

Low-Complexity Sub-Optimal Cell ID Estimation in NB-IoT System

Vincent Savaux ¹ and Matthieu Kanj ²

¹bjç.com

²Affiliation not available

October 30, 2023

Abstract

This paper deals with cell ID estimation in narrowband-internet of things (NB-IoT) system. The cell ID value is carried by the narrowband secondary synchronization signal (NSSS). We suggest a low-complexity sub-optimal estimator, based on the auto-correlation of the received observations. It is up to thirty times less complex than the optimal maximum likelihood (ML) estimator based on cross-correlation. In addition, we present three methods allowing the receiver to take advantage of the different repetitions of the NSSS. They are based on a hard decision after every estimation, a soft combination of the different observations of the NSSS, and an hybrid mix between the two firsts, respectively. The advantages and drawbacks of the presented techniques are stated, and a performance analysis is proposed, which is further discussed through simulations results. It is shown the that different methods reach the performance of ML after several repetitions for a lower overall complexity.

This paper is a preprint of a paper submitted to IET Communications. If accepted, the copy of record will be available at the IET Digital Library.

Low-Complexity Sub-Optimal Cell ID Estimation in NB-IoT System

Vincent Savaux^{1*}, Matthieu Kanj¹

¹ Advanced Connectivity Lab, b<>com, Rennes, France

* E-mail: vincent.savaux@b-com.com

Abstract: This paper deals with cell ID estimation in narrowband-internet of things (NB-IoT) system. The cell ID value is carried by the narrowband secondary synchronization signal (NSSS). We suggest a low-complexity sub-optimal estimator, based on the auto-correlation of the received observations. It is up to thirty times less complex than the optimal maximum likelihood (ML) estimator based on cross-correlation. In addition, we present three methods allowing the receiver to take advantage of the different repetitions of the NSSS. They are based on a hard decision after every estimation, a soft combination of the different observations of the NSSS, and an hybrid mix between the two firsts, respectively. The advantages and drawbacks of the presented techniques are stated, and a performance analysis is proposed, which is further discussed through simulations results. It is shown that different methods reach the performance of ML after several repetitions for a lower overall complexity.

1 Introduction

The machine-type communication market is growing rapidly due to the high utilization of connected devices. The usage of these devices is continuously growing with the expansion of the applications proposed to individuals and industries. Predictions indicate that 1.5 billion of connected devices will be deployed within the next few years, where 70 % of them are for wide area applications [1]. However, even for only wide-area applications, various and different requirements may be needed depending on the design objective of each internet of things (IoT) network.

To follow the needs of the IoT market, the 3rd generation partnership project (3GPP) has introduced a new cellular technology standard called narrowband-internet of things (NB-IoT) to provide IoT services through wide-area cellular networks. This new system was designed to take into account a maximum number of requirements for the IoT services. In particular, a very good indoor coverage, support for a very large number of low-speed connected devices, low latency, very low cost of connectivity, low power consumption and optimized network architecture.

The narrowband-internet of things NB-IoT system is inherited from long term evolution (LTE) system [2], with a reduction of complexity for several features in order to respect the low-cost and low-power constraints of NB-IoT user equipment (UE) modules. The system was conceived to occupy a frequency band of 180 KHz (this corresponds to one resource block in LTE system), and the number of channels and signals was reduced. However, the essential channels and signals remain unchanged, but they were adapted to fit the new NB-IoT frame structure [3–5].

The attachment procedure between a UE and an evolved Node B (eNB) starts with the synchronization process. In NB-IoT, it is split into two successive steps: the UE physically synchronizes with the eNB in time and frequency thanks to the narrowband primary synchronization signal (NPSS) [6–9], and then estimates the eNB's cell ID carried by the narrowband secondary synchronization signal (NSSS) [10–12]. Only few papers deal with the cell ID estimation in NB-IoT. An exhaustive cell ID search (among 504 possible values) is presented in [10], and simplifications are provided in [11, 12], dividing by four and sixteen the number of operations compared with the exhaustive search. However, it remains that

these methods are based on maximum likelihood (ML), leading to a cross-correlation between the observation of the NSSS and a reference signal.

In this paper, we suggest to use auto-correlation of the NSSS instead of cross-correlation to reduce the complexity of the cell ID estimation process. The method is based on the symmetry of the NSSS samples. It is up to thirty times less complex than the exhaustive ML search. In addition, we describe three methods to take advantage of the NSSS repetitions. The first consists in making estimation attempts every NSSS repetitions until good cell ID estimation. The second one softly combines successive observations of the NSSS before the estimation step. The third one is an hybrid technique that combines the principles of the two firsts techniques. All these methods are compared through simulations, and it is shown that they can reach the performance of the exhaustive ML search after few repetitions of NSSS.

The rest of this paper is organized as follows: in Section 2 we describe the NSSS model and we present the exhaustive cell ID search. Section 3 introduces the proposed estimation method based on the auto-correlation, and Section 4 presents the three techniques to take advantage of the NSSS repetitions. We present and discuss the simulations results in Section 5. Finally, Section 6 concludes this paper.

2 System Model

2.1 NSSS Model

This section is dedicated to the description of the NSSS model [2]. It is assumed that the time-frequency synchronization has been properly performed thanks to NPSS [6–9], in such way that the residual synchronization errors are negligible. Furthermore, since the signal is narrowband, it can be reasonably supposed that the frequency response of the channel is constant over the NSSS resource block of 12 subcarriers. Thus, after demodulation and inverse mapping from resource elements, the $N = 132$ samples that compose the received NSSS sequence can be expressed as

$$y_n = \alpha d_n + w_n, \quad (1)$$

where $n = 0, 1, \dots, 131$, α is the complex channel coefficient, $w_n \sim \mathcal{CN}(0, \sigma^2)$ is the additive white Gaussian noise

(AWGN), and d_n is the transmitted NSSS sequence, which is defined in [2] as

$$d_n = b_q(m)e^{-2j\pi\theta_f n} e^{-j\frac{\pi u n'(n'+1)}{131}}. \quad (2)$$

where $\{b_q(m) \in \{-1, 1\}\}$ is one of the four Hadamard sequences defined in Table 10.2.7.2.1-1 in [2]. The table has been reproduced in Table 1 in the present paper. Moreover, we have

$$\theta_f = \frac{1}{4}\left(\frac{n_f}{2}\right) \mod 4, \quad (3)$$

where n_f is the frame number. Note that, n_f takes only even values since NSSS is only transmitted in even frames. Furthermore, we have:

$$\begin{aligned} n' &= n \mod 131 \\ m &= n \mod 128 \\ u &= N_{ID}^{N_{cell}} \mod 126 + 3 \\ q &= \lfloor \frac{N_{ID}^{N_{cell}}}{126} \rfloor, \end{aligned}$$

with $N_{ID}^{N_{cell}}$ the cell ID of the eNB.

It must be emphasized that θ_f in (3) can be simply rewritten as $\theta_f \in \{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}\}$ according to n_f value. Therefore, the sequence $\{e^{-2j\pi\theta_f n}\}$, $n = 0, \dots, 131$ in (2) is chosen among the following ones:

$$\theta = \begin{cases} \mathbf{o}_{132}, & \text{if } \theta_f = 0 \\ \mathbf{o}_{33} \otimes [1, -j, -1, j], & \text{if } \theta_f = \frac{1}{4} \\ \mathbf{o}_{66} \otimes [1, -1], & \text{if } \theta_f = \frac{1}{2} \\ \mathbf{o}_{33} \otimes [1, j, -1, -j], & \text{if } \theta_f = \frac{3}{4} \end{cases}, \quad (4)$$

where \mathbf{o}_L is the $1 \times L$ vector composed of one, and \otimes is the Kronecker product. We can rewrite (1) in vector form as follows:

$$\begin{aligned} \mathbf{y} &= \alpha \mathbf{d} + \mathbf{w} \\ &= \alpha \underbrace{\mathbf{b}\theta}_{\mathbf{X}(q, \theta_f)} \mathbf{E}(u) + \mathbf{w}, \end{aligned} \quad (5)$$

where $\mathbf{E}(u)$ is the $N \times 1$ vector containing the Zadoff-Chu sequence $E_n(u) = e^{-j\frac{\pi u n'(n'+1)}{131}}$ [13, 14]. \mathbf{b} and $\mathbf{X}(q, \theta_f)$ are the $N \times N$ diagonal matrices composed of the elements b_q and $b_q(m)e^{-2j\pi\theta_f n}$ respectively, and \mathbf{w} is the $N \times 1$ vector of white Gaussian noise such as $w_n \sim \mathcal{CN}(0, \sigma^2)$. In this respect, the estimation of the cell ID results in finding the parameters (u, q, θ_f) from the observation \mathbf{y} .

2.2 Exhaustive Cell ID Search Through ML Estimator

The exhaustive search of the optimal parameters u , q , and θ_f through ML estimator can be expressed as

$$(\hat{u}, \hat{q}, \hat{\theta}_f) = \arg \max_{u, q, \theta_f} C e^{-\frac{1}{\sigma^2} \|\mathbf{y} - \alpha \mathbf{X}(q, \theta_f) \mathbf{E}(u)\|^2}, \quad (6)$$

where C is a constant that can be omitted in the following since it does not depends on (u, q, θ_f) , and $\|\cdot\|$ is the Euclidian norm. Since $|X_n|^2 = 1$ and $|E_n(u)|^2 = 1$ for any $X_n \in \{\pm 1, \pm j\}$, we can develop (6) as

$$\begin{aligned} (\hat{u}, \hat{q}, \hat{\theta}_f) &= \arg \max_{u, q, \theta_f} e^{-\frac{1}{\sigma^2} \|\mathbf{y}\|^2} e^{-\frac{|\alpha|^2 N}{\sigma^2}} \\ &\quad e^{\frac{1}{\sigma^2} \text{Re}\{\mathbf{y}^H \alpha \mathbf{X}(q, \theta_f) \mathbf{E}(u)\}} \\ &= \arg \max_{u, q, \theta_f} \text{Re}\{\mathbf{y}^H \alpha \mathbf{X}(q, \theta_f) \mathbf{E}(u)\}. \end{aligned} \quad (7)$$

We conclude from (7) that the optimal parameters u , q , and θ_f are those that maximize the cross-correlation as follows:

$$(\hat{u}, \hat{q}, \hat{\theta}_f) = \arg \max_{u, q, \theta_f} \text{Re}\left\{ \sum_{n=0}^{131} y_n^* \alpha X_n(q, \theta_f) E_n(u) \right\}. \quad (8)$$

This exhaustive search is the optimal solution in the ML sense but is also the most complex as it requires $N_m = 16 \times 126 \times 132 = 266112$ complex multiplications, where 16 corresponds to $q \times \theta_f$ possible values, and 126 corresponds to u possible values. Furthermore, it requires the *a priori* knowledge of the channel coefficient α . In order to bypass the latter constraint, it can be either assumed that the channel has been estimated by means of NPSS signal, or the ML estimator in (8) can be substituted by a sub-optimal one as

$$\begin{aligned} (\hat{u}, \hat{q}, \hat{\theta}_f) &= \arg \max_{u, q, \theta_f} \left| \sum_{n=0}^{131} y_n^* \alpha X_n(q, \theta_f) E_n(u) \right| \\ &= \arg \max_{u, q, \theta_f} \left| \sum_{n=0}^{131} y_n^* X_n(q, \theta_f) E_n(u) \right|. \end{aligned} \quad (9)$$

Once q and u are estimated, then the cell ID is in turn estimated as

$$\hat{N}_{ID}^{N_{cell}} = 126 \cdot \hat{q} + \hat{u} - 3. \quad (10)$$

In the following, we suggest an alternative sub-optimal ML-based solution where u and (q, θ_f) are sequentially estimated. This will allow to greatly reduce the complexity of cell ID estimation compared with the exhaustive search (8).

3 Suggested Cell ID Estimation

3.1 Description of the Algorithm

The algorithm generally breaks down into two main stages: estimating (q, θ_f) and then u using the estimates of (q, θ_f) . The originality of the suggested method consists in taking advantage of the symmetry property of the ZC sequence $\mathbf{E}(u)$ [13, 14] to estimate the parameters (q, θ_f) thanks to a sub-optimal auto-correlation function. Thus, for any $u = 3, 4, \dots, 128$ and $n = 0, 1, \dots, 130$ the symmetry of the sequence $\mathbf{E}(u)$ with respect to the 66-th sample can be expressed as

$$E_n(u) = E_{130-n}(u). \quad (11)$$

Furthermore, it can be noted that $E_{131}(u) = E_{130}(u) = E_0(u)$. This property is illustrated in Fig. 1 which separately shows the real and the imaginary parts of the ZC sequence $E_n(u)$ for $u = 85$ in Fig. 1-(a) and (b). The symmetry axis in dashed line is shown in order to highlight the symmetry between the elements $E_n(u)$.

The first step of the suggested method is to recover the symmetry property of the ZC sequence from the observation y_n to estimate (q, θ_f) . To this end, we define $z_n(q', \theta'_f)$ as follows:

3. **Estimate** u thanks to (16) or (17) according to the *a priori* knowledge or not of the channel coefficient α , respectively.

4. **Estimate the cell ID** using (10).

The total number of complex multiplications involved in the suggested method is reduced to $8 \times 65 + 132 \times 128 = 17416$, where 8×65 corresponds to (15), and 132×128 corresponds to (16) or (17). Thus, it is of order 15 times less complex than the exhaustive cell ID search through (8) or (9). Moreover, we can still reduce the complexity by using the symmetry of $E_n(u)$ allowing us to factorize elements in the sum in (16) and (17). Therefore, let us consider (16), which can be rewritten as

$$\begin{aligned} \hat{u} = \arg \max_{u, \beta} \operatorname{Re} \left\{ \sum_{n=0}^{64} E_n(u) [y_n^* \alpha (-1)^{n\beta} X_n(\hat{q}, \hat{\theta}_f) \right. \\ \left. + y_{130-n}^* \alpha (-1)^{(130-n)\beta} X_{130-n}(\hat{q}, \hat{\theta}_f)] \right. \\ \left. + E_{65}(u) (y_{65}^* \alpha (-1)^{65\beta} X_{65}(\hat{q}, \hat{\theta}_f)) \right. \\ \left. + E_{131}(u) (y_{131}^* \alpha (-1)^{131\beta} X_{131}(\hat{q}, \hat{\theta}_f)) \right\}. \end{aligned} \quad (18)$$

With this new form, the number of multiplications is divided by two, and thus the complexity is reduced to 8708 operations, *i.e.* thirty times less complex than the exhaustive ML search. However, the suggested technique should feature a reduced performance compared to the exhaustive search, since (15) is a sub-optimal technique in the ML sense due to the auto-correlation. As a consequence, it is interesting to take advantage of the repetitions of NSSS such as designed in NB-IoT. In Section 4, we present different methods that combines the successive observations of NSSS to improve the performance of the suggested cell ID estimation technique.

3.2 Probability of Error of the Estimator

This section investigates the error probability of estimation of the cell ID by means of the suggested method. For a sake of conciseness, we define the events $A = (\hat{q} \neq q \cap \hat{u} = u)$, $B = (\hat{q} = q \cap \hat{u} \neq u)$, and $C = (\hat{q} \neq q \cap \hat{u} \neq u)$. We denote by \mathcal{P} the error probability, which can be defined as

$$\begin{aligned} \mathcal{P} &= \mathbb{P}(\hat{N}_{ID}^{N_{cell}} \neq N_{ID}^{N_{cell}}) \\ &= \mathbb{P}(A \cup B \cup C), \end{aligned} \quad (19)$$

which can be reduced thanks to the inclusion-exclusion principle as

$$\mathcal{P} = \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C), \quad (20)$$

since the probability of any intersection of events (*e.g.* $\mathbb{P}(A \cap B)$) is equal to zero. Furthermore, since the estimation of u follows that of q , it can be reasonably assumed that the probability of the event $\hat{u} = u$ conditionally to $\hat{q} \neq q$ is zero. This mathematically highlights the fact that u cannot be properly estimated if q is not. We then deduce that

$$\begin{aligned} \mathbb{P}(\hat{u} = u | \hat{q} \neq q) &= 0 \\ \Leftrightarrow \mathbb{P}(\hat{u} = u \cap \hat{q} \neq q) &= 0. \end{aligned} \quad (21)$$

We assume that the events in B and C are independent, *i.e.* u can be badly estimated for any good or bad estimation of q (this assumption will further discussed afterward). Then, with the substitution of (21) into (20), we obtain

$$\begin{aligned} \mathcal{P} &= \mathbb{P}(B) + \mathbb{P}(C) \\ &= \mathbb{P}(\hat{q} = q) \mathbb{P}(\hat{u} \neq u) + \mathbb{P}(\hat{q} \neq q) \mathbb{P}(\hat{u} \neq u) \\ &= \mathbb{P}(\hat{u} \neq u), \end{aligned} \quad (22)$$

since $\mathbb{P}(\hat{q} = q) + \mathbb{P}(\hat{q} \neq q) = 1$. On one hand, we deduce from (22) that is sufficient to know $\mathbb{P}(\hat{u} \neq u)$ to evaluate the performance of the cell ID estimator. Otherwise, it could be misinterpreted that the performance of the estimator only depends on the performance of estimation of u in (16) or (17). However, it must be emphasized that the estimators in (16) and (17) also depend on $X_n(\hat{q}, \hat{\theta}_f)$, and are therefore implicitly linked to the performance of estimation of q . On the other hand, it is not possible to derive an analytical closed-form expression of \mathcal{P} because it involves the multiple integrals of multivariate Gaussian distribution.

4 Taking Advantage of NSSS Repetitions

In this section, we suggest three methods to take advantage of the repetitions of NSSS signal in order to improve the estimation of the cell ID. The first one is based on the decision of the cell ID estimator and is qualified as "hard decision" method. The second one is based on the combination of successive observations of the NSSS signals and is qualified as "soft combination". The last one is an hybrid method that combines both soft combination and hard decision.

4.1 Hard Decision

The principle of the method is straightforward: the UE continues estimating the cell ID from successive observations of NSSS while $N_{ID}^{N_{cell}}$ is not properly estimated. The way the UE verifies if the cell ID is correctly estimated is not detailed in this paper, but it can be reminded that the cell ID is used to generate the scrambling sequence of the broadcast signal. Thus, if the UE is not able to properly decode this signal (the cyclic redundancy check is false), then it decides that the estimated cell ID is false, and tries again to estimate $N_{ID}^{N_{cell}}$ during the next NSSS opportunity.

It is possible to evaluate the probability of error for the hard decision method from (22). Thus, we assume that all consecutive cell ID estimation attempts are independent. Then, the new error probability for a UE to properly estimate the cell ID at the N_R -th attempt (knowing that the previous $N_R - 1$ ones are unsuccessful), denoted by $\mathcal{P}_H(N_R)$ is defined as

$$\mathcal{P}_H(N_R) = \mathbb{P}(\hat{u} \neq u)^{N_R}. \quad (23)$$

It must be noticed that, although the process of the hard decision method is very simple, it may be actually computationally expensive, since it requires the demodulation of the broadcast signal in order to verify if the cell ID has been well estimated. Moreover, it may require a large latency, since NSSS signal is transmitted every 20 ms. This will be further illustrated and discussed in Section 5.

4.2 Soft Combination of Observations with Maximum Likelihood

This method softly combines the successive observations of the NSSS signal in order to improve the estimation performance. In order to derive the best way to combine the successive observations, we consider the likelihood function L of two consecutive NSSS signals denoted by \mathbf{y}_1 and \mathbf{y}_2 , before generalizing to N_R repetitions. In this respect, L can be expressed as

$$\begin{aligned}
L &= C. \exp\left(-\frac{1}{\sigma^2} \|\mathbf{y}_1 - \alpha_1 \mathbf{X}(q, \theta_{f,1}) \mathbf{E}(u)\|^2\right) \\
&\quad \times \exp\left(-\frac{1}{\sigma^2} \|\mathbf{y}_2 - \alpha_2 \mathbf{X}(q, \theta_{f,2}) \mathbf{E}(u)\|^2\right) \\
&= C. \exp\left(-\frac{1}{\sigma^2} \left[\|\mathbf{y}_1\|^2 + \|\mathbf{y}_2\|^2 \right. \right. \\
&\quad \left. \left. + \|\alpha_1 \mathbf{X}(q, \theta_{f,1}) \mathbf{E}(u)\|^2 + \|\alpha_2 \mathbf{X}(q, \theta_{f,2}) \mathbf{E}(u)\|^2 \right] \right) \\
&\quad \times \exp\left(\frac{1}{\sigma^2} \left[\operatorname{Re}\left\{ (\alpha_1 \mathbf{X}(q, \theta_{f,1}) \mathbf{E}(u))^H \mathbf{y}_1 \right\} \right. \right. \\
&\quad \left. \left. + \operatorname{Re}\left\{ (\alpha_2 \mathbf{X}(q, \theta_{f,2}) \mathbf{E}(u))^H \mathbf{y}_2 \right\} \right] \right). \quad (24)
\end{aligned}$$

We know from (4) that the difference between $\mathbf{X}(q, \theta_{f,2})$ and $\mathbf{X}(q, \theta_{f,1})$ lies in a phase rotation, such that we can write $\mathbf{X}(q, \theta_{f,2}) = \mathbf{M}_\theta \mathbf{X}(q, \theta_{f,1})$, where \mathbf{M}_θ is a 132×132 diagonal matrix containing the elements $e^{-j\frac{\pi}{2}n}$ for any $n = 0, 1, \dots, 131$. It results that maximum likelihood estimation of (q, u) in (24) leads to

$$\begin{aligned}
(\hat{u}, \hat{q}) &= \arg \max_{u, q} L \\
&= \arg \max_{u, q} \exp\left(\frac{1}{\sigma^2} \left[\operatorname{Re}\left\{ (\mathbf{X}(q, \theta_{f,1}) \mathbf{E}(u))^H \right. \right. \right. \\
&\quad \left. \left. \times (\alpha_1^* \mathbf{y}_1 + \alpha_2^* \mathbf{M}_\theta^H \mathbf{y}_2) \right\} \right] \right). \quad (25)
\end{aligned}$$

In the following, we assume that the channel is almost static, such that we can consider $\alpha_2 \approx \alpha_1$. In that case, (25) can be rewritten similarly to (8) and (9) as

$$\begin{aligned}
(\hat{u}, \hat{q}) &= \arg \max_{u, q} \operatorname{Re}\left\{ \alpha_1^* (\mathbf{X}(q, \theta_{f,1}) \mathbf{E}(u))^H \right. \\
&\quad \left. \times (\mathbf{y}_1 + \mathbf{M}_\theta^H \mathbf{y}_2) \right\}, \quad (26)
\end{aligned}$$

if the channel state is known, or

$$(\hat{u}, \hat{q}) = \arg \max_{u, q} \left| (\mathbf{X}(q, \theta_{f,1}) \mathbf{E}(u))^H (\mathbf{y}_1 + \mathbf{M}_\theta^H \mathbf{y}_2) \right|, \quad (27)$$

if the channel state is unknown. In both cases, it can be deduced from (26) and (27) that a new equivalent observation $\mathbf{y}_1 + \mathbf{M}_\theta^H \mathbf{y}_2$ can be used in cell ID ML estimator, and it is the best way to combine \mathbf{y}_1 and \mathbf{y}_2 . Then, the same new observation can be used to feed the suggested sub-optimal algorithm as well. Thus, we add a "step 0" to the method, which is to define a new observation $\tilde{\mathbf{y}}(2)$ as

$$\tilde{\mathbf{y}}(2) = \frac{\mathbf{y}_1 + \mathbf{M}_\theta^H \mathbf{y}_2}{2}, \quad (28)$$

and the suggested low-complexity algorithm (12)-(17) is performed using $\tilde{\mathbf{y}}(2)$ instead of \mathbf{y} . The same principle can be generalized for any N_R repetitions of the NSSS signal with

$$\tilde{\mathbf{y}}(N_R) = \frac{1}{N_R} \sum_{i=1}^{N_R} (\mathbf{M}_\theta^{i-1})^H \mathbf{y}_i. \quad (29)$$

The performance of the soft combination method can be compared with the error probability (22). To this end, we compare the signal to noise ratio (SNR) corresponding to \mathbf{y} denoted by SNR , and that of the equivalent observation $\tilde{\mathbf{y}}(N_R)$ denoted

by $SNR(N_R)$. Thus, the SNR corresponding to \mathbf{y} can be expressed as

$$SNR = \frac{\mathbb{E}\{\|\mathbf{y}\|^2\}}{\sigma^2} - 1 = \frac{\mathbb{E}\{\|\alpha \mathbf{d}\|^2\}}{\sigma^2}. \quad (30)$$

Since the different additive zero-mean noise components \mathbf{w}_i in \mathbf{y}_i are independent, the equivalent noise variance for $\tilde{\mathbf{y}}(N_R)$ is given by $\tilde{\sigma}^2(N_R)$

$$\begin{aligned}
\tilde{\sigma}^2(N_R) &= \mathbb{E}\left\{ \left\| \frac{1}{N_R} \sum_{i=1}^{N_R} (\mathbf{M}_\theta^{i-1})^H \mathbf{w}_i \right\|^2 \right\} \\
&= \frac{1}{N_R^2} \sum_{i=1}^{N_R} \mathbb{E}\{\|\mathbf{w}_i\|^2\} = \frac{\sigma^2}{N_R}. \quad (31)
\end{aligned}$$

We then deduce that the equivalent SNR can be expressed as

$$SNR(N_R) = N_R \cdot SNR, \quad (32)$$

or equivalently in dB: $SNR_{dB}(N_R) = SNR_{dB} + 10 \log(N_R)$. This result highlights that the equivalent SNR of the combined NSSS signals at the N_R -th repetition is $10 \log(N_R)$ higher than the original SNR at the first repetition, which should therefore improve the estimation performance. We denote by $\mathcal{P}_{ML}(N_R)$ the error probability corresponding to the N_R -th repetition. Based on (22), it can be expressed as

$$\mathcal{P}_{ML}(N_R) = \mathbb{P}(\hat{u} \neq u, N_R), \quad (33)$$

where $\mathbb{P}(\hat{u} \neq u, N_R)$ is the shifted version of \mathcal{P} in (22) in which SNR is substituted by $SNR(N_R)$. This method is less computationally expensive than the hard decision as the cell ID needs to be estimated only after N_R repetitions, and the complexity of the combination (29) is negligible.

4.3 Hybrid Hard decision-ML Combination

The hybrid combination consists in merging the principles of both the hard decision and the ML combination techniques. Thus, similarly to the hard decision method, the UE continues estimating the cell ID from successive observations, but it combines the observations at each repetition as in (29). The expected probability of error, denoted by $\mathcal{P}_{Hy}(N_R)$, can be expressed from (23) and (33) as

$$\mathcal{P}_{Hy}(N_R) = \prod_{i=1}^{N_R} \mathbb{P}(\hat{u} \neq u, i). \quad (34)$$

Since $\mathbb{P}(\hat{u} \neq u, N_R) \leq \mathbb{P}$, we deduce from (34) that $\mathcal{P}_{Hy}(N_R)$ shall be lower than $\mathcal{P}_H(N_R)$. However, the hybrid decision is the most computationally expensive method, as it is a combination of the previous ones. The actual performance of the different methods are further discussed through simulations in next section.

5 Simulations and Discussion

5.1 Simulations Results

5.1.1 Performance Comparison for $N_R = 1$: All simulations results have been obtained using Matlab, where 10^5 independent runs per point have been performed. In Fig. 2, we compare the suggested method with the exhaustive ML cell ID estimation for $N_R = 1$, where the channel is supposed to be known as in (8) and (16), and unknown as in (9) and (17). The considered SNR is in the range $[-15, 0]$ dB. It can be observed that the ML exhaustive cell ID search

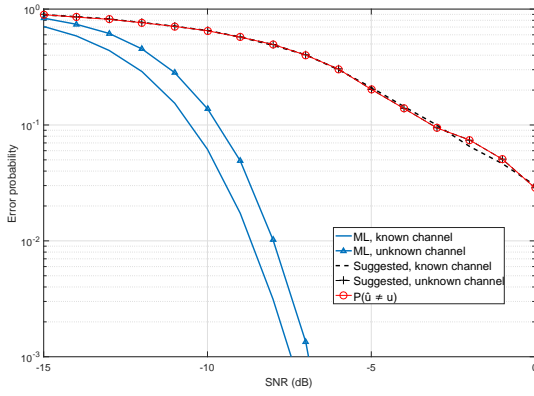


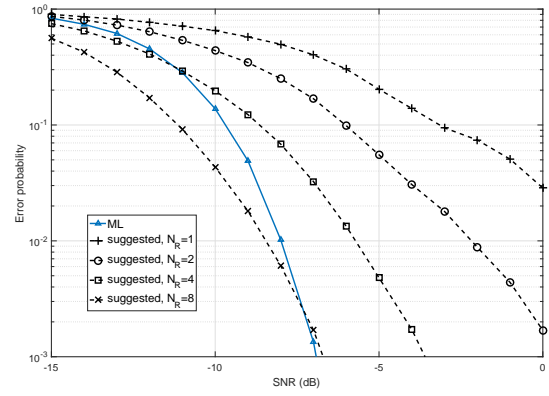
Fig. 2: Performance comparison of exhaustive ML estimation (8) (known channel) (9) (unknown channel) and suggested method (16) (known channel) (17) (unknown channel) for $N_R = 1$.

largely outperforms the suggested method. However, it must be reminded that the ML is 30 times computationally more complex than the sub-optimal technique. This is why NSSS repetitions can be taken into account in order to challenge the ML estimator performance. Moreover, Fig. 2 shows that the performance of the suggested method is the same by considering a known or unknown channel (16) or (17). Therefore in the following, we only consider the case where the channel is unknown. Finally, Fig. 2 allows us to validate the established expression of the probability of error (22). In fact, we can observe that $\mathbb{P}(\hat{u} \neq u)$ actually matches the probability of error of cell ID estimation \mathcal{P} such as expected.

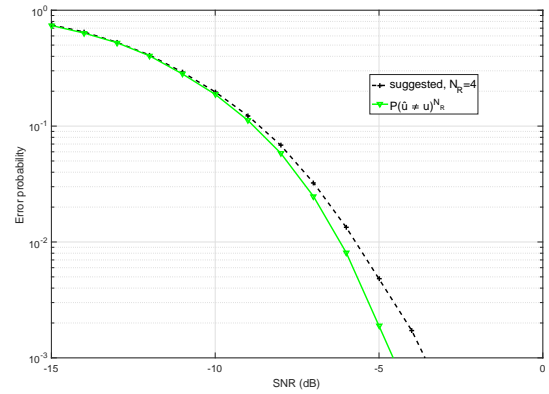
In the following, we compare the exhaustive ML search for $N_R = 1$ with the proposed method and the different combinations. In order to carry out the fairest possible comparison, we examine the performance in term of error probability relatively to the complexity of the methods. Thus, it has been stated that the ML estimator is 30 times more complex than the suggested method. It results that we can consider up to $N_R = 30$ repetitions for our method to reach the complexity of ML. We hereby limit N_R to the set $N_R \in \{1, 2, 4, 8\}$.

5.1.2 Hard Decision Method: Fig. 3 shows the performance of the hard decision technique for SNR in the range $[-15, 0]$ dB. In Fig. 3-(a), we compare the error probability for $N_R \in \{1, 2, 4, 8\}$. It is exhibited that the larger N_R , the better the performance. It can even be noted that the error probability of $\mathcal{P}_H(N_R = 8)$ is lower than that of ML in SNR range $[-15, -7]$ dB, which justifies the use of hard decision method. Furthermore, Fig. 3-(b) compares the expected error probability $\mathcal{P}_H(N_R) = \mathbb{P}(\hat{u} \neq u)^{N_R}$ and that obtained from simulations for arbitrarily chosen $N_R = 4$. It can be observed that $\mathbb{P}(\hat{u} \neq u)^{N_R}$ fits the simulation for $SNR \leq -8$ dB, but underestimates the actual error probability for larger SNR values. We deduce that (23) should be used only as a good lower bound to estimate the error probability. The reason why the theoretical expression does not match the simulations for any SNR value is discussed afterward.

5.1.3 Soft Combination Method: Other series of simulations investigate the performance of the soft combination method in Fig. 4, for $N_R \in \{1, 2, 4, 8\}$. Once again, we observe that the larger N_R , the lower the error probability. The error probability of soft combination with $N_R = 8$ is even largely lower than that of ML for $SNR \leq -8$ dB. Moreover, it should be noticed that the different behaviors for $N_R \in \{2, 4, 8\}$ are shifted versions of the trajectory corresponding to $N_R = 1$. Thus, according to theory in (33), the



(a) Performance comparison with $N_R \in \{1, 2, 4, 8\}$.



(b) Performance analysis (23) compared with simulation for $N_R = 4$.

Fig. 3: Error probability of the suggested cell ID estimation versus SNR using hard decision method.

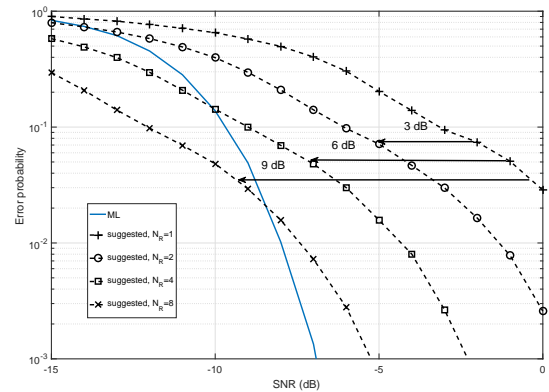


Fig. 4: Error probability of the suggested cell ID estimation versus SNR using soft combination method for $N_R \in \{1, 2, 4, 8\}$.

SNR differences between the curves of $N_R = 1$ and $N_R = 2$, 4, and 8 are of 3, 6, and 9 dB, respectively.

5.1.4 Hybrid Method: In Fig. 5, we compare the performance of the hybrid method with that of the hard decision and the soft combination techniques for $N_R = 4$. It can be seen that in low SNR range (≤ -10 dB), the three techniques outperform the ML estimation, and we have $\mathcal{P}_H(N_R) \leq \mathcal{P}_{ML}(N_R) \leq \mathcal{P}_{Hy}(N_R)$. However, for $SNR \geq -5$ dB, we

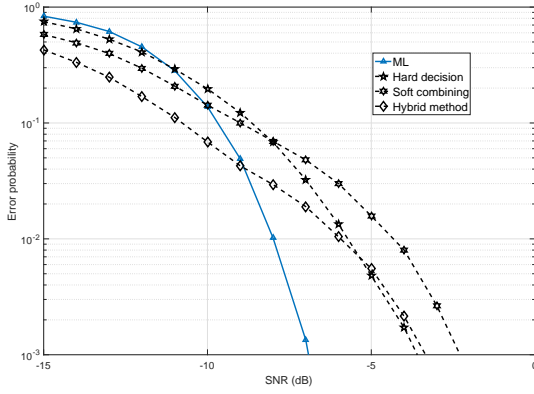


Fig. 5: Error probability of the suggested cell ID estimation versus SNR using soft hybrid method for $N_R = 4$, compared with hard decision and soft combination techniques.

observe that the error probability of the hybrid method is slightly higher than that of the hard decision, which is inconsistent with the analysis deduced from (34). This behavior is hereby discussed.

5.2 Discussion

5.2.1 Behavior Analysis: It has been observed in Figs. 3 and 5 that the expected performance in (23) and (34) match that obtained through simulations for low SNR values (typically ≤ -8 dB), but is not consistent for higher SNR values. This could be due to the assumption of independence of the event in B and C leading to (22). In fact, this holds for low SNR values, but further simulations revealed that for SNR values larger than -5 dB, we have $\mathbb{P}(\hat{u} \neq u | \hat{q} = q) = 0$, i.e. it is unlikely to wrongly estimate u if q is properly estimated. As a consequence, we deduce that

$$\begin{aligned} \mathbb{P}(B) &= \mathbb{P}(\hat{u} \neq u \cap \hat{q} = q) \\ &= \mathbb{P}(\hat{u} \neq u | \hat{q} = q) \mathbb{P}(\hat{q} = q) = 0, \end{aligned} \quad (35)$$

and hence \mathcal{P} in (20) is reduced to $\mathcal{P} = \mathbb{P}(C)$. Furthermore, since $\mathbb{P}(\hat{u} \neq u | \hat{q} = q) = 0$, the errors of cell ID estimation for higher SNR levels are only due to errors of estimation of q , then $\mathcal{P} = \mathbb{P}(\hat{q} \neq q)$.

This result indicate that the error probability of cell ID estimation should be split according to the SNR range: one corresponding to low SNR value where \mathcal{P} is mainly linked to the error of estimation of u , and one for larger SNR values where the errors of cell ID estimation are mainly due to the errors of estimation of q . Thus, under hypothesis of independence between the successive observations of NSSS signal, we could assume that $\mathcal{P}_H(N_R)$ in larger SNR range can be expressed similarly to (23), i.e. $\mathcal{P}_H(N_R) = \mathbb{P}(\hat{q} \neq q)^{N_R}$. This assumption is hereby discussed.

In Fig. 6 we show the error probability of the hard decision method, as well as $\mathbb{P}(\hat{u} \neq u)^{N_R}$ and $\mathbb{P}(\hat{q} \neq q)^{N_R}$ versus SNR for $N_R = 2$. We can observe that, in accordance with Fig. 3-(b), $\mathbb{P}(\hat{u} \neq u)^{N_R}$ fits the cell ID estimation error probability obtained through simulations in low SNR range, whereas $\mathbb{P}(\hat{q} \neq q)^{N_R}$ underestimates the error probability for any SNR value. As a consequence, we cannot estimate $\mathcal{P}_H(N_R) = \mathbb{P}(\hat{q} \neq q)^{N_R}$ such as aforementioned.

This behavior is due to the fact that $\mathbb{P}(\hat{q} \neq q)$ is not independent of the cell ID value. Thus, in presence of noise, it is more likely to misestimate q for some cell ID values than for some others. The reason is that, in absence of noise, the correlation metric in (15) corresponding to $(q', \theta'_f) = (q, \theta_f)$ is equal to

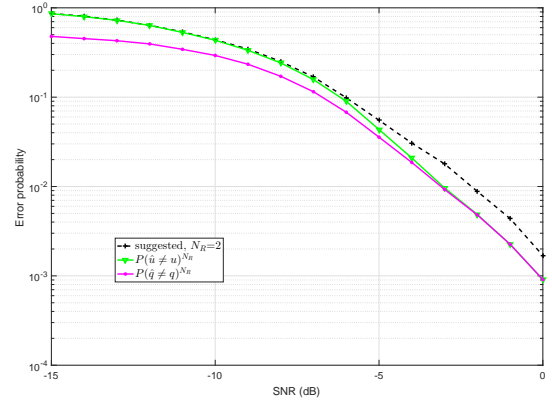


Fig. 6: Error probability of the suggested cell ID estimation versus SNR using hard decision method for $N_R = 2$, compared with $\mathbb{P}(\hat{u} \neq u)^{N_R}$ and $\mathbb{P}(\hat{q} \neq q)^{N_R}$.

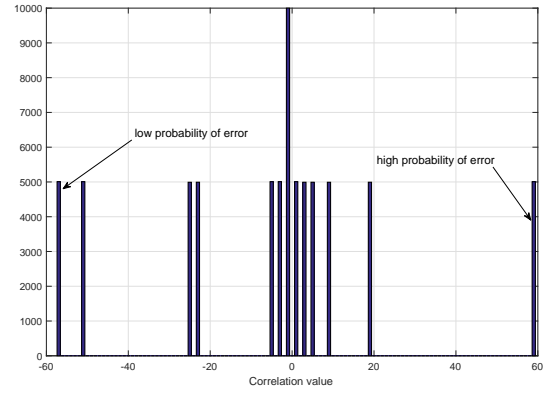


Fig. 7: Histogram of the possible auto-correlation values $\Gamma(q')$.

$$\Gamma(q) = \text{Re} \left\{ \sum_{n=0}^{64} z_{n,0}(q, \theta_f)^* z_{130-n,0}(q, \theta_f) \right\} = 65|\alpha|^2,$$

for any cell ID value. The other possible values of this correlation metric for $(q', \theta'_f) \neq (q, \theta_f)$ are depicted in the histogram in Fig. (7), which have been obtained with 10^4 simulations runs and for $\alpha = 1$. It can be observed that the Γ metric can take different values from $-57|\alpha|^2$ to $59|\alpha|^2$. The negative bins on the left side of Fig. 7 correspond to sequences $z_{n,0}(q', \theta'_f)$ that are "rather" anti-symmetric, the bins close to 0 correspond to sequences $z_{n,0}(q', \theta'_f)$ that are neither anti-symmetric nor symmetric, and the positive bins on the right side of Fig. 7 correspond to sequences $z_{n,0}(q', \theta'_f)$ that are "rather" symmetric, the exact symmetry is only obtained for $(q', \theta'_f) = (q, \theta_f)$ leading to $\Gamma(q) = 65|\alpha|^2$.

We deduce from Fig. (7) that it is likely to erroneously estimate q in presence of noise in particular if a given value of cell ID leads to $\Gamma(q' \neq q) = 59$. In that case, it is also likely to misestimate q in successive repetitions. Therefore, we conclude that there is no straightforward expression of $\mathcal{P}_H(N_R)$. This is the reason why the analysis does not match the results obtained through simulations.

5.2.2 Alternative Estimation Methods: In this section, we briefly present alternative estimation techniques to those presented in Sections 3 and 4. First, the mean square error (MSE)

estimator could be used instead of the auto-correlation (15) to estimate q , which is expressed as

$$(\hat{q}, \hat{\theta}_f) = \arg \min_{q', \theta'_f} \sum_{n=0}^{64} |z_n(q', \theta'_f) - z_{130-n}(q', \theta'_f)|^2. \quad (36)$$

Simulations (not shown in this paper) revealed that the performance of the MSE estimator is the same as the auto-correlation one. Furthermore, it requires the same number of complex multiplications.

Then, it must be stated that the methods presented in Section 4 to take advantage of the repetitions of the NSSS signal are not the only ones that could be used to improve the cell ID estimation using the repetitions. From the previous results, we can suggest two leads that pave the way for further works:

- In order to improve the good estimation of q , it can be suggested to consider the arguments of the two maximal values in (15), if the two maximal values are close, *i.e.* when it is likely to estimate the wrong q value. However, this strategy doubles the complexity of the method, since u should be estimated for the two estimated q values.
- Another alternative is to choose the best combination method according to the SNR (which should be estimated). Thus, the soft combination or hybrid method should be considered in low SNR values, whereas the hard decision method should be chosen in higher SNR ranges.

6 Conclusion

In this paper, we have presented a sub-optimal low-complexity cell ID estimator for NB-IoT, which can be seen as an alternative to the ML exhaustive search. Based on the auto-correlation metric instead of the cross-correlation, it is up to thirty times less complex than the ML. Moreover, we have introduced different methods to take advantage of the NSSS signal repetitions. The first one is based on the hard decision after the estimation step, the second one softly combines the observations before the estimation step, and the third one is an hybrid technique that merges the principles of the two firsts. A performance analysis of the different methods has been suggested, which is further discussed through simulations results. We conclude from the results that the sub-optimal method with combinations of the repetitions can at least reach the performance of the exhaustive ML search for a lower overall complexity. Furthermore, some alternative techniques have been mentioned, paving the way for further works.

7 References

- 1 Ericsson, "Internet of Things forecast," 2016, accessed July 2019. [Online]. Available: <https://www.ericsson.com/en/mobility-report/internet-of-things-forecast>
- 2 3GPP, "3GPP TS 36.211, Physical channels and modulation (Release 14)," 3GPP, Tech. Rep., March 2017.
- 3 R. Ratasuk, N. Mangalvedhe, Y. Zhang, and J.-P. Robert, M. Koskinen, "Overview of Narrowband IoT in LTE Rel-13," in *proc. of CSCN'16*, Berlin, Germany, November 2016, pp. 1 – 7.
- 4 J. Schliez and D. Raddino, "Narrowband Internet of Things - Whitepaper," Rohde & Schwarz, Tech. Rep., 2016.
- 5 Y.-P. E. Wang, X. Lin, A. Adhikary, A. Gröven, Y. Sui, J. Blankenship, Y. Bergman, and H. S. Razaghi, "A Primer on 3GPP Narrowband Internet of Things," *IEEE Communications Magazine*, vol. 55, no. 3, pp. 117 – 123, March 2017.
- 6 H. Kröll, M. Korb, B. Weber, S. Willi, and Q. Huang, "Maximum-Likelihood Detection for Energy-Efficient Timing Acquisition in NB-IoT," in *proc. of WCNCW'17*. San Francisco, CA, March 2017.
- 7 K. Manolakis, D. M. Gutiérrez-Estévez, V. Jungnickel, W. Xu, and C. Drewes, "A Closed Concept for Synchronization and Cell Search in 3GPP LTE Systems," in *proc. of WCNC'09*, Budapest, Hungary, April 2009.
- 8 Qualcomm, "NB-PSS and NB-SSS Design," 3GPP TSG RAN WG1, Sophia-Antipolis, France, Tech. Rep. Tech. Rep. R1-161981, March 2016.
- 9 A. Bhamri, L. Zexian, L. Lindh, and C. Ribeiro, "Primary Synchronization Signal Detection Method for Device-to-Device in LTE-Rel 12 and Beyond," in *proc. of VTC'2015-Fall*, Bosto, MA, USA, September 2015.
- 10 Intel Corporation, "Synchronization and cell search in nb-iot: Performance evaluations," 3GPP TSG RAN WG1, Sophia-Antipolis, France, Tech. Rep. Tech. Rep. R1-161898, April 2016.
- 11 A. Ali and W. Hamouda, "On the Cell Search and Initial Synchronization for NB-IoT LTE Systems," *IEEE Communications Letters*, vol. 21, no. 8, pp. 1843 – 1846, August 2017.
- 12 Y. Li, B., S. Chen, W. Ye, and F. Lin, "A Joint Low-Power Cell Search and Frequency Tracking Scheme in NB-IoT Systems for Green Internet of Things," *Sensors*, vol. 18, no. 10, p. 22, September 2018.
- 13 R. Frank, "Polyphase codes with good nonperiodic correlation properties," *IEEE Transactions on Information Theory*, vol. 9, no. 1, pp. 43 – 45, January 1963.
- 14 D. Chu, "Polyphase codes with good periodic correlation properties," *IEEE Transactions on Information Theory*, vol. 18, no. 4, pp. 531 – 532, July 1972.