Accurate Estimation of the Minimum Primary Channel Activity Time in Cognitive Radio Based on Periodic Spectrum Sensing Observations

Miguel López-Benítez*, Ahmed Al-Tahmeesschi*, Dhaval Patel†
*Department of Electrical Engineering and Electronics, University of Liverpool, United Kingdom
†School of Engineering and Applied Science, Ahmedabad University, India
Email: M.Lopez-Benitez@liverpool.ac.uk

***Abstract*—The minimum duration of idle/busy periods in a primary channel is an important parameter in Cognitive Radio (CR) systems since it determines the minimum amount of time the primary channel will be available for opportunistic transmission (in the case of idle periods) or the minimum amount of time the CR system will have to wait before a busy primary channel can become available again for opportunistic transmission (in the case of busy periods). Recent research has demonstrated that the value of the minimum period plays a critical role in the estimation of primary channel activity statistics, which can be estimated accurately only if the minimum period duration is accurately known. Given the importance of this parameter, this work analyses the problem of estimating the minimum (idle/busy) period in a primary channel based on periodic spectrum sensing observations and proposes novel methods to enable an accurate estimation. The obtained simulation results demonstrate that the proposed methods can provide an accurate estimation of the minimum period regardless of the employed sensing period (i.e., the main aspect that constrains the available time resolution).**

I. Introduction

Given the opportunistic nature of the Dynamic Spectrum Access/Cognitive Radio (DSA/CR) paradigm [1], the behaviour and performance of DSA/CR systems depends on the spectrum occupancy pattern of primary systems. DSA/CR systems therefore need to gain an accurate knowledge of the activity patterns of primary systems to exploit the available spectrum opportunities more efficiently. Statistical information, such as the duration of idle/busy periods of the channel and the underlying distribution, can be exploited by DSA/CR systems in the prediction of future spectrum occupancy trends [2], selection of the most convenient licensed frequency band and radio channel [3], as well as spectrum and radio resource management decisions to minimise interference, optimise system performance and improve spectrum efficiency [4].

DSA/CR systems can estimate the activity statistics of a primary channel (such as the distribution of idle/busy periods) based on spectrum sensing decisions. While the main purpose of spectrum sensing is to detect transmission opportunities, the sequence of binary on/off sensing decisions can also be used to estimate the durations of individual idle/busy periods resulting from primary transmissions. Once a sufficiently large number of individual idle/busy periods have been observed, the DSA/CR system can estimate relevant activity statistics. One of the most commonly sought types of statistical information is the distribution of the durations of idle/busy periods as this provides a complete characterisation of their statistical properties. However, recent research [5] has demonstrated that an accurate estimation of the primary activity statistics, in particular the distribution of idle/busy periods, is possible only if the minimum idle/busy period is know very accurately. The estimated statistical distribution resulting from spectrum sensing observations is in general highly inaccurate (to the point that it could be useless) unless the minimum period duration is accurately known (the reader is referred to [5] for a detailed analysis and discussion on this topic).

The minimum period can be easily estimated from spectrum sensing observations but such estimation is in general inaccurate because the periodicity of spectrum sensing observations (i.e., the sensing period) imposes a fundamental limit on the time resolution to which individual idle/busy periods (including the minimum period) can be observed. This limitation could obviously be overcome by reducing the sensing period, however this may not be possible in practice since the sensing period is typically selected to achieve a certain signal detection performance under specified hardware constraints rather than an accurate estimation of the primary activity statistics. To overcome this limitation, it is necessary to devise methods that can provide an accurate estimation of the minimum primary idle/busy period regardless of the spectrum sensing periodicity. This apparently trivial problem has however important practical implications in the context of DSA/CR systems and, to the best of the authors' knowledge, has not received a rigorous and formal treatment in the literature. To fill this gap, this work proposes several methods to enable and accurate estimation of the minimum activity/inactivity times of a primary channel based on periodic spectrum sensing observations.

The rest of this work is organised as follows. First, Section II provides a formal description of the problem under study and introduces a suitable system model. The methods proposed for the estimation of the minimum period are presented in Section III and their sample size requirements are analytically studied in Section IV. Section V assesses the performance of the proposed methods with numerical and simulation results. Finally, Section VI summarises and concludes the work.

II. Problem Formulation and System Model

Fig. 1 illustrates the estimation of the duration of an idle period based on spectrum sensing (the same principle applies to the estimation of busy periods). The DSA/CR system senses the primary channel periodically with a sensing period of T_s time units (t.u.). Based on the sequence of idle (H_0) and busy (H_1) decisions provided by the employed spectrum sensing method, the DSA/CR system can make an estimation $\widehat{T}_i$ of the true idle/busy periods T_i ($i = 0$ for idle periods, $i = 1$ for busy periods). Note that the estimated periods are integer multiples of the employed sensing period (i.e., $\widehat{T}_i = kT_s$, $k \in \mathbb{N}^+$) and as a result the estimated periods will differ from the true periods, which can in general be assumed to have a continuous domain (i.e., $T_i \in \mathbb{R}^+$). Consequently, the minimum estimated period, denoted as $\widehat{\mu}_i = \min(\widehat{T}_i)$, will in general differ from the true minimum period, denoted as $\mu_i = \min(T_i)$. An obvious solution to this problem would be to reduce the sensing period T_s, which would improve the time resolution to which the periods can be observed, until a sufficiently low sensing period is set such that $\widehat{\mu}_i \approx \mu_i$. However, as discussed in Section I, this solution may not be feasible in most practical cases. In this context, the problem addressed in this paper is how to obtain an accurate estimation of the true minimum value of the idle/busy periods (i.e., μ_i) regardless of the employed sensing period T_s.

To analyse the proposed methods it will be useful first to introduce a model for the estimated periods, $\widehat{T}_i$, as a function of the true periods, T_i, and the employed sensing period, T_s. The periods estimated as shown in Fig. 1 can be expressed as:

$$\widehat{T}_i = \left(\left\lfloor \frac{T_i}{T_s} \right\rfloor + \xi \right) T_s \quad (1)$$

where $\lfloor \cdot \rfloor$ denotes the floor operator and $\xi \in \{0, 1\}$ is a Bernoulli random variable introduced to model the fact that the same original period T_i can lead to two possible estimated periods, either $\widehat{T}_i = \lfloor T_i/T_s \rfloor T_s$ or $\widehat{T}_i = \lceil T_i/T_s \rceil T_s = (\lfloor T_i/T_s \rfloor + 1)T_s$ (where $\lceil \cdot \rceil$ denotes the ceil operator), depending on the relative (random) position of the sensing events with respect to the original period T_i. Notice that the true value of the original period T_i is within the interval $[\lfloor T_i/T_s \rfloor T_s, \lceil T_i/T_s \rceil T_s]$, and the estimation $\widehat{T}_i = \lfloor T_i/T_s \rfloor T_s$ (i.e., $\xi = 0$) will be more likely when T_i is closer to the lower extreme of the interval, while the estimation $\widehat{T}_i = \lceil T_i/T_s \rceil T_s$ (i.e., $\xi = 1$) will be more likely when T_i is closer to the upper extreme of the interval. Taking this into account, the Bernoulli probabilities of the model in (1) can be calculated:

$$P(\xi = 0) = P\left(\widehat{T}_i = \left\lfloor \frac{T_i}{T_s} \right\rfloor T_s\right) = \left\lceil \frac{T_i}{T_s} \right\rceil - \frac{T_i}{T_s} \quad (2a)$$

$$P(\xi = 1) = P\left(\widehat{T}_i = \left\lceil \frac{T_i}{T_s} \right\rceil T_s\right) = \frac{T_i}{T_s} - \left\lfloor \frac{T_i}{T_s} \right\rfloor \quad (2b)$$

Notice that the maximum absolute estimation error will occur when T_i is close to one of the extremes of the interval

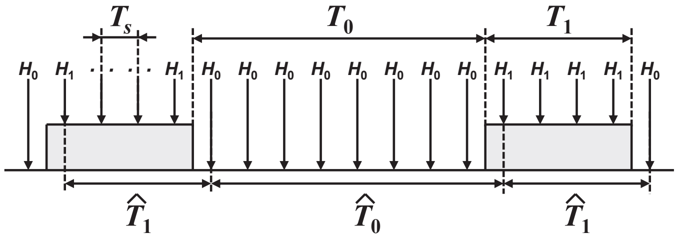

Fig. 1. Estimation of the duration of an idle period based on spectrum sensing.

$[\lfloor T_i/T_s \rfloor T_s, \lceil T_i/T_s \rceil T_s]$ and $\widehat{T}_i$ is the other extreme, hence:

$$\max(|T_i - \widehat{T}_i|) = \left| \left\lfloor \frac{T_i}{T_s} \right\rfloor T_s - \left\lceil \frac{T_i}{T_s} \right\rceil T_s \right| = T_s \quad (3)$$

which shows that the error can be reduced by reducing T_s.

III. Proposed Estimation Methods

Let $\{\widehat{T}_{i,n}\}_{n=1}^N$ be a set of N observed periods $\widehat{T}_i$ of type i ($i = 0$ for idle periods, $i = 1$ for busy periods) estimated as shown in Fig. 1. The DSA/CR system can calculate primary activity statistics based on such set. This section proposes and discusses three different methods to estimate the minimum period from the above mentioned set of observed periods.

A. Direct Estimation Method (DEM)

This method provides an estimation of the minimum period by selecting the minimum value of the set of observed periods:

$$\widehat{\mu}_i^{\text{DEM}} = \min_n \left(\{\widehat{T}_{i,n}\}_{n=1}^N \right) \quad (4)$$

The relation between the minimum period estimated this way and the true minimum period as a function of the employed sensing period can be obtained from (1) by noting that $\min(T_i) = \mu_i$ and $\min(\xi) = 0$:

$$\widehat{\mu}_i^{\text{DEM}} = \min(\widehat{T}_i) = \min\left[\left(\left\lfloor \frac{T_i}{T_s} \right\rfloor + \xi\right) T_s\right] = \left\lfloor \frac{\mu_i}{T_s} \right\rfloor T_s \quad (5)$$

As it can be noted from (5), the estimated minimum is accurate if the employed sensing period is an integer submultiple of the true minimum (i.e., $\widehat{\mu}_i^{\text{DEM}} = \mu_i$ if $T_s = \mu_i/k$ with $k \in \mathbb{N}^+$). However, the estimation provided by this method can in general differ significantly from the true minimum μ_i. The maximum absolute error in (3) for individual periods is also applicable to the DEM-estimated minimum in (5).

It also is worth noting from (5) that the true minimum could be found by means of an exhaustive search. Since $\widehat{\mu}_i^{\text{DEM}} = \mu_i$ when $T_s = \mu_i/k$ ($k \in \mathbb{N}^+$) and $\widehat{\mu}_i^{\text{DEM}} < \mu_i$ otherwise, the true minimum period μ_i could be found by increasing/decreasing progressively the employed T_s and observing the resulting $\widehat{\mu}_i^{\text{DEM}}$ until a local maximum is found by trial-and-error; such maximum would be an accurate estimation of the true minimum μ_i. While this method would eventually find the sought value, it would require testing a large number of values for T_s, which would not be practical in real DSA/CR systems. Therefore, this approach is not further considered here.

B. Curve-Fitting Method (CFM)

Since the floor function is not invertible, the value of the true minimum μ_i cannot be uniquely identified based on a single pair of employed (known) T_s and observed $\widehat{\mu}_i^{\text{DEM}}$. However, by testing a set $\{T_{s,l}\}_{l=1}^{L}$ of L values of the sensing period and noting the set $\{\widehat{\mu}_{i,l}^{\text{DEM}}\}_{l=1}^{L}$ of minimum periods observed for each sensing period (i.e., $\widehat{\mu}_{i,l}^{\text{DEM}} = \lfloor \mu_i / T_{s,l} \rfloor T_{s,l}$), curve-fitting can be used to find the value of μ_i that provides the best fit of (5) to the observed sets $\{T_{s,l}\}_{l=1}^{L}$ and $\{\widehat{\mu}_{i,l}^{\text{DEM}}\}_{l=1}^{L}$.

Since curve-fitting methods are typically based on gradients, the discontinuities of the floor function in (5) can be problematic. This can be solved by replacing the floor function in (5) with its equivalent (continuous) Fourier series [6, eq. (2.1.7)]:

$$\widehat{\mu}_i^{\text{DEM}} = \mu_i - \frac{T_s}{2} + T_s \sum_{m=1}^{\infty} \frac{\sin\left(\frac{2\pi m \mu_i}{T_s}\right)}{\pi m} \tag{6}$$

The expression in (6) can be easily fitted to the observed sets $\{T_{s,l}\}_{l=1}^{L}$ and $\{\widehat{\mu}_{i,l}^{\text{DEM}}\}_{l=1}^{L}$ using conventional curve-fitting methods. The value of μ_i in (6) that provides the best fit represents the CFM-estimated minimum, denoted as $\widehat{\mu}_i^{\text{CFM}}$.

While this method requires testing several values of the sensing period, this does not necessarily require the DSA/CR system to actually modify the employed sensing period, which in some cases might not be practical. Different sensing periods can be emulated by selectively discarding some sensing decisions (e.g., discarding every other sensing decision would be equivalent to employing a sensing period $2T_s$).

C. Method of Moments (MoM)

The DEM and CFM approaches do not make any assumptions on the distribution of the true periods T_i. If a certain distribution model is assumed, then the minimum period can be estimated based on the sample estimates of the distribution moments. This section illustrates this approach assuming that the primary activity periods T_i are exponentially distributed, which is a common assumption in DSA/CR (the method can be easily adapted to any other distributions).

The Probability Density Function (PDF), $f_{T_i}(T)$, and the Cumulative Distribution Function (CDF), $F_{T_i}(T)$, of exponentially distributed periods T_i are respectively given by:

$$f_{T_i}(T) = \begin{cases} 0, & T < \mu_i \\ \lambda_i e^{-\lambda_i (T-\mu_i)}, & T \geq \mu_i \end{cases} \tag{7}$$

$$F_{T_i}(T) = \begin{cases} 0, & T < \mu_i \\ 1 - e^{-\lambda_i (T-\mu_i)}, & T \geq \mu_i \end{cases} \tag{8}$$

where $\mu_i > 0$ (the minimum period) is the distribution location parameter and $\lambda_i > 0$ is the distribution scale parameter.

The mean $\mathbb{E}(T_i)$ and variance $\mathbb{V}(T_i)$ of the distribution are:

$$\mathbb{E}(T_i) = \mu_i + 1/\lambda_i \tag{9}$$

$$\mathbb{V}(T_i) = 1/\lambda_i^2 \tag{10}$$

and the minimum period can therefore be expressed as:

$$\mu_i = \mathbb{E}(T_i) - \sqrt{\mathbb{V}(T_i)} \tag{11}$$

The mean and variance can be estimated based on the sample mean $\widetilde{m}_i$ and (unbiased) sample variance $\widetilde{v}_i$ of the set of observed periods $\{\widehat{T}_{i,n}\}_{n=1}^{N}$ as follows[1]:

$$\widetilde{m}_i = \frac{1}{N} \sum_{n=1}^{N} \widehat{T}_{i,n} \tag{12}$$

$$\widetilde{v}_i = \frac{1}{N-1} \sum_{n=1}^{N} \left(\widehat{T}_{i,n} - \widetilde{m}_i \right)^2 - \frac{T_s^2}{6} \tag{13}$$

and the minimum period can thus be estimated as:

$$\widehat{\mu}_i^{\text{MoM}} = \widetilde{m}_i - \sqrt{\widetilde{v}_i} \tag{14}$$

which represents the MoM-estimated minimum period.

IV. Sample Size Analysis

An important aspect of the proposed methods is how many periods need to be observed (the required sample size N) to provide an estimation of the minimum period. This determines how fast each method can provide an estimation as well as the associated computational cost. This section analyses the sample size requirements for each of the proposed methods.

A. DEM Sample Size

The DEM approach needs at least one instance of $\widehat{\mu}_i^{\text{DEM}} = \lfloor \mu_i / T_s \rfloor T_s$ to be present in the N observed periods. A period with duration $\widehat{T}_i = \lfloor \mu_i / T_s \rfloor T_s$ is observed whenever $T_i \in [\lfloor \mu_i / T_s \rfloor T_s, \lceil \mu_i / T_s \rceil T_s]$ and $\xi = 0$, hence:

$$\begin{aligned} P\left(\widehat{T}_i = \widehat{\mu}_i^{\text{DEM}}\right) &= P\left(\left\lfloor \frac{\mu_i}{T_s} \right\rfloor T_s \leq T_i \leq \left\lceil \frac{\mu_i}{T_s} \right\rceil T_s\right) \mathbb{E}\left(P(\xi = 0)\right) \\ &= P\left(\mu_i \leq T_i \leq \left\lceil \frac{\mu_i}{T_s} \right\rceil T_s\right) \mathbb{E}\left(P(\xi = 0)\right) \chi_0 \\ &= F_{T_i}\left(\left\lceil \frac{\mu_i}{T_s} \right\rceil T_s\right) \mathbb{E}\left(P(\xi = 0)\right) \chi_0 \end{aligned} \tag{15}$$

where $F_{T_i}(\cdot)$ is the CDF of the true periods T_i,

$$\begin{aligned} \mathbb{E}\left(P(\xi = 0)\right) &= \int_T P(\xi = 0)\, f_{T_i}(T)\, dT \\ &= \sum_{m=0}^{\infty} (m+1) \left[F_{T_i}((m+1)T_s) - F_{T_i}(mT_s) \right] - \frac{\mathbb{E}(T_i)}{T_s} \end{aligned} \tag{16}$$

and χ_0 is a correction factor for $P(\xi = 0)$ given by:

$$\chi_0 = \frac{\left\lceil \frac{\mu_i}{T_s} \right\rceil T_s - \mu_i}{T_s} = \left\lceil \frac{\mu_i}{T_s} \right\rceil - \frac{\mu_i}{T_s} \tag{17}$$

Notice that for any arbitrary period T_i the width of the interval $[\lfloor T_i / T_s \rfloor T_s, \lceil T_i / T_s \rceil T_s]$ is T_s. However, around the minimum period μ_i it holds that $T_i \in [\mu_i, \lceil \mu_i / T_s \rceil T_s]$ (since $T_i \geq \mu_i \geq \lfloor \mu_i / T_s \rfloor T_s$) and the width of such interval is $\lceil \mu_i / T_s \rceil T_s - \mu_i$ instead of T_s. Therefore, the probability $P(\xi = 0)$ needs to be scaled accordingly as shown in (17).

[1] The term $-T_s^2/6$ in (13) is needed to correct the effect of the employed sensing period on the sample variance (see [5] for details).

The probability to have at least one instance of $\widehat{\mu}_i^{\text{DEM}}$ in the N observed periods can be obtained from the binomial distribution and is given by:

$$P_{\text{obs}}^{\widehat{\mu}_i^{\text{DEM}}} = 1 - \left[1 - P\left(\widehat{T}_i = \widehat{\mu}_i^{\text{DEM}}\right)\right]^N \tag{18}$$

Given a specified $P_{\text{obs}}^{\widehat{\mu}_i^{\text{DEM}}}$, the required DEM sample size is:

$$N_{\text{DEM}} \geq \frac{\log\left(1 - P_{\text{obs}}^{\widehat{\mu}_i^{\text{DEM}}}\right)}{\log\left(1 - P\left(\widehat{T}_i = \widehat{\mu}_i^{\text{DEM}}\right)\right)} \tag{19}$$

Notice that increasing the sample size will not improve the accuracy itself of the DEM-estimated minimum $\widehat{\mu}_i^{\text{DEM}}$ but the probability that a period $\widehat{\mu}_i^{\text{DEM}}$ is observed (otherwise a longer period $\widehat{T}_i > \widehat{\mu}_i^{\text{DEM}}$ might be selected as the minimum period, thus potentially leading to a more inaccurate estimation).

B. CFM Sample Size

Based on (19), the CFM sample size can be expressed as:

$$N_{\text{CFM}} = \sum_{l=1}^{L} N_{\text{DEM},l} \geq \sum_{l=1}^{L} \frac{\log\left(1 - P_{\text{obs}}^{\widehat{\mu}_i^{\text{DEM}}}\right)}{\log\left(1 - P\left(\widehat{T}_i = \widehat{\mu}_{i,l}^{\text{DEM}}\right)\right)} \tag{20}$$

where $N_{\text{DEM},l}$ is the sample size required for the l-th element of the set $\{\widehat{\mu}_{i,l}^{\text{DEM}}\}_{l=1}^{L}$. If the required sample size is similar for each of the L tested sensing periods, then $N_{\text{CFM}} \approx L \cdot N_{\text{DEM}}$.

C. MoM Sample Size

As opposed to DEM and CFM, the accuracy of the MoM-estimated minimum can be improved by increasing the sample size as a larger N will lead to more accurate sample moments and consequently more accurately estimated minimum period.

A confidence interval of κ standard deviations around the expected value of the estimator $\widehat{\mu}_i^{\text{MoM}}$ in (14) can be defined such that the estimated values are within that interval with a minimum probability ρ (confidence level), i.e., $P(|\widehat{\mu}_i^{\text{MoM}} - \mathbb{E}(\widehat{\mu}_i^{\text{MoM}})| \leq \kappa\sqrt{\mathbb{V}(\widehat{\mu}_i^{\text{MoM}})}) \geq \rho$. Hence, the maximum absolute error of the (unbiased) estimator can be expressed as:

$$\begin{aligned} \varepsilon_{\text{abs}}^{\widehat{\mu}_i^{\text{MoM}}} &= \max\left(\left|\widehat{\mu}_i^{\text{MoM}} - \mathbb{E}(\widehat{\mu}_i^{\text{MoM}})\right|\right) = \max\left(\left|\widehat{\mu}_i^{\text{MoM}} - \mu_i\right|\right) \\ &= \kappa\sqrt{\mathbb{V}\left(\widehat{\mu}_i^{\text{MoM}}\right)} \end{aligned} \tag{21}$$

The expression for κ in (21) can be determined by noting that $\widehat{\mu}_i^{\text{MoM}}$ relies on the sample estimators (12) and (13), which are both unbiased since $\mathbb{E}(\widetilde{m}_i) = \mathbb{E}(T_i)$ and $\mathbb{E}(\widetilde{v}_i) = \mathbb{V}(T_i)$, and can be both assumed to be normally distributed by the central limit theorem. Assuming that $\widehat{\mu}_i^{\text{MoM}}$ is also normally distributed, then the expression $\kappa = \sqrt{2}\,\text{erf}^{-1}(\rho)$ is obtained.

The expression for $\mathbb{V}(\widehat{\mu}_i^{\text{MoM}})$ in (21) can be obtained by computing first the variances of the sample estimators (12) and (13), which are obtained by introducing $\mathbb{V}(\widehat{T}_i) = \mathbb{V}(T_i) + T_s^2/6$ [5, eq. (6)] in [7, eqs. VI.(8) and VI.(10)]:

$$\mathbb{V}(\widetilde{m}_i) = \tfrac{1}{N}\left(\mathbb{V}(T_i) + \tfrac{T_s^2}{6}\right) \tag{22a}$$

$$\begin{aligned} \mathbb{V}(\widetilde{v}_i) = \tfrac{1}{N}\Big(&\mathbb{M}_4(T_i) - \tfrac{N-3}{N-1}[\mathbb{V}(T_i)]^2 + \\ &+ \tfrac{2N}{3(N-1)}T_s^2\mathbb{V}(T_i) + \tfrac{7N+3}{180(N-1)}T_s^4\Big) \end{aligned} \tag{22b}$$

where $\mathbb{M}_4(T_i) = 9/\lambda_i^4$ is the fourth central moment of T_i, and then propagating the errors in (22) through (14):

$$\begin{aligned} \mathbb{V}\left(\widehat{\mu}_i^{\text{MoM}}\right) &= \left(\frac{\partial\widehat{\mu}_i^{\text{MoM}}}{\partial\widetilde{m}_i}\right)^2\mathbb{V}(\widetilde{m}_i) + \left(\frac{\partial\widehat{\mu}_i^{\text{MoM}}}{\partial\widetilde{v}_i}\right)^2\mathbb{V}(\widetilde{v}_i) + \\ &+ 2\left(\frac{\partial\widehat{\mu}_i^{\text{MoM}}}{\partial\widetilde{m}_i}\right)\left(\frac{\partial\widehat{\mu}_i^{\text{MoM}}}{\partial\widetilde{v}_i}\right)\mathbb{C}(\widetilde{m}_i, \widetilde{v}_i) \\ &= \mathbb{V}(\widetilde{m}_i) + \frac{\mathbb{V}(\widetilde{v}_i)}{4\,\mathbb{V}(T_i)} - \frac{1}{N}\frac{\mathbb{M}_3(T_i)}{\sqrt{\mathbb{V}(T_i)}} \end{aligned} \tag{23}$$

where the covariance is given by $\mathbb{C}(\widetilde{m}_i, \widetilde{v}_i) = \mathbb{M}_3(T_i)/N$ [8] and $\mathbb{M}_3(T_i) = 2/\lambda_i^3$ is the third central moment of T_i.

Introducing (23) and $\kappa = \sqrt{2}\,\text{erf}^{-1}(\rho)$ in (21) yields:

$$\begin{aligned} \varepsilon_{\text{abs}}^{\widehat{\mu}_i^{\text{MoM}}} = \text{erf}^{-1}(\rho)\Big[\tfrac{2}{N}\Big(&\tfrac{3N-1}{4(N-1)}\mathbb{V}(T_i) + \tfrac{\mathbb{M}_4(T_i)}{4\,\mathbb{V}(T_i)} + \\ &+ \tfrac{7N+3}{720(N-1)}\tfrac{T_s^4}{\mathbb{V}(T_i)} - \tfrac{\mathbb{M}_3(T_i)}{\sqrt{\mathbb{V}(T_i)}} + \tfrac{2N-1}{N-1}\tfrac{T_s^2}{6}\Big)\Big]^{\frac{1}{2}} \end{aligned} \tag{24}$$

For large N ($N \gg 1$) the terms inside the parenthesis of (24) can be simplified and a solution for N can be obtained:

$$\begin{aligned} N_{\text{MoM}} \geq 2&\left(\frac{\text{erf}^{-1}(\rho)}{\varepsilon_{\text{abs}}^{\widehat{\mu}_i^{\text{MoM}}}}\right)^2 \cdot \\ &\cdot\left(\tfrac{3}{4}\mathbb{V}(T_i) + \tfrac{\mathbb{M}_4(T_i)}{4\mathbb{V}(T_i)} + \tfrac{7T_s^4}{720\mathbb{V}(T_i)} - \tfrac{\mathbb{M}_3(T_i)}{\sqrt{\mathbb{V}(T_i)}} + \tfrac{T_s^2}{3}\right) \end{aligned} \tag{25}$$

Replacing the moments of the exponential distribution in (25):

$$N_{\text{MoM}} \geq 2\left(\frac{\text{erf}^{-1}(\rho)}{\varepsilon_{\text{abs}}^{\widehat{\mu}_i^{\text{MoM}}}}\right)^2\left(\frac{1}{\lambda_i^2} + \frac{7\lambda_i^2 T_s^4}{720} + \frac{T_s^2}{3}\right) \tag{26}$$

N_{MoM} in (26) is the minimum sample size that guarantees a maximum absolute error $\varepsilon_{\text{abs}}^{\widehat{\mu}_i^{\text{MoM}}}$ with a minimum probability ρ. Interestingly, it depends on λ_i and T_s, but not on μ_i itself.

V. Numerical and Simulation Results

This section validates the analytical results developed in this work by means of simulations and evaluates the performance of the three proposed methods. The estimation accuracies are first assessed assuming a sufficiently large sample size (i.e., N large enough to provide the best attainable accuracy). Sample size requirements are analysed afterwards.

Simulations are performed in Matlab by generating a sequence of exponentially distributed random periods T_i with known pre-defined statistics ($\mu_i = 10$ t.u., $\lambda_i = 0.05$), sensing the sequence of periods T_i with a given sensing period T_s in order to calculate the corresponding sequence of estimated periods $\widehat{T}_i$ that would be observed by a DSA/CR system, and finally applying each of the proposed methods to provide an estimation of the minimum period (similar approach as in [9]).

Fig. 2 compares the expression in (5) for the DEM-estimated minimum with simulation results as a function of the sensing period. It can be appreciated, as pointed out in Section III-A, that the provided estimation is accurate ($\widehat{\mu}_i^{\text{DEM}} = \mu_i$) when the sensing period is an integer submultiple of the true minimum (i.e., $T_s = 10/1, 10/2, 10/3, \ldots$) but in general the DEM-estimated minimum period can differ significantly from the true value. A detailed inspection of Fig. 2 also shows that the maximum absolute error is T_s as indicated by (3).

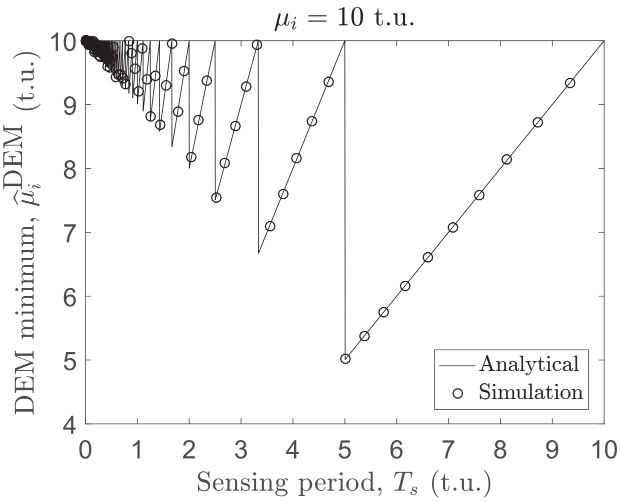

Fig. 2. DEM-estimated minimum as a function of the sensing period.

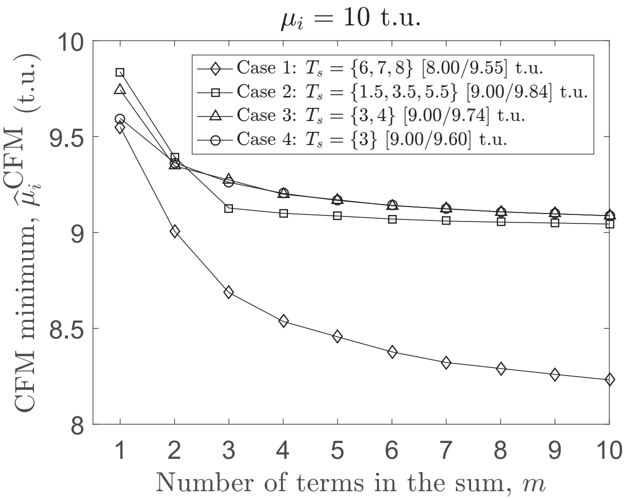

Fig. 3. CFM-estimated minimum as a function of the number of terms.

Fig. 3 shows the CFM-estimated minimum as a function of the number of terms considered in the sum of (6). These results are obtained by fitting (6) to four different pairs of sets $\{T_{s,l}\}_{l=1}^{L}$ and $\{\widehat{\mu}_{i,l}^{\text{DEM}}\}_{l=1}^{L}$. For each pair of sets, the legend shows the set of tested sensing periods $\{T_{s,l}\}_{l=1}^{L}$ and two additional values in the format [*A*/*B*], where *A* is the best DEM-estimation provided individually by any of the tested sensing periods in the set $\{T_{s,l}\}_{l=1}^{L}$ and *B* is the best estimation obtained by means of a non-linear least squares fit (using Matlab's `lsqcurvefit`) of (6) to $\{T_{s,l}\}_{l=1}^{L}$ and $\{\widehat{\mu}_{i,l}^{\text{DEM}}\}_{l=1}^{L}$. It is worth noting that the best CFM estimation is typically obtained, as observed in Fig. 3, when only one term ($m = 1$) is considered in (6). This is because for large m, (6) can reproduce more accurately the discontinuities of the floor function in (5), which leads to a degraded performance of curve-fitting methods (based on gradients). It is also interesting to note that the CFM approach typically achieves better accuracy when the initial value of μ_i for the curve-fitting algorithm is the highest value observed in the sample set $\{\widehat{\mu}_{i,l}^{\text{DEM}}\}_{l=1}^{L}$. As it can be appreciated, the proposed CFM approach can provide accurate estimations of the true minimum period. In Case 1, CFM can reduce the relative error of the DEM-estimated minimum from $(10 - 8)/10 = 20\%$ to $(10 - 9.55)/10 = 4.5\%$, while in the other three cases the relative error is reduced from 10% to 1.6% (Case 2), 2.6% (Case 3), and 4.0% (Case 4). Note that these accurate estimations can be obtained by testing just a few sensing periods, which can be as few as $L = 3$ (cases 1 and 2) or $L = 2$ (case 3) and, in some cases, only $L = 1$ (case 4). This represents a significant reduction compared to the exhaustive search approach discussed in Section III-A.

Fig. 4 shows the MoM-estimated minimum as a function of the sensing period. The results are shown both when the sample variance is corrected as shown in (13) and when it is not corrected (i.e., the term $-T_s^2/6$ is removed from (13)). As it can be appreciated, when the sample moments are computed based on unbiased estimators (12)–(13) and adequately corrected, the MoM approach can provide an accurate (nearly perfect) estimation of the minimum regardless of the employed sensing period. The comparison of Figs. 2, 3 and 4 indicates that MoM is the only method that can provide a very accurate estimation of the minimum period for any sensing period.

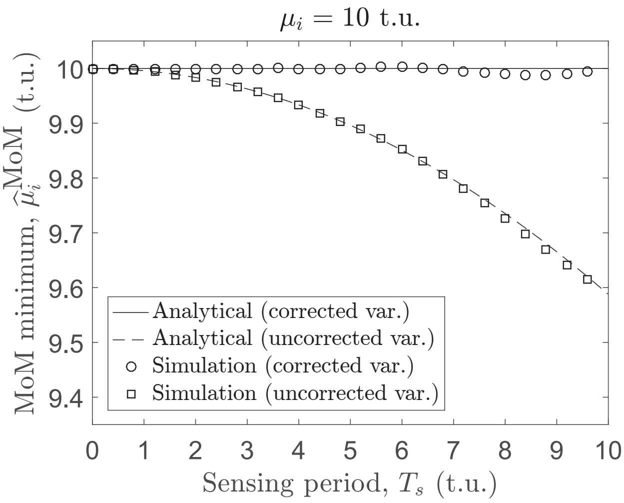

Fig. 4. MoM-estimated minimum as a function of the sensing period.

Fig. 5 shows the DEM sample size as a function of the sensing period. As it can be appreciated, the required sample size increases with the probability $P_{\text{obs}}^{\widehat{\mu}_i^{\text{DEM}}}$. Interestingly, local minima are observed for T_s values that are integer submultiples of the true minimum ($T_s = \mu_i/k$ with $k \in \mathbb{N}^+$). However for slightly higher values the required sample size tends to infinity, since $\chi_0 \approx 0$ and $P\left(\widehat{T}_i = \widehat{\mu}_i^{\text{DEM}}\right) \approx 0$. In general, the sample size required with DEM suffers significant variations depending on the employed sensing period.

Fig. 6 shows the sample size required by CFM for the same four cases shown in Fig. 3. These sample sizes are obtained by adding the individual DEM sample size requirements for each of the tested sensing periods in the set $\{T_{s,l}\}_{l=1}^{L}$ as indicated in (20). As observed, the sample size required by CFM is highly

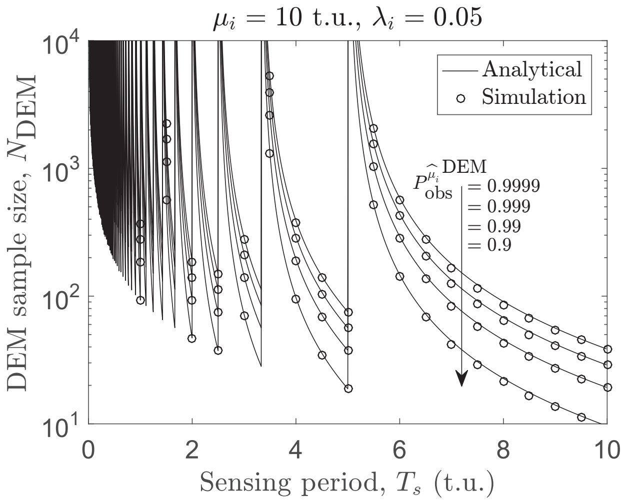

Fig. 5. DEM sample size as a function of the sensing period.

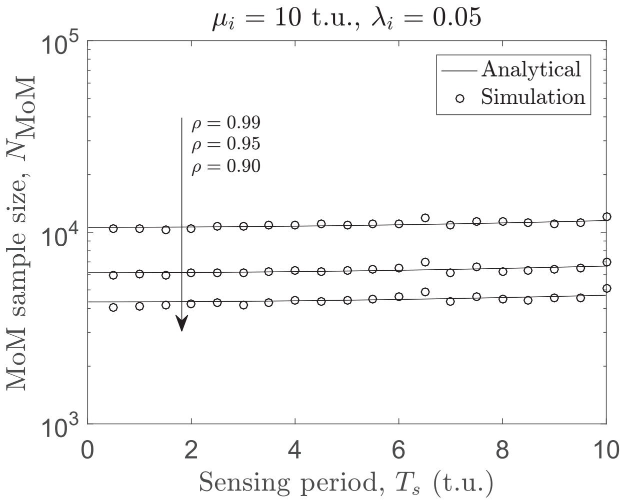

Fig. 7. MoM sample size as a function of the sensing period.

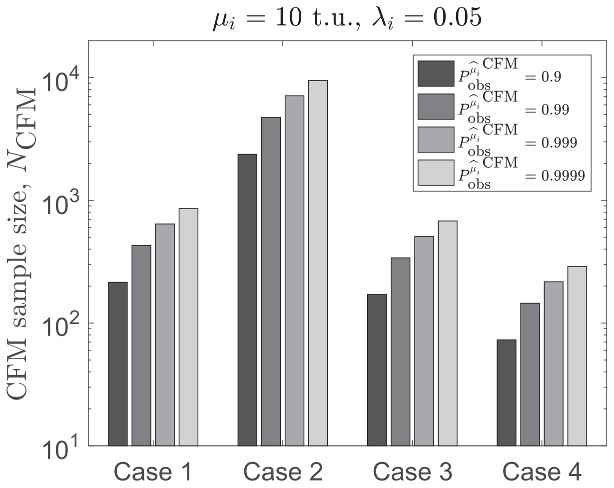

Fig. 6. CFM sample size for each considered case.

variable (one-two orders of magnitude among the considered cases) as it depends on the number L of tested sensing periods as well as the individual sample size required for each of them, which in turn is also highly variable as shown in Fig. 5.

Finally, Fig. 7 shows that the MoM sample size requirement is more stable, although it increases slightly with T_s and more significantly with the desired probability ρ. Comparing Figs. 5, 6 and 7 it can be seen that the better accuracy attained by MoM comes at the expense of larger sample size requirements.

VI. Conclusions

Recent research has demonstrated that an accurate estimation of the minimum primary activity (idle/busy) times plays an important role in the ability to accurately estimate the primary activity statistics from spectrum sensing observations. In this context, this work has proposed and evaluated, both analytically and with simulations, the performance of three estimation methods based on direct estimation (DEM), curve-fitting (CFM) and sample moments (MoM). In general, DEM provides a poor accuracy (except for a few specific cases), which makes of this method an unsuitable option in general. CFM and MoM can provide accurate estimations but establish a tradeoff between provided accuracy and required sample size. In particular, MoM can provide slightly more accurate (nearly perfect) estimations of the true minimum period compared to CFM at the expense of higher sample sizes. In any case, the obtained results have shown that both methods, CFM and MoM, can provide a very accurate estimation of the true minimum period regardless of the employed sensing period.

Acknowledgements

This work was supported by British Council under UKIERI DST Thematic Partnerships 2016-17 (ref. DST-198/2017).

References

[1] M. López-Benítez, "Cognitive radio," in *Heterogeneous cellular networks: Theory, simulation and deployment*. Cambridge University Press, 2013, ch. 13.

[2] Y. Chen and H.-S. Oh, "Spectrum measurement modelling and prediction based on wavelets," *IET Communications*, vol. 10, no. 16, pp. 2192–2198, Oct. 2016.

[3] J. Vartiainen, M. Höyhtyä, J. Lehtomäki, and T. Bräysy, "Priority channel selection based on detection history database," in *Proc. Fifth Int'l. Conf. Cognitive Radio Oriented Wireless Networks & Comms. (CROWNCOM 2010)*, Jun. 2010, pp. 1–5.

[4] K. W. Sung, S.-L. Kim, and J. Zander, "Temporal spectrum sharing based on primary user activity prediction," *IEEE Transactions on Wireless Communications*, vol. 9, no. 12, pp. 3848–3855, Dec. 2010.

[5] A. Al-Tahmeesschi, M. López-Benítez, J. Lehtomäki, and K. Umebayashi, "Accurate estimation of primary user traffic based on periodic spectrum sensing," in *Proc. IEEE Wireless Comms. and Networking Conf. (WCNC 2018)*, Apr. 2018, pp. 1–6, (submitted).

[6] E. C. Titchmarsh and D. R. Heath-Brown, *The theory of the Riemann zeta-function*, 2nd ed. Oxford, 1986.

[7] A. M. Mood, F. A. Graybill, and D. C. Boes, *Introduction to the theory of statistics*, 3rd ed. McGraw-Hill, 1974.

[8] Y. Dodge and V. Rousson, "The complications of the fourth central moment," *American Statistician*, vol. 53, no. 3, pp. 267–269, Aug. 1999.

[9] M. López-Benítez, "Can primary activity statistics in cognitive radio be estimated under imperfect spectrum sensing?" in *Proc. IEEE 24th Int'l. Symp. Pers. Indoor and Mobile Radio Comms. (PIMRC 2013)*, Sep. 2013, pp. 750–755.