Feedback control: two-sided Markov-modulated Brownian motion with instantaneous change of phase at boundaries

Guy Latouche* Giang T. Nguyen†

September 25, 2018

Abstract

We consider a Markov-modulated Brownian motion $\{Y(t), \rho(t)\}$ with two boundaries at 0 and $b > 0$, and allow for the controlling Markov chain $\{\rho(t)\}$ to instantaneously undergo a change of phase upon hitting either of the two boundaries at *semi-regenerative epochs* defined to be the first time the process reaches a boundary since it last hits the other boundary. We call this process a *flexible Markov-modulated Brownian motion.*

Using the recently-established links between stochastic fluid models and Markov-modulated Brownian motions, we determine important characteristics of first exit times of a Markov-modulated Brownian motion from an interval with a regulated boundary. These results allow us to follow a Markov-regenerative approach and obtain the stationary distribution of the flexible process. This highlights the effectiveness of the regenerative approach in analyzing Markov-modulated Brownian motions subject to more general boundary behaviours than the classic regulated boundaries.

Keywords: Fluid queues, Markov-modulated Brownian motion, regenerative processes, finite buffer, stationary distribution, feedback.

*Université Libre de Bruxelles, Département d'informatique, CP 212, Boulevard du Triomphe, 1050 Bruxelles, Belgium, latouche@ulb.ac.be.

†The University of Adelaide, School of Mathematical Sciences, SA 5005, Australia, giang.nguyen@adelaide.edu.au

1 Introduction

We analyze Markov-modulated Brownian motions (MMBMs) restricted to the interval $[0, b]$, the distinguishing feature being that the processes are allowed to undergo an instantaneous change of phase upon hitting either boundaries at *semi-regenerative epochs*. These correspond to the first times the process reaches a boundary after it hits the other boundary. We refer to these processes as *flexible Markov-modulated Brownian motion*.

The simplest example we have in mind is described as follows: consider a buffer of finite capacity b serving as temporary storage for data in a communication network. Assume that its content $X(t)$ evolves in time like a Brownian motion with parameters $\mu < 0$ and $\sigma^2 > 0$. Whenever the buffer gets full, data may be lost. To reduce such losses, additional bandwidth is allocated, or the input stream is throttled, or other measures are taken, such that the mean drift becomes $\mu' < \mu$. Once the buffer is emptied, the process returns to its normal mode of operation, until it gets full again, etc. See Figure 1 for two sample trajectories. The graph at the top depicts a regulated Brownian motion without change of parameters, the one at bottom depicts a flexible version of the process; we have marked with a thick line the interval of time $(0.46, 0.62)$ during which the drift is μ'.

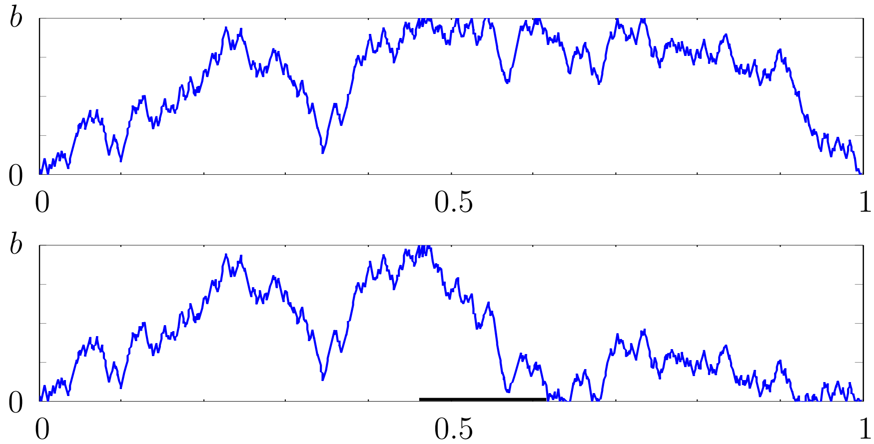

Figure 1: Sample trajectory of a regulated Brownian motion (top) and of a flexible Brownian motion (bottom). The parameters are $b = 4$, $\mu = -1$, $\mu' = -10$, $\sigma^2 = 10$.

In the general formulation, the evolution of the buffer is controlled by a continuous-time Markov chain called the process of phases. Whenever the buffer reaches a boundary for the first time after it has visited the other boundary, the phase is allowed to undergo an instantaneous change. It is clear that the epochs when change may occur at a boundary form a semi-regenerative set of points.

The stationary distribution of a Markov-modulated Brownian motion restricted to a strip $[0, b]$ is well-analyzed, as long as the boundaries are absorbing or regulated as in Ivanovs [14]. Here, we determine the stationary distribution of our flexible MMBMs by following a Markov-regenerative approach. The usefulness of this approach has been repeatedly demonstrated for fluid queues (aka first-order fluid processes) as in da Silva Soares and Latouche [8, 10], Latouche and Taylor [19], Bean and O'Reilly [1]. It has been adapted in Latouche and Nguyen [18] to MMBMs with one *reactive* boundarys, by which we mean any boundary that is not absorbing or regulated. As we shall demonstrate later, the effectiveness of our method extends far beyond the model analyzed here.

To determine the stationary distribution, the key ingredients needed are the expected time spent in an interval $[0, x]$ and in a given phase during an excursion from 0 to b for the process regulated at 0, and from b to 0 for the process regulated at b, as well as the distribution of the phase upon reaching a boundary. To obtain these quantities, we rely on the connections between MMBMs and their approximating Markov-modulated fluid models, exemplified in Latouche and Nguyen [16, 17, 18].

Our results are related to those in Bean *et al.* [2]: the authors analyze sojourn times in specified intervals during various excursions for fluid queues with reactive boundaries. The main differences are that we consider jointly the level and the phase, and that we deal with Markov modulated Brownian motion. We should also mention the results in Breuer [4] about occupation times before a two-sided exit; we discuss later in some detail the connection with our results.

The paper is organized as follows. We give in the next section the technical definition of flexible Markov-modulated Brownian motions and we outline the Markov-regenerative approach for obtaining their stationary distribution. Section 3 includes background material and notation required for the paper. We determine in Sections 4 and 5 the first passage probabilities from one boundary to the other, and the expected time spent during these excursions. In Section 6, we bring all partial results together and determine the stationary distribution of a flexible MMBM. We give three numerical examples in Section 7, we compare in Section 8 our results to the existing literature, and we conclude in Section 9 with a discussion on the applicability of our approach to more complex models.

2 Flexible Markov-modulated Brownian motion

A *free-boundary* Markov-modulated Brownian motion is a two-dimensional Markov process $\{X(t), \kappa(t)\}_{t\geq 0}$ such that

$$X(t) = X(0) + \int_0^t \mu_{\kappa(s)} \, \mathrm{d}s + \int_0^t \sigma_{\kappa(s)} \, \mathrm{d}B(s),$$

where $\{B(t)\}_{t\geq 0}$ is a standard Brownian motion and $\{\kappa(t)\}$ is a continuous-time Markov chain on the state space $\mathcal{E} = \{1, \ldots, m\}$, with generator Q. We denote by $\Delta_\mu = \mathrm{diag}(\mu_i)_{i\in\mathcal{E}}$ and $\Delta_\sigma^2 = \mathrm{diag}(\sigma_i^2)_{i\in\mathcal{E}}$, respectively, the drift and variance matrices of the MMBM. We assume that $\sigma_i > 0$ for all i in order to significantly simplify the presentation.

A Markov-modulated Brownian motion $\{Z(t), \kappa(t)\}_{t\geq 0}$ with two *regulated* boundaries at 0 and at $b > 0$ is defined as

$$Z(t) = X(t) + R_0(t) - R_b(t),$$

where $\{R_0(t)\}_{t\geq 0}$ and $\{R_b(t)\}_{t\geq 0}$ are nonnegative, continuous and almost surely nondecreasing processes; $R_0(t)$ increases only when $Z(t)$ is at zero and $R_b(t)$ only when $Z(t)$ is at b. The regulators $\{R_0(t)\}_{t\geq 0}$ and $\{R_b(t)\}_{t\geq 0}$ are the minimal processes keeping $Z(t)$ in $[0, b]$.

Next, we allow the phase to change as a reaction to $Z(t)$ reaching either level b or level 0. The general idea is that the flexible MMBM evolves like $Z(t)$ during a regeneration interval (θ, θ'), but at time θ' the phase immediately switches to a new value according to the transition probability matrix $P^\bullet$, in case $Z(\theta') = b$, and according to another probability matrix P° in case $Z(\theta') = 0$. The technical details follow.

Our pathwise construction of a flexible MMBM $\{Y(t), \rho(t)\}_{t\geq 0}$ starts with a countably infinite supply of independent copies of regulated MMBMs $\{Z_n(k; t), \kappa_n(k; t)\}$, such that $\kappa_n(k; 0) = k$, for $n \geq 0$ and $k \in \mathcal{E}$. We assume that $Z_n(k; 0) = 0$ if n is even, and $Z_n(k; 0) = b$ for odd n. All processes have the same parameters: lower and upper bounds 0 and b, generator Q, drift and variance matrices Δ_μ and Δ_σ^2. Most of these processes will not be used, but they allow us to maintain independence where our construction requires it.

Without loss of generality, assume that $Y(0) = 0$ and $\rho(0) = i \in \mathcal{E}$, and let $\theta_0 = 0$. We define

$$Y(t) = Z_0(i; t), \qquad \rho(t) = \kappa_0(i; t) \qquad\qquad \text{for } 0 \leq t < \theta_1,$$

where $\theta_1 = \inf\{t > 0 : Z_0(i; t) = b\}$ is the first hitting time to level b. Upon hitting level b, the phase ρ instantaneously changes to some value j according to the transition matrix $P^\bullet$, and we define

$$Y(t) = Z_1(j; t-\theta_1), \qquad \rho(t) = \kappa_1(j; t-\theta_1) \qquad \text{for } \theta_1 \le t < \theta_1 + h_2,$$

where $h_2 = \inf\{t > 0 : Z_1(j;t) = 0\}$. Upon hitting level 0, ρ changes to a new value according to P°, and so on.

In general, starting with $\theta_0 = 0$ and $\rho(0) = i$, we recursively define for $n \ge 0$ the following:

$i_n = \rho(\theta_n)$,

$h_{n+1} = \inf\{t > 0 : Z_n(i_n; t) = b\,\mathbb{1}\{n \text{ is even}\}\}$,

$\theta_{n+1} = \theta_n + h_{n+1}$,

$\rho(\theta_{n+1})$ is obtained from the row $\kappa_n(i_n; h_{n+1})$ of the matrix $P^\bullet$ if n is even, and of P° if n is odd,

and

$$Y(t) = Z_N(\rho(\theta_N); t-S), \qquad \rho(t) = \kappa_N(\rho(\theta_N); t-S), \tag{1}$$

where $N = \arg\max\{s : \theta_s \le t\}$ and $S = \theta_N$. As $\lim_{n\to\infty} \theta_n = \infty$, this defines $\{Y(t), \rho(t)\}$ over the whole interval $[0, \infty)$.

By construction,

$$\begin{aligned} \theta_{2n+1} &= \inf\{t > \theta_{2n} : Y(t) = b\} \quad \text{for } n \ge 0, \\ \theta_{2n} &= \inf\{t > \theta_{2n-1} : Y(t) = 0\} \quad \text{for } n \ge 1, \end{aligned}$$

and $\{\theta_n\}_{n\ge 0}$ forms a set of semi-regenerative points for $\{Y(t), \rho(t)\}$ on the state space $\{0, b\} \times \mathcal{E}$. Because of the discontinuity introduced at the regeneration points, we shall need to define the two limits, from the left and from the right,

$$\rho(\theta_n^-) = \lim_{t\uparrow\theta_n} \rho(t), \qquad \rho(\theta_n^+) = \lim_{t\downarrow\theta_n} \rho(t);$$

the trajectories of $\{\rho(t)\}$ are right-continuous, and so $\rho(\theta_n) = \rho(\theta_n^+)$ for all n.

The semi-Markov kernel $D(\cdot)$ for the transitions between semi-regenerative points is defined as

$$\begin{aligned} D_{(x,i),(y,j)}(t) = \mathrm{P}[\theta_{n+1} - \theta_n \le t, \rho(\theta_{n+1}^-) = j, Y(\theta_{n+1}) = y \mid \\ Y(\theta_n) = x, \rho(\theta_n^-) = i], \end{aligned} \tag{2}$$

where $x, y \in \{0, b\}$ and $i, j \in \mathcal{E}$. By construction, again, the structure of D is

$$D(t) = \begin{bmatrix} 0 & D_0(t) \\ D_b(t) & 0 \end{bmatrix}, \tag{3}$$

where $D_x(\cdot)$ records the transition probabilities from $\rho(\theta_n^-)$ to $\rho(\theta_{n+1}^-)$ given that $Y(\theta_n) = x$.

We need to make some irreducibility assumption at this point, for the arguments that follow to hold. In view of our elementary introductory example, we should not automatically assume, as is usually done, that the generator Q is irreducible: in that example, one may consider that there is one phase and two distinct sets of parameters, one for the even-numbered and one for the odd-numbered intervals. One may also view that there are two phases, one with mean drift μ and one with mean drift μ', and that there is no connection between the two, except through the boundary feedback mechanism. For the time being, we make the following assumption only and we give in Section 6 simple conditions for it to hold.

Assumption 2.1 *The transition matrix D is irreducible. In other words, for any pair of states (x, i) and (y, j), there is a path of positive probability from $(Y(0) = x, \rho(0) = i)$ to $(Y(\theta_n) = y, \rho(\theta_n) = j)$, for some n.*

Next, we define the matrix $\Theta(x)$ of expected sojourn times: for $y = 0$ or b, i and j in $\mathcal{E}$, and $x \geq 0$, the component $\Theta_{y;i,j}(x)$ is the expected time spent by the process in the set $[0, x] \times \{j\}$ during a regeneration interval $[\theta_n, \theta_{n+1})$, conditionally given that $Y(\theta_n) = y$ and $\rho(\theta_n^-) = i$. We display that matrix as

$$\Theta(x) = \begin{bmatrix} \Theta_0(x) \\ \Theta_b(x) \end{bmatrix}.$$

By Çinlar [6, Sect.10.4, Prop.4.9], the joint stationary distribution $\Pi(x)$ of $\{Y(t), \rho(t)\}$ is given by

$$\Pi(x) = (\boldsymbol{d}\,\boldsymbol{\theta})^{-1}\boldsymbol{d}\,\Theta(x), \tag{4}$$

where $\boldsymbol{\theta} = \Theta(\infty)\mathbf{1}$, with $\mathbf{1}$ a column vector of 1s, is the vector of expected length of a regenerative interval, given the initial state, and $\boldsymbol{d}$ is the stationary distribution of the phase immediately before the end of the next interval, that is, $\boldsymbol{d}D(\infty) = \boldsymbol{d}$, and $\boldsymbol{d}\mathbf{1} = 1$.

In Theorem 2.2 below, we express $\Pi(x)$ directly in terms of properties of the regulated Brownian motion $\{Z(t), \kappa(t)\}$. For that purpose, we introduce the transition probability matrices

$$(H_0)_{ij} = \mathrm{P}[\delta_b < \infty, \kappa(\delta_b) = j | Z(0) = 0, \kappa(0) = i], \tag{5}$$
$$(H_b)_{ij} = \mathrm{P}[\delta_0 < \infty, \kappa(\delta_0) = j | Z(0) = b, \kappa(0) = i], \tag{6}$$

where $\delta_x = \inf\{t > 0 : Z(t) = x\}$ is the first passage time to level x, and we define the matrices of expected sojourn times in the interval $[0, x]$ during an excursion from level 0 to level b, and from level b to level 0:

$$(M_0(x))_{ij} = \mathrm{E}[\int_0^{\delta_b} \mathbb{1}\{Z(s) \in [0,x], \kappa(s) = j\} \; \mathrm{d}s \,|\, Z(0) = 0, \kappa(0) = i], \quad (7)$$

$$(M_b(x))_{ij} = \mathrm{E}[\int_0^{\delta_0} \mathbb{1}\{Z(s) \in [0,x], \kappa(s) = j\} \; \mathrm{d}s \,|\, Z(0) = b, \kappa(0) = i]. \quad (8)$$

Theorem 2.2 *The stationary distribution* $\Pi(x)$ *of the flexible Markov-modulated Brownian motion* $\{Y(t), \rho(t)\}$ *is given by* $\Pi(x) = (\boldsymbol{\nu}\,\boldsymbol{m})^{-1}\boldsymbol{\nu} M(x)$, *where* $\boldsymbol{\nu} = \begin{bmatrix}\boldsymbol{\nu}_0 & \boldsymbol{\nu}_b\end{bmatrix}$ *with*

$$\boldsymbol{\nu}_0 = \boldsymbol{\nu}_0 H_0 P^{\bullet} H_b P^{\circ} \qquad \boldsymbol{\nu}_b = \boldsymbol{\nu}_0 H_0 P^{\bullet},$$

$\boldsymbol{\nu}_0$ *being unique up to a multiplicative constant,*

$$M(x) = \begin{bmatrix} M_0(x) \\ M_b(x) \end{bmatrix},$$

and $\boldsymbol{m} = M(b)\mathbf{1}$.

Proof By our definition (2) of the semi-Markov kernel, starting from one of the two boundaries at time θ_n, a new phase is chosen with the corresponding matrix P° or $P^{\bullet}$, and then a process stochastically identical to $\{Z(t)\}$ evolves until it reaches the other boundary. Thus, $D_0(\infty) = P^{\circ} H_0$ and $D_b(\infty) = P^{\bullet} H_b$ or, in matrix form,

$$D(\infty) = \begin{bmatrix} P^{\circ} & 0 \\ 0 & P^{\bullet} \end{bmatrix} \begin{bmatrix} 0 & H_0 \\ H_b & 0 \end{bmatrix}. \quad (9)$$

The stationary probability vector $\boldsymbol{d}$ of $D(\infty)$, written as $\boldsymbol{d} = \begin{bmatrix}\boldsymbol{d}_0 & \boldsymbol{d}_b\end{bmatrix}$, satisfies the equations

$$\boldsymbol{d}_0 = \boldsymbol{d}_b P^{\bullet} H_b, \qquad \boldsymbol{d}_b = \boldsymbol{d}_0 P^{\circ} H_0,$$

or

$$\boldsymbol{d}_0 = \boldsymbol{d}_0 P^{\circ} H_0 P^{\bullet} H_b, \qquad \boldsymbol{d}_b = \boldsymbol{d}_0 P^{\circ} H_0. \quad (10)$$

The matrix $P^{\circ} H_0 P^{\bullet} H_b$ is the transition probability matrix from a phase immediately before a regeneration at level 0 to the phase immediately before the next regeneration at 0; it is irreducible by Assumption 2.1 and so the vector $\boldsymbol{d}_0$ is unique, up to a multiplicative constant.

By the same argument that leads to (9), we conclude that

$$\begin{bmatrix} \Theta_0(x) \\ \Theta_b(x) \end{bmatrix} = \begin{bmatrix} P^{\circ} & 0 \\ 0 & P^{\bullet} \end{bmatrix} \begin{bmatrix} M_0(x) \\ M_b(x) \end{bmatrix},$$

and so, by (4),

$$\Pi(x) = c \begin{bmatrix} \boldsymbol{d}_0 & \boldsymbol{d}_b \end{bmatrix} \begin{bmatrix} P^\circ & 0 \\ 0 & P^\bullet \end{bmatrix} \begin{bmatrix} M_0(x) \\ M_b(x) \end{bmatrix}$$

for some normalizing constant c. The remainder of the proof is immediate once we define $\boldsymbol{\nu}_0 = \boldsymbol{d}_0 P^\circ$ and $\boldsymbol{\nu}_b = \boldsymbol{d}_b \mathrm{P}^\bullet$. □

In consequence of Theorem 2.2, we need only to take into consideration a simple MMBM with two regulated boundaries, and to focus on one excursion from a boundary to the other. This we do in Sections 4 and 5. Before that, we recall some basic properties of MMBMs.

3 Background material and notation

We analyze in Sections 4 and 5 a regulated process controlled by a phase process with an irreducible generator Q and a unique set of parameters $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$. Unlike the process defined in Section 2, there is no reaction of the phase when the buffer reaches a boundary and, to ensure that there is no confusion with the matrices H and M defined in (5–8), we use the symbols $\mathcal{H}$ and $\mathcal{M}$.

Two matrices play an important role in the analysis of MMBMs. They are denoted as U and $\widehat{U}$ in the present paper, and U and $-\widehat{U}$ are solutions of the matrix equation

$$\Delta_\sigma^2 X^2 + 2\Delta_\mu X + 2Q = 0, \tag{11}$$

with U being the minimal solution and $-\widehat{U}$ being the maximal solution, meaning that the eigenvalues of U are the roots of the polynomial $\det(\Delta_\sigma^2 z^2 + 2\Delta_\mu z + 2Q)$ in the negative half complex plane and the eigenvalues of $-\widehat{U}$ are the roots in the positive half-plane (see D'Auria *et al.* [12, Section 4] and Latouche and Nguyen [16, Lemma 5.3]).

Both matrices are irreducible generators: $\widehat{U}$ is the generator of the Markov chain $\{\kappa(\delta_x)\}_{x\ge 0}$ and U is the generator of the Markov chain $\{\kappa(\delta_{-x})\}_{x\ge 0}$. One recognizes three different cases, based on the sign of the mean drift $\boldsymbol{\alpha\mu}$, where $\boldsymbol{\alpha}$ is the stationary distribution of Q ($\boldsymbol{\alpha} Q = \mathbf{0}$, $\boldsymbol{\alpha}\mathbf{1} = 1$):

1. If $\boldsymbol{\alpha\mu} > 0$, then $\widehat{U}\mathbf{1} = \mathbf{0}$ and $U\mathbf{1} \le \mathbf{0}$, with at least one strict inequality, U is nonsingular.

2. If $\boldsymbol{\alpha\mu} < 0$, $U\mathbf{1} = \mathbf{0}$ and $\widehat{U}\mathbf{1} \le \mathbf{0}$, with at least one strict inequality, $\widehat{U}$ is nonsingular.

3. If $\boldsymbol{\alpha\mu} = 0$, both $U\mathbf{1}$ and $\widehat{U}\mathbf{1}$ are equal to $\mathbf{0}$.

We prove in [16, 17] that Markov-modulated Brownian motions can be approximated by a family of fast oscillating free-boundary fluid processes $\{X^\lambda(t), \beta^\lambda(t), \kappa^\lambda(t)\}_{t\geq 0}$ parameterized by $\lambda > 0$, where $\{\beta^\lambda(t), \kappa^\lambda(t)\}$ is a two-dimensional Markov process on the state space $\mathcal{S} = \{1,2\} \times \mathcal{E}$, with generator

$$T^\lambda = \begin{bmatrix} Q - \lambda I & \lambda I \\ \lambda I & Q - \lambda I \end{bmatrix}, \tag{12}$$

and the level process $\{X^\lambda(t)\}$ is driven by the phase $\{\beta^\lambda(t), \kappa^\lambda(t)\}$ as follows

$$X^\lambda(t) = \int_0^t C^\lambda_{\beta^\lambda(u),\kappa^\lambda(u)}\, \mathrm{d}u,$$

with

$$C^\lambda = \begin{bmatrix} \Delta_\mu + \sqrt{\lambda}\Delta_\sigma & \\ & \Delta_\mu - \sqrt{\lambda}\Delta_\sigma \end{bmatrix}. \tag{13}$$

Consider the family of regulated processes $\{Z^\lambda(t), \beta^\lambda(t), \kappa^\lambda(t)\}$ with boundaries at 0 and at $b > 0$, and initial phase $\beta^\lambda(0)$ equal to 1 or 2 with equal probability 0.5, for all λ. By [16, Thm.3.1], the regulated MMBM $\{Z(t), \kappa(t)\}$ defined in Section 2 is the weak limit of the projected process $\{Z^\lambda(t), \kappa^\lambda(t)\}$, and the stationary distribution of the former arises as the limit of that of the latter as $\lambda \to \infty$. In consequence, the sojourn time matrix $\mathcal{M}(x)$ and the first passage probability matrix $\mathcal{H}$ are the limits of the corresponding matrices for the projected process as $\lambda \to \infty$.

We partition the state space into the subsets $\mathcal{S}_+$ and $\mathcal{S}_-$, where $\mathcal{S}_+ = \{(i,j) \in \mathcal{S} : c_{ij} > 0\}$ and $\mathcal{S}_- = \{(i,j) \in \mathcal{S} : c_{ij} < 0\}$. For sufficiently large λ, $\mathcal{S}_+ = \{(1,j) : j \in \mathcal{E}\}$ and $\mathcal{S}_- = \{(2,j) : j \in \mathcal{E}\}$. Several matrices are partitioned in a conformant manner. For instance, we write T^λ as

$$T^\lambda = \begin{bmatrix} T^\lambda_{++} & T^\lambda_{+-} \\ T^\lambda_{-+} & T^\lambda_{--} \end{bmatrix}.$$

Two first passage probabilities are needed in the next section. One is Ψ^λ_b, indexed by $\mathcal{S}_+ \times \mathcal{S}_-$, which records the probability that, starting from level 0 in a state of $\mathcal{S}_+$, Z^λ returns to level 0 before reaching level b; the other matrix, Λ^λ_b is indexed by $\mathcal{S}_+ \times \mathcal{S}_+$ and records the probability that level b is reached before a return to level 0:

$$(\Psi^\lambda_b)_{(1,i)(2,j)} = \mathrm{P}[\delta^\lambda_0 < \delta^\lambda_b, \kappa^\lambda(\delta^\lambda_0) = j \,|\, Z^\lambda(0) = 0, \beta^\lambda(0) = 1, \kappa^\lambda(0) = i], \tag{14}$$
$$(\Lambda^\lambda_b)_{(1,i)(2,j)} = \mathrm{P}[\delta^\lambda_b < \delta^\lambda_0, \kappa^\lambda(\delta^\lambda_b) = j \,|\, Z^\lambda(0) = 0, \beta^\lambda(0) = 1, \kappa^\lambda(0) = i].$$

4 Transition probability matrices

Starting from any state in $\mathcal{S}$ at time 0, the process $\{Z^\lambda(t), \beta^\lambda(t), \kappa^\lambda(t)\}$ is necessarily in a state of $\mathcal{S}_+$ at time δ_b^λ, and the matrix H_0^λ of first passage probability from level 0 to level b has the structure

$$H_0^\lambda = \begin{bmatrix} H_{++}^\lambda \\ H_{-+}^\lambda \end{bmatrix}.$$

(We omit the subscript 0 for the sub-matrices in the following calculations as there is no ambiguity.) Since $\beta^\lambda(0)$ is equal to 1 or 2 with equal probabilities, we have

$$H_0 = \lim_{\lambda\to\infty}(0.5H_{++}^\lambda + 0.5H_{-+}^\lambda). \tag{15}$$

Now, starting in a phase of $\mathcal{S}_-$, the fluid remains at level 0 until it first moves to a phase of $\mathcal{S}_+$, with the transition probability matrix

$$(-T_{--}^\lambda)^{-1}T_{-+}^\lambda = (\lambda I - Q)^{-1}\lambda I = (I - \frac{1}{\lambda}Q)^{-1},$$

so that $H_{-+}^\lambda = (I - \frac{1}{\lambda}Q)^{-1}H_{++}^\lambda$, and we see from (15) that

$$\mathcal{H}_0 = \lim_{\lambda\to\infty} H_{++}^\lambda. \tag{16}$$

Starting from level 0 in a phase of $\mathcal{S}_+$, the fluid queue may move directly to level b without returning to level 0, or it may return to level 0 before having reached level b. Thus,

$$\begin{aligned} H_{++}^\lambda &= \Lambda_b^\lambda + \Psi_b^\lambda H_{-+}^\lambda \\ &= \Lambda_b^\lambda + \Psi_b^\lambda(I - \frac{1}{\lambda}Q)^{-1}H_{++}^\lambda \\ &= (I - \Psi_b^\lambda(I - \frac{1}{\lambda}Q)^{-1})^{-1}\Lambda_b^\lambda \end{aligned}$$

and we find that

$$\mathcal{H}_0 = \lim_{\lambda\to\infty}(I - \Psi_b^\lambda)^{-1}\Lambda_b^\lambda. \tag{17}$$

With this, we are in position to prove the following theorem.

Theorem 4.1 *Consider an MMBM regulated at level 0. The matrix $\mathcal{H}_0$ of first passage probability from level 0 to level b is*

$$\mathcal{H}_0 = (-P_b)^{-1}L_b,$$

where L_b and P_b are solutions of the linear system

$$\begin{bmatrix} L_b & P_b \end{bmatrix} \begin{bmatrix} I & e^{Ub} \\ e^{\widehat{U}b} & I \end{bmatrix} = \Delta_\sigma \begin{bmatrix} -\widehat{U}e^{\widehat{U}b} & U \end{bmatrix}. \tag{18}$$

The matrix P_b is a sub-generator and is nonsingular.

If $\boldsymbol{\alpha\mu} \neq 0$, then the system (18) is nonsingular and its solution may be written as

$$L_b = -\Delta_\sigma(U + \widehat{U})e^{\widehat{U}b}(I - e^{Ub}e^{\widehat{U}b})^{-1} \tag{19}$$

$$P_b = \Delta_\sigma(U + \widehat{U}e^{\widehat{U}b}e^{Ub})(I - e^{\widehat{U}b}e^{Ub})^{-1}. \tag{20}$$

In that case,

$$\mathcal{H}_0 = e^{\widehat{U}b} + (e^{-Ub} - e^{\widehat{U}b})(Ue^{-Ub} + \widehat{U}e^{\widehat{U}b})^{-1}\widehat{U}e^{\widehat{U}b}.$$

If $\boldsymbol{\alpha\mu} = 0$, then the system (18) is singular and one needs the additional equation

$$L_b(b\mathbf{1} - Q^{\#}\boldsymbol{\mu}) - P_bQ^{\#}\boldsymbol{\mu} = \boldsymbol{\sigma} \tag{21}$$

to completely characterize L_b and P_b, where $Q^{\#}$ is the group inverse of Q.

Proof It results from [16, Lemma 5.5] that

$$\Lambda_b^\lambda = \frac{1}{\sqrt{\lambda}}L_b + O(\frac{1}{\lambda}), \qquad \Psi_b^\lambda = I + \frac{1}{\sqrt{\lambda}}P_b + O(\frac{1}{\lambda}),$$

where (L_b, P_b) is a solution of (18) and that P_b is nonsingular. We readily conclude from (17) that $\mathcal{H}_0 = (-P_b)^{-1}L_b$. Furthermore, (19, 20) directly result from [16, Eqn (33) and (34)] when $\boldsymbol{\alpha\mu} \neq 0$.

If $\boldsymbol{\alpha\mu} = 0$, then both e^{Ub} and $e^{\widehat{U}b}$ are stochastic matrices, the coefficient matrix of (18) is singular, and we need an additional equation. D'Auria *et al.* [11] analyze first exit probabilities for the MMBM process $\{X(t), \kappa(t)\}$ and determine the exit probabilities from the interval $[0, b]$

$$P(x, 0) = \mathrm{P}[\delta_0 < \delta_b, \kappa(\delta_0) | X(0) = x],$$
$$P(x, b) = \mathrm{P}[\delta_b < \delta_0, \kappa(\delta_b) | X(0) = x],$$

for $0 \leq x \leq b$. Equations (56, 58) in [11] may be written as

$$\begin{bmatrix} P(x, b) & P(x, 0) \end{bmatrix} \begin{bmatrix} I & e^{Ub} \\ e^{\widehat{U}b} & I \end{bmatrix} = \begin{bmatrix} e^{\widehat{U}(b-x)} & e^{Ux} \end{bmatrix} \tag{22}$$

and

$$P(x,b)((b-x)\mathbf{1}+\boldsymbol{h})+P(x,0)(-x\mathbf{1}+\boldsymbol{h})=\boldsymbol{h}, \tag{23}$$

with $\boldsymbol{h}$ being any solution of the system $Q\boldsymbol{h}=-\boldsymbol{\mu}$.

The matrix Q has one eigenvalue equal to zero and such solutions are of the form $\boldsymbol{h}=-Q^{\#}\boldsymbol{\mu}+c\mathbf{1}$, where c is an arbitrary scalar, and $Q^{\#}$ is the unique solution of the linear system $XQ=I-\mathbf{1}\boldsymbol{\alpha}$, $X\mathbf{1}=\mathbf{0}$; that matrix is called the group inverse of Q (Campbell and Meyer [5]). As Q is a generator, $Q^{\#}$ is also called the deviation matrix of the Markov process with generator Q (Coolen-Schrijner and van Doorn [7]), and one has $Q^{\#}=\int_0^\infty(e^{Qu}-\mathbf{1}\boldsymbol{\alpha})\,\mathrm{d}u$.

In addition, it is shown in [16, Section 6.2] that

$$L_b=\Delta_\sigma\lim_{x\to 0}\frac{\partial}{\partial x}P(x,b) \qquad \text{and} \qquad P_b=\Delta_\sigma\lim_{x\to 0}\frac{\partial}{\partial x}P(x,0).$$

Premultiplying both sides of (22) by Δ_σ and taking the derivative, we obtain(18) as $x\to 0$. Similarly,

$$L_b(b\mathbf{1}+\boldsymbol{h})-\Delta_\sigma P(0,b)\mathbf{1}+P_b\boldsymbol{h}-\Delta_\sigma P(0,0)\mathbf{1}=\mathbf{0},$$

follows from (23). As $P(0,b)=0$ and $P(0,0)=I$, the last equation is identical to (21) if we chose $\boldsymbol{h}=-Q^{\#}\boldsymbol{\mu}$. This completes the proof. $\square$

We may follow a similar line of argument to determine the matrix $\mathcal{H}_b$ of first passage probabilities from the upper boundary to the boundary at level 0, We may also, as an alternative, define the level-reversed process $\{\widehat{X}(t),\kappa(t)\}$, where $\widehat{X}(t)=-Z(t)$. For this process, the fluid rate vector becomes $\widehat{\boldsymbol{\mu}}=-\boldsymbol{\mu}$, the rôles of the matrices U and $\widehat{U}$ are exchanged, and the first passage probability matrix $\widehat{\mathcal{H}}_0$ from 0 to b of the regulated process of $\{\widehat{Z}(t)\}$ is equal to $\mathcal{H}_b$, the first passage probability matrix of $\{Z(t)\}$ from b to 0. The proof of the corollary below is immediate and is omitted.

Corollary 4.2 *Consider an MMBM regulated at level 0 and b. The matrix $\mathcal{H}_b$ of first passage probability from the boundary b to the boundary 0 is*

$$\mathcal{H}_b=(-\widehat{P}_b)^{-1}\widehat{L}_b,$$

where

$$\begin{bmatrix}\widehat{P}_b & \widehat{L}_b\end{bmatrix}\begin{bmatrix}I & e^{Ub}\\ e^{\widehat{U}b} & I\end{bmatrix}=\Delta_\sigma\begin{bmatrix}\widehat{U} & -Ue^{Ub}\end{bmatrix}. \tag{24}$$

The matrix $\widehat{P}_b$ is an irreducible subgenerator and is nonsingular.

If $\boldsymbol{\alpha\mu}\neq 0$, then

$$\begin{aligned}\widehat{L}_b&=-\Delta_\sigma(U+\widehat{U})e^{Ub}(I-e^{\widehat{U}b}e^{Ub})^{-1}\\ \widehat{P}_b&=\Delta_\sigma(\widehat{U}+Ue^{Ub}e^{\widehat{U}b})(I-e^{Ub}e^{\widehat{U}b})^{-1},\end{aligned}$$

and
$$\mathcal{H}_b = e^{Ub} + (e^{-\widehat{U}b} - e^{Ub})(\widehat{U}e^{-\widehat{U}b} + Ue^{Ub})^{-1}Ue^{Ub}.$$
If $\boldsymbol{\alpha\mu} = 0$*, then* $\widehat{L}_b$ *and* $\widehat{P}_b$ *are determined by the system (24) and the additional equation*
$$\widehat{L}_b(b\mathbf{1} + Q^{\#}\boldsymbol{\mu}) + \widehat{P}_b Q^{\#}\boldsymbol{\mu} = \boldsymbol{\sigma}. \tag{25}$$
□

5 Expected time in $[0, x]$ during an excursion

We determine in this section the matrix $\mathcal{M}_0(x)$ of expected sojourn time of a regulated MMBM $\{Z(t), \kappa(t)\}$ during an excursion from 0 to b. It soon becomes clear that to do so, we need to deal at the same time with excursions from b to 0 by the same process. The matrix of expected sojourn time in $[0, x]$ during such an excursion is denoted as $\mathcal{M}_b(x)$.

We define $M_0^\lambda(x)$ to be the matrix of expected sojourn time of the rapidly switching process $\{Z^\lambda(t), \beta^\lambda(t), \kappa^\lambda(t)\}_{t\geq 0}$ in $[0, x]$ during an excursion from 0 to b:
$$\begin{aligned}(M_0^\lambda(x))_{(\ell,i)(k,j)} = \mathrm{E}[\int_0^{\theta^\lambda} \mathbb{1}\{Z^\lambda(s) \in [0, x], \beta^\lambda(s) = k, \kappa^\lambda(s) = j\} \\ |Z^\lambda(0) = 0, \beta^\lambda(0) = \ell, \kappa^\lambda(0) = i],\end{aligned} \tag{26}$$
where $\theta^\lambda = \inf\{t > 0 : Z^\lambda(t) = b\}$ is the first passage time to level b. We partition that matrix as
$$M_0^\lambda(x) = \begin{bmatrix} M_{0;++}^\lambda(x) & M_{0;+-}^\lambda(x) \\ M_{0;-+}^\lambda(x) & M_{0;--}^\lambda(x) \end{bmatrix}$$
and, by an argument similar to the one that leads to (16), we find that $\mathcal{M}_0(x) = \lim_{\lambda\to\infty}(M_{0;++}^\lambda(x) + M_{0;+-}^\lambda(x))$.

Next, we define $M_b^\lambda(x)$ to be the matrix of expected sojourn in $[0, x]$ during an excursion from b to 0 and we partition it as
$$M_b^\lambda(x) = \begin{bmatrix} M_{b;++}^\lambda(x) & M_{b;+-}^\lambda(x) \\ M_{b;-+}^\lambda(x) & M_{b;--}^\lambda(x) \end{bmatrix};$$
one shows that $\mathcal{M}_b(x) = \lim_{\lambda\to\infty}(M_{b;-+}^\lambda(x) + M_{b;--}^\lambda(x))$. To simplify our equations in the remainder of this section, we write
$$M_{0;+}^\lambda = \begin{bmatrix} M_{0;++}^\lambda(x) & M_{0;+-}^\lambda(x) \end{bmatrix} \qquad \text{and} \qquad M_{b;-}^\lambda = \begin{bmatrix} M_{b;-+}^\lambda(x) & M_{b;--}^\lambda(x) \end{bmatrix},$$

and we summarize as follows the discussion above:

$$\begin{bmatrix} \mathcal{M}_0(x) \\ \mathcal{M}_b(x) \end{bmatrix} = \lim_{\lambda\to\infty} \begin{bmatrix} M^\lambda_{0;+}(x) \\ M^\lambda_{b;-}(x) \end{bmatrix} \begin{bmatrix} I \\ I \end{bmatrix}. \tag{27}$$

Theorem 5.1 *If $\boldsymbol{\alpha\mu} \neq 0$, then*

$$\begin{bmatrix} \mathcal{M}_0(x) \\ \mathcal{M}_b(x) \end{bmatrix} = 2 \begin{bmatrix} -P_b^{-1} & \\ & -\widehat{P}_b^{-1} \end{bmatrix} \begin{bmatrix} I & e^{Kb} \\ e^{\widehat{K}b} & I \end{bmatrix}^{-1} \begin{bmatrix} \mathcal{F}(K;x) \\ e^{\widehat{K}(b-x)}\mathcal{F}(\widehat{K};x) \end{bmatrix} \Delta_\sigma^{-1}, \tag{28}$$

where

$$\mathcal{F}(A;x) = \int_0^x e^{Au}\, \mathrm{d}u, \tag{29}$$

and

$$K = \Delta_\sigma U \Delta_\sigma^{-1} + 2\Delta_\sigma^{-2}\Delta_\mu \qquad \textit{and} \qquad \widehat{K} = \Delta_\sigma \widehat{U} \Delta_\sigma^{-1} - 2\Delta_\sigma^{-2}\Delta_\mu.$$

To prove this, we we proceed in three preliminary steps: we express $M^\lambda_{0;+}(x)$ and $M^\lambda_{b;-}(x)$ in terms of exit times from the interval $(0,b)$, next we analyze first passage times for the unregulated fluid process, and we establish a connection between the two. In the final step we prove (28) through a limiting argument.

Step A. We define the matrix $N_0^\lambda(x)$ of sojourn time in $[0,x]$ until $\{X^\lambda(t)\}$ hits *either* level 0 *or* level b, starting from 0 in a phase of $\mathcal{S}_+$:

$$\begin{aligned}(N_0^\lambda(x))_{(1,i);(k,j)} = \mathrm{E}[\int_0^{\delta_0^\lambda \wedge \delta_b^\lambda} \mathbb{1}\{X^\lambda(s) \in [0,x], \beta^\lambda(s) = k, \kappa^\lambda(s) = j\}\, \mathrm{d}s \\ |X^\lambda(0) = 0, \beta^\lambda(0) = 1, \kappa^\lambda(0) = i],\end{aligned} \tag{30}$$

for $(1,i) \in \mathcal{S}_+$, $(k,j) \in \mathcal{S}$.

Lemma 5.2 *The matrix $M^\lambda_{0;+}(x)$ of expected sojourn time in $[0,x]$ during the excursion $[0,\theta^\lambda]$ from level 0 to level b is given by*

$$M^\lambda_{0;+}(x) = (I - \Psi_b^\lambda(-T^\lambda_{--})^{-1}T^\lambda_{-+})^{-1}(N_0^\lambda(x) + \Psi_b^\lambda \begin{bmatrix} 0 & (-T^\lambda_{--})^{-1} \end{bmatrix}), \tag{31}$$

where Ψ_b^λ and $N_0^\lambda(x)$ are defined in (14) and (30), respectively.

Proof We decompose the interval $[0,\theta^\lambda]$ as $[0,\delta_0^\lambda \wedge \delta_b^\lambda] \cup [\delta_0^\lambda \wedge \delta_b^\lambda, \theta^\lambda]$, and obtain

$$M^\lambda_{0;+}(x) = N_0^\lambda(x) + \Psi_b^\lambda \begin{bmatrix} 0 & (-T^\lambda_{--})^{-1} \end{bmatrix} + \Psi_b^\lambda(-T^\lambda_{--})^{-1}T^\lambda_{-+}M^\lambda_{0;+}(x). \tag{32}$$

To justify this, we observe that the process must accumulate time in $[0, x]$ until it hits one of the boundaries; this corresponds to the first term in (32). With probability Ψ_b^λ, the process has returned to level 0, where it accumulates more time (the second term), and then leaves level 0 and accumulates time during the remainder of the excursion (the third term). Equation (31) immediately follows. $\square$

The proof of the next lemma is omitted as it merely mimics the proof of Lemma 5.2. For excursions that start in b, we define a new set of matrices: the transition probability matrices

$$\begin{aligned} \widehat{\Psi}_b^\lambda &= \mathrm{P}[\delta_b^\lambda < \delta_0^\lambda, \kappa^\lambda(\delta_b^\lambda) \,|\, Z^\lambda(0) = b, \beta^\lambda(0) = 2, \kappa^\lambda(0)] \qquad \text{on } \mathcal{S}_- \times \mathcal{S}_+, \quad (33)\\ \widehat{\Lambda}_b^\lambda &= \mathrm{P}[\delta_0^\lambda < \delta_b^\lambda, \kappa^\lambda(\delta_0^\lambda) \,|\, Z^\lambda(0) = b, \beta^\lambda(0) = 2, \kappa^\lambda(0)] \qquad \text{on } \mathcal{S}_- \times \mathcal{S}_-, \end{aligned}$$

and the matrix $N_b^\lambda(x)$ of sojourn time in $[0, x]$ until $\{X^\lambda(t)\}$ hits level 0 or level b, starting from level b in a phase of $\mathcal{S}_-$:

$$\begin{aligned} N_b^\lambda(x) = \mathrm{E}[\int_0^{\delta_0^\lambda \wedge \delta_b^\lambda} & \mathbb{1}\{X^\lambda(s) \in [0, x], \beta^\lambda(s), \kappa^\lambda(s)\} \,\mathrm{d}s \\ & |X^\lambda(0) = b, \beta^\lambda(0) = 2, \kappa^\lambda(0)] \qquad (34) \end{aligned}$$

on $\mathcal{S}_- \times \mathcal{S}$.

Lemma 5.3 *The matrix $M_{b;-}^\lambda(x)$ of expected sojourn time in $[0, x]$ during an excursion from level b to level 0 is given by*

$$M_{b;-}^\lambda(x) = (I - \widehat{\Psi}_b^\lambda(-T_{++}^\lambda)^{-1} T_{+-}^\lambda)^{-1} (N_b^\lambda(x) + \widehat{\Psi}_b^\lambda \begin{bmatrix} 0 & (-T_{++}^\lambda)^{-1} \end{bmatrix})$$

where $\widehat{\Psi}_b^\lambda$ and $N_b^\lambda(x)$ are defined in (33) and (34), respectively. $\square$

Step B. Next, we characterize expected sojourn times during intervals $(0, \delta_0^\lambda)$ or $(0, \delta_b^\lambda)$ for the *unregulated* process $\{X^\lambda(t), \beta^\lambda(t), \kappa^\lambda(t)\}$. We define the matrices

$$\begin{aligned} \Gamma_0^\lambda(x)) = \mathrm{E}[\int_0^{\delta_0^\lambda} & \mathbb{1}\{X^\lambda(s) \in [0, x], \beta^\lambda(s), \kappa^\lambda(s)\} \,\mathrm{d}s \\ & |X^\lambda(0) = 0, \beta^\lambda(0) = 1, \kappa^\lambda(0)], \end{aligned}$$

indexed by $\mathcal{S}_+ \times \mathcal{S}$, and

$$\begin{aligned} \widehat{\Gamma}_b^\lambda(x) = \mathrm{E}[\int_0^{\delta_b^\lambda} & \mathbb{1}\{X^\lambda(s) \in [0, x], \beta^\lambda(s), \kappa^\lambda(s)\} \,\mathrm{d}s \\ & |X^\lambda(0) = b, \beta^\lambda(0) = 2, \kappa^\lambda(0)], \end{aligned}$$

indexed by $\mathcal{S}_- \times \mathcal{S}$. The matrix $\Gamma_0^\lambda(x)$ records the expected sojourn time of $\{X^\lambda(t), \beta^\lambda(t), \kappa^\lambda(t)\}$ in the interval $[0, x]$ during an interval of first return to 0, starting from 0 in a phase of $\mathcal{S}_+$, while $\widehat{\Gamma}_b^\lambda(x)$ corresponds to a first return to b, starting from level b in a phase of $\mathcal{S}_-$.

We can show that

$$\begin{aligned}\Gamma_0^\lambda(x) &= \int_0^x e^{K^\lambda u}\,\mathrm{d}u \left[(C_+^\lambda)^{-1} \quad \Psi^\lambda |C_-^\lambda|^{-1}\right] \\ &= \mathcal{F}(K^\lambda; x)\left[(C_+^\lambda)^{-1} \quad \Psi^\lambda |C_-^\lambda|^{-1}\right],\end{aligned} \tag{35}$$

where

$$K^\lambda = (C_+^\lambda)^{-1} T_{++}^\lambda + \Psi^\lambda |C_-^\lambda|^{-1} T_{-+}^\lambda \tag{36}$$

and Ψ^λ, indexed by $\mathcal{S}_+ \times \mathcal{S}_-$, is the matrix of first return probability from level 0 back to level 0; it is the minimal nonnegative solution of the Riccati equation

$$(C_+^\lambda)^{-1} T_{+-}^\lambda + (C_+^\lambda)^{-1} T_{++}^\lambda \Psi^\lambda + \Psi^\lambda |C_-^\lambda|^{-1} T_{--}^\lambda + \Psi^\lambda |C_-^\lambda|^{-1} T_{-+}^\lambda \Psi^\lambda = 0. \tag{37}$$

For details, we refer to Rogers [21] and Latouche and Nguyen [17]. We give in Appendix A a technical demonstration of (35); a simple justification is that $(e^{K^\lambda u})_{s,s'}$ is, for any states s and s' in $\mathcal{S}_+$, the expected number of crossings of level u in phase s' under taboo of the level 0, given that the process $\{X^\lambda(t)\}$ starts in level 0 and phase s (Ramaswami [20]).

If $\boldsymbol{\alpha}\boldsymbol{\mu} < 0$, then all eigenvalues of K^λ are in $\mathbb{C}_{<0}$, the set of complex numbers with strictly negative real part, and K^λ is nonsingular. If $\boldsymbol{\alpha}\boldsymbol{\mu} \geq 0$, then one eigenvalue is equal to 0, the others are in $C_{<0}$, and K^λ does not have an inverse. Thus, the integral in (35) takes different algebraic forms according to the case.

Lemma 5.4 *If all the eigenvalues of the matrix A are in $\mathbb{C}_{<0}$, then*

$$\mathcal{F}(A; x) = (-A)^{-1}(I - e^{Ax}). \tag{38}$$

If A has all its eigenvalues in $\mathbb{C}_{<0}$, with the exception of one eigenvalue equal to 0, then

$$\mathcal{F}(A; x) = (-A^{\#})(I - e^{Ax}) + x\boldsymbol{v}_a\boldsymbol{u}_a, \tag{39}$$

where $\boldsymbol{v}_a$ and $\boldsymbol{u}_a$ are the right- and left-eigenvectors of A associated to the eigenvalue 0, and $A^{\#}$ is the group inverse of A.

Proof The proof is by verification that both sides of (38) and of (39) are equal for $x = 0$ and have the same derivative with respect to x. $\square$

The matrix $\widehat{\Gamma}_b^\lambda(x)$ is given by

$$\widehat{\Gamma}_b^\lambda(x) = \int_{b-x}^{b} e^{\widehat{K}^\lambda u}\,\mathrm{d}u \begin{bmatrix} |C_-^\lambda|^{-1} & \widehat{\Psi}^\lambda (C_+^\lambda)^{-1} \end{bmatrix}$$
$$= e^{\widehat{K}^\lambda (b-x)} \mathcal{F}(\widehat{K}^\lambda; x) \begin{bmatrix} |C_-^\lambda|^{-1} & \widehat{\Psi}^\lambda (C_+^\lambda)^{-1} \end{bmatrix}, \tag{40}$$

where

$$\widehat{K}^\lambda = |C_-^\lambda|^{-1} T_{--}^\lambda + \widehat{\Psi}^\lambda (C_+^\lambda)^{-1} T_{+-}^\lambda \tag{41}$$

and $\widehat{\Psi}^\lambda$, indexed by $\mathcal{S}_- \times \mathcal{S}$, is the matrix of first return probability from level b back to level b. It is the minimal nonnegative solution of the equation

$$|C_-^\lambda|^{-1} T_{-+}^\lambda + |C_-^\lambda|^{-1} T_{--}^\lambda \widehat{\Psi}^\lambda + \widehat{\Psi}^\lambda (C_+^\lambda)^{-1} T_{++}^\lambda + \widehat{\Psi}^\lambda (C_+^\lambda)^{-1} T_{+-}^\lambda \widehat{\Psi}^\lambda = 0. \tag{42}$$

To prove (40), we define the level-reversed process $\{\widehat{X}^\lambda(t), \beta^\lambda(t), \kappa^\lambda(t)\}$ with fluid rate vector $\widehat{\boldsymbol{\mu}} = -\boldsymbol{\mu}$, and we observe that the time spent in $[0, x]$ by the process $\{X^\lambda(t), \beta^\lambda(t), \kappa^\lambda(t)\}$ during an interval of first return to b has the same distribution as the time spent in $[b-x, b]$ by the level-reversed process during an interval of first return to 0.

We note for future reference that the eigenvalues of $\widehat{H}^\lambda$ are in $\mathbb{C}_{<0}$ if $\boldsymbol{\alpha\mu} > 0$, otherwise the matrix has one eigenvalue equal to 0, with the others in $\mathbb{C}_{<0}$.

Step C. In the third step, we establish a relation between $(N_0^\lambda(x), N_b^\lambda(x))$ and $(\Gamma_0^\lambda(x), \widehat{\Gamma}_b^\lambda(x))$, which leads us to an expression for the matrices $M_{0;+}^\lambda(x)$ and $M_{b;-}^\lambda(x)$ as functions of K^λ, $\widehat{K}^\lambda$, Ψ^λ, and $\widehat{\Psi}^\lambda$.

Lemma 5.5 *The matrix $N^\lambda(x)$ of mean sojourn times in $(0, x)$ during the interval $(0, \delta_0^\lambda \wedge \delta_b^\lambda)$ is a solution of the system*

$$\begin{bmatrix} I & e^{K^\lambda b}\Psi^\lambda \\ e^{\widehat{K}^\lambda b}\widehat{\Psi}^\lambda & I \end{bmatrix} \begin{bmatrix} N_0^\lambda(x) \\ N_b^\lambda(x) \end{bmatrix} = \begin{bmatrix} \Gamma_0^\lambda(x) \\ \widehat{\Gamma}_b^\lambda(x) \end{bmatrix}. \tag{43}$$

where K^λ and $\widehat{K}^\lambda$ are given in (36) and (41).

Proof The proof is similar to that of [9, Lemma 4.1] and we give below its general outline only. First, observe that

$$\Gamma_0^\lambda(x) = N_0^\lambda(x) + \Lambda_b \Gamma^*(x),$$

where $\Gamma^*(x)$ is the matrix of mean sojourn time in the interval $(0, x)$ before the unregulated process $X^\lambda(t)$ first returns to level 0, starting from level b in a phase of $\mathcal{S}_+$. Also,

$$\Gamma^*(x) = \Psi^\lambda(N_b^\lambda(x) + \widehat{\Psi}_b\Gamma^*(x)) = (I - \Psi^\lambda\widehat{\Psi}_b)^{-1}\Psi^\lambda N_b^\lambda(x),$$

and thus

$$\Gamma_0^\lambda(x) = N_0^\lambda(x) + \Lambda_b(I - \Psi^\lambda\widehat{\Psi}_b)^{-1}\Psi^\lambda N_b^\lambda(x).$$

Now, we recognise that

$$\Lambda_b(I - \Psi^\lambda\widehat{\Psi}_b)^{-1} = \Lambda_b(I + \Psi^\lambda\widehat{\Psi}_b + (\Psi^\lambda\widehat{\Psi}_b)^2 + (\Psi^\lambda\widehat{\Psi}_b)^3 + \cdots)$$

is the matrix of expected number of visits to level b in a phase of $\mathcal{S}_+$, starting from 0, before the first return to level 0, and is thus equal to e^{Kb} (Ramaswami [20]). This gives the first equation in (43); the second is similarly proved. $\square$

Remark 5.6 By [9, Lemma 4.2], the coefficient matrix in (43) is nonsingular if $\boldsymbol{\alpha\mu} \neq 0$, and so Lemma 5.5 completely characterizes $N^\lambda(x)$ for fluid processes with non-zero mean drift.

Proof of Theorem 5.1 By [16, Lemmas 5.1 and 5.2], we have

$$\begin{aligned} \Psi^\lambda &= I + O(1/\sqrt{\lambda}), & \widehat{\Psi}^\lambda &= I + O(1/\sqrt{\lambda}), \\ K^\lambda &= K + O(1/\sqrt{\lambda}), & \widehat{K}^\lambda &= \widehat{K} + O(1/\sqrt{\lambda}) \end{aligned}$$

and so, by (35, 40),

$$\begin{bmatrix} \Gamma_0^\lambda(x) \\ \widehat{\Gamma}_b^\lambda(x) \end{bmatrix} = \frac{1}{\sqrt{\lambda}} \begin{bmatrix} \mathcal{F}(K; x) & \\ & e^{\widehat{K}(b-x)}\mathcal{F}(\widehat{K}; x) \end{bmatrix} \begin{bmatrix} \Delta_\sigma^{-1} & \\ & \Delta_\sigma^{-1} \end{bmatrix} \begin{bmatrix} I & I \\ I & I \end{bmatrix} + O(\frac{1}{\lambda})$$

since both $C_+^\lambda = |C_-^\lambda| = \sqrt{\lambda}\Delta_\sigma + O(1)$. Therefore, (43) becomes

$$\begin{aligned} &\begin{bmatrix} I & e^{Kb} \\ e^{\widehat{K}b} & I \end{bmatrix} \begin{bmatrix} N_0^\lambda(x) \\ N_b^\lambda(x) \end{bmatrix} \\ &\quad = \frac{1}{\sqrt{\lambda}} \begin{bmatrix} \mathcal{F}(K; x)\Delta_\sigma^{-1} & \\ & e^{\widehat{K}(b-x)}\mathcal{F}(\widehat{K}; x)\Delta_\sigma^{-1} \end{bmatrix} \begin{bmatrix} I & I \\ I & I \end{bmatrix} + O(\frac{1}{\lambda}). \end{aligned} \tag{44}$$

By [17, Lemma 3.6],

- if $\boldsymbol{\alpha\mu} < 0$, then K has all eigenvalues in $\mathbb{C}_{<0}$ and $\widehat{K}$ has $m-1$ eigenvalues in $\mathbb{C}_{<0}$ and one eigenvalue equal to 0,
- if $\boldsymbol{\alpha\mu} > 0$, then $\widehat{K}$ has all eigenvalues in $\mathbb{C}_{<0}$ and K has $m-1$ eigenvalues in $\mathbb{C}_{<0}$ and one eigenvalue equal to 0.

This entails that one of the two matrices e^{Kb} and $e^{\widehat{K}b}$ has spectral radius equal to one, with the other having spectral radius strictly less than one, so that the left-most matrix in (44) is nonsingular when $\boldsymbol{\alpha\mu} \neq 0$, and

$$\begin{bmatrix} N_0^\lambda(x) \\ N_b^\lambda(x) \end{bmatrix} = \frac{1}{\sqrt{\lambda}} \begin{bmatrix} I & e^{Kb} \\ e^{\widehat{K}b} & I \end{bmatrix}^{-1} \begin{bmatrix} \mathcal{F}(K;x)\Delta_\sigma^{-1} & \\ & e^{\widehat{K}(b-x)}\mathcal{F}(\widehat{K};x)\Delta_\sigma^{-1} \end{bmatrix} \begin{bmatrix} I & I \\ I & I \end{bmatrix} + O(\frac{1}{\lambda}). \tag{45}$$

On the other hand, we have $(I - \Psi_b^\lambda)^{-1} = \sqrt{\lambda}(-P_b)^{-1} + O(1)$ by [16, Lemma 5.5] and $(T_{--}^\lambda)^{-1} = O(1/\lambda)$, $(-T_{--}^\lambda)^{-1}T_{-+}^\lambda = I + O(1/\lambda)$, by definition of T^λ. Thus, (31) may be written as

$$M_{0;+}^\lambda(x) = (\sqrt{\lambda}(-P_b)^{-1} + O(1))N_0^\lambda(x), \tag{46}$$

and similarly,

$$M_{b;-}^\lambda(x) = (\sqrt{\lambda}(-\widehat{P}_b)^{-1} + O(1))N_b^\lambda(x). \tag{47}$$

Equation (28) directly follows from (27, 45, 46, 47). □

It will be useful in Section 6 to have separate expressions for $\mathcal{M}_0(x)$ and $\mathcal{M}_b(x)$. Using

$$\begin{bmatrix} I & e^{Kb} \\ e^{\widehat{K}b} & I \end{bmatrix}^{-1} = \begin{bmatrix} (I - e^{Kb}e^{\widehat{K}b})^{-1} & \\ & (I - e^{\widehat{K}b}e^{Kb})^{-1} \end{bmatrix} \begin{bmatrix} I & -e^{Kb} \\ -e^{\widehat{K}b} & I \end{bmatrix},$$

we easily replace (28) by the pair of equations

$$\mathcal{M}_0(x) = 2(-P_b)^{-1}(I - e^{Kb}e^{\widehat{K}b})^{-1}(\mathcal{F}(K;x) - e^{Kb}e^{\widehat{K}(b-x)}\mathcal{F}(\widehat{K};x))\Delta_\sigma^{-1}, \tag{48}$$

$$\mathcal{M}_b(x) = 2(-\widehat{P}_b)^{-1}(I - e^{\widehat{K}b}e^{Kb})^{-1}e^{\widehat{K}(b-x)}(\mathcal{F}(\widehat{K};x) - e^{\widehat{K}x}\mathcal{F}(K;x))\Delta_\sigma^{-1}. \tag{49}$$

The next corollary is obvious: we merely let $x = b$ in (48, 49).

Corollary 5.7 *If $\boldsymbol{\alpha\mu} \neq 0$, then the expected time spent in the various phases during an excursion from one boundary to the other is given by*

$$\mathcal{M}_0(b) = 2(-P_b)^{-1}(I - e^{Kb}e^{\widehat{K}b})^{-1}(\mathcal{F}(K,b) - e^{Kb}\mathcal{F}(\widehat{K};b))\Delta_\sigma^{-1}$$
$$\mathcal{M}_b(b) = 2(-\widehat{P}_b)^{-1}(I - e^{\widehat{K}b}e^{Kb})^{-1}(\mathcal{F}(\widehat{K},b) - e^{\widehat{K}b}\mathcal{F}(K;b))\Delta_\sigma^{-1}.$$

□

Remark 5.8 In marked contrast to Theorem 4.1 and its Corollary 4.2, Theorem 5.1 does not give an expression for $\mathcal{M}_0(x)$ and $\mathcal{M}_b(x)$ if $\boldsymbol{\alpha\mu} = 0$. In that case, we might write, instead of (28), that $\mathcal{M}(x)$ is to be determined by solving the system

$$\begin{bmatrix} I & e^{Kb} \\ e^{\widehat{K}b} & I \end{bmatrix} \begin{bmatrix} P_b\mathcal{M}_0(x) \\ \widehat{P}_b\mathcal{M}_b(x) \end{bmatrix} = -2 \begin{bmatrix} \mathcal{F}(K;x) \\ e^{\widehat{K}(b-x)}\mathcal{F}(\widehat{K};x) \end{bmatrix} \Delta_\sigma^{-1},$$

plus some additional equation. Unfortunately, this additional equation has eluded us so far.

6 Stationary distribution of a flexible MMBM

We have now obtained all the ingredients necessary to express the stationary distribution of the flexible MMBM $\{Y(t), \rho(t)\}$ once we specify its parameters.

It is natural to expect *some* of the parameters at least to take different values during the two legs of a regeneration cycle, from level 0 to level b and back. We assume that the set $\mathcal{E}$ of phases is made up of two subsets, $\mathcal{E}_u$ and $\mathcal{E}_d$, and that the generator of $\{\rho(t)\}$ is Q partitioned as follows:

$$Q = \begin{bmatrix} Q_u & 0 \\ 0 & Q_d \end{bmatrix}. \tag{50}$$

The idea is that Q_u, on the state space $\mathcal{E}_u$, describes the evolution of the Markov environment during the up-leg, after a regeneration at level 0 until the next regeneration at level b; Q_d on the state space $\mathcal{E}_d$ controls the system during a down-leg, from b to 0.

The other parameters are similarly partitioned and we write $\boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_u & \boldsymbol{\mu}_d \end{bmatrix}$ and $\boldsymbol{\sigma} = \begin{bmatrix} \boldsymbol{\sigma}_u & \boldsymbol{\sigma}_d \end{bmatrix}$. The matrices P° and $P^\bullet$ control the transition from $\mathcal{E}_d$ to $\mathcal{E}_u$ upon hitting level 0 at the end of a down-leg, and from $\mathcal{E}_u$ to $\mathcal{E}_d$ upon hitting b, and we write them as

$$P^\circ = \begin{bmatrix} I & 0 \\ P^\circ_{du} & 0 \end{bmatrix} \qquad \text{and} \qquad P^\bullet = \begin{bmatrix} 0 & P^\bullet_{ud} \\ 0 & I \end{bmatrix}.$$

The identity blocks on the diagonal do not play any role in the calculation to follow, their role is to ensure that P° and $P^\bullet$ are stochastic matrices. We assume that Q_u and Q_d are irreducible, and so Assumption 2.1 is satisfied.

Upon hitting 0 at a regeneration point and after choosing a new phase with the matrix P°, the phase ρ is in $\mathcal{E}_u$. Therefore, the vector $\boldsymbol{\nu}_0$ takes the form $\boldsymbol{\nu}_0 = \begin{bmatrix} \boldsymbol{\nu}_{0;u} & \mathbf{0} \end{bmatrix}$ and, for similar reasons, we have $\boldsymbol{\nu}_b = \begin{bmatrix} \mathbf{0} & \boldsymbol{\nu}_{b;u} \end{bmatrix}$.

As transitions from $\mathcal{E}_u$ to $\mathcal{E}_d$ or from $\mathcal{E}_d$ to $\mathcal{E}_u$ are possible at regeneration points only, the matrices of first passage probabilities from one level to the other have the structure

$$H_0 = \begin{bmatrix} H_{0;u} & 0 \\ 0 & H_{0;d} \end{bmatrix} \qquad \text{and} \qquad H_b = \begin{bmatrix} H_{b;u} & 0 \\ 0 & H_{b;d} \end{bmatrix}$$

The only blocks that one needs to evaluate, however, are $H_{0;u}$ and $H_{b;d}$: the value of $H_{0;d}$ is irrelevant as the process cannot leave level 0 in a phase of $\mathcal{E}_d$ and $H_{b;u}$ is irrelevant as well, for a similar reason.

Obviously, the matrices $M_0(x)$ and $M_b(x)$ have the same structure

$$M_0(x) = \begin{bmatrix} M_{0;u}(x) & 0 \\ 0 & M_{0;d}(x) \end{bmatrix} \qquad \text{and} \qquad M_b(x) = \begin{bmatrix} M_{b;u}(x) & 0 \\ 0 & M_{b;d}(x) \end{bmatrix},$$

and we do not need to evaluate $M_{0;d}(x)$ or $M_{b;d}(x)$.

Altogether, we may re-formulate Theorem 2.2 in a more detailed manner as follows, using Theorem 4.1, Corollary 4.2, and equations (48, 49).

Theorem 6.1 *The stationary distribution* $\Pi(x)$ *of the flexible Markov modulated Brownian motion is given by* $\Pi(x) = (\boldsymbol{\nu}^* \boldsymbol{m}^*)^{-1} \boldsymbol{\nu}^* M^*(x)$*. The vector* $\boldsymbol{\nu}^*$ *is partitioned as* $\boldsymbol{\nu}^* = \begin{bmatrix} \boldsymbol{\nu}_{0;u} & \boldsymbol{\nu}_{b;d} \end{bmatrix}$ *where*

$$\boldsymbol{\nu}_{0;u} = \boldsymbol{\nu}_{0;u} H_{0;u} P^{\bullet}_{ud} H_{b;d} P^{\circ}_{du}, \qquad \boldsymbol{\nu}_{b;d} = \boldsymbol{\nu}_{0;u} H_{0;u} P^{\bullet}_{ud},$$

with

$$H_{0;u} = \mathcal{H}_0|_{Q=Q_u, \boldsymbol{\mu}=\boldsymbol{\mu}_u, \boldsymbol{\sigma}=\boldsymbol{\sigma}_u} \qquad \text{and} \qquad H_{b;d} = \mathcal{H}_b|_{Q=Q_d, \boldsymbol{\mu}=\boldsymbol{\mu}_d, \boldsymbol{\sigma}=\boldsymbol{\sigma}_d}.$$

The matrix $M^*(x)$ *is partitioned as*

$$M^*(x) = \begin{bmatrix} M_{0;u}(x) & \\ & M_{b;d}(x) \end{bmatrix},$$

with

$$M_{0;u}(x) = \mathcal{M}_0(x)|_{Q=Q_u, \boldsymbol{\mu}=\boldsymbol{\mu}_u, \boldsymbol{\sigma}=\boldsymbol{\sigma}_u} \text{ and } M_{b;d}(x) = \mathcal{M}_b(x)|_{Q=Q_d, \boldsymbol{\mu}=\boldsymbol{\mu}_d, \boldsymbol{\sigma}=\boldsymbol{\sigma}_d}.$$

The vector $\boldsymbol{m}^*$ *is given by* $\boldsymbol{m}^* = M^*(b)\mathbf{1}$. □

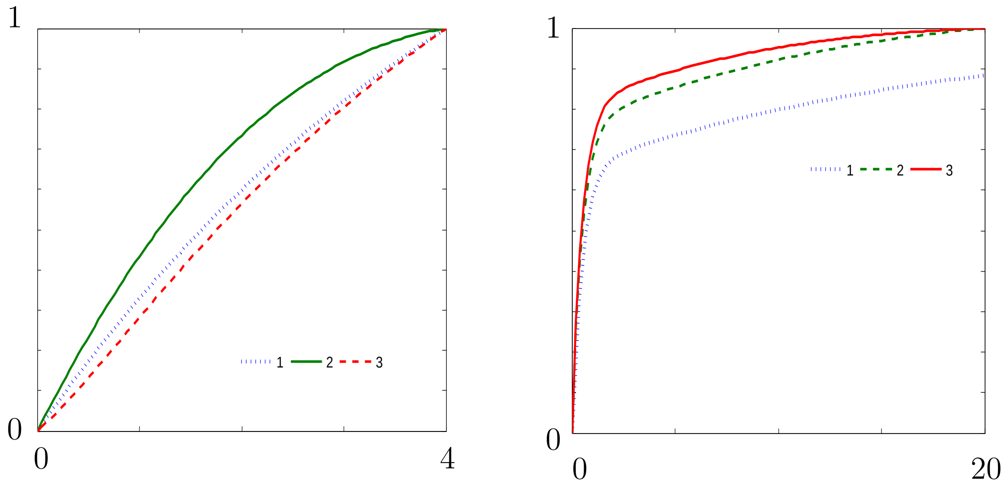

Figure 2: Cumulative stationary distribution functions for the single-phase BM examples (left) and the cyclic-phases examples (right). The parameters are clarified in the text.

7 Illustration

Example 7.1 Single-phase Brownian motion. This is the example given in the introduction: the environmental process has only one phase, possibly characterized by different parameters in alternating intervals between regeneration points. With one phase only, the calculations simplify considerably; if μ is negative, then $U = \widehat{K} = 0$ and $\widehat{U} = K = 2\mu/\sigma^2$. Assuming that both μ_u and μ_d are negative, we obtain from Theorem 6.1 that

$$\Pi(x) = \frac{\mathcal{M}_{0;u}(x) + \mathcal{M}_{b;d}(x)}{\mathcal{M}_{0;u}(b) + \mathcal{M}_{b;d}(b)},$$

with

$$\mathcal{M}_{0;u}(x)+\mathcal{M}_{b;d}(x) = \left(\frac{1}{\mu_d} - \frac{1}{\mu_u}\right) x+\frac{\sigma_d^2}{2\mu_d^2}(1-e^{2\mu_d x/\sigma_d^2})-\frac{\sigma_u^2}{2\mu_u^2}e^{-2\mu_u b/\sigma_u^2}(1-e^{2\mu_u x/\sigma_u^2}).$$

If the parameters in the two types of intervals are equal, then further simplifications yield the well-known truncated exponential distribution

$$\Pi(x) = (1 - e^{2\mu x/\sigma^2})(1 - e^{2\mu b/\sigma^2})^{-1}.$$

The parameters for the three distributions shown on the left of Figure 2 are given in the table below.

Case	line	μ_u	σ_u^2	μ_d	σ_d^2
1	dotted	-1	10	-1	10
2	plain	-1	10	-10	10
3	dashed	-1	10	-1	1

In Case 1, with a single set of parameters, the process is a regulated Brownian motion with two boundaries, in Case 2 the drift is decreased to -10 when the process reaches the upper boundary, and in Case 3 the drift remains the same, but the variance is reduced.

As expected, the buffer content is stochastically smaller in Case 2 than in Case 1: for instance, the 90th percentiles are 2.88 and 3.44 respectively. We had expected that the buffer content would also be smaller in Case 3, our argument being that, with a smaller variance, the negative drift would be better felt, and the buffer content would go down faster. As one sees on Figure 2 this is not the case and the buffer content is slightly larger in Case 3 (the 90th percentile is 3.48).

We give in the table below the expected total duration of excursions from level 0 to level b and back from level b to level 0, and also the proportion of time spent by the process in the regenerative intervals from 0 to b; this quantity is $\Pi_u(b) = (\boldsymbol{\nu}^* \boldsymbol{m}^*)^{-1} \boldsymbol{\nu}_{0;u} M_{0;u}(b) \mathbf{1}$.

Case	$\mathcal{M}_{0;u}(b)$	$\mathcal{M}_{b;d}(b)$	$\Pi_u(b)$
1	2.13	1.25	0.63
2	2.13	0.35	0.86
3	2.13	3.50	0.38

Obviously, the time to move from 0 to b is the same in all cases, and we do observe for Case 2 the effect resulting from switching from $\mu_u = -1$ to $\mu_d = -10$. In Case 3, switching from $\sigma_u^2 = 10$ to $\sigma_d^2 = 1$ increases the expected length of an excursion from b to 0 nearly by a factor 3.

Example 7.2 Cyclic environmental process. In this example, $m = 8$ and the process of phases evolves cyclically from 1 to 8 and back to 1. We take $Q_u = Q_d = \Omega$ with $\Omega_{i,i+1} = \lambda$, for $i = 1, \ldots, 7$, $\Omega_{8,1} = \lambda$, $\Omega_{ii} = -\lambda$ for all i, the other elements are equal to 0. In the three cases to follow, we have $\lambda = 0.1$, so the process moves from one phase to the next in 10 units of time on average. The other parameters are $\mu_i = -1$ for all i, and $\sigma_i = 1$ for all $i \neq 8$ and $\sigma_8 = 10$. Thus, the process is quite regular most of the time but every 80 units of time, on average, the volatility becomes very high during 10 units of time.

In Case 1 (dotted line on the right-hand side graph of Figure 2), the buffer is infinite. One observes the effect of the irregularity, infrequent but very high, of the input process: the stationary expected buffer occupancy is $\mathrm{E}[X_\infty] = 6.69$, but the distribution has a very long tail, with $\mathrm{P}[X_\infty > 20] = 0.12$. Actually, this tail decreases at a rate equal to the maximal eigenvalue of K, equal to -0.14 in the present example.

	1	2	3	4	5	6	7	8
$\mathcal{M}_{0;u}$	109.20	99.20	89.20	79.20	69.20	59.20	49.20	39.24
$\mathcal{M}_{b;d}$ Case 2	19.49	19.45	19.34	19.03	18.22	16.36	12.66	6.34
$\mathcal{M}_{b;d}$ Case 3	2.00	2.00	2.00	2.00	2.00	1.99	1.97	1.53

Table 1: Cyclic environment, moments of first passage time from one boundary to the other.

In Case 2 (dashed line), the buffer is finite, with $b = 20$ and the other parameters are the same as in Case 1. In Case 3 (plain line), $\boldsymbol{\mu}_u$ and $\boldsymbol{\sigma}_u = \boldsymbol{\sigma}_d$ are the same as in Case 2, and $\boldsymbol{\mu}_d = 10\boldsymbol{\mu}_u$. One clearly see that the buffer content is smallest in Case 3, the 90th percentile, for instance, is 5.40, compared to 8.20 in Case 2.

It is interesting to examine in more details the behavior of the two processes. The transition matrix $H_{0;u}$ is the same in both cases and the probability mass is almost exclusively concentrated on the 8th column: the computed values are $(H_{0;u})_{i,8} = 0.9995$ and $(H_{0;u})_{i,1} = 0.0005$, independently of i, the remaining elements of the matrix being negligible.

We give in Table 1 the expected duration of transitions from one boundary to the other. It appears clearly that to reach level b, starting from a phase $i \neq 8$, the process must first move to phase 8, with an expected time equal to $10(8-i)$, and only then get a chance to reach b in a reasonable amount of time. Furthermore, $(\mathcal{M}_{0;u})_8$ is much greater than the expected sojourn time in phase 8. We interpret this as follows: starting from level 0 in phase 8, there is a significant probability that the process reaches level b before switching to phase 1, but it is also possible the system will have to go through one cycle (or more) before eventually reaching level b.

For Case 2, the effect of σ_8 is also noticeable in the expected duration of excursions from b to 0, albeit to a lesser degree; for Case 3, these expected durations are dominated by the large absolute value of $\boldsymbol{\mu}_d$.

Example 7.3 Video streaming application. This example is taken from Gribaudo *et al.* [13]. We use it to illustrate changes in the system characteristics resulting from global changes of the parameters.

There are five states; the video streaming application cycles between *buffering* (States 1 and 3), *playing* (States 2 and 4), and *finishing* (State 5), leaving each state at the rates β_B, β_P and β_F, respectively. The videos are being played in a loop: when a video is finished, the application starts another one. Video streaming packets arrive at rate λ_L with variance γ_L when the network is congested (States 1 and 3), at rate λ_H with variance γ_H

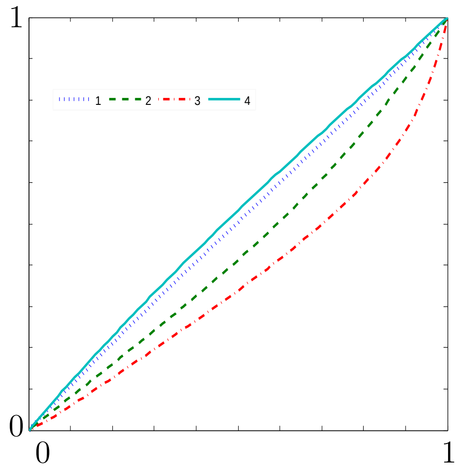

Figure 3: Cumulative stationary distribution functions for Video Streaming Application

otherwise (States 2 and 4). The packets are decoded at rate δ with variance γ.

The transition matrix for the phase process is

$$Q = \begin{bmatrix} * & \beta_B & \alpha_{LH} & 0 & 0 \\ 0 & * & 0 & \alpha_{LH} & \beta_P \\ \alpha_{HL} & 0 & * & \beta_B & 0 \\ 0 & \alpha_{HL} & 0 & * & \beta_P \\ \beta_F p_1 & 0 & \beta_F p_3 & 0 & * \end{bmatrix}$$

where the diagonal elements are such that $Q\mathbf{1} = \mathbf{0}$, and the rates and variance vectors are

$$\boldsymbol{\mu} = \begin{bmatrix} \lambda_L & \lambda_L - \delta & \lambda_H & \lambda_H - \delta & -\delta \end{bmatrix},$$
$$\boldsymbol{\sigma}^2 = \begin{bmatrix} \gamma_L & \gamma_L + \gamma & \gamma_H & \gamma_H + \gamma & \gamma \end{bmatrix}.$$

The distribution shown as a dotted line on Figure 3 is that of the regulated MMBM, with parameters chosen from [13]: the buffer size is 1MB and is the unit of volume, the time unit is 1 second, the parameters values are given in the table below.

$\delta = \gamma = 0.5$	$\lambda_L = \gamma_L = 0.25$	$\lambda_H = \gamma_H = 0.625$
$\alpha_{LH} = \alpha_{HL} = 1/60$	$\beta_B = \beta_F = 0.1$	$\beta_P = 0.03$
$p_1 = \alpha_{HL}/(\alpha_{HL} + \alpha_{LH})$	$p_3 = 1 - p_1$	

The other curves are defined as follows:

Case 2, dashed line — the parameters are the same but the phase transition matrices at regeneration epochs are $P^\circ = P^\bullet = (1/m)\mathbf{1}\,\mathbf{1}^\top$; thus, the phase is sampled at random at the end of each excursion from one boundary to the other.

Case 3, mixed dashed and dotted line — the variances are reduced during excursions from b to 0, with $\boldsymbol{\sigma}_d = 0.1\boldsymbol{\sigma}_u$; all other parameters remain the same and $P^\circ = P^\bullet = I$.

Case 4, plain line — the system cycles 10 times faster through its three stages during excursions from b to 0, with $(\beta_P)_d = (\beta_F)_d = 1$, $(\beta_P)_d = 0.3$; all other parameters remain the same and $P^\circ = P^\bullet = I$.

In addition to the distribution functions in Figure 3, we give below the median of the four distributions, and also the stationary probability $\Pi_u(b)$. We observe that the buffer is more heavily utilized when the variances are reduced (Case 3), that is, it spends half of the time being above level 0.7. We also observe that re-sampling the phases at regeneration epochs (Case 2) has a significant effect.

Case	1	2	3	4
Median	0.50	0.59	0.70	0.46
$\Pi_u(b)$	0.51	0.40	0.19	0.54

8 Comparison with existing literature

A related question is addressed in Breuer [4], where the author analyzes the joint distribution of the random variables

$$\zeta_1(b; x, j) = \int_0^{\delta_0 \wedge \delta_b} \mathbb{1}\{X(s) < x, \kappa(s) = j\}\, \mathrm{d}s$$

and

$$\zeta_2(b; x, j) = \int_0^{\delta_0 \wedge \delta_b} \mathbb{1}\{X(s) > x, \kappa(s) = j\}\, \mathrm{d}s$$

for $j \in \mathcal{E}$ and $0 < x < b$. These are the time spent in $(0, x) \times \{j\}$ and $(x, b) \times \{j\}$ respectively, before the first exit from the interval $(0, b)$.

We need to introduce some more notation. Consider a vector $\boldsymbol{r} \geq \boldsymbol{0}$ indexed by $\mathcal{E}$. Define the matrix $U(\boldsymbol{r})$ as the minimal solution of the matrix equation

$$\Delta_\sigma^2 X^2 + 2\Delta_\mu X + 2(Q - \Delta_r) = 0. \tag{51}$$

For $\boldsymbol{r} = \boldsymbol{0}$, we have $U(\boldsymbol{0}) = U$, the generator introduced at the beginning of Section 3. Similarly, $-\widehat{U}(\boldsymbol{r})$ is the maximal solution of (51).

We further define the random variables

$$T_1(b;x) = \sum_{j \in \mathcal{E}} (\boldsymbol{r}_1)_j \zeta_1(b;x,j), \qquad T_2(b;x) = \sum_{j \in \mathcal{E}} (\boldsymbol{r}_2)_j \zeta_2(b;x,j),$$

where $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ are two nonnegative vectors. The functions

$$E^+_{\boldsymbol{r}_1,\boldsymbol{r}_2}(b;x|a) = \mathrm{E}[e^{-T_1(b;x)-T_2(b;x)} \mathbb{1}\{\delta_b < \delta_0\}, \kappa(\delta_b) | X(0) = a, \kappa(0)]$$

and

$$E^-_{\boldsymbol{r}_1,\boldsymbol{r}_2}(b;x|a) = \mathrm{E}[e^{-T_1(b;x)-T_2(b;x)} \mathbb{1}\{\delta_0 < \delta_b\}, \kappa(\delta_0) | X(0) = a, \kappa(0)],$$

for $0 < a < b$, are the joint Laplace transforms of the ζ_1s and ζ_2s restricted on the exit occurring at the upper or lower boundary, respectively, conditionally given that the process starts from level a in the open interval $(0,b)$.

From [4, Theorem 1, Lemmas 1 and 2], we find after various adaptations to our specific case and some simple manipulations that

$$E^+_{\boldsymbol{r}_1,\boldsymbol{r}_2}(b;x|x) = -2(\widehat{P}_x(\boldsymbol{r}_1) + P_{b-x}(\boldsymbol{r}_2))^{-1} L_{b-x}(\boldsymbol{r}_2) \tag{52}$$

and, by symmetry, that

$$E^-_{\boldsymbol{r}_1,\boldsymbol{r}_2}(b;x|x) = -2(\widehat{P}_x(\boldsymbol{r}_1) + P_{b-x}(\boldsymbol{r}_2))^{-1} \widehat{L}_x(\boldsymbol{r}_1). \tag{53}$$

Here, $\widehat{P}_x(\boldsymbol{r}_1)$ and $\widehat{L}_x(\boldsymbol{r}_1)$ are given by (24) with U, $\widehat{U}$ and b respectively replaced by $U(\boldsymbol{r}_1)$, $\widehat{U}(\boldsymbol{r}_1)$ and x, and $P_{b-x}(\boldsymbol{r}_2)$ and $L_{b-x}(\boldsymbol{r}_2)$ are given by (18) with U, $\widehat{U}$ and b replaced by $U(\boldsymbol{r}_2)$, $\widehat{U}(\boldsymbol{r}_2)$ and $b-x$.

Finally, we define the random variables

$$\xi(y;x,j) = \int_0^{\delta_y} \mathbb{1}\{Z(s) \le x, \kappa(s) = j\}\, \mathrm{d}s,$$

for $y \ge x$, as the time spent in $[0,x] \times \{j\}$ by the *regulated* process until the first passage to level y, and we denote their joint Laplace transform, starting from level 0, by

$$\Xi_{\boldsymbol{r}}(y;x) = \mathrm{E}[e^{-\boldsymbol{r}\boldsymbol{\xi}(y;x)} \kappa(\delta_y) | Z(0 = 0, \kappa(0)].$$

Lemma 8.1 *The joint Laplace transform $\Xi_{\boldsymbol{r}}(y;x)$ of the random variables $\xi(b;x,j)$, $1 \le j \le m$, is given by*

$$\Xi_{\boldsymbol{r}}(b;x) = (I - \Xi_{\boldsymbol{r}}(x;x) E^-_{\boldsymbol{r},\boldsymbol{0}}(b;x|x))^{-1} \Xi_{\boldsymbol{r}}(x;x) E^+_{\boldsymbol{r},\boldsymbol{0}}(b;x|x) \tag{54}$$

for $b > x$, with $\Xi_{\boldsymbol{r}}(x;x) = (-P_x(\boldsymbol{r}))^{-1} L_x(\boldsymbol{r})$.

Proof We decompose the interval $[0, \delta_y]$ into three subintervals:

$$[0, \delta_y] = [0, \delta_x] \cup [\delta_x, \delta^*] \cup [\delta^*, \delta_b],$$

where $\delta^* = \inf\{t > \delta_x : Z(t) = 0 \text{ or } Z(t) = b\}$. The sojourn times in $(0, x) \times \{j\}$ during these intervals are conditionally independent, given the phases, and so we have

$$\Xi_{\boldsymbol{r}}(b; x) = \Xi_{\boldsymbol{r}}(x; x)(E^{+}_{\boldsymbol{r},\boldsymbol{0}}(b; x|x) + E^{-}_{\boldsymbol{r},\boldsymbol{0}}(b; x|x)\Xi_{\boldsymbol{r}}(b; x)),$$

from which (54) follows.

The given expression for $\Xi_{\boldsymbol{r}}(x; x)$ is a consequence of [3, Theorem 1]: we adapt it to our specific case, taking into account the fact that Theorem 1 in [3] is stated for the level-reversed process, and performing some simple manipulations. □

At first, it looks like we might obtain $\mathcal{M}_0(x)$ by differentiating both sides of (54) with respect to $\boldsymbol{r}$ and by evaluating the result at $\boldsymbol{r} = \boldsymbol{0}$. We would need the derivatives with respect to $\boldsymbol{r}$ of the solutions of (51). Details would still need to be worked out and in final analysis, the expressions so obtained would without doubt be much more involved than the very clean expressions given in (28).

9 Conclusion and extensions

In this paper, we have illustrated one useful reason for approximating Markov-modulated Brownian motions with stochastic fluid queues. In particular, the approximation allows for the analysis of MMBMs subject to boundary conditions that are not the traditional regulation. This approach, coupled with the regenerative method, may be adapted easily to other types of feedbacks, such as a combination of absorption, stickiness, and instantaneous change of phase whenever the process hits a boundary.

With the technique developed here, we might analyze systems for which the so-called feedback only lasts for a finite amount of time. For instance, rates change for an exponential amount of time, and then the system resumes its normal mode of operations. The results from Sections 4 to 6 have to be adapted, as the generator for the phase process between two regeneration points is no longer irreducible.

A Expected time under a taboo

Although the statement of Theorem A.1 would seem to be well-known, we do not know of a published formal proof, which is why we include it here.

Consider a fluid queue with generator T for the environmental Markov process and fluid growth rates $\boldsymbol{c}$. We denote by Ψ its matrix of first return probabilities to level 0, and we define $K = C_+^{-1}T_{++} + \Psi|C_-|^{-1}T_{+-}$, where $C = \Delta_c$.

Theorem A.1 *The matrix $\Gamma(x)$ of mean sojourn time in $[0, x]$ before return to the initial level 0, starting from a phase with positive growth, is given by*

$$\Gamma(x) = \mathcal{F}(K; x) \left[C_+^{-1} \quad \Psi|C_-|^{-1} \right], \tag{55}$$

where $\mathcal{F}(K; x) = \int_0^x e^{Ku} \, \mathrm{d}u$.

Proof

We need to consider separately the case when the stationary drift is strictly negative from the case when it is positive or equal to zero.

A. Strictly negative drift In this case, the fluid queue is positive recurrent and the eigenvalues of K are all in $\mathbb{C}_{<0}$.

We define the complementary probability functions

$$[G(x,t)]_{ij} = \mathrm{P}[t < \tau, X(t) > x, \kappa(t) = j \,|\, X(0) = 0, \kappa(0) = i],$$

where τ is the first return time to level 0. Denote by $\overline{\Gamma}(x)$ the mean sojourn time in (x, ∞) before returning to the initial level 0:

$$\overline{\Gamma}(x) = \int_0^\infty G(x, u) \, \mathrm{d}u,$$

obviously, $\Gamma(x) = \overline{\Gamma}(0) - \overline{\Gamma}(x)$. One verifies by the usual argument (Karandikar and Kulkarni [15]) that $G(x, t)$ is the solution to the system of partial differential equations

$$\frac{\partial}{\partial t} G(x,t) + \frac{\partial}{\partial x} G(x,t) C = G(x,t) T \qquad \text{for } x > 0.$$

Integrating both sides with respect to t from 0 to ∞ gives

$$[G(x,t)]_0^\infty + \frac{\partial}{\partial x} \overline{\Gamma}(x) C = \overline{\Gamma}(x) T. \tag{56}$$

Note that

$$\lim_{t \to \infty} G(x,t) = 0, \qquad \lim_{t \to 0} G(x,t) = 0,$$

the first limit is due to the negative drift assumption, which implies that $\tau < \infty$ almost surely, the second holds because the fluid queue does not have enough time to grow beyond x by time t if t is small. Then, it is easy to verify that the solution to (56) is given by

$$\overline{\Gamma}(x) = AK^{-1}e^{Kx}\begin{bmatrix} C_+^{-1} & \Psi|C_-|^{-1} \end{bmatrix} \tag{57}$$

for some matrix A to be determined.

Let us focus on the block $\overline{\Gamma}_{++}(x) = AK^{-1}e^{Kx}C_+^{-1}$. For sufficiently small h, we may write that

$$\overline{\Gamma}_{++}(0) = hC_+^{-1} + \overline{\Gamma}_{++}(h) + o(h). \tag{58}$$

Indeed, the expected time spent above level 0 in a phase j is equal to the time needed to reach level h in that phase, if $\kappa(0) = j$ plus the time spent above h in that phase. The third term in (58) accounts for the time spent in oscillations between 0 and h whenever the fluid drops below h. We get from (58) that

$$\frac{\partial}{\partial x}\overline{\Gamma}_{++}(x)|_{x=0} = -C_+^{-1},$$

from which we conclude that $A = -I$. Thus,

$$\begin{aligned}\Gamma(x) &= (-K)^{-1}(I - e^{Kx})\begin{bmatrix} C_+^{-1} & \Psi|C_-|^{-1} \end{bmatrix} \\ &= \mathcal{F}(K;x)\begin{bmatrix} C_+^{-1} & \Psi|C_-|^{-1} \end{bmatrix} \qquad \text{by Lemma 5.4.}\end{aligned}$$

B. Nonnegative drift In this case, the fluid queue is null-recurrent or transient, and K has one eigenvalue equal to zero.

We may not repeat the argument for Part A because the mean sojourn time in (x, ∞) is infinite, and $\int_0^\infty G(x,u)\,\mathrm{d}u$ diverges for any x.

To get around this problem, we shall kill the process after a random, finite, interval of time and then use a limiting argument. We define ζ to be an exponentially distributed random variable with rate ψ, and

$$\begin{aligned}G'(x,t;\psi) &= \mathrm{P}[t < \min\{\tau,\zeta\}, X(t) > x, \kappa(t) \mid X(0) = 0, \kappa(0)], \\ \Gamma'(x;\psi) &= \int_0^\infty G'(x,u;\psi)\,\mathrm{d}u.\end{aligned}$$

As $\zeta < \infty$ with probability one, $\Gamma'(x;\psi) < \infty$ and we may retrace the steps in Part A. In particular, the matrix $\Gamma'(x;\psi)$ of expected sojourn time in (x,∞) under the taboo of 0 is given by

$$\Gamma'(x;\psi) = (-K_\psi)^{-1}e^{K_\psi x}\begin{bmatrix} C_+^{-1} & \Psi_\psi|C_-|^{-1} \end{bmatrix},$$

where Ψ_ψ is the minimal nonnegative solution of the Riccati equation

$$|C_-^{-1}|T_{-+}^\psi + \Psi_\psi C_+^{-1} T_{++}^\psi + |C_-^{-1}|T_{--}^\psi \Psi_\psi + \Psi_\psi C_+^{-1} T_{+-}^\psi \Psi_\psi = 0,$$

with

$$T^\psi = \begin{bmatrix} T_{++} - \psi I & T_{+-} \\ T_{-+} & T_{--} - \psi I \end{bmatrix},$$

and

$$K_\psi = |C_-^{-1}|T_{--}^\psi + \Psi_\psi C_+^{-1} T_{+-}^\psi.$$

In the limit as $\psi \to 0$, the matrices Ψ_ψ and K_ψ respectively converge to Ψ and K. Then,

$$\begin{aligned} \Gamma(x) &= \lim_{\psi \to 0} (\Gamma'(0;\psi) - \Gamma'(x;\psi)) \\ &= \lim_{\psi \to 0} \left\{ (-K_\psi)^{-1}(I - e^{K_\psi b}) \right\} \begin{bmatrix} C_+^{-1} & \Psi |C_-|^{-1} \end{bmatrix}. \end{aligned}$$

In the last expression, K_ψ converges to K which is singular; thus, we need to exercise some care in evaluating the remaining limit.

The matrix K has one isolated eigenvalue equal to 0, and K_ψ has an isolated, real, maximal eigenvalue ω_ψ which converges to 0 as $\psi \to 0$. Thus, there exist some matrices S and S_ψ such that

$$K = S \begin{bmatrix} J & \\ & 0 \end{bmatrix} S^{-1} \quad \text{and} \quad K_\psi = S_\psi \begin{bmatrix} J_\psi & \\ & \omega_\psi \end{bmatrix} S_\psi^{-1},$$

where J is a matrix with all eigenvalues of K in $\mathbb{C}_{<0}$, J_ψ is a matrix with all eigenvalues of K_ψ with real parts strictly less than ω_ψ, and $S_\psi \to S$ and $J_\psi \to J$ as $\psi \to 0$. We decompose the inverse of K_ψ as

$$K_\psi^{-1} = S_\psi \begin{bmatrix} J_\psi^{-1} & \\ & 0 \end{bmatrix} S_\psi^{-1} + S_\psi \begin{bmatrix} 0 & \\ & \omega_\psi^{-1} \end{bmatrix} S_\psi^{-1}. \tag{59}$$

Note that the first term on the right side of (59) converges to the group inverse $K^\#$ as $\psi \to 0$ [5, Theorem 7.2.1]. Also,

$$\lim_{\psi \to 0} e^{K_\psi x} = \lim_{\psi \to 0} S_\psi \begin{bmatrix} e^{J_\psi x} & \\ & e^{\omega_\psi x} \end{bmatrix} S_\psi^{-1} = S \begin{bmatrix} e^{Jx} & \\ & 1 \end{bmatrix} S^{-1} = e^{Kx}.$$

Thus,

$$\begin{aligned}
&\lim_{\psi\to 0}(-K_\psi)^{-1}(I - e^{K_\psi x}) \\
&\quad = \lim_{\psi\to 0}(-K_\psi)^{-1}S_\psi \begin{bmatrix} I - e^{J_\psi x} & \\ & 1 - e^{\omega_\psi x} \end{bmatrix} S_\psi^{-1} \\
&\quad = -K^{\#}(I - e^{Kx}) - \lim_{\psi\to 0} S_\psi \begin{bmatrix} 0 & \\ & \omega_\psi^{-1}(1 - e^{\omega_\phi x)} \end{bmatrix} S_\psi^{-1} \\
&\quad = -K^{\#}(I - e^{Kx}) + x\boldsymbol{v}\boldsymbol{u},
\end{aligned}$$

where $\boldsymbol{v}$ and $\boldsymbol{u}$ are respectively the right and left eigenvectors of K for the eigenvalue 0. By Lemma 5.4 again, this completes the proof of (55). $\square$

Acknowledgements

The authors thank the Ministère de la Communauté française de Belgique for supporting this research through the ARC grant AUWB-08/13–ULB 5, they acknowledge the financial support of the Australian Research Council through the Discovery Grant DP110101663. Giang Nguyen also acknowledges the support of ACEMS (ARC Centre of Excellence for Mathematical and Statistical Frontiers).

References

[1] N. G. Bean and M. M. O'Reilly. Performance measures of a multi-layer Markovian fluid model. *Annals of Operations Research*, 160:99–120, 2008.

[2] N. G. Bean, M. M. O'Reilly, and P. Taylor. Hitting probabilities and hitting times for stochastic fluid flows. The bounded model. *Probability in the Engineering and Informational Sciences*, 23:121–147, 2009. doi:10.1017/S0269964809000102.

[3] L. Breuer. Exit problems for reflected Markov-modulated Brownian motion. *J. Appl. Probab.*, 49:697–709, 2012.

[4] L. Breuer. Occupation times for Markov-modulated Brownian motion. *Journal of Applied Probability*, 49:549–565, 2012.

[5] S. L. Campbell and C. D. Meyer. *Generalized Inverses of Linear Transformations.* Dover Publications, New York, 1991. Republication.

[6] E. Çinlar. *Introduction to Stochastic Processes.* Prentice–Hall, Englewood Cliffs, NJ, 1975.

[7] P. Coolen-Schrijner and E. A. van Doorn. The deviation matrix of a continuous-time Markov chain. *Probability in the Engineering and Informational Sciences*, 16:351–366, 2002.

[8] A. da Silva Soares and G. Latouche. A matrix-analytic approach to fluid queues with feedback control. *International Journal of Simulation. Systems, Science and Technology*, 6:4–12, 2005.

[9] A. da Silva Soares and G. Latouche. Matrix-analytic methods for fluid queues with finite buffers. *Performance Evaluation*, 63:295–314, 2006.

[10] A. da Silva Soares and G. Latouche. Fluid queues with level dependent evolution. *European J. Oper. Res.*, 196:1041–1048, 2009.

[11] B. D'Auria, J. Ivanovs, O. Kella, and M. Mandjes. First passage process of a Markov additive process with applications to reflection problems. *ArXiv e-prints*, arXiv:1006.2965v1 [math.PR], 2010.

[12] B. D'Auria, J. Ivanovs, O. Kella, and M. Mandjes. Two-sided reflection of Markov-modulated Brownian motion. *Stochastic Models*, 28:316–332, 2012. doi:10.1080/15326349.2012.672285.

[13] M. Gribaudo, D. Manini, B. Sericola, and M. Telek. Second order fluid models with general boundary behaviour. *Annals of Operations Research*, 160:69–82, 2008.

[14] J. Ivanovs. Markov-modulated Brownian motion with two reflecting barriers. *J. Appl. Probab.*, 47:1034–1047, 2010.

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

[16] G. Latouche and G. T. Nguyen. Fluid approach to two-sided Markov-modulated Brownian motion. *Queueing Systems*, 80:105–125, 2015. doi: 10.1007/s11134-014-9432-8.

[17] G. Latouche and G. T. Nguyen. The morphing of fluid queues into Markov-modulated Brownian motion. *Stochastic Systems*, 5:62–86, 2015.
doi: 10.1214/13-SSY133.

[18] G. Latouche and G. T. Nguyen. Slowing time: Markov-modulated Brownian motion with a sticky boundary. *Submitted*, 2015. arXiv:1508.00922.

[19] G. Latouche and P. Taylor. A stochastic fluid model for an ad hoc mobile network. *Queueing Systems*, 63:109–129, 2009.

[20] 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.

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