Performance Analysis of the IEEE 802.11 Power Saving Mode

Rong Zheng, Jennifer C. Hou and Lui Sha
Dept. of Department of Computer Science
University of Illinois at Urbana-Champaign
Urbana, IL 61801
E-mail: \{zheng4,jhou,lrs\}@cs.uiuc.edu

Abstract—Turning the devices to low-power states when they are not in use is an important technique to conserve energy for battery-powered wireless devices. The IEEE 802.11 power saving mode is one of the most popular techniques in wireless LAN and multi-hop wireless networks to coordinate the power states of communication devices. In this paper, we present an analytical characterization of energy consumption, delay and loss rate of the IEEE 802.11 Power Management Mode (PSM) as a function of the traffic load, buffer size and other protocol-specific parameters. Contrary to general beliefs, we observe from the analytical study that the IEEE 802.11 PSM is not very power efficient for light traffic load. Unlike traditional steady-state queuing analysis, the novelty of our approach lies in the use of transient analysis techniques where time-dependent behaviors can be modeled accurately.

I. INTRODUCTION

With the proliferation of portable computing platforms and small wireless devices, wireless networks have received more and more attention as a means of data communication among devices regardless of their physical locations. As wireless devices usually rely on portable power sources such as batteries and the energy cost of devices in idle states is very significant [6], it is in general desirable to turn the radio off when it is not in use, called power management.

The IEEE 802.11 MAC [7] is one of the most widely used medium access protocols in wireless LANs. In addition to services that support data delivery between stations and control LAN access as well as confidentiality, the IEEE 802.11 also specifies power management functions in both the infrastructure network with base station support (BSS) and the independent basic service set (IBSS, a.k.a ad hoc mode). In particular, in the IEEE 802.11 power saving mode (PSM), a node can switch to the low-power state to save power when the interface is idle. Recently, IEEE 802.11 PSM has also been applied in mobile ad hoc networks (MANETs) to coordinate power management states in multi-hop communications[15], [5]. It is shown in [15] that approximately 50% of energy saving can be achieved using the IEEE 802.11 PSM in conjunction with time-out driven policy in multi-hop wireless networks.

Several analytical models for the capacity of IEEE 802.11 MAC protocol have been proposed in [3], [4]. All the theoretical results are derived for IEEE 802.11 MAC operating in active mode, thus the effect of power management is not considered. We present in this paper an analytical characterization of IEEE 802.11 Power Management Mode (PSM) in IBSS. In particular, we are interested in investigating various energy-performance metrics, such as energy consumption, delay and loss rate, as a function of traffic load, buffer size and other protocol-specific parameters.

We believe such a quantitative study can not only provide an evaluation tool for the IEEE 802.11 PSM but also give insights on improving power management designs. For example, we use this model to verify the ns2 implementation for the IEEE 802.11 PSM [12] and find several bugs and implementation deficiency. To demonstrate the extensibility of our model, we further apply our model to analyze systems with multiple homogeneous receivers and systems where the mobile stations can configure different active and sleeping intervals. Compared with traditional steady-state queuing analysis, the novelty of our approach lies in the employment of transient analysis techniques and hence accurately modeling the time-dependent behaviors.

Contrary to general beliefs, we observe from the analytical study that the IEEE 802.11 PSM is not very power efficient for light traffic load. As indicated by the analytical results, it keeps the device active around 50% of the time for 10% traffic load and 18% for traffic load around 1%. This is mainly due to the periodical nature of the IEEE 802.11 PSM. Similar results can be obtained for systems with multiple competing users. This justifies the need for adaptive power saving solutions that operate at finer control granularity.

The rest of the paper is organized as follows. We first present in Section II a succinct overview of the IEEE 802.11 PSM that pertains to the problem considered in this paper. In Section III, we introduce a theoretical model that accounts for the periodical structure of the IEEE 802.11 PSM. The quantitative results for energy consumption, delay and loss rate under the IEEE 802.11 PSM is presented in Section IV. Two extensions of the model are presented in the Section V. Finally, we conclude the paper in Section VI and Section VII with a list of related work and future research avenues.
II. OVERVIEW OF THE IEEE 802.11 POWER SAVING MODE

In the IEEE 802.11 PSM, a node can be in one of two different power modes, i.e., active mode when a node can receive frames at any time and power-save mode (PS) when a node is mainly in low-power state and transits to full powered state subject to the rules described next. The low-power state usually consumes at least an order of magnitude less power than in the active state[6].

In the power-save mode, all nodes in the network are synchronized to wake up periodically to listen to beacon messages. Broadcast/multicast messages or unicast messages to a power-saving node are first buffered at the transmitter and announced during the period when all nodes are awake. The announcement is made via an ad hoc traffic indication message (ATIM) inside a small interval at the beginning of the beacon interval called the ATIM window. If a node receives a directed ATIM frame in the ATIM window (i.e. it is the designated receiver), it sends an acknowledgment and stays awake for the entire beacon interval waiting for data packets to be transmitted. Immediately after the ATIM window, a node can transmit buffered broadcast/multicast frames, data packets and management frames addressed to nodes that are known to be active (by reception of acknowledgment to ATIM frames). Otherwise, the node can switch to the low-power state to conserve energy. In IEEE 802.11, a node’s power management mode is indicated in the frame control field of the MAC header for each packet. The behavior of the IEEE 802.11 PSM is illustrated in Figure 1.

In the IEEE 802.11 PSM, the length of a beacon interval and the size of an ATIM window is configured by the first node that initiates the network in IBSS. A mobile station can choose to wake up every multiples of the beacon intervals for further energy saving.

For the rest of the paper, we use “active interval” to refer to a beacon interval in the IEEE 802.11 PSM when the device is full-powered and “low-power interval” for a beacon interval that a device turns to low-power state after the ATIM window.

III. PERFORMANCE ANALYSIS

In this section, we present an analytical model for the IEEE 802.11 PSM using queuing analysis.

A. System Model

We first consider a pair of transmitter and receiver operating in the IEEE 802.11 distributed coordinating function (DCF) mode. The two nodes are synchronized using beacon messages. Therefore, they can coordinate in the transition of power management states. In Section V, we extend the model to the case of multiple homogeneous competing flows. For the ease of derivation, we assume that packets arrive at the transmitter in compliance with a Poisson arrival process. However, the analysis can be readily extended to other Markov regenerative processes such as Batch Markovian arrival process (BMAP).

We examine the state of the system at the boundary of beacon intervals. A packet that is under service at the boundary of a beacon interval would restart its service at the next interval. This assumption holds true if the service time distribution (or equivalently the packet size distribution) is exponential (due to the memoryless property of the exponential distribution). For non-exponential service times, we claim that as long as the service time is an order of magnitude smaller than the beacon interval, ignoring the packet currently under service would not have a significant impact on the mean value analysis. This claim is verified by the simulation results in Section V.

Packets to a receiver in the low-power state are buffered at the transmitter if the buffer space at the transmitter is sufficient; otherwise, they are subject to drops. At the beginning of each beacon interval, the transmitter examines its buffer. If it has packets for the receiver, the sender informs the receiver to stay active in the next interval of length $b$. Otherwise, both switch to the low-power state for an interval of length $b$. Therefore, the transition of power management states at the sender and receiver is synchronized.

We model the system with a finite buffer size $K$ as a finite state Markov chain sampled at the boundary of beacon intervals\(^1\). The system state at the end of the $i$th beacon interval is given by the number of packets buffered $n_i$. Let $\zeta_i \in \{0, 1\}$ be the power management state of the $i$th beacon interval, with “0” indicating the low-power state and “1” the active state. The power management decision of the IEEE 802.11 PSM can be simply expressed as

1) if $n_i = 0$ then $\zeta_{i+1} = 0$;

2) if $n_i > 0$ then $\zeta_{i+1} = 1$.

Let $A_i$ and $D_i$ denote the number of arrivals and departures in the $i$th interval. Then the transition of $n_i$ is given by

$$n_{i+1} = \begin{cases} \min\{n_i + A_{i+1} - D_{i+1}, K\}, & \text{if } n_i \geq 1, \\ A_{i+1}, & \text{if } n_i = 0. \end{cases}$$

\(^1\)In real implementation, buffer size is usually expressed in $B$ bytes. Therefore, $K$ can be approximated as $K = B/\bar{S}$, where $\bar{S}$ is the average packet size.
Transient analysis for non-exponential service time can be quite complicated. However, in the context of IEEE 802.11 PSM, the time-dependent behavior is of importance since the system cannot reach a steady state within an active interval from medium to high traffic loads.

Let \( N(t) \) be the number of packets in the system at time \( t \), and \( \mathbb{P}(\lambda, t) \) the time-dependent transient transition matrix within an (active/low-power) beacon interval for normalized load \( \lambda \). For ease of presentation, we drop \( \lambda \) in \( \mathbb{P}(\lambda, t) \) and simply write it as \( \mathbb{P}(t) \) unless clarification is needed. In particular, \( \mathbb{P}_{0,m}(t) = \mathbb{P}(N(t) = m|n_i = 0), m = 0, 1, \ldots, K \) denotes the probability that there are \( m \) packets in the system at time \( t \) in a low-power beacon interval. \( \mathbb{P}_{l,m}(t) = \mathbb{P}(N(t) = m|n_i = l), l = 1, 2, \ldots, K \) and \( m = 0, 1, \ldots, K \) gives the probability that there are \( m \) packets in the system at time \( t \) in an active beacon interval \( i \) starting with \( l \) packets. Lastly, \( \mathbb{P} = \mathbb{P}(b) \) denotes the one-step transition matrix for beacon interval of length \( b \).

Logothetis et al. [11] developed a computational technique for obtaining the time-dependent solution of the queue length distribution for a class for Markov regenerative process including M/G/1/K and GI/M/1/K queues. In the case that the service time is exponential, the transition matrix can be written simply as an exponential matrix. In the case of non-exponential service times, one has to resort to renewal theory on Markov regenerative process (MRGP) [10]. We adopt this computation technique in the following derivation and obtain results for both exponential and deterministic service times.

### Exponential Service Time:
For exponential service times, the one-step transition matrix can be written as,

\[
\mathbb{P}_{l,m}(t) = \begin{cases} \frac{e^{(l+1)\lambda t}}{(l+1)!} e^{-\lambda t}, & \text{if } l \geq 1 \\ \frac{1}{m!} e^{-\lambda t}, & \text{if } l = 0, m < K \\ \sum_{j=K}^{\infty} \frac{(\lambda t)^j}{j!} e^{-\lambda t}, & \text{if } l = 0, m = K, \end{cases}
\]  

(1)

where \( Q \) is the generation matrix of the corresponding M/M/1/K queuing system, i.e.,

\[
Q = \begin{bmatrix} -\lambda & \lambda & \mu & \mu -\lambda & \lambda & \mu & \mu -\lambda & \cdots & \cdots \end{bmatrix}.
\]

### Deterministic Service Time:
Transient analysis for non-exponential service time can be quite complicated. However, for deterministic service time, the derivation can be greatly simplified. Let \( \tau \) be the fixed service time of packets. In a low-power interval, due to Poisson arrival assumption, the transition matrix can be trivially written as,

\[
\mathbb{P}_{l,m}(t) = \begin{cases} \frac{(\lambda t)^m}{m!} e^{-\lambda t}, & \text{if } m < K \\ \sum_{j=K}^{\infty} \frac{(\lambda t)^j}{j!} e^{-\lambda t}, & \text{if } m = K. \end{cases}
\]  

(2)

In an active interval, to construct the embedded Markov renewal sequence \{\( N(T_n^+) \), \( T_n \)\}, we define the regeneration points \( T_n \)'s at the service completion instants for an non-empty system and at arrival instants for an empty system (the latter is chosen to simplify the time-domain equations for the deterministic service times). That is,

1. If \( N(T_n^+) = 0 \), define \( T_{n+1} \) as the time instant at which the next arrival occurs.
2. If \( N(T_n^+) \neq 0 \), define \( T_{n+1} \) as the time instant of the next service completion.

Therefore, \( \{N(t), t \geq 0\} \) is an Markov regenerative process (MRGP) with global kernel [10].

\[
K(t) = \begin{bmatrix} 0 & A_0(1,t) & \cdots & A_0(K-2,t) & \cdots & \cdots & \cdots & 0 \\ A_0(1,t) & 0 & \cdots & 0 & \cdots & \cdots & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & \cdots & \cdots & 0 \\ \end{bmatrix}
\]

where \( K_{i,j}(t) = \mathbb{P}(N(T_n^+) = j | N(T_{n-1}^+) = i) \) characterizes the evolution of the MRGP at the regenerative points. \( A_0(l,t) \) and \( S_0(i,t) \) are given by,

\[
A_0(l,t) = \begin{cases} 1 - e^{-\lambda t} & l = 1 \\ 0 & \text{otherwise} \end{cases}
\]

and

\[
S_0(i,t) = \frac{(\lambda \tau)^i e^{-\lambda \tau}}{i!} u(t - \tau).
\]

The local kernel [10] of the MRGP is given by the matrix \( E(t) \), where \( E_{i,j} = \mathbb{P}(N(t) = j, T_1 > t | N(T_0^+) = i) \).

\[
E(t) = \begin{bmatrix} A_0(0,t) & 0 & \cdots & 0 \\ 0 & S_0(0,t) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & S_0(0,t) \\ \end{bmatrix}
\]

(3)

From renewal theory[10], we have

\[
\mathbb{P}_{l,j}(t) = E_{i,j}(t) + \sum_{k \in \Omega} \int_0^t dK_{i,k}(u) \mathbb{P}_{k,j}(t-u).
\]

(3)

It is easy to see that, for \( 0 \leq t \leq \tau \), we have,

\[
\mathbb{P}_{l,j}(t) = \begin{cases} e^{-\lambda (\tau + l - j)} & j \geq t \\ 0 & j < t \end{cases}
\]

(4)

For \( t > \tau \), let \( p_l(t)k_l \) and \( q_l(t) \), \( l = 0, 1, 2, \ldots, K \) be the \( l \)th row of matrices \( \mathbb{P}(t) \), \( K(\infty) \) and \( E(t) \). By plugging in \( K(t) \) and \( E(t) \), Eq. (3) can be rewritten as,

\[
p_l(t) = \begin{cases} k_0(T_{n-1}^+ - t) & 0 \leq l \leq K, t \leq \tau \\ 0 & l = 0, t \geq \tau \end{cases}
\]

Using the backward Euler method, we have,

\[
p_l(t) = \begin{cases} k_0(T_{n-1}^+ - t) & 0 \leq l \leq K, t \leq \tau \\ 0 & l = 0, t \geq \tau \end{cases}
\]
where $\Delta t$ is the discretization step.

Therefore, combining Eq. (2) and Eq. (4), we have

\[
\mathbb{P}_{1,m} = \begin{cases} 
\frac{\varphi_{1,m}(b)}{(\lambda b)^m} e^{-\lambda b}, & \text{if } l \geq 1, \\
\sum_{j=0}^{\infty} \frac{(\lambda b)^j}{j!} e^{-\lambda b}, & \text{if } l = 0, m < K, \\
\sum_{j=K}^{\infty} \frac{(\lambda b)^j}{j!} e^{-\lambda b}, & \text{if } l = 0, m = K,
\end{cases}
\]

as the one-step transition matrix for deterministic service time where $\mathbb{P}_{1,m}(b)$ is the $m$th element of $p_1(b)$.

### C. Steady State Energy-Performance Metrics

In this section, we derive the steady state performance metrics of interest using the previous result on the transition matrix $\mathbb{P}$ and $\mathbb{P}(t)$.

**Probability of the number of packets at the end of each beacon interval:** Let $\pi_k = Pr\{n_i = k\}$ be the probability that there are $k$ packets at the end of the $i$th beacon interval. By solving the Chapman-Kolmogorov equation $\pi = \pi \mathbb{P}$, we can get the steady state distribution (i.e., the distribution of the number of packets) at the end of each beacon interval, where $\mathbb{P}$ is given in Eq. (1) and Eq. (5) for exponential and deterministic service time.

**Blocking probability:** Arrivals to a low-power system have to be buffered till the end of the current beacon interval. Therefore, the blocking probability in low-power system is, $Pr\{B[\zeta = 0]\} = 1 - \sum_{j=0}^{K} \frac{(\lambda b)^j}{j!} e^{-\lambda b}$. On the other hand, the probability that a packet arrives at an active interval starting with $j$ packets and sees $K$ packets in queue is given by $Pr\{B[\zeta = 1, n_i = j]\} = \int_0^b \pi_j(t) dt$, where Poisson arrivals can be thought of a random point process on the time axis. Therefore, the total blocking probability is

\[
P_B = Pr_{asleep} Pr\{B[\zeta = 0]\} + \sum_{j=1}^{K} \frac{\pi_j}{1 - \pi_0} Pr_{awake} \cdot Pr\{B[\zeta = 1, n_i = j]\},
\]

where $Pr_{asleep} = \pi_0$ and $Pr_{awake} = 1 - Pr_{asleep}$ are the probabilities that a packet arrives at a sleeping server and an active server respectively. Consequently, the throughput of the system is $\gamma = \lambda(1 - P_B)$. The carried load is defined as $\rho = \rho(1 - P_B)$.

**Average power:** Let $PW_{tx/rx}$, $PW_{awake}$, $PW_{idle}$ be the power consumed in the transmission/receiving, awake and idle states, respectively. The average power at the sender and receiver can be computed as

\[
P_W^{tx/rx} = Pr_{asleep} PW_{asleep} + Pr_{awake} PW_{tx/rx}
\]

The first term corresponds to the power consumed when the node is put to the low-power state, while the second term gives the power consumed in transmitting or receiving packets. $Pr_{awake} - \rho_c$ gives the idle probability

Now we account for the effect of the length, $\delta$, of an ATIM window (recall that each beacon interval begins with an ATIM window, where the transmitter and receiver coordinates their power management state). The average power can be approximated as

\[
P_W^{tx/rx} = Pr_{asleep}(1 - \delta/b)PW_{asleep} + \rho_c PW_{tx/rx}
+ (Pr_{awake} - \rho_c)PW_{idle}.
\]

The ATIM window in essence reduces the length of a low-power period by $\delta$.

**Average delay:** We consider the FIFO service discipline, i.e., the packets are serviced in the order of their arrivals. If a packet arrives at a low-power system, it has to wait till the end of the beacon interval and the service time of all the packets ahead of it, i.e., $\overline{d}_{asleep} = \frac{1}{b} \int_0^b (\lambda X + (b - t)) dt$. On the other hand, if a packet arrives at an active system, its waiting time is that in a system starting with $i$ packets at the end of the last beacon interval, i.e., $\overline{d}_{awake} = \int_0^b (\pi_i(t) X + \sum_{j=1}^{K-1} \pi_i,j(t) (X + \sum_m \overline{d}_m)) dt$. Therefore, the average delay is given by

\[
\overline{d} = Pr_{asleep} \overline{d}_{asleep} + Pr_{awake} \sum_{i=1}^{K} \frac{\pi_i}{1 - \pi_0} \overline{d}_{awake}.
\]

### IV. Numerical Examples

In this section, we present the numerical results for the IEEE 802.11 under exponential service time to study the impact of the beacon intervals on the energy-efficiency and other performance metrics. We assume the ATIM window takes a fixed proportion of the beacon interval lengths. The service time distribution is deterministic and normalized to one.

Fig. 2 shows the energy consumption of the IEEE 802.11 PSM and energy breakdown into idle states, sleep states and active communication, under different traffic loads and beacon interval values. The energy consumption for the idle, sleep and tx/rx states are $0.83W$, $0.13W$ and $1W$, respectively. The buffer size $K$ is set to 50. From Fig. 2, changing the size of the beacon interval does not have a significant impact on energy consumption. Furthermore, we observe that IEEE 802.11 PSM is not energy efficient under light traffic loads. For example, when the traffic load is as low as 0.1, both the sender and receiver need to awake around 50% of the time on average.

Fig. 3 shows the average delay with respect to the beacon interval length. The average delay consists of three terms, i.e., i) queuing delay, ii) transmission delay and iii) the delay to wake up the interface to the active state if necessary. Not surprisingly, as the beacon interval increases, the delay gets larger since the last term dominates. The decrease in delay under heavy load scenarios is due to packet losses (and thus a smaller queuing delay).

The results for deterministic service times are similar and omitted due to space limits. Interested readers can refer to [14] for more details.

### V. Extension of the Model

In this section, we consider two variations of the basic model proposed in Section III. In the first variation, the lengths of the active interval and the low-power interval are...
set differently. This is feasible in practice as a mobile station can choose to wakeup every multiple of the beacon intervals. The second variation extends the model to the case with multiple contending homogeneous flows. Though simplistic in the assumptions made, the study sheds light on the impact of wireless resource contention on the energy-efficiency.

A. Different Active and Low-Power Intervals

We investigate the benefit of having one additional control knob, i.e., varying the active and low-power interval length. Intuitively, if the low-power interval length is long, the average delay incurred by a packet is longer. On the other hand, since the active interval has higher utilization, the energy efficiency is increased. Therefore, it is interesting to see how different combinations of active and low-power intervals affect the energy-performance trade-off.

Let $b_{on}$ be the length of active interval, $b_{off}$ be the length of low-power interval. Compared to the model presented in Section III, the main difference lies in the transition matrix. Specifically, for exponential service, Eq. (1) is modified to be,

$$P_{l,m} = \begin{cases} \frac{\lambda^m}{m!} e^{-\lambda}, & \text{if } l = 0, m < K, \\ \frac{\lambda^m}{m!} \prod_{j=1}^{m} \sum_{k=m}^{j} \frac{1}{k!} e^{-\lambda}, & \text{if } l = 1, m < K, \\ \frac{\lambda^m}{m!} \prod_{j=1}^{m} \sum_{k=m}^{j} \frac{1}{k!} e^{-\lambda}, & \text{if } l = 0, m = K, \\ \frac{\lambda^m}{m!} \prod_{j=1}^{m} \sum_{k=m}^{j} \frac{1}{k!} e^{-\lambda}, & \text{if } l = 0, m = K, \\ \end{cases}$$

And for deterministic service, Eq. (5) is modified to be,

$$P_{l,m} = \begin{cases} \frac{\lambda^m}{m!} e^{-\lambda}, & \text{if } l = 0, m < K, \\ \frac{\lambda^m}{m!} \prod_{j=1}^{m} \sum_{k=m}^{j} \frac{1}{k!} e^{-\lambda}, & \text{if } l = 1, m < K, \\ \frac{\lambda^m}{m!} \prod_{j=1}^{m} \sum_{k=m}^{j} \frac{1}{k!} e^{-\lambda}, & \text{if } l = 0, m = K, \\ \end{cases}$$

Furthermore, the probability that a packet arrives at a sleeping server is given by $P_{\text{asleep}} = \pi_0 b_{off} / (b_{on} + b_{off})$.

The energy-performance trade-off for different combinations of active and asleep beacon intervals is shown in Figure 4. In Figure 4, each data point corresponds to one traffic load in the ascending order from bottom to top. With $b_{off} = 100ms$, $b_{on} = 50ms$, the highest energy efficiency can be achieved at the expense of larger latency. In comparison, $b_{off} = 50ms$, $b_{on} = 100ms$ minimize the delay for all settings.

B. Multiple Homogeneous Contending Flows

The analysis of multiple contending flows is complicated by the fact that each transmitter’s power management state transition is not independent from one another due to resource contention. For example, a light receiver may be backlogged behind a heavy one and thus “forced” to remain active. Therefore, the number of active nodes at the end of a beacon interval is decided by the service discipline as well as the composition of the backlogged flows. To facilitate the analysis, we make the following simplifying assumptions, i) one flow per node (therefore, we use “node” and “flow” interchangeably), ii) the flows are homogeneous, i.e., with the same traffic load, iii) the service is work-conservative as long as there are active flows in the system, and iv) backlogged active flows are serviced alternately. Note that throughput analysis for CSMA/CA types of MAC protocols for multiple contending flows is a hard problem by itself. In this paper, we make no attempt to quantify the effect of CSMA/CA contention. Instead, we focus on the energy perspective of the IEEE 802.11 PSM.

Let $M$ be the total number of contending flows and $\lambda / M$ be the rate of a single flow. The state at the end of each beacon interval is given by a tuple $(m, k)$, where $m = 0, 1, \ldots, M$ is the number of active flows in the next interval and $k = 0, 1, \ldots, K$ is the total number of buffered messages in the system. A node is active if and only if it has backlogged messages at the end of a beacon interval.

To derive the transition matrix, consider an beacon interval starting with $k$ packets and $m$ active flows. At the end of this beacon interval, the backlogged packets consist of two
parts, i) arrivals from \( M - m \) inactive flows and ii) residual packets from active nodes. Furthermore, the number of active nodes for the next interval is the sum of the newly backlogged flows and active flows which still have remaining packets. The second term is decided by the service/queueing discipline and composition of backlogged flows at the beginning of the beacon interval. Under the assumption that all active flows have equal probability to be backlogged, the conditional probability \( P\{m|k,M\} \) of having \( m \) out of \( M \) different flows given a total \( k \) packets backlogged can be approximated by the following counting arguments.

The number of ways to select \( k \) objects from \( m \) types with repetition allowed, or the number of solutions to \( \sum_{i=1}^{M} x_i = k \) in nonnegative integers is given by \( \binom{k+M-1}{M-1} \). In addition, the number of ways to decompose \( k \) to \( m \) types (as the sum of \( m \) positive integer) is \( \binom{k-1}{m-1} \). Therefore,

\[
P\{m|k,M\} \approx \begin{cases} \binom{M}{m} \binom{k-1}{m-1}, & m \leq k \\ 0, & \text{else} \end{cases} \tag{9}
\]

The transition probability from state \( (m,k) \) to \( (m',k') \) is given by Eq. (10), where \( \lambda_{b} = (M-m)\lambda \) and \( \lambda_{nb} = m\lambda \) are the aggregate loads to the inactive flows and active flow respectively. \( P_{k,k'-1}(\lambda_{nb},b) \) is the transition probability for a single receiver at rate \( \lambda_{nb} \) derived in Section III. The third line in Eq. (10) is due to Bayes Theorem, which sums up all possible combinations of contribution from active flows and inactive flows in the current interval. For finite buffer, the above equation needs to be truncated by boundary conditions. Note that Eq. (10) is exact for the case of a single pair of sender-receiver.

Figure 5 demonstrates the impact of the number of competing flows on the percentage of active beacon intervals. Both active and inactive intervals are set to 100ms. The x-axis gives the aggregated rate of multiple flows with each flow transmitting at the same rate. The model is validated using packet-level ns-2 [13] simulations. In the simulation, multiple flows share a bottleneck of 2Mb. The packet size is 1000 bytes. The active flows are serviced alternately to emulate the wireless channel contention. Power management for each flow follows the same rule as in the IEEE 802.11 PSM. For each setting, there are altogether 10 simulation runs. The variance of the simulation results is shown in error bars in Figure 5.

As shown in Figure 5, the proposed model matches the simulation results very well (\( \leq \pm 2\% \)) for both the case of a single flow and large number of flows (\( \geq 10 \)). However, when the number of flows is small (2 - 5), the deviation of the theoretical results from the simulation results are larger (\( \leq 10\% \)). The error mainly comes from the combinatorial approximation for the number of active flows at the end of a beacon interval. For large number of competing flows, such effect is alleviated due to the randomization in the system. Therefore, the proposed model provides a good approximation for larger number of flows. Another source of inaccuracy of the model is due to the assumption that a packet under-service at the end of a beacon interval restarts its service in the next interval in Section III. The effect of such approximation is more pronounced at the high end of traffic load. When the load is higher, it is more likely a packet is in the middle of service at the end of beacon interval. Thus the percentage of active interval predicted by the model is higher compared to the simulation results. The main observation we have from this set of results is, stochastic multiplexing of multiple competing flows helps improve the energy efficiency. For example, when there are only two flows in the system and each flow sends at 0.9Mbps, the percentage of active interval of each flow is \( \sim 88\% \). On the other hand, if each flow is assigned a separate channel of 1Mb, the percentage of active interval is around 95%.

Though our results are obtained for multiple homogeneous flows, we believe similar observation can be applied to heterogeneous flows. However, the effect of multiplexing on energy consumption may be less pronounced for dominating flows which transmit at much higher rates than the others.

C. Summary

From the previous analysis and numerical results, we observe that at low traffic loads, the IEEE PSM with large beacon intervals saves energy at the expense of large delays. At high traffic load, small beacon intervals are desirable since the packet loss rate is smaller. However, the energy efficiency of the IEEE 802.11 PSM in the light load regime is poor with single or multiple contending flows. Tuning the ratio and length of active and PS interval does not have significant impact in improving the energy efficient due to the overhead of beacon messages. Instead, approaches that proactively switch the power management states at finer control granularity should be considered.

VI. RELATED WORK

To the best of our knowledge, this is the first work to quantify the energy-performance trade-off for the IEEE 802.11 PSM in wireless networks.
There exist several works that address the energy-performance trade-off in wireless networks from different perspectives. Krashinsky et al. [9] identified that using fixed beacon intervals is wasteful in energy while incurring too much delay. They proposed a bounded slow-down (BSD) scheme that essentially probes the round trip time RTT between the request and response progressively. Upon the transmission of a request, the node remains on for \( t_s \) time. To achieve a bounded slowdown of percentage \( p \), the initial awake interval \( t_s = b/p \), where \( b \) is the beacon interval. The solution explores the dependency in two-way traffic but is not applicable to in-bound requests. Jung et al. [8] observed the inefficiency of the IEEE 802.11 PSM due to fixed beacon interval length and proposed a scheme that use a separate DATA window to control the transition to the low-power state in the middle of a beacon interval (as compared to waiting till the end). For fixed traffic load, the approach is similar to using a smaller active interval and longer low-power interval. Therefore, performance of this scheme can be analyzed using our model.

The recent work by Anand et al [2] demonstrates the performance degradation of the IEEE 802.11 PSM for latency-sensitive applications. They proposed a self-tuning power management (STPM) that adapts its behavior to the patterns and intent of applications, the characteristics of the network interface, and the energy usage of the platform. STPM is implemented as a Linux kernel module and is shown to achieve better energy-efficiency for a wide range of application network access patterns. STPM requires modification of applications to provide hints to the power management module. The analysis we conducted in this work targets for the aggregated behavior of a mobile station or multiple competing mobile stations in IBSS mode.

VII. CONCLUSION AND FUTURE WORK

In this paper, we present a transient analysis based study of the energy-performance trade-off of the IEEE 802.11 PSM and its generalized version in wireless networks. We believe such a quantitative study can not only provide an evaluation tool for the IEEE 802.11 PSM but also give insights on improving power management designs. Our main observation is that the IEEE 802.11 PSM in IBSS mode is not energy-efficient for light traffic load. Tuning beacon interval based on the traffic load cannot strike a good balance between energy and performance.

As part of our future work, we are interested in investigating the interaction of power management with channel access contention in IEEE 802.11 networks and devising energy-efficient scheduling in multi-hop wireless networks.

VIII. ACKNOWLEDGMENT

The author would like to thank Prof. P. R. Kumar and Prof. Bruce Hajek for helpful suggestions on the derivations.

REFERENCES