness of corresponding optimization problem, we have proposed to obtain its suboptimal but efficient solution based on graph theory through two phases: 1) clustering for SUEs by recursively partitioning the constructed graph which only requires to estimate the path losses from SBSs to SUEs; and 2) subchannel allocation for the formed clusters through the proposed graph coloring algorithm which only requires to estimate the perfect CSI within each cluster. Each phase required much lower complexity and notably reduced feedback and would be executed by a central entity called HeNB GW in practical communication systems. The computational complexity analysis has demonstrated that the complexity required by the proposed solution is much lower than that of the optimal solution obtained by exhaustive search. Finally, numerical results have shown that the proposed solution provides a performance better than other related schemes and close to the optimal solution.

APPENDIX A
PROOF OF LEMMA 2

According to the definitions of $\mathsf{W}_{\mathcal{C}_i}$ in (18) and cut $(\mathcal{C}_i, \mathcal{V} \setminus \mathcal{C}_i)$ in (20), we have

$$\begin{aligned}\sum_{\mathcal{C}_i \subseteq \mathcal{V}} \mathsf{W}_{\mathcal{C}_i} &= \sum_{\mathcal{C}_i \subseteq \mathcal{V}} \sum_{k,\, l \in \mathcal{C}_i} w_{kl} \\ &= \sum_{\mathcal{C}_i \subseteq \mathcal{V}} \left(\sum_{k \in \mathcal{C}_i,\, m \in \mathcal{V}} w_{km} - \sum_{k \in \mathcal{C}_i,\, q \in \mathcal{V} \setminus \mathcal{C}_i} w_{kq} \right) \\ &= \sum_{\mathcal{C}_i \subseteq \mathcal{V}} \sum_{k \in \mathcal{C}_i,\, m \in \mathcal{V}} w_{km} - \sum_{\mathcal{C}_i \subseteq \mathcal{V}} \text{cut}\,(\mathcal{C}_i, \mathcal{V} \setminus \mathcal{C}_i)\,. \quad (26)\end{aligned}$$

Note that for a given vertex set $\mathcal{V}$, $\sum_{\mathcal{C}_i \subseteq \mathcal{V}} \sum_{k \in \mathcal{C}_i,\, m \in \mathcal{V}} w_{km}$ is a constant, which implies that when $\sum_{\mathcal{C}_i \subseteq \mathcal{V}} \mathsf{W}_{\mathcal{C}_i}$ achieves the maximum value, $\sum_{\mathcal{C}_i \subseteq \mathcal{V}} \text{cut}\,(\mathcal{C}_i, \mathcal{V} \setminus \mathcal{C}_i)$ will achieve the minimum value. Hence the proof is completed. ■

APPENDIX B
PROOF OF THEOREM 1

We first prove that when the weighted adjacency matrix of an arbitrary subgraph partitioned from graph $\mathcal{G}$, e.g., $\mathcal{G}_1$, has at least one element with value $-\mathsf{C}_0$, the edge set corresponding to its minimum cut computed by S-W algorithm must contain at least one edge with weight $-\mathsf{C}_0$. Assume the edge set corresponding to the minimum cut of subgraph $\mathcal{G}_1$ contains no edges with weight $-\mathsf{C}_0$. Therefore, the minimum cut of subgraph $\mathcal{G}_1$ in this case is bigger than 0. However, because $-\mathsf{C}_0$ is a sufficiently small negative number, the cut of $\mathcal{G}_1$ will be much smaller than 0 when there is at least one edge with weight $-\mathsf{C}_0$ in this cut, which contradicts the initial assumption. So the number of edges with weight $-\mathsf{C}_0$ in $\mathcal{G}_1$ decreases as the times of recursively executing steps 2-6 of Algorithm 1 increase. After finite times of recursively partitioning graph $\mathcal{G}$, there will be no edges with weight $-\mathsf{C}_0$ in each obtained subgraph, which implies that the path loss between any two vertices in each obtained subgraph (i.e., a cluster) is smaller than or equal to PL_0 after Algorithm 1. ■

REFERENCES

[1] *Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2017–2022*, Cisco VNI Forecast, Cisco, Vis. Netw. Index, San Jose, CA, USA, Feb. 2019.

[2] A. H. Jafari, D. López-Pérez, M. Ding, and J. Zhang, "Performance analysis of dense small cell networks with practical antenna heights under Rician fading," *IEEE Access*, vol. 6, pp. 9960–9974, 2018.

[3] H. Tullberg, P. Popovski, Z. Li, M. A. Uusitalo, A. Höglund, Ö. Bulakci, M. Fallgren, and J. F. Monserrat, "The METIS 5G system concept: Meeting the 5G requirements," *IEEE Commun. Mag.*, vol. 54, no. 12, pp. 132–139, Dec. 2016.

[4] J. Zhang, Y. Zhang, Y. Yu, R. Xu, Q. Zheng, and P. Zhang, "5G: A tutorial overview of standards, trials, challenges, deployment, and practice," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 6, pp. 1201–1221, Jun. 2017.

[5] J. Chen, X. Ge, and Q. Ni, "Coverage and handoff analysis of 5G fractal small cell networks," *IEEE Trans. Wireless Commun.*, vol. 18, no. 2, pp. 1263–1276, Feb. 2019.

[6] M. Ding, D. López-Pérez, H. Claussen, and M. A. Kaafar, "On the fundamental characteristics of ultra-dense small cell networks," *IEEE Netw.*, vol. 32, no. 3, pp. 92–100, May/Jun. 2018.

[7] J. Shi, Z. Song, and Q. Ni, "Distributed resource allocation assisted by intercell interference mitigation in downlink multicell MC DS-CDMA systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 2, pp. 1250–1266, Feb. 2017.

[8] X. Lu, Q. Ni, W. Li, and H. Zhang, "Dynamic user grouping and joint resource allocation with multi-cell cooperation for uplink virtual MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 6, pp. 3854–3869, Jun. 2017.

[9] E. Hossain, M. Rasti, H. Tabassum, and A. Abdelnasser, "Evolution toward 5G multi-tier cellular wireless networks: An interference management perspective," *IEEE Wireless Commun.*, vol. 21, no. 3, pp. 118–127, Jun. 2015.

[10] Y. Hao, Q. Ni, H. Li, and S. Hou, "Robust multi-objective optimization for EE-SE tradeoff in D2D communications underlaying heterogeneous networks," *IEEE Trans. Commun.*, vol. 66, no. 10, pp. 4936–4949, Oct. 2018.

[11] V. R. Cadambe and S. A. Jafar, "Interference alignment and degrees of freedom of the K-user interference channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3425–3441, Aug. 2008.

[12] N. Zhao, F. R. Yu, M. Jin, Q. Yan, and V. C. M. Leung, "Interference alignment and its applications: A survey, research issues, and challenges," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 3, pp. 1779–1803, 3rd Quart., 2016.

[13] C. M. Yetis, T. Gou, S. A. Jafar, and A. H. Kayran, "On feasibility of interference alignment in MIMO interference networks," *IEEE Trans. Signal Process.*, vol. 58, no. 9, pp. 4771–4782, Sep. 2010.

[14] R. Tresch and M. Guillaud, "Clustered interference alignment in large cellular networks," in *Proc. IEEE 20th Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC)*, Sep. 2009, pp. 1024–1028.

[15] R. Tresch and M. Guillaud, "Performance of interference alignment in clustered wireless ad hoc networks," in *Proc. IEEE ISIT*, Jun. 2010, pp. 1703–1707.

[16] *Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN) (Release 9)*, document TS 36.300, 3rd Generation Partnership Project, Sep. 2013.

[17] S. Chen and R. S. Cheng, "Clustering for interference alignment in multiuser interference network," *IEEE Trans. Veh. Technol.*, vol. 63, no. 6, pp. 2613–2624, Jul. 2014.

[18] N. Zhao, X. Zhang, F. R. Yu, and V. C. M. Leung, "To align or not to align: Topology management in asymmetric interference networks," *IEEE Trans. Veh. Technol.*, vol. 66, no. 8, pp. 7164–7177, Aug. 2017.

[19] Y. Luo, T. Ratnarajah, J. Xue, and F. A. Khan, "Interference alignment in two-tier randomly distributed heterogeneous wireless networks using stochastic geometry approach," *IEEE Syst. J.*, vol. 12, no. 3, pp. 2238–2249, Sep. 2018.

[20] K. Wang, F. R. Yu, L. Wang, J. Li, N. Zhao, Q. Guan, B. Li, and Q. Wu, "Interference alignment with adaptive power allocation in full-duplex-enabled small cell networks," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 3010–3015, Mar. 2019.

[21] C. Qin, C. Wang, D. Pan, W. Wang, and Y. Zhang, "A cross time slot partial interference alignment scheme in two-cell relay heterogeneous networks," *Appl. Sci.*, vol. 9, no. 4, p. 652, 2019.

[22] M. El-Absi, M. Shaat, F. Bader, and T. Kaiser, "Interference alignment with frequency-clustering for efficient resource allocation in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 12, pp. 7070–7082, Dec. 2015.

[23] Y. Meng, J. Li, H. Li, and M. Pan, "A transformed conflict graph-based resource-allocation scheme combining interference alignment in OFDMA Femtocell networks," *IEEE Trans. Veh. Technol.*, vol. 64, no. 10, pp. 4728–4737, Oct. 2015.

[24] H. Zhang, H. Li, and J. H. Lee, "Efficient subchannel allocation based on clustered interference alignment in ultra-dense femtocell networks," *China Commun.*, vol. 14, no. 4, pp. 1–10, Apr. 2017.

[25] J. Ma, S. Zhang, H. Li, N. Zhao, and V. C. M. Leung, "Interference-alignment and soft-space-reuse based cooperative transmission for multi-cell massive MIMO networks," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 1907–1922, Mar. 2018.

[26] R. Brandt, R. Mochaourab, and M. Bengtsson, "Distributed long-term base station clustering in cellular networks using coalition formation," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 2, no. 3, pp. 362–375, Sep. 2016.

[27] J. Xiao, C. Yang, A. Anpalagan, Q. Ni, and M. Guizani, "Joint interference management in ultra-dense small-cell networks: A multi-domain coordination perspective," *IEEE Trans. Commun.*, vol. 66, no. 11, pp. 5470–5481, Nov. 2018.

[28] H. Zhang, H. Li, J. H. Lee, and H. Dai, "QoS-based interference alignment with similarity clustering for efficient subchannel allocation in dense small cell networks," *IEEE Trans. Commun.*, vol. 65, no. 11, pp. 5054–5066, Nov. 2017.
[29] M. Zhou, H. Li, N. Zhao, S. Zhang, and F. R. Yu, "Feasibility analysis and clustering for interference alignment in full-duplex-based small cell networks," *IEEE Trans. Commun.*, vol. 67, no. 1, pp. 807–819, Jan. 2019.
[30] T. Ding, M. Ding, G. Mao, Z. Lin, and D. López-Pérez, "Uplink performance analysis of dense cellular networks with LoS and NLoS transmissions," *IEEE Trans. Wireless Commun.*, vol. 16, no. 4, pp. 2601–2613, Apr. 2017.
[31] A. Abdelnasser, E. Hossain, and D. I. Kim, "Clustering and resource allocation for dense femtocells in a two-tier cellular OFDMA network," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1628–1641, Mar. 2014.
[32] K. Gomadam, V. R. Cadambe, and S. A. Jafar, "A distributed numerical approach to interference alignment and applications to wireless interference networks," *IEEE Trans. Inf. Theory*, vol. 57, no. 6, pp. 3309–3322, Jun. 2011.
[33] S. A. Jafar and S. Shamai (Shitz), "Degrees of freedom region of the MIMO X channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 1, pp. 151–170, Jan. 2008.
[34] M. Stoer and F. Wagner, "A simple min-cut algorithm," *J. ACM*, vol. 44, no. 4, pp. 585–591, Jul. 1997.
[35] D. Brélaz, "New methods to color the vertices of a graph," *Commun. ACM*, vol. 22, no. 4, pp. 251–256, Apr. 1979.
[36] Y. Lin, R. Zhang, C. Li, L. Yang, and L. Hanzo, "Graph-based joint user-centric overlapped clustering and resource allocation in ultradense networks," *IEEE Trans. Veh. Technol.*, vol. 67, no. 5, pp. 4440–4453, May 2018.
[37] R. Wang, J. Wu, and J. Yan, "Resource allocation for D2D-enabled communications in vehicle platooning," *IEEE Access*, vol. 6, pp. 50526–50537, 2018.
[38] S. W. Peters and R. W. Heath, "Interference alignment via alternating minimization," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Apr. 2009, pp. 2445–2448.
[39] *Small Cell Enhancements for E-UTRA and E-UTRAN-Physical Layer Aspects (Release 12)*, document 3GPP, TR 36.872, Dec. 2013.
[40] A. Abdelnasser, E. Hossain, and D. I. Kim, "Tier-aware resource allocation in OFDMA macrocell-small cell networks," *IEEE Trans. Commun.*, vol. 63, no. 3, pp. 695–710, Mar. 2015.
[41] *Further Advancements for E-UTRA Physical Layer Aspects (Release 9)*, document 3GPP, TR 36.814, Mar. 2017.

HAO ZHANG received the B.S. degree in applied mathematics from Chang'an University, Xi'an, China, in 2009, and the M.S. and Ph.D. degrees in communication and information systems from Xidian University, Xi'an, in 2012 and 2017, respectively. From 2014 to 2016, he was a Visiting Ph.D. Student with the Department of Electrical and Computer Engineering, North Carolina State University, Raleigh, NC, USA. He is currently a Postdoctoral Researcher at the School of Marine Science and Technology, Northwestern Polytechnical University, Xi'an. His research interests include interference alignment, clustering, interference management, and resource allocation in dense small cell networks and ship networks under evaporation duct environment.

KUNDE YANG (M'13) received the B.S., M.S., and Ph.D. degrees in underwater acoustic engineering from Northwestern Polytechnical University, Xi'an, China, in 1996, 1999, and 2003, respectively. From 2006 to 2007, he was a Visiting Scholar with the School of Earth and Ocean Sciences, University of Victoria, Victoria, BC, Canada. Since 2003, he has been with the School of Marine Science and Technology, Northwestern Polytechnical University, where he is currently a Full Professor and the Director of the Department of Acoustic and Information Engineering. His research interests include communication for ship networks and microwave propagation in evaporation duct environment.

SHUN ZHANG received the B.S. degree in communication engineering from Shandong University, Jinan, China, in 2007, and the Ph.D. degree in communications and signal processing from Xidian University, Xi'an, in 2013. He is currently an Associate Professor with the State Key Laboratory of Integrated Services Networks, Xidian University. His research interests include MIMO-OFDM systems, relay networks, and detection and parameter estimation theory.