On the Validity of IEEE 802.11 MAC Modeling Hypotheses

Kaidi Huang, Ken R. Duffy, and David Malone

***Abstract*—We identify common hypotheses on which a large number of distinct mathematical models of WLANs employing IEEE 802.11 are founded. Using data from an experimental test bed and packet-level ns-2 simulations, we investigate the veracity of these hypotheses. We demonstrate that several of these assumptions are inaccurate and/or inappropriate. We consider hypotheses used in the modeling of saturated and unsaturated 802.11 infrastructure mode networks, saturated 802.11e networks, and saturated and unsaturated 802.11s mesh networks. In infrastructure mode networks, we find that even for small numbers of stations, common hypotheses hold true for saturated stations and also for unsaturated stations with small buffers. However, despite their widespread adoption, common assumptions used to incorporate station buffers are erroneous. This raises questions about the predictive power of all models based on these hypotheses. For saturated 802.11e models that treat differences in arbitration interframe space (AIFS), we find that the two fundamental hypotheses are reasonable. For 802.11s mesh networks, we find that assumptions are appropriate only if stations are lightly loaded and are highly inappropriate if they are saturated. In identifying these flawed suppositions, this work identifies areas where mathematical models need to be revisited and revised if they are to be used with confidence by protocol designers and WLAN network planners.**

***Index Terms*—Carrier sensing multiple access/collision avoidance (CSMA/CA), hypothesis testng, IEEE 802.11, stochastic models.**

I. Introduction

SINCE its introduction in 1997, IEEE 802.11 has become the *de facto* WLAN standard. Its widespread deployment has led to considerable research effort to gain understanding of its carrier sensing multiple access/collision avoidance (CSMA/CA) algorithm. This includes the use of simulation tools, experiments with hardware, and the building of mathematical models. In particular, analytical models have developed significantly in recent years, and due to the speed with which they can make predictions, they have been proposed as powerful tools to aid protocol designers and WLAN network planners.

Despite the differences in the details of published models, most of them share common hypotheses. In this article, we identify these common assumptions and investigate their validity. This is an important task as authors do not typically check the validity of these assumptions directly, but infer them from the accuracy of model predictions of coarse-grained quantities such as long-run throughput or average MAC delay. If these models are to be used with confidence for the prediction of quantities beyond those validated within published articles, it is necessary that their fundamental hypotheses be sound.

In the present article, we identify the following assumptions that are adopted by numerous authors, often implicitly. For a single station, define $C_k := 1$ if the kth transmission attempt results in a collision, and $C_k := 0$ if it results in a success. For a station in an 802.11 network employing DCF, irrespective of whether it is saturated (always having packets to send) or not, many authors (e.g., [2]–[10]) assume the following.

(A1) The sequence of outcomes of attempted transmissions $\{C_k\}$ forms a stochastically independent sequence.

(A2) The sequence $\{C_k\}$ consists of identically distributed random variables that, in particular, do not depend on station's collision history.

For models where stations have nonzero buffers, in addition define $Q_k := 1$ if there is at least one packet awaiting processing after the kth successful transmission or packet discard, and $Q_k := 0$ if the buffer is empty. Then, the following are commonly assumed (e.g., [11]–[14]).

(A3) The sequence $\{Q_k\}$ consists of independent random variables.

(A4) The sequence $\{Q_k\}$ consists of identically distributed random variables that, in particular, do not depend on backoff stage.

Many authors consider networks employing the 802.11 EDCA in which stations have distinct arbitration interframe space (AIFS) parameters. For a network with two distinct AIFS values, let $\{H_k\}$ denote the sequence of the number of slots during which stations with a lower AIFS value can decrement their counters while stations with the higher AIFS observe the medium as being continuously busy for longer. The commonly adopted assumptions (e.g., [15]–[20]) are the following.

(A5) The sequence $\{H_k\}$ consists of independent random variables.

(A6) Each element of the sequence $\{H_k\}$ is identically distributed, with a distribution that can be identified with one derived in Section VI.

For 802.11s mesh networks, let $\{D_k\}$ denote the interdeparture times of packets from an element of the network. That is, D_k is the difference between the time at which the kth successful transmission and the $(k-1)$th successful transmission occurs from a tagged station. If the station's arrival process is Poisson, then one pair of hypotheses (e.g., [21]) used to enable a tractable mathematical model of 802.11s mesh networks is the following.

Manuscript received October 02, 2008; revised December 10, 2009; accepted May 18, 2010; approved by IEEE/ACM TRANSACTIONS ON NETWORKING Editor V. Misra. This work was supported by Science Foundation Ireland Grant RFP-07-ENEF530. A preliminary report on this work appeared in the *Proceedings of IEEE PIMRC 2008* [1].

The authors are with the Hamilton Institute, National University of Ireland, Maynooth, Ireland (e-mail: ken.duffy@nuim.ie).

Digital Object Identifier 10.1109/TNET.2010.2051335

TABLE I
SUMMARY OF FINDINGS: $\{C_k\}$ COLLISION SEQUENCE; $\{Q_k\}$ QUEUE-OCCUPIED SEQUENCE; $\{H_k\}$ HOLD SEQUENCE; $\{D_k\}$ INTERDEPARTURE TIME SEQUENCE

Assumption	Saturated	Small Buffer	Big Buffer
(A1) $\{C_k\}$ independent	✓ (pairwise)	✓ (pairwise)	✓ (pairwise)
(A2) $\{C_k\}$ identical distributed	✓	✓	×
(A3) $\{Q_k\}$ independent	-	-	✓ (pairwise)
(A4) $\{Q_k\}$ identical distributed	-	-	×
(A5) $\{H_k\}$ independent	✓ (pairwise)	-	-
(A6) $\{H_k\}$ specific distribution	✓	-	-
(A7) $\{D_k\}$ independent	✓ (pairwise)	✓ (pairwise)	✓ (pairwise)
(A8) $\{D_k\}$ exponential distributed	×	only if lightly loaded	only if lightly loaded

(A7) The sequence $\{D_k\}$ consists of independent random variables.

(A8) The elements of $\{D_k\}$ are exponentially distributed.

By performing statistical analysis on large volumes of data collected from an experimental test bed, as described in Section III, we investigate the veracity of the assumptions (A1)–(A4), (A7), and (A8). Due to instrumentation difficulties, we use data from ns-2 packet-level simulations to check (A5) and (A6). Our findings are summarized in Table I.

Autocovariance, Runs Tests, maximum likelihood estimators, and goodness-of-fit statistics lead us to deduce that (A1) and (A2) are reasonable hypotheses for saturated stations and unsaturated stations with small buffers, but are not accurate for unsaturated stations with large buffers. Of greater concern for unsaturated stations with large buffers is that we find (A4) is a dubious and inaccurate assumption. In particular, the queue-busy probability is clearly shown to be a function of backoff stage. For 802.11e models that treat differences in AIFS, we find that (A5) and (A6) are reasonable (albeit not perfect). For 802.11s mesh networks, we find that (A7) and (A8) are appropriate for lightly loaded unsaturated stations, but (A8) fails as stations become heavily loaded. This is particularly true when stations are saturated as interdeparture times coincide with MAC service times. Although these are independent, which validates (A7), they are not distributed exponentially, which contradicts (A8).

Clearly, these findings give rise to serious concerns about the appropriateness of many commonly adopted modeling assumptions. As these hypotheses are inaccurate or inappropriate, it is hard to have confidence in the predictions of models based on them beyond their original validation. Our aim in this article is to draw attention to these deficits and guide the 802.11 modeling community in its ongoing research effort. This is crucial if these models are to be used by network designers.

The rest of this paper is organized as follows. In Section II, we introduce two of the popular 802.11 modeling approaches: p-persistent and mean-field Markov. In Section III, the instrumented test bed used to collect experimental data is introduced. In Section IV, we treat the fundamental decoupling hypotheses, (A1) and (A2), for saturated stations as well as unsaturated stations with both small and large buffers. In Section V, we consider the additional queue decoupling assumptions, (A3) and (A4), that are adopted when treating stations with buffers. In Section VI, the assumptions (A5) and (A6) that lie behind the treatment of different AIFS values in 802.11e are considered. In Section VII, the mesh assumptions (A7) and (A8) that Poisson input gives rise to Poisson output are treated. In Section VIII, we discuss our findings.

II. POPULAR ANALYTIC APPROACHES TO 802.11 DCF AND EDCA

At its heart, the 802.11 CSMA/CA algorithm employs binary exponential backoff (BEB) to share the medium between stations competing for access[1]. As this BEB algorithm couples stations service processes through their shared collisions, its performance cannot be analytically investigated without judiciously approximating its behavior. There are two popular modeling paradigms: the p-persistent approach and the mean-field Markov model approach. The former has a long history in modeling random access protocols, such as Ethernet and Aloha [22], and the latter has its foundations in Bianchi's seminal papers [2], [3]. While these approaches differ in their ideology and details, we shall see that they share basic decoupling hypotheses. Irrespective of the paradigm that is adopted, most authors validate model predictions, but do not directly investigate the veracity of the underlying assumptions.

All of the models we shall consider are based on the assumption of idealized channel conditions where errors occur only as a consequence of collisions. With this environmental conditioning, the key decoupling approximation that enables predictive models in all p-persistent and mean-field analytic models of the IEEE 802.11 random access MAC is that given a station is attempting transmission, there is a fixed probability of collision that is independent of the past.

In p-persistent models, this arises as each station is assumed to have a fixed probability of attempted transmission τ per idle slot that is independent of the history of the station and independent of all other stations. In a network of N identical stations, the likelihood a station does not experience a collision given it is attempting transmission $1-p$ is the likelihood that no other station is attempting transmission in that slot: $1-p=(1-\tau)^{N-1}$. Thus, the sequence of collisions or successes is an independent and identically distributed sequence. In the p-persistent approach, the attempt probability τ is chosen in such a way that the average time to successful transmission matches that in the real system, which is an input to the model. If this average is known, this methodology has been demonstrated to make accurate throughput and average delay predictions [4], [5].

[1]A brief overview of the DCF and EDCA algorithms is given in Appendix I.

In the mean-field approach, the fundamental idea is similar, but the calculation of τ, and thus p, does not require external inputs. One starts by assuming that p is given and each station always has a packet awaiting transmission (the saturated assumption). Then, the backoff counter within the station becomes an embedded semi-Markov process whose stationary distribution can be determined [2], [3]. In particular, the stationary probability that the station is attempting transmission, $\tau(p)$, can be evaluated as an explicit function of p and other MAC parameters ([3, Eq. (7)]). In a network of N stations, as $\tau(p)$ is known, the fixed-point equation $1-p=(1-\tau(p))^{N-1}$ can be solved to determine the "real" p for the network and "real" attempt probability τ. Once these are known, network performance metrics, such as long-run network throughput, can be deduced.

Primarily through comparison with simulations, models based on these assumptions have been shown to make accurate throughput and delay predictions, even for small number of stations. This is, perhaps, surprising as one would expect that the decoupling assumptions would only be accurate in networks with a large number of stations. The p-persistent paradigm has been developed to encompass, for example, saturated 802.11 networks where every station always has a packet to send [4] and saturated 802.11e networks [8]. However, due to its intuitive appeal, its self-contained ability to make predictions, and its predictive accuracy, Bianchi's basic paradigm has been widely adopted for models that expand on its original range of applicability. A selection of these extensions includes: [6], [7], [9], [10], which consider the impact of unsaturated stations in the absence of station buffers and enable predictions in the presence of load asymmetries; [11]–[14], which treat unsaturated stations in the presence of stations with buffers; [15]–[17], which investigate the impact of the variable parameters of 802.11e, including AIFS, on saturated networks; [18]–[20], which treat unsaturated 802.11e networks; [21], which extends the paradigm from single-hop networks to multiple-radio 802.11s mesh networks. Note that the work cited here is a small, selective subcollection within a vast body of literature. To appreciate just how large this literature is, as of November 2009, the p-persistent modeling paper [4] has been cited over 400 times according to ISI Knowledge and over 800 times according to Google Scholar, while the mean-field modeling paper [3] has been cited over 1300 times according to ISI Knowledge and over 3200 times according to Google Scholar.

All of the extensions that we cite are based on the idealized channel assumption as well as the decoupling approximation. Some of these extensions require further additional hypotheses. The purpose of the present article is to dissect these fundamental assumptions to determine the range of the applicability of models based on them.

III. Experimental Apparatus

The experimental apparatus is set up in infrastructure mode. It employs a PC acting as an access point (AP), another PC, and nine PC-based Soekris Engineering net4801 embedded Linux boxes acting as client stations. For every transmitted packet, the client PC records the transmitting timestamp (the time when it receives an ACK), the number of retry attempts experienced, and the absence or presence of another packet in station's buffer, but otherwise behaves as an ordinary client station. All stations are equipped with an Atheros AR5215 802.11b/g PCI card with an external antenna. All stations, including the AP, use a version of the MADWiFi wireless driver modified to allow packet transmissions at a fixed 11-Mb/s rate with RTS/CTS disabled and a specified queue size. The 11-Mb/s rate was selected as in the absence of noise-based losses the MAC's operation is rate-independent, and more observations of transmission can be made at higher rates for an experiment of given real-time duration. The channel on which experiments were conducted was confirmed to be noise-free by use of a spectrum analyzer and by conducting experiments with single transmitter–receiver pairs.

All stations are equipped with a 100-Mb/s wired Ethernet port that is solely used for control of the test bed from a distinct PC. In the experiments, UDP traffic is generated by the Naval Research Laboratory's MGEN in Poisson mode. All UDP packets have a 1000-byte payload and are generated in client stations before being transmitted to the AP. At the AP, tcpdump is used to record traffic details.

Hoeffding's concentration inequality (described in Section IV) was used to determine how many observations were necessary to ensure statistical confidence in estimated quantities. Consequently, saturated and large buffer experiments were run for 2 h while small buffer experiments were run for 4 h.

Care must be exercised when performing experiments with IEEE 802.11 devices. Bianchi *et al.* [23] and Giustiniano *et al.* [24] report on extensive validation experiments that clearly demonstrate that cards from many vendors fail to implement the protocol correctly. The precision of our experimental apparatus was established using the methodology described in [23] with additional statistical tests.

For example, to check if the backoff counters are uniformly distributed, the sequence of transmission times of a single saturated station with fixed packet sizes were recorded. The backoff counter values were inferred from this sequence by evaluating the intertransmission times less the time taken for a packet transmission $(\mathrm{DIFS}+\mathrm{Payload}+\mathrm{SIFS}+\mathrm{ACK})$ and then dividing this quantity by the idle slot length. When the contention window is 32 or 64, Fig. 1 reports on a comparison of the protocol's backoff distribution with empirical distributions based on sample sizes of 8 706 941 and 7 461 686, respectively. With a null hypothesis that the distributions are uniform, Pearson's χ^2-test (described in Appendix II) gives p-values of 0.7437 and 0.2036 so that the null hypothesis would not be rejected.

There is one place where our 802.11 cards do not implement the standard correctly, but it does not impact on our deductions; both ACKTimeout and EIFS are shorter than suggested in the standard. While this must be taken into account when, for example, predicting throughput, it has no impact on the aspects of the MAC's operation that are of interest to us.

For added confirmation, all of the results that are reported here based on experimental data were shadowed in parallel by ns-2-based simulations that gave agreement in every case.[2] Thus, we are satisfied that none of the observations reported

[2]Data from these simulations are not shown due to space constraints.

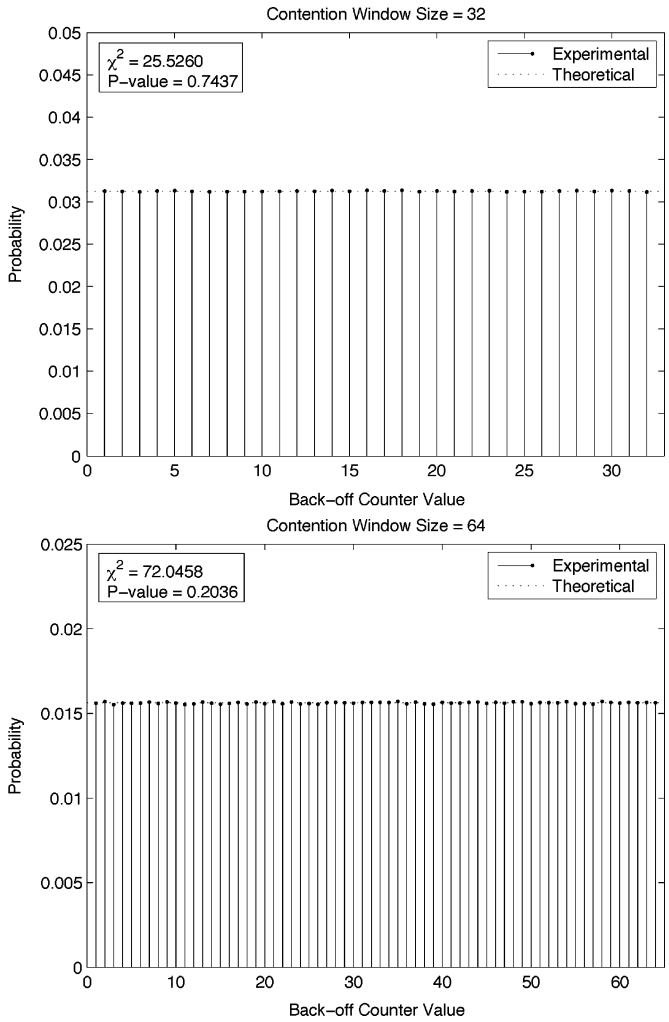

Fig. 1. Comparison of protocol's uniform backoff distribution and empirical distribution for contention windows of size 32 and 64 based on sample sizes of 8 706 941 and 7 461 686, respectively. Pearson's χ^2 does not reject the hypothesis that the distributions are uniform. Experimental data.

in this article are a consequence of peculiarities of the cards, drivers, or experimental environment.

IV. Assumptions (A1) and (A2)

For a single station, define $C_k := 1$ if the kth transmission attempt results in a collision, and $C_k := 0$ if it results in a success. The two key assumptions in [2]–[5] are effectively these: (A1) the sequence $\{C_k\}$ consists of independent random variables; and (A2) the sequence $\{C_k\}$ consists of identically distributed random variables. That is, there exists a fixed collision probability conditioned on attempted transmission, $P(C_k = 1) = p$, that is assumed to be the same for all backoff stages and independent of past collisions or successes.

The assumptions (A1) and (A2) are common across all models developed from the p-persistent and mean-field paradigms. Here, we investigate these for saturated stations, for unsaturated stations with small buffers, and for unsaturated stations with big buffers. All network parameters correspond to standard 11-Mb/s IEEE 802.11b [25].

We begin by investigating (A1), the hypothesized independence of the outcomes (success or collision) in the sequence

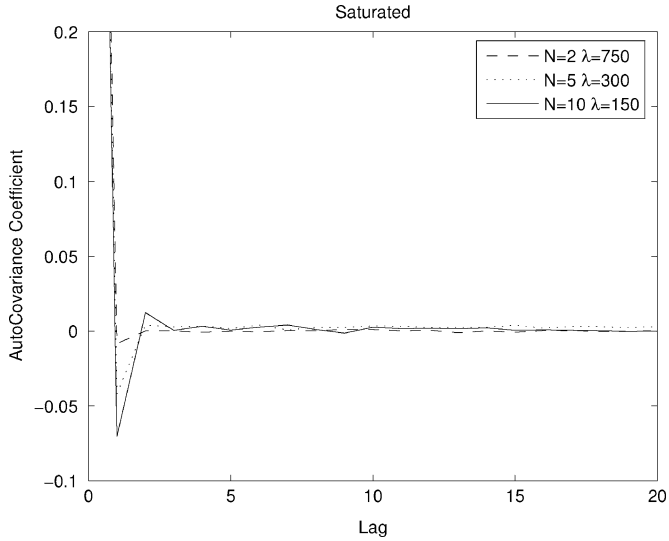

Fig. 2. Saturated collision sequence normalized autocovariances. Experimental data.

TABLE II
Number of Attempted Transmissions $K^{(C)}$. Experimental Data

Number of stations	Saturated	Small buffer	Big buffer
$N = 2$	2,549,550	2,134,187	1,846,049
$N = 5$	1,220,622	975,601	749,295
$N = 10$	711,326	502,955	380,139

of transmission attempts. We draw deductions regarding pair-wise independence from the normalized autocovariance of the sequence $C_1, C_2, \ldots, C_{K^{(C)}}$ obtained experimentally, where $K^{(C)}$ is the number of attempted transmissions that a single tagged station makes during the experiment. For each experiment, $K^{(C)}$ is recorded in Table II. Assuming $\{C_k\}$ is wide-sense stationary, the normalized autocovariance, which is a measure of the dependence in the sequence, is always 1 at lag 0, and if the sequence $\{C_k\}$ consisted of independent random variables, as hypothesized by (A1), then for a sufficiently large sample it would take the value 0 at all positive lags. Nonzero values correspond to apparent dependencies in the data.

Experiments were run for a saturated network with $N = 2$, 5, and 10. As the number of stations is increased, the number of attempts by the tagged station decreases due to the backing-off effects of the MAC. Fig. 2 reports the normalized autocovariances for these sequences at short lags. The plot quickly converges to zero, indicating little dependence in the success per attempt sequence, even for $N = 2$.

The (A1) and (A2) assumptions are also adopted in unsaturated models with small buffers [6], [7], [9], [10] and big buffers [11]–[14]. From experimental data for unsaturated networks large station buffers, Fig. 3 plots the normalized autocovariances of the attempted transmissions for $N = 2$, 5, and 10 with both small and large buffers. As in all unsaturated models that we are aware of, packets arrive at each station as a Poisson process with rate λ packets per second. In the big buffer experiments, the overall network load is kept constant at 500 packets per second, equally distributed amongst the N stations, corresponding to a network-wide offered load of 4.25 Mb/s. In

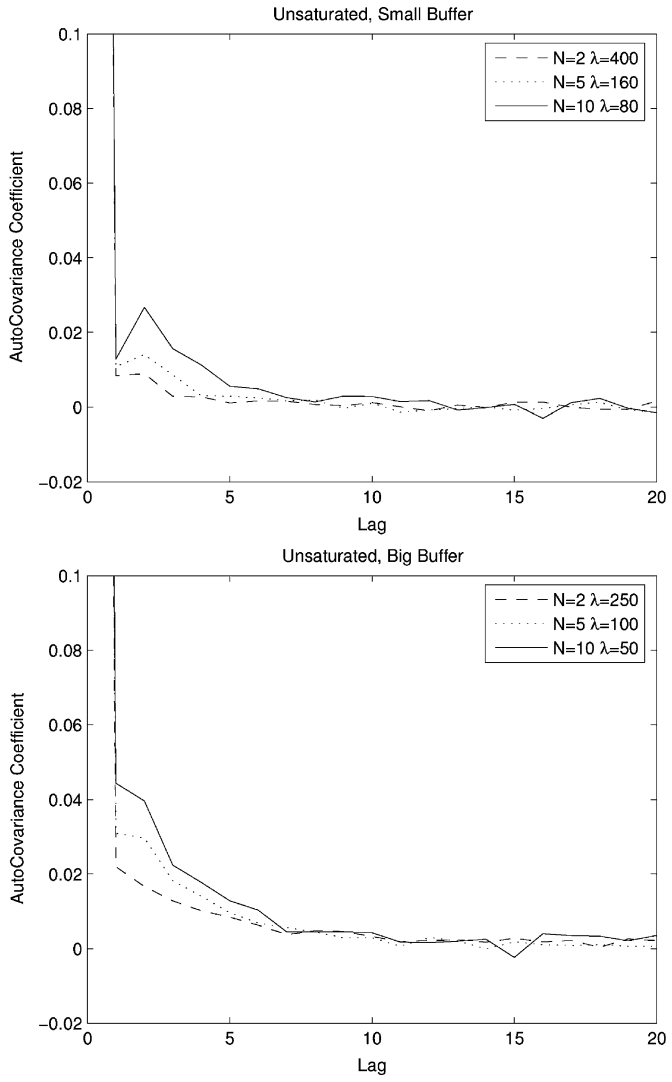

Fig. 3. Unsaturated small and big buffer collision sequence normalized autocovariances. Note the short *y*-range. Experimental data.

the small buffer experiments network load is kept constant at 800 packets per second. Again, we only show short lags as the autocovariance quickly drops to zero, indicating little pair-wise dependency in the $C_1, \ldots, C_{K^{(C)}}$ sequence and supporting the (A1) hypothesis.

We have seen graphs similar to Figs. 2 and 3 for a range of offered loads, which are not shown due to space constraints. These support the (A1) hypothesis that the sequence of collision or success at each attempted transmission is close to being a pair-wise independent one.

Before investigating the (A2) hypothesis on its own, we use the Runs Test (described in Appendix III) to jointly test (A1) and (A2). Given a binary-valued sequence, $C_1, \ldots, C_{K^{(C)}}$, this test's null hypothesis is that it was generated by a Bernoulli sequence of random variables. The test is nonparametric and does not depend on $P(C_1 = 1)$. The Runs Test statistics for each of our nine collision sequences range from 11.6617 for the saturated two-station sequence to 68.5831 for the unsaturated big-buffer two-station sequence. The likelihood that the data was generated by a Bernoulli sequence is a decreasing function of the test value, and if this value is 2.58, there is less than 1% chance that it was generated by a Bernoulli sequence. Even the lower end of the range gives a p-value of 0, leading to rejection of the hypothesis that the collision sequences are i.i.d. The reason for this failure will become apparent when we demonstrate that $P(C_k = 1)$ depends heavily on an auxiliary variable α_k, the backoff stage at which attempt k was made, and that, as is clear from the DCF algorithm, $\{\alpha_k\}$ cannot form an i.i.d. sequence.

To investigate the (A2) hypothesis on its own, we reuse the same collision sequence data $C_1, C_2, \ldots, C_{K^{(C)}}$ with some additional information. For each attempted transmission $k \in \{1, \ldots, K^{(C)}\}$, we record the backoff stage α_k at which it was made. Assume that there is a fixed probability p_i that the tagged station experiences a collision given it is attempting transmission at backoff stage i. A consequence of Assumption (A2) is that $p_i = p$ for all backoff stages i. The maximum likelihood estimator for p_i is given by

$$\hat{p}_i = \frac{\sum_{k=1}^{K^{(C)}} C_k \chi(\alpha_k = i)}{\sum_{k=1}^{K^{(C)}} \chi(\alpha_k = i)} \tag{1}$$

where $\chi(\alpha_k = i) = 1$ if $\alpha_k = i$, and 0 otherwise. The numerator in (1) records the number of collisions at backoff stage i, while the denominator records the total number of attempts at backoff stage i. As $\{C_k\}$ is a sequence of bounded random variables that appear to be nearly independent (although not i.i.d.), we apply Hoeffding's inequality [26] to determine how many samples we need to ensure we have confidence in the estimate $\hat{p}_i$

$$\begin{aligned} &P(|\hat{p}_i - p_i| > x) \\ &\quad = P\left(\left| \sum_{k=1}^{n} (C_k - E(C_k)) \chi(\alpha_k = i) \right| > x \left| \sum_{k=1}^{K^{(C)}} \chi(\alpha_k = i) \right. \right) \\ &\quad \leq 2 \exp\left(-2x \sum_{k=1}^{K^{(C)}} \chi(\alpha_k = i) \right). \end{aligned}$$

Using this concentration inequality, to have at least 95% confidence that $|\hat{p}_i - p_i| \leq 0.01$ requires $\sum_{k=1}^{K^{(C)}} \chi(\alpha_k = i) = 185$ attempted transmissions at backoff stage i. If we have less than 185 observations at backoff stage i, we do not have confidence in the estimate's accuracy, so it is not plotted.

Starting with the saturated networks, Fig. 4 plots the estimates $\hat{p}_i$ for the tagged station as well as the predicted value from [2], [3]. For $N = 2$, we only report backoff stages 0 to 3 due to lack of observations. It can be seen that the $\hat{p}_i$ are similar for all i. To quantify this, with $S := \max \hat{p}_i - \min \hat{p}_i$, $S = 0.01$ for $N = 2$, $S = 0.038$ for $N = 5$, and $S = 0.074$ for $N = 10$. Note that while the estimated values are not identical to those predicted by Bianchi's model, they are close. These observations support the (A2) assumption for saturated stations, even for $N = 2$.

For unsaturated WLANs, we do not include theoretical predictions for comparison as, unlike the saturated setting, there

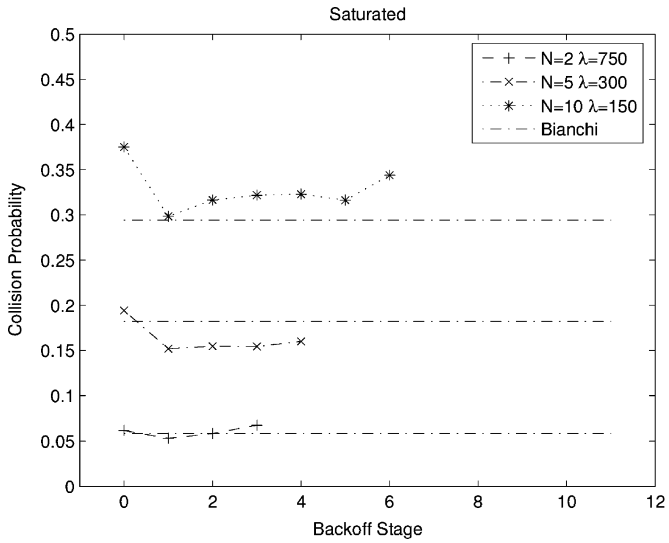

Fig. 4. Saturated collision probabilities. Experimental data.

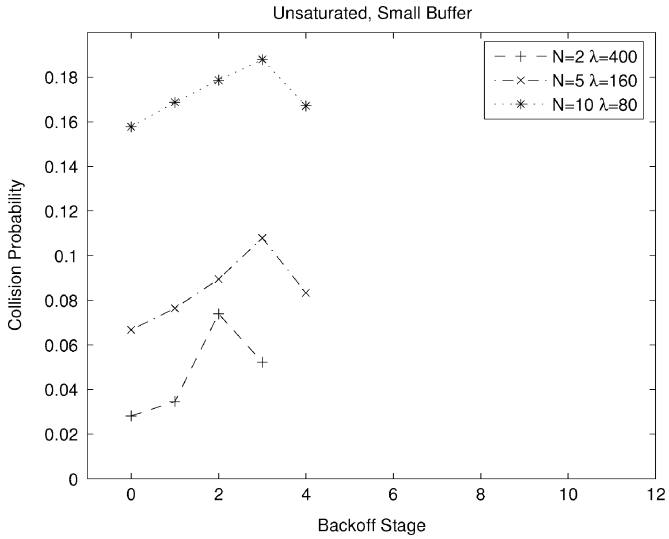

Fig. 5. Unsaturated, small buffer collision probabilities. Experimental data.

are a large range of distinct models to choose from. Plotting the predictions from any single model would not be particularly informative and could reasonably be considered unfair. The significant thing to note is that all of the models we cite assume that $p_i = p$ for all backoff stages i, so that if p_i varies as a function of i, none can provide a perfect match.

Fig. 5 is a plot of the estimates $\hat{p}_i$ for each backoff stage i for the tagged station in the unsaturated three-packet buffer case with $N = 2$,5, and 10, with a network arrival rate of 800 packets per second, equally distributed among the N stations, corresponding to an offered load of 6.8 Mb/s. In comparison to the saturated setting, the absolute variability estimates are similar with S as has been defined, giving $S = 0.045$ for $N = 2$, $S = 0.038$ for $N = 5$, and $S = 0.074$ for $N = 10$. This suggests that (A2) is reasonably appropriate. There is, however, clear structure in the graphs. For each N, the collision probability appears to be dependent on the backoff stage. The collision probability at the first backoff stage is higher than at the

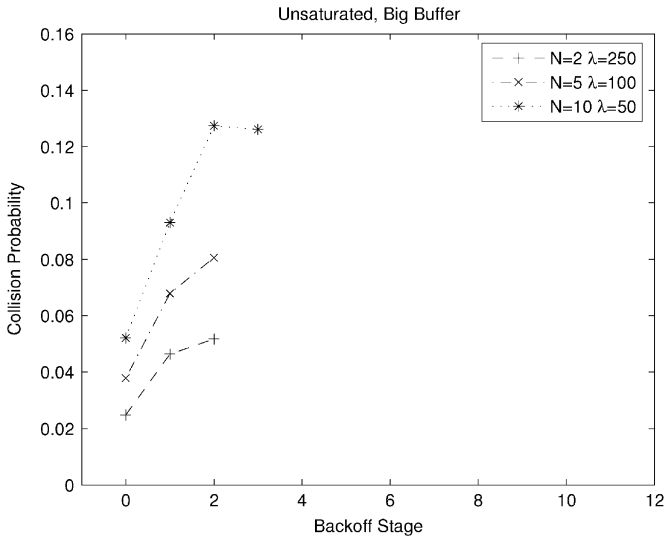

Fig. 6. Unsaturated big-buffer collision probabilities. Experimental data.

zeroth stage. For stations that are unsaturated, we conjecture that this occurs as many transmissions happen at backoff stage 0 when no other station has a packet to send, so collisions are unlikely and $\hat{p}_0$ is small. Conditioning on the first backoff stage is closely related to conditioning that at least one other station has a packet awaiting transmission, giving rise to a higher conditional collision probability at stage 1, so that $\hat{p}_1 > \hat{p}_0$.

Fig. 6 is analogous to Fig. 5, but for stations with 100-packet station buffers. The networks are unsaturated with the queues at each station repeatedly emptying. As with the small buffer case, we again have that $p_1 > p_0$ and conjecture that this occurs for the same reasons. In comparison to the values reported in Figs. 4 and 5, the absolute variability is similar with $S = 0.028$ for $N = 2$, $S = 0.042$ for $N = 5$, and $S = 0.073$ for $N = 10$, but with ϕ being the average of $\{\hat{p}_i\}$, relative variability S/ϕ of the estimates in Fig. 6 is consistently higher than the saturated case: $S/\phi = 0.65$ versus 0.17 for $N = 2, 0.67$ versus 0.23 for $N = 5$, and 0.73 versus 0.22 for $N = 10$. This suggest that (A2) is not as good an approximation in the presence of big station buffers.

In this section, we have investigated the veracity of the assumptions of independence and identical distribution of the outcomes of the collision attempt sequence. The findings are summarized in Table I. In the next section, we consider the additional hypotheses introduced to model buffering.

V. Assumptions (A3) and (A4)

To model stations with buffers serving Poisson traffic, the common idea across various authors, e.g., [11]–[14], is to treat each station as a queueing system where the service time distribution is identified with the MAC delay distribution based on a Bianchi-like model. The assumptions (A1) and (A2) are adopted, so that given conditional collision probability p, each station can be studied on its own, and a standard queueing theory model is used to determine the probability of attempted transmission $\tau(p)$, which is now also a function of the offered load.

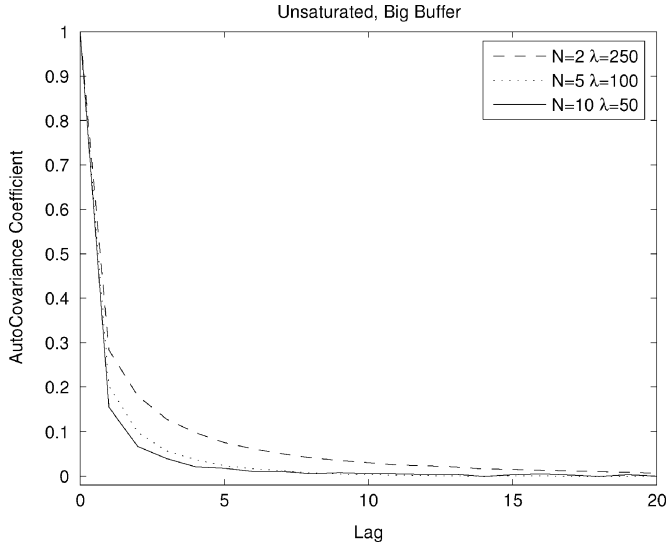

Fig. 7. Unsaturated, big-buffer, queue-non-empty sequence normalized autocovariances. Experimental data.

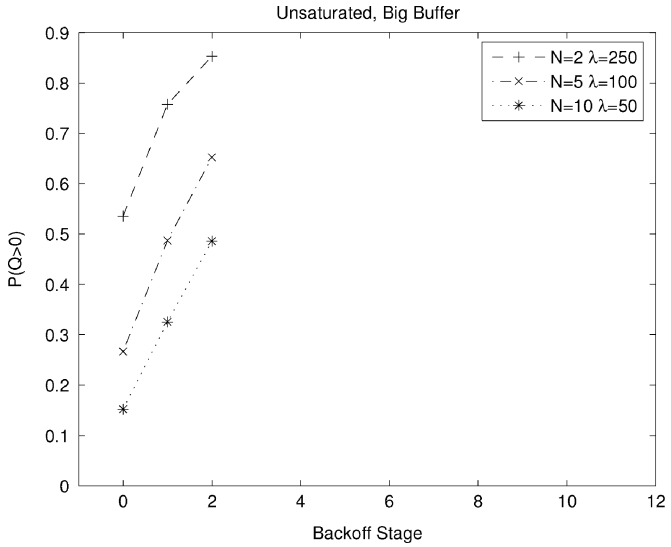

Fig. 8. Unsaturated, big-buffer, queue-non-empty probabilities. Note the large y-range. Experimental data.

For symmetrically loaded stations with identical MAC parameters, the same network coupling equation as used in the saturated system, $1-p=(1-\tau(p))^{N-1}$, identifies the "real" operational conditional collision probability p.

Each time the MAC successfully transmits a packet, it checks to see if there is another packet in the buffer awaiting processing. Define $Q_k := 1$ if there is at least one packet awaiting processing after the kth successful transmission or discard, and $Q_k := 0$ if the buffer is empty. As it is technically challenging to fully model these queueing dynamics while still obtaining tractable equations that can be solved more quickly than a simulation can be run, authors typically employ a second queueing-based decoupling assumption that can be distilled into the following two hypotheses: (A3) the sequence $\{Q_k\}$ consists of independent random variables; and (A4) the sequence $\{Q_k\}$ consists of identically distributed random variables, with $P(Q_k = 1) = q$. The value of q is identified with the steady-state probability that an associated M/G/1 or M/G/1/B queueing system has a nonempty buffer after a successful transmission (e.g., [27]).

Clearly, (A3) and (A4) are more speculative than (A1) and (A2), as both disregard obvious dependencies in the real $Q_1, \ldots, Q_{K^{(Q)}}$ sequence, where $K^{(Q)}$ is the number of successful transmissions from the tagged station. These occur as if there are two or more packets awaiting processing after a successful transmission, there will still be another packet awaiting transmission after the next successful transmission, and in the presence of station buffers, the longer a packet has been awaiting transmission, the more likely it is to have another waiting in its buffer.

To investigate (A3), we evaluate the normalized autocovariance of the empirical sequences $Q_1, \ldots, Q_{K^{(Q)}}$ for $N = 2$, 5, and 10 with $K^{(Q)} = 1\,799\,250, K^{(Q)} = 720\,044$, and $K^{(Q)} = 359\,413$, respectively. These autocovariances are reported in Fig. 7, where it can be seen that there is a small amount of correlation structure, but that by lag 5 this is less than 0.2, and so we do not regard it as significant. As one would expect, this is a function of the load. As stations become more heavily loaded, we have seen this correlation structure become more prevalent, until stations are saturated, whereupon the correlation disappears as $Q_k = 1$ for all k.

To jointly test (A3) and (A4), again we use the Runs Test statistic described in Appendix III. In comparison to the collision sequences, the test statistics are even more extreme with 397.46 for $N = 2$, 171.39 for $N = 5$, and 130.23 for $N = 10$. Thus, this test leads to p-values of 0 in all cases and the rejection of an i.i.d. queue-busy sequence. As with the collision sequences, it will be clear that this happens as $P(Q_k = 1)$ depends on the backoff stage of the kth successful transmission.

To investigate (A4), let β_k denote the backoff stage at the kth successful transmission. With q_i denoting the probability there is another packet awaiting transmission after a successful transmission at backoff stage i, its maximum likelihood estimator is

$$\hat{q}_i = \frac{\sum_{k=1}^{K^{(Q)}} Q_k \chi(\beta_k = i)}{\sum_{k=1}^{K^{(Q)}} \chi(\beta_k = i)}. \tag{2}$$

Although (A3) does not appear to hold at short lags, we can again use Hoeffding's bound to heuristically suggest we need at least 185 observations at a given backoff stage in order to be confident in its accuracy.

Fig. 8 shows these $\hat{q}_i$ estimates for all stations in each network. They show a strong increasing trend as a function of backoff stage. This is as one might expect, given that the longer a packet spends while awaiting successful transmission, the more likely it is that there will be another packet awaiting processing when it is sent. Note that this dependency on backoff stage raises questions over all buffered models that adopt the assumption (A4). These findings are summarized in Table I.

VI. Assumptions (A5) and (A6)

The 802.11e standard, ratified in 2005, enables service differentiation between traffic classes. Each station is equipped

with up to four distinct queues, with each queue effectively being treated as a distinct station. Each queue has its own MAC parameters including minimum contention window ($\mathrm{CW}_{\min}$), retry limit, transmission opportunity (TXOP), and arbitration interframe space (AIFS). To model the first three of these, no additional modeling assumptions are necessary beyond (A1)–(A2) for saturated stations or for unsaturated stations with small buffers (e.g., [10]). For unsaturated stations with large buffers, no additional assumptions are necessary beyond (A1)–(A4) (e.g., [14]). However, to capture the full power of 802.11e's service differentiation, one must model AIFS, and this requires additional innovation and hypotheses [15]–[20].

Consider $N_1 + N_2$ stations, each of which is serving one of two traffic classes with distinct AIFS values: N_1 class one stations with AIFS_1 and N_2 class 2 stations with $\mathrm{AIFS}_2 = \mathrm{AIFS}_1 + D\sigma$, where σ is the length of an idle slot and D is a positive integer. After every attempted transmission, class-1 stations decrement their backoff counters by a minimum of D before class-2 stations see the medium as being idle. If the backoff counter of a class-1 station becomes 0 during these D slots, it attempts transmission, and once it is complete, class-1 stations can decrement their counters by at least another D before class-2 stations see the medium as being idle.

Consider a network of homogeneous saturated class-1 stations and homogeneous class-2 stations. To model the impact of different AIFS values, using the terminology in [15], we have the notion of hold states for class-2 stations. A class-2 station is in a hold state if class-1 stations can decrement their counters while it cannot. As all class-2 stations have the same AIFS value, they all experience the same class-1 preemption. Once in a hold state, they cannot begin to decrement their backoff counters again (while class-1 stations continue to do so) unless all class-1 stations are silent for D consecutive slots.

Given class-2 stations have just entered a hold state, let $H \in \{D, D+1, \ldots\}$ represent the hold time: the number of slots that pass before class-2 stations escape the hold states. Let $\{H_k\}$ denote the sequence of observations of hold times. Implicitly, the commonly adopted assumptions used to treat AIFS are: (A5) the sequence $\{H_k\}$ consists of independent random variables; and (A6) each element of the sequence $\{H_k\}$ is identically distributed with the same distribution as H defined in (3) and which we now derive.

Within the analytic modeling context, this escape from hold states can be formalized mathematically. Let τ_1 denote the stationary probability a class-1 station attempts transmission. Define $P_{s_1} = (1-\tau_1)^{N_1}$, which is the stationary probability that all class-1 stations are silent (no class-1 station is attempting transmission). Let $\{X_n\}$ denote the sequence of hold states. After a transmission, whether successful or not, class-2 stations enter a hold state, and this process starts in hold state $X_0 = 1$. If the medium is idle (no class-1 stations attempt transmission), which happens with probability P_{s_1}, the station moves to hold state $X_1 = 2$, otherwise it is reset to $X_1 = 1$, and so forth. The process stops the first time that the hold state $D+1$ is reached, whereupon all class-2 stations see the medium as being idle and can decrement their counters. This system forms a Markov chain, portrayed in Fig. 9, with $(D+1) \times (D+1)$ the transition matrix

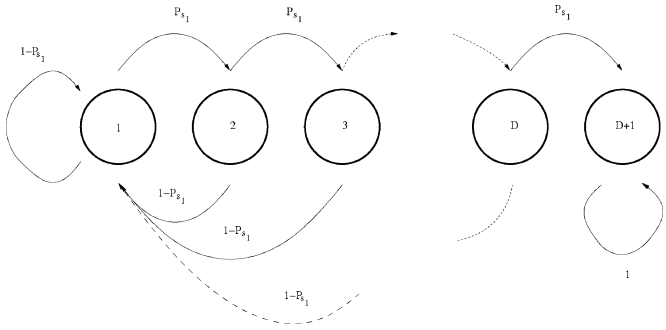

Fig. 9. Markov chain for modeling a difference in AIFS of D slots.

$$\Pi = \begin{pmatrix} 1-P_{s_1} & P_{s_1} & 0 & \ldots & 0 & 0 \\ 1-P_{s_1} & 0 & P_{s_1} & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 1-P_{s_1} & \vdots & \vdots & & 0 & P_{s_1} \\ 0 & \vdots & \vdots & & 0 & 1 \end{pmatrix}.$$

With $X_0 = 1$, we define $H := \inf\{i : X_i = D+1\}$ to be the first time that the D consecutive idle slots are observed. Using the form of the Markov chain, we have

$$\begin{aligned} P(H = i) &= P(X_i = D+1) - P(X_{i-1} = D+1) \\ &= (\Pi^i)_{1,D+1} - (\Pi^{i-1})_{1,D+1}. \end{aligned} \tag{3}$$

Thus, $P(H = i)$ is solely a function of P_{s_1} and can be readily calculated from in (3), albeit not in closed form unless $D = 1$ or $D = 2$.

Due to experimental instrumentation difficulties, the results in this section are based exclusively on ns-2 simulations. In order to determine how many slots the lower class stations have spent in hold states, it is necessary to know the start and finish times of every packet transmitted on the medium. In an experimental setup, no one station is in possession of this information, and due to the time-scales involved, accurate reconciliation of the time-line from data recorded at each station is particularly challenging. In simulation, however, this data is readily accessible.

For a saturated network, the methodology used to measure H is as follows. For each station in the network, a timestamp is recorded at the start and end of every transmission. These timestamps are collected from all stations and combined into a single ordered list: $s_1, e_1, s_2, e_2, \ldots$, where s_i is the start time of the ith packet and e_i is the end time. Define $T_i := (s_{i+1} - e_i - \mathrm{AIFS}_1)/\sigma$ for each $i \geq 1$. The hold times are determined from this sequence by first identifying the indices at which the hold states are delineated: $N_0 := 0$ and $N_{i+1} := \inf\{n \geq N_i : T_n > D\}$ for each $i \geq 1$. The hold states are then the sum of the hold times between transmission, with the final term making a contribution of D slots

$$H_i = \sum_{j=N_i+1}^{N_{i+1}-1} T_j + D$$

where the empty sum is defined to be zero.

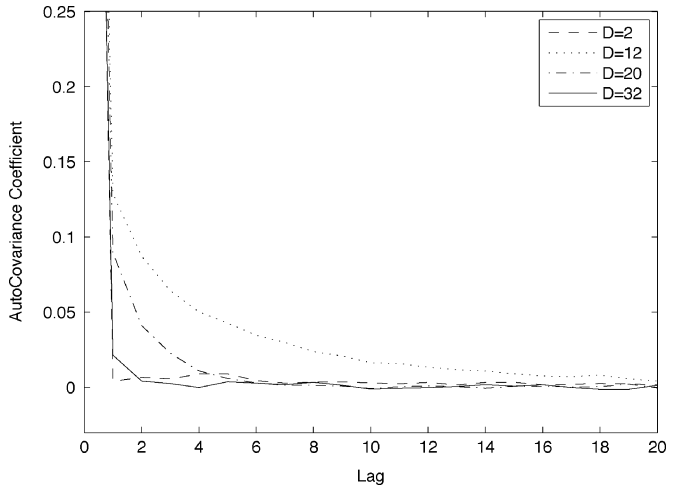

Fig. 10. Autocovariance for period of hold states for class-2 stations in a network of five class-1 and five class-2 saturated station with $D = 2$, 12, 20, and 32. ns-2 data.

To consider the independence assumption (A5), we assume that the sequence of observed hold times $H_1, \ldots, H_{K^{(H)}}$ are wide-sense stationary and plot this data's autocovariance function, thus investigating pair-wise dependence. For a network of five class-1 saturated stations and five class-2 saturated stations that are identical apart from $\mathrm{AIFS}_2 = \mathrm{AIFS}_1 + D\sigma$, where $D = 1$, 2, 4, 8, 12, 16, 20, and 32, these plots suggest that the sequences have little dependence at short lags. Fig. 10 is a representative graph shown for $D = 2$, 12, 20, and 32, with $K^{(H)} = 1\,717\,,545$, $K^{(H)} = 706\,032$, $K^{(H)} = 533\,675$, and $K^{(H)} = 366\,298$, respectively. These D values, particularly the smaller ones, are typical of those proposed for traffic differentiation in the 802.11e standard [28]. As the autocovariance is less than 0.2 by lag 5, the independence assumption (A5) is not unreasonable for differences in AIFS values that are proposed in the standard.

To test the assumption (A6) that the probability density of hold-state idle periods has the form given in (3), rather than use any specific model prediction for the distribution, as it is a function of single parameter P_{s_1}, we estimate P_{s_1} based on the following observation. Given D, if the likelihood that all class-1 stations do not transmit is i.i.d., then $P(H = D) = P_{s_1}^D$. This suggests using the following estimate of P_{s_1}:

$$\left(\frac{\sum_{k=1}^{K^{(H)}} \chi(H_k = D)}{K^{(H)}} \right)^{1/D}.$$

Note that, in using this estimate, we are ensuring that $P(H = D)$ coincides with the empirical observation. However, unless the model is accurate, $P(H = i)$ for $i \neq D$ calculated from (3) need not coincide with the empirically observed value.

Figs. 11 and 12 show a sample of plots for a 10-station network, with five saturated stations in class 1 and five saturated stations in class 2, for $D = 2$ $(K^{(H)} = 5\,152\,635)$ and $D = 12$ $(K^{(H)} = 2\,118\,096)$. Conditioned on having a good estimate of P_{s_1}, the accuracy of the distribution-predicted (3) appears to be remarkable. This apparent accuracy can be explored quantitatively through a test statistic. We use the Kolmogorov–Smirnov test (described in Appendix II), but do not give p-values, as our

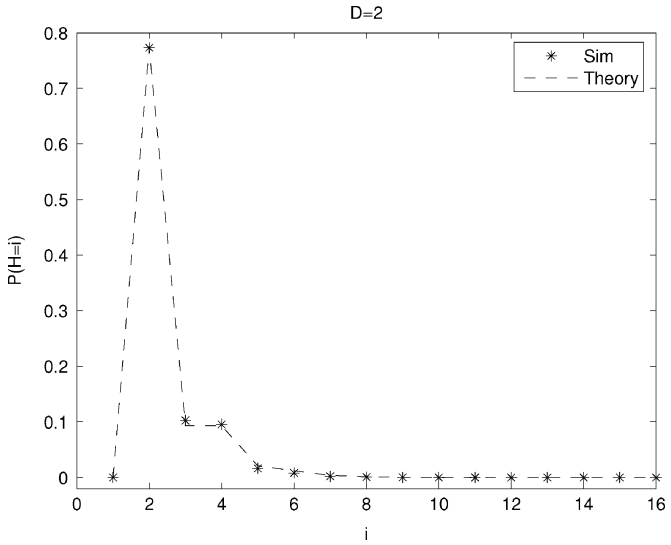

Fig. 11. Empirical and predicted probability density for the length of a hold period for class-2 stations in a network of five class-1 and five class-2 saturated station with $D = 2$. ns-2 data.

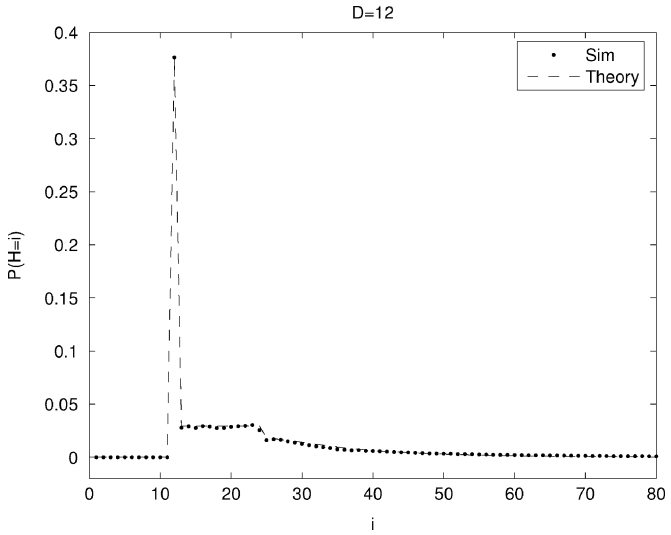

Fig. 12. Empirical and predicted probability density for the length of a hold period for class-2 stations in a network of five class-1 and five class-2 saturated station with $D = 12$. ns-2 data.

distribution is purely discrete. Fig. 13 plots $\sup_k |F_n(k) - F(k)|$ against n for $D = 2$, 4, 8, and 12. It is clear from the graph that the discrepancies are small for moderate values of sample size. However, $\sup_k |F_n(k) - F(k)|$ is not converging to 0 as n becomes large. This suggests that the predicted distribution of H is accurate for all practical purposes, even though the distribution is not a perfect fit. These findings are summarized in Table I.

VII. Assumptions (A7) and (A8)

The 802.11s standard is a draft amendment to enable wireless mesh networks (WMNs). One approach to building a multihop, multiradio mathematical model of a WMN that employs 802.11 is to build on the mean-field Markov ideas, but with more involved coupling that captures medium access dependencies across the mesh. In order to do so, it is necessary to make hypotheses about the stochastic nature of the departures process

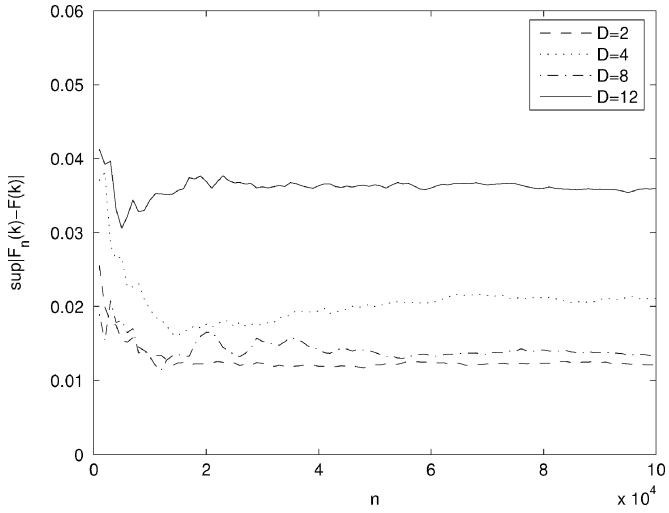

Fig. 13. Largest discrepancy between empirical and predicted distributions, $\sup_k |F_n(k) - F(k)|$, as a function of sample size n. $D = 2$, 4, 8, and 12. ns-2 data.

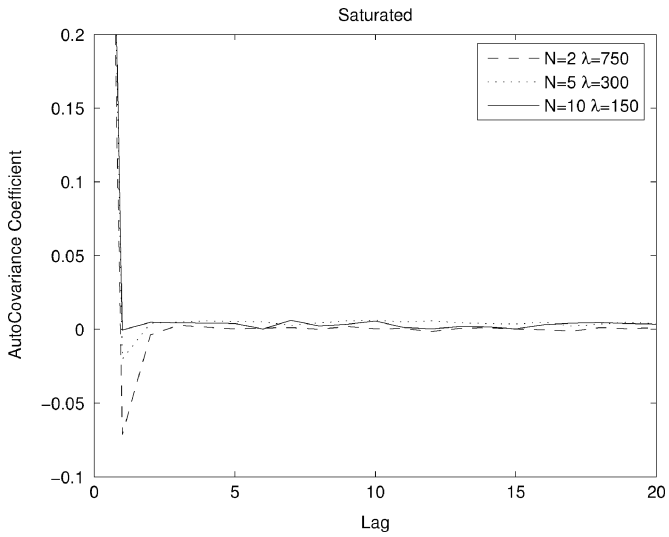

Fig. 14. Saturated interdeparture-time-sequence normalized autocovariances. Experimental data.

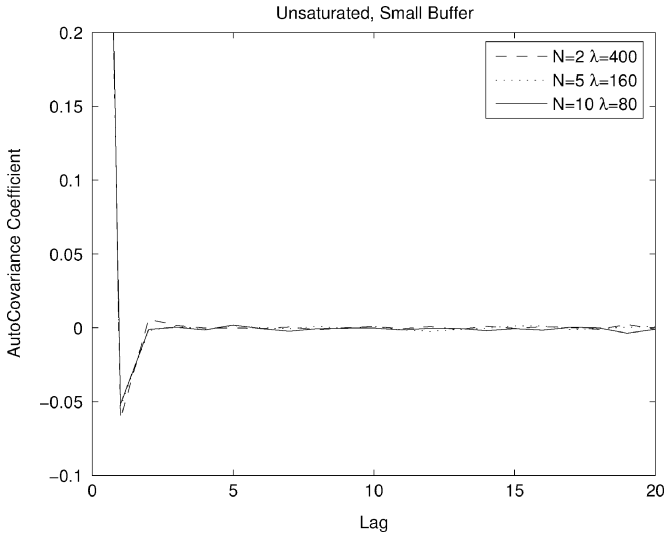

Fig. 15. Unsaturated, small-buffer, interdeparture-time-sequence normalized autocovariances. Experimental data.

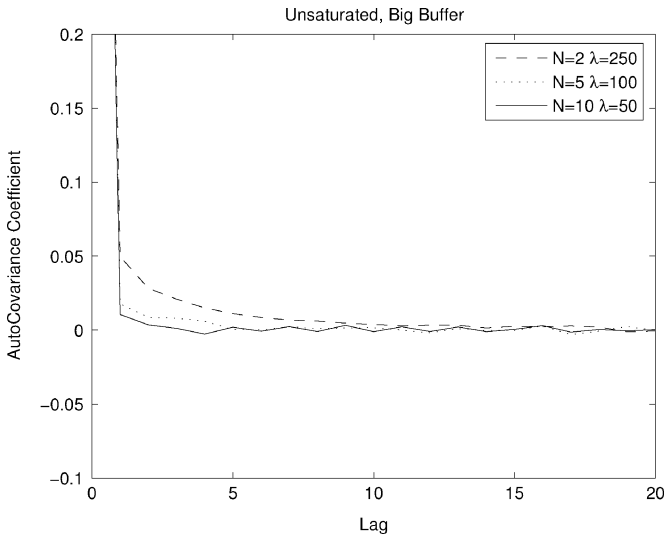

Fig. 16. Unsaturated, big-buffer, interdeparture-time-sequence normalized autocovariances. Experimental data.

from mesh points, as these form the arrivals processes to other parts of the mesh.

Let $\{D_k\}$ denote the interdeparture times of packets from an element of the network. That is, D_k is the difference between the time at which the kth successful transmission and the $(k-1)$th successful transmission occurs from a tagged station. One hypothesis (e.g., [21]) is that if the arrivals process to the station is Poisson, then the departure process is also Poisson. That is: (A7) $\{D_k\}$ is a stochastically independent sequence; and (A8) the elements of $\{D_k\}$ are exponentially distributed.

Having observed $K^{(D)}$ interdeparture times, $D_1, \ldots, D_{K^{(D)}}$, we investigate these hypotheses. These times were recorded in the same experiment as the collision data used in Section IV, so $K^{(D)} = K^{(C)}$ where $K^{(C)}$ is reported in Table II. Figs. 14–16 report the autocovariance for saturated arrivals, unsaturated arrivals with small buffers, and unsaturated arrivals with large station buffers for networks of $N = 2$, 5, and 10 stations, respectively. There is little dependency beyond short lags, suggesting that the independence hypothesis (A7) is not inappropriate.

For hypothesis (A8) and unsaturated stations with large buffers, due to space constraints, we only report the interdeparture time distributions for the $N = 5$ network. These are representative of our observations of other networks sizes. Fig. 17 plots the logarithm of one minus the empirical cumulative distribution function of the interdeparture times from the tagged station. Also plotted is the exponential distribution corresponding to the empirical mean. It can be seen that they overlay each other nearly perfectly, suggesting that (A8) is a good hypothesis in this case. Note that this implies that in the lightly loaded, unsaturated big-buffer setting, the statistics of a Poisson arrivals process is largely unaffected when passing through an 802.11 network element.

Fig. 18 reports the equivalent quantity for the small buffer experiment with $N = 2$, but here the network arrival rate of 800

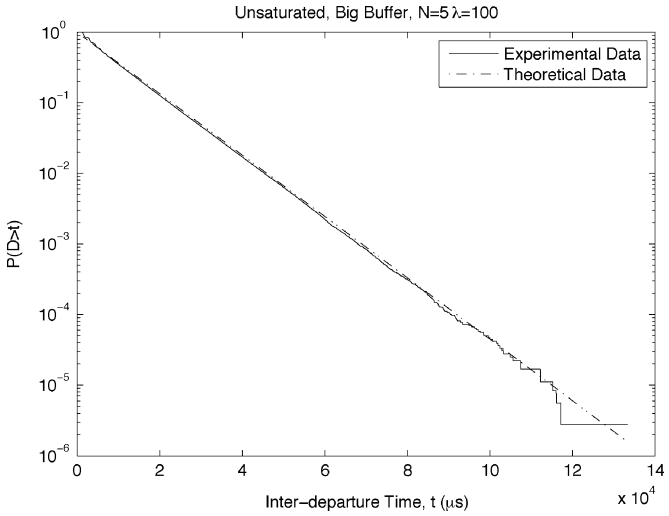

Fig. 17. Unsaturated big-buffer interdeparture time distribution (log y-scale), $N = 5$. Experimental data.

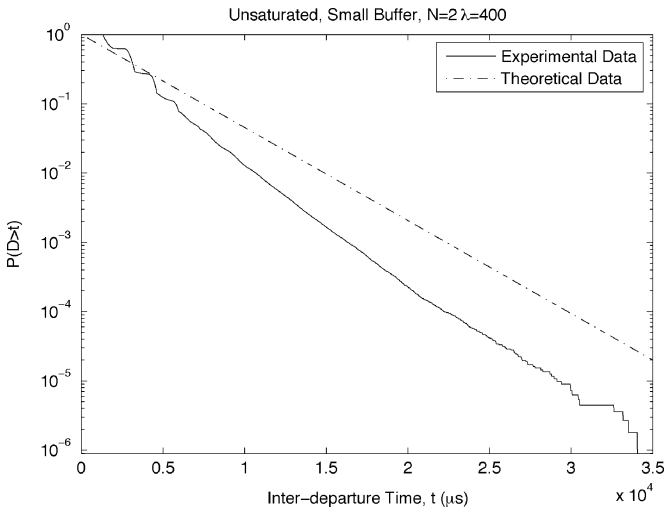

Fig. 18. Unsaturated small-buffer interdeparture time distribution (log y-scale), $N = 2$. Experimental data.

packets per second is chosen to be in the regime near saturation. For $N = 2$, after the successful transmission of a packet, 56% of the time there is another packet awaiting transmission by the time the medium is sensed idle. The effective transition between underloaded and saturated can be seen as short interdeparture times possess the features of MAC service times (convolutions of uniform distributions), whereas longer interdeparture times follow an exponential distribution. With a larger number of stations and the same offered load shared evenly across stations, this effect is less pronounced, and the interdeparture times look exponentially distributed. This effect is independent of the buffering used and can also be observed within big buffer experiments when traffic loads are closer to saturation. Thus, again the evidence supports the assumption (A8) assumption if the network is away from saturation.

However, Fig. 19 reports the same plot for a saturated network. Clearly, the interdeparture times are not exponentially distributed. This is unsurprising as when stations are saturated, the

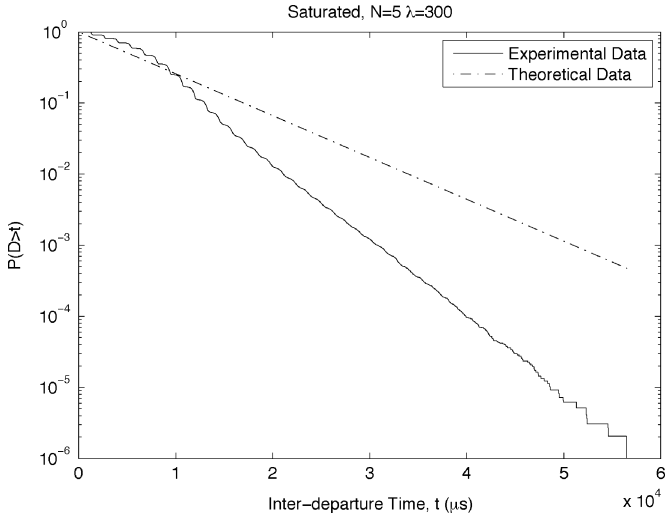

Fig. 19. Saturated interdeparture time distribution (log y-scale), $N = 5$. Experimental data.

interdeparture times correspond to the MAC service times and BEB service times are not well approximated by an exponential distribution. Qualitatively, the $N = 2$ and $N = 10$ networks show the same features, where the (A8) assumption is appropriate for lightly loaded unsaturated networks, but inaccurate for saturated networks. This can be statistically substantiated through the use of a Kolmogorov–Smirnoff test. The null hypothesis that the interdeparture times are exponentially distributed is rejected unless one trims the data by conditioning solely large interdeparture times or if the load is sufficiently light.

VIII. Discussion

Table I summarizes our conclusions. It seems appropriate at this stage to discuss another fundamental assumption: (A0) all stations in the WLAN observe the same sequence of busy and idle slots on the medium. This assumption is a cornerstone of all CSMA/CA models that allow idle slots to be of distinct real-time length from collisions and successful transmissions, as is the case in 802.11 networks, and include collisions in their considerations. Both p-persistent models and all of the mean-field models described here are based on this premise, which is true in the absence of hidden nodes and interfering neighboring WLANs.

In order to model situations where (A0) is false, such as relay topologies that do not have multiple radios and so cannot mitigate interference at noncommunicating distances, new approximations are necessary (e.g., [29]–[32]). Some of these models also use mean-field ideas, usually inspired by percolation theory, and typically assume away the BEB aspect of 802.11. This approach gives insight into scenarios that are distinct from those that are amenable to analysis by the models studied in the present paper, and, as suggested to us by an anonymous referee, the appropriateness of these multihop mean-field approximations would be deserving of an experimental investigation as they facilitate mathematical study of situations that are otherwise analytically intractable. We do not pursue this in the present article due to space constraints.

This validation of the standard decoupling assumptions, (A1) and (A2), for saturated networks helps to explain why the predictions in [2]–[5] are so precise. Even though intuitively one expects the main model assumptions to be valid for large networks, in fact they are accurate even for small networks. As the assumptions are reasonable, deductions from that model should be able to make predictions regarding detailed quality-of-service metrics.

The (A1) assumption continues to hold for both the unsaturated setting with either small or big buffers, suggesting that the attempt sequences have little dependencies. With small buffers, the (A2) assumption that collision probabilities are independent of backoff stage appears to be valid for stations that are not saturated. There is, however, some structure with $p_1 > p_0$, but this is not quantitatively significant. For larger buffers, this discrepancy is more apparent in both relative and absolute terms, suggesting that (A2) is an imprecise approximation in that setting. For large buffer models, this inaccuracy is less dramatic than the failure of the additional queueing decoupling assumption (A3) and (A4).

Our investigations indicate that while (A3) is reasonable at lighter loads, neither (A3) or (A4) are appropriate in general. In particular, contradicting (A4), the probability that the queue is nonempty after a successful transmission is strongly dependent on backoff stage. Despite the apparent inappropriateness of the assumptions (A3) and (A4), models based on them continue to make accurate predictions of goodput and average delay. One explanation is that with an infinite buffer, unless the station is saturated, the goodput corresponds with the offered load. Thus, to have an accurate goodput model, it is only important that the model be accurate when offered load leads a station to be nearly saturated. Once saturated, (A4) is true as Bianchi's model is recovered. Thus, for goodput, the inaccuracy of the approximation (A4) is not significant if this phase transition is predicted by the model. However, one would expect that for more subtle quantities the adoption of (A4) would lead to erroneous deductions. Clearly, caution must be taken when making deductions from big buffer models that incorporate these hypotheses. Extrapolations of that kind from these models should be made with care by network designers.

The hold-state hypotheses that were introduced to incorporate 802.11e AIFS differentiation in saturated networks are labeled (A5) and (A6). The independence hypothesis (A5) appears to be appropriate. The distributional assumption (A6) appears to be accurate for any difference in AIFS once one has a good estimate of the probability that no higher class station attempts transmission in a typical slot. This lends confidence to the use of these models for network design and detailed predictions.

The 802.11s mesh network assumptions (A7) and (A8) hold true for lightly loaded, unsaturated networks where stations can have either large or small buffers. In particular, the output of an unsaturated 802.11 station with Poisson arrivals again appears to be nearly Poisson so long as saturation is not being approached. If the station is saturated, however, the interdeparture times correspond to MAC delays, which are not similar to an exponential distribution. If stations are close to being saturated, short interdeparture times are similar to MAC delays, whereas long interdeparture times correspond to long interarrival times and are Poissonian. The impact of this non-Poisson traffic on the accuracy of unsaturated model predictions must be investigated before they can be used with confidence.

We also make a comment regarding experimentation. It was challenging to emulate the fundamental explicit hypothesis of all the models that we investigated: that of idealized channel conditions where errors occur only as a consequence of collisions. As 802.11 operates in an unlicensed range of the spectrum and other devices are free to operate in this range, these devices lead to interference. There are extensions to the WLAN modeling paradigms that include failed transmissions due to noise on the medium, e.g., [33]. This approach assumes that packet losses due to noise are i.i.d. and independent of all other stochastic processes in the model. Whether this assumption is appropriate is dependent on the particular environment at hand and, clearly, cannot be subject to general validation.

Due to the failure of several of these fundamental hypotheses, clearly there is more work to be done on analytic modeling of 802.11. In particular, models that incorporate buffers at stations are based on flawed hypotheses, but are important for network designers. We suggest that it is an important challenge for the analytic modeling community to revisit and revise models based on these inappropriate assumptions.

Based on the observations in this article, for example, a natural alternative to (A3) and (A4) is to use the approximation that: (A3′) given $\beta_k = i$, $\{Q_k\}$ is an independent sequence; and (A4′) given $\beta_k = i$, $P(Q_k = 1) = q_i$. As a first step in this direction, in [34], negative consequences of adopting the assumptions (A3) and (A4) are identified. A typical validation scenario employed by modelers is to consider a symmetrically loaded network. While this is unlikely to occur in practice, mathematically it leads to homogeneous fixed-point equations whose solutions can be quickly identified by standard numerical techniques. For stations that are asymmetrically loaded, results in [34] demonstrate that a model based on these assumptions provides inaccurate throughput predictions. That this is a consequence of (A3) and (A4) is established by considering the setting where all stations can buffer one packet beyond the MAC, as it is then possible to analyze a model based on (A3′) and (A4′).

Finally, we expect that other researchers will have alternate hypotheses that they wish to check. To facilitate this research, all the data used in the present study is available at: http://www.hamilton.ie/kaidi/.

APPENDIX I
BRIEF OVERVIEW OF 802.11'S BEB ALGORITHM

On detecting the wireless medium to be idle for a period DIFS, each station initializes a counter to a random number selected uniformly in the range $\{0, 1, \ldots, \mathrm{CW}_{\min} - 1\}$. Time is slotted, and this counter is decremented once during each slot that the medium is observed idle. The countdown halts when the medium becomes busy and resumes after the medium is idle again for a period DIFS. Once the counter reaches zero, the station attempts transmission, and if a collision does not occur, it can transmit for a duration up to a maximum period TXOP (defined to be one packet except in the Quality-of-Service MAC ex-

tension 802.11e). If two or more stations attempt to transmit simultaneously, a collision occurs. Colliding stations double their contention window (CW) (up to a maximum value), select a new backoff counter uniformly, and the process repeats. If a packet experiences more collisions than the retry limit (11 in 802.11b), the packet is discarded. After the successful transmission of a packet or after a packet discard, CW is reset to its minimal value $\mathrm{CW}_{\min}$, and a new countdown starts regardless of the presence of a packet at the MAC. If a packet arrives at the MAC after the countdown is completed, the station senses the medium. If the medium is idle, the station attempts transmission immediately; if it is busy, another backoff counter is chosen from the minimum interval. This bandwidth saving feature is called post-backoff. The revised 802.11e MAC enables the values of DIFS (called AIFS in 802.11e), $\mathrm{CW}_{\min}$, and TXOP to be set on a per-class basis for each station. That is, traffic is directed to up to four different queues at each station, with each queue assigned different MAC parameter values.

Appendix II
Testing Goodness of Fit

Given a sequence of observations of independent and identically distributed random variables $X_1, \ldots, X_n$, we wish to test the hypothesis that the $\{X_k\}$ have common distribution F. For discrete-valued random variables taking N distinct values, we use Pearson's χ^2-test (e.g., [35]) Assume a null hypothesis that X_1 has a distribution such that $P(X_1 = i) = f(\vec{\theta}, i)$, where $\vec{\theta}$ is a collection of M parameters estimated from the data. For each possible outcome i, define

$$n_i := \sum_{j=1}^{n} 1_{\{X_j = i\}} \quad \text{and} \quad X^2 := \sum_{i=1}^{N} \frac{(n_i - nf(\vec{\theta}, i))^2}{nf(\vec{\theta}, i)}.$$

For large sample sizes n, the test statistic X^2 has a χ^2 distribution with between $N-1$ and $N-1-M$ degrees of freedom. We use the later, more stringent test to determine the p-value $P(\chi^2(N-1-M) \geq X^2)$.

For real-valued random variables and unbounded random variables with a discrete distribution, we evaluate the Kolmogorov–Smirnov statistic (e.g., [35] and [36]). We can use the former to determine a test for goodness of fit, but exact critical levels are not possible to determine for the latter [37]. Let

$$F_n(k) = \frac{1}{n} \sum_{i=1}^{n} 1_{\{X_i \leq k\}}$$

denote the empirical distribution given n observations. The L^∞ distance $\sup_k |F_n(k) - F(k)|$ is the greatest discrepancy between the two distributions. It is used in the Kolmogorov–Smirnov test based on the observation that if the null hypothesis that $\{X_k\}$ are identically distributed were true, then $\sup_k |F_n(k) - F(k)|$ is of order $n^{-1/2}$ and, in particular, $\sup_k |F_n(k) - F(k)| \to 0$. If F is continuous, we also have the following weak convergence result:

$$\sqrt{n} D_n \Rightarrow \sup_t |B(F(t))|$$

where $B(t)$ is a Brownian bridge [38] from which a p-value can be determined.

Appendix III
Runs Test for Binary Valued Random Variables

If $\{X_n\}$ are binary-valued random variables, then the null hypothesis that the sequence is independent and identically distributed can be efficiently tested using the Runs Test [39], [35]. Given a sequence of observations $X_1, \ldots, X_n$, a run is defined to be a maximal nonempty segment of the sequence consisting of adjacent equal elements. Let R be the number of runs in $X_1, \ldots, X_n$, and define

$$n_0 = \sum_{i=1}^{n} 1_{\{X_i = 0\}}$$
$$\mu = \frac{2n_0(n - n_0)}{n+1}$$
$$\sigma^2 = \frac{(\mu - 1)(\mu - 2)}{n-1}.$$

Then, under the null hypothesis, $Z = (R - \mu)/\sqrt{\sigma^2}$ is asymptotically normally distributed. Thus, given the sequence of observations $X_1, \ldots, X_n$, one evaluates Z, and the p-value for the null hypothesis is $\min(P(N(0,1) \geq Z), P(N(0,1) \leq Z))$, where $N(0,1)$ is a normally distributed random variable.

Acknowledgment

The authors thank the anonymous reviewers for their helpful comments and suggestions.

References

[1] K. D. Huang, K. R. Duffy, D. Malone, and D. J. Leith, "Investigating the validity of IEEE 802.11 MAC modeling hypotheses," in *Proc. IEEE PIMRC*, Cannes, France, 2008, pp. 1–6.

[2] G. Bianchi, "IEEE 802.11—Saturated throughput analysis," *IEEE Commun. Lett.*, vol. 2, no. 12, pp. 318–320, Dec. 1998.

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

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

[5] F. Cali, M. Conti, and E. Gregori, "IEEE 802.11 protocol: Design and performance evaluation of an adaptive backoff mechanism," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 9, pp. 1774–1786, Sep. 2000.

[6] G.-S. Ahn, A. T. Campbell, A. Veres, and L.-H. Sun, "Supporting service differentiation for real-time and best-effort traffic in stateless wireless ad hoc networks (SWAN)," *IEEE Trans. Mobile Comput.*, vol. 1, no. 3, pp. 192–207, Jul.–Sep. 2002.

[7] M. Ergen and P. Varaiya, "Throughput analysis and admission control in IEEE 802.11a," *Mobile Netw. Appl.*, vol. 10, no. 5, pp. 705–716, 2005.

[8] J. Hui and M. Devetsikiotis, "A unified model for the performance analysis of IEEE 802.11e EDCA," *IEEE Trans. Commun.*, vol. 53, no. 9, pp. 1498–1510, Sep. 2005.

[9] K. Duffy, D. Malone, and D. J. Leith, "Modeling the 802.11 distributed coordination function in non-saturated conditions," *IEEE Commun. Lett.*, vol. 9, no. 8, pp. 715–717, Aug. 2005.

[10] D. Malone, K. Duffy, and D. J. Leith, "Modeling the 802.11 distributed coordination function in non-saturated heterogeneous conditions," *IEEE/ACM Trans. Netw.*, vol. 15, no. 1, pp. 159–172, Feb. 2007.

[11] H. Zhai, Y. Kwon, and Y. Fang, "Performance analysis of IEEE 802.11 MAC protocols in wireless LANs," *Wireless Commun. Mobile Comput.*, vol. 4, no. 8, pp. 917–931, 2004.

[12] G. R. Cantieni, Q. Ni, C. Barakat, and T. Turletti, "Performance analysis under finite load and improvements for multirate 802.11," *Comput. Commun.*, vol. 28, no. 10, pp. 1095–1109, 2005.
[13] C. G. Park, D. H. Han, and S. J. Ahn, "Performance analysis of MAC layer protocols in the IEEE 802.11 wireless LAN," *Telecommun. Syst.*, vol. 33, no. 1–3, pp. 233–253, 2006.
[14] K. Duffy and A. J. Ganesh, "Modeling the impact of buffering on 802.11," *IEEE Commun. Lett.*, vol. 11, no. 2, pp. 219–221, Feb. 2007.
[15] B. Li and R. Battiti, "Supporting service differentiation with enhancements of the IEEE 802.11 MAC protocol: Models and analysis," University of Trento, Trento, Italy, Tech. Rep. DIT-03-024, May 2003.
[16] J. W. Robinson and T. S. Randhawa, "Saturation throughput analysis of IEEE 802.11e enhanced distributed coordination function," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 5, pp. 917–928, Jun. 2004.
[17] Z. Kong, D. H. K. Tsang, B. Bensaou, and D. Gao, "Performance analysis of IEEE 802.11e contention-based channel access," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 10, pp. 2095–2106, Dec. 2004.
[18] P. E. Engelstad and O. N. Østerbø, "Non-saturation and saturation analysis of IEEE 802.11e EDCA with starvation prediction," in *Proc. MSWiM*, 2005, pp. 224–233.
[19] P. Clifford, K. Duffy, J. Foy, D. J. Leith, and D. Malone, "Modeling 802.11e for data traffic parameter design," in *Proc. RAWNET*, 2006, pp. 1–10.
[20] X. Chen, H. Zhai, X. Tian, and Y. Fang, "Supporting QoS in IEEE 802.11e wireless LANs," *IEEE Trans. Wireless Commun.*, vol. 5, no. 8, pp. 2217–2227, Aug. 2006.
[21] K. Duffy, D. J. Leith, T. Li, and D. Malone, "Modeling 802.11 mesh networks," *IEEE Commun. Lett.*, vol. 10, no. 8, pp. 635–637, Aug. 2006.
[22] D. P. Bertsekas and R. G. Gallager, *Data Networks*, 2nd ed. Upper Saddle River, NJ: Prentice-Hall, 1991.
[23] G. Bianchi, A. Di Stefano, C. Gianconia, and L. Scalia, "Experimental assessment of the backoff behavior of commercial IEEE 802.11 b network," in *Proc. IEEE INFOCOM*, 2007, pp. 1181–1189.
[24] D. Giustiniano, G. Bianchi, L. Scalia, and I. Tinnirello, "An explanation for unexpected 802.11 outdoor link-level measurement results," in *Proc. IEEE INFOCOM*, 2008, pp. 2432–2440.
[25] *IEEE. Wirless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, IEEE Std 802.11-1997, 1997.
[26] W. Hoeffding, "Probability inequalities for sums of bounded random variables," *J. Amer. Statist. Assoc.*, vol. 58, pp. 13–30, 1963.
[27] S. Asmussen, *Applied Probability and Queues*. New York: Springer-Verlag, 2003, vol. 51.
[28] *IEEE. Wirless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Medium Access Control (MAC) Enhancements for Quality of Service (QoS)*, IEEE Std 802.11e, 2005, edition.
[29] G. Hauksson and M. Alanyali, "Wireless medium access via adaptive backoff: Delay and loss minimization," in *Proc. IEEE INFOCOM*, 2008, pp. 1777–1785.
[30] O. Dousse, M. Durvy, and P. Thiran, "Self-organization properties of CSMA/CA systems and their consequences on fairness," *IEEE Trans. Inf. Theory*, vol. 55, no. 3, pp. 931–943, Mar., 2009.
[31] O. Dousse, P. Thiran, and M. Durvy, "On the fairness of large CSMA networks," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 7, pp. 1093–1104, Sep. 2009.
[32] C. Bordenave, S. Foss, and V. Shneer, "A random multiple access protocol with spatial interactions," *J. Appl. Probab.*, vol. 46, pp. 844–865, 2009.
[33] Q. Ni, T. Li, T. Turletti, and Y. Xiao, "Saturation throughput analysis of error-prone 802.11 wireless networks," *J. Wireless Commun. Mobile Comput.*, vol. 5, no. 8, pp. 945–957, 2005.
[34] K. D. Huang and K. R. Duffy, "On a buffering hypothesis in 802.11 analytic models," *IEEE Commun. Lett.*, vol. 13, no. 5, pp. 312–314, May 2009.
[35] D. O. Wackerly, W. Mendenhall, and R. L. Scheaffer, *Mathematical Statistics With Applications*. Pacific Grove, CA: Duxbury, 2001, vol. 6.
[36] A. N. Pettitt and M. A. Stephens, "The Kolmogorov–Smirnov goodness-of-fit statistic with discrete and grouped data," *Metrika*, vol. 19, no. 2, pp. 205–210, 1977.
[37] W. J. Conover, "A Kolmogorov goodness-of-fit test for discontinuous distributions," *J. Amer. Statist. Assoc.*, vol. 67, pp. 591–596, 1972.
[38] P. Billingsley, *Convergence of Probability Measures*, ser. Probability and Statistics, 2nd ed. New York: Wiley, 1999.
[39] A. M. Mood, "The distribution theory of runs," *Ann. Math. Statist.*, vol. 11, pp. 367–392, 1940.

Kaidi Huang received the B.E. degree in optical communication and engineering and the B.A. degree in law science from the University of Electronic and Scientific Technology of China, Chengdu, China, in 2006, and the M.S. degree with distinction in signal processing and communications from the University of Edinburgh, Edinburgh, U.K., in 2007. He is currently a Ph.D. student at Hamilton Institute, National University of Ireland, Maynooth, Maynooth, Ireland. His research interests include both wireless and optical communication systems.

Ken R. Duffy received the B.A. (mod) and Ph.D. degrees in mathematical science from Trinity College, Dublin, Ireland, in 1996 and 2000, respectively.

Since 2006, he has been a member of faculty with the Hamilton Institute of the National University of Ireland, Maynooth, Maynooth Ireland. His wide-ranging research interests encompass applied probability, mathematical statistics, theoretical immunology, and mathematical applications to stochastic models of communication networks and queueing systems.

David Malone received the B.A. (mod), M.Sc., and Ph.D. degrees in mathematics from Trinity College, Dublin, Ireland, in 1996, 1997, and 2000, respectively.

During his time as a post-graduate, he became a member of the FreeBSD development team. He is a Science Foundation Ireland Stokes Lecturer with Hamilton Institute, National University of Ireland, Maynooth, Maynooth, Ireland. His interests include mathematics of networks, network measurement, IPv6, and systems administration. He is a coauthor of *IPv6 Network Administration* (Sebastopol, CA: O'Reilly, 2005).