J. Appl. Prob. Spec. Vol. **48A,** 309–318 (2011)

NEW FRONTIERS IN APPLIED PROBABILITY

A Festschrift for SØREN ASMUSSEN
Edited by P. GLYNN, T. MIKOSCH and T. ROLSKI

Part 7. Queueing theory and Markov processes

DUALITY RESULTS FOR MARKOV-MODULATED FLUID FLOW MODELS

SOOHAN AHN, *University of Seoul*

Department of Statistics, The University of Seoul, Seoul 130-743, Korea.

VAIDYANATHAN RAMASWAMI, *AT&T Labs-Research*

AT&T Labs-Research, 180 Park Avenue, Florham Park, NJ 07932, USA.
Email address: vram@research.att.com

APPLIED PROBABILITY TRUST
AUGUST 2011

https://doi.org/10.1239/jap/1318940473 Published online by Cambridge University Press

DUALITY RESULTS FOR MARKOV-MODULATED FLUID FLOW MODELS

BY SOOHAN AHN AND VAIDYANATHAN RAMASWAMI

Abstract

We establish some interesting duality results for Markov-modulated fluid flow models. Though fluid flow models are continuous-state analogues of quasi-birth-and-death processes, some duality results do differ by the inclusion of a scaling factor.

Keywords: Fluid flow; time reversal; duality

2010 Mathematics Subject Classification: Primary 60K25
Secondary 90B05; 68M20

1. Introduction

A Markov-modulated fluid flow (MMFF) model is a real-valued stochastic process with piecewise-linear paths whose instantaneous rates of change depend on the state of an associated continuous-time Markov chain called the phase process. In [15], Ramaswami established a strong connection between such processes and quasi-birth-and-death (QBD) models [8], [12], [13], which are discrete state space processes on the integers modulated by a Markov chain and which change by at most one unit up or down in each step. That connection was exploited by Ramaswami in a series of papers [1]–[5], [16] with several coauthors to obtain detailed steady state and transient analyses with powerful algorithms for both bounded and unbounded fluid flows; see the references in those papers for related work by Asmussen [7] and many others. In this paper we explore one further similarity between MMFFs and QBD processes through a set of duality results involving time reversal. Surprisingly, however, there are subtle differences in the duality results, making reliance on mere intuition somewhat hazardous.

In the case of QBD processes, the main duality result [14], which has found many uses [17], is one that relates the density of a first passage time to a downward level in a QBD process and the Markov renewal density in a dual QBD process of taboo visits to an immediately upward level. The dual of a QBD process is obtained by reversing the sequence of phases, changing upward jumps of levels into downward jumps and vice versa [14]; this is tantamount to looking at paths backwards in time [9] (see Figure 1(a)). Specifically, denote by $g(i, j, n)$ the probability that a first passage in the original QBD process from the state $(1, i)$ to level $\mathbf{0}$, which is the set of states of the form $(0, k)$, occurs at step n and through a visit to the specific state $(0, j)$. Denote by $r(j, i, n)$ the probability that at step n the time-reversed QBD process is in state $(1, i)$ avoiding the level $\mathbf{0}$ in the time interval $(0, n]$, given that it started in $(0, j)$ at time 0. Then we can show that $\pi(i)g(i, j, n) = \pi(j)r(j, i, n)$—in matrix notation $G = \Delta_\pi^{-1} R^\top \Delta_\pi$—where π is the steady state probability vector for the phases, Δ_π is a diagonal matrix with π on the diagonal, and $R^\top$ stands for the transpose of the matrix R.

Here, some duality relations for fluid models are shown. They are governed by almost similar, albeit somewhat different, equations involving not only the time reversal of environmental states

https://doi.org/10.1239/jap/1318940473 Published online by Cambridge University Press

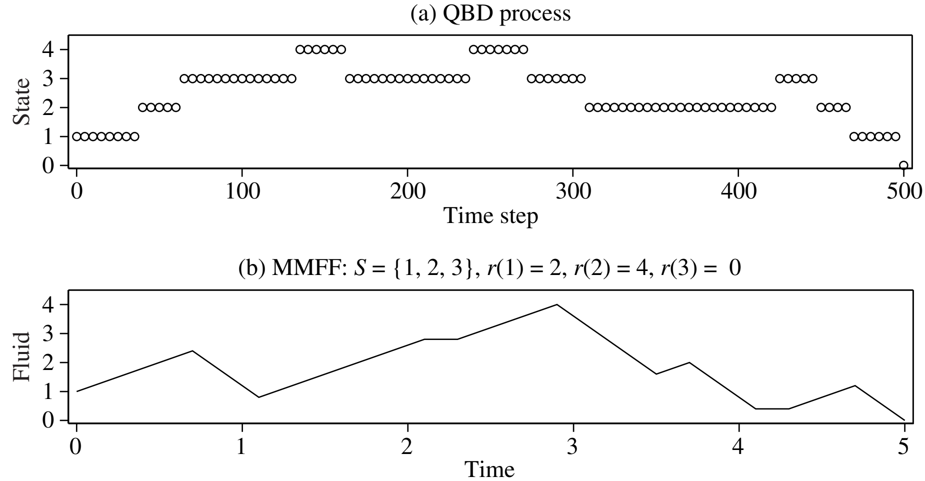

FIGURE 1: Viewed left to right, we have a first passage from level **1** to level **0**; read backwards, ups and downs are reversed, and the reversed path can be viewed as a visit (not necessarily the first) in the time-reversed process from **0** to **1** avoiding **0**.

but also some scale changes. As an example, we obtain

$$\pi(i)c(i)\Psi(i,j,s) = \pi(j)c(j)\Psi^d(j,i,s),$$

where $\Psi(s)$ and $\Psi^d(s)$ are the transforms of the respective return time distributions to level **0** in a fluid model and its time reversal. Note that the corresponding results for a QBD process, whether in discrete or continuous time, would not have a scaling factor (the $c(i)$ terms) in the duality formula. The apparent anomaly of the appearance of the additional scaling factors is perhaps explained by the difference between thinking in terms of 'kinematic states', which only describe the position of a particle, and 'dynamical states', which also include local velocities (see Figure 1(b)). See also the interesting discussion in Etter [11] which shows the relevance particularly in time-reversal arguments of these distinctions, well recognized by quantum physicists. In the QBD process case, the velocities being unity do not show up explicitly in the duality results. An explanation of this can also be given in terms of the change of variables discussed in [4]; see Theorem 3.2.1 and Remark 3.2.3 therein. In any case, ignoring this and relying only on intuition based on the consideration of kinematic states alone, such as in Figure 1(a), can indeed lead to erroneous conclusions.

1.1. The primal model

Assume that $\mathcal{J} = \{J(t),\ t \geq 0\}$ is a continuous-time, irreducible Markov chain with finite state space $S = S_1 \cup S_2 \cup S_3$ and infinitesimal generator Q. Then, when partitioned according to the sets S_i, Q has the form

$$Q = \begin{pmatrix} Q_{11} & Q_{12} & Q_{13} \\ Q_{21} & Q_{22} & Q_{23} \\ Q_{31} & Q_{32} & Q_{33} \end{pmatrix}.$$

Specifically, the submatrix Q_{ij} contains the elements $Q(r,s)$ of the infinitesimal generator with $r \in S_i$ and $s \in S_j$. (Throughout, for any matrix A, we shall denote its elements by $A(i,j)$ or by $[A]_{ij}$, and reserve the notation A_{ij} for the submatrix of A with row indices in S_i and column indices in S_j. Similarly, for a matrix function $A(s)$, we shall denote its (i, j)th element by $A(i, j, s)$ or by $[A(s)]_{ij}$, and reserve the notation $A_{ij}(s)$ for the submatrix of $A(s)$ formed by row indices in S_i and column indices in S_j.)

We assume that $c(i)$, $i \in S_1 \cup S_2$, is a set of positive constants and that a fluid process $\{F(t),\ t \geq 0\}$ evolves in such a way that during sojourn of $\mathcal{J}$ in state $i \in S_1$, the fluid level increases at rate $c(i)$; during sojourn of $\mathcal{J}$ in state $i \in S_2$, the fluid level decreases at (absolute) rate $c(i)$; during sojourn of $\mathcal{J}$ in state $i \in S_3$, the fluid level remains constant. Note that, when $J(t) = i \in S_2$, the instantaneous rate of change of $F(\cdot)$ at t is given by $-c(i) < 0$.

The pair $(\mathcal{F}, \mathcal{J})$ shall denote the bivariate process $\{(F(t), J(t)) \colon t \geq 0\}$ and will be referred to as the primal model with associated parameters (Q, r), where the vector of fluid rates $r = (c_1, -c_2, 0_{|S_3|})$ with 0_n denoting a zero vector of order n.

1.2. The dual model

Let π denote the stationary probability vector of the Markov chain $\mathcal{J}$, that is, π is the unique positive vector satisfying $\pi Q = 0$ and $\pi \mathbf{1} = 1$, where $\mathbf{1}$ is a column vector of 1s of appropriate dimension. We can partition π as $\pi = (\pi_1, \pi_2, \pi_3)$, where π_i is the subvector of components of π for states in the set S_i. Also, we denote by Δ_π a diagonal matrix with π on the diagonal.

Given the primal flow as defined earlier, its dual fluid flow $(\mathcal{F}^d, \mathcal{J}^d)$ is defined as the fluid flow with parameters (Q^d, r^d), where $Q^d = \Delta_\pi^{-1} Q^\top \Delta_\pi$, $r^d = -r$, and $Q^\top$ denotes the transpose of the matrix Q. The Markov chain $\mathcal{J}^d$ governed by Q^d can be considered as the time reversal of the Markov chain $\mathcal{J}$, and, furthermore, the relationship in the fluid rates entails that in the reversal $\mathcal{J}^d$ the sets of ascent and descent have also been interchanged. This corresponds to 'looking at the paths in reverse', similar to what was done in Figure 1.

Our goal is to examine the relationships that govern various first passage times and Markov renewal functions in a primal-dual pair. But before we get to our main results, we set some notation in line with our earlier papers. Let C_i, $i = 1, 2, 3$, and C denote the set of diagonal matrices given by

$$C_1 = \operatorname{diag}(c(i),\ i \in S_1), \qquad C_2 = \operatorname{diag}(c(i),\ i \in S_2), \qquad C_3 = I_{|S_3|},$$

where I_n is an identity matrix of order n, and

$$C = \operatorname{diag}(C_1, C_2, C_3).$$

For a distribution function $\alpha(t, x)$ over t for given x, $\hat{\alpha}(s, x)$ shall denote the Laplace–Stieltjes transform (LST) of α with respect to t. Unless otherwise stated, the argument s of such transforms is assumed to be such that $\operatorname{Re}(s) \geq 0$.

We also use the notation P_{xi} and E_{xi} to denote the conditional probability and conditional expectation, respectively, given that $F(0) = x$ and $J(0) = i$. Finally, ${}_y\mathrm{P}_{xi}$ and ${}_y\mathrm{E}_{xi}$ shall denote the taboo conditional probability and taboo conditional expectation, respectively, given that $F(0) = x$ and $J(0) = i$, and are taken over paths wherein the MMFF avoids the fluid levels $[0, y]$ (except possibly at time 0).

2. Some kernels and first passage times

We recall some key kernels and first passage time results for fluid flows from our prior work. For completeness, we shall occasionally provide a sketch of their proofs but, for ease of reading and for maintaining the flow of ideas, shall use only simple arguments as in [4]; more rigorous and formal proofs can be given using stochastic discretization methods [2].

2.1. Busy period

A basic quantity in the analysis of the fluid flow model $(\mathcal{F}, \mathcal{J})$ is the matrix $\Psi(s)$ of order $|S_1| \times |S_2|$, whose (i, j)th element, $i \in S_1$ and $j \in S_2$, is the LST

$$[\Psi(s)]_{ij} = \mathrm{E}_{0i}[\mathrm{e}^{-s\tau}\chi\{J(\tau) = j\}],$$

where $\tau = \inf\{t > 0\colon F(t) = 0\}$ and $\chi\{A\}$ denotes the indicator function of the set A. The quantity τ is often called the busy period of the fluid flow, and $\Psi(s)$ gives the joint distribution of τ and the phase at τ, given the phase at the start of the busy period. We have, in [3], developed a powerful algorithm to compute the matrix of transforms $\Psi(s)$ in terms of which we can evaluate a very large number of transient results of interest; see the cited references in [1]–[5], and [16]. For other algorithms for $\Psi(s)$, see [10].

2.2. Downward passage times

We can also consider the first passage to fluid level 0 starting at any fluid level $x \geq 0$. To this end, let

$$[\hat{G}(s, x)]_{ij} = \mathrm{E}_{xi}[\mathrm{e}^{-s\tau}\chi\{J(\tau) = j\}],$$

where τ is the first entrance time to level 0. The associated distribution functions $G(t, x)$ give the probabilities of a first passage to 0 before time t. Partitioning $\hat{G}(s, x)$ according to the sets S_i, $i = 1, 2, 3$, we obtain from the fact that fluid level decreases only in states of S_2 the structure

$$\hat{G}(s, x) = \begin{pmatrix} 0 & \hat{G}_{12}(s, x) & 0 \\ 0 & \hat{G}_{22}(s, x) & 0 \\ 0 & \hat{G}_{32}(s, x) & 0 \end{pmatrix}.$$

Furthermore, we can easily establish the following result.

Theorem 1. *For $x > 0$,*

(a) $\hat{G}_{12}(s, x) = \Psi(s)\hat{G}_{22}(s, x)$,

(b) $\hat{G}_{32}(s, x) = (sI - Q_{33})^{-1}Q_{31}\hat{G}_{12}(s, x) + (sI - Q_{33})^{-1}Q_{32}\hat{G}_{22}(s, x)$,

(c) $\hat{G}_{22}(s, x) = \mathrm{e}^{H(s)x}$, *where*

$$H(s) = C_2^{-1}(Q_{22} - sI) + C_2^{-1}Q_{23}(sI - Q_{33})^{-1}Q_{32} + [C_2^{-1}Q_{21} + C_2^{-1}Q_{23}(sI - Q_{33})^{-1}Q_{31}]\Psi(s).$$

Proof. For the proof, we refer the reader to [2] and [4].

Remark 1. Note that, when S_3 is empty, as indeed occurs in many applications, all relevant results are obtained by simply dropping from the formulae all terms involving the set S_3. Also, when S_3 is nonempty, the invertibility of $(sI - Q_{33})$ for $\mathrm{Re}(s) \geq 0$ follows from the irreducibility of the Markov chain $\mathcal{J}$.

2.3. Counts of taboo upward visits

For $x, y \geq 0$, let $N_j(t, x + y)$ denote the number of visits of $(\mathcal{F}, \mathcal{J})$ to the state $(x + y, j)$ in the time interval $[0, t]$. Let

$$[\Phi(t, x, x + y)]_{ij} = {}_x\mathrm{E}_{xi}[N_j(t, x + y)], \qquad t \geq 0,$$

$$[\phi(t, x, x + y)]_{ij} = \frac{\partial}{\partial t}[\Phi(t, x, x + y)]_{ij}, \qquad t > 0.$$

We recognize the matrix $\Phi(t, x, x+y)$ to be a (taboo) Markov renewal kernel counting the expected number of visits to upward levels and $\phi(t, x, x+y)$ to be the associated Markov renewal density. From standard results in Markov renewal theory, we may interpret $[\phi(t, x, x+y)]_{ij}\,\mathrm{d}t$ as the elementary probability that, starting in (x, i), the MMFF visits $(x+y, j)$ in the time interval $(t, t+\mathrm{d}t)$, avoiding the set $[0, x]$ of fluid levels in the time interval $(0, t]$.

Note that, for $y > 0$, the submatrix $\Phi_{ij}(t, x, x+y) = 0$ for $i = 2, 3$ and $j = 1, 2, 3$, because if the MMFF starts off in an environment state in $S_2 \cup S_3$ then the process either stays at the initial fluid level for a positive amount of time or must cross the initial level at least once before hitting $x+y$, and the required taboo visits do not occur. Since the flow rates and transition rates do not depend on the fluid level, we can also assert that

$$\Phi(t, x, x+y) = \Phi(t, 0, y).$$

Now, if we denote by $\hat{\Phi}(s, x)$ the LST of $\Phi(t, 0, x)$ with respect to time t and a transform variable s, then $\hat{\Phi}(s, x)$ has the following partitioned form:

$$\hat{\Phi}(s, x) = \begin{pmatrix} \hat{\Phi}_{11}(s, x) & \hat{\Phi}_{12}(s, x) & \hat{\Phi}_{13}(s, x) \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

We also have the following theorem [2].

Theorem 2. *We have*

(a) $\hat{\Phi}_{11}(s, x) = \mathrm{e}^{K(s)x}$,

(b) $\hat{\Phi}_{12}(s, x) = \mathrm{e}^{K(s)x}\Psi(s)$,

(c) $\hat{\Phi}_{13}(s, x) = [\hat{\Phi}_{11}(s, x)C_1^{-1}Q_{13} + \hat{\Phi}_{12}(s, x)C_2^{-1}Q_{23}](sI - Q_{33})^{-1}$, *where*

$$\begin{aligned} K(s) = {} & C_1^{-1}(Q_{11} - sI) + C_1^{-1}Q_{13}(sI - Q_{33})^{-1}Q_{31} \\ & + \Psi(s)[C_2^{-1}Q_{21} + C_2^{-1}Q_{23}(sI - Q_{33})^{-1}Q_{31}]. \end{aligned} \tag{1}$$

Proof. For the proof, we refer the reader to [2] and [4].

2.4. State distribution at time t

Defining

$$[v(t, x)]_{ij} = \frac{\partial}{\partial x}{}_0\mathrm{P}_{0i}[F(t) \le x,\ J(t) = j], \qquad i \in S_1,\ j \in S,$$

it is easy to see that $[v(t, x)]_{ij}\,\mathrm{d}x$ is the elementary probability that at time t the fluid level is in $(x, x+\mathrm{d}x)$ with phase in j and has arrived there without returning to level 0, given that the process started in state $(0, i)$. We also introduce the associated Laplace transform

$$[\hat{v}(s, x)]_{ij} = \int_0^\infty \mathrm{e}^{-st}[v(t, x)]_{ij}\,\mathrm{d}t.$$

These quantities define, in turn, the matrices $v(t, x)$ and $\hat{v}(s, x)$, respectively, which are all of dimension $|S_1| \times |S|$.

The next theorem provides an important relationship between the state densities given by $\hat{v}(s, x)$ and the Markov renewal densities of $\hat{\Phi}(s, x)$. This was proved as Theorem 3.2.1 in [4]

https://doi.org/10.1239/jap/1318940473 Published online by Cambridge University Press

using a Kolmogorov differential equations approach, and also in [2] using a more rigorous treatment based on stochastic discretization. This result is indeed related to a change-of-variables formula in integration, and underlying it is the simple fact that in phase i, a small $\mathrm{d}t$ increment in time results in a $\mathrm{d}x = \pm c_i\,\mathrm{d}t$ change in the fluid level.

Theorem 3. *For all $x > 0$ and $\mathrm{Re}(s) > 0$, when we partition the matrix $\hat{v}(s,x)$ as*

$$\hat{v}(s,x) = [\hat{v}_{11}(s,x) \mid \hat{v}_{12}(s,x) \mid \hat{v}_{13}(s,x)],$$

according to the sets S_i, $1 \le i \le 3$, the submatrices are given by the following formulae:

$$\hat{v}_{11}(s,x) = \hat{\Phi}_{11}(s,x)C_1^{-1}, \qquad \hat{v}_{12}(s,x) = \hat{\Phi}_{12}(s,x)C_2^{-1}, \qquad \hat{v}_{13}(s,x) = \hat{\Phi}_{13}(s,x).$$

3. Duality

Recall that in the dual process $(\mathcal{F}^d, \mathcal{J}^d)$, the phase process $\mathcal{J}^d$ is the time-reversed version of $\mathcal{J}$ in the primal and has the infinitesimal generator $Q^d = \Delta_\pi^{-1} Q^\top \Delta_\pi$, where π is the steady state probability vector of $\mathcal{J}$ and $\Delta_\pi = \mathrm{diag}(\pi)$. If we consider partitioned submatrices then we can easily observe that, for $i, j = 1, 2, 3$,

$$Q_{ij}^d = \Delta_i^{-1}[Q_{ji}]^\top \Delta_j, \quad \text{where} \quad \Delta_i = \mathrm{diag}(\pi(k),\ k \in S_i). \tag{2}$$

In the dual process, the fluid level decreases at rate $c(i)$ for a phase $i \in S_1$; the fluid level increases at rate $c(i)$ for a phase $i \in S_2$; for $i \in S_3$, the fluid level remains constant. Now, if we define for the dual process the transform matrices $\Psi^d(s)$, $\hat{\Phi}^d(s,x)$, and $\hat{G}^d(s,x)$ as analogues of the transforms $\Psi(s)$, $\hat{\Phi}(s,x)$, and $\hat{G}(s,x)$ of the primal, then we get the following two corollaries immediately from Theorems 1 and 2 applied to the dual process.

Corollary 1. *For $x \ge 0$,*

(a) $\hat{G}_{11}^d(s,x) = \mathrm{e}^{H^d(s)x}$,

(b) $\hat{G}_{21}^d(s,x) = \Psi^d(s)\hat{G}_{11}^d(s,x)$,

(c) $\hat{G}_{31}^d(s,x) = (sI - Q_{33}^d)^{-1}Q_{31}^d\hat{G}_{11}^d(s,x) + (sI - Q_{33}^d)^{-1}Q_{32}^d\hat{G}_{21}^d(s,x)$, *where $H^d(s)$ is given by*

$$H^d(s) = C_1^{-1}(Q_{11}^d - sI) + C_1^{-1}Q_{13}^d(sI - Q_{33}^d)^{-1}Q_{31}^d + [C_1^{-1}Q_{12}^d + C_1^{-1}Q_{13}^d(sI - Q_{33}^d)^{-1}Q_{32}^d]\Psi^d(s). \tag{3}$$

Corollary 2. *For s with $\mathrm{Re}(s) \ge 0$ and $x \ge 0$, we have*

(a) $\hat{\Phi}_{21}^d(s,x) = \mathrm{e}^{K^d(s)x}\Psi^d(s)$,

(b) $\hat{\Phi}_{22}(s,x) = \mathrm{e}^{K^d(s)x}$,

(c) $\hat{\Phi}_{23}(s,x) = [\hat{\Phi}_{21}(s,0,x)C_1^{-1}Q_{13}^d + \hat{\Phi}_{22}(s,0,x)C_2^{-1}Q_{23}^d](sI - Q_{33}^d)^{-1}$, *where $K^d(s)$ and $\Psi^d(s)$ satisfy the following equation:*

$$K^d(s) = C_2^{-1}(Q_{22}^d - sI) + C_2^{-1}Q_{23}^d(sI - Q_{33}^d)^{-1}Q_{32}^d + \Psi^d(s)[C_1^{-1}Q_{12}^d + C_1^{-1}Q_{13}^d(sI - Q_{33}^d)^{-1}Q_{32}^d].$$

A key result we need is the following.

https://doi.org/10.1239/jap/1318940473 Published online by Cambridge University Press

Lemma 1. *For* $\mathrm{Re}(s) \geq 0$, *the matrix* $\Psi(s)$ *satisfies the equation*

$$\begin{aligned} 0 &= C_1^{-1}Q_{12} + C_1^{-1}Q_{13}(sI - Q_{33})^{-1}Q_{32} \\ &\quad + \Psi(s)[C_2^{-1}(Q_{22} - sI) + C_2^{-1}Q_{23}(sI - Q_{33})^{-1}Q_{32}] \\ &\quad + [C_1^{-1}(Q_{11} - sI) + C_1^{-1}Q_{13}(sI - Q_{33})^{-1}Q_{31}]\Psi(s) \\ &\quad + \Psi(s)[C_2^{-1}Q_{21} + C_2^{-1}Q_{23}(sI - Q_{33})^{-1}Q_{31}]\Psi(s). \end{aligned} \tag{4}$$

Furthermore, for $s \geq 0$, *the matrix* $\Psi(s)$ *is the minimal nonnegative solution of (4).*

Proof. We derive (4) using a Markov renewal argument; for a proof of the minimality result, we refer the reader to [10]. Every busy period should end with an interval of time wherein the phase process lies within the set $S_2 \cup S_3$, which is the set of no further ascent for the fluid. Conditioning on the amount of fluid y and the phase at the beginning of such an interval, we can write

$$\Psi(s) = \int_0^\infty \mathrm{e}^{K(s)y} C_1^{-1}(Q_{12} + Q_{13}(sI - Q_{33})^{-1}Q_{32})\mathrm{e}^{D(s)y}\,\mathrm{d}y, \tag{5}$$

where $D(s) = C_2^{-1}(Q_{22} - sI) + C_2^{-1}Q_{23}(sI - Q_{33})^{-1}Q_{32}$. Multiplying both sides by $D(s)$ and performing an integration by parts yields

$$\Psi(s)D(s) = -C_1^{-1}(Q_{12} + Q_{13}(sI - Q_{33})^{-1}Q_{32}) - K(s)\Psi(s),$$

and (4) then follows upon substituting $K(s)$ given in (1).

Applying the above result to the dual process immediately yields the following corollary.

Lemma 2. *For* $\mathrm{Re}(s) \geq 0$, *the matrix* $\hat{\Psi}^d(s)$ *satisfies the equation*

$$\begin{aligned} 0 &= C_2^{-1}Q_{21}^d + C_2^{-1}Q_{23}^d(sI - Q_{33}^d)^{-1}Q_{31}^d \\ &\quad + \Psi^d(s)[C_1^{-1}(Q_{11}^d - sI) + C_1^{-1}Q_{13}^d(sI - Q_{33}^d)^{-1}Q_{31}^d] \\ &\quad + [C_2^{-1}(Q_{22}^d - sI) + C_2^{-1}Q_{23}^d(sI - Q_{33}^d)^{-1}Q_{32}^d]\Psi^d(s) \\ &\quad + \Psi^d(s)[C_1^{-1}Q_{12}^d + C_1^{-1}Q_{13}^d(sI - Q_{33}^d)^{-1}Q_{32}^d]\Psi^d(s). \end{aligned} \tag{6}$$

Furthermore, for $s \geq 0$, $\Psi^d(s)$ *is the minimal nonnegative solution of (6).*

We are now ready to state and prove our first duality result.

Theorem 4. *For s with* $\mathrm{Re}(s) \geq 0$,

$$\Psi(s) = \Delta_1^{-1}C_1^{-1}[\Psi^d(s)]^\top C_2\Delta_2.$$

Proof. Transposing both sides of (6) and using the relations in (2), we obtain

$$\begin{aligned} 0 &= C_1^{-1}Q_{12} + C_1^{-1}Q_{13}(sI - Q_{33})^{-1}Q_{32} \\ &\quad + \Psi^*(s)[C_2^{-1}(Q_{22} - sI) + C_2^{-1}Q_{23}(sI - Q_{33})^{-1}Q_{32}] \\ &\quad + [C_1^{-1}(Q_{11} - sI) + C_1^{-1}Q_{13}(sI - Q_{33})^{-1}Q_{31}]\Psi^*(s) \\ &\quad + \Psi^*(s)[C_2^{-1}Q_{21} + C_2^{-1}Q_{23}(sI - Q_{33})^{-1}Q_{31}]\Psi^*(s), \end{aligned}$$

where $\Psi^*(s) = \Delta_1^{-1}C_1^{-1}(\Psi^d(s))^\top C_2\Delta_2$. This shows that $\Delta_1^{-1}C_1^{-1}(\Psi^d(s))^\top C_2\Delta_2$ is a solution of (4). From the minimality of $\Psi(s)$ for $s \geq 0$, we can see that

$$[\Psi(s)]_{ij} \leq [\Delta_1^{-1}C_1^{-1}(\Psi^d(s))^\top C_2\Delta_2]_{ij} \quad \text{for all } i \in S_1 \text{ and } j \in S_2,\ s \geq 0.$$

Similarly, we can show that $C_2^{-1}\Delta_2^{-1}(\Psi(s))^\top C_1\Delta_1$ is a solution of (6) and also, for all $i \in S_1$ and $j \in S_2$, $[C_2^{-1}\Delta_2^{-1}(\Psi(s))^\top C_1\Delta_1]_{ji} \geq [\Psi^d(s)]_{ji}$, which is equivalent to

$$[\Psi(s)]_{ij} \geq [\Delta_1^{-1}C_1^{-1}(\Psi^d(s))^\top C_2\Delta_2]_{ij} \quad \text{for all } i \in S_1 \text{ and } j \in S_2,\ s \geq 0.$$

The two inequalities together yield

$$\Psi(s) = \Delta_1^{-1}C_1^{-1}(\Psi^d(s))^\top C_2\Delta_2$$

for all $s \geq 0$, and, by an analytic continuation argument, the above holds for all s with $\mathrm{Re}(s) \geq 0$.

Theorem 4 enables us to derive many other duality results of which the following are some examples.

Theorem 5. *For s with $\mathrm{Re}(s) \geq 0$, the matrices $K(s)$ and $H(s)$ are related to the matrices $K^d(s)$ and $H^d(s)$ through the equations*

$$K(s) = \Delta_1^{-1}C_1^{-1}[H^d(s)]^\top C_1\Delta_1 \tag{7}$$

and

$$K^d(s) = \Delta_2^{-1}C_2^{-1}[H(s)]^\top C_2\Delta_2. \tag{8}$$

Proof. Substituting the value of $H^d(s)$ given in (3) into the right-hand side of (7), and using Theorem 4 and the relations given in (2), we see that the matrix $\Delta_1^{-1}C_1^{-1}[H^d(s)]^\top C_2\Delta_2$ equals the right-hand side of (1), which is indeed $K(s)$ by that equation. Equation (8) is immediate by applying (7) to the dual upon noting that the dual of the dual is indeed the primal.

Theorem 6. *For $k = 1, 2, 3$,*

$$\hat{\Phi}_{1k}(s, x) = \Delta_1^{-1}C_1^{-1}[\hat{G}_{k1}^d(s, x)]^\top C_k\Delta_k.$$

Proof. This result is an immediate consequence of Theorem 2, Corollary 1, Theorem 4, and Theorem 5.

The analogue of (5) for the dual yields

$$\Psi^d(s) = \int_0^\infty \mathrm{e}^{K^d(s)y}C_2^{-1}[Q_{21}^d + Q_{23}^d(sI - Q_{33}^d)^{-1}Q_{31}^d]\mathrm{e}^{D^d(s)y}\,\mathrm{d}y,$$

where $D^d(s) = C_1^{-1}(Q_{11}^d - sI) + C_1^{-1}Q_{13}^d(sI - Q_{33}^d)^{-1}Q_{31}^d$. Taking the transpose in the above equation, pre-multiplying the result by $\Delta_1^{-1}C_1^{-1}$, and then post-multiplying by Δ_2C_2, we immediately get from the duality results the equation

$$\Psi(s) = \int_0^\infty \mathrm{e}^{U(s)y}C_1^{-1}[Q_{12} + Q_{13}(sI - Q_{33})^{-1}Q_{32}]\mathrm{e}^{H(s)y}\,\mathrm{d}y, \tag{9}$$

where

$$U(s) = \Delta_1^{-1}C_1^{-1}[D^d(s)]^\top\Delta_1C_1 = C_1^{-1}(Q_{11} - sI) + C_1^{-1}Q_{13}(sI - Q_{33})^{-1}Q_{31}.$$

https://doi.org/10.1239/jap/1318940473 Published online by Cambridge University Press

Equation (9), which forms the starting point of some analysis in [10], is based on conditioning on the first exit time from the set S_1, as opposed to (5), which is based on the last exit time.

We conclude this paper by noting some connections of the duality results with the change-of-variable technique (see [4] and [16]) involved in converting various formulae based on time dynamics into formulae based on space dynamics. Note that from our duality result (7) we easily get the equation

$$\mathrm{e}^{K(s)x} C_1^{-1} = \Delta_1^{-1} (\mathrm{e}^{H^d(s)x} C_1^{-1})^{\top} \Delta_1. \tag{10}$$

With the interpretation of $v_{11}(s, x)\,\mathrm{d}x = \mathrm{e}^{K(s)x} C_1^{-1}\,\mathrm{d}x$ in Theorem 3, note that the above is a comparison of two spatial densities at fixed time points t and involves only a time reversal. Stated in the form of (10), as in the case of QBD processes, the duality result for fluid models also involves only a time reversal. Unlike this, the other duality formulae presented earlier compare two densities over time at fixed points in space, and such comparisons become valid only with an appropriate scaling related to the change of variable. In short, the use of duality results for fluid models requires much greater care.

Acknowledgements

The duality results of Ramaswami [14] for QBD processes were clarified enormously in [9] through a collaboration with Søren Asmussen. The authors thus have an added pleasure in submitting this work in celebration of that highly respected colleague. Soohan Ahn was partially supported by a Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (grant number 20100021831).

References

[1] Ahn, S. and Ramaswami, V. (2003). Fluid flow models and queues—a connection by stochastic coupling. *Stoch. Models* **19,** 325–348.

[2] Ahn, S. and Ramaswami, V. (2004). Transient analysis of fluid flow models via stochastic coupling to a queue. *Stoch. Models* **20,** 71–101.

[3] Ahn, S. and Ramaswami, V. (2005). Efficient algorithms for transient analysis of stochastic fluid flow models. *J. Appl. Prob.* **42,** 531–549.

[4] Ahn, S. and Ramaswami, V. (2006). Transient analysis of fluid flow via elementary level-crossing arguments. *Stoch. Models* **22,** 129–147.

[5] Ahn, S., Badescu, A. L. and Ramaswami, V. (2007). Time dependent analysis of finite buffer fluid flows and risk models with a dividend barrier. *Queueing Systems* **55,** 207–222.

[6] Ahn, S., Jeon, J. and Ramaswami, V. (2005). Steady state analysis of finite fluid flow models using finite QBDs. *Queueing Systems* **49,** 223–259.

[7] Asmussen, S. (1995). Stationary distributions for fluid flow models with or without Brownian noise. *Commun. Statist. Stoch. Models* **11,** 21–49.

[8] Asmussen, S. (2003). *Applied Probability and Queues*, 2nd edn. Springer, New York.

[9] Asmussen, S. and Ramaswami, V. (1990). Probabilistic interpretations of some duality results for the matrix paradigms in queueing theory. *Commun. Statist. Stoch. Models* **6,** 715–733.

[10] Bean, N. G., O'Reilly, M. M. and Taylor, P. G. (2005). Hitting probabilities and hitting times for stochastic fluid flows. *Stoch. Process. Appl.* **115,** 1530–1556.

[11] Etter, T. (2001). Dynamical Markov states and the quantum core. Presentation at the Society of Scientific Exploration Conference, San Diego, CA. Available at http://www.boundaryinstitute.org/bi/articles/DynamicalMarkov.pdf.

[12] Latouche, G. and Ramaswami, V. (1999). *Introduction to Matrix Analytic Methods in Stochastic Modeling.* Society for Industrial and Applied Mathematics, Philadelphia, PA.

[13] Neuts, M. F. (1981). *Matrix-Geometric Solutions in Stochastic Models.* The Johns Hopkins University Press, Baltimore, MD.

[14] Ramaswami, V. (1990). A duality theorem for the matrix paradigms in queueing theory. *Commun. Statist. Stoch. Models* **6,** 151–161.

https://doi.org/10.1239/jap/1318940473 Published online by Cambridge University Press

[15] RAMASWAMI, V. (1999). Matrix analytic methods for stochastic fluid flows. In *Teletraffic Engineering in a Competitive World*, eds D. Smith and P. Key, Elsevier, Amsterdam, pp. 1019–1030.
[16] RAMASWAMI, V. (2006). Passage times in fluid models with application to risk processes. *Methodology Comput. Appl. Prob.* **8,** 497–515.
[17] TAYLOR, P. G. AND VAN HOUDT, B. (2010). On the dual relationship between Markov chains of GI/M/1 and M/G/1 type. *Adv. Appl. Prob.* **42,** 210–225.

SOOHAN AHN, *University of Seoul*

Department of Statistics, The University of Seoul, Seoul 130-743, Korea.

VAIDYANATHAN RAMASWAMI, *AT&T Labs-Research*

AT&T Labs-Research, 180 Park Avenue, Florham Park, NJ 07932, USA. Email address: vram@research.att.com