Elsevier Editorial System(tm) for Journal of the Franklin Institute
Manuscript Draft

Manuscript Number: FI-D-09-00471

Title: On the Asymptotic Queueing Behavior of General AQM Routers

Article Type: Modeling & Simulation in Advanced Comm.

Keywords: AQM, Queueing theory

Corresponding Author: Dr. Dongyu Qiu,

Corresponding Author's Institution: Concordia University

First Author: Dongyu Qiu

Order of Authors: Dongyu Qiu; Ness Shroff

On the Asymptotic Queueing Behavior of General AQM Routers

Dongyu Qiu
ECE, Concordia University
Montreal, QC H3G 1M8, Canada
dongyu@ece.concordia.ca

Ness Shroff
ECE, Ohio State University
Columbus, OH 43210, USA
shroff@ece.osu.edu

Abstract

In this paper, we study the asymptotic behavior of an AQM router serving many AIMD flows. Our model for the AQM router is general and covers most AQM schemes in the current literature. We use a window-based model for the AIMD flows. When the number of AIMD flows is large, we show that the system converges point wise to a limit model. Further, under certain technical conditions, we prove that the system converges uniformly in time and that a steady state exists. We then study the steady state of the system. We show that using appropriate feedback control, the backward accumulation process of net input to the router can be bounded. Under this condition and Gaussian assumption, the queue length distribution can be shown to decay very fast in terms of the buffer size (squared exponential decay). We also provide numerical results to study and support our analytical results.

1 Introduction

TCP [1] is the most dominant transport protocol in the Internet today. It plays an important role in the efficient operation of the network. A key feature of TCP congestion control is the Additive Increase Multiplicative Decrease (AIMD) algorithm, which adjusts the data transfer rate based on feedback information from the routers. A router can use different Active Queue Management (AQM) schemes to generate feedback information. Popular AQM schemes are Random Early Detection (RED) [2], Random Exponential Marking (REM) [3], and Adaptive Virtual Queue (AVQ) [4]. The traditional DropTail TCP mechanism can also be viewed as an AQM scheme, in which packet loss is used as the feedback. Studying the behavior of an AIMD/AQM router is important because it not only helps us understand how the network works, but can also provide us with insight on how to design and improve network performance.

Modeling TCP traffic has long been an important topic. A popular approach was to use Markov chain modeling [5][6], where the congestion window size acts as the state of the Markov chain and the loss rate determines the transition probabilities. Since such modeling is done at the individual flow level, and while there have been efforts to extend the results to more than one flow [7][8], generally speaking, this type of approach is not scalable when the number of flows is large. Another approach is to apply limit theorems to the aggregate TCP traffic [9][10][11][12][13]. In [11], the performance of a RED router with many TCP flows has been considered. It has been shown that when the number of TCP flows N goes to infinity, the congestion window sizes of TCP flows converge to the same distribution and become asymptotically independent of each other. In [11], the normalized queue length (the real queue length over N) was shown to converge to a constant. A more general model was developed in [12] to include both RED and REM. In [13], heterogeneous TCP flows were considered for a RED router.

While the modeling in [11][12][13] shed new insight, it is limited to either RED or REM routers and hence does not help understanding other AQM schemes. Further, the convergence proved in those models is point wise in time. Since in reality, the number of flows is normally in a given range, which doesn't change with time, when time is large, it is not clear whether the limit model is still a good approximation of the real system.

In this paper, we first present a general model for the router that covers virtually all AQM schemes. We extend the results of [11] to this general model. We also prove *uniform convergence* over time under certain conditions (global exponential stability). Hence, our model can be used study the steady state behavior of the system. In addition, we develop a different approach from [11] to study the queueing behavior of the system. We show that if we set the target link utilization to be slightly less than one, the network performance can be significantly improved. When the number of flows is fixed, we show that the overflow probability can be made to decay very fast (squared exponential decay) if the AQM schemes are designed appropriately. Since our results are derived from a general AQM model, they provide directions on how to design and improve AQM schemes.

2 Model

We use a similar model as that in [11], but our model for the AQM router is more general. We consider an AQM router serving N TCP flows. We let $X^{(N)}$ denote the explicit dependence of the quantity X on the number of flows N. The link capacity of the router is NC, where C is a constant. Time is discrete and all TCP flows have the round trip delay of one time unit. We assume that the system begins at time $t = 0$.

The flow control algorithm of a TCP flow is as follows. The size of the congestion window (the amount of unacknowledged packets in the network per round trip) is increased by one packet for the next round trip if all packets transmitted in the current round trip are not marked or dropped. Otherwise, if at least one packet is marked or dropped, the congestion window is halved. This corresponds to the AIMD behavior of TCP.

The router marks or drops packets according to the AQM scheme deployed on the router. At each time slot, a marking/dropping probability is calculated based on the current status of the router. Different AQM schemes may use different parameters to calculate the marking/dropping probability. Generally the probability at time t can be expressed as

$$p^{(N)}(t) = f_t^{(N)}(x^{(N)}(0), \cdots, x^{(N)}(t)), \tag{1}$$

where $x^{(N)}(s)$ is the input data rate to the router at time s. This model is general (i.e., all AQM schemes are special cases) since the total history of the input rate determines everything at the router. However, in practice, routers cannot keep the total history of the input data rate. Instead, the router keeps the information of a finite number of state variables. Let the number of state variables be K and the vector of state variables be $S^{(N)}(t) = [S_1^{(N)}(t), \cdots, S_K^{(N)}(t)]$. Then a more practical marking/dropping function (a special case of the marking/dropping function Eq. (1)) will be

$$p^{(N)}(t) = f^{(N)}(S^{(N)}(t-1), x^{(N)}(t)). \tag{2}$$

$S^{(N)}(t)$ evolves according to

$$S^{(N)}(t) = F^{(N)}(S^{(N)}(t-1), x^{(N)}(t)).$$

Note that $x^{(N)}(t)$ can also be a state variable and hence a component in $S^{(N)}(t)$. While some of the results in this paper can be shown under the more general model in Eq. (1), in this paper, we will mainly focus on the AQM model with state variables given by Eq. (2). It should be noted that this model also covers virtually all practical AQM schemes studied in the literature.

After the marking/dropping probability is calculated, the router marks or drops packets according to $p^{(N)}(t)$. Depending on whether the router uses random number generators or not, we can classify AQM schemes into two categories. Some AQM schemes generate independent $[0, 1]$-uniform random numbers for each packet. If the random number is greater than $p^{(N)}(t)$, the packet will not be marked or dropped. Otherwise, the router marks/drops the packet. Hence, each packet is independently marked or dropped with probability $p^{(N)}(t)$. Examples are RED and REM. Other AQM schemes do not use random number generators. Examples are DropTail and AVQ. In these schemes, at time t, if there are $x^{(N)}(t)$ input packets, the router will mark or drop exactly $p^{(N)}(t)x^{(N)}(t)$ packets. In this case, packets are no longer marked or dropped independently. Our theoretical results will include both type of AQM schemes.

We assume that TCP flows always have enough data to transmit and define $W_i^{(N)}(t)$ to be the congestion window size of flow i at time t. The input data rate to the router at time t will then be $x^{(N)}(t) = \sum_{i=1}^N W_i^{(N)}(t)$. We also assume that the integer $W_i^{(N)}(t)$ is in the range $\{1, \cdots, W_{max}\}$ for some finite integer W_{max}.

We define $M_{i,j}^{(N)}(t+1)$ as the indicator function of the event that the jth packet from source i is not marked/dropped in time slot $[t, t+1)$, i.e.,

$$M_{i,j}^{(N)}(t+1) = 1[V_{i,j}(t+1) > p^{(N)}(t)],$$

where $\{V_{i,j}(t+1), i, j = 1, \cdots; t = 0, 1, \cdots\}$ are $[0, 1]$-uniform random variables (r.v.s) and they may or may not be independent with each other depending on the AQM scheme used. We also define $M_i^{(N)}(t+1)$ as the indicator function of the event that no packet from flow i is marked/dropped in time slot $[t, t+1)$, i.e.,

$$M_i^{(N)}(t+1) = \Pi_{j=1}^{W_i^{(N)}(t)} M_{i,j}^{(N)}(t+1)$$

The evolution of the congestion window for source i can then be described by

$$\begin{aligned} W_i^{(N)}(t+1) &= \min(W_i^{(N)}(t)+1, W_{max}) M_i^{(N)}(t+1) \\ &\quad + \ulcorner \frac{W_i^{(N)}(t)}{2} \urcorner (1 - M_i^{(N)}(t+1)). \end{aligned}$$

where $\ulcorner x \urcorner$ is the smallest integer that is greater than or equals to x.

3 Point Wise Convergence

We first make two assumptions.

$(A1)$ There exist continuous functions f and F such that

$$\begin{aligned} p^{(N)}(t) &= f^{(N)}(S^{(N)}(t-1), x^{(N)}(t)) \\ &= f(\frac{S^{(N)}(t-1)}{N}, \frac{x^{(N)}(t)}{N}) \end{aligned}$$

and

$$\begin{aligned} S^{(N)}(t) &= F^{(N)}(S^{(N)}(t-1), x^{(N)}(t)) \\ &= N \cdot F(\frac{S^{(N)}(t-1)}{N}, \frac{x^{(N)}(t)}{N}) \end{aligned}$$

$(A2)$ The system starts with the initial conditions:

$$S^{(N)}(0) = S, \ W_i^{(N)}(0) = W,$$

for $i = 1, \cdots, N$, where $1 \leq W \leq W_{max}$ is a constant and S is a constant vector.

Assumption $(A1)$ tells us how the marking function and state variables scale when N increases and assumption $(A2)$ simply give us the initial conditions. As should be clear, these are fairly non-restrictive assumptions and can model various AQM schemes, as shown in Section 3.1. The limit behavior of the system is described by the following theorem. In this paper, we use $\xrightarrow{P}_N$ to denote convergence in probability, use $\Rightarrow_N$ to denote convergence in distribution, and use $\xrightarrow{a.s.}_N$ to denote almost surely convergence. Equivalence in distribution between random variables is denoted by $=_{st}$.

Theorem 1 *Under the assumptions $(A1)-(A2)$, for each time t, there exists a constant $p(t)$, a constant vector $S(t)$, and a r.v. $W(t)$ such that:*

1. $p^{(N)}(t) \xrightarrow{P}_N p(t)$, $\frac{S^{(N)}(t)}{N} \xrightarrow{P}_N S(t)$,
   and $W_i^{(N)}(t) \Rightarrow_N W(t)$.

2. *For any function g,*

$$\frac{1}{N}\sum_{i=1}^{N} g(W_i^{(N)}(t)) \to_N \mathbb{E}[g(W(t)].$$

3. *For any integer $I = 1, 2, \cdots$, the r.v.s $\{W_i^{(N)}(t), i = 1, \cdots, I\}$ become asymptotically independent as N goes to infinity.*

4. *For the limit model, we have*

$$\begin{aligned} p(t) &= f(S(t-1), \mathbb{E}[W(t)]), \\ S(t) &= F(S(t-1), \mathbb{E}[W(t)]) \end{aligned} \tag{3}$$

and

$$\begin{aligned} W(t+1) =_{st} \; & min(W(t)+1, W_{max})M(t+1) \\ & + \ulcorner \frac{W(t)}{2} \urcorner (1 - M(t+1)), \end{aligned} \tag{4}$$

where

$$M(t+1) = 1[V(t+1) \leq (1-p(t))^{W(t)}]$$

for i.i.d $[0,1]$-uniform r.v.s $\{V(t+1), t = 0, 1, \cdots\}$.

All proofs are provided in the Appendices.

3.1 Different AQM schemes

As mentioned earlier, our AQM model with state variables is quite general. Next, we will discuss how practical AQM schemes such as DropTail, RED, REM, and AVQ can fit in our model.

For a RED router with infinite buffer, there is only one state variable, the queue size $Q^{(N)}(t)$, and it evolves according to

$$Q^{(N)}(t) = [Q^{(N)}(t-1) + x^{(N)}(t) - NC]^+,$$

where NC is the link capacity. The marking function is then

$$p^{(N)}(t) = f^{(N)}(Q^{(N)}(t-1)) = f(\frac{Q^{(N)}(t-1)}{N}),$$

where $f(\cdot)$ is a continuous function.

For a REM router with infinite buffer, in addition to the queue size $Q^{(N)}(t)$, there is another state variable $P^{(N)}(t)$ called the price and evolves according to

$$\begin{aligned} P^{(N)}(t) = [&P^{(N)}(t-1) + \\ &\gamma(\alpha(Q^{(N)}(t) - NQ^*) + x^{(N)}(t) - NC)]^+, \end{aligned} \tag{5}$$

where γ and α are REM parameters and $Q^* \geq 0$ is the normalized target queue length. The marking function is

$$p^{(N)}(t) = 1 - \Phi^{-\frac{P^{(N)}(t)}{N}},$$

where $\Phi > 1$ is a constant.

For a DropTail router with buffer size NB, the state variable is the queue size $Q^{(N)}(t)$ and it evolves according to

$$Q^{(N)}(t) = [Q^{(N)}(t-1) + x^{(N)}(t) - NC]_0^{NB},$$

where

$$[x]_a^b = \begin{cases} a & \text{if } x < a \\ x & \text{if } a \leq x \leq b \\ b & \text{if } x > b \end{cases}$$

The marking/dropping function is then

$$p^{(N)}(t) = \frac{[Q^{(N)}(t-1) + x^{(N)}(t) - NC - NB]^+}{x^{(N)}(t)}.$$

For an AVQ router with link capacity NC and buffer size NB, the state variable will be the queue length $VQ^{(N)}(t)$ of the virtual queue (with link capacity ρNC, where $\rho < 1$ is a constant),

$$VQ^{(N)}(t) = [VQ^{(N)}(t-1) + x^{(N)}(t) - \rho NC]_0^{NB}.$$

The marking/dropping probability is the dropping probability of the virtual queue

$$p^{(N)}(t) = \frac{[VQ^{(N)}(t-1) + x^{(N)}(t) - \rho NC - NB]^+}{x^{(N)}(t)}.$$

One can easily check that assumption $(A1)$ is satisfied for all AQM schemes discussed and hence Theorem 1 holds.

4 Stability and Uniform Convergence

A shortcoming of Theorem 1 (and also the results of [11]) is that the proved convergence is for each point in time (point-wise convergence). This means that for a given error, the required N may depend on time t. Hence, stronger results such as *uniform convergence* are needed. The convergence is uniform in time t if for any given error ε, there exists a N_0 (not dependent on t), such that when $N > N_0$, the actual error is always less than ε, no matter what the time t is. We believe that uniform convergence is related to the stability of the limit model, which we will discuss next.

Now, for the limit model, let $g_j(t) = \mathbb{P}\{w(t) = j\}$, $G(t) = [g_1(t), \cdots, g_{W_{max}}(t)]$, and $A(t) = [S(t), G(t)]$. Then $A(t)$ is the state vector of the system. For each time t, we have

$$\begin{aligned} x(t) &= [1, \cdots, W_{max}]G^T(t) \\ S(t) &= F(S(t-1), x(t)) \\ p(t) &= f(S(t-1), x(t)) \\ G(t+1) &= T(p(t))G(t), \end{aligned} \tag{6}$$

where $T(p(t))$ is the transition matrix and depends on $p(t)$. We assume this system is stable and let $A^* = [S^*, G^*]$ be the equilibrium point. The system is called global exponential stable [14] if there exist constants $m,\ a\ (0 < a < 1)$, such that

$$||A(t) - A^*|| \leq m a^{t-t_0} ||A(t_0) - A^*|| \tag{7}$$

for any t and t_0 ($t \geq t_0$), where $||\cdot||$ is the Euclid norm. In practice it is desirable to have a stable flow control system. This can be done by correctly choosing the AQM parameters [15][16]. In [16], it has been shown that if the AQM parameters are properly chosen, the system can be made to be global exponential stable.

For the real system, let $n_j^{(N)}(t)$ be the number of flows with window size j at time t and define

$$\begin{aligned} g_j^{(N)}(t) &= \frac{n_j^{(N)}}{N} \\ G^{(N)}(t) &= [g_1^{(N)}(t), \cdots, g_{W_{max}}^{(N)}(t)] \\ A^{(N)}(t) &= [\frac{S^{(N)}(t)}{N}, G^{(N)}(t)]. \end{aligned}$$

Next, we assume $A^{(N)}(t) \in L^2$ and study the convergence in L^2 space. Note that for DropTail and AVQ, since the buffer size is finite, this assumption is always true. For RED, if the buffer size is finite, this assumption is also true. Even if the buffer size is infinite, if the queue distribution decays fast (e.g., exponential decay), we still have this assumption satisfied.

We also assume that function F satisfy the following Lipschitz condition, i.e., there exists a constant $\alpha \geq 0$, such that

$$||F(S(t-1), x_1(t) - F(S(t-1), x_2(t)|| \leq \alpha ||x_1(t) - x_2(t)||, \tag{8}$$

for any $S(t-1)$, $x_1(t)$, and $x_2(t)$. It is easy to see that DropTail, AVQ, RED, and REM all satisfy this condition. The next proposition shows the relationship between the stability and uniform convergence.

Proposition 1 *If the limit model defined by Eq. (6) is global exponential stable with $m = 1$, $A^{(N)}(t) \in L^2$, and F satisfies the Lipschitz condition Eq. (8), then $A^{(N)}(t)$ converges uniformly to $A(t)$ in L^2 as the number of flows N goes to infinity.*

Proposition 1 tells us that if we choose AQM parameters correctly, the system will uniformly converge to the limit model and a steady state exists. Note that global exponential stability is a sufficient, but not necessary condition, for the uniform convergence. Hence, there may be other scenarios in which the system converges uniformly.

Next, we will assume that the limit model is stable and study its steady state behavior, which we believe can give us insight on how different AQM schemes perform.

In [11], the authors studied RED. When N goes to infinity, the normalized queue size $\frac{Q^{(N)}}{N}$ goes to a constant (greater than zero) and the normalized input rate $\frac{x^{(N)}}{N}$ goes to the link capacity C. But the actual queue size goes to infinity, which may not be desirable in practice. From the proof of our theorem, we can see that although the queue size is important to the network performance, it is not necessary that the marking/dropping probability will be related with it. In fact, in our proof, we do not really need information on the queue size at all. It is only in some special cases of AQM schemes (e.g. RED) that the queue size is used as a state variable. Later, we will see that if we decouple the real queue size and the marking/dropping probability, we may get improved performance.

Similarly, in REM, if we assume the limit model is stable, we can easily see that in steady state, the normalized queue size $\frac{Q^{(N)}}{N}$ goes to a constant Q^* and the normalized input rate $\frac{x^{(N)}}{N}$ goes to the link capacity C. Different from RED, the limit of the normalized queue size (Q^*) is chosen by the user and can take zero value. But even if we set $Q^* = 0$, because the link utilization is 1, it is still possible for the actual queue size to be large or even go to infinity, when N goes to infinity.

In a router using DropTail (note that we have finite buffer here), if we assume steady state, we can see that the normalized queue size $\frac{Q^{(N)}}{N}$ goes to the normalized buffer size B and the normalized input rate $\frac{x^{(N)}}{N}$ goes to $\frac{C}{1-p}$, where p is the marking/dropping probability. This tells us that the buffer is always full and the input rate is in fact greater than the link capacity.

In AVQ, the marking/dropping probability is no longer directly related to the real queue size. In steady state, the normalized virtual queue size $\frac{VQ^{(N)}}{N}$ goes to the normalized buffer size B and the normalized input rate $\frac{x^{(N)}}{N}$ goes to $\frac{\rho C}{1-p}$. Remember that ρ is a parameter that we can choose. If we choose ρ to be slightly less than $1 - p$, then the normalized input rate is less than the link capacity C. Hence, the normalized queue size $\frac{Q^{(N)}}{N}$ goes to zero.

From the discussion of different AQM schemes, we can see that different AQM schemes do exhibit different limiting behavior. Our model provides us with an analytical tool to compare different AQM schemes and also sheds light on how to design an appropriate AQM scheme (e.g., decoupling the marking probability from the real queue size). An interesting observation is that if we set the target link utilization to be one (such as RED and REM), the normalized queue length is generally greater than zero and hence the real queue length could be very large when N is large. But once we set target utilization to be less than one, no matter how close it is to one, the normalized queue length becomes zero.

From Theorem 1, we also see that the asymptotic independence of TCP congestion windows does not really come from the randomized marking/dropping (RED, REM), but is mainly because of the large number of TCP flows. Even if a router uses DropTail, when N goes to infinity, the TCP congestion windows of different flows become asymptotically independent.

5 Asymptotic Queueing Behavior

While Theorem 1 tells us what the system converges to, it does not provide us with an explicit distribution for the queue length. For example, from the discussion of the steady state behavior of different AQM schemes in Section 4, we see that the only information we can get about the queue is that the normalized queue length goes to a constant (the constant could be zero). We cannot tell how the queue length is distributed.

When the number of flows are large, there are typically three types of approximations we can use to study the queueing behavior: Central Limit Theorem (CLT) [17], Large Deviation Principles (LDP) [18][19], Moderate Deviation Principles (MDP) [20]. In [20], it has also been shown that when the traffic is heavy, Gaussian approximation performs better than the other two methods. Note that TCP always tries to fully utilize the link capacity if there are enough data to transmit. Hence we expect the AQM router to operate in a heavy traffic mode. Based on this observation, we use the Gaussian approach to study the queueing behavior of the AQM router.

We assume that the target link utilization of the AQM router is slightly less than one and that the system is in steady state. Note that we take a different approach to that of [11]. In [11], the CLT is directly applied to the queue size and tells us how fast the queue size increases when N increase. But that is not what we want in a real network since we want the queue size to be small. This is also the reason why we are interested in AQM routers that have target link utilization less than one. In Section 6, we will see that a very high target link utilization does not mean that the actual link utilization will be high, and may only cause unnecessary workload.

Now, let

$$\boldsymbol{X}^{(N)} = [x^{(N)}(0), x^{(N)}(1), \cdots]$$

be the process of the aggregate input rate. Since $\boldsymbol{X}^{(N)}$ is the sum of many small flows, we expect it to behave like a Gaussian process. Define

$$V_t^{(N)} = \text{Var}\{\sum_{i=0}^{t-1} x^{(N)}(t_0 - i)\}$$

to be the variance of the backward accumulation of $\boldsymbol{X}$ over a time period of t. Since we assume that the system is in steady state, $V_t^{(N)}$ will not depend on t_0. Let the mean value of the congestion window size $\mathbb{E}[W(t)] = \rho C$, where $\rho < 1$ is the target link utilization. Then under the Gaussian assumption, the real queue size distribution can be approximated by

$$\log \mathbb{P}\{Q^{(N)} > b\} \approx -\inf_{t \geq 0} \frac{(b + N(1-\rho)Ct)^2}{2V_t^{(N)}}, \tag{9}$$

where b is queue size.

Rigorous forms and proofs of Eq. (9) can be found in [17][21][20]. When N is large, Eq. (9) can be used as an approximation of the tail probability of the real queue. We can see that when the target link utilization ρ is fixed, $V_t^{(N)}$ plays an important role in determining how fast the queue decays. For a fixed N, we will next simply use V_t instead of $V_t^{(N)}$.

In an open-loop network, V_t generally goes to infinity when t goes to infinity. For example, if the input process is a long range dependent process with Hurst parameter $H \in [1/2, 1)$, we will have $V_t \sim St^{2H}$, when $t \to \infty$, where S is a constant. In our system, if we assume that the marking probability is a fixed constant (i.e., it does not change according to the status of the router) and the router marks/drops packets independently, then the flows are independent of each other and the aggregate input process can be seen as the sum of many independent Markov modulated processes. In this setting, we will have V_t increasing linearly with time t. However in a closed-loop network with appropriately designed AQM schemes, V_t can be bounded and does not go to infinity when t goes to infinity. The intuition behind this is that in a closed-loop network, when the input rate is higher than the average value for a long time, the router will detect it and try to reduce the input rate. While in an open-loop network, there is no mechanism to prevent to the input rate from being higher than the average rate for a long period of time (although with a small probability).

We will use a RED router with link capacity C as an example. Let $x(t)$ be the input rate at time t and assume that $x(t)$ is bounded by X_{max}, where $X_{max} > C$. We also assume that when all packets are marked, $x(t)$ can be decreased to be less than C in a finite time T_0, which is obviously true for AIMD flows. We design the RED scheme such that the marking probability $p(Q(t)) = 1$ if $Q(t) \geq B_h$ and $p(Q(t)) = 0$ if $Q(t) \leq B_l$, where $B_l \leq B_h$ are two constants. We claim that for any t_0 and $T \geq 0$, the net input accumulation $\sum_{j=1}^{T} x(t_0 + j) - CT$ can be bounded by $B_h + (T_0+1)(X_{max} - C)$. The proof is simple. Once the accumulation is greater than B_h, it means the queue size is greater than B_h and all packets will be marked. $x(t)$ will begin to decrease, but it may be still higher than C for at most a time period of T_0. Within this time period, the accumulation may continue increasing, but by at most $(T_0 + 1)(X_{max} - C)$. After that, $x(t)$ will be less than C and the accumulation decreases. Similarly, by correctly choosing B_l, we can derive a lower bound for $\sum_{j=1}^{T} x(t_0 + j) - CT$. This tells us that the net input accumulation is bounded no matter how long the time period is. Hence, the variance V_t will also be bounded.

In REM and AVQ, things are more complicated and it is not clear whether V_t can be bounded. In REM, the net input accumulation itself obviously cannot be bounded because, with exponential marking, the marking probability can never be one and hence, no matter how large the accumulation is, there always a small probability that it will increase. Finding the properties of V_t for a general AQM scheme will be part of our future work. Note that even if V_t is not bounded, because of the feedback mechanism, we expect that it can at least be made to increase slowly, and hence still expect that the queue distribution will decay quickly.

Note that Eq. (9) holds no matter whether $V_t^{(N)}$ is bounded. However, in [21], it has been shown that when $V_t^{(N)}$ is bounded, the queue length distribution decays squared exponentially, i.e.,

$$\mathbb{P}\{Q^{(N)} > b\} \leq e^{-\frac{b^2}{2D^{(N)}}}, \tag{10}$$

when b is large. Note that $D^{(N)}$ is a constant related to N and we assume N is fixed here. Recall that if $V_t^{(N)}$ increases linearly with t, the tail probability decays at most exponentially [17]. Hence Eq. (10) not only gives us an asymptotic upper bound on the queue distribution, but also give us insight on how to design AQM schemes, i.e., try to ensure that $V_t^{(N)}$ is bounded.

6 Numerical Results

We use an *ns2* simulator to simulate a router that serves N TCP flows. The link capacity is $2N$ Mbps and the buffer size is $12N$ packets. Each packet is 1000 bytes. The round trip delay is 10 msec for all flows. AVQ is used as the AQM scheme and the target link utilization is set to be 96%. We run the simulations for $N = 25,\ 50,\ 100$, and 200. Fig 1 shows the distribution of the

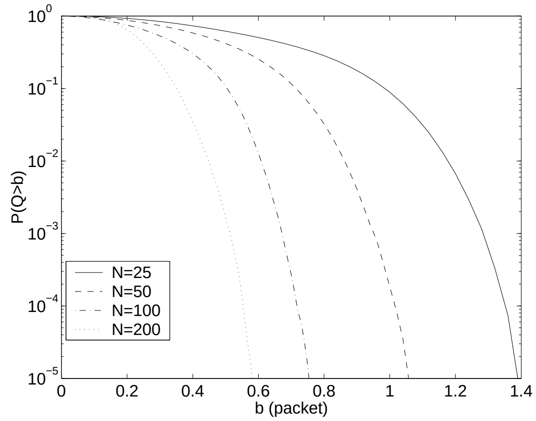

Figure 1: Tail Probability of the Normalized Queue

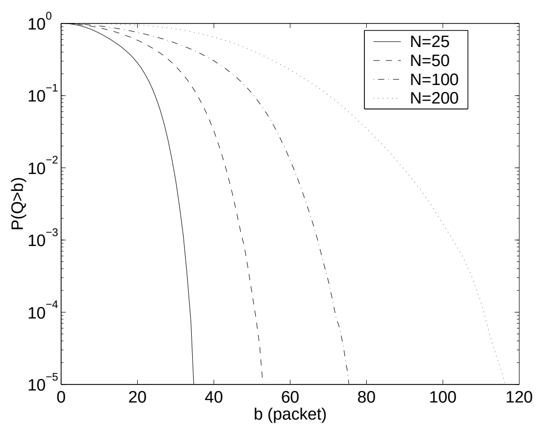

Figure 2: Tail Probability of the Real Queue

normalized queue $\mathbb{P}\{\frac{Q^{(N)}}{N} > b\}$. From Fig 1, we can see that when N increases, the tail probability of the normalized queue decreases. This matches our theoretical result in Theorem 1.

In Fig 2, we show the tail probability of the real queue. Two important observations can be made from this figure. First, the tail probability decays very fast as we discussed in Section 5. This differs significantly with that of an open-loop network, where the decay is at most linear (note that the probability has a log scale). Second, when N increases, the tail probability also increases (for a fixed queue size b). This also differs with what we typically see in an open-loop network. Note that in an open-loop network, under some mild conditions, we can show that when the number of flows increases, the real queue size goes to zero by the *many sources asymptotic results* [19]. In a closed-loop network with feedback control, the traffic is generally smoother than that in an open-loop network [21]. Hence, we expect that the real queue size decreases when the number of flows increases. But this simulation shows that this is in fact not true. It

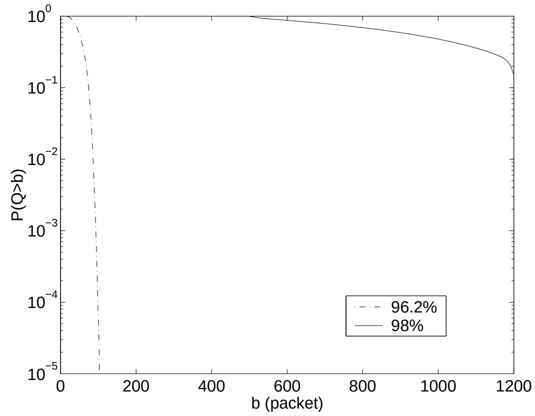

Figure 3: Tail Probability of the Real Queue

also tells us that simply assuming that the marking probability is fixed and treating TCP flows as independent Markov modulated processes may not be appropriate if we want to understand the queueing behavior of a TCP network because it cannot explain the result of this simulation (i.e., why statistical multiplexing does not happen).

In all above simulations, we also measured the actual link utilization and they all equal to the target link utilization 96%. In the next simulation, we show that trying to fully utilize the link capacity may not work well in practice. We fix $N = 100$ and set the target link utilization to be 98% and 96.2% respectively. We measure the actual link utilization and find that it is 96.2% in both simulations. Under the same actual link utilization, we compare their tail probabilities in Fig 3. We can see that when the target link utilization is set to 98%, the tail probability is much higher than the other one. This tells us that setting a very high target link utilization does not guarantee that the actual link utilization is high and may only cause unnecessary workload.

7 Conclusion

In this paper, we have studied the asymptotic behavior of a general AQM router serving many AIMD flows. We first proved that when the number of flows is large, the system converges to a limit model. We then showed that under certain conditions, the convergence is uniform in time and that a steady state exists. Furthermore, we studied the steady state queueing behavior of the AQM router under Gaussian assumption. If the target link utilization is slightly less than one, we find that when the number of flows is fixed, if the AQM scheme is designed appropriately, it is possible to bound the variance of the net input accumulation, and the overflow probability will decay as a squared exponential. These results also provide us with insight on how to design an AQM scheme. Moreover, since our model is general and covers most AQM schemes, it can be used to study and compare different AQM schemes in the current literature.

References

[1] V. Jacobson, "Congestion Avoidance and Control," *SIGCOMM Symposium on Communications Architectures and Protocols*, pp. 314–329, 1988.

[2] S. Floyd and V. Jacobson, "Random Early Detection Gateways for Congestion Avoidance," *IEEE/ACM Transactions on Networking*, vol. 1, no. 4, pp. 397–413, Aug. 1993.

[3] S. Athuraliya, S. Low, V. Li, and Q. Yin, "REM: Active queue management," *IEEE Network*, vol. 15, no. 3, pp. 48–53, 2001.

[4] S. Kunniyur and R. Srikant, "Analysis and Design of an Adaptive Virtual Queue (AVQ) Algorithm for Active Queue Management," in *Proceedings of ACM SIGCOMM*, San Diego, California, USA, August 2001.

[5] J. Padhye, V. Firoiu, D. Towsley, and J. Kurose, "Modeling TCP Throughput: a Simple Model and its Empirical Validation," in *Proceedings of ACM SIGCOMM*, Sept. 1998, pp. 303–314.

[6] E. Altman, K. Avrachenkov, and C. Barakat, "TCP in presence of bursty losses," in *Measurement and Modeling of Computer Systems*, 2000, pp. 124–133.

[7] G. Hasegawa and M. Murata, "Analysis of Dynamic Behaviors of Many TCP Connections Sharing Tail-Drop/RED Routers," in *Proceedings of IEEE GLOBECOM*, 2001.

[8] M. Garetto, R. L. Cigno, M. Meo, and M. A. Marsan, "A Detailed and Accurate Closed Queueing Network Model of Many Interacting TCP Flows," in *Proceedings of IEEE INFOCOM*, 2001, pp. 1706–1715.

[9] S. Shakkottai and R. Srikant, "Mean FDE Models for Internet Congestion Control Under a Many-Flows Regime," Tech. Rep., University of Illinois, 2001.

[10] D. Hong and D. Lebedev, "Many TCP user asymptotic ananlysis of the AIMD model," Tech. Rep., INRIA, 2001.

[11] P. Tinnakornsrisuphap and A. Makowski, "Limit behavior of ecn/red gateways under a large number of tcp flows," in *Proceedings of IEEE INFOCOM*, 2003.

[12] P. Tinnakornsrisuphap and R. J. La, "Characterization of queue fluctuations in probabilistic AQM mechanisms," in *Proceedings of ACM SIGMETRICS*, 2004.

[13] P. Tinnakornsrisuphap and R. J. La, "Asymptotic behavior of heterogeneous tcp flows and red gateway," *IEEE/ACM Transactions on Networking*, vol. 14, no. 1, pp. 108–120, Feb. 2006.

[14] R. M. Murray, Z. Li, and S. S. Sastry, *A Mathematical Introduction to Robotic Manipulation*, CRC Press, 1994.

[15] F. Paganini, J. Doyle, and S. Low, "Scalable Laws for Stable Network Congestion Control," in *Proceedings of the 40th IEEE Conference on Decision and Control*, 2001, vol. 1.

[16] S. Deb and R. Srikant, "Global stability of congestion controllers for the internet," *IEEE Transactions on Automatic Control*, vol. 48, pp. 1055–1060, June 2003.

[17] J. Choe and N. B. Shroff, "Use of the supremum distribution of Gaussian processes in queueing analysis with long-range dependence and self-similarity," *Stochastic Models*, vol. 16, no. 2, Feb 2000.

[18] N. G. Duffield, "Economies of scale in queues with sources having power-law large deviation scaling," *Journal of Applied Probability*, vol. 33, pp. 840–857, 1996.

[19] N. Likhanov and R. Mazumdar, "Cell loss asymptotics for buffers fed with a large number of independent stationary sources," *Journal of Applied Probability*, vol. 36, no. 1, pp. 86–96, 1999.

[20] D. Wischik, "Moderate Deviations in Queueing Theory," Feb. 2001, http://www.statslab.cam.ac.uk/~djw1005/Stats/Research/moddev.html.

[21] D. Qiu and N. B. Shroff, "Queueing Properties of Feedback Flow Control Systems," *IEEE/ACM Transactions on Networking*, vol. 13, no. 1, Feb. 2005.

A Proof of Theorem 1

Our proof of Theorem 1 follows the same approach as that in [11]. But now we focus on the marking/dropping probability rather than the queue size. In addition, we need to deal with AQM schemes that do not use random number generator.

To prove the theorem, for each $t = 0, 1, \cdots$, we will prove the statements **[A:t]**, **[B:t]**, **[C:t]**, and **[D:t]** below.

[A:t] $p^{(N)}(t) \xrightarrow{P}_N p(t)$, $\frac{S^{(N)}(t)}{N} \xrightarrow{P}_N S(t)$, and $p(t)$, $S(t)$ satisfy Eq. (3).

[B:t] $W_i^{(N)}(t) \Rightarrow_N W(t)$ and $W(t)$ evolves according to Eq. (4).

[C:t] For any integer $I = 1, 2, \cdots$, the r.v.s $\{W_i^{(N)}(t), i = 1, \cdots, I\}$ become asymptotically independent as N goes to infinity.

[D:t] For any function g, $\frac{1}{N}\sum_{i=1}^{N} g(W_i^{(N)}(t)) \xrightarrow{P}_N \mathbb{E}[g(W(t)]$.

It is easy to see that statements **[A:t]-[D:t]** hold for $t = 0$. Next, we will prove that if **[A:t]-[D:t]** are true, **[A:t+1]-[D:t+1]** will also be true. Hence, by induction, the statements **[A:t]-[D:t]** hold for all $t = 0, 1, \cdots$. We prove this and hence the theorem by the following lemmas.

Lemma 1 *If* ***[A:t]-[D:t]*** *hold,* ***[B:t+1]*** *also holds.*

Proof: Define $Z_i^{(N)}(t) = (1-p^{(N)}(t))^{W_i^{(N)}(t)}$. The convergence **[A:t]** and **[B:t]** imply the joint convergence $(p^{(N)}(t), W_i^{(N)}(t)) \Rightarrow_N (p(t), W(t))$. Since $(x, w) \to (1-x)^w$ is a continuous mapping, by the Continuous Mapping Theorem, we have,

$$(Z_i^{(N)}(t), W_i^{(N)}(t)) \Rightarrow_N (Z(t), W(t)),$$

where $Z(t) = (1-p(t))^{W(t)}$. Consider an arbitrary bounded mapping $g : \mathbb{N} \to \mathbb{R}$ and define

$$F_g(z, w) = zg(min(w+1, W_{max})) + (1-z)g(\ulcorner \frac{w}{2} \urcorner).$$

Then F_g is bounded and continuous on $[0,1] \times \mathbb{N}$ ($\mathbb{N}$ is topologized according to the usual discrete topology). By the definition of convergence in distribution, we have

$$\lim_{N\to\infty} \mathbb{E}[F_g(Z_i^{(N)}(t), W_i^{(N)}(t))] = \mathbb{E}[F_g(Z(t), W(t))] \tag{11}$$

Let $\mathcal{F}_t$ be the σ-field generated by the r.v.s $\{W(0), V(s), s = 1, \cdots, t\}$. From Eq. (4), we get

$$\mathbb{E}[M(t+1)|\mathcal{F}_t] = (1-p(t))^{W(t)} = Z(t)$$

and

$$\begin{aligned}\mathbb{E}[g(W(t+1))|\mathcal{F}_t] &= Z(t)g(min(W(t)+1, W_{max})) \\ &\quad +(1-Z(t))g(\ulcorner \frac{W(t)}{2} \urcorner) \\ &= F_g(Z(t), W(t)).\end{aligned}$$

So,

$$\begin{aligned}\mathbb{E}[g(W(t+1))] &= \mathbb{E}[\mathbb{E}[g(W(t+1))|\mathcal{F}_t]] \\ &= \mathbb{E}[F_g(Z(t), W(t))].\end{aligned} \tag{12}$$

Let $\mathcal{F}_t^{(N)}$ be the σ-field generated by the r.v.s $\{W_1^{(N)}(0), \cdots,$ $W_N^{(N)}(0), V_{i,j}(s), i, j = 1, 2, \cdots, s = 1, \cdots, t\}$. Then

$$\mathbb{E}[g(W_i^{(N)}(t+1))|\mathcal{F}_t^{(N)}] = F_g(Y_i^{(N)}(t), W_i^{(N)}(t)),$$

where $Y_i^{(N)}(t)$ is the probability that no packet from flow i is marked/dropped at time t under the σ-field $\mathcal{F}_t^{(N)}$. Next, we will prove that

$$Y_i^{(N)}(t) - Z_i^{(N)}(t) \xrightarrow{a.s.}_N 0. \tag{13}$$

When the router marks/drops packets independently, It is easy to show that $Y_i^{(N)}(t) = Z_i^{(N)}(t)$ [11]. Hence, Eq. (13) is true. When packets are not marked/dropped independently (i.e., the router marks/drops exactly $p^{(N)}(t)x^{(N)}(t)$ packets), we have

$$\begin{aligned}Y_i^{(N)}(t) &= \frac{C_{x^{(N)}(t)-W_i^{(N)}(t)}^{p^{(N)}(t)x^{(N)}(t)}}{C_{x^{(N)}(t)}^{p^{(N)}(t)x^{(N)}(t)}} \\ &= \frac{(1-p^{(N)}(t))\cdots(1-p^{(N)}(t)-\frac{W_i^{(N)}(t)-1}{x^{(N)}(t)})}{1\cdots(1-\frac{W_i^{(N)}(t)-1}{x^{(N)}(t)})}.\end{aligned}$$

Remember that

$$Z_i^{(N)}(t) = (1 - p^{(N)}(t))^{W_i^{(N)}(t)}.$$

When $N \to \infty$, we also have $x^{(N)}(t) \to \infty$. Since $W_i^{(N)}(t)$ is always less than W_{max}, we can see that $Y_i^{(N)}(t)(\omega) - Z_i^{(N)}(t)(\omega) \to_N 0$ for any $\omega \in \Omega$ and hence Eq. (13) is true. From the definition of $F_g(z, w)$, it is then easy to show that

$$F_g(Y_i^{(N)}(t), W_i^{(N)}(t)) - F_g(Z_i^{(N)}(t), W_i^{(N)}(t)) \xrightarrow{a.s.}_N 0.$$

So,

$$\lim_{N\to\infty} \mathbb{E}[F_g(Y_i^{(N)}(t), W_i^{(N)}(t)) - F_g(Z_i^{(N)}(t), W_i^{(N)}(t))] = 0,$$

i.e.,

$$\lim_{N\to\infty} \mathbb{E}[g(W_i^{(N)}(t+1))] = \lim_{N\to\infty} \mathbb{E}[F_g(Z_i^{(N)}(t), W_i^{(N)}(t))]. \tag{14}$$

From Eqs. (11),(12), and (14), we have

$$\lim_{N\to\infty} \mathbb{E}[g(W_i^{(N)}(t+1))] = \mathbb{E}[g(W(t+1)].$$

Since g is an arbitrary bounded mapping, this means that

$$W_i^{(N)}(t+1) \Rightarrow_N W(t+1)$$

■

Lemma 2 *If **[A:t]-[D:t]** and **[B:t+1]** hold, **[C:t+1]** also holds.*

Proof: Without loss of generality, we will just prove that $W_1^{(N)}(t+1)$ and $W_2^{(N)}(t+1)$ are asymptotically independent as N goes to infinity. **[B:t]** and **[C:t]** imply the joint convergence

$$(W_1^{(N)}(t), W_2^{(N)}(t)) \Rightarrow_N (W_1(t), W_2(t)),$$

where $W_1(t)$ and $W_2(t)$ are i.i.d r.v.s, each distributed according to $W(t)$. Let g_1 and g_2 be arbitrary bounded mappings. We then have

$$\begin{aligned} &(F_{g_1}(Z_1^{(N)}(t), W_1^{(N)}(t)), F_{g_2}(Z_2^{(N)}(t), W_2^{(N)}(t))) \\ &\Rightarrow_N \quad (F_{g_1}(Z_1(t), W_1(t)), F_{g_2}(Z_2(t), W_2(t))), \end{aligned}$$

where $(Z_1(t), W_1(t)), (Z_2(t), W_2(t))$ are i.i.d. r.v.s. By the definition of convergence in distribution, we have

$$\begin{aligned} &\lim_{N\to\infty} \mathbb{E}[F_{g_1}(Z_1^{(N)}(t), W_1^{(N)}(t))F_{g_2}(Z_2^{(N)}(t), W_2^{(N)}(t))] \\ &= \mathbb{E}[F_{g_1}(Z_1(t), W_1(t))F_{g_2}(Z_2(t), W_2(t))] \\ &= \mathbb{E}[F_{g_1}(Z_1(t), W_1(t))]\mathbb{E}[F_{g_2}(Z_2(t), W_2(t))] \\ &= \mathbb{E}[g_1(W_1(t+1))]\mathbb{E}[g_2(W_2(t+1))] \end{aligned}$$

Next, we will prove

$$\begin{aligned}&\mathbb{E}[g_1(W_1^{(N)}(t+1))g_2(W_2^{(N)}(t+1))|\mathcal{F}_t^{(N)}]\\ -\quad & F_{g_1}(Z_1^{(N)}(t),W_1^{(N)}(t))F_{g_2}(Z_2^{(N)}(t),W_2^{(N)}(t)) \xrightarrow{a.s.}_N 0.\end{aligned} \tag{15}$$

First, we can write

$$\begin{aligned}&\mathbb{E}[g_1(W_1^{(N)}(t+1))g_2(W_2^{(N)}(t+1))|\mathcal{F}_t^{(N)}]\\ =\quad & g_1(\min(W_1^{(N)}(t)+1,W_{max}))\cdot\\ & g_2(\min(W_2^{(N)}(t)+1,W_{max}))Y_{00}^{(N)}(t)\\ & +g_1(\min(W_1^{(N)}(t)+1,W_{max}))g_2(\ulcorner\frac{W_2^{(N)}(t)}{2}\urcorner)Y_{01}^{(N)}(t)\\ & +g_1(\ulcorner\frac{W_1^{(N)}(t)}{2}\urcorner)g_2(\min(W_2^{(N)}(t)+1,W_{max}))Y_{10}^{(N)}(t)\\ & +g_1(\ulcorner\frac{W_1^{(N)}(t)}{2}\urcorner)g_2(\ulcorner\frac{W_2^{(N)}(t)}{2}\urcorner)Y_{11}^{(N)}(t),\end{aligned}$$

where $Y_{00}^{(N)}(t)$ is the probability that no packet from both flow 1 and 2 is marked/dropped at time t under the σ-field $\mathcal{F}_t$ and $Y_{01}^{(N)}(t)$ is the probability that no packet from flow 1 is marked/dropped but at least one packet from flow 2 are marked/dropped etc. Similarly, we can write

$$\begin{aligned}&F_{g_1}(Z_1^{(N)}(t),W_1^{(N)}(t))F_{g_2}(Z_2^{(N)}(t),W_2^{(N)}(t))\\ =\quad & g_1(\min(W_1^{(N)}(t)+1,W_{max}))\cdot\\ & g_2(\min(W_2^{(N)}(t)+1,W_{max}))Z_1^{(N)}(t)Z_2^{(N)}(t)\\ +\quad & g_1(\min(W_1^{(N)}(t)+1,W_{max}))\cdot\\ & g_2(\ulcorner\frac{W_2^{(N)}(t)}{2}\urcorner)Z_1^{(N)}(t)(1-Z_2^{(N)}(t))\\ +\quad & g_1(\ulcorner\frac{W_1^{(N)}(t)}{2}\urcorner)\cdot\\ & g_2(\min(W_2^{(N)}(t)+1,W_{max}))(1-Z_1^{(N)}(t))Z_2^{(N)}(t)\\ +\quad & g_1(\ulcorner\frac{W_1^{(N)}(t)}{2}\urcorner)g_2(\ulcorner\frac{W_2^{(N)}(t)}{2}\urcorner)\cdot\\ & (1-Z_1^{(N)}(t))(1-Z_2^{(N)}(t)),\end{aligned}$$

To prove Eq. (15), we need to prove

$$\begin{aligned}Y_{00}^{(N)}(t)-Z_1^{(N)}(t)Z_2^{(N)}(t) &\xrightarrow{a.s.}_N 0\\ Y_{01}^{(N)}(t)-Z_1^{(N)}(t)(1-Z_2^{(N)}(t)) &\xrightarrow{a.s.}_N 0\\ Y_{10}^{(N)}(t)-(1-Z_1^{(N)}(t))Z_2^{(N)}(t) &\xrightarrow{a.s.}_N 0\\ Y_{11}^{(N)}(t)-(1-Z_1^{(N)}(t))(1-Z_2^{(N)}(t)) &\xrightarrow{a.s.}_N 0\end{aligned} \tag{16}$$

The proof of these equations is very similar. Hence, we will only prove the first one here. When the router marks/dropps packets independently,

$$Y_{00}^{(N)}(t) = Z_1^{(N)}(t) Z_2^{(N)}(t).$$

Hence, Eq. (16) is true. When packets are not marked/dropped independently, we have

$$\begin{aligned} Y_{00}^{(N)}(t) &= \frac{C_{x^{(N)}(t)-W_1^{(N)}(t)-W_2^{(N)}(t)}^{p^{(N)}(t)x^{(N)}(t)}}{C_{x^{(N)}(t)}^{p^{(N)}(t)x^{(N)}(t)}} \\ &= \frac{(1-p^{(N)}(t))\cdots(1-p^{(N)}(t)-\frac{W_1^{(N)}(t)+W_2^{(N)}(t)-1}{x^{(N)}(t)})}{1\cdots(1-\frac{W_1^{(N)}(t)+W_2^{(N)}(t)-1}{x^{(N)}(t)})}. \end{aligned}$$

Note that

$$Z_1^{(N)}(t) Z_2^{(N)}(t) = (1-p^{(N)}(t))^{W_1^{(N)}(t)+W_2^{(N)}(t)}$$

When $N \to \infty$, we also have $x^{(N)}(t) \to \infty$. Since $W_1^{(N)}(t) + W_2^{(N)}(t)$ is always less than $2W_{max}$, we can see that $Y_{00}^{(N)}(t)(\omega) - Z_1^{(N)}(t) Z_2^{(N)}(t)(\omega) \to_N 0$ for any $\omega \in \Omega$ and hence Eq. (16) is true.

Now from Eq. (15), we have

$$\begin{aligned} \lim_{N\to\infty} \mathbb{E}\Big[&\mathbb{E}[g_1(W_1^{(N)}(t+1))g_2(W_2^{(N)}(t+1))|\mathcal{F}_t^{(N)}] \\ - \ & F_{g_1}(Z_1^{(N)}(t), W_1^{(N)}(t)) F_{g_2}(Z_2^{(N)}(t), W_2^{(N)}(t))\Big] = 0, \end{aligned}$$

i.e.,

$$\begin{aligned} &\lim_{N\to\infty} \mathbb{E}[g_1(W_1^{(N)}(t+1))g_2(W_2^{(N)}(t+1))] \\ &= \lim_{N\to\infty} \mathbb{E}[F_{g_1}(Z_1^{(N)}(t), W_1^{(N)}(t)) F_{g_2}(Z_2^{(N)}(t), W_2^{(N)}(t))]. \end{aligned} \tag{17}$$

From Eq. (15) and (17), we can now get

$$\begin{aligned} &\lim_{N\to\infty} \mathbb{E}[g_1(W_1^{(N)}(t+1))g_2(W_2^{(N)}(t+1))] \\ &= \mathbb{E}[g_1(W_1(t+1))]\mathbb{E}[g_2(W_2(t+1))]. \end{aligned}$$

Since g_1 and g_2 are arbitrary bounded mappings, we conclude that $W_1^{(N)}(t+1)$ and $W_2^{(N)}(t+1)$ are asymptotically independent as N goes to infinity. ■

Lemma 3 *If* ***[A:t]-[D:t]****,* ***[B:t+1]****, and* ***[C:t+1]*** *hold,* ***[D:t+1]*** *also holds.*

Proof: Let $g : \mathbf{N} \to \mathbf{R}$ be an arbitrary mapping. Note that $W_1^{(N)}(t), \cdots, W_N^{(N)}(t)$ are exchangeable. We have

$$\begin{aligned} \mathrm{Var}[\frac{1}{N}\sum_{i=1}^{N} g(W_i^{(N)}(t+1))] &= \frac{1}{N}\mathrm{Var}[g(W_1^{(N)}(t+1))] \\ &+ \frac{N-1}{N}\mathrm{Cov}[g(W_1^{(N)}(t+1)), g(W_2^{(N)}(t+1))]. \end{aligned}$$

From **[C:t+1]**, we know that $W_1^{(N)}(t+1)$ and $W_2^{(N)}(t+1)$ are asymptotically independent. Hence,

$$\lim_{N\to\infty} \text{Cov}[g(W_1^{(N)}(t+1)), g(W_2^{(N)}(t+1))] = 0.$$

Combining this with the fact that $g(W_1^{(N)}(t+1))$ is bounded, we can see that

$$\lim_{N\to\infty} \text{Var}[\frac{1}{N}\sum_{i=1}^{N} g(W_i^{(N)}(t+1))] = 0.$$

And hence

$$\frac{1}{N}\sum_{i=1}^{N} g(W_i^{(N)}(t+1)) \xrightarrow{P}_N \mathbb{E}[g(W_1^{(N)}(t+1))]$$

$\blacksquare$

Lemma 4 *If **[A:t]-[D:t]**, **[B:t+1]**, **[C:t+1]**, and **[D:t+1]** hold, **[A:t+1]** also holds.*

Proof: From assumption $(A1)$, we have

$$\begin{aligned} p^{(N)}(t+1) &= f(\frac{S^{(N)}(t)}{N}, \frac{x^{(N)}(t+1)}{N}) \\ \frac{S^{(N)}(t+1)}{N} &= F(\frac{S^{(N)}(t)}{N}, \frac{x^{(N)}(t+1)}{N}) \end{aligned}$$

where f and F are continuous functions. Since

$$\frac{x^{(N)}(t+1)}{N} = \frac{1}{N}\sum_{i=1}^{N} W_i^{(N)}(t+1) \xrightarrow{P}_N \mathbb{E}[W(t+1)],$$

and

$$\frac{S^{(N)}(t)}{N} \xrightarrow{P}_N S(t).$$

It is easy to see that

$$\begin{aligned} p^{(N)}(t+1) &\xrightarrow{P}_N f(S(t), \mathbb{E}[W(t+1)]) = p(t+1), \\ \frac{S^{(N)}(t+1)}{N} &\xrightarrow{P}_N F(S(t), \mathbb{E}[W(t+1)]) = S(t+1) \end{aligned}$$

$\blacksquare$

B Proof of Proposition 1

We first simplify the notation and write Eq. (6) as $A(t+1) = H(A(t))$, where H is a function determined by Eq. (6). From Eq. (6), it is easy to see that H is continuous. Note that at each time step, the real system does not evolve exactly as Eq. (6). It introduces errors at each time step and these errors will be memorized in the state vector. Let $\hat{A}^{(N)}(t+1) = H(A^{(N)}(t))$ be the state vector at time $t+1$ if the real system $A^{(N)}(t)$ evolves exactly as Eq. (6), i.e., no error is introduced in time slot $[t, t+1)$. Then we expect that the real system $A^{(N)}(t+1)$ is close to $\hat{A}^{(N)}(t+1)$ when N is large. This is shown in the next Lemma (note that all lemmas in this section assume the same assumptions as Proposition 1).

Lemma 5 *Under the σ-field $\mathcal{F}_t^{(N)}$, $A^{(N)}(t+1)$ converges to $\hat{A}^{(N)}(t+1)$ in L^2 and the convergence is uniform in $A^{(N)}(t)$.*

Proof: To prove the lemma, we need to prove that for any $A^{(N)}(t)$ and $\varepsilon > 0$, there exists a constant N_0 (not dependent on $A^{(N)}(t)$), such that when $N > N_0$, $||A^{(N)}(t+1) - \hat{A}^{(N)}(t+1)||_t \leq \varepsilon$, where $||\cdot||_t = \sqrt{\mathbb{E}||\cdot|\mathcal{F}_t^{(N)}||^2}$.

We first have a look at state variable $g_i^{(N)}(t+1)$ in $A^{(N)}(t+1)$ and $\hat{g}_i^{(N)}(t+1)$ in $\hat{A}^{(N)}(t+1)$. Note that a flow with window size i at time $t+1$ has at most three possible window sizes at time t: $i-1$, $2i$, and $2i-1$ (recall that an AIMD flow either increases window size by one or halves it). So

$$g_i^{(N)}(t+1) = g_{i-1}^{(N)}(t)T_{i-1,i} + g_{2i}^{(N)}(t)T_{2i,i} + g_{2i-1}^{(N)}(t)T_{2i-1,i}, \tag{18}$$

where $T_{j,i}$ is the proportion of flows with window size j that jump to window size i. For simplicity, we will only consider the first term $g_{i-1}^{(N)}(t)T_{i-1,i}$ here. The other two terms can be treated similarly.

If there is no error introduced at time slot $[t, t+1)$, the jumping probability from $i-1$ to i will be $(1-p^{(N)}(t))^{i-1}$. Next, we will prove $g_{i-1}^{(N)}(t)T_{i-1,i}$ converges to $g_{i-1}^{(N)}(t)(1-p^{(N)}(t))^{i-1}$ uniformly in $g_{i-1}^{(N)}(t)$. The proof is simple, for any give $\varepsilon_1 > 0$, if $g_{i-1}^{(N)}(t) < \varepsilon_1$, then

$$|g_{i-1}^{(N)}(t)T_{i-1,i} - g_{i-1}^{(N)}(t)(1-p^{(N)}(t))^{i-1}| < \varepsilon_1. \tag{19}$$

If $g_{i-1}^{(N)}(t) \geq \varepsilon_1$, then the number of flows with window size $i-1$ at time t will be greater than $N\varepsilon_1$ and we can find a constant N_1 (dependent on ε_1, but not dependent on $p^{(N)}(t)$), such that when $N > N_1$,

$$||T_{i-1,i} - (1-p^{(N)}(t))^{i-1}||_t < \varepsilon_1. \tag{20}$$

We omit the proof of Eq. (20) here, since it can be proved with the similar approach as the proof of Theorem 1. From Eq. (19)(20), we have that if $N > N_1$,

$$||g_{i-1}^{(N)}(t)T_{i-1,i} - g_{i-1}^{(N)}(t)(1-p^{(N)}(t))^{i-1}||_t < \varepsilon_1,$$

no matter what $g_{i-1}^{(N)}(t)$ is.

From Eq. (18), we now have $g_i^{(N)}(t+1)$ converges to $\hat{g}_i^{(N)}(t+1)$ uniformly in $A^{(N)}(t)$ and hence $G^{(N)}(t+1)$ converges to $\hat{G}(N)(t+1)$ uniformly in $A^{(N)}(t)$. We still need to prove $\frac{S^{(N)}(t+1)}{N}$ converges to $\frac{\hat{S}^{(N)}(t+1)}{N}$ uniformly in $A^{(N)}(t)$. This is true because of the following two facts.

First, the normalized input rate $\frac{x^{(N)}(t+1)}{N}$ is a linear combination of $g_i^{(N)}(t+1)$ and hence $\frac{x^{(N)}(t+1)}{N}$ converges to $\frac{\hat{x}^{(N)}(t+1)}{N}$ uniformly in $A^{(N)}(t)$. Second, function F satisfy the Lipschitz condition Eq. (8). Hence from the uniform convergence of $\frac{x^{(N)}(t+1)}{N}$, we can derive the uniform convergence of $\frac{S^{(N)}(t+1)}{N}$. ■

Lemma 5 tells us that when N is large, one step error can be made to be small no matter what the state vector is. The next lemma states the point wise convergence of the system.

Lemma 6 *For each time t, $A^{(N)}(t)$ converges to $A(t)$ in L^2 when N goes to infinity.*

We omit the proof here since it follows the similar approach as the proof of Theorem 1. Note that if all normalized state variables are bounded (e.g., RED with a finite buffer, AVQ, and DropTail), then there exists a constant vector $Y \in L^2$, such that $||A^{(N)}(t)|| \leq Y$ for all N. The convergence in probability (Theorem 1) will simply imply convergence in L^2 (Lemma 6).

Now, we are ready to prove Proposition 1.

Proof: Let $||\cdot||_p = \sqrt{\mathbb{E}||\cdot||^2}$ be the norm in L^2 space. Note that for a constant vector, $||\cdot||_p = ||\cdot||$. We now need to prove that for any $\varepsilon > 0$, there exists a constant N_0, such that when $N > N_0$, $||A^{(N)}(t) - A(t)||_p \leq \varepsilon$ for all time t.

For any $\varepsilon_1 > 0$, since the limit model is global exponential stable, there exists a constant t_0, when $t \geq t_0$, we have

$$||A(t) - A^*|| \leq \varepsilon_1 \tag{21}$$

Note that $A(t)$ and A^* are not random variables. Also, for any sample path of the real system, from the definition of $\hat{A}^{(N)}(t+1)$ and the definition of global exponential stability (m=1), we have

$$||\hat{A}^{(N)}(t+1) - A^*|| \leq a||A^{(N)}(t) - A^*|| \tag{22}$$

for all t.

From Lemma 6, we know that there exists a constant N_1, such that when $N > N_1$,

$$||A^{(N)}(t) - A(t)||_p \leq \varepsilon_1, \tag{23}$$

for all $t \leq t_0$. So

$$||A^{(N)}(t_0) - A^*||_p \leq ||A^{(N)}(t_0) - A(t_0)||_p + ||A(t_0) - A^*||_p \leq 2\varepsilon_1. \tag{24}$$

From Lemma 5, we know that for $\varepsilon_1 > 0$, there exists a constant N_2, such that when $N > N_2$,

$$||A^{(N)}(t+1) - \hat{A}^{(N)}(t+1)||_p \leq (1-a)2\varepsilon_1. \tag{25}$$

Note that Eq. (25) is true for all t and N_2 is not dependent on t.

Now, let $N_0 = \max\{N_1, N_2\}$, we will show that when $N > N_0$, $||A^{(N)}(t) - A^*||_p \leq 2\varepsilon_1$ is true for all $t \geq t_0$. From Eq. (24), we know that it is true for $t = t_0$. Now we assume it is true for t and prove that it is also true for $t+1$ and hence by induction, it is true for all $t \geq t_0$.

$$\begin{aligned} ||A^{(N)}(t+1) - A^*||_p &\leq ||A^{(N)}(t+1) - \hat{A}^{(N)}(t+1)||_p \\ &\quad + ||\hat{A}^{(N)}(t+1) - A^*||_p \\ &\leq (1-a)2\varepsilon_1 + a||A^{(N)}(t) - A^*|| \\ &\leq (1-a)2\varepsilon_1 + 2a\varepsilon_1 = 2\varepsilon_1. \end{aligned} \tag{26}$$

Now, from Eqs (26)(21), we have that when $N > N_0$, for any $t > t_0$,

$$||A^{(N)}(t) - A(t)||_p \leq ||A^{(N)}(t) - A^*|| + ||A(t) - A^*|| \leq 3\varepsilon_1. \tag{27}$$

From Eqs (23)(27), we have that when $N > N_0$,

$$||A^{(N)}(t) - A(t)||_p \leq 3\varepsilon_1. \tag{28}$$

for all t. Simply let $\varepsilon_1 = \frac{\varepsilon}{3}$ and we are done. ■