Adv. Appl. Prob. **38,** 1132–1154 (2006)
Printed in Northern Ireland

RATE OF CONVERGENCE FOR THE 'SQUARE ROOT FORMULA' IN THE INTERNET TRANSMISSION CONTROL PROTOCOL

TEUNIS J. OTT *

Abstract

The 'square root formula' in the Internet transmission control protocol (TCP) states that if the probability p of packet loss becomes small and there is independence between packets, then the stationary distribution of the congestion window W is such that the distribution of $W\sqrt{p}$ is almost independent of p and is completely characterizable. This paper gives an elementary proof of the convergence of the stationary distributions for a much wider class of processes that includes classical TCP as well as T. Kelly's 'scalable TCP'. This paper also gives stochastic dominance results that translate to a rate of convergence.

Keywords: Computer networking; transport protocol; stochastic process; square root law; stationary distribution; stochastic dominance; rate of convergence; coupled stochastic processes

2000 Mathematics Subject Classification: Primary 60F99; 60G10; 60G99; 60K30; 60K99; 68U99

1. Introduction

The paper [11] proposed a class of Internet transport protocols similar to the Internet transmission control protocol (TCP) and used a class of stochastic processes to analyse the performance of these protocols. This class of stochastic processes is defined as follows. Let $(U_n)_{n=0}^{\infty}$ be independent, identically distributed random variables each distributed uniformly in $[0, 1]$. Let p, $0 < p < 1$, be a probability. Define the independent, identically distributed random variables $\chi_{p,n}$ by

$$\chi_{p,n} = \begin{cases} \text{success} & \text{if } U_n \geq p, \\ \text{failure} & \text{if } U_n < p. \end{cases} \tag{1}$$

Furthermore, let the discrete-time, continuous-state-space process $W^*_{p,C,n}$, with $n = 0, 1, 2, \dots, 0 < W^*_{p,C,n} < \infty$, and $0 < p < 1$, be defined by

$$W^*_{p,C,n+1} = \begin{cases} W^*_{p,C,n} + c_1 (W^*_{p,C,n})^{\alpha} & \text{if } \chi_{p,n} = \text{success}, \\ \max(W^*_{p,C,n} - c_2 (W^*_{p,C,n})^{\beta}, C) & \text{if } \chi_{p,n} = \text{failure}, \end{cases} \tag{2}$$

where $\alpha < \beta \leq 1$, $c_1 > 0$, $c_2 > 0$, and $C > 0$. The special case with $\beta = 1$, $\alpha = -1$, $c_1 = 1$, $c_2 = \frac{1}{2}$, and (for example) $C = 1$ models 'classical TCP'. The special case with $\beta = 1$ and $\alpha = 0$ models Kelly's 'scalable TCP'; see [7] and [8]. The paper [11] showed that the more general case, even that in which $0 < \alpha < \beta \leq 1$, is of interest in the study of transport protocols.

Received 7 September 2004; revision received 3 August 2006.
* Postal address: 31 Mountainview Drive, Chester, NJ 07930, USA. Email address: teun@teunisott.com

In [13] it was proven that, for all values $\alpha < \beta \leq 1$, $c_1 > 0$, $c_2 > 0$, $C > 0$, $0 < p < 1$ (and $0 < c_2 < 1$ if $\beta = 1$), the process $W^*_{p,C,n}$ has a unique stationary distribution. The uniqueness of this stationary distribution is derived from the fact that eventually $W^*_{p,C,n} = C$ for some (possibly large) n.

In this paper we will study the case in which $\alpha < \beta = 1$, $c_1 > 0$, and $0 < c_2 < 1$, and we will write $1 - c_2 = b$. In this case we will see that we can drop the 'max$(\cdot, C)$' in (2) (or choose $C = 0$). We will be mostly interested in the case $C = 0$, but after this case has been studied we will observe consequences for the process with $C > 0$.

A process of further interest in this paper is therefore defined by

$$W_{p,n+1} = \begin{cases} W_{p,n} + c_1(W_{p,n})^{\alpha} & \text{if } \chi_{p,n} = \text{success}, \\ W_{p,n} - (1-b)W_{p,n} = bW_{p,n} & \text{if } \chi_{p,n} = \text{failure}, \end{cases} \tag{3}$$

but we will also draw some conclusions for the process $(W^*_{p,C,n})_{n=0}^{\infty}$ defined by

$$W^*_{p,n+1} = \begin{cases} W^*_{p,n} + c_1(W^*_{p,n})^{\alpha} & \text{if } \chi_{p,n} = \text{success}, \\ \max(bW^*_{p,C,n}, C) & \text{if } \chi_{p,n} = \text{failure}. \end{cases}$$

Not surprisingly, as $p \downarrow 0$ the two processes $W_{p,n}$ and $W^*_{p,n}$ become very similar, in a way that will be explained in Section 13.

We always choose $0 < W_{p,0} < \infty$, and therefore have

$$0 < b^n W_{p,0} \leq W_{p,n} < \infty \quad \text{for all } n \geq 0.$$

If $\alpha = -1$, $c_1 = 1$, and $b = \frac{1}{2}$, (3) models the feedback process for the congestion window in TCP; see, e.g. [15] and [16]. In the TCP environment, 'success' stands for the arrival of a 'good' acknowledgement (one which positively acknowledges safe arrival of new and contiguous data), whereas 'failure' stands for the loss of a data packet. For certain values of α, c_1, and b, the process in (3) is a candidate for other, similar, control mechanisms. Of the papers just cited, [15] has existed on the Internet since 1996 and is frequently cited, but has never been published in the open literature. Reference [16] is a rewrite (draft) of [15]. The papers [5] and [6] give more references to literature on this topic.

This paper, and the ones just mentioned, use 'packet time', whereby the progress of time is (essentially) measured by the number of good acknowledgements that have been received. Many other papers studying TCP performance use 'clock time', whereby (apart from during slow start and fast recovery) time is measured in round trip times, or periods of time during which W good acknowledgements are received, usually under the assumption that, during that time, at most one packet gets lost or marked in the sense of [17]. If $\alpha = -1$ then the window increases by almost exactly c_1 maximum segment sizes during such a period. The assumption that at most one reduction of the congestion window occurs during one round trip time is reasonable if $\alpha < 0$ but questionable if $0 \leq \alpha$; see [11]. The paper [15] contains a translation between 'packet time-stationary' and 'clock time-stationary' distributions.

The somewhat overly complicated construction in (1) and (3) to define the process $W_{p,n}$ was chosen because later in this paper there will be a number of stochastic processes 'coupled' to the stochastic process $W_{p,n}$ by being generated by the same sequence of successes and failures.

Writing

$$\zeta_p(t) = p^{1/(1-\alpha)} W_{p,\lfloor t/p \rfloor}, \tag{4}$$

where $\lfloor\cdot\rfloor$ denotes the floor of x, i.e. the largest integer ν with $\nu \le x$, we see that as long as there is 'success' we have

$$\frac{\zeta_p(t+p)-\zeta_p(t)}{p} = c_1\zeta_p(t)^\alpha.$$

Hence, as $p \downarrow 0$ we approach the situation where there is a Poisson process of intensity 1 with 'events' $\dots, \tau_{-1}, \tau_0, \tau_1, \dots$, with a process $\zeta(t)$ defined by

$$\frac{\mathrm{d}}{\mathrm{d}t}\zeta(t) = c_1\zeta(t)^\alpha$$

'between' the events of the Poisson process and by

$$\zeta(\tau^+) = b\zeta(\tau^-)$$

'at' the points of the Poisson process.

If we now define

$$Z(t) = \frac{\zeta(t)^{1-\alpha}}{(1-\alpha)c_1},$$

then

$$\frac{\mathrm{d}}{\mathrm{d}t}Z(t) = 1$$

between the events of the Poisson process and

$$Z(\tau^+) = b^{1-\alpha}Z(\tau^-)$$

at the events of the Poisson process. Henceforth, we write c for $b^{1-\alpha}$.

As in [15] or [16], we see that the process $Z(\cdot)$ has as stationary distribution the distribution of

$$Z = \sum_{k=0}^{\infty} c^k E_k, \tag{5}$$

where $(E_k)_{k=0}^{\infty}$ are independent, identically distributed random variables each exponentially distributed with parameter 1. The distribution of Z in (5) is completely described in [15] and [16]. For example, for all (even complex) ν we have

$$\mathrm{E}[Z^\nu] = \Gamma(\nu+1)\prod_{k=0}^{\infty}\frac{1-c^{\nu+k}}{1-c^k}. \tag{6}$$

If ν is integer then this reduces to

$$\mathrm{E}[Z^k] = \frac{k!}{(1-c)(1-c^2)\cdots(1-c^k)},$$
$$\mathrm{E}[Z^{-k}] = \frac{(1-c)(1-c^2)\cdots(1-c^{k-1})}{(k-1)!\,c^{(k-1)k/2}}\log\left(\frac{1}{c}\right),$$

where k is a positive integer.

In analogy with (4), for $C > 0$ we define

$$\zeta^*_{p,C}(t) = p^{1/(1-\alpha)}W^*_{p,C,\lfloor t/p\rfloor}.$$

https://doi.org/10.1017/S0001867800001488 Published online by Cambridge University Press

In this paper we will study the stationary distribution of the process $(V_{p,n})_{n=0}^{\infty}$ defined by

$$V_{p,n} = \frac{p}{(1-\alpha)c_1}(W_{p,n})^{1-\alpha} = \frac{\zeta_p(pn)^{1-\alpha}}{c_1(1-\alpha)}, \tag{7}$$

as $p \downarrow 0$. Results for $(V_{p,n})_{n=0}^{\infty}$ can be translated into results for the process $(V^*_{p,C,n})_{n=0}^{\infty}$ defined by

$$V^*_{p,C,n} = \frac{p}{(1-\alpha)c_1}(W^*_{p,C,n})^{1-\alpha} = \frac{\zeta^*_{p,C}(pn)^{1-\alpha}}{c_1(1-\alpha)}.$$

These processes evolve as follows:

$$V_{p,n+1} = \begin{cases} V_{p,n}\left(1 + \dfrac{p}{(1-\alpha)V_{p,n}}\right)^{1-\alpha} & \text{if } \chi_{p,n} = \text{success}, \\ cV_{p,n} & \text{if } \chi_{p,n} = \text{failure}, \end{cases} \tag{8}$$

and

$$V^*_{p,C,n+1} = \begin{cases} V^*_{p,C,n}\left(1 + \dfrac{p}{(1-\alpha)V^*_{p,C,n}}\right)^{1-\alpha} & \text{if } \chi_{p,n} = \text{success}, \\ \max\left(cV^*_{p,C,n}, \dfrac{pC^{1-\alpha}}{(1-\alpha)c_1}\right) & \text{if } \chi_{p,n} = \text{failure}. \end{cases} \tag{9}$$

We will first study the stationary distributions of the process $V_{p,n}$, and then find a simple way of translating these results into results on stationary distributions of $V^*_{p,C,n}$. When possible, final results are formulated for the two cases $C = 0$ and $C > 0$ together, i.e. for the processes $\zeta^*_{p,C}(t)$ and $V^*_{p,C,n}$ with $C \geq 0$. In [13] it was proven that if $C > 0$ then the stationary distribution of $V^*_{p,C,n}$ exists and is unique. In [14] it was proven that a stationary distribution for $V_{p,n}$ (i.e. $C = 0$) exists, but the approach in that paper does not prove uniqueness. More will be said on this topic later. In [13] it was proven that, for all (constant) $C \geq 0$, if $p \downarrow 0$ then the process

$$\frac{\zeta^*_{p,C}(t)^{1-\alpha}}{(1-\alpha)c_1}$$

converges weakly to the process $Z(\cdot)$. While the result is obvious, the proof is not.

If $0 \leq \alpha < 1$ then the stationary distribution of $V_{p,n}$ ($C = 0$) is unique; see the final paragraph of this section. A technical problem is that, in the case $\alpha < 0$, we have not (yet) proven the uniqueness of the stationary distribution of $V_{p,n}$. Thus, when we say that the stationary distribution of $V_{p,n}$ converges to the distribution in (5), we really mean that, no matter how we choose the stationary distributions of $V_{p,n}$, as $p \downarrow 0$ they converge to the distribution in (5). In the remainder of the paper we will prove that this is indeed the case, and obtain stochastic dominance results and rate-of-convergence results for the converging stationary distributions. Section 9 contains more discussion of stationary distributions.

Processes of the sort studied in this note have also been studied in, e.g. [5] and [6], where, among other results, the weak convergence of stationary distributions was proved for the clock time process in the case $\alpha = -1$ (the TCP situation). References [5] and [6] involve a different proof technique and do not contain stochastic dominance results or rate-of-convergence results. Reference [11] studied issues pertaining to the stability of feedback protocols (through consideration of relaxation times), both for $\beta = 1$ (the case considered here) and for $\beta < 1$. Among the many other papers on the mathematical analysis of the performance of TCP, we mention only [1], [2], and, in particular, [3], where the case of scalable TCP ($\alpha = 0$) was studied (in clock time). The six papers just cited contain an extensive review of the literature.

Among the results obtained in this paper are the following three theorems.

Theorem 1. *Let V_p have the stationary distribution of the process $(V_{p,n})$, and for some $C > 0$ let $V^*_{p,C}$ have the stationary distribution of the process $(V^*_{p,C,n})$. Then, as $p \downarrow 0$, the distributions of V_p and $V^*_{p,C}$ converge weakly to the distribution of Z.*

The main focus of this paper is on the process $V_{p,n} (C = 0)$. Once results for that case are available there will follow corollaries for the case $C > 0$, i.e. for the process $V^*_{p,C,n}$.

More detailed results describe the rate of convergence in Theorem 1, and indicate that the 'error' in Theorem 1 is $O(p)$. For more detailed results, we must differentiate between the cases $\alpha \le 0$ and $0 \le \alpha < 1$.

Theorem 2. *Let $\alpha \le 0$. For some $C \ge 0$, let $V^*_{p,C}$ have the stationary distribution of $(V^*_{p,C,n})$. Then*

$$\lim_{p\downarrow 0} \mathrm{E}[(V^*_{p,C})^\nu] = \mathrm{E}[Z^\nu]$$

for every ν, $-\infty < \nu < \infty$, and the joint distributions of $(V^*_{p,C}, Z)$ can be chosen such that

$$\limsup_{p\downarrow 0} \mathrm{E}\bigg[\bigg|\frac{Z - V^*_{p,C}}{p}\bigg|^\nu\bigg] < \infty$$

for every $\nu \ge 0$. In other words, every sequence of positive probabilities p converging to 0 has a subsequence of probabilities p_k for which $Err_{p_k} := (Z - V^*_{p_k,C})/p_{p_k}$ converges weakly to a random variable Err. All moments $\mathrm{E}[|Err|^\nu]$, $\nu \ge 0$, of Err are then finite, and $\mathrm{E}[|Err_{p_k}|^\nu]$ converges to $\mathrm{E}[|Err|^\nu]$ for all $\nu \ge 0$.

Theorem 3. *Let $0 \le \alpha < 1$. For some $C \ge 0$, let $V^*_{p,C}$ have the stationary distribution of $(V^*_{p,C,n})$. Then*

$$\lim_{p\downarrow 0} \mathrm{E}[(V^*_{p,C})^\nu] = \mathrm{E}[Z^\nu]$$

for every ν, $0 \le \nu < \infty$, and the joint distribution of $(V_p, V^*_{p,C}, Z)$ can be chosen such that

$$\limsup_{p\downarrow 0} \mathrm{E}\bigg[\frac{|Z - V^*_{p,C}|^k}{p}\bigg] < \infty$$

for every integer $k \ge 1$. In other words, if $0 < \alpha < 1$ then the limiting random variable Err has a finite first moment, but is not claimed to have any higher moments. In the case $C = 0$, the joint distribution can be chosen such that also

$$\mathrm{P}\bigg\{V_p \le \frac{Z}{(1/p)|\log(1-p)|} < Z\bigg\} = 1.$$

The theorems above will first be proven for the process $V_{p,n}$, i.e. the case $C = 0$. Section 13 contains mechanisms to translate results for the process $V_{p,n}$ into results for the process $V^*_{p,C,n}$, i.e. the case $C > 0$. To prove the results for the process $V_{p,n}$, we introduce a number of auxiliary stochastic processes that will provide the desired stochastic bounds and, thus, provide results stronger than the theorems above. Sections 3 and 7 introduce these auxiliary stochastic processes and formulate the stronger results of which the theorems above are corollaries.

In the results above we let $p \downarrow 0$ while C is constant (possibly 0). We can also consider the problem of what happens if $p > 0$ is constant and $C \downarrow 0$. Clearly, in that case (with $V_{p,0} = V^*_{p,C,0}$),

$$\lim_{C \downarrow 0} V^*_{p,C,n} = V_{p,n}.$$

However, the convergence of processes does not always guarantee the convergence of stationary distributions. If $0 \leq \alpha < 1$ then we do have uniqueness of the stationary distribution of $V_{p,n}$ (see the final paragraph of this section), and weak convergence of the stationary distributions of $V^*_{p,C,n}$ to the stationary distribution of $V_{p,n}$; see Section 2.

In Sections 3–12, part of Section 13, and Section 14 we derive and use 'stochastic dominance results' in which various processes defined on the basis of the same sequence $(\chi_{p,n})_{n=0}^{\infty}$ (all with the same value of p) are compared. Only in Section 2 do we compare processes for different values of p.

The uniqueness of the stationary distributions for $0 \leq \alpha < \beta = 1$ follows from the observation that, in this case, if there are two positive starting positions, $W_{p,0,1}$ and $W_{p,0,2}$, which give rise to the processes $(W_{p,n,1})_{n=0}^{\infty}$ and $(W_{p,n,2})_{n=0}^{\infty}$ (with, of course, the same sequences of successes and failures) and satisfy $0 < W_{p,0,1} < W_{p,0,2}$, then $0 < W_{p,n,1} < W_{p,n,2}$ for all n and $W_{p,n,1}/W_{p,n,2}$ is nondecreasing, i.e. it remains the same at failures, increases at successes, and converges to 1 as $n \to \infty$.

2. Stochastic dominance and different values of *p* and *C*

For decreasing values of p, there are more successes and fewer failures, so a natural question to ask is whether $W^*_{p,C,n}$ increases as p decreases (and, say, $W^*_{p,C,0}$ is constant). This is not always true. It is true if $0 \leq \alpha < \beta = 1$, and for $\alpha < 0$ it is still true as long as $C > 0$ is sufficiently large; see [12] for details. Similarly, with p and $W^*_{p,C,0}$ constant, $W^*_{p,C,n}$ is increasing in C as long as $0 \leq \alpha < \beta = 1$. This proves, under these conditions, the convergence of the stationary distributions of $V^*_{p,C,n}$ to the stationary distribution of $V_{p,n}$. For $\alpha < 0$, it can only be proven that $W^*_{p,C,n}$ increases with C for C sufficiently large; see [12] and [14] for details. In this case we cannot conclude that the stationary distributions of $V^*_{p,C,n}$ converge to the stationary distribution of $V_{p,n}$.

3. Method of attack and intermediate results

We define the process $(X_{p,n})_{n=0}^{\infty}$ by

$$X_{p,n+1} = \begin{cases} X_{p,n} + p & \text{if } \chi_{p,n} = \text{success}, \\ cX_{p,n} & \text{if } \chi_{p,n} = \text{failure}, \end{cases} \tag{10}$$

and we give $X_{p,0}$ and $V_{p,0}$ some joint distribution.

Since the processes $(V_{p,n})_{n=0}^{\infty}$ and $(X_{p,n})_{n=0}^{\infty}$ are driven by the same sequence of successes and failures, the processes are dependent and $(X_{p,n})_{n=0}^{\infty}$ and $(V_{p,n})_{n=0}^{\infty}$ have a joint distribution. For this joint distribution we will prove the following result.

Lemma 1. *If $\alpha \leq 0$ and $V_{p,k} \geq X_{p,k}$ for some $k \geq 0$, then $V_{p,n} \geq X_{p,n}$ for all $n \geq k$. If $0 \leq \alpha < 1$ and $V_{p,k} \leq X_{p,k}$ for some $k \geq 0$, then $V_{p,n} \leq X_{p,n}$ for all $n \geq k$. If $\alpha = 0$ and $V_{p,k} = X_{p,k}$ for some $k \geq 0$, then $V_{p,n} = X_{p,n}$ for all $n \geq k$.*

This lemma will be proven in Section 6. We also have the following lemma.

Lemma 2. *The process* $(X_{p,n})_{n=0}^{\infty}$ *has a unique stationary distribution. If* X_p *has this stationary distribution then* X_p *is of the form*

$$X_p = p \sum_{k=0}^{\infty} c^k G_{p,k}, \tag{11}$$

where $(G_{p,k})_{k=0}^{\infty}$ *are independent, identically distributed random variables each geometrically distributed with parameter* $1 - p$*:*

$$\mathrm{P}\{G_{p,k} = n\} = p(1-p)^n.$$

Hence, X_p *has the Laplace–Stieltjes transform*

$$\phi_{X_p}(s) = \mathrm{E}[\exp\{-sX_p\}] = \prod_{k=0}^{\infty} \frac{p}{1-(1-p)\exp\{-pc^k s\}}, \tag{12}$$

and

$$\mathrm{E}[X_p] = \frac{1-p}{1-c}, \qquad \mathrm{var}(X_p) = \frac{1-p}{1-c^2}.$$

Proof. There are at least two obvious ways to prove Lemma 2. The most intuitive proof duplicates the proof given in [15] of the similar result (5) for the process $Z(t)$. This proof works by 'looking back in time'. A less intuitive proof uses the fact that

$$\mathrm{E}[\exp\{-sX_{p,n}\}] = (1-p)\,\mathrm{E}[\exp\{-s(X_{p,n-1}+p)\}] + p\,\mathrm{E}[\exp\{-scX_{p,n-1}\}].$$

Assuming that $X_{p,n-1}$ and $X_{p,n}$ have the same distribution leads to a recursion that proves (12) and, thus, Lemma 2.

Equation (10) trivially shows that

$$\mathrm{E}[X_{p,n}] = \frac{1-p}{1-c} + (1-p(1-c))^n \left(\mathrm{E}[X_{p,0}] - \frac{1-p}{1-c}\right),$$

which, if $\mathrm{E}[X_{p,0}]$ is finite, proves tightness of the family of random variables $(X_{p,n})_{n=0}^{\infty}$. It is easy to extend this to the case in which $\mathrm{E}[X_{p,0}]$ is not finite. In addition, if we have two different initial values, $X_{p,0,1}$ and $X_{p,0,2}$, then with identical sequences of successes and failures we have

$$X_{p,n,1} - X_{p,n,2} = c^{N(n)}(X_{p,0,1} - X_{p,0,2}) \tag{13}$$

for all $n \geq 0$, where $N(n)$ is the number of failures among $\chi_0, \dots, \chi_{n-1}$. This proves that the distribution of $X_{p,n}$ becomes independent of $X_{p,0}$ and n. The details are left to the reader.

Lemmas 1 and 2 provide initial stochastic bounds for stationary distributions of the process $(V_{p,n})_{n=0}^{\infty}$. For $\alpha < 0$, they provide a stochastic lower bound, proving that $V = 0$ is not a critical point. For $0 < \alpha < 1$, they provide a stochastic upper bound, proving that $V = \infty$ is not a critical point. It will be proven that these bounds are asymptotically tight as $p \downarrow 0$.

Before providing the lacking bounds, we show that the stationary distribution of $X_{p,n}$ is very close to the distribution of the random variable Z in (5). It is clear that that random variable Z has Laplace–Stieltjes transform

$$\phi_Z(s) = \mathrm{E}[\exp\{-sZ\}] = \prod_{k=0}^{\infty} \frac{1}{1+c^k s}, \tag{14}$$

and that

$$\mathrm{E}[Z] = \frac{1}{1-c}, \qquad \mathrm{var}(Z) = \frac{1}{1-c^2}.$$

It is immediately obvious that, as $p \downarrow 0$, the Laplace–Stieltjes transform (12) converges to the Laplace–Stieltjes transform (14). Thus, the distribution of X_p converges weakly to the distribution of Z. However, more can be said, as follows.

Lemma 3. *The random variables X_p, in (11), and Z, in (5), can be given a joint distribution for which, with probability* 1,

$$\max\left(0, \frac{Z}{(1/p)\log(1/[1-p])} - \frac{p}{1-c}\right) < X_p \le \frac{Z}{(1/p)\log(1/[1-p])} < Z. \tag{15}$$

Hence, for this joint distribution,

$$\mathrm{E}\left[\left|\frac{Z}{(1/p)\log(1/[1-p])} - X_p\right|^{\nu}\right] = \mathrm{E}\left[\left(\frac{Z}{(1/p)\log(1/[1-p])} - X_p\right)^{\nu}\right] \le \left(\frac{p}{1-c}\right)^{\nu}$$

for $\nu \ge 0$.

We remind the reader (see, e.g. [9]) that the random variable S is *stochastically smaller* than the random variable T if $\mathrm{P}\{S \le x\} \ge \mathrm{P}\{T \le x\}$ for *all* x. This is the case if and only if there is a joint distribution of S and T for which $\mathrm{P}\{S \le T\} = 1$. Thus, we have proven that X_p is stochastically smaller than $Z/[(1/p)\log(1/[1-p])]$ and that the distributions are almost the same. This type of argument will be used several times in what follows.

Proof of Lemma 3. If E is an exponentially distributed random variable with parameter 1 (and, therefore, expected value 1), then the random variable H_p defined by

$$H_p = np \quad \text{for} \quad n \log\left(\frac{1}{1-p}\right) \le E < (n+1)\log\left(\frac{1}{1-p}\right) \tag{16}$$

has the property that

$$\mathrm{P}\{H_p = np\} = p(1-p)^n.$$

Thus, if $(E_k)_{k=0}^{\infty}$ are independent, identically distributed random variables each exponentially distributed with parameter 1, we define $(H_{p,k})_{k=0}^{\infty}$ to be functions of $(E_k)_{k=0}^{\infty}$ as in (16), and we define

$$Z = \sum_{k=0}^{\infty} c^k E_k, \qquad X_p = \sum_{k=0}^{\infty} c^k H_{p,k},$$

then Z and X_p have the required marginal distributions and (15) holds with probability 1.

The method of attack is now clear. We have proven that, for small p, X_p and Z have almost the same distribution (including a rate-of-convergence result). In fact, the distribution of X_p is even closer to the distribution of $Z/[(1/p)\log(1/[1-p])]$. Remaining to be proven is that the (or any) stationary distribution of $V_{p,n}$ must, for small p, be very close to the distribution of X_p. To prove this result we need different approaches in the respective cases $\alpha < 0$ and $0 < \alpha < 1$.

In the case $0 < \alpha < 1$, we will use the methods of linear programming and duality; see Sections 11 and 12. In the case $\alpha < 0$, the approach will be as follows. We define

$$c(\alpha) = \frac{1}{(1+|\alpha|)^2} \sup_{0<z\leq 1} \left(\frac{(1+z)^{1+|\alpha|} - 1 - (1+|\alpha|)z}{z^2} \right),$$
$$d(\alpha) = \frac{1}{(1+|\alpha|)^{1+|\alpha|}} \sup_{0<x\leq 1} ((1+x)^{1+|\alpha|} - x^{1+|\alpha|} - x^{|\alpha|}(1+|\alpha|)),$$

and then define the function $f_{p,\alpha}(\cdot)\colon (0,\infty) \to (0,\infty)$ by

$$f_{p,\alpha}(v) = \begin{cases} \dfrac{c(\alpha)p^2}{v} & \text{if } \dfrac{p}{1+|\alpha|} \leq v < \infty, \\ \dfrac{d(\alpha)p^{1+|\alpha|}}{v^{|\alpha|}} & \text{if } 0 < v < \dfrac{p}{1+|\alpha|}. \end{cases} \tag{17}$$

Next, we define the stochastic process $(Y_{p,n})_{n=0}^{\infty}$ by

$$Y_{p,n+1} = \begin{cases} Y_{p,n} + p + f_{p,\alpha}(X_{p,n}) & \text{if } \chi_{p,n} = \text{success}, \\ cY_{p,n} & \text{if } \chi_{p,n} = \text{failure}, \end{cases} \tag{18}$$

where $X_{p,0}$, $V_{p,0}$, and $Y_{p,0}$ are given some joint distribution. Since the processes $X_{p,n}$, $V_{p,n}$, and $Y_{p,n}$ are all defined on the basis of the same sequence of successes and failures, their joint distribution is well defined.

As in (13), we observe that if we have two starting values, $Y_{p,0,1}$ and $Y_{p,0,2}$, for the process $Y_{p,n}$, but identical sequences of successes and failures, identical values for $X_{p,0}$ and, therefore, identical sequences $(X_{p,n})_{n=0}^{\infty}$, then

$$Y_{p,n,1} - Y_{p,n,2} = c^{N(n)}(Y_{p,0,1} - Y_{p,0,2}).$$

Therefore, if the process $(X_{p,n}, Y_{p,n})$ has a stationary distribution, that stationary distribution is unique.

Lemma 4. *If $\alpha \leq 0$ (and $c(\alpha)$ and $d(\alpha)$ are chosen as above) and, for some k, $X_{p,k} \leq V_{p,k} \leq Y_{p,k}$, then*

$$X_{p,n} \leq V_{p,n} \leq Y_{p,n} \quad \text{for all } n \geq k. \tag{19}$$

Lemma 4 will be proven in Section 7.

Equations (18) and (19) trivially show that if (X_p, V_p, Y_p) has the joint stationary distribution of the process $(X_{p,n}, V_{p,n}, Y_{p,n})$ and $\mathrm{E}[Y_p] < \infty$, then

$$\mathrm{P}\{0 < X_p < V_p < Y_p < \infty\} = 1, \qquad \mathrm{E}[Y_p - X_p] = \frac{(1-p)}{(1-c)p}\mathrm{E}[f_{p,\alpha}(X_p)]. \tag{20}$$

This makes it necessary to compute $\mathrm{E}[f_{p,\alpha}(X_p)]$. It should be noted that, while $X_{p,n}$ is guaranteed to have a stationary distribution (which is unique), there is no guarantee that $Y_{p,n}$ has a stationary distribution. If $(X_{p,n}, Y_{p,n})$ has a joint stationary distribution, then it is unique, $(X_{p,n}, V_{p,n}, Y_{p,n})$ has a joint stationary distribution, and (20) holds. A necessary and sufficient condition for the existence of a stationary distribution with $\mathrm{E}[Y_p] < \infty$ is that $\mathrm{E}[f_{p,\alpha}(X_p)] < \infty$; see [12] for details.

Lemma 5. *If $\alpha < 0$ and X_p has the distribution in (12), then*

$$\mathrm{E}[f(X_p)] = p^2\bigg(c(\alpha)\,\mathrm{E}\bigg[\frac{1}{X_p}\,\mathbf{1}\bigg(X_p \geq \frac{p}{1+|\alpha|}\bigg)\bigg] + \bigg(\frac{p}{c}\bigg)^{|\alpha|} d(\alpha)\,\mathrm{E}\bigg[\frac{1}{X_p^{|\alpha|}}\,\mathbf{1}\bigg(X_p < \frac{p}{c(1+|\alpha|)}\bigg)\bigg]\bigg),$$

where $\mathbf{1}(\cdot)$ is the indicator function.

Lemma 5 makes it necessary to study $\mathrm{E}[(X_p)^\nu]$ for $\nu < 0$, and will be proven in Section 8.

As simple corollary of Lemma 3 and (6) (and the Helly–Bray theorem; see, e.g. [10, pp. 180–185]), we see that, for all $\nu \geq 0$, $\mathrm{E}[X_p^\nu] \leq \mathrm{E}[Z^\nu] < \infty$ and

$$\lim_{p\downarrow 0} \mathrm{E}[X_p^\nu] = \mathrm{E}[Z^\nu]. \tag{21}$$

We will see that (practically speaking) (21) also holds for $\nu < 0$, with the restriction that p must be sufficiently small to guarantee that $\mathrm{E}[X_p^\nu] < \infty$.

Lemma 6. *For every $\nu > 0$, there exists a $p(\nu)$, $0 < p(\nu) < 1$, and a $B(\nu)$, $0 < B(\nu) < \infty$, such that*

$$\mathrm{E}[X_p^{-\nu}] \leq B(\nu) \quad \text{for all } p,\ 0 < p \leq p(\nu).$$

Lemma 6 will be proven in Section 5.

Corollary 1. *Equation (21) holds for all ν, $-\infty < \nu < \infty$.*

Corollary 2. *For $\alpha < 0$ and all p, $0 < p \leq \min(p(1), p(|\alpha|))$, $(X_{p,n}, Y_{p,n})$ has a unique stationary (joint) distribution and there exists a $D(\alpha)$, $0 < D(\alpha) < \infty$, such that if (X_p, Y_p) has this joint stationary distribution, then*

$$\mathrm{P}\{0 < X_p \leq Y_p < \infty\} = 1, \qquad \mathrm{E}[Y_p - X_p] \leq pD(\alpha). \tag{22}$$

Remark 1. In [14] (extended version) it was proved, using an alternative method, that in fact $\mathrm{E}[(X_p)^{-\nu}] < \infty$ if and only if $0 < p < c^\nu$. That alternative method has not produced uniform upper bounds $B(\nu)$.

4. The process $(\log((1-\alpha)V_{p,n}/p))_{n=0}^{\infty}$

We first derive two results (Lemmas 7 and 8) that neither depend on, nor are used in the proofs of, the results of Section 3. We rewrite (8) as

$$\log\frac{(1-\alpha)V_{p,n+1}}{p} = \begin{cases} \log\dfrac{(1-\alpha)V_{p,n}}{p} + (1-\alpha)\log\bigg(1+\dfrac{p}{(1-\alpha)V_{p,n}}\bigg) & \text{if } \chi_{p,n} = \text{success}, \\ \log\dfrac{(1-\alpha)V_{p,n}}{p} + \log c = \log\dfrac{(1-\alpha)V_{p,n}}{p} - |\log c| & \text{if } \chi_{p,n} = \text{failure}, \end{cases} \tag{23}$$

and study the consequences of doing so separately for $\alpha \leq 0$ and $0 \leq \alpha < 1$.

https://doi.org/10.1017/S0001867800001488 Published online by Cambridge University Press

First we consider the case $\alpha \le 0$. In this case, immediately after every success we have

$$\log \frac{(1+|\alpha|)V_{p,n+1}}{p} \ge (1+|\alpha|)\log(1+|\alpha|) - |\alpha|\log|\alpha| \ge 0.$$

Define the stochastic process $N_{p,n}$ by

$$N_{p,n+1} = \begin{cases} 0 & \text{if } \chi_{p,n} = \text{success}, \\ N_{p,n} - |\log c| & \text{if } \chi_{p,n} = \text{failure}. \end{cases}$$

It is clear that, after the first success, $N_{p,k} \le \log((1+|\alpha|)V_{p,k}/p)$ always. It is also clear that the process $N_{p,n}$ is stationary, with stationary distribution

$$\mathrm{P}\{N_p = -k|\log c|\} = p^k(1-p), \qquad \mathrm{E}[N_p] = -\frac{p}{1-p}|\log c|.$$

Lemma 7. *If $\alpha \le 0$ then, for any stationary distribution of $V_{p,n}$,*

$$\mathrm{P}\left\{\log \frac{(1+|\alpha|)V_p}{p} \le -k|\log c|\right\} \le p^k$$

for every nonnegative integer k, and

$$\mathrm{E}\left[\left|\log \frac{(1+|\alpha|)V_p}{p}\right| \mathbf{1}\left(\log \frac{(1+|\alpha|)V_p}{p} \le 0\right)\right] < \mathrm{E}[|N_p|] = \frac{p}{1-p}|\log c| < \infty,$$

where $\mathbf{1}(\cdot)$ is the indicator function.

Next we consider the case $0 \le \alpha < 1$, in which, immediately after every success, we have

$$\log \frac{(1-\alpha)V_{p,n+1}}{p} > \alpha \log \frac{(1-\alpha)V_{p,n}}{p}.$$

Here we define the stochastic process $M_{p,n}$ by

$$M_{p,n+1} = \begin{cases} \alpha M_{p,n} & \text{if } \chi_{p,n} = \text{success}, \\ M_{p,n} - |\log c| & \text{if } \chi_{p,n} = \text{failure}. \end{cases}$$

It is clear that if $M_{p,n} \le \log((1-\alpha)V_{p,n}/p)$ for some n, then this holds for all $k \ge n$. Also, the process $M_{p,n}$ is stationary. It is easily seen (see the similar result for the process $X_{p,n}$ in the previous section) that the stationary distribution has the form

$$M_p = -|\log c| \sum_{k=0}^{\infty} \alpha^k Q_k,$$

where the Q_k are independent, identically geometrically distributed random variables with

$$\mathrm{P}\{Q_k = n\} = p^n(1-p).$$

Thus,

$$\mathrm{E}[\exp\{-sM_p\}] = \prod_{k=0}^{\infty} \frac{1-p}{1-p\exp\{s|\log c|\alpha^k\}},$$

https://doi.org/10.1017/S0001867800001488 Published online by Cambridge University Press

and this holds for $s < |\log p|/|\log c|$. It follows that

$$\mathrm{E}[M_p] = -\frac{p|\log c|}{(1-p)(1-\alpha)},$$

and we thus have the following result.

Lemma 8. *If* $0 \le \alpha < 1$ *then, for the stationary distribution of* $V_{p,n}$,

$$\mathrm{P}\left\{\frac{(1-\alpha)V}{p} \le x\right\} \le \mathrm{P}\{M_p \le \log x\}$$

and

$$\mathrm{E}\left[\left|\log\frac{(1-\alpha)V_p}{p}\right| \mathbf{1}\left(\log\frac{(1-\alpha)V_p}{p} \le 0\right)\right] < \mathrm{E}[|M_p|] = \frac{p|\log c|}{(1-p)(1-\alpha)} < \infty.$$

Next we combine the Lemmas 7 and 8 with the results of the previous section.

Lemma 9. *If either* $0 \le \alpha < 1$ *or both* $\alpha < 0$ *and* $0 < p \le \min(p(1), p(|\alpha|))$, *where* $p(\nu)$ *is as defined in Lemma 6, then, for every stationary distribution of the process* $V_{p,n}$,

$$\mathrm{E}\left[\left|\log\frac{(1-\alpha)V_p}{p}\right|\right] < \infty.$$

Proof. For $\log((1-\alpha)V_p/p) \le 0$, the result has been proven in this section. For the case $\log((1-\alpha)V_p/p) > 0$, we use results from the previous section. For $\alpha < 0$ and $0 < p \le \min(p(1), p(|\alpha|))$, we have $(\log V_p) < V_p \le Y_p$, and Y_p has a finite first moment. For $0 \le \alpha < 1$, we have $\log V_p < V_p \le X_p$, and X_p has a finite first moment.

Remark 2. A minimal modification of the proofs of the previous results also shows tightness of the distributions of $X_{p,n}$, $V_{p,n}$, and $Y_{p,n}$, at least in the case where either $0 \le \alpha < 1$ or both $\alpha < 0$ and $0 < p \le \min(p(1), p(|\alpha|))$.

Lemma 9 and (23) together yield our next theorem, the proof of which we thus omit.

Theorem 4. *If either* $0 \le \alpha < 1$ *or both* $\alpha < 0$ *and* $0 < p \le \min(p(1), p(|\alpha|))$, *where* $p(\nu)$ *is as defined in Lemma 6, then, for every stationary distribution of the process* $V_{p,n}$,

$$\mathrm{E}\left[\log\left(1 + \frac{p}{(1-\alpha)V_p}\right)\right] = \frac{p|\log c|}{(1-p)(1-\alpha)}. \tag{24}$$

Theorem 4 will be used in Sections 11 and 12, to treat the case $0 < \alpha < 1$.

5. The proof of Lemma 6

In order to prove Lemma 6, we observe that if Q is any nonnegative random variable with Laplace transform $\phi_Q(s) = \mathrm{E}[\exp\{-sQ\}]$, then, for all $\nu > 0$,

$$\int_0^\infty s^{\nu-1}\phi_Q(s)\,\mathrm{d}s = \Gamma(\nu)\,\mathrm{E}\left[\frac{1}{Q^\nu}\right].$$

(No proof needed.)

We will also use the following lemma for the Laplace transform (12).

Lemma 10. *If* $0 < p < \frac{1}{3}$, $0 < rp < \frac{1}{3}$, *and* $0 < r < s$, *then*

$$\phi_{X_p}(s) < (1 + \tfrac{1}{2}r)^{-[\log(s) - \log(r)]/\log(1/c)}.$$

Once we have Lemma 10, Lemma 6 is proven as follows.

Proof of Lemma 6. Choose a $\nu > 0$ and then choose an $r > 0$ such that

$$\log\left(1 + \frac{r}{2}\right) > \nu \log \frac{1}{c}, \quad \text{i.e.} \quad 1 + \frac{r}{2} > \left(\frac{1}{c}\right)^{\nu}.$$

Write

$$\int_0^\infty s^{\nu-1}\phi_{X_p}(s)\,\mathrm{d}s = \int_0^r s^{\nu-1}\phi_{X_p}(s)\,\mathrm{d}s + \int_r^\infty s^{\nu-1}\phi_{X_p}(s)\,\mathrm{d}s. \tag{25}$$

In the first integral on the right-hand side of (25), we use the fact that $0 < \phi_{X_p}(s) < 1$. In the second integral, $s > r$ as long as $0 < p < \frac{1}{3}\min(1, 1/r)$, Lemma 10 can be used, and straightforward arithmetic gives

$$\int_0^\infty s^{\nu-1}\phi_{X_p}(s)\,\mathrm{d}s < \frac{r^\nu \log(1 + r/2)}{\nu[\log(1 + r/2) - \nu\log(1/c)]}.$$

This proves Lemma 6.

The proof of Lemma 10 is next.

Proof of Lemma 10. We choose any $s > 0$ and any r, $0 < r < s$. Next, we choose K_0 to be the smallest integer k for which $c^k s < r$. (K_0 thus depends on r and s.) Since $0 < c < 1$ and $0 < r < s$, $K_0 > 0$ and

$$c^{K_0 - 1} \geq r/s > c^{K_0}.$$

We have

$$\phi_{X_p}(s) < \prod_{k=0}^{K_0-1} \frac{p}{1 - (1-p)\exp\{-pc^k s\}}. \tag{26}$$

We will derive an upper bound for the right-hand side of (26). For $0 \leq k \leq K_0 - 1$, we have

$$c^k ps \geq rp$$

and, hence,

$$\exp\{-c^k ps\} \leq \exp\{-rp\} < 1 - rp + \frac{(rp)^2}{2}.$$

Therefore,

$$1 - (1-p)\exp\{-c^k ps\} > 1 - (1-p)\left(1 - rp + \frac{(rp)^2}{2}\right) = p\left(1 + (1-p)r\left(1 - \frac{rp}{2}\right)\right).$$

We check that, since $0 < rp < \frac{1}{3}$ and $0 < p \leq \frac{1}{3}$,

$$1 - (1-p)\exp\{-c^k ps\} > p(1 + \tfrac{5}{6}(1-p)r) > p(1 + \tfrac{5}{9}r) > p(1 + \tfrac{1}{2}r) > p > 0.$$

Hence,

$$\phi_{X_p}(s) < (1 + \tfrac{1}{2}r)^{-K_0}.$$

Since $c^{K_0} < r/s$, we have $K_0 > [\log(s) - \log(r)]/\log(1/c)$ and, thus,

$$\phi_{X_p}(s) < (1 + \tfrac{1}{2}r)^{-[\log(s) - \log(r)]/\log(1/c)}.$$

This completes the proof of Lemma 10.

https://doi.org/10.1017/S0001867800001488 Published online by Cambridge University Press

6. The proof of Lemma 1

We define the function $R_{p,\alpha}(v)$, $v > 0$, by

$$R_{p,\alpha}(v) = v\left(1 + \frac{p}{(1-\alpha)v}\right)^{1-\alpha} - v - p. \tag{27}$$

Whenever it does not lead to confusion, we write $R(\cdot)$ for $R_{p,\alpha}(\cdot)$. With $V_{p,n}$ as in (7), we rewrite (8) as

$$V_{p,n+1} = \begin{cases} V_{p,n} + p + R(V_{p,n}) & \text{if } \chi_{p,n} = \text{success}, \\ cV_{p,n} & \text{if } \chi_{p,n} = \text{failure}. \end{cases} \tag{28}$$

With the substitution $z = p/[(1-\alpha)v]$, $R(v)$ can be rewritten as

$$\frac{p}{(1-\alpha)z}((1+z)^{1-\alpha} - 1 - (1-\alpha)z).$$

With the substitution $x = (1-\alpha)v/p = 1/z$, $R(v)$ can be rewritten as

$$\frac{p}{(1-\alpha)}(x^{\alpha}(1+x)^{1-\alpha} - x - (1-\alpha)). \tag{29}$$

For $\alpha < 0$, this shows that

$$0 < R_{p,\alpha}(v) \leq f_{p,\alpha}(v) < \infty$$

for all $0 < v < \infty$, where $f_{p,\alpha}(\cdot)$ is as in (17), that

$$R(v) \sim \frac{|\alpha| p^2}{2(1+|\alpha|)v} \quad \text{for } v \uparrow \infty, \qquad R(v) \sim \frac{p^{1+|\alpha|}}{(1+|\alpha|)^{1+|\alpha|} v^{|\alpha|}} \quad \text{for } v \downarrow 0,$$

and that $R(\cdot)$ is monotone (decreasing, from ∞ at 0 to 0 at ∞) and convex.

For $0 < \alpha < 1$, it shows that $R(v) < 0$ for all v, $0 < v < \infty$, that

$$R(v) \sim -\frac{\alpha p^2}{2(1-\alpha)v} \quad \text{for } v \uparrow \infty, \qquad R(v) + p \sim \frac{p^{1-\alpha} v^{\alpha}}{(1-\alpha)^{1-\alpha}} \quad \text{for } v \downarrow 0,$$

and that $R(\cdot)$ is monotone (increasing, from $-p$ at 0 to 0 at ∞) and concave.

For $\alpha = 0$, of course $R(v) \equiv 0$.

To prove the monotonicity and convexity for $\alpha < 0$, take the derivative with respect to x in (29), write $A = (1+x)/x$ (hence, $1 < A < \infty$), in terms of which the derivative equals $(1+|\alpha|)A^{|\alpha|} - |\alpha| A^{1+|\alpha|} - 1$, and prove that this expression equals 0 for $A = 1$, i.e. $x = \infty$, and is decreasing in A for $A > 1$, i.e. increasing in x for $x > 0$. Proving the other items is similar or easier.

An interesting special case is where $\alpha = -1$ (the TCP case), where $R(v) = p^2/(4v)$. Here (10), (28), and the results for the function R above prove Lemma 1, and (28) and (10) show that, for the joint stationary distribution of X_p and V_p,

$$\mathrm{E}[V_p - X_p] = \frac{1-p}{p(1-c)} \mathrm{E}[R(V_p)].$$

https://doi.org/10.1017/S0001867800001488 Published online by Cambridge University Press

For $\alpha \leq 0$ this becomes

$$\mathrm{E}[V_p - X_p] = \mathrm{E}[|V_p - X_p|] = \frac{1-p}{p(1-c)}\,\mathrm{E}[R(V_p)],$$

while for $0 < \alpha < 1$ it becomes

$$\mathrm{E}[X_p - V_p] = \mathrm{E}[|X_p - V_p|] = \frac{1-p}{p(1-c)}\,\mathrm{E}[|R(V_p)|]. \tag{30}$$

What we are going to do next amounts, conceptually, to finding an upper bound for $\mathrm{E}[|R(V_p)|]$. In the case $0 < \alpha < 1$, this is exactly what we will do. This case will be treated in Section 11, using the material of Section 4. In the case $\alpha < 0$, we will do something more complicated, based on the fact that, for $\alpha < 0$, V_p is stochastically larger (in fact larger with probability 1) than X_p. To handle this case we will use, in Section 7, the process $Y_{p,n}$ introduced in (18). First we introduce another auxiliary process.

7. The proof of Lemma 4

Throughout this section we have $\alpha \leq 0$. We define

$$\Lambda_{p,n+1} = \begin{cases} \Lambda_{p,n} + p + R(X_{p,n}) & \text{if } \chi_{p,n} = \text{success}, \\ c\Lambda_{p,n} & \text{if } \chi_{p,n} = \text{failure}. \end{cases}$$

Because the function $R(\cdot)$ is positive and decreasing, and because $0 < R(v) \leq f(v)$, it follows that if, for some n,

$$X_{p,n} \leq V_{p,n} \leq \Lambda_{p,n} \leq Y_{p,n} \tag{31}$$

(where $Y_{p,n}$ is as in (18)), then this holds for all $n + m \geq n$ and, therefore, for the stationary joint distribution, if any. This result is often used in the situation where all the processes in (31) have the same initial value. Lemma 4 is a corollary to (31).

As part of the proof of the statement above, we noted that if (31) holds, then

$$R(Y_{p,n}) \leq R(\Lambda_{p,n}) \leq R(V_{p,n}) \leq R(X_{p,n}) \leq f(X_{p,n}), \tag{32}$$

and this inequality holds for the joint stationary distribution and for the corresponding moments. Lemma 5, which will be proven in the next section, has as a consequence the fact that, for the stationary distribution there, $\mathrm{E}[f(X_p)] = O(p^2)$ (as $p \downarrow 0$), this result also holds for the other expected values in (32).

A more refined analysis in [12] showed that if $\alpha < 0$ then in fact

$$\mathrm{E}[R(X_p)] = p^2 \frac{|\alpha|\,|\log c|}{2(1+|\alpha|)} + O(p^3), \qquad p \downarrow 0.$$

Similar results can be derived for higher moments and moments of derivatives of $R(\cdot)$.

8. The proof of Lemma 5

Throughout this section we have $\alpha \leq 0$. For $\mathrm{E}[f(X_p)]$ we have

$$\mathrm{E}[f(X_p)] = c(\alpha)p^2\,\mathrm{E}\bigg[\frac{1}{X_p}\mathbf{1}\bigg(X_p \geq \frac{p}{1+|\alpha|}\bigg)\bigg] + d(\alpha)p^{1+|\alpha|}\,\mathrm{E}\bigg[\frac{1}{X_p^{|\alpha|}}\mathbf{1}\bigg(X_p < \frac{p}{1+|\alpha|}\bigg)\bigg].$$

https://doi.org/10.1017/S0001867800001488 Published online by Cambridge University Press

If $\alpha \leq -1$ then this is enough to prove the actual goal, (22). If $-1 < \alpha < 0$ then one more trick is needed to first prove Lemma 5.

Since X_p can be written as in (11) and $0 < c < 1$, $X_p < p$ implies that $G_{p,0} = 0$. The latter event has probability p. The conditional distribution of X_p given that $G_{p,0} = 0$ is the same as the unconditional distribution of cX_p. Hence, if $A > 1$ and $-\infty < \nu < \infty$, then

$$\mathrm{E}\bigg[X_p^\nu \mathbf{1}\bigg(X_p < \frac{p}{A}\bigg)\bigg] = pc^\nu \mathrm{E}\bigg[X_p^\nu \mathbf{1}\bigg(X_p < \frac{p}{cA}\bigg)\bigg]. \tag{33}$$

Thus,

$$\mathrm{E}[f(X_p)] = p^2\bigg(c(\alpha)\mathrm{E}\bigg[\frac{1}{X_p}\mathbf{1}\bigg(X_p \geq \frac{p}{1+|\alpha|}\bigg)\bigg] + \bigg(\frac{p}{c}\bigg)^{|\alpha|} d(\alpha)\mathrm{E}\bigg[\frac{1}{X_p^{|\alpha|}}\mathbf{1}\bigg(X_p < \frac{p}{c(1+|\alpha|)}\bigg)\bigg]\bigg).$$

This proves Lemma 5 and, thus, (22).

This completes the proofs of the results of Sections 3 and 4.

9. Stationary distributions and the main results

Let $F_{V_p}(\cdot)$ be a stationary distribution function for the process $V_{p,n}$ (with $F_{V_p}(0) = 0$) and let $V_{p,0}$ have this distribution. Choose $X_{p,0} = V_{p,0}$ and, if $\alpha \leq 0$, also choose $Y_{p,0} = \Lambda_{p,0} = V_{p,0}$. In the latter case, we now have

$$X_{p,n} \leq V_{p,n} \leq \Lambda_{p,n} \leq Y_{p,n} \quad \text{for all } n.$$

By letting $n \to \infty$ and using the results from the previous sections we now obtain, for $\alpha \leq 0$, a result stronger than Theorem 1.

Theorem 5. *If $\alpha \leq 0$ then, for all $p \leq \min(p(1), p(|\alpha|))$, there exists a stationary distribution of the process $(X_{p,n}, V_{p,n}, \Lambda_{p,n}, Y_{p,n})$ as in Section 7, and if $(X_p, V_p, \Lambda_p, Y_p)$ has such a joint stationary distribution then it has a joint distribution with the random variable Z (where (X_p, Z) has a joint distribution as in Lemma 3), $\mathrm{P}\{X_p \leq V_p \leq \Lambda_p \leq Y_p\} = 1$, and*

$$\mathrm{E}[|V_p - X_p|] = \mathrm{E}[V_p - X_p] \leq \mathrm{E}[Y_p - X_p] \leq pD(\alpha). \tag{34}$$

To prove Theorem 2 (with $C = 0$) we need to strengthen (34). In the next section we will prove that, in the situation of Theorem 5, for every $k \in \{1, 2, \dots\}$ there exists a $D_k(\alpha) < \infty$ such that

$$\mathrm{E}[|Y_p - X_p|^k] = \mathrm{E}[(Y_p - X_p)^k] \leq p^k D_k(\alpha) \quad \text{for } 0 < p \leq \min(p(k), p(k|\alpha|)). \tag{35}$$

In the case $0 < \alpha < 1$ we have a similar result.

Theorem 6. *If $0 \leq \alpha < 1$ then*

$$\mathrm{E}[|X_p - V_p|] = \mathrm{E}[X_p - V_p] \leq \frac{p\alpha|\log c|}{(1-c)(1-\alpha)},$$

and a similar result can easily be derived for $\mathrm{E}[|Z - V_p|] = \mathrm{E}[Z - V_p]$.

The proof starts the same way as the proof of Theorem 5 (give V_p the stationary distribution) but then uses a different approach; see Section 11. The result for higher moments that completes the proof of Theorem 3 will be given in Section 12.

https://doi.org/10.1017/S0001867800001488 Published online by Cambridge University Press

10. Higher moments for $\alpha \le 0$

In this section we prove the 'higher moments' version of Theorem 3. For every $k \in \{1, 2, \dots\}$, we choose a function $c_k(\alpha) \ge c(\alpha)$ and define

$$f_{k,p,\alpha}(x) = \begin{cases} \dfrac{c_k(\alpha)p^2}{x} & \text{if } \dfrac{c^{k-1}p}{1+|\alpha|} \le x < \infty, \\ \dfrac{d(\alpha)p^{1+|\alpha|}}{x^{|\alpha|}} & \text{if } 0 < x < \dfrac{c^{k-1}p}{1+|\alpha|}. \end{cases}$$

(Note that $d(\alpha)$ need not change.) The only constraint on $c_k(\alpha)$ is that

$$f_{k,p,\alpha}(x) \ge f_{p,\alpha}(x) \ge R_{p,\alpha}(x) \quad \text{for all } x > 0.$$

Clearly, it is easy to find such functions $c_k(\alpha)$. As long as it does not lead to confusion, we write $f_k(\cdot)$ for $f_{k,p,\alpha}(\cdot)$.

Next we define the processes $(Y_{k,p,n})_{n=0}^{\infty}$ by

$$Y_{k,p,n+1} = \begin{cases} Y_{k,p,n} + p + f_k(X_{p,n}) & \text{if } \chi_{p,n} = \text{success}, \\ cY_{k,p,n} & \text{if } \chi_{p,n} = \text{failure}, \end{cases}$$

with $X_{p,0} \le Y_{p,0} \le Y_{k,p,0}$. Clearly, then $X_{p,n} \le Y_{p,n} \le Y_{k,p,n}$ for all $n \ge 0$. We will prove that there exist functions $D_k(\alpha)$, $0 < D_k(\alpha) < \infty$ such that, for $0 < p \le \min(p(k), p(k|\alpha|))$, the stationary distributions satisfy

$$\mathrm{E}[(Y_p - X_p)^k] \le \mathrm{E}[(Y_{k,p} - X_p)^k] \le p^k D_k(\alpha).$$

For $\mathrm{E}[f_k(X_p)^k]$ we have

$$\begin{aligned} \mathrm{E}[f_k(X_p)^k] = c_k(\alpha)^k p^{2k} \,\mathrm{E}\Big[\frac{1}{(X_p)^k} \mathbf{1}\Big(X_p \ge \frac{c^{k-1}p}{1+|\alpha|}\Big)\Big] \\ + d(\alpha)^k p^{k(1+|\alpha|)} \,\mathrm{E}\Big[\frac{1}{X_p^{k|\alpha|}} \mathbf{1}\Big(X_p < \frac{c^{k-1}p}{1+|\alpha|}\Big)\Big]. \end{aligned}$$

For $\alpha \le -1$, this result is good enough for our purposes, and it is in fact unnecessary to introduce the new processes $(Y_{k,p,n})_{n=0}^{\infty}$. For $-1 < \alpha < 0$, another step is needed. By repeated use of (33), we have

$$\begin{aligned} \mathrm{E}[f_k(X_p)^k] = p^{2k}\Big(c_k(\alpha)^k \,\mathrm{E}\Big[\frac{1}{(X_p)^k} \mathbf{1}\Big(X_p \ge \frac{c^{k-1}p}{1+|\alpha|}\Big)\Big] \\ + \Big(\frac{p}{c^k}\Big)^{k|\alpha|} d(\alpha)^k \,\mathrm{E}\Big[\frac{1}{X_p^{k|\alpha|}} \mathbf{1}\Big(X_p < \frac{p}{c(1+|\alpha|)}\Big)\Big]\Big). \end{aligned}$$

In other words, there is a function $h_k(\alpha) < \infty$ for which

$$\mathrm{E}[f_k(X_p)^k] < p^{2k} h_k(\alpha)^k \tag{36}$$

if $0 < p \le \min(p(k), p(k|\alpha|))$.

Now, we have

$$Y_{k,p,n+1} - X_{p,n+1} = \begin{cases} Y_{k,p,n} - X_{p,n} + f_k(X_{p,n}) & \text{if } \chi_{p,n} = \text{success}, \\ c(Y_{k,p,n} - X_{p,n}) & \text{if } \chi_{p,n} = \text{failure}. \end{cases} \tag{37}$$

Therefore, for the stationary distributions, we have

$$(1 - pc^k)\,\mathrm{E}[(Y_{k,p} - X_p)^k] = (1-p)\,\mathrm{E}[(Y_{k,p} - X_p + f_k(X_p))^k].$$

Using (36) and the Minkowski inequality (see, e.g. [10, p. 156]), this gives

$$(1 - pc^k)\,\mathrm{E}[(Y_{k,p} - X_p)^k] \le (1-p)(\mathrm{E}[(Y_{k,p} - X_p)^k]^{1/k} + p^2 h_k(\alpha))^k$$

if $p > 0$ is sufficiently close to 0. This implies that

$$\left(\left(\frac{1-pc^k}{1-p}\right)^{1/k} - 1\right)\mathrm{E}[(Y_{k,p} - X_p)^k]^{1/k} \le p^2 h_k(\alpha)$$

for sufficiently small p. This immediately proves that

$$\mathrm{E}[(Y_{k,p} - X_p)^k]^{1/k} < \frac{pk}{1-c^k} h_k(\alpha)$$

if $p > 0$ is sufficiently close to 0. This completes the proof of (35).

11. The case $0 \le \alpha < 1$: linear programming and duality

Throughout this section we have $0 \le \alpha < 1$, so (30) holds:

$$\mathrm{E}[X_p - V_p] = \mathrm{E}[|X_p - V_p|] = \frac{1-p}{p(1-c)}\,\mathrm{E}[|R(V_p)|].$$

We also have the constraint (24):

$$\mathrm{E}\left[\log\left(1 + \frac{p}{(1-\alpha)V_p}\right)\right] = \frac{p|\log c|}{(1-p)(1-\alpha)}. \tag{38}$$

Thus, we can obtain an upper bound for $\mathrm{E}[|X_p - V_p|]$ by solving the following linear programming problem: find the supremum, *Sup*, of

$$\mathrm{E}[|R(V)|] = \mathrm{E}\left[V + p - V\left(1 + \frac{p}{(1-\alpha)V}\right)^{1-\alpha}\right],$$

taken over all nonnegative random variables V for which (38) holds.

In order to use notation similar to that in, say, [4], we write

$$B = \frac{p|\log c|}{(1-p)(1-\alpha)}.$$

The dual (see, e.g. [4, Chapter 6]) of the linear programming problem above is to find the infimum, *Inf*, of

$$\mu_1 + \mu_2 B,$$

taken over all multipliers μ_1 and μ_2 for which, for all $v,\ 0 < v < \infty$,

$$\mu_1 + \mu_2 \log\left(1 + \frac{p}{(1-\alpha)v}\right) \geq v + p - v\left(1 + \frac{p}{(1-\alpha)v}\right)^{1-\alpha}. \tag{39}$$

The multipliers μ_1 and μ_2 are allowed to take on any value (negative, zero, or positive), and μ_1 is the multiplier for the constraint $\mathrm{P}\{0 < V < \infty\} = 1$ and μ_2 is the multiplier for the constraint (38).

Readers who do not like linear programming problems with continuously many primal variables and continuously many dual constraints can take the appropriate limit of a problem with finitely many primal variables and finitely many dual constraints. (Allow the random variable V to have support only in a finite number of points v_k; $\pi_k = \mathrm{P}\{V = v_k\}$ then becomes the primal variable.)

By setting $x = p/[(1-\alpha)v]$, we rewrite the constraint (39) as follows: for all $0 < x < \infty$,

$$\mu_1 + \mu_2 \log(1+x) \geq \frac{p}{(1-\alpha)x} + p - \frac{p}{(1-\alpha)x}(1+x)^{1-\alpha}. \tag{40}$$

We obtain a dual feasible solution by setting $\mu_1 = 0$ and $\mu_2 = p\alpha$. To prove that, for these values of μ_1 and μ_2, (40) indeed holds for all $x,\ 0 < x < \infty$, takes straightforward arithmetic.

Thus, we know that

$$\mathrm{E}[|R(V_p)|] \leq Sup = Inf \leq \mu_2 B = \frac{p^2\alpha|\log c|}{(1-p)(1-\alpha)}$$

or

$$\mathrm{E}[|X_p - V_p|] \leq \frac{p\alpha|\log c|}{(1-c)(1-\alpha)}.$$

12. The linear programming approach with higher moments

We would of course also like to use the linear programming approach of the previous section for higher moments of $X_p - V_p$ in the situation $0 \leq \alpha < 1$. This attempt does lead to a result, but one weaker than Theorem 2.

For the νth moment, $\nu > 1$, we first want to find the supremum of

$$\mathrm{E}\left[\left(V + p - V\left(1 + \frac{p}{(1-\alpha)V}\right)^{1-\alpha}\right)^{\nu}\right]$$

subject to the constraint that V is a nonnegative random variable with

$$\mathrm{E}\left[\log\left(1 + \frac{p}{(1-\alpha)V}\right)\right] = \frac{p|\log c|}{(1-p)(1-\alpha)}. \tag{41}$$

As before, the right-hand side of (41) is denoted by B. The dual of this linear programming problem is to find the infimum of

$$\mu_1 + \mu_2 B$$

subject to the constraint

$$\mu_1 + \mu_2 \log\left(1 + \frac{p}{(1-\alpha)u}\right) \geq \left(u + p - u\left(1 + \frac{p}{(1-\alpha)u}\right)^{1-\alpha}\right)^{\nu}$$

for all $0 < u < \infty$.

https://doi.org/10.1017/S0001867800001488 Published online by Cambridge University Press

To construct a dual feasible solution, we choose $\mu_1 = 0$ and

$$\mu_2 = \sup_{0<u<\infty} \frac{(u + p - u(1 + p/[(1-\alpha)u])^{1-\alpha})^\nu}{\log(1 + p/[(1-\alpha)u])}. \tag{42}$$

As $u \downarrow 0$, the right-hand side of (42) goes to 0. To study other values of u, we make the substitutions $p/[(1-\alpha)u] = x$ and $u = p/[(1-\alpha)x]$. Equation (42) then becomes

$$\mu_2 = \frac{p^\nu}{(1-\alpha)^\nu} \sup_{0<x<\infty} \frac{(1 + (1-\alpha)x - (1+x)^{1-\alpha})^\nu}{x^\nu \log(1+x)}.$$

We define

$$M_\nu = \frac{1}{(1-\alpha)^{\nu+1}} \sup_{0<x<\infty} \frac{(1 + (1-\alpha)x - (1+x)^{1-\alpha})^\nu}{x^\nu \log(1+x)}. \tag{43}$$

As $x \downarrow 0$, the right-hand side of (43) behaves like some constant times $x^{\nu-1}$. Hence, the supremum in (43) is a maximum and is positive and finite. It depends on α and ν, but is independent of p and c. We now have

$$\mathrm{E}[|R(V_p)|^\nu] \le \mathit{Sup} = \mathit{Inf} \le \frac{p^{\nu+1}}{1-p} |\log c| M_\nu \tag{44}$$

for all p, $0 < p < 1$. Unfortunately, the right-hand side of (44) behaves like $p^{\nu+1}$, not like $p^{2\nu}$ as in (36). (In (36), $\nu > 1$ had to be an integer $\nu = k \ge 2$). Proceeding as in the argument immediately following (37) (where $\nu = k \ge 2$ had to be an integer) we now only find that if $0 < \alpha < 1$ then, for every integer $k \ge 2$ and every p, $0 < p < 1$,

$$\mathrm{E}\bigg[\frac{(X_p - V_p)^k}{p}\bigg] = \mathrm{E}\bigg[\frac{|X_p - V_p|^k}{p}\bigg] < \frac{k^k |\log c| M_k}{(1-p)(1-c^k)^k}.$$

The $C = 0$ case of Theorem 3 now easily follows.

13. Results for $C > 0$

In most of the previous sections we studied processes $V_{p,n}$, i.e. where $\alpha < \beta = 1$, $0 < c_2 < 1$, and $C = 0$. In this section we deal with the case in which $C > 0$ and either $p \downarrow 0$ while $C > 0$ remains constant or $C \downarrow 0$ while $p > 0$ remains constant. We define the process $(X^*_{p,C,n})_{n=0}^\infty$, similar to $V^*_{p,C,n}$ in (9), by

$$X^*_{p,C,n+1} = \begin{cases} X^*_{p,C,n} + p & \text{if } \chi_{p,n} = \text{success}, \\ \max\bigg(cX^*_{p,C,n}, \dfrac{pC^{1-\alpha}}{(1-\alpha)c_1}\bigg) & \text{if } \chi_{p,n} = \text{failure}. \end{cases}$$

The key result that makes the case where $C > 0$ is constant easy to handle is our next lemma.

Lemma 11. *If, for some n,*

$$X_{p,n} \le \max\bigg(X_{p,n}, \frac{pC^{1-\alpha}}{(1-\alpha)c_1}\bigg) \le X^*_{p,C,n} \le X_{p,n} + \frac{pC^{1-\alpha}}{(1-\alpha)c_1},$$

then this holds for all $k \ge n$.

The proof is straightforward and is left as an exercise for the reader.

Among other things, this proves that, while $\mathrm{E}[X_p^\nu] = \infty$ for sufficiently large ν, if $C > 0$ then $\mathrm{E}[(X^*_{p,C})] < \infty$ for all values of ν.

We have the following result similar to Lemma 1.

Lemma 12. *If $\alpha \le 0$ and $V^*_{p,C,k} \ge X^*_{p,C,k}$ for some $k \ge 0$, then this holds for all $n \ge k$. If $0 \le \alpha < 1$ and $V^*_{p,C,k} \le X^*_{p,C,k}$ for some $k \ge 0$, then this holds for all $n \ge k$. If $\alpha = 0$ and $V^*_{p,C,k} = X^*_{p,C,k}$ for some $k \ge 0$, then this holds for all $n \ge k$.*

Proof. The proof makes use of rewriting (9) as

$$V^*_{p,C,n+1} = \begin{cases} V^*_{p,C,n} + p + R_{p,\alpha}(V^*_{p,C,n}) & \text{if } \chi_{p,n} = \text{success}, \\ \max\left(cV^*_{p,C,n}, \dfrac{pC^{1-\alpha}}{(1-\alpha)c_1}\right) & \text{if } \chi_{p,n} = \text{failure}, \end{cases}$$

with $R_{p,\alpha}(\cdot)$ as in (27). The remainder of the proof is left to the reader.

In the case $0 \le \alpha < 1$, we are now finished. In this case, $R(\cdot)$ is an increasing function, so the joint distribution of V_p, $V^*_{p,C}$, $X^*_{p,C}$, and X_p is such that

$$V_p \le V^*_{p,C} \le X^*_{p,C} \le X_p + \frac{pC^{1-\alpha}}{(1-\alpha)c_1}, \tag{45}$$

and we know that, while $V_p \le X_p$, they are close in the sense of Theorem 6, or in the sense of the material in Section 12 if that approach is preferred.

If $0 \le \alpha < 1$ and $C \downarrow 0$ while $p > 0$ is constant, we see that the distribution of $V^*_{p,C}$ weakly converges (and stochastically decreases) to the distribution of V_p. This follows from the stochastic monotonicity observations made in Section 2. We have been unable to prove a similar weak convergence as $C \downarrow 0$ while $p > 0$ is constant and $\alpha < 0$.

In the case $\alpha < 0$, we define the processes $(\Lambda^*_{p,C,n})_{n=0}^\infty$ and $(\Lambda^{**}_{p,C,n})_{n=0}^\infty$ by

$$\Lambda^*_{p,C,n+1} = \begin{cases} \Lambda^*_{p,C,n} + p + R_{p,\alpha}(X^*_{p,C,n}) & \text{if } \chi_{p,n} = \text{success}, \\ \max\left(c\Lambda^*_{p,C,n}, \dfrac{pC^{1-\alpha}}{(1-\alpha)c_1}\right) & \text{if } \chi_{p,n} = \text{failure}, \end{cases}$$

$$\Lambda^{**}_{p,C,n+1} = \begin{cases} \Lambda^{**}_{p,C,n} + p + R_{p,\alpha}(X^*_{p,C,n}) & \text{if } \chi_{p,n} = \text{success}, \\ c\Lambda^{**}_{p,C,n} & \text{if } \chi_{p,n} = \text{failure}. \end{cases}$$

Henceforth, we choose the initial values $X_{p,0}$, $V_{p,0}$, $X^*_{p,C,0}$, $V^*_{p,C,0}$, $\Lambda_{p,0}$, $\Lambda^*_{p,C,0}$, and $\Lambda^{**}_{p,C,0}$ such that the following inequalities hold for all n, instead of for values of n greater than or equal to some k. Since $X^*_{p,n} \le V^*_{p,n}$ and $R(\cdot)$ is a positive, decreasing function,

$$X^*_{p,C,n} \le V^*_{p,C,n} \le \Lambda^*_{p,C,n}$$

for all $n \ge 0$. Similarly, since $X^*_{p,C,n} \ge X_{p,n}$, we also have

$$\Lambda^{**}_{p,C,n} \le \Lambda_{p,n}$$

https://doi.org/10.1017/S0001867800001488 Published online by Cambridge University Press

for all $n \ge 0$. By the same argument as (45), we have

$$\Lambda^{**}_{p,C,n} \le \Lambda^{*}_{p,C,n} \le \Lambda^{**}_{p,C,n} + \frac{pC^{1-\alpha}}{(1-\alpha)c_1}$$

for all $n \ge 0$. For the joint stationary distribution we therefore have

$$X_p \le X^*_{p,C} \le V^*_{p,C} \le \Lambda^*_{p,C} \le \Lambda^{**}_{p,C} + \frac{pC^{1-\alpha}}{(1-\alpha)c_1} \le \Lambda_p + \frac{pC^{1-\alpha}}{(1-\alpha)c_1}.$$

Since we know that $\Lambda_p - X_p$ is nonnegative and small in the sense of Theorem 5, we have proven the desired result for first moments. For higher moments we do not need to call on functions similar to $f_{k,p,\alpha}(\cdot)$, but can directly call on results for the higher moments of $\Lambda_p - X_p$.

The approach above shows that not only does $X^*_{p,C}$ have all moments $\mathrm{E}[(X^*_{p,C})^{\mu}] < \infty$, $-\infty < \nu < \infty$, but so also do $V^*_{p,C}$, $\Lambda^*_{p,C}$, and $\Lambda^{**}_{p,C}$:

$$R_{p,\alpha}(X^*_{p,C}) \le R_{p,\alpha}\bigg(\frac{pC^{1-\alpha}}{(1-\alpha)c_1}\bigg) < \infty,$$

followed by the same argument as in the proof of Lemma 4. We see that if $\alpha < 0$ and $0 < p \le \min(p(k), p(k|\alpha|))$, then, as $C \downarrow 0$ with p constant, the moments $\mathrm{E}[(V^*_{p,C})^{\nu}]$, $0 \le \nu \le k$, remain bounded between $\mathrm{E}[X^{\nu}_p]$ and $\mathrm{E}[(\Lambda_p + pC^{1-\alpha}/[(1-\alpha)c_1])^{\nu}] < \infty$. There is no guarantee that these moments will converge as $C \downarrow 0$. It is conceivable that the set of limit points is some nontrivial subset of $[\mathrm{E}[X^{\nu}_p], \mathrm{E}[\Lambda^{\nu}_p]]$.

14. The special case $\alpha = -1$

The case $\alpha = -1$ is special for several reasons. It is of particular interest because it represents the 'classical TCP' situation. It also admits a significantly simplified proof of the material in Sections 7 and 8, and thereby of Theorems 2 and 5 and, of course, Lemmas 4 and 5.

If $\alpha = -1$ then $R(v) = p^2/(4v)$, which is already of the right form, so we choose $f(v) = R(v)$ so that $Y_p = \Lambda_p$ as in Section 7. Hence,

$$0 < \mathrm{E}[\Lambda_p - X_p] = \frac{p(1-p)}{4(1-c)}\,\mathrm{E}\bigg[\frac{1}{X_p}\bigg],$$

where of course $c = b^{1-\alpha} = b^2$. From Lemma 6 (which still is needed) and Corollary 1, we know that

$$\lim_{p\downarrow 0} \mathrm{E}\bigg[\frac{1}{X_p}\bigg] = \mathrm{E}\bigg[\frac{1}{Z}\bigg] = \log\bigg(\frac{1}{c}\bigg) = 2\log\bigg(\frac{1}{b}\bigg).$$

In the case of TCP, $b = \frac{1}{2}$.

Acknowledgement

I thank Jason Swanson for a stimulating exchange of emails on the topics of this paper as well as [13], and among other things for his observation that the uniqueness of the stationary distribution of V_p, $0 < p < 1$, had not been proven.

References

[1] Altman, E., Avrachenkov, K. and Barakat, C. (2005). A stochastic model of TCP/IP with stationary random losses. *IEEE/ACM Trans. Networking* **13,** 356–369.

[2] Altman, E., Avrachenkov, K., Barakat, C. and Nunez-Queija, R. (2001). State-dependent M/G/1 type queueing analysis for congestion control in data networks. In *Proc. INFOCOM 2001* (Anchorage, AK, April 2001), Vol. 3, pp. 1350–1359.

[3] Altman, E. *et al.* (2005). Analysis of MIMD congestion control algorithm for high speed networks. *Comput. Networks* **48,** 972–989.

[4] Dantzig, G. B. (1963). *Linear Programming and Extensions*. Princeton University Press.

[5] Dumas, V., Guillemin, F. and Robert, P. (2002). A Markovian analysis of additive-increase, multiplicative-decrease algorithms. *Adv. Appl. Prob.* **34,** 85–111.

[6] Guillemin, F., Robert, P. and Zwart, B. (2004). AIMD algorithms and exponential functionals. *Ann. Appl. Prob.* **14,** 90–117.

[7] Kelly, T. (2003). Scalable TCP: improving performance in highspeed wide area networks. *Comput. Commun. Rev.* **33,** 83–91.

[8] Kelly, T. (2004). Engineering flow controls for the Internet. Doctoral Thesis, University of Cambridge. Available at http://www.deneholme.net/tom/papers.

[9] Lehmann, E. L. (1986). *Testing Statistical Hypotheses*, 2nd edn. John Wiley, New York.

[10] Loève, M. (1963). *Probability Theory*, 3rd edn. Van Nostrand, Princeton, NJ.

[11] Ott, T. J. (2005). Transport protocols in the TCP paradigm and their performance. *Telecommun. Systems* **30,** 351–385. Available at http://www.teunisott.com/Papers.

[12] Ott, T. J. (2006). Rate of convergence for the 'square root formula', extended version. Unpublished manuscript. Available at http://www.teunisott.com/Papers.

[13] Ott, T. J. and Swanson, J. (2006). Asymptotic behavior of a generalized TCP congestion avoidance algorithm. Submitted. Available at http://www.teunisott.com/Papers.

[14] Ott, T. J. and Swanson, J. (2006). Stationarity of some processes in transport protocols. To appear in *Performance Eval. Rev.* Extended version available at http://www.teunisott.com/Papers.

[15] Ott, T. J., Kemperman, J. H. B. and Mathis, M. (1996). The stationary behavior of ideal TCP congestion avoidance. Preprint. Available at http://www.teunisott.com/Papers.

[16] Ott, T. J., Kemperman, J. H. B. and Mathis, M. (2004). The stationary behavior of ideal TCP congestion avoidance, revisited. Unpublished manuscript.

[17] Ramakrishnan, K., Floyd, S. and Black, D. (2001). The addition of explicit congestion notification (ECN) to IP. Internet Engineering Task Force RFC 3168. Available at ftp://ftp.rfc-editor.org/in-notes/rfc3168.txt.