Stochastic Geometric Analysis of Handoffs in User-Centric Cooperative Wireless Networks

Wei Bao* and Ben Liang†

*The University of Sydney, Australia, wei.bao@sydney.edu.au
†University of Toronto, Canada, liang@comm.utoronto.ca

***Abstract*—User-centric base station (BS) cooperation has been regarded as an effective solution to improve network coverage and throughput in next-generation wireless systems. However, it also introduces more complicated handoff patterns, which may potentially degrade user performance. In this paper, we aim to quantify the number of handoffs in user-centric cooperative wireless networks. The challenges are two-fold: (1) BSs are spatially randomly deployed, and (2) user-centric BS cooperation further creates complicated network topologies so that it is difficult to track handoffs in the system. We propose a stochastic geometric analysis framework on user mobility, to derive a theoretical expression for the handoff rate experienced by an active user with arbitrary movement trajectory. Furthermore, we characterize the average downlink user data rate under a common non-coherent joint-transmission scheme, which is used to illustrate the tradeoff between handoff rate and data rate in optimizing the cooperative cluster size for each user. Finally, computer simulation is conducted to validate the correctness and usefulness of our analysis.**

I. Introduction

Base station (BS) cooperation is expected to become an important feature in next-generation wireless networks [1]. It allows simultaneous connections from one user to multiple BSs, to significantly enhance the received power level and reduce interference. In addition, compared with the traditional single-BS association mode, users are less likely to enter a dead spot, e.g., near the cell edge, where the received signal-to-interference-plus-noise-ratio (SINR) becomes too low.

A central element in the implementation of BS cooperation is BS clustering, where a set of BSs are selected to cooperatively serve a user. There are two types of BS clustering modes, namely the *disjoint clustering* and *user-centric clustering*. In the disjoint clustering mode, the entire geographical region of the network is partitioned into multiple non-overlapping subregions, and the BSs in each subregion cooperatively serve users within the subregion. In the user-centric clustering mode, each user is served by its individual cluster of neighboring BSs. As the user moves, its BS clusters are updated, so that each BS appears in different clusters. The user-centric clustering mode is more advantageous compared with its disjoint counterpart, since the BS clusters are continuously updated based on the user location, which avoids the low-SINR cluster edges that are artificially created in disjoint clustering. Consequently, in this work, we focus on the user-centric clustering mode.

The majority of the work was done while the first author was affiliated with the University of Toronto, Canada.

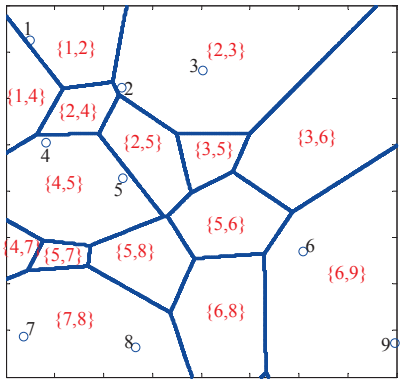
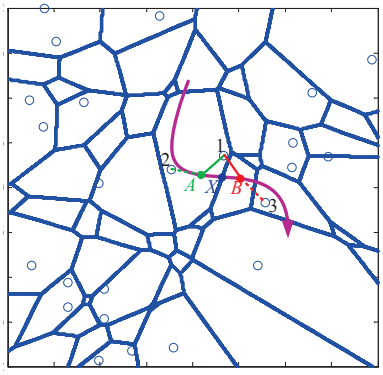


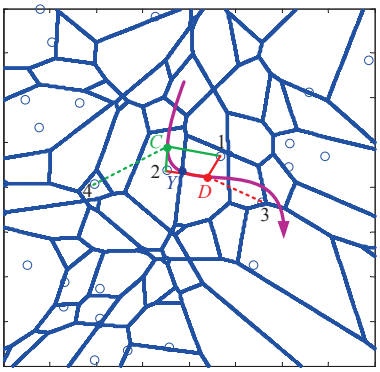
Fig. 1. An example of virtual cells. The numbers 1 to 9 indicate 9 BSs; $\{A, B\}$ indicates the virtual cell region served by BSs A and B.

However, in the presence of user-centric BS cooperation, mobility management becomes more challenging. The handoff patterns are more complicated compared with those in traditional single-BS association systems, since the handoffs now involve changes in terms of a set of multiple BSs. Different from the single-BS association scenario, there are no explicit cells surrounding individual BSs. Instead, we need to characterize the *virtual cell*, which corresponds to the region where a user connects with the same set of BSs. If a user crosses the boundary between two virtual cells, its connected BS set is changed and a handoff is made. However, due to the spatial randomness of BSs as well as user-centric BS cooperation, the virtual cells are generated randomly and irregularly. It is difficult to characterize the virtual cell boundary and to track boundary crossings made by users in the system. Fig. 1 shows an example topology where users are served by two closest BSs. Clearly, previously developed techniques for single-BS handoff analysis [2]–[5] are insufficient to model the complex handoff patterns in such user-centric cooperative wireless networks.

Characterizing the handoff rate can provide important guidelines for system design. For example, optimizing the BS cluster size requires accounting for the impacts of both the handoff rate and the data rate. Consider the example in Fig. 2, where an active user's trajectory is indicated by the



(a) A user is served by two closest BSs. It makes 4 handoffs. For example, it accesses BSs 1 and 2 at A, BSs 1 and 3 at B, and makes a handoff at X.



(b) Same BS locations and user trajectory as above. The user is served by three closest BSs. It makes 6 handoffs. For example, it accesses BSs 1, 2, and 4 at C, BSs 1, 2, and 3 at D, and makes a handoff at Y.

Fig. 2. A comparison between two-BS cooperation and there-BS cooperation. BSs are represented by circles; blue curves show virtual cell boundary; user trajectory is shown as the magenta arrow.

magenta arrow. If the user is served by two cooperative BSs, it experiences 4 handoffs, as shown in Fig. 2(a). In contrast, as shown in Fig. 2(b), if the user is served by three cooperative BSs, it experiences 6 handoffs. However, in the latter case, the user could potentially experience a higher data rate since it is served by one more BS. Thus, an optimal system design must balance the tradeoff between a larger BS cluster to improve data rate and more frequent handoffs, which increases cost and potentially deteriorates service quality.

In this work, we propose a new stochastic geometric analysis framework to quantify the handoff rate in a user-centric cooperative wireless network, where each user is served by its K closest BSs. We model the BSs as a Poisson point process (PPP) to capture their spatial randomness. Our contributions are as follows:

- Through stochastic and analytic geometric analysis, we derive an exact expression for the handoff rate experienced by an active user with arbitrary movement trajectory.
- As a study on the application of the above handoff rate analysis, after calculating the average downlink data rate of users under the non-coherent joint-transmission (NC-JT) scheme, we further investigate the optimal cluster size K, to balance the tradeoff between handoff rate and data rate. We find that the optimal cluster size is asymptotically inversely proportional to the square of the user velocity and inversely proportional to the BS intensity.
- Computer simulation is conducted to validate the correctness and usefulness of our analysis, through which we also show that the handoff rate derived under the PPP assumption provides close approximations even if the BSs are non-PPP distributed.

The rest of this paper is organized as follows. In Section II, we discuss the relation between our work and prior works. In Section III, we present the system model. In Section IV, we present the theoretical analysis on the handoff rate in the system. In Section V, we study the optimal cluster size as an application scenario. In Section VI, we validate our analysis with simulation. Finally, conclusions are given in Section VII.

II. Related Works

In this section, we summarize the prior research in stochastic geometric modeling of cooperative wireless networks, and the existing techniques for handoff analysis.

A. Stochastic Geometric Analysis of Cooperative Wireless Networks

In order to capture the spatial randomness of BSs and users, analysis techniques based on the theory of stochastic geometry have been applied to evaluate performance metrics such as interference distribution, coverage probability, data rate, and throughput in cooperative wireless networks. In these works, the BSs are often assumed to be spatially distributed as a PPP. In [6], a two-BS cooperation model was proposed and the user coverage probability was derived for this model. In [7]–[9], various forms of the disjoint clustering mode of BS cooperation were studied. In [10]–[15], the performance of different joint transmission schemes in user-centric BS cooperation was evaluated, including NC-JT [10], [11], synchronous joint transmission [12], interference nulling [13], and coordinated beamforming [14], [15]. All of these works focused only on networks with stationary users, and thus handoffs were not studied.

B. Handoff Analysis in Wireless Cellular Networks

In the scope of handoff analysis, all previous works concerned only the single-BS association scenario. One well-known category of analysis techniques employ queueing formulation, without explicitly modeling the geometric patterns of cell shapes in the networks [16]–[19]. In these works, cells were modeled as queues containing active users, and handoffs were modeled as unit transfers between queues. Another common category of analysis techniques assume regularly grided cells for mathematical convenience. Examples of such geometric topologies include hexagonal grids [20], [21], square grids [22], and circles overlaying hexagons [23].

To further capture the spatial randomness of network topologies, a seminal study on user mobility was conducted in [2] for a single-tier cellular network with randomly distributed BSs, where the BSs were modeled as a homogeneous PPP, and cell splitting was modeled as a standard Poisson Voronoi. The case of multi-tier cellular networks was considered in [3] and [4], where each tier of BSs was modeled as a homogeneous PPP, and the resultant cell splitting was modeled as a weighted Poisson Voronoi. Further extension to [3] was given in [5], where the BS tiers were modeled as Poisson cluster processes (PCPs), such that their aggregation around highly populated areas could be accommodated. However, the above works considered only single-BS association, which is not applicable to BS cooperative wireless networks.

III. System Model

In this section, we describe the user-centric cooperative network under consideration, clarifying the notions of user handoffs and virtual cells.

A. User-Centric Cooperative Wireless Network

We consider a single-tier wireless system with BSs scattered in the two-dimensional Euclidean space $\mathbb{R}^2$ according to a homogeneous Poisson point process (PPP) Φ with intensity λ. The PPP assumption is commonly adopted in stochastic geometric analysis in the research literature [10]–[15], [24]–[27]. We additionally show through simulation in Section VI that our results derived under the PPP assumption provide close approximations even when BSs are non-PPP distributed.

A general user-centric cooperation scheme where each user is served by its K closest BSs is assumed, where the minimum value of K is $K_{\min} = 2$. Note that we do not need to specify the transmission scheme for handoff analysis, so that this model is applicable to many different scenarios, such as NC-JT, cooperative beamforming, and cloud radio access network (C-RAN).

B. Handoffs in User-Centric Cooperative Wireless Network

A handoff is defined as the event that the connected BS set of an active user is changed. One major goal of this work is to quantify the rate of handoffs of some active user moving in the network. Therefore, we need to characterize the *virtual cell*, defined as the region in which a user is served by the same set of K BSs. Since Φ is a PPP, the overall virtual cells correspond to a Kth-order Poisson Voronoi [6], [14], an example of which is shown in Fig. 1. Let $\mathbf{T}_K^{(1)}$ denote the overall set of cell boundary of the Kth-order Poisson Voronoi. Whenever an active user crosses $\mathbf{T}_K^{(1)}$, the set of connected BSs are changed, and thus a handoff is made. Let $\mathcal{T}_0$ denote the trajectory of the user, which is of finite length. The number of handoffs the user experiences is equal to the number of intersections between $\mathcal{T}_0$ and $\mathbf{T}_K^{(1)}$, which is denoted by $\mathcal{N}(\mathcal{T}_0, \mathbf{T}_K^{(1)})$.

C. Kth-order Poisson Voronoi

We formally define $\mathbf{T}_K^{(1)}$ as follows. Let $\mathcal{C} = \{\mathbf{x}_1, \ldots, \mathbf{x}_K\} \subset \Phi$ denote a set of K BSs. The Kth-order Voronoi cell with respect to the BS set $\mathcal{C}$ is defined as the set of points closer to $\mathbf{x}_1, \ldots, \mathbf{x}_K$ than any other points in Φ, i.e.,

$$\mathcal{V}(\mathcal{C}) = \{\mathbf{y} \in \mathbb{R}^2 \big| \forall \mathbf{x} \in \mathcal{C}, \mathbf{x}' \in \Phi \backslash \mathcal{C}, |\mathbf{y} - \mathbf{x}| \leq |\mathbf{y} - \mathbf{x}'|\}. \tag{1}$$

In other words, the distance from an arbitrary point in $\mathcal{V}(\mathcal{C})$ to any BS in $\mathcal{C}$ is no larger than the distance from the point to any BS not in $\mathcal{C}$. Note that a BS may not be in the cell formed by itself. For example, in Fig. 1, neither BS 6 nor BS 8 is inside the region served by BSs $\{6, 8\}$. We also note that for some $\mathcal{C}$, $\mathcal{V}(\mathcal{C}) = \emptyset$. In Fig. 1, nowhere is served by BSs $\{1, 9\}$.

Thus, $\mathbf{T}_K^{(1)}$ corresponds to the set of points on $\mathbb{R}^2$ which belongs to two different cells:

$$\mathbf{T}_K^{(1)} = \{\mathbf{y} \in \mathbb{R}^2 \big| \exists \mathcal{C} \neq \mathcal{C}', \text{s.t. } \mathbf{y} \in \mathcal{V}(\mathcal{C}) \bigcap \mathcal{V}(\mathcal{C}')\}. \tag{2}$$

Note that $\mathbf{T}_K^{(1)}$ can be determined by Φ, and thus it is a fiber process [28] generated by Φ. Because Φ is stationary and isotropic, $\mathbf{T}_K^{(1)}$ is also stationary and isotropic.

IV. Handoff Rate Analysis

In this section, we present an analytical framework to quantify the handoff rate. First, we rewrite $\mathbf{T}_K^{(1)}$ in a more appropriate form. Second, the handoff rate is derived through analyzing the length intensity of $\mathbf{T}_K^{(1)}$, which is in turn derived through characterizing the area intensity of the Δd-neighborhood of $\mathbf{T}_K^{(1)}$.

Note that in the rest of this paper, we define $\mathcal{B}(\mathbf{x}, r)$ as the disk region $\big\{\mathbf{y} \in \mathbb{R}^2 \big| |\mathbf{x} - \mathbf{y}| \leq r\big\}$, and $\mathcal{B}_c(\mathbf{x}, r)$ as the region $\big\{\mathbf{y} \in \mathbb{R}^2 \big| |\mathbf{x} - \mathbf{y}| \geq r\big\}$.

A. Rewriting Cell Boundary $\mathbf{T}_K^{(1)}$

We first rewrite $\mathbf{T}_K^{(1)}$ in a more appropriate form, which will facilitate the handoff analysis in the subsequent steps.

Theorem 1. $\mathbf{T}_K^{(1)}$ can be rewritten as follows:

$$\begin{aligned}\mathbf{T}_K^{(1)} = \big\{\mathbf{y} \in \mathbb{R}^2 | \exists \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{K-1}, \mathbf{x}_K, \mathbf{x}_K'\} \subset \Phi, \\ \text{s.t. } |\mathbf{z} - \mathbf{y}| \leq |\mathbf{x}_K - \mathbf{y}| = |\mathbf{x}_K' - \mathbf{y}| \leq |\mathbf{x} - \mathbf{y}|, \forall \mathbf{z} \in \\ \{\mathbf{x}_1, \ldots, \mathbf{x}_{K-1}\} \text{ and } \forall \mathbf{x} \in \Phi \backslash \{\mathbf{x}_1, \ldots, \mathbf{x}_{K-1}, \mathbf{x}_K, \mathbf{x}_K'\}\big\}.\end{aligned} \tag{3}$$

See Appendix A for the proof.

Theorem 1 suggests that $\mathbf{T}_K^{(1)}$ is the set of points, whose distances to two BSs are the same, and this distance is greater than or equal to the distances to some arbitrary set of $K-1$ BSs, but is less than or equal to the distances to all the other BSs.

B. Length Intensity and Area Intensity

Handoffs occur at the intersections between an active user's trajectory with $\mathbf{T}_K^{(1)}$. In order to track the number of intersections, we need to first study the intensity of $\mathbf{T}_K^{(1)}$. Higher intensity of $\mathbf{T}_K^{(1)}$ leads to greater opportunities for boundary crossing, and thus higher handoff rate.

Let $\mu_1\big(\mathbf{T}_K^{(1)}\big)$ denote the length intensity of $\mathbf{T}_K^{(1)}$, which is defined as the expected length of $\mathbf{T}_K^{(1)}$ in a unit square.

Because $\mathbf{T}_K^{(1)}$ is stationary and isotropic, the unit square could be arbitrarily picked on $\mathbb{R}^2$. Hence, we have

$$\mu_1\big(\mathbf{T}_K^{(1)}\big) = \mathbb{E}\left(\left|\mathbf{T}_K^{(1)} \bigcap [0,1)^2\right|_1\right), \tag{4}$$

where $|L|_1$ denotes the length of a collection of curves L (i.e., one-dimensional Lebesgue measure of L).

In order to derive $\mu_1\big(\mathbf{T}_K^{(1)}\big)$, we need to introduce the Δd-extended cell boundary of $\mathbf{T}_K^{(1)}$, denoted by $\mathbf{T}_K^{(2)}(\Delta d)$, which is defined as

$$\mathbf{T}_K^{(2)}(\Delta d) = \left\{\mathbf{y} \in \mathbb{R}^2 \,\middle|\, \exists \mathbf{x} \in \mathbf{T}_K^{(1)},\ \text{s.t. } |\mathbf{x}-\mathbf{y}| < \Delta d\right\}. \tag{5}$$

In other words, $\mathbf{T}_K^{(2)}(\Delta d)$ is the Δd-neighborhood of $\mathbf{T}_K^{(1)}$. A point is in $\mathbf{T}_K^{(2)}(\Delta d)$ if and only if its (shortest) distance to $\mathbf{T}_K^{(1)}$ is less than Δd.

The area intensity of $\mathbf{T}_K^{(2)}(\Delta d)$ is defined as the expected area of $\mathbf{T}_K^{(2)}(\Delta d)$ in a unit square:

$$\mu_2\big(\mathbf{T}_K^{(2)}(\Delta d)\big) = \mathbb{E}\left(\left|\mathbf{T}_K^{(2)}(\Delta d) \bigcap [0,1)^2\right|_2\right), \tag{6}$$

where $|S|_2$ denotes the area of some region S (i.e., two-dimensional Lebesgue measure of S).

Note that $\mathbf{T}_K^{(2)}(\Delta d)$ is stationary and isotropic. As a result, given a reference user located at $\mathbf{0}$, the area intensity of $\mathbf{T}_K^{(2)}(\Delta d)$ is equal to the probability that the reference user at $\mathbf{0}$ is in $\mathbf{T}_K^{(2)}(\Delta d)$.

$$\mu_2\big(\mathbf{T}_K^{(2)}(\Delta d)\big) = \mathbb{P}(\mathbf{0} \in \mathbf{T}_K^{(2)}(\Delta d)). \tag{7}$$

The probability in (7) is analytically tractable, which will be presented in the next subsection.

C. Derivations of Area Intensity of $\mathbf{T}_K^{(2)}(\Delta d)$

In this subsection, we present the derivation of $\mathbb{P}\left(\mathbf{0} \in \mathbf{T}_K^{(2)}(\Delta d)\right)$. First, we study the probability that the reference user at $\mathbf{0}$ is in $\mathbf{T}_K^{(2)}(\Delta d)$, given the distance between $\mathbf{0}$ and its Kth closest BS. We observe the following theorem:

Theorem 2. Suppose the reference user is located at $\mathbf{0}$, the distance between the reference user and its Kth closest BS is R_K. The conditional probability of $\mathbf{0} \in \mathbf{T}_K^{(2)}(\Delta d)$ given $R_K = r_0$ is

$$\mathbb{P}\left(\mathbf{0} \in \mathbf{T}_K^{(2)}(\Delta d) | R_K = r_0\right) = 8\lambda \Delta d r_0 + \mathcal{O}(\Delta d^2). \tag{8}$$

See Appendix B for the proof.

Second, through deconditioning on R_K, we can derive the unconditioned probability that the reference user at $\mathbf{0}$ is in $\mathbf{T}_K^{(2)}(\Delta d)$.

Theorem 3. The area intensity of $\mathbf{T}_K^{(2)}(\Delta d)$ is

$$\begin{aligned}\mu_2\big(\mathbf{T}_K^{(2)}(\Delta d)\big) =& \mathbb{P}(\mathbf{0} \in \mathbf{T}_K^{(2)}(\Delta d)) \\ =& \frac{8\Gamma\left(\frac{1}{2}+K\right)\sqrt{\lambda}\Delta d}{\Gamma(K)\sqrt{\pi}} + \mathcal{O}(\Delta d^2),\end{aligned} \tag{9}$$

where $\Gamma(\cdot)$ denotes the Gamma function.

See Appendix C for the proof.

D. From Area Intensity to Handoff Rate

We first derive the length intensity of $\mathbf{T}_K^{(1)}$ from the area intensity of $\mathbf{T}_K^{(2)}(\Delta d)$:

$$\mu_1\big(\mathbf{T}_K^{(1)}\big) = \lim_{\Delta d \to 0} \frac{\mu_2\big(\mathbf{T}_K^{(2)}(\Delta d)\big)}{2\Delta d} \tag{10}$$

$$= \frac{4\Gamma\left(\frac{1}{2}+K\right)\sqrt{\lambda}}{\Gamma(K)\sqrt{\pi}}, \tag{11}$$

where (10) is obtained by noting the relationship between the total length of a collection of curves in $\mathbb{R}^2$ and the total area of their Δd-neighborhood [29, Section 3.2].

Second, we note that the expected number of intersections between an arbitrary curve and a stationary and isotropic fiber process in $\mathbb{R}^2$ is $\frac{2}{\pi}$ multiplied by both the length of the curve and the length intensity of the fiber process [28, Section 9.3]. Therefore, the expected number of intersections between an arbitrary user's trajectory $\mathcal{T}_0$ and $\mathbf{T}_K^{(1)}$ (i.e., handoffs) is given by

$$\mathbb{E}\left(\mathcal{N}(\mathcal{T}_0, \mathbf{T}_K^{(1)})\right) = \frac{2}{\pi}\mu_1\big(\mathbf{T}_K^{(1)}\big)|\mathcal{T}_0|_1, \tag{12}$$

where $|\mathcal{T}_0|_1$ denotes the length of $\mathcal{T}_0$.

Finally, let v denote the instantaneous velocity of an active user, and $H(K,v)$ denote its handoff rate given K and v. Then, from (11)-(12) we have

$$H(K,v) = \frac{8\Gamma\left(\frac{1}{2}+K\right)\sqrt{\lambda}}{\Gamma(K)\pi\sqrt{\pi}} v. \tag{13}$$

E. One Useful Property of Handoffs

Whenever a user makes a handoff, the user is at $\mathbf{T}_K^{(1)}$. From (3), we know that the distances between the user to two reference BSs are the same, and this reference distance is greater than or equal to the distances to some arbitrary set of $K-1$ BSs, but is less than or equal to the distances to all the other BSs. Since BSs are randomly distributed on the two-dimensional space, the probability that the reference distance is exactly equal to the distance between the user to any BSs other than the two reference BSs is 0. Thus, with probability 1, the handoff is made only between the two reference BSs, and all the other BSs are not involved. Therefore, we can conclude that with probability 1, a handoff is a soft handoff where only one of the K connected BSs is changed. In this case, (13) is equivalent to the soft handoff rate where only one of the K connected BSs is changed. Other types of handoff rates, where more than one BSs are changed, are all 0.

Note that the handoffs in the user-centric clustering scenario are quite different with those in the disjoint clustering scenario. In the disjoint clustering scenario, the entire cluster of BSs are changed when the user crosses the cluster boundary.

V. Downlink User Data Rate Analysis and Optimal BS Cluster Size

In this section, we present an application scenario of the above handoff rate analysis. We first study the downlink user data rate under the NC-JT scheme. Then, we discuss the

optimal cluster size K that balances the handoff rate and the data rate. Note that we focus on the NC-JT scheme because it is one of the most commonly adopted cooperative transmission schemes in practical systems [30], and it is easily implemented since the tight synchronization of joint signal transmission is not required [10].

Previous works such as [10], [11] derived the downlink user data rate in non-closed form with multiple levels of integrations, which brings great difficulty to design the optimal cluster size K. In this work, we propose an alternative method, where a constant term with respect to K is ignored in the data rate analysis, and the optimal K is then derived in a simplified way.

A. Non-Coherent Joint Transmission Model

In this subsection, we briefly present the channel model and the NC-JT scheme. In addition to the general user-centric model presented in Section III, we make additional assumptions as follows.

We assume that each user and BS is equipped with a single-antenna. Each BS transmits at power level P. If a BS is located at $\mathbf{x}$, then the received power at $\mathbf{y}$ is $\frac{Ph_{\mathbf{x},\mathbf{y}}}{|\mathbf{x}-\mathbf{y}|^{\alpha}}$, where $\alpha > 2$ is the pathloss exponent, $|\mathbf{x}-\mathbf{y}|^{\alpha}$ is the propagation loss function, and $h_{\mathbf{x},\mathbf{y}}$ is the normalized fast fading term. Corresponding to common Rayleigh fading with power normalization, $h_{\mathbf{x},\mathbf{y}}$ is independently exponentially distributed with unit mean. After assigning K closest BSs to a user, NC-JT is implemented in the downlink transmission, so that the user receives a non-coherent sum of multiple copies of the useful signal transmitted by the K cooperative BSs, and BSs not in the cooperation set generate interference to the user [10], [11]. In addition, we focus on the interference limited scenario, where the noise is negligible.

B. Data Rate Analysis

In this subsection, we study the average user data rate via stochastic geometric analysis. Due to the stationarity of BSs, we focus on the average performance of a reference user located at $\mathbf{0}$, which is equivalent to the average user performance in the system [25].

Without loss of generality, we assume that the user is operated on a unit frequency bandwidth. Following the discussion in [10], [11], under NC-JT, the signal to interference ratio (SIR) at the reference user is expressed as

$$\mathrm{SIR}(K) = \frac{\sum_{\mathbf{x}\in\Phi_K} |\mathbf{x}|^{-\alpha} h_{\mathbf{x},\mathbf{0}}}{\sum_{\mathbf{x}\in\Phi_K^c} |\mathbf{x}|^{-\alpha} h_{\mathbf{x},\mathbf{0}}}, \tag{14}$$

where Φ_K corresponds to the point process of the K (closest) cooperative BSs, and Φ_K^c corresponds to the point process of the other non-cooperative BSs. Let $S(K) \triangleq \sum_{\mathbf{x}\in\Phi_K} |\mathbf{x}|^{-\alpha} h_{\mathbf{x},\mathbf{0}}$ be the received signal power from the K cooperative BSs, and $I(K) \triangleq \sum_{\mathbf{x}\in\Phi_K^c} |\mathbf{x}|^{-\alpha} h_{\mathbf{x},\mathbf{0}}$ be the sum interference caused by non-cooperative BSs. Following conventional stochastic geometric analysis, we study the worst case scenario where the interference is summed over all non-cooperative BSs [10]–[12], [14]. Then, the average data rate of the reference user is

$$\begin{aligned} R(K) &= \mathbb{E}\left[\log_2(1+\mathrm{SIR}(K))\right] \\ &= \mathbb{E}\left[\log_2(S(K)+I(K))\right] - \mathbb{E}\left[\log_2(I(K))\right]. \end{aligned} \tag{15}$$

Note that we have

$$S(K) + I(K) = \sum_{\mathbf{x}\in\Phi} |\mathbf{x}|^{-\alpha} h_{\mathbf{x},\mathbf{0}}, \tag{16}$$

which is a term irrelevant to K. Since we aim to derive the optimal K, the term $\mathbb{E}\left[\log_2(S(K)+I(K))\right] \triangleq C_0$ can be regarded as a constant and is ignored in the subsequent analysis.

In the next step, we study $\mathbb{E}\left[\log_2(I(K))\right]$. However, this term is still difficult to characterize. Therefore, we resort to analyzing its upper bound using Jensen's inequality:

$$\mathbb{E}\left[\log_2(I(K))\right] \leq \log_2(\mathbb{E}[I(K)]). \tag{17}$$

Correspondingly, we focus on a lower bound of the average data rate as follows:

$$R'(K) = C_0 - \log_2(\mathbb{E}[I(K)]). \tag{18}$$

As shown in Sections V-C and V-D, the characterization of $\log_2(\mathbb{E}[I(K)])$ instead of $\mathbb{E}\left[\log_2(I(K))\right]$ will lead to a simple closed-form expression, which can then be used to search for the optimal K in a simplified manner. The steps (15)-(18) differentiate our work with previous research literature, such as [10], [11], where the user data rate is expressed in non-closed form with multiple levels of integrations, which brings great difficulty in subsequent optimization of K. Furthermore, as shown in Section VI, the values of $\mathbb{E}\left[\log_2(I(K))\right]$ are close to those of $\log_2(\mathbb{E}[I(K)])$ over a wide range of parameter settings. Therefore, the approximation of $\mathbb{E}\left[\log_2(I(K))\right]$ by $\log_2(\mathbb{E}[I(K)])$ is an important simplification step in deriving the optimal K.

C. Derivation of $\mathbb{E}\left[I(K)\right]$

The overall interference is summed over all BSs outside the set of K closest BSs to $\mathbf{0}$. Given the distance from the reference user to its Kth closest BS $R_K = r_0$, the point process of Φ_K^c is a PPP with intensity λ in the range $\mathcal{B}_c(\mathbf{0}, r_0)$. Therefore, the conditional average interference can be computed as

$$\mathbb{E}[I(K)|R_K = r_0] = \lambda \int_{\mathcal{B}_c(\mathbf{0},r_0)} |\mathbf{x}|^{-\alpha} \mathrm{d}\mathbf{x} \tag{19}$$

$$= 2\pi\lambda \int_{r_0}^{\infty} r^{1-\alpha} \mathrm{d}r = 2\pi\lambda \frac{r_0^{2-\alpha}}{\alpha-2}. \tag{20}$$

Then, through deconditioning on R_K, and considering (20) and (32), we have

$$\mathbb{E}[I(K)] = \int_0^{\infty} 2\pi\lambda \frac{r_0^{2-\alpha}}{\alpha-2} \frac{2(\lambda\pi r_0^2)^K}{r_0\Gamma(K)} \exp(-\lambda\pi r_0^2) \mathrm{d}r_0 \tag{21}$$

$$= \frac{2\pi^{\frac{\alpha}{2}} \lambda^{\frac{\alpha}{2}} \Gamma(K+1-\frac{\alpha}{2})}{(\alpha-2)\Gamma(K)}. \tag{22}$$

Finally, the term $\log_2(\mathbb{E}[I(K)])$ can be derived accordingly from (22).

D. Optimal Cluster Size

In this subsection, we investigate the optimal cluster size based on the handoff rate study in Section IV and the data rate study in Sections V-B and V-C. Let K^* denote the optimal cluster size. K^* is an integer greater than or equal to $K_{\min}$.

In order to quantify the tradeoff between user date rate and handoff cost, we consider their weighted sum. Let W_1 be the utility value for one bit of data transmission, and W_2 be the cost for one handoff. Note that we assign the same cost value to all handoffs because each of them is a soft handoff where only one of the K connected BSs is changed with probability 1, as shown in Section IV-E. Consequently, the overall average utility of a user being served by K BSs is

$$\begin{aligned} U(K) &= W_1 R'(K) - W_2 H(K) \\ &= W_1 C_0 - W_1 \log_2(e) \ln\left(\frac{2\pi^{\frac{\alpha}{2}}\lambda^{\frac{\alpha}{2}}\Gamma(K+1-\frac{\alpha}{2})}{(\alpha-2)\Gamma(K)}\right) \\ &\quad - W_2 \frac{8\Gamma\left(\frac{1}{2}+K\right)\sqrt{\lambda}}{\Gamma(K)\pi\sqrt{\pi}} v. \end{aligned} \tag{23}$$

We define $L(K) \triangleq \ln\left(\frac{2\pi^{\frac{\alpha}{2}}\lambda^{\frac{\alpha}{2}}\Gamma(K+1-\frac{\alpha}{2})}{(\alpha-2)\Gamma(K)}\right)$ and $H(K) \triangleq \frac{8\Gamma\left(\frac{1}{2}+K\right)\sqrt{\lambda}}{\Gamma(K)\pi\sqrt{\pi}} v$. K^* is the integer that maximizes $U(K)$, or equivalently, minimizes $W_1 \log_2(e) L(K) + W_2 H(K)$. We also define

$$\begin{aligned} \Delta H(K) &\triangleq H(K+1) - H(K) \\ &= W_2 \frac{\Gamma(K+\frac{1}{2})}{\Gamma(K+1)} \frac{1}{2} \frac{8\sqrt{\lambda}}{\pi\sqrt{\pi}} v, \end{aligned} \tag{24}$$

and

$$\begin{aligned} \Delta L(K) &\triangleq L(K+1) - L(K) \\ &= W_1 \log_2(e) \ln\left(\frac{K+1-\frac{\alpha}{2}}{K}\right). \end{aligned} \tag{25}$$

It is straightforward to show that $\Delta H(K)$ is positive and $\Delta L(K)$ is negative, so that $H(K)$ is an increasing function and $L(K)$ is a decreasing function. In the next step, in order to derive the optimal K^*, we focus on the term $-\frac{\Delta H(K)}{\Delta L(K)}$. We note that $-\frac{\Delta H(K)}{\Delta L(K)} < 1$ implies that $H(K) + L(K)$ is decreasing at K and $-\frac{\Delta H(K)}{\Delta L(K)} > 1$ implies that $H(K)+L(K)$ is increasing at K. Also, we have the following theorem:

Theorem 4. $-\frac{\Delta H(K)}{\Delta L(K)}$ is an increasing function of K.

See Appendix D for the proof.

In addition to Theorem 4, we notice that $-\frac{\Delta H(K)}{\Delta L(K)} \to \infty$ when K is sufficient large, and $-\frac{\Delta H(K)}{\Delta L(K)} \to 0$ when K approaches $\frac{\alpha}{2} - 1$. Let $\widetilde{K}^*$ be the solution to $-\frac{\Delta H(K)}{\Delta L(K)} = 1$, then we have the following conclusion:

- If $-\frac{\Delta H(K_{\min})}{\Delta L(K_{\min})} \geq 1$, then $H(K) + L(K)$ is an increasing function and $K^* = K_{\min}$.

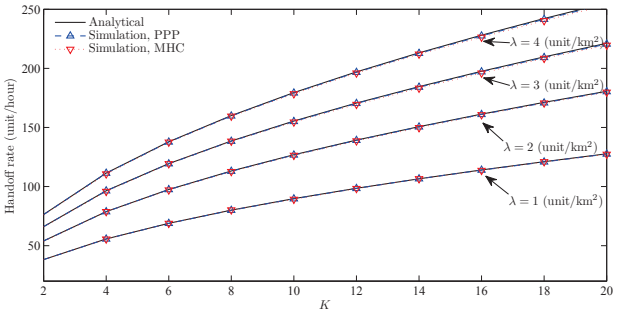


Fig. 3. Handoff rate under different K.

- Otherwise, $H(K) + L(K)$ is firstly decreasing and then increasing, and K^* is equal to $\lfloor \widetilde{K}^* \rfloor$ or $\lceil \widetilde{K}^* \rceil$, whichever that minimizes $H(K) + L(K)$.

E. Asymptotic Value of $\widetilde{K}^*$

From the previous subsection, we see that $\widetilde{K}^*$ is the solution to $-\frac{\Delta H(K)}{\Delta L(K)} = 1$, which can be derived through a simple numerical search. However, it is not in closed form. For deeper insights in characterizing K^*, We are interested in further seeking an approximated expression of $\widetilde{K}^*$ in closed form.

First, we observe the asymptotic values of $\Delta L(K)$ and $\Delta H(K)$, when K is large, are:

$$\Delta L(K) \simeq -W_1 \log_2(e) \frac{\left(\frac{\alpha}{2}-1\right)}{K}, \tag{26}$$

and

$$\Delta H(K) \simeq \frac{W_2}{\sqrt{K}} \frac{4\sqrt{\lambda}}{\pi\sqrt{\pi}} v. \tag{27}$$

Then, by substituting the right-hand sides of (26) and (27) into $-\frac{\Delta H(K)}{\Delta L(K)} = 1$, we derive the asymptotic value of $\widetilde{K}^*$, denoted by $\widehat{K}^*$, as follows:

$$\widehat{K}^* = \frac{W_1^2 \left(\log_2(e)\right)^2 \left(\frac{\alpha}{2}-1\right)^2 \pi^3}{16 W_2^2 \lambda v^2}. \tag{28}$$

The expression (28) suggests that the optimal cluster size is asymptotically inversely proportional to the square of the user velocity v and inversely proportional to the BS intensity λ. Note that in Section VI, we further show through numerical study that even if $\widetilde{K}^*$ is not large, the values of $\widehat{K}^*$ are still close to those of $\widetilde{K}^*$.

VI. Simulation Study

In this section, we present simulation studies to validate the accuracy and usefulness of our proposed analysis. In each round of simulation, BSs are generated on a 20 km $\times$ 20 km square. Then, we randomly generate 5 waypoints $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_5$ in the central 10 km $\times$ 10 km square. The four line segments $\mathbf{X}_1\mathbf{X}_2, \mathbf{X}_2\mathbf{X}_3, \ldots, \mathbf{X}_4\mathbf{X}_5$ form the trajectory of an active user in one round of simulation. By tracking which set of BSs the user is connected to along its trajectory, we can track handoffs of the user in this round of simulation. Each data point is averaged over 2000 simulation rounds.

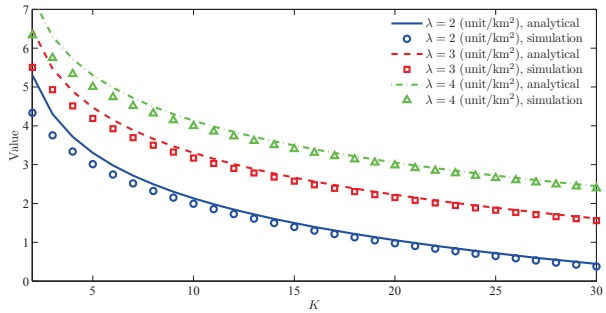


Fig. 4. Comparison of simulated $\mathbb{E}\left[\log_2(I(K))\right]$ and analytical $\log_2(\mathbb{E}[I(K)])$.

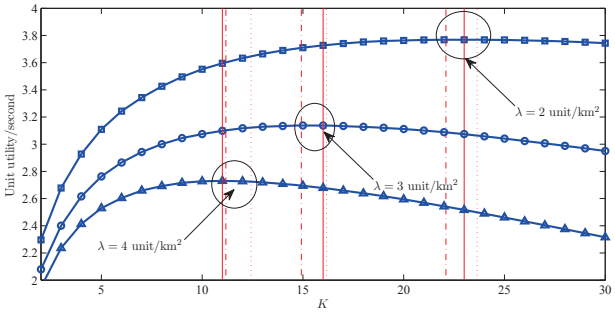


Fig. 5. User utility under different K. The solid, dotted, and dashed dotted vertical lines correspond to the optimal K^* derived by simulation, the analytical values of $\widetilde{K}^*$ and $\widehat{K}^*$ respectively.

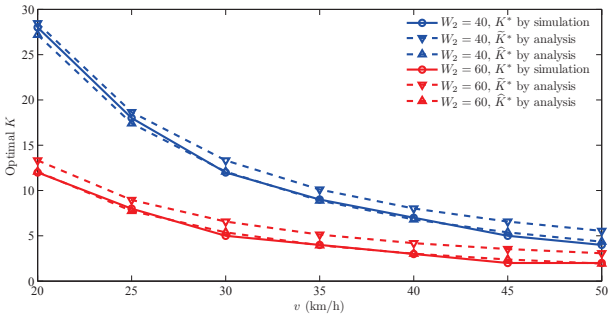


Fig. 6. Optimal K under different user velocity v.

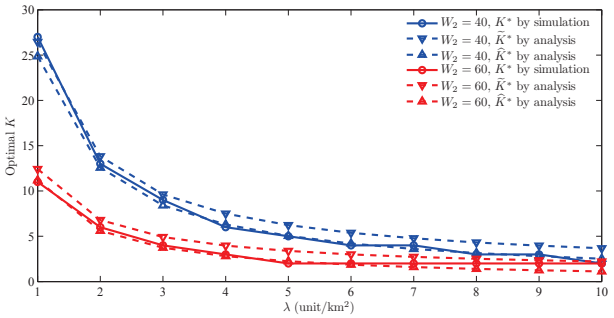


Fig. 7. Optimal K under different BS intensity λ.

We first study the handoff rates under different K and λ values in Fig. 3. The BS intensity is set to $\lambda = 1$, 2, 3, and 4 units/km^2 respectively, and the user velocity is $v = 20$ km/h. As discussed in Section III-A, we are also interested in testing the scenario where BSs are non-PPP distributed. Thus in the simulation, we also consider the case where BSs are distributed as a Matérn hard core (MHC) point process.

The MHC point process has been regarded as an alternative point process in the research literature to counter the drawback of PPP modeling of wireless cellular networks [31], [32]. It can additionally capture the reality that two BSs are unlikely to be located very close to each other. In our simulation, BSs are generated as an MHC point process as follows: First, we generate a PPP with intensity λ'. Each point in the PPP is associated with a "mark", which is independently uniformly distributed on $[0, 1]$. A point is retained in the point process if its mark is the largest among all the points within a distance D from it (or there are no other points within this range); otherwise, the point is removed from the point process. The remaining points form an MHC point process. Note that the distance between any two points in the point process is no less than D. The equivalent BS intensity is $\lambda = \frac{1-e^{-\pi D^2 \lambda'}}{\pi D^2}$. Under the MHC setting, we set $D = 0.1$ km, $\lambda' = 1.0160$, 2.0656, 3.1509, and 4.2746 units/km^2 respectively, in order to maintain the equivalent BS intensity at $\lambda = 1$, 2, 3, and 4 units/km^2 respectively.

Fig. 3 validates that the handoff rate is an increasing and concave function of K, which matches the expression (13). Also, when the BSs are PPP distributed, the handoff rates obtained from simulation match well with our proposed analysis, validating the correctness of our analytical derivations in Section IV. Furthermore, even if the BSs are distributed as an MHC point process, the simulated handoff rates are still very close to those under the PPP assumption. Therefore, our proposed analysis is still useful to provide close approximations of handoff rates when BSs are more realistically distributed.

In Fig. 4, we show a comparison between the values of $\mathbb{E}\left[\log_2(I(K))\right]$ through simulation and those of $\log_2(\mathbb{E}[I(K)])$ derived in Sections V-B and V-C. Fig. 4 shows that the gap between $\mathbb{E}\left[\log_2(I(K))\right]$ and $\log_2(\mathbb{E}[I(K)])$ is small, so that we can use $\mathbb{E}\left[\log_2(I(K))\right]$ to approximate $\log_2(\mathbb{E}[I(K)])$ in this work.

In Fig. 5, we present the simulated user utility $W_1 R(K) - W_2 H(K)$ under different values of the cluster size K. The network parameters are as follows: $\alpha = 4$, $W_1 = 1$, $W_2 = 30$, $v = 36$ km/h, and $P = 30$ dBm. The simulated results are plotted for $\lambda = 2$, 3, and 4 unit/km^2 respectively. The results validate that the simulated optimal solutions are close to $\widetilde{K}^*$, illustrating the effectiveness of our analysis in Section V. In addition, the approximated solutions $\widehat{K}^*$ are also close to both the simulated optimal solutions and $\widetilde{K}^*$ values, illustrating the usefulness of the simplified expression (28).

In Fig. 6, we study the optimal cluster size K under different user velocity v, and in Fig. 7, we study the optimal K under different BS intensity λ. In these figures, we set $\alpha = 4$, $W_1 = 1$, and $P = 30$ dBm. In Fig. 6, we additionally set $\lambda = 3$ unit/km^2, and in Fig. 7, we additionally set $v = 36$ km/h. The figures validate that both the analytical values of

$\widetilde{K}^*$ and $\widehat{K}^*$ are close to the simulated optimal K^* under a wide range of user velocities and BS intensities. In addition, since the values of $\widehat{K}^*$ derived in (28) are close to the simulated results, it suggests that the optimal cluster size is approximately inversely proportional to the square of the user velocity v and inversely proportional to the BS intensity λ even if K is not large.

VII. Conclusions

In this work, we provide a theoretical framework to study the handoffs in cooperative wireless networks. Through our proposed stochastic geometric analysis, we capture the irregularly shaped network topology introduced by randomly distributed BSs and user-centric cooperation. The analytical expression for the handoff rate experienced by an active user with arbitrary movement trajectory is derived. Based on this result, we also propose an optimal cluster size formulation considering both the handoff rate and the data rate. We observe that when the common NC-JT scheme is employed, the optimal cluster size can be derived through solving a simple equation in closed form, which is shown to be asymptotically inversely proportional to the square of the user velocity and inversely proportional to the BS intensity. Computer simulation is conducted, validating the correctness and usefulness of our analytical results.

Appendix

A. Proof of Theorem 1

Proof. For simplicity, we define $\mathbf{T}_K^{(1)'}$ as the right-hand side of (3). We show that $\mathbf{T}_K^{(1)'}$ and $\mathbf{T}_K^{(1)}$ are equivalent through the following two steps:

Step 1: $\mathbf{y} \in \mathbf{T}_K^{(1)'} \Rightarrow \mathbf{y} \in \mathbf{T}_K^{(1)}$.

Suppose $\mathbf{y} \in \mathbf{T}_K^{(1)'}$, then $\exists\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{K-1}, \mathbf{x}_K, \mathbf{x}'_K\}$, such that $|\mathbf{z}-\mathbf{y}| \leq |\mathbf{x}_K-\mathbf{y}| = |\mathbf{x}'_K-\mathbf{y}| \leq |\mathbf{x}-\mathbf{y}|, \forall \mathbf{z} \in \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{K-1}\}$ and $\forall \mathbf{x} \in \Phi\backslash\{\mathbf{x}_1, \ldots, \mathbf{x}_{K-1}, \mathbf{x}_K, \mathbf{x}'_K\}$. Let $\mathcal{C}_1 = \{\mathbf{x}_1, \ldots, \mathbf{x}_{K-1}, \mathbf{x}_K\}$, and $\mathcal{C}_2 = \{\mathbf{x}_1, \ldots, \mathbf{x}_{K-1}, \mathbf{x}'_K\}$. Following the definition in (1), we have $\mathbf{y} \in \mathcal{V}(\mathcal{C}_1)$ and $\mathbf{y} \in \mathcal{V}(\mathcal{C}_2)$. Therefore, by the definition of $\mathbf{T}_K^{(1)}$ in (2), $\mathbf{y}$ is at the boundary of $\mathcal{V}(\mathcal{C}_1)$ and $\mathcal{V}(\mathcal{C}_2)$ and thus $\mathbf{y} \in \mathbf{T}_K^{(1)}$.

Step 2: $\mathbf{y} \in \mathbf{T}_K^{(1)} \Rightarrow \mathbf{y} \in \mathbf{T}_K^{(1)'}$.

Suppose $\mathbf{y} \in \mathbf{T}_K^{(1)}$. First, following the definition in (2), $\exists \mathcal{C}'_1$ and $\mathcal{C}'_2$, such that $\mathbf{y} \in \mathcal{V}(\mathcal{C}'_1)$ and $\mathbf{y} \in \mathcal{V}(\mathcal{C}'_2)$. Let $\mathcal{C}'_1 = \{\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_n, \mathbf{u}_{n+1}, \ldots, \mathbf{u}_K\}$, and $\mathcal{C}'_2 = \{\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_n, \mathbf{v}_{n+1}, \ldots, \mathbf{v}_K\}$, where $\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_n$ are the common elements in $\mathcal{C}'_1$ and $\mathcal{C}'_2$. Note that at least one element in $\mathcal{C}'_1$ is different from that in $\mathcal{C}'_2$, thus $n < K$.

In the second step, we compare the distances between $|\mathbf{u}_i - \mathbf{y}|$ and $|\mathbf{v}_j - \mathbf{y}|, \forall i, j \in \{n+1, n+2, \ldots, K\}$. Since $\mathbf{y} \in \mathcal{V}(\mathcal{C}'_1)$, $\mathbf{u}_i \in \mathcal{C}'_1$ and $\mathbf{v}_j \notin \mathcal{C}'_1$, we have $|\mathbf{u}_i - \mathbf{y}| \leq |\mathbf{v}_j - \mathbf{y}|$ according to the definition of $\mathcal{V}(\mathcal{C}'_1)$ in (1). Similarly, since $\mathbf{y} \in \mathcal{V}(\mathcal{C}'_2)$, $\mathbf{v}_j \in \mathcal{C}'_2$ and $\mathbf{u}_i \notin \mathcal{C}'_2$, we have $|\mathbf{v}_j - \mathbf{y}| \leq |\mathbf{u}_i - \mathbf{y}|$. Therefore, we can conclude that $|\mathbf{v}_j - \mathbf{y}| = |\mathbf{u}_i - \mathbf{y}|, \forall i, j \in \{n+1, n+2, \ldots, K\}$. Finally, we have $|\mathbf{z}_k - \mathbf{y}| \leq |\mathbf{v}_j - \mathbf{y}| = |\mathbf{u}_i - \mathbf{y}| \leq |\mathbf{x} - \mathbf{y}|$, $\forall k \in \{1, 2, \ldots, n\}$, $\forall i, j \in \{n+1, n+2, \ldots, K\}$, and $\forall \mathbf{x} \in \Phi\backslash(\mathcal{C}'_1 \bigcup \mathcal{C}'_2)$.

Let $\mathbf{x}_1 = \mathbf{z}_1, \mathbf{x}_2 = \mathbf{z}_2, \ldots, \mathbf{x}_n = \mathbf{z}_n, \mathbf{x}_{n+1} = \mathbf{u}_{n+1}, \ldots, \mathbf{x}_K = \mathbf{u}_K, \mathbf{x}'_K = \mathbf{v}_K$, then we have $|\mathbf{z} - \mathbf{y}| \leq |\mathbf{x}_K - \mathbf{y}| = |\mathbf{x}'_K - \mathbf{y}| \leq |\mathbf{x} - \mathbf{y}|, \forall \mathbf{z} \in \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{K-1}\}$ and $\forall \mathbf{x} \in \Phi\backslash\{\mathbf{x}_1, \ldots, \mathbf{x}_{K-1}, \mathbf{x}_K, \mathbf{x}'_K\}$. Therefore, we have proved that $\mathbf{y} \in \mathbf{T}_K^{(1)'}$.

□

B. Proof of Theorem 2

Proof. Without loss of generality, we assume the Kth closest BS is located at $\mathbf{x}_K = (r_0, 0)$. Note that there are no BSs other than the K BSs located within $\mathcal{B}(\mathbf{0}, r_0)$, where $\mathcal{B}(\mathbf{0}, r_0)$ is defined at the beginning of Section IV.

Following Theorem 1, $\mathbf{0} \in \mathbf{T}_K^{(1)}$ if and only if there is some point $\mathbf{x}'_K$, such that the perpendicular bisector of the line segment $\mathbf{x}_K\mathbf{x}'_K$ passes $\mathbf{0}$. Since $\mathbf{T}_K^{(2)}(\Delta d)$ is the Δd-neighborhood of $\mathbf{T}_K^{(1)}$, $\mathbf{0} \in \mathbf{T}_K^{(2)}(\Delta d)$ if and only if the distance between $\mathbf{0}$ to the perpendicular bisector of the line segment $\mathbf{x}_K\mathbf{x}'_K$ is smaller than Δd. Then, following Case 3 in the proof of Theorem 1 in [3], $\mathbf{0} \in \mathbf{T}_K^{(2)}(\Delta d)$ if and only if there is some $\mathbf{x}'_K$ located within the following ring region, where (r, θ) denotes the polar coordinate in $\mathbb{R}^2$:

$$S(\Delta d) = \left\{ (r,\theta) \middle| r \geq r_0 \text{ and } \left|r^2 - r_0^2\right| < 2\Delta d\sqrt{r_0^2 + r^2 - 2r_0 r\cos\theta} \right\}. \quad (29)$$

The area of $S(\Delta d)$ is

$$|S(\Delta d)|_2 = 8\Delta d r_0 + \mathcal{O}(\Delta d^2). \quad (30)$$

Given the K closest BSs, the point process of the other BSs, denoted as Φ_K^c, is a PPP with intensity 0 in $\mathcal{B}(\mathbf{0}, r_0)$ and intensity λ in $\mathcal{B}_c(\mathbf{0}, r_0)$, due to the strong Markovian property of a PPP. $\mathbb{P}\left(\mathbf{0} \in \mathbf{T}_K^{(2)}(\Delta d) | R_K = r_0\right)$ is equal to the probability that there is at least one point of Φ_K^c in $S(\Delta d)$. Thus we have

$$\begin{aligned} &\mathbb{P}\left(\mathbf{0} \in \mathbf{T}_K^{(2)}(\Delta d) | R_K = r_0\right) \\ =&1 - \exp\left(-\lambda |S(\Delta d)|_2\right) \\ =&8\lambda\Delta d r_0 + \mathcal{O}(\Delta d^2), \end{aligned} \quad (31)$$

which completes the proof. □

C. Proof of Theorem 3

Proof. The probability density function of the distance between the reference user and its Kth closest BS R_K is [33]

$$f_K(r) = \frac{2(\lambda\pi r^2)^K}{r\Gamma(K)} \exp(-\lambda\pi r^2). \quad (32)$$

Then, we have

$$\begin{aligned} &\mathbb{P}(\mathbf{0} \in \mathbf{T}_K^{(2)}(\Delta d)) \\ =&\int_0^\infty \mathbb{P}\left(\mathbf{0} \in \mathbf{T}_K^{(2)}(\Delta d) | R = r_0\right) f_K(r_0) \mathrm{d}r_0 \end{aligned}$$

$$=\frac{8\Gamma\left(\frac{1}{2}+K\right)\sqrt{\lambda}\Delta d}{\Gamma(K)\sqrt{\pi}}+\mathcal{O}(\Delta d^2), \tag{33}$$

which completes the proof. □

D. Proof of Theorem 4

Proof. We aim to show that

$$-\frac{\Delta H(K+1)}{\Delta L(K+1)}+\frac{\Delta H(K)}{\Delta L(K)}>0, \tag{34}$$

which is equivalent to

$$\Delta H(K)\Delta L(K+1)-\Delta L(K)\Delta H(K+1)>0. \tag{35}$$

Following the definitions in (24) and (25), (35) is equivalent to

$$\ln\left(\frac{K+2-\frac{\alpha}{2}}{K+1}\right)-\frac{K+\frac{1}{2}}{K+1}\ln\left(\frac{K+1-\frac{\alpha}{2}}{K}\right)>0. \tag{36}$$

Through Taylor expansion, we have

$$\ln\left(\frac{K+2-\frac{\alpha}{2}}{K+1}\right)=-\sum_{n=1}^{\infty}\frac{\left(\frac{\alpha}{2}-1\right)^n}{n(K+1)^n}, \tag{37}$$

and

$$\frac{K+\frac{1}{2}}{K+1}\ln\left(\frac{K+1-\frac{\alpha}{2}}{K}\right)=-\frac{K+\frac{1}{2}}{K+1}\left(\sum_{n=1}^{\infty}\frac{\left(\frac{\alpha}{2}-1\right)^n}{nK^n}\right). \tag{38}$$

Since $\frac{\alpha}{2}-1>0$ (i.e., $\alpha>2$), $\forall n\geq 1$, we have

$$-\frac{\left(\frac{\alpha}{2}-1\right)^n}{n(K+1)^n}+\frac{K+\frac{1}{2}}{K+1}\left(\frac{\left(\frac{\alpha}{2}-1\right)^n}{nK^n}\right)>0. \tag{39}$$

Therefore, (36) is verified, which completes the proof. □

REFERENCES

[1] J. Andrews, S. Buzzi, W. Choi, S. Hanly, A. Lozano, A. Soong, and J. Zhang, "What will 5G be?" *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.

[2] X. Lin, R. Ganti, P. Fleming, and J. Andrews, "Towards understanding the fundamentals of mobility in cellular networks," *IEEE Trans. on Wireless Communications*, vol. 12, no. 4, pp. 1686–1698, Apr. 2013.

[3] W. Bao and B. Liang, "Stochastic geometric analysis of user mobility in heterogeneous wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 10, pp. 2212–2225, Oct. 2015.

[4] S. Sadr and R. Adve, "Handoff rate and coverage analysis in multi-tier heterogeneous networks," *IEEE Trans. on Wireless Communications*, vol. 14, no. 5, pp. 2626–2638, May 2015.

[5] W. Bao and B. Liang, "Handoff rate analysis in heterogeneous wireless networks with Poisson and Poisson cluster patterns," in *Proc. of ACM MobiHoc*, Hangzhou, China, Jun. 2015.

[6] F. Baccelli and A. Giovanidis, "A stochastic geometry framework for analyzing pairwise-cooperative cellular networks," *IEEE Trans. on Wireless Communications*, vol. 14, no. 2, pp. 794–808, Feb. 2015.

[7] K. Huang and J. Andrews, "An analytical framework for multicell cooperation via stochastic geometry and large deviations," *IEEE Trans. on Information Theory*, vol. 59, no. 4, pp. 2501–2516, Apr. 2013.

[8] S. Akoum and R. W. Heath, "Interference coordination: Random clustering and adaptive limited feedback," *IEEE Trans. on Signal Processing*, vol. 61, no. 7, pp. 1822–1834, Apr. 2013.

[9] A. Giovanidis, L. D. A. Corrales, and L. Decreusefond, "Analyzing interference from static cellular cooperation using the nearest neighbour model," in *Proc. of IEEE Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, Mumbai, India, May 2015.

[10] R. Tanbourgi, S. Singh, J. Andrews, and F. Jondral, "A tractable model for noncoherent joint-transmission base station cooperation," *IEEE Trans. on Wireless Communications*, vol. 13, no. 9, pp. 4959–4973, Jul. 2014.

[11] W. Nie, F.-C. Zheng, X. Wang, S. Jin, and W. Zhang, "Energy efficiency of cross-tier base station cooperation in heterogeneous cellular networks," arXiv:1406.1867 [cs.IT], 2014.

[12] G. Nigam, P. Minero, and M. Haenggi, "Coordinated multipoint joint transmission in heterogeneous networks," *IEEE Trans. on Communications*, vol. 62, no. 11, pp. 4134–4146, Nov. 2014.

[13] C. Li, J. Zhang, M. Haenggi, and K. Letaief, "User-centric intercell interference nulling for downlink small cell networks," *IEEE Trans. on Communications*, vol. 63, no. 4, pp. 1419–1431, Apr. 2015.

[14] N. Lee, D. Morales-Jimenez, A. Lozano, and R. W. Heath, "Spectral efficiency of dynamic coordinated beamforming: A stochastic geometry approach," *IEEE Transactions on Wireless Communications*, vol. 14, no. 1, pp. 230–241, Jan. 2015.

[15] P. Xia, C.-H. Liu, and J. Andrews, "Downlink coordinated multi-point with overhead modeling in heterogeneous cellular networks," *IEEE Trans. on Wireless Communications*, vol. 12, no. 8, pp. 4025–4037, Aug. 2013.

[16] F. Ashtiani, J. Salehi, and M. Aref, "Mobility modeling and analytical solution for spatial traffic distribution in wireless multimedia networks," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 10, pp. 1699 – 1709, Dec. 2003.

[17] Y. Chen, J. Kurose, and D. Towsley, "A mixed queueing network model of mobility in a campus wireless network," in *Proc. of IEEE INFOCOM*, Orlando, FL, Mar. 2012.

[18] A. Farbod and B. Liang, "Structured admission control policies in heterogeneous wireless networks with mesh underlay," in *Proc. of IEEE INFOCOM*, Rio de Janeiro, Brazil, Apr. 2009.

[19] W. Bao and B. Liang, "Insensitivity of user distribution in multicell networks under general mobility and session patterns," *IEEE Trans. on Wireless Communications*, vol. 12, no. 12, pp. 6244–6254, Dec. 2013.

[20] T. S. Rappaport, *Wireless Communications: Principles and Practice*, 2nd ed. Prentice Hall, 2002.

[21] A. Anpalagan and I. Katzela, "Overlaid cellular system design, with cell selection criteria for mobile wireless users," in *Proc. of IEEE Canadian Conference on Electrical and Computer Engineering*, Edmonton, Canada, May 1999.

[22] N. Shenoy and B. Hartpence, "A mobility model for cost analysis in integrated cellular/WLANs," in *Proc. of International Conference on Computer Communications and Networks*, Chicago, IL, Oct. 2004.

[23] A. Hasib and A. Fapojuwo, "Mobility model for heterogeneous wireless networks and its application in common radio resource management," *IET Communications*, vol. 2, no. 9, pp. 1186–1195, Oct. 2008.

[24] M. Haenggi, J. Andrews, F. Baccelli, O. Dousse, and M. Franceschetti, "Stochastic geometry and random graphs for the analysis and design of wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 7, pp. 1029 – 1046, Sep. 2009.

[25] F. Baccelli and B. Blaszczyszyn, "Stochastic geometry and wireless networks, volume 1: Theory," *Foundations and Trends in Networking*, vol. 3, no. 3-4, pp. 249 – 449, 2009.

[26] ——, "Stochastic geometry and wireless networks, volume 2: Applications," *Foundations and Trends in Networking*, vol. 4, no. 1-2, pp. 1–312, 2009.

[27] M. Haenggi, *Stochastic Geometry for Wireless Networks*. Cambridge University Press, 2012.

[28] D. Stoyan, W. Kendall, and J. Mecke, *Stochastic Geometry and Its Applications*, 2nd ed. Wiley, 1995.

[29] H. Federer, *Geometric Measure Theory*. Springer, 1969.

[30] 3GPP-TR-36.819, "Coordinated multi-point operation for LTE physical layer aspects," Sep. 2013.

[31] A. M. Ibrahim, T. ElBatt, and A. El-Keyi, "Coverage probability analysis for wireless networks using repulsive point processes," in *Proc. of IEEE PIMRC*, London, UK, Sep. 2013.

[32] A. Guo and M. Haenggi, "Spatial stochastic models and metrics for the structure of base stations in cellular networks," *IEEE Trans. on Wireless Communications*, vol. 12, no. 11, pp. 5800–5812, Nov. 2013.

[33] M. Haenggi, "On distances in uniformly random networks," *IEEE Trans. on Information Theory*, vol. 51, no. 10, pp. 3584–3586, Oct. 2005.