IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2020.DOI

Multiuser Detection for Downlink Communication in LoRa-like Networks

ANGESOM ATAKLITY TESFAY[1], ERIC PIERRE SIMON[1], IDO NEVAT[2] AND LAURENT CLAVIER.[1,3]

[1]University of Lille, CNRS, UMR 8520 - IEMN, F-59000, Lille, France (e-mail: firstname.name@univ-lille.fr)
[2]TUMCREATE, Singapore 138602
[3]IMT Lille Douai, France

Corresponding author: Angesom Ataklity Tesfay (e-mail: firstname.name@univ-lille.fr).

ABSTRACT Internet of Things (IoT) technology has become ubiquitous in a multitude of applications and its use is growing. However, the expansion of IoT faces a major difficulty: scalability, that is, very dense deployment of communicating devices is currently limited. In long-range networks, such as LoRa, the downlink is critical because it limits the number of acknowledgements that can be sent, and consequently reliability. It also limits the possibility to update the devices, which could be critical when they are deployed for decades. To overcome those problems, we propose a solution, inspired by Non Orthogonal Multiple Access (NOMA) techniques, to increase by at least one order of magnitude the number of devices that can be addressed. While the approach differentiates the devices by the power allocated to them, it differs from the vast majority of previous works on power domain NOMA because it does not require interference cancellation. Instead, it benefits from the spectrum spreading of the modulation scheme (chirp spread spectrum), where, at the end of the decoding phase, the information carried by a symbol is found in the position of a peak in the Fourier domain. In the vast majority of cases, the information from different users results in different peak positions, not creating any interference. In that sense, we get closer to avoidance schemes such as time or frequency hopping, but without using a code. In this paper, we propose a new solution for NOMA in the power domain that does not suffer from the limitations induced by interference cancellation residues. The proposed system, including preamble detection and channel estimation, is presented and evaluated by simulations. We demonstrate that our proposed scheme increases the number of devices by one order of magnitude compared to the current system which allows addressing only one user at a time and maintains full compatibility with the LoRa physical layer standard.

INDEX TERMS Chirp Spread Spectrum, LoRa, Multiuser detector, Power allocation, Scalability

I. INTRODUCTION

THE Internet of Things (IoT) is rapidly growing with more than 20 billion devices expected to be connected in the next 5 years, including more than 2 billion connected through Low Power Wide Area Networks (LPWAN) [1]. Many applications, such as smart metering [2], [3], parking space monitoring [4], [5], building monitoring [6], and smart agriculture [7], [8] require low power operations and wide communication coverage. LPWAN provides a good solution for such requirements and LoRa is one of those protocols which can enable such methods. Still, challenges remain in order to provide robust and reliable communication links and face the huge increase in the number of communicating devices.

A. RELATED WORKS

So far the researches on LoRaWAN scalability have mainly concerned the uplink and the Medium Access Control (MAC) layer. Challenges of the radio channel access and the scalability are discussed in [9]. Real world deployments of IoT networks and LoRaWAN experimental analyses have been conducted. The impact of an adaptive data rate is illustrated as well as the impact of duty cycle restrictions on the LoRa network. Scalability limits of LoRaWAN have been studied for instance in [10], [11].

It is limited because of the regulatory constraints on the channel and the use of random channel access rather than because of LoRa technical limitations. In [11], it is shown that the network can support only up to 120 nodes with a data extraction rate over 0.9 and using a single gateway, a

Spreading Factor (SF) of 12, a bandwidth of 125 kHz and forward error correction with Code Rate $4/5$. The authors proposed a dynamic channel parameter setting (transmission power, SF, bandwidth, and code rate) with multiple gateways and showed a possible increase by one order of magnitude of the number of users.

At the PHY layer, less work has been done and open problems still remain. Still in the uplink, capture effect is analyzed in [9], [12] and gives the idea of separating users based on the received power. The authors in [13] proposed to decode two desynchronized LoRa-like signals, received concurrently in the same channel, with the same SF. The approach is based on estimating the time shift between the two received signals but the decoding works only if the signal that has the highest power is received first. In [14] a different approach is also proposed that decodes superposed LoRa signals using the chirps timing information.

Those ideas have been further developed and extended using a Serial Interference Cancellation (SIC) scheme to decode signals from multiple users in the uplink [12], [15], [16]. This leads to a significant increase in the possible number of devices connected to the network. For instance, in [16], a SIC receiver allows a receiver to recover multiple signals transmitted in the same time slot (but not necessarily synchronized) and with the same SF. A complete receiver structure, with detection of packets, channel estimation, detection of symbols, and interference cancellation is proposed and shown to support 20 times more nodes than the classic LoRa receiver. Practical implementation issues remain, however, to be addressed.

In [17] it is shown that multiple gateways and directional antennas may be used to increase devices' density by reducing the number of devices connected to a single antenna. This approach can be combined with our proposal, which only considers the Physical (PHY) layer.

The solutions proposed in the previously mentioned papers concentrate solely on the uplink direction. However, it has been shown [18] that the duty cycle restricts both the scalability and the network's reliability. In [19] an event-driven simulator is implemented with multiple gateways and shows that the network performance is significantly impacted by the downlink duty cycle restrictions. The duty cycle is indeed identified as the main scalability limiting factor in LoRaWAN [20], [21].

However, to date, still only a few works tackle this important aspect of scalability issues of LoRa downlink traffic. The authors in [22] present and evaluate gateway selection solutions to improve the performance of LoRaWAN downlink communications in terms of throughput. They consider gateway selection algorithms based on load balancing and received signal strength indicator with different deployment scenarios.

Van et al. [23] conducts an analysis of the downlink frame issue using the ns-3 network simulator. The authors illustrate the main reasons behind the gateways congestion by emphasizing the duty cycle limitation as well as the half-duplex problem. They also propose a multi-gateway system and assess the improvements based on lower duty cycle saturation as well as load balancing among the gateways. However, there is no explanation of how the load balancing is performed.

A comprehensive analysis of the effects of downlink traffic on LoRaWAN capacity was provided in recent paper [24]. The authors describe how the gateway's half-duplex mode and the sequential transmission leads to the duty cycle saturation for downlink traffic. They also propose a multi-gateway deployment, parallel transmission of downlink frames on the same channel but with different orthogonal SF, and load balancing among the deployed multiple gateways.

B. CONTRIBUTION

In the previously cited papers, no solution is proposed to improve the situation at the PHY layer. In comparison to the uplink, the downlink benefits from a crucial advantage: the transmission to all users originates from the same access point, allowing accurate synchronization of signals and power allocation. Leveraging on this, we propose a power-domain Non Orthogonal Multiple Access (NOMA) scheme [25] that limits the impact of the duty cycle, half-duplex mode, and sequential transmission of a LoRa network. Besides, the complexity at the receiver level must remain low, prohibiting the use of many advanced methods, so that a SIC should be avoiding or would significantly limit the possible number of simultaneous transmissions in practical systems. Taking benefit from the Chirp Spread Spectrum (CSS) modulation, the main novelty in our work is to propose a power domain NOMA scheme without the need of a SIC receiver. In addition to reducing the complexity, it also avoids the residual error resulting from signal cancellation, which means a significantly easier and less limited scheme to be implemented in real systems.

The main contributions of the paper can be summarized as follows:

1) We propose a superposition transmission scheme for synchronized (downlink) CSS modulation which does not require SIC.
2) We derive the Maximum Likelihood (ML) optimal Multi User Detector (MUD). Due to its high complexity, it cannot be implemented in practice, we therefore develop an approximation via the Cross Entropy Method (CEM). Still too complex for end devices, this receiver serves us as a reference.
3) We develop a suboptimal detection scheme which significantly reduces the computation time compared with the CEM, specifically designed for low cost devices. This scheme does not require interference cancellation. It benefits from the collision avoidance inherent in CSS. This is a significant advantage of the method since it avoids the limitations due to the propagation of residual errors after each cancellation. Preamble detection and channel estimation scheme are included in our simulation results.

4) We propose Power Allocation (PA) schemes to minimize error probabilities and increase fairness between users having good or bad channels.

Our scheme is very different from traditional NOMA, both for the decoding approach and the PA scheme. Indeed, the information in the decoded signal is carried by a frequency, i.e., a peak in the Fourier domain. Most of the time, users carry different information so that their peaks fall at different positions, avoiding each other and the interference does not accumulate. The scheme is more similar to time or frequency hopping than to superposition coding or Direct Sequence Code Division Multiple Access and does not require a specific code. As far as we know, this is the first time such a NOMA scheme is proposed. A few papers have studied downlink NOMA without SIC. In [26], a Pulse-Amplitude modulation is studied and the Gray labeling without SIC very close to a system using SIC. In [27], a downlink NOMA technique without SIC is proposed using an algebraic lattice to design modulations that guarantee all users to achieve full diversity gain. These approaches, although not relying on SIC, use a superposition coding scheme, and try to maximize capacity or diversity gain. Besides, they are limited in the number of users and only consider two. On the contrary, the CSS that is used in LoRa is not maximizing the transmission rate but our proposal can handle many more users. The avoidance rather than superposition approach avoids the cancellation residue that significantly limits the possible number of simultaneous transmissions.

We keep our analysis at the PHY layer level, although characterizing in terms of goodput and latency could also certainly highlight the benefits of our method. However, it would be necessary to include channel coding schemes, protocol solutions, and re-transmission strategies that go beyond the scope of this paper.

C. ORGANIZATION OF THE PAPER

This paper is organized as follows: the description of LoRa technology is provided in section II. In section III, the proposed multiple user transmission and reception schemes are presented. Two power allocation strategies are described in section III-D and section IV analyzes the simulation results. Section V concludes the paper.

II. LORA TECHNOLOGY

In this section, we provide a short background for the downlink protocol proposed in the widely used LoRaWAN.

A. LORAWAN DOWNLINK CAPABILITIES

LoRa defines a PHY Layer protocol and LoRaWAN is an open standard defining a MAC protocol [28]. The downlink capabilities differ depending on the node class (A, B or C) [29]:

- Class A devices are of the lowest cost and energy consuming nodes. They initiate the communication with a pure ALOHA medium access. A transmission is followed by two short downlink windows to receive a response from the gateway. If the downlink traffic is received in the first window, the second is disabled.
- Class B devices allow additional downlink traffic. They are synchronized using periodic beacons sent by the gateway in order to open an additional receive window regardless of prior successful uplink transmissions. This is achieved at the expense of some additional power consumption in the end nodes.
- Class C devices are always listening and, consequently, can receive packets at any time. They usually require a permanent power source and can be used as access points.

For battery operated class A or B devices and with an expected lifetime of several years, the downlink is very limited in terms of the quantity of information that can be transmitted because a very limited number of slots are available. However, this is not the only limitation. LoRa operates in the license-free Industrial, Scientific and Medical (ISM) radio band, and consequently, suffer severe limitations with respect to channel access. Because the devices do not listen to the channel before transmitting, the devices must adhere to the duty cycle regulation imposed by the regulatory bodies, such as the European Telecommunications Standards Institute (ETSI) [30]. This restriction limits for a transmitting device the time it can occupy the channel, for instance, 1% in Europe ($868-868.6$ MHz band). This duty cycle is significant for the downlink transmission and the gateway is extremely affected, since after receiving multiple uplink frames, it is not able to send downlink frames to all of the devices. This constraint limits the capacity of downlink transmission in terms of the maximum traffic supported and, consequently, impacts the scalability of LoRaWAN networks.

B. CHIRP SPREAD SPECTRUM MODULATION

LoRa is based on CSS modulation [31], [32]. This modulation is defined by its SF, ranging from 7 to 12 (SF 6 also exists but corresponds in fact to a very different modulation scheme and smaller SF also exists in the 2.4 GHz band). It provides a trade-off between rate and communication range for a fixed Bandwidth (B) [33]. The symbol consists in a linear frequency change over the symbol duration T_s, where $T_s = 2^{SF}T$, $T = 1/B$. There are SF bits in one symbol and increasing SF by 1 doubles the symbol duration with only one added bit.

The instantaneous frequency can be described as the derivative of the phase $\varphi(t)$, i.e., $f(t) = \frac{1}{2\pi}\frac{d\varphi(t)}{dt}$. The raw upchirp $x(t)$ is then defined by its instantaneous frequency $f(t) = Bt/T_s$ and given by

$$x(t) = e^{2\imath\pi\frac{B}{2T_s}t^2} \quad \text{for } t \in \Big[-\frac{T_s}{2}, \frac{T_s}{2}\Big[. \tag{1}$$

The transmitted chirp of user i at time qT_s, $q \in \{0, Q-1\}$, where Q is the number of symbols transmitted in a packet, is generated by applying left-shifting of the raw up-chirp by $\delta_q^{(i)} = m_q^{(i)}T$, as shown in Fig. 1. The transmitted data symbol is represented by $m_q^{(i)} \in \{0, ..., 2^{SF}-1\}$. The coded

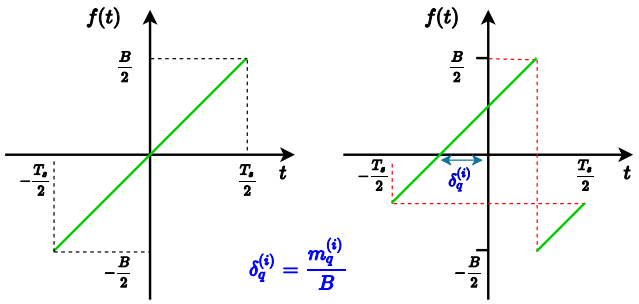
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3034973, IEEE Access

chirp of user i associated with the q-th symbol is

$$x_q^{(i)}(t) = \begin{cases} \exp\left(2\imath\pi\left[\frac{B}{2T_s}t^2 + \frac{m_q^{(i)}}{T_s}t\right]\right), \\ \qquad t \in \left[-\frac{T_s}{2}, \frac{T_s}{2} - \delta_q^{(i)}\right[, \\ \exp\left(2\imath\pi\left[\frac{B}{2T_s}t^2 + \left(\frac{m_q^{(i)}}{T_s} - B\right)t\right]\right), \\ \qquad t \in \left[\frac{T_s}{2} - \delta_q^{(i)}, \frac{T_s}{2}\right[. \end{cases} \quad (2)$$

Finally, the complex envelope of the transmitted signal of the ith user is

$$x^{(i)}(t) = \sum_{q=0}^{Q-1} x_q^{(i)}(t - qT_s). \quad (3)$$



(a) Raw Upchirp (b) Coded chirp

FIGURE 1: Raw chirp and Coded chirp associated with $m_q^{(i)}$.

C. POWER CONSIDERATIONS

The maximum transmit power in an ISM band is restricted. For LoRa (at 868 MHz), in Europe, this maximum is set to 14 dBm for uplink and 27 dBm for downlink. The noise level of a receiver at room temperature is

$$N_0(\text{dBm}) = -174 + 10\log_{10}(B) + \text{NF}, \quad (4)$$

where the first term is the thermal noise in 1 Hz of bandwidth and can only be affected by changing the temperature of the receiver. NF is the receiver noise figure which depends on the hardware implementation and a typical 6 dB noise figure is considered [33]. If we consider $B = 250$ kHz, the noise power density at the receiver is -114 dBm.

III. PROPOSED MULTIUSER SCHEME

In this section, we present our proposed scheme. We first present the transmitter design and the wireless channel model. We then present our method for the receiver, starting from the optimal detector formulation via the Maximum Likelihood Detector (MLD) which we show has a combinatorial complexity. To overcome this limitation, we propose an algorithm which approximates the MLD with a much lower complexity that is based on the CEM [34]. To further reduce the computational complexity, we propose a low-complexity multiuser detector that only searches for the closest peak to the expected one. It is to be noted that the CEM based and the low-complexity detectors will be used as references for comparison. We did not found any competitive method in the literature nor theoretical bounds for a method equivalent to our NOMA proposal. The only one which would be relevant would be a pure ALOHA scheme or a perfect Time Division Multiple Access (TDMA) approach. Even if it is impractical in real networks, this latter approach will be used also as a comparison but our scheme significantly outperforms both.

We will also propose in this section two PA schemes. The first one avoids ambiguities when two users transmit the same information at the same time. The second distributes the powers to ensure a constant distance between the power of the desired peak and that of the nearest lower peak to improve equity.

A. TRANSMITTED SIGNAL

To avoid the limitation due to the duty cycle, we propose to transmit N frames simultaneously, with the same SF and on the same frequency band. It is possible with Class B devices that can be synchronized and in receive mode during the same time frame. The objective is then to design a communication strategy that allows us to superimpose N users in the duration of a single packet.

The idea is to generate information streams for N end-devices, modulate them using the CSS scheme, then add all signals with different allocated powers to form a single packet. A preamble and a common header are added at the packet start. The information about the number of users and the PA scheme is added in the header. At the receiver side, the receivers select and decode the signal which corresponds to their allocated power. The combined transmitted signal is

$$x(t) = \sum_{i=1}^{N} \sqrt{p^{(i)}}\, x^{(i)}(t), \quad (5)$$

where $p^{(i)}$ is the power allocated to user i.

B. CHANNEL MODEL

We consider one cell of radius R with the gateway placed at the center. A large number of devices are uniformly distributed within the cell and the gateway has to send information to N of them. The distance from end-device i to the gateway is denoted by $d^{(i)}$. The propagation channel is considered block and flat fading, so a single constant coefficient throughout the packet's duration. We consider path loss and Rayleigh multipath fading χ_i. The signal amplitude decays with increasing distance according to ${d^{(i)}}^{-\eta/2}$, where η is the path loss exponent. The channel attenuation (in amplitude) is expressed as $h^{(i)} = {d^{(i)}}^{-\eta/2} \cdot \chi^{(i)}$.

In the following, the user we are trying to decode is denoted by j and the signal it receives is

$$r^{(j)}(t) = h^{(j)} \sum_{i=1}^{N} \sqrt{p^{(i)}}\, x^{(i)}(t) + w^{(j)}(t), \quad (6)$$

where $w^{(j)}(t)$ is a complex Gaussian noise and $h^{(j)}$ the channel between the access point and device j.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3034973, IEEE Access

C. RECEIVER DESIGN

The receiver acts in two steps: preamble detection and demodulation process.

First, the correlation between the received signal and the known preamble sequence is calculated to detect the packet and its start. This detection is easy because the preamble is common to all the users, hence the power dedicated to it is large. This step also allows us to estimate the user's channel from the maximum value of the correlation.

Second, we are interested in the decoding of symbol q. We sample the received signal and note the samples

$$r_q^{(j)}[n] = r^{(j)}(nT + qMT),$$

where n is limited to the set $\Omega = \{-\frac{2^{SF}}{2}, \dots, \frac{2^{SF}}{2} - 1\}$. We have

$$r_q^{(j)}[n] = h^{(j)} \sum_{i=1}^{N} \sqrt{p^{(i)}}\, x_q^{(i)}[n] + w_q^{(j)}[n], \tag{7}$$

where $w_q^{(j)}[n] \sim \mathcal{N}_{\mathbb{C}}(0, \sigma_n^2)$ is a complex Gaussian thermal noise (discussed in section II-C) and $x_q^{(i)}[n] = x_q^{(i)}(nT)$.

The received samples $r_q^{(j)}[n]$ are then multiplied by the conjugate complex form of the sampled upchirp. The sampled upchirp is defined for $n \in \Omega$ and denoted by $x[n] = x(nT)$. The signal corresponding to the q-th symbol after dechirping is written as

$$\begin{aligned} y_q^{(j)}[n] &= r_q^{(j)}[n] x^*[n], \\ &= h^{(j)} \sum_{i=1}^{N} \sqrt{p^{(i)}} e^{2\imath\pi \frac{m_q^{(i)}}{2^{SF}} n} + \tilde{w}_q^{(j)}[n]. \end{aligned} \tag{8}$$

To make a decision, we use the ML estimator. It is equivalent to work with $r_q^{(j)}[n]$, $y_q^{(j)}[n]$ or even in the Fourier domain. We did not get any tractable solution in the time domain so, after compensating for the channel gain multiplying $y_q^{(j)}[n]$ by $h^{(j)*}$ (the $*$ denotes the complex conjugate -in the following formulations, we assume a perfect channel estimation but in the simulation part the estimated channel will be used), we perform a Fast Fourier transform (FFT), as the traditional LoRa receiver:

$$\begin{aligned} Z_q^{(j)}[k] &= \mathrm{Re}\left\{ \sum_{n=-2^{SF-1}}^{2^{SF-1}-1} \left(h^{(j)*} y_q^{(j)}[n] \right) e^{-2\imath\pi \frac{nk}{2^{SF}}} \right\}, \\ &= |h^{(j)}|^2 \sum_{i=1}^{N} \sqrt{p^{(i)}} \delta[k - m_q^{(i)}] + W_q^{(j)}[k]. \end{aligned} \tag{9}$$

where $W_q^{(j)}[k] \sim \mathcal{N}_{\mathbb{C}}(0, |h^{(j)*}|^2 \sigma_n^2/2)$ is the FFT of the noise, $\delta[.]$ is Kronecker delta function, $\delta[n] = 1$ for $n = 0$ and $\delta[n] = 0$ for $n \neq 0$. We observe that the obtained vector presents peaks at the positions corresponding to the source symbols ($\delta[k - m_q^{(i)}]$).

The basic LoRa detector searches for the strongest peak to decode a single user. This observation will be used in the receiver architectures we propose.

Based on (9), we note

$$\mathbf{Z}_q^{(j)} = |h^{(j)}|^2 \mathbf{X}_q^{(j)} + \mathbf{W}_q^{(j)}, \tag{10}$$

which is the 2^{SF}-dimensional received vector after FFT. We want to implement a multiuser detection scheme. To do so, we express the log likelihood function. $\mathbf{m}_q = \{m_q^{(1)}, \dots, m_q^{(N)}\}$ is the information vector and we assume an independent and identically distributed noise vector $\mathbf{W}_q^{(j)}$:

$$\begin{aligned} \Lambda &= \log \mathbb{P}\left(\mathbf{Z}_q^{(j)} \,\middle|\, h^{(j)}, \mathbf{m}_q \right), \\ &= \log \prod_{k=0}^{2^{SF}-1} \mathbb{P}\left(Z_q^{(j)}[k] \,\middle|\, h^{(j)}, \mathbf{m}_q \right), \\ &= \sum_{k=0}^{2^{SF}-1} \log \mathbb{P}\left(|h^{(j)}|^2 X_q^{(j)}[k] + W^{(j)}[k] \,\middle|\, h^{(j)}, \mathbf{m}_q \right). \end{aligned} \tag{11}$$

Because all operations are linear and given $\mathbf{m}_q$, $|h^{(j)}|^2 X_q^{(j)}[k] + W_q^{(j)}[k]$ is a Gaussian random variable with mean $|h^{(j)}|^2 X_q^{(j)}[k]$, variance σ_n^2 and we have

$$\begin{aligned} \widehat{\mathbf{m}}_q &= \underset{\mathbf{m}_q \in \mathcal{Q}}{\mathrm{argmax}} \sum_{k=0}^{2^{SF}-1} \log \left(\frac{1}{\sqrt{\pi |h^{(j)}|^2 \sigma_n^2}} e^{-\frac{\left(Z_q^{(j)}[k] - |h^{(j)}|^2 X_q^{(j)}[k] \right)^2}{|h^{(j)}|^2 \sigma_n^2}} \right) \\ &= \underset{\mathbf{m}_q \in \mathcal{Q}}{\mathrm{argmin}} \|\mathbf{Z}_q^{(j)} - |h^{(j)}|^2 \mathbf{X}_q^{(j)}\|^2, \end{aligned} \tag{12}$$

where $\mathcal{Q} := \left[0, \dots, 2^{SF} - 1\right]^N$ is the set of all possible symbols. As expected in the Gaussian independent noise case, maximizing the likelihood function Λ is equivalent to minimizing the Euclidean distance between the transmitted signal $\mathbf{X}_q^{(j)}$ and the received one $\mathbf{Z}_q^{(j)}$. However, this problem does not give an analytical expression for the solution. The difficulty is that $m_q^{(i)}$ can take any integer value between 0 and $2^{SF} - 1$, meaning that with N users, the MLD is required to evaluate $2^{N.SF}$ possible source combinations which is impractical.

One common approach to overcoming the combinatorial complexity is relaxing the problem by assuming that m has a real-valued support. However, in our model this relaxation would not simplify the solution due to the nonlinear structure of the likelihood function, neither in time nor in frequency. We therefore develop an alternative solution which is based on a Monte Carlo sampling technique, known as the CEM [34].

1) Cross Entropy Multiuser Detector

The CEM is a flexible Monte Carlo technique, which was originally developed for rare-event probability estimation, solving combinatorial, continuous, constrained, and noisy optimization problems [35].

The basic idea is to generate a set of candidate solutions ($\mathbf{m}_q$ in our case consisting of N integers in $\{0, \dots, 2^{SF} - 1\}$), select the best possible candidates, update the generating

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3034973, IEEE Access

rule and iterate until convergence is obtained. One important step is the possible solution generation: what distribution for $\mathbf{m}_q$ should be chosen? Let $f_{\mathbf{m}}(.)$ be the Probability Mass Function (PMF) of $\mathbf{m}_q$. The proposed cross entropy algorithm is presented in algorithm 1.

Steps 1 to 3 define some parameters: the number of sequences we generate at each iteration, the number of sequences we keep to update the distribution and a parameter that controls the speed of the convergence. Along with two parameters chosen for the initialisation of the PMF of $\mathbf{m}_q$, these parameters are important and could be optimized because they represent a compromise between the complexity burden and the accuracy of the algorithm. We chose parameters that ensure a good convergence rather than a reduced complexity to perform close to the true ML.

Steps 4 to 9 initialize the PMF of $\mathbf{m}_q$. All values are possible but we give a slightly higher probability to the dominant peaks. This reduces the necessary number of iterations. An example is seen in Fig. 2, where the initial $f_{\mathbf{m}}$ is represented. The same PMF is used for each user.

Step 10 starts the main loop. We set the end of iterations when for each user the probability of a given value is at least 0.85. This probability is set empirically.

From steps 11 to 17, we generate N_{seq} random sequences $\tilde{\mathbf{m}}_q$ according to $f_{\mathbf{m}}$ and the corresponding decoded vector $\tilde{\mathbf{Z}}_q^{(j)}[k]$. This requires the channel estimate $\hat{h}^{(j)}$. The distance with the true received sequence is also calculated. We chose $N_{seq} = 2000$ to ensure enough variety in the generated sequences and a good convergence of the algorithm.

Steps 18 and 19 select the N_{keep} sequences leading to the closest received vectors form the truly received one. We chose $N_{keep} = 100$, also ensuring a good convergence of the algorithm. These sequences will be used in step 20 to update $f_{\mathbf{m}}$ by reinforcing the weights on symbols that have been generated in the set of selected sequences. A parameter δP is needed for this purpose and is empirically set to 0.003, which has been shown to be a good compromise. A larger value increases the convergence speed but also the number of wrong decisions. The process is illustrated in Fig. 3 and Fig. 4, where we show the CEM values of $f_{\mathbf{m}}$ per user after 15 and 30 iterations.

In practice, the CEM-MUD still incurs too high a computational complexity because it needs to generate many samples and keep generating for several rounds in order to converge to the true values. Therefore, we propose a new multiuser approach based on peak detection and collision studies.

Remark: We implemented in algorithm 1 the full MUD. In fact, when implemented at a specific receiver, the algorithm can simply focus on the desired user. The stopping rule can then be adapted (only the probability concerning the desired user has to exceed the threshold).

Algorithm 1 CEM-MUD for downlink LoRa.

Input: Received vector $\mathbf{Z}_q^{(j)}$, $\hat{h}^{(j)}$, N
Output: Decoded vector $\mathbf{m}_q$

1: $N_{seq} \leftarrow 2000$ (*Number of generated sequences*)
2: $N_{keep} \leftarrow 100$ (*Number of selected sequences*)
3: $\delta P \leftarrow 0.003$ (*Update parameter*)
4: $[p, l] \leftarrow$ find peak values and indices of $\mathbf{Z}_q^{(j)}$
5: $u \leftarrow$ sort p in descending order
6: $pos \leftarrow l(u(1:N))$ % *Select the N strongest peaks*
7: $f_{\mathbf{m}} \leftarrow N \times 2^{SF}$ matrix with all elements equal to 0.01
8: $f_{\mathbf{m}}(:, pos) \leftarrow 0.05$
9: $f_{\mathbf{m}} \leftarrow f_{\mathbf{m}} / \sum f_{\mathbf{m}}$ ($\sum f_{\mathbf{m}}$ *is the sum of all elements of $f_{\mathbf{m}}$ to normalize and have a probability mass function.*)
10: **while** min of max of each line of $f_{\mathbf{m}}$ less than 0.85 **do**
11: **for** $idx = 1$ to N_{seq} **do**
12: Generate a vector $\tilde{\mathbf{m}}_q$ according to $f_{\mathbf{m}}$
13: Generate the source vector $\tilde{\mathbf{X}}_q$ from $\tilde{\mathbf{m}}_q$ and (3)
14: $\tilde{\mathbf{R}}_q^{(\mathbf{j})}[n] \leftarrow \hat{h}^{(j)} \tilde{\mathbf{X}}_q[n]$
15: $\tilde{\mathbf{Z}}_q^{(j)}[k] \leftarrow \mathrm{Re}\{FFT(\hat{h}^{(l)*} \tilde{\mathbf{R}}_q^{(\mathbf{j})}[n])\}$
16: $d(idx) \leftarrow \|\mathbf{Z}_q^{(j)} - \tilde{\mathbf{Z}}_q^{(j)}[k]\|^2$
17: **end for**
18: $d \leftarrow$ sort d in ascending order
19: Keep N_{keep} sequences with the smallest d
20: Update $f_{\mathbf{m}}$ by adding δP at each position given by the selected sequences and normalizing.
21: **end while**
22: **return** $\tilde{\mathbf{m}}_q$ corresponding to the smallest d.

2) Proposed Low-Complexity Multiuser Detector

To reduce the complexity we propose a simpler and more direct strategy. In fact, when a good number of peaks is found and because the power allocation scheme is known, it is straightforward to find the one corresponding to the desired user. The CSS in fact allows an interferer avoidance scheme in the Fourier domain at the receiver. In such a case, it is not necessary to cancel interference from stronger users. However, collisions, when two users transmit the same information at the same time, make things more difficult, and if not handled properly, may significantly limit the capabilities of the approach. Our proposed algorithm has the following two steps:

1) *Peak detection using a threshold:* the goal is to find the peaks including the one from the desired user and the larger ones. For instance, if the desired user is user j (users are ordered from the strongest to the weakest allocated power), we define a low enough threshold that will allow us to detect the j strongest peaks but high enough such that the weaker peaks will not be detected.
2) *Peak selection using a search method:* if exactly j peaks are detected, we choose the closest one from the expected received amplitude and its position gives the information of the desired user. Similarly, if more than j peaks are detected, weaker peaks have probably collided and we choose the closest one from the expected received amplitude. Finally, if we detect less than j peaks, it means that collisions occurred between the j strongest users. In that case, we analyze all the possible

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3034973, IEEE Access

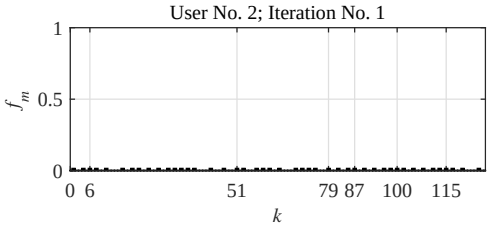


FIGURE 2: CEM-MUD as per Algorithm 1, with $N = 8$ users and $\mathrm{SF} = 7$. This figure represent $f_{\mathbf{m}} = \mathbb{P}(m = k)$ at the beginning of the first iteration. Only one user is represented because the PMF $f_{\mathbf{m}}$ is the same for all users.

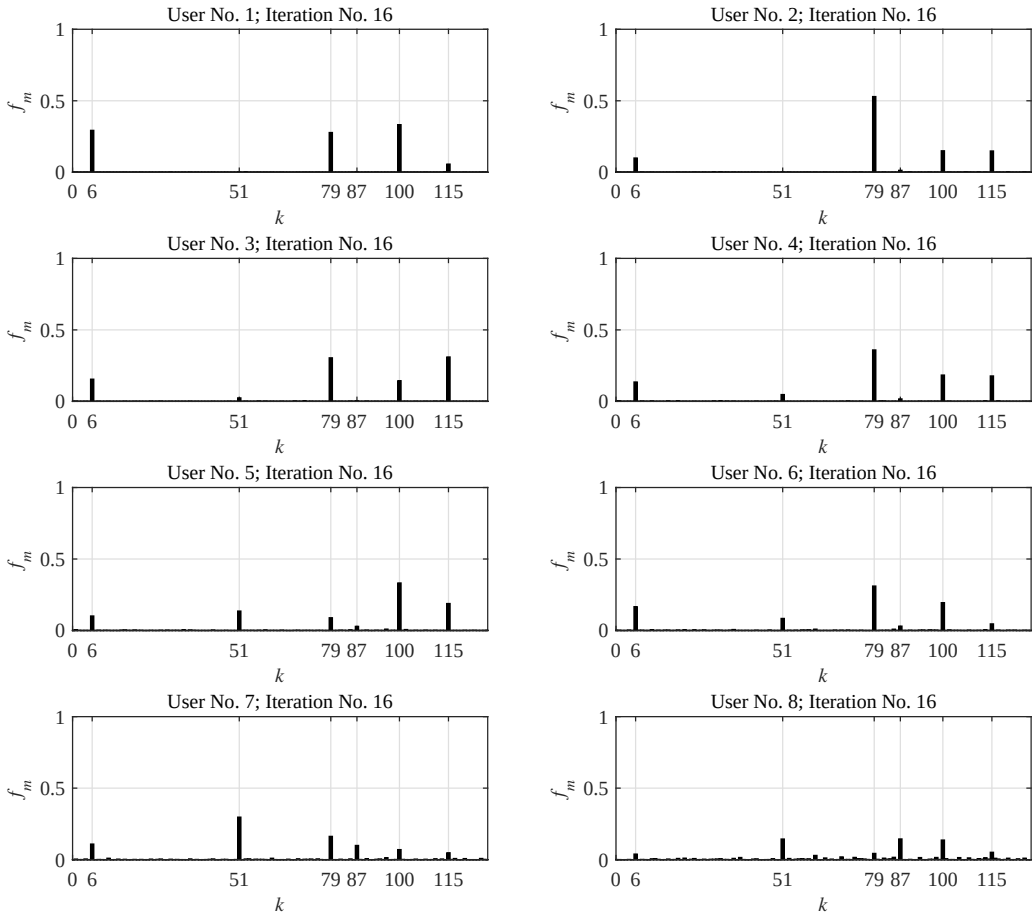


FIGURE 3: CEM-MUD as per Algorithm 1, with $N = 8$ users and $\mathrm{SF} = 7$. In this example two collisions occur ($\mathbf{m}_q = \{6, 79, 115, 79, 100, 100, 51, 87\}$). The figures represent $f_{\mathbf{m}}$ at the beginning of the 16th iteration. We can see that if values tend to gain in probability, the decision cannot be made yet.

collisions cases to choose the most likely and make a decision.

Both steps (choosing the threshold and making a decision when less than j peaks are detected) will be described in the following.

Remark: In the second step when we detect exactly j peaks, there is a case where we can miss the information of the desired user, that is when two or more weaker users collided and result in a peak larger than or equal to the desired user's peak but this desired peak is below the threshold or also collided with another peak. This case is very rare and can be generally avoided by the power allocation scheme described in III-D1.

Threshold definition: Recall that we have N users ordered from the strongest allocated power to the weakest one. We consider user j that we want to decode. We fix a threshold in order to detect the j strongest peaks but not the $N - j$ weakest ones. The expected amplitude for user i at receiver j is $|h^{(j)}|\sqrt{p^{(i)}}$. The objective for choosing the optimal threshold is by maximizing

$$p_1 = \mathbb{P}\left(|h^{(j)}|\sqrt{p^{(j)}} + W_p^j[k] > \gamma\right),$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3034973, IEEE Access

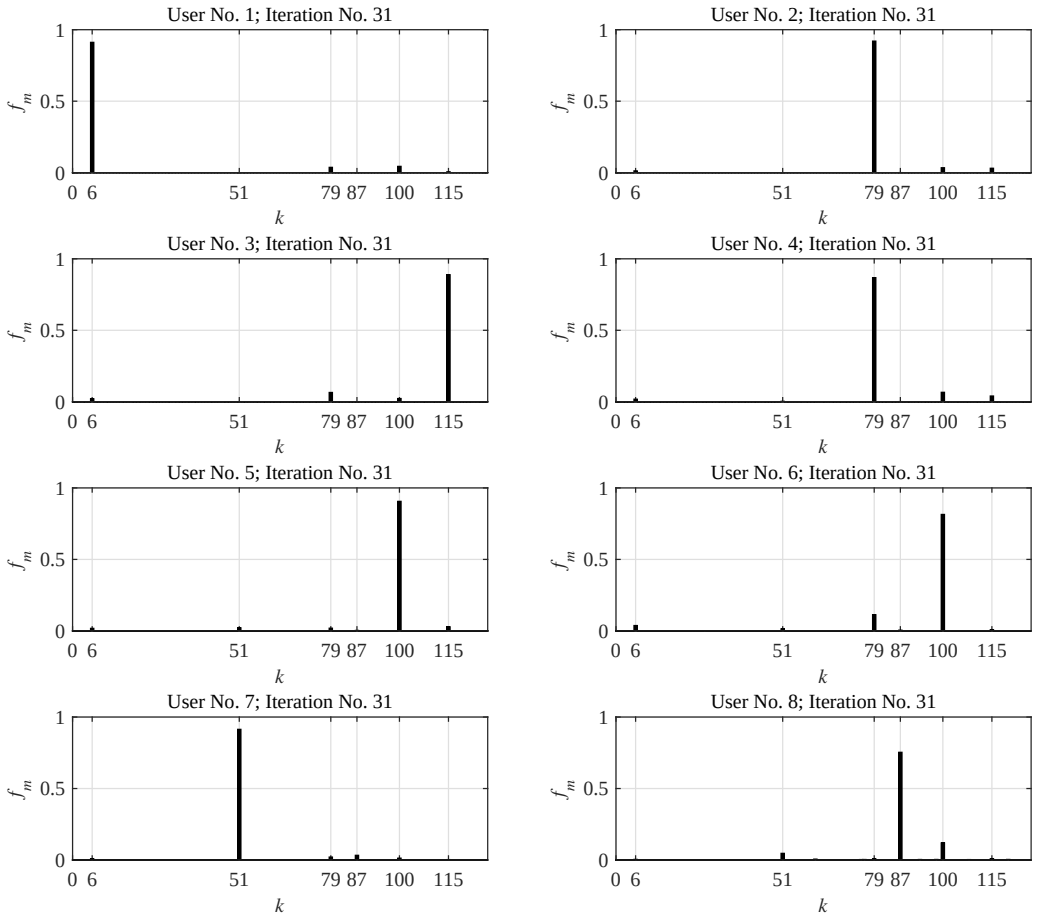


FIGURE 4: CEM-MUD as per Algorithm 1, with $N = 8$ users and $\mathrm{SF} = 7$. In this example two collisions occur ($\mathbf{m}_q = \{6, 79, 115, 79, 100, 100, 51, 87\}$). The figures represent $f_{\mathbf{m}}$ at the beginning of the 31st iteration. We can see that for each user the probability is clearly converging towards a single value equal to one, all the others being zero. We also see that the decision is good despite the two collisions that occur between users 2 and 4 and users 5 and 6.

and, at the same time, to minimize

$$p_0 = \mathbb{P}\left(|h^{(j)}|\sqrt{p^{(j+1)}} + W_p^j[k'] > \gamma\right).$$

where $W_p^j[k]$ is the Gaussian noise and k and k' denote the frequencies where the peaks fall. If user j is the weakest one ($j = N$) we set $p^{(N+1)} = 0$.

$$\gamma_* = \underset{\gamma\in\gamma}{\operatorname{argmax}} \left[\mathbb{P}\left(|h^{(j)}|\sqrt{p^{(j)}} + W_p^j[k] > \gamma\right), \right.$$
$$\left. 1 - \mathbb{P}\left(|h^{(j)}|\sqrt{p^{(j+1)}} + W_p^j[k'] > \gamma\right)\right] \tag{13}$$

We have

$$p_1 = \mathsf{Q}\left(\frac{\gamma - h^{(j)}\sqrt{p^{(j)}}}{\sigma_n}\right), \tag{14}$$

$$p_0 = \mathsf{Q}\left(\frac{\gamma - h^{(j)}\sqrt{p^{(j+1)}}}{\sigma_n}\right), \tag{15}$$

where $\mathsf{Q}(.)$ is the Q-function [36]. We have consequently a multi-objective optimization problem and the Pareto front is easy to obtain. As a starting point, we will choose

$$\gamma = h^{(j)}\frac{\sqrt{p^{(j+1)}} + \sqrt{p^{(j)}}}{2}, \tag{16}$$

but this parameter could be optimized in order to give more weight to p_1.

Making a decision: Let N_{pk} be the number of peaks above the threshold γ. The number of expected peaks is $N_{exp} = j$. The decision rule is the following:

- If $N_{pk} \geq N_{exp}$, we assume no collision between strong peaks and select the peak that has the closest value to the expected one.
- If $N_{pk} < N_{exp}$, we assume a collision occurred. We scan for all possible combinations between the users 1 to j. Let $\breve{m}$ be such a combination. We create a vector adding the amplitudes of the peaks that collide. We ordered the resulting values (including those that did not collide) and calculate the Euclidean distance with the ordered detected peaks. Scanning all possible combinations, we minimize the euclidean distance in

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3034973, IEEE Access

(12) to select the most probable one and deduce the estimated information of the desired user.

Algorithm: The resulting proposed solution is given in algorithm 2.

Step 1 determines the threshold and step 2 finds the peaks above the threshold. Steps 3 and 4 initialize the value of the decoded symbol (m_0) and the minimum distance (M_0).

Steps 5 to 16 are for the case when the number of peaks is less than the threshold: collision occurred. In that case, we scan all possible combinations giving the calculated number of collisions. If the vector received for a generated sequence is closer to the truly received vector (distance less than M_0), we update the values of m_0 and M_0. The number of combinations to scan can become important when the number of users and the number of collisions get large (see table 1) and alternative strategies should be found for these rare events.

In the case when no collision is detected (the number of detected peaks is at least the one we expected), we choose the peak with the closest amplitude to the one we expect (step 18).

Algorithm 2 Proposed Multiuser Detection Algorithm (single symbol)

Input: Received vector $\mathbf{Z}_q^{(j)}$, $\hat{h}^{(j)}$, N, $p^{(j)}$, $p^{(j+1)}$
Output: Decoded symbol $m_q^{(j)}$
1: Calculate γ according to (16)
2: $u \leftarrow \mathbf{Z}_q^{(j)} > \gamma$ *%Get the position of peaks larger than* γ.
3: $m_0 \leftarrow 0$ *% Value of the selected combination*
4: $M_0 \leftarrow \|Z[u]\|^2$ *% Initialize the distance*
5: **if** $length(u) < j$ **then**
6: **while** another combination of $j - length(u)$ collisions exist **do**
7: Choose a combination $\tilde{\mathbf{m}}_q$ with m the symbol of user j
8: Generate the source vector $\tilde{\mathbf{X}}_q$ from $\tilde{\mathbf{m}}_q$ and (3)
9: $\tilde{\mathbf{R}}_q^{(j)}[n] \leftarrow \hat{h}^{(j)}\tilde{\mathbf{X}}_q[n]$
10: $\tilde{\mathbf{Z}}_q^{(j)}[k] \leftarrow \mathrm{Re}\{FFT(\hat{h}^{(j)*}\tilde{\mathbf{R}}_q^{(j)}[n])\}$
11: **if** $\|\mathbf{Z}_q^{(j)}[u] - \tilde{\mathbf{Z}}_q^{(j)}[u]\|^2 < M_0$ **then**
12: $m_0 \leftarrow m$
13: $M_0 \leftarrow \|\mathbf{Z}_q^{(j)}[u] - \tilde{\mathbf{Z}}_q^{(j)}[u]\|^2$
14: **end if**
15: **end while**
16: $m_q^{(j)} \leftarrow m_0$
17: **else**
18: $m_q^{(j)} \leftarrow \underset{u}{\mathrm{argmin}} \left(\|Z_q^{(j)}[u] - |\hat{h}^{(j)}|^2\sqrt{p^{(j)}}\|^2 \right)$
19: **end if**
20: **return** $m_q^{(j)}$

3) Direct Peak Detection

As a reference, we also implement a detection based on finding the peak with the closest amplitude to the expected one. This receiver is the simplest one but does not provide the means to resolve collisions. The decoding process with N users is described as follows: search for the N strongest peaks; select the peak that minimizes the distance between its amplitude and the expected value $|\hat{h}^{(j)}|\sqrt{p^{(j)}}$.

D. POWER ALLOCATION SCHEME

The power allocation allows us to differentiate the different users at the receiver. In order to optimize the NOMA scheme, we attribute the largest power to the user with the worst channel [37], [38]. We can then allocate power based on two different objectives: (*a*) to suppress ambiguities when collisions occur and (*b*) increase fairness between the different users.

1) Power Allocation 1: Suppressing ambiguities

The goal is to avoid that colliding users give rise to a peak with an amplitude equal to another user or a combination of other colliding users. We first order the users from the strongest to the weakest based on an estimate of the channels from a previous uplink ($|\hat{h}^{(1)}| > |\hat{h}^{(2)}| > \cdots > |\hat{h}^{(N)}|$). For user j we allocate the power

$$p_j = \frac{2^{j-1}}{\sum_{i=1}^{N} 2^{i-1}} p_t, \tag{17}$$

where p_t is the total power transmitted by the access point. This guarantees that whatever collision occurs two peaks can not have the same amplitude at the receiver side. However, this results in a significant proportion of the available power to be allocated to further users and the amplitude difference gets small for the closest ones.

2) Power Allocation 2: Fair Spacing

We order the users based on an estimate of the channels from a previous uplink from the weakest to the strongest ($|\hat{h}^{(1)}| < |\hat{h}^{(2)}| < \cdots < |\hat{h}^{(N)}|$). We note $h^{(0)} = 0$. We want that whatever the receiver j, the gap between the peak amplitudes of user j and of the weaker user $j-1$ has always the same value c,

$$|\hat{h}^{(j)}|\sqrt{p^{(j)}} - |\hat{h}^{(j)}|\sqrt{p^{(j-1)}} = c. \tag{18}$$

We show in appendix **A** that the power allocated to user j in that case is

$$p^{(j)} = p_t \frac{A_j^2}{\sum_{i=1}^{N} A_i^2}, \tag{19}$$

where

$$A_i = \sum_{l=1}^{i} \frac{1}{|\hat{h}^{(l)}|}. \tag{20}$$

The drawback of this algorithm is that it depends on the channel coefficients. The access point can estimate them on previous uplinks but the quality of the scheme will depend on the time coherence of the channels. We assume the time coherence of the channels is long in many static IoT cases but the impact of time evolution needs to be further investigated. Theoretical performance analysis as the one in [39] for the uplink would have been a significant added value. However, contrary to other works on NOMA, performance is condi-

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3034973, IEEE Access

tioned on collisions and gaps between allocated powers rather than on Signal to Interference Ratio (SIR), where interference is the sum of all interfering signals. It introduces combinatorial problems (probability of collisions, who collides, and so on) but we have not yet solved the problem.

IV. RESULTS

This section has four main objectives: first, we show that the MUD allows us to significantly increase the capacity of a LoRa network. Second, we confirm that the proposed algorithm delivers performance close to the CEM-MUD. Third, we check the proposed PA schemes, including the performance of the users according to their position (in terms of channel gain) in the group of users simultaneously addressed. Fourth, we provide an analysis of the computational complexity.

A. SIMULATION SETUP

We define a circular region inside which users are uniformly distributed. The radius of the circle is $R = 4,\ 5,\ 10$ km for SF $= 7,\ 8,\ 10$ respectively, and we choose the channel attenuation coefficient $\eta = 3.5$. A complex Gaussian random variable with mean 0 and variance 1 is drawn for each user to model the Rayleigh fading and obtain the channel coefficient as described in section III-B: $h^{(i)} = {d^{(i)}}^{-\eta/2} \cdot \chi^{(i)}$.

However, to be connected to the network with a given SF, the received power in the uplink has to be greater than the receiver sensitivity R_s (we assume a transmit power of 14 dBm). For instance, $R_s = -121.5,\ -124,\ -129$ dBm for SF $= 7,\ 8,\ 10$ respectively, and $B = 250$ kHz. Users that do not respect this condition are discarded and drawn again until N devices are connected. We consider that the access point has a sufficient number of users to address, but no selection scheme is implemented (users are independent and identically distributed). We fix this number and the transmitted packet is the superposition of all corresponding signals.

The noise level of the receiver, as discussed in Section II-C, is -114 dBm. The complete chirp spreading is simulated in baseband and the coding of the information is done for all users. We use a total transmit power of 27 dBm.

Monte Carlo simulations are used to evaluate the performance of the proposed scheme. The parameter settings for the cross entropy decoder are as follows: we generate $N_{seq} = 2000$ sequences and keep $N_{keep} = 100$ sequences for the probability mass function update.

Packet detection and channel estimation are performed using correlation with the known preamble. Common to all users and transmitted with the full power, it does not generate errors and the channel estimation is accurate. We present the performance evaluation in terms of the average symbol error rate (SER). Error correcting codes and a user selection scheme should be included to derive a higher layer Key Performance Indicator. This remains out of the scope of this paper because we are only interested in the PHY layer.

B. PERFORMANCE OF THE THREE RECEIVERS

Figs. 5a and 5b illustrate the SER of a single user with additive Gaussian noise (different SNR, which corresponds to different radio link quality) in the presence of, respectively, 4 and 9 interfering users. The spreading factor is SF $= 7$, $B = 250$ kHz, and $R_s = -121.5$ dBm. Power allocation scheme 2 was used in these simulations. Slightly different results are obtained with scheme 1 but without changing the trends and conclusions. As a baseline, we also added the single-user case, as a traditional LoRa link would do. It shows the impact of the other users and the presence of erroneous decisions due to peak ambiguities or collisions. At equivalent Signal to Noise Ratio (SNR) we see, as expected, a significant gain is SER (one decade at -15 dB in the case of 5 simultaneous users). The MUD performance remains however attractive.

The CEM-MUD has the best performance. It shows that the multiuser scheme can be implemented and significantly increases the number of users in the networks: 10 simultaneous users can be supported. These brute results are very encouraging, especially because we did not consider any user clustering scheme. If the network is highly populated, which is when our proposal is interesting, we should search for the best way to group users and ensure optimal performance.

The CEM-MUD significantly outperforms the simplest receiver that only searches for the close peak amplitude. This shows that it is important to have a multiuser approach and implement a solution to account for collisions.

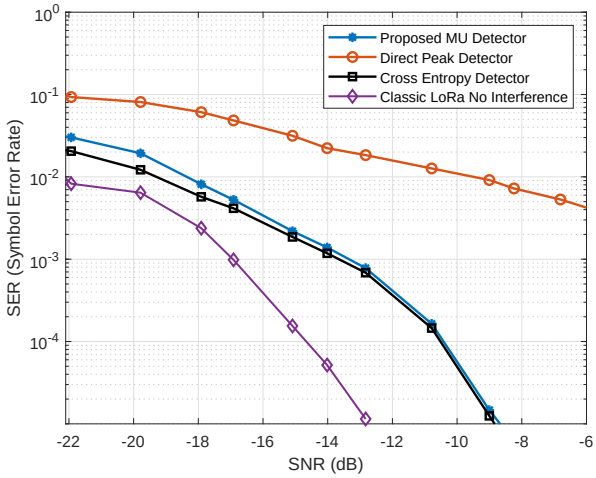
Finally, our proposed receiver and the cross entropy method have almost similar performance, especially at high SNR. This allows a solution for a low complexity implementation of a multiuser detector.

C. POWER ALLOCATION PERFORMANCE

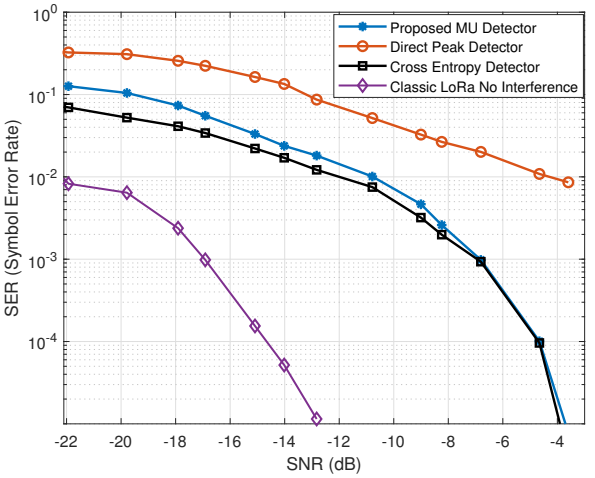
Figs. 6a and 7a show the symbol error rate of a single user with an additive Gaussian noise, SNR$= -10$ and -12 dB respectively, and an increasing number $(N-1)$ of interfering users with $N = 3, \ldots, 13$. In both cases, the power allocation scheme 2 (fair allocation) exhibits better performance. This is more significant in Fig. 7a when the SNR is smaller so that the user is in the weak users. This can be observed from Figs. 6b and 7b, where the histogram of the position of the desired user when users are ordered from the weakest channel (strongest allocated power) to the best channel. For instance, in the case with $N = 12$, it is seen that the mean position is between 3 and 4 in Fig. 7b (SNR$= -12$ dB) when it is between 10 and 11 in Fig. 6b (SNR$= -10$ dB). This latter case sees a larger benefit with the second power allocation scheme, which comes from the fairness approach and the fact that for the first allocation scheme, the gap between amplitudes is small for the users with the good channels.

This analysis is confirmed in Figs. 8 and 9 where the average symbol error rate is plotted for different SF (SF $= 7, 8, 9,$ and 10) and the two power allocation schemes. These two plots also show that the proposed MUD exhibits good performance whatever the SF. As a comparison, the actual

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3034973, IEEE Access

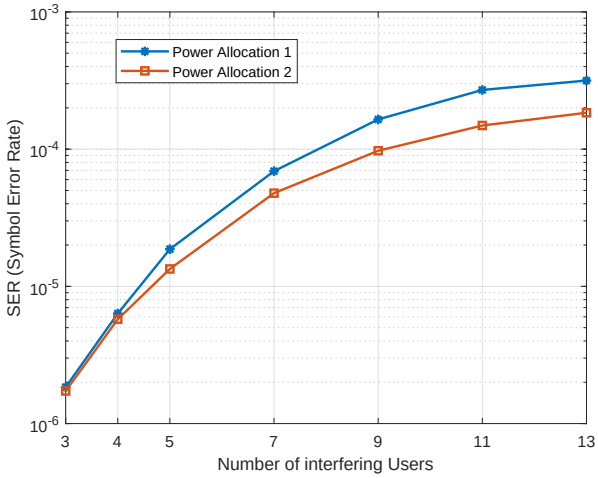


(a) Performance of a single user in the presence of 4 interfering users.

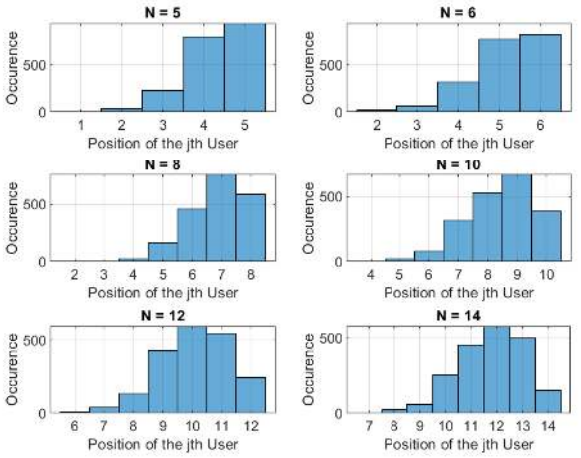


(b) Performance of a single user in the presence of 9 interfering users.

FIGURE 5: SER for different SNR values of a single user, when SF = 7, and $B = 250$ kHz.



(a) SER of a single user in the presence $N - 1$ interfering users.



(b) Position of the user of interest (jth) compared to the interfering users.

FIGURE 6: Performance of the proposed MUD when SNR $= -10$ dB, SF = 8, and $B = 250$ kHz.

implementation of LoRa allows us to address only one user at a time. This means that we can increase by an order of magnitude the number of users that can be addressed in a single time slot.

D. FAIRNESS

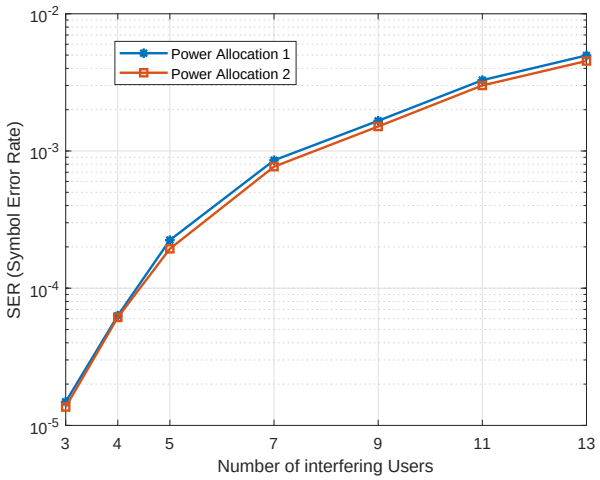
The previous section analyzed the mean error (among all users) but did not consider the SER difference between users with a good or a bad channel.

The SER difference between individual users is shown in Fig. 10. Users are randomly chosen at each round but errors are added depending on their order (from the worst to the best channel). User one for instance is the one with the worst channel, so the one with the highest allocated power. Both power allocation schemes are considered, and we use $N = 7$ and 13. In the first power allocation scheme, ambiguities are avoided but the gap between allocated powers is rapidly decreasing and does not take the channel into account. For the last users (in fact users with the best channel) the gap between allocated powers is not sufficient resulting in an increased SER. On the contrary, the second scheme offers a fairer allocation, slightly increasing the SER for the far users (the first ones) but keeping a more constant value and significantly improving performance in comparison to the other allocation scheme for the last users.

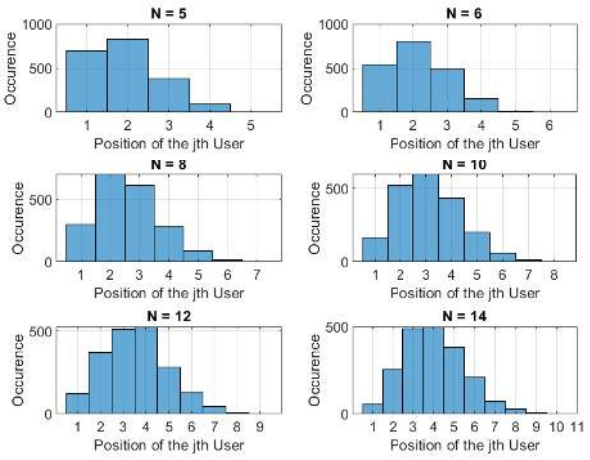
E. COMPUTATIONAL COMPLEXITY ANALYSIS

We compare the computational complexity of the three decoding algorithms namely Direct peak detection, Cross Entropy, and proposed receivers. Our evaluation is considering

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3034973, IEEE Access



(a) SER of a single user in the presence $N-1$ interfering users.



(b) Position of the user of interest (jth) compared to the interfering users.

FIGURE 7: Performance of the proposed MUD when SNR $= -12$ dB, SF = 8, and $B = 250$ kHz.

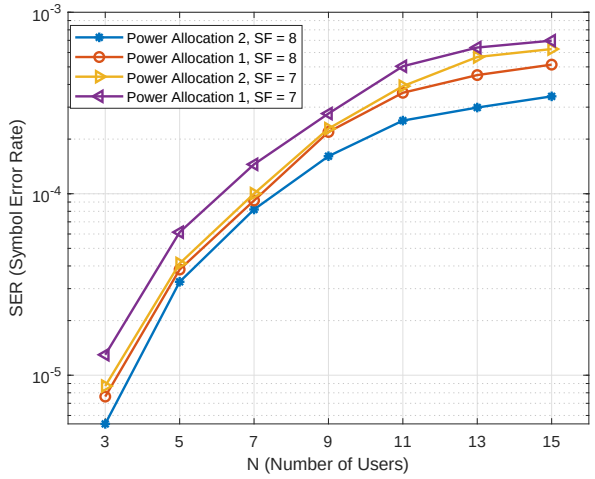


FIGURE 8: SER of the proposed MUD for different N, Noise level of -114 dBm, SF=7, 8, and $B = 250$ kHz.

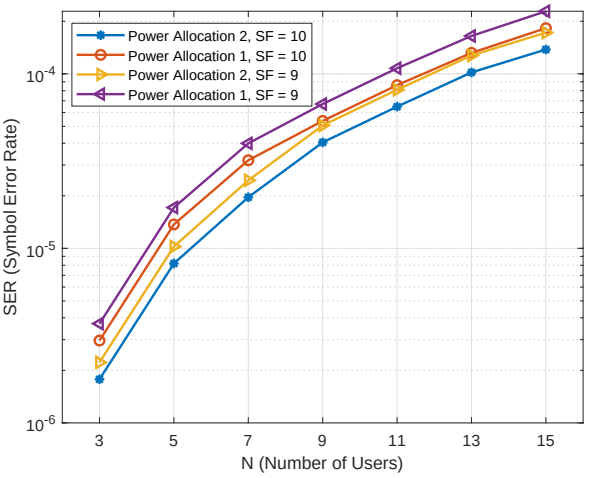


FIGURE 9: SER of the proposed MUD for different N, Noise level of -114 dBm, SF=9, 10, and $B = 250$ kHz.

the decoding of a single symbol. This study gives a broad idea of the complexity but we did not try to optimize the algorithm implementation.

1) Direct peak detection

For each symbol, we repeat the following operations:

1) Multiply by the downchirp (2^{SF} multiplications),
2) FFT (complexity $\mathcal{O}(2^{SF}.\log(2^{SF}))$.

The final complexity for this simple receiver is $\mathcal{O}(2^{SF}.\log(2^{SF}))$.

2) CEM

If we assume that generating a sample from a given PMF and updating a PMF are both $\mathcal{O}(1)$, the complexity of the CEM algorithm depends on the number of operations that are iterated. These operations as depicted in algorithm 1 and their associated complexity for a single symbol are

1) Line 12 (Generate $\tilde{\mathbf{m}}_q$) is $\mathcal{O}(N)$
2) Line 13 (Generate the source vector $\tilde{\mathbf{X}}_q$) is $\mathcal{O}(N.2^{SF})$
3) Line 14 ($\tilde{\mathbf{R}}_q^{(\mathbf{j})}[n] \leftarrow \hat{h}^{(j)}\tilde{\mathbf{X}}_q[n]$) is $\mathcal{O}(2^{SF})$
4) Line 15 (FFT) is $\mathcal{O}(2^{SF}\log(2^{SF}))$
5) Line 16 (distance) is $\mathcal{O}(2^{SF})$

The listed operations are repeated $N_{seq}.N_{ite}$ times, where N_{ite} is the number of iteration before convergence which was observed to be around 30. The PMF updating involves only additions and one normalization and will be $\mathcal{O}(N_{keep})$.

Therefore, the overall complexity is dominated by the 2^{SF} repeated N times in line 13 and $\log(2^{SF})$ in the FFT calculation done for each tested sequence. These figures are of similar order and fixed. In the end we can state that the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3034973, IEEE Access

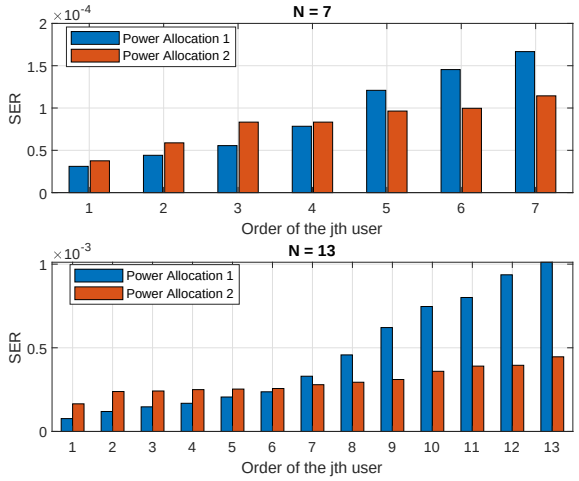


FIGURE 10: SER of individual users when $N = 7$ and 13, SF = 8, Noise level of -114 dBm, and $B = 250$ kHz.

complexity of the CEM is $\mathcal{O}(N_{seq}.N_{ite}.N.2^{SF})$.

3) Proposed method

In our proposal, after peak detection, the complexity is equal to the first approach as long as the number of the detected peaks is equal to or larger than the number of expected peaks.

In the case where fewer peaks are detected, the complexity will be given by the FFT operation that we will need to repeat M_c times, M_c being the number of possible collision combinations we have to scan (steps 6-15 in algorithm 2). For a reasonable number of users and collisions, this number remains low (see table 1). So, finally, the overall complexity is $\mathcal{O}(\mathbb{E}[M_c].2^{SF}\log(2^{SF}))$.

In fact, the time needed to decode a symbol is random and depends on M_c, which is directly linked with the number of users N and collisions N_c occurring. It is easily seen that the probability to have no collision with N users independently and uniformly selecting a number in $\{1, ..., 2^{SF}\}$ is

$$\mathbb{P}\{No\ Collision\} = \frac{\prod_{i=0}^{N-1}(2^{SF} - i)}{2^{N.SF}} \qquad (21)$$

It is, for instance, 0.697 when $SF = 7$ and $N = 10$ and goes up to 0.961 when $SF = 8$ and $N = 5$. In the case when a collision occurs, the number of combinations to be scanned is presented in table 1 and can become very large in the rare cases when N is larger than 10 and N_c larger than 4 for instance.

4) Comparison

It is clearly seen from the previous analysis that the CEM is much more complex than the two others. Our proposed approach is only complex in some rare cases when the number of considered users is large and several collisions occur. Probably it will not be possible to implement these cases in low-cost devices and alternative strategies to address them will be necessary.

To complement our study, table 2 compares the mean computational time R_t (in seconds) of the three receivers, which is required to decode a single user's packet when the length of the packet is $Q = 100$ symbols, and $B = 250$ kHz, and SNR $= -12.5$ dB. To evaluate the differences we used a MATLAB-based software implementation of the digital part of the receiver on a standard computer, but it gives us a general idea about the computational complexity.

TABLE 1: Number of different combinations M_c when N_c collisions occur leading to N_{pk} observed peaks when N_{exp} were expected.

N \ N_c	1	2	3	4	5
1	0	0	0	0	0
2	1	0	0	0	0
3	3	1	0	0	0
4	6	7	1	0	0
5	10	25	15	1	0
6	15	65	90	31	1
7	21	140	350	301	63
8	28	266	1050	1701	966
9	36	462	2646	6951	7700
10	45	750	5880	22827	42525
11	55	1155	11880	63987	179487
12	66	1705	22275	159027	627396
13	78	2431	39325	359502	1899612
14	91	3367	66066	752752	5135130
15	105	4550	106470	1479478	12662650

TABLE 2: Average computational time R_t (in second) to decode a single user (jth) when SF $= 7$, and N users.

N	jth user position	Direct Peak	Proposed	CEM-MUD
5	5	0.0046	0.0059	5.5179
10	8	0.0062	0.0114	17.3731
15	13	0.0077	0.0233	32.6277
20	16	0.0093	0.4254	52.3609

As expected, the ordered peak receiver is much faster at the price of significantly degraded performance. On the other hand, the CEM-MUD has much better performance but the complexity is highly increased. The computation time is multiplied by 1200 for 5 users and the coefficient increases with the number of users.

Our proposed receiver has performance close to the CEM-MUD with a significantly reduced computation time. Indeed, this time is more than 900 times less than the CEM-MUD except when the number of simultaneous packets becomes large (20 in table 2).

V. CONCLUSION

In this paper, we proposed a new joint multiuser receiver for downlink LoRa networks to face the scalability issue in the downlink. Our proposal is inspired by power domain NOMA but does not require the use of a SIC receiver. Doing so, we avoid the limitation of NOMA resulting from the residues

remaining after each cancellation. A second concern for using NOMA in IoT is the low-cost end devices that cannot support computational complexity. The ML does not give an analytical solution and cannot be used. We proposed a sub-optimal approach based on the CEM, which is efficient but remains too complex. Consequently, we derived a simplified method that allows resolving collision but keeps complexity low. We have shown that instead of one single packet per time slot, we could transmit more than 10 packets per time slots and even more, keeping the symbol error rate (without error correcting codes) below 10^{-3}.

.

APPENDIX A POWER ALLOCATION 2

Let us consider N users with estimated channels $\hat{h}^{(i)}$, $i = 1, \cdots, N$ order from the weakest to the strongest channel: $|\hat{h}^{(1)}| < |\hat{h}^{(2)}| < \cdots < |\hat{h}^{(\mathrm{N})}|$. We note $\hat{h}^{(0)} = 0$. We want that whatever the receiver j, the gap between the peak amplitudes of user j and $j-1$ is always the same, c. We denote p_t the total transmit power. We write $h^{(i)} = |\hat{h}^{(i)}|$ for the rest of this annex for lighter expressions. We want

$$\begin{cases} h^{(1)}\sqrt{p^{(1)}} - 0 & = c \\ h^{(2)}\sqrt{p^{(2)}} - h^{(2)}\sqrt{p^{(1)}} & = c \\ \vdots & \\ h^{(\mathrm{N})}\sqrt{p^{(\mathrm{N})}} - h^{(\mathrm{N})}\sqrt{p^{(\mathrm{N}-1)}} & = c \\ p^{(1)} + p^{(2)} + p^{(3)} + \cdots + p^{(\mathrm{N})} & = p_t \end{cases}$$

We note

$$A_i = \left(\sum_{j=1}^{i} \frac{1}{h^{(j)}}\right)^2, \tag{22}$$

Starting with the weakest user ($i = 1$) we have

$$p^{(1)} = \left(\frac{c}{h^{(1)}}\right)^2 = c^2 A_1, \tag{23}$$

If we assume

$$p^{(l)} = c^2 A_l, \tag{24}$$

then from

$$h^{(l+1)}\sqrt{p^{(l+1)}} - h^{l+1)}\sqrt{p^{(l)}} = c, \tag{25}$$

and using (24), we have

$$\begin{aligned} \sqrt{p^{(l+1)}} &= \frac{c}{h^{(l+1)}} + \sqrt{p^{(l)}}, \\ &= c\left(\frac{1}{h^{(l+1)}} + \sum_{j=1}^{l} \frac{1}{h^{(j)}}\right), \\ &= c\left(\sum_{j=1}^{l+1} \frac{1}{h^{(j)}}\right), \\ &= c\sqrt{A_{l+1}}. \end{aligned} \tag{26}$$

which proves that (24) is true for all $i = 1, \cdots, N$.

We then use the final equation in (22)

$$P_t = \sum_{j=1}^{N} p^{(j)} = c^2 \sum_{j=1}^{N} A_j, \tag{27}$$

which gives

$$c = \sqrt{\frac{p_t}{\sum_{j=1}^{N} A_j}}. \tag{28}$$

We finally obtain $p^{(l)} = p_t \dfrac{A_l^2}{\sum_{i=1}^{N} A_i^2}$ which is (19).

ACKNOWLEDGMENT

This work has been partially supported by IRCICA USR CNRS 3380, COST ACTION CA15104 - IRACON and the French National Agency for Research (ANR) under grant ANR-16-CE25-0001 - ARBURST.

REFERENCES

[1] L. Knud Lasse, "State of the iot 2018: Number of iot devices," Aug 2018, (Accessed on 05/22/2020). [Online]. Available: https://iot-analytics.com/state-of-the-iot-update-q1-q2-2018-number-of-iot-devices-now-7b/

[2] N. Shah and S. Sundar, "Smart electric meter using lora protocols and lot applications," in 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA), 2018, pp. 1178–1180.

[3] Y. Bagariang, M. I. Nashiruddin, and N. Mufti Adriansyah, "Lora-based iot network planning for advanced metering infrastructure in urban, sub-urban and rural scenario," in 2019 International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), 2019, pp. 188–193.

[4] R. Vishnubhotla, P. S. Rao, A. Ladha, S. Kadiyala, A. Narmada, B. Ronanki, and S. Illapakurthi, "Zigbee based multi-level parking vacancy monitoring system," in 2010 IEEE International Conference on Electro/Information Technology, 2010, pp. 1–4.

[5] Semtech, "Iot-enabled smart parking management: Lora technology application brief," (Accessed on 05/23/2020). [Online]. Available: https://www.semtech.com/uploads/technology/LoRa/app-briefs/Semtech_Cities_ParkingManagement_AppBrief.pdf

[6] J. Jose and T. Sasipraba, "Indoor air quality monitors using iot sensors and lpwan," in 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), 2019, pp. 633–637.

[7] R. K. Kodali, S. Yerroju, and S. Sahu, "Smart farm monitoring using lora enabled iot," in 2018 Second International Conference on Green Computing and Internet of Things (ICGCIoT), 2018, pp. 391–394.

[8] K. I. Wang, S. Wu, A. Ivoghlian, Z. Salcic, A. Austin, and X. Zhou, "Lws: A lorawan wireless underground sensor network simulator for agriculture applications," in 2019 IEEE SmartWorld, Ubiquitous Intelligence Computing, Advanced Trusted Computing, Scalable Computing Communications, Cloud Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI), 2019, pp. 475–482.

[9] C. Pham, A. Bounceur, L. Clavier, U. Noreen, and M. Ehsan, "Radio channel access challenges in lora low-power wide-area networks," in LPWAN Technologies for IoT and M2M Applications, B. S. Chaudhari and M. Zennaro, Eds. Academic Press, Jan 2020, pp. 65–102.

[10] O. Georgiou and U. Raza, "Low power wide area network analysis: Can lora scale?" IEEE Wireless Communications Letters, vol. 6, no. 2, pp. 162–165, 2017.

[11] M. C. Bor, U. Roedig, T. Voigt, and J. M. Alonso, "Do lora low-power wide-area networks scale?" in Proceedings of the 19th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems, ser. MSWiM '16. New York, NY, USA: Association for Computing Machinery, 2016, pp. 59–67. [Online]. Available: https://doi.org/10.1145/2988287.2989163

[12] U. Noreen, L. Clavier, and A. Bounceur, "Lora-like css-based phy layer, capture effect and serial interference cancellation," in European Wireless 2018; 24th European Wireless Conference, 2018, pp. 1–6.

[13] B. Laporte-Fauret, M. A. Ben Temim, G. Ferré, D. Dallet, B. Minger, and L. Fuché, "An enhanced lora-like receiver for the simultaneous reception of two interfering signals," in 2019 IEEE 30th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), 2019, pp. 1–6.

[14] N. E. Rachkidy, A. Guitton, and M. Kaneko, "Decoding superposed lora signals," in 2018 IEEE 43rd Conference on Local Computer Networks (LCN), 2018, pp. 184–190.

[15] M. A. B. Temim, G. Ferré, B. Laporte-Fauret, D. Dallet, B. Minger, and L. Fuché, "An enhanced receiver to decode superposed lora-like signals," IEEE Internet of Things Journal, pp. 1–1, 2020.

[16] A. A. TESFAY, E. P. SIMON, G. FERRÉ, and L. CLAVIER, "Serial interference cancellation for improving uplink in LoRa-like networks," in 2020 IEEE 31st Annual International Symposium on Personal, Indoor and Mobile Radio Communications: Track 1 - PHY and Fundamentals (PIMRC'20 Track 1 PHY & Fundamentals), London, United Kingdom (Great Britain), Aug. 2020, pp. 1–6.

[17] T. Voigt, M. Bor, U. Roedig, and J. Alonso, "Mitigating inter-network interference in lora networks," in Proceedings of the 2017 International Conference on Embedded Wireless Systems and Networks, USA, 2017, p. 323âĂŞ328.

[18] F. Adelantado, X. Vilajosana, P. Tuset-Peiro, B. Martinez, J. Melia-Segui, and T. Watteyne, "Understanding the limits of lorawan," IEEE Communications Magazine, vol. 55, no. 9, pp. 34–40, Sep. 2017.

[19] M. Centenaro, L. Vangelista, and R. Kohno, "On the impact of downlink feedback on lora performance," in 2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), Oct 2017, pp. 1–6.

[20] A. Pop, U. Raza, P. Kulkarni, and M. Sooriyabandara, "Does bidirectional traffic do more harm than good in lorawan based lpwa networks?" in GLOBECOM 2017 - 2017 IEEE Global Communications Conference, 2017, pp. 1–6.

[21] K. Mikhaylov, J. Petaejaejaerv, and A. Pouttu, "Effect of downlink traffic on performance of lorawan lpwa networks: Empirical study," in 2018 IEEE 29th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), Sep. 2018, pp. 1–6.

[22] S. Abboud, N. el Rachkidy, A. Guitton, and H. Safa, "Gateway selection for downlink communication in lorawan," in 2019 IEEE Wireless Communications and Networking Conference (WCNC), 2019, pp. 1–6.

[23] F. Van den Abeele, J. Haxhibeqiri, I. Moerman, and J. Hoebeke, "Scalability analysis of large-scale lorawan networks in ns-3," IEEE Internet of Things Journal, vol. 4, no. 6, pp. 2186–2198, Dec 2017.

[24] V. Di Vincenzo, M. Heusse, and B. Tourancheau, "Improving downlink scalability in lorawan," in ICC 2019 - 2019 IEEE International Conference on Communications (ICC), May 2019, pp. 1–7.

[25] S. M. R. Islam, N. Avazov, O. A. Dobre, and K. Kwak, "Power-domain non-orthogonal multiple access (noma) in 5g systems: Potentials and challenges," IEEE Communications Surveys Tutorials, vol. 19, no. 2, pp. 721–742, 2017.

[26] S. Shieh, C. Lin, Y. Huang, and C. Wang, "On gray labeling for downlink non-orthogonal multiple access without sic," IEEE Communications Letters, vol. 20, no. 9, pp. 1721–1724, 2016.

[27] M. Qiu, Y. Huang, and J. Yuan, "Downlink non-orthogonal multiple access without sic for block fading channels: An algebraic rotation approach," IEEE Transactions on Wireless Communications, vol. 18, no. 8, pp. 3903–3918, 2019.

[28] Patent, "Low power long range transmitter," Jan 2014. [Online]. Available: https://patents.google.com/patent/EP2763321A1

[29] T. M. W. 1.0, "What is LoRaWAN," LoRa Alliance, Tech. Rep., 2015. [Online]. Available: https://lora-alliance.org/resource-hub/what-lorawanr

[30] ETSI, "Etsi en 300 220-1 v2.4.1: Electromagnetic compatibility and radio spectrum matters (erm); short range devices (srd); radio equipment to be used in the 25 mhz to 1 000 mhz frequency range with power levels ranging up to 500 mw; part 1: Technical characteristics and test methods," European Telecommunications Standards Institute (ETSI), Tech. Rep., 2012.

[31] B. Reynders and S. Pollin, "Chirp spread spectrum as a modulation technique for long range communication," in 2016 Symposium on Communications and Vehicular Technologies (SCVT), 2016, pp. 1–5.

[32] L. Vangelista, "Frequency shift chirp modulation: The lora modulation," IEEE Signal Processing Letters, vol. 24, no. 12, pp. 1818–1821, 2017.

[33] Semtech, "AN1200.22: LoRa Modulation Basics," Semtech Corporation, Tech. Rep., 2015.

[34] R. Y. Rubinstein, "Optimization of computer simulation models with rare events," pp. 89– 112, 1997.

[35] D. Ernst, M. Glavic, G. Stan, S. Mannor, and L. Wehenkel, "The cross-entropy method for power system combinatorial optimization problems," in 2007 IEEE Lausanne Power Tech, 2007, pp. 1290–1295.

[36] J. G. Proakis and M. Saleh, Digital Communications 5th Edition. McGraw Hill, 2007.

[37] L. Dai, B. Wang, Y. Yuan, S. Han, C. I, and Z. Wang, "Non-orthogonal multiple access for 5g: solutions, challenges, opportunities, and future research trends," IEEE Communications Magazine, vol. 53, no. 9, pp. 74–81, 2015.

[38] H. Tabassum, M. S. Ali, E. Hossain, M. J. Hossain, and D. I. Kim, "Uplink vs. downlink noma in cellular networks: Challenges and research directions," in 2017 IEEE 85th Vehicular Technology Conference (VTC Spring), 2017, pp. 1–7.

[39] J. M. d. S. SantâĂŹAna, A. Hoeller, R. D. Souza, H. Alves, and S. Montejo-SÃąnchez, "Lora performance analysis with superposed signal decoding," IEEE Wireless Communications Letters, pp. 1–1, 2020.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3034973, IEEE Access

ANGESOM A. TESFAY received the B.Sc. in Electrical & Computer Engineering from University of Mekelle, Ethiopia, and a masters degree in Telecommunications Engineering from University of Trento, Italy, in 2017. He is currently a Ph.D. student in signal processing at the IEMN (CNRS UMR 8520 - *Institute of Electronics, Microelectronics and Nanotechnology*) and IRCICA (CNRS USR 3380 - *Research Institute on software and hardware devices for information and Advanced communication*), University of Lille. His research interests include IoT networks, the physical layer of wireless communications, and sensor networks.

ERIC P. SIMON received the Masters degree in electronics engineering from the Superior School of Electronics (ESCPE), Lyon, France, in 1999, and the Ph.D. degree in signal processing and communications from the National Polytechnic Institute of Grenoble (INPG), France, in 2004. During 2005, he was a Teaching Assistant at the INPG and the following year he joined one of France Telecom R&D Laboratories as a Postdoctoral Fellow. He is currently an Associate Professor at the Institute of Electronics, Microelectronics and Nanotechnology (IEMN), TELICE (Telecommunications, Interference and Electromagnetic Compatibility) Group, University of Lille, France. His main research interests are in mobile communications and carrier and symbol synchronization.

IDO NEVAT received the B.Sc. degree in electrical engineering from the Technion-Israel Institute of Technology, Haifa, Israel, in 1998, and the Ph.D. degree in electrical engineering from the University of New South Wales, Sydney, NSW, Australia, in 2010. Between 2010 and 2013, he was a Postdoctoral Research Fellow with the Wireless and Networking Technologies Laboratory, CSIRO, Australia. Between 2013 and 2016, he was a Scientist with the Institute for Infocomm Research, Singapore. Since 2017, he has been a team leader with TUM-CREATE, Singapore and a PI of the Cooling Singapore project. His main research interests include statistical signal processing, machine learning, and Bayesian statistics.

LAURENT CLAVIER received his Ph.D. degree in signal processing in TELECOM Bretagne (now IMT Atlantique) in Brest, France and the HDR degree from Lille University, France, in 2009. He is, since October 2011, Professor in Mines-Telecom institute (IMT Lille Douai), with IEMN (UMR CNRS 8520 - *Institut d'Electronique de Microélectronique et de Nanotechnologie*) and IRCICA (USR CNRS 3380 - *Institut de Recherche sur les composants logiciels et matériels pour et l' Information et la Communication Avancée*). His research activities concern digital communications and the physical layer of wireless networks for IoT, more specifically energy-autonomous sensor networks. He is particularly interested in the interference model and impact in ultra-dense wireless networks.