Technical Research Report

On the behavior of ECN/RED gateways under a large number of TCP flows: Limit theorems

by Peerapol Tinnakornsrisuphap, Armand M. Makowski

CSHCN TR 2005-3
(ISR TR 2005-80)

On the behavior of ECN/RED gateways under a large number of TCP flows: Limit theorems *†

Peerapol Tinnakornsrisuphap ‡
and
Armand M. Makowski §

Abstract

We consider a stochastic model of an ECN/RED gateway with competing TCP sources sharing the capacity. As the number of competing flows becomes large, the asymptotic queue behavior at the gateway can be described by a simple recursion and the throughput behavior of individual TCP flows becomes asymptotically independent.

In addition, a Central Limit Theorem complement is presented, yielding a more accurate characterization of the asymptotic queue. These results suggest a scalable yet accurate model of the complex large-scale stochastic feedback system, and crisply reveal the sources of queue fluctuations.

*This work was prepared through collaborative participation in the Communications and Networks Consortium sponsored by the U. S. Army Research Laboratory under the Collaborative Technology Alliance Program, Cooperative Agreement DAAD19-01-2-0011. The U. S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation thereon. This material is based upon work supported by the Space and Naval Warfare Systems Center- San Diego under Contract No: N66001-00-C-8063.

†The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U. S. Government.

‡Qualcomm Inc. 5775 Morehouse Drive San Diego, CA, 92121. Email: peerapol@qualcomm.com

§Department of Electrical and Computer Engineering, and Institute for Systems Research, University of Maryland, College Park, MD 20742. Email: armand@isr.umd.edu

Key words: Limit Theorems, CLT correction, TCP congestion-control, RED AQM mechanism

1 Introduction

One of the key mechanisms for operating the best-effort service Internet is the *additive increase/multiplicative decrease* (AIMD) congestion-control mechanism in the transmission control protocol (TCP) [8]. The AIMD algorithm enables TCP congestion-control to be robust under diverse conditions. Unfortunately, its self-clocking feedback mechanism is highly nonlinear and induces complex behavior in network traffic. There has been a number of studies trying to gain insights into this complex behavior.

While the relationship between the throughput of a single TCP flow and its round-trip and loss probability is fairly well understood [1, 12, 13, 15], these models are not suitable for the analysis of *many* TCP flows competing for the bandwidth of a link. Typically, with each TCP flow modeled in great details, the size of the state space for the model explodes when the number of flows becomes large, and the analysis then becomes intractable. Even numerical calculations or simulations of such models are very complicated and become computationally prohibitive, thereby providing no additional advantages over full-scale (system) simulation with existing simulation packages (e.g., NS [14]). To be sure, certain simplifying assumptions could be made, but it is not clear from the onset which details can be omitted without reducing the predictive power of the model.

To make matters worse, recent developments in Active Queue Management (AQM) techniques have introduced additional complexity in transport protocols. The development of AQM was prompted by the observation that with simple Tail-Drop gateways, TCP congestion-control leads to undesirable behavior, i.e., global synchronization [21]. When several TCP flows compete for bandwidth in a Tail-Drop gateway, it has been observed experimentally that packets from many flows are usually discarded simultaneously, resulting in a poor utilization of the network. AQM mechanisms such as Random Early Detection (RED) [6] and Explicit Congestion Notification (ECN) [5] have been proposed to help alleviate this problem by randomly marking packets with probability depending on queue size. TCP, in turn, triggers its multiplicative decrease mechanism upon receiving marked packets. This allows each TCP flow to react early to the growing congestion, thereby avoiding heavy congestion and preventing global synchronization. As can easily be imagined, the introduction of AQM further exacerbates the difficulty of understanding issues associated with buffer behavior and aggregate TCP traffic.

The interaction between TCP and AQM mechanisms are typically investigated in the framework of utility maximization problems [10, 11]. While there are many advantages of viewing the system as a utility maximization problem and of modeling TCP traffic as distributed algorithms for solving it, there are definite drawbacks as well to this approach. The most glaring is the absence of the probabilistic nature of AQM mechanisms in the model – the signaling mechanisms from AQM to TCP flows are usually modeled as feedback gains or penalty functions in this framework. Additionally, while the solution to the maximization problem might accurately describe the steady-state solution of TCP, the distributed solution does not necessarily capture the short-term dynamics of TCP well due to the absence of the packet-level window mechanism of TCP. The transient behavior of the network traffic is much more crucial in the actual system due to the highly heterogeneous and time-varying nature of Internet.

Recently, there has been a growing interest in *macroscale* modeling of TCP flows, as opposed to microscale models where each TCP flow is modeled in detail. Macroscale TCP models can be developed by systematically applying limit theorems to derive a limiting traffic model when the number of TCP flows is large. The potential benefits of doing so are three-fold. First, model simplification (with the promise of scalability) typically occurs when applying limit theorems, with irrelevant details filtered out without relying on ad-hoc assumptions. Second, limit theorems are central to the modern Theory of Probability, and as such have been the focus of a huge literature that contains a large number of results and techniques. Given this large body of knowledge, it is reasonable to expect the existence of suitable limit theorems (under very weak assumptions) which can be applied to the situation of interest. Finally, in the networking context, resource allocation problems are most pressing in networks operating at high utilization, e.g., when the number of users is large in relation to available resources. In such a scenario, the limiting model will become increasingly more accurate as the number of users increases.

Limit theorems for a bottleneck queue under a large number of rate-controlled TCP-like flows have been recently considered by a number of authors [2, 7, 16]; a survey of the relevant literature can be found in [18].

While these models already suggest some of the preliminary results to be expected from aggregating a large number of TCP flows, they all lack the notion of "packets" and the explicit window mechanism of TCP congestion-control which relies on packet-level operations. Further, they only consider the mean asymptotic queue size which lacks the finer description of the queue distribution necessary for network dimensioning. In this paper, we incorporate the TCP window mechanism explicitly to model the ECN-capable TCP congestion-control mechanism competing for bandwidth in a RED gateway. Our model uses RED as the AQM mechanism because it is the simplest and most widely-deployed AQM mechanism. However, the analytical results in this paper can be generalized to a large class of generic window-based congestion-control mechanisms and probabilistic AQM schemes, e.g., see the recent extensions in [17].

We establish several asymptotics when the number of flows is large, namely a Law of Large Numbers (LLNs) for the aggregate traffic into the RED buffer and a basic limit theorem for the normalized queue size [Section 3]. We sharpen these results with a Central Limit Theorem (CLT) complement [Section 4]. These results were announced mostly without proofs in the conference paper [19]. They crisply reveal the relationship between RED buffers and the probabilistic marking mechanism in RED. The CLT result presented here is the first of its kind in the literature on TCP/AQM modeling and can assist in the network dimensioning problem by establishing a probability distribution on the buffer utilization in RED gateways.

The paper is organized as follows. In Section 2 the model is described in detail, and a first set of asymptotic results are presented in Section 3, and Section 4 contains the Central Limit Theorem complement. A discussion of the results is given in Section 5. The analysis starts with some useful facts in Section 6. This is followed by an outline of the proofs of the the LLNs and CLT in Sections 7 and 8, respectively. The details of the proof of the LLNs are provided in Appendix A. The proof of the CLT complement is more elaborate and occupies the remainder of the paper; it is presented throughout Sections 9–14. Additional technical details have been relegated to Appendices B and C.

A word on the notation in use: Vectors are understood as *row* vectors. Equivalence in law or in distribution between random variables (rvs) is denoted by $=_{st}$. The indicator function of an event A is simply $\mathbf{1}[A]$, and we use $\xrightarrow{P}_n$ (resp. $\Longrightarrow_n$) to denote convergence in probability (resp. weak convergence or convergence in distribution) with n going to infinity. For any scalar x, we use the notation $(x)^+$ to represent $\max(x, 0)$.

2 A discrete-time model

2.1 A brief review of TCP + ECN/RED dynamics

We focus on a given flow transiting through a (bottleneck) node and on the size of its congestion window, i.e., the amount of unacknowledged packets in the network per round-trip. This size is dynamically adjusted by the following TCP congestion-control algorithm [8]: In a round-trip, if all the packets transmitted are not marked, then the size of the congestion window is increased by one packet for the next round-trip. On the other hand, if at least one packet is marked in the round-trip, the congestion window is halved. The probability that the router will mark packets in the buffer depends on the average queue length at the time of packet arrival. The average queue length is calculated by an exponential average filter with large memory to prevent RED from reacting too fast. As a result, consecutive incoming packets into RED are marked with almost identical probability. With this in mind, we now construct a model whose dynamics are similar in spirit to the dynamics of TCP + ECN/RED.

2.2 The discrete-time model

Time is assumed discrete and slotted in contiguous timeslots of duration equal to the round-trip delay of TCP connections. We consider N traffic sources, all transmitting through a bottleneck RED gateway with ECN enabled in both TCP and RED. The bottleneck RED gateway has capacity NC packets/slot for some positive constant C. The RED buffer is modeled as an infinite queue, so that no packet losses occur due to buffer overflow, and congestion-control is achieved solely through the random marking algorithm in the RED gateway.

Fix $N = 1, 2, \ldots$ and $t = 0, 1, \ldots$. We write $X^{(N)}$ to indicate the explicit dependence of the quantity X on the number N of connections.

2.3 Dynamics

Suppose that each of the N TCP sources has an infinite amount of data to transmit and that in each timeslot it transmits as much as allowed by its congestion window in that timeslot. So, for $i = 1, \ldots, N$, let $W_i^{(N)}(t)$ be an integer-valued rv that encodes the number of packets generated by source i (and hence its congestion window) at the beginning of the timeslot $[t, t+1)$. The integer $W_i^{(N)}(t)$ is assumed to be in the range $\{1, \ldots, W_{\max}\}$ for some finite integer $W_{\max}$, i.e.,

$$W_i^{(N)}(t) \in \{1, \ldots, W_{\max}\} \tag{1}$$

with $W_{\max} \geq 2$ to avoid trivial and uninteresting situations.

Given that N sources are active, the total number of packets which are accepted into the RED buffer at the beginning of timeslot $[t, t+1)$ is given by

$$A^{(N)}(t) = \sum_{i=1}^{N} W_i^{(N)}(t). \tag{2}$$

If $Q^{(N)}(t)$ denotes the number of packets in the buffer at the beginning of the timeslot $[t, t+1)$, then $Q^{(N)}(t) + A^{(N)}(t)$ packets are available for transmission in that timeslot. Since the outgoing link operates at the rate of NC packets/timeslot, $\left(Q^{(N)}(t) + A^{(N)}(t) - NC\right)^+$ packets will not be transmitted during timeslot $[t, t+1]$, and remain in the buffer, their transmission being deferred to subsequent timeslots. The number $Q^{(N)}(t+1)$ of packets in the buffer at the beginning of the timeslot $[t+1, t+2)$ is then given by

$$Q^{(N)}(t+1) = \left(Q^{(N)}(t) - NC + A^{(N)}(t)\right)^+ . \tag{3}$$

Upon arrival at the RED gateway, each packet from source i may be marked according to a random marking algorithm (to be specified shortly). We represent this possibility by the $\{0,1\}$-valued rv $M_{i,j}^{(N)}(t+1)$ (with $j = 1, ..., W_i^{(N)}(t)$) with the interpretation that $M_{i,j}^{(N)}(t+1) = 0$ (resp. $M_{i,j}^{(N)}(t+1) = 1$) if the jth packet from source i is marked (resp. not marked) in the RED buffer. Next we introduce the rvs

$$M_i^{(N)}(t+1) = \prod_{j=1}^{W_i^{(N)}(t)} M_{i,j}^{(N)}(t+1) \tag{4}$$

so that $M_i^{(N)}(t+1) = 1$ (resp. $M_i^{(N)}(t+1) = 0$) corresponds to the event that no packet (resp. at least one packet) from source i has been marked in timeslot $[t, t+1)$. The evolution of the window mechanism for source i can now be described through the recursion

$$\begin{aligned} W_i^{(N)}(t+1) &= \min\left(W_i^{(N)}(t) + 1, W_{\max}\right) M_i^{(N)}(t+1) \\ &\quad + \min\left(\lceil \frac{W_i^{(N)}(t)}{2} \rceil, W_{\max}\right)(1 - M_i^{(N)}(t+1)). \end{aligned} \tag{5}$$

This equation emulates the interaction between TCP and RED as follows: If no packet from source i is marked in the timeslot $[t, t+1)$, then the congestion window size in the next timeslot is increased by one packet. On the other hand, if one or more packets are marked in the timeslot $[t, t+1)$, then the congestion window in the next timeslot is reduced by half. The size of the congestion window is limited by the maximum window size $W_{\max}$ [1].

2.4 Statistical assumptions

In order to fully specify the model, we need to specify the *joint* statistics of the rvs

$$\{M_{i,j}^{(N)}(t+1), M_i^{(N)}(t),\ i = 1, \ldots, N;\ j = 1, 2, \ldots;\ t = 0, 1, \ldots\}.$$

To do so we introduce the collection of i.i.d. $[0,1]$-*uniform* rvs $\{V_i(t+1),\ V_{i,j}(t+1),\ i,j = 1, \ldots;\ t = 0, 1, \ldots\}$ [2] which are assumed independent of the rvs $Q^{(N)}(0)$ and $W_1^{(N)}(0), \ldots, W_N^{(N)}(0)$. We also introduce a mapping $f^{(N)} : \mathbb{R}_+ \to [0,1]$ which acts as the *marking probability* function of the RED gateway.

The process by which packets are marked is described first: For each $i = 1, \ldots, N$, we define the marking rvs

$$M_{i,j}^{(N)}(t+1) = \mathbf{1}\left[V_{i,j}(t+1) > f^{(N)}(Q^{(N)}(t))\right], \quad j = 1, 2, \ldots \tag{6}$$

so that the rv $M_{i,j}^{(N)}(t+1)$ is the indicator function of the event that the jth packet from source i is *not* marked in the timeslot $[t, t+1)$. Thus, in

[1] If $W_i^{(N)}(0)$ lies in the range $\{1, \ldots, W_{\max}\}$ for each $i = 1, \ldots, N$, then so does $W_i^{(N)}(t)$ for each $t = 0, 1, \ldots$ and the minimum with $W_{\max}$ in the second term of (5) can be omitted.

[2] The need for the sequence $\{V_i(t+1),\ i = 1, \ldots;\ t = 0, 1, \ldots\}$ will become apparent at a later stage in the discussion.

a round-trip, each packet coming into the router is marked with identical (conditional) probability which depends only on the queue length at the beginning of the timeslot. This model approximates the case where the memory of the queue averaging mechanism is long, which is the case for the recommended parameter settings of RED [6].

3 Asymptotics via the Law of Large Numbers

The first result of the paper consists in the asymptotics for the normalized buffer content as the number N of sources becomes large. This result, contained in Theorem 3.1 below, is discussed under the following assumptions (A1)-(A2):

(A1) There exists a continuous function $f : \mathbb{R}_+ \to [0,1]$ such that for each $N = 1, 2, \ldots$,

$$f^{(N)}(x) = f(N^{-1}x), \quad x \geq 0;$$

(A2) For each $N = 1, 2, \ldots$, the dynamics (3) and (5) start with the conditions

$$Q^{(N)}(0) = 0 \quad \text{and} \quad W_i^{(N)}(0) = W, \quad i = 1, \ldots, N$$

for some integer W in the range $\{1, \ldots, W_{\max}\}$.

Assumption (A1) is a structural condition. Since we are interested in a "snapshot" of the dynamics when N flows exists in the system, then f is just a surrogate function representing the average contribution that each flow has on the marking probability. Meanwhile, Assumption (A2) is made essentially for technical convenience as it implies that for each $N = 1, 2, \ldots$ and all $t = 0, 1, \ldots$, the rvs $W_1^{(N)}(t), \ldots, W_N^{(N)}(t)$ are *exchangeable*. This assumption can be omitted but at the expense of a more cumbersome discussion.

Theorem 3.1. *Assume (A1)-(A2) to hold. Then, for each $t = 0, 1, \ldots$, there exist a (non-random) constant $q(t)$ and an $\{1, \ldots, W_{\max}\}$-valued rv $W(t)$ such that the following holds:*

(i) The convergence results

$$\frac{Q^{(N)}(t)}{N} \xrightarrow{P}_N q(t) \quad \text{and} \quad W_1^{(N)}(t) \Longrightarrow_N W(t) \tag{7}$$

take place;

(ii) For any function $g : \mathbb{N} \to \mathbb{R}$,

$$\frac{1}{N} \sum_{i=1}^{N} g(W_i^{(N)}(t)) \xrightarrow{P}_N \mathbf{E}\left[g(W(t))\right]. \tag{8}$$

(iii) For any integer $I = 1, 2, \ldots$, the rvs $\{W_i^{(N)}(t),\ i = 1, \ldots, I\}$ become asymptotically independent as N becomes large, with

$$\lim_{N\to\infty} \mathbf{P}\left[W_i^{(N)}(t) = k_i,\ i = 1, \ldots, I\right] = \prod_{i=1}^{I} \mathbf{P}\left[W(t) = k_i\right] \tag{9}$$

for any $k_1, \ldots, k_I$ in $\mathbb{N}$

Moreover, with initial conditions $q(0) = 0$ and $W(0) = W$, it holds that

$$q(t+1) = (q(t) - C + \mathbf{E}\left[W(t)\right])^{+} \tag{10}$$

and

$$\begin{aligned} W(t+1) \quad =_{st} \quad & \min\left(W(t) + 1, W_{\max}\right) M(t+1) \\ & + \min\left(\lceil \frac{W(t)}{2} \rceil, W_{\max}\right)(1 - M(t+1)) \end{aligned} \tag{11}$$

where

$$M(t+1) = \mathbf{1}\left[V(t+1) \le (1 - f(q(t)))^{W(t)}\right] \tag{12}$$

for i.i.d. $[0,1]$-uniform rvs $\{V(t+1),\ t = 0, 1, \ldots\}$.

A proof of Theorem 3.1 is given in Section 7. As will become apparent from the discussion given there, Theorem 3.1 readily flows from a Weak Law of Large Numbers [Claim (ii)] for the triangular array

$$\{W_i^{(N)}(t),\ i = 1, \ldots, N; N = 1, 2, \ldots\}.$$

We close this section with a simple but useful consequence of Claim (i) of Theorem 3.1, namely that

$$f\left(\frac{Q^{(N)}(t)}{N}\right) \xrightarrow{P}_N f(q(t)) \tag{13}$$

under the continuity assumption on f.

4 A Central Limit complement

In this section, we present a Central Limit Theorem (CLT) which complements the limiting results obtained earlier. The discussion is carried out under the same setup as in Section 3, but with Assumption (A1) strengthened to read as Assumption (A1b), where

(A1b) Assumption (A1) holds with mapping $f : \mathbb{R}_+ \to [0,1]$ which is continuously differentiable, i.e., its derivative $f' : \mathbb{R}_+ \to \mathbb{R}$ exists and is continuous.

Fix $t = 0, 1, \ldots$. With the notation of Theorem 3.1, define

$$K(t) := C - q(t) - \mathbf{E}\left[W(t)\right]. \tag{14}$$

We can interpret $K(t)$ as the asymptotic residual capacity per user in the timeslot $[t, t+1)$. Moreover, for each $N = 1, \ldots$, set

$$L_0^{(N)}(t) := \frac{Q^{(N)}(t)}{N} - q(t) \tag{15}$$

and

$$L_{\mathrm{avg}}^{(N)}(t) := \frac{1}{N} \sum_{i=1}^{N} \left(W_i^{(N)}(t) - \mathbf{E}\left[W(t)\right] \right). \tag{16}$$

The result will be given a more compact form by using the mapping $\Phi : \mathbb{R}^2 \to \mathbb{R}$ given by

$$\Phi(K, x) := \begin{cases} 0 & \text{if } K > 0 \\ x^+ & \text{if } K = 0 \\ x & \text{if } K < 0 \end{cases} \tag{17}$$

for arbitrary x in $\mathbb{R}$.

Theorem 4.1. *Assume (A1b)-(A2) to hold. Then, for each* $t = 0, 1, \ldots$, *there exists an* $\mathbb{R}^2$*-valued rv* $(L_0(t), L_{\mathrm{avg}}(t))$ *such that the convergence*

$$\sqrt{N} \left(L_0^{(N)}(t), L_{\mathrm{avg}}^{(N)}(t) \right) \Longrightarrow_N (L_0(t), L_{\mathrm{avg}}(t)) \tag{18}$$

holds. Moreover, the distributional recurrence

$$L_0(t+1) =_{st} \Phi(K(t), L_0(t) + L_{\mathrm{avg}}(t)) \tag{19}$$

holds.

The convergence (18) suggests the distributional approximation

$$Q^{(N)}(t) \simeq_{st} Nq(t) + \sqrt{N}L_0(t) \tag{20}$$

and

$$\sum_{i=1}^{N} W_i^{(N)}(t) \simeq_{st} N\mathbf{E}\left[W(t)\right] + \sqrt{N}L_{\text{avg}}(t) \tag{21}$$

for large N. Given the interpretation of $K(t)$ as the asymptotic residual capacity per user in timeslot $[t, t+1)$, if there exists extra capacity for the average user rate to increase ($K(t) > 0$), then there is no fluctuation in the limiting queue. On the other hand, when there is congestion ($K(t) < 0$), the fluctuation has a (non-trivial) limiting distribution which can be characterized. Some technical difficulties arise in the special case $K(t) = 0$.

The proof of Theorem 4.1 is given at the end of Section 8, and relies on showing that some key convergence statements propagate over time. In the process we prove a lot more. In particular, if we specialize (40), one of the by-product of this analysis, to the mapping $g : \mathbb{N} \to \mathbb{R}$ given by $g(w) = w$, we find that $L_{\text{avg}}(t+1)$ is of the form

$$L_{\text{avg}}(t+1) =_{st} \Xi(t) - f'(q(t))R(t)L_0(t) + Y(t+1) \tag{22}$$

where $R(t)$ is a constant, $Y(t+1)$ is a zero-mean Gaussian rv independent of the pair of rvs $(L_0(t), \Xi(t))$ and the statistics of the rv $\Xi(t)$ are determined by $q(0), q(1), \ldots, q(t-1)$. The variance of $Y(t+1)$ is given in (41) (with $\hat{g}(x) = x$) and it follows from (64) (with $h(w) = w$) that

$$R(t) = \mathbf{E}\left[W(t)^2\left(1 - f(q(t))\right)^{W(t)-1}\right].$$

5 Discussion

Theorems 3.1 and 4.1 show that the dynamics of the queue at time t can be approximated by the recursion (20), where both $q(t)$ and $L_0(t)$ can be evaluated independently of the number of users. The approximation becomes more accurate as the number of users becomes large. This limiting model is "scalable" in that it does not suffer from state space explosion, nor does it require any ad-hoc assumptions in the analysis.

Claim (iii) of Theorem 3.1 also states that the dependency at each timeslot between the window size of each TCP connection becomes negligible under a large number of flows. This claim is in line with the commonly held belief and simulation results suggesting that "RED breaks global synchronization" [6].

A closer inspection into the distributional relation (22) reveals that the fluctuation $L_{\text{avg}}(t+1)$ in the input traffic during timeslot $[t+1, t+2)^3$ is composed of the following three distinct components:

(i) The term $-f'(q(t))R(t)L_0(t)$ represents the fluctuations caused by the discrepancy between the feedback information from RED to TCP sources $f^{(N)}\left(Q^{(N)}(t)\right)$ and the limiting feedback information $f(q(t))$. This uncertainty in feedback information manifests itself as $-f'(q(t))R(t)L_0(t)$ and can be explained by the well-known *Delta Method* [Section 12]. As the slope of the feedback function increases, the magnitude of fluctuation due to this component increases as well. This supports the observation that the magnitude of queue size oscillation at RED gateways increases with the slope of the marking probability function of RED mechanism [4];

(ii) The term $Y(t+1)$ represents the fluctuations caused by the difference between the (conditional) Bernoulli rvs representing feedback information available at the sources, i.e., the rvs $M_i^{(N)}(t+1), i = 1, 2, \ldots, N$, and the desired feedback information $(f^{(N)}(Q^{(N)}(t)))$ at RED. Recall that a TCP source can only react to whether the RED gateway marks a packet from this source or not in the previous round-trip. This binary nature of the feedback information imposes a limited feedback information granularity, and induces fluctuations in the input traffic. These fluctuations cannot be captured without taking into account the detailed packet-level operations of the congestion-control mechanism.

(iii) The term $\Xi(t)$ represents the fluctuations in the previous timeslots, i.e., from timeslot $[0,1)$ upto $[t-1,t)$, being propagated over to timeslot $[t+1, t+2)$.

Both the Monte-Carlo simulations of the model and the NS-2 simulations in [19] suggest that the limiting behavior of the queue follows the results in Theorem 3.1. Furthermore, the standard deviation of the normalized queue size at steady-state decreases with a rate which appears to be consistent with the prediction of the Central Limit Theorem, i.e., the standard deviation decreases with a rate $\frac{1}{\sqrt{N}}$.

[3] From (19) we recall that $L_{\text{avg}}(t+1)$ will be compounded into $L_0(t+2)$ if $K(t+1) \leq 0$, i.e., the fluctuations in the input traffic will cause the queue to fluctuate in the next timeslot.

6 Some useful relations

To facilitate the presentation of the proof of Theorem 3.1, we begin with several simple yet helpful facts. Throughout the discussion, for each $t = 0, 1, \ldots$, we shall find it useful to write

$$Z(t) := (1 - f(q(t)))^{W(t)} = \gamma(t)^{W(t)} \tag{23}$$

with

$$\gamma(t) := 1 - f(q(t)). \tag{24}$$

Consider an arbitrary mapping $g : \mathbb{N} \to \mathbb{R}^p$ for some positive integer p : With g we associate the bounded mappings $g^\star, g_\star : \mathbb{N} \to \mathbb{R}^p$ given by

$$g^\star(w) := g(\min(w + 1, W_{\max})), \quad w \in \mathbb{N} \tag{25}$$

and

$$g_\star(w) := g(\min\left(\lceil \frac{w}{2} \rceil, W_{\max}\right)), \quad w \in \mathbb{N}. \tag{26}$$

Fix $i = 1, \ldots, N$ and $t = 0, 1, \ldots$. It follows from (5) that

$$\begin{aligned} & g(W_i^{(N)}(t+1)) \\ = \quad & M_i^{(N)}(t+1) g^\star(W_i^{(N)}(t)) + (1 - M_i^{(N)}(t+1)) g_\star(W_i^{(N)}(t)). \end{aligned} \tag{27}$$

If $\mathcal{F}_t$ denotes the σ-field generated by the rvs

$$\{Q^{(N)}(0), W_i^{(N)}(0), V_i(s), V_{i,j}(s),\ i, j = 1, 2, \ldots;\ s = 1, \ldots, t\},$$

then the rvs $Q^{(N)}(t)$ and $W_i^{(N)}(t)$ $(i = 1, \ldots, N)$ are all $\mathcal{F}_t$-measurable. Hence, under the enforced independence assumptions, it holds that

$$\mathbf{E}\left[M_{i,j}^{(N)}(t+1) \middle| \mathcal{F}_t \right] = 1 - f^{(N)}(Q^{(N)}(t)), \quad j = 1, 2, \ldots$$

so that

$$\mathbf{E}\left[M_i^{(N)}(t+1) \middle| \mathcal{F}_t \right] = Z_i^{(N)}(t) \tag{28}$$

by conditional independence, where we have set

$$Z_i^{(N)}(t) := \left(1 - f^{(N)}(Q^{(N)}(t))\right)^{W_i^{(N)}(t)} = \gamma^{(N)}(t)^{W_i^{(N)}(t)} \tag{29}$$

with

$$\gamma^{(N)}(t) := 1 - f^{(N)}(Q^{(N)}(t)). \tag{30}$$

It readily follows from (27) that

$$\begin{aligned}\mathbf{E}\left[g(W_i^{(N)}(t+1))\middle|\mathcal{F}_t\right] &= Z_i^{(N)}(t)g^\star(W_i^{(N)}(t)) + (1-Z_i^{(N)}(t))g_\star(W_i^{(N)}(t)) \\ &= F_g(Z_i^{(N)}(t), W_i^{(N)}(t)) \qquad (31)\end{aligned}$$

where the mapping $F_g : [0,1] \times \mathbb{N} \to \mathbb{R}$ is associated with g through

$$F_g(z,w) = zg^\star(w) + (1-z)g_\star(w), \quad z \in [0,1],\ w \in \mathbb{N}. \qquad (32)$$

Upon taking expectations on both sides of (31), we finally obtain

$$\mathbf{E}\left[g(W_i^{(N)}(t+1))\right] = \mathbf{E}\left[F_g(Z_i^{(N)}(t), W_i^{(N)}(t))\right]. \qquad (33)$$

7 A proof of Theorem 3.1 (Outline)

For each $t = 0, 1, \ldots$, the statements **[A:t]**, **[B:t]**, **[C:t]** and **[D:t]** below refer to the following convergence statements:

[A:t] For some non-random $q(t)$, it holds that

$$\frac{Q^{(N)}(t)}{N} \xrightarrow{P}_N q(t); \qquad (34)$$

[B:t] For some $\{1, \ldots, W_{\max}\}$-valued rv $W(t)$, it holds that

$$W_1^{(N)}(t) \Longrightarrow_N W(t); \qquad (35)$$

[C:t] For any integer $I = 1, 2, \ldots$, the rvs $\{W_i^{(N)}(t),\ i = 1, \ldots, I\}$ become asymptotically independent with large N as described by (9) and the rv $W(t)$ is the one occurring in **[B:t]**;

[D:t] For any mapping $g : \mathbb{R} \to \mathbb{R}$, the convergence (8) holds with the rv $W(t)$ occurring in **[B:t]**.

Through a series of lemmas, we shall prove the validity of the statements **[A:t]**–**[D:t]** for all $t = 0, 1, \ldots$. We do so by induction on t and in the process we establish Theorem 3.1.

Lemma 7.1. *Under (A1), if* **[A:t]** *and* **[B:t]** *hold for some* $t = 0, 1, \ldots$, *then* **[B:t+1]** *holds with* $W(t+1)$ *related to* $W(t)$ *by (11).*

Lemma 7.2. *Under (A1), if* [**A:t**] *and* [**D:t**] *hold for some* $t = 0, 1, \ldots$, *then* [**A:t+1**] *also holds.*

Lemma 7.3. *Under (A1)–(A2), if* [**A:t**], [**B:t**] *and* [**C:t**] *hold for some* $t = 0, 1, \ldots$, *then* [**C:t+1**] *also holds.*

Lemma 7.4. *Under (A1)–(A2), if* [**A:t**], [**B:t**] *and* [**C:t**] *hold for some* $t = 0, 1, \ldots$, *then* [**D:t**] *holds.*

Lemmas 7.1–7.4 are proved in Appendix A. We now conclude with a proof of Theorem 1: Under (A1)-(A2) the statements [**A:t**]–[**D:t**] trivially hold for $t = 0$ with $q(0) = 0$ and $W(0) = W$. Moreover, if [**A:t**]–[**C:t**] hold for some $t = 0, 1, \ldots$, then so do the statements [**A:t+1**]–[**C:t+1**] by Lemmas 7.1, 7.2 and 7.3. Finally, both statements [**D:t**] and [**D:t+1**] hold by virtue of Lemma 7.4. Consequently, the statements [**A:t**]–[**D:t**] do hold for all $t = 0, 1, \ldots$ by induction, and the validity of Claims (i)-(iii) of Theorem 3.1 is established. The proof of Lemma 7.2 also shows (10), while (11)-(12) are already contained in Lemma 7.1.

We close with a result that builds on, and strengthens, Claim (ii) of Theorem 3.1. This result will be used in proving Theorem 4.1 later on, and its proof is available in Appendix B.

Proposition 7.5. *Assume (A1)-(A2) to hold. Then, for each* $t = 0, 1, \ldots$, *and any function* $g : \mathbb{N} \to \mathbb{R}$, *it holds that*

$$\frac{1}{N} \sum_{i=1}^{N} g(W_i^{(N)}(t)) Z_i^{(N)}(t)^\ell \xrightarrow{P}_N \mathbf{E}\left[g(W(t))Z(t)^\ell\right] \tag{36}$$

for each integer $\ell = 1, 2, \ldots$.

8 A proof of Theorem 4.1 (Outline)

As in the proof of Theorem 3.1, we proceed by induction on t with the help of a series of technical facts. Again the discussion is facilitated by introducing a number of auxiliary convergence statements to be propagated in time. With the aim to simplify the presentation, for arbitrary mapping $g : \mathbb{N} \to \mathbb{R}^p$ with positive integer p, we define

$$L_g^{(N)}(t) := \frac{1}{N} \left(\sum_{i=1}^{N} g(W_i^{(N)}(t)) - \mathbf{E}\left[g(W(t))\right] \right), \quad N = 1, 2, \ldots$$

for each $t = 0, 1, \ldots$. The rv $L^{(N)}_{\text{avg}}(t)$ corresponds to the choice $g(w) = w$.

For each $t = 0, 1, \ldots$, we introduce the auxiliary convergence statements **[E:t]** and **[F:t]**, where

[E:t] For arbitrary mapping $g : \mathbb{N} \to \mathbb{R}^p$ with positive integer p, there exists an $\mathbb{R}^{p+1}$-valued rv $(L_0(t), L_g(t))$ such that the joint convergence

$$\sqrt{N} \left(L^{(N)}_0(t), L^{(N)}_g(t) \right) \Longrightarrow_N (L_0(t), L_g(t)) \tag{37}$$

takes place;

[F:t] For arbitrary mapping $g : \mathbb{N} \to \mathbb{R}^p$ with positive integer p, there exists an $\mathbb{R}^{p+1}$-valued rv $(L_0(t+1), L_g(t))$ such that the joint convergence

$$\sqrt{N} \left(L^{(N)}_0(t+1), L^{(N)}_g(t) \right) \Longrightarrow_N (L_0(t+1), L_g(t)) \tag{38}$$

takes place, with

$$(L_0(t+1), L_g(t)) =_{st} \left(\Phi\left(K(t), L_0(t) + L_{\text{avg}}(t)\right), L_g(t)\right) \tag{39}$$

with mapping Φ defined at (17).

These convergence statements propagate in time as the discussion now shows:

Proposition 8.1. *Under (A1b)–(A2), if* **[E:t]** *holds for some* $t = 0, 1, \ldots$, *then* **[F:t]** *holds.*

Proposition 8.2. *Under (A1b)–(A2), if* **[E:t]** *holds for some* $t = 0, 1, \ldots$, *then* **[E:t+1]** *also holds.*

In the process of establishing Proposition 8.2, we obtain the following characterization of the limit (37) in **[E:t+1]**: For arbitrary mapping $g : \mathbb{N} \to \mathbb{R}^p$, the limiting $\mathbb{R}^{p+1}$-valued rv $(L_0(t+1), L_g(t+1))$ admits the following decomposition

$$\begin{aligned} & (L_0(t+1), L_g(t+1)) \\ =_{st} \; & \left(\Phi\left(K(t), L_0(t) + L_{\text{avg}}(t)\right), L_{\hat{g}_t}(t) - f'(q(t)) R_g(t) L_0(t) + Y_{\hat{g}}(t+1)\right) \end{aligned} \tag{40}$$

where $R_g(t)$ is the element in $\mathbb{R}^p$ introduced at (64) (with $h = g$) and the $\mathbb{R}^p$-valued rv $Y_{\hat{g}}(t+1)$ is a zero-mean Gaussian rv with covariance matrix

$$\Sigma_{\hat{g}}(t+1) := \mathbf{E}\left[\hat{g}(W(t))'\hat{g}(W(t))Z(t)(1-Z(t))\right]. \tag{41}$$

Moreover, this Gaussian rv is independent of the rvs $L_0(t)$, $L_{\mathrm{avg}}(t)$ and $L_{\hat{g}_t}(t)$, the mappings $\hat{g}$ and $\hat{g}_t$ being defined at (47) and (48), respectively.

We complete the proof of Theorem 4.1 by an easy induction argument: For $t = 0$, [**E:t**] trivially holds since for each $N = 1, 2, \ldots$, we have $W_i^{(N)}(0) = W$ for all $i = 1, \ldots, N$ and $W(0) = W$. It is now plain from Proposition 8.1 and Proposition 8.2 that [**E:t**] and [**F:t**] both hold for all $t = 0, 1, \ldots$, and Theorem 4.1 is established.

9 A proof of Proposition 8.1

Pick an arbitrary mapping $g : \mathbb{N} \to \mathbb{R}^p$ with positive integer p. Fix $t = 0, 1, \ldots$ and $N = 1, 2, \ldots$. We begin by noting that under [**E:t**], the mapping g being *arbitrary*, it is the case that there exists an $\mathbb{R}^{p+2}$-valued rv $(L_0(t), L_{\mathrm{avg}}(t), L_g(t))$ such that the joint convergence

$$\sqrt{N}\left(L_0^{(N)}(t), L_{\mathrm{avg}}^{(N)}(t), L_g^{(N)}(t)\right) \Longrightarrow_N (L_0(t), L_{\mathrm{avg}}(t), L_g(t)) \tag{42}$$

takes place. Indeed it suffices to use [**E:t**] with the mapping $w \to (w, g(w))$.

As we seek to identify $L_0(t+1)$, we rewrite the limiting recursion (10) in the form

$$q(t+1) = (q(t) - C + \mathbf{E}\left[W(t)\right])^+ = (-K(t))^+$$

with $K(t)$ given by (14). Combining this observation with the queue dynamics (3) gives

$$\begin{aligned}
L_0^{(N)}(t+1) &= \left(\frac{Q^{(N)}(t)}{N} - C + \frac{1}{N}\sum_{i=1}^{N} W_i^{(N)}(t)\right)^+ - (-K(t))^+ \\
&= \max\left(L_0^{(N)}(t) + \frac{1}{N}\sum_{i=1}^{N} W_i^{(N)}(t) - \mathbf{E}\left[W(t)\right], K(t)\right) - K(t)^+ \\
&= \max\left(L_0^{(N)}(t) + L_{\mathrm{avg}}^{(N)}(t), K(t)\right) - K(t)^+
\end{aligned}$$

so that

$$
\begin{aligned}
& \sqrt{N} L_0^{(N)}(t+1) \\
= \quad & \max\left(\sqrt{N}\left(L_0^{(N)}(t) + L_{\mathrm{avg}}^{(N)}(t)\right), \sqrt{N}K(t)\right) - \sqrt{N}K(t)^+.
\end{aligned} \tag{43}
$$

By the Continuous Mapping Theorem, we already conclude from (42) that

$$
\sqrt{N}\left(L_0^{(N)}(t) + L_{\mathrm{avg}}^{(N)}(t)\right) \Longrightarrow_N L_0(t) + L_{\mathrm{avg}}(t). \tag{44}
$$

Three cases emerge depending on the sign of $K(t)$. If $K(t) = 0$, then (43) reduces to

$$
\sqrt{N} L_0^{(N)}(t+1) \quad = \quad \left(\sqrt{N}\left(L_0^{(N)}(t) + L_{\mathrm{avg}}^{(N)}(t)\right)\right)^+
$$

and the convergence (44) yields

$$
\sqrt{N} L_0^{(N)}(t+1) \Longrightarrow_N \left(L_0(t) + L_{\mathrm{avg}}(t)\right)^+
$$

again by the Continuous Mapping Theorem.

If $K(t) < 0$, then (43) reduces to

$$
\sqrt{N} L_0^{(N)}(t+1) = \max\left(\sqrt{N}\left(L_0^{(N)}(t) + L_{\mathrm{avg}}^{(N)}(t)\right), -\sqrt{N}|K(t)|\right)
$$

and the convergence (44) yields

$$
\sqrt{N} L_0^{(N)}(t+1) \Longrightarrow_N L_0(t) + L_{\mathrm{avg}}(t)
$$

since $|K(t)| > 0$ guarantees $\lim_{N\to\infty} \sqrt{N}|K(t)| = \infty$.

Finally, if $K(t) > 0$, then (43) reduces to

$$
\sqrt{N} L_0^{(N)}(t+1) = \max\left(\sqrt{N}\left(L_0^{(N)}(t) + L_{\mathrm{avg}}(t)^{(N)}(t)\right) - \sqrt{N}K(t), 0\right)
$$

and the convergence (44) yields $\sqrt{N} L_0^{(N)}(t+1) \Longrightarrow_N 0$ since $\lim_{N\to\infty} \sqrt{N}K(t) = \infty$.

It follows from (42) that

$$
\sqrt{N}\left(L_0^{(N)}(t) + L_{\mathrm{avg}}^{(N)}(t), L_g^{(N)}(t)\right) \Longrightarrow_N \left(L_0(t) + L_{\mathrm{avg}}(t), L_g(t)\right), \tag{45}
$$

and the first part of the proof readily leads to (38) with the identification (39). This completes the proof of Proposition 8.1. ■

10 A key decomposition

To establish Proposition 8.2, we start with an arbitrary mapping $g : \mathbb{N} \to \mathbb{R}^p$ for some positive integer p, and recall the definitions (25) and (26) of the induced mappings $g^\star, g_\star : \mathbb{N} \to \mathbb{R}^p$.

Fix $N = 1, 2, \ldots$, $i = 1, \ldots, N$ and $t = 0, 1, \ldots$. Making use of (27) and (11) (in a similar fashion), we get

$$\begin{aligned}
& g(W_i^{(N)}(t+1)) - \mathbf{E}\left[g(W(t+1))\right] \\
= \;& M_i^{(N)}(t+1) g^\star(W_i^{(N)}(t)) + (1 - M_i^{(N)}(t+1)) g_\star(W_i^{(N)}(t)) \\
& - \mathbf{E}\left[M(t+1) g^\star(W(t)) + (1 - M(t+1)) g_\star(W(t))\right] \\
= \;& g_\star(W_i^{(N)}(t)) - \mathbf{E}\left[g_\star(W(t))\right] \\
& + \left(g^\star(W_i^{(N)}(t)) - g_\star(W_i^{(N)}(t))\right) M_i^{(N)}(t+1) \\
& - \mathbf{E}\left[\left(g^\star(W(t)) - g_\star(W(t))\right) M(t+1)\right].
\end{aligned} \tag{46}$$

With the mapping $\hat{g} : \mathbb{N} \to \mathbb{R}^p$ defined by

$$\hat{g}(w) = g^\star(w) - g_\star(w), \quad w \in \mathbb{N} \tag{47}$$

we obtain the decomposition

$$\begin{aligned}
& g(W_i^{(N)}(t+1)) - \mathbf{E}\left[g(W(t+1))\right] \\
= \;& g_\star(W_i^{(N)}(t)) - \mathbf{E}\left[g_\star(W(t))\right] \\
& + \hat{g}(W_i^{(N)}(t)) \left(M_i^{(N)}(t+1) - Z_i^{(N)}(t)\right) \\
& + \hat{g}(W_i^{(N)}(t)) \left(Z_i^{(N)}(t) - \gamma(t)^{W_i^{(N)}(t)}\right) \\
& + \hat{g}(W_i^{(N)}(t)) \gamma(t)^{W_i^{(N)}(t)} - \mathbf{E}\left[\hat{g}(W(t)) Z(t)\right]
\end{aligned}$$

as we note that

$$\mathbf{E}\left[\left(g^\star(W(t)) - g_\star(W(t))\right) M(t+1)\right] = \mathbf{E}\left[\hat{g}(W(t)) Z(t)\right]$$

since, in analogy with (28), we also have

$$\mathbf{E}\left[M(t+1) \middle| \mathcal{F}_t\right] = Z(t) \quad a.s.$$

by virtue of (12), (23) and (24).

Finally, defining the mapping $\hat{g}_t : \mathbb{N} \to \mathbb{R}^p$ by

$$\hat{g}_t(w) := g_\star(w) + \hat{g}(w) \cdot \gamma(t)^w, \quad w \in \mathbb{N} \tag{48}$$

we find

$$\begin{aligned} & g(W_i^{(N)}(t+1)) - \mathbf{E}\left[g(W(t+1))\right] \\ = \; & \hat{g}_t(W_i^{(N)}(t)) - \mathbf{E}\left[\hat{g}_t(W(t))\right] \\ & + \hat{g}(W_i^{(N)}(t)) \left(M_i^{(N)}(t+1) - Z_i^{(N)}(t) \right) \\ & + \hat{g}(W_i^{(N)}(t)) \left(Z_i^{(N)}(t) - \gamma(t)^{W_i^{(N)}(t)} \right). \end{aligned} \tag{49}$$

This decomposition forms the basis for the forthcoming analysis. The subsequent sections discuss the needed asymptotics for each of the three terms of (49).

11 A conditional CLT

The second term of (49) gives rise to a conditional CLT which we now develop: For any mapping $h : \mathbb{N} \to \mathbb{R}^q$ with positive integer q, we set

$$Y_h^{(N)}(t+1) := \frac{1}{N} \sum_{i=1}^{N} h(W_i^{(N)}(t)) \left(M_i^{(N)}(t+1) - Z_i^{(N)}(t) \right), \quad N = 1, 2, \ldots$$

for each $t = 0, 1, \ldots$. We begin with the case $q = 1$.

Proposition 11.1. *Assume (A1b)-(A2) to hold and consider an arbitrary mapping $h : \mathbb{N} \to \mathbb{R}$. Then, for each $t = 0, 1, \ldots$, it holds that*

$$\mathbf{E}\left[\exp\left(j\theta\sqrt{N} Y_h^{(N)}(t+1) \right) \Big| \mathcal{F}_t \right] \xrightarrow{P}_N e^{-\frac{\theta^2}{2}\sigma_h(t+1)}, \quad \theta \in \mathbb{R} \tag{50}$$

with

$$\sigma_h(t+1) := \mathbf{E}\left[h(W(t))^2 Z(t)(1 - Z(t)) \right]. \tag{51}$$

The proof of Proposition 11.1 is given in Appendix C. By using the standard Cramér-Wold device [3, Thm. 7.7, p. 49] we obtain the following analog in higher dimensions:

Corollary 11.2. *Assume (A1b)-(A2) to hold and consider an arbitrary mapping $h : \mathbb{N} \to \mathbb{R}^q$ for some positive integer q. Then, for each $t = 0, 1, \ldots$, it holds that*

$$\mathbf{E}\left[\exp\left(j\sqrt{N}Y_h^{(N)}(t+1)\theta'\right)\Big|\mathcal{F}_t\right] \xrightarrow{P}_N e^{-\frac{1}{2}\theta\Sigma_h(t+1)\theta'}, \quad \theta \in \mathbb{R}^q \tag{52}$$

with $q \times q$ covariance matrix $\Sigma_h(t+1)$ given by

$$\Sigma_h(t+1) := \mathbf{E}\left[h(W(t))'h(W(t))Z(t)(1-Z(t))\right]. \tag{53}$$

We conclude with the following crucial by-products: For some $t = 0, 1, \ldots,$ and an arbitrary positive integer r, consider the situation where a sequence of $\mathbb{R}^r$-valued rvs $\{\Lambda^{(N)}(t),\ N = 1, 2, \ldots\}$ weakly converges, say

$$\Lambda^{(N)}(t) \Longrightarrow_N \Lambda(t) \tag{54}$$

for some limiting $\mathbb{R}^r$-valued rv $\Lambda(t)$. If for each $N = 1, 2, \ldots$, the rv $\Lambda^{(N)}(t)$ is $\mathcal{F}_t$-measurable, then Corollary 11.2 readily implies [3, Thm. 3.2, p. 21] the *joint* convergence

$$(\sqrt{N}Y_h^{(N)}(t+1), \Lambda^{(N)}(t)) \Longrightarrow_N (Y_h(t+1), \Lambda(t)) \tag{55}$$

for some zero-mean Gaussian rv $Y_h(t+1)$ with covariance matrix $\Sigma_h(t+1)$, where the rv $Y_h(t+1)$ is taken to be *independent* of the rv $\Lambda(t)$. In particular, this applies to the mapping $h = \hat{g}$ appearing in the second term of (49).

12 The Delta Method

The contributions of the last term in the decomposition (49) are handled by the *Delta Method* [20, Thm. 3.1, p. 26]. While this result is often associated with the Central Limit Theorem, we now state the version to be used here.

Proposition 12.1. *Let $f : \mathbb{R}_+ \to [0, 1]$ be a differentiable mapping with derivative $f' : \mathbb{R}_+ \to \mathbb{R}$ continuous at $x = q(t)$. If for some $t = 0, 1, \ldots$, the convergence*

$$\sqrt{N}\left(\frac{Q^{(N)}(t)}{N} - q(t)\right) \Longrightarrow_N L_0(t) \tag{56}$$

takes place with some rv $L_0(t)$, then

$$\sqrt{N}\left(f\left(\frac{Q^{(N)}(t)}{N}\right) - f(q(t))\right) \Longrightarrow_N f'(q(t))L_0(t). \tag{57}$$

We do not require *a priori* that the limit in (56) be normally distributed. Although the result is well known, a proof is nevertheless provided below in order to validate several *joint* convergence statements which turn out to be crucial for proving Proposition 8.2.

Proof. Fix $t = 0, 1, \ldots$ and $N = 1, 2, \ldots$. We start with the observation that

$$\begin{aligned}
&\sqrt{N}\left(f(\frac{Q^{(N)}(t)}{N}) - f(q(t))\right) \\
= \quad & \sqrt{N}\int_{q(t)}^{\frac{Q^{(N)}(t)}{N}} \left(f'(x) - f'(q(t))\right) dx + \sqrt{N}\left(\frac{Q^{(N)}(t)}{N} - q(t)\right) f'(q(t)) \\
= \quad & \sqrt{N}U^{(N)}(t) + \sqrt{N}L_0^{(N)}(t)f'(q(t))
\end{aligned} \tag{58}$$

where we have set

$$U^{(N)}(t) := \int_{q(t)}^{\frac{Q^{(N)}(t)}{N}} \left(f'(x) - f'(q(t))\right) dx.$$

The desired conclusion (57) will readily follow if we show the convergence

$$\sqrt{N}U^{(N)}(t) \xrightarrow{P}_N 0. \tag{59}$$

To that end, fix $\varepsilon > 0$ arbitrary. For any $\delta > 0$, we have

$$\begin{aligned}
\mathbf{P}\left[\sqrt{N}\left|U^{(N)}(t)\right| > \varepsilon\right] \quad \leq \quad & \mathbf{P}\left[\sqrt{N}\left|U^{(N)}(t)\right| > \varepsilon,\ \left|\frac{Q^{(N)}(t)}{N} - q(t)\right| \leq \delta\right] \\
& + \mathbf{P}\left[\left|\frac{Q^{(N)}(t)}{N} - q(t)\right| > \delta\right].
\end{aligned} \tag{60}$$

By the continuity of f' at $x = q(t)$, for each $\eta > 0$, there exists $\delta(\eta) > 0$ such that $|f'(x) - f'(q(t))| \leq \eta$ whenever $|x - q(t)| < \delta(\eta)$ in $\mathbb{R}_+$. Now fix $\eta > 0$

and pick $\delta > 0$ in the range $(0, \delta(\eta))$. Thus, on the event $[\left|\frac{Q^{(N)}(t)}{N} - q(t)\right| \leq \delta]$, we find that

$$\sqrt{N}\left|U^{(N)}(t)\right| \leq \sqrt{N}\eta\left|\frac{Q^{(N)}(t)}{N} - q(t)\right|.$$

Reporting this fact into the inequality (60) we obtain

$$\begin{aligned} & \mathbf{P}\left[\sqrt{N}\left|U^{(N)}(t)\right| > \varepsilon\right] \\ \leq\ & \mathbf{P}\left[\sqrt{N}\left|\frac{Q^{(N)}(t)}{N} - q(t)\right| > \frac{\varepsilon}{\eta}\right] + \mathbf{P}\left[\left|\frac{Q^{(N)}(t)}{N} - q(t)\right| > \delta\right] \end{aligned} \tag{61}$$

Letting N go to infinity in (61) and using the convergence (7) and (56), we get

$$\limsup_{N\to\infty} \mathbf{P}\left[\sqrt{N}\left|U^{(N)}(t)\right| > \varepsilon\right] \leq \mathbf{P}\left[L_0(t) > \frac{\varepsilon}{\eta}\right]. \tag{62}$$

The desired conclusion (59) is now immediate upon letting $\eta > 0$ go to zero in this last inequality since its left-hand side is independent of η. ■

13 Using the Delta Method

For any mapping $h : \mathbb{N} \to \mathbb{R}^q$ with positive integer q, set

$$X_h^{(N)}(t) := \frac{1}{N}\sum_{i=1}^{N} h(W_i^{(N)}(t))\left(\gamma^{(N)}(t)^{W_i^{(N)}(t)} - \gamma(t)^{W_i^{(N)}(t)}\right), \quad N = 1, 2, \ldots$$

for each $t = 0, 1, \ldots$. The relevant limiting result is presented next.

Proposition 13.1. *Assume (A1b)-(A2) to hold and consider an arbitrary mapping $h : \mathbb{N} \to \mathbb{R}^q$ for some positive integer q. If for some $t = 0, 1, \ldots$, the convergence (56) holds with some rv $L_0(t)$, then it holds that*

$$\sqrt{N}X_h^{(N)}(t) \Longrightarrow_N -f'(q(t))R_h(t)L_0(t) \tag{63}$$

with

$$R_h(t) := \mathbf{E}\left[W(t)\left(1 - f(q(t))\right)^{W(t)-1} h(W(t))\right]. \tag{64}$$

Proof. Fix $N = 1, 2, \ldots$ and $i = 1, \ldots, N$. From (24) and (30), we observe that

$$\begin{aligned}
& \gamma^{(N)}(t)^{W_i^{(N)}(t)} - \gamma(t)^{W_i^{(N)}(t)} \\
= & -W_i^{(N)}(t) \int_{f(q(t))}^{f(\frac{Q^{(N)}(t)}{N})} (1-y)^{W_i^{(N)}(t)-1} \, dy \\
= & -W_i^{(N)}(t) \Delta_i^{(N)}(t) \\
& - W_i^{(N)}(t) \left(f\left(\frac{Q^{(N)}(t)}{N} \right) - f(q(t)) \right) (1 - f(q(t)))^{W_i^{(N)}(t)-1} \quad (65)
\end{aligned}$$

where we have set

$$\Delta_i^{(N)}(t) := \int_{f(q(t))}^{f(\frac{Q^{(N)}(t)}{N})} \left[(1-y)^{W_i^{(N)}(t)-1} - (1 - f(q(t)))^{W_i^{(N)}(t)-1} \right] \, dy.$$

Consequently,

$$\begin{aligned}
\sqrt{N} X_h^{(N)}(t) &= -\sqrt{N} \left(\frac{1}{N} \sum_{i=1}^{N} h(W_i^{(N)}(t)) W_i^{(N)}(t) \Delta_i^{(N)}(t) \right) \quad (66) \\
& \quad - \left(\frac{1}{N} \sum_{i=1}^{N} h_t^\star(W_i^{(N)}(t)) \right) \cdot \sqrt{N} \left(f\left(\frac{Q^{(N)}(t)}{N} \right) - f(q(t)) \right)
\end{aligned}$$

where the mapping $h_t^\star : \mathbb{N} \to \mathbb{R}^p$ is defined by

$$h_t^\star(w) := w h(w) \left(1 - f(q(t))\right)^{w-1}, \quad w \in \mathbb{N}.$$

Using the inequality (B.7), we find that

$$\begin{aligned}
|\Delta_i^{(N)}(t)| &\leq \left(W_i^{(N)}(t) - 1 \right) \left| \int_{f(q(t))}^{f(\frac{Q^{(N)}(t)}{N})} |y - f(q(t))| \, dy \right| \\
&\leq W_{\max} \left| \int_{f(q(t))}^{f(\frac{Q^{(N)}(t)}{N})} |y - f(q(t))| \, dy \right| \\
&= \frac{W_{\max}}{2} \left| f\left(\frac{Q^{(N)}(t)}{N} \right) - f(q(t)) \right|^2. \quad (67)
\end{aligned}$$

Thus,

$$\begin{aligned}
&\sqrt{N}\left\|\frac{1}{N}\sum_{i=1}^{N} h(W_i^{(N)}(t))W_i^{(N)}(t)\Delta_i^{(N)}(t)\right\| \\
\leq\ & \frac{W_{\max}}{2}\sqrt{N}\left|f\left(\frac{Q^{(N)}(t)}{N}\right)-f(q(t))\right|^2\left(\frac{1}{N}\sum_{i=1}^{N}\|h(W_i^{(N)}(t))\|W_i^{(N)}(t)\right).
\end{aligned}$$

Let N go to infinity. Combining the convergence (13) with Proposition 12.1 already yields

$$\sqrt{N}\left|f\left(\frac{Q^{(N)}(t)}{N}\right)-f(q(t))\right|^2 \xrightarrow{P}_N 0.$$

It now follows by Claim (ii) of Theorem 3.1 (applied to the mapping $w \to w\|h(w)\|$) that

$$\sqrt{N}\left\|\frac{1}{N}\sum_{i=1}^{N} h(W_i^{(N)}(t))W_i^{(N)}(t)\Delta_i^{(N)}(t)\right\| \xrightarrow{P}_N 0. \tag{68}$$

Therefore, in view of the decomposition (66) and of the convergence (68), the convergence (63) will hold provided we can show

$$\sqrt{N}\left(f\left(\frac{Q^{(N)}(t)}{N}\right)-f(q(t))\right)\left(\frac{1}{N}\sum_{i=1}^{N} h_t^\star(W_i^{(N)}(t))\right) \Longrightarrow_N f'(q(t))R_h(t)L_0(t).$$

This convergence follows from Proposition 12.1 once we note that

$$\frac{1}{N}\sum_{i=1}^{N} h_t^\star(W_i^{(N)}(t)) \xrightarrow{P}_N R_h(t)$$

by Claim (ii) of Theorem 3.1 (applied to the mapping $h_t^\star$). ■

A careful inspection of the proofs of Propositions 12.1 and 13.1 reveals that a somewhat stronger statement is true: For some $t = 0, 1, \ldots$, consider the situation where a sequence of $\mathbb{R}^r$-valued rvs $\{\tilde{\Lambda}^{(N)}(t),\ N = 1, 2, \ldots\}$ is weakly convergent as in (54) *together* with (56), i.e., we have the *joint* convergence

$$\left(\sqrt{N}L_0^{(N)}(t), \tilde{\Lambda}^{(N)}(t)\right) \Longrightarrow_N \left(L_0(t), \tilde{\Lambda}(t)\right). \tag{69}$$

for some limiting $\mathbb{R}^r$-valued rv $\tilde{\Lambda}(t)$. Then, the convergence (63) can be extended to read

$$\left(\sqrt{N}X_h^{(N)}(t), \tilde{\Lambda}^{(N)}(t)\right) \Longrightarrow_N \left(-f'(q(t))R_h(t)L_0(t), \tilde{\Lambda}(t)\right). \tag{70}$$

14 A proof of Proposition 8.2

Pick an arbitrary mapping $g : \mathbb{N} \to \mathbb{R}^p$ with positive integer p. Fix $t = 0, 1, \ldots$ and $N = 1, 2, \ldots$, and go back to the basic decomposition (49): With the notation introduced earlier, we note that

$$L_g^{(N)}(t+1) = L_{\hat{g}_t}^{(N)}(t) + X_{\hat{g}}^{(N)}(t) + Y_{\hat{g}}^{(N)}(t+1) \tag{71}$$

where $\hat{g}$ and $\hat{g}_t$ are the mappings $\mathbb{N} \to \mathbb{R}^p$ defined earlier at (47) and (48). Also recall (43) from the proof of Proposition 8.1 to the effect that

$$\begin{aligned} & \sqrt{N}L_0^{(N)}(t+1) \qquad (72) \\ = \quad & \max\left(\sqrt{N}\left(L_0^{(N)}(t) + L_{\mathrm{avg}}^{(N)}(t)\right), \sqrt{N}K(t)\right) - \sqrt{N}K(t)^+. \end{aligned}$$

By Corollary 11.2, we already have

$$\sqrt{N}Y_{\hat{g}}^{(N)}(t+1) \Longrightarrow_N Y_{\hat{g}}(t+1) \tag{73}$$

for some zero-mean Gaussian rv $Y_{\hat{g}}(t+1)$ with covariance matrix $\Sigma_{\hat{g}}(t+1)$ given by (53) (with $h = \hat{g}$). However, the rvs $L_0^{(N)}(t)$, $L_{\hat{g}_t}^{(N)}(t)$ and $X_{\hat{g}}^{(N)}(t)$ are each $\mathcal{F}_t$-measurable for each $N = 1, 2, \ldots$. Therefore, having in mind the comments following Corollary 11.2, we introduce the $\mathbb{R}^{2p+2}$-valued rv $\Lambda^{(N)}(t)$ defined by

$$\Lambda^{(N)}(t) := \sqrt{N}\left(L_0^{(N)}(t), L_{g_t}^{(N)}(t), X_{\hat{g}}^{(N)}(t)\right)$$

where the mapping $g_t : \mathbb{N} \to \mathbb{R}^{p+1}$ is given by

$$g_t(w) := (w, \hat{g}_t(w)), \quad w \in \mathbb{N}.$$

We will have established the convergence part of Proposition 8.2 if we can show the *joint* convergence

$$\Lambda^{(N)}(t) \Longrightarrow_N \Lambda(t) := (L_0(t), L_{g_t}(t), X_{\hat{g}}(t)) \tag{74}$$

for some $\mathbb{R}^{2p+2}$-valued rv $(L_0(t), L_{g_t}(t), X_{\hat{g}}(t))$ to be determined. Indeed, as indicated at the end of Section 11, this would readily imply

$$\left(\Lambda^{(N)}(t), \sqrt{N} Y_{\hat{g}}^{(N)}(t+1)\right) \Longrightarrow_N (\Lambda(t), Y_{\hat{g}}(t+1)) \tag{75}$$

with the Gaussian rv $Y_{\hat{g}}(t+1)$ being taken independently of $\Lambda(t)$, and the conclusion

$$\sqrt{N}\left(L_0^{(N)}(t+1), L_g^{(N)}(t+1)\right) \Longrightarrow_N (L_0(t+1), L_g(t+1)) \tag{76}$$

is now straightforward by applying the Continuous Mapping Theorem on the relations (72) and (71).[4]

We need only identify the limiting rv in (74) and (76): Under **[E:t]**, it is already the case that

$$\sqrt{N}(L_0^{(N)}(t), L_{g_t}^{(N)}(t)) \Longrightarrow_N (L_0(t), L_{g_t}(t)). \tag{77}$$

Next, with

$$\tilde{\Lambda}^{(N)}(t) = \sqrt{N}(L_0^{(N)}(t), L_{g_t}^{(N)}(t)),$$

the strengthening (69)-(70) of Proposition 13.1 leads to (70) in the form

$$\left(\tilde{\Lambda}^{(N)}(t), \sqrt{N} X_h^{(N)}(t)\right) \Longrightarrow_N (L_0(t), L_{g_t}(t), -f'(q(t)) R_g(t) L_0(t)). \tag{78}$$

In other words, joint convergence (74) takes place with the limiting rv identified. ■

References

[1] E. Altman, K. Avrachenkov, and C. Barakat. TCP in presence of bursty losses. In *Proceedings of SIGMETRICS*, 2000.

[2] F. Baccelli, D.R. McDonald, and J. Reynier. A mean-field model for multiple TCP connections through a buffer implementing RED. Technical report, INRIA, April 2002.

[4] See the proof of Proposition 8.1.

[3] P. Billingsley. *Convergence of Probability Measures*. John Wiley & Sons, New York (NY), 1968.

[4] V. Firoiu and M. Borden. A study of active queue management for congestion control. In *Proceedings of IEEE INFOCOM*, 2000.

[5] S. Floyd. TCP and explicit congestion notification. *Computer Communication Review*, 24(5):10–23, October 1994.

[6] S. Floyd and V. Jacobson. Random Early Detection gateways for congestion avoidance. *IEEE/ACM Transactions on Networking*, 1(4):397–413, August 1995.

[7] D. Hong and D. Lebedev. Many TCP user asymptotic analysis of the AIMD model. Technical Report RR-4229, INRIA, July 2001.

[8] V. Jacobson. Congestion avoidance and control. In *Proceedings of SIGCOMM'88 Symposium*, pages 314–332, August 1988.

[9] A.F. Karr. *Probability*. Springer-Verlag, 1993.

[10] F. Kelly, A.K. Maulloo, and D.K.H. Tan. Rate control in communication networks: shadow prices, proportional fairness and stability. *Journal of the Operational Research Society*, 49:237–252, 1998.

[11] S.H. Low. A duality model of TCP and queue management algorithms. In *Proceedings of ITC Specialist Seminar on IP Traffic Measurement, Modeling and Management*, September 2000.

[12] M. Mathis, J. Semske, J. Mahdavi, and T. Ott. The macroscopic behavior of TCP congestion avoidance algorithm. *Computer Communication Review*, 27(3), July 1997.

[13] V. Misra, W. Gong, and D. Towsley. Stochastic differential equation modeling and analysis of TCP windowsize behavior. In *Proceedings of Performance*, 1999.

[14] UCB/VINT/LBNL Network Simulator (ns) version 2.0. http://www.isi.edu/nsnam.

[15] J. Padhye, V. Firoiu, D. Towsley, and J. Kurose. Modeling TCP Reno performance: A simple model and its empirical validation. *IEEE/ACM Transactions on Networking*, April 2000.

[16] S. Shakkottai and R. Srikant. How good are deterministic fluid models of Internet congestion control? In *Proceedings of IEEE INFOCOM*, New York, NY, June 2002.

[17] P. Tinnakornsrisuphap and R.J. La. Characterization of queue fluctuations in probabilistic AQM mechanisms. *ACM SIGMETRICS Performance Evaluation Review*, 32(1), June 2004.

[18] P. Tinnakornsrisuphap and A.M. Makowski. TCP traffic modeling via limit theorems. Technical Report, University of Maryland, 2002.

[19] P. Tinnakornsrisuphap and A.M. Makowski. Limit behavior of ECN/RED gateways under a large number of TCP flows. In *Proceedings of IEEE INFOCOM*, San Francisco, CA, April 2003.

[20] A.W. van der Vaart. *Asymptotic Statistics*. Cambridge University Press, 1998.

[21] L. Zhang and D. Clark. Oscillating behavior of network traffic: A case study simulation. *Internetworking: Research and Experience*, 1(2):101–112, 1990.

Appendices

A A proof of Lemmas 7.1–7.4

A.1 A proof of Lemma 7.1

Together, the convergence statements [**A:t**] and [**B:t**] imply the joint convergence $(N^{-1}Q^{(N)}(t), W_1^{(N)}(t)) \Longrightarrow_N (q(t), W(t))$ [9, Thm. 5.28, p. 150]. Next the continuity of the mapping f implies that of the mapping $(x, w) \to (1 - f(x))^w$ on $\mathbb{R}_+ \times (0, \infty)$, so that

$$(Z_1^{(N)}(t), W_1^{(N)}(t)) \Longrightarrow_N (Z(t), W(t)) \tag{A.1}$$

by the Continuous Mapping Theorem [9, Thm. 5.29, p. 150] with $Z(t)$ defined at (23).

Consider (33) for an arbitrary mapping $g : \mathbb{N} \to \mathbb{R}$, and observe that the mapping F_g defined by (32) is bounded and continuous on $[0,1] \times \mathbb{N}$.[5] Consequently, the Continuous Mapping Theorem can again be invoked to yield

$$F_g(Z_1^{(N)}(t), W_1^{(N)}(t)) \Longrightarrow_N F_g(Z(t), W(t)), \tag{A.2}$$

whence

$$\lim_{N\to\infty} \mathbf{E}\left[F_g(Z_1^{(N)}(t), W_1^{(N)}(t))\right] = \mathbf{E}\left[F_g(Z(t), W(t))\right] \tag{A.3}$$

by the Bounded Convergence Theorem [9, Thm. 4.16, p. 108]. Combining (33) and (A.3) we get

$$\lim_{N\to\infty} \mathbf{E}\left[g(W_1^{(N)}(t+1))\right] = \mathbf{E}\left[F_g(Z(t), W(t))\right].$$

On the other hand, the rvs $W(t)$ and $V(t+1)$ are independent, and inspection of the dynamics (11) and (12) reveals that

$$\mathbf{E}\left[g(W(t+1))\right] = \mathbf{E}\left[F_g(Z(t), W(t))\right]. \tag{A.4}$$

The mapping g being arbitrary, it follows immediately that $W_1^{(N)}(t+1) \Longrightarrow_N W(t+1)$ for some $\{1, \ldots, W_{\max}\}$–valued rv $W(t+1)$ defined through (11) and (12). ■

A.2 A proof of Lemma 7.2

Fix $N = 1, 2, \ldots$. From **[A:t]** and **[D:t]** (with $g(x) = x$), we conclude that

$$\frac{Q^{(N)}(t)}{N} - C + \frac{1}{N}\sum_{i=1}^{N} W_i^{(N)}(t) \xrightarrow{P}_N q(t) - C + \mathbf{E}\left[W(t)\right]$$

and the continuity of the function $x \to x^+$ implies the convergence

$$\frac{Q^{(N)}(t+1)}{N} \xrightarrow{P}_N q(t+1)$$

[5]This continuity is with respect to the product topology on $[0,1] \times \mathbb{N}$ where $\mathbb{N}$ is topologized according to the usual discrete topology.

with non-random $q(t+1)$ since

$$\frac{Q^{(N)}(t+1)}{N} = \left(\frac{Q^{(N)}(t)}{N} - C + \frac{1}{N} \sum_{i=1}^{N} W_i^{(N)}(t) \right)^+$$

The dynamics (10) giving $q(t+1)$ is immediate. ■

A.3 A proof of Lemma 7.3

Fix a positive integer I. The rvs $V_1(t+1), \ldots, V_I(t+1)$ are i.i.d. $[0,1]$-uniform rvs which are independent of $\mathcal{F}_t$. Thus, upon making use of (5), we see that the rvs $W_1^{(N)}(t+1), \ldots, W_I^{(N)}(t+1)$ are mutually independent given $\mathcal{F}_t$. Consequently, for arbitrary mappings $g_1, \ldots, g_I : \mathbb{N} \to \mathbb{R}$, we get

$$\begin{aligned}
\mathbf{E}\left[\prod_{i=1}^{I} g_i(W_i^{(N)}(t+1)) \Big| \mathcal{F}_t \right] &= \prod_{i=1}^{I} \mathbf{E}\left[g_i(W_i^{(N)}(t+1)) \Big| \mathcal{F}_t \right] \\
&= \prod_{i=1}^{I} F_{g_i}(Z_i^{(N)}(t), W_i^{(N)}(t))
\end{aligned}$$

with the help of (31) and (32).

Now it follows from (9) in **[C:t]** that the joint convergence

$$(W_1^{(N)}(t), \ldots, W_I^{(N)}(t)) \Longrightarrow_N (W_1(t), \ldots, W_I(t))$$

holds with limiting rvs $W_1(t), \ldots, W_I(t)$ which are i.i.d. rvs, each distributed according to $W(t)$. As in the proof of Lemma 7.1, the arguments leading to the convergence (A.2) also lead to

$$\begin{aligned}
&(F_{g_1}(Z_1^{(N)}(t), W_1^{(N)}(t)), \ldots, F_{g_I}(Z_I^{(N)}(t), W_I^{(N)}(t))) \\
&\Longrightarrow_N (F_{g_1}(Z_1(t), W_1(t)), \ldots, F_{g_I}(Z_I(t), W_I(t)))
\end{aligned}$$

where the limiting rvs $(Z_1(t), W_1(t)), \ldots, (Z_I(t), W_I(t))$ are i.i.d. rvs each distributed according to the pair $(Z(t), W(t))$. Therefore, by the Bounded

Convergence Theorem,

$$\begin{aligned}
\lim_{N\to\infty} \mathbf{E}\left[\prod_{i=1}^{I} g_i(W_i^{(N)}(t+1))\right] &= \lim_{N\to\infty} \mathbf{E}\left[\prod_{i=1}^{I} F_{g_i}(Z_i^{(N)}(t), W_i^{(N)}(t))\right] \\
&= \mathbf{E}\left[\prod_{i=1}^{I} F_{g_i}(Z_i(t), W_i(t))\right] \\
&= \prod_{i=1}^{I} \mathbf{E}\left[F_{g_i}(Z_i(t), W_i(t))\right] \\
&= \prod_{i=1}^{I} \mathbf{E}\left[g_i(W_i(t+1))\right] \qquad \text{(A.5)}
\end{aligned}$$

where the last equality made use of the relation (A.4). The desired result **[C:t+1]** now follows from (A.5) given that the mappings $g_1, \ldots, g_I$ are arbitrary. ■

A.4 A proof of Lemma 7.4

Pick a mapping $g : \mathbb{N} \to \mathbb{R}$. We begin by observing that under (A2) the rvs $W_i^{(N)}(t), \ldots, W_N^{(N)}(t)$ are exchangeable. As a result, we get

$$\begin{aligned}
&\operatorname{var}\left[\frac{1}{N}\sum_{i=1}^{N} g(W_i^{(N)}(t))\right] \\
&= N^{-2}\sum_{i=1}^{N} \operatorname{var}[g(W_i^{(N)}(t))] + N^{-2}\sum_{i,j=1,i\neq j}^{N} \operatorname{cov}[g(W_i^{(N)}(t)), g(W_j^{(N)}(t))] \\
&= N^{-1}\operatorname{var}[g(W_1^{(N)}(t))] + \frac{N-1}{N}\operatorname{cov}[g(W_1^{(N)}(t)), g(W_2^{(N)}(t))].
\end{aligned}$$

Let N go to infinity in this last relation. The validity of **[C:t]** and the Bounded Convergence Theorem already imply

$$\lim_{N\to\infty} \operatorname{cov}[g(W_1^{(N)}(t)), g(W_2^{(N)}(t))] = \operatorname{cov}[g(W_1(t)), g(W_2(t))] = 0$$

by asymptotic independence. On the other hand, $|g(W_1^{(N)}(t))| \leq G$ where $G := \max\{|g(x)|,\ x = 1, \ldots, W_{\max}\}$ so that $\limsup_{N\to\infty} \operatorname{var}[g(W_1^{(N)}(t))] \leq$

G^2. Combining these observations, we readily see that

$$\lim_{N\to\infty} \text{var}\left[\frac{1}{N}\sum_{i=1}^{N} g(W_i^{(N)}(t))\right] = 0,$$

whence

$$\frac{1}{N}\sum_{i=1}^{N} g(W_i^{(N)}(t)) - \mathbf{E}\left[\frac{1}{N}\sum_{i=1}^{N} g(W_i^{(N)}(t))\right] \xrightarrow{P}_N 0$$

by Chebyshev's Inequality. This last convergence is equivalent to

$$\frac{1}{N}\sum_{i=1}^{N} g(W_i^{(N)}(t)) - \mathbf{E}\left[g(W_1^{(N)}(t))\right] \xrightarrow{P}_N 0$$

by exchangeability, and the desired convergence result (8) is now immediate once we remark under [**B:t**] that $\lim_{N\to\infty} \mathbf{E}\left[g(W_1^{(N)}(t))\right] = \mathbf{E}\left[g(W(t)\right]$. ■

B A proof of Proposition 7.5

Fix $t = 0, 1, \ldots$ and $\ell = 1, 2, \ldots$. Also fix $N = 1, 2, \ldots$ and $i = 1, \ldots, N$. We have

$$\begin{aligned} & g(W_i^{(N)}(t))Z_i^{(N)}(t)^\ell \\ = \; & g(W_i^{(N)}(t))\left(\left(1 - f\left(\frac{Q^{(N)}(t)}{N}\right)\right)^{\ell W_i^{(N)}(t)} - (1 - f(q(t)))^{\ell W_i^{(N)}(t)}\right) \\ & + g_{t,\ell}(W_i^{(N)}(t)) \end{aligned} \tag{B.6}$$

with mapping $g_{t,\ell} : \mathbb{R} \to \mathbb{R}$ given by

$$g_{t,\ell}(w) := g(w)\left(1 - f(q(t))\right)^{\ell w}, \quad w \in \mathbb{N}.$$

Next, for any pair a, b in $[0, 1]$, we have

$$|a^p - b^p| = p\left|\int_a^b t^{p-1}dt\right| \le p|b - a| \tag{B.7}$$

for each $p = 1, 2, \ldots$, so that

$$\begin{aligned} & \left| \left(1 - f\left(\frac{Q^{(N)}(t)}{N}\right)\right)^{\ell W_i^{(N)}(t)} - (1 - f(q(t)))^{\ell W_i^{(N)}(t)} \right| \\ \leq \; & \ell W_i^{(N)}(t) \left| f\left(\frac{Q^{(N)}(t)}{N}\right) - f(q(t)) \right|. \end{aligned}$$

Therefore,

$$\begin{aligned} & \left| \frac{1}{N} \sum_{i=1}^{N} g(W_i^{(N)}(t)) \left(Z_i^{(N)}(t)^\ell - (1 - f(q(t)))^{\ell W_i^{(N)}(t)} \right) \right| \\ \leq \; & \ell \left| f\left(\frac{Q^{(N)}(t)}{N}\right) - f(q(t)) \right| \left(\frac{1}{N} \sum_{i=1}^{N} W_i^{(N)}(t) \left| g(W_i^{(N)}(t)) \right| \right) \end{aligned}$$

and from the convergence (13), we get

$$\frac{1}{N} \sum_{i=1}^{N} g(W_i^{(N)}(t)) \left(Z_i^{(N)}(t)^\ell - (1 - f(q(t)))^{\ell W_i^{(N)}(t)} \right) \xrightarrow{P}_N 0 \tag{B.8}$$

since

$$\frac{1}{N} \sum_{i=1}^{N} W_i^{(N)}(t) |g(W_i^{(N)}(t))| \xrightarrow{P}_N \mathbf{E}\left[W(t)|g(W(t))|\right]$$

by Claim (ii) of Theorem 3.1.

The conclusion is now immediate from the decomposition (B.6), the convergence (B.8) and the convergence

$$\frac{1}{N} \sum_{i=1}^{N} g_{t,\ell}(W_i^{(N)}(t)) \xrightarrow{P}_N \mathbf{E}\left[g_{t,\ell}(W(t))\right]$$

obtained from Claim (ii) of Theorem 3.1. ■

C A proof of Proposition 11.1

We rely on the following two technical lemmas; their proofs are omitted in the interest of brevity.

Lemma C.1. *For any x in $\mathbb{R}$, the Taylor series expansion*

$$e^{jx} = 1 + jx - \frac{x^2}{2} + R(x) \tag{C.9}$$

holds, with complex-valued remainder term $R(x)$ satisfying

$$|R(x)| \le \frac{|x|^3}{6}. \tag{C.10}$$

Lemma C.2. *Consider the array of complex-valued rvs $\{C_{N,i},\ i = 1, \ldots, N;\ N = 1, 2, \ldots\}$ with $|C_{N,i}| < 1$ for $i = 1, \ldots, N$. If $\max_{i=1,\ldots,N} |C_{N,i}| \to_N 0$ a.s. and $\sum_{i=1}^N C_{N,i} \xrightarrow{P}_N \lambda$, then*

$$\prod_{i=1}^N (1 - C_{N,i}) \xrightarrow{P}_N e^{-\lambda}. \tag{C.11}$$

The proof of Proposition 11.1 can now proceed: Fix $N = 1, 2, \ldots$ and θ arbitrary in $\mathbb{R}$. By conditional independence and making use of (28), we find that

$$\begin{aligned} & \mathbf{E}\left[\exp\left(j\theta\sqrt{N}Y_h^{(N)}(t+1)\right) \Big| \mathcal{F}_t\right] \\ = & \prod_{i=1}^N \mathbf{E}\left[\exp\left(j\frac{\theta}{\sqrt{N}}\left(h(W_i^{(N)}(t))\left(M_i^{(N)}(t+1) - Z_i^{(N)}(t)\right)\right)\right) \Big| \mathcal{F}_t\right] \\ = & \prod_{i=1}^N \left(1 - C_i^{(N)}(t)\right) \end{aligned} \tag{C.12}$$

with

$$\begin{aligned} C_i^{(N)}(t) \ = \ & Z_i^{(N)}(t)\left[1 - \exp\left(j\frac{\theta}{\sqrt{N}}h(W_i^{(N)}(t))(1 - Z_i^{(N)}(t))\right)\right] \\ & + (1 - Z_i^{(N)}(t))\left[1 - \exp\left(-j\frac{\theta}{\sqrt{N}}h(W_i^{(N)}(t))Z_i^{(N)}(t)\right)\right] \end{aligned} \tag{C.13}$$

for each $i = 1, \ldots, N$.

In view of these remarks, the desired result (50) is now a simple consequence of Lemma C.2 provided the required conditions can be shown to hold, namely

$$\lim_{N\to\infty} \max_{i=1,\ldots,N} |C_i^{(N)}(t)| = 0 \quad a.s. \tag{C.14}$$

and

$$\sum_{i=1}^{N} C_i^{(N)}(t) \xrightarrow{P}_N \frac{\theta^2}{2}\sigma_h(t+1) \tag{C.15}$$

with $\sigma_h(t+1)$ given by (51).

To check these conditions, we first invoke Lemma C.1 to write

$$C_i^{(N)}(t) = Z_i^{(N)}(t)B_{i,1}^{(N)}(t) + (1 - Z_i^{(N)}(t))B_{i,2}^{(N)}(t) \tag{C.16}$$

with

$$\begin{aligned} B_{i,1}^{(N)}(t) &= -j\frac{\theta}{\sqrt{N}}h(W_i^{(N)}(t))(1 - Z_i^{(N)}(t)) \\ &\quad + \frac{\theta^2}{2N}h(W_i^{(N)}(t))^2(1 - Z_i^{(N)}(t))^2 + \beta_{i,1}^{(N)}(t;\theta) \end{aligned} \tag{C.17}$$

and

$$\begin{aligned} B_{i,2}^{(N)}(t) &= j\frac{\theta}{\sqrt{N}}h(W_i^{(N)}(t))Z_i^{(N)}(t) \\ &\quad + \frac{\theta^2}{2N}h(W_i^{(N)}(t))^2 Z_i^{(N)}(t)^2 + \beta_{i,2}^{(N)}(t;\theta) \end{aligned} \tag{C.18}$$

where the remainder terms $\beta_{i,1}^{(N)}(t;\theta)$ and $\beta_{i,2}^{(N)}(t;\theta)$ in these expressions satisfy

$$\max\left(|\beta_{i,1}^{(N)}(t;\theta)|, |\beta_{i,2}^{(N)}(t;\theta)|\right) \leq K\frac{|\theta|^3}{6\sqrt{N^3}}$$

for some positive constant K, say $K := \max\{|h(w)|^3, w = 1, \ldots, W_{\max}\}$.

Thus, condition (C.14) trivially holds. To establish (C.15), report (C.17) and (C.18) into (C.16). Simplifying the resulting expression, we find

$$C_i^{(N)}(t) = \frac{\theta^2}{2N}h(W_i^{(N)}(t))^2 Z_i^{(N)}(t)(1 - Z_i^{(N)}(t)) + \gamma_i^{(N)}(t;\theta) \tag{C.19}$$

with remainder term

$$\gamma_i^{(N)}(t;\theta) := Z_i^{(N)}(t)\beta_{i,1}^{(N)}(t;\theta) + (1 - Z_i^{(N)}(t))\beta_{i,2}^{(N)}(t;\theta).$$

Obviously,

$$|\gamma_i^{(N)}(t;\theta)| \leq K\frac{|\theta|^3}{6\sqrt{N^3}}. \tag{C.20}$$

Next, we write

$$\sum_{i=1}^{N} C_i^{(N)}(t) = \frac{\theta^2}{2}\left(\frac{1}{N}\sum_{i=1}^{N} h(W_i^{(N)}(t))^2 Z_i^{(N)}(t)(1-Z_i^{(N)}(t))\right) + \Gamma^{(N)}(t)$$

where we have set

$$\Gamma^{(N)}(t) := \frac{1}{N}\sum_{i=1}^{N} \gamma_i^{(N)}(t;\theta).$$

By repeated application of Proposition 7.5 (with $\ell = 1$ and $\ell = 2$) we get

$$\frac{1}{N}\sum_{i=1}^{N} h(W_i^{(N)}(t))^2 Z_i^{(N)}(t)(1-Z_i^{(N)}(t)) \xrightarrow{P}_N \sigma_h(t+1)$$

with $\sigma_h(t+1)$ given by (51), while $\lim_{N\to\infty}\Gamma^{(N)}(t) = 0$ a.s. by virtue of (C.20). The desired conclusion (C.15) is obtained and the proof of Proposition 11.1 is complete. ■