On the Long-Run Behavior of Equation-Based Rate Control

Milan Vojnović and Jean-Yves Le Boudec

To appear in ACM/IEEE Transactions on Networking, April 2005

***Abstract*— We consider unicast equation-based rate control, where, at some points in time, a sender adjusts its rate to f(p,r), where p is an on-line estimate of the loss-event rate observed by this source, r of the average round-trip time, and f is a TCP throughput formula. Conventional wisdom holds that such a source would be TCP-friendly, that is, it would not attain a larger long-run average send rate than a TCP source under the same operating conditions. Our goal is to identify the key factors that determine whether, and how far, this is true. We point out that it is important to breakdown the TCP-friendliness condition into sub-conditions and study them separately. One sub-condition is conservativeness (throughput not larger than f(p,r)). The conservativeness is primarily influenced by some convexity properties of the function f, and a covariance property of the loss process. In many cases, these conditions result in conservativeness, in some cases, excessive conservativeness. Another sub-condition is that the source experiences a loss-event rate that is not smaller than that of TCP. We show two limit cases for which the last sub-condition, respectively, does and does not hold. We show that in the latter situation, the outcome can be a significant non-TCP-friendliness. The claims suggested by our analysis are verified by numerical examples, simulations, Internet and lab experiments. Our findings should help us better understand when to expect the source to be TCP-friendly, or in contrast, non-TCP-friendly. On the basis of our analysis and empirical evaluations we observe that TCP-friendliness is difficult to verify, whereas conservativeness is easier.**

I. Introduction

We consider an adaptive sender of packets that employs unicast equation-based rate control. It means that the send rate is controlled as follows: The sender uses the function $f(p, r)$ that maps a loss-event rate p and an average round-trip time r to a send rate; the sender computes on-line estimates of the loss-event rate and the average round-trip time and plugs-in those estimates into the function f; the sender sets its send rate to such computed values primarily at loss-events. The function f is assumed to be a TCP throughput formula. An example equation-based rate control protocol is TFRC [7]. Given that the send rate is adjusted at *some* points in time by computing the send rate using the TCP throughput formula, conventional wisdom holds that equation-based rate control is TCP-friendly. TCP-friendliness is an axiom that requires adaptive non-TCP sources to attain a long-run time-average send rate (we call throughput) not larger than TCP would attain under the same circumstances [5]. If so, a source is said to be TCP-friendly. If there were convergence, then the throughput x of the source would satisfy $x = f(p, r)$. In practice, though, the control is required to be responsive, so there is no convergence. Thus, the non-linearity of the function f and the randomness of the estimators of p and r leave little hope that we have exactly $x = f(p, r)$. We show when to expect an equation-based rate control to be TCP-friendly, or non-TCP-friendly. Most of our results are based on analysis; the claims suggested by our analysis are verified through numerical examples, simulations, Internet and lab experiments. Next, we summarize our main findings.

A. Breakdown into Sub-Conditions

We find that it is important to breakdown the TCP-friendliness condition into sub-conditions and study them separately. We give the arguments after introducing the sub-conditions. There are four sub-conditions whose conjunction implies TCP-friendliness. The first sub-condition we term *conservativeness*; it means that the source attains a throughput that is not larger than the TCP throughput formula that is used by the source, evaluated at the loss-event rate and average round-trip time as seen by *this* protocol. Note that conservative and TCP-friendly are not the same; it is perfectly possible that our source and TCP would experience different loss-event rates and average round-trip times. The second sub-condition is that the loss-event rate as experienced by our source is not smaller than TCP would experience. The third sub-condition is that the average round-trip time as observed by our source is not smaller than

Part of the results was presented at ACM SIGCOMM'02, Pittsburgh, PA, 2002, and ITC-17, Salvador da Bahia, Brazil, 2001.

Author Affiliations: Milan Vojnović and Jean-Yves Le Boudec are with EPFL, CH-1015 Lausanne, Switzerland, <milan.vojnovic@epfl.ch>, <leboudec@epfl.ch>

that of TCP. The fourth, and last, sub-condition is that TCP attains a throughput that is at least as large as the TCP throughput formula evaluated at the loss-event rate and average round-trip time experienced by TCP. Common practice is to evaluate TCP-friendliness of a protocol by directly comparing the throughputs of this protocol and that of a competing TCP. We point out that this is not a good practice because it may hide the cause of an observed throughput deviation. It is thus important for protocol designers to breakdown the TCP-friendliness condition and study the individual sub-conditions separately. Failing to do so may lead a protocol designer to an improper adjustment of the protocol with the intention to correct for an observed throughput deviation. We illustrate this by an example. Suppose that in a set of experiments a protocol designer observes that her equation-based rate control protocol attains a larger throughput than TCP. The designer then adjusts the protocol by choosing a throughput function that is smaller than the TCP throughput function used in the experiments for a factor that would compensate for the observed throughput deviation. Upon repeating the experiments, the designer may observe that the fix makes the two protocols attain almost the same throughputs, and thus the fix solved the problem. Suppose that for the given set of experiments, the breakdown of the TCP-friendliness condition would reveal that the throughput deviation was due to the equation-based rate control seeing a much smaller loss-event rate than TCP. Now, suppose that our designer runs another set of experiments, but now for those new experiments the loss-event rates of the two protocols are fairly near. Having made the fix, the designer may observe that now her protocol attains an excessivelly smaller throughput than TCP. The reason is that the fix has been done in an ad-hoc manner, by not directly trying to understand and correct for the deviation of the loss-event rates, which in the example was supposed to be the major cause of the deviation.

B. Conservativeness

We give conditions under which the control is conservative, or not. Before announcing our results, we first define our source in some more detail. We assume that our source uses an unbiased estimator of the reciprocal of the loss-event rate $1/p$. Moreover, we assume that the estimator is defined as a moving-average of a fixed number of loss-event intervals as observed by the source. We call a loss-event interval the number of packets sent by our source between two successive loss-events experienced by the source. In particular, the definition of the estimator accommodates a protocol such as TFRC. We distinguish between *basic* and *comprehensive* control. The former adjusts the send rate only at loss-events, whereas the latter increases the send rate on time intervals when the loss-event interval becomes sufficiently large. The comprehensive control reflects what is implemented in TFRC. Some of our conditions are on some analytical properties of the function f. We verify our conditions on the following cases for function f: (1) the square-root formula (we call SQRT), (2) the throughput formula (we call PFTK-standard), and (3) a simplified variant of PFTK-standard (we call PFTK-simplified). We consider PFTK-simplified because it is recommended by the TFRC proposed standard [7]. We also consider PFTK-standard because PFTK-simplified is a simplification of that function, and SQRT that is the limit of both PFTK formulae for rare losses. Our analysis reveals that the properties of the long-run behavior is indeed influenced by the chosen throughput function. We did not consider TCP throughput formulae for short-lived transfers, e.g. [4], given that our focus is on long-run behavior of long-lived connections.

In our analysis of the conservativeness, we assume the estimator of the expected round-trip time is fixed to the mean round-trip time; we account for variability of the round-trip times in our empirical evaluations. For the basic control, we find fairly exhaustive results. First, conservativeness is strongly influenced by the convex or concave nature of two functionals of f and the joint probability law of the loss-event intervals. If (C1) the loss statistics is such that the estimator $\hat{\theta}_n$ of the expected loss-event interval and the next loss-event interval θ_n are slightly positively or negatively correlated, then the control is conservative. We find empirical evidence in [18] that suggests (C1) in practice to be true; we also show later our own empirical confirmation. Further, the larger the variability of the loss-event interval estimator, the more conservative the control is. Both of these effects are more pronounced for PFTK formulae than for SQRT. With PFTK, this causes the control to be excessively conservative for a large loss-event rate. This explains the *throughput-drop* with the loss-event rate, which was empirically observed for TFRC [6], [17], [2], the cause of which was so far unknown. Second, if the covariance condition (C1) does not hold, then the results may be radically different, and strongly depend on the nature of the function f. We identified one case of practical importance that results in non-conservativeness. If (C2c) the correlation of the send rate just after a loss-event and the time it takes until the next loss-event is non-negative, then for PFTK and a large loss-event rate, the control is systematically non-conservative. For PFTK with a small or moderate loss-event rate, this does not occur. The effect is due to the convexity property of the function f, that holds differently in these cases. An example protocol that may conform to the foregoing assumptions is an audio source that has a fixed packet send rate and adjusts its send rate by varying the packet lengths [3]. These findings are exact for the basic control; for the comprehensive control, we pose them as claims and verify them by experiments.

C. Comparison of the Loss-event Rates

We analyze two limit cases under which the sub-condition on the order of the loss-event rates, respectively, does and does not hold. The former limit case is for a source that has a negligible effect on the state of the network, and it is *driven* by a congestion process that captures the network dynamics. In this situation, the problem boils down to the analysis of the *sampling* of the congestion process by the source. Given that our source is expected to be more sluggish in adaptation and to have a smoother send rate than TCP, our analysis suggests that the source would observe a larger loss-event rate than TCP. The latter limit case is when a small number of equation-based rate control and TCP sources compete for a bottleneck. We consider perhaps the simplest model whereby our source and a TCP-like source each run alone over a link with a fixed capacity. Our analysis yields that the TCP-like source observes a loss-event rate that is 1.7 times larger than the loss-event rate observed by the equation-based rate controlled source. This suggests that, in general, our source would see a smaller loss-event rate than TCP when a few equation-based rate control and TCP sources compete for a bottleneck. We verify this by simulations and Internet measurements, and find a fairly good agreement with the suggestion of our simple model. We emphasize that in the present situation, the deviation of the loss-event rates is the major cause of the throughput deviation that makes equation-based rate control be significantly non-TCP-friendly.

D. Other Sub-Conditions

We study through empirical evaluations only, the sub-conditions on the comparison of the average round-trip times and obedience of TCP to the throughput formula. We did observe some deviation of the average round-trip times of TFRC and TCP in our experiments, but in most cases we observed this not to have a major influence on TCP-friendliness. We observed cases when TCP attained a smaller throughput than given by PFTK formulae. In particular, we observed a tendency of this to hold for a few sources competing for a bottleneck.

E. TCP-friendliness is Difficult to Verify

On the basis of our analysis and empirical evaluations, we find that *TCP-friendliness is difficult to verify*. There are several factors that determine whether the control is TCP-friendly, or not. The different factors may lead to different directions. In contrast, and on a more positive side, our analysis and empirical results suggest that *conservativeness is easier to analyze and verify*. The conservativeness condition allows us to study a source that uses equation-based rate control in *isolation* for a driving loss process, and come up with simple conditions on analytical properties of the throughput function and statistics of the loss process that imply conservativeness or non-conservativeness.

F. Organization of the Paper

The paper is organized as follows. Section II describes our assumptions and notations. Section III gives our analysis results that tell when the control is conservative, or non-conservative. We summarize our findings in the form of two claims in Section III-C. Section IV discusses the other sub-conditions of the TCP-friendliness breakdown. Firstly, we give analytical arguments that in a limit case lead us to conclude that equation based rate control would not see a smaller loss-event rate than TCP. Secondly, we pose a claim that describes a different situation in which the order of the loss-event rates is opposite. In Section V we validate our findings by designed numerical experiments, ns-2, lab, and Internet experiments. Section VI concludes. All the proofs are given in the appendix.

II. Additional Assumptions and Notations

We consider an adaptive sender with the send rate at time t equal to $X(t)$. We assume that $X(t)$ can be described by a stationary ergodic process, and thus equate the time-average with the expected value

$$\bar{x} = \lim_{t\to\infty} \frac{1}{t}\int_0^t X(s)ds = \mathbf{E}[X(0)].$$

Index n refers to the nth loss-event. We use the following additional notation. T_n is the instant at which a loss-event labeled n is detected by the sender. $S_n = T_{n+1} - T_n$ is an inter loss-event time. $X_n = X(T_n)$ is the rate set at the nth loss-event. θ_n is the number of packets sent in $[T_n, T_{n+1})$. Following TFRC, we call θ_n, the loss-event interval. Think of both S_n and θ_n as of the loss-event intervals, however, the former measured in seconds, and the latter in packets.

We study the long-run behavior of the control, and hence, it is more convenient to work under the convention $\ldots T_{-1} < T_0 \le 0 < T_1 < \ldots$. The instant 0 is an arbitrary point in time. We assume the point process of loss-events has a finite non-null intensity λ. The quantity λ is the loss-event rate, the expected number of loss events on an arbitrary unit time interval. With $\mathbf{E}_N^0$ we denote the expectation with respect to the Palm probability $\mathbf{P}_N^0$, associated to the points of loss-events. This is the expectation as seen at an arbitrary loss event instant T_n, as opposed to the standard expectation $\mathbf{E}$, which is as seen at an arbitrary point in time.

The loss-event rate as observed by the source is

$$p = \frac{1}{\mathbf{E}[\theta_0]}. \tag{1}$$

The quantity p is also a loss-event rate. Informally speaking, it is the fraction of loss-events one would observe in a number of packets sent over a long time interval. Let $\hat{\theta}_n$ be an estimator of the expected loss-event interval in packets, computed at T_n. We assume, as is done in TFRC:

(E) $\hat{\theta}_n$ is an unbiased estimator of $1/p$.

Moreover, we assume that $\hat{\theta}_n$ is defined as a moving-average of the loss-event intervals, for some positive-valued weights $(w_1, w_2, \ldots, w_L)$,

$$\hat{\theta}_n = \sum_{l=1}^{L} w_l \theta_{n-l}. \tag{2}$$

Note that, by (E), the weights sum up to unity.[1] TFRC uses this type of loss-event interval estimator, for a particular setting of the weights, with w_l all equal for $1 \le l \le L/2$, else w_l linearly decreases with l.

In our analysis we assume the round-trip times are fixed to their mean value. Hence, for convenience of notation, we re-define $f(p)$ be a positive-valued non-increasing function of the loss-event rate p.

A. Basic control

The basic control is defined as follows. For $t \in [T_n, T_{n+1})$, $n \in \mathbb{Z}$,

$$X(t) = f\left(\frac{1}{\hat{\theta}_n}\right). \tag{3}$$

B. Comprehensive control

We add an additional control law to the basic control (3), and call the resulting system the comprehensive control. The mechanism reflects the send rate increase in absence of loss-events, as found in TFRC [7].

[1] As an aside, note that $\mathbf{E}[1/\hat{\theta}_0] \ge p$, and thus $1/\hat{\theta}_n$ is a biased estimator of p. This follows as a direct application of Jensen's inequality and (1).

Let $\theta(t)$ be the number of packets sent since the last loss-event observed at t. Then we define the comprehensive control as follows, for $t \in [T_n, T_{n+1})$, $n \in \mathbb{Z}$,

$$\begin{aligned} X(t) &= f\left(\frac{1}{\hat{\theta}(t)}\right), \qquad (4)\\ \hat{\theta}(t) &= (w_1\theta(t) + \sum_{l=1}^{L-1} w_{l+1}\theta_{n-l})\mathbf{1}_{A_t} + (1-\mathbf{1}_{A_t})\hat{\theta}_n. \end{aligned}$$

Here

$$A_t = \left\{\theta(t) > \frac{1}{w_1}[\hat{\theta}_n - \sum_{l=1}^{L-1} w_{l+1}\theta_{n-l}]\right\},$$

where $\mathbf{1}_{A_t} = 1$, if A_t is true, else $\mathbf{1}_{A_t} = 0$.

In other words, at an instant t, the loss-event interval estimator $\hat{\theta}(t)$ is updated with $\theta(t)$, if that increases the value of the estimator. If this is not the case, then $\hat{\theta}(t)$ is fixed to $\hat{\theta}_n$. Note that if the condition A_t is true, that is $\theta(t)$ is sufficiently large, the control (4) increases its send rate.

Note that the send rate dynamics is such that, if $\hat{\theta}_{n+1} \le \hat{\theta}_n$, then $X(t) = f(1/\hat{\theta}_n)$, all $t \in [T_n, T_{n+1})$. Else, for $\hat{\theta}_{n+1} > \hat{\theta}_n$, the send rate is $X(t) = f(1/\hat{\theta}_n)$, for $t \in [T_n, T_n + U_n]$, and then the rate increases according to (4) for $t \in (T_n + U_n, T_{n+1})$. Here, from the definition of A_t,

$$U_n = \frac{1}{w_1 f\left(\frac{1}{\hat{\theta}_n}\right)}\left[\hat{\theta}_n - \sum_{l=1}^{L-1} w_{l+1}\theta_{n-l}\right].$$

C. Some functions $x \to f(x)$ used in the Internet

We use the following loss-throughput formulae. We first display the simplest one ("the square-root") which we refer to as SQRT [10]

$$f(p) = \frac{1}{c_1 r\sqrt{p}}, \tag{5}$$

where c_1 is a positive constant, r is an event-average of the round-trip time; the event-average by sampling the round-trip times once in a round-trip time round.

We next display another well-known function f (Equation (30) in [12]); we refer to as PFTK-standard

$$f(p) = \frac{1}{c_1 r\sqrt{p} + q\min[1, c_2\sqrt{p}](p + 32p^3)}, \tag{6}$$

for a positive constant c_2. The parameter q is a value of the TCP retransmit timeout parameter. We also consider a simplified version of the last formula; we call, PFTK-simplified

$$f(p) = \frac{1}{c_1 r\sqrt{p} + qc_2(p^{3/2} + 32p^{7/2})}. \tag{7}$$

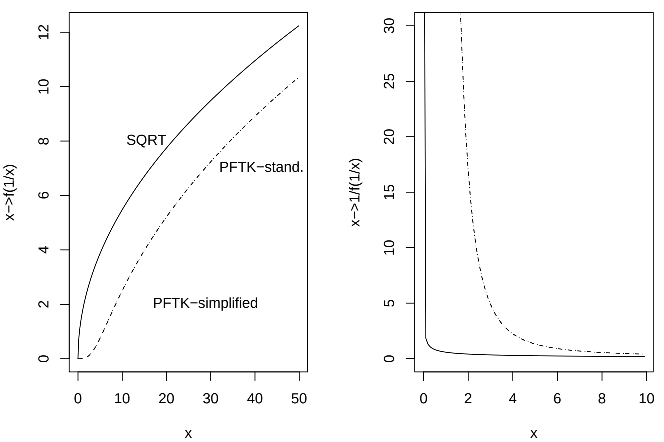

Fig. 1. Functions of interest (Left) $x \to f(1/x)$ and (Right) $x \to 1/f(1/x)$, for SQRT, PFTK-standard, and PFTK-simplified, with $r = 1$, $q = 4r$; the curves for PFTK formulae overlap. Values of x close to 0 correspond to heavy losses. The plots on the right side indicate that the convexity condition (F1) in Theorem 1 would be satisfied in all three cases, but this is strictly true only for SQRT and PFTK-simplified; it also illustrates that convexity is much more pronounced for PFTK-simplified than for SQRT. The left plots illustrate that the concavity condition (F2) of Theorem 2 is true for SQRT; for PFTK-standard and PFTK-simplified it holds only for small loss-event rates; for heavy loss (x small), the curves are convex and thus the opposite condition (F2c) holds.

In the formulae, $c_1 = \sqrt{2b/3}$ and $c_2 = 3/2\sqrt{3b/2}$. b is the number of the packets acknowledged by a single acknowledgment; in practice, typically $b = 2$.

PFTK formulae are de-facto standard. PFTK-simplified is the formula recommended in TFRC standard proposal [7]. (With $q = 4r$, as a recommendation.) Note that, for $p \leq 1/c_2^2$, (7) is equal to (6), else, it is smaller.

We use the above particular formulae in our examples. Note that most of our findings apply to other functions f as well.

III. WHAT MAKES THE CONTROL CONSERVATIVE OR NOT

We first give a throughput formula on which we build our analysis.

A. Throughput formulae

Proposition 1: The throughput of the basic control (3) is

$$\mathbf{E}[X(0)] = \frac{\mathbf{E}[\theta_0]}{\mathbf{E}[\frac{\theta_0}{f(\frac{1}{\hat{\theta}_0})}]}. \qquad (8)$$

Comment: We believe it is instructive to re-write the throughput expression (8) as

$$\mathbf{E}[X(0)] = \frac{1}{\mathbf{E}[\frac{1}{f\left(\frac{1}{\hat{\theta}_0}\right)}]} \frac{1}{1 + \frac{\mathbf{cov}[\theta_0, f(1/\hat{\theta}_0)]}{\mathbf{E}[\theta_0]\mathbf{E}[f(1/\hat{\theta}_0)]}}.$$

The expression reveals that, in part, it is the convexity nature of $x \to 1/f(1/x)$ that would determine conservativeness. The second term in the above display may or may not have a significance. In a particular case, when the loss-event interval estimator and the next sample of the loss-event interval are stochastically independent, then the second term in the above display is equal to one. In this particular case, it is the convexity nature of $x \to 1/f(1/x)$ that entirely determines whether the control is conservative, or not.

For the comprehensive control, we have a bound.

Proposition 2: The throughput of the comprehensive control (4) is such that

$$\mathbf{E}[X(0)] \geq \frac{\mathbf{E}[\theta_0]}{\mathbf{E}[\frac{\theta_0}{f(\frac{1}{\hat{\theta}_0})}]}.$$

The bound implies, if we know the basic control is non-conservative, then we know the comprehensive control is non-conservative, as well. The converse is not true.

An exact throughput expression may be obtained for the comprehensive control for particular functions f. We obtain exact throughput expression for SQRT and PFTK-simplified functions f, as given next.

Proposition 3: Let f be either SQRT ($c_2 = 0$) or PFTK-simplified. The throughput of the comprehensive control is

$$\mathbf{E}[X(0)] = \frac{\mathbf{E}[\theta_0]}{\mathbf{E}[\frac{\theta_0}{f(\frac{1}{\hat{\theta}_0})}] - \mathbf{E}[V_0 \mathbf{1}_{\hat{\theta}_1 > \hat{\theta}_0}]}, \qquad (9)$$

where

$$V_n = \frac{1}{w_1}\Big[-2c_1 r(\hat{\theta}_{n+1}^{\frac{1}{2}} - \hat{\theta}_n^{\frac{1}{2}}) + 2c_2 q(\hat{\theta}_{n+1}^{-\frac{1}{2}} - \hat{\theta}_n^{-\frac{1}{2}}) + \frac{64}{5}c_2 q(\hat{\theta}_{n+1}^{-\frac{5}{2}} - \hat{\theta}_n^{-\frac{5}{2}}) + (\hat{\theta}_{n+1} - \hat{\theta}_n)\frac{1}{f(1/\hat{\theta}_n)}\Big].$$

Note that, in view of the above propositions and the definition of $\hat{\theta}_n$ (2), the throughput of both basic and comprehensive control is expressed in terms of the expected values of some functions of a sequence of $L + 1$ loss-event intervals, $\theta_0, \theta_{-1}, \ldots, \theta_{-L}$. Therefore, knowing the joint probability law of $\theta_0, \theta_{-1}, \ldots, \theta_{-L}$ would, at least in theory, enable us to compute the throughput, and explain how the “correlation structure” of the loss process plays a role.

B. Conditions for the basic control to be conservative

Consider the basic control. We give exact sufficient conditions for conservativeness, or non-conservativeness. The results have interest of their own; they suggest the key factors that may cause conservativeness.

1) Sufficient conditions for the basic control to be conservative:

Theorem 1: Assume that

(F1) $x \to \frac{1}{f(1/x)}$ is convex,

(C1) $\mathbf{cov}[\theta_0, \hat{\theta}_0] \leq 0$.

Then the basic control (3) is conservative.

Interpretation: The convexity condition (F1) is satisfied by the SQRT loss-throughput formula, and by PFTK-simplified; it is not satisfied by PFTK-standard, but almost (we will come back to this in a few lines). This is straightforward to demonstrate and can also be seen in Figure 1. The figure also shows that convexity is much more pronounced for PFTK formulae, and thus, we should expect stronger conservativeness with PFTK than with SQRT formula (this is confirmed numerically later).

Condition (C1) is true in particular when the covariance is 0, which happens when successive loss-event intervals are (stochastically) independent. There are indications in [18] that this may be true, and the theorem says that this would lead to a conservative behavior. We show later, in Section V, some experimental evidence that indicate (C1) to be mostly true in today's Internet (see Figure 10).

We give the following more explicit statement, which gives a bound on the throughput, for $\mathbf{cov}[\theta_0, \hat{\theta}_0]p^2 < -f(p)/(f'(p)p)$,

$$\mathbf{E}[X(0)] \leq f(p) \frac{1}{1 + \frac{f'(p)p}{f(p)} \mathbf{cov}[\theta_0, \hat{\theta}_0]p^2}. \tag{10}$$

This shows that, in most cases, if the covariance is positive but small, there cannot be a significant non-conservativeness of the basic control.

The theorem says more. Remember that $\hat{\theta}_n$ is an incremental estimator of the Palm expectation of the loss-event interval, $1/p$, built on the past information available up to the loss-event n, whereas θ_n is the true next loss-event interval. Both have the same expectation, as we assumed that $\hat{\theta}_n$ is unbiased. However, this does not mean that $\hat{\theta}_n$ is a good *predictor* of θ_n. This depends on the joint statistics, in particular the autocovariance of the loss process. The covariance of θ_n and $\hat{\theta}_n$ reflects how good a predictor $\hat{\theta}_n$ is. Condition (C1) means that $\hat{\theta}_n$ is a bad predictor, and, maybe surprisingly, the theorem suggests that this leads to a conservative behavior. Conversely, consider now a hypothetical case where the loss process goes into phases, with slow transitions. Then the loss-event interval becomes highly predictable; the theorem does not say that this alone will make the control non-conservative. However, this may really happen; we gave an example in [16], Section 3.4. We give a perhaps more realistic example in Section III-B.2.

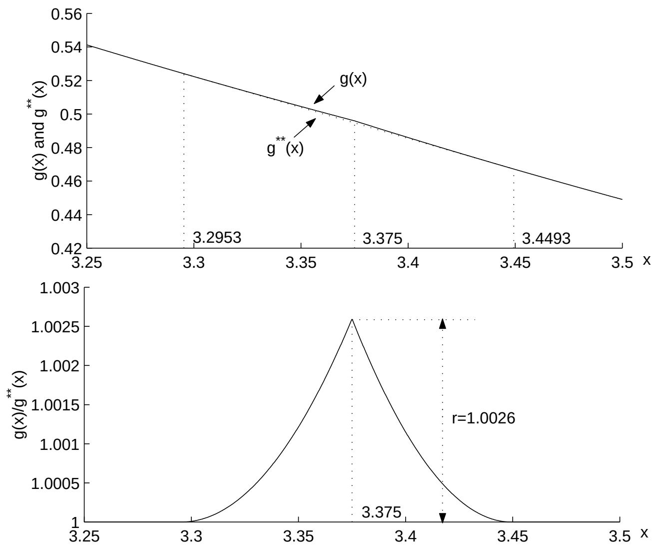

Fig. 2. The top figure shows $g(x) := 1/f(1/x)$ when $f()$ is PFTK-standard and its convex closure (dotted line). On the interval shown in the top figure, g^{**} is equal to the tangent common to both ends of the graph. Outside the interval it is equal to g. $g()$ is not strictly speaking convex, but almost. The bottom figure shows the ratio g/g^{**}, which is bounded by $r = 1.0026$.

Note that $\hat{\theta}_n$ is the moving-average estimator in (2), and thus

$$\mathbf{cov}[\theta_0, \hat{\theta}_0] = \sum_{l=1}^{L} w_l \mathbf{cov}[\theta_0, \theta_{-l}]. \tag{11}$$

In other words, it is a weighted sum of the autocovariance of the loss-event intervals at lags 1 to L.

The following corollary was shown in the discussion above.

Corollary 1: If the convexity condition (F1) holds and the loss-event intervals are stochastically independent, then the basic control (3) is conservative.

When convexity is almost true: The convexity condition (F1) is not true for PFTK-standard (because of the $\min$ term), but almost, as we see now. For a function $x \to g(x)$, we quantify its deviation from convexity by the ratio to its convex closure

$$r = \sup_x \left\{ \frac{g(x)}{g^{**}(x)} \right\}.$$

The convex closure $g^{**}(x)$ is the largest convex function that lower bounds $g(x)$; it is obtained by applying convex conjugation twice [14]. Figure 2 shows $g(x) = 1/f(1/x)$ for PFTK-standard and its convex closure; here, we have $r = 1.0026$.

Proposition 4: Assume that the loss-throughput formula f is such that $1/f(1/x)$ deviates from convexity by a ratio r, and that (C1) holds. Then the basic control (3) cannot overshoot by more than a factor equal to r.

Thus, considering that a fraction of a percent is more than reasonably accurate, we can conclude that for any practical purpose, we can act as if PFTK-standard would satisfy the convexity condition (F1).

2) When the sufficient conditions do not hold: We give a different set of conditions, that provide additional insights. This new set of conditions applies to some cases where Theorem 1 does not apply.

Theorem 2: Assume that

(F2) $x \to f(x)$ is concave,

(C2) $\mathbf{cov}[X_0, S_0] \leq 0$.

Then the basic control (3) is conservative.

Conversely, if

(F2c) $x \to f(x)$ is strictly convex,

(C2c) $\mathbf{cov}[X_0, S_0] \geq 0$,

(V) the loss-event estimator $\hat{\theta}_n$ has non-zero variance.

Then the basic control (3) is non-conservative.

Interpretation: The concavity condition (F2) is true for the SQRT formula. In contrast, the PFTK-standard and PFTK-simplified are such that concavity (F2) is true for rare losses, but convexity (F2c) is true for frequent losses; see Figure 1, left graph. The covariance condition (C2) is between X_n, the send rate at the occurrence of the nth loss-event, and S_n, a time until the next loss-event. If the loss process is memory-less and independent of the activity of our source, then the duration S_n of the loss interval is negatively correlated with the send rate X_n in the given interval (since S_n is counted in real time, not per packet); in such cases, condition (C2) is true, and the basic control is conservative, as long as losses are rare to moderate (or if the SQRT formula is used). This part of Theorem 2 complements Theorem 1.

Consider now the second part of Theorem 2. Assume that $\{S_n\}_n$, a sequence of loss-event intervals counted in real time, is independent of the send rate. This may happen for example for an audio source that modulates its send rate with a fixed packet send rate, but varying the packet lengths, and if the packet dropping probability in routers is independent of packet lengths; for instance, with RED operating in the packet mode. Then (C2c) holds, with equality. Now assume also that PFTK-standard is used, and the network setting happens to be such that the loss-event interval θ_n is mostly in the region where PFTK-standard is convex (that is, heavy losses). The theorem says that the basic control is non-conservative, except in a degenerate case where there is no randomness in the system, i.e. the loss-event interval estimator has converged to a fixed value. We show simulations that illustrate this case in Section V.

Another example is for a more traditional sender such as TFRC, but when the loss process goes through phases (for example, the network paths used by the connection oscillate between congestion and no congestion), and the send rate roughly follows the phases; that is to say, it is responsive at the timescale of the loss process. Then, when the network is in a congestion phase, X_n is most often small, and because of congestion, S_n is small. In such a case, condition (C2c) may be true and the basic control may *not* be conservative. In Section V-C we show such cases.

Viewpoint matters: The first part of Theorem 2 illustrates well the importance of the Feller paradox-type of arguments used in this paper; also known as "bus stop" paradox. The send rate $X(t)$ is updated only at loss-event instances. Consider an observer who picks a point in time at random. This observer is more likely to fall in a large inter loss-event interval S_n. Given that S_n is negatively correlated with X_n, it is thus more likely that on average our observer will see a smaller rate than another observer that would sample the send rate at loss-event instants. From this we conclude $\mathbf{E}[X(0)] \leq \mathbf{E}_N^0[X(0)]$. Now, the concavity assumption (F2), by Jensen's inequality, shows in turn that $\mathbf{E}_N^0[X(0)] \leq f(p)$. Finally, it follows $\mathbf{E}[X(0)] \leq f(p)$, the control is conservative.

Note that the correlation condition (C2) in Theorem 1 is implied by the condition that the conditional expected duration S_n, given the send rate X_n, decreases with X_n. That is

(C3) $\mathbf{E}[S_0|X_0 = x]$ is non-increasing with x.

This is a direct consequence of Harris' inequality[2] that (C3) implies the negative correlation condition (C2). A result based on the above conditional expectation found in [15] is a special case of the first part of Theorem 2.

Of course, we should expect that the combination of (C2c) and (V) implies that (C1) does not hold. This is shown to hold in [16], Appendix A.1 therein.

It is legitimate to wonder whether Theorem 1 is derived from Theorem 2 or vice versa. This does not seem to be the case; see [16], Appendix A.1. Note, however, if the concavity condition (F2) holds, then the convexity condition (F1) necessarily also holds. The converse is not true.

C. What this tells us

The analytical results in the previous section are for the basic control. We expect the comprehensive control to give a slightly higher throughput, since it differs by an additional increase during a long loss-event interval.

[2] Harris' inequality says that if $f(x)$ and $g(x)$ are non-decreasing functions, and X is one random variable, then the covariance of $f(X)$ and $g(X)$ is non-negative. See, for example [1], p. 225.

This motivates us to pose as assumptions the following analysis. We confirm our claims later by experiments.

Claim 1: Assume that the loss-event interval θ_n and the loss-event interval estimator $\hat{\theta}_n$ are slightly positive or negatively correlated. Consider the region where the loss-event interval estimator $\hat{\theta}_n$ takes its values.

- The more convex $1/f(1/x)$ is in this region, the more conservative the control is.
- The more variable $\hat{\theta}_n$ is, the more conservative the control is.

Application: For protocols like TFRC, we expect the condition to hold in many practical cases (see [18], also, our experimental evidence in Section V). For the three functions we consider in this paper, $x \to 1/f(1/x)$ is more convex for small x, that is, for a large loss-event rate p. Thus, the control should be more conservative with high loss than low loss. This effect is more pronounced for PFTK-standard (6) and PFTK-simplified (7), which are convex and very steep for large p, than for SQRT. This explains the observed drop in throughput for the control, with PFTK and heavy losses.

The "variability" of $\hat{\theta}_n$ depends on the variability of the loss-event intervals, and can be controlled by the value of the window of the moving-average estimator, L. With some appropriate setting of the weights $w_1, w_2, \ldots, w_L$, the larger the window of the estimator L, the smaller the variability of the estimator $\hat{\theta}_n$. We should find that for a larger L the control becomes less conservative.

The second claim concerns a case where the conditions in Claim 1 do not hold.

Claim 2: Assume that the duration in real time of the loss-event interval S_n and the send rate X_n are negatively or non-correlated.

- If $f(1/x)$ is concave in the region where the loss-event interval estimator $\hat{\theta}_n$ takes its values, the control tends to be conservative.

Conversely, assume S_n and X_n are positively or non-correlated.

- If $f(1/x)$ is strictly convex in the region where the loss-event interval estimator $\hat{\theta}_n$ takes its values, and $\{\theta_n\}_n$ is not fixed to a constant, the control is *non-conservative*.

In both cases, the more variable $\hat{\theta}_n$ is, the more pronounced the effect is.

Application: We expect to have a close to zero correlation for adaptive audio applications such as [3] when packet losses in RED routers are independent of packet length. Thus, depending on which convexity condition holds, we will find one or the other outcome. For SQRT, the control should always be conservative. The same holds for PFTK with light to moderate losses. The *opposite* holds for either PFTK formulae with heavy losses.

IV. Other Conditions for TCP-friendliness

A. Comparison of the Loss-Event Rates

We first consider how the loss-event rates seen by an equation-based rate control and TCP would compare. Unlike the problem of conservativeness, this problem does not allow us to conclude about the order of the loss-event rates, unless some further assumptions are made. Note that the relation of the loss-event rates seen by two different protocols would depend on the interaction of the protocols sharing a network. Yet, a claim can be made in a limit case.

1) Many-Sources Limit: Assume that senders in a network are driven by a congestion process $Z(t)$ that evolves in real time, $t \in \mathbb{R}$. This is an approximation that fits with the case of a sender with negligible influence on a global network. Assume the congestion process takes values on E, a countable state space. The state transitions are clocked by a point process $\ldots < T'_{-1} < T'_0 \leq 0 < T'_1 < \ldots$. We assume this point process to be stationary and has a finite non-null intensity λ'. Let N' be the associated counting process. Let $\pi_i := \mathbf{P}[Z(0) = i]$ be the steady-state probability that the congestion process is in the state $i \in E$. Define

$$p_i = \frac{1}{\mathbf{E}_N^0[\theta_0 | Z(0) = i]}.$$

This is the loss-event rate, given the congestion process is in the state $i \in E$. Let, also, $\bar{x}_i = \mathbf{E}[X(0)|Z(0) = i]$ be the time-average send rate, given the congestion process is in the state i. We show in the appendix

$$p_i = \frac{\sum_{i \in E} b_i p_i \bar{x}_i \pi_i}{\sum_{i \in E} b_i \bar{x}_i \pi_i}, \tag{12}$$

where

$$b_i = \frac{\mathbf{E}_{N'}^0[\sum_{n \in \mathbb{Z}} \theta_n \mathbf{1}_{[0, T'_1)}(T_n) | Z(0) = i]}{\mathbf{E}_{N'}^0[\int_0^{T'_1} X(s) ds | Z(0) = i]}.$$

In the limit case, as $\frac{\lambda'}{\lambda_i} \to 0$, for all $i \in E$, $b_i \to 1$. Here, by definition, $\lambda_i = 1/\mathbf{E}_N^0[S_0 | Z(0) = i]$ is the intensity of the loss-events in real-time, given the congestion process is in the state $i \in E$. The limit corresponds to a separation of timescales; we assume the congestion process evolves more slowly than the timescale of the control (remember that the control is clocked by the loss-events). We base our further discussion on the loss-event rate, in the foregoing limit case,

$$p \to \frac{\sum_{i \in E} p_i \bar{x}_i \pi_i}{\sum_{i \in E} \bar{x}_i \pi_i}. \tag{13}$$

If our source is non-adaptive (Poisson) then $\bar{x}_i = \bar{x}$ is independent of i. The resulting loss-event rate $p'' = \sum_{i \in E} \pi_i p_i$ can be thought of as the time-average of the "network" loss-event rate. Except for some possible aliasing effects, it should be close to what a constant bit rate (CBR) source would experience. Now if, like TCP, our source is very responsive, that is, follows the congestion process closely, then $\bar{x}_i$ depends on i in the following way: $\bar{x}_i$ is large for "good" states (p_i small) and small for bad states (p_i large). Thus, we should have a smaller p. For TCP, this is confirmed by measurements in [13]. The more responsive the source is, the more pronounced this should be. TCP is expected be more responsive than our adaptive sender, whose responsiveness depends on the averaging window L. We summarize this as follows; see Figure 7 for an illustration.

Claim 3: In the many-sources regime, the loss-event rates of TCP (p'), an equation based-rate control source (p), and a non-adaptive source (Poisson) (p'') should be in the relation

$$p' \leq p \leq p''.$$

The more responsive an equation-based rate controlled source is, the closer p should be to p'.

2) A Few Competing Senders: Consider a setting when a few TCP and equation-based rate control senders compete for a bottleneck.

Claim 4: In a situation when a few senders compete for a bottleneck, TCP may experience a larger loss-event rate than a competing equation-based rate control sender.

We support the claim by simple analysis that goes as follows; the claim is also verified by our experiments. Consider a link of a fixed capacity $c > 0$. Suppose the link is used by *one* sender that exercises a send rate control. Assume the round-trip time of the source is fixed to 1. A loss-event is experienced by the sender, whenever its send rate exceeds or it is equal to the link capacity. First, consider an additive-increase/multiplicative-decrease (AIMD) sender with additive-increase and multiplicative-decrease parameters $\alpha > 0$ and $0 < \beta < 1$, respectively. It is well-known that the loss-throughput function of our AIMD sender is equal to

$$f(p) = \sqrt{\frac{\alpha(1+\beta)}{2(1-\beta)}} \frac{1}{\sqrt{p}}.$$

The loss-event rate can be computed to be

$$p' = \frac{2\alpha}{(1-\beta^2)c^2}.$$

Now, consider the link is used by an equation-based rate control sender that adjusts its send rate by the comprehensive control with the function $f(\cdot)$ as defined above. Assume the send rate of the sender converges to the fixed-point r, then, we have

$$p = \frac{\alpha(1+\beta)}{2(1-\beta)c^2}.$$

The final outcome is

$$\frac{p'}{p} = \frac{4}{(1-\beta)^2}.$$

For a TCP-like setting ($\beta = 1/2$), the ratio of the loss-event rates is equal to $16/9$ (about 1.7778), which is a significant deviation. Our numerical simulations of one AIMD and one equation-based rate control competing for a link of a fixed capacity (not displayed due to space limitations) indicate that the deviation of the loss-event rates does hold, but it is somewhat less pronounced.

B. Obedience of TCP to its Throughput Formula

TCP throughput may not always conform well to a TCP throughput formula. The reason is at least twofold: TCP throughput formulae are derived under various simplifying assumptions, and a TCP implementation may not verify all the assumptions made in deriving a formula. In particular, we claim that when a TCP sender competes for a bottleneck with a few other senders, TCP may attain a smaller throughput than given by PFTK throughput formula. The claim is entirely based on our experimental work. Our conjecture is that when TCP compete for a bottleneck and the TCP window takes large values, the window increase over time is typically sub-linear. This is in contrast to the modeling assumption of the throughput formulae considered in this paper that the window increase is linear with time.

V. Experimental Results

A. Setup of the Experiments

1) Numerical Experiments: We designed numerical experiments with the objective to validate Claim 1 (the validation results given below). We take the loss-event intervals $\{\theta_n\}_n$ as an independent, identically distributed sequence of random variables. The probability density function of θ_0 is chosen to be $\mu(x) = a\exp(-a(x-x_0))$, $x \geq x_0 \geq 0$. (In other words, θ_0 is equal in distribution to a sum of the constant x_0, and a random variable with exponential distribution (a).) It can be checked that the expected value of θ_0 is $1/p = x_0 + 1/a$ and the coefficient of variation $\mathbf{cv}[\theta_0]^2 = (1/a)/(x_0 + 1/a)$. The chosen distribution has two degrees of freedom x_0 and a, which allows us to fix the coefficient of variation and vary p, and the other way around, fix p, and then

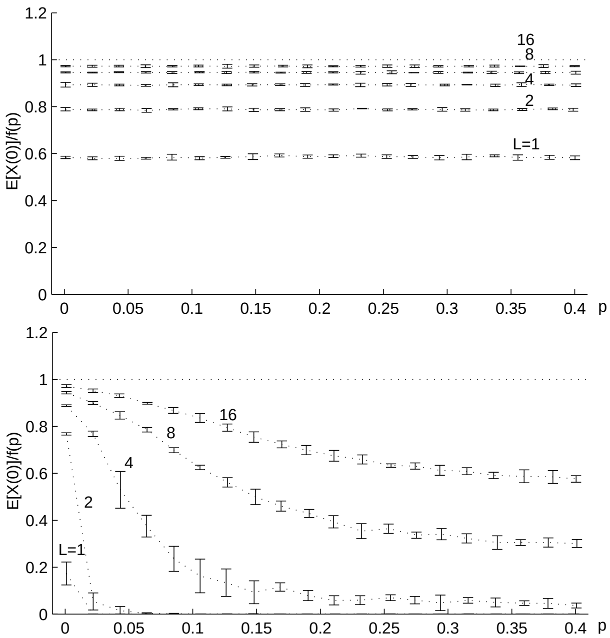

Fig. 3. Normalized throughput $\bar{x}/f(p)$ versus p, for the basic control with $\mathbf{cv}[\theta_0]$ fixed to $1-1/1000$. (Left) SQRT, (Right) PFTK-simplified with $q = 4r$. The estimator weights are as of TFRC, with length L. The same qualitative results remain to hold for the comprehensive control, but the effects are less pronounced; we omit to show this due to space limitations, see [16], Figure 4.

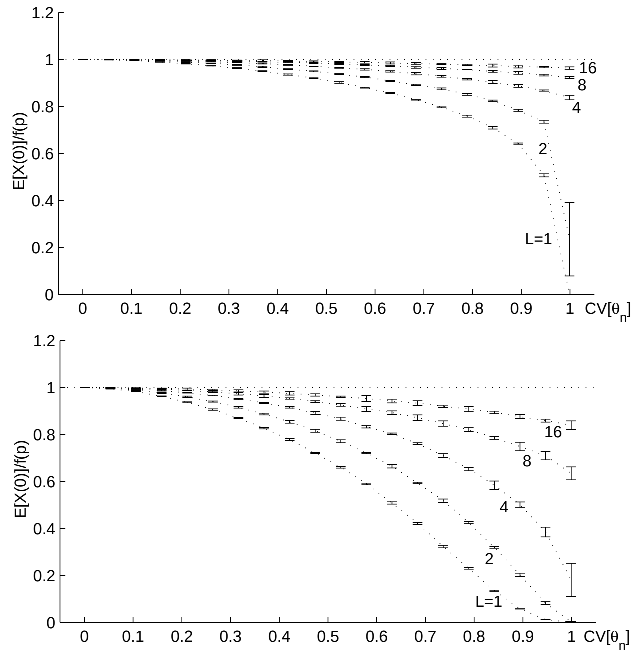

Fig. 4. Normalized throughput $\bar{x}/f(p)$ of the basic control versus the coefficient of variation of $\{\theta_n\}_n$, with p fixed to (Left) $1/100$, (Right) $1/10$. Function f is PFTK-simplified with $q = 4r$. The estimator weights set as of TFRC.

vary the coefficient of variation. A desirable feature is that some higher-order statistics remain intact to the value of (x_0, a), for instance, the skewness and kurtosis parameters are equal to 2 and 6, respectively.[3] We consider SQRT and PFTK-simplified functions f. Note that our numerical experiments are designed such that both hypotheses of Theorem 1 hold (also of Claim 1).

2) ns-2 Experiments: We designed packet-level simulations in ns-2, with TFRC as implemented in ns-2. The ns-2 implementation of TFRC is with the control law of the comprehensive control as defined in this paper. The setup of our ns-2 experiments is as follows. We consider a link shared by TFRC and TCP Sack1 connections. The link runs RED active queue management with a rate of 15 Mb/s, the buffer length, minimum buffer threshold, and maximum buffer threshold set to 5/2, 1/4, and 5/4 times the bandwidth-delay product, respectively. Other RED parameters were set to their ns-2 default values. The round-trip time is about 50 ms.

3) Lab Experiments: The design of the experiments is to some extent similar to the previously-described ns-2 experiments. We designed a network of two Linux PCs acting as routers through which the same number of TCP and TFRC connections is established. We ran one TCP and one TFRC test sender on two separate Linux PCs. Other senders were run on a third Linux PC. All connections terminated at a single receiver Linux PC. The first router on the forward path was connected to the second router via a 10 Mb/s hub, whereas the other connections were through 100 Mb/s Ethernet switches. Hence, the first router was configured to be a bottleneck. On the outgoing interface in the forward path of the first router we configured the queue discipline to be either DropTail or RED. We considered DropTail with the buffer length 64 and 100 packets. For RED, we fixed configuration parameters to match, respectively, the buffer length, minimum and maximum buffer thresholds of $5/2U$, $3/20U$, and $5/4U$, where $U = 62500$ B. The constant of the queue exponential smoothing was targeted to 0.002, and the drop probability at the maximum buffer threshold to $1/10$. The RED was configured not in the “gentle” mode, because this was not possible with the traffic control module of the Linux kernel. On the second router we ran NIST Net network emulator [11], solely for the purpose to adding a fixed propagation delay of 25 ms, in both forward and backward direction to all connections. We used TCP as implemented in Linux kernel 2.4.18.x, and an experimental user-space implementation of TFRC [9], which we adjusted to conform to the latest TFRC specification, and also to conform to most of the hypotheses of our analysis.

[3]Skewness and kurtosis parameters of a probability distribution quantify its skewness and sharpness.

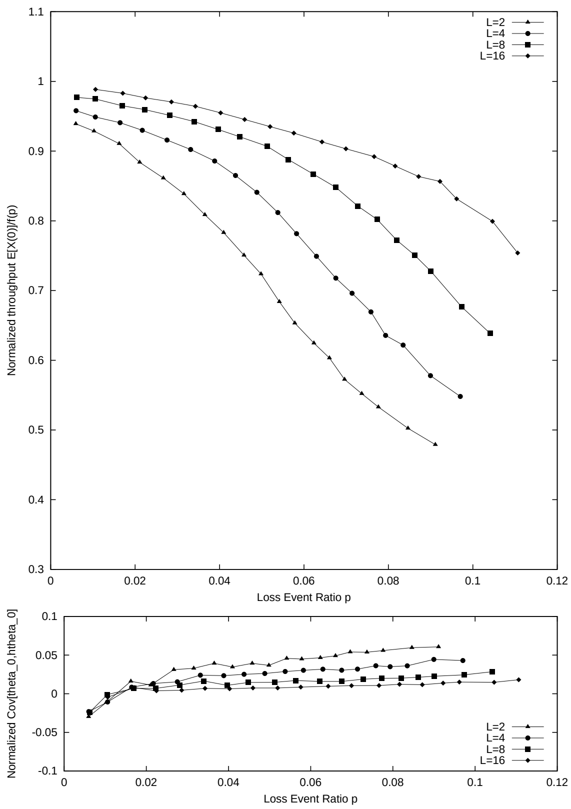

Fig. 5. The upper graph shows the normalized throughput $\bar{x}/f(p)$ attained by TFRC versus the loss-event rate p. The lower graph shows $\mathbf{cov}[\hat{\theta}_0, \theta_0]p^2$ versus p. Function f is PFTK-standard. For SQRT, and PFTK-simplified, the results are qualitatively similar, for the benefit of space we do not show them in this paper, see [16] for the corresponding plots.

In the lab experiments, we disabled the comprehensive control element in TFRC. The function f is PFTK-standard, $L = 8$. The measurement data was collected by direct probing in the TFRC implementation, for TCP, the quantities of interest were inferred from a tcpdump output. Lab experiments enabled us to run the system over a wide range of the loss-event rates, which would be impractical in the Internet. Given that our goal was to evaluate long-run behavior, we fixed the duration of an experiment in order to have a reasonable number of loss-events, in order to expect empirical estimates to converge close to the respective expected values. We fixed an experiment duration to 2500 s. We truncated the data collected for the initial 200 s, and computed some empirical estimates over a consecutive sequence of 6 bins over the remainder interval of an experiment. The expected inter loss-event time was roughly 2.5 s. We ran a sequence of 11 designed experiments with the same number of TCP and TFRC connections, equal to $(1, 2, 4, 6, 9, 12, 16, 20, 25, 30, 36)$.

TABLE I

SOME FACTS ABOUT OUR RECEIVER HOSTS AND CONNECTIONS TO THEM FROM EPFL.

Receiver	Mb/s	Hops	RTT	SACK	OS
INRIA	100	13	30	n	FreeBSD 4.1
UMASS	100	15	97	y	Linux 2.4.19
KTH	10	20	46	y	Linux 2.4.19
UMELB	10	24	350	y	Linux 2.2.14-tsc

The second column shows access rates of the receivers.
The RTT's are rounded estimates in ms by traceroute [8].

4) Internet Experiments: We designed our Internet experiments by setting up the senders in the same manner as with our lab experiments. All the senders were running Linux kernel 2.4.19.x, connected to a 100 Mb/s Ethernet at EPFL. The receivers were setup at either 100 Mb/s or 10 Mb/s Ethernet, all on the university campus networks. Some facts about the receivers are summarized in Table I. We also ran some Internet experiments with a receiver at EPFL connected by a 56 kb/s cable-modem. We made sure TCP window scaling was enabled and tuned TCP write and read socket buffer lengths to prevent TCP being limited by its receiver advertised window. Our Internet experiments were with the comprehensive control element of TFRC enabled. The function f is PFTK-standard, $L = 8$. We fixed the duration of our experiments as follows: (INRIA) 3600 s (12 bins), (KTH) 2 replicas of 1800 s (1 bin), (UMASS) 2 replicas of 1800 s (6 bins), and (UMELB) 3600 s (6 bins). The respective expected inter loss-event times were roughly 1 s, 40 s, 5 s, and 10 s. We ran 6 designed experiments with the same number of TCP and TFRC connections, equal to $(1, 2, 4, 6, 8, 10)$.

B. Validation of Claim 1

1) Numerical Experiments: First, we fix $\mathbf{cv}[\theta_0] = 1 - 1/1000$, and vary p. See Figure 3, which shows the normalized throughput $\bar{x}/f(p)$ versus p, for the basic control. Recall the first statement of Claim 1– the more convex the function $1/f(1/x)$ is in the region where the loss-event estimator takes its values, the more conservative the control is. Now, recall that for SQRT and PFTK-simplified (see Figure 1, right graph), $x \rightarrow 1/f(1/x)$ is convex and becomes steeper for smaller x (the smaller the x, the larger the loss-event rate p). Hence, the first statement of our Claim 1 tells us that the larger the p, the more conservative the control is. We verify this to be true for the basic control with PFTK-simplified, see Figure 3. The same result, but for the comprehensive control (not showed, see [16], Figure 4), is qualitatively the same, but the conservativeness is

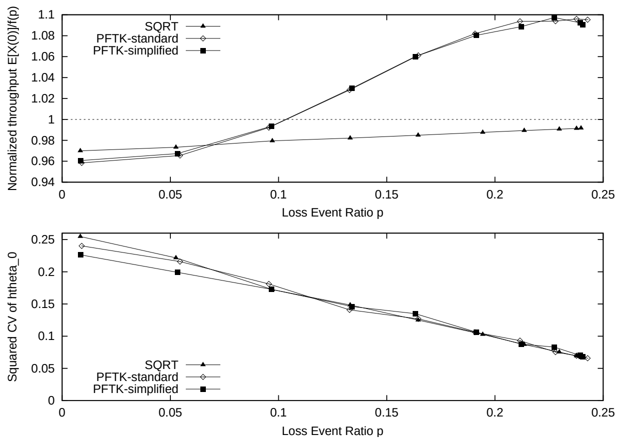

Fig. 6. (Top) The normalized throughput $\bar{x}/f(p)$ versus the loss-event rate of a sender with a constant packet send rate, but controlled packet lengths. The connection traverses a loss module, with a fixed packet drop probability–Bernoulli dropper. $L = 4$. (Bottom) the squared coefficient of variation of $\hat{\theta}_0$.

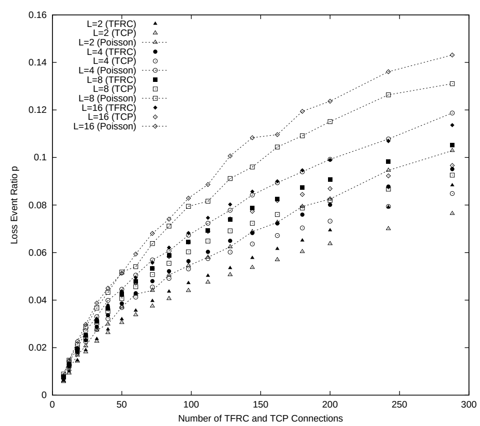

Fig. 7. Loss-event rates as experienced by TFRC, TCP, and Poisson connections versus N (number of TFRC and number TCP connections in one bottleneck). We have $p' \leq p \leq p''$ as expected. Also, the smoother the TFRC flows (larger L), the larger the loss-event rate.

somewhat less pronounced. For SQRT, in Figure 3, we observe that the normalized throughput is invariant to the value of p. It can be shown that for SQRT function f, if the distribution of $p\theta_0$ does not depend on p, then the normalized throughput does not depend on p. The last property holds for the probability density function of θ_0 taken in our example, which we omit to show due to the space limitations.

Second, we fix p, and vary $\mathbf{cv}[\theta_0]$. We show the normalized throughput of the basic control versus $\mathbf{cv}[\theta_0]$ in Figure 4, for (Top graph) $p = 1/100$, and (Bottom graph) $p = 1/10$. The results validate the second statement of Claim 1; the larger the variability of the loss-event interval estimator, the more conservative the control

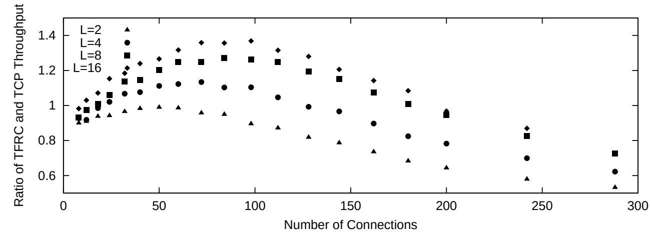

Fig. 8. The ratio of throughputs attained by TFRC and TCP Sack1 versus the number of connections.

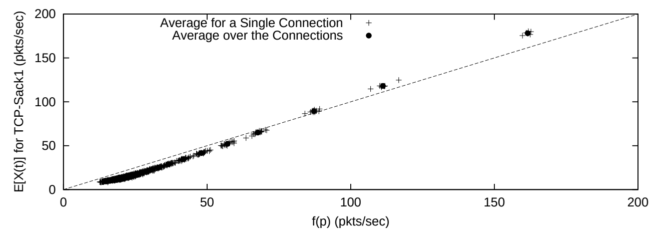

Fig. 9. TCP Sack 1 versus PFTK-standard formula. Throughput is below the throughput predicted by the formula, except for large throughputs.

is. The second statement of Claim 1 is, in fact, also, confirmed by the results in Figure 3, where we notice, the smaller the value of the L, the more conservative the control is. (Smaller value of L means smaller smoothing of the loss-event interval estimator.)

2) ns-2 Experiments: We now give more realistic experiments to validate our Claim 1. We show the estimated normalized throughput in Figure 5. The results are qualitatively the same as with our numerical experiments discussed earlier, and thus, they confirm our Claim 1.

3) Lab Experiments: See Figure 18 and Figure 19, the leftmost graph. The empirical results indicate, the larger the loss-event rate, the stronger the conservativeness. Note that the covariance condition in Claim 1 seems to hold, as indicated in Figure 10, the leftmost graph, for lab experiments. As an aside, note that the observed estimates compare well to the results obtained by the ns-2 simulation (Figure 5). Recall that, for TFRC in our lab experiments, $L = 8$.

4) Internet Experiments: For the Internet experiments, see Figure 12-Figure 15, the conservativeness or non-conservativeness is mostly negligible. Note that this is to be expected, given that the loss-event rates are small.

C. Validation of Claim 2

1) ns-2 experiments: We consider a sender that sends packets at regular time intervals, periodically each 20 ms. The sender runs equation-based rate control by varying

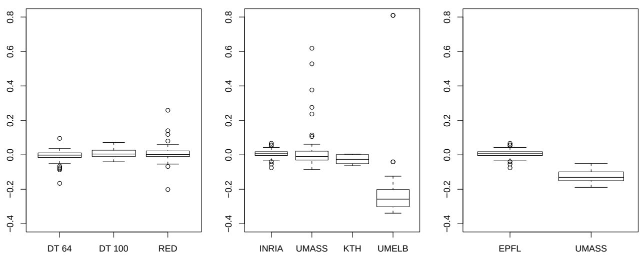

Fig. 10. $\mathbf{cov}[\theta_0, \hat{\theta}_0]p^2$ for (Left) Lab experiments (Middle) Internet experiments (Right) Internet experiments with the receiver connected via a cable-modem at EPFL. Observe that the normalized covariance is mostly near to zero. It is noticeably negative for UMELB, and UMASS sending to the cable-modem. With UMELB, we observed the loss-events occurring in batches, which results in a negative normalized covariance.

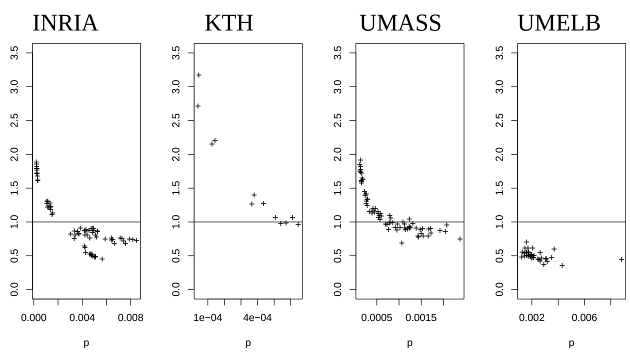

Fig. 11. Internet experiments: check is TFRC TCP-friendly. The graphs show the ratio of $\bar{x}$ and $\bar{x}'$, respectively, the throughputs of TFRC and TCP, versus p. The values not larger than one indicate TCP-friendliness, else, non-TCP-friendliness.

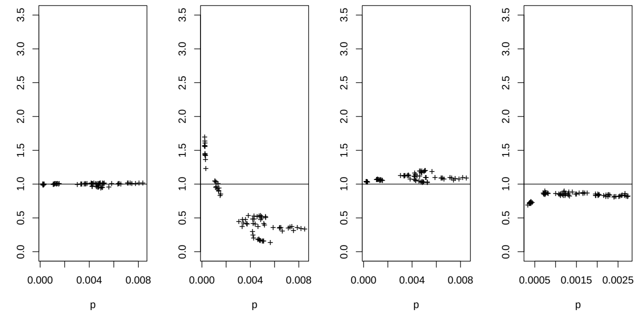

Fig. 12. INRIA. Breakdown of the TCP-friendliness condition into: (Left-to-Right) the ratio of $\bar{x}$ and $f(p, r)$; the ratio of p' and p; the ratio of r' and r; the ratio of $\bar{x}'$ and $f(p', r')$.

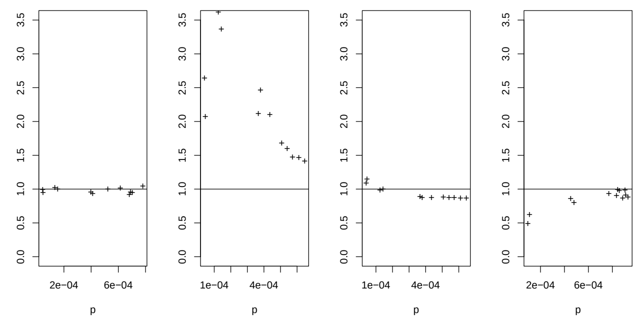

Fig. 13. KTH. Same as in Figure 12.

packet lengths. The sender has a connection established through a loss module that drops a packet with a fixed probability p (Bernoulli dropper). For such a sender, we have that the covariance of the send rate at a loss-event instant, and the time it takes until the next loss-event is equal to zero. Hence, by Claim 2, we expect our sender to be conservative for $x \to f(x)$ concave, and conversely, non-conservative for $x \to f(x)$ convex. We confirm the claim to be true, see Figure 6. The results in Figure 6 are for $L = 4$; for $L = 8$ the results are qualitatively the same, but the effects are less pronounced because of the larger smoothing of the loss-event interval estimator (a plot was shown in [16]).

D. Validation of Claim 3

1) ns-2 experiments: We come back to our ns-2 experiments in Section V-B. We evaluated the loss-event rates as seen by TFRC, TCP, and Poisson senders in the experiments, see Figure 7. The experiments validate Claim 3.

E. Validation of Claim 4

1) ns-2 experiments: See in Figure 17 that TCP experiences a larger loss-event rate than TFRC.

2) Internet experiments: See Figure 12 and Figure 14, the second graph in the row.

F. Breakdown of the TCP-friendliness Condition

1) ns-2 experiments: We review the ns-2 experiments in Section V-B to check whether TFRC is TCP-friendly. From Figure 8, we reveal that the answer is negative. Note that in some of the experiments, TFRC is non-TCP-friendly despite the fact that we observed it is conservative (Figure 5) and it sees larger loss-event rate than TCP (Figure 7). From Figure 9, we observe that the TCP in the experiments does not conform well to PFTK-standard formula. In some cases, TCP a attains smaller throughput than given by the formula.

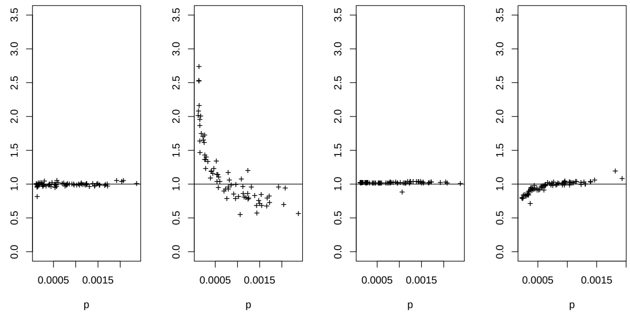

Fig. 14. UMASS. Same as in Figure 12.

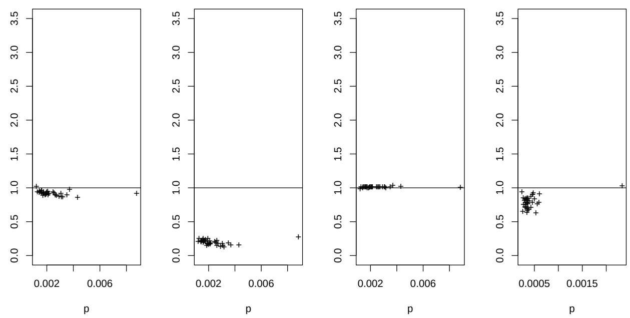

Fig. 15. UMELB. Same as in Figure 12.

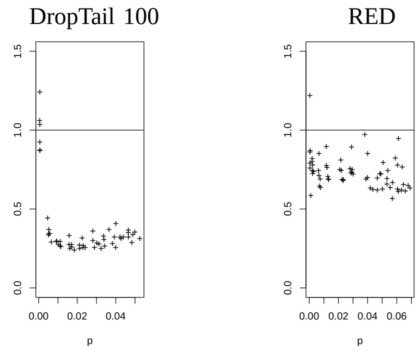

Fig. 16. Lab experiments: check is TFRC TCP-friendly. The graphs show the ratio of $\bar{x}$ and $\bar{x}'$, respectively, the throughputs of TFRC and TCP, versus p.

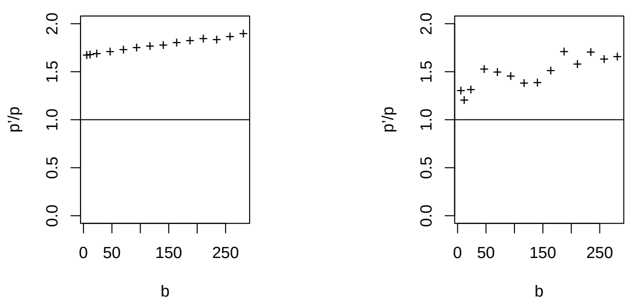

Fig. 17. The ratio of the loss-event rates as observed by TCP and TFRC over a DropTail bottleneck with a buffer length of b packets: (Left) Either one TCP or one TFRC in isolation over the bottleneck, and (Right) One TCP and one TFRC over the bottleneck. Evidently, TFRC experiences smaller loss-event rate than TCP.

2) Internet Experiments: Check whether TFRC is TCP-friendly in Figure 11. For INRIA, KTH, and UMASS, the answer is negative. More specifically, for small loss-event rates (which corresponds to a few competing TCP and TFRC senders), TFRC can be significantly non-TCP-friendly. We find the causes of the observed non-TCP-friendliness in the breakdown of the TCP-friendliness condition, see Figure 12, 13, 14. One cause is that the loss-event rate of TCP, p', can be significantly larger than the loss-event rate of TFRC, p. See the second plot from the left in Figure 12, 13, 14. Another cause is that, in the regime of a few senders competing, TCP attains smaller throughput than given by the formula, see the rightmost plot in Figure 12, 13, 14.

In summary, upon our Internet and lab experiments, we observe:

- In our experiments, the loss-event rate is a dominant factor that determine TCP-friendliness, or non-TCP-friendliness.

VI. CONCLUSION

Our study should help designers of equation-based rate controls better understand the trade-offs that have to be made. First, it is important to separately verify the four factors: (1) conservativeness, (2) TCP loss-event rate versus this protocol's loss-event rate, (3) TCP average round-trip time versus this protocol's average round-trip time, (4) TCP's obedience to its throughput formula. Failing to do so blurs the cause of an observed excessive TCP-friendliness or non-TCP-friendliness, and may lead a protocol designer to an improper protocol adjustment. Second, we should be aware of the strong dependency on the nature of the function f; SQRT behaves differently than PFTK. If PFTK is used, under the conditions of Claim 1, a very pronounced conservativeness should be expected for heavy loss. Under some very specific conditions (rate modulated by variable packet length), the opposite may hold (Claim 2). It still remains to know whether there exists in practice a loss process statistics that would drive the control to a significant non-conservativeness.

Our analysis and experiments demonstrate that TCP-

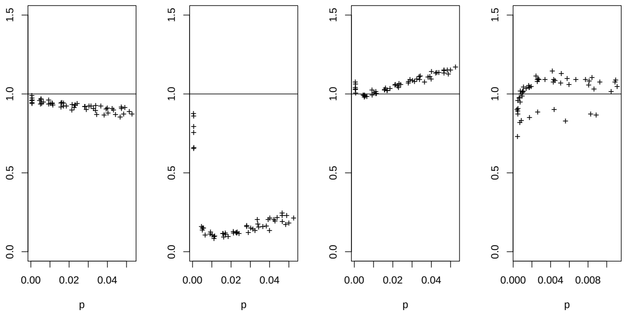

Fig. 18. DropTail 100. Same as in Figure 12.

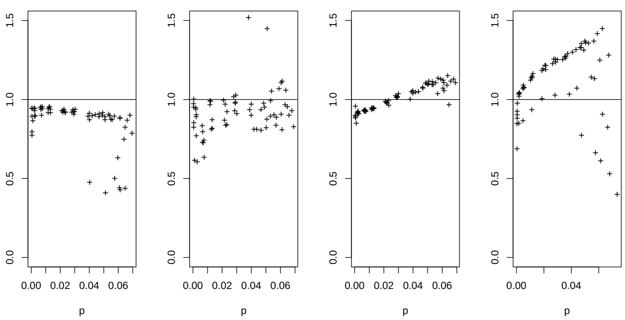

Fig. 19. RED. Same as in Figure 12.

friendliness is difficulty to verify, in contrast to conservativeness, which is easier. Conservativeness as a design objective is less restrictive and would allow for the design of more effective controls. It would not constrain the controls to be TCP-friendly in the cases where TCP performs poorly (few connections in a single bottleneck), while guaranteeing a safe behaviour and fair sharing in the presence of congestion. Exploring the full implications of such a design alternative is for further study.

ACKNOWLEDGMENTS

Chadi Barakat (INRIA), Gunnar Karlsson, Ignacio M. Ivars (KTH), Daniel R. Figueiredo (UMASS), Darryl Veitch (UMELB), generously provided us accounts at hosts at their sites, and offered us support. We thank them all. We thank Chadi Barakat for allowing us to use his code to analyze a tcpdump output and Jörg Widmer for discussions on the TFRC experimental implementation. Marc-Andre Luthi and Božidar Radunović helped us with setting up our lab and Internet measurements at EPFL. We acknowledge the assistance of Christian Latesch and Thorsten Müller.

REFERENCES

[1] Francois Baccelli and Pierre Brémaud. *Elements of Queueing Theory*, volume 26. Applications of Mathematics, Springer-Verlag, 1991.
[2] Deepak Bansal, Hari Balakrishnan, Sally Floyd, and Scott Shenker. Dynamic behavior of slowly-responsive congestion control algorithms. In *Proc. of ACM Sigcomm'01*, San Diego, California, USA, August 2001.
[3] C. Boutremans and J.-Y. Le Boudec. Adaptive delay aware error control for internet telephony. In *Proc. of 2nd IP-Telephony Workshop*, pages 81–92, Columbia University, New York, April 2001.
[4] Neal Cardwell, Stefan Savage, and Thomas Anderson. Modeling TCP latency. In *Proc. of the IEEE INFOCOM'2000*, Tel-Aviv, Israel, March 2000.
[5] Sally Floyd and Kevin Fall. Promoting the use of end-to-end congestion control in the Internet. *IEEE/ACM Trans. on Networking*, 7(4):458–472, August 1999.
[6] Sally Floyd, Mark Handley, Jitendra Padhye, and Jörg Widmer. Equation-based congestion control for unicast applications. In *Proc. of the Sigcomm'00*, pages 43–56, 2000.
[7] Mark Handley, Jitendra Padhye, Sally Floyd, and Jörg Widmer. TCP Friendly Rate Control (TFRC) Protocol Specification, IETF internet-draft, Janunary 2003. RFC 3448, ftp://ftp.isi.edu/in-notes/rfc3448.txt.
[8] V. Jacobson. ftp://ftp.ee.lbl.gov/traceroute.tar.z, 1989.
[9] Page maintaned by Jörg Widmer. Implementation of the TCP-Friendly Congestion Control Protocol (TFRC), February 2000. http://www.icir.org/tfrc.
[10] Matthew Mathis, Jeffrey Semke, Jamshid Mahdavi, and Teunis Ott. The Macroscopic Behavior of the TCP Congestion Avoidance Algorithm. *Computer Communication Review*, 27(3), July 1997.
[11] Network Emulation Package NIST Net, 2003. http://is2.antd.nist.gov/itg/nistnet.
[12] Jitendra Padhye, Victor Firoiu, Don Towsley, and Jim Kurose. Modeling TCP Reno Performance: A Simple Model and its Empirical Validation. *IEEE/ACM Trans. on Networking*, 8(2):133–145, 2000.
[13] Vern Paxson. End-to-End Internet Packet Dynamics. *IEEE/ACM Trans. on Networking*, pages 277–292, June 1999.
[14] Rockafellar R. T. *Convex Analysis*. Princeton University Press, Princeton, 1970.
[15] Milan Vojnović and Jean-Yves Le Boudec. Some observations on equation-based rate control. In *Proc. of ITC-17*, pages 173–184, Salvador, Bahia, Brazil, December 2001.
[16] Milan Vojnović and Jean-Yves Le Boudec. On the long-run behavior of equation-based rate control. In *Proc. of ACM Sigcomm 2002*, Pittsburgh, PA, August 2002.
[17] Y. Richard Yang, Min Sik Kim, and Simon S. Lam. Transient Behaviors of TCP-friendly Congestion Control Protocols. In *Proc. of IEEE Infocom'2001*, March 2001.
[18] Yin Zhang, Nick Duffield, Vern Paxson, and Scott Shenker. On the Constancy of Internet Path Properties. In *Proc. of ACM Sigcomm Internet Measurement Workshop*, November 2001.

APPENDIX

A. Proof of Proposition 1

We commence with the Palm inversion formula [1],

$$\mathbf{E}[X(0)] = \lambda \mathbf{E}_N^0[\int_0^{T_1} X(s)ds]. \tag{14}$$

Note that (14), is a mean-value formula. It is a "cycle" formula, we can interpret it as the expected number of data sent between two successive loss-events divided with the expected time between two successive loss-events. However, it is important to remember the expected values are with respect to the Palm probability, that is as seen at the instants of loss-events.

For the basic control this gives

$$\mathbf{E}[X(0)] = \frac{\mathbf{E}[X_0 S_0]}{\mathbf{E}[S_0]}. \tag{15}$$

From (3), $\theta_n = X_n S_n$ and $X_n = f(1/\hat{\theta}_n)$. Hence, $S_n = \frac{\theta_n}{f(1/\hat{\theta}_n)}$. Combining the last three identities into (15) we obtain (8).

B. Proof of Proposition 2

Define $Y_n := \theta_n / S_n$, all $n \in \mathbb{Z}$. A physical meaning of Y_n is the average send rate over the interval $[T_n, T_{n+1})$, for some fixed $n \in \mathbb{Z}$. Again from Palm inversion formula

$$\mathbf{E}[X(0)] = \frac{\mathbf{E}[\theta_0]}{\mathbf{E}[\frac{\theta_0}{Y_0}]}$$

Now, by definition of the comprehensive control, $X(t) \geq X_n$, for all $t \in [T_n, T_{n+1})$, and hence, $Y_n \geq X_n$. Replacing Y_0 in the above display with its lower bound X_0, and recalling that by definition $X_n = f(1/\hat{\theta}_n)$, $n \in \mathbb{Z}$, recovers the asserted lower bound.

C. Proof of Proposition 3

Case 1: ($\hat{\theta}_{n+1} \leq \hat{\theta}_n$). In this case $\theta_n = X_n S_n$, and hence, $S_n = \frac{\theta_n}{f(1/\hat{\theta}_n)}$.

Case 2: ($\hat{\theta}_{n+1} > \hat{\theta}_n$) In this case, for $T_n \leq t \leq T_n + U_n$, $\theta(t) = t f(1/\hat{\theta}_n)$. Else, for $T_n + U_n < t < T_n + S_n$,

$$\theta(t) = \theta(T_n + U_n) + \int_{T_n+U_n}^{t} X(s) ds.$$

By definition of the comprehensive control ((4)), we obtain the ordinary differential equation, $T_n + U_n \leq t < T_n + S_n$, $\theta(T_n + U_n) = (\hat{\theta}_n - W_n)/w_1$,

$$\frac{d\theta(t)}{dt} = f\left(\frac{1}{w_1\theta(t) + W_n}\right), \tag{16}$$

where $W_n = \sum_{l=1}^{L-1} w_{l+1}\theta_{n-l}$.

Now, we solve $\theta(T_n + S_n -) = \theta_n$ for S_n. To that end, we solve (16) for PFTK-simplified formula (7). Plugging PFTK-simplified function f into (16), and a simple rearrangement, we obtain

$$\begin{aligned} c_1 r \int_{T_n+U_n}^{T_n+S_n} \frac{d\theta(t)}{\sqrt{w_1\theta(t) + W_n}} &+ \\ +c_2 q \int_{T_n+U_n}^{T_n+S_n} \frac{d\theta(t)}{\sqrt{(w_1\theta(t) + W_n)^3}} &+ \\ +32 c_2 q \int_{T_n+U_n}^{T_n+S_n} \frac{d\theta(t)}{\sqrt{(w_1\theta(t) + W_n)^7}} &= S_n - U_n \end{aligned}$$

Use the substitution $y = w_1\theta(t) + W_n$. Note that $d\theta(t) = dy/w_1$ and that with this substitution the boundaries of the integrals, $T_n + U_n$ and $T_n + S_n$, respectively, are equal to $\hat{\theta}_n$ and $\hat{\theta}_{n+1}$. We re-write the last display as

$$\begin{aligned} S_n &= U_n + \frac{c_1 r}{w_1} \int_{\hat{\theta}_n}^{\hat{\theta}_{n+1}} \frac{dy}{\sqrt{y}} + \frac{c_2 q}{w_1} \int_{\hat{\theta}_n}^{\hat{\theta}_{n+1}} \frac{dy}{\sqrt{y^3}} + \\ &\quad + \frac{32 c_2 q}{w_1} \int_{\hat{\theta}_n}^{\hat{\theta}_{n+1}} \frac{dy}{\sqrt{y^7}} \end{aligned}$$

Solving the elementary integrals, we obtain

$$\begin{aligned} S_n &= U_n + \frac{2c_1 r}{w_1}(\hat{\theta}_{n+1}^{1/2} - \hat{\theta}_n^{1/2}) - 2\frac{c_2 q}{w_1}(\hat{\theta}_{n+1}^{-1/2} - \\ &\quad - \hat{\theta}_n^{-1/2}) - \\ &\quad - \frac{64}{5}\frac{c_2 q}{w_1}(\hat{\theta}_{n+1}^{-5/2} - \hat{\theta}_n^{-5/2}) \end{aligned}$$

For convenience of notation, let $B_n := S_n - U_n$. Recall,

$$U_n = \frac{\hat{\theta}_n - W_n}{w_1 f(\frac{1}{\hat{\theta}_n})}$$

Finally, we have

$$\begin{aligned} S_n &= \frac{\theta_n}{f(\frac{1}{\hat{\theta}_n})}\mathbf{1}_{\hat{\theta}_{n+1} \leq \hat{\theta}_n} + \left(B_n + \frac{\hat{\theta}_n - W_n}{w_1 f(\frac{1}{\hat{\theta}_n})}\right)\mathbf{1}_{\hat{\theta}_{n+1} > \hat{\theta}_n} \\ &= \frac{\theta_n}{f(\frac{1}{\hat{\theta}_n})} + \left(B_n - \frac{\hat{\theta}_{n+1} - \hat{\theta}_n}{w_1 f(\frac{1}{\hat{\theta}_n})}\right)\mathbf{1}_{\hat{\theta}_{n+1} > \hat{\theta}_n} \\ &= \frac{\theta_n}{f(\frac{1}{\hat{\theta}_n})} - V_n \mathbf{1}_{\hat{\theta}_{n+1} > \hat{\theta}_n} \end{aligned}$$

The last identity follows directly by definitions of V_n and B_n. It remains only to use Palm inversion formula $\mathbf{E}[X(0)] = \mathbf{E}[\theta_0]/\mathbf{E}[S_0]$, and plug-in the expression of S_n displayed above to show (9).

D. Proof of Theorem 1

Define $g(x) := \frac{1}{f(\frac{1}{x})}$. Also call $m = \frac{1}{p}$, thus $\mathbf{E}[\theta_0] = \mathbf{E}[\hat{\theta}_0] = m$. From Equation (8), conservativeness is equivalent to

$$\mathbf{E}[\theta_0 g(\hat{\theta}_0)] \geq m g(m) \tag{17}$$

Function g is convex, thus is above its tangents:

$$g(x) \geq (x-m)g'(m) + g(m).$$

Apply the above to $x = \hat{\theta}_0$, multiply by θ_0 and take the expectation. After some calculus, this shows Equation (10).

Now f is decreasing. Since $\mathbf{cov}^0_N[\theta_0, \hat{\theta}_0] \leq 0$, it follows from Equation (10) that the control is conservative.

E. Proof of Proposition 4

Use the same notation as in the proof of Theorem 1. By Equation (8) the ratio of throughput to $f(p)$ is equal to

$$\rho := \frac{mg(m)}{\mathbf{E}\left[\theta_0 g(\hat{\theta}_0)\right]}. \tag{18}$$

Now we have

$$g^{**}(x) \leq g(x) \leq r g^{**}(x).$$

and thus $\rho \leq r$.

F. Proof of Theorem 2

Use the same notation as in the proof of Theorem 1.

Part 1: By (C2)

$$\mathbf{E}[\theta_0 g(\hat{\theta}_0)] \geq \frac{m}{\mathbf{E}\left[\frac{1}{g(\hat{\theta}_0)}\right]}, \tag{19}$$

now (F2) means that $\frac{1}{g}$ is concave, thus by Jensen's inequality:

$$\mathbf{E}\left[\frac{1}{g(\hat{\theta}_0)}\right] \leq \frac{1}{g(\mathbf{E}[\hat{\theta}_0])}, \tag{20}$$

which combined with the previous equation shows that the control is conservative.

Part 2: By (C2c) and (F2c) we have the reverse inequalities in Equation (19) and Equation (20), but the inequality is strict in Equation (20) because convexity is strict and $\hat{\theta}_n$ is not a degenerate random variable.

G. Derivation of Equation (12)

We begin from Equation (1). Let $\pi^0_i := \mathbf{P}^0_{N'}[Z(0) = i]$, $i \in E$. By Neveu's exchange formula ([1], Sec. 3.3.4) and simple conditioning

$$\begin{aligned} p &= \frac{1}{\mathbf{E}[\theta_0]} \\ &= \frac{\mathbf{E}^0_{N'}[\sum_{n\in\mathbb{Z}} \mathbf{1}_{[0,T'_1)}(T_n)]}{\mathbf{E}^0_{N'}[\sum_{n\in\mathbb{Z}} \theta_n \mathbf{1}_{[0,T'_1)}(T_n)]} \\ &= \frac{\sum_{i\in E} \mathbf{E}^0_{N'}[\sum_{n\in\mathbb{Z}} \mathbf{1}_{[0,T'_1)}(T_n)|Z(0)=i]}{\sum_{i\in E} \mathbf{E}^0_{N'}[\sum_{n\in\mathbb{Z}} \theta_n \mathbf{1}_{[0,T'_1)}(T_n)|Z(0)=i]} \end{aligned} \tag{21}$$

We show that the above is equivalent to Equation (12).

As an application of Palm inversion formula to $X(0)\mathbf{1}_{Z(0)=i}$, we obtain

$$\bar{x}_i = \mathbf{E}[X(0)|Z(0)=i] = \frac{\mathbf{E}^0_{N'}[\int_0^{T'_1} X(s)ds|Z(0)=i]}{\mathbf{E}^0_{N'}[T'_1|Z(0)=i]},$$

where we also use (obtained by another application of the Palm inversion formula to $\mathbf{1}_{Z(0)=i}$)

$$\pi_i = \mathbf{P}[Z(0)=i] = \frac{\mathbf{E}^0_{N'}[T'_1|Z(0)=i]}{\mathbf{E}^0_{N'}[T'_1]}\pi^0_i.$$

By a similar argument, from Neveu's exchange formula applied to $\theta_0 \mathbf{1}_{Z(0)=i}$, we have

$$\frac{1}{p_i} = \mathbf{E}^0_N[\theta_0|Z(0)=i] = \frac{\mathbf{E}^0_{N'}[\sum_{n\in\mathbb{Z}} \theta_n \mathbf{1}_{[0,T'_1)}(T_n)|Z(0)=i]}{\mathbf{E}^0_{N'}[\sum_{n\in\mathbb{Z}} \mathbf{1}_{[0,T'_1)}(T_n)|Z(0)=i]},$$

where we use the identity obtained by Neveu's exchange formula applied to $\mathbf{1}_{Z(0)=i}$,

$$\mathbf{P}^0_N[Z(0)=i] = \frac{\mathbf{E}^0_{N'}[\sum_{n\in\mathbb{Z}} \mathbf{1}_{[0,T'_1)}(T_n)|Z(0)=i]}{\mathbf{E}^0_{N'}[\sum_{n\in\mathbb{Z}} \mathbf{1}_{[0,T'_1)}(T_n)]}\pi^0_i.$$

Finally, by plugging the above expressions for $\bar{x}_i$, π_i, and p_i into Equation (12) we recover Equation (21).