Cluster Control and Energy Consumption Minimization for Cooperative Prediction Based Spectrum Sensing in Cognitive Radio Networks

Dawei Nie, Wenjuan Yu, *Member, IEEE*, Qiang Ni, *Senior Member, IEEE*, Haris Pervaiz, *Member, IEEE*, and Geyong Min

***Abstract*—Spectrum sensing is a key technique for dynamically detecting available spectrum in cognitive radio networks (CRNs), which can introduce high resource demands such as energy consumption. In this paper, we propose a novel cluster-based cooperative sensing-after-prediction scheme where a learning cluster and a sensing cluster are jointly considered to perform cooperative prediction and sensing efficiently. This enables us to skip the complex physical sensing to reduce the demands when the spectrum availability can be simply predicted using cooperative prediction. Furthermore, the clustering is flexible, in order to meet different performance requirements. We then formulate two optimization problems to minimize the total number of users in the two clusters or to minimize the total energy consumption, to meet different performance requirements, while in both cases guaranteeing the system accuracy requirement and individual energy constraints. To solve the two challenging integer programming problems, the unconstrained problems are mathematically solved first by relaxing the integer variable and fixing the cluster size. Such analytical solutions serve as a foundation for solving the original optimization problems. Then, two low-complexity search algorithms are proposed to achieve the global optimum, as they can obtain the same performance with exhaustive search. Simulation results validate the accuracy of the derived analytical expressions and demonstrate that the total energy consumption and the number of users contributing to learning and sensing can be greatly reduced by applying our optimized clustered sensing-after-prediction scheme.**

***Index Terms*—Spectrum sensing, cognitive radio networks (CRNs), energy consumption, spectrum prediction.**

I. Introduction

The scarcity of useful radio spectrum and energy consumption issue in cellular networks are getting severe since the 6G communication networks are expected to support more challenging applications and meet far more stringent requirements such as Tbps data throughput, sub-ms latency, ultra-dense networks, and ultra-low energy consumption [2]–[5]. These challenges can be tackled by using empowering technologies such as Terahertz (THz) and sub-THz communications, joint communication and sensing, and efficient spectrum re-utilization [3], [4], [6]. To address the severe spectrum scarcity issue, it requires the effective usage of the existing low-band, mid-band and high-band spectrum resources. Furthermore, future 6G networks are expected to support the Internet of Everything (IoE) and Artificial Intelligence (AI) services as an intelligent communication system. This calls for revolutionary enabling technologies that can achieve the deep integration of communication, AI, computing and sensing [7]. Cognitive radio (CR), which facilitates effective spectrum exploration and exploitation and inherently has a close relationship with machine learning (ML), is considered to be one of the key enablers for AI-empowered 6G networks [8].

Mitola firstly proposed the concept of cognitive radio in 2000, which was later identified as the candidate for Opportunistic Spectrum Sharing (OSA) application by FCC in December 2013 [9], [10]. After that, significant attention has been drawn to cognitive radio networks (CRNs) since they can greatly improve the spectrum utilization by allowing secondary users (SUs) to share the licensed spectrum with primary users (PUs) [11]–[13]. Cognitive radios sense the surrounding radio frequency (RF) environment and make decisions as an intelligent wireless communication system [14]–[18]. To improve the sensing accuracy, cooperative sensing (CS) can be applied by utilizing multiple SUs to sense collaboratively and fuse all the sensing results at a fusion center (FC). It has been shown that the sensing performance can be greatly improved with an increase in the number of cooperative SUs [19]. However, one drawback is that the increase in the number of cooperative SUs results in the

This work was supported by the European Union's Horizon 2020 SANCUS project under the grant number GA952672. Part of this work was presented at the IEEE ICC 2022 DDINS workshop [1]. (Corresponding author: Qiang Ni.)

D. Nie, W. Yu, Q. Ni, and H.B. Pervaiz are with the School of Computing and Communications, InfoLab21, Lancaster University, LA1 4WA, U.K. (Email: d.nie, w.yu8, q.ni, h.b.pervaiz@lancaster.ac.uk)

G. Min is with the Department of Computer Science, University of Exeter, EX4 4QF, U.K. (Email: g.min@exeter.ac.uk)

increased energy consumption.

Low energy consumption is a key requirement of 6G communications. In CRNs, energy consumption is also a bottleneck issue because SUs are normally battery-powered wireless devices. When more SUs participate in CS, more energy is consumed. Given the limited battery capacity, spending more energy in spectrum sensing will result in less energy available for the other functions [20]. To effectively reduce energy consumption while maintaining system performance, studies have been conducted on the usage of alternative approaches in spectrum exploration, for example, spectrum prediction using ML techniques [21], [22]. Empirical studies have shown that in real-life spectrum usage, certain patterns exist [23]. This provides an opportunity to use ML-based prediction methods for identifying underutilized spectrum, leading to a reduced total energy consumption without sacrificing the system accuracy. It has been shown that the accuracy of spectrum prediction is normally close to and sometimes even exceeds the accuracy of spectrum sensing, depending on the randomness of environment and the accuracy of ML algorithms [16], [24].

Some recent studies have started to focus on the joint prediction-and-sensing scheme for CRNs. A three-step cooperative sensing and prediction algorithm was proposed in [25], where a spectrum prediction method based on the historical sensing information was used. By applying the joint prediction-sensing approach, the overhead and complexity of spectrum sensing were reduced. To improve the accuracy of spectrum prediction, Abdul Rahim et al [26] proposed an approach called "Exponential Weighted Moving Average (Holt's EWMA)". A sensing approach was combined with spectrum prediction to reduce the interference caused by the SU. In order to improve the accuracy of radio spectrum sensing, a spectrum prediction model for multi-channel was designed in [27]. The joint Long and Short Term Memory (LSTM) and Convolutional Neural Network (CNN) were used for model design. Different soft and hard fusion methods for performing cooperative spectrum prediction with trained LSTM-based local predictors were studied in [28]. The results in [28] indicated that the LSTM-based predictors have better accuracy performance than the traditional fusion methods. These studies mainly investigated the performance in terms of system accuracy and reliability after applying the joint prediction-and-sensing scheme, while energy consumption was not their focus. Moreover, in these studies, the number of users participating in sensing and prediction is fixed. A flexible clustering approach is more suitable for future CRNs to cope with various performance requirements of SUs, compared to the traditional fixed clustering approach.

Taking energy consumption into consideration, some studies have studied the joint prediction-sensing scheme with substantial probability models to obtain the accuracy expressions and aimed to improve system performance. A joint prediction-sensing model was introduced in [21] that utilized a parallel fusion-based cooperative spectrum prediction scheme to minimize errors and increase efficiency for energy-constrained CRNs. The accuracy was illustrated in [21] using simulation results, however the exact analytical expressions were not given. A spectrum sharing model based on spectrum prediction and sensing was proposed in [22]. The authors investigated a joint optimization design of transmit beamforming at the secondary base station (BS) with energy and sensing time constraints, where the probability models in [22] were from the perspective of correlation between single-user and cooperative results. These studies show the benefits of using joint prediction-sensing scheme in energy-constrained CRNs. However, the clustering of users, the minimization of the number of users participating in CS, and the tradeoff between system performance and total energy consumption were not studied in the literature.

In this paper, we propose a novel joint cluster-based sensing-after-prediction scheme where cluster size can be dynamically optimized, while adopting an adjustable decision threshold which decides when the scheme shall not accept the prediction result and opt for sensing instead. We formulate two optimization problems from two different perspectives in order to adapt different scenarios. First, we aim to find the minimum number of participating users that is required for the cluster-based CRN to maintain the system accuracy without violating each user's energy constraint. It will be demonstrated that this flexible approach can reduce the total energy consumption while minimizing the total number of participating users, in comparison to the fixed cluster size design in the literature. However, this optimization does not guarantee that the total energy consumption of the system is minimized. Hence, for the CRNs with strict energy constraints, we formulate a second optimization problem that aims to minimize the total energy consumption of the cluster-based CRN under the constraints of system accuracy and SU peak energy consumption. These two optimization problems have different purposes and are suitable in different scenarios. The analysis and simulations in this paper are based on the statistical analysis of the proposed model. The overall contributions of this paper are summarized as follows:

- Focusing on the proposed sensing-after-prediction

scheme, we derive the exact analytical expressions for accuracy and energy consumption by utilizing the probability models of system decisions. Then, the analytical optimal solution of decision threshold to the unconstrained accuracy maximization is derived after relaxing the integer variable and fixing cluster size, which will lay the foundation for solving the two constrained optimization problems.

- Two independent optimization problems are formulated to optimize the proposed scheme in order to adapt to different scenarios. The first problem, denoted by P1, addresses the minimization of the number of users contributing to the learning and sensing process while supporting the required accuracy and energy performance. The second problem, P2, focuses on the total energy consumption minimization under the same constraints, which is more suitable for energy-constrained scenarios.
- To solve P1 and P2, two low-complexity search algorithms, named UBS-1 and UBS-2, are proposed, by leveraging the analytical results and the optimal solutions to the unconstrained problems. Simulation results show that UBS-1 and UBS-2 perform optimally, by comparing with exhaustive search. The performance of P1 and P2 is further analyzed, by considering different prediction accuracy and utilizing the searched optimal results. It is found that though the optimal solutions to P1 and P2 are for totally different scenarios, they achieve similar performance in a high prediction accuracy environment.
- Comprehensive comparisons between the proposed scheme and some benchmark schemes are conducted. Simulation results demonstrate that the energy consumption and participated users of CRN are greatly reduced by applying the proposed optimizations, while the trade-offs between accuracy and energy performance are analyzed. Furthermore, even with the same cluster size, the proposed scheme can still achieve a minimum energy consumption that is much lower than benchmark schemes with the same requirements.

The remainder of this paper is organised as follows. Section II introduces the system model and the cluster-based sensing-after-prediction scheme. In Section III, we provide the optimization problem formulations and the theoretical analysis on accuracy and energy consumption for the proposed scheme. The optimal solutions are provided in Section IV where the Pseudocodes of the proposed UBS-1 and UBS-2 algorithms are also given. Simulation results are included in Section V, followed by conclusions summarized in Section VI.

II. System Model

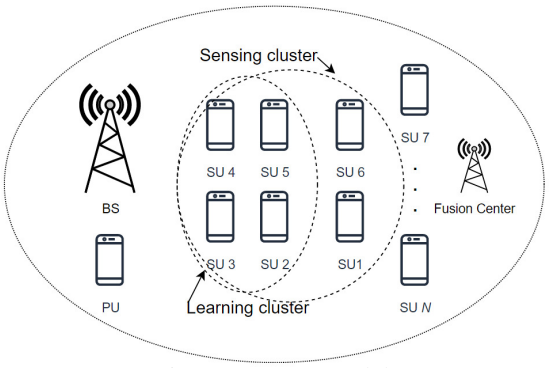


Fig. 1: System model.

Consider a CRN that consists of one PU, N SUs, and one FC. The system model is shown in Fig. 1. It is assumed that the primary channel has a regular usage pattern, which conforms to the observation of empirical study [23]. There are two possible channel states for the PU, i.e., busy and idle. P_{B} and P_{I} are the probabilities of PU being busy and idle respectively, where $P_{\mathrm{B}} + P_{\mathrm{I}} = 1$. Two clusters of SUs are responsible for voting for the channel state decision. A learning cluster $\mathcal{N}_1$ and a sensing cluster $\mathcal{N}_2$ are formed by n_1 SUs and n_2 SUs, respectively, where $n_1, n_2 \leq N$.[1] PU and SUs are assumed to be synchronized. Time division multiple access (TDMA) is assumed in the system. Each time slot T is divided into three sub-slots, namely Prediction Stage, Reaction Stage and Transmission Stage. The slot structure of the proposed system is shown in Table I, where detailed explanations of stages are given below.

A. Prediction Stage

During the Prediction Stage, each SU in the learning cluster $\mathcal{N}_1$, e.g., the SU i, predicts PU channel state using ML techniques based on channel usage histories, and reports a binary decision D_i, $\forall i \in \mathcal{N}_1$, to FC to perform cooperative prediction (CP). D_i being 0 represents that the SU i predicts that PU channel will be idle, while D_i equals 1 means that the SU i predicts that the PU channel is busy. The prediction results of SUs in the learning cluster $\mathcal{N}_1$ are fused together to obtain the final cooperative decision according to the logic rule (1) after

[1] In order to minimize the number of cognitive SUs participating in both clusters so as to reduce resource demands, it is preferable that each set is a subset of the other, i.e., $\mathcal{N}_1 \cap \mathcal{N}_2 = \mathcal{N}_1$ when $n_2 \geq n_1$, and $\mathcal{N}_1 \cap \mathcal{N}_2 = \mathcal{N}_2$ when $n_2 \leq n_1$.

TABLE I: Time and energy structure of one time slot.

Prediction	Reaction	Transmission
e_{M}	DM: 0	0 if channel busy
	IM: e_{S}	e_{T} if channel idle
$\longleftarrow t_{\mathrm{P}} \longrightarrow$	$\longleftarrow t_{\mathrm{R}} \longrightarrow$	$\longleftarrow t_{\mathrm{T}} \longrightarrow$

all the 1-bit decisions, i.e., D_i, are received at the FC. Let $\mathcal{S}$ be the state of the fused prediction decision.

$$\mathcal{S} = \begin{cases} \text{DM} & \mathcal{D} \geq \sigma \\ \text{DM} & \mathcal{D} \leq n_1 - \sigma \\ \text{IM} & \text{otherwise} \end{cases}, \tag{1}$$

where $\mathcal{D} = \sum_{i \in \mathcal{N}_1} D_i$. σ is a decision threshold that distinguishes between determinate (DM) and indeterminate (IM) states. This threshold indicates that if at least σ SUs' prediction results are identical, the system will enter the DM state and take what the majority of SUs (at least σ SUs) agree on as the final result. Hence, it is natural to assume $\sigma \geq \left\lceil \frac{n_1+1}{2} \right\rceil$, in order to guarantee when entering the DM state, at least half of the users have made the same prediction.

In [23], it was observed that the primary channels exhibit regular usage patterns. For example, the analysis of a real-life GSM downlink channel usage shows that every weekday exhibits high usage around noon but the peak usage occurs around 20:00 in the evening [23]. Inspired by this usage behaviour of channel, the system in our work is assumed to be in a long-term stable PU behavior environment (e.g., school classrooms) with a regular PU pattern but some randomness, while the requirements of the secondary users may vary. In such a stable environment, where the PU behavior follows regular patterns, it may not be necessary for SUs to perform sensing-based prediction if an acceptable accuracy performance can be achieved by using PU behavior history for prediction. Hence, in our system, the ML prediction scheme relies solely on the usage history of the PU channel, and each SU has an identical dataset of PU behavior history and prediction algorithm. This implies that the probability of ML prediction being correct for each SU, i.e., p, remains unaffected by the current state of the primary channel, whether it is idle or busy. Therefore, without loss of generality, we assume that the accuracy of the ML prediction for each SU is fixed for a given time period, and it remains the same for all SUs. Hence, the probability that at least σ prediction results of n_1 SUs are correct is given by

$$P_{\text{acc}} = \sum_{l=\sigma}^{n_1} \binom{n_1}{l} p^l (1-p)^{n_1-l}. \tag{2}$$

Recall (1) indicates that if at least σ prediction results are the same, the system will be in DM state and accept the CP result. Because the CP result can be either correct (with a probability of P_{acc}) or wrong, the probability of entering DM state can be calculated as the probability of at least σ predictions being right plus the probability of at least σ predictions being wrong. Hence, the probability of entering DM state, i.e., P_{DM}, and the accuracy of the CP result, i.e., $Q_{\text{DM}} = \frac{P_{acc}}{P_{\text{DM}}}$, are expressed in (3a) and (3b), respectively. The probability of the system in IM state is then given by $P_{\text{IM}} = 1 - P_{\text{DM}}$.

$$P_{\text{DM}} = \sum_{l=\sigma}^{n_1} \left[\binom{n_1}{l} p^l (1-p)^{n_1-l} + \binom{n_1}{l} p^{n_1-l} (1-p)^l \right], \tag{3a}$$

$$Q_{\text{DM}} = \frac{\sum_{l=\sigma}^{n} \binom{n_1}{l} p^l (1-p)^{n_1-l}}{\sum_{l=\sigma}^{n_1} \left[\binom{n_1}{l} p^l (1-p)^{n_1-l} + \binom{n_1}{l} p^{n_1-l} (1-p)^l \right]}. \tag{3b}$$

B. Reaction Stage

As mentioned above, when the system is in DM state, it indicates that it believes the prediction result is accurate. Then, all SUs will stay silent during the Reaction Stage and the transmission decision is made solely based on the CP result. On the other hand, if the system state is IM, it indicates that the system is not confident with the obtained CP result. Hence, independent spectrum sensing will be performed by each SU in the sensing cluster $\mathcal{N}_2$. The energy spent at each SU for performing sensing is denoted by e_{S}. The time duration of Reaction Stage is t_{R}.

There are two key metrics to evaluate the performance of spectrum sensing, namely the probability of detection (P_d) and the probability of false alarm (P_f), respectively. P_d indicates the probability that the SU declares the primary channel is occupied when the primary channel is indeed busy. P_f indicates the probability that the SU declares the primary channel is occupied when the primary channel is idle. P_d and P_f are respectively given by [11]

$$P_d = Q_u(\sqrt{2\gamma}, \sqrt{\lambda}), \tag{4}$$

$$P_f = \frac{\Gamma\left(u, \frac{\lambda}{2}\right)}{\Gamma(u)}, \tag{5}$$

where λ and γ represent energy detection threshold and signal-to-noise ratio (SNR), respectively. u refers to the time-bandwidth product of the energy detector. Γ is the incomplete gamma function, given by $\Gamma(a,x) = \int_x^\infty t^{a-1} e^{-t}\, dt$. Q indicates the generalized Marcum Q-function, given by $Q_u(a,x) =$

$\frac{1}{a^{u-1}} \int_x^\infty t^u e^{-\frac{t^2+a^2}{2}} I_{u-1}(at)\, dt$, where $I(\cdot)$ is the modified Bessel function.

After performing the independent spectrum sensing, each SU reports its binary decision to the FC. At the FC, all binary decisions are fused together according to a fusion rule. A voting rule is used for CS decision in our system. For example, the primary channel is sensed to be busy if at least q sensing results of n_2 SUs declare the channel as occupied. Therefore, the detection probability and the false alarm probability of CS decision are respectively given by

$$Q_d = P(\mathcal{H}_1|H_1) = \sum_{l=q}^{n_2} \binom{n_2}{l} P_d^l \left(1-P_d\right)^{n_2-l}, \quad (6)$$

$$Q_f = P(\mathcal{H}_1|H_0) = \sum_{l=q}^{n_2} \binom{n_2}{l} P_f^l \left(1-P_f\right)^{n_2-l}, \quad (7)$$

where $\mathcal{H}_0$ represents that the primary user is predicted to be absent and $\mathcal{H}_1$ represents that primary user is predicted to be present. H_0 and H_1 represent the channel state, where the probabilities are given by $P(H_0) = P_{\mathrm{I}}$ and $P(H_1) = P_{\mathrm{B}}$. Let Q_{S} denote the accuracy of CS. We can get that

$$\begin{aligned} Q_{\mathrm{S}} =& P(\mathcal{H}_1 H_1) + P(\mathcal{H}_0 H_0) \\ =& P(H_1)P(\mathcal{H}_1|H_1) + P(H_0)P(\mathcal{H}_0|H_0) \\ =& P_{\mathrm{B}} Q_d + P_{\mathrm{I}}(1-Q_f). \end{aligned} \quad (8)$$

C. Transmission Stage

Transmission action is decided based on the CP or CS decision obtained in the first two stages. If the system believes that the PU channel state is idle, the transmission will start by consuming the energy e_{T} and occupying the time period t_{T}. The decision tree diagram of the sensing-after-prediction scheme is illustrated in Fig. 2. The system decision accuracy, denoted by Q, can be finally obtained and given by

$$Q = P_{\mathrm{DM}} Q_{\mathrm{DM}} + P_{\mathrm{IM}} \left(P_{\mathrm{B}} Q_d + P_{\mathrm{I}} \left(1-Q_f\right)\right). \quad (9)$$

Recall that $Q_{\mathrm{S}} = P_{\mathrm{B}} Q_d + P_{\mathrm{I}}(1-Q_f)$. Hence, (9) can be simplified by using (8), yielding $Q = P_{\mathrm{DM}} Q_{\mathrm{DM}} + P_{\mathrm{IM}} Q_{\mathrm{S}}$. It clearly reflects the system decision accuracy depends on the prediction accuracy and the sensing accuracy. Note that the system accuracy is impacted by the learning cluster size n_1, the sensing cluster size n_2, as well as the decision threshold σ. To better reflect that, Q will be written as $Q(\sigma, n_1, n_2)$ in the following sections. Assume that a majority rule is used for CS, i.e., $q = \left\lceil \frac{n_2+1}{2} \right\rceil$ in (6) and (7). Then, by inserting (3a), (3b), (6) and (7) into (9), the analytical expression of system accuracy can be finally given by (10). To verify the accuracy of (10), Monte-Carlo simulations will be provided in Section V.

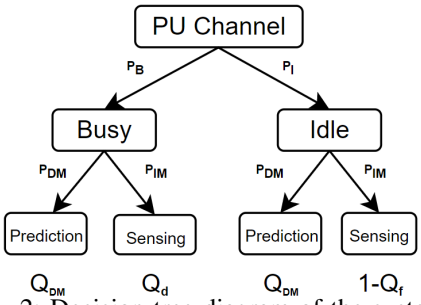


Fig. 2: Decision tree diagram of the system.

III. Problem Formulation and Theoretical Analysis

Based on the proposed cluster-based sensing-after-prediction scheme and the system accuracy analysis given in Section II, in Section III we will formulate the optimization problems, with the aim of reducing the number of participating users and the total energy consumption for CRNs while maintaining system accuracy and energy performance. To solve the challenging integer programming problems, in this section we will first focus on the unconstrained problems and derive the optimal solutions for them. The provided analysis can serve as a foundation for the ultimate solutions to the original constrained optimization problems.

A. Problem Formulation

The transmission states of the system are summarized in Table II, where Ts indicates that a transmission is scheduled consuming the energy e_{T}, while NT represents no transmission. In Table II, Q_i represents the probability of cooperative sensing correctly detecting the idle state of primary channel, i.e., $P(\mathcal{H}_0|H_0)$. From (7), we can get that $Q_i = 1 - Q_f$. When the primary channel is busy, the miss detection probability of cooperative sensing is denoted by Q_m, equal to $1 - Q_d$. Recall that Q_{DM} represents the correct probability of CP. Hence, the error probability of the CP decision is given by $Q_{\mathrm{FM}} = 1 - Q_{\mathrm{DM}}$.

In each time slot, the total energy consumption of the proposed scheme is evidently the summation of consumed energy during the Prediction Stage, Reaction Stage and Transmission Stage. The energy consumption of transmission is assumed to be fixed, and thus is omitted in the discussions below for simplicity. Assume that there are in total L slots during one experimental period. According to the probabilities provided in Table

$$
\begin{aligned}
Q\left(\sigma, n_1, n_2\right) = & \sum_{l=\sigma}^{n_1} \binom{n_1}{l} p^l (1-p)^{n_1-l} \\
& + \left(P_{\mathrm{B}} \sum_{l=q}^{n_2} \binom{n_2}{l} P_d^l \left(1-P_d\right)^{n_2-l} + P_{\mathrm{I}} \left(1 - \sum_{l=q}^{n_2} \binom{n_2}{l} P_f^l \left(1-P_f\right)^{n_2-l} \right) \right) \\
& \times \left(1 - \sum_{l=\sigma}^{n_1} \binom{n_1}{l} p^l (1-p)^{n_1-l} - \sum_{l=\sigma}^{n_1} \binom{n_1}{l} p^{n_1-l} (1-p)^l \right).
\end{aligned} \tag{10}
$$

TABLE II: Transmission states of the system.

Channel condition	Prediction state	Energy cost per slot			Status	Probability
		Prediction	Reaction	Transmission		
Busy	DM	e_{M}	0	NT, 0	Correct silence	$P_{\mathrm{B}} P_{\mathrm{DM}} Q_{\mathrm{DM}}$
			0	Ts, e_{T}	Fault transmission	$P_{\mathrm{B}} P_{\mathrm{DM}} Q_{\mathrm{FM}}$
	IM		e_{S}	NT, 0	Correct silence	$P_{\mathrm{B}} P_{\mathrm{IM}} Q_d$
			e_{S}	Ts, e_{T}	Fault transmission	$P_{\mathrm{B}} P_{\mathrm{IM}} Q_m$
Idle	DM		0	NT, 0	Miss transmission	$P_{\mathrm{I}} P_{\mathrm{DM}} Q_{\mathrm{FM}}$
			0	Ts, e_{T}	Correct transmission	$P_{\mathrm{I}} P_{\mathrm{DM}} Q_{\mathrm{DM}}$
	IM		e_{S}	NT, 0	Miss transmission	$P_{\mathrm{I}} P_{\mathrm{IM}} Q_f$
			e_{S}	Ts, e_{T}	Correct transmission	$P_{\mathrm{I}} P_{\mathrm{IM}} Q_i$

II, it can be calculated that the consumed energy, in Joule, for performing prediction and sensing during one experimental period is

$$
\begin{aligned}
& E_{\mathrm{T}}(\sigma, n_1, n_2) = n_2 L P_{\mathrm{IM}} e_{\mathrm{S}} + n_1 L e_{\mathrm{M}} \\
& = n_2 L \left(1 - \sum_{l=\sigma}^{n_1} \binom{n_1}{l} p^l (1-p)^{n_1-l} \right. \\
& \left. - \sum_{l=\sigma}^{n_1} \binom{n_1}{l} p^{n_1-l} (1-p)^l \right) e_{\mathrm{S}} + n_1 L e_{\mathrm{M}},
\end{aligned} \tag{11}
$$

where e_{M} and e_{S} are the slot-wise energy spent at a SU during prediction and sensing, respectively.

Because the SUs are all battery-powered devices, it is important to limit the peak energy consumption of each device to extend their battery life. In our cluster-based sensing-after-prediction scheme, the maximum energy consumption of a typical SU occurs when the user, denoted by n_p, is in the learning cluster as well as in the sensing cluster, i.e., $n_p \in \mathcal{N}_1 \cap \mathcal{N}_2$. Its peak energy consumption, denoted by $E(\sigma, n_1)$, can be given by

$$
\begin{aligned}
& E(\sigma, n_1) = L P_{IM} e_{\mathrm{S}} + L e_{\mathrm{M}} \\
& = L \left(1 - \sum_{l=\sigma}^{n_1} \binom{n_1}{l} p^l (1-p)^{n_1-l} \right. \\
& \left. - \sum_{l=\sigma}^{n_1} \binom{n_1}{l} p^{n_1-l} (1-p)^l \right) e_{\mathrm{S}} + L e_{\mathrm{M}}.
\end{aligned} \tag{12}
$$

After obtaining the above expressions, we can then start to formulate the optimization problems. As mentioned in the previous sections, two optimization problems will be formulated, namely P1 and P2. P1 aims to minimize the number of participating SUs under the constraints of energy consumption and accuracy requirement, while P2 focuses on the minimization of total energy consumption in the system, under the same constraints. Thus, P1 can be formulated as

$$
\begin{aligned}
\text{P1:} \quad & \min_{n_1, n_2, \sigma} \quad M(n_1, n_2) && \text{(13a)} \\
& \text{s.t. } Q(\sigma, n_1, n_2) \geq \epsilon, && \text{(13b)} \\
& \quad E(\sigma, n_1) \leq E_{\max}, && \text{(13c)} \\
& \quad \left\lceil \frac{n_1+1}{2} \right\rceil \leq \sigma \leq n_1, && \text{(13d)} \\
& \quad 1 \leq n_1 \leq N, && \text{(13e)} \\
& \quad 1 \leq n_2 \leq N, && \text{(13f)} \\
& \quad n_1, n_2, N, \sigma \in \mathbb{Z}, && \text{(13g)}
\end{aligned}
$$

where $M(n_1, n_2)$ is the total number of SUs participating in learning and sensing, i.e., $M(n_1, n_2) = n_1 + n_2$. ϵ and $E_{\max}$ represent the minimum accuracy requirement and the maximum energy value, respectively. For P2, we minimize the total energy consumption $E_{\mathrm{T}}(n_1, n_2, \sigma)$, under the constraints of energy consumption and system accuracy, formulated as

$$
\begin{aligned}
\text{P2:} \quad & \min_{n_1, n_2, \sigma} \quad E_{\mathrm{T}}(n_1, n_2, \sigma) && \text{(14a)} \\
& \text{s.t. } Q(\sigma, n_1, n_2) \geq \epsilon, && \text{(14b)} \\
& \quad E(\sigma, n_1) \leq E_{\max}, && \text{(14c)}
\end{aligned}
$$

$$\left\lceil \frac{n_1+1}{2} \right\rceil \leq \sigma \leq n_1, \tag{14d}$$

$$1 \leq n_1 \leq N, \tag{14e}$$

$$1 \leq n_2 \leq N, \tag{14f}$$

$$n_1, n_2, N, \sigma \in \mathbb{Z}. \tag{14g}$$

B. Theoretical Analysis

The formulated optimization problem P1 and P2 can be categorized as integer programming problems which are challenging to solve. To solve them, we first provide the theoretical analysis for the system accuracy and the total energy consumption that will reveal the properties of constraints and objective functions in P1 and P2.

1) Accuracy Analysis

Let us first focus on the unconstrained accuracy maximization problem where the learning and sensing cluster sizes are assumed to be fixed. The threshold σ is the only variable, in this case. We first relax the integer variable σ by treating it as a continuous one. The optimal threshold σ_0^* that maximizes the system accuracy $Q(\sigma, n_1, n_2)$ for the fixed cluster sizes n_1 and n_2 can be obtained by taking the first derivative of $Q(\sigma, n_1, n_2)$ with respect to σ. According to [11], we get that $\frac{\partial Q(\sigma,n_1,n_2)}{\partial \sigma} \approx Q(\sigma+1) - Q(\sigma)$, yielding

$$\frac{\partial Q(\sigma, n_1, n_2)}{\partial \sigma} = (Q_{\mathrm{S}} - 1)\binom{n_1}{\sigma} p^{\sigma}(1-p)^{n_1-\sigma} + Q_{\mathrm{S}}\binom{n_1}{\sigma} p^{n_1-\sigma}(1-p)^{\sigma}. \tag{15}$$

By setting $\frac{\partial Q(\sigma,n_1,n_2)}{\partial \sigma} = 0$, we get the expression of σ_0^* as follows.

$$\sigma_0^* = \begin{cases} \left\lceil \frac{1}{2} \log_{\frac{p}{1-p}} \frac{Q_{\mathrm{S}}}{1-Q_{\mathrm{S}}} \left(\frac{p}{1-p}\right)^{n_1} \right\rceil & \text{when } n_1 > 1, \\ 1 & \text{when } n_1 = 1. \end{cases} \tag{16}$$

Note that σ_0^* should be in the range of $\left\lceil \frac{n_1+1}{2} \right\rceil \leq \sigma_0^* \leq n_1$ to be feasible. By substituting (16) into the feasible range, it results in the following requirements of Q_{S}:

$$\begin{cases} \alpha \geq Q_{\mathrm{S}} > 0.5 & \text{when } n_1 \text{ is even and } p > 0.5, \\ \alpha \leq Q_{\mathrm{S}} < 0.5 & \text{when } n_1 \text{ is even and } p < 0.5, \\ \alpha \geq Q_{\mathrm{S}} > 1-p & \text{when } n_1 \text{ is odd and } p > 0.5, \\ \alpha \leq Q_{\mathrm{S}} < 1-p & \text{when } n_1 \text{ is odd and } p < 0.5, \end{cases} \tag{17}$$

where $\alpha = \frac{p^{n_1}}{(1-p)^{n_1}+p^{n_1}}$ and $p \neq 0.5$. The detailed proof can be found in Appendix A. Hence, if (17) is satisfied, σ_0^* that maximizes the system accuracy can be calculated using (16), assuming n_1 and n_2 are fixed. If the condition (17) is not met, the accuracy expression becomes a monotonic function and the maximum accuracy is achieved at endpoints, i.e, $\sigma_0^* = n_1$ or $\sigma_0^* = \left\lceil \frac{n_1+1}{2} \right\rceil$.

2) Energy Analysis

Correspondingly, we can also obtain the first derivative of total energy consumption $E_T(\sigma, n_1, n_2)$, given by

$$\frac{\partial E_{\mathrm{T}}(\sigma, n_1, n_2)}{\partial \sigma} \approx n_2 e_{\mathrm{S}} L \left[\binom{n_1}{\sigma} p^{\sigma}(1-p)^{n_1-\sigma} + \binom{n_1}{\sigma} p^{n_1-\sigma}(1-p)^{\sigma} \right], \tag{18}$$

which is always greater than or equal to 0. It means that the total energy consumption increases or remains flat with decision threshold σ. This conclusion will also be verified in Section V-B. As a result, by assuming the fixed n_1 and n_2, when there is no constraint, the optimal value of σ for minimizing energy consumption is obtained at $\left\lceil \frac{n_1+1}{2} \right\rceil$.

IV. Algorithm Design

The analytical results given in Section III-B reveal the properties of accuracy and energy consumption, which provide the optimal solutions to maximize the system accuracy and minimize the total energy consumption when there is no constraint and when the cluster sizes are fixed. Based on what have been analyzed, in this section we propose the ultimate solutions to solve the optimization problems P1 and P2. We start from analyzing the individual consumed energy $E(\sigma, n_1)$ in order to obtain the range of σ that is constrained by the per-user maximum energy consumption. Assuming that n_1 is a fixed value, we find $E(\sigma, n_1)$ has the following properties:

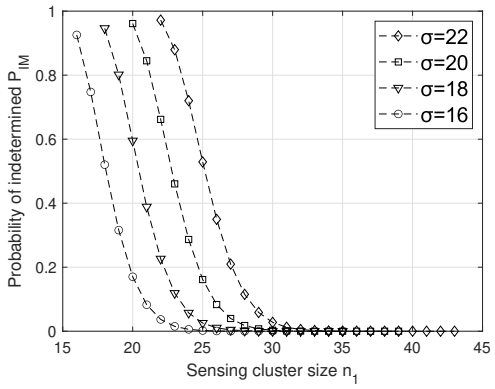
- Property 1: $E(\sigma+1, n_1) \geq E(\sigma, n_1)$. This is obtained by applying $\frac{\partial E(\sigma,n_1)}{\partial \sigma} \approx E(\sigma+1, n_1) - E(\sigma, n_1)$, yielding $\frac{\partial E(\sigma,n_1)}{\partial \sigma} \approx e_S L \left[\binom{n_1}{\sigma} p^{\sigma}(1-p)^{n_1-\sigma} + \binom{n_1}{\sigma} p^{n_1-\sigma}(1-p)^{\sigma}\right]$ which is always greater than or equal to 0. It indicates that $E(\sigma, n_1)$ monotonically increases or stays stable with the increase in σ.
- Property 2: $E(\sigma, n_1 - 1) \geq E(\sigma, n_1)$. To illustrate this, Fig. 3(a) is plotted which shows that P_{IM} decreases with n. Since $E(\sigma, n_1)$ monotonically increases with P_{IM}, therefore in our system, $E(\sigma, n_1)$ becomes smaller with a larger n.
- Property 3: $E(\sigma, n_1) \geq E(\sigma - 1, n_1 - 1) \geq E(\sigma, n_1 + 1)$. This can be confirmed by Fig. 3(b).
- Property 4: $E(n_1, n_1) > E(n_1 - 1, n_1 - 1)$. The detailed proof of this property is provided in Appendix B. By recursively using this property, we can conclude that the upper bounds for each n_1 follow $E(N, N) > E(N-1, N-1) > E(N-2, N-2) > \cdots > E(1, 1)$.

By leveraging these function properties, for any given n_1, an Upper-Bound Search (UBS) method can be designed to find the feasibility region of σ constrained by the energy consumption limit. The design strategies of UBS are summarized as below:

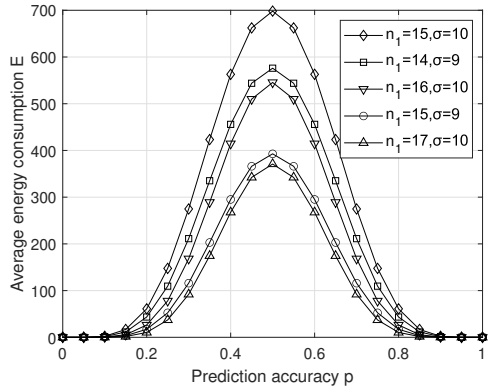
1) We first find the smallest value i that satisfies $E(i,i) > E_{\max}, \forall i \in [1,N]$. According to Property 4, we can note that for any $n_1 \in [1, i-1]$, we have $E(n_1,n_1) \leq E_{\max}$ satisfying the maximum energy constraint. According to Property 1, we know that $E(\sigma,n_1)$ monotonically increases or stays stable with the increase in σ. Hence, for any $n_1 \in [1,i-1]$, the feasibility region of σ satisfying the energy consumption constraint is $\sigma = \left\lceil \frac{n_1+1}{2} \right\rceil : n_1$.
2) Since i is the smallest integer that satisfies $E(i,i) > E_{\max}, \forall i \in [1,N]$, according to Property 2, we have that $E(i,i+1) \leq E_{\max}$. Furthermore, according to Property 3, we have $E(N-1,N) \geq \cdots \geq E(i+1,i+2) \geq E(i,i+1)$. Hence, we can continue the search by increasing the parameter values until we find those violating the energy limit, i.e., $E(i+l,i+1+l) > E_{\max}$, where $1 \leq l \leq N-i-1$. Then, we can see that for any $n_1 \in [i,i+l]$, the feasibility region of σ satisfying the energy consumption limit is $\left[\left\lceil \frac{n_1+1}{2} \right\rceil, n_1 - 1\right]$.
3) By reusing Property 2, we know that $E(i+l,i+2+l) \leq E_{\max}$. Hence, we can use the same strategy to search starting from $E(i+l,i+2+l)$ until we find the next violating point, i.e., $E(i+l+l_1,i+2+l+l_1) > E_{\max}$, where $1 \leq l_1 \leq N-i-2-l$. Then, we can see that for any $n_1 \in [i+1+l, i+1+l+l_1]$, the feasibility region of σ satisfying the energy consumption limit is $\left[\left\lceil \frac{n_1+1}{2} \right\rceil, n_1 - 2\right]$. Let c_t denote the number of times that the (σ,n_1) pair violates the energy consumption constraint, i.e., $E(\sigma,n_1) > E_{\max}$. From the above discussions, we can note that by using our search method, for any learning cluster number $\in [n_1, n_1 + l_{c_t-1}]$, the feasibility region of σ satisfying the energy consumption limit can be given by $\left[\left\lceil \frac{n_1+1}{2} \right\rceil, n_1 - c_t\right]$.

A. Upper-Bound Search Algorithm (UBS)

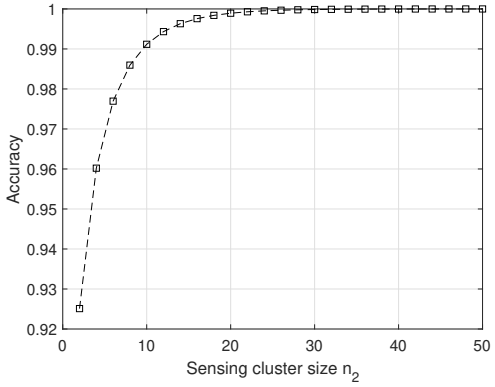
After obtaining the optimal solutions to the unconstrained problems and finding the feasibility region of σ under the energy consumption constraint, we can start searching the optimal solution (n_1^*, n_2^*, σ^*) to P1. Recall that P1 aims to minimize the number of SUs in learning and sensing clusters, i.e., $n_1 + n_2$, while satisfying accuracy and energy constraints. As shown in the Pseudocode of UBS-1, given in Algorithm 1, each n_2 value, where $n_2 \in [1,N]$, needs to be evaluated.



(a) P_{IM} versus n_1



(b) E versus p with different (n_1,σ) values



(c) Q_{S} versus n_2

Fig. 3: Properties of different functions.

Then, for any n_2, the search of n_1 stops if there is σ meeting both constraints (13b) and (13c). Since P1 has an accuracy requirement, hence for any given n_1 and n_2, it is important to first check if the maximum accuracy obtained for the unconstrained accuracy maximization

Algorithm 1 UBS-1

Input: N, ϵ, $E_{\max}$, P_d, P_f, p, P_{B}, P_{I}

Set $n_1 \leftarrow 1$, $n_2 \leftarrow 1$, $\sigma \leftarrow 1$, $\mathcal{F} = \emptyset$, $I = []$, $\mathcal{Q}_I = \emptyset$
while $n_2 \leq N$ **do**
  while $n_1 \leq N$ **do**
    if $E(\sigma, n_1) > E_{\max}$ **then**
      $c_t \leftarrow c_t + 1$;
    end if
    Set $I = \left[\left\lceil \frac{n_1+1}{2} \right\rceil, \cdots, n_1 - c_t\right]$
    if $Q(\sigma_0^*, n_1, n_2) \geq \epsilon$ **then**
      if $\sigma_0^* \in I$ **then**
        Set $\mathcal{F} \leftarrow \mathcal{F} \bigcup \{(n_1, n_2, \sigma_0^*)\}$;
        break
      else
        Insert $\sigma = (I_1, I_{\text{end}})$ into (10) to calculate $Q(\sigma, n_1, n_2)$;
        Set $\mathcal{Q}_I \leftarrow \{Q(\sigma, n_1, n_2) | \sigma = (I_1, I_{\text{end}}), Q(\sigma, n_1, n_2) \geq \epsilon\}$;
        if $\mathcal{Q}_I \neq \emptyset$ **then**
          $\mathcal{F} \leftarrow \mathcal{F} \bigcup \{(n_1, n_2, \sigma) | \sigma \leftarrow \arg\max \mathcal{Q}_I\}$;
          break
        end if
      end if
    end if
    $n_1 = n_1 + 1$;
    $\sigma = n_1 - c_t$;
  end while
  $n_2 = n_2 + 1$;
end while

Output: $(n_1^*, n_2^*, \sigma^*) \leftarrow \underset{(n_1, n_2, \sigma) \in \mathcal{F}}{\arg\min} \; n_1 + n_2$

problem, given in Section III-B1, can satisfy the accuracy requirement (13b). If that does not hold, it means for the considered n_1 and n_2, there is no such σ satisfying the constraint (13b). Hence, we should continue the search by increasing the values of n_1 or n_2. If that holds, it means that for the given n_1 and n_2, there is at least one σ that can satisfy the constraint (13b). Then, we need to check if these σ values satisfying (13b) also meet the constraint (13c). According to the analysis given at the beginning of Section IV, for any given n_1 and n_2, we can find the range of σ constrained by the energy consumption limit, meaning that the constraint (13c) can be now converted into the feasible range of σ. Hence, we just need to check if these σ values satisfying (13b) also fall in the feasible range constrained by (13c). Note that for any n_1 and n_2, only those σ values that can satisfy both (13b) and (13c) are possible solutions to P1. It is possible that for a given (n_1, n_2) pair, there are multiple σ values satisfying both constraints. In this case, we pick the σ that achieves the highest accuracy $Q(\sigma, n_1, n_2)$. Each feasible tuple (n_1, n_2, σ) that meets all constraints (13b)-(13g) is stored in a matrix $\mathcal{F}$. Among all the feasible solutions, the optimal solution (n_1^*, n_2^*, σ^*) is the one that achieves the minimum $n_1 + n_2$.

Let us analyze the complexity of UBS-1. The inner loop "while $n_1 \leq N$" is designed to find the first minimal n_1 that meets the requirements. Since the range of σ for each n_1 is $\sigma = \left\lceil \frac{n_1+1}{2} \right\rceil : n_1$, the complexity of $Q(\sigma, n_1, n_2)$ and $E(\sigma, n_1)$ for each inner loop step is $O(n_1)$. The inner loop is stopped when algorithms find the first optimal solution from $n_1 = 1$ to N, while the maximum complexity of the inner loop becomes $O(N^2)$. The outer loop "while $n_2 \leq N$" is designed to explore the sensing cluster size of each n_1. Each n_2 needs to search its corresponding minimal n_1 value in the inner loop in order to find the minimal $n_1 + n_2$. The outer loop "while $n_2 \leq N$" will not stop until $n_2 = N$ and the complexity of the outer loop will be $O(N)$. Since there are two nested loops from 1 to n_1 and from 1 to N respectively, their maximum complexity will be $O(N^3)$. If the exhaustive search is used for optimization, all n_1, n_2, σ need to be searched in three nested loops. Moreover, the computation of each σ still includes the totalization of $\sigma = \left\lceil \frac{n_1+1}{2} \right\rceil : n_1$. Since the upper bound of each variable is N, the complexity becomes a constant $O(N^4)$. It can be noticed that the complexity of our proposed algorithm is lower than the exhaustive search.

Recall that P2 aims to find the (n_1^*, n_2^*, σ^*) that minimize the total energy consumption E_{T} of the scheme, while satisfying accuracy and energy constraint. As shown in the Pseudocode of UBS-2, given in Algorithm 2, each n_2 value in $n_2 \in [1, N]$ needs to be evaluated. For any n_2, the search for each n_1 and its total energy consumption E_{T} is established. According to the analysis given at the beginning of Section IV, same as P1, for any given n_1 and n_2, we can find the range of σ constrained by the energy limit, meaning that the constraint (14c) can be now converted into the feasible range of σ. Since P2 has the same accuracy requirement as P1, hence for any given n_1 and n_2, it is important to check if the accuracy obtained for each σ in the feasible range can satisfy the accuracy requirement, i.e., we need to check if these σ values in the feasible range constrained by (14c) also satisfy (14b). The σs that do not satisfy the accuracy requirement need to be dropped. The range of σ that meets all constraints is stored in a matrix I. Recall that in Section III-B2, we have shown that E_{T} is a non-

Algorithm 2 UBS-2

Input: N, ϵ, $E_{\max}$, P_d, P_f, p, P_{B}, P_{I}

Set $n_1 \leftarrow 1$, $n_2 \leftarrow 1$, $\sigma \leftarrow 1$, $\mathcal{F} = \emptyset$, $I = []$, $\mathcal{Q}_I = \emptyset$
while $n_2 \leq N$ **do**
  while $n_1 \leq N$ **do**
    if $E(\sigma, n_1) > E_{\max}$ **then**
      $c_t \leftarrow c_t + 1$;
    end if
    Set $I = \left[\left\lceil \frac{n_1+1}{2} \right\rceil, \cdots, n_1 - c_t\right]$
    Insert each $\sigma \in I$ into (10) to calculate $Q(\sigma, n_1, n_2)$;
    Set $\mathcal{Q}_I \leftarrow \{Q(\sigma, n_1, n_2) | \sigma = I, Q(\sigma, n_1, n_2) \geq \epsilon\}$;
    Update $I \leftarrow \{\sigma | Q(\sigma, n_1, n_2) \geq \epsilon, \sigma = I\}$;
    if $\mathcal{Q}_I \neq \emptyset$ **then**
      Insert $\sigma = I_1$ into (11) to calculate $E_{\mathrm{T}}(\sigma, n_1, n_2)$;
      $\mathcal{F} \leftarrow \mathcal{F} \bigcup \{E_{\mathrm{T}}(\sigma, n_1, n_2) | \sigma \leftarrow I_1\}$;
    end if
    $n_1 = n_1 + 1$;
    $\sigma = n_1 - c_t$;
  end while
  $n_2 = n_2 + 1$;
end while

Output: $(n_1^*, n_2^*, \sigma^*) \leftarrow \underset{E_{\mathrm{T}}(n_1, n_2, \sigma) \in \mathcal{F}}{\arg\min} E_{\mathrm{T}}(n_1, n_2, \sigma)$

decreasing function with σ. Hence, for a given n_1 and n_2, the σ^* for the minimum total energy consumption will be the first value of the obtained range, i.e., $\sigma^* = I_1$. We store the minimum E_{T} that meets all constraints (14b)-(14g) for this specific (n_1, n_2, I_1) in a matrix $\mathcal{F}$. Finally, among all feasible solutions, the optimal solution (n_1^*, n_2^*, σ^*) is the one that achieves the minimum E_{T}. In the worst case, the complexity of search algorithm P2 will become $O(N^4)$, which is the same as exhaustive search. However, in most cases, the calculation steps of our algorithm are much less than exhaustive search because the searching of each σ is avoided.

Regarding the feasibility of this system in real-world settings, we would like to provide the following discussions. Firstly, the assumption of regular PU channel usage patterns in this work is based on real-life measurement data collected from a school building, which is considered a relatively stable environment [23]. Therefore, the prediction model does not need to be frequently retrained, as long as the PU exhibits a stable frequency of presence during specific times of the day. In some extreme scenarios, PU behavior may drastically change, but this will not happen frequently. Thus, the possibility of retraining our prediction model in real-time (or in a short period of time) is low. Secondly, the requirements of SUs are not updated in real-time. Instead, it is reasonable to allow SUs to report their energy consumption and accuracy requirements at the beginning of regular intervals. After collecting the updated requirements of SUs, the algorithm will be invoked to generate optimal decisions for the subsequent interval. Thirdly, the algorithm does not select specific SUs to include in a cluster. Instead, the algorithm determines the optimal clustering strategy and calculates the optimal number of SUs required with its optimal threshold in learning and sensing clusters. As a result, the arrival or departure of an individual SU will not trigger the re-execution of our algorithm in real time. Even if an SU departs from a cluster, we only need to add another SU to maintain the optimal number of SUs in the cluster. Therefore, in conclusion, even with a large number of users, calculations only need to be performed once within a given time period under a static environment. It is thus worth sacrificing processing time to reduce energy consumption by adjusting the parameters n_1, n_2, and σ.

V. Simulation Results

In this section, we will simulate the proposed sensing-after-prediction scheme in a CRN and compare it with the other benchmark schemes. Simulations will be also provided to validate the derived analytical expressions of system accuracy and energy consumption. We will simulate and show the expressions of system accuracy in Section V-A. The optimal threshold values to the unconstrained problems given in Section III-B are also simulated and compared with exhaustive search results. Then, in Section V-B, we will simulate and show the energy consumption of different schemes. The trade-offs between energy performance and accuracy are also discussed after deriving the expression of energy efficiency. In Section V-C, to validate the search algorithms UBS-1 and UBS-2, we compare their search results with that of exhaustive search, under the necessary constraints of system accuracy and energy consumption. Finally, we provide performance analysis about optimizations P1 and P2.

For the simulations, we assume that $e_{\mathrm{S}} = 100$ mJ, $e_{\mathrm{M}} = 10$ mJ, $L = 1000$, $P_f = 0.1$ and $P_d = 0.775$ [29], unless otherwise indicated. We also have that $P_{\mathrm{B}} = 1/6$ and $P_{\mathrm{I}} = 5/6$, inspired by [23] and [21]. It is also assumed that the accuracy requirement $\epsilon = 0.995$ and the total number of SUs in the network is 50, i.e., $N = 50$.

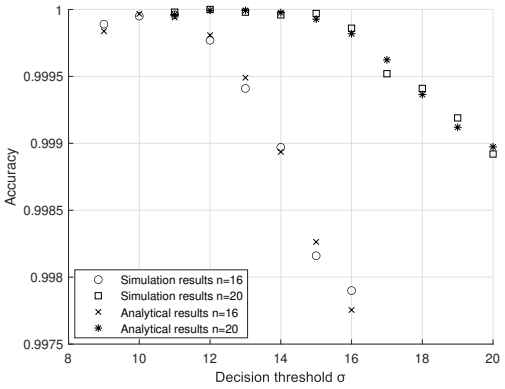
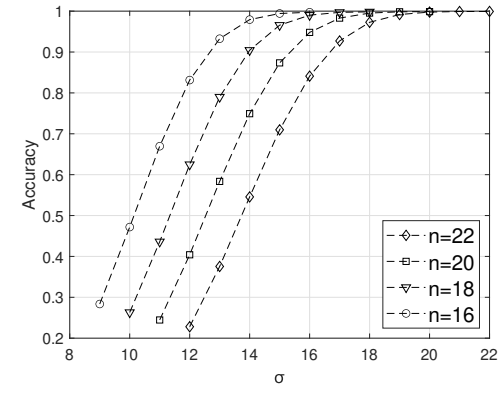


Fig. 4: Comparison of accuracy between analytical results and Monte-Carlo simulations.

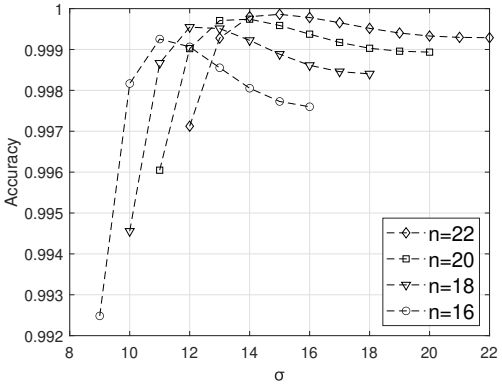
The ML prediction accuracy p is 0.85 in Section V-A and V-B unless otherwise indicated. Since Section V-A and V-B do not involve optimization results, the learning and sensing cluster sizes n_1 and n_2 are set to be identical in these two subsections, i.e. $n_1 = n_2 = n$, in order to clearly show the results and avoid excessive parameters.

A. Accuracy performance

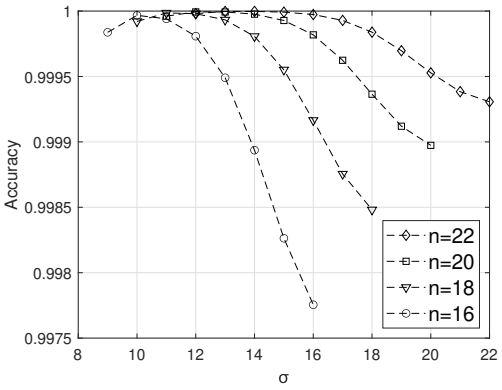
In this subsection, we aim to validate the derived analytical expressions of system accuracy, given in Section II and Section III-B1. Fig. 4 plots the accuracy versus the threshold σ where the analytical results using the derived analytical expression (10) and Monte Carlo results are both shown. It can be observed that the analytical results closely match with Monte Carlo simulations. It confirms the accuracy of the derived expression (10). Recall that when we analyze the optimal solution to the unconstrained accuracy maximization problem in Section III-B1, we find that the optimal solution depends on the values of n and p, where the detailed conditions are given in (17). In order to validate our analysis in Section III-B1, Fig. 5 plots the system accuracy versus the threshold σ for different cluster sizes. In order to show the impact of (17), different values of p are used. From figure 5(b) and 5(c), it can be observed that there exists an optimal threshold for each given cluster size n where the system reaches its maximum accuracy. When $p = 0.4$, the condition (17) is not satisfied. Figure 5(a) shows that in this case, the accuracy becomes a monotonic function, which confirms our analysis. Then, we want to provide simulations to validate the accuracy of the derived analytical optimal solution to the unconstrained accuracy maximization problem discussed in Section III-B1. Fig. 6 is thus plotted showing the optimal threshold versus the cluster size. It can be noticed that the



(a) $p = 0.4$



(b) $p = 0.75$



(c) $p = 0.85$

Fig. 5: System accuracy versus σ with different n and p values.

analytical results of optimal threshold perfectly match with the exhaustive search results.

In order to show the advantage of the proposed sensing-after-prediction scheme, Fig. 7(a) plots the accu-

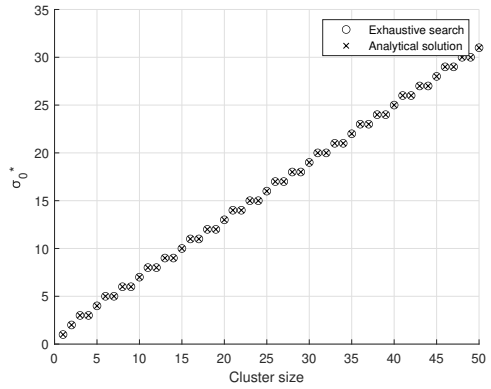
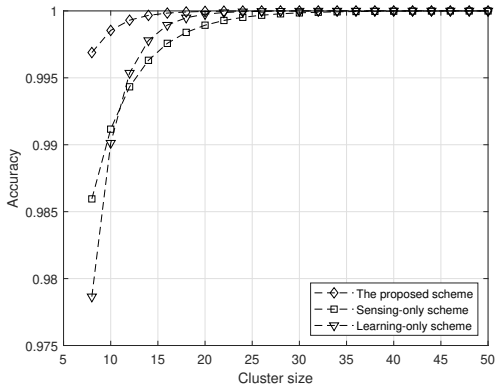
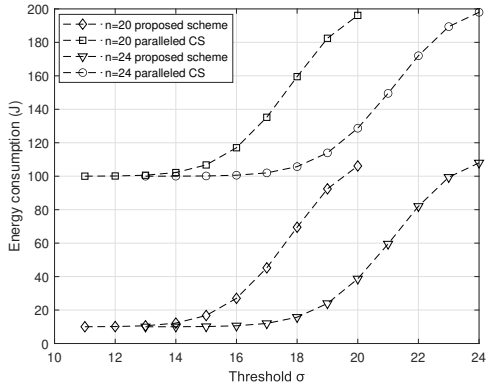


Fig. 6: Comparison between analytical solution and exhaustive search for the unconstrained optimal σ_0^*.



(a) Accuracy comparison between the proposed scheme and two benchmark schemes.



(b) Energy comparison between the proposed scheme and benchmark scheme.

Fig. 7: Comparison between the proposed scheme and benchmark schemes.

racy versus the cluster size for the proposed scheme and two benchmark schemes. The two benchmark schemes are the CS-only scheme or the CP-only scheme respectively. As their names suggest, each of them only uses a single method (sensing or prediction) instead of our joint approach to identify the underutilized spectrum. From this figure, it can be noticed that the accuracy of our sensing-after-prediction scheme is higher than the two benchmark schemes. Furthermore, it can be shown that the CS scheme performs better than the CP scheme when n is small. The reason is when $P_f = 0.1$ and $P_{\rm I} = 5/6$, at small values of n, a low P_f and a high $P_{\rm I}$ will cause $1 - Q_f$ dominate $Q_{\rm S}$, resulting in an amplified influence of $p < 1 - P_f$. When the cluster size n becomes larger, CP performs better than CS. In fact, all schemes reach a similar level when n is very large but our sensing-after-prediction scheme remains the highest accuracy. Fig. 7(b) shows the energy performance of the proposed scheme (cooperative learning-sensing) and parallel CS scheme (cooperative sensing only). The parallel CS scheme is assumed to use the same decision model and optimizations as the proposed scheme but it replaces the ML prediction phase of the proposed scheme with spectrum sensing. Fig. 7(b) illustrates that under the same condition, the scheme only uses spectrum sensing consumes more energy than the proposed scheme. Moreover, the energy consumption of both schemes is increased with decision threshold σ.

B. Energy performance

In this section, we aim to evaluate the energy performance and the trade-offs of the proposed scheme. Energy efficiency (EE) of proposed scheme is also derived and analysed in this section. Finally, we will simulate and show the energy performance of different schemes. Because we set $n_1 = n_2 = n$, it means all users are assumed to perform both CP and CS.

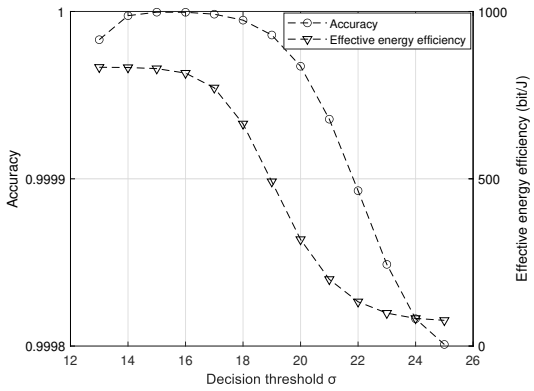
Apart from energy consumption, EE is another critical reflection of energy performance. We will show the performance of the adopted sensing-after-prediction scheme in terms of both metrics and discuss the tradeoff between energy performance and accuracy performance. Let us first derive the expression of EE. According to the probabilities provided in Table II, it can be calculated that the probability of data being successfully transmitted is $P_{\rm succ} = P_{\rm I}\left(P_{\rm DM}Q_{\rm DM} + P_{\rm IM}Q_s\right)$. Let r indicate the data rate of the channel. The amount of data transmitted, denoted by $C_{\rm T}$, equals $rt_{\rm T}$ bits. After obtaining the above equations, the total amount of data successfully transmitted per user can be obtained and expressed as $C_{\rm succ} = P_{\rm succ}LC_{\rm T}$ bits.

The effective energy efficiency (bit/J) per user is defined as the ratio of data successfully transmitted

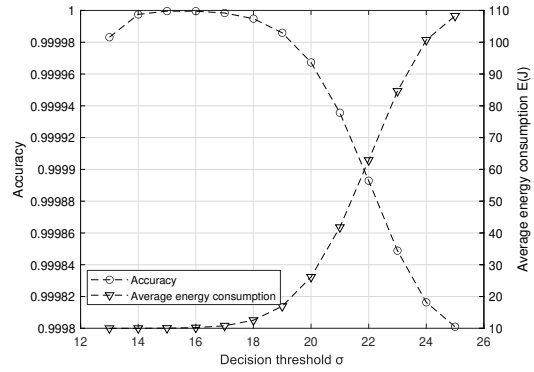
to the energy spent during the sensing-after-prediction process, yielding $\eta = \frac{C_{\text{succ}}}{E} = \frac{rt_{\text{T}}P_{\text{I}}(P_{\text{DM}}Q_{\text{DM}}+P_{\text{IM}}Q_s)}{e_{\text{M}}+P_{\text{IM}}e_{\text{S}}}$. The data rate r is set to be 1Mb/s and the transmission stage duration $t_{\text{T}} = 10$ms. Fig. 8(a) shows that the effective energy efficiency monotonically decreases with decision threshold σ, while Fig. 8(b) shows that the energy consumption monotonically increases with σ. The maximum accuracy and minimum energy consumption cannot be achieved by one same σ, which lead to our optimization problem. On the other hand, maximum efficiency and minimum energy consumption cannot be achieved by one same σ, which is another trade-off.

In order to show the energy performance of the proposed scheme, we simulate and compare the energy consumption between different schemes. Fig. 9(a) and Fig. 9(b) illustrate the average and total energy consumption versus the cluster size n for the proposed scheme and benchmark schemes. In Fig. 9(a), it is assumed that there are no energy and accuracy constraints. The two benchmark schemes are cooperative spectrum prediction model (CPM) in [30] and the CS-only [11] scheme. CPM is a joint sensing-prediction scheme where in spectrum prediction phase, SUs perform CP using HMM/MLP model. However, cluster size and DM/IM state control of sensing and learning cluster are not considered in CPM. Because there is no accuracy constraint in this section, the optimal threshold is used to reach maximum accuracy. Fig. 9(a) shows that the average energy consumption per SU of our proposed sensing-after-prediction scheme is much smaller than two benchmark schemes at large cluster size. Recall that the average energy consumption of the proposed scheme is $E = LP_{\text{IM}}e_{\text{S}} + Le_{\text{M}}$. In comparison, for the CS-only scheme, the average energy consumption is given by $E_{S_1} = Le_{\text{S}}$. When the cluster size becomes larger, the probability of entering the IM state, i.e., P_{IM}, is reduced. Hence, in our proposed scheme, there is a smaller probability that SUs need to perform sensing, resulting in a smaller energy consumption.

Fig. 9(b) is plotted to illustrate that the total energy consumption for the sensing-after-prediction scheme is also smaller than the other benchmark schemes under the same requirements. Using maximum cluster size $\max(n_1, n_2)$ as the x-axis is because it can show the flexibility of the proposed scheme and can be compatible with other benchmark schemes to perform comparisons. In each cluster size, we find the optimal n_1, n_2 combination of different schemes for minimizing the total energy consumption under the same requirements. Similar energy optimizations are performed for the semi-fixed scheme, the parallel CS scheme and the proposed



(a) Accuracy and effective energy efficiency versus σ

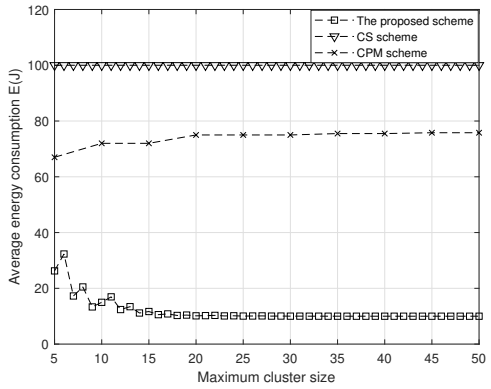


(b) Accuracy and average energy consumption versus σ

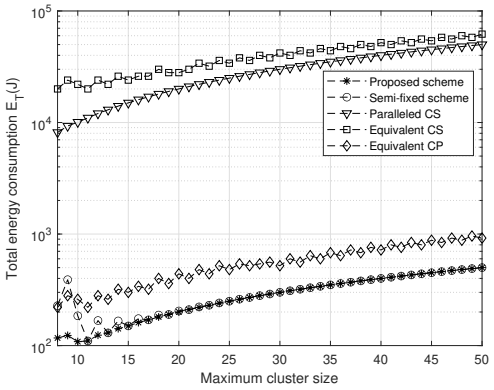
Fig. 8: Trade-offs between energy performance and system accuracy.

scheme. It is assumed that the semi-fixed scheme only exists one cluster and each SU in the cluster performs CS and CP, i.e., $n_1 = n_2$. From Fig. 9(b), it can be observed that the total energy consumption of the proposed scheme is lower than the semi-fixed scheme at smaller cluster sizes and performs similarly at a larger cluster sizes. Thus, the proposed scheme consumes less energy due to the flexibility of n_1, n_2.

On the other hand, benchmark schemes "equivalent cooperative spectrum sensing (CS)" and "equivalent cooperative prediction (CP)" are plotted in Fig. 9(b). The equivalent scheme aims to illustrate that under the same conditions and requirements, the amount of energy consumed by CP and CS schemes in order to reach the same required performance as our proposed scheme. Fig. 9(b) shows that the proposed scheme consumes much less energy than the equivalent CS and CP schemes. Finally, Fig. 9(b) plots a parallel CS scheme as a benchmark. The parallel scheme is assumed to use the same decision

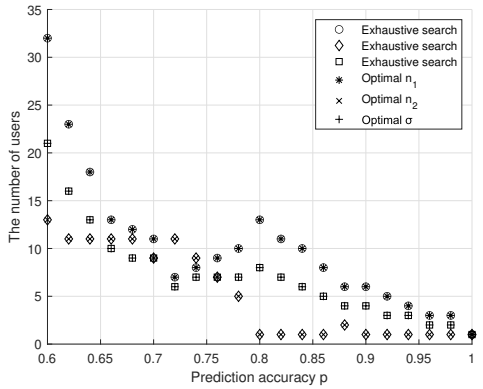


(a) Average energy consumption versus cluster size n

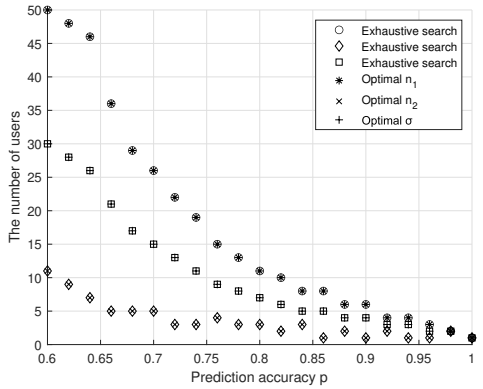


(b) Total energy consumption versus maximum cluster size n

Fig. 9: The comparison of energy consumption between the proposed scheme and benchmark schemes.



(a) Optimal solutions to P1



(b) Optimal solutions to P2

Fig. 10: The comparison of solutions between UBS algorithms and exhaustive search.

model as the proposed scheme but it replaces the ML prediction phase of the proposed scheme with spectrum sensing. From Fig. 9(b), it can be observed that under the same conditions and requirements, the proposed scheme consumes much less energy than parallel CS scheme. Moreover, the parallel scheme still consumes less energy than the equivalent CS scheme because the decision structure of parallel scheme is same as the proposed scheme.

C. Model comparison

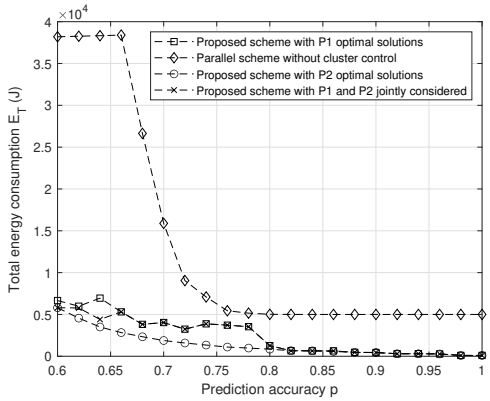
In this section, we will simulate the proposed search algorithms and compare their optimal solutions with exhaustive search. Then, we will show the performance of the proposed optimizations and compare it with a benchmark scheme.

Recall that the optimization problem P1 aims to find the minimum number of participating users and P2 aims to minimize the total energy consumption, solved by search algorithms UBS-1 and UBS-2, respectively. Table III and Fig. 10 are provided to compare the simulation results of the proposed search algorithms and exhaustive search. From a practical point of view, the prediction accuracy p is set to be larger than 0.6. It can be noticed that the obtained optimal solutions of UBS-1 and UBS-2 perfectly match with the exhaustive search results under different values of p. From Fig. 10, it can be observed that n_1 or n_2 does not always decrease with p. This is because the optimization aims to minimize the summation, i.e., $n_1 + n_2$. Furthermore, it can be noticed that n_1 and n_2 reduce to 1 when prediction accuracy becomes large. This is because CP can now easily meet the accuracy requirement hence the demand for performing spectrum sensing becomes relatively low when p is large.

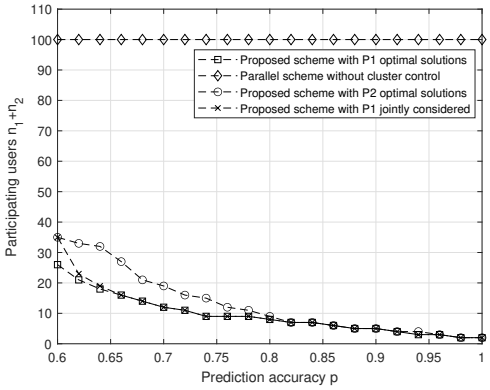
Due to the fact that the power consumption of performing sensing is normally larger than conducting

TABLE III: Optimal (n_1, n_2, σ) obtained using our algorithms and exhaustive search.

	ML accuracy p	0.6	0.7	0.8	0.9	1
For P1	UBS-1	(32,13,21)	(11,9,9)	(13,1,8)	(6,1,4)	(1,1,1)
	Exhaustive search	(32,13,21)	(11,9,9)	(13,1,8)	(6,1,4)	(1,1,1)
	Total users in clusters	42	20	14	7	2
For P2	UBS-2	(50,11,30)	(26,5,15)	(11,3,7)	(6,1,4)	(1,1,1)
	Exhaustive search	(50,11,30)	(26,5,15)	(11,3,7)	(6,1,4)	(1,1,1)
	Total users in clusters	66	31	14	7	2



(a) Total energy consumption versus p



(b) Total participated users versus p

Fig. 11: Performance comparison between the proposed scheme with optimal solutions and benchmark schemes.

learning, it can be observed from Table III that a smaller sensing cluster can reduce the power consumption significantly. In our proposed scheme, the probability of sensing cluster performing spectrum sensing is determined by n_1 and its threshold σ, while the energy consumption and the system accuracy are determined by $n_1, n_2,$ and σ. If the accuracy requirement is low, small values of n_1, n_2 will be preferentially chosen by our algorithm and the power consumption will be reduced significantly. However, in some extreme cases such as low prediction accuracy or high accuracy requirement scenarios, where more sensing users are needed to achieve the accuracy requirement, a very small sensing cluster size would not be feasible. In this case, our algorithm will pick a cluster size that is as small as possible while satisfying the accuracy requirements.

In order to demonstrate the performance of our proposed cluster-based scheme after applying cluster control and threshold optimizations, Figure 11 presents the total energy consumption and the number of participating users, while adopting optimal solutions to P1 and P2. Additionally, the figure also includes a parallel prediction-sensing scheme and another scheme where P1 and P2 are jointly considered. In the joint scheme, total energy consumption is considered as a constraint when minimizing the total participating users. In Figure 11, a total energy consumption limitation of 6000J and an accuracy requirement of 0.98 are assumed for the joint scheme. Similar to the scheme discussed in section V. B, the parallel scheme assumes a fixed cluster size with $n_1 = n_2 = 50$, and the optimal threshold that maximizes the system accuracy is used.

Figure 11(a) demonstrates that the energy consumption decreases with prediction accuracy p after adopting the optimal solutions for P1 and P2. The energy consumption obtained with P2 solutions is either less than or equal to that of P1, as P2 aims to minimize the total energy consumption while P1 focuses on minimizing the number of required users. Notably, the total energy consumption of our proposed cluster-based scheme is consistently lower than that of the parallel scheme. When the prediction accuracy is low, e.g., 0.65, by adopting our optimal P2 solution, the total energy consumption can be reduced by 87.5%. In contrast, the joint scheme serves as a compromise between P1 and P2, and as depicted in Figure 11(a), its energy consumption falls between P1 and P2, with closer proximity to P1 in the high prediction accuracy interval. Additionally, it can be observed from Figure 11(a) that the performance of using P1 solutions and the parallel scheme are similar when $p \in [0.72, 0.75]$. This is because n_2 needs to be

slightly larger than n_1 in order to satisfy the accuracy requirement, and the required energy limitation can be easily met, resulting in a sacrifice of total energy consumption when minimizing $n_1 + n_2$.

Figure 11(b) shows that the number of participating users decreases with p when the optimal solutions for P1 and P2 are adopted. The number of participating users obtained with optimal P1 solutions is not greater than that of P2, as P1 aims to minimize the number of participating users while P2 focuses on minimizing the total energy consumption. As depicted in Figure 11(b), the number of required users in our proposed cluster-based scheme is significantly lower than that of the parallel scheme. When the prediction accuracy is high, e.g., 0.9, by adopting our optimal P1 solution, the total participating users can be reduced by 95%. Similarly, in the joint scheme, which serves as a compromise between P1 and P2, it can be observed from Figures 11(a) and 11(b) that the number of participating users is sacrificed to meet the energy consumption requirement in a low prediction accuracy range. Furthermore, the optimal solutions for P1 and P2 provide similar performance in a high prediction accuracy range, as shown in Figure 11(b).

VI. Conclusions

This paper proposed a novel cluster-based sensing-after-prediction scheme with flexible cluster size where the learning cluster and sensing cluster are jointly considered/optimized. Theoretical analysis were conducted and analytical expressions were derived for system accuracy and energy consumption. With the aims of reducing the total number of participating users and total energy consumption while guaranteeing the system accuracy requirement and individual energy constraints, the optimization problems P1 and P2 were formulated. To solve it effectively, the unconstrained problems were mathematically solved first by relaxing the integer variable and fixing the cluster size, the analytical solutions of which serve as a foundation for solving the original optimization problems. Finally, two low-complexity search algorithms UBS-1 and UBS-2 were developed to solve the two optimization problems P1 and P2 and validated by comparing with exhaustive search. Simulation results validated the accuracy of the derived analytical expressions and demonstrated that the total energy consumption and the number of users contributing to learning and sensing can be greatly reduced by applying our proposed scheme. Furthermore, it is also shown that the proposed scheme outperforms benchmark schemes. Moreover, the simulation results revealed that optimizations P1 and P2 achieve a similar performance in high prediction accuracy region.

Appendix A
Proof of condition (17)

Note that σ_0^* needs to satisfy a feasible range $\left\lceil \frac{n_1+1}{2} \right\rceil \leq \sigma_0^* \leq n_1$, where $n_1 \in \mathbb{Z}$. After obtaining the expression of σ_0^* in (16), it can be substituted into the range, yielding

$$\left\lceil \frac{n_1+1}{2} \right\rceil \leq \left\lceil \frac{1}{2} \log_{\frac{p}{1-p}} \frac{Q_{\mathrm{S}}}{1-Q_{\mathrm{S}}} \left(\frac{p}{1-p} \right)^{n_1} \right\rceil \leq n_1. \tag{19}$$

Let us first focus on the first inequality, i.e., $\left\lceil \frac{n_1+1}{2} \right\rceil \leq \sigma_0^*$. When n_1 is odd, we have $\left\lceil \frac{n_1+1}{2} \right\rceil = \frac{n_1+1}{2}$, meaning that $\frac{n_1-1}{2} < \frac{n_1+1}{2} \leq \sigma_0^*$. Thus, we get $\log_{\frac{p}{1-p}} \frac{Q_{\mathrm{S}}}{1-Q_{\mathrm{S}}} \left(\frac{p}{1-p} \right)^{n_1} > n_1 - 1$. When n_1 is even, we have $\left\lceil \frac{n_1+1}{2} \right\rceil = \frac{n_1+2}{2}$, meaning that $\frac{n_1}{2} < \frac{n_1+2}{2} \leq \sigma_0^*$. Thus, we get $\log_{\frac{p}{1-p}} \frac{Q_{\mathrm{S}}}{1-Q_{\mathrm{S}}} \left(\frac{p}{1-p} \right)^{n_1} > n_1$.

For the second inequality,i.e., $\sigma_0^* \leq n_1$, by assuming $p \neq Q_{\mathrm{S}}$, the inequality can be written as $\frac{1}{2} \log_{\frac{p}{1-p}} \frac{Q_{\mathrm{S}}}{1-Q_{\mathrm{S}}} \left(\frac{p}{1-p} \right)^{n_1} \leq n_1$. It is known that the monotonicity of a logarithmic function depends on its base. For a function $\log_a x$, it is monotonically increasing if $a > 1$ and monotonically decreasing if $0 < a < 1$. Hence, we note that $\log_{\frac{p}{1-p}} \frac{Q_{\mathrm{S}}}{1-Q_{\mathrm{S}}} \left(\frac{p}{1-p} \right)^{n_1}$ is monotonically increasing if $p > 0.5$ and monotonically decreasing if $p < 0.5$. Finally, by combining all the results, we have (17).

Appendix B
Proof of Property 4 in Section IV

It can be confirmed from Property 1 that the $E(\sigma, n_1)$ monotonically increases or stays stable with the increase in σ. For each n_1, $E(\sigma, n_1)$ reaches its upper bound when $\sigma = n_1$. When $\sigma = n_1$, we have

$$E(n_1, n_1) = \left(1 - \left(p^{n_1} + (1-p)^{n_1}\right)\right) Le_S, \tag{20}$$

$$\frac{\partial E(n_1, n_1)}{\partial n_1} = -Le_S \left[p^{n_1} \ln p + (1-p)^{n_1} \ln(1-p)\right]. \tag{21}$$

Since $\ln p$ and $\ln(1-p)$ are always less than 0 when $0 < p < 1$, hence $\frac{\partial E(n_1, n_1)}{\partial n_1}$ is always larger than or equal to 0, which means that $E(n_1, n_1)$ will increase or stay stable when n_1 increases. As a result, $E(\sigma, n_1)$ reaches its upper bound of $E(n_1, n_1)$ when $\sigma = n_1$ and reaches its maximum value $E(N, N)$ when $\sigma = N$ and $n_1 = N$.

REFERENCES

[1] D. Nie, W. Yu, Q. Ni, and H. Pervaiz, "Optimization for prediction-driven cooperative spectrum sensing in cognitive radio networks," in *2022 IEEE International Conference on Communications Workshops (ICC Workshops)*, 2022, pp. 277–282.

[2] M. A. Uusitalo, P. Rugeland, M. R. Boldi, E. C. Strinati, P. Demestichas, M. Ericson, G. P. Fettweis, M. C. Filippou, A. Gati, M.-H. Hamon, M. Hoffmann, M. Latva-Aho, A. Pärssinen, B. Richerzhagen, H. Schotten, T. Svensson, G. Wikström, H. Wymeersch, V. Ziegler, and Y. Zou, "6g vision, value, use cases and technologies from european 6g flagship project hexa-x," *IEEE Access*, vol. 9, pp. 160 004–160 020, 2021.

[3] W. Lu, P. Si, G. Huang, H. Han, L. Qian, N. Zhao, and Y. Gong, "Swipt cooperative spectrum sharing for 6g-enabled cognitive iot network," *IEEE Internet of Things Journal*, vol. 8, no. 20, pp. 15 070–15 080, 2021.

[4] F. Tang, Y. Kawamoto, N. Kato, and J. Liu, "Future intelligent and secure vehicular network toward 6g: Machine-learning approaches," *Proceedings of the IEEE*, vol. 108, no. 2, pp. 292–307, 2020.

[5] W. Yu, C. H. Foh, A. U. Quddus, Y. Liu, and R. Tafazolli, "Throughput analysis and user barring design for uplink noma-enabled random access," *IEEE Transactions on Wireless Communications*, vol. 20, no. 10, pp. 6298–6314, 2021.

[6] S. Chen, B. Ren, Q. Gao, S. Kang, S. Sun, and K. Niu, "Pattern division multiple access—a novel nonorthogonal multiple access for fifth-generation radio networks," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 4, pp. 3185–3196, Apr. 2017.

[7] K. B. Letaief, Y. Shi, J. Lu, and J. Lu, "Edge artificial intelligence for 6g: Vision, enabling technologies, and applications," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 1, pp. 5–36, 2022.

[8] C. Bernardos, M. Uusitalo, C. Anton, D. Artuñedo, P. Demestichas, G. Fettweis, V. Frascolla, A. Kaloxylos, H. Koumaras, J. Magen, P. Merino, T. Norp, P. Rugeland, I. Tomkos, R. Aguiar, J. Chico, A. Bourdoux, D. Ayed, A. Azcorra, and Y. Zou, "European vision for the 6g network ecosystem," 06 2021.

[9] J. Mitola and G. Q. Maguire, "Cognitive radio: making software radios more personal," *IEEE Personal Commun.*, vol. 6, no. 4, pp. 13–18, Aug. 1999.

[10] M. Kaur and A. Kaur, "Cognitive radio spectrum sharing techniques: A review," *Int. J. Comput. Sci. Inf. Technol.*, vol. 6, no. 09756–9646, p. 3089–3090, 2015.

[11] W. Zhang, R. K. Mallik, and K. B. Letaief, "Optimization of cooperative spectrum sensing with energy detection in cognitive radio networks," *IEEE Trans. Wirel. Commun.*, vol. 8, no. 12, p. 5761–5766, 2009.

[12] W. Yu, A. U. Quddus, S. Vahid, and R. Tafazolli, "Opportunistic spectrum access in support of ultra-reliable and low latency communications," in *2018 IEEE Globecom Workshops (GC Wkshps)*, 2018, pp. 1–7.

[13] Q. Ni and C. Zarakovitis, "Nash bargaining game theoretic scheduling for joint channel and power allocation in cognitive radio systems," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 1, pp. 70–81, Jan 2012.

[14] Y. Pei, Y. C. Liang, K. C. Teh, and K. H. Li, "Sensing-throughput tradeoff for cognitive radio networks: A multiple-channel scenario," *IEEE Int. Symp. Pers. Indoor Mob. Radio Commun. PIMRC*, vol. 7, no. 4, p. 1326–1337, 2009.

[15] Y.-C. Liang, K.-C. Chen, G.-Y. Li, and P. Mahonen, "Cognitive radio networking and communications: an overview," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 7, pp. 3386 – 3407, Sept. 2011.

[16] T. Cody and P. A. Beling, "Heterogeneous transfer in deep learning for spectrogram classification in cognitive communications," in *2021 IEEE Cognitive Communications for Aerospace Applications Workshop (CCAAW)*, 2021, pp. 1–5.

[17] S. Aghabeiki, C. Hallet, N. E.-R. Noutehou, N. Rassem, I. Adjali, and M. Ben Mabrouk, "Machine-learning-based spectrum sensing enhancement for software-defined radio applications," in *2021 IEEE Cognitive Communications for Aerospace Applications Workshop (CCAAW)*, 2021, pp. 1–6.

[18] P. Nimudomsuk, M. Sanguanwattanaraks, K. Srisomboon, and W. Lee, "A performance comparison of spectrum sensing exploiting machine learning algorithms," in *2021 18th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, 2021, pp. 102–105.

[19] G. Ganesan and Y. G. Li, "Cooperative spectrum sensing in cognitive radio networks," *Proc. IEEE Symp. New Frontiers Dynamic Spectrum Access Networks (DySPAN'05)*, pp. 137–143, Nov 2005.

[20] M. E. Tanab and W. Hamouda, "Resource allocation for underlay cognitive radio networks: A survey," *IEEE Communications Surveys Tutorials*, vol. 9, no. 2, 2017.

[21] P. Chauhan, S. K. Deka, B. C. Chatterjee, and N. Sarma, "Cooperative spectrum prediction-driven sensing for energy constrained cognitive radio networks," *IEEE Access*, vol. 9, pp. 26 107 – 26 118, Feb. 2021.

[22] V.-D. Nguyen and O.-S. Shin, "Cooperative prediction-and-sensing-based spectrum sharing in cognitive radio networks," *IEEE Transactions on Cognitive Communications and Networking*, vol. 4, no. 1, Mar. 2018.

[23] M. Wellens, J. Riihijärvi, and P. Mähönen, "Empirical time and frequency domain models of spectrum use," *Physical Communication*, vol. 2, no. 1, pp. 10–32, 2009, cognitive Radio Networks: Algorithms and System Design.

[24] M. K. Giri and S. Majumder, "Extreme learning machine based cooperative spectrum sensing in cognitive radio networks," in *2020 7th International Conference on Signal Processing and Integrated Networks (SPIN)*, 2020, pp. 636–641.

[25] G. Hongjian, L. Yang, A. Chunyan, and H. Biyao, "A three-step cooperative spectrum sensing algorithm based on historical sensing information prediction," in *2021 IEEE 6th International Conference on Signal and Image Processing (ICSIP)*, 2021, pp. 1198–1203.

[26] V. C. Abdul Rahim and S. Chris Prema, "A combined spectrum prediction and sensing approach for cognitive radios," in *2021 IEEE 18th India Council International Conference (INDICON)*, 2021, pp. 1–6.

[27] L. Zhang and M. Jia, "Accurate spectrum prediction based on joint lstm with cnn toward spectrum sharing," in *2021 IEEE Global Communications Conference (GLOBECOM)*, 2021, pp. 1–6.

[28] N. Radhakrishnan, S. Kandeepan, X. Yu, and G. Baldini, "Soft fusion based cooperative spectrum prediction using lstm," in *2021 15th International Conference on Signal Processing and Communication Systems (ICSPCS)*, 2021, pp. 1–7.

[29] H. Wang, L. Lightfoot, and T. Li, "On pry-layer security of cognitive radio: Collaborative sensing under malicious attacks," *2010 44th Annual Conference on Information Sciences and Systems (CISS)*, Mar. 2010.

[30] N. Shaghluf and T. A. Gulliver, "Spectrum and energy efficiency of cooperative spectrum prediction in cognitive radio networks," *Wireless Netw.*, vol. 25, pp. 1–10, Mar 2018.