Cooperative Base Station Coloring for Pair-wise Multi-Cell Coordination

Jeonghun Park, Student Member, IEEE, Namyoon Lee, Student Member, IEEE, and Robert W. Heath Jr., Fellow, IEEE

Abstract

A semi-static base station (BS) coordination strategy exploits predefined multiple BS cluster patterns for cooperative transmissions to improve the cell-edge user throughput. This paper proposes an efficient BS cluster pattern for pair-wise semi-static BS coordination, where edge users are guaranteed to be protected from dominant out-of-cluster interference in an irregular cellular network. The key idea is that each BS cluster is formed by using the 2nd-order Voronoi region of BSs’ locations and formed BS clusters are assigned to multiple cluster patterns by using the edge-coloring. The main advantage of the proposed BS cluster pattern is that every user is guaranteed to be served by the two closest BSs irrespective of a BS deployment scenario. With the proposed coordination strategy, analytical expressions for the rate distribution and the ergodic spectral efficiency are derived as a function of relevant system parameters in a non-random network model with irregular BS locations. A lower bound on the ergodic spectral efficiency is characterized for a random network where locations of BSs and users are modeled by using a homogeneous Poisson point process. Through simulations, the analytical expressions are verified, and the performance is compared with that of conventional methods. Our major finding is that the proposed BS cluster pattern provides considerable performance gains in both the rate coverage probability and the ergodic spectral efficiency for cell-edge users compared to conventional strategies.

J. Park, N. Lee and R. W. Heath Jr. are with the Wireless Networking and Communication Group (WNCG), Department of Electrical and Computer Engineering, The University of Texas at Austin, TX 78712, USA. (E-mail: {jeonghun, namyoon.lee, rheath} @utexas.edu)

This research was supported by a gift from Huawei Technologies Co. Ltd.
I. INTRODUCTION

A. Motivation

Coordination among base stations (BSs) is a powerful approach for mitigating inter-cell interference in cellular systems [1]. It has shown that the sum spectral efficiency scales linearly with the signal-to-noise ratio (SNR) when all the BSs are coordinated [2]. In practice, however, coordination with a large number of BSs may not be feasible due to excessive overheads associated with the coordination, e.g., complexity, channel estimation and channel feedback [3]. One practical solution for implementing multicell coordination is to form a BS cluster so that a limited number of BSs are coordinated to control intra-cluster interference with a reasonable amount of overhead [3], [4]. Unfortunately, when clustering-based BSs coordination is applied in a static way [5], [6], the performance of a user is mainly limited by unmanageable out-of-cluster interference [3], [5], [6].

To protect a user effectively from out-of-cluster interference, a semi-static multicell coordination strategy was proposed in [7], [8]. The key idea is to use different predefined BS cluster patterns over multiple phases. This approach provides an opportunity for users to communicate with their favorable set of BSs in a particular phase, improving the cluster edge-user throughput [7], [8]. The performance of the semi-static multicell coordination strategy depends on how the BS cluster patterns are determined. The problem of designing the BS cluster pattern is therefore important, yet it is difficult in general cellular networks with irregular BS locations. The reason is that it becomes unclear how to form multiple BS cluster patterns so that every user obtains equivalent opportunities to communicate with their favorable set of BSs. In this paper, we provide a solution to the cluster pattern design problem for pair-wise BS coordination to improve the edge-user performance in cellular downlink networks with irregular BSs’ locations.

B. Related Work

The most efficient way to protect a user from out-of-cluster interference is dynamic BS clustering, where a BS cluster is formed in a user-centric way depending on the users’ conditions, e.g., distances to interfering BSs and channel quality. Cooperative strategies for dynamic BS clustering have been studied in [9]–[20]. In [9]–[12], an algorithm for making a BS cluster in a user-centric way was proposed. For instance, given the users’ conditions such as distances or
channel conditions, an algorithm selects BSs to include in a BS cluster so that the joint achievable throughput of the formed BS cluster is maximized. In [13], [14], a resource allocation method to enhance sum spectral efficiency was proposed for dynamic BS clustering. By jointly allocating resources to each BS cluster depending on their channel qualities, the sum spectral efficiency in a BS cluster can be improved. In [15]–[17], a beamforming strategy was proposed in a dynamic BS clustering assumption for complexity reduction or sum spectral efficiency improvement. In [18]–[20], the performance of dynamic BS clustering was characterized in a network with randomly located base stations. In general, the main limitations of dynamic clustering are the overheads in dynamically forming BS clusters and the high complexity of the user scheduling [21]. For example, in [15], [16], the transmitted data is sent to other BSs through backhaul within the same BS cluster, but the BS cluster changes by users’ conditions. A centralized unit [22], therefore, should control the backhaul traffic whenever the instantaneous channel conditions of each user change, which causes high complexity. In [18], [19], each user selects a serving BS set, in which users should be carefully scheduled to prevent conflict, i.e., the situation where multiple users want the same BS simultaneously.

Motivated by overcoming the drawbacks of dynamic clustering, in [23], [24] semi-dynamic clustering was proposed, where dynamic BS clustering is applied under a limited number of candidate cooperative BSs. Although semi-dynamic clustering reduces complexity from selecting a BS cluster compared to fully dynamic ways, overheads from complex user scheduling are still required. To refine this, in [7], [8], semi-static BS clustering was proposed. The main idea of semi-static clustering is to use multiple predefined cluster patterns so that edge users have a chance to be protected from dominant out-of-cluster interference. Since used BS cluster patterns are predefined, required complexity is decreased compared to dynamic or semi-dynamic clustering.

The existing semi-static BS clustering approaches [7], [8], however, mainly focused on network topologies where the BSs’ locations are regularly placed on a grid, and in those network models, it is trivial to design BS cluster patterns. In a more practical network model where BSs’ locations are irregular, it is not clear how to form a BS cluster and map the formed BS cluster into multiple cluster patterns. This is mainly because when the network size is large, jointly considering all the possible BS clusters to make BS cluster patterns is infeasible. This leads to necessity of a general rule for forming a BS cluster and constructing BS cluster patterns in an irregular network topology with which every user in the network is guaranteed to be protected from the dominant
interference, i.e., the interference from the closest interfering BS.

C. Contributions

In this paper, we propose a cluster pattern for pair-wise multicell coordination in a cellular downlink network whose BSs’ locations are irregular. The goal of the proposed BS cluster pattern is to ensure that every user in the network communicates with their two closest BSs regardless of irregular BS deployment scenarios. To do this, we leverage the 2nd-order Voronoi region, defined as

$$\mathcal{V}_2(d_i, d_j) = \{ d \in \mathbb{R}^2 | \{d - d_i \leq d - d_k \} \cap \{d - d_j \leq d - d_k \}, \forall k \in \mathbb{Z}^+, k \neq i, j \},$$

i.e., the set of closer points to $d_i$ and $d_j$ than any other points. The core feature of the 2nd-order Voronoi region is that it characterizes an area where any point in that area has the minimum distances to $\{d_i, d_j\}$. If a BS pair located at $\{d_i, d_j\}$ forms a BS cluster and serves users in $\mathcal{V}_2(d_i, d_j)$, those users are guaranteed to communicate with their two closest BSs. To apply this strategy to the whole network, we tessellate a network plane into the 2nd-order Voronoi regions of every feasible pair of BS locations and serve users in each region by the two closest BSs. This strategy, however, can cause BS conflicts, i.e., more than two users from different regions want to communicate with the same BS simultaneously. To prevent this, BS clusters are assigned into multiple cluster patterns and each cluster pattern uses different time-frequency resource. We explain the proposed idea more rigorously in Section II.

With the proposed BS cluster patterns, we characterize the performance of the proposed BS coordination strategy in two different network models. First we consider a network where the BSs’ locations are fixed. In such a model, under the premise that multi-user coordinated beamforming (CBF) \cite{25} is applied to mitigate intra-cluster interference, we derive analytical expressions for the rate coverage probability and the ergodic spectral efficiency as a function of relevant system parameters: 1) distances between BSs and users, 2) the number of antennas per BS, 3) the number of selected users, 4) signal-to-noise ratio (SNR), and 5) the pathloss exponent. Next, we consider a random network model where a location of each BS is modeled by using a homogeneous PPP. Under the interference-limited assumption, we derive a lower bound on the ergodic spectral efficiency as a function of 1) the number of antennas per BS, 2) the number of
selected users, and 3) the pathloss exponent when the proposed cluster pattern is applied in a random network.

To demonstrate efficiency of the proposed strategy, the rate coverage probability and the ergodic spectral efficiency are compared with those obtained from other ways of coordination strategies: single cell operation, fractional frequency reuse and a random BS clustering method. One main finding is that the proposed cluster pattern substantially improves the rate outage probability and the ergodic spectral efficiency of cell-edge users compared to those of the existing approaches. The gain comes from the fact that the proposed cluster pattern ensures that users communicate with their associated BSs without the dominant interference.

There are two main points of novelty in the proposed clustering strategy: generality and uniformity. The generality means that the proposed clustering is able to be applied in any BS deployment scenario. The uniformity means that with the proposed clustering, all the users in the network are guaranteed to communicate with their two closest BSs, with the result that the edge-user performance is enhanced by protecting from the dominant interference. To the best of authors’ knowledge, there has been no work that achieved the two points simultaneously.

The remainder of the paper is organized as follows. In Section II, the system model including the proposed cluster pattern is explained. In Section III, the rate coverage probability is characterized and the ergodic spectral efficiency is obtained in Section IV. In Section V, the proposed cluster pattern applied in a random network is characterized, followed by Section VI which provides comparisons with simulations. Section VII concludes the paper.

II. SYSTEM MODEL

In this section, we introduce the system model that we exploit in this paper. We first describe the network model, the pair-wise coordination model, and the user association assumption used in this paper. Then we explain the proposed cluster pattern by using a toy example, and extend it to a general network model in the next subsection. Signal model and performance metrics are specified in the following subsection.

A. Network Model

We consider a non-random downlink cellular network where each BS has $N$ antennas. As illustrated in Fig. 1 (a), BSs are located at irregular points on a two dimensional plane. We
denote each location of a BS as \( d_i \in \mathbb{R}^2 \) for \( i \in \mathbb{Z}^+ \), so that \( \mathcal{N} = \{ d_i | i \in \mathbb{Z}^+ \} \) includes all the locations of the BS on the plane. A BS located at \( d_i \) for \( i \in \mathbb{Z}^+ \) has its own 1st-order Voronoi region \( \mathcal{V}_1(d_i) \), and each 1st-order Voronoi region tessellates the network plane into \( |\mathcal{N}| \) regions.

\[ \text{B. Pair-wise Multicell Coordination} \]

In this paper, we focus on the case where only two adjacent BSs form a cooperative cluster. The rationale behind focusing on pair-wise BS coordination is that it is highly representative for characterizing coordination gain without causing too much overhead. For example, in [18], it was shown that a BS cluster including more than two BSs decreased the ergodic spectral efficiency when considering signaling overhead. This is because including more BSs into a BS cluster requires more overhead for estimating channel coefficients of the BSs in the cluster, which decreases the ratio of the transmitted data in a packet. This degrades the spectral efficiency. For pair-wise multicell coordination, a BS located at \( d_i \) is able to coordinate with a BS located at \( d_j \) if \( \lambda(\mathcal{V}_2(d_i, d_j)) \neq 0 \), where \( \lambda(\cdot) \) is the Lebesgue measure in 2-dimensional space.

\[ \text{C. User Association} \]

It is assumed that \( 2K \) single antenna users are selected in each 2nd-order Voronoi region irrespective of the size of the region. Among them, \( K \) users are associated with the BS composing the corresponding 2nd-order Voronoi region while the other \( K \) users are associated with the other
BS. For instance, if $2K$ users are in $V_2(d_i, d_j)$, $K$ users in $V_1(d_i) \cap V_2(d_i, d_j)$ are associated with the BS located at $d_i$ and $K$ users in $V_1(d_j) \cap V_2(d_i, d_j)$ are associated with the BS located at $d_j$. This underlying assumption is justified by assuming that the number of users on a field is far larger than the number of deployed BSs, so that there are at least $2K$ users in each 2nd-order Voronoi region satisfying our assumption. More specific user scheduling algorithms are beyond the scope of this paper.

**D. The Proposed Clustering Model - Toy Example**

In this subsection, we explain the proposed clustering model with a motivational example. To this end, suppose that BS $i$ is located at $d_i$ in a finite area $\mathcal{A} \subset \mathbb{R}^2$ for $i \in \{0, 1, 2, 3\}$, as illustrated in Fig. 2. By exploiting the notion of the 2nd-order Voronoi region, it is possible to create $\binom{4}{2} = 6$ different pair-wise cooperative areas $V_2(d_i, d_j)$ for $i \neq j$ and $i, j \in \{0, 1, 2, 3\}$ such that all points within the given region $V_2(d_i, d_j)$ have the same two nearest BSs, i.e., BS $i$ and BS $j$. In the proposed strategy, users located in $V_2(d_i, d_j)$ are served by a BS pair located at $\{d_i, d_j\}$, so that they are guaranteed to communicate with the two closest BSs. Next, once each of 2nd-order Voronoi region is defined, multiple cluster patterns $\mathcal{P}_\ell$ for $\ell = \{1, ..., L\}$ are created, each of which contains different 2nd-order Voronoi regions and $\bigcup_{\ell=1}^{L} \mathcal{P}_\ell = \mathcal{A}$. Each
cluster pattern uses different time-frequency resources, so that there is no BS conflict between two different cluster patterns. In the same cluster pattern, however, BS conflicts still can occur. For instance, if \( V_2(d_0, d_1) \in \mathcal{P}_1 \) and also \( V_2(d_0, d_2) \in \mathcal{P}_1 \), BS 0 is conflicted, and it cannot serve all the associated users. To avoid this, one possible example is to create six cluster patterns such that \( \mathcal{P}_1 = \{ V_2(d_0, d_1) \} \), \( \mathcal{P}_2 = \{ V_2(d_0, d_2) \} \), \( \mathcal{P}_3 = \{ V_2(d_0, d_3) \} \), \( \mathcal{P}_4 = \{ V_2(d_1, d_2) \} \), \( \mathcal{P}_5 = \{ V_2(d_1, d_3) \} \), and \( \mathcal{P}_6 = \{ V_2(d_2, d_3) \} \). This example, however, requires too many time-frequency resources, i.e., 6. It is even larger than the case where each BS uses different time-frequency resources without BS coordination, i.e., 4. A better example is to create three BS cluster patterns such that \( \mathcal{P}_1 = \{ V_2(d_0, d_1), V_2(d_2, d_3) \} \), \( \mathcal{P}_2 = \{ V_2(d_0, d_2), V_2(d_1, d_3) \} \), and \( \mathcal{P}_3 = \{ V_2(d_0, d_3), V_2(d_1, d_2) \} \). With this example, BS conflicts are prevented and the required number of time-frequency resources is reduced to 3. To figure out how much performance gain is achieved by using this cluster pattern, assume that a user is in cell-edge region where the interference chiefly comes from the closest interfering BS. When no BS coordination is applied, the performance of the user might be severely degraded by the dominant interference which has almost same power with the desired signal. When each BS has its own separate resource, the dominant interference can be mitigated but it requires 4 time-frequency resources. By applying the proposed pattern, the dominant interference is mitigated with 3 time-frequency resources, which can lead to the performance improvement.

E. The Proposed Clustering Model - General Network

We extend the proposed clustering concept to a general network. At first, we illustrate in Fig. 3 how the proposed BS clusters are formed in a general network. Focusing on a BS located at \( d_0 \) and its adjacent BSs at \( d_j \) for \( j \in \{1, 2, ..., 5\} \), users in \( V_2(d_0, d_j) \) for \( j \in \{1, 2, ..., 5\} \) are served through a BS pair \( \{d_0, d_j\} \) by using the proposed clustering strategy. After forming BS clusters, we design cluster patterns \( \mathcal{P}_\ell \) for \( \ell \in \{1, 2, ..., L\} \) where \( \mathcal{P}_\ell = \{ V_2(d_i, d_j) | i, j \in \mathbb{Z}^+ \} \). From the intuition of the toy example, we summarize the conditions for a desirable cluster pattern.

1) \( \bigcup_{\ell=1}^{L} \mathcal{P}_\ell = \mathbb{R}^2 \): (all users in the network are covered).
2) Any two \( V_2(d_i, d_j) \in \mathcal{P}_\ell, V_2(d_v, d_w) \in \mathcal{P}_\ell \) for \( \ell \in \{1, ..., L\}, i \neq j \neq v \neq w \): (to avoid BS conflicts).
3) Small \( L \) satisfying 1) and 2): (to minimize the use of unnecessary resources).
Fig. 3. An illustration of cooperative areas characterized by the 2nd-order Voronoi region (denoted as dotted red line) focusing on the BS located at $d_0$. Each cooperative area marked by different shade patterns is assigned to a different cluster pattern to avoid BS conflicts.

In a network whose BSs’ locations are irregular, it is challenging to figure out $L$ and map each of the BS clusters into appropriate cluster patterns satisfying the above conditions. To solve this, we construct a graph $G (\mathcal{N})$ whose vertex is $d_i \in \mathcal{N}$ and an edge is made by the Delaunay triangulation defined in $\mathcal{N}$. The Delaunay triangulation for $\mathcal{N}$ is a triangulation of a plane, where each vertex of a triangle is $d_i \in \mathcal{N}$. The important condition of the Delaunay triangulation is that when drawing the circumcircle of a triangle made by the Delaunay triangulation, there should be no point of $\mathcal{N}$ inside that circumcircle. The example of $G (\mathcal{N})$ is illustrated in Fig. 1 (b), when $\mathcal{N}$ is corresponding to Fig. 1 (a). In $G (\mathcal{N})$, an edge between $d_i$ and $d_j$, i.e., $E (d_i, d_j)$, characterizes the existence of $\mathcal{V}_2 (d_i, d_j)$ since $\lambda (\mathcal{V}_2 (d_i, d_j)) \neq 0$ if and only if $d_i$ and $d_j$ are connected in $G (\mathcal{N})$ [26]. For this reason, designing cluster patterns, i.e., mapping $\mathcal{V}_2 (d_i, d_j)$ into a cluster pattern, is equivalent to mapping $E (d_i, d_j)$ into a cluster pattern. When considering the cluster patterns as “colors,” the cluster pattern design problem is equivalent to the edge-coloring in graph theory [27]. The goal of edge-coloring is to assign colors to the edges of the graph so that any two edges sharing the same vertex have different colors, while minimizing the required number of colors to cover the whole graph. By using edge-coloring for $G (\mathcal{N})$, we are able to
create cluster patterns $\mathcal{P}_t$, $\ell \in \{1, 2, ..., L\}$ that satisfies the above conditions.

Now we study the relationship between the cluster pattern design and the edge-coloring specifically. At first, we characterize $L$, which is the required number of different time-frequency resources. Thanks to the equivalence between designing cluster patterns and the edge-coloring, $L$ is equal to the minimum required number of colors for the edges. The minimum required number of colors is characterized in the following Theorem.

**Theorem 1** (Vizing’s theorem, [28]). *A simple planar graph of maximum degree $\Delta$ has chromatic index $\Delta$ or $\Delta + 1$ in general.*

*Proof:* See the reference [28].

The chromatic index is known as the required minimum number of colors and the degree is the number of edges connected to the corresponding vertex. It has been also proven that if $\Delta > 6$, the chromatic index of the given graph is $\Delta$ [29]. Therefore, if $G(\mathcal{N})$ has $\Delta \leq 6$, $L = \Delta + 1$ is enough to avoid BS conflicts, while if $\Delta > 6$, $L = \Delta$ is enough. Next, we provide two examples to show how to create the cluster patterns for a given graph. The first example, described in Fig. 4 (a), is a symmetric network where every vertex of $G(\mathcal{N})$ has the same degree, 6. This example is fitted to a case where each BS is deployed with a guard distance
for preventing that two BSs are located very close together. Specifically, each BS is randomly located within a certain circle, and a center of each circle is located at a hexagonal grid. For this particular example, it can be shown that $L = 6$. According to the description in Fig. 4 (a), the cluster pattern $P_\ell$ for $\ell \in \{1, 2, ..., 6\}$ are designed as

$$
\begin{align*}
P_1 &= \{V_2(d_0, d_1), V_2(d_2, d_{10}), V_2(d_3, d_4), \cdots\}, \\
P_2 &= \{V_2(d_0, d_2), V_2(d_3, d_{10}), V_2(d_4, d_{11}), \cdots\}, \\
P_3 &= \{V_2(d_0, d_3), V_2(d_1, d_2), V_2(d_4, d_5), \cdots\}, \\
P_4 &= \{V_2(d_0, d_4), V_2(d_2, d_3), V_2(d_5, d_6), \cdots\}, \\
P_5 &= \{V_2(d_0, d_5), V_2(d_1, d_6), V_2(d_7, d_8), \cdots\}, \\
P_6 &= \{V_2(d_0, d_6), V_2(d_4, d_8), V_2(d_3, d_{11}), \cdots\}, \\
\end{align*}
$$

(2)

and $\bigcup_{\ell=1}^6 P_\ell = \mathbb{R}^2$. By using each cluster pattern $P_\ell$ for $\ell = \{1, 2, ..., 6\}$ in a different time-frequency resource, no BS conflict occurs. Since the union of the six cluster patterns covers the whole plane $\mathbb{R}^2$, all the users in the network can be served through the proposed cluster pattern, and are guaranteed to communicate without the dominant interference.

Now we look an example Fig. 4 (b), which describes an asymmetric network, i.e., each vertex of $G(N)$ can have different numbers of degrees. This example is fitted to a case where each BS is deployed with a more random manner than the symmetric network. Since $\Delta = 7$ for this particular case, $L = 7$ is enough. Similar to the previous example, the cluster pattern $P_\ell$ for $\{1, 2, ..., 7\}$ is

$$
\begin{align*}
P_1 &= \{V_2(d_0, d_1), V_2(d_5, d_8), V_2(d_6, d_7), \cdots\}, \\
P_2 &= \{V_2(d_2, d_3), V_2(d_0, d_4), V_2(d_6, d_8), \cdots\}, \\
P_3 &= \{V_2(d_0, d_5), \cdots\}, \\
P_4 &= \{V_2(d_4, d_5), V_2(d_0, d_8), \cdots\}, \\
P_5 &= \{V_2(d_7, d_8), \cdots\}, \\
P_6 &= \{V_2(d_2, d_4), \cdots\}, \\
P_7 &= \{V_2(d_3, d_4), \cdots\},
\end{align*}
$$

(3)
and $\bigcup_{\ell=1}^{\ell} \mathcal{P}_\ell = \mathbb{R}^2$. No BS conflict occurs as long as a different time-frequency resource is assigned to a different cluster pattern.

Observe that although the general edge-coloring problem is NP-complete, complexity is not important since the proposed cluster patterns are for semi-static BS coordination. Unless the geometry of the BSs changes, the cluster patterns are preserved. Conventionally, the BS geometry would not be changed on the order of months or years.

### F. Signal Model

We focus on describing the signal model for a user associated to the BS located at $\mathbf{d}_0$. We denote this user as the tagged user. We also assume that the tagged user is located on the origin. This assumption is easily generalized to arbitrary user location by shifting each BS’s location to $\mathbf{d}_i - \mathbf{u}$ for $i \in \mathbb{Z}^+$. Let $\mathcal{V}_2(\mathbf{d}_0, \mathbf{d}_j)$, a BS pair $\{\mathbf{d}_0, \mathbf{d}_j\}$ forms a cluster according to the proposed cluster pattern. Let $\mathcal{V}_2(\mathbf{d}_0, \mathbf{d}_j) \in \mathcal{P}_\ell$ and $C_\ell = \{\mathbf{d}_0, \mathbf{d}_j\}$. In each BS located at $\mathbf{d}_i$, an information symbol vector $\mathbf{s}^{\ell}_i \in \mathbb{C}^K$, where $\mathbf{s}^{\ell}_i = [s^{\ell}_i, 1, \ldots, s^{\ell}_i, K]^T$ for $i \in \mathbb{Z}^+$ is encoded to be transmitted each user.

The average power of this transmit symbol vector satisfies $\mathbb{E}[\|s^{\ell}_i\|^2] \leq P$. When transmitting the symbol vector, a linear beamforming matrix $\mathbf{V}^{\ell}_i = [\mathbf{v}^{\ell}_i, 1, \ldots, \mathbf{v}^{\ell}_i, K]$, whose $\mathbf{v}^{\ell}_i, k \in \mathbb{CN}$ and $\|\mathbf{v}^{\ell}_i, k\| = 1$ for $k = 1, \ldots, K$ is used. The observation from the tagged user is given by

$$y^{\ell} = \left(\|\mathbf{d}_0\|^{-\beta/2} (\mathbf{h}^{\ell}_0)^T \mathbf{V}^{\ell}_0 \mathbf{s}^{\ell}_0 + \|\mathbf{d}_j\|^{-\beta/2} (\mathbf{h}^{\ell}_j)^T \mathbf{V}^{\ell}_j \mathbf{s}^{\ell}_j \right)^{\text{desired signal}} + \sum_{\mathbf{d}_v \in \mathcal{N}_\ell \setminus C_\ell} \|\mathbf{d}_v\|^{-\beta/2} (\mathbf{h}^{\ell}_v)^T \mathbf{V}^{\ell}_v \mathbf{s}^{\ell}_v + \mathbf{n}^{\ell},$$

where $\mathbf{h}^{\ell}_i \in \mathbb{CN}$ is the channel coefficient vector from the BS at $\mathbf{d}_i$ to the tagged user. To facilitate analysis, the channel coefficients are assumed to have independent and identically distributed (IID) complex Gaussian entries with zero mean and unit variance, i.e., $\mathcal{CN}(0, 1)$. The distant dependent pathloss is with a reference at 1m and has an exponent of $\beta$ and the additive Gaussian noise is $\mathbf{n}^{\ell}$ that follows $\mathcal{CN}(0, \sigma^2)$. $\mathcal{N}_\ell \subseteq \mathcal{N}$ is a set including BSs’ locations whose the 2nd-order Voronoi region is included in the cluster pattern $\mathcal{P}_\ell$, i.e., $\mathcal{N}_\ell = \{\mathbf{d}_i | \mathcal{V}_2(\mathbf{d}_i, \mathbf{d}_j) \in \mathcal{P}_\ell, \forall i, j\}$. The received signal can be separated into three parts: the desired signal from the associated BS at $\mathbf{d}_0$, the intra-cluster interference from a BS set $C_\ell \setminus \mathbf{d}_0$, and the unmanageable out-of-cluster interference, respectively.
Before transmission of an information symbol, pilot symbols are exploited to learn the downlink channel coefficient vector $h^0_\ell$ and $h^j_\ell$ if $d_j \in \mathcal{C}_\ell$, i.e., the BS is included in the corresponding BS cluster. Then, the user sends the obtained CSI back to the corresponding BSs via an error-free feedback link. CSI at the transmitter (CSIT) within the same BS cluster is allowed. To mitigate intra-cluster interference, CBF is employed by using obtained CSIT. Assuming that our tagged user’s index is $k$, the beamforming matrix $V^\ell_0$ and $V^\ell_j$ for $d_j \in \mathcal{C}_\ell$ are designed to satisfy the following condition.

$$\begin{align*}
\text{maximize} & : \left| (h^\ell_0)^T v^\ell_{0,k} \right|^2 \\
\text{subject to} & : \left| (h^\ell_0)^T v^\ell_{0,k'} \right|^2 = 0 \text{ for } k' \neq k \\
& \left| (h^j_\ell)^T v^\ell_{j,k''} \right|^2 = 0 \text{ for } d_j \in \mathcal{C}_\ell \text{ and } 1 \leq k'' \leq K,
\end{align*}$$

where the constraints of multi-user CBF are for nullifying the inter-user interference and for removing intra-cluster interference, respectively. Under the conventional zero forcing (ZF) conditions, it is always possible to find such $V^\ell_0$ and $V^\ell_j$ if $2K \leq N$ with high enough probability if each channel coefficient is IID.

**G. Performance Metrics**

If perfect CSIT within the same BS cluster is allowed, then the inter-user interference and intra-cluster interference terms in (4) are zero. Assuming that $\tilde{h}^\ell_0$ indicates a modified channel coefficient after multiplying with a beamforming matrix $V^\ell_0$ and $s^\ell_{0,0}$ is the information symbol for the tagged user, we have the following modified received signal

$$\tilde{y}^\ell = \|d_0\|^{-\beta/2} \tilde{h}^\ell_0 s^\ell_{0,0} + \sum_{d_v \in \mathcal{N}_\ell \setminus \mathcal{C}_\ell} \|d_v\|^{-\beta/2} (h^\ell_v)^T V^\ell_v s^\ell_v + n^\ell. \quad (7)$$

The signal-to-interference and noise (SINR) of the information symbol $s^\ell_{0,0}$ is

$$\text{SINR}_{\ell} = \frac{\|d_0\|^{-\beta} \left| \tilde{h}^\ell_0 \right|^2}{\sum_{d_v \in \mathcal{N}_\ell \setminus \mathcal{C}_\ell} \|d_v\|^{-\beta} \left| (h^\ell_v)^T V^\ell_v \right|^2 + \sigma^2/P} = \frac{\left| \tilde{h}^\ell_0 \right|^2}{I_\ell + \|d_0\|^{-\beta}/\text{SNR}}. \quad (8)$$
where \( I_\ell = \sum_{d_v \in N_\ell \setminus C_\ell} (\|d_v\| / \|d_0\|)^{-\beta} \left| (h_0^\ell)^T V_\ell^v \right|^2 \), and \( P/\sigma^2 = \text{SNR} \). Given the system assumptions, the rate coverage probability is given as

\[
P^\ell (\text{SNR}, \|d_0\|, D_\ell, N, K, \beta, \gamma) = \mathbb{P} \left[ \log_2 (1 + \text{SINR}_\ell) > \gamma \right],
\]

where \( D_\ell = \{ d_v | d_v \in N_\ell \setminus C_\ell \} \).

where \( \gamma \) is the rate threshold. The ergodic spectral efficiency is then

\[
R^\ell (\text{SNR}, \|d_0\|, D_\ell, N, K, L, \beta) = \mathbb{E} \left[ \frac{1}{L} \log_2 (1 + \text{SINR}_\ell) \right],
\]

where \( D_\ell = \{ d_v | d_v \in N_\ell \setminus C_\ell \} \).

The pre-log term \( 1/L \) is used because \( 1/L \) time-frequency resources are used for serving the tagged user.

### III. Rate Coverage Analysis

In this section, we derive the rate coverage probability in a closed form. To this end, we first introduce the following lemma. Lemma 1 provides distributions of the desired channel gain and the interference power.

**Lemma 1.** The desired channel gain \( \left| \tilde{h}_0^\ell \right|^2 \) follows Chi-squared distribution with \( 2 (N - 2K + 1) \) degrees of freedom. The out-of-cluster interference \( I_\ell \) follows weighted sum of Chi-squared distribution with \( 2K \) degrees of freedom.

**Proof:** See [18] and reference therein.

By using this Lemma, we derive the rate coverage. The following Theorem is the main result of this section.

**Theorem 2.** In the interference limited regime, when the BSs form a BS coordination set according to the proposed cluster pattern, the rate coverage probability of the tagged user is

\[
P^\ell (\text{SNR} \to \infty, \|d_0\|, D_\ell, N, K, \beta, \gamma) = \sum_{m=0}^{N-2K} \frac{(2\gamma - 1)^m}{m!} (-1)^m \frac{\partial^m}{\partial s^m} \prod_{d_v \in D_\ell} \left( \frac{1}{1 + s (\|d_v\| / \|d_0\|)^{-\beta}} \right)^K,
\]

(11)
where \( s = 2^\gamma - 1 \) and \( \gamma \) is the rate threshold.

**Proof:** As defined in (8), the rate coverage probability is rewritten as

\[
\mathbb{P} \left[ \log_2 \left( 1 + \text{SINR}_\ell \right) > \gamma \right] = \mathbb{P} \left[ \frac{\hat{h}_0^\ell}{I_\ell} > (2^\gamma - 1) \right]
\]

\[
= \mathbb{P} \left[ \hat{h}_0^\ell > (2^\gamma - 1) I_\ell \right]
\]

\[
\overset{(b)}{=} \mathbb{E} _{I_\ell} \left[ \sum_{m=0}^{N-2K} \frac{(2^\gamma - 1)^m}{m!} (I_\ell)^m e^{-(2^\gamma - 1)I_\ell} \right], \tag{12}
\]

where (a) comes from that \( \|d\|^\beta / \text{SNR} \to 0 \) when \( \text{SNR} \to \infty \), and (b) follows Lemma 1 and the complement cumulative distribution function of a Chi-squared random variable with degrees of freedom \( 2 \left( N - 2K + 1 \right) \). Using the derivation of the Laplace transform, i.e.,

\[
\mathbb{E} \left[ X^m e^{-sX} \right] = \left( -1 \right)^m \frac{d^m}{ds^m} L_X(s) \bigg|_{s=2^\gamma - 1}, \tag{13}
\]

where \( L_X(s) \) is the Laplace transform of \( I_\ell \), defined as

\[
L_{I_\ell}(s) = \prod_{d_v \in \mathcal{D}_\ell} \left( \frac{1}{1 + s \left( \|d_v\| / \|d_0\| \right)^{-\beta}} \right)^K, \tag{14}
\]

which completes the proof. \( \blacksquare \)

Computing the rate coverage probability (8) is not easy since computing the derivative of the product of \( \left( 1/1 + s \left( \|d_v\| / \|d_0\| \right)^{-\beta} \right)^K \) results in many terms. To provide a simpler form of the rate coverage probability, the following Corollary is used to approximate the rate coverage probability.

**Corollary 1.** The rate coverage probability (8) can be approximated as in the following form.

\[
P^\ell \left( \text{SNR} \to \infty, \|d_0\|, \mathcal{D}_\ell, N, K, \beta, \gamma \right) \approx \sum_{m=0}^{N-2K} \frac{(2^\gamma - 1)^m}{m!} (-1)^m \frac{\partial^m}{\partial s^m} \left( e^{-\sum_{d_v \in \mathcal{D}_\ell \setminus d_{\text{min}} s \left( \|d_v\| / \|d_0\| \right)^{-\beta}} \right)^K, \tag{15}
\]

where \( s = 2^\gamma - 1 \) and \( d_{\text{min}} = \arg \min_{d_v \in \mathcal{D}_\ell} \|d_v\| \).
Proof: Using the Taylor expansion of the exponential function $e^x = 1 + x + x^2/2! + \cdots$, when $x$ is small enough, $e^x \approx 1 + x$, therefore $e^{-x} \approx 1/(1 + x)$. Then we have

$$\prod_{d_v \in D_\ell} \left( \frac{1}{1 + s ( \|d_v\| / \|d_0\|)^{-\gamma}} \right)^K = \left( \frac{1}{1 + s ( \|d_{\min}\| / \|d_0\|)^{-\gamma}} \right)^K \prod_{d_v \in D_\ell \setminus d_{\min}} \left( \frac{1}{1 + s ( \|d_v\| / \|d_0\|)^{-\gamma}} \right)^K \approx \left( \frac{1}{1 + s ( \|d_{\min}\| / \|d_0\|)^{-\gamma}} \right)^K \left( e^{-\sum_{d_v \in D_\ell \setminus d_{\min}} s ( \|d_v\| / \|d_{\min}\|)^{-\beta}} \right)^K,$$

which completes the proof.

Using the Taylor expansion, we approximate a product of $\left(1/1 + s ( \|d_v\| / \|d_0\|)^{-\beta}\right)^K$ in an exponential form, so that we are able to calculate the derivative of (15) more simply than that of (8). Now we verify that the approximated form of the rate coverage probability (15) is reasonable by comparing to the Monte-Carlo simulation results. For the simulations, we assume a symmetric network model, which is equal to Fig. 4 (a). As illustrated in Fig. 5 the analytical approximation tightly matches the simulated rate distribution over entire rate threshold region and for different numbers of BS antennas and users.
IV. ERGODIC SPECTRAL EFFICIENCY ANALYSIS

In this section, we derive an exact expression of the ergodic spectral efficiency in an integral form and provide a lower bound on the ergodic spectral efficiency in a closed form.

A. Exact Characterization

Our exact expression is derived by using Lemma 2, which yields a general expression of the ergodic spectral efficiency in terms of moment generating functions.

**Lemma 2.** Let $x_1, \ldots, x_N, y_1, \ldots, y_M$ be arbitrary non-negative random variables. Then

$$
\mathbb{E} \left[ \ln \left( 1 + \frac{\sum_{n=1}^{N} x_n}{\sum_{m=1}^{M} y_m + 1} \right) \right] = \int_0^\infty \frac{M_y(z) - M_{x,y}(z)}{z} \exp (-z) \, dz,
$$

(17)

where $M_y(z) = \mathbb{E}\left[ e^{-z \sum_{m=1}^{M} y_m} \right]$ and $M_{x,y}(z) = \mathbb{E}\left[ e^{-z \left( \sum_{n=1}^{N} x_n + \sum_{m=1}^{M} y_m \right)} \right]$.

**Proof:** See Lemma 1 in reference [30].

Leveraging Lemma 2, an exact expression of the ergodic spectral efficiency is obtained as in the following Theorem.

**Theorem 3.** When the BSs form a BS coordination set according to the proposed cluster pattern, the ergodic spectral efficiency of the tagged user is

$$
R^t \left( \text{SNR}, \|d_0\|, \mathcal{D}_t, N, K, L, \beta \right) = \frac{\log_2(e)}{4L} \int_0^\infty \exp \left( -z \frac{\|d_0\|^2}{\text{SNR}} \right) \prod_{d_v \in \mathcal{D}_t} \left( \frac{1}{1 + z \left( \|d_v\| / \|d_0\| \right)^{-\beta}} \right)^K \left( 1 - \left( \frac{1}{1 + z} \right)^{N-2K+1} \right) \, dz.
$$

(18)

**Proof:** We start by characterizing the moment generating function of the desired channel gain. Since the desired channel gain is a Chi-squared random variable with $2(N-2K+1)$ degrees of freedom from Lemma 1, its moment generating function is given by

$$
\mathcal{M}_S(z) = \mathbb{E}\left[ e^{-z|\tilde{h}_0|^2} \right]
= \left( \frac{1}{1 + z} \right)^{N-2K+1}.
$$

(19)
Similarly, the moment generating function of the interference is
\[ M_{I_\ell}(z) = \prod_{d_v \in D_\ell} \left( \frac{1}{1 + z \left( \|d_v\| / \|d_0\| \right)^{-\beta}} \right)^K. \]  
(20)

By applying Lemma 2, the Theorem is completed.

Given distances from the tagged user to each BS, the ergodic spectral efficiency of the tagged user is obtained by numerically calculating (18).

B. A Lower Bound Characterization

We also derive a lower bound on the ergodic spectral efficiency with a closed form which is useful in obtaining insight on how system parameters change the ergodic spectral efficiency. To derive a lower bound, the following Lemma is claimed.

**Lemma 3.** For any non-negative independent random variables \( S \) and \( I \),
\[ \mathbb{E} \left[ \log_2 \left( 1 + \frac{S}{I+1} \right) \right] \geq \log_2 \left( 1 + \frac{e^{\mathbb{E} \left[ \ln S \right]}}{\mathbb{E} \left[ I+1 \right]} \right). \]  
(21)

**Proof:**
\[ \mathbb{E} \left[ \log_2 \left( 1 + \frac{S}{I+1} \right) \right] \geq \mathbb{E}_S \left[ \mathbb{E} \left[ \log_2 \left( 1 + \frac{S}{I+1} \right) \left| S \right| \right] \right] \]
\[ \overset{(a)}{=} \mathbb{E} \left[ \log_2 \left( 1 + \frac{e^{\mathbb{E} \left[ \ln S \right]}}{\mathbb{E} \left[ I+1 \right]} \right) \right] \]
\[ \overset{(b)}{=} \log_2 \left( 1 + \frac{e^{\mathbb{E} \left[ \ln S \right]}}{\mathbb{E} \left[ I+1 \right]} \right) \]  
(22)

where (a) follows Jensen’s inequality and \( \log_2 \left( 1 + \frac{1}{x} \right) \) is a convex function of non-negative variable \( x \), and (b) also follows Jensen’s inequality and \( \log_2 (1 + e^y) \) is a convex function of non-negative variable \( y \).

Applying Lemma 3, a lower bound of the ergodic spectral efficiency is obtained in the following Theorem.

**Theorem 4.** When the BSs form a BS coordination set according to the proposed cluster pattern, the ergodic spectral efficiency of the tagged user is lower bounded as
\[ R_\ell^e (\text{SNR}, \|d_0\|, D_\ell, N, K, L, \beta) \geq \frac{1}{L} \log_2 \left( 1 + \frac{\exp \left( \psi (N - 2K + 1) \right)}{K \sum_{d_v \in D_\ell} \left( \|d_v\| / \|d_0\| \right)^{-\beta} + \|d_0\|^\beta / \text{SNR}} \right), \]  
(23)
where $\psi(z)$ is the digamma function defined as $\Gamma'(z)/\Gamma(z)$, where $\Gamma(z)$ is the gamma function defined as

$$\Gamma(z) = \int_0^\infty x^{z-1}e^{-x}dx.$$  \hfill (24)

**Proof:** First, we characterize $E[\ln |\tilde{h}_0|^2]$. Since $|\tilde{h}_0|^2 \sim \chi^2(2(N-2K+1))$,

$$E[\ln |\tilde{h}_0|^2] = \psi(N-2K+1),$$  \hfill (25)

where $\psi(z)$ is the digamma function defined as $\Gamma'(z)/\Gamma(z)$. Next, we compute $E[I_\ell]$. Since the interference is a sum of independent Chi-squared random variables with different weights corresponding to the ratio of the distances, we have

$$E[I_\ell] = K \sum_{d_v \in D_\ell} \left(\frac{\|d_v\|}{\|d_0\|}\right)^{-\beta}.$$  \hfill (26)

By applying Lemma (3), we complete the proof. Verification of the obtained lower bound is provided in Section VI.

Theorem 4 gives intuition about how the proposed clustering strategy affects the ergodic spectral efficiency. For example, from the approximation

$$\log_2(1+x) \approx x \log_2 e$$  \hfill (27)

when $x \approx 0$, the following approximation is allowed in a low SINR regime.

$$\frac{1}{L} \log_2 \left(1 + \frac{\exp(\psi(N-2K+1))}{K \sum_{d_v \in D_\ell} (\|d_v\|/\|d_0\|)^{-\beta} + \|d_0\|^\beta/\text{SNR}} \right) \approx \frac{\log_2 e}{L} \cdot \frac{\exp(\psi(N-2K+1))}{K \sum_{d_v \in D_\ell} (\|d_v\|/\|d_0\|)^{-\beta} + \|d_0\|^\beta/\text{SNR}}.$$  \hfill (28)

By using the proposed cluster pattern, one dominant denominator of (28) is reduced, which leads to significant increase of the ergodic spectral efficiency.

V. ERGODIC SPECTRAL EFFICIENCY IN RANDOM NETWORKS

So far, we characterize the performance when the BSs are coordinated with the proposed cluster patterns in a fixed network model. In this section, we analyze the performance when applying the proposed cluster pattern in a random network. We consider a downlink cellular network where BSs equipped with $N$ antennas are distributed according to a homogeneous PPP,
\( \Phi = \{ \mathbf{d}_i | i \in \mathbb{N} \} \) with density \( \lambda \) on the plane \( \mathbb{R}^2 \). The \( k \)-th closest BS from the origin is denoted as \( \mathbf{d}_k \). Single antenna users are also distributed as a homogeneous PPP, \( \Phi_U = \{ \mathbf{u}_i | i \in \mathbb{N} \} \) with density \( \lambda_U \), which is independent with \( \Phi \). We assume that \( \lambda_U \gg \lambda \) so that there exist at least \( K \) users in each region with high enough probability. Under this assumption, \( K \) users are selected in each In this network model, the proposed BS clustering strategy is applied in the same way with the case of a fixed network model, i.e., construct a graph \( G(\Phi) \) by the Delaunay triangulation defined in \( \Phi \) and solve the corresponding edge-coloring problem. Under the assumption that the proposed clustering strategy, i.e., each BS cluster is formed by using the 2nd-order Voronoi region and time-frequency resources are allocated by solving the corresponding edge-coloring, is employed, we derive a lower bound on the ergodic spectral efficiency.

One should note that the randomness of the network model in this section is not a core part of the proposed strategy. Rather, it is a tool for analyzing the ergodic spectral efficiency performance.

### A. Performance Metric

Without loss of generality, we focus on the typical user located at the origin, and assume that it is associated with the BS located at \( \mathbf{d}_1 \). Under the premise that the typical user is in \( \mathcal{V}_2(\mathbf{d}_1, \mathbf{d}_2) \in \mathcal{P}_\ell \), the typical user is served by a BS coordination pair \( \{ \mathbf{d}_1, \mathbf{d}_2 \} \). Similarly to the fixed network case, we assume that \( K \) users in \( \mathcal{V}_2(\mathbf{d}_1, \mathbf{d}_2) \) are selected to be associated with the BS located at \( \mathbf{d}_1 \), resulting in that there are \( 2K \) users in each 2nd-order Voronoi region. After forming the cluster set, the typical user is protected from interference by applying CBF as in the fixed network model case. The SINR of the typical user is given by

\[
\text{SINR}_\ell = \frac{\|\mathbf{d}_1\|^{-\beta} \left| \tilde{h}_1^\ell \right|^2}{\sum_{\mathbf{d}_i \in \Phi \setminus B(0, \|\mathbf{d}_2\|)} \|\mathbf{d}_i\|^{-\beta} \left| \left( \mathbf{h}_i^\ell \right)^T \mathbf{V}_i^\ell \right|^2 + 1 / \text{SNR}}, \tag{29}
\]

where \( \tilde{h}_1^\ell \) is the modified channel coefficient from the BS at \( \mathbf{d}_1 \) to the typical user after applying CBF, \( \mathbf{h}_i^\ell \) is the channel coefficient vector from the BS at \( \mathbf{d}_i \) to the typical user, and \( \mathbf{V}_i^\ell \) is a beamforming matrix of the BS at \( \mathbf{d}_i \), respectively. The entries of the channel coefficient vector follows \( \mathcal{CN}(0, 1) \). From (29), the ergodic spectral efficiency is defined as

\[
R^\ell(\text{SNR}, N, K, \beta) = \mathbb{E} \left[ \frac{1}{L} \log_2 \left( 1 + \text{SINR}_\ell \right) \right], \tag{30}
\]
where the pre-log term $1/L$ is used because $1/L$ time-frequency resources are exploited to serve the typical user.

### B. Lower Bound on the Ergodic Spectral Efficiency

Now we derive a lower bound on the ergodic spectral efficiency. To do this, we need the following Lemma. Lemma 4 provides probability density function (PDF) of distances of the closest BS and $k$-th closest BS to the origin.

**Lemma 4.** The joint probability density function (PDF) of $\|d_1\|$ and $\|d_k\|$ is

$$f_{\|d_1\|,\|d_k\|}(r_1, r_k) = \begin{cases} \frac{4(\lambda \pi)^k}{(k-2)!} r_1 r_k (r_k^2 - r_1^2)^{k-2} e^{-\lambda \pi r_k^2} & \text{if } r_1 \leq r_k \\ 0 & \text{otherwise.} \end{cases}$$

(31)

**Proof:** See Appendix C in [18].

Leveraging Lemma 4, Theorem 5 is claimed. For analytical tractability, we assume the interference limited regime.

**Theorem 5.** When BSs’ locations are distributed by a homogeneous PPP and the BSs form a BS coordination set according to the proposed cluster pattern the ergodic spectral efficiency of the tagged user is lower bounded as

$$R^T(N, K, \beta) \geq \mathbb{E} \left[ \frac{1}{L} \log_2 \left( 1 + \frac{(\beta^2 - 4)}{8K} \exp \left( \psi(N - 2K + 1) \right) \right) \right],$$

(32)

where $\psi(z)$ is the digamma function, defined as $\Gamma'(z)/\Gamma(z)$, where $\Gamma(z)$ is the gamma function defined as

$$\Gamma(z) = \int_0^\infty x^{z-1} e^{-x} dx.$$ 

(33)

**Proof:** The ergodic spectral efficiency (30) is written by

$$\mathbb{E} \left[ \frac{1}{L} \log_2 \left( 1 + \text{SIR}_t \right) \right] = \mathbb{E} \left[ \frac{1}{L} \right] \mathbb{E} \left[ \log_2 \left( 1 + \text{SIR}_t \right) \right]$$

$$\geq \mathbb{E} \left[ \frac{1}{L} \right] \log_2 \left( 1 + \frac{e^{\mathbb{E}[|h_i^T|^2]}}{\mathbb{E}[\|d_1\|^\beta \sum_{d_i \in \Phi \setminus B(0, k d_2)} \|d_i\|^{-\beta} (h_i^T V_i)^2]} \right),$$

(34)
where (a) follows Lemma 3. Now, we obtain the expectation in the nominator and denominator of (34). Since the interference is mitigated by CBF, as in the case of the deterministic square grid network model, $\left| \tilde{h}_\ell \right|^2$ follows $\chi^2 (2 (N - 2K + 1))$, which leads to

$$E \left[ \ln \left( \left| \tilde{h}_\ell \right|^2 \right) \right] = \psi (N - 2K + 1),$$

(35)

where $\psi (\cdot)$ is the digamma function. Next, we calculate the expectation of the interference part. We write the expectation of the denominator of the SIR as

$$E \left[ \left\| d_1 \right\|^\beta \sum_{d_i \in \Phi \setminus B(0,\left\| d_2 \right\|)} \left\| d_i \right\|^{-\beta} \left| (h_i^T V_i^\ell) \right|^2 \right] \overset{(a)}{=} K E \left[ \left\| d_1 \right\|^\beta \sum_{d_i \in \Phi \setminus B(0,\left\| d_2 \right\|)} \left\| d_i \right\|^{-\beta} \right]$$

$$= K E_{r_1,r_2} \left[ E_{\Phi \setminus B(0,r_2)} \left[ r_1^\beta \sum_{d_i \in \Phi \setminus B(0,r_2)} \left\| d_i \right\|^{-\beta} \left\| d_1 \right\| = r_1, \left\| d_2 \right\| = r_2 \right] \right]$$

$$\overset{(b)}{=} K E_{r_1,r_2} \left[ r_1^\beta 2\pi \lambda \int_{r_2}^{\infty} r^1 - \beta d r \right]$$

$$\overset{(c)}{=} \frac{2K \pi \lambda}{\beta - 2} \int_{r_2=0}^{\infty} \int_{r_1=0}^{r_2} 4 \left( \lambda \pi \right)^2 e^{-\lambda \pi r_2^2} r_1^\beta r_2^{3-\beta} d r_1 d r_2$$

$$= \frac{8K}{\beta^2 - 4}$$

(36)

where (a) follows $\left| (h_i^T V_i^\ell) \right|^2 \sim \chi^2 (2K)$ and it is independent to $\Phi$, (b) follows the Campbell’s theorem, and (c) follows that the joint PDF in Lemma. The joint PDF of $\left\| d_1 \right\|$ and $\left\| d_2 \right\|$ can be directly obtained from Lemma 4. Plugging (35) and (36) into (34) the proof is completed. 

Since it is challenging to find $E [1/L]$ analytically in a homogeneous PPP, it is possible to use numerically found $E [1/L]$. For instance, we find that $E [1/L] \approx 0.09$ by Monte-Carlo simulations when $\lambda = 3 \times 10^{-7}$.

### C. Edge-cutting Algorithm

In an asymmetric network, it is possible that too many time-frequency resources are required for the proposed clustering. For instance, we numerically obtain that the required number of time-frequency resources is $\approx 11$ when considering a network where each BS’s location is distributed according to a homogeneous PPP with $\lambda = 3 \times 10^{-7}$. Therefore, the ergodic spectral efficiency is decreased by $\approx 0.09$. Even though the proposed clustering improves the rate coverage probability performance, the ergodic spectral efficiency can be degraded due to the excessive time-frequency
use. To relieve this, a simple algorithm called edge-cutting algorithm is proposed. The key idea of the edge-cutting algorithm is cutting edges of $G(\mathcal{N})$ so that the maximum number of degrees of $G(\mathcal{N})$ is forced to be $\Delta_{EC} < \Delta$. Specifically, in $G(\mathcal{N})$, if a vertex $d_i \in \mathcal{N}$ has the number of degree $D(d_i) > \Delta_{EC}$, $D(d_i) - \Delta_{EC}$ edges are randomly selected and cut. Repeating this, we can construct the cut-graph $G^\text{cut}(\mathcal{N})$ whose $\Delta = \Delta_{EC}$, and this decreases the required number of time-frequency resources to $\Delta_{EC}$ (See Theorem 1). Note that the constructed cut-graph $G^\text{cut}(\mathcal{N})$ is a subgraph of $G(\mathcal{N})$. Assuming that the edge $E(d_i, d_j)$ is cut in $G^\text{cut}(\mathcal{N})$, users in $\mathcal{V}_2(d_i, d_j)$ are not served while exploiting $G^\text{cut}(\mathcal{N})$. By alternatively applying various cut-graphs $G^\text{cut,1}(\mathcal{N})$, $G^\text{cut,2}(\mathcal{N})$, $\cdots$, all the regions are fairly covered. By exploiting this algorithm, the overall network performance can be improved by saving time-frequency resources. In the later section, we will compare the performance of the proposed clustering with the edge-cutting algorithm and find that the proposed clustering with appropriate $\Delta_{EC}$ improves the ergodic spectral efficiency. The step-by-step description of the edge-cutting algorithm is described in the following using a MATLAB style notation.

**Algorithm 1** Edge-cutting algorithm

**Input:** $G(\mathcal{N})$, $N_{\text{bs}} = |\mathcal{N}|$, $\Delta_{EC}$.

**Initialize:** $E \in \mathbb{N}^{N_{\text{bs}} \times N_{\text{bs}}}$, $E(i, j) = 1$ if and only if $d_i \in \mathcal{N}$ and $d_j \in \mathcal{N}$ is connected by an edge in $G(\mathcal{N})$, otherwise $E(i, j) = 0$.

**for** $i = 1 : N_{\text{bs}}$ **do**

**if** $\text{sum}(E(:, i)) > \Delta_{EC}$ **then**

 Remove$_{\text{list}} = \text{find}(E(:, i) == 1)$

 $E(\text{Remove$_{\text{list}}$} (1 : \text{sum}(E(:, i) - \Delta_{EC})), i) = 0$

**end if**

**end for**

$E_{\text{cut}} = E$

In $G_{\text{cut}}(\mathcal{N})$, $d_i$ and $d_j$ are connected if and only if $E_{\text{cut}}(i, j) = 1$.

**Output:** $G_{\text{cut}}(\mathcal{N})$
VI. PERFORMANCE COMPARISONS

A. Baseline Methods

For the performance comparison, we consider the two following baseline methods: single cell operation, fractional frequency reuse, and random BS clustering. We briefly explain as follows:

1) Single Cell Operation: In this strategy, no interference management is employed. Each BS communicates with selected users through a ZF beamformer. It is considered as a reference strategy for performance comparisons.

2) Fractional Frequency Reuse: In this strategy, interference is managed by using different orthogonal frequency sub-bands in each cell. With frequency reuse factor $\delta$, the total frequency resources are equally divided into $\delta$ sub-bands which are orthogonal to each other, and each cell region is covered by one of the sub-bands. Since allocating the sub-bands with a fixed pattern is difficult due to irregular BSs’ locations, we assume that each BS randomly selects the sub-bands. This strategy reflects an interference management technique without BS coordination.

3) Random Clustering: In this strategy, each BS randomly selects one of the adjacent BSs and forms a BS cluster with the selected BS. In each BS cluster, the intra-cluster interference is nullified by using CBF. Since a BS is formed in a random manner, there is a non-zero possibility that a user is exposed to the dominant interference.

Fig. 6 illustrates the baseline methods applied in a square grid network where each BS is located on a square grid. In single cell operation described in Fig. 6 (a), the user associated with BS 1 receives interference from all the BSs since there is no interference management technique. In fractional frequency reuse described in Fig. 6 (b), the interference only comes from BS 4, since BS 2 and BS 3 (marked by darker shade) are allocated by different sub-band from the sub-band of BS 1. In random clustering describe in Fig. 6 (c), BS 1 forms a BS cluster with BS 2 (marked by green dotted line), so that the interference from BS 2 is mitigated. In particular, in random clustering, there is a non-zero possibility that the performance of the user is degraded by the dominant interference. This is mainly because a BS cluster is formed by a random manner. For instance, in Fig. 6 (c), the dominant interference comes from BS 3 since BS 3 is the closest interfering BS for the user, but BS 1 forms a BS cluster with BS 2, causing the dominant interference still degrades the user performance.
Fig. 6. Baseline methods including (a) single cell operation, (b) fractional frequency reuse, and (c) random clustering. A blue line indicates desirable signal from the associated BS and a red dotted line indicates interference. In (b), BS 2 and BS 3 marked by grey shade use a different sub-band by applying FFR, resulting in that interference comes from only BS 4 that uses the same sub-band with the associated BS. In (c), BS 1 forms a BS cluster (denoted as a green dotted line) with BS 2 so that intra-cluster interference is mitigated, resulting in that interference only comes from BS 3 and BS 4.

B. Simulation Results

For the simulations, we assume two fixed network models. The first network model is a symmetric network, where every $d_i$ for $d_i \in \mathcal{N}$ has the same degree 6. For constructing a symmetric network, each BS is randomly located in a circle whose radius is 600m, and a center of each circle is located at hexagonal grid. A realization of this network is shown in Fig. 4 (a). The second network model is an asymmetric network, where each BS is dropped according to a homogeneous PPP with $\lambda = 3 \times 10^{-7}$ therefore each $d_i$ can have different degrees. A realization of this network is shown in Fig. 4 (b). For the proposed clustering strategy, it is assumed that the edge-coloring in each network is perfectly solved, so that every user in the network communicates with two closest BSs without BS conflicts. In our example, the symmetric network has $L = 6$, while the asymmetric network has $L = 11$ before applying the edge-cutting algorithm. The pathloss exponent is set to be $\beta = 4$, which is typical for a terrestrial wireless environment. We uniformly drop the tagged user on the plane, and shift $d_i - u$ for $d_i \in \mathcal{N}$ where $u$ is the location of the tagged user. Then we focus on the tagged user located at the origin after shifting. The locations of the closest BS and the second closest BSs are denoted by $d_0$ and $d_1$, respectively. The number of BS antennas and selected users are set to be $N = 3$ and $K = 1$. The frequency reuse factor $\delta$ for the fractional frequency reuse strategy is set to be 3, which is a common value for a conventional cellular network.

First, we compare the rate coverage probability. Fig. 8 illustrates the rate coverage probability
Fig. 7. The realization of the considered network models. Both cases are generated by repulsive point process. Specifically, in (a), i.e., a symmetric network, each BS is randomly located in a circle whose radius is 600m, and a center is located at hexagonal grid. In (b), i.e., an asymmetric network model, each BS is uniformly dropped according to a homogeneous PPP with $\lambda = 3 \times 10^{-7}$.

performance under the assumption of interference limited environment. Fig. 8 (a) shows results for the symmetric network model, while Fig. 8 (b) shows results for the asymmetric network model. To understand the edge-user performance, we focus on the rate outage probability at threshold 1bps/Hz. As observed in Fig. 8 (a), in a symmetric network, the proposed clustering has 0.04 rate outage probability, so that the rate outage probability is reduced by 60% and 80% compared to random clustering and fractional frequency reuse, respectively. In an asymmetric network, the rate outage probability of the proposed clustering is 0.157, so that the rate outage probability is reduced by 23.8% and 59% compared to random clustering and fractional frequency reuse, respectively. In both networks, the proposed clustering improves the rate outage performance compared to the baseline methods.

Now we compare the ergodic spectral efficiency. To focus on the cell edge-user performance, we assume that a user is placed $||\mathbf{d}_0|| / ||\mathbf{d}_1|| = 1$ by tuning $\mathbf{d}_0$, where $\mathbf{d}_0$ and $\mathbf{d}_1$ are locations of the associated BS and the closest interfering BS. This setting helps to shed light on how much performance gain is achieved to the worst case (edge) user by using the proposed clustering [32]. For fair comparison, the sum ergodic spectral efficiency per one cell is considered and it is assumed that served users in the cell has the same condition with the tagged user. Fig. 9 shows the ergodic spectral efficiency depending on received SNR. As SNR increases, the slope is separated into two regimes, called the DoF regime and the saturation regime, as introduced in [3].
In the DoF regime, increasing SNR brings significant improvement in terms of ergodic spectral efficiency, while the ergodic spectral efficiency does not increase substantially in the saturation regime. To show the performance gain of the proposed clustering depending on the each regime, we select two SNR, which are SNR = 0dB and SNR = 20dB. For a symmetric network case, the proposed clustering provides better ergodic spectral efficiency performance in both of the DoF regime and the saturation regime. Specifically, compared to each of random clustering and fractional frequency reuse, we get 15% and 240% performance gain at SNR = 0dB, and 27% and 81% performance gain at SNR = 20dB. For the asymmetric network case, as observed in Fig. 9 (b), the performance of the proposed clustering without the edge-cutting algorithm is below that of random clustering. As mentioned in Remark 1, this is mainly because it is possible that the proposed clustering demands too many time-frequency resources in an asymmetric network, e.g., $L = 11$ in our example. To solve this, the edge-cutting algorithm with $\Delta_{EC} \in \{3, 7, 10\}$ is applied. As shown in Fig. 9 (b), the edge-cutting algorithm with $\Delta_{EC} = 7$ outperforms the baseline methods, i.e., 15.5% performance gain compared to random clustering and at 48.7% performance gain compared to fractional frequency reuse are achieved at SNR = 20dB. The edge-cutting algorithm with $\Delta_{EC} \in \{3, 10\}$, however, cannot improve the performance. The reason behind this is explained as follows: Basically, the edge-cutting algorithm aims to reduce the time-frequency resources by sacrificing some of users that causes excessive time-frequency resource use. While the performance gain from the reduction of the required time-frequency resources is negligible when $\Delta_{EC} = 10$, too many users are out-of-service when $\Delta_{EC} = 3$. This is why the edge-cutting with $\Delta_{EC} \in \{3, 10\}$ cannot bring the performance improvement. When $\Delta_{EC} = 7$, large amount of time-frequency resources are saved while there are relatively small number of sacrificed users by cutting edges, which leads to the performance improvement. As in this observation, since $\Delta_{EC}$ has a substantial influence to the performance of the edge-cutting algorithm, $\Delta_{EC}$ should be carefully chosen for applying the edge-cutting algorithm. The optimal $\Delta_{EC}^*$ that provides the highest ergodic spectral efficiency can be numerically obtained by using simulations. Considering an extreme case, i.e., $\Delta_{EC} \to 0$, almost no user is served in the network, resulting in that the ergodic spectral efficiency goes to 0.
Fig. 8. Rate coverage probability comparison result for $N = 3$, $K = 1$, and $\beta = 4$ in the symmetric network (a) and in the asymmetric network (b).

Fig. 9. Ergodic spectral efficiency of an edge-user for $N = 3$, $K = 1$, and $\beta = 4$ in a symmetric network (a) and in an asymmetric network (b). The performance of the proposed clustering with the edge-cutting algorithm is denoted as a dashed line.

VII. CONCLUSION

In this paper, we proposed cluster patterns for semi-static BS clustering. The core feature of the proposed cluster pattern is to form a BS cluster according to the 2nd-order Voronoi regions, so that users are guaranteed to communicate the two closest BSs. To avoid BS conflicts, we use the notion of edge-coloring in graph theory. We assign different time-frequency resources
(colors) to each edge of the graph induced by the Delaunay triangulation. Employing the proposed cluster patterns, every user in the plane is guaranteed to communicate with the two closest BSs without the dominant interference. In a fixed network with irregular BSs’ locations, we provided analytical expressions of the ergodic spectral efficiency and coverage probability in terms of the number of antennas per BS, the number of users, distance, SNR, and the pathloss exponent. We also applied the proposed cluster pattern in a random network, where BSs locations and users locations are modeled by using a homogeneous PPP. We derived a lower bound of the ergodic spectral efficiency was derived as a function of the pathloss exponent and the number of associated users. Through Monte-Carlo simulations, we showed that the proposed clustering outperforms other conventional strategies. Future work should consider the application of more advanced coordination strategies going beyond coordinated beamforming.

REFERENCES

[1] D. Lee, H. Seo, B. Clerckx, E. Hardouin, D. Mazzarese, S. Nagata, and K. Sayana, “Coordinated multipoint transmission and reception in lte-advanced: deployment scenarios and operational challenges,” IEEE Comm. Mag., vol. 50, no. 2, pp. 148–155, Feb. 2012.
[2] O. Somekh, O. Simeone, Y. Bar-Ness, A. Haimovich, and S. Shamai, “Cooperative multicell zero-forcing beamforming in cellular downlink channels,” IEEE Trans. Info. Th., vol. 55, no. 7, pp. 3206–3219, Jul. 2009.
[3] A. Lozano, R. W. Heath, and J. G. Andrews, “Fundamental limits of cooperation,” IEEE Trans. Info. Th., vol. 59, no. 9, pp. 5213–5226, Sep. 2013.
[4] J. Zhang, R. Chen, J. Andrews, A. Ghosh, and R. W. Heath, “Networked MIMO with clustered linear precoding,” IEEE Trans. Wireless Comm., vol. 8, no. 4, pp. 1910–1921, Apr. 2009.
[5] K. Huang and J. Andrews, “An analytical framework for multicell cooperation via stochastic geometry and large deviations,” IEEE Trans. Info. Th., vol. 59, no. 4, pp. 2501–2516, Apr. 2013.
[6] D. Gesbert, S. Hanly, H. Huang, S. Shamai Shitz, O. Simeone, and W. Yu, “Multi-cell MIMO cooperative networks: A new look at interference,” IEEE Jour. Select. Areas in Comm., vol. 28, no. 9, pp. 1380–1408, Dec. 2010.
[7] J.-Y. Hwang, J. Kim, T. Kim, and Y. Han, “A periodic frequency band rotation scheme for multi-cell coordination clustering,” IEEE Comm. Lett., vol. 15, no. 9, pp. 956–958, Sep. 2011.
[8] H. Purmehdi, R. Elliott, W. Krzymien, and J. Melzer, “Rotating clustering with simulated annealing user scheduling for coordinated heterogeneous mimo cellular networks,” in Proc. IEEE Int. Conf. on Comm., Jun. 2014, pp. 5293–5298.
[9] A. Papadogiannis, D. Gesbert, and E. Hardouin, “A dynamic clustering approach in wireless networks with multi-cell cooperative processing,” in Proc. IEEE Int. Conf. on Comm., May 2008, pp. 4033–4037.
[10] A. Papadogiannis and G. Alexandropoulos, “The value of dynamic clustering of base stations for future wireless networks,” in Proc. IEEE Int. Conf. on Fuzzy Sys., Jul. 2010, pp. 1–6.
[11] R. Weber, A. Garavaglia, M. Schulist, S. Brueck, and A. Dekorsy, “Self-organizing adaptive clustering for cooperative multipoint transmission,” in Proc. IEEE Veh. Technol. Conf., May 2011, pp. 1–5.
[12] A. Giovanidis, J. Krolikowski, and S. Brueck, “A 0-1 program to form minimum cost clusters in the downlink of cooperating base stations,” in *Proc. IEEE Wireless Comm. and Net. Conf.*, Apr. 2012, pp. 940–945.

[13] M. Rahman and H. Yanikomeroglu, “Multicell downlink ofdm subchannel allocations using dynamic intercell coordination,” in *Proc. IEEE Glob. Comm. Conf.*, Nov. 2007, pp. 5220–5225.

[14] M. Rahman, H. Yanikomeroglu, and W. Wong, “Interference avoidance with dynamic inter-cell coordination for downlink lte system,” in *Proc. IEEE Wireless Comm. and Net. Conf.*, Apr. 2009, pp. 1–6.

[15] C. Ng and H. Huang, “Linear precoding in cooperative MIMO cellular networks with limited coordination clusters,” *IEEE Jour. Select. Areas in Comm.*, vol. 28, no. 9, pp. 1446–1454, Dec. 2010.

[16] J. Kim, H.-W. Lee, and S. Chong, “Virtual cell beamforming in cooperative networks,” *IEEE Jour. Select. Areas in Comm.*, vol. 32, no. 6, pp. 1126–1138, Jun. 2014.

[17] S. Nagaraj, M. Raghavendra, P. Fleming, and M. Honig, “Multi-cell distributed interference cancellation for co-operative pico-cell clusters,” in *Proc. IEEE Glob. Comm. Conf.*, Dec. 2012, pp. 4193–4199.

[18] N. Lee, D. Morales-Jimenez, A. Lozano, and R. Heath, “Spectral efficiency of dynamic coordinated beamforming: A stochastic geometry approach,” *IEEE Trans. Wireless Comm.*, vol. 14, no. 1, pp. 230–241, Jan. 2015.

[19] C. Li, J. Zhang, M. Haenggi, and K. B. Letaief, “User-centric intercell interference nulling for downlink small cell networks,” *ArXiv*, 2014. [Online]. Available: http://arxiv.org/abs/1405.4395

[20] R. Tanbourgi, S. Singh, J. Andrews, and F. Jondral, “A tractable model for noncoherent joint-transmission base station cooperation,” *IEEE Trans. Wireless Comm.*, vol. 13, no. 9, pp. 4959–4973, Sep. 2014.

[21] S. Ramprashad and G. Caire, “Cellular vs. Network MIMO: A comparison including the channel state information overhead,” in *Proc. IEEE Int. Symp. on Pers., Indoor and Mobile Radio Comm.*, Sep. 2009, pp. 878–884.

[22] China Mobile Research Institute, “C-RAN: The road toward Green RAN,” *White Paper*, Oct. 2011.

[23] W. Mennerich and W. Zirwas, “User centric coordinated multi point transmission,” in *Proc. IEEE Veh. Technol. Conf.*, Sep. 2010, pp. 1–5.

[24] Y. Du and G. de Veciana, “Wireless networks without edges: dynamic radio resource clustering and scheduling,” in *Proc. IEEE Int. Conf. on Comp. and Comm. (INFOCOM)*, Apr. 2014, p. to appear.

[25] C.-B. Chae, I. Hwang, R. Heath, and V. Tarokh, “Interference aware-coordinated beamforming in a multi-cell system,” *IEEE Trans. Wireless Comm.*, vol. 11, no. 10, Oct. 2012.

[26] I. Fischer and C. Gotsman, “Fast approximation of high-order Voronoi diagrams and distance transforms on the GPU,” *J. Graphics Tools*, vol. 11, no. 4, pp. 39–60, 2006.

[27] D. B. West *et al.*, *Introduction to graph theory*. Prentice hall Upper Saddle River, 2001, vol. 2.

[28] V. Vizing, “On an estimate of the chromatic class of a p-graph,” *Diskret. Analiz*, no. 3, pp. 25–30, 1964.

[29] D. P. Sanders and Y. Zhao, “Planar graphs of maximum degree seven are class i,” *Journal of Combinatorial Theory, Series B*, vol. 83, no. 2, pp. 201 – 212, 2001.

[30] K. Hamdi, “A useful lemma for capacity analysis of fading interference channels,” *IEEE Trans. Comm.*, vol. 58, no. 2, pp. 411–416, Feb. 2010.

[31] J. Andrews, F. Baccelli, and R. Ganti, “A tractable approach to coverage and rate in cellular networks,” *IEEE Trans. Comm.*, vol. 59, no. 11, pp. 3122–3134, Nov. 2011.

[32] G. Nigam, P. Minero, and M. Haenggi, “Coordinated multipoint joint transmission in heterogeneous networks,” vol. 62, no. 11, pp. 4134–4146, Nov. 2014.