The morphing of fluid queues into Markov-modulated Brownian motion

Guy Latouche

Université libre de Bruxelles
Département d'Informatique
e-mail: guy.latouche@ulb.ac.be

Giang Nguyen*

The University of Adelaide
School of Mathematical Sciences
e-mail: giang.nguyen@adelaide.edu.au

Abstract: Ramaswami showed recently that standard Brownian motion arises as the limit of a family of Markov-modulated linear fluid processes. We pursue this analysis with a fluid approximation for Markov-modulated Brownian motion. We follow a Markov-renewal approach and we prove that the stationary distribution of a Markov-modulated Brownian motion reflected at zero is the limit from the well-analyzed stationary distribution of approximating linear fluid processes. Thus, we provide a new approach for obtaining the stationary distribution of a reflected MMBM without time-reversal or solving partial differential equations. Our results open the way to the analysis of more complex Markov-modulated processes.

Key matrices in the limiting stationary distribution are shown to be solutions of a matrix-quadratic equation, and we describe how this equation can be efficiently solved.

AMS 2000 subject classifications: 60J25, 60J65, 60B10.
Keywords and phrases: Markov-modulated linear fluid models, Markov-modulated Brownian motion, weak convergence, stationary distribution, computational methods.

1. Introduction

Our purpose is to construct and analyse a family of fluid queues converging to Markov-modulated Brownian motion (MMBM) with the intention of adapting, to the analysis of MMBM, tools and methods which have been developed in the context of fluid queues.

Fluid queues are two-dimensional processes $\{X(t), \varphi(t) : t \geq 0\}$, where $\{\varphi(t) :$

*The authors had interesting discussions with V. Ramaswami over the questions examined in this paper. They thank three anonymous referees for their careful analysis and insightful suggestions, which helped streamline some of the proofs and improve the overall presentation of the paper. They also thank the Ministère de la Communauté française de Belgique for funding this research through the ARC grant AUWB-08/13–ULB 5, and acknowledge the financial support of the Australian Research Council through the Discovery Grant DP110101663.

$t \geq 0\}$ is a continous-time Markov chain on a finite state space $\mathcal{S}$,

$$X(t) = X(0) + \int_0^t c_{\varphi(t)} \, \mathrm{d}t,$$

and c_i for $i \in \mathcal{M}$ are arbitrary real numbers. These are also known as *Markov-modulated linear fluid processes*, with X referred to as the *fluid level* and φ as the *phase*: during intervals of time where the phase φ remains in some state $i \in \mathcal{M}$, the fluid level varies linearly at the rate c_i. The associated process reflected at zero is denoted by $\{\widehat{X}(t), \varphi(t) : t \geq 0\}$, where

$$\widehat{X}(t) = X(t) - \inf_{0 \leq v \leq t} X(v),$$

assuming that $\widehat{X}(0) = 0$. Also referred to as first-order fluid processes, these stochastic models are useful when the relevant rates of change can be well-described by their first moments.

Markov-modulated Brownian motion (MMBM), or the family of second-order fluid processes, takes into account first and second moments of the rates of change. In particular, the fluid level $Y(t)$ of a Markov-modulated Brownian motion $\{Y(t), \kappa(t) : t \geq 0\}$ is a Brownian motion with drift μ_i and variance σ_i^2 during time intervals where $\kappa(t) = i$; one sometimes writes that

$$Y(t) = Y(0) + \int_0^t c_{\kappa(t)} \, \mathrm{d}t + \int_0^t \sigma_{\kappa(t)} \, \mathrm{d}W(t)$$

where $\{W(t)\}$ is a standard Brownian motion.

Three papers appeared in close succession on the stationary distribution of MMBM reflected at zero: Rogers [25], Asmussen [4], and Karandikar and Kulkarni [17]. The focus in the third paper is on solving partial differential equations, and it is not of further concern to us in the present paper. In [4, 25], on the other hand, the authors obtain the stationary distribution in a form which is suitable for calculations with linear algebraic procedures. These results crucially depend on the technique of reversing time, a method already used in Loynes [20] whereby the stationary distribution of the process is obtained from the distribution of the maximum of a random walk with negative drift. More recent work for obtaining the stationary distribution of Markov-modulated Lévy processes with reflecting boundaries, as in Asmussen and Kella [5], Ivanovs [16], D'Auria *et al.* [11], and D'Auria and Kella [12], also uses the reverse-time approach.

For fluid processes without a Brownian component, another line of investigation was open in Ramaswami [23], based on renewal-type arguments similar to the ones used in the analysis of quasi-birth-and-death processes (Neuts [21, Chapter 3], Latouche and Ramaswami [19, Chapter 6]). This eventually led, in addition to interesting algorithmic procedures, to theoretical developments for fluid processes in finite time, and for systems with more complex interactions between phase and level. There, the flow of time is not reversed and this creates a significant difference in the methods of analysis.

Our intention is to establish a link between the results for fluid queues and those for MMBM. In Section 2, we extend the argument from Ramaswami [24] and define a parameterised family of linear fluid processes that converge weakly, as the parameter tends to infinity, to a Markov-modulated Brownian motion. Ahn and Ramaswami [1] are independently using matrix-analytic methods to analyze Markov-modulated Brownian motions, with different approaches to ours. We determine in Sections 3 the limiting structure of key matrices and quadratic matrix equations, and we establish the connection between the stationary distribution so obtained, and the one which follows from the time-reversed approach. We present in Section 4 a computational procedure for solving the quadratic equation efficiently.

2. Markov-modulated Brownian motion

We show here that a family of linear fluid processes converges weakly to a Markov-modulated Brownian motion $\{Y(t), \kappa(t) : t \geq 0\}$, where the phase process κ is a Markov chain with state space $\mathcal{M} = \{1, \ldots, m\}$, and Y is a Brownian motion with drift μ_i and variance σ_i^2 whenever $\kappa(t) = i \in \mathcal{M}$. We denote by D the drift matrix $\mathrm{diag}(\mu_1, \ldots, \mu_m)$, by V the variance matrix $\mathrm{diag}(\sigma_1^2, \ldots, \sigma_m^2)$, and by Q the generator of κ, and we assume that Q is irreducible.

Assumption 2.1. *At time 0, the level $Y(0)$ is equal to 0, and the initial phase $\kappa(0)$ has the stationary distribution $\boldsymbol{\alpha}$ of Q ($\boldsymbol{\alpha} Q = \mathbf{0}$ and $\boldsymbol{\alpha}\mathbf{1} = 1$).*

Formally, for $t \geq 0$ the process $Y(t)$ can be defined recursively as

$$Y(t) = Y(T) + \sum_{i \in \mathcal{M}} \mathbb{1}_{\{\kappa(T)=i\}} \{Y_i(t) - Y_i(T)\}, \tag{1}$$

where $Y(0) = 0$, the random variable T is the last jump epoch of κ before t ($T = 0$ if there has yet to be a jump), the process Y_i is a Brownian motion with mean μ_i and variance σ_i^2, and $\mathbb{1}_{\{\cdot\}}$ denotes the indicator function. The processes Y_i for al $i \in \mathcal{M}$ and φ are assumed to be mutually independent.

We construct the family of fluid processes $\{L_\lambda(t), \beta_\lambda(t), \varphi_\lambda(t) : t \geq 0\}$ as follows: the phase process is a two-dimensional Markov chain $(\beta_\lambda(t), \varphi_\lambda(t))$ with state space $\mathcal{S} = \{(k, i) : k \in \{1, 2\} \text{ and } i \in \mathcal{M}\}$ and generator

$$T_\lambda = \begin{bmatrix} Q - \lambda I & \lambda I \\ \lambda I & Q - \lambda I \end{bmatrix},$$

where the entries of T_λ follow the lexicographic ordering of $\{1, 2\} \times \mathcal{M}$, and I denotes an appropriately-sized identity matrix. Whenever ambiguity might arise, we write I_n to denote the $n \times n$ identity matrix. The fluid rate matrix C_λ is given by

$$C_\lambda = \begin{bmatrix} D + \sqrt{\lambda}\Theta & \\ & D - \sqrt{\lambda}\Theta \end{bmatrix}, \quad \text{where } \Theta = \sqrt{V}.$$

Assumption 2.2. *At time* 0*, the level* $L_\lambda(0)$ *is equal to* 0*, and the initial phases* $\beta_\lambda(0)$ *and* $\varphi_\lambda(0)$ *have their respective stationary distributions* $\boldsymbol{\gamma} = (1/2, 1/2)$ *and* $\boldsymbol{\alpha}$*, their joint distribution is* $\boldsymbol{p} = \boldsymbol{\gamma} \otimes \boldsymbol{\alpha}$*.*

Intuitively speaking, we duplicate the state space $\mathcal{M}$ in the Markov-modulated Brownian motion $\{Y(t), \kappa(t)\}$ and the auxiliary process $\beta_\lambda(t)$ keeps track of which copy is in use. Note that for λ sufficiently large, the phases in the copy with $\beta_\lambda(t) = 1$ have all positive rates while the phases in the other copy have all negative rates. With this construction, we show that the conditional moment generating function of $\{L_\lambda(t), \varphi_\lambda(t)\}$ converges to that of $\{Y(t), \kappa(t)\}$.

Remark 2.3. Based on the recursive definition (1) of Y, an alternative interpretation for our fluid-based approximation is that for each phase $i \in \mathcal{M}$ we approximate the process Y_i by a two-phase fluid process $\{L^i_\lambda(t), \beta^i_\lambda(t)\}$ the phase β^i_λ is a Markov chain with state space $\mathcal{S}^i = \{1, 2\}$ and generator

$$T^i_\lambda = \begin{bmatrix} -\lambda & \lambda \\ \lambda & -\lambda \end{bmatrix},$$

and the fluid rate matrix C^i_λ is given by

$$C^i_\lambda = \begin{bmatrix} \mu_i + \sqrt{\lambda}\sigma_i & \\ & \mu_i - \sqrt{\lambda}\sigma_i \end{bmatrix},$$

the approximating processes being independent.

Denote by $\Delta_Y(s)$ the $m \times m$ Laplace matrix exponent of $\{Y(t), \kappa(t)\}$, by $\tilde{\Delta}_\lambda(s)$ the $2m \times 2m$ Laplace matrix exponent of $\{L_\lambda(t), \beta_\lambda(t), \varphi_\lambda(t)\}$, and by $\Delta_\lambda(s)$ the $m \times m$ Laplace matrix exponent of $\{L_\lambda(t), \varphi_\lambda(t)\}$. These matrices are such that

$$\begin{aligned}
[e^{\Delta_Y(s)t}]_{ij} &= \mathrm{E}[e^{sY(t)} \mathbb{1}_{\{\kappa(t)=j\}} | Y(0) = 0, \kappa(0) = i], \\
[e^{\tilde{\Delta}_\lambda(s)t}]_{(k,i)(k',j)} &= \mathrm{E}[e^{sL_\lambda(t)} \mathbb{1}_{\{\beta_\lambda(t)=k', \varphi_\lambda(t)=j\}} | L_\lambda(0) = 0, \beta_\lambda(0) = k, \varphi_\lambda(0) = i],
\end{aligned}$$

and

$$[e^{\Delta_\lambda(s)t}]_{ij} = \mathrm{E}[e^{sL_\lambda(t)} \mathbb{1}_{\{\varphi_\lambda(t)=j\}} | L_\lambda(0) = 0, \varphi_\lambda(0) = i]. \tag{2}$$

By Asmussen and Kella [5], the Laplace matrix exponent of a Markov-modulated Lévy process $\{Z(t), \xi(t)\}$ with jumps is given by

$$\Delta_Z(s) = \mathrm{diag}(\phi_1(s), \ldots, \phi_p(s)) + Q \circ R(s),$$

where $\phi_i(s)$ is the Laplace exponent of an unmodulated Lévy process with parameters defined for phase $i \in \{1, \ldots, p\}$, Q is the phase-transition matrix of $\xi(t)$, R is the matrix with components $[R(s)]_{ij} = \mathrm{E}[e^{sW_{ij}}]$, which are the Laplace transforms of the jumps W_{ij} for $Z(t)$ when $\xi(t)$ moves from i to j, and $\circ$ indicates the Hadamard product.

As the Laplace exponent of an unmodulated Brownian motion with drift μ and variance σ^2 is given by $\mu s + \sigma^2 s^2/2$, one can verify that

$$\begin{aligned}
\Delta_Y(s) &= sD + (s^2/2)V + Q, \\
\tilde{\Delta}_\lambda(s) &= sC_\lambda + T_\lambda \\
&= I \otimes M + s\sqrt{\lambda} J \otimes \Theta + \lambda G \otimes I,
\end{aligned}$$

where $M = sD + Q$, and

$$e^{\Delta_\lambda(s)t} = (\boldsymbol{\gamma} \otimes I) e^{\tilde{\Delta}_\lambda(s)t} (\mathbf{1} \otimes I), \tag{3}$$

where $\mathbf{1}$ is an appropriately-sized column vector of ones.

The next lemma gives a technical property of the matrix exponential, it is used in the proof of Theorem 2.5.

Lemma 2.4. *Let S be the block-partitioned matrix*

$$S = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}$$

where S_{11} and S_{22} are matrices of order m_1 and m_2, respectively. Denote by $H(t)$ the north-west quadrant of order m_1 of e^{St}:

$$H(t) = \begin{bmatrix} I_{m_1 \times m_1} & 0 \end{bmatrix} e^{St} \begin{bmatrix} I_{m_1 \times m_1} \\ 0 \end{bmatrix}.$$

The matrix $H(t)$ is the solution of

$$H(t) = e^{S_{11}t} + \int_0^t \int_v^t e^{S_{11}(t-u)} S_{12} e^{S_{22}(u-v)} S_{21} H(v) \, \mathrm{d}u \, \mathrm{d}v. \tag{4}$$

Proof. We decompose S as the sum $S = S_A + S_E$, with

$$S_A = \begin{bmatrix} S_{11} & S_{12} \\ 0 & S_{22} \end{bmatrix} \qquad \text{and} \qquad S_E = \begin{bmatrix} 0 & 0 \\ S_{21} & 0 \end{bmatrix} = \begin{bmatrix} 0 \\ I \end{bmatrix} S_{21} \begin{bmatrix} I & 0 \end{bmatrix}.$$

By Higham [15, Equations (10:13) and (10:40)], we obtain

$$e^{St} = e^{S_A t} + \int_0^t e^{S_A(t-v)} S_E e^{S(v)} \, \mathrm{d}v.$$

and

$$e^{S_A t} = \begin{bmatrix} e^{S_{11}t} & \int_0^t e^{S_{11}(t-u)} S_{12} e^{S_{22}u} \, \mathrm{d}u \\ 0 & e^{S_{22}t} \end{bmatrix}$$

Thus,

$$\begin{aligned}
H(t) &= \begin{bmatrix} I & 0 \end{bmatrix} e^{S_A t} \begin{bmatrix} I \\ 0 \end{bmatrix} + \int_0^t \begin{bmatrix} I & 0 \end{bmatrix} e^{S_A(t-v)} S_E e^{S(v)} \begin{bmatrix} I \\ 0 \end{bmatrix} \, \mathrm{d}v \\
&= e^{S_{11}t} + \int_0^t \begin{bmatrix} I & 0 \end{bmatrix} e^{S_A(t-v)} \begin{bmatrix} 0 \\ I \end{bmatrix} S_{21} \begin{bmatrix} I & 0 \end{bmatrix} e^{S(v)} \begin{bmatrix} I \\ 0 \end{bmatrix} \, \mathrm{d}v \\
&= e^{S_{11}t} + \int_0^t \int_0^{t-v} e^{S_{11}(t-v-u)} S_{12} e^{S_{22}u} S_{21} H(v) \, \mathrm{d}u \, \mathrm{d}v
\end{aligned}$$

which proves (4). □

Theorem 2.5. *The conditional moment generating function of* $\{L_\lambda(t), \varphi_\lambda(t)\}$ *converges to that of* $\{Y(t), \kappa(t)\}$*, that is,*

$$\lim_{\lambda\to\infty} (\boldsymbol{\gamma} \otimes I) e^{\tilde{\Delta}_\lambda(s)t} (\mathbf{1} \otimes I) = e^{\Delta_Y(s)t}. \tag{5}$$

Proof. We proceed in three steps. First, we observe that

$$\tilde{\Delta}_\lambda^k(s)(\mathbf{1} \otimes I) = \mathbf{1} \otimes A_k + \boldsymbol{e} \otimes B_k, \qquad \text{for } k \geq 0, \tag{6}$$

where $\boldsymbol{e} = (1, -1)^{\mathrm{T}}$ and

$$\begin{bmatrix} A_k \\ B_k \end{bmatrix} = \Upsilon^k \begin{bmatrix} I \\ 0 \end{bmatrix}, \tag{7}$$

with

$$\Upsilon = \begin{bmatrix} M & s\sqrt{\lambda}\Theta \\ s\sqrt{\lambda}\Theta & M - 2\lambda I \end{bmatrix}.$$

The proof of (6) is by induction: that equation trivially holds for $k = 0$, with $A_0 = I$ and $B_0 = 0$ and, if it also holds for a given value of k, then we easily verify that $\tilde{\Delta}_\lambda^{k+1}(s)(\mathbf{1} \otimes I) = \mathbf{1} \otimes A_{k+1} + \boldsymbol{e} \otimes B_{k+1}$ with

$$A_{k+1} = MA_k + s\sqrt{\lambda}\Theta B_k, \qquad B_{k+1} = s\sqrt{\lambda}\Theta A_k + (M - 2\lambda I)B_k,$$

or

$$\begin{bmatrix} A_{k+1} \\ B_{k+1} \end{bmatrix} = \begin{bmatrix} M & s\sqrt{\lambda}\Theta \\ s\sqrt{\lambda}\Theta & M - 2\lambda I \end{bmatrix} \begin{bmatrix} A_k \\ B_k \end{bmatrix}.$$

Equation (7) readily follows.

To simplify the notation, we define $H_\lambda(t) = e^{\Delta_\lambda(s)t}$ for the remainder of the proof. By (3), we have

$$\begin{aligned} H_\lambda(t) &= \frac{1}{2}(\mathbf{1}^{\mathrm{T}} \otimes I) \sum_{k=0}^{\infty} \frac{t^k}{k!} \tilde{\Delta}_\lambda^k(s)(\mathbf{1} \otimes I) \\ &= \sum_{k\geq 0} \frac{t^k}{k!} A_k \qquad\qquad \text{from (6)} \\ &= \begin{bmatrix} I & 0 \end{bmatrix} \exp \Upsilon t \begin{bmatrix} I \\ 0 \end{bmatrix}. \end{aligned}$$

By Lemma 2.4, $H_\lambda(t)$ is a solution of

$$H_\lambda(t) = e^{Mt} + s^2\lambda \int_0^t \left\{ \int_v^t e^{M(t-u)} \Theta e^{(M-2\lambda I)(u-v)} \, \mathrm{d}u \right\} \Theta H_\lambda(v) \, \mathrm{d}v. \tag{8}$$

Integrating by parts the inner integral, we find

$$\begin{aligned}
&\lambda \int_v^t e^{M(t-u)} \Theta e^{(M-2\lambda I)(u-v)} \, \mathrm{d}u \\
&\quad = \Big[\lambda e^{M(t-u)} \Theta e^{(M-2\lambda I)(u-v)} (M-2\lambda I)^{-1}\Big]_v^t \\
&\qquad + \lambda \int_v^t M e^{M(t-u)} \Theta e^{(M-2\lambda I)(u-v)} (M-2\lambda I)^{-1} \, \mathrm{d}u \\
&\quad = \Theta e^{(M-2\lambda I)(t-v)} (1/\lambda M - 2I)^{-1} - e^{M(t-v)} \Theta (1/\lambda M - 2I)^{-1} \\
&\qquad + \int_v^t M e^{M(t-u)} \Theta e^{(M-2\lambda I)(u-v)} (1/\lambda M - 2I)^{-1} \, \mathrm{d}u,
\end{aligned}$$

which converges to $1/2 e^{M(t-v)} \Theta$ as $\lambda \to \infty$. From (8), we conclude that

$$H_\infty(t) = e^{Mt} + (s^2/2) \int_0^t e^{M(t-v)} V H_\infty(v) \, \mathrm{d}v,$$

and therefore

$$\frac{\mathrm{d}}{\mathrm{d}t} H_\infty(t) = (M + (s^2/2)V) H_\infty(t) = \Delta_Y(s) H_\infty(t).$$

This completes the proof. ☐

The Laplace matrix exponent uniquely characterizes the finite-dimensional distributions of the process and therefore Theorem 2.5 implies the following result.

Corollary 2.6. *The finite-dimensional distributions of* $\{L_\lambda(t), \varphi_\lambda(t)\}$ *converge to the finite-dimensional distributions of the Markov-modulated Brownian motion* $\{Y(t), \kappa(t)\}$. ☐

Theorem 2.7. *The family* $\{L_\lambda(t), \varphi_\lambda(t) : t \geq 0\}$ *is tight.*

Proof. In this proof, we use alternative interpretation of our fluid-based approximation as outlined in Remark 2.3.

The proof of [24, Theorem 5] shows that for each $i \in \mathcal{M}$ the family of marginal processes $\{L_\lambda^i\}$ is tight on any compact interval of $[0, \infty)$. By [26, Corollary 7], we can extend this result to show that the family of marginal processes $\{L_\lambda^i\}$ is tight on $[0, \infty)$.

As the family of $\{L_\lambda^i\}$ converges weakly to the Brownian motion $\{Y^i\}$ with μ_i and σ_i, so do the processes $\{\sup_{s<t} |L_\lambda^i(t)|\}$ to the corresponding supremum process $\{\sup_{s<t} |Y^i(t)|\}$, by the continuous mapping theorem. Thus, the family $\{\sup_{s<t} |Y^i(t)|\}$ is tight on $[0, \infty)$.

Next, we define $\widetilde{L}_\lambda(t)$ for $t \geq 0$ as

$$\widetilde{L}_\lambda(t) = \sum_{i=1}^m \sup_{s \leq t} |L_\lambda^i(t)|. \tag{9}$$

As L_λ is stochastically dominated by $\widetilde{L}_\lambda$, the tightness property of the family of $\{L_\lambda\}$ follows from the tightness property of the family $\{\sup_{s<t} L^i_\lambda(t)\}$ for all $i \in \mathcal{S}$. $\square$

The next theorem follows from Corollary 2.6 and Theorem 2.7.

Theorem 2.8. *The processes $\{L_\lambda(t), \varphi_\lambda(t) : t \geq 0\}$ converge weakly to the Markov-modulated Brownian motion $\{Y(t), \kappa(t) : t \geq 0\}$.* $\square$

3. Stationary distribution

We consider again the Markov-modulated Brownian motion $\{Y(t), \kappa(t) : t \geq 0\}$ described in Section 2, but with a reflection at level zero. The reflected process is denoted as $\{\widehat{Y}(t), \kappa(t) : t \geq 0\}$, with

$$\widehat{Y}(t) = Y(t) - \inf_{0 \leq v \leq t} Y(v).$$

Furthermore, we define the reflected fluid process $\{\widehat{L}_\lambda(t), \beta_\lambda(t), \varphi_\lambda(t) : t \geq 0\}$, where

$$\widehat{L}_\lambda(t) = L_\lambda(t) - \inf_{0 \leq v \leq t} L_\lambda(v).$$

For notational convenience, we define $\varepsilon = 1/\sqrt{\lambda}$. With this, the reflected fluid processes is written as $\{\widehat{L}_\varepsilon(t), \beta_\varepsilon(t), \varphi_\varepsilon(t) : t \geq 0\}$ and our purpose is to show that the stationary distribution of $\{\widehat{Y}(t), \kappa(t)\}$ is the limit, as $\varepsilon \to 0$, of the stationary distribution of the reflected fluid process $\{\widehat{L}_\varepsilon(t), \varphi_\varepsilon(t)\}$. We emphasize that the processes $\{\widehat{L}_\lambda(t), \varphi_\lambda(t)\}$ and $\{\widehat{L}_\varepsilon(t), \varphi_\varepsilon(t)\}$ are the same. The change in subscripts only reflects the notational change in our perturbation parameter.

Assumption 3.1. *The mean drift $\boldsymbol{\alpha} D \mathbf{1}$ is negative, so that all reflected processes are positive recurrent.*

Assumption 3.2. *The variance σ_i^2 is positive for all $i \in \mathcal{M}$. This assumption ensures the existence of Θ^{-1}, which we need later on.*

The following result is a direct corollary of Theorem 2.8, by the Skorokhod mapping theorem.

Corollary 3.3. *The processes $\{\widehat{L}_\varepsilon(t), \varphi_\varepsilon(t) : t \geq 0\}$ weakly converge as $\varepsilon \to 0$ to the reflected Markov-modulated Brownian motion $\{\widehat{Y}(t), \kappa(t) : t \geq 0\}$.* $\square$

We denote the stationary distribution vector of $\{\widehat{L}_\varepsilon(t), \beta_\varepsilon(t), \varphi_\varepsilon(t)\}$ by $\boldsymbol{F}_\varepsilon(x)$ and the associated stationary density vector by $\boldsymbol{\pi}_\varepsilon(x)$, with components

$$[F_\varepsilon(x)]_{ki} = \lim_{t \to \infty} \mathrm{P}[\widehat{L}_\varepsilon(t) \leq x, \beta_\varepsilon(t) = k, \varphi_\varepsilon(t) = i],$$

and $[\pi_\varepsilon(x)]_{ki} = \mathrm{d}/\mathrm{d}x[F_\varepsilon(x)]_{ki}$, for $k \in \{1,2\}$ and $i \in \mathcal{M}$, and we partition the generator and the fluid rate matrices as

$$T_\varepsilon = \begin{bmatrix} T_\varepsilon^{++} & T_\varepsilon^{+-} \\ T_\varepsilon^{-+} & T_\varepsilon^{--} \end{bmatrix} \qquad \text{and} \qquad C_\varepsilon = \begin{bmatrix} C_\varepsilon^{+} & 0 \\ 0 & C_\varepsilon^{-} \end{bmatrix},$$

where

$$\begin{aligned} T_\varepsilon^{++} &= T_\varepsilon^{--} = Q - (1/\varepsilon)^2 I, & C_\varepsilon^{+} &= D + (1/\varepsilon)\Theta, \\ T_\varepsilon^{+-} &= T_\varepsilon^{-+} = (1/\varepsilon)^2 I, & C_\varepsilon^{-} &= D - (1/\varepsilon)\Theta. \end{aligned}$$

We assume that ε is sufficiently small that the diagonal elements of C_ε^+ are all positive, and those of C_ε^- are all negative, and we write $|C_\varepsilon^-|$ for the matrix of absolute values of the elements of C_ε^-.

Let $\xi_\varepsilon^+(x) = \inf\{t < \infty : L_\varepsilon(t) > x\}$ and $\xi_\varepsilon^-(x) = \inf\{t < \infty : L_\varepsilon(t) < x\}$ be the first passage times to the level x, respectively from below and from above, of the unbounded process L_ε. A key component of the stationary distribution of $\{\widehat{L}_\varepsilon(t), \beta_\varepsilon(t), \varphi_\varepsilon(t)\}$ is the matrix Ψ_ε of first passage probability from above, that is,

$$\begin{aligned} (\Psi_\varepsilon)_{ij} = \mathrm{P}[\xi_\varepsilon^-(x) < \infty, \beta_\varepsilon(\xi_\varepsilon^-(x)) = 2, \varphi_\varepsilon(\xi_\varepsilon^+(x)) = j \\ |L_\varepsilon(0) = x, \beta_\varepsilon(0) = 1, \varphi_\varepsilon(0) = i] \end{aligned}$$

for i and j in $\mathcal{M}$, and any level x. Similarly, Ψ_ε^* is the matrix of first passage probabilities from below, from $(x,2,i)$ to $(x,1,j)$, for i and j in $\mathcal{M}$.

The stationary distribution is given in the literature under various slightly different forms; here, we use the one from Govorun *et al.* [14, Theorem 2.1]:

$$\boldsymbol{F}_\varepsilon(0) = \begin{bmatrix} \mathbf{0} & \boldsymbol{\zeta}_\varepsilon \end{bmatrix}, \tag{10}$$

$$\boldsymbol{\pi}_\varepsilon(x) = \boldsymbol{\zeta}_\varepsilon T_\varepsilon^{-+} e^{K_\varepsilon x} \begin{bmatrix} (C_\varepsilon^+)^{-1} & \Psi_\varepsilon |C_\varepsilon^-|^{-1} \end{bmatrix} \quad \text{for } x > 0, \tag{11}$$

where

$$K_\varepsilon = (C_\varepsilon^+)^{-1} T_\varepsilon^{++} + \Psi_\varepsilon |C_\varepsilon^-|^{-1} T_\varepsilon^{-+}, \tag{12}$$

and $\boldsymbol{\zeta}_\varepsilon$ is the unique solution of

$$\boldsymbol{\zeta}_\varepsilon (T_\varepsilon^{--} + T_\varepsilon^{-+} \Psi_\varepsilon) = \mathbf{0}, \tag{13}$$

$$\boldsymbol{\zeta}_\varepsilon \mathbf{1} + \boldsymbol{\zeta}_\varepsilon T_\varepsilon^{-+} (-K_\varepsilon)^{-1} \{(C_\varepsilon^+)^{-1} \mathbf{1} + \Psi_\varepsilon |C_\varepsilon^-|^{-1} \mathbf{1}\} = 1. \tag{14}$$

Probabilistically, $\boldsymbol{\zeta}_\varepsilon$ is up to a multiplicative constant the stationary distribution of the process censored at level 0, and $e^{K_\varepsilon x}$ is the matrix of expected number of crossings of level x in the various phases $(1,i)$, starting from level 0, before the first return to level 0.

In view of (11), we need to analyze Ψ_ε, K_ε and ζ_ε as $\varepsilon \to 0$, and it is obvious from (12) and (13, 14) that we should focus on the matrix Ψ_ε first. The next lemma is the key to our analysis. One might expect (15, 16) to have a simple proof but the one we have is lengthy and tedious. We place it in Appendix to preserve the flow of the paper. Lemma 3.6 and Theorem 3.7 easily follow.

Lemma 3.4. *For $\varepsilon \geq 0$,*

$$\Psi_\varepsilon = I + \varepsilon\Psi_1 + O(\varepsilon^2), \tag{15}$$
$$\Psi^*_\varepsilon = I + \varepsilon\Psi^*_1 + O(\varepsilon^2), \tag{16}$$

where $\Theta^{-1}\Psi_1$ and $-\Theta^{-1}\Psi^*_1$ are both solutions to

$$X^2 + 2V^{-1}DX + 2V^{-1}Q = 0, \tag{17}$$

and are irreducible. Furthermore, the roots $\theta_1, \theta_2, \ldots, \theta_{2m}$ of the polynomial

$$\gamma(z) = \det(z^2 I + 2zV^{-1}D + 2V^{-1}Q)$$

associated to (17), numbered in increasing order of their real parts, satisfy the inequalities

$$Re(\theta_1) \leq \cdots \leq Re(\theta_{m-1}) < \theta_m = 0 < Re(\theta_{m+1}) \leq \cdots \leq Re(\theta_{2m}). \tag{18}$$

Finally, $\Theta^{-1}\Psi_1$ has one eigenvalue equal to zero and $m-1$ eigenvalues with strictly negative real parts, and it is the unique such solution; $-\Theta^{-1}\Psi^*_1$ has m eigenvalues with strictly positive real parts, and is the unique such solution. □

Remark 3.5. Let $\tau^{\pm}_x = \inf\{t < \infty : \pm Y(t) > x\}$ be the first passage times to the corresponding levels x and $-x$ of the unbounded process $Y(t)$. Under Assumption 3.2 that $\sigma_i > 0$ for all $i \in \mathcal{M}$, it is easy to confirm that $\Theta^{-1}\Psi_1$ and $\Theta^{-1}\Psi^*_1$ are the same as, respectively, the generators Λ^- and Λ^+ of the time-changed processes $\kappa(\tau^-_x)$ and $\kappa(\tau^+_x)$ in Ivanovs [16], and $\Theta^{-1}\Psi^*_1$ is the same as the matrix $U(\gamma)$ for $\gamma = 0$ in Breuer [10].

By Lemma 3.4, the matrices Ψ_1 and Ψ^*_1 are uniquely identified through (17). We now turn to the matrix K_ε, and, to complete the picture, to the matrix K^*_ε defined as

$$K^*_\varepsilon = \Psi^*_\varepsilon (C^+_\lambda)^{-1} T^{+-}_\lambda + |C^{-1}_\lambda|^{-1} T^{--}_\lambda. \tag{19}$$

Lemma 3.6. *For $\varepsilon \geq 0$,*

$$K_\varepsilon = K_0 + O(\varepsilon), \tag{20}$$
$$K^*_\varepsilon = K^*_0 + O(\varepsilon), \tag{21}$$

where $K_0 = \Psi_1\Theta^{-1} + 2V^{-1}D$ and $K^*_0 = \Psi^*_1\Theta^{-1} - 2V^{-1}D$. The matrices $-K_0$ and K^*_0 are solutions of the equation

$$X^2 + 2XV^{-1}D + 2\Theta^{-1}Q\Theta^{-1} = 0, \tag{22}$$

and are irreducible. Furthermore, K_0 has m eigenvalues with strictly negative real parts, and it is the unique such solution; K^*_0 has one eigenvalue equal to zero and $m-1$ eigenvalues with strictly negative real parts, and is the unique such solution.

Proof. We write (12) as

$$\begin{aligned}\varepsilon K_\varepsilon &= -(\Theta + \varepsilon D)^{-1}(I - \varepsilon^2 Q) + \Psi_\varepsilon(\Theta - \varepsilon D)^{-1} \\ &= -(\Theta^{-1} - \varepsilon V^{-1} D + O(\varepsilon^2))(I - \varepsilon^2 Q) \\ &\quad + (I + \varepsilon \Psi_1 + O(\varepsilon^2))(\Theta^{-1} + \varepsilon V^{-1} D + O(\varepsilon^2))\end{aligned}$$

by (40), (41) and (15),

$$= \varepsilon(2V^{-1}D + \Psi_1 \Theta^{-1}) + O(\varepsilon^2),$$

which proves (20); equation (21) follows in a similar manner. It is easy to verify that K_0 and $-K_0^*$ both satisfy (22), of which the associated polynomial is

$$\begin{aligned}\Xi(z) &= z^2 I + 2zV^{-1}D + 2\Theta^{-1}Q\Theta^{-1} \\ &= \Theta^{-1}\Gamma(z)\Theta^{-1} \qquad \text{with } \Gamma(z) \text{ defined in Proposition A.4} \\ &= (zI + K_0)\Theta(zI - \Theta^{-1}\Psi_1)\Theta^{-1}\end{aligned}$$

by (52), after some simple manipulations. This, together with Lemma 3.4, shows that the eigenvalues of $-K_0$ are the roots $\theta_{m+1}, \ldots, \theta_{2m}$ of $\gamma(z)$ with strictly positive real parts. Finally, using (53) we write

$$\Xi(z) = (zI - K_0^*)\Theta(zI + \Theta^{-1}\Psi_1^*)\Theta^{-1},$$

and conclude that the eigenvalues of K_0^* are the roots θ_1 to θ_m. Finally, as Ψ_1 and Ψ_1^* are irreducible, and Θ, V, and D are diagonal matrices, we conclude that K_0 and K_0^* are irreducible, and this completes the proof. □

The next theorem states that the limit, as $\lambda \to \infty$, of the stationary distributions of the approximating fluid processes is indeed the stationary distribution of the limiting process $\{Y(t), \kappa(t)\}$. We prove this result by showing that the limiting distribution (23) coincides with the stationary distribution of $\{Y(t), \kappa(t)\}$ as obtained by Asmussen [4, Theorem 2.1 and Corollary 4.1].

Theorem 3.7. *The limiting distribution of $\{\widehat{L}_\lambda(t), \widehat{\varphi}_\lambda(t)\}$ converges, as λ goes to infinity, to the stationary distribution of $\{Y(t), \kappa(t)\}$, and is given by*

$$\lim_{\varepsilon \to 0} \boldsymbol{\pi}_\varepsilon(x)(\mathbf{1} \otimes I) = 2\boldsymbol{\zeta}_1 e^{K_0 x}\Theta^{-1}, \tag{23}$$

$$\lim_{\varepsilon \to 0} \boldsymbol{F}_\varepsilon(0)(\mathbf{1} \otimes I) = \mathbf{0}, \tag{24}$$

where

$$\boldsymbol{\zeta}_1 \Psi_1 = \mathbf{0}, \tag{25}$$

$$2\boldsymbol{\zeta}_1(-K_0)^{-1}\Theta^{-1}\mathbf{1} = 1. \tag{26}$$

Proof. The solution of (13) is of the form $\boldsymbol{\zeta}_\varepsilon = \boldsymbol{\zeta}_0 + \varepsilon\boldsymbol{\zeta}_1 + o(\varepsilon)$ ([18, Theorem 5.4]) and (14) becomes

$$\begin{aligned}&\{\boldsymbol{\zeta}_0 + \varepsilon\boldsymbol{\zeta}_1 + o(\varepsilon)\}\mathbf{1}\\ &+ \{\boldsymbol{\zeta}_0 + \varepsilon\boldsymbol{\zeta}_1 + o(\varepsilon)\}(1/\varepsilon)(-K_\varepsilon)^{-1}\{(\varepsilon D + \Theta)^{-1} + \Psi_\varepsilon(\Theta - \varepsilon D)^{-1}\}\mathbf{1} = 1. \quad (27)\end{aligned}$$

We equate the coefficients of $1/\varepsilon$ on both sides of (27), using (15, 20) and we find that

$$-2\boldsymbol{\zeta}_0 K_0^{-1}\Theta^{-1}\mathbf{1} = 0.$$

Equation (10) implies that $\boldsymbol{\zeta}_\varepsilon \geq \mathbf{0}$ and, by continuity, $\boldsymbol{\zeta}_0 \geq \mathbf{0}$. Furthermore, Θ^{-1} is a diagonal matrix with strictly positive diagonal. Finally, K_0 is irreducible and has eigenvalues with strictly negative real part by Lemma 3.6, so that $\int_0^\infty e^{K_0 u}\,\mathrm{d}u$ converges to $-K_0^{-1}$ and is strictly positive. This implies that $\boldsymbol{\zeta}_0 = \mathbf{0}$, which proves (24). Using $\boldsymbol{\zeta}_0 = \mathbf{0}$, and equating coefficients of ε on both sides of (13) gives (25), while equating the coefficients of ε^0 on both sides of (27) leads to (26).

Gathering everything together, we obtain from (11)

$$\boldsymbol{\pi}_\varepsilon(x) = (1/\varepsilon)\{\varepsilon\boldsymbol{\zeta}_1 + o(\varepsilon)\}e^{K_\varepsilon x}\left[(\varepsilon D + \Theta)^{-1} \quad \Psi_\varepsilon(\Theta - \varepsilon D)^{-1}\right], \qquad (28)$$

from which (23) follows. Equation (24) is a direct consequence of (10).

To verify that the limiting distribution in Theorem 3.7 is the stationary density vector $\boldsymbol{g}(x)$ of the Markov-modulated Brownian motion, we use Asmussen [4]. By [4, Theorem 2.1 and Corollary 4.1],

$$\boldsymbol{g}(x) = [e^{\Lambda x}(-\Lambda\mathbf{1})]^{\mathrm{T}}\Delta_{\boldsymbol{\alpha}}, \qquad (29)$$

where $\boldsymbol{\alpha}$ is the stationary distribution vector of Q, $\Delta_{\boldsymbol{\alpha}} = \mathrm{diag}(\boldsymbol{\alpha})$, and Λ is a defective generator matrix satisfying

$$(1/2)V\Lambda^2 - D\Lambda + \Delta_{1/\boldsymbol{\alpha}}Q^{\mathrm{T}}\Delta_{\boldsymbol{\alpha}} = 0. \qquad (30)$$

Define $Z = \Delta_{1/\boldsymbol{\alpha}}\Lambda^{\mathrm{T}}\Delta_{\boldsymbol{\alpha}}$ and rewrite (29) as

$$\boldsymbol{g}(x) = -\mathbf{1}^{\mathrm{T}}\Delta_{\boldsymbol{\alpha}}Z\Delta_{1/\boldsymbol{\alpha}}e^{\Delta_{\boldsymbol{\alpha}}Z\Delta_{1/\boldsymbol{\alpha}}x}\Delta_{\boldsymbol{\alpha}} = -\boldsymbol{\alpha}Ze^{Zx}. \qquad (31)$$

By (30), we find that

$$(1/2)Z^2V - ZD + Q = 0$$

and that $\Theta^{-1}Z\Theta$ is a solution of (22). It is similar to Z and so to Λ^{T}, therefore, the eigenvalues of $\Theta^{-1}Z\Theta$ all have strictly negative real parts and we have

$$K_0 = \Theta^{-1}Z\Theta. \qquad (32)$$

Substituting (32) into (23) gives $\lim_{\varepsilon\to 0}\boldsymbol{\pi}_\varepsilon(\mathbf{1}\otimes I) = 2\boldsymbol{\zeta}_1\Theta^{-1}e^{Zx}$. Finally, it is straightforward to verify that $\boldsymbol{\zeta}_1 = -\boldsymbol{\alpha}(\Theta^{-1}D + (1/2)\Theta\Psi_1\Theta^{-1})$, and consequently

$$\lim_{\varepsilon\to 0}\boldsymbol{\pi}_\varepsilon(\mathbf{1}\otimes I) = \boldsymbol{g}(x). \qquad (33)$$

□

Remark 3.8. An alternative way to show that the stationary distribution of the approximating fluid process $\{\widehat{L}_\lambda(t), \widehat{\varphi}_\lambda(t)\}$ converges, as $\lambda \to \infty$, to the stationary distribution of the limiting Markov-modulated Brownian motion $\{Y(t), \kappa(t)\}$ is via the maximum representation of the relevant processes. Asmussen [4] derives the stationary distribution, both for fluid queues and for the MMBM in this manner, linking these to the distribution of the maximum of the time-reversed process. Following the arguments in Enikeeva *et al.* [13] and in Stenflo [22], one might show that there is continuity of the maximum distributions of the backward processes, as $\lambda \to \infty$, and consequently obtain the continuity of stationary distributions.

This would lead to a time reversal-based proof of convergence. As stated in the introduction, we aim at following the forward-time approach and so obtain a different representation of the stationary distribution. In addition, we obtain limiting properties for key matrices, and these results will be proved useful in future work.

4. Computational procedure

Theorem 3.7 indicates that the matrix Ψ_1 is the central ingredient in evaluating the stationary distribution of the Markov-modulated Brownian motion $\{Y(t), \varphi(t)\}$. We describe here how to use the splitting property (18) in numerically solving for Ψ_1 and Ψ_1^*.

Bini and Gemignani [6] consider quadratic matrix equations $C + AX + BX^2 = 0$ where the roots of the associated polynomial $\det(C + zA + z^2B)$ are split by a circle in $\mathbb{C}$, half being inside the disk and half outside. The problem in [6] is to find the minimal solution, that is, the solution matrix with all eigenvalues inside the disk.

In our case, the roots are split between the negative and the positive half-planes and we need to apply some transformation, such as the one described in Bini *et al.* [8] and based on the inverse Möbius mapping [3, Chapter 2.1]

$$w(z) = \frac{z-1}{z+1}. \tag{34}$$

This inverse mapping applies the open unit disk $|z| < 1$ onto the negative half-plane $\mathbb{C}_-$, the unit circle $|z| = 1$ minus the point $z = -1$ onto the imaginary axis $\mathbb{C}_0$, the outside $|z| > 1$ of the closed unit disk onto the positive half-plane $\mathbb{C}_+$, and the imaginary axis $\mathbb{C}_0$ onto the unit circle $|w| = 1$ minus the point $w = 1$.

Now, define $W(Z) = (Z - I)(Z + I)^{-1}$. Instead of solving $P(X) = X^2 + 2V^{-1}DX + 2V^{-1}Q = 0$ for $\Theta^{-1}\Psi_1$, we solve $H(Z) = 0$, where

$$\begin{aligned}
H(Z) &= P(W(Z))(I + Z)^2 \\
&= P((Z - I)(Z + I)^{-1})(I + Z)^2 \\
&= (I + 2V^{-1}D + 2V^{-1}Q)Z^2 - 2(I - 2V^{-1}Q)Z + I \\
&\quad - 2V^{-1}D + 2V^{-1}Q.
\end{aligned}$$

The roots of $\det(H(z))$ are given by $\omega_i = w^{-1}(\tau_i) = (1+\tau_i)/(1-\tau_i)$ for $i = 1, \ldots, 2m$, and satisfy the splitting property

$$0 \leq |\omega_1|, \ldots, |\omega_{m-1}| < \omega_m = 1 < |\omega_{m+1}|, \ldots, |\omega_{2m}|. \tag{35}$$

We note that Bini and Geminiani [6] requires $|\omega_i| > 0$ for all i, but as seen in [9, Section 8.3], the weak inequality suffices.

Define Z_0 as the solution of $H(Z) = 0$ such that $\mathrm{sp}(Z_0) \leq 1$. The matrix $W(Z_0) = (Z_0 - I)(Z_0 + I)^{-1}$ is a solution of $P(X) = 0$ with all eigenvalues in $\{\mathrm{Re}(z) \leq 0\}$, and so $\Theta^{-1}\Psi_1 = (Z_0 - I)(Z_0 + I)^{-1}$.

Several iterative algorithms to compute Z_0 are discussed in [6]. Some have superlinear convergence, such as Cyclic reduction [7], Logarithmic reduction [19, Chapter 8], subspace iteration [2] or Graeffe iteration [7]. This means that the approximation error at the ith iteration is $O(\sigma^{2^i})$ with $\sigma = 1/|\omega_{m+1}| < 1$. These algorithms are globally convergent but they are not self-correcting. Since the coefficient matrices in $P(X)$ are of mixed signs, one might prefer the algorithm developed in [6]: it is self-correcting and the approximation error is $O(\sigma^{i2^k})$ for arbitrary k, which makes it arbitrarily fast.

As for Ψ_1^*, if we define Z_1 as the root of $H(Z)$ such that all of its eigenvalues are outside the closed unit disk, then $W(Z_1)$ has all its eigenvalues in the half-plane $\mathbb{C}_+$, and $\Theta^{-1}\Psi_1^* = -W(Z_1)$. The algorithms in [6], however, do not seem to be well adapted to the computation of Z_1, and we suggest to use a different transformation, in order to bring the eigenvalues of $-\Theta^{-1}\Psi_1^*$ inside the unit disk. This transformation is based on Möbius' mapping [3, Chapter 2.1]

$$z(w) = \frac{1+w}{1-w}. \tag{36}$$

This mapping applies the open unit disk $|w| < 1$ onto the positive half-plane $\mathbb{C}_+$, the unit circle $|w| = 1$ minus the point $w = 1$ onto the imaginary axis $\mathbb{C}_0 = \{z : \mathrm{Re}(z) = 0\}$ and the outside $|w| > 1$ of the closed unit disk onto the negative half-plane $\mathbb{C}_-$, finally, it applies the imaginary axis $\mathbb{C}_0$ onto the unit circle $|z| = 1$ minus the point $z = -1$.

Now, define $Z(W) = (I + W)(I - W)^{-1}$. Instead of solving $H(Z) = 0$ we solve $Q(W) = 0$ for the matrix solution W_1 with eigenvalues inside the unit disk, where

$$\begin{aligned} Q(W) &= P(Z(W))(I - W)^2, \\ &= (I - 2V^{-1}D + 2V^{-1}Q)W^2 + 2(I - 2V^{-1}Q)W + I + 2V^{-1}D + 2V^{-1}Q, \end{aligned}$$

and so $\Theta^{-1}\Psi_1^* = -(I + W_1)(I - W_1)^{-1}$.

Appendix A: Proof of Lemma 3.4

The proof goes in four main steps. The matrix Ψ_ε is the stochastic (or sub-stochastic) solution of the Riccati equation

$$(C_\varepsilon^+)^{-1}T_\varepsilon^{+-} + (C_\varepsilon^+)^{-1}T_\varepsilon^{++}\Psi_\varepsilon + \Psi_\varepsilon|C_\varepsilon^-|^{-1}T_\varepsilon^{--} + \Psi_\varepsilon|C_\varepsilon^-|^{-1}T_\varepsilon^{-+}\Psi_\varepsilon = 0$$

(Rogers [25]), equation that we write as

$$(1/\varepsilon)(\varepsilon D+\Theta)^{-1}+(1/\varepsilon)(\varepsilon D+\Theta)^{-1}(\varepsilon^2 Q-I)\Psi_\varepsilon \\ +(1/\varepsilon)\Psi_\varepsilon|\varepsilon D-\Theta|^{-1}(\varepsilon^2 Q-I)+(1/\varepsilon)\Psi_\varepsilon|\varepsilon D-\Theta|^{-1}\Psi_\varepsilon=0 \tag{37}$$

Thus, Ψ_ε is a solution of $\mathcal{F}_\varepsilon(X)=0$, where

$$\mathcal{F}_\varepsilon(X)=(\varepsilon D+\Theta)^{-1}+(\varepsilon D+\Theta)^{-1}(\varepsilon^2 Q-I)X \\ +X|\varepsilon D-\Theta|^{-1}(\varepsilon^2 Q-I)+X|\varepsilon D-\Theta|^{-1}X.$$

For $\varepsilon=0$, we see that $\mathcal{F}_0(I)=0$. It is tempting to invoke the Implicit Function Theorem and claim that Ψ_ε is an analytic function of ε in a neighborhood of $\varepsilon=0$. Unfortunately, the operator $\partial/\partial X\mathcal{F}_\varepsilon(X)$ is singular at the point $(\varepsilon=0, X=I)$, the Implicit Function Theorem does not apply, and we follow a longer, more tortuous path.

Proposition A.1. *For $\varepsilon\geq 0$,*

$$\Psi_\varepsilon=I+\Phi_\varepsilon \quad \textit{where} \ \lim_{\varepsilon\to 0}\Phi_\varepsilon=0, \tag{38}$$

$$\Psi^*_\varepsilon=I+\Phi^*_\varepsilon \quad \textit{where} \ \lim_{\varepsilon\to 0}\Phi^*_\varepsilon=0. \tag{39}$$

Proof. We note that

$$\begin{aligned}(\varepsilon D+\Theta)^{-1}&=\{I-\varepsilon\Theta^{-1}(-D)\}^{-1}\Theta^{-1}\\ &=\{I-\varepsilon\Theta^{-1}D+\varepsilon^2(\Theta^{-1}D)^2+O(\varepsilon^3)\}\Theta^{-1}\\ &=\Theta^{-1}-\varepsilon V^{-1}D+\varepsilon^2\Theta^{-1}V^{-1}D^2+O(\varepsilon^3),\end{aligned} \tag{40}$$

and similarly

$$|\varepsilon D-\Theta|^{-1}=\Theta^{-1}+\varepsilon V^{-1}D+\varepsilon^2\Theta^{-1}V^{-1}D^2+O(\varepsilon^3), \tag{41}$$

for ε sufficiently small. Thus, (37) implies that

$$(1/\varepsilon)\{\Theta^{-1}-\varepsilon V^{-1}D+O(\varepsilon^2)\}(I+\varepsilon^2 Q\Psi_\varepsilon-\Psi_\varepsilon) \\ +(1/\varepsilon)\Psi_\varepsilon\{\Theta^{-1}+\varepsilon V^{-1}D+O(\varepsilon^2)\}(\varepsilon^2 Q-I+\Psi_\varepsilon)=0, \tag{42}$$

which can be reorganized as $\mathcal{G}(\varepsilon)+\mathcal{H}(\varepsilon)=0$, where

$$\mathcal{G}(\varepsilon)=(1/\varepsilon)(\Theta^{-1}-\Theta^{-1}\Psi_\varepsilon-\Psi_\varepsilon\Theta^{-1}+\Psi_\varepsilon\Theta^{-1}\Psi_\varepsilon) \tag{43}$$

and

$$\mathcal{H}(\varepsilon)=(1/\varepsilon)\{\varepsilon^2\Theta^{-1}Q\Psi_\varepsilon+(-\varepsilon V^{-1}D+O(\varepsilon^2))(I+\varepsilon^2 Q\Psi_\varepsilon-\Psi_\varepsilon) \\ +\varepsilon^2\Psi_\varepsilon\Theta^{-1}Q+\Psi_\varepsilon(\varepsilon V^{-1}D+O(\varepsilon^2))(\varepsilon^2 Q-I+\Psi_\varepsilon)\}. \tag{44}$$

The matrix $\mathcal{H}(\varepsilon)$ is bounded and therefore $\mathcal{G}(\varepsilon)$ too remains bounded as $\varepsilon\to 0$.

Now, we observe that Ψ_ε belongs to the compact set $\{M : M \geq 0, M\mathbf{1} \leq \mathbf{1}\}$ of (sub)stochastic matrices; therefore, for every sequence $\{\Psi_\varepsilon\}_{\varepsilon\to 0}$ there exist subsequences that converge. Let $\bar{\Psi}$ be the limit of one such convergent subsequence, and $\{\varepsilon_i\}_{i=1,2,\ldots}$ be a sequence such that $\varepsilon_i \to 0$ and $\Psi_{\varepsilon_i} \to \bar{\Psi}$ as $i \to \infty$. Since $\mathcal{G}(\varepsilon_i)$ remains bounded as $i \to \infty$, necessarily

$$\begin{aligned}\lim_{i\to\infty}(\Theta^{-1} - \Theta^{-1}\Psi_{\varepsilon_i} - \Psi_{\varepsilon_i}\Theta^{-1} + \Psi_{\varepsilon_i}\Theta^{-1}\Psi_{\varepsilon_i}) &= \lim_{i\to\infty}(I - \Psi_{\varepsilon_i})\Theta^{-1}(I - \Psi_{\varepsilon_i}) \\ &= (I - \bar{\Psi})\Theta^{-1}(I - \bar{\Psi}) \\ &= 0,\end{aligned}$$

and thus $\bar{\Psi} = I$. This follows from the facts that $\Theta^{-1}(I-\bar{\Psi})$ is a nilpotent matrix, that the trace of every nilpotent matrix is zero, and that $\bar{\Psi}$ is a (sub)stochastic matrix while Θ^{-1} is a strictly positive diagonal matrix.

All convergent subsequences having the same limit, the conclusion is that Ψ_ε converges to I as $\varepsilon \to 0$, and (38) follows. The proof of (39) is by analogous arguments. ☐

Proposition A.2. *For $\varepsilon > 0$, the matrices Φ_ε and Φ^*_ε are irreducible with nonnegative off-diagonal elements, strictly negative diagonal elements, $\Phi_\varepsilon\mathbf{1} \leq \mathbf{0}$ and $\Phi^*_\varepsilon\mathbf{1} \leq \mathbf{0}$. In addition, under Assumption 3.1, $\Phi_\varepsilon\mathbf{1} = \mathbf{0}$ and $\Phi^*_\varepsilon\mathbf{1} < \mathbf{0}$.*

In short, Φ_ε is an irreducible generator and Φ^*_ε is an irreducible subgenerator.

Proof. As we assume that the fluid queue is irreducible, Ψ_ε is an irreducible (sub)stochastic matrix for all $\varepsilon \geq 0$. Thus, we conclude from (38) that Φ_ε is irreducible and that its off-diagonal elements are nonnegative. Furthermore, since $\Psi_\varepsilon\mathbf{1} \leq \mathbf{1}$, this implies that $\Phi_\varepsilon\mathbf{1} \leq \mathbf{0}$, so that its diagonal elements are strictly negative and Φ_ε is a generator. The same argument holds for Φ^*_ε.

Under Assumption 3.1, the matrix Ψ_ε is stochastic and the matrix Ψ^*_ε is strictly substochastic, and the last claim follows. ☐

Proposition A.3. *The matrices $(1/\varepsilon)\Phi_\varepsilon$ and $(1/\varepsilon)\Phi^*_\varepsilon$ are bounded. Denoting by $\bar{\Psi}_1$ and $\bar{\Psi}^*_1$ the limits of any converging subsequences of $(1/\varepsilon)\Phi_\varepsilon$ and $(1/\varepsilon)\Phi^*_\varepsilon$ respectively, both $\bar{\Psi}_1$ and $-\bar{\Psi}^*_1$ are solutions of the equation*

$$(\Theta^{-1}Y)^2 + 2V^{-1}D\Theta^{-1}Y + 2V^{-1}Q = 0, \tag{45}$$

and are irreducible.

Proof. Substituting (38) into (43, 44) gives us

$$\begin{aligned}\mathcal{G}(\varepsilon) &= (1/\varepsilon)\{\Phi_\varepsilon\Theta^{-1}\Phi_\varepsilon\} \\ \mathcal{H}(\varepsilon) &= \varepsilon\Theta^{-1}Q(I + \Phi_\varepsilon) + [-V^{-1}D + O(\varepsilon)](\varepsilon^2 Q + \varepsilon^2 Q\Phi_\varepsilon - \Phi_\varepsilon) \\ &\quad + \varepsilon(I + \Phi_\varepsilon)\Theta^{-1}Q + (I + \Phi_\varepsilon)[V^{-1}D + O(\varepsilon)](\varepsilon^2 Q + \Phi_\varepsilon).\end{aligned}$$

Clearly, $\lim_{\varepsilon\to 0}\mathcal{H}(\varepsilon) = 0$ and this implies that $\lim_{\varepsilon\to 0}(1/\varepsilon)\{\Phi_\varepsilon\Theta^{-1}\Phi_\varepsilon\} = 0$ since $\mathcal{G}(\varepsilon) + \mathcal{H}(\varepsilon) = 0$. Divide both sides of that equation by ε and obtain

$$(1/\varepsilon)^2\Phi_\varepsilon\Theta^{-1}\Phi_\varepsilon + 2\Theta^{-1}Q + \mathcal{R}(\varepsilon) = 0 \tag{46}$$

where

$$\mathcal{R}(\varepsilon) = \Theta^{-1}Q\Phi_\varepsilon + [-V^{-1}D + O(\varepsilon)](\varepsilon Q + \varepsilon Q\Phi_\varepsilon - (1/\varepsilon)\Phi_\varepsilon) + \Phi_\varepsilon\Theta^{-1}Q + (I + \Phi_\varepsilon)[V^{-1}D + O(\varepsilon)]\{\varepsilon Q + (1/\varepsilon)\Phi_\varepsilon\}. \quad (47)$$

Now, there are three possible cases.

Case 1: $(1/\varepsilon)\Phi_\varepsilon \to 0$ as $\varepsilon \to 0$. Then, $\mathcal{R}(\varepsilon) \to 0$ and taking the limit of both sides of (46) as $\varepsilon \to 0$ leads to $2V^{-1}Q = 0$, which is not true.

Case 2: $(1/\varepsilon)\Phi_\varepsilon$ is unbounded in any neighborhood of $\varepsilon = 0$. Then, there exists a sequence $\{\varepsilon_k\}_{k=1,2...}$ such that $\varepsilon_k \to 0$ and $\max_{ij} |\Phi_{\varepsilon_k}|_{ij}/\varepsilon_k \to \infty$. In this case, we may write $(1/\varepsilon)\Phi_\varepsilon = u_\varepsilon B_\varepsilon$, where u_ε is a scalar function such that $\lim_{k\to\infty} u_{\varepsilon_k} = \infty$ while B_{ε_k} remains bounded and does not converge to zero: since Φ_ε is an irreducible generator, its maximum element is on the diagonal and we take $u_\varepsilon = \max_j |\Phi_\varepsilon|_{jj}/\varepsilon$, then B_ε is an irreducible generator with at least one diagonal element equal to -1, and with $|B_{ij}| \leq 1$ for all i and j.

Next, for ε in the sequence $\{\varepsilon_k\}$, we replace $(1/\varepsilon)\Phi_\varepsilon$ in (46) by $u_\varepsilon B_\varepsilon$ and divide both sides of the equation by u_ε^2 to obtain

$$\begin{aligned} &B_\varepsilon\Theta^{-1}B_\varepsilon + (2/u_\varepsilon^2)\Theta^{-1}Q + (\varepsilon/u_\varepsilon)\Theta^{-1}QB_\varepsilon \\ &\quad + [-V^{-1}D + O(\varepsilon)](\varepsilon/u_\varepsilon^2 Q + \varepsilon^2/u_\varepsilon QB_\varepsilon - (1/u_\varepsilon)B_\varepsilon) \\ &\quad + ((1/u_\varepsilon)I + \varepsilon B_\varepsilon)[V^{-1}D + O(\varepsilon)]((\varepsilon/u_\varepsilon)Q + B_\varepsilon) = 0, \end{aligned} \quad (48)$$

This implies that

$$\lim_{k\to\infty} B_{\varepsilon_k}\Theta^{-1}B_{\varepsilon_k} = 0, \quad (49)$$

Now, take any converging subsequence of B_ε and denote its limit as B. By construction, the trace of $\Theta^{-1}B_\varepsilon$ is at most equal to $\min_j(-\sigma_j^{-1}) < 0$, independently of ε. Thus, the trace of $\Theta^{-1}B$ is strictly negative, the matrix $\Theta^{-1}B$ is not nilpotent, and $\Theta^{-1}B\Theta^{-1}B \neq 0$, which contradicts (49).

Case 3: $(1/\varepsilon)\Phi_\varepsilon$ is bounded and does not converge to 0. Then, from (47)

$$\mathcal{R}(\varepsilon) = (1/\varepsilon)2V^{-1}D\Phi_\varepsilon + \mathcal{R}^*(\varepsilon)$$

where $\mathcal{R}^*(\varepsilon)$ goes to 0 as ε goes to zero. This allows us to rewrite (46) as

$$(1/\varepsilon^2)\Phi_\varepsilon\Theta^{-1}\Phi_\varepsilon + 2\Theta^{-1}Q + (1/\varepsilon)2V^{-1}D\Phi_\varepsilon + \mathcal{R}^*(\varepsilon) = 0$$

and to conclude that

$$\lim_{\varepsilon\to 0}\{(1/\varepsilon^2)\Phi_\varepsilon\Theta^{-1}\Phi_\varepsilon + 2\Theta^{-1}Q + (1/\varepsilon)2V^{-1}D\Phi_\varepsilon + \mathcal{R}^*(\varepsilon)\} = 0. \quad (50)$$

Since $(1/\varepsilon)\Phi_\varepsilon$ is bounded, there exist subsequences $\{\varepsilon_k\}_{k=1,2,...}$ such that $\varepsilon_k \to 0$ and such that $(1/\varepsilon_k)\Phi_{\varepsilon_k} \to \bar{\Psi}_1$. We take in (50) the limit along such a subsequence and conclude that $\bar{\Psi}_1$ is a solution of (45). The same approach is followed for Φ_ε^*.

Now, assume that $\bar{\Psi}_1$ is reducible. We may write

$$\Theta^{-1}\bar{\Psi}_1 = \begin{bmatrix} M_A & 0 \\ M_{AB} & M_B \end{bmatrix},$$

possibly after a permutation of rows and columns, where M_A and M_B are square matrices. As $\Theta^{-1}\bar{\Psi}_1$ is a solution of (45), we are led to conclude that

$$Q = \begin{bmatrix} Q_A & 0 \\ Q_{AB} & Q_B \end{bmatrix}$$

which contradicts our assumption that Q is irreducible. Thus, $\bar{\Psi}_1$ is irreducible, and so is $\bar{\Psi}_1^*$ by the same argument. □

Proposition A.4. *Consider the matrix equation*

$$VX^2 + 2DX + 2Q = 0 \tag{51}$$

and its associated matrix polynomial $\Gamma(z) = Vz^2 + 2Dz + 2Q$.

Under Assumption 3.1, $\det\Gamma(z)$ *has one root equal to zero,* $m-1$ *roots with strictly negative real parts, and* m *roots with strictly positive real parts.*

Proof. Take $\{\varepsilon_k\}_{k=1,2,\ldots}$ to be a subsequence such that $\varepsilon_k \to 0$ and $(1/\varepsilon_k)\Phi_{\varepsilon_k} \to \bar{\Psi}_1$. By Proposition A.3, $\Theta^{-1}\bar{\Psi}_1$ is an irreducible generator and a solution of (51), and we write

$$\begin{aligned} \Gamma(z) &= z^2 V + 2Dz + 2Q - V((\Theta^{-1}\bar{\Psi}_1)^2 + 2V^{-1}D\Theta^{-1}\bar{\Psi}_1 + 2V^{-1}Q) \\ &= (Vz + V\Theta^{-1}\bar{\Psi}_1 + 2D)(zI - \Theta^{-1}\bar{\Psi}_1). \end{aligned} \tag{52}$$

We conclude that all eigenvalues of $\Theta^{-1}\bar{\Psi}_1$ are roots of $\det\Gamma(z)$. As we assume that the fluid queue is positive recurrent, $\Phi_\varepsilon \mathbf{1} = \mathbf{0}$ for all ε by Proposition A.2 and so $\Theta^{-1}\bar{\Psi}_1\mathbf{1} = \mathbf{0}$. Hence, $\det\Gamma(z)$ has at least one root equal to zero, and at least $m-1$ roots with strictly negative real parts.

In a similar manner, we take a subsequence $\{\varepsilon_k^*\}_{k=1,2,\ldots}$ such that $\varepsilon_k^* \to 0$ and $(1/\varepsilon_k^*)\Phi^*_{\varepsilon_k^*} \to \bar{\Psi}_1^*$, and we show that

$$\Gamma(z) = (Vz - V\Theta^{-1}\bar{\Psi}_1^* + 2D)(zI + \Theta^{-1}\bar{\Psi}_1^*). \tag{53}$$

Therefore, all eigenvalues of $-\Theta^{-1}\bar{\Psi}_1^*$ are roots of $\det\Gamma(z)$ and, since $\bar{\Psi}_1^*\mathbf{1} < \mathbf{0}$, this shows that $\det\Gamma(z)$ has at least m roots with strictly positive real part.

The polynomial $\det\Gamma(z)$ has at most $2m$ roots, which concludes the proof. □

Now we are ready to conclude. By the properties of the roots of $\det\Gamma(z)$ given in Proposition A.4, (17) has one unique solution suitable for the role of $\bar{\Psi}_1$ and another unique solution suitable for the role of $\bar{\Psi}_1^*$. Consequently, all convergent subsequences give the same limit Ψ_1 for $(1/\varepsilon_k)\Phi_{\varepsilon_k}$, and Ψ_1^* for $(1/\varepsilon_k)\Phi^*_{\varepsilon_k}$. □

References

[1] S. Ahn and V. Ramaswami. An approach to the stationary and transient analysis on the Markov modulated fluid model. *Private communication*, 2012.

[2] N. Akar and K. Sohraby. An invariant subspace approach in M/G/1 and G/M/1 type Markov chains. *Comm. Statist. Stochastic Models*, 13:381–416, 1997.

[3] T. M. Apostol. *Modular Functions and Dirichlet Series in Number Theory.* New York: Springer–Verlag, 1997.

[4] S. Asmussen. Stationary distributions for fluid flow models with or without Brownian noise. *Communications in Statistics: Stochastic Models*, 11(1):21–49, 1995.

[5] S. Asmussen and O. Kella. A multi-dimensional martingale for Markov additive processes and its applications. *Adv. Appl. Probab.*, 32:376–393, 2000.

[6] D. A. Bini and L. Gemignani. Solving quadratic matrix equations and factoring polynomials: new fixed point iterations based on Schur complements of Toeplitz matrices. *Numer. Linear Algebra Appl.*, 12:181–189, 2005.

[7] D. A. Bini, L. Gemignani, and B. Meini. Computations with infinite Toeplitz matrices and polynomials. *Linear Algebra Appl.*, 343-344:21–61, 2002.

[8] D. A. Bini, G. Latouche, and B. Meini. Solving matrix polynomial equations arising in queueing problems. *Linear Algebra Appl.*, 340:225–244, 2002.

[9] D. A. Bini, G. Latouche, and B. Meini. *Numerical Methods for Structured Markov Chains.* Numerical Mathematics and Scientific Computation. Oxford University Press, Oxford, 2005.

[10] L. Breuer. First passage times for Markov additive processes with positive jumps of Phase-Type. *Journal of Applied Probability*, 45:778–799, 2008.

[11] B. D'Auria, J. Ivanovs, O. Kella, and M. Mandjes. Two-sided reflection of Markov-modulated Brownian motion. *Stoch. Models*, 28(2):316–332, 2012.

[12] B. D'Auria and O. Kella. Markov modulation of a two-sided reflected Brownian motion with application to fluid queues. *Stoch. Proc. Appl.*, 122(4):1566–1581, 2012.

[13] E. Enikeeva, V. Kalashnikov, and D. Rusaityte. Continuity estimates for ruin probabilities. *Scand. Actuarial J.*, 2001(1):18–39, 2001.

[14] M. Govorun, G. Latouche, and M.-A. Remiche. Stability for fluid queues: characteristic inequalities. *Stoch. Models*, 29(1):64–88, 2013.

[15] N. J. Higham. *Functions of Matrices: Theory and Computation.* SIAM, 2008.

[16] J. Ivanovs. Markov-modulated Brownian motion with two reflecting barriers. *Journal of Applied Probability*, 47:1034–1047, 2010.

[17] R. L. Karandikar and V. Kulkarni. Second-order fluid flow models: Reflected Brownian motion in a random environment. *Oper. Res*, 43:77–88, 1995.

[18] T. Kato. *Perturbation Theory for Linear Operators*. Springer-Verlag, New York, 2nd edition, 1976.

[19] G. Latouche and V. Ramaswami. *Introduction to Matrix Analytic Methods in Stochastic Modeling*. ASA-SIAM Series on Statistics and Applied Probability. SIAM, Philadelphia PA, 1999.

[20] R. M. Loynes. Stationary waiting-time distribution for single-server queues. *Ann. Math. Statist.*, 33:1323–1339, 1962.

[21] M. F. Neuts. *Matrix-Geometric Solutions in Stochastic Models: An Algorithmic Approach*. The Johns Hopkins University Press, Baltimore, MD, 1981.

[22] Örjan Stenflo. Ergodic theorems for Markov chains represented by iterated function systems. *Bulletin of the Polish Academy of Sciences: Mathematics*, 49:427–443, 2001.

[23] V. Ramaswami. Matrix analytic methods for stochastic fluid flows. In D. Smith and P. Hey, editors, *Teletraffic Engineering in a Competitive World (Proceedings of the 16th International Teletraffic Congress)*, pages 1019–1030. Elsevier Science B.V., Edinburgh, UK, 1999.

[24] V. Ramaswami. A fluid introduction to Brownian motion and stochastic integration. In G. Latouche, V. Ramaswami, J. Sethuraman, K. Sigman, M. Squillante, and D. Yao, editors, *Matrix-Analytic Methods in Stochastic Models*, volume 27 of *Springer Proceedings in Mathematics & Statistics*, chapter 10, pages 209–225. Springer Science, New York, NY, 2013.

[25] L. C. G. Rogers. Fluid models in queueing theory and Wiener-Hopf factorization of Markov chains. *Ann. Appl. Probab.*, 4:390–413, 1994.

[26] W. Whitt. Weak convergence of probability measures on the function space $C[0,\infty)$. *Annals of Mathematical Statistics*, 41(2):939–944, 1970.