Spatio-Temporal Correlation in the Performance of Cache-Enabled Cellular Networks

Shankar Krishnan

Thesis submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Master of Science
in
Electrical Engineering

Harpreet S. Dhillon, Chair
Walid Saad, Co-Chair
R. Michael Buehrer

June 15, 2016
Blacksburg, Virginia

Keywords: Stochastic Geometry, Correlation, Caching, Mobility, Poisson point process

Spatio-Temporal Correlation in the Performance of Cache-Enabled Cellular Networks

Shankar Krishnan

(ABSTRACT)

Exact characterization and performance analysis of wireless networks should incorporate dependencies or correlations in space and time, i.e., study how the network performance varies spatially and temporally while having prior information about the performance at previous locations and time slots. This spatio-temporal correlation in wireless networks is usually characterized by studying metrics such as joint coverage probability at two spatial locations/time slots or spatio-temporal correlation coefficient. While developing models and analytical expressions for studying the two extreme cases of spatio-temoral correlation - i) uncorrelated scenario and ii) fully correlated scenario are easier, studying the intermediate case is non-trivial. In this thesis, we develop realistic and tractable analytical frameworks based on random spatial models (using tools from stochastic geometry) for modeling and analysis of correlation in cellular networks.

With an ever increasing data demand, caching popular content in the storage of small cells (*small cell caching*) or the memory of user devices (*device caching*) is seen as a good alternative to offload demand from macro base stations and reduce backhaul loads. After providing generic results for traditional cellular networks, we study two applications exploiting spatio-temporal correlation in cache-enabled cellular networks. First, we determine the optimal cache content to be stored in the cache of a small cell network that maximizes the *hit probability* and minimizes the *reception energy* for the two extreme cases of correlation. Our results concretely demonstrate that the optimal cache contents are significantly different for the two correlation scenarios, thereby indicating the need of correlation-aware caching strategies. Second, we look at a *distributed caching scenario* in user devices and show that spatio-temporal correlation (user mobility) can be exploited to improve the network performance (in terms of coverage probability and local delay) significantly.

Spatio-Temporal Correlation in the Performance of Cache-Enabled Cellular Networks

Shankar Krishnan

(GENERAL AUDIENCE ABSTRACT)

Researchers in the wireless domain are always on the look out to try and improve the network performance of wireless systems and networks. A common approach undertaken is to exploit diversity in wireless channels, either temporal diversity by retransmitting a packet multiple times for higher reliability or use multiple antennas to exploit the spatial diversity. However the gains in actual systems may be significantly lower due to dependencies or correlation across temporal as well as spatial domains. Ignoring this spatio-temporal correlation in the performance analysis of wireless networks may lead to incorrect characterization of wireless networks and often results in overestimating the network performance. To give an example, consider a scenario where a node receives a packet in error during a certain transmission. It can be seen that it is more likely that the packet is received again in error compared to a successful reception in the immediate time slots. This is because of the similar spatial distribution of transmitters or interferers (local neighbourhood) across the two time slots.

Incorporating this spatio-temporal correlation, we develop an analytical framework to study the network performance of cellular networks in terms of correlation-based metrics such as *joint coverage probability* and *interference correlation coefficient*. Having looked at a generic cellular network, we then study two applications exploiting spatio-temporal correlation in *cache-enabled cellular networks* - i) determine the optimal *cache* in small cell networks and ii) study effect of user mobility in distributed caching networks. Both the applications develop analytical models for studying cache-enabled networks and show that spatio-temporal correlation can be exploited to improve the overall network performance.

Acknowledgments

First and foremost, my heartfelt thanks to Dr. Harpreet Dhillon, my advisor and mentor through my grad school life. This thesis would not have been possible if not for your invaluable guidance in every single step of my research, showing new ways to tackle problems when I was stuck and your continued motivation to push the boundaries. You have always set the standards high and this experience working under you for the last two years has definitely changed my life in many ways. Thanks again!

I also thank Dr.Walid Saad for his valuable suggestions towards this thesis and Dr.Michael Buehrer for agreeing to be a part of my thesis defense committee.

Next, a big shout out to all my Durham-470 mates. This place has been my second home in Blacksburg and the research would not have been easier without you all. Special thanks to Mehrnaz, who has been a mentor and a great friend guiding me through all the highs and lows of research. To all my lab mates: Chiranjib, Mustafa, Nisanth, Priyabrata, Surabhi, Tapan and Vishnu, the last two years would not have been special without your support and fun times. Will definitely miss the group meetings. Also to Akshay, Deven, Mehdi, Mohammad, Omid, Raghu and other folks in Wireless@VT, it was great knowing you all and being a big part of my grad life.

Also a big thanks to my buddies at home : Karthik, Kaushik and Rahul for always being there during fun and hard times. Finally, a special thanks to my parents and sister for being just a Skype call away and making me not miss home.

Contents

List of Figures

Chapter 1

Introduction

Performance analysis of wireless networks via analytical modeling enables assessment of large, complex systems before actual prototyping. There have been primarily three approaches to model and analyze the performance of wireless networks. The first approach assume no network geometry and ignore dependencies in space and time to model wireless channels and user traffic. This approach is not realistic as it ignores all dependencies in the wireless network. The second approach assume a fixed network geometry [1], [2], e.g., hexagonal cells in traditional cellular networks or a regular lattice. This approach yields results that are only valid for this network geometry, hence lack in generality and are also sometimes hard to analyze. The third approach assume a spatial stochastic model for the node locations [3], [4] and usually yields general and accurate results by averaging over the likely network topologies.

Stochastic geometry has recently emerged as a popular tool for the modeling and analysis of large-scale wireless communication systems, such as wireless ad hoc and cellular networks [5, 6]. Irrespective of the type of wireless network being considered, the main idea behind these analyses is to model the locations of both the transmitters (Txs) and receivers (Rxs) as point processes, often *independent* homogeneous Poisson point processes (PPP) [5], and study *first-order* performance metrics, such as coverage probability, average rate, or interference

distribution, as observed by a randomly chosen Rx, termed typical Rx [5]. In a PPP, the number of nodes in a certain region of area A is Poisson distributed with parameter λA, where λ is the intensity. Also, the number of nodes in disjoint regions is mutually independent. However, all these first order results examine the network at one location and time instant and then take an average to obtain the temporal or spatial statistics. Finer characterization of a wireless network performance requires the joint analysis of observations made at two different spatial locations and/or different time slots. In other words, one should also look at answering *second order* questions such as the joint distribution of interference at two spatial locations or the joint success probability over two transmissions. These questions are about dependencies and *correlations* in the network. They are important but frequently ignored, either explicitly or implicitly.

Most of the prior art either assumes no dependence or complete correlation in the user locations over time. This captures only the extremes of either no mobility (complete correlation), or highly mobile users (little or no correlation). In many realistic settings, however, there is some finite mobility and it is therefore important to study the network performance when there is nontrivial correlation in the transmitter locations. Addressing these second-order properties with nontrivial correlation is important and although challenging can be analyzed and studied, which we do over the course of this thesis.

This introductory chapter is divided into three main parts. The first part, spanning Sections 1.1 and 1.2 provides a background on the different types of correlation in wireless networks and discusses the challenges behind studying non-trivial correlation in mobile wireless networks (where nodes are mobile), specially cellular networks. The second part, covered in Section 1.3 provides a broad overview of cache-enabled cellular networks including small-cell and device caching and briefly discusses different applications of studying spatio-temporal correlation in cache-enabled cellular networks. Section 1.4 forms the third part, where the key contributions of this thesis are summarized.

1.1 Correlation in wireless networks

In wireless networks, the quality of a radio link varies over time. Such time varying nature of the wireless environment is mainly caused by two factors. First, mobility of user devices and obstacles changes the multipath propagation w.r.t. time. As a result, there are fluctuations in received signal powers over time. Second, the data traffic sent by users vary over time resulting in fluctuations of interference powers and therefore the signal-to-interference ratio at a given receiver. Performance analysis of wireless networks requires a comprehensive understanding of these fluctuations of received signal/interference over time and space, also referred as *temporal* and *spatial correlation* respectively. Informally speaking, correlation indicates what difference to expect between two signal samples, i.e., how well can one value be estimated when knowing the other. We now briefly discuss the different types of correlation in wireless networks.

1.1.1 Temporal correlation

The temporal correlation measures the *similarity* of one signal over time. Consider a receiver equipped with a single antenna, then the temporal correlation explores the similarity in received signal or interference powers at the same receiver across different time slots. The signal or interference power levels in different time slots are correlated, since the transmitters/ interferers are chosen from the same set of nodes over different time slots. Several communication techniques and protocols exploit this temporal correlation of wireless signals to design re-transmission techniques. More specifically, the temporal correlation of a signal tells us how long to wait until the environment's characteristics change significantly, such that a retransmission makes sense. Usually, signals received over two time slots separated by a time span much longer than the channel coherence time have little to no correlation. The basic idea in retransmissions is as follows: if a transmission fails at a certain time slot, it might well succeed at a later time slot. However due to temporal correlation, it is observed

that a transmission is more likely to fail in a given time slot if there were failures in the previous time slots. Hence, it is important to design smart and correlation-aware retransmission schemes, where the network doesn't retransmit immediately following a failure or have a few silence slots before retransmission.

1.1.2 Spatial correlation

Correlation between different receive antennas

Consider a wireless network where a typical receiver is equipped with co-located multiple antennas. There exists a spatial correlation among the different antennas as the signals received by the different antennas come from the same source of transmitters. The spatial seperation between the multiple antennas decides the extent of correlation in this scenario. Short antenna separations increase the spatial correlation as adjacent antennas will receive similar signal components. Also, rich multipath propagation decreases the spatial correlation by spreading the signal such that multipath components are received from many different spatial directions.

Correlation between different receivers

This refers to the spatial correlation between different receivers which are not co-located i.e. among receivers separated by a larger distance (atleast a few wavelengths apart). Again, spatially proximate receivers are highly correlated with the degree of correlation increasing with decreasing inter-receiver separation. With recent wireless networks utilizing relay and cooperative transmissions, studying this type of spatial correlation, say between a spatially separated transmitter and relay, is important to characterize the network performance accurately.

1.2 Spatio-temporal correlation in mobile networks

In wireless networks, mobility of the desired transmitter induces spatial diversity and the mobility of the interferers decreases the correlation among the spatial locations. Therefore, node mobility affects the spatio-temporal correlation of interference and outage and hence affects the network performance in Poisson networks. Since interference is the main performance-limiting factor in most wireless networks, we primarily study and characterize the spatio-temporal correlation of interference in wireless networks.

There are three major sources of randomness that affect the interference in wireless networks. The first is multipath fading, which is the time variation of the channel strength due to small time-scale effects. Secondly, the data traffic sent by devices and base stations varies over time resulting in fluctuations of the received interference powers. Thirdly, interference also depends on the spatial location of interferers i.e. their *node placement* in the wireless network. In this thesis, we only focus on the the correlation caused by the spatial distribution of interferers, with i.i.d. assumptions for multipath fading and data traffic statistics.

Let us now look how node mobilities affect the spatio-temporal correlation in wireless networks. The correlation between node locations in different time slots is zero only if a completely new realization of the node placement is drawn in each time slot. Network models assuming independent realizations are impractical since the node velocities cannot be infinite. If the node placement follows a certain type of distribution such as a PPP in each time slot and the nodes do not have infinite mobility, the node locations in different time slots are correlated. An extreme case is a static but random network, where the nodes positions are completely correlated, since the nodes do not move after their initial placement. For realisitic scenarios, the locations of a node always show a certain degree of correlation in different time slots, since the node speed is finite. In this thesis, we provide generalized expressions for the network performance for finite user mobility in addition to the two extreme cases: static and highly mobile networks.

In cellular networks, the interference received by a mobile user comes from the set of randomly located and temporally/spatially correlated BSs. In this thesis, we focus on the interference correlation caused by finite mobility of a user and the resulting network performance in cellular networks. Developing a tractable model for studying this correlation for a finitely mobile user is challenging from different aspects, the details of which are mentioned through the course of this thesis. Apart from traditional cellular networks, studying this spatio-temporal correlation for a mobile user also finds other applications. One relevant application is in cache-enabled cellular networks, where different tiers of cellular network are equipped with storage units (caches) and stored with the popular files of the network, thereby limiting the access to the core network through backhaul links. The quantification of spatio-temporal correlation is important for such cache-enabled networks as it can help determine several design aspects to optimize the network performance. Some of them studied in this thesis include determining the optimal cache for maximizing the cache *hit probability* for a mobile user and exploiting spatio-temporal correlation (user mobility) to improve the coverage and local delay of a *distributed caching* network. Further details on cache-enabled cellular networks are discussed next.

1.3 Cache-enabled Cellular networks

Providing quality broadband experience *everywhere at all times* is widely regarded as one of the key objectives (also termed *use-cases*) for future 5G networks [7, 8]. Among other improvements, this requires a significant increase in the capacity of current cellular networks. One reasonable way to achieve this is by reusing spectrum more aggressively through dense deployment of several types of low-power base stations, collectively called *small cell base stations (SCBSs)*. The resulting network is referred to as an ultra-dense heterogeneous cellular network (HetNet) [9]. While, in principle, this addresses the capacity problem, providing reliable backhaul to each SCBS in an ultra-dense HetNet becomes prohibitively expensive. There are two complementary approaches usually suggested to overcome this

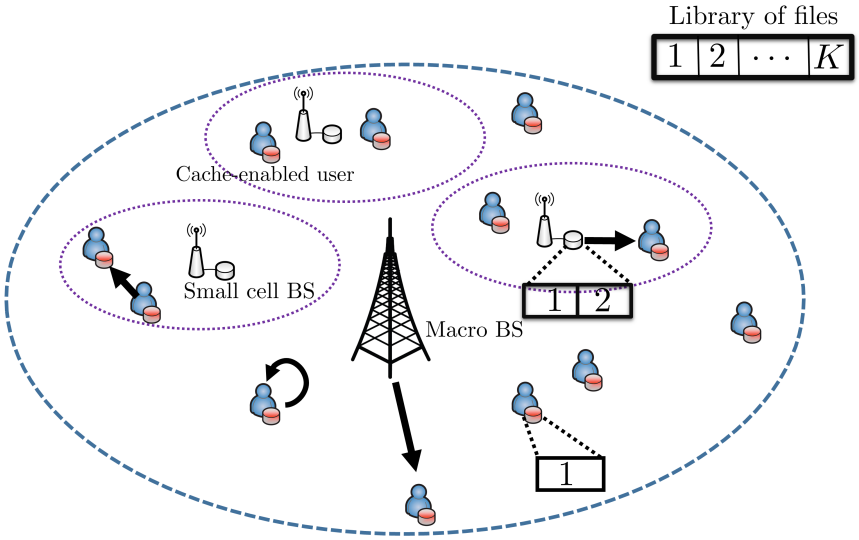

Figure 1.1: Cache-enabled Heterogeneous Network, where a macro base station is overlaid with cache-enabled small cells and cache-enabled user terminals.

challenge: (i) self-backhauling (also called wireless backhauling) [10], and (ii) small cell caching or device caching [11–14]. Fig. 1.1 depicts a cache-enabled cellular network, which comprises of macro base stations overlaid with cache-enabled small cells and user terminals with storage units alloted for caching. The latter is the main focus of Chapter 3 and Chapter 4 in this thesis.

1.3.1 Small cell caching

The basic idea behind small cell caching is to download popular content (mainly video) automatically in the cache of SCBSs at off-peak hours, which can then be delivered to the users during peak hours. There are several factors that enable this idea. First, the cost of memory is decreasing sharply, which makes it realistic to provide dedicated small form-factor high-capacity cache storage units for each SCBS. Second, high degree of correlation has been observed in the video demand [15], which means caching content in the SCBSs

avoids repeated downloads of the same content from the core network, thus resulting in a much more efficient use of backhaul. Third, having content of interest at the *network edge* reduces latency, which is another key use-case for 5G networks [7]. On the flip side, the key challenge in designing such a network is to determine the content that should be cached at each SCBS. Due to the amount of content available on the Internet, this problem seems hopeless at first. However, it has been observed in several studies that only a small fraction of the total content, termed *popular content*, is accessed by a large fraction of users, while a large amount of content remains unpopular and sparsely requested [12]. This allows one to focus on the *library* of popular files, where each file is endowed with a *popularity distribution*, which is empirically determined to follow Zipf's law [15, 16].

Small cell caching problem has two main dimensions: (i) learning the *library* of popular files, and (ii) determining the subset of library content to be placed on each SCBS in order to optimize the system performance. In this thesis, we focus on the latter assuming that the library of popular files has already been determined. Note that while the library of popular files is a small fraction of the total content available on the Internet, it is still, in general, not possible to cache the whole library at each SCBS due to the limited capacities of cache storage units. Therefore, one can only cache a subset of the library on each SCBS. This means that even if the file requested by a particular user is a part of the library of popular files, the user may still not be able to download it because it may not be cached at any SCBS located in the user's vicinity. In this thesis, we show that this problem is naturally addressed if the user is mobile and is allowed to access the cache multiple times, thereby exploiting the spatio-temporal correlation in the cached network.

1.3.2 Device caching

Increasing spatio-temporal correlation in the data demand makes it attractive to cache popular content directly on the user devices so that it can be delivered on demand to the neighboring devices through device-to-device (D2D) communications. By caching content

directly on the devices, and by exploiting D2D communication between devices, the devices themselves can form an effective Content Distribution Network (CDN). D2D communication has the advantage of a limited investment requirement, since it can be run on the existing population of smart phones and devices. However, D2D is not without its challenges. Key among these are: user privacy and security, resource management and interference control, and effective distribution of popular content among the devices. Clearly the full potential of D2D communications can only be realized if the content desired by the user is available on another device in proximity. Thus the storage of content in various devices should be tailored according to content popularity.

Due to the limited storage capacity, each device can only cache a part of the popular content resulting in a *distributed caching* network. In a distributed caching device-to-device (D2D) network, user's file of interest is cached as several portions in the storage of other devices in the network and obtained successively through D2D communication. A higher likelihood of dominant interferers for the file portions cached farther away from the user's location presents a bottleneck in obtaining the entire file. In other words, the low coverage probability and large delay in obtaining the farther cached file portion in a distributed caching network limits the network performance. However, we again show that user mobility addresses this issue and significantly improves network performance.

1.4 Contributions

There are two main challenges in studying spatio-temporal correlation in cache-enabled cellular networks : (i) to develop system models that capture the effect of correlation (temporal correlation due to re-transmissions or spatial correlation due to user mobility) in the performance of a typical user in cellular networks and (ii) to develop analytical frameworks to study metrics like coverage probability, local delay and reception energy of a typical user in a cache-enabled cellular network. This thesis develops a comprehensive framework to tackle

these challenges. The main contributions are summarized below.

1.4.1 Spatio-temporal correlation in cellular networks

In Chapter 2, we provide an exact analytical framework to characterize the spatio-temporal interference correlation in cellular networks and the network performance in terms of joint coverage probability at two spatial locations separated by distance v. Incorporating the possibility of handoff when a user moves in a cellular network, we derive exact expressions for spatio-temporal correlation coefficient and joint coverage probability for any distance separation v. The exact analysis is not straightforward and involves a careful treatment of the neighbourhood of the two spatial locations and the resulting handoff scenarios. Our analysis shows that the interference correlation and joint coverage probability decreases with v, with complete correlation for a static user and an uncorrelated (independent) scenario for the highly mobile user. Further insights are also provided by studying the effect of few network/channel parameters such as BS density and path loss on the interference correlation.

1.4.2 Effect of retransmissions on optimal cache in cache-enabled small cell networks

In Chapter 3, we focused on the problem of determining the *optimal* cache for a cache-enabled small cell network. Unlike prior approaches where the focus was on determining the optimal cache content assuming the users to be static, we focused on the case where the user is mobile, and can hence access its file of interest at multiple possible locations along its trajectory. Our results concretely demonstrated that the optimal cache contents are significantly different for the mobile and static cases. While the static case tends to cache the most popular files in the library on each small cell, mobility de-emphasizes the importance of popularity of files and allows the small cells to cache content in a more balanced way. The overall network performance (in terms of hit probability and reception energy) was shown to be much better

for the mobile case compared to the static case. This is due to the fact that when the user is mobile it comes across more *unique* small cells and hence more unique caches, which increases the probability of it being close to the small cell that has its file of interest.

1.4.3 Analyzed network performance of D2D networks with distributed caching

In Chapter 4, a distributed caching device-to-device (D2D) network is considered, where a user's file of interest is cached as several portions in the storage of other devices in the network and obtained successively through D2D communication. A higher likelihood of dominant interferers for the file portions cached farther away from the user's location presents a *bottleneck* in obtaining the entire file. With device locations modeled as a PPP, we derive new analytical expressions for the coverage probability and *local delay* of the file cached farther from the typical device in a 2-file distributed caching network as a function of the distance v moved by the user from its initial location. For a uniform caching network with a maximum storage of one file portion per device, we show that mobility significantly improves network performance. Also, two equivalent viewpoints for the analysis of coverage probability and average delay in a distributed caching network are discussed for varying cache probabilities and device activity factors.

1.5 Organization

The technical contributions of this thesis are covered in Chapters 2 through 4. Chapter 2 develops an analytical framework for studying spatio-temporal correlation in any cellular network with or without caching capabilities. Exact analytical expression for measuring the network performance in terms of joint coverage probability and interference correlation coefficient are developed. Dealing with only extreme cases of correlation (fully correlated

and uncorrelated scenarios), Chapter 3 studies the network performance in terms of hit probability and reception energy of a cache-enabled small cell network. An optimal cache that maximizes the hit probability and minimizes the reception energy is also developed by solving an appropriate optimization problem. The final contribution of this thesis is presented in Chapter 4, where we present a D2D network model with distributed caching and study the network performance in terms of coverage probability and local delay. The effect of mobility (correlation) on the performance of this distributed cache network is also studied by using tools developed in Chapter 2. The thesis concludes with Chapter 5, which summarizes the key contributions and discusses promising future directions of research in cache-enabled wireless networks.

Chapter 2

Spatio-temporal correlation in cellular networks

Stochastic geometry has recently emerged as a popular tool for the modeling and analysis of large-scale wireless communication systems, such as wireless ad hoc and cellular networks [5, 6]. Irrespective of the type of wireless network being considered, the main idea behind these analyses is to model the locations of both the transmitters (Txs) and receivers (Rx) as point processes, often *independent* homogeneous Poisson point processes, and study *first-order* performance metrics, such as coverage probability, average rate, or interference distribution, as observed by a randomly chosen Rx, termed typical Rx [5]. While this approach of confining the analysis to a single observation point as well as almost always a single time-frequency resource *slice* is useful to get first-order insights, it is not sufficient for the characterization of spatio-temporal dependence (correlation) in the performance of a randomly chosen user in a wireless network. This requires the joint analysis of the observations made at two different spatial locations and/or different time-frequency resource *slices* which is known to be significantly more challenging than the more popular approach described above. As discussed next in detail, while this dependence has already been characterized for ad hoc networks, the same is not true for cellular networks for which the analysis is known to be

far more challenging. Developing new tools to facilitate the exact characterization of spatio-temporal interference correlation as well as joint coverage probability in cellular networks is hence the main goal of this chapter. Please see [17] for a recently submitted journal version of this chapter.

2.1 Related work and Motivation

Correlation in interference observed at different locations or the same location at different times has been studied extensively for wireless networks, *albeit* almost exclusively in the context of wireless *ad hoc* networks. In this ad hoc network setup, wireless nodes are usually assumed to transmit independently according to a random access scheme, such as ALOHA, with a certain transmit probability [5, 6]. Interference is spatially correlated as it originates from the same set of transmitters even in the presence of independent fading. Similarly, interference is also temporally correlated since a subset of the same set of nodes transmit in different time slots. The authors in [18] first characterized this interference correlation in ad hoc networks in terms of spatio-temporal correlation coefficient and showed that it is directly proportional to the random access probability and inversely proportional to the second moment of fading power gain. The authors in [19] extended this work and derived the joint probabilities of successful packet delivery at multiple receivers under spatially and temporally dependent interference. They showed that interference correlation significantly impacts the packet delivery probability. Along the same lines, [20] investigated interference correlation in multi-antenna receivers and showed that a diversity loss occurs due to interference correlation. Other related works include studying the network performance (in terms of outage probability and diversity order) for a decode-and-forward, cooperative relaying system [21–24] under spatially and temporally correlated interference. The effect of interference correlation on the performance of multi-antenna communication systems under Maximal Ratio Combining is also studied recently in [25, 26].

For the rest of this discussion, we note that the prior art on interference correlation can be broadly classified into two categories. The first line of work deals with the study of temporal correlation in *static networks*, where nodes are static or have low mobility. Most of the works discussed above fall in this category. Additionally, considering the temporal correlation of interference over different time slots, [27] derived closed form expressions for joint outage/success probability over n time slots (transmissions) and showed that temporal diversity gain due to retransmissions diminishes with high interference correlation. Along the same lines, [28] investigated temporal correlation in the interference power for a correlated fading (flat fading, Rayleigh block fading) and correlated user traffic (slotted ALOHA) scenario. A key step in these studies is the characterization of moments of conditional success probability (conditioned on the point process), which have also been used recently for the derivation of *meta distribution* of the signal to interference ratio in both ad hoc and cellular networks in [29, 30].

The second line of work deals with mobile networks, where mobility of nodes introduces randomness and thereby decorrelates interference across space and time. If the nodes are considered highly mobile ($v \to \infty$) in each time slot, the set of possible interferers also change rapidly and thus the interference becomes completely uncorrelated over time. However in real life scenarios, the nodes in a wireless network have finite mobility [31] and it is thus important to study the interference correlation in finite mobile networks. For a mobile ad hoc network, the authors in [32] studied temporal correlation in interference and outage under different mobility models such as Random Way point, Random Walk and Discrete-time Brownian motion. Specifically, they characterized correlation coefficient for interference and showed that correlation decreases with the increase in the mean speed of the nodes. More recently, [33] captured the effect of mobility on interference and outage correlation in finite networks (network with finite boundaries) and showed that interference correlation is location-dependent, being higher close to the network edge. Capturing the correlation in user locations in wireless networks, the authors in [34] studied interference correlation in clustered networks (modeled as Matern or Thomas cluster process) and showed that

clustering of interferer locations enhances interference correlation.

Although there has been substantial work quantifying interference correlation in wireless ad hoc networks, there has been limited work studying correlation in cellular networks. Taking into account the temporal/spatial correlation in cellular networks, [35] analyzed the benefits of inter-cell interference coordination (ICIC) with BS coordination and intra-cell diversity (ICD) with selection combining in cellular networks. Their analysis of ICD showed that a diversity gain can always be obtained in a cellular setting with strongest BS association, in contrast with the conclusion drawn from ad hoc networks in [27]. Studying spatiotemporal cooperation among BSs in a heterogeneous cellular network, [36] studied different diversity exploiting techniques such as joint transmission and base station silencing. The authors in [37] studied the correlation of interference and link successes in heterogeneous cellular network with multiple tiers (different transmit powers and BS densities) and different cell association policies. However their analysis is not accurate as the authors also include the received power from the serving BS in the interference analysis. In this chapter, we provide the exact analysis of interference correlation in cellular networks.

Before we describe our contributions in the next subsection, we provide a brief overview of the key differences in the analysis of ad hoc and cellular networks and explain why cellular analysis is more challenging. Ad hoc networks are usually modeled as Poisson bipolar networks [38] with fixed link distance between the transmitter and the node of interest (receiver). The interference field can therefore be modeled as an infinite homogeneous PPP. However, in cellular networks, the serving BS has to be chosen based on some cell association strategy, such as maximum average received power based [39] or highest instantaneous SINR based [40]. This has an important implication on the modeling of interference field. In order to fix this key idea, let us consider *maximum average received power based cell association* in a single-tier cellular network in which the typical user connects to its closest BS. For this association strategy, the interference comprises of all active BSs farther than the closest BS (serving BS) and hence the location of this serving BS plays an important role in the analysis of any performance metric that depends upon the received power of the desired signal and/or

the interfering signal. Moreover, to study interference correlation in two spatial locations (a finite distance apart), it is important to characterize the distance of the serving BS at both spatial locations. Also note that there is a certain probability that the serving BS at the second location is the same as the one in the first location. This dependence among serving BSs in cellular networks is characterized by the *handoff rate* [41], which is the probability that the user is handed off to a new serving BS as it moves from one location to another. This characterization of handoff is not required in ad hoc networks. In this work, we present exact characterization of both interference correlation as well as the joint coverage probability while incorporating all such dependencies. To the best of our knowledge, we are first ones to characterize these correlation-based metrics for cellular networks.

2.2 Contributions and Outcomes

2.2.1 Distance distribution of serving BS at two spatial locations

Incorporating spatio-temporal corrrelation, we study the network performance when a typical user moves a distance v from its initial location in a cellular network. The user follows closest BS association i.e. connects to the closest BS at both user locations. We first identify that the user displacement can result in two scenarios: i) *No Handoff*: the user is associated with the same BS at both the user locations (same BS is the closest at both locations), and ii) *Handoff*: A different BS is closer to the user at the second location and therefore handoff occurs. For both these scenarios, we provide joint distribution of the distances from the two locations to their respective closest BSs.

2.2.2 Spatio-temporal interference correlation coefficient for cellular networks

After characterizing distance distributions, we first study the spatio-temporal interference correlation coefficient of a typical user in a cellular network under closest BS association policy by deriving expressions for mean, variance and first order cross moments of interference at two user locations in the network. We then show that interference correlation decreases with the distance v between the two spatial locations, becoming uncorrelated at locations far apart. As a special case, we derive an easy-to-use expression for the temporal interference correlation coefficient, i.e., when the user is static at a given spatial location. The temporal correlation coefficient in cellular networks is shown to exhibit a *bell curve* relationship w.r.t. BS density and decrease with path loss for large v. This *bell curve trend* does not exist in the case of ad hoc networks.

2.2.3 Exact analysis of joint coverage probability at two spatial locations

Joint coverage probability at two spatial locations completely characterizes the dependence in link successes (coverage) at the two locations in the network. We develop exact expressions for the joint coverage probability for two spatial locations (separated by a distance v) under two handoff strategies: i) *handoff skipping* and ii) *conventional handoffs*. With handoff skipping, a user initially connected to its closest BS at a given location continues to be associated with the same BS irrespective of the distance it moves and skips all possible handoffs to a new serving BS. Handoff skipping is more relevant to ultra-dense networks, where a user can save handoff delays/overheads by skipping certain handoffs. Next, we look at the conventional handoff scenario where a user always switches its association to the closest BS as it moves along in the network. For both handoff strategies, we show that joint coverage probability at two spatial locations decreases with the separation in the two

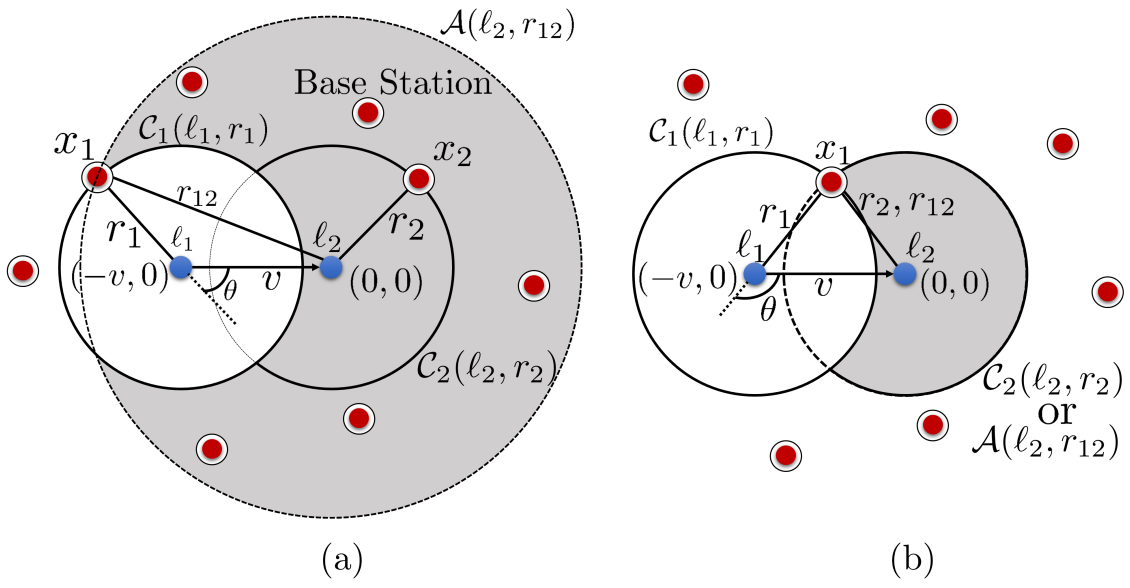

Figure 2.1: System model when a typical user (denoted by *blue dot*) moves from location 1 (ℓ_1) to location 2 (ℓ_2). (a) *Handoff scenario*: user is handed off to a new serving BS x_2 at ℓ_2 as it is closer than the serving BS x_1 at ℓ_1, and (b) *No Handoff scenario*: user is served by the same BS x_1 at both locations ℓ_1 and ℓ_2 and therby no handoffs occur.

locations (i.e. decrease in correlation).

2.3 System Model

We consider a cellular network where the base stations are modeled as a homogeneous PPP Φ of intensity λ. As noted already, our main objective is to study how the interference experienced by a typical user is correlated across two spatial locations in this network. At any spatial location in the network, the typical user follows the *closest BS association policy* i.e. connects to its closest BS as that maximizes its average received power. Also, we study how the interference is correlated temporally as the typical user is subject to interference from the same subset of BSs across time.

2.3.1 System Setup and Key assumptions

Consider a typical user that connects to its closest BS $x_1 \in \Phi$ at location 1 (ℓ_1) as shown in Fig. 2.1. The user now moves a distance v at an angle θ to the serving BS x_1 at ℓ_1 and shifts to location 2 (ℓ_2). With this user displacement, two scenarios can arise : i) a different BS $x_2 \in \Phi$ is located closer to the user at location 2 than the serving BS x_1 at location 1 (Fig. 2.1 (a)) or ii) the BS x_1 (serving BS at location 1) is still the closest BS to the user at location 2 (Fig. 2.1 (b)). The first scenario corresponds to the case where handoff occurs to a new serving BS x_2 while the second scenario results in no handoff and a continued connection to the BS x_1. In this work, we study how the interference is correlated across these two user locations ℓ_1 and ℓ_2. Before that, we need to define the distance distributions of the serving BS at the two spatial locations and understand the above mentioned scenarios (handoff and no handoff) in a bit more detail.

Let R_1 and R_2 be the random variables denoting the distance of the serving (closest) BS to the user at locations 1 and 2 respectively, with r_1 and r_2 being their realizations. As can be seen from Fig. 2.1, the user's location 2 is at a distance $r_{12} = \sqrt{r_1^2 + v^2 + 2r_1 v\cos\theta}$ (using law of cosines) from BS x_1. Let Θ be a uniform random variable in $[0, \pi]$ denoting the angle of user displacement as the user moves from ℓ_1 to ℓ_2 with θ being its realization. Therefore its PDF is given as $f_\Theta(\theta) = 1/\pi$. Denote by $\mathcal{C}_1(\ell_1, r_1)$ - a circle centered at ℓ_1 with radius r_1 and $\mathcal{C}_2(\ell_2, r_2)$ - a circle centered at ℓ_2 with radius r_2. And let $\mathcal{A}(\ell_2, r_{12})$ be a circle centered at ℓ_2 with a radius of r_{12}, which will be used for the handoff analysis. Before that, we state an intermediate result which will be useful in our analysis.

Definition 1. *Consider two intersecting circles $\mathcal{C}_1(\ell_1, r_1)$ and $\mathcal{A}(\ell_2, r_{12})$ with centers separated by a distance v and an angle θ as shown in Fig. 2.1. The shaded region in the figure is denoted by $\mathbf{C}(\ell_1, r_1, \ell_2, r_{12}, v)$ and its area given as*

$$|\mathbf{C}(\ell_1, r_1, \ell_2, r_{12}, v)| = r_{12}^2\left[\pi - \theta + \sin^{-1}\left(\frac{v\sin\theta}{r_{12}}\right)\right] - r_1^2(\pi - \theta) + r_1 v\sin\theta. \tag{2.1}$$

Proof. The area of the shaded region is

$$\begin{aligned}|\mathbf{C}(\ell_1, r_1, \ell_2, r_{12}, v)| &= |\mathcal{A}(\ell_2, r_{12}) \setminus \mathcal{A}(\ell_2, r_{12}) \cap \mathcal{C}_1(\ell_1, r_1)| \\ &= \pi r_{12}^2 - |\mathcal{A}(\ell_2, r_{12}) \cap \mathcal{C}_1(\ell_1, r_1)|\end{aligned}$$

where $|.|$ denotes the area, and the result follows by using the area of intersection of two circles as done in [41, Theorem 1]. ■

As the user moves from ℓ_1 to ℓ_2, whether handoff occurs (Fig. 2.1(a)) or not (Fig. 2.1(b)) is dictated by the existence of a BS within the circle $\mathcal{A}(\ell_2, r_{12})$. Let H be the event that handoff occurs as the user moves from ℓ_1 to ℓ_2 and $\bar{H}$ be the complementary event (no handoffs occur). The probability of handoff $\mathbb{P}(H)$ is derived in [41, Theorem 1] for a similar setup and is restated below in Lemma 1 for completeness. We use a shorthand notation $\mathcal{C}_1$, $\mathcal{C}_2$ and $\mathcal{A}$ for denoting the circles defined before for the sake of simplicity.

Lemma 1. *Conditioned on r_1 and θ, the probability of handoff as the user moves a distance v at an angle θ from location 1 to 2 in a PPP of BS density λ is*

$$\mathbb{P}(H|r_1, \theta) = 1 - \exp\left(-\lambda\left(r_{12}^2\left[\pi - \theta + \sin^{-1}\left(\frac{v\sin\theta}{r_{12}}\right)\right] - r_1^2(\pi - \theta) + r_1 v \sin\theta \right)\right). \quad (2.2)$$

Proof. From Fig. 2.1, for a typical user initially connected to its closest BS x_1 at distance r_1 and moving to a new location ℓ_2 at distance r_{12} from BS x_1, a handoff does not occur if there is no BS closer than r_{12} to the user at ℓ_2, hence:

$$\begin{aligned}1 - \mathbb{P}(H|r_1, \theta) &\overset{(a)}{=} \mathbb{P}(N(|\mathcal{A}|) = 1 | N(|\mathcal{A} \cap \mathcal{C}_1|) = 1) \\ &= \mathbb{P}(N(|\mathcal{A} \setminus \mathcal{A} \cap \mathcal{C}_1|) = 0) \\ &= \exp(-\lambda\, |\mathbf{C}(\ell_1, r_1, \ell_2, r_{12}, v)|)\end{aligned}$$

where (a) follows because only one BS x_1 lies in the region $\mathcal{A}$ for a handoff scenario and the result follows by using $|\mathbf{C}(\ell_1, r_1, \ell_2, r_{12}, v)|$ from Definition 1. ■

We now state some observations about the system model depending on the user displacement v in Remarks 1 and 2.

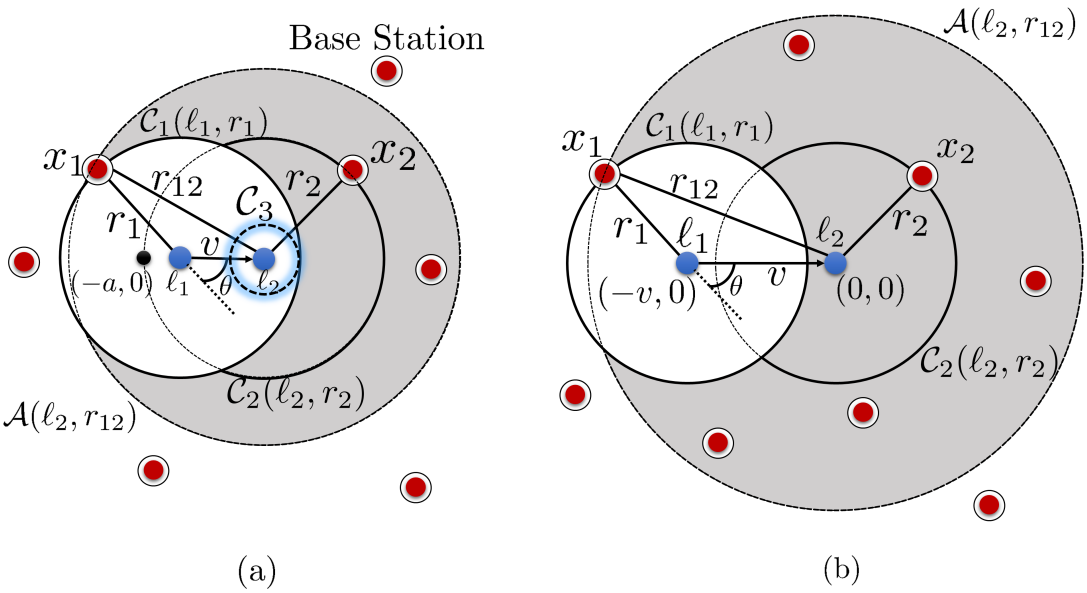

Figure 2.2: Different scenarios based on user displacement v from ℓ_1 to ℓ_2. (a) Scenario 1 ($v < r_1$), where $\mathcal{C}_3(\ell_2, r_1 - v)$ is an exclusion zone and (b) Scenario 2 ($v \geq r_1$), where there is no exclusion zone.

Remark 1. *Conditioned on the occurence of handoff (Fig. 2.1 (a)) when a user moves a distance v from ℓ_1 to ℓ_2 , one of the following 3 cases arises for circles $\mathcal{C}_1$ and $\mathcal{C}_2$: (i) Case 1: Disjoint circles ($v \geqslant r_1 + r_2$), (ii) Case 2: Intersecting circles ($r_2 - r_1 < v < r_1 + r_2$), and (iii) Case 3: Engulfed circles ($v \leqslant r_2 - r_1$). This insight will be useful for the analysis.*

Remark 2. *For the system model shown in Fig. 2.1, there exists two scenarios based on the distance v moved by the user from ℓ_1 to ℓ_2: i) Scenario 1 ($v < r_1$) and ii) Scenario 2 ($v \geq r_1$) (see Fig. 2.2). For scenario 1, even after user displacement, the user is still inside $\mathcal{C}_1$ and hence no BS lies can lie within a distance $r_1 - v$ from ℓ_2 i.e. the closest BS is atleast a distance of $r_1 - v$ from ℓ_2 or $r_2 \geq r_1 - v$. In scenario 2, the user moves a larger distance ($v \geq r_1$) and no such condition exist for the serving distance R_2 i.e. $r_2 \geq 0$. We define $z_1 = \max(0, r_1 - v)$ and circle $\mathcal{C}_3(\ell_2, z_1)$ to handle the two scenarios together, which will be discussed later.*

As stated before, R_1 is the distance of the closest BS from location 1 (closest point of PPP Φ from ℓ_1). Hence the distribution of R_1 is given by the null probability of the PPP Φ and is thus given as $f_{R_1}(r_1) = 2\lambda\pi r_1 e^{-\lambda\pi r_1^2}$ [5]. The distribution of R_2, the distance of the serving BS at location 2 depends whether a handoff occurs or not when user moves from ℓ_1 to ℓ_2. If there is no handoff, the serving BS at location 2 is same as the one at location 1. In that scenario, $r_2 = r_{12} = \sqrt{r_1^2 + v^2 + 2r_1 v\cos\theta}$ as evident from Fig. 2.1(b) and its distribution can be obtained accordingly from the distributions of R_1 and Θ. However when handoff occurs, the distribution of R_2 is not straightforward and is derived next in Lemma 2.

Lemma 2. *Conditioned on r_1, θ and the occurrence of a handoff (event H) as the user moves a distance v at an angle θ from ℓ_1 to ℓ_2 in a PPP of intensity λ as shown in Fig. 2.1, the CDF of distance R_2 of the serving BS at location 2 is given as*

$$F_{R_2|H}(r_2|H, r_1, \theta) = \begin{cases} \frac{1-\exp(-\lambda\,|\mathbf{C}(\ell_1, r_1, \ell_2, r_2, v)|)}{1-\exp(-\lambda\,|\mathbf{C}(\ell_1, r_1, \ell_2, r_{12}, v)|))}, & r_2 \in [z_1, r_{12}] \\ 0, & \textit{otherwise} \end{cases},$$

where $z_1 = \mathtt{max}(0, r_1 - v)$.

Proof. From Remark 2 and Fig. 2.1(a) (handoff scenario), it can be concurred that the distance R_2 between the new serving BS x_2 and the user at ℓ_2 is greater than $z_1 = \max(0, r_1 - v)$. Also, it can not be farther than r_{12} because otherwise it would not be closer than the serving BS x_1 at location 1. Hence $R_2 \in [\max(0, r_1 - v), r_{12}]$. Conditioned on the occurence of handoff, the CDF of R_2 is thus given as:

$$\begin{aligned} \mathbb{P}(R_2 \le r_2|H, r_1, \theta) &\overset{(a)}{=} \frac{1}{\mathbb{P}(H|r_1, \theta)}\big[1 - \mathbb{P}(N\big(\mathbf{C}(\ell_1, r_1, \ell_2, r_2, v)\big) = 0)\big] \\ &= \frac{1}{\mathbb{P}(H|r_1, \theta)}\big[1 - \exp(-\lambda\,|\mathbf{C}(\ell_1, r_1, \ell_2, r_2, v)|)\big] \end{aligned} \tag{2.3}$$

where (a) follows because conditioned on the presence of no BSs inside $\mathcal{C}_1$, the distribution of R_2 is dictated by the presence of no BSs in the region $\mathbf{C}(\ell_1, r_1, \ell_2, r_2, v) = |\mathcal{C}_2 \setminus \mathcal{C}_2 \cap \mathcal{C}_1|$ (same logic as Lemma 1) and the final result follows by using $P(H|r_1, \theta)$ from Lemma 1. ■

2.3.2 Channel Model

We assume all BSs transmit with unit power and consider Rayleigh fading links with mean power gain of unity. We assume the fading gains across all links to be spatially and temporally independent and the fading coefficient between two nodes x and y at any time slot k as $h_{xy}(k)$. We consider the following bounded path-loss function $g(x)$ for the large-scale fading,

$$g(x) = \frac{1}{\epsilon + \|x\|^{\alpha}}, \text{ where } \epsilon > 0. \tag{2.4}$$

We next define the signal and interference experienced by a typical user at two spatial locations in two different time slots for the channel model described above. Say the typical user located at ℓ_1 connects to its closest BS x_1 at time slot t_1. At a different time slot t_2, the user now located at ℓ_2 connects to its closest BS x_2. Let $S(t_1, \ell_1)$ and $S(t_2, \ell_2)$ be the received powers from the serving BSs x_1 and x_2 at time slots t_1 and t_2 (user locations ℓ_1 and ℓ_2) respectively. Then

$$S(t_1, \ell_1) = h_{x_1\ell_1}(t_1)g(x_1 - \ell_1) \tag{2.5}$$

$$S(t_2, \ell_2) = h_{x_2\ell_2}(t_2)g(x_2 - \ell_2). \tag{2.6}$$

The interference power at time slots t_1 and t_2 with the user at locations ℓ_1 and ℓ_2 is denoted by $I(t_1, \ell_1)$ and $I(t_2, \ell_2)$ respectively, which are given as

$$I(t_1, \ell_1) = \sum_{x \in \Phi, x \neq x_1} h_{x\ell_1}(t_1)g(x - \ell_1) \tag{2.7}$$

$$I(t_2, \ell_2) = \sum_{x \in \Phi, x \neq x_2} h_{x\ell_2}(t_2)g(x - \ell_2). \tag{2.8}$$

The Signal-to-Interference Ratio (`SIR`) at time slot t_1 (user located at ℓ_1) is denoted by $\texttt{SIR}(t_1, \ell_1)$ and is thus given as

$$\texttt{SIR}(t_1, \ell_1) = \frac{S(t_1, \ell_1)}{I(t_1, \ell_1)} = \frac{h_{x_1\ell_1}(t_1)g(x_1 - \ell_1)}{\sum_{x \in \Phi, x \neq x_1} h_{x\ell_1}(t_1)g(x - \ell_1)}. \tag{2.9}$$

Similarily, `SIR` at time slot t_2 (user located at ℓ_2) is given as

$$\texttt{SIR}(t_2, \ell_2) = \frac{S(t_2, \ell_2)}{I(t_2, \ell_2)} = \frac{h_{x_2\ell_2}(t_2)g(x_2 - \ell_2)}{\sum_{x \in \Phi, x \neq x_2} h_{x\ell_2}(t_2)g(x - \ell_2)}. \tag{2.10}$$

For this system setup, we first study the interference correlation at two spatial locations ℓ_1 and ℓ_2 (time slots t_1 and t_2) in terms of spatio-temporal correlation coefficient in Section 2.4 which is defined below:

$$\zeta_I(\ell_1, \ell_2) = \frac{\mathbb{E}[I(t_1, \ell_1)I(t_2, \ell_2)] - \mathbb{E}[I(t_2, \ell_2)]^2}{\mathbb{E}[I(t_2, \ell_2)^2] - \mathbb{E}[I(t_2, \ell_2)]^2}. \tag{2.11}$$

Then, we study correlation in link successes at the two user locations ℓ_1 and ℓ_2 in terms of joint coverage probability in Section 2.5, which is formally defined next.

Definition 2. *(Joint Coverage Probability) It is defined as the probability that the* `SIR` *at user's spatial location ℓ_1 (time slot t_1) and spatial location ℓ_2 (time slot t_2) both exceed a certain SIR target (threshold) T. In this chapter, it is denoted by $\mathrm{P_c}(\ell_1, \ell_2)$ and is mathematically defined as*

$$\mathrm{P_c}(\ell_1, \ell_2) = \mathbb{P}(\mathtt{SIR}(t_1, \ell_1) > T, \mathtt{SIR}(t_2, \ell_2) > T). \tag{2.12}$$

2.4 Spatio-temporal Interference Correlation

The joint interference statistics at two spatial locations/time slots in a wireless network capture the effect of spatial/temporal correlation and is helpful in characterizing the network performance. However, deriving such joint statistics is usually not straightforward. The authors in [42] analyzed the joint temporal statistics of interference for a single-hop communication link by deriving the joint characteristic function of interference which follows a multivariate symmetric alpha stable distribution. The expressions, though not obtained in closed form, facilitated the evaluation of different performance metrics such as local delay, average network throughput and transmission capacity in the low-outage regime. In [32], the authors study outage correlation in mobile ad hoc networks and state that the direct evaluation of the joint distribution of correlated interference at two time slots is impractical and hence provide lower and upper bounds for the same. In this work, we derive the exact joint statistics of interference and coverage observed at two user locations. In this Section, we

focus on the characterization of spatio-temporal interference correlation coefficient, whereas in Section Section 2.5 we will study joint coverage probability.

In this work, we consider two spatial locations of the user at a distance v apart in two different time slots, which w.l.o.g. are taken as time slot 1 and 2 i.e. $t_1 = 1$ and $t_2 = 2$. As shown in Fig. 2.1, the two user locations are taken to be $\ell_1 = (-v, 0)$ and $\ell_2 = (0, 0)$. We use the shorthand notation $I(1)$ and $I(2)$ to denote the interference at time slots 1 and 2 (spatial locations ℓ_1 and ℓ_2) and using (2.7) and (2.8), it is given as

$$I(1) = \sum_{x \in \Phi, x \neq x_1} h_x(1) g(x - v) \tag{2.13}$$

$$I(2) = \sum_{x \in \Phi, x \neq x_2} h_x(2) g(x), \tag{2.14}$$

where $h_x(1)$ and $h_x(2)$ denote the fading coefficients between a node $x \in \Phi$ and the user at locations ℓ_1 and ℓ_2 respectively. Again for simplicity, a shorthand notation $S(1)$ and $S(2)$ is used to denote the received power from the serving BSs x_1 and x_2 at time slots 1 and 2 respectively. From (2.5) and (2.6), $S(1) = h_{x_1}(1) g(x_1 - v)$ and $S(2) = h_{x_2}(2) g(x_2)$. Let the total received power (signal plus interference) at time slots 1 and 2 be denoted by $I_{\mathrm{t}}(1)$ and $I_{\mathrm{t}}(2)$ and is given as

$$I_{\mathrm{t}}(1) = \sum_{x \in \Phi} h_x(1) g(x - v) = I(1) + h_{x_1}(1) g(x_1 - v) \tag{2.15}$$

$$I_{\mathrm{t}}(2) = \sum_{x \in \Phi} h_x(2) g(x) = I(2) + h_{x_2}(2) g(x_2). \tag{2.16}$$

The expression of total received powers $I_{\mathrm{t}}(1)$ and $I_{\mathrm{t}}(2)$ for our setup is the same as the interference experienced by a typical user in an ad-hoc network with unit random access channel probability. The mean, second moment and first order cross moment of interference in such an ad-hoc network at time slots 1 and 2 is derived in [18] and will be useful in our

analysis. The expressions are given below.

$$\mathbb{E}[I_{\mathrm{t}}(1)] = \mathbb{E}[I_{\mathrm{t}}(2)] = \lambda \int_{\mathbb{R}^2} g(x)dx \tag{2.17}$$

$$\mathbb{E}[I_{\mathrm{t}}(1)^2] = \mathbb{E}[I_{\mathrm{t}}(2)^2] = E[h^2]\lambda \int_{\mathbb{R}^2} g^2(x)dx + \lambda^2 \left(\int_{\mathbb{R}^2} g(x)dx \right)^2 \tag{2.18}$$

$$\mathbb{E}[I_{\mathrm{t}}(1)I_{\mathrm{t}}(2)] = \lambda \int_{\mathbb{R}^2} g(x-v)g(x)dx + \lambda^2 \left(\int_{\mathbb{R}^2} g(x)dx \right)^2 \tag{2.19}$$

We perform a similar analysis to determine the spatio-temporal correlation of interference in a cellular network with closest-BS association policy. By definition (from Equation (2.11)), the spatio-temporal interference correlation coefficient for a typical user at spatial locations ℓ_1 and ℓ_2 (at time slots 1 and 2) is hence given as

$$\zeta_I(\ell_1, \ell_2) = \frac{\mathbb{E}[I(1)I(2)] - \mathbb{E}[I(2)]^2}{\mathbb{E}[I(2)^2] - \mathbb{E}[I(2)]^2}. \tag{2.20}$$

The mean, second moment and first order cross moment of interference for the typical user is derived next to evaluate the expression for spatio-temporal correlation coefficient given by (2.20). The mean of interference at time slot 2 is

$$\begin{aligned} \mathbb{E}[I(2)] &\overset{(a)}{=} \mathbb{E}[I_{\mathrm{t}}(2) - h_{x_2}(2)g(x_2)] \\ &\overset{(b)}{=} \mathbb{E}[I_{\mathrm{t}}(2)] - \mathbb{E}[g(x_2)], \end{aligned} \tag{2.21}$$

where (a) follows from the definition of $I_{\mathrm{t}}(2)$ in (2.16) and (b) results as $\mathbb{E}[h] = 1$. From (2.21),

$$\begin{aligned} \mathbb{E}[I(2)]^2 &= \mathbb{E}[I_{\mathrm{t}}(2)]^2 + \mathbb{E}[g(x_2)]^2 - 2\mathbb{E}[I_{\mathrm{t}}(2)]\mathbb{E}[g(x_2)] \\ &\overset{(c)}{=} \mathbb{E}[I_{\mathrm{t}}(2)]^2 + \mathbb{E}[g(x_2)]^2 - 2\lambda\mathbb{E}[g(x_2) \int_{\mathbb{R}^2} g(x)dx], \end{aligned} \tag{2.22}$$

where (c) follows by using (2.17) and the property of expectations i.e. $\mathbb{E}[cX] = c\mathbb{E}[X]$ where c is a constant.

The second moment of the interference can be now computed as

$$\begin{aligned}
\mathbb{E}[I(2)^2] &= \mathbb{E}[(I_{\mathrm{t}}(2) - h_{x_2}(2)g(x_2))^2] \\
&= \mathbb{E}[I_{\mathrm{t}}(2)^2] + \mathbb{E}[h^2]\mathbb{E}[g^2(x_2)] - 2\mathbb{E}[h_{x_2}(2)g(x_2)I_{\mathrm{t}}(2)] \\
&= \mathbb{E}[I_{\mathrm{t}}(2)^2] + \mathbb{E}[h^2]\mathbb{E}[g^2(x_2)] - 2\mathbb{E}[h_{x_2}(2)g(x_2)(h_{x_2}(2)g(x_2) + I(2))] \\
&\overset{(d)}{=} \mathbb{E}[I_{\mathrm{t}}(2)^2] - \mathbb{E}[h^2]\mathbb{E}[g^2(x_2)] - 2\mathbb{E}\big[h_{x_2}(2)g(x_2)\sum_{x\in\Phi, x\neq x_2} h_x(2)g(x)\big] \\
&\overset{(e)}{=} \mathbb{E}[I_{\mathrm{t}}(2)^2] - \mathbb{E}[h^2]\mathbb{E}[g^2(x_2)] - 2\lambda\mathbb{E}\big[g(x_2)\int_{\mathbb{R}^2\setminus\mathcal{C}_2}^{\infty} g(x)\,\mathrm{d}x\big], \qquad (2.23)
\end{aligned}$$

where (d) is obtained by using the expression of $I(2)$ in (2.14) and (e) follows from Campbell's law and the spatial independence of fading links. Here $\mathcal{C}_2$ denotes the circle centered at l_2 i.e. origin and radius r_2 as shown in Fig. 2.1.

Now, in order to evaluate correlation coefficient defined by (2.20), we are left to evaluate $E[I(1)I(2)]$, which we do next.

$$\begin{aligned}
&\mathbb{E}[I(1)I(2)] = \mathbb{E}[(I_{\mathrm{t}}(1) - h_{x_1}(1)g(x_1 - v))(I_{\mathrm{t}}(2) - h_{x_2}(2)g(x_2))] \\
&= \mathbb{E}[I_{\mathrm{t}}(1)I_{\mathrm{t}}(2)] + \mathbb{E}[h_{x_1}(1)h_{x_1}(2)g(x_1 - v)g(x_2)] - \mathbb{E}[h_{x_1}(1)g(x_1 - v)I_{\mathrm{t}}(2)] - \mathbb{E}[h_{x_2}(2)g(x_2)I_{\mathrm{t}}(1)] \\
&= \mathbb{E}[I_{\mathrm{t}}(1)I_{\mathrm{t}}(2)] + \mathbb{E}[g(x_1 - v)g(x_2)] - \mathbb{E}[h_{x_1}(1)g(x_1 - v)(h_{x_2}(2)g(x_2) + I(2))] \\
&\qquad - \mathbb{E}[h_{x_2}(2)g(x_2)(h_{x_1}(1)g(x_1 - v) + I(1))] \\
&= \mathbb{E}[I_{\mathrm{t}}(1)I_{\mathrm{t}}(2)] - \mathbb{E}[g(x_1 - v)g(x_2)] - \underbrace{\mathbb{E}[h_{x_1}(1)g(x_1 - v)I(2)]}_{T_1} - \underbrace{\mathbb{E}[h_{x_2}(2)g(x_2)I(1)]}_{T_2} \qquad (2.24)
\end{aligned}$$

The expressions of T_1 and T_2 are further simplified by proceeding as below.

$$\begin{aligned}
T_1 &= \mathbb{E}[h_{x_1}(1)g(x_1-v)I(2)] \\
&\overset{(f)}{=} \mathbb{E}[h_{x_1}(1)g(x_1-v)I(2),\bar{H}] + \mathbb{E}[h_{x_1}(1)g(x_1-v)I(2),H] \\
&\overset{(g)}{=} \mathbb{E}\big[h_{x_1}(1)g(x_1-v)\sum_{x\in\Phi,x\neq\{x_1,x_2\}} h_x(2)g(x),\bar{H}\big] + \\
&\qquad \mathbb{E}\big[h_{x_1}(1)g(x_1-v)\big(h_{x_1}(2)g(x_1)+\sum_{x\in\Phi,x\neq\{x_1,x_2\}} h_x(2)g(x)\big),H\big] \\
&= \mathbb{E}\big[g(x_1-v)\sum_{x\in\Phi,x\neq\{x_1,x_2\}} g(x)\big] + \mathbb{E}[g(x_1-v)g(x_1),H] \\
&\overset{(h)}{=} \lambda\mathbb{E}\left[g(x_1-v)\int_{\mathbb{R}^2\setminus(\mathcal{C}_1\cup\mathcal{C}_2)} g(x)\,\mathrm{d}x\right] + \mathbb{E}[g(x_1-v)g(x_1),H], \qquad (2.25)
\end{aligned}$$

where (f) follows by splitting the expectation into two possible scenarios (no handoff and handoff). This step is taken to consider the interference in the second time slot $I(2)$ appropriately. In case of handoff, interference from BS x_1 also needs to be considered at ℓ_2 whereas in case of no handoff, the BS x_1 continues to be the serving BS at location 2 and hence should not be considered as a part of $I(2)$. Step (g) follows from the above argument and considering the interference from x_1 only in the handoff scenario. Step (h) follows by applying Campbell's law and observing that interference excluding BSs x_1 and x_2 is equivalent to considering interference outside $\mathcal{C}_1$ and $\mathcal{C}_2$, where $\mathcal{C}_1$ and $\mathcal{C}_2$ are as shown in Fig. 2.1.

Proceeding similar to T_1, we obtain

$$T_2 = \lambda\mathbb{E}\left[g(x_2)\int_{\mathbb{R}^2\setminus(\mathcal{C}_1\cup\mathcal{C}_2)} g(x-v)\,\mathrm{d}x\right] + \mathbb{E}[g(x_2)g(x_2-v),H]. \qquad (2.26)$$

Now substituting various moment expressions given by (2.22), (2.23), and (2.24) in (2.20), we get a general expression for the spatio-temporal correlation coefficient as a function of v. While it is not straightforward to gain analytical insights from the final expression (given its complexity), we will revisit this general case in the Numerical Results section (Section 2.6). In the rest of this Section, we focus on the more tractable case of $v = 0$, which corresponds

to the static user, i.e., $l_1 = l_2$. We will mainly study the effect of BS density λ on the resulting temporal interference correlation coefficient $\zeta_I(\ell_1, \ell_1)$, whose expression is given next in Theorem 1.

Theorem 1. *The temporal interference correlation coefficient ($v = 0$) of the typical user at time slots* 1 *and* 2*, where the path loss function $g(x)$ is given by (2.4) is*

$$\zeta_I(\ell_1, \ell_1) = \frac{\lambda \int\limits_{\mathbb{R}^2} g^2(x)\,\mathrm{d}x - a(x_2) + 2\lambda\mathbb{E}\big[g(x_2) \int\limits_{B(0,r_2)} g(x)\,\mathrm{d}x\big]}{\mathbb{E}[h^2]\lambda \int\limits_{\mathbb{R}^2} g^2(x)\,\mathrm{d}x - b(x_2) + 2\lambda\mathbb{E}\big[g(x_2) \int\limits_{B(0,r_2)} g(x)\,\mathrm{d}x\big]}, \tag{2.27}$$

where $a(x_2) = \mathbb{E}[g^2(x_2)] + \mathbb{E}[g(x_2)]^2$, $b(x_2) = \mathbb{E}[h^2]\mathbb{E}[g^2(x_2)] + \mathbb{E}[g(x_2)]^2$, $r_2 = \|x_2\|$ and $f_{R_2}(r_2) = 2\lambda\pi r_2 \exp(-\lambda\pi r_2^2)$.

Proof. For a static user ($v = 0$), the serving BS in time slots 1 and 2 are the same i.e. $x_1 = x_2$ and is simply the closest BS to the user's location. As a result, there is no handoff to a different serving BS and therefore $\mathbb{P}(H|r_1, \theta) = 0$. The distance distribution of this serving BS in time slots 1 and 2 is therefore given by the null probability of PPP Φ as $f_{R_1}(r_1) = 2\lambda\pi r_1 e^{-\lambda\pi r_1^2}$ and $f_{R_2}(r_2) = 2\lambda\pi r_2 e^{-\lambda\pi r_2^2}$. From (2.25),

$$T_1 \overset{(a)}{=} \lambda\mathbb{E}\big[g(x_1) \int_{\mathbb{R}^2\setminus(\mathcal{C}_1\cup\mathcal{C}_2)} g(x)\,\mathrm{d}x\big] \overset{(b)}{=} \lambda\mathbb{E}\big[g(x_2) \int_{\mathbb{R}^2\setminus\mathcal{C}_2} g(x)\,\mathrm{d}x\big], \tag{2.28}$$

where (a) follows because the second term in (2.25) goes to zero (no handoff for static user ($v = 0$)) and (b) follows as $x_1 = x_2$ and $\mathcal{C}_1 = \mathcal{C}_2$. Using the same argument in (2.26), we obtain

$$T_2 = \lambda\mathbb{E}\big[g(x_2) \int_{\mathbb{R}^2\setminus\mathcal{C}_2} g(x)\,\mathrm{d}x\big]. \tag{2.29}$$

Therefore we obtain $T_1 = T_2$ and substituting their expression in (2.24), we get

$$\begin{aligned}\mathbb{E}[I(1)I(2)] &= \mathbb{E}[I_{\mathrm{t}}(1)I_{\mathrm{t}}(2)] - \mathbb{E}[g(x_1)g(x_2)] - 2\lambda\mathbb{E}\big[g(x_2) \int_{\mathbb{R}^2\setminus\mathcal{C}_2} g(x)\,\mathrm{d}x\big] \\ &= \mathbb{E}[I_{\mathrm{t}}(1)I_{\mathrm{t}}(2)] - \mathbb{E}[g^2(x_2)] - 2\lambda\mathbb{E}\big[g(x_2) \int_{\mathbb{R}^2\setminus\mathcal{C}_2} g(x)\,\mathrm{d}x\big]\end{aligned} \tag{2.30}$$

Substituting the expressions of mean, second moment and first order cross moment of interference from (2.21), (2.23) and (2.30) in the definition of correlation coefficient in (2.20), we obtain the final result. ■

For this static user scenario, we now provide asymptotic results on the effect of BS density λ on the temporal interference correlation coefficient $\zeta_I(\ell_1, \ell_1)$.

Corollary 1. $\lim\limits_{\lambda\to 0}\ \zeta_I(\ell_1,\ell_1) = \frac{1}{\mathbb{E}[h^2]}$ *and* $\lim\limits_{\lambda\to\infty}\ \zeta_I(\ell_1,\ell_1) = \frac{1}{\mathbb{E}[h^2]}$.

Proof. As stated before, the distance of the serving BS for the typical user (static) in both time slots is Rayleigh distributed with its distribution given as $f_{R_2}(r_2) = 2\lambda\pi r_2 \exp(-\lambda\pi r_2^2)$. Hence, the serving (closest) BS is located at a mean distance $\mathbb{E}[R_2] = \overline{R_2} = 0.5/\sqrt{\lambda}$.

For $\lambda \to \infty$, we have $\overline{R_2} \to 0$. This asserts that in a highly dense network, the serving BS is located very close to the typical user (origin) and hence the integral in the expression of $\zeta_I(\ell_1, \ell_1)$ in Theorem 1 vanishes to zero. Therefore,

$$\lim_{\lambda\to\infty} \zeta_I(\ell_1,\ell_1) = \lim_{\lambda\to\infty}\ \frac{\lambda \int\limits_{\mathbb{R}^2} g^2(x)\,\mathrm{d}x}{\mathbb{E}[h^2]\lambda \int\limits_{\mathbb{R}^2} g^2(x)\,\mathrm{d}x} = \frac{1}{\mathbb{E}[h^2]}. \tag{2.31}$$

Similarly for $\lambda \to 0$, we have $\overline{R_2} \to \infty$ i.e. the closest BS x_2 (or x_1) is located very far away from the typical user in a sparsely dense network.

$$\lim_{\lambda\to 0} \zeta_I(\ell_1,\ell_1) = \lim_{\lambda\to 0}\ \frac{\mathbb{E}[g^2(x_1)] + \mathbb{E}[g(x_1)]^2}{\mathbb{E}[h^2]\mathbb{E}[g^2(x_1)] + \mathbb{E}[g(x_1)]^2} \overset{(a)}{=} \frac{1}{\mathbb{E}[h^2]}, \tag{2.32}$$

where (a) follows because $\mathbb{E}[g(x_1)]^2 \ll \mathbb{E}[g^2(x_1)]$ due to i) Jensen's inequality and ii) the monotonically decreasing behaviour of $g(x)$. ■

The above result gives insights on the temporal interference correlation in cellular networks under closest BS-association policy for two extreme cases of BS density. The result for the asymptotic cases is the same as an ad-hoc network scenario ($\zeta_I = 1/\mathbb{E}[h^2]$). The spatio-temporal interference correlation coefficient for an ALOHA ad-hoc network is derived in [18]

and stated below for ALOHA parameter $p = 1$.

$$\zeta_I^{\mathtt{ad}}(\ell_1, \ell_2) = \frac{\mathbb{E}[I_{\mathrm{t}}(1)I_{\mathrm{t}}(2)] - \mathbb{E}[I_{\mathrm{t}}(2)]^2}{\mathbb{E}[I_{\mathrm{t}}(2)^2] - \mathbb{E}[I_{\mathrm{t}}(2)]^2} = \frac{\int_{\mathbb{R}^2} g(x)g(x-v)\,\mathrm{d}x}{\mathbb{E}[h^2]\int_{\mathbb{R}^2} g^2(x)\,\mathrm{d}x} \tag{2.33}$$

As noted earlier, we will revisit the general case of $v > 0$ as a part of numerical results in Section 2.6, where we will provide further insights by comparing the spatio-temporal interference correlation coefficient for an ad hoc network $\zeta_I^{\mathtt{ad}}(\ell_1, \ell_2)$ and cellular network $\zeta_I(\ell_1, \ell_2)$.

2.5 Joint Coverage Probability

As studied in Section 2.4, there is correlation in the interference powers among different spatial locations of the user in a cellular network. Consider a *static user* scenario where a typical user is static at a given spatial location in the network for multiple time slots and connects to its closest BS (serving BS) in each time slot. Due to the temporal correlation of interference, it is seen that if the user is in outage (1−coverage) of the serving BS in a given time slot, there is a higher probability that the user will be also be in outage in the future time slots [27]. From this arises the need for correlation-aware retransmission schemes where the BSs do not re-transmit (or remains silent) for certain time slots if an outage is encountered. A suitable metric which measures the correlation in coverage (or outage) in different time slots (or spatial locations) is the *joint coverage probability*, which is defined formally in Definition 2.

In cellular networks, as discussed in Section 3.3, there can either be a handoff to a new serving BS or no handoff as a typical user moves from one location to another. The joint coverage probability of a typical user in the two spatial locations hence depends on the two handoff scenarios. Although handoffs in cellular networks are critical in providing a user with the best serving BS at any given spatial location, excessive handoffs can also result in overheads and handoff delays [43]. This is a more pertinent issue in ultra-dense networks [44], where a large density of BSs may result in unnecessary handoffs i.e. a handoff to a closer BS even

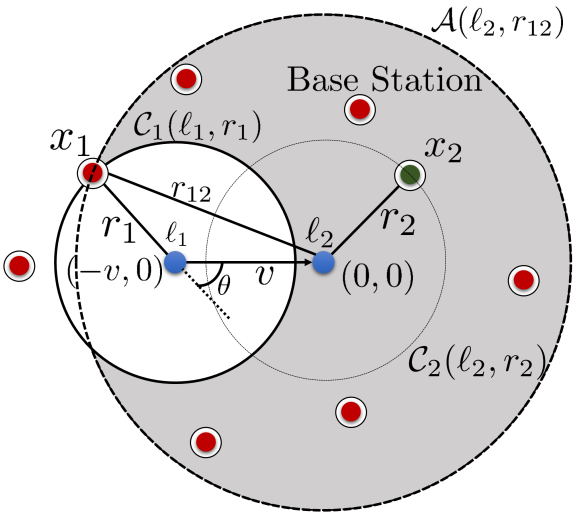

Figure 2.3: System Model when a typical user (denoted by *blue dot*) moves from ℓ_1 to ℓ_2 for different handoff strategies. i) When *handoff skipping* is used, the user skips handoffs to closer BSs and remain connected to the BS x_1 (serving BS at ℓ_1) at ℓ_2. ii) For *conventional handoffs*, the user at ℓ_2 is handed off to the closest BS x_2 (shown in green) at ℓ_2.

though continued connection to the previous serving BS meet the QoS requirements. Hence, for completeness, we study *handoff skipping* [45], where a user skips certain handoffs and remain connected to the same serving BS after moving a certain distance. In this section, we first study the joint coverage probability of a typical user with handoff skipping and then move on to a more *conventional handoff scenario*, where handoffs occur as soon as a user is closer to a new BS. In contrast to prior works which just study joint coverage probability for extreme cases of correlation ($v = 0$ and $v \to \infty$, which respectively correspond to the static and *highly* mobile user scenarios), we derive new analytical results for the joint coverage probability for the more relevant case of finite mobility, where $0 \leq v < \infty$.

2.5.1 Joint Coverage Probability With Handoff Skipping

Fig. 2.3 depicts handoff skipping scenario where user remains connected to BS x_1 after moving from ℓ_1 to ℓ_2 even in the presence of other closer BSs (i.e. it skips handoff to those BSs). In case of conventional handoffs, the user would have been handed off to the closest BS x_2 at ℓ_2. In this section, we first derive the joint coverage probability for the handoff skipping scenario in Theorem 2.

Theorem 2. *When handoff skipping is used, the joint coverage probability of a typical user at two locations ℓ_1 and ℓ_2 separated by a distance v as shown in Fig. 2.3 in a PPP of BS density λ is given as:*

$$\mathbb{P}(\mathtt{SIR}_1 > T, \mathtt{SIR}_2 > T) = \mathbb{E}_{R_1,\Theta,\Gamma}\Bigg[\exp\Bigg(-2\pi\lambda\int_{z_1}^{\infty} F_1(r_1, r, \gamma, \theta)\, r\, \mathrm{d}r\Bigg) \exp\Bigg(2\lambda\int_{|v-r_1|}^{v+r_1} \cos^{-1}\Big(\frac{r^2+v^2-r_1^2}{2rv}\Big) F_1(r_1, r, \gamma, \theta)\, r\, \mathrm{d}r\Bigg)\Bigg], \tag{2.34}$$

where $F_1(r_1, r, \gamma, \theta) = 1 - \frac{1}{(1+Tr_1^{\alpha}(r^2+v^2-2rvcos\gamma)^{-\alpha/2})(1+Tr_{12}^{\alpha}r^{-\alpha})}$, $r_{12} = \sqrt{r_1^2+v^2+2r_1v\cos\theta}$ *and* $z_1 = \mathtt{max}(0, r_1 - v)$.

Proof. Appendix A.1. ■

Having obtained the expression of joint coverage probability with handoff skipping, we now move on to study the joint coverage under conventional handoffs. More insights on the joint coverage probability will be provided through numerical results in Section 2.6.

2.5.2 Joint Coverage Probability With Conventional Handoffs

In this subsection, we derive the joint coverage probability in two spatial locations of a typical user with conventional handoffs. Considering the two scenarios possible when a user moves a distance v at an angle θ from ℓ_1 to ℓ_2 (as shown in Fig. 2.1), we derive the joint coverage

probability for both scenarios individually (no handoff and handoff) to obtain the total joint coverage probability. From total probability theorem, the joint coverage probability $\mathrm{P_c}(\ell_1, \ell_2)$ at two spatial locations ℓ_1 and ℓ_2 is given as

$$\mathbb{P}(\mathtt{SIR}_1 > T, \mathtt{SIR}_2 > T) = \mathbb{P}(\mathtt{SIR}_1 > T, \mathtt{SIR}_2 > T, \bar{H}) + \mathbb{P}(\mathtt{SIR}_1 > T, \mathtt{SIR}_2 > T, H). \quad (2.35)$$

We first derive the expression of the first term i.e. the joint coverage probability under no handoff scenario in Theorem 3 with its proof given in Appendix A.2.

Theorem 3. *Under a no handoff scenario (Fig. 2.1 (b)), the joint coverage probability of a typical user at the two spatial locations ℓ_1 and ℓ_2 in a PPP of BS density λ (i.e. same serving BS at ℓ_1 and ℓ_2) is given as*

$$\mathbb{P}(\mathtt{SIR}_1 > T, \mathtt{SIR}_2 > T, \bar{H}) = \mathbb{E}_{R_1,\Theta,\Gamma}\Bigg[\mathbb{P}(\bar{H}|r_1,\theta)\exp\big(-2\pi\lambda\int_{r_{12}}^{\infty} F_1(r_1,r,\gamma,\theta)\, r\,\mathrm{d}r\big) \exp\big(2\lambda\int_{r_{12}}^{v+r_1}\cos^{-1}\big(\frac{r^2+v^2-r_1^2}{2rv}\big)F_1(r_1,r,\gamma,\theta)\, r\,\mathrm{d}r\big)\Bigg], \quad (2.36)$$

where $F_1(r_1, r, \gamma, \theta)$ and r_{12} are defined in Theorem 2 and $\mathbb{P}(\bar{H}|r_1,\theta) = 1-\mathbb{P}(H|r_1,\theta)$ is given by (2.2).

We now derive the joint coverage probability under the handoff scenario in Theorem 4 with its proof given in Appendix A.3.

Theorem 4. *The joint coverage probability of a typical user at two spatial locations ℓ_1 and ℓ_2 in a PPP of BS density λ under a handoff scenario (Fig. 2.1(a)) i.e. different serving BS at both locations (BS x_1 at ℓ_1 and BS x_2 at ℓ_2) is given as*

$$\mathbb{P}(\mathtt{SIR}_1 > T, \mathtt{SIR}_2 > T, H) = \mathbb{E}_{R_1,R_2,\Theta,\Gamma}\Bigg[\mathbb{P}(H|r_1,\theta)\frac{1}{1+Tr_1^{\alpha}\|x_2-v\|^{-\alpha}}\frac{1}{1+Tr_2^{\alpha}\|x_1\|^{-\alpha}} \exp\big(-2\pi\lambda\int_{r_2}^{\infty}F_2(r_1,r_2,r,\gamma)\, r\,\mathrm{d}r\big)\exp\big(\lambda\mathcal{B}_1(r_1,r_2,\gamma)\big)\Bigg], \quad (2.37)$$

where $F_2(r_1,r_2,r,\gamma) = 1-\frac{1}{(1+Tr_1^{\alpha}(r^2+v^2-2rvcos\gamma)^{-\alpha/2})(1+Tr_2^{\alpha}r^{-\alpha})}$ and $\mathcal{B}_1(r_1,r_2,\gamma)$ given by (A.2).

Having obtained the joint coverage probability in two user locations ℓ_1 and ℓ_2 separated by any distance v, we now study the joint coverage for two extreme cases: i) static user ($v = 0$) and ii) highly mobile user ($v \to \infty$). The results are provided below.

Corollary 2. *(Static user) The joint coverage probability for a static user* $\mathrm{P_c}(\ell_1, \ell_1)$ *i.e. when a typical user remains at the same spatial location* ($\ell_1 = \ell_2$) *for 2 different time slots is given as:*

$$\mathrm{P_c}(\ell_1, \ell_1) = \frac{1}{{}_2F_1\left(2, -\frac{2}{\alpha}; 1 - \frac{2}{\alpha}; -T\right)}, \tag{2.38}$$

where ${}_2F_1(a, b; c; z)$ *is the Gauss hypergeometric function,* $\alpha > 2$ *is the path loss exponent and* T *is the* SIR *threshold.*

Proof. As the user is static ($v = 0$) during both time slots, there is no handoff to a different serving BS at time slot 2, i.e., the user remains connected to the same BS it was connected in time slot 1. This can also be verified from Lemma 1 that $\mathbb{P}(H|r_1, \theta) = 0$ for $v = 0$. As there is no handoff, the joint coverage probability under handoff scenario is zero i.e. $\mathbb{P}(\mathtt{SIR}_1 > T, \mathtt{SIR}_2 > T, H) = 0$ (from Theorem 3). Therefore from (2.35), the joint coverage probability for a static user, $\mathrm{P_c}(\ell_1, \ell_1)$

$$= \mathbb{P}(\mathtt{SIR}_1 > T, \mathtt{SIR}_2 > T, \bar{H}) \tag{2.39}$$

$$\overset{(a)}{=} \mathbb{E}_{R_1,\Theta,\Gamma}\left[\exp\big(-2\pi\lambda \int_{r_1}^{\infty} F(r_1, r, \gamma, \theta)\, r\, \mathrm{d}r\big) \exp\big(2\lambda \int_{r_1}^{r_1} \cos^{-1}\big(\frac{r^2 + v^2 - r_1^2}{2rv}\big) F(r_1, r, \gamma, \theta)\, r\, \mathrm{d}r\big)\right]$$

$$\overset{(b)}{=} \mathbb{E}_{R_1}\left[\exp\big(-2\pi\lambda \int_{r_1}^{\infty} \left(1 - \frac{1}{(1 + Tr_1^{\alpha} r^{-\alpha})^2}\right) r\, \mathrm{d}r\right]$$

where (a) follows from (2.36) and using $r_{12} = r_1$ and $P(\bar{H}|r_1, \theta) = 1$, and (b) results by using the definition of $F(r_1, r, \gamma, \theta)$ from Theorem 2 and substituting $v = 0$. The final result follows by deconditioning w.r.t. r_1 using $f_{R_1}(r_1) = 2\lambda\pi r_1 e^{-\lambda\pi r_1^2}$, some algebraic manipulations and the definition of Gauss hypergeometric function. ■

Corollary 3. *(Highly mobile user) The joint coverage probability for a highly mobile user ($v \to \infty$) i.e. when user moves a large distance between ℓ_1 and ℓ_2 is given as*

$$\lim_{v\to\infty} \mathrm{P_c}(\ell_1, \ell_2) = \left(\frac{1}{1+\rho(T,\alpha)}\right)^2, \tag{2.40}$$

where $\rho(T,\alpha) = T^{2/\alpha}\int_{T^{-2/\alpha}}^{\infty} \frac{\mathrm{d}u}{1+u^{\alpha/2}}$.

Proof. For a highly mobile user ($v \to \infty$), there is always handoff as the user moves a large distance from ℓ_1 to ℓ_2 i.e. $P(H|r_1,\theta) = 1$ from (2.2). Also, $F_{R_2}(r_2) = 1 - e^{-\lambda\pi r_2^2}$ from Lemma 2 as $|\mathbf{C}(\ell_1, r_1, \ell_2, r_2, v)| = \pi r_2^2$ using Definition 1 ($v \to \infty$ correspond to disjoint circle case as per Remark 1). Using the above results in Appendix A.3, we obtain $\lim_{v\to\infty} \mathbb{P}(\mathtt{SIR}_1 > T, \mathtt{SIR}_2 > T, \bar{H})$

$$= \lim_{v\to\infty} \mathbb{E}_{R_1,R_2}\left[\exp\big(-T{r_1}^{\alpha}\sum_{x\in\Phi\setminus\{x_1\}} h_x(1)\|x-v\|^{-\alpha}\big)\exp\big(-T{r_2}^{\alpha}\sum_{x\in\Phi\setminus\{x_2\}} h_x(2)\|x\|^{-\alpha}\big)\right]$$

$$\overset{(a)}{=} \mathbb{E}_{R_1}\left[\exp\big(-T{r_1}^{\alpha}\sum_{y\in\Phi'\setminus\{x_1\}} h_y(1)\|y\|^{-\alpha}\big)\right]\mathbb{E}_{R_2}\left[\exp\big(-T{r_2}^{\alpha}\sum_{x\in\Phi\setminus\{x_2\}} h_x(2)\|x\|^{-\alpha}\big)\right]$$

$$\overset{(b)}{=} \mathbb{E}_{R_1}\left[\exp\left(-\lambda\int_{r_1}^{\infty}\left(1-\frac{1}{1+Tr_1^{\alpha}u^{-\alpha}}\right)u\,\mathrm{d}u\right)\right]\mathbb{E}_{R_2}\left[\exp\left(-\lambda\int_{r_2}^{\infty}\left(1-\frac{1}{1+Tr_2^{\alpha}v^{-\alpha}}\right)v\,\mathrm{d}v\right)\right]$$

where (a) follows from the fact that under $v \to \infty$, two different instances of the point process are observed at the two locations, which allows us to distribute the expectation across the two terms. For notational simplicity, we denote the *translated* PPP as $\Phi' = \{x - v\}$. Step (b) follows as $h_y(1) \sim \exp(1)$, $h_x(2) \sim \exp(1)$ and using the PGFL of PPP Φ and Φ'. The final result follows by deconditioning w.r.t. R_1 and R_2 after some change of variables and algebraic manipulations. ■

2.6 Results and Discussion

In this section, we validate the accuracy of the analytical results (interference correlation coefficient and joint coverage probability) by means of simulations. In all simulations, we set the SIR threshold, T as 0 dB and path loss exponent $\alpha = 4$, unless mentioned otherwise.

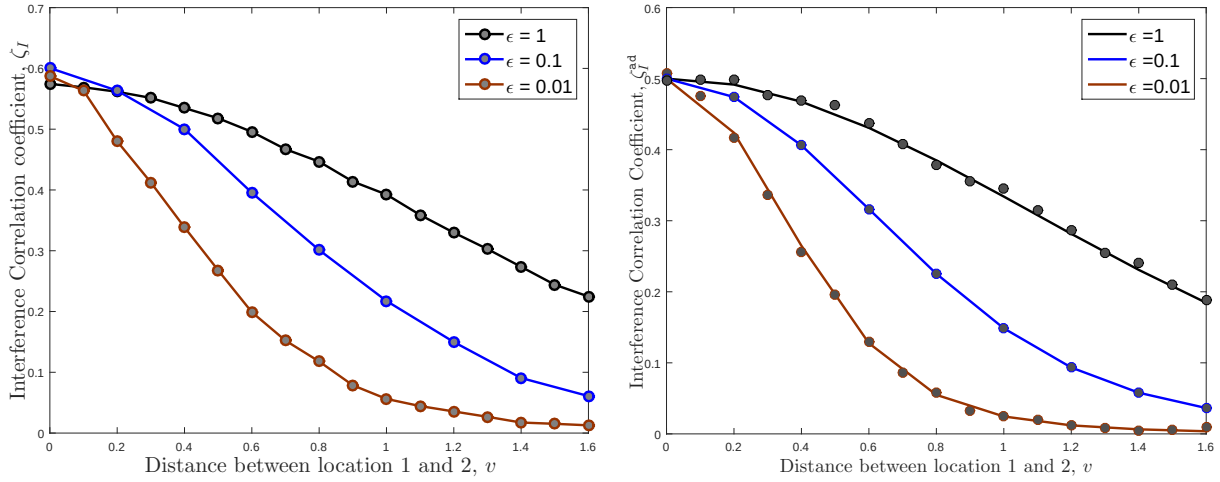

Figure 2.4: Effect of v on interference correlation coefficient in (*left*) cellular networks and (*right*) ad hoc networks. Here, ϵ is the path-loss function parameter.

2.6.1 Spatio-temporal interference correlation coefficient

Effect of distance v

Fig. 2.4 (*left*) and (*right*) plot the interference correlation coefficient between two spatial locations ℓ_1 and ℓ_2 separated by a distance v for cellular networks and ad hoc networks respectively. The interference correlation coefficient decreases with the distance between the two spatial locations. This coincides with our intuition that the set of interferers for two closeby user locations are similar resulting in a higher correlation, while independent interferers for spatial locations far apart result in lower interference correlation. We observe that the correlation coefficient attains the maximum vale for $v = 0$ (same spatial location or static user) which corresponds to the temporal correlation coefficient as derived in Theorem 1. For large v, the correlation coefficient approaches to zero signifying uncorrelated interference powers for far away spatial locations.

Effect of path loss function parameter ϵ

As evident from Fig. 2.4 (*right*), interference correlation in ad hoc networks decreases with higher path loss i.e. lower path loss function parameter ϵ. With higher path loss, the interference is dominated more by the transmitters closer to the user and therefore the correlation among interferers decreases overall. However, as seen from Fig. 2.4 (*left*), interference correlation in cellular networks does not exhibit an even trend with ϵ. This is because the interference in cellular networks depends on the choice of the serving BS (closest BS) at any given spatial location as well as the path loss function. For small v, there is a higher probability of connecting to the same BS at both locations and thereby is a major factor in deciding interference correlation at the two spatial locations. Hence there is no such trend of interference correlation coefficient with ϵ for small v. However for large v, the two locations ℓ_1 and ℓ_2 are far apart (different serving BSs), which means the interference correlation depends primarily on the path loss function and decreases with ϵ like in ad hoc networks.

Effect of BS density λ

Fig. 2.5 plots the effect of BS density λ on the temporal interference correlation coefficient ($v = 0$) in cellular networks. It can be seen from the figure that the correlation coefficient exhibits a bell-curve trend w.r.t. BS density λ i.e. interference correlation increases with BS density, attains a peak and then decreases with further increase in BS density. This behaviour is not observed for ad hoc networks, where the temporal interference correlation is independent of node density. However in cellular networks, this bell curve trend signifies a non-intuitive result that there is a certain BS density λ^* for which interference correlation is maximized and this density λ^* varies w.r.t the path loss parameter ϵ. As ϵ increases, the large-scale path loss $g(x)$ decreases, thereby requiring a lower BS density λ^* to attain a high interference correlation. For $\epsilon \to 0$ (singular path-loss function), it requires an extremely large BS density λ^* for maximum interference correlation and thereby interference correlation

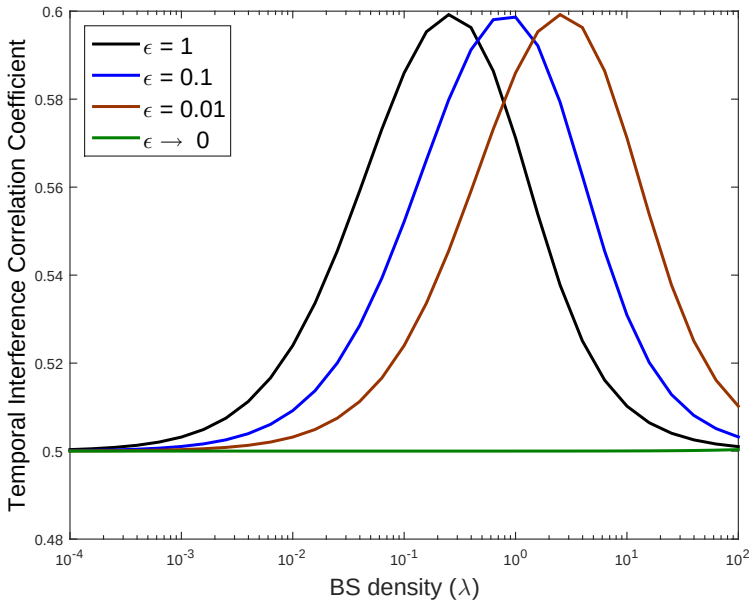

Figure 2.5: Effect of BS density λ on temporal interference correlation coefficient ($v = 0$) for a Rayleigh fading channel ($\mathbb{E}[h^2] = 2$).

does not change w.r.t. λ and remains at $1/\mathbb{E}[h^2] = 0.5$ for a Rayleigh fading channel.

2.6.2 Joint coverage probability

Fig. 2.6 plots the joint coverage probability of a typical user with conventional handoffs at two spatial locations ℓ_1 and ℓ_2 seperated by a distance v. As evident from the figure, the joint coverage probability decreases with distance v between the two spatial locations. This is explained by the decrease in interference correlation with distance v as was seen in Fig. 2.4. With higher correlation, there is a higher chance of being in coverage at the second location ℓ_2 given that the user is in coverage at the first spatial location ℓ_1. However in an uncorrelated scenario (far away spatial locations, i.e., $v \to \infty$), the coverage probability at ℓ_2 is independent of the coverage at ℓ_1, which means the joint coverage probability is simply the product of individual coverage probabilities. This trend is evident from Fig. 2.6, where the joint coverage probability at ℓ_1 and ℓ_2 decreases from a completely correlated scenario

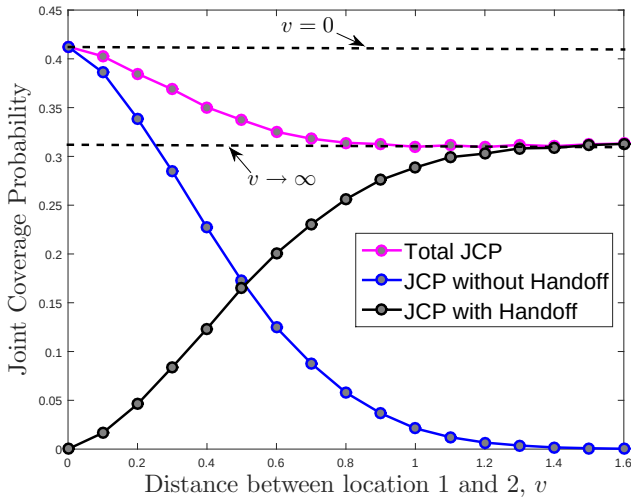

Figure 2.6: Effect of v on joint coverage probability (conventional handoff) in two spatial locations ℓ_1 and ℓ_2 ($\lambda = 1, \epsilon = 0, T = 0$ dB).

($v = 0$) and approaches an uncorrelated scenario for large v.

Fig. 2.7 compares the joint coverage probability of a typical user with handoff skipping and conventional handoffs. When handoff skipping is used, the joint coverage decreases rapidly with v compared to a conventional handoff scenario. This is because of the increase in the number of interfering BSs located closer to the user than the farther located serving BS (due to handoff skipping, user connects to the same serving BS even after displacement). Although joint coverage probability decreases rapidly when handoffs are skipped, we can avoid handoffs till a certain user displacement if the QoS is tolerable, which will naturally reduce excessive handoff delays.

2.7 Summary

In this chapter, we provided first comprehensive framework to study spatio-temporal correlation in the interference power as well as the joint coverage probability as observed at two spatial locations in a cellular network. Considering *closest BS association* policy for the

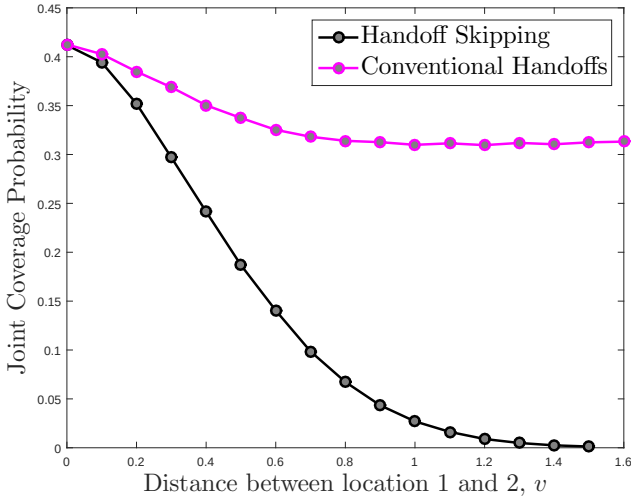

Figure 2.7: Effect of v on joint coverage probability with handoff skipping and conventional handoffs ($\lambda = 1, \epsilon = 0, T = 0$ dB).

user at both the locations, we first characterized distributions of the distances from the two locations to their respective serving BSs. Using these results, we then derived expressions for the mean, second moment and first order cross moment of interference that ultimately led to the derivation of spatio-temporal interference correlation coefficient. As expected, interference correlation was shown to decrease with increasing distance between the two locations, eventually approaching zero when the distance between the two locations approached infinity. In order to study correlation in link successes at two spatial locations at a finite distance apart, we then derived exact results for the joint coverage probability at these two user locations under *handoff skipping* and *conventional handoff* strategies. As expected, joint coverage decreased with distance (less correlation) becoming independent at very far distances. As evident from the analysis, the analytical framework required to characterize these correlations in cellular networks is significantly more complex than its ad hoc network counterpart due to the need to carefully handle cell association policies that complicates the characterization of interference field as observed from the two spatial locations.

This work has many possible extensions. First and foremost, although this work provides exact analysis of the joint coverage probability and spatio-temporal interference correlation

coefficient, the resulting expressions are not in closed form and require numerical integrations as is usually the case in most of the stochastic geometry-based analyses. While it is important to perform exact analyses for mathematical completeness as well as to complement (in some cases, circumvent) the system level simulations, it is equally important to extend such works and derive easy-to-use approximations and bounds that enable the readers to draw even better insights. From the system model side, we assumed two arbitrary spatial locations that were a distance v apart. One possible extension could be to endow this separation with a distribution and study how interference correlation varies across multiple spatial locations thereby modeling an actual mobile user.

Chapter 3

Small cell Caching: Effect of Retransmissions on Optimal caching

In Chapter 2, we developed an analytical framework for studying spatio-temporal correlation in cellular networks. In this chapter, we provide one example of exploiting spatio-temporal correlation in cache-enabled cellular networks, specifically small cell networks. Aggressive reuse of spectrum through dense deployment of *small cell base stations (SCBSs)* and caching popular content in their storage can help address growing capacity demands and reduce backhaul loads [46]. The basic idea behind small cell caching is to download popular content (mainly video) automatically in the cache of SCBSs at off-peak hours, which can then be delivered to the users during peak hours. However, the key challenge in designing such a cache-enabled network is to determine the content that should be cached at each SCBS. Among the entire content available in the internet, only a small fraction of the total content, termed *popular content*, is accessed by a large fraction of users [15], thus allowing us to focus primarily on the *library* of popular files to cache in SCBSs. However due to limited capacities of cache storage units, one can only cache a subset of the library on each SCBS. This necessitates the need to look into optimal caching of the popular content for maximum utilization of these caches and improve the overall network performance.

3.1 Related work and Motivation

Optimal caching of popular content in the storage of cache-enabled wireless networks has been studied for quite some time and can be broadly classified into two categories. The first line of work considers a scenario where each user has access (is in coverage) to at most one cache and the optimal cache placement results in storing the most popular files in the storage of all caches. To give an example, the authors in [47] studied optimal cache placement in a Device-to-Device (D2D) assisted wireless caching network and showed that the most popular contents need to be cached more often in the network to maximize the offloading probabilities in this two-tier wireless caching system. The second line of work considers a multi-coverage scenario (user is covered by multiple caches) and studies optimal cache placement while exploiting the multiple covered caches as a larger distributed cache. Using tools from stochastic geometry, [48] shows that it is not always optimal to cache the most popular content everywhere in a multi-coverage scenario and proposes a probabilistic placement policy to maximize the user's *hit probability*. Recently, the authors in [49] also studied optimal caching in heterogeneous cellular networks with different cache capabilities in each tier of the network and showed that an optimal content placement is significantly better than storing the most popular content everywhere. Despite increasing interest in the analysis of cache-enabled networks, all the prior works consider a single transmission scenario, which is not quite realistic because of the provision of retransmissions in actual systems. Therefore in this work, we consider the effect of retransmissions on the optimal caching policy for static and mobile user scenarios. More details are provided next.

3.2 Contributions and Outcomes

In this work, we determine the optimal caching policies that maximize hit probability for a typical user (static or mobile), which tries to receive its file of interest from the cache of SCBSs within a predefined number of maximum transmissions n. Here hit probability is the

probability that the user is able to successfully receive the file within n transmissions. As expected, it is seen that the hit probability increases with the number of retransmissions. We also study *reception energy* and develop optimal caching policies minimizing the same. In contrast to prior works on optimal caching, we also determine the optimal caching policies in static and mobile scenarios, and show that the optimal solutions in the two cases are significantly different. While the static case tends to cache the most popular files from the library on each SCBS, mobility de-emphasizes the importance of popularity of files and allows the SCBSs to cache content in a more *balanced* way.

3.3 System Model

We consider an ultra-dense cache-enabled SCBS network as illustrated in Fig. 3.1 and model the locations of the SCBSs by a homogeneous Poisson Point Processes (PPP) Φ with density λ [40]. We assume each SCBS can cache a maximum of L files, with $L = 1$ taken throughout this work for the simplicity of exposition. The total number of files in the library is denoted by K. We denote by $\mathtt{P}_{\mathrm{R}_i}$ the probability that the i^{th} file, $\mathcal{F}_i$, will be requested. We order the files based on their popularity, i.e., $i = 1$ and $i = K$ correspond to the most popular and least popular files, respectively. The popularity of the files in the library is assumed to follow Zipf's law [15], i.e. $\mathtt{P}_{\mathrm{R}_i} = i^{-\gamma}/\sum_{j=1}^{K} j^{-\gamma}$, where $\gamma > 0$ is the Zipf parameter which governs the skewness of the popularity distribution. We also assume that each SCBS caches file $\mathcal{F}_i$ with probability b_i independently of the other SCBSs. Therefore, $\sum_{i=1}^{K} b_i = 1$.

3.3.1 Retransmissions

We consider a typical user that attempts to download its file of interest from the cache of SCBSs in an ultra dense network as shown in Fig. 3.1. It is quite likely that a single attempt to download the file is unsuccessful, either because none of the SCBSs in the user's vicinity has cached the file (each SCBS caches only a fraction of the library) or due to

poor channel conditions. In this work, we hence study the effect of retransmissions on the network performance for both static (same user location during all retransmission attempts) and mobile user scenarios. For the analysis of mobile users, we consider the popular *infinite mobility model* [50–52] in which a user is assumed to experience an independent realization of the point process for each transmission. This model is quite relevant for ultra dense networks where even a small displacement of a user may take it to a completely new *local neighborhood* of SCBSs.

For the setup studied in this paper, we assume that a typical user (can be mobile or static) tries to receive its file of interest $\mathcal{F}_i$ from the cache of SCBSs for a maximum of n transmissions (or $n-1$ retranmissions). The signal to interference ratio (SIR) received at the typical user in the k^{th} transmission can be expressed as $\texttt{SIR}_{i,k} = \frac{h_{xk}\|x\|^{-\alpha}}{\sum\limits_{y\in\Phi\setminus\{x\}} h_{yk}\|y\|^{-\alpha}}$, where $\{h_{xk}, h_{yk}\} \sim \exp(1)$ denote the Rayleigh fading channel gains from the serving device $x \in \Phi$ and interferer $\{y\}$ in the k^{th} transmission, and $\|\cdot\|^{-\alpha}$ is standard power-law pathloss with exponent $\alpha > 2$. Here, we assume that the fading gains are independent across transmission attempts.

3.3.2 Cell Selection Policy

To choose a serving SCBS for the typical user, a straightforward choice would be to connect to the SCBS that maximizes its average received power at the receiver. This corresponds to the closest SCBS to the typical user. We term this policy as policy 1 and denote it by $\mathcal{P}1$. While this policy is meaningful in general, it suffers from a disadvantage in cache-enabled networks: the closest SCBS is not guaranteed to have the file requested by the user. Therefore this policy is termed as a *cache-agnostic policy*. To address this, we also study a *cache-aware policy* called policy 2, denoted by $\mathcal{P}2$, where the user connects to the closest SCBS that has the file it needs. This requires the knowledge of the cache of nearby SCBSs, which can be obtained by a centralized mechanism where macrocells can assist the users to connect to the SCBS which contains its file of interest. However since this SCBS may not always be in the vicinity of the user, the file transfer may not necessarily succeed due to

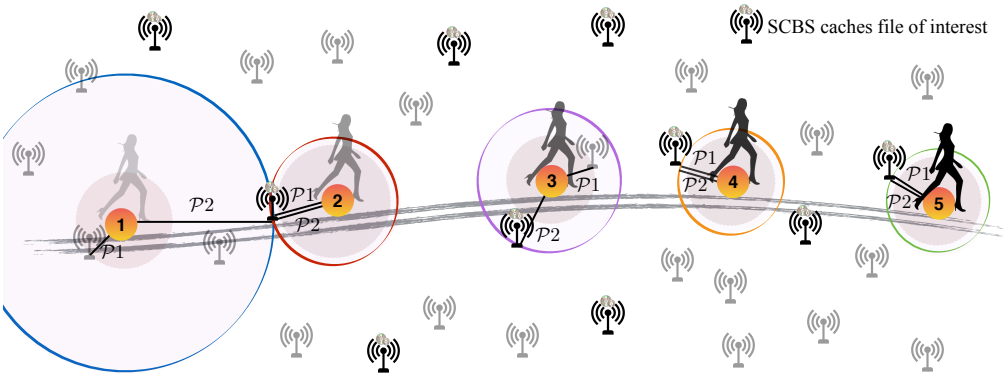

Figure 3.1: User tries to obtain the file of interest from the network using Policies 1 and 2 ($\mathcal{P}1$ and $\mathcal{P}2$) while moving from location 1 through 5. Under $\mathcal{P}1$, user connects to the SCBS providing highest average received power (closest). Under $\mathcal{P}2$, user connects to the closest SCBS that has its file of interest.

poor channel conditions. Therefore, there are tradeoffs involved in choosing the cell selection policies.

For this system model, we introduce the optimal caching problem that maximizes the *hit probability* under the two cell selection policies $\mathcal{P}1$ and $\mathcal{P}2$ in the next section.

3.4 Hit Probability

In a given transmission, a file is successfully received only when the user is in *coverage* of a SCBS which has the file of interest in its cache. For a given modulation and coding threshold T, the coverage probability of file $\mathcal{F}_i$ in the k^{th} transmission is given by $\mathbb{P}(\texttt{SIR}_{i,k} > T)$. Let S_i be the event that file $\mathcal{F}_i$ is successfully received within n transmissions with the *success probability* denoted by $p_{\mathtt{s}_i}^{1|n} = \mathbb{P}(S_i)$. We assume i.i.d. fading over the n transmissions and hence the success probability of file $\mathcal{F}_i$ in each transmission is the same and denoted by $p_{\mathtt{s}_i}$. If all the n transmissions in receiving the file are unsuccessful, the user is said to be in outage from file $\mathcal{F}_i$. Let the outage probability of file $\mathcal{F}_i$ in n transmissions be given by

$p_{\mathtt{o}_i}^{1|n} = 1 - p_{\mathtt{s}_i}^{1|n}$. The performance metric of interest is the *hit probability (HP)*, which we define mathematically as the sum of the probabilities of successfully receiving each file in the library within n transmissions, weighted by their corresponding request probabilities and is expressed below.

$$\mathtt{P}_{\mathtt{hit}} = \sum_{i=1}^{K} \mathtt{P}_{\mathrm{R}_i} p_{\mathtt{s}_i}^{1|n} = \sum_{i=1}^{K} \mathtt{P}_{\mathrm{R}_i} (1 - p_{\mathtt{o}_i}^{1|n}). \tag{3.1}$$

Our goal is maximize the hit probability which will be defined for two scenarios - i) mobile user and ii) static user in the next two subsections. The optimal caching probabilities $\{b_i\}$ that maximize the HP for the two cell selection policies $\mathcal{P}1$ and $\mathcal{P}2$ will also be discussed for both scenarios.

3.4.1 Mobile user

Under infinite mobility assumption, the probability of file success (or outage) in each transmission is independent of the previous transmissions. Hence, the outage (or success) probability of file $\mathcal{F}_i$ in n transmissions is simply given as the product of outage (or success) probabilities in each transmission i.e. $p_{\mathtt{o}_i}^{1|n} = (1 - p_{\mathrm{s}_i})^n$. The optimization problem to maximize the hit probability can be hence formulated as:

$$\begin{aligned} &\max_{\{b_i\}} \ \sum_{i=1}^{K} \mathtt{P}_{\mathrm{R}_i} (1 - (1 - p_{\mathrm{s}_i})^n), \\ &\text{s.t.} \ \sum_{i=1}^{K} b_i = 1 \text{ and } 0 \leq b_i \leq 1, \ i = 1, \dots K. \end{aligned} \tag{3.2}$$

Policy 1

As described in section 3.3, under policy 1 user connects to the closest SCBS as that maximizes its average received signal strength. A successful reception of the file hence depends on the probability that it is available in the cache of the closest SCBS and that the user is in *coverage* of the closest SCBS. Mathematically, the probability of successfully receiving

file $\mathcal{F}_i$ in a given transmission is thus the product of its caching probability b_i and coverage probability under policy 1 denoted by $p_{\mathrm{c}_i}^{(1)}$ i.e. $p_{\mathrm{s}_i} = b_i p_{\mathrm{c}_i}^{(1)}$. It is worth noting that this coverage probability $p_{\mathrm{c}_i}^{(1)}$(when user connects to the closest SCBS) is independent of the density of SCBSs under interference limited regime and has been derived in [53] for a similar cellular downlink problem. The probability of successfully receiving file $\mathcal{F}_i$ in a given transmission under policy 1 is therefore,

$$p_{\mathrm{s}_i} = b_i p_{\mathrm{c}_i}^{(1)} = \frac{b_i}{1+\rho_1(T,\alpha)}, \tag{3.3}$$

$$\text{where } \rho_1(T,\alpha) = T^{2/\alpha} \int_{T^{-2/\alpha}}^{\infty} \frac{\mathrm{d}u}{1+u^{\alpha/2}}. \tag{3.4}$$

It can be shown that the optimization problem (3.2) is concave and hence the Karush-Kuhn-Tucker (KKT) conditions provide necessary and sufficient conditions for optimality. The Lagrangian function corresponding to Problem (3.2) is

$$\begin{aligned}\mathcal{L}(\mathbf{b},\nu,\boldsymbol{\mu},\mathbf{w}) = &\sum_{i=1}^{K} \mathtt{P}_{\mathrm{R}_i}(1-(1-b_i p_{\mathrm{c}_i}^{(1)})^n) + \nu\Big(\sum_{i=1}^{K} b_i - 1\Big) \\ &- \sum_{i=1}^{K} \mu_i b_i + \sum_{i=1}^{K} w_i(b_i-1),\end{aligned}$$

where $\boldsymbol{\mu}, \mathbf{w} \in \mathbb{R}_+^K$ and $\nu \in \mathbb{R}$.

Let $\mathbf{b}^*, \nu^*, \boldsymbol{\mu}^*$ and $\mathbf{w}^*$ be primal and dual optimal. The KKT conditions for Problem (3.2) state that

$$\sum_{i=1}^{K} b_i^* = 1, \tag{3.5}$$

$$0 \leq b_i^* \leq 1 \quad , \forall i = 1, \dots K, \tag{3.6}$$

$$\mu_i^* \geq 0 \quad , \forall i = 1, \dots K, \tag{3.7}$$

$$w_i^* \geq 0 \quad , \forall i = 1, \dots K, \tag{3.8}$$

$$\mu_i^* b_i^* = 0 \quad , \forall i = 1, \dots K, \tag{3.9}$$

$$w_i^*(b_i^* - 1) = 0 \quad , \forall i = 1, \dots K, \tag{3.10}$$

$$\begin{aligned} &\mathtt{P}_{\mathrm{R}_i} n (1 - b_i^* p_{\mathrm{c}_i}^{(1)})^{n-1} p_{\mathrm{c}_i}^{(1)} \\ &+ \nu^* - \mu_i^* + w_i^* = 0 \quad , \forall i = 1, \dots K. \end{aligned} \tag{3.11}$$

The optimal cache placement under policy 1 is given next.

Theorem 5. *The optimal caching probability of file $\mathcal{F}_i$ denoted by b_i^*, that maximizes the HP (with a maximum n transmissions) for a mobile user under policy 1, is given by,*

$$b_i^* = \begin{cases} 1, & \nu^* < -\mathtt{P}_{\mathrm{R}_i} n p_{\mathrm{c}_i}^{(1)} \\ 0, & \nu^* > -\mathtt{P}_{\mathrm{R}_i} n p_{\mathrm{c}_i}^{(1)} (1 - p_{\mathrm{c}_i}^{(1)})^{n-1} \\ \frac{1}{p_{\mathrm{c}_i}^{(1)}} \left[1 - \left(\frac{-\nu^*}{\mathtt{P}_{\mathrm{R}_i} n p_{\mathrm{c}_i}^{(1)}}\right)^{\frac{1}{n-1}}\right], & \textit{otherwise} \end{cases} \tag{3.12}$$

where ν^* can be obtained as the solution of the constraint $\sum_{i=1}^{K} b_i^* = 1$.

Proof. From (3.9), (3.10) and (3.11), we have

$$w_i^* = b_i^* [-\mathtt{P}_{\mathrm{R}_i} n (1 - b_i^* p_{\mathrm{c}_i}^{(1)})^{n-1} p_{\mathrm{c}_i}^{(1)} - \nu^*], \tag{3.13}$$

which when inserted into (3.10) gives

$$b_i^*(b_i^* - 1)[-\mathtt{P}_{\mathrm{R}_i} n (1 - b_i^* p_{\mathrm{c}_i}^{(1)})^{n-1} p_{\mathrm{c}_i}^{(1)} - \nu^*] = 0. \tag{3.14}$$

From (3.14), we see that $0 < b_i^* < 1$ only if,

$$\nu^* = -\mathtt{P}_{\mathrm{R}_i} n (1 - b_i^* p_{\mathrm{c}_i}^{(1)})^{n-1} p_{\mathrm{c}_i}^{(1)}. \tag{3.15}$$

Since we know that $0 \leq b_i^* \leq 1$, this implies that

$$\nu^* \in [-\mathtt{P}_{\mathrm{R}_i} n p_{\mathrm{c}_i}^{(1)}, -\mathtt{P}_{\mathrm{R}_i} n p_{\mathrm{c}_i}^{(1)} (1 - p_{\mathrm{c}_i}^{(1)})^{n-1}]. \tag{3.16}$$

It can be seen that for $\nu^* < -\mathtt{P}_{\mathrm{R}_i} n p_{\mathrm{c}_i}^{(1)}, w_i^* > 0$ and hence from (3.10), $b_i^* = 1$. Similarly if $v^* > -\mathtt{P}_{\mathrm{R}_i} n p_{\mathrm{c}_i}^{(1)} (1 - p_{\mathrm{c}_i}^{(1)})^{n-1}$, we have $\mu_i^* > 0$ and hence from (3.9), $b_i^* = 0$. ■

In order to provide intuition, we specialize the above result to the simple case of only 2 files in the library ($K = 2$).

Corollary 4. *The optimal value (b_1^*, b_2^*) obtained by solving the optimization problem (3.2) for $K = 2$ is*

$$b_1^* = \begin{cases} 1, & n < 1 + \frac{\gamma}{\log_2\left(\frac{1}{1-p_{\mathrm{c}_i}^{(1)}}\right)} \\ \frac{a-1+p_{\mathrm{c}_i}^{(1)}}{(a+1)p_{\mathrm{c}_i}^{(1)}}, & \textit{otherwise} \end{cases}, \tag{3.17}$$

where $a = 2^{\frac{\gamma}{n-1}}$, γ is the Zipf parameter and $b_2^* = 1 - b_1^*$.

From Corollary 4, we can see that it is optimal to cache only the most popular file $\mathcal{F}_1$ till a certain number of transmissions, however an optimal cache (cache both $\mathcal{F}_1$ and $\mathcal{F}_2$) exists as the number of transmissions increase. Also, for $n = 1$ (single transmission), it is always optimal to cache the most popular file in a 2-file library scenario.

Policy 2

In this policy, the user connects to the closest SCBS which has the file of interest in its cache. The key difference in the mathematical formulation under this policy compared to $\mathcal{P}1$ is that the success probability is not weighted by caching probability of the file as the user is always connected to the SCBS that has the file of interest in its cache. Hence, the success probability of obtaining file $\mathcal{F}_i$ under policy 2 is the same as its coverage probability denoted by $p_{\mathrm{c}_i}^{(2)}$, which has been derived for a similar scenario in [54, Theorem 1] and is

stated below.

$$p_{\mathrm{s}_i} = p_{\mathrm{c}_i}^{(2)} = \frac{b_i}{b_i + \rho_1(T,\alpha) + (1-b_i)\rho_2(T,\alpha)}, \tag{3.18}$$

$$\text{where } \rho_2(T,\alpha) = T^{2/\alpha} \int_0^{T^{-2/\alpha}} \frac{\mathrm{d}u}{1+u^{\alpha/2}}, \tag{3.19}$$

and $\rho_1(T,\alpha)$ is defined in (3.4).

The solution of the optimization problem (3.2) under policy 2 can be determined on the same lines as Theorem 5 (policy 1) by using the success probability p_{s_i} defined above. The optimal solution is stated below in Theorem 6.

Theorem 6. *The optimal caching probability of file $\mathcal{F}_i$ denoted by b_i^*, that maximizes the hit probability (with a maximum n transmissions) for a mobile user under policy 2, is given by,*

$$b_i^* = \begin{cases} 1, & \nu^* < \frac{-\mathtt{P}_{\mathrm{R}_i} n}{C} \\ 0, & \nu^* > \frac{-\mathtt{P}_{\mathrm{R}_i} nC(B+C-1)^{n-1}}{(B+C)^{n+1}} \\ \phi(\nu^*), & \textit{otherwise} \end{cases}, \tag{3.20}$$

where $\phi(v^*)$ is the solution over b_i of

$$\frac{\mathtt{P}_{\mathrm{R}_i} nC((B-1)b_i + C)^{n-1}}{(Bb_i + C)^{n+1}} + \nu^* = 0, \tag{3.21}$$

$B = 1 - \rho_2(T,\alpha)$, $C = \rho_1(T,\alpha) + \rho_2(T,\alpha)$ and ν^* can be obtained as the solution of the constraint $\sum_{i=1}^{K} b_i^* = 1$.

The success probability p_{s_i} for policy 2 (given by (3.18)), being more complicated than policy 1 (given by (3.3)) makes the optimal solution in this case harder to obtain. This requires solving the polynomial equalities of the form (3.21), which may not give simple closed form solutions. Therefore, we limit our further discussion in providing the optimal solution only for the extreme cases ($n = 1$ and $n \to \infty$) for a user following policy 2.

Corollary 5. *(Single transmission - Policy 2) The optimal caching probability of file $\mathcal{F}_i$ denoted by b_i^*, that maximizes the HP for a mobile user under $\mathcal{P}2$ with a maximum of one*

attempt ($n = 1$), is given by,

$$b_i^* = \left[\frac{\sqrt{\frac{\mathtt{P}_{\mathrm{R}_i}}{\epsilon}} - (\rho_1(T,\alpha) + \rho_2(T,\alpha))}{1 - \rho_2(T,\alpha)}\right]^+, i = 1, \dots K, \tag{3.22}$$

where $[x]^+ = \mathtt{max}(0, x)$, $\sqrt{\epsilon} = \frac{\sum_{i=1}^{K^*} \sqrt{\mathtt{P}_{\mathrm{R}_i}}}{(K^*-1)\rho_1(T,\alpha)+K^*\rho_2(T,\alpha)+1}$ *and* K^*, $1 \leq K^* \leq K$, *satisfies the constraint that* $0 \leq b_i^* \leq 1$. *Here* $\rho_1(T,\alpha)$ *and* $\rho_2(T,\alpha)$ *are defined in (3.4) and (3.19) respectively.*

Proof. The result follows by substituting $n = 1$ (single attempt) in Theorem 6, solving for v^* using the constraint $\sum_{i=1}^{K} b_i^* = 1$ and simple mathematical manipulation. The detailed proof is skipped due to space constraints. ■

Corollary 6. *(Large number of transmissions) For the mobile user scenario (policy 1 and 2) with the number of transmissions approaching infinity (asymptotically), it is optimal to cache the files in the library evenly i.e.* $\lim_{n\to\infty} {b_i}^* = \frac{1}{K}$, *where K is the total number of popular files in the library.*

Proof. As evident from (3.3) and (3.18), $p_{\mathtt{s}_i}$ is a monotonically increasing function of b_i for both caching policies, say $p_{\mathtt{s}_i} = f(b_i)$. The optimization function for the mobile user defined in (3.2) is hence generalized as:

$$\max_{\{b_i\}} \sum_{i=1}^{K} \mathtt{P}_{\mathrm{R}_i}(1 - (1 - f(b_i))^n), \text{ s.t. } \sum_{i=1}^{K} b_i = 1. \tag{3.23}$$

Taking the derivative of (3.23) w.r.t. b_i for $i = 1, \dots, K-1$ individually to obtain the maxima, we get $\frac{1-f(b_i^*)}{1-f(b_K^*)} = \left(-\frac{f'(b_K^*)P_{R_K}}{f'(b_i^*)P_{R_i}}\right)^{\frac{1}{n-1}}$, $\quad \forall\, i = 1, \dots K-1$. For $n \to \infty$, we hence obtain $f(b_i^*) = f(b_K^*)$, or equivalently $b_i^* = b_K^*$, $i = 1, \dots, K-1$. With $\sum_{i=1}^{K} b_i^* = 1$, it is therefore optimal to cache evenly i.e. $b_i^* = 1/K$ for a mobile user with a large number of retransmissions. ■

3.4.2 Static user

In this scenario, the user is static at a given location in the network and tries to receive the file of interest within n transmissions. As the user is static, it remains in the same local neighbourhood of transmitters across the n transmissions resulting in temporal correlation of the interferers. As a result, the probability of successfully receiving the file in a given transmission depends on the success probability in the previous transmissions. The probability that file $\mathcal{F}_i$ is in coverage atleast once in n transmissions, denoted by $p_{\mathtt{c}_i}^{1|n}$ (or the complement of not being in coverage for all n transmissions) is given as

$$\begin{aligned} p_{\mathtt{c}_i}^{1|n} &= 1 - \mathbb{P}(\cap_{k=1}^{n} \mathtt{SIR}_{i,k} < T) \\ &\overset{(a)}{=} \sum_{k=1}^{n} \binom{n}{k} (-1)^{k+1} \mathtt{P}_{i,k}, \end{aligned} \tag{3.24}$$

where (a) follows from the inclusion-exclusion principle and $\mathtt{P}_{i,k} = \mathbb{P}(\cap_{k=1}^{n} \mathtt{SIR}_{i,k} > T)$ is defined as the joint coverage probability of file $\mathcal{F}_i$ in k transmissions.

As was the case in the mobile user scenario, we define an optimization problem for maximizing the hit probability for the two caching policies $\mathcal{P}1$ and $\mathcal{P}2$ for a static user.

Policy 1

Taking into account that the file of interest $\mathcal{F}_i$ is cached only with probability b_i in the closest SCBS, we multiply coverage probability $p_{\mathtt{c}_i}^{1|n}$ by b_i to obtain the success probability in n transmissions i.e. $p_{\mathtt{s}_i}^{1|n} = b_i p_{\mathtt{c}_i}^{1|n}$. Using (3.24) in the above result and substituting in (3.1), we obtain the following optimization problem.

$$\max_{\{b_i\}} \sum_{i=1}^{K} \mathtt{P}_{\mathrm{R}_i} \sum_{k=1}^{n} \binom{n}{k} (-1)^{k+1} b_i \mathtt{P}_{i,k}^{(1)}, \text{ s.t} \qquad \sum_{i=1}^{K} b_i = 1, \tag{3.25}$$

where $\mathtt{P}_{i,k}^{(1)}$, the joint coverage probability of file $\mathcal{F}_i$ in k transmissions under $\mathcal{P}1$ is stated below with its proof given henceforth, i.e. $\mathtt{P}_{i,k}^{(1)} =$

$$\int_0^\infty \exp\Bigg(-2\pi\lambda \int_{r_1}^\infty \Bigg(1 - \Big(\frac{u^\alpha}{Tr^\alpha + u^\alpha}\Big)^k\Bigg) u \mathrm{d}u\Bigg) f_{R_1}(r_1)\mathrm{d}r_1, \tag{3.26}$$

where R_1 denotes the distance of the closest SCBS from the typical user (or distance of the closest point of the PPP Φ of intensity λ). The distribution of R_1 is hence given from the null probability of a PPP as $f_{R_1}(r_1) = 2\lambda\pi r_1 e^{-\lambda\pi r_1^2}$ [5].

Proof. For a static user scenario, the typical user connects to the closest SCBS located at distance R_1 for all the k transmissions. From definition, the joint coverage probability of file $\mathcal{F}_i$ in k transmissions is given as

$$\begin{aligned}
\mathtt{P}_{i,k}^{(1)} &= \mathbb{E}_{R_1}[\mathbb{P}(\bigcap_{j\in\{1\ldots k\}} \mathtt{SIR}_{i,j} > T|r_1) \\
&\overset{(a)}{=} \mathbb{E}_{R_1}\Bigg[\mathbb{P}\Bigg(\bigcap_{j\in\{1\ldots k\}} \frac{h_{xj} r_1^{-\alpha}}{\sum_{y\in\Phi\setminus\{x\}} h_{yj}\|y\|^{-\alpha}} > T\Big|r_1\Bigg)\Bigg] \\
&\overset{(b)}{=} \mathbb{E}_{R_1}\Bigg[\prod_{j=1}^{k} \exp\Bigg(-Tr_1^\alpha \sum_{y\in\Phi\setminus\{x\}} h_{yj}\|y\|^{-\alpha}\Bigg)\Bigg] \\
&\overset{(c)}{=} \mathbb{E}_{R_1}\Bigg[\prod_{y\in\Phi\setminus\{x\}} \Bigg(\frac{1}{1+Tr_1^\alpha\|y\|^{-\alpha}}\Bigg)^k\Bigg] \\
&\overset{(d)}{=} \mathbb{E}_{R_1}\Bigg[\exp\Bigg(-2\pi\lambda\int_{r_1}^\infty \Bigg(1 - \Bigg(\frac{1}{1+Tr_1^\alpha u^{-\alpha}}\Bigg)^k\Bigg) u\,\mathrm{d}u\Bigg)\Bigg]
\end{aligned}$$

where (a) follows from the definition of $\mathtt{SIR}_{i,j}$ in Equation (**??**), (b) results from the fact that $h_{xj} \sim \exp(1)$ and the i.i.d. fading assumption across the k attempts resulting in simply the product of each term, (c) follows from $h_{yj} \sim \exp(1)$. Step (d) is obtained from the PGFL of the PPP Φ and converting from Cartesian to polar coordinates. The final result now follows by deconditioning w.r.t. R_1 using its distribution $f_{R_1}(r_1) = 2\lambda\pi r_1 e^{-\lambda\pi r_1^2}$. ■

Policy 2

In policy 2, as the user is always connected to the SCBS that has the file of interest in its cache, the success probability of obtaining $\mathcal{F}_i$ is the same as its coverage probability i.e $p_{\mathtt{s}_i}^{1|n} = p_{c_i}^{1|n}$. Thereby, simply using (3.24) in (3.1), we obtain the following optimization problem for policy 2.

$$\max_{\{b_i\}} \sum_{i=1}^{K} \mathtt{P}_{\mathrm{R}_i} \sum_{k=1}^{n} \binom{n}{k} (-1)^{k+1} \mathtt{P}_{i,k}^{(2)}, \text{ s.t} \qquad \sum_{i=1}^{K} b_i = 1, \tag{3.27}$$

where $\mathtt{P}_{i,k}^{(2)}$, the joint coverage probability of file $\mathcal{F}_i$ in k transmissions under policy 2 is derived by proceeding similar to policy 1 and is stated below.

$$\mathtt{P}_{i,k}^{(2)} = \int_0^\infty \exp\left(-2\pi(1-b_i)\lambda \int_0^\infty \left(1 - \left(\frac{u^\alpha}{Tr_2{}^\alpha + u^\alpha}\right)^k\right) u \mathrm{d}u\right)$$
$$\exp\left(-2\pi b_i \lambda \int_{r_2}^\infty \left(1 - \left(\frac{u^\alpha}{Tr_2{}^\alpha + u^\alpha}\right)^k\right) u \mathrm{d}u\right) f_{R_2}(r_2) \mathrm{d}r_2, \tag{3.28}$$

where R_2 denotes the distance of the closest SCBS that has the file of interest $\mathcal{F}_i$ in its cache. As file $\mathcal{F}_i$ is cached with probability b_i in the network, the distribution of R_2 is thus given by the closest point of the PPP of intensity $b_i\lambda$ and its distribution is given by $f_{R_2}(r_2) = 2b_i\lambda\pi r_2 e^{-b_i\lambda\pi r_2^2}$. The key difference in the analysis of policy 2 is that the interference field is now divided into two regions: i) Interference from those SCBSs with the file of interest $\mathcal{F}_i$ in their cache, constituting a PPP of intensity $b_i\lambda$ outside a radius r_2 (closest distance of $\mathcal{F}_i$) and ii) Interference from all other SCBSs (not having $\mathcal{F}_i$) constituting a PPP of intensity $(1-b_i)\lambda$ in $\mathbb{R}^2$.

The optimal caching probabilities of a static user under policy 1 and 2 i.e. the solutions of the optimization problem (3.25) and (3.27) can be obtained by proceeding similar to the mobile user scenario (Theorem 5 and 6). Further insights on the optimal caching probabilities for a static user under both caching policies are provided in section 3.6. Next, we study the optimal caching probabilities $\{b_i\}$ that minimizes the *total reception energy* for the two cell selection policies $\mathcal{P}1$ and $\mathcal{P}2$ in the next section.

3.5 Reception Energy

A user trying to access a file from the cache of SCBS has to expend a certain amount of energy for the reception of the file. This *reception energy* is directly proportional to the amount of time it takes for the user to receive the file of interest, termed as *reception delay*. The reception delay of file $\mathcal{F}_i$ denoted by D_{R_i} is related to its reception energy E_{R_i} by a proportionality constant c as $E_{R_i} = cD_{R_i}$ which is normalized as 1 in this chapter. See [55] for a more comprehensive power model (including backhaul, transmit power) in cache-enabled networks. In this work, we just focus on *reception energy* as that encompasses one of the major components of power consumption in cache-enabled networks.

As was in the case of hit probability, we define a metric called Total Reception Energy (TRE), which is the sum of reception energies of each file in the library weighted by their corresponding request probabilities and is expressed below.

$$\texttt{TRE} = c\sum_{i=1}^{K} \texttt{P}_{\text{R}_i} D_{R_i}.$$

Our goal is to minimize the total reception energy (TRE) of the popular files in the network. The optimization function is thus mathematically formulated as follows :

$$\min_{\{b_i\}} \; c\sum_{i=1}^{K} \texttt{P}_{\text{R}_i} D_{R_i}, \qquad \text{s.t.} \; \sum_{i=1}^{K} b_i = 1. \tag{3.29}$$

Let M be the random variable denoting the number of transmissions required for successful reception of file $\mathcal{F}_i$ and T_o be the duration of a time slot corresponding to one attempt. The (average) reception delay of file $\mathcal{F}_i$ is hence given as follows.

$$D_{R_i} = T_o\mathbb{E}[M] \overset{(a)}{=} T_o\sum_{k=1}^{n} \mathbb{P}(M \geq k) \tag{3.30}$$

where (a) is an alternative definition of the expectation of a random variable. Note that $\mathbb{P}(M \geq k)$ correspond to the probability that it requires atleast k transmissions to receive file $\mathcal{F}_i$ and is equal to the probability that file reception was unsuccessful in the previous $k-1$

transmissions. In (3.30), we take expectation only till $k = n$ as n is the maximum number of transmissions allowed to a typical user. We again look at the two scenarios of mobile user and static user and derive expressions for reception delay in each case. For brevity, we limit our focus to $\mathcal{P}2$ as similar expressions can be derived for policy 1 as well.

3.5.1 Mobile user

As discussed in section 3.4 for the mobile user scenario, the probability that the file reception is unsuccessful for k transmissions is simply given as the product of outage probabilities in each transmissions (*independent transmissions*). Thereby, the probability that it requires atleast k transmissions to receive file $\mathcal{F}_i$ is $\mathbb{P}(M \geq k) = \mathbb{P}\{\text{failure in } k-1 \text{ transmissions}\} = (1 - p_{\mathrm{s}_i})^{k-1}$ for $k = 1, 2, ..n$. Using the above result in (3.30), the reception delay of file $\mathcal{F}_i$ for a mobile user is given as:

$$\begin{aligned} D_{R_i} &= T_o \sum_{k=1}^{n} \mathbb{P}(M \geq k) = T_o \sum_{k=1}^{n} (1 - p_{\mathrm{s}_i})^{k-1} \\ &= \frac{T_o[1 - (1 - p_{\mathrm{s}_i})^n]}{p_{\mathrm{s}_i}} \end{aligned} \tag{3.31}$$

where p_{s_i}, the success probability of obtaining file $\mathcal{F}_i$ under policy 2 is given by (3.18).

3.5.2 Static user

As discussed in section 3.4 for a static user, the probability of successful reception of a file in a given transmission depends on the outcome of the previous transmissions. Proceeding similar to the hit probability analysis, we derive the average reception delay of a file under policy 2 in Theorem 7.

Theorem 7. *The average reception delay of file $\mathcal{F}_i$ for a static user under policy 2 is*

$$D_{R_i} = T_o \sum_{k=1}^{n} (-1)^{k+1} \binom{n}{k} \mathtt{P}^{(2)}_{i,k-1} \tag{3.32}$$

Proof. The probability that it takes atleast k transmissions to receive $\mathcal{F}_i$ or the probability of failure in $k-1$ transmissions for a static user scenario (temporally correlated) under policy 2 follows from (3.24) as
$\mathbb{P}(\cap_{j=1}^{k-1}\mathtt{SIR}_{i,j} < T) = 1 - \sum_{j=1}^{k-1}\binom{k-1}{j}(-1)^{j+1}\mathtt{P}_{i,j}^{(2)}$. Substituting this result in (3.30), we get

$$D_{R_i} = T_o \sum_{k=1}^{n}\left(1 - \sum_{j=1}^{k-1}\binom{k-1}{j}(-1)^{j+1}\mathtt{P}_{i,j}^{(2)}\right) \tag{3.33}$$

The final result follows by using the properties of combinations. ■

As was in the case of hit probability, the optimal caching probabilities that minimizes the total reception energy under policy 2 can be derived by proceeding similar to Theorem 6. To gain insights, we however provide the optimal solution in an asymptotic scenario (large number of attempts n).

Corollary 7. *Asymptotically, the optimal caching probabilities for minimizing the TRE for a mobile user following policy 2 is* $\lim_{n\to\infty} {b_i}^* = \frac{\sqrt{P_{R_i}}}{\sum_{j=1}^{K}\sqrt{P_{R_j}}}$.

Proof. From (3.31) with $n \to \infty$, $D_{R_i} = T_o/p_{\mathtt{s}_i}$. Using the expression of $p_{\mathtt{s}_i}$ for policy 2 from (3.18) and substituting in (3.29), we obtain the following optimization problem for TRE.

$$\min_{\{b_i\}} \; c\sum_{i=1}^{K}\mathtt{P}_{\mathrm{R}_i}T_o\frac{(b_i + \rho_1(T,\alpha) + (1-b_i)\rho_2(T,\alpha))}{b_i} \tag{3.34}$$

Taking the derivative of the objective function w.r.t b_i, for $i = 1, 2, ..., K-1$ to obtain the minima, we get $\frac{\mathtt{P}_{\mathrm{R}_i}}{\mathtt{P}_{\mathrm{R}_K}} = \frac{b_i^2}{b_K^2}$. The optimal value b_i^* follows by using the above relation and the constraint $\sum_{i=1}^{K} b_i^* = 1$. ■

3.6 Results and Discussion

For the purpose of numerical results, we consider a library of $K = 2$ files with the Zipf parameter $\gamma = 1.2$ and an SIR threshold $\beta = 0\,\mathtt{dB}$. Note that even though we have results for

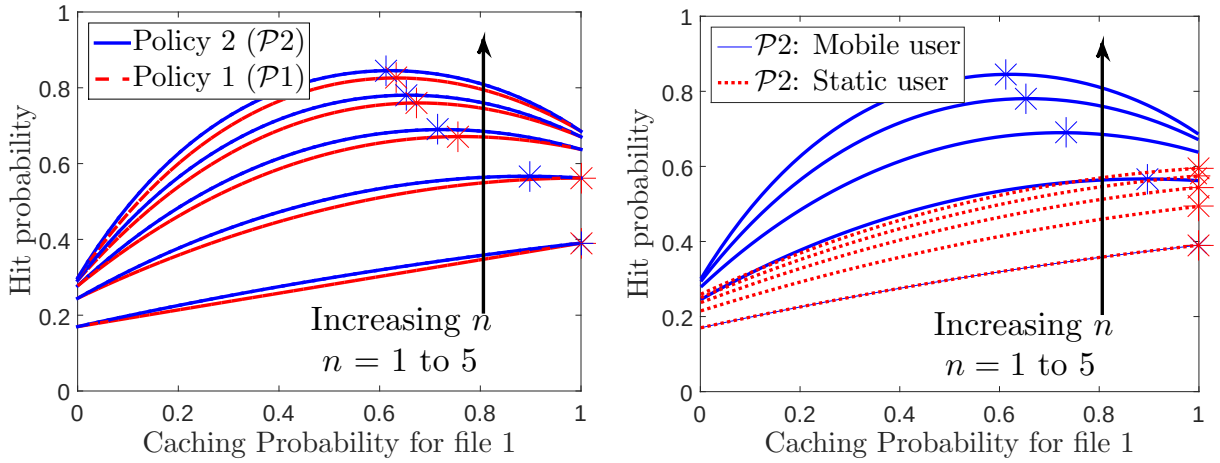

Figure 3.2: Effect of cache gathering policy (policy 1 and 2) and mobility (static and mobile user) on the hit probability for varying number of transmissions.

any generic K, we will focus on a 2-file library for the sake of exposition. It is to be noted that in all figures, the asterisk denotes the optimal caching probability.

Comparison of policies $\mathcal{P}1$ and $\mathcal{P}2$

Fig. 3.2 (*left.*) compares the two policies in terms of HP for different caching probabilities of file 1, $\mathcal{F}_1$ and varying number of transmissions n. It can be seen that the optimal caching probabilities are shifted slightly towards the left in case of policy $\mathcal{P}2$ and has a slightly higher HP compared to policy $\mathcal{P}1$. This behaviour can be attributed to the policy mechanism itself, with $\mathcal{P}2$ exhibiting a higher HP due to the available information of the cache locations. Also, the shift in optimal caching probabilities assert that it is not necessary to cache the most popular file (file $\mathcal{F}_1$) predominantly in the network and allows the SCBSs to cache the lesser requested files to a comparatively larger extent. Finally, as expected, the HP for both policies $\mathcal{P}1$ and $\mathcal{P}2$ increase with the increasing number of transmissions n.

Comparison of the mobile and static cases as a function of n for policy $\mathcal{P}2$

Fig. 3.2 (*right.*) compares the HP for the mobile and static cases when a user following policy $\mathcal{P}2$ has a maximum n transmissions to obtain the file of interest in a 2-file library case. While in the mobile case the retransmissions are made at different locations, in the static case the location of the user is not changed and hence all the retransmissions are made to the same user location. However, we assume that the fading gains are independent across time slots in which these retransmissions are made. The temporal diversity due to independent fading gains increases the HP with n, even for the static case. This trend is evident in Fig. 3.2 (*right.*), where the static case is shown by the dotted lines. As expected, the HP for the mobile user is significantly higher compared to the static user and the gap widens further with the increasing n. Similar observations can be for policy $\mathcal{P}1$ as well.

3.7 Summary

In this chapter, we have studied the effect of retransmissions on the the network performance (in terms of hit probability) for both static and mobile user scenarios in a cache-enabled SCBS network. Intuitively, if the user is allowed to access the caches multiple times (due to retransmissions), it becomes more likely that it will obtain the requested file, either due to better channel conditions or a new neighbourhood of cached SCBSs (due to user mobility). The analytical expressions developed also agree with the intuition and show an increase in hit probability with retransmissions. For a maximum n transmission attempts, we also determine the optimal cache probabilities for both static and mobile user scenarios. The optimal solutions demonstrate that while it is optimal to cache the most popular content for a static user, SCBSs can cache content in a more balanced way for mobile users.

Chapter 4

Distributed Caching in Device-to-Device Networks

Direct communication between proximate devices, termed device-to-device communications, is changing the way we envision, model and design cellular networks. While these networks have classically adopted a *content-agnostic* design approach, the presence of D2D links opens up a completely new possibility of *content-centric* operation. In particular, the devices can now cache popular content locally, which can then be delivered on demand to nearby devices. Owing to the limited storage capacities of the devices, all the popular content cannot be cached on each device, which means that a typical device may have to download content from multiple neighboring devices resulting in a *distributed* caching network. As the typical device tries to access content/file portions cached farther away from it, the number of potential dominant interferers increases, which presents a bottleneck in the content reception, possibly leading to a poor system performance. This degradation in the performance of D2D links has not been analyzed before and is the main focus of this chapter.[1]

Caching popular content on the user devices and delivering it asynchronously to other prox-

[1] A part of this chapter has been published as a conference paper in [54] and reprinted here with permissions from IEEE.

imate devices via D2D communication helps offload traffic from cellular network [56]. However, with the increasing popularity of ultra-HD videos (larger file sizes), it will become increasingly difficult to cache the entire file of interest in a single mobile device. Even though storage costs are decreasing, only a small fraction of user device's memory may be allotted for caching files. This has led to the consideration of *distributed storage* regime [57], where the file of interest is stored as multiple portions among different devices in the network. While this has mostly been explored from the information theoretic perspective, e.g., see [58], we study this problem from wireless communications perspective. In particular, for the file portions cached geographically farther from the typical user's location, there is a likelihood of having dominant interferers closer than the serving device, which present a *bottleneck* in the reception of the complete file [54]. In this chapter, we use tools from stochastic geometry to quantify the effect of user mobility on this bottleneck in terms of both coverage probability and *local delay* [32]. More details are provided next.

4.1 Related Work and Motivation

Content duplication or caching reduces peak traffic by placing relevant content on various devices across the network during off-peak hours. The information theoretic formulation of this problem was introduced recently in [59] where an order-optimal coded caching scheme was also proposed. There has also been a lot of work on exploiting caching in small cell networks to alleviate backhaul load [60] and utilize content reuse through D2D communication [61]. The idea of *distributed caching* is studied in [62], where femtocell-like base stations called *helpers* form a wireless distributed caching network and an optimal caching strategy is proposed. Due to the irregular node locations and inherent randomness in the cache availability, *stochastic geometry* [5] turns out to be a very relevant tool for the analysis of these networks. Recent works on distributed caching in D2D networks based on stochastic geometry are primarily focused on finding the optimal content placement distribution. In [63], the authors find the optimal caching distribution that maximizes the probability of content

delivery, assuming user demands are modeled as Zipf distribution. The authors in [57] evaluate the expected cost of obtaining a complete file from clients having cached parts of the file in their storage. The expected cost is shown to increase with distance between client and storage devices for both coded and uncoded allocation mechanisms.

There exist some prior works in the literature focusing on user mobility in cache-enabled networks which can be broadly classified into two main categories. The first line of work deals with a *storage allocation problem*, where content is proactively stored in the memory of small cells or devices along the user's mobility pattern to support seamless connectivity. The authors in [64] consider a distributed storage system, which stores different *chunks* of a movie requested by a moving user in different Wi-Fi APs to support live streaming of the movie. An optimal storage allocation scheme which minimizes the access probability of files from the macro base-station in a coded cache-enabled network when the user mobility is modeled as discrete-time Markov model is dealt in [65]. The second line of work deals with studying key metrics such as coverage probability or local delay of a moving user in *random* wireless network. The authors in [32] studied finite mobility in Poisson networks; derived various bounds of local delay for different mobility and transmission models and showed that user mobility helps reduce the local delay. While studying the impact of mobility of transmitters in a multicast D2D network, [66] showed that mobility increases the mean number of *covered* receivers. In this work, we build upon this observation and show that mobility increases the chance of a user being closer to a device that has cached its file of interest in a distributed storage scenario. The key difference behind our and these existing works is that we provide new analytical results on these key metrics capturing the user's local neighborhood as it moves around in a distributed caching network. Further details are provided next.

4.2 Contributions and Outcomes

We develop a new model for a distributed caching D2D network in which the content of interest for the typical user is partitioned into several portions. Modeling the device locations as a Poisson Point Process (PPP), we derive easy-to-use expressions for the coverage probability and average delay experienced by a typical user in receiving each file portion. As expected, while the file portions that are cached closer to the device can be received fairly reliably, the D2D link performance degrades significantly while receiving farther file portions due to the presence of several potential dominant interferers. This exposes a performance tradeoff: while it is desirable to partition the content of interest into several portions due to storage constraints, it is not beneficial for the D2D link performance. Two different yet equivalent viewpoints, *file-type* based and *distance* based, are considered for the analysis of coverage probability and average delay for varying cache probabilities and device activity factors.

Assuming a storage capacity of one file portion per device, we study the effect of user mobility on the coverage probability and average local delay experienced by a typical device while obtaining cached file portions. Specifically, we study a 2-file distributed caching system, where a typical user successfully receives one file portion at its initial location and receives the other file portion at its next location, which is assumed to be distance $v \geq 0$ away. Modeling the device locations as a PPP, we first derive distance distributions for receiving the second file portion (termed *farther* file portion) at the second location. The exact analysis at the second location is not straightforward and requires the knowledge of the *local neighborhood* as observed at the first location. After carefully incorporating this information in the form of asymmetric *exclusion zone* with respect to the second location, we derive tractable expressions for the coverage probability and local delay of receiving the farther cached file portion (the one that was not received at the first location) for different levels of mobility. Our results concretely demonstrate that coverage probability at the second location increases with user mobility and asymptotically approaches an *independent*

scenario, where the coverage probability of a file portion is independent of its geographical location in the network. Lower local delays also result with user mobility.

4.3 System Model

4.3.1 System Setup

Device locations are modeled as a homogeneous PPP Φ with intensity λ. We consider a 2-file distributed caching system where each device has either file A or file B cached independently with probabilities $p_A = p$ and $p_B = 1 - p$. Here file A and B correspond to two portions of a larger file requested by the *typical* device. See Fig. 4.1 for an illustration of this 2-file distributed caching setup with caching probabilities $p_A = 0.3$ and $p_B = 0.7$. As evident from the figure, file B is cached on more devices owing to its higher cache probability compared to file A. This setup can also be extended similarly to a K-file distributed caching case. For this system model, we are interested in finding the coverage and average delay experienced while receiving each file type. Independent thinning of the original PPP Φ results in two independent PPPs, Φ_A for file A and Φ_B for file B. Conditioned on the serving link, each interferer is assumed to be active independently with probability q. This activity factor captures the fact that all the devices in the network may not always be active.

4.3.2 Channel Model

For the wireless channels, we assume distance-dependent power-law pathloss with exponent α and Rayleigh fading. For a typical user located at the origin, the power received at the user from a device $x \in \Phi$ is $P = P_t h_x \|x\|^{-\alpha}$, where P_t is the transmit power, $h_x \sim \exp(1)$ models Rayleigh fading, and $\alpha > 2$ is the pathloss exponent. To define the interference power, we need an additional binary random variable t_y, which takes value 1 with probability q and 0 otherwise. For this setup, the received signal to interference ratio (`SIR`) at the typical device

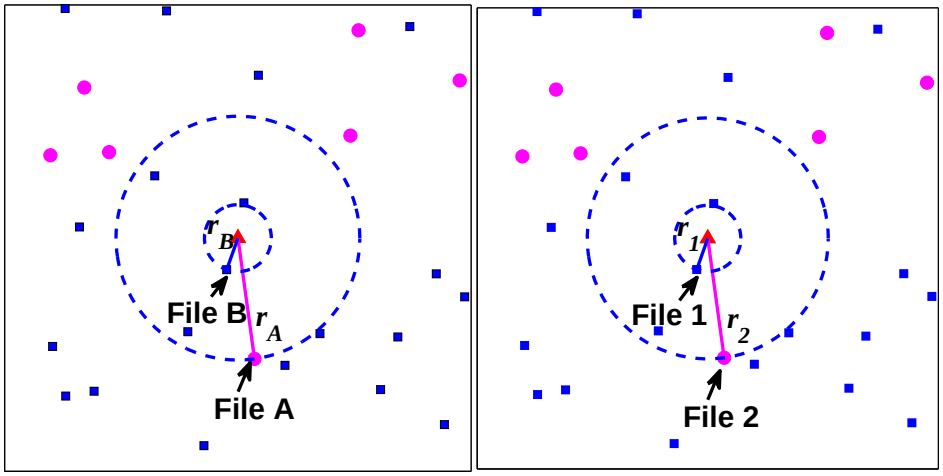

Figure 4.1: Two-file distributed caching setup ($p_A = 0.3$ and $p_B = 0.7$). *(left)* File-type based viewpoint. *(right)* Distance-based viewpoint.

can be expressed as

$$\texttt{SIR} = \frac{P_t h_x \|x\|^{-\alpha}}{\sum\limits_{y \in \Phi \setminus \{x\}} t_y P_t h_y \|y\|^{-\alpha}} = \frac{h_x \|x\|^{-\alpha}}{\sum\limits_{y \in \Phi \setminus \{x\}} t_y h_y \|y\|^{-\alpha}}. \tag{4.1}$$

We consider out-of-band D2D due to which the interference from the cellular network does not show up in the received SIR. Quite reasonably, the network is assumed to be interference limited, as a result of which the thermal noise is ignored in comparison to the interference power. This is usually the case in dense deployments where such a D2D network would be most relevant.

4.4 Coverage Probability Analysis

A device engaged in D2D communication is said to be in coverage if the received SIR at that device is greater than a coding and modulation specific threshold T. For Rayleigh fading, it is well known that the coverage probability, $p_c = P(\texttt{SIR} > T)$, can be simply expressed in terms of Laplace transform of interference I as follows [5]:

$$p_c = \mathbb{P}(h_x > TI \| \, x \, \|^{\alpha}) \stackrel{(a)}{=} \mathcal{L}_I(T \| \, x \, \|^{\alpha}), \tag{4.2}$$

where (a) follows from the fact that $h_x \sim \exp(1)$. In this section, we discuss two different viewpoints for evaluating the coverage probability of obtaining the two files of interest.

4.4.1 File-Type based Viewpoint

In a file-type based viewpoint, the coverage probability of obtaining file of each type (A or B) is determined individually by connecting to the closest device that caches the file of that type. Fig. 4.1 (*left*) depicts this file-type based viewpoint, with r_A and r_B denoting the distances of the typical user to the closest device with file A and B, respectively. These two devices will act as *serving devices* to deliver files A and B.

Coverage probability of file A

Let us consider a scenario where the typical user tries to obtain file A cached in the network and connects to the closest device located at distance r_A. Since this device is simply the closest device of the PPP Φ_A of intensity $p_A\lambda$, the distribution of R_A can be found from its null probability as $f_{R_A}(r_A) = 2\pi p_A \lambda r_A e^{-p_A\lambda\pi r_1^2}$ [5].

The coverage probability of obtaining file A can be determined by considering interference from inside and outside the circle of radius r_A separately. We denote the outside and inside interference powers as I_{out} and I_{ins}, respectively.

Lemma 3. *The Laplace transform of interference at a typical user from all active devices outside the circle of radius r_A in a PPP of intensity λ and activity factor q is given by:*

$$\mathcal{L}_{I_{\text{out}}}(T{r_A}^{\alpha}) = \exp(-\pi q\lambda {r_A}^2 \rho_1(T,\alpha)) \tag{4.3}$$

$$\textit{where} \qquad \rho_1(T,\alpha) = T^{2/\alpha} \int_{T^{-2/\alpha}}^{\infty} \frac{1}{1+u^{\alpha/2}}\,\mathrm{d}u. \tag{4.4}$$

Proof. This result follows directly from Theorem 2 in [67] for the no-noise case by observing that the interferers lying outside a circle of radius r_A with device activity factor q is equivalent

to a PPP of intensity $q\lambda$ (outside the circle). ■

The inside interference field is also handled in a similar way by observing that the all devices inside the circle of radius r_A by definition cache only file type B and hence constitute a PPP of intensity $(1-p_A)\lambda$ with an activity factor q.

Lemma 4. *The Laplace transform of interference at the typical user from devices located inside the circle of radius r_A in a PPP of intensity λ and activity factor q is given by:*

$$\mathcal{L}_{I_{\text{ins}}}(T{r_A}^{\alpha}) = \exp(-\pi(1-p_A)q\lambda {r_A}^2 \rho_2(T,\alpha)) \tag{4.5}$$

$$\textit{where} \qquad \rho_2(T,\alpha) = T^{2/\alpha} \int_0^{T^{-2/\alpha}} \frac{1}{1+u^{\alpha/2}}\,\mathrm{d}u. \tag{4.6}$$

Using the expressions of outside and inside interference from Lemma 1 and 2 respectively and exploiting the fact that I_{out} and I_{ins} are independent random variables, we multiply them to obtain the Laplace transform of total interference and hence the coverage probability of obtaining the closest file A.

Theorem 8. *The coverage probability of obtaining file A cached with probability p_A in a PPP network with intensity λ and activity factor q is*

$$\begin{aligned} p_c^{(A)} &= \int_0^{\infty} \mathcal{L}_{I_{\text{out}}}(T{r_A}^{\alpha}) \mathcal{L}_{I_{\text{ins}}}(T{r_A}^{\alpha}) f_{R_A}(r_A)\,\mathrm{d}r_A \\ &= \frac{p_A}{p_A + q[\rho_1(T,\alpha) + (1-p_A)\rho_2(T,\alpha)]} \end{aligned} \tag{4.7}$$

By symmetry, the coverage probability of obtaining file B (from the closest devices that caches it), $p_c^{(B)}$, is obtained by using its caching probability p_B in Theorem 1.

4.4.2 Distance based Viewpoint

A distance-based viewpoint of the coverage probability discussed in the current section provides a different perspective of the distributed caching problem by emphasizing on the proximity of each file type to the typical user. For this viewpoint, we label the closest file cached

to the user (can be file A or B) as *file 1* and the file type cached "farther" from the typical user as *file 2*. Fig. 4.1 (*right*) depicts this distance-based viewpoint, with r_1 and r_2 denoting the distance of the typical user to the device with file 1 and 2 respectively. This viewpoint of the cached network leads to two possible scenarios as defined below.

$\mathcal{X}$: File A is file 1 (occurs with probability p_A).

$\mathcal{Y}$: File B is file 1 (occurs with probability p_B).

The distribution of R_1, the distance of the device having file 1 from the typical user, is determined from the null probability of the PPP Φ of intensity λ as $f_{R_1}(r_1) = 2\pi\lambda r_1 e^{-\lambda\pi r_1^2}$ [5]. Conditioned on a certain file type located at r_1, the distribution of R_2 is dictated by the presence of no devices caching the other file type between r_1 and r_2 and is given as

$$f_{R_2|R_1}(r_2|r_1) = \begin{cases} 0 & r_2 < r_1 \\ 2\lambda_2\pi r_2 e^{-\lambda_2\pi(r_2^2 - r_1^2)} & r_2 > r_1 \end{cases}$$

where $\lambda_2 = p_A\lambda$ with probability p_A and $p_B\lambda$ otherwise.

Coverage probability of file 1

Assuming that the typical user connects to the device located at distance r_1 (containing file 1), the coverage probability is determined by computing the Laplace transform of interference from devices located outside a circle of radius r_1 constituting a PPP of intensity λ and activity factor q. This is equivalent to considering just the outside interference ($p_A = 1$) in Theorem 1.

Corollary 8. *The coverage probability of obtaining the closest file in a PPP (file 1) of intensity λ and activity factor q is*

$$P_c^{(1)} = \frac{1}{1 + q\rho_1(T, \alpha)} \tag{4.8}$$

Coverage probability of file 2

Having obtained the closest file portion (file 1) located at r_1, the user now tries to obtain file 2 located at a distance r_2. For the coverage probability analysis of file 2, we just focus on subcase $\mathcal{Y}$ (file A is file 2). Subcase $\mathcal{X}$ can be handled in a similar way. The total coverage probability of obtaining file 2 is derived by applying total probability theorem to the two subcases as

$$P_c^{(2)} = p_A P_c^{(X)} + p_B P_c^{(Y)}, \tag{4.9}$$

where $P_c^{(X)}$ and $P_c^{(Y)}$ denote the conditional coverage probability of obtaining file 2 in subcases $\mathcal{X}$ and $\mathcal{Y}$, respectively. For subcase $\mathcal{Y}$, file B is located at the closest device of the PPP at distance r_1 and file A is located at a device farther away at distance r_2. Conditioned on r_1 and r_2, the coverage probability of obtaining file 2 in subcase $\mathcal{Y}$ can be determined by dividing the total interference field into three regions as described below.

I_1: Singleton interference from device $x \in \Phi_B$ at distance r_1

$$I_1 = t_x h_x r_1^{-\alpha} \tag{4.10}$$

I_2: Interference from all devices with file B except the singleton $\{x\}$ at distance r_1

$$I_2 = \sum_{y \in \Phi_B \setminus \{x\}} t_y h_y \|y\|^{-\alpha} \tag{4.11}$$

I_3: Interference from all devices with file A except the serving device $y \in \Phi_A$ at distance r_2

$$I_3 = \sum_{z \in \Phi_A \setminus \{y\}} t_z h_z \|z\|^{-\alpha} \tag{4.12}$$

Lemma 5. *(Interference from singleton set) The Laplace transform of interference at a typical user from the closest device in the PPP Φ (located at distance r_1) with the serving device at r_2 is given by:*

$$\mathcal{L}_{I_1}(T{r_2}^{\alpha}) = (1-q) + \frac{q}{1 + T(\frac{r_2}{r_1})^{\alpha}} \tag{4.13}$$

Proof. By definition, the Laplace transform of interference is

$$\begin{aligned}\mathcal{L}_{I_1}(s) &= \mathbb{E}[e^{-st_x h_x r_1^{-\alpha}}] \\ &\overset{(a)}{=} (1-q) + q\mathbb{E}[e^{-sh_x r_1^{-\alpha}}] \overset{(b)}{=} (1-q) + \frac{q}{1+sr_1^{-\alpha}},\end{aligned}$$

where (a) follows from the fact that the interferer $x \in \Phi_B$ located at r_1 is active with a probability q, and (b) results from $h_x \sim \exp(1)$. The final result now follows by substituting $s = Tr_2^{\alpha}$.

Lemma 6. *The Laplace transform of interference at a typical user from devices cached with file B outside a radius r_1 with the serving device at r_2 (subcase $\mathcal{Y}$) is given by:*

$$\mathcal{L}_{I_2}(T{r_2}^{\alpha}) = \exp\left(-2\pi q p_B \lambda \int_{r_1}^{\infty} \left(\frac{T}{T+(\frac{r}{r_2})^{\alpha}}\right) r \,\mathrm{d}r\right). \tag{4.14}$$

Proof. This result also follows from Theorem 2 in [67] for the no noise case by observing that there are no interferers cached with file B lying inside a circle of radius r_1 and that the interferers lying outside r_1 with device activity factor q form an equivalent PPP of intensity $p_B q\lambda$. ■

Interference from all devices caching file A outside the circle of radius r_2 with the serving device at r_2 is handled exactly as the outside interference from Lemma 1 and hence the Laplace transform of interference can be represented as

$$\mathcal{L}_{I_3}(T{r_2}^{\alpha}) = \exp(-\pi p_A q \lambda {r_2}^2 \rho_1(T,\alpha)). \tag{4.15}$$

Using the expressions of interference from Lemmas 9 and 10 and equation (4.15), and exploiting the fact that all interference powers are independent random variables, we combine them to obtain the total interference and hence the coverage probability of obtaining file 2 in subcase $\mathcal{Y}$.

Theorem 9. *The coverage probability of obtaining file 2 at the typical device in a PPP of intensity λ and activity factor q for subcase $\mathcal{Y}$ is*

$$P_c^{(Y)} = \int_0^{\infty}\int_{r_1}^{\infty} \mathcal{L}_{I_1}(T{r_2}^{\alpha})\mathcal{L}_{I_2}(T{r_2}^{\alpha})\mathcal{L}_{I_3}(T{r_2}^{\alpha}) f_{R_2|R_1}(r_2|r_1) f_{R_1}(r_1)\,\mathrm{d}r_2\,\mathrm{d}r_1 \tag{4.16}$$

4.5 Average Delay

The average delay (also called *local delay*) is defined as the mean number of time slots required by the transmitter to successfully transmit to the receiver. Conditioned on the realization of a PPP Φ, the conditional success probability $p_{c|\Phi} = \mathbb{P}(\mathtt{SIR} > T|\Phi)$ is given by the event that the receiver is successfully connected to its assigned transmitter. If the receiver fails to decode the file, it is retransmitted in the next scheduled transmission slot. Hence, the distribution of the conditional local delay is geometric with mean $p_{c|\Phi}$ [50], taking the expectation of which over different PPP realizations yields the average delay. The average delay is hence given by

$$\mathbb{E}[D_I] = \mathbb{E}_\Phi \left[\frac{1}{\mathbb{P}(\mathtt{SIR} > T|\Phi)} \right] = \mathbb{E}_\Phi \left[\frac{1}{\mathcal{L}_I(s|\Phi)} \right] \tag{4.17}$$

4.5.1 File-Type based Analysis

The two different viewpoints considered in the analysis of coverage probability are also examined for the case of average delay. The mean number of time slots required to successfully transmit from the closest device with file A to the typical device is denoted as $\mathbb{E}[D_A]$ and derived in Lemma 7.

Lemma 7. *The average delay in communicating with the closest device with file A in a PPP*

of intensity λ and activity factor q is

$$\mathbb{E}[D_A] = \frac{p_A}{p_A - q[\rho_3(T,\alpha) + (1-p_A)\rho_4(T,\alpha)]},$$

$$\textit{where} \qquad \rho_3(T,\alpha) = T^{2/\alpha} \int_{T^{-2/\alpha}}^{\infty} \frac{1}{1-q+u^{\alpha/2}}\,\mathrm{d}u, \tag{4.18}$$

$$\textit{and} \qquad \rho_4(T,\alpha) = T^{2/\alpha} \int_{0}^{T^{-2/\alpha}} \frac{1}{1-q+u^{\alpha/2}}\,\mathrm{d}u. \tag{4.19}$$

Proof. By definition of the average delay (4.17), we have

$$\mathbb{E}[D_A] = \mathbb{E}_{\Phi}\left[\frac{1}{\mathcal{L}_{I_{\text{out}}}(Tr_A{}^{\alpha}|\Phi)\mathcal{L}_{I_{\text{ins}}}(Tr_A{}^{\alpha}|\Phi)}\right] \tag{4.20}$$

The result follows directly from Theorem 1 in [68] using suitable limits of integration for the outside and inside interference. ■

By symmetry, the average delay in communicating with the closet device with file B, $\mathbb{E}[D_B]$, can be obtained by using its corresponding caching probability p_B in Lemma 7.

4.5.2 Distance based Analysis

The distance-based analysis gives the mean number of time slots required for successfully communicating with the closest devices caching files 1 and 2, which are denoted by $\mathbb{E}[D_1]$ and $\mathbb{E}[D_2]$ respectively. $\mathbb{E}[D_1]$ is obtained by considering only the outside interference in Lemma 7 ($p_A = 1$) and is given next.

Corollary 9. *The average delay in communicating with the closest device with file 1 in a PPP of intensity λ and activity factor q is*

$$\mathbb{E}[D_1] = \frac{1}{1 - q\rho_3(T,\alpha)}.$$

Theorem 10. *The average delay in communicating with the closest device cached with file 2 for a given λ and q is*

$$E[D_2] = pE[D_2]^{(X)} + (1-p)E[D_2]^{(Y)} \tag{4.21}$$

where

$$E[D_2]^{(Y)} = \int_0^\infty \int_{r_1}^\infty \frac{1}{1-q+\frac{q}{1+sr_1^{-\alpha}}} \exp\left(2\pi p_B q\lambda \int_{r_1}^\infty \frac{r\,\mathrm{d}r}{1-q+\frac{r^\alpha}{s}}\right) \exp(\pi p_A q\lambda r_2^2 \rho_3(T,\alpha)) \\ f_{R_2|R_1}(r_2|r_1) f_{R_1}(r_1)\,\mathrm{d}r_2\,\mathrm{d}r_1.$$

Proof. See Appendix B.1. ■

As evident in Fig. 4.5, we have the following intuitive relation:

$$\mathbb{E}[D_1] + \mathbb{E}[D_2] = \mathbb{E}[D_A] + \mathbb{E}[D_B]. \tag{4.22}$$

The above equality demonstrates the equivalence of the two viewpoints undertaken in this chapter. The file-type based viewpoint provides us with simpler expressions of average delay for each file type (Lemma 7) compared to the more complex results in the distance-based viewpoint (Theorem 10). However, the distance based viewpoint gives a better handle on the worst case coverage probability and average delay i.e. when the file of interest is cached "farther" away.

4.6 Effect of User Mobility on Coverage

A user initially located at $(-v, 0)$, termed *location 1*, connects to its closest device and successfully receives the file portion cached by that device (can be file A or B). This user then moves distance v to *location 2* (taken to be the origin) and receives the other file portion from the closest device that has the other file portion in its cache. Our goal is to study the

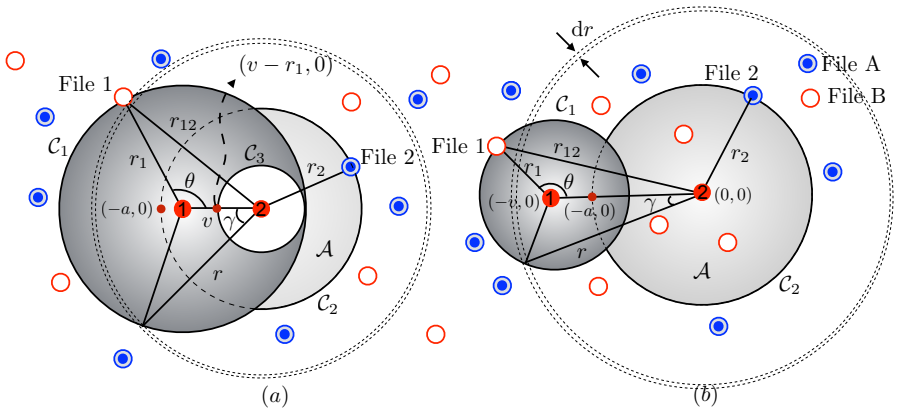

Figure 4.2: System model. (a) Scenario 1 $(v < r_1)$ and (b) Scenario 2 $(v > r_1)$. A user at location 1 $(-v, 0)$ receives file 1 and moves a distance v to location 2 $(0, 0)$, where it receives file 2. Subcase $\mathcal{Y}$ is shown (File B is File 1).

coverage probability at *location 2* as a function of v ($v = 0$ models *static* case studied in [54]). We label the closest file portion cached to the user at location 1 (can be file A or B) as *file 1* and the other file portion as *file 2*. Fig. 4.2 depicts this system setup with R_1 and R_2 denoting the distances of the user at locations 1 and 2 to the closest device with file 1 and file 2, respectively (r_1, r_2 denote realizations of R_1, R_2). Also let us define two circles: $\mathcal{C}_1$ centered at location 1 with radius r_1 and $\mathcal{C}_2$ centered at location 2 with radius r_2. When user moves a distance v from location 1 to 2, one of the following three cases arises: (i) *Case 1:* Disjoint circles ($v \geqslant r_1 + r_2$), (ii) *Case 2:* Intersecting circles ($r_2 - r_1 < v < r_1 + r_2$), and (iii) *Case 3:* Engulfed circles ($v \leqslant r_2 - r_1$). Following intermediate result will be useful for our analysis.

Definition 3. *Consider two partially-overlapping circles with radii r_1 and r_2 with centers separated by distance v, where $r_2 - r_1 < v < r_1 + r_2$, as shown in Fig. 4.2. The lightly shaded region $\mathcal{A}$ is called a lune and its area is [69, Equation (12.76)]*

$$\mathcal{A}_{\texttt{lune}} = \pi r_2^2 + \frac{1}{2}\sqrt{\left[(r_1+v)^2 - r_2^2\right]\left[r_2^2 - (r_1-v)^2\right]} - r_1^2 \cos^{-1}\left(\frac{r_1^2 + v^2 - r_2^2}{2vr_1}\right) - r_2^2 \cos^{-1}\left(\frac{r_2^2 + v^2 - r_1^2}{2vr_2}\right).$$

4.6.1 Distance distribution

As defined before, R_1 is the random variable denoting the distance of the closest file portion (file 1) from location 1. It is by definition the distance from location 1 to the closest point of a PPP with intensity λ. Its distribution can be found from the null probability of a PPP as $f_{R_1}(r_1) = 2\pi\lambda r_1 e^{-\lambda\pi r_1^2}$ [5]. For the system setup studied in this chapter, there exist two possible subcases: (i) $\mathcal{X}$: File A is file 1 (occurs with probability p_A) and (ii) $\mathcal{Y}$: File B is file 1 (occurs with probability p_B). The subcase $\mathcal{Y}$ is depicted in Fig. 4.2.

As there exists no device cached with either file portion within a distance r_1 from location 1, $\mathcal{C}_1$ can be interpreted as an *exclusion zone* in the interference field. Conditioned on a certain file portion located at a distance r_1, the distribution of R_2 is hence dictated by the presence of no devices caching the other file portion in $\mathcal{C}_2 \setminus \mathcal{C}_1$. In Fig. 4.2, the lightly shaded region $\mathcal{A}$ represents $\mathcal{C}_2 \setminus \mathcal{C}_1$, which depends on the distance v moved by the user between location 1 and 2. As per the definition of the three cases, $\mathcal{A}$ is represented by the entire circle $\mathcal{C}_2$ in case of disjoint circles (case 1), a *lune* in case of intersecting circles (case 2) or an annular region between circles $\mathcal{C}_1$ and $\mathcal{C}_2$ in case of engulfed circles (case 3). The area of $\mathcal{A}$ is mathematically expressed in Lemma 8.

Before stating Lemma 8, it is worth defining two scenarios: termed *scenarios 1* and *2*, based on the distance v moved by the user from location 1 to 2. In scenario 1, distance v moved by the user is smaller than the serving distance at location 1 i.e. $v < r_1$. As a result, the user is still inside circle $\mathcal{C}_1$ (Fig.4.2 *(a)*) and hence no device can lie within a distance $r_1 - v$ from location 2. Thus, the closest device with file 2 is located atleast a distance of $r_1 - v$ away from location 2 ($r_2 \geq r_1 - v$). In scenario 2, the user moves a larger distance ($v > r_1$), as a result of which the user moves out of the circle $\mathcal{C}_1$ (Fig.4.2 *(b)*). Hence there exists no such condition for the serving distance R_2 of file 2 in scenario 2. The above mathematical conditions for R_2 based on the two scenarios are handled appropriately by defining $z_1 = \mathtt{max}(0, r_1 - v)$ as its lower limit. Also let us define a circle $\mathcal{C}_3$ centered at location 2 and a radius of z_1, which will be used later in the analysis. It is to be noted that $\mathcal{C}_3$ converges to a point for scenario

2 ($v > r_1$) and hence not shown in the corresponding figure. The conditional distribution of R_2 is now derived next in Lemma 8.

Lemma 8. *For a given v, the conditional distribution of distance R_2 from location 2, conditioned on r_1 is*

$$f_{R_2|R_1}(r_2|r_1) = \frac{\mathrm{d}}{\mathrm{d}r_2}(1 - e^{-\lambda_2|\mathcal{A}|}), \qquad r_2 \geq \mathtt{max}(0, r_1 - v)$$

$$|\mathcal{A}| = \begin{cases} \pi r_2^2, & \mathtt{max}(0, r_1 - v) \leq r_2 \leq |v - r_1| \\ \mathcal{A}_{\mathtt{lune}}, & |v - r_1| < r_2 < v + r_1 \\ \pi(r_2^2 - r_1^2), & r_2 \geq v + r_1 \end{cases}$$

where $\lambda_2 = p_A\lambda$ with probability p_A and $p_B\lambda$ otherwise.

Proof. Conditioned on the presence of no device caching the other file portion (PPP of intensity λ_2) within a distance r_1 from location 1 (or equivalently in $\mathcal{C}_1$), the complementary CDF of R_2 is given by

$$\bar{F}_{R_2|R_1}(r_2|r_1) = \mathbb{P}\big(N(|\mathcal{C}_2|) = 0 \big| N(|\mathcal{C}_1|) = 0\big) = \mathbb{P}(N(|\mathcal{C}_2 \setminus \mathcal{C}_1|) = 0) \overset{(a)}{=} \exp(-\lambda_2|\mathcal{A}|)$$

where $|.|$ denotes the area, $N(.)$ is the number of files of other type in the specified area, and (a) results from the null probability of a PPP with intensity λ_2. The result now follows by differentiating the CDF and using appropriate values for the area of $\mathcal{A}$. ■

4.6.2 Coverage probability of file 2

A user is said to be in coverage of a certain file portion if the received $\mathtt{SIR}$ at that user from a device caching that file portion is greater than a coding and modulation specific threshold T i.e. coverage probability $P_c = \mathbb{P}(\mathtt{SIR} > T)$. For the coverage probability analysis of file 2, we just focus on subcase $\mathcal{Y}$ (file B is file 1). Result for subcase $\mathcal{X}$ will follow immediately by swapping the variables. The total coverage probability of obtaining file 2 is derived by applying total probability theorem to the two subcases as $P_{c_2} = p_A P_{c_2}^{(\mathcal{X})} + p_B P_{c_2}^{(\mathcal{Y})}$, where

$P_{c_2}^{(\mathcal{X})}$ and $P_{c_2}^{(\mathcal{Y})}$ denote the conditional coverage probability of obtaining file 2 in subcases $\mathcal{X}$ and $\mathcal{Y}$, respectively. Conditioned on R_1 and R_2, the coverage probability $P_{c_2}^{(\mathcal{Y})}$ can be determined by dividing the total interference field into three regions as described below.

I_1: Interference experienced at location 2 due to the transmission of the device $x \in \Phi_B$ (has file B) that was the *serving* device for location 1. As shown in Fig. 4.2, this device is at distance r_{12} from location 2. The interference power is

$$I_1 = t_x h_x r_{12}^{-\alpha}. \tag{4.23}$$

I_2: Interference at location 2 from all devices with file B except the singleton $\{x\}$ at distance r_{12}. This interference field is essentially Φ_B with an *asymmetric* exclusion zone $\mathcal{C}_1$ created by the exclusion of $\{x\}$. The interference power is

$$I_2 = \sum_{y \in \Phi_B \setminus \mathcal{C}_1} t_y h_y \|y\|^{-\alpha}. \tag{4.24}$$

I_3: Interference at location 2 from all devices with file A except the serving device from Φ_A at distance r_2. As there exists no device with file A in $\mathcal{C}_1$, this interference is equivalent to considering interference from Φ_A outside *exclusion zone* $\mathcal{C}_1 \cup \mathcal{C}_2$. The interference power is

$$I_3 = \sum_{z \in \Phi_A \setminus (\mathcal{C}_1 \cup \mathcal{C}_2)} t_z h_z \|z\|^{-\alpha}. \tag{4.25}$$

Due to the Rayleigh fading assumption, coverage probability in general can be expressed in terms of the Laplace transform of the interference power distribution [5, 67]. Due to the independence of the three interference terms defined above, the Laplace transform of the distribution of total interference can be expressed as the product of the Laplace transforms of the three terms. Using this, the coverage probability of obtaining file 2 in subcase $\mathcal{Y}$ can be expressed as follows.

Theorem 11. *The coverage probability of obtaining file 2 at location 2 in a PPP of intensity λ for subcase $\mathcal{Y}$ is*

$$P_{c_2}^{(\mathcal{Y})} = \int_0^\infty \int_{z_1}^\infty \int_0^\pi \mathcal{L}_{I_1|R_1,\Theta}(T{r_2}^\alpha | r_1, \theta) \mathcal{L}_{I_2|R_1}(T{r_2}^\alpha | r_1)$$
$$\mathcal{L}_{I_3|R_1,R_2}(T{r_2}^\alpha | r_1, r_2) f_{R_2|R_1}(r_2|r_1) f_{R_1}(r_1) f_\Theta(\theta) \, \mathrm{d}\theta \, \mathrm{d}r_2 \, \mathrm{d}r_1$$

where the conditional Laplace transforms of I_1, I_2 and I_3 are derived below in Lemmas 9, 10 and 11 respectively.

Proof. From the definition of coverage probability,

$$\begin{aligned}
P_{c_2}^{(\mathcal{Y})} &= \mathbb{E}_{R_1,R_2,\Theta}\big[\mathbb{P}(\mathtt{SIR} > T|r_1,r_2,\theta)\big] = \mathbb{E}_{R_1,R_2,\Theta}\big[P\big(hr_2^{-\alpha} > T(I_1+I_2+I_3)\mid r_1,r_2,\theta\big)\big] \\
&\overset{(a)}{=} \mathbb{E}_{R_1,R_2,\Theta}\big[\mathbb{E}[e^{-s(I_1+I_2+I_3)}|r_1,r_2,\theta]\big] \\
&\overset{(b)}{=} \int_0^\infty \int_{z_1}^\infty \int_0^\pi \mathbb{E}[e^{-sI_1}|r_1,\theta]\mathbb{E}[e^{-sI_2}|r_1]\mathbb{E}[e^{-sI_3}|r_1,r_2] f_{R_2|R_1}(r_2|r_1) f_{R_1}(r_1) f_\Theta(\theta)\,\mathrm{d}\theta\,\mathrm{d}r_2\,\mathrm{d}r_1
\end{aligned}$$

where (a) results from $h \sim \exp(1)$ and defining $s = Tr_2^\alpha$. Step (b) follows from the independence of the three interference powers and deconditioning w.r.t. R_1, R_2 and Θ, where Θ is a uniform random variable in $[0,\pi]$ i.e. $f_\Theta(\theta) = 1/\pi$. From (4.23), (4.24) and (4.25), it can be seen that while I_1 and I_2 depend on just r_1, I_3 is a function of both r_1 and r_2. The result now follows by using the definition of Laplace transform for the interference powers and conditioning them accordingly. ■

The Laplace transform of the distribution of interference I_1 from a *singleton* is dealt first in the following Lemma.

Lemma 9. *Given v, the conditional Laplace transform of interference I_1 from a* singleton *defined in* (4.23) *is*

$$\mathcal{L}_{I_1|R_1,\Theta}(s|r_1,\theta) = 1 - q + \frac{q}{1 + s(r_1^2 + v^2 - 2r_1 v\cos\theta)^{\frac{-\alpha}{2}}}.$$

Proof. By definition, the Laplace transform of interference is

$$\mathcal{L}_{I_1|R_1,\Theta}(s|r_1,\theta) = \mathbb{E}[e^{-st_x h_x r_{12}^{-\alpha}}] \overset{(a)}{=} 1 - q + q\mathbb{E}[e^{-sh_x r_{12}^{-\alpha}}] \overset{(b)}{=} 1 - q + \frac{q}{1 + sr_{12}^{-\alpha}},$$

where (a) follows from the fact that the interferer $x \in \Phi_B$ located at r_{12} is active with a probability q, and (b) results from $h_x \sim \exp(1)$. The final result follows by using the law of cosines in which $r_{12}^2 = r_1^2 + v^2 - 2r_1 v\cos\theta$ (see Fig. 4.2). ■

The conditional Laplace transform of interference I_2 is derived next with its proof provided in Appendix B.2. The key is in handling the asymmetric exclusion zone $\mathcal{C}_1$ carefully. Interested readers can refer to [70] for more details on how such exclusion zones can be handled.

Lemma 10. *Given v, the conditional Laplace transform of I_2 under subcase $\mathcal{Y}$ defined in* (4.24) *is*

$$\mathcal{L}_{I_2|R_1}(s|r_1) = \exp\left(-2p_B q\lambda \left(\int_{z_1}^{\infty} \frac{\pi r \,\mathrm{d}r}{1+\frac{r^\alpha}{s}} - \int_{|v-r_1|}^{v+r_1} \frac{f(r,r_1) r \,\mathrm{d}r}{1+\frac{r^\alpha}{s}} \right) \right),$$
$$\text{where } z_1 = \mathtt{max}(0, r_1 - v)\,, f(r,r_1) = \cos^{-1}\left(\frac{r^2+v^2-r_1^2}{2rv}\right).$$

The conditional Laplace transform of interference I_3 is computed similarly with its proof provided in Appendix B.3.

Lemma 11. *Given v, the conditional Laplace transform of I_3 under subcase $\mathcal{Y}$ is given by*

$$\mathcal{L}_{I_3|R_1,R_2}(s|r_1,r_2) = \exp\left(-p_A q\lambda \int_{r_2}^{\infty} \frac{2\pi r \,\mathrm{d}r}{1+\frac{r^\alpha}{s}} \right) \exp\left(p_A q\lambda \mathcal{B}(r_1,r_2,v) \right), \tag{4.26}$$
$$\text{where } \mathcal{B}(r_1,r_2,v) \text{ is given by } (B.2)$$

It is to be noted that the coverage probability analysis for a moving user conducted in this work is a generalization of the work in [54] for a static user ($v = 0$). Using the above results, we can also study the asymptotic coverage probability of file 2 when locations 1 and 2 are far apart ($v \to \infty$). In this case, $I_1 \to 0$, and the asymmetric exclusion zone $\mathcal{C}_1$ does not appear in I_2 and I_3. The final result is given below.

Corollary 10. *For large user mobility ($v \to \infty$), the asymptotic coverage probability of file 2 is given by:*

$$P_{c_2} \to p_A \mathtt{P_c}(p_B) + p_B \mathtt{P_c}(p_A), \tag{4.27}$$

$$\textit{where } \mathtt{P_c}(p) = \frac{p}{p + q[\rho_1(T,\alpha) + (1-p)\rho_2(T,\alpha)]}, \tag{4.28}$$

$$\rho_1(T,\alpha) = T^{\frac{2}{\alpha}} \int_{T^{-\frac{2}{\alpha}}}^{\infty} \frac{\mathrm{d}u}{1+u^{\frac{\alpha}{2}}},\ \rho_2(T,\alpha) = T^{\frac{2}{\alpha}} \int_{0}^{T^{-\frac{2}{\alpha}}} \frac{\mathrm{d}u}{1+u^{\frac{\alpha}{2}}}. \tag{4.29}$$

Proof. For subcase $\mathcal{Y}$ and $v \to \infty$, $\mathcal{L}_{I_1|R_1,\Theta}(Tr_2^{\alpha}|r_1,\theta) = 1$, $\mathcal{L}_{I_2|R_1}(Tr_2^{\alpha}|r_1) = \exp\big(-2\pi p_B q\lambda \int_0^{\infty} \frac{r\,\mathrm{d}r}{1+\frac{r^{\alpha}}{Tr_2^{\alpha}}}\big)$ and $\mathcal{L}_{I_3|R_1,R_2}(Tr_2^{\alpha}|r_1,r_2) = \exp\big(-2\pi p_A q\lambda \int_{r_2}^{\infty} \frac{r\,\mathrm{d}r}{1+\frac{r^{\alpha}}{Tr_2^{\alpha}}}\big)$. Also $v \to \infty$ corresponds to case 1 (disjoint circles) defined in our setup and thus $f_{R_2|R_1}(r_2|r_1) = 2p_A\lambda\pi r_2 \exp(-p_A\lambda\pi r_2^2)$ (from Lemma 1). Plugging the above values in Theorem 11 and defining $\rho_1(T,\alpha)$ and $\rho_2(T,\alpha)$ as in (4.29), we obtain $P_{c_2}^{(\mathcal{Y})} = \mathtt{P_c}(p_A)$. Similarly, it can be shown that $P_{c_2}^{(\mathcal{X})} = \mathtt{P_c}(p_B)$ for $v \to \infty$. The final result follows by applying total probability theorem to the two subcases $\mathcal{X}$ and $\mathcal{Y}$. ■

4.6.3 "Local delay" at location 2

The average delay (also called *local delay*) [32] is defined as the mean number of time slots required by the transmitter to successfully transmit a packet to the receiver. Proceeding similar to the coverage probability analysis, the local delay in communicating with the closest device cached with file 2 for the typical user at location 2 is derived by applying total probability theorem to the two subcases $\mathcal{X}$ and $\mathcal{Y}$ as $D_2 = pD_2^{(\mathcal{X})} + (1-p)D_2^{(\mathcal{Y})}$. Here $D_2^{(\mathcal{X})}$ and $D_2^{(\mathcal{Y})}$ denote the average delay of obtaining file 2 in subcases $\mathcal{X}$ and $\mathcal{Y}$, respectively. Again, we focus only on subcase $\mathcal{Y}$ as subcase $\mathcal{X}$ can be handled by swapping the variables.

Conditioned on R_1, R_2, Θ and Φ, denote $\mathcal{S}$ as the *success* event that the user at location 2 is successfully connected to the closest device cached with file 2 under subcase $\mathcal{Y}$. For an $\mathtt{SIR}$ threshold T, the conditional success probability is hence denoted as $P(\mathcal{S}) = \mathbb{P}(\mathtt{SIR} > T|r_1,r_2,\theta,\Phi)$. If the user fails to decode the file, it is retransmitted in the next scheduled transmission slot. Accordingly, the number of time slots needed until success for a given R_1, R_2, Θ and Φ is a geometrically distributed random variable with mean $P(\mathcal{S})^{-1}$. Taking the

expectation with respect to R_1, R_2, Θ and Φ yields the local delay of file 2 under subcase $\mathcal{Y}$:

$$
\begin{aligned}
D_2^{(\mathcal{Y})} &= \mathbb{E}_{R_1,R_2,\Theta,\Phi}\left[\frac{1}{\mathbb{P}(\mathtt{SIR} > T|r_1,r_2,\theta,\Phi)}\right] \\
&= \mathbb{E}_{R_1,R_2,\Theta,\Phi}\left[\frac{1}{P\left(h > Tr_2^{\alpha}(I_1+I_2+I_3) \mid r_1,r_2,\theta,\Phi\right)}\right] \\
&\overset{(a)}{=} \mathbb{E}_{R_1,R_2,\Theta,\Phi}\left[\frac{1}{\mathbb{E}[e^{-sI_1}|r_1,\theta,\Phi]\mathbb{E}[e^{-sI_2}|r_1,\Phi]\mathbb{E}[e^{-sI_3}|r_1,r_2,\Phi]}\right] \\
&\overset{(b)}{=} \int_0^{\infty}\int_{z_1}^{\infty}\int_0^{\pi} \underbrace{\mathbb{E}_{\Phi}\left[\frac{1}{\mathcal{L}_{I_1|R_1,\Theta}(s|r_1,\theta,\Phi)}\right]}_{D_{I_1}} \underbrace{\mathbb{E}_{\Phi}\left[\frac{1}{\mathcal{L}_{I_2|R_1}(s|r_1,\Phi)}\right]}_{D_{I_2}} \\
&\quad \underbrace{\mathbb{E}_{\Phi}\left[\frac{1}{\mathcal{L}_{I_3|R_1,R_2}(s|r_1,r_2,\Phi)}\right]}_{D_{I_3}} f_{R_2|R_1}(r_2|r_1) f_{R_1}(r_1) f_{\Theta}(\theta)\,\mathrm{d}\theta\,\mathrm{d}r_2\,\mathrm{d}r_1
\end{aligned}
$$

where (a) results from $h \sim \exp(1)$ and defining $s = Tr_2^{\alpha}$. (b) follows from the independence of the three interference powers (I_1, I_2 and I_3) and deconditioning w.r.t. R_1, R_2 and Θ. The final result for the local delay of obtaining file 2 in subcase $\mathcal{Y}$ is now presented in the next Theorem with the details of the rest of the steps provided in Appendix B.4.

Theorem 12. *The local delay of obtaining file 2 at location 2 in a PPP of intensity λ for subcase $\mathcal{Y}$ is*

$$
D_2^{(\mathcal{Y})} = \int_0^{\infty}\int_{z_1}^{\infty}\int_0^{\pi} D_{I_1} D_{I_2} D_{I_3} f_{R_2|R_1}(r_2|r_1) f_{R_1}(r_1) f_{\Theta}(\theta)\,\mathrm{d}\theta\,\mathrm{d}r_2\,\mathrm{d}r_1
$$

$$
\begin{aligned}
&\textit{where } D_{I_1} = \left(1 - q + \frac{q}{1 + s(r_1^2 + v^2 - 2r_1 v\cos\theta)^{-\alpha/2}}\right)^{-1}, \\
&D_{I_2} = \exp\left(2p_B q\lambda\left(\int_{z_1}^{\infty} \frac{\pi r\,\mathrm{d}r}{1 - q + \frac{r^{\alpha}}{s}} - \int_{|v-r_1|}^{v+r_1} \frac{f(r,r_1) r\,\mathrm{d}r}{1 - q + \frac{r^{\alpha}}{s}}\right)\right) \textit{ and} \\
&D_{I_3} = \exp\left(p_A q\lambda \int_{r_2}^{\infty} \frac{2\pi r\,\mathrm{d}r}{1 - q + \frac{r^{\alpha}}{s}}\right) \exp\Big(-p_A q\lambda \mathcal{B}_1(r_1,r_2,v)\Big).
\end{aligned}
$$

where $f(r,r_1)$ is defined in Lemma 10 and $\mathcal{B}_1(r_1,r_2,v)$ given by (B.3).

4.7 Results and Discussion

4.7.1 Effect of caching probability

The coverage probability of a file increases with its caching probability in the network. This is evident from Fig. 4.3, where the coverage probability of obtaining the files of interest are plotted for the two different viewpoints. As expected, file B with its higher caching probability of $p_B = 0.7$ has a better coverage than file A with $p_A = 0.3$. The figure also depicts the sharp decrease in the coverage probability of obtaining the farther cached file 2 compared to obtaining file 1 (around 17x lower at 0 dB). Similar trends can be observed for the case of average delay, with the higher cached file B requiring a lower number of time slots on average for successful transmission compared to the lower cached file A.

4.7.2 Effect of activity factor

Fewer the number of devices active in the network, lesser is the interference and better the coverage. This is the classic tradeoff between more aggressive frequency reuse (and hence higher throughput) and higher interference power. As evident from Fig. 4.4, the plots indicate a significant increase in coverage probability with decrease in the activity factors.

4.7.3 Critical SIR threshold

The mean number of time slots required for a successful file transmission increases with SIR threshold resulting in unsuccessful transmissions after a certain SIR threshold, defined as the cut-off threshold T_c. This is evident from Fig. 4.5 which also indicates the increasing cut-off threshold with the file's caching probability in the network. As was the case for coverage probability, it can be seen that lower delay values are encountered for lower activity factors due to a reduced interference power.

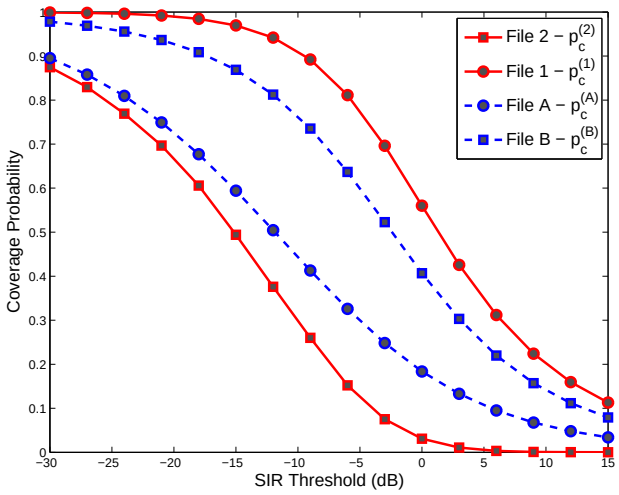

Figure 4.3: Coverage probability in distance-based viewpoint vs. file-type based viewpoint ($p_A = 0.3$, $q = 1$ and $\alpha = 4$).

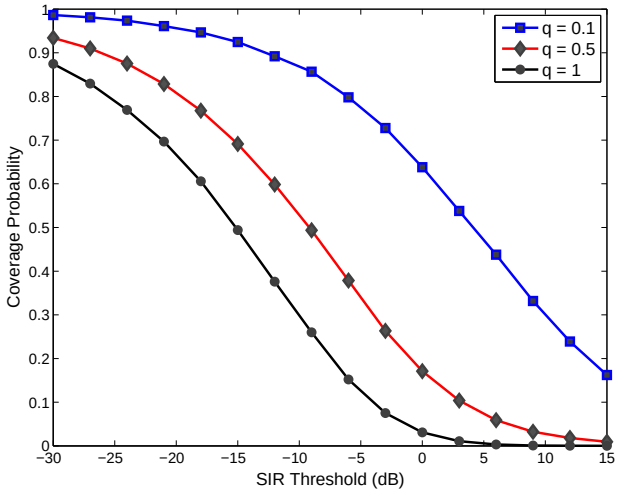

Figure 4.4: Coverage probability of file 2 for different activity factors q ($p_A = 0.3$ and $\alpha = 4$).

4.7.4 Effect of user mobility

Fig. 4.6 plots the coverage probability and average delay of file 2 for various levels of mobility in a 2-file distributed caching system. As evident from the figure, mobility increases the coverage probability of obtaining the farther cached file portion (file 2). For a static user, the low coverage of file 2 is due to the presence of a dominant interferer (file 1) located closer

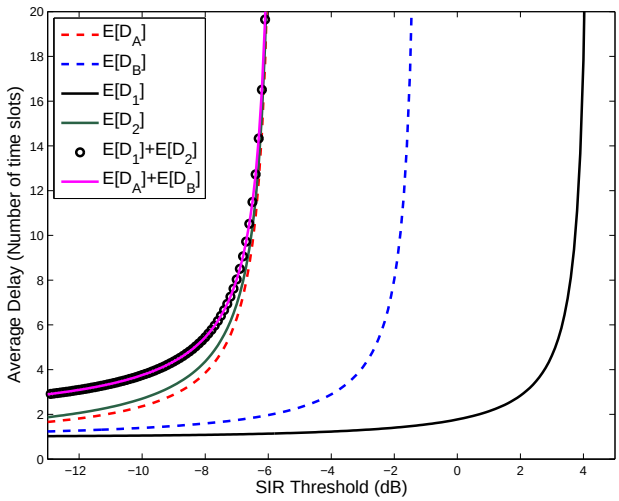

Figure 4.5: Average delay for the file-type based and distance based viewpoints ($p_A = 0.3, q = 0.5, \alpha = 4$).

to the user than the serving device (file 2). With user mobility, the likelihood of having a dominant interferer that is located closer to the serving device reduces, which improves the coverage of file 2. The above observation also explains the lower delay values encountered with user mobility as shown in Fig. 4.6 (*right*). Also, it can be observed that coverage probability decreases with activity factor; the corresponding plots are not shown here due to the lack of space.

4.8 Summary

The effect of mobility on the coverage probability and average delay of the farther cached file portion in a 2-file distributed caching D2D network is studied in this chapter. The novelty lies in the *exact* mobility-aware analysis that has to capture the information of the local neighborhood of nodes at the original location of the device. Our analysis concretely shows that user mobility improves this coverage probability and average delay in a uniform caching network. This improvement is a result of lower likelihood of having dominant interferer

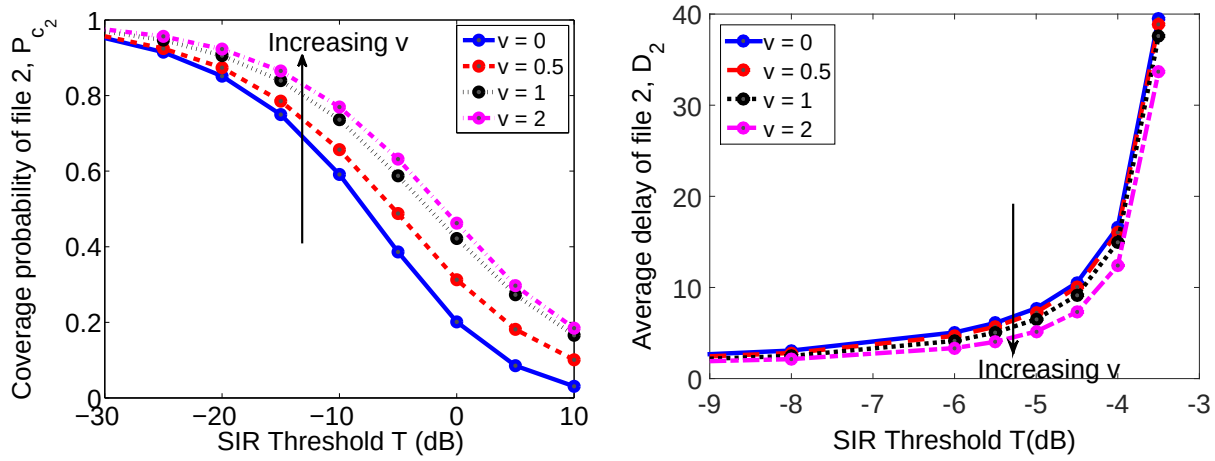

Figure 4.6: Effect of mobility v on the coverage probability and average delay of file 2 ($p_A = 0.5, q = 0.5, \alpha = 4,\ \lambda = 1$).

closer than the serving device.

Chapter 5

Conclusion

5.1 Summary

In this thesis, we modeled and analyzed spatio-temporal correlation in cellular networks and then studied specific applications exploiting spatio-temporal correlation in cache-enabled networks. Although there has been prior literature studying the extreme cases of correlation (uncorrelated and completely correlated scenarios), there is very limited work in studying the network performance under finite spatial and temporal correlation. The prime reason behind the limited literature is the difficulty involved in developing tractable analytical expressions for finite correlation scenario, as it involves careful examination of the neighbourhood at different user locations and time slots. In this thesis, using tools from stochastic geometry, we provide a simple analytical framework to study the performance of cellular network in terms of key metrics such as joint coverage probability and local delay for a moving user in a cellular network. Some of the key conclusions are mentioned below.

In Chapter 2, we provide an exact analytical framework to characterize the spatio-temporal interference correlation in cellular networks and the network performance in terms of joint coverage probability at two spatial locations separated by distance v. Our analysis shows

that the interference correlation and joint coverage probability decreases with v, with complete correlation for a static user and an uncorrelated (independent) scenario for the highly mobile user. Further insights are also provided by studying the effect of few network/channel parameters such as BS density and path loss on the interference correlation.

In Chapter 3, we studied an application of exploiting spatio-temporal correlation in cache-enabled small cell networks to determine correlation-aware optimal cache content. We studied this by considering a static user (completely correlated scenario) and a highly mobile user (uncorrelated scenario) in a a cache-enabled network for different retransmission attempts. Our results showed that while the static case tends to cache the most popular files in the library on each small cell, mobility de-emphasizes the importance of popularity of files and allows the small cells to cache content in a more balanced way. The overall network performance (in terms of hit probability and reception energy) was shown to be much better for the mobile case compared to the static case. This is due to the fact that when the user is mobile it comes across more *unique* small cells and hence more unique caches, which increases the probability of it being close to the small cell that has its file of interest.

In Chapter 4, we studied another application of exploiting spatio-temporal correlation by considering a distributed caching device-to-device (D2D) network, where a user's file of interest is cached as several portions in the storage of other devices in the network and obtained successively thorugh D2D communications. Although the file portions cached farther away present a bottle neck in the reception of entire file (due to more dominant interferers close-by), we show that user mobility (correlation) significantly improves network performance. More specifically, the coverage probability and average delay in a distributed caching network are shown to improve with user mobility and also with increase in cache probabilities of the file portions.

5.2 Future Directions

We now briefly discuss future research directions and possible extensions of the work in this thesis below.

Extend the model to a HetNet scenario to study cross-correlation across tiers

The models studied in this thesis that capture the effect of spatio-temporal correlation or user mobility in cellular networks are very simplistic and primarily studied for a single-tier cellular network. We studied the network performance considering a uniform set of BSs, either small cells (Chapter 3) or user devices (Chapter 4). Considering the densification of next generation cellular networks into a multi-tier HetNet having macrocells, small cells and other low power BSs with all tiers operating together, an interesting follow-up work would be to extend this work into a HetNet scenario. Considering the interference across tiers and different design parameters for each tier (different transmit powers or `SIR` thresholds) will lead to a better understanding of cross-tier correlation in heterogeneous cellular networks.

Develop bounds and approximations for analytical expressions

In Chapter 2, we studied interference and coverage correlation at two spatial locations (a certain distance apart) in cellular networks. Although the current work provides an exact analysis of joint coverage probability and spatio-temporal correlation coefficient, the expressions are complex and involve a lot of numerical integrations. An good follow up work would be developing approximations and bounds for the analytical expressions enabling the reader to get better analytical insights.

Extension of the work on optimal caching

In Chapter 3, we determined optimal cache content for a cache-enabled small cell network for a static and highly mobile user. This work has several promising extensions. First, we can extend this work to a multi-coverage scenario, where a typical user (static or mobile) is covered by multiple caches at a given spatial location, instead of just considering the single best cache (closest cache which has the file of interest) as done in Chapter 3. The corresponding optimal cache content and hit probability can be derived accordingly. Second, it is important to consider device caching (Chapter 4) along with the small cell caching studied in Chapter 3. Jointly optimizing the cache for devices and small cells, with different constraints on their cache storage capacities, is a meaningful problem to pursue. Third, users usually form physical clusters, e.g., see [71], which means it is important to determine optimal caching strategies for clustered networks. Also, the work on optimal caching for a mobile user in this thesis was to expose the fundamental principles using simple mobility models. It would be interesting to perform similar analysis using actual mobility traces and actual data for the SCBS locations.

Develop an advanced model for cache-enabled networks

For the sake of exposition and analytical tractability, we sometimes considered simplistic approximations which may not be relevant in real-life scenarios. A follow-up work would be extending the analytical framework studied in this thesis to model the actual scenarios. Few of them are mentioned next. In Chapter 3, we assumed the cache storage size of any SCBS as unity i.e. it can cache only one file. A simple extension would be to consider a larger cache storage for each SCBS and develop optimal cache by solving the appropriate optimization problem. Also, in Chapter 4, we studied the network performance of a 2-file distributed caching setup, where a a large requested file is split into two file portions and cached in the storage of user devices. An extension of this work can analyze a K-file distributed caching setup and study network metrics such as coverage probability and local delay.

Appendix A

A.1 Proof of Theorem 2

In a network where handoff skipping is used, the joint coverage probability of a typical user which moves a distance v from ℓ_1 to ℓ_2 is given as :

$$\mathbb{P}(\mathtt{SIR}_1 > T, \mathtt{SIR}_2 > T)$$

$$= \mathbb{E}_{R_1,\Theta}[\mathbb{P}(\mathtt{SIR}_1 > T, \mathtt{SIR}_2 > T | r_1, \theta)]$$

$$\stackrel{(a)}{=} \mathbb{E}_{R_1,\Theta}[P(h_{x_1}(1){r_1}^{-\alpha} > TI(1), h_{x_1}(2){r_{12}}^{-\alpha} > TI(2)|r_1, \theta)]$$

$$\stackrel{(b)}{=} \mathbb{E}_{R_1,\Theta}\Big[\exp\big(-T{r_1}^{\alpha}\sum_{x\in\Phi\setminus\{x_1\}} h_x(1)\|x-v\|^{-\alpha}\big)\exp\big(-T{r_{12}}^{\alpha}\sum_{x\in\Phi\setminus\{x_1\}} h_x(2)\|x\|^{-\alpha}\big)\Big]$$

$$= \mathbb{E}_{R_1,\Theta}\Big[\prod_{x\in\Phi\setminus\{x_1\}} \exp\big(-Tr_1^{\alpha}h_x(1)\|x-v\|^{-\alpha}\big)\exp(-Tr_{12}^{\alpha}h_x(2)\|x\|^{-\alpha})\Big]$$

$$\stackrel{(c)}{=} \mathbb{E}_{R_1,\Theta}\Big[\exp\Big(-\lambda\int_{\mathbb{R}^2\setminus\mathcal{C}_1} 1-\frac{1}{(1+Tr_1^{\alpha}\|x-v\|^{-\alpha})(1+Tr_{12}^{\alpha}\|x\|^{-\alpha})}\Big)\Big]$$

$$\stackrel{(d)}{=} \mathbb{E}_{R_1,\Theta,\Gamma}\Big[\exp\Big(-\lambda\int_{\mathbb{R}^2\setminus(\mathcal{C}_3\cup(\mathcal{C}_1\setminus\mathcal{C}_3))} \underbrace{\Big(1-\frac{1}{(1+Tr_1^{\alpha}(r^2+v^2-2rvcos\gamma)^{-\alpha/2})(1+Tr_{12}^{\alpha}r^{-\alpha})}\Big)}_{F_1(r_1,r,\gamma,\theta)} r\,\mathrm{d}r\Big)\Big]$$

$$\stackrel{(e)}{=} \mathbb{E}_{R_1,\Theta,\Gamma}\Big[\exp\Big(-\lambda\int_{\mathbb{R}^2\setminus C_3} F_1(r_1,r,\gamma,\theta)r\,\mathrm{d}r\Big)\exp\Big(\lambda\int_{C_1\setminus C_3} F_1(r_1,r,\gamma,\theta)r\,\mathrm{d}r\Big)\Big]$$

$$\stackrel{(f)}{=} \mathbb{E}_{R_1,\Theta,\Gamma}\Big[\exp\Big(-2\pi\lambda\int_{z_1}^{\infty} F_1(r_1,r,\gamma,\theta)\,r\,\mathrm{d}r\Big)\exp\Big(2\lambda\int_{|v-r_1|}^{v+r_1}\cos^{-1}\big(\frac{r^2+v^2-r_1^2}{2rv}\big)F_1(r_1,r,\gamma,\theta)\,r\,\mathrm{d}r\Big)\Big]$$

where (a) follows from the definition of $\texttt{SIR}_1$ and $\texttt{SIR}_2$ in (2.9) and (2.10) respectively and considering that the serving BS at ℓ_2 is at a distance r_{12} (same serving BS x_1 at ℓ_1 and ℓ_2 due to handoff skipping). Step (b) follows from the definition of $I(1)$ and $I(2)$ in (2.13) and (2.14) respectively and the spatial independence of the fading links, (c) results as $h_x(1) \sim \exp(1)$, $h_x(2) \sim \exp(1)$ and observing that the interference from the PPP Φ except x_1 is equivalent to considering an exclusion zone $\mathcal{C}_1$ in the network (as no BSs lie inside $\mathcal{C}_1$). Step (d) follows by converting the integral from Cartesian to polar coordinates where Γ is a uniform RV in $[0, \pi]$ and denote the angle of interferer $x \in \Phi$ w.r.t. user at ℓ_2. We also express $\mathcal{C}_1$ as the union of $\mathcal{C}_3$ and $\mathcal{C}_1 \setminus \mathcal{C}_3$ in (d), where $\mathcal{C}_3 = B(0, z_1)$ and $z_1 = \max(0, r_1 - v)$ (See Fig. 2.2). We split the integral into the two regions in (e), while the final result follows by using the law of cosines and appropriate limits of integration for the two regions (See [31] for a similar proof).

A.2 Proof of Theorem 3

The joint coverage probability of a typical user under no handoff scenario (event $\bar{H}$) is given as

$$
\begin{aligned}
&\mathbb{P}(\mathtt{SIR}_1 > T, \mathtt{SIR}_2 > T, \bar{H}) \\
&= \mathbb{E}_{R_1,\Theta}[\mathbb{P}(\mathtt{SIR}_1 > T, \mathtt{SIR}_2 > T, \bar{H}|r_1,\theta)] \\
&= \mathbb{E}_{R_1,\Theta}[P(h_{x_1}(1){r_1}^{-\alpha} > TI(1), h_{x_1}(2){r_2}^{-\alpha} > TI(2)|\bar{H}, r_1,\theta)\mathbb{P}(\bar{H}|r_1,\theta)] \\
&= \mathbb{E}_{R_1,\Theta}\Big[\exp\big(-T{r_1}^{\alpha}\sum_{x\in\Phi\setminus\{x_1\}} h_x(1)\|x-v\|^{-\alpha}\big)\exp\big(-T{r_{12}}^{\alpha}\sum_{x\in\Phi\setminus\{x_1\}} h_x(2)\|x\|^{-\alpha}\big)\mathbb{P}(\bar{H}|r_1,\theta)\Big] \\
&= \mathbb{E}_{R_1,\Theta}\Big[\mathbb{P}(\bar{H}|r_1,\theta)\prod_{x\in\Phi\setminus\{x_1\}}\exp\big(-Tr_1^{\alpha}h_x(1)\|x-v\|^{-\alpha}\big)\exp(-Tr_{12}^{\alpha}h_x(2)\|x\|^{-\alpha})\Big] \\
&\overset{(a)}{=} \mathbb{E}_{R_1,\Theta}\Big[\mathbb{P}(\bar{H}|r_1,\theta)\exp\Big(-\lambda\int_{\mathbb{R}^2\setminus\mathcal{C}_1\cup\mathcal{C}_2} 1-\frac{1}{(1+Tr_1^{\alpha}\|x-v\|^{-\alpha})(1+Tr_{12}^{\alpha}\|x\|^{-\alpha})}\Big)\Big] \\
&= \mathbb{E}_{R_1,\Theta,\Gamma}\Big[\mathbb{P}(\bar{H}|r_1,\theta)\exp\Big(-\lambda\int_{\mathbb{R}^2\setminus\mathcal{C}_1\cup\mathcal{C}_2}\underbrace{\Big(1-\frac{1}{(1+Tr_1^{\alpha}(r^2+v^2-2rvcos\gamma)^{-\alpha/2})(1+Tr_{12}^{\alpha}r^{-\alpha})}\Big)}_{F_1(r_1,r,\gamma,\theta)} r\,\mathrm{d}r\Big)\Big] \qquad (A.1)\\
&\overset{(b)}{=} \mathbb{E}_{R_1,\Theta,\Gamma}\Big[\mathbb{P}(\bar{H}|r_1,\theta)\exp\Big(-\lambda\int_{\mathbb{R}^2\setminus C_2} F_1(r_1,r,\gamma,\theta)r\,\mathrm{d}r\Big)\exp\Big(\lambda\int_{C_2\setminus C_1} F_1(r_1,r,\gamma,\theta)r\,\mathrm{d}r\Big)\Big] \\
&\overset{(c)}{=} \mathbb{E}_{R_1,\Theta,\Gamma}\Big[\mathbb{P}(\bar{H}|r_1,\theta)\exp\big(-2\pi\lambda\int_{r_{12}}^{\infty} F_1(r_1,r,\gamma,\theta)\,r\,\mathrm{d}r\big)\exp\big(2\lambda\int_{r_{12}}^{v+r_1}\cos^{-1}\big(\frac{r^2+v^2-r_1^2}{2rv}\big)F_1(r_1,r,\gamma,\theta)\,r\,\mathrm{d}r\big)\Big]
\end{aligned}
$$

where (a) follows by observing that the interference from PPP Φ except the serving BS x_1 in a no handoff scenario (see Fig. 2.1 (b)) is equivalent to considering an exclusion zone $\mathcal{C}_1 \cup \mathcal{C}_2$ in the network (no BSs lie inside $\mathcal{C}_1 \cup \mathcal{C}_2$). Step (b) follows by splitting the integral into two integration regions, while the final step (c) follows by using appropriate limits of integration for the two regions and using the law of cosines in the second integral.

A.3 Proof of Theorem 4

By definition, the joint coverage probability of a typical user under handoff scenario (event H) is given as:

$$
\begin{aligned}
&\mathbb{P}(\mathtt{SIR}_1 > T, \mathtt{SIR}_2 > T, H) \\
&= \mathbb{E}_{R_1,\Theta}[\mathbb{P}(\mathtt{SIR}_1 > T, \mathtt{SIR}_2 > T, H | r_1, \theta)] \\
&\overset{(a)}{=} \mathbb{E}_{R_1,\Theta}[P(h_{x_1}(1) {r_1}^{-\alpha} > TI(1), h_{x_2}(2) {r_2}^{-\alpha} > TI(2) | H, r_1, \theta)\mathbb{P}(H|r_1,\theta)] \\
&= \mathbb{E}_{R_1,R_2|H,\Theta}\Bigg[\exp\big(-T{r_1}^{\alpha}\sum_{x\in\Phi\setminus\{x_1\}} h_x(1)\|x-v\|^{-\alpha}\big)\exp\big(-T{r_2}^{\alpha}\sum_{x\in\Phi\setminus\{x_2\}} h_x(2)\|x\|^{-\alpha}\big)\mathbb{P}(H|r_1,\theta)\Bigg] \\
&\overset{(b)}{=} \mathbb{E}_{R_1,R_2|H,\Theta}\Bigg[\mathbb{P}(H|r_1,\theta)\exp(-T{r_1}^{\alpha}h_{x_2}(1)\|x_2-v\|^{-\alpha})\exp(-T{r_2}^{\alpha}h_{x_1}(2)\|x_1\|^{-\alpha}) \\
&\qquad \prod_{x\in\Phi\setminus(\{x_1\}\cup\{x_2\})} \exp\big(-Tr_1^{\alpha}h_x(1)\|x-v\|^{-\alpha}\big)\exp(-Tr_2^{\alpha}h_x(2)\|x\|^{-\alpha})\Bigg] \\
&\overset{(c)}{=} \mathbb{E}_{R_1,R_2|H,\Theta}\Bigg[\mathbb{P}(H|r_1,\theta)\frac{1}{1+Tr_1^{\alpha}\|x_2-v\|^{-\alpha}}\frac{1}{1+Tr_2^{\alpha}\|x_1\|^{-\alpha}} \\
&\qquad \exp\Bigg(-\lambda\int_{\mathbb{R}^2\setminus\mathcal{C}_1\cup\mathcal{C}_2} 1-\frac{1}{(1+Tr_1^{\alpha}\|x-v\|^{-\alpha})(1+Tr_2^{\alpha}\|x\|^{-\alpha})}\Bigg)\Bigg] \\
&= \mathbb{E}_{R_1,R_2|H,\Theta,\Gamma}\Bigg[\mathbb{P}(H|r_1,\theta)\frac{1}{1+Tr_1^{\alpha}\|x_2-v\|^{-\alpha}}\frac{1}{1+Tr_2^{\alpha}\|x_1\|^{-\alpha}} \\
&\qquad \exp\Bigg(-\lambda\int_{\mathbb{R}^2\setminus\mathcal{C}_1\cup\mathcal{C}_2} \underbrace{\Bigg(1-\frac{1}{(1+Tr_1^{\alpha}(r^2+v^2-2rvcos\gamma)^{-\alpha/2})(1+Tr_2^{\alpha}r^{-\alpha})}\Bigg)}_{F_2(r_1,r_2,r,\gamma)} r\,\mathrm{d}r\Bigg)\Bigg] \\
&= \mathbb{E}_{R_1,R_2|H,\Theta,\Gamma}\Bigg[\mathbb{P}(H|r_1,\theta)\frac{1}{1+Tr_1^{\alpha}\|x_2-v\|^{-\alpha}}\frac{1}{1+Tr_2^{\alpha}\|x_1\|^{-\alpha}} \\
&\qquad \exp\Bigg(-\lambda\int_{\mathbb{R}^2\setminus C_2} F_2(r_1,r_2,r,\gamma) r\,\mathrm{d}r\Bigg)\exp\Bigg(\lambda\underbrace{\int_{C_1\setminus C_2} F_2(r_1,r_2,r,\gamma) r\,\mathrm{d}r}_{\mathcal{B}_1(r_1,r_2,\gamma)}\Bigg)\Bigg] \\
&\overset{(d)}{=} \mathbb{E}_{R_1,R_2|H,\Theta,\Gamma}\Bigg[\mathbb{P}(H|r_1,\theta)\frac{1}{1+Tr_1^{\alpha}\|x_2-v\|^{-\alpha}}\frac{1}{1+Tr_2^{\alpha}\|x_1\|^{-\alpha}} \\
&\qquad \exp\Bigg(-2\pi\lambda\int_{r_2}^{\infty} F_2(r_1,r_2,r,\gamma)\, r\,\mathrm{d}r\Bigg)\exp\Big(\lambda\mathcal{B}_1(r_1,r_2,\gamma)\Big)\Bigg]
\end{aligned}
$$

where (a) follows by using the definition of $\mathtt{SIR}_1$ and $\mathtt{SIR}_2$ for a handoff scenario (serving BS x_1 at distance r_1 at ℓ_1 and serving BS x_2 at distance r_2 at ℓ_2) and conditioning w.r.t.

event H (occurence of handoff). Step (b) follows by splitting the interference into three parts - i) First term corresponds to interference from BS x_2 at ℓ_1, ii) Second term signifies the interference from BS x_1 at ℓ_2 and iii) third term corresponds to interference from all BSs $x \in \Phi$ except x_1 and x_2 (equivalent to exclusion zone of $\mathcal{C}_1 \cup \mathcal{C}_2$). Step (c) results by taking expectation over the fading links and applying Campbell's law to the interference in the third term, while the final step (d) follows by applying suitable limits of integration to the different integration regions. The limits of the integration region $\mathcal{C}_1 \setminus \mathcal{C}_2$ depends on the three cases (disjoint, intersecting or engulfed) as defined in Remark 1, with its lower limit a (See Fig. 2.2) taking values $v - r_1, r_2$ for cases 1 and 2 respectively. The integration region is zero for case 3 as $\mathcal{C}_1 \setminus \mathcal{C}_2 = \phi$ ($\mathcal{C}_1$ is engulfed inside $\mathcal{C}_2$). See [31] for a similar treatment. The integral $\mathcal{B}_1(r_1, r_2, \gamma)$ is summarized below.

$$\mathcal{B}_1(r_1, r_2, \gamma) = \begin{cases} \int\limits_{v-r_1}^{v+r_1} 2\cos^{-1}\left(\frac{r^2+v^2-r_1^2}{2rv}\right) F(r_1, r_2, r, \gamma)\, r\, \mathrm{d}r, & \text{case 1} \\ \int\limits_{r_2}^{v+r_1} 2\cos^{-1}\left(\frac{r^2+v^2-r_1^2}{2rv}\right) F(r_1, r_2, r, \gamma)\, r\, \mathrm{d}r, & \text{case 2} \\ 0, & \text{case 3} \end{cases} \tag{A.2}$$

Appendix B

B.1 Proof of Theorem 10

Conditioned on r_1 and r_2, the average delay of obtaining file 2 is dictated by the average delay from three interference regions individually (I_1, I_2 and I_3) and is given as

$$\begin{aligned}
E[D_2]^{(Y)} &= \int_0^\infty \int_{r_1}^\infty \mathbb{E}[D_{I_1}]\mathbb{E}[D_{I_2}]\mathbb{E}[D_{I_3}] \\
&\qquad f_{R_2|R_1}(r_2|r_1) f_{R_1}(r_1)\,\mathrm{d}r_2\,\mathrm{d}r_1 \qquad \text{(B.1)} \\
\mathbb{E}[D_{I_1}] &= \mathbb{E}_\Phi\left[\frac{1}{1-q+q\mathbb{E}[e^{-sh_x r_1^{-\alpha}}]}\right] \overset{(a)}{=} \frac{1}{1-q+\frac{q}{1+sr_1^{-\alpha}}} \\
\mathbb{E}[D_{I_2}] &= \mathbb{E}_\Phi\left[\prod_{y\in\Phi_B\setminus(\{x\}\cup B(0,r_1))} \frac{1}{1-q+q\mathbb{E}[e^{-sh_y\|y\|^{-\alpha}}]}\right] \\
&\overset{(b)}{=} \exp\left(-2\pi p_B\lambda\int_{r_1}^\infty\left(1-\frac{1}{1-q+\frac{q}{1+sr^{-\alpha}}}\right) r\,\mathrm{d}r\right) \\
&= \exp\left(2\pi p_B q\lambda\int_{r_1}^\infty \frac{r\,\mathrm{d}r}{1-q+\frac{r^\alpha}{s}}\right) \\
\mathbb{E}[D_{I_3}] &= \mathbb{E}_\Phi\left[\prod_{z\in\Phi_A\setminus B(0,r_2)} \frac{1}{1-q+q\mathbb{E}[e^{-sh_z\|z\|^{-\alpha}}]}\right] \\
&\overset{(c)}{=} \exp\left(-2\pi p_A\lambda\int_{r_2}^\infty\left(1-\frac{1}{1-q+\frac{q}{1+sr^{-\alpha}}}\right) r\,\mathrm{d}r\right) \\
&\overset{(d)}{=} \exp(\pi p_A q\lambda r_2^2\rho_3(T,\alpha))
\end{aligned}$$

where (a) follows from the fact that $h_x \sim \exp(1)$, (b) and (c) result from the PGFL of the PPP Φ while considering the activity factor q of the interferers, and (d) follows by rearranging a

few terms and defining $\rho_3(T,\alpha)$ as in (4.19). The result now follows by substituting $s = Tr_2^{\alpha}$.

B.2 Proof of Lemma 10

The conditional Laplace transform of interference I_2 (all file B devices except the *singleton*) is written as

$$\mathcal{L}_{I_2|R_1}(s|r_1) = \mathbb{E}\Big[\exp\Big(-s\sum_{y\in\Phi_B\setminus\mathcal{C}_1} t_y h_y \|y\|^{-\alpha}\Big)\Big] \overset{(a)}{=} \mathbb{E}\Big[\prod_{y\in\Phi_B\setminus\mathcal{C}_1} 1-q+\frac{q}{1+s\|y\|^{-\alpha}}\Big]$$
$$\overset{(b)}{=} \exp\Big(-p_B q\lambda \int\limits_{\mathbb{R}^2\setminus(\mathcal{C}_3\cup(\mathcal{C}_1\setminus\mathcal{C}_3))} \frac{\mathrm{d}y}{1+\frac{\|y\|^{\alpha}}{s}}\Big)$$
$$\overset{(c)}{=} \exp\Big(-p_B q\lambda\Big(\int\limits_{\mathbb{R}^2\setminus\mathcal{C}_3} \frac{\mathrm{d}y}{1+\frac{\|y\|^{\alpha}}{s}} - \int\limits_{\mathcal{C}_1\setminus\mathcal{C}_3} \frac{\mathrm{d}y}{1+\frac{\|y\|^{\alpha}}{s}}\Big)\Big)$$
$$\overset{(d)}{=} \exp\Big(-p_B q\lambda\Big(\int\limits_{z_1}^{\infty} \frac{2\pi r\,\mathrm{d}r}{1+\frac{r^{\alpha}}{s}} - \int\limits_{|v-r_1|}^{v+r_1} \frac{2\cos^{-1}\big(\frac{r^2+v^2-r_1^2}{2rv}\big) r\,\mathrm{d}r}{1+\frac{r^{\alpha}}{s}}\Big)\Big)$$

where (a) follows from $h_y \sim \exp(1)$ while considering the activity factor q of the interferers, (b) results from the PGFL of the PPP Φ_B and expressing $\mathcal{C}_1$ as the union of $\mathcal{C}_3$ and $\mathcal{C}_1 \setminus \mathcal{C}_3$, where $\mathcal{C}_3 = \mathbf{b}(0, z_1)$ and $z_1 = \mathtt{max}(0, r_1 - v)$ (See Fig. 4.2), and (c) follows by splitting the integral into the two regions. Further, we obtain (d) by converting the integral from Cartesian to polar coordinates and using the law of cosines in which $r^2 + v^2 - 2rv\cos\gamma = r_1^2$ (see Fig. 4.2). The lower limit of the integration region of $\mathcal{C}_1 \setminus \mathcal{C}_3$ (dark shaded region in Fig. 4.2) takes values of $r_1 - v$ and $v - r_1$ for scenarios 1 and 2 respectively, therefore we use the absolute value $|v - r_1|$ as the lower limit to capture both scenarios.

B.3 Proof of Lemma 11

Proceeding similar to (a) in Appendix B.2, the conditional Laplace transform of I_3, $\mathcal{L}_{I_3|R_1,R_2}(s|r_1,r_2)$

$$
\begin{aligned}
&= \mathbb{E}\Bigg[\prod_{z\in\Phi_A\setminus(\mathcal{C}_1\cup\mathcal{C}_2)} 1-q+\frac{q}{1+s\|z\|^{-\alpha}}\Bigg] \\
&\overset{(a)}{=} \exp\Bigg(-p_A q\lambda\Bigg(\int\limits_{\mathbb{R}^2\setminus\mathcal{C}_2}\frac{\mathrm{d}z}{1+\frac{\|z\|^\alpha}{s}}-\int\limits_{\mathcal{C}_1\setminus\mathcal{C}_2}\frac{\mathrm{d}z}{1+\frac{\|z\|^\alpha}{s}}\Bigg)\Bigg) \\
&\overset{(b)}{=} \exp\Bigg(-p_A q\lambda\int\limits_{r_2}^{\infty}\frac{2\pi r\,\mathrm{d}r}{1+\frac{r^\alpha}{s}}\Bigg)\exp\Bigg(p_A q\lambda\underbrace{\int\limits_{\mathcal{C}_1\setminus\mathcal{C}_2}\frac{2f(r,r_1)\,r\,\mathrm{d}r}{1+\frac{r^\alpha}{s}}}_{\mathcal{B}(r_1,r_2,v)}\Bigg)\Bigg)
\end{aligned}
$$

where (a) results from the PGFL of the PPP Φ_A and splitting the integral into two regions, (b) follows by converting the integral from Cartesian to polar coordinates and using the law of cosines with $f(r,r_1)$ defined in Lemma 10. The integral in the second term of (b) depends on the integration region $\mathcal{C}_1\setminus\mathcal{C}_2$ with its lower limit (denoted by a in Fig. 4.2) taking values $\{v-r_1,r_2\}$ for cases 1 and 2 . The integration region is zero for case 3 as $\mathcal{C}_1\setminus\mathcal{C}_2=\phi$ ($\mathcal{C}_1$ is engulfed inside $\mathcal{C}_2$). The integral $\mathcal{B}(r_1,r_2,v)$ is summarized below.

$$
\mathcal{B}(r_1,r_2,v)=\begin{cases}\int_{v-r_1}^{v+r_1}\frac{2f(r,r_1)\,r\,\mathrm{d}r}{1+\frac{r^\alpha}{s}}, & \text{case 1}\\ \int_{r_2}^{v+r_1}\frac{2f(r,r_1)\,r\,\mathrm{d}r}{1+\frac{r^\alpha}{s}}, & \text{case 2}\\ 0, & \text{case 3}\end{cases} \tag{B.2}
$$

B.4 Proof of Theorem 12

Conditioned on R_1, R_2 and Θ, the local delay of obtaining file 2 is dictated by the local delay from three interference regions individually (D_{I_1}, D_{I_2} and D_{I_3}). By definition

$$D_{I_1} \overset{(a)}{=} \mathbb{E}_{\Phi}\left[\frac{1}{1-q+q\mathbb{E}[e^{-sh_x {r_{12}}^{-\alpha}}]}\right] \overset{(b)}{=} \frac{1}{1-q+\frac{q}{1+sr_{12}^{-\alpha}}}$$

$$D_{I_2} = \mathbb{E}_{\Phi}\left[\frac{1}{\mathcal{L}_{I_2|R_1}(s|r_1,\Phi)}\right] \overset{(c)}{=} \mathbb{E}_{\Phi}\left[\prod_{y\in\Phi_B\setminus\mathcal{C}_1}\frac{1}{1-q+q\mathbb{E}[e^{-sh_y\|y\|^{-\alpha}}]}\right]$$

$$\overset{(d)}{=} \exp\left(-2\pi p_B\lambda\int_{\mathbb{R}^2\setminus(\mathcal{C}_3\cup(\mathcal{C}_1\setminus\mathcal{C}_3))}\left(1-\frac{1}{1-q+\frac{q}{1+sr^{-\alpha}}}\right)r\,\mathrm{d}r\right)$$

$$D_{I_3} \overset{(e)}{=} \mathbb{E}_{\Phi}\left[\prod_{z\in\Phi_A\setminus(\mathcal{C}_1\cup\mathcal{C}_2)}\frac{1}{1-q+q\mathbb{E}[e^{-sh_z\|z\|^{-\alpha}}]}\right]$$

$$\overset{(f)}{=} \exp\left(p_A q\lambda\int_{r_2}^{\infty}\frac{2\pi r\,\mathrm{d}r}{1-q+\frac{r^{\alpha}}{s}}\right)\exp\Big(-p_A q\lambda\mathcal{B}_1(r_1,r_2,v)\Big),$$

$$\text{where } \mathcal{B}_1(r_1,r_2,v) = \begin{cases} \int_{v-r_1}^{v+r_1}\frac{2f(r,r_1)\,r\,\mathrm{d}r}{1-q+\frac{r^{\alpha}}{s}}, & \text{case 1} \\ \int_{r_2}^{v+r_1}\frac{2f(r,r_1)\,r\,\mathrm{d}r}{1-q+\frac{r^{\alpha}}{s}}, & \text{case 2} \\ 0, & \text{case 3} \end{cases} \tag{B.3}$$

where (a) follows by considering the activity factor q of the interferers, (b) results from $h_x \sim \exp(1)$ and the final expression of D_{I_1} obtained by using $r_{12}^2 = r_1^2 + v^2 - 2r_1 v\cos\theta$. Proceeding similar to (a) and (b), we obtain (c) while (d) results from the PGFL of the PPP Φ_B and converting the integral from Cartesian to polar coordinates. The final expression of D_{I_2} is obtained by proceeding similar to the last step in Appendix B.2. Steps (e) and (f) follow by proceeding similar to Appendix B.3.

Bibliography

[1] X. Liu and M. Haenggi, "Throughput analysis of fading sensor networks with regular and random topologies," *EURASIP J. Wirel. Commun. Netw.*, vol. 2005, no. 4, pp. 554–564, September 2005.

[2] K. Hong and Y. Hua, "Throughput analysis of large wireless networks with regular topologies," *EURASIP Journal on Wireless Communications and Networking*, vol. 2007, no. 1, pp. 1–11, 2007.

[3] J. G. Andrews, R. K. Ganti, M. Haenggi, N. Jindal, and S. Weber, "A primer on spatial modeling and analysis in wireless networks," *IEEE Communications Magazine*, vol. 48, no. 11, pp. 156–163, November 2010.

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

[5] M. Haenggi, *Stochastic Geometry for Wireless Networks*. New York: Cambridge University Press, 2013.

[6] F. Baccelli and B. Blaszczyszyn, *Stochastic Geometry and Wireless Networks*. NOW: Foundations and Trends in Networking, 2010.

[7] Dino Flore, "RAN workshop on 5G: Chairman Summary", RAN workshop on 5G, Sept. 2015. Available online: goo.gl/dwaefk.

[8] Ericsson, "5G Use Cases", white paper, 2015. Available online: http://goo.gl/53OuaG.

[9] I. Hwang, B. Song, and S. S. Soliman, "A holistic view on hyper-dense heterogeneous and small cell networks," *IEEE Commun. Magazine*, vol. 51, no. 6, pp. 20–27, 2013.

[10] H. S. Dhillon and G. Caire, "Wireless backhaul networks: Capacity bound, scalability analysis and design guidelines," *IEEE Trans. on Wireless Commun.*, vol. 14, no. 11, pp. 6043 – 6056, Nov. 2015.

[11] N. Golrezaei, A. F. Molisch, A. G. Dimakis, and G. Caire, "Femtocaching and device-to-device collaboration: A new architecture for wireless video distribution," *IEEE Commun. Magazine,*, vol. 51, no. 4, pp. 142–149, 2013.

[12] E. Bastug, M. Bennis, and M. Debbah, "Living on the edge: The role of proactive caching in 5G wireless networks," *IEEE Commun. Magazine*, vol. 52, no. 8, pp. 82–89, 2014.

[13] K. Shanmugam, N. Golrezaei, A. Dimakis, A. Molisch, and G. Caire, "Femtocaching: Wireless content delivery through distributed caching helpers," *IEEE Trans on Information Theory*, vol. 59, no. 12, pp. 8402–8413, Dec. 2013.

[14] J. Li, Y. Chen, Z. Lin, W. Chen, B. Vucetic, and L. Hanzo, "Distributed caching for data dissemination in the downlink of heterogeneous networks," *IEEE Trans. on Commun.*, vol. 63, no. 10, pp. 3553–3568, Oct. 2015.

[15] M. Cha, H. Kwak, P. Rodriguez, Y.-Y. Ahn, and S. Moon, "I Tube, You Tube, Everybody Tubes: Analyzing the world's largest user generated content video system," in *Proc. of the 7th ACM SIGCOMM conf. on Internet measurement*, 2007, pp. 1–14.

[16] M. E. J. Newman, "Power laws, Pareto distributions and Zipf's law," *Contemporary Physics*, 2005.

[17] S. Krishnan and H. S. Dhillon, "Spatio-temporal interference correlation and joint coverage in cellular networks," submitted to *IEEE Trans. on Wireless Commun.*, June 2016. Available online: https://arxiv.org/abs/1606.05332.

[18] R. K. Ganti and M. Haenggi, "Spatial and temporal correlation of the interference in aloha ad hoc networks," *IEEE Communications Letters*, vol. 13, no. 9, pp. 631–633, Sept 2009.

[19] U. Schilcher, S. Toumpis, A. Crismani, G. Brandner, and C. Bettstetter, "How Does Interference Dynamics Influence Packet Delivery in Cooperative Relaying?" in *Proc. of the 16th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM '13)*, 2013.

[20] M. Haenggi, "Diversity loss due to interference correlation," *IEEE Communications Letters*, vol. 16, no. 10, pp. 1600–1603, October 2012.

[21] R. Tanbourgi, H. Jäkel, and F. K. Jondral, "Cooperative relaying in a poisson field of interferers: A diversity order analysis," in *Proc. ISIT*, July 2013.

[22] A. Crismani, S. Toumpis, U. Schilcher, G. Brandner, and C. Bettstetter, "Cooperative relaying under spatially and temporally correlated interference," *IEEE Trans. on Vehicular Technology*, vol. 64, no. 10, pp. 4655–4669, Oct 2015.

[23] Y. Zhou and W. Zhuang, "Performance analysis of cooperative communication in decentralized wireless networks with unsaturated traffic," *IEEE Trans.on Wireless Commun.*, vol. 15, no. 5, pp. 3518–3530, May 2016.

[24] A. Altieri, L. R. Vega, C. G. Galarza, and P. Piantanida, "Cooperative strategies for interference-limited wireless networks," in *Proc. ISIT*, July 2011.

[25] R. Tanbourgi, H. S. Dhillon, J. G. Andrews, and F. K. Jondral, "Dual-branch mrc receivers under spatial interference correlation and nakagami fading," *IEEE Trans. on Commun.*, vol. 62, no. 6, pp. 1830–1844, June 2014.

[26] ——, "Effect of spatial interference correlation on the performance of maximum ratio combining," *IEEE Trans. on Wireless Commun.*, vol. 13, no. 6, pp. 3307–3316, June 2014.

[27] M. Haenggi and R. Smarandache, "Diversity polynomials for the analysis of temporal correlations in wireless networks," *IEEE Trans. on Wireless Commun.*, vol. 12, no. 11, pp. 5940–5951, November 2013.

[28] U. Schilcher, C. Bettstetter, and G. Brandner, "Temporal correlation of interference in wireless networks with Rayleigh block fading," *IEEE Trans. on Mobile Computing*, vol. 11, no. 12, pp. 2109–2120, Dec 2012.

[29] R. K. Ganti and J. G. Andrews, "Correlation of link outages in low-mobility spatial wireless networks," in *Proc. Asilomar*, Nov 2010, pp. 312–316.

[30] M. Haenggi, "The meta distribution of the SIR in poisson bipolar and cellular networks," *IEEE Trans. on Wireless Commun.*, vol. 15, no. 4, pp. 2577–2589, April 2016.

[31] S. Krishnan and H. S. Dhillon, "Effect of user mobility on the performance of device-to-device networks with distributed caching," submitted to *IEEE Trans. on Vehicular Technology*, 2016, available online: https://arxiv.org/abs/1604.07088.

[32] Z. Gong and M. Haenggi, "The local delay in mobile poisson networks," *IEEE Transactions on Wireless Communications*, vol. 12, no. 9, pp. 4766–4777, September 2013.

[33] K. Koufos and C. P. Dettmann, "Temporal correlation of interference and outage in mobile networks with correlated mobility in finite regions," 2016, available online: http://www.maths.bris.ac.uk/ macpd/sen/TempCorrInterf.pdf.

[34] J. Wen, M. Sheng, K. Huang, and J. Li, "Analysis of interference correlation in non-poisson networks," submitted to *IEEE Globecom*, 2016, available online: https://arxiv.org/abs/1604.04166.

[35] X. Zhang and M. Haenggi, "A Stochastic Geometry Analysis of Inter-Cell Interference Coordination and Intra-Cell Diversity," *IEEE Trans. on Wireless Commun.*, vol. 13, no. 12, pp. 6655–6669, Dec 2014.

[36] G. Nigam, P. Minero, and M. Haenggi, "Spatiotemporal cooperation in heterogeneous cellular networks," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 6, pp. 1253–1265, June 2015.

[37] J. Wen, M. Sheng, B. Liang, X. Wang, Y. Zhang, and J. Li, "Correlations of interference and link successes in heterogeneous cellular networks," in *Proc. IEEE Globecom*, Dec 2015, pp. 1–6.

[38] F. Baccelli, B. Blaszczyszyn, and P. Muhlethaler, "Stochastic analysis of spatial and opportunistic aloha," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 7, pp. 1105–1119, September 2009.

[39] 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 Commun.*, vol. 11, no. 10, pp. 3484–3495, October 2012.

[40] H. S. Dhillon, R. K. 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, April 2012.

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

[42] K. Gulati, R. K. Ganti, J. G. Andrews, B. L. Evans, and S. Srikanteswara, "Characterizing decentralized wireless networks with temporal correlation in the low outage regime," *IEEE Trans. on Wireless Commun.*, vol. 11, no. 9, pp. 3112–3125, September 2012.

[43] M. Kassar, B. Kervella, and G. Pujolle, "An overview of vertical handover decision strategies in heterogeneous wireless networks," *Computer Communications*, vol. 31, no. 10, pp. 2607 – 2620, 2008.

[44] I. Hwang, B. Song, and S. S. Soliman, "A holistic view on hyper-dense heterogeneous and small cell networks," *IEEE Communications Magazine*, vol. 51, no. 6, pp. 20–27, June 2013.

[45] R. Arshad, H. ElSawy, S. Sorour, T. Y. Al-Naffouri, and M.-S. Alouini, "Handover management in dense cellular networks: A stochastic geometry approach," in *Proc. IEEE ICC*, May 2016.

[46] M. K. E. Bastug, M. Bennis and M. Debbah, "Cache-enabled small cell networks: modeling and tradeoffs," *EURASIP Journal on Wireless Communications and Networking*, vol. 2015, no. 1, pp. 1–11.

[47] J. Rao, H. Feng, C. Yang, Z. Chen, and B. Xia, "Optimal caching placement for D2D assisted wireless caching networks," submitted to *ICC 2016*, available online: http://arxiv.org/abs/1510.07865.

[48] B. Blaszczyszyn and A. Giovanidis, "Optimal geographic caching in cellular networks," in *Proc. IEEE ICC*, June 2015.

[49] B. Serbetci and J. Goseling, "On optimal geographical caching in heterogeneous cellular networks," submitted to *ISIT 2016*, available online: http://arxiv.org/abs/1601.07322.

[50] M. Haenggi, "The local delay in Poisson networks," *IEEE Trans. on Info. Theory*, vol. 59, no. 3, pp. 1788–1802, 2013.

[51] D. C. Chen, T. Q. S. Quek, and M. Kountouris, "Backhauling in heterogeneous cellular networks: Modeling and tradeoffs," *IEEE Trans. on Wireless Commun.*, vol. 14, no. 6, pp. 3194–3206, June 2015.

[52] C. Jarray and A. Giovanidis, "The effects of mobility on the hit performance of cached D2D networks," in *Proc. WIOPT-SPASWIN, Arizona, USA*, 2016.

[53] J. G. Andrews, F. Baccelli, and R. K. Ganti, "A tractable approach to coverage and rate in cellular networks," vol. 59, no. 11, pp. 3122–3134, Nov. 2011.

[54] 2015 IEEE. Reprinted, with permission, from Shankar Krishnan and H. S. Dhillon, "Distributed Caching in Device-to-Device Networks: A Stochastic Geometry Perspective," in *Proc., IEEE Asilomar*, 2015.

[55] B. Perabathini, E. Bastug, M. Kountouris, M. Debbah, and A. Conte, "Caching at the edge: A green perspective for 5G networks," in *Proc. IEEE ICCW*, June 2015.

[56] N. Golrezaei, A. F. Molisch, A. G. Dimakis, and G. Caire, "Femtocaching and device-to-device collaboration: A new architecture for wireless video distribution," *IEEE Communications Magazine*, vol. 51, no. 4, pp. 142–149, April 2013.

[57] E. Altman, K. Avrachenkov, and J. Goseling, "Distributed storage in the plane," in *Proc. IFIP Networking Conference*, 2014.

[58] N. Golrezaei, A. G. Dimakis, and A. F. Molisch, "Wireless device-to-device communications with distributed caching," in *Proc. IEEE ISIT*, Cambridge, MA, July 2012.

[59] M. A. Maddah-Ali and U. Niesen, "Fundamental limits of caching," *IEEE Trans. on Info. Theory*, vol. 60, no. 5, pp. 2856 – 2867, May 2014.

[60] E. Bastug, M. Bennis, and M. Debbah, "Cache-enabled small cell networks: Modeling and tradeoffs," in *Proc. Int. Symp. on Wireless Commun. Systems (ISWCS)*, 2014.

[61] M. Ji, G. Caire, and A. F. Molisch, "Wireless device-to-device caching networks: Basic principles and system performance," *IEEE Journal on Sel. Areas in Commun.*, to appear.

[62] N. Golrezaei, K. Shanmugam, A. G. Dimakis, A. F. Molisch, and G. Caire, "Femtocaching: Wireless video content delivery through distributed caching helpers," in *Proc. IEEE INFOCOM*, 2012.

[63] D. Malak and M. Al-Shalash, "Optimal caching for device-to-device content distribution in 5G networks," in *Proc. IEEE Globecom Workshops*, Dec 2014.

[64] S. K. Dandapat, S. Pradhan, N. Ganguly, and R. Roy Choudhury, "Sprinkler: Distributed content storage for just-in-time streaming," in *Proc. ACM Workshop on Cellular Networks: Operations, Challenges, and Future Design (CellNet'13)*, Jun 2013.

[65] K. Poularakis and L. Tassiulas, "Exploiting user mobility for wireless content delivery," in *Proc. IEEE ISIT*, Istanbul, Turkey, July 2013.

[66] X. Lin, R. Ratasuk, A. Ghosh, and J. G. Andrews, "Modeling, analysis, and optimization of multicast device-to-device transmissions," *IEEE Transactions on Wireless Communications*, vol. 13, no. 8, pp. 4346–4359, Aug 2014.

[67] J. G. Andrews, F. Baccelli, and R. K. Ganti, "A tractable approach to coverage and rate in cellular networks," *IEEE Trans. Commun.*, vol. 59, no. 11, pp. 3122–3134, 2011.

[68] W. Nie, Z. Yi, F. Zheng, W. Zhang, and T. O'Farrell, "HetNets with random DTX scheme: Local delay and energy efficiency," *IEEE Trans. on Veh. Technology*, to appear.

[69] Z. Han and K. J. R. Liu, *Resource Allocation for Wireless Networks: Basics, Techniques, and Applications*. Cambridge, U.K.: Cambridge University Press, 2008.

[70] Z. Yazdanshenasan, H. S. Dhillon, M. Afshang, and P. H. J. Chong, "Poisson hole process: Theory and applications to wireless networks," submitted to *IEEE Trans. on Wireless Commun.*, 2015, available online: http://arxiv.org/abs/1601.01090.

[71] M. Afshang, H. S. Dhillon, and P. H. J. Chong, "Fundamentals of cluster-centric content placement in cache-enabled device-to-device networks," *IEEE Trans. on Communications*, vol. PP, no. 99, pp. 1–1, 2016.