Time-frequency ALOHA-like random access: A scalability study of Low-Power Wide-Area Networks for IoT using stochastic geometry

Steeve Zozor*, Zhuocheng Li*,†, Quentin Lampin† and Jean-Marc Brossier*
* GIPSA-Lab
University of Grenoble Alpes, Grenoble, France
Email: (zhuocheng.li,steeve.zozor,jean-marc.brossier)@gipsa-lab.grenoble-inp.fr
† OrangeLabs
Meylan, France
Email: (zhuocheng1.li,quentin.lampin)@orange.com

Abstract—We study the multiple access interferences in a random time-frequency scheme which encompasses the classical pure Aloha protocol. Such an access, additionally to a single hop communication or star topology scheme where sensors communicate with a single base station, allows for a low energy consumption, but the cost to pay is the risk of collision. To finely study the collisions, we propose an approach to model random access schemes based on stochastic geometry, in which the total time-frequency resources is a two-dimensional plane, and transmissions from sensor nodes are represented by time-frequency patterns randomly drawn in the time-frequency plane. We study levels of overlapping between the information packets: provided that the overlap is not too strong, the packets are not necessarily lost when they collide. Fixing a maximal allowed probability of packet loss, we are then able to derive the capacity necessarily lost when they collide. We finally illustrate our approach in an application in the cases of wide-band and ultra-narrow band transmissions.

Index Terms—Random time-frequency access channel; 2D Aloha; network load-capacity; stochastic geometry; IoT; LPWAN.

I. INTRODUCTION

These last decades have seen a huge development of Wireless Sensor Networks (WSN), used in a vast area of applications [1], [2]. Among the various topologies of networks designed to transmit the information from the sensor to a receiver, one can mention Low Power Wide-Area Networks (LPWAN) which will probably be one of the most important technologies of the future to interconnect sensors for the Internet of Things (IoT) or machine to machine (M2M) applications [3]. In these kinds of applications, a large number of sensor nodes deployed at an urban scale generate and transmit a relatively small amount of data directly to the base station or gateway, in a so-called star topology or single-hop communication scheme [2].

Among the various Medium Access Control (MAC) protocols in time, as detailed for instance in [2, chap. 6] or in [4, chap. 6], we concentrate here on that based on random access, also called Aloha. This approach has the advantage to impose less constraints on sensor nodes which are often low-cost, low-complexity devices [5], [6]. The pure aloha access has also the advantage to be low energy consumption since it does not need to request the authorization to emit, it does not need to wait for accessing to the medium, it does not need to sense the medium before sending its information packet, it does not need any synchronization to access the medium at a given time slot [4, 2]. The cost to pay is the risk of collision between information packets [4, 2, 7].

The study of time access find its interest when the whole available frequency band is used by any packet of information, for instance in a spreading spectrum approach [4, 2, 7, 8]. In such a context, due to the jitter of the oscillators [11], [12], the access can be viewed as random in frequency. Indeed, even if the relative uncertainty on the frequency carrier can be tiny, when this carrier is very high frequency, in ultranarrow band transmission, the uncertainty can be of the same order of magnitude as the bandwidth of the packets. Moreover, for an application with a huge number of sensors, one can imagine that the allowed frequency band is not enough to be slotted for all the sensors [13] and thus a pure frequency random access can be privileged (as for the Aloha approach in terms of time access). This approach has also the advantage to avoid sensing the channel to chose a frequency carrier. Again, the cost to pay for random frequency access is the risk of frequency collision, i.e., an overlap between the packets’ bands.

If the collision between packets in the Aloha protocol is well studied [2], as far as we known, a very few studies concern the collision in a joint random time and frequency access [13], [14], [10]. In these works, in general, the frequency access is slotted so that the random access is not “pure”. Moreover, in [10] for instance, the authors fix the time to study the
frequency overlap/collision, so that the problem is somewhat similar to that of the Aloha protocol, except that the study is in the frequency domain instead of in the time domain.

In this paper, we focus on the pure time-frequency access to the medium. To this end, we provide a stochastic geometry model to analyze the resources sharing problem and collision phenomenon. In our model, we represent the total available time-frequency resources by a two-dimensional plane, unlike in common stochastic geometry models, where it’s usually the geographical domain of wireless networks that is modeled by a Euclidean plane [15], [16]. We can imagine slicing time up in periods, thus each period is a limited resource for all the sensors. The medium being also limited in frequency, transmissions from nodes form certain random patterns in the time-frequency plane.

Collision occurs when transmissions overlap in the time-frequency domain. We assume that a collision does not necessarily cause the packets loss (due to the capture effect for instance [17]). Therefore, we define some outage probabilities by fixing thresholds on certain geometrical quantities chosen to reflect the level of interference introduced by the collision. In some contexts, we can assimilate these outage probabilities to the probability of not satisfying some target Signal-to-Interference-plus-Noise Ratio (SINR). In the scenario that we consider, the outage probabilities are intrinsically related to the number of sensor nodes, thus our approach allows to conversely derive the load-capacity of the network, i.e., the maximal achievable load given an acceptable outage probability.

Note that we do not make any restriction on the base station’s capability to recover transmissions when there is a collision, this is related to the choice of different physical layer technologies that we decide not to discuss. In section IV, we describe the simplified model that we use to characterize the time-frequency random access to the medium. Then, in section IV, firstly we introduce some geometrical quantities and a geometry-related SINR expression to represent the interference introduced by collisions. Then, we introduce the corresponding normalized quantities. Finally, we introduce the outage probabilities, i.e., probabilities that these quantities exceed some thresholds. Since these probabilities are related to the number of sensor nodes, by choosing an acceptable outage probability, we can derive conversely the load-capacity of the network. In section V we then focus on the probabilistic study of these quantities, via a stochastic geometrical approach. We characterize the interference introduced by collisions by relating the surface of overlapping patterns to the SINR. We will also provide a general outage formula through which we derive upper and lower bounds of the network capacity. Finally, we give in section V some illustrations of our approach to characterize two specific applications.

II. THE “CARDS TOSSING” MODEL

We consider in the following the scenario of $N \geq 1$ sensor nodes or senders $I_k$ sharing the same limited time-frequency resource to send their readings in packets to a single base station. This scenario applies for instance when the nodes send their information with the same period and a single period is considered.

Let us denote by $I_0$ the user of interest, by $I_k$, $k \geq 1$ the interferers and by $g_k(t,f)$ the packet sent by user $I_k$. To fix the notation, and without loss of generality (provided that appropriate shifts are made), the total time-frequency resources shared among transmissions will be denoted by $[0; T] \times [0; F]$.

In our model we are focusing on the time-frequency supports of the senders’ messages regardless of how they are encoded. We make the following assumptions:

1) We assume that messages from all the senders have the same time-frequency support in the shape of a rectangle, since they are limited in time and bandwidth. We denote by $\Delta t$ the time duration and by $\Delta f$ the occupied bandwidth. Thus $g_k$ takes the form

$$g_k(t,f) = m_k(t,f) 1_{I_k}(t,f),$$

where $1_A$ stands for the indicator function of set $A$ and $I_k = [t_k : t_k + \Delta t] \times [f_k : f_k + \Delta f]$ denotes here the time-frequency support of node $I_k$ (for the sake of simplicity, we use the same notation $I_k$ for the sender itself, its transmission, and the time-frequency support of its message); $t_k$ is the initial time of transmission and $f_k$ the lowest frequency of the information packet (see figures 1 and 2).

2) As we are considering random time-frequency access scenario and we assume that there is no cooperation between the nodes, the couples $(t_k, f_k)$ are independent and identically distributed. We assume that they are continuous with a probability density function (pdf).

3) We assume that, given support $I_k$, the time-frequency energy of the information packet is uniformly distributed over this support, i.e.,

$$\mathbb{E}[|g_k(t,f)|^2 | I_k] = \rho_k 1_{I_k}(t,f)$$

4) The packets sent by the users are assumed to have the same energy density when they reach the base station, i.e.,

$$\rho_k = \rho \quad \forall k$$

5) We assume that the channel is only corrupted by a zero-mean time-frequency white noise $\xi(t,f)$ of energy density denoted $\gamma_0$, i.e.,

$$\mathbb{E}[|\xi(t,f)|^2] = \gamma_0$$

Assumptions 2 and 5 are common in telecommunication and thus reasonable.

Assumption 3 is probably not satisfied in practice. However, the distribution of the energy over $I_k$ is highly dependent on specific applications. Thus, this assumption can also be viewed as reasonable in the sense that our study aims to be as general as possible, regardless a specific technique distribution of the information over $I_k$. 


The load-capacity is intimately linked to the probability that other packets are lost (see figure 1 and 2). In this case, i.e., when their time-frequency supports intersect with each other, collisions between the transmission of interest and one or more interfering transmissions \( I_k \) without considering other channel phenomena that could compromise the transmission.

Intuitive thinking could be that when two packets collide, i.e., when their time-frequency supports intersect with each other, both packets are lost (see figure 1 and 2). In this case, the load-capacity is intimately linked to the probability that \( I_0 \) intersect at least one interferer \( I_k \), i.e., the probability that \( I_0 \cap \bigcup_{k=1}^{N-1} I_k \neq \emptyset \). However, depending on the physical layer protocols chosen by different applications, a small overlap, let us say \( s \), does not necessarily compromise the transmission. Note that a transmission can be overlapped by one or more other interfering transmissions. In the analogy with the game of players tossing cards onto a table, a given player wishes that its own card is not (too much) covered by the other ones so that he can recognize it after the tossing game.

In what follows, we do not discuss how to choose the thresholds nor the acceptable outage probability, which are usually related to technological details. We just provide a general mathematical tool for possible industrial or academic use.

### III. Quantities of Interest and Load-Capacities

As just mentioned, interference is closely linked to the overlap between \( I_0 \) and the interferers. At least two points of view have to be envisaged, depending on the context of the communication protocol. Let us recall again that we do not focus on a particular protocol, so that we do not differentiate the locations where interferers corrupt \( I_0 \) but focus on amount of interference.

#### A. Quantities of interest

To be precise, in the sequel, we focus on the following quantities.

- The surface of the union of the intersections between \( I_0 \) and the interfering transmissions, i.e.,
  \[ S_u = \mu \left( I_0 \cap \bigcup_{k=1}^{N-1} I_k \right), \]
  where \( \mu \) denotes the surface measure. This quantity is interesting for instance when \( I_0 \) supports a "composite" information. For instance, in scenarios where subsurfaces of \( I_0 \) contain information without redundancy between them, that is when a given subsurface is overlapped, all its information is lost.
- The sum of surfaces of intersections between transmission \( I_0 \) and other interfering transmissions, i.e.,
  \[ S_{\Sigma} = \sum_{k=1}^{N-1} S_k \quad \text{with} \quad S_k = \mu \left( I_0 \cap I_k \right). \]
  This quantity is interesting when \( I_0 \) contains for instance an integrated information that is spread over the whole surface of \( I_0 \). Hence the loss of the information is linked to the energy of interference, so we consider the sum instead of the union of the intersections.
- The maximal surface of intersection,
  \[ S_{\max} = \max_k S_k \]
  is also interesting since it bounds the previous quantities as given by the following inequalities
  \[ S_{\max} \leq S_u \leq S_{\Sigma}, \]
  Meanwhile, in scenario where \( N\Delta t \Delta f \ll TF \), one can guess that the probability of collision is weak, and that of having more that one collision is weaker, so we roughly have \( S_{\max} \approx S_u \approx S_{\Sigma} \).
- The SINR, that we define as the ratio between the energy of the message of interest, \( \int_{I_0} \mathbb{E}[|g_0(t,f)|^2] \, df \) and the sum of that of the interferers \( \int_{I_0 \cap I_k} \mathbb{E}[|g_k(t,f)|^2] \, df \) plus noise. This quantity is closely related to \( S_{\Sigma} \) since, from the previous assumptions, we have
  \[ \text{SINR} = \frac{\rho \Delta t \Delta f}{\rho S_{\Sigma} + \gamma_0 \Delta t \Delta f}. \]
  For all quantities, the dependence on \( N \) is omitted in the notation for the sake of simplicity.

\( S_u, S_{\Sigma}, S_{\max} \) and SINR being random, firstly we must evaluate their probability distributions which are necessary to characterize of the load-capacity.

#### B. Normalized quantities

The problem being scale invariant (e.g., invariant by multiplying \( T \) and \( \Delta t \) by the same factor, etc.), it is not the absolute quantities such as \( T, \Delta t \), etc., but the normalized quantities that are of interest.

Therefore we introduce the following normalized quantities,

\[ N_t = \frac{T}{\Delta t}, \quad N_f = \frac{F}{\Delta f}, \quad \eta = \frac{N}{N_t N_f} \]
where $N_f$ and $N_t$ are respectively the number of available frequency carriers and time slots. Thus, $N_tN_f$ represents the maximal achievable load in the network, assuming perfect allocation of each resource fragment, and $\eta$ is thus the loading rate of the total available time-frequency resources.

In what follows, we naturally assume that $N_t \geq 1$ and $N_f \geq 1$ (in the Aloha scenario, $N_f = 1$) and also that $N_tN_f$ is very large.

We normalize the quantities introduced in section II with respect to the surface of the time-frequency support of transmissions

$$X = \frac{S}{\Delta f \Delta T}$$

where $X$ (resp. $S$) stands for $X_k$ (resp. $S_k$), $X_{\cup}$ (resp. $S_{\cup}$), $X_\Sigma$ (resp. $S_\Sigma$) and $X_{\max}$ (resp. $S_{\max}$). Thus, the SINR can be recast as

$$\text{SINR} = \frac{1}{X_\Sigma \text{SNR}^{-1}} \text{ with } \text{SNR} = \frac{\rho}{\gamma_0}$$

where SNR represents the signal-to-noise ratio that characterizes the transmission in the absence of interferences.

C. Performance parameters: load-capacities

Depending on the applications, one may be interested by characterizing the communication for a given loading rate $\eta$, or deriving conversely the loading rate to satisfy certain performance requirements. In the former case, we are interested by the following outage probabilities $\Pr[X > x]$ and $\Pr[\text{SINR} < \zeta]$. In the latter case, we provide the following loading rate capacities

$$\eta_X(x, p) = \arg\max_{\eta} \Pr[X > x] \leq p$$

and

$$\eta_S(\zeta, p, \text{SNR}) = \arg\max_{\eta} \Pr[\text{SINR} < \zeta] \leq p$$

where $x$ (resp. $\zeta$) is supposed to be a chosen threshold (resp. target SINR) and $p$ an acceptable outage probability (the load-capacities being $N_tN_f\eta$ in the above mentions senses).

Note that from relation (4), clearly we have

$$\eta_S(\zeta, p, \text{SNR}) = \eta_X(\zeta - \text{SNR}^{-1}, p)$$

This expression clearly shows that the SNR obviously plays an important role in the derivation of load-capacity in terms of SINR.

From (1) we have the following series of inequalities

$$\Pr[X_{\max} > x] \leq \Pr[X_{\cup} > x] \leq \Pr[X_{\Sigma} > x]$$

(8)

(with equality for $x = 0$, where both probabilities reduced to the probability of collision, and also for $x = 1$ for the first inequality), the loading rate capacities in terms of these three geometrical quantities thus verify the following inequalities

$$\eta_{X_{\max}}(x, p) \geq \eta_{X_{\cup}}(x, p) \geq \eta_{X_{\Sigma}}(x, p)$$

(9)

Finally, note that when the threshold $x$ and the acceptable outage probability $p$ are both small, one can suppose that the probability of having more than two interferers is very low. Therefore, we have $X_{\Sigma} \approx X_{\cup} \approx X_{\max}$. In this situation, the loading rate capacities $\eta_X$ are roughly equal.

In the analogy with the game of players tossing cards onto a table, $x$ could represent the maximal overlapping rate still allowing to recognize a card and $p$ the maximal expected probability of not being able to recognize it. Hence, $\eta_X$ represents the maximal loading rate permitting to achieve this goal, i.e., $N_tN_f\eta_X$ the maximal number of players when the sizes of the cards and of the table are given.

Let us evaluate the previously introduced quantities, namely $\Pr[X > x]$ and the loading rate capacities $\eta_X$.

IV. Probabilistic study

To evaluate the distributions of $X$ and SINR, and thus the load-capacities, we proceed in two steps. Firstly, we reduce the context to $N = 2$ emitters. The probability distributions can then be derived in closed forms. When $N > 2$, the analytic study of the probability distributions is difficult, except for $X_{\max}$. In particular, when dealing with $X_{\cup}$, the study appears to be “combinatorially difficult”. However, as previously seen, its distribution can be bounded by that of $X_{\max}$. As an alternative, we will also propose an algorithmic approach to numerically evaluate this distribution by means of Monte Carlo (MC) simulations. As we will see, when $\eta$ is small, the distributions of the $X$’s are very close, as previously explained. Thus, the analytical expression of $\Pr[X_{\max} > x]$ allows to give insights on the load-capacity in the proposed “cards tossing” scenario.

A. The case of two competitors

A simple look at the geometrical configuration plotted figure 2 allows to see that the normalized surface writes

$$X_k = (1 - \tau_k)(1 - \varphi_k)I_0[0; 1](\tau_k, \varphi_k)$$

(10)

where $\tau_k = \frac{|t_k - t_0|}{\Delta t}$ and $\varphi_k = \frac{|f_k - f_0|}{\Delta f}$ are the normalized absolute time and frequency difference between emission $I_0$ and $I_k$ and are defined over $[0; N_t - 1]$ and $[0; N_f - 1]$ respectively.

Thus, for $x \in [0; 1)$, one have $\Pr[X_k > x] = \Pr[(1 - \tau_k)(1 - \varphi_k) > x]$. Remind that the $(t_k, f_k)$ are assumed
continuous, thus the $(\tau_k, \varphi_k)$ are continuous as well. For the sake of brevity, we will omit index $k$ in the sequel. The dependence on $k$ is clear from the use of $X_k$; moreover, from the assumption of identical distributions for the $(t_k, f_k)$, the $(\tau_k, \varphi_k)$ have also identical distributions. Denoting by $p_{r,\varphi}$ the pdf of $(\tau_k, \varphi_k)$, we immediately get,

$$\Pr[X_k > x] = \int_0^1 \left( \int_0^{1-x} p_{r,\varphi}(u, v) \, dv \right) \, du \quad (11)$$

where the bound of the first integral is due to the fact that when $u \geq 1 - x$, $1 - \frac{u}{1-u} \leq 0$ and thus the inner integral is zero. The probability of collision $p_c$ is thus given by $x = 0$, i.e.,

$$p_c = P_{r,\varphi}(1, 1) \quad (12)$$

where $P_{r,\varphi}$ is the cumulative distribution function (cdf) of $(\tau_k, \varphi_k)$.

In the particular case where the $t_k$ and $f_k$ are independent, $\tau_k$ is independent of $\varphi_k$ as well, i.e., $p_{r,\varphi}(u, v) = p_r(u) p_{\varphi}(v)$. The expression simplifies as

$$\Pr[X_k > x] = \int_0^{1-x} P_{\varphi} \left( 1 - \frac{x}{1-u} \right) p_r(u) \, du \quad (13)$$

and the probability of collision turns to be

$$p_c = P_r(1) P_{\varphi}(1) \quad (14)$$

where $P_r$ and $P_{\varphi}$ stands for the cdf of $\tau_k$ and of $\varphi_k$ respectively (remind that $\tau_k \geq 0$ and $\varphi_k \geq 0$ are continuous, and thus $P_r(0) = P_{\varphi}(0) = 0$). Note that both formulas (11) and (13) admit a symmetrical expression by exchanging the role of $\tau_k$ and $\varphi_k$.

In the context of spreading spectrum, $f_0 = f_k = 0$ so that $\varphi_k = 0$ and thus $P_{\varphi}(v) = 1_{\mathbb{R}_+}(v)$: one can easily check that Eqs. (13) and (14) reduce to

$$\Pr[X_k > x] = P_r(1-x) \quad (15)$$

and

$$p_c = P_r(1) \quad (16)$$

respectively.

Given a probability distribution for $(t_k, f_k)$ (resp. $t_k$ and $f_k$), it is straightforward to determine that of $(\tau_k, \varphi_k)$ (resp. $\tau_k$ and $\varphi_k$) thanks to the independence of the $(t_k, f_k)$, and thus to analytically or numerically determine the probabilities given by Eqs. (11)-(16). For instance, for $t_k$ and $f_k$ independent, and uniformly distributed over $[0; T - \Delta t]$ and $[0; F - \Delta f]$ respectively, and when $N_f > 2$ and $N_t > 2$ (e.g., for an ultra-narrow band approach), some long algebra leads to

$$\Pr[X_k > x] = \frac{(a + b x)(1-x) + (c + x)x}{(N_t - 1)^2(N_f - 1)^2} \quad (17)$$

with

$$\begin{cases} a = (2N_t - 3)(2N_f - 3) \\ b = 9 - 2N_t - 2N_f \\ c = 2(N_t - 2)(N_f - 2) \end{cases}$$

for eq. (13). In a wide-band approach, $N_f = 1$ (and thus $N_t > 2$), one obtains for eq. (15)

$$\Pr[X_k > x] = \frac{(2N_t - 3 + x)(1-x)}{(N_t - 1)^2} \quad (18)$$

B. From $N = 2$ to $N > 2$ competitors

When more than two nodes get access to the medium, the evaluation of the statistics is drastically more complicated. A simple look at figure I shows that a combinatorial number of configurations is possible giving the same surface lost by the emitter. From the independence of the access of the $I_k$s to the channel, the $X_k$ are independent and thus the probability distribution of $X_{max}$ can easily be computed. That of $X_S$ also admits a formal expression although this one is quite complicated and quite difficult to exploit for our purpose. Finally, the distribution of $X_{ij}$ is too much complicated to investigate due to the combinatorial configurations leading to the same union of surfaces. In this case, and for $X_S$ as well, the distributions can be analyzed by mean of MC simulations and thanks to algorithms of the geometrical computation to evaluate $X_{ij}$ for a given configuration.

1) An analytical approach: Let us denote by $P_X(x) = \Pr[X \leq x]$ the cdf of $X$, where $X$ stands respectively for $X_k$, $X_{max}$, $X_S$ and $X_{ij}$.

From, $\Pr[X_{max} \leq x] = \Pr[\cap_k(X_k \leq x)]$, and the independence of the $X_k$, we immediately have $P_{X_{max}}(x) = \prod_{k=1}^N P_{X_k}(x)$. In the considered context of identical distributions, we achieve to the expression

$$\Pr[X_{max} > x] = 1 - \left( 1 - \Pr[X_k > x] \right)^{N-1} \quad (19)$$

where $\Pr[X_k > x]$ is given by Eqs. (11)-(13)-(15) according to the context. Some straightforward algebra leads then to the channel loading rate capacity,

$$\eta_{X_{max}}(x, p) = \frac{\ln(1-p)}{\ln\left(1 - \Pr[X_k > x]\right)} \quad (20)$$
Note now that $P_{X_k}$ is discontinuous at $x = 0$ since $p_c = \Pr[X_k > 0] \neq 1$. However, $P_{X_k}$ can be written as
\[ P_{X_k} = (1 - p_c)\mathbb{1}_{\mathbb{R}^+} + p_c Q \]
where $Q$ is then a continuous cdf. In other words, by denoting by $C_k$ a Bernoulli variable representing the collision between $I_k$ and $I_0$, $\Pr[C_k = 1] = 1 - \Pr[C_k = 0] = p_c$, and by $Y_k$ a continuous random variable of cdf $Q$, independent of $C_k$, one can write
\[ X_k = C_k Y_k \]
where $\mathbb{1}$ stands for the equality in distribution. This stochastic representation allows to evaluate the exact distribution $P_{X_k}$. Writing $C, Y$ and $c$ as column vectors of components $C_k, Y_k$ and $c_k$ respectively, this cdf resolves to
\[ P_{X_k}(x) = \sum_{c \in \{0,1\}^{N-1}} \Pr[c^t Y \leq x \mid c = c] \Pr[C = c] \]
In this stochastic representation, the $C_k$ being independent and identically distributed, and the $Y_k$ as well, independent on the $C_k$, one immediately have
\[ P_{X_k}(x) = \sum_{n=0}^{N-1} \binom{N-1}{n} p_c^n (1-p_c)^{N-1-n} \Pr\left[ \sum_{k=1}^{n} Y_k \leq x \right] \]
with the convention $\sum_{n=0}^{0} = 0$. This finally gives
\[ P_{X_k} = (1 - p_c)^{N-1} \mathbb{1}_{\mathbb{R}^+} + (N - 1) p_c (1 - p_c)^{N-2} Q \]
\[ + Q \sum_{n=2}^{N-1} \binom{N-1}{n} p_c^n (1-p_c)^{N-1-n} q^{(n-1)*} \]
(21)
where $*$ stands for the convolution, $\mathbb{1}^*$ stands for the $l$ times convolution, and $q$ is the pdf associated with the cdf $Q$.

Although exact, this expression is generally very difficult to evaluate, especially when $N$ is large. To overcome this difficulty, one may approximate $q^{(n-1)*}$ by a Gaussian, approximation that asymptotically holds from the central limit theorem (when $n$ is large). However, the approximate expression of $P_{X_k}$ becomes a mixture of Gaussians, which remains complicated to evaluate analytically. The approximation of $X_k$ by a Gaussian asymptotically holds. However, the behavior of the tails of the distribution (especially when $x$ is small) is lost in such an approximation. Moreover, if $p_c$ is very small the “asymptotic” regime may be valid for very large $N$.

As previously evoked, dealing with $X_{ij}$, the study of the distribution $P_{ij}(x)$ is too difficult due to the combinatorial number of geometrical configurations. For instance, when $N = 3$, $S_{ij} = \mu(I_0 \cap (I_1 \cup I_2)) = \mu(I_0 \cap I_1) + \mu(I_0 \cap I_2) - \mu(I_0 \cap I_1 \cap I_2)$. The distribution must be studied by inclusion-exclusion relations, that is drastically difficult.

In both cases $X_{ij}$ and $X_{ij}^\prime$ one may bound the distribution by that of $X_{max}$ (and the load-capacity as well). However, in the following subsection, we will see how distribution $P_{X_{ij}}$ can be estimated by means of MC simulations, and how $P_{X_{ij}^\prime}$ can be estimated via the same snapshots of MC simulations.

2) Combined Monte Carlo and Klee’s approach: In this section, we then propose a Monte Carlo (MC) approach to estimate the distribution $P_{X_{ij}}$: a large number of configurations of emissions $\{I_k\}_{k=0,\ldots,N-1}$ are randomly drawn, and for each, $X_{ij}$ is evaluated, that allows to estimate the cdf $P_{ij}$.

The problem is then to be able to evaluate $X_{ij}$ for a given configuration of “cards tossed on the table”. This problem is very close to what is known as the Klee’s measure problem [13, 19, 20, 21, 22]. The problem of Klee was to design a fast algorithm that can evaluate the area of the union of rectangles. Here, instead of evaluating $X_{ij} = \Delta t \Delta f - (\mu (\cup_{k=0}^{n-1} I_k) - \mu (\cup_{k=0}^{n-1} I_k^*))$, the approach is adapted to our problem. The principle is the following:

1) We search for the $I_{ik}, k \geq 1$ that partially or fully cover $I_0$ and we restrain the problem to these $I_k$;

2) We decompose $I_0$ into subrectangles by cutting $I_0$ vertically with the borders of the $I_k$, i.e., the lines $t = t_k$ when $t_0 < t_k < t_0 + \Delta t$, and the lines $t = t_k + \Delta t$ when $t_0 < t_k + \Delta t < t_0 + \Delta t$. The same cut is made horizontally in frequency domain.

3) Each subrectangle is thus either completely covered by an $I_k$, or completely uncovered. Hence, $S_{ij}$ is initialized to 0 and is then updated subrectangle by subrectangle by the corresponding elementary area.

This principle is illustrated figure [3] Practically, the borders of the rectangles $I_k$ are sorted by a quick sort or a $kD$-tree algorithm and in the same step of sorting one can detect if the border/Ik covers I0 and up-to-date surface $S_{ij}$ accordingly.

Note that this approach has various advantages:

- $S_{ij}$ can be evaluated in the same loop by taking into account the number of $I_k$ covering the subrectangle currently treated.
- The approach is not restricted to rectangles with the same size, but is general to any set of rectangles, whatever their sizes.
- A nonuniform repartition of the energy ($\mathbb{E}[|g_k(t,f)|^2]$ nonconstant on $I_k$) and different energies for each $I_k$ can also be taken into account by integrating the energy of the $I_k$ in the subrectangle.

To avoid repeating MC simulations for various values of $N$, one can fix a maximal number $N_m$ and estimate $P_{X_{ij}}$ and $P_{X_{ij}^\prime}$ for any $N \leq N_m$ in the same MC snapshot. To this end, one records the indexes of the $I_k$ that fall in each subrectangle, let say $k_1 < \cdots < k_l$, and updates the surfaces $S_{ij}$ (resp. $S_{ij}^\prime$) corresponding to all $N \geq k_1$ (resp. to all $N \geq k_1$, and once again to all $N \geq k_2$,

V. SOME NUMERICAL ILLUSTRATIONS AND RESULTS

Let us now illustrate how our approach can characterize the medium access in terms of load-capacity. To this end, we consider two scenarios as follows:

1) In the wide band context, “rectangle” is to be understood as “segments” and the frequency step does not exist.
Fig. 3. Schematic principle of the algorithmic approach to solve the Klee’s measure problem. $I_0$ is cut into elementary rectangles using the borders of the $I_k$ that intersect $I_0$, which is represented by dotted lines here. $S_{ij}$ is then the sum of the elementary surfaces represented in gray. The darkness of the gray indicates the number of $I_k$ that fall in each elementary rectangle (e.g., here from 1 for the brighter to 3 for the darker), illustrating that counting how much $I_k$ cover the elementary surface one can up-to-date $S_{ij}$ in the same loop.

- An ultra narrow band system, where the characteristics are chosen to be: $\Delta t = 6$ s and $\Delta f = 150$ Hz, for $T = 30'$ and $F = 48$ kHz.
- A spreading spectrum approach where we choose $\Delta t = 0.45$ s, $T = 12$ hours and $F = 125$ kHz ($\Delta f = F$).

In both configurations, the maximal number of time-frequency slots is $N_t N_f = 96000$ and thus the example is comparable in terms of allowable resource. The characteristics chosen in these two examples enter in the specifications of technologies for LPWAN in star topology communication and random access to the channel, namely Sigfox® and LoRa® respectively; they are thus highly realistic.

Figure 4 gives for both examples the probability $\Pr[X > x]$ versus $x$, for various values of the load $\eta$. The dots, given for $x = 0$, represent the probabilities of collision. These curves illustrate that more flexibility in the exigence, i.e., allowing some recovery, assuming that in such a case the information is not totally lost, the probability to loose a packet of information naturally decreases, as obviously expected. Clearly, these probabilities satisfy the order relation established by eq. 6. Moreover, the characteristics of $X_{\max}$, $X_{\Sigma}$ and $X_{\Delta}$ are very close when the loading rate $\eta$ is not too large. These curves illustrate that, practically, the quantity $X_{\max}$ is an accurate approximation of any of the $X$s. The loading rate capacity $\eta X_{\max}$ having a closed-form expression, we are thus able to accurately assess the load-capacity without recourse to MC simulations, in a reasonable range for the acceptable probability of loss (and the allowed covering of $I_0$).

Figure 5 represents these probabilities, but now in terms of the loading rate $\eta$, fixing the allowed recovery surface $x$. This representation of the probabilities vs $\eta$ leads to the channel loading rate capacity by fixing $\Pr[X > x] = p$ (see eq. 5). The curves are plotted in a range that seems reasonable in terms of application, where a low probability of loss is prescribed. These curves confirm that in such a range, the three quantities $X$ lead to the same results, and thus to the same load-capacity, with a high accuracy.

Fig. 4. Probabilities $\Pr[X > x]$ versus $x$ for several loading rates $\eta$, for $X_{\max}$ (solid line), $X_{\Sigma}$ (dashed-dotted line) and $X_{\Delta}$ (dashed line). Figure (a) concerns the ultra narrow band example and figure (b) is for the spreading spectrum approach. The dots, given for $x = 0$ represent the probability of collision $p_c$.

Fig. 5. Probabilities $\Pr[X > x]$ versus $\eta$ for several allowed recovery surfaces $x$, for $X_{\max}$ (solid line), $X_{\Sigma}$ (dashed-dotted line) and $X_{\Delta}$ (dashed line). The three curves are almost indistinguishable, that validate the accurate approximation $X_{\Sigma} \approx X_{\Delta} \approx X_{\max}$. Figure (a) concerns the ultra narrow band example and (b) is for the spreading spectrum approach.

Figure 6 gives the loading rate capacities $\eta X$ versus $x$ for some outage probabilities. These curves are deduced from the previous ones by fixing the outage probability. Moreover, these results clearly confirm the accuracy of $\eta X_{\max}$ as an approximation of the loading rate capacities in a useful range for $x$ and $p$ for real-world applications. More than its accuracy, this approximation has also the great advantage to have a closed form, and thus it is very useful to tune a sensor network once the target values of $x$ and $p$ are fixed.

Fig. 6. Loading rate capacities as defined by Eq. 7 versus $x$ for several outage probabilities $p$, for $X_{\max}$ (solid line), $X_{\Sigma}$ (dashed-dotted line) and $X_{\Delta}$ (dashed line). The three curves can be deduced from the previous ones, and again validate the accuracy of the approximation $\eta X_{\Sigma} \approx \eta X_{\Delta} \approx \eta X_{\max}$. Figure (a) concerns the ultra narrow band example and (b) is for the spreading spectrum approach.
Finally, figure [7] illustrates why $X_{\text{max}}$ gives an accurate approximation of both $X_{\cup}$ and $X_{\Sigma}$. From the independence of the $I_k$, the (random) number $N_c$ of $I_k$ that collide with $I_0$ follows a binomial law of parameters $N-1$ and $p_c$ (probability of collision). In the reasonable range of interest in terms of maximal probability of packet loss, the probability of collision $p_c$ is very low. Thus, $Pr[N_c \geq 2]$ is relatively low, even when $\eta$ is not very low: the probability $Pr[X_{\cup} = X_{\text{max}}]$ and $Pr[X_{\Sigma} = X_{\text{max}}]$ are thus close to 1.

VI. CONCLUDING REMARKS

In this paper, we studied the interferences in a two-dimensional time-frequency random access to a medium. This approach can find application for instance in LPWAN. The proposed approach is different from common stochastic geometry models where the geographical domain of a sensors network is modeled as a Euclidean plane, and the locations of sensor nodes modeled as a point process. With the power control assumption, the geometry of the locations of the nodes, especially the distances from the base station to sensor nodes can be put aside. We model the available time-frequency resource as a two-dimensional plane where the transmissions from different sensor nodes are randomly distributed. Our approach encompasses the common Aloha access as a limit case.

In random access protocols, the collision effect is commonly considered as binary: when a collision occurs, the packets are considered to be lost. The “pure” random time-frequency access allows to study more subtly the collision’s effect in the sense that if two packets are not too much overlapped, the packets are not necessarily lost. We thus provided some quantities to refine the description of the level of interference introduced by collisions. Besides, fixing an acceptable probability of packet loss, our approach allows to determine the capacity of such a channel in terms of the maximal number of sensor nodes supported by the channel not to exceed this outage probability. This is what we called the load capacity.

Note that we focused on the MAC layer without discussing the physical layer technological details, which are necessary to fix the thresholds and target SINR. Moreover, the relaxation of the simplified assumption on the time-frequency supports of the packets or on the energy repartition is under study. A more flexible model could thus be used to combine our approach with that dealing with the geographical repartition of the sensor, taking into account path-loss effects.

REFERENCES

[1] J. Yick, B. Mukherjee, and D. Ghosal. Wireless sensor network survey. Computer Networks, 52(12):2292–2330, August 2008.
[2] W. Dargie and C. Poellabauer. Fundamentals of Wireless Sensor Networks. Wiley, 2010.
[3] J. Höller, V. Tsatsis, C. Mulligan, S. Karnouskos, S. Avesand, and D. Boyle. From Machine-to-Machine to the Internet of Things: Introduction to a New Age of Intelligence. Academic Press, Oxford, 2014.
[4] V. K. Garg. Wireless Communication and Networking. Morgan Kaufmann, Amsterdam, 2007.
[5] N. Abramson. The ALOHA system - another alternative for computer communications. In Proceedings of the American Federation of Information Processing Societies (AFIPS): 70 Fall Joint Computer Conference, pages 281–285, Houston, Texas, USA, 17-19 November 1970.
[6] N. Abramson. The ALOHANet – surfing for wireless data. IEEE Communications Magazine, 47(12):21–25, December 2009.
[7] A. S. Tanenbaum and D. J. Wetherall. Computer Networks. Prentice Hall, Boston, 5th edition, 2011.
[8] A. J. Viterbi. Spread spectrum communications – myths and realities. IEEE Communications Magazine, 17(3):11–18, May 1979.
[9] H. R. Walker. VPSK and VMSK modulation transmit digital audio and video at 15 bits/sec/Hz. IEEE Transactions on Broadcasting, 43(1):96–103, March 1997.
[10] M. T. Do, C. Goursaud, and J.-M. Gorce. On the benefits of random FDMA schemes in ultra narrow band networks. In Proceedings of the 12th IEEE International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt), pages 672–677, Hammamet, Tunisia, 12-16 May 2014.
[11] C. T.-C. Nguyen. MEMS technology for timing and frequency control. IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control, 54(2):251, February 2007.
[12] J. Lim, H. Kim, T. N. Jackson, K. Choi, and D. Kenny. An ultra-compact and low-power oven-controlled crystal oscillator design for precision timing applications. IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control, 57(9):1906–1914, September 2010.
[13] D. Buranapanichkit and Y. Andreopoulos. Distributed time-frequency division multiple access protocol for wireless sensor networks. IEEE Wireless Communications Letters, 1(5):440–443, October 2012.
[14] M. Sakr, A. Al-Moghazy, H. Abou-Bakr, and M. Fikri. Hybrid DS-FH packet acquisition for frequency hopped random multiple access. In Proceedings of the 2012 Japan-Egypt Conference on Electronics, Communications and Computers, pages 133–137, Alexandria, Egypt, March, 6-9 2012.
[15] F. Baccelli and Blaszczyszyn. Stochastic geometry and wireless networks: Volume I theory. Foundations and Trends® in Networking, 3(3-4):249–449, 2010.
[16] F. Baccelli and Blaszczyszyn. Stochastic geometry and wireless networks: Volume II applications. Foundations and Trends® in Networking, 4(1-2):1–312, 2010.
[17] K. Whitehouse, A. Woo, F. Jiang, J. Polastre, and D. Culler. Exploiting the capture effect for collision detection and recovery. In Proceedings of the 2nd IEEE Workshop on Embedded Networked Sensors, pages 45–52, Sydney, Australia, May 30–31 2005.
[18] V. Klee. Can the measure of $\cup_{i}^{n}[a_i, b_i]$ be computed in less than $O(n \log n)$ steps? The American Mathematical Monthly, 84(4):284–285, April 1977.
[19] J. L. Bentley. Algorithms for Klee’s rectangle problems. Unpublished manuscript, University of Pennsylvania, Philadelphia, PA, USA, 1977.
[20] J. L. Bentley and D. Wood. An optimal worst case algorithm for reporting intersections of rectangles. IEEE Transactions on Computers, 29(7):571–577, July 1980.
[21] J. van Leeuwen ans D. Wood. The measure problem for rectangular ranges in $d$-space. Journal of Algorithms, 2(3):282–300, September 1981.
[22] T. M. Chan. Klee’s measure problem made easy. In Proceedings of the 54th IEEE Symposium on Foundations of Computer Science (FOCS), pages 410–419, Berkeley, CA, USA, 26-29 October 2013.