Analysis of the Shortest Queue First service discipline with two classes

Fabrice Guillemin
Orange Labs Networks
2 Avenue Pierre Marzin
Lannion, France
fabrice.guillemin@orange.com

Alain Simonian
Orange Labs Networks
38-40 Rue du Général Leclerc
92794 Issy-les-Moulineaux, France
alain.simonian@orange.com

ABSTRACT

To address the problem of buffer bloat causing latency for time sensitive flows in the Internet, we introduce the Shortest Queue First (SQF) algorithm. This service discipline consists of serving the flow with the least number of backlogged bytes in a buffer. Considering a system with two flows and assuming exponentially distributed service times and Poisson arrivals, we propose a method of computing the Laplace transforms of the workloads in the two virtual queues of the system. We notably derive empty queue probabilities and queue asymptotics. As desired, the analytic results show that SQF performs implicit Head of Line priority for the flow with the smallest traffic intensity.

Categories and Subject Descriptors

C.2 [**Computer-Communication Networks**]: Network Architecture and Design.

General Terms

Performance.

Keywords

Quality of Service; Probabilistic Models; Complex Analysis.

1. INTRODUCTION

It is well known that large buffers in the Internet tend to degrade the quality of experience perceived by end users [9]. While such buffers are necessary to absorb random fluctuations of packet arrival rates, delay sensitive flows such as VoIP may experience extra large delays, resulting in poor quality for the user. This situation holds also for ACK flows associated with data connections such video downloads on TCP. This phenomenon occurs, in particular, for uplink buffers in home gateways, especially when the bit rate of the uplink is small (say, ADSL uplinks with capacity less than 1 Mbit/s).

Measurements from Orange ADSL networks in France show that around 5% to 10 % of ADSL customers suffer from saturation of the uplink, leading to severe quality degradation (see [1] for a discussion of that phenomenon). Recognizing that delay sensitive flows are in general smooth (i.e., with evenly spaced out packet arrivals), it has been proposed in [2, 4] to sort packets out of the home gateway buffer (or any buffer associated with a potentially congested link) according to the so-called "Shortest Queue First" (SQF) policy. With SQF, a virtual buffer is associated with each active flow (i.e., with bytes in the buffer) and SQF serves the queue with the least number of backlogged bytes.

Many simulation studies (see [3, 4] for instance) have been performed to investigate the performance of the SQF policy. This algorithm has also been implemented in the Orange home gateway (known as "Livebox") [2] and a field trial with residential customers has been rolled-out in France (see [1]). All these studies conclude that SQF is an efficient algorithm to reduce latency of time sensitive applications for ADSL customers with small uplink bit rates.

Beyond simulation and laboratory tests, the goal of this paper is to better understand how SQF is performing. We intend, in particular, to give mathematical results explaining what is observed in practice. The analysis, however, reveals quite intricate because this system falls into the class of coupled queues. While such queues are investigated for some packet routing policies upon arrivals, such as the well-known Join the Shortest Queue (JSQ) policy (see [6] for a complete analysis of two asymmetric queues), very few systems depending on the service discipline have been considered in the literature. To the best knowledge of the authors, only the Longest Queue First (LQF) has been analysed [5].

To investigate the performance of SQF by means of a mathematical model, we consider a SQF system with two parallel queues, labelled queues $\sharp 1$ and $\sharp 2$ (we here only consider two parallel queues for mathematical tractability; of course, SQF accomodates an arbitrary number of parallel queues). The SQF service policy then processes data as follows: let $U_1(t)$ (resp. $U_2(t)$) denote the workload in queue $\sharp 1$ (resp. queue $\sharp 2$) at any time t, including the remaining amount of work of the packet possibly in service; then

ValueTools'13, December 10 – 12 2013, Turin, Italy

VALUETOOLS 2013, December 10-12, Torino, Italy

DOI 10.4108/icst.valuetools.2013.254388

- Queue ♯1 (resp. queue ♯2) is served if $U_1(t) \neq 0$, $U_2(t) \neq 0$ and $U_1(t) \leq U_2(t)$ (resp. if $U_1(t) \neq 0$, $U_2(t) \neq 0$ and $U_2(t) < U_1(t)$);
- If only one of the queues is empty, the non-empty queue is served;
- If both queues are empty, the server remains idle until the next job arrival.

In the following, we assume that packets arrive according to Poisson processes and have exponentially distributed volumes. Under that Markovian setting, the Laplace transform F of the joint queue occupancy for either JSQ or LQF has been shown to depend on the solutions to Riemann-Hilbert boundary value problems. We were not able to reduce the analysis of SQF to such a classical framework; as detailed below, the corresponding Laplace transform F is instead derived by means of an iterative process which enables us to express it as a series involving the iterates of two simple functions h_1 and h_2. By studying the location of the singularities of that series, key performance metrics (empty queue probabilities and queue asympotics) can then be obtained. The symmetric case, where input rates and service rates are identical for both queues, has already been solved in [8] by using a similar approach. The general asymmetric case considered here allows, however, a better understanding of the SQF performance; as expected, our mathematical analysis eventually shows that the SQF discipline indeed favors less congested queues.

The organization of this paper is as follows. In Section 2, we recall some basic results obtained in [8]; in Section 3, we solve the functional equations verified by the Laplace transforms of interest. Their analytic continuation is addressed in Section 4; the poles with the smallest modules are then identified, which is used in Section 5 to derive queue asymptotics. Concluding remarks are presented in Section 6. Additional technical details can be found in report [11].

2. PRELIMINARY RESULTS

Throughout this paper, we assume that jobs arrive at queue ♯1 (resp. queue ♯2) according to a Poisson process with mean arrival rate λ_1 (resp. λ_2); we let $\lambda = \lambda_1 + \lambda_2$. Their respective service times are i.i.d. exponential random variables with mean $1/\mu_1$ (resp. $1/\mu_2$). Let then $\varrho_1 = \lambda_1/\mu_1$ (resp. $\varrho_2 = \lambda_2/\mu_2$) be the mean load of queue ♯1 (resp. queue ♯2) and $\varrho = \varrho_1 + \varrho_2$ be the total load of the system.

2.1 Functional equations

In the following, we assume that $\varrho < 1$ so that the system has a stationary regime; we then denote by U_1 and U_2 the workloads in queue ♯1 and ♯2 in that stationary regime, respectively. As established in [8], under the assumption of Poisson arrivals and exponential service times, the pair $(U_1(t), U_2(t))$ representing the workload in both queues at any time t defines a Markov process whose infinitesimal generator is easily expressed in terms of system parameters. The characterization of distribution of (U_1, U_2) by means of that generator then enables us to determine that distribution via its Laplace transform.

Specifically, define the Laplace transforms F_1 and G_1 by

$$\left\{\begin{array}{l} F_1(s_1, s_2) = \mathbb{E}\left[e^{-s_1 U_1 - s_2 U_2} \mathbb{1}_{\{0<U_1<U_2\}}\right], \\ G_1(s_1) = \mathbb{E}\left[e^{-s_1 U_1} \mathbb{1}_{\{0=U_2<U_1\}}\right] \end{array}\right. \quad (1)$$

for $\Re(s_1 + s_2) \geq 0$ and $\Re(s_1) \geq 0$, respectively where $\mathbb{1}_E$ is the indicator function of the set E. Similarly, define the Laplace transforms F_2 and G_2 by

$$\left\{\begin{array}{l} F_2(s_1, s_2) = \mathbb{E}\left[e^{-s_1 U_1 - s_2 U_2} \mathbb{1}_{\{0<U_2<U_1\}}\right], \\ G_2(s_2) = \mathbb{E}\left[e^{-s_2 U_2} \mathbb{1}_{\{0=U_1<U_2\}}\right] \end{array}\right. \quad (2)$$

for $\Re(s_1 + s_2) \geq 0$ and $\Re(s_2) \geq 0$, respectively. The Laplace transform F of the pair (U_1, U_2) is then given by

$$\begin{aligned} F(s_1, s_2) = 1 - \varrho + \\ F_1(s_1, s_2) + G_1(s_1) + F_2(s_1, s_2) + G_2(s_2) \quad (3) \end{aligned}$$

for $\Re(s_1) \geq 0$ and $\Re(s_2) \geq 0$. Further introduce the so-called kernels K_1 and K_2 by setting

$$\left\{\begin{array}{l} K_1(s_1, s_2) = s_1 - \dfrac{\lambda_1 s_1}{s_1 + \mu_1} - \dfrac{\lambda_2 s_2}{s_2 + \mu_2}, \\ K_2(s_1, s_2) = s_2 - \dfrac{\lambda_1 s_1}{s_1 + \mu_1} - \dfrac{\lambda_2 s_2}{s_2 + \mu_2} \end{array}\right. \quad (4)$$

together with

$$\left\{\begin{array}{l} J_1(s_1) = (1 - \varrho)\left(\lambda - \dfrac{\lambda_1 \mu_1}{s_1 + \mu_1}\right) - \psi_2(0), \\ J_2(s_2) = (1 - \varrho)\left(\lambda - \dfrac{\lambda_2 \mu_2}{s_2 + \mu_2}\right) - \psi_1(0), \end{array}\right. \quad (5)$$

with $\psi_j(0) = \lim_{s_j \to \infty} s_j G_j(s_j)$ for $j \in \{1, 2\}$ (it is shown in [11] that $\psi_1(0) + \psi_2(0) = \lambda(1 - \varrho)$). As detailed below, the auxiliary functions M_1 and M_2 defined by

$$\left\{\begin{array}{l} M_1(z) = G_2(2z + \mu_1) + F_1(-\mu_1, 2z + \mu_1), \\ M_2(z) = G_1(2z + \mu_2) + F_2(2z + \mu_2, -\mu_2) \end{array}\right. \quad (6)$$

also intervene in the course of the problem resolution. We precisely have the following result [8, Proposition 3.1].

Proposition 1. *Transforms F_1 and G_2 (resp. F_2, G_1) satisfy the coupled functional equations*

$$K_1 F_1 + K_2 G_2 = J_2 + H, \quad K_2 F_2 + K_1 G_1 = J_1 - H \quad (7)$$

over domain $\mathbf{\Omega} = \{(s_1, s_2) \in \mathbb{C} \times \mathbb{C} \mid \Re(s_1) > 0, \ \Re(s_2) > 0\}$, where

$$H(s_1, s_2) = \frac{\lambda_1 \mu_1}{\mu_1 + s_1} M_1(z) - \frac{\lambda_2 \mu_2}{\mu_2 + s_2} M_2(z) \quad (8)$$

for $(s_1, s_2) \in \mathbf{\Omega}$, with $z = (s_1 + s_2)/2$. In particular, transform G_1 satisfies

$$(s_1 - s_2) G_1(s_1) = J_1(s_1) - H(s_1, s_2) \quad (9)$$

for all $(s_1, s_2) \in \mathbf{\Omega}$ such that $K_2(s_1, s_2) = 0$; similarly, transform G_2 satisfies

$$(s_2 - s_1) G_2(s_2) = J_2(s_2) + H(s_1, s_2) \quad (10)$$

for all $(s_1, s_2) \in \mathbf{\Omega}$ such that $K_1(s_1, s_2) = 0$.

2.2 Analytic continuation

Determining a larger analyticity domain than $\mathbf{\Omega}$ for Laplace transforms F_1, F_2, G_1, G_2 and auxiliary functions M_1, M_2 will enable us to obtain an extended validity domain for functional equations (7), (9) and (10) stated in Proposition 1, which is a key ingredient for their resolution. Such an extension is performed by relying on the related queues with Head of Line (HoL) discipline.

Specifically, let $\overline{U}_j(t)$, $j \in \{1,2\}$, denote the workload in queue $\sharp j$ when the other queue has HoL priority. Under the assumption $\varrho < 1$, this HoL system is stable and the Laplace transform of the workload $\overline{U}_j$ in the stationary regime [10] is given by

$$\mathbb{E}\left(e^{-s_1\overline{U}_1}\right) = \frac{(1-\varrho)s_1\xi_2^+(s_1)}{\lambda_1(1-b_1(s_1))(s_1-\xi_2^+(s_1))}, \tag{11}$$

where $b_1(s_1)$ is the Laplace transform of the service time, $\xi_2^+(s_1)$ is the solution in variable s_2 to $K_2(s_1,s_2)=0$; symmetrically, $\mathbb{E}\big[e^{-s_2\overline{U}_2}\big]$ is obtained by exchanging index 1 in 2 in expression (11), and replacing $\xi_2^+(s_1)$ by the solution $\xi_1^+(s_2)$ in variable s_1 to equation $K_1(s_1,s_2)=0$. Functions ξ_1^+ and ξ_2^+ are specified in the following lemma, whose proof relies on the resolution of elementary quadratic equations and is therefore omitted.

Lemma 1. ***a)*** *For $\iota \in \{1,2\}$, let*

$$D_\iota(s_\iota) = (\mu_\iota(\mu_{\bar\iota}-\lambda_{\bar\iota}) + (\mu_{\bar\iota}-\lambda_\iota-\lambda_{\bar\iota})s_\iota)^2 + 4\lambda_\iota\mu_{\bar\iota}s_\iota(\mu_\iota+s_\iota)$$

where $\bar\iota = 3-\iota$. For given $s_\iota \in \mathbb{C}$, equation $K_{\bar\iota}(s_\iota,s_{\bar\iota})=0$ has two solutions

$$\xi_{\bar\iota}^{\pm}(s_\iota) = \frac{-((\mu_{\bar\iota}-\lambda_{\bar\iota})\mu_\iota + (\mu_{\bar\iota}-\lambda_\iota-\lambda_{\bar\iota})s_\iota) \pm \sqrt{D_\iota(s_\iota)}}{\bar\iota(s_\iota+\mu_\iota)} \tag{12}$$

(note that $\xi_{\bar\iota}^+(0)=0$ and $\xi_{\bar\iota}^-(0)=\lambda_{\bar\iota}-\mu_{\bar\iota}$). Function $\xi_{\bar\iota}^-$ (resp. $\xi_{\bar\iota}^+$) has an analytic (resp. meromorphic) extension to the cut plane $\mathbb{C}\setminus[\zeta_\iota^-,\zeta_\iota^+]$, where

$$\zeta_\iota^\pm = -\mu_\iota\frac{(\sqrt{\mu_{\bar\iota}}\mp\sqrt{\lambda_{\bar\iota}})^2}{\lambda_\iota+(\sqrt{\mu_{\bar\iota}}\mp\sqrt{\lambda_{\bar\iota}})^2}. \tag{13}$$

b) *For $\mu_1 \neq \mu_2$, the roots $s \neq 0$ of $K_1(s,s)=K_2(s,s)=s$ are that of quadratic polynomial*

$$P(s) = s^2 + (\mu_1+\mu_2-\lambda)s + \mu_1\mu_2(1-\varrho); \tag{14}$$

these roots are real negative, $\sigma_0^- < \sigma_0^+ < 0$. For $\mu_1 = \mu_2$, the only root $s \neq 0$ of $K(s,s)=s$ is $\sigma_0 = -\mu(1-\varrho)$.

On the basis of the above result, we can determine the analyticity domain for the Laplace transform of $\overline{U}_1$ and $\overline{U}_2$.

Lemma 2. *Transform $s_1 \mapsto \mathbb{E}(e^{-s_1\overline{U}_1})$ is analytic in the domain $\{s_1 \in \mathbb{C} \mid \Re(s_1) > \widetilde{s}_1\}$ where*

$$\widetilde{s}_1 = \begin{cases} \sigma_0^+ \ \textit{if}\ (2\mu_2-\mu_1)\sqrt{\varrho_2}+\mu_1 \leq \mu_2+\lambda_1+\lambda_2 & (\mathrm{I}^+), \\ \zeta_1^+ \ \textit{if}\ (2\mu_2-\mu_1)\sqrt{\varrho_2}+\mu_1 > \mu_2+\lambda_1+\lambda_2 & (\mathrm{I}^-). \end{cases}$$

Similarly, transform $s_2 \mapsto \mathbb{E}(e^{-s_2\overline{U}_2})$ is analytic in the domain $\{s_2 \in \mathbb{C} \mid \Re(s_2) > \widetilde{s}_2\}$ where

$$\widetilde{s}_2 = \begin{cases} \sigma_0^+ \ \textit{if}\ (2\mu_1-\mu_2)\sqrt{\varrho_1}+\mu_2 \leq \mu_1+\lambda_1+\lambda_2 & (\mathrm{II}^+), \\ \zeta_2^+ \ \textit{if}\ (2\mu_1-\mu_2)\sqrt{\varrho_1}+\mu_2 > \mu_1+\lambda_1+\lambda_2 & (\mathrm{II}^-). \end{cases}$$

Proof. By Lemma 1.a, expression (11) defines a meromorphic transform in the cut plane $\mathbb{C}\setminus[\zeta_1^-,\zeta_1^+]$ since function ξ_2^+ is itself a meromorphic function in this domain; its possible poles are the solutions to $s_1 = \xi_2^+(s_1)$, that is, $s_1 = \sigma_0^\pm$ as defined in Lemma 1.b (we obviously have $\zeta_1^+ \leq \sigma_0^+$).

To localize such poles, consider the curves $s_2 = \xi_2^+(s_1)$ and $s_2 = \xi_2^-(s_1)$ in the real plane (O, Os_1, Os_2) for $s_1 > \zeta_1^+$. Basic calculus shows that the graph of function $s_1 \mapsto \xi_2^+(s_1)$ (resp. $s_1 \mapsto \xi_2^-(s_1)$) is concave (resp. convex) for $s_1 > \zeta_1^+$. These graphs and the vertical axis Os_2 delineate a convex domain and the straight line $s_1 = s_2$ intersects that domain at either $s_1 = s_2 = 0$ or $s_1 = s_2 = \sigma_0^+$ (see Fig.1). The latter intersection point then belongs to the upper branch $s_2 = \xi_2^+(s_1)$ if and only if $\sigma_0^+ \geq a_2^+ = \xi_2^\pm(\zeta_1^+)$, in which case σ_0^+ is a pole for expression (11). Conversely, condition $\sigma_0^+ < a_2^+$ ensures that σ_0^+ is not a pole for (11) and that its smallest singularity is consequently ζ_1^+.

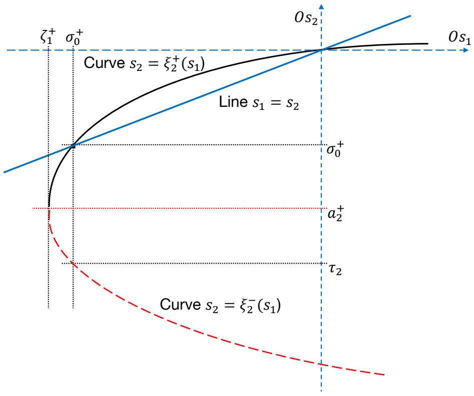

Figure 1: Curves $s_2 = \xi_2^+(s_1)$ (black) and $s_2 = \xi_2^-(s_1)$ (dotted red) and their intersection with line $s_1 = s_2$ under condition (I^+).

Quadratic polynomial $P(s)$ has roots $\sigma_0^- < \sigma_0^+ < 0$. Condition $\sigma_0^+ \geq a_2^+$ is then equivalent to $P(a_2^+) \leq 0$, which reduces to $(2\mu_2-\mu_1)\sqrt{\varrho_2}+\mu_1-\mu_2-\lambda_1-\lambda_2 \leq 0$. We can then conclude that threshold $\widetilde{s}_1$ equals either σ_0^+ or ζ_1^+ according to condition (I^+) or (I^-), as claimed. A similar proof holds for threshold $\widetilde{s}_2$. $\square$

By using Lemma 2 and the stochastic domination property $U_1 \leq \overline{U}_1$ (resp. $U_2 \leq \overline{U}_2$) when queue $\sharp 2$ has HoL priority (resp. when queue $\sharp 1$ has HoL priority), we can now deduce the extended analyticity domains of transforms F_1, G_1 and F_2, G_2.

Corollary 1. *Given the constant $\widetilde{s}_2$ defined in Lemma 2, transform F_1 can be analytically extended to the domain*

$\widetilde{\mathbf{\Omega}}_1 = \{(s_1,s_2) \in \mathbb{C} \times \mathbb{C} \mid \ \Re(s_2) > \max(\widetilde{s}_2, \widetilde{s}_2 - \Re(s_1))\}$ *and transform* G_2 *can be analytically extended to the domain* $\widetilde{\omega}_2 = \{s_2 \in \mathbb{C} \mid \ \Re(s_2) > \widetilde{s}_2\}$.

Given constant $\widetilde{s}_1$ *defined in Lemma 2,* F_2 *can be extended to* $\widetilde{\mathbf{\Omega}}_2 = \{(s_1,s_2) \in \mathbb{C} \times \mathbb{C} \mid \ \Re(s_1) > \max(\widetilde{s}_1, \widetilde{s}_1 - \Re(s_2))\}$ *and* G_1 *can be extended to* $\widetilde{\omega}_1 = \{s_1 \in \mathbb{C} \mid \ \Re(s_1) > \widetilde{s}_1\}$.

As a direct consequence of Corollary 1, we deduce that both functions $z \mapsto M_1(z)$ and $z \mapsto M_2(z)$ introduced in (6) are analytic for $\Re(z) > \max(\widetilde{s}_1, \widetilde{s}_2)/2$.

3. SOLVING FUNCTIONAL EQUATIONS

The objective of this section is to show that functions M_1 and M_2 verify a two-dimensional functional equation that can be solved in terms of a series expansion. This notably allows us to compute empty queue probabilities.

3.1 Underlying cubic equations

As stressed above, algebraic equations $K_1(s_1,s_2) = 0$ and $K_2(s_1,s_2) = 0$ play a central role in solving Equations (9)-(10), enabling one variable, say s_1, to be expressed in terms of s_2 or conversely. As algebraic curves in the (O, Os_1, Os_2) plane, these equations represent rational cubics (i.e., cubics with a rational parametrization). In the following, we will define a suitable variable change $(s_1,s_2) \mapsto (w,z)$ so that variable z parametrize both cubics and a pair of algebraic functions $h_1 : z \mapsto h_1(z)$, $h_2 : z \mapsto h_2(z)$ be defined so as to write equations (9)-(10) in an iterative form.

Define the variable change $(s_1,s_2) \mapsto (w,z)$ by

$$2w = s_1 - s_2, \quad 2z = s_1 + s_2. \tag{15}$$

Equations $K_1(s_1,s_2) = 0$ and $K_2(s_1,s_2) = 0$ are then equivalent to $K_1(z+w, z-w) = 0$ and $K_2(z+w, z-w) = 0$, respectively, with

$$\left\{ \begin{array}{l} K_1(z+w, z-w) = \dfrac{-R_1(w,z)}{(w+z+\mu_1)(-w+z+\mu_2)}, \\ K_2(z+w, z-w) = \dfrac{+R_2(w,z)}{(w+z+\mu_1)(-w+z+\mu_2)} \end{array} \right. \tag{16}$$

and where $R_1(w,z)$ and $R_2(w,z)$ are the cubic polynomials

$$R_1(w,z) = \sum_{k=0}^{3} R_{1k}(z) w^{3-k}, \quad R_2(w,z) = \sum_{k=0}^{3} R_{2k}(z) w^{3-k}$$

with respective coefficients

$$\left\{ \begin{array}{l} R_{10}(z) = 1, \ \ R_{11}(z) = -(\lambda - \mu_1 + \mu_2 - z), \\ R_{12}(z) = \lambda_1\mu_2 - \lambda_2\mu_1 - \mu_1\mu_2 - 2\mu_2 z - z^2, \\ R_{13}(z) = -zP(z), \end{array} \right. \tag{17}$$

and

$$\left\{ \begin{array}{l} R_{20}(z) = 1, \ \ R_{21}(z) = \lambda + \mu_1 - \mu_2 - z, \\ R_{22}(z) = \lambda_2\mu_1 - \lambda_1\mu_2 - \mu_1\mu_2 - 2\mu_1 z - z^2, \\ R_{23}(z) = zP(z), \end{array} \right. \tag{18}$$

$P(z)$ being given by (14). The roots in variable w of polynomials $R_1(w,z)$ and $R_2(w,z)$ satisfy the following properties; the proof invokes simple algebraic arguments and can be found in [11].

LEMMA 3. *For* $z > 0$*, polynomial* $R_1(\bullet, z)$ *has 3 real roots* $\alpha_1(z)$, $\beta_1(z)$, $\gamma_1(z)$ *with* $\alpha_1(z) < -z < \beta_1(z) < z < \gamma_1(z)$*. Similarly, polynomial* $R_2(\bullet, z)$ *has 3 real roots* $\alpha_2(z)$, $\beta_2(z)$, $\gamma_2(z)$ *with* $\alpha_2(z) < -z < \beta_2(z) < z < \gamma_2(z)$.

PROOF. We calculate $R_1(-z,z) = 2\lambda_2\mu_1 z$ and $R_1(z,z) = -2\mu_2 z(2z - \lambda_1 + \mu_1)$ so that for given $z > 0$, $R_1(-z,z) > 0$ and $R_1(z,z) < 0$ since $\lambda_1 < \mu_1$. Further, the 3rd degree polynomial $R_1(\bullet, z)$ has limits $R_1(-\infty, z) = -\infty$ and $R_1(+\infty, z) = +\infty$. For given $z > 0$, real polynomial $R_1(\bullet, z)$ has therefore 3 distinct real roots $\alpha_1(z)$, $\beta_1(z)$, $\gamma_1(z)$ which can be ordered as $\alpha_1(z) < -z < \beta_1(z) < z < \gamma_1(z)$. Similar arguments hold for $R_2(\bullet, z)$. ☐

LEMMA 4. *For given* $z > 0$*, the largest solution* s_2 *and* s_1 *to equations*

$$z = \frac{s_2 + \xi_1^-(s_2)}{2}, \quad z = \frac{s_1 + \xi_2^-(s_1)}{2} \tag{19}$$

are $s_2 = z - \alpha_1(z)$ *and* $s_1 = z + \gamma_2(z)$*, respectively. For such values of* s_2 *and* s_1*, define function* h_1 *and* h_2 *by*

$$h_1(z) = \frac{s_2 + \xi_1^+(s_2)}{2}, \quad h_2(z) = \frac{s_1 + \xi_2^+(s_1)}{2}. \tag{20}$$

respectively. We then have

$$h_1(z) > z, \quad h_2(z) > z \tag{21}$$

for all $z > 0$.

PROOF. By Lemma 1, there are two distinct solutions s_1 to equation $K_1(s_1,s_2) = 0$ for given $s_2 > 0$, namely $\xi_1^-(s_2)$ and $\xi_1^+(s_2)$, its smallest solution being $s_1 = \xi_1^-(s_2)$. As $K_1(s_1,s_2) = K_1(z+w, z-w)$ and the zeros of $K_1(z+w, z-w)$ in w are that of $R_1(w,z)$ by (16), we can set $s_1 = z + \epsilon_1(z)$ and $s_2 = z - \epsilon_1(z)$ where $\epsilon_1(z) \in \{\alpha_1(z), \beta_1(z), \gamma_1(z)\}$. For given $z > 0$, Lemma 3 entails that the smallest value of s_1 corresponds to $\epsilon_1(z) = \alpha_1(z)$, hence $s_2 = z - \alpha_1(z)$. For that value of s_2, the minimal (resp. maximal) value of the sum $s_1 + s_2$ corresponds to $s_1 = \xi_1^-(s_2)$ (resp. $s_1 = \xi_1^+(s_2)$), the corresponding extreme values being $2z = \xi_1^-(s_2) + s_2$ and $2h_1(z) = \xi_1^+(s_2) + s_2$. By construction, we thus have $h_1(z) > z$ for each given $z > 0$. The discussion for $h_2(z)$ follows a similar pattern. ☐

Property (21) will prove essential for ensuring the convergence of series defining the final solution to Equations (9)-(10).

3.2 Functional equations for M_1 and M_2

We can now specify the functional equations verified by functions M_1 and M_2 and complete their resolution. Consider any root $\epsilon_j(z) \in \{\alpha_j(z), \beta_j(z), \gamma_j(z)\}$ of polynomial $R_j(\cdot, z)$, $j \in \{1,2\}$, defined in Lemma 3 for real $z > 0$. We let

$$q_1(z; \epsilon_j(z)) = \frac{\lambda_1\mu_1}{\mu_1 + z + \epsilon_j(z)}, \quad q_2(z; \epsilon_j(z)) = \frac{\lambda_2\mu_2}{\mu_2 + z - \epsilon_j(z)} \tag{22}$$

and write $q_1(z; \epsilon_j(z)) = q_1(\epsilon_j)$ and $q_2(z; \epsilon_j(z)) = q_2(\epsilon_j)$ for short without mentioning the current argument z of ϵ_j. Given $\beta_1 = \beta_1(z)$ and $\beta_2 = \beta_2(z)$, we also set

$$\beta_1^* = \beta_1(z_1^*), \quad \beta_2^* = \beta_2(z_2^*)$$

where $z_1^* = h_1(z)$ and $z_2^* = h_2(z)$ are defined by (20); in a similar manner, we write $q_1(\beta_1^*) = q_1(z_1^*, \beta_1(z_1^*))$ and $q_2(\beta_2^*) = q_2(z_2^*, \beta_2(z_2^*))$. Defining the 2×1 column vector

$$\mathbf{M}(z) = \left(M_1(z) \;\; M_2(z)\right)^T ,$$

we have the following result.

Proposition 2. *Function $\mathbf{M}$ verifies*

$$\mathbf{M}(z) = Q_1(z) \cdot \mathbf{M}\left(h_1(z)\right) + Q_2(z) \cdot \mathbf{M}\left(h_2(z)\right) + \mathbf{L}(z) \quad (23)$$

for all $z > 0$, where the 2×2 matrices

$$Q_1(z) = k_1(z)\Pi_1(z), \quad Q_2(z) = k_2(z)\Pi_2(z)$$

are defined by factors

$$k_1(z) = \frac{1}{D(z)} \frac{s_2 - \xi_1^-(s_2)}{s_2 - \xi_1^+(s_2)}, \;\; k_2(z) = \frac{1}{D(z)} \frac{s_1 - \xi_2^-(s_1)}{s_1 - \xi_2^+(s_1)}$$

with $s_2 = z - \alpha_1(z)$, $s_1 = z + \gamma_2(z)$ and

$$D(z) = 4\lambda_1\mu_1\lambda_2\mu_2 \frac{(\mu_1 + \mu_2 + 2z)\alpha_1\gamma_2(\alpha_1 - \gamma_2)}{R_1(\gamma_2, z)R_2(\alpha_1, z)}, \qquad (24)$$

by matrices

$$\Pi_1 = \begin{pmatrix} -q_2(\gamma_2)q_1(\beta_1^*) & q_2(\gamma_2)q_2(\alpha_1) \\ -q_1(\gamma_2)q_1(\beta_1^*) & q_1(\gamma_2)q_2(\alpha_1) \end{pmatrix},$$

$$\Pi_2 = \begin{pmatrix} q_2(\alpha_1)q_1(\gamma_2) & -q_2(\alpha_1)q_2(\beta_2^{**}) \\ q_1(\alpha_1)q_1(\gamma_2) & -q_1(\alpha_1)q_2(\beta_2^{**}) \end{pmatrix},$$

and where vector $\mathbf{L}(z) = \left(L_1(z) \;\; L_2(z)\right)^T$ is given by

$$L_1 = \frac{1}{D}\left[q_2(\alpha_1)\left(\frac{\xi_2^-(s_1) - \xi_2^+(s_1)}{s_1 - \xi_2^+(s_1)}\right) J_1(s_1) - q_2(\gamma_2)\left(\frac{\xi_1^+(s_2) - \xi_1^-(s_2)}{s_2 - \xi_1^+(s_2)}\right) J_2(s_2)\right]$$

and

$$L_2 = \frac{1}{D}\left[q_1(\alpha_1)\left(\frac{\xi_2^-(s_1) - \xi_2^+(s_1)}{s_1 - \xi_2^+(s_1)}\right) J_1(s_1) - q_1(\gamma_2)\left(\frac{\xi_1^+(s_2) - \xi_1^-(s_2)}{s_2 - \xi_1^+(s_2)}\right) J_2(s_2)\right].$$

Proof. Note that $s_2 + \xi_1^+(s_2)$ and $s_2 + \xi_1^-(s_2)$ are positive for large enough real s_2. Successively taking $s_1 = \xi_1^+(s_2)$ and $s_1 = \xi_1^-(s_2)$ in Equation (10), we obtain

$$(s_2 - \xi_1^+(s_2))G_2(s_2) = J_2(s_2) + \frac{\lambda_1\mu_1}{\mu_1 + \xi_1^+(s_2)} M_1(h_1(z)) - \frac{\lambda_2\mu_2}{\mu_2 + s_2} M_2(h_1(z)) \quad (25)$$

after definition (20) for $h_1(z)$, and

$$(s_2 - \xi_1^-(s_2))G_2(s_2) = J_2(s_2) + \frac{\lambda_1\mu_1}{\mu_1 + \xi_1^-(s_2)} M_1(z) - \frac{\lambda_2\mu_2}{\mu_2 + s_2} M_2(z) \quad (26)$$

with $s_2 = z - \alpha_1$ and $\xi_1^-(s_2) = z + \alpha_1$ by Lemma 4. Equating then the common value of $G_2(s_2)$ from the above equations, we have

$$q_1(\alpha_1)M_1(z) - q_2(\alpha_1)M_2(z) = \frac{\xi_1^+(s_2) - \xi_1^+(s_2)}{s_2 - \xi_1^+(s_2)} J_2(s_2) + \frac{\lambda_1\mu_1(s_2 - \xi_1^-(s_2))}{(\mu_1 + \xi_1^+(s_2))(s_2 - \xi_1^+(s_2))} M_1(h_1(z)) - \frac{\lambda_2\mu_2(s_2 - \xi_1^-(s_2))}{(\mu_2 + s_2)(s_2 - \xi_1^+(s_2))} M_2(h_1(z)) \quad (27)$$

for large enough real s_2 and with $q_1(\alpha_1)$ and $q_2(\alpha_1)$ defined in (22) for $\epsilon_1 = \alpha_1$. Similarly, applying Equation (9) to $s_2 = \xi_2^+(s_1)$ and $s_2 = \xi_2^-(s_1)$ for $z = s_1 + \xi_2^-(s_1) \geq 0$ for large enough real s_1, we obtain a second equation which, together with Equation (27), gives the linear system

$$V\mathbf{M}(z) = \mathbf{N}_0 + V_1\mathbf{M}(h_1(z)) + V_2\mathbf{M}(h_2(z)) \qquad (28)$$

for vector $\mathbf{M}(z)$, with 2×2 matrix

$$V = \begin{pmatrix} q_1(\alpha_1) & -q_2(\alpha_1) \\ q_1(\gamma_2) & -q_2(\gamma_2) \end{pmatrix},$$

some diagonal matrices V_1, V_2 and some 2×1 vector $\mathbf{N}_0$; linear system (28) can then be solved for $\mathbf{M}(z)$ in terms of $\mathbf{M}(h_1(z))$ and $\mathbf{M}(h_2(z))$, provided that matrix V above has non-zero determinant $D = \det V$. In fact, applying definition (22) for coefficients $q_1(\gamma_2)$, $q_2(\gamma_2)$ and $q_1(\alpha_1)$, $q_2(\alpha_1)$, we calculate

$$\begin{aligned} D &= q_1(\gamma_2)q_2(\alpha_1) - q_2(\gamma_2)q_1(\alpha_1) \\ &= \frac{\lambda_1\mu_1}{\mu_1 + \gamma_2 + z}\frac{\lambda_2\mu_2}{\mu_2 - \alpha_1 + z} - \frac{\lambda_2\mu_2}{\mu_2 - \gamma_2 + z}\frac{\lambda_1\mu_1}{\mu_1 + \alpha_1 + z} \\ &= \frac{\lambda_1\mu_1\lambda_2\mu_2(\mu_1 + \mu_2 + 2z)(\alpha_1 - \gamma_2)}{(\mu_1 + \gamma_2 + z)(\mu_2 - \gamma_2 + z)(\mu_1 + \alpha_1 + z)(\mu_2 - \alpha_1 + z)}; \end{aligned}$$

As $R_1(w, z) + R_2(w, z) = -2w(w + z + \mu_1)(-w + z + \mu_2)$ for any pair (w, z), we can write

$$(\mu_1 + \alpha_1 + z)(\mu_2 - \alpha_1 + z) = -\frac{R_2(\alpha_1, z)}{2\alpha_1} \qquad (29)$$

since $R_1(\alpha_1, z) = 0$; we similarly write

$$(\mu_1 + \gamma_2 + z)(\mu_2 - \gamma_2 + z) = -\frac{R_1(\gamma_2, z)}{2\gamma_2}; \qquad (30)$$

determinant $D = \det V$ then reduces to expression (24) and is consequently non-zero for $z > 0$ in view of Lemma 3. Solving then system (28) for $\mathbf{M}(z)$ in terms of $\mathbf{M}(h_1(z))$ and $\mathbf{M}(h_2(z))$ readily provides functional relation (23). $\square$

Using functional equation (23), we can now obtain a series expansion for $\mathbf{M}(z)$. As detailed below, that expansion involves the semi-group $< h_1; h_2 >$ generated by h_1 and h_2, that is, the set of all iterates

$$h = h_{i_1} \circ h_{i_2} \circ ... \circ h_{i_k}$$

for any $k \in \mathbb{N}$ and $\mathbf{i}_k = (i_1, ..., i_k) \in \{1, 2\}^k$ ($\circ$ denotes the composition of functions and we set $h = \mathrm{Id}$ for $k = 0$).

THEOREM 1. *The column vector* $\mathbf{M}$ *is given by the series expansion*

$$\mathbf{M}(z) = \sum_{k=0}^{+\infty} \sum_{\mathbf{i}_k \in \{1,2\}^k} \mathbb{T}_{\mathbf{i}_k}(z) \cdot \mathbf{L}\left(h_{\mathbf{i}_k}(z)\right) \tag{31}$$

for all $z > 0$, *with* $h_{i,j,...,\ell} = h_i \circ h_j \circ ... \circ h_\ell$ *and where*

$$\mathbb{T}_{\mathbf{i}_k} = Q_{i_k} Q_{i_{k-1}}(h_{i_k}) ... Q_{i_1}(h_{i_2,...,i_k})$$

is a product matrix, with matrices Q_1 *and* Q_2 *introduced in Proposition 2 (by convention, that product reduces to the unit matrix* Id *for* $k = 0$, *and we set* $Q_{i_{k-\ell}}(h_{i_{k-\ell+1},...,i_k}) = \mathrm{Id}$ *for* $k \geq 1$ *and* $\ell = 0$*).*

PROOF. For given $z > 0$, let $\mathbf{M}_k(z)$ denote the generic term at order $k \geq 0$ of series (31). Applying then iteratively functional equation (23) to order $K \geq 1$, we obtain

$$\mathbf{M}(z) = \sum_{0 \leq k \leq K} \mathbf{M}_k(z) + \mathbf{E}^{(K)}(z)$$

where the remainder term $\mathbf{E}^{(K)}(z)$ is equal to

$$\sum_{\mathbf{i}_{K+1} \in \{1,2\}^{K+1}} \prod_{\ell=0}^{K} Q_{i_{K+1-\ell}}(h_{i_{K-\ell+2},...,i_{K+1}}(z)) \cdot \mathbf{M}\left(h_{\mathbf{i}_{k+1}}(z)\right).$$

and with the notation $\mathbf{i}_{K+1} = (i_1, \ldots, i_{K+1})$.

We now show that $\mathbf{E}^{(K)}(z) \to 0$ as $K \uparrow +\infty$. As detailed in [8], Theorem 5.1, Property (21) for h_1 and h_2 implies that the sequence of iterated $h_1 \circ ... \circ h_1(z)$, K times, (resp. $h_2 \circ ... \circ h_2(z)$, K times) of function h_1 (resp. function h_2) tends to $+\infty$ when $K \uparrow +\infty$. As a consequence, any iterated $h_{i_1,...,i_K,i_{K+1}}(z)$ tends to $+\infty$ when $K \uparrow +\infty$. On the other hand, following definition (8), functions M_1 and M_2 are bounded in the neighborhood of infinity since Laplace transforms F_1, F_2 and G_1, G_2 vanish at infinity; the sequence $\mathbf{M}\left(h_{i_1,...,i_K,i_{K+1}}(z)\right)$, $(i_1, ..., i_K, i_{K+1}) \in \{1,2\}^{K+1}$, $K \geq 0$, is consequently bounded.

By arguments similar to that of [8], Theorem 5.1, letting $z \uparrow +\infty$ implies that $z + \alpha_1(z)$ (resp. $z - \alpha_1(z)$) tends to σ_1^- (resp. $+\infty$) and $z + \gamma_2(z)$ (resp. $z - \gamma_2(z)$) tends to $+\infty$ (resp. to σ_2^-) where

$$\left\{ \begin{aligned} \sigma_1^{\pm} &= \frac{\lambda - \mu_1 \pm \sqrt{(\lambda - \mu_1)^2 + 4\lambda_2\mu_1}}{2}, \\ \sigma_2^{\pm} &= \frac{\lambda - \mu_2 \pm \sqrt{(\lambda - \mu_2)^2 + 4\lambda_1\mu_2}}{2}. \end{aligned} \right. \tag{32}$$

It follows that $\Pi_1 = \Pi_1(z)$ and $\Pi_2 = \Pi_2(z)$ defined in Proposition 2 are such that

$$\begin{aligned} \Pi_1(z) &\to \begin{pmatrix} -\dfrac{\lambda_1\mu_1}{\mu_1 + \sigma_1^+} \times \dfrac{\lambda_2\mu_2}{\mu_2 + \sigma_2^-} & 0 \\ 0 & 0 \end{pmatrix}, \\ \Pi_2(z) &\to \begin{pmatrix} 0 & 0 \\ 0 & -\dfrac{\lambda_1\mu_1}{\mu_1 + \sigma_1^-} \times \dfrac{\lambda_2\mu_2}{\mu_2 + \sigma_2^+} \end{pmatrix} \end{aligned}$$

as $z \uparrow +\infty$. On the other hand, the definitions of factors $k_1(z)$ and $k_2(z)$ given in Proposition 2 give in turn $D(z)k_1(z) \to 1$ and $D(z)k_2(z) \to 1$. Besides, identities (29)-(30) entail that $R_2(\alpha_1, z) \sim -(-2z)(\mu_1 + \sigma_1^-)2z$ and $R_1(\gamma_2, z) \sim -(2z)2z(\mu_2 + \sigma_2^-)$. Using the above estimates, definition (24) of $D(z)$ then gives

$$\begin{aligned} D(z) &\sim 4\lambda_1\mu_1\lambda_2\mu_2 \frac{2z \times (-z)z(-2z)}{(-4(\mu_2 + \sigma_2^-)z^2)(4(\mu_1 + \sigma_1^-)z^2)} \\ &= -\frac{\lambda_1\lambda_2\mu_1\mu_2}{(\mu_1 + \sigma_1^-)(\mu_2 + \sigma_2^-)}. \end{aligned}$$

The previous estimates therefore show that matrices $Q_1(z)$ and $Q_2(z)$ tend to

$$\begin{pmatrix} r_1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ 0 & r_2 \end{pmatrix}$$

respectively, where $r_j = (\mu_j + \sigma_j^-)/(\mu_j + \sigma_j^+)$. The non-zero element r_j is > 0 and < 1 since $0 < \mu_j + \sigma_j^- < \mu_j + \sigma_j^+$ for each $j \in \{1, 2\}$. Using expressions (32) of σ_1^- and σ_1^+ and writing

$$\begin{aligned} r_1 &= \frac{\mu_1 + \sigma_1^-}{\mu_1 + \sigma_1^+} \\ &= \frac{4\varrho_1}{(\varrho_1 + m\varrho_2 + 1 + \sqrt{(\varrho_1 + m\varrho_2 - 1)^2 + 4m\varrho_2})^2} \end{aligned}$$

we further note that r_1 is a decreasing function of the ratio $m = \mu_2/\mu_1$, and equals ϱ_1 for $m = 0$; we thus deduce that $r_1 \leq \varrho_1$ and similarly $r_2 \leq \varrho_2$. Coming back to the definition of remainder $\mathbf{E}^{(K)}(z)$ above, the above arguments therefore imply that

$$\mathbf{E}^{(K)}(z) = O\left(\sum_{n_1 + n_2 = K+1} r_1^{n_1} r_2^{n_2}\right) = O(r_1 + r_2)^K = O(\varrho^K)$$

for large K, where $\varrho = \varrho_1 + \varrho_2 < 1$. Remainder $\mathbf{E}^{(K)}(z)$ therefore tends to 0 for increasing K; as $\mathbf{M}(z)$ is finite for any $z > 0$ by the existence of the stationary distribution, we conclude that expansion (31) holds for such values of z. ☐

As a first conclusion, the derivation for functions M_1 and M_2 in Theorem 1 enables us to obtain transform G_2 by equality (26), together with transform G_1 via a similar equality; since M_1 and M_2 also determine H via (8), transforms F_1 and F_2 are then derived from Equations (7) which in turn determines the complete solution F. Following (31), however, solution $\mathbf{M}$ linearly depends on vector $\mathbf{L}$ and is therefore a linear combination of functions J_1 and J_2 introduced in (5). The latter still depend on unknown constants $\psi_1(0)$ and $\psi_2(0)$ which can be determined as follows. First write $\mathbf{L}(z) = \mathbf{L}^{(0)}(z) + \psi_1(0)\mathbf{L}^{(1)}(z) + \psi_2(0)\mathbf{L}^{(2)}(z)$ as a linear combination of of $\psi_1(0)$ and $\psi_2(0)$ with

$$\begin{aligned} \mathbf{L}^{(0)}(z) = \frac{-(1 - \varrho)}{D(z)} \Bigg[&\frac{\xi_2^+(s_1) - \xi_2^-(s_1)}{s_1 - \xi_2^+(s_1)}(\lambda - \lambda_1 b_1(s_1))\mathbf{e}_2(z) \\ &\frac{\xi_1^+(s_2) - \xi_1^-(s_2)}{s_2 - \xi_1^+(s_2)}(\lambda - \lambda_2 b_2(s_2))\mathbf{e}_1(z) \Bigg], \end{aligned}$$

and

$$\begin{aligned} \mathbf{L}^{(1)}(z) &= \frac{1}{D(z)} \frac{\xi_1^+(s_2) - \xi_1^-(s_2)}{s_2 - \xi_1^+(s_2)} \mathbf{e}_1(z), \\ \mathbf{L}^{(2)}(z) &= \frac{1}{D(z)} \frac{\xi_2^+(s_1) - \xi_2^-(s_1)}{s_1 - \xi_2^+(s_1)} \mathbf{e}_2(z) \end{aligned}$$

with $\mathbf{e}_1(z) = (q_2(\gamma_2) \quad q_1(\gamma_2))^T$, $\mathbf{e}_2(z) = (q_2(\alpha_1) \quad q_1(\alpha_1))^T$ and where $s_1 = z + \gamma_2(z)$, $s_2 = z - \alpha_1(z)$. For $i \in \{0,1,2\}$, let now $\mathcal{L}^{(i)}$ denote the 2×1 vector satisfying the functional equation

$$\mathcal{L}^{(i)}(z) = Q_1(z) \cdot \mathcal{L}^{(i)}\left(h_1(z)\right) + Q_2(z) \cdot \mathcal{L}^{(i)}(h_2(z)) + \mathbf{L}^{(i)}(z)$$

for $z > 0$, whose solution is given by Theorem 1 as

$$\mathcal{L}^{(i)}(z) = \sum_{k=0}^{+\infty} \sum_{\mathbf{i}_k \in \{1,2\}^k} \mathbb{T}_{\mathbf{i}_k}(h_{\mathbf{i}_k}(z)) \cdot \mathbf{L}^{(i)}\left(h_{\mathbf{i}_k}(z)\right) \qquad (33)$$

so that $\mathbf{M}(z) = \mathcal{L}^{(0)}(z) + \psi_1(0)\mathcal{L}^{(1)}(z) + \psi_2(0)\mathcal{L}^{(2)}(z)$.

Proposition 3. *For each $i \in \{0,1,2\}$, denote by $\mathcal{L}_j^{(i)}$, $j \in \{1,2\}$, the components of vector $\mathcal{L}^{(i)}(z)$ defined by expansion (33).*

a) *Constant $\psi_1(0)$ is then given by*

$$\psi_1(0) = \frac{\lambda_1(1-\varrho) + \psi_{11} + \lambda(1-\varrho)\psi_{12}}{1 + \psi_{13}}$$

with

$$\psi_{1,1} = \lambda_1 \mathcal{L}_1^{(0)}(0) - \lambda_2 \mathcal{L}_2^{(0)}(0),$$
$$\psi_{1,2} = \lambda_1 \mathcal{L}_1^{(2)}(0) - \lambda_2 \mathcal{L}_2^{(2)}(0),$$
$$\psi_{1,3} = -\lambda_1 \mathcal{L}_1^{(1)}(0) + \lambda_1 \mathcal{L}_1^{(2)}(0) + \lambda_2 \mathcal{L}_2^{(1)}(0) - \lambda_2 \mathcal{L}_2^{(2)}(0)$$

and $\psi_2(0) = \lambda(1-\rho) - \psi_1(0)$.

b) *The empty queue probabilities are given by*

$$\mathbb{P}(U_1 = 0) = 1 - \varrho + G_2(0), \quad \mathbb{P}(U_2 = 0) = 1 - \varrho + G_1(0) \quad (34)$$

with

$$G_2(0) = \lim_{s_2 \downarrow 0} \frac{1}{s_2 - \xi_1^+(s_2)} \left[J_2(s_2) + \frac{\lambda_1 \mu_1 M_1(h_1(z))}{\mu_1 + \xi_1^+(s_2)} - \frac{\lambda_2 \mu_2 M_2(h_1(z))}{\mu_2 + s_2} \right]$$

where $h_1(z) = (s_2 + \xi_1^+(s_2))/2$, and

$$G_1(0) = \lim_{s_1 \downarrow 0} \frac{1}{s_1 - \xi_2^+(s_1)} \left[J_1(s_1) - \frac{\lambda_1 \mu_1 M_1(h_2(z))}{\mu_1 + s_1} + \frac{\lambda_2 \mu_2 M_2(h_2(z))}{\mu_2 + \xi_2^+(s_1)} \right]$$

where $h_2(z) = (s_1 + \xi_2^+(s_1))/2$, respectively.

Proof. **a)** By (8), we have $H(0,0) = \lambda_1 M_1(0) - \lambda_2 M_2(0)$; besides, (5) gives $J_2(0) = \lambda_1(1-\varrho) - \psi_1(0)$. Applying equation (10) for $s_1 = s_2 = 0$ and invoking the finiteness of $G_2(0)$ consequently implies that $J_2(0) + H(0,0) = 0$. Reduce then the latter equation to $\lambda_1 M_1(0) - \lambda_2 M_2(0) = \psi_1(0) - \lambda_1(1-\varrho)$ and combine it with identity $\psi_1(0) + \psi_2(0) = \lambda(1-\varrho)$; solving for both $\psi_1(0)$ and $\psi_2(0)$ provides the announced formulas.

b) Write $\mathbb{P}(U_1 = 0) = \mathbb{P}(U_1 = U_2 = 0) + \mathbb{P}(U_1 = 0 < U_2)$ and $\mathbb{P}(U_1 = U_2 = 0) = 1 - \varrho$; identity (34) then follows by definition (2) of G_2. Now, to calculate $G_2(0)$, apply Equation (25) for $G_2(s_2)$ with $s_2 = 0$; as $\xi_1^+(0) = 0$ and since $G_2(0)$ is finite, $G_2(0)$ is necessarily equal to the limit of the quotient expressed above. *Mutatis mutandis*, the same derivation pattern holds for $\mathbb{P}(U_2 = 0)$ and $G_1(0)$. $\square$

In Figure 2, we depict the variations of empty queue probabilities $\mathbb{P}(U_1 = 0)$ and $\mathbb{P}(U_2 = 0)$ as a function of ϱ_1, assuming the total load ϱ is fixed. Implementing formulas of Proposition 3 was easily performed with Mathematica software tool by using tree structures, as numerous iterations are necessary for computing infinite sums and products. We note that for small load ϱ_1, probability $\mathbb{P}(U_1 = 0)$ is close enough to probability $1 - \varrho_1$ that would be obtained if a fixed HoL priority scheme were applied (with queue ♯1 having highest priority). A similar situation holds for queue ♯2 when load ϱ_2 decreases. This confirms the interest of the SQF discipline to favor traffic flows with the least intensity.

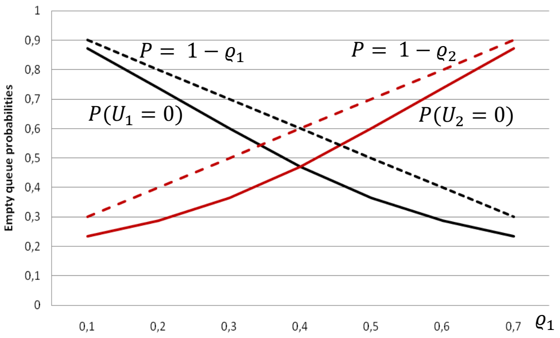

Figure 2: *Empty queue probabilities* $\mathbb{P}(U_1 = 0)$ *and* $\mathbb{P}(U_2 = 0)$ *for varying load* ϱ_1 *(*$\mu_1 = \mu_2 = 1$ *and constant total load* $\varrho_1 + \varrho_2 = 0.8$*).*

4. SMALLEST MODULE SINGULARITIES

Recall that the smallest singularity of the Laplace transform of a real random variable determines the decay rate of its distribution at infinity. We here extend the analyticity domain of functions M_1 and M_2 in order to determine the singularities with smallest module for transforms G_1 and G_2, and thus for transform F; the results of the present section will be applied in Section 5 to variables U_1 and U_2.

4.1 Analytic continuation of function M

A property is said to hold *generically* if it does for almost all $(\lambda_1, \lambda_2, \mu_1, \mu_2)$ in $\mathbb{R}^4$ with respect to Lebesgue measure. We first assert the following preliminary result; its proof is technical and is not detailed here (see [11]).

Theorem 2. *Given polynomial $R_j(w,z)$, $j \in \{1,2\}$, defined in (17)-(18), let $\Delta_j(z)$, $j \in \{1,2\}$, denote the discriminant of $R_j(w,z)$ in variable w.*

a) *Discriminant $\Delta_j(z)$ has generically four distinct roots $\eta_j^{(1)}$, ..., $\eta_j^{(4)}$, two of those roots being real negative and the two others non real (complex conjugate). Let then $\eta_j^{(1)}$, $\eta_j^{(2)}$ denote the two real roots.*

b) *Algebraic function α_1 (resp. h_1) is analytic (resp. meromorphic) in the cut-plane $\mathbb{C} \setminus [\eta_1^{(1)}, \eta_1^{(2)}]$. Symmetrically, algebraic function γ_2 (resp. h_2) is analytic (resp. meromorphic) in the cut-plane $\mathbb{C} \setminus [\eta_2^{(1)}, \eta_2^{(2)}]$.*

Recall that M_1 and M_2 are analytic at least on the half-plane $\{z \in \mathbb{C} \mid \Re(z) > \max(\widetilde{s}_1, \widetilde{s}_2)/2\}$. Using Theorem 2, we can now extend the analyticity domain of $\mathbf{M} = (M_1, M_2)$ according to the exclusive conditions (I^+), (I^-), (II^+) and (II^-) introduced in Lemma 2.

PROPOSITION 4. *Let $\tau_1 = \xi_1^-(\sigma_0^+)$ (resp. $\tau_2 = \xi_2^-(\sigma_0^+)$).*
Function $\mathbf{M}$ can be analytically extended to the half-plane $\mathbf{V}_M$ defined by
a1. *$\mathbf{V}_M = \{z \in \mathbb{C} \mid \Re(z) > \frac{1}{2}\max(\sigma_0^+ + \tau_1, \sigma_0^+ + \tau_2)\}$ if conditions (I^+) and (II^+) hold;*
a2. *$\mathbf{V}_M = \{z \in \mathbb{C} \mid \Re(z) > \max(\frac{1}{2}(\sigma_0^+ + \tau_1), \eta_2^{(1)}\}$ if conditions (I^-) and (II^+) hold;*
a3. *$\mathbf{V}_M = \{z \in \mathbb{C} \mid \Re(z) > \max(\eta_1^{(1)}, \frac{1}{2}(\sigma_0^+ + \tau_2))\}$ if conditions (I^+) and (II^-) hold;*
a4. *$\mathbf{V}_M = \{z \in \mathbb{C} \mid \Re(z) > \max(\eta_1^{(1)}, \eta_2^{(1)}\}$ if conditions (I^-) and (II^-) hold.*
In the above defined domains $\mathbf{V}_M$, the smallest abscissa is always smaller than σ_0^+.

PROOF. By using Equation (26) for $G_2(s)$ and a similar equation for $G_1(s)$, we obtain

$$M_1(z) = \frac{-\lambda_2\mu_2}{E(z)}\left[\frac{(s_1 - \xi_2^-(s_1))G_1(s_1) - J_1(s_1)}{\mu_2 + s_2} + \frac{(s_2 - \xi_1^-(s_2))G_2(s_2) - J_2(s_2)}{\mu_2 + \xi_2^-(s_1)}\right]$$

and

$$M_2(z) = \frac{-\lambda_1\mu_1}{E(z)}\left[\frac{(s_2 - \xi_1^-(s_2))G_2(s_2) - J_2(s_2)}{\mu_1 + s_1} + \frac{(s_1 - \xi_2^-(s_1))G_1(s_1) - J_1(s_1)}{\mu_1 + \xi_1^-(s_2)}\right],$$

where

$$\frac{E(z)}{\lambda_1\mu_1\lambda_2\mu_2} = \frac{1}{(\mu_1 + s_1)(\mu_2 + s_2)} - \frac{1}{(\mu_1 + \xi_1^-(s_2))(\mu_2 + \xi_2^-(s_1))}$$

and where s_1 and s_2 depend on z according to $s_1 = z + \gamma_2(z)$ and $s_2 = z - \alpha_1(z)$, respectively.

Simple algebraic arguments show that denominator $E(z)$ cannot vanish for $\Re(z) > \max(\eta_1^{(1)}, \eta_2^{(1)})$ (see [11] for details). Besides, we note by Theorem 2.b, that $z \mapsto \alpha_1(z)$ (resp. $z \mapsto \gamma_2(z)$) is analytic on $\mathbb{C}$ cut along the segment joining its real ramification points, namely the real negative roots $\eta_1^{(1)}$, $\eta_1^{(2)}$ (resp. $\eta_2^{(1)}$, $\eta_2^{(2)}$) of discriminant $\Delta_1(z)$ (resp. $\Delta_2(z)$). We hereafter assume that, for instance, inequalities $\eta_1^{(2)} < \eta_1^{(1)} < 0$ and $\eta_2^{(2)} < \eta_2^{(1)} < 0$ hold. In addition, by Lemma 1, ξ_2^- (resp. ξ_1^-) is analytic on $\mathbb{C} \setminus [\zeta_1^-, \zeta_1^+]$ (resp. $\mathbb{C} \setminus [\zeta_2^-, \zeta_2^+]$) and by Corollary 1, G_1 (resp. G_2) is analytic on $\widetilde{\omega}_1 = \{s_1 \in \mathbb{C} \mid \Re(s_1) > \widetilde{s}_1\}$ (resp. on $\widetilde{\omega}_2 = \{s_2 \in \mathbb{C} \mid \Re(s_2) > \widetilde{s}_2\}$), with $\widetilde{s}_1$ and $\widetilde{s}_2$ defined in Lemma 2.

From the expressions of $M_1(z)$ and $M_2(z)$ above and the latter properties, we deduce that $\mathbf{M}$ is analytic at any point z such that $\Re(z) > \max(\eta_1^{(1)}, \eta_2^{(1)})$ and

$$z - \alpha_1(z) > \max(\zeta_2^+, \widetilde{s}_2) \;\;\&\;\; z + \gamma_2(z) > \max(\zeta_1^+, \widetilde{s}_1). \quad (35)$$

According to which pair of conditions amongst (I^+), (II^+), (I^-) and (II^-) holds, the values in the right-hand sides of inequalities (35) are given in Table 1.

Table 1: Values of $\max(\zeta_2^+, \widetilde{s}_2)$ and $\max(\zeta_1^+, \widetilde{s}_1)$.

Case	$\max(\zeta_2^+, \widetilde{s}_2)$	$\max(\zeta_1^+, \widetilde{s}_1)$
a1. (I$^+$), (II$^+$)	σ_0^+	σ_0^+
a2. (I$^-$) (II$^+$)	σ_0^+	ζ_1^+
a3. (I$^+$) (II$^-$)	ζ_2^+	σ_0^+
a4. (I$^-$) (II$^-$)	ζ_2^+	ζ_1^+

Let us for instance consider case **a.1** (see Table 1). By examining the graphs of functions $s_1 \mapsto \frac{1}{2}(s_1 + \xi_2^{\pm}(s_2))$ and $s_2 \mapsto \frac{1}{2}(s_2 + \xi_1^{\pm}(s_1))$ (see [11], Figure 5), we can verify that Condition (35) is satisfied for $z > \frac{1}{2}\max(\sigma_0^+ + \tau_1, \sigma_0^+ + \tau_2)$ only. The other cases in Table 1 are treated in a similar manner. ☐

4.2 Smallest module singularities of G_1 and G_2

Corollary 1 ensures that G_1 (resp. G_2) has no singularity in $\{s \in \mathbb{C} \mid \Re(s) > \widetilde{s}_1\}$ (resp. in $\{s \in \mathbb{C} \mid \Re(s) > \widetilde{s}_2\}$) where thresholds $\widetilde{s}_1$ and $\widetilde{s}_2$ are specified in Lemma 2. Latter Proposition 4 will now enable us to specify the smallest singularity of transforms G_1 and G_2.

THEOREM 3. *Let constants*

$$\begin{cases} r_{0,1} = \dfrac{1}{1 - \xi_2^{+\prime}(\sigma_0^+)}\left[J_1(\sigma_0^+) - H(\sigma_0^+, \sigma_0^+)\right], \\[2ex] r_{0,2} = \dfrac{1}{1 - \xi_1^{+\prime}(\sigma_0^+)}\left[J_2(\sigma_0^+) + H(\sigma_0^+, \sigma_0^+)\right], \end{cases}$$

and

$$\begin{cases} r_1^+ = \dfrac{\sqrt{D_{0,1}(\zeta_1^+ - \zeta_1^-)}}{2(\mu_1 + \zeta_1^+)(\zeta_1^+ - a_2^+)}\left[G_1(\zeta_1^+) - \dfrac{\lambda_2\mu_2 M_2(z_2^+)}{(\mu_2 + a_2^+)^2}\right], \\[2ex] r_2^+ = \dfrac{\sqrt{D_{0,2}(\zeta_2^+ - \zeta_2^-)}}{2(\mu_2 + \zeta_2^+)(\zeta_2^+ - a_1^+)}\left[G_2(\zeta_2^+) - \dfrac{\lambda_1\mu_1 M_1(z_1^+)}{(\mu_1 + a_1^+)^2}\right] \end{cases}$$

with $D_{0,1} = 4\lambda_1\lambda_2 + (\mu_2 + \lambda_1 - \lambda_2)^2$, $z_2^+ = (\zeta_1^+ + a_2^+)/2$, and ζ_1^+, ζ_1^- given in (13) (resp. $D_{0,2} = 4\lambda_1\lambda_2 + (\mu_1 + \lambda_2 - \lambda_1)^2$, $z_1^+ = (\zeta_2^+ + a_1^+)/2$, and ζ_2^+, ζ_2^- obtained from (13) by permuting indexes 1 and 2).

*In **Cases a1, a2, a3** and **a4** of Proposition 4, the singularities with smallest module of transforms G_1 and G_2 are defined by*
a1. *a simple pole at $s_1 = \sigma_0^+$ for G_1 (resp. a simple pole at $s_2 = \sigma_0^+$ for G_2) with residue $r_{0,1}$ (resp. residue $r_{0,2}$);*
a2. *an algebraic singularity with order 1 at $s_1 = \zeta_1^+$ for G_1 (resp. a simple pole at $s_2 = \sigma_0^+$ for G_2) with residue r_1^+ (resp. residue $r_{0,2}$);*
a3. *a simple pole at $s_1 = \sigma_0^+$ for G_1 (resp. an algebraic singularity with order 1 at $s_2 = \zeta_2^+$ for G_2) with residue $r_{0,1}$ (resp. residue r_2^+);*
a4. *an algebraic singularity with order 1 at $s_1 = \zeta_1^+$ for G_1 (resp. an algebraic singularity with order 1 at $s_2 = \zeta_2^+$ for G_2) with residue r_1^+ (resp. residue r_2^+).*

PROOF. Consider the following cases:

- ***Case*** (II^+). As $s_2 \to \sigma_0^+$, we have $\xi_1^+(s_2) \to \sigma_0^+$ while $h_1(z) = \frac{1}{2}(s_2 + \xi_1^+(s_2)) \to \sigma_0^+$. Following Proposition 4.**a1**, functions $M_1 \circ h_2$ and $M_2 \circ h_2$ are analytic at $z = \sigma_0^+$ since $\frac{1}{2}\max(\sigma_0^+ + \tau_1, \sigma_0^+ + \tau_2) < \sigma_0^+$. By Corollary 1, $G_2(s_2)$ has presently no singularity for $\Re(s_2) > \widetilde{s}_2 = \sigma_0^+$; we then conclude from equation (25) that G_2 has a simple pole at $s = \sigma_0^+$ with residue $r_{0,2}$;
- ***Case*** (II^-). Letting $s_2 \to \sigma_0^+$, we have $\xi_1^+(s_2) \to \tau_1$ and thus $h_1(z) = \frac{1}{2}(s_2 + \xi_1^+(s_1)) \to z_{0,1}$ where $2z_{0,1} = \sigma_0^+ + \tau_1$. Proposition 4 then ensures that $M_1 \circ h_1$ and $M_2 \circ h_1$ are analytic at $z = z_{0,1}$ since $z_{0,1} > \max(z_{0,2}, \eta_1^{(2)})$ (in fact, we clearly have $z_{0,1} > \eta_1^{(2)}$ by examining the graph of function $s_2 \mapsto \frac{1}{2}(s_2 + \xi_1^+(s_2))$; besides, our assumption $\tau_2 < \sigma_0^+ < \tau_1$ implies $z_{0,1} > z_{0,2}$). We conclude from (25) and the latter discussion that σ_0^+ is not a singularity of G_2. Furthermore, second condition (35) ensures that M_1 and M_2 are analytic at any z for which $\Re(z - \alpha_1(z)) > \max(\zeta_2^+, \widetilde{s}_2) = \zeta_2^+$; by equation (25) again, we conclude that G_2 is analytic at any point s_2 for which $\Re(s_2) > \zeta_2^+$.

To specify the nature of point ζ_2^+ for G_2, use then formula (12) for $\xi_1^+(s_2)$, where discriminant $D_2(s_1)$ is written as $D_1(s_1) = D_{0,2}(s_2 - \zeta_2^-)(s_1 - \zeta_2^+)$; besides, note that $z = (s_2 + \xi_1^-(s_2))/2$ tends to $z_1^+ = (\zeta_2^+ + a_1^+)/2$ as $s_2 \to \zeta_2^+$, where $a_1^+ = \xi_1^{\pm}(\zeta_2^+)$; by Equation (25), we then obtain $G_2(s_2) = G_2(\zeta_2^+) + r_2^+(s_2 - \zeta_2^+)^{1/2} + o(s_2 - \zeta_2^+)^{1/2}$ after some simple algebra, with constants

$$r_2^+ = \frac{E_{0,2}}{\zeta_2^+ - a_1^+}\left[G_2(\zeta_2^+) - \frac{\lambda_1 \mu_1 M_1(z_1^+)}{(\mu_1 + a_1^+)^2}\right]$$

and $E_{0,2} = [D_{0,2}(\zeta_2^+ - \zeta_2^-)]^{1/2}/2(\mu_2 + \zeta_2^+)$. We conclude that the singularity with smallest module of G_2 is ζ_2^+, an algebraic singularity with order 1 and residue r_2^+.

- Cases (I^+) and (I^-) for transform G_2 are similarly treated, *mutatis mutandis*.

Mixed cases ***a1***, ***a2***, ***a3*** and ***a4*** are then readily derived from the above discussion. □

5. LARGE QUEUE ASYMPTOTICS

Once the smallest singularities of G_1 and G_2 have been determined, we can eventually specify the tail behavior for the distribution of workloads U_1 and U_2. Applying definition (3) to $s_2 = 0$ first gives the Laplace transform of U_1 as

$$F(s_1,0) = 1 - \varrho + F_1(s_1,0) + G_1(s_1) + F_2(s_1,0) + G_2(0). \quad (36)$$

THEOREM 4. *For large u_1, we have*

$$\mathbb{P}(U_1 > u_1) \sim \begin{cases} -\dfrac{(\sigma_0^+ + \mu_1)r_{0,1}}{\lambda_1 \sigma_0^+} \cdot e^{\sigma_0^+ u_1} & \text{if } (I^+) \text{ holds,} \\ \dfrac{(\zeta_1^+ + \mu_1)r_1^+}{2\lambda_1 \zeta_1^+ \sqrt{\pi}} \cdot \dfrac{e^{\zeta_1^+ u_1}}{u_1^{3/2}} & \text{if } (I^-) \text{ holds} \end{cases} \quad (37)$$

with residues $r_{0,1}$ and r_1^+ introduced in Theorem 3.

For large u_2, asymptotics for $\mathbb{P}(U_2 > u_2)$ is similarly derived, replacing conditions (I^+) and (I^-) by (II^+) and (II^-), respectively, and exchanging indexes 1 and 2 in (37).

PROOF. Assume first that (I^+) holds. By Prop. 4, the expression (36) of $H(s_1,0)$ in terms of M_1 and M_2 shows that $s_1 \mapsto H(s_1,0)$ is analytic for $\Re(s_1) > \max(\sigma_0^+ + \tau_1, \sigma_0^+ + \tau_2)$ and $\Re(s_1) > \max(2\eta_1^{(1)}, \sigma_0^+ + \tau_2)$, both conditions encompassing point $s_1 = \sigma_0^+$. We then deduce from (36) that the singularity with smallest module of $F(s_1,0)$ is at $s_1 = \sigma_0^+$ with leading term

$$F(s_1,0) \sim -\frac{K_2(s_1,0)}{K_1(s_1,0)}G_1(s_1) + G_1(s_1) = \frac{s_1 + \mu_1}{\lambda_1}G_1(s_1) \quad (38)$$

since by definition $K_1(s_1,0)/K_2(s_1,0) = -(s_1 + \mu_1 - \lambda_1)/\lambda_1$ and the root $\lambda_1 - \mu_1$ of $K_1(s_1,0)$ is less than σ_0^+ (in fact, $P(\lambda_1 - \mu_1) = -\lambda_1\lambda_2 < 0$ by (14)). By Theorem 3, the point $s_1 = \sigma_0^+$ is a simple pole for G_1 with residue $r_{0,1}$; applying [7, Th. 25.2, p 237] to asymptotics (38), we derive the asymptotics for $\mathbb{P}(U_1 > u)$ with large u, as claimed.

Assume now that (I^-) holds. By examining the graph of the function $s_1 \mapsto \frac{1}{2}(s_1 + \xi_2^+(s_1))$, it is easily checked that

$$\frac{\zeta_1^+ + a_2^+}{2} > \sigma_0^+ > \max(\eta_1^{(1)}, \eta_2^{(1)})$$

and hence $\zeta_1^+/2 > \sigma_0^+ > \max(\eta_1^{(1)}, \eta_2^{(1)})$ since $a_2^+ < 0$. By Proposition 4 and the expression (36) of $H(s_1,0)$ in terms of M_1 and M_2, the latter inequalities ensure that function $s_1 \mapsto H(s_1,0)$ is analytic at $s_1 = \zeta_1^+$. It then follows from (36) that the singularity with smallest module of $F(s_1,0)$ is at $s_1 = \zeta_1^+$ with leading term provided by (38) again, so that

$$F(s_1,0) - F(\zeta_1^+,0) \sim \frac{\zeta_1^+ + \mu_1}{\lambda_1}\left[G_1(s_1) - G_1(\zeta_1^+)\right] \quad (39)$$

near $s_1 = \zeta_1^+$. By Theorem 3, $s_1 = \zeta_1^+$ is an algebraic singularity for G_1 with residue r_1^+; (39) then gives the asymptotics $F(s_1,0) - F(\zeta^+,0) \sim r_1(s_1 - \zeta_1^+)^{1/2}$ as $s_1 \to \zeta_1^+$ with constant $r_1 = (\zeta_1^+ + \mu_1)r_1^+/\lambda_1$. Applying again [7, Th. 25.2, p 237] to the latter asymptotics, we then derive that $\mathbb{P}(U_1 > u) \sim \kappa_1 e^{\zeta_1^+ u}/u^{3/2}$ for large u with prefactor κ_1 given by $\kappa_1 = -r_1/\zeta_1^+\Gamma(-1/2) = r_1/2\zeta_1^+\sqrt{\pi}$, as claimed. □

Theorem 4 eventually specifies the tail behavior for the respective distributions of workloads U_1 and U_2, depending on parameter values λ_1, μ_1, λ_2, μ_2 with $\varrho_1 + \varrho_2 < 1$; in a summarized form, we have shown that $\mathbb{P}(U_1 > u_1) = O(e^{\widetilde{s}_1 u_1})$ and $\mathbb{P}(U_2 > u_2) = O(e^{\widetilde{s}_2 u_2})$ with decay rates $\widetilde{s}_1$ and $\widetilde{s}_2$ given by Lemma 2. Specifically, Case ***a1*** gives exponential decay at infinity to both queues with identical rate σ_0^+, while last Case ***a4*** corresponds to sub-exponential decays with respective rate ζ_1^+ and ζ_2^+; finally, Case ***a2*** and Case ***a3*** correspond to mixed exponential / sub-exponential behaviors.

In Fig.3, we draw the regions of the (ϱ_1, ϱ_2)-plane corresponding to cases **a1**, **a2**, **a3** and **a4** for either $\mu_1 = \mu_2$ (with $\lambda_1 \neq \lambda_2$) or $\lambda_1 = \lambda_2$ (with $\mu_1 \neq \mu_2$); it is easily verified from Lemma 2 that the symmetric boundary curves have equations $\varrho_2 = h(\varrho_1)$ and $\varrho_1 = h(\varrho_2)$ with $h(r) = \sqrt{r}(1 - \sqrt{r})$ if $\mu_1 = \mu_2$, and $h(r) = r(1 - \sqrt{r})/(1 - 2\sqrt{r} + 2r)$ if $\lambda_1 = \lambda_2$, respectively. For instance, assume that $\mu_1 = \mu_2$ and queue

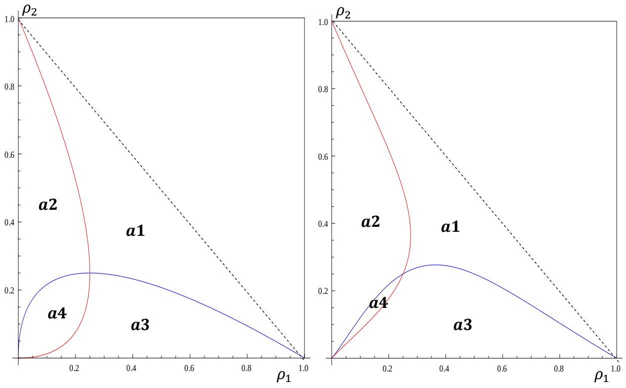

Figure 3: ***Regions of the*** $(O, \varrho_1, \varrho_2)$***-plane associated with Cases a1, a2, a3, a4 for asymmetric queues with*** $\lambda_1 \neq \lambda_2$ ***and*** $\mu_1 = \mu_2$ ***(left) or*** $\lambda_1 = \lambda_2$ ***and*** $\mu_1 \neq \mu_2$ ***(right).***

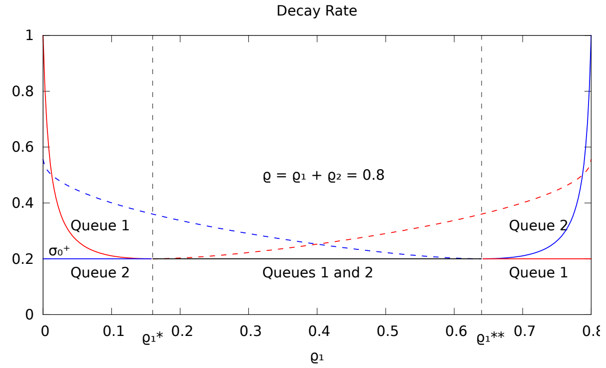

Figure 4: ***Decay rates associated with Queue*** ♯1 ***and Queue*** ♯2 ***as a function of load*** ϱ_1 $(\varrho_1 + \varrho_2 = 0.8)$***.***

♯1 receives low intensity traffic, i.e. ϱ_1 tends to 0; by Fig.3, this corresponds to

- either **Case *a2***, where we have $\mathbb{P}(U_1 > u_1) = O(e^{\zeta_1^+ u_1})$ and $\mathbb{P}(U_2 > u_2) = O(e^{\sigma_0^+ u_2})$ with $-\zeta_1^+ > -\sigma_0^+$, so that queue ♯1 is smaller than queue ♯2 regarding the sharpness of distribution tails;
- or **Case *a4***, where we have $\mathbb{P}(U_1 > u_1) = O(e^{\zeta_1^+ u_1})$ and $\mathbb{P}(U_2 > u_2) = O(e^{\zeta_2^+ u_2})$. By formulas (13), supplementary condition $\zeta_1^+ < \zeta_2^+$ easily reduces to $\varrho_1 < \varrho_2$, which is clearly fulfilled for small ϱ_1; queue ♯1 is then smaller than queue ♯2 in the same sense.

The dynamic SQF discipline consequently provides priority to the queue with less traffic intensity, as motivated by its definition.

As a final illustration, we draw the values of decay rates associated with queues ♯1 and ♯2 as continuous functions of the system load. Assuming $\mu_1 = \mu_2 = 1$ with total load $\varrho = 0.8$, the decay rate associated with queue ♯1 (resp. queue ♯2) then equals $-\zeta_1^+ = f(\varrho_1)$ with

$$f(r) = \frac{(1 - \sqrt{\varrho - r})^2}{r + (1 - \sqrt{\varrho - r})^2}$$

(resp. $-\sigma_0^+ = 1 - \varrho$) by formula (13) as long as (ϱ_1, ϱ_2) remains in the region **a2**, that is, for $0 < \varrho_1 < \varrho_1^* = \varrho(1-\varrho)$; both queues have decay rate $-\sigma_0^+$ within region **a1**, that is, for $\varrho_1^* < \varrho_1 < \varrho_1^{**} = \varrho^2$; finally, queue ♯1 (resp. queue ♯2) has decay rate $-\sigma_0^+$ (resp. $-\zeta_2^+ = f(\varrho - \varrho_1)$) within region **a3**, that is, for $\varrho_1^{**} < \varrho_1 < \varrho$ (see Fig.4; the applicable values of decay rates are that of curves with thick lines).

6. CONCLUSION

As a generalization to the static HoL priority scheme, the SQF discipline provides a dynamic scheme for controlling traffic congestion in favor of less congested queues. Within the Markovian framework, its mathematical analysis involves a challenging new setting, namely functional equations whose solutions expand as a series involving the semi-group generated by two algebraic functions h_1 and h_2; such functions prove to be naturally attached to a pair of rational cubics. The resolution method developed in this paper has enabled us to derive essential performance characteristics such as empty queue probabilities and queue asymptotics.

7. REFERENCES

[1] N. Benameur, F. Guillemin, and L. Muscariello. Latency reduction in home access gateways with Shortest Queue First. In *Proc. ISOC Workshop on Reducing Internet Latency*, London, September 2013.

[2] T. Bonald and L. Muscariello. Shortest Queue First: implicit service differentiation through per-flow scheduling. Demo. at IEEE LCN, 2009.

[3] T. Bonald, L. Muscariello, and N. Ostallo. Self-prioritization of Audio and Video Traffic. *Proceedings of ICC 2011*, 2011.

[4] G. Carofiglio and L. Muscariello. On the impact of TCP and per-flow scheduling on Internet performance. *IEEE/ACM Transactions on Networking*, 2011.

[5] J. Cohen. A two-queue, one-server model with priority for the longest queue. *Queueing Systems: Theory and Applications*, 2:261 – 283, 1987.

[6] J. Cohen. Analysis of the asymmetrical shortest two-server queuing model. *J. of Appl. Math. and Stoch. Analysis*, 11(2):115 – 162, 1998.

[7] G. Doetsch. *Einführung in Theorie und Anwendung der Laplace Transformation.* Birkhäuser, 1958.

[8] F.Guillemin and A.Simonian. Stationary analysis of the Shortest queue First service policy. *Queueing Systems: Theory and Applications*, 2013.

[9] J. Gettys. Bufferbloat: Dark buffers in the Internet. *IEEE Internet Computing*, 15(3):95 – 96, 2100.

[10] F. Guillemin and R. Mazumdar. Rate conservation laws for multidimensional processes of bounded variation with application to priority queueing systems. *Methodology and Computing in Applied Probability*, 6:135–149, 2004.

[11] F. Guillemin and A. Simonian. Stationary analysis of the SQF service policy: the asymmetric case. Preprint available at http://arxiv.org/abs/1305.3496.