Experimental Evidence for Heavy Tailed Interference in the IoT

Laurent Clavier, Troels Pedersen, Ignacio Rodriguez, Mads Lauridsen and Malcolm Egan

***Abstract*—5G and beyond sees an ever increasing density of connected things. As not all devices are coordinated, there are limited opportunities to mitigate interference. As such, it is crucial to characterize the interference in order to understand its impact on coding, waveform and receiver design. While a number of theoretical models have been developed for the interference statistics in communications for the IoT, there is very little experimental validation. In this paper, we address this key gap in understanding by performing statistical analysis on recent measurements in the unlicensed 863 MHz to 870 MHz band in different regions of Aalborg, Denmark. In particular, we show that the measurement data suggests the distribution of the interference power is heavy tailed, confirming predictions from theoretical models.**

***Index Terms*—Interference, IoT, statistical models, subexponential distributions, heavy tails.**

I. Introduction

Fifth generation sees an ever increasing density of wirelessly connected devices as the Internet of Things (IoT) emerges. Combined with the recent trends towards non-orthogonal multiple access (NOMA), stringent constraints on cost, energy and computational capabilities and grant free access schemes with low coordination between devices, interference management is a key challenge. However, characterizing the interference in unlicensed bands is a non-trivial issue. One reason is the high level of heterogeneity within the network, ranging from access protocols to PHY-layer design. In particular, time-on-air, symbol duration, bandwidth and waveform selection can significantly differ from one radio access technology to another. To address this issue, a number of different probability models for the interference have been proposed sharing a key observation: *additive Gaussian noise is a poor model.*

In this paper, we analyze recent measurements in Aalborg of interference in the unlicensed European ISM 863 MHz to 870 MHz band, which is duty cycle limited and not based on listen before talk [1]. Our analysis validates key features of interference statistics predicted from theoretical models. In particular, analysis via the converging variance test and of the empirical survival function shows that the empirical distribution of the interference from the measurement data is heavy tailed. An important family of tractable heavy tailed models are the α-stable models and we also provide evidence to suggest these models are consistent with the measurement data.

Laurent Clavier is with IMT Lille Douai, Univ. Lille, CNRS, UMR 8520 - IEMN, USR 3380, France; Troels Pedersen and Ignacio Rodriguez are with Dept. Electronic Systems, Aalborg University, Denmark; Mads Lauridsen with Nokia Bell Labs Aalborg; Malcolm Egan with INRIA, CITI, France

Section II sets up the notation and give the necessary background on interference models. Section III describes the expected properties of impulsive interference. Section IV presents the measurement setup, states the main results and open the discussion for the appropriate model.

II. Theoretical Interference Models

Interference is the consequence of concurrent, at least partially, transmissions from different devices (called interferers) on the same channel observed by a given receiver. As Fig. 1 illustrates, we consider a situation where a receiver is surrounded by a set of interferers, the locations of which forms a point process denoted as Ω. Interferer $i \in \Omega$ sits

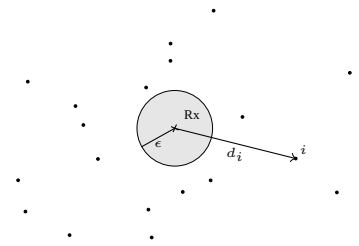

Fig. 1. Receiver (Rx) surrounded by interferers (dots). Interferer i is at distance d_i from the receiver and sits outside the guard zone of radius ϵ.

at distance d_i from the receiver. Furthermore, the signal of interferer i is subject to a path loss $\ell(d_i)$ given by the positive path loss function ℓ. At a particular time-frequency (t, f) bin, the interference can be expressed as a complex number

$$X_{t,f} = \sum_{i \in \Omega} \ell(d_i) Q_{i,t,f}. \tag{1}$$

$Q_{i,t,f}$ includes propagation effects (e.g., multipath and shadowing) as well as the baseband emission of interferer i at time t and frequency f. In existing work, the path loss function $\ell(\cdot)$ takes many different forms, but is often given by $\ell_{\gamma,\epsilon}(d) = d^{-\gamma}\mathbf{1}_{\{d \geq \epsilon\}}$, where γ is the path loss exponent and ϵ is the guard zone radius; that is, no interferer can be closer than ϵ from the given receiver.

A key challenge is to characterize the probability distribution of the interference, $X_{t,f}$ or the corresponding interference power $\mathcal{I}_{t,f} = |X_{t,f}|^2$. Here, we characterize $\mathcal{I}_{t,f}$ by its probability distribution function $F(x) = \mathbb{P}(X < x)$ or equivalently its survival $S(x) = 1 - F(x)$.

Dating back to Middleton [2], stochastic geometry has played an important role in characterizing interference statistics. In this approach, interferers are located according to a

homogeneous Poisson point process, and non-Gaussian interference distributions naturally arise. For example, when the network radius is finite and the guard zone radius, ϵ, is non-zero, the interference distribution is known to be Middleton [2], [3]. While exact, the Middleton model is difficult to work with analytically (e.g., for the purpose of receiver design) and a number of approximate models have been proposed; e.g., the Gaussian mixture model [4] or the ϵ-contaminated model [5].

On the other hand, when interferers are located according to a binomial point process, Weber and Andrews have shown that the resulting interference amplitude is subexponential [6][1] In this case, the tail behaviour is dominated by the strongest interferer: F is subexponential if for independent and identically distributed (i.i.d.) X_i with common survival function S, $S\left(X_1 + X_2 + \cdots + X_n\right) \to nS\left(X_1\right)$ as $x \to +\infty$.

In the case of the interference power, a detailed study has been carried out by Haenggi and Ganti [8]. In particular for interferers located according to a Poisson point process, they showed that the distribution is heavily dependent on the path loss attenuation coefficient γ. For example, also shown by Win and Pinto [9], in a network with infinite radius and no guard zone ($\epsilon = 0$), the interference power has the totally skewed α-stable distribution, where α depends on γ.

III. Classes of Heavy Tailed Distributions

Several important performance metrics, such as outage probability, strongly depend on the probability that the interference is large. That is, rare events in the tail of the interference power distribution play a key role. One reason that the Gaussian distribution fails to be a good model for the interference amplitude is that the tail probability decay rapidly. On the other hand, models such as the α-stable distribution have much slower decay in the tails.

Heavy tailed distributions are defined in general as distributions with tails that decay slower than the exponential distribution [10]. Formally, a distribution is heavy tailed if for any $M > 0$ and $t > 0$, the survival function does not satisfy

$$S(x) \leq M \exp(-tx), \ \forall x > 0. \tag{2}$$

Failing to satisfy (2) means that the moment generating function does not exist, and hence some higher order moments, even sometimes the variance or the mean are not finite.

A natural question is whether measurements are consistent with heavy tailed models, which are able to capture slow decay of the tails. That is, are these models—predicted in certain regimes by theory—good approximations of the real behavior of the interference, despite some moments not being finite? To address this question we consider three subclasses of heavy tailed models and their corresponding statistical tests.

A. Fat Tailed Distributions

A distribution with infinite variance is said to be fat tailed. Given a set of observations, the hypothesis that the variance of the underlying distribution is finite can be tested using the *converging variance test* [11, Section 5.5]. Assuming that the process is a first order stationary and ergodic time-frequency process, we estimate the variance by computing the sample variance using all available time and frequency measurements as

$$\sigma_n^2 = \frac{1}{n} \sum_{k=1}^{n} \left(\mathcal{I}_k - \overline{\mathcal{I}}_n\right)^2, \tag{3}$$

where $\overline{\mathcal{I}}_n = \frac{1}{n}\sum_{k=1}^{n} \mathcal{I}_k$. If $\mathcal{I}$ has finite variance, then σ_n should rapidly converge to a finite value as n increases. On the other hand, if the variance is infinite or very large, the convergence should not be obvious and features such as large jumps can be present as n increases.

B. Subexponential Tail Decay

Another class of heavy tailed distributions are those with *subexponential tail decay*. In particular, a distribution F is said to have subexponential tail decay if there exists some $\gamma > 0$ such that its survival function satisfies

$$S(x) = x^{-\frac{1}{\kappa}} L(x), \text{ for } x > 0, \tag{4}$$

where κ is called the tail index and L is a slowly varying function satisfying $\lim_{t\to\infty} \frac{L(tx)}{L(t)} = 1$. Plotting $S(x)$ against $\log(x)$ yields, for subexponentially decaying F, a straight line with slope $-1/\kappa$ for x large. An exponentially decaying distribution has $\kappa = 0$ which leads to an abrupt decrease in the curve as $\log x$ increases.

Again, assuming that the interference process is first order stationary and ergodic, we estimate the marginal distribution by computing the empirical distribution function as $\hat{F}(x) = \frac{1}{n}\sum_{k=1}^{n} \mathbf{1}_{\{X_k \leq x\}}$ and the empirical survival function as $\hat{S}(x) = 1 - \hat{F}(x)$.

C. Non-Gaussian α-Stable Distributions

The α-stable distributions are a special case of fat tailed distributions when $0 < \alpha < 2$ where the Gaussian is obtained with $\alpha = 2$. The distribution function of an α-stable random variable is described by four parameters: the characteristic exponent $0 < \alpha \leq 2$; the scale parameter $\gamma \in \mathbb{R}_+$; the skew parameter $\beta \in [-1, 1]$; and the shift parameter $\delta \in \mathbb{R}$. As such, a common notation for an α-stable random variable X is $X \sim S_\alpha(\gamma, \beta, \delta)$. In general, α-stable random variables do not have closed-form PDFs, but are usually represented by their characteristic function [12, Eq. 1.1.6]. As noted in Section II, the α-stable distributions arise as the interference distribution for Poisson point process models with no guard zones ($\epsilon = 0$).

The parameters of the α-stable can be estimated using classical methods, such as fitting the four parameters to the empirical characteristic function estimated from data as proposed by Koutrouvelis [13]. The quality of the fit between the measurement data and the estimated α-stable distribution can be evaluated by generating samples from an α-stable distribution based on the estimated parameters. The quantiles of the measurement samples are then plotted against those of the generated samples, i.e. a quantile-quantile (QQ) plot. If the measurement data and the generated samples are from the same distribution, the QQ plot will be close to the line $y = x$.

[1]We also note that the subexponential family has been used to model the transmit power distribution in [7].

IV. Measurement Data Analysis

We now turn to analyze the measured interference data first reported in [1], where received power measurements were performed at five distinct locations in Aalborg (Denmark): 1) downtown shopping area, 2) a business park with office buildings, 3) hospital complex, 4) industrial area consisting of industrial production facilities and office buildings, and 5) residential area with single-family houses. At each location, measurements were performed at street level by using a radio network scanner equipped with an omni-directional antenna for a period of 2 hours. The entire on-air RF activity in the 868 MHz ISM band (863 MHz to 870 MHz) was recorded with a 7 kHz bin resolution in frequency and 200 ms sampling time yielding a sensitivity level of approximately –115 dBm. The setup and measurements are further detailed in [1].

We reduce the data by aggregating data in time-frequency windows of 200 ms and 126 kHz which fits a LoRa scheme. This yields a sequence of interference samples $\mathcal{I}_{1,1},\ldots,\mathcal{I}_{N_t,N_f}$, with N_t and N_f being the number of time and frequency samples, respectively. Examples of measurements in a particular frequency band as a function of time are presented in Fig. 2.

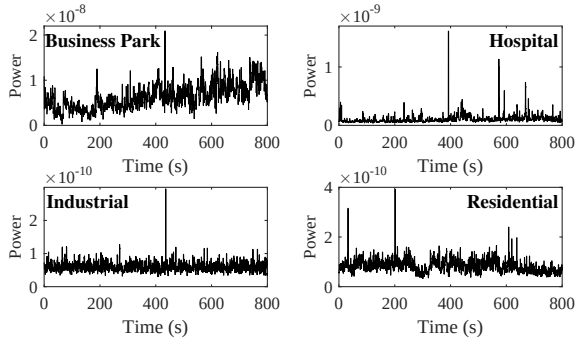

Fig. 2. Example of interference samples measured in the different areas.

A. Fat Tails: Converging Variance Test

We first test whether the data set is consistent with fat tails; i.e., infinite variance, or at least some infinite moments. To illustrate the variance convergence test, we plot in Fig. 3 the results on one finite variance distributions (chi-square with two degrees of freedom) and two subexponential distributions from the α-stable family (one with $\alpha = 1.8$, slightly impulsive, and one with $\alpha = 1.2$, more impulsive). For the Chi-square, as expected, the estimated variance converges rapidly as the sample size increases. For the two non-Gaussian α-stable distributions, there is no clear convergence even for very large sample sizes up to $n = 200\,000$. For measured data obtained at all five locations, the convergence tests also reported in Fig. 3 shows no clear convergence. This is observation consistent with fat tailed models.

B. Subexponentail Tail Decay

In Figs. 4, we plot the log empirical survival function versus $\log(x)$ for the shopping, business and residential areas,

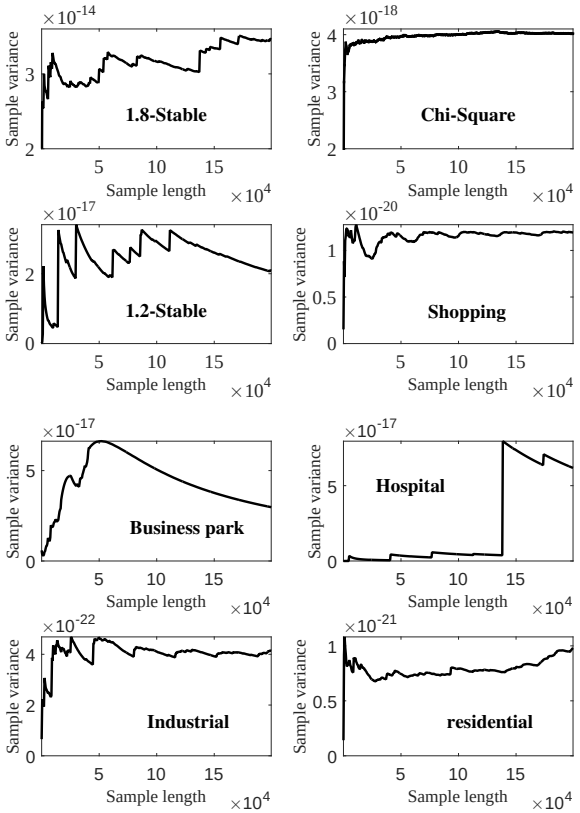

Fig. 3. Converging variance test for chi-square and α-stable distributions along with the measured data set (shopping area, bussiness park, hospital, industrial and residental areas).

respectively. To highlight the slow decay of the tails, we also plot curves corresponding to the Gaussian distribution (exponential decay) and three α-stable distributions (subexponential decay). Although the tests in Figs. 4 are only visual tests, they clearly confirm that, on the measurement sets, interference is consistent with subexponentially decaying tails, even if the end of the decay in the residential case requires further studies. This is also consistent with the results of the converging variance test in Fig. 3.

C. Non-Gaussian α-Stable Models

The fit between the measurement data set and estimated α-stable models is examined by means of the QQ plots in Figs. 5 and 6. It appears that the α-stable model fits well the data as shown for two different locations. Other locations are not presented due to space constraints but give similar results. Slight deviations can appear for quantiles under 5% or above 95%, but are difficult to interpret: only few samples are present and these are affected mostly by the particular measurement environment nearby the receiver.

D. Discussion

Our analysis is only based on a visual inspection. It clearly shows that heavy tail models have a role to play in modeling interference in ISM bands. Distributions with some infinite

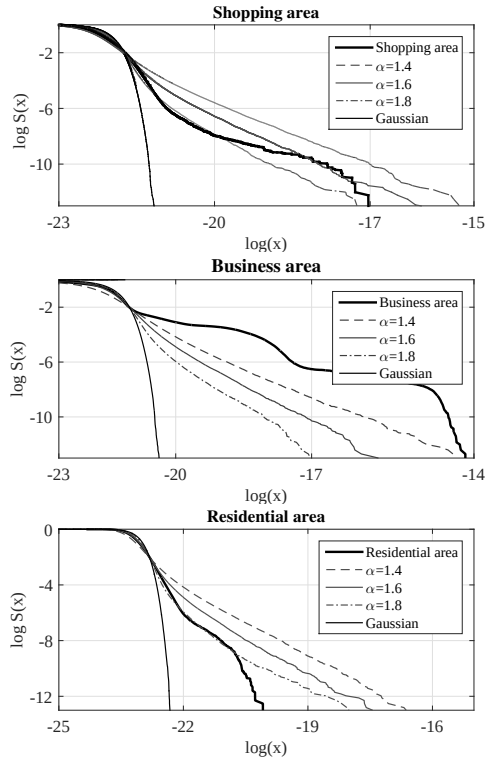

Fig. 4. Log empirical survival function $\log \hat{S}(x)$ as a function of $\log(x)$ in the shopping, business and residential areas, where $\hat{S}$ is the empirical survival function obtained from the measurement data.

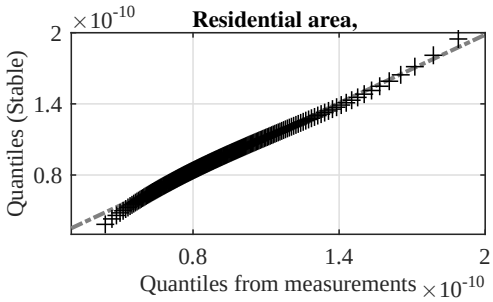

Fig. 5. QQ plot of interference samples from the residential area— more than 1.5 hours (28140 instants) and in the band 863 MHz to 868 MHz, BW = 126 MHz (39 bands)— versus simulated from an α-stable ($\alpha = 1.69, \beta = 1$).

moments of order two or larger could also be attractive and accurate solutions because they can represent some sudden changes in the interference that is more difficult to capture with distributions having all their moments finite. For instance, the α-stable distributions appear to fit well the measured data.

However, further measurements are required to properly explain the deviations in the tails and identify the best adapted models. Besides the data set is limited (one city, five locations and a receiver at the ground level). In IoT network it is to be expected that the access point will more likely be at higher altitudes which could impact the interference statistics. Finally we did not analyze time or frequency dependence which is another important issue for future study.

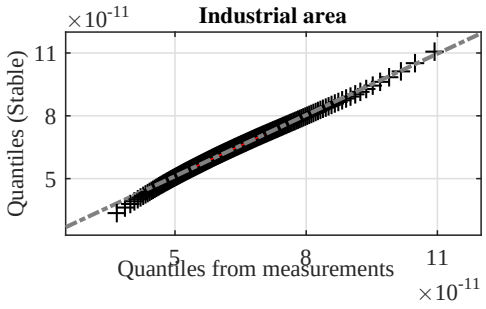

Fig. 6. QQ plot of interference samples from the industrial area— nearly 2 hours (35911 instants), in the band from 863 MHz to 865.5 MHz, BW = 126 MHz (19 bands)— versus simulated from an α-stable ($\alpha = 1.7, \beta = 1$).

V. Conclusions

We find that the measurement data, obtained at five different locations, confirm the hypothesis of the heavy tailed nature of the distribution of interference power. While there is an abundance of theoretical studies of interference statistics, the measurements in Aalborg are—to the best of our knowledge—the first to clearly validate the heavy tailed nature of the interference in the context of IoT communications. Interference models are the key to designing efficient coding and decoding strategies as well as efficiently adapting channel access and the network topology. As such, the measurement data suggests that there is a need to reconsider the utility of Gaussian models in network design for the IoT.

References

[1] M. Lauridsen, B. Vejlgaard, I. Z. Kovacs, H. Nguyen, and P. Mogensen, "Interference measurements in the European 868 MHz ISM Band with focus on LoRa and SigFox," in *2017 IEEE Wireless Commun. and Netw. Conf. (WCNC)*, pp. 1–6, March 2017.

[2] D. Middleton, "Statistical-physical models of electromagnetic interference," *IEEE Tran. Electromagnetic Compatibility*, vol. EMC-19, pp. 106–127, Aug. 1977.

[3] D. Middleton, "Non-Gaussian noise models in signal processing for telecommunications: new methods and results for class A and class B noise models," *IEEE Trans. Inf. Theory*, vol. 45, pp. 1129–1149, May 1999.

[4] N. Guney, H. Deliç, and M. Koca, "Robust detection of ultra-wideband signals in non-gaussian noise," *IEEE Trans. Microwave Theory Tech.*, vol. 54, pp. 1724–1730, June 2006.

[5] O. Alhussein, I. Ahmed, J. Liang, and S. Muhaidat, "Unified analysis of diversity reception in the presence of impulsive noise," *IEEE Trans. Veh. Technol.*, vol. 66, pp. 1408–1417, Feb. 2017.

[6] S. Weber and J. Andrews, "Transmission capacity of wireless networks," in *Foundations and Trends in Networking*, no. 2-3 in 5, NOW Publishers, 2012.

[7] A. J. Ganesh and G. L. Torrisi, "Large deviations of the interference in a wireless communication model," *IEEE Trans. Inf. Theory*, vol. 54, pp. 3505–3517, Aug. 2008.

[8] M. Haenggi and R. Ganti, "Interference in large wireless networks," *Foundations and Trends in Networking*, vol. 3, no. 2, pp. 127–248, 2009.

[9] M. Win, P. Pinto, and L. Shepp, "A mathematical theory of network interference and its applications," *Proc. IEEE*, vol. 97, pp. 205–230, Feb. 2009.

[10] G. Peters and P. Shevchenko, *Advances in Heavy Tailed Risk Modeling: A Handbook of Operational Risk*. Hoboken, Wiley, 2015.

[11] C. L. Nikias and M. Shao, *Signal processing with alpha-stable distributions and applications*. Wiley-Interscience, 1995.

[12] G. Samorodnitsky and M. Taqqu, *Stable Non-Gaussian Random Processes : Stochastic Models with Infinite Variance*. Chapmann and Hall, 1994.

[13] I. A. Koutrouvelis, "Regression-type estimation of the parameters of stable laws," *J. Am. Stat. Assoc.*, vol. 75, pp. 918–928, 1980.