Energy-Delay Analysis of MAC Protocols in Wireless Networks*

Shih Yu Chang, Wayne E. Stark and Achilleas Anastasopoulos

Electrical Engineering and Computer Science Department

University of Michigan

Ann Arbor, MI 48109-2122.

email: {shihyuc,stark,anastas}@umich.edu

Abstract

In this paper the tradeoff between energy and delay for wireless networks is studied. A network using a request-to-send (RTS) and clear-to-send (CTS) type medium access control (MAC) protocol is considered. We first determine the average delay incurred and the average energy consumed when the effects of an imperfect channel are incorporated in the model. In order to incorporate the relation between packet error probability, energy, and delay, we use the reliability function bounds on the error probability of a coded system. Further, we present a generic framework that allows us to obtain the joint statistics of energy and delay through their joint generating function. Several important design tradeoffs are studied from the joint generating function, such as the average energy with an outage delay constraint. This framework allows us to optimize over the system parameters for various objective functions, such as average delay. An approximation method is also proposed to calculate the average energy and average delay analytically. This approximation is found to be quite accurate for a wide range of lengths.

Index Terms—**Delay effects, crosslayer design, protocol analysis, wireless LAN (WLAN), random codes.**

I. INTRODUCTION

Recently there has been considerable interest in the design and performance evaluation of wireless local area networks (WLANs). Some WLANs must operate solely on battery power. In such cases it is

important to consider energy consumption in the system design and analysis. It is possible to reduce energy consumption by increasing delay incurred. Two critical components of a wireless network are the medium access control (MAC) protocol and the physical layer (PHY). The MAC protocol resolves conflicts between users attempting to access the channel. Generally users make reservations for transmissions in a decentralized way. Thus there is some amount of delay in accessing the channel and there is energy used in reserving the channel. An important component of the PHY layer is forward error control coding, which mitigates the effect of channel noise at the receiver. By transmitting redundant bits in addition to information bits, error control coding reduces the energy needed for transmission at the expense of increased delay.

There are many MAC protocols that have been developed for wireless voice and data communication networks. Typical examples include the time-division multiple access (TDMA), code-division multiple access (CDMA), and contention-based protocols such as IEEE 802.11 [1], [2]. In this paper, we adopt the MAC protocol used in the 802.11 standard. There are two basic techniques to access the medium in the 802.11 standard. The first one called the distributed coordination function (DCF), is employed when there is no centralized coordinator in the system to assign the medium to users in the network. The DCF is a random access scheme, based on carrier sense multiple access with collision avoidance (CSMA/CA). When packets collide or have errors, the transmitter performs a random backoff before retransmitting the packets. Another MAC method in the 802.11 standard called point coordinator function (PCF), is used when there is a centralized coordinator to coordinate the access of the medium. In this paper we focus on the DCF protocol for accessing the medium.

In order to combat the problem of hidden terminals [3], which occurs when some stations in the network are unable to detect each other, the 802.11 protocol uses a mechanism known as request-to-send/clear-to-send (RTS/CTS). Before transmitting the data packet, the source station sends a request-to-send (RTS) packet. If the RTS packet is received correctly at the destination station (there is no collision with other RTS packets sent by other competing stations and the receiver correctly decodes the packet) the destination station broadcasts a clear-to-send (CTS) packet. If the CTS packet is successfully received by the source station, the channel reservation is successful and the source station will begin to send data and wait for the acknowledgement (ACK) packet. The source station detects an unsuccessful channel reservation by the lack of a correct CTS packet.

Previous research on performance evaluation of 802.11 has been carried out by two methods. Crow [4], [5] and Weinmiller et al. [6] used computer simulations to evaluate the network throughput. In [7], [8], [9], the system performance was evaluated by an analytical model. Bianchi [10] used a simple but accurate

model that characterizes the random exponential backoff protocol. These papers did not incorporate channel noise in the analysis, which is an important factor in wireless network. Although Hadzi-Velkov and Spasenovski [11] considered the effects of packet errors in the analysis, they did not relate the packet error probability to the energy used and the number of redundant bits used for error control coding. The motivation for this paper is to understand the role of energy and codeword length (number of redundant bits) at the PHY layer have on the total energy and delay of the network. The contributions in this paper are as follows:

1) By modifying the approach proposed in [11], which combines the MAC layer and the PHY layer problem together, we determine average energy and delay of a successful packet transmission by taking the packet error probability into consideration.

2) We propose a state diagram representation the operation of the MAC layer and obtain the joint generating function of the energy and delay by incorporating the effects of the PHY layer. This is a universal approach and could be apply to other MAC protocols.

3) We optimize numerically over code rate to have minimum average transmission delay over different packet by introducing the random coding bound to represent the packet error probability. By changing the signal-to-noise ratio, the energy-delay tradeoff curves for minimum delay are obtained.

4) We propose an approximation method to express the energy-delay tradeoff curves analytically. The comparison of the energy-delay tradeoff curves evaluated from this approximation method with the exact energy-delay tradeoff curves (from numerical optimization) indicates that this approximation method is extremely accurate especially when the number of information bits per packet is large.

The remaining of this paper is organized as follows. In section II, we give a brief description for the protocol used in our analysis and introduce the system assumptions. In section III, we first analyze the average energy and delay for the system in subsection III-B. Then in subsection III-C, we discuss our system state diagram and utilize it to derive the joint generating function of energy and delay. The average energy with outage delay constraint and average delay with outage energy constraint are analyzed with generating function. Finally, the proposed approximation method for energy-delay tradeoff curves is given in subsection IV. We demonstrate that the energy and delay relationship with random coding under AWGN channel through numerical method in section V. Section VI gives the conclusion and future research.

II. System Description

The wireless networks that we analyze here have the following network layer specifications. First, each station with a fixed position can hear (detect and decode) the transmission of $n - 1$[1] other stations in the network. Second, stations always have a packet ready to transmit. Third, each station uses the 802.11 MAC protocol. At the PHY layer, a packet of K information bits is encoded into a packet of N coded symbols. It is assumed that the receivers have no multiple-access capability (i.e., they can only receive one packet at a time) and they cannot transmit and receive simultaneously. The packet error probability depends on the parameters K, N, and E_c/N_0, where E_c is the received coded symbol energy and N_0 is the one-sided power spectral density level of the thermal noise at the receiver.

In the following we give a brief description of the most salient features of the IEEE 802.11 MAC protocol (more details can be found in [1] and [2]). When a station is ready to transmit a packet, it senses the channel for DIFS seconds. If the channel is sensed idle, the transmission station picks a random number j, uniformly distributed in $\{0, 1, \ldots, W_i - 1\}$, where $W_i = 2^i W$ is the contention window (CW) size, i is the contention stage (initially $i = 0$), and W is the minimum CW size. A backoff time counter begins to count down with an initial value j: it decreases by one for every idle slot of duration σ seconds (also referred to as the standard slot) as long as the channel is sensed idle, stops the count down when the channel is sensed busy, and reactivates when the channel is sensed idle again. The station transmits an RTS packet when the counter counts down to zero. After transmitting the RTS packet, the station will wait for a CTS packet from the receiving station. If there is a collision of the RTS packet with other competing stations or a transmission error occurs in the RTS or CTS packet, the transmitting station doubles the CW size (increases the contention stage i by one) and picks another random number j as before. If there are no collisions or errors in the RTS and CTS packets, the station begins to transmit the data packet and waits for an acknowledgment (ACK) packet. However, if the data or the ACK packet is not successfully received, the CW size will also be doubled (the contention stage i will increase by one) and the transmitting station will join the contention period again. The contention stage is reset (i is set to zero) when the transmitting station receives an ACK correctly. A time diagram indicating the sequence of these events is depicted in Fig. 1. It is also noted that there is a maximum CW size (or equivalently, a maximum contention stage, m); when the transmitter is in this maximum stage and needs to join the contention period again, it does not increase further the CW, but picks a random number in $\{0, 1, \ldots, W_m - 1\}$.

---

[1]n is the density of stations for each station's neighbor.

III. ENERGY-DELAY ANALYSIS

In this section, we analyze the energy and delay characteristics of the wireless networks described above. The delay T_d of each data packet is defined as the time duration from the moment the backoff procedure is initiated until DIFS seconds after the ACK packet is received correctly by the transmitting station, as shown in Fig. 1. Similarly, the energy E_t is defined as the energy consumed by both transmitting and receiving stations in the duration of T_d. Without loss of the generality, for notational simplicity we assume that the propagation loss between transmitter and receiver is one (0 dB). We also assume that the propagation time is negligible. In this paper, we only consider the energy consumption for packet transmission and omit the energy required for signal processing and channel sensing. The system parameter SIFS is defined as the time between the end of a packet reception, say RTS and the beginning of a packet transmission, say CTS. This time includes the time required for decoding a packet and other processing functions at the receiver.

A. Three Nonlinear System Equations

In order to analyze energy and delay relationships, we need to define two random processes to characterize the backoff counter state and the CW size. The first process $b(\tau)$ represents the backoff time counter for the reference station (this is the station for which we evaluate energy and delay). The second random process $s(\tau)$ is used to represent the CW stage $i \in \{0, 1, \ldots, m\}$ of the station at time τ. In order to analyze the energy and delay characteristics it is sufficient to only consider the time instances the backoff counter (and CW stage) changes value. To this end we further define the discrete-time random processes $b_t = b(\tau_t)$ and $s_t = s(\tau_t)$, where τ_t is the time instance of the t-th change in value of $b(\tau)$. A realization of these random processes is shown in Fig. 2. In this realization, $b_3 = 4$ is the value of the counter just before being frozen and $b_4 = 3$ is the value of the counter after the channel has been sensed idle for DIFS seconds, and the counter becomes active again.

The first assumption made in this analysis is that the event of packet collision is independent of past collisions and thus independent of the contention stage. The second assumption is that the packet collision is identical for all states (values of b_t and s_t) of a user. As verified in [10], these two assumptions are extremely accurate when the number of stations in the network is large (say greater then 10). As a result of the above assumptions, the two dimensional process (s_t, b_t) is a discrete-time Markov chain. With the assumption that errors in transmission can occur only due to collisions, the one-step transition

probabilities developed in [10] are given by $P\{i, k \mid j, l\} = P\{s_{t+1} = i, b_{t+1} = k \mid s_t = j, b_t = l\}$ with

$$P\{i, k \mid i, k+1\} = \quad 1, \quad 0 \le k \le W_i - 2, 0 \le i \le m \tag{1a}$$

$$P\{0, k \mid i, 0\} = \quad \frac{1-p_c}{W_0}, \quad 0 \le k \le W_0 - 1, 0 \le i \le m \tag{1b}$$

$$P\{i, k \mid i-1, 0\} = \quad \frac{p_c}{W_i}, \quad 0 \le k \le W_i - 1, 1 \le i \le m \tag{1c}$$

$$P\{m, k \mid m, 0\} = \quad \frac{p_c}{W_m}, \quad 0 \le k \le W_m - 1, \tag{1d}$$

where p_c is the conditional collision probability, i.e., the probability of a collision given a packet transmitted on the channel. The first equation in (1a) corresponds to the decrement of the backoff counter at the beginning of each time slot. The second equation accounts for the fact that a new packet following a successful packet transmission starts at contention stage $i = 0$, and thus the backoff counter is initially uniformly chosen in the range of $\{0, 1, \ldots, W_0 - 1\}$. The other two cases describe the system evolution after an unsuccessful transmission. As described in the third equation, when an unsuccessful transmission occurs at contention stage $i - 1$, the contention stage increases and the backoff counter is initialized with a uniformly chosen value in the range $\{0, 1, \ldots, W_i\}$. Finally, the last case models the fact that the contention stage is not increased in subsequent packet transmissions when the contention window size reaches the maximum.

By modifying the Markov chain model described above, we can take into account packets errors as shown in Fig. 3. We denote the error probability of the four kinds of packets in the system as $P_{e,RTS}$, $P_{e,CTS}$, $P_{e,DT}$ and $P_{e,ACK}$. We assume that the channel is memoryless between packets. These probabilities depend on the particular channel, coding, modulation etc (a specific example will be given in Section V). A successful packet transmission requires that the RTS, CTS, DT, and ACK packets are received correctly. Let p_{ce} denote the probability of the complement of this event, i.e., collision in the RTS packet or error in any of the packets. This is also the transition probability from one contention stage to the next in the two-dimensional Markov chain, as shown in Fig. 3. Using similar assumptions as in [10] for the packet collision probability p_c, and since the events of packet collision and packet error are independent, the probability of p_{ce} can be expressed as

$$p_{ce} = p_c + (1 - p_c)[P_{e,RTS} + (1 - P_{e,RTS})P_{e,CTS} + (1 - P_{e,RTS})(1 - P_{e,CTS})P_{e,DT}$$

$$+ (1 - P_{e,RTS})(1 - P_{e,CTS})(1 - P_{e,DT})P_{e,ACK}]. \tag{2}$$

Following the derivation in [10], we can evaluate the stationary probability $P(i, j)$ of each state (i, j) of the Markov chain. Let p_{tx} be the probability of a transmitting station sending an RTS packet during each

backoff slots. The transmission probability p_{tx} and the collision probability p_c can be related as in [10]

$$p_{tx} \triangleq \sum_{i=0}^{m} P(i,0) = \frac{2(1-2p_{ce})}{(1-2p_{ce})(W+1) + p_{ce}W(1-(2p_{ce})^m)} \tag{3}$$

$$p_c = 1 - (1-p_{tx})^{n-1}. \tag{4}$$

From the above three nonlinear equations, (2)–(4), the probabilities p_{ce}, p_{tx} and p_c can be evaluated numerically. Another important probability that will be used in our analysis later is the probability of a transmission of an RTS packet from exactly one of the remaining $n-1$ stations given that at least one of the remaining stations is transmitting. It is denoted by p_{tx_1} and can be expressed as

$$p_{tx_1} = \frac{(n-1)p_{tx}(1-p_{tx})^{n-2}}{p_c}. \tag{5}$$

B. Mean System Energy Consumption and Delay

In this section, we analyze average energy and delay of our system. The average delay analysis is presented first. In each CW stage $i \in \{0,1,2,\dots,m\}$, the initial value of the backoff counter has mean of $W_i/2$, so that the average number of deferred slots (include standard slots and interrupted slots) before a retransmission attempt is $W_i/2$. The average number of consecutive standard slots n_s between two consecutive interrupted slots due to the transmission of the remaining $n-1$ stations can be evaluated as

$$n_s = \sum_{i=0}^{\infty} i(1-p_c)^i p_c = \frac{1}{p_c} - 1. \tag{6}$$

We define a renewal cycle as the period between two consecutive transmission of the $n-1$ remaining stations includes multiple consecutive standard slots and an interrupted slot as shown in Fig. 1. There are four possible cases in an interrupted slot. The first case is that the RTS packet from the remaining $n-1$ stations suffers collision or packet error. The second case is that the RTS packet from the remaining $n-1$ stations is correctly received and collision free but there is an error in the CTS packet transmission. The third case is that the RTS packet from the remaining $n-1$ stations is correctly received and no collision occurs, moreover there is no error in the CTS packet. However, there is an error happened in the DT packet transmission. The last case is to extend case three with a correct DT packet transmission so there is an ACK packet transmission. The average duration of a renewal cycle can be expressed as

$$
\begin{aligned}
T_{rc} &= n_s\sigma + [(1 - p_{tx_1}) + p_{tx_1}P_{e,RTS})]T_{s,RTS} \\
&\quad + [p_{tx_1}(1 - P_{e,RTS})P_{e,CTS}]T_{s,CTS} \\
&\quad + [p_{tx_1}(1 - P_{e,RTS})(1 - P_{e,CTS})P_{e,DT}]T_{s,DT} \\
&\quad + [p_{tx_1}(1 - P_{e,RTS})(1 - P_{e,CTS})(1 - P_{e,DT})]T_{s,ACK}, \quad (7)
\end{aligned}
$$

and

$$
\begin{aligned}
T_{s,RTS} &= N_{RTS}T_b + T_{DIFS} & (8a) \\
T_{s,CTS} &= (N_{RTS} + N_{CTS})T_b + T_{SIFS} + T_{DIFS} & (8b) \\
T_{s,DT} &= (N_{RTS} + N_{CTS} + N_{DT})T_b + 2T_{SIFS} + T_{DIFS} & (8c) \\
T_{s,ACK} &= (N_{RTS} + N_{CTS} + N_{DT} + N_{ACK})T_b + 3T_{SIFS} + T_{DIFS}, & (8d)
\end{aligned}
$$

where T_b is the time duration to transmit each coded bit, and T_{DIFS}, T_{SIFS} are system delay parameters defined by the standard.

Since the retransmission attempt of a reference station in contention stage i is on the average preceded by $W_i/[2(n_s + 1)] = W_i p_c/2$ renewal cycles of $n - 1$ remaining stations, the average time duration between two consecutive retransmissions of the reference station is $W_i p_c T_{rc}/2$. If the reference station fails to receive the ACK packet correctly, i.e., it is not a successful transmission, there are four possible cases of unsuccessful transmission. The previous three cases are the same as the first three cases of the above interrupted slot. But the last case will be modified as that the reference station receives ACK packet incorrectly. The average duration while the reference station itself occupies the channel during each unsuccessful retransmission attempt, denoted as T_{coe}, is

$$
\begin{aligned}
T_{coe} &= \frac{1}{p_{ce}}[(p_c + (1 - p_c)P_{e,RTS})T_{s,RTS} \\
&\quad + (1 - p_c)(1 - P_{e,RTS})P_{e,CTS}T_{s,CTS} \\
&\quad + (1 - p_c)(1 - P_{e,RTS})(1 - P_{e,CTS})P_{e,DT}T_{s,DT} \\
&\quad + (1 - p_c)(1 - P_{e,RTS})(1 - P_{e,CTS})(1 - P_{e,DT})P_{e,ACK}T_{s,ACK}]. \quad (9)
\end{aligned}
$$

The average time before the reference station makes its $(i+1)$-th retransmission attempt, denoted as $T_{r,i}$,

is

$$T_{r,i} = \sum_{k=0}^{i}(\frac{W_k}{2}p_cT_{rc}) + iT_{coe} \tag{10a}$$

$$= \begin{cases} iT_{coe} + W(2^{i+1} - 1)\frac{p_cT_{rc}}{2} & \text{for } 0 \le i \le m - 1 \\ iT_{coe} + W(2^{m+1} - 1 + 2^m(i - m))\frac{p_cT_{rc}}{2} & \text{for } m \le i. \end{cases} \tag{10b}$$

Finally, the average delay of a successful data packet transmission

$$\overline{T}_d = T_{s,ACK} + \sum_{i=0}^{\infty}(1 - p_{ce})p_{ce}^i T_{r,i} \tag{11a}$$

$$= T_{s,ACK} + T_{coe}\frac{p_{ce}}{1 - p_{ce}}$$

$$+T_{rc}[\frac{p_cW}{2}\frac{1 - 2(2p_{ce})^m(1 - p_{ce}) + p_{ce}^m(1 - 2p_{ce})}{1 - 2p_{ce}} + \frac{p_cW}{2}(2^{m+1} - 1 + \frac{2^m p_{ce}}{1 - p_{ce}})p_{ce}^m] \tag{11b}$$

The average transmission energy E_t of each data packet is the total energy consumed in the transmitting and receiving stations from the moment the backoff procedure is initiated until the moment the ACK packet is received by the transmitting station correctly. We will consider only the energy spent in packet transmission and assume that the energy consumption for each packet transmission is NE_c for a packet of length N. The four possibilities of consuming energy when the reference station fails to receive the ACK packet correctly are the same in the previous delay analysis. The average energy consumption can be expressed as

$$E_{coe} = \frac{1}{p_{ce}}[(p_c + (1 - p_c)P_{e,RTS})E_{s,RTS}$$

$$+(1 - p_c)(1 - P_{e,RTS})P_{e,CTS}E_{s,CTS}$$

$$+(1 - p_c)(1 - P_{e,RTS})(1 - P_{e,CTS})P_{e,DT}E_{s,DT}$$

$$+(1 - p_c)(1 - P_{e,RTS})(1 - P_{e,CTS})(1 - P_{e,DT})P_{e,ACK}E_{s,ACK}], \tag{12}$$

where

$$E_{s,RTS} = N_{RTS}E_c \tag{13a}$$

$$E_{s,CTS} = (N_{RTS} + N_{CTS})E_c \tag{13b}$$

$$E_{s,DT} = (N_{RTS} + N_{CTS} + N_{DT})E_c \tag{13c}$$

$$E_{s,ACK} = (N_{RTS} + N_{CTS} + N_{DT} + N_{ACK})E_c. \tag{13d}$$

In the above equations, $E_{s,RTS}$, $E_{s,CTS}$, $E_{s,DT}$, and $E_{s,ACK}$ denote the energy consumed for the four above mentioned cases. Hence, the average transmitting energy for a successful data packet transmission is

$$\overline{E}_t = E_{s,ACK} + \sum_{i=0}^{\infty} i(1 - p_{ce})p_{ce}^i E_{coe} \tag{14a}$$

$$= E_{s,ACK} + E_{coe}\frac{p_{ce}}{1 - p_{ce}}. \tag{14b}$$

C. Joint Generating Function of Energy and Delay

Our goal in this section is to obtain the joint generating function of energy and delay for a successful data packet transmission, which can be expressed as

$$G_s(X,Y) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} Pr(T_d = i\Delta_t, E_t = j\Delta_e)X^i Y^j, \tag{15}$$

where Δ_t, Δ_e are parameters that determine the resolution of our analysis (we choose $\Delta_t = T_b$ and $\Delta_e = E_c$ in the remaining part of our analysis). We further define the quantities $N_{SIFS} = T_{SIFS}/\Delta_t$, $N_{DIFS} = T_{DIFS}/\Delta_t$, and $N_\sigma = \sigma/\Delta_t$ make the additional assumption that N_{SIFS}, N_{DIFS} and N_σ are integers.

Based on the protocol description in Section II, the state flow diagram shown in Fig. 4 can be obtained. Each transition from state A to state B in this diagram is labelled with the conditional joint generating function of the additional energy and delay incurred when the protocol makes a transition from state A to state B. For instance, the generating function G_{rs} corresponding to a channel reservation success is

$$G_{rs}(X,Y) = (1 - p_c)(1 - P_{e,RTS})(1 - P_{e,CTS})X^{N_{RTS}+N_{CTS}+2N_{SIFS}}Y^{N_{RTS}+N_{CTS}}, \tag{16}$$

where $(1 - p_c)(1 - P_{e,RTS})(1 - P_{e,CTS})$ is the probability of not having a collision and having a correct reception of the RTS and CTS packets, $[N_{RTS} + N_{CTS} + N_{SIFS} + T_{DIFS}] \times T_b$ is the delay incurred during this transition, and $[N_{RTS} + N_{CTS}] \times E_c$ is the corresponding energy consumed. Similarly, the generating function G_{ts} corresponding to a data packet transmission success once the channel is reserved can be expressed as

$$G_{ts}(X,Y) = (1 - P_{e,DT})(1 - P_{e,ACK})X^{N_{DT}+N_{ACK}+N_{SIFS}+N_{DIFS}}Y^{N_{DT}+N_{ACK}}. \tag{17}$$

The generating functions associated with failure to either reserve a channel or to transmit a data packet are products of two generating functions. Each product contains a factor that is independent of the particular CW stage, and a factor that depends on the contention stage i. We first describe the factors that

are independent of the CW stage. The generating function G_{rf}, corresponding to a channel reservation failure is given by

$$
\begin{aligned}
G_{rf}(X,Y) & = [p_c + (1-p_c)P_{e,RTS}]X^{N_{RTS}+N_{DIFS}}Y^{N_{RTS}} \\
& \quad + (1-p_c)(1-P_{e,RTS})P_{e,CTS}X^{N_{RTS}+N_{CTS}+N_{SIFS}+N_{DIFS}}Y^{N_{RTS}+N_{CTS}}, \quad (18)
\end{aligned}
$$

the meaning of which is that a failure can be due to either a collision/RTS transmission error, or a CTS transmission error. Similarly, after reserving the channel, a data packet transmission failure is either due to a data packet transmission error, or an acknowledgement packet transmission error, which is captured by the generating function G_{tf} as follows

$$
\begin{aligned}
G_{tf}(X,Y) & = P_{e,DT}X^{N_{DT}+N_{DIFS}}Y^{N_{DT}} \\
& \quad + (1-P_{e,DT})P_{e,ACK}X^{N_{DT}+N_{ACK}+N_{SIFS}+N_{DIFS}}Y^{N_{DT}+N_{ACK}}. \quad (19)
\end{aligned}
$$

We now evaluate the generating functions denoted by $G_{bs,i}$ of the state diagram. This generating function characterizes the delay for the transmitting station from the instant of starting the backoff procedure to the instant that the backoff counter reaches to zero at stage i. We do not need to consider energy consumption here since the reference station is not transmitting during this period. Thus, $G_{bs,i}$ is not a function of the variable Y. The probability of a busy slot due to the transmission of other stations is p_c and this event is independent and identical for each slot from our previous assumptions. At contention stage i, the range of possible backoff slots is from 1 to $2^i W$. Let j be the backoff slot chosen uniformly from the above range. The number of the occupied slots in these j slots is binomialy distributed with parameters (j, p_c). Hence, the generating function, $G_{bs,i}$, is

$$
G_{bs,i}(X) = \sum_{j=1}^{2^i W} \frac{1}{2^i W} \sum_{k=0}^{j} \binom{j}{k} [(1-p_c)X^{N_\sigma}]^{j-k}(p_c G_{oc}(X))^k. \quad (20)
$$

where $G_{oc}(X)$ is the generating function of the delay due to an occupied slot. We define an occupied slot as a slot when the transmitting station senses the channel is busy due to the transmission of one of the $n-1$ remaining stations. In an occupied slot, there are four possible cases. The first case is when two or more RTS packets from the remaining $n-1$ stations collide or have a packet error. The second case is when the RTS packet is correctly received without collision but there is an error in the CTS packet transmission. The third case is when the RTS packet is correctly received without collision and there is no error in the CTS packet but there is an error in the DT packet transmission. The last case is when there is a correct reception of RTS, CTS, and DT and either correct or erroneous reception of the

ACK packet. Note that the time duration of successful or unsuccessful ACK transmission are the same. Therefore, the corresponding generating function is

$$
\begin{aligned}
G_{oc}(X) &= [(1 - p_{tx_1}) + p_{tx_1} P_{e,RTS}] X^{N_{RTS}+N_{DIFS}} \\
&\quad + p_{tx_1}(1 - P_{e,RTS}) P_{e,CTS} X^{N_{RTS}+N_{CTS}+N_{SIFS}+N_{DIFS}} \\
&\quad + p_{tx_1}(1 - P_{e,RTS})(1 - P_{e,CTS}) P_{e,DT} X^{N_{RTS}+N_{CTS}+N_{DT}+N_{SIFS}+N_{DIFS}} \\
&\quad + p_{tx_1}(1 - P_{e,RTS})(1 - P_{e,CTS})(1 - P_{e,DT}) X^{N_{RTS}+N_{CTS}+N_{DT}+N_{ACK}+N_{SIFS}+N_{DIFS}}.
\end{aligned}
\tag{21}
$$

From the state diagram and Mason's gain formula [12], we obtain the following backward recursive equations for the generating function G_s.

$$
f_m(X,Y) = \frac{G_{rs} G_{ts}}{1 - G_{rf} G_{bs,m} - G_{rs} G_{tf} G_{bs,m}}
\tag{22a}
$$

$$
f_{i-1}(X,Y) = G_{rs} G_{ts} + (G_{rf} + G_{rs} G_{tf}) G_{bs,i} f_i, \quad \text{for } i = 1, 2, \cdots, m
\tag{22b}
$$

$$
G_s(X,Y) = G_{bs,0} f_0
\tag{22c}
$$

where i is the index of the CW stage from 1 to m.

From the joint generating function, the average energy and delay can be easily evaluated as

$$
\overline{T}_d = \Delta_t \frac{\partial G_s}{\partial X} \Big|_{X=Y=1},
\tag{23a}
$$

$$
\overline{E}_t = \Delta_e \frac{\partial G_s}{\partial Y} \Big|_{X=Y=1}.
\tag{23b}
$$

Below, we will utilize the generating function to derive the average energy with an outage delay constraint.

D. Average Energy with Delay Constraint

We define the transform, dennoted X^+, for a sequence of numbers $(n_0, n_1, n_2, \cdots)$ by

$$
X^+(x) \triangleq \sum_{k=1}^{\infty} n_k x^k.
\tag{24}
$$

The average energy with delay constraint γ_d is calculated as follows. Let $n_d = \gamma_d/\Delta_t$ be the normalized delay constraint (which we assume is an integer). Then

$$\mathbf{E}[E_t \mid T_d \le \gamma_d] \quad = \quad \sum_{j=0}^{\infty} j\Delta_e Pr(E_t = j\Delta_e \mid T_d \le \gamma_d) \tag{25a}$$

$$= \quad \Delta_e \frac{\sum_{j=0}^{\infty} j \sum_{i=0}^{n_d} Pr(T_d = i\Delta_t, E_t = j\Delta_e)}{Pr(T_d \le \gamma_d)} \tag{25b}$$

$$= \quad \Delta_e \frac{\sum_{i=0}^{n_d} \sum_{j=0}^{\infty} j Pr(T_d = i\Delta_t, E_t = j\Delta_e)}{Pr(T_d \le \gamma_d)} \tag{25c}$$

$$= \quad \Delta_e \frac{\sum_{i=0}^{n_d} \sum_{j=0}^{\infty} j Pr(T_d = i\Delta_t, E_t = j\Delta_e)}{\sum_{i=0}^{n_d} Pr(T_d = i\Delta_t)} \tag{25d}$$

In order to evaluate the denominator of (25) we proceed as follows. We let $s_l = \sum_{i=0}^{l} Pr(T_d = i\Delta_t)$ and $S(X) = \sum_{l=0}^{\infty} s_l X^l$. Then we have

$$s_l - s_{l-1} \quad = \quad \sum_{j=0}^{\infty} Pr(T_d = l\Delta_t, E_t = j\Delta_e), \qquad l = 1, 2, \dots. \tag{26}$$

Multiplying both sides by X^l, summing over l, and taking into account that $s_0 = \sum_{j=0}^{\infty} Pr(T_d = 0\Delta_t, E_t = j\Delta_e)$, we have

$$S(X)(1 - X) \quad = \quad \sum_{l=0}^{\infty} \sum_{j=0}^{\infty} Pr(T_d = l\Delta_t, E_t = j\Delta_e) X^l$$

$$= \quad G_s(X, Y) \mid_{Y=1}. \tag{27}$$

Thus,

$$S(X) = \frac{G_s(X, Y) \mid_{Y=1}}{1 - X}. \tag{28}$$

Therefore, the denominator of (25) is the coefficient of term in $S(X)$ that has degree equal to n_d and can be found by inverse transforming X^+.

The numerator of (25) can also be expressed with system generating function G_s. By setting $u_l = \sum_{j=0}^{\infty} j Pr(T_d = l\Delta_t, E_t = j\Delta_e)$ and taking the transform of $\{u_l\}$ with dummy variable X, we get

$$U(X) \quad = \quad \frac{\partial G_s(X, Y)}{\partial Y} \mid_{Y=1}. \tag{29}$$

Following a similar procedure as in (26) and (27), the numerator becomes

$$\sum_{i=0}^{n_d} u_i = \Delta_e \frac{\frac{\partial G_s(X,Y)}{\partial Y} \mid_{Y=1}}{1 - X}. \tag{30}$$

Finally, $\mathbf{E}[E_t \mid T_d \le \gamma_d]$ can be written as

$$\mathbf{E}[E_t \mid T_d \le \gamma_d] \quad = \quad \Delta_e \frac{\text{coef}(\frac{\frac{\partial G_s(X,Y)}{\partial Y} \mid_{Y=1}}{1-X}, X^{n_d})}{\text{coef}(\frac{G_s(X,Y) \mid_{Y=1}}{1-X}, X^{n_d})}, \tag{31}$$

where $\mathrm{coef}(f(X), X^i)$ is the coefficient of the term X^i in the polynomial expansion of $f(X)$. By a similar approach as above, we can also determine the average delay with an energy constraint. The quantity $\mathbf{E}[T_d \mid E_t \leq n_e \Delta_e]$ can be written as

$$\mathbf{E}[T_d \mid E_t < n_e \Delta_e] = \Delta_t \frac{\mathrm{coef}(\frac{\frac{\partial G_s(X,Y)}{\partial X}|_{X=1}}{1-Y}, Y^{n_e})}{\mathrm{coef}(\frac{G_s(X,Y)|_{X=1}}{1-Y}, Y^{n_e})}. \tag{32}$$

IV. ENERGY-DELAY OPTIMIZATION

In order to evaluate the energy and delay characteristics, we need to have a relation between error probability, energy and delay for each type of packet in the system. We will use the reliability function bounds for a memoryless channel to determine the packet error probability. Let K be the number of information bits in a packet (specified by the 802.11 standard) and N be the number of coded symbols for this packet. Then there exist an encoder and decoder for which the packet error probability $P_{e,K,N}$ is bounded by

$$P_{e,K,N} \leq 2^{K-NR_0}, \tag{33}$$

where R_0 is the cutoff rate. For an additive white Gaussian noise (AWGN) channel using binary input the cutoff rate is $R_0 = 1 - \log_2(1 + e^{-E_c/N_0})$, where E_c/N_0 is the received signal-to-noise ratio per coded symbol. An interesting observation is that there is an optimal N for minimum delay. For small N the packet error probability of a random code will be large. Large packet error probability will increase the system delay and energy due to the high chance of packet retransmission. On the other hand, if N is large, the packet error probability for a random code will be small but the transmission time will be large. Our goal is to find the best N for each kind of packet to minimize system delay. To achieve this, we fix E_c/N_0, and optimize delay (average or energy constrained) over the corresponding packet lengths N_{RTS}, N_{CTS}, N_{DT} and N_{ACK}. Based on these optimal values, we evaluate the energy consumption (average or delay constrained). Repeating the above procedure for different E_c/N_0, we get a curve quantifying the tradeoff between energy and delay. For the case when the number of information bits is large, we propose an approximate method to determine the energy and delay tradeoff curves mentioned in above. The numerical results of energy and delay curves will be given in Section V. In this section, we will concentrate in deriving the approximate equations for energy and delay tradeoff curves.

The first step of this approximation is to derive the average delay $\overline{T}_d$ when there is only one transmitting station in the network. Some approximations are made to obtain the estimates of the packet error probabilities that achieve minimum average delay, $\overline{T}_d$. We then evaluate the corresponding packet lengths

by solving (33) for N (used as an equality) with the optimal packet error probabilities. After getting the optimal packet error probabilities and lengths, the average energy $\overline{E}_t$ and delay $\overline{T}_d$ are evaluated according to Section III-A and III-B.

We begin by considering the average delay with only one transmitting station that uses the same protocol as described in Section II. This means that there are no RTS packet collisions in the medium. We need to modify the analysis for the single station case by excluding the effects of packet collision. Therefore, the packet collision probability p_c becomes zero. We still use p_{ce} to represent the probability from one stage to next. An unsuccessful transmission can happen only due to packet errors. Now the probability p_{ce} can be rewritten as

$$
\begin{aligned}
p_{ce} \quad = \quad & P_{e,RTS} + (1 - P_{e,RTS})P_{e,CTS} + (1 - P_{e,RTS})(1 - P_{e,CTS})P_{e,DT} \\
& + (1 - P_{e,RTS})(1 - P_{e,CTS})(1 - P_{e,DT})P_{e,ACK}.
\end{aligned}
\tag{34}
$$

There are four possible cases of an unsuccessful transmission. The first case is packet error of the RTS packet. The second case is that the RTS packet is correctly received but there is an error in the CTS packet transmission. The third case is that the RTS packet is correctly received, there is no error in the CTS packet, but, there is an error happened in the DT packet transmission. The last case is that the RTS, CTS and DT packets are successful but there is an error at the ACK packet. The time duration of each above cases is $T_{s,RTS}, T_{s,CTS}, T_{s,DT}$ and $T_{s,ACK}$. The average duration while the reference station itself occupies the channel during each unsuccessful retransmission attempt, denoted as T_{coe}, is

$$
\begin{aligned}
T_{coe} \quad = \quad & \frac{1}{p_{ce}}[P_{e,RTS}T_{s,RTS} + (1 - P_{e,RTS})P_{e,CTS}T_{s,CTS} \\
& + (1 - P_{e,RTS})(1 - P_{e,CTS})P_{e,DT}T_{s,DT} \\
& + (1 - P_{e,RTS})(1 - P_{e,CTS})(1 - P_{e,DT})P_{e,ACK}T_{s,ACK}].
\end{aligned}
\tag{35}
$$

The average elapsed time $T_{r,i}$ before the reference station makes its $(i+1)th$ retransmission attempt is

$$
\begin{aligned}
T_{r,i} \quad = \quad & (\sum_{k=0}^{i} \frac{W_k \sigma}{2}) + iT_{coe} \\
= \quad & \begin{cases} iT_{coe} + \frac{W\sigma(2^{i+1}-1)}{2} & \text{for } 0 \le i \le m-1 \\ iT_{coe} + \frac{W\sigma(2^{m+1}-1+2^m(i-m))}{2} & \text{for } m \le i. \end{cases}
\end{aligned}
\tag{36}
$$

Observed that (36) is not derived from (10) by substituting $p_c = 0$. The reason is that (10) is derived under the assumption of large n while (36) is derived separately for $n = 1$ since the large n approximation

does not hold. Finally, the average system delay for successful transmission is

$$
\begin{aligned}
\overline{T}_d &= T_{s,ACK} + \sum_{i=0}^{\infty}(1 - p_{ce})p_{ce}^{i}T_{r,i} \\
&= T_{s,ACK} + T_{coe}\frac{p_{ce}}{1 - p_{ce}} \\
&\quad + \frac{W\sigma}{2}[\frac{1 - 2(2p_{ce})^m(1 - p_{ce}) + p_{ce}^m(1 - 2p_{ce})}{1 - 2p_{ce}} + (2^{m+1} - 1 + \frac{2^m p_{ce}}{1 - p_{ce}})p_{ce}^m].
\end{aligned}
\tag{37}
$$

If we assume that the optimal error probability of each packet is quite small (i.e., $P_{e,RTS} \ll 1$, $P_{e,CTS} \ll 1$, $P_{e,DT} \ll 1$ and $P_{e,ACK} \ll 1$), then p_{ce} will also be very small compared to 1, as can be seen from the (34). The validity of this assumption will be verified later. The average system delay in (37) can be approximated by keeping the first order terms in P_{ce} as

$$
\overline{T}_d \approx \widetilde{T}_d = T_{s,ACK} + T_{coe}\frac{p_{ce}}{1 - p_{ce}} + \frac{W\sigma}{2}(1 + 2p_{ce}).
\tag{38}
$$

Using the assumption that the number of information bits K is large for each kind of packet, T_{SIFS} and T_{DIFS} can be ignored in the evaluation of $T_{s,ACK}$ and T_{coe}, resulting in

$$
\begin{aligned}
T_{s,ACK} &= (N_{RTS} + N_{CTS} + N_{DT} + N_{ACK})T_b + 3T_{SIFS} + T_{DIFS} \\
&\approx (N_{RTS} + N_{CTS} + N_{DT} + N_{ACK})T_b,
\end{aligned}
\tag{39}
$$

and

$$
\begin{aligned}
T_{coe} &= \frac{1}{p_{ce}}[P_{e,RTS}T_{s,RTS} + (1 - P_{e,RTS})P_{e,CTS}T_{s,CTS} \\
&\quad + (1 - P_{e,RTS})(1 - P_{e,CTS})P_{e,DT}T_{s,DT} \\
&\quad + (1 - P_{e,RTS})(1 - P_{e,CTS})(1 - P_{e,DT})P_{e,ACK}T_{s,ACK}] \tag{40a} \\
&= \frac{1}{p_{ce}}[P_{e,RTS}(N_{RTS}T_b + T_{DIFS}) + (1 - P_{e,RTS})P_{e,CTS}((N_{RTS} + N_{CTS})T_b + T_{SIFS} + T_{DIFS}) \\
&\quad + (1 - P_{e,RTS})(1 - P_{e,CTS})P_{e,DT}((N_{RTS} + N_{CTS} + N_{DT})T_b + 2T_{SIFS} + T_{DIFS}) \\
&\quad + (1 - P_{e,RTS})(1 - P_{e,CTS})(1 - P_{e,DT})P_{e,ACK}((N_{RTS} + N_{CTS} + N_{DT} + N_{ACK})T_b \\
&\quad + 3T_{SIFS} + T_{DIFS})] \tag{40b} \\
&\approx \frac{1}{p_{ce}}[P_{e,RTS}N_{RTS}T_b + (1 - P_{e,RTS})P_{e,CTS}(N_{RTS} + N_{CTS})T_b \\
&\quad + (1 - P_{e,RTS})(1 - P_{e,CTS})P_{e,DT}(N_{RTS} + N_{CTS} + N_{DT})T_b \\
&\quad + (1 - P_{e,RTS})(1 - P_{e,CTS})(1 - P_{e,DT})P_{e,ACK}(N_{RTS} + N_{CTS} + N_{DT} + N_{ACK})T_b]. \tag{40c}
\end{aligned}
$$

By using (39) and (40) in (38), the approximate average delay of the system becomes

$$
\begin{aligned}
\widetilde{T}_d &= (N_{RTS} + N_{CTS} + N_{DT} + N_{ACK})T_b + \frac{1}{1 - p_{ce}}[P_{e,RTS}N_{RTS}T_b \\
&\quad + (1 - P_{e,RTS})P_{e,CTS}(N_{RTS} + N_{CTS})T_b \\
&\quad + (1 - P_{e,RTS})(1 - P_{e,CTS})P_{e,DT}(N_{RTS} + N_{CTS} + N_{DT})T_b \\
&\quad + (1 - P_{e,RTS})(1 - P_{e,CTS})(1 - P_{e,DT})P_{e,ACK}(N_{RTS} + N_{CTS} + N_{DT} + N_{ACK})T_b] \\
&\quad + \frac{W\sigma}{2}(1 + 2p_{ce}) \qquad\qquad (41a)
\end{aligned}
$$

$$
\begin{aligned}
&\approx (N_{RTS} + N_{CTS} + N_{DT} + N_{ACK})T_b + [P_{e,RTS}N_{RTS}T_b \\
&\quad + P_{e,CTS}(N_{RTS} + N_{CTS})T_b + P_{e,DT}(N_{RTS} + N_{CTS} + N_{DT})T_b \\
&\quad + P_{e,ACK}(N_{RTS} + N_{CTS} + N_{DT} + N_{ACK})T_b] \\
&\quad + \frac{W\sigma}{2}(1 + 2(P_{e,RTS} + P_{e,CTS} + P_{e,DT} + P_{e,ACK})), \qquad (41b)
\end{aligned}
$$

where we have used the assumption that the optimal packet error probabilities are small.

Our goal is to find the packet lengths N^*_{RTS}, N^*_{CTS}, N^*_{DT} and N^*_{ACK} that minimize the average system delay $\widetilde{T}_d$. After taking the partial derivative with respect to the four packet lengths and setting them equal to zero, we get four packet error probabilities that achieve minimum average delay. We will derive the optimal packet probability of ACK as an example. The other three optimal packet error probabilities can be obtained in a similar way. We have

$$
\begin{aligned}
\frac{\partial \widetilde{T}_d}{\partial N_{ACK}} &= T_b(1 + P_{e,ACK}) - W\sigma P_{e,ACK}R_0 \ln(2) \\
&\quad - P_{e,ACK}R_0 \ln(2)(N_{RTS} + N_{CTS} + N_{DT} + N_{ACK}) \\
&= T_b(1 + P_{e,ACK}) - W\sigma P_{e,ACK}R_0 \ln(2) \\
&\quad + P_{e,ACK}[\ln(2^{-N_{RTS}R_0}) + \ln(2^{-N_{CTS}R_0}) + \ln(2^{-N_{DT}R_0}) + \ln(2^{-N_{ACK}R_0})] \\
&= T_b(1 + P_{e,ACK}) - W\sigma P_{e,ACK}R_0 \ln(2) \\
&\quad + P_{e,ACK}[\ln(2^{-N_{RTS}R_0}2^{K_{RTS}}2^{-K_{RTS}}) + \ln(2^{-N_{CTS}R_0}2^{K_{CTS}}2^{-K_{CTS}}) \\
&\qquad + \ln(2^{-N_{DT}R_0}2^{K_{DT}}2^{-K_{DT}}) + \ln(2^{-N_{ACK}R_0}2^{K_{ACK}}2^{-K_{ACK}})] \\
&= T_b(1 + P_{e,ACK}) - W\sigma P_{e,ACK}R_0 \ln(2) + P_{e,ACK}(\ln(P_{e,RTS}2^{-K_{RTS}}) \\
&\quad + \ln(P_{e,CTS}2^{-K_{RTS}}) + \ln(P_{e,DT}2^{-K_{DT}}) + \ln(P_{e,ACK}2^{-K_{ACK}}))
\end{aligned}
$$

$$= T_b(1 + P_{e,ACK}) - W\sigma P_{e,ACK} R_0 \ln(2)$$

$$+ P_{e,ACK}[\ln P_{e,RTS} - K_{RTS}\ln(2) + \ln P_{e,CTS} - K_{CTS}\ln(2)$$

$$+ \ln P_{e,DT} - K_{DT}\ln(2) + \ln P_{e,ACK} - K_{ACK}\ln(2)]$$

$$\approx T_b - W\sigma P_{e,ACK} R_0 \ln(2) - P_{e,ACK}\ln(2)(K_{RTS} + K_{CTS} + K_{DT} + K_{ACK}). \qquad (42)$$

By setting (42) equal to zero, we obtain the optimal probability for the ACK packet as

$$P^*_{e,ACK} \approx \frac{1}{\ln(2)(K_{RTS} + K_{CTS} + K_{DT} + K_{ACK} + W\sigma R_0/T_b)}. \qquad (43a)$$

Using a similar derivation as above, the other optimal packet error probabilities are given as

$$P^*_{e,RTS} \approx \frac{T_b}{T_b + \ln(2)(T_b K_{RTS} + W R_0)} \qquad (43b)$$

$$P^*_{e,CTS} \approx \frac{T_b}{T_b + \ln(2)(T_b(K_{RTS} + K_{CTS}) + W R_0)} \qquad (43c)$$

$$P^*_{e,DT} \approx \frac{T_b}{T_b + \ln(2)(T_b(K_{RTS} + K_{CTS} + K_{DT}) + W R_0)}. \qquad (43d)$$

Note that the optimal packet error probabilities in (43) are consistent with the assumption that they are small compared to 1 when K_{RTS}, K_{CTS}, K_{DT} and K_{ACK} are large. Furthermore, we can also solve for the optimal packet lengths by using (33) as an equality, thus obtaining

$$N^*_{RTS} \approx \frac{K_{RTS} + \log_2(\ln(2)(K_{RTS} + W\sigma R_0/T_b))}{R_0} \qquad (44a)$$

$$N^*_{CTS} \approx \frac{K_{CTS} + \log_2(\ln(2)((K_{RTS} + K_{CTS}) + W\sigma R_0/T_b))}{R_0} \qquad (44b)$$

$$N^*_{DT} \approx \frac{K_{DT} + \log_2(\ln(2)((K_{RTS} + K_{CTS} + K_{DT}) + W\sigma R_0/T_b))}{R_0} \qquad (44c)$$

$$N^*_{ACK} \approx \frac{K_{ACK} + \log_2(\ln(2)(K_{RTS} + K_{CTS} + K_{DT} + K_{ACK}) + W\sigma R_0/T_b))}{R_0}. \qquad (44d)$$

For the single user case, we can derive the approximate energy-delay values analytically from (43) and (44). Since p^*_{ce} is evaluated from (34) with (43) and $T^*_{s,ACK}$, T^*_{coe}, $E^*_{s,ACK}$ and E^*_{coe} can be determined from (8), (12), (13), and (35) with (44), the energy-delay values can be analytically obtained as:

$$(\widehat{T}^*_d, \widehat{E}^*_t) = (T^*_{s,ACK} + \frac{p^*_{ce}}{1 - p^*_{ce}} T^*_{coe} + B^*_{single}, E^*_{s,ACK} + \frac{p^*_{ce}}{1 - p^*_{ce}} E^*_{coe}), \qquad (45)$$

where B^*_{single} is

$$B^*_{single} = \frac{W\sigma}{2}[\frac{1 - 2(2p^*_{ce})^m(1 - p^*_{ce}) + p^{*m}_{ce}(1 - 2p^*_{ce})}{1 - 2p^*_{ce}} + (2^{m+1} - 1 + \frac{2^m p^*_{ce}}{1 - p^*_{ce}})p^{*m}_{ce}]. \qquad (46)$$

For the multiuser case, the probabilities of p^*_{ce}, p^*_c, p^*_{tx} and $p^*_{tx_1}$ will be evaluated according to (2), (3), (4) and (5) with (43). By using the optimal probabilities of p^*_{ce}, p^*_c, p^*_{tx} and $p^*_{tx_1}$ and the optimal packet

lengths obtained from (44), the approximate energy-delay values for the multiuser case can be obtained in a similar manner to the single user case as

$$(\widehat{T}_d^*, \widehat{E}_t^*) = (T_{s,ACK}^* + \frac{p_{ce}^*}{1 - p_{ce}^*}T_{coe}^* + B_{multi}^*T_{rc}^*, E_{s,ACK}^* + \frac{p_{ce}^*}{1 - p_{ce}^*}E_{coe}^*), \qquad (47)$$

where B_{multi}^* is

$$B_{multi}^* = \frac{p_c^*W\sigma}{2}[\frac{1 - 2(2p_{ce}^*)^m(1 - p_{ce}^*) + p_{ce}^{*m}(1 - 2p_{ce}^*)}{1 - 2p_{ce}^*} + (2^{m+1} - 1 + \frac{2^m p_{ce}^*}{1 - p_{ce}^*})p_{ce}^{*m}]. \quad (48)$$

V. Numerical results

In this section we present and discuss our numerical results for various scenarios and system parameters. Table I summarizes the system parameters used to obtain the numerical results. In all plots, both energy and delay are normalized with respect to K_{DT} to allow for a fair comparison. In addition, we normalize the energy by N_0 since all results depend on the ratio of energy to noise power spectral density.

In Fig. 5, we plot the average energy and delay tradeoff curves for different K_{DT}. The interesting observation in Fig. 5 is that while both energy and delay increase with K_{DT}, the normalized energy and delay decrease. The explanation is that the delay or energy consumption of the overhead transmissions, such as RTS, CTS and ACK, become less significant for larger data packet lengths.

To study the effect of different users, we fix the data packet length as $K_{DT} = 6400$ and plot the average energy-delay curves for different n. We can see from Fig. 6 that both energy and delay increase with the number of users. We also plot the curve for the single user case for the same protocol (802.11) and the simple ARQ protocol. These two curves are lower bounds to the multiuser curves. Note that the minimum normalized average energy achieved at high delay approaches 1.42 dB, as predicted by the channel capacity.

The delay variance can be calculated from the generating function as $\sigma_d^2 = \Delta_t^2 \frac{\partial^2 G_s}{\partial X^2}\mid_{X=Y=1} + \Delta_t\overline{T}_d - \overline{T}_d^2$. In Fig. 7, we plot the normalized standard deviation of the delay vs. the normalized average delay for different n and K_{DT}. We observe that the normalized standard deviation of the delay decreases as K_{DT} increases. It is also clear from the figure that the normalized standard deviation is grown in proportion to the normalized average delay with a factor of 1.5.

The average energy-delay tradeoff curves with different delay constraints are shown in Fig. 8. We fix E_c/N_0 and optimize $\overline{T}_d$ over $N_{RTS}, N_{CTS}, N_{DT}, N_{ACK}$. Let $N_{RTS}^*, N_{CTS}^*, N_{DT}^*, N_{ACK}^*$ be the packet lengths that minimize $\overline{T}_d$. The system generating function with these packet lengths and E_c/N_0 is denoted as G_s^*. If outage delay probability is given, we can evaluate both outage delay γ_d and $\mathbf{E}[E_t \mid T_d < \gamma_d]$ from G_s^*. Repeating the above procedure for different E_c/N_0, we obtain the average energy-delay tradeoff

curve with a delay constraint. Fig. 8 shows that when the outage delay probability increases, the range of outage delay decreases and conditional average energy also decreases. For comparison, we also plot the average energy vs. average delay curve. We see that the average energy used can be quite different depending on whether an average delay or a strict delay requirement is imposed.

The average energy-delay tradeoff curves with different energy constraints is given in Fig. 9. Given an outage energy probability constraint, we can evaluate both the outage energy γ_e and $\mathbf{E}[T_d \mid E_t < \gamma_e]$. Then using the same procedure as above we obtain the average energy-delay tradeoff curve with an energy constraint. The observations made for Fig. 8 apply to Fig. 9 as well.

In Fig. 10, we plot energy-delay curves of under different K_{DT} with $n = 10$ users in the network. It is observed that the approximate method is very accurate over a wide range of K_{DT}.

VI. CONCLUSION

In this paper we obtained the joint distribution of energy and delay for packet transmission using RTS/CTS-type MAC protocols taking into account the PHY layer. By representing the protocol using a state diagram we derived the joint generating function of energy and delay. The generating function was used to obtain various statistics such as the average energy and delay, the average energy with a delay constraint, etc. This approach allows us to optimize the performance over the block lengths used for different packets. Finally, a very accurate analytical approximation for the minimum average delay and corresponding delay was derived that enabled us to get the optimal energy-delay tradeoff curves analytically. We believe that our methods can be extended to the analysis of other protocols, including segmentation of data packets or retransmission of data segments without channel reservation.

REFERENCES

[1] *IEEE Standard for 802.11 Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, Std., Nov. 1997.

[2] *IEEE Standard for 802.11 Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, Std., Nov. 1999.

[3] L. Kleinrock and F. Tobagi, "Packet switching in radio channels, part II-the hidden terminal problem in carrier sense multiple access and the busy tone solution," *IEEE Trans. Communications*, vol. 33, no. 12, pp. 1417–1433, Dec. 1975.

[4] B. P. Crow, "Performance evaluaion of the IEEE 802.11 wireless local area network protocol," Master's thesis, University of Arizona, Tucson, AZ, 1996.

[5] G. Bianchi, L. Fratta, and M. Oliveri, "Performance analysis of IEEE 802.11 CSMA/CA medium access control protocol," in *Proc. IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications*, Taipei, Taiwan, Oct. 1996, pp. 407–411.

[6] J. Weinmiller, M. Schlager, A. Festag, and A. Wolisz, "Performance study of access control in wireless LANs IEEE 802.11 DFWMAC and ETSI RES 10 HIPERLAN," *Mobile Networks and Applications*, vol. 2, no. 1, pp. 55–67, June 1997.

[7] T. S. Ho and K. C. Chen, "Performance evaluation and enhancement of the CAMA/CA MAC protocol for 802.11 wireless LANs," in *Proc. IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications*, Taipei, Taiwan, Oct. 1996, pp. 392–396.

[8] H. S. Chhaya and S. Gupta, "Performance modeling of asynchronous data transfer methods of IEEE 802.11 MAC protocol," *Wireless Networks*, vol. 3, no. 3, pp. 217–234, Jan. 1997.

[9] F. Cali, M. Conti, and E. Gregori, "IEEE 802.11 wireless LAN: capacity analysis and protocol enhancement," in *Proc. INFOCOM.*, San Francisco, CA, Mar. 1998, pp. 142–149.

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

[11] Z. Hadzi-Velkov and B. Spasenovski, "Saturation throughput-delay analysis of IEEE 802.11 DCF in fading channel," in *Proc. International Communications Conf.*, vol. 1, Sept. 2004, pp. 121–126.

[12] G. F. Franklin, J. D. Powell, and A. Emami-Naeini, Eds., *Feedback Control of Dynamic Systems*. Addision Wesely, 1994.

TABLE I

SYSTEM PARAMETERS FOR NUMERICAL RESULTS

K_{RTS}	128 bits
K_{CTS}	128 bits
K_{ACK}	128 bits
Channel Bit Rate, $1/T_b$	1 Mbit/s
Slot Duration, σ	50μs
Maximum Contention Stage, m	5
Minimum Contention Window Size, W	8 slots
T_{SIFS}	28μs
T_{DIFS}	128μs
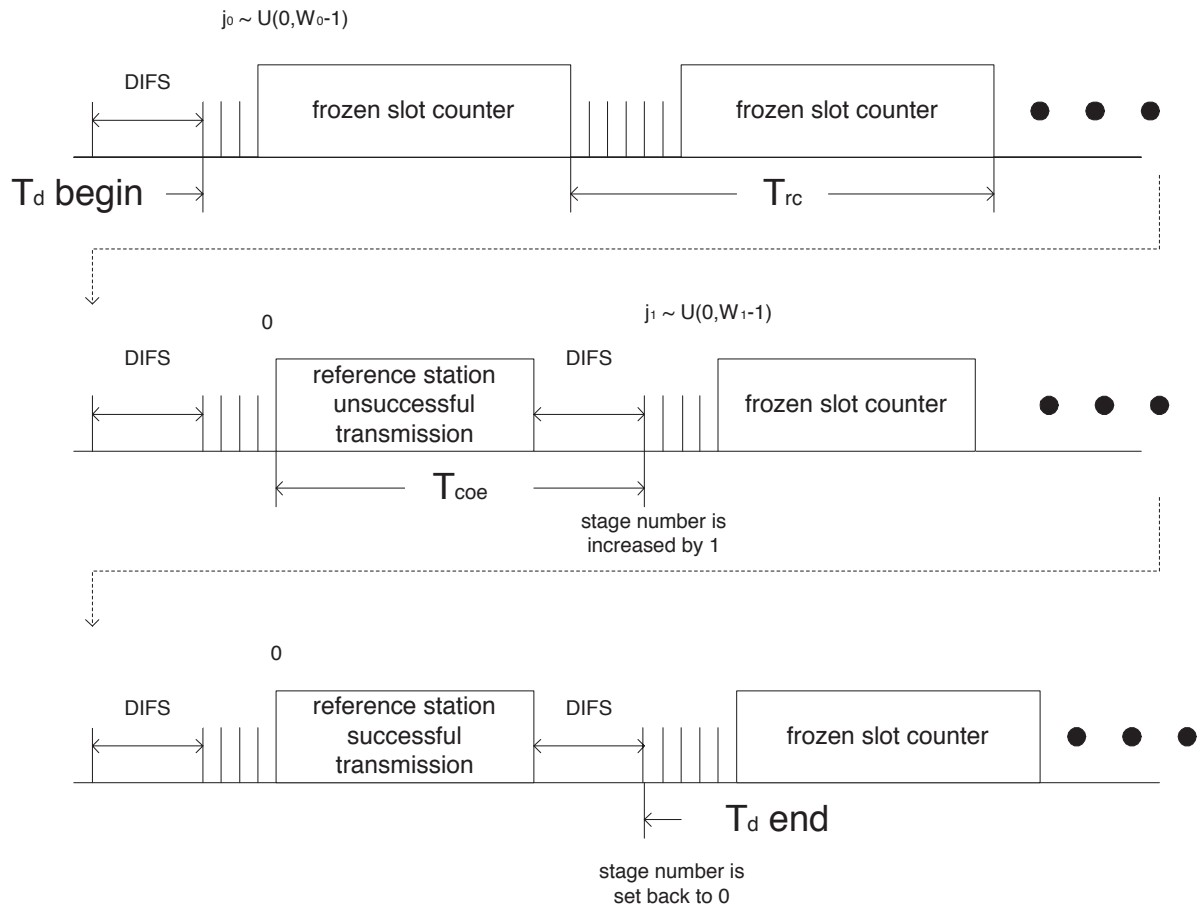
Fig. 1.   Timing diagram for the protocol under investigation. The numbers j_0 and j_1 represent random numbers at stage $i = 0$, and $i = 1$, respectively.
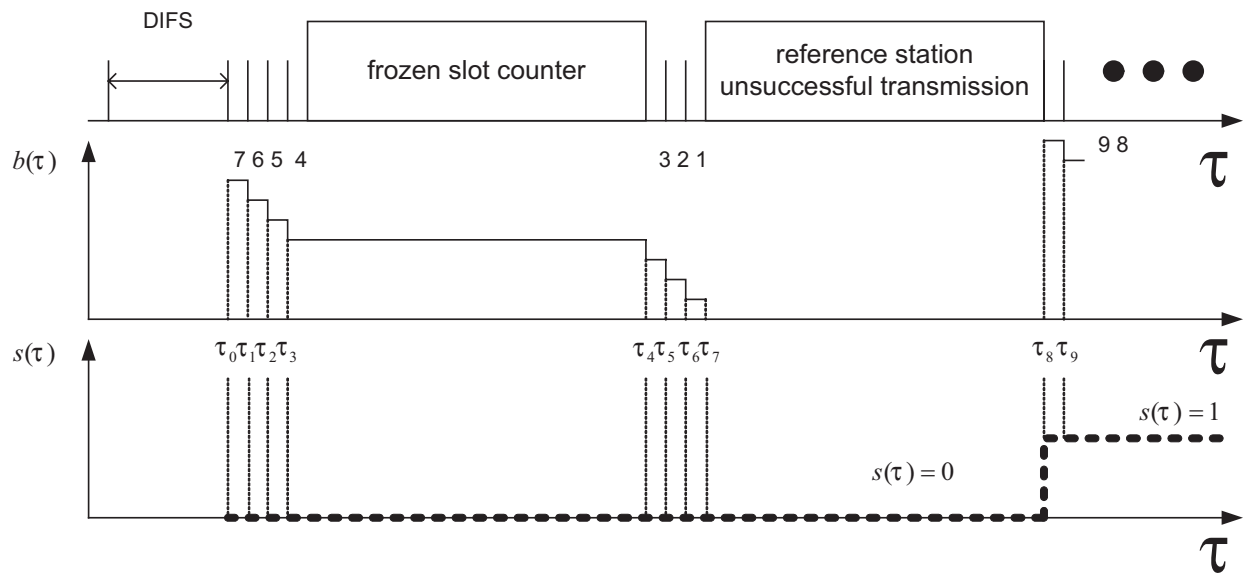
Fig. 2. Illustration of the random processes, $b(\tau)$, $s(\tau)$ and their discrete-time counterparts b_t, s_t.
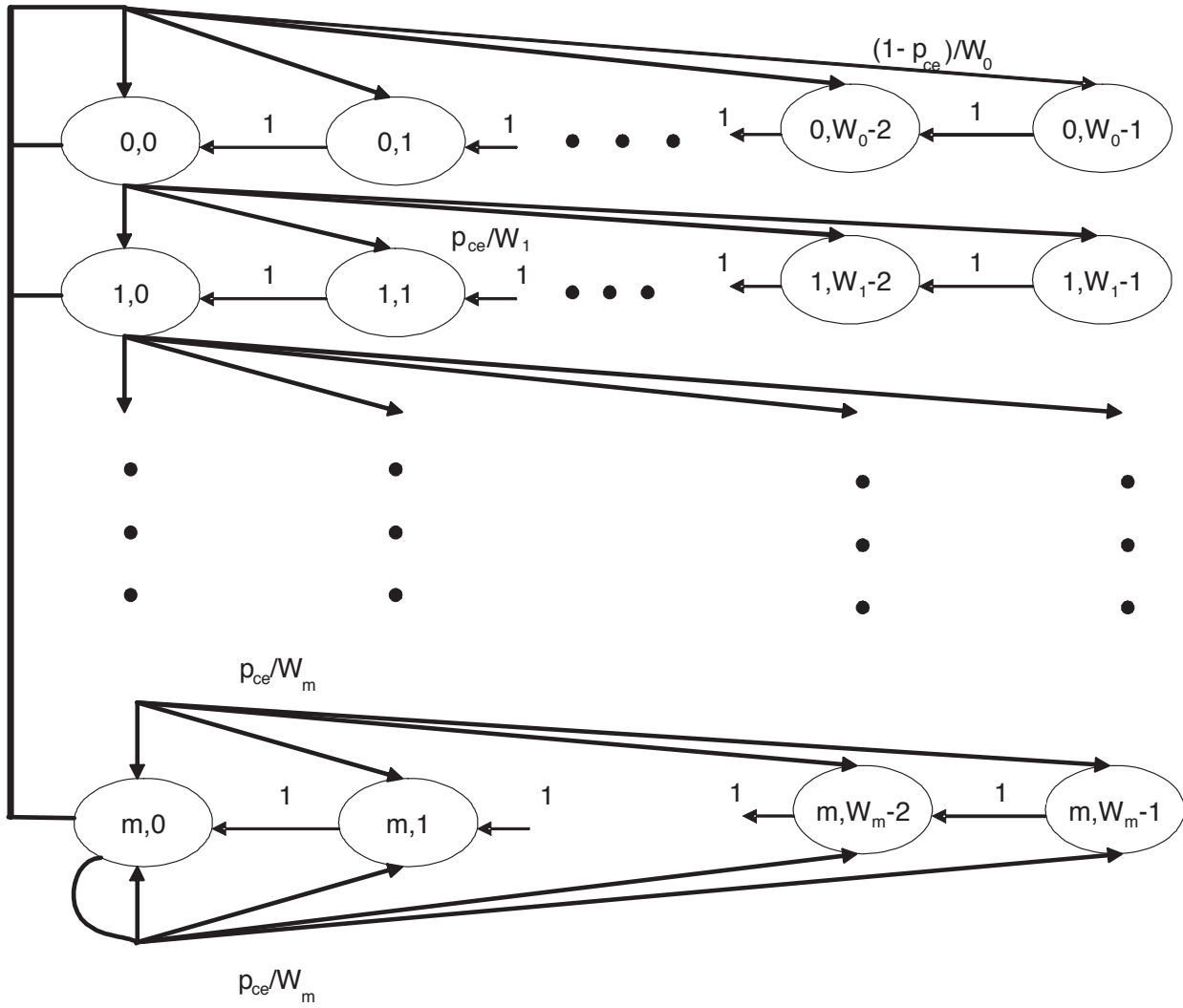
Fig. 3.   Markov chain for backoff counter and contention window stage.
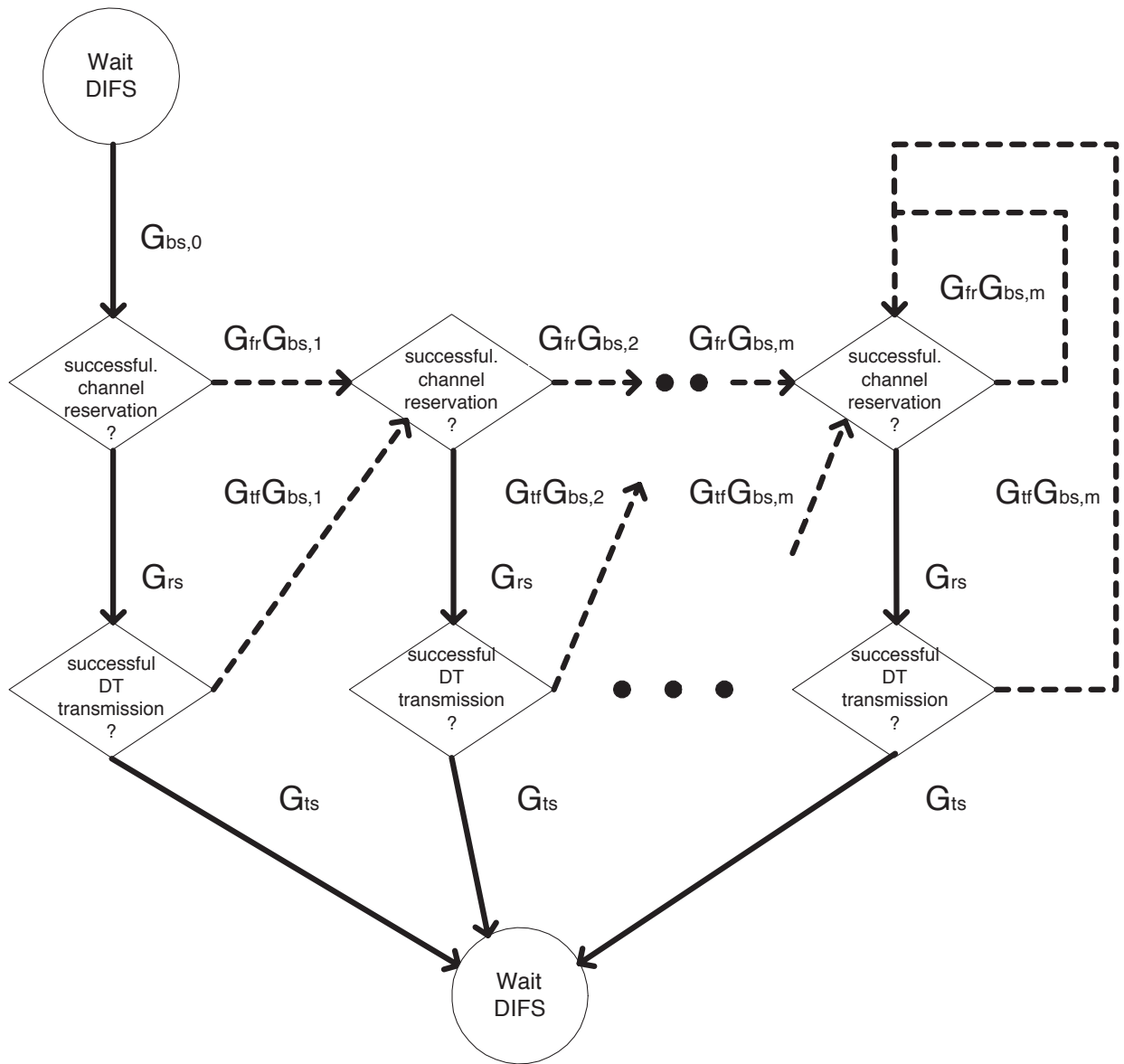
Fig. 4.   State diagram representation of the 802.11 MAC protocol. Transform variables X and Y are omitted for simplicity.
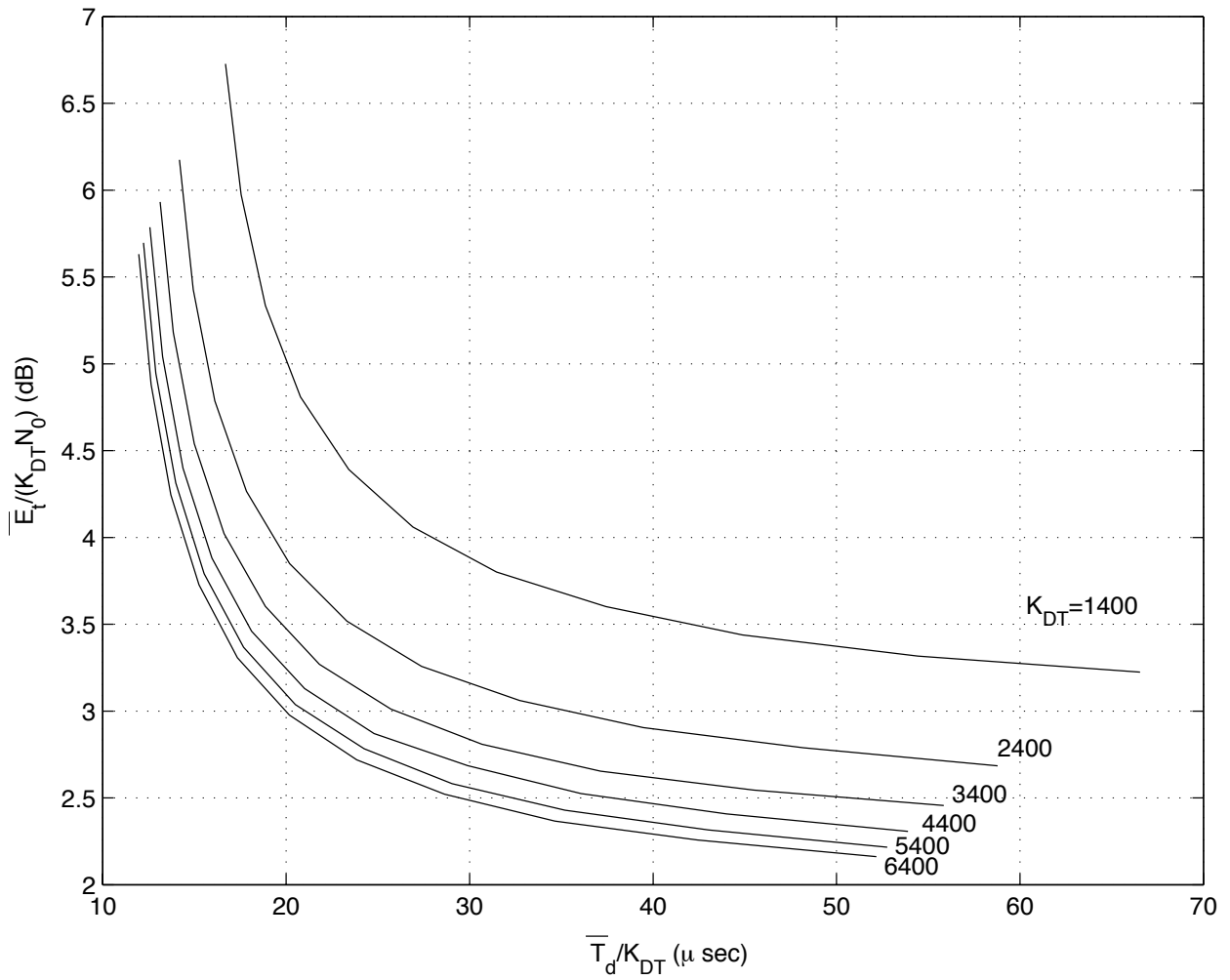
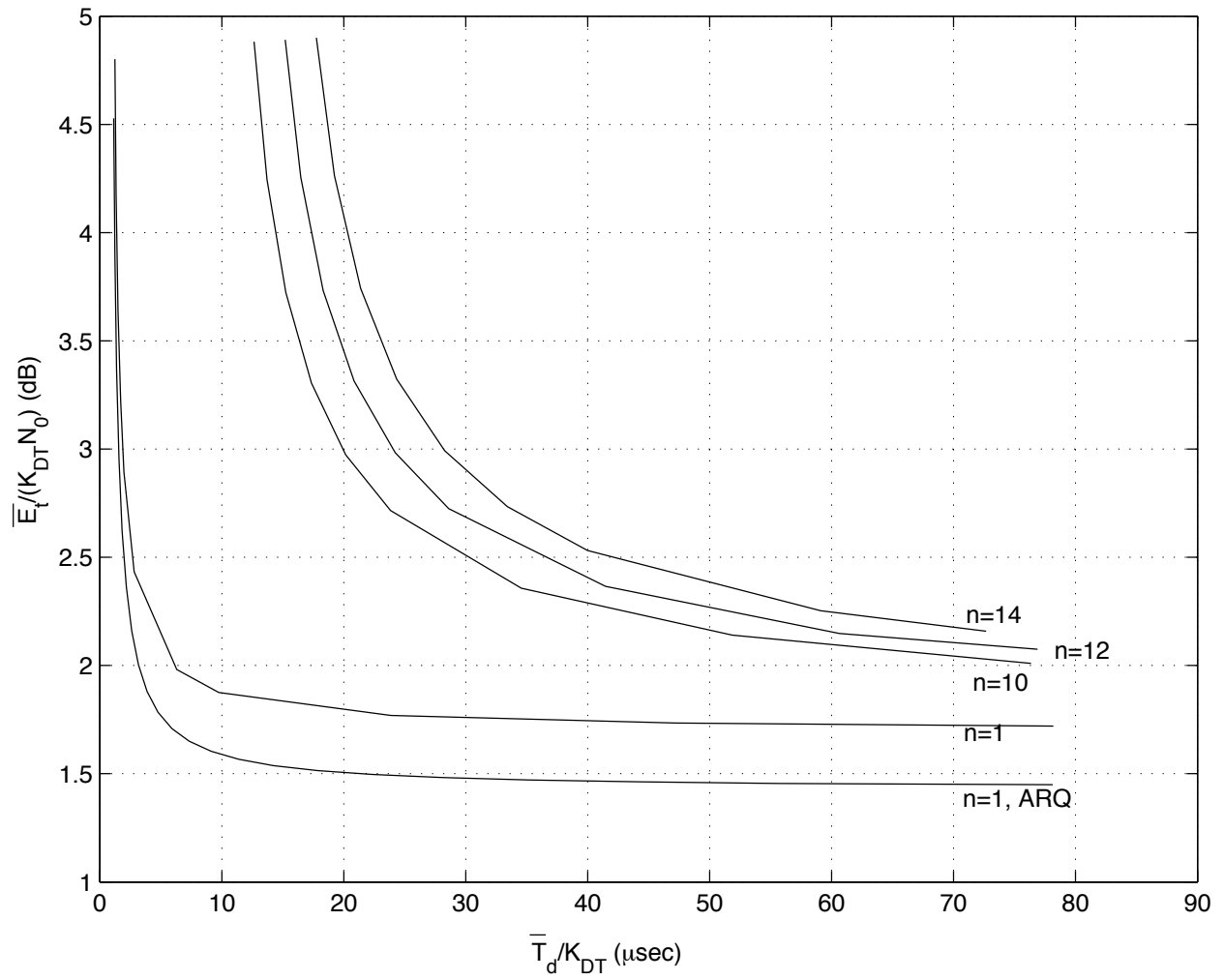Fig. 5. Energy-delay curves for different K_{DT} with $n = 10$.

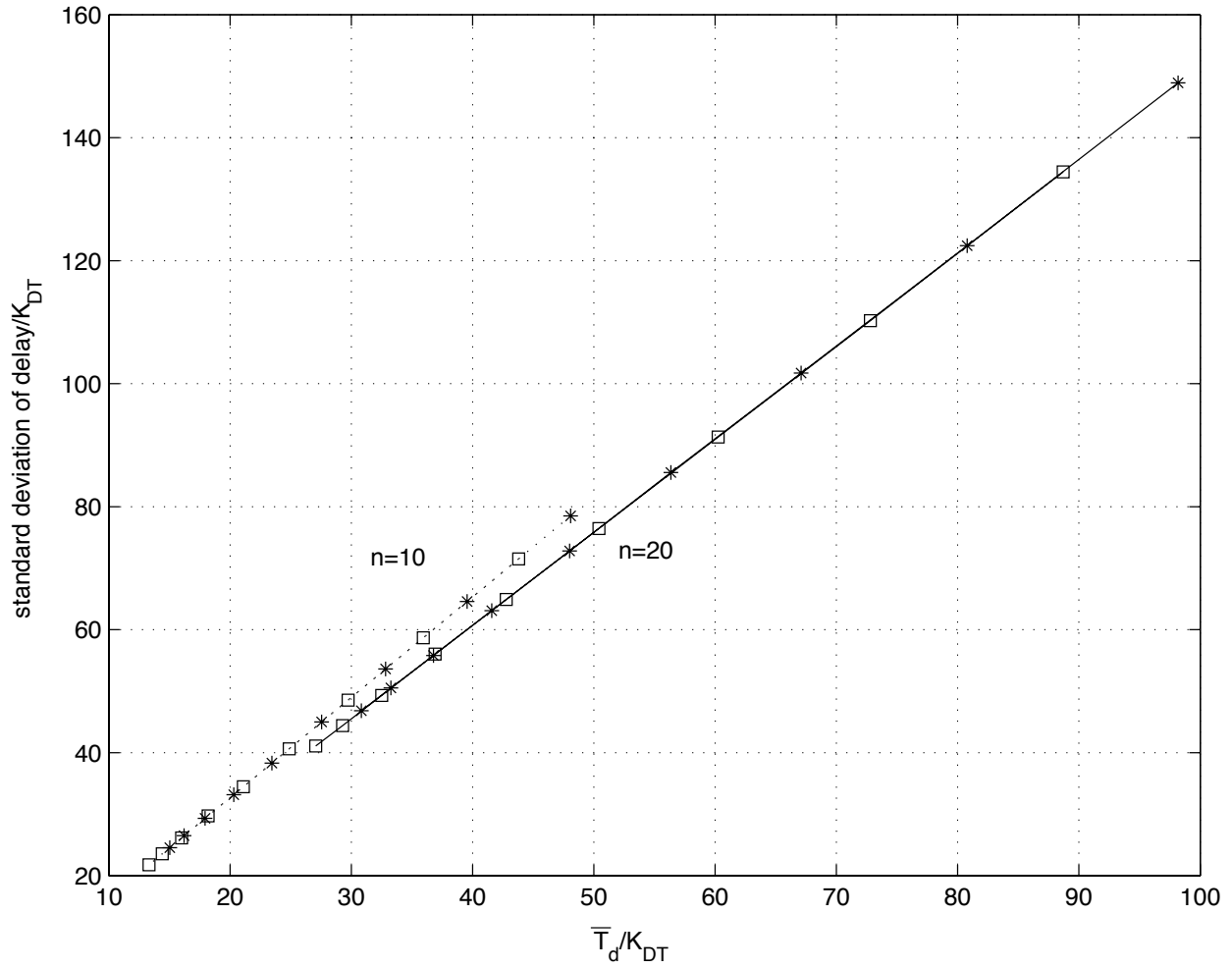Fig. 6.   Energy-delay curves for different number of users n with $K_{DT} = 6400$.

Fig. 7. Normalized standard deviation of delay vs. average delay curves for different n and K_{DT}. Lines with a square symbol represent the case of $K_{DT} = 4400$ and lines with a star symbol represent the case of $K_{DT} = 2400$.
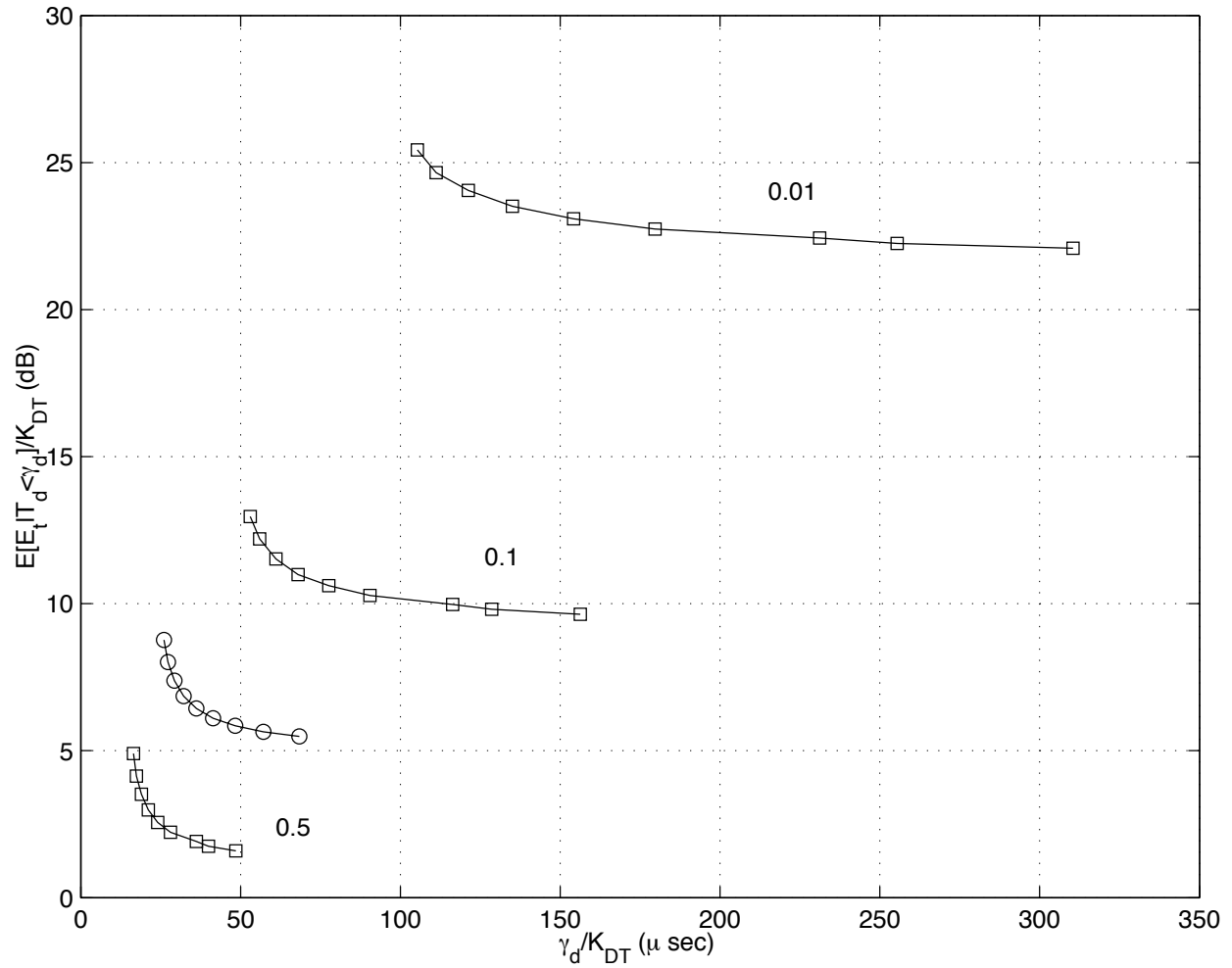
Fig. 8. Energy-delay curves for various outage delay probabilities and different values for the outage probability $Pr(T_d > \gamma_d)$ ($m = 1, n = 10, W = 8, K_{DT} = 640$). The average energy and delay curve (circle symbol) is also shown for comparison.
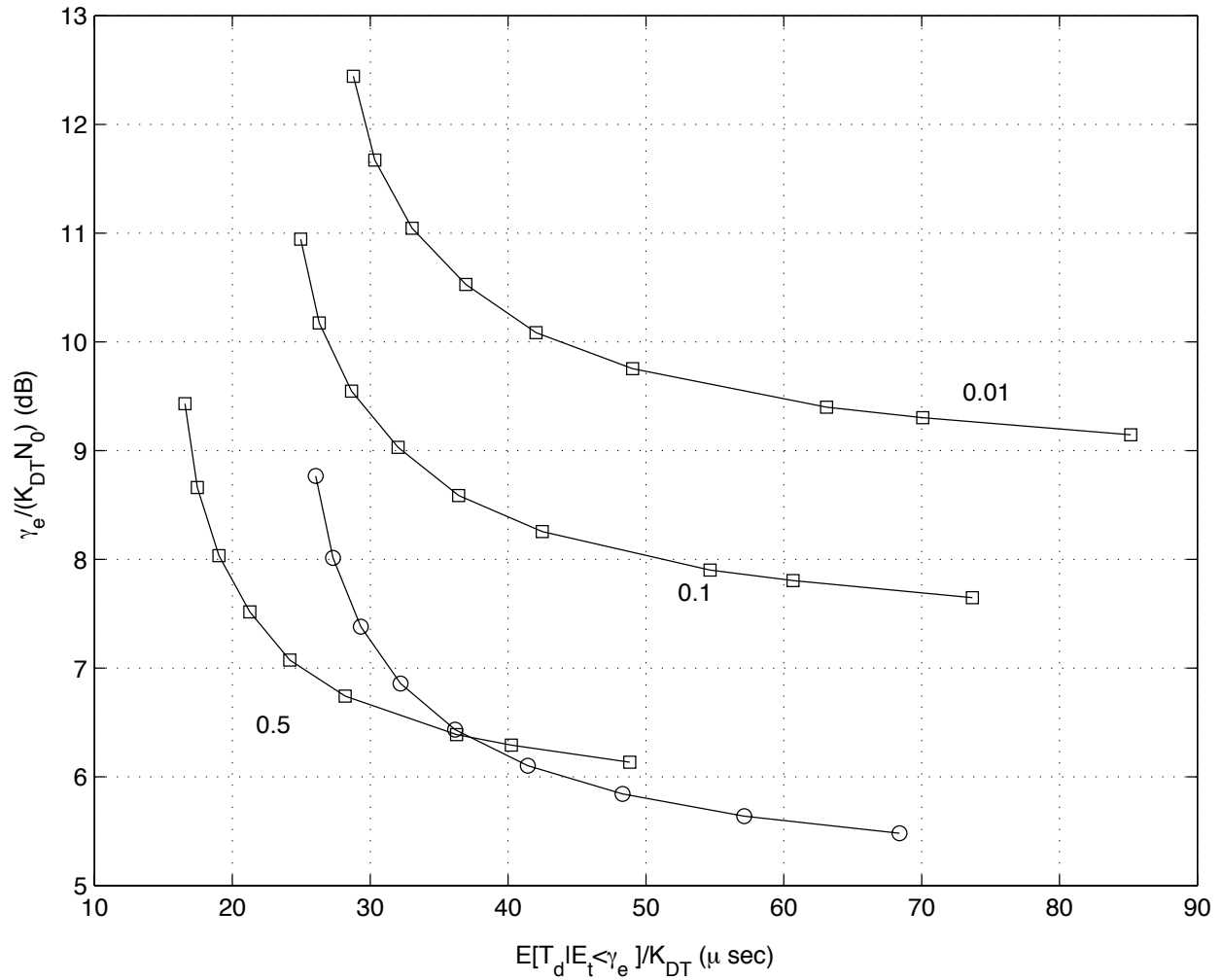
Fig. 9. Energy-delay curves for various outage energy probabilities and different values for the outage probability $Pr(E_t > \gamma_e)$ ($m = 1, n = 10, W = 8, K_{DT} = 640$). The average energy and delay curve (circle symbol) is also shown for comparison.
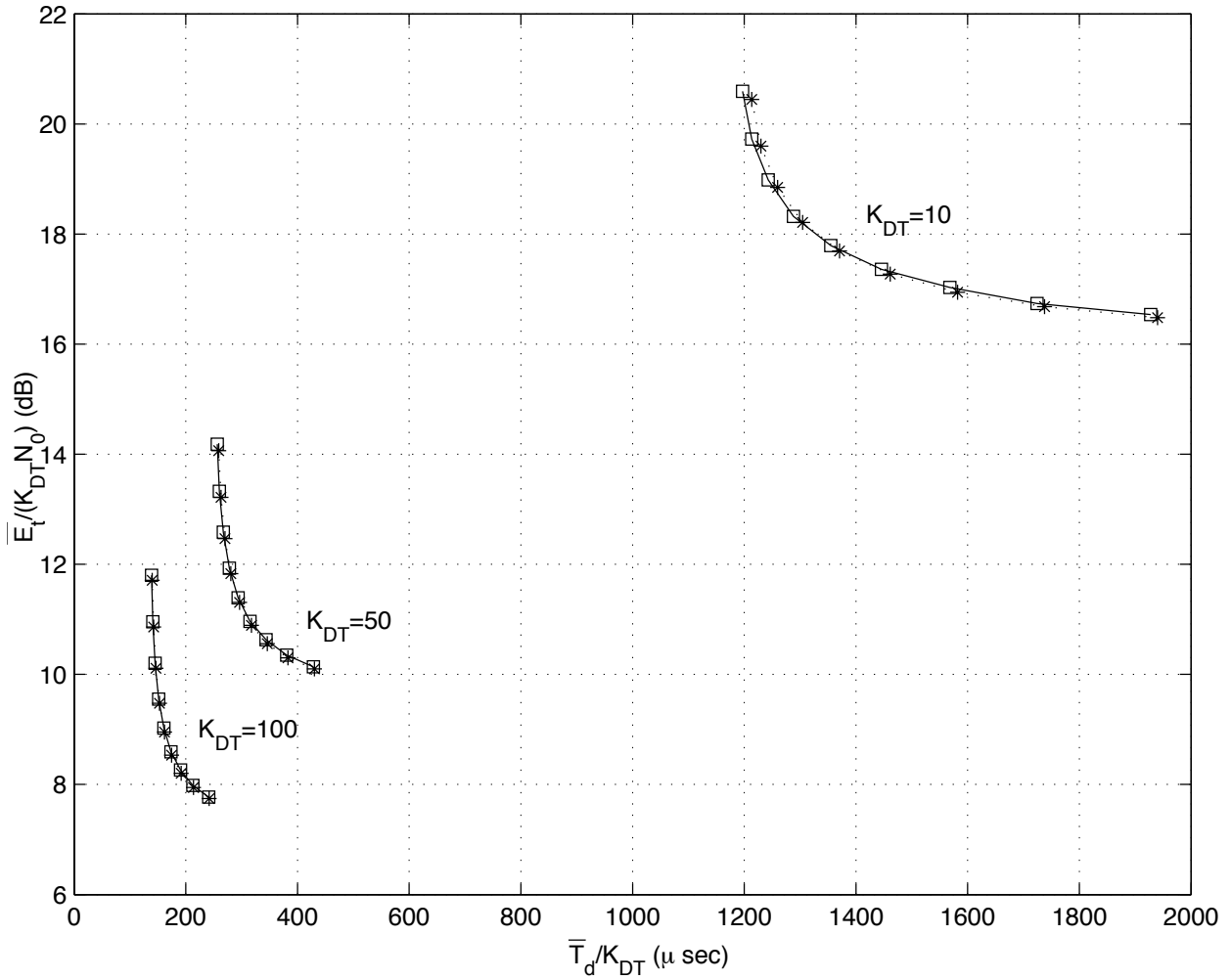
Fig. 10. Energy-delay tradeoff curves evaluated from the approximation method (star symbol) and the exact energy-delay tradeoff curves (square symbol) with $n = 10$