ADAM: An Adaptive Access Mechanism for NB-IoT Systems in the 5G Era

SUBIN NARAYANAN[ID]**, (Graduate Student Member, IEEE), DIMITRIS TSOLKAS, NIKOS PASSAS, AND LAZAROS MERAKOS**
Department of Informatics and Telecommunications, National and Kapodistrian University of Athens, 157 72 Athens, Greece
Corresponding author: Subin Narayanan (snarayanan@di.uoa.gr)

ABSTRACT The technology of Narrow Band Internet of Things (NB-IoT) is expected to contribute towards the support of massive Machine Type Communications (mMTC), i.e., one of the fundamental performance objectives of 5G networks. However, that call for massiveness, requests a thorough revisit on the NB-IoT access procedure. Considering this, the performance of the NB-IoT access procedure is analysed under different network densities and for various combinations of the related configuration parameters. The analysis led to a useful reference table with 3GPP compliant configurations that maximize the access success probability in an NB-IoT system. The table enables an adaptive access mechanism, where the access parameters are adjusted according to the expected network density. The evaluation results quantify the gain that adaptiveness can provide to the performance of the access procedure.

INDEX TERMS NB-IoT, 5G, massive MTC, random access, cellular IoT.

I. INTRODUCTION

The term machine type communication (MTC) refers to communication among terminals or devices where data are exchanged without human intervention. Different from the existing human type communication (HTC), the MTC paradigm is characterized by periodic or infrequent small data transmissions, which mainly originate at the end-devices.

Already, various solutions in the literature deal with service provisioning for MTC. For example, OneM2M [1], a collaboration of eight standardisation organizations (SDOs), works towards a common machine to machine (M2M) service layer that can be readily embedded within various hardware devices, and promises to connect myriads of them with M2M application servers worldwide. In the same direction, 3rd Generation Partnership Project (3GPP) has released service requirements for MTC (since the Rel. 14 [2]). Also, 5G-PPP (5G Infrastructure Public Private Partnership) has adopted a service-oriented approach, introducing the term vertical industries, where MTC services is one of the key vertical sectors.

The associate editor coordinating the review of this manuscript and approving it for publication was Wei-Wen Hu[ID].

From the network perspective, the specifications for enabling MTC are also available [3]–[7]. Two major network deployments have been coined: i) the direct use of the mobile cellular infrastructure [8], and ii) the independent use of infrastructures deployed by service providers. Here, we consider the former approach for two main reasons. First, because it can provide a high level of reliability and security, due to the use of licensed spectrum controlled by the operator, and second, because it is based on an already mature (with global coverage) infrastructure that evolves towards the 5G and beyond networks.

For the target deployment approach, the so-called Cellular IoT (CIoT) system has been presented, since 3GPP Rel. 13 [9]. Three CIoT technologies have been introduced, namely, the Extended Coverage GSM (EC-GSM), the Narrow-Band IoT (NB-IoT), and the enhanced MTC (eMTC). The EC-GSM was deployed in the GSM network, while NB- IoT and eMTC are integrated in the LTE/4G network. These CIoT technologies are expected to provide support to the wide range of MTC services with diverse requirements, foreseen in the 5G and beyond networks [10]–[12].

Currently, NB-IoT is leading the course of the technologies for MTC in the licensed spectrum [13], [14], and it is expected

to be the major player for the support of massive MTC in the future. However, accommodating a massive number of MTC devices imposes new technical challenges on the PHY and MAC layers, such as the need for advanced access schemes, control overhead reduction mechanisms, energy efficiency, coverage enhancements, cost-effectiveness, and security [15], [16].

Focusing on the access procedure an enormous number of MTC devices simultaneously try to access the network by randomly choosing transmission resources of a limited licensed spectrum portion. Inevitably, the massiveness of the devices leads to a collision in the resource selection, which in turn degrades the overall system performance [17], [18]. The, so called, collision-effect defines a well-studied problem, known from the ad hoc networks, which causes a delay in the attachment of a device to the system. This delay can, in turn, prolong the service establishment and, potentially, lead to a connection failure. To improve system performance, we revisit the challenge of mitigating the collision-effect in the NB-IoT random access channel by jointly optimizing various resource control factors (configurable access parameters) provided by the NB-IoT Radio Access Network (RAN).

The rest of the paper is organized as follows. Section II, includes related state of the art and lists the contributions of this paper. The random access (RA) procedure used in NB-IoT systems is described in Section III. In the same section, the configurable access parameters, which can be tuned for optimising the RA procedure, are identified. The proposed system model is presented in section IV. Section V includes a performance analysis based on the system model adopted. The proposed adaptive RA procedure is presented and evaluated in section VI. Finally, conclusions and insights are included in section VII.

II. RELATED WORK AND CONTRIBUTIONS

Several PHY and MAC layer solutions have been proposed to address the massive access scenario in NB-IoT systems. A related comprehensive study with focus on the MAC layer solutions can be found in [19]. Here, we cluster the related approaches in two major categories. The first category includes the contributions that propose changes in the standardised RA procedure, while the second includes approaches that propose optimisations without affecting the format of the RA procedure specified by 3GPP.

For the first category, in [20] and [21], the standardised four-stage access protocol [22] is replaced by a one-stage protocol (data delivery and access notification are done in a single stage) and a two-stage access protocol (first access attempt and second data delivery), respectively, to expedite the short data transmissions. Another approach proposes the so called signature-based access [23], [24], where the preamble, which corresponds to the device's signature, is spread over multiple Physical Random- Access Channels (PRACH). This approach reduces the signaling overhead but suffers from the problem of false positive detection of the devices. In [25], [26], the authors proposed an efficient protocol for

small data transmission through the connection request message of the RA procedure. In their method, user plane establishment is avoided for the Protocol Data Unit (PDU) session by exchanging the DL/UL data between the devices and Mobility Management Entity (MME) through Non-Access Stratum (NAS) signalling; in this way, signalling overhead is reduced. A low latency RA procedure using compression sensing (CS) is proposed in [27], [28]. In the conventional RA procedure, the Zadoff–Chu (ZC) sequence limits the number of orthogonal preambles. However, in the CS-based method proposed in [27], [28], non-orthogonal preambles can be used. This, in turn, reduces collisions as well as the signaling required to complete the RA process.

As it has been revealed from the studies described above, in a massive access environment, guaranteeing the use of separate channels or preambles is an extremely challenging task. Considering this, more disruptive approaches have emerged, such as those based on the concept of network coding [29], [30]. Network coding techniques using random linear coefficients and other coding features for massive MTC applications are presented in [31]. Multiple optimization models have been presented for Physical-layer Network Coding (PNC) schemes for random access [32] and energy efficient relay-based massive MTC networks [33]. Further with the advancements in LTE-A and 5G environments where devices can also participate in a cooperative out-band device to device (D2D) communications, MAC protocols can be benefited from network coding principles. Authors in [35] present an adaptive cooperative network coding based MAC protocol for D2D data exchange using idle devices in the network as relays. The performance analysis shows that the proposed model outperforms state of the art schemes and a cross network analysis has been done to study the tradeoffs in terms of energy efficiency, network parameters, and scheduling policies.

The approaches in the second category improve the RA performance by configuring the network access factors, such as the back-off window, the Access Barring (AB) factor, the retransmission and repetition factors, and the resource scheduling mechanism. In this context, the back-off window and AB parameters are considered as the most useful tools to alleviate congestion and, thus, collisions in the massive access scenario.

Prioritized access with a distinct back-off interval for MTC specific traffic is used in [36]. The applications are grouped in different classes and the RACH is virtually divided and allocated to them. Prioritization is achieved by using different back-off interval for each traffic type. In [37], the MTC devices use a larger back-off interval compared to that used by HTC devices; thus, the impact of MTC over HTC is reduced.

A dynamic Access Barring (AB) method for bursty MTC traffic is proposed in [38], where contending devices are barred from transmitting in a slot with a certain probability, which depends on the traffic load in each slot. However, in [38], it is assumed that the base station knows the exact number of MTC devices contending in each slot.

A differentiated AB mechanism for low latency NB-IoT applications is proposed in [39]. The devices are classified into two groups (delay-sensitive and delay-insensitive) and, under the massive access scenario, the proposed algorithm bars only the devices with delay-insensitive services. In the same direction, a traffic aware AB factor determination is proposed in [40] where the AB factor is varied in each RA slot based on the total number of contending devices in a RA slot. In [41], a cooperative AB mechanism is proposed, where the base station shares information about its traffic load. The AB factor is determined at each base station considering the congestion level in the neighboring cell. The cooperative AB mechanism reduces the access delay by 30% compared to the static AB mechanism. A 3GPP standard compatible probabilistic retransmission method to reduce the preamble collision is proposed in [42], where a device with an uplink grant for RA uses a probability value to decide whether or not to transmit the uplink signaling/data. The authors in [43] analyses the trade-off between the repetition and retransmission mechanisms and how it affects the RA success probability of a single cell NB-IoT system.

The collisions during the massive access have a direct impact on the energy efficiency of the MTC devices. Considering this aspect, the authors in [45] proposed a two-dimensional channel coding and link adaptation scheme by selecting the optimal combination of Modulation and Coding Scheme (MCS) as well as the subframe repetition number considering the reliability and throughput of the system. The authors in [46] proposed a two-step procedure for link adaptation in NB-IoT. In their scheme, an inner loop link adaptation is used to guarantee transmission block error rate by optimizing the repetition number, and an outer loop link adaptation is to select the MCS level and repetition number. In the same direction, the authors in [47] propose a bandwidth coupled repeated transmission method to enhance the energy efficiency of the MTC system. Finally, a scheduler for the NB-IoT system is proposed in [48], [49] as a mechanism to improve the average access delay and ensure the energy efficiency of the system.

On top of the research effort listed above where the target is the fine tuning of one or two access configuration parameters, we provide a performance analysis on the joint effect of several configuration parameters to the performance of RA procedure, and propose an adaptive mechanism for applying the results of the analysis in a real system. More precisely the contributions are summarized as follows:

- A comprehensive study on the NB-IoT radio access network, focusing on the RA procedure, is presented. The main attributes that affect RA performance have been revealed from this study.
- We modeled the random access procedure of a three CE level NB-IoT system as a multi-channel multi-band slotted aloha system. Considering the intra-cell, and inter-cell interference of a multi-cell NB-IoT system with subframe repetition enabled for coverage
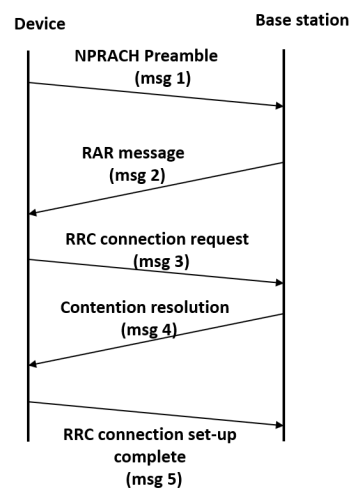

FIGURE 1. Random-access procedure in NB-IoT.

extension, we derived the expression for RACH success probability and access latency.
- Performance analysis of the above-mentioned system for 3GPP-defined values of access barring factor, repetition number, retransmission number, and back-off window size (configurable access parameters). The outcome of this study is a look-up table that provides the optimal combination of the configurable access parameters for a given traffic load which maximizes the access success probability.
- An ADAptive Mechanism (ADAM) for optimizing the RA procedure in NB-IoT systems is proposed. ADAM uses the verified results in the look-up table, to enable an on-line configuration of the RA performance-affecting factors.

III. RANDOM ACCESS IN NB-IoT

For accessing the NB-IoT network, the NB-IoT devices perform a handshaking approach with the base station [51], [56], [58], [59]. Practically, the exchange of five messages is required, as depicted in Fig. 1. The first message is sent from a device towards the base station and includes the, so called, preamble message (Msg. 1 in Fig. 1). The base station responds with a Radio Access Response (RAR) message (Msg. 2 in Fig. 1) and, then, three Radio Resource Control (RRC) messages are exchanged (Msg. 3, 4 and 5 in Fig. 1). The most challenging part of that procedure is the preamble transmission, since it is based on a random access process, during which NB-IoT devices compete for accessing the network.

The entire procedure is performed in a 180KHz spectrum portion, which is structured as depicted in Fig. 2. In the frequency domain 48 subchannels are considered, while in the time domain, symbol groups (time slots) are defined, where each one of them is composed of 5 symbols and a cyclic prefix. The total number of 48 subchannels is actually the result of dividing the 180 KHz of the available bandwidth by 3.75 KHz, which is the subcarrier spacing for the random

FIGURE 2. NB-IoT preamble structure.

access procedure in NB-IoT systems [52], [53], [55]. Each device has the potential to periodically transmit a preamble by using the resource portion of 180 KHz, also referred to as Radio Access Opportunity (RAO). Within an RAO, the preambles are repeated to increase the coverage as shown in Fig. 2. A description of repetition mechanism in the NB-IoT is presented in the next section.

A. THE CONFIGURATION FACTORS

1) REPETITIONS

The resource unit that is used for a preamble message transmission is defined by four continues symbol groups distributed in four different sub-channels. The same preamble message might repeat multiple times within a RAO, based on the channel conditions experienced by the NB-IoT device. Practically, reference signal measurements (during the cell search process) categorise the devices into three Coverage Enhancement (CE) levels. Those CE levels are defined based on the maximum coupling loss (MCL) experienced, using the following thresholds [56], [57]:

- CE-0: MCL≤144 dB
- CE-1: 144 dB < MCL≤155.7 dB
- CE-2: 155.7 dB < MCL ≤ 164 dB

In the case that a device belongs to CE-0, the repetition number is chosen from the set {1, 2}; otherwise (for CE-1 and CE-2 levels) the repetition number is chosen from the set {4, 8, 16, 32, 64, 128}. The resource units that carry the repetitions are pseudo-randomly selected, based on a frequency hopping formula. At a specific CE level, the time needed for a preamble transmission (duration of a single RAO) equals the time needed for completing the number of the repetitions selected for the specific CE level. Here, the preamble repetition number is denoted by *Rep*. Thus, considering that the four-group preamble lasts for 5.6 ms, the RAO duration in the CE level *b* is as follows:

$$t_b = Rep \times 5.6 \ ms \qquad (1)$$

where

$$Rep = \begin{cases} \{1, 2\} & \text{for CE-0} \\ \{4, 8, 16, 32, 64, 128\} & \text{for CE-1, CE-2} \end{cases}$$

2) ACCESS BARRING

The NB-IoT system has an access barring mechanism that can postpone a device's preamble transmission for a period of time. This means, that when a device is ready to transmit a preamble, it is allowed to proceed immediately with the transmission only if the access barring mode is deactivated. Practically, in the case that the base station detects that the number of recent preamble receptions is higher than a threshold, it requests all the devices to switch to the access barring mode. In this mode, each device generates a random value between 0 and 1, and only if the generated value is lower than a barring factor p_{AB}, the transmission is performed. Otherwise, the device is barred from the RA procedure for a duration t_{AB} given by:

$$t_{AB} = (0.7 + 0.6 \times (\text{generated random value}) \times T_{AB}) \qquad (2)$$

where $T_{AB} \in \{4, 8, 16, \ldots, 512 \ \text{sec}\}$, as defined by the 3GPP [49], [58]. The device generates a new random value in the next available RAO, and this continues until the generated value is less than p_{AB}. In this way, the number of contending devices in a RAO can be controlled. To apply the AB (Access Barring) method the base station broadcasts the barring factor p_{AB}, $0 \leq p_{AB} \leq 1$ and barring time, T_{AB}, every Transmission Time Interval (TTI) [58]. Both these parameters (p_{AB} and T_{AB}) define the AB configurable factor [68].

3) BACK-OFF WINDOW

While the AB mechanism provides an on demand control of the contention in the system, a conventional back-off mechanism is also applied prior any transmission. Upon the RA failure, the device delays it's next RA attempt for a maximum time duration given by the back-off window size. The back-off window size, denoted here by BW, is broadcast by the base station, and it can be adjusted according to the device density in each RAO. In this way the number of contending devices is distributed over time, and thus, the congestion in the RAO is reduced.

4) RETRANSMISSIONS

As soon as a devices sends a preamble, the device expects to receive the RAR message within a time window which has a pre-configured duration (set by the base station with broadcast messages). This window is called contention resolution window, and it is denoted here as T_{CR}. If a device does not receive the RAR message within this time window, the RA procedure is considered as failed, and restarts. The reasons for a fail are the following:

- **Collision.** It occurs if two or more devices simultaneously use the same sub-channel for the transmission of the preamble, resulting in unsuccessful decoding of the received signals at the base station.
- **Insufficient channel conditions.** For a single transmission in a preamble group, there is the possibility that the base station may not receive/decode the message successfully due to insufficient channel conditions (low received power at the base station).

After a fail, the procedure can restart multiple times based on the re-transmission factor, which defines the maximum number of re-transmissions allowed. Actually, thresholds for the maximum number of allowed re-transmissions in all the CE levels as well as for each CE level, are defined. We denote by TM_b (where $b \in \{0,1,2\}$), the maximum number of preamble transmission attempts that an NB-IoT device can perform at a CE level b, and by TM_G the maximum number of attempts an NB-IoT can perform in the entire system (such that $TM_b \leq TM_G$). The value of TM_G is limited, since there are also service-level latency bounds, denoted here by T_{max} [61]. A device should add up its transmission attempts in all the CE levels, when checking for TM_G. The absence of a RAR message after all the re-transmissions may refer to poor channel condition. To compensate with this situation the devices may decide to move to the next higher CE level (where more repetitions are used) as shown in Fig. 3.

FIGURE 3. Preamble transmission in NB-IoT.

5) SUB-CHANNELS SPLIT
The 48 sub-channels that are in total available for the preamble transmission are distributed among the CE levels. Since each CE level is characterized by a different repetition number, the time-correlated mutual interference among the transmitted preamble symbol groups is different for each CE level. Hence, depending upon the traffic, the sub-channels can be distributed among CE levels in such a way that the combined effect of preamble collision and preamble mutual interference is minimized in the NB-IoT system. For the analysis here, we denote the number of sub-channels in CE-0, CE-1 and CE-2, by $SC0$, $SC1$, and $SC2$, respectively.

B. RANDOM ACCESS TIMING
The valid values for the above mentioned configuration factors are summarised in Table 1. Fig. 3 illustrates the timing of RA attempts for multiple devices in a three CE-level NB-IoT system, together with coupling loss calculation for each CE level. As depicted in Fig. 3, within a given T_{max}, the maximum number of RAOs in the CE level b is as follows:

$$N_b = \left\lfloor \frac{T_{max}}{T_b} \right\rfloor \tag{3}$$

where T_b is the RAO periodicity in the CE level b. The value of T_b should be large enough so that the device will be informed within the T_b period whether its RA attempt was

TABLE 1. Configurable access parameters (notation and values space).

Parameter	Notation	Values Space
Repetition number	Rep	1, 2, 4, 8, 16, 32, 64, 128
Access barring factor	$AB:\{p_{AB}, T_{AB}\}$	p_{AB}: $0 \leq p_{AB} \leq 1$ and $T_{AB} \in \{4,8,16...\}$ (in sec)
Back-off window size	BW	2^k, $k \in \{0, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19\}$ (in ms)
Retransmission number in CE level b	TM_b	$TM_b \leq N_b$ where N_b is the maximum number of RAOs within the given Latency T_{max}
Subcarrier split	$SC:\{SC0, SC1, SC2\}$	A total of 48 subchannels split into thee portions, one for each CE level, $SC0$, $SC1$, $SC2$, such as $SC0 + SC1 + SC2 = 48$

successful or not. Thus:

$$T_b > t_b + T_{RAR} + T_{CR} \tag{4}$$

where, t_b represents the NB-preamble duration in the CE level b, T_{RAR} represents the duration of the RAR message response window, and T_{CR} represents the duration of the contention resolution window.

IV. SYSTEM MODEL
In this section, we elaborate on how the above-listed configurable access parameters affect the access success probability in the RA procedure. For the estimation of the access success probability a set of deployment and functional considerations are made.

A. DEPLOYMENT AND FUNCTIONAL CONSIDERATIONS
The system under consideration is a multi-cell and multi-user NB-IoT system, deployed over an area of size A, where the base stations are distributed following a Homogeneous Poisson Point Process (HPPP), with intensity λ_B. Similarly, NB-IoT devices are distributed following HPPP, with intensity λ_D. We consider an independent and identically distributed (i.i.d) Rayleigh fading channel, where the channel power gain is assumed to be an exponentially distributed random variable with unit mean and unit variance.

Since, the spectrum is divided into three portions, each one allocated to the devices of a single CE level (based on the sub-channel split factor), the intensity of NB-IoT devices in CE level b is denoted by λ_{D_b}, such as $\lambda_D = \sum_{b=0}^{2} \lambda_{D_b}$. One step further, in a single cell (area served by a single base station) and for a specific CE level, the expected number of devices in the unit area is defined by λ_{D_b}/λ_B.

In our model, we assume that the deployment characteristics of the system change periodically every T_{max} seconds (latency requirement of NB-IoT service). Thus, at the beginning of each T_{max} period the expected number of devices and base stations are estimated by the realisation of the above-mentioned HPPPs. Also, in a single RAO within the T_{max} period, a portion of the total λ_{D_b}/λ_B devices in a cell will attempt access. We assume that the distribution of those devices in the various RAOs within the T_{max} period can be

described by a time-limited Beta distribution. Considering this, the expected number of devices in the unit area that will transmit their preamble for access during the nth RAO is given by:

$$\lambda_{D_b}(n) = \lambda_{D_b}/\lambda_B \int_n^{n+1} B(n)dn \qquad (5)$$

where $B(n)$ is the Beta distribution, and is given by:

$$B(n) = \frac{n^{\Omega-1}(N_b - n)^{\beta-1}}{N_b^{\Omega+\beta-1} Beta(\Omega, \beta)}, \qquad \Omega, \beta > 0 \qquad (6)$$

where $Beta(\Omega, \beta)$ is the Beta function, N_b is the number of RAOs in CE level b during the T_{max} period, and the parameters Ω and β are tuning factors of the Beta function.

Based on the random access procedure, in the case that the AB factor is enabled at the nth RAO, the base station commands each device to spare its transmissions and attempt access with probability p_{AB} for a period t_{AB}. Consequently, (5) is reshaped as follows:

$$\lambda_{D_b}^{AB}(n) = \begin{cases} p_{AB}\lambda_{D_b}/\lambda_B \int_n^{n+1} B(n)dn & \text{if AB enabled} \\ \lambda_{D_b}/\lambda_B \int_n^{n+1} B(n)dn & \text{if AB disabled} \end{cases} \qquad (7)$$

B. PREAMBLE SUCCESS PROBABILITY-NO REPETITIONS, NO RETRANSMISSIONS

As explained in Section III, a preamble occupies one sub-channel, from the set of SC_b that are available, to send each one of the four symbol groups. In this subsection, we analyse the expected Signal-to-Interference-plus-Noise Ratio (SINR) at a base station and calculate the success probability of the preamble transmission. The analysis is valid for any sub-channel of the RAO and any CE level; thus, we define the expected number of devices that transmit their preamble in a single sub-channel of any RAO as follows.

$$\lambda_{pream}(n, b) = \lambda_{D_b}^{AB}(n)/SC_b, \forall n, \forall b \qquad (8)$$

A preamble can be received at the base station successfully if its SINR is above a given threshold, denoted here by γ_{th}. The probability that the preamble (which is actually carried in a single transmission since no repetitions are assumed) is received at the base station with SINR that exceeds the SINR threshold, is provided as follows:

$$P_{Preamble}(n, b) = P(\frac{\rho h}{m(I_{Inter} + I_{Intra} + \sigma^2)} > \gamma_{th})$$
$$= P(h > \frac{m\gamma_{th}}{\rho}(I_{Inter} + I_{Intra} + \sigma^2)) \qquad (9)$$

where $m = 4$ (since a preamble consist of four symbol groups), γ_{th} is the SINR threshold at the base station, σ^2 is the noise power, ρ is the full path-loss inversion power control threshold, I_{Intra} represents the aggregate intra-cell interference, I_{Inter} represents the aggregate inter-cell interference, and h is the channel power gain. As mentioned earlier, h is

an exponentially distributed random variable with unit mean and unit variance. Hence,

$$P(h > \frac{m\gamma_{th}}{\rho}(I_{Inter} + I_{Intra} + \sigma^2))$$
$$= E[exp(-\frac{\gamma_{th}m}{\rho}(I_{Inter} + I_{Intra} + \sigma^2))] \qquad (10)$$

where $E[\cdot]$ denotes the expected value. Considering that the quantity $E[exp(-x)]$ defines the Laplace Transform of a variable x, we can reshape (10) as follows:

$$E[exp(-\frac{\gamma_{th}m}{\rho}(I_{Inter} + I_{Intra} + \sigma^2))]$$
$$= exp(\frac{-m\gamma_{th}\sigma^2}{\rho})L_{intra}(\frac{-m\gamma_{th}\sigma^2}{\rho})L_{inter}(\frac{-m\gamma_{th}\sigma^2}{\rho}) \qquad (11)$$

where $L_I(\cdot)$ denotes the Laplace Transform of the aggregate interference I [44], [64]. Below we derive the Laplace Transform of the inter-cell and intra-cell interference.

1) LAPLACE TRANSFORM OF AGGREGATED INTER-CELL INTERFERENCE

The Laplace Transform of the aggregate inter-cell interference received at the base station is given by:

$$L_{inter}(\frac{-m\gamma_{th}\sigma^2}{\rho})$$
$$= exp(-2(\gamma_{th})^{\frac{2}{\alpha}}\lambda_{pream}(n, b) \int_{\gamma_{th}^{\frac{-1}{\alpha}}}^{\infty} 1 - \frac{1}{(1 + y^{-\alpha})^m}ydy) \qquad (12)$$

where α is the path-loss parameter. The proof of (12) is given in Appendix A.

2) LAPLACE TRANSFORM OF AGGREGATE INTRA-CELL INTERFERENCE

Let us denote by F the number of intra-cell interfering NB-IoT devices in a specific cell. In practice, F represents the portion of the devices attempting in a single RAO (defined in (8)) that have selected the same sub-carrier for their preamble transmission. We denote this portion by $P_{IN}(F = Z)$. Assuming that the probability Mass Function (PMF) of the number of interfering intra-cell devices is derived by the Monte Carlo Method [69], $P_{IN}(F = Z)$ quantity is given in (13) below.

$$P_{IN}(F = Z) = \frac{c^{c+1}\Gamma(Z + c + 1)(\lambda_{pream}(n, b))^Z}{\Gamma(c + 1)\Gamma(Z + 1)(\lambda_{pream}(n, b))^{Z+c+1}} \qquad (13)$$

where $\Gamma(\cdot)$ is the gamma function, and $c = 3.575$ is a constant related to the approximate PMF of the cells defined by the deployment process. The, so called, Voronoi cells assume that the area is portioned in such a way that each device is attached to a base station, and no device is located outside the coverage of a base station. The Laplace Transform of the aggregate intra-cell interference conditioned on $F = Z$ number of interfering devices is given by:

$$L_{inter}(\frac{-m\gamma_{th}\sigma^2}{\rho}) = \frac{1}{(1 + \gamma_{th})^{mZ}} \qquad (14)$$

The proof of (14) is given in the Appendix B. By substituting equations (14) and (12) in (11), we get the probability that a preamble is successfully received at the base station. The result is given below.

$$P_{Preamble|Z}(n, b)$$
$$= P(h > \frac{m\gamma_{th}}{\rho}(I_{Inter} + I_{Intra} + \sigma^2)|F = Z) \quad (15)$$

$$= \frac{exp(-\frac{m\gamma_{th}\sigma^2}{\rho} - 2(\gamma_{th})^{\frac{2}{\alpha}}\lambda_{pream}(n, b)\int_{\gamma_{th}^{\frac{-1}{\alpha}}}^{\infty} 1 - \frac{1}{(1+y^{-\alpha})^m}y dy)}{(1 + \gamma_{th})^{mZ}} \quad (16)$$

C. PREAMBLE SUCCESS PROBABILITY-REPETITIONS ENABLED, NO RETRANSMISSIONS

In the NB-IoT system, a preamble transmission success happens if at least one of the repetitions is successful. Accordingly, the probability that at least one of the *Rep* repetitions is received at the base station successfully, conditioned on Z number of active intra-cell interfering devices is given by:

$$P_{RAO}(Rep, n, b) = 1 - \prod_{r=1}^{Rep}(1 - P_{Preamble|Z}(n, b)) \quad (17)$$

Since only one preamble transmission of those that compete in a single subcarrier will be decoded at the base station, the success probability for a specific/target device is calculated as follows.

$$P'_{RAO}(Rep, n, b)$$
$$= \sum_{Z=0}^{\infty} P_{IN}(F = Z)P_{RAO}(Rep, n, b)(1 - P_{RAO}(Rep, n, b))^Z \quad (18)$$

D. PREAMBLE SUCCESS PROBABILITY-REPETITIONS AND RETRANSMISSIONS ENABLED

As explained in Section III, the collided devices in an RAO will perform uniform back-off mechanism by randomly choosing a back-off window value from the available set $[0, BW]$. Consequently, the calculation of the expected number of devices in the unit area that will transmit their preamble for access during the nth RAO, should be updated every RAO to include the re-transmissions, i.e, the number of devices that have experienced a collision and come back for a new access attempt after the back off process. Thus, for the success probability in each RAO the $\lambda_{pream}(n, b)$ should be updated by the $\lambda'_{pream}(n, b)$ quantity, which incorporates the impact of back-off process and it is defined as follows.

$$\lambda'_{pream}(n, b) = \lambda_{D_b}^{AB,BW}(n)/SC_b, \forall n, \forall b \quad (19)$$

where,

$$\lambda_{D_b}^{AB,BW}(n)$$
$$= \lambda_{D_b}^{AB}(n) + \sum_{k=1}^{n-1} \psi_{k,n-1}^b \lambda_{D_b}^{AB}(k)(1 - P_{RAO}(Rep, k, b)) \quad (20)$$

where, $\psi_{(k,n-1)}^b$ is the probability that an NB-IoT device that collided in the kth RAO will fall in the transmission window of the nth RAO for its next RA attempt, and it is given by:

$$\psi_{k,n-1}^b = \begin{cases} \frac{1}{\left\lfloor \frac{BW}{T_b} \right\rfloor} & \text{if } k + \left\lfloor \frac{BW}{T_b} \right\rfloor < n \\ 0 & \text{if } k + \left\lfloor \frac{BW}{T_b} \right\rfloor \geq n \end{cases} \quad (21)$$

E. SYSTEM SUCCESS PROBABILITY AND ACCESS LATENCY

Let us denote by $\lambda_{D_b,Success}^{AB,BW}(n)$ the expected number of devices in an unit area which has successfully transmitted preamble in the nth RAO. This quantity is given by:

$$\lambda_{D_b,Success}^{AB,BW}(n) = \lambda_{D_b}^{AB,BW}(n) \times P_{RAO}(Rep, n, b) \quad (22)$$

Considering the parallel transmissions in the three CE levels, we define the system success probability P_{Access}, as given by:

$$P_{Access} = \frac{\sum_{b=0}^{2}\sum_{n=1}^{N_b} \lambda_{D_b,Success}^{AB,BW}(n)}{\sum_{b=0}^{2}\lambda_{D_b}} \quad (23)$$

We define access latency D_{Access} as the average time required for an NB-IoT device to complete the RA procedure. D_{Access} can be expressed as follows:

$$D_{Access} = \frac{\sum_{b=0}^{2}\sum_{n=1}^{N_b} \lambda_{D_b,Success}^{AB,BW}(n) \times T_b \times (n-1)}{P_{Access} \times \sum_{b=0}^{2}\lambda_{D_b}} \quad (24)$$

V. PERFORMANCE ANALYSIS

In this section, we adopt the model presented in the previous section and quantify the performance of the random access procedure in terms of system success probability and latency. The system under evaluation is described by the set of parameters depicted in Table 2. We have developed a realistic simulation framework to simulate the 3GPP defined NB-IoT random access procedure. The framework captures the randomness of 1) RA slot selection by the device, 2) Preamble selection and transmission, 3) Back-off window selection, and 4) Access Barring (AB) factor and timer selection. Our stochastic geometry-based analysis are verified using this simulation framework. The analysis is performed in three steps.

- First, we select a set of values for the configurable access parameters (explained in Section III, and listed in Table 1) and calculate the performance by varying the device arrival process within a T_{max} period.
- The second performance analysis step includes the quantification of the system success probability in a CE level for different network densities and various realisations of the configurable access parameters. The major result is a set of optimal values for the configurable access parameters for different network densities, summarised in Table 4.
- In the third step, we examine how the decision on the split of the available subcarriers (SC) to the CE levels

TABLE 2. System level parameters.

Parameters	Values
Coverage level	CE-0 (143 dBm), CE-1(149 dBm), CE-2 (158 dBm)
SNR threshold at base station	-4.7 dB
Area	300 km2
Path-loss parameter	3
T_{max}	10 sec
T_b	640 ms, b = {0, 1, 2}
base station noise figure	5 dB
Power control at device	Full path-loss inversion power control ρ = -120 dBm (CE0), ρ = -126 dBm (CE1), ρ = -135 dBm (CE2)
Distribution of devices	Homogeneous Poisson Point Process with $\lambda_{D,b}$
Distribution of base station	Homogeneous Poisson Point Process with λ_B
Density of devices	λ_D/λ_B scalable in the range [100-1000]
Channel model	i.i.d Rayleigh fading channel
Channel bandwidth	180 KHz
Subcarrier spacing	3.75 KHz (Maximum of 48 subcarriers)
Repetition number	Rep = 1 (CE-0), Rep = 4 (CE-1), Rep = 32 (CE-2)

affects the performance. Based on this set of evaluations, a decision factor is proposed for splitting the SC to the CE levels, leading to most efficient 3GPP-valid allocations (depicted in Table 5).

The outcome of the performance analysis (i.e., Tables 4 and 5) are used as look-up tables for the adaptive mechanism proposed in Section VI, where the network update online the configurable access parameters to maximize the access success probability.

A. IMPACT OF THE DEVICE ARRIVAL PROCESS

A quantification of the system success probability for different tuning parameters of the time limited beta distribution in a single CE system (CE-0) is shown in Fig. 4. Three realisations of the beta distribution are adopted. The first one is based on the recommendations from 3GPP [63], where the values $\Omega = 3$, $\beta = 4$ are proposed for configuring the beta distribution. The second one refers to the uniform arrival of the devices, i.e., the beta distribution with values $\Omega = 1$, $\beta = 1$; and the third one, with values $\Omega = 1$, $\beta = 50$ to model a peak/bursty traffic. Since the devices are allowed to access the network at the start of each RAO, the device arrival with $\Omega = 1$, $\beta = 50$ behaves as the one-shot arrival process, where all the devices try to access the network in the very first RAO.

As shown in Fig. 4, at high device density values the system success probability with one-shot arrival is lowest compared to the uniform arrival and Beta distribution (with $\Omega = 3$, and $\beta = 4$). This is because with uniform arrival and Beta distribution with $\Omega = 3$, and $\beta = 4$, the access is spread over T_{max}, while with one-shot arrival, the access is spread only over back-off window (depicted as BW) duration starting from the first RAO. On the other hand, when the network density is low, the NB-IoT system has enough resources (within the T_{max}) to accommodate all the devices in the cell. Since
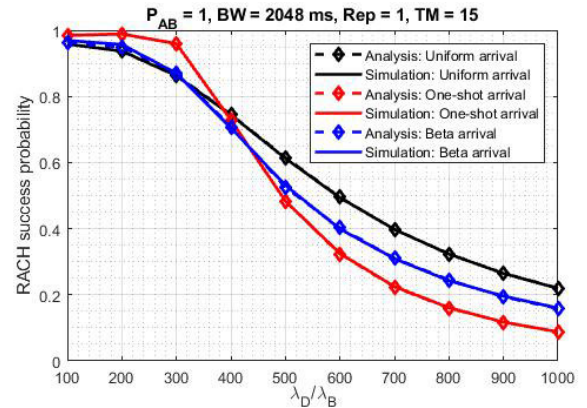
FIGURE 4. Comparison of RA success probability in a multi-cell, single CE level NB-IoT system within a time bound of T_{max}, for different arrival process.

TABLE 3. The mean relative percentage error between analytical and simulation results for single CE scenario.

RAN values	Relative percentage error (%)	
	P_{Access}	D_{Access}
$Rep = 2, BW = 4096$ ms, $SC = 48, TM = 15, p_{AB} = 1$	0.96	1.28
$Rep = 4, BW = 4096$ ms, $SC = 48, TM = 15, p_{AB} = 1$	1.02	1.47
$Rep = 8, BW = 4096$ ms, $SC = 48, TM = 15, p_{AB} = 1$	1.11	1.52
$BW = 4096$ ms, $SC = 42, Rep = 4, TM = 15, p_{AB} = 1$	1.04	1.11
$BW = 8192$ ms, $SC = 42, Rep = 4, TM = 15, p_{AB} = 1$	0.92	1.56
$TM = 4, SC = 48, BW = 4096$ ms, $Rep = 4, p_{AB} = 1$	0.95	1.30
$TM = 12, SC = 48, BW = 4096$ ms, $Rep = 4, p_{AB} = 1$	1.10	1.42
$p_{AB} = 1, SC = 48, BW = 2048$ ms, $Rep = 4, TM = 15$	1.07	1.13
$p_{AB} = 0.5, SC = 48, BW = 2048$ ms, $Rep = 4, TM = 15$	0.98	1.01
$p_{AB} = 0.1, SC = 48, BW = 2048$ ms, $Rep = 4, TM = 15$	1.07	1.44

the one-shot arrival spreads access only over the duration of back-off window size, more re-transmissions are possible within the T_{max}, which in turn increases the system success probability. Since under massive access scenario, one-shot arrival imposes more challenges at the network side, the one-shot arrival has been adopted for the evaluations presented in the following sections.

B. OPTIMAL CONFIGURATIONS WITHIN EACH CE LEVEL

Here, we quantitatively evaluate the relation of the network density with the system success probability and access latency under different values for the configurable parameters. The accuracy of analysis expressed using the mean relative percentage error in Table 3. To facilitate the analysis, we define a parameter called *Resource Parameter (RP)*, as given in (25). The RP gives the average amount of the RA resources that are expected to be available for a device in a CE level within each cell. This is an indicator on the expected
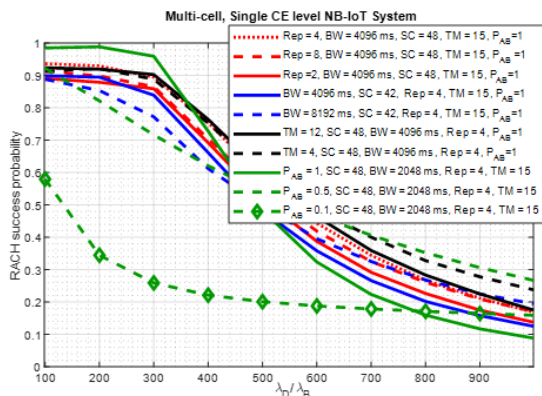
FIGURE 5. Comparison of RA success probability for a different choice of RAN parameters in single CE level.

FIGURE 6. Comparison of access latency for a different choice of RAN parameters in single CE level.

level of saturation of the resources, defined by the following ratio:

$$RP_b = \frac{N_b \times SC_b \times \lambda_B}{\lambda_{D_b}} \qquad (25)$$

Consequently, low network load, leads to $RP_b > 1$; while, for high load (more devices expected than the amount of resources), it holds that $RP_b \leq 1$.

1) IMPACT OF THE BACK OFF WINDOW SIZE
In a congested network, a larger BW offloads the traffic by spreading the access attempts from the devices over time. On the other hand, at low network load, access collisions can be resolved through retransmissions, and a smaller BW provides a higher number of retransmission opportunities within the T_{max}. This behavior is validated in Fig. 5, where as it can be observed, a larger BW yields to better performance when the $RP_b \leq 1$ (high network load), while a smaller BW provides better performance when the $RP_b > 1$ (low network load).

2) IMPACT OF THE REPETITION NUMBER
The repetition number is set to a value in such a way that the preamble detection probability at the base station is at least 99% [65]. If the repetition value is less than this initial value, the probability of signal outage increases, while if the repetition value is much higher than this value, any increase in the system success probability is not guaranteed. This relation is depicted in Fig. 5. For the scenario depicted in Fig.5, the appropriate value for the number of petitions is $Rep = 4$. A lower value of Rep (e.g., $Rep = 2$) decreases the system success probability as the preamble detection probability is decreased at the base station. On the other hand, a high increase on the number of repetitions, e.g., $Rep = 8$, does not reflect any increase in the success probability, since the extra transmissions cause collisions in the system. Moreover, the access latency increases as the repetitions increase, since the base station has to wait for all the repeated subframes to start decoding the preamble as shown in Fig. 6.
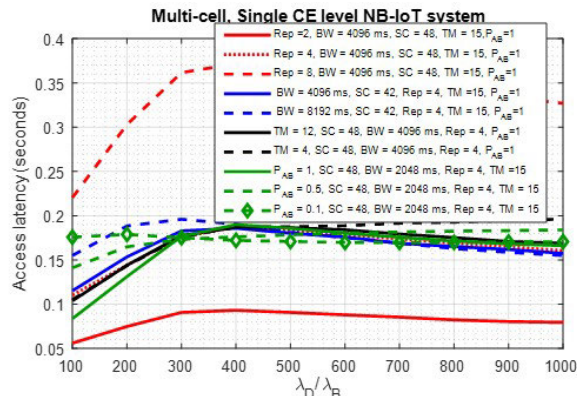
3) IMPACT OF THE AB FACTOR
For the analysis of AB values on system success probability and access latency, we set T_{AB} to 4 seconds (which is the lowest value defined by the 3GPP for T_{AB}) so that barred devices will get maximum possible RA attempts within the T_{max}. It is depicted in Fig. 5 that a large value of AB factor (p_{AB}) at a low network load, and a small value of AB factor at high network load yields to better system success probability. Given the network load is low, the devices are not required to be barred from the RA attempt as the network has enough resources to support the devices. In this scenario, applying a small value of AB factor reduces the number of devices attempting the RA, which in turn reduces the number of successfully accessed devices, even though there is no collision. On the other hand, when the network is congested, it is required to bar the devices from attempting the RA to reduce the preamble collision. Hence a small value of AB factor yields to the better success probability.

4) IMPACT OF THE NUMBER OF RETRANSMISSIONS
A low number of TM is favored in a congested network as it results in better system success probability as shown in Fig. 5. Allocating a lower TM removes the device from contending for access after performing the given number of TM, thus the device density is reduced with time, which in turn increases the system success probability.

5) SUMMARISING THE PERFORMANCE ANALYSIS OUTCOME
Considering that each CE level can be assigned 12 or 24 SCs (since the maximum for the three CE levels is 48) [44], we find out the optimal combinations of AB factor, BW, and TM for each CE level for various device densities, as listed in Table 4. We have also filtered the values so as to list in Table 4 the ones that are compliant to the valid set of values defined by 3GPP [67]. As results, for a service-level latency requirement T_{max} of 10 seconds and a RA periodicity of 640 ms [61], the search space for BW is limited to $BW \in \{1024$ ms, 2048 ms, 4096 ms, and 8192 ms$\}$, and the maximum value of TM is limited to 15.

TABLE 4. Optimal values of the configurable RAN parameters for different network densities.

λ_D/λ_B		Success probability in a CE level	Subcarrier allocation(SC) per CE	RP_b	Repetition Value	AB factor	Back-off Window	Retransmission number (TM)
100	CE-0	0.97	24	10.8	1	1	1024 ms	14
		0.68	12	5.40	1	1	4096 ms	12
	CE-1	0.66	12	5.40	4	1	4096 ms	13
		0.89	24	10.8	4	1	4096 ms	15
	CE-2	0.51	12	5.40	32	1	4096 ms	13
		0.75	24	10.8	32	1	1024 ms	12
200	CE-0	0.69	24	5.40	1	1	4096 ms	13
		0.37	12	2.70	1	0.70	8192 ms	12
	CE-1	0.34	12	2.70	4	0.75	4096 ms	11
		0.65	24	5.40	4	0.95	4096 ms	15
	CE-2	0.26	12	2.70	32	0.70	4096 ms	6
		0.54	24	5.40	32	0.95	4096 ms	7
300	CE-0	0.50	24	3.60	1	0.85	4096 ms	15
		0.25	12	1.80	1	0.55	8192 ms	9
	CE-1	0.23	12	1.80	4	0.55	4096 ms	11
		0.46	24	3.60	4	0.85	4096 ms	4
	CE-2	0.17	12	1.80	32	0.55	8192 ms	8
		0.35	24	3.60	32	0.80	4096 ms	15
400	CE-0	0.37	24	2.70	1	0.65	8192 ms	13
		0.19	12	1.35	1	0.45	8192 ms	11
	CE-1	0.17	12	1.35	4	0.50	8192 ms	9
		0.34	24	2.70	4	0.75	8192 ms	7
	CE-2	0.13	12	1.35	32	0.45	8192 ms	7
		0.26	24	2.70	32	0.70	4096 ms	7
500	CE-0	0.30	24	2.16	1	0.60	8192 ms	15
		0.15	12	1.08	1	0.40	8192 ms	9
	CE-1	0.14	12	1.08	4	0.35	8192 ms	9
		0.27	24	2.16	4	0.60	8192 ms	9
	CE-2	0.11	12	1.08	32	0.35	8192 ms	5
		0.21	24	2.16	32	0.60	8192 ms	4
600	CE-0	0.25	24	1.80	1	0.55	8192 ms	14
		0.13	12	0.90	1	0.35	8192 ms	7
	CE-1	0.13	12	0.90	4	0.30	8192 ms	9
		0.23	24	1.80	4	0.55	8192 ms	6
	CE-2	0.09	12	0.90	32	0.30	8192 ms	9
		0.17	24	1.80	32	0.55	8192 ms	6
700	CE-0	0.11	12	0.75	1	0.30	8192 ms	12
		0.22	24	1.53	1	0.55	8192 ms	6
	CE-1	0.20	24	1.53	4	0.50	8192 ms	9
		0.10	12	0.25	4	0.75	8192 ms	15
	CE-2	0.08	12	0.75	32	0.25	8192 ms	14
		0.15	24	1.53	32	0.50	8192 ms	5
800	CE-0	0.11	12	0.69	1	0.25	8192 ms	12
		0.19	24	1.35	1	0.50	8192 ms	10
	CE-1	0.17	24	1.35	4	0.45	8192 ms	12
		0.09	12	0.69	4	0.30	8192 ms	9
	CE-2	0.07	12	0.69	32	0.20	8192 ms	14
		0.13	24	1.35	32	0.45	8192 ms	4
900	CE-0	0.08	12	0.60	1	0.20	8192 ms	12
		0.15	24	1.20	1	0.45	8192 ms	10
	CE-1	0.17	24	1.20	4	0.40	8192 ms	14
		0.08	12	0.60	4	0.15	192 ms	14
	CE-2	0.06	12	0.60	32	0.15	8192 ms	13
		0.12	24	1.20	32	0.40	8192 ms	11
1000	CE-0	0.09	12	0.54	1	0.20	8192 ms	7
		0.15	24	1.08	1	0.40	8192 ms	7
	CE-1	0.14	24	1.08	4	0.35	8192 ms	7
		0.07	12	0.54	4	0.15	8192 ms	8
	CE-2	0.06	12	0.54	32	0.15	8192 ms	8
		0.11	24	1.08	32	0.35	8192 ms	9

An exhaustive search based approach has been followed to derive Table 4. In the approach, the entries in the table for given action a (which describes the random access at a given RP_b), at a state s (which is given by the combination of [$AB factor$, BW, TM]) is given by $Q(s, a)$ as:

$$Q(s, a) = Q_{pre}(s, a) + [R(s, a) + (Q_{max}(s, a) - Q_{pre}(s, a))] \tag{26}$$

where $Q_{pre}(s, a)$ is the value in the previous state, $Q_{max}(s, a)$ is the maximum expected value in the current state, $R(s, a)$ is the reward for the action performed in the current state.

The table values are initialized to zero at the beginning of the procedure. The first and last state is given by the vectors [0.05, 1024 ms, 1] and [1, 8192 ms, 15] respectively. We arrange states in such a way that for a given value of AB factor, and BW, the value of TM is varied from the first value to last value. In the next step, we increment the value of BW for a given AB, then vary the value of TM from the first value to last value, and so on. The reward value $R(s, a)$ is set to P_{Access} if the P_{Access} of the current state is higher than previous state, and $R(s, a)$ is set to $-P_{Access}$ if the P_{Access} of the current state is lower than previous state. For a given

value of *ABfactor*, and *BW*, if the P_{Access} of the current state is lower than the immediate previous state, we skip the rest of the values of *TM* in the iteration, and the values for all the skipped states are entered as $-\infty$. In the next step, we increase the value of *BW*, and the process continues until the last state. Table 4 is formed by choosing the combination of *AB* factor, *BW*, and *TM* that yields the maximum value $Q(s, a)$ for a given RP_b.

C. PROACTIVE SC ALLOCATION AMONG THE CE LEVELS

As mentioned above, the possible distributions of SC among the CE levels, are {24 12 12}, {12 24 12}, and {12 12 24}. However, considering the access procedure, the allocation of the SCs to the CE levels is performed proactively based on the load/device density at each CE level and then the further configurations per CE level can be applied based on the reference Table 4. To examine the impact of the SC allocation on the performance, we consider four cases for the distribution of the load/device density at each CE level, namely:

- Case 1: Equal device distribution among CE levels (e.g., CE-0: CE-1: CE-2 = 1:1:1)
- Case 2: CE-2 is more crowded than CE-0 and CE-1 (e.g., device distribution among CE levels, CE-0: CE-1: CE-2 = 1:1:2)
- Case 3: CE-1 is more crowded than CE-0 and CE-2 (e.g., device distribution among CE levels, CE-0: CE-1: CE-2 = 1:2:1)
- Case 4: CE-0 is more crowded than CE-1 and CE-2 (e.g., device distribution among CE levels, CE-0: CE-1: CE-2 = 2:1:1)

For the performance analysis we choose the extreme values (worst case scenario) for the configurable parameters, i.e., 1024 ms for the back off window *BW*, 15 for the number of allowed retransmissions *TM*, and 1 for the AB factor p_{AB}. The results of the simulations performed are provided in the following subsections, while the summarised results that define the proposed SC allocations for different distribution of devices among the three CE levels are given in Table 5.

To facilitate the analysis of the subcarrier allocation, we define RP_G, where RP_G is given by:

$$RP_G = \frac{SC_{tot} \times \lambda_B}{\lambda_D} \sum_{b=0}^{2} \frac{N_b}{3} \qquad (27)$$

where $SC_{tot} = 48$ (total number of subcarriers allocated to the NB-IoT system). The RP_G gives the average amount of RA resources expected to available for a device in a three CE level system at the beginning of the RA procedure.

1) DEVICES EQUALLY SPLIT AMONG THE CE LEVELS

As shown in Fig. 7, given the fact that the device density is low, preamble collisions are not severe such that the collisions can be resolved by retransmissions. Since the preamble repetitions are less in the CE-0, the time-correlated interference between the repeated preambles is low in the CE-0 compared
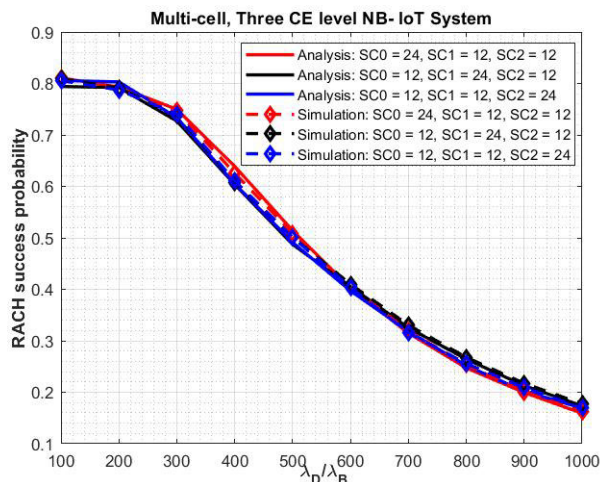
FIGURE 7. RA success probability for different combination of subcarrier allocation, when the device distribution among the CE levels is CE-0: CE-1: CE-2 = 1:1:1.

to higher CE levels (CE-1 and CE-2). Hence, allocating more SCs to CE-0 yields to a higher system success probability. On the other hand, under congestion (when $RP_G \leq 1$), allocating more resources to the CE-1 is favored as it results in better system success probability. This is because the access failure in the CE-0 is mainly due to preamble collision only as the signal outage probability is very low. Hence, reducing the resources to CE-0 will not drastically reduce the number of devices accessing the network as the devices will be able to access the network through the retransmission mechanism. But for CE-1 and CE-2, access failure is due to the combined effect of preamble collision and signal outage. Hence, reducing the resources will result in more unsuccessful RA attempts in the CE-1 and CE-2, compare to the CE-0 due to the extra impact of signal outage. To choose between CE-1, and CE-2, allocating more resources to CE-1 is preferred over CE-2 as the CE-1 uses lesser resources due to the lower number of preamble repetitions allowed in CE-1.

2) ONE CE LEVEL IS MORE CROWDED COMPERED TO THE OTHERS

From the study depicted in Fig. 8 - 10, it is revealed that when $RP_G > 1$, allocating more SCs to a CE level that has the larger number of devices is favored as it yields to the better system success probability. It is a straightforward result, since providing more SCs to a CE-level which has a larger number of contending devices offloads the traffic in that CE-level, which in turn, increases the system success probability. The interesting result is when $RP_G \leq 1$. In that case, the proper choice seems to be to allocate more SCs to the CE-level which has a lower number of devices as shown in Fig. 8 - 10. This is because, when $RP_G \leq 1$, the network congestion is high such that with the given resources collisions cannot be resolved. In this situation, allocating more resources to a CE level with lower device density can yield to a better system success probability. Again, also in this scenario, to choose between two CE-levels with the same number of devices,
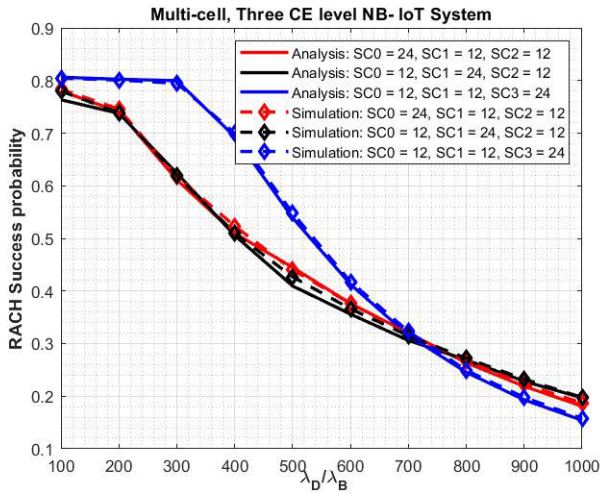
FIGURE 8. RA success probability for different combination of subcarrier allocation, when the device distribution among the CE levels is CE-0: CE-1: CE-2 = 1:1:2.

FIGURE 9. RA success probability for different combination of subcarrier allocation, when the device distribution among the CE levels is CE-0: CE-1: CE-2 = 1:2:1.
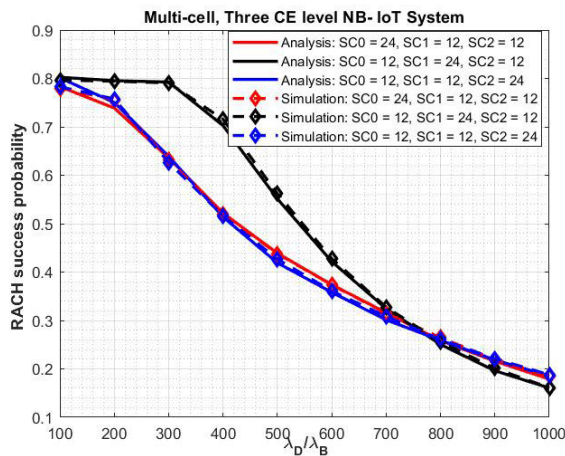
the allocation of more SCs to the lower CE-level is preferred, as it consumes lesser resources due to the lower number of preamble repetitions.

VI. ADAPTIVE RANDOM ACCESS MECHANISM-ADAM
In this section, we exploit the outcome from the performance analysis to propose an ADAptive RA Mechanism, called hereafter ADAM. The mechanism updates the configurable access parameters in each TTI based on the value of the RP_b parameter. In the ADAM, the AB factor, back-off window size, and (re)-transmission number that are send to the devices by the base station in every TTI (in order to tune their RA process) are the ones that corresponds to the best match (closest value) of RP_b estimated by the base station in each TTI to the RP_b values those listed in Table 4.

The method to calculate RP_b value for RA attempts in a TTI is as follows. Let there be H_b RAO within a TTI in the CE-level b, and L_{TTI} be the number of TTI transmission from the base station within T_{max}. The H_b, and L_{TTI}
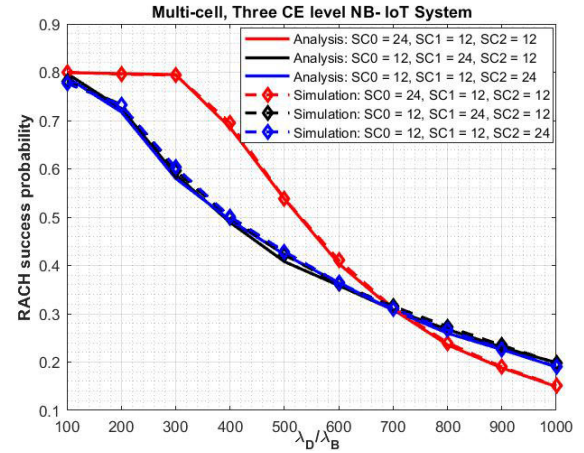

FIGURE 10. RA success probability for different combination of subcarrier allocation, when the device distribution among the CE levels is CE-0: CE-1: CE-2 = 2:1:1.

TABLE 5. The proposed allocation of subcarriers for a different distribution of devices in the CE-levels.

Device distribution among CE levels	RP_G	SC0	SC1	SC2
Equal distribution of devices	>1	24	12	12
	≤ 1	12	24	12
CE-0 is crowded than CE-1 and CE-2	>1	24	12	12
	≤ 1	12	24	12
CE-1 is crowded than CE-0 and CE-2	>1	12	24	12
	≤ 1	24	12	12
CE-2 is crowded than CE-0 and CE-1	>1	12	12	24
	≤ 1	24	12	12

are determined using:

$$H_b = \left\lfloor \frac{D_{TTI}}{T_b} \right\rfloor \tag{28}$$

$$L_{TTI} = \left\lfloor \frac{T_{max}}{D_{TTI}} \right\rfloor \tag{29}$$

Let's consider the start of the jth TTI, and the average number of device in a unit area, performing the RA in the future TTIs j till L_{TTI} is given by $\lambda_{D_b,TTI}^{AB,BW}(j)$:

$$\lambda_{D_b,TTI}^{AB,BW}(j) = \sum_{i=(j-1 \times H_b)+1}^{L_{TTI} \times H_b} \lambda_{D_b}^{AB}(i) + \lambda_{D_b,TTI}^{'AB,BW}(i) \tag{30}$$

where $\lambda_{D_b,TTI}^{'AB,BW}(i)$ is given by:

$$\lambda_{D_b,TTI}^{'AB,BW}(i) = \sum_{k=1}^{j-1 \times H_b} \psi_{k,i-1}^b \lambda_{D_b}^{AB}(k)(1 - P'_{RAO}(Rep,k,b)) \tag{31}$$

Here, $\psi_{(k,i-1)}^b$ is the probability that an NB-IoT device that collided in the kth RAO will fall in the transmission window of the ith RAO for its next RA attempt, and it is given in (21).

$$\lambda_{D_b,TTI}^{AB,BW}(j) = \sum_{n=U_j}^{U_j+H_b} \lambda_{D_b}^{AB,BW}(n) \tag{32}$$

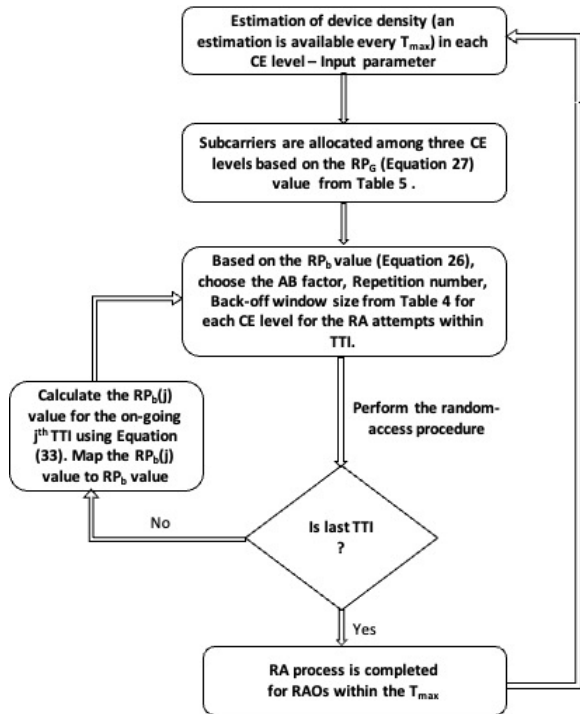$$RP_b(j) = \frac{(L_{TTI} - j) \times H_b \times SC_b \times \lambda_B}{\lambda_{D_b,TTI}^{AB,BW}(j)} \tag{33}$$

FIGURE 11. A flow chart of the proposed adaptive RA mechanism.

FIGURE 12. RA success probability with non-adaptive and adaptive RA mechanism applied on single CE- level (CE-0) NB-IoT system in a multi-cell scenario.

FIGURE 13. RA success probability with non-adaptive and adaptive RA mechanism applied on single CE- level (CE-1) NB-IoT system in a multi-cell scenario.

More precisely, $RP_b(j)$ is a ratio of total amount of the RA resources available for the future transmission (denoted as $(L_{TTI} - j) \times H_b \times SC_b$) to the sum of average number of the devices contending for the RA in the future TTIs (denoted as $\lambda_{D_b,TTI}^{AB,BW}(j)/\lambda_B$) in a single Voronoi cell.

A flow chart of the ADAM mechanism is shown in Fig. 11. The main input in the beginning of each T_{max} period is an estimation of the device density for each CE level. The RP_G value is calculated from the given device density and the total number of subcarriers available to the NB-IoT system. This RP_G value, and the device distribution among the CE levels are used as index to Table 5 in order to define the best allocation of SCs to the CE levels. Then a loop process for each TTI is performed. In the beginning of jth TTI, the $RP_b(j)$ value is updated (depending on the updated value λ_{D_b}, and the remaining number of RAOs within the T_{max}) and the values of AB factor, back-off window, and (re)-transmission number are selected for each CE level (w.r.t. RP_b in Table 4). The base station broadcast the selected values at the beginning of each TTI, to tune the RA procedure within the TTI.

The complexity of the above mentioned algorithm is defined by the initial complexity of creating the look-up table, i.e., Table 4 (Section V.B.5). Hence, the upper bound of complexity is given by the product of the maximum number of states of $ABfactor$, BW, and TM. With the given Table 4, the base station chooses the optimal combination of $ABfactor$, BW, and TM based on the $RP_b(j)$ value which is calculated using (33). The complexity of calculating the $RP_b(j)$ is constant, O(1), since $RP_b(j)$ calculation involves one summation operation, three multiplication operations, and one division operation.

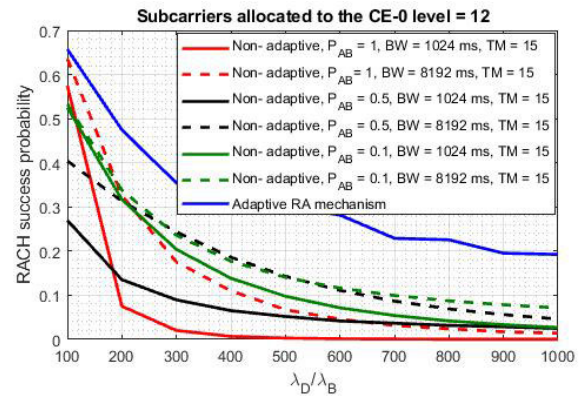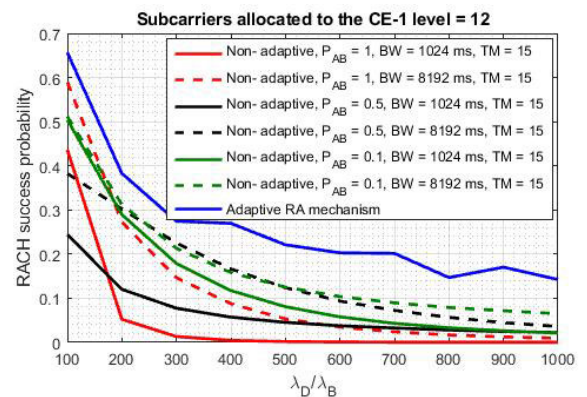## A. ADAM's PERFORMANCE AT CE LEVEL

For the quantification of the adaptive RA performance for CE-0, CE-1, and CE-2, the one-shot arrival process for the attempting devices is considered. Fig. 12 – 14, include the simulation results. As it can be observed, in all the three CE-levels, the proposed adaptive RA mechanism outperforms the conventional non-adaptive RA mechanism in terms of system success probability. This is because, at the beginning of each TTI, the base station broadcast the optimal combination RAN parameters (chosen from Table 4) for RA attempts within the TTI to maximize the system success probability. Also, we can observe that for the same amount of RA resources, the system success probability of a higher CE level is less than the lower CE level. This is because, at higher CE-levels, the number of repetitions is higher, compared to lower CE levels, which in turn increases the time-correlated interference between the preambles in the higher CE levels. This interference decreases the detection probability of the transmitted preambles at the base station.

B. ADAM's PERFORMANCE AT SYSTEM LEVEL

In this section, we applied the adaptive RA mechanism on the three CE-level NB-IoT systems as shown in Fig. 15. We choose the scenario where the devices are distributed equally over the three CE levels, and the corresponding SC

TABLE 6. The mean relative percentage error between analytical and simulation results for adaptive and non-adaptive RA.

| RAN values	Relative percentage error (%) Single CE level (Subcarrier allocation per CE = 12)			Relative percentage error (%) Three CE-level
	P_{Access}			P_{Access}
	CE-0	CE-1	CE-2	Three CE level
Non-adaptive: $p_{AB} = 1$, $BW = 1024$ ms, $TM = 15$	0.89	1.04	1.06	0.92
Non-Adaptive: $p_{AB} = 1$, $BW = 8192$ ms, $TM = 15$	1.03	0.98	1.05	1.04
Non-Adaptive: $p_{AB} = 0.5$, $BW = 1024$ ms, $TM = 15$	1.08	1.02	1.17	0.87
Non-Adaptive: $p_{AB} = 0.5$, $BW = 8192$ ms, $TM = 15$	0.98	1.15	1.06	0.92
Non-Adaptive: $p_{AB} = 0.1$, $BW = 1024$ ms, $TM = 15$	1.07	0.93	1.02	0.83
Non-Adaptive: $p_{AB} = 0.1$, $BW = 8192$ ms, $TM = 15$	0.92	1.09	1.32	1.09
Adaptive RA mechanism	0.96	0.88	1.06	0.85
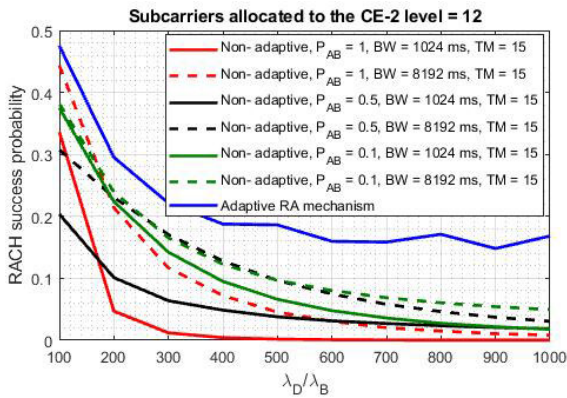
FIGURE 14. RA success probability with non-adaptive and adaptive RA mechanism applied on single CE- level (CE-2) NB-IoT system in a multi-cell scenario.
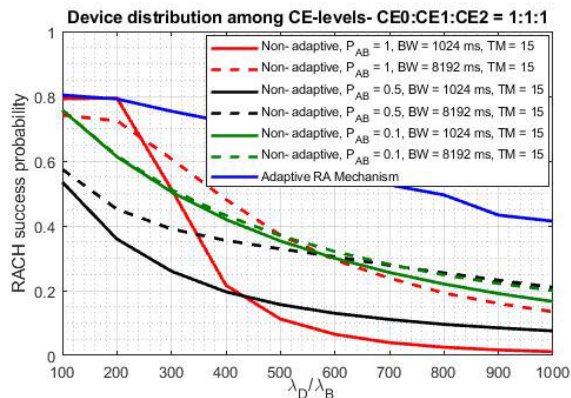

FIGURE 15. RA success probability with non-adaptive and adaptive RA mechanism applied on a three CE level NB-IoT system in a multi-cell scenario.

FIGURE 16. Performance comparison of ADAM on a three CE level NB-IoT system in a multi-cell scenario.

distribution is applied based on our results listed in Table 5. Again, at the beginning of each TTI, the values of back-off window size, AB factor, and (re)-transmission number for RA attempts within the TTI for each CE levels is updated based on the RP value listed in Table 4. Similar to single CE level system, the adaptive RA mechanism yields to better system success probability compared to legacy RA procedure for any value of λ_D/λ_B in the NB-IoT system with three CE levels.

As we can observe from Fig. 15, non-adaptive access with a high value of P_{AB}, and a low value of BW yields results that are comparable with the proposed adaptive RA mechanism when the network congestion is low. This is because when
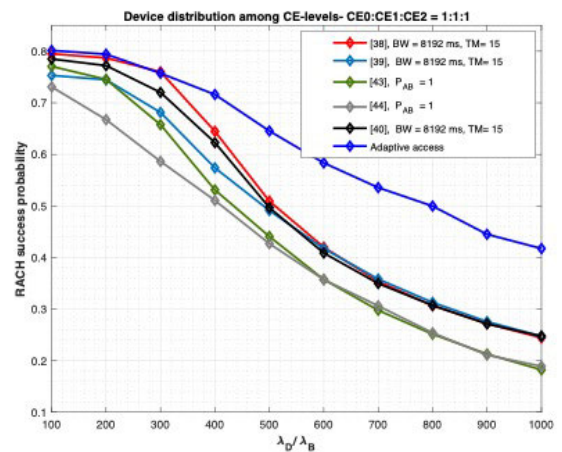
the network is less congested, a high value of P_{AB}, and a low value of BW provide the maximum RA opportunities for the devices. Based on studies in Section V.B, when the network is congested, it is preferred to have a low value of P_{AB}, and a high value of BW to spread the RA attempts from the devices. Even with this choice, using constant values of P_{AB}, and BW throughout all RA slots cannot provide optimal performance as shown in Fig. 15. Since the ADAM optimally configures the configurable access parameter in each TTI based on the estimated device densities in a RA slot, the ADAM outperforms all the combinations of non-adaptive RA mechanisms especially when the network is congested.

The accuracy of analysis expressed using mean relative percentage error as given in Table 6.

C. PERFORMANCE COMPARISON WITH ADAM AT SYSTEM LEVEL

Finally, we compare the performance of the proposed adaptive RA mechanism with similar works, and the result is shown in Fig. 16. In [38]–[40], dynamic access barring mechanisms based on the traffic load is presented. The ADAM outperforms the [38]–[40] because the effect of the back-off window, and retransmission is not considered in the literature. The ADAM also yields better results compared [43], [44] as the effect of the access barring factor is not considered in the literature. Moreover, the optimized values for the

RAN parameters in [43], [44] are set at the beginning of the RA procedure, and hence these parameters are not adaptive according to the time varying traffic in each RA slot.

VII. CONCLUSION

In this paper, we have proposed an adaptive random access mechanism for NB-IoT systems to maximize the access success probability for a given amount of resources. Assuming a multi-cell NB-IoT system with three coverage enhancement layers (CE-Levels) and SINR-based collisions, we developed an analytical model of the NB-IoT random access procedure and derived the expressions for the access success probability and access latency. Building upon the analytical model, 3GPP compliant configurations have been extracted to compose a look-up table with the values that maximise access success probability. A key outcome from our study is the need for holistic decision making mechanisms at the RAN level to address the massiveness of the devices in NB-IoT systems. Indeed, solutions that focus on adapting a single parameter (or a small subset of them) ignore the performance potential that a combined selection of the values can provide. Our performance analysis revealed that by combining in a controlled way the appropriate values for the involved configurable access parameters yields impressive performance improvement.

APPENDIX A

The derivation of the Laplace Transform of the inter-cell interference is given below. The process is based on many related works ([50], [71], [72]) that have revealed the steps for the derivation of the quantity $L_{inter}(\frac{-m\gamma_{th}\sigma^2}{\rho})$.

$$L_{inter}(\frac{-m\gamma_{th}\sigma^2}{\rho})$$

$$\stackrel{a}{=} E_{F_{Inter}}[\prod_{x_i \in F_{Inter}} E_{P_i}E_{h_i}[e^{-(-sP_i(\sum_{k=1}^{m} h_i^k)||x_i||^{-a})}]]$$

$$\stackrel{b}{=} exp(-2\pi \frac{\lambda_D}{SC} \int_{(\frac{P}{\rho})^{\frac{1}{\alpha}}}^{\infty} E_P E_h[1 - e^{-sP(\sum_{k=1}^{m} h^k)x^{-\alpha}}]xdx)$$

$$\stackrel{c}{=} exp(-2\pi \frac{\lambda_D}{SC} \int_{(\frac{P}{\rho})^{\frac{1}{\alpha}}}^{\infty} E_P[1 - \frac{1}{(1 + sPx^{-\alpha})^m}]xdx)$$

$$\stackrel{d}{=} exp(-2(\gamma_{th})^{\frac{2}{\alpha}}\lambda_{pream} \int_{\gamma_{th}^{\frac{-1}{\alpha}}}^{\infty} 1 - \frac{1}{(1 + y^{-\alpha})^m}ydy)$$

where F_{inter} represents the set of active MTC interfering devices, $s = (\frac{-m\gamma_{th}\sigma^2}{\rho})$, P_i is the transmit power of the ith interfering MTC device, x_i and h_i are distance and channel power gain from the ith interfering MTC device to the base station. $E_x[\cdot]$ denoted the expected value of the random variable x. The step (a) is obtained by following the independence between h_i, and P_i and by taking average with respect to $h_i^1 \ldots . h_i^m$, step (b) is obtained from the PGFL (probability generating functional) of the HPPP, step (c) by taking the Laplace Transform of h, and step (d) is deduced by changing

the variables $(\frac{x}{sP})^{\frac{1}{\alpha}}$ to y and by representing the moment of the transmitted power $E[P^{\frac{2}{\alpha}}]$ by $\rho^{\frac{2}{\alpha}}/\pi \lambda_B$ [70].

APPENDIX B

The derivation of the Laplace Transform of the intra-cell interference is given below.

$$L_{intra}(\frac{-m_b\gamma_{th}\sigma^2}{\rho_b}) = E[exp(-s \sum_{j\in F_{Intra}} \rho \sum_{k=1}^{m} h_j^k)]$$

$$= E[\prod_{j=1}^{Z} exp(-s\rho \sum_{k=1}^{m} h_j^k)] \stackrel{a}{=} (\frac{1}{1 + s\rho})^{mZ}$$

where F_{intra} represents the set of active intra-cell interfering MTC devices. The step (a) is obtained by taking average with respect to $h_j^1 \ldots . h_j^m$.

REFERENCES

[1] OneM2M. *Standards For M2M and the Internet Of Things*. Accessed: Jan. 3, 2020. [Online]. Available: http://www.onem2m.org

[2] *Service Requirements for Machine-Type Communications (MTC)*, Standard 3GPP TS 22.368, V16.0.0, Jul. 2020. [Online]. Available: http://www.3gpp.org

[3] *Machine-to-Machine Communications (M2M); M2M Service Requirements*, Standard ETSI TS 102 689, V2.1.1, Aug. 2013. [Online]. Available: https://www.etsi.org

[4] *Security Aspects of Machine-Type Communications (MTC) and Other Mobile Data Applications Communications Enhancements*, Standard 3GPP TS 33.187, V15.1.0, Jun. 2018. [Online]. Available: http://www.3gpp.org

[5] *Machine-to-Machine communications (M2M); Functional Architecture*, Standard ETSI TS 102 690, V2.1.1, 2013. [Online]. Available: https://www.etsi.org

[6] *Study on Alternatives to E.164 for Machine-Type Communications (MTC)*, Standard 3GPP TS 22.988, V15.0.0, Jul. 2019. [Online]. Available: http://www.3gpp.org

[7] *TSP Interface Protocol Between the MTC Interworking Function (MTC-IWF) and Service Capability Server (SCS)*, Standard 3GPP TS 29.368, V15.1.0, Sep. 2019. [Online]. Available: http://www.3gpp.org

[8] U. Raza, P. Kulkarni, and M. Sooriyabandara, "Low power wide area networks: An overview," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 2, pp. 855–873, Jan. 2017, doi: 10.1109/COMST.2017.2652320.

[9] *Study on Architecture Enhancements for Cellular Internet of Things (CIoT)*, Standard 3GPP TR 23.720, V13.0.0, Mar. 2016. [Online]. Available: http://www.3gpp.org

[10] *Physical Layer Procedures*, Standard 3GPP TS 36.213, V15.4.0, Jan. 2019. [Online]. Available: http://www.3gpp.org

[11] *NR: Physical Layer Procedures for Control*, Standard 3GPP TS 38.213, V16.0.0, Dec. 2019. [Online]. Available: http://www.3gpp.org

[12] *Release 14 Description—Summary of Rel-14 Work Items*, Standard 3GPP TR 21.914, V14.0.0, Jun. 2018. [Online]. Available: http://www.3gpp.org

[13] P. Salva-Garcia, J. M. Alcaraz-Calero, Q. Wang, J. B. Bernabe, and A. Skarmeta, "5G NB-IoT: Efficient network traffic filtering for multi-tenant IoT cellular networks," *Secur. Commun. Netw.*, vol. 2018, pp. 1–21, Dec. 2018.

[14] 5G Americas. (Jul. 2019). *5G—The Future of IoT*. Accessed: Jan. 14, 2020. [Online]. Available: https://www.5gamericas.org/wp-content/uploads/2019/07/

[15] 5G Americas. *IoT Deployment*. Accessed: Nov. 23, 2019. [Online]. Available: https://www.5gamericas.org/resources/deployments/deployments-iot/

[16] C. Bockelmann, N. Pratas, H. Nikopour, K. Au, T. Svensson, C. Stefanovic, P. Popovski, and A. Dekorsy, "Massive machine-type communications in 5G: Physical and MAC-layer solutions," *IEEE Commun. Mag.*, vol. 54, no. 9, pp. 59–65, Sep. 2016, doi: 10.1109/MCOM.2016.7565189.

[17] A. Rajandekar and B. Sikdar, "A survey of MAC layer issues and protocols for machine-to-machine communications," *IEEE Internet Things J.*, vol. 2, no. 2, pp. 175–186, Apr. 2015, doi: 10.1109/JIOT.2015.2394438.

[18] N. Jiang, Y. Deng, A. Nallanathan, X. Kang, and T. Q. S. Quek, "Analyzing random access collisions in massive IoT networks," *IEEE Trans. Wireless Commun.*, vol. 17, no. 10, pp. 6853–6870, Oct. 2018, doi: 10.1109/TWC.2018.2864756.

[19] C. Bockelmann, N. K. Pratas, G. Wunder, S. Saur, M. Navarro, D. Gregoratti, G. Vivier, E. De Carvalho, Y. Ji, C. Stefanovic, and P. Popovski, "Towards massive connectivity support for scalable mMTC communications in 5G networks," *IEEE Access*, vol. 6, pp. 28969–28992, 2018, doi: 10.1109/ACCESS.2018.2837382.

[20] S. Saur and M. Centenaro, "Radio access protocols with multi-user detection for URLLC in 5G," in *Proc. 23th Eur. Wireless Conf.*, May 2017, pp. 1–6.

[21] S. Saur, A. Weber, and G. Schreiber, "Radio access protocols and preamble design for machine type communications in 5G," in *Proc. 49th Asilomar Conf. Signals, Syst. Comput.*, Nov. 2015, pp. 8–12, doi: 10.1109/ACSSC.2015.7421049.

[22] T. P. C. de Andrade, C. A. Astudillo, L. R. Sekijima, and N. L. S. da Fonseca, "The random access procedure in long term evolution networks for the Internet of Things," *IEEE Commun. Mag.*, vol. 55, no. 3, pp. 124–131, Mar. 2017, doi: 10.1109/MCOM.2017.1600555CM.

[23] N. K. Pratas, S. Pattathil, C. Stefanovic, and P. Popovski, "Massive machine-type communication (mMTC) access with integrated authentication," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2017, pp. 1–6, doi: 10.1109/ICC.2017.7997466.

[24] M. Mohammadkarimi, M. A. Raza, and O. A. Dobre, "Signature-based nonorthogonal massive multiple access for future wireless networks: Uplink massive connectivity for machine-type communications," *IEEE Veh. Technol. Mag.*, vol. 13, no. 4, pp. 40–50, Dec. 2018, doi: 10.1109/MVT.2018.2869425.

[25] D. Sijabat, R. Harwahyu, and R.-G. Cheng, "Energy-efficiency of RACH-based small data transmission scheme in LTE networks," in *Proc. 40th Int. Conf. Telecommun. Signal Process. (TSP)*, Jul. 2017, pp. 106–109, doi: 10.1109/TSP.2017.8075946.

[26] S.-M. Oh and J. Shin, "An efficient small data transmission scheme in the 3GPP NB-IoT system," *IEEE Commun. Lett.*, vol. 21, no. 3, pp. 660–663, Mar. 2017, doi: 10.1109/LCOMM.2016.2632128.

[27] F. Monsees, M. Woltering, C. Bockelmann, and A. Dekorsy, "A potential solution for MTC: Multi-carrier compressed sensing multi-user detection," in *Proc. 49th Asilomar Conf. Signals, Syst. Comput.*, Nov. 2015, pp. 18–22, doi: 10.1109/ACSSC.2015.7421051.

[28] H. Seo, J.-P. Hong, and W. Choi, "Low latency random access for sporadic MTC devices in Internet of Things," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 5108–5118, Jun. 2019, doi: 10.1109/JIOT.2019.2896620.

[29] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1204–1216, Jul. 2000, doi: 10.1109/18.850663.

[30] S. Zhang, S. C. Liew, and P. P. Lam, "Hot topic: Physical-layer network coding," in *Proc. 12th Annu. Int. Conf. Mobile Comput. Netw. (MobiCom)*, 2006, pp. 358–365, doi: 10.1145/1161089.1161129.

[31] M. Navarro, D. Gregoratti, A. Pastore, E. Kosmatos, P. Demestichas, and S. Pfletschinger, "PLNC decoding: Enabler for massive MTC in 5G networks," in *Proc. 15th Int. Symp. Wireless Commun. Syst. (ISWCS)*, Aug. 2018, pp. 1–6, doi: 10.1109/ISWCS.2018.8491212.

[32] Z. Sun, L. Yang, J. Yuan, and D. W. K. Ng, "Physical-layer network coding based decoding scheme for random access," *IEEE Trans. Veh. Technol.*, vol. 68, no. 4, pp. 3550–3564, Apr. 2019, doi: 10.1109/TVT.2019.2896617.

[33] T. Lv, Z. Lin, P. Huang, and J. Zeng, "Optimization of the energy-efficient relay-based massive IoT network," *IEEE Internet Things J.*, vol. 5, no. 4, pp. 3043–3058, Aug. 2018, doi: 10.1109/JIOT.2018.2829827.

[34] S. Malathy, V. Porkodi, A. Sampathkumar, M. H. D. N. Hindia, K. Dimyati, V. Tilwari, F. Qamar, and I. S. Amiri, "An optimal network coding based backpressure routing approach for massive IoT network," *Wireless Netw.*, vol. 26, no. 5, pp. 3657–3674, Jul. 2020.

[35] E. Datsika, A. Antonopoulos, N. Zorba, and C. Verikoukis, "Cross-network performance analysis of network coding aided cooperative outband D2D communications," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 3176–3188, May 2017, doi: 10.1109/TWC.2017.2675887.

[36] T.-M. Lin, C.-H. Lee, J.-P. Cheng, and W.-T. Chen, "PRADA: Prioritized random access with dynamic access barring for MTC in 3GPP LTE-A networks," *IEEE Trans. Veh. Technol.*, vol. 63, no. 5, pp. 2467–2472, Jun. 2014, doi: 10.1109/TVT.2013.2290128.

[37] M. S. Ali, E. Hossain, and D. I. Kim, "LTE/LTE-A random access for massive machine-type communications in smart cities," *IEEE Commun. Mag.*, vol. 55, no. 1, pp. 76–83, Jan. 2017, doi: 10.1109/MCOM.2017.1600215CM.

[38] S. W. H. Shah, A. T. Riaz, and K. Iqbal, "Congestion control through dynamic access class barring for bursty MTC traffic in future cellular networks," in *Proc. Int. Conf. Frontiers Inf. Technol. (FIT)*, Dec. 2018, pp. 176–181, doi: 10.1109/FIT.2018.00038.

[39] Y. Miao, Y. Tian, J. Cheng, M. S. Hossain, and A. Ghoneim, "RADB: Random access with differentiated barring for latency-constrained applications in NB-IoT network," *Wireless Commun. Mobile Comput.*, vol. 2018, pp. 1–9, Jan. 2018, doi: 10.1155/2018/6210408.

[40] L. Wang, W. Wang, X. Hu, and T. Xie, "Optimization of large-scaled random access congestion control oriented to narrow band Internet of Things," *J. Phys., Conf. Ser.*, vol. 1570, Jun. 2020, Art. no. 012089.

[41] S.-Y. Lien, T.-H. Liau, C.-Y. Kao, and K.-C. Chen, "Cooperative access class barring for machine-to-machine communications," *IEEE Trans. Wireless Commun.*, vol. 11, no. 1, pp. 27–32, Jan. 2012, doi: 10.1109/TWC.2011.111611.110350.

[42] C. A. Astudillo, F. H. S. Pereira, and N. L. S. da Fonseca, "Probabilistic retransmissions for the random access procedure in cellular IoT networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2019, pp. 1–7.

[43] R. Harwahyu, R.-G. Cheng, W.-J. Tsai, J.-K. Hwang, and G. Bianchi, "Repetitions versus retransmissions: Tradeoff in configuring NB-IoT random access channels," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 3796–3805, Apr. 2019, doi: 10.1109/JIOT.2019.2891366.

[44] R. Harwahyu, R.-G. Cheng, C.-H. Wei, and R. F. Sari, "Optimization of random access channel in NB-IoT," *IEEE Internet Things J.*, vol. 5, no. 1, pp. 391–402, Feb. 2018, doi: 10.1109/JIOT.2017.2786680.

[45] E. Migabo, K. Djouani, and A. Kurien, "An energy-efficient and adaptive channel coding approach for narrowband Internet of Things (NB-IoT) systems," *Sensors*, vol. 20, no. 12, p. 3465, Jun. 2020, doi: 10.3390/s20123465.

[46] C. Yu, L. Yu, Y. Wu, Y. He, and Q. Lu, "Uplink scheduling and link adaptation for narrowband Internet of Things systems," *IEEE Access*, vol. 5, pp. 1724–1734, 2017, doi: 10.1109/ACCESS.2017.2664418.

[47] P. Andres-Maldonado, P. Ameigeiras, J. Prados-Garzon, J. J. Ramos-Munoz, J. Navarro-Ortiz, and J. M. Lopez-Soler, "Analytic analysis of narrowband IoT coverage enhancement approaches," in *Proc. Global Internet Things Summit (GIoTS)*, Jun. 2018, pp. 1–6.

[48] R. Boisguene, S.-C. Tseng, C.-W. Huang, and P. Lin, "A survey on NB-IoT downlink scheduling: Issues and potential solutions," in *Proc. 13th Int. Wireless Commun. Mobile Comput. Conf. (IWCMC)*, Jun. 2017, pp. 547–551.

[49] B.-Z. Hsieh, Y.-H. Chao, R.-G. Cheng, and N. Nikaein, "Design of a UE-specific uplink scheduler for narrowband Internet-of-Things (NB-IoT) systems," in *Proc. 3rd Int. Conf. Intell. Green Building Smart Grid (IGBSG)*, Apr. 2018, pp. 1–5.

[50] N. Jiang, Y. Deng, M. Condoluci, W. Guo, A. Nallanathan, and M. Dohler, "RACH preamble repetition in NB-IoT network," *IEEE Commun. Lett.*, vol. 22, no. 6, pp. 1244–1247, Oct. 2018, doi: 10.1109/LCOMM.2018.2793274.

[51] *Study on Architecture Enhancements for Cellular Internet of Things (CIoT)*, Standard 3GPP TR 23.720, V13.0.0, Mar. 2016. [Online]. Available: http://www.3gpp.org

[52] *Cellular System Support for Ultra-Low Complexity and Low Throughput Internet of Things*, Standard 3GPP TR 45.820, V13.1.0, Dec. 2018. [Online]. Available: http://www.3gpp.org

[53] M. El Soussi, P. Zand, F. Pasveer, and G. Dolmans, "Evaluating the performance of eMTC and NB-IoT for smart city applications," in *Proc. IEEE 54th Int. Conf. Commun. (ICC)*, Kansas City, MO, USA, May 2018, pp. 1–7, doi: 10.1109/ICC.2018.8422799.

[54] Y.-P. E. Wang, X. Lin, A. Adhikary, A. Grovlen, Y. Sui, Y. Blankenship, J. Bergman, and H. S. Razaghi, "A primer on 3GPP narrowband Internet of Things," *IEEE Commun. Mag.*, vol. 55, no. 3, pp. 117–123, Mar. 2017, doi: 10.1109/MCOM.2017.1600510CM.

[55] S. Narayanan, D. Tsolkas, N. Passas, and L. Merakos, "NB-IoT: A candidate technology for massive IoT in the 5G era," in *Proc. IEEE 23rd Int. Workshop Comput. Aided Model. Design Commun. Links Netw. (CAMAD)*, Sep. 2018, pp. 1–6, doi: 10.1109/CAMAD.2018.8514963.

[56] *Radio Resource Control Protocol Specification*, Standard 3GPP TS 36.331, V16.1.1, Jul. 2020. [Online]. Available: http://www.3gpp.org

[57] O. Liberg, M. Sundberg, E. Wang, J. Bergman, and J. Sachs, *Cellular Internet Of Things-Technologies, Standards, and Performance*. 1st ed. Amsterdam, The Netherlands: Elsevier, 2018.

[58] *Medium Access Control Protocol Specifications*, Standard 3GPP TS 36.321, V16.1.0, Jul. 2020. [Online]. Available: http://www.3gpp.org.

[59] X. Lin, A. Adhikary, and Y.-P. E. Wang, "Random access preamble design and detection for 3GPP narrowband IoT systems," *IEEE Wireless Commun. Lett.*, vol. 5, no. 6, pp. 640–643, Dec. 2016, doi: 10.1109/LWC.2016.2609914.

[60] S. Cho, H. Kim, and G. Jo, "Determination of optimum threshold values for NPRACH preamble detection in NB-IoT systems," in *Proc. 10th Int. Conf. Ubiquitous Future Netw. (ICUFN)*, Jul. 2018, pp. 616–618, doi: 10.1109/ICUFN.2018.8436646.

[61] *Study on Scenarios and Requirements for Next Generation Access Technologies*, Standard 3GPP TS 38.913, V16.0.0, Jul. 2020. [Online]. Available: http://www.3gpp.org.

[62] *Backoff Enhancements for RAN Overload Control*, document ZTE R2-111916, Apr. 2011. [Online]. Available: http://www.3gpp.org.

[63] *RAN Improvements for Machine-type Communications*, Standard 3GPP TR 37.868, V11.0.0, Sep. 2011. [Online]. Available: http://www.3gpp.org.

[64] N. Jiang, Y. Deng, X. Kang, and A. Nallanathan, "Random access analysis for massive IoT networks under a new spatio-temporal model: A stochastic geometry approach," *IEEE Trans. Commun.*, vol. 66, no. 11, pp. 5788–5803, Nov. 2018, doi: 10.1109/TCOMM.2018.2854275.

[65] *Base Station (BS) Radio Transmission and Reception*, Standard 3GPP TS 36.104, V16.5.0, Mar. 2020. [Online]. Available: http://www.3gpp.org

[66] (Jan. 2017). *Coverage Analysis LTE-M Category M1, V1.0*. Accessed: Jul. 15, 2019. [Online]. Available: http://www.coverageanalysisoflte-m.com/

[67] *Service Accessibility*, Standard 3GPP TS 22.011, V13.5.0, Mar. 2016. [Online]. Available: http://www.3gpp.org

[68] L. Tello-Oquendo, I. Leyva-Mayorga, V. Pla, J. Martinez-Bauset, J. R. Vidal, V. Casares-Giner, and L. Guijarro, "Performance analysis and optimal access class barring parameter configuration in LTE-A networks with massive M2M traffic," *IEEE Trans. Veh. Technol.*, vol. 67, no. 4, pp. 3505–3520, Apr. 2018, doi: 10.1109/TVT.2017.2776868.

[69] J.-S. Ferenc and Z. Néda, "On the size distribution of Poisson Voronoi cells," *Phys. A, Statist. Mech. Appl.*, vol. 385, no. 2, pp. 518–526, Nov. 2007.

[70] H. ElSawy and E. Hossain, "On stochastic geometry modeling of cellular uplink transmission with truncated channel inversion power control," *IEEE Trans. Wireless Commun.*, vol. 13, no. 8, pp. 4454–4469, Aug. 2014, doi: 10.1109/TWC.2014.2316519.

[71] C.-H. Lee and M. Haenggi, "Interference and outage in Poisson cognitive networks," *IEEE Trans. Wireless Commun.*, vol. 11, no. 4, pp. 1392–1401, Apr. 2012, doi: 10.1109/TWC.2012.021512.110131.

[72] Z. Yazdanshenasan, H. S. Dhillon, M. Afshang, and P. H. J. Chong, "Poisson hole process: Theory and applications to wireless networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7531–7546, Nov. 2016, doi: 10.1109/TWC.2016.2604799.

DIMITRIS TSOLKAS received the Ph.D. degree from the Department of Informatics and Telecommunications, National and Kapodistrian University of Athens (NKUA). He has long experience in research and development and project management, working for a plethora of EC-funded research and innovation actions. He also has teaching experience as a Lecturer and an Assistant Instructor with the Department of Informatics and Telecommunications, NKUA, and the Department of Computer Science and Engineering, University of Ioannina. He is currently a Senior Research Fellow at NKUA. So far, he has published more than 40 articles in peer-reviewed journals, international conferences, and book chapters, while he has served as a reviewer for high-quality journals and magazines. With key expertise in 5G network functions optimization and end-to-end network emulation and testing, his current research interests include wireless/mobile communication systems, radio access networks design and analysis, and user experience management in service provisioning.

NIKOS PASSAS received the Diploma degree (Hons.) from the Department of Computer Engineering, University of Patras, Greece, in 1992, and the Ph.D. degree from the Department of Informatics and Telecommunications, University of Athens, Greece, in 1997. He is currently a member of the Teaching Staff with the Department of Informatics and Telecommunications, University of Athens, where he as a Group Leader of the Green, Adaptive and Intelligent Networking (GAIN) Research Group inside the department. Over the years, he has participated and coordinated a large number of national and European research projects. He has published more than 120 papers in peer-reviewed journals and international conferences and has also published one book and 11 book chapters. He has served as a Guest Editor and Technical Program Committee Member for prestigious magazines and conferences, such as *IEEE Wireless Communications Magazine*, *Wireless Communications and Mobile Computing* journal, IEEE Vehicular Technology Conference, IEEE PIMRC, and IEEE GLOBECOM.

LAZAROS MERAKOS received the Diploma degree in electrical and mechanical engineering from the National Technical University of Athens, Greece, in 1978, and the M.S. and Ph.D. degrees in electrical engineering from the State University of New York, Buffalo, in 1981 and 1984, respectively. He has been on the Faculty of the University of Connecticut (1983–1986), and Northeastern University, Boston (1986–1994). Since 1994, he has been a Professor with the Department of Informatics and Telecommunications, National and Kapodistrian University of Athens (NKUA), where he has been the Scientific Director of the Network Operations and Management Center. Since June 2020, he has been serving as the Chair for the newly founded Department of Digital Industry Technologies, NKUA. He has been the Principal Investigator for NKUA in many EU funded research and development projects in the area of mobile/wireless networks. His research interests include network technologies, services, and applications, where he has authored more than 300 publications in international journals and conferences. He was a recipient of the Guanella Award for the Best Paper presented at the 1994 International Zurich Conference on Mobile Communications and the Best Paper Award in the 2008 IEEE International Symposium on Communications, Control and Signal Processing. He is the Chairman of the Board of the Greek Universities Network, a non-profit organization for the design and development of advanced ICT services for the 25 universities in Greece.

SUBIN NARAYANAN (Graduate Student Member, IEEE) received the B.Tech. degree in electronics and communication engineering and the M.Tech. degree in wireless communications from Cochin University of Science of Technology, Kerala, India, in 2012 and 2015, respectively, and the Ph.D. degree in RF engineering from TU Eindhoven, in 2017. He is currently working as a Researcher with the Department of Informatics and Telecommunication, National and Kapodistrian University of Athens, Greece. His research interests include but are not limited to 5G-IoT, sidelink communications, cellular IoT, and antenna engineering.