J. Appl. Prob. **44,** 618–635 (2007)
Printed in England

ASYMPTOTIC BEHAVIOR OF A GENERALIZED TCP CONGESTION AVOIDANCE ALGORITHM

TEUNIS J. OTT,* *Rutgers University*
JASON SWANSON,** *University of Wisconsin-Madison*

Abstract

The transmission control protocol (TCP) is a transport protocol used in the Internet. In Ott (2005), a more general class of candidate transport protocols called 'protocols in the TCP paradigm' was introduced. The long-term objective of studying this class is to find protocols with promising performance characteristics. In this paper we study Markov chain models derived from protocols in the TCP paradigm. Protocols in the TCP paradigm, as TCP, protect the network from congestion by decreasing the 'congestion window' (i.e. the amount of data allowed to be sent but not yet acknowledged) when there is packet loss or packet marking, and increasing it when there is no loss. When loss of different packets are assumed to be independent events and the probability p of loss is assumed to be constant, the protocol gives rise to a Markov chain $\{W_n\}$, where W_n is the size of the congestion window after the transmission of the nth packet. For a wide class of such Markov chains, we prove weak convergence results, after appropriate rescaling of time and space, as $p \to 0$. The limiting processes are defined by stochastic differential equations. Depending on certain parameter values, the stochastic differential equation can define an Ornstein–Uhlenbeck process or can be driven by a Poisson process.

Keywords: Weak convergence; stochastic differential equation; stationary distribution; TCP/IP; congestion avoidance

2000 Mathematics Subject Classification: Primary 60F05
Secondary 60G10; 60H10; 60J05

1. Introduction

The congestion avoidance algorithm of the TCP is designed to prevent network congestion during the transmission of data over a computer network. It does this by controlling the congestion window, i.e. the amount of data 'transmitted but not yet acknowledged' by a sender. What follows is a simplified description of a more general class of transport protocols.

Under appropriate units, the congestion window W determines the maximum amount of data that a source can send without acknowledgement. The 'TCP paradigm' (see [34]) is a class of protocols that includes the TCP (and other transport protocols). For each protocol in the TCP paradigm there are two functions, $\text{incr}(\cdot)$ and $\text{decr}(\cdot)$. If, while the congestion window equals W, a packet is found to be lost (or marked, under explicit congestion control (ECN); see [19] and [44]), then the congestion window is reduced by $\text{decr}(W)$. However, the congestion window is never reduced below some fixed minimum value $\ell \geq 0$. If the packet is not lost, then the congestion window is increased by $\text{incr}(W)$. For protocols in the TCP paradigm,

Received 20 March 2007; revision received 21 June 2007.
* Postal address: WINLAB, Rutgers University, New Brunswick, NJ 07930, USA.
Email address: ott@winlab.rutgers.edu
** Postal address: Mathematics Department, University of Wisconsin-Madison, 480 Lincoln Drive, Madison, WI 53706-1388, USA. Email address: swanson@math.wisc.edu

https://doi.org/10.1239/jap/1189717533 Published online by Cambridge University Press

$\text{incr}(W) = c_1 W^{\alpha}$ and $\text{decr}(W) = c_2 W^{\beta}$. In the special case of TCP, we have $c_1 = 1$, $\alpha = -1$, $c_2 = \frac{1}{2}$, and $\beta = 1$. Another special case of interest is when $\alpha = 0$ and $\beta = 1$. This is the algorithm which Kelly calls 'scalable TCP' in [22] and [23].

Let W_n denote the size of the congestion window after the transmission of the nth packet, or, more accurately, after receipt of the nth 'good' acknowledgement. Let χ_n be the indicator function of the event that the nth packet is lost, or, more accurately, that there is a loss between the $(n-1)$th and nth 'good' acknowledgement. We shall assume that the χ_ns are independent and identically distributed (i.i.d.). In particular, we are assuming that $p = \mathrm{P}(\chi_n = 1)$ is a constant that does not change with time. Under these assumptions, we are led to the parameterized family of Markov processes

$$W_{p,n+1} = (W_{p,n} + c_1 W_{p,n}^{\alpha}(1 - \chi_{p,n+1}) - c_2 W_{p,n}^{\beta} \chi_{p,n+1}) \vee \ell, \tag{1.1}$$

where $a \vee b$ denotes the maximum of a and b. We place the following assumptions on the various parameters in the model:

- $\{\chi_{p,n}\}_{n=1}^{\infty}$ is an i.i.d. sequence of $\{0, 1\}$-valued random variables,
- $p = \mathrm{P}(\chi_{p,n} = 1)$,
- $c_1 > 0$ and $c_2 > 0$,
- $-\infty < \alpha < \beta \leq 1$ and $\ell \geq 0$,
- if $\beta = 1$ then $c_2 < 1$, and
- if $\beta < 1$ then $\ell > 0$.

We will frequently drop the dependence on p from our notation and simply refer to the processes $\{\chi_n\}$ and $\{W_n\}$.

We are interested in studying the asymptotic behavior of $\{W_n\}$ as $p \to 0$. To this end, we define the continuous-time process

$$Z_p(t) = p^{\gamma} W_{\lfloor tp^{-\nu} \rfloor}, \tag{1.2}$$

where $\gamma = (\beta - \alpha)^{-1}$ and $\nu = (1 - \alpha)\gamma$. For the case in which $\beta = 1$, we will show that Z_p converges weakly as $p \to 0$ to the process Z defined by

$$Z(t) = Z(0) + c_1 \int_0^t Z(s)^{\alpha} \, \mathrm{d}s - c_2 \int_0^t Z(s-) \, \mathrm{d}N(s), \tag{1.3}$$

where N is a unit rate Poisson process, independent of $Z(0) = \lim Z_p(0)$. (Note that this is the conjecture given in [34, p. 362].) We will also show that, when $\ell > 0$, the stationary distributions of the discrete-time Markov chains $\{p^{\gamma} W_n\}$ converge weakly to the unique stationary distribution of Z. Questions about the convergence of the stationary distributions when $\beta = 1$, as well as the rate of convergence, are addressed in [36] and [38] using techniques that differ from those used in this paper.

For the case in which $\beta < 1$, we will show that Z_p converges to the process ζ defined by

$$\zeta(t) = \zeta(0) + \int_0^t (c_1 \zeta(s)^{\alpha} - c_2 \zeta(s)^{\beta}) \, \mathrm{d}s, \tag{1.4}$$

where $\zeta(0) = \lim Z_p(0)$. With the exception of the initial condition, the process ζ is entirely deterministic. The convergence of Z_p to ζ is therefore a law of large numbers type of result.

Hence, for the case in which $\beta < 1$, we can extend our analysis and study the fluctuations of Z_p around this central tendency. Unfortunately, it will not suffice to center Z_p by ζ. We must rather define

$$\zeta_p(t) = \zeta_p(0) + \int_0^t (c_1(1-p)\zeta_p(s)^\alpha - c_2\zeta_p(s)^\beta)\,\mathrm{d}s, \tag{1.5}$$

where $\zeta_p(0) \to \zeta(0)$, and consider the processes

$$\xi_p(t) = p^{-\tau}(Z_p(t) - \zeta_p(t)), \tag{1.6}$$

where $\tau = (\nu - 1)/2$. We will show that ξ_p converges weakly as $p \to 0$ to the process ξ defined by

$$\xi(t) = \xi(0) + \int_0^t (c_1\alpha\zeta(s)^{\alpha-1} - c_2\beta\zeta(s)^{\beta-1})\xi(s)\,\mathrm{d}s \tag{1.7}$$

$$- c_2\int_0^t \zeta(s)^\beta\,\mathrm{d}B(s), \tag{1.8}$$

where B is a Brownian motion and $\xi(0) = \lim \xi_p(0)$.

A special case of this last result is worth mentioning. For each $p \in [0, 1)$, define

$$c_p = \left(\frac{c_1(1-p)}{c_2}\right)^\gamma, \tag{1.9}$$

so that $\zeta_p(t) = c_p$ is an invariant solution to (1.5). Also, $\zeta(0) = \lim \zeta_p(0) = c_0$ is an invariant solution to (1.4). Hence, for an appropriate choice of $Z_p(0)$, ξ_p converges to the Ornstein–Uhlenbeck process defined by

$$\mathrm{d}\xi = -\mu\xi\,\mathrm{d}t + \sigma\,\mathrm{d}W, \tag{1.10}$$

where $W = -B$,

$$\mu = c_2\beta\left(\frac{c_1}{c_2}\right)^{\gamma(\beta-1)} - c_1\alpha\left(\frac{c_1}{c_2}\right)^{\gamma(\alpha-1)} = (\beta-\alpha)c_1^{-(1-\beta)/(\beta-\alpha)}c_2^{(1-\alpha)/(\beta-\alpha)},$$

and

$$\sigma = c_2\left(\frac{c_1}{c_2}\right)^{\gamma\beta} = c_1^{\beta/(\beta-\alpha)}c_2^{-\alpha/(\beta-\alpha)}.$$

(Note that this is the conjecture given in [34, p. 364].) We will also show that the stationary distributions of the discrete-time Markov chains, $\{p^{-\tau}(p^\gamma W_n - c_p)\}$, converge weakly to the unique stationary distribution of the above Ornstein–Uhlenbeck process.

It should be remarked that in this paper we use so-called ‘packet time’. That is, the progress of time is expressed in the number of data packets sent, or, more accurately, the number of good acknowledgements received. Several papers analyzing TCP use ‘clock time,’ where the progress of time is expressed in the number of round trip times (RTTs) elapsed. If the congestion window is the only limit on the ‘flight size’ (i.e. the amount of data transmitted by the source for which no acknowledgement has yet been received), all packets contain one maximum segment size (MSS) of data, and the congestion window is expressed in MSSs, then clock time, t_C, and packet time, t_P, are related by $\mathrm{d}t_\mathrm{P} = W\,\mathrm{d}t_\mathrm{C}$, where W denotes the size of the congestion window. Stationary distributions for ‘packet time’ and ‘clock time’ are related but are not the same. The relationship is given in [39].

2. Related work

When results like those in this paper are applied to the 'classical TCP', which has $\alpha = -1$ and $\beta = 1$, they predict throughput for a (large) TCP flow in the order of $1/\sqrt{p}$. This is called the 'square root law' for TCP, and original papers in this area were often identified with the square root law for TCP. Work in this area started with [39], which among other things gave the stationary distribution of the limit process for the case in which $\beta = 1$, and the relationship between 'packet time stationary distributions' and 'clock time stationary distributions'. This work gave the stationary distribution of the limit process $W_{p,n}$ for $p \downarrow 0$ and assumed the weak convergence results which strictly speaking were not proven until [36] and this paper. The paper [39] was presented at a workshop of the IFIP WG7.3 during Performance 1996 in Lausanne (October 1996) and also in a DIMACS workshop at Rutgers University in November 1996.

Another paper of historical interest is [33], which was presented in a workshop at ENS, Paris 2000. This paper first explicitly formulated the conjectures proven in this paper. It later appeared, in rewritten form, as [34].

In a nondistributional sense, some of the $1/\sqrt{p}$ results had been anticipated in [25].

The first papers identified with the 'square root law' that made it into the open literature were [1], [21], [27], [32], [41], and [42]. Of these, [32] was the first to use the language of stochastic differential equations. It used clock time and assumed that the probability of a drop in an RTT is independent of the size of the congestion window, i.e. the drop-probability per packet is roughly inversely proportional to the size of the window.

An extensive bibliography and discussion of previous work can be found in [16], which, among other things, includes a study of the effect of a congestion window limited by a send window or receive window (through the advertised window).

The first papers to use 'clock time' were [1] and [32]. Other papers to use clock time were [17] and [20].

Another paper of particular interest is [15], which uses stochastic differential equations, in clock time, to study joint evolution of RTT and congestion window size. The parameters of the two-dimensional stochastic differential equation were obtained from Internet measurements, not from postulating a particular behavior of sources and routers.

Other papers worth mentioning are [6], where (as in [15]) the RTT depends on the flightsize, [2], which is an ambitious attempt to build an all-encompassing model where many flows keep each others' RTTs and drop probabilities in equilibrium, [8], which analyzes the performance of scalable TCP ($\alpha = 0$, $\beta = 1$), [3], [4], [5], [7], [9], [10], and [26].

The papers [11], [12], [13], [14], [28], [29], [30], [31], and [40] use 'square root law results' and include analysis with, for example, drop probabilities that depend on the current size of the congestion window. The dependence was modeled by assuming ECN and a queuelength in the router which is a simple function of the flightsize.

The conjectures proven in this paper are formulated in [34], within which a number of other results linked to 'practicality' of control schemes were also obtained, such as relaxation times, typical numbers of dropped or marked packets per RTT, etc.

An alternative proof of the stationarity of the processes $(W_{p,n})_{n=0}^{\infty}$ studied in this paper is given in [38].

For a more complete review of the literature, the reader is referred to [16].

Among papers of possible future interest are [35] and [37]. In [35] a start is made with investigating the impact on stability of one RTT delay in the feedback. Ott and Kemperman [37] studied the transient behavior of the limit process we obtained in the case in which $\beta = 1$,

https://doi.org/10.1239/jap/1189717533 Published online by Cambridge University Press

and thus, insofar as limit results apply, can be used to predict the amount of clock time it takes to transfer a very large file using the file transfer protocol or similar protocol.

3. Main results

We first consider the case in which $\beta = 1$ and begin by cataloging some properties of the limit process Z.

Lemma 3.1. *If $Z(0) > 0$ almost surely (a.s.) then the stochastic differential equation (1.3) has a unique solution Z. With probability 1, $Z(t) > 0$ for all $t \geq 0$. Moreover, if $\tau = \inf\{t \geq 0\colon Z(t) = c_0\}$, where c_0 is given by (1.9), then $\tau < \infty$ a.s.*

Proof. For each realization of the Poisson process, (1.3) can be solved deterministically and the solution is unique. Let

$$T = \inf\{t \geq 0\colon Z(t) \notin (0, \infty)\}.$$

Since Z decreases only at the jump times of the Poisson process, and, with probability 1, these jump times have no accumulation points, it follows that $T = \infty$ a.s.

To show that $\tau < \infty$ a.s., it will suffice to assume that $Z(0) = x > 0$ is deterministic. We first consider the case in which $x \leq c_0$. Suppose that $\tau(\omega) = \infty$. Then $Z(t, \omega) < c_0$ for all $t \geq 0$. Find $u > r$ such that $u - r > \gamma c_2^{-1}$ and $N(u, \omega) = N(r, \omega)$. Then, for all $t \in (r, u]$,

$$Z(t, \omega) = Z(r, \omega) + c_1 \int_r^t Z(s, \omega)^\alpha \, \mathrm{d}s.$$

Since the solution to this integral equation is unique

$$Z(t, \omega) = (c_1(1 - \alpha)(t - r) + Z(r, \omega)^{1-\alpha})^\gamma.$$

Therefore,

$$c_0 > Z(u, \omega) > (c_1(1 - \alpha)(u - r))^\gamma > c_0$$

is a contradiction. Hence, $\tau < \infty$ a.s.

Next we consider the case in which $x > c_0$. Define

$$\sigma_1 = \inf\{t \geq 0\colon Z(t) < c_0\} \quad \text{and} \quad \sigma_2 = \inf\{t \geq \sigma_1\colon Z(t) = c_0\},$$

so that $\tau \leq \sigma_2$, and it will suffice to show that $\sigma_2 < \infty$ a.s. Fix $L > x$ and define $\rho = \inf\{t \geq 0\colon Z(t) \notin [c_0, L]\}$. Suppose that $\rho(\omega) = \infty$. Then $Z(t, \omega) \in [c_0, L]$ for all $t \geq 0$. Let

$$K = \inf\{u^\alpha\colon c_0 \leq u \leq L\} > 0.$$

Find $u > r$ such that $u - r > (L - c_0)/(c_1 K)$ and $N(u, \omega) = N(r, \omega)$. Then

$$L \geq Z(u, \omega) = Z(r, \omega) + c_1 \int_r^u Z(s, \omega)^\alpha \, \mathrm{d}s \geq c_0 + c_1(u - r)K > L$$

is a contradiction. Hence, $\rho < \infty$ a.s.

Now, observe that

$$Z(t \wedge \rho) = x + \int_0^{t\wedge\rho} (c_1 Z(s)^\alpha - c_2 Z(s)) \, \mathrm{d}s - c_2 \int_0^{t\wedge\rho} Z(s-) \, \mathrm{d}M(s),$$

where $M(t) = N(t) - t$ is the compensated Poisson process. If $s < t \wedge \rho$, where $a \wedge b$ denotes the minimum of a and b, then $Z(s) \geq c_0 = (c_1/c_2)^\gamma$. This implies that $c_1 Z(s)^\alpha - c_2 Z(s) \leq 0$. Since M is a martingale, $\mathrm{E}[Z(t \wedge \rho)] \leq x$. Letting $t \to \infty$ gives $\mathrm{E}[Z(\rho)] \leq x$. Hence, $\mathrm{P}(Z(\rho) = L) \leq x/L$. Note that either $Z(\rho) = L$ or $Z(\rho) < c_0$. Therefore,

$$\mathrm{P}(\sigma_1 = \infty) \leq \mathrm{P}(Z(\rho) = L) \leq \frac{x}{L}.$$

Letting $L \to \infty$ shows that $\sigma_1 < \infty$ a.s.

As in Theorem V.6.35 of [43], Z is a strong Markov process. Therefore,

$$\mathrm{P}(\sigma_2 = \infty) = \mathrm{E}[\mathrm{P}^{Z(\sigma_1)}(\tau = \infty)].$$

But $Z(\sigma_1) < c_0$, and we have already shown that $\mathrm{P}^x(\tau = \infty) = 0$ for all $x \leq c_0$. Hence, $\sigma_2 < \infty$ a.s.

We are now prepared to state our main results for the case in which $\beta = 1$. If μ_p and μ are Borel measures on a metric space S, then the notation $\mu_p \Rightarrow \mu$ will mean that μ_p converges weakly to μ as $p \to 0$, that is, $\int_S f \, \mathrm{d}\mu_p \to \int_S f \, \mathrm{d}\mu$ as $p \to 0$ for all bounded, continuous $f \colon S \to \mathbb{R}$. If X_p and X are S-valued random variables, then $X_p \Rightarrow X$ will mean that $PX_p^{-1} \Rightarrow PX^{-1}$. When X_p and X are processes, we will take our metric space to be $D_{\mathbb{R}^d}[0, \infty)$, the space of càdlàg functions (i.e. functions that are right continuous with left limits) from $[0, \infty)$ to $\mathbb{R}^d$, with the Skorohod metric. See [18, pp. 116–154] for details.

Theorem 3.1. *Suppose that $\beta = 1$. Let the processes Z_p be given by (1.2) and suppose that $Z_p(0) \Rightarrow Z(0)$, where $Z(0) > 0$ a.s. Let Z be the unique solution to (1.3). Then $Z_p \Rightarrow Z$.*

Theorem 3.2. *Suppose that $\beta = 1$ and $\ell > 0$. Then the Markov chain $\{W_n\}$ has a unique stationary distribution. Moreover, the process Z given by (1.3) has a unique stationary distribution η on $(0, \infty)$. For each $p > 0$, let η_p be the stationary distribution for the Markov chain $\{p^\gamma W_n\}$. Then $\eta_p \Rightarrow \eta$.*

For some results on stationary distributions for the case in which $\beta = 1$ and $\ell = 0$, see [36] and [38]. For the case in which $\beta < 1$, we need some preliminary definitions. Assume that, for all $p \in (0, 1)$, the processes $\{W_{p,n}\}$ are defined on the same probability space $(\Omega, \mathcal{F}, \mathrm{P})$. Define the σ-algebra by

$$\mathcal{F}_0 = \sigma(W_{p,0} \colon 0 < p < 1) \vee \mathcal{N}, \tag{3.1}$$

where $\mathcal{N}$ denotes the collection of events $D \in \mathcal{F}$ with $\mathrm{P}(D) = 0$.

Theorem 3.3. *Suppose that $\beta < 1$. Let the processes Z_p be given by (1.2). Suppose that $Z_p(0) \Rightarrow \zeta(0)$, where $\zeta(0) > 0$ a.s. Let ζ be the unique solution to (1.4). Then $Z_p \Rightarrow \zeta$. Moreover, if $Z_p(0) \to \zeta(0)$ in probability, then $Z_p \to \zeta$ in probability.*

Theorem 3.4. *Suppose that $\beta < 1$. Let the processes Z_p be given by (1.2). For each $p \in (0, 1)$, let $\zeta_p(0)$ be a strictly positive random variable defined on $(\Omega, \mathcal{F}, \mathrm{P})$. Assume that $\zeta_p(0)$ is $\mathcal{F}_0$-measurable and $Z_p(0) - \zeta_p(0) \to 0$ in probability. Define ζ_p and ξ_p by (1.5) and (1.6), respectively.*

Suppose that there exists a pair of random variables $(\xi(0), \zeta(0))$, defined on $(\Omega, \mathcal{F}, \mathrm{P})$, such that $\zeta(0) > 0$ a.s. and $(\xi_p(0), \zeta_p(0)) \Rightarrow (\xi(0), \zeta(0))$. Let B be a standard Brownian motion independent of $(\xi(0), \zeta(0))$ and define the processes ζ and ξ by (1.4) and (1.7), respectively. Then $(\xi_p, \zeta_p) \Rightarrow (\xi, \zeta)$.

Theorem 3.5. *Suppose that* $\beta < 1$. *Then the Markov chain* $\{W_n\}$ *has a unique stationary distribution. For each* $p > 0$, *let* η_p *be the stationary distribution for the Markov chain* $\{p^{-\tau}(p^\gamma W_n - c_p)\}$. *Then* $\eta_p \Rightarrow \eta$, *where* η *is the stationary distribution of the Ornstein–Uhlenbeck process given by (1.10).*

4. General definitions

Define

$$\Lambda_n = (\ell - W_{n-1} - c_1 W_{n-1}^\alpha (1-\chi_n) + c_2 W_{n-1}^\beta \chi_n) \vee 0,$$

so that

$$W_{n+1} = W_n + c_1 W_n^\alpha - (c_1 W_n^\alpha + c_2 W_n^\beta)\chi_{n+1} + \Lambda_{n+1}.$$

If we let $W(t) = W_{\lfloor t \rfloor}$, where $\lfloor a \rfloor$ denotes the greatest integer less than or equal to a, then we can rewrite this recursive relation as the integral equation

$$W(t) = W(0) + c_1 \int_0^t W(s-)^\alpha \,\mathrm{d}m(s) - \int_0^t (c_1 W(s-)^\alpha + c_2 W(s-)^\beta)\,\mathrm{d}S(s) + L(t),$$

where

$$m(t) = \lfloor t \rfloor, \qquad S(t) = \sum_{j=1}^{\lfloor t \rfloor} \chi_j, \quad \text{and } L(t) = \sum_{j=1}^{\lfloor t \rfloor} \Lambda_j.$$

Using (1.2), it is then easy to see that

$$\begin{aligned} Z_p(t) = Z_p(0) &+ c_1 \int_0^t Z_p(s-)^\alpha \,\mathrm{d}m_p(s) - c_1 p \int_0^t Z_p(s-)^\alpha \,\mathrm{d}S_p(s) \\ &- c_2 \int_0^t Z_p(s-)^\beta \,\mathrm{d}S_p(s) + L_p(t), \end{aligned} \tag{4.1}$$

where

$$m_p(t) = p^\nu m(tp^{-\nu}), \qquad S_p(t) = p^{\nu-1} S(tp^{-\nu}), \quad \text{and } L_p(t) = p^\gamma L(tp^{-\nu}).$$

Note that, if we define the filtration

$$\mathcal{F}_t^p = \mathcal{F}_0 \vee \sigma(\chi_{p,j} \colon j \le \lfloor tp^{-\nu} \rfloor),$$

then m_p, S_p, and L_p are all $\{\mathcal{F}_t^p\}$-adapted.

Define the $\mathbb{R}^2$-valued càdlàg $\{\mathcal{F}_t^p\}$-semimartingale by

$$Y_p = (m_p, S_p)^T$$

and define the function $G_p \colon \mathbb{R}^2 \to \mathbb{R}$ by

$$G_p(x) = (c_1 x^\alpha, -c_1 p x^\alpha - c_2 x^\beta)\, \mathbf{1}_{\{x>0\}}.$$

Then (4.1) becomes

$$Z_p(t) = Z_p(0) + \int_0^t G_p(Z_p(s-))\,\mathrm{d}Y_p(s) + L_p(t).$$

To show that Z_p converges as $p \to 0$, we will apply the methods of [24]. This approach, however, comes with two technical difficulties. The first is the presence of the local time term L_p; the second is the fact that G_p may have a singularity at the origin. To deal with these issues, we introduce the process Z_p^ε, defined as the unique solution to

$$Z_p^\varepsilon(t) = Z_p(0) + \int_0^t G_p^\varepsilon(Z_p^\varepsilon(s-))\,\mathrm{d}Y_p(s), \tag{4.2}$$

where $G_p^\varepsilon = G_p(\varepsilon)\,\mathbf{1}_{(-\infty,\varepsilon)} + G_p\,\mathbf{1}_{[\varepsilon,\infty)}$. To quantify the sense in which Z_p and Z_p^ε are close, we define the functional $h_\varepsilon\colon D_{\mathbb{R}^d}[0,\infty) \to [0,\infty]$ by

$$h_\varepsilon(x) = \inf\{t \geq 0\colon |x(t)| \wedge |x(t-)| \leq \varepsilon\},$$

and the stopping times $\tau_p(\varepsilon) = h_\varepsilon(Z_p^\varepsilon)$, and we observe that

$$L_p = 0 \quad \text{and} \quad Z_p = Z_p^\varepsilon \quad \text{on } [0, \tau_p(\varepsilon \vee p^\gamma \ell)). \tag{4.3}$$

By (3.5.2) of [18], if two càdlàg functions x and y agree on the interval $[0,t)$, then $d(x,y) \leq \mathrm{e}^{-t}$, where d is the metric on $D_{\mathbb{R}^d}[0,\infty)$.

5. Convergence of Z_p

In this section, we will prove Theorems 3.1 and 3.3 by applying the methods of [24] to the processes Z_p^ε given by (4.2). Therefore, we must define the processes to which they converge for the cases in which $\beta = 1$ and $\beta < 1$.

Let $G(x) = (c_1 x^\alpha, -c_2 x^\beta)\,\mathbf{1}_{\{x>0\}}$ and $G^\varepsilon = G(\varepsilon)\,\mathbf{1}_{(-\infty,\varepsilon)} + G\,\mathbf{1}_{[\varepsilon,\infty)}$, and note that $G_p^\varepsilon \to G^\varepsilon$ uniformly on compacts as $p \to 0$. Let N be a unit rate Poisson process, define

$$Y(t) = (t, N(t))^T \quad \text{and} \quad y(t) = (t,t)^T,$$

and let Z^ε and ζ^ε be the unique solutions to

$$Z^\varepsilon(t) = Z(0) + \int_0^t G^\varepsilon(Z^\varepsilon(s-))\,\mathrm{d}Y(s), \tag{5.1}$$

$$\zeta^\varepsilon(t) = \zeta(0) + \int_0^t G^\varepsilon(\zeta^\varepsilon(s-))\,\mathrm{d}y(s),$$

where $Z(0)$ and N are independent. Note that if $\beta = 1$, then $Z^\varepsilon = Z$ on $[0, h_\varepsilon(Z^\varepsilon))$ and $h_\varepsilon(Z^\varepsilon) = h_\varepsilon(Z) \to \infty$ a.s. as $\varepsilon \to 0$. Hence, $d(Z^\varepsilon, Z) \leq \exp(-h_\varepsilon(Z)) \to 0$ a.s. That is, $Z^\varepsilon \to Z$ a.s. in $D_{\mathbb{R}}[0,\infty)$. Similarly, if $\beta < 1$, then $\zeta^\varepsilon = \zeta$ on $[0, h_\varepsilon(\zeta^\varepsilon))$, $h_\varepsilon(\zeta^\varepsilon) = h_\varepsilon(\zeta) \to \infty$ a.s., and $\zeta^\varepsilon \to \zeta$ a.s. in $D_{\mathbb{R}}[0,\infty)$.

We will show that $Z_p^\varepsilon \Rightarrow Z^\varepsilon$ and $\zeta_p^\varepsilon \Rightarrow \zeta^\varepsilon$. To pass from this to the conclusions of Theorems 3.1 and 3.3, we will need the following lemma, which is easily proved using the Prohorov metric. (See [18, Section 3.1].)

Lemma 5.1. *Let (S,d) be a complete and separable metric space. Let $\{X_p\}_{p>0}$ be a family of S-valued random variables and suppose that, for each ε, there exists a family $\{X_p^\varepsilon\}_{p>0}$ such that*

$$\limsup_{p\to 0} \mathrm{E}[d(X_p, X_p^\varepsilon)] \leq \delta_\varepsilon,$$

where $\delta_\varepsilon \to 0$ as $\varepsilon \to 0$. Also, suppose that, for each ε, there exists Y^ε such that $X_p^\varepsilon \Rightarrow Y^\varepsilon$ as $p \to 0$. Then there exists an X such that $X_p \Rightarrow X$ and $Y^\varepsilon \Rightarrow X$.

https://doi.org/10.1239/jap/1189717533 Published online by Cambridge University Press

Proof of Theorem 3.1. Suppose that $\beta = 1$, Z_p is given by (1.2), and $Z_p(0) \Rightarrow Z(0)$, where $Z(0) > 0$ a.s. Let Z be the solution to (1.3).

Let Z_p^ε and Z^ε be as given by (4.2) and (5.1), respectively. We first show that $Z_p^\varepsilon \Rightarrow Z^\varepsilon$. Recall that $G_p^\varepsilon \to G^\varepsilon$ uniformly on compacts. Also, observe that $S_p \Rightarrow N$; see, for example, Problem 7.1 of [18]. Hence, since $Z_p(0)$ and Y_p are independent, $(Z_p(0), Y_p) \Rightarrow (Z(0), Y)$ in $D_{\mathbb{R}^3}[0,\infty)$. Hence, by Theorem 5.4 of [24], it will suffice to show that Y_p has a semimartingale decomposition, $Y_p = M_p + A_p$, into a martingale part and a bounded variation part such that, for each $t \geq 0$,

$$\sup_p \mathrm{E}[[M_p]_t + T_t(A_p)] < \infty, \tag{5.2}$$

where $[M_p]_t$ is the quadratic variation process of M_p and $T_t(A_p)$ is the total variation of A_p on the interval $[0, t]$. For this, define

$$\tilde{S}_p(t) = S_p(t) - m_p(t) = p^{\nu-1} \sum_{j=1}^{\lfloor tp^{-\nu} \rfloor} (\chi_j - p),$$

so that $\tilde{S}_p$ is an $\{\mathcal{F}_t^p\}$-martingale. Note that $T_t(m_p) = m_p(t)$ and

$$\mathrm{E}[\tilde{S}_p]_t = p^{2\nu-2} \sum_{j=1}^{\lfloor tp^{-\nu} \rfloor} \mathrm{E}\,|\chi_j - p|^2 = p^{2\nu-2} \lfloor tp^{-\nu} \rfloor p(1-p) \leq tp^{\nu-1}. \tag{5.3}$$

Since $\beta = 1$ implies that $\nu = 1$, this verifies (5.2) and shows that $Z_p^\varepsilon \Rightarrow Z^\varepsilon$.

By passing to a subsequence, we can assume that there exists a $[0,\infty]$-valued random variable $\sigma(\varepsilon)$ such that $(Z_p^\varepsilon, h_\varepsilon(Z_p^\varepsilon)) \Rightarrow (Z^\varepsilon, \sigma(\varepsilon))$. By (4.3), we have

$$\begin{aligned}
\limsup_{p\to 0} \mathrm{E}[d(Z_p, Z_p^\varepsilon)] &\leq \limsup_{p\to 0} \mathrm{E}[\exp(-\tau_p(\varepsilon \vee p^\gamma \ell))] \\
&= \limsup_{p\to 0} \mathrm{E}[\exp(-h_\varepsilon(Z_p^\varepsilon))] \\
&= \mathrm{E}[\exp(-\sigma(\varepsilon))].
\end{aligned}$$

We claim that $\mathrm{E}[\exp(-\sigma(\varepsilon))] \leq \mathrm{E}[\exp(-h_\varepsilon(Z^\varepsilon))]$. To see this, let us assume that, by the Skorohod representation theorem (see Theorem 3.1.8 of [18]), $(Z_p^\varepsilon, h_\varepsilon(Z_p^\varepsilon)) \to (Z^\varepsilon, \sigma(\varepsilon))$ a.s. Then $h_\varepsilon(Z^\varepsilon) \leq \sigma(\varepsilon)$ a.s., which proves the claim.

Since $h_\varepsilon(Z^\varepsilon) = h_\varepsilon(Z) \to \infty$ a.s. as $\varepsilon \to 0$, we can apply Lemma 5.1 to conclude that $Z_p \Rightarrow Z$.

Proof of Theorem 3.3. Suppose that $\beta < 1$, Z_p is given by (1.2), and $Z_p(0) \Rightarrow \zeta(0)$, where $\zeta(0) > 0$ a.s. Let ζ be the solution to (1.4).

Note that $\beta < 1$ implies that $\nu > 1$. Hence, (5.3) implies that (5.2) is satisfied and $\tilde{S}_p \to 0$ in probability. Therefore, $(Z_p(0), Y_p) \Rightarrow (Z(0), y)$ in $D_{\mathbb{R}^3}[0,\infty)$. By Theorem 5.4 of [24], $Z_p^\varepsilon \Rightarrow \zeta^\varepsilon$. By Corollary 5.6 of [24], if $Z_p(0) \to \zeta(0)$ in probability, then $Z_p^\varepsilon \to \zeta^\varepsilon$ in probability. By the same argument as above, this implies that Z_p converges to ζ in distribution or in probability, respectively.

6. Fluctuations of Z_p

In this section we prove Theorem 3.4. Let us first recall the setting of that theorem. We have $\beta < 1$ and Z_p given by (1.2). Recall that the processes Z_p are all defined on the same probability space $(\Omega, \mathcal{F}, \mathrm{P})$. For each $p > 0$, $\zeta_p(0)$ is an $\mathcal{F}_0$-measurable random variable, where $\mathcal{F}_0$ is given by (3.1), such that $\zeta_p(0) > 0$ a.s. and $Z_p(0) - \zeta_p(0) \to 0$ in probability. The processes ζ_p and ξ_p are then given by (1.5) and (1.6), respectively.

To apply the methods of [24], we wish to write ξ_p as the solution to a stochastic differential equation. By (1.5) and (4.1), we have

$$\begin{aligned}\xi_p(t) = \xi_p(0) &+ c_1(1-p)\int_0^t p^{-\tau}(Z_p(s-)^\alpha - \zeta_p(s)^\alpha)\,\mathrm{d}m_p(s)\\ &- c_2\int_0^t p^{-\tau}(Z_p(s-)^\beta - \zeta_p(s)^\beta)\,\mathrm{d}S_p(s) - c_2\int_0^t \zeta_p(s)^\beta\,\mathrm{d}B_p(s) + R_p(t), \end{aligned} \tag{6.1}$$

where

$$B_p(t) = p^{-\tau}(S_p(t) - m_p(t)) = p^{(\nu-1)/2}\sum_{j=1}^{\lfloor tp^{-\nu}\rfloor}(\chi_j - p)$$

and

$$\begin{aligned} R_p(t) = p^{-\tau}&\int_0^t (c_1(1-p)\zeta_p(s)^\alpha - c_2\zeta_p(s)^\beta)\,\mathrm{d}(m_p(s) - s)\\ &- c_1 p\int_0^t Z_p(s-)^\alpha\,\mathrm{d}B_p(s) + p^{-\tau}L_p(t). \end{aligned} \tag{6.2}$$

Given a real number r, let us define the continuous function $F_r\colon (0,\infty)^2 \to \mathbb{R}$ by

$$F_r(x,y) = \frac{x^r - y^r}{x-y}\,\mathbf{1}_{\{x\neq y\}} + ry^{r-1}\,\mathbf{1}_{\{x=y\}}\,.$$

Using this, (6.1) becomes

$$\begin{aligned}\xi_p(t) = \xi_p(0) &+ c_1(1-p)\int_0^t \xi_p(s-)\mathcal{D}_p^\alpha(s-)\,\mathrm{d}m_p(s)\\ &- c_2\int_0^t \xi_p(s-)\mathcal{D}_p^\beta(s-)\,\mathrm{d}S_p(s) - c_2\int_0^t \zeta_p(s)^\beta\,\mathrm{d}B_p(s) + R_p(t),\end{aligned} \tag{6.3}$$

where $\mathcal{D}_p^r = F_r(Z_p, \zeta_p)$.

Proof of Theorem 3.4. Suppose that there exists a pair of random variables $(\xi(0), \zeta(0))$, defined on $(\Omega, \mathcal{F}, \mathrm{P})$, such that $\zeta(0) > 0$ a.s. and $(\xi_p(0), \zeta_p(0)) \Rightarrow (\xi(0), \zeta(0))$. By the Skorohod representation theorem (see, for example, Theorem 2.1.8 of [18]), we can assume, without loss of generality, that $(\xi_p(0), \zeta_p(0)) \to (\xi(0), \zeta(0))$ a.s. Since the map that takes a point $x > 0$ to the unique solution of (1.5) with $\zeta_p(0) = x$ is continuous, $\zeta_p \to \zeta$ in probability and $(\xi_p(0), \zeta_p) \Rightarrow (\xi(0), \zeta)$. Also, since F_r is continuous, $\mathcal{D}_p^r \to r\zeta(\cdot)^{r-1}$ in probability.

Let

$$\begin{aligned}\mathcal{U}_p(t) &= \xi_p(0) - c_2\int_0^t \zeta_p(s)^\beta\,\mathrm{d}B_p(s) + R_p(t), \quad \text{and}\\ \mathcal{Y}_p(t) &= c_1(1-p)\int_0^t \mathcal{D}_p^\alpha(s-)\,\mathrm{d}m_p(s) - c_2\int_0^t \mathcal{D}_p^\beta(s-)\,\mathrm{d}S_p(s),\end{aligned}$$

so that (6.3) becomes

$$\xi_p(t) = \mathcal{U}_p(t) + \int_0^t \xi_p(s-)\,\mathrm{d}\mathcal{Y}_p(s). \tag{6.4}$$

We will apply the methods of [24] to this integral equation.

We first show that $R_p \to 0$ in probability. By the martingale central limit theorem (Theorem 7.1.4 of [18]), $B_p \Rightarrow B$, where B is a standard Brownian motion; by Theorem 3.3, $Z_p \to \zeta$ in probability; and by (5.3), $\{B_p\}$ satisfies (5.2). Hence, by Theorem 2.2 of [24],

$$c_1 p \int_0^t Z_p(s-)^\alpha \,\mathrm{d}B_p(s) \to 0$$

in probability. By (4.3), $p^{-\tau} L_p = 0$ on $[0, h_{p^\gamma \ell}(Z_p))$. Since $h_{p^\gamma \ell}(Z_p) \to \infty$ in probability, $p^{-\tau} L_p \to 0$ in probability.

For the final term of (6.2), note that $p^{-\tau}|m_p(t) - t| \le p^{\nu-\tau}$ and $\nu - \tau = (\nu+1)/2 > 0$. Hence, $p^{-\tau}(m_p(t) - t) \to 0$ uniformly. Let $f_p(s) = c_1(1-p)\zeta_p(s)^\alpha - c_2\zeta_p(s)^\beta$. Since $\zeta_p \to \zeta$ in probability, we can pass to a subsequence and assume that $\zeta_p \to \zeta$ uniformly on $[0, t]$, a.s. By (1.5), this implies that $\zeta_p' \to \zeta'$ uniformly on $[0, t]$, where the prime denotes differentiation with respect to t. Hence, f_p and f_p' converge uniformly. Integrating by parts, we have

$$p^{-\tau}\int_0^t f_p(s)\,\mathrm{d}(m_p(s) - s) = p^{-\tau} f_p(t)(m_p(t) - t) - p^{-\tau}\int_0^t (m_p(s) - s) f_p'(s)\,\mathrm{d}s,$$

which tends to 0 uniformly and completes the proof that $R_p \to 0$ in probability.

It now follows from Theorem 5.2 of [24] that $(\mathcal{U}_p, \mathcal{Y}_p, \zeta_p) \Rightarrow (\mathcal{U}, \mathcal{Y}, \zeta)$, where

$$\mathcal{U}(t) = \xi(0) - c_2\int_0^t \zeta(s)^\beta \,\mathrm{d}B(s), \quad \text{and}$$

$$\mathcal{Y}(t) = c_1\int_0^t \alpha\zeta(s)^{\alpha-1}\,\mathrm{d}s - c_2\int_0^t \beta\zeta(s)^{\beta-1}\,\mathrm{d}s,$$

and B is a standard Brownian motion independent of $(\xi(0), \zeta(0))$. By Remark 2.5 of [24], we may apply Theorem 5.4 of [24] to (6.4) and conclude that $(\xi_p, \zeta_p) \Rightarrow (\xi, \zeta)$, where ξ is the unique solution to (1.7).

7. Stationary distributions

In this section we prove Theorems 3.2 and 3.5. For this, we make time continuous in a slightly different manner than before. Let N be a unit rate Poisson process independent of $\{W_n\}$ and let $X(t) = W_{N(t)}$. Then X is a continuous-time Markov chain on $E = [\ell, \infty)$ with generator

$$A\varphi(x) = p(\varphi(x - g(x)) - \varphi(x)) + (1-p)(\varphi(x + c_1 x^\alpha) - \varphi(x)),$$

where $g(x) = (c_2 x^\beta) \wedge (x - \ell)$. When $\beta = 1$, we will study the process

$$\hat{Z}_p(t) = p^\gamma X(tp^{-1}),$$

whereas when $\beta < 1$, we will consider

$$\hat{\xi}_p(t) = p^{-\tau}(p^\gamma X(tp^{-\nu}) - c_p),$$

https://doi.org/10.1239/jap/1189717533 Published online by Cambridge University Press

where c_p is given by (1.9). It is easy to see that a probability measure is a stationary distribution for $\{p^\gamma W_n\}$ or $\{p^{-\tau}(p^\gamma W_n - c_p)\}$ if and only if it is a stationary distribution for $\hat{Z}_p$ or $\hat{\xi}_p$, respectively.

Lemma 7.1. *If* $\ell > 0$, *then* $\{W_n\}$ *has a unique stationary distribution.*

Proof. It will suffice to show that X has a unique stationary distribution. Let $\varphi(x) = x$ so that

$$A\varphi(x) = -pg(x) + (1-p)c_1 x^\alpha.$$

Since $g(x) = c_2 x^\beta$ for sufficiently large x, $A\varphi$ is bounded above and $A\varphi(x) \to -\infty$ as $x \to \infty$. By Lemmas 4.9.5 and 4.9.7 of [18], the family of probability measures $\{\mu_t\}_{t\geq 1}$ defined by

$$\mu_t(\Gamma) = \frac{1}{t}\int_0^t P^x(X(s) \in \Gamma)\,\mathrm{d}s$$

is relatively compact. By Theorem 4.9.3 of [18], any subsequential weak limit of $\{\mu_t\}$ is a stationary distribution for X.

To show that the stationary distribution is unique, it will suffice to show that, for all $x \in E$,

$$\tau = \inf\{t \geq 0 \colon X(t) = \ell\} < \infty, \qquad \mathrm{P}^x\text{-a.s.}$$

(See, for example, Problem 4.36 of [18].) Let $x \in E$ be arbitrary and let $\varepsilon > 0$. Choose M such that $\mu_t([\ell, M]) \geq 1 - \varepsilon$ for all $t \geq 0$. Note that there exists a $K > 0$ such that $\mathrm{P}^y(\tau < \infty) \geq K$ for all $y \in [\ell, M]$.

Define the stopping times $\tau_0 = 0$ and

$$\tau_{j+1} = \inf\{t \geq \tau_j + 1 \colon X(t) \leq M\},$$

and note that $\tau_j \to \infty$ a.s. By the strong Markov property,

$$\begin{aligned}\mathrm{P}(\tau = \infty, \tau_j < \infty) &= \mathrm{E}[\mathbf{1}_{\{\tau \geq \tau_j, \tau_j < \infty\}}\,\mathrm{P}^{X(\tau_j)}(\tau = \infty)] \\ &\leq (1-K)\,\mathrm{P}(\tau \geq \tau_j, \tau_j < \infty).\end{aligned}$$

Letting $j \to \infty$ shows that $\mathrm{P}(\{\tau = \infty\} \cap D) = 0$, where D is the event that $\tau_j < \infty$ for all j. Note that

$$\mathbf{1}_{D^c} \leq \liminf_{t\to\infty} \frac{1}{t}\int_0^t \mathbf{1}_{\{X(s) > M\}}\,\mathrm{d}s.$$

Hence, by Fatou's Lemma, $\mathrm{P}(D^c) \leq \liminf_{t\to\infty} \mu_t((M, \infty)) \leq \varepsilon$. Therefore, $\mathrm{P}(\tau = \infty) = \mathrm{P}(\{\tau = \infty\} \cap D^c) \leq \varepsilon$. Since ε was arbitrary, $\tau < \infty$ P^x-a.s. and the stationary distribution is unique.

Proof of Theorem 3.2. In what follows, C and K will denote strictly positive, finite constants that do not depend on p and may change value from line to line.

Suppose that $\beta = 1$, $\ell > 0$, and η_p is the stationary distribution for $\{p^\gamma W_n\}$. Then η_p is the stationary distribution for $\hat{Z}_p$, which is a continuous-time Markov chain on $E_p = [p^\gamma \ell, \infty)$ with generator

$$A_p\varphi(x) = \varphi(x - p^\gamma g(p^{-\gamma}x)) - \varphi(x) + p^{-1}(1-p)(\varphi(x + pc_1x^\alpha) - \varphi(x)).$$

https://doi.org/10.1239/jap/1189717533 Published online by Cambridge University Press

Let $\varphi(x) = x + x^{-1}$, so that

$$A_p\varphi(x) = -p^\gamma g(p^{-\gamma}x) + (1-p)c_1x^\alpha + \frac{p^\gamma g(p^{-\gamma}x)}{x(x - p^\gamma g(p^{-\gamma}x))} - \frac{(1-p)c_1x^\alpha}{x(x+pc_1x^\alpha)}.$$

Since $x \mapsto 1 + pc_1x^{\alpha-1}$ is decreasing,

$$1 + pc_1x^{\alpha-1} \le 1 + pc_1(p^\gamma\ell)^{\alpha-1} = 1 + c_1\ell^{\alpha-1}$$

for all $x \in E_p$. Hence,

$$A_p\varphi(x) \le -p^\gamma g(p^{-\gamma}x) + Cx^\alpha + \frac{p^\gamma g(p^{-\gamma}x)}{x(x - p^\gamma g(p^{-\gamma}x))} - Kx^{\alpha-2},$$

whenever $p < \frac{1}{2}$.

If $x \ge p^\gamma\ell/(1-c_2)$, then $g(p^{-\gamma}x) = c_2p^{-\gamma}x$ and

$$A_p\varphi(x) \le -Kx + Cx^\alpha + Cx^{-1} - Kx^{\alpha-2}.$$

If $x < p^\gamma\ell/(1-c_2)$, then $g(p^{-\gamma}x) = p^{-\gamma}x - \ell$ and

$$A_p\varphi(x) \le Cx^\alpha + \frac{x - p^\gamma\ell}{xp^\gamma\ell} - Kx^{\alpha-2} \le Cx^\alpha + (p^\gamma\ell)^{-1} - Kx^{\alpha-2}.$$

But in this case, $(p^\gamma\ell)^{-1} < Cx^{-1}$. Therefore, it follows that

$$A_p\varphi(x) \le C - Kx - Kx^{\alpha-2}$$

for all $x \in E_p$.

Let $\varepsilon > 0$. Define

$$L = \sup_{p<1/2}\ \sup_{x\in E_p} A_p\varphi(x) < \infty$$

and let $m = L(1-\varepsilon)/\varepsilon$. Choose $M > 0$ such that $x \notin [M^{-1}, M]$ implies that $A_p\varphi(x) < -m$ for all $p < \frac{1}{2}$. By Corollary 4.9.8 of [18],

$$\eta_p([M^{-1}, M]) \ge \eta_p(\{x : A_p\varphi(x) \ge -m\}) \ge \frac{m}{L+m} = 1 - \varepsilon.$$

The family of measures $\{\eta_p\}$ is therefore relatively compact on $(0,\infty)$. By passing to a subsequence, we can assume that $\eta_p \Rightarrow \eta$ for some probability measure η on $(0,\infty)$.

Now let $p^\gamma W_0$ have distribution η_p and let Z_p be given by (1.2). By Theorem 3.1, $Z_p \Rightarrow Z$, where Z satisfies (1.3) with $\mathrm{P}\,Z(0)^{-1} = \eta$. Fix $t_1 \le \cdots \le t_n$. Then

$$\begin{aligned}(Z_p(t_1), \dots, Z_p(t_n)) &= p^\gamma(W_{\lfloor t_1p^{-1}\rfloor}, \dots, W_{\lfloor t_np^{-1}\rfloor})\\ &\overset{\mathrm{D}}{=} p^\gamma(W_0, W_{\lfloor t_2p^{-1}\rfloor - \lfloor t_1p^{-1}\rfloor}, \dots, W_{\lfloor t_np^{-1}\rfloor - \lfloor t_1p^{-1}\rfloor})\\ &= (Z_p(0), Z_p(t_2 - t_1), \dots, Z_p(t_n - t_1)) + \varepsilon,\end{aligned}$$

where the components of $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)$ are given by $\varepsilon_j = Z_p(h_j) - Z_p(t_j - t_1)$ and $h_j = (\lfloor t_jp^{-1}\rfloor - \lfloor t_1p^{-1}\rfloor)p$. Note that $h_j \to t_j - t_1$ as $p \to 0$ and, for fixed t, Z is almost surely continuous at t. Hence, $\varepsilon \to 0$ a.s., which gives

$$(Z_p(t_1), \dots, Z_p(t_n)) \Rightarrow (Z(0), Z(t_2 - t_1), \dots, Z(t_n - t_1)).$$

But,

$$(Z_p(t_1), \dots, Z_p(t_n)) \Rightarrow (Z(t_1), \dots, Z(t_n)),$$

so Z is a stationary process, and η is a stationary distribution for Z. The uniqueness of η follows from Lemma 3.1.

For the proof of Theorem 3.5, note that $\hat{\xi}_p$ is a continuous-time Markov chain on $E_p = [p^{-\tau}(p^{\gamma}\ell - c_p), \infty)$ with generator

$$\begin{aligned} A_p\varphi(x) = {} & p^{-\nu+1}(\varphi(x - p^{\gamma-\tau}g(p^{\tau-\gamma}x + p^{-\gamma}c_p)) - \varphi(x)) \\ & + p^{-\nu}(1-p)(\varphi(x + p^{\gamma-\tau}c_1(p^{\tau-\gamma}x + p^{-\gamma}c_p)^{\alpha}) - \varphi(x)). \end{aligned} \tag{7.1}$$

We will use the same argument as in the proof of Theorem 3.2; however, this time we will use the Lyapunov function $\varphi(x) = |x|^r$, where r is sufficiently large. Our key estimate on $A_p\varphi(x)$ is given in the following lemma and is valid as long as $|x|$ is not too large.

Lemma 7.2. *Suppose that $\beta < 1$. Let $\varphi(x) = |x|^r$, where $r \geq 2$, and let A_p be given by (7.1). Let $0 < \delta < M < \infty$ be arbitrary. Then there exist a $p_0 > 0$ and strictly positive, finite constants C and K such that*

$$A_p\varphi(x) \leq C - K|x|^r$$

for all $p \leq p_0$ and all $x \in E_p$ satisfying $\delta \leq p^{\tau}x + c_p \leq M$.

Proof. For notational simplicity, let us define $y_p(x) = p^{\tau}x + c_p$ so that

$$\begin{aligned} A_p\varphi(x) = {} & p^{-\nu+1}(\varphi(x - p^{\gamma-\tau}g(p^{-\gamma}y_p)) - \varphi(x)) \\ & + p^{-\nu}(1-p)(\varphi(x + p^{\gamma-\tau}c_1(p^{-\gamma}y_p)^{\alpha}) - \varphi(x)). \end{aligned}$$

Either $g(x) = c_2x^{\beta}$ or $g(x) < c_2x^{\beta}$. Note that there exists an $x_0 > \ell$ such that $g(x) = c_2x^{\beta}$ if and only if $x \geq x_0$. Hence, if $g(p^{-\gamma}y_p) < c_2(p^{-\gamma}y_p)^{\beta}$, then $p^{-\gamma}y_p < x_0$, which implies that $x < p^{-\tau}(p^{\gamma}x_0 - c_p)$. If p is sufficiently small, this implies that $x < 0$. Since φ is decreasing on $(-\infty, 0]$, it follows that

$$\begin{aligned} A_p\varphi(x) \leq {} & p^{-\nu+1}(\varphi(x - p^{\gamma-\tau-\gamma\beta}c_2y_p^{\beta}) - \varphi(x)) \\ & + p^{-\nu}(1-p)(\varphi(x + p^{\gamma-\tau-\gamma\alpha}c_1y_p^{\alpha}) - \varphi(x)) \end{aligned}$$

for all $x \in E_p$.

Observe that

$$\begin{aligned} |\varphi(z) - \varphi(x) - \varphi'(x)(z-x)| &= \left| \int_x^z (z-u)\varphi''(u)\,\mathrm{d}u \right| \\ &\leq C|z-x|^2(|x|^{r-2} + |z|^{r-2}) \\ &\leq C|x|^{r-2}|z-x|^2 + C|z-x|^r. \end{aligned}$$

Hence,

$$\begin{aligned} A_p\varphi(x) \leq {} & -\varphi'(x)p^{-\tau}(p^{-\nu+1+\gamma-\gamma\beta}c_2y_p^{\beta} - p^{-\nu+\gamma-\gamma\alpha}c_1(1-p)y_p^{\alpha}) \\ & + C|x|^{r-2}(p^{-\nu+1+2\gamma-2\tau-2\gamma\beta}c_2^2y_p^{2\beta} + p^{-\nu+2\gamma-2\tau-2\gamma\alpha}c_1^2y_p^{2\alpha}) \\ & + C(p^{-\nu+1+r\gamma-r\tau-r\gamma\beta}c_2^r y_p^{r\beta} + p^{-\nu+r\gamma-r\tau-r\gamma\alpha}c_1^r y_p^{r\alpha}). \end{aligned}$$

We can simplify these exponents by observing that

$$\begin{aligned}
-\nu + \gamma - \gamma\alpha &= 0, \\
-\nu + 1 + \gamma - \gamma\beta &= 0, \\
-\nu + 2\gamma - 2\tau - 2\gamma\alpha &= 1, \\
-\nu + 1 + 2\gamma - 2\tau - 2\gamma\beta &= 0, \\
-\nu + 1 + r\gamma - r\tau - r\gamma\beta &= \tau(r-2), \\
-\nu + r\gamma - r\tau - r\gamma\alpha &= r - 1 + \tau(r-2).
\end{aligned}$$

Thus,

$$\begin{aligned}
A_p\varphi(x) \le {} & -\varphi'(x)p^{-\tau}(c_2 y_p^\beta - c_1(1-p)y_p^\alpha) + C|x|^{r-2}(y_p^{2\beta} + p y_p^{2\alpha}) \\
& + C(p^{\tau(r-2)} y_p^{r\beta} + p^{r-1+\tau(r-2)} y_p^{r\alpha}).
\end{aligned}$$

Since $\varphi'(x)$ and $c_2 y_p^\beta - c_1(1-p)y_p^\alpha$ have the same sign, this gives

$$\begin{aligned}
A_p\varphi(x) \le {} & -r|x|^{r-1}p^{-\tau}|c_2 y_p^\beta - c_1(1-p)y_p^\alpha| + C|x|^{r-2}(y_p^{2\beta} + p y_p^{2\alpha}) \\
& + C(p^{\tau(r-2)} y_p^{r\beta} + p^{r-1+\tau(r-2)} y_p^{r\alpha}) \qquad (7.2)
\end{aligned}$$

for all $x \in E_p$.

If $r \ge 2$ and $\delta \le y_p \le M$, then

$$A_p\varphi(x) \le -r|x|^{r-1}p^{-\tau}c_2 y_p^\alpha |y_p^{\beta-\alpha} - c_p^{\beta-\alpha}| + C|x|^{r-2} + C.$$

By the mean value theorem,

$$\begin{aligned}
\psi_p(x) &\le -K|x|^{r-1}p^{-\tau}|y_p - c_p| + C|x|^{r-2} + C \\
&= -K|x|^r + C|x|^{r-2} + C,
\end{aligned}$$

which completes the proof.

The following two lemmas provide the necessary estimates on $A_p\varphi$ in the extreme regimes.

Lemma 7.3. *Suppose that $\beta < 1$. Let $\varphi(x) = |x|^r$, where $r \ge 2$, and let A_p be given by (7.1). Then there exist $p_0 > 0$, $M < \infty$, and $K > 0$, such that*

$$A_p\varphi(x) \le -K|x|^{(r-1)\wedge(r-1+\beta)}$$

for all $p \le p_0$ and all $x \in E_p$ satisfying $p^\tau x + c_p > M$.

Proof. Let $p \le p_0$ and $y_p = p^\tau x + c_p > M$. If p_0 is sufficiently small and M is sufficiently large, then $x \ge Kp^{-\tau}$ and $y_p \le x$. By (7.2),

$$\begin{aligned}
A_p\varphi(x) &\le -K|x|^{r-1}y_p^\beta + C|x|^{r-2}y_p^{2\beta} + C y_p^{r\beta} \\
&= -|x|^{r-1}y_p^\beta(K - C|x|^{-1}y_p^\beta - C|x|^{-r+1}y_p^{\beta(r-1)}).
\end{aligned}$$

If $\beta \le 0$, then, for sufficiently small p,

$$A_p\varphi(x) \le -|x|^{r-1}y_p^\beta(K - C|x|^{-1} - C|x|^{-r+1}) \le -K|x|^{r-1+\beta}.$$

https://doi.org/10.1239/jap/1189717533 Published online by Cambridge University Press

If $\beta > 0$, then

$$A_p\varphi(x) \leq -|x|^{r-1} y_p^{\beta}(K - C|x|^{\beta-1} - C|x|^{(\beta-1)(r-1)}),$$

so, for sufficiently small p, $A_p\varphi(x) \leq -K|x|^{r-1} y_p^{\beta} \leq -K|x|^{r-1}$.

Lemma 7.4. *Suppose that $\beta < 1$. Let $\varphi(x) = |x|^r$, where $r \geq 2$, and let A_p be given by (7.1). Then there exist $p_0 > 0$, $\delta > 0$, and $K > 0$ such that*

$$A_p\varphi(x) \leq -K|x|^{r\wedge(r-2\alpha/(1-\beta))}$$

for all $p \leq p_0$ and all $x \in E_p$ satisfying $p^{\tau}x + c_p < \delta$.

Proof. Let $p \leq p_0$ and $y_p = p^{\tau}x + c_p < \delta$. Note that since $x \in E_p$, $y_p \geq p^{\gamma}\ell$. If p_0 and δ are sufficiently small, then $x < 0$ and $Kp^{-\tau} \leq |x| \leq Cp^{-\tau}$. By (7.2), for sufficiently small δ,

$$\begin{aligned} A_p\varphi(x) &\leq -|x|^r y_p^{\alpha}(K|y_p^{\beta-\alpha} - c_p^{\beta-\alpha}| - C(p^{2\tau} y_p^{2\beta-\alpha} + p^{2\tau+1} y_p^{\alpha}) \\ &\qquad - C(p^{\tau r+\tau(r-2)} y_p^{r\beta-\alpha} + p^{\tau r+r-1+\tau(r-2)} y_p^{r\alpha-\alpha})) \\ &\leq -|x|^r y_p^{\alpha}(K - C(p^{2\tau} y_p^{2\beta-\alpha} + p^{2\tau(r-1)} y_p^{r\beta-\alpha}) \\ &\qquad - C(p^{2\tau+1} y_p^{\alpha} + p^{(2\tau+1)(r-1)} y_p^{\alpha(r-1)})). \end{aligned}$$

Let us first estimate the term $p^{2\tau} y_p^{2\beta-\alpha}$. If $2\beta-\alpha \geq 0$, then $p^{2\tau} y_p^{2\beta-\alpha} \leq Cp^{2\tau}$. If $2\beta-\alpha < 0$, then $p^{2\tau} y_p^{2\beta-\alpha} \leq Cp^{2\tau+\gamma(2\beta-\alpha)}$. Note that $2\tau + \gamma(2\beta-\alpha) = \gamma + 1$. Hence, for all values of α and β, there exist some $s > 0$ such that $p^{2\tau} y_p^{2\beta-\alpha} \leq p^s$.

Similarly, for the remaining terms in the above inequality, we observe that

$$\begin{gathered} 2\tau(r-1) + \gamma(r\beta-\alpha) = (2\tau+\gamma\beta)(r-1) + 1 = \gamma(r-1) + 1, \\ 2\tau + 1 + \gamma\alpha = \gamma, \quad \text{and} \quad (2\tau+1)(r-1) + \gamma\alpha(r-1) = \gamma(r-1). \end{gathered}$$

Therefore, if p_0 is sufficiently small, then $A_p\varphi(x) \leq -K|x|^r y_p^{\alpha}$. If $\alpha < 0$, then $A_p\varphi(x) \leq -K|x|^r$. If $\alpha \geq 0$, then

$$A_p\varphi(x) \leq -K|x|^r p^{\gamma\alpha} \leq -K|x|^{r-\gamma\alpha/\tau}.$$

Since $\gamma\alpha/\tau = 2\alpha/(1-\beta)$, this completes the proof.

Proof of Theorem 3.5. Suppose that $\beta < 1$ and that η_p is the stationary distribution for $\{p^{-\tau}(p^{\gamma}W_n - c_p)\}$. Then η_p is the stationary distribution for $\hat{\xi}_p$. Let $\varphi(x) = |x|^r$, where $r \geq 2$. By Lemmas 7.2, 7.3, and 7.4, if r is sufficiently large, there exist a $p_0 > 0$ and strictly positive, finite constants C and K such that

$$A_p\varphi(x) \leq C - K|x|^s$$

for some $s > 0$ and all $p \leq p_0$ and $x \in E_p$. As in the proof of Theorem 3.2, this implies that the family of measures $\{\eta_p\}$ is relatively compact on $\mathbb{R}$. By passing to a subsequence, we can assume that $\eta_p \Rightarrow \eta$ for some probability measure η on $\mathbb{R}$.

Let $p^{-\tau}(p^{\gamma}W_0 - c_p)$ have distribution η_p, let Z_p be given by (1.2), and let ξ_p be given by (1.6) with $\zeta_p \equiv c_p$. Note that $\xi_p(0)$ converges in distribution, so $p^{\tau}\xi_p(0) = Z_p(0) - \zeta_p(0) \to 0$ in probability. Hence, by Theorem 3.4, $\xi_p \Rightarrow \xi$, where ξ satisfies (1.10) with $\mathrm{P}\,\xi(0)^{-1} = \eta$. As in the proof of Theorem 3.2, ξ is a stationary process, so η is the stationary distribution for ξ.

Acknowledgements

The authors gratefully acknowledge Tom Kurtz, Ruth Williams, and Timo Seppäläinen for their support and helpful ideas. Thanks also to an anonymous referee for his or her helpful suggestions. This work was supported in part by the VIGRE grant of University of Wisconsin-Madison.

References

[1] Altman, E., Avrachenkov, K. and Barakat, C. (2000). A stochastic model of TCP/IP with stationary random losses. *ACM SIGCOMM Comput. Commun. Rev.* **30,** 231–242.

[2] Altman, E., Avrachenkov, K. and Barakat, C. (2002). TCP network calculus: the case of large delay-bandwidth product. In *Proc. IEEE INFOCOM 2002*, pp. 417–426.

[3] Altman, E., Avrachenkov, K. and Prabhu, B. (2005). Fairness in MIMD congestion control algorithms. *Proc. IEEE INFOCOM 2005*, pp. 1350–1361.

[4] Altman, E., Barakat, C. and Ramos, V. M. (2005). Analysis of AIMD protocols over paths with variable delay. *Comput. Commun.* **28,** 1605–1617.

[5] Altman, E., Jimenez, T. and Kofman, D. (2004). DPS queues with stationary ergodic service times and the performance of TCP in overload. In *Proc. IEEE INFOCOM 2004*, pp. 975–983.

[6] Altman, E., Avrachenkov, K., Barakat, C. and Nunez-Queija, R. (2001). TCP modeling in the presence of nonlinear window growth. In *Proc. ITC-17 2001*, pp. 883–894.

[7] Altman, E., Avrachenkov, K., Kherani, A. and Prabhu, B. (2005). Performance analysis and stochastic stability of congestion control protocols. In *Proc. IEEE INFOCOM 2005*, pp. 1316–1327.

[8] Altman, E. *et al.* (2004). Analysis of scalable TCP. In *Proc. IEEE HSNMC 2004*, pp. 51–62.

[9] Baccelli, F., McDonald, D. and Reynier, J. (2002). A mean-field model for multiple TCP connections through a buffer implementing RED. *Performance Evaluation* **49,** 77–97.

[10] Baccelli, F., Chaintreau, A., de Vleeschauwer, D. and McDonald, D. (2004). A mean-field analysis of short lived interacting TCP flows. In *Proc. ACM SIGMETRICS 2004*, ACM, New York, pp. 343–354.

[11] Baras, J., Misra, A. and Ott, T. J. (1999). The window distribution of multiple TCPs with random loss queues. In *Proc. Globecomm'99*, pp. 1714–1726.

[12] Baras, J., Misra, A. and Ott, T. J. (2000). Generalized TCP congestion avoidance and its effect on bandwidth sharing and variability. In *Proc. Globecomm 2000*, pp. 329–337.

[13] Baras, J., Misra, A. and Ott, T. J. (2000). Using drop-biasing to stabilize the occupancy of random drop queues with TCP traffic. In *Proc. ICCS 2000*.

[14] Baras, J., Misra, A. and Ott, T. J. (2002). Predicting bottleneck bandwidth sharing by generalized TCP flows. *Comput. Networks* **40,** 557–576.

[15] Bohacek, S. (2003). A stochastic model of TCP and fair video transmission. In *Proc. IEEE INFOCOM 2003*, pp. 1134–1144.

[16] Budhiraja, A., Hernandez-Campos F., Kulkarni, V. G. and Smith, F. D. (2004). Stochastic differential equation for TCP window size: analysis and experimental validation. *Prob. Eng. Inf. Sci.* **18,** 111–140.

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

[18] Ethier, S. N. and Kurtz, T. G. (1986). *Markov Processes: Characterization and Convergence*. John Wiley, New York.

[19] Floyd, S. (1994). TCP and explicit congestion notification. *ACM Comput. Commun. Rev.* **21,** 8–23.

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

[21] Hollot, C., Misra, V., Towsley, D. and Gong, W. B. (2001). A control theoretic analysis of RED. In *Proc. IEEE INFOCOM 2001*, pp. 1510–1519.

[22] Kelly, C. T. (2003). Scalable TCP: improving performance in high speed wide area networks. *ACM SIGCOMM Comput. Commun. Rev.* **32,** 83–91.

[23] Kelly, C. T. (2004). Engineering flow controls in the internet. Doctoral Thesis, Cambridge University.

[24] Kurtz, T. G. and Protter, P. (1991). Weak limit theorems for stochastic integrals and stochastic differential equations. *Ann. Prob.* **19,** 1035–1070.

[25] Lakshman, T. V. and Madhow, U. (1997). The performance of networks with high bandwidth-delay products and random loss. *IEEE/ACM Trans. Networking* **5,** 336–350.

[26] Marquez, R., Altman, E. and Sole-Alvarez, S. (2004). Modeling TCP and high speed TCP: a nonlinear extension to AIMD mechanisms. In *Proc. IEEE HSNMC 2004*, pp. 132–143.

[27] MATHIS, M., SEMKE, J., MAHDAVI, J. AND OTT, T. J. (1997). The macroscopic behavior of the TCP congestion avoidance algorithm. *ACM SIGCOMM Comput. Commun. Rev.* **27,** 67–82.
[28] MISRA, A. AND OTT, T. J. (1999). The window distribution of idealized TCP congestion avoidance with variable packet loss. In *Proc. IEEE INFOCOM 1999*, pp. 1564–1572.
[29] MISRA, A. AND OTT, T. J. (2001). Effect of exponential averaging on the variability of a RED queue. In *Proc. ICC 2001*, pp. 1817–1823.
[30] MISRA, A. AND OTT, T. J. (2001). Jointly coordinating ECN and TCP for rapid adaptation to varying bandwidth. In *Proc. MILCOM 2001*, pp. 719–725.
[31] MISRA, A. AND OTT, T. J. (2003). Performance sensitivity and fairness of ECN-aware 'modified TCP'. *J. Performance Evaluation*, **53,** 255–272.
[32] MISRA, V., GONG, W. B. AND TOWSLEY, D. (1999). Stochastic differential equation modeling and analysis of TCP-windowsize behavior. In *Proc. IFIP WG 7.3 Performance 1999*.
[33] OTT, T. J. (1999). ECN protocols and the TCP paradigm. Preprint. Available at www.teunisott.com.
[34] OTT, T. J. (2005). Transport protocols in the TCP paradigm and their performance. *Telecommun. Systems* **30,** 351–385.
[35] OTT, T. J. (2006). On the Ornstein–Uhlenbeck process with delayed feedback. Preprint. Available at www.teunisott.com.
[36] OTT, T. J. (2006). Rate of convergence for the 'square root formula'. *Adv. Appl. Prob.* **38,** 1132–1154.
[37] OTT, T. J. AND KEMPERMAN, J. H. B. (2007). The transient behavior of processes in the TCP paradigm. Work in Progress. Available at www.teunisott.com.
[38] OTT, T. J. AND SWANSON, J. (2006). Stationarity of some processes in transport protocols. *SIGMETRICS Perf. Eval. Rev.* **34,** 30–32.
[39] OTT, T. J., KEMPERMAN, J. H. B. AND MATHIS, M. (1996). The stationary behavior of idealized TCP congestion behavior. Preprint. Available at www.teunisott.com.
[40] OTT, T. J., LAKSHMAN, T. V. AND WONG, L. H. (1999). SRED: stabilized RED. In *Proc. IEEE INFOCOM 1999*, pp. 1346–1355.
[41] PADHYE, J., FIROIU, V., TOWSLEY, D. AND KUROSE, J. (1998). Modeling TCP throughput: a simple model and its empirical validation. In *Proc. ACM SIGCOMM 1998*, pp. 303–314.
[42] PADHYE, J., FIROIU, V., TOWSLEY, D. AND KUROSE, J. (2000). Modeling of TCP Reno performance: a simple model and its empirical validation. *IEEE/ACM Trans. Networking* **8,** 133–145.
[43] PROTTER, P. E. (2004). *Stochastic Integration and Differential Equations*, 2nd edn. Springer, Berlin.
[44] RAMAKRISHNAN, K. K., FLOYD, S. AND BLACK, D. (2001). The addition of explicit congestion control (ECN) to IP. In *Proc. IETF RFC 3168 2001*.