

# Fog-Based Detection for Random-Access IoT Networks with Per-Measurement Preambles

Rahif Kassab\*, Osvaldo Simeone\* and Petar Popovski†

\*King's Communications, Learning, and Information Processing (KCLIP) Lab, King's College London, London, United Kingdom

†Department of Electronic Systems, Aalborg University, Aalborg, Denmark

Emails: \*{rahif.kassab,osvaldo.simeone}@kcl.ac.uk, †petarp@es.aau.dk

**Abstract**—Internet of Things (IoT) systems may be deployed to monitor spatially distributed quantities of interests (QoIs), such as noise or pollution levels. This paper considers a fog-based IoT network, in which active IoT devices transmit measurements of the monitored QoIs to the local edge node (EN), while the ENs are connected to a cloud processor via limited-capacity fronthaul links. While the conventional approach uses preambles as metadata for reserving communication resources, here we consider assigning preambles directly to measurement levels across all devices. The resulting Type-Based Multiple Access (TBMA) protocol enables the efficient remote detection of the QoIs, rather than of the individual payloads. The performance of both edge and cloud-based detection or hypothesis testing is evaluated in terms of error exponents. Cloud-based hypothesis testing is shown theoretically and via numerical results to be advantageous when the inter-cell interference power and the fronthaul capacity are sufficiently large.

**Index Terms**—Random Access, IoT, Fog-RAN, Hypothesis Testing

## I. INTRODUCTION

The density of connected wireless devices is expected to continue growing as 5G and beyond-5G systems are deployed, especially for Internet-of-Things (IoT) services supported by massive Machine-Type Communications [1]. This motivates the investigation of access schemes that support high device densities without penalizing the end-to-end performance for specific IoT services. In this paper, we address this problem by considering the fog-radio access network deployment illustrated in Fig. 1, in which IoT devices monitor distributed Quantities of Interest (QoIs), such as noise or pollution levels. The devices access the network through their local Edge Nodes (ENs), e.g., access points, which are in turn connected via fronthaul links to a cloud processor. Devices are interrogated periodically from the corresponding EN, and they only transmit their measurements of the QoIs if active. The goal of the network is to detect the distributed QoIs based on hypothesis testing at either ENs or at the cloud.

A conventional approach would prescribe a random access protocol, such as ALOHA, through which devices communicate individual payloads to the local ENs. In case two or more devices select the same preamble during the random access phase, a collision would occur and no information would be

This work has received funding from the European Research Council (ERC) under the European Union Horizon 2020 research and innovation program (grant agreements 725731 and 648382).

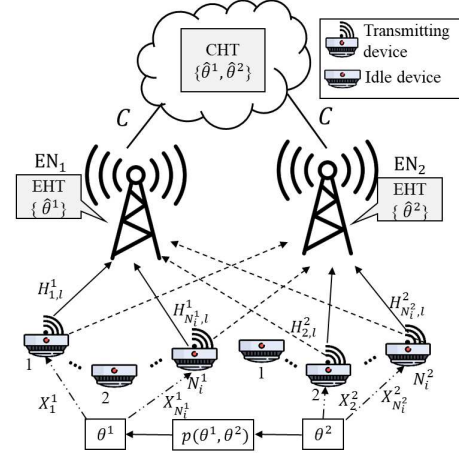


Fig. 1: Multi-cell fog-based IoT network aimed at estimating correlated distributed quantities of interest (QoIs) through either edge or cloud processing.

delivered to the network. As recognized for *single-cell systems* in [2]–[4], when the goal is estimating a common QoI measured at multiple devices, the requirement of distinct preambles per active user for a successful transmission is unnecessary. In such a case, it is in fact potentially more efficient to assign a specific preamble to each measurement level: in this way, devices making the same measurement contribute energy to the same preamble, potentially reinforcing its detection signal-to-noise ratio. The estimate of the QoI can then be obtained from the histogram of the received measurements [2]–[4]. The outlined protocol can be considered as a form of joint source-channel coding, and is known as Type-Based Multiple Access (TBMA).

In this paper, we investigate the performance of TBMA in a multi-cell fog-based system, as seen in Fig. 1 and detailed in Sec. II and Sec. III. A key new aspect of this type of network deployments is that detection, via hypothesis testing, of the distributed QoIs can be either carried out locally at the ENs or centrally at the cloud. Edge detection is impaired by inter-cell interference, while cloud detection is subject to fronthaul capacity constraints. In contrast to recent works that considered distributed hypothesis testing over wireless channels [5], [6], the goal in this paper is to detect the value of the QoI and not the joint distribution of all QoIs. The error

exponent analysis presented in this paper (Sec. IV) provides insights into the performance comparison between edge and cloud processing, and the presented numerical results (Sec. V) validate our findings. Additional results can be found in the extended version of this paper [7].

**Notation:** Lower-case and upper case bold characters represent vectors and matrices respectively.  $\mathbf{A}^T$  denotes the transpose of matrix  $\mathbf{A}$ .  $|\mathbf{A}|$  denotes the determinant of matrix  $\mathbf{A}$ .  $A(i, j)$  denotes the element of  $\mathbf{A}$  located at the  $i$ -th row and  $j$ -th column.  $\mathcal{CN}(x|\mu, \sigma^2)$  is the probability density function (pdf) of a complex Gaussian random variable (RV) with mean  $\mu$  and standard deviation  $\sigma$ .  $C(f_1||f_2)$  represents the Chernoff information for the probability distributions  $f_1$  and  $f_2$ . Given  $a < b$ ,  $[a, b]$  represents the segment of values between  $a$  and  $b$ .  $\langle a(t), b(t) \rangle = \int a(t)b(t)dt$  represents the correlation as applied to the given correlation interval.

## II. SYSTEM AND SIGNAL MODEL

*System Model:* As illustrated in Fig. 1, we study a multi-cell fog-based IoT system that aims at detecting QoIs, such as pollution level, based on measurements received from IoT devices. There are  $K$  cells, with a single-antenna EN and multiple IoT devices per cell. We assume that each QoI is described in each cell  $c \in \{1, \dots, K\}$  by a Random Variable (RV)  $\theta^c$ . RVs  $\{\theta^c\}_{c=1}^K$  are arbitrarily correlated across cells, and each device in cell  $c$  makes a noisy measurement of  $\theta^c$ . In this paper, we assume for simplicity of notation and analysis that each QoI can take two possible values, denoted as  $\theta_0$  and  $\theta_1$ .

The IoT devices are interrogated periodically by their local EN during  $L$  collection intervals, which are synchronized across all cells. In each collection interval, a random number of devices in each cell  $c$  transmit their measurements in the uplink using a grant-free access protocol based on TBMA [2]. Mathematically, in any collection interval  $l = 1, \dots, L$ , each IoT device in cell  $c$  is active probabilistically, independently of the observation being sensed, so that the total number  $N_l^c$  of devices active in collection interval  $l$  in cell  $c$  is a Poisson RV with mean  $\lambda$ . All devices share the same spectrum and hence their transmissions generally interfere, both within the same cell and across different cells.

We compare two different architectures to perform hypothesis testing in order to detect the QoIs: (i) *Edge-based Hypothesis Testing (EHT)*: Estimation of each QoI  $\theta^c$  is done locally at the EN in cell  $c$  based on the uplink signals received from the IoT devices, producing a local estimate  $\hat{\theta}^c$  (see Fig. 1); and (ii) *Cloud-based Hypothesis Testing (CHT)*: The ENs are connected with orthogonal finite-capacity digital fronthaul links to a cloud processor with fronthaul capacity of  $C$  [bit/s/Hz]. Each EN forwards the received signal upon quantization to the cloud processor using the fronthaul link. Unlike conventional C-RAN systems, here the goal is for the cloud to estimate all QoIs  $\{\theta^c\}_{c=1}^K$  (see Fig. 1).

*Signal Model:* When active, an IoT device  $i$  in cell  $c$  during the  $l$ -th collection makes a measurement  $X_{i,l}^c$ . We assume that the measurement takes values in an alphabet  $\{1, 2, \dots, M\}$  of

size  $M$ . The distribution of each observation  $X_{i,l}^c$  depends on the underlying QoI as

$$\Pr[X_{i,l}^c = m | \theta^c = \theta_0] = p_0^c(m) \quad \text{and} \quad \Pr[X_{i,l}^c = m | \theta^c = \theta_1] = p_1^c(m), \quad (1)$$

for  $m = 1, \dots, M$ . In words, devices in cell  $c$  make generally noisy measurements with  $\theta^c$ -dependent distributions  $p_0^c(\cdot)$  and  $p_1^c(\cdot)$ . When conditioned on QoIs  $\{\theta^c\}$ , measurements  $X_{i,l}^c$  are i.i.d. across all values of the cell index  $c$ , device index  $i$ , and the collection index  $l$ .

We denote by  $H_{i,l}^c \sim \mathcal{CN}(\mu_H, \sigma_H^2)$  the flat-fading Ricean fading channel, with mean  $\mu_H$  and variance  $\sigma_H^2$ , from device  $i$  to the EN in the same cell  $c$  during collection interval  $l$ ; and by  $G_{i,l}^{c,c'} \sim \mathcal{CN}(\mu_G, \sigma_G^2)$ , with mean  $\mu_G$  and variance  $\sigma_G^2$ , the flat-fading Ricean fading channel from device  $i$  in cell  $c' \neq c$  to the EN in cell  $c$  during collection interval  $l$ . All channels are assumed i.i.d. across indices  $i, l$  and  $c$ .

## III. COMMUNICATION PROTOCOL AND METRICS

In this section, we detail the communication protocol and the performance metrics used.

### A. Communication Protocol

Within the available bandwidth and time per-collection interval, as in [3], we assume the presence of  $M$  orthogonal waveforms  $\{\phi_m(t), m = 1, \dots, M\}$ , or preambles, with unit energy. According to TBMA, each waveform  $\phi_m(t)$  encodes the value  $m \in \{1, \dots, M\}$  of the observations of a device. The signal transmitted by a device  $i$  in cell  $c$  that is active in interval  $l$  is then given as  $S_{i,l}^c(t) = \sqrt{E_s} \phi_{X_{i,l}^c}(t)$ , that is, we have  $S_{i,l}^c(t) = \sqrt{E_s} \phi_m(t)$  if the observed signal is  $X_{i,l}^c(t) = m$ , where  $E_s$  is the transmission energy of a device per collection interval. Devices observing the same value  $m$  hence transmit using the same waveform. As a result, the spectral resources required by TBMA scale with the number  $M$  of observations values rather than with the total amount of packets sent by all the active devices, which may be much larger than  $M$ .

The received signal at the EN in cell  $c$  during the  $l$ -th collection can be written as

$$Y_l^c(t) = \sum_{i=1}^{N_l^c} H_{i,l}^c S_{i,l}^c(t) + \sum_{c'=1, c' \neq c}^K \sum_{i=1}^{N_l^{c'}} G_{i,l}^{c,c'} S_{i,l}^{c'}(t) + W_l^c(t), \quad (2)$$

where  $W_l^c(t) \sim \mathcal{CN}(0, W_0)$  is white Gaussian noise, i.i.d. over  $l$  and  $c$ , with power  $W_0$ . The first term in (2) represents the contribution from the IoT devices in the same cell  $c$ , while the second term represents the contribution from devices from the remaining cells  $c'$ . We emphasize that contributions related to the same preamble from different devices are not necessarily added coherently, but they only contribute to the average received energy for the preamble.

Given the orthogonality of the waveforms  $\{\phi_m(t)\}_{m=1}^M$ , a demodulator based on a bank of matched filters can be implemented at each EN without loss of optimality [2] (see [8] for extensions). After matched filtering of the received signal

with all waveforms  $\{\phi_m(t)\}_{m=1}^M$  each EN  $c$  obtains the  $M \times 1$  vector

$$\mathbf{Y}_l^c = \frac{1}{\sqrt{E_s}} [\langle \phi_1(t), Y_l^c(t) \rangle, \dots, \langle \phi_M(t), Y_l^c(t) \rangle]^\top$$

$$= \sum_{i=1}^{N_l^c} H_{i,l}^c \mathbf{e}_{X_{i,l}^c} + \sum_{\substack{c'=1 \\ c' \neq c}}^K \sum_{i=1}^{N_l^{c'}} G_{i,l}^{c,c'} \mathbf{e}_{X_{i,l}^{c'}} + \mathbf{W}_l, \quad (3)$$

where  $\mathbf{W}_l$  is a vector with i.i.d.  $\mathcal{CN}(0, \text{SNR}^{-1})$  elements, with  $\text{SNR} = E_s/W_0$ ; and  $\mathbf{e}_m$  represents an  $M \times 1$  unit vector with all zero entries except in position  $m$ .

For detection of the QoIs, we study both EHT and CHT:

**EHT:** Each EN  $c$  produces an estimate  $\hat{\theta}^c$  of the RV  $\theta^c$  based on the received signals  $\mathbf{Y}_l^c$  for all collection intervals  $l = 1, \dots, L$ , where  $\mathbf{Y}_l^c$  is given in (3).

**CHT:** Each EN  $c$  compresses the received signals  $\{\mathbf{Y}_l^c\}_{l=1}^L$  across all  $L$  collection intervals and sends the resulting compressed signals  $\{\hat{\mathbf{Y}}_l^c\}_{l=1}^L$  to the cloud. The cloud carries out joint detection of all QoIs  $\{\theta^c\}_{c=1}^K$  producing estimates  $\{\hat{\theta}^c\}_{c=1}^K$ .

## B. Performance Metrics

The performance of CHT and EHT will be evaluated in terms of the error exponent that describes the scaling of the joint error probability  $P_e$  as a function of the number  $L$  of collections. The joint error probability is given by

$$P_e = \Pr[\cup_{c=1}^K \{\hat{\theta}^c \neq \theta^c\}], \quad (4)$$

where  $\hat{\theta}^c$  is the estimate of the QoI  $\theta^c$  obtained at EN  $c$  or at the cloud, for EHT and CHT respectively. From large deviation theory, the detection error probability  $P_e$  decays exponentially as [9]

$$P_e = \exp(-LE + o(L)) \quad \text{with } L \rightarrow \infty, \quad (5)$$

where  $o(L)/L \rightarrow 0$  as  $L \rightarrow \infty$ , for some error exponent  $E$ . We will hence be interested in the rest of this paper in computing analytically the error exponent  $E$  for EHT and CHT.

## IV. ASYMPTOTIC PERFORMANCE

In this section, we derive the error exponent  $E$  in (5) for the optimal detection when the number of collection intervals  $L$  grows to infinity. In order to simplify the analysis, as in [2], we will take the assumption of large average number of active devices, i.e., of large  $\lambda$ . This scenario is particularly relevant for mMTC [1].

### A. Edge-based Hypothesis Testing

With EHT, each EN in cell  $c$  performs the binary test

$$\mathcal{H}_0^c : \theta^c = \theta_0 \text{ versus } \mathcal{H}_1^c : \theta^c = \theta_1 \quad (6)$$

based on the available received signals  $\mathbf{Y}^c = \{\mathbf{Y}_l^c\}_{l=1}^L$  in (3). The optimum Bayesian decision rule that minimizes the probability of error at each EN chooses the hypothesis with the Maximum A Posteriori (MAP) probability. The error exponent

$E$  in (5) using EHT can be lower bounded as shown in the following proposition.

**Proposition 1:** Under the optimal Bayesian detector, the error exponent  $E$  in (5) in the large- $\lambda$  regime and for any  $0 < \rho < 1$  is lower bounded as  $E \geq E^{\text{edge}} = \min_{c \in \{1, \dots, K\}} E^c$ , where

$$E^c = \min_{\mathbf{k} \in \{0,1\}^{K-1}} \max_{\alpha \in [0,1]} \left[ \frac{1}{2} \sum_{m=1}^M \log \left( \frac{\alpha \Sigma_{0,\mathbf{k}}^c(m, m) + (1-\alpha) \Sigma_{1,\mathbf{k}}^c(m, m)}{(\Sigma_{0,\mathbf{k}}^c(m, m))^\alpha (\Sigma_{1,\mathbf{k}}^c(m, m))^{1-\alpha}} \right) \right. \\ \left. + \frac{\alpha(1-\alpha)}{2} \sum_{m=1}^M \frac{(\mu_{0,\mathbf{k}}^c(m) - \mu_{1,\mathbf{k}}^c(m))^2}{(\alpha \Sigma_{0,\mathbf{k}}^c(m, m) + (1-\alpha) \Sigma_{1,\mathbf{k}}^c(m, m))} \right] \quad (7)$$

with  $\mathbf{k} = \{k_{c'}\}_{c'=1, c' \neq c}^K$

$$\mu_{k_c, \mathbf{k}}^c(m) = \mu_H \lambda p_{k_c}^c(m) + \mu_G \lambda \sum_{c=1, c \neq c'}^K p_{k_{c'}}^c(m), \quad (8)$$

and

$$\Sigma_{k_c, \mathbf{k}}^c(m, m) = \sigma_H^2 \lambda p_{k_c}^c(m) + \sigma_G^2 \lambda \sum_{c'=1, c' \neq c}^K p_{k_{c'}}^c(m) + \frac{1}{\text{SNR}}. \quad (9)$$

**Proof:** In a manner similar to [2, Theorem 3], the proof of the above theorem relies on the Central Limit Theorem (CLT) with random number of summands [9, p. 369] and on the Chernoff Information [9]. We refer to the Appendix for more details.  $\square$

The term in (7) being optimized over  $\mathbf{k}$  corresponds to the Chernoff information [9, Chapter 11] for the binary hypothesis test between the distributions of the received signal  $\mathbf{Y}_l^c$  under hypotheses  $\theta^c = \theta_0$  and  $\theta^c = \theta_1$  when  $\theta^{c'} = \theta_{k_{c'}}$ . In fact, for large values of  $\lambda$ , when  $\theta^c = \theta_{k_c}$  and  $\theta^{c'} = \theta_{k_{c'}}$ , the received signal  $\mathbf{Y}_l^c$  in (3) can be shown to be approximately distributed as  $\mathcal{CN}(\mu_{k_c, \mathbf{k}}^c, \Sigma_{k_c, \mathbf{k}}^c)$ , with mean vector  $\mu_{k_c, \mathbf{k}}^c = [\mu_{k_c, \mathbf{k}}^c(1), \dots, \mu_{k_c, \mathbf{k}}^c(M)]^\top$  and diagonal covariance matrix  $\Sigma_{k_c, \mathbf{k}}^c$  with diagonal elements  $\Sigma_{k_c, \mathbf{k}}^c(m, m)$ .

### B. Cloud-based Hypothesis Testing

The cloud tackles the  $2^K$ -ary hypothesis testing problem of distinguishing among hypotheses  $\mathcal{H}_{k_1, \dots, k_K} : (\theta^1, \dots, \theta^K) = (\theta_{k_1}, \dots, \theta_{k_K})$  for  $k_c \in \{0, 1\}$  on the basis of the quantized signals  $\{\hat{\mathbf{Y}}_l\}_{l=1}^L$  received from both ENs on the fronthaul links.

Following a standard approach, see, e.g., [10], the impact of fronthaul quantization is modeled as an additional quantization noise. In particular, the signal received at the cloud from EN  $c$  can be written accordingly as  $\hat{\mathbf{Y}}_l^c = \mathbf{Y}_l^c + \mathbf{Q}_l^c$ , where  $\mathbf{Q}_l^c$  represents the quantization noise vector. As in most prior references (see, e.g., [10]), the quantization noise vector  $\mathbf{Q}_l^c$  is assumed to have i.i.d. elements being normally distributed with zero mean and variance  $\sigma_{q^c}^2$ . Furthermore, from rate-distortion theory, the fronthaul capacity constraint implies the following inequality, for each EN  $c$

$$MC \geq I(\mathbf{Y}_l^c; \hat{\mathbf{Y}}_l^c). \quad (10)$$

This is because the number of bits available to transmit each measurement  $\hat{\mathbf{Y}}_l^c$  is given by  $C$  bits per symbol, or, equivalently, per orthogonal spectral resource; that is,  $MC$  bits in total for all  $M$  resources. From (10), one can in principle derive

the quantization noise power  $\sigma_{q^c}^2$ . However, evaluating the mutual information in (10) directly is difficult due to the non-Gaussianity of the received signals  $\mathbf{Y}_l^c$ . To tackle this issue, we bound the mutual information term in (10) using the property that the Gaussian distribution maximizes the differential entropy under covariance constraints [9], obtaining the following Lemma. In what follows, we denote  $\mathbf{k} = \{k_c\}_{c=1}^K$ .

*Lemma 1:* The quantization noise power can be upper bounded as  $\sigma_{q^c}^2 \leq \bar{\sigma}_{q^c}^2$ , where  $\bar{\sigma}_{q^c}^2$  is obtained by solving the non-linear equation

$$MC = \frac{1}{2} \sum_{m=1}^M \log \left( \frac{\sum_{\mathbf{k} \in \{0,1\}^K} \Pr(\theta^1 = \theta_{k_1}, \dots, \theta^K = \theta_{k_K}) \Sigma_{\mathbf{k}}^c(m, m) + \sigma_{q^c}^2}{(\sigma_{q^c}^2)^M} \right). \quad (11)$$

with  $\Sigma_{\mathbf{k}}^c$  given in equation (9).

*Proof:* See [7, Appendix A] for details.  $\square$

*Proposition 2:* Under optimal detection, the error exponent  $E$  in (5) in the large- $\lambda$  regime for CHT can be lower bounded as  $E \geq E^{\text{cloud}} = \min_{\mathbf{k} \in \{0,1\}^K} E_{\mathbf{k}}$ , where

$$E_{\mathbf{k}} = \min_{\mathbf{k}' \neq \mathbf{k}} \max_{\alpha \in [0,1]} \left[ \frac{1}{2} \log \frac{|\alpha \Sigma_{\mathbf{k}} + (1-\alpha) \Sigma_{\mathbf{k}'}|}{|\Sigma_{\mathbf{k}}|^\alpha |\Sigma_{\mathbf{k}'}|^{1-\alpha}} + \frac{\alpha(1-\alpha)}{2} (\boldsymbol{\mu}_{\mathbf{k}} - \boldsymbol{\mu}_{\mathbf{k}'})^T (\alpha \Sigma_{\mathbf{k}} + (1-\alpha) \Sigma_{\mathbf{k}'})^{-1} \right. \\ \left. \times (\boldsymbol{\mu}_{\mathbf{k}} - \boldsymbol{\mu}_{\mathbf{k}'}) \right], \quad (12)$$

where the entries of the  $KM \times 1$  vector  $\boldsymbol{\mu}_{\mathbf{k}}$  are defined as

$$\mu_{\mathbf{k}}(m') = \mu_{\mathbf{k}}^c(m) \quad \text{for } m' = (c-1)M, \dots, cM \quad (13)$$

with  $\mu_{\mathbf{k}}^c(m)$  defined in (8), and the entries of the  $2M \times 2M$  covariance matrix  $\Sigma_{\mathbf{k}}$  given as

$$\Sigma_{\mathbf{k}}(m', m') = \Sigma_{\mathbf{k}}^1(m, m) + \sigma_{q^c}^2 \quad \text{for } m' = (c-1)M, \dots, cM, \\ \Sigma_{\mathbf{k}}(m', M+m') = \Sigma_{\mathbf{k}}(M+m', m') = \\ p_{k_c}^c(m)(1-p_{k_c}^c(m))\lambda\mu_H\mu_G + p_{k_c'}^c(m)(1-p_{k_c'}^c(m))\lambda\mu_H\mu_G \\ \text{for } m = 1, \dots, M, \text{ and } m' = (c-1)M, \dots, cM, \quad (14)$$

where  $\Sigma_{\mathbf{k}}^c(m, m)$  is defined in (9) and all other entries of matrix  $\Sigma_{\mathbf{k}}$  are zero.

*Proof:* The proof follows in a manner similar to *Proposition 1* and uses Sanov's Theorem [9, p. 362] as detailed in [7, Appendix C].  $\square$

The term in (12) being optimized over  $\mathbf{k}' \in \{0,1\}^K$  corresponds to the Chernoff information for the binary test between the distribution of the signal received at the cloud under hypotheses  $\mathcal{H}_{\mathbf{k}} : (\theta^1 = \theta_{k_1}, \dots, \theta^K = \theta_{k_K})$  and  $\mathcal{H}_{\mathbf{k}'} : (\theta^1 = \theta_{k'_1}, \dots, \theta^K = \theta_{k'_K})$ . As discussed above, for large  $\lambda$ , the signal received at the cloud under hypothesis  $\mathcal{H}_{\mathbf{k}}$  is approximately distributed as  $\mathcal{CN}(\boldsymbol{\mu}_{\mathbf{k}}, \Sigma_{\mathbf{k}})$ , where the elements of the mean vector  $\boldsymbol{\mu}_{\mathbf{k}}$  and covariance matrix  $\Sigma_{\mathbf{k}}$  are described in (13) and (14).

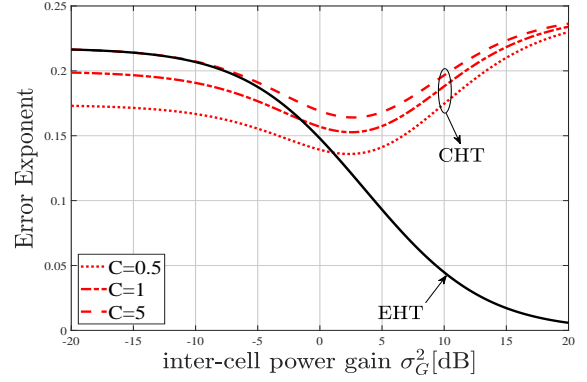


Fig. 2: Error exponent for EHT and CHT as function of the inter-cell power gain  $\sigma_G^2$  ( $\mu_H = 1, \sigma_H^2 = 1, \mu_G = 0, \lambda = 4$ , and SNR = -1 dB).

### C. Edge vs Cloud-Based Hypothesis Testing

In this section, we prove that the performance of CHT is superior to EHT as long as the inter-cell channel power gain power  $\sigma_G^2$  is sufficiently large. The main result can be summarized in the following theorem.

*Theorem 1:* The error exponents derived in *Proposition 1* and *Proposition 2* satisfy the following limits

$$\lim_{\sigma_G^2 \rightarrow \infty} E^{\text{edge}} = 0 \quad \text{and} \quad \lim_{\sigma_G^2 \rightarrow \infty} E^{\text{cloud}} > 0. \quad (15)$$

*Proof:* The proof can be found in [7, Appendix D].  $\square$

*Theorem 1* implies that, for high inter-cell power gains, EHT leads to vanishing error exponent, while this is not the case for CHT. This demonstrates that the performance of EHT is inter-cell interference limited, while this is not the case for CHT. In practice, as shown via numerical results in Sec. V, fairly low interference levels are sufficient for CHT to outperform EHT.

## V. NUMERICAL RESULTS

In this section, we provide numerical simulations to evaluate the performance of both CHT and EHT. Unless specified otherwise, we fix the following values for the parameters  $\mu_H = 1, \sigma_H^2 = 1, \mu_G = 0, \lambda = 4$  and  $K = 2$  cells. The joint distribution of QoIs is defined as

$$f(\theta^c, \theta^{c'}) = \frac{\rho}{2} 1_{\{\theta^c = \theta^{c'}\}} + \frac{1-\rho}{2} 1_{\{\theta^c \neq \theta^{c'}\}} \quad (16)$$

where  $0 \leq \rho \leq 1$  represents a “correlation” parameter that measures the probability that the two QoIs have the same value, i.e.,  $\rho = \Pr[\theta^c = \theta^{c'}]$ . Note that under (16), both values of the QoI are equiprobable, i.e.,  $\Pr(\theta^c = \theta_j) = 0.5$  for  $j \in \{0,1\}$  and  $c \in \{1,2\}$ . Furthermore, when  $\rho = 0.5$ , the two QoIs are independent.

In Fig. 2, we plot the error exponent for both EHT and CHT with different values of  $C$  as function of the inter-cell power gain  $\sigma_G^2$ . As  $\sigma_G^2$  increases, the performance of edge detection is seen to decrease, since interference from the other cell is treated as noise by the edge. In contrast, inline with the theoretical results in *Theorem 1*, CHT is able to benefit from a sufficiently

large inter-cell interference due to centralized processing. We note that, the same U-shaped behavior is observed for the uplink throughput in C-RAN as function of the inter-cell interference [11]. Furthermore, a larger fronthaul capacity leads to an improved detection performance, since measurements are received at the cloud with a better resolution.

In Fig. 3, we plot the error exponent as function of the fronthaul capacity  $C$ . For low values of  $C$ , EHT outperforms CHT since in this regime, the quantization noise is large and thus measurements are received with low resolution. In contrast, CHT outperforms EHT for high enough values of  $C$ .

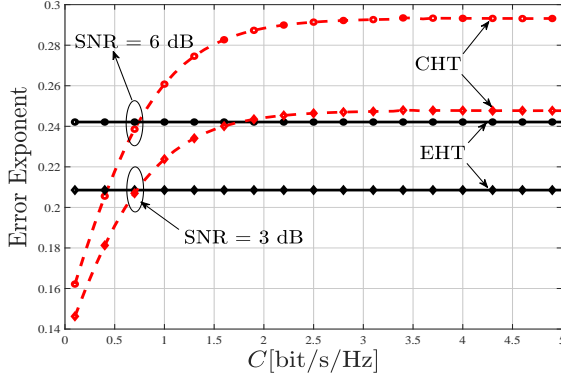


Fig. 3: Error exponent for EHT and CHT as function of the fronthaul capacity  $C$  ( $\mu_H = 1, \sigma_H^2 = 1, \mu_G = 0, \sigma_G^2 = 0$ , and  $\lambda = 4$ ).

## VI. CONCLUSIONS

This paper considers the problem of detection of Quantities of Interest (QoIs) at the edge or the cloud of a fog-based IoT network. The performance of cloud-based detection was demonstrated analytically and via numerical results to be superior to edge-based detection for sufficiently high fronthaul capacity and inter-cell interference. As for future research directions, we mention the study of the coexistence of heterogeneous IoT services with different service requirements.

## APPENDIX: PROOF OF PROPOSITION 1

From the union bound  $P_e \leq \sum_{c=1}^K P_e^c$  with  $P_e^c = \Pr[\hat{\theta}^c \neq \theta^c]$  and the identity  $P_e^c = \frac{1}{2^{K-1}} \sum_{\mathbf{k} \in \{0,1\}^{K-1}} \Pr[\hat{\theta}^c \neq \theta^c | \{\theta^{c'} = \theta_{k'_c}, \text{ for } c' \neq c \in \{1, \dots, K\}\}]$ , we directly obtain the lower bound on the error exponent

$$E \geq \min_{c \in \{1, \dots, K\}} \min_{\mathbf{k} \in \{0,1\}^{K-1}} E^c, \quad (17)$$

where  $E^c = -\lim_{L \rightarrow \infty} \frac{1}{L} \log \Pr[\hat{\theta}^c \neq \theta^c | \{\theta^{c'} = \theta_{k'_c}, \text{ for } c' \neq c \in \{1, \dots, K\}\}]$  is the error exponent for detection of QoI  $\theta^c$  conditioned on the condition  $\theta^{c'} = \theta_{k'_c}$  for  $c' \neq c$ . Under optimal Bayesian detection, the error exponent  $E_k^c$  is given by the Chernoff information [9, Chapter 11] as

$$E^c = C(f_{0,\mathbf{k}}(\mathbf{Y}_l^c), f_{1,\mathbf{k}}(\mathbf{Y}_l^c)), \quad (18)$$

where we have denoted  $f_{k_c,\mathbf{k}}(\mathbf{Y}_l^c) = f(\mathbf{Y}_l^c | \theta^c = \theta_{k'_c}, \{\theta^{c'} = \theta_{k'_c}, \text{ for } c' \neq c \in \{1, \dots, K\}\})$  for brevity. Computing the error exponent in (18) requires finding the distributions  $f_{k_c,\mathbf{k}}(\mathbf{Y}_l^c)$ .

Following [2], this can be approximated by a Gaussian distribution in the regime of large  $\lambda$  thanks to the Central Limit Theorem (CLT) with random number of summands [12, p. 369]. In particular, referring to [2] for details, we can conclude that, when  $\lambda \rightarrow \infty$ , the conditional distribution  $f_{k_c,\mathbf{k}}(\mathbf{Y}^c)$  tends in distribution to  $\mathcal{CN}(\boldsymbol{\mu}_{k_c,\mathbf{k}}, \boldsymbol{\Sigma}_{k_c,\mathbf{k}})$ , where  $\boldsymbol{\mu}_{k_c,\mathbf{k}}$  and  $\boldsymbol{\Sigma}_{k_c,\mathbf{k}}$  are the mean vector and covariance matrix respectively when  $\theta^c = \theta_{k'_c}$  and  $\theta^{c'} = \theta_{k'_c}$  and are defined in (8) and (9).

The Chernoff Information between two Gaussian distributions can be obtained by maximizing over  $\alpha \in [0, 1]$  the  $\alpha$ -Chernoff information defined as [13]

$$\begin{aligned} C_\alpha(f_{0,\mathbf{k}}(\mathbf{Y}_l^c), f_{1,\mathbf{k}}(\mathbf{Y}_l^c)) = & \frac{1}{2} \log \frac{|\alpha \boldsymbol{\Sigma}_{0,\mathbf{k}} + (1-\alpha) \boldsymbol{\Sigma}_{1,\mathbf{k}}|}{|\boldsymbol{\Sigma}_{0,\mathbf{k}}|^\alpha |\boldsymbol{\Sigma}_{1,\mathbf{k}}|^{1-\alpha}} + \frac{\alpha(1-\alpha)}{2} (\boldsymbol{\mu}_{0,\mathbf{k}} - \boldsymbol{\mu}_{1,\mathbf{k}})^\top \\ & \times (\alpha \boldsymbol{\Sigma}_{0,\mathbf{k}} + (1-\alpha) \boldsymbol{\Sigma}_{1,\mathbf{k}})^{-1} (\boldsymbol{\mu}_{0,\mathbf{k}} - \boldsymbol{\mu}_{1,\mathbf{k}}). \end{aligned} \quad (19)$$

By plugging in (17) and (19) the expressions of  $\boldsymbol{\mu}_{k_c,\mathbf{k}}$  and  $\boldsymbol{\Sigma}_{k_c,\mathbf{k}}$  and using (18) we obtain the desired result.  $\square$

## REFERENCES

- [1] Z. Zhang, Y. Xiao, Z. Ma, M. Xiao, Z. Ding, X. Lei, G. K. Karagiannidis, and P. Fan, "6G Wireless Networks: Vision, Requirements, Architecture, and Key Technologies," *IEEE Vehicular Technology Magazine*, vol. 14, no. 3, pp. 28–41, 2019.
- [2] A. Anandkumar and L. Tong, "Type-Based Random Access for Distributed Detection Over Multiaccess Fading Channels," *IEEE Trans. Signal Process.*, vol. 55, no. 10, pp. 5032–5043, Sept. 2007.
- [3] G. Mergen and L. Tong, "Type Based Estimation over Multiaccess Channels," *IEEE Trans. Signal Process.*, vol. 54, no. 2, pp. 613–626, Feb. 2006.
- [4] K. Liu and A. M. Sayeed, "Type-Based Decentralized Detection in Wireless Sensor Networks," *IEEE Trans. Signal Process.*, vol. 55, no. 5, pp. 1899–1910, May 2007.
- [5] N. Weinberger, Y. Kochman, and M. Wigger, "Exponent Trade-off for Hypothesis Testing Over Noisy Channels," in *IEEE Inter. Symp. Inf. Th. (ISIT)*, July 2019, pp. 1852–1856.
- [6] S. Salehkalaibar and M. Wigger, "Distributed Hypothesis Testing with Variable-Length Coding," *arXiv preprint arXiv:1910.08261*, 2019.
- [7] R. Kassab, O. Simeone, and P. Popovski, "Information-Centric Grant-Free Access for IoT Fog Networks: Edge vs Cloud Detection and Learning," *CoRR*, vol. abs/1907.05182, 2019. [Online]. Available: <http://arxiv.org/abs/1907.05182>
- [8] J. Dommel, Z. Utkovski, S. Stanczak, and O. Simeone, "Joint Source-Channel Coding and Bayesian Message Passing Detection for Grant-Free Radio Access in IoT," *arXiv preprint arXiv:1910.11704*, 2019.
- [9] T. M. Cover and J. A. Thomas, *Elements of information theory*. John Wiley & Sons, 2012.
- [10] T. Quek, M. Peng, O. Simeone, and W. Yu, *Cloud Radio Access Networks: Principles, Technologies, and Applications*. Cambridge University Press, 2017.
- [11] O. Simeone, N. Levy, A. Sanderovich, O. Somekh, B. M. Zaidel, H. V. Poor, and S. Shamai, "Cooperative Wireless Cellular Systems: An Information-Theoretic View," *Foundations and Trends in Communications and Information Theory*, vol. 8, no. 1-2, pp. 1–177, 2012.
- [12] P. Billingsley, *Probability and measure*. John Wiley & Sons, 2008.
- [13] F. Nielsen, "Chernoff information of exponential families," *arXiv preprint arXiv:1102.2684*, 2011.