

# Group Discovery Time in Device-to-Device (D2D) Proximity Services (ProSe) Networks

David Griffith, Aziza Ben Mosbah, and Richard Rouil  
National Institute of Standards and Technology  
Gaithersburg, Maryland 20899–8920  
Email: david.griffith@nist.gov

**Abstract**—Device-to-device (D2D) communications for Long Term Evolution (LTE) networks relies on a discovery process to enable User Equipment (UE) to determine which D2D applications and services are supported by neighboring UEs. This is especially important for groups of UEs that operate outside the coverage area of any base station. The amount of time required for discovery information to reach every UE in a group depends on the number of UEs in the group and the dimensions of the discovery resource pool associated with the Physical Sidelink Discovery Channel (PSDCH); an additional factor is the half-duplex property of current UEs. In this paper, we use a Markov chain to characterize the performance of Mode 2 direct discovery. The resulting analytical model gives the distribution of the time for a UE to discover all other UEs in its group. We validate the model using Monte Carlo and network simulations.

## I. INTRODUCTION

Proximity Services (ProSe) for Long Term Evolution (LTE) was developed by 3rd Generation Partnership Project (3GPP) for Device-to-Device (D2D) communications. The standard will allow User Equipments (UEs) to communicate directly with other UEs that are within range by using a portion of the channel known as the sidelink (SL). ProSe covers UEs that are in the coverage area of an evolved Node-B (eNB), in which case the eNB can coordinate SL resource allocation, but it also supports communication between UEs that are out-of-coverage with respect to any eNB. This affects public safety applications, which may involve deployment of personnel to remote areas, or deployment in disaster areas where the infrastructure has been destroyed.

UEs use discovery messages to exchange information regarding their D2D applications and capabilities. Discovery messages use the Physical Sidelink Discovery CHannel (PSDCH). In this paper, we consider Mode 2 discovery, where PSDCH resources are not allocated to individual UEs but are available for all UEs to use; this mode applies to the out-of-coverage case. UEs pick PSDCH resources randomly, which creates the risk of message loss due to collisions when more than one UE selects a given resource. Allocating more resources to the PSDCH reduces the collision rate, but it

**Disclaimer:** Certain commercial software packages are identified in this paper in order to specify the experimental procedure adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the software packages identified are necessarily the best available for the purpose.

also reduces the bandwidth for data transmission. In order to properly size the discovery resource pool and to determine the maximum size of a group of D2D UEs that the network can support, it is important to develop models that allow operators to accurately measure the performance of the PSDCH.

This work extends our previous study of the PSDCH [1], which used an analytical model to get the discovery message transmission probability that maximizes the message success probability between two UEs in a single period. Sarret et al. used simulations to get the time for a UE to be discovered by all other UEs in a group, considering both half-duplex and full duplex UEs [2]. Lin et al. examined two D2D discovery schemes and used a Markov chain to model a backoff procedure for an individual UE [3]. Zhang and Liu examined hopping patterns in the discovery resource pool for half-duplex UEs and used simulations to compare their performance with respect to the discovery rate and cumulative number of discovered UEs [4].

In this paper, we develop a Markov chain-based analytical model that gives the distribution of the time for a single UE in a D2D group to discover all other UEs in the group. First, we describe the Physical Sidelink Discovery Channel (PSDCH) and the assumptions behind our model in Section II. In Section III, we derive an analytical expression for the elements of the Markov chain's state transition probability matrix, which we use to obtain the cumulative distribution function (CDF) for the group discovery time. In Section IV, we validate the theoretical model from Section III using both Monte Carlo simulations in Matlab and network simulations in NS3, and we demonstrate the model's use by obtaining the maximum UE group size that allows a UE to discover all of its peers within a required number of periods with a given probability. We discuss extensions for the model and summarize our work in Section V. In Table I, we provide a list of the symbols that we use in this paper.

## II. MODELING THE DISCOVERY RESOURCE POOL

In this section, we describe D2D discovery message transmission and discuss the assumptions that underlie the analysis in Section III. We assume that UEs transmit discovery messages during every occurrence of the PSDCH discovery resource pool, which repeats periodically with period  $P$  [5, Clause 14.3.3], as described in [6, Clause 8.3]. A pool resource

is a single transport block, which is composed of a pair of adjacent Physical Resource Blocks (PRBs) that occupy the same subframe [5, Clause 14.3.1]. Discovery messages can be repeated in a given period up to four times (i.e., one initial transmission followed by zero, one, two, or three retransmissions); the number of transmissions is given by the parameter  $N_{SLD}^{TX}$  [5, Clause 14.3.1].

Increasing  $N_{SLD}^{TX}$  decreases the number of available resources in the pool, since the resource pool's effective dimensions in the frequency and time domains are respectively  $N_f = \lfloor M_{RB}^{PSDCH,RP}/2 \rfloor$  and  $N_t = \lfloor L_{PSDCH}/N_{SLD}^{TX} \rfloor$ , where  $M_{RB}^{PSDCH,RP}$  is the number of PRBs in the pool and  $L_{PSDCH}$  is the number of subframes spanned by the pool in the time domain. The total number of resources in the pool is thus  $N_r = N_f N_t$ .

We assume that all UEs are half-duplex. We also assume that the UEs use Mode 2 discovery [5, Clause 14.3.1]. In each period, every UE, independently of every other UE, generates a uniformly random resource index  $n_{PSDCH} \in \{0, 1, \dots, N_r - 1\}$ , which maps to a unique set of PRB and subframe indices via the following equations:

$$a_j^{(i)} = ((j-1)\lfloor N_f/N_{SLD}^{TX} \rfloor + \lfloor n_{PSDCH}/N_t \rfloor) \bmod N_f \quad (1a)$$

$$b_j^{(i)} = n_{PSDCH} \bmod N_t, \quad (1b)$$

where  $i$  is the period index and  $1 \leq j \leq N_{SLD}^{TX}$ , so that  $j$  indexes the transmission attempts in the  $i$ th period. The parameters  $a_j^{(i)}$  and  $b_j^{(i)}$  map to the subframe and PRB pair indices  $l_{N_{SLD}^{TX}, b_j^{(i)}+j-1}^{PSDCH}$  and  $(m_{2a_j^{(i)}}^{PSDCH}, m_{2a_j^{(i)}+1}^{PSDCH})$ , respectively. A UE can throttle its message transmissions by generating a  $[0, 1]$ -uniform random variate and transmitting if the variate is less than a defined threshold [7, Clause 5.15.1.1]. We assume that the threshold is one; arbitrary thresholds are part of future work.

Due to the mapping scheme, as we discussed in [1], when  $N_{SLD}^{TX} > 1$ , two UEs that pick resource indices that produce identical values for  $b_j^{(i)}$  in a given period,  $i$ , will transmit in the same subframes for all  $N_{SLD}^{TX}$  transmissions during that period. Thus we can model the discovery resource pool as shown in Fig. 1, where each element of the grid is associated with a unique value of  $n_{PSDCH}$ .

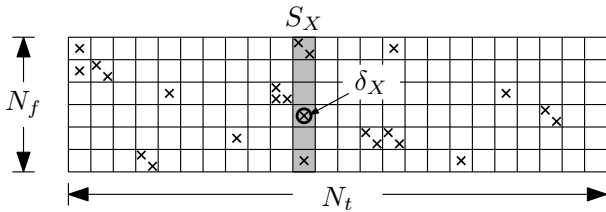


Fig. 1. The discovery resource pool model, showing transmissions from UEs in the D2D group,  $\mathcal{G}$ , and indicating the location of the discovery message  $\delta_X$  from a UE of interest, UE  $X$ , and the set of subframes used by UE  $X$ ,  $S_X$  (Fig. 3 from [1]).

Let  $\mathcal{G}$  denote a group of D2D-capable UEs; the number of UEs in  $\mathcal{G}$  is  $N_u$ , which we assume is constant. We assume that the area occupied by  $\mathcal{G}$  is small enough that every UE

in  $\mathcal{G}$  is able to receive transmissions from every other UE, and that if a UE picks a resource that no other UE picks, its message will be received by all other UEs. In practice, channel effects will introduce a message loss probability, which we are incorporating into the next generation of this model. We also assume that when two or more UEs pick the same resource, the mutual interference will prevent any collided message's being received by other UEs. In practice, some collided messages may be received by some UEs if the Signal to Interference Ratio (SIR) at the receiver is high enough. There has been some work on modeling the effect of SIR on the discovery process, notably the work by Kang and Kang [8] and Bagheri et al. [9]. However, Kang and Kang compute the average number of devices discovered in a given number of periods, while we obtain the CDF for the number of periods to discover all devices in a group, and Bagheri et al. do not consider the half-duplex effect in their analysis.

### III. ANALYTICAL MODEL

We characterize the time for a randomly chosen UE, which we denote as UE  $X$ , to discover all other UEs in its group. We use a discrete-time Markov chain whose time index  $t$  indicates the number of PSDCH periods that have elapsed since the starting time, and whose single state variable,  $N_D[t]$ , is the number of UEs that have been discovered by UE  $X$  at the end of the  $t$ th period. Also, we denote the number of undiscovered UEs at the end of the  $t$ th period as  $N_U[t] = (N_u - 1) - N_D[t]$ . The range of possible values for  $N_D[t]$  is  $0 \leq N_D[t] \leq N_u - 1$ ; the starting state is  $N_D[0] = 0$  (because UE  $X$  has not yet discovered any of the other  $(N_u - 1)$  UEs in the group) and the Markov chain's eventual ending state is  $\lim_{n \rightarrow \infty} N_D[t] = N_u - 1$ , which is the Markov chain's sole absorbing state.

#### A. The Markov Chain Model

We define the state probability vector for  $N_D[t]$  to be  $\pi[t] = [\pi_0[t], \dots, \pi_{N_u}[t]]$ , where  $\pi_i[t] = \Pr\{N_D[t] = i\}$  for  $i = 0, 1, 2, \dots, N_u - 1$ . Since we start with  $N_D[0] = 0$  discovered UEs,  $\pi[0] = [1, 0, 0, \dots, 0]$ , and  $\lim_{n \rightarrow \infty} \pi[t] = [0, 0, \dots, 0, 1]$ .

1) *The State Transition Matrix*: We define the state transition matrix to be  $\mathbf{T} = [T_{i,j}]$ , where the probability of transitioning from State  $i$  to State  $j$  is  $T_{i,j} = \Pr\{N_D[t] = j \mid N_D[t-1] = i\}$ , for  $i, j \in \{0, 1, 2, \dots, N_u - 1\}$ . Because  $N_D[t]$ , the number of UEs discovered by UE  $X$ , never decreases,  $T_{i,j} = 0$  when  $i > j$ . For  $i \leq j$ , a transition from State  $i$  to State  $j$  occurs when UE  $X$  discovers  $(j - i)$  UEs during the  $t$ th period. Thus

$$\begin{aligned} T_{i,j} &= \Pr\{D_a[t] = j - i \mid N_D[t-1] = i\} \\ &= \Pr\{D_a[t] = j - i \mid N_U[t-1] = N_u - 1 - i\}, \end{aligned} \quad (2)$$

where  $D_a[t]$  is the number of UEs discovered by UE  $X$  during the  $t$ th period, and  $D_a[0] = 0$ .

In Eq. (3), we show  $\mathbf{T}$  when  $N_u = 4$ . The value of  $N_D[t-1]$  is to the left of the corresponding row and the value of  $N_D[t]$  is above the corresponding column. Since  $N_D[t] = (N_u - 1)$  UEs

$$\mathbf{T} = \begin{array}{c|ccc|c} & 0 & 1 & 2 & 3 \\ \hline 0 & \Pr\{D_a[t]=0 | N_D[t-1]=0\} & \Pr\{D_a[t]=1 | N_D[t-1]=0\} & \Pr\{D_a[t]=2 | N_D[t-1]=0\} & \Pr\{D_a[t]=3 | N_D[t-1]=0\} \\ 1 & 0 & \Pr\{D_a[t]=0 | N_D[t-1]=1\} & \Pr\{D_a[t]=1 | N_D[t-1]=1\} & \Pr\{D_a[t]=2 | N_D[t-1]=1\} \\ 2 & 0 & 0 & \Pr\{D_a[t]=0 | N_D[t-1]=2\} & \Pr\{D_a[t]=1 | N_D[t-1]=2\} \\ \hline 3 & 0 & 0 & 0 & \underbrace{\Pr\{D_a[t]=0 | N_D[t-1]=3\}}_{=1} \end{array} \quad (3)$$

TABLE I  
LIST OF SYMBOLS

Symbol	Definition
$\Pr\{A\}$	Probability of event $A$
$E\{Z\}$	Expected value of random variable $Z$
$L_{PSDCH}$	Number of subframes spanned by the PSDCH
$M_{RB}^{PSDCH,RP}$	Number of PRBs occupied by the PSDCH
$l_i^{PSDCH}$	$i$ th subframe in the PSDCH
$m_j^{PSDCH}$	$j$ th PRB slot in the PSDCH
$N_r$	Number of resources in discovery pool
$N_f$	Number of PRB pairs in discovery pool
$N_t$	Number of subframe sets in discovery pool
$n_{PSDCH}$	Discovery resource index
$P$	PSDCH period duration
$\mathcal{G}$	The set of UEs in a given D2D group
$N_u$	Number of UEs in D2D group $\mathcal{G}$
UE $X$	Randomly chosen UE of interest from $\mathcal{G}$
$\delta_X$	Discovery message sent by UE $X$
$S_X$	Set of subframes occupied by $\delta_X$
$\mathcal{N}(A)$	Number of occurrences of event $A$
$n$	Number of collided discovery messages
$t$	Index indicating the $t$ th PSDCH period
$P_C(n   N_r; N_u)$	Probability of $n$ collisions given $N_u$ UEs using a pool with $N_r$ resources
$N_U[t]$	Number of UEs in $\mathcal{G}$ undiscovered by UE $X$ at the end of the $t$ th period
$N_D[t]$	Number of UEs in $\mathcal{G}$ discovered by UE $X$ at the end of the $t$ th period
$M_D, M_U$	Values of $N_D[t-1]$ and $N_U[t-1]$ respectively
$m_d, m_u$	Number of UEs out of $M_D$ and $M_U$ with discovery messages not in $S_X$
$D_a[t]$	Number of UEs in $\mathcal{G}$ discovered by UE $X$ during $t$ th period
$\nu$	Value taken by $D_a[t]$
$N_{\mathcal{G} \rightarrow X}$	Number of periods for UE $X$ to discover all UEs in $\mathcal{G}$
$\mathbf{T}$	Markov state transition matrix
$T_{i,j}$	$(i, j)$ th element of $\mathbf{T}$
$\mathbf{Q}$	Sub-matrix of $\mathbf{T}$
$\mathbf{N}$	Fundamental matrix of $\mathbf{T}$
$n_{i,j}$	$(i, j)$ th element of $\mathbf{N}$
$\hat{F}_{N_{\mathcal{G} \rightarrow X}}[n]$	Estimated CDF of $N_{\mathcal{G} \rightarrow X}$
$\hat{\sigma}_{N_{\mathcal{G} \rightarrow X}}$	Standard deviation of error in $\hat{F}_{N_{\mathcal{G} \rightarrow X}}[n]$

is an absorbing state, i.e.,  $\Pr\{D_a[t] = 0 | N_D[t-1] = N_u - 1\} = 1$  as shown in Eq. (3).

2) *Mean Time to Absorption:* For  $t > 0$ ,  $\pi_j[t] = \sum_{i=0}^{N_u} \pi_i[t-1]T_{i,j}$ , and  $\boldsymbol{\pi}[t] = \boldsymbol{\pi}[t-1]\mathbf{T}$ , so that  $\boldsymbol{\pi}[t] = \boldsymbol{\pi}[0]\mathbf{T}^n$ . We can get the CDF of  $N_{\mathcal{G} \rightarrow X}$  since the  $(0, N_u - 1)$ th element of  $\mathbf{T}^n$  is

$$\begin{aligned} (\mathbf{T}^n)_{0, N_u-1} &= \Pr\{N_D[t] = N_u - 1 | N_D[0] = 0\} \\ &= \Pr\{N_{\mathcal{G} \rightarrow X} \leq n\}. \end{aligned} \quad (4)$$

$N_D[t] = N_u - 1$  is the lone absorbing state for this Markov chain (i.e.,  $T_{N_u-1, N_u-1} = 1$ ), and all other states are transient (i.e.,  $T_{i,i} < 1$  for  $i \neq N_u - 1$ ). Because the absorbing state is reachable from all other states, this is an absorbing Markov chain. Given that we start in State  $N_D[0] = 0$ , we can determine the distribution of  $N_{\mathcal{G} \rightarrow X}$ .

To get an expression for the mean number of periods to reach the absorbing state, we use the chain's fundamental matrix, which we derive using the approach given by Grinstead and Snell [10, Section 11.2]. We start by partitioning  $\mathbf{T}$  as follows:

$$\mathbf{T} = \begin{bmatrix} \mathbf{Q} & \mathbf{r} \\ \mathbf{0}_{N_u-1} & 1 \end{bmatrix}, \quad (5)$$

where  $\mathbf{Q}$  is a  $(N_u - 1) \times (N_u - 1)$  matrix whose  $(i, j)$ th element is  $T_{i,j}$ ,  $\mathbf{r}$  is a length- $(N_u - 1)$  column vector whose  $i$ th element is  $r_i = \Pr\{D_a[t] = (N_u - 1) - i | N_D[t-1] = i\}$ , and  $\mathbf{0}_{N_u-1}$  is a length- $(N_u - 1)$  all-zero row vector. Thus

$$\mathbf{T}^2 = \begin{bmatrix} \mathbf{Q}^2 & (\mathbf{I} + \mathbf{Q})\mathbf{r} \\ \mathbf{0}_{N_u-1} & 1 \end{bmatrix},$$

where  $\mathbf{I}$  is the  $(N_u - 1) \times (N_u - 1)$  identity matrix. In general, by recursion and the matrix form of the geometric series,

$$\mathbf{T}^k = \begin{bmatrix} \mathbf{Q}^k & (\mathbf{I} - \mathbf{Q}^k)(\mathbf{I} - \mathbf{Q})^{-1}\mathbf{r} \\ \mathbf{0}_{N_u-1} & 1 \end{bmatrix}, \quad (6)$$

where  $\mathbf{Q}^0 = \mathbf{I}$ . The fundamental matrix is

$$\mathbf{N} = [n_{i,j}] = \sum_{\ell=0}^{\infty} \mathbf{Q}^{\ell} = (\mathbf{I} - \mathbf{Q})^{-1}, \quad (7)$$

where  $n_{i,j} = \sum_{\ell=0}^{\infty} \Pr\{N_D[\ell] = j | N_D[0] = i\}$ , is the mean number of times the chain visits transient State  $j$  given that it started in transient State  $i$ . Thus the mean of the total number of transient state visits after starting in state  $N_D[0] = 0$  is

$$E\{N_{\mathcal{G} \rightarrow X}\} = \sum_{j=0}^{N_u-2} n_{0,j}, \quad (8)$$

which is also the mean number of PSDCH periods required to reach the absorbing state.

### B. The Probability of Collided Discovery Messages

We define  $P_C(n | N_r; N_u)$  to be the probability that  $N_u$  UEs using a pool of  $N_r$  resources experience  $n$  collisions, where we assume that a collision occurs if two or more discovery messages occupy a given resource. (For all values of  $N_r$  and  $N_u$ ,  $P_C(1 | N_r; N_u) = 0$ .) Each UE independently picks a resource at random, which is equivalent to sampling with replacement  $N_u$  times from the set  $\{1, 2, \dots, N_r\}$ . The number of ways for  $N_u$  UEs to choose a set of resources is  $N_r^{N_u}$ . We consider two cases, based on the ratio of the number of resources to the number of UEs.

1)  $N_r \geq N_u$ : The probability that no collisions occur,  $P_C(0 | N_r; N_u)$ , is the probability that the UEs will choose resources so that only one UE uses each utilized resource in the pool. The number of ways that this can happen is  $(N_r)(N_r - 1)(N_r - 2) \cdots (N_r - N_u + 1)$ . Thus,

$$P_C(0 | N_r; N_u) = \frac{N_r! / (N_r - N_u)!}{N_r^{N_u}} = \frac{N_u!}{N_r^{N_u}} \binom{N_r}{N_u}. \quad (9)$$

For  $2 \leq n \leq N_u$ , we use the occupancy vector  $\mathbf{x}$  for the set of resources. The occupancy vector as defined by Feller [11] is the ordered length- $N_r$  vector  $\mathbf{x} = [x_1, x_2, \dots, x_{N_r}]$ , which indicates the allocation of UEs among the set of resources without identifying which UEs have chosen a particular resource. When there are  $n$  collisions,  $\mathbf{x}$  has the form

$$\mathbf{x} = \underbrace{[0, 0, \dots, 0]}_{N_r - (N_u - n) - s}, \underbrace{[1, 1, \dots, 1]}_{N_u - n}, \underbrace{[k_1, k_2, \dots, k_s]}_s,$$

where  $s$  is the number of resources that are occupied by two or more UEs. The set of occupancy numbers associated with collided UEs thus form an occupancy sub-vector  $\mathbf{k} = [k_1, k_2, \dots, k_s]$  that has the following two properties:

$$k_1 + k_2 + \dots + k_s = n \quad (10a)$$

$$2 \leq k_1 \leq k_2 \leq \dots \leq k_s \quad (10b)$$

Next, we define the vector  $\mathbf{d}(\mathbf{k}) = [d_1, d_2, \dots, d_{\mathcal{U}(\mathbf{k})}]$ , where  $\mathcal{U}(\mathbf{k}) \in \{1, 2, \dots, s\}$  is the number of distinct elements of  $\mathbf{k}$ , and  $d_i$  is number of occurrences of the  $i$ th distinct element of  $\mathbf{k}$ . For example, if the number of collisions is  $n = 17$ , then one possible occupancy vector is  $\mathbf{x} = [0, 0, \dots, 0, 1, 1, \dots, 1, 2, 2, 2, 3, 5, 5]$ . In this case,  $\mathbf{k} = [2, 2, 2, 3, 5, 5]$ , whose distinct elements are 2, 3, and 5; thus  $\mathcal{U}(\mathbf{k}) = 3$ , and  $\mathbf{d}(\mathbf{k}) = [3, 1, 2]$ .

The number of ways that  $N_r$  resources can be arranged into  $\mathcal{U}(\mathbf{k}) + 2$  groups, where each resource in a given group has been chosen by the same number of UEs, is

$$\frac{N_r!}{(N_r - (N_u - n) - s)! (N_u - n)! d_1! d_2! \cdots d_{\mathcal{U}(\mathbf{k})}!} \quad (11)$$

and the number of ways that the  $N_u$  UEs can be arranged into  $(N_u - n) + s$  groups, where each group corresponds to an occupied resource, is

$$\frac{N_u!}{\underbrace{1! 1! \cdots 1!}_{N_u - n} k_1! k_2! \cdots k_s!} = \frac{N_u!}{k_1! k_2! \cdots k_s!}. \quad (12)$$

Taking the product of Eq. (11) and Eq. (12), multiplying both numerator and denominator by  $n!$ , and simplifying, we get the number of ways that the  $N_u$  UEs in  $\mathcal{G}$  can choose resources so that there are  $n$  collisions that produce the length- $s$  occupancy sub-vector  $\mathbf{k} = [k_1, k_2, \dots, k_s]$ :

$$\frac{N_r! \binom{N_u}{n} \binom{n}{k_1, k_2, \dots, k_s}}{(N_r - (N_u - n) - s)! d_1! d_2! \cdots d_{\mathcal{U}(\mathbf{k})}!}. \quad (13)$$

To get the probability of  $n$  collisions, we must divide the number of ways to arrange  $N_u$  resource choices such that there are  $n$  collisions by  $N_r^{N_u}$ . We get the numerator by summing over all possible occupancy sub-vectors  $\mathbf{k}$  that produce  $n$  collisions. The sub-vector length,  $s$ , varies from  $s = 1$  (in which case  $\mathbf{k} = [n]$ ) to  $s = \lfloor n/2 \rfloor$ , since the longest possible occupancy sub-vector is  $\mathbf{k} = [2, 2, \dots, 2, 2]$  if  $n$  is even or  $\mathbf{k} = [2, 2, \dots, 2, 3]$  if  $n$  is odd. For a given occupancy sub-vector length  $s$ , we sum over all sub-vectors  $[k_1, k_2, \dots, k_s]$  that satisfy Eqs. (10a) and (10b).

Summing Eq. (13) over all possible occupancy sub-vectors, and dividing the result by  $N_r^{N_u}$ , we get the probability of  $n$  collisions:

$$\begin{aligned} P_C(n | N_r; N_u) &= \frac{\binom{N_u}{n}}{N_r^{N_u}} \sum_{s=1}^{\lfloor n/2 \rfloor} \frac{N_r!}{(N_r - (N_u - n) - s)!} \\ &\quad \times \sum_{\substack{s \\ \sum_{i=1}^s k_i = n \\ 2 \leq k_1 \leq k_2 \leq \dots \leq k_s}} \frac{1}{d_1! d_2! \cdots d_{\mathcal{U}(\mathbf{k})}!} \binom{n}{k_1, k_2, \dots, k_s} \end{aligned} \quad (14)$$

for  $2 \leq n \leq N_u$ , when  $N_u \leq N_r$ .

To get the number of occupancy sub-vectors  $\mathbf{k}$ , we note that creating  $\mathbf{k}$  is analogous to distributing  $n$  objects into  $s$  bins by first putting one object into each bin, and then distributing the remaining  $(n - s)$  objects among the  $s$  bins so that at least one object goes into each bin. The number of ways to do this is the number of ways to partition a set of  $n - s$  identical objects into  $s$  non-empty subsets, which is the set partition number  $\Pi(n - s, s)$ . There is no closed form expression for this number; we must use the following recurrence relation from Martin [12, p. 35]:

$$\Pi(i, j) = \Pi(i - 1, j - 1) + \Pi(i - j, j) \quad (15)$$

where  $\Pi(i, j)$  is the number of ways to partition  $i$  indistinguishable objects into  $j$  non-empty, indistinguishable groups, where  $\Pi(i, i) = \Pi(i, 1) = 1$  for all  $i$ , and  $\Pi(i, j) = 0$  for  $j > i$  [12, p. 35].

2)  $N_r < N_u$ : For the case where there are more UEs than resources, the number of collisions cannot be zero, since it is impossible to distribute the UEs in such a way that there is one UE per resource. The minimum number of collisions occurs when the occupancy vector has the form

$$\mathbf{x} = \underbrace{[1, 1, \dots, 1]}_{N_r - 1}, N_u - (N_r - 1),$$

so that  $\mathbf{k} = [N_u - N_r + 1]$ . Thus  $P_C(n | N_r; N_u)$  is non-zero for  $N_u - N_r + 1 \leq n \leq N_u$ ; and when  $N_u > N_r$ ,  $N_u - N_r + 1 \geq 2$ . Secondly, the maximum possible length for  $\mathbf{k}$  is constrained by the fact that the number of initial zeros in the occupancy vector,  $N_r - (N_u - n) - s$ , cannot be negative, i.e.,

$$s \leq N_r - N_u + n. \quad (16)$$

Applying Eq. (10b) to Eq. (10a) gives  $2s \leq n$ . Thus, when  $N_r \geq N_u$ , Eq. (16) always holds. However, when  $N_r < N_u$ , Eq. (16) becomes an additional constraint on the length of  $\mathbf{k}$  (e.g., if  $n = N_u$ , then  $s \leq N_r$ ), so that the maximum value of  $s$  is  $\min(N_r - N_u + n, \lfloor n/2 \rfloor)$ . Summing Eq. (13) over all possible occupancy sub-vectors and dividing by  $N_r^{N_u}$ , we get the following general expression for  $P_C(n | N_r; N_u)$ :

$$\begin{aligned} P_C(n | N_r; N_u) &= \left( \binom{N_u}{n} / N_r^{N_u} \right) \sum_{s=1}^{\min(N_r - N_u + n, \lfloor n/2 \rfloor)} \frac{N_r!}{(N_r - (N_u - n) - s)!} \\ &\quad \times \sum_{\substack{\sum_{i=1}^s k_i = n \\ 2 \leq k_1 \leq k_2 \leq \dots \leq k_s}} \frac{1}{d_1! d_2! \dots d_{\mathcal{U}(\mathbf{k})}!} \binom{n}{k_1, k_2, \dots, k_s}, \end{aligned} \quad (17)$$

for  $\max(2, N_u - N_r + 1) \leq n \leq N_u$ .

If we compare Eq. (14) and Eq. (17), we see that if  $N_r \geq N_u$ , then  $\min(N_r - N_u + n, \lfloor n/2 \rfloor) = \lfloor n/2 \rfloor$ , so that Eq. (17) gives  $P_C(n | N_r; N_u)$  for both cases.

### C. The Markov State Transition Probabilities

With  $P_C(n | N_r; N_u)$  in hand, we can get the elements of  $\mathbf{T}$ . We now derive  $\Pr\{D_a[t] = \nu | N_D[t-1] = M_D\}$ , the probability that UE  $X$  discovers  $n$  UEs in the  $t$ th period given UE  $X$  has discovered  $M_D$  UEs already. First, we condition on the set of events where  $m_d \leq M_D$  discovered UEs and  $m_u \leq M_U$  undiscovered UEs do not choose resources in  $S_X$ , the set of subframes in which UE  $X$  transmits<sup>1</sup>, so that their discovery messages can be received by UE  $X$ . If  $m_u < \nu$ ,  $\Pr\{D_a[t] = \nu | m_d + m_u \text{ not in } S_X\} = 0$ . Also,

$$\begin{aligned} &\Pr\{m_d + m_u \text{ not in } S_X\} \\ &= \Pr\{m_d \text{ not in } S_X\} \Pr\{m_u \text{ not in } S_X\} \\ &= \binom{M_D}{m_d} \left(1 - \frac{1}{N_t}\right)^{m_d} \left(\frac{1}{N_t}\right)^{M_D - m_d} \\ &\quad \times \binom{M_U}{m_u} \left(1 - \frac{1}{N_t}\right)^{m_u} \left(\frac{1}{N_t}\right)^{M_U - m_u}. \end{aligned} \quad (18)$$

Applying both of these conditions, we get

$$\begin{aligned} &\Pr\{D_a[t] = \nu | N_D[t-1] = M_D\} \\ &= \sum_{m_d=0}^{M_D} \sum_{m_u=\nu}^{M_U} \Pr\{D_a[t] = \nu | m_d + m_u \text{ not in } S_X\} \\ &\quad \times \binom{M_D}{m_d} \binom{M_U}{m_u} \left(1 - \frac{1}{N_t}\right)^{m_d + m_u} \left(\frac{1}{N_t}\right)^{M_D + M_U - m_d - m_u}. \end{aligned} \quad (19)$$

<sup>1</sup>Note that  $M_U = (N_u - 1) - M_D$ .

We evaluate  $\Pr\{D_a[t] = \nu | m_d + m_u \text{ not in } S_X\}$  in Eq. (19) by conditioning on the value of  $\rho$ , the number of resources occupied by the  $m_d$  discovered UEs that do not choose resources in  $S_X$ , where  $0 \leq \rho \leq m_d$  (we denote the event “ $m_d$  discovered UEs occupy  $\rho$  resources” as  $\{m_d \Rightarrow \rho\}$ ). First, we prove the following lemma.

*Lemma 1:* Given an event  $A$  that depends on mutually independent events  $B_1$  and  $B_2$ , if there exists a set of events  $\{C_i\}_{i=1}^N$  that are mutually independent ( $C_i \cap C_j = \emptyset$  if  $i \neq j$ ) such that for  $i = 1, 2, \dots, N$ ,  $C_i \subseteq B_1$  and  $C_i \cap B_2 = \emptyset$ , then  $\Pr\{A | B_1 \cap B_2\} = \sum_{i=1}^N \Pr\{A | B_2 \cap C_i\} \Pr\{C_i | B_1\}$ .

*Proof:*

$$\begin{aligned} &\Pr\{A | B_1 \cap B_2\} \\ &= \frac{\Pr\{A \cap B_1 \cap B_2\}}{\Pr\{B_1 \cap B_2\}} = \sum_{i=1}^N \frac{\Pr\{A \cap B_1 \cap B_2 | C_i\} \Pr\{C_i\}}{\Pr\{B_1\} \Pr\{B_2\}} \\ &= \sum_{i=1}^N \frac{\Pr\{A \cap B_1 \cap B_2 \cap C_i\} \Pr\{B_1 \cap C_i\} \Pr\{B_2\}}{\Pr\{B_1 \cap B_2 \cap C_i\} \Pr\{B_1\} \Pr\{B_2\}} \\ &= \sum_{i=1}^N \Pr\{A | B_1 \cap B_2 \cap C_i\} \Pr\{C_i | B_1\}. \end{aligned}$$

Since  $B_1 \cap C_i = C_i$  for  $i = 1, 2, \dots, N$ ,  $\Pr\{A | B_1 \cap B_2\} = \sum_{i=1}^N \Pr\{A | B_2 \cap C_i\} \Pr\{C_i | B_1\}$ . ■

There are  $\tilde{N}_r = N_r - N_f$  resources available to the  $m_d + m_u$  UEs that did not choose resources in  $S_X$ . Using Lemma 1, we get

$$\begin{aligned} &\Pr\{D_a[t] = \nu | m_d + m_u \text{ not in } S_X\} \\ &= \sum_{\rho=0}^{m_d} \Pr\{D_a[t] = \nu | \{m_d \Rightarrow \rho\} \cap \{m_u \text{ not in } S_X\}\} \\ &\quad \times \Pr\{m_d \Rightarrow \rho | m_d \text{ not in } S_X\}, \end{aligned} \quad (20)$$

where  $\Pr\{m_d \Rightarrow \rho | m_d \text{ not in } S_X\}$  is the probability that the occupancy vector for  $\tilde{N}_r$  resources has the form

$$\underbrace{[0, 0, \dots, 0]}_{\tilde{N}_r - \rho}, \underbrace{[k_1, k_2, \dots, k_\rho]}_{d_1, d_2, \dots, d_{\mathcal{U}(\mathbf{k})}}$$

and where  $\mathcal{U}(\mathbf{k})$  is the number of distinct elements of  $\mathbf{k} = [k_1, k_2, \dots, k_\rho]$ ,  $d_i$  is the number of times the  $i$ th distinct element of  $\mathbf{k}$  appears in  $\mathbf{k}$ , and  $\mathbf{k}$  is subject to the following restrictions:

$$k_1 + k_2 + \dots + k_\rho = m_d \quad (21a)$$

$$1 \leq k_1 \leq k_2 \leq \dots \leq k_\rho \quad (21b)$$

The number of ways to divide  $\tilde{N}_r$  resources into  $\mathcal{U}(\mathbf{k}) + 1$  groups of sizes  $\tilde{N}_r - \rho, d_1, d_2, \dots, d_{\mathcal{U}(\mathbf{k})}$  is

$$\binom{\tilde{N}_r}{\tilde{N}_r - \rho, d_1, d_2, \dots, d_{\mathcal{U}(\mathbf{k})}} = \frac{\tilde{N}_r!}{(\tilde{N}_r - \rho)! d_1! d_2! \dots d_{\mathcal{U}(\mathbf{k})!}},$$

and the number of ways to divide  $m_d$  discovered UEs into groups of size  $k_1, k_2, \dots, k_\rho$  is

$$\binom{m_d}{k_1, k_2, \dots, k_\rho} = \frac{m_d!}{k_1! k_2! \dots k_\rho!}.$$

The total number of ways to get this particular value of  $\mathbf{k}$  is the product of the two multinomials. By tallying the number of ways that all occupancy vectors with  $\rho$  non-zero elements occur, and then dividing by  $\tilde{N}_r^{m_d}$ , the total number of possible outcomes, we get the probability that  $m_d$  discovered UEs that have chosen resources outside  $S_X$  occupy  $\rho$  resources:

$$\Pr\{m_d \Rightarrow \rho \mid m_d \text{ not in } S_X\} = \frac{\rho!}{\tilde{N}_r^{m_d}} \binom{\tilde{N}_r}{\rho} \sum_{\substack{\sum_{i=1}^{\rho} k_i = m_d \\ 1 \leq k_1 \leq k_2 \leq \dots \leq k_\rho}} \frac{\binom{m_d}{k_1, k_2, \dots, k_\rho}}{d_1! d_2! \dots d_{\mathcal{U}(\mathbf{k})}!}. \quad (22)$$

There are three special cases that apply to Eq. (22):

- If  $m_d = 0$ , then  $\Pr\{m_d \Rightarrow \rho \mid m_d \text{ not in } S_X\} = 1$  for  $\rho = 0$  and  $\Pr\{m_d \Rightarrow \rho \mid m_d \text{ not in } S_X\} = 0$  for  $\rho > 0$ .
- If  $m_d = 1$ , then  $\Pr\{m_d \Rightarrow \rho \mid m_d \text{ not in } S_X\} = 1$  for  $\rho = 1$  and  $\Pr\{m_d \Rightarrow \rho \mid m_d \text{ not in } S_X\} = 0$  for  $\rho \neq 1$ .
- If  $m_d > 0$ , then  $\Pr\{m_d \Rightarrow \rho \mid m_d \text{ not in } S_X\} = 0$  for  $\rho = 0$ .

The next step in the derivation of  $\mathbf{T}$  is to develop an expression for  $\Pr\{D_a[t] = \nu \mid \{m_d \Rightarrow \rho\} \cap \{m_u \text{ not in } S_X\}\}$  in Eq. (20). UE  $X$  will detect one of the  $m_u$  undiscovered UEs that did not choose resources in  $S_X$  if the undiscovered UE does not choose the same resource as any other UEs, either discovered or undiscovered. To compute  $\Pr\{D_a[t] = \nu \mid \{m_d \Rightarrow \rho\} \cap \{m_u \text{ not in } S_X\}\}$ , we first condition on the number of undiscovered UEs out of  $m_u$  that use the  $\rho$  resources being used by the  $m_d$  discovered UEs that chose resources that are not in  $S_X$ . We define  $N_{O_{u,d}}$  to be the number of undiscovered UEs not in  $S_X$  that overlap with discovered UEs. Since  $\{N_{O_{u,d}} = \ell\} \subseteq \{m_u \text{ not in } S_X\}$ , by Lemma 1,

$$\Pr\{D_a[t] = \nu \mid \{m_d \Rightarrow \rho\} \cap \{m_u \text{ not in } S_X\}\} = \sum_{\ell=0}^{m_u} \Pr\{D_a[t] = \nu \mid \{m_d \Rightarrow \rho\} \cap \{N_{O_{u,d}} = \ell\}\} \times \Pr\{N_{O_{u,d}} = \ell \mid \{m_d \Rightarrow \rho\}\}. \quad (23)$$

We first compute  $\Pr\{N_{O_{u,d}} = \ell \mid \{m_d \Rightarrow \rho\}\}$ . The event  $\{N_{O_{u,d}} = \ell\}$  occurs when  $\ell$  out of  $m_u$  undiscovered UEs that are not in  $S_X$  choose resources that are among the  $\rho$  out of  $N_r$  resources that are occupied by the  $m_d$  discovered UEs that did not choose resources in  $S_X$ . The probability that one of the  $m_u$  undiscovered UEs picks one of the  $\rho$  resources, given that it is not in  $S_X$ , is  $(\rho/N_r)/(\tilde{N}_r/N_r) = \rho/\tilde{N}_r$ . Thus, the probability that  $\ell$  of the  $m_u$  UEs picked one of the  $\rho$  resources is

$$\Pr\{N_{O_{u,d}} = \ell \mid \{m_d \Rightarrow \rho\}\} = \binom{m_u}{\ell} \left(\frac{\rho}{\tilde{N}_r}\right)^\ell \left(1 - \frac{\rho}{\tilde{N}_r}\right)^{m_u - \ell}. \quad (24)$$

The last component of the state transition probability that remains to be derived is  $\Pr\{D_a[t] = \nu \mid \{m_d \Rightarrow \rho\} \cap \{N_{O_{u,d}} = \ell\}\}$  in Eq. (23). Since  $\ell$  undiscovered UEs out of  $m_u$  chose the same resources as some of the  $m_d$  discovered UEs, there

are  $m_u - \ell$  undiscovered UEs that UE  $X$  could discover. If UE  $X$  discovers  $\nu$  UEs out of the  $m_u - \ell$  undiscovered UEs, where  $0 \leq \nu \leq m_u - \ell$ , then there were  $m_u - \ell - \nu$  collisions among the  $m_u - \ell$  undiscovered UEs as they chose from the set of  $\tilde{N}_r - \rho$  resources. The probability of this number of collisions is  $P_C(m_u - \ell - \nu \mid \tilde{N}_r - \rho; m_u - \ell)$ . Thus we have

$$\Pr\{D_a[t] = \nu \mid \{m_d \Rightarrow \rho\} \cap \{N_{O_{u,d}} = \ell\}\} = P_C(m_u - \ell - \nu \mid \tilde{N}_r - \rho; m_u - \ell). \quad (25)$$

Note that for the collision probability to be non-zero,  $0 \leq \ell \leq m_u - \nu$ .

Using Eq. (24) and Eq. (25), and applying the limits on the values of  $\ell$ , we can write Eq. (23) as

$$\Pr\{D_a[t] = \nu \mid \{m_d \Rightarrow \rho\} \cap \{m_u \text{ not in } S_X\}\} = \sum_{\ell=0}^{m_u - \nu} \binom{m_u}{\ell} \left(\frac{\rho}{\tilde{N}_r}\right)^\ell \left(1 - \frac{\rho}{\tilde{N}_r}\right)^{m_u - \ell} \times P_C(m_u - \ell - \nu \mid \tilde{N}_r - \rho; m_u - \ell). \quad (26)$$

Inserting Eq. (22) and Eq. (26) into Eq. (20) gives

$$\Pr\{D_a[t] = \nu \mid m_d + m_u \text{ not in } S_X\} = \sum_{\rho=0}^{m_d} \sum_{\ell=0}^{m_u - \nu} \binom{m_u}{\ell} \left(\frac{\rho}{\tilde{N}_r}\right)^\ell \left(1 - \frac{\rho}{\tilde{N}_r}\right)^{m_u - \ell} \times P_C(m_u - \ell - \nu \mid \tilde{N}_r - \rho; m_u - \ell) \times \frac{\rho!}{\tilde{N}_r^{m_d}} \binom{\tilde{N}_r}{\rho} \sum_{\substack{\sum_{i=1}^{\rho} k_i = m_d \\ 1 \leq k_1 \leq k_2 \leq \dots \leq k_\rho}} \frac{\binom{m_d}{k_1, k_2, \dots, k_\rho}}{d_1! d_2! \dots d_{\mathcal{U}(\mathbf{k})}!}. \quad (28)$$

Finally, applying Eq. (28) to Eq. (19), and rearranging the order of summation, we get Eq. (28). Using Eq. (28) in Eq. (2), we can generate the Markov chain's state transition matrix  $\mathbf{T}$  by letting  $\nu = j - i$  and  $M_D = N_u - 1 - i$  for the set of ordered pairs  $\{(i, j) \mid j = i, i + 1, \dots, N_u - 1; i = 0, 1, \dots, N_u - 1\}$ , using the indexing scheme shown in Eq. (3). We can get  $E\{N_{G \rightarrow X}\}$  by applying Eq. (5), Eq. (7), and Eq. (8), and we can get the CDF of  $N_{G \rightarrow X}$  by using Eq. (6) and taking the  $(0, N_u - 1)$ th element of  $\mathbf{T}^k$ .<sup>2</sup>

#### IV. NUMERICAL RESULTS

In this section, we validate the theoretical results from Section III, using both Monte Carlo simulations in Matlab and simulation of a group of UEs in NS3. We also use the theoretical model to determine the maximum group size the allows a single UE to discover all other members of the group within a given amount of time, with a given level of certainty.

For the Monte Carlo simulations, we used a resource pool composed of  $N_t = 5$  subframe sets and  $N_f = 10$  PRB pairs. The Monte Carlo simulation consisted of  $N_{\text{runs}} = 50$  runs,

<sup>2</sup>We note that the computational cost associated with these results can be significant. A desktop using an Intel Xeon CPU with a 3.3 GHz clock rate and 16 GBytes of Random Access Memory (RAM) produced Fig. 2 in about 3000 s, with the Monte Carlo results taking up approximately 2 s.

$$\begin{aligned}
& \Pr\{D_a[t] = \nu \mid N_D[t-1] = M_D\} \\
&= \sum_{m_d=0}^{M_D} \sum_{m_u=\nu}^{M_U} \binom{M_D}{m_d} \binom{M_U}{m_u} \left(1 - \frac{1}{N_t}\right)^{m_d+m_u} \left(\frac{1}{N_t}\right)^{M_D+M_U-m_d-m_u} \sum_{\rho=0}^{m_d} \frac{\rho!}{\tilde{N}_r^{m_d}} \binom{\tilde{N}_r}{\rho} \left[ \sum_{\substack{\sum_{i=1}^{\rho} k_i = m_d \\ 1 \leq k_1 \leq k_2 \leq \dots \leq k_\rho}} \frac{\binom{m_d}{k_1, k_2, \dots, k_\rho}}{d_1! d_2! \dots d_U(\mathbf{k})!} \right] \\
&\quad \times \sum_{\ell=0}^{m_u-\nu} \binom{m_u}{\ell} \left(\frac{\rho}{\tilde{N}_r}\right)^\ell \left(1 - \frac{\rho}{\tilde{N}_r}\right)^{m_u-\ell} P_C(m_u - \ell - n \mid \tilde{N}_r - \rho; m_u - \ell).
\end{aligned} \tag{28}$$

with  $N_{\text{trials}} = 100$  trials per run. The UE group contained  $N_u = 51$  UEs, including UE  $X$ . Each trial consisted of a sequence of periods in which the number of undiscovered UEs was initialized to  $N_u - 1$  and the number of discovered UEs was set to zero. In each period, a  $N_f \times N_t$  matrix was populated with randomly placed messages associated with the discovered and undiscovered UEs that were represented using complex numbers, each of whose real part was the number of undiscovered UEs and whose complex part was the number of discovered UEs. One column was chosen at random to contain  $\delta_X$ ; all messages in this column were lost by setting all column elements to zero. The simulation determined the number of successful discoveries of new UEs by counting the matrix elements whose real part was equal to unity. The simulation then adjusted the numbers of discovered and undiscovered UEs and moved to the next period, using the reduction of the number of undiscovered UEs to zero as the stopping criterion.

We plot the theoretical and simulation-based CDFs together in Fig. 2. Because  $N_{\mathcal{G} \rightarrow X}$  is a discrete random variable, the CDF assumes the staircase form seen in the figure, with point discontinuities indicated by pairs of closed circles and open circles that are connected by dashed lines. Fig. 2 shows uncertainty in the simulation results using two methods. We compute the estimated CDF at a given index value  $n$  as

$$\begin{aligned}
\hat{F}_{N_{\mathcal{G} \rightarrow X}}[n] &= \frac{1}{N_{\text{runs}}} \sum_{i=1}^{N_{\text{runs}}} \hat{F}_{N_{\mathcal{G} \rightarrow X}, i}[n] \\
&= \frac{1}{N_{\text{runs}}} \sum_{i=1}^{N_{\text{runs}}} \frac{\mathcal{N}(\{N_{\mathcal{G} \rightarrow X, i, j}\}_{j=1}^{N_{\text{trials}}} \leq n)}{N_{\text{trials}}}, \tag{29}
\end{aligned}$$

where  $\hat{F}_{N_{\mathcal{G} \rightarrow X}, i}[n]$  is the estimate of  $F_{N_{\mathcal{G} \rightarrow X}}[n]$  based on the trials that compose the  $i$ th run,  $\{N_{\mathcal{G} \rightarrow X, i, j}\}_{j=1}^{N_{\text{trials}}}$  is the set of simulation outputs generated during the  $i$ th run, and where  $\mathcal{N}(\{N_{\mathcal{G} \rightarrow X, i, j}\}_{j=1}^{N_{\text{trials}}} \leq n)$  is the number of simulation trial outputs during the  $i$ th run that were less than or equal to  $n$ . From the set of simulation results, we can create a set of pointwise 95 % confidence intervals which appear as light gray bars in the figure, and whose limits for a given value of  $n$  are  $\hat{F}_{N_{\mathcal{G} \rightarrow X}}[n] \pm 1.96 \hat{\sigma}_{N_{\mathcal{G} \rightarrow X}} / \sqrt{N_{\text{runs}}}$ , where  $\hat{\sigma}_{N_{\mathcal{G} \rightarrow X}}^2$  is the variance of the simulation results and is

$$\hat{\sigma}_{N_{\mathcal{G} \rightarrow X}}^2 = \frac{1}{N_{\text{runs}} - 1} \sum_{i=1}^{N_{\text{runs}}} (\hat{F}_{N_{\mathcal{G} \rightarrow X}, i}[n] - \hat{F}_{N_{\mathcal{G} \rightarrow X}}[n])^2.$$

Fig. 2 also shows a confidence envelope based on Masart's refinement of the Dvoretzky-Kiefer-Wolfowitz (DKW) inequality, which states that, given a set of random variates  $\{X_i\}_{i=1}^N$  that are drawn from a distribution with CDF  $F_X$  and that produce an empirical CDF  $\hat{F}_X$ , the probability that the true and empirical CDFs are separated by more than  $\varepsilon \in \mathbb{R}^+$  over their entire support has the following upper bound [13]:

$$\Pr\{\sup_{x \in \mathbb{R}} |\hat{F}_X(x) - F_X(x)| > \varepsilon\} \leq 2e^{-2N\varepsilon^2}. \tag{30}$$

Note that the upper bound on the excursion probability in Eq. (30) decreases to zero for any  $\varepsilon > 0$  as  $N$  increases, and we can easily extend the inequality to discrete random variables, since the resulting CDF is still defined over the whole real line.

The 95 % confidence interval associated with the DKW inequality is actually a confidence envelope over the entire domain of interest, whose upper and lower bounds are defined by the empirical CDF plus or minus the error offset  $\varepsilon$ . Since we are interested in interval bounds that result in the theoretical curve lying entirely within the confidence envelope with 95 % probability, we can set the upper bound in Eq. (30) equal to 0.05, the probability that the theoretical CDF deviates from the empirical CDF by more than  $\varepsilon$ , and then solve for  $\varepsilon$ , giving

$$\varepsilon = \sqrt{\log(2/0.05)/(N_{\text{runs}} N_{\text{trials}})}, \tag{31}$$

since the empirical CDF is constructed from  $N = N_{\text{runs}} N_{\text{trials}}$  variates. We plot the resulting confidence envelope  $\hat{F}_{N_{\mathcal{G} \rightarrow X}}[n] \pm \varepsilon$  in Fig. 2 using dark gray bars. Note that the pointwise 95 % confidence intervals are contained within the 95 % confidence envelope, which is a looser interval because it covers the entire domain. Note also that the pointwise confidence intervals are widest near the median and become narrower in the distribution's tails.

The resulting plots in Fig. 2 show excellent agreement between the theoretical and simulation results, with the theoretical curve lying entirely within the narrower pointwise confidence intervals over the entire domain. We also used Eq. (8) to get the associated theoretical expected value of  $N_{\mathcal{G} \rightarrow X}$ , which is  $E\{N_{\mathcal{G} \rightarrow X}\} = 11.5566$  PSDCH periods. The 95 % confidence interval for the corresponding estimated expected value of  $N_{\mathcal{G} \rightarrow X}$  that we obtained from the Monte Carlo simulations is  $E\{N_{\mathcal{G} \rightarrow X}\} = 11.5988 \pm 0.1007$  PSDCH periods, which also indicates close agreement.

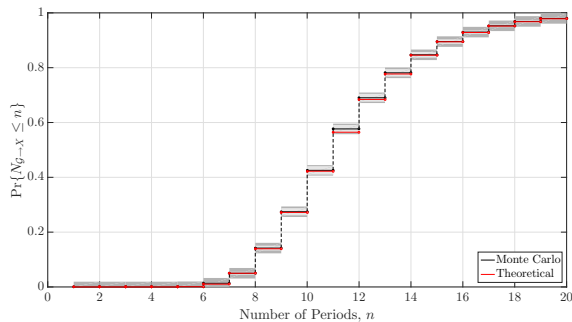


Fig. 2. CDF plots using Eq. (28) and corresponding Monte Carlo simulation results, with pointwise and envelope 95 % confidence intervals shown, plotted versus  $n$ , the number of PSDCH periods, for  $N_t = 5$  subframe sets,  $N_f = 10$  PRB pairs, and  $N_u = 51$  UEs.

For the validation using NS3, we examined six scenarios. We examined two resource pool configurations:  $N_f = 5$  PRB pairs and  $N_f = 10$  PRB pairs, with  $N_t = 10$  subframe sets in both cases. For each resource pool configuration, we considered three D2D group sizes: 10 UEs, 20 UEs, and 30 UEs. For each scenario, we performed 10 runs, with 500 trials per run. In each trial, we generated a set of uniformly distributed UEs within an area sufficiently small to that each UE could receive messages from every other UE (e.g., a 10000 m<sup>2</sup> (100 m × 100 m) square). Each UE transmitted a discovery message in every period using a randomly chosen pool resource, and we recorded the number of periods required for the UE to discover all of the other UEs in the group. We used these results to generate a sample CDF for each run, and we averaged the ensemble over all runs to produce our estimate of the CDF of  $N_{G \rightarrow X}$ .

We show the results for the six cases in Fig. 3, with 95 % confidence envelopes for each case, and the corresponding theoretical CDF plotted in each subfigure for comparison. These NS3 results also agree closely with the theoretical model. The results also illustrate the effect of increasing the size of the discovery resource pool. Comparing Figs. 3a-3c with Figs. 3d-3f shows that doubling  $N_r$  produces a noticeable leftward shift in the CDF for each group size, and that the CDFs for the various group sizes are closely spaced when  $N_r = 100$  resources, while the CDF associated with  $N_u = 30$  UEs indicates degraded performance due to increased collisions by UEs using the smaller pool.

Using the CDF, we can calculate the maximum group size that achieves a desired level of performance. Using the Monte Carlo simulations that we used to produce Fig. 2, we plot in Fig. 4 the 50 %, 90 %, and 99 % quantiles of  $N_{G \rightarrow X}$  versus  $N_u$  for  $N_r = 1000$  resources and  $N_t = 20$  subframe sets, and show 95 % confidence intervals in the figure. Since  $N_{G \rightarrow X}$  is a discrete random variable, a given design constraint will produce a range of values for  $N_u$ , as shown in the figure. However, if we require that a randomly chosen UE discover all other UEs in the group within a certain number of PSDCH periods with a given probability, then we would use the largest

value of  $N_u$  that satisfies this condition. For example, requiring all UEs to be discovered within 7 PSDCH periods with a probability of 0.9 allows a maximum group size between 200 UEs and 300 UEs, using Fig. 4.

To determine the maximum group size with greater precision, we would generate a plot like Fig. 5, which shows the 90 % quantile produced by Monte Carlo simulation, with 95 % confidence intervals shown. Because of the uncertainty in these results, the estimated maximum group size is (322–340) UEs, since 322 UEs is the largest value of  $N_u$  whose confidence interval is restricted to 7 PSDCH periods, and 340 UEs is the largest value of  $N_u$  whose confidence interval includes 7 PSDCH periods. A conservative design would use the lower end of the range as the upper bound.

## V. SUMMARY AND FUTURE WORK

In this paper, we developed a Markov chain model of the discovery process at a single UE and obtained closed form expressions for the state transition probabilities. Using these in the fundamental matrix allows us to produce the distribution of the number of PSDCH periods required for a UE to discover all of the other UEs in its group, assuming all devices are half-duplex. We validated our results using two approaches: Monte Carlo simulations in Matlab and simulations of groups of UEs in NS3. We showed how to use the model to obtain the quantiles of the discovery time as a function of the group size and pool parameters, which allows one to determine the maximum number of UEs that can use a given resource pool while ensuring that the probability that all UEs are discovered within a given time is below a desired threshold. We also showed that the half-duplex effect means that performance improvements come from adding subframe sets to pools rather than PRBs; future full-duplex UEs will allow performance improvements by expanding pools in either domain.

As we noted previously, this model assumes that messages are lost only during collisions, and that collisions always produce losses. A future version of the model will incorporate the effect of path loss, fading, and shadowing, and will allow for partial recovery of collided messages in high-SIR cases.

## REFERENCES

- [1] D. Griffith and F. Lyons, "Optimizing the UE transmission probability for D2D direct discovery," in *2016 IEEE Global Telecommunications Conference (GLOBECOM 2016)*, December 2016.
- [2] M. G. Sarret, G. Berardinelli, N. H. Mahmood, B. Soret, and P. Mogenssen, "Can full duplex reduce the discovery time in D2D communication?" in *2016 International Symposium on Wireless Communication Systems (ISWCS)*, Sept 2016, pp. 27–31.
- [3] Z. Lin, L. Du, Z. Gao, L. Huang, X. Du, and M. Guizani, "Analysis of discovery and access procedure for D2D communication in 5G cellular network," in *2016 IEEE Wireless Communications and Networking Conference (WCNC)*, April 2016, pp. 1–6.
- [4] Q. Zhang and D. Liu, "On the hopping pattern design for D2D discovery," in *IEEE 80th Vehicular Technology Conference (VTC2014-Fall)*, Sept 2014, pp. 1–6.
- [5] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures," 3rd Generation Partnership Project (3GPP), TS 36.213 V12.7.0, September 2015. [Online]. Available: [http://www.3gpp.org/ftp/Specs/archive/36\\_series/36.213/36213-c70.zip](http://www.3gpp.org/ftp/Specs/archive/36_series/36.213/36213-c70.zip)



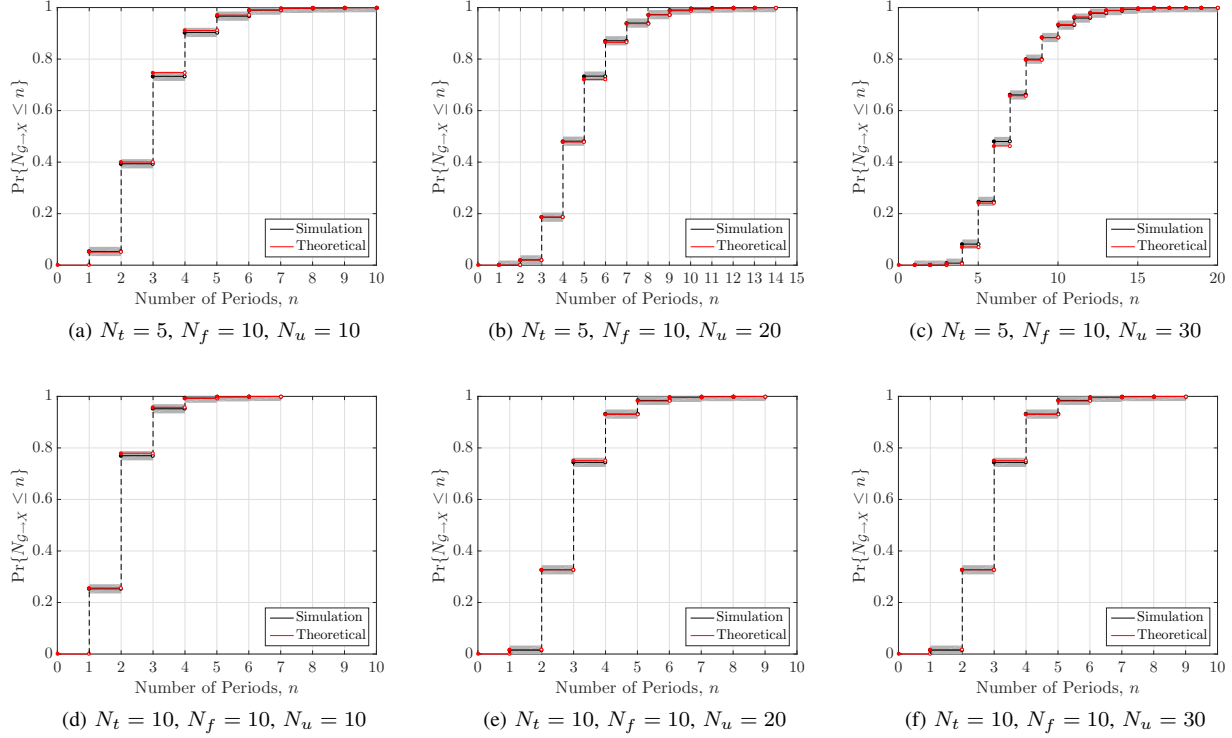


Fig. 3. CDF plots using Eq. (28) and corresponding NS3 simulation results, with envelope 95 % confidence intervals shown, plotted versus  $n$ , the number of PSDCH periods.

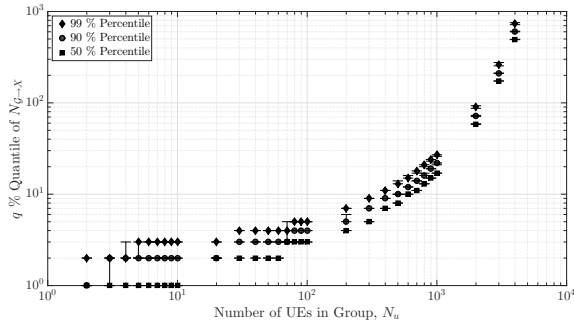


Fig. 4. Plot of the 50 %, 90 %, and 99 % quantiles of  $N_{G \rightarrow X}$  generated from Monte Carlo simulations, plotted versus UE group size,  $N_u$ . In both plots,  $N_r = 1000$  resources and  $N_t = 20$  subframe sets. Simulation results include 95 % confidence intervals.

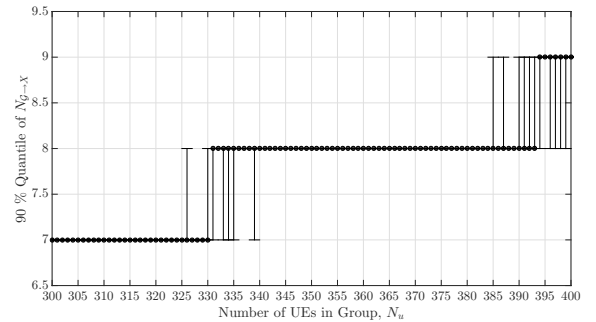


Fig. 5. Plot of the 90 %, quantile of  $N_{G \rightarrow X}$  generated from Monte Carlo simulations, plotted versus UE group size,  $N_u$ , where  $N_r = 1000$  resources and  $N_t = 20$  subframe sets. Simulation results include 95 % confidence intervals.

- [6] —, “Study on LTE Device to Device Proximity Services; Radio Aspects,” 3rd Generation Partnership Project (3GPP), TR 36.843 V12.0.1, March 2014. [Online]. Available: [http://www.3gpp.org/ftp/Specs/archive/36\\_series/36.843/36843-c01.zip](http://www.3gpp.org/ftp/Specs/archive/36_series/36.843/36843-c01.zip)
- [7] —, “Evolved Universal Terrestrial Radio Access (E-UTRA); Medium Access Control (MAC) protocol specification,” 3rd Generation Partnership Project (3GPP), TS 36.321 V12.7.0, September 2015. [Online]. Available: [http://www.3gpp.org/ftp/Specs/archive/36\\_series/36.321/36321-c70.zip](http://www.3gpp.org/ftp/Specs/archive/36_series/36.321/36321-c70.zip)
- [8] H. Kang and C. Kang, “Performance analysis of device-to-device discovery with stochastic geometry in non-homogeneous environment,” in *2014 Int. Conf. Information and Communication Technology Convergence (ICTC)*, October 2014, pp. 407–412.
- [9] H. Bagheri, P. Sartori, V. Desai, B. Classon, M. Al-Shalash, and A. Soong, “Device-to-device proximity discovery for LTE systems,” in *2015 IEEE Int. Conf. Communication Workshop (ICCW)*, June 2015, pp. 591–595.
- [10] C. M. Grinstead and J. L. Snell, *An Introduction to Probability: Second Revised Edition*. Providence, RI: American Mathematical Society, 1997.
- [11] W. Feller, *An Introduction to Probability Theory and Its Applications, Volume I*. New York: John Wiley & Sons, Inc., 1968.
- [12] G. E. Martin, *Counting: the Art of Enumerative Combinatorics*. New York: Springer-Verlag, 2001.
- [13] P. Massart, “The tight constant in the Dvoretzky-Kiefer-Wolfowitz inequality,” *The Annals of Probability*, vol. 18, no. 3, pp. 1269–1283, July 1990. [Online]. Available: <http://dx.doi.org/10.1214/aop/1176990746>