Optimal Power Allocation With Statistical QoS Provisioning for D2D and Cellular Communications Over Underlaying Wireless Networks

Wenchi Cheng, Xi Zhang, *Senior Member, IEEE*, and Hailin Zhang, *Member, IEEE*

***Abstract*—By enabling two adjacent mobile devices to establish a direct link, device-to-device (D2D) communication can increase the system throughput over *underlaying wireless networks*, where D2D and cellular communications coexist to share the same radio resource. Traditional D2D schemes mainly focus on maximizing the system throughput without taking into account the quality-of-service (QoS) provisioning. To overcome this problem, we develop a framework to investigate the impact of delay-QoS requirement on the performance of D2D and cellular communications in underlaying wireless networks. Then, we propose the optimal power allocation schemes with statistical QoS provisioning for the following two channel modes: 1). co-channel mode based underlaying wireless networks where D2D devices and cellular devices share the same frequency-time resource; 2). orthogonal-channel mode based underlaying wireless networks where the frequency-time resource is partitioned into two parts for D2D devices and cellular devices, respectively. Applying our proposed optimal power allocations into D2D based underlaying wireless networks, we obtain the maximum network throughput subject to a given delay-QoS constraint for above-mentioned two underlaying wireless network modes, respectively. Also conducted is a set of numerical and simulation results to evaluate our proposed QoS-driven power allocation schemes under different delay-QoS requirements.**

***Index Terms*—Underlaying wireless networks, statistical quality-of-service (QoS) provisioning, power allocation, device-to-device (D2D) communication, cellular communication.**

I. Introduction

Underlaying device-to-device (D2D) and cellular communication, as a promising, but challenging, technical approach to enhance the spectrum efficiency of wireless cellular networks, has been paid much research attention recently [1]–[3]. In underlaying wireless networks, instead of transmission through base station (BS) or access point (AP), two devices in proximity of each other may communicate directly as a D2D communication pair, bypassing BS or AP. This kind of communication and device are called D2D communication and D2D device, respectively. D2D communication allows either the central controlling or the distributed controlling. For the scenario with central controlling, the BS or the AP manages the switching between the D2D communication and the traditional cellular communication[1] for D2D devices [2]. For the scenario with distributed controlling, the D2D devices manage the D2D communication and the traditional cellular communication by themselves [4]. Due to its easy synchronization and implementation, we focus on the central controlling mechanism in this paper.

In underlaying wireless networks, D2D devices can use three transmission modes [3], [5]–[7]: 1). The co-channel mode where the D2D devices use the same frequency-time resource as the cellular devices; 2). The orthogonal-channel mode where the D2D devices use the partial orthogonal frequency-time resource with the cellular devices; 3). The cellular mode where the D2D devices communicate with each other via a BS as the traditional transmission in wireless cellular networks. The co-channel mode and the orthogonal-channel mode can be uniformly called the D2D mode. It is worth to employ the D2D mode when the co-channel mode or the orthogonal-mode can increase the system throughput as compared with that of the cellular mode. Otherwise, it is better for the D2D devices to choose the cellular mode.

Since the transmit power of the D2D devices determines the received signal-to-noise ratio (SNR) of the D2D communication and the interference of the cellular communication, many power allocation strategies have been proposed to increase the system throughput or decrease the power consumption of underlaying wireless networks [8]–[11]. With the co-channel mode, the authors of [8] proposed the joint power allocation and mode selection scheme to maximize the power efficiency of the D2D communication. With the orthogonal-channel mode, the authors of [9] developed the power optimization scheme with joint subcarrier allocation, adaptive modulation, and mode selection in an OFDMA system with D2D communication. Also with the orthogonal-channel mode, the authors in [10] analyzed the maximum achievable transmission capacity of the D2D communication in heterogeneous networks with multi-bands. In [11], the authors proposed the optimal centralized power allocation strategies for both the co-channel mode and the orthogonal-channel mode. All these schemes are efficient for underlaying wireless networks without taking the delay-quality-of-service (delay-QoS) requirements into account. To

This work was supported in part by the U.S. National Science Foundation under Grant CNS-205726 and the U.S. National Science Foundation CAREER Award under Grant ECCS-0348694.

Xi Zhang is with the Networking and Information Systems Laboratory, Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77843, USA (e-mail: xizhang@ece.tamu.edu).

Wenchi Cheng and Hailin Zhang are with the State Key Laboratory of Integrated Services Networks, Xidian University, Xi'an, 710071, China (e-mails: chengwenchi1986@gmail.com; hlzhang@xidian.edu.cn).

Digital Object Identifier 0000000000000000

[1]D2D devices can communicate with each other through either the D2D communication which bypasses the BS (or AP) or the traditional cellular communication which goes through BS (or AP).

guarantee real-time transmission of time-sensitive traffic and the transmission reliability, it is necessary to take delay-QoS requirements into account which is very crucial for underlaying wireless networks.

Because of the nature of the time-varying channels, the deterministic quality-of-service (QoS) is usually difficult to guarantee for real-time transmission in wireless networks. Consequently, the statistical QoS guarantee, in terms of QoS exponent and effective capacity, has become an important alternative to support real-time wireless communications in wireless networks [12]–[15]. Effective capacity is defined as the maximum constant arrival rate which can be supported by the service rate to guarantee the specified QoS exponent θ. The effective capacity characterizes the system throughput with different delay-QoS requirements. For real-time traffic such as video conference, a stringent delay-bound needs to be guaranteed and the effective capacity turns to be the outage capacity. On the other hand, the non-real-time traffic such as data disseminations demands high throughput while a loose delay constraint is imposed and the effective capacity turns to be the ergodic capacity. Thus, the previous works on D2D networks are only efficient for the non-real-time traffic where the delay constraint is very loose [1]–[11]. To support statistical QoS provisioning real-time services over underlaying wireless networks, in this paper we propose the optimal power allocation schemes to maximize the system throughput for real-time traffics with different QoS requirements in D2D and cellular communications over the underlaying wireless networks.

The rest of this paper is organized as follows. Section II describes the system model, where we design the QoS-guaranteed underlaying wireless networks. We also derive a sufficient condition to guarantee that using D2D and cellular communications can achieve larger effective capacity than that of only using the cellular communication. Section III formulates the optimization problems for the QoS-guaranteed co-channel mode. The QoS-driven power allocation scheme is also developed to maximize the effective capacity of the co-channel mode. Section IV formulates the optimization problems for the QoS-guaranteed orthogonal-channel mode. We also develop the optimal power allocation scheme for the QoS-guaranteed orthogonal-channel mode. Section V simulates and evaluates our proposed power allocation schemes for the QoS-guaranteed underlaying wireless networks with different delay-QoS requirements. The paper concludes with Section VI.

II. System Models

In this paper, we consider the underlaying wireless network, which is defined as the network structure where D2D and cellular communications coexist to share the same radio resources. Fig. 1 shows an example of underlaying wireless network, consisting of a number of important components with different operating modes to be defined in the following sections. Using the defined underlaying wireless network structure, its components, and operating modes, we develop the framework where the statistical delay-QoS provisionings can be applied to efficiently support the real-time services in the D2D and cellular communications.

A. The Underlaying Wireless Network

Referring to the example of underlaying wireless networks structure shown in Fig. 1, we define its components and operating modes as follows.

A-1. D2D Device and Cellular Device

We define the *D2D devices* as the mobile terminals which can implement the direct communication (D2D communication) with each other without using a BS. On the other hand, the *cellular devices* refer to the mobile terminals which can only communicate through a BS. A pair of D2D communication parties forms a *D2D pair*. In contrast, the communication going through a BS is called the cellular communication. A number of D2D devices in proximity of each other can be clustered into a *D2D group*. If a D2D device attempts to send a signal to another D2D device within the same D2D group, it needs to determine whether to transmit it bypassing or going through a BS. If a D2D device wants to communicate with the mobile device outside this D2D group (which can be a cellular device or a D2D device in another D2D group), it must transmit the signal through a BS.

A-2. The D2D-Cellular Underlaying Unit

A D2D-cellular underlaying unit is defined as a unit which is composed of a D2D pair, a cellular device, and the BS. For an example as shown in Fig. 1, a D2D-cellular underlaying unit consists of three devices: D1, D2, and D3, where D2 and D3 are two D2D devices and form a D2D pair; D1 represents a cellular device. We denote the power gains of the channels from D1 to the BS, from D2 to D3, from D2 to the BS, from D1 to D3, and from the BS to D3 by γ_1, γ_2, γ_3, γ_4, and γ_5, respectively.[2] Therefore, an underlaying wireless network is partitioned into a number of D2D-cellular underlaying units. This implies that in our underlaying wireless network model, each mobile terminal uniquely belongs to one D2D-cellular underlaying unit, playing a role of either D2D device or cellular device, and each D2D device uniquely belongs to a D2D group. Thus, we can focus on the D2D-cellular underlaying unit to investigate the D2D and cellular communications in underlaying wireless networks. We assume a flat fading channel model for all wireless channels. All channel power gains follow the stationary block fading model, where they remain unchanged within a time frame having the fixed length T, but vary independently across different time frames. The frame duration T is assumed to be less than the fading coherence time, but sufficiently long so that the information-theoretic assumption of infinite code-block length is meaningful. We set the assigned bandwidth to one D2D-cellular underlaying unit as B. We assume the circularly symmetric complex Gaussian noise with the variance σ^2 for all wireless channels.

A-3. Co-Channel Mode, Orthogonal-Channel Mode, and Cellular Mode

A D2D-cellular underlaying unit can operate under one of three different modes: the *co-channel mode*, the *orthogonal-*

[2]For the special case $\gamma_1 = 0$, there is only D2D communications in the D2D-cellular underlaying unit. For the special case that $\gamma_2 = 0$, there is only standard cellular communications in the D2D-cellular underlaying unit. Both of these two scenarios are considered to be the trivial cases and thus not discussed in this paper.

channel mode, and the *cellular mode*, respectively. The *co-channel mode* is defined as the mode where the D2D devices use the same frequency-time resource with the cellular devices. The *orthogonal-channel mode* is defined as the mode where the D2D devices use the part of orthogonal frequency-time resource with the cellular devices as the *orthogonal-channel mode*. The *cellular mode* is defined as the mode that the D2D devices communicate with each other via a BS. Notice that D2D communications can only be applied under the co-channel mode or the orthogonal-channel mode. However, under the cellular mode, all devices (D2D and cellular devices) must use the standard cellular communications through a BS.

All the devices in one D2D-cellular underlaying unit (consisting of D1, D2, and D3) are controlled by the BS. Thus, the channel state information (CSI) can be transferred through the channels between the BS and each device (a D2D device or a cellular device). For the co-channel mode, the BS can get the CSI γ_3 from D2 and the CSI γ_1 from D1. The BS can also obtain the CSIs γ_2 and γ_4 from D3 through the channel between D3 and the BS corresponding to γ_5. After having obtained the CSIs γ_1, γ_2, γ_3, and γ_4, the BS can send γ_1, γ_2, γ_3, and γ_4 to D1 and D2. For the orthogonal-channel mode, the BS can get the CSI γ_1 from D1. D2 can obtain the CSI γ_2 through the feedback channel from D3 to D2 directly and send the CSI γ_2 to the BS. After having obtained the CSIs γ_1 and γ_2, the BS can send γ_1 and γ_2 to D1 and D2. For the cellular mode, the BS can get the CSIs γ_1 and γ_3 from D1 and D2, respectively. Then, the BS can send γ_1 and γ_3 to D1 and D2.

B. Statistical Delay-QoS Provisionings Over Underlaying Wireless Networks

To support statistical QoS provisioning real-time services over underlaying wireless networks, we need to formally define the measure for the statistical QoS provisioning with D2D and cellular communications in one D2D-cellular underlaying unit.

Based on large deviation principle, the author of [16] showed that with sufficient conditions, the queue length process $Q(t)$ converges in distribution to a random variable $Q(\infty)$ such that

$$-\lim_{Q_{th}\to\infty}\frac{\log(\Pr\{Q(\infty)>Q_{th}\})}{Q_{th}}=\theta \tag{1}$$

where Q_{th} is the queue length bound and the parameter $\theta>0$ is a real-valued number. The parameter θ, which is called the *QoS exponent*, indicates the exponential decay rate of the delay-bound QoS violation probabilities. A larger θ corresponds to a faster decay rate, which implies that the system can provide a more *stringent* QoS requirement. A smaller θ leads to a slower decay rate, which implies a *looser* QoS requirement. Asymptotically, when $\theta\to\infty$, this implies that the system cannot tolerate any delay, which corresponds to the very stringent QoS constraint. On the other hand, when $\theta\to 0$, the system can tolerate an arbitrarily long delay, which corresponds to the very loose QoS constraint. The QoS exponent θ specified by Eq. (1) characterizes the delay-QoS requirement of underlaying wireless networks. Since there may be different traffics with various delay-QoS requirements in underlaying wireless networks, we can devise a general scheme for any θ varying from 0 to ∞, reflecting the variation from the very loose QoS constraint to the very stringent QoS constraint.

Then, we further define the sequence $\{R[k], k=1,2,...\}$ as the data service-rate, which is a discrete-time stationary and ergodic stochastic process. The parameter k represents the time frame index with a fixed time-duration equal to T. The $R[k]$ changes from frame to frame and $S[t]\triangleq\Sigma_{k=1}^{t}R[k]$ represents the partial sum of the service process. The Gartner-Ellis limit of $S[t]$, expressed as $\Lambda_C(\theta)=\lim_{t\to\infty}(1/t)\log\left(\mathbb{E}\left\{e^{\theta S[t]}\right\}\right)$, is a convex function differentiable for all real-valued θ, where $\mathbb{E}\{\cdot\}$ denotes the expectation. Inspired by the principle of effective bandwidth [17], the authors in [12] defined effective capacity as the maximum constant arrival rate which can be supported by the service rate to guarantee the specified QoS exponent θ. If the service-rate sequence $R[k]$ is stationary and time uncorrelated, the effective capacity can be written as [14]

$$\mathcal{C}(\theta)=-\frac{1}{\theta}\log\left(\mathbb{E}\left\{e^{-\theta R[k]}\right\}\right). \tag{2}$$

For simplicity, we assume that the delay-bound QoS exponent θ for D2D devices and cellular devices in the underlaying wireless network are the same. Our purpose is to maximize the effective capacity of the underlaying wireless network.

B-1. Effective Capacity for Co-Channel Mode in D2D-Cellular Underlaying Unit

Using the co-channel mode, we can derive the instantaneous transmission rate for the D2D-cellular underlaying unit, denoted by $R_1\left(P_1(\boldsymbol{\nu}),P_2(\boldsymbol{\nu})\right)$, as follows:

$$\begin{aligned}&R_1\left(P_1(\boldsymbol{\nu}),P_2(\boldsymbol{\nu})\right)\\&=BT\left[f_1\left(P_1(\boldsymbol{\nu}),P_2(\boldsymbol{\nu})\right)+f_2\left(P_1(\boldsymbol{\nu}),P_2(\boldsymbol{\nu})\right)\right]\\&=BT\left[\log_2\left(1+\frac{P_1(\boldsymbol{\nu})\gamma_1}{\sigma^2+P_2(\boldsymbol{\nu})\gamma_3}\right)\right.\\&\qquad\left.+\log_2\left(1+\frac{P_2(\boldsymbol{\nu})\gamma_2}{\sigma^2+P_1(\boldsymbol{\nu})\gamma_4}\right)\right],\end{aligned} \tag{3}$$

where $\boldsymbol{\nu}\triangleq(\gamma_1,\gamma_2,\gamma_3,\gamma_4,\theta)$ is defined as the QoS-based CSI for the D2D-cellular underlaying unit with the co-channel mode; $P_1(\boldsymbol{\nu})$ and $P_2(\boldsymbol{\nu})$ denote the instantaneous transmit power of D1 and D2, respectively; $f_1\left(P_1(\boldsymbol{\nu}),P_2(\boldsymbol{\nu})\right)$ and $f_2\left(P_1(\boldsymbol{\nu}),P_2(\boldsymbol{\nu})\right)$ are the instantaneous spectrum efficiencies for cellular communication and D2D communication in the D2D-cellular underlaying unit with the co-channel mode, respectively. To simplify the mathematical expression, we define the full channel state information (FCSI) for the co-channel mode based D2D communication as $\boldsymbol{\gamma}\triangleq(\gamma_1,\gamma_2,\gamma_3,\gamma_4)$.

From Eqs. (2) and (3), we can obtain the effective capacity of the D2D-cellular underlaying unit using the co-channel mode, denoted by $C_1(P_1(\boldsymbol{\nu}),P_2(\boldsymbol{\nu}),\theta)$, as follows:

$$\begin{aligned}&C_1\left(P_1(\boldsymbol{\nu}),P_2(\boldsymbol{\nu}),\theta\right)=-\frac{1}{\theta}\log\left(\mathbb{E}_{\boldsymbol{\gamma}}\left\{e^{-\theta R_1(P_1(\boldsymbol{\nu}),P_2(\boldsymbol{\nu}))}\right\}\right)\\&=-\frac{1}{\theta}\log\left(\mathbb{E}_{\boldsymbol{\gamma}}\left\{e^{-\beta\left[\log\left(1+\frac{P_1(\boldsymbol{\nu})\gamma_1}{\sigma^2+P_2(\boldsymbol{\nu})\gamma_3}\right)+\log\left(1+\frac{P_2(\boldsymbol{\nu})\gamma_2}{\sigma^2+P_1(\boldsymbol{\nu})\gamma_4}\right)\right]}\right\}\right),\end{aligned} \tag{4}$$

where $\beta = (\theta TB)/\log 2$ is the normalized QoS exponent and $\mathbb{E}_{\boldsymbol{x}}\{\cdot\}$ represents the expectation over $\boldsymbol{x}$.

B-2. Effective Capacity for Orthogonal-Channel Mode in D2D-Cellular Underlaying Unit

We assume that the bandwidth are equally partitioned for the D2D communication and the cellular communication. Then, we can derive the instantaneous transmission rate for the D2D-cellular underlaying unit using the orthogonal-channel mode, denoted by $R_2\left(P_1(\widetilde{\boldsymbol{\nu}}_1), P_2(\widetilde{\boldsymbol{\nu}}_2)\right)$, as follows:

$$R_2\left(P_1(\widetilde{\boldsymbol{\nu}}_1), P_2(\widetilde{\boldsymbol{\nu}}_2)\right) = \frac{BT}{2}\Bigg[\log_2\left(1+P_1(\widetilde{\boldsymbol{\nu}}_1)\gamma_1\right) + \log_2\left(1+P_2(\widetilde{\boldsymbol{\nu}}_2)\gamma_2\right)\Bigg], \quad (5)$$

where $\widetilde{\boldsymbol{\nu}}_1 \triangleq (\gamma_1, \theta)$ and $\widetilde{\boldsymbol{\nu}}_2 \triangleq (\gamma_2, \theta)$ are defined as the QoS-based CSI for D1 and D2 with the orthogonal-channel mode, respectively; D1 and D2 use the instantaneous transmit power $P_1(\widetilde{\boldsymbol{\nu}}_1)$ and $P_2(\widetilde{\boldsymbol{\nu}}_2)$. The FCSI for the orthogonal-channel mode based D2D communication is defined as $\widetilde{\boldsymbol{\gamma}} \triangleq (\gamma_1, \gamma_2)$.

From Eqs. (2) and (5), we can obtain the effective capacity of the D2D-cellular underlaying unit using the orthogonal-channel mode, denoted by $C_2\left(P_1(\widetilde{\boldsymbol{\nu}}_1), P_2(\widetilde{\boldsymbol{\nu}}_2), \theta\right)$, as follows:

$$C_2\left(P_1(\widetilde{\boldsymbol{\nu}}_1), P_2(\widetilde{\boldsymbol{\nu}}_2), \theta\right) = -\frac{1}{\theta}\log\left(\mathbb{E}_{\widetilde{\boldsymbol{\gamma}}}\left\{e^{-\theta R_2\left(P_1(\widetilde{\boldsymbol{\nu}}_1), P_2(\widetilde{\boldsymbol{\nu}}_2)\right)}\right\}\right) = -\frac{1}{\theta}\log\left(\mathbb{E}_{\widetilde{\boldsymbol{\gamma}}}\left\{e^{-\frac{\beta}{2}\left(\log\left[(1+P_1(\widetilde{\boldsymbol{\nu}}_1)\gamma_1)(1+P_2(\widetilde{\boldsymbol{\nu}}_2)\gamma_2)\right]\right)}\right\}\right). \quad (6)$$

B-3. Effective Capacity for Cellular Mode in D2D-Cellular Underlaying Unit

If only using cellular communications, D1 occupies $B/2$ frequency band during the whole time frame. D2 sends its signal to D3 through the BS using the other $B/2$ frequency band. In the first half frame, D2 sends its signal the BS (uplink). In the second half frame, the BS sends the signal from D2 to D3 (downlink). *Since the power supply of the BS is powerful, we assume that the downlink can always achieve larger transmission rate than the uplink.* Therefore, we can derive the instantaneous transmission rate for the cellular mode, denoted by $R_3\left(P_1(\widetilde{\boldsymbol{\nu}}_1), P_2(\widetilde{\boldsymbol{\nu}}_3)\right)$, as follows:

$$\begin{aligned} R_3\left(P_1(\widetilde{\boldsymbol{\nu}}_1), P_2(\widetilde{\boldsymbol{\nu}}_3)\right) &= \frac{BT}{2}\left[\log_2\left(1+P_1(\widetilde{\boldsymbol{\nu}}_1)\gamma_1\right)\right] + \frac{BT}{4}\min\left\{\log_2\left(1+P_2(\widetilde{\boldsymbol{\nu}}_3)\gamma_3\right), R_{\text{down}}\right\} \\ &= \frac{BT}{2}\left[\log_2\left(1+P_1(\widetilde{\boldsymbol{\nu}}_1)\gamma_1\right)\right] + \frac{BT}{4}\left[\log_2\left(1+P_2(\widetilde{\boldsymbol{\nu}}_3)\gamma_3\right)\right], \end{aligned} \quad (7)$$

where $\widetilde{\boldsymbol{\nu}}_3 \triangleq (\gamma_3, \theta)$ is defined as the QoS-based CSI for D2 with the cellular mode; R_{down} is the transmission rate of the channel from the BS to D3 (downlink). The FCSI for the cellular mode based D2D communication is defined as $\boldsymbol{\gamma}_c \triangleq (\gamma_1, \gamma_3)$.

From Eqs. (2) and (7), we can obtain the effective capacity of the D2D-cellular underlaying unit using the cellular mode, denoted by $C_3\left(P_1(\widetilde{\boldsymbol{\nu}}_1), P_2(\widetilde{\boldsymbol{\nu}}_3), \theta\right)$, as follows:

$$C_3\left(P_1(\widetilde{\boldsymbol{\nu}}_1), P_2(\widetilde{\boldsymbol{\nu}}_3), \theta\right) = -\frac{1}{\theta}\log\left(\mathbb{E}_{\boldsymbol{\gamma}_c}\left\{e^{-\theta R_3\left(P_1(\widetilde{\boldsymbol{\nu}}_1), P_2(\widetilde{\boldsymbol{\nu}}_3)\right)}\right\}\right) = -\frac{1}{\theta}\log\left(\mathbb{E}_{\boldsymbol{\gamma}_c}\left\{e^{-\beta\left(\log\left[(1+P_1(\widetilde{\boldsymbol{\nu}}_1)\gamma_1)^{\frac{1}{2}}+\log(1+P_2(\widetilde{\boldsymbol{\nu}}_3)\gamma_3)^{\frac{1}{4}}\right]\right)}\right\}\right). \quad (8)$$

C. The Sufficient Condition for D2D Devices to Use the D2D Communication

It is expected that with D2D communication, the wireless networks can achieve larger effective capacity than that of only using cellular communication. From the analyses of Sections II-B-2 and II-B-3, we can derive that if the SNR of the channel between D2 and D3 is larger than that between D2 and the BS, we can obtain that

$$C_2\left(P_1(\widetilde{\boldsymbol{\nu}}_1), P_2(\widetilde{\boldsymbol{\nu}}_2), \theta\right) > C_3\left(P_1(\widetilde{\boldsymbol{\nu}}_1), P_2(\widetilde{\boldsymbol{\nu}}_3), \theta\right). \quad (9)$$

With D2D communication, the maximum effective capacity can be obtained as follows:

$$C_{D2D} = \max\left\{C_1\left(P_1(\boldsymbol{\nu}), P_2(\boldsymbol{\nu}), \theta\right), C_2\left(P_1(\widetilde{\boldsymbol{\nu}}_1), P_2(\widetilde{\boldsymbol{\nu}}_2), \theta\right)\right\} \geq C_2\left(P_1(\widetilde{\boldsymbol{\nu}}_1), P_2(\widetilde{\boldsymbol{\nu}}_2), \theta\right). \quad (10)$$

The cellular and D2D devices share the same frequency band with the co-channel mode. If the co-channel interference is efficiently suppressed, the co-channel mode can achieve larger effective capacity than that of the orthogonal-channel mode. How to achieve the larger effective capacity between the co-channel mode and the orthogonal-channel is not the focus of this paper. However, we can still derive a *sufficient condition* to guarantee that the D2D-cellular underlay unit achieves larger effective capacity with D2D communication than that of only using cellular communication as follows:

$$\mathbb{E}_{\gamma_2}\left\{\gamma_2\right\} > \mathbb{E}_{\gamma_3}\left\{\gamma_3\right\}. \quad (11)$$

In the following two sections, we formulate the optimization problems to maximize the effective capacity of the co-channel mode and the orthogonal-channel mode, respectively. We solve the optimization problems to obtain the QoS-driven power allocation schemes corresponding to the co-channel mode and the orthogonal-channel mode, respectively.

III. QoS-Driven Power Allocation Schemes for the Co-Channel Mode

For the co-channel mode, we first formulate the optimization problem to maximize the effective capacity of the D2D-cellular underlaying unit. Then, we develop the QoS-driven power allocation scheme for the co-channel mode.

A. Power Allocation Optimization Problem for the Co-Channel Mode

For a fixed QoS exponent θ, our goal is to maximize $C_1\left(P_1(\boldsymbol{\nu}), P_2(\boldsymbol{\nu}), \theta\right)$ under the total power constraint $\mathbb{E}_{\boldsymbol{\gamma}}[P_1(\boldsymbol{\nu})+P_2(\boldsymbol{\nu})] \leq \overline{P}$, where $\overline{P}$ is the total power constraint for the D2D-cellular underlaying unit. Therefore, we can formulate the optimization problem for the co-channel mode as follows:

$$\boldsymbol{P1}: \quad \max_{\left(P_1(\boldsymbol{\nu}), P_2(\boldsymbol{\nu})\right)}\left\{C_1\left(P_1(\boldsymbol{\nu}), P_2(\boldsymbol{\nu}), \theta\right)\right\} = \min_{\left(P_1(\boldsymbol{\nu}), P_2(\boldsymbol{\nu})\right)}\left\{\mathbb{E}_{\boldsymbol{\gamma}}\left[e^{-\theta R_1\left(P_1(\boldsymbol{\nu}), P_2(\boldsymbol{\nu})\right)}\right]\right\}$$

$$\text{s.t.}: \quad 1).\ \mathbb{E}_{\boldsymbol{\gamma}}[P_1(\boldsymbol{\nu}) + P_2(\boldsymbol{\nu})] \leq \overline{P}; \quad (12)$$

$$2).\ P_1(\boldsymbol{\nu}) \geq 0 \text{ and } P_2(\boldsymbol{\nu}) \geq 0. \quad (13)$$

It is highly desirable that ***P1*** is a convex optimization problem. To study the convexity/concavity of ***P1***, we give the following lemma to illustrate the relationship between the convexity/concavity of ***P1*** and the convexity/concavity of $R_1\left(P_1(\boldsymbol{\nu}), P_2(\boldsymbol{\nu})\right)$.

Lemma 1: If $R_1\left(P_1(\boldsymbol{\nu}), P_2(\boldsymbol{\nu})\right)$ is strictly concave on the space spanned by $(P_1(\boldsymbol{\nu}), P_2(\boldsymbol{\nu}))$, ***P1*** is a strictly convex optimization problem.

Proof: Using the following three facts: 1). $\mathbb{E}_{\boldsymbol{\gamma}}(\cdot)$ is a linear function; 2). $e^{-\theta x}$ is convex if x is concave; 3). $R_1\left(P_1(\boldsymbol{\nu}), P_2(\boldsymbol{\nu})\right)$ is strictly concave on the space spanned by $(P_1(\boldsymbol{\nu}), P_2(\boldsymbol{\nu}))$, we can show that $\min_{(P_1(\boldsymbol{\nu}),P_2(\boldsymbol{\nu}))}\left\{\mathbb{E}_{\boldsymbol{\gamma}}\left[e^{-\theta R_1(P_1(\boldsymbol{\nu}),P_2(\boldsymbol{\nu}))}\right]\right\}$ is strictly convex on the space spanned by $(P_1(\boldsymbol{\nu}), P_2(\boldsymbol{\nu}))$. On the other hand, $\mathbb{E}_{\boldsymbol{\gamma}}[P_1(\boldsymbol{\nu}) + P_2(\boldsymbol{\nu})]$, $P_1(\boldsymbol{\nu})$, and $P_2(\boldsymbol{\nu})$ are all linear functions on the space spanned by $(P_1(\boldsymbol{\nu}), P_2(\boldsymbol{\nu}))$. Therefore, Lemma 1 follows. ■

Based on Lemma 1, we can study the convexity/concavity of $R_1\left(P_1(\boldsymbol{\nu}), P_2(\boldsymbol{\nu})\right)$ to show the convexity/concavity of ***P1***. However, notice that $R_1\left(P_1(\boldsymbol{\nu}), P_2(\boldsymbol{\nu})\right)$ is not strictly concave on the space spanned by $(P_1(\boldsymbol{\nu}), P_2(\boldsymbol{\nu}))$, we need to convert $R_1\left(P_1(\boldsymbol{\nu}), P_2(\boldsymbol{\nu})\right)$ to be a new concave function while keeping that $R_1\left(P_1(\boldsymbol{\nu}), P_2(\boldsymbol{\nu})\right)$ is very close to the new concave function.

Let us rewrite $R_1(P_1(\boldsymbol{\nu}), P_2(\boldsymbol{\nu}))$ as follows:

$$\begin{aligned}
&R_1(P_1(\boldsymbol{\nu}), P_2(\boldsymbol{\nu})) \\
&= BT\left(\log_2\left[P_1(\boldsymbol{\nu})\gamma_1 + P_2(\boldsymbol{\nu})\gamma_3 + \sigma^2\right] - \log_2\left[\sigma^2 + P_2(\boldsymbol{\nu})\gamma_3\right]\right) \\
&+ BT\left(\log_2\left[P_2(\boldsymbol{\nu})\gamma_2 + P_2(\boldsymbol{\nu})\gamma_4 + \sigma^2\right] - \log_2\left[\sigma^2 + P_1(\boldsymbol{\nu})\gamma_4\right]\right).
\end{aligned} \tag{14}$$

Eq. (14) shows that $R_1(P_1(\boldsymbol{\nu}), P_2(\boldsymbol{\nu}))$ is a sum of functions with the d.c. structure, which is defined as the difference of two concave functions. Generally, optimization problems which have the d.c. structure are well-known as NP-hard problems and often difficult to directly derive the global optimal solutions [18]. However, with proper relaxation and transformation, some optimization problems which have the d.c. structure can be converted to strictly convex optimization problems [19]–[21].

To convert the function $R_1(P_1(\boldsymbol{\nu}), P_2(\boldsymbol{\nu}))$ to be a concave function, the function $\log_2(1+z)$ where $z \geq 0$ is of our interest. A great deal of lower-bounds which are of concave property can be used to relax $\log_2(1+z)$. We need to derive and apply a very tight relaxation of $R_1(P_1(\boldsymbol{\nu}), P_2(\boldsymbol{\nu}))$ to obtain the optimal solution of problem ***P1***. The Lemma 2 that follows gives a very tight lower-bound for the function $\log_2(1+z)$.

Lemma 2: A very tight lower-bound of $\log_2(1+z)$ where $z \geq 0$ is given by

$$\alpha \log_2 z + \beta. \tag{15}$$

The lower-bounds coefficients α and β are chosen as

$$\begin{cases} \alpha = \frac{z_0}{1+z_0}; \\ \beta = \log_2(1+z_0) - \frac{z_0}{1+z_0}\log_2 z_0, \end{cases} \tag{16}$$

where $z_0 \in [0, \infty)$ is a positive real-valued number. At $z = z_0$, the lower-bound $\alpha \log_2 z + \beta$ is equal to $\log_2(1+z)$.

Figure 2 shows the tightness between $\log_2(1+z)$ and its lower-bound $\alpha \log_2 z + \beta$. We can see that the lower-bound $\alpha \log_2 z + \beta$ is very close to $\log_2(1+z)$ and attains $\log_2(1+z)$ at z_0.

Based on Lemma 2, we can relax $R_1\left(P_1(\boldsymbol{\nu}), P_2(\boldsymbol{\nu})\right)$ as Eq. (17), where α_1 and β_1 are the lower-bounds coefficients corresponding to $f_1\left(P_1(\boldsymbol{\nu}), P_2(\boldsymbol{\nu})\right)$; α_2 and β_2 are the lower-bounds coefficients corresponding to $f_2\left(P_1(\boldsymbol{\nu}), P_2(\boldsymbol{\nu})\right)$. In Eq. (17), we use the logarithmic change of variables $\widetilde{P}_1(\boldsymbol{\nu}) = \log(P_1(\boldsymbol{\nu}))$ and $\widetilde{P}_2(\boldsymbol{\nu}) = \log(P_2(\boldsymbol{\nu}))$ to convert the function with the d.c. structure to the function with the log-sum-exp structure [22, Sec. 3.1.5], which is defined as the logarithm for the sum of a number of exponential functions. The general form of the function with the log-sum-exp structure can be derived as follows:

$$g(x, y) = \log\left(a + be^x + ce^y\right), \tag{18}$$

where a, b, and c are constant coefficients. We have 1). $g(\widetilde{P}_1(\boldsymbol{\nu}), 0) = \log(\sigma^2 + e^{\widetilde{P}_1(\boldsymbol{\nu})}\gamma_3)$, where $a = \sigma^2$, $b = \gamma_3$, and $c = 0$; 2). $g(0, \widetilde{P}_2(\boldsymbol{\nu})) = \log(\sigma^2 + e^{\widetilde{P}_2(\boldsymbol{\nu})}\gamma_4)$, where $a = \sigma^2$, $b = 0$, and $c = \gamma_4$. The Hessian matrix of $g(x, y)$, denoted by M_g, can be derived as follows:

$$M_g = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} = \begin{bmatrix} \frac{e^x b + e^x e^y bc}{(a + e^x b + e^y c)^2} & \frac{-e^x e^y bc}{(a + e^x b + e^y c)^2} \\ \frac{-e^x e^y bc}{(a + e^x b + e^y c)^2} & \frac{e^y c + e^x e^y bc}{(a + e^x b + e^y c)^2} \end{bmatrix}. \tag{19}$$

Then, we have $g_{11} > 0$ and $(g_{11}g_{22} - g_{12}g_{21}) > 0$. By using the Sylvesters criterion [23], we can show that $g(x, y)$ is a strictly convex function on the space spanned by (x, y). Thus, $g(\widetilde{P}_1(\boldsymbol{\nu}), 0)$ and $g(0, \widetilde{P}_2(\boldsymbol{\nu}))$ are strictly convex on the space spanned by $(\widetilde{P}_1(\boldsymbol{\nu}), \widetilde{P}_2(\boldsymbol{\nu}))$.

Then, using the property that a nonnegative, nonzero weighted sum of strictly convex (concave) functions is strictly convex (concave) [22, Sec. 3.2.1], we can show that $R_1(\widetilde{P}_1(\boldsymbol{\nu}), \widetilde{P}_2(\boldsymbol{\nu}))$ is a strictly concave function on the space spanned by $(\widetilde{P}_1(\boldsymbol{\nu}), \widetilde{P}_2(\boldsymbol{\nu}))$.

Now, we can convert problem ***P1*** to be a new problem, denoted by ***P2***, as follows:

$$\begin{aligned}
\boldsymbol{P2}: &\min_{\left(\widetilde{P}_1(\boldsymbol{\nu}), \widetilde{P}_2(\boldsymbol{\nu})\right)} \left\{\mathbb{E}_{\boldsymbol{\gamma}}\left[e^{-\theta R_1\left(\widetilde{P}_1(\boldsymbol{\nu}), \widetilde{P}_2(\boldsymbol{\nu})\right)}\right]\right\} \\
&= \min_{\left(\widetilde{P}_1(\boldsymbol{\nu}), \widetilde{P}_2(\boldsymbol{\nu})\right)} \left\{\mathbb{E}_{\boldsymbol{\gamma}}\Bigg(\exp\Bigg(-\beta\Big[\alpha_1\widetilde{P}_1(\boldsymbol{\nu}) + \alpha_2\widetilde{P}_2(\boldsymbol{\nu}) \right. \right. \\
&\left. \left. -\alpha_1 \log\left(\sigma^2 + e^{\widetilde{P}_2(\boldsymbol{\nu})}\gamma_3\right) - \alpha_2 \log\left(\sigma^2 + e^{\widetilde{P}_1(\boldsymbol{\nu})}\gamma_4\right)\Big]\Bigg)\Bigg)\right\} \\
&\text{s.t.}: \quad \mathbb{E}_{\boldsymbol{\gamma}}\left[e^{\widetilde{P}_1(\boldsymbol{\nu})} + e^{\widetilde{P}_2(\boldsymbol{\nu})}\right] \leq \overline{P}.
\end{aligned} \tag{20}$$

Because the objective function of ***P2*** is strictly convex and the function on the left-hand side of the inequality constraint (Eq. (20)) is convex on the space spanned by $(\widetilde{P}_1(\boldsymbol{\nu}), \widetilde{P}_2(\boldsymbol{\nu}))$, we can prove that ***P2*** is a strictly convex optimization problem.

B. QoS-Driven Optimal Power Allocations for D1 and D2 Under the Co-Channel Mode

We use the powerful Lagrangian method to derive the QoS-driven power allocations for D1 and D2 with the co-channel mode. The Lagrangian function of ***P2***, denoted by

$$
\begin{aligned}
R_1\left(P_1(\boldsymbol{\nu}),P_2(\boldsymbol{\nu})\right) &= BT\left[\alpha_1\log_2\left(\frac{P_1(\boldsymbol{\nu})\gamma_1}{\sigma^2+P_2(\boldsymbol{\nu})\gamma_3}\right)+\beta_1+\alpha_2\log_2\left(\frac{P_2(\boldsymbol{\nu})\gamma_2}{\sigma^2+P_1(\boldsymbol{\nu})\gamma_4}\right)+\beta_2\right]\\
&= \frac{BT}{\log 2}\Big[\alpha_1\log\left(P_1(\boldsymbol{\nu})\right)+\alpha_1\log\left(\gamma_1\right)-\alpha_1\log\left(\sigma^2+P_2(\boldsymbol{\nu})\gamma_3\right)+\beta_1\log 2\\
&\quad+\alpha_2\log\left(P_2(\boldsymbol{\nu})\right)+\alpha_2\log(\gamma_2)-\alpha_2\log\left(\sigma^2+P_1(\boldsymbol{\nu})\gamma_4\right)+\beta_2\log 2\Big]\\
&= \frac{BT}{\log 2}\Big[\alpha_1\widetilde{P}_1(\boldsymbol{\nu})+\alpha_1\log\left(\gamma_1\right)-\alpha_1\log\left(\sigma^2+e^{\widetilde{P}_2(\boldsymbol{\nu})}\gamma_3\right)+\beta_1\log 2\\
&\quad+\alpha_2\widetilde{P}_2(\boldsymbol{\nu})+\alpha_2\log(\gamma_2)-\alpha_2\log\left(\sigma^2+e^{\widetilde{P}_1(\boldsymbol{\nu})}\gamma_4\right)+\beta_2\log 2\Big]=R_1(\widetilde{P}_1(\boldsymbol{\nu}),\widetilde{P}_2(\boldsymbol{\nu})),
\end{aligned}\tag{17}
$$

$L_1(\widetilde{P}_1(\boldsymbol{\nu}),\widetilde{P}_2(\boldsymbol{\nu});\lambda)$, can be formulated as follows:

$$
\begin{aligned}
&L_1\left(\widetilde{P}_1(\boldsymbol{\nu}),\widetilde{P}_2(\boldsymbol{\nu});\lambda\right)=\mathbb{E}_{\boldsymbol{\gamma}}\left\{J_1\left(\widetilde{P}_1(\boldsymbol{\nu}),\widetilde{P}_2(\boldsymbol{\nu});\lambda\right)\right\}\\
&\triangleq\mathbb{E}_{\boldsymbol{\gamma}}\Bigg\{\exp\Bigg(-\beta\Big[\alpha_1\widetilde{P}_1(\boldsymbol{\nu})-\alpha_1\log\left(\sigma^2+e^{\widetilde{P}_2(\boldsymbol{\nu})}\gamma_3\right)\\
&\qquad+\alpha_2\widetilde{P}_2(\boldsymbol{\nu})-\alpha_2\log\left(\sigma^2+e^{\widetilde{P}_1(\boldsymbol{\nu})}\gamma_4\right)\Big]\Bigg)\Bigg\}\\
&\qquad+\lambda\left(\mathbb{E}_{\boldsymbol{\gamma}}\left[e^{\widetilde{P}_1(\boldsymbol{\nu})}+e^{\widetilde{P}_2(\boldsymbol{\nu})}\right]-\overline{P}\right),
\end{aligned}\tag{21}
$$

where the Lagrangian multiplier λ is associated with the constraint given by Eq. (20). We denote the optimal Lagrangian multiplier as λ^*. Since ***P2*** is a strictly convex optimization problem, the optimal Lagrangian multiplier λ^* is also the optimal solution of ***P2***'s dual problem ***P2-Dual***, which is formulated as follows:

$$
\begin{aligned}
\textbf{\textit{P2-Dual}}: \quad &\max_{\lambda}\left\{\widetilde{L}_1\left(\lambda\right)\right\}\\
&\text{s.t.}: \ \lambda\geq 0,
\end{aligned}\tag{22}
$$

where $\widetilde{L}_1\left(\lambda\right)$ is the Lagrangian dual function defined by

$$
\widetilde{L}_1\left(\lambda\right)=\min_{\left(\widetilde{P}_1(\boldsymbol{\nu}),\widetilde{P}_2(\boldsymbol{\nu})\right)}\left\{L_1\left(\widetilde{P}_1(\boldsymbol{\nu}),\widetilde{P}_2(\boldsymbol{\nu});\lambda\right)\right\}.\tag{23}
$$

Following convex optimization theory, $\widetilde{L}_1\left(\lambda\right)$ is a concave function over λ. Thus, we can use the subgradient method [22] to track the optimal Lagrangian multiplier λ^* as follows:

$$
\lambda^*=\left[\lambda^*+\epsilon_\lambda\left(\mathbb{E}_{\boldsymbol{\gamma}}\left[e^{\widetilde{P}_1(\boldsymbol{\nu})}+e^{\widetilde{P}_2(\boldsymbol{\nu})}\right]-\overline{P}\right)\right]^+,\tag{24}
$$

where ϵ_λ is a positive real-valued number arbitrarily close to 0; $[\widetilde{y}]^+$ represents the maximum value between $\widetilde{y}$ and 0. Because of the concavity of $\widetilde{L}_1\left(\lambda\right)$, the iteration of Eq. (24) will converge to the optimal λ^*. Then, taking the obtained optimal λ^* into

$$
\begin{cases}
\frac{\partial L_1\left(\widetilde{P}_1(\boldsymbol{\nu}),\widetilde{P}_2(\boldsymbol{\nu});\lambda\right)}{\partial\widetilde{P}_1(\boldsymbol{\nu})}\big|_{\lambda=\lambda^*}=0;\\
\frac{\partial L_1\left(\widetilde{P}_1(\boldsymbol{\nu}),\widetilde{P}_2(\boldsymbol{\nu});\lambda\right)}{\partial\widetilde{P}_2(\boldsymbol{\nu})}\big|_{\lambda=\lambda^*}=0,
\end{cases}\tag{25}
$$

we can derive the optimal solutions for problem ***P2*** as $\widetilde{P}_1^*(\boldsymbol{\nu})$ and $\widetilde{P}_2^*(\boldsymbol{\nu})$ corresponding to $\widetilde{P}_1(\boldsymbol{\nu})$ and $\widetilde{P}_2(\boldsymbol{\nu})$, respectively. Using the exponential variation $P_1^*(\boldsymbol{\nu})=\exp(\widetilde{P}_1^*(\boldsymbol{\nu}))$ and $P_2^*(\boldsymbol{\nu})=\exp(\widetilde{P}_2^*(\boldsymbol{\nu}))$, we can derive the optimal solutions for problem ***P1*** as $P_1^*(\boldsymbol{\nu})$ and $P_2^*(\boldsymbol{\nu})$ corresponding to $P_1(\boldsymbol{\nu})$ and $P_2(\boldsymbol{\nu})$, respectively, which are also the QoS-driven power allocations for D1 and D2, respectively. Then, taking the QoS-driven power allocations $P_1^*(\boldsymbol{\nu})$ and $P_2^*(\boldsymbol{\nu})$ into Eq. (2), we can derive the maximum effective capacity for the co-channel mode.

C. Suboptimal Power Allocation Schemes Under the Co-Channel Mode

In this subsection, we derive two suboptimal power allocation schemes with the co-channel mode: 1). The power allocation scheme for the co-channel mode without QoS provisioning; 2). The power allocation scheme for the co-channel mode with fixed QoS exponent ($\theta\to\theta_0$).

- The power allocation scheme for the co-channel mode without QoS provisioning: Without QoS provisioning, the power allocations for D1 and D2, denoted by $P_1^o(\boldsymbol{\nu})$ and $P_2^o(\boldsymbol{\nu})$, respectively, can be obtained as follows:

$$
\begin{cases}
P_1^o(\boldsymbol{\nu})=\lim\limits_{\theta\to 0}P_1^*(\boldsymbol{\nu});\\
P_2^o(\boldsymbol{\nu})=\lim\limits_{\theta\to 0}P_2^*(\boldsymbol{\nu}).
\end{cases}\tag{26}
$$

  Plugging Eq. (26) into Eq. (2), we can derive the corresponding effective capacity as follows:

$$
\begin{aligned}
&C_1\left(P_1^o(\boldsymbol{\nu}),P_2^o(\boldsymbol{\nu}),\theta\right)\\
&=-\frac{1}{\theta}\log\Bigg(\mathbb{E}_{\boldsymbol{\gamma}}\Bigg\{\exp\Bigg(-\beta\Bigg[\log\left(1+\frac{P_1^o(\boldsymbol{\nu})\gamma_1}{\sigma^2+P_2^o(\boldsymbol{\nu})\gamma_3}\right)\\
&\qquad+\log\left(1+\frac{P_2^o(\boldsymbol{\nu})\gamma_2}{\sigma^2+P_1^o(\boldsymbol{\nu})\gamma_4}\right)\Bigg]\Bigg)\Bigg\}\Bigg).
\end{aligned}\tag{27}
$$

- The power allocation scheme for the co-channel mode with fixed QoS exponent ($\theta\to\theta_0$): In this scheme, we use the power allocations corresponding to that $\theta\to\theta_0$. Clearly, this scheme is suboptimal when the required QoS exponent of the traffic is not equal to θ_0. The power allocations of this scheme, denoted by $P_1^s(\boldsymbol{\nu})$ and $P_2^s(\boldsymbol{\nu})$, respectively, can be obtained as follows:

$$
\begin{cases}
P_1^s(\boldsymbol{\nu})=\lim\limits_{\theta\to\theta_0}P_1^*(\boldsymbol{\nu});\\
P_2^s(\boldsymbol{\nu})=\lim\limits_{\theta\to\theta_0}P_2^*(\boldsymbol{\nu}).
\end{cases}\tag{28}
$$

  Plugging Eq. (28) into Eq. (2), we can derive the corre-

sponding effective capacity as follows:

$$C_1\left(P_1^s(\boldsymbol{\nu}),P_2^s(\boldsymbol{\nu}),\theta\right) = -\frac{1}{\theta}\log\left(\mathbb{E}_{\boldsymbol{\gamma}}\left\{\exp\left(-\beta\left[\log\left(1+\frac{P_1^s(\boldsymbol{\nu})\gamma_1}{\sigma^2+P_2^s(\boldsymbol{\nu})\gamma_3}\right)+\log\left(1+\frac{P_2^s(\boldsymbol{\nu})\gamma_2}{\sigma^2+P_1^s(\boldsymbol{\nu})\gamma_4}\right)\right]\right)\right\}\right). \quad (29)$$

IV. QoS-Driven Power Allocation Schemes for The Orthogonal-Channel Mode

For the orthogonal-channel mode, we first directly formulate the effective capacity maximization problem as a convex optimization problem. Then, we develop the QoS-driven power allocation scheme for the orthogonal-channel mode.

A. Power Allocation Optimization Problem for the Orthogonal-Channel Mode

For a fixed QoS exponent θ, we can formulate the optimization problem for the orthogonal-channel mode as follows:

$$\begin{aligned}\boldsymbol{P3}: \quad & \min_{(P_1(\widetilde{\boldsymbol{\nu}}_1),P_2(\widetilde{\boldsymbol{\nu}}_2))}\left\{\mathbb{E}_{\widetilde{\boldsymbol{\gamma}}}\left[e^{-\theta R_2(P_1(\widetilde{\boldsymbol{\nu}}_1),P_2(\widetilde{\boldsymbol{\nu}}_2))}\right]\right\}\\ = \quad & \min_{(P_1(\widetilde{\boldsymbol{\nu}}_1),P_2(\widetilde{\boldsymbol{\nu}}_2))}\left\{\mathbb{E}_{\widetilde{\boldsymbol{\gamma}}}\left(\left[(1+P_1(\widetilde{\boldsymbol{\nu}}_1)\gamma_1)(1+P_2(\widetilde{\boldsymbol{\nu}}_2)\gamma_2)\right]^{-\frac{\beta}{2}}\right)\right\}\end{aligned}$$

$$\text{s.t.}: \quad 1).\ \mathbb{E}_{\widetilde{\boldsymbol{\gamma}}}\left[P_1(\widetilde{\boldsymbol{\nu}}_1)+P_2(\widetilde{\boldsymbol{\nu}}_2)\right]\le\overline{P}; \quad (30)$$

$$2).\ P_1\left(\widetilde{\boldsymbol{\nu}}_1\right)\ge 0; \quad (31)$$

$$3).\ P_2\left(\widetilde{\boldsymbol{\nu}}_2\right)\ge 0. \quad (32)$$

It is clear that $R_2(P_1(\widetilde{\boldsymbol{\nu}}_1),P_2(\widetilde{\boldsymbol{\nu}}_2))$ is strictly concave on the space spanned by $(P_1(\widetilde{\boldsymbol{\nu}}_1),P_2(\widetilde{\boldsymbol{\nu}}_2))$. On the other hand, the functions on the left-hand side of all inequality constraints (Eqs. (30), (31), and (32)) are all linear functions. Therefore, we can obtain that ***P3*** is a strictly convex optimization problem.

B. QoS-Driven Optimal Power Allocations for D1 and D2 Under the Orthogonal-Channel Mode

In order to obtain the optimal solutions of ***P3***, we first construct the Lagrangian function of ***P3***, denoted by $L_2(P_1(\widetilde{\boldsymbol{\nu}}_1),P_2(\widetilde{\boldsymbol{\nu}}_2))$, as follows:

$$\begin{aligned}&L_2(P_1(\widetilde{\boldsymbol{\nu}}_1),P_2(\widetilde{\boldsymbol{\nu}}_2)) = \mathbb{E}_{\widetilde{\boldsymbol{\gamma}}}\{J_2(P_1(\widetilde{\boldsymbol{\nu}}_1),P_2(\widetilde{\boldsymbol{\nu}}_2))\}\\ &\triangleq \mathbb{E}_{\widetilde{\boldsymbol{\gamma}}}\left(\left[(1+P_1(\widetilde{\boldsymbol{\nu}}_1)\gamma_1)(1+P_2(\widetilde{\boldsymbol{\nu}}_2)\gamma_2)\right]^{-\frac{\beta}{2}}\right)-\widetilde{\mu}_1P_1\left(\widetilde{\boldsymbol{\nu}}_1\right)\\ &\quad -\widetilde{\mu}_2P_2\left(\widetilde{\boldsymbol{\nu}}_2\right)+\widetilde{\lambda}\left[\mathbb{E}_{\widetilde{\boldsymbol{\gamma}}}\left(P_1(\widetilde{\boldsymbol{\nu}}_1)+P_2(\widetilde{\boldsymbol{\nu}}_2)\right)-\overline{P}\right],\end{aligned} \quad (33)$$

where $\widetilde{\lambda}$, $\widetilde{\mu}_1$, and $\widetilde{\mu}_2$ are Lagrangian multipliers associated with the constraints given by Eqs. (30), (31), and (32), respectively. Then, taking the derivatives of $J_2(P_1(\widetilde{\boldsymbol{\nu}}_1),P_2(\widetilde{\boldsymbol{\nu}}_2))$ with respective to $P_1(\widetilde{\boldsymbol{\nu}}_1)$ and $P_2(\widetilde{\boldsymbol{\nu}}_2)$, and setting the derivatives equal to zero, respectively, we can obtain Eq. (34), where $p_\Gamma(\widetilde{\boldsymbol{\gamma}})$ is the probability density function (PDF) of the channel SNR for the orthogonal-channel mode. Using the complementary slackness condition [22, Sec. 5.5.2], we can obtain $[\widetilde{\mu}_1P_1(\widetilde{\boldsymbol{\nu}}_1)]=0$ and $[\widetilde{\mu}_2P_2(\widetilde{\boldsymbol{\nu}}_2)]=0$. Thus, we can derive the optimal solutions for ***P3*** by analyzing the following two cases. **Case I:** $P_1(\widetilde{\boldsymbol{\nu}}_1)>0$ and $P_2(\widetilde{\boldsymbol{\nu}}_2)>0$. For this case, with the complementary slackness condition, we have $\widetilde{\mu}_1=0$ and $\widetilde{\mu}_2=0$. Then, taking $\widetilde{\mu}_1=0$ and $\widetilde{\mu}_2=0$ into Eq. (34), and setting

$$\frac{\partial J_2(P_1(\widetilde{\boldsymbol{\nu}}_1),P_2(\widetilde{\boldsymbol{\nu}}_2))}{\partial P_1(\widetilde{\boldsymbol{\nu}}_1)}=\frac{\partial J_2(P_1(\widetilde{\boldsymbol{\nu}}_1),P_2(\widetilde{\boldsymbol{\nu}}_2))}{\partial P_2(\widetilde{\boldsymbol{\nu}}_1)}, \quad (35)$$

we can obtain

$$\frac{\gamma_1}{1+P_1(\widetilde{\boldsymbol{\nu}}_1)\gamma_1}=\frac{\gamma_2}{1+P_2(\widetilde{\boldsymbol{\nu}}_1)\gamma_2}. \quad (36)$$

Taking Eq. (36) into Eq. (34), we have

$$\begin{cases}P_1(\widetilde{\boldsymbol{\nu}}_1)=\frac{1}{\left(\frac{2\widetilde{\lambda}}{\beta}\right)^{\frac{1}{\beta+1}}(\gamma_1\gamma_2)^{\frac{\beta}{2(\beta+1)}}}-\frac{1}{\gamma_1};\\ P_2(\widetilde{\boldsymbol{\nu}}_2)=\frac{1}{\left(\frac{2\widetilde{\lambda}}{\beta}\right)^{\frac{1}{\beta+1}}(\gamma_1\gamma_2)^{\frac{\beta}{2(\beta+1)}}}-\frac{1}{\gamma_2}.\end{cases} \quad (37)$$

Case II: $[P_1(\widetilde{\boldsymbol{\nu}}_1)P_2(\widetilde{\boldsymbol{\nu}}_2)]=0$. The equality $[P_1(\widetilde{\boldsymbol{\nu}}_1)P_2(\widetilde{\boldsymbol{\nu}}_2)]=0$ implies that at least one of $P_1(\widetilde{\boldsymbol{\nu}}_1)$ and $P_2(\widetilde{\boldsymbol{\nu}}_2)$ is equal to zero, leading to three different scenarios described as follows:

- **Scenario I:** $P_1(\widetilde{\boldsymbol{\nu}}_1)=0$ and $P_2(\widetilde{\boldsymbol{\nu}}_2)>0$. In this scenario, we have $\widetilde{\mu}_2=0$. Then, by solving $[\partial J_2(P_1(\widetilde{\boldsymbol{\nu}}_1),P_2(\widetilde{\boldsymbol{\nu}}_2))/\partial P_2(\widetilde{\boldsymbol{\nu}}_2)]|_{P_1(\widetilde{\boldsymbol{\nu}}_1)=0}=0$, we can obtain

$$\begin{cases}P_1(\widetilde{\boldsymbol{\nu}}_1)=0;\\ P_2(\widetilde{\boldsymbol{\nu}}_2)=\frac{1}{\left(\frac{2\widetilde{\lambda}}{\beta}\right)^{\frac{2}{\beta+2}}\gamma_2^{\frac{\beta}{\beta+2}}}-\frac{1}{\gamma_2}.\end{cases} \quad (38)$$

- **Scenario II:** $P_1(\widetilde{\boldsymbol{\nu}}_1)>0$ and $P_2(\widetilde{\boldsymbol{\nu}}_2)=0$. In this scenario, we have $\widetilde{\mu}_1=0$. Then, by solving $[\partial J_2(P_1(\widetilde{\boldsymbol{\nu}}_1),P_2(\widetilde{\boldsymbol{\nu}}_2))/\partial P_1(\widetilde{\boldsymbol{\nu}}_1)]|_{P_2(\widetilde{\boldsymbol{\nu}}_2)=0}=0$, we can obtain

$$\begin{cases}P_1(\widetilde{\boldsymbol{\nu}}_1)=\frac{1}{\left(\frac{2\widetilde{\lambda}}{\beta}\right)^{\frac{2}{\beta+2}}\gamma_1^{\frac{\beta}{\beta+2}}}-\frac{1}{\gamma_2};\\ P_2(\widetilde{\boldsymbol{\nu}}_2)=0.\end{cases} \quad (39)$$

- **Scenario III:** $P_1(\widetilde{\boldsymbol{\nu}}_1)=0$ and $P_2(\widetilde{\boldsymbol{\nu}}_2)=0$. In this scenario, there is no strategy that can satisfy the strict inequalities $P_1(\widetilde{\boldsymbol{\nu}}_1)>0$ and $P_2(\widetilde{\boldsymbol{\nu}}_2)>0$.

Based on the optimal power allocations for **Case I** and **Case II**, we develop the following scheme to obtain the QoS-driven power allocations for the orthogonal-channel mode as follows (for the orthogonal-channel mode, the optimal power allocations for D1 and D2 are denoted by $P_1^*(\widetilde{\boldsymbol{\nu}}_1)$ and $P_2^*(\widetilde{\boldsymbol{\nu}}_2)$, respectively):

QoS-driven power allocation scheme for the orthogonal-channel mode:

1) Calculating $P_1(\widetilde{\boldsymbol{\nu}}_1)$ and $P_2(\widetilde{\boldsymbol{\nu}}_2)$ by using Eq. (37);
2) **If** ($P_1(\widetilde{\boldsymbol{\nu}}_1)>0$ and $P_2(\widetilde{\boldsymbol{\nu}}_2)>0$)
3) $P_1^*(\widetilde{\boldsymbol{\nu}}_1)=P_1(\widetilde{\boldsymbol{\nu}}_1)$ and $P_2^*(\widetilde{\boldsymbol{\nu}}_2)=P_2(\widetilde{\boldsymbol{\nu}}_2)$;
4) **Else if** ($P_1(\widetilde{\boldsymbol{\nu}}_1)\le 0$ and $P_2(\widetilde{\boldsymbol{\nu}}_2)>0$)
5) Calculating $P_1(\widetilde{\boldsymbol{\nu}}_1)$ and $P_2(\widetilde{\boldsymbol{\nu}}_2)$ by using Eq. (38);
6) **Else if** ($P_2(\widetilde{\boldsymbol{\nu}}_2)>0$)

$$\begin{cases} \frac{\partial J_2(P_1(\widetilde{\boldsymbol{\nu}}_1),P_2(\widetilde{\boldsymbol{\nu}}_2))}{\partial P_1(\widetilde{\boldsymbol{\nu}}_1)} = -\frac{\beta\gamma_1}{2}\left(1+P_1(\widetilde{\boldsymbol{\nu}}_1)\gamma_1\right)^{-\frac{\beta}{2}-1}\left(1+P_2(\widetilde{\boldsymbol{\nu}}_2)\gamma_2\right)^{-\frac{\beta}{2}} p_\Gamma(\widetilde{\boldsymbol{\gamma}}) - \widetilde{\mu}_1 + \widetilde{\lambda} p_\Gamma(\widetilde{\boldsymbol{\gamma}}) = 0; \\ \frac{\partial J_2(P_1(\widetilde{\boldsymbol{\nu}}_1),P_2(\widetilde{\boldsymbol{\nu}}_2))}{\partial P_2(\widetilde{\boldsymbol{\nu}}_1)} = -\frac{\beta\gamma_2}{2}\left(1+P_2(\widetilde{\boldsymbol{\nu}}_1)\gamma_1\right)^{-\frac{\beta}{2}-1}\left(1+P_1(\widetilde{\boldsymbol{\nu}}_2)\gamma_2\right)^{-\frac{\beta}{2}} p_\Gamma(\widetilde{\boldsymbol{\gamma}}) - \widetilde{\mu}_2 + \widetilde{\lambda} p_\Gamma(\widetilde{\boldsymbol{\gamma}}) = 0 \end{cases} \tag{34}$$

7) $P_1^*(\widetilde{\boldsymbol{\nu}}_1) = P_1(\widetilde{\boldsymbol{\nu}}_1)$ and $P_2^*(\widetilde{\boldsymbol{\nu}}_2) = P_2(\widetilde{\boldsymbol{\nu}}_2)$;
8) **Else**
9) $P_1^*(\widetilde{\boldsymbol{\nu}}_1) = 0$ and $P_2^*(\widetilde{\boldsymbol{\nu}}_2) = 0$;
10) **end**
11) **end**
12) **Else if** $(P_1(\widetilde{\boldsymbol{\nu}}_1) > 0$ and $P_2(\widetilde{\boldsymbol{\nu}}_2) \leq 0)$
13) Calculating $P_1(\widetilde{\boldsymbol{\nu}}_1)$ and $P_2(\widetilde{\boldsymbol{\nu}}_2)$ by using Eq. (39);
14) **Else if** $(P_1(\widetilde{\boldsymbol{\nu}}_2) > 0)$
15) $P_1^*(\widetilde{\boldsymbol{\nu}}_1) = P_1(\widetilde{\boldsymbol{\nu}}_1)$ and $P_2^*(\widetilde{\boldsymbol{\nu}}_2) = P_2(\widetilde{\boldsymbol{\nu}}_2)$;
16) **Else**
17) $P_1^*(\widetilde{\boldsymbol{\nu}}_1) = 0$ and $P_2^*(\widetilde{\boldsymbol{\nu}}_2) = 0$;
18) **end**
19) **end**
20) **Else**
21) $P_1^*(\widetilde{\boldsymbol{\nu}}_1) = 0$ and $P_2^*(\widetilde{\boldsymbol{\nu}}_2) = 0$;
22) **end**

Then, taking the QoS-driven power allocations $P_1^*(\widetilde{\boldsymbol{\nu}}_1)$ and $P_2^*(\widetilde{\boldsymbol{\nu}}_2)$ into Eq. (2), we can derive the maximum effective capacity for the orthogonal-channel mode.

C. Suboptimal Power Allocation Schemes Under the Orthogonal-Channel Mode

For the orthogonal-channel mode, we propose two suboptimal power allocation schemes as follows:

- The power allocation scheme for the orthogonal-channel mode without QoS provisioning: Without QoS provisioning, the power allocations for D1 and D2, denoted by $P_1^o(\widetilde{\boldsymbol{\nu}}_1)$ and $P_2^o(\widetilde{\boldsymbol{\nu}}_2)$, respectively, can be obtained as follows:

$$\begin{cases} P_1^o(\widetilde{\boldsymbol{\nu}}_1) = \lim\limits_{\theta\to 0} P_1^*(\widetilde{\boldsymbol{\nu}}_1); \\ P_2^o(\widetilde{\boldsymbol{\nu}}_2) = \lim\limits_{\theta\to 0} P_2^*(\widetilde{\boldsymbol{\nu}}_2). \end{cases} \tag{40}$$

Plugging Eq. (40) into Eq. (2), we can derive the corresponding effective capacity as follows:

$$\begin{aligned} &C_2\left(P_1^o(\widetilde{\boldsymbol{\nu}}_1), P_2^o(\widetilde{\boldsymbol{\nu}}_2), \theta\right) \\ &= -\frac{1}{\theta}\log\left(\mathbb{E}_{\widetilde{\boldsymbol{\gamma}}}\left\{e^{-\frac{\beta}{2}(\log[(1+P_1^o(\widetilde{\boldsymbol{\nu}}_1)\gamma_1)(1+P_2^o(\widetilde{\boldsymbol{\nu}}_2)\gamma_2)])}\right\}\right). \end{aligned} \tag{41}$$

- The power allocation scheme for the orthogonal-channel mode with fixed QoS exponent ($\theta \to \theta_0$): The power allocations of this scheme, denoted by $P_1^s(\widetilde{\boldsymbol{\nu}}_1)$ and $P_2^s(\widetilde{\boldsymbol{\nu}}_2)$, respectively, can be obtained as follows:

$$\begin{cases} P_1^s(\widetilde{\boldsymbol{\nu}}_1) = \lim\limits_{\theta\to\theta_0} P_1^*(\widetilde{\boldsymbol{\nu}}_1); \\ P_2^s(\widetilde{\boldsymbol{\nu}}_2) = \lim\limits_{\theta\to\theta_0} P_2^*(\widetilde{\boldsymbol{\nu}}_2). \end{cases} \tag{42}$$

Plugging Eq. (42) into Eq. (2), we can derive the corresponding effective capacity as follows:

$$\begin{aligned} &C_2\left(P_1^s(\widetilde{\boldsymbol{\nu}}_1), P_2^s(\widetilde{\boldsymbol{\nu}}_2), \theta\right) \\ &= -\frac{1}{\theta}\log\left(\mathbb{E}_{\widetilde{\boldsymbol{\gamma}}}\left\{e^{-\frac{\beta}{2}(\log[(1+P_1^s(\widetilde{\boldsymbol{\nu}}_1)\gamma_1)(1+P_2^s(\widetilde{\boldsymbol{\nu}}_2)\gamma_2)])}\right\}\right). \end{aligned} \tag{43}$$

D. QoS-Guaranteed Power Allocation Under the Cellular Mode

For QoS-guaranteed cellular mode, we formulate the optimization problem to maximize the effective capacity given by Eq. (8) as follows:

$$\begin{aligned} \boldsymbol{P4}: \quad & \min_{(P_1(\widetilde{\boldsymbol{\nu}}_1),P_2(\widetilde{\boldsymbol{\nu}}_3))} \left\{\mathbb{E}_{\boldsymbol{\gamma}_c}\left[e^{-\theta R_3(P_1(\widetilde{\boldsymbol{\nu}}_1),P_2(\widetilde{\boldsymbol{\nu}}_3))}\right]\right\} \\ = \quad & \min_{(P_1(\widetilde{\boldsymbol{\nu}}_1),P_2(\widetilde{\boldsymbol{\nu}}_3))} \left\{\mathbb{E}_{\boldsymbol{\gamma}_c}\left[\left(1+P_1(\widetilde{\boldsymbol{\nu}}_1)\gamma_1\right)^{-\frac{\beta}{2}} \right.\right. \\ & \qquad \left.\left. \left(1+P_2(\widetilde{\boldsymbol{\nu}}_3)\gamma_3\right)^{-\frac{\beta}{4}}\right]\right\} \end{aligned}$$

$$\text{s.t.}: \quad 1).\ \mathbb{E}_{\boldsymbol{\gamma}_c}\left[P_1(\widetilde{\boldsymbol{\nu}}_1) + P_2(\widetilde{\boldsymbol{\nu}}_3)\right] \leq \overline{P}; \tag{44}$$

$$2).\ P_1\left(\widetilde{\boldsymbol{\nu}}_1\right) \geq 0; \tag{45}$$

$$3).\ P_2\left(\widetilde{\boldsymbol{\nu}}_3\right) \geq 0. \tag{46}$$

We can derive the optimal solutions for problem ***P4*** in the way similar to that solving problem ***P3***. Thus, we can solve problem ***P4*** for the following four cases:

Case A: $P_1(\widetilde{\boldsymbol{\nu}}_1) > 0$ and $P_2(\widetilde{\boldsymbol{\nu}}_3) > 0$. We can derive the optimal solutions of problem ***P4*** for this case as follows:

$$\begin{cases} P_1(\widetilde{\boldsymbol{\nu}}_1) = \frac{2}{\left(\frac{2^{\frac{\beta+4}{2}}\widetilde{\lambda}_c}{\beta}\right)^{\frac{4}{3\beta+4}}\left(\gamma_1^2\gamma_3\right)^{\frac{\beta}{3\beta+4}}} - \frac{1}{\gamma_1}; \\ P_2(\widetilde{\boldsymbol{\nu}}_3) = \frac{1}{\left(\frac{2^{\frac{\beta+4}{2}}\widetilde{\lambda}_c}{\beta}\right)^{\frac{4}{3\beta+4}}\left(\gamma_1^2\gamma_3\right)^{\frac{\beta}{3\beta+4}}} - \frac{1}{\gamma_3}, \end{cases} \tag{47}$$

where $\widetilde{\lambda}_c$ is the Lagrangian multiplier associated with the constraint given by Eq. (44).

Case B: $P_1(\widetilde{\boldsymbol{\nu}}_1) = 0$ and $P_2(\widetilde{\boldsymbol{\nu}}_3) > 0$. We can derive the optimal solutions of problem ***P4*** for this case as follows:

$$\begin{cases} P_1(\widetilde{\boldsymbol{\nu}}_1) = 0; \\ P_2(\widetilde{\boldsymbol{\nu}}_3) = \frac{1}{\left(\frac{4\widetilde{\lambda}_c}{\beta}\right)^{\frac{4}{\beta+4}}\gamma_3^{\frac{\beta}{\beta+4}}} - \frac{1}{\gamma_3}. \end{cases} \tag{48}$$

Case C: $P_1(\widetilde{\boldsymbol{\nu}}_1) > 0$ and $P_2(\widetilde{\boldsymbol{\nu}}_3) = 0$. We can derive the optimal solutions of problem ***P4*** for this case as follows:

$$\begin{cases} P_1(\widetilde{\boldsymbol{\nu}}_1) = \frac{1}{\left(\frac{2\widetilde{\lambda}_c}{\beta}\right)^{\frac{2}{\beta+2}}\gamma_1^{\frac{\beta}{\beta+2}}} - \frac{1}{\gamma_3}; \\ P_2(\widetilde{\boldsymbol{\nu}}_3) = 0. \end{cases} \tag{49}$$

Case D: $P_1(\widetilde{\boldsymbol{\nu}}_1) = 0$ and $P_2(\widetilde{\boldsymbol{\nu}}_3) = 0$. In this case, there is no strategy that can satisfy the strict inequalities $P_1(\widetilde{\boldsymbol{\nu}}_1) > 0$ and $P_2(\widetilde{\boldsymbol{\nu}}_3) > 0$.

Based on the solutions to problem ***P4*** specified by **Case A**, **Case B**, **Case C**, and **Case D**, we can obtain the optimal QoS-driven power allocations for the cellular mode using the similar method which is used to derive the QoS-driven power allocations for the orthogonal-channel mode specified by Section IV-B. The detailed derivation is omitted due to lack of space.

V. Simulation Evaluations

We conduct simulation experiments to evaluate the performance of our proposed optimal power allocation schemes for QoS-guaranteed underlaying wireless networks. Throughout our simulations, we set the bandwidth for the D2D-cellular underlaying unit as $B = 100$ KHz, the time frame length as $T = 2$ ms, and the noise variance as $\sigma^2 = 1$. All the channels' amplitudes follow independent Rayleigh distribution. The average power degradation of each channel is determined by $\overline{\gamma} = K(d_0/d)^{\eta}$ [24], where d is the transmission distance, d_0 is the reference distance, K is a unitless constant corresponding to the antenna characteristics, and η is the path loss exponent. In our simulations, we set $d_0 = 1$ m and $\eta = 3$. Furthermore, we choose K such that $\overline{\gamma} = 0$ dB at $d = 100$ m. Also, we set the average power constraint for the D2D-cellular underlaying unit as $\overline{P} = 2$ W.

Figure 3 shows an instance of positions of the BS, the cellular device, and the D2D pair, where the distance between D2 and D3 is smaller than the distance between D2 and the BS, reflecting that Eq. (11) holds. This kind of topology is obtained by arbitrarily choosing the positions for D1 and D2 within the region that the distance between D1 (and D2) and the BS is not larger than 80 m. The position for D3 is arbitrarily chosen except guaranteeing that the distance between D2 and D3 is *less* than the distance between D2 and the BS.

Based on the topology specified by Fig. 3, Figs. 4 and 5 compare the effective capacity of our proposed QoS-driven power allocation schemes with the effective capacity of a number of suboptimal power allocation schemes for the co-channel mode and the orthogonal-channel mode, respectively. The effective capacity of using the QoS-guaranteed cellular mode is also depicted for comparison in Figs. 4 and 5, respectively. As illustrated in Figs. 4 and 5, the QoS-guaranteed co-channel/orthogonal-channel mode can achieve larger effective capacity than the effective capacity of the QoS-guaranteed cellular mode. This is because when the average SNR of the channel between D2 and D3 is larger than the average SNR of the channel between D2 and the BS, our proposed QoS-driven power allocation schemes can efficiently obtain the channel average SNR gain of using the channel between D2 and D3 instead of using the channel between D2 and the BS. From Figs. 4 and 5, we can observe that the effective capacity of using the co-channel/orthogonal-channel mode without QoS provisioning is always less that the effective capacity of using our proposed QoS-guaranteed co-channel/orthogonal-channel mode under different delay-QoS requirements. However, when the delay-QoS is very loose ($\theta \to 0$), the effective capacity of using the co-channel/orthogonal-channel mode without QoS provisioning is very close to the effective capacity of using our proposed QoS-guaranteed co-channel/orthogonal-channel mode. This is consistent with the fact that the effective capacity turns to the ergodic capacity when the delay-QoS is very loose ($\theta \to 0$). We can also observe that if we always use the QoS-driven power allocation scheme with a fixed $\theta_0 = 10^{-2}$, we can only achieve the maximum effective capacity for the traffic with delay-QoS requirement $\theta_0 = 10^{-2}$ while cannot obtain the maximum effective capacity for the traffic with other different delay-QoS requirements. Comparing Fig. 4 and Fig. 5, we can find that the QoS-guaranteed co-channel mode can achieve larger effective capacity than the QoS-guaranteed orthogonal-channel mode with the topology shown as in Fig. 3.

Also based on Fig. 3, Fig. 6 shows that the effective capacity of the co-channel mode decreases as the distance between D2 and D3 increases, where we assume that the distance between D2 and D3 increases from 21.32 m (the distance between D2 and D3 in Fig. 3) to 46.40 m (the distance between D2 and the BS in Fig. 3). The upper bold line corresponds to the effective capacity of the QoS-guaranteed co-channel mode in Fig. 4 and the bottom bold line corresponds to $\mathbb{E}_{\gamma_2}\{\gamma_2\} = \mathbb{E}_{\gamma_3}\{\gamma_3\}$, respectively. The effective capacity corresponding to $\mathbb{E}_{\gamma_2}\{\gamma_2\} = \mathbb{E}_{\gamma_3}\{\gamma_3\}$ is larger than the effective capacity of the QoS-guaranteed cellular mode because the D2D communication enjoys double frequency-time resource as compared with that of the cellular communication. The effective capacity orthogonal-channel mode follows the similar decreasing trend and we omit the corresponding performance evaluation here.

Figure 7 shows another instance of positions of the BS, the cellular device, and the D2D pair, where the distance between D2 and D3 is larger than the distance between D2 and the BS, reflecting that Eq. (11) is not satisfied. The positions of D1 and D2 are arbitrarily chosen within the region that the distance between D1 (and D2) and the BS is not larger than 80 m. The position for D3 is arbitrarily chosen except guaranteeing that the distance between D2 and D3 is *larger* than the distance between D2 and the BS.

Based on the topology shown in Fig. 7, Fig. 8 shows the effective capacity of our proposed QoS-guaranteed co-channel mode, our proposed QoS-guaranteed orthogonal-channel mode, and the QoS-guaranteed cellular mode. As shown in Fig. 8, both the QoS-guaranteed co-channel mode and the QoS-guaranteed orthogonal-channel mode cannot achieve larger effective capacity than the QoS-guaranteed cellular mode. This is mainly because the average SNR of the channel between D2 and D3 is less than the average SNR of the channel between D2 and the BS. As a result, in such a topology where Eq. (11) is not satisfied, it is undesirable to cluster D2 and D3 as a D2D group. Instead, it is desirable to choose the QoS-guaranteed cellular mode.

VI. Conclusions

Based on the framework we developed to analyze the impact of different delay-QoS requirements on underlaying wireless networks, we formulated the effective capacity optimization problems for the QoS-guaranteed co-channel mode and the QoS-guaranteed orthogonal-channel mode, respectively. We developed the QoS-driven optimal power allocation schemes to maximize the effective capacity with the co-channel mode and the orthogonal-channel mode based underlaying wireless networks. We also proposed a number of suboptimal power allocation schemes to compare their obtained effective capacity with the effective capacity of our developed QoS-guaranteed co-channel/orthogonal-channel mode. The obtained simulation results show that under different QoS requirements, when

the average SNR of the channel between two D2D devices is larger than the average SNR of the channel between the D2D device (with transmit signal) and the BS, our developed QoS-driven optimal power allocation schemes for co-channel/orthogonal-channel mode can achieve larger effective capacity than the effective capacity of QoS-guaranteed cellular mode for the underlaying wireless networks.

REFERENCES

[1] K. Doppler, M. Rinne, C. Wijting, C. B. Ribeiro, and K. Hugl, "Device-to-device communication as an underlay to LTE-advanced networks," *IEEE Communications Magazine*, vol. 47, no. 12, pp. 42–49, 2009.

[2] H. Min, W. Seo, J. Lee, S. Park, and D. Hong, "Reliability improvement using receive mode selection in the device-to-device uplink period underlaying cellular networks," *IEEE Transactions on Wireless Communications*, vol. 10, no. 2, pp. 413–418, 2011.

[3] C. Yu, K. Doppler, C. B. Ribeiro, and O. Tirkkonen, "Resource sharing optimization for device-to-device communication underlaying cellular networks," *IEEE Transactions on Wireless Communications*, vol. 10, no. 8, pp. 2752–2763, 2011.

[4] G. Fodor and N. Reider, "A distributed power control scheme for cellular network assisted D2D communications," in *IEEE 2011 Global Telecommunications Conference (GLOBECOM)*, 2011, pp. 1–6.

[5] Z. Liu, T. Peng, S. Xiang, and W. Wang, "Mode selection for device-to-device (D2D) communication under LTE-advanced networks," in *IEEE 2012 International Conference on Communications (ICC)*, 2012, pp. 5563–5567.

[6] S. Hakola, T. Chen, J. Lehtomaki, and T. Koskela, "Device-to-device (D2D) communication in cellular network - performance analysis of optimum and practical communication mode selection," in *IEEE 2010 Wireless Communications and Networking Conference (WCNC)*, 2010, pp. 1–6.

[7] S. Xu and H. Wang, "Transmission mode selection and communication establishment in the hybrid device-to-device and cellular networks," in *2012 Fourth International Conference on Ubiquitous and Future Networks (ICUFN)*, 2012, pp. 156–161.

[8] M. Jung, K. Hwang, and S. Choi, "Joint mode selection and power allocation scheme for power-efficient device-to-device (D2D) communication," in *IEEE 75th Vehicular Technology Conference (VTC Spring)*, 2012, pp. 1–5.

[9] X. Xiao, X. Tao, and J. Lu, "A QoS-aware power optimization scheme in OFDMA systems with integrated device-to-device (D2D) communications," in *IEEE 2011 Vehicular Technology Conference (VTC Fall)*, 2011, pp. 1–5.

[10] Z. Liu, H. Chen, T. Peng, and W. Wang, "Optimal density and power allocation of D2D communication under heterogeneous networks on multi-bands with outage constraints," in *IEEE 2013 International Conference on Computing, Networking and Communications (ICNC)*, 2013, pp. 1179–1183.

[11] C. Yu, O. Tirkkonen, K. Doppler, and C. Ribeiro, "Power optimization of device-to-device communication underlaying cellular communication," in *IEEE 2009 International Conference on Communications (ICC)*, 2009, pp. 1–5.

[12] D. Wu and R. Negi, "Effective capacity: a wireless link model for support of quality of service," *IEEE Transactions on Wireless Communications*, vol. 2, no. 4, pp. 630–643, July 2003.

[13] X. Zhang and Q. Du, "Cross-layer modeling for QoS-driven multimedia multicast/broadcast over fading channels in mobile wireless networks," *IEEE Communications Magazine*, vol. 45, no. 8, pp. 62–70, Aug. 2007.

[14] J. Tang and X. Zhang, "Quality-of-service driven power and rate adaptation over wireless links," *IEEE Transactions on Wireless Communications*, vol. 6, no. 8, pp. 3058–3068, Aug. 2007.

[15] W. Cheng, X. Zhang, and H. Zhang, "Joint spectrum and power efficiencies optimization for statistical QoS provisionings over SISO/MIMO wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 5, pp. 903–915, 2013.

[16] C. S. Chang, "Stability, queue length, and delay of deterministic and stochastic queueing networks," *IEEE Transactions on Automatic Control*, vol. 39, no. 5, pp. 913–931, May 1994.

[17] C. S. Chang, *Performance Guarantees in Communication Networks*, Spinger-Verlag London, 2000.

[18] R. Horst and H. Tuy, *Global Optimization: Deterministic Approaches*, 2nd edition Berlin, Germany: Springer-Verlag, 1993.

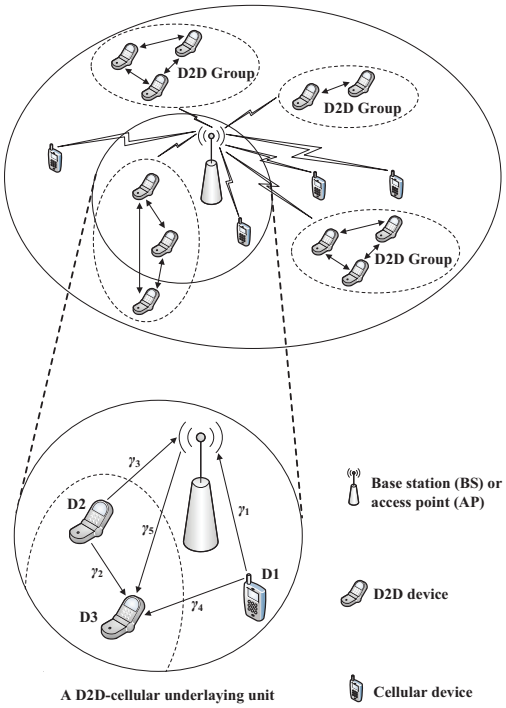

Fig. 1. The underlaying wireless network with cellular devices, the BS, and the D2D groups. The D2D groups consist of a set of D2D devices. The zoom-in figure on the bottom left corner shows the detailed topology of a D2D-cellular underlaying unit. The underlaying wireless network is partitioned into a number of D2D-cellular underlaying units. This implies that in our underlaying wireless network model, each mobile terminal uniquely belongs to one D2D-cellular underlaying unit, playing a role of either D2D device or cellular device, and each D2D device uniquely belongs to a D2D group.

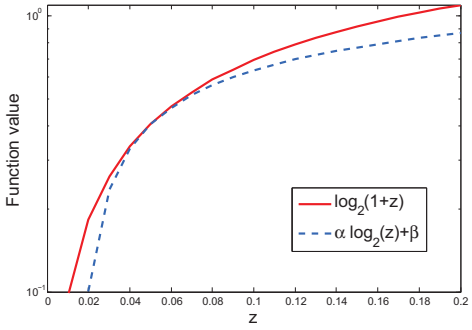

Fig. 2. The function $\log_2(1+z)$ and its tightest lower-bound $\alpha \log_2(z)+\beta$ with $z_0 = 0.05$.

[19] M. Chiang and J. Bell, "Balancing supply and demand of bandwidth in wireless cellular networks: utility maximization over powers and rates," in *IEEE 2004 INFOCOM*, 2004, vol. 4, pp. 2800–2811.

[20] M. Chiang, "Balancing transport and physical layers in wireless multihop networks: Jointly optimal congestion control and power control," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 1, pp. 104–116, 2005.

[21] J. Papandriopoulos, S. Dey, and J. S. Evans, "Optimal and distributed protocols for cross-layer design of physical and transport layers in manets," *IEEE/ACM Transactions on Networking*, vol. 16, no. 6, pp. 1392–1405, June 2008.

[22] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.

[23] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, 1985.

[24] A. J. Goldsmith, *Wireless Communications*, Cambridge University Press, 2005.

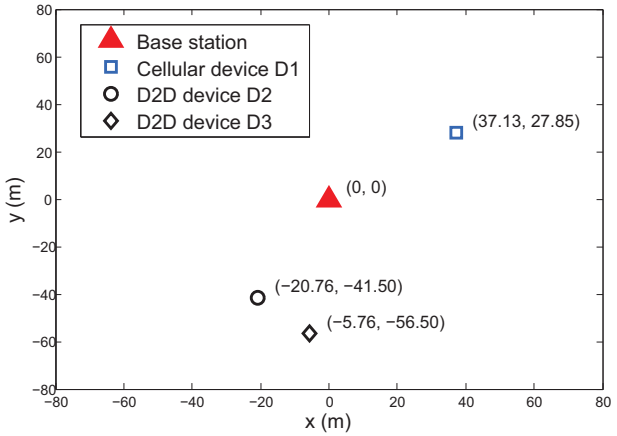

Fig. 3. Instance 1: An instance of positions of the BS, the cellular device, and the D2D pair. The coordinates of the BS, the cellular device D1, the D2D device D2, and the D2D device D3 are (0, 0), (37.13, 27.85), (-20.76, -41.50), and (-5.76, -56.50), respectively.

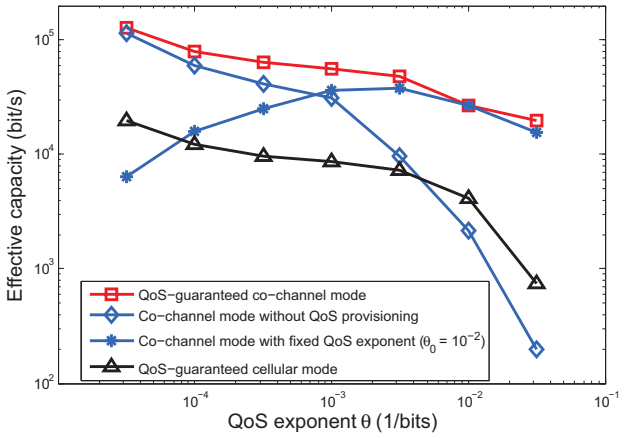

Fig. 4. Comparison of the effective capacity (corresponding to Instance 1) using our proposed QoS-driven power allocation scheme for the co-channel mode, the power allocation scheme for the co-channel mode without QoS provisioning, the power allocation scheme for the co-channel mode with fixed QoS exponent ($\theta_0 = 10^{-2}$), and the QoS-driven power allocation scheme for the cellular mode.

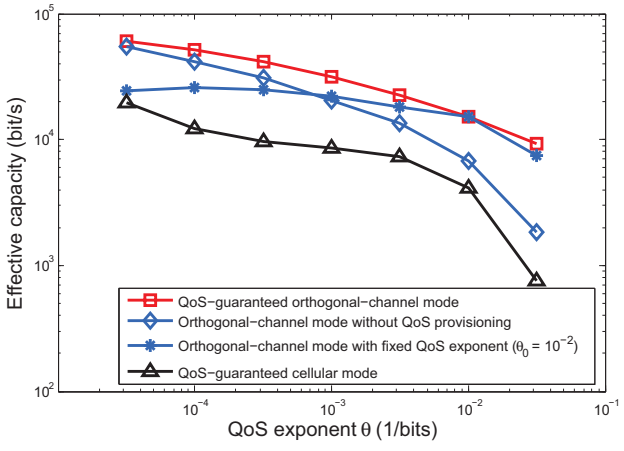

Fig. 5. Comparison of the effective capacity (corresponding to Instance 1) using our proposed QoS-driven power allocation scheme for the orthogonal-channel mode, the power allocation scheme for the orthogonal-channel mode without QoS provisioning, the power allocation scheme for the orthogonal-channel mode with fixed QoS exponent ($\theta_0 = 10^{-2}$), and the QoS-driven power allocation scheme for the cellular mode.

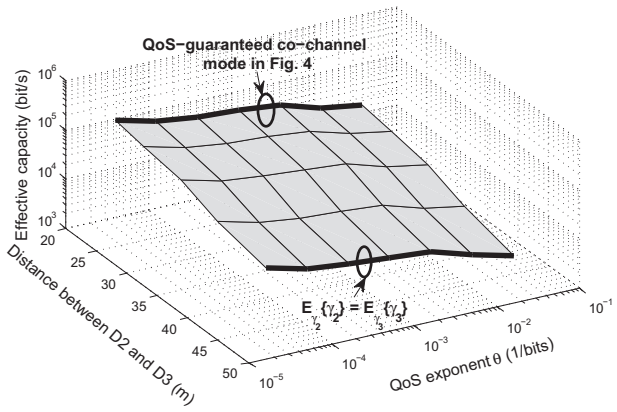

Fig. 6. The effective capacity of the co-channel mode versus the QoS exponent and the distance between the D2D pair.

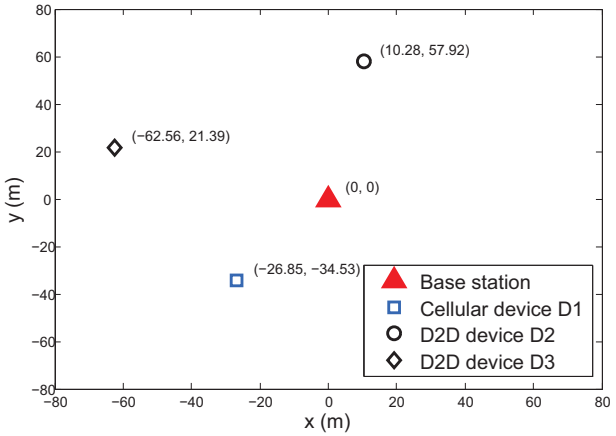

Fig. 7. Instance 2: An instance of positions of the BS, the cellular device, and the D2D pair. The coordinates of the BS, the cellular device D1, the D2D device D2, and the D2D device D3 are (0, 0), (-26.85, -34.53), (10.28, 57.92), and (-62.56, 21.39), respectively.

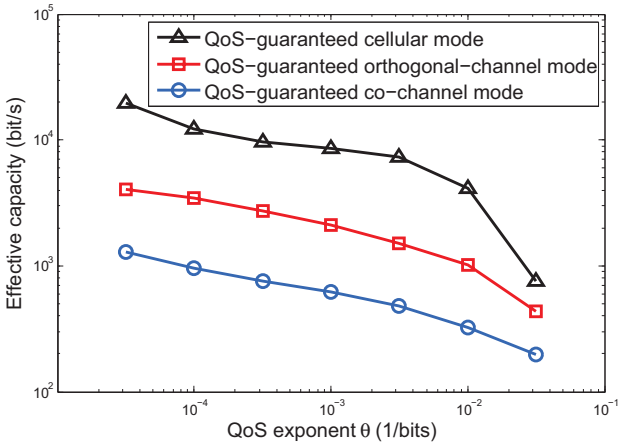

Fig. 8. Comparison of the effective capacity (corresponding to Instance 2) using the QoS-driven power allocation scheme for the cellular mode, our proposed QoS-driven power allocation scheme for the orthogonal-channel mode, and our proposed QoS-driven power allocation scheme for the co-channel mode.