On the Performance of Channel Occupancy Rate Estimation with Iterative Threshold Setting in Cognitive Radios

Ming He, Ming Jin, *Member, IEEE*, Qinghua Guo, *Senior Member, IEEE*, Miguel López-Benítez, *Senior Member, IEEE*, and Wei Zhang, *Fellow, IEEE*

***Abstract*—With the knowledge of channel occupancy rate (COR), cognitive radios can significantly improve their performance in exploring and exploiting spectrum holes. However, most existing COR estimators suffer from overestimation or underestimation even at high signal-to-noise ratios (SNRs). The iterative threshold-setting algorithm (ITA) is promising to address this issue. In this work, we revisit ITA and provide a thorough theoretical analysis of ITA. First, we prove that ITA converges to the true COR with a sufficiently large number of traffic samples. Then, we investigate its convergence when the number of traffic samples is small, and show that ITA deviates from the true COR especially at low SNRs. To address this issue, we analyze the upper bound of the number of traffic samples required to achieve a certain estimation error, and further propose an improved ITA (iITA). The proposed iITA enables us to achieve a prespecified estimation accuracy by adaptively adjusting the number of traffic samples. Extensive simulation results are provided, which validate our analyses and demonstrate the superior performance of ITA and iITA compared to state-of-the-art COR estimators.**

***Index Terms*—Cognitive radio, channel occupancy rate, convergence mechanism, iterative algorithm, primary user traffic.**

I. Introduction

DYNAMIC spectrum access (DSA) is promising to improve the spectrum efficiency and alleviate the problem of spectrum shortage. Cognitive radio is one of key technologies for DSA [1]. In cognitive radio (CR), the knowledge of primary user (PU) traffic can be used to significantly improve the performance of functionalities (such as spectrum sensing and channel selection) of secondary users (SUs), thereby leading to more efficient spectrum utilization [2]–[5]. A key parameter characterizing PU traffic is the channel occupancy rate (COR), which represents the occupancy rate of a licensed channel by PUs. It is shown in [6] that the performance of spectrum sensing depends on COR, and in [7] that the sensing performance can be improved by using prior information of COR. In [8], it is shown that the probability of connectivity among SUs can be improved approximately by 50% with the knowledge of COR.

Many DSA schemes assume the perfect knowledge of the COR of PUs [9]. However, the exact knowledge of COR is normally unavailable at SUs, and COR estimation needs to be performed by using measured energy levels from multiple sensing durations. A variety of COR estimators have been proposed [10]–[19]. In [10], the holding time lengths of PU states (ON and OFF) are estimated with the assumption of perfect spectrum sensing, which are then used to calculate the COR. The impacts of sample size and imperfect spectrum sensing on estimating the holding time lengths of PU states are analyzed in [11] and [12], respectively. In [13], a maximum likelihood (ML) estimator was employed to estimate the COR of PUs based on traffic samples (a traffic sample is an estimated energy level of a sensing duration). The issue with the ML estimator is that the computational complexity increases with the number of traffic samples quickly, and it requires the probability density functions of ON/OFF durations of PUs, which, however, are often unavailable [14]–[16]. Different from the ML estimator based on traffic samples, the method in [17] takes the average of the binary decision results of spectrum sensing from a number of sensing durations to estimate the COR. This method was called averaging estimator with uniform sampling (AE-US) [18] and has been applied to the estimation of channel occupancy rate in [19]. The decision threshold for spectrum sensing in AE-US is often set based on a target false-alarm probability (P_f), which results in severe overestimation or underestimation of the COR when P_f is not properly set [2]. In [2], an improved channel occupancy rate (iCOR) estimation method was proposed to alleviate the issue of overestimation or underestimation by transforming the result of AE-US linearly with the use of P_f. In [18], an AE with non-uniform sampling (AE-NS) was proposed, where the first and last traffic sample intervals are larger than others, then, an AE with weighted samples (AE-WS) was proposed, where the weights of the first and last traffic samples are larger than the others. It was shown in [18] that

This work was partially supported by the Natural Science Foundation of China under Grants 61871246 and 61971249, the Australian Research Council's Project funding scheme under LP160101244, the Zhejiang Provincial Natural Science Foundation for Distinguished Young Scholars under Grant LR21F010001, the Zhejiang Provincial Natural Science Foundation under Grand LY18F010008, and the Ningbo Key Program of Natural Science Funds under Grant 202003N4013.

M. He and M. Jin are with the Faculty of Electrical Engineering and Computer Science, Ningbo University, Ningbo 315211, China. (e-mail: hem_nbu@163.com, jinming@nbu.edu.cn).

Q. Guo is with the School of Electrical, Computer and Telecommunications Engineering, University of Wollongong, Wollongong, NSW 2522, Australia, (e-mail: qguo@uow.edu.au).

M. López-Benítez is with the Department of Electrical Engineering and Electronics, University of Liverpool, Liverpool L69 3GJ, U.K., and also with the ARIES Research Centre, Antonio de Nebrija University, 28040 Madrid, Spain (e-mail: m.lopez-benitez@liverpool.ac.uk).

W. Zhang is with School of Electrical Engineering and Telecommunications, The University of New South Wales, Sydney, NSW 2052, Australia (e-mail: w.zhang@unsw.edu.au).

AE-NS and AE-WS outperform AE-US when the number of traffic samples is small. However, AE-NS and AE-WS are designed and configured with the assumption that the PU periods are exponentially distributed, so the sample intervals and the weighting coefficients need to be recalculated for other distributions. In [20], a rank order filtering (ROF) based COR estimator is proposed by identifying the states of primary users at each sampling time instant. However, when the SNR is low and the spectrum occupancy is high, the performance of ROF noticeably degrades. Another way to acquire COR is estimating the average length of holding time of channel states [12], [21]–[23]. In [12] and [21], estimated channel states under imperfect spectrum sensing are reconstructed by fitting the average length of channel holding time with the estimate of a COR estimator, when both false-alarm and miss-detection probabilities are known. In [22], a deep learning network is trained based on a training data set for estimating COR, and the work in [23] analyses the performance of methods proposed in [10] under the assumption of exponentially distributed holding time of channel states.

A drawback of the aforementioned estimators is that the estimate of COR is significantly biased even at high signal-to-noise ratios (SNRs) and with a large number of traffic samples. To address this issue, in this work, we reformulate the COR estimation problem by minimizing the mean-squared error (MSE), and show that, with a sufficiently large number of traffic samples, minimizing MSE leads to a minimized estimation bias. This further leads to the iterative threshold-setting algorithm (ITA) which was originally proposed in [24]. We conduct a thorough theoretical analysis of ITA. Inspired by the analysis, we propose an improved ITA (iITA), where the number of traffic samples is adaptively adjusted to meet the requirement of estimation accuracy. The main contributions of this work are summarized as follows.

(1) We prove that ITA converges to the true COR with an arbitrary initial value, when the number of traffic samples is sufficiently large. We also reveal that an asymptotically unbiased estimate of COR can be achieved by adaptively choosing a proper value of decision threshold in spectrum sensing. Moreover, the convergence behavior of ITA is visualized using evolution trajectory by examining the cumulative distribution function (CDF) of traffic samples and a function obtained by letting the expectation of the estimation bias be zero. It is shown that the two functions have a single intersection, which corresponds to the true COR.

(2) To investigate the impact of a small number of traffic samples on the performance of ITA, we analyze the convergence behavior of ITA from geometrical and statistical analyses based on empirical CDF (ECDF). It is revealed that ITA still converges with a small number of traffic samples, but the convergence point depends on the initial value of COR and may deviate from the true COR severely.

(3) The upper bound of the number of traffic samples required to achieve a prespecified estimation accuracy is derived. Based on the result, an improved ITA (iITA) estimator is proposed, where the traffic sample number can be easily obtained through trial and error.

(4) Extensive simulation results are provided, which validate the theoretical analysis and demonstrate the superior performance of ITA and iITA, compared to existing COR estimators. The two estimators can be implemented in cognitive radio systems to acquire COR estimate efficiently, thereby improving the system performance.

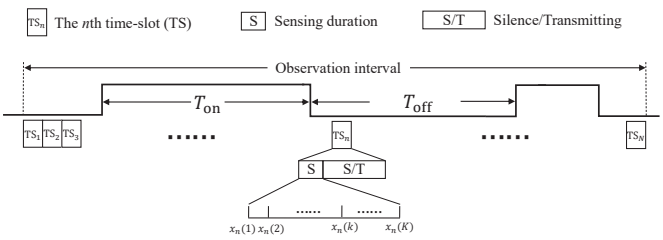


Fig. 1. Diagrammatic sketch of PU states and time-slots of SUs.

In terms of theoretical implication, we reveal that asymptotically unbiased estimation of COR can be achieved by adaptively choosing a proper value of decision threshold in spectrum sensing. Through the thorough analysis of ITA, the convergence behavior of ITA is fully understood, which further leads to new perspective on algorithm design and the iITA algorithm. In terms of practical implication, the implementation of the efficient algorithms ITA and iITA in cognitive radio systems allows us to acquire accurate COR estimates (especially the proposed iITA to deal with small number of traffic samples), therefore significantly improving the performance of cognitive radio systems.

The remainder of this paper is organized as follows. In Section II, the signal model is introduced. In Section III, after revealing that the estimation bias dominates the estimation accuracy of the averaging estimator, we reformulate the COR estimation problem by minimizing the estimation bias, leading to the ITA. In Section IV, we investigate the convergence mechanism of ITA for both large and small number of traffic samples, analyze the number of traffic samples required to achieve a pre-set estimation accuracy, and an improved ITA is proposed. Simulation results are provided in Section V, followed by conclusions drawn in Section VI.

II. Signal Model

Assume that the PU traffic follows a two-state (i.e., OFF and ON) random process. Let T_{OFF} and T_{ON} be the off and on durations of PUs, respectively. We assume that T_{OFF} and T_{ON} are independent random variables with unknown distributions. Let u be the COR of the PU traffic during an observation of several time-slots. Each SU time-slot consists of a sensing duration and a silence/transmitting duration. The change of PU state within a time-slot of SUs may cause interference from SUs. To avoid unacceptable interference to PU transmissions, it is required that the PU state keeps unchanged within a whole time-slot duration for most time-slots. This requirement can be satisfied by assuming that the length of an SU time-slot is far smaller than the OFF and ON durations of the PU traffic,

and so is the length of sensing durations. This is illustrated in Fig. 1. Since a small amount of time-slots within which the PU state may change almost have no effect on the following analysis, to simplify the derivation, we only consider the case where the PU state keeps unchanged within a whole time-slot duration in the following.

Assume that an observation interval consists of N time-slots, i.e., N sensing durations of an SU. Let $x_n(k)$ be the kth sample of the received signal at the SU from the nth sensing duration. It can be expressed as

$$x_n(k) = z_n h_n(k) s_n(k) + w_n(k) \tag{1}$$

where z_n indicates the PU state at the nth time-slot, and we use $z_n = 1$ to represent that the PU is active and $z_n = 0$, otherwise; $h_n(k)$ denotes the Rayleigh channel coefficient from the PU to the SU, and it follows a complex Gaussian distribution with mean zero and variance σ_h^2, i.e., $h_n(k) \sim \mathcal{CN}(0, \sigma_h^2)$; $s_n(k)$ denotes the PU signal, and it is assumed that $s_n(k)$ has an average power of E_s; $w_n(k)$ denotes additive white complex Gaussian noise with mean zero and variance σ_w^2, i.e., $w_n(k) \sim \mathcal{CN}(0, \sigma_w^2)$. Moreover, we assume that $h_n(k)$, $s_n(k)$ and $w_n(k)$ are independent of each other. The SNR is defined as $\text{SNR} = \sigma_h^2 E_s / \sigma_w^2$. Let E_0 and E_1 denote the average power of received signals at the SU under $z_n = 0$ and $z_n = 1$, respectively. Then, we have $E_0 = \sigma_w^2$ and $E_1 = \sigma_h^2 E_s + \sigma_w^2$. The COR which needs to be estimated is defined as $u = \text{Prob}(z_n = 1)$.

Without loss of generality, we employ energy detection (ED) for spectrum sensing. The test-statistic of ED at the nth sensing duration is given by [28]

$$T_{\text{ED}}^n = \frac{1}{K} \sum_{k=0}^{K-1} |x_n(k)|^2 \tag{2}$$

where K is the total number of samples within a sensing duration. Let $\hat{z}_n$ be the decision of the PU state at the nth sensing duration of the SU. Then, the false-alarm probability is given by

$$\begin{aligned} P_f &= \text{Prob}(T_{\text{ED}}^n > \lambda | z_n = 0) \\ &= \text{Prob}(\hat{z}_n = 1 | z_n = 0) \end{aligned} \tag{3}$$

where λ denotes the decision threshold of ED. The detection probability is given by

$$\begin{aligned} P_d &= \text{Prob}(T_{\text{ED}}^n > \lambda | z_n = 1) \\ &= \text{Prob}(\hat{z}_n = 1 | z_n = 1). \end{aligned} \tag{4}$$

Further, we can obtain that [25]–[27]

$$P_f = Q\left(\frac{\lambda - E_0}{\frac{1}{\sqrt{K}} E_0} \right) \tag{5}$$

and

$$P_d = Q\left(\frac{\lambda - E_1}{\frac{1}{\sqrt{K}} E_1} \right) \tag{6}$$

where $Q(\cdot)$ is the complementary function of the cumulative distribution for a standard Gaussian random variable. Note that P_f and P_d are monotonically decreasing functions with respect to the decision threshold λ.

The decision results $\{\hat{z}_n, n = 1, 2, \cdots, N\}$ are often used to estimate the COR u. However, existing estimators may give rise to significant estimate deviation. In the following, we investigate the reason and introduce ITA-based COR estimator and its modified version to overcome the problem.

III. Channel Occupancy Rate Estimation with Iterative Threshold Setting

Fig. 1 shows a diagrammatic sketch of PU states and time-slots of an SU. In each time-slot, the SU makes a decision on the PU state in the sensing duration and then transmits or keeps silent according to the decision in the following S/T duration. The decisions from N successive sensing durations are employed to estimate the COR. As each decision $\hat{z}_n$ takes on $\hat{z}_n = 1$ when PU is claimed to exist, and $\hat{z}_n = 0$ when PU is claimed to be absent, according to [18], the average of N decisions can be used for estimating the COR as

$$\hat{u} = \frac{1}{N} \sum_{n=1}^{N} \hat{z}_n. \tag{7}$$

If we assume that nearly perfect sensing performance is achieved, i.e., $P_f \approx 0$ and $P_d \approx 1$, $\hat{u}$ is a good estimate of u. However, sensing errors are inevitable, which can degrade the estimation performance significantly.

We analyze the bias and MSE of the estimator in (7). Conditioned on the latent variable z_n, $\hat{z}_n$ follows a Bernoulli distribution with probabilities[1]

$$\text{Prob}(\hat{z}_n = 1) = uP_d + (1-u)P_f \tag{8}$$

and

$$\text{Prob}(\hat{z}_n = 0) = u(1-P_d) + (1-u)(1-P_f). \tag{9}$$

Therefore, the variance of $\hat{z}_n$ can be obtained as

$$\mathbb{V}[\hat{z}_n] = \left(uP_d + (1-u)P_f\right)\left(u(1-P_d) + (1-u)(1-P_f)\right). \tag{10}$$

Considering that the elements in $\{\hat{z}_n, \forall n\}$ conditioned on the latent variables $\{z_n, \forall n\}$ are independent and identically distributed [29], we can obtain that $\hat{u}$ follows a scaled binomial distribution and its expectation and variance are respectively given by

$$\mathbb{E}[\hat{u}] = uP_d + (1-u)P_f \tag{11}$$

and

$$\mathbb{V}[\hat{u}] = \frac{1}{N}\left(uP_d + (1-u)P_f\right)\left(u(1-P_d) + (1-u)(1-P_f)\right) \tag{12}$$

where $\mathbb{E}[\cdot]$ and $\mathbb{V}[\cdot]$ denote the expectation and variance operators, respectively. Then, the estimation bias and MSE can be respectively obtained as

$$e_{\text{bias}} = \mathbb{E}[\hat{u}] - u = uP_d + (1-u)P_f - u \tag{13}$$

and

$$e_{\text{MSE}} = \mathbb{V}[\hat{u}] + e_{\text{bias}}^2. \tag{14}$$

With a sufficiently large N, we can obtain from (12) that $\mathbb{V}[\hat{u}] \approx 0$. Thus,

$$e_{\text{MSE}} \approx e_{\text{bias}}^2. \tag{15}$$

Algorithm 1 ITA for estimating COR

Initialization:

1: Select a value for $u^{(0)} \in (0,1)$.
2: $l = 0$ (index).
3: **repeat**
4: $l \leftarrow l+1$.
5: Obtain $\lambda^{(l)}$ by solving $u^{(l-1)}P_d(\lambda) + (1-u^{(l-1)})P_f(\lambda) - u^{(l-1)} = 0$ with the bisection algorithm.
6: Obtain $u^{(l)}$ using (7) or (19) with decisions by comparing T_{ED}^n $(n = 1, 2, \cdots, N)$ with $\lambda^{(l)}$.
7: **until** $|u^{(l)} - u^{(l-1)}| \leq \epsilon$

Output: $\lambda^{(l)}$ is the optimized decision threshold, and $u^{(l)}$ is an estimate of the COR.

According to (15), the estimation bias dominates the MSE. Hence, we focus on the analysis of the estimation bias of the estimator (7). In [18], the performance of the estimator is analyzed for the two cases of imperfect sensing with $P_f = 0.05$, $P_d = 0.95$ and $P_f = 0.1$, $P_d = 0.9$. For the former case, we can obtain that the estimation bias $e_{\text{bias}} = 0.05 - 0.1u$ and for the latter, $e_{\text{bias}} = 0.1 - 0.2u$. These indicate that the estimation bias e_{bias} is small only when u is close to 0.5, otherwise the absolute estimation bias e_{bias} is high when u is far away from 0.5. More generally, they indicate that the estimation bias is small only for specific u when P_d and P_f are determined.

Next, we introduce the ITA which can decrease the estimate bias to be sufficiently small. Note that P_f and P_d are monotonically decreasing functions with respect to the decision threshold λ, and the range of e_{bias} in (13) is $[-u, 1-u]$. Hence, e_{bias} can be reduced to zero by optimizing λ. Once an unbiased estimate is achieved, i.e., $e_{\text{bias}} = 0$, (13) gives

$$e_{\text{bias}} = uP_d(\lambda) + (1-u)P_f(\lambda) - u = 0. \tag{16}$$

Equation (16) involves two variables u and λ. It is noted that, with a fixed value of u, $uP_d(\lambda) + (1-u)P_f(\lambda) - u$ is a monotonically decreasing function of λ. Hence, with the true value of the COR u^*, there is a unique solution of λ to (16), which is denoted by λ^*. As u^* is unknown, we cannot determine λ^*. Note that, with N traffic samples, one can obtain an estimate of COR with (7) by comparing the traffic samples with a threshold. This leads to the ITA [24] for estimating the COR. In other words, we start with any initial value of $u \in (0,1)$ and determine the value of λ and estimate u alternately. The algorithm is summarized in **Algorithm** 1. It is interesting that the iterative algorithm converges to u^* and λ^* for a sufficiently large N, which will be proved in the next Section. It is noted that the optimization of the decision threshold here is purely to achieve an unbiased estimate of the COR and this may cause the variation of P_f with SNR as shown in Sec. V-B. It should be distinguished from the decision threshold in spectrum sensing to achieve a pre-specified P_f.

The iteration stopping criterion is given by [30]

$$|u^{(l)} - u^{(l-1)}| \leq \epsilon \tag{17}$$

[1]The condition "$\cdot|z_n$" will be omitted without confusion in the following.

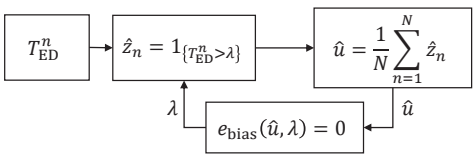


Fig. 2. Illustration of ITA.

where l denotes the iteration index and ϵ is a small number. In the iterative algorithm, the solution of λ to $u^{(l)}P_d(\lambda) + (1-u^{(l)})P_f(\lambda) - u^{(l)} = 0$ can be obtained by the bisection algorithm, as $u^{(l)}P_d(\lambda) + (1-u^{(l)})P_f(\lambda) - u^{(l)}$ is a monotonically decreasing function with respect to λ.

The iterative process of ITA is illustrated in Fig. 2, where $1_{\{\cdot\}}$ denotes an indicator function. We can see that ITA consists of three main operation modules: spectrum sensing, averaging decisions and calculating decision threshold.

IV. Performance Analysis of ITA and Improved ITA

In this section, we first analyze the convergence of the ITA with a sufficiently large number of traffic samples, and investigate the evolution trajectory of the estimation with iterations. For a small number of traffic samples and in the low SNR regime, we find that the iterative algorithm may get stuck at a point far from the true value. To address this issue, we examine the possible range to which the ITA converges from the perspective of probability framework, and then investigate the asymptotic performance of the estimator. After that, we analyze the number of traffic samples required to guarantee an estimation accuracy, and then propose an improved ITA which can achieve an expected estimation accuracy by adaptively adjusting the number of traffic samples. Finally, the computational complexity of ITA is analyzed.

A. Performance Analysis with Large Sample Size and Evolution Trajectory

In **Algorithm** 1, with $u^{(l-1)}$, the lth iteration produces $\lambda^{(l)}$ and $u^{(l)}$. First, $\lambda^{(l)}$ is obtained by solving

$$u^{(l-1)} = u^{(l-1)}P_d(\lambda^{(l)}) + (1-u^{(l-1)})P_f(\lambda^{(l)}), \tag{18}$$

and the solution $\lambda^{(l)}$ is unique. Then, $u^{(l)}$ is obtained by comparing T_{ED}^n $(n = 1, 2, \cdots, N)$ with $\lambda^{(l)}$, i.e.,

$$u^{(l)} = \frac{1}{N}\sum_{n=1}^{N} \hat{z}_n^{(l)} \tag{19}$$

where $\hat{z}_n^{(l)} = 1$ for $T_{\text{ED}}^n > \lambda^{(l)}$, and $\hat{z}_n^{(l)} = 0$ otherwise. With the number of traffic samples N, we can obtain that

$$\begin{aligned} u^{(l)} &= \text{Prob}(z_n = 1)\text{Prob}(\hat{z}_n = 1 | z_n = 1) \\ &\quad + \text{Prob}(z_n = 0)\text{Prob}(\hat{z}_n = 1 | z_n = 0) \\ &= u^* \hat{P}_d(\lambda^{(l)}) + (1-u^*)\hat{P}_f(\lambda^{(l)}) \end{aligned} \tag{20}$$

where

$$\hat{P}_d(\lambda^{(l)}) = \frac{1}{Nu^*}\sum_{n} 1_{\{T_{\text{ED}}^n > \lambda^{(l)} | z_n = 1\}} \tag{21}$$

and

$$\hat{P}_f(\lambda^{(l)}) = \frac{1}{N(1-u^*)} \sum_n 1_{\{T_{\text{ED}}^n > \lambda^{(l)} | z_n = 0\}} \tag{22}$$

with $1_{\{\cdot\}}$ being an indicator function.

Theorem 1: The ITA algorithm (**Algorithm** 1) converges to (λ^*, u^*) for a sufficiently large N.

Proof 1: First, we prove the convergence of the algorithm for the case $u^{(0)} > u^*$.

It can be easily virified that $\lim_{N\to+\infty} \hat{P}_d(\lambda^{(l)}) = P_d(\lambda^{(l)})$ and $\lim_{N\to+\infty} \hat{P}_f(\lambda^{(l)}) = P_f(\lambda^{(l)})$. Hence, with a sufficiently large N, (20) can be rewritten as

$$u^{(l)} = u^* P_d(\lambda^{(l)}) + (1-u^*) P_f(\lambda^{(l)}) \tag{23}$$

where $P_f(\lambda^{(l)})$ and $P_d(\lambda^{(l)})$ can be calculated by (5) and (6).

We rewrite (16) as

$$g(\lambda) = f(u) \tag{24}$$

where on the left-hand side

$$g(\lambda) = \frac{1 - P_d(\lambda)}{P_f(\lambda)} \tag{25}$$

is a monotonically increasing function with respect to λ and on the right-hand side

$$f(u) = \frac{1-u}{u} \tag{26}$$

is a monotonically decreasing function with respect to u. Note that (24) holds with (λ^*, u^*), i.e.,

$$g(\lambda^*) = f(u^*). \tag{27}$$

Without loss of generality, we let

$$u^{(l-1)} > u^* \tag{28}$$

for an arbitrary $l \geq 1$. Then, we can obtain $\lambda^{(l)}$ by solving

$$g(\lambda) = f(u^{(l-1)}). \tag{29}$$

Because $g(\lambda)$ is a monotonically increasing function, $f(u)$ is a monotonically decreasing function and $u^{(l-1)} > u^*$, we can obtain that

$$\lambda^{(l)} < \lambda^* \tag{30}$$

which indicates that λ is underestimated at the lth iteration (compared with λ^*). Further, the underestimated decision threshold $\lambda^{(l)}$ will produce higher P_f and P_d (compared with $P_f(\lambda^*)$ and $P_d(\lambda^*)$) and then an overestimated COR $u^{(l)}$, i.e.,

$$u^{(l)} > u^*. \tag{31}$$

This implies that $u^{(l-1)}$ and $u^{(l)}$ are at the same side of u^*, and both of them are larger than u^* for any l.

Next, we prove that $u^{(l)}$ decreases with respect to l. Note that the average power of received signals at an SU is statistically smaller under $\mathcal{H}_0$ (i.e., E_0) than $\mathcal{H}_1$ (i.e., E_1). Hence, following (5) and (6), we can obtain that

$$P_d(\lambda^{(l)}) > P_f(\lambda^{(l)}). \tag{32}$$

From (18) and (20), we can obtain that

$$u^{(l-1)} - u^{(l)} = (u^{(l-1)} - u^*) \left(P_d(\lambda^{(l)}) - P_f(\lambda^{(l)}) \right). \tag{33}$$

Note that the right-hand side of (33) is larger than zero. Thus, for $u^{(l-1)} > u^*$, we have

$$u^{(l-1)} > u^{(l)}. \tag{34}$$

Combining (31) and (34) gives

$$u^{(l-1)} > u^{(l)} > u^*. \tag{35}$$

This indicates that $u^{(l)}$ decreases with the iteration number l, and it converges to some value. Therefore,

$$\lim_{l\to+\infty} \left(u^{(l)} - u^{(l-1)} \right) = 0. \tag{36}$$

With the above derivations, we will show that $u^{(l)}$ converges to u^*. As

$$P_d(\lambda^{(l)}) - P_f(\lambda^{(l)}) > 0, \tag{37}$$

according to (33) and (36), we can obtain that

$$\lim_{l\to+\infty} \left(u^{(l-1)} - u^* \right) = 0. \tag{38}$$

which indicates that $u^{(l)}$ converges to u^*. Then, $\lambda^{(l)}$ converges to λ^*.

Similarly, we can also prove the convergence of the algorithm for the case $u^{(0)} < u^*$, which is shown in Appendix A.

■

In the COR estimator, λ and u are alternatively calculated based on an initial value u. It can be obtained that the optimal value (u^*, λ^*) is the intersection of the two functions

$$u = u P_d(\lambda) + (1-u) P_f(\lambda) \tag{39}$$

and

$$u = u^* P_d(\lambda) + (1-u^*) P_f(\lambda). \tag{40}$$

The former one is for calculating λ in each iteration as in (18) and the latter represents obtaining an updated estimate of u by comparing a sufficiently large number of traffic samples with a decision threshold λ. Fig. 3 shows the curves of the functions in (39) and (40) where $u^* = 0.3$, $K = 100$ and SNR $= -10$dB, and the evolution trajectory of λ and u with iteration from an initial value of $u = 0.9$ is represented by dashed lines with arrow heads. It can be observed from the trajectory that the required iteration number is small to achieve an estimate with a small error.

B. Impact of Small Traffic Sample Size

The COR u with respect to λ in (40) can be regarded as an empirical cumulative distribution function with a sufficiently large number of traffic samples (i.e., $N \to +\infty$). With a more practical value of N, (40) becomes

$$u = u^* \hat{P}_d(\lambda) + (1-u^*) \hat{P}_f(\lambda) \tag{41}$$

where $\hat{P}_d(\lambda)$ and $\hat{P}_f(\lambda)$ are defined in (21) and (22). Hence, the curve of (41) is not a deterministic one as that of (40). Equation (41) represents a stepped line, which may have multiple intersections with the curve of (39). Fig. 4 shows the curves of the functions in (39) and (41) when $u^* = 0.3$, $K = 100$, $N = 100$ and SNR $= -10$dB. Multiple intersections around

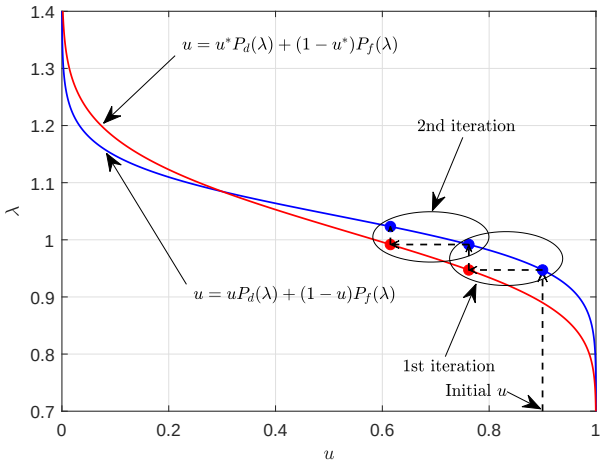


Fig. 3. Curves of the functions in (39) and (40) when $u^* = 0.3$, $K = 100$ and SNR $= -10$dB. Dashed lines show the evolution trajectory of λ and u with iteration from an initial value of $u = 0.9$.

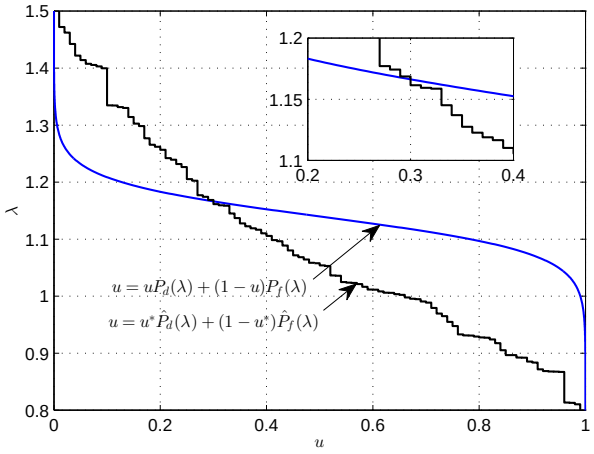


Fig. 5. Curves of the functions in (39) and (41) when $u^* = 0.3$, $K = 100$, $N = 100$ and SNR $= -5$dB. The stepped line comes from a realization of 100 traffic samples, and intersections vary with different realizations.

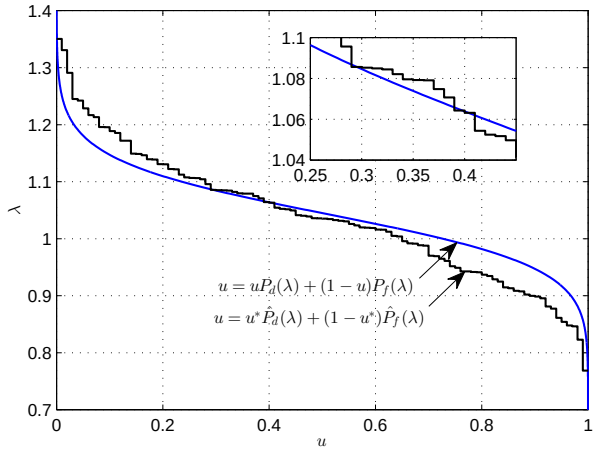


Fig. 4. Curves of the functions in (39) and (41) when $u^* = 0.3$, $K = 100$, $N = 100$ and SNR $= -10$dB. The stepped line comes from a realization of 100 traffic samples, and intersections vary with different realizations.

$u^* = 0.3$ can be found in Fig. 4, which indicates that the algorithm converges to $u = 0.4$ when the initial value of u is larger than 0.4, while converges to $u = 0.27$ when the initial value of u is smaller than 0.27. By increasing SNR to -5dB, Fig. 5 shows that the curves of the functions in (39) and (41) have only one intersection at $u = 0.3$. Hence, the algorithm will converge to $u = 0.3$ with an initial value of u no matter larger or smaller than 0.3.

It has been proved in Sec. IV-A that the ITA converges to u^* with a sufficiently large N. However, it has been shown that, with a small or more practical number of traffic samples and at low SNRs, the estimator may not converge to u^*. Next we examine the possible range to which the ITA converges from the perspective of probability framework, and then investigate the asymptotic performance of the estimator.

First, we consider the case of $u^{(0)} > u^*$. Given that $u^{(l-1)}$ has been obtained from the $(l-1)$th iteration, we can derive the probability

$$\eta = \text{Prob}\left(u^{(l)} < u^{(l-1)}\right) \tag{42}$$

where η denotes the confidence of the event $u^{(l)} < u^{(l-1)}$. Equation (42) means that the estimate $u^{(l)}$ at the lth iteration moves from $u^{(l-1)}$ toward (rather than the inverse direction) u^* with a probability of η.

As $u^{(l-1)}$ has been given at the $(l-1)$th iteration, we can get $\lambda^{(l)}$ with (18) and then $u^{(l)}$ with (19) at the lth iteration. With the central limit theorem [31], it can be obtained that $u^{(l)}$ approximately follows a Gaussian distribution with mean $\mathbb{E}[u^{(l)}]$ and variance $\mathbb{V}[u^{(l)}]$, i.e.,

$$u^{(l)} \sim \mathcal{N}\left(\mathbb{E}[u^{(l)}], \mathbb{V}[u^{(l)}]\right). \tag{43}$$

Similar to (11) and (12), the mean and the variance of $u^{(l)}$ are respectively given by

$$\mathbb{E}[u^{(l)}] = u^* P_d(\lambda^{(l)}) + (1-u^*) P_f(\lambda^{(l)}) \tag{44}$$

and

$$\begin{aligned}\mathbb{V}[u^{(l)}] = \frac{1}{N}&\left(u^* P_d(\lambda^{(l)}) + (1-u^*) P_f(\lambda^{(l)})\right) \\ &\times \left(u^*(1-P_d(\lambda^{(l)})) + (1-u^*)(1-P_f(\lambda^{(l)}))\right).\end{aligned} \tag{45}$$

Moreover, (18) leads to

$$u^{(l-1)} = \frac{P_f(\lambda^{(l)})}{1 - P_d(\lambda^{(l)}) + P_f(\lambda^{(l)})}. \tag{46}$$

$$\text{Prob}\left(u^{(l)} < u^{(l-1)}\right) = 1 - Q\left(\frac{\sqrt{N}\left(\frac{P_f(\lambda^{(l)})}{1-P_d(\lambda^{(l)})+P_f(\lambda^{(l)})} - u^* P_d(\lambda^{(l)}) - (1-u^*) P_f(\lambda^{(l)})\right)}{\sqrt{\left(u^* P_d(\lambda^{(l)}) + (1-u^*) P_f(\lambda^{(l)})\right)\left(u^*(1-P_d(\lambda^{(l)})) + (1-u^*)(1-P_f(\lambda^{(l)}))\right)}}\right) \tag{47}$$

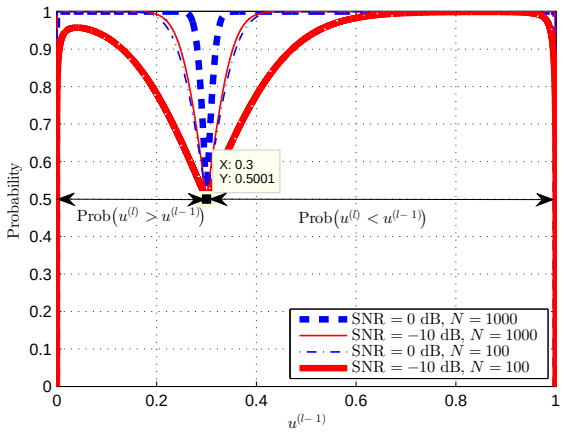


Fig. 6. Probability that $u^{(l)}$ at the lth iteration moves toward u^* from different $u^{(l-1)}$ of the $(l-1)$th iteration when $u^* = 0.3$ and $K = 100$.

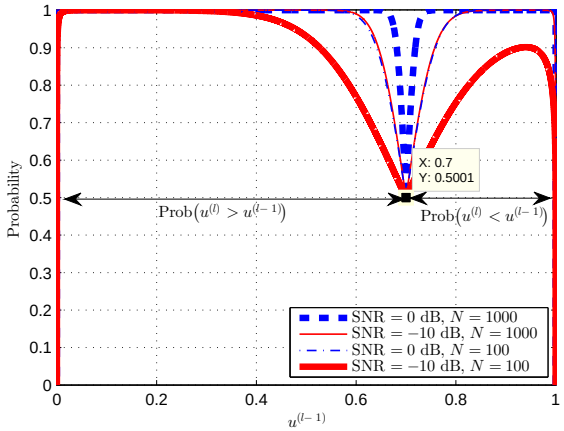


Fig. 7. Probability that $u^{(l)}$ at the lth iteration moves toward u^* from different $u^{(l-1)}$ of the $(l-1)$th iteration when $u^* = 0.7$ and $K = 100$.

Combining (42)-(46) gives (47) at the bottom of this page.

For the case of $u^{(0)} < u^*$, the same result as in (47) can also be derived for Prob $\left(u^{(l)} > u^{(l-1)}\right)$.

Fig. 6 shows the probability that $u^{(l)}$ at the lth iteration moves toward u^* from different $u^{(l-1)}$ of the $(l-1)$th iteration when $u^* = 0.3$ and $K = 100$. It is observed from Fig. 6 that the probability approaches zero as $u^{(l-1)}$ is close to zero or one. This implies that we should not choose $u^{(0)}$ too close to zero or one. For other values of $u^{(l-1)}$, the probability of $u^{(l-1)}$ moving toward u^* is high when $u^{(l-1)}$ is far away from u^*. In other words, ITA will converge to a value far from u^* with a low probability, especially for $N = 1000$. Around u^*, a notch of probability whose value is less than one is formed, that indicates the range of possible estimates of the ITA. Moreover, it is depicted that the width of the notch is narrowed by increasing SNR or the number of traffic samples. This implies that the average estimation error can be decreased by increasing SNR or the number of traffic samples. Fig. 7 shows the probability that $u^{(l)}$ at the lth iteration moves toward u^* from different $u^{(l-1)}$ of the $(l-1)$th iteration when $u^* = 0.7$ and $K = 100$. The same conclusions as those in Fig. 6 can also be drawn.

When N goes to infinity, based on the above analyses, we can obtain that $\lambda^{(l)}$ converges to λ^* with a probability of one. Let $u^{(l)}$ be the final estimate of COR, i.e., $\hat{u}$, then, from (44) we have

$$\mathbb{E}[\hat{u}] = u^* P_d(\lambda^*) + (1-u^*) P_f(\lambda^*) = u^* \tag{48}$$

which implies that the expectation of the estimate equals its true value when N goes to infinity, i.e., the estimate obtained by the ITA is asymptotically unbiased.

C. Required Traffic Samples With Constraint of Estimation Error

We have demonstrated the convergence of ITA with large sample size N and analyzed the impact of N. Here we analyze the upper bound on N in order to achieve a pre-set estimation accuracy.

As the number of traffic samples is N and the averaging estimator is used, all the possible estimates of u^* are $\{\frac{1}{N}, \frac{2}{N}, \cdots, \frac{N-1}{N}\}$. From (47) we can obtain the probability that the next iteration gets closer to u^* when the current iteration takes a value from $\{\frac{1}{N}, \frac{2}{N}, \cdots, \frac{N-1}{N}\}$.

Given $u^{(0)}$ with $u^{(0)} < u^* - \alpha$, a lower bound of probability that the estimate eventually falls within $[u^* - \alpha, u^* + \alpha]$ is given by

$$\begin{aligned} &\text{Prob}\left(u^* - \alpha < \hat{u} < u^* + \alpha\right) \\ &= \prod_{u_i^{(0)} \in U} \text{Prob}\left(u_i^{(0)} < u_i^{(1)}\right) \end{aligned} \tag{49}$$

where α is a tolerable estimation error, $u_i^{(1)}$ represents the estimate from $u_i^{(0)}$ after one iteration, and $U = \{u^{(0)}, u^{(0)} + \frac{1}{N}, u^{(0)} + \frac{2}{N}, \cdots, u^* - \alpha\}$.

Similarly, a lower bound of probability that the estimate eventually falls within $[u^* - \alpha, u^* + \alpha]$ when $u^{(0)} > u^* + \alpha$ is given by

$$\begin{aligned} &\text{Prob}\left(u^* - \alpha < \hat{u} < u^* + \alpha\right) \\ &= \prod_{u_i^{(0)} \in V} \text{Prob}\left(u_i^{(0)} > u_i^{(1)}\right) \end{aligned} \tag{50}$$

where $V = \{u^{(0)}, u^{(0)} - \frac{1}{N}, u^{(0)} - \frac{2}{N}, \cdots, u^* + \alpha\}$.

Since (47) is a monotonically increasing function with respect to N, (49) and (50) also monotonically increase with N. Given α, $u^{(0)}$ and Prob $\left(u^* - \alpha < \hat{u} < u^* + \alpha\right)$, we can obtain N by solving (49) or (50) with the bisection algorithm. It is noted that in the calculation of N, we assume that the iteration experiences all possible estimates. In practice, we do not need such a large number of iterations. Hence, the calculated N is larger than the one actually required and is an upper bound.

Fig. 8 shows the required N of the upper bound and Monte Carlo simulation results at different SNR when $u^* = 0.3$, $\alpha = 0.05$ and Prob $\left(u^* - \alpha < \hat{u} < u^* + \alpha\right) = 0.9$. It is observed from Fig. 8 that the upper bound of N is larger than the one obtained from Monte Carlo simulations, but the gap between them is not that big. This is consistent with what we described

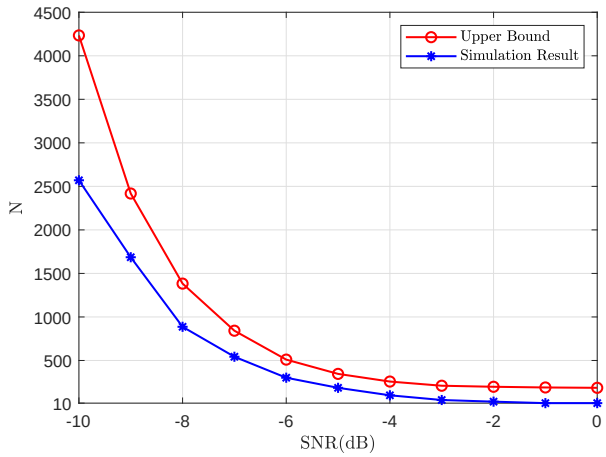


Fig. 8. Companions of Ns obtained through analysis (upper bound) and Monte Carlo simulations at different SNR when $u^* = 0.3$ and $K = 100$.

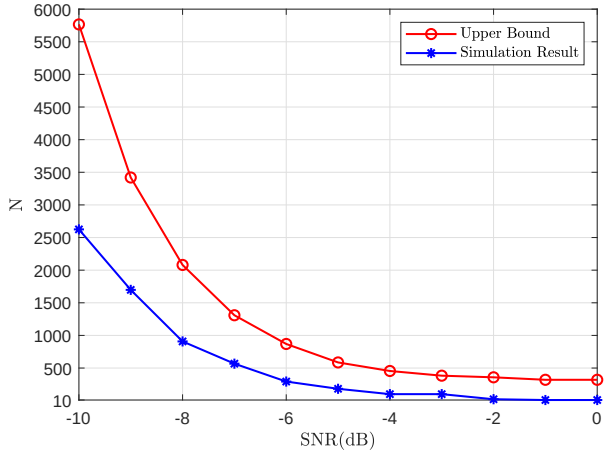


Fig. 9. Comparisons of Ns obtained through analysis (upper bound) and Monte Carlo simulations at different SNR when $u^* = 0.7$ and $K = 100$.

above and shows that the value of N obtained through the theoretical analysis about N is reasonable and valid. Fig. 9 shows the upper bound of N and its Monte Carlo simulation results at different SNR when $u^* = 0.7$. The same conclusions as those in Fig. 8 can be drawn.

D. Improved ITA

Based on the convergence analysis of ITA and the analysis of the required sample number N, it can be concluded that ITA can achieve a good performance with a large N, but its performance will be degraded when the number of traffic samples is small at low SNR. To effectively control the estimation error of COR, a modified version of the ITA is proposed by introducing a variable N.

Although we have known the upper bound of N, the use of the upper bound will cause a waste of resources as the bound may be considerably larger than N. To avoid this, we propose a new method that can find an appropriate number of N based on the given tolerable estimation error α. Given a traffic sample size N_t at current time t, we can estimate the COR by using ITA. Hence, one can estimate COR at the time t, i.e., $\hat{u}_t$, with the initial value being previous estimate $\hat{u}_{t-1}$

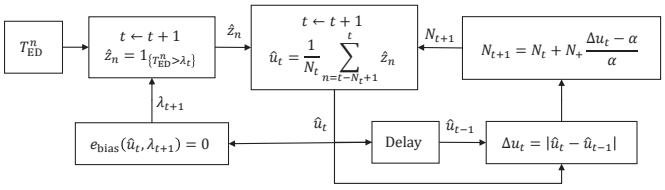


Fig. 10. Illustration of iITA.

at time $t-1$. The difference between the current estimate $\hat{u}_t$ and the previous estimate $\hat{u}_{t-1}$ is given by

$$\Delta u_t = |\hat{u}_t - \hat{u}_{t-1}|. \tag{51}$$

After getting Δu_t, we can recalculate the traffic sample size for the next time $t+1$ as

$$N_{t+1} = N_t + N_+ \frac{\Delta u_t - \alpha}{\alpha} \tag{52}$$

where N_+ is a positive integer.

Fig. 10 shows the iterative process of the iITA, where we can see that iITA includes additional modules for adaptively adjusting the number of traffic samples, compared to ITA.

E. Computational Complexity

Next, we analyze the complexity of the iterative algorithm of ITA and iITA. Note that $P_d(\lambda)$ and $P_f(\lambda)$ can be easily obtained from a look-up table of the CDF for a Gaussian distribution. Thus, only 2 multiplications and 3 additions are required to obtain the value of $u^{(l-1)}P_d(\lambda)+(1-u^{(l-1)})P_f(\lambda)-u^{(l-1)}$. Therefore, with a tolerance error δ relative to the initial interval, the maximum iteration number of the bisection algorithm is given by [30]

$$N_{bisec} = \log_2 \frac{1}{\delta}. \tag{53}$$

Let N_l be the maximum iteration number for $u^{(l)}$. Then the COR estimator requires at most $2N_lN_{bisec}$ multiplications and $3N_lN_{bisec}$ additions. Taking $\delta = 10^{-3}$ and $N_l = 12$ as an example, at most 240 multiplications and 360 additions are required for the iterative algorithm. Hence, the algorithm has a low computational complexity.

V. Simulation Results

Extensive simulation results are provided to demonstrate the superiority of the ITA and the improved estimator, compared to state-of-the-art estimators. The noise variance σ_w^2 and the Rayleigh channel variance σ_h^2 are all set to one. Then, the PU average power can be calculated with SNR $\times\ \sigma_w^2/\sigma_h^2$. Considering that the performance of energy detection (which is employed for spectrum sensing in our proposed method) depends on SNR (no matter what combination of PU average power and Rayleigh channel variance is), in the simulations, we keep the channel variance σ_h^2 unchanged and adjust the PU average power to achieve different SNRs. Moreover, the lengths of ON and OFF durations are randomly drawn from generalized Pareto distributions [10]. The location, scale and shape parameters of the distribution are set as $\rho_{\mathrm{ON}} = \rho_{\mathrm{OFF}} = 1$,

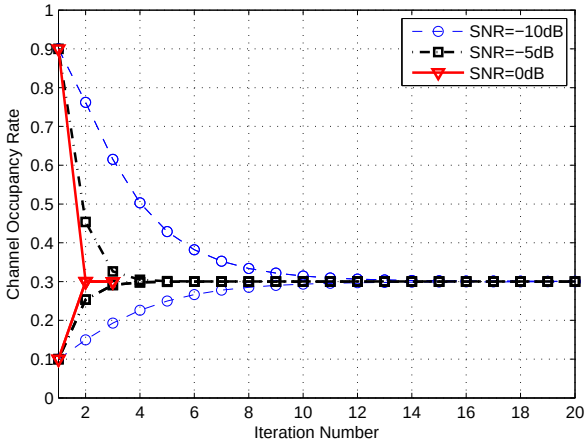


Fig. 11. Convergence of the proposed algorithm for $u^* = 0.3$ at different SNR.

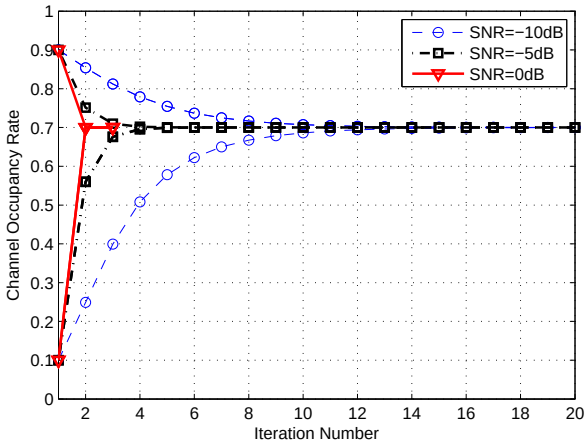


Fig. 12. Convergence of the proposed algorithm for $u^* = 0.7$ at different SNR.

$\beta_{\text{ON}} = 3$, $\beta_{\text{OFF}} = 1$ and $\xi_{\text{ON}} = \xi_{\text{OFF}} = 0.5$ for $u^* = 0.7$, and $\rho_{\text{ON}} = \rho_{\text{OFF}} = 1$, $\beta_{\text{ON}} = 1$, $\beta_{\text{OFF}} = 3$ and $\xi_{\text{ON}} = \xi_{\text{OFF}} = 0.5$ for $u^* = 0.3$. The unit of the parameters for the generalized Pareto distribution is time-slot, i.e., the primary user has the minimum holding time of one time-slot but has no limit on the maximum holding time.

A. Convergence Speed

Fig. 11 shows the estimated COR with iteration number for $u^* = 0.3$ at different SNR. In the simulations, $K = 100$ and $N = 10000$. For each SNR, different initial values of $u^{(0)}$ (e.g., $u^{(0)} = 0.1$ and 0.9) are selected. It can be observed from Fig. 11 that the ITA algorithm converges for all initial values for this group of traffic samples. Moreover, at $\text{SNR} = 0\text{dB}$, the algorithm converges with only one iteration. With the same parameters but $u^* = 0.7$, Fig. 12 shows the estimated COR with iteration number. It can be observed that the iterative algorithm still converges fast.

B. Performance Evaluation with u^*

In this subsection, we evaluate the performances of different COR estimators when $u^* = 0.3$ or $u^* = 0.7$. RMSE (Root Mean Square Error) and bias are adopted as criteria.

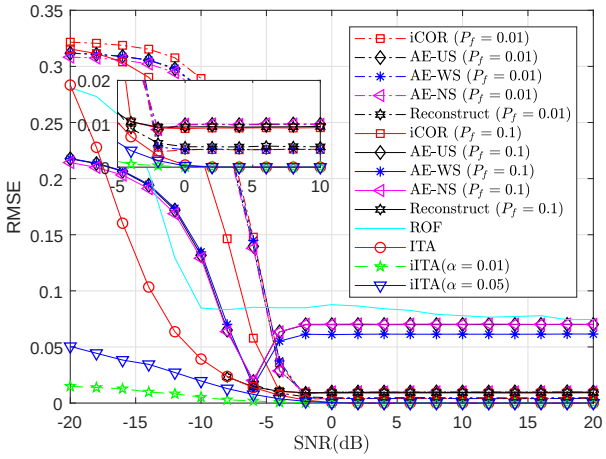


Fig. 13. RMSE of various estimators with different SNR when $u^* = 0.3$, $K = 100$ and $N = 1000$.

Fig. 13 shows the RMSE of the estimated COR versus SNR for various estimators. In the simulations, $u^* = 0.3$, $K = 100$ and $N = 1000$, and in each Monte Carlo trial, the initial value of $u^{(0)}$ is independently and randomly drawn from an uniform distribution within the range of $(0, 1)$. For the improved ITA (iITA), we set $\alpha = 0.01$ and 0.05, $N_0 = 1000$, $N_+ = 100$ and we also set a lower bound for N as $N_{\min} = 100$. In Fig. 13, the RMSE of AE-US [17], AE-NS [18] and AE-WS [18] with $P_f = 0.1$ increases when $\text{SNR} > -6\text{dB}$. This is because $\hat{u}$ of AE-US, AE-NS or AE-WS with $P_f = 0.1$ is already overestimated at $\text{SNR} = -6\text{dB}$, and increased SNR (leading to higher P_d) makes the overestimation more severe. It can be observed that iCOR [2] has better performance than AE-US, AE-NS and AE-WS only when SNR is high, and this observation is the same as in [2]. It can be also observed that the reconstruction algorithm [12], [21] does not work when $\text{SNR} < -8\text{dB}$ as the constructed false-alarm probability becomes negative. ITA has better performance than the others except the iITA when $\text{SNR} > -16\text{dB}$. When $\text{SNR} < -16\text{dB}$, the performance of ITA degrades. This is because the range of possible convergence values is large at low SNRs, and uniform selection of initial value leads to performance degradation. It can be seen that the RMSE of iITA with $\alpha = 0.01$ and $\alpha = 0.05$ are around 0.01 and 0.05 respectively at low SNR.

With the same parameter setting as in Fig. 13, Fig. 14 shows the bias of the estimated COR versus SNR for various estimators. It can be observed that, with the increase of the SNR, the COR is firstly underestimated and then overestimated by these estimators except ITA and iITA, while the bias of ITA and iITA approaches zero. iCOR has smaller bias than that of AE-US, AE-NS and AE-WS when $\text{SNR} > -5\text{dB}$, while AE-US, AE-NS and AE-WS have smaller bias than iCOR when $\text{SNR} < -5\text{dB}$. Overestimation is observed at low SNR for ITA, because the randomly initialized value $u^{(0)}$ is larger than u^* with a high probability. Also we can see that iITA has smaller absolute bias than ITA. When $\text{SNR} > -10\text{dB}$, it can be observed that the bias of ITA and iITA is close to zero. Moreover, it can be obtained from Fig. 14 that ITA and iITA

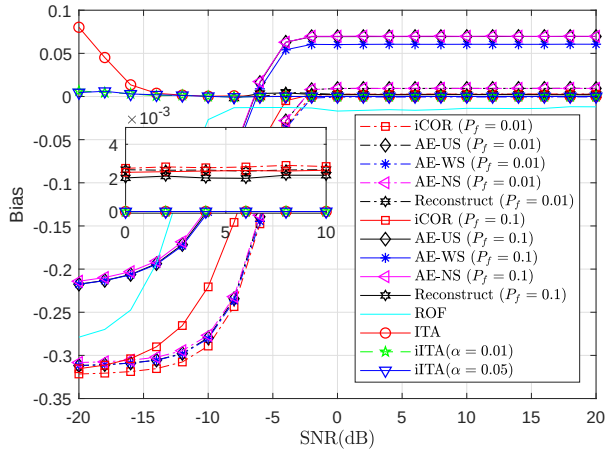


Fig. 14. Bias of various estimators with different SNR when $u^* = 0.3$, $K = 100$ and $N = 1000$.

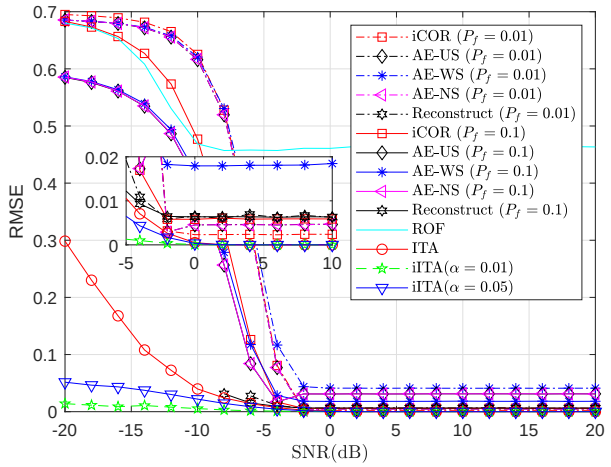


Fig. 15. RMSE of various estimators with different SNR when $u^* = 0.7$, $K = 100$ and $N = 1000$.

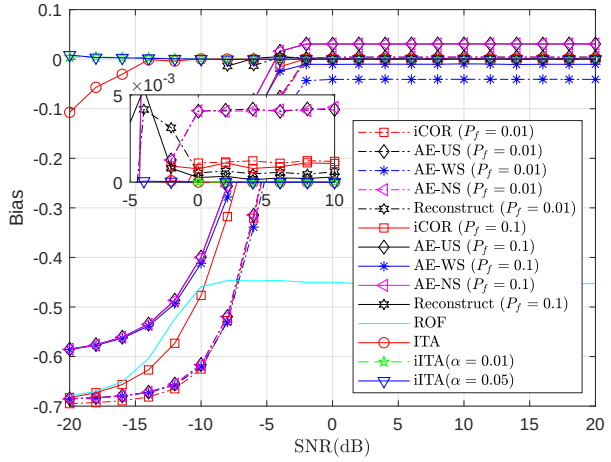


Fig. 16. Bias of various estimators with different SNR when $u^* = 0.7$, $K = 100$ and $N = 1000$.

have the smallest bias (absolute value) among the estimators.

Fig. 15 shows the RMSE of the estimated COR versus SNR for various estimators for $u^* = 0.7$ when $K = 100$ and $N = 1000$. The parameters of iITA are the same as those

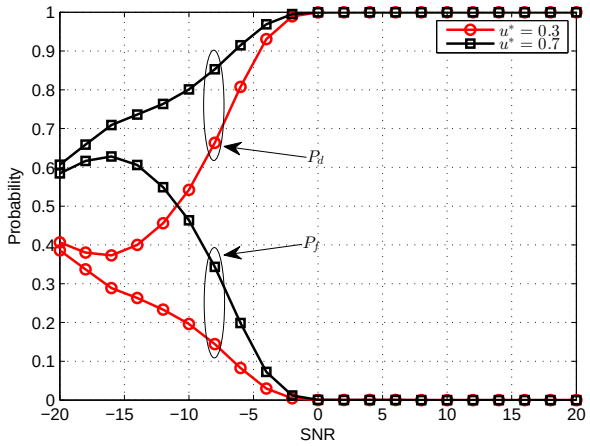


Fig. 17. Achieved P_f and P_d of the proposed method at different SNR for $u^* = 0.3$ and $u^* = 0.7$ when $K = 100$ and $N = 1000$.

in Fig. 13. In each Monte Carlo trial, the initial value of $u^{(0)}$ is independently and randomly drawn from an uniform distribution with the range of $(0, 1)$. In the simulation results, the RMSE of AE-US, AE-NS and AE-WS with $P_f = 0.1$ increases when SNR > -4dB. This is because $\hat{u}$ of AE-US, AE-NS or AE-WS with $P_f = 0.1$ is already overestimated at SNR $= -4$dB, and increased SNR (leading to higher P_d) leads to a more severe overestimation. It can be observed that iCOR has better performance than AE-US, AE-NS and AE-WS only when SNR is high. ITA method has better performance than the others. While the performance of ITA is not very good when SNR is low, and from this figure and Fig. 13 we can see that iITA greatly improves the performance at low SNRs by reducing α.

With the same parameter setting as in Fig. 15, Fig. 16 shows the bias of the estimated COR of $u^* = 0.7$ versus SNR for various estimators. It can be observed again that, while SNR increases, other methods experience firstly underestimation and then overestimation. In contrast, the bias of ITA approaches zero. Underestimation occurs at low SNR for ITA, because the randomly initial value $u^{(0)}$ is smaller than u with a high probability. Also, iITA delivers better permanence than ITA when SNR is very low.

Fig. 17 shows the achieved P_f and P_d of ITA at different SNRs for $u^* = 0.3$ and $u^* = 0.7$ when $K = 100$ and $N = 1000$. It can be observed that $P_f = 0$ and $P_d = 1$ when SNR$\geq$ 0dB. This explains that the RMSE of ITA is close to zero when SNR$\geq$ 0dB, while the RMSEs of other methods cannot approach zero as they have a constant nonzero P_f.

Fig. 18 shows the RMSE for various estimators with different N at SNR $= -10$dB. In the simulations, $u^* = 0.3$, $K = 100$, $N_0 = 1000$, $N_+ = 100$ and $N_{\min} = 100$. For iITA, the RMSE is obtained with with different N_0. Considering that the iCOR with $P_f = 0.01$ does not work for $N < 100$, we only evaluate it for $N \geq 100$. It can be seen that AE-WS, AE-NS and AE-US have similar performance when N is large. The performance of all estimators improves with N except iITA, and the performance gap between other methods and ITA increases with N. From Fig. 18 we can see that whatever

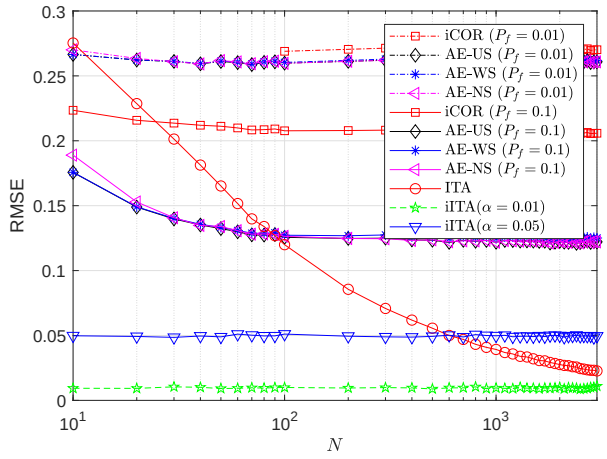


Fig. 18. RMSE for various estimators with different N at SNR $= -10$dB when $u^* = 0.3$ and $K = 100$.

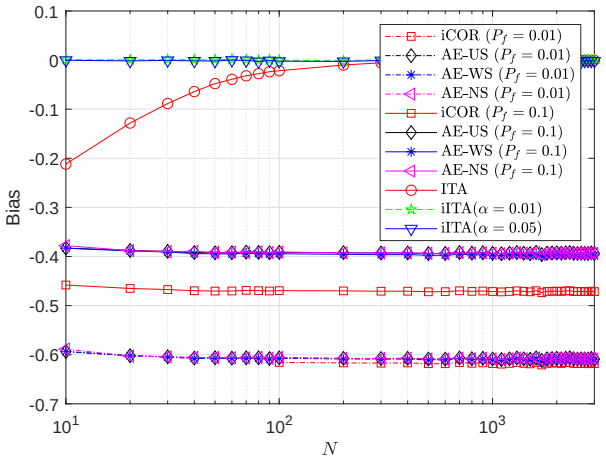


Fig. 21. Bias for various estimators with different N at SNR $= -10$dB when $u^* = 0.7$ and $K = 100$.

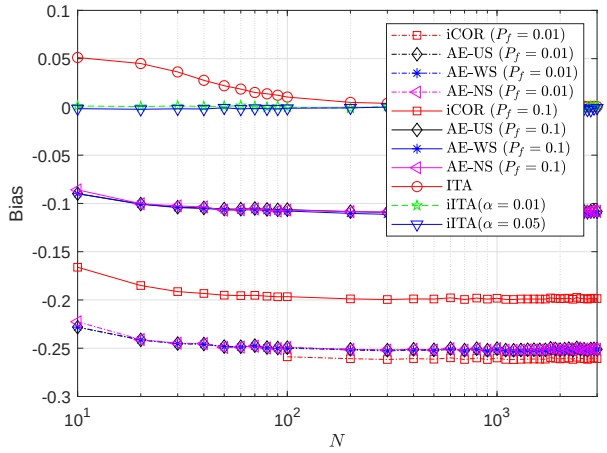


Fig. 19. Bias for various estimators with different N at SNR $= -10$dB when $u^* = 0.3$ and $K = 100$.

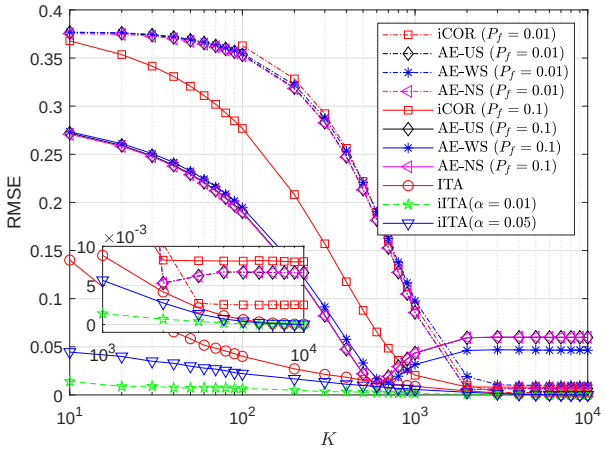


Fig. 22. RMSE for various estimators with different K at SNR $= -10$dB when $u^* = 0.3$ and $N = 1000$.

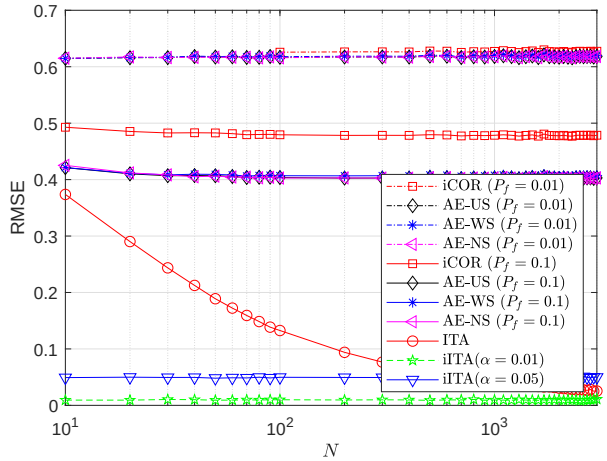


Fig. 20. RMSE for various estimators with different N at SNR $= -10$dB when $u^* = 0.7$ and $K = 100$.

N_0 is, the RMSE of iITA with $\alpha = 0.01$ and $\alpha = 0.05$ is about 0.01 and 0.05 respectively. This is because we have preset tolerable estimation errors and iITA can automatically adjust the traffic sample size to achieve the expected performance.

With the same parameter settings, Fig. 19 shows the bias for various estimators with different N at SNR $= -10$dB. It can be observed that the bias of ITA and iITA approaches zero quickly, while this is not the case for other methods. Moreover, the absolute bias of the other methods increases slightly when N is small.

Fig. 20 shows the RMSE for various estimators with different N at SNR $= -10$dB when $u^* = 0.7$, $K = 100$, $N_0 = 1000$, $N_+ = 100$ and $N_{\min} = 100$. For iITA, the RMSE is obtained with different N_0. Considering that iCOR with $P_f = 0.01$ does not work for $N < 100$, we only evaluate it for $N \geq 100$. It can be seen from Fig. 20 that AE-WS, AE-NS and AE-US have similar performance when N is large. The performance of all estimators improves with N except iITA, and the performance of ITA improves faster than other methods. For iITA, the same conclusion as those for Fig. 18 can be drawn. The corresponding results for the bias of these estimators are presented in Fig. 21. The same conclusion as from Fig. 19 can be drawn from Fig. 21.

Fig. 22 shows the RMSE for various estimators with different K at SNR $= -10$dB when $u^* = 0.3$, $N = 1000$, $N_0 = 1000$, $N_+ = 100$ and $N_{\min} = 100$, and Fig. 23 shows the

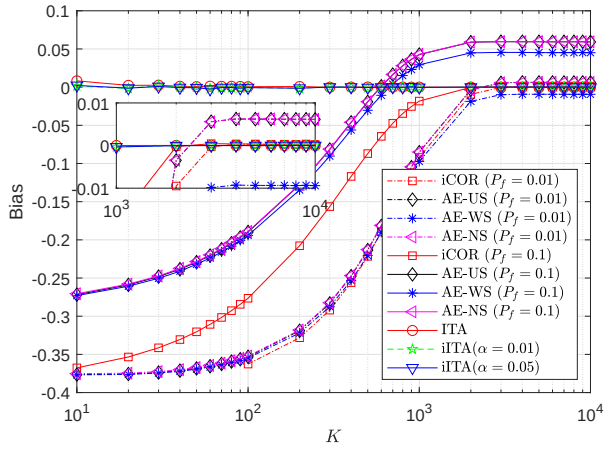


Fig. 23. Bias for various estimators with different K at SNR $= -10$dB when $u^* = 0.3$ and $N = 1000$.

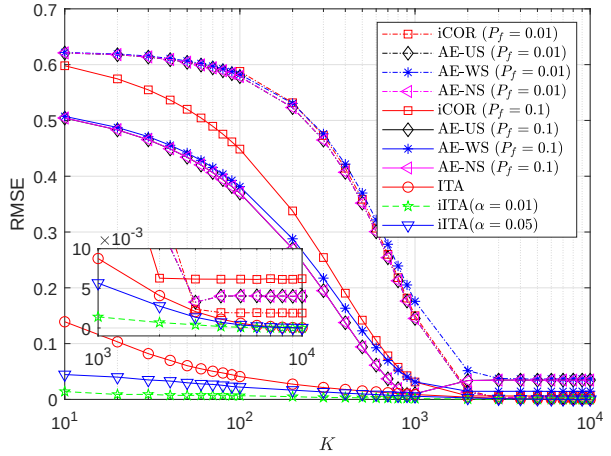


Fig. 24. RMSE for various estimators with different K at SNR $= -10$dB when $u^* = 0.7$ and $N = 1000$.

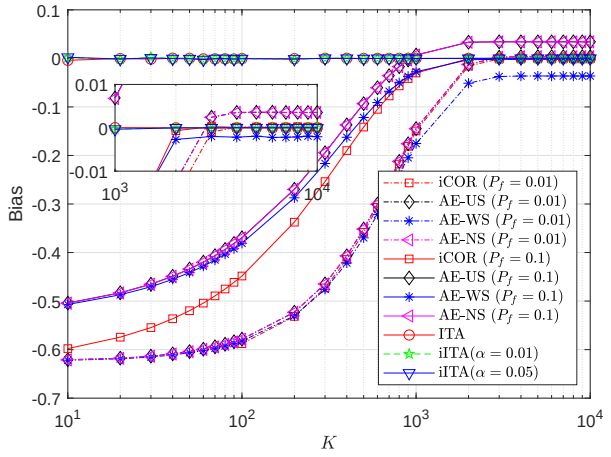


Fig. 25. Bias for various estimators with different K at SNR $= -10$dB when $u^* = 0.7$ and $N = 1000$.

bias for various estimators with different K. It can be seen that the RMSEs of AE-WS, AE-NS and AE-US with P_f decrease and then increase with K, since the COR is underestimated with small K but overestimated with large K (as shown in Fig. 23). Fig. 22 shows that the RMSEs of other methods decrease with K when $K \leq 1000$, and ITA has smaller RMSE than other methods. The RMSE of iITA with $\alpha = 0.01$ and $\alpha = 0.05$ is around 0.01 and 0.05 respectively when K is small and then decreases with K. Fig. 23 shows that the bias of ITA and iITA goes to zero with K.

Fig. 24 shows the RMSE for various estimators with different K at SNR $= -10$dB when $u^* = 0.7$, $N = 1000$, $N_0 = 1000$, $N_+ = 100$ and $N_{\min} = 100$, and Fig. 25 shows the bias for various estimators with different K. It can be seen that the RMSEs of other methods decrease with K when $K \leq 1000$, and ITA has smaller RMSE, and for iITA, the same conclusion as those for Fig. 22 can be drawn. Fig. 25 shows that the bias of ITA and iITA goes to zero with K. In contrast, AE-US, AE-NS and AE-WS overestimate COR if K increases further.

VI. Conclusion

In this paper, we have given an thorough analysis of the performance of COR estimation with iterative threshold setting. We have proved the convergence of ITA in the case with a sufficiently large number of traffic samples, and investigated its convergence mechanism with small number of traffic samples. We also analyze the upper bound on of the number of traffic numbers that is required to achieve a pre-set estimation accuracy. To improve its performance at low SNRs, we have proposed an improved estimator. Numerical results have shown that ITA delivers significantly better performance than the state-of-the-art ones, and the improved ITA can deliver better performance even at low SNRs by adaptively adjusting the pre-set estimation accuracy.

In some scenarios, the minimum length of the holding time of channel states may be several time-slots [12], [21]–[23]. This feature has not been exploited by both ITA and iITA. Incorporating this feature of channel states will further improve the estimation performance of COR, which is an interesting direction to explore.

Appendix A
Convergence for the case of $u^{(0)} < u^*$

Without loss of generality, we let

$$u^{(l-1)} < u^* \tag{A.1}$$

for an arbitrary $l \geq 1$. Then, we can obtain $\lambda^{(l)}$ by solving

$$g(\lambda) = f(u^{(l-1)}). \tag{A.2}$$

Considering that $g(\lambda)$ is a monotonically increasing function, $f(u)$ is a monotonically decreasing function and $u^{(l-1)} < u^*$, we can obtain that

$$\lambda^{(l)} > \lambda^*. \tag{A.3}$$

Hence, $\lambda^{(l)}$ will produce an underestimated COR $u^{(l)}$, i.e.,

$$u^{(l)} < u^*. \tag{A.4}$$

Next, we prove that $u^{(l)}$ increases with respect to l. Note that the average power of received signals at an SU is statistically smaller under $\mathcal{H}_0$ than $\mathcal{H}_1$. Hence, we can obtain that

$$P_d(\lambda^{(l)}) > P_f(\lambda^{(l)}). \tag{A.5}$$

From (18) and (20), we can obtain that

$$u^{(l-1)} - u^{(l)} = (u^{(l-1)} - u^*) \left(P_d(\lambda^{(l)}) - P_f(\lambda^{(l)}) \right). \quad \text{(A.6)}$$

Note that the right-hand side of (A.6) is smaller than zero. Thus, we have

$$u^{(l-1)} < u^{(l)} \quad \text{(A.7)}$$

for $u^{(l-1)} < u^*$.

Combing (A.4) and (A.7) gives

$$u^{(l-1)} < u^{(l)} < u^*. \quad \text{(A.8)}$$

This implies that $u^{(l)}$ increases with the iteration number l, and it converges to some value, and

$$\lim_{l \to +\infty} \left(u^{(l)} - u^{(l-1)} \right) = 0. \quad \text{(A.9)}$$

As

$$P_d(\lambda^{(l)}) - P_f(\lambda^{(l)}) > 0, \quad \text{(A.10)}$$

according to (A.6), we can obtain that

$$\lim_{l \to +\infty} \left(u^* - u^{(l-1)} \right) = 0. \quad \text{(A.11)}$$

which indicates that $u^{(l)}$ converges to u^*. Then, $\lambda^{(l)}$ converges to λ^*.

REFERENCES

[1] Q. Zhao and B. M. Sadler, "A survey of dynamic spectrum access," *IEEE Signal Process. Mag.*, vol. 24, no. 3, pp. 79–89, May 2007.

[2] J. J. Lehtomäki, M. López-Benítez, K. Umebayashi and M. Juntti, "Improved channel occupancy rate estimation," *IEEE Trans. Commun.*, vol. 63, no. 3, pp. 643–654, Mar. 2015.

[3] M. M. Rashid, M. J. Hossain, E. Hossain, and V. K. Bhargava, "Opportunistic spectrum scheduling for multiuser cognitive radio: A queueing analysis," *IEEE Trans. Wireless Commun.*, vol. 8, no. 10, pp. 5259–5269, Oct. 2009.

[4] M. A. Al-Jarrah, A. Al-Dweik, S. S. Ikki and E. Alsusa, "Spectrum-occupancy aware cooperative spectrum sensing using adaptive detection," *IEEE Systems Journal*, vol. 14, no. 2, pp. 2225–2236, June. 2020.

[5] H. Jiang, L. Lai, R. Fan, and H. V. Poor, "Optimal selection of channel sensing order in cognitive radio," *IEEE Trans. Wireless Commun.*, vol. 8, no. 1, pp. 297–307, Jan. 2009.

[6] B. Soni, D. K. Patel, Z. G. Ding, Y. L. Guan and S. Sun, "On sensing performance of multi-antenna mobile cognitive radio conditioned on primary user activity statistics," *IEEE Trans. Wireless Commun.*, vol. 21, no. 5, pp. 3381–3394, May 2022.

[7] K. Umebayashi, K. Hayashi and J. J. Lehtomäki, "Threshold-setting for spectrum sensing based on statistical information," *IEEE Commun. Lett.*, vol. 21, no. 7, pp. 1585–1588, Jul. 2017.

[8] D. K. Patel, S. Kavaiya, Z. G. Ding, Y. L. Guan and S. Sun, "Impact of primary user activity statistics in cognitive vehicular networks," *IEEE Trans. Veh. Technol.*, vol. 71, no. 3, pp. 2859–2873, Mar. 2022.

[9] X. Y. Wang and P. H. Ho, "Gossip-enabled stochastic channel negotiation for cognitive radio ad hoc networks," *IEEE Trans. Mobile Comput.*, vol. 10, no. 11, pp. 1632–1645, Nov. 2011.

[10] M. López-Benítez, A. Al-Tahmeesschi, D. K. Patel, J. Lehtomäki and K. Umebayashi, "Estimation of primary channel activity statistics in cognitive radio based on periodic spectrum sensing observations," *IEEE Trans. Wireless Commun.*, vol. 18, no. 2, pp. 983–996, Feb. 2019.

[11] A. Al-Tahmeesschi, M. López-Benítez, D. K. Patel, J. Lehtomäki and K. Umebayashi, "On the sample size for the estimation of primary activity statistics based on spectrum sensing," *IEEE Trans. Cognitive Commun. Network.*, vol. 5, no. 1, pp. 59–72, Mar. 2019.

[12] O. H. Toma, M. López-Benítez, D. K. Patel and K. Umebayashi, "Estimation of primary channel activity statistics in cognitive radio based on imperfect spectrum sensing," *IEEE Trans. Commun.*, vol. 68, no. 4, pp. 2016–2031, Apr. 2020.

[13] C. Liu and D. Cabric, "Prediction of Erlang-2 distributed primary user traffic for dynamic spectrum access," *IEEE Wireless Commun. Lett.*, vol. 4, no. 5, pp. 481–484, Oct. 2015.

[14] M. Wellens, J. Riihijärvi and P. Mähönen, "Empirical time and frequency domain models of spectrum use," *Phys. Commun.*, vol. 2, no. 1/2, pp. 10–32, Mar. 2009.

[15] R.A. Halaseh and D. Dahlhaus, "Continuous hidden Markov model based interference aware cognitive radio spectrum occupancy prediction," in *IEEE 27th Ann. Int. Symposium on Personal, Indoor, and Mobile Radio Commun.*, Valencia, 2016, pp. 1–6.

[16] S. Geirhofer, L. Tong and B. M. Sadler, "A measurement-based model for dynamic spectrum access in WLAN channels," in *IEEE Mil. Commun. Conf.*, Washington, DC, 2006, pp. 1-7.

[17] H. Kim and K. G. Shin, "Efficient discovery of spectrum opportunities with MAC-layer sensing in cognitive radio networks," *IEEE Trans. Mobile Comput.*, vol. 7, no. 5, pp. 533–545, May 2008.

[18] W. Gabran, C. H. Liu, P. Pawelczak and D. Cabric, "Primary user traffic estimation for dynamic spectrum access," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 3, pp. 544–558, Mar. 2013.

[19] J. J. Lehtomäki, R. Vuohtoniemi and K. Umebayashi, "On the measurement of duty cycle and channel occupancy rate," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 11, pp. 2555–2565, Nov. 2013.

[20] J. Nikonowicz, A. Mahmood and M. Gidlund, "A blind signal samples detection algorithm for accurate primary user traffic estimation," *Sensors*, vol. 20, no. 15, pp. 1–11, Jul. 2020.

[21] O. H. Toma, M. López-Benítez, D. K. Patel and K. Umebayashi, "Reconstruction algorithm for primary channel statistics estimation under imperfect spectrum sensing," in *IEEE Wireless Commun. Network. Conf. (WCNC)*, Seoul, 2020, pp. 1–5.

[22] O. H. Toma and M. López-Benítez, "Traffic learning: a deep learning approach for obtaining accurate statistical information of the channel traffic in spectrum sharing systems," *IEEE Access*, vol. 9, pp. 124324–124336, Sep. 2021.

[23] M. López-Benítez, O. H. Toma, D. K. Patel and K. Umebayashi, "Methods for fast estimation of primary activity statistics in cognitive radio systems," in *IEEE Wireless Commun. Network. Conf. (WCNC)*, Seoul, 2020, pp. 1–5.

[24] M. López-Benítez and J. Lehtomäki, "Energy detection based estimation of primary channel occupancy rate in cognitive radio," in *IEEE Wireless Commun. Network. Conf. (WCNC)*, Doha, 2016, pp. 1–6.

[25] M. Jin, Q. Guo, J. Xi, Y. M. Li and Y. H. Li, "On spectrum sensing of OFDM signals at low SNR: New detectors and asymptotic performance," *IEEE Trans. Signal Process.*, vol. 65, no. 12, pp. 3218–3233, Jun. 2017.

[26] L. Huang, C. Qian, Y. Xiao and Q. T. Zhang, "Performance analysis of volume-based spectrum sensing for cognitive radio," *IEEE Trans. Wireless Commun.*, vol. 14, no. 1, pp. 317–330, Jan. 2015.

[27] L. Huang, Y. Xiao and Q. T. Zhang, "Robust spectrum sensing for noncircular signal in multiantenna cognitive receivers," *IEEE Trans. Signal Process.*, vol. 63, no. 2, pp. 498–511, Jan. 2015.

[28] P. B. Gohain, S. Chaudhari and V. Koivunen,"Cooperative energy detection with heterogeneous sensors under noise uncertainty: SNR wall and use of evidence theory," *IEEE Trans. Cognitive Commun. Network.*, vol. 4, no. 3, pp. 473–485, Sept. 2018.

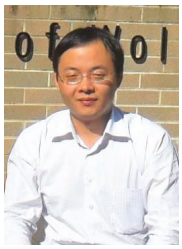
[29] C. Bishop, Pattern recognition and machine learning, New York, Springer, 2006.

[30] M. S. Bazaraa, H. D. Sherali, C. M. Shetty,"Nonlinear programming: Theory and algorithms," 3rd Ed., Hoboken, NJ: John Wiley and Sons Inc., 2006.

[31] G. E. P. Box, W. G. Hunter, J. S. Hunter,"Statistics for experimenters: Design, innovation and discovery," 2nd Ed., Chicester, UK: John Wiley and Sons Ltd, 2005.

Ming He received the B.E. degree in telecommunications engineering, in 2017, from Jiangxi University of Science and Technology, Ganzhou, China. He received the M.S. degree in electronic engineering from Ningbo University, Ningbo, China, where he is currently working toward the Ph.D. degree in telecommunications engineering. His research interests include parameter estimation and spectrum access.

Ming Jin received the B.E. degree in electronic engineering and the Ph.D. degree in signal and information processing from Xidian University, Xi'an, China, in 2005 and 2010, respectively. From 2013 to 2014, he was an Associate Researcher in the School of Electrical, Computer and Telecommunications Engineering, University of Wollongong, Wollongong, NSW, Australia. He is currently a Professor in the Faculty of Electrical Engineering and Computer Science, Ningbo University, Ningbo, China. His current research interests include cognitive radio, resource optimization and machine learning.



Qinghua Guo (S'07CM'08CSM'18) received the B.E. degree in electronic engineering and the M.E. degree in signal and information processing from Xidian University, Xi'an, China, in 2001 and 2004, respectively, and the Ph.D. degree in electronic engineering from the City University of Hong Kong, Hong Kong, China, in 2008. He is currently an Associate Professor with the School of Electrical, Computer and Telecommunications Engineering, University of Wollongong, Wollongong, NSW, Australia, and an Adjunct Associate Professor with the School of Engineering, The University of Western Australia, Perth, WA, Australia. His research interests include signal processing, telecommunications, radar and optical sensing. He was a recipient of the Australian Research Council's inaugural Discovery Early Career Researcher Award in 2012.

Miguel López-Benítez (Senior Member, IEEE) received the BSc and MSc degrees (both with Distinction) from Miguel Hernández University, Elche, Spain in 2003 and 2006, respectively, and the PhD degree from the Technical University of Catalonia, Barcelona, Spain in 2011. From 2011 to 2013, he was a Research Fellow with the Centre for Communication Systems Research, University of Surrey, Guildford, UK. In 2013, he became a Lecturer (Assistant Professor) with the Department of Electrical Engineering and Electronics, University of Liverpool, UK, where he has been a Senior Lecturer (Associate Professor) since 2018. His research interests are in the field of wireless communications.

Wei Zhang (S'01-M'06-SM'11-F'15) received the Ph.D. degree from The Chinese University of Hong Kong in 2005. Currently, he is a Professor at the School of Electrical Engineering and Telecommunications, the University of New South Wales, Sydney, Australia. His current research interests include UAV communications, 5G and beyond. He received 6 best paper awards from IEEE conferences and ComSoc technical committees. He was elevated to Fellow of the IEEE in 2015 and was an IEEE ComSoc Distinguished Lecturer in 2016-2017. He is Vice President of IEEE Communications Society.

Within the IEEE ComSoc, he has taken many leadership positions including Member-at-Large on the Board of Governors (2018-2020), Chair of Wireless Communications Technical Committee (2019-2020), Vice Director of Asia Pacific Board (2016-2021), Editor-in-Chief of IEEE Wireless Communications Letters (2016-2019), Technical Program Committee Chair of APCC 2017 and ICCC 2019, Award Committee Chair of Asia Pacific Board and Award Committee Chair of Technical Committee on Cognitive Networks.

In addition, he has served as a member in various ComSoc boards/standing committees, including Journals Board, Technical Committee Recertification Committee, Finance Standing Committee, Information Technology Committee, Steering Committee of IEEE Transactions on Green Communications and Networking and Steering Committee of IEEE Networking Letters. Currently, he serves as an Area Editor of the IEEE Transactions on Wireless Communications and the Editor-in-Chief of Journal of Communications and Information Networks. Previously, he served as Editor of IEEE Transactions on Communications, IEEE Transactions on Wireless Communications, IEEE Transactions on Cognitive Communications and Networking, and IEEE Journal on Selected Areas in Communications C Cognitive Radio Series.