Adv. Appl. Prob. **43,** 847–874 (2011)
Printed in Northern Ireland

STATIONARITY AND CONTROL OF A TANDEM FLUID NETWORK WITH FRACTIONAL BROWNIAN MOTION INPUT

CHIHOON LEE,* *Colorado State University*
ANANDA WEERASINGHE,** *Iowa State University*

Abstract

We consider a stochastic control model for a queueing system driven by a two-dimensional fractional Brownian motion with Hurst parameter $0 < H < 1$. In particular, when $H > \frac{1}{2}$, this model serves to approximate a controlled two-station tandem queueing model with heavy-tailed ON/OFF sources in heavy traffic. We establish the weak convergence results for the distribution of the state process and construct an explicit stationary state process associated with given controls. Based on suitable coupling arguments, we show that each state process couples with its stationary counterpart and we use it to represent the long-run average cost functional in terms of the stationary process. Finally, we establish the existence result of an optimal control, which turns out to be independent of the initial data.

Keywords: Stochastic control; controlled queueing system; heavy traffic theory; fractional Brownian motion; tandem queue; long-range dependence; self-similarity

2010 Mathematics Subject Classification: Primary 60K25; 68M20; 90B22
Secondary 90B18

1. Introduction

Empirical evidence of long-range dependence and self-similarity of the underlying data in several queueing systems has been observed and analyzed [11], [30], [32], [33]. One simple concrete explanation for this kind of phenomena is the behavior of superposition of many ON/OFF sources (also known as 'packet trains' [18]) with strictly alternating ON and OFF periods. It has been shown that long-range dependence and self-similarity signatures of network traffic are successfully described by stochastic models associated with fractional Brownian motion, abbreviated as FBM hereafter, with the Hurst parameter H greater than $\frac{1}{2}$ (see [15], [20], [26], [29], and [30, Chapters 7.2 and 8.7]). It is well known that such models exhibit both of these statistical features and, therefore, understanding the behavior and control of these stochastic models are of significant interest. However, the highly non-Markovian nature of FBM makes it more difficult to analyze the control problems related to such models.

In this paper we focus on a controlled queueing system driven by a two-dimensional FBM with Hurst parameter $0 < H < 1$, and it serves to approximate a controlled two-station tandem queueing model with ON/OFF sources (when $H > \frac{1}{2}$). Tandem systems can be seen in many applications, such as storage systems and high-speed communication networks, from router architectures to protocol stacks [14], [24]. Our work is motivated by the recent article of

Received 4 May 2010; revision received 11 February 2011.
* Postal address: Department of Statistics, Colorado State University, Fort Collins, CO 80523, USA.
Email address: chihoon@stat.colostate.edu
** Postal address: Department of Mathematics, Iowa State University, Ames, IA 50011, USA.

https://doi.org/10.1239/aap/1316792672 Published online by Cambridge University Press

Delgado [6], which obtained a reflected FBM model as a limiting process for fluid models with heavy-tailed ON/OFF sources in heavy traffic. We are interested in the optimal control of such reflected FBM models (we describe our connection to the work of [6] in detail in Section 3). To this end, we introduce the notion of 'thin control' related to such models with ON/OFF sources in heavy traffic and obtain a limiting controlled fluid queue driven by a two-dimensional FBM. For ordinary Brownian networks, this notion was used in [1]. This leads us to consider a drift rate control problem of a tandem fluid queueing network fed by an FBM at each station. Our analysis allows these FBMs to be correlated with a constant correlation coefficient. For a related one-dimensional controlled queue with FBM input, we refer the reader to [10]; our work extends that of [10] to the two-dimensional situation. The probability estimates for maximum workload of a one-dimensional queue fed by FBM were obtained in [9] and [35].

Our contributions are two-fold. We consider a state process represented by a two-dimensional reflected FBM model with Hurst parameter $0 < H < 1$. First, we show that, under suitable moment conditions on initial data, any state process couples with an explicitly described stationary state process and this coupling time has finite moments. Second, we establish the existence of an optimal control for a related long-term average cost minimization problem. Despite the non-Markovian behavior of the FBM, this optimal control is independent of the initial data. These results are meant as a first step towards the further analysis of networks with general topology, where the nodes are operating under advanced scheduling and routeing disciplines in a heavy traffic environment.

There are only a few stochastic control problems for models driven by FBM that are addressed in the literature. The linear quadratic regulator problem is addressed in [16] and [19]. A stochastic maximum principle was developed and applied to several stochastic control problems in [3]. We refer the reader to [16] and Chapter 9 of [4] for further examples of such control problems. In contrast with the models considered in the aforementioned references, the model described here is motivated by queueing applications in heavy traffic and involves processes with state constraints. In Section 3 we discuss a concrete example of a queueing network which leads to our model.

We begin with the definitions of multidimensional FBM and reflected FBM. We closely follow the notation of [6]. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathrm{P})$ be a given filtered probability space. A stochastic process $\boldsymbol{B}_H = \{\boldsymbol{B}_H(t) = (B_1(t), \dots, B_J(t))^\top,\ t \geq 0\}$ defined on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathrm{P})$ is called a J-dimensional FBM of Hurst parameter $H \in (0, 1)$, starting from the origin in $\mathbb{R}^J$, with drift vector $\boldsymbol{\vartheta} \in \mathbb{R}^J$ and associated matrix $\boldsymbol{\Lambda}$, if it is a continuous Gaussian process with $\boldsymbol{B}_H(0) = \boldsymbol{0}$, P-almost surely (P-a.s.) with $\mathrm{E}[\boldsymbol{B}_H(t)] = \boldsymbol{\vartheta} t$ for all $t \geq 0$, and its covariance function is given by

$$\operatorname{cov}(\boldsymbol{B}_H(t), \boldsymbol{B}_H(s)) = \mathrm{E}[(\boldsymbol{B}_H(t) - \boldsymbol{\vartheta} t)(\boldsymbol{B}_H(s) - \boldsymbol{\vartheta} s)^\top] = \Upsilon_H(s,t)\boldsymbol{\Lambda}$$

for all $s, t \geq 0$. Here $\boldsymbol{\Lambda}$ is a $J \times J$ nonnegative definite matrix and

$$\Upsilon_H(s,t) = \tfrac{1}{2}(t^{2H} + s^{2H} - |t-s|^{2H}). \tag{1.1}$$

Also, it is assumed that $\boldsymbol{B}_H$ is adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$. We will say that $\boldsymbol{B}_H$ is a J-dimensional FBM with associated data $(\boldsymbol{0}, H, \boldsymbol{\vartheta}, \boldsymbol{\Lambda})$.

Next, let $\boldsymbol{X}_0$ be an $\mathcal{F}_0$-measurable, $\mathbb{R}^J$-valued random vector with $\mathrm{E}|\boldsymbol{X}_0| < \infty$, defined on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathrm{P})$. We introduce the process $\{\boldsymbol{X}_H(t) \colon t \geq 0\}$ by

$$\boldsymbol{X}_H(t) = \boldsymbol{X}_0 + \boldsymbol{B}_H(t) \quad \text{for all } t \geq 0,$$

where $\boldsymbol{B}_H$ is a J-dimensional FBM with associated data $(\boldsymbol{0}, H, \boldsymbol{\vartheta}, \boldsymbol{\Lambda})$. Note that the process

https://doi.org/10.1239/aap/1316792672 Published online by Cambridge University Press

$(\boldsymbol{X}_H(t))_{t\geq 0}$ is adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$, $\boldsymbol{X}_H(0) = \boldsymbol{X}_0$ with $\mathrm{E}[\boldsymbol{X}_H(t) - \boldsymbol{X}_H(0)] = \boldsymbol{\vartheta} t$ for all $t \geq 0$, and its covariance matrix $\operatorname{cov}(\boldsymbol{X}_H(t), \boldsymbol{X}_H(s))$ is given by $\Upsilon_H(s, t)\boldsymbol{\Lambda}$ for all $s, t \geq 0$. We will say that $\boldsymbol{X}_H$ is a J-dimensional FBM with associated data $(\boldsymbol{X}_0, H, \boldsymbol{\vartheta}, \boldsymbol{\Lambda})$.

The following definition of a reflected FBM slightly generalizes that of [6] to allow random initial data. The stationary process $(\boldsymbol{Z}^*(t))_{t\geq 0}$ we obtain in Theorem 4.2 below turns out to be a reflected FBM (RFBM) with random initial data $\boldsymbol{Z}^*(0)$.

Definition 1.1. An RFBM on $S = \mathbb{R}_+^J$ associated with the data $(\boldsymbol{Z}_0, H, \boldsymbol{\vartheta}, \boldsymbol{\Lambda}, R)$ that starts from $\boldsymbol{Z}_0 \in S$ is a continuous J-dimensional process $\boldsymbol{Z}$, defined on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathrm{P})$ such that

(i) $\boldsymbol{Z}_0$ is an $\mathcal{F}_0$-measurable, S-valued random vector with $\mathrm{E}|\boldsymbol{Z}_0| < \infty$,

(ii) $\boldsymbol{X}_H$ is a J-dimensional FBM adapted to $(\mathcal{F}_t)_{t\geq 0}$ with associated data $(\boldsymbol{Z}_0, H, \boldsymbol{\vartheta}, \boldsymbol{\Lambda})$,

(iii) $\boldsymbol{Z}(t) = \boldsymbol{X}_H(t) + \boldsymbol{R}\boldsymbol{L}(t) \in S$ for all $t \geq 0$, P-a.s., where $\boldsymbol{R}_{J\times J}$ is the 'reflection matrix', and the process $(\boldsymbol{Z}, \boldsymbol{L})$ is adapted to $(\mathcal{F}_t)_{t\geq 0}$,

(iv) $\boldsymbol{L}$ is a J-dimensional process satisfying $L_j(0) = 0$ for $j = 1, \ldots, J$, P-a.s. For each $j = 1, \ldots, J$, L_j is continuous and nondecreasing, and L_j can increase only when $\boldsymbol{Z}(\cdot)$ is on the face $F_j = \{\boldsymbol{x} \in S : x_j = 0\}$, *i.e.* $\int_0^t \mathbf{1}_{\{Z_j(s)\neq 0\}}\, \mathrm{d}L_j(s) = 0$ for all $t \geq 0$.

For our model of the tandem queueing network with two stations, the reflection matrix is given by

$$\boldsymbol{R} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}.$$

In this case, given a filtration $(\mathcal{F}_t)_{t\geq 0}$ and a two-dimensional FBM $\boldsymbol{X}_H$ adapted to $(\mathcal{F}_t)_{t\geq 0}$ with associated data $(\boldsymbol{Z}_0, H, \boldsymbol{\vartheta}, \boldsymbol{\Lambda})$, an explicit construction of the process $(\boldsymbol{Z}, \boldsymbol{L})$ adapted to $(\mathcal{F}_t)_{t\geq 0}$ is carefully described in (2.3)–(2.6). The pathwise uniqueness of $(\boldsymbol{Z}, \boldsymbol{L})$ also follows from these equations and the uniqueness property of the one-dimensional reflection map. For a general reflection matrix $\boldsymbol{R}$ associated with a J-dimensional FBM, suitable assumptions on $\boldsymbol{R}$ to guarantee the strong existence and pathwise uniqueness of such a reflected process $\boldsymbol{Z}$ satisfying Definition 1.1(i)–(iv) are carefully described in Section 2 of [6]. We also refer the reader to Theorem 2 of [2] and Proposition 4.2 of [31] for related results on existence and pathwise uniqueness of $\boldsymbol{Z}$.

To get an idea of RFBM introduced in the above definition, we note that it behaves like an FBM in the interior of the orthant S and it is confined to the orthant by instantaneous 'reflection' at the boundary ∂S. For each j, the jth column of the reflection matrix $\boldsymbol{R}$ gives the direction of the reflection on the jth face F_j.

Here, we consider a tandem fluid queueing network with two stations $j = 1, 2$ (see Figure 1). At each station, an FBM input with Hurst parameter $0 < H < 1$ is added and these FBMs are allowed to be correlated. The constant drift rates u_1 and u_2 at these two stations are considered as control terms. Our interest here is to establish the existence of optimal controls which guarantee the minimization of an appropriate long-term average cost functional.

In our analysis, the key ingredient in the proof is a coupling method. It helps us to analyze the behavior of a controlled state process represented by an RFBM with data $(\boldsymbol{Z}_0, H, -\boldsymbol{u}, \boldsymbol{\Lambda}, \boldsymbol{R})$, where $\boldsymbol{u} = (u_1, u_2)^\top$, $\boldsymbol{R}$ is a given 2×2 reflection matrix (see Section 2 for more details), and the Hurst parameter $0 < H < 1$. We show that the two-dimensional RFBM with initial data $\boldsymbol{Z}_0$ eventually couples with the RFBM with initial data $\mathbf{0}$. Typically, such a coupling argument works with Markov processes. In our case, the main reason for validity of the

https://doi.org/10.1239/aap/1316792672 Published online by Cambridge University Press

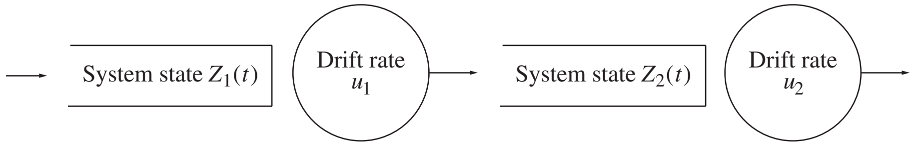

FIGURE 1: Two queues in tandem with FBM input to each station and controllable static drift rates.

coupling arguments is based on the uniqueness results related to the reflection map (also known as the Skorokhod map or the regulator map [13, Chapter 2.2], [30, Chapter 13.5]). A similar coupling method was used in the one-dimensional problem addressed in [10]. Our (coupling) techniques are different from those employed in [3], [8], [12], and [16].

This coupling argument leads us to establish the existence and uniqueness of a stationary state process for a given control (u_1, u_2). The construction of our stationary state process is explicit (see Theorem 4.2). For a tandem fluid queueing network with a general input (with stationary increments) fed only at the first station, the existence of a stationary state process was established in [5]. In our model, there is a noisy input modeled by FBM at each station and our results complement the work of [5]. We also obtain the estimates for the tail distribution of a two-dimensional stationary process. In the discrete setting, when the interarrival time and service time sequences are stationary, the stability of a system of queues in series was investigated in [21] and [22]. Our stability arguments in Theorem 4.1(a) complement the results in [21] and [22]. We refer the reader to [7], [9], and [25] for tail asymptotics of a one-dimensional queue length process with FBM input. We use the existence and uniqueness of this stationary process to show that the pay-off from the long-run average cost functional depends only on the control (u_1, u_2) and is independent of the initial data. Further analysis of the cost functional $I(u_1, u_2)$ enables us to establish the existence of an optimal control (u_1^*, u_2^*), which minimizes the cost functional over all available strategies.

The organization of the paper is as follows. In Section 2 we carefully describe our model in (2.3)–(2.6) and introduce the long-run average cost functional in (2.7). In Section 3 we provide a description of a sequence of ON/OFF network models whose limit of suitably scaled workload processes satisfies our model. This example is based on Delgado's work [6]. To obtain the controlled model of Section 2 as the limiting model, we also introduce the notion of 'thin control' for the ON/OFF queueing network in heavy traffic. Section 4 is devoted to the weak convergence results in Theorem 4.1 for the distribution of the state process with initial data (0, 0). We also construct an explicit stationary state process associated with given control (u_1, u_2) in Theorem 4.2. In Section 5 we introduce the above described coupling method and show that the arbitrary state process $\boldsymbol{Z}$ coalesces with the stationary state process $\boldsymbol{Z}^*$. We also obtain finite moment bounds for this coalescing time. The main result of this section is given in Theorem 5.1. In Section 6, by combining the results of Sections 4 and 5 we represent the long-run average cost functional in terms of the stationary state process and establish the existence of an optimal control (u_1^*, u_2^*) in Theorem 6.1. Furthermore, it turns out that this optimal control (u_1^*, u_2^*) is independent of the initial data. We indicate the generalization of our results to a tandem queueing network that consists of n stations in Section 7. In particular, we describe the distribution of the stationary process.

The following notation is used. Denote the set of real numbers by $\mathbb{R}$ and nonnegative real numbers by $\mathbb{R}_+$. Let $\mathbb{R}^d$ be the d-dimensional Euclidean space endowed with the usual Euclidean norm. For a given matrix $\boldsymbol{M}$, denote by $\boldsymbol{M}^\top$ its transpose and by M_i the ith row of $\boldsymbol{M}$. Let $\boldsymbol{I} = \boldsymbol{I}_{K\times K}$ denote the identity matrix for some K. When it is clear from the

https://doi.org/10.1239/aap/1316792672 Published online by Cambridge University Press

context, we will omit the subscript. For a set $A \subseteq \mathbb{R}^d$, denote its boundary by ∂A. When $\sup_{0\le s\le t} |f_n(s) - f(s)| \to 0$ as $n \to \infty$ for all $t \ge 0$, we say that $f_n \to f$ uniformly on compact sets. By '$\stackrel{\mathrm{D}}{=}$' and '$\stackrel{\mathrm{D}}{\to}$' we denote equality and convergence in distribution, respectively. The class of continuous functions $f\colon X \to Y$ is denoted by $C(X, Y)$. Inequalities for vectors are interpreted componentwise. We will denote generic constants by $K_1, K_2, \dots,$ and their values may change from one proof to another.

2. Model

Let $\boldsymbol{W}_H = (W_1, W_2)^\top$ be a two-dimensional FBM with data $(\boldsymbol{0}, H, \boldsymbol{0}, \boldsymbol{\Lambda})$, where

$$\boldsymbol{\Lambda} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$$

with $|\rho| < 1$ and Hurst parameter $0 < H < 1$. It is assumed that there exists a complete right-continuous filtration $(\mathcal{F}_t)_{t\ge0}$ such that $\boldsymbol{W}_H$ is adapted to this filtration. We begin with a two-dimensional controlled state process $\{\boldsymbol{Q}(t) = (Q_1(t), Q_2(t))^\top\}_{t\ge0}$ which is an RFBM. Such a state process satisfying (2.1) below will be obtained as a heavy traffic limit of a controlled ON/OFF network in Section 3. The process $\{\boldsymbol{Q}(t)\}_{t\ge0}$ takes values in the state space $S = [0,\infty) \times [0,\infty)$ and it can be written as

$$\boldsymbol{Q}(t) = \boldsymbol{Q}(0) + \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} \boldsymbol{W}_H(t) - \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} t + \begin{pmatrix} \nu_1 & 0 \\ -\nu_2 & \nu_3 \end{pmatrix} \boldsymbol{Y}(t) \tag{2.1}$$

for all $t \ge 0$, where the initial data $\boldsymbol{Q}(0) = (Q_1(0), Q_2(0))^\top \in S$ and $\boldsymbol{Q}(0)$ is an $\mathcal{F}_0$-measurable random vector such that $\mathrm{E}|\boldsymbol{Q}(0)| < \infty$. The constant control vector is given by $\boldsymbol{\theta} = (\theta_1, \theta_2)^\top$, where $\theta_1 > 0$ and $\theta_2 > 0$ are constants. Also, $\sigma_i > 0$ and $\nu_j > 0$ are constants for $i = 1, 2$ and $j = 1, 2, 3$. The two-dimensional process $\{\boldsymbol{Y}(t) = (Y_1(t), Y_2(t))^\top\}_{t\ge0}$ satisfies $\boldsymbol{Y}(0) = \boldsymbol{0}$, $Y_i(\cdot)$ is nondecreasing with continuous paths and $\int_0^\infty Q_i(t)\,\mathrm{d}Y_i(t) = 0$ for $i = 1, 2$. Next, we reduce (2.1) to a simpler model given by (2.3) and (2.4) below for further analysis. Consider the constant matrix

$$\boldsymbol{K} = \frac{1}{\sigma_1} \begin{pmatrix} 1 & 0 \\ 0 & \dfrac{\nu_1}{\nu_2} \end{pmatrix}$$

and multiply (2.1) by $\boldsymbol{K}$ to obtain

$$\boldsymbol{K}\boldsymbol{Q}(t) = \boldsymbol{K}\boldsymbol{Q}(0) + \boldsymbol{K} \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} \boldsymbol{W}_H(t) - \boldsymbol{K} \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} t + \boldsymbol{K} \begin{pmatrix} \nu_1 & 0 \\ -\nu_2 & \nu_3 \end{pmatrix} \boldsymbol{Y}(t). \tag{2.2}$$

We define

$$\boldsymbol{Z}(t) = \boldsymbol{K}\boldsymbol{Q}(t), \qquad \boldsymbol{Z}(0) = \boldsymbol{K}\boldsymbol{Q}(0), \qquad L_1(t) = \frac{\nu_1}{\sigma_1} Y_1(t), \qquad L_2(t) = \frac{\nu_1\nu_3}{\nu_2\sigma_1} Y_2(t),$$

and the new constant control vector $\boldsymbol{u} = \boldsymbol{K}\boldsymbol{\theta}$. Then our model (2.2) can be written in the form

$$Z_1(t) = Z_1(0) + W_1(t) - u_1 t + L_1(t), \tag{2.3}$$

$$Z_2(t) = Z_2(0) + \sigma W_2(t) - u_2 t - L_1(t) + L_2(t), \tag{2.4}$$

where $\sigma = \nu_1\sigma_2/\nu_2\sigma_1 > 0$. Thus, the process $\{\boldsymbol{Z}(t) = (Z_1(t), Z_2(t))^\top\}_{t\ge0}$ also takes values in the two-dimensional orthant S. The process $\boldsymbol{Z}$ is adapted to the filtration $(\mathcal{F}_t)_{t\ge0}$. Note

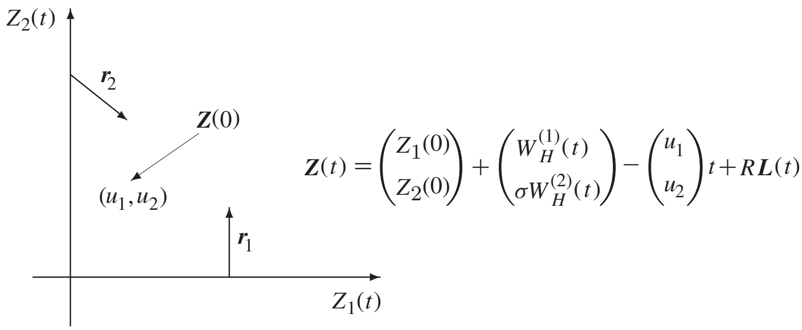

FIGURE 2: A reflected FBM in the first quadrant with drift vector $\boldsymbol{u} = (u_1, u_2)^\top$ and reflection matrix $\boldsymbol{R} = [\boldsymbol{r}_1, \boldsymbol{r}_2]$, $\boldsymbol{r}_1 = (1, 0)^\top$, $\boldsymbol{r}_2 = (1, -1)^\top$.

that, for $i = 1, 2$, $L_i(0) = 0$, and $L_i(\cdot)$ is nondecreasing with continuous paths and satisfies $\int_0^\infty Z_i(t)\, \mathrm{d}L_i(t) = 0$. The picture depicted in Figure 2 is useful for a visualization of the two-dimensional state process $\boldsymbol{Z} = (Z_1, Z_2)^\top$ in (2.3) and (2.4). Since $\boldsymbol{Z}(0) = \boldsymbol{KQ}(0)$, the random vector $\boldsymbol{Z}(0)$ is $\mathcal{F}_0$-measurable, $\boldsymbol{Z}(0) \in S$, and $\mathrm{E}\,|\boldsymbol{Z}(0)| < \infty$. Hence, the process $\boldsymbol{Z}$ is an RFBM with associated data $(\boldsymbol{KQ}(0), H, -\boldsymbol{u}, \boldsymbol{\Lambda}, \boldsymbol{R})$, where

$$\boldsymbol{\Lambda} = \begin{pmatrix} 1 & \rho\sigma \\ \rho\sigma & \sigma^2 \end{pmatrix} \quad \text{and} \quad \boldsymbol{R} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}.$$

In the later sections, we will assume a suitable moment condition on $\boldsymbol{Z}(0)$. Using the properties of the Skorokhod map (see, e.g. [30, p. 439]), we can write

$$L_1(t) = \max\Big\{0, \max_{s\in[0,t]} (u_1 s - W_1(s) - Z_1(0))\Big\}, \tag{2.5}$$

$$L_2(t) = \max\Big\{0, \max_{s\in[0,t]} (u_2 s - \sigma W_2(s) + L_1(s) - Z_2(0))\Big\}. \tag{2.6}$$

Note that, for each $j = 1, 2$, $L_j(t)$ represents the cumulative idle time in the station j during $[0, t]$.

The process $\boldsymbol{Z} = (Z_1, Z_2)^\top$ can be considered as the workload process of a two-station tandem queueing system, where the controlled queue is fed by a fractional Brownian motion to each station. In Section 3 we provide a concrete example based on the recent work of Delgado [6]. For a chosen constant control $\boldsymbol{u} = (u_1, u_2)^\top$ with $u_1 > 0, u_2 > 0$, and $\mathcal{F}_0$-measurable initial data $\boldsymbol{Z}(0) = \boldsymbol{KQ}(0) \in S$, with $\mathrm{E}\,|\boldsymbol{Z}(0)| < \infty$, the corresponding state process $\boldsymbol{Z}$ is an RFBM with associated data $(\boldsymbol{Z}(0), H, -\boldsymbol{u}, \boldsymbol{\Lambda}, \boldsymbol{R})$, where $\boldsymbol{\Lambda}$ and $\boldsymbol{R}$ are as given in the previous paragraph. Associated with this controlled state process $\boldsymbol{Z}$, the controller is faced with a cost structure consisting of the following three additive components during a time interval $[t, t + \mathrm{d}t]$:

(i) a control cost $h(\boldsymbol{u})\, \mathrm{d}t$,

(ii) a state dependent holding cost $C(\boldsymbol{Z}(t))\, \mathrm{d}t$, and

(iii) a penalty of $p_1\, \mathrm{d}L_1(t) + p_2\, \mathrm{d}L_2(t)$ for the idle times at two stations.

Here $p_1 \geq 0$ and $p_2 \geq 0$ are constants, and h and C are nonnegative continuous functions satisfying some basic assumptions. In the long-run average cost minimization problem (also

https://doi.org/10.1239/aap/1316792672 Published online by Cambridge University Press

known as the *ergodic control problem*), the controller's goal is to minimize the cost functional

$$
\begin{aligned}
I(\boldsymbol{u}, \boldsymbol{Z}(0)) &\equiv \limsup_{T\to\infty} \frac{1}{T}\,\mathrm{E}\biggl[\int_0^T [h(\boldsymbol{u}) + C(\boldsymbol{Z}(t))]\,\mathrm{d}t + \int_0^T [p_1\,\mathrm{d}L_1(t) + p_2\,\mathrm{d}L_2(t)]\biggr] \\
&= h(\boldsymbol{u}) + \limsup_{T\to\infty} \frac{1}{T}\,\mathrm{E}\biggl[\int_0^T C(\boldsymbol{Z}(t))\,\mathrm{d}t + p_1 L_1(T) + p_2 L_2(T)\biggr]. \qquad (2.7)
\end{aligned}
$$

The functions h and C satisfy the following standing assumptions.

(H1) The function $h\colon S \to [0, \infty)$ is continuous, with $h(0, 0) = 0$ and is increasing to $+\infty$ in each variable as the variable tends to ∞.

(H2) The function $C\colon S \to [0, \infty)$ is also continuous with $C(0, 0) = 0$ and nondecreasing in each variable, and $\lim_{x+y\to\infty} C(x, y) = \infty$.

(H3) The function C satisfies the following polynomial growth condition:

$$0 \le C(x, y) \le K(1 + |x|^m + |y|^m)$$

for some constants $K > 0$ and $m \ge 1$. These constants are independent of x and y.

The polynomial growth condition in (H3) for the running cost function is quite common in the stochastic control problem related to Brownian networks.

3. Controlled two-station fluid models with ON/OFF sources

Here we provide a brief description of a sequence of concrete network models in which the limit of suitably scaled workload processes satisfies a controlled RFBM model. This example is based on Delgado's work [6], whose notation we use throughout this section. It should be noted that in [6] a more general model is considered, whereas our example in this section is related to a tandem queue with two service stations. The novel feature here is the introduction of a 'thin control' using the heavy traffic condition.

Consider a sequence of controlled queueing networks indexed by (N, r), where $N \ge 1$ is an integer-valued parameter and $r > 0$ is a real-valued parameter. Each network consists of two stations ($j = 1, 2$) and there is a single server at each station (recall Figure 1). In the (N, r)th network, there are N input sources for each station (e.g. N users connected to the server) and each user stays connected to the server for a random ON period with distribution function F_1, and stays off during a random OFF period of time with distribution function F_2. It is assumed that, for each user, these 'ON' periods and 'OFF' periods are independent of each other. For each $i = 1, 2$, assume that $1 - F_i(x) \sim c_i x^{-\beta_i}$ for large x, where $1 < \beta_i < 2$, and c_i and β_i are positive constants. Hence, each F_i has finite mean $\tilde{\mu}_i$ and infinite variance. In the jth station, ON and OFF periods of the nth user are described by

$$
U_j^{(n)}(t) = \begin{cases} 1 & \text{if the } n\text{th source is 'ON' at time } t, \\ 0 & \text{if the } n\text{th source is 'OFF' at time } t. \end{cases}
$$

Assume that if all the sources are 'ON' then fluid would arrive at station j at a deterministic rate α_j^N for $j = 1, 2$.

Next, let $\boldsymbol{P} = (p_{k\ell})_{2\times 2}$ represent the 'routeing matrix' of the network. Here $p_{k\ell}$ is the proportion of fluid that leaves station k and goes next to station ℓ, and $1 - \sum_{k=1}^{2} p_{k\ell} \ge 0$ is the

https://doi.org/10.1239/aap/1316792672 Published online by Cambridge University Press

proportion of fluid that leaves the network after being served at station k. In our tandem queue example,

$$\boldsymbol{P} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

The quantities $\tilde{\mu}_1, \tilde{\mu}_2, \alpha_1^N, \alpha_2^N$, and $\boldsymbol{P}$ are considered as system primitives. We let $c = \tilde{\mu}_1/(\tilde{\mu}_1 + \tilde{\mu}_2)$, $\boldsymbol{\alpha}^N = (\alpha_1^N, \alpha_2^N)^\top$ and $\mathbb{Q} = (\boldsymbol{I} - \boldsymbol{P}^\top)^{-1}$. First, we assume that

$$\lim_{N\to\infty} \boldsymbol{\alpha}^N = \boldsymbol{\alpha},$$

where $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)^\top$ for some $\alpha_1 > 0$ and $\alpha_2 > 0$.

Then, following [6], we can compute the long-run fluid rate vector $\boldsymbol{\lambda}^N$ which satisfies the traffic equation $\boldsymbol{\lambda}^N = c\mathbb{Q}\boldsymbol{\alpha}^N$. Then $\boldsymbol{\lambda}^N$ is given by $\lambda_1^N = c\alpha_1^N$ and $\lambda_2^N = c(\alpha_1^N + \alpha_2^N)$. Furthermore,

$$\lim_{N\to\infty} \boldsymbol{\lambda}^N = \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} c\alpha_1 \\ c(\alpha_1 + \alpha_2) \end{pmatrix}. \tag{3.1}$$

For the (N, r)th system, it is assumed that the controlled deterministic service rate at station j is given by

$$\mu_j^N(r) = \lambda_j^N \left(1 + \frac{1}{\sqrt{N}}\theta_j(r)\right) \quad \text{for } j = 1, 2, \tag{3.2}$$

where the control variables $\theta_j(r)$ are positive bounded continuous functions and

$$\lim_{r\to\infty} \theta_j(r) = 0.$$

More precisely, we assume that

$$\lim_{r\to\infty} r^{1-H}\boldsymbol{\theta}(r) = \lim_{r\to\infty} \begin{pmatrix} r^{1-H}\theta_1(r) \\ r^{1-H}\theta_2(r) \end{pmatrix} = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}, \tag{3.3}$$

where $\theta_1 > 0$ and $\theta_2 > 0$ are constants and $H = \frac{1}{2}(3 - \min\{\beta_1, \beta_2\}) \in (\frac{1}{2}, 1)$. The mean service time at station j is given by $m_j^N(r) = 1/\mu_j^N(r)$ for $j = 1, 2$, and the corresponding mean service time matrix $\boldsymbol{M}^N(r)$ is given by $\boldsymbol{M}^N(r) = \operatorname{diag}(m_1^N(r), m_2^N(r))$. Note that $\lim_{N\to\infty} \mu_j^N(r) = \lambda_j$ and $\lim_{N\to\infty} \boldsymbol{M}^N(r) = \boldsymbol{M}$, where $\boldsymbol{M} = \operatorname{diag}(\lambda_1^{-1}, \lambda_2^{-1})$ and those limits are uniform on compact sets (in r).

Next, we introduce the fluid-traffic intensity vector of the (N, r)th network by

$$\boldsymbol{\rho}^N(r) = \boldsymbol{M}^N(r)\boldsymbol{\lambda}^N. \tag{3.4}$$

It is easy to see that

$$\lim_{N\to\infty} \boldsymbol{\rho}^N(r) = \boldsymbol{e} \equiv \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Then we observe the 'heavy traffic' condition

$$\lim_{N\to\infty} \sqrt{N}(\boldsymbol{\rho}^N(r) - \boldsymbol{e}) = -\boldsymbol{\theta}(r), \tag{3.5}$$

which readily follows from (3.1), (3.2), and (3.4). We remark that the effect of the control $\boldsymbol{\theta}(r)$ in the service rate is of order $1/\sqrt{N}$ and also (3.3) holds. Such controls are called 'thin controls' (see [1]). In comparison, the heavy traffic condition in [6] assumes that $\boldsymbol{\theta}(r)$ is identically zero and their network is not controlled.

To compute the workload process (i.e. the amount of time required for the server to complete processing of all fluids in the queue (or being served)), next we compute the matrix $\boldsymbol{R}^N(r) = \boldsymbol{I} - \boldsymbol{M}^N(r)\boldsymbol{P}^\top \boldsymbol{M}^N(r)^{-1}$ as in Lemma 1 of [6]. It can be easily seen that

$$\lim_{N\to\infty} \boldsymbol{R}^N(r) = \boldsymbol{R} \equiv \begin{pmatrix} 1 & 0 \\ -\dfrac{\lambda_1}{\lambda_2} & 1 \end{pmatrix},$$

which is independent of r. Then, we define the cumulative external fluid that arrives at station j during $[0, t]$ by

$$E_j^N(t) = \alpha_j^N \int_0^t \frac{1}{N}\biggl(\sum_{n=1}^N U_j^{(n)}(s)\biggr)\,\mathrm{d}s \quad \text{for } j = 1, 2.$$

The aggregated cumulative external fluid-traffic process is given by

$$\{\boldsymbol{E}^N(t) = (E_1^N(t), E_2^N(t))^\top\}_{t\ge 0}.$$

The two-dimensional workload process $\{\boldsymbol{Z}_r^N(t)\}_{t\ge0}$ and the cumulative idle time process $\{\boldsymbol{L}_r^N(t)\}_{t\ge0}$ of the (N, r)th network satisfy

$$\boldsymbol{Z}_r^N(t) = \boldsymbol{R}^N(r)\boldsymbol{M}^N(r)\mathbb{Q}\boldsymbol{E}^N(t) - \boldsymbol{R}^N(r)\boldsymbol{e}t + \boldsymbol{R}^N(r)\boldsymbol{L}_r^N(t)$$

for $t \ge 0$. Next, we introduce the scaled processes associated with the (N, r)th network. Let

$$\widehat{\boldsymbol{Z}}_r^N(t) \equiv \sqrt{N}\frac{\boldsymbol{Z}^N(rt)}{r^H \mathcal{L}^{1/2}(r)}, \qquad \widehat{\boldsymbol{E}}_r^N(t) \equiv \sqrt{N}\frac{\boldsymbol{E}^N(rt) - crt\boldsymbol{\alpha}^N}{r^H \mathcal{L}^{1/2}(r)},$$
$$\widehat{\boldsymbol{L}}_r^N(t) \equiv \sqrt{N}\frac{\boldsymbol{L}^N(rt)}{r^H \mathcal{L}^{1/2}(r)},$$

where $\mathcal{L}(r)$ is a positive, slowly varying function at ∞, as defined in Section 3.3 of [6]. Then, following the discussions in [6], these scaled processes are related by

$$\widehat{\boldsymbol{Z}}_r^N(t) = \widehat{\boldsymbol{X}}_r^N(t) + \boldsymbol{R}^N(r)\widehat{\boldsymbol{L}}_r^N(t), \tag{3.6a}$$

where

$$\widehat{\boldsymbol{X}}_r^N(t) = \boldsymbol{R}^N(r)\boldsymbol{M}^N(r)\mathbb{Q}\widehat{\boldsymbol{E}}_r^N(t) + \frac{\sqrt{N}}{r^H}\boldsymbol{R}^N(r)(\boldsymbol{\rho}^N(r) - \boldsymbol{e})rt \tag{3.6b}$$

and, for each j,

$$\widehat{L}_{r,j}^N(0) = 0, \qquad \int_0^\infty \widehat{Z}_{r,j}^N(s)\,\mathrm{d}\widehat{L}_{r,j}^N(s) = 0. \tag{3.6c}$$

Associated with the scaled processes of the (N, r)th network, we consider a long-run average cost minimization problem with linear control costs and the corresponding cost functional

$$I_r^N(\boldsymbol{\theta}(r), \widehat{\boldsymbol{Z}}_r^N(0)) = h(r^{1-H}\boldsymbol{\theta}(r)) + \limsup_{T\to\infty}\frac{1}{T}\,\mathrm{E}\biggl[\int_0^T C(\widehat{\boldsymbol{Z}}_r^N(t))\,\mathrm{d}t + \boldsymbol{p}\cdot\widehat{\boldsymbol{L}}_r^N(T)\biggr], \tag{3.7}$$

where $\boldsymbol{p} = (p_1, p_2)^\top$ is a constant vector. The function h is linear and C satisfies the assumptions in Section 2. Using Theorem 1 of [6], we can approximate the scaled system

https://doi.org/10.1239/aap/1316792672 Published online by Cambridge University Press

in (3.6) by RFBM and we can minimize the associated cost functional of the limiting RFBM system. To obtain the limiting workload process, we introduce different types of convergence and their notation as described in [6]. We denote the convergence in distribution in $C[0,\infty)$ by '$\mathcal{D}$-lim' and the convergence of finite-dimensional distributions by '$\widetilde{\lim}$'. Following the proof of Theorem 1 of [6] together with (3.3), (3.5), and (3.6), we obtain

$$\mathcal{D}\text{-}\lim_{r\to\infty} \widetilde{\lim_{N\to\infty}} \widehat{\boldsymbol{X}}_r^N(\cdot) = \boldsymbol{X}(\cdot), \quad \text{where} \quad \boldsymbol{X}(t) = \boldsymbol{R}\boldsymbol{M}\mathbb{Q}\widetilde{\boldsymbol{B}}_H(t) - \boldsymbol{R}\boldsymbol{\theta}t \quad \text{for all } t \geq 0,$$

where $\widetilde{\boldsymbol{B}}_H$ is a two-dimensional FBM with $H = (3 - \min\{\beta_1, \beta_2\})/2$ and covariance matrix $\boldsymbol{\Lambda} = \operatorname{diag}(\sigma_{\lim}^2\alpha_1^2, \sigma_{\lim}^2\alpha_2^2)$, and $\sigma_{\lim}^2 > 0$ is as in [6, p. 196]. Note that $\frac{1}{2} < H < 1$. An easy computation shows that $\boldsymbol{R}\boldsymbol{M}\mathbb{Q} = \operatorname{diag}(\lambda_1^{-1}, \lambda_2^{-1})$. Therefore, using the continuous mapping theorem, as in [6], we obtain the limiting system as an RFBM given by

$$\boldsymbol{Z}(t) = \begin{pmatrix} \sigma^2\alpha_1^2/\lambda_1 & 0 \\ 0 & \sigma^2\alpha_2^2/\lambda_2 \end{pmatrix} \boldsymbol{B}_H(t) - \begin{pmatrix} u_1 t \\ u_2 t \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ -\lambda_1/\lambda_2 & 1 \end{pmatrix} \boldsymbol{L}(t), \tag{3.8}$$

where $\boldsymbol{B}_H(\cdot)$ is a standard two-dimensional FBM, $u_1 = \theta_1, u_2 = \theta_2 - \lambda_1\theta_1/\lambda_2$, $\boldsymbol{L}(\cdot)$ represents the idle time process, and, for each j, $L_j(\cdot)$ is nondecreasing, $L_j(0) = 0$, and $\int_0^t Z_j(t)\,\mathrm{d}L_j(t) = 0$. Hence, we see that, with a superimposed ON–OFF input source and controllable services times for the queueing system, a suitably scaled workload process in the limit satisfies the model in (3.8), which is essentially (2.1). With cost structure (3.7) for the queueing network problem in mind, we intend to study in this paper a formal fractional Brownian control problem by imposing the cost functional for the limiting model (3.8) as

$$I(\boldsymbol{u}, \boldsymbol{Z}(0)) = h(\boldsymbol{u}) + \limsup_{T\to\infty} \frac{1}{T} \mathrm{E}\bigg[\int_0^T C(\boldsymbol{Z}(t))\,\mathrm{d}t + \boldsymbol{p}\cdot\boldsymbol{L}(T)\bigg], \tag{3.9}$$

where $\boldsymbol{p} = (p_1, p_2)^\top$ is a constant vector. Note that $h(\boldsymbol{u}) = u_1 + u_2 = \alpha_2\theta_1/(\alpha_1 + \alpha_2) + \theta_2 > 0$. However, we do not attempt to solve the underlying queueing control problem in this paper. A solution to the limiting control problem with cost functional (3.9) provides useful insights into the queueing network control problem with associated cost functional (3.7).

4. Weak convergence and stationarity

Recall the model described by (2.3)–(2.6). If the initial data $(Z_1(0), Z_2(0)) = (0, 0)$ then the corresponding processes Z_1^0 and Z_2^0 can be written as

$$Z_1^0(t) = W_1(t) - u_1 t + L_1^0(t) \tag{4.1}$$

and

$$Z_2^0(t) = \sigma W_2(t) - u_2 t - L_1^0(t) + L_2^0(t) \quad \text{for } t \geq 0, \tag{4.2}$$

where $\sigma > 0$ is a constant, W_1 and W_2 are correlated FBMs with constant correlation coefficient $\rho \in [-1, 1]$ and Hurst parameter $0 < H < 1$. Furthermore, we have

$$L_1^0(t) = \max_{s\in[0,t]} (u_1 s - W_1(s)) \tag{4.3}$$

and

$$L_2^0(t) = \max_{s\in[0,t]} (u_2 s - \sigma W_2(s) + L_1^0(s)) \quad \text{for } t \geq 0. \tag{4.4}$$

We define the vector-valued process $\boldsymbol{Z}^0$ by $(\boldsymbol{Z}^0(t) = (Z_1^0(t), Z_2^0(t))^\top)_{t\geq 0}$, and next we establish the weak convergence of $\boldsymbol{Z}^0$ and identify its limiting distribution $\boldsymbol{Z}^0(\infty)$.

Theorem 4.1. *The following results hold.*

(a) *Assume that* $u_1 > 0$ *and* $u_1 + u_2 > 0$. *The process* $(\mathbf{Z}^0(t))_{t\geq 0}$ *converges weakly to the random vector* $\mathbf{Z}^0(\infty) = (Z_1^0(\infty), Z_2^0(\infty))^\top$ *as* $t \to \infty$, *where* $Z_1^0(\infty)$ *and* $Z_2^0(\infty)$ *satisfy*

$$Z_1^0(\infty) = \sup_{0\leq s<\infty} \{W_1(s) - u_1 s\} \tag{4.5}$$

and

$$Z_1^0(\infty) + Z_2^0(\infty) = \sup_{0\leq r\leq s<\infty} \{(W_1(s) - u_1 s) + (\sigma W_2(r) - u_2 r)\}. \tag{4.6}$$

The random vector $\mathbf{Z}^0(\infty)$ *has a proper distribution function. (More precisely,* $Z_1^0(\infty) < \infty$ *a.s. if* $u_1 > 0$ *and* $Z_2^0(\infty) < \infty$ *a.s. when* $u_1 > 0$ *and* $u_1 + u_2 > 0$.*)*

(b) *When* $u_1 > 0$ *and* $u_2 > 0$, *the tail distribution of* $\mathbf{Z}^0(\infty)$ *satisfies*

$$\lim_{z\to\infty} z^{2H-2} \log \mathrm{P}[Z_1^0(\infty) \geq z] = -\theta^*(u_1) \tag{4.7}$$

and

$$\limsup_{z\to\infty} z^{2H-2} \log \mathrm{P}[Z_2^0(\infty) \geq z] \leq -\frac{1}{\sigma^2}\theta^*(u_2), \tag{4.8}$$

where

$$\theta^*(u) = \frac{u^{2H}}{2H^{2H}(1-H)^{2-2H}} > 0 \quad \textit{for } u > 0.$$

Proof. (a) First we consider the case in which $u_1 > 0$ and $u_2 > 0$. Consider W_1 and W_2, which are correlated FBMs with constant correlation coefficient $\rho \in [-1, 1]$. To construct such a process, we begin with two independent FBMs Y_1 and Y_2, and let, for all $t \geq 0$,

$$W_1(t) = Y_1(t), \qquad W_2(t) = \rho Y_1(t) + \bar{\rho} Y_2(t), \tag{4.9}$$

where $\bar{\rho} = \sqrt{1-\rho^2}$. To prove part (a), our first step is to show that

$$\begin{aligned}(Z_1^0(T), &Z_1^0(T) + Z_2^0(T)) \\ &\stackrel{\mathrm{D}}{=} \Big(\max_{0\leq s\leq T} \{W_1(s) - u_1(s)\}, \max_{0\leq r\leq s\leq T} \{W_1(s) - u_1 s + \sigma W_2(r) - u_2 r\}\Big)\end{aligned} \tag{4.10}$$

for each $T > 0$, where Z_1^0 and Z_2^0 satisfy (4.1) and (4.2), respectively. We keep $T > 0$ fixed and define $(B_1(s), B_2(s))_{0\leq s\leq T}$ by

$$B_1(s) = W_1(T) - W_1(T-s), \qquad B_2(s) = W_2(T) - W_2(T-s), \tag{4.11}$$

for each $0 \leq s \leq T$. Then, it is easy to verify that each $(B_i(s))_{0\leq s\leq T}$ is a one-dimensional FBM for $i = 1, 2$ (see Exercise 5.1.1 of [28, p. 286]). Using (4.9) in (4.11), we also have

$$\mathrm{E}[B_1(t)B_2(s)] = \mathrm{E}[W_1(t)W_2(s)] = \rho \Upsilon_H(s, t),$$

where $\Upsilon_H(s, t)$ is as in (1.1). Since both $(B_1(s), B_2(s))_{0\leq s\leq T}$ and $(W_1(s), W_2(s))_{0\leq s\leq T}$ are Gaussian processes with the same mean and covariance function, they induce the same measure μ_W on $C[0, T]$. The point here is that even though $(B_1(s), B_2(s))$ depends on T in (4.11), the measure induced on $C[0, T]$ is the same as that of $(W_1(s), W_2(s))_{0\leq s\leq T}$.

Using (4.1)–(4.4), we can write $Z_1^0(T)$ and $Z_2^0(T)$ in the form

$$Z_1^0(T) = \max_{0\le r\le T}\{B_1(T-r) - u_1(T-r)\}$$

and

$$\begin{aligned} Z_1^0(T) + Z_2^0(T) &= \max_{0\le r\le s\le T}\{B_1(T-r) - u_1(T-r) + \sigma B_2(T-s) - u_2(T-s)\} \\ &= \max_{0\le v\le t\le T}\{B_1(t) - u_1 t + \sigma B_2(v) - u_2 v\}. \end{aligned}$$

We have used the time substitutions $t = T - r$ and $v = T - s$ in the last equality. Since $(B_1, B_2) \stackrel{\mathrm{D}}{=} (W_1, W_2)$ on $C[0, T]$ for each $T > 0$, the desired equality (4.10) follows. We let

$$M_1(T) = \max_{0\le s\le T}\{W_1(s) - u_1 s\}$$

and

$$M_2(T) = \max_{0\le r\le s\le T}\{W_1(s) - u_1 s + \sigma W_2(r) - u_2 r\}.$$

Since W_i has stationary and ergodic increments, we have $\lim_{T\to\infty} W_i(T)/T = 0$ a.s. for $i = 1, 2$ (see, e.g. [23] and [28, Chapter 5.1] for additional properties and a more detailed description of FBM). Thus, $M_1(T) < \infty$ and $M_2(T) < \infty$ a.s. Clearly, $(M_1(T), M_2(T)) \to (M_1(\infty), M_2(\infty))$ as $T \to \infty$ a.s. and $M_i(\infty) < \infty$ for $i = 1, 2$. Hence, we can conclude that

$$(Z_1^0(T), Z_1^0(T) + Z_2^0(T)) \xrightarrow{\mathrm{D}} (M_1(\infty), M_2(\infty)) \quad \text{as } T \to \infty,$$

and, as a consequence, we have

$$(Z_1^0(T), Z_2^0(T)) \xrightarrow{\mathrm{D}} (M_1(\infty), M_2(\infty) - M_1(\infty)) \quad \text{as } T \to \infty.$$

This completes the proof of part (a) for the case in which $u_1 > 0$ and $u_2 > 0$. For the case in which $u_1 > 0$ and $u_1 + u_2 > 0$, we pick an $\varepsilon > 0$ so that $\min\{u_1, u_1 + u_2\} > \varepsilon > 0$. We intend to show that the right-hand side of (4.6) is finite. Using (4.6), we obtain

$$\begin{aligned} &\sup_{0\le r\le s}\{(W_1(s) - u_1 s) + (\sigma W_2(r) - u_2 r)\} \\ &\qquad \le \sup_{0\le r\le s}\{(W_1(s) - \varepsilon s) + (\sigma W_2(r) - (u_1 + u_2 - \varepsilon) r)\} \\ &\qquad \le \sup_{s\ge 0}(W_1(s) - \varepsilon s) + \sup_{s\ge 0}(\sigma W_2(s) - (u_1 + u_2 - \varepsilon) s). \end{aligned}$$

From the above proof for the case in which $u_1 > 0$ and $u_2 > 0$, the right-hand side of the last inequality is finite a.s. Hence, it follows that $Z_1^0(\infty) + Z_2^0(\infty) < \infty$ a.s. and this completes the proof of part (a).

(b) The result in (4.7) was shown in [7], [9], and [25]. To prove (4.8), we begin with $Z_2^0(t)$ as in (4.2). Using (4.4), we note that

$$L_2^0(t) \le L_1^0(t) + \max_{s\in[0,t]}(u_2 s - \sigma W_2(s))$$

https://doi.org/10.1239/aap/1316792672 Published online by Cambridge University Press

and, hence, by (4.2) we have

$$\begin{aligned}
Z_2^0(t) &\le \sigma W_2(t) - u_2 t + \max_{s\in[0,t]}(u_2 s - \sigma W_2(s)) \\
&= \max_{s\in[0,t]}\{\sigma(W_2(t) - W_2(s)) - u_2(t-s)\} \\
&= \max_{s\in[0,t]}\{\sigma B_2(s) - u_2 s\} \quad \text{(using (4.11))} \\
&\stackrel{\mathrm{D}}{=} \max_{s\in[0,t]}\{\sigma W_2(s) - u_2 s\}.
\end{aligned}$$

Thus,

$$\mathrm{P}[Z_2^0(T) \ge z] \le \mathrm{P}[Y(\infty) \ge z],$$

where $Y(\infty) = \sup_{0\le s}\{\sigma W_H(s) - u_2 s\} = \sigma \sup_{0\le s}\{W_H(s) - u_2 s/\sigma\}$ and W_H is a one-dimensional FBM with Hurst parameter H. Using (4.7), we can estimate $\mathrm{P}[Y(\infty) \ge z]$. Hence, result (4.8) follows.

When the drift rates u_1 and u_2 satisfy the condition $u_1 > 0 > u_2$ with $u_1 + u_2 > 0$ (i.e. $u_1 > u_1 + u_2 > 0$), we can replace (4.8) with a weaker upper bound as described below.

Corollary 4.1. *Assume that $u_1 > u_1 + u_2 > 0$. Then (4.7) holds and instead of (4.8) the following estimate holds for $Z_2^0(\infty)$:*

$$\limsup_{z\to\infty} z^{2H-2} \log \mathrm{P}[Z_2^0(\infty) \ge z] \le -\frac{1}{2^{2(1-H)}}\theta^*\left(\frac{u_1+u_2}{1+\sigma^{1/H}}\right), \tag{4.12}$$

where $\theta^*(\cdot)$ is given in Theorem 4.1(b).

Proof. The estimate (4.7) remains valid since $u_1 > 0$. It remains to estimate $\mathrm{P}[Z_2^0(\infty) \ge z]$. Since $Z_2^0(\infty) \le Z_1^0(\infty) + Z_2^0(\infty)$, we estimate $\mathrm{P}[Z_1^0(\infty) + Z_2^0(\infty) \ge z]$. We pick $0 < \varrho < u_1 + u_2$ and introduce $\varsigma = u_1 + u_2 - \varrho/\sigma > 0$. Then, using (4.6), we obtain

$$Z_1^0(\infty) + Z_2^0(\infty) \le \sup_{t\ge 0}(W_1(t) - \varrho t) + \sigma \sup_{t\ge 0}(W_2(t) - \varsigma t).$$

Hence,

$$\mathrm{P}[Z_1^0(\infty) + Z_2^0(\infty) \ge z] \le \mathrm{P}\left[\sup_{t\ge 0}(W_1(t) - \varrho t) \ge \frac{z}{2}\right] + \mathrm{P}\left[\sup_{t\ge 0}(W_2(t) - \varsigma t) \ge \frac{z}{2\sigma}\right].$$

Since $\varrho > 0$ and $\varsigma > 0$, we can use (4.7) and a straightforward calculation using the above estimate to obtain

$$\limsup_{z\to\infty} z^{2H-2} \log \mathrm{P}[Z_1^0(\infty) + Z_2^0(\infty) \ge z] \le -\frac{1}{2^{2(1-H)}}\bar{\theta}(\varrho, \sigma), \tag{4.13}$$

where

$$\bar{\theta}(\varrho, \sigma) = \min\left\{\theta^*(\varrho), \frac{1}{\sigma^{2(1-H)}}\theta^*\left(\frac{u_1+u_2-\varrho}{\sigma}\right)\right\}.$$

Using the expression for $\theta^*(\cdot)$ in Theorem 4.1(b), we observe that

$$\bar{\theta}(\varrho, \sigma) = \frac{1}{2H^{2H}(1-H)^{2(1-H)}} \min\left\{\varrho^{2H}, \left(\frac{u_1+u_2-\varrho}{\sigma^{1/H}}\right)^{2H}\right\},$$

where $0 < \varrho < u_1 + u_2$. It is straightforward to compute the maximum value of $\bar{\theta}(\varrho, \sigma)$

when $0 < \varrho < u_1 + u_2$, and it is achieved at $\varrho = \varrho^* = (u_1 + u_2)/(1 + \sigma^{1/H})$. Moreover, $\bar{\theta}(\varrho^*, \sigma) = \theta^*((u_1 + u_2)/(1 + \sigma^{1/H}))$. Since $\mathrm{P}[Z_2^0(\infty) \geq z] \leq \mathrm{P}[Z_1^0(\infty) + Z_2^0(\infty) \geq z]$, then, as a consequence of the above estimates, we obtain (4.12).

Next, we intend to establish the existence of a stationary process on the same probability space on which $\boldsymbol{Z}^0$ is defined. The coupling arguments in the next section will establish the uniqueness in law for this stationary process (see Corollary 5.1 below). For a tandem network with two stations and only one random input process with stationary ergodic increments at the first station, the existence of a unique stationary process was established in [5]. In our situation, there are two input noise processes, which are correlated and have stationary ergodic increments. The following result complements the work of [5] and provides a more explicit description of the stationary process.

Theorem 4.2. *Let $u_1 > 0$ and $u_1 + u_2 > 0$. Then there is a probability space $(\Omega, \mathcal{F}, \mathrm{P})$ supporting $\boldsymbol{Z}^0$ as described in (4.1)–(4.4) and a stationary process $\boldsymbol{Z}^* = (Z_1^*, Z_2^*)$, which satisfies the following equations with respect to the same FBMs W_1 and W_2 with correlation coefficient $\rho \in [-1, 1]$. For all $t \geq 0$,*

$$Z_1^*(t) = Z_1^*(0) + W_1(t) - u_1 t + L_1^*(t), \tag{4.14}$$

$$Z_2^*(t) = Z_2^*(0) + \sigma W_2(t) - u_2 t - L_1^*(t) + L_2^*(t). \tag{4.15}$$

Here

$$L_1^*(0) = L_2^*(0) = 0,$$

and $L_1^*(t)$ and $L_2^*(t)$ are nondecreasing, continuous processes adapted to the filtration of $\{\boldsymbol{W}_H = (W_1, W_2)^\top\}$, which also satisfy

$$\int_0^\infty Z_i^*(t)\,\mathrm{d}L_i^*(t) = 0 \quad \textit{for } i = 1, 2. \tag{4.16}$$

Let $g(x) = \mathrm{e}^{x^\alpha}$, where $0 < \alpha < 2(1 - H)$. Then $\mathrm{E}[g(\boldsymbol{Z}^*(t))] < \infty$ and, consequently, $\mathrm{E}\,|\boldsymbol{Z}^*(t)|^N < \infty$ for every $N \geq 1$.

Proof. We begin with two independent two-sided FBMs, Y_1 and Y_2, defined on some probability space $(\Omega, \mathcal{F}, \mathrm{P})$ (cf. [23]). Thus, Y_1 and Y_2 are defined for all $-\infty < t < +\infty$. We let $W_1 = Y_1$ and $W_2 = \rho Y_1 + \bar{\rho} Y_2$, where $\bar{\rho} = \sqrt{1 - \rho^2}$. Then $\boldsymbol{W}_H(t) = (W_1(t), W_2(t))^\top$ is defined for all $-\infty < t < +\infty$ and $\boldsymbol{W}_H(0) = \boldsymbol{0}$.

Next, we consider the two-dimensional ‘free process’

$$(W_1(t) - u_1 t, (W_1(t) - u_1 t) + (\sigma W_2(t) - u_2 t))^\top$$

for all $-\infty < t < +\infty$. Introduce the two-dimensional process $(X(t), Y(t))^\top$ using the reflection map described below. We write

$$X(t) = W_1(t) - u_1 t - \inf_{-\infty < s \leq t} \{W_1(s) - u_1 s\},$$

$$Y(t) = W_1(t) - u_1 t + \sigma W_2(t) - u_2 t - \inf_{-\infty < s \leq r \leq t} \{W_1(s) - u_1 s + \sigma W_2(r) - u_2 r\},$$

for all $-\infty < t < +\infty$. Using the fact that $\lim_{|t| \to \infty} |W_i(t)|/|t| = 0$ a.s., it is easy to check that $X(t)$ and $Y(t)$ are finite for every $-\infty < t < +\infty$.

We intend to show that $(X(t), Y(t)) \stackrel{\mathrm{D}}{=} (X(0), Y(0))$ for all $t \geq 0$. Fix $t > 0$, and note that we can write

$$X(t) = \sup_{-\infty < s \leq t} [W_1(t) - W_1(s) - u_1(t-s)], \tag{4.17}$$

$$Y(t) = \sup_{-\infty < s \leq r \leq t} [W_1(t) - W_1(s) - u_1(t-s) + \sigma(W_2(t) - W_2(r)) - u_2(t-r)]. \tag{4.18}$$

For $i = 1, 2$, let $B_i(s) = W_i(t) - W_i(t-s)$ for all $s \geq 0$. Then it is straightforward to check that B_1 and B_2 are also one-dimensional FBMs with the same correlation coefficient ρ. It is important to note that B_1 and B_2 depend on t by their definitions. Substituting B_1 and B_2 into (4.17) and (4.18), and then using the time substitutions $\tilde{s} = t - s \geq 0$ and $\tilde{r} = t - r \geq 0$, we obtain

$$X(t) = \max_{0 \leq \tilde{s} \leq t} (B_1(\tilde{s}) - u_1 \tilde{s}) \quad \text{and} \quad Y(t) = \max_{0 \leq \tilde{r} \leq \tilde{s} \leq t} [(B_1(\tilde{s}) - u_1 \tilde{s}) + (\sigma B_2(\tilde{r}) - u_2 \tilde{r})].$$

We observe that $Y(t) \geq X(t)$ for all t, by letting $\tilde{r} = 0$. We recall that

$$\{(B_1(s), B_2(s)) \colon s \geq 0\} \stackrel{\mathrm{D}}{=} \{(W_1(s), W_2(s)) \colon s \geq 0\},$$

and, therefore, we conclude that

$$\begin{pmatrix} X(t) \\ Y(t) \end{pmatrix} \stackrel{\mathrm{D}}{=} \begin{pmatrix} \sup_{0 \leq s < \infty} (W_1(s) - u_1 s) \\ \sup_{0 \leq r \leq s < \infty} [(W_1(s) - u_1 s) + (\sigma W_2(r) - u_2 r)] \end{pmatrix}. \tag{4.19}$$

Note that the right-hand side of (4.19) is independent of t and, hence, $(X(t), Y(t))$ is a stationary process. In particular, $(X(t), Y(t)) \stackrel{\mathrm{D}}{=} (Z_1^0(\infty), Z_1^0(\infty) + Z_2^0(\infty))$ for all $t \geq 0$, where $Z_1^0(\infty)$ and $Z_2^0(\infty)$ are given in (4.5) and (4.6). Next, we define

$$\begin{pmatrix} Z_1^*(t) \\ Z_2^*(t) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} X(t) \\ Y(t) \end{pmatrix} \quad \text{for all } t \geq 0.$$

Then, clearly, $\boldsymbol{Z}^* = (Z_1^*, Z_2^*)^\top$ is also a stationary process. Since $Y(t) \geq X(t) \geq 0$ for all t, we have $Z_1^*(t) \geq 0$ and $Z_2^*(t) \geq 0$ for all $t \geq 0$. We let

$$\widetilde{L}_1^*(t) = \sup_{-\infty < s \leq t} (u_1 s - W_1(s)),$$

$$\widetilde{L}_2^*(t) = \sup_{-\infty < s \leq r \leq t} (u_1 s - W_1(s) + u_2 r - \sigma W_2(r)).$$

Note that $Z_1^*(0) = \widetilde{L}_1^*(0)$ and $Z_2^*(0) = \widetilde{L}_2^*(0) - \widetilde{L}_1^*(0)$. Define the processes $L_1^*(\cdot)$ and $L_2^*(\cdot)$ by

$$L_1^*(t) = \max\Big\{0, \max_{s \in [0,t]} (u_1 s - W_1(s) - Z_1^*(0))\Big\},$$

$$L_2^*(t) = \max\Big\{0, \max_{s \in [0,t]} (u_2 s - \sigma W_2(s) + L_1^*(s) - Z_2^*(0))\Big\},$$

for all $t \geq 0$. Then, clearly,

$$L_1^*(t) = \widetilde{L}_1^*(t) - Z_1^*(0) \quad \text{and} \quad L_2^*(t) = \widetilde{L}_2^*(t) - Z_1^*(0) - Z_2^*(0)$$

hold for all $t \geq 0$. Now, it is straightforward to check that the above defined processes $(Z_i^*(t), L_i^*(t))$ for $i = 1, 2$ satisfy (4.14)–(4.16).

https://doi.org/10.1239/aap/1316792672 Published online by Cambridge University Press

Let $g(x) = \mathrm{e}^{x^\alpha}$, where $0 < \alpha < 2(1-H)$. To show that $\mathrm{E}[g(\boldsymbol{Z}^*(t))] < \infty$, we observe that $|\boldsymbol{Z}^*(t)| \le Z_1^*(t) + Z_2^*(t) = Y(t)$ for all $t \ge 0$. Using (4.6) and (4.19), $Y(t) \stackrel{\mathrm{D}}{=} Z_1^0(\infty) + Z_2^0(\infty)$ for all $t \ge 0$. When $u_1 > 0$ and $u_2 > 0$, we can employ the tail distribution bounds (4.7) and (4.8) in Theorem 4.1 to conclude that $\mathrm{E}[g(\boldsymbol{Z}^*(t))] < \infty$. If $u_1 > u_1 + u_2 > 0$, we can use (4.7) and the tail estimate in Corollary 4.1 to obtain $\mathrm{E}[g(\boldsymbol{Z}^*(t))] < \infty$. Hence, as a consequence, $\mathrm{E}[Z_1^0(\infty) + Z_2^0(\infty)]^N < \infty$ for each $N \ge 1$. This completes the proof.

Remark 4.1. Consider the two-sided filtration $(\mathcal{F}_t : -\infty < t < \infty)$ defined by

$$\mathcal{F}_t = \sigma(\{\boldsymbol{W}_H(s) : -\infty < s \le t\})$$

for each $-\infty < t < \infty$ and allow each $\mathcal{F}_t$ to have all the null sets. Here $\boldsymbol{W}_H$ is the two-sided, two-dimensional FBM introduced in the above proof. Then it is evident that the stationary process $\boldsymbol{Z}^*$ is adapted to the filtration $(\mathcal{F}_t)_{t\ge 0}$.

5. A coupling time result

Consider the probability space $(\Omega, \mathcal{F}, \mathrm{P})$ described in Theorem 4.2. Let $(\mathcal{F}_t)_{t\ge 0}$ be the filtration described in Remark 4.1. Then $\boldsymbol{W}_H = (W_1, W_2)^\top$ is adapted to $(\mathcal{F}_t)_{t\ge 0}$ and $\boldsymbol{Z}^*(0)$ is $\mathcal{F}_0$-measurable, where $\boldsymbol{Z}^*$ is the stationary process defined on $(\Omega, \mathcal{F}, \mathrm{P})$. Henceforth, all our processes are defined on $(\Omega, \mathcal{F}, \mathrm{P})$ and adapted to the filtration $(\mathcal{F}_t)_{t\ge 0}$. The one-dimensional FBMs W_1 and W_2 are correlated with a constant correlation coefficient ρ. Next, recall that our model $\boldsymbol{Z} = (Z_1, Z_2)^\top$ is described by

$$Z_1(t) = Z_1(0) + W_1(t) - u_1 t + L_1(t) \tag{5.1}$$

and

$$Z_2(t) = Z_2(0) + \sigma W_2(t) - u_2 t - L_1(t) + L_2(t) \quad \text{for } t \ge 0. \tag{5.2}$$

Here $\sigma > 0$ is a constant, and $Z_1(0)$ and $Z_2(0)$ are nonnegative, $\mathcal{F}_0$-random variables which satisfy the condition

$$\mathrm{E}[Z_1(0) + Z_2(0)] < \infty. \tag{5.3}$$

Recall that the nondecreasing processes $L_1(\cdot)$ and $L_2(\cdot)$ are given by

$$L_1(t) = \max\Big\{0, \max_{s\in[0,t]} (u_1 s - W_1(s) - Z_1(0))\Big\}, \tag{5.4}$$

$$L_2(t) = \max\Big\{0, \max_{s\in[0,t]} (u_2 s - \sigma W_2(s) + L_1(s) - Z_2(0))\Big\}, \tag{5.5}$$

for all $t \ge 0$. Also, we define the processes $\widetilde{L}_1(\cdot)$ and $\widetilde{L}_2(\cdot)$ by

$$\widetilde{L}_1(t) = Z_1(0) + L_1(t) = \max\Big\{Z_1(0), \max_{s\in[0,t]} (u_1 s - W_1(s))\Big\}, \tag{5.6}$$

$$\begin{aligned}\widetilde{L}_2(t) &= Z_1(0) + Z_2(0) + L_2(t)\\ &= \max\Big\{Z_1(0) + Z_2(0), \max_{s\in[0,t]} (u_2 s - \sigma W_2(s) + \widetilde{L}_1(s))\Big\}.\end{aligned} \tag{5.7}$$

Then (5.1) and (5.2) can be written as

$$\begin{aligned}Z_1(t) &= W_1(t) - u_1 t + \widetilde{L}_1(t),\\ Z_2(t) &= \sigma W_2(t) - u_2 t - \widetilde{L}_1(t) + \widetilde{L}_2(t) \quad \text{for all } t \ge 0.\end{aligned}$$

It is evident that $\widetilde{L}_1(0) = Z_1(0)$, $\widetilde{L}_2(0) = Z_1(0) + Z_2(0)$, $\widetilde{L}_j(\cdot)$ is nondecreasing with continuous paths, and $\int_0^\infty Z_j(t)\,\mathrm{d}\widetilde{L}_j(t) = 0$ for $j = 1, 2$.

Our aim in this section is to show the existence of a stopping time $\tau \geq 0$ such that $\mathbf{Z}(t) = \mathbf{Z}^0(t)$ for all $t \geq \tau$ and $\mathrm{E}[\tau] < \infty$. Here, $\{\mathbf{Z}^0(t)\}_{t\geq 0}$ is the process described in (4.1) and (4.2). Furthermore, we show that if $\mathrm{E}[Z_1(0) + Z_2(0)]^N < \infty$ for some $N \geq 1$ then $\mathrm{E}[\tau^N] < \infty$. From these results, it also follows that the stationary process $\mathbf{Z}^*$ in (4.14) and (4.15) is unique in law. Our first lemma is a variant of Proposition 4.1 of [10], and the difference here is that we allow $Z_1(0)$ to be a random variable.

Lemma 5.1. *Assume that (5.3) holds. Let the processes L_1^0 and $\widetilde{L}_1$ be as in (4.3) and (5.6), respectively. Then there is a stopping time τ_1 such that $\widetilde{L}_1(t) = L_1^0(t)$ for all $t \geq \tau_1$ and* $\mathrm{E}[\tau_1] < \infty$. *In addition, if we assume that* $\mathrm{E}[Z_1(0)]^N < \infty$ *for some $N \geq 1$ then* $\mathrm{E}[\tau_1^N] < \infty$.

Proof. We begin by introducing the stopping time τ_1 and then showing that $\mathrm{E}[\tau_1^N] < \infty$ if $\mathrm{E}[Z_1(0)]^N < \infty$ for some $N \geq 1$. This establishes both parts of the lemma. Let

$$\tau_1 = \inf\{t \geq 0\colon L_1^0(t) \geq Z_1(0)\}, \tag{5.8}$$

where the infimum over an empty set is defined to be ∞. Note that

$$\mathrm{E}[\tau_1^N] = N\int_0^\infty t^{N-1}\,\mathrm{P}[\tau_1 > t]\,\mathrm{d}t = N\int_0^\infty t^{N-1}\,\mathrm{P}[L_1^0(t) < Z_1(0)]\,\mathrm{d}t, \tag{5.9}$$

and

$$\begin{aligned}\mathrm{P}[L_1^0(t) < Z_1(0)] &= \mathrm{P}\Big[\max_{s\in[0,t]}(u_1 s - W_1(s)) < Z_1(0)\Big] &(5.10)\\ &\leq \mathrm{P}[W_1(t) + Z_1(0) > u_1 t]\\ &\leq \mathrm{P}\Big[W_1(t) > \frac{u_1}{2}t\Big] + \mathrm{P}\Big[Z_1(0) > \frac{u_1}{2}t\Big]. &(5.11)\end{aligned}$$

For $t > 0$, $\mathcal{Z} \equiv W_1(t)/t^H$ is a standard normal random variable. For $y > 0$, it is known that

$$\mathrm{P}[\mathcal{Z} > y] \leq \frac{1}{\sqrt{2\pi}}\frac{1}{y}\,\mathrm{e}^{-y^2/2}$$

and, hence, for $t > 0$, we have

$$\mathrm{P}\Big[W_1(t) > \frac{u_1}{2}t\Big] = \mathrm{P}\Big[\mathcal{Z} > \frac{u_1}{2}t^{1-H}\Big] \leq \frac{1}{\sqrt{2\pi}}\frac{2}{u_1 t^{1-H}}\exp\Big(-\frac{u_1^2}{8}t^{2(1-H)}\Big). \tag{5.12}$$

Using (5.9)–(5.12), we have

$$\begin{aligned}\mathrm{E}[\tau_1^N] &\leq 1 + N\int_1^\infty t^{N-1}\,\mathrm{P}[L_1^0(t) < Z_1(0)]\,\mathrm{d}t\\ &\leq 1 + N\int_1^\infty t^{N-1}\Big(\mathrm{P}\Big[W_1(t) > \frac{u_1}{2}t\Big] + \mathrm{P}\Big[\frac{2}{u_1}Z_1(0) > t\Big]\Big)\,\mathrm{d}t\\ &\leq 1 + \frac{2}{u_1}\frac{N}{\sqrt{2\pi}}\int_1^\infty t^{N+H-2}\exp\Big(-\frac{u_1^2}{8}t^{2(1-H)}\Big)\,\mathrm{d}t + \Big(\frac{2}{u_1}\Big)^N\mathrm{E}[Z_1(0)]^N.\end{aligned}$$

The above integral is finite since $H < 1$ and, hence, $\mathrm{E}[\tau_1^N] < \infty$ by the assumption that $\mathrm{E}[Z_1(0)]^N < \infty$. This completes the proof.

https://doi.org/10.1239/aap/1316792672 Published online by Cambridge University Press

In the next lemma, we consider any stopping time τ with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$.

Lemma 5.2. *Let $u_i > 0$ for $i = 1, 2$, and let τ be any stopping time such that* $\mathrm{E}[\tau^N] < \infty$ *for some $N \geq 1$. Then*

$$\mathrm{E}\Big[\max_{s\in[0,\tau]} |u_1 s - W_1(s)|^N\Big] + \mathrm{E}\Big[\max_{s\in[0,\tau]} |u_2 s - \sigma W_2(s)|^N\Big] \leq C_N[1 + \mathrm{E}[\tau^N]], \tag{5.13}$$

where $C_N > 0$ is a constant which depends only on u_1, u_2, and N.

Proof. For simplicity, we only show that $\mathrm{E}[\max_{s\in[0,\tau]} |u_2 s - \sigma W_2(s)|^N] \leq K_N[1+\mathrm{E}[\tau^N]]$, where $K_N > 0$ is a constant that depends only on u_2 and N. An estimate for $\mathrm{E}[\max_{s\in[0,\tau]} |u_1 s - W_1(s)|^N]$ can be obtained along the same lines of the following proof. We begin with the fact that

$$\max_{s\in[0,\tau]} |u_2 s - \sigma W_2(s)| \leq u_2\tau + \sigma \max_{s\in[0,\tau]} |W_2(s)| \quad \text{a.s.}$$

Hence, for each $N \geq 1$,

$$\mathrm{E}\Big[\max_{s\in[0,\tau_1]} |u_2 s - \sigma W_2(s)|^N\Big] \leq 2^{N-1}\Big[u_2^N\,\mathrm{E}[\tau^N] + \sigma^N\,\mathrm{E}\Big[\max_{s\in[0,\tau]} |W_2(s)|^N\Big]\Big]. \tag{5.14}$$

For $0 < H < 1$, it is known from Corollary 3.1 of [34] (see also Theorem 1.2 of [27], Exercise 5.1.5 of [28], and the recent article [17] for the analysis of a related martingale) that

$$\mathrm{E}\Big[\max_{s\in[0,\tau]} |W_2(s)|^N\Big] \leq C_{N,H}\,\mathrm{E}[\tau^{NH}], \tag{5.15}$$

where $C_{N,H} > 0$ is a constant that depends only on N and H. Since $0 < NH < N$, $\mathrm{E}[\tau^{NH}] \leq 1 + \mathrm{E}[\tau^N]$. Now combining this estimate with (5.14) and (5.15) yields the desired result.

Lemma 5.3. *Assume that* $\mathrm{E}[Z_1(0)+Z_2(0)]^N < \infty$ *for some $N \geq 1$ and that A is an $\mathcal{F}_0$-random variable such that* $\mathrm{E}[A^N] < \infty$ *for some $N \geq 1$. Let τ_1 be the stopping time defined in (5.8). Define $M(t) = \max_{s\in[0,t]}(u_2 s - \sigma W_2(s))$ for all $t \geq 0$. Then there exists a stopping time $\tau_2 > \tau_1$ such that*

$$(u_2\tau_2 - \sigma W_2(\tau_2)) \geq A + M(\tau_1) \quad \textit{a.s.} \tag{5.16}$$

and

$$\mathrm{E}[\tau_2^N] < \infty. \tag{5.17}$$

Proof. By Lemma 5.2, $M(\tau_1)$ is finite a.s. We let

$$\tau_2 = \inf\{t \geq 0 \colon u_2 t - \sigma W_2(t) \geq A + M(\tau_1)\}, \tag{5.18}$$

where we set the infimum over an empty set to be ∞. Since $\lim_{t\to\infty}(u_2 t - \sigma W_2(t)) = +\infty$ a.s., τ_2 is also finite a.s. By the definition of $M(\tau_1)$, it clearly follows that $\tau_2 > \tau_1$ a.s. Next, we show that τ_2 is indeed a stopping time. Since $\tau_2 > \tau_1$ a.s. and τ_1 is a stopping time, we have, for fixed $s > 0$, $\{\tau_2 > s\} \cap \{\tau_1 \geq s\} = \{\tau_1 \geq s\} \in \mathcal{F}_s$. On the other hand,

$$\begin{aligned}\{\tau_2 > s\} \cap \{\tau_1 < s\} &= \{M(s) < A + M(\tau_1),\ \tau_1 < s\} \\ &= \{M(s) < A + M(\tau_1)\mathbf{1}_{\{\tau_1 < s\}},\ \tau_1 < s\}.\end{aligned}$$

Observe that $M(\cdot)$ is a nonnegative, continuous, nondecreasing process adapted to $(\mathcal{F}_t)$. Therefore, we find (by standard discrete approximation) that $M(\tau_1)\mathbf{1}_{\{\tau_1<s\}}$ is an $\mathcal{F}_s$-random variable. Recall that A is an $\mathcal{F}_0$-random variable. Hence, $\{\tau_2 > s\} \cap \{\tau_1 < s\} \in \mathcal{F}_s$ and, consequently, the result in (5.16) follows.

To establish (5.17), note that

$$\mathrm{E}[\tau_2^N] = N \int_0^\infty t^{N-1}\,\mathrm{P}[\tau_2 > t]\,\mathrm{d}t \le 1 + N \int_1^\infty t^{N-1}\,\mathrm{P}[\tau_2 > t]\,\mathrm{d}t, \tag{5.19}$$

and, from (5.18),

$$\begin{aligned}\mathrm{P}[\tau_2 > t] &= \mathrm{P}[A + M(\tau_1) + \sigma W_2(t) \ge u_2 t] \\ &\le \mathrm{P}\Big[A \ge \frac{u_2}{3}t\Big] + \mathrm{P}\Big[M(\tau_1) \ge \frac{u_2}{3}t\Big] + \mathrm{P}\Big[W_2(t) \ge \frac{u_2}{3\sigma}t\Big].\end{aligned} \tag{5.20}$$

Since $\mathrm{E}[A^N] < \infty$, we have

$$N \int_0^\infty t^{N-1}\,\mathrm{P}\Big[A \ge \frac{u_2 t}{3}\Big]\,\mathrm{d}t < \infty.$$

The assumption that $\mathrm{E}[Z_1(0) + Z_2(0)]^N < \infty$ implies that $\mathrm{E}[\tau_1^N] < \infty$ by Lemma 5.1. Next, we can employ Lemma 5.2 with the stopping time τ_1 to conclude that $\mathrm{E}[M(\tau_1)]^N < \infty$. Therefore,

$$N \int_0^\infty t^{N-1}\,\mathrm{P}\Big[M(\tau_1) \ge \frac{u_2}{3}t\Big]\,\mathrm{d}t < \infty. \tag{5.21}$$

Finally,

$$\mathrm{P}\Big[W_2(t) \ge \frac{u_2}{3\sigma}t\Big] \le \frac{1}{\sqrt{2\pi}}\,\frac{3\sigma}{u_2 t^{1-H}}\exp\Big(-\frac{u_2^2}{18\sigma^2}t^{2(1-H)}\Big)$$

holds for $t > 0$ as in (5.12) and, hence, we can conclude that

$$N \int_1^\infty t^{N-1}\,\mathrm{P}\Big[W_2(t) \ge \frac{u_2}{3\sigma}t\Big]\,\mathrm{d}t < \infty. \tag{5.22}$$

Combining (5.19)–(5.22) it follows that $\mathrm{E}[\tau_2^N] < \infty$. This completes the proof.

Next, we employ the above three lemmas to prove the following proposition.

Proposition 5.1. *Assume that* $\mathrm{E}[Z_1(0) + Z_2(0)]^N < \infty$ *for some* $N \ge 1$*, and choose* $A = 1 + Z_1(0) + Z_2(0) > 0$ *in Lemma 5.3. Let the stopping times* τ_1 *and* τ_2 *be as defined in (5.8) and (5.18), respectively. Then, for all* $t \ge \tau_2$*,*

(a) $\max_{s\in[\tau_1,t]}(u_2 s - \sigma W_2(s) + L_1^0(s)) \ge \max\{L_2^0(\tau_1), \widetilde{L}_2(\tau_1)\}$ *a.s., and*

(b) $\widetilde{L}_1(t) = L_1^0(t)$, $\widetilde{L}_2(t) = L_2^0(t)$, *and* $\mathrm{E}[{\tau_2}^N] < \infty$.

Proof. By the definition of τ_2 in (5.18) we have

$$u_2\tau_2 - \sigma W_2(\tau_2) \ge A + \max_{s\in[0,\tau_1]}(u_2 s - \sigma W_2(s)) \ge Z_1(0) + Z_2(0) \quad \text{a.s.}$$

Since $\tau_2 \geq \tau_1$, when $t \geq \tau_2$, we obtain

$$\begin{aligned}\max_{s\in[\tau_1,t]}(u_2 s - \sigma W_2(s) + L_1^0(s)) &\geq u_2\tau_2 - \sigma W_2(\tau_2) + L_1^0(\tau_2)\\ &\geq A + \max_{s\in[0,\tau_1]}(u_2 s - \sigma W_2(s) + L_1^0(s))\\ &\geq Z_2(0) + \max_{s\in[0,\tau_1]}(u_2 s - \sigma W_2(s) + \widetilde{L}_1(s)).\end{aligned}$$

Also, we obtain

$$\max_{s\in[\tau_1,t]}(u_2 s - \sigma W_2(s) + L_1^0(s)) \geq u_2\tau_2 - \sigma W_2(\tau_2) \geq Z_1(0) + Z_2(0) \quad \text{a.s.}$$

Hence, it follows that

$$\max_{s\in[\tau_1,t]}(u_2 s - \sigma W_2(s) + L_1^0(s)) \geq \max\{L_2^0(\tau_1), \widetilde{L}_2(\tau_1)\} \quad \text{a.s.}$$

for all $t \geq \tau_2$. Therefore, part (a) follows.

For part (b), by Lemma 5.1 we already know that $\widetilde{L}_1(t) = L_1^0(t)$ for $t \geq \tau_1$. Using this fact, whenever $t \geq \tau_2$, we can write

$$\begin{aligned}\widetilde{L}_2(t) &= \max\Big\{Z_1(0) + Z_2(0),\ \max_{s\in[0,\tau_1]}(u_2 s - \sigma W_2(s) + \widetilde{L}_1(s)),\\ &\qquad \max_{s\in[\tau_1,t]}(u_2 s - \sigma W_2(s) + \widetilde{L}_1(s))\Big\}\\ &= \max\Big\{\widetilde{L}_2(\tau_1),\ \max_{s\in[\tau_1,t]}(u_2 s - \sigma W_2(s) + \widetilde{L}_1(s))\Big\}\\ &= \max_{s\in[\tau_1,t]}(u_2 s - \sigma W_2(s) + \widetilde{L}_1(s)).\end{aligned}$$

Therefore, $\widetilde{L}_2(t) \leq L_2^0(t)$ whenever $t \geq \tau_2$. On the other hand,

$$\begin{aligned}L_2^0(t) &= \max\Big\{L_2^0(\tau_1),\ \max_{s\in[\tau_1,t]}(u_2 s - \sigma W_2(s) + L_1^0(s))\Big\}\\ &= \max_{s\in[\tau_1,t]}(u_2 s - \sigma W_2(s) + L_1^0(s)) \quad \text{for all } t \geq \tau_2.\end{aligned}$$

This yields $\widetilde{L}_2(t) = L_2^0(t)$ for all $t \geq \tau_2$. We have already established $\mathrm{E}[{\tau_2}^N] < \infty$ in Lemma 5.3. This completes the proof.

We now state and prove the main coupling time result in this section.

Theorem 5.1. *Let $\mathbf{Z}$ be a process that satisfies (5.1) and (5.2) with $\mathrm{E}[Z_1(0) + Z_2(0)]^N < \infty$ for some $N \geq 1$. Then the following statements hold.*

(a) *There exists a stopping time τ such that $\mathbf{Z}(t) = \mathbf{Z}^0(t)$ for all $t \geq \tau$ and $\mathrm{E}[\tau^N] < \infty$, where $\mathbf{Z}^0$ is the process that satisfies (4.1) and (4.2).*

(b) *There exists a stopping time $\widehat{\tau}$ such that $\mathbf{Z}(t) = \mathbf{Z}^*(t) = \mathbf{Z}^0(t)$ for all $t \geq \widehat{\tau}$ and $\mathrm{E}[\widehat{\tau}^N] < \infty$, where $\mathbf{Z}^*$ is the stationary process described in Theorem 4.2.*

https://doi.org/10.1239/aap/1316792672 Published online by Cambridge University Press

Proof. Part (a) clearly follows from Proposition 5.1 and the representation of the process $\boldsymbol{Z}$ in (5.1)–(5.7). To establish part (b), note that

$$Z_1^*(0) + Z_2^*(0) \overset{\mathrm{D}}{=} Z_1^0(\infty) + Z_2^0(\infty),$$

where $Z_1^0(\infty)$ and $Z_2^0(\infty)$ are described in (4.5) and (4.6). Using the tail estimates in (4.7) and (4.8), it clearly follows that

$$\mathrm{E}[Z_1^*(0) + Z_2^*(0)]^n < \infty \quad \text{for every } n \geq 1.$$

Consequently, we can apply part (a) and, thus, there is a stopping time $\tau^* > 0$ such that $\boldsymbol{Z}^*(t) = \boldsymbol{Z}^0(t)$ for all $t \geq \tau^*$. Let $\boldsymbol{Z}$ be any other process, which satisfies (5.1) and (5.2) with $\mathrm{E}[Z_1(0) + Z_2(0)]^N < \infty$ for some $N \geq 1$. Then there is a stopping time τ satisfying part (a). Let $\widehat{\tau} = \tau + \tau^*$. Then

$$\boldsymbol{Z}(t) = \boldsymbol{Z}^*(t) = \boldsymbol{Z}^0(t) \quad \text{for all } t \geq \widehat{\tau}.$$

Since $\mathrm{E}[\tau^N] < \infty$ and $\mathrm{E}[(\tau^*)^N] < \infty$, it follows that $\mathrm{E}[\widehat{\tau}^N] < \infty$. This completes the proof.

The following corollary is an immediate consequence of the above theorem.

Corollary 5.1. *Let $\widetilde{\boldsymbol{Z}}^*$ be any other stationary process that satisfies (4.14)–(4.16) with the moment condition* $\mathrm{E}[\widetilde{Z}_1^*(0) + \widetilde{Z}_2^*(0)] < \infty$*. Then there is a stopping time $\widetilde{\tau}$ such that $\widetilde{\boldsymbol{Z}}^*(t) = \boldsymbol{Z}^*(t)$ for all $t \geq \widetilde{\tau}$ and* $\mathrm{E}[\widetilde{\tau}] < \infty$*. Hence, the stationary process $\boldsymbol{Z}^*$ is unique in law.*

6. Cost minimization

In this section we analyze the cost structure described in (2.7) and address the associated cost minimization problem. Our state process $\boldsymbol{Z} = (Z_1, Z_2)^\top$ is adapted to $(\mathcal{F}_t)_{t\geq 0}$, and it satisfies (5.1) and (5.2). We assume that the initial data $\boldsymbol{Z}(0)$ satisfies the moment condition (5.3). For the cost minimization problem with cost functional $I(\boldsymbol{u}, \boldsymbol{Z}(0))$ in (2.7), the running cost function C satisfies assumptions (H1)–(H3) given in Section 2. Henceforth, we say that a state process $\boldsymbol{Z}$ is an *admissible state process* if the initial data $\boldsymbol{Z}(0)$ satisfies the moment condition

$$\mathrm{E}\,|Z_1(0) + Z_2(0)|^{m+1} < \infty, \quad \text{where } m \geq 1 \text{ is as in (H3).} \tag{6.1}$$

To address the cost minimization problem, first we show that the cost functional $I(\boldsymbol{u}, \boldsymbol{Z}(0))$ described in (2.7) is independent of the initial data $\boldsymbol{Z}(0)$, and we obtain a representation for it using the stationary distribution $\boldsymbol{Z}^*$ of Theorem 4.2. This representation will be used to address the cost minimization problem. We begin with the following lemma.

Lemma 6.1. *Let the process $\boldsymbol{Z}$ satisfy (5.1), (5.2), and (5.3). Then*

$$\lim_{T\to\infty} \frac{1}{T}\,\mathrm{E}[L_1(T)] = u_1, \tag{6.2}$$

$$\lim_{T\to\infty} \frac{1}{T}\,\mathrm{E}[L_2(T)] = u_1 + u_2. \tag{6.3}$$

Proof. Since $\mathrm{E}[Z_1(0)] < \infty$ from condition (5.3), conclusion (6.2) can be obtained by following the proof of Lemma 3.1 of [10]. For (6.3), we begin with the definition of the standard one-dimensional reflection mapping (Skorokhod map) $\Gamma\colon C([0,\infty),\mathbb{R}) \to C([0,\infty),\mathbb{R})$, which is defined as

$$\Gamma(f)(t) = f(t) + \max\Big\{0, \max_{s\in[0,t]}(-f(s))\Big\} \quad \text{for } f \in C([0,\infty),\mathbb{R}) \text{ and } t \geq 0.$$

https://doi.org/10.1239/aap/1316792672 Published online by Cambridge University Press

Then we have $Z_2(t) = \Gamma(Z_2(0) + \sigma W_2 - u_2 e - L_1)(t)$, where $e(t) \equiv t$ for all $t \geq 0$. Since $u_2 t + L_1(t)$ is nonnegative and nondecreasing in t, we have

$$Z_2(0) + \sigma W_2(t) - u_2 t - L_1(t) \leq Z_2(0) + \sigma W_2(t).$$

Therefore, from the basic properties of the Skorokhod map (see, for instance, [30, p. 439]), we have

$$0 \leq Z_2(t) \leq \Gamma(Z_2(0) + \sigma W_2)(t) \leq 2\Big(Z_2(0) + \max_{s\in[0,t]} \sigma |W_2(s)|\Big). \tag{6.4}$$

Hence,

$$\begin{aligned} 0 &\leq \frac{\mathrm{E}[Z_2(T)]}{T} \\ &\leq \frac{2}{T}\Big(\mathrm{E}[Z_2(0)] + \mathrm{E}\Big[\max_{s\in[0,T]} \sigma |W_2(s)|\Big]\Big) \\ &\leq \frac{2}{T}(\mathrm{E}[Z_2(0)] + K_1 T^H) \\ &\to 0 \end{aligned}$$

as $T \to \infty$. Here $K_1 \in (0, \infty)$ is a generic constant independent of T (see [28, p. 296]). Since

$$L_2(T) = Z_2(T) - Z_2(0) - \sigma W_2(T) + u_2 T + L_1(T), \qquad \mathrm{E}[W_2(T)] = 0,$$

and using (6.2), we obtain $\lim_{T\to\infty}(1/T)\,\mathrm{E}[L_2(T)] = u_1 + u_2$.

Remark 6.1. The estimate in (6.4) also implies that $\lim_{T\to\infty} Z_2(T)/T = 0$ a.s. A similar estimate in [10] can be used to show that $\lim_{T\to\infty} Z_1(T)/T = 0$ a.s. Consequently, $\lim_{T\to\infty} L_1(T)/T = u_1$ a.s. Using this with (6.4) also leads to $\lim_{T\to\infty} L_2(T)/T = u_1 + u_2$ a.s.

Proposition 6.1. *Let* $\boldsymbol{Z}$ *be an admissible state process that satisfies (5.1), (5.2), and (6.1). Then*

$$\lim_{T\to\infty} \frac{1}{T}\,\mathrm{E}\int_0^T C(Z_1(t), Z_2(t))\,\mathrm{d}t = \mathrm{E}[C(Z_1^0(\infty), Z_2^0(\infty))]. \tag{6.5}$$

Proof. Since $\mathrm{E}|Z_1(0) + Z_2(0)|^{m+1} < \infty$, we can use Theorem 5.1 to conclude that there exists a stopping time $\widehat{\tau}$ such that $\boldsymbol{Z}(t) = \boldsymbol{Z}^*(t)$ for all $t \geq \widehat{\tau}$ and $\mathrm{E}[\widehat{\tau}^{m+1}] < \infty$. Furthermore, $\boldsymbol{Z}^*$ is a stationary process and $(Z_1^*(t), Z_2^*(t)) \overset{\mathrm{D}}{=} (Z_1^0(\infty), Z_2^0(\infty))$ for all $t \geq 0$. Therefore, to establish (6.5), it suffices to show that

$$\lim_{T\to\infty} \frac{1}{T}\,\mathrm{E}\int_0^{\widehat{\tau}} C(Z_1(t), Z_2(t))\,\mathrm{d}t = 0. \tag{6.6}$$

This will follow if $\mathrm{E}\int_0^{\widehat{\tau}} C(Z_1(t), Z_2(t))\,\mathrm{d}t < \infty$, and, hence, we establish this fact in the argument below.

Using (H3), there is an $m \geq 1$ such that $0 \leq C(x, y) \leq K(1 + |x + y|^m)$ for all $x \geq 0$ and $y \geq 0$, where $K > 0$ is a constant. Therefore,

$$\mathrm{E}\int_0^{\widehat{\tau}} C(Z_1(t), Z_2(t))\,\mathrm{d}t \leq K\,\mathrm{E}\Big[\Big(1 + \max_{t\in[0,\widehat{\tau}]} |Z_1(t) + Z_2(t)|^m\Big)\widehat{\tau}\Big].$$

We have, by Theorem 5.1, $\mathrm{E}[\widehat{\tau}^{m+1}] < \infty$ and, thus, the right-hand side of the above inequality is finite, if we establish that

$$\mathrm{E}\Big[\Big(\max_{t\in[0,\widehat{\tau}]} |Z_1(t) + Z_2(t)|^m\Big)\widehat{\tau}\Big] < \infty. \tag{6.7}$$

But, using Hölder's inequality with $p = (m+1)/m$ and $q = m+1$, we have

$$\mathrm{E}\Big[\Big(\max_{t\in[0,\widehat{\tau}]} |Z_1(t) + Z_2(t)|^m\Big)\widehat{\tau}\Big] \leq \Big(\mathrm{E}\Big[\max_{t\in[0,\widehat{\tau}]} |Z_1(t) + Z_2(t)|^{m+1}\Big]\Big)^{m/(m+1)} (\mathrm{E}[\widehat{\tau}^{m+1}])^{1/(m+1)}.$$

Since $\mathrm{E}[\widehat{\tau}^{m+1}] < \infty$, it remains to establish that $\mathrm{E}(\max_{t\in[0,\widehat{\tau}]} |Z_1(t) + Z_2(t)|^{m+1}) < \infty$ to guarantee (6.7). Using (5.1)–(5.5), we obtain

$$\begin{aligned} \mathrm{E}\Big[&\max_{t\in[0,\widehat{\tau}]} |Z_1(t) + Z_2(t)|^{m+1}\Big] \\ &\leq K_2\Big(\mathrm{E}|Z_1(0) + Z_2(0)|^{m+1} + \mathrm{E}\Big[\max_{t\in[0,\widehat{\tau}]} |u_2 s - \sigma W_2(s)|^{m+1}\Big] \\ &\qquad + \mathrm{E}\Big[\max_{t\in[0,\widehat{\tau}]} |u_1 s - W_1(s)|^{m+1}\Big]\Big), \end{aligned} \tag{6.8}$$

where $K_2 > 0$ is a generic constant which may depend on m. Following the proof of Lemma 5.2 and using the fact that $\mathrm{E}[\widehat{\tau}^{m+1}] < \infty$, we have

$$\mathrm{E}\Big[\max_{t\in[0,\widehat{\tau}]} |u_2 s - \sigma W_2(s)|^{m+1}\Big] < \infty \quad \text{and} \quad \mathrm{E}\Big[\max_{t\in[0,\widehat{\tau}]} |u_1 s - W_1(s)|^{m+1}\Big] < \infty.$$

Therefore, we can conclude that the left-hand side of (6.8) is finite and this establishes (6.7). Hence, (6.6) follows. To complete the proof, we should check that $\mathrm{E}[C(Z_1^0(\infty), Z_2^0(\infty))]$ is finite. But, this directly follows from the tail distribution asymptotics of $\mathbf{Z}^0(\infty)$ described in (4.7) and (4.8). This completes the proof.

Let us introduce $F(u_1, u_2) \equiv \mathrm{E}[C(Z_1^0(\infty), Z_2^0(\infty))]$ for all $u_1 > 0$ and $u_2 \geq 0$. We intend to establish the continuity of F on the domain

$$\mathcal{D} \equiv \{(u_1, u_2) \colon u_1 > 0 \text{ and } u_2 \geq 0\}.$$

To help the arguments in the next proposition, we introduce the following notation. Let the random variables $G(u)$ and $H(u, v)$ be defined by

$$G(u) = \sup_{0\leq s}(W_1(s) - us), \qquad H(u, v) = \sup_{0\leq r\leq s} [(W_1(s) - us) + (\sigma W_2(r) - vr)]$$

for all $u > 0$ and $v \geq 0$. By Theorem 4.1, $Z_1^0(\infty) = G(u_1)$ and $Z_1^0(\infty) + Z_2^0(\infty) = H(u_1, u_2)$. Also, note that $H(u, v)$ is finite if $u > 0$ and $u + v \geq 0$, as noted in Theorem 4.1(a).

We also define the function $\widehat{C}$ on the set $\{(x, y) \colon y \geq x \geq 0\}$ by

$$\widehat{C}(x, y) = C(x, y - x). \tag{6.9}$$

Hence, it follows that

$$F(u, v) = \mathrm{E}[\widehat{C}(G(u), H(u, v))] \tag{6.10}$$

for all $u > 0$ and $v \geq 0$. For each $u > 0$ and $v \geq 0$, $F(u, v)$ is finite. If $v > 0$, this follows

https://doi.org/10.1239/aap/1316792672 Published online by Cambridge University Press

from the fact that

$$H(u,v) \le G(u) + \sup_{0\le r}(\sigma W_2(r) - vr) < \infty,$$

the polynomial growth condition of C in (H3), and the tail estimates (4.7) and (4.8). If $v = 0$, a very similar argument using the estimate in the proof of Theorem 4.1(a) guarantees the finiteness of $F(u, v)$.

Proposition 6.2. *Under the assumptions of Proposition 6.1, the following statements hold.*

(a) *The function $F(u, v)$ is continuous on the domain $\mathcal{D} = \{(u, v)\colon u > 0$ and $v \ge 0\}$.*

(b) *The function $F(u, v)$ is decreasing in the variable v and $\lim_{(u,v)\to(0,b)} F(u, v) = \infty$ for each $b > 0$.*

Proof. Let $(a, b) \in \mathcal{D}$. We pick $\delta > 0$ such that $0 < 3\delta < a$. To show the continuity of F at (a, b), we pick a sequence (a_n, b_n) which converges to (a, b) as $n \to \infty$. Without loss of generality, we assume that $a_n > 3\delta$ for all n. Our first step is to show that $G(a_n) \to G(a)$ a.s. and $H(a_n, b_n) \to H(a, b)$ a.s. as n tends to ∞. Since $\lim_{t\to\infty} W^{(i)}(t)/t = 0$ a.s. for $i = 1, 2$, there exists a $T_1(\omega) > 0$ such that $\max\{W_1(t) - \delta t, \sigma W_2(t) - \delta t\} < 0$ for all $t \ge T_1(\omega)$. We let

$$T_0(\omega) = \max\left\{T_1(\omega), \frac{1}{\delta} \max_{t\in[0,T_1(\omega)]}(\sigma W_2(t) - \delta t)\right\},$$

and $T_0(\omega) > 0$ is finite. Then clearly it follows that

$$G(a_n) = \max_{0\le s\le T_0(\omega)}(W_1(s) - a_n s) \quad \text{and} \quad G(a) = \max_{0\le s\le T_0(\omega)}(W_1(s) - as).$$

From this, it is evident that $|G(a_n) - G(a)| \le T_0(\omega)|a_n - a|$ and, thus, $G(a_n) \to G(a)$ as n tends to ∞.

Next, we consider

$$H(a_n, b_n) = \sup_{0\le r\le s}[(W_1(s) - a_n s) + (\sigma W_2(r) - b_n r)].$$

Since $a_n > 3\delta > 0$ and $b_n \ge 0$, we obtain the following estimates. For any $s > T_0(\omega)$ and $r \le s$,

$$\begin{aligned}(W_1(s) - a_n s) + (\sigma W_2(r) - b_n r) &\le W_1(s) - 3\delta s + \sigma W_2(r)\\ &\le (W_1(s) - \delta s) - \delta s + (\sigma W_2(r) - \delta r)\\ &\le -\delta s + \max_{r\in[0,T_1(\omega)]}(\sigma W_2(r) - \delta r)\\ &< 0.\end{aligned}$$

Therefore, we can write

$$H(a_n, b_n) = \max_{0\le r\le s\le T_0(\omega)}[(W_1(s) - a_n s) + (\sigma W_2(r) - b_n r)]$$

and, similarly,

$$H(a, b) = \max_{0\le r\le s\le T_0(\omega)}[(W_1(s) - as) + (\sigma W_2(r) - br)].$$

Hence, we obtain

$$|H(a_n, b_n) - H(a, b)| \le T_0(\omega)(|a_n - a| + |b_n - b|).$$

Since $T_0(\omega) > 0$ is finite, we have $H(a_n, b_n) \to H(a, b)$ as $n \to \infty$.

Our second step is to obtain an integrable upper bound for $C(G(a_n), H(a_n, b_n) - G(a_n))$. Since $C(x, y)$ is nondecreasing in each variable, we have

$$0 \le C(G(a_n), H(a_n, b_n) - G(a_n)) \le C(G(a_n), H(a_n, b_n)).$$

Since $a_n > 3\delta > 0$ and $b_n \ge 0$, we also have

$$0 \le G(a_n) \le G(\delta) \quad \text{and} \quad 0 \le H(a_n, b_n) \le G(\delta) + \sup_{0\le r}(\sigma W_2(r) - \delta r).$$

Therefore,

$$0 \le C(G(a_n), H(a_n, b_n) - G(a_n)) \le C(G(\delta), G(\delta) + \widetilde{G}(\delta)), \tag{6.11}$$

where $\widetilde{G}(\delta) = \sup_{0\le r}(\sigma W_2(r) - \delta r)$. Using the tail estimates (4.7) and (4.8), and the polynomial growth condition of C in (H3), it follows that

$$\mathrm{E}[C(G(\delta), G(\delta) + \widetilde{G}(\delta))] < \infty. \tag{6.12}$$

In our third step, we apply the dominated convergence theorem to establish the continuity of F at (a, b). Since $C(x, y)$ is continuous, using our first step above, we have

$$C(G(a_n), H(a_n, b_n) - G(a_n)) \to C(G(a), H(a, b) - G(a))$$

a.s. as $n \to \infty$. Next, using (6.11), (6.12), and the aforementioned almost-sure convergence together with the dominated convergence theorem, we conclude that

$$F(a_n, b_n) = \mathrm{E}[C(G(a_n), H(a_n, b_n) - G(a_n))] \to F(a, b) = \mathrm{E}[C(G(a), H(a, b) - G(a))]$$

as $n \to \infty$. This completes the proof of part (a).

For part (b), observe that if $b_1 < b_2$ then $H(a, b_2) < H(a, b_1)$. Since $C(x, y)$ is increasing in the variable y, we obtain $F(a, b_2) \le F(a, b_1)$, whenever $b_2 > b_1$. Finally, we intend to compute the limit $\lim_{(u,v)\to(0,b)} F(u, v)$ for $b > 0$. Let $b > 0$ and $(a_n, b_n) \to (0, b)$ as $n \to \infty$. We can simply assume that a_n is decreasing to 0 as n tends to ∞. Hence, $G(a_n)$ is increasing to ∞. Next, using assumption (H2) for the cost function C, we can conclude that

$$\lim_{(a_n,b_n)\to(0,b)} F(a_n, b_n) = \infty.$$

This completes the proof.

Remark 6.2. Using (6.9) and (6.10), it can be shown that $F(u, v)$ is decreasing in both variables u and v under the assumption that $\partial C(x, y)/\partial x \ge \partial C(x, y)/\partial y$ for all x and y.

Our next theorem is the main result of this section.

Theorem 6.1. *Let $\mathbf{Z}$ be an admissible process that satisfies (5.1) and (5.2) with control $\boldsymbol{u} = (u_1, u_2)$ in $\mathcal{D}$. Then, the following results hold.*

(a) *The cost functional $I(\boldsymbol{u}, \mathbf{Z}(0))$ described in (2.7) is independent of $\mathbf{Z}(0)$ and has the representation*

$$I(u_1, u_2) \equiv I(\boldsymbol{u}, \mathbf{Z}(0)) = h(u_1, u_2) + p_1 u_1 + p_2(u_1 + u_2) + F(u_1, u_2), \tag{6.13}$$

where $F(u_1, u_2)$ is as defined in (6.10).

https://doi.org/10.1239/aap/1316792672 Published online by Cambridge University Press

(b) *There exists an optimal control* $\boldsymbol{u}^* = (u_1^*, u_2^*)$ *in* $\mathcal{D}$ *such that*

$$I(u_1^*, u_2^*) = \inf_{\boldsymbol{u}\in\mathcal{D}} I(u_1, u_2).$$

Moreover, the process $\boldsymbol{Z}^*$ *defined as in Theorem 4.2 with control* $\boldsymbol{u}^* = (u_1^*, u_2^*)$ *is an optimal stationary process.*

Proof. Part (a) follows directly from the definition of $I(\boldsymbol{u}, \boldsymbol{Z}(0))$ in (2.7), and the results obtained in Lemma 6.1, Proposition 6.1, and Proposition 6.2. For part (b), with representation (6.13) for $I(u_1, u_2)$ in hand, we have $I(u_1, u_2)$ is finite and continuous on the domain $\mathcal{D} \equiv \{(u_1, u_2)\colon u_1 > 0,\ u_2 \geq 0\}$. Also, from Proposition 6.2 and representation (6.13), we have

$$\lim_{u_1+u_2\to\infty} I(u_1, u_2) = +\infty, \qquad \lim_{(u_1,u_2)\to(a,0)} I(u_1, u_2) = I(a, 0),$$
$$\lim_{(u_1,u_2)\to(0,b)} I(u_1, u_2) = +\infty,$$

for any $a > 0$ and $b > 0$. In the following argument we consider the stationary state process $\boldsymbol{Z}_{\boldsymbol{u}}^*$ associated with control $\boldsymbol{u} \in \mathcal{D}$ as described in Theorem 4.2. Then it automatically satisfies the assumed moment condition for the initial data since the tail estimates (4.7) and (4.8) imply the finiteness of all the moments of $|\boldsymbol{Z}_{\boldsymbol{u}}^*(0)|$. Consequently, $\boldsymbol{Z}_{\boldsymbol{u}}^*$ is an admissible state process. Any state process with nonrandom initial data also satisfies the moment condition for initial data and is hence admissible.

Next, consider a control (a_0, b_0) in $\mathcal{D}$ with $a_0 > 0$ and $b_0 > 0$. We keep (a_0, b_0) fixed. Let $M \equiv I(a_0, b_0)$, which is finite, and define the set $\mathcal{D}_0 \subset \mathcal{D}$ by

$$\mathcal{D}_0 \equiv \{(u_1, u_2) \in \mathcal{D}\colon I(u_1, u_2) \leq M\}.$$

Then $\inf_{\mathcal{D}} I(u_1, u_2) = \inf_{\mathcal{D}_0} I(u_1, u_2)$. With the above described limits and properties of $I(u_1, u_2)$, it clearly follows that $\mathcal{D}_0$ is a bounded set. Now let $\{(a_n, b_n)\colon n \geq 1\}$ be a sequence in $\mathcal{D}_0$ such that $I(a_n, b_n) \to \inf_{\mathcal{D}_0} I(u_1, u_2)$ as $n \to \infty$. Thus, $\{(a_n, b_n)\colon n \geq 1\}$ has a convergent subsequence. Therefore, we simply assume that $(a_n, b_n) \to (u_1^*, u_2^*)$ as n tends to ∞. Hence, $u_1^* \geq 0$ and $u_2^* \geq 0$. Clearly, $u_1^* > 0$ since $\lim_{(u_1,u_2)\to(0,b)} I(u_1, u_2) = +\infty$. Therefore, there exists (u_1^*, u_2^*) in $\mathcal{D}$ such that $I(u_1^*, u_2^*) = \inf_{\mathcal{D}} I(u_1, u_2)$. This completes the proof.

7. Concluding remarks

Our methods can be readily extended to the case of a tandem queueing network with n stations, where $n \geq 2$. Let $W_1, \ldots, W_n$ be possibly correlated one-dimensional FBMs with constant correlation coefficients. Then we can represent the n-dimensional state process $\{\boldsymbol{Z}(t) = (Z_1(t), \ldots, Z_n(t))^\top\}_{t\geq 0}$, where $\boldsymbol{Z}(t) \in \mathbb{R}_+^n$, by

$$Z_i(t) = Z_i(0) + \sigma_i W_i - u_i t - L_{i-1}(t) + L_i(t)$$

for $i = 1, \ldots, n$. Here $\sigma_i > 0$ and $u_i > 0$ are constants (with $\sigma_1 \equiv 1$), and the constant $u_i > 0$ represents the controllable drift rate at the ith station. Also, $L_0(t) \equiv 0$ for all $t \geq 0$ and the local time process L_i (corresponding to Z_i) is a continuous, nondecreasing process which increases only when $Z_i(t) = 0$. That is, $\int_0^\infty Z_i(t)\, \mathrm{d}L_i(t) = 0$ a.s. and $Z_i(t) \geq 0$ for all $t \geq 0$. In this situation we can obtain a stationary state process $\boldsymbol{Z}^*$ and conclude the existence of an optimal control vector $\boldsymbol{u}^* = (u_1^*, \dots, u_n^*)^\top$ following our methods in the previous sections. Moreover, the distribution of $\boldsymbol{Z}^*(t)$ can be explicitly described as follows. Let

$$\xi_i = \sup_{0 \le s_1 \le \dots \le s_i} \left(\sum_{j=1}^{i} \sigma_j W_j(s_j) - u_j s_j \right) \quad \text{for } i = 1, \dots, n.$$

Then, for all $t \ge 0$,

$$\boldsymbol{Z}^*(t) \stackrel{\mathrm{D}}{=} \boldsymbol{R}(\xi_1, \dots, \xi_n)^\top,$$

where

$$\boldsymbol{R} = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ -1 & 1 & \cdots & 0 & 0 \\ 0 & -1 & 1 & \cdots & 0 \\ \vdots & \cdots & -1 & 1 & 0 \\ 0 & \cdots & 0 & -1 & 1 \end{pmatrix}.$$

Acknowledgements

The authors would like to thank the associate editor and the anonymous referee for their helpful comments and suggestions. The work of Chihoon Lee was partially supported by the Simons Foundation, under grant number 209658. The work of Ananda Weerasinghe was supported in part by the Army Research Office, under grant number W911NF0710424. We are grateful for their support.

References

[1] Ata, B. and Kumar, S. (2005). Heavy traffic analysis of open processing networks with complete resource pooling: asymptotic optimality of discrete review policies. *Ann. Appl. Prob.* **15,** 331–391.

[2] Bernard, A. and el Kharroubi, A. (1991). Régulations déterministes et stochastiques dans le premier 'orthant' de $\mathbf{R}^n$. *Stoch. Stoch. Rep.* **34,** 149–167.

[3] Biagini, F., Hu, Y., Øksendal, B. and Sulem, A. (2002). A stochastic maximum principle for processes driven by fractional Brownian motion. *Stoch. Process. Appl.* **100,** 233–253.

[4] Biagini, F., Hu, Y., Øksendal, B. and Zhang, T. (2008). *Stochastic Calculus for Fractional Brownian Motion and Applications*. Springer, London.

[5] Cooper, W. L., Schmidt, V. and Serfozo, R. F. (2001). Skorohod-Loynes characterizations of queueing, fluid, and inventory processes. *Queueing Systems* **37,** 233–257.

[6] Delgado, R. (2007). A reflected fBm limit for fluid models with ON/OFF sources under heavy traffic. *Stoch. Process. Appl.* **117,** 188–201.

[7] Duffield, N. G. and O'Connell, N. (1995). Large deviations and overflow probabilities for the general single-server queue, with applications. *Math. Proc. Camb. Phil. Soc.* **118,** 363–374.

[8] Duncan, T. E. (2007). Some stochastic systems with a fractional Brownian motion and applications to control. In *Proc. American Control Conf.* (New York, July 2007), pp. 1110–1114.

[9] Duncan, T. E. and Jin, Y. (2008). Maximum queue length of a fluid model with an aggregated fractional Brownian input. In *Markov Processes and Related Topics: A Festschrift for Thomas G. Kurtz* (Inst. Math. Statist. Collections **4**), Institute of Mathematical Statistics, Beachwood, Ohio, pp. 235–251.

[10] Ghosh, A. P., Roitershtein, A. and Weerasinghe, A. (2010). Optimal control of a stochastic network driven by a fractional brownian motion input. *Adv. Appl. Prob.* **42,** 183–209.

[11] Gong, W.-B., Liu, Y., Misra, V. and Towsley, D. (2005). Self-similarity and long range dependence on the internet: a second look at the evidence, origins and implications. *Comput. Networks* **48,** 377–399.

[12] Hairer, M. (2005). Ergodicity of stochastic differential equations driven by fractional Brownian motion. *Ann. Prob.* **33,** 703–758.

[13] Harrison, J. M. (1985). *Brownian Motion and Stochastic Flow Systems*. John Wiley, New York.

[14] Harrison, J. M. and Shepp, L. A. (1984). A tandem storage system and its diffusion limit. *Stoch. Process. Appl.* **16,** 257–274.

[15] Heath, D., Resnick, S. and Samorodnitsky, G. (1998). Heavy tails and long range dependence in ON/OFF processes and associated fluid models. *Math. Operat. Res.* **23,** 145–165.
[16] Hu, Y. and Zhou, X. Y. (2005). Stochastic control for linear systems driven by fractional noises. *SIAM J. Control Optimization* **43,** 2245–2277.
[17] Hu, Y., Nualart, D. and Song, J. (2009). Fractional martingales and characterization of the fractional Brownian motion. *Ann. Prob.* **37,** 2404–2430.
[18] Jain, R. and Routhier, S. A. (1986). Packet trains: measurements and a new model for computer network traffic. *IEEE J. Sel. Areas Commun.* **4,** 986–995.
[19] Kleptsyna, M. L., Le Breton, A. and Viot, M. (2003). About the linear-quadratic regulator problem under a fractional Brownian perturbation. *ESAIM Prob. Statist.* **7,** 161–170.
[20] Konstantopoulos, T. and Lin, S.-J. (1996). Fractional Brownian approximations of queueing networks. In *Stochastic Networks* (New York, 1995, Lecture Notes Statist. **117**), Springer, New York, pp. 257–273.
[21] Loynes, R. M. (1962). The stability of a queue with non-independent interarrival and service times. *Proc. Camb. Phil. Soc.* **58,** 497–520.
[22] Loynes, R. M. (1964). The stability of a system of queues in series. *Proc. Camb. Phil. Soc.* **60,** 569–574.
[23] Mandelbrot, B. B. and Van Ness, J. W. (1968). Fractional Brownian motions, fractional noises and applications. *SIAM Rev.* **10,** 422–437.
[24] Mandjes, M., Mannersalo, P. and Norros, I. (2007). Gaussian tandem queues with an application to dimensioning of switch fabric interfaces. *Comput. Networks* **51,** 781–797.
[25] Massoulie, L. and Simonian, A. (1999). Large buffer asymptotics for the queue with fractional Brownian input. *J. Appl. Prob.* **36,** 894–906.
[26] Norros, I. (1995). On the use of fractional Brownian motion in the theory of connectionless networks. *IEEE J. Sel. Areas Commun.* **13,** 953–962.
[27] Novikov, A. and Valkeila, E. (1999). On some maximal inequalities for fractional Brownian motions. *Statist. Prob. Lett.* **44,** 47–54.
[28] Nualart, D. (2006). *The Malliavin Calculus and Related Topics*, 2nd edn. Springer, Berlin.
[29] Taqqu, M. S., Willinger, W. and Sherman, R. (1997). Proof of a fundamental result in self-similar traffic modeling. *Comput. Commun. Rev.* **27,** 5–23.
[30] Whitt, W. (2002). *Stochastic-Process Limits*. Springer, New York.
[31] Williams, R. J. (1998). An invariance principle for semimartingale reflecting Brownian motions in an orthant. *Queueing Systems* **30,** 5–25.
[32] Willinger, W., Paxson, V. and Taqqu, M. S. (1998). Self-similarity and heavy tails: structural modeling of network traffic. In *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*, Birkhäuser, Boston, pp. 27–53.
[33] Willinger, W., Taqqu, M. S., Sherman, R. and Wilson, D. V. (1997). Self-similarity through high-variability: statistical analysis of Ethernet LAN traffic at the source level. *IEEE/ACM Trans. Networking* **5,** 71–86.
[34] Yan, L. (2004). Maximal inequalities for the iterated fractional integrals. *Statist. Prob. Lett.* **69,** 69–79.
[35] Zeevi, A. J. and Glynn, P. W. (2000). On the maximum workload of a queue fed by fractional Brownian motion. *Ann. Appl. Prob.* **10,** 1084–1099.