

Performance Analysis and Optimization of Cache-Assisted CoMP for Clustered D2D Networks

Ramy Amer, *Student Member, IEEE*, Hesham ElSawy, *Senior Member, IEEE*,
 Jacek Kibilda, *Member, IEEE*, M. Majid Butt, *Senior Member, IEEE*,
 and Nicola Marchetti, *Senior Member, IEEE*

Abstract—Caching at mobile devices and leveraging cooperative device-to-device (D2D) communications are two promising approaches to support massive content delivery over wireless networks while mitigating the effects of interference. To show the impact of cooperative communication on the performance of cache-enabled D2D networks, the notion of device clustering must be factored in to convey a realistic description of the network performance. In this regard, this paper develops a novel mathematical model, based on stochastic geometry and an optimization framework for cache-assisted coordinated multi-point (CoMP) transmissions with clustered devices. Devices are spatially distributed into disjoint clusters and are assumed to have a surplus memory to cache files from a known library, following a random probabilistic caching scheme. Desired contents that are not self-cached can be obtained via D2D CoMP transmissions from neighboring devices or, as a last resort, from the network. For this model, we analytically characterize the offloading gain and rate coverage probability as functions of the system parameters. An optimal caching strategy is then defined as the content placement scheme that maximizes the offloading gain. For a tractable optimization framework, we pursue two separate approaches to obtain a lower bound and a provably accurate approximation of the offloading gain, which allows us to obtain optimized caching strategies. Remarkably, if we replace the obtained expression for offloading gain with its lower bound, we can find a suboptimal caching strategy that is not only described via analytical formulas but can also show an improvement over the state-of-the-art caching schemes. Results reveal that cooperative transmission becomes more appealing in denser D2D caching networks and adverse interference conditions, which is the case of the imminent internet of things (IoT) and massive machine type communications era.

Index Terms—Coordinated multi-point (CoMP), probabilistic caching, offloading gain, clustered D2D communication.

1 INTRODUCTION

1.1 Background

THE proliferation of advanced mobile devices such as smartphones together with the popularity of video streaming causes tremendous growth of data traffic in cellular networks. To address this challenge, the cellular industry is promoting the deployment of heterogeneous networks that are composed of different types of small base stations (BSs), such as micro-, pico-, and femto-BSs [2]. Different types of small BSs are generally characterized by their transmission powers, coverage areas, and by whom they are deployed. For instance, the typical coverage range of a micro-BS is less than two kilometres. A pico-BS covers 200 meters or less whereas the range of a femto-BS is on the order of 10 meters. Micro- and pico- BSs can be deployed by the network operator while femto-BSs can be deployed by the end users.

The deployment of heterogeneous networks results in a higher density of spatial reuse of radio resources and thus in higher overall network throughput (i.e., the achievable rates). However, deploying a dense heterogeneous network comes with its own challenges. One such challenge is the deployment cost associated with connecting all the small cells to the backbone network with fast links, or the performance degradation accompanied by the capacity-limited backhauls that connect these deployed BSs to the core of the network. Besides, co-existence between small BSs and conventional macro base stations causes additional inter-cell interference when spectrum resources are shared. To address these challenges, caching popular content in advance at the network edge, e.g., at mobile devices or BSs, has been envisioned as a promising technique to relieve the backhaul congestion and improve user quality-of-service (QoS) [3] and [4]. By caching, we mean storing the frequently demanded content at the network edge so as to remove the heavy burden and frequent requests on the (limited) backhauls. Broadly speaking, caching can be performed at the levels, ranging from, mobile devices to the level of small and macro BSs.

In addition to caching, coordinated multipoint (CoMP) transmission has been proposed to mitigate interference and increase network coverage and cell-edge throughput [5]. Cooperative communication through CoMP allows multiple transmission or reception of the same data from multiple BSs/devices in order to improve the spatial diversity.

*The material in this paper is published in part to IEEE WCNC 2019 [1].
 Ramy Amer and Jacek Kibilda and Nicola Marchetti are with CONNECT
 Centre for Future Networks, Trinity College Dublin, Ireland. Email: {ramy,
 kibildj, nicola.marchetti}@tcd.ie.
 Hesham ElSawy is with King Fahd University of Petroleum and Minerals
 (KFUPM), Saudi Arabia. Email: hesham.elsawy@kfupm.edu.sa.
 M. Majid Butt is with Nokia Bell Labs, France, and CONNECT Centre for
 Future Networks, Trinity College Dublin, Ireland. Email: Majid.Butt@tcd.ie.
 This publication has emanated from research conducted with the financial
 support of Science Foundation Ireland (SFI) and is co-funded under the
 European Regional Development Fund under Grant Numbers 13/RC/2077
 and 14/US/I3110.*

This essentially enables seamless co-existence of multiple in-band transmissions and allows efficient spectrum sharing and frequency reuse in such heterogeneous and D2D-enabled networks. In this paper, we are particularly interested in caching on mobile devices [6] and [7], together with cooperative transmission to boost the network traffic offloading and to overcome the performance degradation caused by co-channel interference [8].

Architecture of device caching exploits the large storage available in modern smartphones to cache multimedia files that might be highly demanded by end devices in the network [6]. Devices can then exchange this multimedia content stored on their local storage with nearby devices. Since the distance between a requesting device and the device that stores a requested content is small in most cases, device-to-device (D2D) communication is commonly used for content transmission [9]. As more than one device might cache the same content, the signal-to-interference ratio (SIR) can be improved by joint transmission of the same cached content, which we refer to as cooperative communication, e.g., via CoMP transmission. For a detailed review of the main results and the literature on CoMP transmission, the reader is referred to [10] and [8].

Because of their respective advantages, both caching and cooperative transmission can be jointly adopted in many practical scenarios. For example, ensuring reliable delivery of ultra-high-definition streaming and virtual reality (VR) applications over wireless networks is very challenging due to the stringent QoS requirements [11]. Leveraging D2D CoMP transmissions of pre-downloaded frames from multiple devices to a requesting device, helps reduce communication delays and improve the perceived QoS. Proximity marketing, which is a wireless content advertising system associated with a particular place, might be another use case that leverages both caching and CoMP transmissions [12]. In particular, exploiting CoMP transmissions to send pre-cached advertising content could lead to increasing the transmission range and mitigating interference among different operators of proximity marketing systems. Such cooperatively served contents are then delivered to individuals who wish to receive them, provided that they have the necessary equipment to do so [12]. Motivated by the aforementioned discussions, it is important to study the role of cooperative transmission for cache-enabled D2D networks. The next section is devoted to summarizing relevant works in the literature that adopt cooperative transmission in wireless caching.

1.2 Related Works

The joint adoption of wireless caching and collaborative transmissions, where BSs (or devices) cooperatively serve a content, is widely adopted in literature [13]–[19]. For instance, the authors in [13] proposed a combined caching scheme whereby part of the cache space is reserved for caching the most popular content, that is then cooperatively served from multiple BSs. Moreover, in [14], the authors investigated the trade-off between content diversity gain, i.e., serving diverse content, and cooperative gain, i.e., jointly transmitting the same content. In [15], the authors proposed cooperative transmissions for cache-enabled small cell networks to reduce the backhaul cost and delay. Meanwhile,

the authors in [16] employed content caching at wireless relays to improve the overall performance of collaborative relayings for a network consisting of one source, one destination, and multiple relay nodes.

Following a similar approach, employing cooperative content delivery in D2D caching networks is discussed in [17]–[19]. For instance, the authors in [17] proposed a multiple devices to a single device content delivery method via D2D communication. Moreover, an opportunistic cooperation strategy for D2D transmission is proposed in [18] to mitigate interference among D2D links. Combining coded caching along with CoMP transmissions is recently studied in [19], wherein redundantly stored data at caching helpers is utilized to combat wireless channel impairments due to channel fading and interference.

While interesting, the works in [17]–[19] did not consider the notion of device clustering, which is quite fundamental to the D2D network architecture [20] and [21]. In this regard, the authors in [22] developed a stochastic geometry-based model to characterize the performance of content delivery in a clustered D2D network whose devices are distributed according to a Poisson cluster process (PCP). Similarly, the authors in [23] proposed a cluster-centric content placement scheme for PCPs, where the content of interest is cached closer to the cluster center. A clustered process such as PCP essentially consists of two kinds of point processes, namely, the parent and daughter processes. If the underlying parent point process follows a standard Poisson point process (PPP), the whole process is called PCP [24]. The works in [22] and [23] assumed PCPs with equal number of devices per clusters. Meanwhile, the authors in [25] proposed hybrid caching strategies to reduce the energy cost of D2D transmitters, where the location of these transmitters is modeled as a Gaussian Poisson process (GPP). For the GPP, each cluster has *only* one or two points, with probabilities p and $1 - p$, respectively. If a cluster consists of one point, that point is at the parent's location. If it has two points, one is at the parent's position (cluster center), and the other is uniformly distributed on a circle of certain radius centered at the parent point. However, while the clustering nature of D2D communication is considered in the prior works [22] and [23], these works assumed that contents are pre-cached, i.e., there was no study of the caching problem. Moreover, modeling clustered D2D networks by means of GPPs, as done in [25], is limited by two facts: (i) The distance between transmitting and receiving devices within the same cluster is not captured by this model. (ii) The number of devices per cluster is assumed to be constant, particularly, fixed to only one (or two) device(s) per cluster. Furthermore, the content popularity and caching schemes in [25] were assumed to be the same for all clusters. However, in practice, users in different clusters might have different interests of files. For instance, users in a library might be interested in a different set of content from that of users in a pub.

The offloading gain of a clustered D2D caching network modeled by Thomas cluster process (TCP) is maximized in [26] by joint optimization of content caching and channel access. Moreover, the authors in [27] showed that the average service delay can be efficiently reduced by jointly optimizing content caching and bandwidth partitioning. In [28], the authors proposed an efficient caching scheme for

clustered D2D networks achieving minimum energy consumption. While these works studied the caching problem for clustered D2D networks, they only considered non-cooperative transmissions of requested contents.

Compared with this prior art [13]–[23] and [25]–[28], this paper conducts performance analysis and statistical optimization of cache-assisted cooperative transmissions for a clustered D2D network. In particular, we characterize and optimize the offloading gain of a network of spatially clustered devices that adopt CoMP transmissions and probabilistic caching. By maximizing the statistically-averaged offloading gain, our approach efficiently provides optimal averaged performance on a long-time scale to reduce signaling and processing overheads [29]. Moreover, our model effectively captures the stochastic nature of channel fading and the clustered, yet random, network topology aspects, which have not been studied in the literature, particularly in the context of caching and CoMP transmission. *To the best of our knowledge, this paper provides the first rigorous analysis of cache-assisted CoMP transmissions for D2D caching networks whose devices are modeled by a TCP.*

1.3 Contributions

The main contributions of this paper are summarized as follows:

- We propose a cooperative transmission scheme via D2D communications for clustered cache-enabled networks, whereby a device can be collaboratively served from multiple devices within the same cluster. We analytically characterize the offloading gain and rate coverage probability for the proposed network.
- Given the complexity of the obtained rate coverage probability expression, we propose a tractable lower bound. We use this bound to prove that the interference power seen by the typical device of a TCP can be upper-bounded by the interference power seen by the typical device of a PPP with density that is the product of the TCP cluster density and the average number of devices per cluster.
- To further improve tractability and computational efficiency, we propose to approximate the signal power received from cooperative transmissions by two components: nearest and mean received power terms. Using Chebyshev’s inequality, we prove that this approximation is remarkably tight and helps to reduce the original formulation to single integral.
- We use these closed-form expressions to define two suboptimal caching solutions for the offloading gain maximization problem. Ultimately, we show considerable improvements in the offloading gain under the optimized caching strategies compared with benchmark caching techniques.

The rest of this paper is organized as follows. Section 2 and 3 present the system model and offloading gain characterization, respectively. The rate coverage analysis is conducted in Section 4, and the optimized caching probabilities are obtained in Section 5. Numerical results are then presented in Section 6, and conclusions are drawn in Section 7.

2 SYSTEM MODEL

2.1 Network Model

We consider a D2D caching network in which devices are spatially distributed into disjoint clusters. The devices are assumed to have surplus memory that can be used to store content such as video files. Such a cached content is needed either for future use or to participate in content sharing with other devices within the same cluster. For this network, we model the location of the devices with a TCP composed of parent and daughter points. An TCP is generated by taking a parent homogeneous PPP and daughter Gaussian PPP, one per parent, and translating the daughter processes to the position of their parent [24]. The cluster process is then the union of all the daughter points. Importantly, the cluster centers are virtual points that determine the point around which cluster members are distributed. In other words, a cluster center resembles a geographical reference point for the spatial locations of the D2D devices within the same cluster. The cluster devices, however, are those cluster members (daughter points) that are scattered around the corresponding cluster centers.

Let us denote the parent point process by $\Phi_p = \{\mathbf{x}_1, \mathbf{x}_2, \dots\}$, where $\mathbf{x}_i = \{x_1, x_2\} \in \mathbb{R}^2$, and $i \in \mathcal{N}$. Further, let (Φ_i) be a family of finite point sets representing the untranslated daughter Gaussian PPPs, denoted as Φ_c , i.e., untranslated clusters. The cluster process is then the union of the translated clusters:

$$\Phi \triangleq \cup_{i \in \mathcal{N}} \mathbf{x}_i + \Phi_i. \quad (1)$$

The parent and daughter points are referred, respectively, to as cluster centers and cluster members. We assume that the cluster centers are distributed according to a PPP Φ_p of density is λ_p . We also assume that, for Gaussian PPPs, the cluster members are normally scattered with variance $\sigma^2 \in \mathbb{R}$ around their cluster centers [24]. Given this normal scattering of daughter points, the probability distribution function (PDF) of the cluster member location relative to its cluster center is given by

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{\|\mathbf{y}\|^2}{2\sigma^2}\right), \quad \mathbf{y} \in \mathbb{R}^2, \quad (2)$$

where $\mathbf{y} \in \mathbb{R}^2$ is the device location relative to its cluster center, $\|\cdot\|$ is the Euclidean norm. Within each cluster, the number of cluster members is a Poisson random variable (RV) with a certain mean. For instance, if the average number of devices per cluster is \bar{n} , the cluster intensity will be:

$$\lambda_c(\mathbf{y}) = \frac{\bar{n}}{2\pi\sigma^2} \exp\left(-\frac{\|\mathbf{y}\|^2}{2\sigma^2}\right), \quad \mathbf{y} \in \mathbb{R}^2. \quad (3)$$

Accordingly, the intensity of the entire process Φ will be $\lambda = \bar{n}\lambda_p$.

2.2 Content Popularity and Probabilistic Caching

Similarly to [25], we assume two kind of devices co-exist within the same cluster, namely, *content clients* and *content providers*. In particular, the devices that can perform proactive caching and provide content delivery are called content providers while those requesting content, that also have caching capability, are called content clients. We assume that each device has a surplus memory of size M files

designated for caching content from a known file library \mathcal{F} . The total number of files is $N_f > M$ and the set of content indices is denoted as $\mathcal{F} = \{1, 2, \dots, N_f\}$. These N_f files represent the content catalog that all the devices in a cell may request, which are indexed in a descending order of popularity. We assume that the probability that the m -th content is requested follows the standard Zipf distribution as widely-adopted in the literature, which is given by [30]

$$q_m = \left(m^\beta \sum_{k=1}^{N_f} k^{-\beta} \right)^{-1}, \quad (4)$$

where β is a parameter reflecting how skewed the popularity distribution is. The larger β , the fewer files that are responsible for the majority of requests [30]; by definition, $\sum_{m=1}^{N_f} q_m = 1$. Moreover, we assume that the content popularity may vary across clusters. For instance, users in a library may be interested in an entirely different set of files from the users in a sports center. Therefore, the Zipf distribution models the per-cluster popularity of files. Such a discrepancy of contents of interest can be captured by having different popularity indexes β per different clusters, i.e., different concentration rates. Ranks of popular contents can be also different among different clusters. This discrepancy of popular files implies that the content request and, consequently, caching design models vary across clusters. Such a cluster-specific popularity can be seen as a direct generalization of the individual user preferences that is modeled in [31].

The cluster-specific popularity model necessitates the design of content placement on a per-cluster basis. Hence, within each cluster, we assume a random content placement scheme in which file m is cached independently at each cluster device according to the probability c_m , with $0 \leq c_m \leq 1, \forall m \in \{1, \dots, N_f\}$. To avoid duplicate caching of the same file within the memory of the same device, we follow the probabilistic caching (PC) approach proposed in [32], which requires that $\sum_{m=1}^{N_f} c_m = M$. It is worth mentioning that the PC is a standard caching technique that is widely adopted in the literature [33]–[35].

2.3 Content Request and Delivery Model

Enabling seamless video delivery over cellular networks implies stringent QoS requirements. However, the performance of wireless networks, especially D2D communications, is limited by interference and the effects of small scale fading. Therefore, cooperative communication turns to be more appealing as a prominent interference mitigation tool. We hence allow multiple devices to jointly serve their cached content to a common device within the same cluster via non-coherent CoMP transmission. The underlying reason of assuming a non-coherent transmission is that it is hard to estimate the channel state information (CSI) for the D2D communications. We consider out-of-band D2D communication system, i.e., there is no cross-interference between the cellular network and D2D communication. All devices are equipped with a single transmit-receive isotropic antenna, and they have no CSI from the device they are serving. Furthermore, each D2D transmission uses all the available bandwidth. Transmitted signals experience single-slope path loss with attenuation exponent $\alpha > 2$ and

small scale fading, which we model as an independently and identically distributed (i.i.d.) complex Gaussian RV with zero mean and unit variance. In our study, we are particularly focused on developing a general model that can be then tailored for any physical layer transmission scheme. For instance, if the mobile devices use perfectly-synchronized orthogonal frequency-division multiplexing (OFDM) symbols with sufficiently large cyclic prefix (CP), the inter-symbol interference (ISI) can be effectively cancelled out. Moreover, since we assume static devices, there would be neither doppler effect nor inter-carrier interference (ICI) when adopting OFDM and proper channel selections.

Due to the cost of participating in content caching and delivery, e.g., battery consumption and memory utilization, not all content providers can be active in all time slots. Hence, within each cluster, we assume that content providers can be available for content delivery with probability $p \in [0, 1]$. In other words, among \bar{n} average number of devices per cluster, only average of $p\bar{n}$ devices are willing to participate in content delivery and caching. Further, we assume a BS-assisted D2D link setup scheme, where the transmissions of different files in different clusters are orchestrated by the BS [36]. In details, a *content client* first sends its request to its geographically closest BS, which knows the active content providers within the same cluster, their cached files, as well as their locations. If there are active content providers caching the requested file, the BS then establishes direct CoMP D2D links between the content client and the set of active content providers for this requested file. Requests for contents are assumed to be of negligible size, so that they do not add signalling overhead and are always successfully decoded at the BS.

Within each cluster, we assume a content client device whose distance to its cluster center is drawn from a Rayleigh distribution of scale parameter σ , according to the TCP definition [24]. Throughout time, content clients in different clusters may request files $i \in \mathcal{F}$ with a probability following the assumed *per-cluster Zipf distribution* in Eq. (4). Since each cluster has its own library, a given content client may either be served via D2D connection(s) from active provider(s) within the same cluster or, as a last resort, via the nearest geographical BS. To recap, under the proposed transmission and caching schemes, one content client per cluster is cooperatively served at a time from neighboring active content providers while being interfered only from active providers in other (remote) clusters (i.e., inter-cluster interference).

Notice that, according to the independent thinning theorem [24, Theorem 2.36], active devices within the same cluster form a Gaussian PPP Φ_{cp} whose intensity function is $\lambda_{cp}(\mathbf{y}) = p\lambda_c(\mathbf{y})$. Similarly, the set of active providers that cache a desired content m are modeled as a Gaussian PPP Φ_{cpm} with the intensity function given by $\lambda_{cpm}(\mathbf{y}) = c_m p \lambda_c(\mathbf{y})$. Hence, within each cluster, the number of active devices and the number of active devices caching content m are Poisson RVs of means $p\bar{n}$ and $c_m p\bar{n}$, respectively. By definition, $\Phi_{cpm} \subseteq \Phi_{cp} \subseteq \Phi_c$.

The main advantages of D2D caching networks lie in alleviating the burden of the backhaul links and improving the network spectral efficiency. To leverage these features, it is crucial to intelligently cache and deliver contents to max-

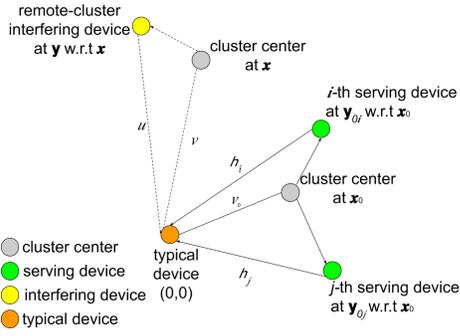


Fig. 1. Illustration of the representative cluster and one interfering cluster, where $\{x_0, y_{0i}, y_{0j}, x, y\} \in \mathbb{R}^2$ and $\{v_0, h_i, h_j, v, u\} \in \mathbb{R}$.

imize the percentage of offloaded traffic from the network core to the edge. The offloading gain is widely-adopted as a key performance metric to quantify this percentage [13], [14], and [25]. Specifically, the offloading gain is defined as the probability of obtaining a desired content either from the self-cache or via D2D communication with a received SIR greater than a target threshold. Hence, our target in the next section is to characterize and maximize the offloading gain of the proposed CoMP-assisted D2D caching network.

3 OFFLOADING GAIN CHARACTERIZATION

Given stationarity of the parent process and independence of the daughter process, we can conduct our analysis for the representative cluster, which is an arbitrary cluster whose center is located at $x_0 \in \Phi_p$, and a *typical client*, which is a randomly selected member of the representative cluster that requests the content. We are particularly interested in the network averaged offloading gain by focusing on the performance of this typical device. The study of the network overall capacity and size is also an interesting problem but is beyond the scope of this paper. Without loss of generality, we assume the typical client is located at the origin $(0, 0) \in \mathbb{R}^2$.

When some active content providers jointly transmit a desired content m , the signal received at the typical client consists of two main components: the desired signal that represents the joint non-coherent transmissions from active providers that cache content m in the representative (local) cluster, and the interference component that is created by other active providers in remote clusters. This can be formally stated as:

$$y_m = \underbrace{\sum_{y_{0i} \in \Phi_{cpm}} \sqrt{\gamma_d} G_{y_{0i}} \|x_0 + y_{0i}\|^{-\alpha/2} s_{y_{0i}}}_{\text{desired signal}} + \underbrace{\sum_{x \in \Phi_p^!} \sum_{y \in \Phi_{cp}} \sqrt{\gamma_d} G_y \|x + y\|^{-\alpha/2} s_y + z}_{\text{inter-cluster interference}},$$

where $i \in \{1, \dots, |\Phi_{cpm}|\}$, $G_{y_{0i}}$ denotes the power fading between an active provider at $y_{0i} \in \Phi_{cpm}$ relative to its cluster center at x_0 and the typical client, see Fig. 1; γ_d denotes the D2D transmission power, and $s_{y_{0i}}$ is the symbol jointly transmitted by the active providers $y_{0i} \in \Phi_{cpm}$.

$\Phi_p^! = \Phi_p \setminus \{x_0\}$ denotes the set of remote clusters, and $\Phi_{cp} \subseteq \Phi_c$ represents the set of active devices in a remote cluster centered at $x \in \Phi_p^!$. Finally, z denotes the standard additive white Gaussian noise.

Note that representing the set of inter-cluster interferers as Φ_{cp} corresponds to the worst case interference scenario, when all active devices in a remote cluster are caching the required content of their own-cluster content client. This bound is in line with our analysis and the underlying network model, particularly, the assumption of different content popularity and placement per clusters. This is because, based on this bound, the inter-cluster interference power, and correspondingly, the per cluster cache design will be independent of the content demand in other clusters.

We focus on the interference-limited regime, thus omitting the thermal noise. This assumption of negligible thermal noise power compared to the interference power is widely-adopted in the literature, especially for clustered D2D networks [22], [23], and [25]. Moreover, accounting for the noise power is a direct extension of the analysis in this paper. Assuming unit power Gaussian symbols, the received SIR at the typical client when downloading content m is given by

$$\text{SIR}_m = \frac{\gamma_d \left| \sum_{y_{0i} \in \Phi_{cpm}} G_{y_{0i}} \|x_0 + y_{0i}\|^{-\alpha/2} \right|^2}{I_{\text{out}}}, \quad (5)$$

where I_{out} is the sum of interfering signal power associated with the downloading of content m , given by:

$$I_{\text{out}} = \gamma_d \left| \sum_{x \in \Phi_p^!} \sum_{y \in \Phi_{cp}} G_y \|x + y\|^{-\alpha/2} \right|^2.$$

Finally, the offloading gain can be formally stated as:

$$\mathbb{P}_o(\mathbf{c}) = \sum_{m=1}^{N_f} q_m c_m + q_m (1 - c_m) \Upsilon_m, \quad (6)$$

where $\mathbf{c} = \{c_1, \dots, c_m, \dots, c_{N_f}\}$, and $\Upsilon_m = \mathbb{P}(\text{SIR}_m > \vartheta)$ is the rate coverage probability for content m , i.e., the probability that the received SIR via CoMP transmissions is larger than a target threshold ϑ , which we characterize in the sequel. In Eq. (6), the first term corresponds to the event of serving a desired content from local memory, i.e., self-cache [37]. The second term represents the joint event that the desired content is not locally cached while being cached and downloadable from active providers in the same cluster, with an SIR greater than the target threshold ϑ .

4 RATE COVERAGE PROBABILITY ANALYSIS

Our objective in this section is to analytically characterize the offloading gain. In particular, we first derive the exact expression of $\mathbb{P}_o(\mathbf{c})$ as a function of the system parameters. Then, we seek lower bound and approximation of the rate coverage probability Υ_m that will result in easy-to-compute expressions of the offloading gain, and provide useful system design insights.

In the case of CoMP transmissions, active providers in the representative cluster jointly transmit the requested content to the typical client. The received power at the typical client is the sum of the received signal powers from

active providers, and hence, the rate coverage probability Υ_m is:

$$\Upsilon_m = \mathbb{P} \left[\frac{\gamma_d \left| \sum_{\mathbf{y}_{0i} \in \Phi_{cpm}} G_{\mathbf{y}_{0i}} \|\mathbf{x}_0 + \mathbf{y}_{0i}\|^{-\alpha/2} \right|^2}{I_{\text{out}}} \geq \vartheta \right]. \quad (7)$$

Since $G_{\mathbf{y}_{0i}}$ are i.i.d. complex Gaussian RVs, we get

$$\left| \sum_{\mathbf{y}_{0i} \in \Phi_{cpm}} \|\mathbf{x}_0 + \mathbf{y}_{0i}\|^{-\alpha/2} G_{\mathbf{y}_{0i}} \right|^2 \sim \exp \left(\frac{1}{\sum_{\mathbf{y}_{0i} \in \Phi_{cpm}} \|\mathbf{x}_0 + \mathbf{y}_{0i}\|^{-\alpha}} \right). \quad (8)$$

Hence, from (7) and (8), we have

$$\begin{aligned} \Upsilon_m &= \mathbb{E} \left[\exp \left(- \frac{\vartheta (I_{\text{out}})}{\gamma_d S_{\Phi_{cpm}}} \right) \right] \\ &\stackrel{(a)}{=} \mathbb{E} \left[\mathcal{L}_{I_{\text{out}}}(t) \Big| S_{\Phi_{cpm}} = s_{\Phi_{cpm}} \right], \end{aligned} \quad (9)$$

where $S_{\Phi_{cpm}} = \sum_{\mathbf{y}_{0i} \in \Phi_{cpm}} \|\mathbf{x}_0 + \mathbf{y}_{0i}\|^{-\alpha}$ is a RV that can be physically interpreted as the received signal power from the active providers devices $\mathbf{y}_{0i} \in \Phi_{cpm}$ subject to path loss only (as we have already averaged over the fading based on the PDF in Eq. (8)), assuming normalized power. (a) follows from the Laplace transform of the interference I_{out} evaluated at $t = \frac{\vartheta}{\gamma_d S_{\Phi_{cpm}}}$. We derive the Laplace transform of interference in the following Lemma to compute the rate coverage probability, and correspondingly, the offloading gain.

Lemma 1. *Laplace transform of the inter-cluster interference, conditioned on a realization of the active providers for content m in the representative cluster, is given by*

$$\mathcal{L}_{I_{\text{out}}}(t) = \exp \left(- 2\pi\lambda_p \int_{v=0}^{\infty} \left(1 - e^{-p\bar{n}\zeta(v,t)} \right) v \, dv \right), \quad (10)$$

where $t = \frac{\vartheta}{\gamma_d S_{\Phi_{cpm}}}$, $\zeta(v,t) = \int_{u=0}^{\infty} \frac{t\gamma_d}{u^\alpha + t\gamma_d} f_{U|V}(u|v) \, du$, $f_{U|V}(u|v) = \text{Rice}(u; v, \sigma)$ is the Rician PDF modeling the distance $U = \|\mathbf{x} + \mathbf{y}\|$ between an interfering device at \mathbf{y} relative to its cluster center at $\mathbf{x} \in \Phi_p$ and the origin $(0, 0)$, conditioned on $V = \|\mathbf{x}\| = v$.

Proof: Please refer to Appendix A. \square

We continue by characterizing the joint serving distance distribution. For a given realization $S_{\Phi_{cpm}} = s_{\Phi_{cpm}}$, let us assume that there are k active providers in the representative cluster. Let us also denote joint distances from the typical client (origin) to the k content providers in the representative cluster, centered at \mathbf{x}_0 , as $\mathbf{H}_k = \{H_1, \dots, H_k\}$. Then, conditioning on $\mathbf{H}_k = \mathbf{h}_k$, where $\mathbf{h}_k = \{h_1, \dots, h_k\}$, the conditional (i.e., on k) PDF of the joint serving distance is denoted as $f_{\mathbf{H}_k}(\mathbf{h}_k)$. Hence, conditioning on k , we can express the rate coverage probability as

$$\Upsilon_{m|k} = \mathbb{E} \left[\mathcal{L}_{I_{\text{out}}}\left(t = \frac{\vartheta}{\gamma_d \sum_{i=1}^k h_i^{-\alpha}}\right) \Big| S_{\Phi_{cpm}} = s_{\Phi_{cpm}} \right], \quad (11)$$

where $s_{\Phi_{cpm}} = \sum_{i=1}^k h_i^{-\alpha}$, $h_i = \|\mathbf{x}_0 + \mathbf{y}_{0i}\|$, and $\mathbf{y}_{0i} \in \Phi_{cpm}$. Since a content provider i in the representative cluster centered at \mathbf{x}_0 has its coordinates in \mathbb{R}^2 chosen independently from a Gaussian distribution with standard deviation σ , then, by definition, the distance from such a content provider to the origin, denoted as $h_i = \|\mathbf{x}_0 + \mathbf{y}_{0i}\|$, has Rician distribution $f_{H_i|V_0}(h_i|v_0) = \text{Rice}(h_i; v_0, \sigma)$. Since also the content providers and the typical client have their locations sampled from a normal distribution with variance σ^2 relative to their cluster center \mathbf{x}_0 , then, by definition, the statistical distance distribution between any two points, e.g., from the i -th content provider to the typical client, follows Rayleigh distribution with scale parameter $\sqrt{2}\sigma$, written as

$$f_{H_i}(h_i) = \text{Rayleigh}(h_i, \sqrt{2}\sigma) = \frac{h_i}{2\sigma^2} e^{-\frac{h_i^2}{4\sigma^2}}. \quad (12)$$

If the serving distances from the typical client to the different points of the cluster were independent from each other, $f_{\mathbf{H}_k}(\mathbf{h}_k)$ would simply be the product of k independent PDFs, each of which having $f_{H_i}(h_i) = \text{Rayleigh}(h_i, \sqrt{2}\sigma)$. However, there is a correlation between the serving distances due to the common factor \mathbf{x}_0 in the serving distance equation $h_i = \|\mathbf{x}_0 + \mathbf{y}_{0i}\|$ with $\mathbf{y}_{0i} \in \Phi_{cpm}$, see also Fig. 1. For analytical tractability, similar to [22] and [27], we neglect this correlation. Hence, the conditional PDF of the joint serving distance $f_{\mathbf{H}_k}(\mathbf{h}_k)$ can be obtained from

$$f_{\mathbf{H}_k}(\mathbf{h}_k) = \prod_{i=1}^k \frac{h_i}{2\sigma^2} e^{-\frac{h_i^2}{4\sigma^2}}. \quad (13)$$

Such a negligible correlation assumption yields an upper bound on the exact rate coverage probability of the typical device. Moreover, if the typical device is assumed to be at the cluster center, there would be no correlation between the underlying serving distances, e.g., see [38] and [39], and (13) will represent the exact joint serving distance PDF. Conditioning on having k active content providers, i.e., $s_{\Phi_{cpm}} = \sum_{i=1}^k h_i^{-\alpha}$, the rate coverage probability will be given by $\Upsilon_{m|k} =$

$$\int_{\mathbf{h}_k=0}^{\infty} \mathcal{L}_{I_{\text{out}}}\left(\frac{\vartheta}{\gamma_d \sum_{i=1}^k h_i^{-\alpha}} \Big| S_{\Phi_{cpm}} = s_{\Phi_{cpm}}\right) f_{\mathbf{H}_k}(\mathbf{h}_k) \, d\mathbf{h}_k. \quad (14)$$

Given that Φ_{cpm} is a Gaussian PPP, the number of active content providers for content m is a Poisson RV with mean $c_m p \bar{n}$. Therefore, the probability that there are k content providers is equal to $\frac{(p \bar{n} c_m)^k e^{-p \bar{n} c_m}}{k!}$. Invoking this along with (10), (13), and (14) into (6), $\mathbb{P}_o(\mathbf{c})$ is given as

$$\begin{aligned} \mathbb{P}_o(\mathbf{c}) &= \sum_{m=1}^{N_f} q_m \left(c_m + (1 - c_m) \cdot \sum_{k=1}^{\infty} \frac{(p \bar{n} c_m)^k e^{-c_m p \bar{n}}}{k!} \right. \\ &\quad \left. \int_{\mathbf{h}_k=0}^{\infty} e^{-2\pi\lambda_p \int_{v=0}^{\infty} \left(1 - e^{-p \bar{n} (1 - \zeta(v,t))} \right) v \, dv} \prod_{i=1}^k \frac{h_i}{2\sigma^2} e^{-\frac{h_i^2}{4\sigma^2}} \, d\mathbf{h}_k \right). \end{aligned} \quad (15)$$

Since the obtained expression in Eq. (15) involves multi-fold integrals and summations, this renders the calculation of the rate coverage probability computationally complex. Furthermore, the offloading gain maximization problem turns to be intractable. Therefore, in the sequel, we focus

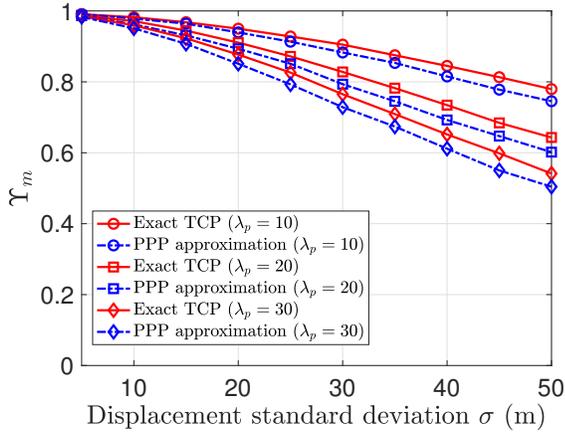


Fig. 2. The lower bound on Υ_m based on (16) versus displacement standard deviation σ is plotted for various parent PPP densities λ_p ($\bar{n} = 20$, $p = 0.5$, $c_m = 0.5$). "Exact TCP" in the legend refers to the exact performance for the TCP while "PPP approximation" refers to the lower bound based on Theorem 1.

on tight bound and approximation of the rate coverage probability that will result in easy-to-compute expressions that also enable us to formulate a tractable optimization problem to maximize the offloading gain.

4.1 Lower Bound on Offloading Gain

Next, we obtain a tractable lower bound on the offloading gain based on an upper bound on the interference power. By replacing the exponential interference formula in the Laplace transform expression of inter-cluster interference in (10) by the first and second terms of its Taylor's series expansion, we obtain an upper bound on the interference power. Accordingly, this bound on the Laplace transform of inter-cluster interference yields a lower bound on the rate coverage probability Υ_i .

Theorem 1. *Laplace transform of interference derived in (10) can be bounded by*

$$\mathcal{L}_{I_{\text{out}}}(t) \approx \exp\left(-\pi p \bar{n} \lambda_p t^{2/\alpha} \Gamma(1+2/\alpha) \Gamma(1-2/\alpha)\right), \quad (16)$$

and, correspondingly, a lower bound on the offloading gain is given by

$$\mathbb{P}_\sigma^\sim(c) = \sum_{m=1}^{N_f} q_m \left(c_m + (1-c_m) \cdot \sum_{k=1}^{\infty} \frac{(c_m p \bar{n})^k e^{-c_m p \bar{n}}}{k!} \times \int_0^\infty e^{-\pi p \bar{n} \lambda_p \left(\frac{\vartheta}{\sum_{i=1}^k h_i^{-\alpha}}\right)^{2/\alpha} \Gamma(1+2/\alpha) \Gamma(1-2/\alpha)} \prod_{i=1}^k \frac{h_i}{2\sigma^2} e^{-\frac{h_i^2}{4\sigma^2}} dh_k \right) \quad (17)$$

Proof: Please refer to Appendix B. \square

Remark 1. The obtained expression in Eq. (16) boils down to the Laplace transform of a PPP with intensity $\bar{n}\lambda_p$. This shows that the inter-cluster interference of a TCP with density of clusters λ_p and average number of devices per cluster \bar{n} , i.e., with intensity $\bar{n}\lambda_p$, is upper bounded by that of a PPP of the same intensity.

In Fig. 2, we plot the exact expression and its lower bound, based on (16), of the rate coverage probability versus displacement variance σ for various parent PPP Φ_p densities λ_p . The derived lower bound is considerably tight when

both σ and λ_p are relatively small. Also, it is noticeable that Υ_m monotonically decreases with both σ and λ_p , which reflects the fact that the desired signal is weaker when the distance between content providers and the typical client is larger, and the effect of inter-cluster interference increases when the density of clusters increases, respectively. When λ_p and σ increase, the obtained lower bound becomes no longer tight, however, it still represents a reasonable bound on the exact Υ_m .

Having obtained a lower bound on the offloading gain, next, we seek further approximation by replacing the desired signal power with the sum of the signal power from the nearest device signal and the mean power received from all other active content providing devices.

4.2 Serving Power Approximation

From (8), $S_{\Phi_{cpm}} = \sum_{\mathbf{y}_{0i} \in \Phi_{cpm}} \|\mathbf{x}_0 + \mathbf{y}_{0i}\|^{-\alpha}$ represents a RV that models the intended signal power from content providers $\mathbf{y}_{0i} \in \Phi_{cpm}$ subject to path loss only. Next, we adopt the so-called *mean plus nearest approximation* to approximate this intended signal power $S_{\Phi_{cpm}}$ as a sum of two terms. More specifically, the first term $\|\mathbf{x}_0 + \mathbf{y}_{01}\|^{-\alpha}$ is the received signal power from the nearest active provider, where $\mathbf{y}_{01} = \text{argmin}_{\mathbf{y}_{0i} \in \Phi_{cpm}} \{\|\mathbf{x}_0 + \mathbf{y}_{0i}\|\}$. In addition, the second term is the statistical average of the RV $S_{\Phi_{cpm}^!}$, where $S_{\Phi_{cpm}^!} = \sum_{\mathbf{y}_{0i} \in \Phi_{cpm} \setminus \mathbf{y}_{01}} \|\mathbf{x}_0 + \mathbf{y}_{0i}\|^{-\alpha}$ is the overall signal power received from all other active providers conditioning on the nearest serving distance $H_1 = h_1 = \|\mathbf{x}_0 + \mathbf{y}_{01}\|$. The main motivation behind this approximation is that if the RV $S_{\Phi_{cpm}^!}$ is concentrated around its mean, it can be effectively approximated by this mean $\mathbb{E}[S_{\Phi_{cpm}^!} | h_1]$. The tightness of this approximation is investigated in the sequel.

As we will see, this approximation yields an easy way to obtain the rate coverage probability while also being tight. This approach has been similarly adopted to circumvent intractable analysis in the stochastic geometry literature, see, e.g., [40]. Starting from the Laplace transform expression in Eq. (9), we approximate $S_{\Phi_{cpm}}$ as

$$S_{\Phi_{cpm}} \approx \|\mathbf{x}_0 + \mathbf{y}_{01}\|^{-\alpha} + \mathbb{E}[S_{\Phi_{cpm}^!} | \mathbf{y}_{01}], \quad (18)$$

where $\Phi_{cpm}^! = \Phi_{cpm} \setminus \mathbf{y}_{01}$, and $S_{\Phi_{cpm}^!} = \sum_{\mathbf{y}_{0i} \in \Phi_{cpm}^!} \|\mathbf{x}_0 + \mathbf{y}_{0i}\|^{-\alpha}$. Next, we derive the distribution of nearest serving distance $h_1 = \|\mathbf{x}_0 + \mathbf{y}_{01}\|$. Then, we prove the concentration of the proposed approximation using Chebyshev's inequality. Finally, given the distance distribution to the nearest device $f_{H_1}(h_1)$, and the derived formula for $\mathbb{E}[S_{\Phi_{cpm}^!} | H_1 = h_1]$, an approximation for $S_{\Phi_{cpm}}$ is obtained based on (18).

Lemma 2. *The PDF of the distance from the typical client to the nearest active provider in Φ_{cm} is given by*

$$f_{H_1}(h_1) = c_m p \bar{n} \int_{v_0=0}^{\infty} f_{V_0}(v_0) f_{H_1|V_0}(h_1|v_0) \times e^{-c_m p \bar{n} \int_0^{h_1} f_{H|V_0}(h|v_0) dh} dv_0, \quad (19)$$

and using Jensen's inequality, it can be approximated by

$$f_{H_1}(h_1) \approx \frac{c_m p \bar{n} h_1 \exp\left(-c_m p \bar{n} \left(1 - \exp\left(\frac{-h_1^2}{4\sigma^2}\right)\right) - \frac{h_1^2}{4\sigma^2}\right)}{2\sigma^2}. \quad (20)$$

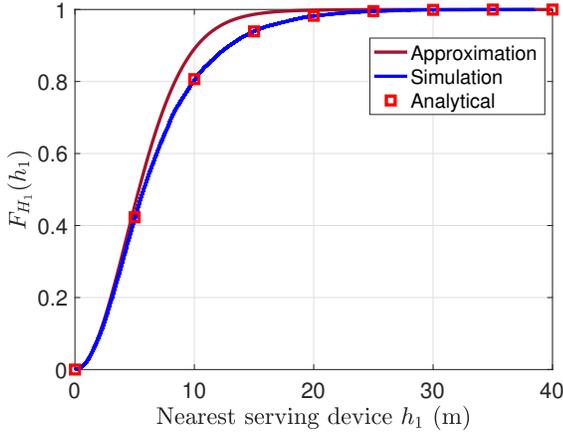


Fig. 3. The derived nearest serving distance CDF in Eq. (39) is plotted and compared with simulation and Jensen's inequality-based approximation in Eq. (40) ($\bar{n} = 20$, $\sigma = 10$ m, $c_m = 0.5$, $p = 1$).

Proof: The proof is provided in Appendix C. \square

The accuracy of the derived CDF $F_{H_1}(h_1)$ in Eq. (39) and its approximation based on Jensen's inequality in Eq. (40) (see Appendix C) are verified in Fig. 3. It is clear from (20) that the distance to the nearest active content provider statistically decreases as c_m or p increase, i.e., when there is a high probability of having active and caching content providers within the local cluster. The distance is also more likely to decrease as \bar{n} increases since a congested cluster has shorter distance between the content client and providers.

Next, we show that approximating the desired signal by its nearest and conditional mean components, see (18), yields an accurate yet tractable expressions for the rate coverage probability and offloading gain.

Proposition 1. *For scenarios of practical interest, the proposed approximation for $S_{\Phi_{cpm}^!}$ in Eq. (18) is a tractable yet remarkably tight bound, and hence, it introduces a reasonable approximation for the rate coverage probability and offloading gain.*

Proof: The proof of the proposition relies on calculating the concentration bounds for $S_{\Phi_{cpm}^!}$. In other words, we will show that the RV $S_{\Phi_{cpm}^!}$ concentrates around its mean. For that purpose, we use Chebyshev's inequality that can be formulated as

$$\mathbb{P}\left(\left|S_{\Phi_{cpm}^!} - \mathbb{E}\left[S_{\Phi_{cpm}^!}\right]\right| > a\right) \leq \frac{\text{Var}\left[S_{\Phi_{cpm}^!}\right]}{a^2}, \quad (21)$$

for $a > 0$, where $\text{Var}\left[S_{\Phi_{cpm}^!}\right]$ is the variance of $S_{\Phi_{cpm}^!}$. We start by calculating the conditional variance $\text{Var}\left[S_{\Phi_{cpm}^!}|h_1\right]$ and mean $\mathbb{E}\left[S_{\Phi_{cpm}^!}|h_1\right]$ in the next two Lemmas.

Lemma 3. *The variance of the signal power received from all active providers except for the nearest device, subject to path loss only, and conditioned on the distance to the nearest active provider $H_1 = h_1$, is expressed as*

$$\text{Var}\left[S_{\Phi_{cpm}^!}|H_1 = h_1\right] = c_m p \bar{n} \Gamma\left(-2\alpha + 1, \frac{h_1^2}{4\sigma^2}\right), \quad (22)$$

Proof: The proof can be found in Appendix D. \square

Lemma 4. *The average over the signal power received from all active content providers except for the nearest one, subject to path*

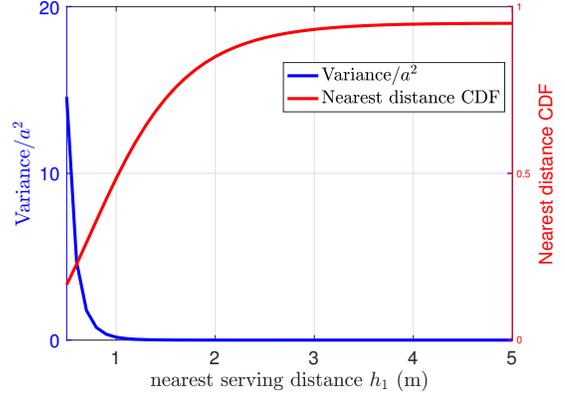


Fig. 4. Nearest serving distance CDF $F_{H_1}(h_1)$ (right side y-axis) and $\text{Var}\left[S_{\Phi_{cpm}^!}\right]/a^2$ (left side y-axis) are plotted versus the nearest serving distance h_1 ($\sigma = 1$ m, $a = 1$, $c_m = 0.6$, $p = 0.5$, $\bar{n} = 10$).

loss only, and conditioned on the distance to the nearest active provider $H_1 = h_1$, is expressed as

$$\mathbb{E}\left[S_{\Phi_{cpm}^!}|H_1 = h_1\right] = \frac{c_m p \bar{n}}{2\sigma^2} \left[\frac{\exp\left(-\frac{h_1^2}{4\sigma^2}\right)}{2h_1^2} - \frac{\Gamma\left(0, \frac{h_1^2}{4\sigma^2}\right)}{8\sigma^2} \right]. \quad (23)$$

Proof: We can write the conditional mean as

$$\begin{aligned} \mathbb{E}\left[S_{\Phi_{cpm}^!}|H_1 = h_1\right] &= \mathbb{E}\left[\sum_{\mathbf{y}_{0i} \in \Phi_{cpm}^!} \|\mathbf{x}_0 + \mathbf{y}_{0i}\|^{-\alpha}\right] \\ &\stackrel{(a)}{=} c_m p \bar{n} \int_{\mathbb{R}^2} \frac{1}{\|\mathbf{x}_0 + \mathbf{y}_{0i}\|^\alpha} f_{\mathbf{Y}_{0i}}(\mathbf{y}_{0i}) d\mathbf{y}_{0i}. \end{aligned} \quad (24)$$

where (a) follows from the mean and variance for PPPs [24, Corollary 4.8], along with the Gaussian PPP assumption $\Phi_{cpm}^!$. Following the same methodology as in Appendix D, the conditional mean can be directly obtained. Hence, Lemma 4 is proven. \square

As an illustrative example, it is reasonable to assume a limited cache-size per device, which triggers $c_m < 1$, mean number of devices per cluster \bar{n} from 5 to 10 devices, and small displacement standard deviation σ from 1 m to 10 m. In such setup, we can observe the tightness of our approximation in Fig. 4. Particularly, we plot the term $\text{Var}\left[S_{\Phi_{cpm}^!}\right]/a^2$, measuring how much $S_{\Phi_{cpm}^!}$ deviates from its mean, along with the CDF of nearest serving distance $F_{H_1}(h_1)$ versus the nearest distance h_1 . From the figure, we first note that $\text{Var}\left[S_{\Phi_{cpm}^!}\right]/a^2$ is almost zero when the nearest active provider is farther than $\sigma = 1$ m, which happens with high probability (from the CDF $F_{H_1}(h_1)$). Moreover, $\text{Var}\left[S_{\Phi_{cpm}^!}\right]/a^2$ is larger than zero when the distance to the nearest active provider is shorter than σ , which happens with small probability (from the CDF $F_{H_1}(h_1)$). This shows that the "mean plus nearest" approximation can yield a remarkably tight bound on $S_{\Phi_{cpm}^!}$, and correspondingly on $\mathbb{P}_o(c)$ for scenarios of practical interest. Hence, the proof of Proposition 1 is completed. \square

Piecing everything together, we get a tight approximation on $\mathbb{P}_o(c)$ as follows. We start with (9) with the

substitution $s_{\Phi_{cpm}} \approx h_1^{-\alpha} + \mathbb{E}[S_{\Phi_{cpm}} | H_1 = h_1]$, where $\mathbb{E}[S_{\Phi_{cpm}} | H_1 = h_1]$ is derived in Eq. (23). Then, we proceed by calculating Laplace transform of inter-cluster interference before averaging over the nearest distance h_1 using the nearest distance PDF $f_{H_1}(h_1)$ in Eq. (19). The approximated offloading gain is formally characterized in the next corollary. Recall that this approximation is based on replacing the intended signal power subject to path loss only by two components, namely, the received signal power from the nearest active provider and the statistical mean over the received power from other content providers.

Corollary 1. *A tight approximation of the offloading gain can be calculated from*

$$\begin{aligned} \mathbb{P}_o^{\approx}(c) &= \sum_{m=1}^{N_f} q_m \left(c_m + (1 - c_m) \times \right. \\ &\quad \left. \int_{h_1=0}^{\infty} e^{-2\pi\lambda_p \int_{v=0}^{\infty} (1 - e^{-p\bar{n}\zeta(v,t)}) v dv} f_{H_1}(h_1) dh_1 \right) \quad (25) \\ &\stackrel{(a)}{\approx} \sum_{m=1}^{N_f} q_m \left(c_m + (1 - c_m) p\bar{n} c_m \int_{h_1=0}^{\infty} \frac{h_1}{2\sigma^2} \times \right. \\ &\quad \left. e^{-2\pi\lambda_p \int_{v=0}^{\infty} (1 - e^{-p\bar{n}\zeta(v,t)}) v dv} \times \right. \\ &\quad \left. e^{-c_m p\bar{n} \left(1 - e^{-\frac{h_1^2}{4\sigma^2}} \right) - \frac{h_1^2}{4\sigma^2}} dh_1 \right), \quad (26) \end{aligned}$$

where

$$t = \frac{\vartheta/\gamma_d}{h_1^{-\alpha} + \frac{c_m p\bar{n}}{2\sigma^2} \left[e^{-\frac{h_1^2}{4\sigma^2}} - \frac{\Gamma(0, \frac{h_1^2}{4\sigma^2})}{8\sigma^2} \right]}. \quad (27)$$

Proof. The above result follows from the nearest plus mean approximation in Eq. (18), along with the conditional mean expression obtained in Eq. (23); (a) follows from the approximated nearest serving distance PDF obtained in Eq. (20). \square

Note that the exponential term inside the integral of (25) is a function of h_1 since $t = \frac{\vartheta}{\gamma_d s_{\Phi_{cpm}}}$. It is worth mentioning that with such an approximation, replacing $s_{\Phi_{cpm}}$ with its nearest plus conditional mean approximation converts the multi-fold integral over \mathbf{h}_k in Eq. (15) to a single integral over h_1 . Furthermore, the effect of having a random number of active providers is implicitly involved in the nearest distance PDF as well as in the conditional mean term. This explains why the condition of having k active content providers in the representative cluster no longer exists in the approximated rate coverage probability expression.

In Fig. 5, we plot the obtained formulas for the rate coverage probability Υ_m in Eq. (25) and Eq. (26) versus the displacement variance σ for various density of clusters λ_p . The derived bound is remarkably tight for the practical values of λ_p and σ . It starts to slightly diverge only for a system with considerably large λ_p and σ . It is also clear that the approximated expression of the nearest distance PDF in Eq. (20) bounds that in Eq. (19) very tightly.

It is worth mentioning that the importance of the obtained approximation $\mathbb{P}_o^{\approx}(c)$ in Eq. (25) is two-fold. Firstly, it provides an easy-to-compute approximation for the rate coverage probability, as shown in Fig. 5, and correspondingly, the offloading gain. Secondly, it allows us to solve

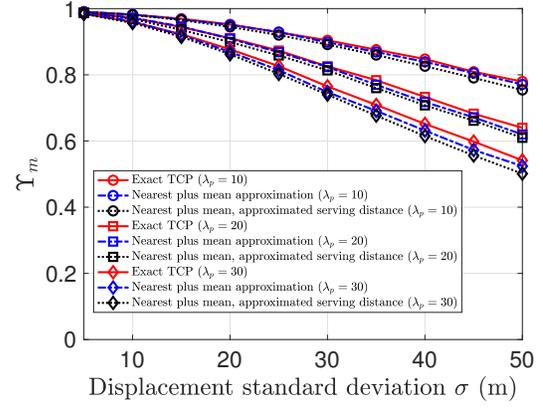


Fig. 5. The derived approximations of Υ_m in Eq. (25) and (26) are plotted versus the displacement standard deviation σ for various parent PPP density λ_p ($\bar{n} = 20$, $p = 0.5$, $c_m = 0.5$). "Nearest plus mean approximation" in the legend refers to the performance based on the exact nearest serving distance PDF in Eq. (19).

for an optimized caching probability, by maximizing the approximated offloading gain, as will be clear shortly. In this regard, it is quite important to examine the achievable performance when a requested content is downloaded only from the nearest active provider. We refer to this as *nearest content provider (NCP) transmission scheme*, where its corresponding rate coverage probability is characterized in the next corollary.

Corollary 2. *The rate coverage probability for the NCP transmission scheme is expressed as*

$$\begin{aligned} \Upsilon_m &= \int_0^{\infty} e^{-2\pi\lambda_p \int_{v=0}^{\infty} (1 - e^{-\bar{n}\zeta(v,t)}) v dv} f_{H_1}(h_1) dh_1 \\ &\approx \frac{c_m p\bar{n}}{2\sigma^2} \int_0^{\infty} h_1 e^{-2\pi\lambda_p \int_{v=0}^{\infty} (1 - e^{-\bar{n}\zeta(v,t)}) v dv} \times \\ &\quad e^{-c_m p\bar{n} \left(1 - e^{-\frac{h_1^2}{4\sigma^2}} \right) - \frac{h_1^2}{4\sigma^2}} dh_1, \end{aligned}$$

where $t = \frac{\vartheta}{\gamma_d h_1^{-\alpha}}$, and $\zeta(v, t)$ is defined in Lemma 1.

Having obtained a lower bound and approximation of the rate coverage probability, next, we employ the obtained results to compute optimized caching probabilities c in order to maximize the offloading gain.

5 OPTIMIZED CACHING PROBABILITIES

We first formulate and solve the offloading gain maximization problem based on the mean plus nearest approximation in Eq. (26). Then we pursue another approach, based on the derived lower bound in Eq. (17), to obtain a low complexity solution.

5.1 Optimized Caching Based on the approximation

Here, we aim at maximizing the approximated offloading gain obtained in Corollary 1. We formulate the offloading gain maximization problem as

$$\mathbf{P1:} \quad \max_c \quad \mathbb{P}_o^{\approx}(c) \quad (28)$$

$$\text{s.t.} \quad \sum_{n=1}^{N_f} c_m = M, \quad c_m \in [0, 1] \quad (29)$$

Since the integral in the approximated offloading gain expression in Eq. (25) depends on the caching probability c_m ,

and c_m exists as a complex exponential term in the nearest serving distance PDF $f_{H_1}(h_1)$, it is hard to analytically characterize (e.g., show concavity of) the objective function or find a tractable expression for the caching probability c_m . In order to tackle this, similar to [41], we introduce a \mathbf{c} -independent integral by substituting the caching probability with an arbitrary caching probability \mathbf{c}^0 . Denoting this \mathbf{c} -independent integral by $I^{\mathbf{c}^0}$, it can be easily verified that

$$\mathbb{P}_o^{\approx \mathbf{c}^0}(\mathbf{c}) = \sum_{m=1}^{N_f} q_m \left(c_m + (1 - c_m) p \bar{n} c_m I^{\mathbf{c}^0} \right), \quad (30)$$

is concave in \mathbf{c} . From the Karush-Kuhn-Tucker (KKT) conditions, the optimized caching probability maximizing $\mathbb{P}_o^{\approx \mathbf{c}^0}(\mathbf{c})$ under the constraint (29) is given by

$$c_m^* = \left[0.5 - \frac{v^* - q_m}{2q_m p \bar{n} I^{\mathbf{c}^0}} \right]^+, \quad (31)$$

where v^* satisfies the maximum cache constraint $\sum_{i=1}^{N_f} \left[0.5 - \frac{v^* - q_m}{2q_m p \bar{n} I^{\mathbf{c}^0}} \right]^+ = M$, and $[x]^+ = \max(x, 0)$. For the arbitrary caching probability \mathbf{c}^0 , a locally optimal caching probability can be adopted, which can be computed via, e.g., interior point method [42]. Nonetheless, we can increase the probability of finding the optimal solution of **P1** by using the interior point method with multiple random initial values and then picking the solution with the highest offloading gain. We refer to this caching probability in Eq. (31) as the solution from convex approximation. However, to obtain a caching policy of lower complexity, we maximize a special case of the lower bound $\mathbb{P}_o^{\sim}(\mathbf{c})$ in the sequel.

5.2 Optimized Caching Based on the Lower Bound

Although $\mathbb{P}_o^{\sim}(\mathbf{c})$ characterized in Theorem 1 is simpler to compute compared to $\mathbb{P}_o(\mathbf{c})$, it is still challenging to obtain the optimal caching probability due to the summation and multi-fold integration in Eq. (17). Similar to [14], we consider a special case when downloading content from one active content provider, for which the offloading gain maximization problem turns out to be convex.

One Content Provider: Next, we solve for the optimized caching probability when considering one serving content provider (instead of k in Eq. (17)). Starting from (17) with $t = \frac{\vartheta h^\alpha}{\gamma_d}$ for one serving provider, and the void probability of a Poisson RV, we get

$$\Upsilon_m = \left(1 - \frac{(p \bar{n} c_m)^0}{0!} e^{-c_m p \bar{n}} \right) \times \int_{h=0}^{\infty} e^{-\pi p \bar{n} \lambda_p (\vartheta h^\alpha)^{2/\alpha} \Gamma(1+2/\alpha) \Gamma(1-2/\alpha)} \frac{h}{2\sigma^2} e^{-\frac{h^2}{4\sigma^2}} dh. \quad (32)$$

Solving the integral in Eq. (32), and substituting in Eq. (17), we get $\mathbb{P}_o^{\sim 1}(\mathbf{c})$ written as

$$\mathbb{P}_o^{\sim 1}(\mathbf{c}) = \sum_{m=1}^{N_f} q_m \left(c_m + (1 - c_m) (1 - e^{-c_m p \bar{n}}) \frac{1}{\mathcal{Z}(\vartheta, \alpha, \sigma)} \right), \quad (33)$$

TABLE 1
Simulation Parameters

Description	Parameter	Value
Path loss exponent	α	4
SIR threshold	ϑ	0 dB
Density of clusters	λ_p	30 km ⁻²
Displacement standard deviation	σ	30 m
Average number of devices per cluster	\bar{n}	6
Library size	N_f	12
Device cache size	M	2
Access probability	p	0.5

where $\mathcal{Z}(\vartheta, \alpha, \sigma) = 4\sigma^2 \pi p \bar{n} \lambda_p \vartheta^{2/\alpha} \Gamma(1+2/\alpha) \Gamma(1-2/\alpha) + 1$. Hence, the optimized caching probability can be computed by solving the following problem

$$\mathbf{P2:} \quad \max_{\mathbf{c}} \quad \mathbb{P}_o^{\sim 1}(\mathbf{c}) \quad (34)$$

$$\text{s.t.} \quad \sum_{n=1}^{N_f} c_m = M, \quad c_m \in [0, 1] \quad (35)$$

The optimal solution for **P2** is formulated in the following Lemma.

Lemma 5. *The lower bound on the offloading gain $\mathbb{P}_o^{\sim 1}(\mathbf{c})$ in Eq. (33) is concave w.r.t. the caching probability, and the optimal probabilistic caching $\underline{\mathbf{c}}^*$ for **P2** is given by*

$$\underline{c}_m^* = \begin{cases} 1 & , v^* < q_m - \frac{q_m(1-e^{-p\bar{n}})}{\mathcal{Z}} \\ 0 & , v^* > q_m + \frac{p\bar{n}q_m}{\mathcal{Z}} \\ \psi(v^*) & , \text{otherwise,} \end{cases}$$

where $\psi(v^*)$ is the solution of

$$v^* = q_m + \frac{q_m}{\mathcal{Z}} (p\bar{n}(1 - \underline{c}_m^*) e^{-p\bar{n}\underline{c}_m^*} - (1 - e^{-p\bar{n}\underline{c}_m^*})),$$

that satisfies $\sum_{m=1}^{N_f} \underline{c}_m^* = M$, and $\mathcal{Z} = \mathcal{Z}(\vartheta, \alpha, \sigma)$ for ease of presentation.

Proof. It is easy to show the concavity of the objective function $\mathbb{P}_o^{\sim 1}(\mathbf{c})$ by confirming that Hessian matrix w.r.t. the caching variables is negative semi-definite. Also, the constraints are linear, which imply that the necessity and sufficiency conditions for optimality exist. The detailed proof of finding $\underline{\mathbf{c}}^*$ is omitted for brevity. \square

Noticeably, the offloading gain for this special case, specifically when considering one content provider, resembles a strict lower bound on the exact offloading gain [14]. However, given the complexity of the original optimization problem, we sought a suboptimal caching solution by optimizing a convex lower bound. Similar approaches are used in the literature, especially when dealing with complex optimization problems like ours, see, e.g., [14], [41], and [43]. When substituting this suboptimal solution in in Eq. (15), it provides useful insights into the system design and also attains considerable performance improvements over traditional caching schemes, as quantified in Section 6.

6 NUMERICAL RESULTS

In this section, we evaluate the performance of our proposed cache-assisted CoMP transmissions for clustered D2D networks. Unless otherwise specified, results are obtained for the parameters shown in Table 1, which represent typical

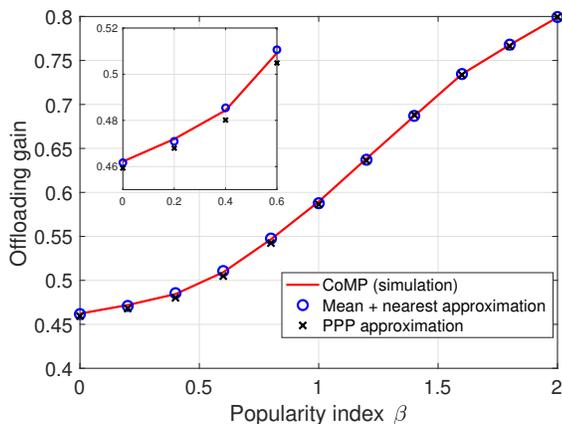


Fig. 6. The exact offloading gain (simulation) based on CoMP transmissions is compared to PPP-based lower bound ($\mathbb{P}_o^{\sim}(c)$), and mean plus nearest-based approximation ($\mathbb{P}_o^{\approx}(c)$), versus the popularity of files β under the Zipf caching scheme.

values used in many previous works.¹ We refer to both solutions based on convex approximation in Eq. (31) and the suboptimal caching of Lemma 5 as optimized PC. We first verify the accuracy of the derived bound and approximation of the offloading gain. Then, we compare the achievable performance of our proposed PC and CoMP transmission, with conventional caching and transmission schemes.

6.1 Exact Offloading Gain Versus Approximation and Lower Bound

Fig. 6 verifies the accuracy of the obtained lower bound and approximation of the offloading gain in Eq. (17) and Eq. (25), respectively. It is clear that both the derived lower bound and approximation are tight to the exact offloading gain. Moreover, we observe that averaging over the request and caching probabilities, i.e., q_m and c_m , makes the adopted lower bound and approximation of the offloading gain tighter than those for the rate coverage probability Υ_m shown earlier in Fig. 2 and Fig. 5, respectively. As shown in Fig. 6, the offloading gain monotonically increases with the popularity of files β . This is because when β is large, only a small portion of content undergoes most of the demand, which can be cached among the cluster devices.

6.2 Comparison with Other Caching Schemes

Fig. 7 compares the offloading gain of the proposed PC with other benchmark schemes, namely, Zipf caching (Zipf), caching popular files (CPF), and random caching (RC) against the popularity of files β . For CPF, the M -most popular files are cached among each cluster device, see, e.g. [6]. Similarly, for RC, contents to be cached are randomly chosen according to a uniform distribution irrespective of the popularity as in [6], while for Zipf caching, contents

1. In Table 1, we assume relatively small device cache and content library sizes. These values, which are very close to those used in [14], [25], [27], [41], and [43], are reasonable in the study of communication and caching aspects of D2D content delivery networks. Other works in the literature, e.g., [44], considered a much larger size of file library, however, their objective was to conduct the scaling analysis of caching networks.

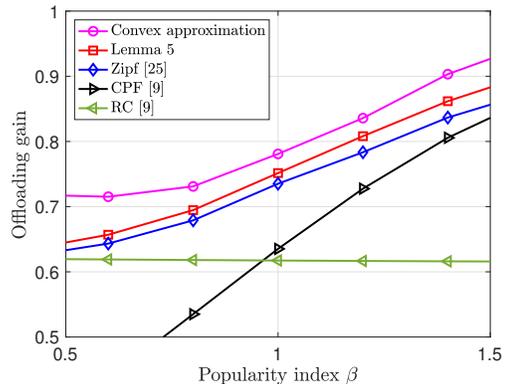


Fig. 7. The offloading gain versus the popularity of files β under different caching schemes ($N_f = 40$, $M = 8$).

are probabilistically cached according to their popularity as given in (4), see, e.g., [30]. Moreover, in Fig. 7, "convex approximation" refers to the caching solution characterized in Eq. (31), whereas "Lemma 5" refers to the caching solution characterized in Lemma 5. All the caching schemes are evaluated under CoMP transmissions. We first see that the offloading gains under the optimized PC schemes outperform conventional caching schemes. Moreover, the PC based on convex approximation in Eq. (31) is superior to the suboptimal solution of Lemma 5. As β increases, except for RC, the offloading gain increases and gradually, the optimized PC, Zipf, and CPF schemes tend to achieve the same performance. This shows that when a small portion of files becomes highly demanded, i.e., for higher β , the optimal caching probability is attained via caching popular files among all cluster devices.

To show the prominent effect of the network geometry on the optimized caching probability, we plot the histograms of the solution of Lemma 5 for three different cases in Fig. 8. These three cases are ranging from a sparse network (small σ and λ_p), a relatively dense network (medium σ and λ_p), and a highly dense network (large σ and λ_p). Note that smaller σ results in higher desired signal power, while smaller λ_p yields lower inter-cluster interference power as clusters become sparser. The first case in Fig. 8(a) represents a sparse system with small values of λ_p and σ , i.e., sufficiently good transmission conditions. In this case, we see that the optimized caching probability tends to be more uniform taking advantage of hitting a large number of files while being served in favorable transmission conditions. The second case in Fig. 8(b) represents a system with relatively good transmission conditions, i.e., medium values of σ and λ_p . It is clear from the histogram that the optimized caching solution tends to be more skewed than in the first case. The third case in Fig. 8(c) is then for a highly dense network with large values of both σ and λ_p . Clearly, the caching probability tends to be very skewed, which implies that caching popular files is an appropriate choice for such a highly dense network, i.e., a network with poor transmission conditions. Summing up, the results in Fig. 8 reveal interesting dependence of the optimized caching probability in Lemma 5 on the network geometry.

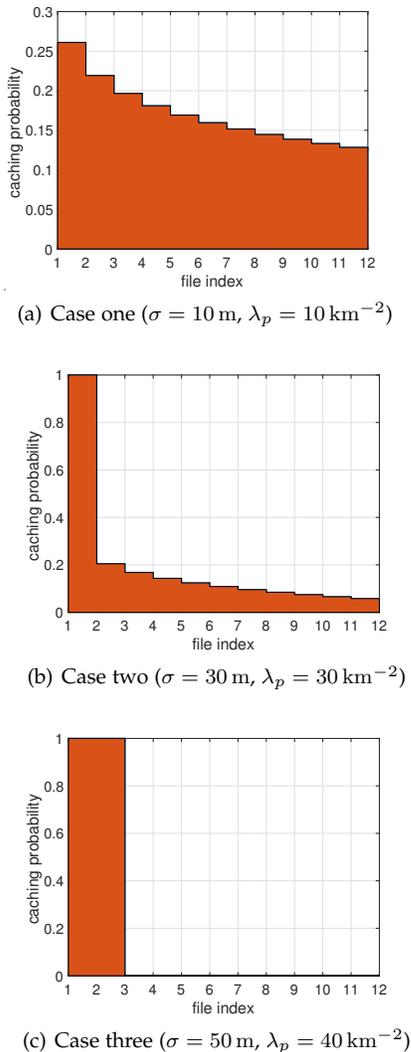


Fig. 8. Histogram of the caching solution \underline{c}^* is plotted for different network geometries ($\beta = 0.4$).

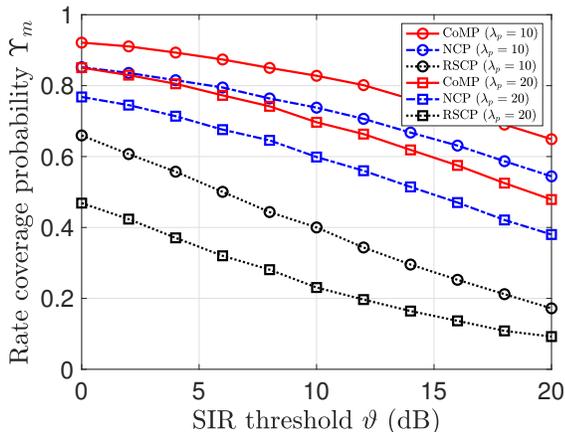


Fig. 9. The rate coverage probability versus the SIR threshold ϑ for different transmission schemes ($\bar{n} = 20$, $c_m = 0.5$).

6.3 Comparison with Other Transmission Schemes

To quantify how much CoMP transmission can improve the achievable performance, we here compare the rate coverage

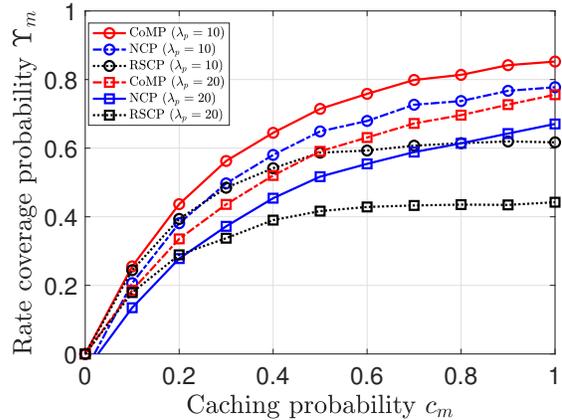


Fig. 10. The rate coverage probability versus the caching probability c_m for different transmission schemes ($\bar{n} = 10$, $\vartheta = 5 \text{ dB}$).

probability Υ_m for three transmission schemes, namely, CoMP, NCP, and randomly-selected content provider (RSCP). Recall that for the NCP scheme, the requested content is served from the the nearest active provider to the content client within the same cluster while for the RSCP scheme, an active provider is chosen at random to serve the desired content. Firstly, Fig. 9 plots the exact rate coverage probability versus SIR threshold ϑ for different density of clusters λ_p . Intuitively, CoMP transmissions achieves higher rate coverage probability than those of the other schemes. In particular, at high SIR threshold, allowing CoMP transmissions can provide up to 300% improvement in the rate coverage probability compared to the RSCP scheme. Moreover, for all schemes, the rate coverage probability is seen to decrease as λ_p increases since higher interference power is encountered at the typical client when D2D clusters are denser.

In light of this comparison, Fig. 10 plots the rate coverage probability against the caching probability c_m for the three transmission schemes. As shown in Fig. 10, as the content availability increases, i.e., higher c_m , the rate coverage probability improves. Besides, Fig. 10 illustrates that while the rate coverage probability for the RSCP transmission scheme tends to flatten when c_m further increases, it continues increasing for CoMP and NCP transmission schemes. This is attributed to the fact that, for the NCP scheme, the serving distance is more likely to decrease with the increase of c_m , and hence the corresponding performance improves. Similarly, for the CoMP scheme, the transmission diversity improves with c_m since the average number of active and caching providers increases.

Finally, Fig. 11 shows the effect of the average number of devices per cluster \bar{n} on the rate coverage probability Υ_m (left hand side y-axis). Fig. 11 also presents the amount of consumed energy per content request in one cluster (right hand side y-axis). Fig. 11 first shows that as \bar{n} increases, the rate coverage probability for CoMP and NCP schemes increases. This is due to the fact that, for the NCP scheme, the nearest serving distance is more likely to decrease for congested clusters. Moreover, for the CoMP scheme, the transmission diversity also improves with \bar{n} . However, for the RSCP scheme, the rate coverage probability first in-

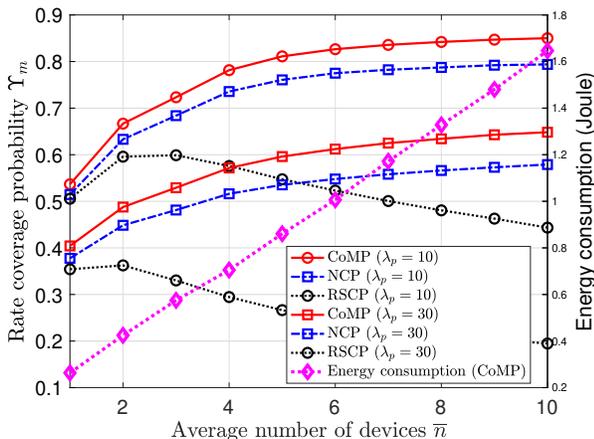


Fig. 11. The rate coverage probability (left hand side y-axis) and the amount of consumed energy per content request in one cluster (right hand side y-axis) versus the average number of devices per cluster \bar{n} ($\vartheta = 8$ dB, $c_m = 1$, $P_d = 20$ dBm, $\bar{S} = 5$ MBytes, $W = 5$ MHz).

increases driven by the increasing probability of finding a requested content within the local cluster. Then, the rate coverage probability turns to decrease as \bar{n} further increases. This degradation is attributed to the fact that while the requested content can be found within the local cluster with high probability for higher \bar{n} , the effect of inter-cluster interference also grows with \bar{n} . The latter then becomes the dominant factor that drives the rate coverage probability Υ_m down. Please notice that as \bar{n} increases, the number of collaborative devices in other remote clusters correspondingly increases, which yields higher interference power. Furthermore, the rate coverage probability is shown to decrease for all schemes when λ_p increases driven by the growing effects of interference. Besides, while the performance of RSCP is very sensitive to the density of clusters, especially when the average number of devices is high, cooperative transmission attains substantially better performance. We hence conclude that for such highly dense D2D networks and adverse interference conditions, cooperative transmission becomes more appealing. From Fig. 11, we also see that the consumed energy monotonically increases with the average number of devices per cluster \bar{n} . This is attributed to the fact as the number of content providers increases with \bar{n} , more energy will be consumed to serve a content. The average energy consumption per content request in one cluster is calculated from $E_m = q_m c_m \bar{n} \bar{E}$, with $\bar{E} = \frac{P_d \bar{S}}{R_d}$ being the consumed energy per requested content per content provider; P_d denotes the device transmission power, \bar{S} is the average content size, and R_d is the D2D average transmission rate that is calculated numerically from $R_d = W \log_2(1 + \vartheta) \Upsilon_m$.

7 CONCLUSION

In this paper, we have conducted performance analysis and content placement optimization for cache-assisted CoMP transmissions in clustered D2D networks. In particular, we have characterized the rate coverage probability and offloading gain as functions of the network parameters, namely, the density of clusters, average number of devices per cluster, and the content popularity and placement schemes. Then, we have sought simple yet tight lower

bound and approximation of the rate coverage probability and offloading gain. Based on the obtained results, we have shown that the inter-cluster interference of a TCP is upper bounded by that of a PPP of the same intensity. Moreover, we have formulated the corresponding offloading gain maximization problem and obtained optimized caching probabilities based on the proposed lower bound and approximation. Results showed that allowing CoMP transmissions can attain up to 300% improvement in the rate coverage probability compared to the RSCP scheme. Finally, we conclude by showing that the proposed optimized PC results in a considerable improvement of the offloading gain over conventional caching schemes.

REFERENCES

- [1] R. Amer, H. ElSawy, J. Kibilda, M. M. Butt, and N. Marchetti, "Cooperative transmission and probabilistic caching for clustered D2D networks," in *IEEE Wireless Communications and Networking Conference (WCNC)*, Marrakech, Morocco, April 2019.
- [2] V. Chandrasekhar, J. G. Andrews, and A. Gatherer, "Femtocell networks: a survey," *Proc. of IEEE Communications magazine*, vol. 46, no. 9, 2008.
- [3] E. Bastug, M. Bennis, and M. Debbah, "Living on the edge: The role of proactive caching in 5G wireless networks," *IEEE Communications Magazine*, vol. 52, no. 8, pp. 82–89, Aug 2014.
- [4] R. Amer, M. M. Butt, and N. Marchetti, "Caching at the edge in low latency wireless networks," *Wireless Automation as an Enabler for the Next Industrial Revolution*, pp. 209–240, 2020.
- [5] P. Marsch and G. P. Fettweis, *Coordinated Multi-Point in Mobile Communications: from theory to practice*. Cambridge University Press, 2011.
- [6] R. Amer, M. M. Butt, M. Bennis, and N. Marchetti, "Inter-cluster cooperation for wireless D2D caching networks," *IEEE Transactions on Wireless Communications*, vol. 17, no. 9, pp. 6108–6121, September 2018.
- [7] —, "Delay analysis for wireless D2D caching with inter-cluster cooperation," in *IEEE Global Communications Conference (GLOBECOM)*, Singapore, Dec. 2017.
- [8] G. Nigam, P. Minero, and M. Haenggi, "Coordinated multipoint joint transmission in heterogeneous networks," *IEEE Transactions on Communications*, vol. 62, no. 11, pp. 4134–4146, 2014.
- [9] M. Ji, G. Caire, and A. F. Molisch, "Fundamental limits of caching in wireless D2D networks," *IEEE Transactions on Information Theory*, vol. 62, no. 2, pp. 849–869, Feb 2016.
- [10] M. Sawahashi, Y. Kishiyama, A. Morimoto, D. Nishikawa, and M. Tanno, "Coordinated multipoint transmission/reception techniques for LTE-advanced [coordinated and distributed MIMO]," *IEEE Wireless Communications*, vol. 17, no. 3, pp. 26–34, June 2010.
- [11] C. Chaccour, R. Amer, B. Zhou, and W. Saad, "On the reliability of wireless virtual reality at terahertz (THz) frequencies," in *10th IFIP International Conference on New Technologies*, Spain, June. 2019.
- [12] K. E. Hunter, S. A. Sprigg, N. J. Meijers, C. S. Wurster, and P. E. Jacobs, "Retail proximity marketing," Nov. 7 2013, uS Patent App. 13/833,110.
- [13] Z. Chen, J. Lee, T. Q. S. Quek, and M. Kountouris, "Cooperative caching and transmission design in cluster-centric small cell networks," *IEEE Transactions on Wireless Communications*, vol. 16, no. 5, pp. 3401–3415, May 2017.
- [14] S. H. Chae, T. Q. S. Quek, and W. Choi, "Content placement for wireless cooperative caching helpers: A tradeoff between cooperative gain and content diversity gain," *IEEE Transactions on Wireless Communications*, vol. 16, no. 10, pp. 6795–6807, Oct 2017.
- [15] W. C. Ao and K. Psounis, "Distributed caching and small cell cooperation for fast content delivery," in *Proceedings of the 16th ACM International Symposium on Mobile Ad Hoc Networking and Computing*. ACM, 2015, pp. 127–136.
- [16] G. Zheng, H. A. Suraweera, and I. Krikidis, "Optimization of hybrid cache placement for collaborative relaying," *IEEE Communications Letters*, vol. 21, no. 2, pp. 442–445, 2017.
- [17] A. Daghaj, H. Zhu, and J. Wang, "Content delivery analysis in multiple devices to single device communications," *IEEE Transactions on Vehicular Technology*, 2018.

- [18] B. Chen, C. Yang, and G. Wang, "High throughput opportunistic cooperative device-to-device communications with caching," *IEEE transactions on vehicular technology*, vol. 66, no. 8, pp. 7527–7539, 2017.
- [19] G. Kim, B. Hong, W. Choi, and H. Park, "Mds coded caching leveraged by coordinated multi-point transmission," *IEEE Communications Letters*, 2018.
- [20] Y. Zhang, E. Pan, L. Song, W. Saad, Z. Dawy, and Z. Han, "Social network aware device-to-device communication in wireless networks," *IEEE Transactions on Wireless Communications*, vol. 14, no. 1, pp. 177–190, January 2015.
- [21] X. Hu, L. Meng, and A. D. Striegel, "Evaluating the raw potential for device-to-device caching via co-location," *Procedia Computer Science*, vol. 34, pp. 376–383, 2014.
- [22] M. Afshang, H. S. Dhillon, and P. H. J. Chong, "Modeling and performance analysis of clustered device-to-device networks," *IEEE Transactions on Wireless Communications*, vol. 15, no. 7, pp. 4957–4972, July 2016.
- [23] —, "Fundamentals of cluster-centric content placement in cache-enabled device-to-device networks," *IEEE Transactions on Communications*, vol. 64, no. 6, pp. 2511–2526, June 2016.
- [24] M. Haenggi, *Stochastic geometry for wireless networks*. Cambridge University Press, 2012.
- [25] N. Deng and M. Haenggi, "The benefits of hybrid caching in gauss-poisson D2D networks," *IEEE Journal on Selected Areas in Communications*, vol. 36, no. 6, pp. 1217–1230, June 2018.
- [26] R. Amer, M. M. Butt, and N. Marchetti, "Optimizing joint probabilistic caching and channel access for clustered D2D networks," *IEEE Wireless Communication Letter*, to appear, 2020.
- [27] R. Amer, H. Elsawy, M. M. Butt, E. A. Jorswieck, M. Bennis, and N. Marchetti, "Optimized caching and spectrum partitioning for D2D enabled cellular systems with clustered devices," in *IEEE Transactions on Communications*, 2020.
- [28] R. Amer, M. M. Butt, H. Elsawy, M. Bennis, J. Kibilda, and N. Marchetti, "On minimizing energy consumption for D2D clustered caching networks," in *IEEE Global Communications Conference (GLOBECOM)*, December 2018, pp. 1–6.
- [29] H. Elsawy, H. Dahrouj, T. Y. Al-naffouri, and M. Alouini, "Virtualized cognitive network architecture for 5G cellular networks," *IEEE Communications Magazine*, vol. 53, no. 7, pp. 78–85, July 2015.
- [30] L. Breslau, P. Cao, L. Fan, G. Phillips, and S. Shenker, "Web caching and zipf-like distributions: Evidence and implications," in *Eighteenth Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, vol. 1, 1999, pp. 126–134.
- [31] M. Lee, A. F. Molisch, N. Sastry, and A. Raman, "Individual preference probability modeling and parameterization for video content in wireless caching networks," *IEEE/ACM Transactions on Networking*, vol. 27, no. 2, pp. 676–690, April 2019.
- [32] B. Blaszczyszyn and A. Giovanidis, "Optimal geographic caching in cellular networks," in *Proc. of IEEE International Conference on Communications (ICC)*, London, UK, 2015.
- [33] Z. Chen, N. Pappas, and M. Kountouris, "Probabilistic caching in wireless D2D networks: Cache hit optimal versus throughput optimal," *IEEE Communications Letters*, vol. 21, no. 3, pp. 584–587, March 2017.
- [34] B. Blaszczyszyn and A. Giovanidis, "Optimal geographic caching in cellular networks," in *IEEE International Conference on Communications (ICC)*, London, UK, June 2015, pp. 3358–3363.
- [35] S. H. Chae and W. Choi, "Caching placement in stochastic wireless caching helper networks: Channel selection diversity via caching," *IEEE Trans. Wireless Communications*, vol. 15, no. 10, pp. 6626–6637, Oct 2016.
- [36] M. Lee and A. F. Molisch, "Caching policy and cooperation distance design for base station-assisted wireless D2D caching networks: Throughput and energy efficiency optimization and tradeoff," *IEEE Transactions on Wireless Communications*, vol. 17, no. 11, pp. 7500–7514, November 2018.
- [37] M. Naslcheraghi, M. Afshang, and H. S. Dhillon, "Modeling and performance analysis of full-duplex communications in cache-enabled D2D networks," in *2018 IEEE International Conference on Communications (ICC)*, May 2018, pp. 1–6.
- [38] R. Amer, W. Saad, H. Elsawy, M. Butt, and N. Marchetti, "Caching to the sky: Performance analysis of cache-assisted CoMP for cellular-connected UAVs," in *Proc. of the IEEE Wireless Communications and Networking Conference (WCNC)*, Marrakech, Morocco, April. 2019.
- [39] R. Amer, W. Saad, and N. Marchetti, "Mobility in the sky: Performance and mobility analysis for cellular-connected UAVs," *IEEE Transactions on Communications*, pp. 1–1, 2020.
- [40] M. A. Abd-Elmagid, M. A. Kishk, and H. S. Dhillon, "Joint energy and SINR coverage in spatially clustered RF-powered IoT network," *IEEE Transactions on Green Communications and Networking*, vol. 3, no. 1, pp. 132–146, March 2019.
- [41] B. Chen, C. Yang, and Z. Xiong, "Optimal caching and scheduling for cache-enabled D2D communications," *IEEE Communications Letters*, vol. 21, no. 5, pp. 1155–1158, May 2017.
- [42] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.
- [43] S. H. Chae, J. Y. Ryu, T. Q. S. Quek, and W. Choi, "Cooperative transmission via caching helpers," in *IEEE Global Communications Conference (GLOBECOM)*, Dec 2015, pp. 1–6.
- [44] N. Golrezaei *et al.*, "Base-station assisted device-to-device communications for high-throughput wireless video networks," *IEEE Transactions on Wireless Communications*, vol. 13, no. 7, pp. 3665–3676, 2014.
- [45] M. Afshang, C. Saha, and H. S. Dhillon, "Nearest-neighbor and contact distance distributions for thomas cluster process," *IEEE Wireless Communications Letters*, vol. 6, no. 1, pp. 130–133, February 2017.



Ramy Amer received the MSc degree in Electrical Engineering from Alexandria University, Egypt, in 2016. Since 2016, he is pursuing his PhD degree in wireless communication, especially, wireless caching, under the supervision of Dr Nicola Marchetti at CONNECT centre, Trinity College Dublin, Ireland. His research interests include cross-layer design, Cognitive Radio, Energy Harvesting, Stochastic Geometry, and Reinforcement Learning. Prior to CONNECT, he was also an assistant lecturer for the switching department, National Telecommunication Institute of Egypt (NTI), Cairo, Egypt, where he conducted professional training both at the national and international levels. He was a visiting scholar at Wireless at Virginia Tech, with Prof Walid Saad's group from September 2018 to March 2019. He is the recipient of best paper award from IFIP NTMS in 2019 and the IEEE student travel grant from WCNC in 2019. He is recognized as an Exemplary Reviewer by the IEEE TRANSACTIONS ON COMMUNICATIONS 2019. He is also certified as a Cisco instructor and has other Cisco data and voice certificates. He worked as a part-time instructor for many national and international training centres, e.g., New-Horizon Egypt and Fast-Lane KSA.



Hesham Elsawy (S'10 – M'14 – SM'17) is an assistant professor at King Fahd University of Petroleum and Minerals (KFUPM), Saudi Arabia. Prior to that, he was a postdoctoral Fellow at King Abdullah University of Science and Technology (KAUST), Saudi Arabia, a research assistant at TRTech, Winnipeg, MB, Canada, and a telecommunication engineer at the National Telecommunication Institute, Egypt. Dr. Elsawy received his Ph.D. degree in Electrical Engineering from the University of Manitoba, Canada,

in 2014, where he received several academic awards, including the NSERC Industrial Postgraduate Scholarship during the period of 2010–2013, and the TRTech Graduate Students Fellowship in the period of 2010–2014. He co-authored three award-winning papers that are recognized by the IEEE COMSOC Best Survey Paper Award, the Best Scientific Contribution Award to the IEEE International Symposium on Wireless Systems 2017, and the Best Paper Award in Small Cell and 5G Networks (SmallNets) Workshop of the 2015 IEEE International Conference on Communications (ICC). He is the recipient of the IEEE ComSoc Outstanding Young Researcher Award for Europe, Middle East, & Africa Region in 2018. He is recognized as an exemplary reviewer by the IEEE Transactions on Communications for the three years 2014–2016, by the IEEE Transactions on Wireless Communications in 2017 and 2018, and by the IEEE Wireless Communications Letters in 2018.



Jacek Kibilda received the M.Sc. degree from the Poznan University of Technology, Poznan, Poland, in 2008, and the Ph.D. degree from Trinity College, The University of Dublin, Dublin, Ireland, in 2016. He is currently a Research Fellow with CONNECT, Trinity College, The University of Dublin. His research interests include architectures and models for future mobile networks.



Nicola Marchetti Dr. Nicola Marchetti is currently Assistant Professor in Wireless Communications at Trinity College Dublin, Ireland. He performs his research under the Trinity Information and Complexity Labs (TRICKLE) and the Irish Research Centre for Future Networks and Communications (CONNECT). He received the PhD in Wireless Communications from Aalborg University, Denmark in 2007, and the M.Sc. in Electronic Engineering from University of Ferrara, Italy in 2003. He also holds an M.Sc. in

Mathematics which he received from Aalborg University in 2010. His collaborations include research projects in cooperation with Nokia Bell Labs and US Air Force Office of Scientific Research, among others. His research interests include Adaptive and Self-Organizing Networks, Complex Systems Science for Communication Networks, PHY Layer, Radio Resource Management. He has authored 130 journals and conference papers, 2 books and 8 book chapters, holds 3 patents, and received 4 best paper awards.



M. Majid Butt (S'07 – M'10 – SM'15) received the MSc degree in Digital Communications from Christian Albrechts University, Kiel, Germany, in 2005, and the PhD degree in Telecommunications from the Norwegian University of Science and Technology, Trondheim, Norway, in 2011. He is an Assistant Professor at University of Glasgow as well as an adjunct Assistant Professor at Trinity College Dublin, Ireland. Before that, he has held senior researcher positions at Trinity College Dublin, Ireland and Qatar University. He

is recipient of Marie Curie Alain Bensoussan postdoctoral fellowship from European Research Consortium for Informatics and Mathematics (ERCIM). He held ERCIM postdoc fellow positions at Fraunhofer Heinrich Hertz Institute, Germany, and University of Luxembourg. Dr. Majid's major areas of research interest include communication techniques for wireless networks with focus on radio resource allocation, scheduling algorithms, energy efficiency and cross layer design. He has authored more than 50 peer reviewed conference and journal publications in these areas. He has served as TPC chair for various communication workshops in conjunction with IEEE WCNC, ICUWB, CROWNCOM, IEEE Greencom and Globecom. He is a senior member of IEEE and serves as an associate editor for IEEE Access journal and IEEE Communication Magazine since 2016.

APPENDIX A

PROOF OF LEMMA 1

In the following, by saying $u \in \Phi_{cp}$, we mean that $\mathbf{y} \in \Phi_{cp}$, where $u = \|\mathbf{x} + \mathbf{y}\|$.

$$\begin{aligned} \mathcal{L}_{I_{\text{out}}}(t) &= \mathbb{E} \left[e^{-t\gamma_d \sum_{\Phi_p^!} \sum_{u \in \Phi_{cp}} G_u u^{-\alpha}} \right] \\ &= \mathbb{E}_{\Phi_p} \left[\prod_{\Phi_p^!} \mathbb{E}_{\Phi_{cp}, G_u} e^{-t\gamma_d \sum_{u \in \Phi_{cp}} G_u u^{-\alpha}} \right] \end{aligned} \quad (36)$$

$$= \mathbb{E}_{\Phi_p} \left[\prod_{\Phi_p^!} \mathbb{E}_{\Phi_{cp}} \prod_{u \in \Phi_{cp}} \mathbb{E}_{u, G_u} e^{-t\gamma_d G_u u^{-\alpha}} \right], \quad (37)$$

where $G_u = G_{\mathbf{y}}$ for ease of exposition; from the Rayleigh fading assumption, we get

$$\begin{aligned} \mathcal{L}_{I_{\text{out}}}(t) &= \mathbb{E}_{\Phi_p} \left[\prod_{\Phi_p^!} \mathbb{E}_{\Phi_{cp}} \prod_{u \in \Phi_{cp}} \mathbb{E}_u \frac{1}{1 + t\gamma_d u^{-\alpha}} \right] \\ &\stackrel{(a)}{=} \mathbb{E}_{\Phi_p} \left[\prod_{\Phi_p^!} \exp \left(-p\bar{n} \int_{u=0}^{\infty} \left(1 - \frac{1}{1 + t\gamma_d u^{-\alpha}} \right) f_U(u|v) du \right) \right], \end{aligned}$$

where (a) follows from the probability generating functional (PGFL) of the Gaussian PPP Φ_{cp} . Notice that, in step (a), the PGFL of the Gaussian PPP Φ_{cp} is adopted for the intensity function given in polar coordinates rather than Cartesian coordinates, i.e., $v = \|\mathbf{x}\|$ and $u = \|\mathbf{x} + \mathbf{y}\|$. Substituting $\int_{u=0}^{\infty} \left(1 - \frac{1}{1 + t\gamma_d u^{-\alpha}} \right) f_U(u|v) du = \int_{u=0}^{\infty} \frac{t\gamma_d}{u^\alpha + t\gamma_d} f_U(u|v) du = \zeta(v, t)$, we get

$$\begin{aligned} \mathcal{L}_{I_{\text{out}}}(t) &= \mathbb{E}_{\Phi_p} \left[\prod_{\Phi_p^!} \exp(-p\bar{n}\zeta(v, t)) \right] \\ &\stackrel{(b)}{=} \exp \left(-2\pi\lambda_p \int_{v=0}^{\infty} \left(1 - \exp(-p\bar{n}\zeta(v, t)) \right) v dv \right), \end{aligned} \quad (38)$$

where (b) follows from the PGFL of the PPP Φ_p . Hence, Lemma 1 is proven.

APPENDIX B

PROOF OF THEOREM 1

By conditioning on $S_{\Phi_{cpm}} = s_{\Phi_{cpm}} = \sum_{i=1}^k h_i^{-\alpha}$, we derive a bound on Laplace transform of inter-cluster interference based on Taylor's series expansion. Starting from equation (37) in Appendix A, we have

$$\begin{aligned} \mathcal{L}_{I_{\text{out}}}(t|k) &= \mathbb{E}_{\Phi_p} \left[\prod_{\Phi_p^!} \mathbb{E}_{\Phi_{cp}} \prod_{u \in \Phi_{cp}} \mathbb{E}_{u, G_u} \exp(-tG_u u^{-\alpha}) \right] \\ &\stackrel{(a)}{=} \mathbb{E}_{G_u} \exp \left(-2\pi\lambda_p \int_{v=0}^{\infty} \left(1 - \exp(-p\bar{n}(1 - \zeta'(v, t))) \right) v dv \right) \\ &\stackrel{(b)}{\approx} \mathbb{E}_{G_u} \exp \left(-2\pi\lambda_p \int_{v=0}^{\infty} \left(1 - (1 - p\bar{n}(1 - \zeta'(v, t))) \right) v dv \right) \\ &= \exp \left(-2\pi p\bar{n}\lambda_p \left(\overbrace{\int_{v=0}^{\infty} v dv}^{J(t)} - \mathbb{E}_{G_u} \int_{v=0}^{\infty} \zeta'(v, t) v dv \right) \right), \end{aligned}$$

where $\zeta'(v, t) = \int_{u=0}^{\infty} e^{-t\gamma_d G_u u^{-\alpha}} f_U(u|v) du$ and $G_u = G_{\mathbf{y}}$ for ease of notation; (a) follows from tracking the proof of Lemma 1 up until equation (38), (b) follows from Taylor

series expansion for exponential function $e^{-x} \approx 1 - x$ when x is small. It is worth mentioning that the obtained $\mathcal{L}_{I_{\text{out}}}(t|k)$ in the above is Laplace transform of an upper bound on the interference. Correspondingly, the resulting rate coverage probability Υ_m and offloading gain $\mathbb{P}_o^{\sim}(c)$ are lower bounds on their exact values. We proved in [28, Lemma 2] that $J(t) = \frac{(t\gamma_d)^{2/\alpha}}{2} \Gamma(1 + 2/\alpha) \Gamma(1 - 2/\alpha)$, which proves (16). Plugging the result obtained in Eq. (16) into (6) yields the lower bound on the offloading gain in Eq. (17), which completes the proof.

APPENDIX C

PROOF OF LEMMA 2

With reference to Fig. 1, the nearest serving distance h_1 is defined as the distance from the typical client at $(0, 0)$ to its nearest provider within the same cluster. Following [45], we define the point generating function (PGF) of the number of active clients that cache content m within a ball $\mathbf{b}(o, h_1)$ with radius h_1 and centered around the origin o as:

$$\begin{aligned} G_N(\vartheta) &= \mathbb{E} \left[\vartheta^{\sum_{\mathbf{y}_{0i} \in \Phi_{cpm}} \mathbf{1}\{\|\mathbf{x}_0 + \mathbf{y}_{0i}\| < h_1\}} \right] \\ &= \mathbb{E}_{\Phi_c, \mathbf{x}_0} \prod_{\mathbf{y}_{0i} \in \Phi_{cpm}} \left[\vartheta^{\mathbf{1}\{\|\mathbf{x}_0 + \mathbf{y}_{0i}\| < h_1\}} \right] \\ &\stackrel{(a)}{=} \mathbb{E}_{\mathbf{x}_0} \exp \left(-c_m p\bar{n} \int_{\mathbb{R}^2} \left(1 - \vartheta^{\mathbf{1}\{\|\mathbf{x}_0 + \mathbf{y}_{0i}\| < h_1\}} \right) f_{\mathbf{Y}_{0i}}(\mathbf{y}_{0i}) d\mathbf{y}_{0i} \right) \\ &\stackrel{(b)}{=} \mathbb{E}_{\mathbf{x}_0} \exp \left(-c_m p\bar{n} \int_{\mathbb{R}^2} \left(1 - \vartheta^{\mathbf{1}\{\|\mathbf{z}_0\| < h_1\}} \right) f_{\mathbf{Y}_{0i}}(\mathbf{z}_0 - \mathbf{x}_0) d\mathbf{z}_0 \right), \end{aligned}$$

where $\mathbf{1}\{\cdot\}$ is the indicator function, and $\mathbf{x}_0 \in \mathbb{R}^2$ is a RV modeling the location of representative cluster center relative to the origin o , with a realization $\mathbf{X}_0 = \mathbf{x}_0$; (a) follows from the PGFL of the PPP Φ_{cpm} along with its intensity function $c_m p\bar{n} f_{\mathbf{Y}_{0i}}(\mathbf{y}_{0i})$, and (b) follows from change of variables $\mathbf{z}_0 = \mathbf{x}_0 + \mathbf{y}_{0i}$. By converting Cartesian coordinates to polar coordinates with $h = \|\mathbf{z}_0\|$, we get $G_N(\vartheta) =$

$$\begin{aligned} \mathbb{E}_{V_0} \exp \left(-c_m p\bar{n} \int_{h=0}^{\infty} \left(1 - \vartheta^{\mathbf{1}\{h < h_1\}} \right) f_{H|V_0}(h|v_0) dh \right) \\ \stackrel{(c)}{=} \mathbb{E}_{V_0} \exp \left(-c_m p\bar{n} \int_{h=0}^{h_1} (1 - \vartheta) f_{H|V_0}(h|v_0) dh \right) \stackrel{(d)}{=} \\ \int_{v_0=0}^{\infty} f_{V_0}(v_0) \exp \left(-c_m p\bar{n} \int_{h=0}^{h_1} (1 - \vartheta) f_{H|V_0}(h|v_0) dh \right) dv_0, \end{aligned}$$

where $V_0 \in \mathbb{R}$ is a RV modeling the distance from representative cluster's center to the origin o , with a realization $V_0 = v_0 = \|\mathbf{x}_0\|$; (c) follows from the definition of the indicator function $\mathbf{1}\{h < h_1\}$, and (d) follows from unconditioning over v_0 . To clarify how the normal distribution $f_{\mathbf{Y}_{0i}}(\mathbf{z}_0 - \mathbf{x}_0)$ is converted to the Rician distribution $f_{H|V_0}(h|v_0)$, consider first the representative cluster centered at $\mathbf{x}_0 \in \Phi_p$, with a distance $v_0 = \|\mathbf{x}_0\|$ from the origin. A randomly-selected active provider belonging to the representative cluster has its coordinates in \mathbb{R}^2 chosen independently from Gaussian distributions with standard deviation σ . Then, by definition, the distance h from such an active provider to the origin has Rician PDF denoted as $f_{H|V_0}(h|v_0)$. Recall that $f_{V_0}(v_0) = \text{Rayleigh}(v_0, \sigma)$ from the definition of Gaussian PPP.

Now, the CDF of nearest serving distance $F_{H_1}(h_1)$ can be derived as

$$F_{H_1}(h_1) = 1 - G_N(0) = 1 - \int_{v_0=0}^{\infty} f_{V_0}(v_0) \exp\left(-c_m p \bar{n} \int_0^{h_1} f_{H|V_0}(h|v_0) dh\right) dv_0. \quad (39)$$

Applying Leibniz integral rule, we obtain the nearest distance PDF as

$$\begin{aligned} f_{H_1}(h_1) &= -\frac{\partial}{\partial h_1} \int_{v_0=0}^{\infty} f_{V_0}(v_0) e^{-c_m p \bar{n} \int_0^{h_1} f_{H|V_0}(h|v_0) dh} dv_0 \\ &= -\int_{v_0=0}^{\infty} f_{V_0}(v_0) \frac{\partial}{\partial h_1} e^{-c_m p \bar{n} \int_0^{h_1} f_{H|V_0}(h|v_0) dh} dv_0 = c_m p \bar{n} \times \\ &\int_0^{\infty} f_{V_0}(v_0) \frac{\partial}{\partial h_1} \left[\int_0^{h_1} f_{H|V_0}(h|v_0) dh \right] e^{-c_m p \bar{n} \int_0^{h_1} f_{H|V_0}(h|v_0) dh} dv_0 \\ &= c_m p \bar{n} \int_{v_0=0}^{\infty} f_{V_0}(v_0) f_{H_1|V_0}(h_1|v_0) e^{-c_m p \bar{n} \int_0^{h_1} f_{H|V_0}(h|v_0) dh} dv_0, \end{aligned}$$

The distance PDF $f_{H_1}(h_1)$ can be calculated numerically from (19). However, a tractable yet accurate approximation can be obtained using Jensen's inequality as follows:

$$\begin{aligned} f_{H_1}(h_1) &= \frac{\partial}{\partial h_1} F_{H_1}(h_1) \\ &\stackrel{(a)}{\approx} \frac{\partial}{\partial h_1} \left(1 - e^{-c_m p \bar{n} \int_0^{h_1} \int_0^{\infty} f_{V_0}(v_0) f_{H|V_0}(h|v_0) dv_0 dh} \right) \\ &\stackrel{(b)}{=} \frac{\partial}{\partial h_1} \left(1 - \exp\left(-c_m p \bar{n} \left(1 - \exp\left(-\frac{h_1^2}{4\sigma^2}\right)\right)\right) \right) \\ &= \frac{c_m p \bar{n} h_1 e^{-c_m p \bar{n} \left(1 - e^{-\frac{h_1^2}{4\sigma^2}}\right)} - \frac{h_1^2}{4\sigma^2}}{2\sigma^2}, \end{aligned} \quad (40)$$

where (a) follows from Jensen's inequality applied to the CDF $F_{H_1}(h_1)$, and (b) follows from $\int_0^{h_1} \int_{v_0=0}^{\infty} f_{V_0}(v_0) f_{H|V_0}(h|v_0) dv_0 dh = 1 - \exp\left(-\frac{h_1^2}{4\sigma^2}\right)$. This completes the proof.

APPENDIX D PROOF OF LEMMA 3

The conditional variance $\text{Var}\left[S_{\Phi_{cpm}^!} | H_1 = h_1\right]$ can be expressed as

$$\begin{aligned} \text{Var}\left[S_{\Phi_{cpm}^!} | H_1 = h_1\right] &= \text{Var}\left[\sum_{\mathbf{y}_{0i} \in \Phi_{cpm}^!} \|\mathbf{x}_0 + \mathbf{y}_{0i}\|^{-\alpha}\right] \\ &\stackrel{(a)}{=} \int_{\mathbb{R}^2} \frac{1}{\|\mathbf{x}_0 + \mathbf{y}_{0i}\|^{2\alpha}} c_m p \bar{n} f_{\mathbf{Y}_{0i}}(\mathbf{y}_{0i}) d\mathbf{y}_{0i} \\ &\stackrel{(b)}{=} c_m p \bar{n} \int_{\mathbb{R}^2} \frac{1}{\|\mathbf{z}_0\|^{2\alpha}} f_{\mathbf{Y}_{0i}}(\mathbf{z}_0 - \mathbf{x}_0) d\mathbf{z}_0, \end{aligned} \quad (42)$$

where (a) follows from the mean and variance for PPPs [24, Corollary 4.8], along with the Gaussian PPP assumption Φ_{cpm} ; (b) follows from the substitution $\mathbf{z}_0 = \mathbf{x}_0 + \mathbf{y}_{0i}$, where $\{\mathbf{x}_0, \mathbf{y}_{0i}, \mathbf{z}_0\} \in \mathbb{R}^2$. By converting the Cartesian coordinates to polar coordinates, where $h = \|\mathbf{z}_0\|$, and unconditioning over v_0 , we get

$$\begin{aligned} \text{Var}\left[S_{\Phi_{cpm}^!} | H_1 = h_1\right] &= c_m p \bar{n} \int_{h_1}^{\infty} h^{-2\alpha} f_{H|V_0}(h|v_0) dh \\ &= c_m p \bar{n} \int_{v_0=0}^{\infty} f_{V_0}(v_0) \int_{h=h_1}^{\infty} h^{-2\alpha} f_{H|V_0}(h|v_0) dh dv_0 \end{aligned}$$

$$\stackrel{(c)}{=} c_m p \bar{n} \int_{h=h_1}^{\infty} h^{-2\alpha} \int_{v_0=0}^{\infty} f_{V_0}(v_0) f_{H|V_0}(h|v_0) dv_0 dh, \quad (43)$$

where (c) follows from changing the order of integration. Finally, we proceed as follows: $\text{Var}\left[S_{\Phi_{cpm}^!} | H_1 = h_1\right] =$

$$\begin{aligned} &c_m p \bar{n} \int_{h=h_1}^{\infty} h^{-2\alpha} \int_0^{\infty} \frac{v_0}{\sigma^2} e^{-\frac{v_0^2}{2\sigma^2}} \frac{h}{\sigma^2} e^{-\frac{-(v_0^2+h^2)}{2\sigma^2}} I_0\left(\frac{hv_0}{\sigma^2}\right) dv_0 dh \\ &= \frac{c_m p \bar{n}}{\sigma^2} \int_{h=h_1}^{\infty} h^{1-2\alpha} \int_{v_0=0}^{\infty} \frac{v_0}{\sigma^2} e^{-\frac{v_0^2}{2\sigma^2}} e^{-\frac{-(v_0^2+h^2)}{2\sigma^2}} I_0\left(\frac{hv_0}{\sigma^2}\right) dv_0 dh \\ &= \frac{c_m p \bar{n}}{\sigma^2} \int_{h=h_1}^{\infty} h^{1-2\alpha} e^{-\frac{h^2}{4\sigma^2}} \int_{v_0=0}^{\infty} \frac{v_0}{\sigma^2} e^{-\frac{v_0^2}{\sigma^2}} I_0\left(\frac{hv_0}{\sigma^2}\right) dv_0 dh \\ &\stackrel{(d)}{=} \frac{c_m p \bar{n}}{2\sigma^2} \int_{h_1}^{\infty} h^{1-2\alpha} e^{-\frac{h^2}{4\sigma^2}} dh \stackrel{(e)}{=} c_m p \bar{n} \int_{\frac{h_1^2}{4\sigma^2}}^{\infty} \tau^{-2\alpha} e^{-\tau} d\tau \end{aligned} \quad (44)$$

$$\stackrel{(f)}{=} c_m p \bar{n} \Gamma\left(-2\alpha + 1, \frac{h_1^2}{4\sigma^2}\right), \quad (45)$$

where (d) follows from solving the inner integrals of (44), (e) follows from the substitution $\tau = \frac{h^2}{4\sigma^2}$, and (f) follows from solving the integration of (45), where $\Gamma(\cdot, \cdot)$ denotes the upper incomplete gamma function. This completes the proof.