# Rate-Based versus Queue-Based Models of Congestion Control


Supratim Deb
Laboratory for Information & Decision Systems
Massachusetts Institute of Technlogy
Cambridge, MA-02139
supratim@mit.edu

R. Srikant
Coordinated Science Lab. and Dept. ECE
University of Illinois at Urbana-Champaign
Urbana, IL-61801
rsrikant@uiuc.edu



## ABSTRACT

Mathematical models of congestion control capture the congestion indication mechanism at the router in two different ways: rate-based models, where the queue-length at the router does not explicitly appear in the model, and queue-based models, where the queue length at the router is explicitly a part of the model. Even though most congestion indication mechanisms use the queue length to compute the packet marking or dropping probability to indicate congestion, we argue that, depending upon the choice of the parameters of the AQM scheme, one would obtain a rate-based model or a rate-and-queue-based model as the deterministic limit of a stochastic system with a large number of users. We also consider the impact of implementing AQM schemes in the real queue or a virtual queue. If an AQM scheme is implemented in a real queue, we show that, to ensure that the queuing delays are negligible compared to RTTs, one is forced to choose the parameters of a AQM scheme in a manner which yields a rate-based deterministic model. On the other hand, if the AQM scheme is implemented in a virtual queue, small-queue operation is achieved independent of the choice of the parameters, thus showing a robustness property of virtual queue-based schemes.



## Categories and Subject Descriptors

H.1.0 [**Information Systems**]: Models and Principles

## General Terms

Design, Performance

## Keywords

Congestion Control, AQM parameters, Virtual Queue




## 1. INTRODUCTION

Deterministic fluid-flow models have been widely used [24, 20, 4] to describe congestion control and active queue management (AQM) schemes in the Internet. These models capture the mean behavior of the congestion controlled sources. All of these models use a packet marking (or packet dropping) function to describe the fraction of packets marked (or dropped) at a link. Depending upon the model, the marking function is either a function of the queue length [17, 15] or a function of the instantaneous arrival rate at the router [18, 24]. In this paper, we consider AQM schemes where the router decides the fraction of packets to be marked based on the occupancy level of a real or virtual queue [14]. We study appropriate models for the deterministic marking function based on the parameters of the virtual queue based AQM scheme.

Judiciously designed AQM schemes can lead to small queue lengths at the routers by providing early warning about incipient congestion to the sources accessing the network. Such mechanisms include RED [13], REM [3], and PI [17]. All of these mechanisms compute a marking or dropping probability based on some function of the queue-length at the router. In [14], the authors propose a marking scheme based on a virtual queue which operates at a capacity slightly smaller than that of the real queue and adds a packet whenever there is an arrival into the real-queue. An Adaptive Virtual-Queue (AVQ) mechanism to drive the utilization towards a given value was proposed and analyzed in [23, 22]. It has been shown that the mean behavior of congestion controlled sources can be described by deterministic fluid-flow model [29, 10, 30], which can then be used to design the AQM parameters. However, the models in [29, 10] start by assuming that the marking probability is a function of the arrival rate at the link, or assume that the parameters of the AQM scheme are chosen such that the deterministic model explicitly includes the queue length [29, 30]. In this paper, we derive deterministic models which may or may not contain the queue length in the limit, depending upon choice of the AQM parameters.

We start with a stochastic model of a single link accessed by many congestion-controlled flows. We study a limiting regime where the system capacity is scaled with the number of congestion-controlled flows. Randomness in the congestion-controlled Internet may be due to many reasons:

- unresponsive flows which do not respond to congestion indication,

- the probabilistic nature of packet marking by an AQM scheme,
- asynchronous updates among sources,
- the inability to precisely model window flow control mechanism, and
- the initial ramp-up phase (for example, *slow start* in TCP flow control) of the congestion control mechanism.

In addition to deriving deterministic model of the average flow behavior under a limiting regime where the number of sources is large, we also derive stochastic model to capture the behavior of any individual flow. We use the stochastic model to further study the performance of rate-based and queue-based models of AQM schemes.

When AQM schemes are not used, the models in [2, 27, 6] capture the randomness in the Internet, by assuming that these disturbances are independent of the arrival rate from the TCP sources. In [30], starting with a discrete-time stochastic model of a single link implementing RED which is accessed by many TCP sources, the authors have derived limiting deterministic and stochastic models of the behavior of TCP flows. With a real-queue based marking, the queue-length was shown to be $O(N)$, where $N$ is the number of flows in the system and the capacity of the link is $Nc$. In addition to capturing the suitable limiting regimes under different choices of AQM parameters, our objective is to obtain models for congestion control/AQM schemes which achieve queue lengths that are $O(1)$, thus leading to a negligible queuing-delays as compared to RTTs. In [28, 19, 30], continuous-time stochastic models (OU processes) were proposed for TCP flows to capture the variance due to probabilistic marking. These models capture the deviations in the users' arrival rate at times scales that are comparable to the sources' RTTs. In this paper, we also study the queue-length dynamics at the routers at a time scale which is of the order of the time required to process a packet at the router.

For the purposes of this paper, we primarily study REM as the AQM scheme at the router. As we will comment later, the results should be applicable to a large class of queue-length based AQM schemes such as RED and its variants.

Our main contributions are as follows.

1. Depending upon the manner in which a parameter in REM is scaled with the number of flows, we argue that the limiting deterministic/stochastic model of the congestion-controlled link would capture the AQM behavior using either a rate-based or a jointly rate-and-queue-based marking function. The choice of the appropriate model for the marking function is critical in designing the parameters of the congestion control/AQM scheme.
2. We show that a virtual-queue based AQM scheme is very robust to the choice of parameter in terms of attaining a low-loss, low-delay and high-utilization operation. On the other hand, with real-queue based marking, the choice of parameter is critical if we want the queuing delay to be negligible when compared to the propagation delay. The analytical models are also corroborated through simulations.

The rest of the paper is structured as follows. The model and the problem statement is provided in Section 2. In Section 3, we discuss two different regimes of AQM parameters and obtain the appropriate models for each of these regimes. The discussion in this section is heuristic, and we only provide arguments to show that these limits are plausible. In a longer version of the paper we provide formal proofs of the limiting models. In Section 4 we discuss the equilibrium and the stability properties of a proportionally fair controller in the two different regimes. In Section 5, we precisely show the assumption which leads to deterministic models of TCP widely used in the literature and show the equilibrium values of the mean rate. In Section 6, we obtain the variance of the arrival process in the link time-scale. In Section 7, we validate our design and analysis with simulation results. Finally, we provide concluding remarks in Section 8.

# 2. BASIC FRAMEWORK AND PROBLEM STATEMENT

Our model is that of a single bottleneck link being accessed by many congestion controlled flows and uncontrolled flows. We consider discrete-time rate update model for the flows to model *TCP* and *proportionally fair* congestion controllers. The rate update mechanisms are described later in this section. The delay in the forward and the reverse path is $r/2$ sec so that the round-trip propagation delay is $r$ sec. The number of flows in the system is $N$, which is also the scaling parameter. We consider a sequence of such systems indexed by $N$, where in the $N^{th}$ system, there are $N$ congestion controlled flows accessing the link. Further, in the $N^{th}$ system there are $N$ uncontrolled flows accessing the link. The link capacity is scaled as $Nc$ packets per second. In this paper we consider AQM schemes which mark packets based on the queue length of a real or a virtual queue. If the marking is based on virtual queue, we denote the capacity of the virtual queue in the $N^{th}$ system by $N\tilde{c} = N\theta c$. For modeling purposes, we assume that the marking is based on the virtual-queue. We can recover the case with real-queue by setting $\theta = 1$. We consider a slotted system, where the length of a time-slot is the time to serve one packet in the virtual queue which is $1/(N\tilde{c})$ s. Further, we denote by $\hat{q}^{(N)}[k]$ the length of the virtual queue at the $k^{th}$ slot in the $N^{th}$ system. The evolution of the virtual-queue can be described by the following.

$$\hat{q}^{(N)}[k+1] = (\hat{q}^{(N)}[k] - 1)^{+} + A_c^{(N)}[k] + \sum_{i=1}^{N} (e_i^{(N)}(\tfrac{1}{N\tilde{c}})) \quad (1)$$

In the above, $A_c^{(N)}[k]$ denotes the number of arrivals due to the controlled flows in a slot and $e_i^{(N)}(t)$ is a stationary stochastic process denoting the number of arrivals due to an uncontrolled flow in a time interval of length $t$ sec in the $N^{th}$ system. We also denote by $e^{(N)}(t)$, the average number of arrivals due to the uncontrolled flows in time $t$ sec so that

$$e^{(N)}(t) = \frac{1}{N} \sum_{i=1}^{N} e_i^{(N)}(t) \ .$$

## 2.1 Rate update mechanism

In the following, we assume that all the rates are measured in packets per second. Let $y_i^{(N)}[k]$ denote the flow rate of the $i^{th}$ controlled flow at time slot $k$ in the $N^{th}$ system. Further, let $x^{(N)}[k]$ denote the average flow rate of the controlled flows through the link at time $k$, and so

$$x^{(N)}[k] = \frac{1}{N}\sum_{i=1}^{N} y_i^{(N)}[k] \ .$$

In a window based implementation, the rate update interval of the congestion-controlled flows is taken to be the time between two successive ACKs, which is roughly equal to the inverse of the equilibrium rate seen by the controlled flow. Since $N\tilde{c}$ is the capacity of the virtual queue, the rate update interval will be approximately equal to $1/\tilde{c}$. Thus $y_i^{(N)}[k]$ is updated once every $N$ slots in the $N^{th}$ system. We also introduce the variable $\tilde{y}_{i,l}^{(N)}$ to denote the rate of the $i^{th}$ source in the $N^{th}$ system after the $l^{th}$ update and so

$$\tilde{y}_{i,l}^{(N)} = y_i^{(N)}[Nl].$$

Denote by $\tilde{x}_l^{(N)}$ the following quantity.

$$\tilde{x}_l^{(N)} = \frac{1}{N}\sum_{i=1}^{N} \tilde{y}_l^{(N)} \ .$$

Denote by $M_i^{(N)}[N(k+1)]$ the number of packets of source $i$ marked by the link in the slot interval $[Nk+1, N(k+1)]$. Let $\tau$ denote the rate update interval of the controlled flows so that $\tau = 1/\tilde{c}$. Since the slot length in the $N^{th}$ system is $1/(N\tilde{c})$, the round trip propagation delay of the controlled flows in the $N^{th}$ system is clearly $Nd$ slots, where $d = r/\tau$. Note that $d$ can also be viewed as the number of updates of source $i$ in a round trip time. For simplicity we will assume that $d$ is an even integer.

Before we introduce the congestion control models with stochasticity due to different sources of randomness, we first have a look at the deterministic models of congestion control used in the literature. This will help us to understand the quantity which in reality cannot be modeled in a deterministic fashion. Suppose, in the interval $[Nk+1, N(k+1)]$, the link marks $p_{det}[N(k+1)]$ fraction of the packets. Then, the deterministic rate update model with a proportionally fair controller is given by

$$y_i^{(N)}[N(k+1)] - y_i^{(N)}[Nk] = \kappa[w - y_i^{(N)}[N(k-d+1)]p_{det}[N(k+1-\tfrac{d}{2})]] \ .$$

However, the quantity $y_i^{(N)}[N(k-d+1)]p_{det}[N(k+1-\frac{d}{2})]$, which denotes the rate at which marked packets are received, is not deterministic. In fact, $y_i^{(N)}[N(k-d+1)]p_{det}[N(k+1-\frac{d}{2})]$ should be replaced by number of marked packets received (which is a random variable) in the rate update interval divided by the length of the rate update interval.

We note that the number of marked packets received by source $i$ in the interval $[Nk+1, N(k+1)]$ (i.e., between the $k^{th}$ and the $(k+1)^{th}$ rate update interval) is simply the random variable $M_i^{(N)}[N(k+1-d/2)]$, since the one way propagation delay is $Nd/2$ slots. With that in mind, we next describe two popular congestion controllers in the following.

• **Proportionally fair controller:** With proportionally fair congestion controller, the update of the $i^{th}$ source is described by

$$y_i^{(N)}[N(k+1)] - y_i^{(N)}[Nk] = \kappa(w - \frac{1}{\tau}M_i^{(N)}[N(k+1-\tfrac{d}{2})]) \ ,$$

$$k = 0, 1, 2 \ldots . \quad (2)$$

$$y_i^{(N)}[Nk] = y_i^{(N)}[Nk+l] \quad \text{for } l \in \{0, 1, 2, \ldots N-1\} \quad (3)$$

Since there are $d$ updates in a round trip time, the marks received between the $k^{th}$ and the $(k+1)^{th}$ update is a fraction of packets sent after the $(k-d+1)^{th}$ update (and before the $(k-d+2)^{th}$ update) which are $\tau y_i^{(N)}[N(k-d+1)]$ in number. If we substitute

$$M_i^{(N)}[N(k+1-\tfrac{d}{2})] = \tau p_{det}[N(k+1-\tfrac{d}{2})]y_i^{(N)}[N(k+1-d)] \ ,$$

which is essentially replacing $M_i^{(N)}[N(k+1-\frac{d}{2})]$ with its mean, we get the familiar form of a *proportionally fair* controller. However, $M_i^{(N)}[Nk]$ is random and is determined by the marking mechanism at the router, and we will discuss a model for this in the next section. We note that, if the updated rates become negative, the updated rates should be projected into the positive axis. However, for the purposes of the derivations in this paper, we do not show this explicitly in the rate update of the individual flows. It was shown in [29] for a continuous time version of the system with unresponsive flows but without probabilistic marking that, in a many flows regime, the average flow rate behaves as though the non-negativity constraint is ignored. Thus, in deriving the behavior of the average flow rate, we ignore the non-negativity constraint.

• **TCP-like or minimum potential delay controller:** We use the following model for the rate update of TCP:

$$y_i^{(N)}[N(k+1)] - y_i^{(N)}[Nk] = \kappa(w - \frac{1}{\tau}y_i^{(N)}[Nk]M_i^{(N)}[N(k+1-\tfrac{d}{2})]) \ . \quad (4)$$

The standard TCP is recovered with $w = 1/r^2$. Again, if we substitute $M_i^{(N)}[N(k+1-\frac{d}{2})]$ with its mean so that

$$M_i^{(N)}[N(k+1-\tfrac{d}{2})] = \tau p_{det}[N(k+1-\tfrac{d}{2})]y_i^{(N)}[N(k+1-d)] \ ,$$

we get the rate update model for congestion avoidance phase of TCP as proposed in [24].

## 2.2 Queue-based marking

In a queue-based marking mechanism, packets are marked based on the queue length [13, 3, 17, 1, 32, 25]. The queue length could be the length of the real queue or the length of a virtual queue. The notion of a virtual queue was first introduced in [14]. A packet is added to the virtual queue whenever there is a packet arrival into the real queue, but packets are drained from the virtual queue at a rate $\tilde{C} = N\tilde{c} = N\theta c$ ($\theta \leq 1$). If $\theta < 1$, the capacity of the virtual queue is smaller than the capacity of the real queue. Thus, marking based on the queue length at the virtual queue helps to detect incipient congestion. Throughout the paper, we use $\hat{q}$ to denote the length of virtual-queue, and $q$ to denote the length of the real-queue. We also use $f(\hat{q})$ to denote the marking profile at the link which depends on the length of the virtual queue $\hat{q}$.

As mentioned before, we will assume that marking is based on the length of the virtual queue. We can recover the case of marking based on the real queue by letting $\theta = 1$. In the case of virtual-queue-based marking, the RTT can be well-approximated by the propagation delay since the queuing delay will be negligible (we also demonstrate this later in our results). However, in the case of real-queue-based marking, the RTT is the sum of the propagation delay and the queuing delay, which is variable. Nevertheless, for our modeling purposes, in that case, we still approximate the RTT by a constant (equal to the sum of the propagation delay and the equilibrium queuing delay).

In a virtual-queue based REM, a packet is marked with probability

$$f(\hat{q}) = 1 - \exp(-\gamma^{(N)}\hat{q})$$

when the length of the virtual queue is $\hat{q}$. The critical parameter in REM is $\gamma^{(N)}$. In this paper, we consider two regimes of parameters to study the robustness of a virtual-queue mechanism to the choice of parameters. In the context of REM, we consider $\gamma^{(N)} = \gamma/N$ and $\gamma^{(N)} = \gamma$.

In addition to REM, the analysis can be applied to other AQM schemes like RED [12]. The marking profile for RED is given by

$$f(\hat{q}) = \min(1,\ \hat{\alpha}(\hat{q} - \hat{q}_{min})^{+})$$

where, $\hat{\alpha}$ and $\hat{q}_{min}$ are parameters of the marking function. Thus, our analysis can be applied to RED with the choice of $\hat{\alpha}, \hat{q}_{min}$ as $\hat{\alpha} = \alpha/N, \hat{q}_{min} = Nq_{min}$ and, $\hat{\alpha} = \alpha, \hat{q}_{min} = q_{min}$.

## 2.3 Model for packet arrivals at the router

The queuing phenomenon at the router happens at a faster time scale compared to the rate update mechanism at the sources. As mentioned earlier, the sources' rate updates occur approximately once every $\tau = 1/\tilde{c}$ s, whereas the time to process a packet at the router is $1/Nc$ s. Due to the features of the window flow control mechanism that we have not modeled, there can be a lot variability in the number of packets received at the router at time scales of the order of $1/Nc$. For example, the window implementation of flow control implies that the rate updates of the sources are not in regular intervals (we use $\tau = 1/\tilde{c}$ as roughly the mean update interval at the sources). Further, the sources are not synchronized; thus, the packets from a source can arrive at the link at any point of time within a slot. We make the following assumption to capture the variability of the arrival rate into the router due to these additional factors not reflected in the discrete-time rate update model.

ASSUMPTION 1. *(Arrival process at the link) If the arrival rate of the $i^{th}$ controlled flow is $x_i$, the number of packets arriving at the link over a time interval of length $1/(N\tilde{c})$ is Poisson distributed with mean $x_i/(N\tilde{c})$.* □

Our simulation results later in the paper support this assumption. Please note that the above assumption does not assume that the arrival rates due to the controlled flows are Poisson. The arrival rates of the controlled flows are time-varying quantitities and are updated once every $\tau$ time. Assumption 1 models the packet arrival at the link between any two updates of the controlled flow rate.

## 2.4 Problem statement

We attempt to answer the following question in this paper.

- With REM as a marking mechanism, given a choice of $\gamma^{(N)}$ which can be either $\gamma^{(N)} = \gamma/N$ or $\gamma^{(N)} = \gamma$, what is the suitable model for the marking function to predict the steady state properties like mean and variance of the arrival process into the link?

We will also provide design rules to guarantee low-loss and low queuing-delay at the link, along with high utilization. Specifically, we are interested in developing models to provide guidelines to answer the following questions:

- Suppose the system capacity is fixed. How should we choose the parameters of the marking function so that the probability of buffer overflow is smaller than a given target?
- Suppose the parameters of the marking function are fixed. What should be the link capacity so that the probability of buffer overflow is smaller than a given target?

# 3. QUEUE-BASED MARKING AND SCALING REGIMES

We recall that REM [3] marks a packet with probability $(1 - \exp(-\gamma^{(N)}\hat{q}))$, if there are $\hat{q}$ packets in the virtual queue. The crucial parameter in REM mechanism is $\gamma^{(N)}$. Prescriptions for the values of $\gamma^{(N)}$ are provided in [5] based on extensive simulations. Roughly speaking, a larger value of $\gamma^{(N)}$ leads to faster convergence of the rate-control mechanism, but at the cost of smaller utilization at the link. To this end, we consider two choices for the parameter $\gamma^{(N)}$, $\gamma^{(N)} = \gamma/N$ and $\gamma^{(N)} = \gamma$.

We describe the stochastic models with proportionally fair controller and TCP-like controller both.

## 3.1 REM with $\gamma^{(N)} = \gamma/N$

Let $\overline{q}^{(N)}[k] = \hat{q}^{(N)}[k]/N$ for the purpose of this subsection. We also call $\overline{q}^{(N)}$ the normalized virtual-queue length. With $\gamma^{(N)} = \gamma/N$, in this parameter regime, the marking probability can be expressed as a function of $\overline{q}^{(N)}$ as follows:

$$p(\overline{q}^{(N)}) = 1 - \exp(-\gamma\overline{q}^{(N)})\ . \quad (5)$$

### 3.1.1 *Proportionally fair controller*

The number of marks received by a source between two updates in this parameter regime depends on $\overline{q}^{(N)}[k]$ over a rate update interval. We show that the number of marks received by a source in this regime can be expressed as a function of $\overline{q}^{(N)}$ at instants $Nk$ $(k = 1, 2, \ldots)$ and the average flow rate of the controlled flows. To derive the marking function, we assume that the rate-update epochs of controlled flows are synchronized. Consider a time interval of length $\tau$ (the rate update interval of the controlled flows) over which the rates of the controlled flows do not change. To this end, let $x$ be the average flow rate of the controlled flows, and let $x_i$ be the average flow rate of the $i^{th}$ flow into the link in the time interval under consideration. Now, note that the interval of length $\tau$ s has $N$ slots. Let those slots be indexed by $m, m+1, m+2, \ldots m+N-1$. Recall that, by Assumption 1, the number of arrivals from the $i^{th}$ controlled flow in a slot is Poisson$(x_i/(N\tilde{c}))$. Further, let $q_0$ be the value of $\overline{q}^{(N)}[\cdot]$ at the beginning of the slot under consideration. In other words $\overline{q}^{(N)}[m] = q_0$. We want to derive the fraction of packets marked in this slot as a function of $x$ and $q_0$.

In deriving the model, we assume that the number of arrivals from the $N$ uncontrolled flows in a slot is Poisson$(a/\tilde{c})$ for simplicity, where $a$ is the mean flow rate of any of the $N$ uncontrolled flows. This assumption can be relaxed, but is reasonable in the following sense. It is well-known that, if there are $N$ identical stationary point processes, then the process obtained by superposing the $N$ processes and dilating the time scale by a factor of $N$ goes to a Poisson process in the limit of large $N$ [8]. Thus, the assumption that the number of packets from the $N$ controlled flows over a time-interval of length $1/(N\tilde{c})$ is Poisson, is reasonable.

We are interested in obtaining the number of packets from the $i^{th}$ source which are marked over the time interval of length $\tau$. The virtual queue length process evolves as discrete-time queue with service rate one per slot and the number of arrivals in a slot distributed as Poisson$((x+a)/\tilde{c})$. Let $m_i^{(N)}[k]$ denote the number of packets of source $i$ marked in slot $k$ ($k = m, m+1, m+2, \ldots m+N-1$). If the packets are marked upon arrival at the virtual-queue, then $m_i^{(N)}[k]$ is distributed as Poisson$(p(\overline{q}^{(N)}[k])x_i/(Nc))$ which follows from the fact that a probabilistic splitting of a Poisson random variable gives a Poisson random variable. Here $p(\overline{q}^{(N)}[k])$ is the marking probability which in this parameter regime is given by (5). We are interested in finding the distribution of $M_i^{(N)} = \sum_{k=m}^{m+N-1} m_i^{(N)}[k]$. Define $q^{(N)}(t)$ as follows:

$$q^{(N)}(t) = \overline{q}^{(N)}[\lfloor N\tilde{c}t \rfloor + m] = \frac{\hat{q}^{(N)}[\lfloor N\tilde{c}t \rfloor + m]}{N} \ . \tag{6}$$

It is well known that [7], in the limit of large $N$, the process $q^{(N)}(t)$ behaves like a deterministic process (in the interval under consideration). Specifically, we have

$$\lim_{N\to\infty} q^{(N)}(t) = q(t) \text{ u.o.c.,} \tag{7}$$

$$q(t) = (q_0 + t(x+a-\tilde{c}))^+ \tag{8}$$

In the above, the convergence is in an almost sure sense and is uniformly over compact sets.

We now apply the Poisson limit theorem for triangular arrays as given in [8]. Suppose we can verify the following conditions:

$$\lim_{N\to\infty} \sum_{k=m}^{m+N-1} \Pr(m_i^{(N)}[k] \geq 2) = 0 \tag{9}$$

$$\lim_{N\to\infty} \sum_{k=m}^{m+N-1} \Pr(m_i^{(N)}[k] \geq 1) = \mu \ . \tag{10}$$

Then, $M_i^{(N)}$ converges weakly (in distribution) to Poisson$(\mu)$ for large $N$ [8]. Condition (9) can be easily verified. To verify (10), note the following:

$$\begin{aligned}
&\lim_{N\to\infty} \sum_{k=m}^{m+N-1} \Pr(m_i^{(N)}[k] \geq 1) \\
&= \lim_{N\to\infty} \sum_{k=m}^{m+N-1} (1 - \exp(-\frac{x_i}{N\tilde{c}} p(\overline{q}^{(N)}[k]))) \\
&= \lim_{N\to\infty} \sum_{k=0}^{N-1} (1 - \exp(-\frac{x_i}{N\tilde{c}} p(\overline{q}^{(N)}[k+m]))) \\
&= \lim_{N\to\infty} \sum_{k=0}^{N-1} (\frac{x_i}{N\tilde{c}} p(\overline{q}^{(N)}[k+m]) + O\left(\frac{1}{N^2}\right)) \\
&= x_i \lim_{N\to\infty} \frac{1}{N\tilde{c}} \sum_{k=0}^{N-1} p(\overline{q}^{(N)}[k+m]) \\
&= x_i \lim_{N\to\infty} \frac{1}{N\tilde{c}} \sum_{k=0}^{N-1} p(q^{(N)}\left(\frac{k}{N\tilde{c}}\right)) \\
&= x_i \lim_{N\to\infty} \frac{1}{N\tilde{c}} \sum_{k=0}^{N-1} p(q\left(\frac{k}{N\tilde{c}}\right)) \\
&+ x_i \lim_{N\to\infty} \frac{1}{N\tilde{c}} \sum_{k=0}^{N-1} [p(q^{(N)}\left(\frac{k}{N\tilde{c}}\right)) - p(q\left(\frac{k}{N\tilde{c}}\right))] \ .
\end{aligned}$$

The second term in the preceding expression goes to zero, since $q^{(N)}(t)$ converges to $q(t)$ uniformly over the interval $(0, \tau)$ and further since the marking profile, $p(\cdot)$ is bounded from above by 1. The first term in the expression converges to an appropriate integral and so we have the following in the limit of large $N$:

$$\begin{aligned}
&\lim_{N\to\infty} \sum_{k=m}^{m+N-1} \Pr(m_i^{(N)}[k] \geq 1) \\
&= x_i \int_0^{1/\tilde{c}} p(q(s))ds \\
&= x_i \int_0^{1/\tilde{c}} p((q_0 + s(x+a-\tilde{c}))^+)ds \\
&= \tau x_i p_q(q_0, x)
\end{aligned}$$

In the above,

$$\begin{aligned}
p_q(q_0, x) = 1 - \exp(-\gamma q_0) &\frac{\exp\left(\gamma \min\left(q_0, 1 - \frac{x+a}{\tilde{c}}\right)\right) - 1}{\gamma\left(1 - \frac{x+a}{\tilde{c}}\right)} \\
&- \left(1 - \frac{\min\left(q_0, 1 - \frac{x+a}{\tilde{c}}\right)}{1 - \frac{x+a}{\tilde{c}}}\right)
\end{aligned} \tag{11}$$

We thus have

$$M_i^{(N)} \sim \tau x_i p_q(q_0, x) \ .$$

Since $\tau x_i$ is the mean number of packets sent by source $i$ in the rate update interval, $p_q(q_0, x)$ can be viewed as the equivalent marking function in this parameter regime, which is a function of the normalized virtual-queue length $\overline{q}^{(N)}$ at instants $Nk$ ($k = 1, 2, \ldots$) and the average rate $x$.

REMARK 1. *In the typical operating regime, $\gamma$ is a small number and the average rate of the controlled flows is close to the target equilibrium value of $\tilde{c} - a$ and so the equivalent marking function can be approximated by*

$$p_q(q, x) \approx 1 - \exp(-\gamma q) \ . \tag{12}$$

*Thus the marking function depends solely on the normalized queue length at the boundaries of the rate update interval. This modeling assumption has been widely used in the literature.* □

Having derived the behavior of the equivalent marking function, we now derive models for the individual flow rates and the average flow rate as $N \to \infty$.

Since the fraction of packets marked depends on the normalized virtual queue length through the values at slots $Nk$ for $k = 1, 2, 3 \ldots$, the system evolution with a proportionally fair controller can be given by the following for large $N$.

$$y_i^{(N)}[N(k+1)] - y_i^{(N)}[Nk] = \kappa \left[w - \frac{1}{\tau} M_i^{(N)}[N(k+1-\tfrac{d}{2})]\right] \tag{13}$$

where

$M_i^{(N)}[Nk] \sim$
$\text{Poisson}(\tau y_i^{(N)}[N(k-\frac{d}{2})]p_q(\overline{q}^{(N)}[Nk], x^{(N)}[N(k-\frac{d}{2})]))$ .

We now derive models for the mean rate $x^{(N)}[Nk]$ and $y_i^{(N)}[Nk]$ for large $N$. Denote by $\tilde{x}_l^{(N)}$ the process $x^{(N)}[Nl]$ and by $\tilde{y}_{i,l}^{(N)}$ the process $y_i^{(N)}[Nl]$. These are the quantities denoting the values after the $l^{th}$ update. Also, denote by $\tilde{q}_l^{(N)}$ the process $\overline{q}^{(N)}[Nl]$. Clearly, $\tilde{q}_l^{(N)}$ denotes the value of $\overline{q}^{(N)}[k]$ at the boundaries of the rate-update intervals of the sources.

Suppose $\tilde{x}_l^{(N)}$ converges to $\tilde{x}_l$ for large $N$ in an appropriate sense. Note that, using (7)-(8), we can describe the values of $\overline{q}^{(N)}[k]$ at the boundaries of the rate-update intervals. Thus, $\tilde{q}_l^{(N)} \to \tilde{q}_l$ for large $N$, where

$$\tilde{q}_{(l+1)} = (\tilde{q}_l + \tau(x_{(l-\frac{d}{2})} + a - \tilde{c}))^+ \tag{14}$$

Then, we further have from (13), $\tilde{y}_{i,l}^{(N)} \to \tilde{y}_l$ for large $N$, where

$$\tilde{y}_{(l+1)} - \tilde{y}_l = \kappa\left[w - \frac{1}{\tau}M_{(l+1-\frac{d}{2})}\right] \tag{15}$$

where

$$M_l \sim \text{Poisson}(\tau\tilde{y}_{(l-\frac{d}{2})}p_q(\tilde{q}_l, \tilde{x}_{(l-\frac{d}{2})}))\ .$$

Further, since $\tilde{y}_{i,l}^{(N)} \to \tilde{y}_l$,

$$\tilde{x}_l^{(N)} = \frac{1}{N}\sum_{i=1}^{N}\tilde{y}_{i,l}^{(N)} \to \mathbb{E}[\tilde{y}_l] \tag{16}$$

by the law of large numbers. Since, $\tilde{x}_l^{(N)} \to \tilde{x}_l$, we have $\mathbb{E}[\tilde{y}_l] = \tilde{x}_l$. We thus have the following limiting equation for $\tilde{x}_l$ upon taking expectations in the update of $\tilde{y}_l$:

$$\tilde{x}_{(l+1)} - \tilde{x}_l = \kappa\left[w - \frac{1}{\tau}\mathbb{E}[M_{(l+1-\frac{d}{2})}]\right]$$

$$\Rightarrow \tilde{x}_{(l+1)} - \tilde{x}_l = \kappa\left[w - \tilde{x}_{(l+1-d)}p_q(\tilde{q}_{(l+1-\frac{d}{2})}, \tilde{x}_{(l+1-d)})\right]\ .$$

Thus, the complete set of equations describing the system is given by the following:

$$\tilde{y}_{(l+1)} - \tilde{y}_l = \kappa\left[w - \frac{1}{\tau}M_{(l+1-\frac{d}{2})}\right] \tag{17}$$

$$M_l \sim \text{Poisson}(\tau\tilde{y}_{(l-\frac{d}{2})}p_q(\tilde{q}_l, \tilde{x}_{(l-\frac{d}{2})})) \tag{18}$$

$$\tilde{x}_{(l+1)} - \tilde{x}_l = \kappa\left[w - \tilde{x}_{(l+1-d)}p_q(\tilde{q}_{(l+1-\frac{d}{2})}, \tilde{x}_{(l+1-d)})\right] \tag{19}$$

$$\tilde{q}_{(l+1)} = (\tilde{q}_l + \tau(x_{(l-\frac{d}{2})} + a - \tilde{c}))^+\ . \tag{20}$$

The above equations completely characterize the individual flow behavior $\tilde{y}_l$ and also the average flow behavior $\tilde{x}_l = \mathbb{E}[\tilde{y}_l]$ for all $l$. As noted before, we have ignored that the updated rates $\tilde{y}_l$'s are always projected into the positive real line. However, this can be taken into account to model the behavior of $\tilde{y}_l$'s. It was shown in [29] for a continuous time version of the system with unresponsive flows but without probabilistic marking that, in a many flows regime, the average flow rate behaves as though the non-negativity constraint is ignored.

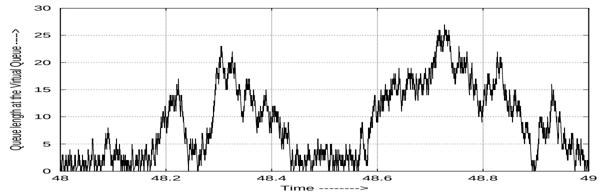


**Figure 1: Virtual queue length with marking function $f(\hat{q}) = 1 - \exp(-0.0075\hat{q})$. $N = 25$ flows, $C = 2000$ packets/s, $RTT = 0.1$ s.**

### 3.1.2 TCP-like controller

We now show similar models with a TCP-like controller. The equivalent marking function, $p_q(q, x)$, as a function of the normalized queue length at the boundaries of the rate update interval and the average flow rate of the controlled flows takes a similar form as in Section 3.1.1. However, the rate update equations are different in this case. Starting with (4) and going through similar calculations as in Section 3.1.1, one can argue that, as $N \to \infty$, $\tilde{x}_l^{(N)} \to \tilde{x}_l$, $\tilde{y}_{i,l}^{(N)} \to \tilde{y}_l$ and $\tilde{q}_l^{(N)} \to \tilde{q}_l$, where

$$\tilde{y}_{(l+1)} - \tilde{y}_l = \kappa\left[w - \frac{1}{\tau}\tilde{y}_{(l+1-d)}M_{(l+1-\frac{d}{2})}\right] \tag{21}$$

$$M_l \sim \text{Poisson}(\tau\tilde{y}_{(l-\frac{d}{2})}p_q(\tilde{q}_l, \tilde{x}_{(l-\frac{d}{2})})) \tag{22}$$

$$\tilde{x}_{(l+1)} - \tilde{x}_l = \kappa\left[w - \mathbb{E}[\tilde{y}_l\tilde{y}_{(l+1-d)}]p_q(\tilde{q}_{(l+1-\frac{d}{2})}, \tilde{x}_{(l+1-d)})\right] \tag{23}$$

$$\tilde{q}_{(l+1)} = (\tilde{q}_l + \tau(x_{(l-\frac{d}{2})} + a - \tilde{c}))^+\ . \tag{24}$$

The model above is nearly identical to the one in [30] except for the fact that we distinguish between the queue length and number of packets in a window. This distinction is important when using this model to predict queue lengths accurately.

## 3.2 REM with $\gamma^{(N)} = \gamma$

We now consider a regime when $\gamma^{(N)}$ is not scaled with the number of flows in the system. We comment that, in [33], equivalent marking function in a similar regime has been obtained for a continuous time model of rate update schemes (so that the rate updates can be descibed by differential equations) that only takes randomness due to uncontrolled flows into account.

First, we show some simulation results to indicate that the virtual-queue length process fluctuates a lot in this scaling-regime. The behavior of the virtual queue length in this regime for a particular choice of parameters is shown in Figure 1. Note that the virtual-queue length process fluctuates a lot and can hit zero multiple times in a slot (rate update interval). Thus, it is reasonable to expect that the behavior of the system cannot simply be described in terms of the queue-length at the boundaries of the rate-update intervals.

### 3.2.1 Proportionally fair controller

We argue that the appropriate model for the marking function in this parameter regime is a rate-based model and explicitly derive the model. Our starting point is as in the previous subsection. We assume that the rate-update epochs of controlled flows are synchronized. Consider a time interval of length $\tau$ (the rate update interval of the controlled flows) over which the rates of the controlled flows do not change. Let $x$ be the average flow rate of the controlled flows, and let $x_i$ be the average flow rate of the $i^{th}$ flow in the time interval under consideration. By Assumption 1, the number of arrivals from the $i^{th}$ controlled flow in a slot is Poisson$(x_i/(N\tilde{c}))$. We assume that the number of arrivals from the $N$ uncontrolled flows in a slot is Poisson$(a/\tilde{c})$.

We are interested in obtaining the number of packets from the $i^{th}$ source which are marked over the time interval of length $\tau$. Index the slots by $m, m+1, m+2 \ldots m+N-1$. The virtual queue length process evolves as discrete-time queue with service rate one per slot and the number of arrivals in a slot distributed as Poisson$((x+a)/\tilde{c})$. The number of packets of source $i$ marked in slot $k$, $m_i^{(N)}[k]$, is distributed as Poisson$(f(\hat{q}^{(N)}[k]x_i/(Nc)))$. Here $f(\hat{q})$ is the marking probability which in in this parameter regime is given by

$$f(\hat{q}) = 1 - \exp(-\gamma\hat{q}) \ .$$

We are interested in finding the distribution of $M_i^{(N)} = \sum_{k=1}^{N\tilde{c}\tau} m_i^{(N)}[k]$. We again apply the Poisson limit theorem for triangular arrays as given in [8]. We want to verify the conditions given by (9)-(10). The condition given by (9) can be easily verified. To verify (10), we proceed in a manner similar to that in the previous subsection.

$$\begin{aligned}
&\lim_{N\to\infty} \sum_{k=m}^{m+N-1} \Pr(m_i^{(N)}[k] \geq 1) \\
&= \lim_{N\to\infty} \sum_{k=m}^{m+N-1} (1 - \exp(-\frac{x_i}{N\tilde{c}} f(\hat{q}^{(N)}[k]))) \\
&= \lim_{N\to\infty} \sum_{k=0}^{N-1} (\frac{x_i}{Nc} f(\hat{q}^{(N)}[k+m]) + O\left(\frac{1}{N^2}\right)) \\
&= \tau x_i \lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} f(\hat{q}^{(N)}[k+m]) \\
&= x_i \tau \mathbb{E}_\pi[f(\hat{q}_\infty)]
\end{aligned}$$

The last step follows from the ergodic theorem and $\hat{q}_\infty$ denotes the stationary behavior of the queue-length process $\hat{q}^{(N)}[\cdot]$. Thus, we conclude that the number of marks received by the $i^{th}$ controlled flow, when it has a flow rate $x_i$, is Poisson$(\tau x_i \mathbb{E}_\pi[f(\hat{q}_\infty)])$. Further, the expectation is with respect to the steady-state queue-length distribution $\pi$ of a discrete time queue with Poisson arrivals in each time-unit. It can be shown from [21] that, $\pi$ has a transform domain representation given by

$$\tilde{\Pi}(z) = \mathbb{E}_\pi[z^{\hat{q}_\infty}] = \frac{\exp(\rho(z-1))(1-\rho)(1-z)}{\exp(\rho(z-1)) - z} \ , \tag{25}$$

where $\rho = \min(1, (x+a)/\tilde{c})$. The above expression can be used to compute $\mathbb{E}_\pi[f(\hat{q}_\infty)]$ for any given marking function. In the case of REM, the equivalent marking profile $\overline{p}(x)$ is given by

$$\begin{aligned}
\overline{p}(x) &= \mathbb{E}_\pi[1 - \exp(-\gamma\hat{q}_\infty)] \\
&= 1 - \frac{\exp\left(\frac{x+a}{\tilde{c}}(e^{-\gamma}-1)\right)\left(1-\frac{x+a}{\tilde{c}}\right)^+(1-e^{-\gamma})}{\exp\left(\frac{x+a}{\tilde{c}}(e^{-\gamma}-1)\right) - e^{-\gamma}} \ .
\end{aligned}$$

We have

$$M_i^{(N)} \sim \tau x_i \overline{p}(x) \ ,$$

and hence, the equivalent marking function $\overline{p}(x)$ in this parameter regime is simply a function of the average rate.

We can again derive the limiting models for the mean rate $x^{(N)}[Nk]$ and $y_i^{(N)}[Nk]$ for large $N$ in a similar vain as in Section *3.1.1*. Let $\tilde{x}_l^{(N)}$ converge to the stochastic process $\tilde{x}_l$ for large $N$. Then, $\tilde{y}_{i,l}^{(N)}$ also converges to $\tilde{y}_l$ for all $i$ and further $\mathbb{E}[\tilde{y}_l] = \tilde{x}_l$ by strong law of large numbers. The behavior of the system can be described by the following:

$$\tilde{x}_{(l+1)} - \tilde{x}_l = \kappa\left[w - \tilde{x}_{(l+1-d)}\overline{p}(\tilde{x}_{(l+1-d)})\right] \tag{26}$$

$$\tilde{y}_{(l+1)} - \tilde{y}_l = \kappa\left[w - \frac{1}{\tau}M_l\right] \tag{27}$$

$$M_l \sim \text{Poisson}(\tau\tilde{y}_{(l+1-d)}\overline{p}(\tilde{x}_{(l+1-d)})) \tag{28}$$

The above equations can be used to completely characterize $\tilde{x}_l = \mathbb{E}[\tilde{y}_l]$ and all the moments of $\tilde{y}_l$ for all $l$ and also in the steady state.

### 3.2.2 TCP-like controller

The equivalent marking function, $\overline{p}(x)$, as a function of the the average flow rate of the controlled flows takes a similar form. Again, as $N \to \infty$, if $\tilde{x}_l^{(N)} \to \tilde{x}_l$, $\tilde{y}_{i,l}^{(N)} \to \tilde{y}_l$ and $\tilde{q}_l^{(N)} \to \tilde{q}_l$, the following set of equations completely describe the system:

$$\tilde{y}_{(l+1)} - \tilde{y}_l = \kappa\left[w - \frac{1}{\tau}\tilde{y}_{(l+1-d)}M_{(l-\frac{d}{2})}\right] \tag{29}$$

$$\tilde{x}_{(l+1)} - \tilde{x}_l = \kappa\left[w - \mathbb{E}[\tilde{y}_l\tilde{y}_{(l+1-d)}]\overline{p}(\tilde{x}_{(l+1-d)})\right] \tag{30}$$

## 4. EQUILIBRIUM AND STABILITY OF MEAN RATE WITH PROPORTIONALLY FAIR CONTROLLER

In this section we study the mean limiting behavior $\tilde{x}_l$ in the two parameter regimes.

In the regime $\gamma^{(N)} = \gamma/N$, recall that the difference equation describing the evolution of the mean rate is as follows:

$$\begin{aligned}
\tilde{x}_{(l+1)} - \tilde{x}_l &= \kappa\left[w - \tilde{x}_{(l+1-d)}p_q(\tilde{q}_{(l+1-\frac{d}{2})}, \tilde{x}_{(l+1-d)})\right] \\
\tilde{q}_{(l+1)} &= (\tilde{q}_l + \tau(\tilde{x}_{(l-\frac{d}{2})} + a - \tilde{c}))^+ \ .
\end{aligned}$$

It immediately follows that the equilibrium point $(x^*, q^*)$ is given by

$$x^* = \tilde{c} - a = \theta c - a, \quad w = x^* p(q^*)$$

since $p_q(\tilde{q}_l, \tilde{c} - a) = p(\tilde{q}_l)$ (in the case of REM $p(\tilde{q}_l) = 1 - \exp(-\gamma\tilde{q}_l)$). However, the equilibrium can be reached only if the delay-difference equation is stable. This depends on the choice of $\gamma$. For a given value of $\gamma$, one can proceed with the difference equation and verify the stability of a linearized model numerically. However, to have a better insight on how we should choose $\gamma$ to ensure stability, we derive conditions for a continuous time deterministic model to be stable. If the step size of the discrete-time model is small, then we can expect the conditions derived from this model to be reasonably accurate. To this end, consider the system given by

$$\frac{dx(t)}{dt} = k[w - x(t-r)p_q(\tilde{q}(t-r), x(t-r))]$$
$$\frac{d\tilde{q}(t)}{dt} = (x(t) + a - \theta c)I_{(\tilde{q}(t)>0)} + \ (x(t) + a - \theta c)^+ I_{(\tilde{q}(t)=0)}$$

In the above, $x(t)$ is a continuous time version of $\tilde{x}_l$ and $k$ can be viewed as $k = \kappa/\tau$. We now find conditions for the local stability of the system. Linearizing the above system and denoting by $u(t)$ the shifted rate $(u(t) - x^*)$, and by $v(t)$ the shifted normalized virtual queue length $(\tilde{q}(t) - q^*)$ we get,

$$\frac{du(t)}{dt} = -kp_1(q^*, x^*)u(t-r) - \ kx^* p_2(q^*)v(t-r) \quad (31)$$
$$\frac{dv(t)}{dt} = u(t) \ . \quad (32)$$

In the above,

$$p_1(q^*, x^*) = 1 - \exp(-\gamma q)(1 - \tfrac{\gamma x^*}{2\tilde{c}})$$

and $p_2(q^*) = \gamma \exp(-\gamma q^*)$. We now provide condition for the system to be stable. The outline of the derivation using the Nyquist criterion is given in the Appendix. A simplified stability criterion can be obtained, if we also suppose that $x^* r \geq p_1(q^*, x^*)/(\gamma \exp(-\gamma q^*))$. This means that the equilibrium window size is not too small. Then, the system given by (31)-(32) is stable if the parameters satisfy

$$(krp_1(q^*, x^*))^2 + \frac{(kx^* p_2(q^*)r^2)^2}{4\pi^2} \leq 4\pi^2 \ ,$$

and the system is unstable if

$$(krp_1(q^*))^2 + \frac{4(kx^* p_2(q^*)r^2)^2}{81\pi^2} \geq \frac{81\pi^2}{4} \ .$$

Note that, in so far as $p_1(q^*, x^*)$ is very small, the stability condition is roughly equivalent to

$$kx^* r^2 \gamma < 4\pi^2 \ . \quad (33)$$

Thus, stability requires that $k$ and $\gamma$ both to be inversely scaled with $r$.

We now consider the parameter regime $\gamma^{(N)} = \gamma$ where the equivalent marking function is a rate based one. We have the following difference equation for $\tilde{x}_l$:

$$\tilde{x}_{(l+1)} - \tilde{x}_l = \kappa \left[ w - \tilde{x}_{(l+1-\frac{d}{2})} \overline{p}(\tilde{x}_{(l+1-\frac{d}{2})}) \right] \ . \quad (34)$$

It immediately follows that the equilibrium point $x^*$ is given by the solution of

$$w = x^* \overline{p}(x^*) \ .$$

One can obtain the stability condition of the above system numerically. However, a simplified condition for the local-stability can be obtained from the of the corresponding continuous time version of $\tilde{x}_l$ which evolves as follows:

$$\frac{dx(t)}{dt} = k[w - x(t-r)\overline{p}(x(t-r))] \ .$$

A sufficient condition for the local stability [18] of the above system is given by

$$kr(\overline{p}(x^*) + x^* \overline{p}'(x^*)) \leq \frac{\pi}{2} \ .$$

Further, it can be shown that, for $x^* < \tilde{c} - a$, $x^* \overline{p}'(x^*)/\overline{p}(x^*) \leq K_1$ for a suitable $K_1$ which depends on $\gamma$. Thus, a sufficient condition for stability is $kr\overline{p}(x^*) \leq \pi/(2(1 + K_1)$.

# 5. SIMPLIFICATIONS WITH TCP-LIKE CONTROLLER FOR CALCULATING EQUILIBRIUM

In the previous section we have given models to characterize the the behavior of the average and the individual flow rate. In the case of proportionally fair controller, the equilibrium point is easy to compute from the limiting behavior of the mean flow rate. However, in the case of TCP-like controller the equilibrium point of the mean rate is not easy to compute. Note that, in the regime $\gamma^{(N)} = \gamma/N$ the mean rate can be described as follows:

$$\tilde{x}_{(l+1)} - \tilde{x}_l = \kappa \left[ w - \mathbb{E}[\tilde{y}_l \tilde{y}_{(l+1-d)}] p_q(\tilde{q}_{(l+1-\frac{d}{2})}, \tilde{x}_{(l+1-d)}) \right] \ .$$

Clearly, calculating the equilibrium is quite complicated and can only be done along with system of stochastic equations describing the update of $\tilde{y}_l$. Thus we make an additional assumption which leads to simple difference equations describing the behavior of the mean rate. We assume $\tilde{y}_l$ and $\tilde{y}_{(l+1-d)}$ to be uncorrelated so that we can replace the term $\mathbb{E}[\tilde{y}_l \tilde{y}_{(l+1-d)}]$ by $\tilde{x}_l \tilde{x}_{(l+1-d)}$ in (23) and (30). The resulting deterministic equations (which we describe shortly) are also known as deterministic fluid models. Such models have been widely used to study the equilibrium and stability properties of congestion controlled sources and have been shown to predict the mean rate of the system very accurately [24, 17, 18, 31, 26].

In the parameter regime $\gamma^{(N)} = \gamma/N$, we thus have the following difference equation describing the mean behavior $x_l$ if we substitute $\tilde{x}_l \tilde{x}_{(l+1-d)}$ for $\mathbb{E}[\tilde{y}_l \tilde{y}_{(l+1-d)}]$ in (23):

$$\tilde{x}_{(l+1)} - \tilde{x}_l = \kappa \left[ w - \tilde{x}_l \tilde{x}_{(l+1-d)} p_q(\tilde{q}_{(l+1-\frac{d}{2})}, \tilde{x}_{(l+1-d)}) \right]$$
$$\tilde{q}_{(l+1)} = (\tilde{q}_l + \tau(\tilde{x}_{(l-\frac{d}{2})} + a - \tilde{c}))^+ \ .$$

It immediately follows that the equilibrium point $(x^*, q^*)$ is given by

$$x^* = \tilde{c} - a = \theta c - a, \quad w = {x^*}^2 p(q^*) \ .$$

We now consider the parameter regime $\gamma^{(N)} = \gamma$. Again, if we substitute $\tilde{x}_l \tilde{x}_{(l+1-d)}$ for $\mathbb{E}[\tilde{y}_l \tilde{y}_{(l+1-d)}]$ in (30) we get the following:

$$\tilde{x}_{(l+1)} - \tilde{x}_l = \kappa \left[ w - \tilde{x}_l \tilde{x}_{(l+1-d)} \overline{p}(\tilde{x}_{(l+1-d)}) \right] \ .$$

It immediately follows that the equilibrium point $x^*$ is given by the solution of

$$w = {x^*}^2 \overline{p}(x^*) \ .$$

One can also obtain stability conditions with the equations describing the approximate behavior of the average rate.

# 6. VARIANCE OF THE ARRIVAL PROCESS IN THE LINK TIME-SCALE

In this section we find the variance of the arrival process into the link. The variance of the arrival process into the link depends on the time interval over which we study the arrival process. In the following, we obtain the variance of the arrival process over the time required to serve $b$ packets. Thus we are interested in the variance over a time period $b/(Nc)$.

Let $T_b$ be the time required to serve $b$ packets. With $N$ flows in the system $T_b = b/Nc$. For the purposes of this section, we use the term big-slot to denote an interval of length $T_b$. Let $X_b$ denote the following random variable.

$$X_b \quad := \quad \text{Number of packet arrivals from the controlled flows in a big-slot}$$

It can be easily shown that [9], under Assumption 1, It thus follows that the variance of $X_b$, denoted by $\sigma^2_{X_b}$, the variance of $X_b$ is given by

$$\sigma^2_{X_b} \quad = \quad \frac{x^* b}{c} + \frac{\mathrm{Var}(z) b^2}{Nc^2} \ . \tag{35}$$

In the above $z = (\sum_{i=1}^N x_i/N \ - \ x^*)\sqrt{N}$. Under a stationarity assumption one can compute the varaice of $z$ [11]. However, note that the variance of the rate process at the source time-scale plays a smaller role at the link-time scale as the number of flows in the system increases. Thus, the variance of the arrival process into the link over a time interval $T_b$ is given by

$$\sigma^2_b = \sigma^2_{X_b} \ + \ N\sigma^2_{uc}\left(\tfrac{b}{Nc}\right) \ ,$$

where $\sigma^2_{uc}(b/(Nc))$ denotes the variance of the number of arrivals due to any uncontrolled flow in a time interval $T_b$.

The variance due to the uncontrolled flows can be greatly simplified if the arrival process due to each uncontrolled flow can be modeled as a stationary point-process. Let $A_b^{uc(N)}$ be the random variable denoting the total arrivals from the uncontrolled flows in time $T_b = b/(Nc)$. If $e_i(t)$ is the stationary point process denoting the number of arrivals from the $i^{th}$ uncontrolled flow in a time interval of length $t$, then we have

$$A_b^{uc(N)} = \sum_{i=1}^N e_i\left(\tfrac{b}{Nc}\right) \ .$$

Using a result in [8], we have that the process $A_b^{uc(N)}$ converges weakly to Poisson$(ab/c)$, where $a$ is the mean rate of any uncontrolled flow. Thus, for large $N$, the variance of the arrival process into the link over a time interval $T_b$ is given by

$$\sigma^2_b = \frac{(x^* + a)b}{c} \ . \tag{36}$$

Now, suppose we want to find the buffer size needed to guarantee a certain buffer-overflow probability $\delta$. If the buffer size is $b$, one way to guarantee a very small buffer overflow probability is to ensure that the number of arrivals into the link over a period of time $T_b$ is less than $b$ with a very high probability. Note that the round trip delay as expressed in big-slots is $m = rNc/b$ big-slots. Thus, we are trying to make the queuing delay equal to $1/m$ fraction of the the propagation delay and this fraction becomes very small in a large system. A simple rule to choose the buffer size can be $\Pr[X_b + A_b^{uc(N)} \leq b] \leq \delta$. If we further use a Gaussian approximation for the arrival process (the unconditioned arrival process into the link), we have the following simple rule to guarantee a buffer over-flow probability of less than, say $\delta$, with a buffer size of $b$ packets:

$$Q\left(\frac{b - b\frac{x^*+a}{c}}{\sqrt{\sigma_X^2 + N\sigma_u^2(b/Nc)}}\right) < \delta \ .$$

For $\delta = 0.01$, we simply need

$$\frac{b - b\frac{x^*+a}{c}}{\sqrt{\sigma_X^2 + N\sigma_{uc}^2\left(\frac{b}{Nc}\right)}} > 2.33 \ .$$

Again, if the controlled flows are modeled as a stationary point process, the above condition is asymptotically equivalent to

$$b \geq \frac{5.43\eta}{(1-\eta)^2} \ ,$$

where $\eta$ the target utilization of the link given by $\eta = (x^* + a)/c$. The value of $x^*$ can be obtained using the suitable model for marking function depending upon the AQM parameter.

# 7. RESULTS AND DISCUSSION

In this section we show packet-based simulation results to validate some of our observations and results. The purpose of the simulations is the following. First, we demonstrate that the utilization as predicted by the suitable models for marking function with the different scaling of parameter is close to that observed from the simulations. Second, we show the accuracy of the variance of the arrival process as predicted from our models. Third, we want to verify that negligible queue delay can indeed be obtained using a virtual-queue mechanism.

We simulate a single bottleneck link accessed by multiple TCP sources (each with round-trip delay 100 ms), all of which are in the congestion avoidance phase. Apart from the TCP sources we also consider unresponsive flows. We use an ON-OFF model for the uncontrolled flows [16]. The uncontrolled flows toggle between ON and OFF state which are exponentially distributed with mean 0.2 s. In the ON state, an uncontrolled flow sends data at a rate $\rho$ packets/s. In all our simulations with various AQM schemes, we change $N$, the number of TCP sources, which is also the number of uncontrolled flows in the system. The link capacity in all our simulations is $Nc$, where $c = 80$ packets/s. The flow rate $\rho$ of the uncontrolled flows in the ON state is adjusted so that uncontrolled flows deliver a load of 25% into the link. Every simulation result is averaged over 10 runs.

## 7.1 Virtual queue-based marking with REM

We compare average utilization, variance and mean queue length with two parameter scalings of REM: $\gamma^{(N)} = \gamma/N$ and $\gamma^{(N)} = \gamma$. We are interested in comparing the variance of the total arrivals over time to serve $b$ packets in the queue. In our simulation we choose the value of $b$ as $b = 32$. The capacity of the virtual queue is chosen as $0.85Nc$ so that $\theta = 0.85$.

In Figure 2, we compare the average utilization of the link, the variance of the arrival process into the link over a time to serve 32 packets and the mean queue-lengths. The predicted results (from the theory) and the simulations results are shown. In all the cases we have chosen $\gamma = 0.0075$, so that the plots on the left-hand panel in Figure 2 are with marking function $f(\hat{q}) = 1 - \exp(-0.0075\hat{q}/N)$, and

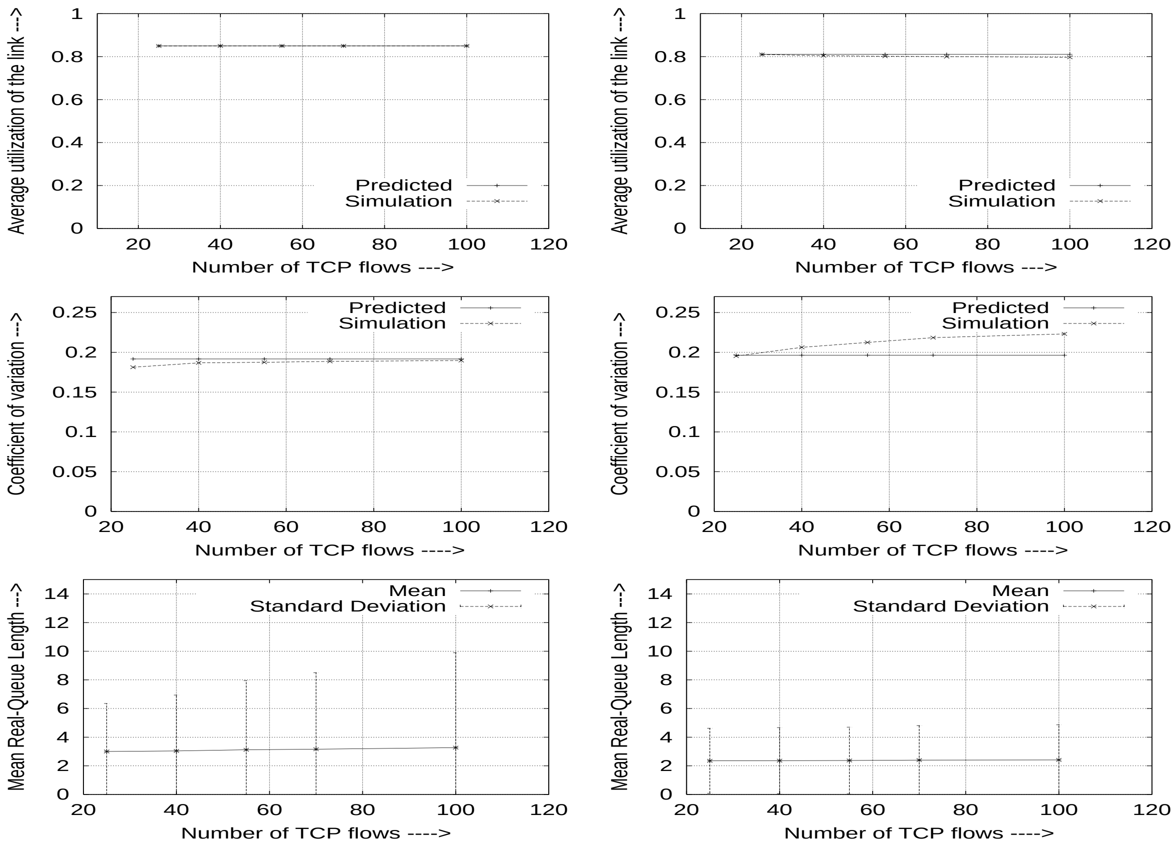


**Figure 2: Comparison of average utilization, coefficient of variation and mean queue length with virtual queue based REM. On the left-hand panel we show plots when $\gamma^{(N)}$ is scaled as $\gamma^{(N)} = \gamma/N$, and the right-hand panel shows plots with $\gamma^{(N)} = \gamma$. In each case $\gamma = 0.0075$.**

the plots on the right hand panel are with marking function $f(\hat{q}) = 1 - \exp(-0.0075\hat{q})$. The value of $\gamma$ satisfies the local stability condition of the corresponding deterministic continuous time model in all the cases. The jointly rate-and-queue-based model or the equivalent rate based model of the marking function as proposed can be used to calculate the equilibrium point of the corresponding deterministic model. Note that the utilization as predicted from the equilibrium point is quite close to the average utilization observed in the simulations. Thus, the rate-based marking model can be used to accurately predict the target utilization of the link. Note that the numerical results of coefficient of variation (from the expression in (36)) is close to that observed from our simulations. In most cases, the standard deviation predicted using the theory is within 10-15% of the value observed from the simulations. We have also shown the mean queue length at the real-queue. In all the cases, the real queue length is very small. Clearly, a virtual queue based marking can lead a low-loss, low-delay operation in two different regimes of parameter.

The robustness of virtual queue based REM with respect to the scaling of $\gamma^{(N)}$ suggests that it may not be required to know the exact number of flows through the link. However, the utilization in the parameter regime $\gamma^{(N)} = \gamma/N$ is easy to set and is equal to $\theta$ ($\theta C$ is the capacity of the virtual queue).

## 7.2 Real-queue based marking with REM

The performance of the system is much more sensitive to $\gamma^{(N)}$ if real-queue-based marking is used instead of virtual-queue-based marking. With real-queue-based marking, if $\gamma^{(N)}$ is scaled as $\gamma^{(N)} = \gamma/N$, the real queue length becomes $O(N)$ which can lead to huge losses and undesirable performance deterioration. However, in a real-queue-based marking, if $\gamma^{(N)} = \gamma$, we can still achieve a low-loss, low-delay operation. Recall that the appropriate models for marking function in these regimes are rate-based models. In Figure 3, we show the different quantities when $\gamma^{(N)} = 0.05$. Note that, the average utilization as predicted by the rate-based model is quite accurate. Thus, with a real-queue based marking the proposed equivalent rate based model can be used to design the parameters for a desired utilization of the link. In the case of $\gamma^{(N)} = \gamma$, as expected, the mean queue length does not grow with $N$.

# 8. CONCLUDING REMARKS

In this paper, we have provided appropriate models for virtual-queue based marking, based on the scaling of parameters in a virtual-queue based AQM scheme. Using virtual queue based REM as an illustration, we have shown that the right kind of marking model can be rate based or virtual-queue length based model depending on the scaling of a certain parameter with the number of flows. Further,

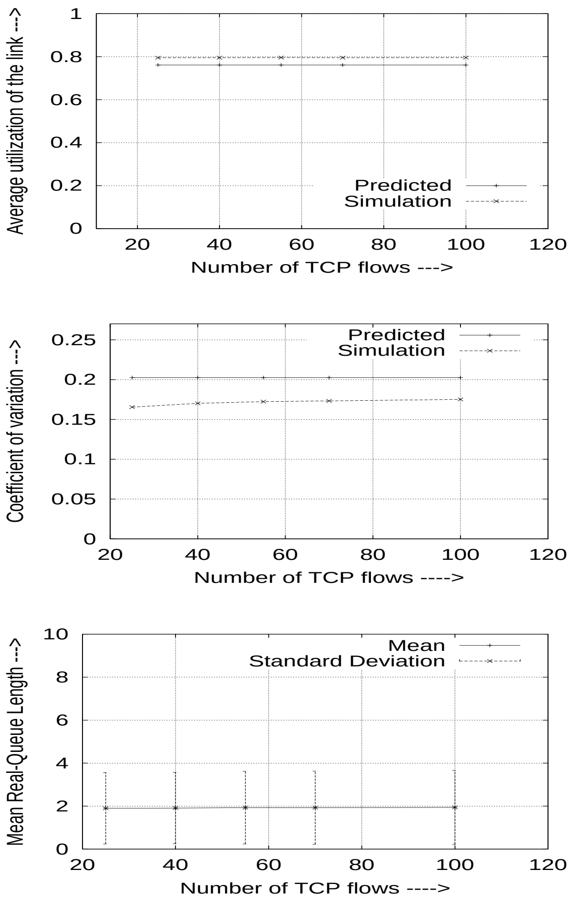


**Figure 3: Comparison of average utilization, coefficient of variation, and mean queue length with real queue-based REM with $\gamma^{(N)} = \gamma = 0.05$.**

depending on the model, the stability conditions can be different. One of the our observations is that: *If a virtual-queue based marking is used, the steady state properties like mean and variance of the arrival process at the link is very robust to the choice of parameter.* The rate-based models quite accurately predict the utilization of the link and thus can be used to design the link the capacity or the parameters of the AQM scheme. One more advantage of virtual-queue based marking is that, if there are $N$ flows accessing a link of capacity $Nc$, a buffer size of $O(1)$ can provide a low-loss, low-delay operation for a given link utilization across different parameter regimes.

## Acknowledgment

Research supported in part by DARPA grant F30602-00-2-0542 and AFOSR grant URI F49620-01-1-0365.

# 9. REFERENCES


[1] T. Alpcan and T. Başar. A utility-based congestion control scheme for internet-style networks with delay. In *Proceedings of IEEE Infocom*, pages 2039–2048, San Francisco, California, March-April 2003.

[2] E. Altman, K. Avrachenkov, C. Barakat, and R. Nunez-Queija. State-dependent $M/G/1$ type queueing analysis for congestion control in data networks. *Computer Networks*, 39(6):789–808, August 2002.

[3] S. Athuraliya, D. E. Lapsley, and S. H. Low. Random early marking for Internet congestion control. In *Proceedings of IEEE GLOBECOM*, pages 1747–1752, Rio de Janeiro, Brazil, 1999.

[4] S. Athuraliya and S. H. Low. Optimization flow control, I: Basic algorithm and convergence. *IEEE/ACM Transactions on Networking*, pages 861–875, December 1999.

[5] S. Athuraliya and S. H. Low. Optimization flow control, II: Implementation. 2000. http://netlab.caltech.edu/.

[6] F. Baccelli and D. Hong. Flow level simulation of large IP networks. In *Proceedings of INFOCOM*, pages 1911–1921, April 2003.

[7] H. Chen and D. D. Yao. *Fundamentals of Queueing Networks*. Springer-Verlag, New York, 2000.

[8] D. J. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes*. Springer-Verlag, New York, 1988.

[9] S. Deb. Deterministic fluid models for internet congestion control. PhD thesis, University of Illinois at Urbana-Champaign, September 2003.

[10] S. Deb, S. Shakkottai, and R. Srikant. Stability and convergence of TCP flows in a many-flows regime. In *Proceedings of INFOCOM*, pages 884–894, San Francisco, CA, April.

[11] S. Deb, and R. Srikant. Stochastic models of proportionally fair congestion controllers. In *Proceedings of IEEE Conference on Decision and Control*, Dec 2003.

[12] S. Floyd. TCP and explicit congestion notification. *ACM Computer Communication Review*, 24:10–23, October 1994.

[13] S. Floyd and V. Jacobson. Random early detection gateways for congestion avoidance. *IEEE/ACM Transactions on Networking*, 1(4):397–413, August 1993.

[14] R. J. Gibbens and F. P. Kelly. Resource pricing and the evolution of congestion control. *Automatica*, 35:1969–1985, 1999.

[15] C. V. Hollot and Y. Chait. Nonlinear stability analysis for a class of TCP/AQM schemes. In *Proceedings of the IEEE Conference on Decision and Control*, pages 2309–2314, Orlando, FL, December 2001.

[16] C. V. Hollot, Y. Liu, V. Mishra, and D. Towsley. Unresponsive flows and AQM performance. In *Proceedings of INFOCOM*, volume 1, pages 85–95, San Francisco, CA, April 2003.

[17] C. V. Hollot, V. Misra, D. Towsley, and W. Gong. On designing improved controllers for AQM routers supporting TCP flows. In *Proceedings of INFOCOM*, pages 1726–1734, Anchorage, Alaska, April 2001.

[18] F. P. Kelly. Models for a self-managed Internet. *Philosophical Transactions of the Royal Society*, A358:2335–2348, 2000.

[19] F. P. Kelly. Fairness and stability of end-to-end congestion control. *European Journal of Control*, 9:149–165, 2003.

[20] F. P. Kelly, A. Maulloo, and D. Tan. Rate control in communication networks: shadow prices, proportional fairness and stability. *Journal of the Operational Research Society*, 49:237–252, 1998.

[21] L. Kleinrock. *Queueing Systems, Volume 1: Theory.* Wiley–Interscience, New York, 1975.

[22] S. Kunniyur and R. Srikant. Analysis and design of an adaptive virtual queue algorithm for active queue management. In *Proceedings of ACM Sigcomm*, pages 123–134, 2001.

[23] S. Kunniyur and R. Srikant. A time-scale decomposition approach to adaptive ECN marking. In *Proceedings of IEEE INFOCOM*, pages 1330–1339, Anchorage, Alaska, April 2001.

[24] S. Kunniyur and R. Srikant. End-to-end congestion control: utility functions, random losses and ECN marks. *IEEE/ACM Transactions on Networking*, to be published. An earlier version appeared in *Proceedings of IEEE INFOCOM, 2000*.

[25] S. Liu, T. Başar, and R. Srikant. Controlling the Internet: A survey and some new results. In *Proceedings of IEEE Conference on Decision and Control*, Maui, Hawaii, December 2003.

[26] S. H. Low, F. Paganini, J. Wang, S. Adlakha, and J. C. Doyle. Dynamics of TCP/RED and a scalable control. In *Proceedings of INFOCOM*, pages 239–248, New York, June 2002.

[27] V. Misra, W. Gong, and D. Towsley. Stochastic differential equation modeling and analysis of TCP windowsize behavior. In *Proceedings of Performance'99*, Istanbul, Turkey, October 1999.

[28] T. J. Ott. ECN protocols and TCP paradigm. *http://web.njit.edu/∼ott/Papers*, 1999.

[29] S. Shakkottai and R. Srikant. Mean FDE models for Internet congestion control. 2002,. *http://www.comm.csl.uiuc.edu/∼srikant.* A shorter version appeared in the *Proceedings of IEEE INFOCOM*, 2002.

[30] P. Tinnakornsrisuphap and A. Makowski. Limit behavior of ECN/RED gateways under a large number of TCP flows. In *Proceedings of INFOCOM*, pages 873–883, San Francisco, CA, April 2003.

[31] G. Vinnicombe. On the stability of networks operating TCP-like congestion control. In *Proceedings of the IFAC World Congress*, Barcelona, Spain, 2002. Also, University of Cambridge Technical Report CUED/F-INFENG/TR.398, *http://www.eng.cam.ac.uk/˜gv.*

[32] J. T. Wen and M. Arcak. A unifying passivity framework for network flow control. In *Proceedings of IEEE Infocom*, pages 1156–1166, San Francisco, California, April 2003.

[33] Y. Yi, S. Deb, and S. Shakkottai. Short queue regime and rate based marking. In *Proceedings of Conference on Information Science and Systems*, Princeton, NJ, March 2004.


# APPENDIX

**Stability Condition of Virtual Queue-Based REM with $\gamma^{(N)} = \gamma/N$:** We now outline the derivation of the stability condition for the system given by (31)-(32). Note that the system under consideration can be expressed as

$$\frac{du}{dt} = -au(t-r) - bv(t-r) \ , \quad \frac{dv}{dt} = u(t) \ .$$

The above system is stable if the roots of

$$\begin{aligned} G(s) &= \det \begin{bmatrix} s + a\exp(-sr) & b\exp(-sr) \\ -1 & s \end{bmatrix} \\ &= s^2 \left(1 + \frac{a\exp(-sr)}{s} + \frac{b\exp(-sr)}{s^2}\right) \end{aligned}$$

lie in the left half of the complex plane. From the Nyquist stability criterion, it is sufficient to show that the Nyquist plot of $L(j\omega)$, where

$$L(s) = \frac{a\exp(-sr)}{s} + \frac{b\exp(-sr)}{s^2}$$

does not enclose the point $(-1+j0)$. Expanding $L(j\omega)$, it can be verified that a sufficient condition for this is

$$h(\omega) = \frac{a\sin(\omega r)}{\omega} + \frac{b\cos(\omega r)}{\omega^2} \leq 1$$

whenever $\tan(\omega r) = \frac{a\omega}{b}, \ \ \omega \neq 0$ . The above is a sufficient condition for stability of the linear system under consideration.

The condition can be simplified if we further assume $x^* r \geq p_1(q^*, x^*)/(\gamma \exp(-\gamma q^*))$, or equivalently $a/(br) \leq 1$.

Now suppose $a/(br) < 1$. It is easy to see $\omega$ satisfying (37) is such that $\omega r \in [\pi, 3\pi/2] \cup [2\pi, 5\pi/2] \cup [3\pi, 7\pi/2] \ldots$. First note that if $\omega r \in [\pi, 3\pi/2] \cup [3\pi, 7\pi/2] \cup [5\pi, 11\pi/2] \ldots$, the values of $\cos(wr)$ and $\sin(\omega r)$ are negative, and so the condition $h(\omega) < 1$ for $\omega$ satisfying (37) is trivially satisfied. It is thus enough to consider $\omega r$ satisfying (37) in the range $\{[2\pi, 5\pi/2] \cup [4\pi, 9\pi/2] \ldots\}$ when the values of $\cos(wr)$ and $\sin(\omega r)$ are positive. Using routine trigonometric manipulations, it can be shown that, under (37), $h(\omega)$ satisfies

$$h(\omega) = \frac{\sqrt{a^2\omega^2 + b^2}}{\omega^2} \ .$$

Since $h(\omega)$ is decreasing in $\omega$, by considering the solution of (37) such that $\omega r \in [2\pi, 5\pi/2]$, it follows that a sufficient condition for local stability is

$$a^2 r^2 + \frac{b^2 r^4}{4\pi^2} \leq 4\pi^2 \ .$$

To find a sufficient condition for the instability of the linearized system when $a/(br) \leq 1$, note that, for $\omega r$ satisfying (37) in the range $[2\pi, 5\pi/2] \cup [4\pi, 9\pi/2]$,

$$h(\omega) \geq \sqrt{\frac{4a^2r^2}{81\pi^2} + \frac{16b^2r^4}{(81\pi^2)^2}} \ . \tag{37}$$

Thus, $h(\omega) > 1$ if

$$a^2 r^2 + \frac{4b^2 r^4}{81\pi^2} \geq \frac{81\pi^2}{4} \ . \tag{38}$$

Further, we also have $\omega r$ satisfying (37) in the range $[\pi, 3\pi/2] \cup [3\pi, 7\pi/2] \cup \ldots$ trivially satisfies $h(\omega) < 1$. Suppose, $\omega_2 r \in [2\pi, 5\pi/2]$, $\omega_3 r \in [3\pi, 7\pi/2]$, $\omega_4 r \in [4\pi, 9\pi/2]$, and $\omega_5 r \in [5\pi, 11\pi/2]$ satisfy (37). By our preceding argument, $h(\omega_2) > 1$, $h(\omega_3) < 1$, $h(\omega_4) > 1$ and $h(\omega_5) < 1$. It is not hard to see that the Nyquist plot of $L(j\omega)$ encircles $(-1+j0)$, and hence, (39) provides a sufficient condition for instability of the linearized system.