Rate of convergence for the "Square Root Formula" in TCP, Extended Version

Teunis J. Ott

September 3, 2006

Abstract

The "square root formula" for TCP (Traffic Control Protocol) in IP states that if the probability p of packet loss becomes small, and there is independence between packets, the stationary distribution of the congestion window W becomes such that the distribution of $W\sqrt{p}$ becomes 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 Tom Kelly's "Scalable TCP". This paper also gives stochastic dominance results that translate to a rate of convergence.

1 Introduction

The paper [3] proposes a class of "TCP–like" Internet Transport Protocols and uses a class of stochastic processes to analyze the performance of these protocols. That class of stochastic processes is defined by:

Let $(U_n)_{n=0}^{\infty}$ be independent, identically distributed random variables, each distributed uniformly $[0, 1]$. Let p be a probability, $0 < p < 1$. Define the i.i.d. random variables $\chi_{p,n}$ by

$$\chi_{p,n} = \begin{cases} \text{success} & \text{if } U_n \geq p \\ \text{failure} & \text{if } U_n < p \end{cases} \tag{1.1}$$

Further, let the discrete time, continuous state space process $W^*_{p,C,n}$ ($n = 0, 1, 2, \cdots$, $0 < W^*_{p,C,n} < \infty, \quad 0 < p < 1$) be defined by

$$W^*_{p,C,n+1} = \begin{cases} W^*_{p,C,n} + c_1 (W^*_{p,C,n})^{\alpha} & \text{if } \chi_{p,n} = \text{success}, \\ \max\left(W^*_{p,C,n} - c_2 (W^*_{p,C,n})^{\beta}, C\right) & \text{if } \chi_{p,n} = \text{failure}, \end{cases} \tag{1.2}$$

where $\alpha < \beta \leq 1,\ c_1 > 0,\ c_2 > 0,\ C > 0$.

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 $\beta = 1,\ \alpha = 0$ models Tom Kelly's "Scalable TCP", see [8, 9].

The paper [3] shows that the more general case, even the case $0 < \alpha < \beta \leq 1$, is of interest in the study of transport protocols.

In [4] it is 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 that 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 $\alpha < \beta = 1,\ c_1 > 0,\ 0 < c_2 < 1$ and we will write $1 - c_2 = b$. In this case we will see that we can drop the "max (..., C)" in (1.2) (or choose $C = 0$). We will be mostly interested in the case $C = 0$, but after the case $C = 0$ has been studied we will observe consequences for the process with $C > 0$.

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

$$W_{p,n+1} = \begin{cases} W_{p,n} + c_1(W_{p,n})^\alpha & \text{if } \chi_{p,n} = \text{success}, \\ W_{p,n} - (1-b)W_{p,n} = bW_{p,n} & \text{if } \chi_{p,n} = \text{failure}. \end{cases} \tag{1.3}$$

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

$$W^*_{p,n+1} = \begin{cases} W^*_{p,n} + c_1(W^*_{p,n})^\alpha & \text{if } \chi_{p,n} = \text{success}, \\ \max(bW^*_{p,C,n}, C) & \text{if } \chi_{p,n} = \text{failure}. \end{cases} \tag{1.4}$$

Not surprising, for $p \downarrow 0$ the two processes $W_{p,n}$ and $W^*_{p,n}$ become "very similar", in a way that will be explained in Section 16.

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

$$0 < b^n W_{p,0} \leq W_{p,n} < \infty \text{ for all } n \geq 0. \tag{1.5}$$

If $\alpha = -1,\ c_1 = 1,\ b = \frac{1}{2}$ (1.3) models the feedback process for the congestion window in TCP, see e.g. [2] and [1]. In the TCP environment, "success" stands for arrival of a "good acknowledgment" (one which positively acknowledges safe arrival of new and contiguous data), whereas "failure" stands for loss of a data packet. For certain values of α, c_1, and b the process in (1.3) is a candidate for similar control mechanisms. Of the papers just cited, [2] has existed on the Internet since 1996 and is frequently cited, but has never been published in the open literature. [1] is a re-write (draft) of [2] to be submitted soon. [13, 14] give more references to literature on this topic.

This paper, and the ones just mentioned, use "packet–time": progress of time is (essentially) measured by the number of "good acknowledgments" that have been received. Many other papers studying TCP performance use "clock–time" where (apart from during slowstart and fast recovery) time is measured in RTTs (Round Trip Times) or periods of time during which W "good" acknowledgments are received, usually under the assumption that during that time at most one packet gets lost or marked. If $\alpha = -1$ the window increases by almost exactly c_1 MSSs during such a period. The assumption that at most one reduction of the congestion window occurs during one RTT is reasonable if $\alpha < 0$ but questionable if $0 \leq \alpha$, see [3]. The paper [2] contains a translation between "packet time stationary" and "clock time stationary" distributions.

The somewhat overly complicated construction (1.1) and (1.3) 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

$$\zeta_p(t) = p^{\frac{1}{1-\alpha}} W_{p,\lfloor \frac{t}{p} \rfloor} \tag{1.6}$$

we see that as long as there is "success" we have

$$\frac{\zeta_p(t+p) - \zeta_p(t)}{p} = c_1(\zeta_p(t))^\alpha. \tag{1.7}$$

Hence, when $p \downarrow 0$ we approach the situation where there is a Poisson Process with intensity 1, with "events" $\cdots, \tau_{-1}, \tau_0, \tau_1, \cdots$, with a process $\zeta(t)$ defined by

$$\frac{d}{dt}\zeta(t) = c_1(\zeta(t))^\alpha \tag{1.8}$$

"in between" the events of the Poisson Process, and

$$\zeta(\tau^+) = b\zeta(\tau^-) \tag{1.9}$$

"in" the points of the Poisson process.

If we now define

$$Z(t) = \frac{(\zeta(t))^{1-\alpha}}{(1-\alpha)c_1} \tag{1.10}$$

then

$$\frac{d}{dt}Z(t) = 1 \tag{1.11}$$

“in between” the events of the Poisson Process, and

$$Z(\tau^+) = b^{1-\alpha} Z(\tau^-) \tag{1.12}$$

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

As in [2] or [1] we see that the process $Z(.)$ has as stationary distribution the distribution of

$$Z = \sum_{k=0}^{\infty} c^k E_k, \tag{1.13}$$

where $(E_k)_{k=0}^{\infty}$ are independent, identically distributed random variables, each exponentially distributed with parameter 1. The distribution of Z in (1.13) is completely described in [2] and in [1]. For example, we have that for all (even complex) ν

$$E[Z^\nu] = \Gamma(\nu+1) \prod_{k=0}^{\infty} \frac{1-c^{\nu+k}}{1-c^k}. \tag{1.14}$$

If ν is integer this reduces to (k is a positive integer):

$$\begin{aligned} E[Z^k] &= \frac{k!}{(1-c)(1-c^2)\cdots(1-c^k)}, \\ E[Z^{-k}] &= \frac{(1-c)(1-c^2)\cdots(1-c^{k-1})}{(k-1)!c^{\frac{1}{2}(k-1)k}} \log\left(\frac{1}{c}\right). \end{aligned} \tag{1.15}$$

Analogous to (1.6) we define, for $C > 0$,

$$\zeta^*_{p,C}(t) = p^{\frac{1}{1-\alpha}} W^*_{p,C,\lfloor \frac{t}{p} \rfloor}. \tag{1.16}$$

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

$$V_{p,n} = \frac{p}{(1-\alpha)c_1} (W_{p,n})^{1-\alpha} = \frac{(\zeta_p(pn))^{1-\alpha}}{c_1(1-\alpha)}. \tag{1.17}$$

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

$$V^*_{p,C,n} = \frac{p}{(1-\alpha)c_1} \left(W^*_{p,C,n}\right)^{1-\alpha} = \frac{(\zeta^*_{p,C}(pn))^{1-\alpha}}{c_1(1-\alpha)}. \tag{1.18}$$

These processes evolve as

$$V_{p,n+1} = \begin{cases} V_{p,n}\left(1+\frac{p}{(1-\alpha)V_{p,n}}\right)^{1-\alpha} & \text{if } \chi_{p,n} = \text{success}, \\ cV_{p,n} & \text{if } \chi_{p,n} = \text{failure}, \end{cases} \tag{1.19}$$

and

$$V^*_{p,C,n+1} = \begin{cases} V^*_{p,C,n}\left(1+\frac{p}{(1-\alpha)V^*_{p,C,n}}\right)^{1-\alpha} & \text{if } \chi_{p,n} = \text{success}, \\ \max(cV^*_{p,C,n}, \frac{pC^{1-\alpha}}{(1-\alpha)c_1}) & \text{if } \chi_{p,n} = \text{failure}. \end{cases} \tag{1.20}$$

We will first study the stationary distributions of the process $V_{p,n}$ and then find a simple way of translating 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 [4] it is proven that if $C > 0$ the stationary distribution of $V^*_{p,C,n}$ exists and is unique.

In [5] it is 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 on this topic later.

In [4] it is proven that for all $C \geq 0$ (constant), if $p \downarrow 0$, the process

$$\frac{(\zeta^*_{p,C}(t))^{1-\alpha}}{(1-\alpha)c_1} \tag{1.21}$$

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

If $0 \leq \alpha < 1$ the stationary distribution of $V_{p,n}$ ($C = 0$) is unique, see the next section. A technical problem is that in the case $\alpha < 0$ we have not proven (yet?) 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 (1.13) we really mean that whatever way we choose the stationary distributions of $V_{p,n}$, for $p \downarrow 0$ they converge to the distribution in (1.13). In the remainder of this 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 11 contains more of a discussion of stationary distributions.

Processes as studied in this note have also been studied in e.g. [13], [14], which among other results prove weak convergence of stationary distributions, for the "clock time process", in the case $\alpha = -1$ (the TCP situation). [13, 14] use a different proof technique and do not have stochastic dominance results or rate of convergence results.

The paper [3] studies issues leading up to stability of feedback protocols (through consideration of relaxation times), both in the situation of this paper ($\beta = 1$) and in the case ($\beta < 1$).

Among the many other papers doing mathematical analysis of the performance of TCP we only mention [10], [11], and in particular [12], which (in "clock time") studies the case of scalable TCP ($\alpha = 0$). The last six papers cited contain an extensive review of the literature.

Among the results obtained in this paper are the Theorems below:

Theorem 1:

Let V_p have a stationary distribution of the process $(V_{p,n})$, and let for some $C > 0$ $V^*_{p,C}$ have the stationary distribution of the process $(V^*_{p,C,n})$. Then for $p \downarrow 0$ 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 $V_{p,n}$, i.e. the case $C = 0$. Once results for that case are available there are corollaries for the case $C > 0$, i.e. 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 $\alpha \leq 0$ and $0 \leq \alpha < 1$:

Theorem 2:

Let $\alpha \leq 0$. Let for some $C \geq 0$ $V^*_{p,C}$ have the stationary distribution of $(V^*_{p,C,n})$. Then for every $-\infty < \nu < +\infty$:

$$\lim_{p\downarrow 0} E[(V^*_{p,C})^\nu] = E[Z^\nu], \tag{1.22}$$

and the joint distributions of $(V^*_{p,C}, Z)$ can be chosen such, that for every $\nu \geq 0$

$$\limsup_{p\downarrow 0} E\left[\left|\frac{Z - V^*_{p,C}}{p}\right|^\nu\right] < \infty, \tag{1.23}$$

In other words, every sequence of positive probabilities p converging to zero has a subsequence p_k for which $Err_{p_k} = \frac{Z - V_{p_k}}{p_k}$ converges weakly to a random variable Err. That random variable Err then has all moments $E[|Err|^\nu]$ $(\nu \geq 0)$ finite, and for all $\nu \geq 0$ $E[|Err_{p_k}|^\nu]$ converges to $E[|Err|^\nu]$, with a similar result for $V^*_{p,C}$.

Theorem 3:

Let $0 \leq \alpha < 1$. Let for some $C \geq 0$ $V^*_{p,C}$ have the stationary distribution of $(V^*_{p,C,n})$. Then for every $0 \leq \nu < +\infty$:

$$\lim_{p\downarrow 0} E[(V^*_{p,C})^\nu] = E[Z^\nu], \tag{1.24}$$

and the joint distribution of $(V_p, V^*_{p,C}, Z)$ can be chosen such, that for every integer $k \geq 1$

$$\limsup_{p\downarrow 0} E\left[\frac{|Z - V^*_{p,C}|^k}{p}\right] < \infty, \tag{1.25}$$

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

$$P\left\{V_p \leq \frac{Z}{\frac{1}{p}|\log(1-p)|} < Z\right\} = 1, \tag{1.26}$$

The Theorems above will first be proven for the process $V_{p,n}$. Section 16 contains mechanisms to translate results for the process $V_{p,n}$ into results for the process $V^*_{p,C,n}$.

For the proofs for the process $V_{p,n}$ we introduce a number of auxiliary stochastic processes that will provide the desired stochastic bounds and thus prove results stronger than the Theorems above.

Sections 4 and 8 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 zero). We also can consider the problem of what happens if $p > 0$ is constant and $C \downarrow 0$. Clearly, in that case (with $V_{p,0} = V^*_{p,C,0}$)

$$\lim_{C \downarrow 0} V^*_{p,C,n} = V_{p,n}. \tag{1.27}$$

However, convergence of processes does not always guarantee convergence of stationary distributions. If $0 \leq \alpha < 1$ we do have uniqueness of the stationary distribution of $V_{p,n}$ (see the end of this introduction) and weak convergence of the stationary distributions of $V^*_{p,C,n}$ to the stationary distribution of $V_{p,n}$, see Section 3.

In the Sections 4 etc we derive and use "stochastic dominance results" where various processes defined on the basis of the same sequence $(\chi_{p,n})_{n=0}^{\infty}$ (all the same p) are compared. Only in Section 3 do we compare processes for different p.

2 Uniqueness of Stationary Distributions

As stated before, in the case with $0 \leq \alpha < 1$ the processes $W_{p,n}$ and $V_{p,n}$ have unique stationary distributions on $(0, \infty)$, even though $C = 0$. It must be noted that as long as $0 < \alpha < 1$ there also is a stationary distribution concentrated in the point $\{0\}$, and of course linear combinations of the two distributions are also stationary.

The existence of a stationary distribution concentrated on $(0, \infty)$ will be proven in the remainder of this paper. In this section we prove the uniqueness.

The uniqueness follows from the observation that if we have two 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 $(W_{p,n,1})_{n=0}^{\infty}$, if $0 \leq \alpha < 0$

and $0 < W_{p,0,2} < W_{p,0,2}$ then $0 < W_{p,n,1} < W_{p,n,2}$ for all n, and $\frac{W_{p,n,1}}{W_{p,n,2}}$ is non–decreasing: remains the same at failures, increases at successes, and converges to one for $n \to \infty$.

Thus, if $0 < \alpha < 1$ the distribution of $W_{p,n}$ becomes independent on $W_{p,0}$ for $n \to \infty$. Once we have proven that there exists a stationary distribution on $(0, \infty)$ we have proven the required result (for $0 \leq \alpha < 1$).

In the case $\alpha = 0$ we can even choose $W_{p,0} = 0$ and still converge to the unique stationary distribution. In this case the stationary distribution of $V_{p,n}$ can be explicitly described, see Section 4.

By [5], if $\alpha < 0$ the processes $W_{p,n}$ and $V_{p,n}$ still have a stationary distribution, but in this case we have been unable top prove uniqueness. While we believe all such stationary distributions to be unique, we can not exclude the possibility that there are combinations of c_1, b, α and p with weird number–theoretic properties for which the set $(0, \infty)$ is divided in non-communicating sets.

These sets might (in theory!) even have periodicities.

A method which we hope could be used to prove uniqueness even with $\alpha < 0$ is described in Section 13.

Simulation can not be used to investigate the (barely) conceivable non-uniqueness. Finite precision would destroy the (barely) conceivable weird number–theoretic properties, and in simulation we would in fact always have a positive (though tiny) $C > 0$.

3 Stochastic dominance and different p and C

With decreasing p there are more successes and fewer failures, so a natural question to ask is whether with decreasing p (and say $W^*_{p,C,0}$ constant) $W^*_{p,C,n}$ is increasing. This seems not to be true in general, but becomes true for many choices of the "minimal value" C in

(1.2). As in the remainder of this paper we have $\beta = 1, \;\; 0 < c_2 < 1$. Some of the results in this Section still hold if $\beta < 1$.

We define:

$$m = \begin{cases} 0 & \text{if } 0 \le \alpha < 1, \\ (c_1|\alpha|)^{\frac{1}{1+|\alpha|}} & \text{if } \alpha \le 0, \end{cases} \tag{3.1}$$

We then have:

Observation: If $C \ge m$ then we have the monotonicity result that if $0 < p_1 \le p_2 < 1$ and $W^*_{p_1,C,0,1} \ge W^*_{p_2,C,0,2}$ (and $(W^*_{p,C,n,i})_{n=0}^{\infty}$ generated from $W^*_{p,C,0,i}$ in the obvious way) then

$$W^*_{p_1,C,n,1} \ge W^*_{p_2,C,n,2} \;\text{ for all }\; n \ge 0. \tag{3.2}$$

Hence we have:

Observation: If $C \ge m$ and $0 < p_1 < p_2 < 1$ then the stationary distribution for p_1 is "stochastically larger" than the stationary distribution for p_2. (See [16] or Section 4 for a discussion of "stochastically larger".)

The requirements that $C \ge m$ may very well be indispensable: If $C < m$ a smaller value of $W^*_{p,C,n}$ now may lead to a larger value of $W^*_{p,C,n+k}$ later on.

It also is clear that (with $W^*_{p,C,0}$ constant) $W^*_{p,C,n}$ is increasing with C for $C \ge m$. Hence, the stationary distribution of $W^*_{p,C,n}$ is stochastically increasing with C if $C \ge m$. If $m = 0$ ($0 \le \alpha < 1$) this implies that the stationary distribution of $W_{p,n}$ is the weak limit of the stationary distributions of $W^*_{p,C,n}$ for $C \downarrow 0$.

In the case $\alpha < 0, \;\; \beta = 1, \;\; 0 < c_2 < 1$ we have that $m > 0$, and for $W_{p,n}$ very small the upward jumps become very large.

Also in this case it "should be possible" to prove uniqueness of the stationary distribution of $W_{p,n}$ by proving that it is the weak limit of the stationary distributions of $W^*_{p,C,n}$ for $C \downarrow 0$ (stochastic monotonicity or not). We have not yet been able to prove this.

4 Method of attack and intermediate results

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

$$X_{p,n+1} = \begin{cases} X_{p,n} + p & \text{if } \chi_{p,n} = \text{success}, \\ cX_{p,n} & \text{if } \chi_{p,n} = \text{failure}. \end{cases} \tag{4.1}$$

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:

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

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

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

This Lemma will be proven in Section 7.

We also have:

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

$$X_p = p \sum_{k=0}^{\infty} c^k G_{p,k}, \tag{4.2}$$

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

$$P\{G_{p,k} = n\} = p(1-p)^n. \tag{4.3}$$

Hence, X_p has Laplace–Stieltjes Transform

$$\phi_{X_p}(s) = E[exp\{-sX_p\}] = \prod_{k=0}^{\infty} \frac{p}{1-(1-p)exp\{-pc^k s\}}, \tag{4.4}$$

and

$$E[X_p] = \frac{1-p}{1-c}, \;\; \mathrm{Var}(X_p) = \frac{1-p}{1-c^2}. \tag{4.5}$$

There are at least two obvious ways to prove Lemma 2. The most intuitive proof duplicates the proof, given in [2], of the similar result (1.13) for the process $(Z(t))$. This proof works by "looking back in time". A less intuitive proof uses the fact that

$$E[exp\{-sX_{p,n}\}] = (1-p)E[exp\{-s(X_{p,n-1}+p)\}] + pE[exp\{-scX_{p,n-1}\}]. \tag{4.6}$$

Assuming that $X_{p,n-1}$ and $X_{p,n}$ have the same distribution leads to a daisy–chain that proves (4.4), and thus Lemma 2.

(4.1) trivially shows that

$$E[X_{p,n}] = \frac{1-p}{1-c} + (1-p(1-c))^n \left(E[X_{p,0}] - \frac{1-p}{1-c} \right), \tag{4.7}$$

which, among other things, proves tightness. In addition, if we have two different initial values $X_{p,0,1}$ and $X_{p,0,2}$ then with the identical sequences of successes and failures we have

$$X_{p,n,1} - X_{p,n,2} = c^{N(n)}(X_{p,0,1} - X_{p,0,2}) \tag{4.8}$$

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

Lemmas 1 and 2 make a start with providing stochastic bounds for stationary distributions of the process $(V_{p,n})_{n=0}^{\infty}$. If $\alpha < 0$ they provide a stochastic lower bound, proving that $V = 0$ is not a critical point. If $0 < \alpha < 1$ 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 if $p \downarrow 0$.

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 (1.13). It is clear that that random variable Z has Laplace–Stieltjes Transform

$$\phi_Z(s) = E[exp\{-sZ\}] = \prod_{k=0}^{\infty} \frac{1}{1+c^k s}, \tag{4.9}$$

and

$$E[Z] = \frac{1}{1-c}, \quad \mathrm{Var}(Z) = \frac{1}{1-c^2}. \tag{4.10}$$

It is immediately obvious that for $p \downarrow 0$ the Laplace–Stieltjes transform (4.4) converges to the Laplace–Stieltjes transform (4.9). Thus, the distribution of X_p converges weakly to the distribution of Z. However, more can be said:

Lemma 3: The random variables X_p in (4.2) and Z in (1.13) can be given a joint distribution for which with probability one

$$\max(0, \frac{Z}{\frac{1}{p}\log\frac{1}{1-p}} - \frac{p}{1-c}) < X_p \le \frac{Z}{\frac{1}{p}\log\frac{1}{1-p}} < Z, \tag{4.11}$$

and hence, (for that joint distribution), for $\nu \ge 0$,

$$E[|\frac{Z}{\frac{1}{p}\log\frac{1}{1-p}} - X_p|^\nu] = E[(\frac{Z}{\frac{1}{p}\log\frac{1}{1-p}} - X_p)^\nu] \le \left(\frac{p}{1-c}\right)^\nu. \tag{4.12}$$

We remind the reader (see e.g. [16]) that the random variable S is **stochastically smaller** than the random variable T if $P\{S \le x\} \ge P\{T \le x\}$ for **all** x. This is the case if and only if there is a joint distribution of S and T for which $P\{S \le T\} = 1$. Thus, we have proven that X_p is stochastically smaller than $\frac{Z}{\frac{1}{p}\log\frac{1}{1-p}}$, and that the distributions are almost the same. This type of argument will be used several times in this paper.

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:

$$\text{if} \quad n\log\left(\frac{1}{1-p}\right) \le E < (n+1)\log\left(\frac{1}{1-p}\right) \quad \text{then} \quad H_p = np \tag{4.13}$$

has the property that

$$P\{H_p = np\} = p(1-p)^n. \tag{4.14}$$

Thus, if $(E_k)_{k=0}^{\infty}$ are i.i.d., exponentially distributed with parameter 1, and we define $(H_{p,k})_{k=0}^{\infty}$ functions of $(E_k)_{k=0}^{\infty}$ as in (4.13), and then define

$$Z = \sum_{k=0}^{\infty} c^k E_k, \quad X_p = \sum_{k=0}^{\infty} c^k H_{p,k}, \tag{4.15}$$

then Z and X_p have the required marginal distributions and (4.11) holds with probability one, etc.

The "Line of Attack" now is clear: we have proven that for p small 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 $\frac{Z}{\frac{1}{p}\log\frac{1}{1-p}}$.

Left to be proven is that the (or "any") stationary distribution of $V_{p,n}$ must, for p small, be very close to the distribution of X_p.

For that result we need different approaches in the cases $\alpha < 0$ and $0 < \alpha < 1$.

In the case $0 < \alpha < 1$ we will use "Linear Programming and Duality", see Sections 14 and 15.

In the case $\alpha < 0$ the approach will be as follows: we define

$$c(\alpha) = \frac{1}{(1+|\alpha|)^2} \sup_{0<z\le 1} \left(\frac{\left((1+z)^{1+|\alpha|} - 1 - (1+|\alpha|)z\right)}{z^2} \right), \tag{4.16}$$

$$d(\alpha) = \frac{1}{(1+|\alpha|)^{1+|\alpha|}} \sup_{0<x\le 1} \left((1+x)^{1+|\alpha|} - x^{1+|\alpha|} - x^{|\alpha|}(1+|\alpha|)\right), \tag{4.17}$$

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

$$f_{p,\alpha}(v) = \begin{cases} \frac{c(\alpha)p^2}{v} & \text{if } \frac{p}{1+|\alpha|} \le v < \infty, \\ \frac{d(\alpha)p^{1+|\alpha|}}{v^{|\alpha|}} & \text{if } 0 < v < \frac{p}{1+|\alpha|}, \end{cases} \tag{4.18}$$

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

$$Y_{p,n+1} = \begin{cases} Y_{p,n} + p + f_{p,\alpha}(X_{p,n}) & \text{if } \chi_{p,n} = \text{success}, \\ cY_{p,n} & \text{if } \chi_{p,n} = \text{failure}, \end{cases} \tag{4.19}$$

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 $Y_{p,n}$ all three are defined on the basis of the same sequence of successes and failures, all these random variables have a well defined joint distribution.

As in (4.8) 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, and identical values for $X_{p,0}$ and therefore identical sequences $(X_{p,n})_{n=0}^{\infty}$, then

$$Y_{p,n,1} - Y_{p,n,2} = c^{N(n)}(Y_{p,0,1} - Y_{p,0,2}). \tag{4.20}$$

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

Lemmas 4 below will be proven in Section 8.

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

$$X_{p,n} \leq V_{p,n} \leq Y_{p,n} \ \text{ for all } \ n \geq k. \tag{4.21}$$

(4.19) and (4.21) trivially show that if (X_p, V_p, Y_p) has a joint stationary distribution of the process $(X_{p,n}, V_{p,n}, Y_{p,n})$ and $E[Y_p] < \infty$ then

$$P\{0 < X_p < V_p < Y_p < \infty\} = 1 \ \text{ and } \ E[Y_p - X_p] = \frac{(1-p)}{(1-c)p} E[f_{p,\alpha}(X_p)]. \tag{4.22}$$

This makes it mandatory to compute $E[f_{p,\alpha}(X_p)]$. It must be noted that while $X_{p,n}$ is guaranteed to have a stationary distribution (and this stationary distribution is unique), there is no guarantee that $Y_{p,n}$ has a stationary distribution. If $Y_{p,n}$ has a stationary distribution then it is unique and $(X_{p,n}, V_{p,n}, Y_{p,n})$ has a joint stationary distribution, and (4.22) holds. We will see that if $\alpha < 0$ there exists a stationary distribution for Y_p with $E[Y_p] < \infty$ if and only if $E[f_{p,\alpha}(X_p)] < \infty$.

Lemma 5 below will be proven in Section 10.

Lemma 5: If $\alpha < 0$ and X_p has the distribution as in (4.4) then

$$E[f(X_p)] = p^2 \left(c(\alpha) E[\frac{1}{X_p} \chi(X_p \geq \frac{p}{1+|\alpha|})] + \left(\frac{p}{c}\right)^{|\alpha|} d(\alpha) E[\frac{1}{X_p^{|\alpha|}} \chi(X_p < \frac{p}{c(1+|\alpha|)})] \right), \tag{4.23}$$

where of course $\chi(.)$ is the indicator function. Lemma 5 makes it necessary to study $E[(X_p)^\nu]$ for $\nu < 0$.

As simple corollary of Lemma 3 and (1.14) (and the Helly-Bray theorem, see e.g. [15]) we see that for all $\nu \geq 0$ $E[X_p^\nu] \leq E[Z^\nu] < \infty$ and

$$\lim_{p \downarrow 0} E[X_p^\nu] = E[Z^\nu]. \tag{4.24}$$

We will see that (practically speaking) (4.24) also holds for $\nu < 0$, with the restriction that now p must be sufficiently small to guarantee that $E[X_p^\nu] < \infty$:

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

$$E[X_p^{-\nu}] \leq B(\nu) \text{ for all } 0 < p \leq p(\nu). \tag{4.25}$$

Lemma 6 will be proven in Section 6. As by-product of the proof of Lemma 9 in Section 5 we will see that

$$E[X_p^{-\nu}] < p^{-\nu} \frac{1-p}{1-pc^\nu} \text{ if } \nu \geq 0 \text{ and } 0 < p < c^\nu. \tag{4.26}$$

However, that result does not produce the uniformity of $B(\nu)$ in p.

Corollary 1: (4.24) holds for all $-\infty < \nu < +\infty$.

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

$$P\{0 < X_p \leq Y_p < \infty\} = 1 \text{ and } E[Y_p - X_p] \leq pD(\alpha) \text{ for all } 0 < p \leq \min(p(1), p(|\alpha|)). \tag{4.27}$$

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

We first derive two results (Lemmas 7 and 8) that neither depend on, nor are used in the proofs of, the results in Section 4.

We re-write (1.19) as

$$\log \frac{(1-\alpha)V_{p,n+1}}{p} = \begin{cases} \log \frac{(1-\alpha)V_{p,n}}{p} + (1-\alpha)\log\left(1+\frac{p}{(1-\alpha)V_{p,n}}\right) & \text{if } \chi_{p,n} = \text{success}, \\ \log \frac{(1-\alpha)V_{p,n}}{p} + \log c = \log \frac{(1-\alpha)V_{p,n}}{p} - |\log c| & \text{if } \chi_{p,n} = \text{failure}. \end{cases} \tag{5.1}$$

We now must study the consequences of (5.1) separately for $\alpha \leq 0$ and $0 \leq \alpha < 1$.

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

$$\log \frac{(1+|\alpha|)V_{p,n+1}}{p} \geq (1+|\alpha|)\log(1+|\alpha|) - |\alpha|\log|\alpha| \geq 0. \tag{5.2}$$

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

$$N_{p,n+1} = \begin{cases} 0 & \text{if } \chi_{p,n} = \text{success}, \\ N_{p,n} - |\log c| & \text{if } \chi_{p,n} = \text{failure}. \end{cases} \tag{5.3}$$

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

$$P\{N_p = -k|\log c|\} = p^k(1-p), \quad E[N_p] = -\frac{p}{1-p}|\log c|. \tag{5.4}$$

Thus,

Lemma 7: If $\alpha \leq 0$ then for any stationary distribution of $V_{p,n}$:

$$P\{\log \frac{(1+|\alpha|)V_p}{p} \leq -k|\log c|\} \leq p^k \tag{5.5}$$

for every nonnegative integer k, and

$$E\left[\left|\log\frac{(1+|\alpha|)V_p}{p}\right|\ \chi\left(\log\frac{(1+|\alpha|)V_p}{p}\leq 0\right)\right] < E[|N_p|] = \frac{p}{1-p}|\log c| < \infty, \tag{5.6}$$

where of course $\chi(.)$ is the indicator function.

Next, we consider the case $0 \leq \alpha < 1$. In this case we have immediately after every success

$$\log\frac{(1-\alpha)V_{p,n+1}}{p} > \alpha\log\frac{(1-\alpha)V_{p,n}}{p}. \tag{5.7}$$

In this case we define the stochastic process $M_{p,n}$ by

$$M_{p,n+1} = \begin{cases} \alpha M_{p,n} & \text{if } \chi_{p,n} = \text{success}, \\ M_{p,n} - |\log c| & \text{if } \chi_{p,n} = \text{failure}. \end{cases} \tag{5.8}$$

It is easily seen that if now $M_{p,n} \leq \log\frac{(1-\alpha)V_{p,n}}{p}$ for some n, then this holds for all $k \geq n$. 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

$$M_p = -|\log c|\sum_{k=0}^{\infty}\alpha^k Q_k, \tag{5.9}$$

where the Q_k are i.i.d geometrically distributed random variables, with

$$P\{Q_k = n\} = p^n(1-p), \tag{5.10}$$

thus

$$E[exp\{-sM_p\}] = \prod_{k=0}^{\infty}\frac{1-p}{1-p\exp\{+s|\log c|\alpha^k\}}, \tag{5.11}$$

and this holds for $s < \frac{|\log p|}{|\log c|}$.

Hence, $E[M_p] = -\frac{p|\log c|}{(1-p)(1-\alpha)}$ and thus we have

Lemma 8: If $0 \leq \alpha < 1$ then for any stationary distribution of $V_{p,n}$:

$$P\{\frac{(1-\alpha)V}{p} \leq x\} \leq P\{M_p \leq \log x\}, \tag{5.12}$$

and

$$E\left[\left|\log\frac{(1-\alpha)V_p}{p}\right|\chi\left(\log\frac{(1-\alpha)V_p}{p}\leq 0\right)\right] < E[|M_p|] = \frac{p|\log c|}{(1-p)(1-\alpha)} < \infty. \tag{5.13}$$

In addition, since (still $0 \leq \alpha < 1$) for $s \geq 0 \quad -s\log\frac{(1-\alpha)V_p}{p}$ is stochastically smaller than $-sM_p$ we have

$$E[exp\{-s\log\frac{(1-\alpha)V_p}{p}\}] = \left(\frac{p}{1-\alpha}\right)^s E[\frac{1}{V_p^s}] \leq$$

$$E[\exp\{-sM_p\}] = \prod_{k=0}^{\infty}\frac{1-p}{1-p\exp\{+s|\log c|\alpha^k\}} \tag{5.14}$$

as long as $0 \leq s < \frac{|\log p|}{|\log c|}$. This gives:

Lemma 9: If $0 \leq \alpha < 1$, if V_p has the stationary distribution of $V_{p,n}$ and $0 \leq s < \frac{|\log p|}{|\log c|}$ then

$$E[\frac{1}{V_p^s}] \leq \left(\frac{1-\alpha}{p}\right)^s \prod_{k=0}^{\infty}\frac{1-p}{1-p\exp\{+s|\log c|\alpha^k\}}. \tag{5.15}$$

Remark: Once Lemma 1 has been proven we know that if $0 \leq \alpha < 1$ the stationary distribution of V_p is stochastically smaller than the stationary distribution of X_p, therefore (for $\nu \geq 0$) $E[\frac{1}{X_p^\nu}] \leq E[\frac{1}{V_p^\nu}]$. The distribution of X_p is the distribution of V_p for $\alpha = 0$. The result above yields that

$$E[X_p^{-\nu}] < p^{-\nu}\frac{1-p}{1-pc^\nu} \text{ if } \nu \geq 0 \text{ and } 0 < p < c^\nu. \tag{5.16}$$

Once also lemma 6 has been proven, Lemma 9 shows that in fact for $p(\nu)$ any $p(\nu) < c^\nu$ satisfies. Lemma 9 does not provide the uniform upper bound $B(\nu)$ required to prove Lemma 6. Existence of the uniform bound will be proven in Section 6.

Next, we combine the Lemmas 7 and 8 with the results in the previous section to obtain:

Lemma 10: If either $(0 \leq \alpha < 1)$ or $(\alpha < 0$ **and** $0 < p \leq \min(p(1), p(|\alpha|s)))$ (where $p(\nu)$ is as in lemma 6) then for every stationary distribution of the process $V_{p,n}$,

$$E[|\log \frac{(1-\alpha)V_p}{p}|] < \infty. \tag{5.17}$$

Proof of Lemma 10: The part for $\log \frac{(1-\alpha)V_p}{p} \leq 0$ has been proven in this section. For the other part we use results from the previous section. If $(\alpha < 0$ **and** $0 < p \leq \min(p(1), p(|\alpha|)))$ we have $(\log V_p) < V_p \leq Y_p$, and Y_p has a finite first moment. If $0 \leq \alpha < 1$ we have $\log V_p < V_p \leq X_p$, and X_p has a finite first moment.

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

Lemma 10 and (5.1) together give:

Theorem 4:

As long as either $(0 \leq \alpha < 1)$ or $(\alpha < 0$ **and** $0 < p \leq \min(p(1), p(|\alpha|)))$ (where $p(\nu)$ is as in lemma 6) then for every stationary distribution of the process $V_{p,n}$:

$$E[\log(1 + \frac{p}{(1-\alpha)V_p})] = \frac{p|\log c|}{(1-p)(1-\alpha)}. \tag{5.18}$$

(Proof not needed.)

Theorem 4 will be used in the Sections 14 and 15, to cover the case $0 < \alpha < 1$.

6 The proof of Lemma 6

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

$$\int_0^\infty s^{\nu-1}\phi_Q(s)ds = \Gamma(\nu)E[\frac{1}{Q^\nu}]. \tag{6.1}$$

(No proof needed.)

We will also use the following Lemma for the Laplace Transform (4.4):

Lemma 11: If $0 < p < 1/3$, $0 < rp < 1/3$, and $0 < r < s$, then

$$\phi_{X_p}(s) < (1 + r/2)^{-\frac{\log(s) - \log(r)}{\log(1/c)}}. \tag{6.2}$$

Once we have Lemma 11, Lemma 6 is proven as follows:

Choose $\nu > 0$, then choose $r > 0$ such, that

$$\log(1 + \frac{r}{2}) > \nu \log \frac{1}{c}, \text{ i.e. } 1 + \frac{r}{2} > \left(\frac{1}{c}\right)^{\nu}. \tag{6.3}$$

Write

$$\int_0^\infty s^{\nu-1}\phi_{X_p}(s)ds = \int_0^r s^{\nu-1}\phi_{X_p}(s)ds + \int_r^\infty s^{\nu-1}\phi_{X_p}(s)ds. \tag{6.4}$$

In the first integral in the RHS of (6.4) use the fact that $0 < \phi_{X_p}(s) < 1$. In the second integral, $s > r$ so as long as $0 < p < \frac{1}{3}\min(1, \frac{1}{r})$, the lemma can be used and straightforward arithmetic gives

$$\int_0^\infty s^{\nu-1}\phi_{X_p}(s)ds < \frac{r^\nu \log(1 + r/2)}{\nu(\log(1 + r/2) - \nu\log(1/c))}. \tag{6.5}$$

This proves Lemma 6. The proof of Lemma 10 is next:

Proof of Lemma 10:

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

$$c^{K_0-1} \geq \frac{r}{s} > c^{K_0}. \tag{6.6}$$

We have

$$\phi_{X_p}(s) < \prod_{k=0}^{K_0-1} \frac{p}{1 - (1-p)exp\{-pc^k s\}}. \tag{6.7}$$

We will derive an upper bound for the right hand side in (6.7). For $0 \leq k \leq K_0 - 1$ we have

$$c^k ps \geq rp, \tag{6.8}$$

and hence

$$exp\{-c^k ps\} \leq exp\{-rp\} < 1 - rp + \frac{(rp)^2}{2}. \tag{6.9}$$

Hence:

$$1-(1-p)exp\{-c^k ps\} > 1-(1-p)(1-rp+\frac{(rp)^2}{2}) = p(1+(1-p)r(1-\frac{rp}{2})). \tag{6.10}$$

We check that since $0 < rp < 1/3$ and $0 < p \leq 1/3$

$$1-(1-p)exp\{-c^k ps\} > p(1+\frac{5}{6}(1-p)r) > p(1+\frac{10}{18}r) > p(1+\frac{1}{2}r) > p > 0. \tag{6.11}$$

Hence,

$$\phi_{X_p}(s) < (1+\frac{1}{2}r)^{-K_0}. \tag{6.12}$$

Since $c^{K_0} < \frac{r}{s}$, $K_0 > \frac{(\log(s) - \log(r))}{\log(\frac{1}{c})}$, and

$$\phi_{X_p}(s) < (1+r/2)^{-\frac{\log(s)-\log(r)}{\log(1/c)}}. \tag{6.13}$$

This completes the proof of Lemma 10, and thereby the proof of Lemma 6.

7 The proof of Lemma 1

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

$$R_{p,\alpha}(v) = v\left(1+\frac{p}{(1-\alpha)v}\right)^{1-\alpha} - v - p. \tag{7.1}$$

Whenever that does not lead to confusion we write $R(.)$ for $R_{p,\alpha}(.)$. With $V_{p,n}$ as in (1.17) we re-write (1.19) as

$$V_{p,n+1} = \begin{cases} V_{p,n} + p + R(V_{p,n}) & \text{if } \chi_{p,n} = \text{success}, \\ cV_{p,n} & \text{if } \chi_{p,n} = \text{failure}. \end{cases} \tag{7.2}$$

With the substitution $z = \frac{p}{(1-\alpha)v}$, $R(v)$ can be re-written as

$$\frac{p}{(1-\alpha)z}\Big((1+z)^{1-\alpha} - 1 - (1-\alpha)z\Big)\,. \tag{7.3}$$

With the substitution $x = \frac{(1-\alpha)v}{p} = \frac{1}{z}$, $R(v)$ can be re-written as

$$\frac{p}{(1-\alpha)}\Big(x^{\alpha}(1+x)^{1-\alpha} - x - (1-\alpha)\Big)\,. \tag{7.4}$$

This shows:

If $\alpha < 0$,

$$0 < R_{p,\alpha}(v) \le f_{p,\alpha}(v) < \infty \tag{7.5}$$

for all $0 < v < \infty$, where $f_{p,\alpha}(.)$ is as in (4.18),

$$R(v) \sim \frac{|\alpha| p^2}{2(1+|\alpha|)v} \ \text{ for } \ v \uparrow \infty, \ \ R(v) \sim \frac{p^{1+|\alpha|}}{(1+|\alpha|)^{1+|\alpha|}\ v^{|\alpha|}} \ \text{ for } \ v \downarrow 0, \tag{7.6}$$

and $R(.)$ is completely monotone (see e.g. [17]), therefore also convex, and decreasing from ∞ at 0 to 0 at ∞).

If $0 < \alpha < 1$, $R(v) < 0$ for all $0 < v < \infty$ and

$$R(v) \sim -\frac{\alpha p^2}{2(1-\alpha)v} \ \text{ for } \ v \uparrow \infty, \ \ R(v) + p \sim \frac{p^{1-\alpha}\ v^{\alpha}}{(1-\alpha)^{1-\alpha}} \ \text{ for } \ v \downarrow 0, \tag{7.7}$$

and $-R(.)$ is completely monotone, therefore also convex, and decreasing, from p at 0 to 0 at ∞).

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

To prove the complete monotonicity: Write $R(.)$ as in (7.4), take the derivative with respect to x, write the result as $\frac{p}{1-\alpha}\left(\alpha(B(x))^{\alpha-1}+(1-\alpha)(B(x))^{\alpha}-1\right)$ where $B(x)=\frac{x}{1+x}$, show that for $0\leq\alpha<1$ the derivative is completely monotone as function of B while for $\alpha\leq 0$ minus the derivative is completely monotone as function of B. Then use the fact that $B(x)$ as function of x is positive and has a completely monotone derivative.

An interesting special case is $\alpha=-1$ (the "TCP" case), where $R(v)=\frac{p^2}{4v}$.

(4.1), (7.2) and the results for the function R above prove Lemma 1.

Conditions that guarantee $E[V_p]$ and $E[V_p-X_p]$ to be finite are given in Section 9. If these expected values are finite, (7.2) and (4.1) show that for the joint stationary distribution of X_p and V_p

$$E[V_p-X_p]=\frac{1-p}{p(1-c)}E[R(V_p)]. \tag{7.8}$$

If $\alpha\leq 0$ this becomes

$$E[V_p-X_p]=E[|V_p-X_p|]=\frac{1-p}{p(1-c)}E[R(V_p)], \tag{7.9}$$

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

$$E[X_p-V_p]=E[|X_p-V_p|]=\frac{1-p}{p(1-c)}E[|R(V_p)|]. \tag{7.10}$$

What we are going to do next amounts, "conceptually", to finding an upper bound to $E[|R(V_p)|]$. In the case $0<\alpha<1$ this is exactly what we will do. It will be done in Section 14 and it will use the material in Section 5.

In the case $\alpha<0$ we will do something more complicated. It will be based on the fact that in that case V_p is stochastically larger (in fact: larger with probability 1) than X_p. To handle this case we will use, in Section 8, the process $Y_{p,n}$ introduced in (4.19). First we introduce another auxiliary process.

8 The case $\alpha < 0$ and the processes $\Lambda_{p,n}^{(k)}$ and $Y_{p,n}$

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

$$\Lambda_{p,n}^{(0)} = X_{p,n} \tag{8.1}$$

and for $k \geq 1$ define

$$\Lambda_{p,n+1}^{(k)} = \begin{cases} \Lambda_{p,n}^{(k)} + p + R(\Lambda_{p,n}^{(k-1)}) & \text{if } \chi_{p,n} = \text{success}, \\ c\Lambda_{p,n}^{(k)} & \text{if } \chi_{p,n} = \text{failure}. \end{cases} \tag{8.2}$$

We often write $\Lambda_{p,n}$ for $\Lambda_{p,n}^{(1)}$. Because the function $R(.)$ is positive and decreasing, and because $0 < R(v) \leq f(v)$, we have ($Y_{p,n}$ is as in (4.19)) that if for some n:

$$X_{p,n} = \Lambda_{p,n}^{(0)} \leq \Lambda_{p,n}^{(2)} \leq \Lambda_{p,n}^{(4)} \leq \cdots \leq V_{p,n} \leq \cdots \leq \Lambda_{p,n}^{(5)} \leq \Lambda_{p,n}^{(3)} \leq \Lambda_{p,n}^{(1)} = \Lambda_{p,n} \leq Y_{p,n} \tag{8.3}$$

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 processes in (8.3) have the same initial value. For example, $V_{p,0}$ with a stationary distribution of the process $(V_{p,n})$ and $X_{p,0} = Y_{p,0} = V_{p,0} = \Lambda_{p,0}^{(k)}$ for all $k \geq 0$.

Lemma 4 is a corollary to (8.3).

We see that a sufficient condition for existence of a stationary joint distribution of this infinite system of stochastic processes is that for the stationary distribution of $X_{p,n}$:

$$E[f(X_p)] < \infty. \tag{8.4}$$

We will see that in fact

$$E[f(X_p)] = O(p^2) \text{ for } p \downarrow 0. \tag{8.5}$$

and

$$E[V_p - X_p] \leq E[\Lambda_p - X_p] = \frac{1-p}{p(1-c)} E[R(X_p)] \leq \frac{1-p}{p(1-c)} E[f(X_p)]. \tag{8.6}$$

and in general

$$E[\Lambda_p^{(k+1)} - \Lambda_p^{(k)}] = \frac{1-p}{p(1-c)} E[R(\Lambda_p^{(k)}) - R(\Lambda_p^{(k-1)})]. \tag{8.7}$$

In Section 13 it will be shown that in fact for the stationary distributions of the processes $\Lambda_p^{(k)}$

$$E[|\Lambda_p^{(k)} - \Lambda_p^{(k-1)}|^\nu]^{\frac{1}{\nu}} = O(p^k) \ \ (p \downarrow 0). \tag{8.8}$$

This almost, but not quite, proves the uniqueness of the stationary distribution of V_p. The problem is that for $k \to \infty$ the constants in the big O functions may go to infinity faster than p^k goes to zero. With some extra work the proofs that lead to (8.8) can be modified to prove that for $p \downarrow 0$ $E[V_p - X_p]$ is $O(p)$ but not $o(p)$.

9 Upper Bounds for various Moments

(4.1) and (8.2) (with $k = 1$) show that

$$\Lambda_{p,n+1} - X_{p,n+1} = \begin{cases} (\Lambda_{p,n} - X_{p,n}) + R(X_{p,n} & \text{if } \chi_{p,n} = \text{success}, \\ c(\Lambda_{p,n} - X_{p,n}) & \text{if } \chi_{p,n} = \text{failure}. \end{cases} \tag{9.1}$$

Hence,

$$E[(\Lambda_{p,n+1} - X_{p,n+1})^\nu] = (1-p)E[((\Lambda_{p,n} - X_{p,n}) + R(X_{p,n}))^\nu] + pc^\nu E[(\Lambda_{p,n} - X_{p,n})^\nu]. \tag{9.2}$$

In this section we will use this to obtain crude, but finite, upper bounds for $E[(\Lambda_{p,n} - X_{p,n})^\nu]$ in the case $\nu > 0$, as long as $E[(R_(X_p))^\nu] < \infty$. Similar results can be obtained for $Y_{p,n} - X_{p,n}$ using $f(X_{p,n})$, etc.

Observation: Let G and H be any two random variables, and let $\nu \geq 0$. Let $\mu > 0$. Then

$$E[|G+H|^\nu] = E[|G+H|^\nu \chi(|H| \leq \mu|G|)] + E[|G+H|^\nu \chi(|H| > \mu|G|)] \leq$$

$$(1+\mu)^\nu E[|G|^\nu \chi(|H| \leq \mu|G|)] + \left(1 + \frac{1}{\mu}\right)^\nu E[|H|^\nu \chi(|H| > \mu|G|)] \leq$$

$$(1+\mu)^\nu E[|G|^\nu] + \left(\frac{1+\mu}{\mu}\right)^\nu E[|H|^\nu]. \tag{9.3}$$

We apply this result to (9.2) with $\nu \geq 0$ and obtain

$$E[(\Lambda_{p,n+1} - X_{p,n+1})^\nu] \leq$$

$$((1-p)(1+\mu)^\nu + pc^\nu)\, E[(\Lambda_{p,n} - X_{p,n})^\nu] + (1-p)\left(\frac{1+\mu}{\mu}\right)^\nu E[(R(X_{p,n})^\nu]. \tag{9.4}$$

We now choose $X_{p,n}$ to have the stationary distribution, so that $E[(R(X_{p,n}))^\nu] = E[(R(X_p))^\nu]$ for all n. We choose $\nu > 0$ and choose $\mu > 0$ such, that $(1-p)(1+\mu)^\nu + pc^\nu < 1$ (that is clearly possible). If now for some n

$$E[(\Lambda_{p,n} - X_{p,n})^\nu] \leq \frac{(1-p)\left(\frac{1+\mu}{\mu}\right)^\nu}{1 - ((1-p)(1+\mu)^\nu + pc)^\nu)} E[(R(X_p))^\nu] \tag{9.5}$$

then (9.5) holds for all $m \geq n$. Hence, if $E[(R(X_p))^\nu] < \infty$ and $E[(\Lambda_{p,0} - X_{p,0})^\nu] < \infty$ then

$$\limsup_{n\to\infty} E[(\Lambda_{p,n} - X_{p,n})^\nu] \leq \frac{(1-p)\left(\frac{1+\mu}{\mu}\right)^\nu}{1 - ((1-p)(1+\mu)^\nu + pc^\nu)} E[(R(X_p))^\nu] < \infty. \tag{9.6}$$

It is easily seen that if $\nu > 0$ and $0 < p \leq p(\nu|\alpha|)$ then $E[(R(X_p))^\nu] < \infty$ (see e.g. the next section) and the results above are applicable.

This section gives very coarse upper bounds to moments of interest, little more that proofs of finiteness. Once finiteness has been established the argument at the end of Section 7 becomes valid, and Sections 10 and 12 give much sharper upper bounds.

10 The proof of Lemma 5

Throughout this Section we have $\alpha \leq 0$. For $E[f(X_p)]$ we have ($\chi(.)$ stands for the indicator function)

$$E[f(X_p)] = c(\alpha)p^2 E[\frac{1}{X_p}\chi(X_p \geq \frac{p}{1+|\alpha|})] + d(\alpha)p^{1+|\alpha|}E[\frac{1}{X_p^{|\alpha|}}\chi(X_p < \frac{p}{1+|\alpha|})]. \quad (10.1)$$

If $\alpha \leq -1$ this is enough to prove the actual goal, (4.27). If $-1 < \alpha < 0$ one more trick is needed to first prove Lemma 5:

Since X_p can be written as in (4.2), and $0 < c < 1$, $X_p < p$ implies that $G_{p,0} = 0$. That 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 $-\infty < \nu < +\infty$

$$E[X_p^\nu \chi(X_p < \frac{p}{A})] = pc^\nu E[X_p^\nu \chi(X_p < \frac{p}{cA})]. \quad (10.2)$$

Thus:

$$E[f(X_p)] = p^2 \left(c(\alpha)E[\frac{1}{X_p}\chi(X_p \geq \frac{p}{1+|\alpha|})] + (\frac{p}{c})^{|\alpha|} d(\alpha)E[\frac{1}{X_p^{|\alpha|}}\chi(X_p < \frac{p}{c(1+|\alpha|)})] \right). \quad (10.3)$$

This proves Lemma 5 and thus (4.27).

This completes the proofs of the results in the Sections 4 and 5.

11 Stationary Distributions and the Main Results

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

$$X_{p,n} = \Lambda_{p,n}^{(0)} \leq \Lambda_{p,n}^{(2)} \leq \Lambda_{p,n}^{(4)} \leq \cdots \leq V_{p,n} \leq \cdots \leq \Lambda_{p,n}^{(5)} \leq \Lambda_{p,n}^{(3)} \leq \Lambda_{p,n}^{(1)} = \Lambda_{p,n} \leq Y_{p,n}. \quad (11.1)$$

By letting $n \to \infty$ and using the results from previous sections we now get, in the case $\alpha \leq 0$, 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 the process $(X_{p,n}, V_{p,n}, \Lambda_{p,n}, Y_{p,n})$ as in Section 8, and if $(X_p, V_p, \Lambda_p, Y_p)$ have such a joint stationary distribution then they have a joint distribution with the random variable Z where also (X_p, Z) have the joint distribution as in Lemma 3, $P\{X_p \leq V_p \leq \Lambda_p \leq Y_p\} = 1$, and

$$E[|V_p - X_p|] = E[V_p - X_p] \leq E[Y_p - X_p] \leq pD(\alpha). \tag{11.2}$$

To prove theorem 2 (with $C = 0$) we need to strengthen (11.2). In the next section we will prove that in the situation of Theorem 5, for every $k \in \{1, 2, \cdots\}$ there exists a $D_k(\alpha) < \infty$ such, that

$$E[|Y_p - X_p|^k] = E[(Y_p - X_p)^k] \leq p^k D_k(\alpha) \text{ for } 0 < p \leq min(p(k), p(k|\alpha|)). \tag{11.3}$$

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

Theorem 6:
If $0 \leq \alpha < 1$ then

$$E[|X_p - V_p|] = E[X_p - V_p] \leq \frac{p\alpha |\log c|}{(1-c)(1-\alpha)}, \tag{11.4}$$

and similar result can easily be derived for $E[|Z - V_p|] = E[Z - V_p]$.

The proof starts the same way (give V_p the stationary distribution) but then uses a different approach, see Section 14. The result for higher moments that completes the proof of Theorem 3 will be given in Section 15

12 Higher Moments if $\alpha \le 0$

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

$$f_{k,p,\alpha}(x) = \begin{cases} \frac{c_k(\alpha)p^2}{x} & \text{if } \frac{c^{k-1}p}{1+|\alpha|} \le x < \infty, \\ \frac{d(\alpha)p^{1+|\alpha|}}{x^{|\alpha|}} & \text{if } 0 < x < \frac{c^{k-1}p}{1+|\alpha|}. \end{cases} \tag{12.1}$$

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

$$f_{k,p,\alpha}(x) \ge f_{p,\alpha}(x) \ge R_{p,\alpha}(x) \;\text{ for all }\; x > 0. \tag{12.2}$$

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

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

$$Y_{k,p,n+1} = \begin{cases} Y_{k,p,n} + p + f_k(X_{p,n}) & \text{if } \chi_{p,n} = \text{success}, \\ cY_{k,p,n} & \text{if } \chi_{p,n} = \text{failure}, \end{cases} \tag{12.3}$$

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 $0 < D_k(\alpha) < \infty$ such, that for $0 < p \le \min(p(k), p(k|\alpha|))$ the stationary distributions satisfy

$$E[(Y_p - X_p)^k] \le E[(Y_{k,p} - X_p)^k] \le p^k D_k(\alpha). \tag{12.4}$$

We write $f_k(.)$ for $f_{k,p,\alpha}(.)$.

For $E[(f_k(X_p))^k]$ we have ($\chi(.)$ stands for the indicator function)

$$E[(f_k(X_p))^k] =$$

$$(c_k(\alpha))^k p^{2k} E[\frac{1}{(X_p)^k}\chi(X_p \ge \frac{c^{k-1}p}{1+|\alpha|})] + (d(\alpha))^k p^{k(1+|\alpha|)} E[\frac{1}{X_p^{k|\alpha|}}\chi(X_p < \frac{c^{k-1}p}{1+|\alpha|})]. \tag{12.5}$$

If $\alpha \leq -1$ this result is good enough for our purposes, and in fact it was unnecessary to introduce the new processes $(Y_{k,p,n})_{n=0}^{\infty}$. If $-1 < \alpha < 0$ another step is needed. By repeated use of (10.2) we have

$$E[(f_k(X_p))^k] =$$
$$p^{2k}\left((c_k(\alpha))^k E[\frac{1}{(X_p)^k}\chi(X_p \geq \frac{c^{k-1}p}{1+|\alpha|})] + \left(\frac{p}{c^k}\right)^{k|\alpha|}(d(\alpha))^k E[\frac{1}{X_p^{k|\alpha|}}\chi(X_p < \frac{p}{c(1+|\alpha|)})]\right). \tag{12.6}$$

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

$$E[(f_k(X_p))^k] < p^{2k}(h_k(\alpha))^k \tag{12.7}$$

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

Now we have

$$Y_{k,p,n+1} - X_{p,n+1} = \begin{cases} (Y_{k,p,n} - X_{p,n}) + f_k(X_{p,n}) & \text{if } \chi_{p,n} = \text{success}, \\ c(Y_{k,p,n} - X_{p,n}) & \text{if } \chi_{p,n} = \text{failure}, \end{cases} \tag{12.8}$$

Therefore, for the stationary distributions we have

$$(1-pc^k)E[(Y_{k,p} - X_p)^k] = (1-p)E[((Y_{k,p} - X_p) + f_k(X_p))^k]. \tag{12.9}$$

Using (12.7) and the Minkowski inequality (see e.g. [15]) this gives

$$(1-pc^k)E[(Y_{k,p} - X_p)^k] \leq (1-p)\left((E[(Y_{k,p} - X_p)^k])^{\frac{1}{k}} + p^2 h_k(\alpha)\right)^k \tag{12.10}$$

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

$$\left(\left(\frac{1-pc^k}{1-p}\right)^{\frac{1}{k}} - 1\right) E[(Y_{k,p} - X_p)^k]^{\frac{1}{k}} \leq p^2 h_k(\alpha) \tag{12.11}$$

for p sufficiently small. This immediately proves that

$$(E[(Y_{k,p} - X_p)^k])^{\frac{1}{k}} < \frac{pk}{1-c^k} h_k(\alpha) \tag{12.12}$$

if $p > 0$ is sufficiently close to zero. This completes the proof of (11.3)

13 The processes $\Lambda_{p,n}^{(k)}$ and higher moments

By the same methods as in the previous Section we prove that if all $\Lambda_p^{(k)}$ have the stationary distributions of the corresponding processes, and ν is a positive integer then

$$E[|\Lambda_p^{(k+1)} - \Lambda_p^{(k)}|^\nu]^{\frac{1}{\nu}} \le \frac{\nu}{p(1-c^\nu)} E[|R(\Lambda_p^k) - R(\Lambda_p^{k-1})|^\nu]^{\frac{1}{\nu}}. \tag{13.1}$$

Using the monotonicity and convexity of $R(.)$ we obtain

$$E[|R(\Lambda_p^k) - R(\Lambda_p^{k-1})|^\nu] \le E[|\Lambda_p^k - \Lambda_p^{k-1}|^\nu |R(X_p)|^\nu] \le E[|\Lambda_p^k - \Lambda_p^{k-1}|^{2\nu}]^{\frac{1}{2}} E[|R'(X_p)|^{2\nu}]^{\frac{1}{2}}. \tag{13.2}$$

Using the same methods as in the previous Section we prove that if p is sufficiently small then

$$E[|R'(X_p)|^\nu] \le p^{2\nu}(g_\nu(\alpha))^\nu, \tag{13.3}$$

where $g_\nu(\alpha)$ does not depend on p. Hence, for p sufficiently small,

$$E[|\Lambda_p^{(k+1)} - \Lambda_p^{(k)}|^\nu]^{\frac{1}{\nu}} \le \frac{\nu p}{1-c^\nu} \sqrt{g_{2\nu}(\alpha)} E[|\Lambda_p^{(k)} - \Lambda_p^{(k-1)}|^{2\nu}]^{\frac{1}{2\nu}}. \tag{13.4}$$

Continuing this process proves (8.8)

14 The case $0 \le \alpha < 1$: LP and Duality

Throughout this section we have $0 \le \alpha < 1$ and therefore have (7.10):

$$E[X_p - V_p] = E[|X_p - V_p|] = \frac{1-p}{p(1-c)} E[|R(V_p)|]. \tag{14.1}$$

We also have the constraint (5.18):

$$E[\log(1 + \frac{p}{(1-\alpha)V_p})] = \frac{p|\log c|}{(1-p)(1-\alpha)}. \tag{14.2}$$

Thus, we can obtain an upper bound for $E[|X_p - V_p|]$ by solving the "Linear Programming Problem": Find the supremum Sup for

$$E[|R(V)|] = E\left[V + p - V\left(1 + \frac{p}{(1-\alpha)V}\right)^{1-\alpha}\right] \tag{14.3}$$

where the supremum is taken over all non-negative random variables V for which (14.2) holds.

In order to create notation similar to that used in say [18] we denote

$$B = \frac{p|\log c|}{(1-p)(1-\alpha)}. \tag{14.4}$$

The dual (see e.g. [18]) of the LP above is: Find the infimum Inf for

$$\mu_1 + \mu_2 B \tag{14.5}$$

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

$$\mu_1 + \mu_2 \log\left(1 + \frac{p}{(1-\alpha)v}\right) \geq \left(v + p - v\left(1 + \frac{p}{(1-\alpha)v}\right)^{1-\alpha}\right). \tag{14.6}$$

The multipliers μ_1 and μ_2 are allowed to take on any value (negative, zero, or positive).

μ_1 is the multiplier for the constraint "$P\{0 < V < \infty\} = 1$", and μ_2 is the multiplier for the constraint (14.2).

Readers who do not like an LP with "continuous many" primal variables and "continuous many" dual constraints can take limits 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, $\pi_k = P\{V = v_k\}$ then becomes the primal variable.)

By setting $x = \frac{p}{(1-\alpha)v}$ we re-write the constraint (14.2) as: for all $0 < x < \infty$,

$$\mu_1 + \mu_2 \log(1+x) \geq \left(\frac{p}{(1-\alpha)x} + p - \frac{p}{(1-\alpha)x}(1+x)^{1-\alpha}\right). \tag{14.7}$$

We obtain a dual feasible solution by setting $\mu_1 = 0$, $\mu_2 = p\alpha$. To prove that for these values of μ_1 and μ_2 (14.7) indeed holds for all $0 < x < \infty$ takes straightforward arithmetic.

Thus, we know that

$$E[|R(V_p)|] \leq Sup = Inf \leq \mu_2 B = \frac{p^2\alpha|\log c|}{(1-p)(1-\alpha)}, \tag{14.8}$$

or

$$E[|X_p - V_p|] \leq \frac{p\alpha|\log c|}{(1-c)(1-\alpha)}. \tag{14.9}$$

15 The LP approach with higher moments

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

In case of higher moments (ν-th moment, with $\nu > 1$) we first want to find the supremum for

$$E\left[\left(V + p - V(1 + \frac{p}{(1-\alpha)V})^{1-\alpha}\right)^{\nu}\right] \tag{15.1}$$

subject to the constraint that V is a non-negative random variable with

$$E\left[\log(1 + \frac{p}{(1-\alpha)V})\right] = \frac{p|\log c|}{(1-p)(1-\alpha)}. \tag{15.2}$$

As before the RHS in (15.2) is denoted as B.

The dual of this LP is: Find the infimum for

$$\mu_1 + \mu_2 B \tag{15.3}$$

subject to

$$\mu_1 + \mu_2 \log(1 + \frac{p}{(1-\alpha)u}) \geq \left(u + p - u(1 + \frac{p}{(1-\alpha)u})^{1-\alpha}\right)^{\nu} \tag{15.4}$$

for all $0 < u < \infty$.

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

$$\mu_2 = \sup_{0<u<\infty} \frac{\left(u + p - u(1 + \frac{p}{(1-\alpha)u})^{1-\alpha}\right)^{\nu}}{\log\left(1 + \frac{p}{(1-\alpha)u}\right)}. \tag{15.5}$$

For $u \downarrow 0$ the RHS in (15.5) goes to zero. To study other values of u we substitute $\frac{p}{(1-\alpha)u} = x, \;\; u = \frac{p}{(1-\alpha)x}$. (15.5) now becomes

$$\mu_2 = \frac{p^{\nu}}{(1-\alpha)^{\nu}} \sup_{0<x<\infty} \frac{(1 + (1-\alpha)x - (1+x)^{1-\alpha})^{\nu}}{x^{\nu}\log(1+x)}. \tag{15.6}$$

We define

$$M_{\nu} = \frac{1}{(1-\alpha)^{\nu+1}} \sup_{0<x<\infty} \frac{(1 + (1-\alpha)x - (1+x)^{1-\alpha})^{\nu}}{x^{\nu}\log(1+x)}. \tag{15.7}$$

For $x \downarrow 0$ the RHS in (15.7) behaves like some constant times $x^{\nu-1}$. Hence, the sup in (15.7) is a max and is positive, finite. It depends on α and ν but is independent of p and c. We now have

$$E[|R(V_p)|^{\nu}] \leq Sup = Inf \leq \frac{p^{\nu+1}}{1-p}|\log c| M_{\nu} \tag{15.8}$$

for all $0 < p < 1$.

Unfortunately, the RHS in (15.8) behaves like $p^{\nu+1}$, not the $p^{2\nu}$ we had in (12.7). (In (12.7) $\nu > 1$ had to be integer, $\nu = k \geq 2$). Proceeding as in (12.8) etc (where $\nu = k \geq 2$ must be integer) we now only get that if $0 < \alpha < 1$ then for every integer $k \geq 2$ and every $0 < p < 1$

$$E\left[\frac{(X_p - V_p)^k}{p}\right] = E\left[\frac{|X_p - V_p|^k}{p}\right] < \frac{k^k |\log c| M_k}{(1-p)(1-c^k)^k} \tag{15.9}$$

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

16 Results for $C > 0$

In most of the previous sections we studied processes $V_{p,n}$, i.e. $\alpha < \beta = 1, \quad 0 < c_2 < 1, \quad C = 0$. In this section we deal with the case $C > 0$ and either $p \downarrow 0$ while $C > 0$, constant, or $C \downarrow 0$ while $p > 0$, constant. Similar to $V^*_{p,C,n}$ in (1.20) we define the process $(X^*_{p,C,n})_{n=0}^{\infty}$ by

$$X^*_{p,C,n+1} = \begin{cases} X^*_{p,C,n} + p & \text{if } \chi_{p,n} = \text{success}, \\ \max(cX^*_{p,C,n}, \frac{pC^{1-\alpha}}{(1-\alpha)c_1}) & \text{if } \chi_{p,n} = \text{failure}. \end{cases} \tag{16.1}$$

The key result that makes the case $C > 0$, constant, easy to handle is

Lemma 12: If for some n

$$X_{p,n} \leq X^*_{p,C,n} \leq X_{p,n} + \frac{pC^{1-\alpha}}{(1-\alpha)c_1}, \tag{16.2}$$

then this holds for all $k \geq n$. The proof is straightforward and is left as an exercise for the reader.

Similar to Lemma 1 we have

Lemma 13: If $\alpha \leq 0$ and for some $k \geq 0$ $V^*_{p,C,k} \geq X^*_{p,C,k}$ then this holds for all $n \geq k$.
If $0 \leq \alpha < 1$ and for some $k \geq 0$ $V^*_{p,C,k} \leq X^*_{p,C,k}$ then this holds for all $n \geq k$.
If $\alpha = 0$ and for some $k \geq 0$ $V^*_{p,C,k} = X^*_{p,C,k}$ then this holds for all $n \geq k$.

The proof makes use of re–writing (1.20) as

$$V^*_{p,C,n+1} = \begin{cases} V^*_{p,C,n} + p + R_{p,\alpha}(V^*_{p,C,n}) & \text{if } \chi_{p,n} = \text{success}, \\ \max(cV^*_{p,C,n}, \frac{pC^{1-\alpha}}{(1-\alpha)c_1}) & \text{if } \chi_{p,n} = \text{failure}, \end{cases} \tag{16.3}$$

with $R_{p,\alpha}(.)$ as in (7.1). The remainder of the proof is left to the reader.

In the case $0 \leq \alpha < 1$ we now are done: In that case $R(.)$ is an increasing function, hence the joint distribution of $V_p, V^*_{p,C}, X^*_{p,C}, X_p$ is such that

$$V_p \le V^*_{p,C} \le X^*_{p,C} \le X_p + \frac{pC^{1-\alpha}}{(1-\alpha)c_1}, \tag{16.4}$$

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 15 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 is stochastically decreasing) to the distribution of V_p. This follows from the stochastic monotonicity observations in Section 3. We have been unable to prove similar weak convergence with $C \downarrow 0$, while $p > 0$ constant for the case $\alpha < 0$.

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

$$\Lambda^*_{p,C,n+1} = \begin{cases} \Lambda^*_{p,C,n} + p + R_{p,\alpha}(X^*_{p,C,n}) & \text{if } \chi_{p,n} = \text{success}, \\ \max(c\Lambda^*_{p,C,n}, \frac{pC^{1-\alpha}}{(1-\alpha)c_1}) & \text{if } \chi_{p,n} = \text{failure}, \end{cases} \tag{16.5}$$

$$\Lambda^{**}_{p,C,n+1} = \begin{cases} \Lambda^{**}_{p,C,n} + p + R_{p,\alpha}(X^*_{p,C,n}) & \text{if } \chi_{p,n} = \text{success}, \\ c\Lambda^{**}_{p,C,n} & \text{if } \chi_{p,n} = \text{failure}. \end{cases} \tag{16.6}$$

Henceforth we choose the initial $X_{p,0}$ etc such, that the following inequalities hold for all n, instead of from some k on.

Since $X^*_{p,n} \le V^*_{p,n}$ and $R(.)$ is a positive, decreasing function,

$$X^*_{p,C,n} \le V^*_{p,C,n} \le \Lambda^*_{p,C,n} \tag{16.7}$$

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

$$\Lambda^{**}_{p,C,n} \le \Lambda_{p,n} \tag{16.8}$$

for all $n \ge 0$. By the same argument as (16.4) we have that

$$\Lambda^{**}_{p,C,n} \le \Lambda^*_{p,C,n} \le \Lambda^{**}_{p,C,n} + \frac{pC^{1-\alpha}}{(1-\alpha)c_1} \tag{16.9}$$

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

$$X_p \leq X^*_{p,C} \leq V^*_{p,C} \leq \Lambda^*_{p,C} \leq \Lambda^{**}_{p,C} + \frac{pC^{1-\alpha}}{(1-\alpha)c_1} \leq \Lambda_p + \frac{pC^{1-\alpha}}{(1-\alpha)c_1}. \tag{16.10}$$

Since we know that $\Lambda_p - X_p$ is non–negative, 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 $f_{k,p,\alpha}(.)$ but can directly call on results for the higher moments of $\Lambda_p - X_p$.

The approach above shows that for $C > 0$ $V^*_{p,C}$ etc have all moments: $R_{p,\alpha}(X^*_{p,C}) \leq R_{p,\alpha}(\frac{pC^{1-\alpha}}{(1-\alpha)c_1}) < \infty$, etc.

We see that if $\alpha < 0$ and $0 < p \leq \min(p(k), p(k|\alpha|))$ and $C \downarrow 0$ (with p constant), then all moments $E[(V^*_{p,C})^\nu]$ $(0 \leq \nu \leq k)$ remain bounded between $E[X^\nu_p]$ and $E[(\Lambda_p + \frac{pC^{1-\alpha}}{(1-\alpha)c_1})^\nu] < \infty$. There is no guarantee that for $C \downarrow 0$ these moments converge. It is conceivable that the set of limitpoints is some nontrivial subset of $[E[X^\nu_p], E[\Lambda^\nu_p]]$.

17 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 allows a significantly simplified proof.

If $\alpha = -1$, $R(v) = \frac{p^2}{4v}$, which already is of the right form, so we choose $f(v) = R(v)$, so that $Y_p = \Lambda_p$ as in Section 8. Hence,

$$0 < E[\Lambda_p - X_p] = \frac{p(1-p)}{4(1-c)} E[\frac{1}{X_p}], \tag{17.1}$$

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

$$\lim_{p\downarrow 0} E[\frac{1}{X_p}] = E[\frac{1}{Z}] = \log(\frac{1}{c}) = 2\log(\frac{1}{b}). \tag{17.2}$$

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

A Appendix, the behavior of $E[R(X_p)]$ etc

It is well known that

$$(1+x)^{1-\alpha} = \sum_{k=0}^{n} \begin{pmatrix} 1-\alpha \\ k \end{pmatrix} x^k + \begin{pmatrix} 1-\alpha \\ n+1 \end{pmatrix} \int_0^x \frac{(n+1)(x-u)^n}{(1+u)^{n+\alpha}} du \tag{A .1}$$

If $n \geq -\alpha$, for example, as in our situation, $\alpha \leq 0$ and $n \geq |\alpha|$, and if also $x \geq 0$, the absolute value of the rest-term in (A .1) is less than

$$\left| \begin{pmatrix} 1-\alpha \\ n+1 \end{pmatrix} \right| x^{n+1}. \tag{A .2}$$

Using this in the expression for $R(v)$, setting $v = X_p$ and taking expected values yields: If $\alpha < 0$ and $n \geq \max(|\alpha|, 2)$ and $0 < p \leq p(n)$ then

$$\left| E[R(X_p)] - \sum_{k=2}^{n} \begin{pmatrix} 1-\alpha \\ k \end{pmatrix} (\frac{p}{1-\alpha})^k E[X_p^{-(k-1)}] \right| \leq$$

$$\left| \begin{pmatrix} 1-\alpha \\ n+1 \end{pmatrix} (\frac{p}{1-\alpha})^{n+1} E[X_p^{-n}] \right| < \infty. \tag{A .3}$$

With Lemma 6 and its corollary, and using (1.15) and the fact that $0 < E[(X_p)^{-1}] - E[Z^{-1}] = O(p) \ \ (p \downarrow 0)$ we obtain: If $\alpha < 0$ then

$$\left| \frac{1}{p^2} E[R(X_p)] - \frac{|\alpha| |\log c|}{2(1+|\alpha|)} \right| = O(p) \ \ (p \downarrow 0). \tag{A .4}$$

For $\frac{d}{dv}R(v)$ we have

$$R'(v) = (1 + \frac{p\alpha}{(1-\alpha)v})(1 + \frac{p}{(1-\alpha)v})^{-\alpha} - 1 \tag{A .5}$$

and in the same way as the result above we obtain for the derivative $R'(.)$ of $R(.)$ that if $\alpha < 0$ then

$$\left| \frac{1}{p^2} E[R'(X_p)] + \frac{|\alpha|(1-c)|\log c|}{2(1+|\alpha|)c} \right| = O(p) \ \ (p \downarrow 0). \tag{A .6}$$

Similar results can be obtained for higher moments and higher derivatives. Such results could be used to rewrite some of the results in this paper.

Acknowledgment: I thank Jason Swanson for a stimulating exchange of Emails on the topics of this paper as well as [4], and among other items for his observation that uniqueness of the stationary stationary distribution of V_p $(0 < p < 1)$ had not been proven.

References

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

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

[3] Ott, T.J. (2005)
Transport Protocols in the TCP Paradigm and their Performance.
Telecommunication Systems **30:**4 pp 351-385, Springer 2005.
Also available at www.teunisott.com/Papers.

[4] Ott, T.J. and Swanson, J. (2006)
Asymptotic Behavior of a Generalized TCP Congestion Avoidance Algorithm.
(Submitted.) www.teunisott.com/Papers.

[5] Ott, T.J. and Swanson, J. (2006)
Stationarity of some processes in Transport Protocols.
Presented at the MAMA workshop, St Malo, France, June 2006.
An extended version is to appear in "Performance Evaluation Review".
Also available at www.teunisott.com/Papers.

[6] Ott, T.J. (2006)
Rate of convergence for the "Square Root Formula.
www.teunisott.com/Papers etc. (A short version of this paper.)

[7] Mathis, M., Semke, J. Mahdavi, J. and Ott, T.J. (1997)
The Macroscopic Behavior of the TCP Congestion Avoidance Algorithm.
Computer Communications Review 27 (3), pp 67 - 82 (July 1997).

[8] Kelly, C.T. (2003)
Scalable TCP: Improving performance in highspeed wide area networks.
ACM SigComm Computer Communications Review **32**(2) pp 83-91 (2003).

[9] Kelly, C.T. (2004)
Engineering Flow Controls in the Internet.
PhD thesis, Cambridge Univ, England. Available at
http://www-lce.eng.cam.ac.uk/~ctk21/papers.

[10] Altman, E., Avrachenkov, K., and Barakat, C. (2000)
A Stochastic Model of TCP/IP with Stationary Random Losses. *ACM SigComm 2000*, also *CCR* **30** no 4, pp 231 - 242, Oct 2000.

[11] Altman, E., Avrachenkov, K., Barakat, C., and Nunez-Queija, R. (2001)
State-Dependent M/G/1 type queueing analysis for congestion control in data net-

works.

Proceedings of IEEE Infocom 2001 Anchorage, Alaska, pp 1350 - 1359.

[12] Altman, E., Avrachenkov, K, Barakat, C., Kherani, A.A. and Prabhu, B.J. (2005) Analysis of MIMD congestion control algorithm for high speed networks. *Computer Networks* **48** (6) pp 972 - 989 (2005).

[13] Dumas, V., Guillemin, F. and Robert, Ph. (2002) A Markovian Analysis of Additive-Increase, Multiplicative Decrease (AIMD) Algorithms. *Adv. Appl. Prob.* **34** (1) pp 85 - 111.

[14] Guillemin, F., Robert, Ph. and Zwart, B. (2004) AIMD Algorithms and exponential functionals. *Ann. Appl. Prob.* **14** (1) pp 90 - 117.

[15] Michel Loève (1963) *Probability Theory* third ed D. van Nostrand.

[16] E.L. Lehmann (1986) *Testing Statistical Hypotheses* second edition John Wiley and sons.

[17] Feller, W. (1966) *An Introduction to Probability Theory and its Applications*, second corrected printing. Wiley, 1966.

[18] George B. Dantzig (1963) *Linear Programming and Extensions* Princeton University Press.