Linear Stability Analysis of FAST TCP Using A New Accurate Link Model

Ao Tang, Krister Jacobsson, Lachlan L. H. Andrew, and Steven H. Low

***Abstract*— A link model which captures the queue dynamics when congestion windows of TCP sources change was recently proposed. By considering both the self-clocking and the link integrator effects, the model is a generalization of the known extreme cases and is more accurate in general. In this paper, we apply this model to the stability analysis of FAST TCP. It is shown that FAST TCP flows over a single link are always linearly stable regardless of delay distribution. This result resolves the notable discrepancy between empirical observations and previous theoretical predictions. The analysis highlights the critical role of self-clocking in TCP stability and the scalability of FAST TCP with respect to delay. The proof technique is new and much less conservative than existing ones.**

I. Introduction

In the current active research field of network congestion control [17], [22], one line of work of fundamental interest is the dynamics of congestion control protocols. Stability is crucial to ensure that the system operating point is indeed the desired equilibrium, which carries various ideal target properties (e.g., efficiency, fairness). Local as well as global behaviors of congestion control protocols have attracted much attention (see e.g., [5], [10], [12], [13], [14], [15], [18], [19], [20], [23] for local stability and [1], [8], [9], [26], [28] for global stability).

Until recently, stability analyses have been based predominantly on rate-based (rather than window based) source models, and, almost without exception, on integrator queue models. In these models, sources control their data rates explicitly and the rate of change of the queueing delay is proportional to the difference between the aggregate incoming traffic and link capacity. Typically, these results show that the system is stable when round trip delays do not exceed some upper bound. This is in line with our intuition that increased feedback delay may have a destabilizing effect on the closed loop system.

There are however features of TCP which play important roles on stability but are not captured in the rate-based/integrator model. Current TCPs are window based algorithms, meaning that each sender controls its window size—an upper bound on the number of packets that have been sent out but not acknowledged. The actual rate of transmission of new packets is controlled or "clocked" by the stream of received acknowledgments (ACKs) by the transmission control that transmits a new packet for each received ACK, thereby keeping the number of outstanding packets constant and equal to the window size. Therefore, sources control the amount of data they inject into the network rather than the rate of doing so. Intuitively, this "*volume control*" is much safer in terms of stability than "*rate control*". Essentially, a new packet is sent into the network only when an acknowledgment is received *and* when the current congestion window allows it. This so called "self-clocking" has the consequence that the queue size traces the changes in the window much faster than the integrator queue model predicts, and also that the sending rate cannot exceed the capacity over long time frames. Both are favorable to stability.

For the case where flows' round trip delays coincide and no non-window-based cross traffic is present, the self-clocking effect is dominant: the total ACK rate cannot exceed the link capacity at any time and hence a window change translates into a proportional change in the queue. Indeed, as an example, homogeneous FAST TCP [27] flows are reported empirically to be stable with any delay, in stark contrast to the prediction of the integrator link model which asserts instability when delay is large enough. This observation motivated the proposal of a static link model to capture queue dynamics under TCP window control [24], [25]. Using the static link model, it has been theoretically shown that FAST TCP flows are always stable for the homogeneous case [6], [24], [25]. As the static model fails to hold in heterogeneous cases, we certainly need a better model to study the general case.

It turns out that the integrator model and the static link model each only emphasizes one side of the story; the former tends lag the true dynamics resulting in conservative stability results, while the latter leads the true dynamics yielding optimistic predictions. After introducing basic notation and preliminary knowledge, we will review in Section II a recent result [11] asserting that a natural combination of these two models leads to a more accurate one. As an illustration of its impact, we then specialize to the problem of the stability of FAST TCP in Section III where using the joint link model, we prove that FAST running over a single link is stable for any heterogeneous delays, and hence resolve the discrepancy between previous experimental results and existing theoretical predictions. Closed loop experiments are also reported where accurate predictions on stability region

A. Tang is with Division of Engineering and Applied Science at Caltech aotang@caltech.edu

K. Jacobsson is with Automatic Control School of Electrical Engineering at the Royal Institute of Technology (KTH) krister.jacobsson@s3.kth.se

L. Andrew is with Computer Science Department at Caltech lachlan@caltech.edu

S. Low is with Electrical Engineering and Computer Science Departments at Caltech slow@caltech.edu

is obtained and verified with packet level simulations[1].

II. Model and Notation

To capture the self-clocking effect in window based congestion control, we avoid working directly with the sources' sending rates as states of the protocol. Instead, we use the sources' window sizes and the bottleneck queue size to represent the state of the closed loop system.

A. Preliminaries

We consider N window-based FAST TCP sources sending over a bottleneck link with capacity c. Let $w_n(t)$ denote the congestion window of source n at time t, $n \in \{1, \ldots, N\}$. Let a packet that is sent by source n at time t appear at the bottleneck queue at time $t + \tau_n^f$. This *forward delay* τ_n^f models the amount of time it takes to travel *from* source n *to* the link, and it accounts for forward latency but not queuing delays. The *backward delay* $\tau_n^b(t)$ is defined in the same manner: it is the time from when a packet arrives at the link to when the corresponding acknowledgment is received at source n. Note that the backward delay includes the queuing delay at the bottleneck queue. The *round trip delay* $\tau_n(t)$ seen by source n is the elapsed time between when a packet is sent and when the corresponding acknowledgment is received; naturally $\tau_n(t) = \tau_n^f + \tau_n^b(t)$. The *latency* of source n is denoted d_n and is defined as the minimum achievable round trip delay, i.e. the round-trip delay when the bottleneck queue is empty.

The queueing delay of the bottleneck link is denoted by $p(t)$, and $c > 0$ is the capacity of the link. The queuing delay observed by the nth source at time t is $q_n(t)$; it relates to the queue delay by $q(t) = p(\tilde{t})$, where $\tilde{t}$ solves $t = \tilde{t} + \tau_n^b(\tilde{t})$.

The bottleneck link may also carry non-window-based traffic such as User Datagram Protocol (UDP) traffic. Let $x_c(t) \in [0, c]$ be the rate (averaged over a suitable time interval) at which non-window-based cross-traffic is sent over the link. This implies that the *available bandwidth*, from the perspective of the window based sources over the link, is $x_a(t) := c - x_c(t)$.

Whenever a time argument of a variable is omitted it represents its equilibrium value, e.g., $p(t)$ corresponds to the queueing delay variable at time t while p is its equilibrium value. When working in discrete time the convention $w(k) := w(t_k)$ will be used.

B. Link models

As described in Section I, previous work differs in how the dynamic map between the window sizes and the buffer size is modeled. Most existing literature on window based congestion control, see [2], [4], [10], [15], [17], [16], assumes that the sending rate is proportional to window size divided by the round-trip time and may further model the queue as a simple integrator, integrating the excess rate at the link, i.e.,
Integrator link model:

$$\dot{p}(t) = \frac{1}{c}\left(\sum_{n=1}^{N} \frac{w_n(t - \tau_n^f)}{d_n + p(t)} + x_c(t) - c\right). \tag{1}$$

The model (1) does not, however, take into account the "self-clocking" characteristic of window based schemes, where the sending rate is regulated by the rate of the received ACKs. To model this phenomenon, an alternative model is used in [24], [27] and implicitly in [21], based on the empirical observation that, due to "self-clocking", transient effects are negligible. The relation between the window size and the buffer size is then described by the algebraic relation[2]
Static link model:

$$\sum_{n=1}^{N} \frac{w_n(t - \tau_n^f)}{d_n + p(t)} = c - x_c(t). \tag{2}$$

In (2), a change in any of the sources' congestion windows w_n results in a proportional change in the queueing delay p one forward delay τ_n^f later. This is in contrast to the integrator model (1) where a window change gives a smooth queue transition. Obviously the two models are fundamentally different. We will now combine their main features and arrive at a more accurate joint model.

1) Joint link model: The joint model is justified heuristically in this section, with the aim of emphasizing the underlying intuition. See [11] for a rigorous derivation, where the same model results from a thorough and detailed analysis of the system at the packet level.

To understand the difference between (and the accuracy of) the integrator model (1) and the static model (2), the key is to examine cases where flows have heterogeneous round trip delays or where non-window-based cross traffic such as UDP exist. Consider the case of N flows are sending over the bottleneck link with constant window sizes, and consider the response to a change in window size by a system initially in equilibrium.

The long term effect of a window change is that the queue integrates the excess rate. This is well known, and captured by the integrator model (1).

The short term effect of a window change is more complex, and often wrongly ignored. Since the link is fully utilized in equilibrium, the queue's immediate response to a window change (that is, transmission control injects an extra packet or discards an ACK) is a proportional change in the queue size; this will occur one forward delay τ_n^f after the window w_n is changed. If there is no cross traffic and the sources' share the same round trip time (RTT), there is no further transient. This is due to the fact that sources' sending rates are auto-regulated by their individual ACK rates which sum up to the capacity; subsequently the queue input rate will equal its output rate (capacity). This is in

[1]To the best of our knowledge, the current status of research on congestion control protocols can provide quantitative results on equilibrium, while for dynamics, most works focus on qualitative study and have not been able to compare predictions with packet level simulations quantitatively.

[2]The original model was presented in discrete time for multiple bottlenecks, here we use its continuous time version used in e.g. [6] and we consider a single bottleneck.

line with the static queue model (2) which neglects transient behavior. However, in the case when there is stationary UDP cross traffic, *sources can affect their ACK rates over time intervals greater than one RTT* (actually it is a function of the amount of cross traffic, see [11]). This is also true when no cross traffic is present but the heterogeneity among sources' RTT is large enough. As individual flows operate on their individual RTT time scales and it takes one RTT before a queue change affects the queue input rate, from the perspective of flows with small round times, flows with larger RTTs can be considered as non-responsive cross traffic and the system is hence transient in this case as well.

Adopting the standard fluid flow approximation, that packets transmitted from a source form a continuous smooth flow, and motivated by the previous discussion we capture both the short and long term behavior in a single model,

Joint link model:

$$\dot{p}(t) = \frac{1}{c}\left(\sum_{n=1}^{N}\left(\frac{w_n(t-\tau_n^f)}{d_n+p(t)} + \dot{w}_n(t-\tau_n^f)\right) + x_c(t) - c\right), \tag{3}$$

which can be seen as a superposition of (1) and (2). The derivative term $\dot{w}(t-\tau_n^f)$ models the immediate proportional change in the queue size due to a window change. Note that it is the window size and its corresponding time derivative *only* that have delayed variable arguments, which furthermore are identical. This is rigorously motivated in the full model derivation in [11]. A similar model was also implicitly used for flow control stability analysis in [3].

Consider the Laplace domain transfer function from the window size of flow n to the queue derived from the linearized version of (3)

$$G_{pw_n}(s) = \frac{\mathcal{L}[p(t)]}{\mathcal{L}[w_n(t)]} = \frac{\frac{1}{c}\left(s+\frac{1}{d_n+p}\right)e^{-s\tau_n^f}}{s+\frac{1}{c}\sum_{i=1}^{N} w_i/(d_i+p)^2}. \tag{4}$$

For the the case with homogeneous delay $d_n = d$ and no cross traffic, applying the identity $\sum_{n=1}^{N} w_n/(d+p) = c$ shows that the transfer function zero and pole will cancel and hence that the map is a pure delay, scaled by $1/c$. This agrees with the description in the previous discussion as well as in the model (2). Moreover, the dynamics will be more distinguishable with increasing heterogeneity among the sources and cross traffic as in (1). Finally, note that (4) is open loop stable, as expected due to self-clocking.

In summary, while the integrator link model may lag the true dynamics and the static link model can lead the true dynamics a lot, the joint model (3) succeeds in modelling the two main system characteristic of "self-clocking", i.e., the short term proportional change as well as the long term integrating effect that are present in the system. For rigorous derivation of the model and open loop validation experiments, one is referred to [11].

III. Linear Stability of FAST TCP

FAST [27] is a high speed TCP variant that uses delay as its main control signal. So far, all experiments with FAST TCP have operated at a stable equilibrium regardless of what the round trip delays are. This section will use the joint link model (3) to show that FAST is indeed locally stable for a single bottleneck with default step size.

A. Window model of FAST TCP

The sending rate of FAST TCP is implicitly adjusted via the congestion window mechanism. Each sender updates its window size according to

$$w_n(k+1) = (1-\gamma_n)w_n(k) + \gamma_n \frac{d_n}{d_n+\hat{q}_n(k)} w_n(k) + \gamma_n \alpha_n \tag{5}$$

where $\alpha \in \mathbb{Z}^+$ and γ_n are protocol parameters [27]. This update is performed *once* every RTT. The queuing delay $q_n(k)$ is *estimated* by the source and the kth estimate is denoted $\hat{q}_n(k)$. The window algorithm operates in time scale of RTTs while the estimator is such that it operates on a time scale of packets; for the case of high bandwidth and latency, these time scales are fundamentally separated. This work will therefore ignore estimator dynamics and use $\hat{q}_n(k) = q_n(k)$ in (5).

To obtain a continuous time approximation of the window control, first rewrite (5) as

$$\begin{aligned}\frac{w_n(k+1)-w_n(k)}{\tau_n(k)} \quad = \quad & \frac{\gamma_n}{\tau_n(k)}\cdot\frac{d_n}{d_n+q_n(k)}w_n(k) \\ & -\frac{\gamma_n}{\tau_n(k)}w_n(k) + \frac{\gamma_n \alpha}{\tau_n(k)}. \end{aligned} \tag{6}$$

Using a first order Euler approximation of the derivative and applying the identity $\tau_n(t) = d_n + q_n(t)$ yields the continuous time window update

$$\dot{w}_n(t) = -\gamma_n \frac{q_n(t)}{(d_n+q_n(t))^2} w_n(t) + \gamma_n \frac{\alpha}{d_n+q_n(t)}. \tag{7}$$

Linearizing (7) around (w, q), and noting that in equilibrium $\alpha_n/q = w_n/(d_n+q)$, gives

$$\dot{\tilde{w}}_n(t) = -\gamma_n \frac{q}{{\tau_n}^2}\tilde{w}_n(t) - \gamma_n \frac{\alpha_n d_n}{q{\tau_n}^2}\tilde{q}_n(t) \tag{8}$$

where $\tilde{w}_n$ and $\tilde{q}_n$ represent perturbed variables. Adopting the standard convention that whenever a round-trip time, or forward and backward delay, appears in a variable argument, it is replaced by its equilibrium value [18], [25], yields

$$\tilde{q}_n(t) = \tilde{p}(t-\tau_n^b) \tag{9}$$

around equilibrium.

B. Loop Gain

Combining the Laplace transform version of the linear window dynamics (8) and the communicated price (queuing delay) (9) with the frequency domain version of the linear queue dynamics (4) and the backward propagation delay, results in the open loop transfer function

$$L(s) = \sum_{n=1}^{N} \mu_n L_n(s) \tag{10}$$

where

$$\mu_n = \frac{x_n}{c} = \frac{\alpha_n}{cq} \tag{11}$$

and

$$L_n(s) = \frac{s + \frac{1}{\tau_n}}{s + \frac{1}{\hat{\tau}_n}} \frac{d_n \gamma_n e^{-\tau_n s}}{\tau_n^2 s + \gamma_n q} \tag{12}$$

where

$$\frac{1}{\hat{\tau}} = \sum_{n=1}^{N} \mu_n \frac{1}{\tau_n}. \tag{13}$$

We can interpret $\hat{\tau}$ as a weighted harmonic mean value of τ_n. In particular, when all flows have equal α_n and hence $\mu_n = 1/N$, $\hat{\tau}$ is the harmonic mean of τ_n[3].

C. Stability analysis

For notational simplicity, we will consider the case $\gamma_n = \gamma$ in the rest of this paper. Intuitively, for given parameters, $q = 0$ is the worst case for stability as each $L_n(s)$ will have larger gain and lose more phase. We now focus on the case of $q = 0$, i.e., the open loop transfer function is

$$L(s) = \sum_i \mu_i \frac{s + \frac{1}{\tau_i}}{s + \frac{1}{\hat{\tau}}} \frac{\gamma e^{-s\tau_i}}{s\tau_i}. \tag{14}$$

Define $H(\omega)$ as the half plane under the line that passes $-1 + j0$ with slope $1/\omega\hat{\tau}$. Formally, we have

$$H(\omega) = \left\{ x \mid \arg(x+1) - \arctan\left(\frac{1}{\omega\hat{\tau}}\right) \in (-\pi, 0) \right\}. \tag{15}$$

Lemma 1. *If* $q = 0$ *and*

$$\sum_{i=1}^{N} \frac{\mu_i}{\tau_i} \left(\cos(\omega\tau_i) - \frac{\sin(\omega\tau_i)}{\omega\tau_i} \right) + \frac{1}{\hat{\tau}\gamma} > 0 \tag{16}$$

then

$$L(j\omega) \in H(\omega). \tag{17}$$

Proof. When $q = 0$, by definition $L(j\omega) \in H(\omega)$ is equivalent to

$$\arg(L(j\omega) + 1) - \arctan\left(\frac{1}{\omega\hat{\tau}}\right) \in (-\pi, 0). \tag{18}$$

Substituting (14) and noting that

$$\arg\left(j\omega + \frac{1}{\hat{\tau}}\right) + \arctan\left(\frac{1}{\omega\hat{\tau}}\right) = \frac{\pi}{2}, \tag{19}$$

condition (18) can be further rewritten as

$$\arg\left(\sum_i \mu_i \left(j\omega + \frac{1}{\tau_i} \right) \frac{\gamma e^{-j\omega\tau_i}}{j\omega\tau_i} + 1 \right) \in \left(-\frac{\pi}{2}, \frac{\pi}{2} \right) \tag{20}$$

[3]The loop gain determined by (10) to (13) can be used to study scenarios when UDP background traffic exist by setting the corresponding $\tau_n = \infty$. Therefore all techniques and results in this paper can be applied to those cases.

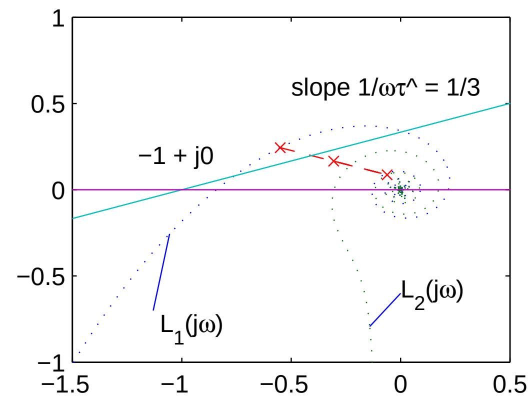


Fig. 1. An example of a line of slope $1/\omega\hat{\tau}$ which bounds $L(j\omega)$, denoted by the center cross. Note that the individual terms $L_n(j\omega)$, denoted by the individual crosses, are not all below this line.

which is equivalent to

$$\begin{aligned} &\mathrm{Re}\left(\sum_i \mu_i \left(j\omega + \frac{1}{\tau_i} \right) \frac{\gamma e^{-j\omega\tau_i}}{j\omega\tau_i} + j\omega + \frac{1}{\hat{\tau}} \right) > 0 \\ \iff &\sum_{i=1}^{N} \frac{\mu_i}{\tau_i} \left(\gamma\cos(\omega\tau_i) - \frac{\gamma\sin(\omega\tau_i)}{\omega\tau_i} \right) + \frac{1}{\hat{\tau}} > 0. \end{aligned}$$

Dividing through by $\gamma > 0$ gives the hypothesis of the lemma. □

The construction used for Lemma 1 is depicted in Figure 1, for $\tau_1 = 1$, $\tau_2 = 5$, $\mu_1 = \mu_2 = 1/2$, at $\omega\hat{\tau} = 3$.

An analogous lemma for the general case of $q \geq 0$ is given in Appendix A.

Remark: The techniques used here are significantly different from ones in the existing literature on linear stability of TCP, in two aspects. First, the usual approach is to find a convex hull that contains all individual $L_n(j\omega)$ curves and then argue any convex combination of them is still contained by the convex hull. See, e.g., [23], [18], [5]. However, the proof of Lemma 1 deals directly with $L(j\omega)$ instead of $L_n(j\omega)$. Second, for each ω, a separate region is found to bound $L(j\omega)$ away from the interval $(-\infty, -1]$. That is, the half plane $H(\omega)$ defined by (15) depends on ω. In existing works, convex regions are typically used to bound the whole curves and hence are independent of ω. One exception is [20], where the frequency range is divided into two parts and different convex regions are used in the two parts. With these two features, it is understandable that the technique presented here can generate tighter bounds, which is necessary for the analysis of this problem. □

We are now ready to prove the following main theorem.

Theorem 2. *If* $\gamma \leq 0.94$, *the model (10) for FAST TCP operating over a single link is locally stable.*

Proof. Define $F(\theta) = \cos(\theta) - \sin(\theta)/\theta$ and denote its

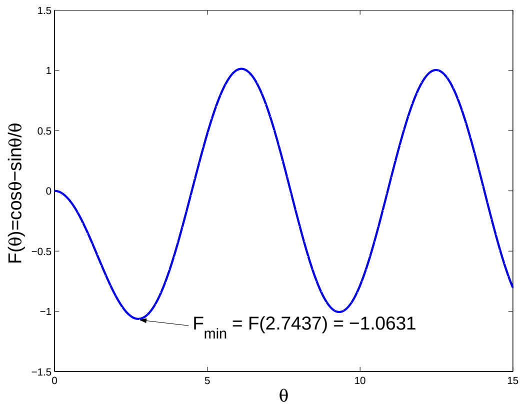


Fig. 2. $F(\theta) = \cos(\theta) - \sin(\theta)/\theta$

minimal value by $F_{\min}$. Then

$$\sum_{i=1}^{N} \frac{\mu_i}{\tau_i} \left(\cos(\omega\tau_i) - \frac{\sin(\omega\tau_i)}{\omega\tau_i} \right) \geq F_{\min} \sum_{i=1}^{N} \frac{\mu_i}{\tau_i} = \frac{F_{\min}}{\hat{\tau}}. \tag{21}$$

By (16), if $\gamma < -1/F_{\min}$, then the condition of Lemma 1 is satisfied and the Nyquist curve cannot encircle -1, since

$$(-\infty, -1] \cap \bigcup_{\omega > 0} H(\omega) = \emptyset.$$

It is straight forward to find $F_{\min}$ is -1.0631 which is attained when $\theta = 2.7437$, the smallest positive solution to $\tan(\theta) = \theta/(1-\theta^2)$. ($F(\theta)$ is plotted in Figure 2). Therefore, the assumption on γ guarantees

$$\gamma < 0.94 < \frac{1}{1.0631} = -\frac{1}{F_{\min}}. \tag{22}$$

□

Remark: Note that a default value of $\gamma = 0.5$ is used for the FAST implementation [27], Therefore Theorem 2 immediately establishes FAST TCP's stability for any pattern of round trip delays. This then also fully explains the fact that FAST TCP has been stable for all experimental cases studied. □

Theorem 2 is proved by finding a uniform bound for all flows' τ_i. If we have more detailed knowledge about the round-trip-delay distribution, we may achieve even better bounds. Let us now consider some special cases.

1) $N = 1$: If there is a single flow, the leading factor of (14) becomes 1 and the joint link model degenerate to the static link model. In this case, FAST is stable for all $\gamma < \pi/2$. In this case, (22) is loose simply because the frequency, $\omega = 2.7437/\tau$, which minimizes $F(\omega\tau)$ does not coincide with a frequency at which the Nyquist plot of $L(s)$ crosses the real axis.

2) $N = 2$: Consider two FAST flows with $\mu_1 = \mu_2 = 0.5$ (corresponding to the current practice that all FAST flows share the same α). Assume $\tau_1 = k\tau_2$ where k measures the heterogeneity. Define

$$F(\theta, k) = \cos(\theta) - \frac{\sin(\theta)}{\theta} + \frac{1}{k}\left(\cos(k\theta) - \frac{\sin(k\theta)}{k\theta} \right)$$

and let its minimal value over θ be $F_{\min}(k)$. It then follows that a sufficient condition for stability is

$$\gamma < \frac{k+1}{k} \frac{1}{-F_{\min}(k)}. \tag{23}$$

Table I lists the stability bounds for a few values of k. It is straightforward to show that the bound first increases and then decreases in k and asymptotically when $k \to \infty$, the bound is again 0.94, the same as the case of $k = 1$.

k	1	1.5	2	5	10	20
Stability bound	0.940	1.052	1.294	1.164	0.947	0.945

TABLE I

MAXIMUM γ FOR STABILITY FOR CASES WITH TWO FAST FLOWS WITH DIFFERENT RTT; HETEROGENEITY, k; AND $q = 0$

3) $N = \infty$: In reality, the link is likely to be shared by many flows. It is then interesting to find the statistical mean value of the stability bound for those scenarios. We will now consider the case of $q = 0$ with many flows with continuously distributed RTTs.

Let $M(\tau) = \sum_{\tau_i \leq \tau} \alpha_i/cq$, and let all τ_i be in the range $\Omega = (\underline{\tau}, \bar{\tau})$, with $\bar{\tau}$ possibly infinite. If there are many flows with RTTs drawn from a continuous distribution, then applying $\mu(\tau) = M'(\tau)$ to (16) gives

$$\int_{\Omega} \frac{\mu(\tau)}{\tau} \left(\cos(\omega\tau) - \frac{\sin(\omega\tau)}{\omega\tau} \right) d\tau + \frac{1}{\gamma} \int_{\Omega} \frac{\mu(\tau)}{\tau} d\tau > 0. \tag{24}$$

Noting that

$$\frac{d}{d\theta} \frac{\sin(\theta)}{\theta} = \frac{\cos(\theta)}{\theta} - \frac{\sin(\theta)}{\theta^2},$$

and setting $\theta = \omega\tau$, (24) becomes

$$\frac{1}{\gamma} > -\frac{\displaystyle\int_{\omega\Omega} \mu\left(\frac{\theta}{\omega}\right) \left(\frac{\sin(\theta)}{\theta}\right)' d\theta}{\displaystyle\int_{\omega\Omega} \frac{\mu(\theta/\omega)}{\theta} d\theta} \tag{25}$$

$$= -\frac{\left[\mu\left(\frac{\theta}{\omega}\right) \frac{\sin(\theta)}{\theta}\right]_{\omega\underline{\tau}}^{\omega\bar{\tau}} + \displaystyle\int_{\omega\Omega} \frac{\mu'(\theta/\omega)}{\omega} \frac{\sin(\theta)}{\theta} d\theta}{\displaystyle\int_{\omega\Omega} \frac{\mu(\theta/\omega)}{\theta} d\theta} \tag{26}$$

where $\omega\Omega = (\omega\underline{\tau}, \omega\bar{\tau})$ and $()'$ denotes derivative. This must hold for all $\omega > 0$.

As an example, assume RTTs follow a uniform distribution. As units of time are arbitrary, this can be modeled without loss of generality as

$$\mu(\tau) = \begin{cases} 1/(k-1) & \tau \in (1, k) \\ 0 & \text{otherwise}, \end{cases} \tag{27}$$

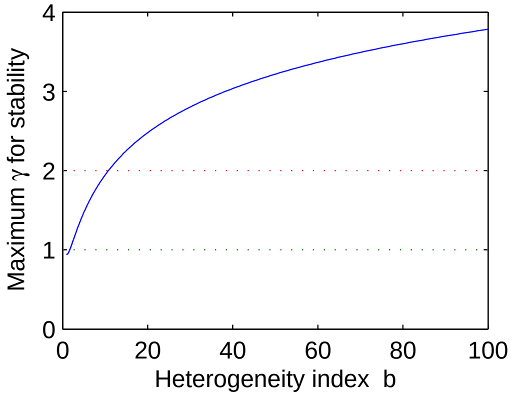


Fig. 3. Maximum value for γ for stability with RTTs in $(1, k)$.

with $k > 1$. In that case, (26) becomes

$$\frac{1}{\gamma} > \max_{\omega>0} \left(-\left[\frac{\sin(\theta)}{\theta}\right]_{\omega}^{\omega k} \Big/ \int_{\omega}^{\omega k} \frac{d\theta}{\theta} \right) \quad (28)$$

$$= \max_{\omega>0} \left(\frac{\sin(\omega)k - \sin(\omega k)}{\omega k \log(k)} \right). \quad (29)$$

The right hand side approaches $-F_{\min}$ as $k \to 1$, in accordance with (22), while for $k > 1$ the bound is strictly looser as shown in Figure 3, and asymptotes to $1/\log(k)$ for large k. Already for $k = 1.69$, the upper bound exceeds 1, while for $k = 10.79$ it becomes 2. If $\gamma > 2$ then the discrete rule (5) becomes intrinsically unstable since each update overshoots by more than the current error, and so it is not helpful to increase γ beyond 2 in the continuous time model.

Note that in this case, $\mu(\tau)$ reflects both the distribution of RTTs and also differences in α values of different flows. If all FAST flows use the same α which is true for the current implementation, then $\mu(\tau)$ is the distribution of RTTs.

D. (De)stabilized FAST TCP: closed loop experiment

Using the new joint link model, we have shown FAST TCP flows with default step sizes is always linear stable over a single link with any delay pattern. This is supported by the empirical experience so far. In this subsection, we will use cases with $\gamma > 1$ to compare stability predictions of all three models. There are two folds of objectives here. First to investigate the critical step size for FAST to maintain stability. This can potentially suggest larger step size for quicker response. Second, via comparing three models' predictions, This closed loop experiment will further strengthen the validation results in [11] where open loop experiments are conducted. The following example is used to serve these purposes.

Two FAST TCP flows share a single link with capacity of 10 000 pkt/s. The propagation delays of the two flows are 400 ms and 700 ms, respectively. Both flows use $\alpha = 50$. The open loop transfer functions for all three models and the critical step size(γ_c) for stability predicted by those models are summarized below. The integrator model predicts a critical step size much smaller than the one from the static model, while the joint model yields a prediction in between as expected.

- Integrator model: $\gamma_c = 1.23$

$$L(s) = \sum_{n=1}^{N} \mu_n \frac{1/\tau_n}{s + 1/\hat{\tau}} \frac{d_n \gamma_n e^{-\tau_n s}}{\tau_n^2 s + \gamma_n q} \quad (30)$$

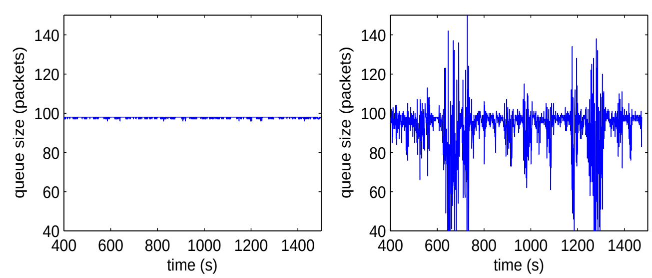
- Static model: $\gamma_c = 1.80$

$$L(s) = \sum_{n=1}^{N} \mu_n \frac{d_n \gamma_n e^{-\tau_n s}}{\tau_n^2 s + \gamma_n q} \quad (31)$$

- Joint model: $\gamma_c = 1.69$

$$L(s) = \sum_{n=1}^{N} \mu_n \frac{s + 1/\tau_n}{s + 1/\hat{\tau}} \frac{d_n \gamma_n e^{-\tau_n s}}{\tau_n^2 s + \gamma_n q} \quad (32)$$

We now report NS-2 packet level simulations [7].[4] Figure 4 shows the queue trajectories with $\gamma = 1.23$ and 1.80, the critical step size predicted by the integrator link model and the static link model. It is clear that the queue is not stable with $\gamma = 1.80$, which means the static model is too optimistic for stability analysis. We further show queue trajectories with $\gamma = 1.65$ and 1.75 in Figure 5. The case with $\gamma = 1.65$ is still stable which suggests that the integrator model is too conservative. Note the queue becomes to oscillate with $\gamma = 1.75$, we see that the critical step size prediction ($\gamma_c = 1.69$) from the new joint model is surprisingly accurate.



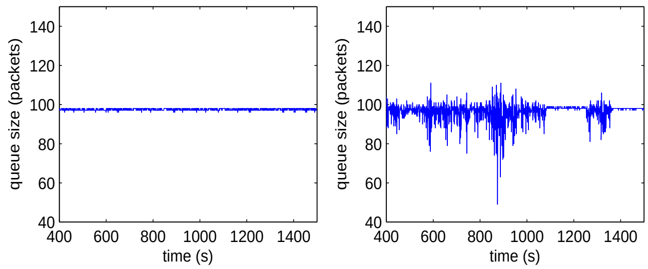
(a) $\gamma = 1.23$ (b) $\gamma = 1.80$

Fig. 4. Queue trajectories with critical step sizes predicted by the integrator link model and the static link model.

IV. Conclusion

We have reviewed a recent proposed link model which captures the queue dynamics when congestion windows of TCP sources change. The model is shown to be much more accurate than existing ones. It agrees with the known static link model when flows' round trip delays are similar, and approximates the standard integrator link model when the heterogeneity of round trip delays is significant. Using this new model, we have shown that FAST TCP is always stable over a single bottleneck link. This extends the existing stability result for homogeneous FAST flows to cases with

[4]To validate the link model, the code was modified to update the window once per RTT, and for modeling simplicity the RTT estimate was evaluated over 0.1 RTT. All queue trajectories are plotted after initial transients, to emphasize the local stability of the congestion avoidance phase.



(a) $\gamma = 1.65$ (b) $\gamma = 1.75$

Fig. 5. Queue trajectories around critical step sizes.

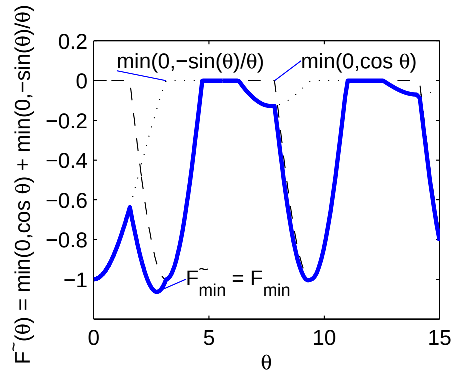


Fig. 6. Plot of $\tilde{F}(\theta)$ along with $\min(0, \cos(\theta))$ and $\min(0, -\sin(\theta)/\theta)$.

heterogeneous delays, and resolves the notable discrepancy between empirical observations and existing theoretical predictions. In particular, we are able to predict the stability region of the closed loop system which is accurately verified by packet level simulations.

References

[1] T. Alpcan and T. Basar. Global stability analysis of an end-to-end congestion control scheme for general topology networks with delay. In *Proceedings of IEEE Conference on Decision and Control*, 2003.

[2] E. Altman, C. Barakat and V. Ramos. Analysis of AIMD protocols over paths with variable delay. In *Proceedings of IEEE Infocom*, 2004.

[3] L. L. H. Andrew, S. V. Hanly and R. G. Mukhtar. CLAMP: A system to enhance the performance of wireless access networks In *Proceedings of IEEE Globecom*, pp. 4142-4147, 2003.

[4] F. Baccelli and D. Hong. AIMD, Fairness and fractal scaling of TCP traffic. In *Proceedings of IEEE Infocom*, 2002.

[5] H. Choe and S. H. Low. Stabilized Vegas. In *Proceedings of IEEE Infocom*, 2003.

[6] J. Choi, K. Koo, J. Lee and S. H. Low. Global stability of FAST TCP in single-link single-source network. In *Proceedings of IEEE Conference on Decision and Control*, 2005.

[7] T. Cui and L. Andrew. FAST TCP simulator module for ns-2, version 1.1 Available ⟨http://www.cubinlab.ee.mu.oz.au/ns2fasttcp⟩.

[8] S. Deb and R. Srikant. Global stability of congestion controllers for the Internet. *IEEE/ACM Transactions on Automatic Control*, 48(6):1055-1060, June 2003.

[9] C. Hollot and Y. Chait. Nonlinear stability analysis for a class of TCP/AQM schemes. In *Proceedings of IEEE Conference on Decision and Control*, 2001.

[10] C. Hollot, V. Misra, D. Towsley and W. Gong. A control theoretic analysis of RED. In *Proceedings of IEEE Infocom*, 2001.

[11] K. Jacobsson, H. Hjalmarsson, and N. Möller. ACK-clock dynamics in network congestion control – an inner feedback loop with implications on inelastic flow impact. In *Proceedings of the 45th IEEE Conference on Decision and Control*, San Diego, USA, December 2006, *to appear*.

[12] R. Johari and D. Tan. End-to-end congestion control for the Internet: delays and stability. *IEEE/ACM Transactions on Networking*, 9(6):818-832, December 2001.

[13] K. Kim, A. Tang and S. H. Low. Design of AQM in supporting TCP based on the well-known AIMD model. In *Proceedings of IEEE Globecom*, 2003.

[14] K. Kim, A. Tang and S. H. Low. A stabilizing AQM based on virtual queue dynamics in supporting TCP with arbitrary delays. In *Proceedings of IEEE CDC*, 2003.

[15] S. Liu, T. Basar and R. Srikant. Pitfalls in the fluid modeling of RTT variations in window-based congestion control. In *Proceedings of IEEE Infocom*, 2005.

[16] Y. Liu, F. L. Presti, V. Misra, D. Towsley, and Y. Gu. Fluid models and solutions for large-scale IP networks In *Proceedings of the 2003 ACM SIGMETRICS*, 2003.

[17] S. H. Low, F. Paganini, and J. C. Doyle. Internet congestion control. *IEEE Control Systems Magazine*, 22(1):28–43, Feb. 2002.

[18] S. H. Low, F. Paganini, J. Wang, and J. C. Doyle. Linear stability of TCP/RED and a scalable control. *Computer Networks Journal*, 43(5):633–647, 2003.

[19] L. Massoulie Stability of distributed congestion control with heterogeneous feedback delays. *IEEE Transactions on Automatic Control*, 47(6): 895-902, June 2002

[20] F. Paganini, Z. Wang, J. C. Doyle and S. H. Low Congestion control for high performance, stability and fairness in general networks. *IEEE/ACM Transactions on Networking*, 13(1):43-56, February 2005.

[21] R. Shorten, F. Wirth and D. Leith Modelling TCP congestion control dynamics in drop-tail environments. *Automatica*, in press.

[22] R. Srikant. *The Mathematics of Internet Congestion Control*. Birkhauser, 2004.

[23] G. Vinnicombe. On the stability of networks operating TCP-like protocols. In *Proceedings of IFAC*, 2002.

[24] J. Wang, A. Tang, and S. H. Low. Local stability of FAST TCP. In *Proceedings of IEEE Conference on Decision and Control*, December 2004.

[25] J. Wang, D. X. Wei, and S. H. Low. Modeling and stability of FAST TCP. In *Proceedings of IEEE Infocom*, 2005.

[26] Z. Wang and F. Paganini. Global stability with time-delay in network congestion control. In *Proceedings of IEEE Conference on Decision and Control*, December 2002.

[27] D. Wei, C. Jin, S. H. Low, and S. Hegde. FAST TCP: motivation, architecture, algorithms, performance. *To appear in IEEE/ACM Transactions on Networking*, 2007.

[28] L. Ying, G. Dullerud and R. Srikant. Global stability of Internet congestion controllers with heterogeneous delays. In *Proceedings of American Control Conference*, 2004.

Appendix A
Bounds on gain for non-negligible queues

For the general case of $q \geq 0$, we will prove a weaker form of Lemma 1 which is still sufficient to prove Theorem 2.

Lemma 3. *Let*

$$\tilde{F}(\theta) = \min(0, \cos(\theta)) + \min\left(0, -\frac{\sin(\theta)}{\theta}\right). \qquad (33)$$

If $\gamma \leq (\pi/2)^2$ *and*

$$\sum_{i=1}^{N} \frac{\mu_i}{\tau_i} \tilde{F}(\omega\tau_i) + \frac{1}{\gamma\hat{\tau}} > 0 \qquad (34)$$

then $L(j\omega) \in H(\omega)$. *Moreover,* $\tilde{F}_{\min} := \min_{\theta \geq 0}(\tilde{F}(\theta)) = F_{\min}$.

The function $\tilde{F}(\theta)$ is shown in Figure 6, along with its two constituent terms.

Proof. By (12), a sufficient condition for $L(j\omega) \in H(\omega)$ is

$$\arg\left(1+\sum_{i=1}^{N}\gamma\mu_i\frac{j\omega+1/\tau_i}{j\omega+1/\hat{\tau}_i}\frac{\tau_i-q}{j\omega\tau_i^2+\gamma q}e^{-j\omega\tau_i}\right) - \arg(j+\omega\hat{\tau}) \in \quad (-\pi,0). \tag{35}$$

Using (19), this is equivalent to

$$\arg\left(\frac{j\omega+\frac{1}{\hat{\tau}}}{\gamma}+\sum_{i=1}^{N}\frac{\mu_i}{\tau_i}\frac{(j\omega+\frac{1}{\tau_i})(\tau_i-q)}{j\omega\tau_i+\gamma q/\tau_i}e^{-j\omega\tau_i}\right)\in\left(-\frac{\pi}{2},\frac{\pi}{2}\right). \tag{36}$$

Thus it is sufficient that the real part of the left hand side be strictly greater than 0 for all ω:

$$\sum_{i=1}^{N}\frac{\mu_i}{\tau_i}\left(1-\frac{b_i}{\gamma}\right)\left(\frac{a_i^2+b_i}{a_i^2+b_i^2}\cos(a_i)-\frac{a_i^2-a_i^2b_i}{a_i^2+b_i^2}\frac{\sin(a_i)}{a_i}\right) + \frac{1}{\gamma\hat{\tau}} > 0 \tag{37}$$

where $a_i = \omega\tau_i$ and $b_i = \gamma q/\tau_i$. Because $q/\tau_i \leq 1$, it is sufficient that

$$\sum_{i=1}^{N}\frac{\mu_i}{\tau_i}\left(A\min(0,\cos(a_i))+B\min\left(0,-\frac{\sin(a_i)}{a_i}\right)\right) + \frac{1}{\gamma\hat{\tau}} > 0 \tag{38}$$

where $A = (1-b_i/\gamma)(a_i^2+b_i)/(a_i^2+b_i^2)$ and $B = (1-b_i/\gamma)(a_i^2-a_i^2b_i)/(a_i^2+b_i^2)$.

Note that $B \in [0,1]$. Also, $A \geq 0$ and

$$A \leq \frac{a_i^2+b_i(1-(a_i^2+b_i)/\gamma)}{a_i^2+b_i^2},$$

giving $A \leq 1$ if $1-(a_i^2+b_i)/\gamma \leq b_i$, which is true if $a_i > \pi/2$, since $\gamma \leq (\pi/2)^2$. Thus $A \in [0,1]$ when $\cos(a_i) \leq 0$. Since each term in (38) is non-positive, it is bounded below by $(\mu_i/\tau_i)\tilde{F}(\omega\tau_i)$ and the first part of the lemma is established.

It remains to show $\tilde{F}_{\min} = F_{\min}$. Now $\tilde{F}_{\min} \leq F_{\min} < 1$. For $\tilde{F}_{\min} < 1$, the minimum must occur when both $\cos(\omega)$ and $\sin(\omega)/\omega)$ are negative as both are bounded in magnitude by 1. When that occurs, $F(\omega) = \tilde{F}(\omega)$, giving $\tilde{F}_{\min} \geq F_{\min}$, and hence the result. □