Quantifying the Impact of Guard Capacity on Session Continuity in 3GPP New Radio Systems

Vyacheslav Begishev, Dmitri Moltchanov, Eduard Sopin, Andrey Samuylov, Sergey Andreev, *Senior Member, IEEE*, Yevgeni Koucheryavy, *Senior Member, IEEE*, and Konstantin Samouylov

Abstract—Dynamic blockage of millimeter-wave (mmWave) radio propagation paths by dense moving crowds calls for advanced techniques to preserve session continuity in the emerging New Radio (NR) systems. To further improve user performance by balancing the new and ongoing session drop probabilities, we investigate the concept of guard capacity – reserving a fraction of radio resources at the NR base stations exclusively for the sessions already accepted for service. To this aim, we develop a detailed mathematical framework that takes into account the key effects in mmWave systems, including the heights of communicating entities, blocker geometry and mobility, modulation and coding schemes and antenna array geometry, as well as radio propagation and queuing specifics. Using our framework, which enables sessions to change their resource requirements during service, we demonstrate that reserving even a small fraction of bandwidth (less than 10%) exclusively for the sessions already accepted by the system allows to enhance session continuity at the expense of a slight growth in the new session drop probability as well as a small decrease in the resource utilization (approximately $5-7\%$). Furthermore, guard capacity is shown to perform better in overloaded conditions and with sessions having high data rate requirements, thus making it particularly useful for the NR systems. Our results indicate that guard capacity is a viable option for improving session continuity that can be used by the network operators in combination with other techniques, such as multi-connectivity, to maintain user experience.

I. INTRODUCTION

3GPP's New Radio (NR) technology preparing to operate in millimeter-wave (mmWave) frequency bands is expected to become an enabling solution for 5G mobile systems by providing sufficient throughputs and lower latencies for the emerging bandwidth-hungry applications. While its standardization process is almost over and various vendors are performing test trials, the focus of academic research is shifting towards improving the system-level performance indicators.

The need for reliable 5G NR operation in dense urban scenarios brings new challenges in system design. Particularly, the blockage of mmWave radio propagation by smaller objects in the channel, such as human bodies, leads to increased dynamics of the received signal strength, which can hardly be compensated even with advanced signal processing techniques.

V. Begishev, E. Sopin, A. Samuylov, and K. Samouylov are with Peoples' Friendship University of Russia (RUDN University), Moscow, Russia. Email: begishev-vo@rudn.ru, sopin-es@rudn.ru, samuylov-ak@rudn.ru, samuylov-ke@rudn.ru.

D. Moltchanov, S. Andreev, and Y. Koucheryavy are with Tampere University, Finland. Email: firstname.lastname@tuni.fi

E. Sopin and K. Samouylov are also with Institute of Informatics Problems, Federal Research Center Computer Science and Control of Russian Academy of Sciences, Moscow, Russia.

Recently, there were multiple attempts to characterize this process for dynamic moving crowds. A blockage model with stationary blockers and a mobile user equipment (UE) has been developed in [1]. A reverse scenario with mobile blockers and a static UE has been detailed in [2], where the authors revealed that blockage events occur at sub-second time scales.

These blockage events may force the NR systems to experience two principally different effects. When the distance from the UE to its serving base station (BS) exceeds a certain value, R_B, the UE suffers from outage conditions in case of blockage, since the received power of the reflected paths is below the signal-to-noise ratio (SNR) threshold. Together with receiver sensitivity and NR radio-layer structure, this distance depends on the propagation conditions, emitted power, and employed antenna arrays at the BS and UE ends. To improve session continuity in these situations, 3GPP has proposed multi-connectivity operation [3]. Accordingly, the UE is allowed to exploit spatial diversity by establishing simultaneous connections to several NR BSs in close proximity. In case of blockage on the currently active link, the UE reroutes its traffic to other unblocked link(s).

When the distance from the UE to the BS is less than R_B, the UE experiences no outage conditions. However, in order to maintain a certain target data rate, more radio resources need to be provided by the BS. In [4], the authors have recently proposed the concept of guard capacity, according to which a certain fraction of radio resources is not available to a new user session; it is being reserved for sessions already accepted by the system. Hence, a session changing its state from LoS to nLoS, and thus requiring more radio resources to maintain the same target data rate, can continue its service. The main benefit of guard capacity is that it is localized at the individual BSs and does not require dense NR deployments, therefore making this solution suitable for early NR market penetration stages. Additionally, it does not demand simultaneous support of multiple BS links by reducing the complexity of the UE implementation. Finally, it can be enabled jointly with multi-connectivity operation, allowing to improve session continuity at arbitrary UE locations.

To characterize the performance gains of guard capacity, the authors in [4] relied upon a number of simplifying assumptions. In particular, the effects of distance between the BS and the UE were disregarded; hence, modulation and coding schemes (MCSs) in blocked and non-blocked states were not distinguishable. Furthermore, at most one state transition from LoS to nLoS or back has been assumed during an active session. These assumptions may lead to underestimation of the actual gains for the proposed solution.

In this work, to rigorously investigate the effects of guard capacity, we develop a powerful mathematical framework that characterizes the performance of the NR BS. Our developed model takes into account the inherent features of mmWave propagation, specifics of antenna arrays at the UE and the BS, as well as details of the NR service process based on resource reservation.

The main contributions of our study significantly extend the past work in this field and are listed as follows.

- Reserving a small fraction of radio resources at the NR BSs (i.e., less than 10%) exclusively for sessions already accepted into the system for service allows to improve session continuity at the expense of a certain increase in the new session drop probability and only slight decrease in the system resource utilization ($5 - 7\%$).
- Although quantitative performance gains of guard capacity are highly sensitive to the system parameters and environmental conditions, this mechanism is particularly useful when the system operates under overload as well as when sessions have high data rate requirements (i.e., 10 Mbps and more), thus making it a preferred choice for NR access systems.
- The gains of guard capacity reduce when the blocker intensity or velocity increases, which implies that this mechanism might be complemented with multi-connectivity operation to further enhance session continuity.

The rest of this paper is organized as follows. In Section II, we provide a brief summary of radio resource management (RRM) in 3GPP NR systems. Further, in Section III, we introduce our system model and its main components. This system is analyzed in Section IV. We identify the operating regions and the associated performance gains of the guard capacity by conducting an extensive numerical evaluation in Section VI. Conclusions are drawn in the last section.

II. BACKGROUND AND RELATED WORK

In this section, we review the RRM essentials in 3GPP NR systems, introduce multi-connectivity and guard bandwidth concepts, as well as survey the recent RRM studies for NR.

A. Radio Resource Management in NR Systems

The NR interface is based on orthogonal frequency-division multiplexing (OFDM) techniques [5]. To support a wide range of deployment scenarios and frequency bands, NR specifies a flexible numerology with subcarrier spacing that ranges from 15 to 240 kHz with a proportional change in the cyclic prefix (CP) duration. Feasible subcarrier spacings can be $15 \times 2l$, where l is a positive/negative integer or zero. In the time domain, NR subframe length is 1 ms, which is composed of 14 OFDM symbols using 15 kHz subcarrier spacing. Since each slot carries control signals/channels in its initial OFDM symbols, NR design potentially allows for prompt resource allocation. Hybrid uplink/downlink division of OFDM symbols in a slot is also feasible. To further support small-size packet transmissions, mini-slots are adopted by 3GPP NR. A mini-slot is the minimum unit for the purposes of resource allocation

[6]. In the frequency domain, the basic scheduling unit is a physical resource block (PRB) that consists of 12 subcarriers. Since NR needs to support multiple subcarrier spacings, it takes advantage of PRBs having different bandwidth ranges.

The RRM functionality in NR systems can be broadly classified into fast and slow options [7]. Fast RRM accepts channel state measurements at the input; it operates over the set of PRBs in the frequency domain and one or several mini-slots in the time domain, while encompassing a set of functions that have tight timing constraints, i.e., per a transmission time interval (TTI). The minimum functionality required for maintaining interoperability is detailed by the 3GPP RAN2 working group in TR 38.21x specifications. Similarly to LTE, the actual scheduling is expected to be vendor-specific. Beam management and transmission scheduling are typical fast RRM functions performed by each NR BS. On the contrary, slow RRM operates over looser timing constraints and includes the necessary mechanisms, such as load balancing, call admission control, connection mobility management, and interference coordination. Similarly to fast RRM functionality, only the mandatory aspects of these schemes have been ratified by 3GPP. Both multi-connectivity and guard capacity mechanisms can be classified as slow RRM techniques.

B. Session Continuity during NR Operation

The features of mmWave propagation may lead to frequent blockage of the signal path between the NR BS and the UE [2]. Depending on the environmental characteristics, system parameters, and BS-to-UE distance, these occlusions may or may not lead to outage. Dynamic nature of objects in the mmWave channel might thus compromise session continuity by negatively affecting the service process of enhanced mobile broadband (eMBB) and ultra-reliable low latency communication (URLLC) services that are expected to be supported by 3GPP NR systems [8], [9]. To alleviate the negative effects of blockage, 3GPP has proposed multi-connectivity operation, also known as macro-diversity [3]. Accordingly, the UE is allowed to maintain multiple connections with different NR BSs in its proximity and dynamically switch between them in case of outage. Note that only a single link is permitted to be active at a time though.

The use of multi-connectivity operation places additional burden on the UEs as they have to perform standard physical and data-link layer NR functions over multiple active links simultaneously. Furthermore, practical implementation of this functionality is only feasible at more mature stages of NR market penetration. As opposed to multi-connectivity, guard capacity mechanisms aim at improving system performance when signal blockage does not result in outage conditions. They primarily target non-elastic sessions that demand constant bit rates throughout their lifetime. Guard capacity does not require dense BS deployments, which implies that it can be implemented at the early NR market penetration stages.

C. Modeling 3GPP NR with Multi-Connectivity

Understanding of mmWave system operation in static blockage environments has been attempted by a number of studies

based on stochastic geometry. The authors in [10] obtained the pdf of signal-to-interference ratio (SIR) for the mmWave systems operating at 28 GHz. The pdfs of interference and SIR in the absence of blockage have been reported in [11]. The study in [12] also neglected the phenomenon of blockage. The moments of interference and SIR for mmWave systems in the presence of blockage have been derived in [13], [14]. One of the main conclusions after this first wave of NR evaluation work is that the use of highly directional antenna radiation and sensitivity patterns produced by massive antenna arrays may indeed lead to nearly noise-limited operation [15].

Recently, research community aimed to characterize the performance benefits brought by multi-connectivity operation. Particularly, the authors in [16] relied upon simulations to demonstrate the data rate improvements where at most two active connections are supported at the UE side. An upper bound on the ergodic rate in the presence of multi-connectivity has been derived in [17] by using stochastic geometry methods. In [18], the authors developed a detailed mathematical framework for quantifying outage and spectral efficiency as functions of the number of simultaneously supported links. Among other conclusions, it was revealed that the number of simultaneous connections required to reduce the outage probability down to acceptable levels (under 1%) is at least two for moderate to high densities of blockers.

In [19], the authors proposed a model to account for the beamsearching time by characterizing the multi-connectivity gains as functions of the number of antenna elements at the BS and the UE sides. To reduce the burden imposed on the UEs, they considered the system where beamtracking is performed solely over the currently active link. These results imply that in order to fully benefit from the use of multi-connectivity the UE needs to have a rather complex implementation, being capable of maintaining multiple simultaneously active links. As one may observe, the gains are only possible in dense BS deployments, which are not expected to appear at the early NR market penetration stages.

The aforementioned studies relied exclusively upon the use of stochastic geometry to assess the performance of NR systems. The well-known feature of stochastic geometry is that it allows to obtain upper bounds on the system performance. At the same time, it completely disregards another crucial component of the packet-based wireless systems – traffic dynamics at the NR BSs. More recently, emerging studies employed various methods of queuing theory to characterize service quality of the NR systems. The performance gains of multi-connectivity operation in terms of user session continuity have been quantified in [20]. The authors considered a crowded square scenario with multiple BSs in close proximity and evaluated several session rerouting strategies. A similar study for vehicular environments has been conducted in [21].

III. SYSTEM MODEL

In this section, we specify our system model. We start by describing the deployment, then proceed with the radio part, which includes propagation, blockage, and antenna related assumptions, and finally introduce the service process in the

presence of guard capacity. Finally, we define the metrics of interest. The notation used in this paper is summarized in Table I.

TABLE I
NOTATION USED IN THIS PAPER.

Parameter	Definition
f_c	Operating frequency
d_E	Effective NR BS coverage
λ	Session arrival intensity per square meter
μ	Session service rate
h_A	NR BS height
h_U	UE height
λ_B	Blocker intensity per square meter
h_B	Height of blockers
r_B	Blocker radius
v	Blocker velocity
τ	Run length in RDM mobility model
ζ	Path loss exponent
N_0	Thermal noise
L_B	Blockage attenuation
$L(x), L_{dB}(x)$	Path loss in linear and dB scales
A_1, A_2	Propagation constants in blocked and non-blocked states
$\theta_m, \theta_{3dB}^{\pm}, \beta$	Parameters of linear antenna array
ω	Half-power beamwidth
P_A	BS transmit power
R	Session rate
π_N	New session drop probability
π_O	Ongoing session drop probability
U	Resource utilization coefficient
N_A, N_U	Number of planar antennas at BS and UE
G_A, G_U	BS transmit and UE receive antenna gains
$F_X(x), f_X(x)$	CDF and pdf of random variable X
N	Number of servers in the queuing model
γ	Guard capacity fraction
B_0	Bandwidth available to new sessions
B	Available bandwidth
$V(t)$	Current amount of bandwidth occupied at NR BS
$X(t), \check{X}(t)$	Exact and aggregated Markov queuing process
$\xi(t)$	Number of sessions in the system at time t
$\delta(t)$	Total amount of resources occupied at time t
$q_k(r)$	Probability that k sessions occupy r resources
$p_r^{(k)}$	Probability that k sessions require r resources
α	Temporal intensity of UE state changes
$\theta(j)$	Heaviside function
$\mathbf{\Phi}$	Infinitesimal generator of the Markov queuing process
$\mathbf{\Lambda_i}, \mathbf{M_i}, \mathbf{\Psi_i}$	Block matrices of infinitesimal generator
$\mathbf{1}$	Vector of ones with appropriate size
$M_{S,B}$	Shadow fading margin in blocked state
p_C	Cell-edge coverage probability
$\sigma_{S,B}, \sigma_{S,nB}$	STD of shadow fading in blocked/non-blocked states
$S_{\min}$	SNR outage threshold
S, S^{dB}	Weighted SNR in linear and dB scales
S_B, S_{nB}	SNR in blockage/non-blockage states
S_B^{dB}, S_{nB}^{dB}	SNR in blockage/non-blockage states in dB
$S_{B,S}^{dB}, S_{nB,S}^{dB}$	SNR in blockage/non-blockage with shadow fading
C_O	Control channel overhead
C_L	Cable losses
M_I	Interference margin
N_F	Noise floor
χ_1, χ_2	Shadow fading in non-blocked and blocked states
D	Distance between UE and BS
$\pi_B(x), \pi_B$	Distance-dependent and mean blockage probabilities
s_j	SNR margins of NR MCSs
m_j	Probability that UE session is assigned to MCS j
$\epsilon(x)$	Intensity of blockers crossing LoS blockage zone
M_i	Areas of zones around the LoS blockage zone
$g(x,y)$	Two-dimensional pdf of blocker locations
$\eta_i(x,y)$	Probability that blocker crosses the LoS blockage zone

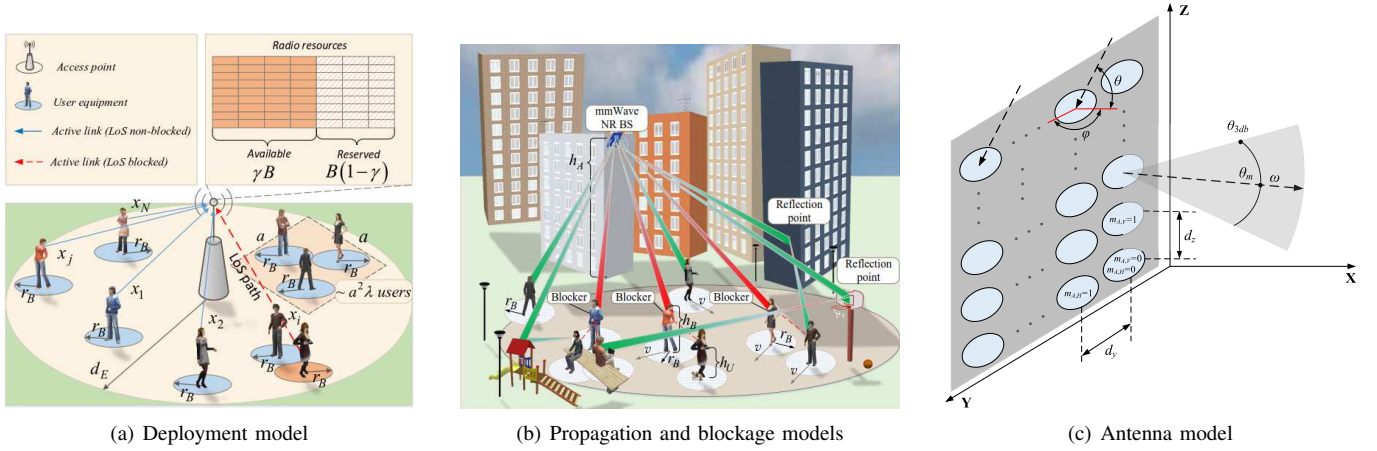(a) Deployment model   (b) Propagation and blockage models   (c) Antenna model

Fig. 1. Illustration of components of the proposed system model.

A. Deployment Model

We consider a single 3GPP NR BS at the height of h_A and a number of pedestrians around it, see Fig. 1(a). The BS has a circular coverage range of radius d_E. We assume that d_E is such that no UEs inside it experience outage conditions, i.e., there is a feasible NR MCS at the distance of d_E. We use the set of MCSs specified for 3GPP NR technology in Release 15 [5]. Note that d_E can be either estimated by using the propagation model specified below or established from the environmental characteristics, e.g., accounting for buildings that naturally limit the NR BS coverage area.

B. Blockage Model

In this paper, we consider a dynamic blockage environment, see Fig. 1(b). The number of pedestrians follows a Poisson distribution with the density of λ_B per square meter. They move within the service area according to a random direction mobility (RDM) model [22] with the speed of v m/s and an exponentially distributed run length with the mean of τ meters. The flux of pedestrians across the cell boundary is assumed to be constant. Pedestrians are modeled as cylinders having height h_B and radius r_B. The height of blockers is assumed to be h_B, $h_B \leq h_U$. In practice, h_B is set to be the average height of humans, 1.7m

C. Propagation Model

The LoS path between the UE and the NR BS might be temporarily occluded by pedestrians, see Fig. 1(b). Depending on the current link state, LoS blocked or non-blocked, and the distance between the NR BS and the UE, the session utilizes an appropriate MCS to maintain reliable data transmission.

The received signal power at the UE can be written as

$$S(x) = \frac{P_T G_A G_U}{N_0 L(x) C_O C_L M_I N_F}, \qquad (1)$$

where P_T is the transmit power, G_A and G_U are the antenna gains at the transmit and receive ends, $L(x)$ is the propagation losses at the distance of x, N_0 is the thermal noise, C_O is the control channel overhead, C_L is the cable losses, M_I is the interference margin, and N_F is the noise floor.

Following 3GPP [23], the path loss measured in dB is

$$L_{dB}(x) = \begin{cases} 32.4 + 21\log(x) + 20\log f_c, & \text{non-blocked,} \\ 47.4 + 21\log(x) + 20\log f_c, & \text{blocked,} \end{cases}$$

where f_c is the operating frequency in GHz and x is the distance between the BS and the UE.

The path loss in (2) can be represented in the linear scale using $A_i x^\zeta$, where A and ζ are propagation coefficients. Introducing coefficients (A_1, ζ_1) and (A_2, ζ_2) that correspond to non-blockage and blockage conditions, we have

$$A_1 = 10^{2\log_{10} f_c + 3.24} C_O C_L M_I N_F, \ \zeta_1 = 2.1,$$
$$A_2 = 10^{2\log_{10} f_c + 4.74} C_O C_L M_I N_F, \ \zeta_2 = 2.1. \qquad (2)$$

Let $\pi_B(x)$ be the UE blockage probability at the separation distance of x. SNR at the receiver can then be written as

$$S(x) = \frac{P_A G_A G_U}{N_0} \left[\frac{\chi_1 x^{-\zeta}[1 - \pi_B(x)]}{A_1} + \frac{\chi_2 x^{-\zeta}\pi_B(x)}{A_2} \right], \quad (3)$$

where χ_1 and χ_2 are random variables (RVs) capturing the shadow fading in non-blockage and blockage states, respectively [23].

In our work, we capture interference from the adjacent NR BSs via interference margin M_I. Recent studies of interference in mmWave systems, e.g., [14], [24], [25], indicated that the use of highly directional transmit and receive antenna radiation patterns in mmWave-based NR drastically reduces the overall interference in the network as compared to 4G microwave systems. However, even in the presence of interference at more mature stages of the NR market penetration, the core part of the said approach remains unchanged. Indeed, the interference negatively affects the received signal-to-interference-plus noise ratio (SINR), and thus the achievable channel capacity, by implying that the radius where blockage leads to outage, R_O, becomes smaller. Hence, all of the UEs that remain within the circle of radius R_O may benefit from the guard capacity provisioning, while the ones outside need to rely upon other mechanisms, such as 3GPP multi-connectivity.

D. Antenna Model

To complete the parametrization of our propagation model, one needs the antenna gains G_A and G_U. We assume linear

antenna arrays at both transmit and receive ends [26]. Half-power beamwidth (HPBW) of the array, ω, is proportional to the number of its elements, N, in the appropriate plane and can be established as shown in Fig. 1(c) by

$$\omega = 2|\theta_m - \theta_{3db}|, \qquad (4)$$

where θ_{3db} is the 3-dB point and θ_m is the location of the array maximum. The latter is computed as $\theta_m = \arccos(-\beta/\pi)$. Assuming $\beta = 0$, we have $\theta_m = \pi/2$. The upper and lower 3-dB points are thus

$$\theta_{3db}^{\pm} = \arccos[-\beta \pm 2.782/(N\pi)]. \qquad (5)$$

For $\beta = 0$, the mean antenna gain over HPBW is then [26]

$$G = \frac{1}{\theta_{3db}^+ - \theta_{3db}^-} \int_{\theta_{3db}^-}^{\theta_{3db}^+} \frac{\sin(N\pi\cos(\theta)/2)}{\sin(\pi\cos(\theta)/2)} d\theta. \qquad (6)$$

E. Traffic Model

The intensity of session arrivals is assumed to be λ sessions per square meter. The number of active UEs in the NR BS coverage area follows a Poisson distribution with the mean of $\lambda\pi r^2$. Hence, the position of each active user is uniformly distributed within the coverage area. The session duration is distributed exponentially with the parameter μ. Upon its arrival, each session is assumed to require a generally distributed amount of resources with the cumulative distribution function (CDF) of $F_R(x)$, $x > 0$, measured in Hz. The traffic is assumed to be non-elastic. We also note that the proposed methodology can be extended to multiple types of sessions having different resource requirements, $F_{R_i}(x)$.

For a particular UE, the resource requirements are not fixed for the entire duration of a session. Each active session present in the system is associated with an external Poisson process having the intensity of α, which models the occlusion of LoS by blockers. We refer to this process as to blockage process. Let $r_i(t_i)$ be the amount of resources requested from the system at time t. Once an occlusion event occurs, the resource requirements of the ongoing session change by redrawing a new sample from $F_R(x)$, $r_{i+1}(t_{i+1})$, to reflect the need for a different amount of resources to maintain the same throughput. Note that $r_{i+1}(t_{i+1})$ may or may not be greater than $r_i(t_i)$.

F. Resource Allocation with Guard Capacity

Each NR BS is assumed to have B resources, measured in Hz. We assume that only a fraction $B_0 = \gamma B$, $\gamma \in (0,1)$, of these resources is made available for new sessions. However, the entire pool of resources is accessible for the sessions already accepted by the system and changing their operating clusters. The difference $B - B_0 = B(1-\gamma)$ is termed guard capacity that is being reserved at every BS for dynamic handovers.

The system is assumed to operate as follows, see Fig. 2. Let $V(t)$, $t > 0$, be the amount of NR BS resources occupied at time t. A new session is accepted by the system if the share of remaining resources $\max(B_0 - V(t), 0)$ is sufficient to handle the incoming request drawn from the session resource request distribution, $F_R(x)$. Otherwise, a new session is considered
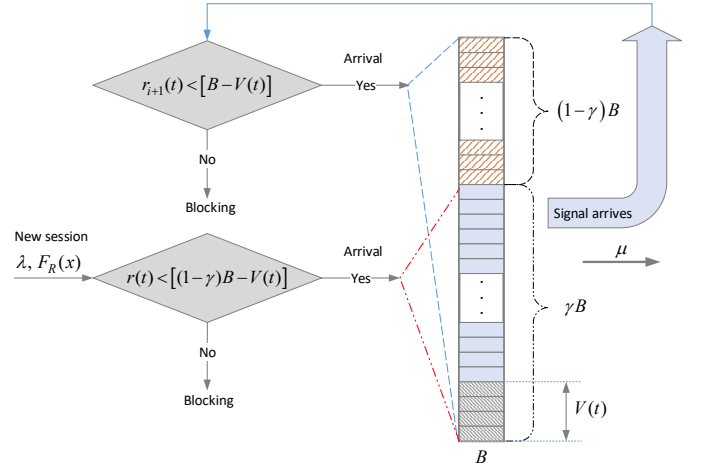

Fig. 2. Illustration of our queuing model.

dropped. Once accepted, the state of the UE this session is associated with may change multiple times from blocked to non-blocked and vice versa during the exponentially distributed service duration.

Whenever the UE state changes, the resource requirements of the session are redrawn from the distribution $F_R(x)$. This session is then considered to be "returning" and upon each return it has access to the entire pool of the available resources B. A session is considered dropped during service if the amount of the requested resources at any of its "returns" is higher than $\max(B - V(t), 0)$. Once a session completes its service or becomes dropped during service, the system reduces the total amount of the occupied resources of each type by the amount of resources allocated to this session.

G. Metrics of Interest

The main parameter of interest is the probability of dropping a session that has been accepted for service, π_O. The lower the value of π_O is, the better the session continuity in the system becomes. Another parameter we consider is the probability of not accepting a new session for service, π_N. Both probabilities are functions of the input parameters introduced above and, as one may expect, there should be a trade-off between these two metrics. We also study the effects of γ on the resource utilization coefficient, $U = \lim_{t \to \infty} V(t)/B$.

In the following sections, we develop a performance evaluation framework capable of quantifying the above metrics of interest. This framework is logically divided into two parts: (i) a queuing model specified in Section IV and (ii) a parametrization model provided in Section V. The former captures the features of radio resource allocation at NR BSs with guard capacity capabilities. At the input, it accepts the CDF of the amount of requested resources, $F_R(x)$, and the temporal intensity of the UE stage changes, α. The parametrization component estimates these two intermediate parameters depending on the scenario geometry as well as the system and environmental characteristics. This convenient separation permits to apply our core NR BS model with guard capacity for various scenarios featured by different types of geometry, dissimilar blocker types and their mobility, etc.

Particularly, for street deployments of NR BSs, one may employ a combination of techniques from [27] and [2] to account for dynamic car blockage on the lanes and human body blockage on the sidewalks, respectively. Our proposed methodology is specifically tailored to the session-level analysis of mmWave NR BSs, where one has to ensure that the traffic load does not exceed the system capacity in the long run. Local fluctuations in the user demand are assumed to be handled at lower layers by taking advantage of buffering, scheduling, and radio-level mechanisms, such as beamforming, beamsteering, and power control [28], [7], [29]. Hence, we focus on the capability of mmWave-based NR BSs with guard capacity mechanisms to handle their offered load.

IV. NR BS MODEL WITH GUARD CAPACITY

In this section, we specify our core modeling construct – the queuing system with random resource requirements and external session interruptions – that captures the dynamics of the blockage and resource reallocation processes. Then, we present a numerical solution algorithm.

A. Queuing Framework

We consider a queuing system with N servers, finite resource pool of size B, fraction of resources $B_0 = B\gamma$ available for new sessions only, and session resource requirement CDF $F_R(x)$. There are no waiting positions in the system, which implies that our model is classified as a pure loss system [30]. Full description of the system introduced in Section III requires a multi-dimensional Markov process $\tilde{X}(t) = \big(\xi(t), (\eta_0(t), \ldots, \eta_{\xi(t)}(t))\big)$, where $\xi(t)$ is the number of sessions in the system and $\eta_i(t)$ is the amount of resources occupied by the i-th session. Let k be the number of sessions in the system at the time $t_i \geq 0$ and denote by $r_1, \ldots, r_k$ the amount of resources occupied by the k-th session.

Consider the interval (t_i, t_{i+1}). First, a new session may arrive into the system with intensity λ and require j resources; if $k \leq N$ and $j \leq B_0 - (r_1 + \ldots + r_k)$ the session is accepted by the system and occupies j resources with the probability of $p_j = P(r_{k+1} = j)$; otherwise, the system drops this session. Second, the i-th session leaves the system by clearing r_i resources. Finally, each session that is currently in service can change its state from LoS to nLoS or vice versa. These state change events occur according to a Poisson process with intensity α independently for all the active sessions in the system. At each event, i-th session releases its previously occupied resources r_i and requests a new amount of resources r_i^* with the probability $p_{r_i^*}$ having the same CDF $F_R(x)$.

Define the steady-state distribution of $\tilde{X}(t)$ as

$$q_k(r_1, .., r_k) = \lim_{t\to\infty} P\{\xi(t) = k, \eta_1(t) = r_1, .., \eta_{\xi(t)}(t) = r_k\},$$
$$q_0 = \lim_{t\to\infty} P\{\xi(t) = 0\}. \tag{8}$$

For the introduced system, one may produce a set of equilibrium equations for $q_k(r_1, .., r_k)$, q_0, in the form of (7). They are obtained as follows. The term on the left-hand side (LHS) of the first equation, $\lambda q_0 \sum_{j=0}^{B_0} p_j$, is the intensity of exits to q_0, where p_j is the probability of occupying j resources

in the system. In the right-hand side (RHS), the first term $\mu \sum_{j=0}^{B_0} q_1(j)$ is the intensity of transitions from state $q_1(j)$ to state q_0 denoting the intensity of service. The second term, $\alpha \sum_{j=0}^{B} q_1(j)\big(1 - \sum_{s=0}^{B} p_s\big)$, captures transitions from state $q_1(j)$ to state q_0 induced by the UE state change, where $\big(1 - \sum_{s=0}^{B} p_s\big)$ accounts for a session drop when the new resource requirement is excessively high. In the LHS of the second equation in (7), $\big(\lambda \sum_{s=0}^{B_0-(r_1,...,r_k)} p_s + k\mu + k\alpha\big)q_k(r_1,...,r_k)$ is the transition intensity from state $q_k(r_1,...,r_k)$ to one of $q_{k-1}(r_1,...,r_{k-1})$ or $q_{k+1}(j, r_2,...,r_k), ..., q_{k+1}(r_1,...,r_k,j)$ states.

Further, in the RHS, the first term $\lambda p_{r_k} q_{k-1}(r_1,...,r_{k-1})$ is the intensity of transitions from state $q_{k-1}(r_1,...,r_{k-1})$ to state $q_k(r_1,...,r_k)$ induced by a new session arrival. The following two terms in the RHS are the transition intensities to state $q_k(r_1,...,r_k)$ associated with session completions. The fourth and the fifth terms describe the transitions to state $q_k(r_1,...,r_k)$ corresponding to the case of the UE state change and the system having sufficient amount of resources to continue its service. The last term captures a transition associated with the UE state change and subsequent session drop due to insufficient system resources. Finally, there are similar terms in the last equation. The difference is that in state $q_N(r_1,...,r_N)$ there can only be transitions associated with a successful session completion.

However, the system description in this form does not allow for a closed-form analytical solution for the steady-state probabilities. Furthermore, numerical algorithms to solve the equations in (7) along with the normalization condition for $q_k(r_1,..,r_k)$, q_0, are not directly applicable, since the number of components in the state of the system is virtually unlimited. To simplify the original system, we employ the state aggregation technique commonly accepted in queuing theory. Denoting by $\delta(t)$ the amount of resources occupied at time t, $\eta_0(t) + \ldots + \eta_{\xi(t)}$, we introduce a new Markov process $X(t) = (\xi(t), \delta(t))$ defined over the state space

$$\mathcal{X} = \bigcup_{k=1}^{N} \mathcal{X}_k, \ \mathcal{X}_k = \{(k,r) : 0 \leq r \leq B, p_r^{(k)} \geq 0\}, \tag{9}$$

where $p_r^{(k)}$ is a k-fold convolution of probabilities p_r, $r \geq 0$. Observe that $p_r^{(k)}$ is interpreted as the probability that k sessions occupy r resources.

We now arrange the states in $\mathcal{X}_k$ by increasing the order of the amount of occupied resources and denote $I(k,r)$ to be the sequence number of state (k,r) in $\mathcal{X}_k$. The steady-state probabilities of the simplified process $X(t)$ are defined as

$$q_0 = \lim_{t\to\infty} P\{\xi(t) = 0, \delta(t) = 0\},$$
$$q_k(r) = \lim_{t\to\infty} P\{\xi(t) = k, \delta(t) = r\}, (k,r) \in \mathcal{X}_k. \tag{11}$$

To complete our specification of the model, we also need to derive the distribution of the amount of resources released by a departing user. Observe that this distribution is different from the CDF of requested resources. The reason is that the former is conditioned on the event of accepting a session into

$$\lambda q_0 \sum_{j=0}^{B_0} p_j = \mu \sum_{j=0}^{B_0} q_1(j) + \alpha \sum_{j=0}^{B} q_1(j) \left(1 - \sum_{s=0}^{B} p_s\right),$$

$$\left(\lambda \sum_{s=0}^{B_0-(r_1,\dots,r_k)} p_s + k\mu + k\alpha\right) q_k(r_1,\dots,r_k) = \lambda p_{r_k} q_{k-1}(r_1,\dots,r_{k-1}) + \mu \sum_{j=0}^{B_0-(r_1,\dots,r_k)} q_{k+1}(j, r_2,\dots,r_k) + \dots$$

$$+ \mu \sum_{j=0}^{B_0-(r_1,\dots,r_k)} q_{k+1}(r_1,\dots,r_k,j) + \alpha \sum_{j=0}^{B-(r_2,\dots,r_k)} p_{r_1} q_k(j, r_2,\dots,r_k) + \dots + \alpha \sum_{j=0}^{B-(r_1,\dots,r_{k-1})} p_{r_k} q_k(r_1,\dots,r_{k-1},j) +$$

$$+ \alpha \left(1 - \sum_{s=0}^{B-(r_1,\dots,r_k)} p_s\right) \left[\sum_{j=0}^{B-(r_1,\dots,r_k)} q_{k+1}(j, r_2,\dots,r_{k+1}) + \dots + \sum_{j=0}^{B-(r_1,\dots,r_k)} q_{k+1}(r_1,\dots,r_k,j)\right],$$

$$k(\mu+\alpha)q_N(r_1,\dots,r_N) = \lambda p_{r_N} q_{N-1}(r_1,\dots,r_{N-1}) + \alpha \sum_{j=0}^{B-(r_1,\dots,r_{N-1})} p_{r_1} q_N(j, r_2,\dots,r_N) + \dots$$

$$+ \alpha \sum_{j=0}^{B-(r_1,\dots,r_{N-1})} p_{r_N} q_N(r_1,\dots,r_{N-1},j). \tag{7}$$

the system. Unconditioning by using the Bayes theorem, the probability that a customer releases j resources is given by

$$P_r(k,j) = \frac{p_r p_{j-r}^{(k-1)}}{p_j^{(k)}}. \tag{12}$$

Similarly, if the new session resource requirements do not exceed the amount of unoccupied of resources, the system changes its state from $q_k(j)$ to $q_k(r)$ with probability $\sum_{i=0}^{\min(j,r)} p_{j-i} p_i^{(k-1)} p_{r-i}/p_j^{(k)}$. Otherwise, the session is blocked and the system moves from state $q_{k+1}(r+j)$ to state $q_k(r)$ with probability $\left(1 - \sum_{s=0}^{B-r} p_s\right) p_j p_r^{(k)}/p_{j+r}^{(k+1)}$. Introducing the Heaviside function,

$$\theta(B_0 - j) = \begin{cases} 0, & j > B_0, \\ 1, & j \le B_0, \end{cases} \tag{13}$$

the system of equilibrium equations (7) characterizing the intensity of transitions of the process $\tilde{X}(t)$ assumes the form of (10), which is illustrated in Fig. 3. The system of equilibrium equations for the simplified system is given by (10), where $(N,r) \in \mathcal{X}_N$ and $1 \le k \le N-1, (k,r) \in \mathcal{X}_k$.

B. Numerical Solution

The steady-state probabilities of the simplified system in (11) have to be calculated by using the equilibrium equations
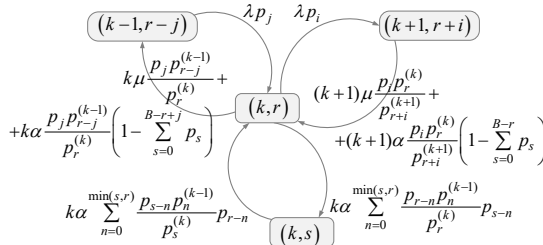

Fig. 3. Illustration of transition rates in our Markov model.

specified in (11) along with the normalization condition

$$q_0 + \sum_{k=1}^{N} \sum_{r=0}^{B} q_k(r) = 1. \tag{14}$$

Observe that the infinitesimal matrix $\mathbf{\Phi} = [a((i,j),(k,r))]$ of the Markov process $X(t)$ has a block tridiagonal structure with the main diagonal blocks $\mathbf{\Psi}_0, \mathbf{\Psi}_1, \dots, \mathbf{\Psi}_N$, upper diagonal blocks $\mathbf{\Lambda}_0, \mathbf{\Lambda}_1, \dots, \mathbf{\Lambda}_N$, and lower diagonal blocks $\mathbf{M}_0, \mathbf{M}_1, \dots, \mathbf{M}_{N-1}$ i.e.,

$$\mathbf{\Phi} = \begin{pmatrix} \mathbf{\Psi}_0 & \mathbf{\Lambda}_1 & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{M}_0 & \mathbf{\Psi}_1 & \mathbf{\Lambda}_2 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_1 & \mathbf{\Psi}_2 & \mathbf{\Lambda}_3 & \mathbf{0} & \dots \\ \dots & \mathbf{0} & \dots & \dots & \dots & \mathbf{0} \\ \mathbf{0} & \dots & \mathbf{0} & \mathbf{M}_{N-2} & \mathbf{\Psi}_{N-1} & \mathbf{\Lambda}_N \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{M}_{N-1} & \mathbf{\Psi}_N \end{pmatrix}. \tag{15}$$

The matrices corresponding to index 0 are defined as

$$\mathbf{\Psi}_0 = -\lambda \sum_{j=0}^{B_0} p_j, \tag{16}$$

$$\mathbf{\Lambda}_0 = (\lambda p_0, \dots, \lambda p_{B_0}, 0, 0, \dots, 0),$$

$$\mathbf{M}_0 = \left(\mu + \alpha \left(1 - \sum_{j=0}^{B} p_j\right), \dots, \mu + \alpha \left(1 - \sum_{j=0}^{B} p_j\right)\right)^T.$$

The elements of $\mathbf{\Psi}_n$ take the following form

$$\psi_n(I(n,i), I(n,j)) = \tag{17}$$

$$= \begin{cases} -\theta(B_0 - j)\lambda \sum_{k=0}^{B_0-i} p_k - n\mu - \\ \quad -n\alpha\left(1 - \sum_{k=0}^{i} \frac{p_{i-k}p_k^{(n-1)}}{p_i^{(n)}} p_{i-k}\right), & i = j, \\ n\alpha \sum_{k=0}^{i} \frac{p_{i-k}p_k^{(n-1)}}{p_i^{(n)}} p_{j-k}, & i < j, \\ n\alpha \sum_{k=0}^{j} \frac{p_{i-k}p_k^{(n-1)}}{p_i^{(n)}} p_{j-k}, & i > j, \end{cases}$$

where $(n,i) \in \mathcal{X}_n, (n,j) \in \mathcal{X}_n, n = 1, 2, \dots, N-1$.

$$\lambda q_0 \sum_{j=0}^{B_0} p_j = \mu \sum_{j:(1,j)\in\mathcal{X}_1} q_1(j) + \alpha \sum_{j:(1,j)\in\mathcal{X}_1} q_1(j) \left(1 - \sum_{s=0}^{B} p_s\right),$$

$$\left(\theta(B_0 - j)\lambda \sum_{j=0}^{B_0-r} p_j + k\mu + k\alpha\right) q_k(r) = \theta(R_0 - j)\lambda \sum_{j\geq 0:(k-1,r-j)\in\mathcal{X}_{k-1}} q_{k-1}(r-j)p_j + (k+1)\mu \sum_{j\geq 0:(k+1,r+j)\in\mathcal{X}_{k+1}} q_{k+1}(r+j)\frac{p_j p_r^{(k)}}{p_{j+r}^{(k+1)}} +$$

$$+ (k+1)\alpha \left(1 - \sum_{s=0}^{B-r} p_s\right) \sum_{j\geq 0:(k+1,r+j)\in\mathcal{X}_{k+1}} q_{k+1}(r+j)\frac{p_j p_r^{(k)}}{p_{j+r}^{(k+1)}} + k\alpha \sum_{j\geq 0:(k,j)\in\mathcal{X}_k} q_k(j) \sum_{i=0}^{min(j,r)} \frac{p_{j-i}\, p_i^{(k-1)}}{p_j^{(k)}} p_{r-i},$$

$$N(\mu + \alpha)q_N(r) = \theta(B_0 - j)\lambda \sum_{j\geq 0:(N-1,j)\in\mathcal{X}_{N-1}} q_{N-1}(r-j)p_j + N\alpha \sum_{j\geq 0:(N,j)\in\mathcal{X}_N} q_N(j) \sum_{i=0}^{min(j,r)} \frac{p_{j-i}\, p_i^{(N-1)}}{p_j^{(N)}} p_{r-i}. \quad (10)$$

The elements of Λ_n are

$$\lambda_n(I(n-1,i), I(n,j)) = \quad (18)$$
$$= \begin{cases} \lambda p_{j-i}, & i \leq j \leq B_0, \\ 0, & i > j \text{ or } j > B_0, \end{cases}$$

where $(n-1,i) \in \mathcal{X}_{n-1}, (n,j) \in \mathcal{X}_n, n = 1, 2, \ldots, N-1$.

The elements of $\mathbf{M}_n$ are given by

$$\mu_n(I(n+1,i), I(n,j)) = \quad (19)$$
$$= \begin{cases} (n+1)\mu \frac{p_{i-j}\, p_j^{(n)}}{p_j^{(n+1)}} + \\ +(n+1)\alpha \left(1 - \sum_{k=0}^{B-j} p_k\right)\frac{p_{i-j}\, p_j^{(n)}}{p_j^{(n+1)}}, & j \leq i, \\ 0, & j > i, \end{cases}$$

where $(n+1,i) \in \mathcal{X}_{n+1}, (n,j) \in \mathcal{X}_n, n = 1, 2, \ldots, N-1$.

The elements of $\mathbf{\Psi}_N$ can be derived as

$$\psi_N(I(N,i), I(N,j)) = \quad (20)$$
$$= \begin{cases} -N\mu - N\alpha \left(1 - \sum_{k=0}^{i} \frac{p_{i-k}\, p_k^{(N-1)}}{p_i^{(N)}} p_{i-k}\right), & i = j, \\ N\alpha \sum_{k=0}^{i} \frac{p_{i-k}\, p_k^{(N-1)}}{p_i^{(N)}} p_{j-k}, & i < j, \\ N\alpha \sum_{k=0}^{j} \frac{p_{i-k}\, p_k^{(N-1)}}{p_i^{(N)}} p_{j-k}, & i > j, \end{cases}$$

where $(N,i) \in \mathcal{X}_N, (N,j) \in \mathcal{X}_N$.

After constructing the infinitesimal generator (15), the solution of the system in question can be obtained by employing any numerical method for solving linear matrix equations in the form $\mathbf{q}^T\mathbf{\Phi} = \mathbf{0}^T$, $\mathbf{q}^T\mathbf{1} = 1$. Since (15) has a block tridiagonal form, various efficient algorithms can be applied [31]. In our calculations, we employed LU-decomposition.

C. Metrics of Interest

Once the steady-state probabilities are obtained, one can proceed with determining the metrics of interest: the new and ongoing session drop probabilities and the system resource utilization. The user-centric metrics are provided straightforwardly from the state-transition diagram illustrated in Fig. 3 as well as the associated equilibrium equations in (10). Particularly, the new session drop probability is

$$\pi_N = 1 - \sum_{n=0}^{N-1} \sum_{r=0}^{B_0} q_n(r) \sum_{j=0}^{B_0-r} p_j. \quad (21)$$

To derive the ongoing session drop probability, we introduce the following notation. Let T be a certain time interval of finite duration and let $\lambda(1 - \pi_N)$ be the effective arrival rate of user sessions i.e., $\lambda(1 - \pi_N)T$ is the mean number of sessions accepted during time T. Observe that the mean number of sessions in the system is given by

$$E[N] = \sum_{n=0}^{N} \sum_{r=0}^{B} n q_n(r). \quad (22)$$

Letting $\alpha E[N]$ be the mean intensity of UE state changes, the intensity of session drops caused by the UE state changes is provided by

$$\nu = \alpha E[N] \sum_{n=1}^{N} \sum_{r=0}^{B} q_n(r) \sum_{j=0}^{r} \frac{p_j p_{r-j}^{(n-1)}}{p_r^{(n)}} \left[1 - \sum_{i=0}^{B-r+j} p_i\right]. \quad (23)$$

Using (23), we establish the ongoing session drop probability as

$$\pi_O = \lim_{T\to\infty} \frac{\nu T}{\lambda(1 - \pi_N)T} = \frac{\nu}{\lambda(1 - \pi_N)}. \quad (24)$$

Finally, the system resource utilization is obtained by summing up all of the states of the system associated with r occupied resources i.e.,

$$U = \sum_{n=1}^{N} \sum_{r=0}^{B} r q_n(r). \quad (25)$$

V. MODEL PARAMETRIZATION

To parametrize our model, one needs to provide (i) effective coverage of the NR BS for a given emitted power as well as transmit and receive antenna configurations, d_E, (ii) CDF of the amount of resources requested from the BS by a new session, $F_R(x)$, and (iii) intensity of the UE state changes, α. In this section, we derive these quantities as functions of the input system parameters.

A. Effective NR BS Coverage

To ensure full coverage by the NR BSs, d_E should be such that no outage occurs at the UE that currently resides in blockage conditions and is located at the distance of d_E from the BS. Let $S_{\min}$ be the SNR outage threshold i.e., $S_{\min}$ is a lower bound of the SNR range corresponding to the lowest

MCS. Using a part of (3) corresponding to the LoS blockage state, we have the following relation

$$S_{\min} = \frac{P_A G_A G_U}{N_0 A_2}(d_E + [h_A - h_U]^2)^{-\zeta/2}, \quad (26)$$

where ζ is the path loss exponent, h_A and h_U are the heights of BS and UE, P_A is the BS transmit power, G_A and G_U are the BS transmit and the UE receive antenna gains, N_0 is the thermal noise, and A_2 is the propagation constant.

Solving (26) with respect to d_E, we obtain

$$d_E = \sqrt{\left(\frac{P_A G_A G_U}{N_0 A_2 S_{\min} M_{S,B}}\right)^{\zeta/2} + (h_A - h_U)^2}, \quad (27)$$

where $M_{S,B}$ is the shadow fading margin in the blocked state, which is computed as follows

$$M_{S,B} = \sqrt{2}\sigma_{S,B}\text{erfc}^{-1}(2p_C), \quad (28)$$

where $\text{erfc}^{-1}(\cdot)$ is the inverse complementary error function, p_C is the cell-edge coverage probability, and $\sigma_{S,B}$ is the standard deviation (STD) of the shadow fading distribution for the LoS blocked state, which is provided in [23].

B. Characterizing Resource Requirements

After obtaining d_E, we can determine the CDFs of the resource requirements. Let S_{nB} be a RV denoting the SNR in non-blocked conditions and $F_{S_{nB}}(x), x > 0$, be its CDF. Recall that the locations of new session arrivals are distributed uniformly within the BS coverage. Hence, the CDF of the distance between UE and BS can be written as

$$F_D(y) = (y^2 - (h_A - h_U)^2)/d_E^2, \quad (29)$$

which is defined over $|h_A - h_U| < y < \sqrt{d_E^2 + (h_A - h_U)^2}$.

Since the propagation model is monotonously decreasing in y, the distribution of SNR can be expressed in terms of the distribution of distance D i.e.,

$$F_{S_{nB}}(y) = 1 - F_D(P_A G_A G_U/N_0 A_1 y^{\zeta/2}). \quad (30)$$

The CDF of RV S_B denoting the SNR in the blocked state is obtained similarly. To determine the overall SNR CDF, we also require the blockage probability. Following [32], the blockage probability at the distance of x is given by

$$\pi_B(x) = 1 - e^{-2\lambda_B r_B(x\frac{h_B - h_U}{h_A - h_U} + r_B)}, \quad (31)$$

which leads to the following weighted expression

$$\pi_B = \int_0^{d_E} p_B(x)\frac{2x}{d_E^2}dx. \quad (32)$$

Further, we capture the effect of shadow fading. We define the CDF of SNR measured in dB as $F_{S_{nB}}^{dB}(y) = F_{S_{nB}}(10^{y/10})$. Recalling that shadow fading is characterized by a lognormal distribution leading to a normal distribution in dB, the RV specifying the SNR distribution is expressed as

$$S_{nB,S}^{dB} = S_{nB}^{dB} + \mathcal{N}(0, \sigma_{S,nB}), \quad (33)$$

where $\mathcal{N}(0, \sigma_{S,nB})$ is a zero-mean normal distribution with the STD of $\sigma_{S,nB}$ that characterizes the shadow fading.
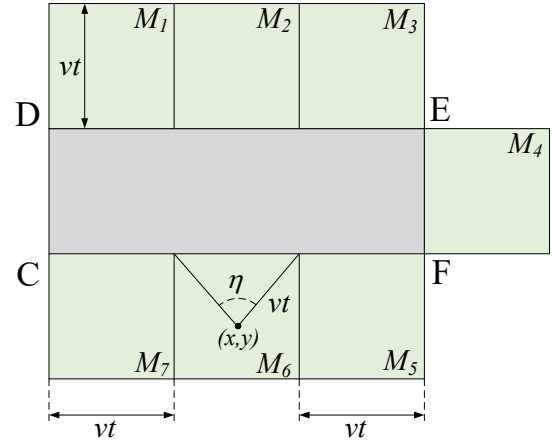

Fig. 4. Translation of spatial blocker intensity into temporal domain.

The SNR CDF capturing both distance-dependent path loss and shadow fading can now be determined as a convolution of $F_{S_{nB}}^{dB}(y)$ and $\mathcal{N}(0, \sigma_S)$, which yields

$$F_{S_{nB,S}}^{dB}(y) = \int_{-\infty}^{\infty} F_{S_{nB}}^{dB}(y+u)\frac{e^{-u^2/2\sigma_S^2}}{\sqrt{2\pi}\sigma_S}du. \quad (34)$$

Unfortunately, (34) cannot be made available in a closed-form via standard RV transformation techniques [33], but it can be represented in terms of an error function, $\text{erf}(\cdot)$, as in (36)

$$A = \frac{P_A 10^{G_A G_U/10}}{f_c^2 10^{3.24 + L_B/10} 10^{\frac{1}{10}(C_O + C_L + M_I + N_F)} N_0}, \quad (35)$$

where C_O is the control channel overhead, C_L is the cable losses, M_I is the interference margin, and N_F is the noise floor.

The SNR CDF $F_{S^{dB}}(y)$ can now be determined by weighing the CDFs corresponding to blocked and non-blocked states with the probabilities π_B and $(1 - \pi_B)$. Let us further denote s_j, $j = 1, 2, .., K$, as the SNR margins of the NR MCS schemes, where K is the MCS number. Further, let m_i be the probability that the UE session is assigned to MCS i. Therefore, we have

$$m_i = Pr\{s_i < s < s_{i+1}\} = F_{S^{dB}}(s_{i+1}) - F_{S^{dB}}(s_i). \quad (37)$$

Once m_i, $i = 1, 2, \ldots, K$, are available, the amount of resources required for a session with data rate R can be derived.

C. Intensity of UE State Changes

To obtain the intensity of UE state changes, one needs to characterize the temporal dynamics of the blockage process. First, we determine the intensity of blockers, $\epsilon(x)$, entering the blockage zone associated with the user located at the distance of x. We specify the area around the blockage zone, as shown in Fig. 4, where moving blockers may cross the blockage zone by occluding the LoS between the UE and the NR BS. We further divide the area around the LoS blockage zone into i, $i = 1, 2, ...7$, zones as demonstrated in Fig. 4. The intensity

$$F_{S_{nB}^{dB}}(y) = \frac{1}{2d_E^2}\left[A^{2/\zeta}10^{-\frac{y}{5\zeta}}e^{\frac{\sigma_{S,nB}^2\log^2(10)}{50\zeta^2}}\left[\mathrm{erf}\left(\frac{50\zeta\log(A) - 25\zeta^2\log\left(d_E^2 + (h_A - h_U)^2\right) + \sigma_{S,nB}^2\log^2(10) - 5\zeta y\log(10)}{5\sqrt{2}\zeta\sigma_{S,nB}\log(10)}\right) - \right.\right.$$

$$\left. - \mathrm{erf}\left(\frac{50\zeta(\log(A) - \zeta\log(h_A - h_U)) + \sigma_{S,nB}^2\log^2(10) - 5\zeta y\log(10)}{5\sqrt{2}\zeta\sigma_{S,nB}\log(10)}\right)\right] + \left(d_E^2 + (h_A - h_U)^2\right) \times$$

$$\times\,\mathrm{erf}\left(\frac{-10\log(A) + 5\zeta\log\left(d_E^2 + (h_A - h_U)^2\right) + y\log(10)}{\sqrt{2}\sigma_{S,nB}\log(10)}\right) - (h_A - h_U)^2 \times$$

$$\left. \times\,\mathrm{erf}\left(\frac{\sqrt{2}(-10\log(A) + 10\zeta\log(h_A - h_U) + y\log(10))}{\sigma_{S,nB}\log(100)}\right) + d_E^2\right]. \tag{36}$$

VI. PERFORMANCE ANALYSIS

of blockers crossing the blockage zone of the UE located at the distance of x from the NR BS can be approximated as

$$\epsilon(x) = \sum_{i=1}^{7}\iint_{M_i} g_i(x,y)Pr\{E\}Pr\{T > 1\}\lambda_B M_i dx dy, \tag{38}$$

where M_i is the area of zone i, $g_i(x,y)$ is the pdf of the locations of blockers in zone i, $g_i(x,y) = 1/M_i$, and $Pr\{T > 1\} = \exp(-1/\tau)$ is the probability that a blocker moves longer than a unit of time without changing its direction.

The probability that a blocker moves towards the LoS blockage zone is $Pr\{E\} = \eta_i(x,y)/2\pi$, where $\eta_i(x,y)$ is the range of movement angles in zone i that leads to crossing the LoS blockage zone. We may now simplify (38) as follows

$$\epsilon(x) = \frac{\lambda e^{-1/\tau}}{2\pi}\sum_{i=1}^{7}\iint_{M_i}\eta_i(x,y)\,dx\,dy, \tag{39}$$

where $\eta_i(x,y)$ is calculated as

$$\eta_i(x,y) = ([x_D - x]/vt),\ i = 1, 3, 5, 7,$$
$$\eta_i(x,y) = 2\cos^{-1}([x_E - x]/vt),\ i = 2, 6,$$
$$\eta_4(x,y) = 2\tan^{-1}([x - x_E]/[y - y_E]), \tag{40}$$

where $x_\cdot, y_\cdot$ are the coordinates of points in Fig. 4.

The average temporal intensity of blockers entering the LoS blockage zone within the coverage area of the NR BS is then

$$\epsilon = \int_0^{d_E}\epsilon(x)\frac{2x}{d_E^2}dx. \tag{41}$$

It has been demonstrated in [34] that the process of meetings between a stationary convex zone and a point moving inside a bounded area according to the RDM is approximately Poisson. Furthermore, the authors in [2] showed that the blockage process is of alternating renewal nature where non-blocked intervals follow an exponential distribution with the parameter ϵ. The distribution of the blocked intervals coincides with the busy period in the $M/GI/\infty$ queuing system that can be calculated numerically for any pdf of the LoS blockage zone residence time of a single blocker [35]. The latter can be found similarly to [2] by observing that due to the properties of the RDM the entry point of a blocker is distributed uniformly over the three sides of the blockage zone. Finally, once the mean blockage and non-blockage intervals are established, one can estimate the intensity of the UE state changes α as the mean of blockage and non-blockage interval intensities.

In this section, we numerically evaluate the performance of the guard capacity concept in 3GPP NR systems. As our performance indicators, we track both user-centric metrics, including new and ongoing session drop probabilities, as well as system-centric metrics that characterize system resource utilization. We first study the response of the system to the input parameters and then proceed with identifying an upper bound on the BS deployment.

The default system parameters are summarized in Table II. We specifically note that the offered traffic load in our system depends on the coverage of a single BS antenna. The latter is a function of the antenna gains and thus the number of antenna elements used at the BS and the UE sides. Hence, in the rest of this section, we employ the arrival intensity of sessions per square meter to parametrize our model. We also

TABLE II
DEFAULT PARAMETERS FOR NUMERICAL ASSESSMENT

Parameter	Value
Carrier frequency, f_c	28 GHz
Bandwidth, B	400 MHz
Height of BS, h_A	10 m
Height of blocker, h_B	1.7 m
Height of UE, h_U	1.5 m
Blocker radius, r_B	0.4 m
Blocker velocity, v	2 m/s
Mean blocker movement duration, $1/\tau$	5 m
SNR blockage threshold, S_B	-9.47 dB
Transmit power, P_T	2 W
Path loss exponent, ζ	2.1
Target outage probability, p_C	0.05
STD of shadow fading, $\sigma_{S,B}, \sigma_{S,nB}$	8.2/4 dB
Blocker intensity, λ_B	0.5 units/m^2
Thermal noise, N_0	-174 dBm/Hz
Blockage attenuation, L_B	20 dB
UE antenna array, $N_U \times 1$	4×1 elements
UE receive gain, G_R	5.57 dBi
Control channel overhead, C_O	1 dB
Cable losses, C_L	2 dB
Interference margin, M_I	3 dB
Noise floor, N_F	7 dB
Considered session data rates, R	{10,20,30} Mbps
Mean session service time, $1/\mu$	30 s
Arrival intensity of sessions, λ	{0.0001,0.01} sess./s
BS planar antenna elements	1, 4, 8, 16, 32, 64, 128, el.
BS vertical antenna elements	4 el.
UE antenna array	4×4 el.
BS inter-site distance, ISD	$3d_E$ m
Guard capacity fraction, γ	$(0.8\dots,1)$

TABLE III
SYSTEM PARAMETERS INDUCED BY BS ANTENNA ARRAYS

Array	Gain, dBi	HPBW, °	d_E, m	10Mbps	20Mbps	30Mbps
128x4	20.58	0.79	686	9.18	18.14	26.82
64x4	17.59	1.59	494	9.19	18.14	26.82
32x4	14.58	3.18	355	9.19	18.15	26.83
16x4	11.57	6.37	255	9.20	18.16	26.84
8x4	8.57	12.75	183	9.20	18.17	26.87
4x4	5.57	25.50	132	9.21	18.20	26.91
1x4	2.58	102.00	95	9.22	18.26	27.00

note that the antenna array at the UE side is assumed to be 4×4 elements throughout this assessment. The system parameters that depend on the number of antenna elements at the BS side are shown in Table III. The 3GPP NR MCSs have been utilized to estimate these parameters [5].

A. Model Validation

We start by validating the proposed analytical framework. To achieve this, we developed a single-purpose simulation environment that accepts the input parameters in Table II, as well as blocker mobility, propagation, and service sub-models, and returns both user- and system-centric metrics of interest. To construct our simulator, we relied upon a discrete-event modeling framework (DES [36]). The beginning of the steady-state period is determined by using an exponentially-weighted moving average test with a smoothing constant of 0.05 [37]. The statistics were collected during the steady-state period by using the method of batch means [38], by sampling the state of the system each 10 seconds of the simulation time.

A comparison of new and ongoing session drop probabilities obtained with the developed mathematical model and computer simulations is shown in Fig. 5 for $\lambda = 1.15E-7$ sessions per square meter, the intensity of blockers equal to $\lambda_B = 0.04$, and the mean session data rate of $E[R] = 10$ Mbps. Since the confidence intervals were always less that 0.01 of the respective absolute values for the level of significance set to 0.95, we only demonstrate point estimates. As one may observe, the simulation data agrees with the analytical results tightly. Similar outcomes have been noted for other input parameters (also in case of the system-centric resource utilization metric). Hence, in what follows, we rely upon our developed analytical model to deliver the target assessment of the guard capacity effects.
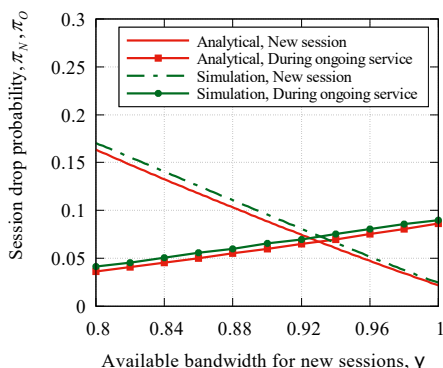

Fig. 5. Comparison of analytical and simulation results.

B. User Performance Assessment

First, we characterize the guard capacity gains for a wide range of input parameters by focusing on the user-centric metrics, including new and ongoing session drop probabilities. Throughout this evaluation, to compare the system response in fair conditions, we assume 32×4 antenna array at the BS having the transmit power of 2 W, which corresponds to the effective coverage of 355 m.

The new and ongoing session drop probabilities as functions of the fraction of bandwidth available for the new sessions are demonstrated in Fig. 6 for the session data rate of 10 Mbps and the blocker intensity of 0.04 blockers per square meter. Analyzing the impact of guard capacity on new and ongoing session drop probabilities as illustrated in Fig. 6(a), one may observe that its use allows to improve session continuity by decreasing the probability that an accepted session is dropped during service. As expected, there is a trade-off between the ongoing and the new session drop probabilities, and the benefits are evident for all the considered session arrival intensities. Importantly, the positive effect of guard capacity for the ongoing session drop probability increases as the arrival intensity of sessions grows; that is, it allows to achieve higher gains in overloaded conditions, which is of critical importance for prospective NR network operators.

Further, consider the effect of session data rate illustrated in Fig. 6(b) for the arrival traffic intensity of 1.1E-7 sessions per square meter and the blocker intensity of 0.04 blockers per square meter. We emphasize several important effects. Firstly, increasing the rate of sessions affects both user-centric metrics negatively. The underlying reason is that higher session rates leave unused more resources in the system when either new or ongoing session is dropped due to insufficient resources available. Secondly and more importantly, the benefits of guard capacity increase in response to growing session rates. Particularly, for the session data rate of 10 Mbps the ongoing session drop probability decreases from approximately 0.42 for $\gamma = 1$ down to around 0.33 for $\gamma = 0.8$. For the same values of γ and $R = 10$ Mbps, the guard capacity gains are marginal. Hence, we may conclude that guard capacity is particularly useful for heavy sessions having extreme data rate requirements. Handling such sessions is a primary target in the future NR deployments.

Blockage phenomenon is known to severely affect the performance of mmWave-based NR systems. Fig. 6(c) demonstrates the effect of the intensity of blockers (human crowd) on new and ongoing session drop probabilities for the mean session data rate of 10 Mbps and the arrival intensity of 1.1E-7 sessions per square meter. First, one may notice a large gap between the ongoing session drop probabilities corresponding to different blocker intensities. In fact, the said probability increases from below 0.03 to over 0.8. The associated difference in the new session drop probabilities is much smaller and does not exceed 0.01 across the entire considered range of γ. The reason for this behavior is that higher values of blocker intensities lead to higher intensity of LoS/nLoS state changes. Therefore, the number of state changes per session service time grows, which leads to higher chances of dropping
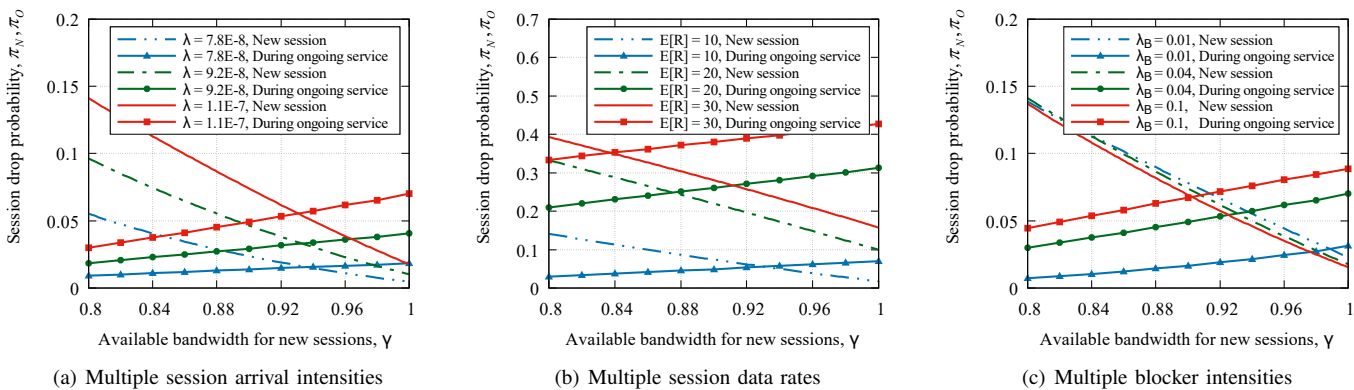
Fig. 6. New and ongoing session drop probabilities for selected system parameters.

an ongoing session. We also note that a qualitative response of user-centric performance metrics remains similar for all the considered values of blocker intensity.

We also emphasize that the impact of blocker velocity has a similar nature to that of the blocker density. In fact, higher velocity leads to higher frequency of state changes between LoS and nLoS, which negatively impacts the ongoing session drop probability. Furthermore, the effect of the BS, UE, and blocker heights is also straightforward, and we do not discuss it in detail here. Particularly, larger BS or blocker heights as well as smaller UE heights increase the perimeter of the LoS blockage zone, thus leading to more frequent LoS/nLoS state changes and eventually resulting in higher ongoing session drop probabilities.

After identifying that guard capacity is particularly useful under high system loads, we take a closer look at the impact of traffic arrival intensity displayed in Fig. 7. Naturally, both new and ongoing session drop probabilities increase as λ grows. For a relatively small arrival intensity, one may observe that the ongoing session drop probabilities almost coincide for different values of γ. However, as the arrival intensity grows, one may notice variation in this parameter. Qualitatively, similar behavior is observed for the new session drop probabilities. However, the gap between curves corresponding to different values of γ is generally larger.

C. System Performance Assessment

By definition, guard capacity targets to exploit the trade-off between the user-centric metrics, particularly, the new and
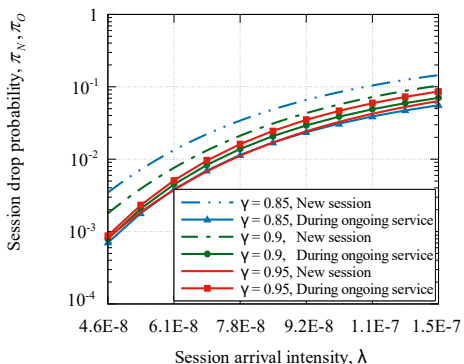

Fig. 7. Session drop probability as a function of λ.

ongoing session drop probabilities. In general, reserving a fraction of bandwidth for the sessions already accepted by the system may compromise the resource utilization. We now proceed with studying the latter.

System resource utilization as a function of session arrival intensity, λ, is illustrated in Fig. 8 for the mean session data rate of 10 Mbps and the blocker intensity of 0.04 blockers per square meter. Clearly, as λ grows, the system resource utilization increases across the entire considered range of the guard capacity values. More importantly, although resource utilization indeed decreases as the value of guard capacity $(1 - \gamma)$ grows, the degradation is not drastic. Particularly, for lower values of λ e.g., 1.5E-8, the resource utilization remains almost unchanged at approximately 0.3 for $\gamma \in (0.8, 1)$. For higher values of session arrival intensity, e.g., 1.1E-7, where guard capacity demonstrates the best user-side gains, see Fig. 6, the degradation is from approximately 0.81 for $\gamma = 1$ down to around 0.73 for $\gamma = 0.8$. This small drop in system-centric performance is explained by that the ongoing sessions occupy the reserved resources as a result of LoS/nLoS state changes. Hence, the use of guard capacity offers user-centric performance improvements at the expense of a slight decrease in the resource utilization of the system.

The impact of the mean session data rate on the system resource utilization is shown in Fig. 8(b) for the two arrival intensities, three mean session rates, and the blocker intensity of 0.04 blockers per square meter. As one may observe, a decline in the system resource utilization is still negligible for all the considered values of the mean session rates. Further, it is important to note that for higher session data rates the system resource utilization increases. The reason behind this behavior is twofold. First, the improvement is partially due to higher offered traffic load, $\rho = E[R]\lambda/\mu$, which increases when either session arrival intensity or mean session rate grows. Another reason is that the sessions with higher mean data rate requirements are characterized by higher resource demands in the nLoS state. Based on this analysis, we may conclude that guard capacity is well-suited for the NR systems that are primarily targeted to accommodate throughput-hungry applications.

Finally, consider the effect of blocker intensity on the system resource utilization as demonstrated in Fig. 8(c) for the mean session data rate of 10 Mbps and the arrival intensity of
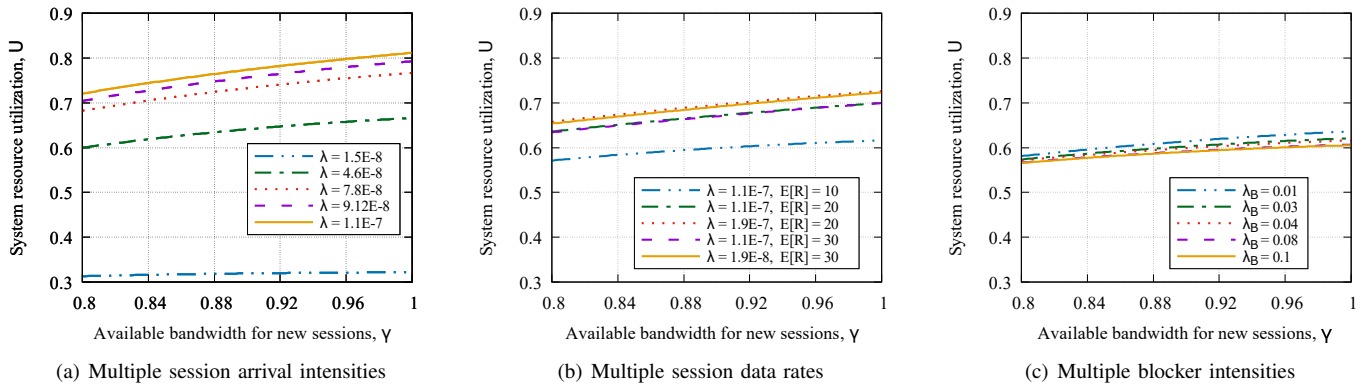
Fig. 8. System resource utilization as function of system parameters.

1.1E-7 sessions per square meter. We observe that an increase in λ_B leads to slight degradation of the system resource utilization across the entire range of guard capacity values. The underlying reason is that the intensity of LoS/nLoS state changes leads to multiple ongoing session drops as one may infer from Fig. 6(c). However, this impact is almost negligible in its absolute values for the considered system parameters. We note that it heavily depends on the relation between the mean session service time and the intensity of UE state changes.

D. BS Deployment Density

Previously, we considered the case of 32×4 antenna array at the BS side. Results for other types of arrays are qualitatively similar. However, as one may recall, an increase in the number

(a) New session drop probability

(b) Ongoing session drop probability

Fig. 9. New and ongoing session drop probabilities.

of transmit antennas at the BS leads to higher transmit gains and thus potentially larger effective cell radius, d_E. The growth of the effective radius in its turn affects the offered traffic load at a single BS. Therefore, for a given session arrival intensity (per square meter) the use of a larger number of antenna elements may threaten the QoS guarantees in terms of new and ongoing session drop probabilities. In this subsection, we demonstrate how our developed framework can be utilized to determine the optimal number of antenna elements and thus the preferred density of BSs for a given session arrival intensity.

To this aim, Fig. 9 displays new and ongoing session drop probabilities as functions of the number of antenna elements for the session arrival intensity of 1.1E-7 sessions per square meter, the mean session data rate of 10 Mbps, and the blocker intensity of 0.04 blockers per square meter. Here, an increase in the number of antenna elements forming a planar antenna radiation pattern leads to a linear decrease in the new session drop probability and a linear increase in the ongoing session drop probability for all the considered values of γ. This is explained by the fact that for the fixed emitted power the area served by a single BS antenna array configuration grows. Recall that the ongoing session drop probability is directly impacted by the effective coverage radius of the NR BS. Indeed, as d_E increases, the perimeter of the LoS zone grows as well, which leads to higher intensity of LoS/nLoS state changes, and thus to higher ongoing session drop probability. As a result, more resources are made available for the new sessions by decreasing the new session drop probability.

Analyzing the data presented in Fig 9 further, we note that the use of guard capacity brings another system trade-off. Particularly, the positive effect of guard capacity on the ongoing session drop probability diminishes as the number of antenna elements forming a directional radiation pattern grows. This is compensated by the associated decrease in the new session drop probability. However, in absolute values, the gain in the new session drop probability is much milder as compared to the degradation in the ongoing session drop probability.

Finally, consider the impact of the number of antenna elements on the system resource utilization as illustrated in Fig. 10. For growing session arrival intensity, the resource utilization slightly decreases for all the considered numbers
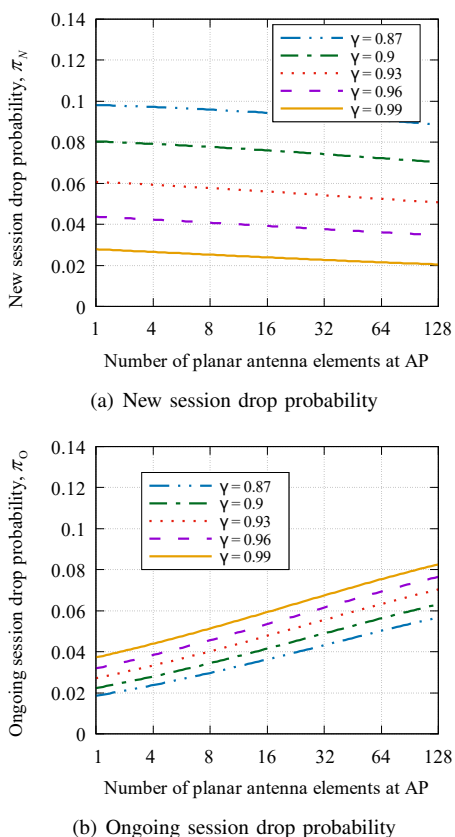
(a) As function of λ
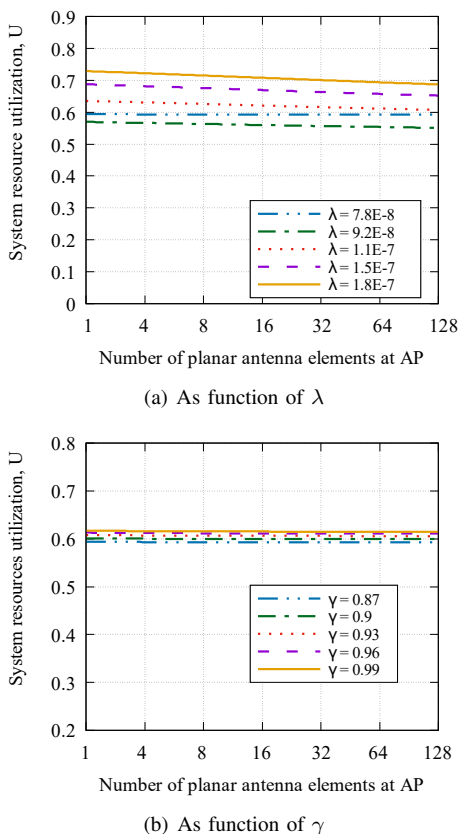
(b) As function of γ

Fig. 10. System resource utilization as function of antenna elements.

continuity and the new session drop probability at the expense of a slight performance degradation in the overall resource utilization. Importantly, the system demonstrates its best performance in overloaded conditions. Furthermore, for any given guard capacity value, the ongoing session drop probability decreases as the mean session data rate grows. This particularly suits 3GPP NR systems that primarily target bandwidth-hungry applications. In practice, reserving a small amount of radio resources exclusively for the sessions accepted by the system may significantly benefit session continuity.

We emphasize that the new and ongoing session drop probabilities are extremely sensitive to the offered load levels. Therefore, to implement the considered guard capacity capabilities under practical spatially- and temporally-varying traffic conditions, the service operator needs to monitor the current loading and adjust the guard capacity parameters accordingly.

Multi-connectivity operation is another mechanism aimed at improving session continuity, which can complement guard capacity should it fail to alleviate the effects of blockage. Particularly, for higher values of blocker intensity and/or their velocity, both user- and system-centric parameters degrade substantially. However, we expect that a combination of guard capacity and multi-connectivity operation will work efficiently in practical NR deployments by accommodating both eMBB and URLLC types of services. Here, multi-connectivity may be utilized to combat outages, while guard capacity can provide the means for supporting the desired session data rates.

of antenna elements as shown in Fig. 10(a) for the mean session data rate of 10 Mbps and the blocker intensity of 0.04 blockers per square meter. It is important that the difference between the curves remains almost intact (the gap between the session arrival rates of 1.8E-7 and $7.8E-7$ is approximately 0.1) for all the considered values of the numbers of antenna elements. At the same time, guard capacity affects the resource utilization insignificantly for all the considered numbers of antenna elements, see Fig. 10(b), since the main effect here is produced by increasing the session arrival intensity.

VII. CONCLUSIONS

In 3GPP NR systems, frequent state changes from LoS to nLoS propagation conditions at the UE side may lead to drops of data sessions that are already accepted for service. To improve session continuity, the concept of guard capacity has been outlined, which is essentially reserving a fraction of radio resources exclusively for the ongoing sessions. To understand the performance trade-offs between the new and ongoing session drop probabilities as well as characterize the effects of guard capacity on the system resource utilization, we developed a detailed mathematical model of the service process at the NR BS. It takes into account the key features of mmWave-based NR technology, including the heights of communicating entities, blocker geometry and mobility, BS MCSs and antenna arrays, as well as LoS and nLoS propagation conditions.

Our numerical results demonstrate that the use of guard capacity offers significant flexibility to balance the session

REFERENCES

[1] A. Samuylov, M. Gapeyenko, D. Moltchanov, M. Gerasimenko, S. Singh, N. Himayat, S. Andreev, and Y. Koucheryavy, "Characterizing spatial correlation of blockage statistics in urban mmWave systems," in *Globecom Workshops (GC Wkshps), 2016 IEEE*. IEEE, 2016, pp. 1–7.
[2] M. Gapeyenko *et al.*, "On the temporal effects of mobile blockers in urban millimeter-wave cellular scenarios," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 11, pp. 10 124–10 138, November 2017.
[3] 3GPP, "NR; Multi-connectivity; Overall description; Stage-2 (Release 15)," 3GPP TS 37.340, December 2017.
[4] D. Moltchanov, A. Samuylov, V. Petrov, M. Gapeyenko, N. Himayat, S. Andreev, and Y. Koucheryavy, "Improving session continuity with bandwidth reservation in mmWave communications," *IEEE Wireless Communications Letters*, vol. 8, no. 1, pp. 105–108, 2019.
[5] 3GPP, "NR; Physical channels and modulation (Release 15)," 3GPP TR 38.211, Dec 2017.
[6] ——, "NR; Multiplexing and channel coding (Release 15)," 3GPP TS 38.212, Dec 2017.
[7] Y. Li, E. Pateromichelakis, N. Vucic, J. Luo, W. Xu, and G. Caire, "Radio resource management considerations for 5G millimeter wave backhaul and access networks," *IEEE Communications Magazine*, vol. 55, no. 6, pp. 86–92, 2017.
[8] C. She, C. Yang, and T. Q. Quek, "Radio resource management for ultra-reliable and low-latency communications," *IEEE Communications Magazine*, vol. 55, no. 6, pp. 72–78, 2017.
[9] C. She, Z. Chen, C. Yang, T. Q. Quek, Y. Li, and B. Vucetic, "Improving network availability of ultra-reliable and low-latency communications with multi-connectivity," *IEEE Transactions on Communications*, vol. 66, no. 11, pp. 5482–5496, 2018.
[10] K. Venugopal and R. Heath, "Millimeter wave networked wearables in dense indoor environments," *IEEE Access*, vol. 4, pp. 1205–1221, Apr. 2016.
[11] M. Di Renzo, "Stochastic geometry modeling and analysis of multi-tier millimeter wave cellular networks," *IEEE Transactions on Wireless Communications*, vol. 14, no. 9, pp. 5038–5057, 2015.

[12] S. Singh, M. N. Kulkarni, A. Ghosh, and J. G. Andrews, "Tractable model for rate in self-backhauled millimeter wave cellular networks," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 10, pp. 2196–2211, 2015.

[13] V. Petrov, M. Komarov, D. Moltchanov, J. Jornet, and Y. Koucheryavy, "Interference Analysis of EHF/THF Communications Systems with Blocking and Directional Antennas," in *Proc. of Global Communications Conference (GLOBECOM)*, 2016, pp. 1–7.

[14] V. Petrov, M. Komarov, D. Moltchanov, J. M. Jornet, and Y. Koucheryavy, "Interference and SINR in millimeter wave and terahertz communication systems with blocking and directional antennas," *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1791–1808, 2017.

[15] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. Soong, and J. C. Zhang, "What will 5G be?" *IEEE Journal on selected areas in communications*, vol. 32, no. 6, pp. 1065–1082, 2014.

[16] M. Polese, M. Giordani, M. Mezzavilla, S. Rangan, and M. Zorzi, "Improved handover through dual connectivity in 5G mmWave mobile networks," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 9, pp. 2069–2084, 2017.

[17] D. Moltchanov, A. Ometov, S. Andreev, and Y. Koucheryavy, "Upper bound on capacity of 5G mmwave cellular with multi-connectivity capabilities," *Electronics Letters*, vol. 54, no. 11, pp. 724–726, 2018.

[18] M. Gapeyenko, V. Petrov, D. Moltchanov, M. R. Akdeniz, S. Andreev, N. Himayat, and Y. Koucheryavy, "On the degree of multi-connectivity in 5g millimeter-wave cellular urban deployments," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 2, pp. 1973–1978, 2019.

[19] M. Gerasimenko, D. Moltchanov, M. Gapeyenko, S. Andreev, and Y. Koucheryavy, "Capacity of multi-connectivity mmwave systems with dynamic blockage and directional antennas," *IEEE Transactions on Vehicular Technology*, 2019.

[20] V. Petrov, D. Solomitckii, A. Samuylov, M. A. Lema, M. Gapeyenko, D. Moltchanov, S. Andreev, V. Naumov, K. Samouylov, M. Dohler *et al.*, "Dynamic multi-connectivity performance in ultra-dense urban mmWave deployments," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 9, pp. 2038–2055, 2017.

[21] V. Petrov, M. A. Lema, M. Gapeyenko, K. Antonakoglou, D. Moltchanov, F. Sardis, A. Samuylov, S. Andreev, Y. Koucheryavy, and M. Dohler, "Achieving end-to-end reliability of mission-critical traffic in softwarized 5G networks," *IEEE Journal on Selected Areas in Communications*, vol. 36, no. 3, pp. 485–501, 2018.

[22] P. Nain, D. Towsley, B. Liu, and Z. Liu, "Properties of random direction models," in *IEEE 24th Annual Joint Conference of the IEEE Computer and Communications Societies.*, vol. 3, March 2005, pp. 1897–1907.

[23] 3GPP, "Study on channel model for frequencies from 0.5 to 100 GHz (Release 14)," 3GPP TR 38.901 V14.3.0, January 2018.

[24] R. Kovalchukov, D. Moltchanov, A. Samuylov, A. Ometov, S. Andreev, Y. Koucheryavy, and K. Samouylov, "Evaluating SIR in 3D millimeter-wave deployments: Direct modeling and feasible approximations," *IEEE Transactions on Wireless Communications*, vol. 18, no. 2, pp. 879–896, 2019.

[25] R. Kovalchukov, A. Samuylov, D. Moltchanov, A. Ometov, S. Andreev, Y. Koucheryavy, and K. Samouylov, "Modeling three-dimensional interference and SIR in highly directional mmWave communications," in *GLOBECOM 2017-2017 IEEE Global Communications Conference*. IEEE, 2017, pp. 1–7.

[26] A. B. Constantine *et al.*, "Antenna theory: analysis and design," *Microstrip Antennas (third edition), John Wiley & Sons*, 2005.

[27] V. Begishev, A. Samuylov, D. Moltchanov, E. Machnev, Y. Koucheryavy, and K. Samouylov, "Connectivity properties of vehicles in street deployment of 3GPP NR systems," in *2018 IEEE Globecom Workshops (GC Wkshps)*. IEEE, 2018, pp. 1–7.

[28] S. Parkvall, E. Dahlman, A. Furuskar, and M. Frenne, "NR: The new 5G radio access technology," *IEEE Communications Standards Magazine*, vol. 1, no. 4, pp. 24–30, 2017.

[29] S.-Y. Lien, S.-L. Shieh, Y. Huang, B. Su, Y.-L. Hsu, and H.-Y. Wei, "5G new radio: waveform, frame structure, multiple access, and initial access," *IEEE Communications Magazine*, vol. 55, no. 6, pp. 64–71, 2017.

[30] D. G. Kendall, "Stochastic processes occurring in the theory of queues and their analysis by the method of the imbedded markov chain," *The Annals of Mathematical Statistics*, pp. 338–354, 1953.

[31] C. D. Meyer, *Matrix analysis and applied linear algebra*. Siam, 2000, vol. 71.

[32] M. Gapeyenko, A. Samuylov, M. Gerasimenko, D. Moltchanov, S. Singh, E. Aryafar, S.-p. Yeh, N. Himayat, S. Andreev, and Y. Koucheryavy, "Analysis of human-body blockage in urban millimeter-wave cellular communications," in *Communications (ICC), 2016 IEEE International Conference on*. IEEE, 2016, pp. 1–7.

[33] S. M. Ross, *Introduction to probability models*. Academic press, 2014.

[34] R. Groenevelt, "Stochastic models for mobile ad hoc networks," INRIA Sophia-Antipolis, PhD thesis, 2005.

[35] D. J. Daley and L. Servi, "Idle and busy periods in stable M/M/k queues," *Journal of applied probability*, vol. 35, no. 4, pp. 950–962, 1998.

[36] B. P. Zeigler, T. G. Kim, and H. Praehofer, *Theory of modeling and simulation*. Academic press, 2000.

[37] H. G. Perros, "Computer simulation techniques: The definitive introduction!" 2009.

[38] G. S. Fishman and L. S. Yarberry, "An implementation of the batch means method," *INFORMS Journal on Computing*, vol. 9, no. 3, pp. 296–310, 1997.