Learning Multiple Primary Transmit Power Levels for Smart Spectrum Sharing

Rui Zhang and Peng Cheng
School of Electrical and Information Engineering
The University of Sydney
Sydney, Australia
{rui.zhang1; peng.cheng}@sydney.edu.au

Zhuo Chen
DATA61
CSIRO
Sydney, Australia
zhuo.chen@csiro.au

Yonghui Li and Branka Vucetic
School of Electrical and Information Engineering
The University of Sydney
Sydney, Australia
{yonghui.li; branka.vucetic}@sydney.edu.au

***Abstract*—Multi-parameter cognition in a cognitive radio network provides a potential avenue to more efficient spectrum usage. In this paper, we propose a two-stage spectrum sharing strategy, where the primary user operates with multiple transmit power levels. Different from the conventional approaches, our method does not require any prior knowledge of the primary transmitter (PT) power characteristics. In the first stage, we use a conditionally conjugate Dirichlet process Gaussian mixture model to capture the multi-level power characteristics inherent in the PT signals, and design a Bayesian inference method to infer the model parameters. In the second stage, we propose a secondary transmitter (ST) prediction-transmission method based on reinforcement learning, which adapts to the PT power variation and strike an excellent tradeoff between the secondary network throughput and the interference to the primary network. The simulation results show the effectiveness of the proposed strategy.**

I. Introduction

The emerging new wireless technologies are fueling an ever-increasing demand for access to the radio frequency spectrum. Cognitive radio (CR) opens a potential communication paradigm to achieve more efficient and flexible spectrum use [1]. A secondary user (SU) with CR capability monitors the spectrum utilization of a primary user (PU) and determines its access to such spectrum accordingly. The SU first senses the surrounding radio spectrum state based on various signal processing methods, where energy detection [2] features low computational and implementable complexity, and is widely adopted. Upon obtaining the radio spectrum state, the SU accesses the licensed spectrum either when the PU is idle [3], or concurrently with the PU following a power control strategy to constrain the interference to the PU [4].

It is worth noting that many contemporary wireless standards, such as IEEE 802.11, GSM, and LTE, have specified multiple transmit power levels to dynamically adapt to the fast changing environment and varying quality of service. However, only limited studies in the literature took this into account, while the majority assume the SU adopts a binary approach in reporting the radio spectrum state as idle or busy. In [5] and [6], the authors proposed an energy detection based multiple hypothesis test to derive the decision thresholds for the multiple power level identification. The results were extended to the scenarios with noise uncertainty [7] and non-Gaussian transmission signals [8]. However, all the methods in [5]–[8] made a fundamental assumption that the ST has the full prior knowledge of the PT transmit power mode, which are unlikely to be available to the ST *a prior*.

In this paper, we propose a two-stage spectrum sharing strategy. In the first stage, based on machine learning theory, we propose a blind spectrum learning method to model the received PT signals at the ST and discover their latent patterns reflecting the power variation. It is blind in the sense that the ST does not require any prior knowledge of the PT transmit power mode. In the second stage, we propose a secondary transmitter (ST) prediction-transmission method, where the prediction part in our method can identify the current PT power level through collecting PT signal samples, and the transmission part adjusts the ST power accordingly. Furthermore, we dynamically determine the interval between predictions, which can be formulated as a partial observable decision problem.

The main contributions of this paper are summarized as follows. We propose a two-stage spectrum sharing strategy for a CR network, enabling the spectrum access when the PT power varies with time in multiple levels. We propose to utilize the Gaussian mixture models to capture the multiple signal levels received by the ST, and carry out the Bayesian inference to estimate the model parameters. We propose a prediction-transmission method for the spectrum access which enables the ST to closely adapt to the PT power variation.

II. System Model

We consider a spectrum sharing CR network with a primary network consisting of a PT and a primary receiver (PR), and a secondary network consisting of a ST and a secondary receiver (SR). Transmission happens simultaneously in both networks sharing the same frequency band. It is assumed that there is no information exchange between the primary and secondary networks. The PT operates with multiple power levels, and the ST attempts to learn the PT transmit power mode and then optimize the spectrum access accordingly.

We propose a novel two-stage spectrum sharing strategy, as illustrated in Fig. 1. Let $\{P_l, l = 1, \cdots, L\}$ be the transmit powers of the PT, where $P_1 < \cdots < P_{L-1}$ and $P_L = 0$ indicates an idle PT. For convenience, hypothesis $\mathcal{H}_l$ indicates that the PT transmits with power P_l. We define the time duration of each hypothesis as a random variable

978-1-5386-8088-9/19/$31.00 ©2019 IEEE

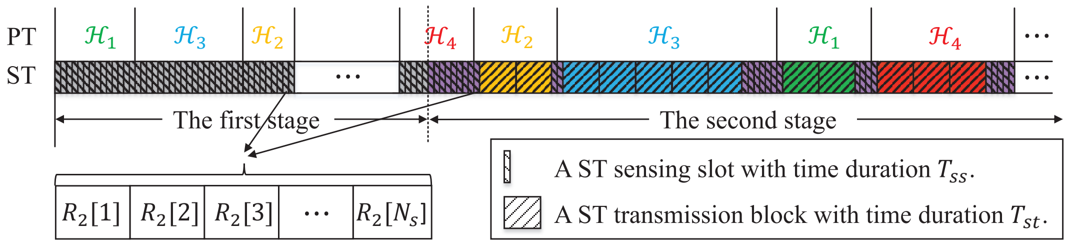


Fig. 1. The proposed two-stage spectrum sharing strategy. The sensing slots of the ST in both stages have the same time duration T_{ss}. The sensing slots in the first stage are used for learning, while that in the second stage are for prediction.

T_{po}, which is usually much larger than that of the ST sensing slot T_{ss} and the ST transmission block T_{st}. In this paper, we consider a time discretized model with a minimum time unit T_{ss}. We denote $\tau_s = T_{st}/T_{ss}$ as the fixed discretized time duration of the ST transmission block, and $\tau_p = T_{po}/T_{ss}$ as the varied discretized PT power level duration. The prior knowledge on the PT transmit power mode, defined as the number of transmit powers L, the exact values P_l, and the prior probability of each hypothesis $\Pr\{\mathcal{H}_l\}$, are unknown to the ST.

In the first stage, the ST samples the received PT signals at a sampling frequency f_s and collects N_s samples in each sensing slot with duration T_{ss}. The ST observes N sensing slots in the first stage and collects a total of NN_s samples. It is assumed that the learning period is reasonably large so that it covers all L hypotheses. Thus, the i-th sample in the n-th sensing slot under hypothesis $\mathcal{H}_l$ can be given by

$$R_n[i] = \sqrt{P_l} s_n[i] + u_n[i],\ \mathcal{H}_l, i = 1, \cdots, N_s; n = 1, \cdots, N, \tag{1}$$

where $\sqrt{P_l} s_n[i]$ is the received primary signal in the n-th sensing slot with average power P_l. Following [9], we assume that $s_n[i]$ is an independent and identical distributed (i.i.d.) random process with mean 0 and variance 1. Without loss of generality, we assume that $s_n[i]$ is a complex PSK modulated signal [9]. In (1), $u_n[i] \sim \mathcal{CN}(0, \sigma_u^2)$ is the additive white Gaussian noise with mean 0 and precision σ_u^2. The test statistic in the n-th sensing slot can be calculated as

$$X_n = \frac{1}{N_s} \sum_{i=1}^{N_s} |R_n[i]|^2, \quad \mathcal{H}_l. \tag{2}$$

When N_s is large, according to the central limit theorem, the distribution of X_n under hypothesis $\mathcal{H}_l$ can be approximated by a Gaussian one, and we have

$$X_n \sim \mathcal{N}\left((\gamma_{st}^l + 1)\sigma_u^2, \frac{1}{N_s}(2\gamma_{st}^l + 1)\sigma_u^4\right), \quad \mathcal{H}_l, \tag{3}$$

where $\gamma_{st}^l = P_l/\sigma_u^2$ is the received signal-to-noise ratio (SNR) at the ST and $\mathcal{N}(\mu, S^{-1})$ denotes the Gaussian distribution with mean μ and precision S. Considering all the hypotheses, we establish that X_n follows a mixed Gaussian distribution

$$X_n \sim \sum_{l=1}^{L} \pi_l \mathcal{N}(\mu_l, S_l^{-1}), \tag{4}$$

where $0 \leqslant \pi_l \leqslant 1$ is the mixing coefficients with $\sum_{l=1}^{L} \pi_l = 1$. In each Gaussian density, $\mathcal{N}(\mu_l, S_l^{-1})$ is a component of the mixture with mean value $\mu_l = (\gamma_{st}^l + 1)\sigma_u^2$ and precision $S_l = \left(\frac{1}{N_s}(2\gamma_{st}^l + 1)\sigma_u^4\right)^{-1}$.

Given the observation set $\mathbf{X} = \{X_1, \cdots, X_N\}$, the proposed Bayesian nonparametric method aims to infer the Gaussian Mixture Models (GMM) parameter set $\{\boldsymbol{\theta}, \boldsymbol{\pi}, L\}$, where $\boldsymbol{\theta} = \{\theta_1, \cdots, \theta_L\}$, $\theta_l = \{\mu_l, S_l\}$ and $\boldsymbol{\pi} = \{\pi_1, \cdots, \pi_L\}$.

For the SU, there is a fundamental tradeoff between two conflicting goals, namely, maximization of its own throughput and minimization of its interference to the PU. It is extremely difficult to optimize this tradeoff in practice when there is no information exchange or cooperation between the PU and SU. To provide a pragmatic solution to this dilemma, we propose a new metric, referred to as the normalized power level alignment (NPLA), which is defined as the time proportion that the ST matches its transmit power level to that of the PT.

Therefore, in the second stage, we propose a prediction-transmission method adapting to the PT power level variation. As shown in Fig. 1, two actions exist: prediction and transmission. In the prediction, the ST can easily identify the current PT power level l, which is jointly determined by the test statistic $X_n, n > N$ in (2) and the inferred GMM parameter set $\{\boldsymbol{\theta}, \boldsymbol{\pi}, L\}$. In the transmission, the ST adjusts its transmit power level k to match the latest identified PT power level l ($k = l$). In the proposed non-periodic prediction-transmission method, the intervals are dynamically determined to enhance the NPLA performance. As shown in Fig. 1, zero intervals are used to track the PT power variation, while long intervals are selected to avoid unnecessary prediction.

III. Spectrum Learning Based on Bayesian Inference

In this section, we focus on the first stage, and introduce a Bayesian nonparametric method to infer the GMM parameter set $\{\boldsymbol{\theta}, \boldsymbol{\pi}, L\}$ based on the signal set $\mathbf{X}$. As L is unknown *a priori*, the traditional methods, such as the K-mean and expectation maximization (EM), are inapplicable. This motivates us to resort to Dirichlet process Gaussian mixture model (DPGMM) [10], which is able to identify the unknown number of Gaussian components. For specific

Bayesian inference, we choose the Markov chain Monte-Carlo based Gibbs sampling method [11] for its simplicity.

A. Dirichlet Process Mixture Model

The Dirichlet distribution is an extension of the Beta distribution for multivariate cases. It represents the probability of K events given that the k-th event x_k $(k=1,\cdots,K)$ has been observed $\alpha_k - 1$ times. The probability density function can be expressed as

$$\mathrm{Dir}(\alpha_1,\cdots,\alpha_K) = \frac{\Gamma\left(\sum_{k=1}^{K}\alpha_k\right)}{\prod_{k=1}^{K}\Gamma(\alpha_k)}\prod_{k=1}^{K}\pi_k^{\alpha_k-1}, \tag{5}$$

where $\Gamma(\cdot)$ denotes the Gamma function, π_k is the probability of the k-th event x_k with $\sum_{k=1}^{K}\pi_k = 1$, and $\pi_k > 0$.

In our application, the event x_k represents the k-th possible PT transmit power level, which can not be observed explicitly. Instead, the explicit observation is the test statistic X_n. As X_n is drawn from a distribution based on event x_k, we introduce the DPMM to define the distribution of X_n. Here, X_n can be regarded as an independent draw from the distribution $F(\theta_n)$, where each $\theta_n = \phi_{z_n}$ is an i.i.d. draw from a base probability distribution G_0. We introduce z_n to explicitly indicate which transmit power level that X_n is associated with, and will be referred to as an indicator hereafter. Mathematically, the DPMM can be expressed as

$$\begin{aligned} X_n|\{\boldsymbol{z},\boldsymbol{\phi}\} \sim F(\phi_{z_n}), \quad p(z_n=k)=\pi_k, \\ \boldsymbol{\pi}|\{\alpha,K\} \sim \mathrm{Dir}(\alpha/K,\cdots\alpha/K), \quad \phi_k|G_0 \sim G_0, \end{aligned} \tag{6}$$

where $\boldsymbol{z} = \{z_1,\cdots,z_N\}$ is the set of indicators, $\boldsymbol{\phi} = \{\phi_1,\cdots,\phi_K\}$ is the set of unique values in $\boldsymbol{\theta}$. Hereafter, K refers to the total number of Gaussian components, and each component consists of the observations that are determined by the ST as having the same transmit power level. Let N_k denote the number of observations assigned to the k-th component, then the distribution of the indicators is

$$p(\boldsymbol{z}|\boldsymbol{\pi}) = \prod_{k=1}^{K}\pi_k^{N_k}. \tag{7}$$

Integrate out the mixing proportions of the product of $p(\boldsymbol{\pi}|\alpha)$ and $p(\boldsymbol{z}|\boldsymbol{\pi})$, the prior on $\boldsymbol{z}$ in terms of α is expressed as [10]

$$p(\boldsymbol{z}|\alpha) = \frac{\Gamma(\alpha)}{\Gamma(N+\alpha)}\prod_{k=1}^{K}\frac{\Gamma(N_k+\alpha/K)}{\Gamma(\alpha/K)}. \tag{8}$$

As all the observations are exchangeable, if we assume that $\boldsymbol{z}_{-n} = \{z_1,\cdots,z_{n-1},z_{n+1},\cdots,z_N\}$ has been obtained, the conditional distribution for the individual indicator is given by

$$p(z_n=k|\boldsymbol{z}_{-n},\alpha) = \frac{p(z_n=k,\boldsymbol{z}_{-n}|\alpha)}{p(\boldsymbol{z}_{-n}|\alpha)} = \frac{N_{-n,k}+\alpha/K}{N-1+\alpha}, \tag{9}$$

where $N_{-n,k}$ is the number of samples excluding X_n in the k-th component.

B. Conditionally Conjugate Dirichlet Process Gaussian Mixture Model

Recall that the observation X_n in (4) follows a mixed Gaussian distribution, X_n can be modeled as a DPGMM and expressed as

$$\begin{aligned} X_n|\{\boldsymbol{z},\boldsymbol{\phi}\} \sim \mathcal{N}(\mu_{z_n}, S_{z_n}^{-1}), \quad (\mu_k, S_k^{-1}) \sim G_0, \\ \boldsymbol{\pi}|\{\alpha,K\} \sim \mathrm{Dir}(\alpha/K,\cdots\alpha/K), \quad p(z_n=k)=\pi_k. \end{aligned} \tag{10}$$

In (10), G_0 represents a prior guess of the distributions of μ_k and S_k^{-1} in the DPGMM. In our case, the distribution of G_0 specifies the prior on the mixture Gaussian distributions parameters $\boldsymbol{\mu} = \{\mu_1,\cdots,\mu_K\}$ and $\boldsymbol{S} = \{S_1,\cdots,S_K\}$, which will lead to an undesirable dependence of μ_k on S_k. To remove such dependency, we modify the original conjugate feature in the DPGMM and introduce the conditionally conjugate version of DPGMM (CCDPGMM) which can be given by [12]

$$\mu_k|\{\lambda,R\} \sim \mathcal{N}(\lambda,R^{-1}), \quad S_k|\{\beta,W\} \sim \mathcal{G}(\beta,W^{-1}), \tag{11}$$

where ξ, R, β and W is the hyperparameter for the DPGMM and $\mathcal{G}(a,b)$ denotes the Gamma distribution with shape parameter a and scale parameter b. To complete the CCDPGMM and capture the features inherent in $\boldsymbol{X}$, we impose vague priors for the hyperparameters following [10],

$$\begin{aligned} \lambda \sim \mathcal{N}(\mu_y, S_y^{-1}), \quad R \sim \mathcal{G}(1,S_y), \\ W \sim \mathcal{G}(1,S_y^{-1}), \quad \beta^{-1} \sim \mathcal{G}(1,1), \end{aligned} \tag{12}$$

where the hyperpriors μ_y and S_y refer to the empirical mean and precision of $\mathbf{X}$, respectively.

C. Inference Using Gibbs Sampling

Given the likelihood of μ_k and S_k in (4) and their priors in (11), we can multiply the priors by the likelihood conditioned on $\boldsymbol{z}$ and obtain the conditional posterior distributions of μ_k and S_k. The conditional posteriors of the hyperparameters can be obtained similarly.

To make the CCDPGMM applicable to the scenario with an infinite number of Gaussian components, we let $K \to \infty$ in (9) and the conditional prior reaches the following limits

$$p(z_n=k|\boldsymbol{z}_{-n},\alpha) = \begin{cases} \dfrac{N_{-n,k}}{N-1+\alpha}, & N_{-n,k} > 0, \\ \dfrac{\alpha}{N-1+\alpha}, & N_{-n,k} = 0. \end{cases} \tag{13}$$

Following [11], we employ the auxiliary variable sampling algorithm, which means adding k_0 auxiliary components in each sampling iteration to represent the effect of the auxiliary components. We combine the priors of $\boldsymbol{z}$ with its likelihood given in (9), the conditional posterior can be given by

$$\begin{aligned} &p(z_n=k'|\boldsymbol{z}_{-n},\mu_{k'},S_{k'},\alpha) \\ &= \begin{cases} \dfrac{qN_{-n,k'}\mathcal{N}(X_n|\mu_{k'},S_{k'})}{N-1+\alpha}, & 1 \leqslant k' \leqslant K', \\ \dfrac{q\alpha\mathcal{N}(X_n|\mu_{k'},S_{k'})}{k_0(N-1+\alpha)}, & K' < k' \leqslant K'+k_0, \end{cases} \end{aligned} \tag{14}$$

where k' is the index of unique components in each iteration during the Gibbs sampling algorithm, q is the appropriate constant for normalization, and K' is the number of active components. To this end, we summarize the sampling algorithm in Algorithm 1.

Algorithm 1 Gibbs sampling algorithm.

Require:
Initial observation set $\mathbf{X}$. Set a component which contains all X_n. Initialize the hyperparameters λ, R, β, and W, the hyperpriors μ_y and S_y, and the indicator set $\boldsymbol{z}$;
Ensure:
The sets $\boldsymbol{z}$, $\boldsymbol{\mu}$, and $\boldsymbol{S}$.

```
Update μ and S conditional on the indicator z and hyperparameters λ, R, β and W;
Update the hyperparameters λ, R, β and W conditional on hyperpriors μ_y and S_y;
for n = 1, 2, ··· , N do
    if z_n ≠ z_n' for all n' ∈ {1, ··· , n−1, n+1, ··· , N} then
        Let z_n = K' + 1.
    end if
    Draw μ_k' and S_k' for k' ∈ {K'+1, ··· , K'+k_0} following (11).
    Update the indicator z_n following (14).
    Discard the empty components.
end for
```

IV. Proposed Prediction-Transmission Spectrum Access Method

In this section, we propose a prediction-transmission method for spectrum access. We first introduce the functions of the prediction and transmission parts. Then, we present the details of how to determine the prediction intervals. As directly optimizing the NPLA performance is intractable, we propose to maximize an expected average utility by imposing reward (penalty) for power level match (mismatch).

A. Functions of the Prediction and Transmission Parts

In the prediction part, with the inferred GMM parameters $\{\boldsymbol{\theta}, \boldsymbol{\pi}, L\}$, the ST can easily identify the current PT power level by a single sensing slot with test statistic X_n, based on the following criterion

$$k = \arg \max_{k \in \{1, \cdots, K\}} \pi_k \mathcal{N}(X_n | \mu_k, S_k). \tag{15}$$

In the transmission part, the ST adjusts its transmit power level k to match the PT power level l, which means $k = l$. Note that K is an estimate of L. In the simulation, we find that the CCDPGMM is able to identify L ($K = L$) with a high probability.

B. Non-periodic Prediction-Transmission Method

The essential question of designing a non-periodic prediction-transmission method is how to determine the prediction intervals. Basically, this needs to find out the distribution of the PT power level duration, and the corresponding observation of each action. If the prediction action is taken, the observation will be the PT power level identified from the received test statistic X_n according to (15). If the transmission action is taken, the observation will be a positive or negative acknowledgment (ACK) received by the ST from the SR. Based on these observations, the ST can infer the PT transmit power level, and then dynamically adjust the prediction intervals.

Without loss of generality, the discretized PT power level duration τ_p of all hypotheses is assumed to follow a Poisson distribution with the same mean value ν. Its cumulative distribution function can be given by

$$F_\nu(\tau) = \frac{\Gamma(\tau + 1, \nu)}{\Gamma(\tau)}, \tag{16}$$

where $\Gamma(\cdot, \cdot)$ denotes the incomplete Gamma function and ν is the mean value of the Poisson distribution, which can be estimated in the first stage. If the PT has been keeping the same power level for time τ immediately after a power level change, the probability that the PT will continue staying in the same power level during the following discretized time duration τ_0 can be expressed as

$$g_\tau(\tau_0) = \frac{1 - F_\nu(\tau + \tau_0)}{1 - F_\nu(\tau)}, \tau_0 \geqslant 1. \tag{17}$$

We define an $L \times L$ transition probability matrix $\mathbf{C}$ for the PT transmit power levels. The element $C_{kj}, k, j \in \{1, \cdots, L\}$ of $\mathbf{C}$ refers to the probability that the PT transfers from the k-th to the j-th transmit power level, and is given by

$$C_{kj} = \begin{cases} \dfrac{\pi_j}{1 - \pi_k}, & k \neq j, \\ 0, & k = j. \end{cases} \tag{18}$$

We also define the vector $\boldsymbol{c}_k$ as the k-th row of matrix $\mathbf{C}$.

We also defined an $L \times L$ prediction probability matrix $\mathbf{H}$, with the element $H_{kj} = \Pr\{\hat{\mathcal{H}}_j | \mathcal{H}_k\}, k, j \in \{1, \cdots, L\}$, representing the probability that the PT is operating under hypothesis $\mathcal{H}_k$ while the detection by the ST is in favor of hypothesis $\mathcal{H}_j$. $\hat{\mathcal{H}}_j$ represents that the ST identifies the PT operating in the j-th transmit power level following (15). We also define vector $\boldsymbol{h}_j$ as the j-th column of matrix $\mathbf{H}$.

We denote the prediction action of the ST at time τ as $a_\tau = 0$, and its observation as $O_\tau^E \in \{\mathcal{H}_k\}$. We also denote the transmission action at time τ as $a_\tau = 1$, and its observation as $O_\tau^A \in \{\mathcal{A}(\text{positive ACK}), \mathcal{N}(\text{negtive ACK})\}$. Let $p_{\tau+\tau_0}^k$ denote the conditional probability that the PT keeps operating with the k-th transmit power level at time $\tau + \tau_0$ given $p_0^k = 1$ and $\{a_0, \cdots, a_\tau, O_0, \cdots, O_\tau\}$, where $O_\tau \in \{O_\tau^E, O_\tau^A\}$. Based on Bayesian rule, the probabilities of PT staying in the k-th power level at time $\tau + \tau_0, \tau_0 \in \{1, \tau_s\}$, can be given as follows. When $a_\tau = 0, \tau_0 = 1$, we have

$$p_{\tau+1}^k(O_{\tau+1}^E) = \begin{cases} \dfrac{p_\tau^k g_\tau^E H_{kk}}{p_\tau^k g_\tau^E H_{kk} + (1 - p_\tau^k g_\tau^E) \boldsymbol{c}_k \boldsymbol{h}_k}, & O_{\tau+1}^E = \mathcal{H}_k, \\ \dfrac{p_\tau^k g_\tau^E H_{jk}}{p_\tau^k g_\tau^E H_{jk} + (1 - p_\tau^k g_\tau^E) \boldsymbol{c}_k \boldsymbol{h}_j}, & O_{\tau+1}^E = \mathcal{H}_j, j \neq k, \end{cases} \tag{19}$$

and when $a_\tau = 1, \tau_0 = \tau_s$, we have

$$p_{\tau+\tau_s}^k(O_{\tau+\tau_s}^A) = \begin{cases} \dfrac{p_\tau^k g_\tau^A}{p_\tau^k g_\tau^A + (1 - p_\tau^k g_\tau^A) \sum_{j=1}^{k-1} C_{kj}}, & O_{\tau+\tau_s}^T = \mathcal{A}, \\ 0, & O_{\tau+\tau_s}^T = \mathcal{N}, \end{cases} \tag{20}$$

where we assume the positive and negative ACKs from the SR can be received by the ST error free. In (19) and (20), $g_\tau^E = g_\tau(1)$ and $g_\tau^A = g_\tau(\tau_s)$, where the superscripts E and A represent the prediction and transmission, respectively. The expected utility that the ST obtains at time τ with the PT operating with the k-th transmit power level, which is $r(p_\tau^k, a_\tau, k)$, can be given by

$$r(p_\tau^k, a_\tau, k) = \begin{cases} \left[p_\tau^k g_\tau^A D_k - (1 - p_\tau^k g_\tau^A) \sum_{j=1}^{L} C_{kj} Y_j\right] \tau_s, & a_\tau = 1, \\ 0, & a_\tau = 0, \end{cases} \tag{21}$$

where D_k is the reward that the ST will receive when the ST transmits with the k-th power level and $k = l$, and Y_k is the penalty when $k \neq l$.

An ST access policy $\epsilon = [d_0, \cdots, d_\tau, \cdots]$ maps the ST belief space $\{p_\tau^k, \tau \geqslant 0\}$ to action space $\{a_\tau, \tau \geqslant 0\}$. Thus, the optimal prediction-transmission policy aims to maximize the expected average utility, which can be given by

$$\max_\epsilon \lim_{M' \to \infty} \frac{\sum_{m=M+1}^{M'} \sum_{\tau=0}^{\tau_p^m - 1} r(p_\tau^k, a_\tau, k) a_\tau / (M' - M)}{\sum_{m=M+1}^{M'} \tau_p^m / (M' - M)}. \tag{22}$$

The total utility obtained by the ST during each PT hypothesis is i.i.d.. Thus, by the law of large numbers, the maximization problem in (22) can be rewritten as

$$\max_\epsilon \mathrm{E}\left[\sum_{\tau=0}^{\tau_p - 1} r(p_\tau^k, a_\tau, k) a_\tau\right]. \tag{23}$$

Let $V_\epsilon(0, p_\tau^k, k)$ denote the expected utility that can be achieved in each PT hypothesis following policy ϵ, which is

$$V_\epsilon(\tau = 0, p_\tau^k = 1, k) = E_\epsilon\left[\sum_{\tau=0}^{\tau_p - 1} r(p_\tau^k, a_\tau, k) a_\tau\right]. \tag{24}$$

Then, the maximum utility that can be achieved by the ST in each PT hypothesis as

$$V(\tau = 0, p_\tau^k = 1, k) = \sup_\epsilon V_\epsilon(\tau = 0, p_\tau^k = 1, k), \tag{25}$$

where

$$V(\tau, p_\tau^k, k) = \max\left\{E(\tau, p_\tau^k, k), A(\tau, p_\tau^k, k)\right\}. \tag{26}$$

In (26), $E(\tau, p_\tau^k, k)$ and $A(\tau, p_\tau^k, k)$ are the expected utilities that can be obtained by the ST through prediction and transmission, respectively. We have

$$E(\tau, p_\tau^k, k) = \sum_{j=1}^{L} \Pr\left[O_{\tau+1}^E = \mathcal{H}_j\right] V(\tau + 1, p_{\tau+1}^k(\mathcal{H}_j), k), \tag{27}$$

and

$$\begin{aligned} A(\tau, p_\tau^k, k) = &\Pr\left[O_{\tau+\tau_s}^A = \mathcal{A}\right] V(\tau + \tau_s, p_{\tau+\tau_s}^k(\mathcal{A}), k) \\ &+ r(p_\tau^k, 1, k). \end{aligned} \tag{28}$$

Lemma 1. *$V(\tau, p_\tau^k, k)$ is a convex function increasing with p_τ^k for given τ and k.*

Lemma 2. *$E(\tau, p_\tau^k, k)$ and $A(\tau, p_\tau^k, k)$ are convex functions increasing with p_τ^k for given τ and k.*

The proofs are omitted here due to the limited space.

It is clear that $V(\tau, p_\tau^k, k)$ is derived backward in time domain in (26), (27) and (28). Thus it will be helpful for the derivation of the optimal policy if an upper bound of τ can be established, which is given by

$$\mathcal{T}_k = \min\left\{\tau' : g_\tau^A < \frac{\sum_{j=1}^{L} C_{kj} Y_j}{\sum_{j=1}^{L} C_{kj} Y_j + D_k}, \forall \tau > \tau'\right\}. \tag{29}$$

In each PT hypothesis with the k-th transmit power level, the transmission action will not be taken by the ST after $\mathcal{T}_k$ as $\forall \tau > \mathcal{T}_k, r(1, 1, k) < 0$. Therefore, $V(\tau, 1, k) = 0, \forall \tau > \mathcal{T}_k$. According to [13], we find that $\mathcal{T}_k < \infty$ always holds as τ_p follows a Poisson distribution in (16). Note that $E(\tau, 0, k) = 0$ in (27) and $A(\tau, 0, k) = -\sum_{j=1}^{L} C_{kj} Y_j \tau_s < 0$ in (28). Combining Lemma 2, we define the probability thresholds as

$$p_\tau^{k*} = \min\{p_\tau^k : E(\tau, p_\tau^k, k) \leqslant A(\tau, p_\tau^k, k)\}. \tag{30}$$

Then, we can give the optimized protocol as

$$a_\tau^* = \begin{cases} 0, & p_\tau^k \leqslant p_\tau^{k*} \text{ or } \sum_{\tau_0=1}^{\tau_s - 1} a_{\tau-\tau_0}^* > 0, \\ 1, & \text{others}, \end{cases} \tag{31}$$

where a_τ^* is the optimal action at time τ.

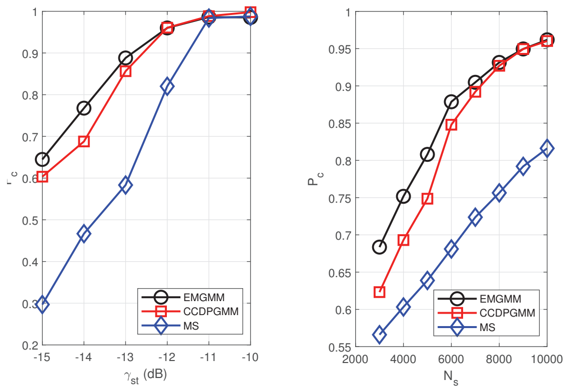
V. Numerical Results

In this section, numerical results are provided to illustrate the advantages of the proposed two-stage spectrum sharing strategy. In the simulation, we set the power level $L = 4$ and the probability of each hypothesis $\Pr(\mathcal{H}_l) = 0.25$. It is assumed that the noise variance $\sigma_u^2 = 1$, the PT transmit powers $P_1 : P_2 : P_3 = 1 : 2 : 3$, and $P_4 = 0$, thus $\gamma_{st}^1 : \gamma_{st}^2 : \gamma_{st}^3 = 1 : 2 : 3$ and $\gamma_{st}^4 = 0$. The average SNR at the ST is defined as $\gamma_{st} = (1/L) \sum_{l=1}^{L} \gamma_{st}^l$. In addition, we set the time duration of a sensing slot $T_{ss} = 2$ ms, the reward $D_k = 1$, and penalty $Y_l = 1$ in (21).

For our proposed strategy, in the prediction part of the second stage, the ST can easily identify the current PT power level based on the inferred $\{\boldsymbol{\theta}, \boldsymbol{\pi}, L\}$ in the first stage and X_n, $n > N$. Note that in addition to the proposed CCDPGMM, there are other learning methods to obtain this parameter set, such as expectation maximization GMM (EMGMM) and mean shift (MS). Among them, the EMGMM is a parametric clustering method requiring the prior knowledge of L, while the CCDPGMM and MS belong to a nonparametric class without the need for such prior knowledge. Taking into account the multiple power levels, we define the probability of correct PT power level prediction as

$$P_c = \sum_{l=1}^{L} \Pr\{\hat{\mathcal{H}}_l | \mathcal{H}_l\} \Pr\{\mathcal{H}_l\}. \tag{32}$$

In Fig. 2, we illustrate P_c for three different learning methods, with respect to different γ_{st} and N_s. It is shown in



(a) $N_s = 10^4$. (b) $\gamma_{st} = -12$ dB.

Fig. 2. The probability of correct PT power level prediction in the second stage (P_c) versus γ_{st} and N_s.

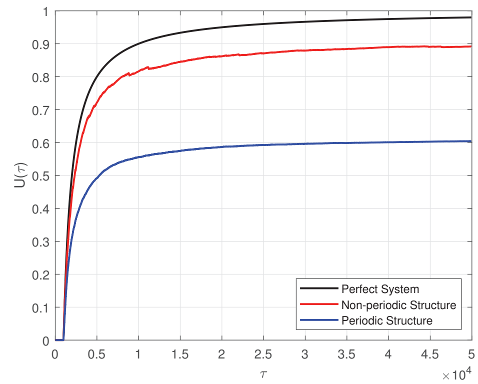


Fig. 3. The NPLA performance $U(\tau)$ for the spectrum sharing strategies using different structures in the second stage with $f_s = 5$ MHz, $N_s = 10^4$, $\tau_s = 2$, $\gamma_{st} = -12$ dB, and $\nu = 50$.

Fig. 2(a) that P_c increases with γ_{st} for all the three methods. This is because the gap between the adjacent transmit powers increases with γ_{st}, rendering them more distinguishable from the perspective of machine learning. Similarly, P_c improves with increasing N_s in Fig. 2(b), because a larger N_s results in a smaller variance of each Gaussian distribution in the mixture model. In Fig. 2, without the prior knowledge of L, DPGMM significantly outperforms MS, particularly for small γ_{st} and large N_s. Meanwhile, despite the lack of the prior knowledge of L, CCDPGMM is only slightly inferior to EMGMM, and the gap diminishes for increased γ_{st} or N_s. This verifies the advantages of the CCDPGMM over other learning methods.

Then, we illustrate the impact of different access strategies in the second stage on the NPLA performance. The NPLA performance is defined as

$$U(\tau) = \frac{\tau_s}{\tau} \sum_{\tau_0=0}^{\tau} a_{\tau_0} \psi\Big(\sum_{\tau_1=0}^{\tau_s-1} |k_{\tau_0+\tau_1} - l_{\tau_0+\tau_1}|\Big), \qquad (33)$$

where k_τ and l_τ denote the ST and PT transmit power levels at time τ, respectively. We have $\psi(x) = 1$ when $x = 0$, and $\psi(x) = 0$ otherwise. Basically, a larger $U(\tau)$ leads to a better tradeoff between the secondary network throughput and the interference to the primary network.

In Fig. 3, we compare $U(\tau)$ of the proposed non-periodic spectrum sharing strategy with a periodic one in the second stage for different τ. Note that CCDPGMM is used in the first stage for all the cases. As an upper bound, we include a perfect system where the ST can always accurately track the PT power level variation. It is shown that $U(\tau)$ of three different structures remains 0 when $\tau \leqslant 10^3$ due to the learning period (i.e., no transmission), and approaches certain positive value when $\tau > 10^3$. Fig. 3 shows that the non-periodic structure outperforms the periodic one, which verifies the benefit of dynamically adjusting the prediction intervals.

VI. Conclusions

In this paper, we proposed a two-stage spectrum sharing strategy, where the PT transmits with multiple power levels. In the first stage, we proposed a CCDPGMM to capture the PT power variation. Then, we designed a Bayesian inference method to infer the model parameters for establishing the PT power profile. In the second stage, we designed a prediction-transmission method to improve the NPLA performance. Finally, we verified the effectiveness of the proposed strategy with numerical results.

References

[1] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201–220, Feb 2005.

[2] T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Communications Surveys Tutorials*, vol. 11, no. 1, pp. 116–130, First 2009.

[3] Y. Chen, Q. Zhao, and A. Swami, "Joint design and separation principle for opportunistic spectrum access in the presence of sensing errors," *IEEE Trans. Inf. Theory*, vol. 54, no. 5, pp. 2053–2071, May 2008.

[4] 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.

[5] F. Gao, J. Li, T. Jiang, and W. Chen, "Sensing and recognition when primary user has multiple transmit power levels," *IEEE Trans. Signal Process.*, vol. 63, no. 10, pp. 2704–2717, May 2015.

[6] Z. Li, S. Cheng, F. Gao, and Y. C. Liang, "Sequential detection for cognitive radio with multiple primary transmit power levels," *IEEE Trans. Commun.*, vol. 65, no. 7, pp. 2769–2780, July 2017.

[7] F. Gao, C. Qian, H. Qian, and T. Zhang, "Sensing and recognition for multiple-primary-power-level scenario with noise uncertainty," *IEEE Trans. Veh. Technol.*, vol. 66, no. 3, pp. 2289–2300, March 2017.

[8] D. Wang, N. Zhang, Z. Li, F. Gao, and X. Shen, "Leveraging high order cumulants for spectrum sensing and power recognition in cognitive radio networks," *IEEE Trans. Wirel. Commun.*, vol. 17, no. 2, pp. 1298–1310, Feb 2018.

[9] Y. C. Liang, Y. Zeng, E. C. Y. Peh, and A. T. Hoang, "Sensing-throughput tradeoff for cognitive radio networks," *IEEE Trans. Wirel. Commun.*, vol. 7, no. 4, pp. 1326–1337, April 2008.

[10] D. Görür and C. Edward Rasmussen, "Dirichlet process Gaussian mixture models: Choice of the base distribution," *Journal of Computer Science and Technology*, vol. 25, no. 4, pp. 653–664, Jul 2010. [Online]. Available: https://doi.org/10.1007/s11390-010-9355-8

[11] R. M. Neal, "Markov chain sampling methods for Dirichlet process mixture models," *J. Comput. Graph. Stat.*, vol. 9, no. 2, pp. 249–265, 2000.

[12] C. E. Rasmussen, "The infinite Gaussian mixture model," in *Advances in neural information processing systems*, 2000, pp. 554–560.

[13] S. Huang, X. Liu, and Z. Ding, "Optimal sensing-transmission structure for dynamic spectrum access," in *IEEE INFOCOM 2009*, April 2009, pp. 2295–2303.