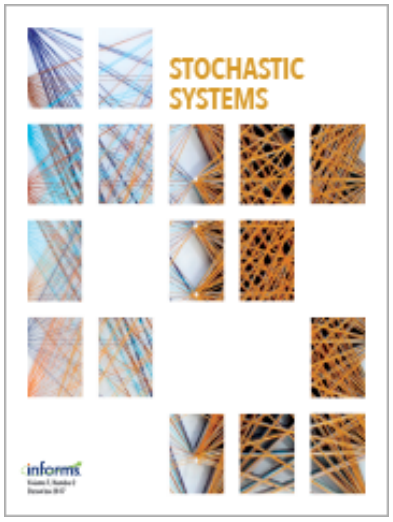
This article was downloaded by: [216.73.217.172] On: 15 November 2026, At: 05:56
Publisher: Institute for Operations Research and the Management Sciences (INFORMS)
INFORMS is located in Maryland, USA



Stochastic Systems

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Queue-based Random-access Algorithms: Fluid Limits and Stability Issues

Javad Ghaderi, Sem Borst, Phil Whiting

To cite this article:
Javad Ghaderi, Sem Borst, Phil Whiting (2014) Queue-based Random-access Algorithms: Fluid Limits and Stability Issues. Stochastic Systems 4(1):81-156. https://doi.org/10.1287/13-SSY104

This work is licensed under a Creative Commons Attribution 4.0 International License. You are free to copy, distribute, transmit and adapt this work, but you must attribute this work as "*Stochastic Systems*. Copyright 2014 The Author(s). https://doi.org/10.1287/13-SSY104, used under a Creative Commons Attribution License: http://creativecommons.org/licenses/by/4.0/."

Copyright © 2014, The author(s)

Please scroll down for article—it is on subsequent pages

With 12,500 members from nearly 90 countries, INFORMS is the largest international association of operations research (O.R.) and analytics professionals and students. INFORMS provides unique networking and learning opportunities for individual professionals, and organizations of all types and sizes, to better understand and use O.R. and analytics tools and methods to transform strategic visions and achieve better outcomes.
For more information on INFORMS, its publications, membership, or meetings visit http://www.informs.org

Stochastic Systems
2014, Vol. 4, No. 1, 81–156
DOI: 10.1214/13-SSY104

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

QUEUE-BASED RANDOM-ACCESS ALGORITHMS: FLUID LIMITS AND STABILITY ISSUES*

By Javad Ghaderi†, Sem Borst§, and Phil Whiting¶

Columbia University†, Alcatel-Lucent Bell Labs §, Eindhoven University of Technology§, and Macquarie University¶

We use fluid limits to explore the (in)stability properties of wireless networks with queue-based random-access algorithms. Queue-based random-access schemes are simple and inherently distributed in nature, yet provide the capability to match the optimal throughput performance of centralized scheduling mechanisms in a wide range of scenarios. Unfortunately, the type of activation rules for which throughput optimality has been established, may result in excessive queue lengths and delays. The use of more aggressive/persistent access schemes can improve the delay performance, but does not offer any universal maximum-stability guarantees.

In order to gain qualitative insight and investigate the (in)stability properties of more aggressive/persistent activation rules, we examine fluid limits where the dynamics are scaled in space and time. In some situations, the fluid limits have smooth deterministic features and maximum stability is maintained, while in other scenarios they exhibit random oscillatory characteristics, giving rise to major technical challenges. In the latter regime, more aggressive access schemes continue to provide maximum stability in some networks, but may cause instability in others. In order to prove that, we focus on a particular network example and conduct a detailed analysis of the fluid limit process for the associated Markov chain. Specifically, we develop a novel approach based on stopping time sequences to deal with the switching probabilities governing the sample paths of the fluid limit process. Simulation experiments are conducted to illustrate and validate the analytical results.

1. Introduction. Emerging wireless mesh networks typically lack any centralized access control entity, and instead vitally rely on the individual nodes to operate autonomously and to efficiently share the medium in a dis-

Received March 2013.

*The work of the first author was supported by a summer project at Alcatel-Lucent and grants NSF CNS-1161404, ARO MURI W911NF-08-1-0233, and ONR N00014-13-1-0038 during his PhD at University of Illinois.

MSC 2010 subject classifications: Primary 60K25, 60K25; secondary 68M20, 90B18, 90B22

Keywords and phrases: Carrier-Sense Multiple Access (CSMA), fluid limits, queue-based strategies, stability issues.

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

tributed fashion. This requires the nodes to schedule their individual transmissions and decide on the use of a shared medium based on knowledge that is locally available or only involves limited exchange of information. A popular mechanism for distributed medium access control is provided by the so-called Carrier-Sense Multiple-Access (CSMA) protocol. In the CSMA protocol each node attempts to access the medium after a certain back-off time, but nodes that sense activity of interfering nodes freeze their back-off timer until the medium is sensed idle. While the CSMA protocol is fairly easy to understand at a local level, the interaction among interfering nodes gives rise to quite intricate behavior and complex throughput characteristics on a macroscopic scale. In recent years relatively parsimonious models have emerged that provide a useful tool in evaluating the throughput characteristics of CSMA-like networks, see for instance [3, 8, 9, 39]. Experimental results in Liew *et al.* [23] demonstrate that these models, while idealized, provide throughput estimates that match remarkably well with measurements in actual systems.

Despite their asynchronous and distributed nature, CSMA-like algorithms have been shown to offer the remarkable capability of achieving the full capacity region and thus match the optimal throughput performance of centralized scheduling mechanisms operating in slotted time [19, 20, 24]. More specifically, any throughput vector in the interior of the convex hull associated with the independent sets in the underlying interference graph can be achieved through suitable back-off rates and/or transmission lengths. Based on this observation, various ingenious algorithms have been developed for finding the back-off rates that yield a particular target throughput vector or that optimize a certain concave throughput utility function in scenarios with saturated buffers [19, 20, 26]. In the same spirit, several effective approaches have been devised for adapting the transmission lengths based on queue length information, and been shown to guarantee maximum stability [18, 29, 34, 35].

Roughly speaking, the maximum-stability guarantees were established under the condition that the activity factors of the various nodes behave as logarithmic functions of the queue lengths. Unfortunately, such activity factors can induce excessive queue lengths and delays, which has triggered a strong interest in developing approaches for improving the delay performance [16, 22, 25, 28, 33]. Motivated by this issue, Ghaderi & Srikant [15] recently showed that it is in fact sufficient for the *logarithms* of the activity factors to behave as logarithmic functions of the queue lengths, divided by an arbitrarily slowly increasing, unbounded function. These results indicate that the maximum-stability guarantees are preserved for activity functions that

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

are essentially linear for all practical values of the queue lengths, although asymptotically the activity rate must grow slower than any positive power of the queue length. A careful inspection reveals that the proof arguments leave little room to weaken the stated growth condition. Since the growth condition is only a sufficient one, however, it is not clear to what extent it is actually a strict requirement for maximum stability to be maintained.

In the present paper we explore the scope for using more aggressive activity functions in order to improve the delay performance while preserving the maximum-stability guarantees. Since the proof methods of [15, 18, 29, 34, 35] do not easily extend to more aggressive activity functions, we will instead adopt fluid limits where the dynamics of the system are scaled in both space and time. Fluid limits may be interpreted as first-order approximations of the original stochastic process, and provide valuable qualitative insight and a powerful approach for establishing (in)stability properties [5, 6, 7, 27].

As observed in [4], qualitatively different types of fluid limits can arise, depending on the structure of the interference graph, in conjunction with the functional shape of the activity factors. For sufficiently *tame* activity functions as in [15, 29, 34, 35], 'fast mixing' is guaranteed, where the activity process evolves on a much faster time scale than the scaled queue lengths. Qualitatively similar fluid limits can arise for more *aggressive* activity functions as well, provided the topology is benign in a certain sense, which implies that the maximum-stability guarantees are preserved in those cases. In different regimes, however, aggressive activity functions can cause 'sluggish mixing', where the activity process evolves on a much slower time scale than the scaled queue lengths, yielding oscillatory fluid limits that follow random trajectories. It is highly unusual for such random dynamics to occur, as in queueing networks typically the random characteristics vanish and deterministic limits emerge on the fluid scale. A few exceptions are known for various polling-type models as considered in [13, 21, 14].

The random nature of the fluid limits gives rise to several complications in the convergence proofs that are not commonly encountered. Since the random-access networks that we consider are fundamentally different from the polling type-models in the above-mentioned references, the fluid limits are qualitatively different as well, and require a substantially different approach to establish convergence. Specifically, we develop an approach based on stopping time sequences to deal with the switching probabilities governing the sample paths of the fluid limit process. While these proof arguments are developed in the context of random-access networks, several key components extend far beyond the scope of the present problem. Hence, we believe that the proof constructs are of broader methodological value in handling

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

random fluid limits and of potential use in establishing both stability and instability results for a wider range of models. For example, the methodology that we develop could be easily applied to prove the stability results for the random capture scheme as conjectured in work of Feuillet *et al.* [12].

The possible oscillatory behavior of the fluid limit itself does not necessarily imply that the system is unstable, and in some situations maximum stability is in fact maintained. In other scenarios, however, the fluid limit reflects that more aggressive activity functions may force the system into inefficient states for extended periods of time and produce instability. We will demonstrate instability for super-linear activity functions, but our proof arguments suggest that it can potentially occur for any activity factor that grows as a positive power of the queue lengths in networks with sufficiently many nodes. In other words, the growth conditions for maximum stability depend on the number of nodes, which seems loosely related to results in [17, 36, 37] characterizing how (upper bounds for) the mean queue length and delay scale as a function of the size of the network.

The remainder of the paper is organized as follows. In Section 2, we present a detailed model description. We introduce fluid limits and discuss the various qualitative regimes in Section 3. We then use the fluid limits to demonstrate the potential instability of aggressive activity functions in Sections 4 and 5. Simulation experiments are conducted in Section 6 to support the analytical results. In Section 7, we make some concluding remarks and identify topics for further research. Appendices at the end of the paper contain proofs of our results.

2. Model description.

Network, interference graph, and traffic model. We consider a network of several nodes sharing a wireless medium according to a random-access mechanism. The network is represented by an undirected graph $G = (V, E)$ where the set of vertices $V = \{1, \ldots, N\}$ correspond to the various nodes and the set of edges $E \subseteq V \times V$ indicate which pairs of nodes interfere. Nodes that are neighbors in the interference graph are prevented from simultaneous activity, and thus the independent sets correspond to the feasible joint activity states of the network. A node is said to be blocked whenever the node itself or any of its neighbors is active, and unblocked otherwise. Define $S \subseteq \{0, 1\}^N$ as the set of incidence vectors of all the independent sets of the interference graph, and denote by $\mathcal{C} = \text{conv}(S)$ the capacity region, with $\text{conv}(\cdot)$ indicating the convex hull operator.

Packets arrive at node i as a Poisson process of rate λ_i. The packet transmission times at node i are independent and exponentially distributed with mean $1/\mu_i$. Denote by $\rho_i = \lambda_i/\mu_i$ the traffic intensity of node i.

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

Let $U(t) \in S$ represent the joint activity state of the network at time t, with $U_i(t)$ indicating whether node i is active at time t or not. Denote by $X_i(t)$ the queue length at node i at time t, i.e., the number of packets waiting for transmission or in the process of being transmitted.

Queue-based random-access mechanism. As mentioned above, the various nodes share the medium in accordance with a random-access mechanism. When a node ends an activity period (consisting of possibly several back-to-back packet transmissions), it starts a back-off period. The back-off times of node i are independent and exponentially distributed with mean $1/\nu_i$. The back-off period of a node is suspended whenever it becomes blocked by activity of any of its neighbors, and only resumed once the node becomes unblocked again. Thus the back-off period of a node can only end when none of its neighbors are active. Now suppose a back-off period of node i ends at time t. Then the node starts a transmission with probability $\phi_i(X_i(t))$, with $\phi_i(0) = 0$, and begins a next back-off period otherwise. When a transmission of node i ends at time t, it releases the medium and begins a back-off period with probability $\psi_i(X_i(t^-))$, or starts the next transmission otherwise, with $\psi_i(1) = 1$. Equivalently, node i may be thought of as activating at an exponential rate $f_i(X_i(t))$, with $f_i(\cdot) = \nu_i\phi_i(\cdot)$, whenever it is unblocked at time t, and de-activating at rate $g_i(X_i(t))$, with $g_i(\cdot) = \mu_i\psi_i(\cdot)$, whenever it is active at time t. For conciseness, the functions $f_i(\cdot)$ and $g_i(\cdot)$ will be referred to as *activation* and *de-activation functions*, respectively.

There are two special cases worth mentioning that (loosely) correspond to random-access schemes considered in the literature before. First of all, in case $\phi_i(X_i) = 1$ and $\psi_i(X_i) = 0$ for all $X_i \geq 1$, node i starts a transmission each time a back-off period ends, and does not release the medium, i.e., continues transmitting until its entire queue has been cleared. This corresponds to the random-capture scheme considered in [12]. In case $\mu_i = 1$, $\nu_i = 1$, $\phi_i(X_i) = 1 - \psi_i(X_i)$, and $\psi_i(X_i) = 1/(1 + r_i(X_i))$, node i may be thought of as becoming (or continuing to be) active with probability $r_i(X_i(t))/(1 + r_i(X_i(t)))$ each time a unit-rate Poisson clock ticks. This roughly corresponds to the scheme considered in [15, 18, 29, 34, 35] based on Glauber dynamics with a 'weight' function $w_i(X_i) = \log(r_i(X_i))$, except that the latter scheme operates with a random round-robin clock, and uses $\tilde{w}_i(X_i) = w_i(X_i) \vee \frac{\epsilon}{2N} w_i(X_{\max})$, with $X_{\max} = \vee_{j=1}^{N} X_j$.

Network dynamics. Under the above-described queue-based schemes, the process $\{(U(t), X(t))\}_{t\geq 0}$ evolves as a continuous-time Markov process with state space $S \times \mathbb{N}_0^N$. Transitions (due to arrivals) from a state (U, X) to $(U, X + e_i)$ occur at rate λ_i, transitions (due to activations) from a state

(U, X) with $X_i \geq 1$, $U_i = 0$, and $U_j = 0$ for all neighbors of node i, to $(U + e_i, X)$ occur at rate $\nu_i f_i(X_i)$, transitions (due to transmission completions followed back-to-back by a subsequent transmission) from a state (U, X) with $U_i = 1$ (and thus $X_i \geq 1$) to $(U, X - e_i)$ occur at rate $\mu_i(1 - g_i(X_i))$, transitions (due to transmission completions followed by a back-off period) from a state (U, X) with $U_i = 1$ (and thus $X_i \geq 1$) to $(U - e_i, X - e_i)$ occur at rate $\mu_i g_i(X_i)$.

We are interested to determine under what conditions the system is stable, i.e., the process $\{(U(t), X(t))\}_{t\geq 0}$ is positive-recurrent. It is easily seen that $(\rho_1, \ldots, \rho_N) < \sigma \in \mathcal{C}$ is a necessary condition for that to be the case. In [15], it is shown that this condition is in fact also sufficient for weight functions of the form $w_i(X_i) = \log(1 + X_i)/y_i(X_i)$, where $y_i(X_i)$ is allowed to increase to infinity at an arbitrarily slow rate. For practical purposes, this means that the function $r_i(X_i)$ is essentially allowed to be linear, except that it must eventually grow to infinity slower than any positive power of X_i. Results in [4] suggest that more aggressive choices of the functions $f_i(\cdot)$ and $g_i(\cdot)$, which translate into functions $r_i(\cdot)$ that grow faster to infinity, can improve the delay performance. In view of these results, we will be particularly interested in such functions $r_i(\cdot)$, where the stability results of [15] do not apply. In order to examine under what conditions the system will remain stable then, we will examine fluid limits for the process $\{(U(t), X(t))\}_{t\geq 0}$ as introduced in the next section.

3. Qualitative discussion of fluid limits. Fluid limits may be interpreted as first-order approximations of the original stochastic process, and provide valuable qualitative insight and a powerful approach for establishing (in)stability properties [5, 6, 7, 27]. In this section we discuss fluid limits for the process $\{(U(t), X(t))\}_{t\geq 0}$ from a broad perspective, with the aim to informally exhibit their qualitative features in various regimes, and we deliberately eschew rigorous claims or proofs.

3.1. *Fluid-scaled process.* In order to obtain fluid limits, the original stochastic process is scaled in both space and time. More specifically, we consider a sequence of processes $\{(U^{(R)}(t), X^{(R)}(t))\}_{t\geq 0}$ indexed by a sequence of positive integers R, each governed by similar statistical laws as the original process, where the initial states satisfy $\sum_{i=1}^{N} X_i^{(R)}(0) = R$ and $X_i^{(R)}(0)/R \to Q_i$ as $R \to \infty$. The process $\{(U^{(R)}(Rt), \frac{1}{R}X^{(R)}(Rt))\}_{t\geq 0}$ is referred to as the fluid-scaled version of the process $\{(U^{(R)}(t), X^{(R)}(t)\}_{t\geq 0}$. Note that the activity process is scaled in time as well but not in space. For compactness, denote $Q^R(t) = \frac{1}{R}X^{(R)}(Rt)$. Any (possibly random) weak limit $\{Q(t)\}_{t\geq 0}$ of the sequence $\{Q^R(t)\}_{t\geq 0}$, as $R \to \infty$, is called a fluid limit.

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

It is worth mentioning that the above notion of fluid limit based on the continuous-time Markov process is only introduced for the convenience of the qualitative discussion below. For all the proofs of fluid limit properties and instability results we will rely on a rescaled linear interpolation of the uniformized jump chain (as will be defined in Appendix A.I), with a time-integral version of the $U(\cdot)$ component. This construction yields convenient properties of the fluid limit paths and allows us to extend the framework of Meyn [27] for establishing instability results for discrete-time Markov chains. (The original continuous-time Markov process has in fact the same fluid limit properties, but this is not directly relevant in any of the proofs.)

The process $\{(U^{(R)}(Rt), \frac{1}{R}X^{(R)}(Rt))\}_{t\geq 0}$ comprises two interacting components. On one hand, the evolution of the (scaled) queue length process $\frac{1}{R}X^{(R)}(Rt)$ depends on the activity process $U^{(R)}(Rt)$. On the other hand, the evolution of the activity process $U^{(R)}(Rt)$ depends on the queue length process $X^{(R)}(Rt)$ through the activation and de-activation functions $f_i(\cdot)$ and $g_i(\cdot)$. In many cases, a separation of time scales arises as $R \to \infty$, where the transitions in $U^{(R)}(Rt)$ occur on a much faster time scale than the variations in $Q^R(t) = \frac{1}{R}X^{(R)}(t)$. Loosely phrased, the evolution of $Q^R(t)$ is then governed by the time-average characteristics of $U^{(R)}(\cdot)$ in a scenario where $Q^R(t)$ is fixed at its instantaneous value.

In other cases, however, the transitions in $U^{(R)}(Rt)$ may in fact occur on a much slower time scale than the variations in $Q^R(t)$, or there may not be a separation of time scales at all. As a result, qualitatively different types of fluid limits can arise, as observed in [4], depending on the mixing properties of the activity process. These mixing properties, in turn, depend on the functional shape of the activation and de-activation functions $f_i(\cdot)$ and $g_i(\cdot)$, in conjunction with the structure of the interference graph G.

3.2. *Fast mixing: Smooth deterministic fluid limits.* We first consider the case of fast mixing. In this case, the transitions in $U^{(R)}(Rt)$ occur on a much faster time scale than the variations in $Q^R(t)$, and completely average out on the fluid scale as $R \to \infty$. Informally speaking, this entails that the mixing time of the activity process in a scenario with fixed activation rates $f_i(Rq_i)$ and de-activation rates $g_i(Rq_i)$ grows slower than R as $R \to \infty$. In order to obtain a rough bound for the mixing time, assume that $f_i(\cdot) \equiv f(\cdot)$, $g_i(\cdot) \equiv g(\cdot)$, and denote $h(x) = f(x)/g(x)$. Further suppose that $h(R) \to \infty$ as $R \to \infty$, and $h(aR)/h(R) \to \hat{h}(a)$ as $R \to \infty$, with $\hat{h}(a) > 0$ for any $a > 0$. The latter assumptions are satisfied, for example, when $h(x) = x^\gamma$, $\gamma > 0$, with $\hat{h}(a) = a^\gamma$, or when $h(x) = \log(x)$ with $\hat{h}(a) \equiv 1$. Without proof, we claim that the mixing time then grows at most at rate $f(R)^{m^*-1}g(R)^{-m^*}$ as

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

$R \to \infty$, with m^* the cardinality of a maximum-size independent set. Thus, fast mixing behavior is guaranteed when $f(\cdot)$ does not grow too fast, $g(\cdot)$ does not decay too fast, or m^* is sufficiently small, e.g.,

(i) $g(x) = g$ and $m^* = 1$;
(ii) $f(x) = x^{1/(m^*-1)-\delta}$, $g(x) = g$, and $m^* \geq 2$;
(iii) $f(x) = f$ and $g(x) \geq x^{-1/m^*+\delta}$;
(iv) $f(x) = f$, $g(x) = 1/\log(1+x)$;
(v) $f(x) = \log(1+x)$ and $g(x) = g$.

As mentioned above, the fluid limit then follows an entirely deterministic trajectory, which is described by a differential equation of the form

$$\frac{\mathrm{d}}{\mathrm{d}t}Q_i(t) = \lambda_i - \mu_i u_i(Q(t)),$$

as long as $Q(t) > 0$ (component-wise), with the function $u_i(\cdot)$ representing the fraction of time that node i is active. We may write

$$u_i(q) = \sum_{s \in S} s_i \pi(s;q),$$

with $\pi(s;q)$ denoting the fraction of time that the activity process resides in state $s \in S$ in a scenario with fixed activation rates $f_j(Rq_j)$ and deactivation rates $g_j(Rq_j)$ as $R \to \infty$. Let $S^* = \{s \in S : \sum_{i=1}^N s_i = m^*\}$ correspond to the collection of all maximum-size independent sets. Under the above-mentioned assumptions,

$$\begin{aligned}\pi(s;q) &= \lim_{R\to\infty} \frac{\prod_{i=1}^N h(Rq_i)^{s_i}}{\sum_{u\in S^*}\prod_{i=1}^N h(Rq_i)^{u_i}} \\ &= \frac{\prod_{i=1}^N \hat{h}(q_i)^{s_i}}{\sum_{u\in S^*}\prod_{i=1}^N \hat{h}(q_i)^{u_i}} = \frac{\exp(\sum_{i=1}^N s_i \log(\hat{h}(q_i)))}{\sum_{u\in S^*}\exp(\sum_{i=1}^N u_i \log(\hat{h}(q_i)))},\end{aligned}$$

for $s \in S^*$, while $\pi(s;q) = 0$ for $s \notin S^*$. In particular, if $h(x) = x^\gamma$, $\gamma > 0$, then

$$\pi(s;q) = \frac{\prod_{i=1}^N q_i^{\gamma s_i}}{\sum_{u\in S^*}\prod_{i=1}^N q_i^{\gamma u_i}} = \frac{\exp(\gamma \sum_{i=1}^N s_i \log(q_i))}{\sum_{u\in S^*}\exp(\gamma \sum_{i=1}^N u_i \log(q_i))},$$

for $s \in S^*$. Also, if $h(x) = \log(1+x)$, then $\pi(s;q) = 1/|S^*|$ for $s \in S^*$.

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

When some of the components of q are zero, i.e., some of the queue lengths are zero at the fluid scale, it is considerably harder to characterize $u_i(q)$, since the competition for medium access from the queues that are zero at the fluid scale still has an impact. It may be shown though that

$$\sum_{i=1}^{N} \rho_i \mathrm{I}_{\{q_i > 0\}} \leq (1-\epsilon) \sum_{i=1}^{N} u_i(q) \mathrm{I}_{\{q_i > 0\}}$$

for some $\epsilon > 0$, assuming that $(\rho_1, \ldots, \rho_N) < \sigma \in \mathcal{C}$. The latter inequality also holds when $q > 0$, noting that then $\sum_{i=1}^{N} u_i(q) = m^*$, while $\sum_{i=1}^{N} \rho_i \leq (1-\epsilon)m^*$ for some $\epsilon > 0$.

We conclude that almost everywhere

$$\begin{aligned} \sum_{i=1}^{N} \frac{1}{\mu_i} \frac{\mathrm{d}Q_i(t)}{\mathrm{d}t} &\leq \sum_{i=1}^{N} (\rho_i - u_i(Q(t))) \mathrm{I}_{\{Q_i(t) > 0\}} \\ &\leq -\epsilon \sum_{i=1}^{N} \rho_i \mathrm{I}_{\{Q_i(t) > 0\}}, \end{aligned}$$

as long as $Q(t) \neq 0$. This means that $Q(t) = 0$ for all $t \geq T$ for some finite $T < \infty$, which implies that the original Markov process is positive-recurrent [5, 7]. This agrees with the stability results in [15, 18, 29, 35, 34] for the case $f(X_i) = 1 - g(X_i)$ and $g(X_i) = 1/(1 + \exp(\tilde{w}(X_i)))$, $\tilde{w}(X_i) = w(X_i) \vee \frac{\epsilon}{2N} w(X_{\max})$ (with the minor differences noted in the previous section), and suggests that these results in fact hold without the need to know the maximum queue size $X_{\max}$.

Of course, in order to convert the above arguments into an actual stability proof, the informal characterization of the fluid limit needs to be rigorously justified. This is a major challenge, and not the real goal of the present paper, since we aim to demonstrate the opposite, namely that more aggressive activity or de-activation functions can cause instability. Strong evidence of the technical complications in establishing the fluid limits is provided by recent work of Robert & Véber [30]. Their work focuses on the simpler case of a single work-conserving resource (which corresponds to a full interference graph in the present setting) without any back-off mechanism, where the service rates of the various nodes are determined by a logarithmic function of their queue lengths.

3.3. *Sluggish mixing: Erratic random fluid limits.* With the above aim in mind, we now turn to the case of sluggish mixing. In this case, the transitions in $U^{(R)}(Rt)$ occur on a much slower time scale than the variations in $Q^R(t)$,

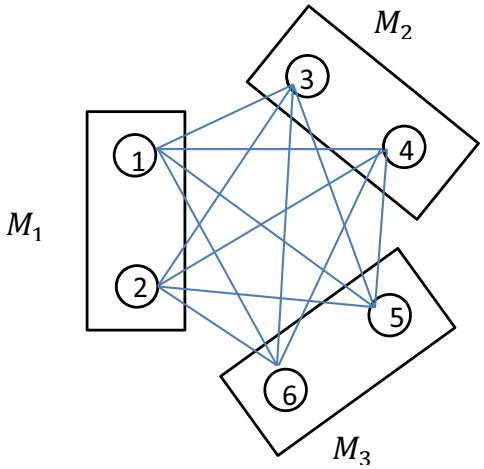


FIG 1. *The* diamond network*: A complete partite graph with* $K = 3$ *components, each containing 2 nodes.*

and vanish on the fluid scale as $R \to \infty$, except at time points where some of the queues hit zero. The detailed behavior of the fluid limit in this case depends delicately on the specific structure of the interference graph G and the shape of the functions $f_i(\cdot)$ and $g_i(\cdot)$. This prevents a characterization in any degree of generality, and hence we focus attention on some particular scenarios.

In order to show that sluggish mixing behavior itself need not imply instability, we first examine a complete K-partite graph as considered in [12], where the nodes can be partitioned into $K \geq 2$ components. All nodes are connected except those belonging to the same component. Figure 1 depicts an example of a complete partite graph with $K = 3$ components, each containing 2 nodes. We will refer to this network as the *diamond network*, since the edges correspond to those of an eight-faced diamond structure, with the node pairs constituting the three components positioned at the opposite ends of three orthogonal axes.

Denote by $M_k \subseteq \{1, \ldots, N\}$ the subset of nodes belonging the k-th component. Once one of the nodes in component M_k is active, other nodes within M_k can become active as well, but none of the nodes in the other components M_l, $l \neq k$, can be active. The necessary stability condition then takes the form $\rho = \sum_{k=1}^{K} \hat{\rho}_k < 1$, with $\hat{\rho}_k = \vee_{i \in M_k} \rho_i$ denoting the maximum traffic intensity of any of the nodes in the k-th component.

Now consider the case that each node operates with an activation function $f(x)$ with $\lim_{x\to\infty} f(x) > 0$ and a de-activation function $g(x) = o(x^{-\gamma})$, with $\gamma > 1$, which subsumes the random-capture scheme with $g(x) \equiv 0$ for all

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

$x \geq 1$ in [12]. Since the de-activation rate decays so sharply, the probability of a node releasing the medium once it has started transmitting with an initial queue length of order R, is vanishingly small, until the queue length falls below order R or the total number of transmissions exceeds order R (but the latter implies the former). Hence, in the fluid limit, a node must completely empty almost surely before it releases the medium. Because of the interference constraints, it further follows that once the activity process enters one of the components, it remains there until all the queues in that component have entirely drained (on the fluid scale), and then randomly switches to one of the other components. For conciseness, the fluid limit process is said to be in an M_k-period during time intervals when at least one of the nodes in component M_k is served at full rate (on the fluid scale).

Based on the above informal observations, we now proceed with a more detailed description of the dynamics of the fluid limit process. We do not aim to provide a proof of the stated properties, since the main goal of the present paper is to demonstrate the potential for instability rather than establish stability. However, the proof arguments that we will develop for a similar but more complicated interference graph in the remainder of the paper, could easily be applied to provide a rigorous justification of the fluid limit and establish the claimed stability results.

Assume that the system enters an M_k-period at time t, then

(a) It spends a time period $T_k(t) = \vee_{i \in M_k} \frac{Q_i(t)}{\mu_i - \lambda_i}$ in M_k.

(b) During this period, the queues of the nodes in M_k drain at a linear rate (or remain zero)

$$Q_i(t+u) = (Q_i(t) + (\lambda_i - \mu_i)u) \vee 0, \ \forall i \in M_k,$$

while the queues of the other nodes fill at a linear rate

$$Q_i(t+u) = Q_i(t) + \lambda_i u, \ \forall i \notin M_k,$$

for all $u \in [0, T_k(t)]$.

(c) At time $t + T_k(t)$, the system switches to an M_l-period, $l \neq k$, with probability

$$p_{kl}(t + T_k(t)) = \lim_{R \to \infty} \frac{\sum_{i \in M_l} f(RQ_i(t + T_k(t)))}{\sum_{l' \neq k,l} \sum_{i \in M_{l'}} f(RQ_i(t + T_k(t)))}.$$

Thus the fluid limit follows a piece-wise linear sample path, with switches between different periods governed by the transition probabilities specified above. Figure 2 depicts an example of the fluid limit sample path for the network of Figure 1 with $f(x) = 1$, $x \geq 1$.

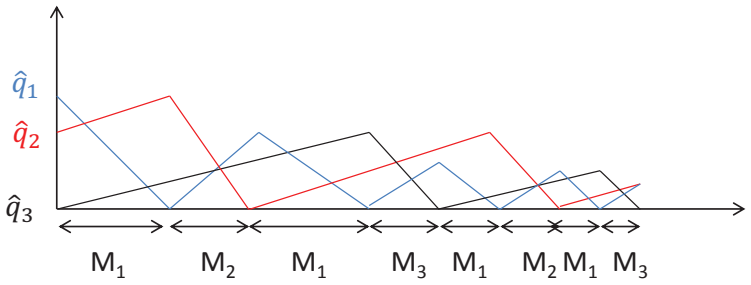


FIG 2. *A fluid limit sample path for the* diamond network *of Figure 1.*

Now define the Lyapunov function $L(t) := \sum_{k=1}^{K} \hat{Q}_k(t)$, with $\hat{Q}_k(t) = \vee_{i \in M_k} Q_i(t)/\mu_i$. Then, $\frac{\mathrm{d}}{\mathrm{d}t} L(t) \leq \sum_{k=1}^{K} \hat{\rho}_k - 1 = \rho - 1 < 0$ almost everywhere when $\rho < 1$, as long as $L(t) > 0$. Therefore, $L(t) = 0$, and hence $Q(t) = 0$, for all $t \geq T$, with $T = \frac{L(0)}{1-\rho} < \infty$, implying stability [5, 7], even though the fluid limit behavior is not smooth at all.

4. Fluid limits for broken-diamond network. In the previous section we discussed qualitative features of fluid limits in various scenarios, and in particular for so-called complete partite graphs. We now proceed to consider a 'nearly' complete partite graph, and will demonstrate that if some of the edges between two components M_k and M_l are removed (thus reducing interference), the network might become unstable for 'aggressive' activation and/or deactivation functions! Specifically, we will consider the *diamond network* of Figure 1, and remove the edge between nodes 4 and 5 to obtain a *broken-diamond network* with an additional component/maximal schedule M_4, as depicted in Figure 3.

The intuitive explanation for the potential instability may be described as follows. Denote $\rho_0 = \rho_1 \vee \rho_2$, and assume $\rho_3 \geq \rho_4$ and $\rho_6 \geq \rho_5$. It is easily seen that the fraction of time that at least one of the nodes 1, 2, 3 and 6 is served, must be no less than $\rho = \rho_0 + \rho_3 + \rho_6$ in order for these nodes to be stable. During some of these periods nodes 4 or 5 may also be served, but not simultaneously, i.e., schedule M_4 cannot be used. In other words, the system cannot be stable if schedule M_4 is used for a fraction of the time larger than $1 - \rho$. As it turns out, however, when the de-activation function is sufficiently aggressive, e.g., $g(x) = \mathrm{o}(x^{-\gamma})$, with $\gamma > 1$, schedule M_4 is in fact persistently used for a fraction of the time that does not tend to 0 as ρ approaches 1, which forces the system to be unstable.

Although the above arguments indicate that invoking schedule M_4 is a recipe for trouble, the reason may not be directly evident from the sys-

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

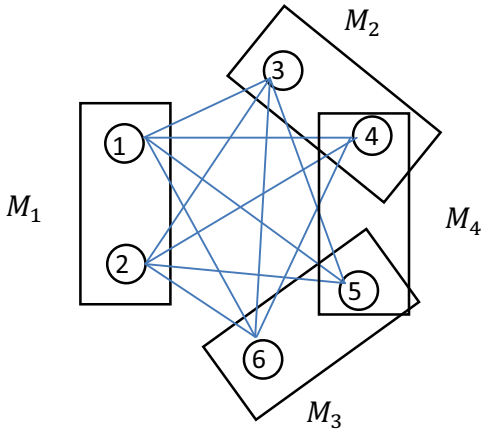


FIG 3. *The* broken-diamond network*, obtained by removing 1 edge from the* diamond network *of Figure 1, yielding an additional schedule* M_4.

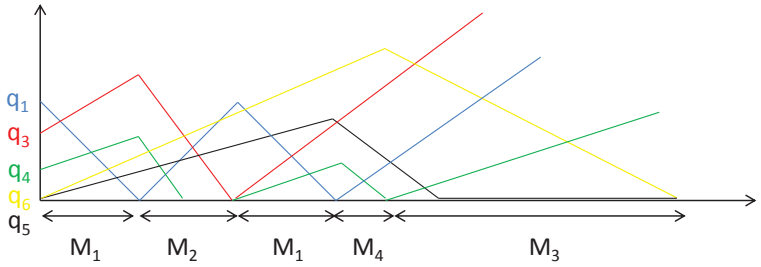


FIG 4. *A fluid limit sample path for the* broken-diamond network *of Figure 3, corresponding to the switching sequence* $M_1 \to M_2 \to M_1 \to M_4 \to M_3$.

tem dynamics, since no obvious inefficiency occurs as long as the queues of nodes 4 and 5 are non-empty. However, the fact that the Lyapunov function $L(t) = \sum_{k=1}^{3} \vee_{i \in M_k} Q_i(t)$ may increase while serving nodes 4 and 5, when $Q_3(t) \geq Q_4(t)$ and $Q_6(t) \geq Q_5(t)$, is already highly suggestive. (Such an increase is depicted in Figure 4 during the M_4-period of the switching sequence $M_1 \to M_2 \to M_1 \to M_4 \to M_3 \to M_1$.) Indeed, serving nodes 4 and 5 may make their queues smaller than those of nodes 3 and 6, leaving these queues to be served by themselves at a later stage, at which point inefficiency inevitably occurs.

In the sequel, the fluid limit process is said to be in a natural state when $Q_3(t) \geq Q_4(t)$ and $Q_6(t) \geq Q_5(t)$, with equality only when both sides are zero. We will assume $\lambda_3 > \lambda_4$ and $\lambda_6 > \lambda_5$, and will show that the process must always reside in a natural state after some finite amount of time. As

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

described above, instability is bound to occur when schedule M_4 is used repeatedly for substantial periods of time while the fluid limit process is in a natural state. Since the process is always in a natural state after some finite amount of time, it is intuitively plausible that such events occur repeatedly with positive probability, but a rigorous proof that this leads to instability is far from simple. Such a proof requires detailed analysis of the underlying stochastic process (in our case via fluid limits), and its conclusion crucially depends on the de-activation function. Indeed, the stability results in [15, 18, 29, 34, 35] indirectly indicate that the *broken-diamond network* is *not* rendered unstable for sufficiently cautious de-activation functions.

Just like for the complete partite graphs, the fluid limit process is said to be in an M_1-period when node 1 or node 2 (or both) is served at full rate. The process is in an M_2- or M_3-period when node 3 or 6 is served at full rate, respectively. The process is in an M_4-period when nodes 4 and 5 are both served at full rate simultaneously.

In Subsection 4.1 we will provide a detailed description of the dynamics of the fluid limit process once it has reached a natural state and entered an M_1-, M_2-, M_3- or M_4-period. The justification for the description follows from a collection of lemmas and propositions which are stated and proved in Appendices A–D, with a high-level outline provided in Subsection 4.2. In Section 5 we will exploit the properties of the fluid limit process in order to prove that the harmful behavior described above indeed occurs for sufficiently aggressive de-activation functions, implying instability of the fluid limit process as well as the original stochastic process.

4.1. *Description of the fluid limit process.* We now provide a detailed description of the dynamics of the fluid limit process once it has reached a natural state and entered an M_1-, M_2-, M_3- or M_4-period. For sufficiently high load, i.e., ρ sufficiently close to 1, a natural state and such a period occur in uniformly bounded time almost surely for any initial state. As will be seen, for de-activation functions $g_i(x) = o(x^{-\gamma})$, with $\gamma > 1$, the fluid limit process then follows similar piece-wise linear trajectories, with random switches, as described in the beginning of Section 4 for complete partite graphs and further illustrated in Figure 4. For notational convenience, we henceforth assume $\mu_i \equiv 1$, so that $\rho_i \equiv \lambda_i$, for all $i = 1, \ldots, N$, and additionally assume activation functions $f_i(x) \equiv 1$, $x \geq 1$, for all $i = 1, \ldots, N$.

4.1.1. *M_1-period.* Assume the system enters an M_1-period at time t, then

(a) It spends a time period $T_1(t) = \frac{Q_1(t)}{1-\rho_1} \vee \frac{Q_2(t)}{1-\rho_2}$ in M_1.

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

(b) During this period, the queues of nodes 1 and 2 drain at a linear rate (or remain zero)

$$Q_i(t+u) = (Q_i(t) - (1-\rho_i)u) \vee 0, \text{ for } i = 1, 2,$$

while the queues of nodes 3, 4, 5, and 6 fill at a linear rate

$$Q_i(t+u) = Q_i(t) + \rho_i u, \text{ for } i = 3, 4, 5, 6,$$

for all $u \in [0, T_1(t)]$. In particular, $Q_1(t + T_1(t)) = Q_2(t + T_1(t)) = 0$.

(c) At time $t+T_1(t)$, the system switches to an M_2-, M_3- or M_4-period with transition probabilities $p_{12} = \frac{3}{8}$, $p_{13} = \frac{3}{8}$, and $p_{14} = \frac{1}{4}$, respectively.

4.1.2. *M_2-period.* Assume that the system enters an M_2-period at time t, then

(a) The system spends a time period $T_2(t) = \frac{Q_3(t)}{1-\rho_3}$ in M_2.

(b) During this period, the queues of nodes 3 and 4 drain (or remain zero)

$$Q_i(t+u) = (Q_i(t) - (1-\rho_i)u) \vee 0, \text{ for } i = 3, 4,$$

while the queues of nodes 1, 2, 5, and 6 fill at a linear rate

$$Q_i(t+u) = Q_i(t) + \rho_i u, \text{ for } i = 1, 2, 5, 6,$$

for all $u \in [0, T_2(t)]$. In particular, $Q_3(t + T_2(t)) = 0$.

(c) At time $t + T_2(t)$, the system switches to an M_1- or M_3-period. Note that $\frac{Q_3(t)}{1-\rho_3} > \frac{Q_4(t)}{1-\rho_4}$ by the assumption that $\lambda_3 > \lambda_4$ and that the process has reached a natural state, so that $Q_3(t) > Q_4(t)$ (since $Q_3(t) = Q_4(t) = 0$ cannot occur at the start of an M_2-period). Thus node 4 has emptied before time $t + T_2(t)$, and remained empty (on the fluid scale) since then, precluding a switch to an M_4-period except for a negligible duration on the fluid scale), only allowing the system to switch to either an M_1- or M_3-period. The corresponding transition probabilities can be formally expressed in terms of certain stationary distributions, but are difficult to obtain in explicit form. Note that in order for any of the nodes 1, 2, 5 or 6 to activate, node 3 must be inactive. In order for nodes 1, 2 or 6 to activate, node 4 must be inactive as well, but the latter is not necessary in order for node 5 to activate. Since node 4 may be active even when it is empty on the fluid scale, it follows that node 5 enjoys an advantage in competing for access to the medium over nodes 1, 2 and 6. While it may be argued that node 4 is active with probability ρ_4 by the time node 3 becomes

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

inactive for the first time, the resulting probabilities for the various nodes to gain access to the medium first do not seem to allow a simple expression.

REMARK 1. If the process had not yet reached a natural state, the case $\frac{Q_3(t)}{1-\rho_3} \leq \frac{Q_4(t)}{1-\rho_4}$ could also arise. In case that inequality is strict, i.e., $\frac{Q_3(t)}{1-\rho_3} < \frac{Q_4(t)}{1-\rho_4}$, the queue of node 4 is still non-empty by time $t+T_2(t)$, simply forcing a switch to an M_4-period with probability 1.
In case of equality, i.e., $\frac{Q_3(t)}{1-\rho_3} = \frac{Q_4(t)}{1-\rho_4}$, however, the situation would be much more complicated, which serves as the illustration for the significance of the notion of a natural state. In order to describe these difficulties, note that the queues of nodes 3 and 4 both empty at time $t+T_2(t)$, barring a switch to an M_4-period, and permitting only a switch to either an M_1- or M_3-period. Just like before, node 5 is the only one able to activate during periods where node 3 is inactive while node 4 is active, and hence enjoys an advantage in competing for access to the medium. In fact, node 5 will gain access to the medium first almost surely if node 3 is the first one to become inactive (in the pre-limit). The probability of that event, and hence the transition probabilities to an M_1- or M_3-period, depends on queue length differences between nodes 3 and 4 at time t that can be affected by the history of the process and are not visible on the fluid scale.

4.1.3. *M_3-period.* The dynamics for an M_3-period are entirely symmetric to those for an M_2-period, and are therefore omitted.

4.1.4. *M_4-period.* Assume that the system enters an M_4-period at time t, then

(a) It spends a time period $T_4(t) = \frac{Q_4(t)}{1-\rho_4} \wedge \frac{Q_5(t)}{1-\rho_5}$ in M_4.
(b) During this period, the queues of nodes 4 and 5 drain at a linear rate

$$Q_i(t+u) = Q_i(t) - (1-\rho_i)u, \text{ for } i = 4, 5,$$

while the queues of nodes 1, 2, 3, and 6 fill at a linear rate

$$Q_i(t+u) = Q_i(t) + \rho_i u, \text{ for } i = 1, 2, 3, 6,$$

$u \in [0, T_4(t)]$. In particular, $Q_4(t+T_4(t)) \wedge Q_5(t+T_4(t)) = 0$.
(c) At time $t+T_4(t)$, the system switches to either an M_2- or M_3-period. In order to determine which of these events can occur, we need to distinguish between three cases, depending on whether $\frac{Q_4(t)}{1-\rho_4}$ is (i) larger than, (ii) equal to, or (iii) smaller than $\frac{Q_5(t)}{1-\rho_5}$.

In case (i), i.e., $\frac{Q_4(t)}{1-\rho_4} > \frac{Q_5(t)}{1-\rho_5}$, we have $Q_4(t+T_4(t)) > 0$, i.e., the queue of node 4 is still non-empty by time $t + T_4(t)$, causing a switch to an M_2-period with probability 1.
In case (ii), i.e., $\frac{Q_4(t)}{1-\rho_4} = \frac{Q_5(t)}{1-\rho_5}$, we have $Q_4(t+T_4(t)) = Q_5(t+T_4(t)) = 0$, i.e., the queues of nodes 4 and 5 both empty at time $t + T_4(t)$. Even though both queues empty at the same time on the fluid scale, there will with overwhelming probability be a long period in the pre-limit where one of the nodes has become inactive for the first time while the other one has yet to do so. Since both nodes 4 and 5 must be inactive in order for nodes 1 and 2 to activate, these nodes have no chance to activate during that period, but either node 3 or node 6 does, depending on whether node 5 or node 4 is the first one to become inactive. As a result, the system cannot switch to an M_1-period, but only to an M_2- or M_3-period. In fact, a switch to M_2 will occur almost surely if node 5 is the first one to become inactive, while a switch to M_3 will occur almost surely if node 4 is the first one to become inactive. The probabilities of these two scenarios, and hence the transition probabilities to M_2 and M_3, depend on queue length differences between nodes 4 and 5 at time t that are affected by the history of the process and are not visible on the fluid scale.
In case (iii), i.e., $\frac{Q_4(t)}{1-\rho_4} < \frac{Q_5(t)}{1-\rho_5}$, we have $Q_5(t + T_4(t)) > 0$, i.e., the queue of node 5 is still non-empty by time $t + T_4(t)$, forcing a switch to an M_3-period with probability 1.

Remark 2. As noted in the above description of the fluid limit process, in cases 2(c), 3(c), and 4(c)(ii) the transition probabilities from an M_2-period to an M_1- or M_3-period, from an M_3-period to an M_1- or M_2-period, and from an M_4- to an M_2- or M_3-period, depend on queue length differences that are affected by the history of the process and are not visible on the fluid scale. Depending on whether or not the initial state and parameter values allow for these cases to arise, it may thus be impossible to provide a probabilistic description the evolution of the resulting fluid limit process, even in terms of its entire own history.

4.2. *Overview of fluid limit proofs.* In Section 4.1, we provided a description of the dynamics of the fluid limit process once it has reached a natural state and entered an M_1, M_2-, M_3 or M_4-period. As was further stated, for ρ sufficiently close to 1, a natural state and such a period occur in uniformly bounded time almost surely for any initial state. The justification for all these properties follows from a series of lemmas and propositions stated and

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

proved in Appendices A–D. In this section we present a high-level outline of the fluid limit statements and proofs.

First of all, recall that the description of the fluid limit process referred to the continuous-time Markov process representing the system dynamics as introduced in Section 2. For all the proofs of fluid limit properties and instability results however we consider a rescaled linear interpolation of the uniformized jump chain (as defined in Appendix A.I). This construction yields convenient properties of the fluid limit paths and allows us to extend the framework of Meyn [27] for establishing instability results for discrete-time Markov chains. (The original continuous-time Markov process has in fact the same fluid limit properties, but this is not directly relevant in any of the proofs.)

The proofs of the fluid limit properties consist of four main parts. Part A identifies several basic properties of the fluid limit paths, and in particular establishes that the queue length trajectory of each of the individual nodes exhibits 'sawtooth' behavior. This fundamental property in fact holds in arbitrary interference graphs, and only requires an exponent $\gamma > 1$ in the backoff probability. Part B of the proof shows a certain dominance property, saying that if all the interferers of a particular node also interfere with some other node that is currently being served at full rate, then the former node must be empty or served at full rate (on the fluid scale) as well. Under the assumption $\lambda_3 > \lambda_4$, $\lambda_5 < \lambda_6$, the dominance property implies that after a finite amount of time the fluid limit process for the broken-diamond network must always reside in a natural state (as defined in Section 4.1). Part C of the proof centers on the M_1-, M_2-, M_3- and M_4-periods, and establishes that at the end of any such period, the process immediately switches to one of the other types of periods with the probabilities indicated in Section 4.1. In particular, it is deduced that an M_4-period cannot be entered from an M_2- or M_3-period, and must always be preceded by an M_1-period once the process has reached a natural state. The combination of the sawtooth queue length trajectories and the switching probabilities provides a probabilistic description of the dynamics of the fluid limit once the process has reached a natural state and entered an M_1-, M_2-, M_3- or M_4-period. Part B already established that the process must always reside in a natural state after a finite amount of time, but it remains to be shown that the process will inevitably enter an M_1-, M_2-, M_3- or M_4-period, which constitutes the final Part D of the proof. The core argument is that interfering empty and nonempty queues can not coexist, since the empty nodes will frequently enter back-off periods, offering the nonempty nodes abundant opportunities to gain access, drain their queues, and cause the empty nodes to build queues in turn.

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

Now we provide additional guidance to the reader regarding the technical results in the appendix. Part A of the proof starts with the simple observation that, by the 'skip-free' property of the original pre-limit process, the sample paths of the interpolated version of the uniformized jump chain are Lipschitz continuous with a uniform Lipschitz constant (see (A.3)). It follows from this property and our initial assumptions that fluid limits exist (see Theorem 3 and the surrounding discussion). In addition, the fluid limit paths (associated with a fluid limit) are also Lipschitz continuous with the same constant, and are thus differentiable almost everywhere with probability one.

We then go on to show the fluid limit paths are determined by a countable set of 'entrance' times and 'exit' times of $(0, \infty)$ with probability one. The exit times for each queue ℓ are defined in terms of countably many sequences of stopping times (see Definition 4, Corollary 2, and the preceding discussion). With these in hand, the entrance times are seen to be random variables, measurable with respect to the pre-T σ-algebras of the exit times. Next we show that between an exit time and its corresponding entrance time the sample path is 'sawtooth'. To do this we show that if a nonempty node (on the fluid scale) receives any amount of service during some time interval, then it must in fact be served at the full rate until it has completely emptied (on the fluid scale), assuming $\gamma > 1$ (see Lemma 8). This implies that when node i is nonempty (on the fluid scale), its queue must either increase at rate λ_i or decrease at rate $1 - \lambda_i$ until it has entirely drained. In other words, the queue length trajectory of each of the individual nodes exhibits sawtooth behavior (Theorem 4).

Part B of the proof pertains to the joint behavior of the fluid limit trajectories of the various queue lengths. First of all, the natural property is proved that whenever a particular node is served, none of its interferers can receive any service (Lemma 3). Second, it is established that whenever a particular node is served, any node whose interferers are a subset of those of the node served, must either be empty or be served at full rate as well (on the fluid scale) (Corollary 3). For example, in the broken-diamond network, whenever node 3 is served, node 4 must either be empty or be served at full rate as well, and similarly for nodes 5 and 6. These two properties combined yield a dominance property, saying that if all the interferers of a particular node also interfere with some other node that is currently being served at full rate, then the former node must be empty or served at full rate (on the fluid scale) as well. In the case of the broken-diamond network, under the assumption $\lambda_3 > \lambda_4$, the queue of node 3 will therefore never be smaller than that of node 4 after some finite amount of time, and similarly

for nodes 4 and 5. Thus the fluid limit process will always reside in a natural state after some finite amount of time. In Theorem 5 it is shown that not only is this the case but also for utilizations ρ sufficiently close to 1 the network is non-empty at the time of first entrance into a natural state.

Part C of the proof focuses on the M_1-, M_2-, M_3- and M_4-periods as described above. Because of the sawtooth behavior, an M_1-period can only end when both nodes 1 and 2 are empty (on the fluid scale). Likewise, an M_2- or M_3-period can only end when node 3 or node 6 is empty, respectively. An M_4-period can only end when node 4 or node 5 (or both) is empty. It is then proven that at the end of an M_1-period, the fluid limit process immediately switches to an M_2-, M_3- or M_4-period with the probabilities specified in Section 4.1 (see Theorem 6 and also Lemma 13 and Corollary 5 which state related results for certain finite dimensional sets of paths). When the process resides in a natural state, an M_2-period is always instantaneously followed by an M_1- or M_3-period, while an M_3-period is always instantaneously followed by an M_1- or M_2-period. In particular, it is concluded that an M_4-period cannot be entered from an M_2- or M_3-period, and must always be preceded by an M_1-period once the process has reached a natural state. In the case of switching from an M_4-period our results only show that the probability of entering M_1 is 0, following entry into a natural state, that is switching is either to an M_2- or M_3-period (see Theorem 8). The actual switching probabilities for these two events may have arbitrary dependencies, however Theorem 8 is all that is needed for our main results.

There is no reason a priori however that the process is guaranteed to actually ever enter an M_1-, M_2-, M_3- or M_4-period. In fact, the process may very well spend time in different kinds of states, but the final Part D of the proof establishes that these kinds of states are transient, and cannot occur once a natural state has been reached, which is forced to happen in a finite amount of time for particular arrival rates as was already shown in Part B. Note that an M_1-, M_2-, M_3- or M_4-period occurs as soon as node 1, node 2, node 3, node 6 or nodes 4 and 5 simultaneously are served at full rate. In other words, the only ways for the process to avoid an M_1-, M_2-, M_3- or M_4-period, are: (i) for node 4 to be served at full rate, but not nodes 3 and 5; (ii) for node 5 to be served at full rate, but not nodes 4 and 6; (iii) for none of the nodes to be served at full rate. Scenario (i) requires node 3 to be empty (on the fluid scale) and node 4 to be nonempty, which can not occur in a natural state. Likewise, scenario (ii) cannot arise in a natural state either. Scenario (iii) requires that every empty node i is served at rate ρ_i (on the fluid scale), while all nonempty nodes are served at rate 0. Such a scenario is not particularly plausible, but a rigorous proof is quite involved

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

(see Theorem 9). Theorem 9 is treated by cases and we outline the arguments in Lemma 16 and Lemma 17, for the case nodes 3,4,5,6 empty, leaving the proofs of these as well as the other cases as an exercise for the reader. The main idea is to introduce the notion of a 'control-swap' (see Definition 8), whereby there is always a positive probability of an M_1-period within a bounded number of control-swaps. Second it is shown that control-swaps occur infinitely often and with negligible fluid time for the network 3,4,5,6 in isolation. Thus our insights rely strongly on the specific properties of the broken-diamond network, and an extension to arbitrary graphs does not seem straightforward.

5. Instability results for broken-diamond network. In Section 4, we provided a detailed description of the dynamics of the fluid limit process once it has reached a natural state and entered an M_1-, M_2-, M_3 or M_4-period. In this section we exploit the properties of the fluid limit process in order to prove that it is unstable for ρ sufficiently close to 1, and then show how the instability of the original stochastic process can be deduced from the instability of the fluid limit process.

5.1. *Instability of the fluid limit process.* In order to prove instability of the fluid limit process, we first revisit the intuitive explanation discussed earlier, see Figure 4 for an illustration. Denote $\rho_0 = \rho_1 \vee \rho_2$, and recall that $\rho_3 \geq \rho_4$ and $\rho_5 \leq \rho_6$ by assumption. Since nodes 1, 2, 3 and 6 are only served during M_1-, M_2- and M_3-periods, and not during M_4-periods, it is easily seen that the fraction of time that the system spends in M_1-, M_2- and M_3-periods must be no less than $\rho = \rho_0 + \rho_3 + \rho_6$ in order for these nodes to be stable. Thus, the system cannot be stable if it spends a fraction of the time larger than $1 - \rho$ in M_4-periods. As it turns out, however, when the de-activation function is sufficiently aggressive, e.g., $g(x) = \mathrm{o}(x^{-\gamma})$, with $\gamma > 1$, M_4-periods in fact persistently occur for a fraction of time that does not tend to 0 as ρ approaches 1, which forces the system to be unstable.

Figure 4 shows a fluid-limit sample path corresponding to the switching sequence $M_1 \to M_2 \to M_1 \to M_4 \to M_3 \to M_1$. The aggregate queue size starts building up in the M_3-period that follows the M_4-period.

In order to prove instability of the fluid limit process, we adopt the Lyapunov function $L(t) = \sum_{k=1}^{3} \vee_{i \in M_k} Q_i(t)$, and will show that the load $L(t)$ grows without bound almost surely. Note that the load $L(t)$ increases during M_4-periods while the process is in a natural state.

In preparation for the instability proof, we first state two auxiliary lemmas. It will be convenient to view the evolution of the fluid limit process, and in particular the Lyapunov function $L(t)$, over the course of *cycles*. The

i-th *cycle* is the period from the start of the $(i-1)$-th M_1-period to the start of the i-th M_1-period *once the fluid limit process has reached a natural state*. Denote by t_i the start time of the i-th cycle, $i = 1, 2, \ldots$. Each t_i is finite almost surely for ρ sufficiently close to 1, and in particular an infinite number of cycles must occur almost surely. In order to see that, recall that the fluid limit process will reach a natural state and enter an M_1, M_2, M_3- or M_4-period in finite time almost surely for any initial state as stated in Section 4.1. The description of the dynamics of the fluid limit process provided in that subsection then implies that M_1-periods and hence cycles must occur infinitely often (and if only finitely many M_1-periods occurred, then at least one of the nodes would in fact never be served again after some finite time, implying that the fluid limit process is unstable regardless).

The next lemma shows that the duration of a cycle and the possible increase in the load over the course of a cycle are linearly bounded in the load at the start of the cycle.

LEMMA 1. *The duration of the i-th cycle and the increase in the load over the course of the i-th cycle, $L(t_{i+1}) - L(t_i)$, are bounded from above by*

$$t_{i+1} - t_i \le C_T L(t_i) \text{ and } L(t_{i+1}) - L(t_i) \le C_L L(t_i),$$

for all $\rho \le 1$, where $C_T = \frac{1}{1-\rho_3-\rho_6}(\frac{1}{1-\rho_0} + \frac{1}{1-(\rho_4 \vee \rho_5)})$ and $C_L = \frac{\rho}{1-(\rho_4 \vee \rho_5)}$.

The proof of the above lemma is presented in Appendix E.

In order to establish that the durations of M_4-periods are non-negligible, it will be useful to introduce the notion of 'weakly-balanced' queues, ensuring that the queues of nodes 4 and 5 are not too small compared to the queues of nodes 3 and 6.

DEFINITION 1. Let $\beta^{\min}$ and $\beta^{\max}$ be fixed positive constants. The queues are said to be weakly-balanced in a given cycle (with respect to $\beta^{\min}$ and $\beta^{\max}$) if $\beta^{\min} \le \frac{Q_3(t)}{Q_5(t)}, \frac{Q_6(t)}{Q_4(t)} \le \beta^{\max}$, with t denoting the time when the M_1-period ends that initiated the cycle.

The next lemma shows that over two consecutive cycles, the queues will be weakly-balanced with probability at least 1/3.

LEMMA 2. *Let*

$$\epsilon = \frac{\rho_2}{2\left(\rho_2 + (\rho_3+\rho_6)\frac{1-(\rho_4 \wedge \rho_5)}{1-(\rho_4 \vee \rho_5)}\right)} \ge \frac{\rho_2}{\rho}\frac{1-(\rho_4 \vee \rho_5)}{1-(\rho_4 \wedge \rho_5)}.$$

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

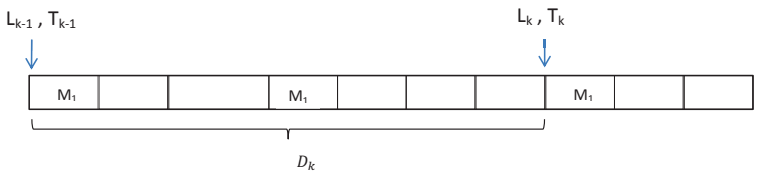


FIG 5. *A cycle D_k consisting of a pair of consecutive cycles.*

Then over two consecutive cycles, with probability at least 1/3, the queues will be weakly-balanced in at least one of these cycles with

$$\beta^{\max} = \frac{(\rho_3 \vee \rho_6) + (1-\rho_2)(1-\epsilon)/\epsilon}{\rho_4 \wedge \rho_5},$$

and $\beta^{\min} = 1/\beta^{\max}$.

The proof of the above lemma is presented in Appendix E.

As suggested by the above lemma, it will be convenient to consider pairs of two consecutive cycles in order to prove instability of the fluid limit process.

Let D_k be the pair of cycles consisting of cycles $2k-1$ and $2k$ as in Figure 5, $k = 1, 2, \ldots$. With minor abuse of notation, denote by $T_k = t_{2k-1}$ the start time of D_k and $L_k = L(T_k)$. Denote by $\Delta T_k = T_{k+1} - T_k$ the duration of D_k and by $\Delta L_k = L_{k+1} - L_k$ the increase in $L(t)$ over the course of D_k.

The next proposition shows that for ρ sufficiently close to 1 the load cannot significantly decrease over a pair of cycles and will increase by a substantial amount with non-zero-probability. We henceforth assume

$$(\rho_1, \rho_2, \rho_3, \rho_4, \rho_5, \rho_6) = \rho(\kappa_1, \kappa_2, \kappa_3, \kappa_3 - \alpha, \kappa_6 - \alpha, \kappa_6),$$

with $(\kappa_1 \vee \kappa_2) + \kappa_3 + \kappa_6 = 1$ and $0 < \alpha < \kappa_3 \wedge \kappa_6$, so that $\rho = \rho_0 + \rho_3 + \rho_6$.

PROPOSITION 1. *Let $C_{LT} = C_T(2 + C_L)$, with C_T and C_L as specified in Lemma 1, $\theta = 1 - (1-\rho)C_{LT}$, $p = 1/12$. Over cycle pairs D_k, $k = 1, 2, \ldots$,*

(i) $\Delta T_k \le C_{LT} L_k$;
(ii) $L(t) \ge \theta L_k$ *for all* $t \in [T_k, T_{k+1}]$;
(iii) $\mathbb{P}\big(L_{k+1} - \theta L_k \ge \delta(\rho)\theta L_k | L_k\big) \ge p$, *with* $\delta(\rho)$ *a constant, depending on* ρ, *and* $\delta(\rho) \uparrow \delta = \frac{1}{\beta^{\max}(1+\beta^{\max})(1+\alpha-(\kappa_3 \wedge \kappa_6))}$, *as* $\rho \uparrow 1$.

PROOF. We first show part (i). Using Lemma 1, we find

$$\Delta T_k = \Delta t_{2k-1} + \Delta t_{2k} \le C_T(L(t_{2k-1}) + L(t_{2k})) \le C_T(2 + C_L)L_k.$$

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

In order to prove part (ii), note that $L(t)$ cannot decrease at a larger rate than $1-\rho$, so that in view of part (i),

$$L(t) \geq L_k - (1-\rho)(t-T_k) \geq L_k - (1-\rho)\Delta T_k \geq (1-(1-\rho)C_{LT})L_k = \theta L_k,$$

for all $t \in [T_k, T_{k+1}]$.

We now turn to part (iii). Suppose that the following event occurs: the queues are weakly-balanced at the end of an M_1-period, say time τ, during D_k (which according to Lemma 2 happens with at least probability 1/3) and the system then enters an M_4-period (which happens with probability 1/4). Recalling that $\rho_3 > \rho_4$, $Q_3(t) \geq Q_4(t)$, $\rho_5 < \rho_6$ and $Q_5(t) \leq Q_6(t)$, we find that during the M_4-period $L(t)$ increases by

$$\rho\left(\frac{Q_4(\tau)}{1-\rho_4} \wedge \frac{Q_5(\tau)}{1-\rho_5}\right) \geq \rho\frac{Q_4(\tau) \wedge Q_5(\tau)}{1-\rho(\kappa_3 \wedge \kappa_6)+\rho\alpha}.$$

Since the queues are weakly-balanced, we deduce

$$\begin{aligned} Q_3(\tau) &\leq \beta^{\max} Q_5(\tau) \leq \beta^{\max} Q_6(\tau) \leq (\beta^{\max})^2 Q_4(\tau), \\ Q_6(\tau) &\leq \beta^{\max} Q_4(\tau) \leq \beta^{\max} Q_3(\tau) \leq (\beta^{\max})^2 Q_5(\tau). \end{aligned}$$

Noting that $Q_1(\tau) = Q_2(\tau) = 0$, we obtain

$$L(\tau) = Q_3(\tau) + Q_6(\tau) \leq (1+\beta^{\max})Q_6(\tau) \leq \beta^{\max}(1+\beta^{\max})Q_4(\tau),$$

and also

$$L(\tau) = Q_3(\tau) + Q_6(\tau) \leq (1+\beta^{\max})Q_3(\tau) \leq \beta^{\max}(1+\beta^{\max})Q_5(\tau).$$

So

$$L(\tau) \leq \beta^{\max}(1+\beta^{\max})(Q_4(\tau) \wedge Q_5(\tau)),$$

and thus the increase in $L(t)$ during the M_4-period is no less than $\delta(\rho)L(\tau)$, with

$$\delta(\rho) = \frac{\rho}{\beta^{\max}(1+\beta^{\max})(1-\rho(\kappa_3 \wedge \kappa_6)+\rho\alpha)}.$$

Using part (i) once again, we conclude that with at least probability 1/12,

$$\begin{aligned} L_{k+1} &\geq L_k + \delta(\rho)L(\tau) - (1-\rho)\Delta T_k \\ &\geq L_k + \delta(\rho)(L_k - (1-\rho)\Delta T_k) - (1-\rho)\Delta T_k \\ &= (1+\delta(\rho))(L_k - (1-\rho)\Delta T_k) \geq (1+\delta(\rho))(L_k - (1-\rho)C_{LT}L_k) \\ &= (1+\delta(\rho))\theta L_k. \end{aligned}$$

□

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

Armed with the above proposition, we now proceed to prove that the fluid limit process is unstable, in the sense that $L(T) \to \infty$ as $T \to \infty$. In fact, $L(T)$ grows faster than any sub-linear function $T^{\frac{1}{m}}$, $m > 1$, as stated in the next theorem.

THEOREM 1. *For any $m > 1$, there exists a constant $\rho^* = \rho^*(\kappa, m) < 1$, such that for all $\rho \in (\rho^*, 1]$,*

$$\limsup_{T\to\infty} \mathbb{E}\left[\frac{T}{L^m(T)}\right] = 0,$$

for any initial state $\mathbf{Q}(0)$ with $||\mathbf{Q}(0)|| = 1$, and $||\cdot||$ denoting the L_1-norm.

PROOF. Consider the cycle pairs D_k, $k = 1, 2, \ldots$, as defined right before Proposition 1. Assume $\rho \in (1-\frac{1}{C_{LT}}, 1]$, so that $\theta \in (0, 1]$ in Proposition 1. For any time $t > T_1$, we can define a stopping time N_t such that $T_{N_t} < t \le T_{N_t+1}$, i.e., t is within the N_t-th cycle pair. (This is possible almost surely, since $T_k \to \infty$ as $k \to \infty$ almost surely, as will be proved below.) Recall that $T_{N_t+1} \le T_{N_t} + C_{LT} L_{N_t}$ and $L(t) \ge \theta L_{N_t}$ by parts (i) and (ii) of Proposition 1, respectively, and trivially $L_{N_t} \le L(0) + \rho T_{N_t} \le 2T_{N_t}$ for t sufficiently large. Thus,

$$\begin{aligned}
\limsup_{t\to\infty} \mathbb{E}\left[tL^{-m}(t)\right] &\le \limsup_{t\to\infty} \mathbb{E}\left[T_{N_t+1}\theta^{-m}L_{N_t}^{-m}\right] \\
&\le \theta^{-m} \limsup_{t\to\infty} \mathbb{E}\left[T_{N_t}L_{N_t}^{-m}\right] + \theta^{-m}C_{LT} \limsup_{t\to\infty} \mathbb{E}\left[L_{N_t}^{-m+1}\right] \qquad (5.1)\\
&\le \theta(1+2C_{LT}) \limsup_{t\to\infty} \mathbb{E}\left[T_{N_t}L_{N_t}^{-m}\right].
\end{aligned}$$

So it suffices to prove that there exists $\rho^* = \rho^*(\kappa, m) < 1$ such that (5.1) is zero for $\rho > \rho^*$, which we now proceed to show.

First of all, by Proposition 1, for any $m > 0$,

$$\begin{aligned}
\mathbb{E}\left[L_{k+1}^{-m}|\mathcal{F}_k\right] &\le (1-p)(\theta L_k)^{-m} + p(\theta(1+\delta)L_k)^{-m} \qquad (5.2)\\
&= \alpha_m L_k^{-m},
\end{aligned}$$

where $\mathcal{F}_k$ is a suitable filtration and $\alpha_m := (1-p)\theta^{-m} + p(\theta(1+\delta))^{-m}$.

Since $\theta(\rho) \to \theta(1) = 1$ and $\delta(\rho) \to \delta(1) = \delta > 0$ as $\rho \uparrow 1$, $\alpha_m(\rho)$ is a continuous function of ρ in the vicinity of 1. Because $\alpha_m(1) < 1$, there must exist a $\rho_m^* = \rho^*(\kappa, m) < 1$ such that $\alpha_m < 1$ for all $\rho > \rho^*$. This shows that, for $\rho > \rho_m^*$, L_k^{-m} is a positive (geometric) supermartingale with parameter $\alpha_m < 1$. Taking expectations on both sides of (5.2) yields

$$\mathbb{E}\left[L_k^{-m}\right] \le \alpha_m^k L_0^{-m}. \qquad (5.3)$$

with $L_0 = L(t_{i_0}) > 0$ as noted earlier. In particular, $\lim_{k\to\infty} \mathbb{E}[L_k^{-m}] = 0$, and $1/L_k \to 0$ almost surely as $k \to \infty$ by Doob's supermartingale-convergence Theorem (page 147 of [31]). This implies that $T_k \to \infty$ almost surely because $L_k \le \rho T_k + 1 \le T_k + 1$. Therefore, the stopping time T_{N_t} is well-defined.

Next, consider the sequence of random variables $T_k L_k^{-m}$ with $m > 1$. Using Proposition 1,

$$\begin{aligned} \mathbb{E}\left[T_k L_k^{-m} | \mathcal{F}_{k-1}\right] &\le (T_{k-1} + C_{LT} L_{k-1}) \mathbb{E}\left[L_k^{-m} | \mathcal{F}_{k-1}\right] \\ &\le (T_{k-1} + C_{LT} L_{k-1}) \alpha_m L_{k-1}^{-m} \\ &= \alpha_m T_{k-1} L_{k-1}^{-m} + \alpha_m C_{LT} L_{k-1}^{-m+1}. \end{aligned} \tag{5.4}$$

Define $\epsilon_k := C_{LT} \alpha_m L_k^{-m+1}$, then, by (5.2) and (5.3), ϵ_k is a positive (geometric) super-martingale with parameter $\alpha_{m-1} < 1$ for $\rho > \rho^*_{m-1} = \rho^*(\kappa, m-1)$. Then, $\sum_{k=1}^{\infty} \mathbb{E}[\epsilon_k] \le C_{LT} \alpha_m \sum_{k=1}^{\infty} \alpha_{m-1}^k < \infty$, which shows that $\lim_{k\to\infty} T_k L_k^{-m} = 0$ almost surely. In particular, define $\alpha := \alpha_m \vee \alpha_{m-1}$ and $\rho^* = \rho^*_m \vee \rho^*_{m-1}$, then taking expectations on both sides of (5.5) yields

$$\mathbb{E}\left[T_k L_k^{-m}\right] \le \alpha \mathbb{E}\left[T_{k-1} L_{k-1}^{-m}\right] + \alpha C_{LT} \alpha^{k-1}, \tag{5.5}$$

which, by induction, shows that

$$\mathbb{E}\left[T_k L_k^{-m}\right] \le \alpha^{k-1} (\mathbb{E}\left[T_1 L_1^{-m}\right] + C_{LT}(k-1)\alpha), \tag{5.6}$$

for $\rho \in (\rho^*, 1]$. Now observe that T_1 is strictly bounded and L_1 is bounded away from zero, since a natural state is reached in finite time, before the system can empty, almost surely. It then follows that $\lim_{k\to\infty} \mathbb{E}[T_k L_k^{-m}] = 0$.

The fact that $T_k L_k^{-m}$ converges in $\mathcal{L}_1$ implies that the sequence of random variables $T_k L_k^{-m}$ is *Uniformly Integrable* (UI) (page 147, Theorem 50.1 of [31]). It therefore follows, by adapting the arguments of Doob's optional sampling theorem (page 159 of [31]), that the family of random variables $\{T_{N_t} L_{T_{N_t}}^{-m}\}$ is also UI. Thus by definition, given $\varepsilon > 0$, there exists K_ε such that

$$\mathbb{E}\left[T_{N_t} L_{N_t}^{-m} \mathrm{I}_{\left\{T_{N_t} L_{N_t}^{-m} \ge K_\varepsilon\right\}}\right] \le \varepsilon, \quad \forall t > 0$$

We deduce

$$\begin{aligned} \mathbb{E}\left[T_{N_t} L_{N_t}^{-m}\right] &\le \sum_{k=1}^{\infty} \mathbb{E}\left[T_k L_k^{-m} \mathrm{I}_{\{N_t = k\}} \mathrm{I}_{\left\{T_{N_t} L_{N_t}^{-m} \le K_\varepsilon\right\}}\right] + \varepsilon \\ &\le K_\varepsilon \mathbb{P}\{N_t \le D\} + A_I \sum_{k=D+1}^{\infty} k\alpha^{k-1} + \varepsilon. \end{aligned}$$

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

for some constant $A_I > 0$. Fixing ε and D, we find that

$$\limsup_{t\to\infty} \mathbb{E}\left[T_{N_t} L_{N_t}^{-m}\right] \leq (D+1)A_I \frac{\alpha^D}{1-\alpha} + \varepsilon$$

by the Monotone Convergence Theorem (Theorem 16.2, page 208 of [2]), and thus, letting $D \to \infty$ and $\varepsilon \to 0$, we have $\limsup_{t\to\infty} \mathbb{E}[T_{N_t} L_{N_t}^{-m}] = 0$ for $\rho > \rho^*$. □

COROLLARY 1. *For any $m > 1$, there exists a constant $\rho^* = \rho^*(\kappa, m) < 1$, such that for all $\rho \in (\rho^*, 1]$,*

$$\liminf_{T\to\infty} \frac{L(T)}{T^{1/m}} = \infty,$$

almost surely for any initial state $\mathbf{Q}(0)$ with $||\mathbf{Q}(0)|| = 1$.

PROOF. Note that for any initial state $\mathbf{Q}(0)$ with $||\mathbf{Q}(0)|| = 1$,

$$\liminf_{T\to\infty} \frac{L(T)}{T^{1/m}} \geq \liminf_{k\to\infty} \frac{\theta L_k}{T_{k+1}^{1/m}},$$

as can be seen from Proposition 1, and so it suffices to show that $\limsup_k T_{k+1} L_k^{-m} = 0$. But $T_{k+1} \leq T_k + C_{LT} L_k$, thus,

$$\limsup_{k\to\infty} T_{k+1} L_k^{-m} \leq \limsup_{k\to\infty} T_k L_k^{-m} + C_{LT} \limsup_{k\to\infty} L_k^{-m+1}. \tag{5.7}$$

The right-hand side is zero because, as we saw in the proof of Theorem 1, both $T_k L_k^{-m}$ and L_k^{-m+1} converge to zero almost surely for $\rho \in (\rho^*(\kappa, m), 1]$. □

5.2. *Instability of the original stochastic process.* In Theorem 1 we established that the fluid limit process in unstable, in the sense that $L(T) \to \infty$ as $T \to \infty$. We now proceed to show how the instability of the original stochastic process can be deduced from the instability of the fluid limit process. The original stochastic process is said to be unstable when $\{(U(t), X(t))\}_{t\geq 0}$ is transient, and $\|X(t)\| \to \infty$ almost surely for any initial state $X(0)$.

We will exploit similar arguments as developed in Meyn [27]. A notable distinction is that the result in [27] requires that a suitable Lyapunov function exhibits strict growth over time. In our setting the fluid limit is random, and the growth behavior as stated in Theorem 1 is not strict, but only in expectation and in an asymptotic sense, which necessitates a somewhat delicate extension of the arguments in [27].

The next theorem states the main result of the present paper, indicating that aggressive deactivation functions cause the network of Figure 3 to be unstable for load values ρ sufficiently close to 1.

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

THEOREM 2. *Consider the network of Figure 3, and suppose that* $f_i(x) \equiv 1$, $x \geq 1$, *and* $g_i(x) = o(x^{-\gamma})$, *with* $\gamma > 1$. *Let* $(\rho_1, \rho_2, \rho_3, \rho_4, \rho_5, \rho_6) = \rho(\kappa_1, \kappa_2, \kappa_3, \kappa_3 - \alpha, \kappa_6 - \alpha, \kappa_6)$, *with* $(\kappa_1 \vee \kappa_2) + \kappa_3 + \kappa_6 = 1$, *and* $0 < \alpha < \kappa_3 \wedge \kappa_6$. *Then there exists a constant* $\rho^\star(\kappa, \alpha) < 1$, *such that for all* $\rho \in (\rho^\star(\kappa, \alpha), 1]$:

$$\lim_{\|X(0)\| \to \infty} \mathbb{P}_{X(0)}\{\liminf_{t \to \infty} \|X(t))\| = \infty\} = 1.$$

Since our Markov Chain is irreducible, the above theorem immediately implies that it is transient. The proof of Theorem 2 relies on similar arguments as developed in the proof of Theorem 3.2 in [27]. The proof details are presented in Appendix E.

REMARK 3. Recall that the class of deactivation functions $g_i(x) = o(x^{-\gamma})$ includes the random-capture scheme with $g(x) \equiv 0$, $x \geq 1$, as considered in [12]. The result in Theorem 2 thus disproves the conjecture that the random-capture scheme is throughput-optimal in arbitrary topologies.

6. Simulation experiments. We now discuss the simulation experiments that we have conducted to support and illustrate the analytical results. Consider the broken-diamond network as depicted Figure 3 and considered in the previous sections. In the simulation experiments, the relative traffic intensities are assumed to be $\kappa_1 = \kappa_2 = 0.4$, $\kappa_3 = 0.4$, and $\kappa_6 = 0.2$ with $\alpha = 0$, for the components M_1, M_2, and M_3, respectively, with a normalized load of $\rho = 0.97$. At each node i, the initial queue size is $X_i(0) = 500$, the activation function is $f_i(x) \equiv 1$, $x \geq 1$, and the de-activation function is $g_i(x) = (1 + x)^{-\gamma}$, where we set $\gamma = 2$.

Figure 6 plots the evolution of the queue sizes at the various nodes over time, and shows that once a node starts transmitting, it will continue to do so until the queue lengths of all nodes in its component have largely been cleared. This characteristic, and the associated oscillations in the queues, strongly mirror the qualitative behavior displayed by the fluid limit.

Although Figure 6 suggests an upward trend in the overall queue lengths, the fluctuations make it hard to discern a clear picture. Figure 7 therefore plots the evolution of the node-average queue size over time, and reveals a distinct growth pattern. Evidently, it is difficult to make any conclusive statements concerning stability/instability based on simulation results alone. However, the saw-tooth type growth pattern in Figure 7 demonstrates strong signs of instability, and corroborates the qualitative growth behavior exhibited by the fluid limit. Indeed, careful inspection of the two figures confirms that the large increments in the node-average queue size occur immediately

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

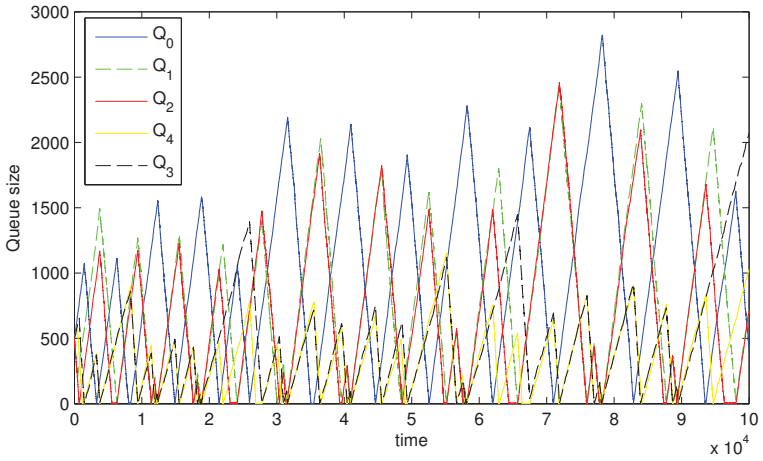


FIG 6. *Queue sizes at the various nodes as function of time for the network of Figure 3.*

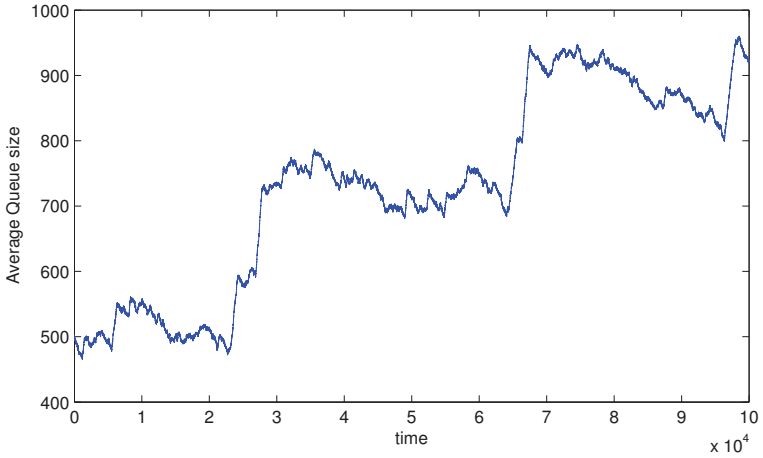


FIG 7. *Node-average queue size as function of time for the network of Figure 3.*

after M_4-periods, exactly as predicted by the fluid limit. We further observe that in between these periods, the node-average queue size tends to follow a slightly downward trend, consistent with the negative drift of rate $(\rho - 1)/3$ in the fluid limit.

7. Concluding remarks and extensions. We have used fluid limits to demonstrate the potential instability of queue-based random-access algorithms. For the sake of transparency, we focused on a specific six-node network and super-linear activity functions ($\gamma > 1$). Similar instability issues can however arise in a far broader class of interference graphs, as we

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

will discuss in Section 7.1 below. The proof arguments further suggest that instability can in fact occur for any activity factor that grows as a positive power ($\gamma \geq 1/k$) of the queue length for network sizes of order k, as will be described in Section 7.2 below.

In terms of backoff probabilities $\frac{1}{1+e^{w(X)}}$ used in [15], our results therefore suggest instability for weights $w(X)$ that grow faster than $\gamma \log X$, for any $\gamma > 0$ (as k can be chosen arbitrarily large). In other words, *the near-logarithmic growth condition* on the weights in [15] is a fundamental limit on the aggressiveness of nodes to ensure maximum stability (throughput optimality) in any general topology.

7.1. *Instability in general interference graphs.* The instability of random access, with aggressive de-activation functions, is not restricted to the broken-diamond network, and can arise in many other interference graphs. Consider a general interference graph $G = (V, E)$. Without loss of generality, we can assume G is connected, because otherwise we can consider each connected subgraph separately. For $\gamma > 1$, the fluid limit sample paths still exhibit the *sawtooth behavior*, i.e., when a node starts transmitting, it does not release the channel until its entire queue is cleared (on the fluid scale). Let $\mathcal{M} = \{M_1, \ldots, M_K\}$ denote the set of maximal independent sets (maximal schedules) of G. We say the network operates in M_i if a subset $W \subseteq M_i$ of nodes are served at full rate (on the fluid scale), and W does not belong to any other maximal schedules M_j, $j \neq i$. Under the random-access algorithm, at any point in time the network operates in one of the maximal schedules and switches to another maximal schedule when one or several of the queues in the current maximal schedule drain (on the fluid scale). More specifically, assume the network operates in a maximal schedule M_i. If M_i interferes with all other maximal schedules, i.e., $M_i \cap M_j = \emptyset$ for all $1 \leq j \leq K$, $j \neq i$, then a transition from M_i to any maximal schedule M_j, $j \neq i$, is possible when all the queues in M_i hit zero (on the fluid scale). On the other hand, if M_i overlaps with a subset of maximal schedules $\mathcal{M}'_i := \{M_j \in \mathcal{M} : M_i \cap M_j \neq \emptyset\}$, then the activity process can make a transition to $M_j \in \mathcal{M}'_i$ when all the queues in $M_i \backslash M_j$ drain (on the fluid scale).

The capacity region of the network is the convex set $\mathcal{C} = \text{conv}(S)$, which is full-dimensional because all the basis vectors of $\mathbb{R}^N$ belong to that set. The incidence vectors of the sets $\mathcal{M}$ correspond to the extreme points of $\mathcal{C}$ as they can not be expressed as convex combinations of other points. Consider a covering of $V = \{1, 2, \ldots, N\}$ using the maximal schedules. Formally, a set cover C of V is a collection of maximal schedules such that $V \subseteq \cup_{M_i \in C} M_i$. A set cover C is minimal if removal of any of the elements $M_i \in C$ leaves some

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

nodes of V uncovered. Consider the class of graphs in which $|C| \leq K-1$ for some minimal set cover C, i.e., we do not need all M_i's for covering V. Without loss of generality, let $\mathcal{M}^* = \{M_1, M_2, \ldots, M_{K^*}\}$ denote such a minimal cover with $K^* \leq K-1$. Consider a (strictly positive) vector of arrival rates $\lambda = \rho \sum_{i=1}^{K^*} \sigma_i 1_{M_i}$ where $\sigma_i > 0$, $1 \leq i \leq K^*$, such that $\sum_{i=1}^{K^*} \sigma_i = 1$, and $0 < \rho < 1$ is the load factor. Hence, a centralized algorithm can stabilize the network by scheduling each $M_i \in \mathcal{M}^*$ for at least a fraction $\rho\sigma_i$ of the time. However, under the random-access algorithm, the network might spend a non-vanishing fraction of time in the schedules $\mathcal{M}\backslash\mathcal{M}^*$, which can cause instability as ρ approaches 1. This phenomenon is easier to observe in graphs with a *unique* minimal set cover $\mathcal{M}^*$ and with a maximal schedule M_1 interfering with all the other maximal schedules, hence $M_1 \in \mathcal{M}^*$.

This means any valid covering of V must contain $\mathcal{M}^*$. Therefore, considering arrival rate vectors of the form $\lambda = \rho \sum_{i=1}^{K^*} \sigma_i 1_{M_i}$, $\sigma_i > 0$, $\sum_{i=1}^{K^*} \sigma_i = 1$, *the only way* to stabilize the network is to use M_i for a time fraction greater than $\rho\sigma_i$. Visits to M_1 have to occur infinitely often, otherwise the network is trivially unstable, and at the end of such visits, a transition to any other maximal schedule is possible, including the schedules in $\mathcal{M}\backslash\mathcal{M}^*$ with positive probability. Then, upon entrance to schedules in $\mathcal{M}\backslash\mathcal{M}^*$, the network spends a positive time in such schedules because the queues in $\mathcal{M}\backslash\{M_1\}$ build up during visits to M_1. Hence, the arguments in the instability proof of the broken-diamond network can be extended to such networks, although a rigorous proof of the fluid limits in such general cases remains a formidable task. Figure 8 shows a few examples of such unstable networks with unique minimal set covers.

7.2. *Instability for de-activation functions with polynomial decay.* Consider any unstable network $G = (V, E)$, for example the broken-diamond network or a graph as described in the previous subsection. Let $\mathcal{I}(i)$ denote the set of neighbors of node i in G. We construct a *k-duplicate graph* $G^{(k)}$, $k \in \mathbb{N}$, of G as follows. For each node $i \in V$, add k duplicate nodes $d_1^{(i)}, \ldots, d_k^{(i)}$ to the graph, with the same arrival rate λ_i and the same initial queue length $X_i(0)$, such that each node $d_j^{(i)}$ is connected to all the neighbors of node i and their duplicates, i.e., $\mathcal{I}(d_j^{(i)}) = \mathcal{I}(i) \cup_{l \in \mathcal{I}(i)} \{d_1^{(l)}, \ldots, d_k^{(l)}\}$, for all $1 \leq j \leq k$. For notational convenience, we define $D_i^{(k)} := \{i, d_1^{(i)}, \ldots, d_k^{(i)}\}$ and call it the duplicate collection of node i. Note that the duplicate graph has the same number of maximal schedules as the original graph. In fact, each maximal schedule $M_i^{(k)}$ of $G^{(k)}$ consists of nodes in the maximal schedule M_i of G and their duplicates, i.e., $M_i^{(k)} = \cup_{l \in M_i} D_l^{(k)}$. Next, we show

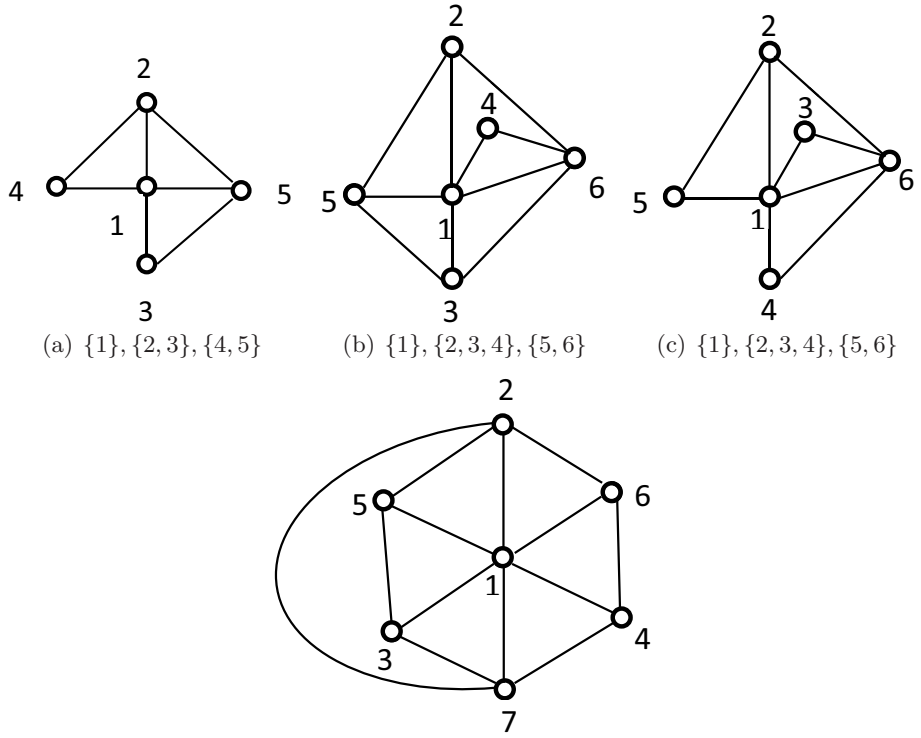


FIG 8. *A few unstable networks with their unique minimal cover $\mathcal{M}^*$ using the maximal schedules.*

that the duplicate graph is unstable for de-activation functions that decay as $o(x^{-\gamma})$, for $\gamma > 1/(k+1)$. Essentially, for such a range of γ, each duplicate collection acts as a super node with $\gamma > 1$, i.e., (i) if one of the queues in a duplicate collection $D_i^{(k)}$ starts growing, all the queues in $D_i^{(k)}$ grow linearly at the same rate λ_i (on the fluid scale), (ii) if a nonempty queue in $D_i^{(k)}$ starts draining, then all the queues in $D_i^{(k)}$ drain at full rate until they all hit zero (on the fluid scale). Then the instability follows from that of the original network, as we can simply regard the duplicate collections as super nodes. An informal proof of claims (i) and (ii) is presented below.

Claim (i) is easy to prove as all the queues in a duplicate collection share the same set of conflicting neighbors and the fact that one of the queues grows, over a small time interval, implies that some conflicting neighbors are transmitting over such interval. To show (ii), note that if one of the queues in the duplicate collection drains over a non-zero time interval, no matter how

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

small the interval is, all the conflicting neighbors must be in backoff for $O(R)$ units of time in the pre-limit process. This guarantees that all the queues in the duplicate collection will start a packet transmission during such interval almost surely. As long as the duplicate collection does not lose the channel, each queue of the collection follows the fluid limit trajectory of an M/M/1 queue. Suppose all the queues of the duplicate collection are above a level ϵ on the fluid scale for some fixed small $\epsilon > 0$. Thus, in the pre-limit process, the amount of time required for the queues to fall below a threshold ϵR is $O(R)$ with high probability as $R \to \infty$. The duplicate collection loses the channel if and only if all $k+1$ nodes in the collection are in backoff and a conflicting node acquires the channel by winning the competition between the backoff timers. The probability that a node goes into backoff at the end of a packet transmission is $O((\epsilon R)^{-\gamma})$, or approximately the fraction of time that a node spends in backoff is $O((\epsilon R)^{-\gamma})$. Therefore, the fraction of time that all $k+1$ nodes of the duplicate collection are simultaneously in backoff is $O((\epsilon R)^{-k\gamma})$ because the nodes in the duplicate collection act independently from each other. Therefore, over an interval of duration $O(R)$, the amount of time that all $k+1$ nodes are in backoff is $O(R^{1-(k+1)\gamma})$, which goes to zero as $R \to \infty$ if $\gamma > 1/(k+1)$. Thus, the nodes in the duplicate collection follow the fluid limits of an M/M/1 queue until their backlog is below ϵ on the fluid scale. Since ϵ could be made arbitrarily small, we can view the duplicate collection as a super node that does not release the channel until its backlog hits zero. This demonstrates the instability of fluid limits for the initial queue lengths described above for the duplicate network.

To rigorously prove instability of the original process using the framework of Meyn [27], we need to show instability of the fluid limit for any initial state. Handling arbitrary initial states for general activity functions and interference graphs is more involved than in the specific broken-diamond network considered here. An alternative option would be to extend the methodology and develop a proof apparatus where it suffices to show instability of the fluid limit for one particular initial state. The framework of Dai [6] offers the advantage that instability of the fluid limit only needs to be shown for an all-empty initial state. However the characterization of the fluid limit for an all-empty initial state appears to involve additional complications.

The above proof arguments suggest that instability can in fact occur for any $\gamma > 0$ as k can be chosen arbitrarily large. This indicates that the growth conditions in Ghaderi & Srikant [15] are sharp in the sense that backoff probabilities of the form $\frac{1}{O(\log(x))}$ are essentially the most aggressive de-activation functions that guarantee maximum stability of queue-based random access in arbitrary graphs.

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

APPENDIX A: FLUID LIMIT PROOFS – PART A

A.I. Prelimit model. We start with the time-homogeneous Markov process $(\mathbf{U}(t), \mathbf{X}(t))$, $t \geq 0$, with state space $\mathcal{S} = S \times \mathbb{N}_0^N$, where $N = 6$ and $S \subseteq \{0,1\}^N$ is the set of feasible activity states, which has already been fully described earlier in Section 2. We recap to state that service times are unit exponential as are backoff periods. In addition, the Poisson arrival processes are determined by the vector of arrival rates $\boldsymbol{\lambda}$ and the probability of backoff is determined as a function of queue length $1/(1+Q)^\gamma$ with $\gamma \in (1, \infty)$. As mentioned earlier, the case $\gamma = \infty$ corresponds to the random capture algorithm, considered in [12].

The fluid limit will not be obtained directly from the above process but rather via the *jump chain* of a *uniformized version* with 'clock ticks' from a Poisson clock with constant rate,

$$\beta \doteq \sum_{\ell=1}^{N} \lambda_\ell + N,$$

independent of state, with null (dummy) events introduced as needed.

With minor abuse of notation, denote by $(\mathbf{U}(n), \mathbf{X}(n)) \in \mathcal{S}$ the state of the jump chain at n-th clock tick. For our subsequent construction, it will be convenient to replace $\mathbf{U}(n)$ with the cumulative state $\mathbf{I}(n) = \sum_{k=0}^{n} \mathbf{U}(n) \in \mathbb{N}_0^N$, which is by definition nondecreasing. It determines and is determined by the sequence $\mathbf{U}(n)$ and the associated jump chain is Markov if the state is altered to be $(\mathbf{I}(n), \mathbf{I}(n-1), \mathbf{X}(n))$ with $\mathbf{I}(-1) = 0$. *Note that the process* $\mathbf{I}(n)$ *counts the number of steps where the queue process is active. It is not a count of the number of service completions by step* n.

From the jump chain, we obtain a continuous stochastic process in $C[0, \infty)$ by linear interpolation and by accelerating time by the factor β. To be specific, at an arbitrary intermediate time $t > 0$ between two clock ticks $t_l = (k-1)/\beta \leq t \leq t_u = k/\beta$, $k \in \mathbb{N}$, the interpolated process takes the values

$$\begin{aligned} \overline{\mathbf{Q}}(t) &\doteq \beta(t_u - t)\mathbf{X}(k-1) + \beta(t - t_l)\mathbf{X}(k), \\ \overline{\mathbf{I}}(t) &\doteq \beta(t_u - t)\mathbf{I}(k-1) + \beta(t - t_l)\mathbf{I}(k). \end{aligned}$$

From this construction we can obtain a sequence of such processes, indexed by $R \in \mathbb{N}$, with the usual fluid limit scaling

$$\left(\mathbf{Q}^R(t), \mathbf{I}^R(t)\right) \doteq \left(\frac{1}{R}\overline{\mathbf{Q}}^{(R)}(Rt), \frac{1}{R}\overline{\mathbf{I}}^{(R)}(Rt)\right). \tag{A.1}$$

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

This is obtained together with a corresponding sequence of initial queue lengths

$$\mathbf{Q}^R(0) = \frac{1}{R}\overline{\mathbf{Q}}^{(R)}(0) \to \mathbf{q}. \tag{A.2}$$

Without loss of generality we may take $||\mathbf{q}|| = 1$ where $||\cdot||$ denotes the $\mathcal{L}_1$ norm.

For every R and time $t \geq 0$, $(\mathbf{Q}^R(t), \mathbf{I}^R(t))$ take values in $E \doteq \mathbb{R}_+^N \times \mathbb{R}_+^N$, which is therefore the *state space* of the process. E has the usual Euclidean metric and associated topology and we will denote the Borel sets by $\mathcal{B}_E$. Furthermore the underlying jump chain $(\mathbf{U}(n), \mathbf{X}(n))_{n\geq 0}$ of the uniformized Markov process satisfies the 'skip-free' property [27], which ensures that the jumps between states are bounded in $\mathcal{L}_1$. It follows that the interpolated paths are Lipschitz continuous with Lipschitz constant $3\beta < \infty$. This property is conferred on the sample paths ω themselves as stated below

$$||\mathbf{Q}^R(t,\omega) - \mathbf{Q}^R(s,\omega)|| + ||\mathbf{I}^R(t,\omega) - \mathbf{I}^R(s,\omega)|| \leq 3\beta\,(t-s) \tag{A.3}$$

which holds $\forall \omega, 0 \leq s < t, R \in \mathbb{N}$. The factor 3 appears since at most two nodes can be active at the same time and at each clock tick at most one queue can be in(de)cremented.

To summarize, the scaled sequence of processes as defined in (A.1) take values in the space $C[0,\infty)$ of continuous paths taking values in E, endowed with the sup-norm topology, and σ-algebra $\mathcal{C}$ generated by the open sets. This is obtained through the usual metric ρ_C as defined in [40], page 6. This space is both separable and complete, see [40] Theorem 2.1, page 7. The probability measure induced on $\mathcal{C}$ by the R-th interpolated process (A.1) is denoted by μ_R so that $\mu_R(A)$ is the probability of an event $A \in \mathcal{C}$.

A.II. Fluid limit. If there is an infinite subsequence, $R_{k_1}, R_{k_2}, \ldots$ such that $\mu_{R_{k_n}} \Rightarrow \mu$, where $\Rightarrow$ denotes weak convergence, then μ is said to be a fluid limit measure. If such a fluid limit exists, then the corresponding process can be defined as follows. Its *state space* is again E with underlying sample space $C[0,\infty)$ and corresponding σ-algebra $\mathcal{C}$ generated by the open sets under the metric, ρ_C, as mentioned earlier. This is the same space as for the sequence of prelimit processes. With the fluid limit measure μ (including the deterministic initial conditions) we have an underlying probability space $(C[0,\infty), \mathcal{C}, \mu)$. The stochastic process, $(\mathbf{Q}, \mathbf{I})$ is the mapping $[0,\infty) \times C[0,\infty) \to E$, with values $(\mathbf{Q}(t,\omega), \mathbf{I}(t,\omega)) \in E$. The curves $(\mathbf{Q}(.,\omega), \mathbf{I}(.,\omega))$ and ω itself are the same. While these definitions are somewhat redundant, nevertheless in what follows, it will be convenient to think

of a sample path as either a point ω or as a random function. Finally, on some occasions, we will use the notation $X \in m\mathcal{C}$ to indicate that $X : C[0,\infty) \to \mathbb{R}$ is measurable.

The proof of the next theorem is standard and follows from Lipschitz continuity, and our assumptions on the initial conditions (see e.g., Theorem 8.3, page 56 of [1], and Lemma 3.1, page 19 of [40]). The details are omitted for brevity.

THEOREM 3. *The sequence of measures μ_R defined on $(C[0,\infty),\mathcal{C})$ is tight.*

Thus, it follows from Prohorov's Theorem (Theorem 6.1, page 37 of [1]) that the sequence μ_R is relatively compact and the fluid limit measure μ must exist. We suppose without loss of generality that $\mu_R \Rightarrow \mu$. The sample paths under μ have the same Lipschitz constant 3β. It follows that the sample paths of μ are differentiable a.e., almost surely [32], Corollary 12, page 109.

Lipschitz continuity also implies that there are only a countable number of closed intervals $[a,b]$, $0 \le a < b$, such that $Q_\ell(a,\omega) = Q_\ell(b,\omega) = 0$, $\ell = 1,\dots,N$, and $Q_\ell(x,\omega) > 0, \forall x \in (a,b)$, $\ell = 1,\dots,N$, holding almost surely.

We denote by $\{\mathcal{F}_t\}_{t\in[0,\infty)}$, $\mathcal{F}_t \subset \mathcal{C}$ the filtration of sub σ-algebras generated by the open sets restricted to the interval $[0,t]$. The process $(\mathbf{Q},\mathbf{I})$ is adapted to $\{\mathcal{F}_t\}_{t\in[0,\infty)}$. (In fact it is $\mathcal{F}_t$-progressive as the process is continuous, see [10], page 50 and Problem 1, page 89.)

By consideration of the weak law of large numbers and the existence of the fluid limit measure μ, it holds that

$$\mathbf{Q}(t) = \mathbf{Q}(0) + \boldsymbol{\lambda} t - \frac{1}{\beta}\mathbf{I}(t),\ t \ge 0. \tag{A.4}$$

This equation can be thought of as an accounting identity. If node ℓ is active for a unit interval, then I_ℓ increases by β, which corresponds (almost surely) to departures at unit rate. During the same period the arrival rate is λ_ℓ of course.

Since $I_\ell(t+h) - I_\ell(t) \le \beta h$ for any node ℓ, and any times $t \ge 0$ and $h > 0$, it follows from (A.4) that

$$Q_\ell(t+h) \ge Q_\ell(t) + \lambda_\ell h - h,\ \mu\ a.s.$$

We now derive an elementary property of the fluid limit process. Before doing so, since many of the events that we consider later are in terms of activity, we adapt the following notation throughout the paper. Given $t \ge 0$,

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

$h > 0$, define

$$J_{=}^{\ell}(t,h,\beta h) := \{\omega : I_\ell(t+h,\omega) - I_\ell(t,\omega) = \beta h\} \tag{A.5}$$

to be the event that node ℓ is being served (at maximum rate) during the interval $[t, t+h]$, i.e., the node is fully active during the given interval. Here the superscript ℓ denotes the node, t time and h duration. The term βh is the amount of activity which must be met with equality here, as indicated by the subscript "=". The subscript "=" may be replaced by $>$, $\geq$, $<$, or $\leq$, depending on the event.

LEMMA 3 (No Conflict Lemma). *Let $\ell_1 \neq \ell_2 \in \{1,\ldots,N\}$ be two neighbors in the interference graph G, and $h > 0$, $t \geq 0$, then*

$$\mu\left\{J_{=}^{\ell_1}(t,h,\beta h) \cap J_{=}^{\ell_2}(t,h,\beta h)\right\} = 0.$$

PROOF. This follows by definition, and the existence of the fluid limit. The event $J_{=}^{\ell_1}(t,h,\beta h) \cap J_{=}^{\ell_2}(t,h,\beta h)$ contradicts the inequality that for all $t \geq 0, h > 0$,

$$[I_{\ell_1}(t+h,\omega) - I_{\ell_1}(t,\omega)] + [I_{\ell_2}(t+h,\omega) - I_{\ell_2}(t,\omega)] \leq \beta h,$$

which holds μ almost surely. □

To obtain more detailed information with respect to the sample paths of μ, we proceed to the construction of sequences of stopping times.

A.III. Sequences of stopping times. We proceed to define stopping times which will be used subsequently to mark the end of M_i-periods and to state various properties of the fluid limit process. In this regard, it is natural to try to make these statements through stopping times which are returns of queues to 0 from being positive. However it appears difficult to do this for a given Q_ℓ in terms of a single sequence of stopping times. This is because such sequences can have limit points. As a brief illustration suppose Q_4 large and Q_3, Q_5 are both small and say that Q_3 and Q_4 are being drained. Anticipating that paths will be piecewise linear, it is readily seen that queue 3 will drain while queue 5 will be built up and then vice versa (alternating between M_1- and M_4-periods). Moreover both queues 3 and 5 will shrink so that bounded infinite sequences of such stopping times can occur. Therefore in what follows we work with *countably many* infinite sequences of stopping times, each one corresponding to the queue being *non-zero within given bounded intervals*.

The following definition is in connection with the amount of time a sample path for Q_ℓ is positive, immediately prior to a time $z > 0$.

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

DEFINITION 2. Given a time $z > 0$, and $v, 0 < v \leq z$, and an $\ell = 1, \ldots, N$, define

$$K^{\ell}_{z,v} \doteq \{\omega : Q_{\ell}(z-s,\omega) > 0,\ \forall s \in (0,v)\}\,.$$

In words, $K^{\ell}_{z,v}$ is the set of sample paths for Q_{ℓ} which are strictly positive in the interval $(z-v, z)$; if $z = 0$, $K^{\ell}_{z,v}$ is taken to be $\emptyset$.

Observe that it could be the case that either $Q_{\ell}(z,\omega) = 0$ or $Q_{\ell}(z-v,\omega) = 0$ (or both) and still $\omega \in K^{\ell}_{z,v}$. Finally note that it is possible for a given ω that no such v can be found, which requires that $Q_{\ell}(z,\omega) = 0$ on account of continuity. It can be shown that

$$K^{\ell}_{z,v} = \cap_{n:2/n<v}[\cup_{m=1}^{\infty}\{\omega : Q_{\ell}(z-q,\omega) \geq 1/m,\ q \in [1/n, v-1/n] \cap \mathbb{Q}\}] \in \mathcal{F}_z,$$

for $z > 0$, where $\mathbb{Q}$ is the set of rational numbers.

Given a time $z \geq 0$ and a path ω, we define the mapping $A^{\ell}(z,\omega) : C[0,\infty) \to [0,z]$ to be $A^{\ell}(z,\omega) \doteq \sup[\{v : \omega \in K^{\ell}_{z,v}\} \cup \{0\}]$, which is the amount of time for which Q_{ℓ} was positive immediately prior to z. By definition, if $z \geq u > 0$, then

$$\left\{\omega : A^{\ell}(z,\omega) \geq u\right\} = K^{\ell}_{z,u},$$

from which it follows that $A^{\ell}(z,\omega) \in \mathcal{F}_z$. So far z has been fixed. However $A^{\ell} : [0,\infty) \times C[0,\infty) \to \mathbb{R}_+$ is a stochastic process carried by the underlying probability space $(C[0,\infty), \mathcal{C}, \mu)$ and $\mathcal{F}_t$-adapted as we have just seen. This process is piecewise linear and left-continuous. (It falls to 0 immediately after Q_{ℓ} returns to 0 from being positive.) It follows that A^{ℓ} is $\mathcal{F}_t$-progressive, [10] Problem 1 page 89. We are now in a position to make the following definition.

DEFINITION 3. Given an $\mathcal{F}_t$ stopping time σ, a node $\ell \in 1, \ldots, N$ and $m \in \mathbb{Z}_0 \doteq \mathbb{Z} - \{0\}$, define $T_{\ell,m}(\omega,\sigma) : C[0,\infty) \times [0,\infty] \to [0,\infty]$ as follows

$$T_{\ell,m}(\omega,\sigma) \doteq \inf\left\{z \geq \sigma(\omega) : Q_{\ell}(z) = 0, A^{\ell}(z,\omega) \in (e_m, f_m]\right\} \leq \infty,$$

where

$$\begin{aligned} f_m &= \frac{1}{m},\ e_m = \frac{1}{m+1};\ \text{for } m \in \mathbb{Z}_0,\ m > 0, \\ f_m &= |m-1|,\ e_m = |m|;\ \text{for } m \in \mathbb{Z}_0,\ m < 0, \end{aligned}$$

where again empty sets have an infinite infimum.

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

In words, $T_{\ell,m}$ is the earliest right-hand end of an open interval, with value z, such that Q_ℓ is positive for a period $A^\ell(z,\omega) \in (e_m, f_m]$, immediately prior to $T_{\ell,m}$. If $z - f_m$ is the first time prior to z that $Q_\ell = 0$, then z is in the set on the RHS. However, if this occurs at $z - e_m$, this is not the case.

For certain choices of σ it will be shown in Lemma 5 that $T_{\ell,m}$ is also an $\mathcal{F}_t$ stopping time. We first state a construction lemma using a sequence of stopping times. These are returns to 0 following a fixed positive interval, in which we wait for a particular event A_k to occur.

LEMMA 4 (Stop and Look Back). *Let $\sigma \geq 0$ be an $\mathcal{F}_t$ stopping time and $a > 0$ a constant. Proposition 1.5, page 54 in Ethier & Kurtz [10] ensures that the following inductively defined sequence is a sequence of $\mathcal{F}_t$ stopping times: $s_0, s_1, s_2, \ldots,$*

$$\begin{aligned} s_0 &\doteq \sigma \\ s_k &\doteq \tau_c(\{0\}, s_{k-1} + a), \ k = 1, 2, \ldots \end{aligned}$$

Here, given an $\mathcal{F}_t$ stopping time $\sigma_1 > 0$, $\tau_c(\{0\}, \sigma_1) = \inf\{t \geq \sigma_1, Q(t,\omega) = 0\}$. Now let $A_k \in \mathcal{F}_{s_k}, k = 1, 2, \ldots$ be a sequence of events in the pre-T σ-algebras of the above stopping time sequence. Finally, define $\tau \doteq s_k$ if A_k occurs for the first time at step k and $\tau = \infty$ otherwise. Then τ is an $\mathcal{F}_t$ stopping time.

Note that we do not check to see if A_k has occurred if $s_k = \infty$ at any stage, as τ is assigned this value regardless.

We now proceed to show the following.

LEMMA 5. *Let $\sigma_0 \geq 0$ be an $\mathcal{F}_t$ stopping time such that $Q_\ell(\sigma_0(\omega), \omega) = 0$ or $\sigma_0 = \infty$ and suppose ℓ, m are given. Let $a = e_m$ and $\sigma \doteq a + \sigma_0$ which is therefore an $\mathcal{F}_t$ stopping time, and $T_{\ell,m}$ be the mapping given in Definition 3. Then $T_{\ell,m}(\omega, \sigma)$ is an $\mathcal{F}_t$ stopping time.*

PROOF. Given σ we will obtain a sequence of stopping times as in the first part of Lemma 4. However, as we have already discussed, A^ℓ is $\mathcal{F}_t$-progressive, from which it follows by Proposition 1.4, page 52 of [10], that

$$A_k := A^\ell(s_k(\omega), \omega) \in m\mathcal{F}_{s_k}, \forall \ k = 1, 2, 3, \ldots$$

Hence τ as defined in Lemma 4 is an $\mathcal{F}_t$ stopping time.

It remains to show that τ coincides with $T_{\ell,m}$ as defined. First suppose $\tau < \infty$, and immediately, $\tau \geq \sigma, Q_\ell(\tau, \omega) = 0, A^\ell(\tau, \omega) \in (e_m, f_m]$ by definition. The fact that there is no earlier time satisfying these conditions follows since

each s_k is a zero of Q_ℓ and the construction rules out that the event could have taken place any earlier. The case $\tau = \infty$ coincides with there being no zero satisfying the required conditions. □

We now make the following recursive definitions.

DEFINITION 4. Given $m \in \mathbb{Z}_0$ and node $\ell \in \{1, \ldots, N\}$, let τ_0 be the first entry of $Q_\ell(t, \omega)$ into 0 (τ_0 is an $\mathcal{F}_t$ stopping time). Then $Z^\ell_{m,0}$ is defined as

$$\begin{aligned} Z^\ell_{m,0} &\doteq T_{\ell,m}(\omega, 0);\ \text{if } Q_\ell(0,\omega) = 0 \\ Z^\ell_{m,0} &\doteq \tau_0;\ \text{if } Q_\ell(0,\omega) > 0,\ \tau_0 \in (e_m, f_m] \\ Z^\ell_{m,0} &\doteq T_{\ell,m}(\omega, \tau_0);\ \text{if } Q_\ell(0,\omega) > 0,\ \tau_0 \notin (e_m, f_m], \end{aligned}$$

and subsequent stopping times are defined as

$$Z^\ell_{m,n} \doteq T_{\ell,m}(\omega, Z^\ell_{m,n-1}),\ n = 1, 2, 3, \ldots.$$

The value of the stopping time is taken to be ∞ if the events do not occur. With obvious notation, we also define

$$A^\ell_{m,n}(\omega) \doteq A^\ell(Z^\ell_{m,n}, \omega)$$

to be the actual amount of time that the queue ℓ is positive prior to $Z^\ell_{m,n}$, and $A^\ell_{m,n} = \infty$ in case $Z^\ell_{m,n} = \infty$. Finally define the time at which Q_ℓ last enters $(0, \infty)$ prior to $Z^\ell_{m,n}$ to be

$$V^\ell_{m,n} \doteq Z^\ell_{m,n} - A^\ell_{m,n},$$

when $Z^\ell_{m,n} < \infty$ and $V^\ell_{m,n} = \infty$ otherwise. Note that $V^\ell_{m,n} \in m\mathcal{F}_{Z^\ell_{m,n}}$ and is thus a non-negative random variable but *not* a stopping time.

The following corollary follows immediately from Lemma 5 and Definition 4.

COROLLARY 2. *$Z^\ell_{m,n}$, $n \in \mathbb{N}_0$ is a strictly increasing sequence of $\mathcal{F}_t$ stopping times, $\forall \ell \in N, m \in \mathbb{Z}_0$.*

This completes our goal of constructing sequences of stopping times for the queue processes. For any node ℓ, by construction and by Lipschitz continuity, it follows that the set of stopping times, $Z^\ell_{m,n}$ determine all intervals where Q_ℓ is positive for any sample path almost surely.

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

For each $m \in \mathbb{Z}_0$, and $\ell \in \{1, \ldots, N\}$, let $\bar{Z}_m^\ell \doteq \sup_n\{Z_{m,n}^\ell : Z_{m,n}^\ell < \infty\}$ to be the supremum of the finite stopping times for positive intervals with duration in $(e_m, f_m]$. If there is an m such that $\bar{Z}_m^\ell = \infty$, then Q_ℓ returns to 0 infinitely often. Otherwise there is a $\bar{Z} > 0, \bar{Z} > \bar{Z}_m^\ell, \forall m \in \mathbb{Z}_0$. In this case, either Q_ℓ remains at 0 as $t \to \infty$, or Q_ℓ never returns to 0.

A.IV. Piecewise linearity and no backoff until empty. So far the backoff exponent $\gamma > 1$ has not been taken into consideration, but from now on it will be. The following lemma bounds the probability, for the jump chain, that node ℓ has a backoff before its queue gets 'small', provided that it was active earlier. Given some number $X_T \in \mathbb{N}$ define,

$$K_{X_T} \doteq \{\exists n : X_\ell(k) \geq X_T, 0 \leq k \leq n, U_\ell(n) = 0, U_\ell(n-1) = 1\}$$

to be the event that X_ℓ has remained above X_T and has had a backoff at step n. We then have the following lemma,

LEMMA 6 (No Early Backoff). *Given any $X_0 > X_T$,*

$$\mathbb{P}\{K_{X_T} | X_\ell(0) \geq X_0\} \leq \frac{1}{1-\lambda_\ell} \sum_{r=X_T}^{\infty} \frac{1}{(1+r)^\gamma} = \varepsilon_{X_T}. \tag{A.6}$$

Note that since the sum in the RHS of (A.6) is convergent, $\varepsilon_{X_T} \downarrow 0$ as $X_T \uparrow \infty$.

PROOF. It is convenient to consider the packets being served in generations. That is given a target packet, suppose that we serve the packets which arrive during its service with preemptive priority up to and including the target packet. This makes no difference to queue behavior as the service times are exponential and we are only interested in the *first occasion* when node ℓ goes into backoff.

Suppose there are $X_T > 0$ packets in queue at the time the service of a given target packet starts. Consider the busy period of this packet, i.e., the time to serve the target packet and the subsequent high-priority packets (without backoff). It is easy to see that the mean number of packet arrivals during this busy period is $\frac{1}{1-\lambda_\ell}$, including the target packet itself.

The service of each packet ends with a random decision to backoff with probability less than $\frac{1}{(1+X_T)^\gamma}$, since the queue length is never shorter than X_T until the target packet has departed. Thus, by the union bound, the probability of a backoff occurring before or immediately after the target packet departs, is less than $\frac{1}{(1+X_T)^\gamma}\frac{1}{1-\lambda_\ell}$. The probability of a backoff, starting with $X_0 > X_T$ packets, and before the queue drops below the level X_T,

is therefore smaller than $\frac{1}{1-\lambda_\ell}\sum_{r=X_T}^{r=X_0}\frac{1}{(1+r)^\gamma}$, which implies the statement of the lemma. □

The following lemma is a stepping stone to showing that any period of activity, no matter how small leads to a positive queue draining at full rate until it is empty. This gives rise to the 'sawtooth' trajectories which we shall formalize later in Theorem 4. First given times $t_2 > t_1 \geq 0$ on the fluid scale, let $B_\ell([t_1,t_2])$ be the event that node ℓ *starts a backoff* in the interval $[t_1,t_2]$. This event occurs in the prelimit process (Q_ℓ^R, I_ℓ^R) if for some jump chain index $n, U_\ell(n) = 1, U_\ell(n+1) = 0$ with $\lfloor Rt_1\beta\rfloor \leq n \leq \lceil Rt_2\beta\rceil$. Let $D_{\ell,\varsigma}([t_1,t_2])$ be the event that $Q_\ell^R(u) > \varsigma$ (or equivalently $\bar{Q}_\ell(Ru) \geq R\varsigma$) for all $u \in [t_1,t_2]$.

LEMMA 7. *Given the above definitions,*

$$\lim_{R\to\infty}\mu_R\left\{B_\ell([t_1,t_2])\cap D_{\ell,\varsigma}([t_1,t_2])\right\} = 0. \tag{A.7}$$

PROOF. First we may suppose node ℓ becomes active at some stage or there is nothing to prove. The lemma then follows from the union bound. Since there are at most $R_{t_1,t_2} \doteq \lceil R(t_2-t_1)\beta\rceil + 2$ departures in the entire interval, the union bound implies that the probability of a backoff is smaller than

$$\mu_R\left\{B_\ell([t_1,t_2])\cap D_{\ell,\varsigma}([t_1,t_2])\right\} \leq \frac{R_{t_1,t_2}}{(1+R\varsigma)^\gamma} \to 0.$$

This completes the proof. □

DEFINITION 5. Given a node ℓ, a time $t \in [0,\infty)$ on the fluid scale, and a queue length $Q_\ell(t) = Q > 0$, we say that t is a point of increase for the activity process of node ℓ if the event

$$P_{t,Q}^\ell \doteq \cap_{M=1}^\infty \left\{J_>^\ell(t,\frac{1}{M},0)\right\}\cap Q_{t,Q}^\ell$$

occurs, with $Q_{t,Q}^\ell \doteq \{\omega : Q_\ell(t,\omega) > Q\}$. In words, node ℓ is active in any arbitrarily small interval $(t, t+1/M)$ and Q_ℓ is greater than Q at time t.

Furthermore, given a time $s \in [0,\infty)$ and $h > 0$, we say that node ℓ is under active, with duration $h > 0$ if $J_<^\ell(s,h,\beta h)$ occurs.

Points of increase rule out that there is a sequence $t_n \downarrow t$ such that $I_\ell(t_n,\omega) = I_\ell(t,\omega)$, as there is activity no matter how small the interval. Under activity means that there was some idling during the interval. Given our choice of γ, it will be shown that a point of increase cannot be followed

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

by a period of under activity until queue ℓ has drained. This is because the probability of even a single backoff once service has begun, is effectively 0 until the queue has drained on the fluid scale.

LEMMA 8. *Suppose $s \in (t, t+Q/(1-\lambda_\ell))$. Then $\forall h, 0 < h < t+Q/(1-\lambda_\ell) - s$, and for all sufficiently large M,*

$$\mu\left\{J^\ell_>(t,1/M,0) \cap Q^\ell_{t,Q} \cap J^\ell_<(s,h,\beta h)\right\} = 0. \tag{A.8}$$

PROOF. Consider the sequence of prelimit processes. We will choose M large enough and ς small enough so that $[s, s+h] \subset (t+1/M, t+(Q-\varsigma)/(1-\lambda_\ell)]$ (for some small constant $\varsigma > 0$). Then for R, M sufficiently large, and then by definition, occurrence of $J^\ell_<(s,h,\beta h)$ for the R-th prelimit process, implies occurrence of $B_\ell([t+1/M, s+h])$. Hence we obtain,

$$\begin{aligned}
&\mu_R\left\{J^\ell_>(t,1/M,0) \cap Q^\ell_{t,Q} \cap J^\ell_<(s,h,\beta h)\right\} \leq \\
&\mu_R\left\{J^\ell_>(t,1/M,0) \cap Q^\ell_{t,Q} \cap B_\ell([t+1/M, s+h])\right\} \leq \mu_{B,R} + \mu_{F,D,R},
\end{aligned}$$

where

$$\mu_{B,R} \doteq \mu_R\left\{J^\ell_>(t,1/M,0) \cap B_\ell\left(t+1/M, s+h\right) \cap D_{\ell,\varsigma}([t, s+h])\right\},$$

and

$$\mu_{F,D,R} \doteq \mu_R\left\{Q^\ell_{t,Q} \cap \left(D_{\ell,\varsigma}([t, s+h])\right)^c\right\}.$$

Thus, in order to prove the lemma, it is sufficient to show that both $\mu_{B,R} \to 0$ and $\mu_{F,D,R} \to 0$, as $R \to \infty$, because then we may conclude that

$$\text{LHS of (A.8)} \leq \liminf \mu_R\left\{J^\ell_>(t,1/M,0) \cap Q^\ell_{t,Q} \cap J^\ell_<(s,h,\beta h)\right\} = 0$$

on applying Theorem 2.1, page 11 in [1] and since the sets $J^\ell_>(t,1/M,0)$, $Q^\ell_{t,Q}$, and $J^\ell_<(s,h,\beta h)$ are all open.

The fact that $\mu_{F,D,R} \to 0$ follows from (A.4) and then by definition of $Q^\ell_{t,Q}$ and additionally by the choice of s, h, ς. As far as $\mu_{B,R}$ is concerned, the event $J^\ell_>(t,1/M,0)$ implies that service has started during the interval $[t, t+1/M]$. On the other hand, the event $B_\ell([t+1/M, s+h])$ implies that at some time in $[t, s+h]$ node ℓ starts to backoff. Setting $t_1 = t$ and $t_2 = s+h$, we may invoke Lemma 7 as by definition the event $D_{\ell,\varsigma}([t, s+h])$ implies Q_ℓ did not go below ς in the interval $[t_1, t_2]$. It follows that $\mu_{B,R} \to 0$ as required. □

The implication of Lemma 8 is that any positive period of transmission, no matter how short, must be followed by full activity until the queue has drained on the fluid scale. This implies that there is no period of under activity, until the queue has drained, with probability 1.

A.V. Piecewise linear paths with probability 1. The aim of this section is to show that the queue sample paths follow a certain bilinear path during the interval prior to the queue becoming zero again. The bilinear path depends on the duration of the interval and on the arrival rate for the given queue.

To make the above statements precise, given $\ell \in \{1, \ldots, N\}$, define the bilinear path $\Phi^{\ell}_{t_0,t_1}$ for the interval $[t_0, t_1]$ to be

$$\Phi^{\ell}_{t_0,t_1}(s) = \begin{cases} \lambda_\ell\,(s - t_0)\,; & t_0 \le s \le s_0, \\ \lambda_\ell\,(s - t_0) - (1 - \lambda_\ell)(s - s_0); & s_0 \le s \le t_1, \end{cases}$$

where $s_0 \doteq t_1 - \lambda_\ell(t_1 - t_0)$. In words, Q_ℓ builds up linearly in the interval $[t_0, s_0]$ at rate λ_ℓ and drains at rate $1 - \lambda_\ell$ in the interval $[s_0, t_1]$.

Given $\eta > 0$, and $\ell \in \{1, \ldots, N\}$, define $\mathbb{1}^{(\eta,\ell)}_{t_0,t_1}$ to be the indicator for the event

$$\left\{\omega : \sup_{s\in[t_0,t_1]} |Q_\ell(s,\omega) - \Phi^{\ell}_{t_0,t_1}(s)| < \eta\right\} \in \mathcal{F}_{t_1}$$

In words, $\mathbb{1}^{(\eta,\ell)}_{t_0,t_1}(\omega) = 1$ iff the absolute difference between $\Phi^{\ell}_{t_0,t_1}$ and the sample path for Q_ℓ is smaller than η in sup-norm over the interval $[t_0, t_1]$.

We now examine the conditional probability that $\mathbb{1}^{(\eta,\ell)}_{V^{\ell}_{m,n},Z^{\ell}_{m,n}}(\omega) = 1$, given $Z^{\ell}_{m,n} < \infty$ and $A^{\ell}_{m,n}$ (the case $Z^{\ell}_{m,n} = \infty$ is irrelevant). Define, $\mathcal{Z}^{\ell}_{m,n} \doteq \sigma(Z^{\ell}_{m,n}, A^{\ell}_{m,n}) \subset \mathcal{C}$ and also $\mathcal{Z}^{\ell,\infty}_{m,n} = \mathcal{Z}^{\ell}_{m,n} \cap \{\omega : Z^{\ell}_{m,n}(\omega) < \infty\}$.

It will be enough to show that the sample paths lie in an arbitrarily small tube around $\Phi^{\ell}_{V^{\ell}_{m,n},Z^{\ell}_{m,n}}$ conditional on $A^{\ell}_{m,n}, Z^{\ell}_{m,n}$ lying in some small rectangle $Z^{(a,b)}_{(s,t)} \doteq \{\omega : A^{\ell}_{m,n}(\omega) \in (a, b], Z^{\ell}_{m,n}(\omega) \in (s, t]\} \in \mathcal{Z}^{\ell,\infty}_{m,n}$.

THEOREM 4. *Given $n \ge 1, m \in \mathbb{Z}_0$, then $\forall \eta > 0$,*

$$\mu\left\{\mathbb{1}^{(\eta,\ell)}_{V^{\ell}_{m,n},Z^{\ell}_{m,n}} = 1 | \mathcal{Z}^{\ell,\infty}_{m,n}\right\} = 1 \ a.s.$$

In words, given the stopping time $Z^{\ell}_{m,n}$ and the time $A^{\ell}_{m,n}$ prior to this that Q_ℓ was positive, $\Phi^{\ell}_{V^{\ell}_{m,n},Z^{\ell}_{m,n}}$ is followed, starting at $V^{\ell}_{m,n}$ and ending at $Z^{\ell}_{m,n}$, with probability 1, under the fluid limit measure μ.

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

PROOF. For any given $\varepsilon > 0$, the sets $Z^{(a,b)}_{(s,t)}$ $0 < s < t$, $0 < a < b$, $0 < t-s, b-a < \varepsilon$, are a π-system, see [42] Lemma 1.6, page 19 (i.e. closed under finite intersections), which generate $\mathcal{Z}^{\ell,\infty}_{m,n}$ and the entire space is a countable union of such sets. Hence by Theorem 10.3, page 163 of [2], we only need to show that

$$\mu\left\{\mathbb{1}^{(\eta,\ell)}_{V^\ell_{m,n},Z^\ell_{m,n}}(\omega) = 1; Z^{(a,b)}_{(s,t)}\right\} = \mu\left\{Z^{(a,b)}_{(s,t)}\right\}, \tag{A.9}$$

for suitably chosen ε given $\eta > 0$.

Let $B^\ell_{m,n}(\omega) \le A^\ell_{m,n}$ be the additional time, following strict entry of Q_ℓ into $(0,\infty)$ at $V^\ell_{m,n}$, until the first point of increase of I_ℓ is reached. $B^\ell_{m,n} \in m\mathcal{F}_{Z^\ell_{m,n}}$ as can be seen on consideration of its definition,

$$B^\ell_{m,n}(\omega) \doteq \inf\left\{u \in (0, A^\ell_{m,n}(\omega)) \cap \mathbb{Q} \ : I_\ell(V^\ell_{m,n}+u,\omega) - I_\ell(V^\ell_{m,n},\omega) > 0\right\}, \tag{A.10}$$

when $Z^\ell_{m,n} < \infty$. By definition of $B^\ell_{m,n}$, Lemma 8 and then (A.4), we may deduce that for $\omega \in Z^{(a,b)}_{(s,t)}$,

$$\begin{aligned} Q_\ell(V^\ell_{m,n}(\omega)+u,\omega) &= \lambda_\ell u,\ u \in [0, B^\ell_{m,n}(\omega)], \\ Q_\ell(V^\ell_{m,n}(\omega)+u,\omega) &= \lambda_\ell B^\ell_{m,n}(\omega) - (1-\lambda_\ell)(u - B^\ell_{m,n}(\omega)), \\ &\qquad u \in [B^\ell_{m,n}(\omega), \frac{B^\ell_{m,n}(\omega)}{1-\lambda_\ell}], \end{aligned}$$

μ almost surely. Moreover $B^\ell_{m,n}(\omega)$ must satisfy

$$t - s + b \ge \frac{B^\ell_{m,n}(\omega)}{1-\lambda_\ell} \ge s - t + a,\ \mu\ a.s.$$

in order to reach 0 in $[s,t]$. Therefore, given any $\eta > 0$, we may choose $\varepsilon_\eta > 0$ such that for all $v \in [s-b, t-a]$, $z \in [s,t]$ with $b-a, t-s < \varepsilon_\eta$

$$\sup_{u \in [v,z]} |Q_\ell(u,\omega) - \Phi^\ell_{v,z}(u)| < \eta,$$

μ almost surely, using Lipschitz continuity. Since $\omega \in Z^{(a,b)}_{(s,t)}$ implies $V^\ell_{m,n} \in [s-b, t-a]$, $Z^\ell_{m,n} \in [s,t]$, we obtain (A.9), for all such a, b, s, t as required. □

A similar result can be obtained when $n = 0$, where the possibility occurs that $Q_\ell(V^\ell_{m,0}) > 0$.

A.VI. Brief discussion of results. Theorem 4 applies to general networks and relies only on the assumption that $\gamma > 1$. The theorem implies that the sample paths are more or less determined given the sequences of

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

stopping times $Z^{\ell}_{m,n}$. Only in the case where the (finite) stopping times have a common upper bound is the process not completely defined, as otherwise the queue returns to 0 infinitely often, determining the path completely. If there is such a bound, either the queue remains at 0, or increases linearly, as there can be no subsequent point of increase of I_ℓ.

Indeed, since there are only countably many stopping times, and since for each finite $Z^{\ell}_{m,n} < \infty$ the queue sample paths follow $\Phi^{\ell}_{\cdot,\cdot,\cdot}$ for some finite interval with probability 1, we may confine sample path realizations to countable successions of such intervals. These either determine the entire sample path; or the queue remains at 0 following the final return; or as the final alternative, the queue remains zero for some interval and then increases linearly at rate λ_ℓ thereafter. We define the set of such sample paths by $P \subset C[0,\infty)$. The probability of any event $F \in \mathcal{C}$ can as well be taken as

$$\mu\{F\} = \mu\{F \cap P\},$$

and, therefore, we suppose that the probability space is defined on $(P, \mathcal{C}_P)$ with topology relativized in the usual way to P which is a subset of $\mathcal{C}[0,\infty)$. This establishes that the queue-length trajectory of each of the individual nodes exhibits *sawtooth* behavior in the fluid limit. This concludes Part A.

In Part B, we will show that we can in fact confine ourselves to a smaller set of paths which reflect the constraints resulting from the underlying interference graph.

APPENDIX B: FLUID LIMIT PROOFS – PART B

B.I. No idling property and zero delay capture. From Lemma 3, it follows that if queue 2 is draining, then queues 3, 4, 5, and 6 are increasing linearly. However, we also expect that queue 1 is either draining or remaining at 0, and this is indeed the case as we now show.

More generally, given a node ℓ, let $\mathcal{I}_\ell$ be the set of its interfering nodes, i.e., the set of its neighbors in the interference graph G. The following lemma shows that if $Q_\ell(s) > Q$, and all its interferers are idle in some interval $[s,t]$ then node ℓ is fully active until its queue drains.

LEMMA 9 (No Idling Property). *Given a node ℓ with interference set $\mathcal{I}_\ell$, an interval $[s,t]$, and a fixed $Q > 0$, define $h^{\ell}_{s,t,Q} \doteq Q/(1-\lambda_\ell) \wedge (t-s)$. Then,*

$$\mu\left\{\cap_{j\in\mathcal{I}_\ell} J^{j}_{=}(s,t-s,0) \cap J^{\ell}_{<}\left(s, h^{\ell}_{s,t,Q}, \beta h^{\ell}_{s,t,Q}\right) \cap Q^{\ell}_{s,Q}\right\} = 0.$$

Recall that:

- *$Q^{\ell}_{s,Q}$ is the event $\{\omega : Q_\ell(s,\omega) > Q\}$.*

- $\cap_{j\in\mathcal{I}_\ell} J^j_=(s,t-s,0)$ *is the event that there is no activity for any node in* $\mathcal{I}_\ell$ *during* $[s,t]$.
- $h^\ell_{s,t,Q}$ *is the time period over which the queue* ℓ *empties.*
- $J^\ell_<(s,h^\ell_{s,t,Q},\beta h^\ell_{s,t,Q})$ *is the events that node* ℓ *is under active in* $[s,s+h^\ell_{s,t,Q}]$.

PROOF. Given $n\in\mathbb{N}$ such that $n>1/(t-s)$, fix an arbitrary ζ, $0<\zeta<\frac{1}{2N}$. (Recall that N is the number of nodes in the network.) Clearly,

$$\cap_{j\in\mathcal{I}_\ell} J^j_=(s,t-s,0)\subset \tilde{D}^\ell_{\zeta,n}:=\cap_{j\in\mathcal{I}_\ell} J^j_<(s,1/n,\zeta/n).$$

Hence, for arbitrary $\epsilon_n>0$ depending on n, to be fixed later,

$$\cap_{j\in\mathcal{I}_\ell} J^j_=(s,t-s,0)\subseteq J^\ell_>(s,1/n,0)\cup\left(J^\ell_<(s,1/n,\epsilon_n)\cap\tilde{D}^\ell_{\zeta,n}\right).$$

Next, observe that for all $n_S\in\mathbb{N}$ sufficiently large,

$$J^\ell_<\left(s,h^\ell_{s,t,Q},\beta h^\ell_{s,t,Q}\right)=\cup_{n>n_S}G_n,$$

with $G_n\doteq J^\ell_<(s+2/n,s+h^\ell_{s,t,Q},\beta(h^\ell_{s,t,Q}-2/n))$. The union bound thus implies that

$$\begin{aligned}\mu\left\{D^\ell_{s,t}\cap J^\ell_<\left(s,h^\ell_{s,t,Q},\beta h^\ell_{s,t,Q}\right)\cap Q^\ell_{s,Q}\right\}&\le\sum_{n>n_S}\mu\left\{G_n\cap Q^\ell_{s,Q}\cap J^\ell_>(s,1/n,0)\right\}\\&\quad+\sum_{n>n_S}\mu\left\{\tilde{D}^\ell_{\zeta,n}\cap J^\ell_<(s,1/n,\epsilon_n)\cap Q^\ell_{s,Q}\right\}.\end{aligned}$$

Provided n_S is sufficiently large, each term in the first sum must be 0, else Lemma 8 is contradicted. To complete the proof, it is therefore sufficient to show that each of the terms in the second sum is 0 as well by suitable choice of ϵ_n. Given n, it is sufficient to find $\epsilon_n>0$ so that

$$\lim_{R\to\infty}\mu_R\left\{\tilde{D}^\ell_{\zeta,n}\cap J^\ell_<(s,1/n,\epsilon_n)\cap Q^\ell_{s,Q}\right\}=0,$$

because $\tilde{D}^\ell_{\zeta,n}$, $J^\ell_<(s,1/n,\epsilon)$, and $Q^\ell_{s,Q}$ are all open, so that Theorem 2.1 on page 11 of [1] implies that

$$\mu\left\{\tilde{D}^\ell_{\zeta,n}\cap J^\ell_<(s,1/n,\epsilon_n)\cap Q^\ell_{s,Q}\right\}=0.$$

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

The event $\tilde{D}^{\ell}_{\zeta,n}$ implies that there must have been at least

$$\frac{R\beta}{n}(1-N\zeta) > \frac{R\beta}{2n} \tag{B.1}$$

steps in the jump chain (if we allow for no overlap between active periods and since $|\mathcal{I}_\ell| < N$) at which all nodes in $\mathcal{I}_\ell$ are in backoff for the interval $[s, s+1/n]$. Also,

$$Q^R_\ell > Q - \frac{\beta}{n} > \varsigma > 0, \tag{B.2}$$

throughout $[s, s+1/n]$ since there can be at most $R\beta/n$ departures.

But if (B.1) occurs, we may suppose that node ℓ becomes active within $R\beta/(4n)$ such steps, as the probability converges to 1 as $R \to \infty$ that it does so. But if we take $0 < \epsilon_n < \beta/(4n)$, the implication is that there is a subsequent backoff. Since (B.2) also occurs, Lemma 7 with $t_1 = s, t_2 = s+1/n$ and ς above shows that the probability of a subsequent backoff goes to 0, which establishes the result. □

Since s, t, Q are arbitrary in Lemma 9, it follows from continuity that node ℓ begins service the instant its interferers become idle, if it has a positive queue length.

Lemmas 3 and 9 carry an implication for the node pairs $(1,2),(3,4),(5,6)$ in our network. We say that node ℓ_1 dominates node ℓ_2, $\ell_1 \neq \ell_2$, if $\mathcal{I}_{\ell_2} \subseteq \mathcal{I}_{\ell_1}$. Hence, if (say) node 3 (the dominant node) is draining, then no other node than 4 may be active as a consequence of Lemma 3. But this implies all interferers of node 4 are inactive. Hence, if $Q_4 > 0$, it will therefore begin to drain immediately, i.e., if node 3 is draining so is node 4. Also if Q_4 becomes 0 before Q_3, then it must remain at 0, until node 3 drains.

This result is formally stated in the following corollary, the proof of which is omitted for brevity.

Given any node $k \in \{1,\dots,N\}$, $Q_k \geq 0$, and time t, define

$$\Psi^k_{t,Q_k}(u) \doteq [Q_k - (u-t)(1-\lambda_k)] \vee 0, \ u \geq t, \tag{B.3}$$

and given $v > t$, let $F^k_{t,v,\eta}$ be the event that $|\Psi^k_{t,Q_k(t,\omega)}(u) - Q_k(u,\omega)| < \eta$ for $u \in [t,v]$.

COROLLARY 3. *Given a node ℓ, let k be any other node with $\mathcal{I}_k \subseteq \mathcal{I}_\ell$. $\forall t \geq 0, Q > 0, \eta > 0$, define $v = t + Q/(1-\lambda_\ell)$, then with $P^\ell_{t,Q}$ as in Definition 5, it holds that,*

$$\mu\left\{P^\ell_{t,Q} \cap \left(F^{(k)}_{t,v,\eta}\right)^c\right\} = 0.$$

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

Corollary 3 implies that μ almost surely the dominated node k follows Ψ^k the moment that dominating node ℓ becomes active.

In case the arrival rates satisfy $\lambda_1 = \lambda_2 = \lambda > 0$, $\lambda_4 = \lambda_5$, $\lambda_6 > \lambda_5$, and $\lambda_3 > \lambda_4$, Corollary 3 may be used to show that the network enters a *natural state* (as defined in Section 5) μ a.s. This result is proved in the following theorem.

THEOREM 5 (Almost Sure Natural State). *Given the initial condition* $\mathbf{Q}(0) = \mathbf{q}$ *with* $||\mathbf{q}|| = 1$, *there exists a* $T_N > 0$ *such that* μ *a.s. for all* $t \geq T_N$,

$$\begin{aligned} Q_3(t) &\geq Q_4(t), \\ Q_6(t) &\geq Q_5(t). \end{aligned}$$

Moreover (recalling the definition of ρ *given in Section 4),* $\exists \rho^* < 1$ *such that for all* $\rho \in [\rho^*, 1)$, $\vee_\ell Q_\ell(T_N) > 0$, *i.e., the network is non-empty at time* T_N.

PROOF. This result follows from Lipschitz continuity and more particularly from the fact that the sample paths are piecewise linear. Hence, apart from a set of measure 0, the derivatives of all queue lengths exist.

Consider now nodes 3 and 4. Where the derivatives exist and $Q_4 > 0$, it holds that

$$\frac{\mathrm{d}Q_3}{\mathrm{d}t} > \frac{\mathrm{d}Q_4}{\mathrm{d}t},$$

since $\lambda_3 > \lambda_4$ and since Q_4 is decreasing at linear rate whenever $Q_4 > 0$ and Q_3 is decreasing at a linear rate, as shown in Lemma 9. We may therefore deduce μ a.s. and where differentiability holds that

$$\frac{\mathrm{d}\left[Q_4(t) - Q_3(t)\right]_+}{\mathrm{d}t} \leq \lambda_4 - \lambda_3 < 0,$$

until some time T_3, such that $[Q_4(t) - Q_3(t)]_+ = 0,\ t \geq T_3$. The same holds for nodes 5 and 6, with corresponding time T_6 and the following inequalities are satisfied,

$$T_3 \leq \frac{[Q_4(0) - Q_3(0)]_+}{\lambda_3 - \lambda_4},\ T_3 \leq \frac{[Q_5(0) - Q_6(0)]_+}{\lambda_6 - \lambda_5}.$$

We may therefore take

$$T_N = T_3 \vee T_6,$$

and by taking worst case values in the above inequalities, we obtain a uniform bound on T_N. This concludes the first part of the lemma.

We now show that T_E, the time to empty, can be taken arbitrarily large. Define $L_P(t) \doteq (Q_1(t) \vee Q_2(t)) + Q_3(t) + Q_6(t)$. Then L_P can be reduced at most at rate 1, since service of nodes (1, 2), 3 and 6 is mutually exclusive, and grows at rate $\rho = \rho_0 + \rho_3 + \rho_6$, which can be made arbitrarily close to

1. Hence $T_E \to \infty$ as $\rho \uparrow 1$ if $L_P(0) > 0$. It can be the case that $L_P(0) = 0$ but then $Q_4(0) + Q_5(0) = 1$, so that $L_P(1/2) = \rho_0/2$ and $T_E \geq \frac{1}{2}(1 + \frac{\rho_0}{1-\rho})$ and again $T_E \to \infty$ as $\rho \uparrow 1$. □

This shows that a non-empty natural state can be reached in finite time, because of the dominance property. Given Theorem 5 we can and will suppose that the state is natural at time 0, without loss of generality.

We define the set of paths which additionally satisfy the constraints of Lemmas 3 and 9 to be $P_L \subset P \subset C[0,\infty)$. As previously, we now restrict the set of sample paths to P_L, so that the probability of an event $F \in \mathcal{C}$ can be determined as $\mu\{F\} = \mu\{F \cap P_L\}$. This concludes Part B.I.

B.II. Discussion. We now give a largely informal description of the paths in P_L. Section 4.1 gives a detailed description of the M_k-periods, $k = 1, 2, 3, 4$. The ends of M_1-, M_2-, and M_3-periods are marked by the corresponding stopping times $Z^{1,2}_{m,n}, Z^3_{m,n}, Z^6_{m,n}$. For M_4-periods, the following construction is needed. (It is needed because $Z^4_{m,n}$ stopping times may be part of an M_2-period, and hence do not mark the end of an M_4-period.)

We first define $P^\ell_{m,n} = V^\ell_{m,n} + B^\ell_{m,n} \in m\mathcal{F}_{Z^\ell_{m,n}}$ to be the time prior to $Z^\ell_{m,n}$ when service begins (recall the definition of $B^\ell_{m,n}$ in (A.10)).

DEFINITION 6. A stopping time $Z^4_{m,n}$ is a $M_4^{(5)}$ stopping time, denoted by $Z^{4,M_4^{(5)}}_{m,n}$ if the following holds:

$$Q_5(Z^4_{m,n} - P^4_{m,n}) \geq Q_4(Z^4_{m,n} - P^4_{m,n}), \tag{B.4}$$

$$I_\ell(Z^4_{m,n} - P^4_{m,n}) = I_\ell(Z^4_{m,n}),\ \ell = 3, 6. \tag{B.5}$$

That $Z^{4,M_4^{(5)}}_{m,n}$ is an $\mathcal{F}_t$ stopping time follows as both the above events lie in $\mathcal{F}_{Z^4_{m,n}}$.

This is consistent with an M_4-period taking place in which queue 4 emptied first (or at the same time as queue 5) by (B.4). If this is a strict inequality then we say this is a strict $M_4^{(5)}$ stopping time. (B.5) ensures that node 5 is being served throughout $[P^4_{m,n}, Z^4_{m,n}]$ as a consequence of Lemma 9. Similarly we may define $Z^{5,M_5^{(4)}}_{m,n}$.

APPENDIX C: FLUID LIMIT PROOFS – PART C

In this part, we will derive the probabilities according to which one period M_i is followed by another M_j with no switching delay (on the fluid scale). We first concentrate on the case of switching out of M_1.

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

We begin with some preliminary results. The first is for measures constructed from closed continuity sets. Given a set of sample paths G, define the improper probability measures

$$\mu_G\{F\} \doteq \mu\{F \cap G\}, \quad \mu_G^{(R)}(F) = \mu_R\{F \cap G\}.$$

The following lemma shows that weak convergence is conferred on $\mu_G^{(R)}$ provided G is a closed μ-continuity set.

LEMMA 10. *Suppose $\mu^{(R)}$ is a sequence of probability measures on a metric space $(\Omega, \mathcal{F})$, such that $\mu^{(R)} \Rightarrow \mu$, where μ is also a probability measure on the same space. Let $G \in \mathcal{F}$ be a closed μ-continuity set. Then it holds that $\mu_G^{(R)} \Rightarrow \mu_G$.*

In particular, the weak convergence definitions (iii), (iv), and (v), in Theorem 2.1, page 11 of [1], all equivalently hold.

Suppose a pair of non-interfering queues in the network are operating in isolation, e.g. queues (1,2). Then each queue will be empty and in fact will then subsequently be empty infinitely often, almost surely. Given that the evolutions of the two queues are independent, it is plausible that the total number of steps in the jump chain for which both queues are backed off together increases to infinity in a period which is negligible on the fluid scale. We will formalize this in Lemma 11 below.

Given a start time taken to be 0, define $W^R(u)$ to be the total number of steps that nodes 1 and 2 are both in backoff, starting at time 0 and ending at time $u > 0$ on the fluid scale, in $(\mathbf{Q}^R(t), \mathbf{I}^R(t))$. Partial periods between one clock tick and the next, at the start and at the end are neglected. The following lemma supposes nodes 1 and 2 are in isolation, so that no other nodes may gain the medium.

LEMMA 11 (Total Backoff). *Given $Q > 0$, define $t \doteq Q/(1-\lambda)$, and suppose that $Q_\ell^R(0) \le Q$, $\ell = 1, 2$, and both nodes are active at time 0. Then for any $Q, \xi > 0$,*

$$\lim_{R\to\infty} \mu_R\left\{W^R(t+2\xi) \ge 2\sqrt{R}\right\} = 1.$$

PROOF. Let $\tau_{0,0}$ be the stopping index in the jump chain for the first occurrence of

$$X_1(\tau_{0,0}) = X_2(\tau_{0,0}) = 0. \tag{C.1}$$

Given any $\xi > 0$, define $p_{\xi,Q}^R \doteq \mathbb{P}\{\tau_{0,0} \le \lfloor \beta(t+\xi)R \rfloor | X_\ell(0) \le RQ, \ell = 1, 2\}$. It will be enough to show that $p_{\xi,Q}^R \to 1$ as $R \to \infty$. To see this, note that any

queue in isolation is positive recurrent, as a consequence of Lemma 6. Thus, the jump chain restricted to nodes 1 and 2 in isolation (i.e., with remaining queues barred from gaining the medium) is also positive recurrent. Let m_0 be the mean number of steps between indices k such that (C.1) is again satisfied. Also let K_ξ^R be the random number of such steps in the next interval of $\lfloor \beta\xi R \rfloor$ steps. It is easily seen from the weak law of large numbers that

$$\lim_{R\to\infty} \mu_R \left\{ K_\xi^R > \frac{\lfloor \beta\xi R \rfloor}{2m_0} \right\} = 1,$$

which implies the statement of the lemma.

Thus, to complete the proof, we just need to show that $p_{\xi,Q}^R \to 1$. Fix $\varepsilon_{X_T} > 0$ and choose $X_T := X_T(\lambda,\gamma) < \infty$ as in (A.6) so that the probability of even a single backoff before either queue reaches X_T is no more than ε_{X_T}. Moreover let $\tau_{T,\ell}, \ell = 1,2$ be the stopping indices for $X_\ell(\tau_{T,\ell}) = X_T$. Then, given any $\eta > 0$, and $\varepsilon_{R,\eta} > 0$, it can be seen that $\tau_{T,1} \vee \tau_{T,2} \le \beta R(t+\eta)$ occurs with probability larger than $1 - 2\varepsilon_{R,\eta} - 2\varepsilon_{X_T}$, with $\varepsilon_{R,\eta} \to 0$ as $R \to \infty$ by the weak law of large numbers.

Next, given any $\varepsilon_L > 0$, there exists a X_L large enough such that $\mathbb{P}\{X_\ell(\tau_{T,\ell}+k) \le X_L\} > 1-\varepsilon_L$ for all $k \in \mathbb{N}$. This follows from the fact that the jump chain in isolation is positive recurrent, and thus the corresponding sequence of infinite probability vectors is tight as they are converging to the steady-state distribution, by the converse to Prohorov's Theorem, [1], page 37. Hence, with probability larger than $1 - 2\varepsilon_{R,\eta} - 2\varepsilon_{X_T} - 2\varepsilon_L$, $X_\ell(\lfloor \beta R(t+\eta) \rfloor) \le X_L$, $\ell = 1,2$.

Moreover, again by the positive recurrence of the isolated jump chain, the mean number of steps for queues 1 and 2 both to become 0, starting from any state with $X_\ell \le X_L$, $\ell = 1,2$, is bounded by some constant $m_L := m_L(X_L) < \infty$. Thus, by Markov's inequality, except with a probability less than $m_L/(\eta R)$, in a further ηR steps both queues will become 0 (and thus inactive).

Finally, given any $\epsilon > 0$, choose X_T and X_L large enough so that $\varepsilon_{X_T} < \epsilon/8$ and $\varepsilon_L < \epsilon/8$ and then R sufficiently large so that $\varepsilon_{R,\eta} < \epsilon/8$ and $m_L/(\eta R) < \epsilon/8$. Hence, with probability larger than $1-\epsilon$, $\tau_{0,0} < (t+2\eta)R$ for all R sufficiently large. Since ϵ and η are arbitrary, the proof is complete. □

C.I. Transition from an M_1-period. In what follows we will further suppose that the queue lengths at nodes 1 and 2 and their activity are both equal, as the following arguments are readily modified where this is not the case. We therefore denote their common queue length as $Q(u) = Q_1(u) = Q_2(u)$ in what follows, and similarly for the activity $I(u) = I_1(u) = I_2(u)$.

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

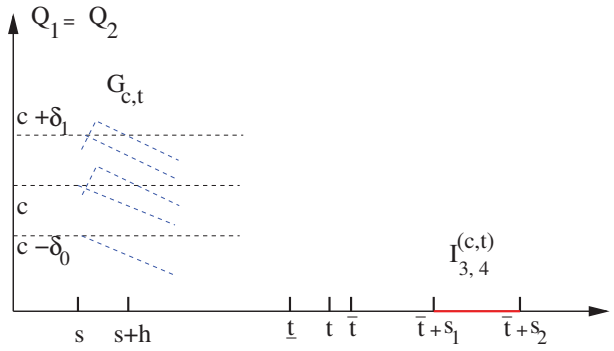


FIG 9. *Sample paths for the sets* $G_{c,t}$ *and* $I_{c,t}^{(3,4)}$.

Finally, in the following t, c and hence s are fixed, and

$$s \doteq t - \tfrac{c}{1-\lambda}, \quad \delta_k \doteq \alpha_k c,\ 0 < \alpha_k < 1,\ k = 0, 1,$$
$$h \doteq \nu c, \qquad \zeta \doteq \chi c,\ \nu > \chi > 0,$$

for some small positive constants α_k, ν, and χ to be determined later. We are now ready to define the following closed set of paths,

$$\begin{aligned} G_{c,t} \doteq &\{\omega : 0 < c - \delta_0 \le Q(s,\omega) \le c + \delta_1\} \\ &\cap \{\omega : I(s+h,\omega) - I(s,\omega) \ge \beta(h - \zeta)\}. \end{aligned} \tag{C.2}$$

Note that $G_{c,t}$ is constructed to correspond to an M_1-period.

Now given $0 < s_1 < s_2$, and $\overline{t}$ (which will be specified later), define

$$I_{c,t}^{(3,4)} \doteq J_{=}^{3}(\overline{t} + s_1, s_2 - s_1, \beta(s_2 - s_1)) \cap J_{=}^{4}(\overline{t} + s_1, s_2 - s_1, \beta(s_2 - s_1)).$$

Note that $I_{c,t}^{(3,4)}$ is a (closed) set of paths for which node 3 (and also node 4) are fully active during the interval $[\overline{t} + s_1, \overline{t} + s_2]$. Similar definitions, using the same s_1, s_2, and $\overline{t}$, can be made for $I_{c,t}^{(4,5)}, I_{c,t}^{(5,6)}$.

The first set of paths, $G_{c,t}$, is illustrated in the dashed lines in Figure 9.

Note all sample paths must pass through the interval $[c - \delta_0, c + \delta_1]$ at time s, but may continue to increase for a brief period at the beginning. After $s + h$ the two queues must be draining at rate $1 - \lambda$ almost surely, as shown in Lemma 8. The red interval to the right indicates periods where one of the other three node pairs are expected to have the medium during the interval $[\overline{t} + s_1, \overline{t} + s_2]$. Only one such pair will be active during this period as a result of the forthcoming construction.

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

The following is the earliest time that queues 1 and 2 can drain if the sample paths are constrained to lie in $G_{c,t}$,

$$\underline{t} \doteq t - \frac{\alpha_0}{1-\lambda}c.$$

As far as additional queue build-up is concerned, under the fluid limit,

$$Q(s+h,\omega) \leq Q(s,\omega) + \lambda h = Q(s,\omega) + \lambda\nu c$$

holds for sample paths in $G_{c,t}$ (see (A.4)). It then follows that queues 1 and 2 will reach 0 under the fluid limit no later than

$$\overline{t} \doteq t + \frac{\alpha_1 + \lambda\nu}{1-\lambda}c,$$

which is the definition for $\overline{t}$. We thus conclude that, under the fluid limit, queues 1 and 2 will reach 0 in the interval $(\underline{t},\overline{t})$ (for the first time after $s+h$ on occurrence of the event $G_{c,t}$). We formalize the above in the following lemma.

LEMMA 12 (Queue Bounds). *Let* $\tau^0_{c,s} \doteq \tau_c(s,\{0\}) = \inf\{t \geq s : Q(t) = 0\}$ *be the first contact time with 0 for* $Q = Q_1 = Q_2$. *Then,*

$$\mu\left\{G_{c,t} \cap \left\{\omega : \tau^0_{c,s}(\omega) \notin [\underline{t},\overline{t}]\right\}\right\} = 0.$$

Additionally, $\forall \ell = 3,4,5,6$,

$$\mu\left\{G_{c,t} \cap \{\omega : Q_\ell(\underline{t},\omega) < \Delta t\lambda_\ell\}\right\} = 0,$$

where

$$\Delta t \doteq \underline{t} - (s+h) = c(1-\alpha_0-\nu(1-\lambda))/(1-\lambda).$$

PROOF. By definition of $G_{c,t}$, $Q(s,\omega) \geq c - \delta_0, \forall \omega \in G_{c,t}$. It follows that Q cannot reach 0 before $\underline{t}$, as sample paths by definition lie in P_L (see Part A.VI, following Theorem 4). A similar argument applies to $\overline{t}$.

For the last part, Lemma 3 shows that nodes 3, 4, 5, and 6 must be idle in the period $[s+h,\underline{t}]$. Since the sample paths are restricted to lie in P_L, it follows that their queues must satisfy the stated inequality at time $\underline{t}$. The proof is complete. □

The time for node ℓ to reach 0 following $\underline{t}$ is therefore at least

$$f_\ell \doteq \Delta t\frac{\lambda_\ell}{1-\lambda_\ell},\ \ell = 3,4,5,6.$$

Clearly $f_\ell \to (c\lambda_\ell)/((1-\lambda)(1-\lambda_\ell))$ as $\alpha_0,\nu \downarrow 0$, and so this expression is bounded from below as $\alpha_0,\alpha_1,\nu > \chi$ are made arbitrarily small. For future

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

use, we define

$$\underline{f} \doteq \Delta t \wedge_{\ell=3}^{6} \lambda_\ell/(1-\lambda_\ell),$$

as a lower bound on the time needed to drain any node $\ell = 3, 4, 5, 6$.

Our results thus far do not rule out the possibility that there is an idle period during which nodes 3, 4, 5, or 6 fail to obtain the medium. In order to make allowance for this, we introduce a period ξc, $\xi > 0$, which comes following queues 1 and 2 draining, and to be definite, we set $\xi c = \underline{f}/8$. Hence, if it is the case that

$$\overline{t} - \underline{t} < \underline{f}/4 \tag{C.3}$$

and that service of node ℓ cannot start before $\underline{t} - \xi c$ and must have started no later than $\overline{t} + \xi c$, then it follows that service will continue throughout the interval $[\overline{t} + \xi c, \overline{t} + \xi c + \underline{f}/2]$. In this case, we may take $s_1 = \underline{f}/4, s_2 = \underline{f}/2$ again to be definite. Further, set $s_3 = \xi c + \underline{f}/2$. To summarize, if (C.3) holds, on occurrence of $G_{c,t}$ and that service of queues 3 and 4 commences in the interval $[\underline{t} - \xi c, \overline{t} + \xi c]$, then the event $I_{c,t}^{(3,4)}$ must take place. The same is true in case service commences for either queue pair $(4, 5)$ or $(5, 6)$ in $[\underline{t} - \xi c, \overline{t} + \xi c]$.

Let $\hat{C}_k, k = 3, 4, 5, 6$, be the residual backoff time for nodes 3, 4, 5, or 6, at time $s + h$, with $\hat{C}_1 = \hat{C}_2 = 0$ as these nodes will be almost surely active. Define S_M to be the number of steps in the jump chain before one of these nodes gains the medium and also define

$$\begin{aligned} W^{(3,4)} &\doteq \left\{\hat{C}_3 < \wedge_{k=4}^{6} \hat{C}_k\right\} \cup \left(\left\{\hat{C}_4 < \hat{C}_3 \wedge \hat{C}_5 \wedge \hat{C}_6\right\} \cap \left\{\hat{C}_3 < \hat{C}_5\right\}\right), \\ C_{c,t,R}^{(3,4)} &\doteq W^{(3,4)} \cap \left\{S_M \leq \sqrt{R}\right\}. \end{aligned}$$

$W^{(3,4)}$ is the event that nodes 3 and 4 win the backoff competition to take the medium first from nodes 1 and 2. Similar definitions can be made for nodes $(4, 5)$ and for nodes $(5, 6)$ in addition. The probabilities of these events are

$$\mathbb{P}\{W^{(3,4)}\} = \frac{3}{8} = \mathbb{P}\{W^{(5,6)}\},\ \mathbb{P}\{W^{(4,5)}\} = \frac{1}{4}, \tag{C.4}$$

as the backoff periods are unit mean i.i.d. exponential random variables. $C_{c,t,R}^{(3,4)}$ is the event that nodes $(3, 4)$ win the backoff competition, and that they do so in no more than $\sqrt{R}$ of the jump chain steps when nodes 1 and 2 are in backoff together.

Next let

$$B_R^{(1,2)} \doteq N_R^{(1,2)}(s + h, \underline{t} - \xi c) \cap \{W^R(s + h, \overline{t} + \xi c) \geq 2\sqrt{R}\}$$

be the intersection of the event $N_R^{(1,2)}(s+h, \underline{t} - \xi c)$ that neither node 1 nor node 2 starts to backoff during the time interval $[s+h, \underline{t} - \xi c]$ and the event $\{W^R(s+h, \overline{t} + \xi c) \geq 2\sqrt{R}\}$ that nodes 1 and 2 operating in isolation would be simultaneously in backoff for a cumulative period of time of at least $2\sqrt{R}$ during the interval $[s+h, \overline{t} + \xi c]$. Informally speaking, the event $B_R^{(1,2)}$ ensures that there is sufficient backoff by nodes 1 and 2 and that they do not begin to backoff while there are a significant number of packets remaining at node 1 or 2.

Next define c_Q to be,

$$c_Q \doteq \frac{s_3 - s_2}{2}(1 - \lambda_3) > 0,$$

which is at least half the content of queues 3 and 4 on the fluid scale at time $\overline{t} + s_2$, given our construction. Further, define the following event

$$Q_R^{(3,4)}(\underline{t}, \overline{t} + s_2) \doteq \left\{\omega : \inf\left\{Q_m^R(u, \omega), u \in [\underline{t}, \overline{t} + s_2]\right\} > c_Q, m = 3, 4\right\} \in \mathcal{F}_{\overline{t}+s_2},$$

for which we obtain the following corollary.

COROLLARY 4.

$$\lim_{R\to\infty} \mu_R \left\{G_{c,t} \cap \left(Q_R^{(3,4)}\right)^c\right\} = 0.$$

PROOF. Lemma 12 implies that for all n sufficiently large,

$$\limsup_{R\to\infty} \mu_R \left\{G_{c,t} \cap \left\{\omega : Q_\ell^R(\underline{t}, \omega) \leq \Delta t \lambda_\ell - 1/n\right\}\right\} = 0, \ \ell = 3, 4,$$

on using Theorem 2.1, page 11 in [1] and the fact that both the above sets are closed. Hence we need only show that,

$$\lim_{R\to\infty} \mu_R \left\{\left\{\omega : Q_\ell^R(\underline{t}, \omega) > \Delta t \lambda_\ell - 1/n, \ell = 3, 4\right\} \cap \left(Q_R^{(3,4)}\right)^c\right\} = 0, \tag{C.5}$$

for sufficiently large n. However (C.5) follows from the weak law of large numbers, from the definition of Δt, c_Q, and the event $Q_R^{(3,4)}$. □

Finally, define $N_R^{(3,4)}(\underline{t}, \overline{t} + s_2)$ to be the event that neither node 3 nor node 4 has a backoff during the time interval $[\underline{t}, \overline{t} + s_2]$ (on the fluid scale). Clearly, equivalent definitions for this and the above corollary can be made for node pairs $(4, 5), (5, 6)$.

In what follows it will be convenient to write $G := G_{c,t}$. Our aim now is to show that no matter what trajectory the fluid limit path followed earlier, if it lies in G so that nodes 1 and 2 almost surely reach 0 in the interval $[\underline{t}, \overline{t}]$, marking the end of an M_1-period, then the probability of the next period depends only on the residual backoff times, which is a Markov property.

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

LEMMA 13. *Suppose that G is a set of paths as defined in (C.2), with parameter values so that (C.3) holds, and is also a μ-continuity set. In addition, let $F \in \mathcal{F}_s$ be an arbitrary closed, finite-dimensional set of paths defined by times s and earlier. It then holds that*

$$\begin{aligned}\mu_G\left\{F \cap I_{c,t}^{(3,4)}\right\} &\geq \frac{3}{8}\mu_G\{F^o\},\\ \mu_G\left\{F \cap I_{c,t}^{(5,6)}\right\} &\geq \frac{3}{8}\mu_G\{F^o\},\\ \mu_G\left\{F \cap I_{c,t}^{(4,5)}\right\} &\geq \frac{1}{4}\mu_G\{F^o\}.\end{aligned}$$

In case F is a μ-continuity set, the interior can be dropped and $\geq$ replaced with equality.

PROOF. We first show the last part of the lemma, assuming the first part to be true. If F is a μ-continuity set, then by definition, $0 = \mu\{\partial F\} \geq \mu\{G \cap \partial F\}$, and it follows that F is a μ_G-continuity set as well. Since the factors sum to 1 and the events on the left are almost surely exclusive as a consequence of Lemma 3, we can now replace the inequality sign with equality.

We move to the first part of the lemma, which we will prove for nodes 3 and 4. The proof for the other node pairs is similar.

First observe that

$$C_{c,t,R}^{(3,4)} \cap B_R^{(1,2)} \cap N_R^{(3,4)}(\underline{t}, \overline{t} + s_2) \subseteq I_{c,t}^{(3,4)},$$

since $C_{c,t,R}^{(3,4)} \cap B_R^{(1,2)}$ implies that nodes 3 and 4 activate before time $\overline{t} + s_1$, while $N_R^{(3,4)}(\underline{t}, \overline{t} + s_2)$ ensures that neither node 3 nor node 4 has a backoff during the time interval $[\underline{t}, \overline{t} + s_2]$. We thus obtain the following chain of inequalities

$$\begin{aligned}\mu_G^{(R)}\left\{F \cap I_{c,t}^{(3,4)}\right\} &\geq \mu_G^{(R)}\left\{F \cap C_{c,t,R}^{(3,4)} \cap B_R^{(1,2)} \cap N_R^{(3,4)}\right\}\\ &\geq \mu_G^{(R)}\left\{F \cap W^{(3,4)}\right\}\\ &\quad - \mu_G^{(R)}\left\{\left(\left\{S_M \leq \sqrt{R}\right\} \cap B_R^{(1,2)} \cap N_R^{(3,4)}\right)^c\right\}\\ &\geq \frac{3}{8}\mu_G^{(R)}\{F\} - \mu_G^{(R)}\left\{S_M > \sqrt{R}\right\}\\ &\quad - \mu_G^{(R)}\left\{\left(B_R^{(1,2)}\right)^c\right\} - \mu_G^{(R)}\left\{\left(N_R^{(3,4)}\right)^c\right\},\end{aligned} \tag{C.6}$$

with $N_R^{(3,4)} \equiv N_R^{(3,4)}(\underline{t}, \overline{t} + s_2)$ for compactness. The first line follows by inclusion, the second using $\mu_G\{A \cap B\} \geq \mu_G\{A\} - \mu_G\{B^c\}$, and the third from (C.4) by independence of the residual backoff times and by using the union bound in conjunction with de Morgan's laws. We now proceed to show that

$$\begin{aligned}
\mu_{G,1}^{(R)} &\doteq \mu_R\left\{S_M > \sqrt{R}\right\} \to 0,\\
\mu_{G,2}^{(R)} &\doteq \mu_R\left\{\left(B_R^{(1,2)}\right)^c \cap G_{c,t}\right\} \to 0,\\
\mu_{G,3}^{(R)} &\doteq \mu_R\left\{\left(N_R^{(3,4)}\right)^c \cap G_{c,t}\right\} \to 0.
\end{aligned}$$

The first limit is immediate.

In order to deal with the second limit, define the event

$$Q_R^{(1,2)}(s+h, \underline{t} - \xi c) \doteq \left\{\omega : \inf\left\{Q_m^R(u,\omega), u \in [s+h, \underline{t} - \xi c]\right\} > \varsigma, m = 1, 2\right\}$$

for some small constant $\varsigma > 0$, and use the upper bound

$$\begin{aligned}
\mu_{G,2}^{(R)} \leq\ & \mu_R\left\{\left(B_R^{(1,2)}\right)^c \cap Q_R^{(1,2)}(s+h, \underline{t} - \xi c) \cap G_{c,t}\right\}\\
&+\mu_R\left\{\left(Q_R^{(1,2)}(s+h, \underline{t} - \xi c)\right)^c \cap G_{c,t}\right\}.
\end{aligned}$$

The limit of the second term is 0 by definition of $\underline{t}$ as the earliest time that queues 1 and 2 can drain under the event $G_{c,t}$ and on making a suitable choice for ς. It suffices then to show that the limit of the first term is 0. In order to prove this, we invoke the definition of the event $B_R^{(1,2)}$ to obtain that the first term is bounded from above by 'A'+'B', where

$$\begin{aligned}
\text{'A'} &= \mu_R\left\{\left(N_R^{(1,2)}(s+h, \underline{t} - \xi c)\right)^c \cap Q_R^{(1,2)}(s+h, \underline{t} - \xi c)\right\},\\
\text{'B'} &= \mu_R\left\{\{W^R(s+h, \overline{t} + \xi c) \leq 2\sqrt{R}\} \cap G_{c,t}\right\}.
\end{aligned}$$

The term 'A' converges to 0 by definition of the events and Lemma 7. Lemma 11 shows that the limit of the term 'B' (i.e., the event there is insufficient backoff by nodes 1 and 2 on occurrence of $G_{c,t}$) is 0.

In order to handle the third limit, we apply the upper bound

$$\mu_{G,3}^{(R)} \leq \mu_R\left\{\left(N_R^{(3,4)}(\underline{t}, \overline{t} + s_2)\right)^c \cap Q_R^{(3,4)}\right\} + \mu_R\left\{\left(Q_R^{(3,4)}\right)^c \cap G_{c,t}\right\}.$$

Lemma 7 shows that the limit of the first term is 0, while the statement of Corollary 4 is that the limit of the second term is 0.

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

Taking limits in (C.6) with respect to R, and using Lemma 10, it follows that

$$\begin{aligned}\mu_G\left\{F \cap I_{c,t}^{(3,4)}\right\} &\geq \frac{3}{8}\limsup_R \mu_G^{(R)}\{F\} \geq \frac{3}{8}\liminf \mu_G^{(R)}\{F^o\} \\ &\geq \frac{3}{8}\mu_G\{F^o\},\end{aligned}$$

where the first inequality follows from the fact that F and $I_{c,t}^{(3,4)}$ are both closed and the third since F^o is open and again from Lemma 10. □

Let μ be the fluid limit measure and proceed to define for any given $t \geq 0$ the following class of sets, the *finite-dimensional continuity rectangles* $\mathcal{K}_{\mu,t}$ which are a subset of the *finite-dimensional sets* $\mathcal{H}_t$.

Definition 7. Define the class of finite closed rectangles $\mathcal{R}$ to be the sets

$$\left(\prod_{j=1}^{N}[q_{j,L}, q_{j,H}]\right)\left(\prod_{j=1}^{N}[r_{j,L}, r_{j,H}]\right) \subset \mathbb{R}_+^N \times \mathbb{R}_+^N,$$

where $q_{j,L} \leq q_{j,H}, r_{j,L} \leq r_{j,H}$, otherwise we obtain the empty set.

Given times $0 \leq t_1 < t_2 < \cdots < t_J \leq t$, define $\pi_{J,t} : C[0,\infty) \to E^J$ to be the (continuous) projection map taking the sample path to its position at times $t_1, \ldots, t_J$,

$$\pi_{K,t}(\omega) = \Big((\mathbf{Q}(t_1,\omega), \mathbf{I}(t_1,\omega)), \ldots, (\mathbf{Q}(t_J,\omega), \mathbf{I}(t_J,\omega)) \Big).$$

Finally, take $\mathcal{R}^J$ to be J-products of closed rectangles. Define $\mathcal{K}_t$ to be sets of the form $\pi_{K,t}^{-1}R_J, R_J \in \mathcal{R}^J$ and finally $\mathcal{K}_{\mu,t} \subset \mathcal{K}_t$ to be those $H \in \mathcal{K}_t$ such that $\mu\{\partial H\} = 0$. Clearly $\mathcal{K}_{\mu,t} \subset \mathcal{K}_t \subset \mathcal{F}_t$.

Returning to Lemma 13, we see that it is satisfied by all sets $F \in \mathcal{K}_{\mu,s}$ with equality since they are by definition closed μ-continuity sets. Furthermore, since the terms on the left and on the right are measures and since $\mathcal{K}_{\mu,s}$ generates $\mathcal{F}_s$, the following corollary holds.

Corollary 5. *$\forall F \in \mathcal{F}_s$, Lemma 13 holds with equality, i.e.,*

$$\begin{aligned}\mu_G\left\{F \cap I_{c,t}^{(3,4)}\right\} &= \frac{3}{8}\mu_G\{F\}, \\ \mu_G\left\{F \cap I_{c,t}^{(5,6)}\right\} &= \frac{3}{8}\mu_G\{F\}, \\ \mu_G\left\{F \cap I_{c,t}^{(4,5)}\right\} &= \frac{1}{4}\mu_G\{F\}.\end{aligned}$$

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

PROOF. First note that the measures on the LHS and RHS are both finite and the entire space lies in a countable union of such sets. Thus both measures are σ-finite, with respect to the sets in $\mathcal{K}_{\mu,s}$. It is readily shown that $\mathcal{K}_{\mu,s}$ is a π-system and $\sigma(\mathcal{K}_{\mu,s}) = \mathcal{F}_s$. Theorem 10.3, page 163 in [2] thus shows that LHS and RHS agree on $\mathcal{F}_s$. □

To continue towards Theorem 6, we now define paths that one of which is followed immediately on completion of a (positive) M_1-period at time s, μ a.s. First define

$$\vartheta_{s,t}^{(\mathbf{Q},M_k)}(u),\ u \in [s,t], k = 2,3,4,$$

to be the path which is at $\mathbf{Q}$ at time s and then follows M_k until time t, e.g., if $k = 1$, queues 1 and 2 are decreasing linearly at rate $(1-\lambda)$ and any other queue $\ell = 3,4,5,6$ is increasing at rate λ_ℓ. Precise definitions we omit as the form of the sample paths have already been discussed. The next definition is for an indicator function that the above path is being followed in an interval $[s, s+h], h > 0$.

$$\mathbb{1}_{M_k,\mathbf{Q}}^{(s,h,\eta)} \doteq \mathbb{1}\left\{\omega : ||\mathbf{Q}(v,\omega) - \vartheta_{s,s+h}^{(\mathbf{Q}(s,\omega),M_k)}(v)|| < \eta, v \in [s,s+h]\right\}. \tag{C.7}$$

In words, M_k is 'followed' for an interval of duration h starting at s to a closeness η.

Note that the result of Corollary 5 applies only to events in some σ-algebra $\mathcal{F}_w$ where $w \geq 0$ is fixed. However, we require that equivalent results be established for all events $F \in \mathcal{F}_{Z_{m,n}^{(1,2)}}$. This issue can be approached as follows.

Given $s < t, a < b$, with $n \in \mathbb{N}_0$ and recalling Definition 4, let

$$C^{(n)} \doteq \left\{\omega : (Z_{m,n}^{(1,2)}, A_{m,n}^{(1,2)}) \in [s,t) \times [a,b)\right\},$$

then it is readily seen that

$$F \cap C^{(n)} \in \mathcal{F}_t,\ \forall F \in \mathcal{F}_{Z_{m,n}^{(1,2)}}$$

Suppose that $s < t$ in the definition of $C^{(n)}$ satisfy $t - s < \lambda_1 e_m$, and $e_m < a < b \leq f_m$ (e_m and f_m are given in Definition 4). Then, it can be seen that for all paths $\omega \in F \cap C^{(n)}$, for any $F \in \mathcal{F}_{Z_{m,n}^{(1,2)}}$, we can find a $w < s$ such that $I_\ell(s,\omega) - I_\ell(w,\omega) = \beta(s-w)$, $\ell = 1,2$, i.e., the queues and activity components constitute a set of parallel lines over the interval $[w,s]$. This suggests that any event in the σ-algebra $\mathcal{F}_{Z_{m,n}^{(1,2)}} \cap C^{(n)}$ is also an event in $\mathcal{F}_w$.

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

The above intuitive argument can be formalized by establishing the existence of an equivalent σ-algebra. We say that the σ-algebra $\mathcal{F}_{Z_{m,n}^{(1,2)}} \cap C^{(n)}$ is *equivalent* to a sub σ-algebra, $\mathcal{H}_w \subset \mathcal{F}_w$, if to each event $H \in \mathcal{F}_{Z_{m,n}^{(1,2)}} \cap C^{(n)}$ there is an event $H_w \in \mathcal{H}_w$ so that $H = H_w$.

LEMMA 14 (Equivalent σ-algebra). *Given $n \in \mathbb{N}_0$, arbitrary $t > s \geq 0$ such that $t - s < \lambda_1 e_m$, and $e_m < a < b \leq f_m$ and arbitrary $w \in (t - \lambda_1 e_m, s)$, there is a σ-algebra, $\mathcal{H}_w \subset \mathcal{F}_w$ equivalent to $\mathcal{F}_{Z_{m,n}^{(1,2)}} \cap C^{(n)}$.*

We omit the proof. Let $t_Z = Z_{m,n}^{(1,2)}$, $h^{(1,2)} = \lambda_1 e_m \times \wedge_{\ell=3}^{6} \lambda_\ell/(1-\lambda_\ell)$, and $\mathbf{Q}_Z = \mathbf{Q}(Z_{m,n}^{(1,2)})$. Next define $\mathbb{1}_{M_k,m,n}^{(\eta,M_1)}$ to be $\mathbb{1}_{M_k,\mathbf{Q}_Z}^{(t_Z,h^{(1,2)},\eta)}$, $k = 2,3,4$ if $Z_{m,n}^{(1,2)} < \infty$. Define $\mathcal{F}_{Z_{m,n}^{(1,2)}}^{\infty} \doteq \mathcal{F}_{Z_{m,n}^{(1,2)}} \cap \{Z_{m,n}^{(1,2)} < \infty\}$ as we are only interested in finite stopping times.

THEOREM 6. *$\forall n \in \mathbb{N}_0, m \in \mathbb{Z}_0$, $\exists \eta_m$ such that $\forall \eta$, $\eta_m > \eta > 0$,*

$$\begin{aligned} \mu\left\{\mathbb{1}_{M_2,m,n}^{(M_1,\eta)} | \mathcal{F}_{Z_{m,n}^{(1,2)}}^{\infty}\right\} &= \frac{3}{8}, \ \mu \ a.s., \\ \mu\left\{\mathbb{1}_{M_3,m,n}^{(M_1,\eta)} | \mathcal{F}_{Z_{m,n}^{(1,2)}}^{\infty}\right\} &= \frac{3}{8}, \\ \mu\left\{\mathbb{1}_{M_4,m,n}^{(M_1,\eta)} | \mathcal{F}_{Z_{m,n}^{(1,2)}}^{\infty}\right\} &= \frac{1}{4}. \end{aligned} \tag{C.8}$$

Since $\eta > 0$ can be taken arbitrarily small, the conclusion is that one of the M_2-, M_3-, or M_4-periods start immediately at $Z_{m,n}^{(1,2)}$ on occurrence of $Z_{m,n}^{(1,2)} < \infty$ and with probabilities determined solely by the residual backoff times.

PROOF. Given $F \in \mathcal{F}_{Z_{m,n}^{(1,2)}}^{\infty}$, we may write $F = \cup_k F_k$ as a countable union of disjoint sets. F_k is obtained by intersection of F with the disjoint sets,

$$C_k \doteq \left\{\omega : (Z_{m,n}^{(1,2)}, A_{m,n}^{(1,2)}) \in (s_k, t_k] \times (a_k, b_k]\right\},$$

where $e_m \leq a_k < b_k \leq f_m$ and $e_m \leq s_k < t_k$ are chosen according to η in a way to be described subsequently. $C_k \cap C_m = \emptyset$, $m \neq k$ is constructed by first choosing the intervals for the stopping time $Z_{m,n}^{(1,2)}$ to be disjoint and then likewise the durations into disjoint semi-open intervals. Thus, $F_k \doteq F \cap C_k \in \mathcal{F}_{t_k}$.

We turn to F_k and will suppose that $t_k - s_k$ is sufficiently small, so that we may find a time $w_k \in (t_k - e_m\lambda_1, s_k)$ as in Lemma 14. w_k will be a constant

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

determined by a_k, b_k, s_k, t_k and η only. For the moment suppose that w_k and η are used to determine constants c_η, t_{w_k} and then a set $G_{c_\eta,t_{w_k}}$, satisfying the conditions of Lemma 13, such that in addition

$$C_k \subset G_{c_\eta,t_{w_k}}, \tag{C.9}$$

with the s in the definition of $G_{c,t}$, see (C.2), taken to be w_k. It can then be seen that the following chain of equalities hold,

$$\begin{aligned}
\mu\left\{F_k \cap I^{(3,4)}_{c_\eta,t_{w_k}}\right\} &= \mu_{G_{c_\eta,t_{w_k}}}\left\{F_k \cap I^{(3,4)}_{c_\eta,t_{w_k}}\right\} \\
&= \mu_{G_{c_\eta,t_{w_k}}}\left\{F_k^{(w_k)} \cap I^{(3,4)}_{c_\eta,t_{w_k}}\right\} \\
&= \frac{3}{8}\mu_{G_{c_\eta,t_{w_k}}}\left\{F_k^{(w_k)}\right\} \\
&= \frac{3}{8}\mu_{G_{c_\eta,t_{w_k}}}\left\{F_k\right\} \\
&= \frac{3}{8}\mu\left\{F_k\right\}
\end{aligned} \tag{C.10}$$

The first equality follows from (C.9), the second from Lemma 14 as there exists a $F_k^{(w_k)} \in \mathcal{F}_{w_k}$ such that $F_k^{(w_k)} = F_k$, the third from Corollary 5 and by definition of $G_{c_\eta,t_{w_k}}$, the fourth equality is again from Lemma 14, and the final one follows again from (C.9). Corresponding results follow for $I^{(4,5)}_{c_\eta,t_{w_k}}$ and $I^{(5,6)}_{c_\eta,t_{w_k}}$. Once one of these events has occurred, μ almost surely the queues corresponding to the active period proceed to empty because they lie in P_L and therefore P. Moreover, η_m is determined depending on the duration (at least $\lambda_1 e_m$) of the M_1-period. η_m is taken sufficiently small, so that if we take any $\eta, \eta_m > \eta > 0$, only the node pair (and corresponding M_k-period) can satisfy the constraints in (C.7), for the interval $[Z^{(1,2)}_{m,n}, Z^{(1,2)}_{m,n} + h^{(1,2)}]$.

The above steps may be taken provided that i) $G_{c_\eta,t_{w_k}}$ is a closed μ-continuity set, ii) $G_{c_\eta,t_{w_k}}$ contains C_k and hence F_k iii) $\overline{t} + s_3 - w_k$ is sufficiently small, so that the paths $\vartheta^{(\mathbf{Q},M_k)}_{\cdot,\cdot}$ satisfy the constraints as in (C.8), and iv) the condition (C.3) must hold so that the conditions of Lemma 13 and also of Corollary 5 are met.

To show that c_η, t_{w_k} and a corresponding $G_{c_\eta,t_{w_k}}$ exist, given $\eta > 0$, set $t_k - s_k = A_1\eta$ and $c_\eta = A_2\eta$, where A_1 and A_2 will be fixed later. Next fix the time $t_{w_k} = (s_k + t_k)/2$. w_k is now determined using $t_{w_k} - w_k = (1-\lambda)^{-1}c_\eta$. A brief calculation shows that α_0, α_1 must be chosen so that

$$\alpha_0, \alpha_1 > \frac{A_1(1-\lambda)}{2A_2}, \tag{C.11}$$

in order that condition ii) above is met.

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

As far as i) is concerned, $G_{c,t}$ is an intersection of two sets, a queue constraint and an activity constraint, so that it is enough to obtain each as a μ-continuity set. With respect to the queue constraint set, at time w_k, there are uncountably many choices for α_0, α_1, which may be taken as close as we like to the constraint in (C.11) given c_η. For the activity set, we may choose ν arbitrarily small and having fixed it, then there are uncountably many choices for $\chi > 0$ which we may also take arbitrarily small. Thus the activity set can also be chosen to be a μ-continuity set. Putting the above together, G may be constructed as a μ-continuity set for given A_1, A_2, η and so that $C_k \subset G$.

We turn to condition iv), where it can be checked that it is satisfied provided that

$$\frac{\alpha_0 + \alpha_1 + \lambda\nu}{1-\lambda} < \frac{1}{4}\left(\frac{1-\alpha_0-\nu(1-\lambda)}{1-\lambda} \wedge_{\ell=3}^{6} \frac{\lambda_\ell}{1-\lambda_\ell}\right), \tag{C.12}$$

which obviously holds by making α_0, α_1, and ν sufficiently small, and then choosing A_1/A_2 sufficiently small according to (C.11).

As far as iii) is concerned, an examination of the construction preceding Lemma 13 shows that $\overline{t} + s_3 - w_k \propto c_\eta = A_2\eta$. We may thus proceed by taking $A_2 > 0$ sufficiently small to ensure that iii) is met and then choose $A_1 > 0$ and $\nu > 0$ sufficiently small so as to meet (C.11) and (C.12), fixing c_η, ν. The rest follows on choice of α_0, α_1, χ.

Note that common choices may be made for A_1, A_2, ν for each C_k and once these are fixed, common values may be chosen for α_0, α_1, χ as there are uncountably many possibilities and only a countable number of choices can have positive probability for any C_k. We have thus shown that a suitable $G_{c,t}$ can be found for each k, given $\eta > 0$.

The rest of the proof follows on summing (C.10) over k, to obtain

$$\begin{aligned} \sum_k \mu\left\{\mathbb{1}_{M_2,\mathbf{Q}_Z}^{(t_Z,h_Z,\eta)}; F_k\right\} &= \frac{3}{8}\sum_k \mu\{F_k\} \\ \mu\left\{\mathbb{1}_{M_2,\mathbf{Q}_Z}^{(t_Z,h_Z,\eta)}; F\right\} &= \frac{3}{8}\mu\{F\}, \end{aligned} \tag{C.13}$$

and similarly for M_3 and M_4. This is the required result as F is arbitrary. □

C.II. Switchover from M_2, M_3, M_4. Here we will only state our results, moreover M_2- and M_3-periods are analogous and so we will only deal with the former. To state our theorem for switching out of a M_2-period, define $\mathbb{1}_{M_k,m,n}^{(M_2,\eta)}, k = 1, 3$ as was done for switching out of M_1. Also define $\mathcal{F}_{Z^3_{m,n}}^{\infty} \doteq \mathcal{F}_{Z^3_{m,n}} \cap \{Z^3_{m,n} < \infty\}$.

THEOREM 7. *$\forall n \in \mathbb{N}_0, m \in \mathbb{Z} - \{0\}$, $\exists \overline{p}, \overline{q} > 0, \overline{p} + \overline{q} = 1$ and $\exists \eta_m$ such that $\forall \eta,\ \eta_m > \eta > 0$*

$$\begin{aligned} \mu\left\{ \mathbb{1}^{(M_2,\eta)}_{M_1,m,n} | \mathcal{F}^{\infty}_{Z^3_{m,n}} \right\} &= \overline{p},\ \mu\ a.s., \\ \mu\left\{ \mathbb{1}^{(M_2,\eta)}_{M_3,m,n} | \mathcal{F}^{\infty}_{Z^3_{m,n}} \right\} &= \overline{q}. \end{aligned} \tag{C.14}$$

The quantities $\overline{p}, \overline{q}$ are determined as follows

$$\begin{aligned} \overline{p} &= \sum_{X=0}^{\infty} \sum_{X_4=0}^{\infty} \sum_{U_4=0,1} b^{(3)}(X) \pi_4^{\infty}(X_4, U_4) c_1^X(X_4, U_4), \\ \overline{q} &= \sum_{X=0}^{\infty} \sum_{X_4=0}^{\infty} \sum_{U_4=0,1} b^{(3)}(X) \pi_4^{\infty}(X_4, U_4) c_5^X(X_4, U_4), \end{aligned} \tag{C.15}$$

where $\pi_4^{\infty}(X_4, U_4)$ is the equilibrium jump chain probability that node 4 is in state (X_4, U_4) when operating in isolation (i.e., when node 4 is the only node in the network). $b^{(3)}(X)$ is the limiting probability as $X_0^3 \uparrow \infty$ that a first backoff of node 3 occurs when $X_3 = X$, service starting with X_0^3 packets. $c_1^X(X_4, U_4)$ is the probability that nodes 1 or 2 first gain the medium when node 3 has a first backoff with $X_3 = X$ packets and the state of node 4 as given. The remaining definitions for $\overline{q}$ are similar. Thus in this case there is no simple formula and $\overline{p}, \overline{q}$ depend on the backoff parameter γ as well as the arrival rates at nodes 3 and 4.

For the case of switching out of M_4 we have the following result, again making the corresponding definitions as in Theorem 6.

THEOREM 8. *For any $Z_{m,n}^{(4,M_4^{(5)})}$ stopping time, there is a $\eta_m > 0$ sufficiently small so that, for all $\eta_m > \eta > 0$*

$$\mu\left\{ \mathbb{1}^{(M_2,\eta)}_{M_4,m,n} \vee \mathbb{1}^{(M_3,\eta}_{M_4,m,n)} \mid \mathcal{F}^{\infty}_{Z_{m,n}^{(4,M_4^{(5)})}} \right\} = 1\ \mu\ a.s., \tag{C.16}$$

and so that

$$\mu\left\{ \mathbb{1}^{(M_1,\eta)}_{M_4,m,n} | \mathcal{F}^{\infty}_{Z_{m,n}^{(4,M_4^{(5)})}} \right\} = 0,\ \mu\ a.s. \tag{C.17}$$

A similar result holds for $Z_{m,n}^{(5,M_4^{(4)})}$ stopping times.

This concludes Part C.

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

APPENDIX D: FLUID LIMIT PROOFS – PART D

In Parts A–C we have established a) 'sawtooth' properties and some constraints on those sample paths, b) what will occur at the end of a given M_k-period, $k = 1, 2, 3, 4$ and c) that a natural state will be entered in finite time before the network can empty almost surely. What has not been shown, is whether any M_1-period would ensue at all. The purpose of this section is to show that M_1-periods will occur μ a.s. following a natural state, provided ρ is sufficiently close to 1.

In fact, establishing this result is not strictly needed to prove instability. If there is a last visit to queues 1 and 2 (which might occur when they are both empty), then these two queues must grow linearly and therefore the fluid system is unstable. Nevertheless, we will show that an infinite sequence of M_1-periods will occur μ almost surely and in strictly bounded time, following T_N to enter a non-empty natural state,

Denote by $\tau_p^{(1,2)} : C[0,\infty] \to [0,\infty]$ as the first point of increase of either I_1, I_2, as in Definition 5, following T_N. It is easily shown that $\tau_p^{(1,2)}$ is a $\mathcal{F}_{t^+}$ stopping time, and corresponds to the start of a positive M_1-period. Our main result is:

THEOREM 9. *There exists a* $0 < T_V < \infty$ *such that* $\tau_p^{(1,2)} < T_N + T_V$, μ *a.s.*

We first show that the issue of occurrence of an M_1-period arises only when there is at least one zero queue. To see this, consider the network at time T_N, and, without loss of generality, suppose $T_N > 0$. If $Q_\ell(T_N) = q_\ell > 0$, $\ell = 1, \ldots, 6$ then continuity implies that this actually holds for some small interval $[T_N - \xi, T_N]$, depending on ω, with $\xi > 0$. During any such interval, one of the M_1-, M_2-, M_3-, or M_4-periods has to be active, with probability 1, as a consequence of Lemma 9. Moreover if the active period is not M_1, then one will follow in bounded time $\tau_p^{(1,2)} < T_N + T_S$, as a consequence of Theorems 6, 7 and 8 and because the subnetwork determined by nodes 3, 4, 5, and 6 is work-conserving once a natural state is entered.

Next, consider the cases where a subset of queues $\{q_1, q_2, q_3, q_6\}$ are 0. Rather than deal with all every such cases, we will consider just one. The approach for the remaining cases will then become apparent. We therefore focus on the case $q_3 = q_4 = q_5 = q_6 = 0$ with $q_1 \vee q_2 > 0$, and show that $\tau_p^{(1,2)}$ is effectively T_N on the fluid scale. Other cases are simpler to address.

For the above case, it will be sufficient to prove the following lemma.

LEMMA 15. *Given* $\epsilon > 0, \eta > 0$, *and* $0 < Q_L < Q_H$, *there exists* $\delta > 0$, *such that* $\liminf_{R\to\infty} \mu_R(\Xi) < \eta$, *where*

$$\Xi \doteq \{\tau_p^{(1,2)} > T_N + \epsilon; Q_\ell^R(T_N) < \delta, \ell = 3,4,5,6, \\ Q_1^R(T_N) \vee Q_2^R(T_N) \in (Q_L, Q_H)\}.$$

It then follows that

$$\mu\left\{\tau_p^{(1,2)} > T_N + \epsilon;\ Q_\ell(T_N) = 0, \ell = 3,4,5,6; Q_1 \vee Q_2 \in (Q_L, Q_H)\right\} = 0,$$

by Theorem 2.1 on page 11 of [1]. Thus on occurrence of $Q_\ell = 0, \ell = 3,4,5,6$ together with $Q_1 \vee Q_2 > 0$ the start of a M_1 period must occur in negligible time after T_N on the fluid scale.

The proof of Lemma 15 is strongly specific to our choice of network and so we only provide a sketch of the proof, which relies on the following definition. Consider the sequence of node activations in the network and in particular node 4 and 5 activations.

DEFINITION 8. A control step, at index k of the jump chain, is $C_{J,k} \in \{4,5\}$. At index 0, it is 4, unless only node 5 is active, in which case it is 5. At any subsequent index $k \geq 1$, the control step is determined as follows: (i) if both nodes 4 and 5 are active at index k, $C_{J,k}$ is the node that has become active first, (ii) if only one is active, $C_{J,k}$ is the active node, (iii) if neither, then $C_{J,k}$ is the last node that was active.

A control swap to node 5 occurs at step k if $C_{J,k-1} = 4$ and $C_{J,k} = 5$, and vice versa for a control swap to node 4. These events are denoted as $4 \to 5$ and $5 \to 4$. The step at which the r-th control swap takes place is a (discrete) stopping time $\tau_r^{(S)}$.

For the subnetwork of nodes 3, 4, 5, and 6 in isolation, it is readily shown that the probability is 1 that an infinite number of control swaps occur, $\tau_r^{(S)} < \infty$, $\tau_1^{(S)} < \tau_2^{(S)} < \cdots$, with corresponding filtration $\{\mathcal{G}_r\}_{r\in\mathbb{N}}$.

Next let S be the stopping time until one of nodes 1 and 2 gain the medium (this event is blocked since we are considering the subnetwork in isolation). Using the properties of our network it can then be shown that

LEMMA 16. $\exists \epsilon > 0$ *and* $N_S \in \mathbb{N}$ *such that* $\forall r \in \mathbb{N}$, $\mathbb{P}\{S < r + N_S \mid \mathcal{G}_r\} > \epsilon$.

The proof relies on showing that once a control swap has taken place, say $4 \to 5$, then within a bounded number of additional control swaps, either it will occur that node 3 has a backoff with $Q_4 = 0$ or node 6 has a backoff with $Q_5 = 0$ and with probability at least $\epsilon > 0$.

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

The result of Lemma 16 implies that $m_S = \mathbb{E}[S] < \infty$ and actually that $\mathbb{P}\{S > rN_S\} < (1-\epsilon)^r$, see [42], Ex. E10.5 page 233, for example. That is, nodes 1 and 2 will gain the medium within a number of control swaps which has finite expectation. It then follows from Markov's inequality, that given any $\eta > 0$, there is a number of control swaps M_η such that $\mathbb{P}\{S > M_\eta\} < \eta$.

Given the above, in proving Lemma 15, it will be enough to show the following lemma.

LEMMA 17. *Furthermore, given $\epsilon > 0, \eta_P > 0$ and any fixed number of control swaps $M \in \mathbb{N}$, let S_M be the total number of steps to complete M control swaps. Then there exists a $\delta > 0$ such that for all R sufficiently large*

$$\mathbb{P}\{S_M > \lceil R\epsilon \rceil | \; Q_\ell^R(0) \leq \delta, \ell = 3, \ldots, 6\} < \eta_P.$$

Lemma 17 is sufficient since we are dealing with the case when queues 3,4,5,6 are $o(R)$ in the prelimit, by assumption. Moreover in proving Theorem 9 we will choose $M := M_\eta$ as above, so that the probability more control swaps are needed can be taken arbitrarily small.

The result of the lemma clearly relies on the supposed initial conditions. It can be demonstrated via a construction. The construction works by determining M intervals so that the probability of a control swap in each is close to 1, and so the total queue length at the start of each interval is small based on the arrivals which might have taken place during previous intervals. The proof is then completed by showing that given any control step at the start of an interval, say 5, a swap $5 \to 4$ will occur with high probability in a number of steps in proportion to the initial condition for the interval. Since only M such intervals are required, a δ can be obtained accordingly.

Other cases are dealt with similarly but there is some positive but bounded delay before an M_1-period occurs. For example if $q_1 = q_2 = q_6 = q_5 = 0$ and $q_3 > 0$, then one shows that a M_2-period occurs in negligible fluid time.

Thus in all the cases, we may can show that M_1-periods occur within bounded fluid time following a natural state. This concludes Part D.

APPENDIX E: ADDITIONAL PROOFS

E.I. Proof of Lemma 1. The proof relies on basic sample path properties of the fluid limit process $\{Q(t)\}$ as described in Subsection 4.1. First of all, the M_1-period that initiates the i-th cycle ends at time $t_i + T_{i1}$, with

$$T_{i1} = \frac{Q_1(t_i)}{1-\rho_1} \vee \frac{Q_2(t_i)}{1-\rho_2} \leq \frac{Q_1(t_i) \vee Q_2(t_i)}{1-\rho_0} \leq \frac{L(t_i)}{1-\rho_0}.$$

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

Define $K(t) = (Q_3(t) \vee Q_4(t)) + (Q_5(t) \vee Q_6(t))$ and recall that $\rho = \rho_0 + \rho_3 + \rho_6$. Then

$$\begin{aligned} K(t_i + T_{i1}) &\leq K(t_i) + (\rho_3 + \rho_6)T_{i1} \\ &\leq L(t_i) - (Q_1(t_i) \vee Q_2(t_i)) + (\rho_3 + \rho_6)\frac{Q_1(t_i) \vee Q_2(t_i)}{1 - \rho_0} \\ &= L(t_i) - \frac{(1-\rho)(Q_1(t_i) \vee Q_2(t_i))}{1 - \rho_0} \\ &= L(t_i) - (1-\rho)T_{i1}, \end{aligned}$$

which may also be seen from the fact that $L(t)$ decreases at a rate $1 - \rho$ or larger during the time interval $[t_i, t_i + T_{i1}]$ and $K(t_i + T_{i1}) = L(t_i + T_{i1})$ since $Q_1(t_i + T_{i1}) = Q_2(t_i + T_{i1}) = 0$.

Define $T_0 = \frac{K(t_i+T_{i1})}{1-\rho_3-\rho_6}$. We distinguish between two cases, depending on whether an M_4-period starts before time $t_i + T_{i1} + T_0$ or not.

If no M_4-period occurs before time $t_i + T_{i1} + T_0$, then $K(t)$ decreases at a rate $1 - \rho_3 - \rho_6$ or larger for all $t \in [t_i + T_{i1}, t_i + T_{i1} + T_0]$ and reaches zero no later than time $t_i + T_{i1} + T_0$, unless an M_1-period intervenes. This implies that the next M_1-period must start no later than time $t_i + T_{i1} + T_0$.

Using the above results, a simple calculation shows that

$$\begin{aligned} t_{i+1} - t_i \leq T_{i1} + T_0 &\leq T_{i1} + \frac{L(t_i) - (1-\rho)T_{i1}}{1 - \rho_3 - \rho_6} \leq \frac{L(t_i)}{(1-\rho_0)(1-\rho_3-\rho_6)} \\ &\leq C_T L(t_i). \end{aligned}$$

Also, $L(t)$ has continuously decreased during the cycle, so $L(t_{i+1}) - L(t_i) \leq 0$.

Now suppose that an M_4-period does start at some time $t_0 \in [t_i + T_{i1}, t_i + T_{i1} + T_0]$, and ends at time u_0.

Since $K(t)$ decreases at a rate $1 - \rho_3 - \rho_6$ or larger during the time interval $[t_i + T_{i1}, t_0]$, it follows that

$$K(t_0) \leq K(t_i + T_{i1}) - (1 - \rho_3 - \rho_6)(t_0 - t_i - T_{i1}).$$

Noting that $Q_4(t_0), Q_5(t_0) \leq K(t_0)$, we conclude that the duration of the M_4-period is no longer than

$$u_0 - t_0 \leq \left(\frac{Q_4(t_0)}{1 - \rho_4} \wedge \frac{Q_5(t_0)}{1 - \rho_5}\right) \leq \frac{K(t_0)}{1 - (\rho_4 \vee \rho_5)}.$$

Since $K(t)$ increases at a rate no larger than $\rho_3 + \rho_6$ during the time interval $[t_0, u_0]$, it follows that

$$K(u_0) \leq K(t_0) + (\rho_3 + \rho_6)(u_0 - t_0).$$

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

The M_4-period will cause queue 4 to empty at some point and become smaller than queue 3, and likewise queue 5 must empty at some point and become smaller than queue 6. Because M_4-periods can no longer be initiated from M_2 and M_3, $K(t)$ decreases at a rate $1-\rho_3-\rho_6$ or larger from time u_0 onward, and reaches zero no later than time $u_0+\frac{K(u_0)}{1-\rho_3-\rho_6}$, unless an M_1-period intervenes. This implies that the next M_1-period must start no later than time $u_0+\frac{K(u_0)}{1-\rho_3-\rho_6}$.

Combining the above results, we obtain

$$\begin{aligned}
t_{i+1}-t_i &\le u_0+\frac{K(u_0)}{1-\rho_3-\rho_6}-t_i\\
&= T_{i1}+(t_0-t_i-T_{i1})+(u_0-t_0)+\frac{K(u_0)}{1-\rho_3-\rho_6}\\
&\le T_{i1}+(t_0-t_i-T_{i1})+\frac{K(t_0)+(u_0-t_0)}{1-\rho_3-\rho_6}\\
&\le T_{i1}+(t_0-t_i-T_{i1})+\left(1+\frac{1}{1-(\rho_4\vee\rho_5)}\right)\frac{K(t_0)}{1-\rho_3-\rho_6}\\
&\le T_{i1}-\frac{t_0-t_i-T_{i1}}{1-(\rho_4\vee\rho_5)}+\frac{(2-(\rho_4\vee\rho_5))K(t_i+T_{i1})}{(1-(\rho_4\vee\rho_5))(1-\rho_3-\rho_6)}\\
&\le T_{i1}+\frac{(2-(\rho_4\vee\rho_5))(L(t_i)-(1-\rho)T_{i1})}{(1-(\rho_4\vee\rho_5))(1-\rho_3-\rho_6)}\\
&\le \frac{(2-(\rho_4\vee\rho_5))L(t_i)+\rho_0(1-(\rho_4\vee\rho_5))T_{i1})}{(1-(\rho_4\vee\rho_5))(1-\rho_3-\rho_6)}\\
&= \frac{L(t_i)}{(1-(\rho_4\vee\rho_5))(1-\rho_3-\rho_6)}+\frac{L(t_i)+\rho_0T_{i1}}{1-\rho_3-\rho_6}\\
&\le \frac{L(t_i)}{(1-(\rho_4\vee\rho_5))(1-\rho_3-\rho_6)}+\frac{L(t_i)}{(1-\rho_0)(1-\rho_3-\rho_6)}\\
&= \frac{L(t_i)}{1-\rho_3-\rho_6}\left(\frac{1}{1-\rho_0}+\frac{1}{1-(\rho_4\vee\rho_5)}\right)=C_TL(t_i).
\end{aligned}$$

Also, $L(t)$ has only increased during the M_4-period at a rate no larger than $\rho=\rho_0+\rho_3+\rho_6$, so

$$L(t_{i+1})-L(t_i)\le\rho(u_0-t_0)\le\frac{\rho K(t_0)}{1-(\rho_4\vee\rho_5)}\le\frac{\rho L(t_i)}{1-(\rho_4\vee\rho_5)}=C_LL(t_i).$$

E.II. Proof of Lemma 2. Denote by t_1 and t_2 the times that the cycles start and by u_1 and u_2 the times that the M_1-periods end. First assume $Q_1(t_1)\vee Q_2(t_1)\le\epsilon L(t_1)$. Then, $(Q_3(t_1)\vee Q_4(t_1))+(Q_5(t_1)\vee Q_6(t_1))\ge(1-2\epsilon)L(t_1)$, so we must have $(Q_3(t_1)\vee Q_4(t_1))\ge(1-2\epsilon)L(t_1)/2$ or $(Q_5(t_1)\vee$

$Q_6(t_1)) \geq (1-2\epsilon)L(t_1)/2$. In the former scenario, with probability 3/8 the M_1-period is followed by an M_2-period, which will last for an amount of time no less than $\frac{Q_3(t_1)}{1-\rho_3} \vee \frac{Q_4(t_1)}{1-\rho_4} \geq \frac{Q_3(t_1)\vee Q_4(t_1)}{1-\rho_4} \geq \frac{(1-2\epsilon)L(t_1)}{2(1-\rho_4)}$. Likewise, in the latter scenario, with probability 3/8 the M_1-period is followed by an M_3-period, which will last for an amount of time no less than $\frac{Q_5(t_1)}{1-\rho_5} \vee \frac{Q_6(t_1)}{1-\rho_6} \geq \frac{Q_5(t_1)\vee Q_6(t_1)}{1-\rho_5} \geq \frac{(1-2\epsilon)L(t_1)}{2(1-\rho_5)}$. Thus, in either scenario, with probability at least 3/8, the time until the start of the next cycle is at least $\frac{(1-2\epsilon)L(t_1)}{2(1-(\rho_4\wedge\rho_5))}$, so that

$$Q_1(t_2) \vee Q_2(t_2) \geq Q_2(t_2) \geq \frac{\rho_2(1-2\epsilon)L(t_1)}{2(1-(\rho_4 \wedge \rho_5))}.$$

Invoking the fact that $L(t_2) \leq C_L L(t_1)$, with C_L as defined in the previous lemma, we find that

$$Q_1(t_2) \vee Q_2(t_2) \geq \epsilon L(t_2),$$

with ϵ as specified in the statement of the lemma.

Now consider a cycle with $Q_1(t_k) \vee Q_2(t_k) \geq \epsilon L(t_k)$, $k = 1, 2$. Then

$$Q_i(u_k) = Q_i(t_k) + \rho_i \frac{Q_1(t_k) \vee Q_2(t_k)}{1-\rho_2}, \text{ for } i = 3, 4, 5, 6.$$

Note that $0 \leq Q_i(t_k) \leq (1-\epsilon)L(t_k)$, $i = 3, 4, 5, 6$, and $\epsilon L(t_k) \leq Q_1(t_k) \vee Q_2(t_k) \leq L(t_k)$. Then it is easily verified that the queues are weakly balanced at time u_k with $\beta^{\min}$ and $\beta^{\max}$ as given in the statement of lemma.

E.III. Proof of Theorem 2. Let $(U(n), X(n))$ denote the jump chain obtained from the continuous-time Markov process by uniformization according to a Poisson clock of rate β as described in Appendix A.I. In order to prove Theorem 2 for the original stochastic process, it suffices to establish a similar result for the jump chain:

$$\lim_{\|X(0)\|\to\infty} \mathbb{P}_{X(0)}\{\liminf_n \|X(n)\| = \infty\} = 1. \tag{E.1}$$

The above result will be established via Theorem 3.1 of [27], which is reproduced below for completeness.

THEOREM 10. *Suppose that for a Markov chain* $\{X(n); n = 0, 1, 2, \dots\}$ *with discrete state space* S*, there exist positive functions* $W(\cdot)$ *and* $\Delta(\cdot)$ *on* S*, and a positive constant* c_0*, such that*

$$\mathbb{E}\,[W(X(n+1))|\mathcal{F}_n] \leq W(X(n)) - \Delta(X(n)), \tag{E.2}$$

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

whenever $X(n) \in S_{c_0} = \{x \in S : W(x) \leq c_0\}$, *with* $\mathcal{F}_n := \sigma(X(0), X(1), \ldots, X(n))$. *Then for all* $x \in S$,

$$\mathbb{P}_x \left\{ \sum_{n=0}^{\infty} \Delta(X(n)) < \infty \right\} \geq 1 - W(x)/c_0.$$

In order to apply the above theorem, we will consider the function $W(x) = \mathbb{E}[\mathcal{W}|X(0) = x]$, where the random variable $\mathcal{W}$ is defined as

$$\mathcal{W} := \sum_{n=0}^{\|X(0)\|T} [1 + \|X(0)\| + a\|X(n)\|]^{-m}$$

for some positive constants a and T to be determined later and $m > 1$. Note that, with minor abuse of notation, $W(X(0) = x, U(0) = u) = W(x)$, i.e., W only depends on the queue and not on the activity vector. The function $W(x)$ may be interpreted as the following approximation to a Lyapunov function for the fluid limit process

$$\|x\|^{m-1} W(x) \approx \mathbb{E}_{\hat{x}} \left[\int_0^T (1 + a\|Q_{\hat{x}}(t/\beta)\|)^{-m} \mathrm{d}t \right] = V(Q_{\hat{x}}(t)), \tag{E.3}$$

with equality when $\|x\| \to \infty$, and $\hat{x} = \frac{x}{\|x\|}$ is the initial state of the fluid limit process. Then it follows from the instability of the fluid limit process that we can choose a and T large enough such that $V(Q_{\hat{x}}(t+r)) < V(Q_{\hat{x}}(t))$ for any $r > 0$ and any initial state $\hat{x}$. This implies that

$$\|x\|^m \mathbb{E}\left[W(X(n+1)) - W(X(n))|\mathcal{F}_n\right] \leq -\text{constant},$$

when $x = X(n)$ and $\|x\|$ is sufficiently large. Thus, we can apply Theorem 10.

The detailed arguments may be described as follows. First of all, note that

$$\begin{aligned} &\mathbb{E}\left[W(X(1)) - W(X(0))|X(0) = x, U(0) = u\right] \\ &\quad = \mathbb{E}\left[\theta^1 \mathcal{W} - \mathcal{W}|X(0) = x, U(0) = u\right], \end{aligned}$$

where θ^1 is the usual backward shift operator on the sample path space [27]. We write $\theta^1 \mathcal{W} - \mathcal{W} = A + B + C$, where

$$\begin{aligned} A &= -[1 + \|X(0)\| + a\|X(0)\|]^{-m}, \\ B &= \sum_{n=1}^{\|X(0)\|T} \left\{ [1 + \|X(1)\| + a\|X(n)\|]^{-m} \right. \\ &\quad \left. - [1 + \|X(0)\| + a\|X(n)\|]^{-m} \right\}, \end{aligned}$$

and

$$C = \sum_{n=\|X(0)\|T+1}^{\|X(1)\|T} [1 + \|X(1)\| + a\|X(n)\|]^{-m}.$$

The term 'A' provides the negative drift and the other terms can be bounded as follows. Using the fact that $\|X(1)\| \geq \|X(0)\| - 1$, and noting that $[\cdot]^{-m}$ is a convex decreasing function, we have

$$B \leq \sum_{n=1}^{\|X(0)\|T} m[\|X(0)\| + a\|X(n)\|]^{-m-1}. \tag{E.4}$$

Multiplying both sides by $\|X(0)\|^m$, we see that

$$\|X(0)\|^m B \leq \frac{m}{\|X(0)\|} \sum_{n=1}^{\|X(0)\|T} \left(1 + a\frac{\|X(n)\|}{\|X(0)\|}\right)^{-m-1}. \tag{E.5}$$

Let $X(0) = x$ and $\hat{x} := x/\|x\|$. For any x, the random variable in the right-hand-side (RHS) of (E.5) is bounded by mT, and hence

$$\limsup_{\|x\|\to\infty} \mathbb{E}_{\hat{x}} \left[\|x\|^m B\right] \leq \mathbb{E}_{\hat{x}} \left[m \int_0^T [1 + a\|Q(s/\beta)\|]^{-m-1} \mathrm{d}s\right],$$

because of the weak limit convergence of $\frac{1}{\|x\|} X^{(\|x\|)}(\|x\|t) \Rightarrow Q(t/\beta)$ over $[0, T]$ and uniform integrability of the random variables of the form RHS of (E.5).

Next, for 'C', it is sufficient to consider the case that $\|X(1)\| = \|X(0)\|+1$, where

$$C \leq T[1 + \|x\| + 1 + a(\|X(\|x\|T)\| - T)]^{-m}.$$

Similarly to 'B', multiplying both sides with $\|x\|^m$ and taking the limit gives

$$\limsup_{\|x\|\to\infty} \mathbb{E}_{\hat{x}} \left[\|x\|^m C\right] \leq \mathbb{E}_{\hat{x}} \left[T[1 + a\|Q(T/\beta)\|]^{-m}\right], \tag{E.6}$$

again, because $\|x\|^m C < T$ (thus, uniform integrability holds) and by the weak limit convergence. Putting the bounds together, we obtain

$$\begin{aligned} \limsup_{\|x\|\to\infty} \|x\|^m \mathbb{E}_{\hat{x}} \left[\theta^1 \mathcal{W} - \mathcal{W}\right] \quad \leq \quad & -(1+a)^{-m} \\ & + m\mathbb{E}_{\hat{x}} \left[\int_0^\infty (1 + aL(s/\beta))^{-m-1} \mathrm{d}s\right] \\ & + \mathbb{E}_{\hat{x}} \left[T(1 + aL(T/\beta))^{-m}\right], \end{aligned}$$

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

because $\|Q(s)\| \geq L(s)$ based on our notation with some initial state $Q(0) = \hat{x}$ such that $\|\hat{x}\| = 1$. Consider the cycle pairs D_k, $k = 1, 2, \ldots$, as defined for Theorem 1. Then,

$$\begin{aligned}
\mathbb{E}_{\hat{x}}\left[\int_0^\infty (1 + aL(s/\beta))^{-m-1} \mathrm{d}s\right] &\leq \beta \mathbb{E}_{\hat{x}}\left[\sum_{k=0}^{\infty} \int_{T_k}^{T_{k+1}} (1 + aL(s))^{-m-1} \mathrm{d}s\right] \\
&\leq \beta \mathbb{E}_{\hat{x}}\left[\sum_{k=0}^{\infty} \int_{T_k}^{T_{k+1}} (1 + a\theta L_k)^{-m-1} \mathrm{d}s\right] \\
&\leq \beta \mathbb{E}_{\hat{x}}\left[\sum_{k=0}^{\infty} \Delta T_k (a\theta L_k)^{-m-1}\right] \\
&\leq \beta C_{LT} (a\theta)^{-m-1} \sum_{k=0}^{\infty} \mathbb{E}_{\hat{x}}\left[L_k^{-m}\right].
\end{aligned}$$

where we have used Proposition 1 (ii), (i) for the second and last inequality respectively. As we saw in the proof of Theorem 1, for $\rho \in (\rho^*, 1]$, $\mathbb{E}[L_k^{-m}] \leq L_0^{-m} \alpha^k$. Therefore,

$$m\mathbb{E}_{\hat{x}}\left[\int_0^\infty (1 + aL(s/\beta))^{-m-1} \mathrm{d}s\right] \leq m\beta C_{LT} (a\theta)^{-m-1} \frac{L_0^{-m}}{1-\alpha}. \tag{E.7}$$

So, we can choose a large enough to ensure that the RHS of (E.7) is less than $\frac{1}{3}(1+a)^{-m}$. Next we show that we can choose T large enough such that

$$\mathbb{E}_{\hat{x}}\left[T[1 + aL(T/\beta)]^{-m}\right] \leq \frac{1}{3}(1+a)^{-m}. \tag{E.8}$$

Note that

$$\mathbb{E}_{\hat{x}}\left[T[1 + aL(T/\beta)]^{-m}\right] \leq a^{-m} \mathbb{E}_{\hat{x}}\left[TL^{-m}(T/\beta)\right], \tag{E.9}$$

and by Theorem 1, $\limsup_{T\to\infty} \mathbb{E}_{\hat{x}}[TL^{-m}(T)] = 0$, for $\rho \in (\rho^*, 1]$. Hence, we can choose T large enough such that (E.8) holds.

Therefore,

$$\limsup_{\|x\|\to\infty} \|x\|^m \mathbb{E}\left[W(X(1)) - W(X(0)) | (X(0), U(0)) = (x, u)\right] \leq -\frac{1}{3}(1+a)^{-m}.$$

This means that there exists a positive constant $\|x_0\|$ such that,

$$\mathbb{E}\left[W(X(1)) - W(X(0)) | X(0) = (x, u)\right] \leq -\frac{1}{6}(1+a)^{-m} \|x\|^{-m},$$

whenever $\|x\| > \|x_0\|$. Let $c_0 = W(x_0) = W(\|x_0\|)$. On the other hand, it follows from (E.3) that $\limsup_{\|x\|\to\infty} W(x) = 0$, which means that S_{c_0} is well-defined and also c_0 can be made arbitrarily small by letting $\|x_0\| \to \infty$. Therefore, the conditions of Theorem (10) are satisfied with $\Delta(x) = \frac{1}{6}(1+a)^{-m}\|x\|^{-m}$. This shows that

$$\mathbb{P}_{X(0)}\left\{\sum_{n=0}^{\infty}\frac{\text{constant}}{\|X(n)\|^m} < \infty\right\} \to 1, \tag{E.10}$$

as $\|X(0)\| \to \infty$, which implies (E.1).

REFERENCES

[1] Billingsley, P. (1968). *Weak Convergence of Probability Measures*, Wiley–Interscience. MR0233396

[2] Billingsley, P. (1995). *An Introduction to Probability and Measure*, Wiley–Interscience, 3rd edition. MR1324786

[3] Boorstyn, R.R., Kershenbaum, A., Maglaris, B., Sahin, V. (1987). Throughput analysis in multihop CSMA packet radio networks. *IEEE Trans. Commun.* **35**, 267–274.

[4] Bouman, N., Borst, S.C., van Leeuwaarden, J.S.H., Proutiere, A. (2011). Backlog-based random access in wireless networks: fluid limits and delay issues. In: *Proc. ITC 23 Conf.*, 39–46.

[5] Dai, J.G. (1995). On positive Harris recurrence of multiclass queueing networks: a unified approach via fluid limit models. *Ann. Appl. Prob.* **5**, 49–77. MR1325041

[6] Dai, J.G. (1996). A fluid limit model criterion for instability of multiclass queueing networks. *Ann. Appl. Prob.* **6**, 751–757. MR1410113

[7] Dai, J.G., Meyn, S.P. (1995). Stability and convergence of moments for multiclass queueing networks via fluid limit models. *IEEE Trans. Aut. Control* **40(11)**, 1889–1904. MR1358006

[8] Durvy, M., Dousse, O., Thiran, P. (2007). Modeling the 802.11 protocol under different capture and sensing capabilities. In: *Proc. Infocom 2007 Conf.*

[9] Durvy, M., Thiran, P. (2006). A packing approach to compare slotted and non-slotted medium access control. In: *Proc. Infocom 2006 Conf.*

[10] Ethier, N., Kurtz, T.G. (1985). *Markov Processes: Characterization and Convergence*, Wiley–Interscience. MR0838085

[11] Feller, W. (1970). *An Introduction to Probability Theory and Its Applications*, Vol. 2, John Wiley and Sons.

[12] Feuillet, M., Proutiere, A., Robert, Ph. (2010). Random capture algorithms: fluid limits and stability. In: *Proc. ITA Workshop.*

[13] Foss, S.G., Kovalevskii, A.P. (1999). A stability criterion via fluid limits and its application to a polling model. *Queueing Systems* **32**, 131–168. MR1720552

[14] Remerova, M., Foss, S.G., Zwart, B. (2014). Random fluid limit of an overloaded polling model. *Advances in Applied Probability* **46(1)**, 76–101.

[15] Ghaderi, J., Srikant, R. (2010). On the design of efficient CSMA algorithms for wireless networks. In: *Proc. CDC 2010 Conf.*

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

[16] Ghaderi, J., Srikant, R. (2012). The impact of access probabilities on the delay performance of Q-CSMA algorithms in wireless networks. *IEEE/ACM Trans. Netw.*, to appear.

[17] Jiang, L., Leconte, M., Ni, J., Srikant, R., Walrand, J. (2011). Fast mixing of parallel Glauber dynamics and low-delay CSMA scheduling. In: *Proc. Infocom 2010 Mini-Conf.* MR2982678

[18] Jiang, L., Shah, D., Shin, J., Walrand, J. (2010). Distributed random access algorithm: scheduling and congestion control. *IEEE Trans. Inf. Theory* **56(12)**, 6182–6207. MR2809994

[19] Jiang, L., Walrand, J. (2008). A distributed CSMA algorithm for throughput and utility maximization in wireless networks. In: *Proc. Allerton 2008 Conf.*

[20] Jiang, L., Walrand, J. (2010). A distributed CSMA algorithm for throughput and utility maximization in wireless networks. *IEEE/ACM Trans. Netw.* **18(3)**, 960–972.

[21] Kovalevskii, A.P., Topchii, V.A., Foss, S.G. (2005). On the stability of a queueing system with uncountable branching fluid limits. *Prob. Inf. Trans.* **41(3)**, 254–279. MR2163853

[22] Lee, C.-H., Eun, D.Y., Yun, S.-Y., Yi, Y. (2012). From Glauber dynamics to Metropolis algorithm: smaller delay in optimal CSMA. In: *Proc. ISIT 2012 Conf.*

[23] Liew, S.C., Kai, C.H., Leung, J., Wong, B. (2010). Back-of-the-envelope computation of throughput distributions in CSMA wireless networks. *IEEE Trans. Mob. Comp.* **9(9)**, 1319–1331.

[24] Liu, J., Yi, Y., Proutière, A., Chiang, M., Poor, H.V. (2008). Towards utility-optimal random access without message passing. *Wireless Commun. Mobile Comput.* **10(1)**, 115–128.

[25] Lotfinezhad, M., Marbach, P. (2011). Throughput-optimal random access with order-optimal delay. In: *Proc. Infocom 2011 Conf.*

[26] Marbach, P., Eryilmaz, A. (2008). A backlog-based CSMA mechanism to achieve fairness and throughput-optimality in multihop wireless networks. In: *Proc. Allerton 2008 Conf.*

[27] Meyn, S.P. (1995). Transience of multiclass queueing networks via fluid limit models. *Ann. Appl. Prob.* **5**, 946–957. MR1384361

[28] Ni, J., Tan, B., Srikant, R. (2010). Q-CSMA: queue-length based CSMA/CA algorithms for achieving maximum throughput and low delay in wireless networks. In: *Proc. Infocom 2010 Mini-Conf.*

[29] Rajagopalan, S., Shah, D., Shin, J. (2009). Network adiabatic theorem: an efficient randomized protocol for content resolution. In: *Proc. ACM SIGMETRICS/Performance 2009 Conf.*

[30] Robert, Ph., Véber, A. (2012). On the fluid limits of a resource sharing algorithm with logarithmic weights. arXiv:1211.5968v1.

[31] Rogers, L.C.G., Williams, D. (1989). *Diffusions, Markov Processes and Martingales, Vol. II*, Wiley, London.

[32] Royden, H.L. (1987). *Real Analysis*, Prentice Hall, 3rd edition. MR1013117

[33] Shah, D., Shin, J. (2010). Delay-optimal queue-based CSMA. In: *Proc. ACM SIGMETRICS 2010 Conf.*

[34] Shah, D., Shin, J. (2012). Randomized scheduling algorithms for queueing networks. *Ann. Appl. Prob.* **22(1)**, 128–171. MR2932544

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.

[35] Shah, D., Shin, J., Tetali, P. (2010). Medium access using queues. In: *Proc. FOCS 2011 Conf.* MR2933306

[36] Shah, D., Tse, D.N.C., Tsitsiklis, J.N. (2011). Hardness of low delay network scheduling. *IEEE Trans. Inf. Theory* **57(12)**, 7810–7817. MR2895362

[37] Subramanian, V., Alanyali, M. (2011). Delay performance of CSMA in networks with bounded degree conflict graphs. In: *Proc. ISIT 2011 Conf.*

[38] Tassiulas, L., Ephremides, A. (1992). Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks. *IEEE Trans. Aut. Contr.* **37**, 1936–1948. MR1200609

[39] Wang, X., Kar, K. (2005). Throughput modeling and fairness issues in CSMA/CA based ad-hoc networks. In: *Proc. Infocom 2005 Conf.*

[40] Whitt, W. (1969). Weak convergence of probability measures on the function space $C[0,\infty)$. Technical Report No. 120, Stanford University. MR0261646

[41] Whitt, W. (1970). Weak convergence of probability measures on the function space $C[0,\infty)$. *Ann. Math. Statistics* **41(3)**, 939–944. MR0261646

[42] Williams, D. (1991). *Probability with Martingales*, Cambridge University Press. MR1155402

500 West 120th Street
New York, NY 10027
USA
E-mail: jghaderi@ee.columbia.edu

600 Mountain Avenue
Murray Hill, NJ 07974-0636
USA
and
P.O. Box 513, 5600 MB Eindhoven
The Netherlands
E-mail: sem@research.bell-labs.com

North Ryde, NSW 2109, Australia
E-mail: philip.witing@mq.edu.au

Downloaded from informs.org by [216.73.217.172] on 15 November 2026, at 05:56 . For personal use only, all rights reserved.