J. Appl. Prob. Spec. Vol. **51A**, 249–265 (2014)

CELEBRATING 50 YEARS OF THE APPLIED PROBABILITY TRUST

Edited by
S. ASMUSSEN, P. JAGERS, I. MOLCHANOV and L. C. G. ROGERS

Part 6. Heavy tails

TWO-NODE FLUID NETWORK WITH A HEAVY-TAILED RANDOM INPUT: THE STRONG STABILITY CASE

SERGEY FOSS, *Heriot Watt University and Novosibirsk State University*
Department of Actuarial Mathematics and Statistics, Heriot-Watt University, Edinburgh EH14 4AS, UK.
Email address: s.foss@hw.ac.uk

MASAKIYO MIYAZAWA, *Tokyo University of Science*
Tokyo University of Science, 2641 Yamazaki, Noda, Chiba, 278-8510, Japan.
Email address: miyazawa@rs.tus.ac.jp

APPLIED PROBABILITY TRUST
DECEMBER 2014

TWO-NODE FLUID NETWORK WITH A HEAVY-TAILED RANDOM INPUT: THE STRONG STABILITY CASE

BY SERGEY FOSS AND MASAKIYO MIYAZAWA

Abstract

We consider a two-node fluid network with batch arrivals of random size having a heavy-tailed distribution. We are interested in the tail asymptotics for the stationary distribution of a two-dimensional workload process. Tail asymptotics have been well studied for two-dimensional reflecting processes where jumps have either a bounded or an unbounded light-tailed distribution. However, the presence of heavy tails totally changes these asymptotics. Here we focus on the case of *strong stability* where both nodes release fluid at sufficiently high speeds to minimise their mutual influence. We show that, as in the one-dimensional case, big jumps provide the main cause for workloads to become large, but now they can have multidimensional features. We first find the weak tail asymptotics of an arbitrary directional marginal of the stationary distribution at Poisson arrival epochs. In this analysis, decomposition formulae for the stationary distribution play a key role. Then we employ sample-path arguments to find the exact tail asymptotics of a directional marginal at renewal arrival epochs assuming one-dimensional batch arrivals.

Keywords: Fluid network; Poisson and renewal arrivals; heavy-tailed distribution of batch size; workload process; stability; strong stability

2010 Mathematics Subject Classification: Primary 60K25; 90B15
Secondary 60E15; 60G17

1. Introduction

Problems concerning tail asymptotics have been studied in queueing networks and related reflecting processes for many years; new developments continue to arise. A key feature of this is an influence of the multiple boundary faces in a multidimensional state space. This requires analysis that differs from the traditional. Recent studies of these multidimensional processes have been done mainly in the *light-tail regime* where no heavy tails arise (see, e.g. [4, 9] and the references therein). Heavy-tail asymptotics have mostly been studied for processes with single boundary faces or for certain monotone characteristics.

It is natural to ask how the presence of heavy tails changes the tail asymptotics in multidimensional reflecting processes in more complex queueing networks. The aim of this paper is to address this problem for the stationary distribution of a continuous-time reflecting process in the two-dimensional nonnegative quadrant. For this, we consider a two-node fluid network with a compound input with either Poisson or renewal arrivals; this is a simple model that nevertheless retains the feature of being a multidimensional reflecting process. It can be viewed as a continuous-time approximation of a generalised Jackson network in which there can be simultaneous arrivals of large batches of customers.

We analyse the tail asymptotics for this fluid network as follows. First we assume the input to be Poisson, and derive two new decomposition formulae for a one-dimensional marginal of the stationary distribution of the workload in an arbitrary direction. These are used to derive stochastic lower and upper bounds for the marginals (Lemmas 4.3 and 4.4). From these bounds, we find the weak tail asymptotics for the stationary marginal distribution in an arbitrary direction, assuming subexponentiality of the batch size distributions (Theorem 2.1). Next we take the sample-path approach to obtain the exact tail asymptotics of the marginal stationary distribution in an arbitrary direction, assuming that the arrival batch sizes are one dimensional and have subexponential distributions, where arrival instants constitute a renewal process (Theorem 2.2). To do so requires new derivations for the corresponding fluid model as in Appendix C.

Our results relate to the tail asymptotics in both generalised Jackson networks and more general classes of max-plus systems, with heavy-tailed distributions of service times and Markovian routeing; these have been studied in [1] (see also [2, 10]). In these papers, the exact tail asymptotics were found only for the 'maximal dater' (time needed to empty the system in the steady state after stopping the input process). In tandem queues, the maximal dater coincides with the stationary sojourn time of a 'typical' customer. But the two notions differ when the routing includes feedback.

There is an extensive literature on feedforward networks with heavy-tailed distributions that include fluid queues with jump inputs, Lévy-driven queues, parallel queues, coupled queues, etc. (see, e.g. [11, 13] and the references therein). Here we consider stochastic fluid networks with feedback loops, for which in general the techniques developed in those papers cannot be applied. We do not restrict our analysis to the case of regularly varying distributions but consider a more general subexponential class. We hope this paper may stimulate further studies of networks with heavy-tailed inputs.

The paper is organised as follows. In Section 2 we introduce a fluid network with random jumps and present our main results and assumptions, including stability conditions. In subsequent sections we assume that a *strong stability condition* holds. For a model with Poisson arrivals, decomposition formulae are obtained in Sections 3 and 4. These results yield stochastic bounds and weak tail asymptotics for the marginal stationary distribution in an arbitrary direction in Section 4. In Section 5 we present sample-path analyses, establishing exact tail asymptotics in the case of one-dimensional jumps. We list some open problems in Section 6. In the appendices we provide auxiliary material that includes two short proofs and an analysis of a corresponding fluid model, and we recall basic definitions and properties of subexponential distributions.

Our results can be generalised to any dimension. The sample-path approach does not require any significant change. For the analytic approach, we need decomposition formulae which may require further restrictions on the model parameters.

2. Fluid network with compound input

Consider a two-node fluid network where nodes $i = 1, 2$ receive an input process $\mathbf{\Lambda}(\cdot) = (\Lambda_1(\cdot), \Lambda_2(\cdot))$, which is a compound process generated by the point process $\{N(\cdot)\}$ and independent, identically distributed (i.i.d.) jumps $\{(J_{1n}, J_{2n}) : n = \dots, -1, 0, 1, 2, \dots\}$, where, for convenience, the sequence $\{(J_{1n}, J_{2n})\}$ is assumed to be doubly infinite. Define the components of $\mathbf{\Lambda}$ by

$$\Lambda_i(0, t] = \sum_{n=1}^{N(0,t]} J_{in}, \qquad t > 0,\ i = 1, 2.$$

Throughout the paper, all vectors are considered as column vectors, but we omit the transpose sign for simplicity. In this section we assume only that the point processes $N_s(\cdot) := \{N(s, s+t]: t \geq 0\}$ converge weakly as $s \downarrow -\infty$ to a stationary point process $\{N^*(\cdot)\}$ on $(0, \infty)$ with finite intensity $\lambda := \mathbb{E}[N^*(0, 1]]$. In subsequent sections we require $\{N(\cdot)\}$ to be either a Poisson process (for the analytic approach) or a renewal process (for the sample-path approach). In what follows, $\boldsymbol{J} := (J_1, J_2)$ denotes a random vector having the same distribution as (J_{1n}, J_{2n}). Denote the joint distribution of (J_1, J_2) by F, with marginals F_i, and set $m_i = \mathbb{E}[J_i]$ and $\alpha_i = \lambda m_i$ for $i = 1, 2$.

Both nodes have infinite capacity buffers and release fluid at respective rates μ_i, $i = 1, 2$. For $i, j = 1, 2$ with $i \neq j$, a proportion p_{ij} of the outflow from node i goes to node j, while the remaining proportion $1 - p_{ij}$ leaves the system. Assume that

$$0 \leq p_{12}p_{21} < 1, \qquad 0 < p_{12} + p_{21}; \tag{2.1}$$

this excludes trivial boundary cases (including parallel queues, when $p_{12} = p_{21} = 0$), and, without loss of generality, we may put $p_{ii} = 0$.

Remark 2.1. We could assume that, in addition to the jump input, both nodes have continuous fluid inputs at rates β_1 and β_2, respectively, say, so that $\boldsymbol{\Lambda}$ is now defined for $t > 0$ by

$$\Lambda_i(0, t] = \beta_i t + \sum_{n=1}^{N(0,t]} J_{in}.$$

Given stability, such a model is reduced to the original model by using smaller release rates, namely, replacing μ_i, $i = 1, 2$, by $\mu_i[1 - (\beta_i + \beta_{3-i}p_{3-i,i})/(1 - p_{12}p_{21})]$. This is readily checked via (2.2) and (2.3) below where the buffer content processes are defined. Accordingly, in this paper we assume that $\beta_1 = \beta_2 = 0$.

Introduce next a buffer content process $\boldsymbol{Z}(t) := (Z_1(t), Z_2(t))$, defined as a nonnegative solution to the equations

$$Z_1(t) = Z_1(0) + \Lambda_1(0, t] + p_{21}(\mu_2 t - Y_2(t)) - \mu_1 t + Y_1(t), \tag{2.2}$$

$$Z_2(t) = Z_2(0) + \Lambda_2(0, t] + p_{12}(\mu_1 t - Y_1(t)) - \mu_2 t + Y_2(t), \tag{2.3}$$

for $t > 0$, where $Y_i(t)$ is the minimal nondecreasing process ensuring that $Z_i(t)$ remains nonnegative. As usual, we assume that sample paths are right continuous and have left-hand limits. Recalling $\boldsymbol{\Lambda}(\cdot)$ and setting $\boldsymbol{\mu} = (\mu_1, \mu_2)$, let

$$\boldsymbol{X}(t) = \boldsymbol{\Lambda}(0, t] - \begin{pmatrix} \mu_1 - \mu_2 p_{21} \\ \mu_2 - \mu_1 p_{12} \end{pmatrix} t = \boldsymbol{\Lambda}(0, t] - \boldsymbol{R}\,\boldsymbol{\mu}\,t, \qquad \boldsymbol{R} = \begin{pmatrix} 1 & -p_{21} \\ -p_{12} & 1 \end{pmatrix}.$$

Then (2.2) and (2.3) may be rewritten as

$$\boldsymbol{Z}(t) = \boldsymbol{Z}(0) + \boldsymbol{X}(t) + \boldsymbol{R}\,\boldsymbol{Y}(t), \qquad t \geq 0, \tag{2.4}$$

where $\boldsymbol{Y}(t) = (Y_1(t), Y_2(t))$. This is the standard definition of a reflecting process (in the nonnegative quadrant $\mathbb{R}_+^2$) for a given process $\boldsymbol{X}(t)$, where $\boldsymbol{Y}(t)$ is a *regulator* such that $Y_i(t)$ increases only when $Z_i(t) = 0$. Here $\boldsymbol{R}$ is a *reflection matrix* (see, e.g. [12, Section 3.5]). The conditions at (2.1) ensure that the inverse $\boldsymbol{R}^{-1}$ exists and is nonnegative. This guarantees the existence of the process $\{\boldsymbol{Z}(t): t \geq 0\}$. We refer to this process as a *two-dimensional fluid network with compound inputs*.

Recall that $\alpha_i = \mathbb{E}[\Lambda_i(0, t]]$, so $\boldsymbol{R}^{-1}\boldsymbol{\alpha}$ is the total inflow rate vector. Hence, the fluid network is stable if and only if

$$\boldsymbol{R}^{-1}\boldsymbol{\alpha} < \boldsymbol{\mu}, \tag{2.5}$$

where the inequality is strict in both coordinates (a formal proof for this stability condition is given in [8]). Write $\boldsymbol{\delta} = \boldsymbol{R}\,\boldsymbol{\mu}$, where $\boldsymbol{\delta} = (\delta_1, \delta_2)$, so that $\mathbb{E}[\boldsymbol{X}(t)] = \boldsymbol{\alpha} - \boldsymbol{\delta}$; writing $\Delta_i = \delta_i - \alpha_i$, $\mathbb{E}[X_i(1)] = -\Delta_i$, and the stability condition (2.5) is equivalent to both

$$\Delta_1 + \Delta_2 p_{21} > 0 \quad \text{and} \quad \Delta_1 p_{12} + \Delta_2 > 0. \tag{2.6}$$

Under this condition, there are in fact two different stability scenarios, namely,

$$\min\{\Delta_1, \Delta_2\} > 0, \tag{2.7}$$

$$\min\{\Delta_1, \Delta_2\} \le 0; \tag{2.8}$$

we refer to these as conditions of *strong stability* and *weak stability*, respectively.

Under condition (2.6), which we assume to hold throughout the paper, there exists a unique stationary distribution for $\boldsymbol{Z}(t)$, $\boldsymbol{\pi}$ say; we let $\boldsymbol{Z} := (Z_1, Z_2)$ denote a random vector subject to $\boldsymbol{\pi}$. We are interested in the tail behaviour of $\mathbb{P}\{c_1 Z_1 + c_2 Z_2 > x\}$ as $x \to \infty$ for a given *directional* vector $\boldsymbol{c} := (c_1, c_2) \ge \boldsymbol{0}$ satisfying $c_1 + c_2 > 0$. In this paper we study this asymptotic behaviour mostly under the strong stability condition; the other case is to be studied in a companion paper [6]. Under the strong stability condition (2.7), both nodes are sufficiently fast to process fluids given the input is always maximal, and the following result holds.

Lemma 2.1. (i) Sample-path majorant. *On any elementary event, consider an auxiliary model of two parallel queues, one at each of the two nodes $i = 1, 2$, and at each of which there is a continuous input of rate $\mu_{3-i} p_{3-i,i}$, release rate μ_i, and jump input process Λ_i. Let $\widetilde{Z}_i(t)$ be the content of node i at time t (i.e. queueing process i). If $\widetilde{Z}_i(0) \ge Z_i(0)$ then $\widetilde{Z}_i(t) \ge Z_i(t)$ for all $t > 0$.*

(ii) Stable majorant. *Assume that the input is a renewal process and (2.7) holds. Then the processes $\widetilde{Z}_i(t)$ admit a unique stationary version, and, under the natural coupling of the input processes, $\widetilde{Z}_i \ge Z_i$ almost surely (a.s.).*

Proof. Between any two jumps, the trajectories of $Z_i(t)$ and $\widetilde{Z}_i(t)$ are Lipschitz, and at any regular point t with $Z_i(t) > 0$ and $\widetilde{Z}_i(t) > 0$, the derivative of Z_i is smaller than that of $\widetilde{Z}_i$. So the inequality $\widetilde{Z}_i(t) \ge Z_i(t)$ is preserved between any two jumps. Since the jumps are synchronous and the jump sizes are equal, an induction argument completes the proof of (i). Then statement (ii) is straightforward.

Thus, we have natural upper bounds under (2.7). Similarly, easy lower bounds can be obtained by cancelling internal flow transfers. However, they hold only for the marginal stationary distributions in the coordinate directions. Even in these cases, it is unclear how they (and, in particular, the lower bounds) can be improved. We will answer these questions assuming that one of the following extra conditions holds.

(A1) $\{N(\cdot)\}$ is a Poisson process at rate λ.

(A2) The point process $\{N(\cdot)\}$ is a renewal process with i.i.d. interarrival times having a general distribution with finite mean $a := 1/\lambda$, and the batch sizes are one dimensional:

$$F(x, y) = p_1 F_1(x) + p_2 F_2(y), \qquad x, y \ge 0.$$

https://doi.org/10.1239/jap/1417528479 Published online by Cambridge University Press

Recall that jumps $\{\boldsymbol{J}_n\}$ are assumed to be i.i.d., so (A1) means that $\{\Lambda(\cdot)\}$ is a compound Poisson process. We need more notation. For each nonzero vector $\boldsymbol{c} \equiv (c_1, c_2) \geq \mathbf{0}$, define the distribution $F_{\boldsymbol{c}}$ by

$$F_{\boldsymbol{c}}(x) = \mathbb{P}\{c_1 J_1 + c_2 J_2 \leq x\}, \qquad x \geq 0,$$

and its integrated distribution $F_{\boldsymbol{c}}^I$ by

$$F_{\boldsymbol{c}}^I(x) = 1 - \frac{1}{m_{\boldsymbol{c}}} \int_x^\infty \mathbb{P}\{c_1 J_1 + c_2 J_2 > y\}\, \mathrm{d}y, \qquad x \geq 0,$$

where $m_{\boldsymbol{c}} = c_1 m_1 + c_2 m_2$. Here $\boldsymbol{c}$ can be normalized as $c_1 + c_2 = 1$, but we do not require it because it does not affect any computation. Distribution functions F and G are *weakly tail equivalent* if, for $\overline{F}(x) = 1 - F(x)$ and $\overline{G}(x) = 1 - G(x)$,

$$0 < \liminf_{x\to\infty} \frac{\overline{F}(x)}{\overline{G}(x)} \leq \limsup_{x\to\infty} \frac{\overline{F}(x)}{\overline{G}(x)} < \infty;$$

denote this $\overline{F}(x) \asymp \overline{G}(x)$. Write $\overline{F}(x) \sim \overline{G}(x)$ and say that F and G are *tail equivalent* (or have the *same tail asymptotics*) if

$$\lim_{x\to\infty} \frac{\overline{F}(x)}{\overline{G}(x)} = 1. \tag{2.9}$$

We can now formulate our main results.

Theorem 2.1. *Assume that (A1) and the strong stability condition (2.7) hold. For each $i = 1, 2$, if F_i^I is subexponential then*

$$\mathbb{P}\{Z_i > x\} \asymp \overline{F}_i^I(x). \tag{2.10}$$

For a fixed $\boldsymbol{c} > \mathbf{0}$, if F_1^I, F_2^I, and $F_{\boldsymbol{c}}^I$ are subexponential, then

$$\mathbb{P}\{c_1 Z_1 + c_2 Z_2 > x\} \asymp \overline{F}_{\boldsymbol{c}}^I(x). \tag{2.11}$$

Remark 2.2. In the particular case that F_1^I and F_2^I have regularly varying tails (see Appendix D for a definition), $F_{\boldsymbol{c}}^I$ also has a regularly varying tail for any $\boldsymbol{c} > \mathbf{0}$. However, subexponentiality of F_1^I and F_2^I does not imply that of $F_{\boldsymbol{c}}^I$ in general; furthermore, there are examples where $F_{\boldsymbol{c}}^I$ is subexponential for some but not all $\boldsymbol{c} > \mathbf{0}$.

Theorem 2.2. *Assume that (A2) and the strong stability condition (2.7) hold. For each $i = 1, 2$, if F_i^I is subexponential then*

$$\mathbb{P}\{Z_i > x\} \sim \frac{\alpha_i}{\Delta_i + \Delta_{3-i} p_{3-i,i}} \overline{F}_i^I(x) \quad \text{as } x \to \infty.$$

For fixed $\boldsymbol{c} > \mathbf{0}$, if both F_1^I and F_2^I are subexponential and $\overline{F}_1(x/c_1) \asymp \overline{F}_2(x/c_2)$, then

$$\mathbb{P}\{c_1 Z_1 + c_2 Z_2 > x\} \sim \frac{\alpha_1}{\Delta_1 + \Delta_2 p_{21}} \overline{F}_1^I\Big(\frac{x}{c_1}\Big) + \frac{\alpha_2}{\Delta_2 + \Delta_1 p_{12}} \overline{F}_2^I\Big(\frac{x}{c_2}\Big). \tag{2.12}$$

Remark 2.3. Asymptotics (2.12) also hold in the two particular cases that (a) both distributions F_i^I have regularly varying tails, and (b) one of the tails F_i^I is negligible with respect to the other. If (a) or (b) holds, we do not need to assume the weak tail-equivalence condition preceding (2.9).

We prove Theorem 2.1 using decomposition formulae obtained in Section 3, and Theorem 2.2 is proved in Section 5. In Section 4 we derive general bounds for the tail probabilities under assumption (A1) without assuming subexponentiality.

3. Decomposition formula under (A1)

In this section we assume that (A1) holds, and derive decomposition formulae for the stationary distribution in terms of moment generating functions. We first derive the stationary balance equation under the stability condition (2.6).

Let $\mathcal{C}^1(\mathbb{R}^2)$ be the set of all functions from $\mathbb{R}^2$ to $\mathbb{R}$ having continuous first-order partial derivatives. Write $f_i'(x_1, x_2) = \partial f(x_1, x_2)/\partial x_i$ for short. It follows from (2.4) that, for $f \in \mathcal{C}^1(\mathbb{R}^2)$, the increment $f(\boldsymbol{Z}(1)) - f(\boldsymbol{Z}(0))$ can be expressed in terms of integrals on $[0, 1]$ with respect to $\mathrm{d}t$, $\Lambda(\mathrm{d}t)$, and $\mathrm{d}Y_i(t)$ (formally by Itô's integral formula). Then, taking expectations with respect to $\boldsymbol{Z}(0)$ subject to the stationary distribution $\boldsymbol{\pi}$ and recalling the stationary version $\boldsymbol{Z} := (Z_1, Z_2)$,

$$\begin{aligned}&\sum_{i=1}^{2}(-\delta_i \mathbb{E}[f_i'(\boldsymbol{Z})]) + \lambda \mathbb{E}[f(\boldsymbol{Z} + \boldsymbol{J}) - f(\boldsymbol{Z})]\\ &+ \mathbb{E}_1[f_1'(0, Z_2) - p_{12} f_2'(0, Z_2)] + \mathbb{E}_2[-p_{21} f_1'(Z_1, 0) + f_2'(Z_1, 0)]\\ &\quad = 0, \end{aligned} \tag{3.1}$$

provided that all expectations are finite, where the jump size vector $\boldsymbol{J}$ is independent of everything else. Here $\mathbb{E}_i$ denotes the expectation with respect to the Palm measure for $\{Y_i(t)\}$, that is, for any bounded measurable function g on $\mathbb{R}_+$,

$$\mathbb{E}_i[g(Z_{3-i})] = \mathbb{E}_{\boldsymbol{\pi}}\bigg[\int_0^1 g(Z_{3-i}(u))\, Y_i(\mathrm{d}u)\bigg], \qquad i = 1, 2,$$

where $\mathbb{E}_{\boldsymbol{\pi}}$ denotes the expectation subject to $\boldsymbol{Z}(0)$ having the stationary distribution $\boldsymbol{\pi}$. These expectations uniquely determine finite measures ν_i on $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$, where $\mathcal{B}(\mathbb{R}_+)$ is the Borel σ-field on $\mathbb{R}_+$. They are called *boundary measures*. Let V_i denote a random variable with the probability distribution $\pi_i^{(0)} := \nu_{3-i}/\nu_{3-i}(\mathbb{R}_+)$.

The stationary equation (3.1) uniquely determines the stationary distribution $\boldsymbol{\pi}$ if it holds for a sufficiently large class of functions f. For this, we may choose a class of exponential functions $f(\boldsymbol{x}) = \mathrm{e}^{\langle \boldsymbol{\theta}, \boldsymbol{x}\rangle}$ on $\mathbb{R}_+^2$ for each $\boldsymbol{\theta} := (\theta_1, \theta_2) \le \boldsymbol{0}$, where $\langle \boldsymbol{a}, \boldsymbol{b}\rangle$ denotes the inner product of vectors $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^2$. Let $\boldsymbol{\delta} = (\delta_1, \delta_2)$, and let

$$\begin{aligned}\varphi(\boldsymbol{\theta}) &= \mathbb{E}[\mathrm{e}^{\langle \boldsymbol{\theta}, \boldsymbol{Z}\rangle}], \qquad \varphi_{3-i}(\theta_i) = \mathbb{E}_{3-i}[\mathrm{e}^{\theta_i Z_i}], \quad i = 1, 2,\\ \widehat{F}(\boldsymbol{\theta}) &= \mathbb{E}[\mathrm{e}^{\langle \boldsymbol{\theta}, \boldsymbol{J}\rangle}], \qquad \kappa(\boldsymbol{\theta}) = \langle \boldsymbol{\delta}, \boldsymbol{\theta}\rangle - \lambda(\widehat{F}(\boldsymbol{\theta}) - 1).\end{aligned}$$

Here $-\kappa(\boldsymbol{\theta})$ is the *Lévy component* of $\boldsymbol{X}(t)$. Then (3.1) is expressible as

$$\kappa(\boldsymbol{\theta})\varphi(\boldsymbol{\theta}) = (\theta_1 - p_{12}\theta_2)\varphi_1(\theta_2) + (\theta_2 - p_{21}\theta_1)\varphi_2(\theta_1), \tag{3.2}$$

as long as $\varphi(\boldsymbol{\theta})$, $\widehat{F}(\boldsymbol{\theta})$, and $\varphi_i(\theta_i)$ are finite. Clearly, (3.2) is always valid for $\boldsymbol{\theta} \le \boldsymbol{0}$. This kind of equation can sometimes be solved explicitly by Wiener–Hopf factorization in the case of independent input streams (see, e.g. [3]).

For computational convenience, we first find $\nu_i(\{0\})$ and $\nu_i(\mathbb{R}_+)$. Clearly, $\nu_i(\{0\}) = \varphi_i(-\infty)$ and $\nu_i(\mathbb{R}_+) = \varphi_i(0)$. Denote the respective traffic intensities at nodes 1 and 2 by

$$\rho_1 = \frac{\alpha_1 + \alpha_2 p_{21}}{\mu_1(1 - p_{12}p_{21})}, \qquad \rho_2 = \frac{\alpha_2 + \alpha_1 p_{12}}{\mu_2(1 - p_{12}p_{21})}.$$

Lemma 3.1. *Under the stability condition* (2.6), *for* $i = 1, 2$,

$$\varphi_i(-\infty) = \mu_i \boldsymbol{\pi}(\mathbf{0}), \qquad \varphi_i(0) = \frac{\Delta_i + \Delta_{3-i} p_{(3-i)i}}{1 - p_{12} p_{21}} = \mu_i(1 - \rho_i). \tag{3.3}$$

In particular, $\boldsymbol{\pi}(\mathbf{0}) = \mathbb{P}\{\mathbf{Z} = \mathbf{0}\} < \min\{1 - \rho_1, 1 - \rho_2\}$, *and, for* $i = 1, 2$,

$$\mathbb{P}\{V_i > x\} = \frac{\delta_{3-i}}{\mu_{3-i}(1 - \rho_{3-i})} \mathbb{P}\{Z_i > x,\ Z_{3-i} = 0\}, \qquad x \geq 0.$$

We defer proof of this lemma to Appendix A. From the equations in (3.3), it is easy to see that $(\varphi_1(0), \varphi_2(0)) = \boldsymbol{\mu} - \boldsymbol{R}^{-1}\boldsymbol{\alpha} > \mathbf{0}$. Also, the second equality for $\varphi_i(0)$ yields another representation for Δ_i, namely,

$$\Delta_1 = \mu_1(1 - \rho_1) - \mu_2 p_{21}(1 - \rho_2), \qquad \Delta_2 = \mu_2(1 - \rho_2) - \mu_1 p_{12}(1 - \rho_1).$$

For each nonzero $\boldsymbol{c} \geq \mathbf{0}$, we now consider the distribution of $c_1 Z_1 + c_2 Z_2$, whose moment generating function is $\varphi(s\boldsymbol{c})$. It is generally hard to find this distribution, so we aim to find its tail asymptotics, namely, the asymptotics of $\mathbb{P}\{c_1 Z_1 + c_2 Z_2 > x\}$ as $x \to \infty$. To this end, we extract the moment generating function $\varphi(s\boldsymbol{c})$ from the stationary equation (3.2); specifically, we express $\varphi(s\boldsymbol{c})$ as a linear combination of the moment generating functions of certain measures, which may include the unknown boundary measures ν_1 and ν_2. This may be viewed as a distributional decomposition that is very useful in finding the asymptotics of $\mathbb{P}\{c_1 Z_1 + c_2 Z_2 > x\}$ when the jump size distributions are heavy tailed.

These arguments may look similar to those used in deriving the Pollaczek–Khinchine formula from the stationary equation of the $\mathrm{M/G/1}$ queue. However, there is a crucial difference arising from the boundary of the state space when the reflecting process is two dimensional. This is in marked contrast to the $\mathrm{M/G/1}$ queue for which the boundary is a single point: this is what facilitates the simple analysis there.

Set $\boldsymbol{\theta} = s\boldsymbol{c}$ in the stationary equation (3.2). Then

$$\begin{aligned}(c_1\delta_1 + c_2\delta_2 - (c_1\alpha_1 + c_2\alpha_2)\widehat{F}^I_{\boldsymbol{c}}(s))\varphi(s\boldsymbol{c}) \\ = (c_1 - p_{12}c_2)\varphi_1(c_2 s) + (c_2 - p_{21}c_1)\varphi_2(c_1 s),\end{aligned} \tag{3.4}$$

where $\widehat{F}^I_{\boldsymbol{c}}$ is the moment generating function of the distribution $F^I_{\boldsymbol{c}}$ defined in Section 2. To single out $\varphi(s\boldsymbol{c})$ in a tractable form, consider its coefficient in (3.4). For positive $r < 1$, let

$$\widehat{S}^{I(r)}_{\boldsymbol{c}}(s) = \frac{1 - r}{1 - r\widehat{F}^I_{\boldsymbol{c}}(s)},$$

which is the moment generating function of the geometric sum with parameter r of i.i.d. random variables having distribution $F^I_{\boldsymbol{c}}$. Let $S^{I(r)}_{\boldsymbol{c}}$ denote a random variable that has the moment generating function $\widehat{S}^{I(r)}_{\boldsymbol{c}}(s)$.

From the strong stability assumption (2.7),

$$c_2\delta_1 + c_2\delta_2 - (c_1\alpha_1 + c_2\alpha_2) = c_1\Delta_1 + c_2\Delta_2 > 0,$$

and, therefore, $r_{\boldsymbol{c}} := (c_1\alpha_1 + c_2\alpha_2)/(c_1\delta_1 + c_2\delta_2) < 1$. Then, the results in Lemma 3.2 below follow directly from (3.4) since

$$\frac{1}{1 - r_{\boldsymbol{c}}} = \frac{c_2\delta_1 + c_2\delta_2}{c_2\Delta_1 + c_2\Delta_2}.$$

Lemma 3.2. *For each nonzero vector $\boldsymbol{c} \geq \mathbf{0}$,*

$$\varphi(s\boldsymbol{c}) = \frac{(c_1 - p_{12}c_2)\varphi_1(c_2 s) + (c_2 - p_{21}c_1)\varphi_2(c_1 s)}{c_2\Delta_1 + c_2\Delta_2} \widehat{S}_{\boldsymbol{c}}^{I(r_{\boldsymbol{c}})}(s), \tag{3.5}$$

and, for $B \in \mathscr{B}(\mathbb{R}_+)$, $\mathbb{P}\{c_1Z_1 + c_2Z_2 \in B\}$ equals

$$\eta_{\boldsymbol{c}}^{(1)}\mu_1(1-\rho_1)\mathbb{P}\{c_2V_2 + S_{\boldsymbol{c}}^{I(r_{\boldsymbol{c}})} \in B\} + \eta_{\boldsymbol{c}}^{(2)}\mu_2(1-\rho_2)\mathbb{P}\{c_1V_1 + S_{\boldsymbol{c}}^{I(r_{\boldsymbol{c}})} \in B\}, \tag{3.6}$$

where V_1 and V_2 are independent of $S_{\boldsymbol{c}}^{I(r_{\boldsymbol{c}})}$, and

$$\eta_{\boldsymbol{c}}^{(1)} = \frac{c_1 - p_{12}c_2}{c_1\Delta_1 + c_2\Delta_2}, \qquad \eta_{\boldsymbol{c}}^{(2)} = \frac{c_2 - p_{21}c_1}{c_1\Delta_1 + c_2\Delta_2}. \tag{3.7}$$

4. Stochastic bounds and weak tail equivalence

In this section we consider upper and lower bounds for $\mathbb{P}\{c_1Z_1 + c_2Z_2 > x\}$ for $\boldsymbol{c} \geq \mathbf{0}$ with $\boldsymbol{c} \neq \mathbf{0}$, provided (A1) holds. For this, we use (3.6), but we must take care with the signs of the coefficients $c_1 - p_{12}c_2$ and $c_2 - p_{21}c_1$. If both signs are negative then $c_1(1 - p_{12}p_{21}) \leq 0$ and $c_2(1 - p_{12}p_{21}) \leq 0$, which contradicts (2.1) and $\boldsymbol{c} \neq \mathbf{0}$. Hence, under (2.7), only the following three cases are possible.

(C0) $c_1 - p_{12}c_2 \geq 0$ and $c_2 - p_{21}c_1 \geq 0$ (in this case, we must have $\boldsymbol{c} > \mathbf{0}$).

(C1) $c_1 - p_{12}c_2 \geq 0$ and $c_2 - p_{21}c_1 < 0$.

(C2) $c_1 - p_{12}c_2 < 0$ and $c_2 - p_{21}c_1 \geq 0$.

Since (C1) and (C2) are symmetric, it is enough to consider only (C0) and (C1).

Since $\eta_{\boldsymbol{c}}^{(1)}\mu_1(1-\rho_1) + \eta_{\boldsymbol{c}}^{(2)}\mu_2(1-\rho_2) = 1$ from (3.6) with $B = \mathbb{R}_+$ and $\eta_{\boldsymbol{c}}^{(i)}$ for $i = 1, 2$ are positive for (C0), we obtain the following lower bound.

Lemma 4.1. *When (C0) holds,*

$$\mathbb{P}\{S_{\boldsymbol{c}}^{I(r_{\boldsymbol{c}})} > x\} \leq \mathbb{P}\{c_1Z_1 + c_2Z_2 > x\} \quad \textit{for all } x > 0.$$

When (C1) holds, we can no longer use (3.6) to obtain a lower bound. Instead we use the following representation (it is proved in Appendix B):

$$\varphi(s\boldsymbol{c}) = \varphi^+(s\boldsymbol{c}) + \frac{\varphi_2(c_1 s)}{\delta_2} + \frac{\varphi_1(c_2 s)}{\delta_1} - d_0. \tag{4.1}$$

Here $\varphi^+(\boldsymbol{\theta}) = \mathbb{E}[\mathrm{e}^{\langle\boldsymbol{\theta},\boldsymbol{Z}\rangle}1_{\{\boldsymbol{Z}>\mathbf{0}\}}]$ and $d_0 = \mu_1\mu_2(1 - p_{12}p_{21})\boldsymbol{\pi}(\mathbf{0})/\delta_1\delta_2$.

Lemma 4.2. *For the case (C1), let $r'_{\boldsymbol{c}} = (c_1\alpha_1 + c_2\alpha_2)/c_1(\delta_1 + \delta_2 p_{21})$; then $0 < r'_{\boldsymbol{c}} < 1$ and*

$$\varphi(s\boldsymbol{c}) = \left[\frac{d_{\boldsymbol{c}}^{(1)}}{\delta_1}\varphi_1(c_2 s) + d_{\boldsymbol{c}}^{(2)}(\varphi^+(s\boldsymbol{c}) - d_0)\right]\widehat{S}_{\boldsymbol{c}}^{I(r'_{\boldsymbol{c}})}(s), \tag{4.2}$$

where

$$d_{\boldsymbol{c}}^{(1)} = \frac{\delta_1(c_1 - p_{12}c_2) + \delta_2(p_{21}c_1 - c_2)}{c_1(\delta_1 + \delta_2 p_{21})(1 - r'_{\boldsymbol{c}})}, \qquad d_{\boldsymbol{c}}^{(2)} = \frac{\delta_2(p_{21}c_1 - c_2)}{c_1(\delta_1 + \delta_2 p_{21})(1 - r'_{\boldsymbol{c}})}.$$

Therefore, for $x \geq 0$,

$$\mathbb{P}\{S_{\boldsymbol{c}}^{I(r'_{\boldsymbol{c}})} > x\} \leq \mathbb{P}\{c_1Z_1 + c_2Z_2 > x\}. \tag{4.3}$$

Proof. Multiply (4.1) by $(p_{21}c_1 - c_2)\delta_2$ and add to (3.5). This yields

$$[c_1(\delta_1 + \delta_2 p_{21}) - (c_1\alpha_1 + c_2\alpha_2)\widehat{F}^I_{\boldsymbol{c}}(s)]\varphi(s\boldsymbol{c}) = \frac{1}{\delta_1}[\delta_1(c_1 - p_{12}c_2) + \delta_2(p_{21}c_1 - c_2)]\varphi_1(c_2 s) + \delta_2(p_{21}c_1 - c_2)(\varphi^+(s\boldsymbol{c}) - d_0)$$

and, hence, (4.2) because

$$c_1(\delta_1 + \delta_2 p_{21}) - (c_1\alpha_1 + c_2\alpha_2) = c_1(\Delta_1 + \Delta_2 p_{21}) + \alpha_2(c_1 p_{21} - c_2) > 0.$$

Since $d^{(1)}_{\boldsymbol{c}} > 0$ and $d^{(2)}_{\boldsymbol{c}} > 0$, the right-hand side of (4.2) represents the convolution of two distributions on $[0, \infty)$, and this leads to (4.3).

We continue to assume that (A1) holds, and consider the tail probability $\mathbb{P}\{c_1 Z_1 + c_2 Z_2 > x\}$ for directional vectors $\boldsymbol{c} \geq \boldsymbol{0}$. First we obtain results for $\boldsymbol{c} = (1, 0)$ but under weaker assumptions.

Lemma 4.3. *Assume that (A1) holds, that the system is stable, and that $\Delta_1 > 0$. Then, for $x > 0$,*

$$\mathbb{P}\{S_1^{I(r_1')} > x\} \leq \mathbb{P}\{Z_1 > x\} \leq \mathbb{P}\{S_1^{I(r_1)} > x\}, \tag{4.4}$$

where $r_1 = \alpha_1/\delta_1$ and $r_1' = \alpha_1/(\delta_1 + \delta_2 p_{21})$.

Proof. The upper bound follows directly from Lemma 2.1 because $\widetilde{Z}_1$ is subject to the stationary workload distribution of the M/G/1 queue (it can also be obtained analytically from (3.5)). We obtained the lower bound earlier in Lemma 4.2.

Remark 4.1. Clearly, $\Delta_1 > 0$ and $\mu_2 - \mu_1 p_{12} = \delta_2 > 0$ imply that

$$\frac{\alpha_1}{\mu_1} < r_1' = \frac{\alpha_1}{\mu_1(1 - p_{12}p_{21})} < r_1 = \frac{\alpha_1}{\mu_1 - \mu_2 p_{21}} < 1.$$

By arguments similar to those used in the proof of Lemma 2.1, $\mathbb{P}\{S_1^{I(p_1)} > x\}$ with $p_1 = \alpha_1/\mu_1$ also provides a lower bound, but the lower bound in (4.4) is tighter because $p_1 < r_1'$.

We next consider the case of nonzero $\boldsymbol{c} \geq \boldsymbol{0}$ for options (C0) and (C1).

Lemma 4.4. *Assume that (A1) holds, let the strong stability condition (2.7) hold, and recall the definition of $\eta^{(i)}_{\boldsymbol{c}}$ given in (3.7). Let $x \geq 0$ and $\boldsymbol{c} \geq \boldsymbol{0}$, $\boldsymbol{c} \neq \boldsymbol{0}$. For the case (C0),*

$$\mathbb{P}\{S_{\boldsymbol{c}}^{I(r_{\boldsymbol{c}})} > x\} \leq \mathbb{P}\{c_1 Z_1 + c_2 Z_2 > x\} \leq \delta_1 \eta^{(1)}_{\boldsymbol{c}} \mathbb{P}\{c_2 S_2^{I(r_2)} + S_{\boldsymbol{c}}^{I(r_{\boldsymbol{c}})} > x\} + \delta_2 \eta^{(2)}_{\boldsymbol{c}} \mathbb{P}\{c_1 S_1^{I(r_1)} + S_{\boldsymbol{c}}^{I(r_{\boldsymbol{c}})} > x\}. \tag{4.5}$$

In the case (C1),

$$\mathbb{P}\{S_{\boldsymbol{c}}^{I(r_{\boldsymbol{c}}')} > x\} \leq \mathbb{P}\{c_1 Z_1 + c_2 Z_2 > x\} \leq \delta_1 \eta^{(1)}_{\boldsymbol{c}} \mathbb{P}\{c_2 S_2^{I(r_2)} + S_{\boldsymbol{c}}^{I(r_{\boldsymbol{c}})} > x\}. \tag{4.6}$$

In both (4.5) and (4.6) the random variables $S_1^{I(r_1)}$, $S_2^{I(r_2)}$, and $S_{\boldsymbol{c}}^{I(r_{\boldsymbol{c}})}$ are mutually independent.

Remark 4.2. For the case (C0), $\boldsymbol{c} > \boldsymbol{0}$, so (4.5) does not contradict (4.4).

Proof of Lemma 4.4. In the case (C0), by Lemma 3.1, Lemma 4.3, and its symmetric version,

$$\varphi_{3-i}(0)\mathbb{P}\{V_i > x\} \leq \delta_{3-i}\mathbb{P}\{Z_i > x\} \leq \delta_{3-i}\mathbb{P}\{S_i^{I(r_i)} > x\}, \qquad i = 1, 2.$$

Hence, Lemma 3.2 yields the upper bound of (4.5), and its lower bound is obtained by Lemma 4.1. In the case (C1), the upper bound of (4.6) is immediate from Lemma 3.2, while we obtained the lower bound earlier in Lemma 4.2.

https://doi.org/10.1239/jap/1417528479 Published online by Cambridge University Press

4.1. Proof of Theorem 2.1

By property (P5) in Appendix D, as $x \to \infty$,

$$\mathbb{P}\{S_1^{I(r_1)} > x\} \sim \frac{r_1}{1-r_1}\overline{F}{}_1^I(x), \qquad \mathbb{P}\{S_1^{I(r_1')} > x\} \sim \frac{r_1'}{1-r_1'}\overline{F}{}_1^I(x).$$

These relations with Lemma 4.3 yield (2.10).

Assume now that all three distributions F_1^I, F_2^I, and $F_{\boldsymbol{c}}^I$ are subexponential. Then

$$\mathbb{P}\{c_i S_i^{I(r_i)} > x\} \sim \frac{r_i}{1-r_i}\overline{F}{}_i^I\left(\frac{x}{c_i}\right), \quad i = 1, 2, \qquad \mathbb{P}\{S_{\boldsymbol{c}}^{I(r_{\boldsymbol{c}})} > x\} \sim \frac{r_{\boldsymbol{c}}}{1-r_{\boldsymbol{c}}}\overline{F}{}_{\boldsymbol{c}}^I(x).$$

Similar asymptotic equivalence relations hold on replacing $r_{\boldsymbol{c}}$ by either $r_{\boldsymbol{c}}'$ or $r_{\boldsymbol{c}}''$, where $r_{\boldsymbol{c}}'' = (c_1\alpha_1 + c_2\alpha_2)/c_2(\delta_2 + \delta_1 p_{12})$. Clearly,

$$\overline{F}{}_{\boldsymbol{c}}^I(x) = \frac{1}{m_{\boldsymbol{c}}}\int_x^\infty \mathbb{P}\{c_1J_1 + c_2J_2 > y\}\,\mathrm{d}y \ge \frac{c_1}{m_{\boldsymbol{c}}}\int_{x/c_1}^\infty \mathbb{P}\{J_1 > y\}\,\mathrm{d}y = \frac{c_1m_1}{m_{\boldsymbol{c}}}\overline{F}{}_1^I\left(\frac{x}{c_1}\right),$$

so

$$\mathbb{P}\{c_1S_1^{I(r_1)} > x\} \le \frac{[1+o(1)]m_{\boldsymbol{c}}r_1}{(1-r_1)c_1m_1}\overline{F}{}_{\boldsymbol{c}}^I(x) \le \frac{[1+o(1)](1-r_{\boldsymbol{c}})m_{\boldsymbol{c}}r_1}{r_{\boldsymbol{c}}(1-r_1)c_1m_1}\mathbb{P}\{S_{\boldsymbol{c}}^{I(r_{\boldsymbol{c}})} > x\}$$

and, finally,

$$\mathbb{P}\{c_1S_1^{I(r_1)} + S_{\boldsymbol{c}}^{I(r_{\boldsymbol{c}})} > x\} \le [1+o(1)]K\overline{F}{}_{\boldsymbol{c}}^I(x).$$

Here the constant $K := (1-r_{\boldsymbol{c}})m_{\boldsymbol{c}}r_1/r_{\boldsymbol{c}}(1-r_1)c_1m_1 + r_{\boldsymbol{c}}/(1-r_{\boldsymbol{c}})$ is positive and finite. These observations and Lemma 4.4 imply (2.11).

5. Proof of Theorem 2.2

Throughout this section, assume that (A2) and the strong stability condition (2.7) hold. Based on the fluid dynamics considered in Appendix C, we provide a lower bound for the tail probabilities assuming only that the integrated tail distributions F_1^I and F_2^I are *long tailed* (see Appendix D for the definition). We derive the exact asymptotics in the case of subexponential distributions.

5.1. Lower bounds

Assume that the system runs in the discrete-time stationary regime starting from time $-\infty$, and that (Z_1, Z_2) is the workload vector observed at time 0 when the 0th batch arrives. In what follows, $\mathrm{LB}(x)$ denotes a lower bound on the probability $\mathbb{P}\{c_1Z_1 + c_2Z_2 > x\}$. Represent the random variables (J_{1n}, J_{2n}) as

$$(J_{1n}, J_{2n}) = \nu_n(\sigma_{1n}, 0) + (1-\nu_n)(0, \sigma_{2n}),$$

where the random variables on the right-hand side are mutually independent, and also independent for distinct n, $\mathbb{P}\{\nu_n = 1\} = 1 - \mathbb{P}\{\nu_n = 0\} = p_1$, and the random variables $\{\sigma_{in}\}$ have distribution F_i. In other words, $\mathbb{P}\{\sigma_{in} \in \cdot\} = \mathbb{P}\{J_{in} \in \cdot \mid J_{in} > 0\}$, where $\mathbb{P}\{J_{in} > 0\} = p_i$. Then $\mathbb{E}[\sigma_{in}] = m_i/p_i$ and the integrated distribution of σ_{in} coincides with F_i^I.

Lemma 5.1. *Let the distributions F_1^I and F_2^I be long tailed. Then, for $c_1, c_2 \geq 0$ with $c_1 + c_2 > 0$,*

$$\mathrm{LB}(x) = [1 + o(1)]\frac{\alpha_1}{\Delta_1 + p_{21}\Delta_2}\overline{F}{}_1^I\left(\frac{x}{c_1}\right) + [1 + o(1)]\frac{\alpha_2}{\Delta_2 + p_{12}\Delta_1}\overline{F}{}_2^I\left(\frac{x}{c_2}\right), \tag{5.1}$$

where, by convention, $x/0 = \infty$ and $\overline{F}{}_i^I(\infty) = 0$ (recall that $\alpha_i = p_i m_i / a$ for $i = 1, 2$).

Remark 5.1. In particular, for $\boldsymbol{c} = (1, 0)$,

$$\frac{r_1'}{1 - r_1'} = \frac{\alpha_1}{\mu_1(1 - p_{12}p_{21}) - \alpha_1} < \frac{\alpha_1}{\Delta_1 + p_{21}\Delta_2} < \frac{\alpha_1}{\Delta_1}.$$

Hence, lower bound (5.1) is tighter than that in (4.4).

Proof of Lemma 5.1. First, apply to our model the arguments from Lemma 2 of [5] (see also Theorem 14 of [1]) to conclude that the lower bound asymptotics in our model are equivalent to those in an auxiliary model with deterministic arrival times $-na$, $n = 1, 2, \ldots$. Second, we can use the strong law of large numbers for the i.i.d. jump sizes and Corollary C.1 to conclude that, for any $\varepsilon > 0$ and $x \to \infty$,

$$\begin{aligned}\mathrm{LB}(x) = {} & [1 + o(1)]p_1\mathbb{P}\Bigg(\bigcup_{n\geq 1}\bigg\{\sigma_{1,-n} > \frac{x}{c_1} + na(\Delta_1 + p_{21}\Delta_2 + \varepsilon)\bigg\}\Bigg) \\ & + [1 + o(1)]p_2\mathbb{P}\Bigg(\bigcup_{n\geq 1}\bigg\{\sigma_{2,-n} > \frac{x}{c_2} + na(\Delta_2 + p_{12}\Delta_1 + \varepsilon)\bigg\}\Bigg).\end{aligned} \tag{5.2}$$

Apply property (P7) from Appendix D to each of the probabilities on the right-hand side of (5.2). Then letting $\varepsilon \downarrow 0$ leads to (5.1).

5.2. Weak tail equivalence

It is known that, for a single-server queue with subexponential service time distributions, the stationary workload is large due to a single large service time (see property (P7) in Appendix D). For a single-server queue, there is no difference in having a single customer with service time J and a batch of customers whose total service time is of size J. Then, from the upper bound in Lemma 2.1 and lower bound (5.1) with $\boldsymbol{c} = (1, 0)$, $\mathbb{P}\{Z_i > x\}$ for stationary Z_i is weakly tail equivalent to $\overline{F}{}_i^I(x)$. A similar result holds for the linear sum $c_1 Z_1 + c_2 Z_2$ if we assume the tails $\overline{F}_1(x)$ and $\overline{F}_2(x)$ to be weakly equivalent. But here we can get more.

5.3. Exact asymptotics

We first consider the exact tail asymptotics for $\boldsymbol{c} = (1, 0)$, applying the 'squeeze principle' (see Theorem 8 of [1]). Thus, we focus on the tail asymptotics for Z_1. We do this in the following two steps.

Step 1. Let the distribution F_1^I be subexponential. Consider the upper bound random variable $\widetilde{Z}_1$ introduced in Lemma 2.1 (it is a one-dimensional stationary workload random variable). It is easy to see that the model does not change if we assume, analogously to the discussion before that lemma, that there is no fluid input from node 2 and the release-cum-service rate for node 1 is $\widetilde{\mu}_1 = \mu_1 - p_{21}\mu_2$. Furthermore, we can rescale time by assuming a unit service rate and that the interarrival times to node 1 are i.i.d. with mean $b_1 = a/\widetilde{\mu}_1 p_1$ (this comes from the geometric argument). We can then speak about a single-server queue with customer n having service time σ_{1n} (instead of a batch of size σ_{1n}). Consider again the discrete-time stationary regime, and let $T_{1,0} = 0 > T_{1,-1} > T_{1,-2} > \cdots$ denote the arrival times of customers $0, -1, -2, \dots$, with i.i.d. interarrival times $t_{1,-n} = T_{1,-n+1} - T_{1,-n}$ and $\mathbb{E}[t_{1,-n}] = b_1$ for $n = 1, 2, \dots$. Then

$$\widetilde{Z}_1 = \widetilde{Z}_1(0) = \max\left\{0, \sup_{n\geq 1} \sum_{i=-n}^{-1} (\sigma_{1i} - t_{1i})\right\} + \sigma_{10}.$$

We now follow the lines of the appendix of [1]. For any $\varepsilon \in (0, b_1 - m_1/p_1)$, consider another stable single-server queue with service times σ_{1n} and constant interarrival times $b_1 - \varepsilon$. Denote the stationary workload in that queue by

$$\widehat{Z}_1 = \max\left\{0, \sup_{n\geq 1} \sum_{i=-n}^{-1} (\sigma_{1i} - (b_1 - \varepsilon)i)\right\} + \sigma_{10}.$$

Then $\widetilde{Z}_1 \leq \widehat{Z}_1 + M_{1\varepsilon}$, where $M_{1\varepsilon} := \sup_n \{\sum_{i=-n}^{-1} (b_1 - \varepsilon - t_{1i})\}_+$ does not depend on $\widehat{Z}_1$ and has a light-tailed distribution.

Since the distribution F_1^I is subexponential, properties (P3), (P6), and (P7) in Appendix D imply that

$$\begin{aligned}\mathbb{P}\{\widehat{Z}_1 + M_{1\varepsilon} > x\} &\sim \mathbb{P}\{\widehat{Z}_1 > x\}\\ &\sim \mathbb{P}\left\{\sup_{n\geq 1} \sum_{i=-n}^{-1} (\sigma_{1i} - (b_1 - \varepsilon)i) > x\right\}\\ &\sim \sum_{n>0} \mathbb{P}\{\sigma_{1,-n} > (b_1 - \varepsilon)n + x\},\end{aligned}$$

where the right-hand side is of the same order as the lower bound $\mathrm{LB}(x)$. Since $Z_1 \leq \widetilde{Z}_1 \leq \widehat{Z}_1 + M_{1\varepsilon}$ a.s., we can write

$$\begin{aligned}\mathbb{P}(Z_1 > x) &= \mathbb{P}\{Z_1 > x,\ \widehat{Z}_1 + M_{1,\varepsilon} > x\}\\ &\sim \sum_n \mathbb{P}\{Z_1 > x,\ \sigma_{1,-n} > (b_1 - \varepsilon)n + x\}. \end{aligned} \tag{5.3}$$

Step 2. If there is only one big jump before time 0 and all other jump sizes are replaced by their means, then it follows from Lemma C.1 that $Z_1 > x$ can occur only if, for some n, the $(-n)$th service time is large. Note that $b_1 < a(\Delta_1 + p_{21}\Delta_2)$. Applying to (5.3) arguments similar to those used in the proof of Theorem 8 of [1], for any γ in $(0, \Delta_1 + p_{21}\Delta_2)$, a single big jump of size $\sigma_{1,-n} < x + an(\Delta_1 + p_{21}\Delta_2 - \gamma)$ is not sufficient for the inequality $Z_1 > x$ to hold, so, for $x \to \infty$, we must have

$$\mathbb{P}\{Z_1 > x) \leq [1 + o(1)] \sum_{n=0}^{\infty} \mathbb{P}\{Z_1 > x,\ \sigma_{1,-n} > x + an(\Delta_1 + p_{21}\Delta_2 - \gamma)\}.$$

Letting $\gamma \to 0$ leads to an upper bound for $\mathbb{P}\{Z_1 > x\}$ that, up to the term $o(1)$, coincides with the lower bound.

By symmetry, a similar result holds for $\boldsymbol{c} = (0, 1)$.

Assume finally that $c_1 > 0$ and $c_2 > 0$. Let $F_1^I(x/c_1)$ and $F_2^I(x/c_2)$ be weakly tail equivalent. Introduce by analogy b_2, $\widehat{Z}_2$, and $M_{2\varepsilon}$. Then, with $M_\varepsilon := c_1 M_{1\varepsilon} + c_2 M_{2\varepsilon}$,

$$c_1 Z_1 + c_2 Z_2 \leq c_1 \widehat{Z}_1 + c_2 \widehat{Z}_2 + M_\varepsilon,$$

where on the right-hand side the three terms are mutually independent, the first two having heavy-tailed distributions, and the last a light-tailed distribution. Then, from condition (3.32) of Theorem 3.33 of [7], the distribution of $c_1\widehat{Z}_1 + c_2\widehat{Z}_2$ is also subexponential and, furthermore,

$$\mathbb{P}\{c_1\widehat{Z}_1 + c_2\widehat{Z}_2 + M_\varepsilon > x\} \sim \mathbb{P}\bigg\{\widehat{Z}_1 > \frac{x}{c_1}\bigg\} + \mathbb{P}\bigg\{\widehat{Z}_2 > \frac{x}{c_2}\bigg\}.$$

Then, similarly to (5.3), we obtain

$$\mathbb{P}\{c_1 Z_1 + c_2 Z_2 > x\} \sim \mathbb{P}\bigg\{Z_1 > \frac{x}{c_1},\ \widehat{Z}_1 > \frac{x}{c_1}\bigg\} + \mathbb{P}\bigg\{Z_2 > \frac{x}{c_2},\ \widehat{Z}_2 > \frac{x}{c_2}\bigg\}, \tag{5.4}$$

and, as before, we can proceed with each term in (5.4) separately to obtain the second statement of Theorem 2.2.

6. Concluding remarks and related open problems

We have obtained a number of results for the tail asymptotics of linear functionals $c_1Z_1+c_2Z_2$ of a stationary two-dimensional workload using two approaches, under different stochastic assumptions. The open questions below are closely related to the results in this paper.

(Q1) The exact asymptotics for $\mathbb{P}\{Z_i > x\}$ are obtained in Theorem 2.2 only when arrival batches are one-dimensional. Can they be obtained in the two dimensional case? Since $\mathbb{P}\{Z_i > x\}$ is weakly tail equivalent to $\overline{F}{}_i^I(x)$, we may conjecture that $\mathbb{P}\{Z_i > x\} \sim K\overline{F}{}_i^I(x)$ for some $K > 0$ under (A1) and subexponentiality of F_i^I. If this is correct, it follows from (3.4) that $\mathbb{P}\{V_i > x\} \sim K'\overline{F}{}_i^I(x)$ for some $K' > 0$. For example, this is satisfied for $i = 1$ since (3.4) with $\boldsymbol{c} = (1, 0)$ implies that, for $x \geq 0$,

$$\alpha_1\mathbb{P}\{J_1^I + Z_1 > x\} - \delta_1\mathbb{P}\{Z_1 > x\} = p_{21}\varphi_2(0)\,\mathbb{P}\{V_1 > x\},$$

where J_1^I is a random variable with distribution F_1^I and independent of Z_1. Once the exact asymptotics for $\mathbb{P}\{V_i > x\}$ are known, we can derive exact asymptotics for $\mathbb{P}\{c_1Z_1 + c_2Z_2 > x\}$ for $\boldsymbol{c} > \boldsymbol{0}$ from (3.6) under the assumptions in Theorem 2.1.

(Q2) Might it be possible to obtain the weak tail asymptotics of Theorem 2.1 using the sample-path approach? It is unclear how to construct a majorant, while a lower bound may be easily given. For example, one may introduce an auxiliary model with smaller batch sizes $(\widehat{J}_1, \widehat{J}_2) = \alpha(J_1, 0) + (1-\alpha)(0, J_2)$, where α is an independent random variable, $\mathbb{P}\{\alpha = 1\} = 1 - \mathbb{P}\{\alpha = 0\} = \frac{1}{2}$, and compare the two models.

(Q3) Find the tail asymptotics for the two-dimensional stationary vector (Z_1, Z_2) or for functionals of that vector which are not linear.

(Q4) Find the tail asymptotics for the stationary sojourn time in a stable generalized Jackson network with heavy-tailed service time distributions. Our asymptotic results provide only lower bounds for that, since the arrival of a large batch to one of the nodes does not delay service in the other node, while this is the case for a single customer with a large service time.

Appendix A. Proof of Lemma 3.1

Dividing (3.2) by θ_1,

$$\bigg(\delta_1 + \delta_2\frac{\theta_2}{\theta_1} - \frac{\lambda}{\theta_1}(\widehat{F}(\boldsymbol{\theta}) - 1)\bigg)\varphi(\boldsymbol{\theta}) = \bigg(1 - p_{12}\frac{\theta_2}{\theta_1}\bigg)\varphi_1(\theta_2) + \bigg(\frac{\theta_2}{\theta_1} - p_{21}\bigg)\varphi_2(\theta_1).$$

https://doi.org/10.1239/jap/1417528479 Published online by Cambridge University Press

Letting $\theta_1 \to -\infty$ in this equation yields

$$\varphi_1(\theta_2) = \delta_1\varphi(-\infty, \theta_2) + p_{21}\varphi_2(-\infty), \tag{A.1}$$

implying that

$$\varphi_1(-\infty) = \delta_1\varphi(-\infty, -\infty) + p_{21}\varphi_2(-\infty).$$

Similarly,

$$\varphi_2(-\infty) = \delta_2\varphi(-\infty, -\infty) + p_{12}\varphi_1(-\infty).$$

Since $\varphi(-\infty, -\infty) = \boldsymbol{\pi}(\mathbf{0})$, solving these equations yields $\varphi_i(-\infty) = \mu_i\boldsymbol{\pi}(\mathbf{0})$. On the other hand, putting $\theta_2 = 0$,

$$\Delta_1 + p_{21}\varphi_2(0) = \varphi_1(0), \qquad \Delta_2 + p_{12}\varphi_1(0) = \varphi_2(0).$$

Solving these equations yields the first equality of (3.3) for $\varphi_i(0)$. The second equality is immediate from the definitions of Δ_i and ρ_i.

Appendix B. Proof of (4.1)

From the definitions of φ and φ^+, we have

$$\varphi(\boldsymbol{\theta}) = \varphi^+(\boldsymbol{\theta}) + \varphi(\theta_1, -\infty) + \varphi(-\infty, \theta_2) + \boldsymbol{\pi}(\mathbf{0}).$$

Substitution in this equation for $\varphi(\theta_1, -\infty)$ and for its symmetric form from (A.1), and then setting $\boldsymbol{\theta} = s\boldsymbol{c}$, yields (4.1) because

$$\delta_1 p_{12}\varphi_1(-\infty) + \delta_2 p_{21}\varphi_2(-\infty) + \delta_1\delta_2\boldsymbol{\pi}(\mathbf{0}) = \mu_1\mu_2(1 - p_{12}p_{21})\boldsymbol{\pi}(\mathbf{0}) = \delta_1\delta_2 d_0.$$

Appendix C. Analysis of a pure fluid model

Assume again, for $i = 1, 2$, that $\Delta_i > 0$, and consider an auxiliary pure fluid model with continuous fluid input rates α_i, service rates μ_i, and transition fractions p_{12} and p_{21} as described in Section 2. We use the same notation $Z_i(t)$ as before, but now for deterministic buffer quantities.

Fix $t > 0$, and assume that the fluid model starts at negative time $-t$ from levels y_i (meaning $Z_i(-t) = y_i$); we want to identify conditions on y_i for

$$c_1 Z_1 + c_2 Z_2 \geq x$$

to hold, where again $c_i \geq 0$ and $c_1 + c_2 = 1$ are given constants, and where $Z_i = Z_i(0)$.

Cases $\boldsymbol{c} = (1, 0)$ *and* $\boldsymbol{c} = (0, 1)$. We want to find conditions for $Z_1 \geq x$ (and then by symmetry for $Z_2 \geq x$). From the monotonicity properties of fluid limits (see, e.g. [10]), under the stability conditions, if the fluid model starts from a nonzero initial value at time $-t$ and if some coordinate, i say, becomes 0, $Z_i(u) = 0$ at time $u > -t$, then it stays at 0, $Z_i(v) = 0$ for all $u \leq v \leq 0$.

Let $L_2 = y_2/\Delta_2$. Suppose first that $L_2 \geq t$. Then at any time instant $u \in (-t, 0)$, and, for $i = 1, 2$,

(i) the input rate to queue i is $\alpha_i + \mu_{3-i}p_{3-i,i}$; and

(ii) the output rate from queue i is μ_i.

https://doi.org/10.1239/jap/1417528479 Published online by Cambridge University Press

Then $Z_1 \geq x$ if and only if

$$y_1 \geq x + t\Delta_1. \tag{C.1}$$

Suppose now that $L_2 < t$. Then, for any $u \in (-t, -t + L_2)$ and $i = 1, 2$,

(iii) the input rate to queue i is $\alpha_i + \mu_{3-i} p_{3-i,i}$; and

(iv) the output rate from queue i is μ_i;

while, for any $u \in (-t + L_2, 0)$,

(v) the input and output rates to/from queue 1 and the input rate to queue 2 are as in (iii) and (iv); but

(vi) the output rate from queue 2 equals the input rate, i.e. it equals $\alpha_2 + \mu_1 p_{12}$.

Then the condition $Z_1 \geq x$ is the same as

$$y_1 - x \geq L_2\Delta_1 + (t - L_2)(\mu_1 - [\alpha_1 + p_{21}(\alpha_2 + \mu_1 p_{12})]) = t\Delta_1 + tp_{21}\Delta_2 - y_2 p_{21}. \tag{C.2}$$

Combining (C.1) and (C.2),

$$y_1 \geq x + t\Delta_1 + p_{21}(t\Delta_2 - y_2)_+.$$

Case $c_1 > 0, c_2 > 0$. Following the same logic as just given, if $L_2 \leq t$ then $Z_2 = 0$, and the condition on y_1 coincides with (C.2) on replacing x by x/c_1, i.e.

$$y_1 \geq \frac{x}{c_1} + t\Delta_1 + p_{21}(\Delta_2 - y_2).$$

Similarly, if $L_1 \leq t$ then $Z_1 = 0$ and

$$y_2 \geq \frac{x}{c_2} + t\Delta_2 + p_{12}(t\Delta_1 - y_1).$$

Otherwise, if both $L_1 > t$ and $L_2 > t$, then $y_1 = Z_1 + t\Delta_1$, $y_2 = Z_2 + t\Delta_2$, and

$$c_1 Z_1 + c_2 Z_2 = c_1(y_1 - t\Delta_1) + c_2(y_2 - t\Delta_2) \geq x.$$

Combining these three cases leads to the following result.

Lemma C.1. *For $i = 1, 2$, let a purely fluid model have input rates α_i, service rates μ_i, and transition fractions $p_{12} > 0$ and $p_{21} > 0$; let $c_i \geq 0$ have $c_1 + c_2 > 0$. Let $t > 0$, and let the system start at time $-t$ from $Z_i(-t) = y_i \geq 0$. If both $\Delta_i > 0$ then the inequality $c_1 Z_1 + c_2 Z_2 \geq x$ holds if and only if*

$$c_1(y_1 - t\Delta_1 - p_{21}(t\Delta_2 - y_2)_+)_+ + c_2(y_2 - t\Delta_2 - p_{12}(t\Delta_1 - y_1)_+)_+ \geq x. \tag{C.3}$$

Corollary C.1. *When $y_1 y_2 = 0$ but $y_1 + y_2 > 0$, (C.3) is equivalent to*

$$\max\{c_1(y_1 - t\Delta_1 - p_{21}t\Delta_2),\ c_2(y_2 - t\Delta_2 - p_{12}t\Delta_1)\} > x,$$

and this last inequality is equivalent to the union of the two events

$$\left\{y_1 > \frac{x}{c_1} + t\Delta_1 + p_{21}t\Delta_2\right\} \cup \left\{y_2 > \frac{x}{c_2} + t\Delta_2 + p_{12}t\Delta_1\right\},$$

where if $c_i = 0$ then the corresponding event is empty.

https://doi.org/10.1239/jap/1417528479 Published online by Cambridge University Press

Appendix D. Heavy-tailed distributions

D.1. Definitions

The distribution F of a positive random variable X is

(D1) *heavy tailed* if $\mathbb{E}[\mathrm{e}^{cX}] := \int_0^\infty \mathrm{e}^{cx}\,\mathrm{d}F(x) = \infty$ for all $c > 0$, and *light tailed* otherwise;

(D2) *long tailed* if $\overline{F}(x) > 0$ for all $x > 0$ and $\overline{F}(x+1)/\overline{F}(x) \to 1$ as $x \to \infty$;

(D3) *subexponential* if, as $x \to \infty$, $\overline{F * F}(x) \sim 2\overline{F}(x)$, or, equivalently, $\mathbb{P}\{X_1 + X_2 > x\} \sim 2\mathbb{P}\{X > x\}$ as $x \to \infty$ (here X_1 and X_2 are two independent copies of X).

(D4) The distribution F of a real-valued random variable X is subexponential if the distribution of $\max\{X, 0\}$ is subexponential.

(D5) The distribution F is *regularly varying* if $\overline{F}(x) = l(x)x^{-k}$, where $k \geq 0$ and the positive function $l(x)$ is *slowly varying* at infinity. Regularly varying distributions are subexponential.

D.2. Key properties

For details, see, e.g. [7].

(P1) Any subexponential distribution is long tailed, and any long-tailed distribution is heavy tailed.

(P2) If distribution F is long tailed then there exists a function $h(x) \to \infty$ as $x \to \infty$ such that $\overline{F}(x + h(x))/\overline{F}(x) \to 1$ as $x \to \infty$.

(P3) If distribution F is long tailed and if $\overline{G}(x) = o(\overline{F}(x))$, then the convolution $F * G$ is also long tailed and $\overline{F * G}(x) \sim \overline{F}(x)$. In particular, if a random variable X has a long-tailed distribution and a random variable Y is nonpositive, then $\mathbb{P}\{X + Y > x\} \sim \mathbb{P}\{X > x\}$ as $x \to \infty$.

(P4) If F is subexponential and $\overline{G}(x) \sim \overline{F}(x)$, then G is subexponential.

(P5) If $X_1, X_2, \ldots$ are i.i.d. with common subexponential distribution F and if τ is a light-tailed counting random variable, then $\sum_{i=1}^{\tau} X_i$ also has a subexponential distribution and $\mathbb{P}\{\sum_{i=1}^{\tau} X_i > x\} \sim \mathbb{E}[\tau]\overline{F}(x)$.

(P6) Let nonnegative random variable X have distribution F, mean m, and integrated distribution $F^I(x) = (1/m)\int_0^x \overline{F}(t)\,\mathrm{d}t$. Then $\overline{F}(x) = o(\overline{F}^I(x))$ if and only if $\overline{F}^I$ is long tailed.

(P7) If $X_1, X_2, \ldots$ are i.i.d. nonnegative random variables with mean m and whose integrated distribution F^I is long tailed, then, for any $c > 0$ and as $x \to \infty$,

$$\mathbb{P}\biggl(\bigcup_{n\geq 1}\{X_n > x + nc\}\biggr) \sim \sum_{n\geq 1} \mathbb{P}\{X_n > x + nc\} \sim \frac{m}{c}\overline{F}^I(x).$$

Here the first equivalence follows from Bonferroni inequalities: for any events $\{A_n\}$ with $\sum_n \mathbb{P}(A_n) < \infty$, $\sum_n \mathbb{P}(A_n) \geq \mathbb{P}(\bigcup_n A_n) \geq \sum_n \mathbb{P}(A_n) - \sum_{n\neq m} \mathbb{P}(A_n A_m)$ and from observing that $\sum_{n\neq m} \mathbb{P}(A_n A_m) = o(\sum_n \mathbb{P}(A_n))$ in our case. The second equivalence follows from the long tailedness of F^I.

(P8) 'The principle of a single big jump' for a single-server queue with i.i.d. service times σ_n with mean m and subexponential integrated tail distribution F^I, and with i.i.d. interarrival times with mean $b > m$. Let Z be the waiting time of customer 0 that arrives in a stationary queue at time 0. Then, as $x \to \infty$,

$$\mathbb{P}\{Z > x\} \sim \mathbb{P}\Big(\bigcup_{n\geq 1}\{\sigma_{-n} > x + n(b-m)\}\Big) \sim \frac{m}{b-m}\overline{F}^I(x),$$

where the second equivalence follows from property (P6).

Acknowledgements

S. Foss thanks both the Tokyo University of Science for hospitality during his visit, where this work was partly done, supported by the Japan Society for the Promotion of Science under grant no. 24310115, and the FIM Institute (ETH Zürich) for their hospitality in February–April 2013 when the work was largely completed.

References

[1] Baccelli, F. and Foss, S. (2004). Moments and tails in monotone-separable stochastic networks. *Ann. Appl. Prob.* **14,** 612–650.
[2] Baccelli, F., Foss, S. and Lelarge, M. (2005). Tails in generalized Jackson networks with subexponential service-time distributions. *J. Appl. Prob.* **42,** 513–530.
[3] Boxma, O. and Ivanovs, J. (2013). Two coupled Lévy queues with independent input. *Stoch. Systems* **3,** 574–590.
[4] Dai, J. G. and Miyazawa, M. (2011). Reflecting Brownian motion in two dimensions: exact asymptotics for the stationary distribution. *Stoch. Systems* **1,** 146–208.
[5] Foss, S. and Korshunov, D. (2006). Heavy tails in multi-server queue. *Queueing Systems* **52,** 31–48.
[6] Foss, S. and Miyazawa, M. (2014). Two-node fluid network with a heavy-tailed random input: the weak stability case. Working paper.
[7] Foss, S., Korshunov, D. and Zachary, S. (2013). *An Introduction to Heavy-Tailed and Subexponential Distributions*, 2nd edn. Springer, New York.
[8] Kella, O. (1996). Stability and nonproduct form of stochastic fluid networks with Lévy inputs. *Ann. Appl. Prob.* **6,** 186–199.
[9] Kobayashi, M. and Miyazawa, M. (2014). Tail asymptotics of the stationary distribution of a two-dimensional reflecting random walk with unbounded upward jumps. *Adv. Appl. Prob.* **46,** 365–399.
[10] Lelarge, M. (2005). Fluid limit of generalized Jackson queueing networks with stationary and ergodic arrivals and service times. *J. Appl. Prob.* **42,** 491–512.
[11] Lieshout, P. and Mandjes, P. (2008). Asymptotic analysis of Lévy-driven tandem queues. *Queueing Systems* **60,** 203–226.
[12] Miyazawa, M. (2011). Light tail asymptotics in multidimensional reflecting processes for queueing networks. *TOP* **19,** 233–299.
[13] Zwart, B., Borst, S. and Mandjes, M. (2004). Exact asymptotics for fluid queues fed by multiple heavy-tailed on-off sources. *Ann. Appl. Prob.* **14,** 903–957.

SERGEY FOSS, *Heriot Watt University and Novosibirsk State University*

Department of Actuarial Mathematics and Statistics, Heriot-Watt University, Edinburgh EH14 4AS, UK.
Email address: s.foss@hw.ac.uk

MASAKIYO MIYAZAWA, *Tokyo University of Science*

Tokyo University of Science, 2641 Yamazaki, Noda, Chiba, 278-8510, Japan.
Email address: miyazawa@rs.tus.ac.jp