36TH ANNUAL ALLERTON CONFERENCE ON COMMUNICATIONS, CONTROL AND COMPUTING – ALLERTON HOUSE, MONTICELLO, IL, SEPT. 23–25, 1998

On channel modeling for delay analysis of packet communications over wireless links

Michele Zorzi and Ramesh R. Rao

Center for Wireless Communications, University of California, San Diego
9500 Gilman Drive, La Jolla, CA 92093-0407
e-mail: {zorzi,rao}@ece.ucsd.edu

Abstract

Higher order characteristics of channel errors are known to have a significant impact on protocol performance. Two state Markov models, which have been widely used to model bursty communication channels may be inadequate for the representation of some time-varying channels. Although a larger number of states will almost surely provide a better representation of the channel, the complexity of the model makes the analysis of higher layer protocols difficult. Furthermore, the quality of a particular channel model has much to do with the nature of the metric being computed.

Although a two state Markov channel model leads to throughput predictions of protocols such as ARQ that agree quite well with simulations over Rayleigh fading channels, such a model is found to be inadequate for predicting delay in an ARQ/queueing system that operates over a fading channel. On the other hand, good matches with simulation results are obtained by modeling the error process by means of a Markov model with one good state and two bad states. The model with two bad states results in a richer, non geometric fade duration and can consequently better model the effect of fading from the perspective of delay. The effect of model parameter values on the goodness of the approximation is also discussed.

1 Introduction

For some time now, we have been studying the effect of wireless channel errors on the performance of protocols [1]–[6]. These studies have shown that higher order characteristics of channel errors have a significant impact on performance and that correlations in channel errors can be exploited to reduce energy consumption. These findings suggest that proper modeling of the channel errors is key in being able to accurately predict the performance of a protocol, and motivated us to reexamine the channel models we have used in the past. Markov models have been widely used to model bursty communication channels [7], from the early work of Gilbert [8] to more recent papers studying their applicability to fading channels [1], [9]–[15].

Although more general than the iid (memoryless) error model, the 2-state channel model [1] may be inadequate for the representation of some time-varying channels. One way to overcome this problem is to enlarge the number of states [11]–[15]. Fritchman [16] investigated a finite-state Markov chain model with more than two states. The state space was then partitioned into two groups, corresponding to error-free states and error states. Fritchman's model has been applied by many researchers to represent error sequences obtained over the fading channels. Although a large number of states will almost surely provide a better representation of the channel, the complexity of the model makes subsequent performance analysis of higher layer protocols increasingly difficult.

The quest to develop models that adequately represent real channels while being mathematically tractable has taken several directions. In [17], the authors explored conditions under which an N-state Markov channel model can be reduced to a 2-state model and ways to stochastically upper and lower bound the original channel model with a 2-state Markovian model. In comparing the quality of the bounds that result it became apparent that the goodness of a particular bounding or modeling technique has much to do with the nature of the metric being computed.

In particular, while studying the throughput of protocols such as ARQ [2] and TCP [18], Markov channel models with two states (one good and one bad), lead to throughput predictions that agree quite well with simulations over fading channels. Is such a model adequate for studying other performance metrics, such as delay, in a model that includes queues?

In a two state model, the duration of the good and bad periods are geometrically distributed. As a result, the distribution of the residual life (time to end of duration) of both bad and good periods will be memoryless and have the same distribution as their respective typical distributions. In fact, iid models also exhibit this property with the additional constraint that the parameters of the two geometric distributions are complements of each other. The difference between an iid model and a two state model is that in the latter case, the two geometric distributions can be parameterized by arbitrary parameters. One may anticipate that realistic channels are unlikely to produce bad and good periods that are memoryless and therefore one is forced to consider higher order chains.

In this paper we study a simple example that reinforces these observations. We consider a simple ARQ/queueing model, operating over a Rayleigh fading channel (as in the classic model described in [19, 20]) and show that good matches with simulation results can be obtained by modeling the error process by means of a Markov model with three states (one good and two bad). The additional flexibility introduced by the third state allows us to adequately predict the complementary distribution of the packet delay for the cases considered.

What appears to be happening over the fading channel is that packets awaiting transmission when the channel is in a deep fade see atypical durations of the residual life of the fade. The model with two bad states results in a richer, non geometric fade duration and can consequently better model the effect of fading from the perspective of delay. The effect of model parameter values on the goodness of the approximation is also discussed.

2 System Model

We assume a wireless link, subject to bursty errors, with a relatively high data rate, of the order of 1 Mbps. The time axis is slotted, and transmissions occur in blocks (packets). We focus on a single pair of communicating users. In each slot, packets are independently generated at the transmitter according to a Bernoulli arrival process with rate λ, and queued in a buffer awaiting transmission.

A simple ARQ scheme is used to counteract the effect of packet errors. A perfect and instantaneous feedback channel is assumed, so that at the end of each slot the transmitter knows exactly whether or not the transmission was successful[1]. Packets which were not successfully received are immediately retransmitted.

The performance metric considered in this study is the *lateness probability*, i.e., the probability that a packet is not successfully delivered within D slots of its arrival. In the presence of a smoothing buffer at the receiver, this is the probability that the jitter experienced by a packet is too large to be absorbed by the buffer itself.

3 The channel error process

The packet error process is derived from Clarke's model for the Rayleigh fading channel [19, 20]. It is assumed here for simplicity that during a packet duration the value of the fading envelope does not change significantly[2]. Furthermore, if the value of the envelope is above a certain threshold, the packet is assumed to be correctly decoded with probability 1, and incorrectly otherwise [1]. Long-term propagation effects (path loss and shadowing) are usually compensated by power control and will be ignored here.

The key fading parameters in this context are the *fading margin*, F, and the *normalized Doppler frequency*, $f_D T$, where f_D is the maximum Doppler shift (mobile speed divided by the carrier wavelength) and T is the duration of a packet (slot). The fading margin is the maximum fading attenuation which still allows correct reception, i.e., $F = 10$ dB means that a packet transmission is successful as long as the value of the fading envelope does not drop more than 10 dB below its mean. The normalized Doppler frequency is directly related to the correlation properties of the error process.

The two-state Markov model can be obtained from Clarke's model by matching the average packet error rate and the conditional probability of a success given a failure, which can be either measured from a simulation or computed following Clarke's analytical model [20].

The computation of the lateness probability for the simple system described above and for the two-state error model has been performed analytically in [3]. However, when comparing the analytical results based on the Markov approximation for the Rayleigh fading channel with the performance obtained from its true simulation, significant differences are observed. As shown in Figure 1, the analytical prediction may be accurate for small values of the delay parameter (which is the less interesting case), whereas the match is much worse in the more interesting right-hand side of the graph. This behavior can be explained by looking at the actual statistics of the length of bursts of good and bad slots. For example, Figure 2 shows the complementary probability distribution of the bad burst length for $f_D T = 0.01$ and some typical values of F. On the scale used, a two-state Markov model would result in straight lines, which is obviously

[1] Extensions to erroneous feedback and finite round-trip are discussed in [3].

[2] This assumption is well verified for all values of the Doppler frequency and data rates considered in this study.

not the case here. In fact, the curves appear to exhibit a two-slope behavior, which cannot be correctly captured by a two-state approximation.

The reason why throughput predictions are accurate is that in evaluating throughput the time span of important events does not extend much beyond the round-trip delay. That is, fitting the channel error process with a Markov process which correctly approximates the bad burst lengths up to just a few slots will provide good results simply because the region where the model is not accurate plays a negligible role in determining the performance. On the other hand, in delay and queueing analysis, in which longer time scales are to be considered, such differences between error model and true behavior may lead to significantly inaccurate performance evaluations.

A look at the statistics of the good burst duration (plotted in Figure 3), on the other hand, reveals that a geometric model for the burst length may be good enough. Evidence in support of this conjecture is given by Figure 4, which reports results of a simulation where the true statistics was used for the bad burst duration, whereas a geometric approximation was used for the good periods. It can be see that the source of error in using a two-state Markov approximation lies in the inaccuracy of the geometric model for the error bad bursts.

4 Higher-order models for the fading channel

The above results demonstrate the need for higher-order error models in applications which require consideration of a larger time scale. This problem can obviously be addressed by increasing the size of the state space of the Markov chain beyond two. However, the beauty of the two-state Markov model does not lie in its accuracy (which is actually not very good in some cases, as just shown), but rather in the ease of handling it analytically when doing performance analysis of protocols running on top of such error processes. The strength of the two-state approach is then the fact that it enables analytical predictions, which make it possible to obtain both quick results and revealing insight.

Based on this point of view, it appears that the straightforward approach of increasing the state space of the chain until some prescribed accuracy of the match is achieved (according to some metric) does not satisfy our main goal, namely, *capturing the essential channel behavior with a minimal number of states.* Therefore, even though accurate channel models can be obtained as in, e.g., [13, 14, 15], we focus in this paper on a much smaller class of models.

Specifically, by again looking at the plots for the burst durations given in Figures 2 and 3, it appears that using just one good channel state may be adequate, whereas a two-slope approximation for the error burst duration would require two bad channel states. In general, we can approximate the complementary distribution of the error burst length, $P[\text{bad burst duration} > x]$, by fitting it with the following function

$$F(x) = \sum_{i=1}^{N} \alpha_i e^{-\beta_i x}, \tag{1}$$

with $\sum_{i=1}^{N} \alpha_i = 1$. Once the parameters of this function have been identified by appropriate fit with the simulation results, on can build a Markov chain with $N+1$ states which has $F(x)$

as the distribution of the bad burst length[3]. A possible choice[4] is a chain where state 0 is the (single) good state, states $1, \ldots, N$ are error states, and the transition matrix is given by

$$P = \begin{pmatrix} p_0 & \alpha_1 q_0 & \alpha_2 q_0 & \cdots & \alpha_N q_0 \\ 1-\beta_1 & \beta_1 & 0 & \cdots & 0 \\ 1-\beta_2 & 0 & \beta_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1-\beta_N & 0 & 0 & \cdots & \beta_N \end{pmatrix}, \tag{2}$$

where $p_0 = 1 - q_0$ can be found by matching the average error rate.

5 Computation of the lateness probability

The delay analysis for the system described in Section 2 has been presented in [3] for the two-state Markov channel model, and is based on the analytical evaluation of the steady-state distribution of the number of packets in the queue and on the recursive evaluation of the number of successful slots in a given time duration. Except for some analytical simplifications which lead to closed-form expressions, the same approach can be applied in this case. In fact, the recursion can be rewritten to account for the increased number of states, whereas the derivation of the steady-state queue length distribution can be obtained in general via matrix-geometric techniques [21], or even in closed form for small N (we will actually focus on $N = 2$ in this study). The interested reader is referred to [3, 21, 22] for more details.

6 Results and discussion

In this section, we present some numerical results obtained based on a three-state Markov approximation for the packet error process induced by the Rayleigh fading. This Markov model has been obtained by fitting the burst length as explained in Section 4, and has been used in the analytical evaluation of the lateness probability. These results are compared to the performance obtained by running the ARQ protocol over the fading channel according to Jakes' simulator [19].

Figure 5 shows the lateness probability for various values of the fading margin and for $f_D T = 0.01$. This corresponds, for example, to the transmission of packets of 1500 bits at a rate of 1 Mbps with carrier frequency 1800 MHz and pedestrian speed. Each plot shows curves for two values of the arrival rate, namely $\lambda = 0.5$ and $\lambda = 0.9$ (other values of λ show similar behavior). For each value of λ, three curves are plotted, i.e., the analytical results for two-state (solid lines) and three-state (dashed lines) Markov models, and the simulated performance (symbols). It is clear that the three-state model provides adequate accuracy, whereas the two-state model may be exceedingly optimistic.

[3]We do not address here in detail the issue of the independence of the dwelling times in different states which is inherent in the Markov characterization.

[4]Since a Markov chain with $N + 1$ states has $N(N + 1)$ independent entries in the transition matrix, whereas the constraints are the parameters of $F(x)$ and the average packet error rate (giving $2N$ equations), it is clear that for $N > 1$ there exists an infinite number of chains which result in the same error statistics.

Other values of the fading rate are considered in Figure 6, where results are plotted for $F = 25$ dB and for $f_DT = 0.002, 0.05$ (to be compared with $f_DT = 0.01$ in Figure 5c). It is seen that the accuracy of the approximation is still satisfactory in most cases, although for some values of the parameters this may not be true, as shown in Figure 7 for $F = 15$ dB and fast fading ($f_DT = 0.05$). In the latter case, the true performance curves change slope twice, and after some value of the delay, the approximate results have the wrong slope and become less and less accurate. Based on what we observed in passing from two to three states, we conjecture that adding another error state may provide the additional degree of freedom necessary to produce that second change of slope. Investigation of this conjecture is left for future study.

We stress the fact that the analytical results presented approximate the simulated performance very closely, which is remarkable given the extreme simplicity of the error model compared to the original Rayleigh fading process. It has been reported in the literature that accurate Markov modeling can be achieved if tens or even hundred of states are used. Here we are getting very good results by using only three!

In order to verify the generality of our findings, we also give another example of application, in which a more complicated protocol is used. Again under the assumption of instantaneous and error-free feedback, we consider an idealized Transport Layer protocol, which can be taken as a representation of TCP. The window adaptation algorithm and the loss recovery mechanism with fast retransmit are as in TCP Tahoe [23]. The protocol is run over a Radio Link Protocol (RLP) which recovers packet errors through retransmission of the erroneous TCP packets. A single connection is assumed, and the TCP packet length is taken to be constant and equal to the radio link packet length.

The radio link protocol is only allowed to transmit a given packet for a finite number of times, after which it will discard the packet and ask the TCP layer to pass down the next one[5]. The discarded packet is therefore considered lost from the TCP viewpoint, and this situation will be handled according to the protocol rules (i.e., if the window size and the channel behavior cause K packets to be successfully received, fast retransmit will occur, otherwise timeout expiration will be waited for). The metric of interest is again the complementary distribution of the packet delay. Bernoulli arrivals are assumed.

Results for $F = 20$ dB and $f_DT = 0.01$ are shown in Figure 8. Again, results obtained by using the two-state and the three-state Markov models are compared with those for the true Rayleigh fading simulation[6]. As before, the two-state model leads to optimistic results, whereas the three-state model gives a better approximation.

From the results of this study, we can draw the following conclusions:

- the assessment of the goodness of an error model must take into account the metric under study, i.e., a model which gives very good approximation for one metric (e.g., throughput) may turn out to be poor for others (e.g., delay)
- even though the error process induced by Rayleigh fading has nontrivial statistical properties and the interaction between this process and the queueing behavior of an ARQ

[5]When asking another packet, the RLP does not signal to the TCP whether the previous one was successfully delivered or dropped due to too many retransmissions, thereby complying with layering requirements.

[6]Even for the Markov channel models, the performance shown has been obtained by simulation, since no analytical approach has been developed for the performance analysis of the TCP/RLP stack.

protocol or a TCP/RLP stack is involved, a surprisingly simple channel model has been found to provide very accurate predictions, making analytical computations possible in some cases

- complicated Markov models with many states, which certainly give a better approximation of the error process but may not give as much insight, are not necessarily to be preferred to simple (though a little less accurate) models

- the goodness of the three-state error model depends also on the parameters involved, as we found that for some values of F and $f_D T$ is better than for others; extension of the procedure to more states seems straightforward, and such a study is left as future work.

References

[1] M. Zorzi, R.R. Rao, L.B. Milstein, "On the accuracy of a first-order Markov model for data block transmission on fading channels," in *Proc. IEEE ICUPC'95*, pp. 211-215, Nov. 1995.

[2] M. Zorzi, R.R. Rao, L.B. Milstein, "ARQ error control on fading mobile radio channels," *IEEE Trans. Veh. Tech.*, vol. 46, pp. 445–455, May 1997.

[3] M. Zorzi, R.R. Rao, "ARQ Error Control for Delay-Constrained Communications on Short-Range Burst-Error Channels," in *Proc. VTC'97*, pp. 1528–32, May 1997. Also, under review for *IEEE Trans. Comm.*

[4] M. Zorzi, R.R. Rao, "On the Statistics of Block Errors in Bursty Channels," *IEEE Trans. Comm.*, vol. 45, pp. 660–667, Jun. 1997.

[5] M. Zorzi, R.R. Rao, "The Role of Error Correlations in the Design of Protocols for Packet Switched Services," in *Proc. 35th Annual Allerton Conference*, pp. 749–758, Sep. 1997.

[6] M. Zorzi, R.R. Rao, "Energy constrained error control for wireless channels," *IEEE Personal Communications Magazine*, vol. 4, pp. 27–33, Dec. 1997.

[7] W. Turin, *Performance analysis of digital transmission systems*, Computer Science Press, NY, 1990.

[8] E.N. Gilbert, "Capacity of a burst-noise channel", *Bell System Tech. Journal*, vol. 39, pp. 1253–1266, Sept. 1960.

[9] F. Swarts, H.C. Ferreira, "Markov characterization of channels with soft decision outputs," *IEEE Trans. Commun.*, vol. COM-41, pp. 678-682, May 1993.

[10] H.S. Wang, "On verifying the first-order Markovian assumption for a Rayleigh fading channel model," in *Proc. ICUPC'94*, pp. 160-164, San Diego, CA, Sep. 1994. Also in *IEEE Trans. Veh. Tech.*, vol. VT-45, pp. 353–357, May. 1996.

[11] H.S. Wang, N. Moayeri, "Finite-state Markov channel — a useful model for radio communication channels", *IEEE Trans. Veh. Technol.*, vol. VT-44, pp. 163-171, Feb. 1995.

[12] M. Zorzi, "Outage and Error Events in Bursty Channels," *IEEE Trans. Comm.*, vol. 46, pp. 349–356, Mar. 1998.

[13] Q. Zhang, S.A. Kassam, "Finite-state Markov model for Rayleigh fading channels," submitted for publication.

[14] W. Turin, R. Van Nobelen, "Hidden Markov modeling of fading channels," in *Proc. VTC98*, pp. 1234–38, May 1998.

[15] Y. Y. Kim, S.Q. Li, "Modeling fast fading channel dynamics for packet data performance analysis," in *Proc. INFOCOM98*, pp. 1292–1300, Apr. 1998.

[16] B. D. Fritchman, "A binary channel characterization using partitioned Markov chains," *IEEE Trans. on Info. Theory*, vol. IT-13, Apr. 1967.

[17] A. M. Chen, Ramesh R. Rao, "On Tractable Wireless Channel Models," in *Proc. PIMRC98*, Paper ID inv017, Boston, MA, September. 1998.

[18] M. Zorzi, A. Chockalingam, R.R. Rao, "Performance Analysis of TCP on Channels with Memory," submitted for publication.

[19] W.C. Jakes, Jr., *Microwave mobile communications*, New York: John Wiley & Sons, 1974.

[20] R.H. Clarke, "A statistical theory of mobile radio reception", *Bell System Tech. Journal*, vol. 47, pp. 957-1000, Jul. 1968.

[21] M.F. Neuts, *Matrix-geometric solutions in stochastic models: an algorithmic approach*, Baltimore: Johns Hopkins University Press, 1981.

[22] R.A. Howard, *Dynamic probabilistic systems*, John Wiley & Sons, 1971.

[23] W. R. Stevens, *TCP/IP Illustrated*, *Volume 1*, Addison Wesley, 1994.

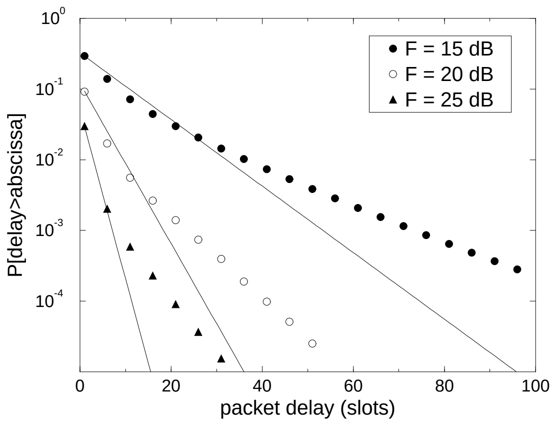


Figure 1: Delay performance of ARQ: comparison between two-state Markov approximation and simulated fading process, $F = 15, 20, 25$ dB, $f_D T = 0.01, \lambda = 0.9$.

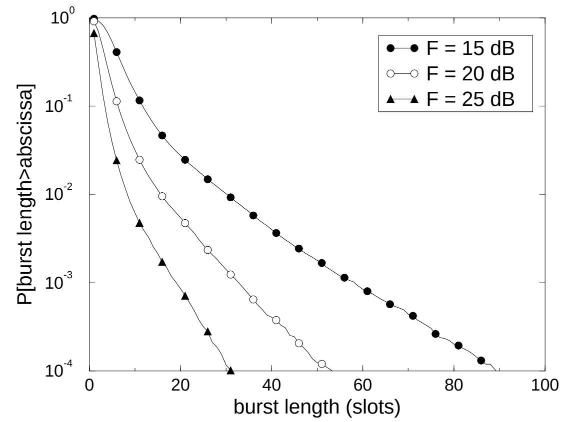


Figure 2: Bad burst length distribution (simulated) for $F = 15, 20, 25$ dB and $f_D T = 0.01$.

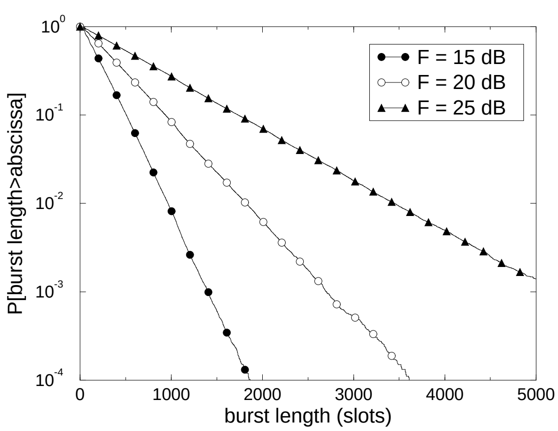


Figure 3: Good burst length distribution (simulated) for $F = 15, 20, 25$ dB and $f_D T = 0.01$.

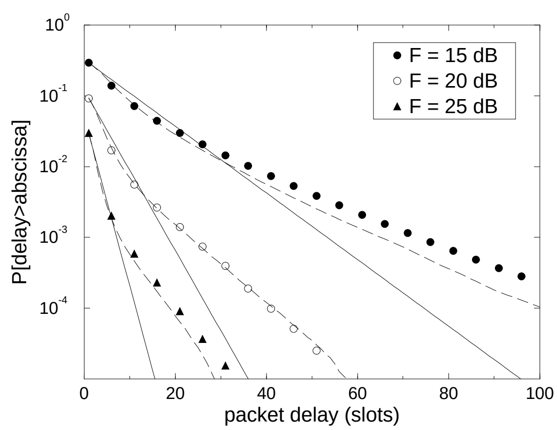
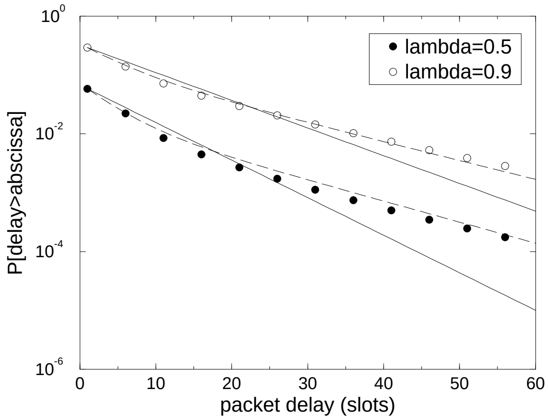
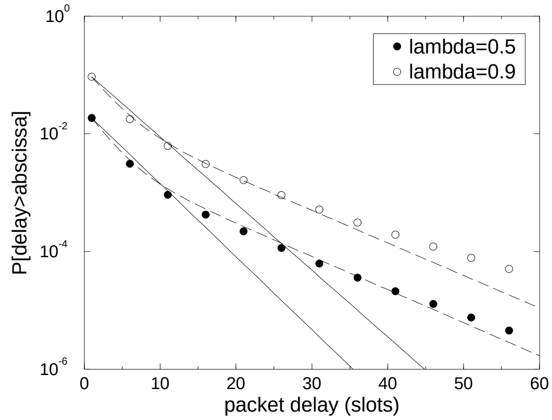


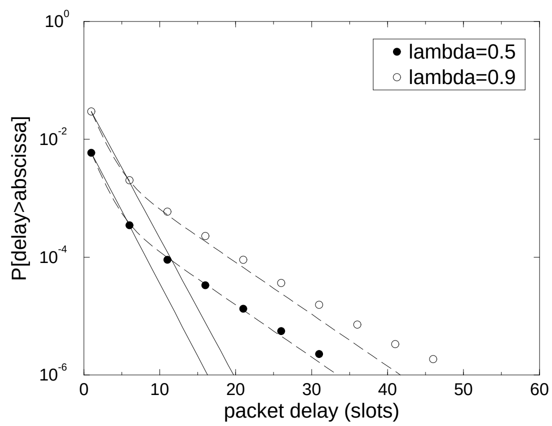
Figure 4: Delay performance of ARQ: comparison between two-state Markov approximation (solid lines), combination of exponential approximation of good bursts and true distribution of bad bursts (dashed lines), and simulated fading process (symbols), $F = 15, 20, 25$ dB, $f_D T = 0.01, \lambda = 0.9$.



(a)

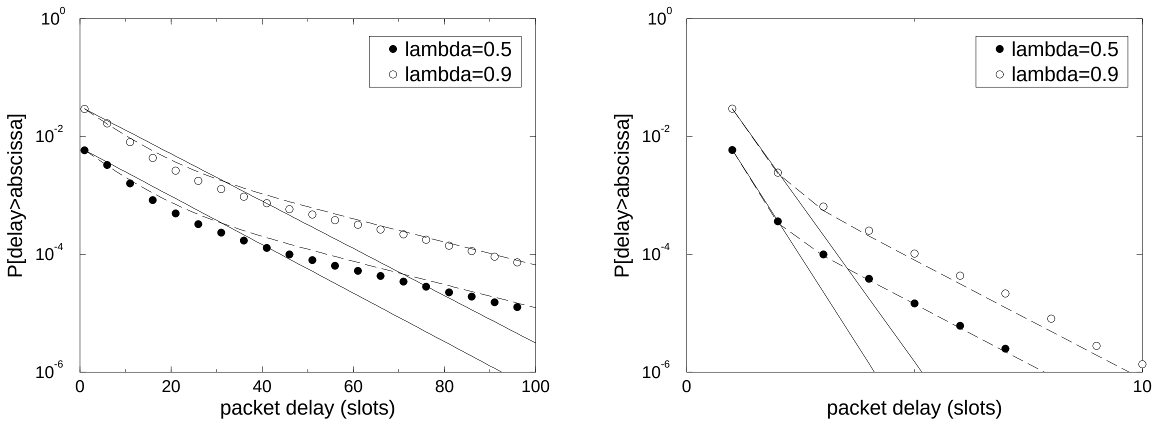


(b)



(c)

Figure 5: Delay performance of ARQ: comparison between two-state Markov approximation (solid lines), three-state Markov approximation (dashed lines) and simulation (symbols), $f_D T = 0.01$, $\lambda = 0.5$ and 0.9. (a) $F = 15$ dB, (b) $F = 20$ dB, (c) $F = 25$ dB.



(a) (b)

Figure 6: Delay performance of ARQ: comparison between two-state Markov approximation (solid lines), three-state Markov approximation (dashed lines) and simulation (symbols), $F = 25$ dB, $\lambda = 0.5$ and 0.9. (a) $f_D T = 0.002$, (b) $f_D T = 0.05$.

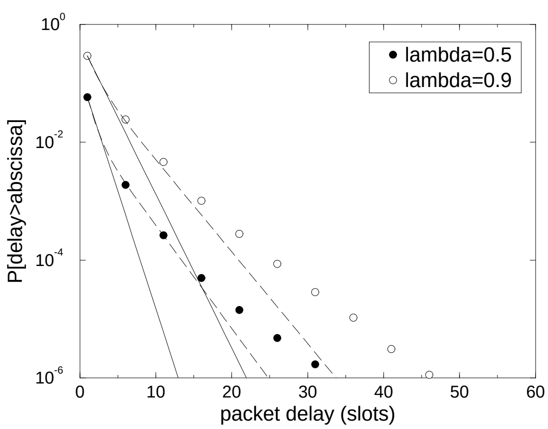


Figure 7: Delay performance of ARQ: comparison between two-state Markov approximation (solid lines), three-state Markov approximation (dashed lines) and simulation (symbols), $F = 15$ dB, $f_D T = 0.05$, $\lambda = 0.5$ and 0.9.

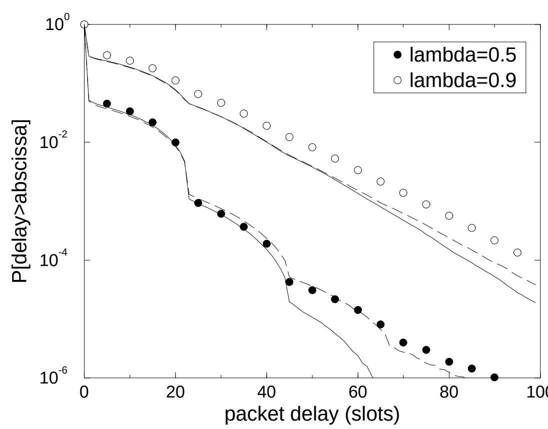


Figure 8: Delay performance of TCP/RLP: comparison between two-state Markov approximation (solid lines), three-state Markov approximation (dashed lines) and simulated fading (symbols), $f_D T = 0.01$, $F = 20$ dB, $\lambda = 0.5$ and 0.9.