Maximizing Connection Density in NB-IoT Networks with NOMA

Shashwat Mishra, Lou Salaun, Chung Shue Chen

To cite this version:

Shashwat Mishra, Lou Salaun, Chung Shue Chen. Maximizing Connection Density in NB-IoT Networks with NOMA. IEEE VTC 2020-Spring, May 2020, Antwerp, Belgium. hal-02571150

HAL Id: hal-02571150

https://hal.archives-ouvertes.fr/hal-02571150

Submitted on 12 May 2020

Maximizing Connection Density in NB-IoT Networks with NOMA

Shashwat Mishra*†, Lou Salaün*, Chung Shue Chen*
*Bell Labs, Nokia Paris-Saclay, 91620 Nozay, France
†Indian Institute of Technology Madras, Chennai, India 600036
Email: ee17s050@smail.iitm.ac.in, lou.salaun@nokia-bell-labs.com, chung_shue.chen@nokia-bell-labs.com

***Abstract*—We address the issue of maximizing the number of connected devices in a Narrowband Internet of Things (NB-IoT) network using non-orthogonal multiple access (NOMA). The scheduling assignment is done on a per-transmit time interval (TTI) basis and focuses on efficient device clustering. We formulate the problem as a combinatorial optimization problem and solve it under interference, rate and sub-carrier availability constraints. We first present the bottom-up power filling algorithm (BU), which solves the problem given that each device can only be allocated contiguous sub-carriers. Then, we propose the item clustering heuristic (IC) which tackles the more general problem of non-contiguous allocation. The novelty of our optimization framework is two-fold. First, it allows any number of devices to be multiplexed per sub-carrier, which is based on the successive interference cancellation (SIC) capabilities of the network. Secondly, whereas most existing works only consider contiguous sub-carrier allocation, we also study the performance of allocating non-contiguous sub-carriers to each device. We show through extensive simulations that non-contiguous allocation through IC scheme can outperform BU and other existing contiguous allocation methods.**

***Index Terms*—NB-IoT, NOMA, Connectivity Maximization, Greedy Heuristic, mMTC, MTCD, 5G**

I. INTRODUCTION

5G is a momentous advance in the area of wireless communication which promises ubiquitous connectivity, high device density and lightning-fast data rates. Of the many important use cases that 5G serves, massive machine type communications (mMTC) has gained huge traction in recent times. mMTC devices are characterized by high device density per unit area, limited data rates, moderate reaction times and small transmit power. These devices are usually cost-effective and deployed in huge number to serve a variety of settings like health-care, traffic monitoring, smart buildings, weather forecast networks and so on. mMTC devices belong to a bigger class of devices generally referred to as machine type communications devices (MTCDs), characterised as one or more communicating entities that do not necessarily need human interaction. mMTC has been already developed as part of 3GPP Release 13 low power wide area (LPWA) technologies, which also includes NB-IoT. 3GPP specifies Cat-NB1 and Cat-NB2 [1] for exceptionally deep coverage and extremely low power applications. NB-IoT is a suitable key technology for 5G mMTC and can operate inside the 5G NR frequency bands similar to LTE today. NB-IoT is a cellular LPWA network technology that typically operates using a bandwidth of 180 kHz. One physical resource block (PRB) of 180 kHz bandwidth has a sub-carrier spacing of 15 kHz in the downlink and an option of using either 3.75 kHz or 15 kHz in the uplink [2]. Uplink transmission also supports multi-tone operation in which a single device can be assigned a contiguous bond of multiple sub-carriers for transmission [2]. NB-IoT uses coverage enhancement techniques such as power boosting in downlink or sub-frame repetition in both uplink and downlink to meet the additional 20 dB maximum coupling loss (MCL) requirement. mMTC focuses on providing connectivity to a large number of devices that transmit small amounts of data at irregular intervals. Taking a look at the numbers, we see that the standards promise a connection density of 10^6 devices/km^2, battery life of up to 10 years and low data rates of the order of 10 kbps [2]. We, therefore, see a pressing need to maximize the number of connected devices per unit area. There is limited literature that addresses this problem of connectivity maximization.

Papers [3]–[5] point out the great untapped potential in using NOMA for wireless access. The issues with uplink scheduling for NB-IoT have been discussed in [6]. The authors of [7] present a framework for connectivity maximization in a device pool consisting of mMTC and ultra-reliable low latency (URLLC) devices. The solution presented is limited by sub-optimal device pairing assumptions. That is, only one URLLC device and one mMTC device can be served on the same sub-carrier. In addition, the URLLC signal is always decoded first and then SIC is performed to decode the mMTC signal, whereas the possibility of having an mMTC-mMTC or URLLC-URLLC device pairing on the same sub-carrier is not explored. Furthermore, each device can only be served on a contiguous set of sub-carriers. In order to enable massive connectivity in 5G and beyond 5G(B5G) networks, it is important to develop non-contiguous sub-carrier allocation schemes. Findings of [8], [9] reflect the efficiency of power domain NOMA for increasing the connection density. In [8], the authors have relaxed the SIC limitations, thus allowing any number of superimposed signals from devices to be successfully decoded, and shown that the system can support more than 5 times the number of connected devices than traditional schemes. However, it should be noted that superposing a huge number of devices in any NOMA system is infeasible in practice due to error propagation issues in SIC and decoding complexity.

Reference [10] suggests a user clustering based approach where devices from two different device classes are bunched together into small groups and then considered for allocation, however, the emphasis of this work is to maximize the total throughput of the network by optimizing the resource allocation of MTC devices and NOMA clustering. Drawing from the insights of the solutions in existing literature, in this paper we propose a connectivity maximization framework which focuses on forming efficient device clusters through proper power allocation with the following salient contributions:

1) We propose a novel power filling based heuristic, the bottom-up power filling (BU) strategy, that is computationally efficient and can accommodate any kind of SIC constraints imposed by the system. When used with multi-tone operation, this method can support contiguous sub-carrier allocation.
2) We develop an item clustering heuristic (IC) which optimizes the clustering of devices on each sub-carrier. IC performs non-contiguous sub-carrier allocation in the multi-tone uplink.
3) We study and compare the performance of our proposed schemes in contiguous and non-contiguous sub-carrier allocation under NOMA with the state-of-the-art solutions as proposed in [7]. We also compare the performance of our proposed algorithms with orthogonal multiple access (OMA) systems. We show that allocation with NOMA using our proposed IC approach can significantly increase the number of connected devices compared to contiguous approach such as BU and other existing solutions.

The rest of this paper is organized as follows. Section II describes the system model and defines the connectivity maximization problem. In Section III, we propose BU and IC algorithms to solve the problem. We evaluate the performance of our proposed solutions through computer simulations in Section IV. Finally, Section V concludes the paper.

II. SYSTEM MODEL AND PROBLEM FORMULATION

We consider a system having a set of D devices, which is denoted by $\mathcal{D} \triangleq \{1, \dots, D\}$, being served by a base station. We aim to maximize the number of connected devices per TTI in the uplink. Device allocation is performed over the bandwidth of one PRB having S sub-carriers. We denote the set of sub-carriers by $\mathcal{S} \triangleq \{1, \dots, S\}$. We consider NOMA and the system can support up to M devices per sub-carrier, where M is a positive integer. Note that M is a system parameter and predefined according to SIC limitations. The transmit power per device is limited to a threshold of P_d, usually referred to as the power budget of the system. N_0 is the power spectral density of noise over the operating bandwidth. The system has a bandwidth of B_s Hz, equally divided into S sub-carriers with each sub-carrier of bandwidth B Hz. The methods showcased in this paper are very generic and can be applied to address a wide variety of use cases and service scenarios. The only requirement for the schemes to function well is the availability of channel state information for devices. In this paper, we will demonstrate the efficacy of our methods under NB-IoT system for MTCDs. We consider both single-tone and multi-tone operation in the uplink i.e. from the user equipment to the base station, defining $\mathcal{S}_\#$ as the set of possible number of sub-carriers that can be given to a device. The actual collection of sub-carriers given to a device d is denoted by the set S_d. It is evident that $|S_d| \in S_\#$. For example, following 3GPP technical report [2] for this paper, we know that $1, 3, 6$ or 12 sub-carriers can be assigned to each device during multi-tone operation. In such case $\mathcal{S}_\# = \{1, 3, 6, 12\}$.

A. NOMA Framework

As aforementioned, we consider a NOMA based system, where each sub-carrier can accommodate a maximum of M devices. The system performs power domain NOMA. The received signal for $k \leq M$ users transmitting simultaneously on the same sub-carrier with messages $x_1, x_2, \dots, x_k$, with transmit powers $p_1, p_2, \dots, p_k$ respectively can be written as:

$$y = \sqrt{g_1 p_1} x_1 + \sqrt{g_2 p_2} x_2 + \ldots + \sqrt{g_k p_k} x_k + \sigma \qquad (1)$$

where g_k represents the magnitude of the channel gain, consisting of both the distance dependent path-loss and the small scale fading, for user k, and σ is the additive white Gaussian noise. Assuming that the received power from x_k satisfies $g_k p_k > g_i p_i$, for $i \in \{1, \dots, k-1\}$, we begin with decoding x_k first. We denote the signal-to-interference-plus-noise ratio (SINR) at receiver for x_k over a bandwidth B by γ_k, which can be expressed as:

$$\gamma_k = \frac{g_k p_k}{\sum\limits_{i=1}^{k-1} g_i p_i + N_0 B}. \qquad (2)$$

Here $\sum_{i=1}^{k-1} g_i p_i$ is the interference power from the other users. As is common notion with power domain NOMA, the strong user can decode both its own signal and the signal for the interferer while the weak user treats the other user's signal as noise and can decode only its own signal. So, at the base station, x_k is first decoded, then subtracted from the superposed signal before decoding for $\{x_1, \dots, x_{k-1}\}$. Iteratively, the received SINR for x_1 can, therefore, be represented as:

$$\gamma_1 = \frac{g_1 p_1}{N_0 B}. \qquad (3)$$

B. Constraints for NOMA Framework

There are three primary constraints to be considered while using the NOMA scheme. We shall denote the set of users not yet served by $\mathcal{D}'$. First, each device can only transmit with a total power budget that does not exceed the threshold P_d. The transmit power of device d on a single sub-carrier, $n \in S_d$, is denoted by $p_{d,n}$. Therefore,

$$p_d = \sum_{n \in S_d} p_{d,n} \qquad (4)$$

where p_d is the total power consumed by a device on the set of all sub-carriers assigned to it. Additionally,

$$0 \leq p_d \leq P_d, \qquad \forall d \in \mathcal{D} \setminus \mathcal{D}'. \qquad (5)$$

Secondly, each device must achieve a minimum critical rate to remain viable in the network. The achievable rate for an arbitrary device d on sub-carrier n can be written as:

$$r_{d,n} = B \log_2 \left(1 + \frac{g_d p_{d,n}}{I_{d,n} + N_0 B} \right), \qquad (6)$$

$I_{d,n}$ is the interference on sub-carrier n in S_d for device d caused by all the other devices occupying it, given by:

$$I_{d,n} \triangleq \sum_{d' \in D \setminus \{d\}} g_{d'} p_{d',n} \tag{7}$$

where d' is such that $g_{d'} p_{d',n} < g_d p_{d,n}$, which reflects the decoding order for SIC. We consider $g_{d,n} = g_d$ for all n since the channel response is not frequency selective over the PRB, where $g_{d,n}$ is the magnitude of the channel gain of device d on sub-carrier n. In (8), we define $\boldsymbol{I}$ as the vector containing net interference on all sub-carriers, where each entry is calculated following (7).

$$\boldsymbol{I} \triangleq (I_{d,1}, \ldots, I_{d,S}) \,. \tag{8}$$

We denote the total achievable rate for device d by r_d such that:

$$r_d = \sum_{n \in S_d} r_{d,n}. \tag{9}$$

To fulfill the quality of service (QoS) requirements of devices being served, we require that

$$r_d \geq R_d, \quad \forall d \in \mathcal{D} \setminus \mathcal{D}' \tag{10}$$

where R_d is the minimum critical rate for each device.

Finally, since we have a limited system bandwidth, we require that:

$$\sum_{d \in \mathcal{D} \setminus \mathcal{D}'} |S_d| B \leq M B_s. \tag{11}$$

C. Problem Formulation

We wish to maximize the number of connected machine type communication devices (MTCDs) that satisfy their QoS, transmit power and bandwidth constraints. Consider a vector of binary entries, $\boldsymbol{z} = (z_1, \ldots, z_k, \ldots, z_D)$. The k^{th} entry of this vector is 1 if $r_k \geq R_d$ and $p_k \leq P_d$, otherwise 0. Now, we have the optimization problem as follows:

$$\begin{aligned} \underset{p_{d,n}, z_d, S_d}{\text{maximize}} \quad & \sum_{i=1}^{D} z_i \\ \text{subject to} \quad & C1: \; z_d \in \{0,1\}, \forall d \in \mathcal{D} \\ & C2: \; r_d \geq z_d R_d, \forall d \in \mathcal{D} \\ & C3: \; 0 \leq z_d p_d \leq P_d, \forall d \in \mathcal{D} \\ & C4: \; \sum_{d} \mathbb{1}_{p_{d,n}} \leq M, n \in \mathcal{S}. \\ & C5: \; \sum_{d} z_d |S_d| B \leq M B_s, \forall d \in \mathcal{D}. \end{aligned} \tag{12}$$

Here, $\mathbb{1}_{p_{d,n}}$ is the indicator function which takes value 1 if $p_{d,n} > 0$ and 0 if $p_{d,n} = 0$. Constraint C4 states the SIC limit, i.e., the allowable maximum number of superimposed devices per sub-carrier in the system.

III. PROPOSED ALGORITHM

A. Bottom-Up Power Based Filling

In the following, we present the bottom-up power filling algorithm (BU) to solve the problem. It is a greedy heuristic scheme. We begin by calculating the power required by the devices to achieve their required data rate and sorting them in the increasing order of this power. The power required by device d can be evaluated as follows:

$$p_{d,n} = (2^{\frac{r_{d,n}}{B}} - 1) \frac{N_0 B + I_{d,n}}{g_d}. \tag{13}$$

We consider only those devices for which $p_d \leq P_d$. For the first round of sub-carrier allocation, all sub-carriers are free of interference, i.e., $\boldsymbol{I} = (0, \ldots, 0)$. Devices from the top of the sorted list are selected to fill up sub-carriers. These devices now act as interferers for subsequent devices. With the knowledge of interference on each sub-carrier, we again calculate the power required by the devices to achieve their data rates using (13). Again, we follow the same allocation strategy of sorting by power as before. With SIC limitation of M devices per sub-carrier, we go through M rounds of allocation. This greedy approach is computationally efficient. But the above allocation scheme cannot check for all possible resource allocation combinations since the allocations on the previous layer are already fixed by the time we move on to the next layer. Note that this shortcoming can be alleviated by the item clustering approach shown in Section III-B.

Algorithm 1 illustrates the proposed scheme. Similar to [7], one may set the transmit power to the maximum permissible P_d. This is reflected by over-riding the condition on line 9 of Algorithm 1 for some rounds of allocation. This consideration is meaningful when we have devices that are rate sensitive and not really concerned about power, for example, URLLC devices. However, we must be cautious of the increased interference incurred.

B. Item Clustering Based Greedy Heuristic

We now introduce a novel approach for maximizing device connection density. Instead of treating each device independently for connectivity consideration, we look at clusters of devices that may be put on any given sub-carrier such that all of the devices in that cluster achieve the data rate requirement. If we choose the SIC limit to be M devices, each device cluster will have a size of at most M. We shall refer to these clusters as *items*. Observe that the set of subsets of $\mathcal{D}$ which have cardinality less than or equal to M is:

$$[\mathcal{D}]^M \triangleq \{X \colon X \subseteq \mathcal{D} \wedge |X| \leq M\}. \tag{14}$$

In single-tone mode, each item corresponds to an element of $[\mathcal{D}]^M$. However, not all elements of $[\mathcal{D}]^M$ are valid items. The items that are valid for allocation are the ones for which any device d present in the item requires a transmit power not exceeding P_d and achieves its data rate requirement R_d. We denote the set of valid items by J.

These definitions can be extended to multi-tone mode. Since the items we form are on a per sub-carrier basis, we add

Algorithm 1 Bottom-Up Power Filling (BU)

```
Input: $B, S, M$ and $\forall d \in \mathcal{D}$, $S_d, g_d, P_d, R_d$
Initialization: $\boldsymbol{I} \leftarrow (0, \dots, 0)$, layer $\leftarrow 1$, $\mathcal{D}' \leftarrow \mathcal{D}$
while layer $\leq M$ do
    $n \leftarrow 1$
    while $\mathcal{D}' \neq \varnothing$ and $n \leq S$ do
        $A \leftarrow$ empty array of dimension $D \times |S_d|$
        for $d \in \mathcal{D}'$ and $s \in S_d$ do
            $A[d, s]$ is computed as the required power to achieve rate $R_d$ using sub-carriers $n$ to $n + s - 1$ with interference $\boldsymbol{I}[n, \dots, n + s - 1]$ according to (13)
        end for
        if $\forall\, d \in \mathcal{D}'$, $s \in S_d$, $A[d, s] > P_d$ then
            break
        end if
        $s^* \leftarrow \min\{s \in S_d \colon \exists\, d \in \mathcal{D}', A[d, s] \leq P_d\}$
        if $n + s^* - 1 > S$ then
            break
        end if
        $d^* \leftarrow \mathrm{argmin}_{d \in \mathcal{D}'}\{A[d, s^*]\}$
        $z_{d^*} \leftarrow 1$
        $\boldsymbol{I}[n, n + s^* - 1] \leftarrow \boldsymbol{I}[n, n + s^* - 1] + g_{d^*} A[d^*, s^*]$
        $n \leftarrow n + s^*$
        $\mathcal{D}' \leftarrow \mathcal{D}' \setminus \{d^*\}$
    end while
    layer $\leftarrow$ layer $+ 1$
end while
Output: $\boldsymbol{z}$
```

in each item the information regarding the number of sub-carriers $|S_d|$ used by each device d. Hence, valid items are the ones for which any device d present in the item requires a transmit power not exceeding $P_d/|S_d|$ and achieves at least rate $R_d/|S_d|$.

For each item j, we define $r_d(j)$ as the rate achieved by device d in item j. Furthermore, we extend the definition of $r_d(\cdot)$ by (15) below to take a set as input instead of a single item. This set function takes the set $\mathcal{K}$ of one or more items as the input and returns the rate achieved by device d for the given items. For example, $r_d(\{i, j\})$ gives the rate achieved by device d considering both items i and j. However, there can be cases where selecting an item can cause a device to be served above the required data rate requirement. We shall refer to such devices as over-served devices. When we make considerations in a system that allows allocation of only one sub-carrier per device, i.e., single-tone allocation, this aspect is insignificant. But when the system supports multi-tone operation, a set of more than one sub-carriers can be given to a single device. In such cases, some items $j \in J$ may have significant efficiency $r_d(\{j\})$ even though most part of $r_d(\{j\})$ contributes to over-serving. Now we define the efficiency of set $\mathcal{K}$ containing item j as follows:

$$r_d(\mathcal{K}) \triangleq \min\left\{\sum_{j \in \mathcal{K}} r_d(j), R_d\right\}. \tag{15}$$

Algorithm 2 Item Clustering Heuristic (IC)

```
Input: $J, S$, and $\forall d \in \mathcal{D}$, $R_d$
 1: Initialization: $n \leftarrow 0$, $\mathcal{K} \leftarrow \varnothing$, $\mathcal{D}' \leftarrow \mathcal{D}$
 2: while $n < S$ or $J \neq \varnothing$ do
 3:     $j^* \leftarrow \underset{j \in J}{\mathrm{argmax}}\ \tilde{r}_d(\mathcal{K} \cup \{j\})$
 4:     $\mathcal{K} \leftarrow \mathcal{K} \cup \{j^*\}$
 5:     $n \leftarrow n + 1$
 6:     if $\tilde{r}_d(\mathcal{K}) = R_d$ for any $d \in \mathcal{D}'$ then
 7:         $z_d \leftarrow 1$
 8:         $J \leftarrow J \setminus \{j\}$, for any item $j$ such that $\tilde{r}_d(\{j\}) > 0$
 9:         $\mathcal{D}' \leftarrow \mathcal{D}' \setminus \{d\}$
10:     end if
11: end while
Output: $\boldsymbol{z}$
```

It is easy to observe that the efficiency for an item is dynamically updated at each iteration according to the previously picked items in $\mathcal{K}$. Intuitively, this decreases the efficiency of an item if including it exceeds R_d for any of the associated devices.

All items are sorted in decreasing order of efficiency. We pick greedily the most efficient item until we run out of sub-carriers or unserved devices (see line 2 of Algorithm 2). When any device d is served with rate R_d, all items related to d can be removed from J (see lines 6-8). This is done to avoid continuing serving device d after it has achieved its data rate requirement. In addition, d is removed from the set of unserved devices $\mathcal{D}'$, and we indicate $z_d = 1$ (see lines 7 and 9).

We have the item clustering heuristic (IC) given in Algorithm 2. Note that variable Variable n represents the number of allocated sub-carriers. It is initialized as 0 and incremented each time when an item is picked (see line 5). The set $\mathcal{K}$ represents the pool of items already taken into consideration. $\mathcal{K}$ is empty at initialization, and it is updated at each iteration. $\mathcal{D}'$ corresponds to the set of unserved devices.

This approach can be applied to both single-tone and multi-tone modes. When operating in multi-tone mode, item clustering will perform non-contiguous sub-carrier allocations since the most efficient item in a given round of allocation may not necessarily contain the same devices as the previously chosen item. Removing the restriction of contiguous sub-carrier allocation allows us to have more items for consideration which in turn increases the possibility of supporting more devices. A significant issue with non-contiguous sub-carrier allocation is that it may render us unable to use SC-FDMA, thereby increasing the peak-to-average power ratio (PAPR) in the uplink. This issue of fragmented bandwidth assignment to a UE has already been addressed in LTE-A systems through the application of clustered DFT-S-OFDM [11]. Hence, when using item clustering, non-contiguous sub-carrier allocation performs better than its contiguous counterpart in terms of number of connected devices with little compromise in the PAPR.

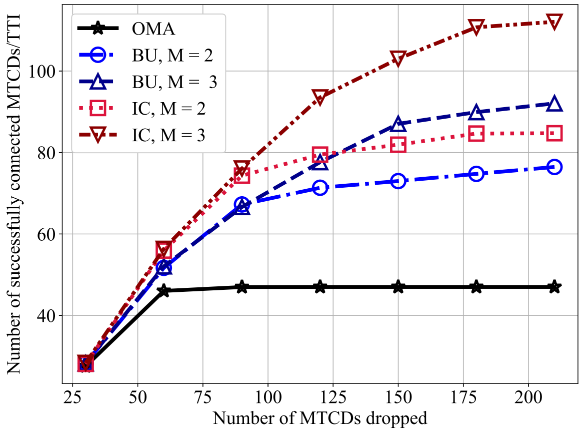

Fig. 1. Number of successfully connected MTCDs in one TTI versus number of MTCDs for single-tone uplink assignment.

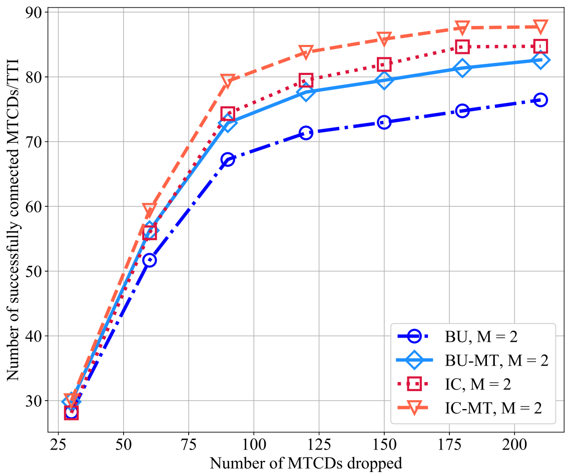

Fig. 2. Number of successfully connected MTCDs in one TTI versus number of MTCDs for multi-tone uplink assignment.

IV. SIMULATION RESULTS AND PERFORMANCE EVALUATION

We evaluate the performance of the proposed algorithms and also show that of conventional orthogonal multiple access (OMA) scheme from 3GPP which can be considered as a baseline. We compare the total number of MTCDs which can be connected with respect to their rate and power constraints in a 3GPP system with one PRB of bandwidth 180 kHz. The system parameters considered for performance evaluation follow the NB-IoT specifications given in [2]. It is assumed that perfect channel state information is available at the base station. For the simulations, we have 48 sub-carriers, each with bandwidth 3.75 kHz for both single and multi-tone operations. MTCDs are distributed uniformly in a hexagonal cell with inter-site distance 500 m [12]. The frequency of operation is taken to be 900 MHz. We consider a single base station with an omnidirectional antenna supporting the MTCDs' demands. Flat fading Rayleigh channel is considered due to narrow system bandwidth. The path loss model follows [2], where the path-loss $PL(D)$ is defined as:

$$PL(D) = 120.9 + 37.6 \log\left(\frac{D}{1000}\right) + L + G \tag{16}$$

where D is the distance between an MTCD and the base station in meter, and L is the indoor penetration loss that is assumed to be 20 dB. G is the antenna gain of -4 dB. 80% of MTCDs are assumed to be indoor, the remaining 20% are outdoor MTCDs. An additive white Gaussian noise with power spectral density -174 dBm/Hz and noise figure of 5 dB are considered. The maximum transmission power, P_d, is set to 23 dBm [2].

Figure 1 shows the performance of BU (Algorithm 1), IC (Algorithm 2), and 3GPP-OMA. All schemes follows single tone operation such that one sub-carrier can be allocated to at most one device. For figures 1 and 2, all devices have to achieve a rate R_d, which is randomly and uniformly distributed in the range $(0.1, 50)$ kbps. It is evident that using NOMA on each sub-carrier drastically increases the connectivity even when using a simple assignment scheme like BU, connecting up to 60% more devices per TTI with $M = 2$ and up to 95% more devices per TTI with $M = 3$. When employing IC, we can connect up to 80% more devices compared to OMA when $M = 2$, and up to 135% more devices compared to OMA when $M = 3$. Figure 2 shows simulation results for multi-tone operation where we compare Algorithms 1 and 2, whose results are denoted by BU-MT and IC-MT respectively. IC performs the best among the schemes when operating in multi-tone mode, connecting up to 6% more devices compared to BU-MT. Besides, we see that BU-MT performs better than BU with single-tone operation, connecting up to 8% more devices. This is because with bottom-up approach, especially when R_d is high, we can often run out of eligible devices for connectivity when using one sub-carrier. Using multi-tone operation, it is possible to connect some additional devices which cannot achieve rate R_d while using one sub-carrier and transmitting with power less than P_d, but can do so using multiple sub-carriers. We do not observe similar significant gains for item-clustering based approach when comparing BU-MT with BU with single-tone since BU with single-tone often already forms a rather exhaustive list of eligible device clusters per sub-carrier while constructing the item pool.

Figure 4 shows the variation of number of connected devices as the required service rate changes. We consider 200 devices requesting connection in one TTI with service rate R_d being random and uniformly distributed in the range $(0.1, R_d^{max})$ kbps. As R_d^{max} increases, we can see that the number of successfully connected devices will decrease gradually. However, still both IC and BU can perform much better than the conventional 3GPP-OMA, connecting up to 100% more devices.

Figure 3 compares BU and IC schemes to state-of-the-art scheme developed by Mostafa et. al. in [7]. For this comparison, we follow the system model and simulation

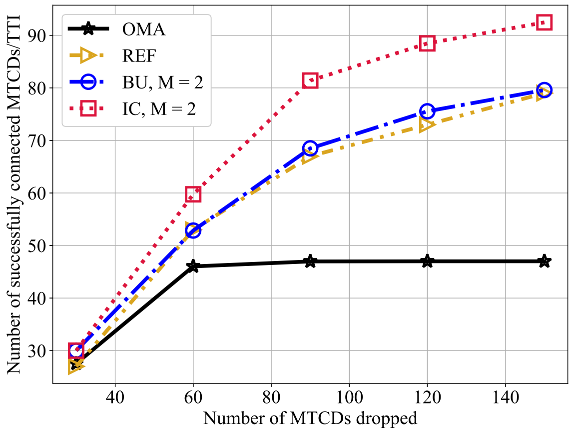

Fig. 3. Comparison of proposed schemes with state-of-the-art techniques

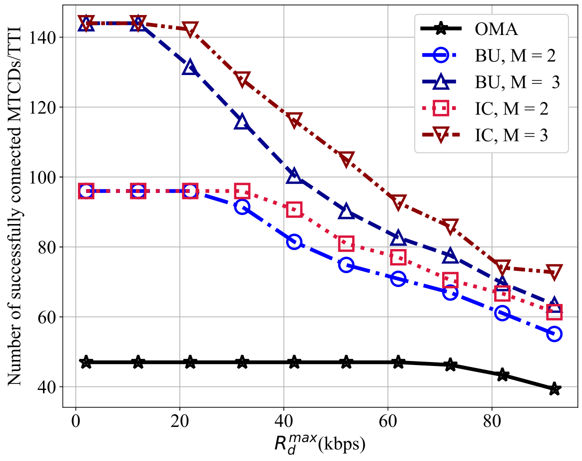

Fig. 4. Number of connected MTCDs with various allowable service rate

parameters of [7]. The devices are dropped randomly and uniformly distributed in a square region of area 1 km^2. Single-tone operation is considered. Two classes of devices are considered, mMTC devices for which $R_d \geq 10$ kbps and URLLC devices for which $R_d \geq 50$ kbps. URLLC devices do not have a transmit power constraint in the model of [7]. Following identical system parameters, we run BU and IC schemes. We can see that BU performs as good as the scheme proposed in [7], which we denote as REF. On the other hand, IC substantially outperforms both BU and REF, connecting up to 55% more devices per TTI.

V. CONCLUSION

In this paper, we investigate the bottom-up power filling (BU) and item clustering (IC) based schemes for maximizing the connection density in NB-IoT with NOMA. BU performs significantly better than both 3GPP-OMA and state-of-the-art scheme REF, connecting up to 55% more devices when compared to the latter. IC dynamically adjusts with the devices being served, preventing over-service of devices and hence increasing the number of connected devices. Note that BU and IC are applicable to both single-tone and multi-tone uplink operation, and are easily implementable and efficient without the need of heavy computations. Additionally, through IC, we see that in multi-tone mode, non-contiguous allocation performs better than contiguous allocation. We evaluate the performance of the proposed schemes over a wide range of required service rate and conclude that IC outperforms OMA and BU at all rates. Future work includes multi-cell user scheduling, transmit power optimization and user fairness considerations.

REFERENCES

[1] 3GPP, “Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) radio transmission and reception,” 3rd Generation Partnership Project (3GPP), Technical Specification (TS) 36.101, 10 2019, version 16.3.0. [Online]. Available: http://www.3gpp.org/DynaReport/36101.htm

[2] 3GPP, “Cellular system support for ultra-low complexity and low throughput Internet of Things,” 3rd Generation Partnership Project (3GPP), Technical Report (TR) 45.820, December 2015, version 13.1.0. [Online]. Available: http://www.3gpp.org/DynaReport/45820.htm

[3] Z. Ding, P. Fan, and H. V. Poor, “Impact of user pairing on 5G non-orthogonal multiple-access downlink transmissions,” *IEEE Transactions on Vehicular Technology*, vol. 65, no. 8, pp. 6010–6023, 2016.

[4] K. Higuchi and A. Benjebbour, “Non-orthogonal multiple access (NOMA) with successive interference cancellation for future radio access,” *IEICE Transactions on Communications*, vol. E98.B, no. 3, pp. 403–414, 2015.

[5] Z. Ding, M. Xu, Y. Chen, M. Peng, and H. V. Poor, “Embracing non-orthogonal multiple access in future wireless networks,” *Frontiers of Information Technology & Electronic Engineering*, vol. 19, no. 3, pp. 322–339, 2018.

[6] 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.

[7] A. E. Mostafa, Y. Zhou, and V. W. S. Wong, “Connectivity maximization for narrowband IoT systems with NOMA,” in *IEEE International Conference on Communications (ICC)*, 2017.

[8] Z. Zhang, Y. Hou, Q. Wang, and X. Tao, “Joint sub-carrier and transmission power allocation for MTC under power-domain NOMA,” in *IEEE International Conference on Communications Workshops (ICC Workshops)*, 2018.

[9] Y. Saito, A. Benjebbour, Y. Kishiyama, and T. Nakamura, “System-level performance of downlink non-orthogonal multiple access (NOMA) under various environments,” in *IEEE 81st Vehicular Technology Conference (VTC Spring)*, 2015.

[10] A. Shahini and N. Ansari, “NOMA aided narrowband IoT for machine type communications with user clustering,” *IEEE Internet of Things Journal*, vol. 6, no. 4, pp. 7183–7191, 2019.

[11] T. Iwai, A. Matsumoto, S. Takaoka, Y. Ogawa, A. Nishio, and D. Imamura, “System performance of clustered DFT-S-OFDM considering maximum allowable transmit power,” in *2011 IEEE International Conference on Communications (ICC)*, June 2011, pp. 1–5.

[12] 3GPP, “Study on scenarios and requirements for next generation access technologies,” 3rd Generation Partnership Project (3GPP), Technical Report (TR) 38.913, July 2018, version 15.0.0. [Online]. Available: http://www.3gpp.org/DynaReport/38913.htm