

# Stochastic Geometric Analysis of User Mobility in Heterogeneous Wireless Networks

Wei Bao, *Student Member, IEEE* and Ben Liang, *Senior Member, IEEE*

**Abstract**—Horizontal and vertical handoffs are important ramifications of user mobility in multi-tier heterogeneous wireless networks. They directly affect the signaling overhead and quality of calls in the system. However, they are difficult to analyze due to the irregularly shaped network topologies introduced by multiple tiers of cells. In this work, a stochastic geometric analysis framework on user mobility is proposed, to capture the spatial randomness and various scales of cell sizes in different tiers. We derive theoretical expressions for the rates of all handoff types experienced by an active user with arbitrary movement trajectory. We also derive the downlink data rate of the user given the set of cell tiers that it is willing to use. Based on these results, we provide guidelines for optimal tier selection under different user velocity, taking both the handoff rates and the data rate into consideration. Empirical study using real user mobility trace data and extensive simulation are conducted, demonstrating the correctness and usefulness of our analysis.

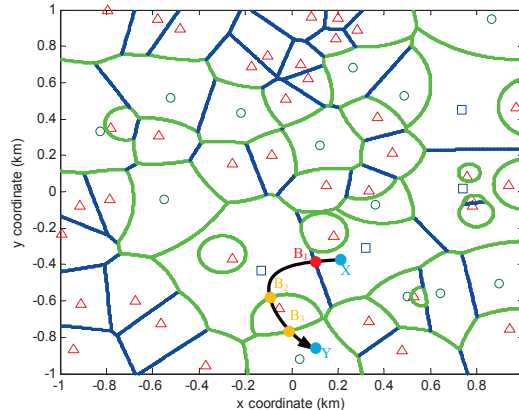
**Index Terms**—Heterogeneous wireless network, mobility, hand-off, stochastic geometry, analytic geometry

## I. INTRODUCTION

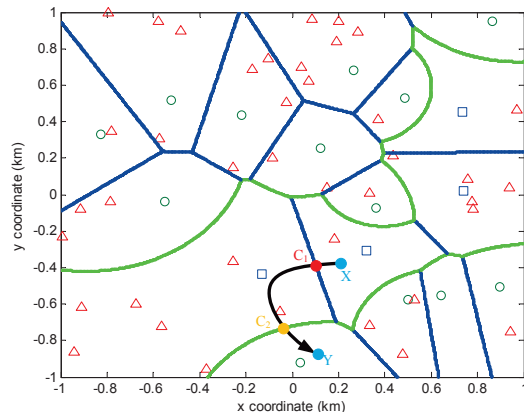
TRADITIONAL single-tier macro-cellular networks provide wide coverage for mobile user equipments (UEs), but they are insufficient to satisfy the exploding demand for high bandwidth access driven by modern mobile traffic, such as multimedia transmissions and cloud computing tasks. One effective means to increase network capacity is to provide more serving stations within a geographical area, i.e., installing a diverse set of small-cells such as femtocells [2] and WiFi hotspots [3], overlaying the macrocells, to form a multi-tier heterogeneous wireless network (HWN). Each small-cell is equipped with a shorter-range and lower-cost base station (BS) or access point (AP), to provide nearby UEs with higher-bandwidth network access with lower power usage, and to offload data traffic from macrocells. The commercial deployment of small-cells has attracted increasing attention in recent years. For example, AT&T Inc. now supplies a femtocell product [4], and it has also deployed WiFi APs in a number of metropolitan areas with dense population [5].

In the presence of multiple tiers of cells, however, mobile UEs may experience internetworking issues among different tiers. In particular, vertical handoffs (i.e., handoffs made between two BSs in different tiers) are introduced [6]. Compared with horizontal handoffs (i.e., handoffs made between two BSs

The authors are affiliated with the Department of Electrical and Computer Engineering, University of Toronto, 10 King's College Road, Toronto, Ontario, Canada (email: {wbao, liang}@comm.utoronto.ca). This work has been supported in part by grants from Bell Canada and the Natural Sciences and Engineering Research Council (NSERC) of Canada. A part of this work has appeared as [1], which contains only the derivation of handoff rates.



(a) A UE starts a call at  $X$  and terminates it at  $Y$ . It experiences one horizontal handoff at  $B_1$  and two vertical handoffs at  $B_2$  and  $B_3$ .



(b) Same BS locations and UE trajectory. Tier-3 BSs are not accessed. The same UE experiences one horizontal handoff at  $C_1$  and one vertical handoff at  $C_2$ .

Fig. 1. An example of a three-tier HWN and selective tier association. Tier-1, 2, and 3 BSs are represented by squares, circles, and triangles respectively; blue curves show intra-tier cell boundaries; green curves show inter-tier cell boundaries.

in the same tier), vertical handoffs have a more complicated impact on both the UEs and the overall system. Additional risks are present during channel setup and tear down when a vertical handoff is made, such as (1) extra traffic latency; (2) additional network signaling; (3) more UE power consumption due to simultaneously active network interface to multiple tiers; and (4) higher risk in call drops or degraded quality of service (QoS) caused by the lack of radio resource after handoffs. Furthermore, vertical handoffs may be classified into inter-RAT (radio access technology) handoffs (e.g., handoffs made between LTE access and WiFi access) and intra-RAT

handoffs, where the former could cause worse performance degradation on UEs [7].

The handoff rate is defined as the expected number of handoffs experienced by one UE per unit time, which directly affects the signaling overhead in the system and QoS of UEs. As a prerequisite to performance evaluation and system design in HWNs, it is essential to quantify the rates of different handoff types. However, a study on handoff rates in HWNs will inevitably be challenged by the irregularly shaped multi-tier network topologies introduced by the small-cell structure. An example topology with three tiers of BSs is shown in Fig. 1. First, BSs are spread irregularly, sometimes in an anywhere plug-and-play manner, leading to a high level of spatial randomness. Second, different tiers of cells are equipped with BSs communicating at different power levels, causing various scales of cell sizes. As a consequence, it is difficult to characterize the cell boundaries and to track boundary crossings made by UEs (i.e., handoffs) in the system. Few previous works have resolved the above challenges.

Characterizing the handoff rates provides important guidelines for system design. One main design concern is the tradeoff between the handoff rates and the data rate. As shown in the example in Fig. 1(a), a UE starts a call at  $X$  and terminates it at  $Y$ . By choosing to access all of tier-1, 2, and 3 cells, it experiences one horizontal handoff at  $B_1$  and two vertical handoffs at  $B_2$  and  $B_3$ . As shown in Fig. 1(b), by choosing to access tier-1 and 2 cells only, it experiences one horizontal handoff at  $C_1$  and only one vertical handoff at  $C_2$ . However, in the latter case, the UE misses the opportunity to access a high-bandwidth tier-3 BS. Thus, a UE could choose to access high-bandwidth small-cell BSs to improve data rate, but this may also lead to more frequent vertical handoffs, which potentially deteriorates the service quality. Therefore, in this paper, we are also motivated to optimize the UE tier selection scheme, by taking both the handoff rates and the data rate into consideration.

In this work, we contribute to user mobility modeling and network access optimization in HWNs by providing new technical tools to quantify the rates of horizontal and vertical handoffs, under random multi-tier BSs, arbitrary user movement trajectory, and flexible user-BS association. A new stochastic geometric analysis framework on user mobility is proposed. In this framework, different tiers of BSs are modeled as Poisson point processes (PPPs) to capture their spatial randomness. To model flexible scaling of cell sizes in different tiers, we consider the biased user association scheme [8], [9], [10], [11], in which each tier of BSs is assigned an association bias value, and a UE is associated with a BS that provides the largest biased received power. Through stochastic and analytic geometric analysis, we derive exact expressions for the rates of all handoff types experienced by an active UE with arbitrary movement trajectory. In addition, as a study on the application of the above handoff rate analysis, after calculating the downlink data rate of an active UE given the set of BS tiers that it chooses to associate with, we further study optimal tier selection by the UE, considering both the handoff rates and the data rate.

We confirm our theoretical analysis through an empirical

study using the Yonsei Trace [12]. The trace provides a large data set, accumulating fine-grained mobility data from commercial mobile phones in an 8-month period. Numerical studies using the empirical trace data set, together with further simulation, demonstrate the correctness and usefulness of our analytical conclusions.

The rest of the paper is organized as follows. In Section II, we discuss the relation between our work and prior works. In Section III, we describe the system model. In Section IV, we present our contributions in handoff rates derivations. In Section V, we discuss the optimal tier selection scheme considering both the handoff rates and the data rate. In Section VI, we present empirical study with the Yonsei Trace as well as simulation. Finally, conclusions are given in Section VII.

## II. RELATED WORKS

Classical mobility modeling and management techniques are limited to homogeneous single-tier networks [13], [14], which do not concern vertical handoffs or tier selection introduced by HWNs. In the following, we present prior related works mainly focusing on HWNs.

### A. Mobility Modeling Based on Queueing Systems

One common category of previous works employ queueing systems to model HWNs. In this case, cells are modeled as queues, active users are modeled as units in the queues, and handoffs correspond to unit transfers among queues. Ghosh *et al.* [15] studied the single-cell scenario using an  $M/G/\infty$  queue. Kirsal *et al.* [16] studied one WLAN cell overlaying one 3G cell, and a two-queue model is proposed accordingly. For multicell scenarios, queueing network models have been employed in [17], [18], [19], [20], [21], [22]. However, none of these works explicitly modeled the geometric patterns of cell shapes in heterogeneous networks.

### B. Geometric Pattern Study

In order to characterize the geometric patterns of network topologies, a second category of works model the shape of cells, mostly in non-random regular grids. Zonoozi and Dassanayake [23] modeled a one-tier cellular network as a hexagonal grid. Anpalagan and Katzela [24] studied a two-tier network by modeling small-cells as hexagons, and each macrocell as a cluster of neighbouring small-cells. Shenoy and Hartpence [25] studied a two-tier network by modeling WLAN small-cells as squares, and macrocells as larger squares, each covering  $5 \times 5$  WLAN cells. Hasib and Fapojuwo [26] studied a two-tier cellular network including one hexagonal macrocell and a predetermined  $N$  circular microcells. Lin *et al.* [27] conducted a pioneering study on the user mobility in one-tier macro cellular network considering randomly distributed BSs. Macrocells were modeled as a standard Poisson Voronoi. However, in [27], the authors did not consider multi-tier BSs with different scales of cell sizes. To the best of our knowledge, ours is the first work studying user mobility in *multi-tier* HWNs that captures their *random* geometric patterns.

### C. Real-world Trace Study

Another important category of related works employ empirical traces to investigate user mobility. Kotz *et al.* [28], [29] studied user mobility patterns on the Dartmouth campus. McNett and Voelker [30] characterized the mobility and access patterns of hand-held PDA users on the UCSD campus, and a campus waypoint model was proposed to characterize the trace. Halepovic and Williamson [31] studied mobility parameters such as the number of calls initiated per user, call inter-arrival time, and the number of cell sites visited per user, based on data traffic traces of a regional CDMA2000 cellular network. Rhee *et al.* [32] concluded that human walk patterns contain statistically similar features observed in Levy walks, based on a large daily GPS trace set accumulated in 5 different places in US and Korea. In [33], empirical study on spatial and temporal mobility patterns of the Yonsei Trace [12] was conducted, in order to predict users' future position precisely. Ficek and Kencl [34] proposed inter-call mobility model to locate users' position between calls based on the trace accumulated in a trip between San Jose and San Francisco. Baumann *et al.* [35] predicted user arrival and residence times in the system through extracting important parameters from the trace accumulated by Nokia Research.

These works based on real-world traces study are practically valuable for system evaluation and design. However, they are insufficient to provide in-depth analytical modeling of handoff and data rates. In our work, we use the Yonsei Trace [12], [33] to demonstrate the correctness and usefulness of our theoretical results.

### D. Handoff and Association Decision Algorithms

Orthogonal to the scope of this work, there is a large body of previous works that study handoff timing algorithms, without considering the random geometric patterns of UEs and BSs. One type of handoff decision algorithms employ a threshold comparison of one or several specific metrics, (e.g., received signal strength, network loading, bandwidth, and so on) to derive handoff decisions [36], [37], [38], [39]. Another type uses dynamic programming (DP) [40] or artificial intelligence techniques (e.g., fuzzy logic [41]) to improve the effectiveness of handoff procedures. In our work, we do not explicitly specify the handoff timing. Instead, we derive handoff rates and tier-level association decisions through stochastic geometric analysis.

Stochastic geometric analysis has been employed to derive tier-association decisions in HWNs [8], [9], [10], [11]. These works assume that the UEs are randomly placed but stationary. They focus on average performance metrics such as the mean data throughput or outage probability. They ignore the movement of UEs and the effect of handoffs, which are the main focus of our work.

## III. SYSTEM MODEL

### A. Multi-tier Network

We consider a HWN with spatially randomly distributed  $K$  tiers of BSs. Let  $\mathcal{K} = \{1, 2, \dots, K\}$ . In order to characterize

the random spatial patterns of BSs, we use the conventional assumption that each tier of BSs independently form a homogeneous Poisson point process (PPP) in two-dimensional Euclidean space  $\mathbb{R}^2$  [8], [9], [10], [42], [43], [44], [45]. Let  $\Phi_k$  denote the PPP corresponding to tier- $k$  BSs, and let  $\lambda_k$  be its intensity.

### B. Biased User Association

Different tiers of BSs transmit at different power levels. Let  $P_k$  be the transmission power of tier- $k$  BSs, which is a given parameter. If  $P_t(\mathbf{x})$ , for  $P_t(\mathbf{x}) \in \{P_1, P_2, \dots, P_K\}$ , is the transmission power from a BS at  $\mathbf{x}$  and  $P_r(\mathbf{y})$  is the received power at  $\mathbf{y}$ , we have  $P_r(\mathbf{y}) = \frac{P_t(\mathbf{x})h_{\mathbf{x},\mathbf{y}}}{\alpha|\mathbf{x}-\mathbf{y}|^\gamma}$ , where  $\alpha|\mathbf{x}-\mathbf{y}|^\gamma$  is the propagation loss function with  $\gamma > 2$ , and  $h_{\mathbf{x},\mathbf{y}}$  is the fast fading term. Corresponding to common Rayleigh fading with power normalization,  $h_{\mathbf{x},\mathbf{y}}$  is independently exponentially distributed with unit mean.

In order to capture various scales of different cell sizes, biased user association is considered [8], [9], [10]. Given that a UE is located at  $\mathbf{y}$ , it associates itself with the BS that provides the maximum *biased received power* as follows:

$$\mathcal{BS}(\mathbf{y}) = \arg \max_{\mathbf{x} \in \Phi_k, \forall k} B_k P_k |\mathbf{x} - \mathbf{y}|^{-\gamma}, \quad (1)$$

where  $\mathcal{BS}(\mathbf{y})$  denotes the location of the BS chosen for the UE,  $P_k|\mathbf{x} - \mathbf{y}|^{-\gamma}$  is the received power from a tier- $k$  BS located at  $\mathbf{x}$ , and  $B_k$  is the association bias, indicating the received power preference of UEs toward tier- $k$  BSs.  $B_k$  may be different in different tiers, mainly because (1) different radio access technologies may require different received power levels, and (2) some tiers could be assigned larger values of  $B_k$ , in order to offload data traffic from other tiers. As a consequence, the resultant cell splitting forms a generalized Dirichlet tessellation, or weighted Poisson Voronoi [46], an example of which is shown in Fig. 1(a). Let  $\mathbf{T}^{(1)}$  denote the overall cell boundaries, and let  $\mathbf{T}_{kj}^{(1)}$  denote the boundaries of tier- $k$  cells and tier- $j$  cells, which is also referred to as *type  $k$ - $j$  cell boundaries* in this paper. Note that  $\mathbf{T}_{kj}^{(1)}$  and  $\mathbf{T}_{jk}^{(1)}$  are equivalent (i.e., type  $k$ - $j$  cell boundaries and type  $j$ - $k$  cell boundaries are equivalent).

Note that for  $B_1, B_2, \dots, B_K$ , their effects remain the same if we multiply all of them by the same positive constant. For presentation convenience, we define  $\beta_{kj} = \left(\frac{P_k B_k}{P_j B_j}\right)^{1/\gamma}$ . Clearly,  $\beta_{kj} = \frac{1}{\beta_{jk}}$ .

Let  $A_k$  denote the probability that a UE associates itself with a tier- $k$  BS. As derived in [8], we have

$$A_k = \frac{\lambda_k (P_k B_k)^{\frac{2}{\gamma}}}{\sum_{j=1}^K \lambda_j (P_j B_j)^{\frac{2}{\gamma}}}. \quad (2)$$

### C. UE Trajectory and Handoff Rate

We aim to study the rates of all handoff types of some active UE moving in the network. Let  $\mathcal{T}_0$  denote the trajectory of the UE, which is of finite length. The number of handoffs the UE experiences is equal to the number of intersections of  $\mathcal{T}_0$  and  $\mathbf{T}^{(1)}$ , which is denoted by  $\mathcal{N}(\mathcal{T}_0, \mathbf{T}^{(1)})$ . In this paper, a handoff made from a tier- $k$  cell to a tier- $j$  cell is called a

type  $k$ - $j$  handoff. The number of type  $k$ - $j$  handoffs is denoted by  $\mathcal{N}_{kj}(\mathcal{T}_0, \mathbf{T}_{kj}^{(1)})$ .

If  $j \neq k$ , a type  $k$ - $j$  (vertical) handoff is **not** equivalent to a type  $j$ - $k$  handoff. When the UE crosses type  $k$ - $j$  boundary, either a type  $k$ - $j$  or a type  $j$ - $k$  handoff is made, depending on the moving direction. Thus, the number of type  $k$ - $j$  plus type  $j$ - $k$  handoffs is equal to the number of intersections of  $\mathcal{T}_0$  and  $\mathbf{T}_{kj}^{(1)}$ , which is denoted by  $\mathcal{N}(\mathcal{T}_0, \mathbf{T}_{kj}^{(1)})$ . In other words, we have  $\mathcal{N}(\mathcal{T}_0, \mathbf{T}_{kj}^{(1)}) = \mathcal{N}_{kj}(\mathcal{T}_0, \mathbf{T}_{kj}^{(1)}) + \mathcal{N}_{jk}(\mathcal{T}_0, \mathbf{T}_{kj}^{(1)})$ .

If  $j = k$ ,  $\mathcal{N}(\mathcal{T}_0, \mathbf{T}_{kk}^{(1)}) = \mathcal{N}_{kk}(\mathcal{T}_0, \mathbf{T}_{kk}^{(1)})$  indicates the number of type  $k$ - $k$  (horizontal) handoffs.

In Section IV, we aim to study the rates of all handoff types, which correspond to the expected numbers of handoffs experienced by the active UE per unit time.

#### IV. HANDOFF RATE ANALYSIS IN MULTI-TIER HWNS

The proposed analysis of handoff rates consists of a progressive sequence of four components, which are described in the following subsections.

##### A. Length Intensity of Cell Boundaries

Handoffs occur at the intersections of the active UE's trajectory with cell boundaries. In order to track the number of intersections, we need to first study the length intensity of cell boundaries  $\mathbf{T}^{(1)}$  (resp.  $\mathbf{T}_{kj}^{(1)}$ ), which is defined as the expected length of  $\mathbf{T}^{(1)}$  (resp.  $\mathbf{T}_{kj}^{(1)}$ ) in a unit square. Higher length intensity of cell boundaries leads to greater opportunities for boundary crossing, and thus higher handoff rates.

The cell boundaries  $\mathbf{T}^{(1)}$  is a fiber process [47] generated by  $\Phi_1, \Phi_2, \dots, \Phi_K$ .  $\mathbf{T}^{(1)}$  also corresponds to the set of points on  $\mathbb{R}^2$ , where a same biased power level is received from two nearby BSs, and this biased received power level is no less than those from any other BSs. Mathematically, we have

$$\mathbf{T}^{(1)} = \left\{ \mathbf{x} \mid \forall k, j \in \mathcal{K}, \exists \mathbf{x}_1 \in \Phi_k, \mathbf{x}_2 \in \Phi_j, \mathbf{x}_1 \neq \mathbf{x}_2, \right. \\ \left. \text{s.t. } P_r = \frac{P_k B_k}{|\mathbf{x}_1 - \mathbf{x}|^\gamma} = \frac{P_j B_j}{|\mathbf{x}_2 - \mathbf{x}|^\gamma}, \text{ and} \right. \\ \left. \forall i \in \mathcal{K}, \mathbf{y} \in \Phi_i, P_r \geq \frac{P_i B_i}{|\mathbf{y} - \mathbf{x}|^\gamma} \right\}. \quad (3)$$

Similarly,  $\mathbf{T}_{kj}^{(1)}$  can be expressed as

$$\mathbf{T}_{kj}^{(1)} = \left\{ \mathbf{x} \mid \exists \mathbf{x}_1 \in \Phi_k, \mathbf{x}_2 \in \Phi_j, \mathbf{x}_1 \neq \mathbf{x}_2, \right. \\ \left. \text{s.t. } P_r = \frac{P_k B_k}{|\mathbf{x}_1 - \mathbf{x}|^\gamma} = \frac{P_j B_j}{|\mathbf{x}_2 - \mathbf{x}|^\gamma}, \text{ and} \right. \\ \left. \forall i \in \mathcal{K}, \mathbf{y} \in \Phi_i, P_r \geq \frac{P_i B_i}{|\mathbf{y} - \mathbf{x}|^\gamma} \right\}. \quad (4)$$

Note that  $\bigcup_{k=1}^K \bigcup_{j=k}^K \mathbf{T}_{kj}^{(1)} = \mathbf{T}^{(1)}$ .

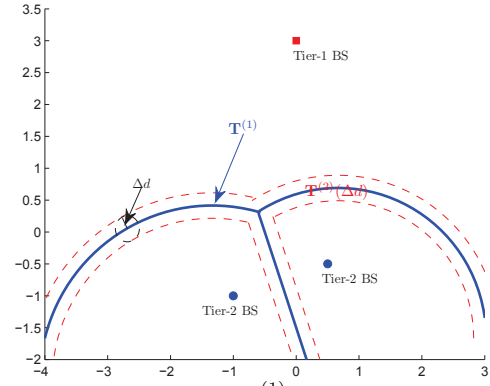


Fig. 2. The blue bold curves show  $\mathbf{T}^{(1)}$ ; and the region within red dashed curves shows  $\mathbf{T}^{(2)}(\Delta d)$ .

Let  $\mu_1(\mathbf{T}^{(1)})$  denote the length intensity of  $\mathbf{T}^{(1)}$ , which is the expected length of  $\mathbf{T}^{(1)}$  in a unit square<sup>1</sup> [47]:

$$\mu_1(\mathbf{T}^{(1)}) = \mathbb{E} \left( \left| \mathbf{T}^{(1)} \cap [0, 1]^2 \right|_1 \right), \quad (5)$$

where  $|L|_1$  denotes the length of  $L$  (i.e., one-dimensional Lebesgue measure of  $L$ ). Similarly, let  $\mu_1(\mathbf{T}_{kj}^{(1)})$  denote the length intensity of  $\mathbf{T}_{kj}^{(1)}$ :

$$\mu_1(\mathbf{T}_{kj}^{(1)}) = \mathbb{E} \left( \left| \mathbf{T}_{kj}^{(1)} \cap [0, 1]^2 \right|_1 \right). \quad (6)$$

Note that we have  $\mu_1(\mathbf{T}^{(1)}) = \sum_{k=1}^K \sum_{j=k}^K \mu_1(\mathbf{T}_{kj}^{(1)})$ .

##### B. $\Delta d$ -Extended Cell Boundaries

It is difficult to directly quantify the one-dimensional measures  $\mu_1(\mathbf{T}^{(1)})$  and  $\mu_1(\mathbf{T}_{kj}^{(1)})$  on the two-dimensional plane. Instead, we first introduce the  $\Delta d$ -extended cell boundaries, which extends the one-dimensional measures to two-dimensional measures.

The  $\Delta d$ -extended cell boundaries of  $\mathbf{T}^{(1)}$ , denoted by  $\mathbf{T}^{(2)}(\Delta d)$  is defined as

$$\mathbf{T}^{(2)}(\Delta d) = \left\{ \mathbf{x} \mid \exists \mathbf{y} \in \mathbf{T}^{(1)}, \text{ s.t. } |\mathbf{x} - \mathbf{y}| < \Delta d \right\}. \quad (7)$$

In other words,  $\mathbf{T}^{(2)}(\Delta d)$  is the  $\Delta d$ -neighbourhood of  $\mathbf{T}^{(1)}$ . A point is in  $\mathbf{T}^{(2)}(\Delta d)$  iff its (shortest) distance to  $\mathbf{T}^{(1)}$  is less than  $\Delta d$ , as shown in Fig. 2. Similarly, we define  $\mathbf{T}_{kj}^{(2)}(\Delta d)$  as the  $\Delta d$ -extended cell boundaries of  $\mathbf{T}_{kj}^{(1)}$  (i.e.,  $\Delta d$ -neighbourhood of  $\mathbf{T}_{kj}^{(1)}$ ):

$$\mathbf{T}_{kj}^{(2)}(\Delta d) = \left\{ \mathbf{x} \mid \exists \mathbf{y} \in \mathbf{T}_{kj}^{(1)}, \text{ s.t. } |\mathbf{x} - \mathbf{y}| < \Delta d \right\}. \quad (8)$$

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

$$\mu_2(\mathbf{T}^{(2)}(\Delta d)) = \mathbb{E} \left( \left| \mathbf{T}^{(2)}(\Delta d) \cap [0, 1]^2 \right| \right), \quad (9)$$

where  $|S|$  denotes the area of  $S$  (i.e., two-dimensional Lebesgue measure of  $S$ ). Similarly, the area intensity of  $\mathbf{T}_{kj}^{(2)}(\Delta d)$  is

$$\mu_2(\mathbf{T}_{kj}^{(2)}(\Delta d)) = \mathbb{E} \left( \left| \mathbf{T}_{kj}^{(2)}(\Delta d) \cap [0, 1]^2 \right| \right). \quad (10)$$

<sup>1</sup>Because  $\Phi_1, \dots, \Phi_K$  are stationary,  $\mathbf{T}^{(1)}$  is also stationary, and thus the unit square could be arbitrarily picked on  $\mathbb{R}^2$ .

Because  $\Phi_1, \Phi_2, \dots, \Phi_K$  are stationary and isotropic,  $\mathbf{T}^{(2)}(\Delta d)$  and  $\mathbf{T}_{kj}^{(2)}(\Delta d)$  are also stationary and isotropic. As a result, given a reference UE located at  $\mathbf{0}$ , the area intensity of  $\mathbf{T}^{(2)}(\Delta d)$  (resp.  $\mathbf{T}_{kj}^{(2)}(\Delta d)$ ) is equal to the probability that the reference UE at  $\mathbf{0}$  is in  $\mathbf{T}^{(2)}(\Delta d)$  (resp.  $\mathbf{T}_{kj}^{(2)}(\Delta d)$ ).

$$\mu_2(\mathbf{T}^{(2)}(\Delta d)) = \mathbb{P}(\mathbf{0} \in \mathbf{T}^{(2)}(\Delta d)), \quad (11)$$

$$\mu_2(\mathbf{T}_{kj}^{(2)}(\Delta d)) = \mathbb{P}(\mathbf{0} \in \mathbf{T}_{kj}^{(2)}(\Delta d)). \quad (12)$$

We observe that the probabilities in (11) and (12) are analytically tractable, which will be presented in the next subsection.

### C. Derivations of the Area Intensities

In this subsection, we present the derivations of  $\mathbb{P}(\mathbf{0} \in \mathbf{T}^{(2)}(\Delta d))$  and  $\mathbb{P}(\mathbf{0} \in \mathbf{T}_{kj}^{(2)}(\Delta d))$ . First, we study the probability that the reference UE at  $\mathbf{0}$  is in  $\mathbf{T}_{kj}^{(2)}(\Delta d)$ , given that it is associated with a tier- $k$  BS at a distance of  $r_0$  from it. By employing both analytic geometric and stochastic geometric tools, we derive the following theorem:

*Theorem 1:* Suppose the reference UE is located at  $\mathbf{0}$ , it is associated with a tier- $k$  BS, and their distance is  $R$ . The conditional probability that  $\mathbf{0} \in \mathbf{T}_{kj}^{(2)}(\Delta d)$  given  $R = r_0$  is

$$\mathbb{P}\left(\mathbf{0} \in \mathbf{T}_{kj}^{(2)}(\Delta d) | R = r_0, \text{tier} = k\right) = 1 - \exp\left(-2\lambda_j \Delta d r_0 \mathcal{F}(\beta_{kj}) + \mathcal{O}(\Delta d^2)\right), \quad (13)$$

where

$$\mathcal{F}(\beta) \triangleq \frac{1}{\beta^2} \int_0^\pi \sqrt{(\beta^2 + 1) - 2\beta \cos(\theta)} d\theta. \quad (14)$$

See Appendix VIII-A for the proof.

Second, through stochastic geometric tools and deconditioning on  $R$ , we can derive the unconditioned probabilities that the reference UE at  $\mathbf{0}$  is in  $\mathbf{T}^{(2)}(\Delta d)$  and in  $\mathbf{T}_{kj}^{(2)}(\Delta d)$ :

*Theorem 2:* The area intensities of  $\mathbf{T}^{(2)}(\Delta d)$  and  $\mathbf{T}_{kj}^{(2)}(\Delta d)$  are:

(a)

$$\begin{aligned} \mu_2(\mathbf{T}^{(2)}(\Delta d)) &= \mathbb{P}(\mathbf{0} \in \mathbf{T}^{(2)}(\Delta d)) \\ &= \sum_{k=1}^K \frac{\lambda_k \left( \sum_{i=1}^K \lambda_i \Delta d \mathcal{F}(\beta_{ki}) \right)}{\left( \sum_{i=1}^K \lambda_i \beta_{ik}^2 \right)^{\frac{3}{2}}} + \mathcal{O}(\Delta d^2). \end{aligned} \quad (15)$$

(b)

$$\begin{aligned} \mu_2(\mathbf{T}_{kj}^{(2)}(\Delta d)) &= \mathbb{P}\left(\mathbf{0} \in \mathbf{T}_{kj}^{(2)}(\Delta d)\right) \\ &= \begin{cases} \frac{\lambda_k \lambda_j \Delta d \mathcal{F}(\beta_{kj})}{\left( \sum_{i=1}^K \lambda_i \beta_{ik}^2 \right)^{\frac{3}{2}}} + \frac{\lambda_j \lambda_k \Delta d \mathcal{F}(\beta_{jk})}{\left( \sum_{i=1}^K \lambda_i \beta_{ij}^2 \right)^{\frac{3}{2}}} + \mathcal{O}(\Delta d^2) & \text{if } k \neq j, \\ \frac{\lambda_k^2 \Delta d \mathcal{F}(1)}{\left( \sum_{i=1}^K \lambda_i \beta_{ik}^2 \right)^{\frac{3}{2}}} + \mathcal{O}(\Delta d^2) & \text{if } k = j. \end{cases} \end{aligned} \quad (16)$$

See Appendix VIII-B for the proof.

### D. From Area Intensities to Handoff Rates

In this subsection, we derive handoff rates from area intensities derived in Theorem 2. This involves two steps: (1) from area intensities  $\mu_2(\mathbf{T}^{(2)}(\Delta d))$  and  $\mu_2(\mathbf{T}_{kj}^{(2)}(\Delta d))$  to length intensities  $\mu_1(\mathbf{T}^{(1)})$  and  $\mu_1(\mathbf{T}_{kj}^{(1)})$ , and (2) from length intensities to handoff rates.

First, we derive the length intensity  $\mu_1(\mathbf{T}^{(1)})$  (resp.  $\mu_1(\mathbf{T}_{kj}^{(1)})$ ) from the area intensity  $\mu_2(\mathbf{T}^{(2)}(\Delta d))$  (resp.  $\mu_2(\mathbf{T}_{kj}^{(2)}(\Delta d))$ ) by taking  $\Delta d \rightarrow 0$ . We have

*Theorem 3:* The length intensities of  $\mathbf{T}^{(1)}$  and  $\mathbf{T}_{kj}^{(1)}$  can be computed as follows:

(a)

$$\mu_1(\mathbf{T}^{(1)}) = \sum_{k=1}^K \frac{\lambda_k \left( \sum_{i=1}^K \lambda_i \mathcal{F}(\beta_{ki}) \right)}{2 \left( \sum_{i=1}^K \lambda_i \beta_{ik}^2 \right)^{\frac{3}{2}}}. \quad (17)$$

(b)

$$\mu_1(\mathbf{T}_{kj}^{(1)}) = \begin{cases} \frac{\lambda_k \lambda_j \mathcal{F}(\beta_{kj})}{2 \left( \sum_{i=1}^K \lambda_i \beta_{ik}^2 \right)^{\frac{3}{2}}} + \frac{\lambda_j \lambda_k \mathcal{F}(\beta_{jk})}{2 \left( \sum_{i=1}^K \lambda_i \beta_{ij}^2 \right)^{\frac{3}{2}}} & \text{if } k \neq j, \\ \frac{\lambda_k^2 \mathcal{F}(1)}{2 \left( \sum_{i=1}^K \lambda_i \beta_{ik}^2 \right)^{\frac{3}{2}}} & \text{if } k = j. \end{cases} \quad (18)$$

See Appendix VIII-C for the proof.

*Remark 1:* Note that, if we consider the single-tier case by taking  $K = 1$ , we have  $\mathcal{F}(1) = 4$ , and  $\mu_1(\mathbf{T}^{(1)}) = \mu_1(\mathbf{T}_{11}^{(1)}) = 2\sqrt{\lambda_1}$ . This matches the length intensity of a standard Poisson Voronoi. See Section 10.6 of [47].

Second, we can derive the expected number of handoffs of an active UE as follows:

*Theorem 4:* Let  $\mathcal{T}_0$  denote an arbitrary UE's trajectory on  $\mathbb{R}^2$  with length  $|\mathcal{T}_0|_1$ . Then, the expected number of intersections of  $\mathcal{T}_0$  and  $\mathbf{T}^{(1)}$  (resp.  $\mathbf{T}_{kj}^{(1)}$ ) are

$$\mathbb{E}\left(\mathcal{N}(\mathcal{T}_0, \mathbf{T}^{(1)})\right) = \frac{2}{\pi} \mu_1(\mathbf{T}^{(1)}) |\mathcal{T}_0|_1, \quad (19)$$

$$\mathbb{E}\left(\mathcal{N}(\mathcal{T}_0, \mathbf{T}_{kj}^{(1)})\right) = \frac{2}{\pi} \mu_1(\mathbf{T}_{kj}^{(1)}) |\mathcal{T}_0|_1, \quad (20)$$

and the expected number of type  $k$ - $j$  handoffs are

$$\mathbb{E}\left(\mathcal{N}_{kj}(\mathcal{T}_0, \mathbf{T}_{kj}^{(1)})\right) = \begin{cases} \frac{1}{2} \mathbb{E}\left(\mathcal{N}(\mathcal{T}_0, \mathbf{T}_{kj}^{(1)})\right) & \text{if } k \neq j, \\ \mathbb{E}\left(\mathcal{N}(\mathcal{T}_0, \mathbf{T}_{kj}^{(1)})\right) & \text{if } k = j. \end{cases} \quad (21)$$

*Proof:*  $\mathbf{T}^{(1)}$  and  $\mathbf{T}_{kj}^{(1)}$  are stationary and isotropic fibre processes with length intensity  $\mu_1(\mathbf{T}^{(1)})$  and  $\mu_1(\mathbf{T}_{kj}^{(1)})$  respectively. The proof follows the conclusions in Section 9.3 of [47]. ■

Note that the expected number of type  $k$ - $j$  handoffs is the same as the expected number of type  $j$ - $k$  handoffs, both of which are equal to half of  $\mathbb{E}\left(\mathcal{N}(\mathcal{T}_0, \mathbf{T}_{kj}^{(1)})\right)$ .

Let  $v$  denote the instantaneous velocity of an active UE,  $H(v)$  denote its overall handoff rate (i.e., sum handoff rate of all types), and  $H_{kj}(v)$  denote its type  $k$ - $j$  handoff rate. Then we have the following Corollary from Theorem 4:

*Corollary 1:*

$$H(v) = \frac{2}{\pi} \mu_1(\mathbf{T}^{(1)}) v, \quad (22)$$

$$H_{kj}(v) = \begin{cases} \frac{1}{\pi} \mu_1(\mathbf{T}_{kj}^{(1)})v & \text{if } k \neq j, \\ \frac{2}{\pi} \mu_1(\mathbf{T}_{kj}^{(1)})v & \text{if } k = j. \end{cases} \quad (23)$$

Note that the above handoff rates are instantaneous rates. Our analysis allows time-varying velocity for the UEs, in which case the handoff rates are also time varying.

## V. UE'S DATA RATE AND TIER SELECTION

Suppose that an active UE chooses only to connect to a set of tiers  $\mathcal{S} \subset \mathcal{K}$ . We assume that  $1 \in \mathcal{S}$  (i.e., the UE always selects tier-1 macrocells). In this section, we first study the average *downlink* data rate of the UE under tier selection  $\mathcal{S}$ . Then, we discuss optimal tier selection taking both the handoff rates and the data rate into consideration.

### A. UE Data Rate

1) *Spectrum Allocation and Coverage Probability*: We assume that the active UE is allocated with a spectrum bandwidth of  $W_k$  if it is associated with a tier- $k$  BS. Different tiers of BSs are allocated separate spectrum, but BSs in the same tier share the same spectrum [48], [49], [50]. As a consequence, we need to characterize the co-tier interference in the system, which will influence the UE's data rate. Following conventional wireless modeling [8], [10], [50], we assume the UE requires a minimum Signal-to-Interference Ratio (SIR)  $T$ . The coverage probability of the UE is defined as the probability that its SIR is no less than  $T$  [44]. If the UE experiences coverage probability  $\mathbb{P}'$  and is allocated with a spectrum bandwidth  $W'$ , its data rate is  $W' \log(1 + T)$  if its SIR is no less than  $T$ , and its data rate is 0 if its SIR is less than  $T$  (i.e., outage occurs). Thus, the overall data rate of the UE is  $W' \log(1 + T)\mathbb{P}'$ . Note that  $\log$  is in base 2 throughout this paper. Also, we have assumed the common scenario where the system is interference limited, such that noise is negligible.

2) *Average UE Data Rate Derivation*: In this subsection, we derive the average UE data rate. An approach similar to one proposed in [11] is used to derive a closed-form expression for the average UE data rate. The main difference is that here a UE may choose an arbitrary subset of the tiers. We present an outline of the derivation for completeness. Interested readers are referred to [11] for more details.

Following (2), the probability that the active UE associates itself with a tier- $k$  ( $k \in \mathcal{S}$ ) BS is

$$A_{k,\mathcal{S}} = \frac{\lambda_k (P_k B_k)^{\frac{2}{\gamma}}}{\sum_{j \in \mathcal{S}} \lambda_j (P_j B_j)^{\frac{2}{\gamma}}}. \quad (24)$$

Given that the UE is associated with a tier- $k$  BS and their distance is  $d$ , the overall interference to it is the sum interference from all tier- $k$  BSs other than the BS associated by the UE:

$$I_{k,\mathcal{S}}(d) = \sum_{\mathbf{x} \in \Phi'_k} \frac{P_k h_{\mathbf{x},\mathbf{0}}}{\alpha |\mathbf{x}|^\gamma}, \quad (25)$$

where  $\Phi'_k$  is the reduced Palm point process [8] corresponding to all tier- $k$  BSs other than the BS associated by the UE. It can be shown that  $\Phi'_k$  is a PPP with intensity 0 in  $\mathcal{B}(\mathbf{0}, d)$  and

intensity  $\lambda_k$  in  $\mathbb{R}^2 \setminus \mathcal{B}(\mathbf{0}, d)$ , where  $\mathcal{B}(\mathbf{x}, r)$  denotes the disk region centered at  $\mathbf{x}$  with radius  $r$  [44].

The distribution of  $I_{k,\mathcal{S}}(d)$  is derived through its Laplace transform as follows:

$$\begin{aligned} \mathcal{L}_{I_{k,\mathcal{S}}}(d, s) &= \mathbf{E} \left[ \exp \left( - \sum_{\mathbf{x} \in \Phi'_k} \frac{s P_k h_{\mathbf{x},\mathbf{0}}}{\alpha |\mathbf{x}|^\gamma} \right) \right] \\ &= \exp \left( -2\pi \lambda_k \int_d^\infty \frac{\frac{s P_k r}{\alpha}}{\frac{s P_k}{\alpha} + r^\gamma} dr \right). \end{aligned} \quad (26)$$

Let  $P_{\text{cov},k,\mathcal{S}}(d)$  denote the conditional coverage probability of the active UE (given  $k$  and  $d$ ). Then,

$$\begin{aligned} P_{\text{cov},k,\mathcal{S}}(d) &= \mathbf{P} \left( \frac{P_k h_{\mathbf{x}_B,\mathbf{0}}}{\alpha d^\gamma} \geq T I_{k,\mathcal{S}}(d) \right) \\ &= \mathcal{L}_{I_{k,\mathcal{S}}}(d, s) \Big|_{s=\frac{T\alpha d^\gamma}{P_k}}, \end{aligned} \quad (27)$$

where  $\mathbf{x}_B$  is the coordinate of the BS associated by the UE, and  $|\mathbf{x}_B| = d$ . (27) is because  $h_{\mathbf{x}_B,\mathbf{0}}$  is exponentially distributed. Substituting (26) into (27), we have

$$\begin{aligned} P_{\text{cov},k,\mathcal{S}}(d) &= \exp \left( -2\pi \lambda_k \int_d^\infty \frac{T d^\gamma r}{T d^\gamma + r^\gamma} dr \right) \\ &\stackrel{t=\frac{r^2}{T^{\frac{2}{\gamma}} d^2}}{=} \exp \left( -\pi \lambda_k T^{\frac{2}{\gamma}} d^2 \int_{(\frac{1}{T})^{\frac{2}{\gamma}}}^\infty \frac{1}{1+t^{\frac{\gamma}{2}}} dt \right). \end{aligned} \quad (28)$$

Note that by doing so, we are able to capture the fact that the UE's data rate is higher if it is closer to its serving BS.

Next, the probability density function (pdf) of the distance between the UE and the BS associated by the UE is

$$f_{k,\mathcal{S}}(d) = \frac{2\pi \lambda_k}{A_{k,\mathcal{S}}} d \exp \left( -\pi d^2 \sum_{j \in \mathcal{S}} \lambda_j \left( \frac{P_j B_j}{P_k B_k} \right)^{\frac{2}{\gamma}} \right) \quad (29)$$

$$= \frac{2\pi \lambda_k}{A_{k,\mathcal{S}}} d \exp \left( -\pi d^2 \frac{\lambda_k}{A_{k,\mathcal{S}}} \right), \quad (30)$$

where (29) is derived in [8], and (30) is by substituting (24) into (29).

Thus, the coverage probability  $P_{\text{cov},k,\mathcal{S}}$  of the UE (given that it is associated with a tier- $k$  BS) can be computed as

$$\begin{aligned} P_{\text{cov},k,\mathcal{S}} &= \int_0^\infty f_{k,\mathcal{S}}(d) P_{\text{cov},k,\mathcal{S}}(d) dd \\ &= \frac{1}{1 + A_{k,\mathcal{S}} C}, \end{aligned} \quad (31)$$

where  $C \triangleq (T)^{\frac{2}{\gamma}} \int_{(\frac{1}{T})^{\frac{2}{\gamma}}}^\infty \frac{1}{1+t^{\frac{\gamma}{2}}} dt$ .

Because the active UE is associated with a spectrum bandwidth of  $W_k$  if it is associated with a tier- $k$  BS, its expected data rate can be computed as

$$\begin{aligned} \bar{R}_{\mathcal{S}} &= \sum_{k \in \mathcal{S}} A_{k,\mathcal{S}} W_k \log(1 + T) P_{\text{cov},k,\mathcal{S}} \\ &= \sum_{k \in \mathcal{S}} \frac{A_{k,\mathcal{S}} R_k}{1 + A_{k,\mathcal{S}} C}, \end{aligned} \quad (32)$$

where  $R_k \triangleq W_k \log(1 + T)$ .<sup>2</sup>

<sup>2</sup>In reality, UE's data rate may also be influenced by its velocity. To capture this effect, we can combine (32) with some empirical formulas (e.g., [51], [52], [53]). The remaining part the optimal tier selection does not change after the modification of (32).



## B. Optimal Tier Selection

Following the derivations in Section IV and V-A, we see that different tier selections lead to different data rates and handoff rates. Let  $\mathcal{C}_{kj}$  be the cost for one type  $k$ - $j$  handoff, and  $\mathcal{U}_R$  be the utility value for one bit data transmission. We assume the UE also pays service charge  $\mathcal{P}_k$  per second when it is associated with a type- $k$  BS. Note that  $\mathcal{C}_{kj}$ ,  $\mathcal{U}_R$ , and  $\mathcal{P}_k$  could be assigned arbitrarily, and  $\mathcal{C}_{kj}$  and  $\mathcal{C}_{jk}$  may be different if  $k \neq j$  [6]. If the UE favors higher data rate, it could assign a larger value for  $\mathcal{U}_R$ ; if it favors lower handoff rates, it could assign larger values for  $\mathcal{C}_{kj}$ .

If the active UE's tier selection is  $\mathcal{S}$ , its overall average utility on data transmission per second is  $\mathcal{U}_R \bar{R}_{\mathcal{S}}$ , overall average service charge per second is  $\mathcal{P}(\mathcal{S}) = \sum_{k \in \mathcal{S}} A_{k,\mathcal{S}} \mathcal{P}_k$ , and overall average expense on handoffs per second is

$$\mathcal{C}(\mathcal{S}, v) = \frac{2}{\pi} v \sum_{k \in \mathcal{S}} \frac{\lambda_k \left( \sum_{i \in \mathcal{S}} \frac{(\mathcal{C}_{ki} + \mathcal{C}_{ik})}{2} \lambda_i \mathcal{F}(\beta_{ki}) \right)}{2 \left( \sum_{i \in \mathcal{S}} \lambda_i \beta_{ik}^2 \right)^{\frac{3}{2}}}, \quad (33)$$

where (33) follows the conclusions of Theorem 3 and Corollary 1. Consequently, the overall average utility per second of tier selection  $\mathcal{S}$  is

$$G(\mathcal{S}, v) = \mathcal{U}_R \bar{R}_{\mathcal{S}} - \mathcal{C}(\mathcal{S}, v) - \mathcal{P}(\mathcal{S}). \quad (34)$$

Finally, the optimal tier selection is

$$\mathcal{S}_{opt} = \arg \max_{\mathcal{S} \in \mathbb{S}} (\mathcal{U}_R \bar{R}_{\mathcal{S}} - \mathcal{C}(\mathcal{S}, v) - \mathcal{P}(\mathcal{S})), \quad (35)$$

where  $\mathbb{S}$  is the set of all possible tier selections. Because the number of tiers  $K$  is usually not high in reality (i.e.,  $K \leq 5$ ), the cardinality of  $\mathbb{S}$ ,  $2^{K-1}$ , is not large. Therefore,  $\mathcal{S}_{opt}$  can be derived through comparing all possible tier selections.

## VI. EXPERIMENTAL STUDY

In this section, our analysis is validated via experimenting with real-world traces and simulations.

### A. Yonsei Trace Data

We use the real-world Yonsei Trace [12] to validate our analytical results. The trace was accumulated from 12 commercial mobile phones during an 8-month period in 2011 in the city of Seoul. An application named SmartDC had been running on the commercial mobile phones equipped with GPS, GSM, and WiFi. For every 2 to 5 minutes, the application collected UE's location information (latitude and longitude), the MAC addresses of surrounding WiFi APs, and the cell IDs of nearby cellular BSs they could detect. Each AP has a unique MAC address and each BS has a unique cell ID. By analyzing the data set, we are able to determine which APs and BSs a UE could detect at the recorded coordinates and time instants. In the following, we regard cellular BSs as tier-1 BSs and APs as tier-2 BSs.

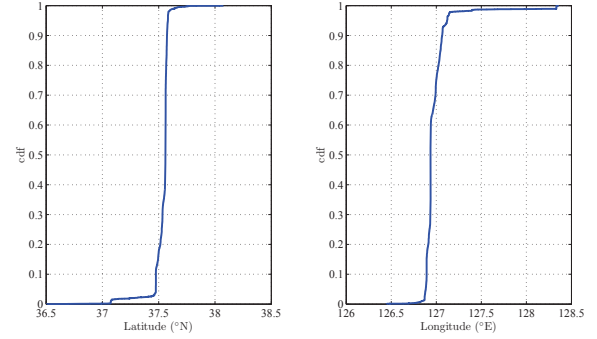


Fig. 3. Cumulative distribution function (cdf) of the latitude and longitude.

### B. Data Processing

1) *Location Approximations of APs and BSs*: As the data set does not explicitly provide the latitudes and longitudes of APs and BSs, we apply the following approach to approximate their locations: for each AP (resp. BS), we list all the coordinates recorded by UEs when they are able to detect the AP (resp. BS). Then, we approximate the location of the AP (resp. BS), by taking the average of these recorded coordinates.

2) *Reference Region*: In order to avoid the edge effect, we define a reference region, in which most recorded coordinates are located. The UEs' trajectories are only accounted inside the reference region. By plotting the cumulative distribution function (cdf) of the latitude (resp. longitude) of all recorded coordinates (shown in Fig. 3), we observe a sharp step upward between  $37.48^\circ N$  and  $37.58^\circ N$  (resp.  $126.9^\circ E$  and  $127.1^\circ E$ ). As a consequence, we employ the rectangle defined by  $37.48^\circ N$ , and  $37.58^\circ N$ ,  $126.9^\circ E$ , and  $127.1^\circ E$  as the reference region.

3) *UE Trajectory*: In the trace data, the coordinates of a UE are recorded only once every few minutes. To recover its full trajectory, we regard it as moving in a straight line at a constant velocity between two consecutive recorded coordinates. Thus, interpolations can be made to determine the coordinate of the UE at any time. Note that only the trajectories inside the reference region are used.

4) *Handoff Rates*: Through the locations of BSs and APs, as well as the UE trajectories, we are able to derive the empirical rates of all handoff types following the biased user association scheme discussed in Section III-B. If we ignore all the APs, we can also derive the empirical handoff rates for the one-tier case.

5) *BS and AP Intensities*: The AP (resp. BS) density is computed as the number of APs (resp. BSs) over the area of the reference region, which is  $455.1$  unit/km<sup>2</sup> (resp.  $52.6$  unit/km<sup>2</sup>). This indicates an urban area with high population and BS densities.

### C. Empirical Results

We compare the handoff rates derived from our analysis and those from our empirical study based on the Yonsei Trace. The empirical handoff rates are derived from the steps in Sections VI-B1 - VI-B4. For the analytical results, we use the BS and AP intensities shown in Section VI-B5 as input parameters.

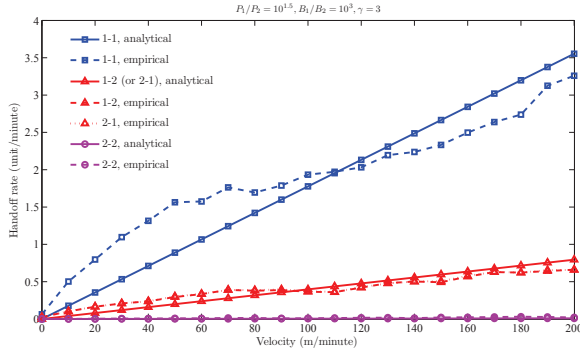


Fig. 4. Two-tier case: comparison of analytical and empirical handoff rates.

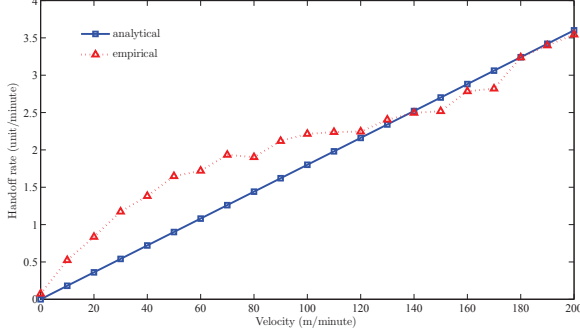


Fig. 5. One-tier case: comparison of analytical and empirical handoff rates.

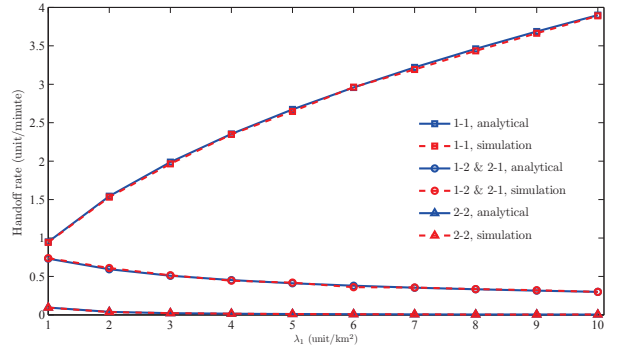
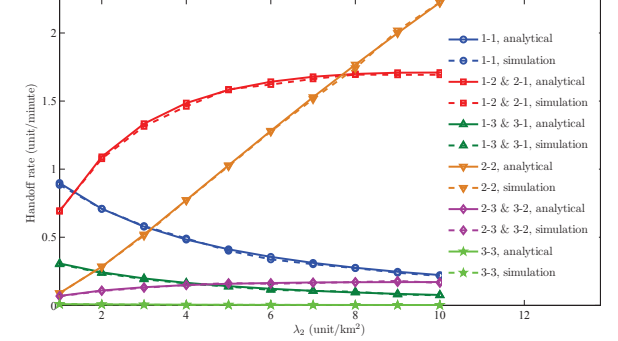
For the two-tier case, the comparison of analytical and empirical handoff rates is shown in Fig. 4. For the one-tier case (by eliminating all the APs), the comparison is shown in Fig. 5. Both figures illustrate the accuracy of our analysis. When the UE's velocity is low, empirical handoff rates are slightly greater than analytical handoff rates. This is because the locations of APs and BSs are not strictly homogeneous distributed (e.g., some APs and BSs are crowded along some streets, or at the center of the urban region). We also observe that UEs with lower velocity are more likely to be sampled in the region with higher AP and BS densities. As a consequence, the empirical handoff rates are higher than those expected by our analytical results.

Fig. 4 and Fig. 5 also show that type 1-1 horizontal handoff rates are almost the same in the one-tier and two-tier cases, but extra type 1-2 and type 2-1 vertical handoffs are introduced in the two-tier case. This agrees with our expectation that adding a second tier of APs brings more vertical handoffs. In addition, as a validation of (21), type 1-2 and type 2-1 handoff rates are almost the same in empirical results.

#### D. Simulation Study

In this subsection, we present simulation results to further demonstrate our analysis in more complex HWNs.

1) *Simulation Setup*: The simulation procedure is as follows: in each round of simulation, two or three tiers of BSs are generated on a 10 km  $\times$  10 km square. Then, we randomly generate 5 waypoints  $X_1, \dots, X_5$  in the central 5 km  $\times$  5 km square (uniformly distributed). The five line segments  $X_1X_2, X_2X_3, \dots, X_4X_5$  construct the trajectory of an active UE. In this way, we derive the simulated handoff rates in this round of simulation. The above procedure is repeated 200

Fig. 6. Two-tier case: handoff rates under different  $\lambda_1$ .Fig. 7. Three-tier case: handoff rates under different  $\lambda_2$ .

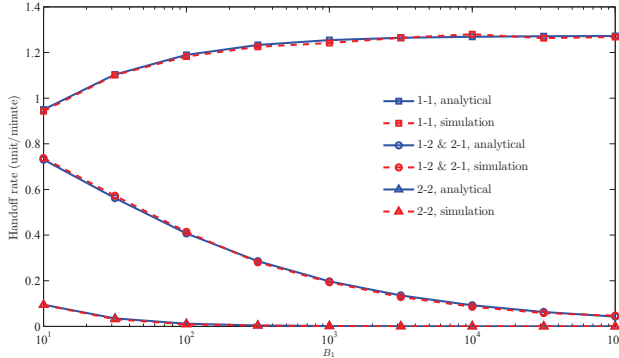
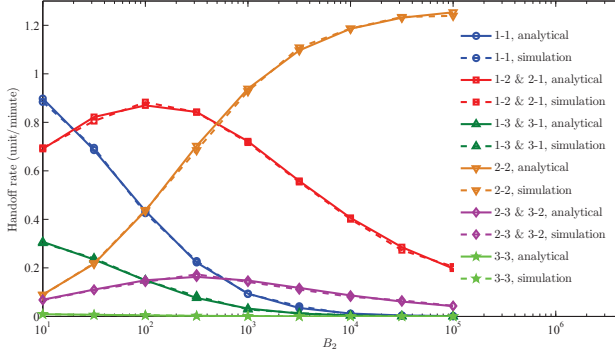
rounds to derive one simulated data point. Note that in this subsection, in order to avoid overlapping in figures, we only show the sum rate of type  $j-k$  and type  $k-j$  ( $k \neq j$ ) handoffs for easier inspection; the individual handoff rates are half of the sum handoff rate.

2) *Handoff Rates under Different BS Intensities*: We study handoff rates under different BS intensities. Fig. 6 shows a two-tier case, with parameters as follows:  $(P_1, P_2) = (30, 20)$  dBm,  $(B_1, B_2) = (1, 1)$ , and  $\lambda_2 = 1$  unit/km<sup>2</sup>. Fig. 7 shows a three-tier case, with parameters as follows:  $(P_1, P_2, P_3) = (30, 20, 10)$  dBm,  $(B_1, B_2, B_3) = (1, 1, 1)$ , and  $(\lambda_1, \lambda_3) = (1, 1)$  unit/km<sup>2</sup>. The parameter values  $\gamma = 3$  and  $v = 60$  km/h are used for both Fig. 6 and Fig. 7.

Fig. 6 illustrates that increasing  $\lambda_1$  leads to higher type 1-1 handoff rate but lower type 2-2 handoff rate. Fig. 7 illustrates that increasing  $\lambda_2$  leads to higher type 2-2 handoff rate but lower type 1-1 and 1-3 & 3-1 handoff rates. Both observations suggest that increasing the BS intensity of one tier causes higher horizontal handoff rate within this tier, but lower handoff rates outside this tier.

3) *Handoff Rates under Different Association Bias Values*: Next, we study handoff rates under different association bias values. Fig. 8 shows a two-tier case, with parameters as follows:  $(P_1, P_2) = (30, 20)$  dBm,  $B_2 = 1$ , and  $(\lambda_1, \lambda_2) = (1, 1)$  unit/km<sup>2</sup>. Fig. 9 shows a three-tier case, with parameters as follows:  $(P_1, P_2, P_3) = (30, 20, 10)$  dBm,  $(B_1, B_3) = (1, 1)$ , and  $(\lambda_1, \lambda_2, \lambda_3) = (1, 1, 1)$  unit/km<sup>2</sup>. The parameter values  $\gamma = 3$  and  $v = 60$  km/h are used for both Fig. 8 and Fig. 9. These figures suggest that, increasing the association bias value of one tier has a similar effect as increasing the BS intensity of this tier, leading to higher horizontal handoff rate within this tier, but lower handoff rates outside this tier.



Fig. 8. Two-tier case: handoff rates under different  $B_1$ .Fig. 9. Three-tier case: handoff rates under different  $B_2$ .

4) *UE's Utility under Different Tier Selections*: Fig. 10 shows simulated utility of an active UE under different velocity values in a two-tier case. The parameters are as follows:  $(P_1, P_2) = (40, 20)$  dBm,  $(B_1, B_2) = (1, 1)$ ,  $(\lambda_1, \lambda_2) = (1, 5)$  units/km<sup>2</sup>,  $\gamma = 3$ ,  $(R_1, R_2) = (2, 5)$  Mbps,  $(C_{11}, C_{12}, C_{21}, C_{22}) = (10, 45, 35, 20)$ ,  $U_R = 1$ ,  $T = 0.5$ , and  $(P_1, P_2) = (0.1, 0.1)$ <sup>3</sup>. The vertical line shows the analytical tier selection threshold on the UE velocity, which is 34.47 km/h. This figure demonstrates that the simulation results agree with the analytical results in tier selection as discussed in Section V. Tier selection  $\{1, 2\}$  is optimal if the UE's velocity is low, but its overall utility decreases faster due to higher handoff expense. Tier selection  $\{1\}$  is optimal if the UE's velocity is greater than 34.47 km/h.

Fig. 11 shows simulated utility of an active UE under different velocity values in a three-tier case. The parameters are as follows:  $(P_1, P_2, P_3) = (40, 20, 10)$  dBm,  $(B_1, B_2, B_3) = (1, 1, 1)$ ,  $(\lambda_1, \lambda_2, \lambda_3) = (1, 5, 20)$  units/km<sup>2</sup>,  $\gamma = 3$ ,  $(R_1, R_2, R_3) = (2, 5, 10)$  Mbps,  $(C_{11}, C_{12}, C_{21}, C_{22}, C_{13}, C_{31}, C_{23}, C_{32}, C_{33}) = (10, 35, 25, 20, 45, 35, 55, 45, 30)$ ,  $U_R = 1$ ,  $T = 0.5$ , and  $(P_1, P_2, P_3) = (0.1, 0.1, 0.1)$ . Two vertical lines show the analytical tier selection thresholds, which are 32.24 km/h and 50.70 km/h respectively. This figure again shows that the simulation results agree with the analytical results in tier selection. When the velocity is in the range  $[0, 32.24)$  km/h, tier selection  $\{1, 2, 3\}$  is optimal; when the velocity is in the range  $[32.24, 50.70)$  km/h, tier selection  $\{1, 2\}$  is optimal; when the velocity is in the range  $[50.70, \infty)$  km/h,

<sup>3</sup> $C_{k_j}$  is in unit utility/handoff,  $U_R$  is in unit utility/Mbit, and  $P_k$  is in unit utility/s in this section.

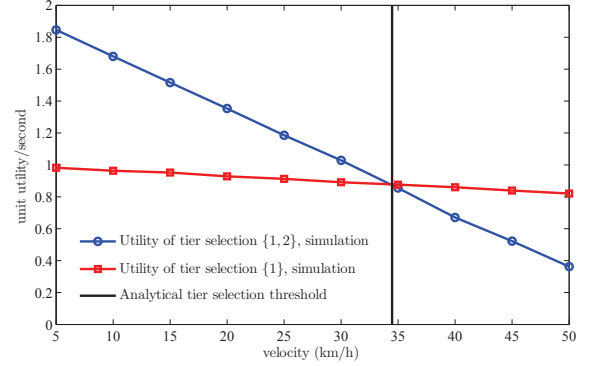


Fig. 10. Two-tier case: overall utility comparison of different tier selections.

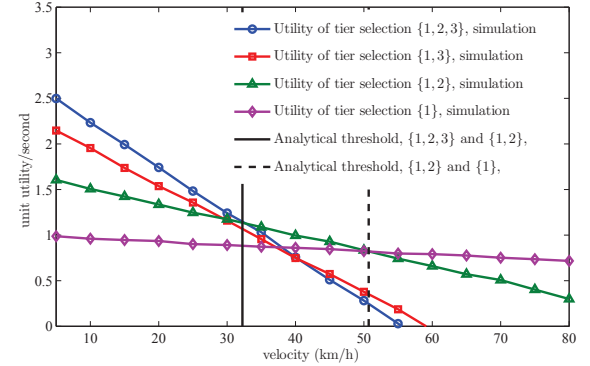


Fig. 11. Three-tier case: overall utility comparison of different tier selections.

tier selection  $\{1\}$  is optimal.

5) *Velocity Threshold*: Fig. 12 and Fig. 13 show the computed velocity thresholds for tier selections, under different BS densities. Fig. 12 shows a two-tier case. The parameters are as follows:  $(P_1, P_2) = (40, 20)$  dBm,  $\lambda_1 = 1$  units/km<sup>2</sup>,  $\gamma = 3$ ,  $(R_1, R_2) = (2, 5)$  Mbps,  $(C_{11}, C_{12}, C_{21}, C_{22}) = (10, 45, 35, 20)$ ,  $T = 0.5$ , and  $(P_1, P_2) = (0.1, 0.1)$ . Fig. 13 shows a three-tier case. The parameters are as follows:  $(P_1, P_2, P_3) = (40, 20, 10)$  dBm,  $(B_1, B_2, B_3) = (1, 1, 1)$ ,  $(\lambda_1, \lambda_2) = (1, 5)$  units/km<sup>2</sup>,  $\gamma = 3$ ,  $(R_1, R_2, R_3) = (2, 5, 10)$  Mbps,  $(C_{11}, C_{12}, C_{21}, C_{22}, C_{13}, C_{31}, C_{23}, C_{32}, C_{33}) = (10, 35, 25, 20, 45, 35, 55, 45, 30)$ ,  $T = 0.5$ , and  $(P_1, P_2, P_3) = (0.1, 0.1, 0.1)$ .

In the two-tier case, increasing  $\lambda_2$  or  $B_2$  improves the average UE data rate (as the UE has higher probability to associate with tier-2 BSs), but it could also cause higher handoff rates. Through our theoretical analysis, we could observe that the latter factor has a stronger effect and the velocity threshold value is lowered if  $\lambda_2$  increases; while the former factor dominates and the velocity threshold value increases if  $B_2$  becomes greater. In addition, increasing  $U_R$  leads to a higher weight in data rate, so the threshold values increase.

In the three-tier case, increasing  $\lambda_3$  could cause a more complicated impact on optimal tier-selection. We observe that the tier selection  $\{1, 2\}$  is broken into two separate regions. When  $\lambda_3$  is small, the velocity range of tier selection  $\{1, 2\}$  is below the velocity range of tier selection  $\{1, 3\}$ , while the former range is above the latter one when  $\lambda_3$  becomes larger. Still, increasing  $U_R$  leads to a higher weight in data rate, thus

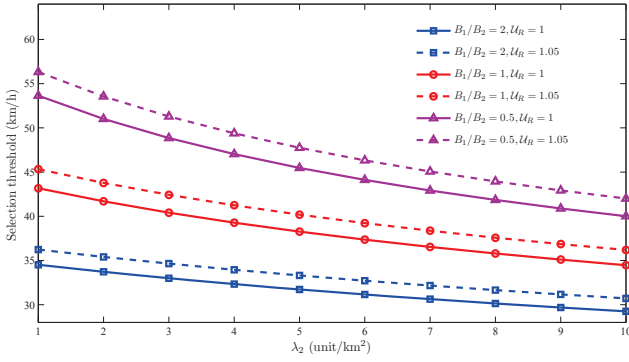


Fig. 12. Two-tier case: tier selection velocity threshold.

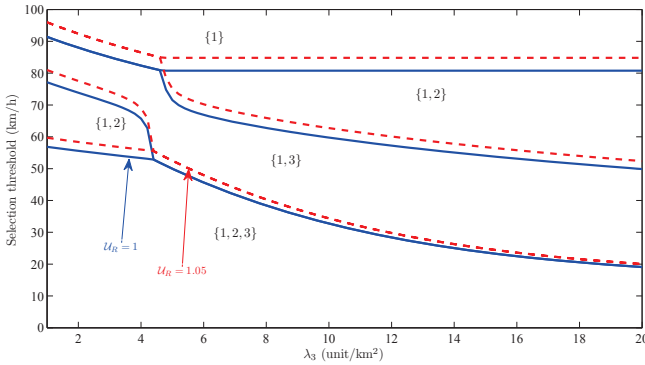


Fig. 13. Three-tier case: tier selection velocity threshold.

the threshold values are increased (i.e., the dashed curves are shifted upward compared with the solid curves).

## VII. CONCLUSIONS

In this work, we provide a theoretical framework to study user mobility in multi-tier HWNs. Through establishing a stochastic geometric framework, we capture the irregularly shaped network topologies introduced by the small-cell structure. Theoretical expressions for the rates of all handoff types experienced by an active UE with arbitrary movement trajectory are derived. In addition, we investigate downlink data rate of the UE under different tier selections. Based on these results, optimal tier selection considering both the handoff rates and the data rate is studied. Empirical study on the Yonsei Trace and extensive simulation are conducted, validating the accuracy and usefulness of our analytical conclusions.

## VIII. APPENDIX

### A. Proof of Theorem 1

*Proof:* In this proof, the tier- $k$  BS with which the reference UE is associated is referred to as the reference BS. Without loss of generality, we assume the reference BS is located at  $\mathbf{r}_0 = (r_0, 0)$ . Note that because the reference UE receives the highest biased power level from the reference BS,  $\forall j \in \mathcal{K}$ , there are no tier- $j$  BSs located within  $\mathcal{B}(\mathbf{0}, \frac{r_0}{\beta_{kj}})$ , where  $\mathcal{B}(\mathbf{x}, r)$  denotes the disk centered at  $\mathbf{x}$  with radius  $r$ , and  $\tilde{\mathcal{B}}(\mathbf{x}, r)$  denotes  $\mathbb{R}^2 \setminus \mathcal{B}(\mathbf{x}, r)$ .

Let  $\mathbf{x}_0 = (x_0, y_0)$  denote the position of some tier- $j$  BS (other than the reference BS if  $j = k$ ). Let  $\mathcal{T}(\mathbf{r}_0, \mathbf{x}_0, \beta_{kj})$  denote the trace satisfying the following condition:

$$\mathcal{T}(\mathbf{r}_0, \mathbf{x}_0, \beta_{kj}) = \left\{ (x, y) \left| \frac{P_k B_k}{((x - r_0)^2 + y^2)^{\gamma/2}} = \frac{P_j B_j}{((x - x_0)^2 + (y - y_0)^2)^{\gamma/2}} \right. \right\}. \quad (36)$$

Note that  $\mathbf{0} \in \mathbf{T}_{kj}^{(2)}(\Delta d)$  is equivalent to that the distance from  $\mathbf{0}$  to the trace  $\mathcal{T}(\mathbf{r}_0, \mathbf{x}_0, \beta_{kj})$  is less than  $\Delta d$ .

In the following, we discuss three cases respectively:  $\beta_{kj} > 1$ ,  $\beta_{kj} = 1$ , and  $\beta_{kj} < 1$ .

**Case 1:**  $\beta_{kj} > 1$ .

In this case, we have

$$\mathcal{T}(\mathbf{r}_0, \mathbf{x}_0, \beta_{kj}) = \left\{ (x, y) \left| \left[ x - \left( \frac{\beta_{kj}^2 x_0 - r_0}{\beta_{kj}^2 - 1} \right) \right]^2 + \left[ y - \frac{\beta_{kj}^2 y_0}{\beta_{kj}^2 - 1} \right]^2 = \frac{\beta_{kj}^2 (r_0^2 + x_0^2 + y_0^2 - 2x_0 r_0)}{(\beta_{kj}^2 - 1)^2} \right. \right\}, \quad (37)$$

which is a circle centered at  $\left( \frac{\beta_{kj}^2 x_0 - r_0}{\beta_{kj}^2 - 1}, \frac{\beta_{kj}^2 y_0}{\beta_{kj}^2 - 1} \right)$  with radius  $\frac{\beta_{kj} \sqrt{(r_0^2 + x_0^2 + y_0^2 - 2x_0 r_0)}}{(\beta_{kj}^2 - 1)}$ . Thus, the distance from  $\mathbf{0}$  to the trace  $\mathcal{T}(\mathbf{r}_0, \mathbf{x}_0, \beta_{kj})$  is

$$d(\mathbf{r}_0, \mathbf{x}_0, \beta_{kj}) = \frac{\left| \sqrt{(\beta_{kj}^2 x_0 - r_0)^2 + (\beta_{kj}^2 y_0)^2} - \sqrt{\beta_{kj}^2 (r_0^2 + x_0^2 + y_0^2 - 2x_0 r_0)} \right|}{(\beta_{kj}^2 - 1)}. \quad (38)$$

Thus,  $\mathbf{0} \in \mathbf{T}_{kj}^{(2)}(\Delta d)$  iff  $d(\mathbf{r}_0, \mathbf{x}_0, \beta_{kj}) < \Delta d$ , or equivalently,  $\mathbf{x}_0 \in \tilde{S}_{kj}(\Delta d)$ , where

$$\tilde{S}_{kj}(\Delta d) = \left\{ \mathbf{x}_0 \mid d(\mathbf{r}_0, \mathbf{x}_0, \beta_{kj}) < \Delta d \right\}. \quad (39)$$

Mathematical manipulations of (39) lead to

$$\tilde{S}_{kj}(\Delta d) = \left\{ (x_0, y_0) \left| \left| (x_0^2 + y_0^2) - \frac{r_0^2}{\beta_{kj}^2} \right| < \frac{\Delta d}{\beta_{kj}^2} \right. \right\}. \quad (40)$$

$$\sqrt{2 \left( \beta_{kj}^4 + \beta_{kj}^2 \right) (x_0^2 + y_0^2) - 8\beta_{kj}^2 x_0 r_0 + 2(\beta_{kj}^2 + 1)r_0^2 + \mathcal{O}(\Delta d^2)}.$$

By converting  $(x_0, y_0)$  into polar coordinates  $(r, \theta)$ , (40) becomes

$$\tilde{S}_{kj}(\Delta d) = \left\{ (r, \theta) \left| \left| r^2 - \frac{r_0^2}{\beta_{kj}^2} \right| < \frac{\Delta d}{\beta_{kj}^2} \right. \right\}. \quad (41)$$

$$\sqrt{2 \left( \beta_{kj}^4 + \beta_{kj}^2 \right) r^2 - 8\beta_{kj}^2 r_0 r \cos(\theta) + 2(\beta_{kj}^2 + 1)r_0^2 + \mathcal{O}(\Delta d^2)}.$$

Note that there are no tier- $j$  BSs located inside  $\mathcal{B}(\mathbf{0}, \frac{r_0}{\beta_{kj}})$ . Let  $S_{kj}(\Delta d) = \tilde{S}_{kj}(\Delta d) \cap \tilde{\mathcal{B}}(\mathbf{0}, r_0/\beta_{kj})$ . As a result,  $\mathbf{0} \in \mathbf{T}_{kj}^{(2)}(\Delta d)$  iff  $\mathbf{x}_0 \in S_{kj}(\Delta d)$ , where

$$S_{kj}(\Delta d) = \left\{ (r, \theta) \left| r \geq \frac{r_0}{\beta_{kj}} \text{ and } \left| r^2 - \frac{r_0^2}{\beta_{kj}^2} \right| < \frac{\Delta d}{\beta_{kj}^2} \right. \right\}. \quad (42)$$

$$\sqrt{2 \left( \beta_{kj}^4 + \beta_{kj}^2 \right) r^2 - 8\beta_{kj}^2 r_0 r \cos(\theta) + 2(\beta_{kj}^2 + 1)r_0^2 + \mathcal{O}(\Delta d^2)}.$$

According to (42),  $S_{kj}(\Delta d)$  corresponds to a “ring” region (shaded area) shown in Fig. 14. We can observe that  $\forall (r, \theta) \in S_{kj}(\Delta d)$ ,  $r = \frac{r_0}{\beta_{kj}} + \mathcal{O}(\Delta d)$ . Substituting it into (42) gives,

$$S_{kj}(\Delta d) = \left\{ (r, \theta) \mid r \geq \frac{r_0}{\beta_{kj}} \text{ and} \right. \quad (43)$$

$$\left. \left| r^2 - \frac{r_0^2}{\beta_{kj}^2} \right| < \frac{\Delta d}{\beta_{kj}^2} \cdot \left[ 2(\beta_{kj}^4 + \beta_{kj}^2) \left( \frac{r_0}{\beta_{kj}} + \mathcal{O}(\Delta d) \right)^2 - 8\beta_{kj}^2 r_0 \left( \frac{r_0}{\beta_{kj}} + \mathcal{O}(\Delta d) \right) \cos(\theta) + 2(\beta_{kj}^2 + 1)r_0^2 + \mathcal{O}(\Delta d^2) \right]^{1/2} \right\},$$

which leads to,

$$S_{kj}(\Delta d) = \left\{ (r, \theta) \mid r \geq \frac{r_0}{\beta_{kj}} \text{ and } \left| r^2 - \frac{r_0^2}{\beta_{kj}^2} \right| < \frac{2\Delta d r_0}{\beta_{kj}^2} \sqrt{(\beta_{kj}^2 + 1) - 2\beta_{kj} \cos(\theta)} + \mathcal{O}(\Delta d^2) \right\}. \quad (44)$$

The area of  $S_{kj}(\Delta d)$  is

$$\begin{aligned} |S_{kj}(\Delta d)| & \quad (45) \\ &= 2 \int_0^\pi \int_{\frac{r_0}{\beta_{kj}}}^{\sqrt{\frac{r_0^2}{\beta_{kj}^2} + \frac{2\Delta d r_0}{\beta_{kj}^2} \sqrt{(\beta_{kj}^2 + 1) - 2\beta_{kj} \cos(\theta)} + \mathcal{O}(\Delta d^2)}} r dr d\theta \\ &= \frac{2\Delta d r_0}{\beta_{kj}^2} \int_0^\pi \sqrt{(\beta_{kj}^2 + 1) - 2\beta_{kj} \cos(\theta)} d\theta + \mathcal{O}(\Delta d^2). \end{aligned}$$

Given the reference UE and BS, it can be shown that  $\Phi_j^4$  is a PPP with intensity 0 in  $\mathcal{B}\left(\mathbf{0}, \frac{r_0}{\beta_{kj}}\right)$  and intensity  $\lambda_j$  in  $\overline{\mathcal{B}}\left(0, r_0/\beta_{kj}\right)$  [44]. Because  $\mathbb{P}\left(\mathbf{0} \in \mathbf{T}_{kj}^{(2)}(\Delta d) \mid R = r_0, \text{tier} = k\right)$  is equal to the probability that there is at least one point of  $\Phi_j$  in  $S_{kj}(\Delta d)$  (i.e., some  $\mathbf{x}_0$  in  $S_{kj}(\Delta d)$ ), we have

$$\begin{aligned} & \mathbb{P}\left(\mathbf{0} \in \mathbf{T}_{kj}^{(2)}(\Delta d) \mid R = r_0, \text{tier} = k\right) \quad (46) \\ &= 1 - \exp(-\lambda_j |S_{kj}(\Delta d)|) \\ &= 1 - \exp(-2\lambda_j \Delta d r_0 \mathcal{F}(\beta_{kj}) + \mathcal{O}(\Delta d^2)), \end{aligned}$$

which completes the proof for Case 1.

**Case 2:**  $\beta_{kj} < 1$ . The proof is similar to that of Case 1.

**Case 3:**  $\beta_{kj} = 1$ .

In this case, we have

$$\mathcal{T}(\mathbf{r}_0, \mathbf{x}_0, 1) = \left\{ (x, y) \mid y_0 \left( y - \frac{y_0}{2} \right) = -(x_0 - r_0) \left( x - \frac{x_0 + r_0}{2} \right) \right\}, \quad (47)$$

which is a line. Thus, the distance from  $\mathbf{0}$  to  $\mathcal{T}(\mathbf{r}_0, \mathbf{x}_0, 1)$  is

$$d(\mathbf{x}_0, \mathbf{r}_0, 1) = \frac{\left| \frac{y_0^2}{2} - \frac{(r_0 - x_0)(r_0 + x_0)}{2} \right|}{\sqrt{(r_0 - x_0)^2 + y_0^2}}. \quad (48)$$

Consequently, similar to (39),  $\mathbf{0} \in \mathbf{T}_{kj}^{(2)}(\Delta d)$  if  $d(\mathbf{x}_0, \mathbf{r}_0, 1) < \Delta d$ , or equivalently,  $\mathbf{x}_0 \in \tilde{S}_{kj}(\Delta d)$ , where

$$\tilde{S}_{kj}(\Delta d) = \left\{ (x_0, y_0) \mid \left| \frac{y_0^2}{2} - \frac{(r_0 - x_0)(r_0 + x_0)}{2} \right| < \Delta d \right\}. \quad (49)$$

<sup>4</sup>If  $k = j$ , it is the reduced Palm point process [8] corresponding to all tier- $k$  BSs other than reference BS.

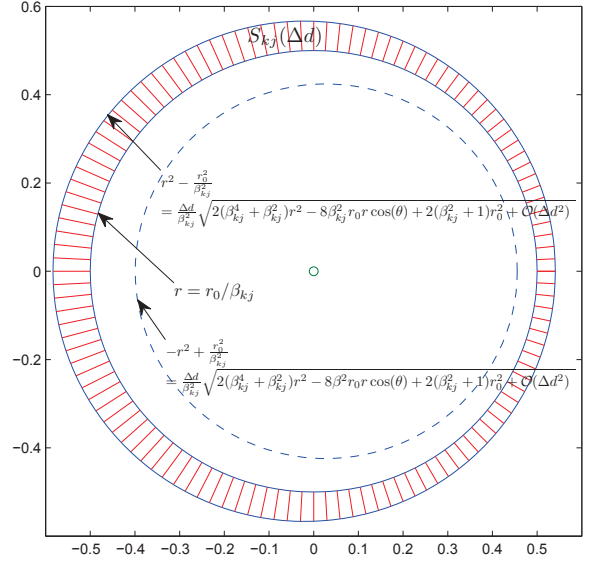


Fig. 14. The region (shaded part) of  $S_{kj}(\Delta d)$ .

After converting  $(x_0, y_0)$  into polar coordinate  $(r, \theta)$ ,

$$\tilde{S}_{kj}(\Delta d) = \left\{ (r, \theta) \mid \left| r^2 - r_0^2 \right| < 2\Delta d \sqrt{r_0^2 + r^2 - 2r_0 r \cos \theta} \right\}, \quad (50)$$

which is a special case of (41) with  $\beta_{kj} = 1$ . Thus, following the same steps as (42)-(46), we can still derive (13), which completes the proof of Case 3. ■

## B. Proof of Theorem 2

*Proof:* (a) Let  $E_i$  denote the event that there is at least one tier- $i$  BS located in  $S_{ki}(\Delta d)$ . Then

$$\begin{aligned} & \mathbb{P}\left(\mathbf{0} \in \mathbf{T}^{(2)}(\Delta d) \mid R = r_0, \text{tier} = k\right) \\ &= 1 - \mathbb{P}\left(\overline{E_1} \cap \overline{E_2} \cap \dots \cap \overline{E_K} \mid R = r_0, \text{tier} = k\right) \\ &= 1 - \exp\left(-\sum_{i=1}^K |S_{ki}(\Delta d)| \lambda_i\right) \\ &= 1 - \exp\left(-\sum_{i=1}^K 2\lambda_i \Delta d r_0 \mathcal{F}(\beta_{ki}) + \mathcal{O}(\Delta d^2)\right) \\ &= \sum_{i=1}^K 2\lambda_i \Delta d r_0 \mathcal{F}(\beta_{ki}) + \mathcal{O}(\Delta d^2). \end{aligned} \quad (51)$$

Furthermore, according to the results in [8], the probability density function (pdf) of the distance between the reference UE and the reference BS is

$$f_k(r_0 \mid \text{tier} = k) = \frac{2\pi\lambda_k}{A_k} r_0 \exp\left(-\pi r_0^2 \sum_{i=1}^K \lambda_i \beta_{ik}^2\right). \quad (52)$$

Also, we have  $\mathbb{P}(\text{tier} = k) = A_k$ , thus

$$\begin{aligned} & \mathbb{P}\left(\mathbf{0} \in \mathbf{T}^{(2)}(\Delta d)\right) \\ &= \sum_{k=1}^K \int_0^\infty \mathbb{P}\left(\mathbf{0} \in \mathbf{T}^{(2)}(\Delta d) \mid R = r_0, \text{tier} = k\right) \end{aligned}$$

$$\begin{aligned}
& \cdot f_k(r_0|\text{tier} = k)\mathbb{P}(\text{tier} = k)dr_0 \\
& = \sum_{k=1}^K \int_0^\infty 2\pi\lambda_k r_0 \exp\left(-\pi r_0^2 \sum_{i=1}^K \lambda_i \beta_{ik}^2\right) \\
& \quad \cdot \left(\sum_{i=1}^K 2\lambda_i \Delta dr_0 \mathcal{F}(\beta_{ki}) + \mathcal{O}(\Delta d^2)\right) dr_0 \\
& = \sum_{k=1}^K \frac{\lambda_k \left(\sum_{i=1}^K \lambda_i \Delta d \mathcal{F}(\beta_{ki}) + \mathcal{O}(\Delta d^2)\right)}{\left(\sum_{i=1}^K \lambda_i \beta_{ik}^2\right)^{\frac{3}{2}}}, \tag{53}
\end{aligned}$$

which completes the proof of (a).

**(b)**

Similar to (53), if  $k \neq j$  we have

$$\begin{aligned}
& \mathbb{P}\left(\mathbf{0} \in \mathbf{T}_{kj}^{(2)}(\Delta d)\right) \\
& = \int_0^\infty \mathbb{P}(\mathbf{0} \in \mathbf{T}_{kj}^{(2)}(\Delta d)|R = r_0, \text{tier} = k) f_k(r_0|\text{tier} = k) \\
& \quad \cdot \mathbb{P}(\text{tier} = k) dr_0 \\
& + \int_0^\infty \mathbb{P}(\mathbf{0} \in \mathbf{T}_{kj}^{(2)}(\Delta d)|R = r_0, \text{tier} = j) f_j(r_0|\text{tier} = j) \\
& \quad \cdot \mathbb{P}(\text{tier} = j) dr_0 \\
& = \frac{\lambda_k (\lambda_j \Delta d \mathcal{F}(\beta_{kj}) + \mathcal{O}(\Delta d^2))}{\left(\sum_{i=1}^K \lambda_i \beta_{ik}^2\right)^{\frac{3}{2}}} + \frac{\lambda_j (\lambda_k \Delta d \mathcal{F}(\beta_{jk}) + \mathcal{O}(\Delta d^2))}{\left(\sum_{i=1}^K \lambda_i \beta_{ij}^2\right)^{\frac{3}{2}}}. \tag{54}
\end{aligned}$$

Otherwise, if  $k = j$  we have

$$\mathbb{P}\left(\mathbf{0} \in \mathbf{T}_{kk}^{(2)}(\Delta d)\right) = \frac{\lambda_k (\lambda_k \Delta d \mathcal{F}(1) + \mathcal{O}(\Delta d^2))}{\left(\sum_{i=1}^K \lambda_i \beta_{ik}^2\right)^{\frac{3}{2}}}, \tag{55}$$

which completes the proof of (b).  $\blacksquare$

### C. Proof of Theorem 3

*Proof:* We have

$$\mu_1(\mathbf{T}^{(1)}) = \lim_{\Delta d \rightarrow 0} \frac{\mu_2(\mathbf{T}^{(2)}(\Delta d))}{2\Delta d} \tag{56}$$

$$= \sum_{k=1}^K \frac{\lambda_k \left(\sum_{i=1}^K \lambda_i \mathcal{F}(\beta_{ki})\right)}{2 \left(\sum_{i=1}^K \lambda_i \beta_{ik}^2\right)^{\frac{3}{2}}}, \tag{57}$$

where (56) follows [54] and Section 3.2 in [55]. Similarly,

$$\begin{aligned}
\mu_1(\mathbf{T}_{kj}^{(1)}) & = \lim_{\Delta d \rightarrow 0} \frac{\mu_2(\mathbf{T}_{kj}^{(2)}(\Delta d))}{2\Delta d} \tag{58} \\
& = \begin{cases} \frac{\lambda_k \lambda_j \mathcal{F}(\beta_{kj})}{2 \left(\sum_{i=1}^K \lambda_i \beta_{ik}^2\right)^{\frac{3}{2}}} + \frac{\lambda_j \lambda_k \mathcal{F}(\beta_{jk})}{2 \left(\sum_{i=1}^K \lambda_i \beta_{ij}^2\right)^{\frac{3}{2}}} & \text{if } k \neq j, \\ \frac{\lambda_k^2 \mathcal{F}(1)}{2 \left(\sum_{i=1}^K \lambda_i \beta_{ik}^2\right)^{\frac{3}{2}}} & \text{if } k = j. \end{cases}
\end{aligned}$$

### REFERENCES

- [1] W. Bao and B. Liang, "Handoff rate analysis in heterogeneous cellular networks: a stochastic geometric approach," in *Proc. of ACM MSWiM*, Montreal, Canada, 2014.
- [2] V. Chandrasekhar, J. Andrews, and A. Gatherer, "Femtocell networks: a survey," *IEEE Communications Magazine*, vol. 46, no. 9, pp. 59–67, Sep. 2008.
- [3] M. Bennis, M. Simsek, A. Czylik, W. Saad, S. Valentin, and M. Debbah, "When cellular meets WiFi in wireless small cell networks," *IEEE Communications Magazine*, vol. 51, no. 6, pp. 44–50, Jun. 2013.
- [4] "AT&T 3G Microcell," <http://www.att.com/standalone/3gmicrocell>.
- [5] "AT&T Wi-Fi hot spot locations," <https://www.att.com/maps/wifi>.
- [6] M. Stemm and R. H. Katz, "Vertical handoffs in wireless overlay networks," *ACM/Springer Mobile Networks and Applications*, vol. 3, no. 4, pp. 335–350, Jan. 1998.
- [7] I. Akyildiz, J. Xie, and S. Mohanty, "A survey of mobility management in next-generation all-IP-based wireless systems," *IEEE Wireless Communications*, vol. 11, no. 4, pp. 16–28, Aug. 2004.
- [8] H.-S. Jo, Y. J. Sang, P. Xia, and J. G. Andrews, "Heterogeneous cellular networks with flexible cell association: A comprehensive downlink SINR analysis," *IEEE Trans. on Wireless Communications*, vol. 11, no. 10, pp. 3484–3495, Oct. 2012.
- [9] S. Singh, H. Dhillon, and J. G. Andrews, "Offloading in heterogeneous networks: Modeling, analysis, and design insights," *IEEE Trans. on Wireless Communications*, vol. 12, no. 5, pp. 2484–2497, May 2013.
- [10] Y. Lin and W. Yu, "Optimizing user association and frequency reuse for heterogeneous network under stochastic model," in *Proc. of IEEE Globecom*, Atlanta, GA, Dec. 2013.
- [11] W. Bao and B. Liang, "Structured spectrum allocation and user association in heterogeneous cellular networks," in *Proc. of IEEE INFOCOM*, Toronto, Canada, Apr. 2014.
- [12] Y. Chon, E. Talipov, H. Shin, and H. Cha, "Mobility data collected by Lifemap monitoring system at Yonsei University in Seoul," Downloaded from <http://crawdad.cs.dartmouth.edu/yonsei/lifemap/>, May 2012.
- [13] I. Akyildiz, J. McNair, J. Ho, H. Uzunalioglu, and W. Wang, "Mobility management in next-generation wireless systems," *Proceedings of the IEEE*, vol. 87, no. 8, pp. 1347–1384, Aug. 1999.
- [14] V. Wong and V. Leung, "Location management for next-generation personal communications networks," *IEEE Network*, vol. 14, no. 5, pp. 18–24, Sep/Oct. 2000.
- [15] A. Ghosh, R. Jana, V. Ramaswami, J. Rowland, and N. Shankaranarayanan, "Modeling and characterization of large-scale Wi-Fi traffic in public hot-spots," in *Proc. of IEEE INFOCOM*, Shanghai, China, Apr. 2011.
- [16] Y. Kirsal, E. Gemikonakli, E. Ever, G. Mapp, and O. Gemikonakli, "An analytical approach for performance analysis of handoffs in the next generation integrated cellular networks and WLANs," in *Proc. of International Conference on Computer Communications and Networks*, Zurich, Switzerland, Aug. 2010.
- [17] 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.
- [18] G. Mohimani, F. Ashtiani, A. Javanmard, and M. Hamdi, "Mobility modeling, spatial traffic distribution, and probability of connectivity for sparse and dense vehicular ad hoc networks," *IEEE Trans. on Vehicular Technology*, vol. 58, no. 4, pp. 1998 – 2007, May 2009.
- [19] 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.
- [20] A. H. Zahran, B. Liang, and A. Saleh, "Mobility modeling and performance evaluation of heterogeneous wireless networks," *IEEE Trans. on Mobile Computing*, vol. 7, no. 8, pp. 1041–1056, Aug. 2008.
- [21] 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.
- [22] 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.
- [23] M. M. Zonoozi and P. Dassanayake, "User mobility modeling and characterization of mobility patterns," *IEEE Journal on Selected Areas in Communications*, vol. 15, no. 7, pp. 1239–1252, Sep. 1997.
- [24] 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.
- [25] 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.
- [26] 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.

- [27] 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.
- [28] D. Kotz and K. Essien, "Analysis of a campus-wide wireless network," in *Proc. of ACM MobiCom*, Atlanta, GA, Sep. 2002.
- [29] T. Henderson, D. Kotz, and I. Abyzov, "The changing usage of a mature campus-wide wireless network," in *Proc. of ACM MobiCom*, Philadelphia, PA, 2004.
- [30] M. M. G. M. Voelker, "Access and mobility of wireless pda users," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 9, no. 2, pp. 40–55, Apr. 2005.
- [31] E. H. C. Williamson, "Characterizing and modeling user mobility in a cellular data network," in *Proc. of ACM PE-WASUN*, Montreal, Canada, Oct. 2005.
- [32] I. Rhee, M. Shin, S. Hong, K. Lee, S. J. Kim, and S. Chong, "On the Levy-walk nature of human mobility," *IEEE/ACM Trans. on Networking*, vol. 19, no. 3, pp. 630–643, Jun. 2011.
- [33] Y. Chon, H. Shin, E. Talipov, and H. Cha, "Evaluating mobility models for temporal prediction with high-granularity mobility data," in *Proc. of IEEE International Conference on Pervasive Computing and Communications (PerCom)*, Lugano, Switzerland, Mar. 2012.
- [34] M. Ficek and L. Kencl, "Inter-call mobility model: A spatio-temporal refinement of call data records using a Gaussian mixture model," in *Proc. of IEEE INFOCOM*, Orlando, FL, Mar. 2012.
- [35] P. Baumann, W. Kleiminger, and S. Santini, "How long are you staying? predicting residence time from human mobility traces," in *Proc. of ACM MobiCom*, Miami, FL, Sep. 2013.
- [36] G. Pollini, "Trends in handover design," *IEEE Communications Magazine*, vol. 34, no. 3, pp. 82–90, Mar. 1996.
- [37] A. H. Zahran, B. Liang, and A. Saleh, "Signal threshold adaptation for vertical handoff in heterogeneous wireless networks," *ACM/Springer Mobile Networks and Applications*, vol. 11, no. 4, pp. 625–640, Aug. 2006.
- [38] M. Liu, Z. Li, X. Guo, and E. Dutkiewicz, "Performance analysis and optimization of handoff algorithms in heterogeneous wireless networks," *IEEE Trans. on Mobile Computing*, vol. 7, no. 7, pp. 846–857, Jul. 2008.
- [39] C. W. Lee, L. M. Chen, M. C. Chen, and Y. S. Sun, "A framework of handoffs in wireless overlay networks based on mobile IPv6," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 11, pp. 2118 – 2128, Nov. 2005.
- [40] E. Stevens-Navarro, Y. Lin, and V. Wong, "An MDP-based vertical handoff decision algorithm for heterogeneous wireless networks," *IEEE Trans. on Vehicular Technology*, vol. 57, no. 2, pp. 1243–1254, Mar. 2008.
- [41] J. Hou and D. O'brien, "Vertical handover-decision-making algorithm using fuzzy logic for the integrated radio-and-OW system," *IEEE Trans. on Wireless Communications*, vol. 5, no. 1, pp. 176–185, Jan. 2006.
- [42] S. Singh and J. G. Andrews, "Joint resource partitioning and offloading in heterogeneous cellular networks," *IEEE Trans. on Wireless Communications*, vol. 13, no. 2, pp. 888 – 901, Feb. 2014.
- [43] H. Dhillon, R. Ganti, F. Baccelli, and J. G. Andrews, "Modeling and analysis of K-tier downlink heterogeneous cellular networks," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 3, pp. 550–560, Apr. 2012.
- [44] 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.
- [45] —, "Stochastic geometry and wireless networks, volume 2: Applications," *Foundations and Trends in Networking*, vol. 4, no. 1-2, pp. 1–312, 2009.
- [46] P. F. Ash and E. D. Bolker, "Generalized Dirichlet tessellations," *Geometriae Dedicata*, vol. 20, no. 2, pp. 209–243, Apr. 1986.
- [47] D. Stoyan, W. Kendall, and J. Mecke, *Stochastic Geometry and Its Applications*, 2nd ed. Wiley, 1995.
- [48] W. C. Cheung, T. Q. S. Quek, and M. Kountouris, "Stochastic analysis of two-tier networks: Effect of spectrum allocation," in *Proc. of IEEE ICASSP*, Prague, Czech Republic, May 2011.
- [49] V. Chandrasekhar and J. G. Andrews, "Spectrum allocation in tiered cellular networks," *IEEE Trans. on Communications*, vol. 57, no. 10, pp. 3059–3068, Oct. 2009.
- [50] W. C. Cheung, T. Q. S. Quek, and M. Kountouris, "Throughput optimization, spectrum allocation, and access control in two-tier femtocell networks," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 3, pp. 561–574, Apr. 2012.
- [51] C. Steger, P. Radosavljevic, and J. Frantz, "Performance of IEEE 802.11b wireless LAN in an emulated mobile channel," in *Proc. of IEEE VTC-Spring*, Jeju, South Korea, Apr. 2003.
- [52] K. Xu, B. T. Garrison, and K.-C. Wang, "Throughput modeling for multi-rate IEEE 802.11 vehicle-to-infrastructure networks with asymmetric traffic," in *Proc. of ACM MSWiM*, Miami Beach, FL, Oct. 2011.
- [53] B. Dusza, C. Ide, and C. Wietfeld, "Interference aware throughput measurements for mobile WiMAX over vehicular radio channels," in *Proc. of IEEE WCNC Workshops*, Paris, France, Apr. 2012.
- [54] A. Chambolle, S. Lisini, and L. Lussardi, "A remark on the anisotropic outer Minkowski content," *Advances in Calculus of Variations*, vol. 7, no. 2, pp. 241–266, Apr. 2014.
- [55] H. Federer, *Geometric Measure Theory*. Springer, 1969.