

Mobile-Traffic-Aware Offloading for Energy- and Spectral-Efficient Large-Scale D2D-Enabled Cellular Networks

Guogang Zhao, Sheng Chen, *Fellow, IEEE*, Lin Qi, *Member, IEEE*, Liqiang Zhao, *Member, IEEE*, and Lajos Hanzo, *Fellow, IEEE*

Abstract—This paper investigates how to enhance the energy and spectral efficiency (ESE) performance of large-scale cellular networks by offloading mobile traffic with the aid of device-to-device (D2D) communication. By appropriately exploiting the D2D-based mobile-traffic offloading mechanism, the users' behaviors and the specific network operating conditions, we develop an ESE evaluation framework for large-scale D2D-enabled cellular networks. This framework enables us to characterize the explicit relationship between the network's ESE and the offloading parameters as well as to quantify the influence of the users' behavior. Explicitly, we quantify the effects of the mobile-traffic intensity, the users' quality of service requirements as well as the base station density and other cellular system parameters on the achievable ESE. Tractable closed-form ESE-expressions are derived for a pair of spectrum sharing schemes, namely, D2D overlay and underlay in-band modes. Furthermore, we apply the analytical results to derive an optimal D2D-enabled mobile-traffic offloading scheme for D2D overlay cellular networks to maximize the network's ESE under a specific maximal cellular user outage and D2D transmitter power constraint. Numerical and simulation results are provided to verify our modeling accuracy and to demonstrate the impact of the system parameters on the achievable ESE.

Index Terms—Device-to-device (D2D) communication, D2D-enabled cellular network, mobile traffic offloading, spectral efficiency, energy efficiency

I. INTRODUCTION

Device-to-device (D2D) communication has been recognized as a significant component of supporting the emerging demands for popular local area services in the fifth-generation (5G) mobile network [1]–[3]. In a D2D-enabled cellular network, mobile devices (MDs) are allowed to directly link to other proximity devices without going through their associated base stations (BSs). Despite the increased interference caused by D2D communications, substantial performance benefits can be obtained, including more efficient resource utilization, extended cellular coverage, and enhanced network throughput.

G. Zhao and L. Qi (ggzhao@s-an.org, ielqi@zzu.edu.cn) are with School of Information Engineering, Zhengzhou University, Zhengzhou, 453001, China.

S. Chen and L. Hanzo (sqc@ecs.soton.ac.uk, lh@ecs.soton.ac.uk) are with School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, UK. S. Chen is also with King Abdulaziz University, Jeddah 21589, Saudi Arabia.

L. Zhao (lqzhao@mail.xidian.edu.cn) is with State Key Laboratory of Integrated Service Networks, Xidian University, Xian 710071, China.

This work was supported in part by European Research Council Advanced Fellow Grant through the Beam Me Up Project, in part by the Engineering and Physical Sciences Research Council under Grants EP/N004558/1 and EP/L018659/1, in part by the Royal Society's Wolfson Research Merit Award, and in part by China Scholarship Council (No. 201506960042).

It has been widely accepted that the evolution to the 5G system should not only dramatically enhance the end-user's quality of service (QoS) and the system capacity but also significantly improve the system's energy efficiency (EE) [4]. Thus, achieving an increased performance by exploiting the energy- and spectral resources more efficiently whilst satisfying specific end-user QoS constraints is a crucial design target in the D2D-enabled cellular network.

It is worth highlighting the main reason for introducing D2D communications into the cellular network in the first place. A substantial amount of cellular spectrum and energy resources are consumed by MDs to repeatedly download an ever-increasing number of common contents from cellular BSs, such as large files containing photos or videos in social networks [5]. According to Cisco [6], the global mobile data traffic grew 63 percent in 2016, reaching 7.2 exabytes per month at the end of 2016, and it is expected to increase sevenfold between 2016 and 2021, reaching 49 exabytes per month by 2021. It is inconceivable that the existing and future cellular networks can cope with such ever-increasing amount of mobile traffic without alternative relief mechanisms. In fact, the mobile traffic offloaded to WiFi already exceeded cellular traffic in 2015 [6]. This problem can be further mitigated by offloading more of the cellular traffic to D2D communication [7]–[9].

The physical operations supporting D2D-based mobile-traffic offloading include proximity discovery, mode selection and resource allocation [10]–[14]. These operations are always coupled with the users' behaviors and with the resource/QoS-constrained physical cellular situations [5]. Therefore, the modelling of the D2D-based mobile-traffic offloading and the evaluation of the overall energy and spectrum efficiency (ESE) performance are quite challenging [20]–[22]. Moreover, it is important to optimize the system performance by taking into account both the temporal and the geographic fluctuations of the users' behaviors as well as the time-varying physical cellular scenarios.

A. Related Works

In terms of resource utilization, such as spectrum sharing and power control, many schemes were proposed to improve the D2D communication performance [10]–[19]. For example, several resource allocation algorithms were proposed in [10] to maximize the throughput, where other system indicators, such as the maximum power constraint and rate constraints, are regarded as the constraints. The approach proposed in the study

[11] maximizes the throughput while considering both the D2D and cellular link qualities as constraints. The authors of [12] developed an energy-efficient spectrum sharing strategy, which jointly considers mode selection, uplink reuse allocation and power control. We also note that the contributions [12]–[19] consider the EE or ESE as a key D2D design indicator. It can be seen that these existing works analyze the EE and spectrum efficiency aspects of D2D communication. However, their analytical results were based on the single-cell scenario with a small number of users and, therefore, can hardly be adopted in the analysis and evaluation of large-scale D2D-enabled cellular networks. Several mobile-traffic offloading methods aimed at different service types were investigated in [20]–[25], but again, these results can only be applied to a local network area.

Performance analysis and optimization of large-scale D2D-enabled cellular networks were considered in [26]–[29]. Specifically, the study [26] proposed two different power control schemes and analyzed the coverage performance in a large-scale D2D-assisted cellular network. A D2D communication framework was proposed in [27] for analyzing the signal to interference plus noise ratio distributions and the average rates. Additionally, the authors of [28] incorporated spectrum sharing and provided a comparative coverage performance analysis of two different sharing modes. Both mode selection and power control were conceived for the uplink of a large-scale D2D-enabled cellular network in [29]. However, these treatises did not analyze the influence of the temporal and geographic fluctuations of mobile traffic. By investigating a wide range of D2D communication operational conditions, including mode selection, varying maximum D2D link distances and user densities, the influence of average user rate on the achievable performance of a large-scale D2D-enabled cellular network is quantified in [30]. However, all the above-mentioned contributions focus on user-centric performance measures, such as coverage probability or achievable system capacity but they do not analyze the ESE of large-scale D2D-enabled cellular networks. Specifically, they all assume a constant power consumption in each BS and ignore the impact of mobile-traffic migration due to the offloading operation on the resource utilization and ESE. As demonstrated in our recent work [31], it is essential to adopt a realistic variable power consumption model for BSs in order to analyze and optimize the ESE of heterogeneous cellular networks.

It is quite a challenge to characterize the relationship between the mobile-traffic migration owing to D2D offloading and resource utilization whilst considering the users' large-scale behaviors. It is even more so, when facilitating offloading in large-scale D2D-enabled cellular networks and aiming for a high network-ESE performance. Hence there is a paucity of literature on this challenging subject.

B. Our Contributions

Against this backcloth, we endeavour to answer the following two fundamental questions: 1. What is the quantitative impact of the users' behavior, the physical cellular operating conditions and the offloading parameters on the network's

ESE? 2. How to optimize the network's ESE by appropriately controlling the volume of geographical mobile-traffic offloaded by D2D communications, whilst satisfying the end-user's QoS requirement and power constraint?

We begin by introducing a D2D-assisted offloading mechanism for large-scale cellular networks, which takes into account both the users' behaviors and the network's operating conditions, such as the D2D offloading radius, the varying geographical mobile-traffic intensity and BS density. Our goal is to determine when and how to offload mobile traffic onto the D2D tier. Based on this offloading mechanism, we then propose an ESE evaluation framework for large-scale D2D-enabled cellular networks, which allows us to analyze the impact of the users' behaviors, the D2D-assisted offloading parameters and the cellular network operating conditions on the achievable network's ESE. Finally, we derive an optimal scenario-aware D2D-enabled offloading solution. Our main contributions are summarized below.

1) *Tractable ESE modeling for D2D-enabled cellular networks*: Based on the offloading mechanism conceived, we develop an ESE evaluation framework for large-scale D2D-enabled cellular networks, which allows a joint analysis of the impacts of user-behaviors, specifically, mobile-traffic intensity and average user rate, as well as of offloading strategy specified by offloading radius, of BS density and other key cellular network parameters. A pair of spectrum sharing modes, namely, the D2D overlay and underlay in-band modes (O mode and U mode), are considered¹. Several analytical results are obtained, including the average transmit power of D2D transmitters and the network-level ESE. Finally, the qualitative impact of mobile-traffic migration between cellular communication and D2D offloading are characterized accurately. The first part of our contributions are valid for large-scale D2D-enabled cellular networks operating either in the U mode or in the O mode.

2) *Energy-spectral-efficient D2D-assisted offloading strategy for O mode*: For O mode, a fundamental operation issue is how the system bandwidth can be partitioned optimally into two orthogonal segments between cellular tier and D2D tier. With the aid of the above-mentioned tractable analytical ESE-tool, we design an optimal D2D mobile-traffic offloading scheme for O-mode D2D enabled cellular networks to maximize the network's ESE under specific cellular user outage performance and D2D transmitter power constraints. By matching the users' behaviors and their cellular operating conditions, this optimal offloading strategy ensures the optimal spectrum sharing between the D2D tier and the cellular tier. An additional advantage of our design is its computational efficiency, since our outage-constrained and power-constrained optimal D2D-enabled offloading solution is readily given in a closed-form.

II. D2D-ENABLED CELLULAR SYSTEM MODEL

We consider a large-scale D2D-enabled cellular network in the downlink (DL), composed of the set of BSs Θ_B , the set

¹In U mode, both cellular tier and D2D tier share the same system bandwidth, while in O mode, they operate in the different system bandwidths orthogonally, i.e., share the system bandwidth orthogonally.

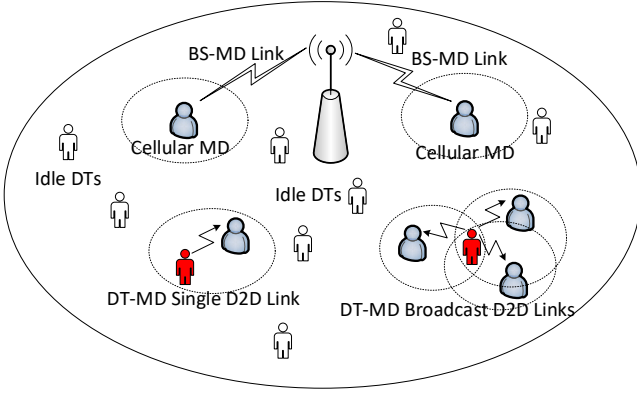


Fig. 1. Illustration of a typical cellular cell in the large-scale D2D-enabled cellular network, including two single BS-MD links, one single DT-MD link and a broadcast DT-MD link. The red DT denotes the activated DT while the white DTs are idle DTs. The dotted circles indicate the offloading decision region with each MD's location as the center.

of MDs Θ_U , which are retrieving their desired contents, and the set of potential D2D transmitters (DTs) Θ_{DT} , which are actually the MDs that currently have no content to retrieve themselves and which have been chosen by the associated BSs to relay other users' contents². Any MD in Θ_U has two service modes: cellular mode and D2D mode. Cellular MDs are associated with their geographically nearest BSs, and D2D MDs can bypass their associated BSs and directly access their DTs in their proximity. DTs are distinguished from MDs in that they either transmit radio signals to the nearby D2D MDs or stay idle. Fig. 1 depicts a typical cellular cell in such a large-scale D2D-enabled cellular network. We model Θ_B , Θ_U and Θ_{DT} with three independent Poisson point processes (PPPs) in the Euclidean plane \mathbb{R}^2 with the intensities λ_B , λ_U and λ_{DT} , respectively. We consider both the large-scale pathloss and small-scale fast fading for each desired or interference signal. More specifically, the large-scale pathloss is modeled by $l = d^{-\alpha}$ with d denoting the distance between the transmitter and receiver and α being the pathloss exponent. Rayleigh-distributed fast-fading is considered. In our analysis, the effect of the noise is ignored, since the interference power is far higher than the noise power in this interference-limited scenario³.

Each MD $u_i \in \Theta_U$ has to make an offloading decision, either retrieving its content in its cellular mode or via its D2D mode. This decision depends on the distribution state of the DTs in its proximity. Let the effective D2D transmission range or radius be ϖ_D , which is defined by

$$\varpi_D^{-\alpha} = \nu \quad [\%], \quad (1)$$

i.e., it is defined as the distance at which the pathloss causes the signal power to drop to ν %. Then, u_i has an offloading

²In this work, we assume that the set of DTs Θ_{DT} has been selected. How to select DTs or D2D relays is a very complicated problem. This is because many factors, including instantaneous link conditions, spectrum sharing mode, resource allocation and interference mitigation technique adopted, must be taken into consideration. Moreover, mobility patterns and social network features can be utilized to select better D2D relays [5], [22], [32]–[34].

³It is well known that in this circumstance, the noise power only has a very marginal effect on the system coverage probability [35] and almost has no influence on the network EE [36].

region $\mathcal{O}(u_i, \varpi_D)$, which is the disc with u_i as the center and the radius ϖ_D . Let $N_{\mathcal{O}_{u_i}}$ denote the number of DTs located inside $\mathcal{O}(u_i, \varpi_D)$. Since the objective is to maximally offload mobile traffic to the D2D tier while maintaining certain user QoS requirements, u_i is in its D2D mode, whenever $N_{\mathcal{O}_{u_i}} \geq 1$, and it has to opt for its cellular mode if $N_{\mathcal{O}_{u_i}} = 0$. Furthermore, a MD u_i with D2D mode always choose the nearest DT inside $\mathcal{O}(u_i, \varpi_D)$ to retrieve its content. Thus, we can define the mobile-traffic offloading probability as $P_{OL} = \Pr(N_{\mathcal{O}_{u_i}} \geq 1)$, which effectively divides all the MDs into the subset of cellular MDs Θ_U^C and the subset of D2D MDs Θ_U^D . We further assume that each DT can serve multiple MDs by broadcasting the offloaded traffic, and the activation of other idle DT will not prevent it from receiving the offloading requests from its proximate potential MDs. This assumption implies that there is no correlation in establishing different DT-MD links. Consequently, Θ_U^C and Θ_U^D follow the two independent thinned PPPs with the densities of $(1 - P_{OL})\lambda_U$ and $P_{OL}\lambda_U$, respectively.

For a typical BS $b_j \in \Theta_B$, we define its service or coverage area by $S_{A,j}$. The maximum number of cellular MDs supported in each cell is N_C . Furthermore, we denote the set of all the MDs under the coverage area $S_{A,j}$ by Θ_U^j . Clearly, $\Theta_U^j \cap \Theta_U^{j'} = \emptyset$ for $j \neq j'$ and $\cup_{b_j \in \Theta_B} \Theta_U^j = \Theta_U$. The MDs of Θ_U^j can be divided into two groups, cellular MDs, collected in the subset $\Theta_U^{C,j} = \{u_i^{C,j}\}_{i=1}^{N_C}$, and D2D MDs, collected in the subset $\Theta_U^{D,j}$. Obviously, $\Theta_U^{C,j} \cup \Theta_U^{D,j} = \Theta_U^j$. The system bandwidth of W [Hz] is reused spatially in every cell.

For the U mode, both the cellular and D2D tiers share the same system bandwidth of W [Hz] in each cell. The total bandwidth of W [Hz] is divided into the N_C subbands to be allocated to the MDs in $\Theta_U^{C,j}$. The same system bandwidth of W [Hz] is divided into $\beta \in \mathbb{N}^+$ subbands in the D2D tier for the D2D pairs to access⁴. Each D2D pair in $\Theta_U^{D,j}$, which is associated with BS b_j , can randomly and independently access one subband to reduce the interference imposed on the other links in the same cell. Note that the probability to access a subband by a D2D pair is $1/\beta$. Therefore, $1/\beta$ can be considered as the access probability of D2D pairs, which measures the intensity of D2D bandwidth access. Since D2D link should supports higher-rate than cellular link, the subband allocated to a cellular MD is narrower than that accessed by a D2D MD, and we have $N_C > \beta$.

For the O mode, the total bandwidth is divided into two orthogonal segments. The fraction of $(1 - \omega)W$ is allocated to the cellular tier, and the other part of ωW is used for the D2D tier, where $0 < \omega < 1$. Specifically, the cellular-tier bandwidth $(1 - \omega)W$ is divided into the N_C subbands for the MDs of $\Theta_U^{C,j}$, while the D2D-tier bandwidth ωW is divided into the β subbands for the D2D pairs of $\Theta_U^{D,j}$. Similarly, the subband allocated to a cellular MD is typically narrower than that for a D2D MD.

First, let us consider a typical BS-MD link between b_j and the i -th cellular MD $u_i^{C,j} \in \Theta_U^{C,j}$, where $1 \leq i \leq N_C$. The

⁴Each D2D MD $u_i^{D,j} \in \Theta_U^{D,j}$ is obviously associated with a DT or DT-MD link. For convenience, we refer to $u_i^{D,j}$ and its associated DT as a D2D pair in $\Theta_U^{D,j}$.

link service rate satisfies

$$R_i^{C,j} = W_C \log_2 \left(1 + \frac{P_i^{C,j} l_i^{C,j} h_i^{C,j}}{I_j^{CC} + I_j^{DC}} \right), \quad (2)$$

where $W_C = W/N_C$ for the U mode and $W_C = (1 - \omega)W/N_C$ for the O mode, $P_i^{C,j} \leq P_C/N_C$ is the required transmit power allocated to the subband of this cellular link with P_C denoting the maximum transmission power of the BS, while $l_i^{C,j} = \|u_i^{C,j} - b_j\|^{-\alpha}$ is the pathloss and $h_i^{C,j}$ denotes the fast-fading power gain of the channel between b_j and $u_i^{C,j}$, which is exponentially distributed with a unity mean, i.e., $h_i^{C,j} \sim \exp(1)$. Furthermore, we consider the worst-case interference, where all interfering BSs transmit at the maximum power P_C/N_C at the subband. Hence, for both the U mode and O mode, the interference coming from all the other BSs' transmissions is given by

$$I_j^{CC} = \sum_{b_m \in \{\Theta_B \setminus b_j\}} \frac{P_C}{N_C} h_i^{C,m} l_i^{C,m}. \quad (3)$$

Similarly, let us assume that all the interfering DTs transmit at the maximum power of P_D/β at a subband, where P_D is the maximum transmission power of DT. Then for the U mode, the worst-case interference coming from all the interfering DTs can be expressed as

$$I_j^{DC} = \sum_{DT_n \in \mathbb{O}(u_i^{C,j}, \varpi_D), (DT_n \wedge u_i^{C,j})} \frac{P_D}{N_C} h_i^{D,n} l_i^{D,n}, \quad (4)$$

where P_D/N_C is the power of the interfering D2D transmissions inside the cellular subband for $u_i^{C,j}$, $l_i^{D,n} = \|u_i^{C,j} - DT_n\|^{-\alpha}$ and $h_i^{D,n}$ are respectively the pathloss and fast fading power gain of the link between DT_n and $u_i^{C,j}$, while $\mathbb{O}(u_i^{C,j}, \varpi_D)$ denotes the exclusive region to the offloading region $\mathbb{O}(u_i^{C,j}, \varpi_D)$ and $(DT_n \wedge u_i^{C,j})$ indicates that the D2D subband for DT_n overlaps with the cellular subband for $u_i^{C,j}$. Note that in the cellular mode, there exists no DT in $\mathbb{O}(u_i^{C,j}, \varpi_D)$. For the O mode, obviously,

$$I_j^{DC} = 0. \quad (5)$$

Because the worst-case interference is considered, the service rate of (2) represents a low bound of the achievable rate.

Second, for the k -th D2D MD $u_k^{D,j}$ located inside $S_{A,j}$, let $DT_n \in \mathbb{O}(u_k^{D,j}, \varpi_D)$ be the nearest DT. Then the service rate of the link between DT_n and $u_k^{D,j}$ satisfies

$$R_k^{D,n} = W_D \log_2 \left(1 + \frac{P_k^{D,n} l_k^{D,n} h_k^{D,n}}{I_n^{CD} + I_n^{DD}} \right), \quad (6)$$

where $W_D = W/\beta$ for the U mode and $W_D = \omega W/\beta$ for the O mode, while $P_k^{D,n} \leq P_D/\beta$ is the required transmit power allocated to the subband of this D2D link. In (6), $l_k^{D,n}$ and $h_k^{D,n}$ are the pathloss and fast fading power gain of the link between DT_n and $u_k^{D,j}$, respectively. For the U mode, the worst-case interference coming from all the cellular transmissions is given by

$$I_n^{CD} = \sum_{b_m \in \Theta_B} \frac{P_C}{\beta} h_k^{C,m} l_k^{C,m}, \quad (7)$$

where P_C/β is the power of the interfering cellular transmissions inside the D2D subband for $u_k^{D,j}$, while $l_k^{C,m}$ and $h_k^{C,m}$ are the pathloss and fast fading power gain of the link between b_m and $u_k^{D,j}$, respectively. For the O mode, clearly,

$$I_n^{CD} = 0. \quad (8)$$

Furthermore, denote $l_k^{D,n'}$ and $h_k^{D,n'}$ as the pathloss and fast fading power gain of the link between $DT_{n'}$ and $u_k^{D,j}$, respectively. Then, for both the U mode and O mode, the worst-case interference coming from all the other D2D transmissions can be expressed as

$$I_n^{DD} = \sum_{n' \neq n, (DT_{n'} \wedge DT_n)} \frac{P_D}{\beta} h_k^{D,n'} l_k^{D,n'}, \quad (9)$$

where $(DT_{n'} \wedge DT_n)$ indicates that $DT_{n'}$ and DT_n are allocated with the same D2D subband. Again, the service rate of (6) represents a low bound of the achievable rate.

Table I lists the main notations used in this paper.

III. ESE EVALUATION OF LARGE-SCALE MOBILE-TRAFFIC-AWARE D2D-ENABLED OFFLOADING

In this section, we first model the large-scale mobile-traffic-aware D2D-assisted offloading by appropriately configuring the offloading region at each MD in both the U and O modes, and then quantify the overall ESE performance based on the network's resource utilization in both the D2D and cellular tiers for a large-scale D2D-enabled cellular network.

TABLE I
LIST OF MAIN NOTATIONS

| Notation | Definition |
|--------------------------------------|--|
| $\Theta_B, \Theta_U, \Theta_{DT}$ | Sets of BSs, MDs, DTs |
| $\lambda_B, \lambda_U, \lambda_{DT}$ | BS density, MD density, DT density |
| ϖ_D | Offloading radius |
| P_{OL} | Offloading probability |
| $S_{A,j}$ | Coverage area of typical BS b_j |
| Θ_U^j | Set of MDs covered by b_j |
| $\Theta_U^{C,j}, \Theta_U^{D,j}$ | Sets of cellular MDs, D2D MDs covered by b_j |
| N_C | Maximum number of cellular MDs served by b_j |
| W | Total system bandwidth in Hz |
| α | Pathloss exponent |
| $1/\beta$ | D2D access probability |
| ω | Faction of system bandwidth for O mode D2D tier |
| $R_i^{C,j}$ | Service rate of typical BS-MD link |
| $R_k^{D,n}$ | Service rate of typical DT-MD link |
| R_C | Average service rate of typical BS-MD link |
| R_{D2D} | Average service rate of typical DT-MD link |
| P_C, P_D | Maximum transmit powers of BSs and DTs |
| $l_i^{C,j}, h_i^{C,j}$ | Pathloss and fading gain between b_j and $u_i^{C,j}$ |
| $l_k^{D,n}, h_k^{D,n}$ | Pathloss and fading gain between DT_n and $u_k^{D,j}$ |
| $P_i^{C,j}$ | Transmit power for the link between b_j and $u_i^{C,j}$ |
| $P_k^{D,n}$ | Transmit power for the link between DT_n and $u_k^{D,j}$ |
| $\eta_{ESE}^O, \eta_{ESE}^U$ | Network ESEs in O and U modes |
| a_E, b_E | Efficiency factors of power amplifiers at BS and DT |
| \mathcal{K} | DT active probability |
| $P_{D_a}^O, P_{D_a}^U$ | Average D2D transmit powers in O and U modes |
| P_{C1}^O, P_{C1}^U | Average transmit powers of b_j to $u_i^{C,j}$ in two modes |
| P_{BS}^O, P_{BS}^U | Aggregate transmit powers of BS b_j in two modes |

$$Z_U^D \geq P_{D_a}^U \triangleq \left(2^{\frac{\beta R_{D2D}}{W}} - 1\right) \left(\frac{2\pi\mathcal{K}P_D\lambda_U P_{OL}}{(\alpha-2)\beta} \left(\frac{P_{OL}}{\pi\lambda_{DT}} - \frac{\varpi_D^2}{\exp(\pi\lambda_{DT}\varpi_D^2)} \right) + \frac{2\lambda_B^{\frac{\alpha}{2}} P_C}{(\alpha-2)\beta\lambda_{DT}^{\frac{\alpha}{2}}} \Gamma\left(2-\frac{\alpha}{2}\right) \gamma\left(\frac{\alpha+2}{2}, \pi\lambda_{DT}\varpi_D^2\right) \right), \quad (10)$$

A. Offloading Region for D2D-Enabled Offloading

Modeling offloading region is the first step in the ESE performance evaluation of a large-scale D2D-enabled offloading system. Therefore, we firstly characterize the effective D2D transmission radius by considering the requirement for the transmit power capability of D2D transmitter, measured in watt. Hence, the following two propositions are given.

Proposition 1: In a large-scale D2D-enabled U-mode based cellular network having a BS density of λ_B , a mobile-traffic intensity of λ_U and a DT density of λ_{DT} , the transmit power capability of D2D transmitters, denoted by Z_U^D [watt], the average D2D user rate R_{D2D} , the D2D access probability $1/\beta$, and the effective D2D transmission radius ϖ_D , which defines the offloading region $\mathbb{O}(u_i, \varpi_D)$ in the U mode, satisfy the inequality given in (10) at the top of the page, where $\alpha > 2$, P_{OL} is the offloading probability given by

$$P_{OL} = \Pr(N_{\mathbb{O}_{u_i}} \geq 1) = 1 - \exp(-\pi\lambda_{DT}\varpi_D^2), \quad (11)$$

$\gamma(a, x)$ is the incomplete Gamma function expressed as

$$\gamma(a, x) = \int_0^x t^{a-1} \exp(-t) dt, \quad (12)$$

and $\Gamma(x) = \int_0^{+\infty} t^{x-1} \exp(-t) dt$, while the DT active probability \mathcal{K} , a function of λ_U and ϖ_D , is given by

$$\mathcal{K} = (\text{Ei}(y) - \text{Ei}(1) - \ln y + 1.3179) \exp(-y) \leq 1, \quad (13)$$

with $y = \pi\lambda_U\varpi_D^2$ and $\text{Ei}(y)$ being the exponential integral function defined by

$$\text{Ei}(y) = - \int_{-y}^{\infty} \frac{\exp(-t)}{t} dt. \quad (14)$$

Clearly, the maximum tolerable D2D transmission radius $\varpi_{D_{\max}}^U$ corresponding to Z_U^D can be numerically obtained by setting $\varpi_D = \varpi_{D_{\max}}^U$ and replacing ‘ \geq ’ with ‘ $=$ ’ in (10).

Proof: See Appendix A. ■

Remark 1: The right-hand side of the inequality (10) defines the D2D transmit power requirement in the U mode.

- 1) Since increasing β leads to exponential increase of $P_{D_a}^U$, the D2D transmit power capability is a monotonically increasing function of β . Thus, using narrower subbands in the D2D tier reduces $P_{D_a}^U$ and hence Z_U^D .
- 2) From (10), $P_{D_a}^U$ is a monotonically increasing function of λ_{DT} . This is because increasing the DT density λ_{DT} directly increases the interference because of more active D2D pairs.
- 3) According to (10), to increase the effective D2D radius ϖ_D requires a corresponding increase in the D2D transmit power capability Z_U^D , which also makes sense.

Proposition 2: In a large-scale D2D-enabled O-mode based cellular network having the total bandwidth ωW in the D2D tier, a mobile-traffic intensity of λ_U and a DT density of λ_{DT} , the D2D transmit power capability Z_O^D [watt], the average user rate R_{D2D} , the D2D access probability $1/\beta$, and the

effective D2D transmission radius ϖ_D , which characterizes the offloading region $\mathbb{O}(u_i, \varpi_D)$ in the O mode, satisfy the following inequality

$$Z_O^D \geq P_{D_a}^O \triangleq \left(2^{\frac{\beta R_{D2D}}{\omega W}} - 1\right) \frac{2\pi\mathcal{K}P_D\lambda_U P_{OL}}{\beta(\alpha-2)} \left(\frac{P_{OL}}{\pi\lambda_{DT}} - \frac{\varpi_D^2}{\exp(\pi\lambda_{DT}\varpi_D^2)} \right), \quad (15)$$

where $\alpha > 2$. Furthermore, the maximum tolerable D2D transmission radius $\varpi_{D_{\max}}^O$ related to Z_O^D can be numerically obtained by setting $\varpi_D = \varpi_{D_{\max}}^O$ and replacing ‘ \geq ’ with ‘ $=$ ’ in (15).

Proof: See Appendix B. ■

Remark 2: Different from the U mode, the D2D transmit power capability Z_O^D in the O mode is independent of λ_B , since the cross-tier interference is eliminated in an orthogonal spectrum sharing. Similar to the case of U mode, to increase the effective D2D transmission radius ϖ_D in the O mode requires a corresponding increase in the D2D transmit power capability Z_O^D .

B. Downlink Transmit Power Analysis

Next we derive the average DL transmit power for a typical BS-MD link and the BS’s aggregate DL transmit power in both the U and O modes.

Proposition 3: In a large-scale D2D-enabled cellular network having a BS density of λ_B with the average cellular user rate R_C , a mobile-traffic intensity of λ_U and a DT density of λ_{DT} with the D2D offloading radius ϖ_D , the average DL transmit power assigned in W_C for a typical BS-MD link in the U mode is given by

$$P_{C1}^U = \frac{\left(2^{\frac{R_C}{W_C}} - 1\right) 2\lambda_B}{(\alpha-2) \exp(-\pi\lambda_{DT}\varpi_D^2) \lambda_U} \left(P_C + \frac{\alpha \Gamma\left(\frac{\alpha}{2}\right) \pi^{\frac{2-\alpha}{2}} P_D \lambda_U \mathcal{K} P_{OL}}{2\lambda_B^{\frac{\alpha}{2}} \varpi_D^{\alpha-2}} \right), \quad (16)$$

where $\alpha > 2$. Furthermore, assuming that all the cellular MDs have an identical user rate of R_C , the BS’s aggregate DL transmit power in the U mode is given as

$$P_{BS}^U = \frac{\left(2P_C\lambda_B^{\frac{\alpha}{2}}\varpi_D^{\alpha-2} + \alpha\Gamma\left(\frac{\alpha}{2}\right)\pi^{\frac{2-\alpha}{2}}P_D\lambda_U\mathcal{K}P_{OL}\right) K^K}{(\alpha-2)\left(K - \left(2^{\frac{R_C}{W_C}} - 1\right)(1-P_{OL})\frac{\lambda_U}{\lambda_B}\right)^K \lambda_B^{\frac{\alpha}{2}}\varpi_D^{\alpha-2}} - \frac{2P_C}{(\alpha-2)} - \frac{\alpha\Gamma\left(\frac{\alpha}{2}\right)\pi^{\frac{2-\alpha}{2}}P_D\lambda_U\mathcal{K}P_{OL}}{(\alpha-2)\lambda_B^{\frac{\alpha}{2}}\varpi_D^{\alpha-2}}, \quad (17)$$

where $K = 3.75$.

Proof: See Appendix C. ■

Remark 3: The influence of λ_{DT} to P_{C1}^U is complicated. Increasing λ_{DT} decreases $\exp(-\pi\lambda_{DT}\varpi_D^2)$ but increases P_{OL} .

From (16), it is seen that P_{C1}^U contains both $\exp(-\pi\lambda_{DT}\varpi_D^2)$ and P_{OL} . However, increasing λ_{DT} also decreases W_C , since the average of W_C is $E[W_C] \approx W\lambda_B(1 - P_{OL})/\lambda_U = W\lambda_B \exp(-\pi\lambda_{DT}\varpi_D^2)/\lambda_U$, which increases P_{C1}^U . Therefore, the overall effect of increasing λ_{DT} is to increase P_{C1}^U . The influence of ϖ_D to P_{C1}^U is also complicated. Increasing ϖ_D decreases $\exp(-\pi\lambda_{DT}\varpi_D^2)$ but increases P_{OL} . However, P_{C1}^U contains $\varpi_D^{2-\alpha}$ with $\alpha > 2$, which decreases as ϖ_D increases. Hence, overall, increasing ϖ_D reduces P_{C1}^U . Moreover, P_{C1}^U contains the factor of $\frac{1}{\lambda_U}$, but increasing $\frac{1}{\lambda_U}$ can increase W_C which will reduce P_{C1}^U . Therefore, the overall effect of increasing $\frac{1}{\lambda_U}$ is to decrease P_{C1}^U .

Proposition 4: In a large-scale D2D-enabled cellular network having a BS density of λ_B with the average cellular user rate R_C , a mobile-traffic intensity of λ_U and a DT density λ_{DT} with the D2D offloading radius ϖ_D , given the total bandwidth in cellular tier $(1 - \omega)W$, the average DL transmit power assigned in W_C for a typical BS-MD link in the O mode is given by

$$P_{C1}^O = \frac{2P_C\lambda_B \left(2^{\frac{R_C}{W_C}} - 1\right)}{(\alpha - 2) \exp(-\pi\lambda_{DT}\varpi_D^2) \lambda_U}. \quad (18)$$

Furthermore, assuming that all the cellular MDs have an identical service rate of R_C , the BS's aggregate DL transmit power in the O mode is given by

$$P_{BS}^O = \frac{2P_C K^K \lambda_B^K (\exp(\pi\lambda_{DT}\varpi_D^2))^K}{(\alpha - 2) \left(K\lambda_B \exp(\pi\lambda_{DT}\varpi_D^2) - \left(2^{\frac{R_C}{(1-\omega)W}} - 1\right) \lambda_U\right)^K} - \frac{2P_C}{(\alpha - 2)}. \quad (19)$$

Proof: The proof is similar to the proof of Proposition 3. In fact, the proof is simpler, as the cross-tier interference $I_j^{DC} = 0$ in the O mode. ■

C. Network ESE Performance Evaluation

We now quantify the network's ESE. First, the network's area power consumptions (APCs) from all the BS-MD links in the U and O modes are given respectively by

$$C_{BS}^U = a_E P_{BS}^U \lambda_B + P_{OM} \lambda_B \text{ [watt/m}^2\text{]}, \quad (20)$$

$$C_{BS}^O = a_E P_{BS}^O \lambda_B + P_{OM} \lambda_B \text{ [watt/m}^2\text{]}, \quad (21)$$

where a_E is the efficiency factor of BS power amplifier and P_{OM} is the static operation power consumption of a BS. Second, the APCs from all the D2D links in the U and O modes are given respectively by

$$C_{D2D}^U = b_E \lambda_{DT}^{\text{act}} P_{Da}^U \text{ [watt/m}^2\text{]}, \quad (22)$$

$$C_{D2D}^O = b_E \lambda_{DT}^{\text{act}} P_{Da}^O \text{ [watt/m}^2\text{]}, \quad (23)$$

where b_E is the efficiency factor of DT power amplifier, and $\lambda_{DT}^{\text{act}} = K\lambda_U P_{OL}$ is the active DT density (see Appendix A). Note that in (20) to (23), the receiver signal processing power consumption is neglected⁵. By using R_C and R_{D2D} to denote

the average user rates in cellular and D2D tiers, respectively, the network's area spectral efficiency (ASE) is given by

$$S = \frac{\lambda_U}{W} ((1 - P_{OL})R_C + P_{OL}R_{D2D}) \text{ [bit/s/Hz/m}^2\text{]}. \quad (24)$$

We are ready to present the ESE performance of the large-scale D2D-enabled cellular network operating in the U and O modes in the following two propositions, respectively.

Proposition 5: For the large-scale D2D-enabled cellular network operating in the U mode with the cellular and D2D spatially average user rates of R_C and R_{D2D} , respectively, as well as having a BS density of λ_B , a mobile-traffic intensity of λ_U , a DT density λ_{DT} with the D2D offloading radius ϖ_D , the BS static operation power consumption P_{OM} , and the BS and DT power amplifier efficiencies of a_E and b_E , respectively, the network's ESE is expressed as

$$\eta_{ESE}^U = \frac{S}{C_{BS}^U + C_{D2D}^U} = \frac{(1 - P_{OL})R_C + P_{OL}R_{D2D}}{W(X^U + Y^U)}, \quad (25)$$

where

$$X^U = \frac{1}{\lambda_U} (a_E P_{BS}^U + P_{OM}) \lambda_B, \quad (26)$$

$$Y^U = \frac{1}{\lambda_U} b_E \lambda_{DT}^{\text{act}} P_{Da}^U. \quad (27)$$

Proposition 6: For the large-scale D2D-enabled cellular network operating in the O mode with the total bandwidth in cellular tier $(1 - \omega)W$, the cellular and D2D spatially average user rates of R_C and R_{D2D} , respectively, as well as having a BS density of λ_B , a mobile-traffic intensity of λ_U , a DT density λ_{DT} with the D2D offloading radius ϖ_D , the BS static operation power consumption P_{OM} , and the BS and DT power amplifier efficiencies of a_E and b_E , respectively, the network's ESE is expressed as

$$\eta_{ESE}^O = \frac{S}{C_{BS}^O + C_{D2D}^O} = \frac{(1 - P_{OL})R_C + P_{OL}R_{D2D}}{W(X^O + Y^O)}, \quad (28)$$

where

$$X^O = \frac{1}{\lambda_U} (a_E P_{BS}^O + P_{OM}) \lambda_B, \quad (29)$$

$$Y^O = \frac{1}{\lambda_U} b_E \lambda_{DT}^{\text{act}} P_{Da}^O. \quad (30)$$

IV. ESE OPTIMIZATION FOR MOBILE-TRAFFIC-AWARE D2D-ENABLED CELLULAR NETWORKS

In this section, we first evaluate the cellular user outage performance in large-scale mobile-traffic-aware D2D-enabled cellular networks for both U and O mods, and then propose a spectrum partitioning scheme for the O mode to maximize the achievable network's ESE under both the cellular-user outage constraint and D2D transmit power constraint.

⁵According to [37], reception generally consumes less power than transmission.

A. Outage Performance

The outage performance of a typical BS-MD link with the coexist of D2D links is presented in the following two propositions for the U and O modes, respectively.

Proposition 7: In a large-scale D2D-enabled cellular network operating in the U mode, given an effective D2D transmission radius of ϖ_D , the outage probability for a typical BS-MD link is given by

$$Q_U = 1 - \exp \left(- \left(\rho_1 + p_c \frac{\lambda_{DT}^{\text{act}}}{\lambda_B} \rho_2 \right) \right), \quad (31)$$

with $p_c = (P_D/P_C)^{\frac{1}{\alpha}}$,

$$\rho_1 = \gamma^{\frac{2}{\alpha}} \int_{\gamma^{-\frac{2}{\alpha}}}^{+\infty} \frac{1}{1+u^{\frac{\alpha}{2}}} du, \quad (32)$$

$$\rho_2 = \gamma^{\frac{2}{\alpha}} \int_{\left(\sqrt{\pi \lambda_B \varpi_D} / p_c \gamma^{1/\alpha}\right)^2}^{+\infty} \frac{1}{1+u^{\frac{\alpha}{2}}} du, \quad (33)$$

$$\gamma = 2^{\frac{R_C}{W_C}} - 1. \quad (34)$$

Proof: See Appendix D. ■

Utilizing $E[W_C] = W \lambda_B / \lambda_U (1 - P_{OL})$, we define the average cellular outage probability in the U mode as

$$\tilde{Q}_U(\varpi_D) = 1 - \exp \left(- \left(\tilde{\rho}_1 + p_c \frac{\lambda_{DT}^{\text{act}}}{\lambda_B} \tilde{\rho}_2 \right) \right), \quad (35)$$

where $\tilde{\rho}_1$ and $\tilde{\rho}_2$ are obtained respectively by replacing γ with $\tilde{\gamma} = 2^{R_C \lambda_U (1-P_{OL}) / W \lambda_B} - 1$ in (33) and (34). The average cellular outage probability is a function of ϖ_D which is explicitly indicated in (35). Note that both $\tilde{\rho}_1$ and $\tilde{\rho}_2$ increase as $\tilde{\gamma}$ increases, while $\tilde{\gamma}$ is a monotonically decreasing function of ϖ_D since $(1 - P_{OL}) = \exp(-\pi \lambda_{DT} \varpi_D^2)$. This indicates that $\tilde{Q}_U(\varpi_D)$ is a monotonically decreasing function of ϖ_D .

By setting $\rho_2 = 0$ in (31), we readily obtain the following result for the O mode.

Proposition 8: In a large-scale D2D-enabled cellular network operating in the O mode with the spectrum partitioning factor ω , given an effective D2D transmission radius of ϖ_D , the outage probability for a typical BS-MD link is given by

$$Q_O = 1 - \exp(-\rho_1). \quad (36)$$

In the O mode, since $E[W_C] = (1-\omega)W \lambda_B / \lambda_U (1 - P_{OL})$, we define the average cellular outage probability as

$$\tilde{Q}_O(\omega, \varpi_D) = 1 - \exp(-\tilde{\rho}_1'), \quad (37)$$

where

$$\tilde{\rho}_1' = \tilde{\gamma}'^{\frac{2}{\alpha}} \int_{\tilde{\gamma}'^{-\frac{2}{\alpha}}}^{+\infty} \frac{1}{1+u^{\frac{\alpha}{2}}} du, \quad (38)$$

with $\tilde{\gamma}' = 2^{R_C \lambda_U (1-P_{OL}) / ((1-\omega)W \lambda_B)} - 1$. Clearly, \tilde{Q}_O is a monotonically increasing function of ω , and it is a monotonically decreasing function of ϖ_D . When investigating the

influence of the spectral partitioning factor ω for given ϖ_D , we will simply denote the average cellular outage probability in the O mode as $\tilde{Q}_O(\omega)$.

B. Optimal Cross-tier Spectrum Partitioning in O Mode

For the O model, the system bandwidth is partitioned orthogonally into the two segments of $(1-\omega)W$ and ωW for the cellular and D2D tiers, respectively. We propose a cross-tier spectrum partitioning scheme in the O mode to maximize the achievable network's ESE under both the cellular-user outage constraint and D2D transmit power constraint. Mathematically, this optimal spectrum partitioning can be formulated as the following optimization problem

$$\begin{aligned} \bar{\omega}^* &= \arg \max_{0 < \omega < 1} \eta_{\text{ESE}}^O, \\ \text{s.t. } &\tilde{Q}_O(\omega) < \theta_O, \\ &P_{D_a}^O(\omega) < Z_D^O, \end{aligned} \quad (39)$$

where θ_O is the cellular user outage threshold, and we have explicitly indicated that the average D2D transmit power $P_{D_a}^O$ is a function of ω .

Since increasing the spectrum portion of the D2D tier decreases the requirement of D2D transmit power, $P_{D_a}^O(\omega)$ is a monotonically decreasing function of ω . But $\tilde{Q}_O(\omega)$ is a monotonically increasing function of ω . Therefore, the constrained feasible region for ω that satisfies both the cellular-user outage constraint and D2D transmit power constraint is $\omega \in [\omega_{D_a}, \omega_{\text{out}}]$, where ω_{D_a} and ω_{out} can be numerically obtained from the following two equations respectively

$$P_{D_a}^O(\omega_{D_a}) = Z_D^O, \quad (40)$$

$$\tilde{Q}_O(\omega_{\text{out}}) = \theta_O. \quad (41)$$

Clearly, if $\omega_{D_a} > \omega_{\text{out}}$, the feasible region of ω is empty, and there exists no solution for the constrained optimization (39).

Proposition 9: Given the effective D2D transmission range of ϖ_D , there exists a unique ω^* that maximizes the network's ESE and it can be obtained numerically from (42), given at the bottom of the page.

Proof: The derivative of $\eta_{\text{ESE}}^{\text{OL}}$ with respect to ω is

$$\frac{\partial \eta_{\text{ESE}}^O}{\partial \omega} = - \left(\frac{\partial X^O}{\partial \omega} + \frac{\partial Y^O}{\partial \omega} \right) \frac{(1-P_{OL}) R_C + P_{OL} R_{D2D}}{W (X^O + Y^O)^2}, \quad (43)$$

in which

$$\frac{\partial X^O}{\partial \omega} = \frac{\frac{2(\ln 2) a_E P_C (1-P_{OL}) R_C 2^{\frac{R_C}{(1-\omega)W}}}{(\alpha-2)(1-\omega)^2 W}}{\left(1 - \frac{\lambda_U (1-P_{OL}) \left(2^{\frac{R_C}{(1-\omega)W}} - 1 \right)}{K \lambda_B} \right)^{K+1}}, \quad (44)$$

$$\frac{\omega^2 a_E P_C (1-P_{OL}) R_C 2^{\frac{R_C}{(1-\omega)W}}}{(1-\omega)^2 2^{\frac{R_{D2D}}{\omega W}} \pi b_E \lambda_{DT}^{\text{act}} K P_D P_{OL} R_{D2D}} = \left(1 - \frac{\lambda_U (1-P_{OL}) \left(2^{\frac{R_C}{(1-\omega)W}} - 1 \right)}{K \lambda_B} \right)^{K+1} \left(\frac{P_{OL}}{\pi \lambda_{DT}} - \frac{\varpi_D^2}{\exp(\pi \lambda_{DT} \varpi_D^2)} \right). \quad (42)$$

$$\frac{\partial Y^O}{\partial \omega} = \left(\frac{\varpi_D^2}{\exp(\pi \lambda_{DT} \varpi_D^2)} - \frac{P_{OL}}{\pi \lambda_{DT}} \right) \times \frac{2(\ln 2) \pi b_E 2^{-\frac{\beta R_{D2D}}{\omega W}} \lambda_{DT}^{\text{act}} \mathcal{K} P_D P_{OL} R_{D2D}}{(\alpha - 2) \omega^2 W}. \quad (45)$$

Clearly, $\frac{\partial \eta_{\text{ESE}}^O}{\partial \omega} = 0$ is equivalent to $\frac{\partial X^O}{\partial \omega} + \frac{\partial Y^O}{\partial \omega} = 0$. Furthermore, it can be proved⁶ that $\frac{\partial X^O}{\partial \omega} \in [0, +\infty)$ is a monotonically increasing function of ω and $\frac{\partial Y^O}{\partial \omega} \in [0, +\infty)$ is a monotonically decreasing function of ω since $\frac{\partial^2 X^O}{\partial \omega^2} > 0$ and $\frac{\partial^2 Y^O}{\partial \omega^2} < 0$. Therefore, there must exist a unique ω^* which makes $\frac{\partial X^O}{\partial \omega} + \frac{\partial Y^O}{\partial \omega} = 0$. ■

We are ready to present the solution $\bar{\omega}^*$.

Theorem 1: Let ω^* be the unconstrained optimal spectrum partitioning solution obtained by solving (42). Assume that the feasible region of ω for the constrained optimization (39) is not empty, i.e., $\omega_{D_a} \leq \omega_{\text{out}}$. Then the optimal solution $\bar{\omega}^*$ for the constrained optimization (39) is given as follows.

Case 1. $\omega^* < \omega_{D_a}$: $\bar{\omega}^* = \omega_{D_a}$.

Case 2. $\omega_{D_a} \leq \omega^* \leq \omega_{\text{out}}$: $\bar{\omega}^* = \omega^*$.

Case 3. $\omega^* > \omega_{\text{out}}$: $\bar{\omega}^* = \omega_{\text{out}}$.

V. MOBILE-TRAFFIC-AWARE D2D-ENABLED OFFLOADING PERFORMANCE EVALUATION

We use numerical (theoretical) and simulation results to evaluate the ESE performance for a large-scale D2D-enabled cellular network. In particular, we investigate the influence of the key parameters, including the effective D2D transmission radius ϖ_D and the number of subbands in the D2D tier β , on the achievable network's ESE. Unless otherwise stated, the network has the total bandwidth $W = 20$ MHz, the maximum BS transmit power $P_C = 20$ watts, the maximum DT transmit power $P_D = 1$ watt, the BS static power consumption $P_{OM} = 20$ watts, the BS and DT power amplifier efficiencies $a_E = b_E = 1$, and the DT density $\lambda_{DT} = 150$ DTs/km². The pathloss exponent is set to $\alpha = 3.8$. The default network parameters are listed in Table II.

TABLE II
DEFAULT NETWORK PARAMETERS

| parameter | value |
|----------------|-------------------------|
| λ_B | 5 BSs/km ² |
| λ_{DT} | 150 DTs/km ² |
| λ_U | 200 MDs/km ² |
| ϖ_D | 60 m |
| β | 20 |
| W | 20 MHz |
| R_C | 0.2 Mb/s |
| R_{D2D} | 2 Mb/s |
| P_C | 20 watts |
| P_{OM} | 20 watts |
| P_D | 1 watts |
| $a_E = b_E$ | 1 |
| α | 3.8 |

A. Impact of D2D Transmission Radius

The D2D transmission radius ϖ_D is one of the most important network parameters that influence the achievable

⁶The proofs of $\frac{\partial^2 X^O}{\partial \omega^2} > 0$ and $\frac{\partial^2 Y^O}{\partial \omega^2} < 0$ are not difficult but are too tedious and hence are omitted here.

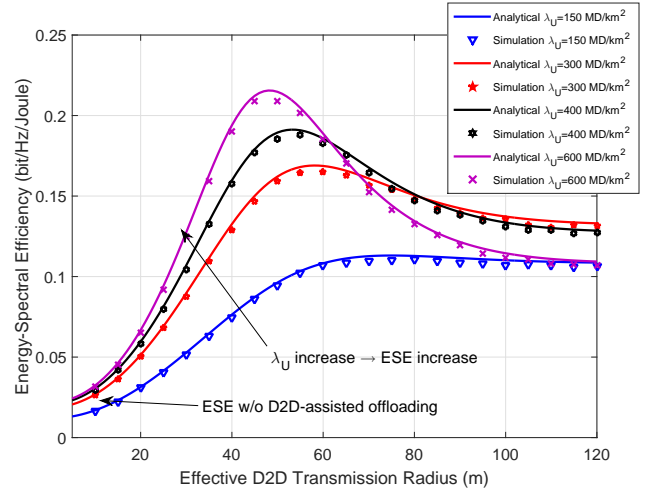


Fig 2. Achievable network ESE performance (curve: analytical and point: simulation) in U mode by varying ϖ_D and λ_U , given $\lambda_B = 5$ BSs/km², $R_C = 0.2$ Mb/s, $R_{D2D} = 2$ Mb/s and $\beta = 20$.

network's ESE. The network's ESE performance in the U and O modes, η_{ESE}^U and η_{ESE}^O , are depicted in Figs. 2 and 3, respectively, by varying ϖ_D as well as given various values of the mobile-traffic intensity λ_U , where the BS density $\lambda_B = 5$ BSs/km², the average cellular service rate $R_C = 0.2$ Mb/s, the average D2D service rate $R_{D2D} = 2$ Mb/s, and $\beta = 20$, while additionally for the O mode network, the D2D-tier bandwidth fraction $\omega = 0.3$. Note that $\varpi_D \rightarrow 0$ corresponds to the network without D2D-enabled offloading, which is indicated by 'ESE w/o D2D-assisted offloading' in Figs. 2 and 3. It can be seen that the analytical values and the simulated results of the network's ESE agree with each other well. This verifies the accuracy of our theoretical ESE modeling. Observe from both Figs. 2 and 3 that the network's ESE increases quite fast with ϖ_D when ϖ_D is relatively small, and after reaching the peak value, the network's ESE becomes decreasing as ϖ_D increases. Also the network's ESE becomes saturated for sufficiently large ϖ_D . This is because when the effective D2D transmission radius is large, most of the cellular mobile-traffic have been offloaded to the D2D tier, and enlarging further

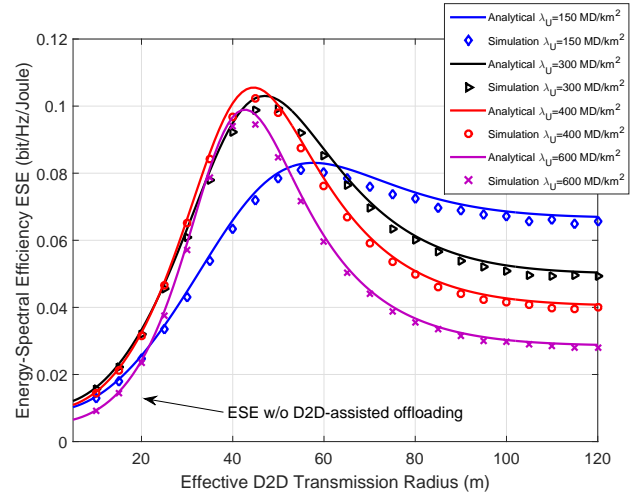


Fig 3. Achievable network ESE performance (curve: analytical and point: simulation) in O mode by varying ϖ_D and λ_U , given $\omega = 0.3$, $\lambda_B = 5$ BSs/km², $R_C = 0.2$ Mb/s, $R_{D2D} = 2$ Mb/s and $\beta = 20$.

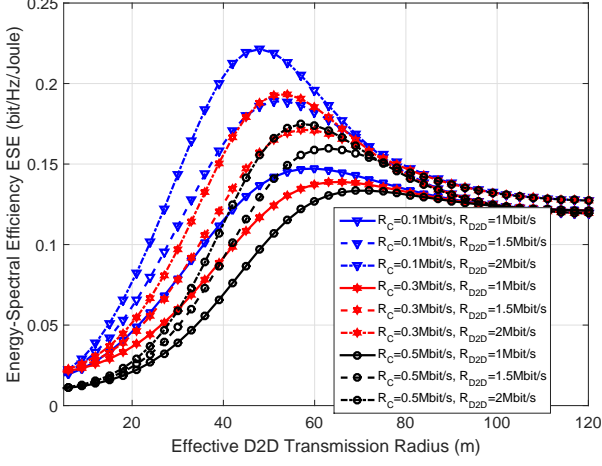


Fig. 4. Achievable network ESE performance in U mode by varying ϖ_D as well as R_C and R_{D2D} , given $\lambda_U = 500$ MDs/km², $\lambda_B = 5$ BSs/km² and $\beta=20$.

ϖ_D will not alter the achievable ESE noticeably. It can also be seen from Fig. 2 that for relatively small ϖ_D , increasing the mobile-traffic intensity λ_U always leads to a higher η_{ESE}^U but the same cannot be said for relatively large ϖ_D . As for the O mode network, the influence of λ_U to η_{ESE}^O is more complicated, which also depends on ϖ_D . Observe from Fig. 3 that for relatively large ϖ_D , increasing λ_U always leads to the decrease in η_{ESE}^O .

We continue to investigate the impacts of the effective D2D transmission radius ϖ_D as well as the cellular and D2D links' service data rates R_C and R_{D2D} by varying ϖ_D as well as R_C and R_{D2D} , and the resulting networks' ESE performance in the U and O modes are depicted in Figs. 4 and 5, respectively, where $\lambda_U = 500$ MDs/km², $\lambda_B = 5$ BSs/km² and $\beta=20$, while additionally for the O mode network, $\omega = 0.5$. Obviously, the influence of ϖ_D to the network's ESE is similar to that depicted in Figs. 2 and 3. We note that for the relatively small ϖ_D , the increase of the D2D tier service rate R_{D2D} leads to the increase of the network's ESE but by contrast, increasing the cellular service

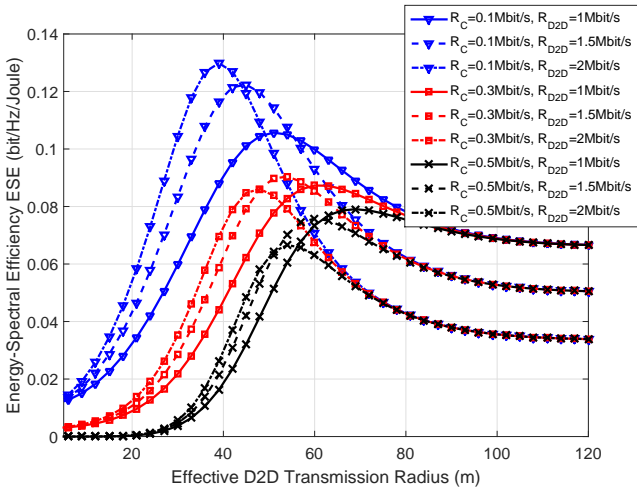


Fig. 5. Achievable network ESE performance in O mode by varying ϖ_D as well as R_C and R_{D2D} , given $\omega = 0.5$, $\lambda_U = 500$ MDs/km², $\lambda_B = 5$ BSs/km² and $\beta=20$.

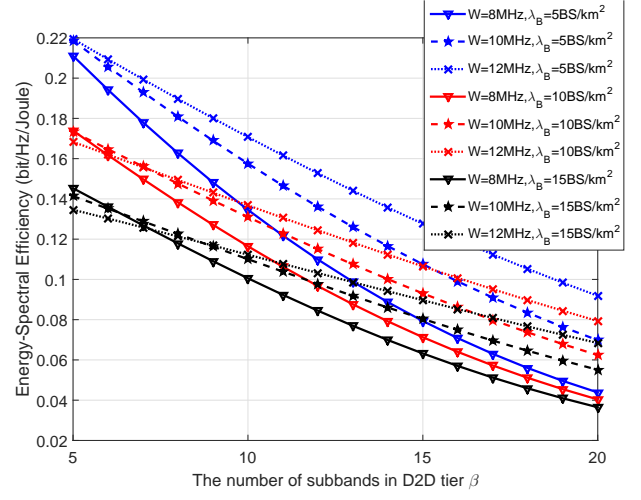


Fig. 6. Achievable network ESE performance in U mode by varying β as well as W and λ_B , given $R_C = 0.3$ Mb/s, $R_{D2D} = 2$ Mb/s and $\lambda_U = 500$ MDs/km².

rate R_C causes the decrease in the network's ESE. This implies that the cellular link is not an energy-spectral efficient link, in comparison with the D2D link, owing to its large distance and scarce spectrum resources. An interesting phenomenon observed from Figs. 4 and 5 is that for the same D2D-link service rate R_{D2D} , the ESEs corresponding to different cellular service rates R_C approach to a same value when ϖ_D is large. This is easy to understand since most of the mobile-traffic have been offloaded to the D2D tier for large ϖ_D and the network's ESE is mainly determined by R_{D2D} .

B. Impact of D2D Access Probability

The D2D access probability $1/\beta$ or the number of subbands in the D2D tier β is another key network parameter that significantly influences the network's achievable ESE performance. The achievable η_{ESE}^U and η_{ESE}^O are depicted in Figs. 6 and 7, respectively, by varying β as well as W and λ_B , where $R_C = 0.3$ Mb/s, $R_{D2D} = 2$ Mb/s and $\lambda_U = 500$ MDs/km², while additionally for the O-mode

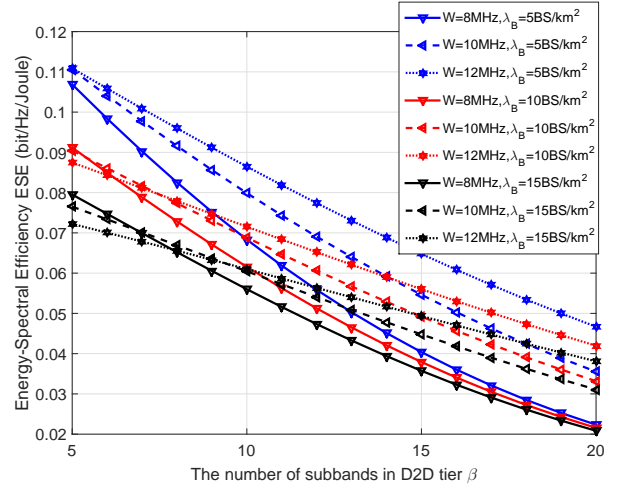


Fig. 7. Achievable network ESE performance in O mode by varying β as well as W and λ_B , given $\omega = 0.5$, $R_C = 0.3$ Mb/s, $R_{D2D} = 2$ Mb/s and $\lambda_U = 500$ MDs/km².

network, $\omega = 0.5$. As expected, increasing the BS density λ_B decreases the achievable ESE in general, since the increased interference as a result of increasing λ_B outweighs the effect of the reduced BS-MD distance. It can be seen that the network's ESE is a monotonically decreasing function of β . As expected, given $\lambda_B = 5 \text{ BS/km}^2$, increasing the system bandwidth W increases the achievable network's ESE, while given $\lambda_B = 10 \text{ BS/km}^2$ and 15 BS/km^2 as well as for not too small β , the network's ESE is an increasing function of W . Also observe that $\eta_{\text{ESE}}^{\text{O}}$ is smaller than $\eta_{\text{ESE}}^{\text{U}}$ in general, suggesting that the underlay operating mode is more spectrum and energy efficient than the overlay operating mode.

C. Performance Evaluation

Fig. 8 quantifies the required average D2D transmit powers of a typical D2D link in both the U and O modes, $P_{\text{D}_a}^{\text{U}}$ and $P_{\text{D}_a}^{\text{O}}$, as the functions of R_D and $R_{\text{D}2\text{D}}$, given $R_C = 0.2 \text{ Mb/s}$, $\lambda_U = 500 \text{ MDs/km}^2$, $\lambda_B = 5 \text{ BSs/km}^2$ and $\beta = 20$, while additionally $\omega = 0.5$ for the O-mode network. In Fig. 8, curves are the theoretical $P_{\text{D}_a}^{\text{U}}$ and $P_{\text{D}_a}^{\text{O}}$ calculated according to Propositions 1 and 2, while points correspond to the simulated results. Clearly, the simulated results agree with the analytical values well. In line with Propositions 1 and 2, the average D2D transmit power increases monotonically with ϖ_D . Obviously, $P_{\text{D}_a}^{\text{U}}$ is smaller than $P_{\text{D}_a}^{\text{O}}$ as more spectrum resource is available to D2D links in the U mode, and the increase of the mobile-traffic intensity increases the required average D2D transmit power as expected.

In Fig. 9, the relationship between the O-mode network's ESE $\eta_{\text{ESE}}^{\text{O}}$ and the fraction of the total bandwidth ω is illustrated. In line with Proposition 9, there exists a unique unconstrained optimal ω^* , which maximizes the achievable network's ESE $\eta_{\text{ESE}}^{\text{O}}$. Furthermore, increasing the mobile-traffic intensity moves the optimal ω^* to the right, which implies that when λ_U is increased, we need allocate more spectrum to the D2D tier in order to attain a higher ESE.

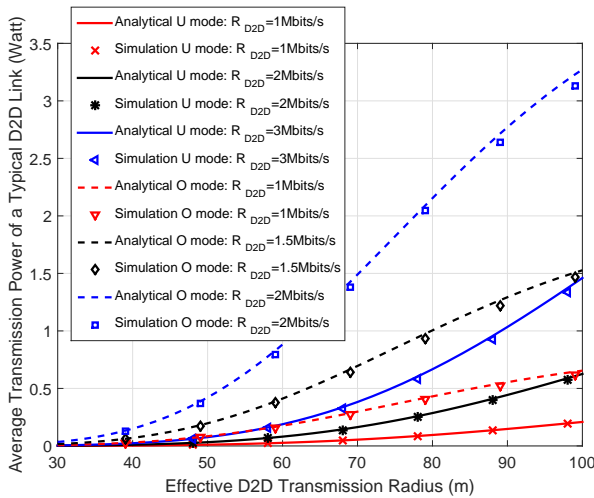


Fig. 8. Average required D2D transmit power (curve: analytical and point: simulation) as the function of ϖ_D and $R_{\text{D}2\text{D}}$, given $\lambda_U = 500 \text{ MDs/km}^2$, $R_C = 0.2 \text{ Mb/s}$, $\lambda_B = 5 \text{ BSs/km}^2$ and $\beta = 20$, while $\omega = 0.5$ for the O-mode network.

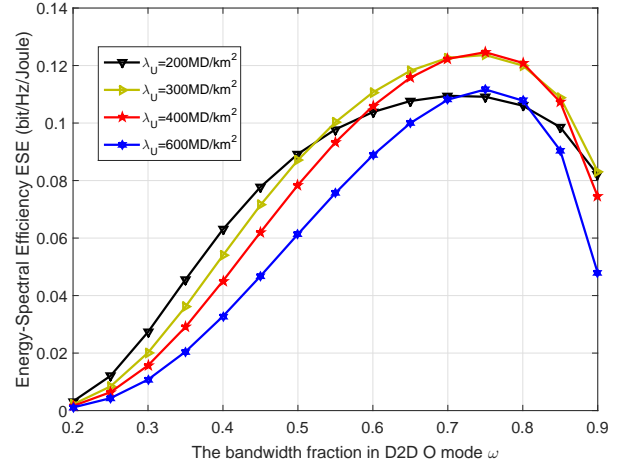


Fig. 9. Achievable network's ESE in O mode as the function of ω and λ_U , given $\varpi_D = 60 \text{ m}$, $R_C = 0.2 \text{ Mb/s}$, $R_{\text{D}2\text{D}} = 2 \text{ Mb/s}$, $\lambda_B = 5 \text{ BS/km}^2$ and $\beta = 20$.

D. Optimal Spectrum Partitioning in O Mode

We now illustrate our optimal spectrum sharing design of the constrained optimization (39) for the large-scale D2D-enabled cellular network operating in the O mode. The network parameters are given in Table II. For this network, the optimal unconstrained spectrum partitioning solution that maximizes the network's ESE $\eta_{\text{ESE}}^{\text{O}}$ is $\omega^* = 0.70$, which can be obtained by solving (42). From Fig. 9, it can also be seen that $\omega^* = 0.70$ corresponds to the peak value of the $\eta_{\text{ESE}}^{\text{O}}$ curve related to $\lambda_U = 200 \text{ MDs/km}^2$ as to be expected. The optimal constrained solutions $\bar{\omega}^*$ obtained according to Theorem 1 are listed in Table III given different values of the cellular outage threshold θ_O and D2D transmit power capacity Z_{O}^{D} .

TABLE III
OPTIMAL SPECTRUM SHARING SOLUTION FOR THE O-MODE NETWORK UNDER CELLULAR USER OUTAGE PERFORMANCE AND D2D TRANSMITTER POWER CONSTRAINTS

| θ_O | Z_{O}^{D} | $\omega_{\text{D}_a}, \omega_{\text{out}}$ | $\bar{\omega}^*$ |
|------------|---------------------------|---|------------------|
| 0.21 | 0.12 | [0.75, 0.80] | 0.75 |
| 0.18 | 0.32 | [0.60, 0.75] | 0.70 |
| 0.10 | 1.70 | [0.40, 0.55] | 0.55 |
| 0.12 | 0.09 | $\omega_{\text{D}_a} = 0.80 > \omega_{\text{out}} = 0.60$ | no solution |

VI. CONCLUSIONS

Based on the practical D2D-assisted offloading protocol conceived for the future 5G cellular network, we have proposed an ESE evaluation framework for large-scale mobile-traffic-aware D2D-enabled cellular networks operating in U and O modes, respectively, which allows us to analyze the impact of user-behaviors, D2D-assisted offloading parameters and cellular network operating conditions on the achievable network's ESE. Since all the our analytical results have been derived in closed-form and their accuracy has been verified in an extensive simulation study, they provide efficient and practical tools for designing and evaluating future D2D enabled cellular networks. With aid of our tractable and efficient ESE analytical tool, we have further designed an

optimal spectrum partitioning scheme for large-scaled D2D-enabled cellular networks operating in O mode to maximize the network's ESE with cellular user outage performance and D2D transmitters' power as constraints. This solution/design is particularly valuable to future D2D enabled cellular networks operating in O mode, because it ensures that the optimal offloading strategy always matches both the users behaviors and cellular operation conditions.

APPENDIX

A. Proof of Proposition 1

Proof: Recall that a typical MD u_i has the offloading region $\mathbb{O}(u_i, \varpi_D)$. A typical activated DT by broadcasting can simultaneously serve one or more MDs who's distances to the DT are less than or equal to ϖ_D . Let $\lambda_{DT}^{\text{act}}$ denote the density of active DTs, which can be expressed as

$$\lambda_{DT}^{\text{act}} = \frac{\Pr(N_{\mathbb{O}_{u_i}} \geq 1) \lambda_U}{E[N_{DT}]}, \quad (46)$$

where $E[N_{DT}]$ is the average number of MDs served by a DT. Let d denote the distance between u_i and its nearest DT. Then the probability density function (PDF) of d is given by $f(d) = 2\pi\lambda_{DT} \exp(-\pi\lambda_{DT}d^2)$. Therefore, the probability for a typical MD u_i is offloaded by DTs can be expressed as

$$\begin{aligned} P_{OL} &= \Pr(N_{\mathbb{O}_{u_i}} \geq 1) = \int_0^{\varpi_D} f(x) dx \\ &= 1 - \exp(-\pi\lambda_{DT}\varpi_D^2). \end{aligned} \quad (47)$$

Since the MDs are distributed according to a PPP, the probability massive function (PMF) of N_{DT} is $f_{DT}(N_{DT}) = (\pi\lambda_U\varpi_D^2)^{N_{DT}} \exp(-\pi\lambda_U\varpi_D^2) / (N_{DT}!)$, e.g., [38], where $\pi\lambda_U\varpi_D^2$ is the circular area covered by a typical DT. Therefore, the density of active DTs can be expressed as

$$\lambda_{DT}^{\text{act}} = \sum_{N_{DT}=1}^{+\infty} \frac{(\pi\lambda_U\varpi_D^2)^{N_{DT}} \lambda_U P_{OL}}{N_{DT} \exp(\pi\lambda_U\varpi_D^2) (N_{DT}!)}. \quad (48)$$

By applying the well-known property of the infinite series (48) and using the notation $y = \pi\lambda_U\varpi_D^2$, $\lambda_{DT}^{\text{act}}$ can be obtained as

$$\begin{aligned} \lambda_{DT}^{\text{act}} &= (\text{Ei}(y) - \text{Ei}(1) - \ln y + 1.3179) \exp(-y) P_{OL} \lambda_U \\ &= \mathcal{K} P_{OL} \lambda_U, \end{aligned} \quad (49)$$

where $\text{Ei}(y) = -\int_{-y}^{\infty} \frac{\exp(-t)}{t} dt$.

From (6), the average transmit power for the typical D2D link associated with the MD $u_k^{D,j}$ in the U mode is given by

$$P_{D_a}^U = E \left[\frac{(2^{\frac{\beta R_{D2D}}{W}} - 1)}{l_k^{D,n} h_k^{D,n}} (I_n^{CD} + I_n^{DD}) \right], \quad (50)$$

where the expectation $E[\cdot]$ is with respect to the distribution of the interferers' layout and the links' channel gains related to the links between the typical MD $u_k^{D,j}$ and its associated DTs [28]. Let $\mathcal{M}_{I_n^{CD}}(s)$ and $\mathcal{M}_{I_n^{DD}}(s)$ be the moment-generating functions (MGFs) of I_n^{CD} and I_n^{DD} , respectively. According

to [39] as well as the well-known property of MGFs $x = -\frac{\partial}{\partial s} \ln \mathcal{M}_x(s)|_{s=0}$, (50) can be expressed as

$$P_{D_a}^U = \frac{2^{\frac{\beta R_{D2D}}{W}} - 1}{r_{dd}^{-\alpha}} \left(-\frac{\partial}{\partial s} \ln \mathcal{M}_{I_n^{CD}} - \frac{\partial}{\partial s} \ln \mathcal{M}_{I_n^{DD}} \right) \Big|_{s=0}, \quad (51)$$

where

$$r_{dd} = \min_n \|u_k^{D,j} - \text{DT}_n\| \leq \varpi_D. \quad (52)$$

Next consider the D2D tier interference I'_{dd} contributed by the PPP distributed DT interferers in the finite annular region with the inner diameter r_{dd} and the outer diameter $r_{\text{out},D}$. The MGF of I'_{dd} is given by

$$\mathcal{M}_{I'_{dd}}(s) = \exp \left(2\pi\lambda_{DT}^{\text{act}} \int_{r_{dd}}^{r_{\text{out},D}} \left(1 - \exp \left(-\frac{sP_D}{\beta y^\alpha} \right) \right) y dy \right). \quad (53)$$

Similarly, consider the cellular tier interference I'_{cd} contributed by the PPP distributed BS interferers in the finite annular region with the inner diameter r_{cd} , where

$$r_{cd} = \min_{l \in \Theta_B} \|u_k^{D,j} - b_l\|, \quad (54)$$

and the outer diameter $r_{\text{out},C}$. The MGF of I'_{cd} is given by

$$\mathcal{M}_{I'_{cd}}(s) = \exp \left(2\pi\lambda_B \int_{r_{cd}}^{r_{\text{out},C}} \left(1 - \exp \left(-\frac{sP_C}{\beta x^\alpha} \right) \right) x dx \right). \quad (55)$$

Thus

$$I'_{dd} = -\frac{\partial \ln \mathcal{M}_{I'_{dd}}(s)}{\partial s} \Big|_{s=0} = \frac{2\pi\lambda_{DT}^{\text{act}} P_D (r_{\text{out},D}^{2-\alpha} - r_{dd}^{2-\alpha})}{(2-\alpha)\beta}, \quad (56)$$

$$I'_{cd} = -\frac{\partial \ln \mathcal{M}_{I'_{cd}}(s)}{\partial s} \Big|_{s=0} = \frac{2\pi\lambda_B P_C (r_{\text{out},C}^{2-\alpha} - r_{cd}^{2-\alpha})}{(2-\alpha)\beta}. \quad (57)$$

Clearly,

$$\lim_{r_{\text{out},D} \rightarrow \infty} I'_{dd} = I_n^{DD}, \quad (58)$$

$$\lim_{r_{\text{out},C} \rightarrow \infty} I'_{cd} = I_n^{CD}. \quad (59)$$

For $\alpha > 2$, plugging (56) to (59) into (51) yields the typical D2D link's average transmit power conditioned on r_{cd} and r_{dd} as

$$\begin{aligned} P_{D_a}^U \Big|_{r_{cd}, r_{dd}} &= \frac{2\pi(2^{\frac{\beta R_{D2D}}{W}} - 1)}{\beta(\alpha - 2)} \left(\lambda_B P_C r_{cd}^{2-\alpha} r_{dd}^\alpha \right. \\ &\quad \left. + \mathcal{K} P_{OL} \lambda_U P_D r_{dd}^2 \right). \end{aligned} \quad (60)$$

Further denote $v = r_{cd}$ and $u = r_{dd}$. By integrating (60) with the PDFs of u and v , we obtain the average transmit power $P_{D_a}^U$ given by

$$\begin{aligned} P_{D_a}^U &= \frac{8\pi^3(2^{\frac{\beta R_{D2D}}{W}} - 1)}{\beta(\alpha - 2)} \int_0^{\varpi_D} \frac{\lambda_{DT} u^{\alpha+1}}{\exp(\pi\lambda_{DT} u^2)} \int_0^{+\infty} \left(\frac{\lambda_B^2 P_C v^{3-\alpha}}{\exp(\pi\lambda_B v^2)} \right. \\ &\quad \left. + \frac{P_D \lambda_B \mathcal{K} \lambda_U u^{2-\alpha} v}{\exp(\pi\lambda_B v^2)} P_{OL} \right) dv du. \end{aligned} \quad (61)$$

After some algebraic manipulations to complete the integration (61), we obtain the expression of $P_{D_a}^U$ which is the righthand side of the inequality (10). This completes the proof. ■

B. Proof of Proposition 2

Proof: The average transmit power for the typical D2D link associated with the MD $u_k^{D,j}$ in the O mode is given by

$$P_{D_a}^O = E \left[\frac{\left(2^{\frac{\beta R_{D2D}}{\omega W}} - 1 \right) I_n^{DD}}{l_k^{D,n} h_k^{D,n}} \right], \quad (62)$$

which is obtained by removing I_n^{CD} and replacing W with ωW in (50). Similar to the derivation of (60), by considering the interfer layout $\overline{\mathcal{O}}(u_k^{D,j}, r_{dd})$ with r_{dd} as given in (52), the average D2D transmit power in the O mode conditioned on r_{dd} can be expressed as

$$P_{D_a}^O \Big|_{r_{dd}} = \left(2^{\frac{\beta R_{D2D}}{\omega W}} - 1 \right) \frac{2\pi \mathcal{K} \lambda_U P_D r_{dd}^2 P_{OL}}{\beta(\alpha - 2)}. \quad (63)$$

Again denote $u = r_{dd}$. By integrating (63) with the PDF of u , we arrive at

$$P_{D_a}^O = \int_0^{\varpi_D} \left(2^{\frac{\beta R_{D2D}}{\omega W}} - 1 \right) \frac{4\pi^2 \mathcal{K} P_D u^3}{\beta(\alpha - 2)} P_{OL} \lambda_U \times \lambda_{DT} \exp(-\pi \lambda_{DT} u^2) du. \quad (64)$$

After some algebraic manipulations to complete the integration (64), we obtain the expression of $P_{D_a}^O$ which is the righthand side of the inequality (15). This completes the proof. ■

C. Proof of Proposition 3

Proof: From (2), the average required transmit power for the typical link between BS b_j and MD $u_i^{C,j}$ in the U mode can be expressed by

$$P_{C,i}^U = E \left[\frac{\left(2^{\frac{R_C}{W_C}} - 1 \right) (I_j^{CC} + I_j^{DC})}{l_i^{C,j} h_i^{C,j}} \right], \quad (65)$$

where the interference I_j^{CC} and I_j^{DC} are given in (3) and (4), respectively, and hence the cellular and D2D interfer layouts, $\Theta_B \setminus b_j$ and $\overline{\mathcal{O}}(u_i^{C,j}, \varpi_D)$, are conditioned with $r_{cc} = \|u_i^{C,j} - b_j\|$. With the aid of the MGFs for I_j^{CC} and I_j^{DC} , we have

$$P_{C,i}^U = \left(2^{\frac{R_C}{W_C}} - 1 \right) r_{cc}^\alpha \left(-\frac{\partial \ln \mathcal{M}_{I_j^{CC}}}{\partial s} - \frac{\partial \ln \mathcal{M}_{I_j^{DC}}}{\partial s} \right) \Big|_{s=0}. \quad (66)$$

Similar to the proof of Proposition 1, define the cellular tier interference I'_{cc} in the corresponding annular region and the

D2D tier interference I'_{dc} in the relevant annular region. Then their MGFs are given respectively by

$$\mathcal{M}_{I'_{cc}} = \exp \left(2\pi \lambda_B \int_{r_{cc}}^{r_{bn,C}} \left(1 - \exp \left(-\frac{s P_C}{N_C x^\alpha} \right) \right) x dx, \quad (67)$$

$$\mathcal{M}_{I'_{dc}} = \exp \left(2\pi \lambda_{DT}^{\text{act}} \int_{\varpi_D}^{r_{bn,D}} \left(1 - \exp \left(-\frac{s P_D}{N_C x^\alpha} \right) \right) x dx. \quad (68)$$

Based on the well-known property of MGFs, we have

$$-\frac{\partial \ln \mathcal{M}_{I'_{cc}}}{\partial s} \Big|_{s=0} = \frac{2\pi \lambda_B P_C (r_{bn,C}^{2-\alpha} - r_{cc}^{2-\alpha})}{(2-\alpha) N_C}, \quad (69)$$

$$-\frac{\partial \ln \mathcal{M}_{I'_{dc}}}{\partial s} \Big|_{s=0} = \frac{2\pi \lambda_{DT}^{\text{act}} P_D (r_{bn,D}^{2-\alpha} - \varpi_D^{2-\alpha})}{(2-\alpha) N_C}. \quad (70)$$

By taking $r_{bn,C} \rightarrow \infty$ and $r_{bn,D} \rightarrow \infty$ in (69) and (70), respectively, and plugging the results into (66), we obtain the average DL transmit power of a typical BS-MD link associated with $u_i^{C,j}$ and conditioned on r_{cc} as

$$P_{C,i}^U \Big|_{r_{cc}} = \frac{2\pi \left(2^{\frac{R_C}{W_C}} - 1 \right)}{N_C (\alpha - 2)} \left(\lambda_B P_C r_{cc}^2 + \mathcal{K} P_{OL} \lambda_U P_D \varpi_D^{2-\alpha} r_{cc}^\alpha \right). \quad (71)$$

Since the PDF of r_{cc} is $f(r_{cc}) = 2\pi \lambda_B r_{cc} \exp(-\pi \lambda_B r_{cc}^2)$, the average required DL transmit power for $u_i^{C,j}$ is

$$P_{C,i}^U = \int_0^{+\infty} P_{C,i}^U \Big|_{r_{cc}=x} f(x) dx. \quad (72)$$

Completing the integral (72) leads to the average DL transmit power given in (16).

Then, we focus on b_j having the coverage area of $S_{A,j} = S_A$ and serving N_C cellular MDs. Since the PMF of $N_C \geq 1$ is given by [38]

$$g(N_C) = \frac{((1 - P_{OL}) \lambda_U S_A)^{N_C}}{(N_C!) \exp((1 - P_{OL}) \lambda_U S_A)}, \quad (73)$$

and all the cellular MDs have the identical rate of R_C , the average aggregate DL transmit power conditioned on the cell size of S_A is given by (74) at the bottom of this page. Noting that the PDF of S_A is given by $f_{S_A}(t) = \lambda_B^K \frac{K^K}{\Gamma(K)} t^{K-1} \exp(-K \lambda_B t)$ with $K = 3.75$ [40], we have

$$P_{BS}^U = \int_0^\infty P_{BS}^U \Big|_{S_A=t} f_{S_A}(t) dt. \quad (75)$$

After some algebraic manipulations to complete the integral (75), we arrive at (17). This completes the proof. ■

$$P_{BS}^U \Big|_{S_A} = \sum_{N_C=1}^\infty \sum_{i=1}^{N_C} g(N_C) P_{C,i}^U = \frac{2}{\alpha-2} \left(P_C + \frac{\alpha \Gamma(\frac{\alpha}{2}) \pi^{\frac{2-\alpha}{2}} P_D \lambda_U \mathcal{K} P_{OL}}{2 \lambda_B^{\frac{\alpha}{2}} \varpi_D^{\alpha-2}} \right) \left(\exp \left(\left(2^{\frac{R_C}{W_C}} - 1 \right) (1 - P_{OL}) \lambda_U S_A \right) - 1 \right). \quad (74)$$

$$Q_j^{C,i} \Big|_{r_{cc}} = 1 - \exp \left(-2\pi \lambda_B \int_{r_{cc}}^\infty \frac{\gamma}{\gamma + \left(\frac{v}{r_{cc}} \right)^\alpha} v dv \right) \exp \left(-2\pi \lambda_{DT}^{\text{act}} p_c \int_{p_c^{-1} \varpi_D}^\infty \frac{\gamma}{\gamma + \left(\frac{v}{r_{cc}} \right)^\alpha} v dv \right). \quad (76)$$

D. Proof of Proposition 7

Proof: The outage probability of a typical cellular user $u_i^{C,j}$ conditioned on r_{cc} is given by (76) at the bottom of the previous page. Considering the worst case that $u_i^{C,j}$ is at the cell edge, we have $r_{cc} = \sqrt{1/\pi\lambda_B}$, since $\pi r_{cc}^2 = 1/\lambda_B$. Plugging it into (76) and with the help of Appendix B of [35] completes the proof. ■

REFERENCES

- [1] L. Wei, R. Q. Hu, Y. Qian, and G. Wu, "Enable device-to-device communications underlying cellular networks: challenges and research aspects," *IEEE Commun. Mag.*, vol. 52, no. 6, pp. 90-96, Jun. 2014.
- [2] A. Asadi, Q. Wang, and V. Mancuso, "A survey on device-to-device communication in cellular networks," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 4, pp. 1801-1819, 2014.
- [3] M. N. Tehrani, M. Uysal, H. Yanikomeroglu, "Device-to-device communication in 5G cellular networks: challenges, solutions, and future directions," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 86-92, May 2014.
- [4] D. Astely, E. Dahlman, G. Fodor, S. Parkvall, and J. Sachs, "LTE release 12 and beyond," *IEEE Commun. Mag.*, vol. 51, no. 7, pp. 154-160, Jul. 2013.
- [5] Y. Li, T. Wu, P. Hui, D. Jin, and S. Chen, "Social-aware D2D communications: qualitative insights and quantitative analysis," *IEEE Commun. Mag.*, vol. 52, no. 6, pp. 150-158, Jun. 2014.
- [6] Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2016-2021, Feb. 7, 2017 [Online]. Available: <http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.html>
- [7] Y. Shen, C. Jiang, T. Q. S. Quek, and Y. Ren, "Device-to-device assisted communications in cellular networks: an energy efficient approach in downlink video sharing scenario," *IEEE Trans. Wireless Commun.*, vol. 15, no. 2, pp. 1575-1587, Feb. 2016.
- [8] J. Du, E. Gelenbe, C. Jiang, H. Zhang, and Y. Ren, "Contract design for traffic offloading and resource allocation in heterogeneous ultra-dense networks," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 11, pp. 2457-2467, Nov. 2017.
- [9] J. Wang, C. Jiang, Z. Bie, T. Q. S. Quek, and Y. Ren, "Mobile data transactions in device-to-device communication networks: pricing and auction," *IEEE Wireless Commun. Lett.*, vol. 5, no. 3, pp. 300-303, Jun. 2016.
- [10] C.-H. Yu, K. Doppler, C. B. Ribeiro, and O. Tirkkonen, "Resource sharing optimization for device-to-device communication underlying cellular networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2752-2763, Aug. 2011.
- [11] G. Yu, L. Xu, D. Feng, R. Yin, G. Y. Li, and Y. Jiang, "Joint mode selection and resource allocation for device-to-device communications," *IEEE Trans. Commun.*, vol. 62, no. 11, pp. 3814-3824, Nov. 2014.
- [12] D. Wu, J. Wang, R. Hu, Y. Cai, and L. Zhou, "Energy-efficient resource sharing for mobile device-to-device multimedia communications," *IEEE Trans. Vehicular Technology*, vol. 63, no. 5, pp. 2093-2103, Jun. 2014.
- [13] D. D. Penda, L. Fu, and M. Johansson, "Mode selection for energy efficient D2D communications in dynamic TDD systems," in *Proc. ICC 2015* (London, UK), Jun. 8-12, 2015, pp. 5404-5409.
- [14] T. D. Hoang, L. B. Le, and T. Le-Ngoc, "Energy-efficient resource allocation for D2D communications in cellular networks," *IEEE Trans. Vehicular Technology*, vol. 65, no. 9, pp. 6972-6986, Sep. 2016.
- [15] Z. Zhou, M. Dong, K. Ota, R. Shi, Z. Liu, and T. Sato, "Game-theoretic approach to energy-efficiency resource allocation in device-to-device underlay communications," *IET Communications*, vol. 9, no. 3, pp. 375-385, 2015.
- [16] Z. Zhou, M. Dong, K. Ota, G. Wang, and L. T. Yang, "Energy-efficient resource allocation for D2D communications underlying cloud-RAN-based LTE-A networks," *IEEE Internet of Things J.*, vol. 3, no. 3, pp. 428-438, Jun. 2016.
- [17] Z. Zhou, M. Dong, K. Ota, J. Wu, and T. Sato, "Energy efficiency and spectral efficiency tradeoff in device-to-device (D2D) communications," *IEEE Wireless Commun. Letters.*, vol. 3, no. 5, pp. 485-488, Oct. 2014.
- [18] A. Bhardwaj and S. Agnihotri, "Energy- and spectral-efficiency trade-off for D2D-multicasts in underlay cellular networks," *IEEE Wireless Commun. Letters.*, vol. 7, no. 4, pp. 546-549, Aug. 2014.
- [19] R. Zhang, Y. Li, C.-X. Wang, Y. Ruan, Y. Fu, and H. Zhang, "Energy-spectral efficiency trade-off in underlying mobile D2D communication: an economic efficiency perspective," *IEEE Trans. Wireless Commun.*, vol. 17, no. 7, pp. 4288-4301, Jul. 2018.
- [20] S. Andreev, A. Pyattaev, K. Johnsson, O. Galinina, and Y. Koucheryavy, "Cellular traffic offloading onto network-assisted device-to-device connections," *IEEE Commun. Mag.*, vol. 52, no. 4, pp. 20-31, Apr. 2014.
- [21] Z. Wang and V. W. S. Wong, "A novel D2D data offloading scheme for LTE networks," in *Proc. ICC 2015* (London, UK), Jun. 8-12, 2015, pp. 3107-3112.
- [22] Y. Li, Z. Wang, D. Jin, and S. Chen, "Optimal mobile content downloading in device-to-device communication underlying cellular networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 7, pp. 3596-3608, Jul. 2014.
- [23] T. Wang, Y. Sun, L. Song, and Z. Han, "Social data offloading in D2D-enhanced cellular networks by network formation games," *IEEE Trans. Wireless Commun.*, vol. 14, no. 12, pp. 7004-7015, Dec. 2015.
- [24] L. Al-Kanj, H. V. Poor, and Z. Dawy, "Optimal cellular offloading via device-to-device communication networks with fairness constraints," *IEEE Trans. Wireless Commun.*, vol. 13, no. 8, pp. 4628-4643, Aug. 2014.
- [25] Y. Li, D. Jin, P. Hui, and Z. Han, "Optimal base station scheduling for device-to-device communication underlying cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 1, pp. 27-40, Jan. 2016.
- [26] N. Lee, X. Lin, J. G. Andrews, and R. W. Heath, Jr., "Power control for D2D underlaid cellular networks: modeling, algorithms and analysis," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 1, pp. 1-13, Jan. 2015.
- [27] Q. Ye, M. Al-Shalash, C. Caramanis, and J. G. Andrews, "Resource optimization in device-to-device cellular systems using time-frequency hopping," *IEEE Trans. Wireless Commun.*, vol. 13, no. 10, pp. 5467-5480, Oct. 2014.
- [28] X. Lin, J. G. Andrews, and A. Ghosh, "Spectrum sharing for device-to-device communication in cellular networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 12, pp. 6727-6740, Dec. 2014.
- [29] H. ElSawy, E. Hossain, and M.-S. Alouini, "Analytical modeling of mode selection and power control for underlay D2D communication in cellular networks," *IEEE Trans. Commun.*, vol. 62, no. 11, pp. 4147-4161, Nov. 2014.
- [30] S. Stefanatos, A. G. Gotsis, and A. Alexiou, "Operational region of D2D communications for enhancing cellular network performance," *IEEE Trans. Commun.*, vol. 14, no. 11, pp. 5984-5997, Nov. 2015.
- [31] G. Zhao, S. Chen, L. Zhao, and L. Hanzo, "Energy-spectral-efficiency analysis and optimization of heterogeneous cellular networks: a large-scale user behaviors perspective," *IEEE Trans. Vehicular Technology*, vol. 67, no. 5, pp. 4098-4112, May 2018.
- [32] Y. Li, S. Su, and S. Chen, "Social-aware resource allocation for device-to-device communications underlying cellular networks," *IEEE Wireless Commun. Lett.*, vol. 4, no. 3, pp. 293-296, Jun. 2015.
- [33] C. Gao, Y. Li, Y. Zhao, and S. Chen, "A two-level game theory approach for joint relay selection and resource allocation in network coding assisted D2D communications," *IEEE Trans. Mobile Computing*, vol. 16, no. 10, pp. 2697-2711, Oct. 2017.
- [34] C. Gao, H. Zhang, X. Chen, Y. Li, D. Jin, and S. Chen, "Impact of selfishness in device-to-device communication underlying cellular networks," *IEEE Trans. Vehicular Technology*, vol. 66, no. 10, pp. 9338-9349, Oct. 2017.
- [35] J. G. Andrews, F. Baccelli, and R. K. Ganti, "A tractable approach to coverage and rate in cellular networks," *IEEE Trans. Commun.*, vol. 59, no. 11, pp. 3122-3134, Nov. 2011.
- [36] S.-R. Cho and W. Choi, "Energy-efficient repulsive cell activation for heterogeneous cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 5, pp. 870-882, May 2013.
- [37] E. Björnson, L. Sanguinetti, J. Hoydis, and M. Debbah, "Optimal design of energy-efficient multi-user MIMO systems: Is massive MIMO the answer?," *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 3059-3075, Jun. 2015.
- [38] R. L. Streit, *Poisson Point Processes: Imaging, Tracking and Sensing*. Springer, 2010.
- [39] S. Srinivasa, *Modeling Interference in Uniformly Random Wireless Networks: Theory and Applications*. MSc Thesis, University of Notre Dame, USA, 2007.
- [40] J. Ferenc and Z. Neda, "On the size distribution of Poisson Voronoi cells," *Physica A: Statistical Mechanics and its Applications*, vol. 385, no. 2, pp. 518-526, 2007.



Sheng Chen (M'90-SM'97-F'08) received his BEng degree from the East China Petroleum Institute, Dongying, China, in 1982, and his PhD degree from the City University, London, in 1986, both in control engineering. In 2005, he was awarded the higher doctoral degree, Doctor of Sciences (DSc), from the University of Southampton, Southampton, UK. From 1986 to 1999, He held research and academic appointments at the Universities of Sheffield, Edinburgh and Portsmouth, all in UK. Since 1999, he has been with the School of Electronics and Computer

Science, the University of Southampton, UK, where he holds the post of Professor in Intelligent Systems and Signal Processing. Dr Chen's research interests include adaptive signal processing, wireless communications, modelling and identification of nonlinear systems, neural network and machine learning, intelligent control system design, evolutionary computation methods and optimisation. He has published over 650 research papers. Dr. Chen is a Fellow of the United Kingdom Royal Academy of Engineering, a Fellow of IET, a Distinguished Adjunct Professor at King Abdulaziz University, Jeddah, Saudi Arabia, and an original ISI highly cited researcher in engineering (March 2004). Professor Chen has 13,000+ Web of Science citations and 27,000+ Google Scholar citations.