IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 18 April 2023, accepted 10 May 2023, date of publication 22 May 2023, date of current version 26 May 2023.

*Digital Object Identifier 10.1109/ACCESS.2023.3278260*

RESEARCH ARTICLE

# Performance Improvement of Random Access by Prioritizing Collisions

**YUNBAE KIM AND SEUNGKEUN PARK, (Member, IEEE)**
Radio Resource Research Section, Electronics and Telecommunications Research Institute (ETRI), Daejeon 34129, South Korea

Corresponding author: Yunbae Kim (doko9gum@etri.re.kr)

This work was supported by the Institute for Information & communications Technology Promotion (IITP) grant funded by the Korea government (MSIT) (2019-0-00964, Development of Incumbent Radio Stations Protection and Frequency Sharing Technology through Spectrum).

**ABSTRACT** The importance of unlicensed spectrum is highlighted in terms of the flexibility of network deployment for various services envisioned in 5G and beyond. Since listen before talk is mandatory for channel access in unlicensed spectrum and it causes an unavoidable waste of resources due to collisions, an efficient random backoff mechanism is required. In the existing backoff schemes that impose waiting penalties on collided packets, a degraded fairness performance is observed. In this work, we analyze how prioritizing collided packets can improve performance compared to existing schemes. To this end, we devise a random backoff scheme called the Collision Priority Backoff (CPB) under the concept of giving priority to collided packets. We apply Bayesian optimization to carefully determine channel access parameters of the CPB to maximize network throughput. Since the optimized access parameters require the number of stations in the network, we also devise an adaptive version of the CPB called the Adaptive CPB (ACPB). We deal with an environment where the number of stations changes as a switching bandit problem, and employ a variant of upper confidence bound policy in the ACPB. Various simulation results validate that the proposed backoff scheme shows high throughput and fairness performance.

**INDEX TERMS** Random access, backoff, Bayesian optimization, switching bandit, fairness, throughput.

## I. INTRODUCTION

5G and beyond envision the expansion of services to various industries along with advances in mobile broadband [1]. For industrial applications, it is essential to design a network that meets the requirements such as low latency, high reliability and security in critical communication scenarios [2]. Typically, these networks are deployed with specific sensors [3] or Internet of Things devices [4], [5]. Since mobile broadband service-oriented licensed spectrum limits the flexibility of network deployment for these various services, the unlicensed spectrum is expected to play a crucial role [6]. A prominent challenge of operating in the unlicensed spectrum is harmonious coexistence with prevalent networks such as Wi-Fi [7]. Listen Before Talk (LBT) is mandatory for channel access in unlicensed spectrum [8]. Unfortunately, LBT causes waste of resources due to inevitable collisions occurring in contention for channel access. Since the degradation of service quality due to such issues becomes exacerbated as the number of devices increases, the need for an efficient backoff mechanism is emerging [9].

A fundamental random backoff scheme applied in IEEE 802.11-based Wi-Fi system is the Binary Exponential Backoff (BEB). In the BEB, the size of the Contention Window (CW) within which a backoff counter value is randomly selected is doubled with every collision. Unlicensed LTE has adopted a very similar LBT scheme to that used in the IEEE 802.11, and this scheme is also considered as a baseline for unlicensed 5G operation [8]. In this scheme, the CW size doubles when the number of negative feedbacks of hybrid automatic repeat request process exceeds a threshold [10]. The IEEE 802.11ax amendment adopts an Orthogonal Frequency-Division Multiple Access (OFDMA) approach to improve throughput [11]. Unlike LTE or 5G, OFDMA operates on top of the legacy LBT scheme and is coordinated by Access Point (AP). In uplink, IEEE 802.11ax allows for random access

The associate editor coordinating the review of this manuscript and approving it for publication was Biju Issac.

through a procedure called OFDMA Backoff (OBO), in addition to scheduled access. The OBO procedure also adopts the approach of doubling the CW size after a transmission failure.

In a seminal paper [12], a widely accepted analytical model for the throughput performance of the BEB was provided, showing that high throughput can be achieved with an appropriate size of the initial CW. On the other hand, in terms of fairness, it has been pointed out that a drastic change in the size of CW in the BEB can lead to performance degradation [13], [14]. This unfairness issue arises from the fact that stations that fail to transmit take, on average, several times longer to successfully transmit compared to stations that succeed at once. Therefore, access technologies in unlicensed spectrum using such backoff schemes are inherently susceptible to unfairness issues, and the same applies to 802.11ax in the contention with legacy stations for triggering OFDMA transmissions or in the uplink access through OBO. In this regard, mitigating the long-wait penalty imposed on collided packets can be a natural approach to improving fairness. In [15], a backoff scheme called the Renewal Access Protocol (RAP) was proposed. In the RAP, each station selects a new backoff counter value according to a particular selection distribution, regardless of the packet transmission result. It was shown in [15] that the RAP using a Poisson selection distribution achieves high throughput and fairness if the distribution has an expectation value appropriate for the number of stations in the network. Collided packets in the RAP are not subject to the doubling penalty as in the BEB, but they still require at least twice the average time to succeed compared to packets that are successfully transmitted at once. So, it can be argued that performance can be further improved by prioritizing retransmission of collided packets.

In this work, we investigate how the performance is improved through a random backoff scheme that prioritizes collided packets. To this end, we devise a random backoff scheme called the Collision Priority Backoff (CPB). The CPB grants priority in a similar manner to the quality of service provisioning in IEEE 802.11e. An IEEE 802.11e station has parallel backoff entities, each of which corresponds to a specific access category (AC), and an AC with higher priority uses backoff parameters, such as small CW sizes, that enable relatively quick access [16]. A CPB station uses different channel access probabilities after a successful transmission and after a collision. Specifically, after a collision, a backoff counter value is selected to increase the access probability to induce fast and successful retransmission, thereby improving fairness.

A key element in designing the CPB is determining the channel access probabilities for high performance. Since the structure of the CPB itself is designed to enhance fairness by giving priority to collided packets, we focus on maximizing throughput. As mentioned in Section III, for the CPB, it is hard to obtain a tractable formula for throughput as in [12] and [15]. As an alternative way, we use the Bayesian optimization technique to find the throughput maximizing access probabilities. The optimal access probabilities depend on the network parameters and the number of stations in the network. We derive simple functions for optimal access probabilities using smoothing methods on Bayesian optimization results for various combinations of network parameter and the number of stations. We validate through simulation that the CPB operating with the access probabilities obtained from these functions shows the desired high performance.

As noted in [12], the network parameters that pertain to the function for the optimal access probabilities are determined within a physical layer. Therefore, we can consider the optimal access probabilities as functions of the number of stations in the network. In the literature, there have been extensive research works on optimizing throughput when the number of stations in the network is known in various random access schemes such as the RAP [15], $p$-persistent IEEE 802.11 [17], and the Idle Sense [18], including the BEB [12]. These works have necessarily been followed by studies estimating the number of stations that changes over time. The number of stations (or the throughput maximizing access parameters) can be estimated by observing network status such as the probability of transmission success or collision and the number of idle slots. Some works take advantage of accurately derived analytical formulas for estimation based on network state observations [17], [18], [19]. However, for high performance, these methods require that all stations agree to use the same tuned access parameters. The Access Mechanism with Optimal Contention Window (AMOCW) given in [20] is an improved version of the Idle Sense and is different from the above schemes. Each station using the AMOCW operates with different access parameters based on its own estimate, and still achieves good network performance. The adaptive version of the RAP called the A-RAP given in [21] is advantageous in complexity over the above techniques in that each station adjusts its access parameters with only its own transmission result. This adjustment method is similar to the BEB, the Quadratic Backoff (QB) [22], or the Exponential Increase and Exponential Decrease (EIED) [23]. The A-RAP estimates the number of stations as well as network state observation-based methods and shows performance comparable to the optimized RAP.

In this work, we also devise an adaptive version of the CPB, called the Adaptive CPB (ACPB). In the ACPB, as in [24], each station continuously estimates the range to which the number of stations belongs, and operates with access probabilities suitable for the estimated range. As observed in [24], this approach can lead to a reduction in estimate variability at the expense of some precision. The range estimation is based on the network status information collected by each station as in [20]. Reinforcement learning can be utilized in tasks that involve observing the time-varying network environment and adaptively determining appropriate actions based on it [25], [26]. We treat the range prediction procedure in the ACPB as a multi-armed bandit problem. Selecting a range and operating with the appropriate access probabilities

corresponds to selecting and playing an arm. The reward for the selected arm is designed to indicate whether the actual number of stations is included in the selected range. Further, we deal with a network environment where the number of stations varies over time as a switching bandit problem in which the reward distribution of each arm is non-stationary. The ACPB adopts some variants of Upper Confidence Bound (UCB) strategy against this time-varying environment. Simulation studies validate our proposed ACPB shows good performance comparable to the CPB with known number of stations in a time-varying network.

The main contributions of this work can be summarized as follows:

- We devise a random backoff scheme that prioritizes collided packets, called the CPB, and define key access parameters (Section II).
- We obtain the channel access probability values that maximize network throughput with the CPB through Bayesian optimization, and derive simple functions of the network parameters and the number of stations that represent these probabilities by applying smoothing. Furthermore, the performance improvement compared to the BEB scheme is verified through numerical analysis (Section III).
- We devise a practical form of the CPB by relaxing the assumption of an ideal situation envisioned by the concept of the CPB (Section IV).
- Leveraging the switching bandit learning theory, we propose an adaptive from of the CPB called the ACPB that operates adaptively in a network environment where the number of stations changes over time (Section V).

## II. DESCRIPTION OF THE CPB

In this section, we describe the CPB scheme. In this work, we consider a fully connected network consisting of homogeneous stations, all of which can sense the transmissions from other stations. All stations are assumed to be saturated, i.e., they always have packets to send, and an ideal channel condition is also assumed as in [12] and [15]. Throughout this work, various forms of the CPB scheme are presented, each with different assumptions, but one assumption applies in all cases. The assumption, called the *busy distinction*, is that all stations can distinguish whether the busy state of the network is due to a transmission success or a collision.[1]

### A. CONCEPT OF THE CPB

Here, we describe the concept of the CPB. With the CPB, the network alternates between two phases: *ordinary* and *special*. In an ordinary phase, all stations decrease their backoff counter values in idle slots as in legacy 802.11. However, when a collision occurs in an ordinary phase, the phase is altered to a special phase. In a special phase, only the collided stations can access the channel, and other stations not involved in the collision freeze their backoff counter values. In other words, the special phase guarantees contention only of the collided stations. The collided stations choose their new backoff counter values from a different distribution than in an ordinary phase to contend in the special phase. If the backoff counter selection distribution in the ordinary phase is set appropriately so that collisions rarely occur, the number of stations participating in a collision can be expected to be small. Therefore, in the special phase, designing a small CW size is advantageous in terms of throughput, and can improve fairness by allowing collided packets to be quickly retransmitted. If one of the colliding stations succeeds in retransmission, this station selects a new backoff counter value according to the backoff counter selection distribution for an ordinary phase and then freezes it until the remaining collided stations complete their successful retransmissions. When the last station among the collided stations has successfully retransmitted, the network enters an ordinary phase again. An example of the operation of the CPB is depicted in Fig. 1.

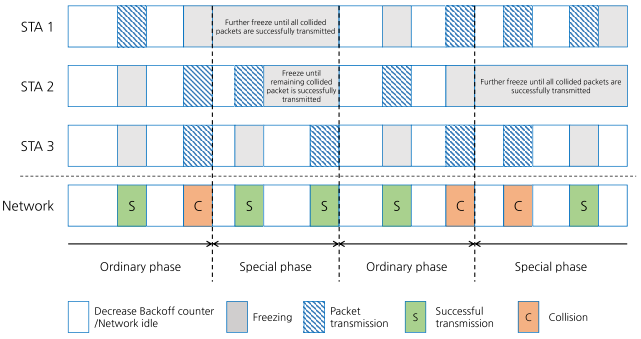



**FIGURE 1. An example of the operation of the CPB in a network consisting of three stations.**

### B. CHANNEL ACCESS PARAMETERS

Here, we formally define channel access parameters. As described above, we need to specify two kinds of backoff counter selection distributions: after a successful transmission and after a collision. The CPB takes advantage of the RAP [15], [21] in that a backoff counter value is chosen from a Poisson selection distribution. A Poisson selection distribution has a single parameter, the mean. Specifically, if the mean of a Poisson selection distribution is $\mu$, then the backoff counter value selected from it is given by Pois $(\mu - 1) + 1$ where Pois$(\mu - 1)$ is a typical Poisson random variable having the probability mass of $(\mu - 1)^k \exp(-\mu + 1)/k!$ for $k \geq 0$.

As in [12], [15], and [28], we consider the backoff process to evolve over slots, where a slot is a standardized time interval (e.g., 9 $\mu s$ in IEEE 802.11ac). In this system, the access probability of a station is defined by the average attempt rate. If the access probability is $\tau$ and the mean of the backoff selection distribution is $\mu$, then we have the relation of

[1] This distinction can be made using the detection of an ACK signal as in [27] or the duration of the network busy state in the case of adopting RTS/CTS mechanism.

$\tau = 1/\mu$ [21], [28]. Now let the access probabilities after a success and a collision be $\tau_s$ and $\tau_c$, respectively. In the following we describe how to find the optimal parameters $(\tau_s, \tau_c)$ for our goal.

## III. DETERMINATION OF OPERATION PARAMETERS

In this section, we find the optimal access probabilities defined above. The CPB structure itself is expected to improve fairness in the form of conceding the opportunity of stations that have already succeeded to collided stations, so we focus on maximizing throughput. Thus, our objective is to optimize the access parameters $(\tau_s, \tau_c)$ for the maximum network throughput.

In the literature, the *decoupling approximation* has been widely used in throughput analysis [12], [15], [28]. As addressed in [28] and [29], it can be expected that the decoupling approximation works when the number of stations accessing the channel is sufficiently large. In an ordinary phase, the decoupling approximation can be applied because all stations are contending for channel access, but in the special phase the situation is somewhat different. A special phase begins with a collision, which usually involves a small number of stations. Also, this number of colliding stations varies randomly. Therefore, network behavior in a special phase requires case-by-case analysis, and it is difficult to derive an exact formula for throughput performance. As an alternative way, we optimize the parameters based on simulation. However, naive approaches such as grid search are not suitable because it is expensive to obtain values through simulation. To tackle this issue, we use Bayesian optimization technique, which is a powerful tool for finding the maximum of an objective function that is difficult to evaluate [30].

### A. OBJECTIVE

Here we examine how to evaluate throughput performance in simulations and derive an objective function to optimize with access parameters.

For given parameters $(\tau_s, \tau_c)$, each station selects a new backoff counter value $\mathrm{Pois}(1/\tau_s - 1) + 1$ after a successful transmission, and selects $\mathrm{Pois}(1/\tau_c - 1) + 1$ after a collision as described in Section II. In one simulation run, let $M_e$, $M_s$, and $M_c$ be the number of idle slots, successful transmissions, and collisions observed in the network, respectively. According to the normalized network throughput defined as the fraction of time the channel is used to successfully transmit payload bits as given in [31], the throughput can be computed by

$$\frac{\sum_{i=1}^{M_s} T_P^{(i)}}{M_e + \sum_{j=1}^{M_c} T_C^{(j)} + \sum_{k=1}^{M_s} T_S^{(k)}}, \tag{1}$$

where $T_P^{(i)}$ is the time length in slots to transmit the payload of a packet, and $T_C^{(j)}$ and $T_S^{(k)}$ are time lengths in slots of busy periods for a collided packet transmission and a successful packet transmission, respectively. Since the stations are homogeneous, we can assume that each duration is independent and identically distributed. Typically, a simulation run observes a sufficiently large number of packet transmissions, we can approximate the throughput as

$$\begin{aligned}\frac{\sum_{i=1}^{M_s} T_P^{(i)}}{M_e + \sum_{j=1}^{M_c} T_C^{(j)} + \sum_{k=1}^{M_s} T_S^{(k)}} &\approx \frac{M_s \mathrm{E}[T_p]}{M_e + M_c \mathrm{E}[T_c] + M_s \mathrm{E}[T_s]} \\ &= \mathrm{E}[T_p] \frac{1}{\frac{M_e + M_c \mathrm{E}[T_c]}{M_s} + \mathrm{E}[T_s]},\end{aligned} \tag{2}$$

where $\mathrm{E}[T_p]$, $\mathrm{E}[T_c]$, and $\mathrm{E}[T_s]$ are average time lengths for the transmissions of the payload of a packet, a collided packet transmission, and a successful packet transmission, respectively. As given in the literature [12], [15], $\mathrm{E}[T_c]$ and $\mathrm{E}[T_s]$ depend on $\mathrm{E}[T_p]$, the time lengths for transmitting headers or ACK frames, and other protocol parameters such as SIFS and DIFS.

Let the number of stations $N$ be given. Also, $\mathrm{E}[T_p]$, $\mathrm{E}[T_s]$, and $\mathrm{E}[T_c]$ are determined by network parameters for the simulation. Then, in view of (2), maximizing throughput is equivalent to maximizing

$$\frac{M_s}{M_e + M_c \mathrm{E}[T_c]},$$

which depends only on $N$ and $\mathrm{E}[T_c]$. In sum, given network parameters containing $\{N, \mathrm{E}[T_c]\}$, a simulation run with the pair of access probabilities $(\tau_s, \tau_c) \in (0, 1)^2$ yields a realization of the objective function

$$f(\tau_s, \tau_c) = \frac{M_s}{M_e + M_c \mathrm{E}[T_c]}. \tag{3}$$

Then the optimization problem is to find

$$\left(\tau_s^*(N, \mathrm{E}[T_c]), \tau_c^*(N, \mathrm{E}[T_c])\right) = \underset{(\tau_s, \tau_c) \in (0,1)^2}{\mathrm{argmax}} f(\tau_s, \tau_c).$$

### B. REVIEW OF BAYESIAN OPTIMIZATION

Here, we briefly review the Bayesian optimization approach with the notations in this work. The definitions and results given here follow [30]. We use a subscript $s : t$ for integers $s \le t$ to denote a sequence of data, for e.g., $y_{1:t} = \{y_1, \ldots, y_t\}$.

Let $\boldsymbol{\tau} = (\tau_s, \tau_c)$. We are interested in finding $\boldsymbol{\tau}^* = \mathrm{argmax}_{\boldsymbol{\tau}} f(\boldsymbol{\tau})$. The objective $f$ is expensive to evaluate in that an evaluation takes a considerable amount time (e.g., due to running simulations). Bayesian optimization can be utilized in such a restricted environment. The Bayesian model effectively utilizes informative priors to describe properties of an objective function, such as the most probable locations of a maximum. Let $P(f)$ be a prior distribution representing the space of possible objective functions. As we accumulate observations $\mathcal{D}_{1:t} = \{\boldsymbol{\tau}_{1:t}, f_{1:t}\}$ where $f_i$ is the evaluation of the objective at a point $\boldsymbol{\tau}_i$, the prior is combined with the likelihood function $P(\mathcal{D}_{1:t}|f)$. Then we obtain the posterior distribution:

$$P(f|\mathcal{D}_{1:t}) \propto P(\mathcal{D}_{1:t}|f)P(f).$$

This posterior distribution is exploited to construct an easy-to-evaluate acquisition function, and the point that maximizes it is determined as the next evaluation location $\boldsymbol{\tau}_{t+1}$.

Although various models can be used for prior, Gaussian Process (GP) priors have been widely used and shown to be well-suited to the task. The GP is defined by the property that any finite combination of points induces a multivariate Gaussian distribution. As a Gaussian distribution is completely specified by its mean and covariance, a GP is completely specified by the mean function $m$ and covariance function $k$:

$$f(\boldsymbol{\tau}) \sim \mathcal{GP}\big(m(\boldsymbol{\tau}), k(\boldsymbol{\tau}, \boldsymbol{\tau}')\big).$$

Then, for any finite points $\{\boldsymbol{\tau}_1, \boldsymbol{\tau}_2, \ldots, \boldsymbol{\tau}_n\}$, we have

$$\begin{bmatrix} f(\boldsymbol{\tau}_1) \\ f(\boldsymbol{\tau}_2) \\ \vdots \\ f(\boldsymbol{\tau}_n) \end{bmatrix} \sim \mathcal{N}\left( \begin{bmatrix} m(\boldsymbol{\tau}_1) \\ m(\boldsymbol{\tau}_2) \\ \vdots \\ m(\boldsymbol{\tau}_n) \end{bmatrix}, \begin{bmatrix} k(\boldsymbol{\tau}_1, \boldsymbol{\tau}_1) & \cdots & k(\boldsymbol{\tau}_1, \boldsymbol{\tau}_n) \\ k(\boldsymbol{\tau}_2, \boldsymbol{\tau}_1) & \cdots & k(\boldsymbol{\tau}_2, \boldsymbol{\tau}_n) \\ \vdots & \ddots & \vdots \\ k(\boldsymbol{\tau}_n, \boldsymbol{\tau}_1) & \cdots & k(\boldsymbol{\tau}_n, \boldsymbol{\tau}_n) \end{bmatrix} \right),$$

where $\mathcal{N}(\boldsymbol{\mu}, \Sigma)$ denotes the multivariate Gaussian distribution with the mean vector $\boldsymbol{\mu}$ and the covariance matrix $\Sigma$. Usually, the mean function $m(\boldsymbol{\tau})$ is chosen to be zero or a constant, but the covariance function is treated more technically. In this work, we also assume a zero mean function, i.e., $m(\boldsymbol{\tau}) = 0$. In fact, the power of GP to express a rich distribution on functions rests solely on the covariance function. In this work, we apply the Squared Exponential (SE) covariance function, which is probably the most widely used [32]. The SE covariance function is given by

$$k(\boldsymbol{\tau}, \boldsymbol{\tau}'; \boldsymbol{\theta}) = \theta_1^2 \exp\left( -\frac{\|\boldsymbol{\tau} - \boldsymbol{\tau}'\|^2}{2\theta_2^2} \right),$$

with nonnegative hyperparameters $\boldsymbol{\theta} = (\theta_1, \theta_2)$.

In Bayesian optimization, one can set the hyperparameter values of the covariance function by maximizing the marginal likelihood. For given observations $\mathcal{D}_{1:t} = \{\boldsymbol{\tau}_{1:t}, f_{1:t}\}$, the GP model gives

$$\mathbf{f} \sim \mathcal{N}(\mathbf{0}, \mathbf{K}(\boldsymbol{\theta})),$$

where $\mathbf{f} = [f_1 \ \cdots \ f_t]^\top$ and

$$\mathbf{K}(\boldsymbol{\theta}) = \begin{bmatrix} k(\boldsymbol{\tau}_1, \boldsymbol{\tau}_1; \boldsymbol{\theta}) & \cdots & k(\boldsymbol{\tau}_1, \boldsymbol{\tau}_t; \boldsymbol{\theta}) \\ k(\boldsymbol{\tau}_2, \boldsymbol{\tau}_1; \boldsymbol{\theta}) & \cdots & k(\boldsymbol{\tau}_2, \boldsymbol{\tau}_t; \boldsymbol{\theta}) \\ \vdots & \ddots & \vdots \\ k(\boldsymbol{\tau}_t, \boldsymbol{\tau}_1; \boldsymbol{\theta}) & \cdots & k(\boldsymbol{\tau}_t, \boldsymbol{\tau}_t; \boldsymbol{\theta}) \end{bmatrix}. \quad (4)$$

And, we have the log of the marginal likelihood as

$$\log p(f_{1:t} | \boldsymbol{\tau}_{1:t}; \boldsymbol{\theta}) = -\frac{1}{2} \mathbf{f}^\top \mathbf{K}(\boldsymbol{\theta})^{-1} \mathbf{f} - \frac{1}{2} \log(\det(\mathbf{K}(\boldsymbol{\theta}))) - \frac{t}{2} \log 2\pi.$$

Then, as a maximum likelihood estimate, we set $\boldsymbol{\theta}$ to the value that maximizes the log-likelihood

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\operatorname{argmax}} \log p(f_{1:t} | \boldsymbol{\tau}_{1:t}; \boldsymbol{\theta}). \quad (5)$$

**TABLE 1. A pseudo-code for Bayesian optimization.**

| |
|---|
| Observe the objective function evaluation $f_1$ at point $\boldsymbol{\tau}_1$. |
| Set the data set $\mathcal{D}_{1:1} = \{\boldsymbol{\tau}_1, f_1\}$. |
| **for** $t = 2, 3, \ldots, T$ **do** |
| Update the GP using $\mathcal{D}_{1:t}$. |
| Update the posterior distribution on $f$ as in (6). |
| Let $\boldsymbol{\tau}_{t+1}$ be the maximizer of the acquisition function. |
| Observe $f_{t+1}$ at $\boldsymbol{\tau}_{t+1}$. |
| Augment the data $\mathcal{D}_{t+1} = \mathcal{D}_t \cup \{\boldsymbol{\tau}_{t+1}, f_{t+1}\}$. |
| Update the GP and posterior using $\mathcal{D}_{1:T}$. |
| Return the point with the largest posterior mean as a solution. |

Now we derive the posterior. Let a sequence of observations $\mathcal{D}_{1:t} = \{\boldsymbol{\tau}_{1:t}, f_{1:t}\}$ be given. Also, suppose that the covariance function $k$ is specified, i.e., $k(\boldsymbol{\tau}, \boldsymbol{\tau}') = k(\boldsymbol{\tau}, \boldsymbol{\tau}'; \hat{\boldsymbol{\theta}})$ with $\hat{\boldsymbol{\theta}}$ in (5). For an arbitrary point $\tilde{\boldsymbol{\tau}}$, the GP prior model gives a joint Gaussian distribution of $\mathbf{f}$ and $\tilde{f} = f(\tilde{\boldsymbol{\tau}})$ as

$$\begin{bmatrix} \mathbf{f} \\ \tilde{f} \end{bmatrix} \sim \mathcal{N}\left( \mathbf{0}, \begin{bmatrix} \mathbf{K}_t & \mathbf{k}_t \\ \mathbf{k}_t^\top & k(\tilde{\boldsymbol{\tau}}, \tilde{\boldsymbol{\tau}}) \end{bmatrix} \right),$$

where $\mathbf{K}_t = \mathbf{K}(\hat{\boldsymbol{\theta}})$ from (4) and $\mathbf{k}_t = \begin{bmatrix} k(\boldsymbol{\tau}_1, \tilde{\boldsymbol{\tau}}) & \cdots & k(\boldsymbol{\tau}_t, \tilde{\boldsymbol{\tau}}) \end{bmatrix}^\top$. Then the posterior distribution is given by $\tilde{f} | \mathcal{D}_{1:t}, \tilde{\boldsymbol{\tau}} \sim \mathcal{N}\left(\mu_t(\tilde{\boldsymbol{\tau}}), \sigma_t^2(\tilde{\boldsymbol{\tau}})\right)$ where

$$\begin{aligned} \mu_t(\tilde{\boldsymbol{\tau}}) &= \mathbf{k}_t^\top \mathbf{K}_t^{-1} \mathbf{f}, \\ \sigma_t^2(\tilde{\boldsymbol{\tau}}) &= k(\tilde{\boldsymbol{\tau}}, \tilde{\boldsymbol{\tau}}) - \mathbf{k}_t^\top \mathbf{K}_t^{-1} \mathbf{k}_t. \end{aligned} \quad (6)$$

Recall that the posterior distribution is used to construct the acquisition function for determining the next evaluation point. The acquisition function is defined such that a high acquisition corresponds to a potentially high value of the objective function due to high predicted value, high uncertainty, or both. In this work, we apply a UCB-based acquisition function. With the expressions in (6), the acquisition function has the form of

$$\mathcal{A}(\boldsymbol{\tau}; \mathcal{D}_t) = \mu_t(\boldsymbol{\tau}) + \kappa \sigma_t(\boldsymbol{\tau}),$$

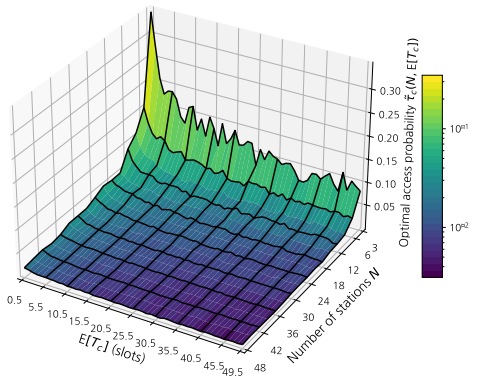
with a tunable $\kappa > 0$ to balance exploitation against exploration. Once the acquisition function is set, the next location $\boldsymbol{\tau}_{t+1}$ to sample is determined as follows:

$$\boldsymbol{\tau}_{t+1} = \underset{\boldsymbol{\tau}}{\operatorname{argmax}} \mathcal{A}(\boldsymbol{\tau}; \mathcal{D}_t).$$

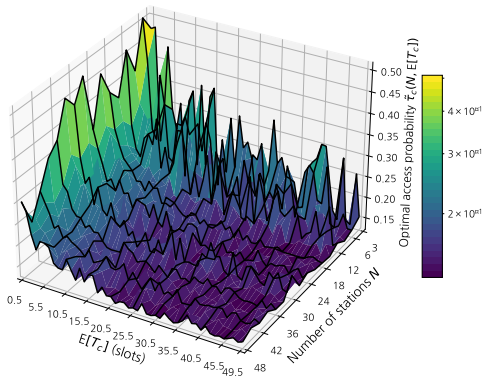
When the budget for a given $T$ function evaluations is exhausted, we can get a final posterior mean function $\mu_T(\boldsymbol{\tau})$. Then, the value that maximizes this posterior mean is proposed as an estimate for $\boldsymbol{\tau}^*$. The series of Bayesian optimization procedures mentioned here are summarized in Table 1.

### C. OPTIMIZATION RESULTS

Here we present the access probabilities obtained by applying the Bayesian optimization technique discussed above. We use the 'gaussian_process' module in Python's scikit-learn to perform the optimization. In the optimization process, we set the UCB acquisition parameter $\kappa$ to 1 and the function evaluation budget $T$ to 20. In order to derive the optimal access probabilities $\boldsymbol{\tau}^* = (\tau_s^*, \tau_c^*)$ for a wide range of cases, we apply




(a) $\tau_s^*$




(b) $\tau_c^*$

**FIGURE 2.** Optimal access probabilities from Bayesian optimization.

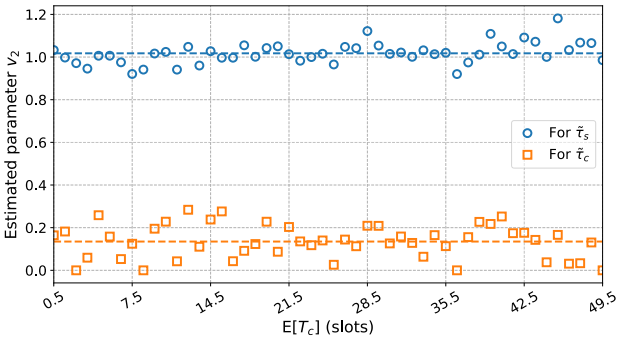



**FIGURE 3.** Estimated parameter $v_2$ when the optimal probability value from Bayesian optimization is fitted by $V$. The dotted line indicates the mean of estimated values for $v_2$.

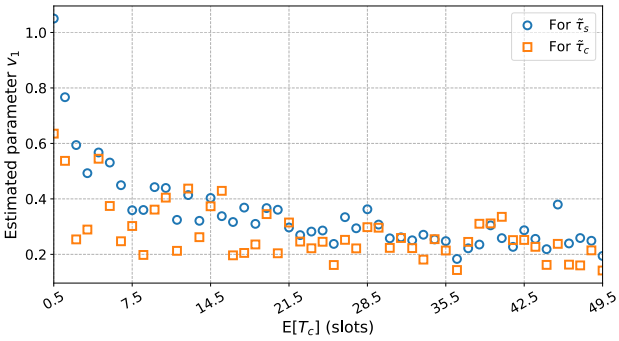



**FIGURE 4.** Estimated parameter $v_1$ when the optimal probability value from Bayesian optimization is fitted by $V$.

the optimization process to each $(N, \mathrm{E}[T_c])$ combination for $N \in \{3, 6, 9, \ldots, 48\}$ and $\mathrm{E}[T_c] \in \{0.5, 1.0, 1.5, \ldots, 49.5\}$. For a combination $(N, \mathrm{E}[T_c])$, each function evaluation at a pair of access probabilities $\boldsymbol{\tau} = (\tau_s, \tau_c)$ is obtained by calculating (3) with $M_s$, $M_c$, and $M_e$ observed in a simulation run over $10^5$ slots. We set the first observation point $\boldsymbol{\tau}_1$ to $(0.1, 0.1)$. In maximizing the acquisition function and finding the largest posterior mean at the last step, we simply apply a grid search method.

The optimization results are shown in Fig. 2. As can be seen in the figure, the optimal access probabilities tend to decrease as $N$ or $\mathrm{E}[T_c]$ increases. However, since the results are wiggly and unstable, we fit them with smooth monotonic curves. The results obtained via Bayesian optimization for each $(N, \mathrm{E}[T_c])$ are denoted as $\tilde{\tau}_s(N, \mathrm{E}[T_c])$ and $\tilde{\tau}_c(N, \mathrm{E}[T_c])$ to distinguish them from the optimal access probabilities to be used in practice.

At first, we consider the optimal access probabilities for given $\mathrm{E}[T_c]$ as functions of $N$. In [12], [28], and [33], it is shown that the optimal access probability for LBT with random backoff approximately inversely proportional to the number of stations $N$. Using this fact, we consider a one-term power series model $V$ as given by

$$V(N; v_1, v_2) = v_1 N^{-v_2}, \tag{7}$$

where $v_1, v_2 > 0$. The parameters $v_1$ and $v_2$ are estimated to minimize the residual sum of squares. Note from Fig. 2 that the higher the probability value, the greater the variability. To mitigate the expected heteroscedasticity, residual sum of squares minimization is applied after log transformation. Specifically, for a given $\mathrm{E}[T_c]$ and a phase $ph \in \{s, c\}$, the values of $v_1$ and $v_2$ are chosen to minimize $\sum_{n=1}^{16}\{\log(\tilde{\tau}_{ph}(3n, \mathrm{E}[T_c])) - \log(v_1) + v_2 \log(3n)\}^2$. Fig. 3 shows the estimated values of parameter $v_2$ in (7). It can be observed that the estimated values are randomly spread around the mean with no discernible trend with respect to $\mathrm{E}[T_c]$. Thus, we assume that the parameter $v_2$ is independent of the values of $\mathrm{E}[T_c]$. Meanwhile, Fig. 4 shows the estimated values of the parameter $v_1$ in (7). In the figure, we can observe that the estimated values for $v_1$ show convex decreasing trends with respect to $\mathrm{E}[T_c]$. Here we consider a two-term power series model to fit $v_1$ as a function of $\mathrm{E}[T_c]$. Taking these together, we fit the optimal probability values from Bayesian optimization with a surface $U$ given by $U(N, \mathrm{E}[T_c]; u_1, u_2, u_3, u_4) = u_4 N^{-u_1}\left(\mathrm{E}[T_c]^{-u_2} + u_3\right)$ with non-negative parameters $u_1$, $u_2$, $u_3$, and $u_4$. As above, the parameter values are estimated to minimize the sum

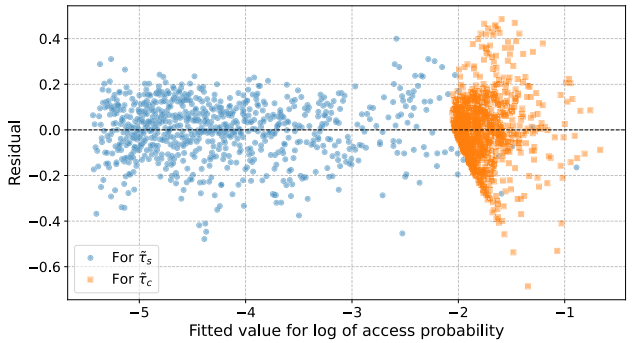



**FIGURE 5.** Residual plot for fitting the log of the optimal probabilities from Bayesian optimization.

**TABLE 2.** Estimated parameters of fitting surfaces for optimal access probabilities.

| Access probability | Parameters | | | |
|---|---|---|---|---|
| | $u_1$ | $u_2$ | $u_3$ | $u_4$ |
| $\tau_s$ | 1.018 | 0.375 | 0 | 0.969 |
| $\tau_c$ | 0.134 | 0.314 | 0.242 | 0.401 |

of squares of the residuals after log transformation. Fig. 5 shows the residual plot for this surface fitting for the log of the optimal probability values from Bayesian optimization. In the figure, we can check the validity of the smoothing surface by observing that most of residuals form roughly a horizontal band and bounce randomly around the zero line. The estimated values for the parameters $u_1$, $u_2$, $u_3$, and $u_4$ are given in Table 2.

In sum, the value we propose as the optimal access probability for each phase is given by the following function:

$$\tau^*_{ph}(N, \mathrm{E}[T_c]) = u_4 N^{-u_1}(\mathrm{E}[T_c]^{-u_2} + u_3), \tag{8}$$

where $ph \in \{s, c\}$ and the parameters $u_1$, $u_2$, $u_3$, and $u_4$ are given in Table 2.

### D. NUMERICAL ANALYSIS

Now we numerically evaluate the throughput and fairness performances of the network to which the CPB with the optimal access probabilities of (8) is applied. Performance evaluations are done through simulation. We use Python to simulate a network with network parameters given in Table 3. For each numerical result, we take the average of 20 runs where each run observes network behavior for $10^6$ slots. In [15], it is shown that the RAP with a Poisson selection distribution having an optimized mean value maximizes throughput given the number of stations $N$. Also, in terms of fairness, the RAP is shown to outperform not only the legacy IEEE 802.11 but also access methods such as the Idle Sense [18] proposed in the literature to improve it. In this regard, we use the RAP's performance optimized for the number of stations as a benchmark for performance comparison. To examine the performance improvement of the CPB compared to the BEB, a recently proposed BEB scheme [34],

**TABLE 3.** Network parameters.

| Parameter | Unit | Value |
|---|---|---|
| Payload | bits | 8184 |
| PHY header | | 128 |
| MAC header | | 272 |
| RTS frame | | PHY header + 160 |
| CTS frame | | PHY header + 112 |
| ACK frame | | PHY header + 112 |
| Time slot | $\mu s$ | 9 |
| SIFS | | 16 |
| DIFS | | 34 |
| Propagation delay | | 1 |
| Data rate | Mbps | 5, 50 |
| $N$ | | 6, 12, 24, 48 |

**TABLE 4.** Network throughput performance.

| Data rate [Mbps] | $N$ | Backoff scheme | | | | |
|---|---|---|---|---|---|---|
| | | CPB | RAP | [34] | PCPB | |
| | | | | | $p_e = 0.1$ | $p_e = 0.01$ |
| 5 | 6 | 0.817 | 0.816 | 0.816 | 0.817 | 0.817 |
| | 12 | 0.816 | 0.815 | 0.814 | 0.816 | 0.816 |
| | 24 | 0.816 | 0.814 | 0.813 | 0.816 | 0.815 |
| | 48 | 0.815 | 0.814 | 0.813 | 0.815 | 0.815 |
| 50 | 6 | 0.531 | 0.526 | 0.53 | 0.531 | 0.53 |
| | 12 | 0.529 | 0.523 | 0.525 | 0.528 | 0.526 |
| | 24 | 0.527 | 0.521 | 0.522 | 0.526 | 0.523 |
| | 48 | 0.525 | 0.52 | 0.521 | 0.524 | 0.522 |

which is optimized for a known number of stations, is also considered for comparison. For this scheme, we assume that all packets belong to the highest priority AC.

The throughput results are given in Table 4. For a simulation run, the throughput performance is computed by the formula in (1). As shown in [15] and [28], the throughput performance of the contention-based backoff scheme depends on the expectations of the backoff counter selection distributions. Furthermore, the maximized throughput performance is not dependent on the number of stations in the network. Table 4 shows the results of using the backoff counter selection distribution with optimized expectation in each scheme, so similar performance is obtained in all cases, as mentioned above. In the case of the CPB, it can be observed that the special phase only for collided packets, which is different from the other two schemes, has a slight advantage on throughput.

As in [15] and [35], we use Jain's fairness index [36] to measure fairness performance. In addition, the sliding window method [37] is used to assess the fairness over time. In this method, a window of size $w$ packets is slid across a sequence of successful transmissions. In a window, let $s_n$ for $1 \le n \le N$ be the number of packets successfully transmitted by station $n$ within this window. Then Jain's fairness index within this window is computed as

$$\frac{\left(\sum_{n=1}^{N} s_n\right)^2}{N \sum_{n=1}^{N} s_n^2}.$$

After sliding the window through the whole sequence, averaging the fairness indices of all windows provides the fairness metric. For a fair comparison of networks with

different number of stations, we use a normalized window size as in [35]. When a network consists of $N$ stations, the normalized window $W$ means that one sliding window monitors $WN$ successful transmissions. Fig. 6 shows fairness performances. In fairness, the differences between the backoff schemes are clearly manifested. As expected, the BEB-based scheme in [34] that imposes the largest waiting penalty on collided packets exhibits degraded performance. In the case of the RAP, moderate penalties for collided packets lead to moderate level of fairness performance. Meanwhile, the CPB, which grants priority to collided packets for prompt retransmission, demonstrates the highest performance. Even in terms of short-term fairness, which corresponds to small window sizes $W = 1$ or 2, the CPB achieves a quite high fairness.

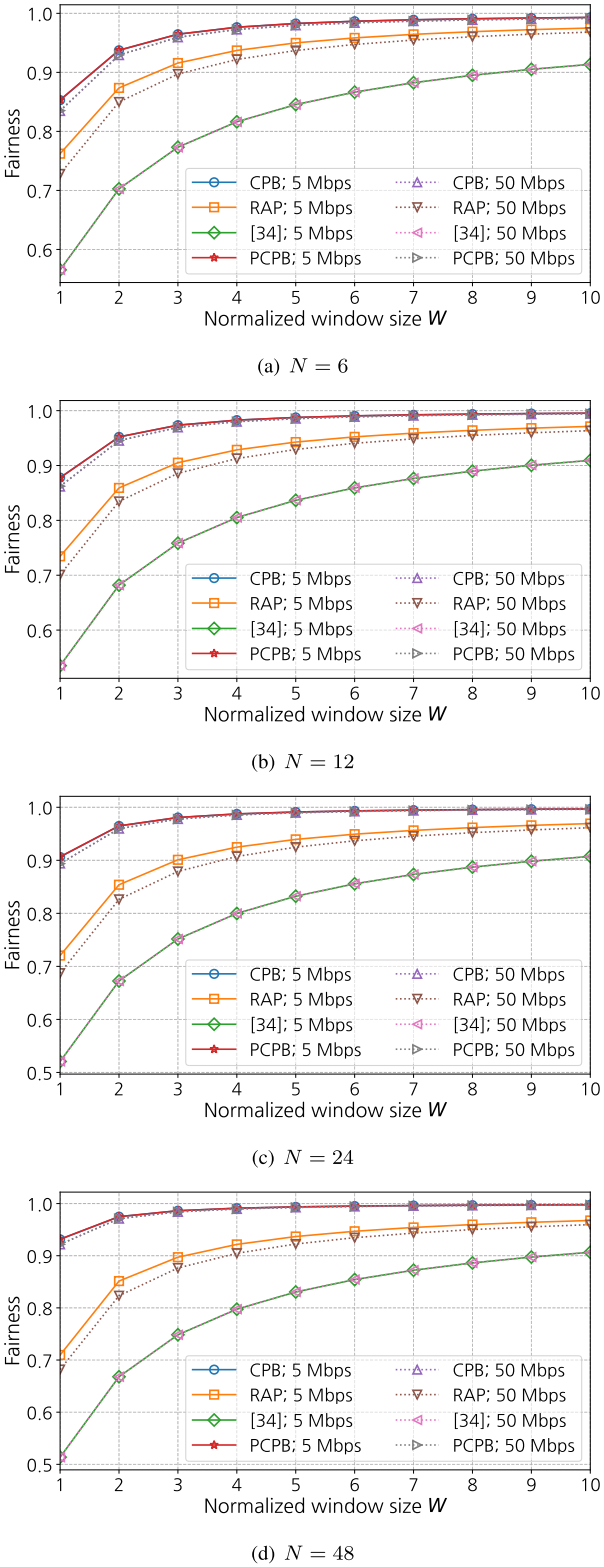
## IV. PRAGMATIC IMPLEMENTATION OF THE CPB

In this section, we provide a practical version of the CPB called the *Pragmatic CPB* (PCPB) by relaxing one of the impractical assumptions on which the CPB is based. Recall that a network adopting the CPB alternates ordinary and special phases. When the network is in an ordinary phase, all stations are able to detect collisions under the busy distinction assumption, so they can recognize the transition to a special phase. However, it is difficult to know how many stations are involved in a collision, and it is even harder to determine when the special phase ends. In this regard, the PCPB removes the assumption that all stations can be aware of the end of the special phase.

As noted in Section III, the decoupling approximation can be applied in the ordinary phase. As in [12] and [15], the number $N_c$ of collided packets involved in a collision can be modeled to follow the probability distribution $P[N_c = v] = \binom{N}{v}(\tau_s^*)^v(1-\tau_s^*)^{N-v}/\{1-(1-\tau_s^*)^N - N\tau_s^*(1-\tau_s^*)^{N-1}\}$ for $2 \le v \le N$. As shown in [12], a collision takes at least the DIFS interval, and DIFS duration is calculated by SIFS + (2 × Slot time) [38], so we have $\mathrm{E}[T_c] \ge 2$. In Fig. 7, we can see that two packets are most likely to cause a collision, and three or fewer packets cause most collisions. This becomes clearer as the value of $\mathrm{E}[T_c]$ increases. From this observation, we assume that collisions are only caused by three or fewer packets, and the probability values for $N_c$ are modified as

$$P(N_c=2) = \frac{\binom{N}{2}(\tau_s^*)^2(1-\tau_s^*)^{N-2}}{1-(1-\tau_s^*)^N - N\tau_s^*(1-\tau_s^*)^{N-1}},$$
$$P(N_c=3) = \frac{\sum_{v=3}^{N}\binom{N}{v}(\tau_s^*)^v(1-\tau_s^*)^{N-v}}{1-(1-\tau_s^*)^N - N\tau_s^*(1-\tau_s^*)^{N-1}}. \quad (9)$$

In the PCPB, every station considers the network to enter a special phase whenever a collision is detected. After entering the special phase, all stations consider this phase to continue until at least two consecutive successful transmissions have been made in the network. Stations not involved in the collision defer decrementing their backoff counters during this period, and the station involved in the collision that successfully retransmitted first also freezes its new backoff counter until the next successful transmission. If the special phase is due to a collision of three packets, this phase should continue after two consecutive successful transmissions until




(a) $N = 6$

(b) $N = 12$

(c) $N = 24$

(d) $N = 48$

**FIGURE 6. Fairness performances of the CPB, RAP, BEB-based scheme in [34], and the PCPB with $p_e = 0.1$.**

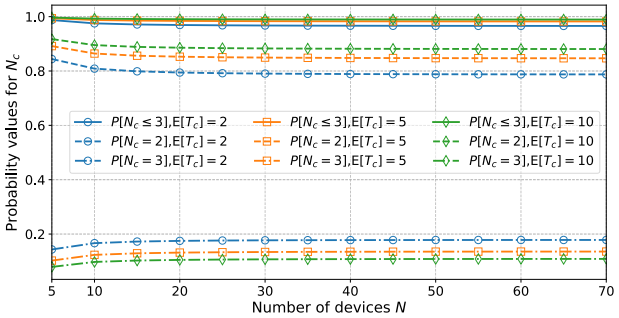



**FIGURE 7.** Probability values for the number of packets involved in a collision.

the last colliding packet is successfully retransmitted. Additional freezing after two consecutive successful transmissions is performed according to a certain criterion described below.

In the PCPB, each station is given an indicator $i_s$ of whether the current network is considered to be in a special phase, and its value is set to 1 whenever a collision is detected. Also, each packet is given an indicator $i_c$ of whether it has suffered a collision. Stations involved in the collision set the $i_c$ values of the corresponding packets to 1. As mentioned in III-A, the collided packets contend for retransmission with backoff counter values selected from a distribution of Pois $(1/\tau_c^* - 1) + 1$. Let $B_v^{(u)}$ be the $u$th largest of the $v$ backoff counter values independently selected in the special phase for $1 \le u \le v$, $v = 2, 3$. When two consecutive successful transmissions are observed after a collision, it can be confirmed that there are no ties among the selected backoff counter values, i.e., $B_2^{(1)} < B_2^{(2)}$ or $B_3^{(1)} < B_3^{(2)} < B_3^{(3)}$. Also, at this time, the special phase is terminated if the number of packets causing the collision is 2, but otherwise, it should be continued for $B_3^{(3)} - B_3^{(2)}$ idle slots. In the PCPB, after observing two consecutive successful transmissions, stations with no packets to retransmit further freeze their backoff counter for a number of slots determined by a specific criterion. Let $n_s$ and $d_s$ be the numbers of consecutive successful transmissions in the network and elapsed idle slots, respectively, after a collision. Also, we let $N_a$ be the number of packets involved in the last collision. When two consecutive successful transmissions is observed ($n_s = 2$) at $d_s = l$, with a given threshold $p_e \in (0, 1)$, the number $v_a(l)$ of slots for additional observation is set as follows:

$$v_a(l) = \max\{v \ge 1 : P(N_a = 3, B_3^{(3)} \ge l + v | n_s = 2, d_s = l) > p_e\}, \tag{10}$$

with the convention that $\max \emptyset = 0$. Then, stations with no packets to retransmit freeze the backoff counter up to $v_a(l)$ slots. If the network remains idle for $v_a(l)$ slots, then the network is considered to have entered an ordinary phase. On the other hand, if a successful transmission is observed before $v_a(l)$ slots, it is immediately determined that the special phase ends at this point. Therefore, the larger $p_e$ is, the more aggressively stations contend for transmission.

For stations to operate using the PCPB, probability values in (10) are required. The probability in (10) can be written as

$$P\left(N_a = 3, B_3^{(3)} \ge l + v | n_s = 2, d_s = l\right) = \frac{P(n_s = 2, d_s = l, N_a = 3, B_3^{(3)} \ge l + v)}{P(n_s = 2, d_s = l)}. \tag{11}$$

Note that the denominator and numerator in (11) are further written as

$$P(n_s = 2, d_s = l) = P\left(B_2^{(2)} = l | N_a = 2\right) P(N_a = 2) + P\left(B_3^{(2)} = l | N_a = 3\right) P(N_a = 3), \tag{12}$$

and

$$P(n_s = 2, d_s = l, N_a = 3, B_3^{(3)} \ge l + v) = P\left(B_3^{(2)} = l, B_3^{(3)} \ge l + v | N_a = 3\right) P(N_a = 3), \tag{13}$$

respectively. Let $\lambda = 1/\tau_c^* - 1$. Then, we have

$$P\left(B_2^{(2)} = l | N_a = 2\right) = 2\left(\sum_{i=0}^{l-2} \frac{\lambda^i e^{-\lambda}}{i!}\right) \frac{\lambda^{l-1} e^{-\lambda}}{(l-1)!}, \tag{14}$$

$$P\left(B_3^{(2)} = l | N_a = 3\right) = 3!\left(\sum_{i=0}^{l-2} \frac{\lambda^i e^{-\lambda}}{i!}\right) \frac{\lambda^{l-1} e^{-\lambda}}{(l-1)!} \left(\sum_{j=l}^{\infty} \frac{\lambda^j e^{-\lambda}}{j!}\right), \tag{15}$$

$$P\left(B_3^{(2)} = l, B_3^{(3)} \ge l + v | N_a = 3\right) = 3!\left(\sum_{i=0}^{l-2} \frac{\lambda^i e^{-\lambda}}{i!}\right) \frac{\lambda^{l-1} e^{-\lambda}}{(l-1)!} \left(\sum_{j=l+v-1}^{\infty} \frac{\lambda^j e^{-\lambda}}{j!}\right). \tag{16}$$

Substituting (14)-(16) into (12) and (13), and substituting these into (11), we have

$$P\left(N_a = 3, B_3^{(3)} \ge l + v | n_s = 2, d_s = l\right) = \frac{3\sum_{j=l+v-1}^{\infty} \frac{\lambda^j e^{-\lambda}}{j!}}{\frac{P(N_a=2)}{P(N_a=3)} + 3\sum_{j=l}^{\infty} \frac{\lambda^j e^{-\lambda}}{j!}}. \tag{17}$$

In (17), we need to specify the probabilities $P(N_a = 2)$ and $P(N_a = 3)$. We consider two cases of collision occurrence, one within an ordinary phase and one within a special phase. For the first case, the probabilities $P(N_a = 2)$ and $P(N_a = 3)$ are equal to $P(N_c = 2)$ and $P(N_c = 3)$ given in (9), respectively. As can be seen in (17), we only need the ratio of $P(N_a = 2)$ to $P(N_a = 3)$ to derive $v_a(l)$, which is calculated as

$$\frac{P(N_a = 2)}{P(N_a = 3)} = \frac{P(N_c = 2)}{P(N_c = 3)} = \frac{\binom{N}{2}(\tau_s^*)^2(1 - \tau_s^*)^{N-2}}{\sum_{v=3}^{N} \binom{N}{v}(\tau_s^*)^v(1 - \tau_s^*)^{N-v}}. \tag{18}$$

For the second case, the ratio of $P(N_a = 2)$ to $P(N_a = 3)$ is calculated as (30), and the derivation for this is presented in

**TABLE 5. Backoff mechanism of a station with the PCPB.**

| | |
|---|---|
| **1**: | Select a backoff counter value for a packet to transmit, BC $\leftarrow$ Pois $\left(\frac{1}{\tau_s^*}-1\right)+1$, |
| **2**: | Set $i_s, i_c, i_p, n_s, d_s, d_a, v_a$ to 0. |
| **3**: | Repeat for each slot, |
| **4**: | **if** the current backoff counter value BC $>0$, listen the channel, |
| **5**: | **if** the channel is sensed idle, |
| **6**: | **if** it is assumed to be in the ordinary phase ($i_s=0$), decrease the backoff counter by 1 (BC $\leftarrow$ BC $-1$), |
| **7**: | **else**, count 1 on $d_s$ ($d_s \leftarrow d_s+1$), |
| **8**: | **if** the collision indicator $i_c=1$, BC $\leftarrow$ BC $-1$, |
| **9**: | **else**, freeze the backoff counter (BC $\leftarrow$ BC), |
| **10**: | **if** there have been two consecutive successes ($n_s=2$), count 1 on the number of elapsed idle slots after two consecutive successes ($d_a \leftarrow d_a+1$), |
| **11**: | **if** $d_a \geq v_a$, reset $i_s, n_s, d_s, d_a, v_a$ to 0, |
| **12**: | **else** (the channel is sensed busy), |
| **13**: | **if** it is due to a success, |
| **14**: | **if** $i_s=1$, count 1 on $n_s$ ($n_s \leftarrow n_s+1$), |
| **15**: | **if** $n_s=2$, set $d_a \leftarrow 0$, |
| **16**: | **if** $i_p=0$, set $v_a \leftarrow v_a(d_s)$ with (18), |
| **17**: | **else**, set $v_a \leftarrow v_a(d_s)$ with (30), |
| **18**: | **if** $v_a=0$, reset $i_p, d_s, n_s, d_a$ to 0, |
| **19**: | **else if** $n_s>2$, reset $i_p, d_s, n_s, d_a, v_a$ to 0, |
| **20**: | **else** (the channel is busy due to a collision), |
| **21**: | **if** $i_s=0$, set $i_p \leftarrow 0$, |
| **22**: | **else**, set $i_p \leftarrow 1$, |
| **23**: | set $i_s \leftarrow 1$ and reset $d_s, n_s, d_a, v_a$ to 0, |
| **24**: | **else** (BC $=0$), transmits a packet, |
| **25**: | **if** it is successful, reset $i_c \leftarrow 0$ and BC $\leftarrow$ Pois $\left(\frac{1}{\tau_s^*}-1\right)+1$, |
| **26**: | do steps **14-19**, |
| **27**: | **else** (collided), set $i_c \leftarrow 1$ and BC $\leftarrow$ Pois $\left(\frac{1}{\tau_c^*}-1\right)+1$, |
| **28**: | do steps **21-23**. |

Appendix A. An indicator $i_p$ can be introduced to distinguish between these two cases during the operation. The overall operation process of the PCPB is described in Table 5.

The throughput and fairness performances of the PCPB are also shown in Table. 4 and Fig. 6, respectively. In terms of throughput, the PCPB shows slightly lower performance than the CPB. In addition, it can be seen that the PCPB shows higher throughput when $p_e$ is larger, that is, when stations aggressively compete with more weight on the possibility of collision by two packets rather than three or more. In what follows, we adopt the case of $p_e=0.1$ as the operation of the PCPB. As shown in Fig. 6, the PCPB shows similar fairness performance to the CPB, exceeding those of the optimized RAP and the BEB-based scheme in [34].

The approach of giving priority to collided packets also works favorably from a delay perspective. The detailed analysis on this is provided in Appendix B.

## V. ADAPTIVE OPERATION OF THE PCPB

Typically, it is hard for every station to know exactly how many stations are in the network, and even the number changes over time. In this section, we devise a version of the PCPB called the ACPB that adaptively operates in an environment where the number of stations in the network is time-varying. For a given $\mathrm{E}[T_c]$, we first quantize the optimal access probabilities, which are regarded as functions of $N$, by step functions having $L$ levels. Then each component of the step function represents a range of $N$ and the level each station utilizes as an access probability in that range. After the quantization, each station estimates a range in which the number of stations in the network is included among the $L$ ranges, and operates with access probabilities which are valid in that range. For the range estimation procedure, we regard it as a multi-armed bandit problem. Each range is regarded as an arm, and when a station operates with the access probabilities corresponding to the selected range, it obtains a reward value from RTS frames received during the operation. This reward is designed to indicate whether the currently selected range includes the actual number of stations in the network.

### A. QUANTIZATION OF OPTIMAL ACCESS PROBABILITIES

Here we provide the quantization procedure. Let the value of $\mathrm{E}[T_c]$ be given since it can be determined independently of $N$. We first quantize the optimal access probability $\tau_s^*(N)$, as a function of $N$, of the ordinary phase. Here, we consider the domain to be continuous and denote the function as $\tau_s^*(z)$. We rewrite $\tau_s^*(z)$ in (8) by $\tau_s^*(z) = \eta z^{-u_1}$ where $\eta = u_4(\mathrm{E}[T_c]^{-u_2}+u_3)$, and $u_1, u_2, u_3$, and $u_4$ are given in Table 2 for $\tau_s$. The goal of quantization is to find a step function $s(z)$ having $L$ levels, $s(z)=\sum_{l=1}^{L-1}\alpha_l \mathbb{1}_{[n_{l-1},n_l)}(z)+\alpha_L \mathbb{1}_{[n_{L-1},n_L]}(z)$ where $\mathbb{1}_A$ is an indication function of set $A$, that minimizes the distance from $\tau_s^*(z)$ with respect to the $L^2$ norm $\|\cdot\|_2$. We set the range of values, $[n_0, n_L]$, that $N$ can take by considering practical possibilities. When a station selects the $l$th range, it accesses the channel with probability $\alpha_l$ in an ordinary phase. Since a backoff counter value corresponding to the probability $\alpha_l$ has the mean of $1/\alpha_l$, this station observes $1/\alpha_l$ idle slots on average until the first transmission. As will be described later, a reward for the selected range is determined from the information analyzed in the RTS frames received while operating with the access probability corresponding to this range. Since different access probabilities cause different numbers of RTS frames to be analyzed on average, an imbalance may occur in the information necessary for reward calculation. To handle this, we consider a specific form for the levels $\{\alpha_l\}_{1\leq l\leq L}$ of the step function. We set $\alpha_l=\alpha/2^{l-1}$ with some $\alpha>0$ for $1\leq l\leq L$. Then, if a station selecting the $l$th range analyzes the received RTS frames until $2^{L-l}$ successful transmissions, the similar amount of information can be gathered at all arms. Now, the quantization task is to find $\alpha$ and partition $n_1<n_2<\cdots<n_{L-1}$ that minimize the following distance between $\tau_s^*(z)$ and $s(z)$,

$$\begin{aligned}\|\tau_s^*-s\|_2^2 &= \int_{n_0}^{n_L}(\tau_s^*(z)-s(z))^2dz \\ &= \int_{n_0}^{n_L}\Bigg(\eta z^{-u_1}-\sum_{l=1}^{L-1}\frac{\alpha}{2^{l-1}}\mathbb{1}_{[n_{l-1},n_l)}(z) \\ &\qquad -\frac{\alpha}{2^{L-1}}\mathbb{1}_{[n_{L-1},n_L]}(z)\Bigg)^2dz \\ &= \sum_{l=1}^{L}\int_{n_{l-1}}^{n_l}\left(\eta z^{-u_1}-\frac{\alpha}{2^{l-1}}\right)^2dz. \end{aligned} \tag{19}$$

**TABLE 6. Quantization of the optimal access probability $\tau_s^*$ in an ordinary phase.**

| |
|---|
| Calculate $\mathrm{E}[T_c]$ from network parameters. |
| With the optimal parameters, $u_2$, $u_3$, and $u_4$, calculate $\eta$, $\eta \leftarrow u_4(\mathrm{E}[T_c]^{-u_2} + u_3)$. |
| Set the number of levels $L$, and the minimal and maximal numbers of stations $n_0$ and $n_L$, respectively, that meet the condition, $n_0^{-u_1} > 2^{L-1} n_L^{-u_1}$. |
| Set the initial value $\alpha_{\text{old}}$, $\alpha_{\text{old}} \leftarrow \frac{\tau_s^*(n_0) + 2^{L-1}\tau_s^*(n_L)}{2}$. |
| **Step 1** Find the optimal partition for given $\alpha_0$. For $1 \le l \le L-1$, $n_l \leftarrow \left(\frac{3\alpha_{\text{old}}}{\eta 2^{l+1}}\right)^{-1/u_1}$. |
| **Step 2** Find the optimal $\alpha_{\text{new}}$ for given partition $\{n_l\}_{1\le l\le L-1}$. For $\alpha \in [\eta 2^{L-1} n_L^{-u_1}, \eta n_0^{-u_1}]$, $\alpha_{\text{new}} \leftarrow \text{argmin}_\alpha \sum_{l=1}^{L} \int_{n_{l-1}}^{n_l} \left(\eta z^{-u_1} - \frac{\alpha}{2^{l-1}}\right)^2 dz$. |
| **While** $\lvert\alpha_{\text{new}} - \alpha_{\text{old}}\rvert$ exceeds a given threshold, update $\alpha_{\text{old}} \leftarrow \alpha_{\text{new}}$, and do **Step 1** and **Step 2**. |

To find the values for $s(z)$, we adopt the method given in [39]. We begin with an initial value for $\alpha$. Since $\tau_s^*(z)$ is monotone decreasing, $\alpha$ should satisfy $\alpha \le \tau_s^*(n_0)$ and $\alpha/2^{L-1} \ge \tau_s^*(n_L)$. So, $n_0$, $n_L$, and $L$ should be carefully set so that

$$n_0^{-u_1} > 2^{L-1} n_L^{-u_1} \tag{20}$$

is satisfied. With $n_0$, $n_L$, and $L$ satisfying the condition (20), we set the initial value to $\{\tau_s^*(n_0) + 2^{L-1}\tau_s^*(n_L)\}/2$. In this case with fixed monotone levels, $\alpha_1 > \alpha_2 > \cdots > \alpha_L$, for a monotone target, $\tau_s^*(z)$, it is shown in [39] that an optimal partition $n_1 < n_2 < \cdots < n_{L-1}$ is given implicitly by

$$\tau_s^*(n_l) = \frac{\alpha_l + \alpha_{l+1}}{2},$$

for $1 \le l \le L-1$. With $\alpha$ for the levels, we have

$$\begin{aligned} \tau_s^*(n_l) &= \frac{y_l + y_{l+1}}{2} \\ \Leftrightarrow\ \eta n_l^{-u_1} &= \frac{3\alpha}{2^{l+1}} \\ \Leftrightarrow\ n_l &= \left(\frac{3\alpha}{\eta 2^{l+1}}\right)^{-1/u_1}. \end{aligned}$$

Thus, the optimal partition is given by

$$n_l = \left(\frac{3\alpha_0}{\eta 2^{l+1}}\right)^{-1/u_1}, \quad 1 \le l \le L-1.$$

Now, with the fixed partition, the next candidate value for $\alpha$ can be easily obtained by minimizing (19) because it is a quadratic function of $\alpha$. We repeat these two steps until the successive difference in $\alpha$ does not exceed a certain threshold to find the optimal partition and levels. Since each step is applied in the way of decreasing the distance (19), convergence is guaranteed. These procedures are described in Table 6.

In the case of quantizing the optimal access probability $\tau_c^*(N)$ for the special phase, we use the partition $n_1 < n_2 < \cdots < n_{L-1}$ obtained from quantization on $\tau_s^*(N)$. Let a step function for this case be

$$c(z) = \sum_{l=1}^{L-1} \beta_l \mathbb{1}_{[n_{l-1}, n_l)}(z) + \beta_L \mathbb{1}_{[n_{L-1}, n_L]}(z).$$

Similarly to above, we assume a continuous domain and denote the target as $\tau_c^*(z)$. By abuse of notation, we use $u_1$ and $\eta$ in the same way as above with different values for $\tau_c^*$ given in Table. 2. When a partition $n_1 < n_2 < \cdots < n_{L-1}$ is given, it is shown in [39] that the optimal levels $\{\beta_l\}_{1\le l\le L}$ are given by

$$\beta_l = \frac{1}{n_l - n_{l-1}} \int_{n_{l-1}}^{n_l} \tau_c^*(z) dz, \quad 1 \le l \le L.$$

More specifically, we have

$$\begin{aligned} \beta_l &= \frac{1}{n_l - n_{l-1}} \int_{n_{l-1}}^{n_l} \eta z^{-u_1} dz \\ &= \frac{\eta}{n_l - n_{l-1}} \cdot \frac{1}{-u_1 + 1} \left(n_l^{-u_1+1} - n_{l-1}^{-u_1+1}\right), \end{aligned}$$

for $1 \le l \le L$.

### B. RANGE ESTIMATION THROUGH BANDIT PROBLEM

As mentioned above, we consider the operation of a station repeatedly select one of the $L$ ranges as an $L$-armed bandit problem. When a station selects a range and operates with the associated levels as the access probabilities, a reward value indicating whether the actual number of stations in the network is within this range should be given.

In RTS/CTS mechanism, a station that has a packet to send transmits an RTS frame before sending it. Since an RTS frame contains the transmitter address, as utilized in [20], each station can distinguish different stations in the network from the received RTS frames. Consider a station contending for channel access by selecting the $l$th range. This station operates with access probabilities $\alpha/2^{l-1}$ and $\beta_l$ for the ordinary and special phases, respectively. The reward for choosing this range is calculated from RTS frames received until it successfully transmits $2^{L-l}$ packets. Such a setting allows a similar number of received RTS frames to be utilized in all ranges when calculating the reward. Let $N_R$ be the number of different stations analyzed from the RTS frames, including this station. If $N_R$ falls within the selected range, a reward of 1 is given; otherwise, a reward of 0 is given. Furthermore, $N_R$ takes on the role of $N$ in the calculation of (18) during the operation of PCPB for the selected arm. This process can be easily performed compared to the scheme in [20], which requires recording the number of successful transmissions for each detected station.

Under the condition that the reward distributions of the arms are stationary, the goal is to maximize the expected total reward over a certain time period. In this stationary bandit problems, it is known that UCB strategy [40] achieves optimal performance. Let $A_t$ be the chosen arm and $R_t(l)$ be the reward for arm $l$, $1 \le l \le L$ at time step $t$. From a station's perspective, the unit of one time step depends on the

**TABLE 7. UCB algorithm.**

| |
|---|
| **for** $t = 1, 2, \ldots, L$ **do** |
| play arm $A_t = t$; |
| **for** $t = L+1, \ldots$ **do** |
| play arm $A_t = \text{argmax}_{1 \le l \le L} [Q_t(l) + c_t(l)]$. |

arm selected at that time. If $A_t = l$, the reward $R_t(l)$ is given when the station completes its $2^{L-l}$ successful transmissions. At each time step $t$, the UCB strategy computes the average reward as

$$Q_t(l) = \frac{1}{H_t(l)} \sum_{s=1}^{t-1} R_s(l) \mathbb{1}_{\{A_t = l\}}, \tag{21}$$

where $H_t(l) = \sum_{s=1}^{t-1} \mathbb{1}_{\{A_t = l\}}$, and the exploration bonus, which is the size of the one-sided confidence interval for exploration, as

$$c_t(l) = \rho_U \sqrt{\frac{\log t}{H_t(l)}}, \tag{22}$$

for some parameter $\rho_U > 0$. Then the index of this strategy for arm $l$ is configured by $Q_t(l) + c_t(l)$. Using these notations, the algorithm for UCB strategy to play arms is given in Table 7.

When the number of stations in the network changes over time, the reward distribution of each arm is not stationary. The switching bandit problem takes this form of non-stationarity into account by allowing the distributions of rewards to change at unknown time instants. Noting that the UCB strategy achieves optimal performance in the stationary bandit problems, some adaptations of UCB to the switching bandit problems have been proposed in the literature. *Discounted UCB* (DUCB) [41] relies on a discount factor $\delta \in (0, 1)$. At each time step $t$, this policy computes the discounted average reward as

$$Q_t(\delta, l) = \frac{1}{H_t(\delta, l)} \sum_{s=1}^{t-1} \delta^{t-1-s} R_s(l) \mathbb{1}_{\{A_t = l\}}, \tag{23}$$

where $H_t(\delta, l) = \sum_{s=1}^{t-1} \delta^{t-1-s} \mathbb{1}_{\{A_t = l\}}$, and the exploration bonus as

$$c_t(\delta, l) = \rho_D \sqrt{\frac{\log h_t(\delta)}{H_t(\delta, l)}},$$

where $h_t(\delta) = \sum_{l=1}^{L} H_t(\delta, l)$, for some parameter $\rho_D > 0$. Then the index of this strategy for arm $l$ is configured by $Q_t(\delta, l) + c_t(\delta, l)$. The algorithm for playing arms in discounted UCB is the same as in the UCB strategy given in Table 7, except that $Q_t(l)$ and $c_t(l)$ are replaced with $Q_t(\delta, l)$ and $c_t(\delta, l)$, respectively.

Another variant of UCB is *sliding-window UCB* (SWUCB) [42]. This strategy only considers recent play results over a fixed-time horizon $\omega$ for arm selection. At each time step $t$, it computes the local average reward as

$$Q_t(\omega, l) = \frac{1}{H_t(\omega, l)} \sum_{s=t-\omega}^{t-1} R_t(l) \mathbb{1}_{\{A_t = l\}}, \tag{24}$$

where $H_t(\omega, l) = \sum_{s=t-\omega}^{t-1} \mathbb{1}_{\{A_t = l\}}$, and the exploration bonus as

$$c_t(\omega, l) = \rho_S \sqrt{\frac{\log \min(t, \omega)}{H_t(\omega, l)}}$$

for some parameter $\rho_S > 0$. As above, the index for arm $l$ is constructed by $Q_t(\omega, l) + c_t(\omega, l)$, and the algorithm for this strategy is obtained by replacing $Q_t(l)$ and $c_t(l)$ with $Q_t(\omega, l)$ and $c_t(\omega, l)$, respectively, in Table 7.

### C. NUMERICAL RESULTS

Here, we numerically evaluate the performances of the ACPB developed in this section. As in Section III-D, we perform simulation with network parameters in Table 3. For the simulation, we set the number of ranges for the quantization to $L = 4$. Also, the range of the number of stations $N$ is set to [5, 70], i.e., $n_0 = 5$ and $n_L = 70$. The threshold for the quantization is set to $10^{-7}$.

Note that the ACPB requires no information on the number of stations in the network and operates in a fully distributed manner. So we compare the performance of the ACPB with three backoff schemes, the A-RAP$^+$ [21], the AMOCW [20], and the Probability-based Opportunity Dynamic Adaptation (PODA) [43], which operate completely distributed. In these schemes, each station estimates the number of stations in the network and adjusts the backoff parameters based on this. In the A-RAP$^+$, each station uses only its own transmission results, and it is shown in [21] that the A-RAP$^+$ achieves superior performance than other schemes, such as the BEB, the QB, and the EIED, of this type. The AMOCW is an improved version of the Idle Sense technique, and each station using it monitors the network status as in the ACPB. The PODA is an enhanced version of the BEB, where each station tunes its CW adaptively by estimating the number of stations from its actual backoff probability.

In the case of the ACPB, the simulation starts with each station randomly selecting one of the four ranges uniformly. For the A-RAP$^+$ or AMOCW, at the start of each simulation run, each station randomly selects an initial estimate of the number of stations in the network uniformly from [5, 70]. The simulation for PODA starts with CW of 16 for all stations. At first, we consider a static environment in which the number of stations in the network is fixed. In this case, we consider that the ACPB operates based on UCB, which is suitable for the static environment, given in Table 7. For the value of $\rho_U$ in (22), we examine the performance for each of the values in $\{10^{-3}, 10^{-1}, 10\}$. For each simulation scenario, the performance result is obtained as the average of those from 20 runs.

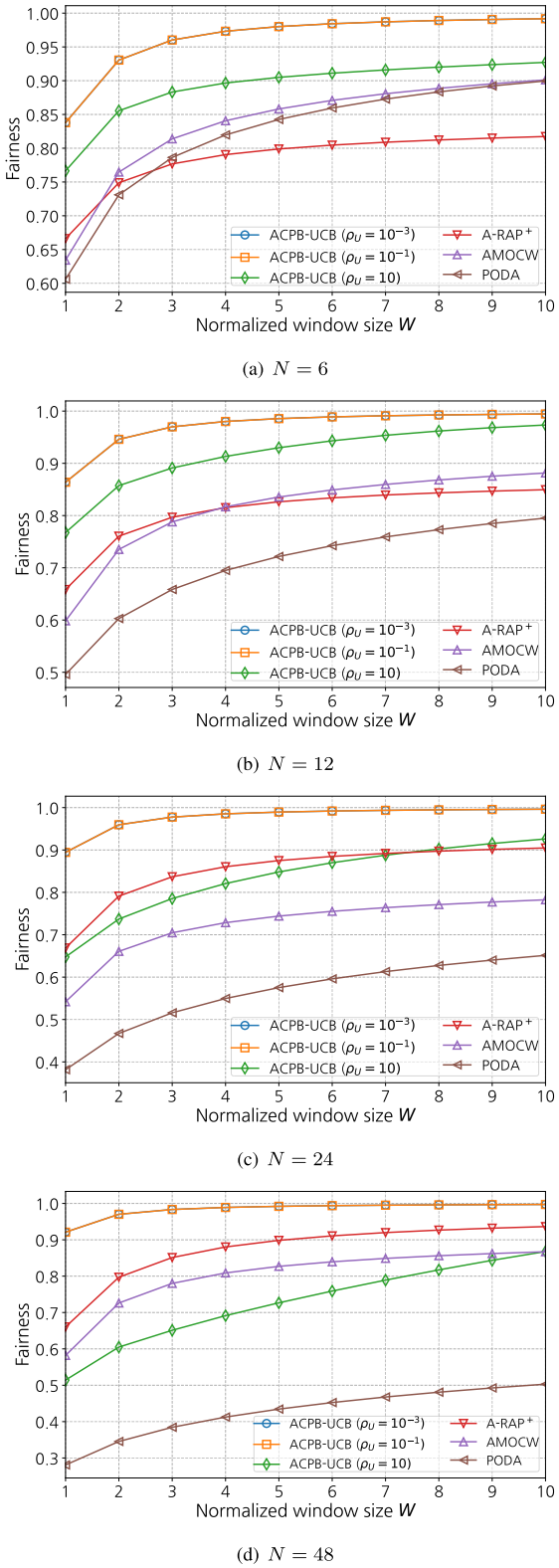
The throughput performance is provided in Table 8. The case where $\rho_U$ is $10^{-3}$ shows the same results as the case

**TABLE 8. Network throughput performance of adaptive schemes.**

| Data rate [Mbps] | $N$ | Backoff scheme | | | | |
|---|---|---|---|---|---|---|
| | | ACPB-UCB ($\rho_U$) | | A-RAP$^+$ | AMOCW | PODA |
| | | $10^{-1}$ | 10 | | | |
| 5 | 6 | 0.817 | 0.816 | 0.814 | 0.808 | 0.813 |
| | 12 | 0.816 | 0.816 | 0.814 | 0.81 | 0.815 |
| | 24 | 0.816 | 0.815 | 0.814 | 0.81 | 0.814 |
| | 48 | 0.815 | 0.811 | 0.814 | 0.81 | 0.813 |
| 50 | 6 | 0.531 | 0.529 | 0.521 | 0.488 | 0.504 |
| | 12 | 0.528 | 0.528 | 0.522 | 0.497 | 0.519 |
| | 24 | 0.526 | 0.525 | 0.521 | 0.495 | 0.523 |
| | 48 | 0.524 | 0.515 | 0.52 | 0.496 | 0.522 |

where $\rho_U$ is $10^{-1}$, and it is omitted from this table. As can be seen in the table, AMOCW achieves the lowest performance in all cases, while PODA achieves the second or third best performance among all schemes depending on the case. On the other hand, the ACPB with $\rho_U = 10^{-3}$ or $10^{-1}$ achieves the highest throughput in all cases. Also, the ACPB shows similar throughput results to the optimized PCPB given in Table 4 even though it operates without the information on the number of stations. This verifies the effectiveness of the proposed range estimation mechanism.

Fig. 8 shows fairness performance of the adaptive schemes when the 'Data rate' is set to 50Mbps. Since the result when the 'Data rate' is 5Mbps is similar to that in Fig. 8, it is omitted here. As can be seen in the figure, the ACPB with small $\rho_U$ ($10^{-3}$ or $10^{-1}$) show higher fairness performance than other access methods, which is similar to the results of the PCPB in Fig. 6 with a difference within about 1%. In the ACPB, the larger the $\rho_U$ value, the more weight is placed on the exploration. Thus, in the static environment currently under consideration, performance degradation is observed when $\rho_U = 10$. The A-RAP$^+$ and AMOCW show different performance patterns depending on the number of stations. The 'phase' element used in the A-RAP$^+$ is designed to adjust access probabilities with lower variability than the AMOCW, resulting in a more stable but slow adaptation. An abrupt adjustment of AMOCW could be detrimental from a short-term fairness perspective. Recall that the optimal access probability is inversely proportional to the number of stations. Thus, when the number of stations is small, even small fluctuations in the station number estimation can result in greater variations in access probability compared to cases where there are many stations. Fig. 9 shows the ratio of slots where the channel access probability of a station deviates from the mean for different levels of deviation during the total simulation for the A-RAP$^+$, AMOCW, and PODA. As expected, the A-RAP$^+$ shows a relatively high deviation rate in the case of a low number of stations. In this case, some stations operate for a relatively long time with different access probabilities, resulting in lower fairness compared to other schemes from a long-term perspective. On the other hand, as the number of stations increases, the variability in access probability decreases, so the A-RAP$^+$ exhibits high fairness. In the case of the PODA, it can be observed that the unfairness




(a) $N = 6$

(b) $N = 12$

(c) $N = 24$

(d) $N = 48$

**FIGURE 8. Fairness performances of adaptive schemes when the 'Data rate' is 50Mbps.**

caused by the waiting penalty of the BEB-based mechanism becomes more prominent as the number of stations increases, and the performance degradation becomes more severe.

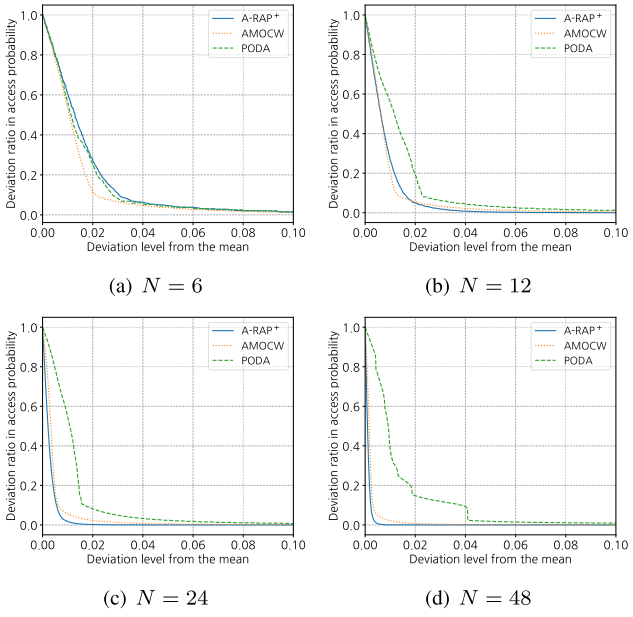



**FIGURE 9.** Ratio of slots in which the channel access probability of a station deviates from the mean versus several levels of deviation during the total simulation.

Now we consider an environment in which the number of stations in the network changes over time. As discussed in Section V-B, DUCB and SWUCB based operation can be applied in response to such an environment. For the simulation, we consider a scenario in which the number of stations in the network changes in the order of 24, 6, 48, and 12 for every $2.5 \times 10^5$ slots. For $\delta$ in (23) and $\omega$ in (24), values belonging to $\{0.5, 0.7, 0.9\}$ and $\{10, 30, 50\}$, respectively, are considered. For $\rho_D$ and $\rho_S$, we consider values belonging to $\{10^{-3}, 10^{-1}, 10\}$ as in $\rho_U$. With the network parameters in Table 3, the number of stations 24, 6, 48, and 12 correspond to ranges 3, 1, 4, and 2, respectively. Fig. 10 shows, for each scheme, the change over time of the average of the argmax of the estimated average rewards ($\{Q_t(l)\}_l$ in (21) for UCB, $\{Q_t(\delta, l)\}_l$ in (23) for DUCB, and $\{Q_t(\omega, l)\}_l$ in (24)) for ranges of all stations. In Fig. 10-(a), it can be seen that the UCB scheme cannot respond promptly to changes in the number of stations because the estimated average reward is calculated as the average of all rewards so far. Specifically, when the number of stations changes from 24 to 6 (from the first interval to the second interval), the accumulated reward of the remaining 6 stations tends to remain maximal in range 3, and only responds slowly to the change when the degree of exploration is high ($\rho_U = 10$). For the second change from range 1 to 4, there are many new stations entering to the network, and these lead to an increase in the average of the argmax. In the last interval, the larger the value of $\rho_U$, the faster the response to the change is observed, but it can be checked that the history in the second and third intervals is reflected in the average reward of the stations.

On the other hand, DUCB and SWUCB operate in a way that weights more on recent observations by adjusting the

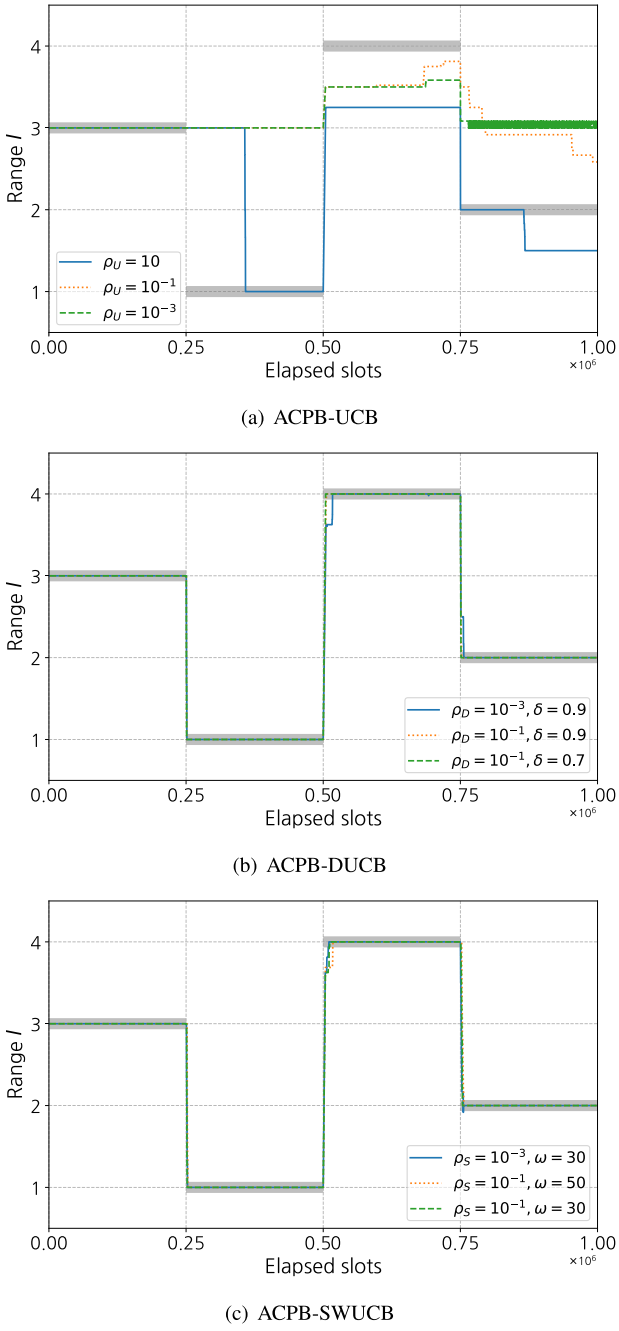
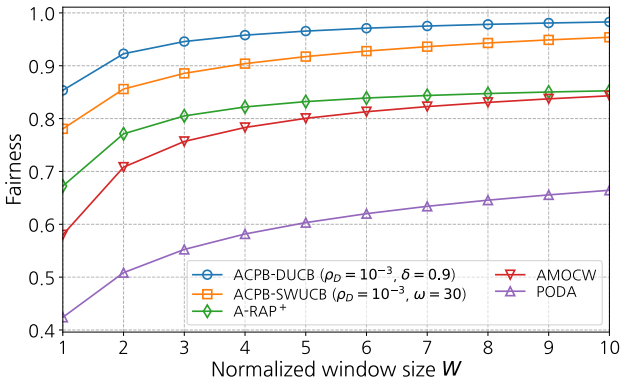



**FIGURE 10.** Averages of argmax of the estimated average rewards for ranges of all stations over time. The thick gray line indicates the range to which the actual number of stations belongs in each interval. The results are for the case where the 'Data rate' is 50Mbps.

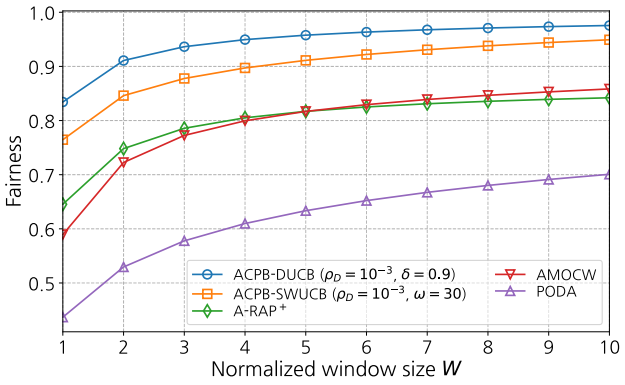
parameters $\delta$ and $\omega$, respectively, so a quick response to abrupt changes in network environment can be expected. Fig. 10-(b) shows the results for the three cases that achieve the highest throughput in the DUCB-based ACPB. As can be seen in the figure, the DUCB-based ACPB responds quickly to sudden changes in the number of stations in all cases. Although it is not clearly distinguished in the plot, at the boundaries between intervals, it can be checked that a relatively slow response to changes in the number of stations is

**TABLE 9. Network throughput performance of adaptive schemes when the number of stations changes.**

| Backoff scheme | Data rate [Mbps] | |
|---|---|---|
| | 5 | 50 |
| ACPB-DUCB ($\rho_D = 10^{-3}, \delta = 0.9$) | 0.816 | 0.527 |
| ACPB-SWUCB ($\rho_S = 10^{-3}, \omega = 30$) | 0.815 | 0.526 |
| A-RAP$^+$ | 0.814 | 0.520 |
| AMOCW | 0.809 | 0.494 |
| PODA | 0.814 | 0.518 |




(a) Data rate = 5Mbps




(b) Data rate = 50Mbps

**FIGURE 11. Averages of fairness performance measured for each interval.**

shown when the strategy places greater emphasis on history (high $\delta$) and less on exploration (low $\rho_D$). In 10-(c), as above, the results for the three cases where the highest throughput is achieved in the SWUCB-based ACPB. From the figure, we can see that the SWUCB-based ACPB also responds quickly and accurately to changes in the number of stations, as in the case of the DUCB-based ACPB. Note from (24) that $\omega$ means the length of the history used to estimate the average rewards of the arms. As expected, at the boundaries between intervals, it can be checked that a faster response to changes is made when a low $\omega$ values is used.

From now on, for the ACPB, we consider only the cases where DUCB or SWUCB is applied. Table 9 shows the network throughput achieved over the entire $10^6$ slots for adaptive backoff schemes under consideration. In the case of the ACPB, the results are shown for the parameter combination that yields the maximum throughput performance. As can be seen in the table, the ACPB achieves higher throughput than other techniques. Furthermore, the ACPB shows similar performance to the PCPB, given in Table 4, operating in an environment where the number of stations is static and known. Regarding fairness, Fig. 11 shows the average of fairness performance measured for each of the four intervals. From this figure, we can see that the ACPB achieves higher fairness performance then other schemes. The results for other schemes show similar patterns to those in the static environment given in Fig. 8. The most degraded performance is observed in the BEB-based PODA where the waiting penalty imposed on collided packets is the highest.

## VI. CONCLUSION

In this paper, we devised a random backoff scheme called the CPB which is based on the concept of giving priority to collided packets. In the CPB, two different backoff behaviors are defined in the ordinary and special phases. The special phase is designed to induce contention of only collided packets. We analyzed how much performance improvement can be achieved by prioritizing collisions through optimizing the backoff behavior of the CPB compared to existing schemes such as the BEB, which imposes a waiting penalty. For a given number of stations and network parameters, optimal values for the access probabilities were derived through Bayesian optimization. We also devised a practical form of the CPB called the PCPB, which operates probabilistically by relaxing the ideal assumptions inherent in the concept of the CPB. In a static environment where the number of stations is fixed and known, it was verified through simulation studies that the PCPB operating with the optimized access probabilities outperforms the existing schemes in term of throughput and fairness. Considering a more realistic environment in which the number of stations is time-varying, we also devised an adaptive version of the CPB called the ACPB. We formulated the operation of the ACPB as a multi-armed bandit problem through quantization of the optimal access probability. We applied UCB-based strategies to the ACPB, inspired by switching bandit learning theory which is suitable for a time-varying environment. Simulation results verified that our proposed ACPB not only outperforms the existing adaptive schemes but also achieves comparable performance to the optimized PCPB in a static environment.

## APPENDIX A
## DERIVATION OF THE PROBABILITIES ON THE NUMBER OF PACKETS INVOLVED IN A COLLISION THAT OCCURRED WITHIN A SPECIAL PHASE

Here, we derive the probabilities $P(N_a = 2)$ and $P(N_a = 3)$ when the collision occurs within a special phase. The special phase starts when a collision occurs in the ordinary phase. Therefore, collisions in the special phase occur when some of the stations trying to retransmit select the same backoff counter values. Let $K_c^{(2)}$ and $K_c^{(3)}$ be the numbers of collisions

caused by two and three packets within a special phase, respectively. Then we approximate the values of $P(N_a = 2)$ and $P(N_a = 3)$ as

$$P(N_a = 2) \approx \frac{\mathrm{E}\left[K_c^{(2)}\right]}{\mathrm{E}\left[K_c^{(2)}\right] + \mathrm{E}\left[K_c^{(3)}\right]},$$
$$P(N_a = 3) \approx \frac{\mathrm{E}\left[K_c^{(3)}\right]}{\mathrm{E}\left[K_c^{(2)}\right] + \mathrm{E}\left[K_c^{(3)}\right]}. \quad (25)$$

Note that

$$\mathrm{E}\left[K_c^{(2)}\right] = P(N_c = 2)\mathrm{E}\left[K_c^{(2)}\middle|N_c = 2\right] + P(N_c = 3)\mathrm{E}\left[K_c^{(2)}\middle|N_c = 3\right],$$
$$\mathrm{E}\left[K_c^{(3)}\right] = P(N_c = 3)\mathrm{E}\left[K_c^{(3)}\middle|N_c = 3\right]. \quad (26)$$

To derive formulas for the expected values in (26), we define the following probabilities:

$$p_2 = P(\exists \text{ a tie between two independent Pois}(\lambda)) = \sum_{i=0}^{\infty}\left(\frac{\lambda^i e^{-\lambda}}{i!}\right)^2,$$
$$p_{33} = P(\exists \text{ a tie for all three independent Pois}(\lambda)) = \sum_{i=0}^{\infty}\left(\frac{\lambda^i e^{-\lambda}}{i!}\right)^3,$$
$$p_{32} = P(\exists \text{ a tie in two of the three independent Pois}(\lambda)) = 3\sum_{i=0}^{l}\left(\frac{\lambda^i e^{-\lambda}}{i!}\right)^2\sum_{j=l+1}^{\infty}\frac{\lambda^j e^{-\lambda}}{j!} + 3\sum_{i=0}^{l}\frac{\lambda^i e^{-\lambda}}{i!}\left(\sum_{j=l+1}^{\infty}\left(\frac{\lambda^j e^{-\lambda}}{j!}\right)^2\right).$$

Now we derive conditional expectations in (26). If the special phase starts with a collision due to two packets, a collision occurs with probability $p_2$ and the phase ends with probability $1 - p_2$. Thus, $K_c^{(2)}$ given $N_c = 2$ is geometrically distributed, and hence we have

$$\mathrm{E}\left[K_c^{(2)}\middle|N_c = 2\right] = \frac{p_2}{1 - p_2}. \quad (27)$$

Similarly, $K_c^{(3)}$ given $N_c = 3$ is also geometrically distributed with the success probability $1 - p_{33}$. Hence we have

$$\mathrm{E}\left[K_c^{(3)}\middle|N_c = 3\right] = \frac{p_{33}}{1 - p_{33}}. \quad (28)$$

For $K_c^{(2)}$ given $N_c = 3$, with probability $p_{32}$, it follows the same distribution as $K_c^{(2)}$ given $N_c = 2$ after one count, and with probability $p_{33}$, it shows a renewed distribution with no counting. Thus, it follows that

$$\mathrm{E}\left[K_c^{(2)}\middle|N_c = 3\right] = p_{32}\left(1 + \mathrm{E}\left[K_c^{(2)}\middle|N_c = 2\right]\right) + p_{33}\mathrm{E}\left[K_c^{(2)}\middle|N_c = 3\right],$$

**TABLE 10. Average medium access delay (in milliseconds).**

| Data rate [Mbps] | $N$ | Backoff scheme | | |
|---|---|---|---|---|
| | | PCPB | RAP | [34] |
| 5 | 6 | 12.004 | 12.205 | 12.123 |
| | 12 | 24.055 | 24.079 | 24.301 |
| | 24 | 48.197 | 48.578 | 48.826 |
| | 48 | 96.266 | 96.918 | 96.461 |
| 50 | 6 | 1.844 | 1.865 | 1.853 |
| | 12 | 3.72 | 3.79 | 3.763 |
| | 24 | 7.484 | 7.526 | 7.515 |
| | 48 | 14.987 | 15.065 | 15.354 |

and hence we have

$$\mathrm{E}\left[K_c^{(2)}\middle|N_c = 3\right] = \frac{p_{32}}{(1 - p_2)(1 - p_{33})}. \quad (29)$$

As observed in (17), we only need the ratio $P(N_a = 2)/P(N_a = 3)$ to derive $v_a$ in (10). From (25), the ratio is approximated by $\mathrm{E}[K_c^{(2)}]/\mathrm{E}[K_c^{(3)}]$. Then, substituting (27)-(29) into (26), the ratio is calculated as

$$\frac{P(N_a = 2)}{P(N_a = 3)} = \frac{\frac{P(N_c=2)}{P(N_c=3)}\frac{p_2}{1-p_2} + \frac{p_{32}}{(1-p_2)(1-p_{33})}}{\frac{p_{33}}{1-p_{33}}}. \quad (30)$$

## APPENDIX B
## DELAY ANALYSIS

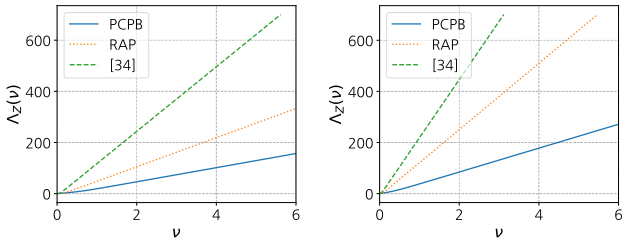
The total end-to-end packet delay consists of queuing delay and medium access delay. Queuing delay refers to the amount of time that a data packet waits in the system until it begins contending for channel access. Medium access delay refers to the amount of time between the initiation of contention for channel access and the reception of a positive acknowledgment for a transmission. Here, we compare the delay performance of the PCPB, RAP, and BEB-based scheme in [34] when the number of stations is known. All considered backoff schemes are assumed to be optimized in terms of throughput.

First, we investigate the medium access delay. Let $Z$ denote a random variable representing the medium access delay of a packet. As shown in Fig. 6, all backoff schemes provide long-term fairness in that their fairness indices converges to 1 as the sliding window size increases. Thus, one can speculate that the mean of $Z$ is inversely proportional to throughput. Table 10 and 11 show the average and standard deviation of medium access delay calculated from 10000 samples of $Z$ for each case in the simulations performed in Sections III and IV, respectively. The optimized throughput in Table 4 appears to be similar across all techniques, so the average medium access delay also shows similar results as speculated. On the other hand, delay jitter exhibits significant differences between backoff schemes, as shown in Table 11. Specifically, the higher the waiting penalty imposed on the collided packets, the higher the delay variability.

To analyze queuing delay, we tag a station and consider the worst case for this tagged station where all other stations are saturated. In [44], leveraging effective bandwidth theory, it is shown that the queue overflow probability in steady state

**TABLE 11. Standard deviation of medium access delay (in milliseconds).**

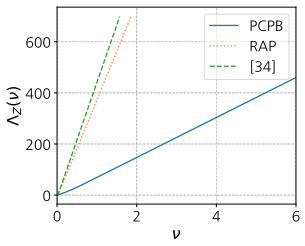
| Data rate [Mbps] | $N$ | Backoff scheme | | |
|---|---|---|---|---|
| | | PCPB | RAP | [34] |
| 5 | 6 | 3.259 | 6.178 | 12.077 |
| | 12 | 4.713 | 12.687 | 24.477 |
| | 24 | 6.611 | 26.789 | 48.949 |
| | 48 | 9.282 | 52.958 | 99.292 |
| 50 | 6 | 0.567 | 1.057 | 1.829 |
| | 12 | 0.841 | 2.335 | 3.769 |
| | 24 | 1.202 | 4.624 | 7.616 |
| | 48 | 1.694 | 9.683 | 16.214 |



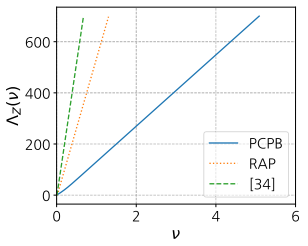

(a) $N = 6$, Data rate = 5Mbps



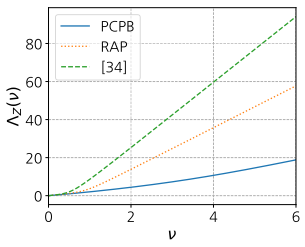

(b) $N = 12$, Data rate = 5Mbps



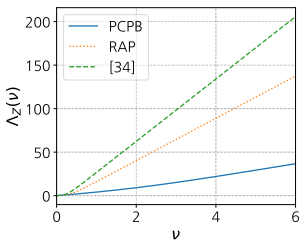

(c) $N = 24$, Data rate = 5Mbps



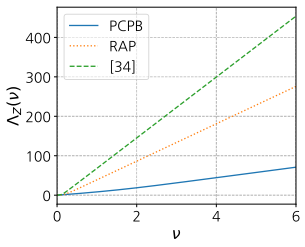

(d) $N = 48$, Data rate = 5Mbps



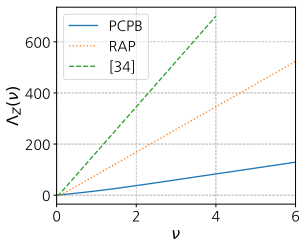

(e) $N = 6$, Data rate = 50Mbps




(f) $N = 12$, Data rate = 50Mbps




(g) $N = 24$, Data rate = 50Mbps




(h) $N = 48$, Data rate = 50Mbps

**FIGURE 12. Cumulant generating functions of medium access delay for the PCPB, RAP, and BEB-based scheme in [34].**

decreases as the mean and cumulant generation function of $Z$ decrease. Since the mean of $Z$ is fixed as throughput is optimized, we focus on the cumulant generating function.

The cumulant generating function $\Lambda_Z(\nu)$ of $Z$ is defined by, for $\nu \in \mathbb{R}$,

$$\Lambda_Z(\nu) := \log \mathrm{E}\left[e^{\nu Z}\right].$$

Fig. 12 shows the results for $\Lambda_Z(\nu)$ calculated with samples of $Z$ from the simulation for each backoff scheme as above. As can be seen in the figure, the PCPB has a much smaller cumulant generating function than other schemes. In other words, the PCPB can maintain a more stable system by having a lower queue overflow probability than other schemes.

In sum, the PCPB that gives priority to collided packets rather than waiting penalty achieves good delay performance in terms of low delay jitter and stable queue management.

## REFERENCES


[1] A. Yadav and O. A. Dobre, "All technologies work together for good: A glance at future mobile networks," *IEEE Wireless Commun.*, vol. 25, no. 4, pp. 10–16, Aug. 2018.

[2] D. Kozma, P. Varga, and G. Soós, "Supporting digital production, product lifecycle and supply chain management in industry 4.0 by the arrowhead framework—A survey," in *Proc. IEEE 17th Int. Conf. Ind. Informat. (INDIN)*, vol. 1, Jul. 2019, pp. 126–131.

[3] C. Aggarwal and B. B. Gupta, "A survey of civilian applications of WSN and security protocols," *Int. J. Sci. Res. Comput. Sci. Eng.*, vol. 6, no. 3, pp. 56–66, Jun. 2018.

[4] A. Tewari and B. B. Gupta, "Secure timestamp-based mutual authentication protocol for IoT devices using RFID tags," *Int. J. Semantic Web Inf. Syst.*, vol. 16, no. 3, pp. 20–34, Jul. 2020.

[5] B. Sejdiu, F. Ismaili, and L. Ahmedi, "Integration of semantics into sensor data for the IoT: A systematic literature review," *Int. J. Semantic Web Inf. Syst.*, vol. 16, no. 4, pp. 1–25, Oct. 2020.

[6] P. Varga, J. Peto, A. Franko, D. Balla, D. Haja, F. Janky, G. Soos, D. Ficzere, M. Maliosz, and L. Toka, "5G support for industrial IoT applications—Challenges, solutions, and research gaps," *Sensors*, vol. 20, no. 3, p. 828, Feb. 2020.

[7] Q. Cui, Y. Gu, W. Ni, and R. P. Liu, "Effective capacity of licensed-assisted access in unlicensed spectrum for 5G: From theory to application," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 8, pp. 1754–1767, Aug. 2017.

[8] M. Hirzallah, M. Krunz, B. Kecicioglu, and B. Hamzeh, "5G new radio unlicensed: Challenges and evaluation," *IEEE Trans. Cognit. Commun. Netw.*, vol. 7, no. 3, pp. 689–701, Sep. 2021.

[9] R. Bajracharya, R. Shrestha, and H. Jung, "Future is unlicensed: Private 5G unlicensed network for connecting industries of future," *Sensors*, vol. 20, no. 10, p. 2774, May 2020.

[10] *Physical Layer Procedures for Shared Spectrum Channel Access*, Standard TS 37.213 v17.4.0, 3GPP, 3GPP, Sophia Antipolis, France, Dec. 2022.

[11] E. Khorov, A. Kiryanov, A. Lyakhov, and G. Bianchi, "A tutorial on IEEE 802.11ax high efficiency WLANs," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 1, pp. 197–216, 1st Quart., 2019.

[12] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 3, pp. 535–547, Mar. 2000.

[13] N. Vaidya, A. Dugar, S. Gupta, and P. Bahl, "Distributed fair scheduling in a wireless LAN," *IEEE Trans. Mobile Comput.*, vol. 4, no. 6, pp. 616–629, Nov./Dec. 2005.

[14] S.-R. Ye and Y.-C. Tseng, "A multichain backoff mechanism for IEEE 802.11 WLANs," *IEEE Trans. Veh. Technol.*, vol. 55, no. 5, pp. 1613–1620, Sep. 2006.

[15] Y. Kim and G. Hwang, "Design and analysis of medium access protocol: Throughput and short-term fairness perspective," *IEEE/ACM Trans. Netw.*, vol. 23, no. 3, pp. 959–972, Jun. 2015.

[16] S. Mangold, S. Choi, G. R. Hiertz, O. Klein, and B. Walke, "Analysis of IEEE 802.11e for QoS support in wireless LANs," *IEEE Wireless Commun.*, vol. 10, no. 6, pp. 40–50, Dec. 2003.

[17] F. Cali, M. Conti, and E. Gregori, "Dynamic tuning of the IEEE 802.11 protocol to achieve a theoretical throughput limit," *IEEE/ACM Trans. Netw.*, vol. 8, no. 6, pp. 785–799, Dec. 2000.

[18] M. Heusse, F. Rousseau, R. Guillier, and A. Duda, "Idle sense: An optimal access method for high throughput and fairness in rate diverse wireless LANs," in *Proc. Conf. Appl., Technol., Archit., Protocols Comput. Commun.*, Aug. 2005, pp. 121–132.
[19] D.-J. Deng, C.-H. Ke, H.-H. Chen, and Y.-M. Huang, "Contention window optimization for ieee 802.11 DCF access control," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 5129–5135, Dec. 2008.
[20] S. Chun, D. Xianhua, L. Pingyuan, and Z. Han, "Adaptive access mechanism with optimal contention window based on node number estimation using multiple thresholds," *IEEE Trans. Wireless Commun.*, vol. 11, no. 6, pp. 2046–2055, Jun. 2012.
[21] Y. Oh, Y. Kim, J. Kim, G. Hwang, and S. Park, "A new autonomous adaptive MAC protocol in wireless networks," *IEEE Access*, vol. 6, pp. 15155–15169, 2018.
[22] X. Sun and L. Dai, "Backoff design for IEEE 802.11 DCF networks: Fundamental tradeoff and design criterion," *IEEE/ACM Trans. Netw.*, vol. 23, no. 1, pp. 300–316, Feb. 2015.
[23] N.-O. Song, B.-J. Kwak, J. Song, and L. E. Miller, "Enhancement of IEEE 802.11 distributed coordination function with exponential increase exponential decrease backoff algorithm," in *Proc. 57th IEEE Semiannual Veh. Technol. Conf. (VTC-Spring)*, Apr. 2003, pp. 2775–2778.
[24] Y. Kim and S. Park, "Performance analysis and fair coexistence of heterogeneous radio networks for multiple devices with different channel access schemes," *IEEE Access*, vol. 8, pp. 73398–73419, 2020.
[25] Y. Shi, Q. Cui, W. Ni, and Z. Fei, "Proactive dynamic channel selection based on multi-armed bandit learning for 5G NR-U," *IEEE Access*, vol. 8, pp. 196363–196374, 2020.
[26] P. Vijayakumar, D. L. Jegatha, and S. C. Rajkumar, "Deep reinforcement learning-based pedestrian and independent vehicle safety fortification using intelligent perception," *Int. J. Softw. Sci. Comput. Intell.*, vol. 14, no. 1, pp. 1–33, Mar. 2022.
[27] Y. Yu, T. Wang, and S. C. Liew, "Deep-reinforcement learning multiple access for heterogeneous wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 6, pp. 1277–1290, Jun. 2019.
[28] A. Kumar, E. Altman, D. Miorandi, and M. Goyal, "New insights from a fixed-point analysis of single cell IEEE 802.11 WLANs," *IEEE/ACM Trans. Netw.*, vol. 15, no. 3, pp. 588–601, Jun. 2007.
[29] J.-w. Cho, J.-Y. L. Boudec, and Y. Jiang, "On the asymptotic validity of the decoupling assumption for analyzing 802.11 MAC protocol," *IEEE Trans. Inf. Theory*, vol. 58, no. 11, pp. 6879–6893, Nov. 2012.
[30] E. Brochu, V. M. Cora, and N. de Freitas, "A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning," 2010, *arXiv:1012.2599*.
[31] G. Bianchi, "IEEE 802.11-saturation throughput analysis," *IEEE Commun. Lett.*, vol. 2, no. 12, pp. 318–320, Dec. 1998.
[32] C. E. Rasmussen and C. K. Williams, *Gaussian Processes for Machine Learning*, vol. 1. Cambridge, MA, USA: MIT Press, 2006.
[33] Y. Kim, G. Hwang, J. Um, S. Yoo, H. Jung, and S. Park, "Throughput performance optimization of super dense wireless networks with the renewal access protocol," *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3440–3452, May 2016.
[34] M. A. Salem, I. F. Tarrad, M. I. Youssef, and S. M. A. El-kader, "An adaptive EDCA selfishness-aware scheme for dense WLANs in 5G networks," *IEEE Access*, vol. 8, pp. 47034–47046, 2020.
[35] C. Vlachou, J. Herzen, and P. Thiran, "Fairness of MAC protocols: IEEE 1901 vs. 802.11," in *Proc. IEEE 17th Int. Symp. Power Line Commun. Appl.*, Mar. 2013, pp. 58–63.
[36] R. Jain, D.-M. W. Chiu, and W. R. Hawe, "A quantitative measure of fairness and discrimination for resource allocation in shared computer system," Eastern Res. Lab., Digit. Equip. Corp., Hudson, MA, USA, Tech. Rep. TR-301, 1984, vol. 38.
[37] C. E. Koksal, H. Kassab, and H. Balakrishnan, "An analysis of short-term fairness in wireless media access protocols," in *Proc. ACM SIGMETRICS*, 2000, pp. 118–119.
[38] *Local and Metropolitan Area Networks-Specific Requirements—Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, Standard 802.11-2020 Dec. 2020.
[39] T. Kämpke, "Optimal and near optimal quantization of integrable functions," *Comput. Math. with Appl.*, vol. 40, nos. 10–11, pp. 1315–1347, Nov. 2000.
[40] P. Auer, N. Cesa-Bianchi, and P. Fischer, "Finite-time analysis of the multiarmed bandit problem," *Mach. Learn.*, vol. 47, no. 2, pp. 235–256, 2002.
[41] L. Kocsis and C. Szepesvári, "Discounted UCB," in *Proc. 2nd PASCAL Challenges Workshop*, Apr. 2006, pp. 51–134.
[42] A. Garivier and E. Moulines, "On upper-confidence bound policies for switching bandit problems," in *Proc. Int. Conf. Algorithmic Learn. Theory*, 2011, pp. 174–188.
[43] J. T. Liew, F. Hashim, A. Sali, M. F. A. Rasid, and A. Jamalipour, "Probability-based opportunity dynamic adaptation (PODA) of contention window for home M2M networks," *J. Netw. Comput. Appl.*, vol. 144, pp. 1–12, Oct. 2019.
[44] Y. Kim and G. Hwang, "Delay analysis and optimality of the renewal access protocol," *Ann. Oper. Res.*, vol. 252, no. 1, pp. 41–62, May 2017.

**YUNBAE KIM** received the B.Sc. and Ph.D. degrees in mathematical sciences from the Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Republic of Korea, in 2009 and 2015, respectively. Since November 2015, he has been with the Electronics and Telecommunications Research Institute (ETRI), Daejeon. His research interests include performance evaluation and optimization of wireless communication networks.

**SEUNGKEUN PARK** (Member, IEEE) received the B.S. and M.S. degrees in applied statistics from Korea University, Seoul, South Korea, in 1991 and 1993, respectively, and the Ph.D. degree in information communication engineering from the University of Chungbuk, Cheongju-si, South Korea, in 2004. He is currently an Assistant Vice President of the Electronics and Telecommunications Research Institute, Daejeon, South Korea. His current research interests include communication theory and spectrum management.