Abstract—Cache-enabled Device-to-Device (D2D) communication is widely recognized as one of the key components of the emerging fifth generation (5G) cellular network architecture. However, conventional half-duplex (HD) transmission may not be sufficient to provide fast enough content delivery over D2D links and to meet strict latency targets of emerging D2D applications. In-band full-duplex (FD), with its capability of allowing simultaneous transmission and reception, can improve spectral efficiency and reduce latency by providing more content delivery opportunities. In this paper, we consider a finite network of D2D nodes in which each node is endowed with FD capability. We first carefully list all possible operating modes for an arbitrary device and use it to compute the number of devices that are actively transmitting at any given time. We then characterize network performance in terms of the success probability, which depends on the content availability, signal-to-interference ratio (SIR) distribution, as well as the operating mode of the D2D receiver. Our analysis concretely demonstrates that the caching dictates the system performance in lower target SIR thresholds whereas interference dictates the performance at the higher target SIR thresholds.

Index Terms—D2D, stochastic geometry, BPP, caching, half-duplex, full-duplex, power control.

I. INTRODUCTION

In the 5G evolution, cache-enabled D2D technology has drawn widespread attention because of its potential to improve system performance, enhance user experience [1]. The main idea behind this technology is to use the local storage of the user devices to store popular content and deliver it asynchronously to proximate devices through D2D communications whenever they need it. Fundamental limits of the wireless D2D caching systems is analyzed in [2]. One of the schemes being proposed for the 5G mobile communications systems is that of single channel FD, which allows simultaneous transmission and reception on the same channel at the same time. In a cache-enabled D2D networks, FD radios can promise more advantages in comparison with its HD counterpart [3]–[5], by providing more content delivery opportunities, improving spectral efficiency, and reducing end-to-end delay.

The performance analysis of the cache-enabled D2D communications requires considerations of the caching mechanism not only in the caching performance, but also in the SIR distribution. More precisely, the caching mechanism determines the possible transmitting users and the user operating modes (HD/FD) for the receiver of interest. This consideration leads to more challenging analysis compared to the case in which the SIR distribution is assumed to be independent of the caching mechanism due to two main reasons. First, the performance strictly depends on the D2D links formed due to content cached in the user devices as well as their demands. For instance, two users can initiate D2D link, provided at least one of them finds its desired content in the other’s cache, and the experienced SIR at the receiver exceeds some predefined target SIR thresholds. Second, depending upon the cached content and the user demand, an arbitrary node can operate in either HD or FD mode. Even when an arbitrary node operates in FD, it does not necessarily form the more intuitive bi-directional FD (BFD) link in which two devices exchange data with each other. Another possibility is a three node FD (TNFD) collaboration in which an intermediate node can receive its desired content from one node and concurrently serve some other node using content stored in its cache. These challenges in the analysis of cache-enabled D2D networks with FD capability have not yet been addressed, which is the main focus of this work.

As evident from the above discussion already, an arbitrary user can operate in different operating modes depending upon the caches contents and random demands. All possible operating models are illustrated in Fig. 1 and will be discussed in more detail in the next Section. In the existing literature, e.g., see [6]–[11], the focus is always on the performance analysis of an arbitrary node when it is obtaining content from a proximate node. This implicitly means that the received of interest is always assumed in half-duplex mode which corresponds to half-duplex receiver (HDRX) case in Fig. 1. As will be discussed in the sequel, the random operating modes also affect the SIR distribution. Therefore, the existing works focus on a very specific case out of all possible scenarios that could occur in a D2D network in which the nodes are endowed with FD capability. In this paper, we overcome this shortcoming by modeling these operating modes and their impact on the system performance accurately. More details about the main contributions are provided next.

Contributions. Despite Poisson point process (PPP) modeling in [3], we model a cache-enabled finite network using Binomial point process (BPP). We carefully list all possible operating modes when users have the FD capability. We then derive closed form expressions for the probabilities that an arbitrarily selected user is operating in one of these modes. These probabilities are used to derive the probability mass function (PMF) of the number of nodes that actively transmit at any given time. Using this PMF, we then characterize the SIR distribution for the HD/FD receiver of interest in the presence of power control as well as the success probability.
II. SYSTEM MODEL

We consider a finite network consisting of fixed $N$ number of users forming a BPP inside a disk $b(o, R) \subset \mathbb{R}^2$ with radius $R$. Users have the capability of FD communication. In other words, these nodes are assumed to located uniformly at random independently of each other over the disk. Denoting the sequence $\{y_i\} \equiv \Phi$ as the locations of the users, the probability density function (PDF) of each element $y_i$ is

$$f(y_i) = \begin{cases} \frac{1}{\pi R^2} & : \|y_i\| \leq R, \\ 0 & : \text{o.w.} \end{cases}$$

A. Caching Model

Denote the library of popular contents with size $m$ as $L = \{c_k\}_{k=1}^m$. Each content has an associated popularity score, which is characterized by the requests from users. Each user has a unique identity $u_k$, $k \in \{1, 2, \ldots, N\}$. To determine which contents are cached in which user devices, we use optimal caching policy [1], [12]. According to this policy, users cache a content from the library in such a way that the sequence of the contents $\{c_k\}_{k=1}^N$ are cached in users $\{u_k\}_{k=1}^N$, which means that there is no overlap between cached contents in user devices. Each user has the capability of storing multiple contents, however, for the sake of simplicity, we consider single content caching per user, nevertheless, multiple contents consideration can be considered as an extension of this work. The popularity of content $c_k$ is equivalent to the probability of requesting content $c_k$, which is cached in $u_k$. The request probability is denoted by $\rho_k$ and defined by

$$\rho_k = \Upsilon(\kappa, \gamma_r, m),$$

where $\Upsilon(\cdot)$ is the popularity distribution, and the parameter $\gamma_r$ is the skew exponent and characterizes the popularity distribution by controlling popularity of the contents for a given library size $m$. Each user randomly requests a content from the library according to popularity distribution given by $\Upsilon(\cdot)$ A pair of users can potentially initiate a D2D communication, providing that one of them finds its desired content in the other user. Based on the information of the caches and users’ requests, there will be different operating modes for an arbitrary user, which are introduced next.

B. Modeling User Operating Modes

There are six different possible operating modes for an arbitrary node as shown in Fig. 1 Definitions of the operating modes are as follows.

- **Self-Request (SR):** an arbitrary user can find its desired content on its own cache. We denote $P_{SR}$ as the probability of this mode.

- **Self-Request and HD Transmission (SR-HDTX):** an arbitrary user can find its desired content on its own cache, and concurrently can serve for other users' demand. We denote $P_{SR-HDTX}$ as the probability of this mode.

- **Full-Duplex Transceiver (FDTR):** an arbitrary user can find its desired content on its vicinity through D2D link, and concurrently can serve for other users’ demand. We denote $P_{FDTR}$ as the probability of this mode. This case can be divided into two different configurations as follows:
  - **Bi-Directional Full-Duplex (BFD):** an arbitrary user can concurrently exchange content with another user. We denote $P_{BFD}$ as the probability of this mode.
  - **Three-Node Full-Duplex (TNFD):** an arbitrary user can concurrently receive and transmit from and to different user. We denote $P_{TNFD}$ as the probability of this mode.

- **Half-Duplex Transmitter (HDTX):** an arbitrary user cannot find its desired content either on its vicinity or on its own cache, however, it can serve for other users' demand. We denote $P_{HDTX}$ as the probability of this mode.

- **Half-Duplex Receiver (HDRX):** an arbitrary user can receive its desired content via D2D link, and there is no user(s) that demand(s) for a content that is cached in this user. We denote $P_{HDRX}$ as the probability of this mode.

- **Hitting Outage (HO):** an arbitrary user cannot find its desired content neither on its vicinity nor its own cache, and there is no user(s) that demand(s) for the content that is cached in this user. We denote $P_{HO}$ as the probability of this mode.

One can say: $P_{FDTR} = P_{BFD} + P_{TNFD}$ and

$$P_{SR} + P_{SR-HDTX} + P_{FDTR} + P_{HDTX} + P_{HDRX} + P_{HO} = 1.$$
C. Channel Model and System Key Assumptions

In our system, all D2D pairs are sharing the same time/frequency resources. Fig. 2 illustrates example of D2D links along with the serving and interference links. For the channel model, standard power-law path loss model is considered in which the signal power decays at the rate of \( r^{-\alpha} \), where \( \alpha > 2 \) is the path loss exponent. Independent Rayleigh fading with unit mean, i.e., \( h \sim \exp(1) \) is assumed between any D2D pair. We employ full channel inversion for the power control scheme, which means the transmitting user compensates the experienced pathloss by \( Z_0^\alpha \), where \( Z_0 \) is the serving distance as shown in Fig. 2. We assume imperfect self-interference cancellation with residual power ratio 0 \( \leq \beta \leq 1 \) for the FD radios in concurrent transmission and reception over the same time/frequency. We also assume that the background noise is negligible compared to the interference and is hence ignored. We denote \( x_0 \in b(o, R) \) as the location of an arbitrary receiver within the disk. Now, according to the D2D operating modes described in II-B we have two types of the receivers in system, i.e., HDRX and FDTR. Denoting \( y_0 \) as the location of the serving node, \( x_0 \) as the location of receiver of interest, \( \|y_0 - x_0\| \) as the distance between the serving and the receiver of interest, \( x_i \) as the receiver of the interfering node, \( \|y_i - x_0\| \) as the distance between the interfering node and the receiver of interest, and \( \|y_i - x_i\| \) as the distance between the interfering node and its respective receiver located at \( x_i \in b(o, R) \), the \( \text{SIR} \) at the receiver type \( \delta \in \{\text{HDRX, FDTR}\} \), denoted by \( \text{SIR}_\delta \) can be defined as

\[
\text{SIR}_\delta = \frac{h_0}{\sum_{y_i \in N \setminus y_0} h_i \|y_i - x_0\|^{-\alpha} \|y_i - x_i\|^\alpha + \vartheta}, \tag{2}
\]

where \( \vartheta = 1_{\delta} \beta \|y_0 - x_0\|^{\alpha} \) and \( 1_{\delta} \) is an indicator function defined by \( 1_{\delta} = \begin{cases} 1 & ; \delta = \text{FDTR} \\ 0 & ; \delta = \text{HDRX} \end{cases} \).

III. ANALYZING OPERATING MODE PROBABILITIES

The following Theorem, provides the closed form expression for the probabilities of all possible operating modes described in subsection II-B

**Theorem 1.** The probabilities of all possible operating modes for an arbitrary user are

\[
\mathcal{P}_{\text{SR}} = \frac{1}{N} \sum_{\kappa=1}^{N} \rho_\kappa (1 - \rho_\kappa)^{N-1}, \tag{3}
\]

\[
\mathcal{P}_{\text{SR-HDTR}} = \frac{1}{N} \sum_{\kappa=1}^{N} \rho_\kappa \left( 1 - (1 - \rho_\kappa)^{N-1} \right), \tag{4}
\]

\[
\mathcal{P}_{\text{FDTR}} = \frac{1}{N} \sum_{\kappa=1}^{N} (\mathcal{P}_{\text{hit}} - \rho_\kappa) \left( 1 - (1 - \rho_\kappa)^{N-1} \right), \tag{5}
\]

\[
\mathcal{P}_{\text{FBD}} = \frac{1}{N} \sum_{\kappa=1}^{N} (\mathcal{P}_{\text{hit}} - \rho_\kappa) \rho_\kappa, \tag{6}
\]

\[
\mathcal{P}_{\text{TNFD}} = \frac{1}{N} \sum_{\kappa=1}^{N} (\mathcal{P}_{\text{hit}} - \rho_\kappa) \left( 1 - \rho_\kappa - (1 - \rho_\kappa)^{N-1} \right), \tag{7}
\]

\[
\mathcal{P}_{\text{HDRX}} = \frac{1}{N} \sum_{\kappa=1}^{N} (\mathcal{P}_{\text{hit}} - \rho_\kappa) (1 - \rho_\kappa)^{N-1}, \tag{8}
\]

\[
\mathcal{P}_{\text{HDTX}} = \frac{1}{N} \sum_{\kappa=1}^{N} (1 - \mathcal{P}_{\text{hit}}) \left( 1 - (1 - \rho_\kappa)^{N-1} \right), \tag{9}
\]

\[
\mathcal{P}_{\text{HO}} = \frac{1}{N} \sum_{\kappa=1}^{N} (1 - \mathcal{P}_{\text{hit}}) (1 - \rho_\kappa)^{N-1}, \tag{10}
\]

where \( \rho_\kappa \) is given in eq. \( (6) \), and \( \mathcal{P}_{\text{hit}} = \sum_{\kappa=1}^{N} \rho_\kappa \), which is the hitting probability.

**Proof:** See Appendix A.

IV. SUCCESS PROBABILITY ANALYSIS

The key intermediate analysis for the interference modeling and success probability is derivation of the distance distributions associated with the serving distance \( Z \) and the interfering distance \( W \) shown in Fig. 2.

A. Distance Distribution

First, we aim to derive the PDF of distance \( Z \) between the serving node and the receiver of interest. It is worth to note that probabilities of the operating modes, which are obtained in Theorem 1 determine the number of transmitters as mentioned in Corollary 1. Since these probabilities are not function of the distances between D2D pairs or the radius of the disk \( R \), hence, the serving node and the receiver of interest can be any arbitrary user among \( N \) users. This means that the serving node and the receiver of interest are chosen uniformly at random. Let us denote \( N_t \) as the number of concurrently transmitting nodes, \( Z_0 = \|y_0 - x_0\| \) as the distance of the serving link, \( \mathcal{W} = \{W_i = \|y_i - x_0\|\}_{i=1:N_t-1} \) as the set of distances from interfering nodes to the receiver of interest, and \( Z = \{Z_i = \|y_i - x_i\|\}_{i=1:N_t-1} \) as the set of distances from the interfering nodes to their respective receivers. We can infer that distance of the receiver of interest \( \|x_0\| \) is common factor in \( Z_0 \) and \( \mathcal{W} \), and distance of the interfering node \( \|y_i\| \) is common factor in \( W \) and \( Z \). By conditioning on \( \|x_0\| \), the elements in \( \mathcal{W} \) and \( Z_0 \) become independent, and by conditioning on \( \|y_i\| \), the elements in \( W \) and \( Z \) become independent. These independences will be used in the analysis of the Laplace transform of the interference field, which is a key to the analysis of the success probability. The
Lemma 1. The conditional PDF of the distance $Z_\lambda$ for a given $q \in \{\|x_0\|, \|y_i\|\}$ can be written as

$$f_{Z_\lambda}(z\lambda|q) = \begin{cases} f_{Z_{\lambda,1}}(z\lambda|q) : 0 \leq z\lambda \leq R - q, \\ f_{Z_{\lambda,2}}(z\lambda|q) : R - q < z\lambda \leq R + q, \end{cases}$$ (12)

where, $q = \{\|x_0\| : \lambda = 0, \|y_i\| : \lambda = i\}$, $f_{Z_{\lambda,1}}(z\lambda|q) = \frac{2z\lambda}{R^2}$, and $f_{Z_{\lambda,2}}(z\lambda|q) = \frac{2z\lambda}{\pi^2} \arccos\left(\frac{5z^2 + q^2 - R^2}{2qz}\right)$.

Proof: The proof is available in [1] Theorem 2.3.6.

Now, we need to derive the PDF of distance $W_i$. Similar to the previous Lemma, the elements $\|x_0\|$ and $\|y_i\|$ are common factors for $W$ and $Z$. Conditioning on $t = \|y_i\|$ and $v = \|x_0\|$ is sufficient to get a set of i.i.d. distances for the elements in $W$ and $Z$. The following Lemma provides the conditional PDF of the distance $W_i$.

Lemma 2. The conditional PDF of the distance $W_i$ for given $v$ and $t$, denoted by $f_{W_i}(w_i|t,v)$, is defined by

$$f_{W_i}(w_i|t,v) = \frac{1}{\pi} \frac{w_i/v\sqrt{v-t}}{\sqrt{1 - \left(\frac{w_i^2 + t^2}{2vt}\right)^2}}, \quad |v-t| < w_i < v+t,$$ (13)

Proof: The proof is available in [6].

For the PDFs of distances $V$ and $T$, it can be easily shown that: $f_V(v) = \frac{2v}{\pi^2}$ and $f_T(t) = \frac{2t}{\pi^2}$.

B. Laplace Transform of the Interference

By using the distance distributions obtained in subsection [IV-A], we aim to obtain the Laplace transform of the interference field through the following Lemma.

Lemma 3. The Laplace transform of the interference field at the receiver of interest, by considering the operating mode $\delta \in \{\text{HDRX, FDTR}\}$ denoted by $L_{\delta}(s)$, is given by

$$L_{\delta}(s) = \int_0^R \int_0^R \mathcal{M}(v,t) \frac{e^{st}}{R^2} dv dt,$$ (14)

where $\mathcal{M}(v,t)$ is given by eq. (15). $\mathcal{F}(w_i, t, z_0) = \int_{|v-t|}^{R-t} f_{Z_{\lambda,1}}(z_i|t)f_{W_i}(w_i|v,t) dz_i dv_i$, $\mathcal{G}(w_i, t, z_0) = \int_{|v-t|}^{R-t} f_{Z_{\lambda,2}}(z_i|t)f_{W_i}(w_i|v,t) dz_i dw_i$, and $\mathcal{F} = \exp(-\eta \gamma r)^{m/\eta \gamma} = \frac{\exp(-s\lambda^2 \zeta z_0^2)}{1 + s\lambda^2 \zeta \eta \gamma r^m}$.

Proof: See appendix [B].

C. Success Probability

Definition 1 (Success Probability). The probability that an arbitrary node can capture its desired content either i) through its own cache or ii) through a D2D communication in its vicinity where the experienced SIR value exceeds some predefined threshold $\theta$.

Theorem 2. The success probability for an arbitrary node is denoted by $P_s(N, \gamma_r, \theta)$ and given by

$$P_s(N, \gamma_r, \theta) = P_s,\text{cache}(N, \gamma_r) + P_s,\text{SIR}(N, \gamma_r, \theta),$$ (16)

where $P_s,\text{cache}(N, \gamma_r) = \frac{1}{\pi} P_{\text{hit}}$, $P_s,\text{SIR}(N, \gamma_r, \theta) = \sum_{n_t=1}^{N} Q_t(N, \gamma_r, \theta) f_{N_t}(n_t)$, $Q_t(N, \gamma_r, \theta) = P_{\text{HDRX}_t, \text{FDTR}_t}(\theta) P_{\text{FDTR}_t, \text{FDTR}_t}(\theta)$, $f_{N_t}(n_t) = \left(\frac{N}{n_t}\right) (P_{\text{TX}})^{n_t} (1 - P_{\text{TX}})^{N-n_t}$, and $P_{\text{TX}}$ is given by Corollary [1].

Proof: See appendix [C].

V. RESULTS AND DISCUSSION

For the popularity distribution, we use Zipf distribution, which is an special case of the Riemann Zeta function and widely used in the existing works [1]. [6] [11]. This distribution is defined by $\rho_\kappa = \sum_{n=1}^{\infty} \frac{\kappa^n}{\zeta(n-\kappa)}$. Here, we assume $m = 1000$, $\alpha = 4$, and $\beta = 10^{-5}$.

1) Impact of $N$: Fig. [3] demonstrates the effects of $N$ on the success probability. Although the increasing number of the nodes can lead to the interference rising, in contrary, it can also cause more successful content capturing due to more available contents. Hence, explicitly, there is a tradeoff between the contents availability and the interference, namely, more users, the higher is the contents availability, and the higher is the interference. In this tradeoff, domination of the content availability accounts for the lower target SIR thresholds, because the complementary cumulative distribution function (CCDF) of the SIR in the lower thresholds for all values of the $N$ is almost the same. But, the only involving factor in lower thresholds is the caching related parameters, which is the content availability. When the threshold increases, the interference dominates against the content availability due to increasing the number of transmitters. This tradeoff is more accentuated for the higher values of the $N$, i.e., the slope of the curves are steeping when $N$ increases, due to increasing the number of transmitters.

2) Impact of Zipf exponent: Fig. [2] illustrates the effects of Zipf exponent $\gamma_r$ on the success probability. Technically, the higher values of the $\gamma_r$ means more redundancy in the users’ demands, that is to say, few number of the contents accounts for the majority of requests. Correspondingly, caching performance dominates against SIR distribution, when $\gamma_r$ increases. More precisely, caching dictates the system performance in lower target SIR thresholds whereas interference dictates the performance at the higher target SIR thresholds.

VI. CONCLUDING REMARKS

In this paper, we have derived the probabilities of different user operating modes in closed form expressions. By using tools from the stochastic geometry, we analyzed success probability for an arbitrary node, in the presence of the power control scheme. The performance analysis captures both possible operating modes through the strongly interdependent cache parameters. We left derivation of the closed form bound or approximations for the success probability as a promising future work. Exploiting more realistic caching policy, namely random caching policy, can be considered as an extension to this work. Moreover, in this paper and all related works, the channel for the FD links is assumed to be reciprocal and...
\[ M(v, t) = \int_{0}^{R-v} \left( F(v, t, z_0) + G(v, t, z_0) \right)^{N_t-1} f_{Z_0,1}(z_0|v)dz_0 + \int_{-R+v}^{R-v} \left( F(v, t, z_0) + G(v, t, z_0) \right)^{N_t-1} f_{Z_0,2}(z_0|v)dz_0, \]  

(15)

Fig. 3. Coverage Probability versus SIR threshold $\theta$ for different values of $N$, and $\gamma_r = 1.2$, $R = 30$.

Fig. 4. Coverage Probability versus SIR threshold $\theta$ for different values of $\gamma_r$, and $N = 20$, $R = 40$.

exchanging data on a FD nodes is assumed to be symmetric. FD communications with asymmetric data and dissimilar channel is an open problem, and has worthy of investigation in the cache-enabled FD-D2D networks.

APPENDIX

A. Proof for Theorem 7

From the observations in Fig. 1 we can infer that the probability for each case at an arbitrary user $u_k$ depends on two different events; i) request of the user $u_k$. We denote this event by $A$ and the probability of this event by $P_{A,k}$. ii) requests from other users to the content which is cached in user $u_k$. We denote this event by $B$ and the probability of this event by $P_{B,k}$. Joint probability of both events, i.e., $P(A, B)$ gives the probability of the operating mode $\Delta$ for a specific node $u_k$. Since the requests at all users are independent from each other, we can say $P(A, B) = P(A)P(B) = P_{A,k}P_{B,k}$. Now, by using the law of total probability, the probability of operating mode $\Delta \in \{ \text{SR, SR-HDTX, FDTR, BFD, TNFD, HDTX, HDRX, HO} \}$ denoted by $P_{\Delta}$ for an arbitrary node can be defined by

\[ P_{\Delta} = \sum_{\kappa=1}^{N} P_{\Delta|\kappa}P_{\kappa} = P_{\kappa} = \frac{1}{N}, \]  

(17)

where, $P_{\kappa} = \frac{1}{N}$ is the probability of choosing an arbitrary user among $N$ users uniformly at random. Due to lack of space, we provide the proof for the FDTR mode, however, the approach for the other modes is the same. Now, let us define two binary random variables $X_k$ and $H_{\mu,k}$ for $u_k$ as follows.

\[ X_k = \begin{cases} 0 & \text{if } u_k \text{ cannot find its desired content} \\ 1 & \text{if } u_k \text{ can find its desired content} \end{cases} \]  

(18)

\[ H_{\mu,k} = \begin{cases} 0 & \text{if } u_\mu \text{ do not demands for content } c_k \\ 1 & \text{if } u_\mu \text{ demands for content } c_k \end{cases} \]  

(19)

The probability $Pr(X_k = 1)$ is equivalent to the situation that $u_k$ demands for a content, which is cached by other node in its vicinity, i.e., $Pr(X_k = 1) = Pr(\mu_\text{hit} = \mu_k)$, which corresponds to the parameter $P_{\kappa} = \frac{1}{N}$, i.e.,

\[ P_{\kappa} = Pr(X_k = 1), \]  

(20)

and the probability $Pr(H_{\mu,k} = 0)$ is equivalent to

\[ Pr(H_{\mu,k} = 0) = 1 - \rho_k. \]  

(21)

Now, the parameter $P_{\kappa} = \frac{1}{N}$ is equivalent to the probability that there is at least one node demands content $c_k$, hence

\[ P_{\kappa} = 1 - \left( \frac{1}{N} \right)^{N-1}, \]  

(22)

where (a) follows the fact that the requests at all users are independent from each other and (b) follows directly using the eq. (21). By substituting the eqs. (20) and (22) in eq. (17), we can get the final expression in eq. (7).

B. Proof for Lemma 3

According to Fig. 2 denoting $Z_0 = \|y_0 - x_0\|$, $Z_i = \|y_i - x_0\|$, and $W_i = \|y_i - x_0\|$, the Laplace transform of
the interference $I = \sum_{y_i \in \Phi \setminus y_0} \left( h_i Z_i W_i - \alpha + 1/b \beta Z_0 \right)$ is

$$
\mathcal{L}_{I, \delta}(s) = \mathbb{E} \left[ e^{-sIT} \right] = \mathbb{E} \left[ \prod_{y_i \in \Phi \setminus y_0} \exp \left( -s \left( h_i Z_i W_i - \alpha + 1/b \beta Z_0 \right) \right) \right]
$$

where $(a)$ follows from eqs. (23) and (26) in eq. (16) completes the proof.

The success probability for the receiver of interest denoted by $Q_s(\gamma, r, \gamma)$ depends on its operation mode, which is either HDRX or FDTR. Hence, we have

$$
Q_s(\gamma, r, \gamma) = \sum_{\delta \in \{\text{HDRX}, \text{FDTR}\}} \mathbb{P}_{SIR} \geq \theta
$$

where $(a)$ follows substitution from eq. (24). The expression in eq. (23) contains the parameter $N_t$, which is the number of transmitters and is random. From the Fig. 1 we can infer that the distribution of the random variable $N_t$ depends on the probability of the transmitting operation modes, which is defined in Corollary [1] i.e., $P_{TX}$. The number of transmitters $N_t$ is a Binomial random variable and its probability mass function (PMF) is $F_{N_t}(n_t) = \binom{N}{n_t} P_{TX}^{n_t} (1 - P_{TX})^{N-n_t}$.

Now, the final expression for the second part of the success probability $P_{SIR}(\gamma, r, \gamma)$ can be obtained by taking expectation of $Q_s(\gamma, r, \gamma)$ over the random variable $N_t$, i.e.,

$$
P_{SIR}(\gamma, r, \gamma) = \mathbb{E}_{N_t} [Q_s(\gamma, r, \gamma)]
$$

which completes the proof for the second part. Finally, substituting eqs. (23) and (26) in eq. (16) completes the proof.

REFERENCES

[1] N. Golrezaei, et al., "Femtocaching and device-to-device collaboration: A new architecture for wireless video distribution," IEEE Commun. Magazine, vol. 51, no. 4, pp. 124 - 149, April 2013.
[2] M. Ji, G. Caire, and A. F. Molisch, "Fundamental limits of caching in wireless D2D networks," IEEE Trans. on Inf. Theory, vol. 62, no. 2, pp. 849 - 869, Feb. 2016.
[3] M. Naslcheraghi, S. A. Ghorashi, and M. Shikh-Bahaei, "Performance Analysis of Inband FD-D2D Communications with Imperfect SI Cancellation for Wireless Video Distribution," arXiv preprint arXiv:1706.06142 (2017).
[4] M. Naslcheraghi, S. A. Ghorashi, and M. Shikh-Bahaei, "FD device-to-device communication for wireless video distribution," IET Communications, vol. 11, no. 7, pp. 1074 - 1081, Jan. 2017.
[5] C. Yang, et al., "Modeling and Analysis for Cache-enabled Cognitive D2D Communications in Cellular Networks," IEEE Globecom, 2016.
[6] A. Afzal, et al., "On the analysis of cellular networks with caching and coordinated device-to-device communication," IEEE ICC, 2016.
[7] Z. Chen, N. Pappas, and M. Kountouris, "Probabilistic caching in wireless D2D networks: Cache hit optimal versus throughput optimal," IEEE Communications Letters 21.3 (2017): 584-587.
[8] X. Song, et al. "Cache-Enabled Device to Device Networks With Contention-Based Multimedia Delivery," IEEE Access 2017:3228-3239.
[9] D. Malak, M. Al-Shalash, and J. G. Andrews, "Optimizing content caching to maximize the density of successful receptions in device-to-device networking," IEEE Trans. on Commun., vol. 64, no. 10, pp. 4365 - 4380, 2016.
[10] A. Mathai, An introduction to geometrical probability: distributional aspects with applications, Vol. 1. CRC Press, 1999.