Modeling and Analysis of IEEE 802.11ax OFDMA-based Random Access

Yang Hang, Der-Jiunn Deng, *Member, IEEE,* and Kwang-Cheng Chen, *Fellow, IEEE*

***Abstract*—With the progressive increase of dense WiFi networks deployment, IEEE 802.11ax, the task group aimed at High Efficiency WLAN (HEW) in dense scenario, makes revolutionary modifications on both MAC and PHY layer. Especially one feature on MAC layer is OFDMA-Based Random Access (OBRA) mechanism. In this paper, we develop a bi-dimensional Markov process to model of the OBRA and derive a closed-form expression of system efficiency and access delay. Our simulation results verify the accuracy of the model. Finally, the effects of system parameters, including the number of resource units (RUs) for random access, initial and maximum contention window, are evaluated.**

***Index Terms*—OFDMA-based Random Access, Multi-User PHY, OFDMA, 802.11ax, High Efficiency WLAN**

I. Introduction

During last decade, IEEE 802.11 has achieved a great success with enormous WiFi networks deployed for its high throughput and relative simplicity of implementation. According to Cisco Visual Network index [**?**], the global mobile traffic will experience an eightfold increase within 2015-2020, reaching 30.6 EB per month by 2020. Therefore, the *dense scenarios*, where there are plenty of stations (STAs) associating with one access point (AP) or multiple basic service sets (BSSs) in a limited area, are becoming a common scenario.

Previous IEEE 802.11 amendments made fewer improvement on MAC layer compared with that on PHY layer. A series of standards (802.11b/g/n/ac) have evolved to respond to the increasing WLAN demand, mainly by means of raising data rate on the physical layer (PHY) from 2 Mbps to 7 Gbps [**?**]. However, user experience of WiFi networks does not improve enough with the data rate, especially in the dense scenario. That is because few enhancement has been made to improve MAC efficiency. The MAC is always based on distributed coordination function (DCF) in legacy 802.11, which is a random access protocol on a Single-User (SU) PHY.

MAC efficiency mainly degrades with two causes, the overhead of control signaling and packet collisions. Much effort has been conducted in legacy 802.11 to reduce the overhead of control signaling, such as Reduced Inter-Frame Spacing (RIFS), frame aggregation, etc [**?**]. However, in the dense scenario, collision is becoming the major component of bandwidth waste. The collision comes inherently from instability of DCF [**?**]. In addition, an unfair queueing problem worsens the instability of DCF.

The unfair queueing problem results from both DCF and star topology of BSS. When the WiFi network is seen with queueing model, each STA and AP are modeled as an individual queue, and the shared channel as the server. Since the BSS operates in a star topology, AP as the coordinator and n STAs nearby, all STAs download and upload data from and to AP, which are referred to as down-link (DL) traffic and up-link (UL) traffic respectively. On the other hand, each STA and AP has the same chance to access the channel. Thus the AP has chance of $1/(n+1)$ to access channel while it shares the DL traffic loading, which accounts for more than $1/2$ of total traffic. The queue model of WiFi network is, therefore, an unfair queueing between DL and UL transmissions.

Study group 802.11ax, aimed at HEW in the dense scenarios, thus modifies the MAC thoroughly by substituting DCF with centralized control. On the PHY layer, OFDMA is adopted by 802.11 for the first time to implement both the DL and UL Multi-User (MU) channel, which means AP could communicate with multiple STAs simultaneously. On the MAC layer, the AP is supposed to schedule both DL and UL transmissions, which means AP does not need to contend with STAs. As a result, instability of DCF should be mitigated and the unfair queueing problem does not exist anymore. A more complicated MAC scheme is required for 802.11ax compared with DCF in legacy 802.11. One of the MAC schemes, OFDMA-based random access (OBRA), is proposed in 802.11ax as random access for delivering short frames, like bandwidth request. The OBRA mechanism proposes a three-way handshake procedure to implement random access: 1) AP initiates the OBRA by transmitting an announce frame at first; 2) STAs receiving the the announcement then contends with Aloha and binary exponential backoff mechanism. 3) finally AP responds with ACK confirming which STAs contend successfully. Details of the mechanism will be illustrated in Section III.

Since in [**?**], the 2-dimension Markov chain model has been demonstrated simple and accurate to model DCF, we extend the Markov chain model to the OBRA mechanism. Though the OBRA differs much from DCF, including from SU channel to MU channel, from distributed MAC scheme to centralized MAC scheme, and from carrier sense to Aloha, our work shows that the Markov chain model still works well by some modifications adapting to these differences. And thereby, we are able to evaluate the system efficiency and access delay.

The paper is organized as follows. Section II displaies previous related works. Some necessary features of 802.11ax are illustrated in Section III, including MU-OFDMA and OFDMA-based random access procedure. Section IV contains the system model and performance analysis. Then Section V shows simulation results, along with analysis results to validate

the model. In Section VI, additional considerations on optimal performance are carried out and the effects of some important system parameters are evaluated. Conclusion remark is finally given in Section VII.

II. Related Works

Evaluation of 802.11 MAC efficiency is a fertile field along with the evolution of 802.11, from DCF to EDCA [?] exploiting QoS guarantee. However, the fundamental mechanism is still CSMA/CA. Various methods were proposed during the past 20 years. In [?] [?] and [?], assuming backoff time follows a geometric distribution, an analytical formula for the protocol capacity namely throughput of DCF is derived and [?] proposed a distributed algorithm to tune its backoff algorithm at run-time. In [?], Bianchi proposed a simple and accurate model, which is a discrete-time 2-dimention Markov chain model, to compute the throughput of DCF under saturated condition and ideal channel assumptions. Afterward, this model is extended to evaluate throughput and delay performance of DCF or EDCA in plenty of scenarios, which validate the accuracy of the model. In [?] and [?], the authors consider unsaturated condition by introducing a new idle state. The impact of channel induced errors and capture effects is also evaluated in [?]. In [?] and [?], the authors extend Bianchi's model accounting most features of EDCA in 802.11e under ideal channel condition.

Standardization of next generation WLAN 802.11ax [?] is almost complete, exploiting therein MU-OFDMA PHY and a centralized MAC scheme. However, random access is important to efficiently handle initial access, bursty traffic, and short packets. The OBRA is thus proposed in [?] as a multi-channel slotted Aloha with exponential backoff algorithm, transforming SU to MU channel.

Multi-channel slotted Aloha is implemented by 802.11 WLAN for the first time, while it has been adopted by cellular networks for a while to perform initial association and to request transmission resource. One of few works [?] generalizes CSMA/CA to MU-OFDMA for 802.11 WLAN, which is very different from the OBRA defined in 802.11ax [?]. Most focus on cellular networks, like 3GPP LTE-A and IEEE 802.16. In the literature, works follow the research of single channel slotted Aloha [?], working on stability [?], closed-form of throughput [?] [?] [?], collision probability [?] [?], and access delay [?] [?] [?] of multi-channel slotted Aloha. [?] stabilizes multi-channel slotted Aloha by pseudo-Bayesian algorithm. Collision resolution including uniform backoff and exponential backoff is investigated, [?] [?] comparing the impact of both two backoff algorithms, while [?] considering only exponential backoff. And [?] designs a 1-persistent type retransmission that avoids backoff to achieve a fast access. Different from above works on steady-state performance, Wei [?] [?] conducts a transient-performance analysis of OFDMA system with busty traffic.

As far as we know, no closed-form of performance analysis of the OBRA has been derived, and the Markov chain model has never been used to model the multi-channel slotted Aloha.

TABLE I: Maximum number of RUs for each channel width

RU type	CBW20	CBW40	CBW80	CBW80+80 and CBW160
26-tone RU	9	18	37	74
52-tone RU	4	8	16	32
106-tone RU	2	4	8	16
242-tone RU	1	2	4	8
484-tone RU	N/A	1	2	4
996-tone RU	N/A	N/A	1	2
2×996 tone RU	N/A	N/A	N/A	1

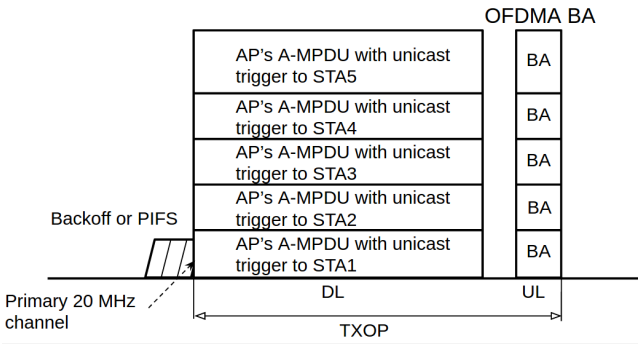


Fig. 1: MU DL of 802.11ax

III. System Overview

Some necessary features of 802.11ax to better understand the OBRA mechanism are illustrated in this section. First, MU-OFDMA is firstly adopted by 802.11 to implement MU DL and MU UL. To support MU UL, study group 802.11ax proposes trigger-based MU UL procedure, which means UL transmission could be scheduled by AP, shifting distributed coordination to centralized control. Then OFDMA-based random access (OBRA) is illustrated in the following subsection. For more features of 802.11ax, refer to [?] [?].

A. MU-OFDMA

OFDM has been proposed as one of the prime schemes for MU communications. In such OFDM-based systems, the total bandwidth is divided into multiple sub-channels so that multiple access can be accommodated in an OFDMA. Though MU PHY has been implemented in 802.11n and 802.11ac with MU-MIMO, only MU DL transmission is realized. MU-OFDMA implements both MU DL and MU UL. Especially for MU UL transmission, which is more difficult to implement, trigger-based MU UL is thus proposed.

With MU PHY, the original SU 20 MHz channel could be specified more fine-grained and be also aggregated to a wider channel to meet various bandwidth demands. The resource unit (RU), which can be accessed by one STA, is specified as Table I. For example, the smallest RU is 26-tone, with which a 20 MHz could be separated into 9 subchannels. Also multiple 20 MHz channels can be aggregated to improve system throughput, which is referred to as *Channel Bonding*. It is worth mentioning that every transmission of MU should end at the same time. That means padding is required for shorter packets.

In respect of MAC layer, first, for MU DL transmission, AP transmits DL packets to multiple stations simultaneously as in

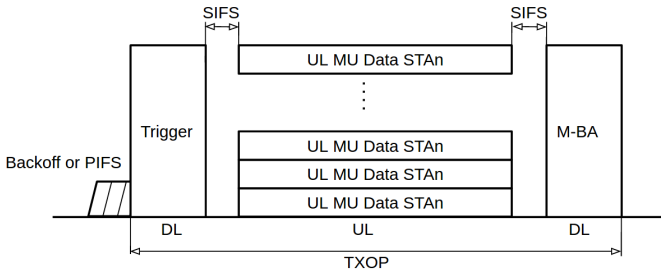


Fig. 2: Trigger-based MU UL of 802.11ax

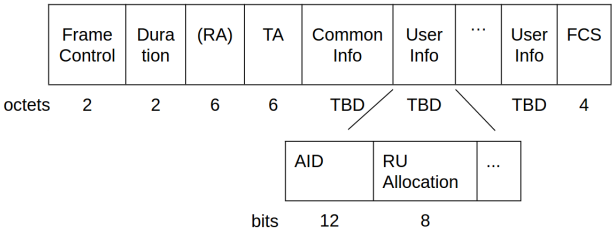


Fig. 3: Trigger Frame format

Fig. 1. Secondly, MU UL transmission is more complicated because WiFi network is not a well-synchronized system. A trigger-based MU UL transmission is thus issued in Fig. 2. A brand new control frame, trigger frame (TF), is created to be transmitted by AP to initiate the UL transmission. STAs could not transmit UL packets until they receive a TF which allocates RU for the STA or specifies RUs for random access. Afterwards, AP responds with ACK frame. All the procedure form a three-way handshake UL transmission. The trigger frame format is in Fig. 3. Since the standard is in progress, some fields remain to be determined (TBD). In the field *User Info*, subfield *AID* represents the identification of STA and subfield *RU Allocation* represents the RU allocated to the STA. Especially, *AID* with value 0 means the RU is for random access.

B. OBRA

An example of TF-based MU UL transmission with OBRA, illustrated in Fig. 4, is divided into two steps, one as random access procedure and the other UL data transmission. The random access procedure, namely the OBRA, is initiated by AP to collect traffic information named buffer state report (BSR) from stations. Thereby in the next step, the AP could allocate RUs to STAs by transmitting a TF containing RU allocation. Then the STAs, receiving the TF with resource allocation information, transmit UL data packets. And AP at last responds ACK. Therefore, TFs in the example have two variants, one for random access and the other for resource allocation. When traffic information is collected, the allocation is beyond the scope of this paper. Only step one is our interest.

Details of OBRA mechanism is as Fig. 5. To initialize a random access procedure, AP first transmits a TF announcing RUs for random access by setting the AID of the RUs to 0. Attempting STAs, whose buffers are not empty, maintain a backoff counter named OFDMA Backoff (OBO), which are randomly generated among range $[0, OCW]$. Then the OBO

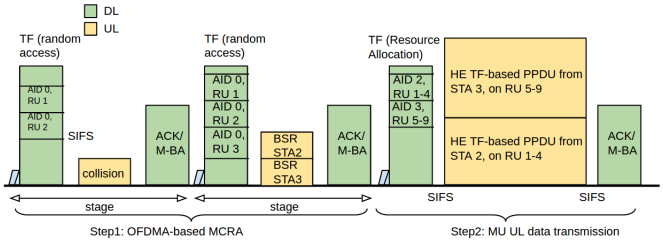


Fig. 4: An example of Trigger-based MU UL transmission with OBRA

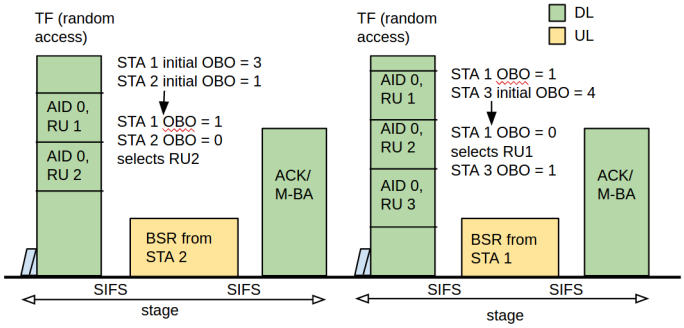


Fig. 5: Illustration of OBRA

subtracts the value of M once receiving the TF, otherwise freezes, where M is calculated by sum the number of RUs whose AID equals 0. When the OBO reaches 0, the STA will randomly select a RU from those whose AIDs equal 0 to transmit a request after short inter-frame spacing (SIFS). After that, AP responds with a block-ACK indicating which STAs contend successfully. The period of the whole three-way handshaking is named a *stage*. A successful stage means at least one STA contend successfully to transmit a request in a stage. It is worth noting that the stage in this paper, which is specified from standard [**?**], is a concept of the time interval, not the backoff stage in other papers. To avoid confusion of the two meanings, backoff stage is replaced with *backoff level* in this paper. When STAs fail to contend, the OCW will be doubled until OCW reaches OCW_{max}, which means the backoff level increases one step once a failure until the highest level.

In terms of implementation, system parameters of the OBRA are configured dynamically by AP, including OCW_{min}, OCW_{max}, M, where OCW_{min}, OCW_{max} represent the minimum and the maximum contention window, and M as the number of RUs for random access. Two of critical parameters OCW_{min}, OCW_{max} are configured in Random Access Parameter Set (RAPS) element contained in beacon frame sent by AP. Check field *OCW Range* in RAPS element as in Fig. 6, $OCW_{min} = 2^{EOCW_{min}} - 1$, $OCW_{max} = 2^{EOCW_{max}} - 1$. M is obtained from TF by sum the number of RUs whose *AID* equal 0. To simplify following analysis, we issue another parameter m, *maximum backoff level*, so that $OCW_{max} = (OCW_{min} + 1) * 2^m - 1$. The configuration of system parameter is absolutely different from that of legacy 802.11, where all the parameters are predefined in each STA's hardware. OBRA is thus more flexible compared with DCF.

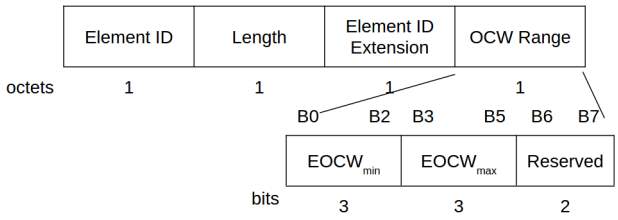


Fig. 6: Random Access Parameter Set (RAPS) Element

IV. System Model

Markov chain model has been verified of its accuracy of depicting the steady state behavior of DCF based on the assumption that at each request transmission, each request frame collides with constant and independent probability p regardless of the number of retransmission suffered [**?**]. Although some differences exist between OBRA and DCF, following shows the accuracy of Markov chain model for OBRA mechanism.

The analysis is divided into two parts. First is the Markov chain model to estimate the packet transmission probability τ and conditional collision probability p. Secondly, we evaluate some metrics given τ, including the number of successful contending stations in a stage n_s, self-defined system efficiency *eff*, and access delay of a STA D. Table II is a list of all parameters and notations.

TABLE II: Parameters and Notations Interpretation

Notations	Meaning
n	Number of stations
OCW_{min} (W_0)	Minimum OFDMA contention window
OCW_{max} (W_m)	Maximum OFDMA contention window
M	Number of RUs for random access
m	Maximum backoff level
p	Packet collision probability
τ	Station's transmission probability
n_s	Number of successful stations in a stage
D	Access delay, number of stages for a station to contend successfully

A. Packet Transmission Probability

Consider a single BSS with an AP and n of STAs under saturation condition, which means each station is an attempting station. Also, the ideal channel is assumed so that collision happens only if more than one station transmits at the same RU. Only UL BSR transmission (bandwidth request) is concerned, while DL transmission and UL data transmission are ignored here.

To model OBRA with the discrete-time Markov chain model, the concept of time is supposed to be adapted. In DCF, time unit of the model corresponds with slot, while in OBRA time unit of the model is stage, a three-way handshake. Thus the delay will be measured in the number of stages.

Similar to Bianchi's work, let $\{s(t), b(t)\}$ be the bi-dimension process, where $s(t)$ denotes the backoff level $(0, \cdots, m)$, and $b(t)$ denotes backoff counter $(0, \cdots, W_i)$. With the discrete and integer time scale, t and $t+1$ corresponds to beginnings of two consecutive stages. $\{b(t)\}$ is not a Markov process because the state of the current stage depends on the history of transmission instead of only the previous stage. The bi-dimensional process $\{s(t), b(t)\}$ is a Markov chain. The key assumption is still necessary that at each request transmission, and regardless of the number of retransmission suffered, each request frame collides with constant and independent probability p. With the independence assumption, p will be a constant. The Markov chain is able to be conducted as in Fig. 7.

Another important difference between the two Markov chain models of OBRA and DCF, is clarified here. Since the station of DCF senses the carrier before transmitting, it will freeze its backoff counter and stay at the state if channel is sensed busy. However in the OBRA, because the time unit, a stage, contains a period for exactly a packet transmission, stations certainly subtract M from the OBO only if they receive a TF for random access. Stations of the OBRA thus stay in one state for a period of exactly single stage.

Some modifications are mentioned here to adapt to differences between OBRA and DCF. First, in a row of states, as OBO subtracts the value of M rather than 1, stations transfer to states M-step ahead. Second, since states with $b(t) \leq M$ will decrease to 0 at the current state, which means stations could access RUs, we could merge these states into one state, denoted by $\{i, T\}$. T is an integer set of $[0, M]$.

Let's assume $P\{i_1, k_1 | i_0, k_0\} = P\{s(t+1) = i_1, b(t+1) = k_1 | s(t) = i_0, b(t) = k_0\}$. In this Markov Chain, the only non null one-step transition probabilities are

$$\begin{cases} P\{i,T|i,k\} = 1 & k \in [M+1, 2M] \quad i \in [0,m] \\ P\{i,k-M|i,k\} = 1 & k \in [2M+1, W_i] \quad i \in [0,m] \\ P\{0,k|i,T\} = \frac{1-p}{W_0+1} & k \in [M+1, W_0] \quad i \in [0,m] \\ P\{0,T|i,T\} = \frac{(1-p)(M+1)}{W_0+1} & \quad i \in [0,m] \\ P\{i,k|i-1,T\} = \frac{p}{W_i+1} & k \in [M+1, W_i] \quad i \in [1,m] \\ P\{i,T|i-1,T\} = \frac{p(M+1)}{W_i+1} & \quad i \in [1,m] \\ P\{m,k|m,T\} = \frac{p}{W_m+1} & k \in [M+1, W_m] \\ P\{m,k|m,T\} = \frac{p(M+1)}{W_m+1} & \end{cases} \tag{1}$$

The first and second equations in (1) accounts for the fact that the backoff counter maintained by stations will subtract M, the number of RUs for random access. The third and fourth equations represent that after a successful contention, stations will reset the contention window size to initial window size and uniformly generate a backoff value among $[0, W_0]$, since T is an integer set $[0, M]$, the transition probability to states $\{i, T\}$ is $M+1$ times of that to states $\{i, k\}$. For the fifth and sixth equations, they represent when a failure contention occurs, the contention window size will be doubled, $W_i = 2W_{i-1}+1$. The last two equations are the case of failure at the maximum backoff level. We assume no packets are discarded, repeating retransmitting until success.

Let $b_{i,k} = \lim_{t\to\infty} P\{s(t) = i, b(t) = k\},\ i \in [0, m],\ k \in [0, W_i]$, be the stationary distribution of the Markov chain. Then we show how to obtain transmission probability τ and conditional collision probability p. First, for states with $b(t) =$

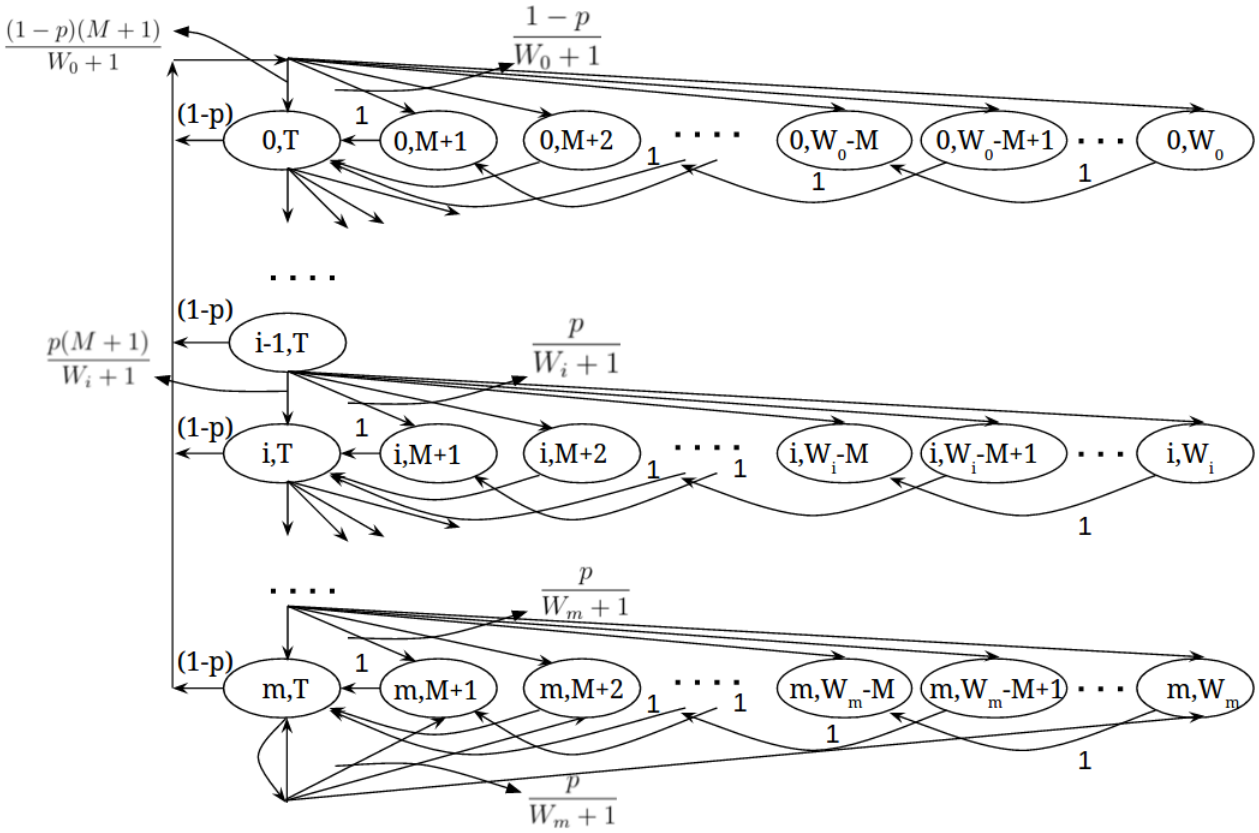


Fig. 7: Markov Chain model for the backoff window size

T, in which stations will transmit requests or BSRs in current stage,

$$\begin{aligned} b_{i-1,T}\cdot p = b_{i,T} \rightarrow b_{i,T} = p^i b_{0,T}, \quad 0 \le i < m \\ b_{m-1,T}\cdot p = (1-p) b_{m,T} \rightarrow b_{m,T} = \frac{p^m}{1-p} b_{0,T}. \end{aligned} \tag{2}$$

Then, the other states could be expressed by states with $b(t) = T$:

$$b_{i,k} = \begin{cases} (\lfloor \frac{W_0-k}{M} \rfloor + 1)\frac{(1-p)}{W_0+1}\sum_{i=0}^{m} b_{i,T}, \ M+1 \le k \le W_0, \ i = 0 \\ (\lfloor \frac{W_i-k}{M} \rfloor + 1)\frac{p}{W_i+1} b_{i-1,T}, \ M+1 \le k \le W_i, \ 0 < i < m \\ (\lfloor \frac{W_m-k}{M} \rfloor + 1)\frac{p}{W_m+1}(b_{m-1,T} + b_{m,T}), \\ \qquad\qquad M+1 \le k \le W_m, \quad i = m \end{cases} \tag{3}$$

The first and third subequation in (3) are first and last row of states in 7 except the first column state with $b(t) = T$. The second subquation is the states in the middle rows without the states in the first column. From (2), $\sum_{i=0}^{m} b_{i,T} = \frac{b_{0,T}}{1-p}$ is obtained; whereby each row of states in Fig. 7 could be obtained by summing the subequations of (3) respectively to get (4).

$$\begin{cases} \sum_{k=M+1}^{W_0} b_{0,k} = \\ \quad \frac{b_{0,T}}{W_0+1}\left(-\frac{M}{2}\left\lfloor \frac{W_0}{M} \right\rfloor^2 + \left(W_0 - \frac{M}{2}\right)\left\lfloor \frac{W_0}{M} \right\rfloor\right) \\ \sum_{i=1}^{m-1}\sum_{k=M+1}^{W_i} b_{i,k} = \\ \quad \frac{b_{0,T}}{W_0+1}\left(\frac{p}{2}\right)^i\left(-\frac{M}{2}\left\lfloor \frac{W_i}{M} \right\rfloor^2 + \left(W_i - \frac{M}{2}\right)\left\lfloor \frac{W_i}{M} \right\rfloor\right) \\ \sum_{k=M+1}^{W_m} b_{m,k} = \\ \quad \frac{b_{0,T}}{W_0+1}\frac{(\frac{p}{2})^m}{1-p}\left(-\frac{M}{2}\left\lfloor \frac{W_m}{M} \right\rfloor^2 + \left(W_m - \frac{M}{2}\right)\left\lfloor \frac{W_m}{M} \right\rfloor\right) \end{cases} \tag{4}$$

Each subequation in (4) has the same term: $-\frac{M}{2}\left\lfloor \frac{W_i}{M} \right\rfloor^2 + \left(W_i - \frac{M}{2}\right)\left\lfloor \frac{W_i}{M} \right\rfloor$. To simplify the expression, let $X_i = -\frac{M}{2}\left\lfloor \frac{W_i}{M} \right\rfloor^2 + \left(W_i - \frac{M}{2}\right)\left\lfloor \frac{W_i}{M} \right\rfloor$. Then, sum the steady state probability of all states to get (6). Therefore, τ, the probability of a station transmitting a request at a stage, could be expressed by

$$\tau = \sum_{i=0}^{m} b_{i,T} = \frac{b_{0,T}}{1-p} = \frac{W_0+1}{W_0+1+(1-p)X_0+(1-p)\sum_{i=1}^{m-1} X_i\left(\frac{p}{2}\right)^i + X_m\left(\frac{p}{2}\right)^m} \tag{7}$$

For $m = 0, M = 1$ which is SU PHY with a fixed contention window, check (5), the terms containing $X_i, i > 0$ will

$$1 = \sum_{i=0}^{m}\sum_{k=0}^{W_i} b_{i,k} = \frac{b_{0,T}}{W_0+1}\left(X_0 + \sum_{i=1}^{m-1} X_i \left(\frac{p}{2}\right)^i + X_m \frac{\left(\frac{p}{2}\right)^m}{1-p}\right) + \frac{b_{0,T}}{1-p} \tag{5}$$

$$= b_{0,T}\left(\frac{(1-p)X_0 + (1-p)\sum_{i=1}^{m-1} X_i \left(\frac{p}{2}\right)^i + X_m \left(\frac{p}{2}\right)^m + W_0 + 1}{(W_0+1)(1-p)}\right) \tag{6}$$

disappear, and $b_{0,T}/(1-p)$ degrades to $b_{0,T}$. As a result, (6) degrades to

$$1 = b_{0,T}\left(\frac{W_0+1+X_0}{W_0+1}\right). \tag{8}$$

Thereby,

$$\tau = b_{0,T} = \frac{W_0+1}{W_0+1+X_0} = \frac{2(W_0+1)}{W_0^2+W_0+2}, \tag{9}$$

which is different from that of CSMA/CA with constant window size [**?**], where $\tau = \frac{2}{W_0+1}$. That is because stations of CSMA/CA freezes backoff counter when sensing busy channel while no such freeze mechanism of backoff counter exists in OBRA.

On the other hand, conditional collision probability p has another relation with the probability τ. With ideal channel assumption that the collision happens only if at least one of other stations select the same RU, we have

$$p = 1 - \left(1 - \frac{\tau}{M}\right)^{n-1}. \tag{10}$$

Rewrite (10) as $\tau^\star = \left(1-(1-p)^{\frac{1}{n-1}}\right)M$. To obtain probability τ and p, we need to find solutions to a group of the two equations 7 and 10. $\tau^\star(p)$ is a monotonically increasing function. Though $\tau(p)$ is hard to determine the monotonicity from the expression of equation 7 with respect to p, the function 7 is estimated monotonically decreasing by numerical method. Also, $\tau(0) = \frac{W_0+1}{W_0+1+X_0} > \tau^\star(0) = 0$. And $\tau(1) < \tau^\star(1) = M$. Therefore, the only solution could be found by numerical method.

B. Random Access Efficiency

With the transmission probability τ, performance of OBRA mechanism could be easily evaluated. First, the expected number of stations who contend successfully at a stage, denoted by $E[n_s]$. Extending n_s, a system efficiency is defined as another important metric. Also, the access delay denoted by D, referred to as number of stages required for a station to contend successfully, could be derived.

1) n_s and System Efficiency: What we are concerned about is that how many stations contend successfully at a stage. Given τ and p, the probability that a station contend successfully in a stage is firstly derived: $P_{s\_station} = \tau(1-p)$. Then, with (10), $E[n_s]$ is easily expressed as follows.

$$E[n_s] = nP_{s\_station} = n\tau(1-p) = n\tau(1-\frac{\tau}{M})^{n-1} \tag{11}$$

Furthermore, normalizing n_s, *system efficiency* here is defined as the ratio of the number of successful contending stations in a stage to the number of RUs for random access in a stage.

$$\textit{eff}\ (\tau) = \frac{E[n_s]}{M} = \frac{n\tau(1-\frac{\tau}{M})^{n-1}}{M} \tag{12}$$

2) Access Delay: D represents the number of stages required for a station to contend successfully in a stage. Because of saturation assumption, a new request arrives once one previous request is transmitted successfully. Thus queueing waiting time is not considered here. In this way, the access delay D follows the geometric distribution with parameter $P_{s\_station}$, which is obtained just now. Then the expected value of access delay of request frame, $E[D]$, is

$$E[D] = \frac{1}{\tau(1-\frac{\tau}{M})^{n-1}}. \tag{13}$$

V. Model Validation

The same with scenario for analysis, only OBRA mechanism, i.e., Fig. 5, is concerned. Simulation runs with three-way handshake stage by stage. We run simulations for 1,000,000 stages with a variety of parameter sets $\{M, OCW_{min}, OCW_{max}\}$ and collects the information of the two variables, the number of successful attempt STAs n_s and expected access delay D. The values of results from both analysis and simulation are given in Fig. 8 and Table III. The results show that the Markov model precisely predicts the steady state behavior of the OBRA mechanism.

TABLE III: Analysis versus simulation: n_s and access delay with $m = 3, M = 9, OCW_{min} = 15$

n_s	analysis	simulation
$n = 1$	0.72727	0.72728
$n = 5$	2.23001	2.22335
$n = 10$	2.88954	2.88546
$n = 20$	3.29798	3.29857
delay	analysis	simulation
$n = 1$	1.37500	1.37499
$n = 5$	2.24214	2.24886
$n = 10$	3.46075	3.46565
$n = 20$	6.06432	6.06323

VI. Performance Evaluation

With above analysis, we could conveniently evaluate performance, including system efficiency and access delay, and the effects of system parameters. Since the model acts quite precisely, model analysis is solicited to evaluate the performance later.

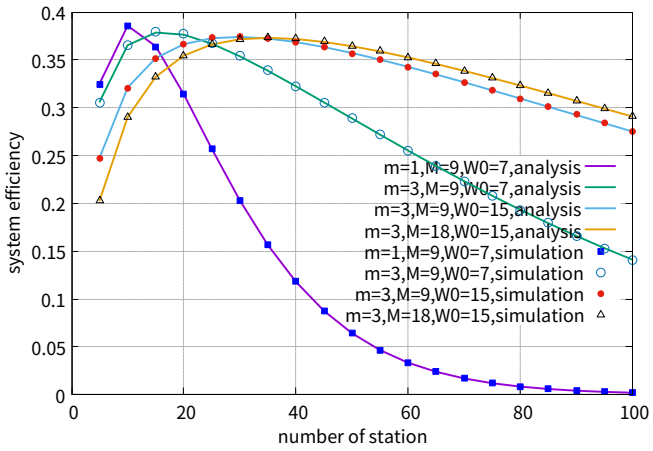


Fig. 8: System efficiency: Analysis versus Simulation

A. Maximum System Efficiency and Minimum Access Delay

With the system efficiency given in (12), take the derivative with respect to τ, and find the extreme point, $\tau^\star = M/n$. Since $\tau \in [0,1]$, $\tau^\star = min\{1, M/n\}$. In the dense scenario, i.e., n is large, then $\tau^\star = M/n$. The system efficiency thus is

$$\mathit{eff}\ (\tau^\star) = (1 - \frac{1}{n})^{n-1} \tag{14}$$

Then the maximum n_s is

$$E[n_s]^\star = M \cdot \mathit{eff}\ (\tau^\star) = M(1 - \frac{1}{n})^{n-1}. \tag{15}$$

The limit of system efficiency, based on infinite n, is

$$\lim_{n\to\infty} \mathit{eff}\ (\tau^\star) = \lim_{n\to\infty} (1 - \frac{1}{n})^{n-1} = \frac{1}{e} \tag{16}$$

which is the same with efficiency of SU slotted Aloha [**?**].

With the delay analysis given in (13), take the derivative with respect to τ, and find the extreme point, $\tau^\star = M/n$. Also, $\tau^\star = min\{1, M/n\}$. When $n \geq M$, the minimum access delay is

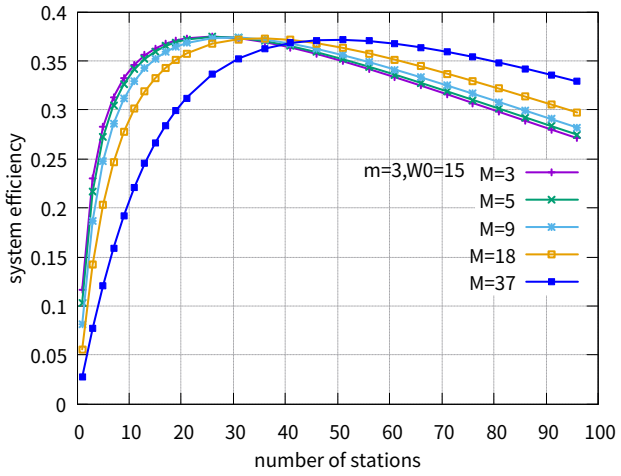
$$D(\tau^\star) = \frac{n}{M(1 - \frac{1}{n})^{n-1}}. \tag{17}$$

From above analysis, we find that the maximum system efficiency and minimum access delay are both obtained by the $\tau^\star = min\{1, M/n\}$. What's more, optimal system efficiency is independent with M, while M affects access delay. The larger M is, the shorter the access delay will be. It indicates that when AP allocates RUs for random access, the AP could allocate as many as possible.

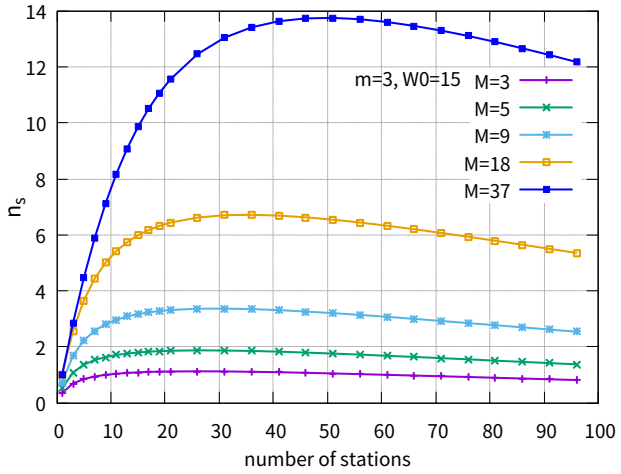
Since τ and p are obtained by solving group equations (7) and (10), it's hard to give a closed-form expression of τ only depending on system parameters, M, W_0, m and n. The only way to catch the insights of the OBRA is to tune the system parameters $\{M, W_0, W_m\}$ one by one.

B. RUs for Random Access M

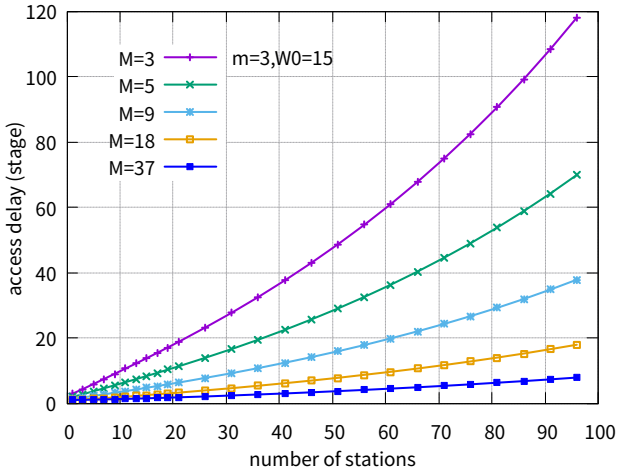
(14) has indicated that M, has nothing to do with optimal system efficiency. However, n_s and D are proportional and inversely proportional to M respectively according to (15) and (17). Following analysis results validate the statement.



(a) System efficiency versus number of stations



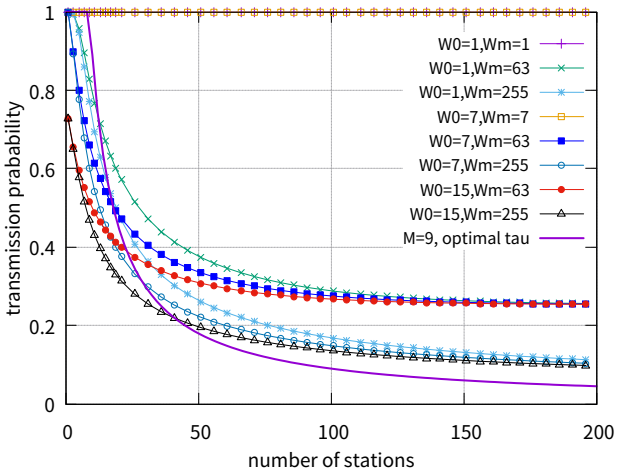
(b) Number of successful stations in a single stage versus number of stations



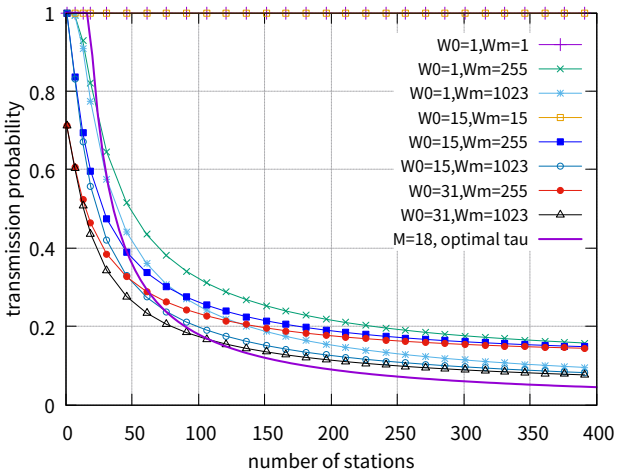
(c) Access delay versus number of stations

Fig. 9: Configure M

In Fig. 9a, the maximum system efficiency from various cases are almost the same, approaching $1/e$. The only difference is "where" the optimal point locates. Practically, the other two metrics, n_s and D, are closely related to M.

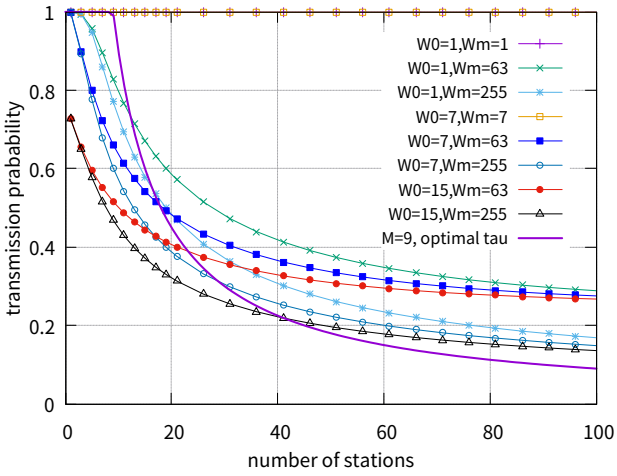


(a) Case 1, given $M = 9$

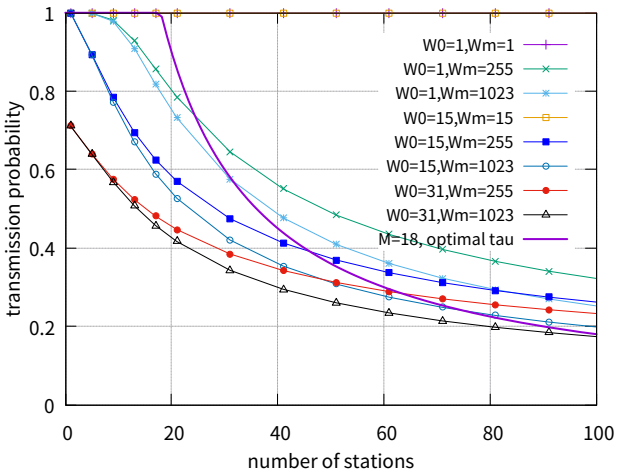


(b) Case 2, given $M = 18$

Fig. 10: Transmission probability versus number of stations



(a) Case 1, given $M = 9$



(b) Case 2, given $M = 18$

Fig. 11: Details of transmission probability versus number of stations when $n \leq 100$

Fig. 9b shows that larger M results more stations contend successfully in a stage. Moreover, Fig. 9c shows that larger M markedly decrease the access delay. Above all, when AP allocates RUs for random access, AP should allocate as many RUs as possible.

C. OCW_{min}, OCW_{max}

Fig. 10 shows case 1 ($M = 9$) and case 2 ($M = 18$), so that the rules we get are more convinced. In the figure, the purple line without point depicts the $\tau^\star$ that $\tau^\star = min\{1, M/n\}$. With above analysis and effects of M, we tune remaining parameters OCW_{min}, OCW_{max} so that τ approaches the optimal line. To see the trend of the line clearly, larger n is configured so that some rules could be obtained as follows.

First, the OCW_{min} or W_0, determines τ with small n. The larger W_0 is, the lower τ is at $n = 1$. That's why cases in Fig. 10 have two different start points.

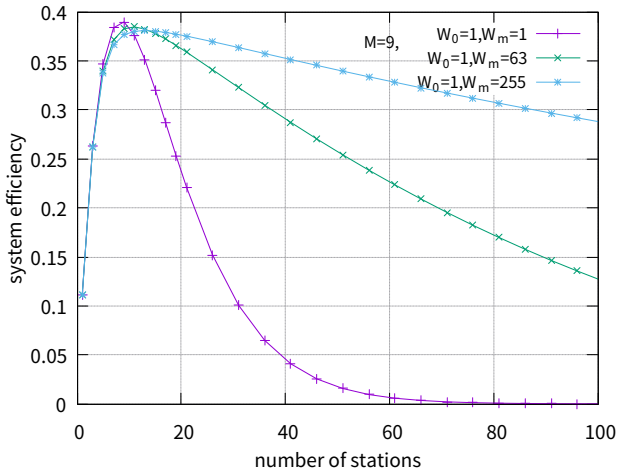
Secondly, $m = 0$ results in constant τ, which is consistent with (9) that τ does not depend on n. A special case, $W_0 < M$ for scenario $n \leq M$, results in constant τ equals to 1 regardless of n. It perfectly matches $\tau^\star$ at $n \leq M$ as $\tau^\star = 1$ for the scenario $n \leq M$. Thus, if given $n \leq M$, the optimal configuration will be $OCW_{max} = OCW_{min} < M$.

Thirdly, performance of the dense scenario strongly depends on OCW_{max}. It determines the limit of the τ, i.e., where the line of τ will converge. And both the two figures in Fig. 10 correspond with the above statement. And larger W_m causes lower τ, which is closer to $\tau^\star$ when n is large.

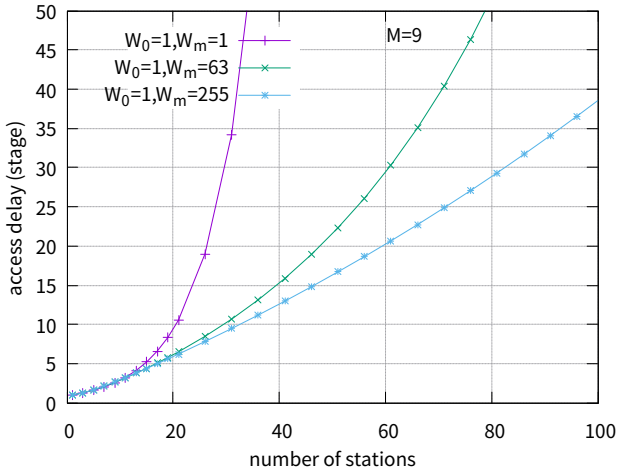
Then, two subfigures as in Fig. 11 are generated from Fig. 10 with n ranges from $[0, 100]$ to observe practical scenarios more clearly. Another observation whereby is that when $W_0 = 1$, line of τ will have a flat start, which happens to match better with $\tau^\star$ at $n \leq M$.

All above observations are summed up that W_0 has significant influence on scenario of $n \leq M$, while W_m counts when n is large, namely the dense scenario. In the next two subsections, the system efficiency and access delay under different parameter configurations are practically evaluated, whereby above observed effects of parameters could be conformed.

1) Configure OCW_{max}: With above rough rules, the effect of OCW_{max} is first evaluated by setting different OCW_{max}

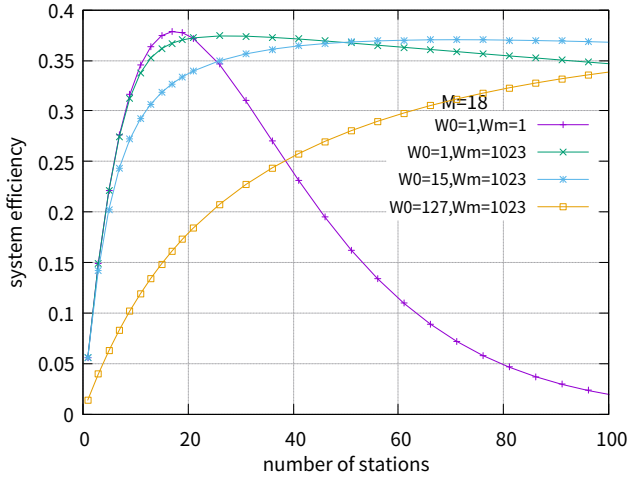


(a) System efficiency versus number of stations

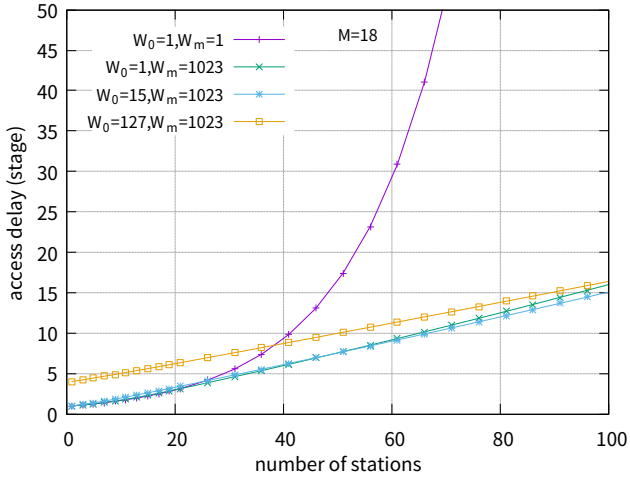


(b) Access delay versus number of stations

Fig. 12: Example of Configuring OCW_{max}, given $M = 9$



(a) System efficiency versus number of stations



(b) Access delay versus number of stations

Fig. 13: Example of Configuring OCW_{min}, given $M = 18$

while fixing $OCW_{min} = 1$ and $M = 9$. In Fig. 12a, three cases which has the same OCW_{min} are selected to clearly display the effect of OCW_{max} on system efficiency. From the figure, it is evident that larger OCW_{max} is much better for system efficiency when $n \geq M$. The result corresponds to the above rules obtained from the effect of parameters on τ. Additionally, larger OCW_{max} results in shorter access delay.

2) Configure OCW_{min}: To estimate the effect of OCW_{min}, the performance of different configurations of OCW_{min} while fixing large $OCW_{max} = 1023$, and $M = 18$. First, OCW_{min} determines the start of the probability τ, and it has a significant influence on the scenario of $n \leq M$. From Fig. 13a and 13b, we find larger $OCW_{min} = 127$ has lower system efficiency and larger access delay. Secondly, as stated in Section VI-C that $W_0 = W_m \leq M$ is the perfect configuration in the scenario of $n \leq M$, it is validated in Fig. 13 that maximum system efficiency and minimum access delay are achieved with the configuration in the scenario of $n \leq M$. However, the configuration of small OCW_{min} and large OCW_{max} almost achieves as good performance as the perfect configuration.

To this end, all observed rules are listed as follows.

1 For M: the larger the better

2 For OCW_{min}, OCW_{max}:
- OCW_{min} counts when $n \leq M$. Smaller OCW_{min} is better.
- OCW_{max} counts when $n > M$. Larger OCW_{max} is better.

- Special case: for $n \leq M$ $OCW_{max} = OCW_{min} < M$ is the optimal configuration.

VII. Conclusion

In this paper, we are the first to propose a Markov chain model conducting saturation analysis of the OBRA of 802.11ax. The simple Markov chain model is validated that it could accurately depict the steady state behavior of the OBRA. Thereby, closed-form expressions of system efficiency and access delay are derived. Finally, it is observed that performance strongly depends on the system parameters. Larger number of RUs or subchannels for random access results in more

successful contending stations in a stage. The initial contention window has significant influence when only a few stations exist, while maximum contention window counts in the dense scenario.

Different from DCF of legacy 802.11, the OBRA mechanism is more flexible, with system parameters dynamically configured. A real-time algorithm is required to configure the system parameters dynamically. This paper takes the first step to catch some insight from the steady state behavior of the OBRA mechanism. In the future, transient analysis is necessary to generate a configuration algorithm.

REFERENCES

Yang Hang Biography text here.

Der-Jiunn Deng Biography text here.

Kwang-Cheng Chen Biography text here.