

# Ergodic Rate Control Problem for Single Class Queueing Networks

Amarjit Budhiraja <sup>\*</sup>, Arka P. Ghosh <sup>†</sup> and Chihoon Lee

**Abstract:** We consider critically loaded single class queueing networks with infinite buffers in which arrival and service rates are state (i.e., queue length) dependent and may be dynamically controlled. An optimal rate control problem for such networks with an ergodic cost criterion is studied. It is shown that the value function (i.e., optimum value of the cost) of the rate control problem for the network converges, under a suitable heavy traffic scaling limit, to that of an ergodic control problem for certain controlled reflected diffusions. Furthermore, we show that near optimal controls for limit diffusion models can be used to construct asymptotically near optimal rate control policies for the underlying physical queueing networks. The expected cost per unit time criterion studied here is given in terms of an unbounded holding cost and a linear control cost (“cost for effort”). Time asymptotics of a related uncontrolled model are studied as well. We establish convergence of invariant measures of scaled queue length processes to that of the limit reflecting diffusions. Our proofs rely on infinite time horizon stability estimates that are uniform in control and the heavy traffic parameter, for the scaled queue length processes. Another key ingredient, and a result of independent interest, in the proof of convergence of value functions is the existence of *continuous* near optimal feedback controls for the diffusion control model.

**AMS 2000 subject classifications:** primary 93E20; secondary 60H30, 60J25, 60K25, 93E15.

**Keywords:** ergodic cost, rate control for queueing networks, heavy traffic, reflected diffusions, Skorohod problem, diffusion approximations, invariant measures.

## 1. Introduction

We study Jackson networks with state dependent and dynamically controlled arrival and service rates. The network consists of  $K$  service stations, each of which has an associated infinite capacity buffer. Arrivals of jobs can be from outside the system and/or from internal routing. Upon completion of service at a station, the customer is routed to one of the other service stations (or exits the system) according to a probabilistic routing matrix. Network is assumed to be in heavy traffic in a sense that is made precise in Section 2. Roughly speaking, one considers a sequence of networks with identical topology, parametrized by  $n \in \mathbb{N}$ . Instantaneous arrival and service rates in the  $n$ -th network are queue length dependent and are of order  $\mathcal{O}(n)$ .

---

<sup>\*</sup>Research supported in part by the Army Research Office (Grant W911NF-0-1-0080)

<sup>†</sup>Research supported by NSF grant DMS-0608669.

Additionally, a system manager can exercise service rate controls of order  $\mathcal{O}(n^{1/2})$ . The heavy traffic assumption says that traffic intensity at each station is of the order  $1 - \mathcal{O}(1/n^{1/2})$ . In an uncontrolled setting, diffusion approximations of suitably scaled queue length processes for such networks have been studied in [24] (see also [20]). The limit stochastic process is a reflected diffusion in the nonnegative orthant. Dependence of arrival and service rates on queue length processes leads to drift and diffusion coefficients for the limit model that are state dependent. Existence and uniqueness of solutions to such reflected stochastic differential equations (SDE) follows from the classical theory and well understood regularity properties of the Skorohod map (cf. [11]).

For the controlled setting considered here, a formal analysis along similar lines leads to controlled reflected diffusions with control entering linearly in the drift coefficient. Goal of this work is to use such a formal diffusion approximation in order to obtain provably asymptotically optimal rate control policies for the underlying physical networks. We are concerned with an average long term cost per unit time criterion (see (2.10)). Cost function is a sum of two terms; the first term measures the cost for holding customers in the buffer, while the second term is the cost for exercising control. The holding cost is allowed to have a polynomial growth (as a function of queue lengths) while the control cost is assumed to be linear. Since the cost criterion involves an infinite time horizon, additional stability conditions are needed to ensure the finiteness of the cost (i.e. the expression in (2.10)). In our work these conditions are manifested in the definition of the class of admissible controls. The definition in particular ensures that the controls are “uniformly stabilizing”, viz. all polynomial moments of the queue length process are bounded uniformly in time, the scaling parameter and the choice of admissible control sequence. This stability property, given in Proposition 3.2, is a key ingredient in our proofs. For a description of the set of admissible controls and a discussion of stability considerations that lead to such a formulation, see below (2.1).

Our main result, Theorem 2.8, relates the value function (i.e. the optimum value of the cost) of the queueing rate control problem with the value function associated with an ergodic cost problem for the formal controlled diffusion limit. Cost function for the diffusion model is analogous to that for the physical networks (see (2.15)). It is shown in Theorem 2.8 that, under certain conditions, the value function for the  $n$ -th queueing network converges to that for the diffusion model, as  $n \rightarrow \infty$ . The theorem also shows that using an appropriately chosen  $\epsilon$ -optimal feedback control for the diffusion model, one can construct an asymptotically  $\epsilon$ -optimal rate control policy for the physical queueing network. Thus the theorem describes the precise sense in which the diffusion model, which is originally justified only formally, is a rigorous asymptotic approximation for the queueing model. Rate control problems with an ergodic cost criterion for single class open queueing networks in heavy traffic have been studied in several papers [1, 13, 17–19]. The first two papers concern a one dimensional problem (i.e. a single server queue) and provide explicit expressions/recipes for asymptotically optimal policies. The paper [19] (see also Theorem 2.1, Chapter 9 of [18]) studies a general  $K$  dimensional Jackson type network with finite buffers. In this setting, due to compactness of the state space, all stability issues become relatively easier. The paper [17] considers a very general setting, with unbounded buffers, that includes both rate control and scheduling control problems. However, the paper does not establish the convergence of the value functions.

One of the main steps in the proof of our main result is Theorem 2.7 which establishes existence of an  $\epsilon$ -optimal continuous feedback control for the limit diffusion model. Existence of optimal feedback controls for the class of ergodic diffusion control problems studied here has been established in [7]. Our proof of Theorem 2.7 starts from such an optimal control  $b^*$  and constructs a sequence  $\{b_n\}$  of continuous feedback controls in a manner such that an appropriately chosen measure of the set  $\{b_n \neq b^*\}$  converges to 0 as  $n \rightarrow \infty$ . We show in Theorem 5.2 that, as long as initial distributions converge, the solution  $X^n$  of the reflected SDE that uses  $b_n$  as the feedback control converges weakly to  $X$  which solves the equation with  $b_n$  replaced with  $b^*$ . Proof of this result relies on certain representations and estimates for transition probability densities of reflected Brownian motions (see (7.38)) that are of independent interest. Once weak convergence of  $X^n$  to  $X$  is established, proof of Theorem 2.7 is an immediate consequence of the linearity of the control cost. Next, an  $\epsilon$ -optimal continuous feedback control  $b^\epsilon$ , obtained from Theorem 2.7, is used to define a rate control policy for the  $n$ -th network, for each  $n \in \mathbb{N}$ . The associated costs are shown to converge in Theorem 2.6 to the cost associated with  $b^\epsilon$  in the diffusion control problem. As a corollary of this result one obtains that the value function for the limit diffusion model is an upper bound for any limit point of the sequence of value functions for the controlled queueing networks. In Theorem 2.5, we establish the reverse inequality. As an immediate consequence of these we have the main result, Theorem 2.8, of the paper. Proofs of Theorem 2.5 and 2.6 use functional occupation measure methods developed in [19]. The main idea is to represent the cost associated with the  $n$ -th network in terms of integrals with respect to occupation measures associated with various stochastic processes that constitute the dynamical description of the network. Using stability estimates of Section 3 one can establish tightness of these occupation measures and then classical martingale characterization methods can be applied to identify the limit points in an appropriate manner.

Stability estimates obtained in Section 3 are also useful for the study of an uncontrolled setting. Using Proposition 3.1, we show in Section 4 that, under appropriate conditions, stationary distributions of scaled queue length processes (which are strong Markov processes when there are no controls) converge to the unique stationary distribution of the limit reflecting diffusion model. For the setting where the arrival and service rates are constant, such a result has been proved (for general service/inter-arrival time distributions) under certain exponential integrability conditions on the network primitives in [12]; and more generally with only square integrability conditions on the primitives in [9]. To our knowledge the current paper is the first one to treat a setting with state dependent rates.

The paper is organized as follows. We begin, in Section 2, with problem formulation and assumptions. Set of admissible service rate controls, the ergodic cost criterion of interest and definition of value function for the queueing system are introduced. We then describe the formal controlled diffusion approximation to this queueing system that arises in the heavy traffic limit. We introduce an analogous cost criterion and the associated value function for this limit model. Main results of the work, namely Theorems 2.5, 2.6, 2.7, 2.8 and 2.9, are presented next. Rest of the paper is devoted to the proofs of these results. Theorem 2.8 is an immediate consequence of Theorems 2.5, 2.6 and 2.7 and its proof appears in Section 2. Section 3 presents some key stability and tightness results. Theorem 2.9, that concerns an uncontrolled setting, is proved in

Section 4. Theorem 2.7 is at the heart of our analysis and is treated next in Section 5. Section 6 proves Theorems 2.5 and 2.6.

The following notation will be used. For a metric space  $X$ , let  $\mathcal{B}(X)$  be the Borel  $\sigma$ -field on  $X$  and  $\mathcal{P}(X)$  the collection of all probability measures on  $X$ . The Dirac measure at the point  $x$  is denoted by  $\delta_x$ . For  $\mu \in \mathcal{P}(X)$  and a  $\mu$ -integrable function  $f$ , we will denote  $\int_X f(x)\mu(dx)$  as  $\langle \mu, f \rangle$ . The set of positive integers is denoted by  $\mathbb{N}$ , the set of real numbers by  $\mathbb{R}$  and nonnegative real numbers by  $\mathbb{R}_+$ . For sets  $A, B \subseteq \mathbb{R}^d$ ,  $dist(A, B)$  will denote the distance between two sets, i.e.,  $\inf\{|x - y| : x \in A, y \in B\}$ . For a given matrix  $M$  denote by  $M'$  its transpose and by  $M^i$  the  $i$ -th row of  $M$ . Let  $\mathbb{I} = \mathbb{I}_{K \times K}$  denote the identity matrix for some  $K$ . When clear from the context, we will omit the subscript. For  $a, b \in \mathbb{R}$ , let  $a \vee b \doteq \max\{a, b\}$  and  $a \wedge b \doteq \min\{a, b\}$ . The convergence in distribution of random variables (with values in some Polish space)  $\Phi_n$  to  $\Phi$  will be denoted as  $\Phi_n \Rightarrow \Phi$ . With an abuse of notation weak convergence of probability measures (on some Polish space)  $\mu_n$  to  $\mu$  will also be denoted as  $\mu_n \Rightarrow \mu$ . When  $\sup_{0 \leq s \leq t} |f_n(s) - f(s)| \rightarrow 0$  as  $n \rightarrow \infty$ , for all  $t \geq 0$ , we say that  $f_n \rightarrow f$  uniformly on compact sets. The class of continuous functions  $f : X \rightarrow Y$  is denoted by  $C(X, Y)$ , real continuous bounded functions on  $X$  by  $C_b(X)$ . For  $f \in C_b(X)$ , define  $\|f\|_\infty \doteq \sup_{x \in X} |f(x)|$ . Let  $D(I, Y)$ , when  $I$  is  $[0, 1]$  or  $[0, \infty)$ , denote the class of right continuous functions with having left limit defined from  $I$  to  $Y$ , equipped with the usual Skorohod topology. A sequence  $\{Z_n\}$  of  $D([0, \infty), X)$  valued random variables is said to be  $C$ -tight if and only if the measures induced by the  $Z_n$  on  $(D([0, \infty), X), \mathcal{B}(D([0, \infty), X)))$  form a tight sequence and any weak limit point of the sequence has values in  $C([0, \infty), X)$  almost surely. Finally, let  $C(X), C[0, 1], C[0, \infty)$  denote  $C(X, \mathbb{R}), C([0, 1], \mathbb{R}), C([0, \infty), \mathbb{R})$ , respectively. Similarly,  $D[0, 1], D[0, \infty)$  will denote  $D([0, 1], \mathbb{R}), D([0, \infty), \mathbb{R})$ , respectively. A stochastic process will be denoted interchangeably by  $\{Z(t)\}$  or  $\{Z_t\}$ . Inequalities for vectors are interpreted componentwise. We will denote generic constants by  $c_1, c_2, \dots$ , and their values may change from one proof to another.

## 2. Problem formulation and main results

**Network structure.** We begin with a rough description which will subsequently be made precise. Consider a sequence of networks indexed by  $n$ . Each network has a similar structure, in particular there are  $K$  service stations, with the  $i$ -th station denoted by  $P_i, i \in \mathcal{K} \doteq \{1, \dots, K\}$ . We assume that each station has an infinite capacity buffer. Arrivals of jobs can be from outside the system and/or from internal routing. Upon completion of service at station  $P_i$  a customer is routed to other service stations (or exits the system) according to a probabilistic routing matrix  $\mathbb{P} = (p_{ij})_{i,j \in \mathcal{K}}$ . At every station the jobs are assumed to be processed by First-In-First-Out discipline. We assume that the network is open, that is, any customer entering the network eventually leaves it (this is made precise in Assumption 2.1(i)). We allow arrival and service rates to be time varying random processes. They may be described as deterministic functions of the state processes or, more generally, given as nonanticipative control processes. Although quite general nonanticipative controls are allowed, one finds (see Theorem 2.7) that asymptotically near optimal performance can be achieved using continuous Markov controls, i.e. controls that are given as a continuous function of the state.

A precise description of the model is as follows. Let  $\lambda_i^n, \mu_i^n, n \geq 1, i \in \mathbb{K}$  be measurable functions from  $\mathbb{R}_+^K \rightarrow \mathbb{R}_+$ . We write  $\lambda^n = (\lambda_1^n, \dots, \lambda_K^n)'$ ;  $\mu^n$  is defined similarly. These functions will be used to define state dependent arrival and service rates. External arrivals are assumed to occur any for  $i \in \mathbb{K}_e$ , where  $\mathbb{K}_e$  is a subset of  $\mathbb{K}$ . Thus  $\lambda_i^n(x) = 0$  for all  $n \in \mathbb{N}, x \in S \doteq \mathbb{R}_+^K$  and  $i \in \mathbb{K}^c \doteq \mathbb{K} \setminus \mathbb{K}_e$ . Let  $R \doteq [\mathbb{I} - \mathbb{P}']$ . This matrix plays a central role in heavy traffic theory of Jackson networks and in the study of stability properties of diffusions approximating such networks [14, 15]. Let  $a^n(\cdot) \doteq \lambda^n(\cdot) - R\mu^n(\cdot)$ , and set  $\hat{a}^n \doteq \frac{1}{\sqrt{n}}a^n$ . The following is our standing assumption.

**Assumption 2.1.**

- (i) Spectral radius of  $\mathbb{P}$  is strictly less than 1 and  $p_{ii} = 0$  for all  $i \in \mathbb{K}$ .
- (ii) For each  $n \in \mathbb{N}$ ,  $\lambda^n, \mu^n \in C(S)$ .
- (iii) For some  $\kappa_1 \in (0, \infty)$ ,  $|\lambda^n(x)| \leq n\kappa_1, |\mu^n(x)| \leq n\kappa_1$  for all  $n \geq 1$  and  $x \in S$ .
- (iv) There exists a constant  $\kappa_2 \in (0, \infty)$  such that  $\sup_{x \in S} |\hat{a}^n(x)| \leq \kappa_2$ .
- (v) There exists  $a \in C_b(S)$  such that  $\frac{a^n(\sqrt{n}\cdot)}{\sqrt{n}} \rightarrow a(\cdot)$  uniformly on compact sets as  $n \rightarrow \infty$ .
- (vi) There exist Lipschitz functions  $\lambda, \mu \in C_b(\mathbb{R}^K)$ , such that  $\frac{\lambda^n(\sqrt{n}\cdot)}{n} \rightarrow \lambda(\cdot), \frac{\mu^n(\sqrt{n}\cdot)}{n} \rightarrow \mu(\cdot)$  uniformly on compact sets, as  $n \rightarrow \infty$ . Furthermore,  $\lambda = R\mu$ .
- (vii)  $\inf_{i \in \mathbb{K}_e} \inf_{x \in S} \inf_n \frac{\lambda_i^n(\sqrt{n}x)}{n} = \underline{\lambda} > 0, \inf_{i \in \mathbb{K}} \inf_{x \in S} \inf_n \frac{\mu_i^n(\sqrt{n}x)}{n} = \underline{\mu} > 0$ .
- (viii) For each  $i \in \mathbb{K} \setminus \mathbb{K}_e$ , there exists  $j \in \mathbb{K}_e$  such that  $\mathbb{P}^m(j, i) > 0$  for some  $m \in \mathbb{N}$ .

Part (i) of the assumption says that the network is open. It in particular implies ([11, 14]) the existence of a regular Skorohod map associated with the network data (see Proposition 2.3), which is a key ingredient for many estimates in this paper. Conditions (ii)-(vii) describe the *heavy traffic* parameter regime that is considered here. We note that the arrival and service rates,  $\lambda^n, \mu^n$  are  $\mathcal{O}(n)$ . Many works that study diffusion approximations for queueing systems (eg. [14], [18]) formulate the underlying heavy traffic condition in terms of rates that are  $\mathcal{O}(1)$  and then consider scaled processes where time is scaled up by a factor of  $n$  while space is scaled down by a factor of  $\sqrt{n}$ . In the scaling considered here (see (2.7)), which is adapted from [24] (see also [20]), the time parameter is left unchanged and the factor of  $n$  is absorbed in the arrival and service rates. A simple example of rate functions satisfying (ii)-(vii) is given as  $\lambda^n(x) = n\lambda^*(x/\sqrt{n}), \mu^n(x) = n\mu^*(x/\sqrt{n})$ , where  $\lambda^*, \mu^*$  are nonnegative bounded Lipschitz functions, bounded away from 0, such that  $\lambda^* = R\mu^*$ . Parts (vii) and (viii) are nondegeneracy conditions that ensure that the diffusion coefficient in the limiting reflected diffusion model is uniformly positive definite (cf. (2.13)). Condition (viii) can be assumed without loss of generality since if it fails for some  $i \in \mathbb{K} \setminus \mathbb{K}_e$ , one can consider a reduced model that is obtained by omitting station  $P_i$ .

In addition to state dependence of service rates, we will allow the modulation of these rates by a controller. Class of allowable controls will be such that the corresponding queueing system is stable. To be precise we consider the following set of controls: Fix  $\delta_0, M \in (0, \infty)$ . Let

$$\bar{\alpha} \doteq \sup_{x \in S, i \in \mathbb{K}} [R^{-1}a(x)]_i. \quad (2.1)$$

Define  $\Lambda^{\delta_0} \doteq \{u \in \mathbb{R}^K : \bar{\alpha}\mathbf{1} - u \leq -\delta_0\mathbf{1}, |u| \leq M\}$ . Let  $\Lambda_n^{\delta_0} \doteq \{u \in \mathbb{R}^K : \frac{u}{\sqrt{n}} \in \Lambda\}$ . We

will suppress  $\delta_0$  from the notation unless needed. A control for the  $n$ -th network will be a stochastic process with values in  $\Lambda_n$ . Such a process will enter in the state dynamics (see (2.2)) as an additive change to the nominal service rate  $\mu^n$ . Thus in the controlled model, service rates are allowed to be  $\mathcal{O}(\sqrt{n})$  of the nominal rates. Although such perturbations approach 0, as  $n \rightarrow \infty$ , under the fluid scaling, they can have a significant impact on asymptotic system performance viewed under the diffusion scaling. Nonanticipative properties of the control are made precise in the next paragraph. The lower bound requirement on  $u \in \Lambda_n$  can be interpreted as a condition on the minimal service rate at which the system is allowed to operate. If the effective service rate  $\mu_i^n(Q^n(s)) + U_i^n(s)$  (see (2.2)) is too small, one expects the corresponding queue lengths to blow up. Indeed, for the diffusion limit model  $Z$  in (2.14), for an uncontrolled, constant coefficient setting, the condition  $\bar{\alpha} < 0$  is known to be necessary and sufficient for positive recurrence of  $Z$  (see [15]). The requirement that in the controlled setting a limit model control  $U$  should satisfy  $\bar{\alpha}\mathbf{1} - U(s) < 0$  is achieved by similar stability considerations (cf. [7]). This, in view of the scaling, motivates the definition of the control set  $\Lambda_n$  introduced above. In particular the above lower bound constraint on controls is at the heart of the stability estimate of Proposition 3.2.

We now describe the precise mathematical model. Fix  $n \in \mathbb{N}$ . Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a probability space on which are given unit rate independent Poisson processes,  $N_i, N_{ij}$ ,  $i \in \mathbb{K}, j \in \mathbb{K} \cup \{0\}$ . For  $i \in \mathbb{K}$ ,  $N_i$  will be used to define the stream of jobs entering the  $i$ -th buffer and for  $i, j \in \mathbb{K}$ ,  $N_{ij}$  will be used to represent the flow of jobs to buffer  $j$  from buffer  $i$ . For  $i \in \mathbb{K}$  and  $j = 0$ ,  $N_{ij}$  will be associated with jobs that leave the system after service at station  $P_i$ . Precise state evolution is described below.

Let  $U^n = (U_1^n, U_2^n, \dots, U_K^n)'$  be a  $\Lambda_n$  valued measurable process.  $U^n$  will represent the service rate control in the system. Under the control  $U^n$ , the state of the system  $Q^n = (Q_1^n, Q_2^n, \dots, Q_K^n)'$  is given by the following equation.

$$\begin{aligned} Q_i^n(t) &= Q_i^n(0) + N_i \left( \int_0^t \lambda_i^n(Q^n(s)) ds \right) \\ &\quad + \sum_{j=1}^K N_{ji} \left( p_{ji} \int_0^t 1_{\{Q_j^n(s) > 0\}} [\mu_j^n(Q^n(s)) + U_j^n(s)] ds \right) \\ &\quad - \sum_{j=0}^K N_{ij} \left( p_{ij} \int_0^t 1_{\{Q_i^n(s) > 0\}} [\mu_i^n(Q^n(s)) + U_i^n(s)] ds \right), \quad i \in \mathbb{K}, \end{aligned} \tag{2.2}$$

where  $Q^n$  represents the queue length vector process obtained under the (rate) control process  $U^n$ . In the above display  $p_{i0} \doteq 1 - \sum_{j=1}^K p_{ij}$ . We require that for some filtration  $\{\mathcal{F}_t\}$  on  $(\Omega, \mathcal{F}, \mathbb{P})$ ,  $U^n$  is  $\{\mathcal{F}_t\}$ -progressively measurable,  $Q^n$  is  $\{\mathcal{F}_t\}$ -adapted, and  $M_{ij}^n$ ,  $i, j \in \mathbb{K} \cup \{0\}$ ,

defined below are  $\{\mathcal{F}_t\}$  martingales.

$$\begin{aligned} M_{0i}^n(t) &\doteq N_i \left( \int_0^t \lambda_i^n(Q^n(s)) ds \right) - \int_0^t \lambda_i^n(Q^n(s)) ds, \\ M_{ij}^n(t) &\doteq N_{ij} \left( p_{ij} \int_0^t 1_{\{Q_i^n(s) > 0\}} [\mu_i^n(Q^n(s)) + U_i^n(s)] ds \right) \\ &\quad - p_{ij} \int_0^t 1_{\{Q_i^n(s) > 0\}} [\mu_i^n(Q^n(s)) + U_i^n(s)] ds. \end{aligned}$$

From Assumption 2.1, it follows that the martingales defined above are square integrable. We will refer to a  $U^n$  satisfying above conditions as an admissible control and denote by  $\mathcal{A}_n$  the collection of all admissible controls. We remark that, although one can assume from the point of view of applications that all controls and controlled processes are given on a single probability space, there is some subtlety in defining an appropriate filtration. In general, a filtration for which the above stated adaptedness and martingale properties hold would depend on the choice of the control (and the corresponding state process). For a Markov control one can take the filtration to be the one generated by  $\{M_{ij}^n, Q_i^n; i, j\}$ .

Letting  $M_i^n \doteq \sum_{j=0}^K M_{ji}^n - \sum_{j=0}^K M_{ij}^n$ , we can rewrite the evolution (2.2) as

$$\begin{aligned} Q_i^n(t) &= Q_i^n(0) + \int_0^t \left[ \lambda_i^n(Q^n(s)) + \sum_{j=1}^K p_{ji} \mu_j^n(Q^n(s)) - \mu_i^n(Q^n(s)) \right] ds \\ &\quad + \int_0^t \left[ \sum_{j=1}^K p_{ji} U_j^n(s) - U_i^n(s) \right] ds + M_i^n(t) + [RY^n(t)]_i, \end{aligned} \quad (2.3)$$

where

$$Y_i^n(t) = \sum_{j=0}^K N_{ij} \left( p_{ij} \int_0^t 1_{\{Q_i^n(s)=0\}} [\mu_i^n(Q^n(s)) + U_i^n(s)] ds \right), \quad i \in \mathcal{K}.$$

Note that  $Y_i^n$  is a right continuous with having left limits, nondecreasing  $\{\mathcal{F}_t\}$  adapted process and  $Y_i^n$  increases only when  $Q_i^n(t) = 0$ , i.e.,  $\int_0^\infty 1_{\{Q_i^n(t) \neq 0\}} dY_i^n(t) = 0$  a.s.

The state evolution can be described succinctly by the following equation:

$$Q^n(t) = Q^n(0) + \int_0^t a^n(Q^n(s)) ds - \int_0^t R U^n(s) ds + M^n(t) + R Y^n(t). \quad (2.4)$$

The above dynamics can equivalently be described in terms of a Skorohod map as described below.

**Definition 2.2.** Let  $\psi \in D([0, \infty), \mathbb{R}^K)$  be given with  $\psi(0) \in S$ . Then  $(\phi, \eta) \in D([0, \infty), \mathbb{R}^K) \times D([0, \infty), \mathbb{R}^K)$  solves the Skorohod problem for  $\psi$  with respect to  $S$  and  $R$  if and only if the following hold:

- (i)  $\phi(t) = \psi(t) + R\eta(t) \in S$ , for all  $t \geq 0$ ;

- (ii)  $\eta$  satisfies, for  $i \in \mathbb{K}$ , (a)  $\eta_i(0) = 0$ , (b)  $\eta_i$  is nondecreasing, and (c)  $\eta_i$  can increase only when  $\phi$  is on the  $i$ -th face of  $S$ , that is,  $\int_0^\infty 1_{\{\phi_i(s) \neq 0\}} d\eta_i(s) = 0$ .

Let  $D_S([0, \infty), \mathbb{R}^K) \doteq \{\psi \in D([0, \infty), \mathbb{R}^K) : \psi(0) \in S\}$ . On the domain  $D \subset D_S([0, \infty), \mathbb{R}^K)$  on which there is a unique solution to the Skorohod problem (SP) we define the Skorohod map (SM)  $\Gamma$  as  $\Gamma(\psi) \doteq \phi$ , if  $(\phi, R^{-1}[\phi - \psi])$  is the unique solution of the SP posed by  $\psi$ . Define the map  $\Gamma_1 : D \rightarrow D([0, \infty), \mathbb{R}^K)$  by  $\Gamma_1(\psi) \doteq R^{-1}[\phi - \psi]$ . The following result (see [11, 14]) gives the regularity of the Skorohod map defined by the data  $(S, R)$ . It is a consequence of Assumption 2.1 (i).

**Proposition 2.3.** *The Skorohod map (SM) is well defined on all of  $D_S([0, \infty), \mathbb{R}^K)$  (that is,  $D = D_S([0, \infty), \mathbb{R}^K)$ ) and the SM is Lipschitz continuous in the following sense. There exists a constant  $L \in (1, \infty)$  such that for all  $\psi_1, \psi_2 \in D_S([0, \infty), \mathbb{R}^K)$ :*

$$\sup_{0 \leq t < \infty} \{|\Gamma(\psi_1)(t) - \Gamma(\psi_2)(t)| + |\Gamma_1(\psi_1)(t) - \Gamma_1(\psi_2)(t)|\} < L \sup_{0 \leq t < \infty} |\psi_1(t) - \psi_2(t)|. \quad (2.5)$$

The dynamics in (2.4) can now equivalently be described in terms of the SM as follows:

$$Q^n = \Gamma(\tilde{Q}^n), \quad Y^n = \Gamma_1(\tilde{Q}^n), \quad (2.6)$$

where  $\tilde{Q}^n(t) \doteq \tilde{Q}^n(0) + \int_0^t a^n(Q^n(s)) ds - \int_0^t R U^n(s) ds + M^n(t)$ . For asymptotic analysis we will consider processes under the diffusion scaling. Given a stochastic process  $X^n$ , we will denote by  $\hat{X}^n$  the process defined as

$$\hat{X}^n(t) \doteq X^n(t)/\sqrt{n}, \quad t \geq 0, \quad (2.7)$$

and refer to it as the diffusion-scaled form of  $X^n$ . With this notation, we have from (2.6) that

$$\hat{Q}^n(t) = \Gamma \left( \hat{Q}^n(0) + \int_0^t \hat{a}^n(\sqrt{n}\hat{Q}^n(s)) ds - \int_0^t R \hat{U}^n(s) ds + \hat{M}^n(\cdot) \right) (t). \quad (2.8)$$

When  $\hat{Q}^n(0) \equiv x$ , we will sometimes write the corresponding scaled state process as  $\hat{Q}_x^n$ .

In this work we are concerned with an ergodic cost problem associated with the sequence of controlled queueing systems  $\{Q^n\}_{n \geq 1}$ . Let  $k(\cdot)$  be a continuous function from  $S$  to  $\mathbb{R}_+$  and  $c \in \mathbb{R}^K$  be fixed. We will assume that for some  $m_0 \in (0, \infty)$

$$\sup_{x \in S} \frac{k(x)}{1 + |x|^{m_0}} < \infty. \quad (2.9)$$

Define for each  $n \in \mathbb{N}$ ,  $x \in S$ ,  $T \in (0, \infty)$ , the ‘average cost’ on time interval  $[0, T]$  for the control  $U^n \in \mathcal{A}_n$  by

$$J_T^n(U^n, x) \doteq \frac{1}{T} \mathbb{E} \int_0^T \left[ k(\hat{Q}_x^n(s)) + c \cdot \hat{U}^n(s) \right] ds.$$

Asymptotic average cost is defined as

$$J^n(U^n, x) \doteq \limsup_{T \rightarrow \infty} J_T^n(U^n, x) = \limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \int_0^T \left[ k(\hat{Q}_x^n(s)) + c \cdot \hat{U}^n(s) \right] ds. \quad (2.10)$$

Define the value function  $V_n$  for the  $n$ -th queueing system as

$$V_n(x) \doteq \inf_{U^n \in \mathcal{A}_n} J^n(U^n, x). \quad (2.11)$$

The main goal of this paper is to show that the value function and near optimal control policies in the  $n$ -th network, for large  $n$ , can be approximated by the same quantities in an associated diffusion control problem. Towards this goal, we now introduce a diffusion control problem.

Define for  $x \in S$ , a  $[K \times [K + K(K + 1)]]$ -dimensional matrix  $\Sigma(x)$  as

$$\Sigma(x) \doteq (A(x), B_1(x), \dots, B_K(x)), \quad (2.12)$$

where  $A$  and  $B_i, i \in \mathbb{K}$  are  $K \times K$  and  $K \times (K + 1)$  matrices given as follows. For  $x \in S$ ,

$$A(x) = \text{diag}(\sqrt{\lambda_1(x)}, \dots, \sqrt{\lambda_K(x)}), \quad B_i(x) = (B_i^0(x), B_i^1(x), \dots, B_i^K(x)),$$

where  $B_i^0(x) \doteq -\sqrt{\mu_i(x)}\mathbf{1}_i$  and for  $j \in \mathbb{K}, j \neq i$ ,  $B_i^j(x) \doteq \mathbf{1}_{ij}\sqrt{p_{ij}\mu_i(x)}$ . Here  $\mathbf{1}_i$  is a  $K$ -dimensional vector with 1 at the  $i$ -th coordinate and 0 elsewhere;  $\mathbf{1}_{ij}$  is a  $K$ -dimensional vector with -1 at the  $i$ -th; 1 at the  $j$ -th coordinates; and zeroes elsewhere. Finally,  $B_i^i(x) \doteq 0$ . It is easy to see (cf. Appendix) that due to Assumption 2.1 (vii) and (viii),  $\Sigma(x)\Sigma(x)'$  is uniformly nondegenerate. More precisely,

$$\text{there exists a } \kappa \in (0, \infty) \text{ such that for all } v \in \mathbb{R}^K, x \in S, v'(\Sigma(x)\Sigma(x)')v \geq \kappa v'v. \quad (2.13)$$

One can then find a Lipschitz  $K \times K$  matrix valued function  $\sigma(x)$  such that  $\Sigma(x)\Sigma(x)' = \sigma(x)\sigma(x)'$ . (See Theorem 5.2.2 in [23].)

Let  $B$  be a  $K$ -dimensional standard Brownian motion given on some filtered probability space  $(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{\mathbb{P}}, \{\widehat{\mathcal{F}}_t\})$ . We will denote  $(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{\mathbb{P}}, \{\widehat{\mathcal{F}}_t\}, B)$  by  $\Xi$  and refer to it as a system. Denote by  $\mathcal{A}(\Xi)$  the collection of all  $\{\widehat{\mathcal{F}}_t\}$  progressively measurable  $\Lambda$ -valued processes. We will refer to such processes as admissible controls (for the diffusion control problem). For  $U \in \mathcal{A}(\Xi)$  and  $x \in S$ , let  $Z \equiv Z^{x,U}$  be the unique solution of

$$Z(t) = \Gamma \left( x + \int_0^t a(Z_s)ds - \int_0^t RU(s)ds + \int_0^t \sigma(Z_s)dB_s \right) (t). \quad (2.14)$$

Ergodic cost for this limit diffusion model is given as follows:

$$J(U, x) \doteq \limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \int_0^T [k(Z(s)) + c \cdot U(s)] ds. \quad (2.15)$$

The value is defined as

$$V \doteq \inf_{\Xi} \inf_{U \in \mathcal{A}(\Xi)} J(U, x). \quad (2.16)$$

The above formulation of the control problem, where one optimizes over all controls given on an arbitrary system, is common in the theory of stochastic control for diffusions (cf. [5]). A key reason to allow for an arbitrary system is that if one replaces  $U(s)$  in (2.14) by  $\beta(Z(s))$ , where  $\beta$  is a measurable map from  $S$  to  $\Lambda$ , then the resulting equation in general may not admit

a strong solution. However, due to uniform nondegeneracy of  $\sigma\sigma'$ , a unique weak solution to (2.14) is ensured. Such maps, also referred to as Markov controls, are some of the most basic forms of actions a system manager can exercise and thus it is desirable to allow for them in the problem formulation. This motivates the definition in (2.16). We note that the expression in (2.16) does not depend on  $x$  (cf. [7]).

We denote by  $BM(S, \Lambda)$  the space of all bounded measurable maps from  $S$  to  $\Lambda$ . For  $b \in BM(S, \Lambda)$ , let  $U_b$  be an admissible control for the limit diffusion model, given on some system  $\Xi$  as:

$$U_b(t) = b(Z_t), \quad Z_t = \Gamma \left( x - \int_0^t [Rb(Z_s) - a(Z_s)] ds + \int_0^t \sigma(Z_s) dB_s \right) (t), \quad t \geq 0. \quad (2.17)$$

Existence of such a  $\Xi$  and  $Z$  and the uniqueness in law of  $Z$  are consequences of uniform nondegeneracy of  $\sigma$ .

The following result follows from [7] and [8]. Some comments on the proof are given in the Appendix.

**Theorem 2.4.** (i) Fix  $b \in BM(S, \Lambda)$ . Then  $Z$  defined by (2.17) is a strong Markov process that admits a unique invariant probability measure  $\eta_b$ .

(ii) There is a  $\varpi \in (0, \infty)$  such that  $\sup_{b \in BM(S, \Lambda)} \int_S e^{\varpi|x|} \eta_b(dx) < \infty$ .

(iii) For each  $b \in BM(S, \Lambda)$ ,  $J(U_b) \equiv J(U_b, x) = \langle \eta_b, k_b \rangle$ , where  $k_b = k + c \cdot b$ . In particular, the cost is the same for all initial states  $x$ .

(iv) There is a  $b^* \in BM(S, \Lambda)$  such that  $V = J(U_{b^*}) = \langle \eta_{b^*}, k_{b^*} \rangle$ .

We now present the main results of this work.

**Theorem 2.5.** For every bounded sequence  $\{x_n\} \subseteq S$ ,  $\liminf_n V_n(x_n) \geq V$ .

For  $b \in BM(S, \Lambda)$ , let  $U_b^n \in \mathcal{A}_n$  be defined as  $U_b^n(s) \doteq \sqrt{nb}(\widehat{Q}_x^n(s))$ , where  $\widehat{Q}_x^n$  is given by (2.8) with  $U^n$  on the right side of (2.8) replaced by  $U_b^n$ . Note that we have suppressed the dependence of  $U_b^n$  and  $U_b$  on  $x$  in the notation.

**Theorem 2.6.** Let  $\tilde{b} : S \rightarrow \Lambda$  be continuous. Let  $\{x_n\} \subseteq S$  be a bounded sequence. Then  $J^n(U_{\tilde{b}}^n, x_n) \rightarrow J(U_{\tilde{b}})$  as  $n \rightarrow \infty$ .

**Theorem 2.7.** Suppose that for all  $x \in S$ ,  $\sigma(x) \equiv \sigma$ . Then, given  $\epsilon > 0$ , there exists a continuous  $b : S \rightarrow \Lambda$  such that  $J(U_b) \leq V + \epsilon$ .

As an immediate consequence of Theorems 2.5, 2.6 and 2.7 we get the following main result.

**Theorem 2.8.** Suppose that for all  $x \in S$ ,  $\sigma(x) \equiv \sigma$ . Let  $\{x_n\} \subseteq S$  be a bounded sequence. Then as  $n \rightarrow \infty$ ,  $V_n(x_n) \rightarrow V$ . Also, for every  $\epsilon > 0$ , there exists a continuous  $b : S \rightarrow \Lambda$  such that

$$0 \leq \limsup_{n \rightarrow \infty} [J^n(U_b^n, x_n) - V_n(x_n)] \leq \epsilon. \quad (2.18)$$

Note that Theorem 2.5, which gives a lower bound on the asymptotic value function (i.e.  $\liminf_n V_n$ ), does not require  $\sigma$  to be constant. However, the upper bound is only derived under the condition  $\sigma(x) \equiv \sigma$ . The main reason for this restrictive assumption is that the approach taken in this paper crucially relies on the existence of a near optimal continuous Markov control for the limit diffusion control problem (i.e. Theorem 2.7). The proof for the existence of such a control uses certain representations and estimates for transition probability densities of reflected diffusions (see (7.38)) which we currently are unable to establish for a general state dependent setting. For an alternative approach to the construction of a near optimal continuous Markov control, we refer the reader to [18] (see pages 362-363 and Theorem 4.4.5 therein) where a model with a compact state space is studied. We believe that the construction given in the current paper and the related transition probability estimates are of independent interest.

**Proof of Theorem 2.8.** From Theorem 2.5

$$\liminf_{n \rightarrow \infty} V_n(x_n) \geq V.$$

Also, for fixed  $\epsilon > 0$ , we have from Theorem 2.7 that for some continuous  $b : S \rightarrow \Lambda$ ,  $J(U_b) \leq V + \epsilon$ . Let  $U_b^n$  be as defined above Theorem 2.6. Then from Theorem 2.6

$$\limsup_{n \rightarrow \infty} V_n(x_n) \leq \limsup_{n \rightarrow \infty} J^n(U_b^n, x_n) = J(U_b) \leq V + \epsilon.$$

We get on combining the inequalities in the above two displays that  $V_n(x_n) \rightarrow V$ , as  $n \rightarrow \infty$ , and that (2.18) holds. ■

As a side consequence of estimates used in the proofs of Theorems 2.5-2.7, we obtain the following result on convergence of steady state distributions of  $\hat{Q}^n$ .

**Theorem 2.9.** *Suppose that  $\bar{\alpha} < 0$ . Let  $\hat{Q}^n$  be given by (2.8) with  $\hat{U}^n \equiv 0$ . Then the Markov process  $\hat{Q}^n$  admits a stationary probability distribution. Furthermore, Markov process  $Z$  given by (2.14) with  $U \equiv 0$  admits a unique stationary probability distribution  $\pi$ . Finally, if  $\pi_n$  is an arbitrary stationary law for  $\hat{Q}^n$ ,  $n \geq 1$  then  $\pi_n \Rightarrow \pi$  as  $n \rightarrow \infty$ .*

We remark that the condition  $\bar{\alpha} < 0$  ensures that  $0 \in \Lambda^{\delta_0}$  for some  $\delta_0 > 0$ . Thus  $\hat{U}^n \equiv 0$  is an admissible control with such a choice of  $\delta_0$ . As noted in the introduction, a result similar to Theorem 2.9 for the constant coefficient setting (i.e.  $\sigma(x) \equiv \sigma$ ,  $a(x) \equiv a$ ) was established in [12] (see also [9]).

We start with proving some stability results in Section 3. Section 4 is devoted to the proof of Theorem 2.9. Proof of Theorem 2.7 is given in Section 5. Theorems 2.5 and 2.6 are proved in Section 6.

### 3. Some stability results

Following stability properties are key ingredients in the proofs.

**Proposition 3.1.** *Let  $\widehat{Q}_x^n$  be defined by (2.8) with  $\widehat{Q}_x^n(0) \equiv x \in S$  and  $U^n \in \mathcal{A}_n$ . Then there exists a  $t_0 \in (0, \infty)$  such that for all  $t \geq t_0$ ,*

$$\lim_{|x| \rightarrow \infty} \sup_n \sup_{U^n \in \mathcal{A}_n} \frac{1}{|x|^2} \mathbb{E} \left( |\widehat{Q}_x^n(t|x)|^2 \right) = 0.$$

**Proof.** Fix  $x \in S$  and  $U^n \in \mathcal{A}_n$ . Write (2.8) as,  $\widehat{Q}_x^n(t) \equiv \Gamma \left( x + r^n(\cdot) + \widehat{M}^n(\cdot) \right) (t)$ , where

$$r^n(t) \doteq \int_0^t \widehat{a}^n(\sqrt{n}\widehat{Q}^n(s)) ds - \int_0^t R\widehat{U}^n(s) ds \equiv \int_0^t \bar{b}^n(s) ds, t \geq 0.$$

Define  $Z_x^n(t) \doteq \Gamma(x + r^n(\cdot))(t)$ ,  $t \geq 0$ . Using the Lipschitz property of the Skorohod map (Proposition 2.3), we have

$$|\widehat{Q}_x^n(t) - Z_x^n(t)| \leq L \sup_{0 \leq s \leq t} |\widehat{M}^n(s)|, \text{ for all } t \geq 0. \quad (3.1)$$

Next, denoting by  $\mathcal{C} \doteq \{v \in \mathbb{R}^K : R^{-1}v \leq 0\}$  we see, from Assumption 2.1 and the discussion below it, that there exists a  $\delta \in (0, \infty)$  satisfying

$$\inf_n \inf_s \text{dist}(\bar{b}^n(s), \partial\mathcal{C}) \geq \delta. \quad (3.2)$$

Thus for all  $n \geq 1$  and  $s \geq 0$ ,  $\bar{b}^n(s) \in \mathcal{C}_\delta \doteq \{v \in \mathcal{C} : \text{dist}(v, \partial\mathcal{C}) \geq \delta\}$ . For  $q_0 \in S$  denote by  $\mathcal{T}(q_0)$  the collection of all trajectories  $\psi : [0, \infty) \rightarrow S$  of the form

$$\psi(t) \doteq \Gamma \left( q_0 + \int_0^t \xi(s) ds \right) (t), \quad t \geq 0,$$

where  $\xi : [0, \infty) \rightarrow \mathbb{R}^K$  is a measurable map satisfying for all  $t \in [0, \infty)$ ,  $\int_0^t |\xi(s)| ds < \infty$ ,  $\xi(t) \in \mathcal{C}_\delta$ . Define

$$T(q_0) \doteq \sup_{\psi \in \mathcal{T}(q_0)} \inf \{t \in [0, \infty) : \psi(t) = 0\}. \quad (3.3)$$

Lemma 3.1 of [2] shows that

$$T(q_0) \leq \frac{4L^2}{\delta} |q_0|, \text{ and for all } \psi \in \mathcal{T}(q_0), \psi(t) = 0 \text{ for all } t \geq T(q_0).$$

Combining this observation with (3.2) we now have that  $Z_x^n(t) = 0$ , for all  $t \geq \bar{D}|x|$ , where  $\bar{D} \doteq \frac{4L^2}{\delta}$ . Using this in (3.1) we now see that

$$|\widehat{Q}_x^n(t|x)| \leq L \sup_{0 \leq s \leq t|x|} |\widehat{M}^n(s)|, \quad (3.4)$$

for all  $t \geq \bar{D}$  and for all initial conditions  $x$ .

Next we obtain an estimate on the second moment of the right side of (3.4). Noting that  $\widehat{M}_i^n$  is an  $\{\mathcal{F}_t\}_{t \geq 0}$  square integrable martingale, one gets using Doob's inequality that

$$\begin{aligned} \mathbb{E} \sup_{0 \leq s \leq t} |\widehat{M}_i^n(s)|^2 &\leq 4\mathbb{E}|\widehat{M}_i^n(t)|^2 \leq \frac{c_1}{n} \mathbb{E} \left[ \int_0^t \left[ \lambda_i^n(Q^n(s)) + \sum_{j=1}^K \mu_j^n(Q^n(s)) + U_i^n(s) \right] ds \right] \\ &\leq c_2(1+t), \end{aligned} \quad (3.5)$$

where the last inequality follows from Assumption 2.1 (iii). (Here  $c_1, c_2 \in (0, \infty)$  are independent of  $n$ .) Applying this estimate in (3.4) we now have that for all  $t \geq \bar{D}$  and  $x \in S$ ,

$$\mathbb{E}|\widehat{Q}_x^n(t|x)|^2 \leq c_3(1+t|x|). \quad (3.6)$$

The result now follows on choosing  $t_0 = \bar{D}$ . ■

The proof of the following result is analogous to that of Lemma 4.4 of [2]. For completeness, we give a sketch in the Appendix. For  $M \in (0, \infty)$ , let  $S_M \doteq \{x \in S : |x| \leq M\}$ .

**Proposition 3.2.** *Let  $\widehat{Q}_x^n$  be defined by (2.8) with  $\widehat{Q}^n(0) \equiv x \in S$  and  $U^n \in \mathcal{A}_n$ . Then there exists a  $\varpi_1 \in (0, \infty)$  such that*

$$\sup_{n \geq 1} \sup_{U^n \in \mathcal{A}_n} \sup_{x \in S_M} \sup_{t \geq 0} \mathbb{E} e^{\varpi_1 |\widehat{Q}_x^n(t)|} < \infty, \text{ for all } M > 0.$$

*In particular, for every  $M \in (0, \infty)$ ,  $\{\widehat{Q}_x^n(t) : n \geq 1, t > 0, x \in S_M\}$  is a tight family of random variables.*

#### 4. Uncontrolled case: Convergence of invariant distributions

This section will be concerned with the uncontrolled setting, namely, the case where  $U^n \equiv 0$ . Throughout this section we will make the assumption that  $\bar{\alpha} < 0$  and so  $0 \in \Lambda^{\delta_0}$ , for some  $\delta_0 > 0$ . Also  $\widehat{Q}^n$  will be given by (2.8) with  $U^n = 0$ . It is well known that  $\widehat{Q}^n$  is a strong Markov process with state space  $S$ . Furthermore, Proposition 3.2 shows that for each  $n \geq 1$  and  $x \in S$ ,  $\{\widehat{Q}_x^n(t) : t \geq 0\}$  forms a tight family of random variables. In particular, the Markov process  $\widehat{Q}^n$  admits an invariant probability distribution. We denote such a distribution by  $\pi_n$ . We now present a heavy traffic limit theorem from [24].

Given  $\rho \in \mathcal{P}(S)$ , let  $Z_\rho$  be a continuous  $\{\overline{\mathcal{F}}_t\}$  adapted process given on some filtered probability space  $(\overline{\Omega}, \overline{\mathcal{F}}, \{\overline{\mathcal{F}}_t\}, \overline{\mathbb{P}})$ , supporting a  $K$ -dimensional standard Brownian motion  $\overline{B}$ , such that  $Z_\rho$  solves

$$Z_\rho(t) = \Gamma \left( Z_\rho(0) + \int_0^t a(Z_\rho(s)) ds + \int_0^t \sigma(Z_\rho(s)) d\overline{B}_s \right) (t), \quad \overline{\mathbb{P}} \circ Z_\rho^{-1}(0) = \rho.$$

Existence and uniqueness in law of such a process follows from the uniform nondegeneracy and Lipschitz continuity of  $\sigma$ . The following Theorem follows from Theorem 1 of [24].

**Theorem 4.1.** *Assume that  $\widehat{Q}^n(0)$  converges in distribution to some  $\rho \in \mathcal{P}(S)$ . Then  $\widehat{Q}^n$  converges weakly in  $D([0, \infty), S)$  to  $Z_\rho$ .*

**Remarks on the Proof.** The proof follows that of Theorem 1 of [24]. We note that conditions in [24] are somewhat different from those considered here. The paper [24] imposes a restrictive form of state dependence on the rates  $\lambda^n, \mu^n$ ; namely it requires for  $x = (x_1, \dots, x_K)' \in S$ ,  $\lambda_i^n(x) = \bar{\lambda}_i^n(x_i)$ ,  $\mu_i^n(x) = \bar{\mu}_i^n(x_i)$  for suitable maps  $\bar{\lambda}_i^n, \bar{\mu}_i^n$  from  $\mathbb{R}_+$  to  $\mathbb{R}_+$ . However, [24] imposes weaker convergence requirements on these rates than Assumption 2.1 (v), (vi) of the current paper (see A2, A4 in [24]). Also, the limit process in [24] is expressed in terms of a  $[K + K(K + 1)]$  dimensional Brownian motion and the diffusion coefficient matrix  $\Sigma$ . One finds that the proofs in [24] can be carried over to the current setting with only minor modifications. In particular, noting that  $\Sigma\Sigma' = \sigma\sigma'$ , standard martingale representation arguments show that by suitably augmenting the probability space one can describe the limit using a  $K$  dimensional Brownian motion and with the diffusion coefficient matrix  $\sigma$ .

The collection  $\{Z_\rho : \rho \in \mathcal{P}(S)\}$  describes a strong Markov process, stability properties of which have been studied in [2]. In particular, recalling that  $R^{-1}a(x) \leq -\delta_0$  for all  $x \in S$ , we have from Theorem 2.2 of [2], that the above Markov process has a unique invariant probability measure. Henceforth, this probability measure will be denoted by  $\pi$ . We remark here that Theorem 2.2 of [2] was stated under an additional Lipschitz assumption on the drift vector  $a$ . However, an examination of the proof shows that this assumption is not needed for the proof (cf. [7]). In fact, due to the uniform nondegeneracy of  $\sigma$ , the result continues to hold if  $a$  is merely bounded and measurable.

We now present the proof of Theorem 2.9. We begin with the following propositions, the proofs of which, being very similar to the proof of Theorem 3.5 of [9], are relegated to the Appendix. For  $\bar{\delta} \in (0, \infty)$  and a compact set  $C \subseteq S$ , let  $\tau_C^n(\bar{\delta}) \doteq \inf\{t \geq \bar{\delta} : \widehat{Q}^n(t) \in C\}$ .

**Proposition 4.2.** *For  $f : S \rightarrow \mathbb{R}_+$ ,  $\bar{\delta} \in (0, \infty)$ , and a compact set  $C \subseteq S$ , define*

$$G_n(x) \doteq \mathbb{E} \left[ \int_0^{\tau_C^n(\bar{\delta})} f(\widehat{Q}_x^n(t)) dt \right], \quad x \in S.$$

*If  $\sup_n G_n$  is everywhere finite and bounded on  $C$ , then there exists a  $\bar{\kappa} \in (0, \infty)$  such that for all  $n \in \mathbb{N}$ ,  $t > 0$ ,  $x \in S$*

$$\frac{1}{t} \mathbb{E}[G_n(\widehat{Q}_x^n(t))] + \frac{1}{t} \int_0^t \mathbb{E}[f(\widehat{Q}_x^n(s))] ds \leq \frac{1}{t} G_n(x) + \bar{\kappa}. \quad (4.1)$$

**Proposition 4.3.** *For some constants  $c, \bar{\delta} \in (0, \infty)$ , and a compact set  $C \subseteq S$ ,*

$$\sup_n \mathbb{E} \left[ \int_0^{\tau_C^n(\bar{\delta})} (1 + |\widehat{Q}_x^n(t)|) dt \right] \leq c(1 + |x|^2), \quad x \in S. \quad (4.2)$$

**Proof of Theorem 2.9.** The proof is similar to that of Theorems 3.1 and 3.2 of [9]. To keep the presentation self contained we provide the details. Since  $\pi$  is the unique invariant

probability measure for  $\{Z_\rho : \rho \in \mathcal{P}(S)\}$ , we have from Theorem 4.1 that it suffices to establish the tightness of the family  $\{\pi_n\}$ . We apply Proposition 4.2 with  $f(x) \doteq 1 + |x|$  for  $x \in S$  and  $\bar{\delta}$ ,  $C$  as in Proposition 4.3. To prove the desired tightness it suffices to show that for all  $n \in \mathbb{N}$ ,  $\langle \pi_n, f \rangle \doteq \int_S f(x) \pi_n(dx) \leq \bar{\kappa}$ . Note that for any nonnegative, real measurable function  $\xi$  on  $S$ ,

$$\int_S \mathbb{E}[\xi(\widehat{Q}_x^n(t))] \pi_n(dx) = \langle \pi_n, \xi \rangle. \quad (4.3)$$

Fix  $k \in \mathbb{N}$  and  $t \in (0, \infty)$ . Let  $G_n^k(x) \doteq G_n(x) \wedge k$  and

$$\Psi_n^k(x) \doteq \frac{1}{t} G_n^k(x) - \frac{1}{t} \mathbb{E}[G_n^k(\widehat{Q}_x^n(t))].$$

From (4.3), we have that  $\int_S \Psi_n^k(x) \pi_n(dx) = 0$ . Let  $\Psi_n(x) \doteq \frac{1}{t} G_n(x) - \frac{1}{t} \mathbb{E}[G_n(\widehat{Q}_x^n(t))]$ . Monotone convergence theorem yields that  $\Psi_n^k(x) \rightarrow \Psi_n(x)$  as  $k \rightarrow \infty$ . Next we will show that  $\Psi_n^k$  is bounded from below uniformly in  $n$  and  $k$ . If  $G_n(x) \leq k$ ,

$$\Psi_n^k(x) = \frac{1}{t} G_n^k(x) - \frac{1}{t} \mathbb{E}[G_n^k(\widehat{Q}_x^n(t))] \geq \frac{1}{t} G_n(x) - \frac{1}{t} \mathbb{E}[G_n(\widehat{Q}_x^n(t))] \geq -\bar{\kappa},$$

where the last inequality follows from (4.1). On the other hand, if  $G_n(x) \geq k$

$$\Psi_n^k(x) = \frac{1}{t} k - \frac{1}{t} \mathbb{E}[G_n^k(\widehat{Q}_x^n(t))] \geq 0, \quad (4.4)$$

where the second inequality follows on noting that  $G_n^k \leq k$ . Hence  $\Psi_n^k(x) \geq -\bar{\kappa}$  for all  $x \in S$ . By an application of Fatou's Lemma we conclude that

$$\int_S \Psi_n(x) \pi_n(dx) \leq \liminf_{k \rightarrow \infty} \int_S \Psi_n^k(x) \pi_n(dx) = 0. \quad (4.5)$$

From (4.1),  $\Psi_n(x) \geq \frac{1}{t} \int_0^t \mathbb{E}[f(\widehat{Q}_x^n(s))] ds - \bar{\kappa}$ . Integrating with respect to  $\pi_n$  and combining that with (4.5), we have

$$0 \geq \int_S \Psi_n(x) \pi_n(dx) \geq \frac{1}{t} \int_0^t \int_S \mathbb{E}[f(\widehat{Q}_x^n(s))] \pi_n(dx) ds - \bar{\kappa}.$$

Using (4.3) once more we now have that  $\langle \pi_n, f \rangle \leq \bar{\kappa}$ . This completes the proof.  $\blacksquare$

## 5. Proof of Theorem 2.7

This section is devoted to the proof of the Theorem 2.7. Thus throughout this section we assume  $\sigma(x) \equiv \sigma$ . From Theorem 2.4 we have that there exists a measurable map  $b^* : S \rightarrow \Lambda$  such that  $V = \int_S [k(x) + c \cdot b^*(x)] \eta_{b^*}(dx)$ . Also, from the same theorem, if  $b : S \rightarrow \Lambda$  is a measurable map and  $U$  is replaced with  $U_b$  in (2.14), the corresponding state process is a strong Markov process that admits a unique stationary probability distribution  $\eta_b$  and  $J(U_b, x) = \langle \eta_b, k_b \rangle$  for all  $x \in S$ . Thus in order to prove the theorem it suffices to show that there is a sequence of continuous maps  $b_n : S \rightarrow \Lambda$  such that  $\eta_{b_n} \Rightarrow \eta_{b^*}$  and  $\langle \eta_{b_n}, k_{b_n} \rangle \Rightarrow \langle \eta_{b^*}, k_{b^*} \rangle$ . We begin with the following lemma. Let  $\gamma_g \in \mathcal{P}(S)$  be defined as  $\gamma_g(A) \doteq \bar{c} \int_A e^{-\frac{|x|^2}{2}} dx$ , where  $A \in \mathcal{B}(S)$  and  $\bar{c}$  is the normalization constant.

**Lemma 5.1.** *For each  $n \in \mathbb{N}$ , there exist  $b_n \in C(S, \Lambda)$  and compact sets  $A_n \subseteq S$  such that  $b_n$  is continuous,*

$$\{x \in S : b^*(x) \neq b_n(x)\} \subseteq A_n^c \quad \text{and} \quad \gamma_g(A_n^c) \leq \frac{1}{2^{n+1}}. \quad (5.1)$$

**Proof.** From Lusin's Theorem (see, e.g., Theorem 2.24 in [22]) one can find a continuous function  $\hat{b}_n : S \rightarrow \mathbb{R}^K$  such that (5.1) is satisfied with  $b_n$  replaced by  $\hat{b}_n$ . Note that  $\Lambda$  is a convex closed subset of  $\mathbb{R}^K$ . Let  $\Pi_\Lambda : \mathbb{R}^K \rightarrow \Lambda$  be the projection map. Then  $\Pi_\Lambda$  is a Lipschitz function. The result follows on setting  $b_n(x) \doteq \Pi_\Lambda(\hat{b}_n(x))$ ,  $x \in S$ . ■

Let  $\{\rho_n\} \subseteq \mathcal{P}(S)$  be such that  $\rho_n \Rightarrow \rho$ , for some  $\rho \in \mathcal{P}(S)$ . Given on some filtered probability space  $\Xi_n \doteq (\Omega^n, \mathcal{F}^n, \{\mathcal{F}_t^n\}, \mathbb{P}^n)$ , let  $X^n$  be the unique weak solution of

$$X^n(t) = \Gamma \left( X^n(0) + \int_0^t (a - Rb_n)(X^n(s)) ds + \sigma W^n(\cdot) \right) (t), \quad X^n(0) \sim \rho_n, \quad (5.2)$$

where  $b_n$  is as in Lemma 5.1 and  $W^n$  is a  $K$ -dimensional  $\{\mathcal{F}_t^n\}$  standard Brownian motion. Let

$$Y^n(t) \doteq \Gamma_1 \left( X^n(0) + \int_0^t (a - Rb_n)(X^n(s)) ds + \sigma W^n(\cdot) \right) (t), \quad (5.3)$$

where  $\Gamma_1(\cdot)$  is defined below Definition 2.2. Expectation operator with respect to the probability measure  $\mathbb{P}^n$  will be denoted by  $\mathbb{E}^n$ .

**Theorem 5.2.** *Let  $\rho_n, \rho, b_n$  and  $(X^n, Y^n)$  be as above. Then  $(X^n, Y^n) \Rightarrow (X, Y)$  as a sequence of  $C([0, \infty), \mathbb{R}^K \times \mathbb{R}^K)$ -valued random variables, where  $X$  is a continuous  $\{\bar{\mathcal{F}}_t\}$  adapted process, on some filtered probability space  $(\bar{\Omega}, \bar{\mathcal{F}}, \{\bar{\mathcal{F}}_t\}_{t \geq 0}, \bar{\mathbb{P}})$ , such that  $\bar{\mathbb{P}} \circ X_0^{-1} = \rho$  and*

$$\begin{aligned} X(t) &= \Gamma \left( X(0) + \int_0^t (a - Rb^*)(X(s)) ds + \sigma W(\cdot) \right) (t), \quad t \geq 0 \\ Y(t) &= \Gamma_1 \left( X(0) + \int_0^t (a - Rb^*)(X(s)) ds + \sigma W(\cdot) \right) (t), \quad t \geq 0, \end{aligned} \quad (5.4)$$

where  $W$  is a standard Brownian motion adapted to  $\{\bar{\mathcal{F}}_t\}$ .

Proof of the above theorem is given immediately after the proof of Theorem 2.7.

**Proof of Theorem 2.7.** It suffices to show that, as  $n \rightarrow \infty$ ,

$$\langle \eta_{b_n}, k_{b_n} \rangle \rightarrow \langle \eta_{b^*}, k_{b^*} \rangle. \quad (5.5)$$

From Lemma 7.1 of [7], the family  $\{\eta_{b_n} : n \geq 1\}$  is tight. Let  $\eta^0$  be a limit point of the sequence  $\{\eta_{b_n}\}$ . Denote the subsequence, along which  $\eta_{b_n}$  converges to  $\eta^0$ , again by  $\{\eta_{b_n}\}$ .

Let  $X^n, Y^n$  be as in (5.2), (5.3) with  $\rho_n$  there replaced by  $\eta_{b_n}$ . From Theorem 5.2 we get that  $X^n \Rightarrow X$ , where  $X$  is given by (5.4) with  $\rho$  replaced by  $\eta^0$ . Since  $X^n$  is stationary, the

same is true for  $X$ . Thus from Theorem 3.4 of [7] we get that  $\eta^0 = \eta_{b^*}$ . This proves  $\eta_{b_n} \Rightarrow \eta_{b^*}$  as  $n \rightarrow \infty$ . Next the uniform boundedness of  $a - Rb_n$  and (5.3) imply that

$$\sup_n \mathbb{E}^n \left( \sup_{0 \leq t \leq 1} |Y^n(t)|^2 \right) < \infty.$$

In particular, for all  $t \in [0, 1]$  we have  $|\mathbb{E}^n Y^n(t) - \overline{\mathbb{E}} Y(t)| \rightarrow 0$  as  $n \rightarrow \infty$ , where  $\overline{\mathbb{E}}$  denotes expectation with respect to the probability measure  $\overline{\mathbb{P}}$ . Also,

$$\left| \mathbb{E}^n \int_0^t a(X^n(s)) ds - \overline{\mathbb{E}} \int_0^t a(X(s)) ds \right| \rightarrow 0, \text{ as } n \rightarrow \infty.$$

Thus taking expectation in (5.2), (5.4), and using the stationarity of  $X^n$  and  $X$ , we get as  $n \rightarrow \infty$

$$\int_S b_n(x) \eta_{b_n}(dx) \rightarrow \int_S b^*(x) \eta_{b^*}(dx).$$

This proves (5.5) and the result follows. ■

For a Polish space  $\mathcal{T}$ , denote by  $\mathcal{M}(\mathcal{T})$  the space of subprobability measures on  $(\mathcal{T}, \mathcal{B}(\mathcal{T}))$  with the usual topology of weak convergence. Let  $G \doteq [0, 1] \times S \times \Lambda$ , where  $\Lambda$  is as introduced in Section 2.

**Proof of Theorem 5.2.** It suffices to prove the result with  $[0, \infty)$  replaced by  $[0, T]$  where  $T > 0$  is arbitrary. Without loss of generality we can assume  $T = 1$ . We first consider the case  $\rho_n = \delta_{x_n}$  and  $\rho = \delta_x$ , where  $x_n, x \in S$  and  $x_n \rightarrow x$  as  $n \rightarrow \infty$ . For  $t \in [0, 1]$ , define  $m_t^n \in \mathcal{M}(G)$  as

$$m_t^n(A \times B \times C) \doteq \int_0^t 1_A(s) 1_B(X^n(s)) 1_C(b_n(X^n(s))) ds,$$

where  $A \in \mathcal{B}[0, 1]$ ,  $B \in \mathcal{B}(S)$ ,  $C \in \mathcal{B}(\Lambda)$ . Note that  $\{m_t^n\}_{0 \leq t \leq 1}$  is a continuous stochastic process, with values in  $\mathcal{M}([0, 1] \times S \times \Lambda)$ , defined on the filtered probability space  $\Xi_n$ . Furthermore,  $\int_0^t b_n(X^n(s)) ds = \int_G u m_t^n(ds dx du)$  and thus

$$X^n(t) = X^n(0) - R \int_G u m_t^n(ds dx du) + \int_0^t a(X^n(s)) ds + \sigma W^n(t) + RY^n(t).$$

Since  $\Lambda$  is compact and  $a$  is bounded, we can find  $c_1 \in (0, \infty)$  such that for all  $n \geq 1$  and  $0 \leq s \leq t < \infty$ ,

$$|X^n(t) - X^n(s)| + |Y^n(t) - Y^n(s)| \leq c_1 (|t - s| + w_{W^n}(|t - s|)),$$

where for  $g \in C([0, \infty), \mathbb{R}^K)$  and  $\delta > 0$ ,  $w_g(\delta) \doteq \sup_{0 \leq s \leq t < \infty, |t-s| \leq \delta} |g(t) - g(s)|$ . This along with the tightness of  $\{X^n(0)\}_{n \geq 1}$  gives that  $(X^n, Y^n)$  are tight in  $C([0, 1], \mathbb{R}^K \times \mathbb{R}^K)$ . Let  $f \in C_b(G)$ . Then for  $0 \leq s < t < 1$  we have  $|\langle m_t^n, f \rangle - \langle m_s^n, f \rangle| \leq \|f\|_\infty |t - s|$ . This along with the observation  $m_0^n = 0$  shows that  $\{m_t^n\}_{n \geq 1}$  is a tight family of  $C([0, 1], \mathcal{M}(G))$  valued random variables. Thus  $\{(X^n, Y^n, W^n, m^n)\}_{n \geq 1}$  is a tight family of  $C([0, 1], \mathcal{E}_0)$  valued random variables, where  $\mathcal{E}_0 \doteq \mathbb{R}^{3K} \times \mathcal{M}(G)$ .

Consider a weak limit point  $(X, Y, W, m)$  of the sequence  $\{(X^n, Y^n, W^n, m^n)\}_{n \geq 1}$ . Abusing notation, we will denote the subsequence once more with the superscript  $n$ . Denote by  $(\Omega^*, \mathcal{F}^*, \mathbb{P}^*)$  the probability space on which all the limit processes are given, and denote the expectation with respect to this probability as  $\mathbb{E}^*$ . Then  $W$  is a Wiener process and  $(X, Y)$  satisfy  $\mathbb{P}^*$ -a.s.,

$$\begin{aligned} X(t) &= X(0) - R \int_G um_t(ds dx du) + \int_0^t a(X(s))ds + \sigma W(t) + RY(t), \\ Y(t) &= \Gamma_1 \left( X(0) - R \int_G um_t(ds dx du) + \int_0^t a(X(s))ds + \sigma W(\cdot) \right) (t), \end{aligned} \quad (5.6)$$

Also  $\mathbb{P}^* \circ X_0^{-1} = \rho$ .

Define  $\mathcal{F}_t^* = \sigma\{(X(s), Y(s), W(s), m(s)) : 0 \leq s \leq t\}$ . We now show that  $W$  is an  $\{\mathcal{F}_t^*\}$  martingale. It suffices to show for all  $p \geq 1$ ,  $0 \leq t_1 \leq t_2 \leq \dots \leq t_p \leq s \leq t \leq 1$ ,

$$\mathbb{E}^*(\psi(\mathcal{Z}_{t_i}, i \leq p)[W_t - W_s]) = 0, \quad (5.7)$$

where  $\psi \in C_b(\mathcal{E}_0^p, \mathbb{R})$  is arbitrary and  $\mathcal{Z}_t = (X_t, Y_t, W_t, m_t)$ . Left side of (5.7) can be expressed as

$$\lim_{n \rightarrow \infty} \mathbb{E}^n(\psi(\mathcal{Z}_{t_i}^n, i \leq p)[W_t^n - W_s^n]),$$

where  $\mathcal{Z}^n \doteq (X^n, Y^n, W^n, m^n)$ . However, the last expression is clearly 0 since  $W^n$  is an  $\{\mathcal{F}_t^n\}$  martingale and  $\mathcal{Z}^n$  is  $\{\mathcal{F}_t^n\}$  adapted.

We will now argue that for all  $t \in [0, 1]$ ,

$$\int_G um_t(ds dx du) = \int_0^t b^*(X(s))ds, \quad \text{a.s. } \mathbb{P}^*. \quad (5.8)$$

This along with weak uniqueness of solutions to (5.4) will prove the result. We will denote the marginal of  $m_t$  on the  $(i, j)$ -th coordinate by  $\hat{m}_t^{i,j}$ . For example,  $\hat{m}_t^{1,3}$  is a sub-probability measure on  $[0, 1] \times \Lambda$  defined as  $\hat{m}_t^{1,3}(A \times B) = m_t(A \times S \times B)$ . Similarly, marginals of  $m_t$  on the  $i$ -th coordinate will be denoted by  $\hat{m}_t^i$ .

Since  $m_t^n((s_1, s_2] \times S \times \Lambda) = (s_2 \wedge t) - (s_1 \wedge t)$  for all  $0 \leq s_1 \leq s_2 \leq 1$  a.s.  $\mathbb{P}^n$ , we have that  $\hat{m}_t^1(s_1, s_2] = (s_2 \wedge t) - (s_1 \wedge t)$  for all  $0 \leq s_1 \leq s_2 \leq 1$  a.s.  $\mathbb{P}^*$ . Thus for  $f \in C_b[0, 1]$ ,  $\int_0^1 f(s)\hat{m}_t^1(ds) = \int_0^t f(s)ds$ .

Next, for  $f \in C_b(S)$  we have  $\int_G f(x)m_t^n(ds dx du) = \int_0^t f(X^n(s))ds$ . Thus

$$\int_G f(x)m_t(ds dx du) = \int_0^t f(X(s))ds, \quad \text{for all } f \in C_b(S), t \in [0, 1], \text{ a.s. } \mathbb{P}^*. \quad (5.9)$$

Disintegrating, we can write

$$\int_0^t \left( \int_S f(x)k_t(s, dx) \right) ds = \int_0^t f(X(s))ds,$$

where  $k_t : \Omega^* \times [0, t] \rightarrow \mathcal{P}(S)$  is a measurable map satisfying

$$\int_A k_t(\omega, s, B) ds = \hat{m}_t^{1,2}(\omega, A \times B) = \int_A 1_B(X(s)) ds, \quad \forall A \times B \subseteq \mathcal{B}([0, 1] \times S), \text{ a.s. } \mathbb{P}^*.$$

Thus

$$k_t(\omega, s, dx) = \delta_{X_s(\omega)}(dx) \text{ a.e. } (s, \omega) \text{ } [\hat{m}_t \otimes \mathbb{P}^*]. \quad (5.10)$$

Recall the definition of  $A_n$  given in (5.1) and the definition of  $\gamma_g$  at the beginning of this section. Define  $B_n \doteq \bigcup_{m=n}^{\infty} A_m^c$  and  $E_n \doteq B_n^c$ . Then

$$\gamma_g(E_n) \geq 1 - \frac{1}{2^n} \quad \text{for all } n \geq 1 \quad (5.11)$$

and  $b^*(x) = b_n(x) = b_{n+1}(x) = \dots$  for all  $x \in E_n$ . Let  $F_1 \subseteq S$  be a compact set such that  $\{x_n\} \subseteq F_1$ . Fix  $\epsilon > 0$ . Let  $F$  be a compact set such that  $F_1 \subseteq F \subseteq S$  and

$$\sup_n \sup_{0 \leq t \leq 1} \mathbb{P}^n[X^n(t) \in F^c] < \frac{\epsilon}{2}. \quad (5.12)$$

For  $\delta > 0$ , define  $F^\delta \doteq F \cap \{x \in S : \text{dist}(x, \partial S) \geq \delta\}$ . For  $t > 0$ , let  $p(t, x, z)$  be the transition probability density function of  $X_0^x(t) \doteq \Gamma(x + \sigma W(\cdot))(t)$  as given in Theorem 7.1. Then from this theorem, we have that for each  $\delta > 0$ , there is a function  $\Psi_\delta : [0, 1] \rightarrow \mathbb{R}_+$  and  $\alpha > 0$  such that

$$\sup_{x \in F, z \in F^\delta} p(t, x, z) \leq \Psi_\delta(t), \quad t \in (0, 1), \quad (5.13)$$

$$\int_0^1 e^{-\alpha/t} \Psi_\delta(t) dt < \infty.$$

Henceforth, fix  $\delta > 0$ . Let  $Q_n^{x_n}, Q_0^{x_n}$  be probability measures induced by  $X^n, X_0^{x_n}$  on  $C([0, 1], S)$ . Then by Girsanov's theorem, uniform boundedness of  $b_n$  and nondegeneracy of  $\sigma$ , there exists a  $\theta \in (0, \infty)$  such that for all  $t \in [0, 1]$

$$Q_n^{x_n} \circ \pi_t^{-1}(A) \leq \theta \sqrt{Q_0^{x_n} \circ \pi_t^{-1}(A)}, \quad \forall n \geq 1 \text{ and } A \in \mathcal{B}(S), \quad (5.14)$$

where  $\pi_t$  is the usual coordinate map from  $C([0, 1], S)$  to  $S$ . Since the Lebesgue measure on  $S$  is absolutely continuous with respect to  $\gamma_g$ , we have from Theorem 7.1 (see (7.38)) and (5.11) that there is  $n_0 \in \mathbb{N}$  such that

$$\lambda(E_{n_0}^c \cap F) \int_0^1 e^{-\alpha/t} \Psi_\delta(t) dt < \frac{\epsilon^2}{4\theta^2}, \quad (5.15)$$

where  $\lambda$  is the Lebesgue measure on  $S$ .

Note that (5.15) together with (5.13) implies that for all  $x \in F, t \in [0, 1]$

$$\mathbb{E}^* \int_{[0, t]} e^{-\alpha/s} 1_{E_{n_0}^c \cap F^\delta}(X_0^x(s)) ds \leq \frac{\epsilon^2}{4\theta^2}. \quad (5.16)$$

Let  $S_\delta \doteq \{x \in S : \text{dist}(x, \partial S) < \delta\}$ . From (5.14) and (5.12) we now have that

$$\begin{aligned}
\int_0^1 e^{-\alpha/2s} Q_n^{x_n} \circ \pi_s^{-1}(E_{n_0}^c) ds &= \int_0^1 e^{-\alpha/2s} Q_n^{x_n} \circ \pi_s^{-1}(E_{n_0}^c \cap F) ds + \frac{\epsilon}{2} \\
&\leq \theta \int_0^1 e^{-\alpha/2s} \sqrt{Q_0^{x_n} \circ \pi_s^{-1}(E_{n_0}^c \cap F)} ds + \frac{\epsilon}{2} \\
&\leq \theta \left( \int_0^1 e^{-\alpha/s} Q_0^{x_n} \circ \pi_s^{-1}(E_{n_0}^c \cap F) ds \right)^{1/2} + \frac{\epsilon}{2} \\
&\leq \theta \left( \int_0^1 e^{-\alpha/s} Q_0^{x_n} \circ \pi_s^{-1}(E_{n_0}^c \cap F^\delta) ds \right)^{1/2} \\
&\quad + \theta \left( \int_0^1 e^{-\alpha/s} Q_0^{x_n} \circ \pi_s^{-1}(S_\delta) ds \right)^{1/2} + \frac{\epsilon}{2} \\
&\leq \epsilon + \ell(x_n, \delta), \tag{5.17}
\end{aligned}$$

where  $\ell(x_n, \delta) = \theta \left( \int_0^1 e^{-\alpha/s} Q_0^{x_n} \circ \pi_s^{-1}(S_\delta) ds \right)^{1/2}$  and the last step follows from (5.16). From Feller property of  $\{X_0^x\}$  and the fact that  $X_0^x(t)$  has a density for every  $t > 0$ , we have that

$$\ell(x_n, \delta) \rightarrow \ell(x, \delta) \text{ as } n \rightarrow \infty. \tag{5.18}$$

Let  $\bar{m}_t^n(A \times B) \doteq m_t^n(A \times B \times \Lambda)$ ,  $A \in \mathcal{B}([0, t])$ ,  $B \in \mathcal{B}(S)$ . For  $n, n_0 \in \mathbb{N}$ ,  $n \geq n_0$  and  $f \in C(\Lambda)$ ,  $h \in C[0, 1]$ ,

$$\begin{aligned}
\int_G e^{-\alpha/2s} h(s) f(u) m_t^n(ds dx du) &= \int_{[0,1] \times S} e^{-\alpha/2s} h(s) f(b_n(x)) \bar{m}_t^n(ds dx) \\
&= \int_{[0,1] \times E_{n_0}} e^{-\alpha/2s} h(s) f(b_{n_0}(x)) \bar{m}_t^n(ds dx) + \int_{[0,1] \times E_{n_0}^c} e^{-\alpha/2s} h(s) f(b_n(x)) \bar{m}_t^n(ds dx).
\end{aligned}$$

Thus we have

$$\begin{aligned}
&\left| \int_G e^{-\alpha/2s} h(s) f(u) m_t^n(ds dx du) - \int_{[0,1] \times S} e^{-\alpha/2s} h(s) f(b_{n_0}(x)) \bar{m}_t^n(ds dx) \right| \\
&\leq 2 \|f\|_\infty \|h\|_\infty \int_{[0,1] \times E_{n_0}^c} e^{-\alpha/2s} \bar{m}_t^n(ds dx). \tag{5.19}
\end{aligned}$$

From (5.17), the expected value of left side of (5.19) is bounded above by  $2 \|f\|_\infty \|h\|_\infty [\epsilon + \ell(x_n, \delta)]$ . Now, letting  $n \rightarrow \infty$  in (5.19) we obtain from (5.18)

$$\begin{aligned}
&\mathbf{E}^* \left| \int_G e^{-\alpha/2s} h(s) f(u) m_t(ds dx du) - \int_{[0,1] \times S} e^{-\alpha/2s} h(s) f(b_{n_0}(x)) \bar{m}_t(ds dx) \right| \\
&\leq 2 \|f\|_\infty \|h\|_\infty [\epsilon + \ell(x, \delta)],
\end{aligned}$$

where  $\bar{m}_t = \hat{m}_t^{1,2}$ . Thus noting that  $b_{n_0}(x) = b^*(x)$  on  $E_{n_0}$  we get

$$\begin{aligned}
&\mathbf{E}^* \left| \int_G e^{-\alpha/2s} h(s) f(u) m_t(ds dx du) - \int_{[0,1] \times S} e^{-\alpha/2s} h(s) f(b^*(x)) \bar{m}_t(ds dx) \right| \\
&\leq 2 \|f\|_\infty \|h\|_\infty \left[ \epsilon + \ell(x, \delta) + \mathbf{E}^* \int_{[0,1] \times E_{n_0}^c} e^{-\alpha/2s} \bar{m}_t(ds dx) \right]. \tag{5.20}
\end{aligned}$$

Since  $E_{n_0}^c$  is an open set, we have from (5.17) and (5.18)

$$\mathbb{E}^* \int_{[0,1] \times E_{n_0}^c} e^{-\alpha/2s} \bar{m}_t(ds dx) \leq \liminf_{n \rightarrow \infty} \mathbb{E}^n \int_{[0,1] \times E_{n_0}^c} e^{-\alpha/2s} \bar{m}_t^n(ds dx) < \epsilon + \ell(x, \delta).$$

Noting that  $\ell(x, \delta) \rightarrow 0$  as  $\delta \rightarrow 0$ , we have, letting  $\epsilon \rightarrow 0$ ,  $\delta \rightarrow 0$  in (5.20), for all  $t \in (0, 1)$ ,  $h \in C[0, 1]$  and  $f \in C(\Lambda)$ ,

$$\int_G e^{-\alpha/2s} h(s) f(u) m_t(ds dx du) = \int_{[0,1] \times S} e^{-\alpha/2s} h(s) f(b^*(x)) \bar{m}_t(ds dx) \quad \text{a.e. } \mathbb{P}^*.$$

Combining this with (5.10) we now obtain for all  $t \in [0, 1]$ ,

$$\mathbb{P}^* \otimes m_t \{(\omega, s, X_s(\omega), b^*(X_s(\omega))) : \omega \in \bar{\Omega}, s \in [0, t]\} = 1.$$

As an immediate consequence, we obtain (5.8). This proves the result for the case  $\rho_n = \delta_{x_n}$  and  $\rho = \delta_x$ .

Next let  $\rho_n, \rho$  be arbitrary such that  $\rho_n \Rightarrow \rho$ . Let  $\widehat{Q}_n^{\rho_n}$  (resp.  $\widehat{Q}^\rho$ ) be the measure induced by  $(X^n, Y^n)$  (resp.  $(X, Y)$ ) on  $C([0, 1], S \times \mathbb{R}^K)$ , where  $(X^n, Y^n)$  is as in (5.2), (5.3) and  $(X, Y)$  as in (5.4). We will write  $\widehat{Q}_n^{\rho_n}$  as  $\widehat{Q}_n^{x_n}$  when  $\rho_n = \delta_{x_n}$ . Similarly we will write  $\widehat{Q}^x$  for  $\widehat{Q}^\rho$  when  $\rho = \delta_x$ . We then have from the first part of the proof that for all  $f \in C_b(C([0, 1], S \times \mathbb{R}^K))$

$$\sup_{x \in F} |\langle f, \widehat{Q}_n^{x_n} \rangle - \langle f, \widehat{Q}^x \rangle| \rightarrow 0, \quad \text{as } n \rightarrow \infty, \text{ whenever } x_n \rightarrow x,$$

which implies  $\sup_{x \in F} |\langle f, \widehat{Q}_n^x \rangle - \langle f, \widehat{Q}^x \rangle| \rightarrow 0$ , as  $n \rightarrow \infty$ , for all compact subset  $F \subseteq S$ .

Using this along with the continuity of the map  $x \mapsto \langle f, \widehat{Q}^x \rangle$ , for  $f \in C_b(C([0, 1], S \times \mathbb{R}^K))$ , and the weak convergence of  $\rho_n$  to  $\rho$ , we have, as  $n \rightarrow \infty$ ,

$$|\langle f, \widehat{Q}_n^{\rho_n} \rangle - \langle f, \widehat{Q}^\rho \rangle| \leq \int_S |\langle f, \widehat{Q}_n^x \rangle - \langle f, \widehat{Q}^x \rangle| \rho_n(dx) + \left| \int_S \langle f, \widehat{Q}^x \rangle \rho_n(dx) - \int_S \langle f, \widehat{Q}^x \rangle \rho(dx) \right| \rightarrow 0.$$

The result follows. ■

## 6. Convergence of value functions

In this section we present the proofs of Theorems 2.5 and 2.6. The proofs rely on the functional occupation measure approach developed in [19]. We begin with some notation and definitions.

For a Polish space  $E$ , let  $E_{\text{path}} \doteq D([0, \infty), E)$ . Note that  $E_{\text{path}}$  is a Polish space with the usual Skorohod topology. For an  $E$ -valued stochastic process  $\{z(t) : t \geq 0\}$  with paths in  $D([0, \infty), E)$ , define,  $E_{\text{path}}$  valued stochastic processes  $\{z_p(t) : t \geq 0\}$ ,  $\{\Delta z_p(t) : t \geq 0\}$  with paths in  $D([0, \infty), E_{\text{path}})$  as

$$z_p(t, \omega)(s) \doteq z(t + s, \omega), \quad \Delta z_p(t, \omega)(s) \doteq z(t + s, \omega) - z(t, \omega), \quad s, t \geq 0.$$

We rewrite (2.8) as follows

$$\begin{aligned}\widehat{Q}^n(t) &= \widehat{Q}^n(0) + \int_0^t \widehat{a}^n(\sqrt{n}\widehat{Q}^n(s))ds - \int_0^t R\widehat{U}^n(s)ds + \widehat{M}^n(t) + R\widehat{Y}^n(t) \\ &\equiv \widehat{Q}^n(0) + \widehat{A}^n(t) - R\widehat{C}^n(t) + \widehat{M}^n(t) + R\widehat{Y}^n(t).\end{aligned}\quad (6.1)$$

Suppose that  $\widehat{Q}^n(0) = x_n$  for  $n \geq 1$  and  $\sup_n |x_n| \leq M$ . Let  $\{\widehat{H}_p^n(t), t \geq 0\}$  be a stochastic process with paths in  $D([0, \infty), \mathbb{S}_{\text{path}})$  defined as

$$\widehat{H}_p^n(t) \doteq (\widehat{Q}_p^n(t), \Delta\widehat{A}_p^n(t), \Delta\widehat{C}_p^n(t), \Delta\widehat{M}_p^n(t), \Delta\widehat{Y}_p^n(t)),$$

where  $\mathbb{S} \doteq (S \times \mathbb{R}^{4K})$ . Note that we have suppressed the dependence of the processes on  $x_n$  in our notation. Then for any  $t, s \geq 0$ , (6.1) can be rewritten as

$$\widehat{Q}_p^n(t)(s) = \widehat{Q}^n(t+s) = \Gamma\left(\widehat{Q}_p^n(t) + \Delta\widehat{A}_p^n(t) - R\Delta\widehat{C}_p^n(t) + \Delta\widehat{M}_p^n(t)\right)(s). \quad (6.2)$$

Recall, we say a family of random variables is tight, if the corresponding collection of laws (i.e., induced measures) is tight.

**Proposition 6.1.** *Let  $(n_\ell, T_\ell)_{\ell=1}^\infty$  be a sequence such that  $n_\ell \rightarrow \infty$ ,  $T_\ell \rightarrow \infty$  as  $\ell \rightarrow \infty$ . Let  $\{\mathcal{Q}_\ell : \ell \geq 1\}$  be a sequence of  $\mathcal{P}(\mathbb{S}_{\text{path}})$  valued random variables defined as*

$$\mathcal{Q}_\ell(F) \doteq \frac{1}{T_\ell} \int_0^{T_\ell} 1_F(\widehat{H}_p^{n_\ell}(s))ds,$$

where  $F \in \mathcal{B}(\mathbb{S}_{\text{path}})$ . Then  $\{\mathcal{Q}_\ell : \ell \geq 1\}$  is a tight family of random variables.

Proof of the proposition is based on the following well known result (see, e.g., Theorem 5.4 in [19]).

**Lemma 6.2.** *Suppose that the sequence  $\{\mathbb{E}\mathcal{Q}_n\}_{n \geq 1}$  is tight family of probability measures on  $\mathbb{S}_{\text{path}}$ . Then  $\{\mathcal{Q}_n : n \geq 1\}$  is a tight family of random variables.*

**Proof of Proposition 6.1.** From Lemma 6.2, it suffices to show that the collection

$$\{\widehat{H}_p^{n_\ell}(t) : \ell \geq 1, t \geq 0\} \quad (6.3)$$

is a tight family of  $\mathbb{S}_{\text{path}}$  valued random variables. Tightness of  $\{(\Delta\widehat{A}_p^{n_\ell}(t), \Delta\widehat{C}_p^{n_\ell}(t)) : \ell \geq 1, t \geq 0\}$  is immediate on recalling the uniform boundedness of  $\widehat{a}^n$  and the compactness of  $\Lambda$ . Next we argue the tightness of

$$\{\Delta\widehat{M}_p^{n_\ell}(t) : \ell \geq 1, t \geq 0\}. \quad (6.4)$$

Note that for  $\ell \geq 1, t \geq 0$ ,  $\Delta\widehat{M}_p^{n_\ell}(t)$  is a martingale (with respect to its own filtration). Furthermore, if  $\tau$  is a bounded stopping time (with respect to this filtration), we have that for  $\beta > 0$ ,

$$\mathbb{E}|\Delta\widehat{M}_p^{n_\ell}(t)(\tau + \beta) - \Delta\widehat{M}_p^{n_\ell}(t)(\tau)|^2 \leq c_1\beta, \quad (6.5)$$

for some  $c_1 \in (0, \infty)$ . The constant  $c_1$  can be chosen to be uniform in  $\ell$  and  $t$ . (It may depend on the upper bound on  $\tau$ .) The estimate in (6.5) follows exactly as the one in (3.5) and so the proof

is omitted. Using Aldous's criterion (see, e.g., Theorem 16.10 of [4]) we now have the tightness of the family in (6.4). In fact we have shown that the family  $\{(\Delta \widehat{A}_p^{n_\ell}(t), \Delta \widehat{C}_p^{n_\ell}(t)), \Delta \widehat{M}_p^{n_\ell}(t) : \ell \geq 1, t \geq 0\}$  is  $C$ -tight. Next from Proposition 3.2 we have that

$$\{\widehat{Q}^{n_\ell}(t) : \ell \geq 1, t \geq 0\} \text{ is tight.} \quad (6.6)$$

The tightness of  $\{\widehat{Q}_p^{n_\ell}(t) : \ell \geq 1, t \geq 0\}$  now follows on combining the above with (6.2) and recalling Lipschitz property of  $\Gamma$  given in Proposition 2.3. Similarly, using the Lipschitz property of  $\Gamma_1$ , one has the tightness of  $\{\widehat{Y}_p^{n_\ell}(t) : \ell \geq 1, t \geq 0\}$ . This proves the tightness of (6.3) and the result follows.  $\blacksquare$

Suppose  $\{\mathcal{Q}_\ell : \ell \geq 1\}$  along some subsequence  $\{\ell_m\}_{m \geq 1}$  converges in distribution to  $\widetilde{\mathcal{Q}}$  defined on some probability space  $(\Omega_0, \mathcal{F}_0, \mathbb{P}_0)$ . We will denote expectation under  $\mathbb{P}_0$  by  $\mathbb{E}_0$ , a generic element of  $\Omega_0$  by  $\omega_0$  and write  $\widetilde{\mathcal{Q}}(\omega_0)$  as  $\widetilde{\mathcal{Q}}^{\omega_0}$ . Denote the canonical coordinate process on  $\mathbb{S}_{\text{path}}$  as  $H^* = (Q^*, A^*, C^*, M^*, Y^*)$ . With an abuse of notation, we will also denote a typical element of  $\mathbb{S}_{\text{path}}$  by the same symbol. The following theorem describes the law of  $H^*$  under  $\widetilde{\mathcal{Q}}^{\omega_0}$  for  $\mathbb{P}_0$  a.e.  $\omega_0$ .

**Theorem 6.3.**

(i) For  $\mathbb{P}_0$ -almost every  $\omega_0$ , we have the following properties:

(1) For all  $t \geq 0$ ,

$$Q^*(t) = \Gamma(Q^*(0) + A^*(\cdot) - RC^*(\cdot) + M^*(\cdot))(t), \quad \widetilde{\mathcal{Q}}^{\omega_0} \text{ a.s.} \quad (6.7)$$

(2)  $H^*$  is stationary under  $\widetilde{\mathcal{Q}}^{\omega_0}$  in the following sense: The probability measure

$$(\widetilde{\mathcal{Q}}^{\omega_0}) \circ (Q_p^*(t), \Delta A_p^*(t), \Delta C_p^*(t), \Delta M_p^*(t), \Delta Y_p^*(t))^{-1}$$

is the same for every  $t \geq 0$ .

(3) Under  $\widetilde{\mathcal{Q}}^{\omega_0}$ ,  $M^*$  is a continuous square integrable  $\mathcal{G}_t$ -martingale, where  $\mathcal{G}_t \doteq \sigma\{H^*(s) : s \leq t\}$ .

(4)  $\langle M^*, M^* \rangle_t = \int_0^t [\sigma(Q^*(s))\sigma(Q^*(s))'] ds$ ,  $t \geq 0$ , a.s.  $\widetilde{\mathcal{Q}}^{\omega_0}$ .

(5)  $A^*(t) = \int_0^t a(Q^*(s)) ds$ , for all  $t \geq 0$ , a.s.  $\widetilde{\mathcal{Q}}^{\omega_0}$ .

(6) There is a  $\mathcal{G}_t$  progressively measurable process  $U^*$  with values in  $\Lambda$  such that  $C^*(t) = \int_0^t U^*(s) ds$ , for all  $t \geq 0$ , a.s.  $\widetilde{\mathcal{Q}}^{\omega_0}$ .

(ii) Suppose that  $0 \in \Lambda$  and  $\widehat{Q}^n$  is defined by (6.1) with  $U^n(s) = 0$  for all  $s \geq 0$ . Then conclusions of part (i) continue to hold with  $C^*$  in (6.7) replaced by 0.

The proof of the above result is similar to that of Theorem 6.3 of [19]. For completeness we give a sketch in the Appendix.

**Proof of Theorem 2.5.** In order to prove the result, it suffices to show that for every sequence  $(T_\ell, n_\ell)$  such that  $T_\ell \rightarrow \infty, n_\ell \rightarrow \infty$  as  $\ell \rightarrow \infty$  and arbitrary  $U^{n_\ell} \in \mathcal{A}_{n_\ell}$ ,

$$\liminf_{\ell \rightarrow \infty} \frac{1}{T_\ell} \mathbb{E} \int_0^{T_\ell} [k(\widehat{Q}^{n_\ell}(s)) + c \cdot \widehat{U}^{n_\ell}(s)] ds \geq V. \quad (6.8)$$

From Theorem 2.4 (iv) we have that  $V = \lim_{m \rightarrow \infty} \langle \eta_{b^*}, k \wedge m + c \cdot b^* \rangle$ . Thus in proving (6.8) we can assume without loss of generality that  $k$  is bounded.

Define  $\mathcal{K} : \mathbb{S}_{\text{path}} \rightarrow \mathbb{R}$  as  $\mathcal{K}(H^*) \doteq k(Q^*(0)) + c \cdot C^*(1)$ . Then for  $\ell \geq 1$ ,

$$\frac{1}{T_\ell} \int_0^{T_\ell} [k(\widehat{Q}^{n_\ell}(s)) + c \cdot \widehat{U}^{n_\ell}(s)] ds = \int_{\mathbb{S}_{\text{path}}} \mathcal{K}(H^*) d\mathcal{Q}_\ell(H^*) + \delta_\ell,$$

where

$$\begin{aligned} \delta_\ell &= \frac{1}{T_\ell} \left[ \int_0^1 c \cdot \widehat{U}^{n_\ell}(s)(1-s) ds - \int_{T_\ell}^{T_\ell+1} c \cdot \widehat{U}^{n_\ell}(s)(T_\ell - s + 1) ds \right] \\ &\leq \frac{c_1}{T_\ell} \rightarrow 0, \text{ as } \ell \rightarrow \infty. \end{aligned}$$

By a usual subsequential argument, we can assume without loss of generality that  $\mathcal{Q}_\ell$  converges in distribution to  $\widetilde{\mathcal{Q}}$  as given in Theorem 6.3. Thus

$$\liminf_{\ell \rightarrow \infty} \mathbb{E} \frac{1}{T_\ell} \int_0^{T_\ell} [k(\widehat{Q}^{n_\ell}(s)) + c \cdot \widehat{U}^{n_\ell}(s)] ds = \mathbb{E}_0 \int_{\mathbb{S}_{\text{path}}} \mathcal{K}(H^*) d\widetilde{\mathcal{Q}}(H^*). \quad (6.9)$$

Using stationarity (Theorem 6.3 (i)(2)) and noting that under  $\widetilde{\mathcal{Q}}^{\omega_0}$ ,  $C^*(0) = 0$ , a.s., the right side of (6.9) can be written as

$$\mathbb{E}_0 \left[ \lim_{T \rightarrow \infty} \int_{\mathbb{S}_{\text{path}}} \left( \frac{1}{T} \int_0^T \{k(Q^*(s)) + c \cdot [C^*(s+1) - C^*(s)]\} ds \right) d\widetilde{\mathcal{Q}}(H^*) \right].$$

Using Theorem 6.3 (i)(6), the last expression is the same as

$$\mathbb{E}_0 \left[ \lim_{T \rightarrow \infty} \int_{\mathbb{S}_{\text{path}}} \left( \frac{1}{T} \int_0^T \{k(Q^*(s)) + c \cdot U^*(s)\} ds \right) d\widetilde{\mathcal{Q}}(H^*) \right]. \quad (6.10)$$

By appealing to Martingale Representation Theorem (see e.g. [16], Proposition 6.2) one has, for a.e.  $\omega_0$ , by suitably augmenting the filtered probability space  $(\mathbb{S}_{\text{path}}, \mathcal{B}(\mathbb{S}_{\text{path}}), \{\mathcal{G}_t\}, \widetilde{\mathcal{Q}}^{\omega_0})$ , that  $M_t^* = \int_0^t \sigma(Q^*(s)) dW(s)$ ,  $t \geq 0$ , a.e.  $\widetilde{\mathcal{Q}}^{\omega_0}$ , where  $W$  is a standard  $K$ -dimensional Brownian motion on the augmented filtered probability space. Thus the expression inside the expectation operator in (6.10) represents the cost under some admissible control (and some initial condition) for the diffusion control problem of Section 2. Hence the expression in (6.10) can be bounded below by  $V$ . This proves (6.8) and Theorem 2.5 follows.  $\blacksquare$

We now proceed to the proof of Theorem 2.6.

**Proof of Theorem 2.6.** From Theorem 2.4 we have that when  $U$  is replaced with  $U_{\tilde{b}}$  in (2.14), the corresponding state process is a strong Markov process which admits a unique stationary probability distribution  $\eta_{\tilde{b}}$ . Furthermore, for all  $x \in S$   $J(U_{\tilde{b}}, x) = \langle \eta_{\tilde{b}}, k_{\tilde{b}} \rangle$ . Let  $(T_\ell, n_\ell)_{\ell \geq 1}$  be, as before, a sequence such that  $T_\ell \rightarrow \infty$  and  $n_\ell \rightarrow \infty$  as  $\ell \rightarrow \infty$ . In order to prove the theorem, it suffices to show that, as  $\ell \rightarrow \infty$ ,

$$\mathbb{E} \frac{1}{T_\ell} \int_0^{T_\ell} [k(\widehat{Q}^{n_\ell}(s)) + c \cdot \widehat{U}_{\tilde{b}}^{n_\ell}(s)] ds \rightarrow \langle \eta_{\tilde{b}}, k_{\tilde{b}} \rangle, \quad (6.11)$$

where  $\widehat{Q}^{n_\ell}$  is defined as in (6.1) with  $\widehat{U}^{n_\ell}$  there replaced by  $\widehat{U}_{\tilde{b}}^{n_\ell}$ . Note that the left side of (6.11) can be rewritten as

$$\mathbb{E} \int_{\mathbb{S}_{\text{path}}} k_{\tilde{b}}(Q^*(0)) d\mathcal{Q}_\ell(H^*). \quad (6.12)$$

From Proposition 6.1, we have that  $\{\mathcal{Q}_\ell\}_{\ell \geq 1}$  is a tight family of random variables. Once more one can assume by a subsequential argument that  $\mathcal{Q}_\ell$  converges to some  $\tilde{\mathcal{Q}}$ , defined on some probability space  $(\Omega_0, \mathcal{F}_0, \mathbb{P}_0)$ . From Proposition 3.2 we have that

$$\sup_{\ell \geq 1} \sup_{s \in (0, \infty)} \mathbb{E} |\widehat{Q}^{n_\ell}(s)|^{2m_0} < \infty$$

where  $m_0$  is as in (2.9). Thus the expression in (6.12) converges to

$$\mathbb{E}_0 \int_{\mathbb{S}_{\text{path}}} k_{\tilde{b}}(Q^*(0)) d\tilde{\mathcal{Q}}(H^*). \quad (6.13)$$

We will now apply part (ii) of Theorem 6.3 with  $(a^n, a)$  replaced by  $(a_1^n, a_1)$ , where  $a_1^n(x) \doteq a^n(x) - \sqrt{n}R\tilde{b}(\frac{x}{\sqrt{n}})$ ,  $a_1(x) \doteq a(x) - R\tilde{b}(x)$ ,  $x \in S$  and  $\Lambda$  replaced by  $\tilde{\Lambda}_1$ , where  $\tilde{\Lambda}_1$  is defined analogously to  $\Lambda$  with  $\bar{\alpha}$  replaced by  $\bar{\alpha}_1 \doteq \sup_{x \in S, i \in \mathbb{K}} [R^{-1}a_1(x)]_i$ . Since  $\tilde{b}(x) \in \Lambda$  for all  $x$ , we have that  $0 \in \tilde{\Lambda}_1$ . Thus from part (ii) of Theorem 6.3, for  $\mathbb{P}_0$ -almost every  $\omega_0$ , (1) to (6) in that theorem hold with  $a$  in (5) replaced by  $a_1$  and  $U^*$  in (6) replaced by 0. Once more by suitably augmenting the filtered probability space  $(\mathbb{S}_{\text{path}}, \mathcal{B}(\mathbb{S}_{\text{path}}), \{\mathcal{G}_t\}, \tilde{\mathcal{Q}}^{\omega_0})$  one has the representation, for  $\mathbb{P}_0$ -almost every  $\omega_0$ ,

$$Q^*(t) = \Gamma \left( Q^*(0) + \int_0^\cdot a_1(Q^*(s)) ds + \int_0^\cdot \sigma(Q^*(s)) dW(s) \right) (t), \quad a.s. \tilde{\mathcal{Q}}^{\omega_0},$$

for some  $K$ -dimensional standard Brownian motion  $W$ . Recalling the stationarity property from part (2) of Theorem 6.3 and the uniqueness property of the invariant measure  $\eta_{\tilde{b}}$ , we see that  $Q^*(0)$  has law  $\eta_{\tilde{b}}$  under  $\tilde{\mathcal{Q}}^{\omega_0}$ , for  $\mathbb{P}_0$ -a.e.  $\omega$ . Thus  $\int_{\mathbb{S}_{\text{path}}} k_{\tilde{b}}(Q^*(0)) d\tilde{\mathcal{Q}}(H^*) = \langle \eta_{\tilde{b}}, k_{\tilde{b}} \rangle$  a.e.  $\omega_0$ . Using this observation in (6.13) we have (6.11) and thus the result follows.  $\blacksquare$

## 7. Appendix

**Proof of Claim (2.13).** It is easy to check that for all  $v = (v_1, \dots, v_K)' \in \mathbb{R}^K$  and all  $i, j \in \mathbb{K}$ ,  $v' B_i B_i' v \geq p_{ij} v_j^2 \underline{\mu}$ . Also, for  $j \in \mathbb{K}_e$ ,  $v' A A' v \geq \underline{\lambda} v_j^2$ . Defining  $\kappa_j = \underline{\lambda} 1_{j \in \mathbb{K}_e} +$

$\frac{1}{4}\mu \max_{i \in \mathbb{K}} \{p_{ij}\} 1_{j \notin \mathbb{K}_e}$  we see that for all  $j \in \mathbb{K}$ ,  $v'AA'v \geq \kappa_j v_j^2$ . The claim in (2.13) now follows on defining  $\kappa \doteq \min_{j \in \mathbb{K}} \{\kappa_j\} / K$ .

■

**Comments on the Proof of Theorem 2.4.** Part (i) is Theorem 3.2 of [7]. Finiteness of  $\int_S e^{\varpi|x|} \eta_b(dx)$ , for some  $\varpi \in (0, \infty)$ , for each fixed  $b \in \text{BM}(S, \Lambda)$ , is established in Corollary 5.11 of [8] under the additional assumption that  $Rb - a$  is a Lipschitz function. The Lipschitz assumption is only used in order to ensure unique solutions and the proof of exponential integrability of invariant measures goes through unchanged for the more general setting considered here. Exponential integrability, uniformly over  $b \in \text{BM}(S, \Lambda)$ , as stated in (ii) is a consequence of the fact that  $\delta, \beta, b_2$  and  $C_2$  in Lemma 5.9 of [8] can be chosen (independent of  $b \in \text{BM}(S, \Lambda)$ ) so that the estimate (5.6) of the cited paper holds uniformly over  $b \in \text{BM}(S, \Lambda)$ . Part (iii) is immediate from the above exponential integrability (cf. Lemma 4.1 of [7]). Part (iv) has been established in Theorem 3.4 of [7] for the case where  $k$  is bounded. The main reason for this restrictive condition there was the lack of availability of exponential integrability estimates which were later established in [8]. Using (ii), the proof of (iv) for the general setting considered here is carried out exactly as in [7].

**Proof of Proposition 3.2.** The proof follows along the lines of Lemma 4.4 of [2] and thus only a sketch will be provided. Let  $T$  be as in (3.3). Fix  $\Delta, t \in (0, \infty)$ . Along the lines of Lemma 4.1 of [2] one can show that

$$T(\widehat{Q}_x^n(t + \Delta)) \leq [T(\widehat{Q}_x^n(t)) - \Delta]^+ + c_1 \sum_{i=1}^K \sum_{j=0}^K \sup_{0 \leq s \leq \Delta} |\widehat{M}_{ij}^n(t + s) - \widehat{M}_{ij}^n(t)|, \quad (7.1)$$

where  $c_1 = 2LC$  and  $C$  is as in Lemma 3.1 of [2]. Fix  $q \in \mathbb{N}$  and let

$$S_n(\Delta) \doteq \{j \in \{1, 2, \dots, q-1\} : T(\widehat{Q}_x^n(t)) \leq \Delta, \text{ for some } t \in [(j-1)\Delta, j\Delta]\}.$$

Define  $m = \max\{j : j \in S_n(\Delta)\}$  if  $S_n(\Delta)$  is nonempty. We set  $m = 0$  otherwise. From (7.1) we have, along the lines of Lemma 4.4 of [2],

$$\begin{aligned} T(\widehat{Q}_x^n(q\Delta)) &\leq T(x) + 2\Delta + \sum_{j=m}^q (2c_1\nu_j^n - \Delta) \\ &\leq T(x) + 2\Delta + \max_{1 \leq i \leq q} \sum_{j=i}^q (2c_1\nu_j^n - \Delta), \end{aligned}$$

where for  $1 \leq l \leq q$ ,

$$\nu_l^n \doteq \sum_{i=1}^K \sum_{j=0}^K \sup_{0 \leq s \leq \Delta} |\widehat{M}_{ij}^n((l-1)\Delta + s) - \widehat{M}_{ij}^n((l-1)\Delta)|.$$

Thus for  $M_0, \alpha \in (0, \infty)$

$$\mathbb{P}(T(\widehat{Q}_x^n(q\Delta)) \geq M_0) \leq \exp\{\alpha(T(x) + \Delta - M_0)\} \sum_{i=1}^q \frac{\mathbb{E}(\exp\{2c_1\alpha \sum_{j=i}^q \nu_j^n\})}{\exp\{\alpha(q-i)\Delta\}}. \quad (7.2)$$

We claim that there are positive constants  $\gamma_0, \Delta, \eta, n_0$  such that

$$\sup_{n \geq n_0} \sup_{U^n \in \mathcal{A}_n} \sup_{l \in \mathbb{N}} e^{-\gamma_0 \Delta} \mathbb{E}(e^{\gamma_0 c_2 \nu_l^n} \mid \mathcal{F}_{(l-1)\Delta}) \leq e^{-\eta \Delta}, \quad (7.3)$$

where  $c_2 = 2c_1$ . Given the claim holds, we now have setting  $\alpha = \gamma_0$  in (7.2), for  $n \geq n_0$ ,

$$\mathbb{P}(T(\widehat{Q}_x^n(q\Delta)) \geq M_0) \leq \frac{\exp\{\gamma_0(T(x) + \Delta - M_0)\}}{1 - \exp(-\eta \Delta)}.$$

In view of Lemma 3.1 of [2] we then have that for all  $\gamma_1 < \gamma_0$

$$\sup_{n \geq n_0} \sup_{U^n \in \mathcal{A}_n} \sup_{q \in \mathbb{N}} \sup_{x \in S_M} \mathbb{E} e^{\gamma_1 |\widehat{Q}_x^n(q\Delta)|} < \infty. \quad (7.4)$$

Using Theorem 3.2 of [2] we have, following arguments in Lemma 4.4 of the same paper, that for  $t \in [(q-1)\Delta, q\Delta)$

$$|\widehat{Q}_x^n(t)| \leq L(|\widehat{Q}_x^n((q-1)\Delta)| + \nu_q^n).$$

The result now follows on combining the above display with (7.4) and (7.3).

Finally we prove the claim in (7.3). It suffices to prove (7.3) with  $\nu_l^n$  replaced with  $\nu_{l;i,j}^n$  for each  $(i, j)$ , where  $\nu_{l;i,j}^n = \sup_{0 \leq s \leq \Delta} |\widehat{M}_{ij}^n((l-1)\Delta + s) - \widehat{M}_{ij}^n((l-1)\Delta)|$ . We only consider the case  $l = 1$ . The proof will show that the case of a general  $l$  follows similarly and in fact the bound is uniform in  $l$ . We only consider the case  $j \neq 0$ . Proof for  $j = 0$  follows along similar lines. Without loss of generality, we assume that  $p_{ij} \neq 0$ . From Assumption 2.1 (iii), (viii) and definition of  $\Lambda$ , we can find  $n_0, c_3, c_4 \in (0, \infty)$  such that for all  $x \in S, u \in \Lambda_n, n \geq n_0$ ;

$$nc_3 \leq p_{ij}(\mu_i^n(x) + u_i) \leq nc_4. \quad (7.5)$$

For  $s > 0$ , let  $\tau(s)$  be an  $\{\mathcal{F}_t\}_{t \geq 0}$  stopping time such that

$$p_{ij} \int_0^{\tau(s)} [\mu_i^n(Q^n(t)) + U_i^n(t)] dt = s. \quad (7.6)$$

Let

$$\widetilde{M}_{ij}^n(s) \doteq M_{ij}^n(\tau(s)), \quad s \geq 0. \quad (7.7)$$

Then  $\widetilde{M}_{ij}^n \stackrel{\mathcal{L}}{=} N_0$ , where  $N_0$  is a unit rate compensated Poisson process. (See Theorem T16, Chapter II in [6].)

Note that (7.5)-(7.7) yield  $\sqrt{n} \nu_{ij}^n \leq \sup_{0 \leq s \leq nc_3 \Delta} |\widetilde{M}_{ij}^n(s)|$ . Therefore, for  $\gamma \in (0, \infty)$  we have  $\mathbb{E} e^{\gamma c_2 \sqrt{n} \nu_{ij}^n} \leq \mathbb{E} e^{\gamma c_2 \widehat{\nu}^n}$ , where  $\widehat{\nu}^n \doteq \sup_{0 \leq s \leq nc_3 \Delta} |N_0(s)|$ . Applying Doob's maximal inequality for submartingales, we get

$$\mathbb{E} e^{\gamma c_2 \widehat{\nu}^n} \leq 4 \mathbb{E} e^{\gamma c_2 |N_0(nc_3 \Delta)|}.$$

Assume that  $\gamma$  is small enough so that  $\gamma c_2 < 1$ . Then a straightforward calculation shows that with  $c_5 \doteq \frac{e}{2} c_2^2 c_3$ ,  $\mathbb{E} e^{\gamma c_2 \widehat{\nu}^n} \leq 8e^{c_5 \Delta n \gamma^2}$ . Thus for all  $\gamma \in (0, \infty)$  such that  $\gamma c_2 < 1$ ,

$$e^{-\sqrt{n} \gamma \Delta} \mathbb{E} e^{\gamma c_2 \sqrt{n} \nu_{ij}^n} \leq 8e^{c_5 \Delta n \gamma^2} e^{-\sqrt{n} \gamma \Delta}.$$

Now choose  $\gamma = \frac{\gamma_0}{\sqrt{n}}$ ,  $\gamma_0 \in (0, \frac{1}{c_2})$ . Then

$$e^{-\gamma_0 \Delta} \mathbb{E} e^{\gamma_0 c_2 \nu_{ij}^n} \leq 8 e^{c_5 \Delta \gamma_0^2} e^{-\gamma_0 \Delta}.$$

Now choose  $\gamma_0$  sufficiently small and  $\Delta$  sufficiently large so that for some  $\eta \in (0, 1)$ ,

$$\frac{\log 8}{\Delta} + c_5 \gamma_0^2 - \gamma_0 \leq -\eta.$$

Thus (7.3) (with  $\nu_{ij}^n$  replaced by  $\nu_{ij}^n$ ) holds with such a choice of  $\gamma_0, \Delta, \eta$ . The claim follows.  $\blacksquare$

**Proof of Proposition 4.2.** The proof is adapted from Proposition 5.4 of [10]. We begin by showing that:

$$\text{For all } m \in \mathbb{N}, \quad \mathbb{E} \left[ \int_0^{\tau_C^n(m\bar{\delta})} f(\widehat{Q}_x^n(t)) dt \right] \leq G_n(x) + b_1 m \bar{\delta}, \quad x \in S, \quad (7.8)$$

where  $b_1 \doteq \sup_n \sup_{x \in C} G_n(x) / \bar{\delta}$ . The proof is by induction. For  $m = 1$ , the inequality in (7.8) holds trivially. Suppose now that (7.8) holds for  $m = k \in \mathbb{N}$ . In what follows, instead of indicating the dependence on the initial condition as a subscript to  $\widehat{Q}^n$ , we will indicate it in the expectation operation. For example,  $\mathbb{E}[f(\widehat{Q}_x^n(t))]$  will be written as  $\mathbb{E}_x[f(\widehat{Q}^n(t))]$ , etc. Using the strong Markov property of  $\widehat{Q}^n$  we have that

$$\begin{aligned} \mathbb{E}_x \left[ \int_0^{\tau_C^n((k+1)\bar{\delta})} f(\widehat{Q}^n(t)) dt \right] &\leq \mathbb{E}_x \left[ \int_0^{\tau_C^n(\bar{\delta})} f(\widehat{Q}^n(t)) dt \right] \\ &\quad + \mathbb{E}_x \left[ \mathbb{E}_{\widehat{Q}^n(\tau_C^n(\bar{\delta}))} \left[ \int_0^{\tau_C^n(k\bar{\delta})} f(\widehat{Q}^n(t)) dt \right] \right] \\ &\leq G_n(x) + \sup_{x \in C} \mathbb{E}_x \left[ \int_0^{\tau_C^n(k\bar{\delta})} f(\widehat{Q}^n(t)) dt \right] \\ &\leq G_n(x) + \sup_n \sup_{x \in C} G_n(x) + b_1 k \bar{\delta}, \end{aligned} \quad (7.9)$$

where the last inequality follows from the induction hypothesis. Proof of (7.8) now follows on noting that the right side of (7.9) coincides with  $G_n(x) + b_1(k+1)\bar{\delta}$ . Using the monotonicity in  $m$  of the expression on the left side of (7.8) we now have that for all  $t \geq \bar{\delta}$ ,

$$\mathbb{E}_x \left[ \int_0^{\tau_C^n(t)} f(\widehat{Q}^n(s)) ds \right] \leq G_n(x) + 2b_1 t. \quad (7.10)$$

Note that (7.10) is trivially satisfied for all  $t < \bar{\delta}$ . Thus (7.10) holds for all  $t \geq 0$ . Using the strong Markov property once again, we have (see proof of Proposition 5.4 of [10] for analogous

arguments)

$$\begin{aligned}
\mathbb{E}_x[G_n(\widehat{Q}^n(t))] &\leq G_n(x) - \int_0^t \mathbb{E}_x \left[ f(\widehat{Q}^n(s)) \right] ds \\
&\quad + \mathbb{E}_x \left[ \mathbb{E}_{\widehat{Q}^n(\tau_C^n(\delta))} \left[ \int_0^{\tau_C^n(t)} f(\widehat{Q}^n(s)) ds \right] \right] \\
&\leq G_n(x) - \int_0^t \mathbb{E}_x \left[ f(\widehat{Q}^n(s)) \right] ds + \sup_n \sup_{x \in C} \mathbb{E}_x \left[ \int_0^{\tau_C^n(t)} f(\widehat{Q}^n(s)) ds \right] \\
&\leq G_n(x) - \int_0^t \mathbb{E}_x \left[ f(\widehat{Q}^n(s)) \right] ds + \sup_n \sup_{x \in C} G_n(x) + 2b_1 t, \tag{7.11}
\end{aligned}$$

where the last inequality follows from (7.10). We obtain (4.1) by dividing both sides in (7.11) by  $t$  and setting  $\bar{\kappa} \doteq \left(\frac{2}{\delta} + 1\right) \sup_n \sup_{x \in C} G_n(x)$ .  $\blacksquare$

**Proof of Proposition 4.3.** By Proposition 3.1, there exists  $\bar{L} \in (0, \infty)$  such that with  $C \doteq \{x \in S : |x| \leq \bar{L}\}$ ,

$$\sup_n \mathbb{E}[\widehat{Q}_x^n(t_0|x)^2] \leq \frac{1}{2}|x|^2, \quad \forall x \in C^c, \tag{7.12}$$

where  $t_0$  is as in Proposition 3.1. Let  $\bar{\delta} \doteq t_0 \bar{L}$  and set  $\tau_C^n(\bar{\delta}) \equiv \tau^n \doteq \inf\{t \geq \bar{\delta} : |\widehat{Q}_x^n(t)| \leq \bar{L}\}$ . Define a sequence of stopping times  $\sigma_m$  as

$$\sigma_0 \doteq 0, \quad \sigma_m = \sigma_{m-1} + t_0[|\widehat{Q}_x^n(\sigma_{m-1})| \vee \bar{L}], \quad m \in \mathbb{N}.$$

Note that the dependence of these stopping times on  $n$  and  $x$  has been suppressed in notation. Also, let  $m_0^n \doteq \min\{m \geq 1 : |\widehat{Q}_x^n(\sigma_m)| \leq \bar{L}\}$ . Define

$$\widehat{G}_n(x) \doteq \mathbb{E} \left[ \int_0^{\tau^n} (1 + |\widehat{Q}_x^n(t)|) dt \right], \quad x \in S.$$

Then

$$\widehat{G}_n(x) \leq \mathbb{E} \left[ \int_0^{\sigma_{m_0^n}} (1 + |\widehat{Q}_x^n(t)|) dt \right] = \sum_{k=0}^{\infty} \mathbb{E} \left[ \int_{\sigma_k}^{\sigma_{k+1}} (1 + |\widehat{Q}_x^n(t)|) dt \mathbf{1}_{k < m_0^n} \right]. \tag{7.13}$$

Let  $\mathcal{F}_t \doteq \sigma\{\widehat{Q}_x^n(s) : 0 \leq s \leq t\}$  (we suppress  $n$  and  $x$  in the notation). We claim that for some constant  $c_1 \in (0, \infty)$ , and for all  $n, k \in \mathbb{N}$ ,  $x \in S$ ,

$$\mathbb{E} \left[ \int_{\sigma_k}^{\sigma_{k+1}} (1 + |\widehat{Q}_x^n(t)|) dt \middle| \mathcal{F}_{\sigma_k} \right] \mathbf{1}_{k < m_0^n} \leq c_1 \left( 1 + |\widehat{Q}_x^n(\sigma_k)|^2 \right) \mathbf{1}_{k < m_0^n}. \tag{7.14}$$

The claim is proved below (7.17). Assuming that the claim holds and using this estimate in (7.13) we get by suitable conditioning

$$\sup_n \widehat{G}_n(x) \leq c_1 \sup_n \mathbb{E} \left[ \sum_{k=0}^{m_0^n-1} \left( 1 + |\widehat{Q}_x^n(\sigma_k)|^2 \right) \right]. \tag{7.15}$$

Next note that  $\{\widehat{Q}_x^n(\sigma_k)\}_{k \geq 1}$  is a Markov chain with the one step transition kernel

$$\check{P}_n(x, A) \doteq P_n^{t_0(|x| \vee \bar{L})}(x, A), \quad x \in S, \quad A \in \mathcal{B}(S),$$

where  $P_n^t$  is the transition probability kernel for the Markov process  $\widehat{Q}^n$ . Using (3.6) and (7.12) one has for some constant  $c_2 \in (1, \infty)$ ,

$$\sup_n \int_S \check{P}_n(x, dy) |y|^2 \leq |x|^2 - \frac{1}{2}|x|^2 + c_2 1_{[0, \bar{L}]}(|x|). \quad (7.16)$$

Using Theorem 14.2.2 of [21] we have that

$$\sup_n \mathbb{E} \sum_{k=0}^{m_0^n - 1} [1 + |\widehat{Q}_x^n(\sigma_k)|^2] \leq 3 \left[ |x|^2 + 2c_2 \sup_n \mathbb{E} \sum_{k=0}^{m_0^n - 1} 1_C(\widehat{Q}_x^n(\sigma_k)) \right] = 3 \left[ |x|^2 + 2c_2 1_{[0, \bar{L}]}(|x|) \right], \quad (7.17)$$

where the equality follows from the fact that whenever  $1 \leq k \leq m_0^n - 1$ ,  $|\widehat{Q}_x^n(\sigma_k)| > \bar{L}$  (we assume without loss of generality  $\bar{L} > 2$ ). The inequality (4.2) now follows on using the above estimate in (7.15).

Thus it only remains to prove the claim in (7.14). By an application of strong Markov property this is equivalent to showing for some  $c_3 \in (0, \infty)$  and all  $n, x$

$$\mathbb{E} \left[ \int_0^{\sigma_1} (1 + |\widehat{Q}_x^n(t)|) dt \right] \leq c_3 (1 + |x|^2). \quad (7.18)$$

From definition of  $\sigma_1$ , we see that

$$\sigma_1 \leq c_4(1 + |x|) \quad (7.19)$$

for some constant  $c_4 \in (0, \infty)$ . With notation introduced in proof of Proposition 3.1, we have

$$\begin{aligned} \sup_n \mathbb{E} \sup_{0 \leq t \leq \sigma_1} |\widehat{Q}_x^n(t)| &= \sup_n \mathbb{E} \sup_{0 \leq t \leq \sigma_1} \left| \Gamma \left( x + \widehat{M}^n(\cdot) + r^n(\cdot) \right) (t) - Z_x^n(t) + Z_x^n(t) \right| \\ &\leq \sup_n L \mathbb{E} \sup_{0 \leq t \leq \sigma_1} |\widehat{M}^n(t)| + \sup_n \mathbb{E} \sup_{0 \leq t \leq \sigma_1} |Z_x^n(t)|. \end{aligned} \quad (7.20)$$

Using the boundedness assumption (iv) in Assumption 2.1 and (7.19) we see that for some  $c_5 \in (0, \infty)$

$$\mathbb{E} \sup_{0 \leq t \leq \sigma_1} |Z_x^n(t)| \leq c_5(1 + |x|), \quad \forall x \in S.$$

Using this estimate along with (3.5) and (7.19) in (7.20) we now have for some  $c_6 \in (0, \infty)$  and  $x \in S$ ,

$$\mathbb{E} \left[ \int_0^{\sigma_1} (1 + |\widehat{Q}_x^n(t)|) dt \right] \leq c_6(1 + |x|^2). \quad (7.21)$$

This proves (7.14) and the result follows.  $\blacksquare$

**Proof of Theorem 6.3.** We will only prove (i) since (ii) follows in an analogous fashion. For simplicity of notation, we will write  $\tilde{Q}$  to denote  $\tilde{Q}^{\omega_0}$  for the proof. For  $x \in \mathbb{R}^K$ , let  $|x|^* \doteq |x| \wedge 1$ . In order to prove (1), it suffices to show that for all  $t \geq 0$

$$\int_{\mathbb{S}_{\text{path}}} |\psi(t)|^* d\tilde{Q}(H^*) = 0, \quad \text{a.s.}, \quad (7.22)$$

where  $\psi(t) \doteq Q^*(t) - \Gamma(Q^*(0) + A^*(\cdot) - RC^*(\cdot) + M^*(\cdot))(t)$ . Note that

$$\int_{\mathbb{S}_{\text{path}}} [j(Q^*) + j(M^*)] d\mathcal{Q}^\ell(H^*) \leq \frac{1}{\sqrt{n_\ell}},$$

where for  $z \in D([0, \infty), \mathbb{R}^K)$ ,  $j(z) \doteq \sup_{t>0} |z(t) - z(t-)|$ . Thus

$$\tilde{\mathcal{Q}}(\mathbb{S}_{\text{path}}^0) = 1, \text{ a.s.}, \quad (7.23)$$

where  $\mathbb{S}_{\text{path}}^0 \doteq C([0, \infty), S \times \mathbb{R}^{4K})$ . In particular,

$$\mathbf{E}_0 \int_{\mathbb{S}_{\text{path}}} |\psi(t)|^* d\tilde{\mathcal{Q}}(H^*) = \lim_{\ell \rightarrow \infty} \mathbf{E} \int_{\mathbb{S}_{\text{path}}} |\psi(t)|^* d\mathcal{Q}^\ell(H^*) = 0. \quad (7.24)$$

Next  $\int_{\mathbb{S}_{\text{path}}} |\psi(t)|^* d\mathcal{Q}^\ell(H^*)$  is equal to

$$\frac{1}{T_\ell} \int_0^{T_\ell} \left| \hat{Q}_p^{n_\ell}(s)(t) - \Gamma \left( \hat{Q}_p^{n_\ell}(s)(0) + \Delta \hat{A}_p^{n_\ell}(s)(\cdot) - R \Delta \hat{C}_p^{n_\ell}(s)(\cdot) + \Delta \hat{M}_p^{n_\ell}(s)(\cdot) \right) (t) \right|^* ds,$$

which from (6.2) is 0 a.s. The equality in (7.22) now follows on combining (7.24) with the above observation. This proves (1).

For  $t \geq 0$  and  $H^* \in \mathbb{S}_{\text{path}}$ , define  $\hat{H}_p^* \in \mathbb{S}_{\text{path}}$  as  $\hat{H}_p^*(t) \doteq (Q_p^*(t), \Delta A_p^*(t), \Delta C_p^*(t), \Delta M_p^*(t), \Delta Y_p^*(t))$ . In order to prove (2), it suffices to show that for all  $f \in C_b(\mathbb{S}_{\text{path}})$ ,  $t \geq 0$ ,

$$\int_{\mathbb{S}_{\text{path}}} [f(\hat{H}_p^*(t)) - f(\hat{H}_p^*(0))] d\tilde{\mathcal{Q}}(H^*) = 0, \text{ a.e. } \omega_0 [\mathbb{P}_0]. \quad (7.25)$$

Note that left side above is the limit (in distribution), as  $\ell \rightarrow \infty$ , of

$$\int_{\mathbb{S}_{\text{path}}} [f(H_p^*(t)) - f(H_p^*(0))] d\mathcal{Q}^\ell(H^*). \quad (7.26)$$

The absolute value of the expression in (7.26) can be written as

$$\left| \frac{1}{T_\ell} \int_0^{T_\ell} [f(\hat{H}_p^{n_\ell}(t+s)) - f(\hat{H}_p^{n_\ell}(s))] ds \right| \leq \frac{2\|f\|_\infty t}{T_\ell} \rightarrow 0, \text{ as } \ell \rightarrow \infty.$$

This proves (7.25) and thus (2) follows.

We now consider (3). Sample path continuity of  $M^*$ , under  $\tilde{\mathcal{Q}}$ , has already been shown in (7.23). From Assumption 2.1 (iii), for any  $t \geq 0$  and  $p \geq 1$ ,

$$\sup_{\ell \geq 1} \sup_{s \geq 0} \mathbf{E} |\widehat{M}^{n_\ell}(t+s) - \widehat{M}^{n_\ell}(s)|^p \doteq \mathcal{N}(p, t) < \infty. \quad (7.27)$$

This in particular shows that for all  $t \geq 0$ ,  $\mathbf{E}_0 \int_{\mathbb{S}_{\text{path}}} |M^*(t)|^2 d\tilde{\mathcal{Q}}(H^*) < \infty$ . In order to prove (3) it suffices now to show that for all  $k \geq 1$ ,  $0 \leq t_1 < t_2 < \dots < t_k \leq s \leq t$ ,  $\psi \in C_b(\mathbb{S}^k)$ ,

$$\mathbf{E}_0 \left| \int_{\mathbb{S}_{\text{path}}} \psi(H^*(t_1), \dots, H^*(t_k)) [M^*(t) - M^*(s)] d\tilde{\mathcal{Q}}(H^*) \right|^2 = 0. \quad (7.28)$$

In view of (7.27) and weak convergence of  $\tilde{Q}^{n\ell}$  to  $\tilde{Q}$ , the expression on the left side of (7.28) is

$$\lim_{\ell \rightarrow \infty} \mathbb{E} \left| \frac{1}{T_\ell} \int_0^{T_\ell} \psi(\hat{H}^{n\ell}(u)(t_1), \dots, \hat{H}^{n\ell}(u)(t_k)) [\hat{M}^{n\ell}(u+t) - \hat{M}^{n\ell}(u+s)] du \right|^2. \quad (7.29)$$

Denote the expression inside the time integral by  $\Lambda(u)$ . Then the above can be rewritten as

$$\lim_{\ell \rightarrow \infty} \mathbb{E} \left[ \frac{2}{T_\ell^2} \int_0^{T_\ell} \int_0^v \Lambda(u) \cdot \Lambda(v) dudv \right]. \quad (7.30)$$

Define the sets

$$L_0 \doteq \{(u, v) \in [0, T_\ell]^2 : 0 \leq u - v < t - s\}, \quad L_1 \doteq \{(u, v) \in [0, T_\ell]^2 : u - v > t - s\}.$$

Using the fact that  $\hat{M}^{n\ell}(t)$  is  $\mathcal{F}_t^n \doteq \sigma\{\hat{H}^{n\ell}(s) : s \leq t\}$  martingale, we see that  $\mathbb{E}[\Lambda(u) \cdot \Lambda(v)] = 0$  for all  $(u, v) \in L_1$ . Thus the expression in (7.30) is the same as

$$\frac{2}{T_\ell^2} \int_{L_0} \mathbb{E}[\Lambda(u) \cdot \Lambda(v)] dudv. \quad (7.31)$$

Using (7.27) the expression in (7.31) can be bounded by

$$\frac{2\|\psi\|_\infty^2 \mathcal{N}(2, t-s) T_\ell(t-s)}{T_\ell^2}, \quad (7.32)$$

which approaches 0 as  $\ell \rightarrow \infty$ . This proves the expression in (7.29) is zero. Thus (7.28) holds and (3) follows.

Proof of (4) is very similar to that of (3) and so we only give a sketch. One needs to establish (7.28) with  $M^*$  replaced by the  $\mathbb{R}^{K \times K}$  valued stochastic process  $N^*$  defined as

$$N^*(t) \doteq M^*(t)[M^*(t)]' - \int_0^t \sigma(Q^*(s))\sigma(Q^*(s))' ds.$$

In order for this it suffices to show that the expression in (7.29), with  $\hat{M}^{n\ell}$  replaced by  $\hat{N}^{n\ell}$ , approaches 0 as  $\ell \rightarrow \infty$  where

$$\begin{aligned} \hat{N}^{n\ell}(t) &\doteq \hat{M}^{n\ell}(t)[\hat{M}^{n\ell}(t)]' - \int_0^t \sigma(\hat{Q}^{n\ell}(s))\sigma(\hat{Q}^{n\ell}(s))' ds \\ &= \hat{M}^{n\ell}(t)[\hat{M}^{n\ell}(t)]' - \int_0^t \Sigma(\hat{Q}^{n\ell}(s))\Sigma(\hat{Q}^{n\ell}(s))' ds. \end{aligned}$$

Let  $\Sigma^n$  be as in (2.12) with  $\lambda, \mu$  there replaced by  $\lambda^n, \mu^n$ . Let  $\tilde{\Sigma}^n(x) \doteq \frac{1}{\sqrt{n}}\Sigma^n(\sqrt{n}x)$ . Then

$$\tilde{N}^n(t) \doteq \hat{N}^n(t) - \int_0^t [(\Sigma\Sigma^*)(\hat{Q}^n(s)) - (\tilde{\Sigma}^n(\tilde{\Sigma}^n)^*)(\hat{Q}^n(s))] ds$$

is an  $\{\hat{\mathcal{F}}^n(t)\}$  martingale. We claim that

$$\mathbb{E} \frac{1}{T_\ell} \int_0^{T_\ell} \left( \int_{u+s}^{u+t} |(\Sigma\Sigma^*)(\hat{Q}^{n\ell}(r)) - (\tilde{\Sigma}^{n\ell}(\tilde{\Sigma}^{n\ell})^*)(\hat{Q}^{n\ell}(r))| dr \right) du \rightarrow 0, \text{ as } \ell \rightarrow \infty. \quad (7.33)$$

Once (7.33) is established, the proof of (4) reduces to showing that the expression in (7.30) is zero, where  $\Lambda(u)$  is defined by replacing  $\widehat{M}^{n_\ell}$  with  $\widehat{N}^{n_\ell}$ . The latter result is shown upon following steps similar to those leading to (7.32) and therefore its proof is omitted. We now prove (7.33). Let  $f_\ell : \mathbb{S}_{\text{path}} \rightarrow \mathbb{R}$  be defined as

$$f_\ell(H^*) \doteq \sup_{s \leq r \leq t} \left| (\Sigma \Sigma^*)(Q^*(r)) - (\widetilde{\Sigma}^{n_\ell}(\widetilde{\Sigma}^{n_\ell})^*)(Q^*(r)) \right|.$$

Then  $f_\ell(H^*) \rightarrow 0$  as  $\ell \rightarrow \infty$  uniformly on compact subsets of  $\mathbb{S}_{\text{path}}$ . Since  $\mathcal{Q}^\ell$  converges to  $\widetilde{\mathcal{Q}}$ , we have

$$\mathbb{E} \int_{\mathbb{S}_{\text{path}}} f_\ell(H^*) d\mathcal{Q}^\ell(H^*) \rightarrow 0. \quad (7.34)$$

Finally, the expression on left side of (7.33) is bounded above by

$$\mathbb{E} \frac{(t-s)}{T_\ell} \int_0^{T_\ell} f_\ell(\widehat{H}^{n_\ell}(u)) du = (t-s) \mathbb{E} \int_{\mathbb{S}_{\text{path}}} f_\ell(H^*) d\mathcal{Q}^\ell(H^*).$$

Combining this with (7.34), we have (7.33). This proves (4).

We now consider (5). It suffices to show that for all  $t \geq 0$ ,

$$\mathbb{E}_0 \int \left| A^*(t) - \int_0^t a(Q^*(s)) ds \right| d\widetilde{\mathcal{Q}}(H^*) = 0. \quad (7.35)$$

Let  $\widetilde{a}^n(x) \doteq \frac{a^n(\sqrt{nx})}{\sqrt{n}}$ ,  $x \in S$ . From Assumption 2.1 (v),  $\widetilde{a}^n \rightarrow a$  uniformly on compact sets. Let  $g_n : \mathbb{S}_{\text{path}} \rightarrow \mathbb{R}$  be defined as  $g_n(H^*) \doteq \sup_{0 \leq s \leq t} |(\widetilde{a}^n - a)(Q^*(s))|$ . Then

$$\mathbb{E} \int_{\mathbb{S}_{\text{path}}} g_{n_\ell}(H^*) d\mathcal{Q}^\ell(H^*) \rightarrow 0, \text{ as } \ell \rightarrow \infty. \quad (7.36)$$

Now the left side of (7.35) can be written as

$$\begin{aligned} & \limsup_{\ell \rightarrow \infty} \mathbb{E} \frac{1}{T_\ell} \int_0^{T_\ell} \left| \widehat{A}^{n_\ell}(t+u) - \widehat{A}^{n_\ell}(u) - \int_u^{t+u} a(\widehat{Q}^{n_\ell}(s)) ds \right| du \\ & \leq t \limsup_{\ell \rightarrow \infty} \mathbb{E} \left| \int_{\mathbb{S}_{\text{path}}} g_{n_\ell}(H^*) d\mathcal{Q}^\ell(H^*) \right|, \end{aligned}$$

where the inequality follows on recalling the representation (6.1). The last expression is 0 in view of (7.36) and thus (7.35) follows. This proves (5).

To prove (6), it suffices to show that for all  $0 < s < t < \infty$ ,

$$\mathbb{E}_0 \int \text{dist} \left( \frac{C^*(t) - C^*(s)}{t-s}, \Lambda \right) d\widetilde{\mathcal{Q}}(H^*) = 0, \quad (7.37)$$

where for  $x \in \mathbb{R}^K$ ,  $\text{dist}(x, \Lambda) \doteq \inf_{y \in \Lambda} |x - y|$ . However, (7.37) is immediate on using weak convergence of  $\mathcal{Q}^\ell$  to  $\widetilde{\mathcal{Q}}$  and recalling that

$$\frac{\widehat{C}^{n_\ell}(t+u) - \widehat{C}^{n_\ell}(s+u)}{t-s} \in \Lambda$$

a.s. for all  $\ell \geq 1$  and  $u \geq 0$ . ■

### 7.1. An estimate on transition probability densities.

Following is the main result of this section which follows from results in [3]. For the sake of readability, detailed arguments are provided. Writing  $\mathcal{C} = C([0, 1], \mathbb{R}^K)$ , let  $\mathbb{P}_x$  be the probability measure on  $(\mathcal{C}, \mathcal{B}(\mathcal{C}))$  under which the canonical coordinate process  $\{W(t), 0 \leq t \leq 1\}$  is a Brownian motion starting at  $x$  with covariance matrix  $\sigma\sigma'$ . Let  $Z(t) \doteq \Gamma(W)(t)$ ,  $Y(t) \doteq \Gamma_1(W)(t)$ ,  $0 \leq t \leq 1$ . For a closed set  $F \subset S$  and  $\delta > 0$ , let  $F^\delta \doteq F \cap \{x \in S : \text{dist}(x, \partial S) \geq \delta\}$ .

**Theorem 7.1.** *There is a measurable function  $p : [0, 1] \times S \times S \rightarrow [0, \infty)$  such that:*

(i)

$$\mathbb{P}_x(Z(t) \in A) = \int_A p(t, x, z) dz, \quad 0 \leq t \leq 1, \quad x \in S, \quad A \in \mathcal{B}(S).$$

(ii) *For each  $\delta > 0$ , and a compact set  $F \subset S$ , there is a measurable function  $\Psi_\delta : [0, 1] \rightarrow \mathbb{R}_+$  and  $\alpha > 0$  such that*

$$\sup_{x \in F, z \in F^\delta} p(t, x, z) \leq \Psi_\delta(t), \quad t \in (0, 1), \quad (7.38)$$

$$\int_0^1 e^{-\alpha/t} \Psi_\delta(t) dt < \infty.$$

**Proof.** Let  $\Gamma(t, x, z)$  be the transition density function of a Brownian motion starting at  $x$  and covariance  $a = \sigma\sigma'$ , i.e.,

$$\Gamma(t, x, z) = |2\pi at|^{-1/2} \exp \left[ -\frac{(x-z)' a^{-1} (x-z)}{2t} \right],$$

where  $|A|$  denotes the determinant of a  $K \times K$  matrix  $A$ . Let  $K(t) \doteq RY(t)$ . For a bounded open (relative to  $S$ ) set  $G \subseteq S$ ,  $t > 0$ , define

$$p_G(t, x, z) \doteq \Gamma(t, x, z) - \mathbb{E}_x \left[ 1_{[0, t]}(\tau_G) \Gamma(t - \tau_G, Z(\tau_G), z) \right] \\ + \mathbb{E}_x \left[ \int_0^{t \wedge \tau_G} \langle \nabla \Gamma(t - r, Z(r), z), dK(r) \rangle \right] \quad (7.39)$$

where  $x \in \bar{G}$ ,  $z \in G \cap S^\circ$ ,  $\tau_G = \inf\{r \in (0, \infty) : Z(r) \notin G\}$  and  $\nabla \Gamma(t, \zeta, z)$  denotes the gradient of  $\Gamma$  in the variable  $\zeta$ . Henceforth we will abbreviate  $\tau_G$  as  $\tau$ .

Note that there exist  $c_1, c_2 \in (0, \infty)$  such that for all  $x, z \in \mathbb{R}^K$ ,  $0 < t \leq 1$ ,

$$|D_x^\alpha \Gamma(t, x, z)| \leq c_1 t^{-\left(\frac{K+|\alpha|}{2}\right)} \exp \left[ \frac{-c_2 |z-x|^2}{t} \right], \quad (7.40)$$

where  $\alpha = (\alpha_1, \dots, \alpha_K)$  is a multi-index with  $0 \leq |\alpha| \leq 2$  where  $|\alpha| = \alpha_1 + \dots + \alpha_K$  and  $D_x^\alpha$  denotes differentiation with respect to  $x$ . Also

$$\sup_{0 < s \leq 1} \frac{e^{-c/s}}{s^m} < \infty \quad \text{for all } m \in \mathbb{R}, \quad c \in (0, \infty). \quad (7.41)$$

Furthermore, since  $z \in G \cap S^\circ$ ,  $\text{dist}(z, \partial G) > 0$  and  $\text{dist}(z, \partial S) > 0$ . Combining these observations with the fact that  $Y(t)$  increases only when  $Z(t) \in \partial S$ , we see that the expectations on the right side of (7.39) are well defined and finite. Thus  $p_G(t, x, z)$  is well-defined.

We claim that for any continuous function  $f$  with compact support  $A$  in  $G \cap S^\circ$ ,  $x \in \bar{G}$  and  $t \in (0, 1]$

$$\int_S f(z) p_G(t, x, z) dz = \mathbb{E}_x \left[ 1_{[t, \infty)}(\tau) f(Z(t)) \right]. \quad (7.42)$$

Before proving the claim we show how it implies the desired result. We have from (7.42) that for all  $t > 0$ ,  $x \in \bar{G}$  and  $z \in G \cap S^\circ$ ,  $p_G(t, x, z) \geq 0$ . Furthermore, for any Borel set  $E \subseteq G \cap S^\circ$ ,  $t > 0$ , and  $x \in \bar{G}$ ,

$$P_x(Z(t) \in E, \tau \geq t) = \int_G 1_E(z) p_G(t, x, z) dz. \quad (7.43)$$

From Feller property of  $Z$  it follows that for all  $t > 0$ ,  $x \mapsto p_G(t, x, z)$  is continuous on  $\bar{G}$ . Combining this with the bound in (7.41) and observations below it, we see that for each  $t > 0$ ,  $(x, z) \mapsto p_G(t, x, z)$  is continuous on  $\bar{G} \times (G \cap S^\circ)$ . Defining  $p_G(t, \cdot, \cdot)$  to be zero outside  $\bar{G} \times (G \cap S^\circ)$ , we have that  $p_G(t, \cdot, \cdot)$  is jointly measurable on  $S \times S$  for all  $t > 0$ .

For  $n \in \mathbb{N}$ , let  $B(0; n)$  be the open ball centered at 0 with radius  $n$ . Let  $G_n \doteq B(0; n) \cap S$  and define  $p(t, x, z) \doteq \lim_{n \rightarrow \infty} p_{G_n}(t, x, z)$ . Note that for fixed  $t, x, z$ ,  $p_{G_n}(t, x, z)$  is nondecreasing in  $n$ , so the above limit is well defined. By definition  $(x, z) \mapsto p(t, x, z)$  is a jointly measurable nonnegative function on  $S \times S$ . Furthermore, observe that for  $C \in \mathcal{B}(S)$

$$\begin{aligned} P_x[Z(t) \in C] &= P_x[Z(t) \in C \cap S^\circ] \\ &= \lim_{n \rightarrow \infty} P_x[Z(t) \in C \cap S^\circ \cap G_n, \tau_{G_n} \geq t] \\ &= \lim_{n \rightarrow \infty} \int_S 1_C(z) p_n(t, x, z) dz = \int_S 1_C(z) p(t, x, z) dz. \end{aligned} \quad (7.44)$$

This proves part (i) of the theorem.

We now consider part (ii) of the theorem. Fix  $n_0$  large enough so that  $B(0; n_0) \supset F$ . It suffices to prove (7.38) with  $p$  replaced with  $p_{G_n}$ ,  $n \geq n_0$  and with a function  $\Psi_\delta$  that is independent of the choice of  $n$ . Note that for  $z \in F^\delta$

$$\begin{aligned} p_{G_n}(t, x, z) &\leq |2\pi at|^{-1/2} + \mathbb{E}_x \int_0^t c_1(t-r)^{-(K+1)/2} \exp\{-c_2\delta^2/(t-r)\} d|K|(r) \\ &\leq |2\pi at|^{-1/2} + c_3(\delta) \mathbb{E}_x |K|(t), \end{aligned}$$

where  $|K|$  denotes the scalar total variation process for  $K$ . In the above display, the first inequality follows from (7.40) while the second inequality follows from (7.41). Standard estimates show that  $\sup_{0 \leq t \leq 1} \sup_{x \in F} \mathbb{E}_x |K|(t) = c_4(F) < \infty$ . Now (7.38) with  $p$  replaced with  $p_{G_n}$  follows on taking  $\Psi_\delta(t) = |2\pi at|^{-1/2} + c_3(\delta) c_4(F)$ .

Finally we establish the claim in (7.42). Denote the left hand side of (7.42) by  $u(t, x)$ . We begin by showing that

$$\left\{ u(t - (\tau \wedge r), Z(\tau \wedge r)) : 0 \leq r < t \right\} \text{ is a } P_x\text{-martingale w.r.t. } \{\mathcal{F}_{\tau \wedge r} : 0 \leq r < t\}, \quad (7.45)$$

where  $\{\mathcal{F}_t\}$  is the canonical filtration on  $(\mathcal{C}, \mathcal{B}(\mathcal{C}))$ . For  $0 \leq r_1 \leq r_2 < t$ , let  $\mathcal{F}_1 \doteq \mathcal{F}_{\tau \wedge r_1}$ ,  $\mathcal{F}_2 \doteq \mathcal{F}_{\tau \wedge r_2}$  and  $\xi_1 \doteq Z(\tau \wedge r_1)$ ,  $\xi_2 \doteq Z(\tau \wedge r_2)$ . We abbreviate  $\tau \wedge r_i$  by  $\tau_i$ ,  $i = 1, 2$ . We first show that

$$\begin{aligned} & \mathbb{E}_x \left[ \int_{S^\circ} f(z) \Gamma(t - \tau_2, \xi_2, z) dz \middle| \mathcal{F}_1 \right] \\ &= \int_{S^\circ} f(z) \Gamma(t - \tau_1, \xi_1, z) dz + \int_{S^\circ} f(z) \mathbb{E} \left[ \int_{\tau_1}^{\tau_2} \langle \nabla \Gamma(t - r, Z(r), z), dK(r) \rangle \middle| \mathcal{F}_1 \right] dz. \end{aligned} \quad (7.46)$$

By applying Itô's formula to  $\Gamma(t - (\tau \wedge r), Z(\tau \wedge r), z)$ , for  $x \in S^\circ$ , over  $[r_1, r_2]$ , we have

$$\begin{aligned} & \Gamma(t - \tau_2, \xi_2, z) - \Gamma(t - \tau_1, \xi_1, z) \\ &= \int_{\tau_1}^{\tau_2} \left( -\frac{\partial}{\partial t} \Gamma(t - r, Z(r), z) + \mathcal{L} \Gamma(t - r, Z(r), z) \right) dr \\ &+ \int_{\tau_1}^{\tau_2} \langle \nabla \Gamma(t - r, Z(r), z), dK(r) \rangle + \int_{\tau_1}^{\tau_2} \langle \nabla \Gamma(t - r, Z(r), z), \sigma dW(r) \rangle, \end{aligned}$$

where  $\mathcal{L} \Gamma(t, x, z) = \left( \frac{1}{2} \sum_{i,j=1}^K a_{ij} \frac{\partial^2}{\partial x_i \partial x_j} \right) \Gamma(t, x, z)$ . Conditioning on  $\mathcal{F}_1$  and taking expectation on both sides, we have on using Kolmogorov's backward equation,  $\frac{\partial}{\partial t} \Gamma(t, x, z) = \mathcal{L} \Gamma(t, x, z)$ , that

$$\mathbb{E} [\Gamma(t - \tau_2, \xi_2, z) | \mathcal{F}_1] = \Gamma(t - \tau_1, \xi_1, z) + \mathbb{E} \left[ \int_{\tau_1}^{\tau_2} \langle \nabla \Gamma(t - r, Z(r), z), dK(r) \rangle \middle| \mathcal{F}_1 \right].$$

Equality (7.46) now follows from the above equality via an application of Fubini's Theorem.

Next we show that

$$\{U_{\tau \wedge r}, \mathcal{F}_{\tau \wedge r} : 0 \leq r < t\} \text{ is a martingale,} \quad (7.47)$$

where  $U_r \doteq \int p_1(t - r, Z(r), z) f(z) dz$ ,  $0 \leq r < t$  and

$$p_1(t - r, x, z) \doteq \mathbb{E}_x \left[ 1_{[0, t-r)}(\tau) \Gamma(t - r - \tau, Z(\tau), z) \right].$$

From strong Markov property of  $Z$  we have that

$$\mathbb{E}_x \left[ 1_{[0, t)}(\tau) \Gamma(t - \tau, Z(\tau), z) \middle| \mathcal{F}_{\tau \wedge r} \right] = p_1(t - \tau \wedge r, Z(\tau \wedge r), z). \quad (7.48)$$

Thus from (7.48)

$$U_{\tau \wedge r} = \mathbb{E} \left[ \int_{S^\circ} f(z) 1_{[0, t)}(\tau) \Gamma(t - \tau, Z(\tau), z) dz \middle| \mathcal{F}_{\tau \wedge r} \right]$$

which proves (7.47).

Next let  $p_2(t, x, z) \doteq \mathbb{E}_x \left[ \int_0^{t \wedge \tau} \langle \nabla \Gamma(t - r, Z(r), z), dK(r) \rangle \right]$ . Using strong Markov property again, for  $r \leq t$ ,

$$p_2(t - \tau \wedge r, Z(\tau \wedge r), z) = \mathbb{E}_x \left[ \int_{\tau \wedge r}^{\tau \wedge t} \langle \nabla \Gamma(t - u, Z(u), z), dK(u) \rangle \middle| \mathcal{F}_{\tau \wedge r} \right].$$

Thus

$$\mathbb{E} [p_3(t - \tau_2, \xi_2, z) | \mathcal{F}_1] - p_3(t - \tau_1, \xi_1, z) = -\mathbb{E} \left[ \int_{\tau_1}^{\tau_2} \langle \nabla \Gamma(t - u, Z(u), z), dK(u) \rangle \middle| \mathcal{F}_1 \right].$$

Integrating over  $S^\circ$ , we have

$$\begin{aligned} & \mathbb{E}_x \left[ \int_{S^\circ} f(z) p_3(t - \tau_2, \xi_2, z) dz \middle| \mathcal{F}_1 \right] - \mathbb{E}_x \left[ \int_{S^\circ} f(z) p_3(t - \tau_1, \xi_1, z) dz \right] \\ &= - \int_{S^\circ} \mathbb{E}_x \left[ \int_{\tau_1}^{\tau_2} \langle \nabla \Gamma(t - u, Z(u), z), dK(u) \rangle \middle| \mathcal{F}_1 \right] f(z) dz. \end{aligned} \quad (7.49)$$

Proof of (7.45) now follows on combining (7.39), (7.46), (7.47), and (7.49).

Next observe that (7.42) holds trivially for  $x \in \partial G$ . Finally, for any  $x \in G$  and continuous  $f$  with compact support  $A \subseteq G \cap S^\circ$

$$\begin{aligned} \int_{S^\circ} f(z) p_G(t, x, z) dz &= \lim_{r \uparrow t} \mathbb{E}_x [u(t - \tau \wedge r, Z(\tau \wedge r))] \\ &= \lim_{r \uparrow t} \mathbb{E}_x \left[ \int_A f(z) p_G(t - \tau \wedge r, Z(\tau \wedge r), z) dz \right] \\ &= \lim_{r \uparrow t} \mathbb{E}_x \left[ \int_A f(z) 1_{[t, \infty)}(\tau) p_G(t - \tau \wedge r, Z(\tau \wedge r), z) dz \right] \\ &\quad + \lim_{r \uparrow t} \mathbb{E}_x \left[ \int_A f(z) 1_{[0, t)}(\tau) p_G(t - \tau \wedge r, Z(\tau \wedge r), z) dz \right] \\ &\equiv I_1 + I_2, \end{aligned} \quad (7.50)$$

where the first equality follows on using the martingale property (7.45). Next note that from (7.40) and (7.41) (and observations below these equations), for all compact set  $A \subseteq G \cap S^\circ$ ,

$$\sup_{z \in A} \sup_{\tilde{z} \in \partial G} \sup_{0 \leq s \leq t} \Gamma(s, \tilde{z}, z) < \infty, \quad (7.51)$$

and

$$\sup_{z \in A} \sup_{\tilde{z} \in \partial S} \sup_{0 \leq s \leq t} |\nabla \Gamma(s, \tilde{z}, z)| < \infty. \quad (7.52)$$

Combining the uniform estimates (7.51), (7.52) with the Feller property of  $Z$ , we have by an application of dominated convergence theorem: If  $x \in G, y(s) \in G$  are such that  $y(s) \rightarrow x$  as  $s \rightarrow t$ , then

$$\lim_{s \uparrow t} \int_{S^\circ \cap G} f(z) p_G(t - s, y(s), z) dz = f(x). \quad (7.53)$$

Next noting that  $\sup_{0 \leq s \leq t} \sup_{x \in \bar{G}, z \in A} |p_G(s, x, z) - \Gamma(s, x, z)| < \infty$ , we have via another application of dominated convergence theorem that

$$\begin{aligned} I_1 &= \mathbb{E}_x \left[ 1_{[t, \infty)}(\tau) \lim_{r \uparrow t} \int_{S^\circ} f(z) p_G(t - \tau \wedge r, Z(\tau \wedge r), z) dz \right] \\ &= \mathbb{E}_x \left[ 1_{[t, \infty)}(\tau) f(Z(t)) \right], \end{aligned} \quad (7.54)$$

$$\begin{aligned}
I_2 &= \mathbb{E}_x \left[ \mathbf{1}_{[0,t)}(\tau) \lim_{r \uparrow t} \int_{S^\circ} f(z) p_G(t - \tau \wedge r, Z(\tau \wedge r), z) dz \right] \\
&= \mathbb{E}_x \left[ \mathbf{1}_{[0,t)}(\tau) \int_{S^\circ} f(z) p_G(t - \tau, Z(\tau), z) dz \right] = 0,
\end{aligned} \tag{7.55}$$

where the second equality in (7.54) follows from (7.53) and the last inequality is a consequence of the observation that  $p_G(t, x, z) = 0$  for all  $t > 0$ ,  $x \in \partial G$ ,  $z \in A$ . The result now follows on combining (7.50), (7.54), and (7.55). ■

## References

- [1] B. Ata, J. M. Harrison, and L. A. Shepp. Drift rate control of a Brownian processing system. *Ann. Appl. Probab.*, 15(2):1145–1160, 2005.
- [2] R. Atar, A. Budhiraja, and P. Dupuis. On positive recurrence of constrained diffusion processes. *Ann. Probab.*, 29(2):979–1000, 2001.
- [3] S. Balaji and S. Ramasubramanian. Asymptotics of reflecting diffusions in an orthant. *Proc. Intl. Conf. Held at Cochin, India.*, 1997.
- [4] P. Billingsley. *Convergence of probability measures, Second Edition*. Wiley Series in Probability and Statistics. John Wiley and Sons, Inc., New York, 1999.
- [5] V. Borkar. *Optimal control of diffusion processes*. Pitman Research Notes in Mathematics Series, 203. Longman Scientific & Technical, 1989.
- [6] P. Brémaud. *Point processes and queues*. Springer-Verlag, New York, 1981. Martingale dynamics, Springer Series in Statistics.
- [7] A. Budhiraja. An ergodic control problem for constrained diffusion processes: Existence of optimal Markov control. *SIAM J. Control Optim.*, 42(2):532–558, 2003.
- [8] A. Budhiraja and C. Lee. Long time asymptotics for constrained diffusions in polyhedral domains. *Stochastic Process. Appl.*, 117(8): 1014–1036, 2007.
- [9] A. Budhiraja and C. Lee. Stationary distribution convergence for generalized Jackson networks in heavy traffic. *To appear in Math. Oper. Res.*, 2009.
- [10] J. G. Dai and S. P. Meyn. Stability and convergence of moments for multiclass queueing networks via fluid limit models. *IEEE Transactions on Automatic Control*, 40:1889–1904, 1995.
- [11] P. Dupuis and H. Ishii. On Lipschitz continuity of the solution mapping to the Skorohod problem, with applications. *Stochastics*, 35:31–62, 1991.
- [12] D. Gamarnik and A. Zeevi. Validity of heavy traffic steady-state approximations in generalized Jackson networks. *Ann. Appl. Probab.* 16:56–90, 2006.
- [13] A. P. Ghosh and A. Weerasinghe. Optimal buffer size and dynamic rate control for a queueing network with impatient customers in heavy traffic. *Submitted*, 2008.  
[http://www.public.iastate.edu/~apghosh/reneing\\_queue.pdf](http://www.public.iastate.edu/~apghosh/reneing_queue.pdf)
- [14] J. M. Harrison and M. I. Reiman. Reflected Brownian motion on an orthant. *Ann. Probab.*, 9(2):302–308, 1981.
- [15] J. M. Harrison and R. J. Williams. Brownian models of open queueing networks with homogeneous customer populations. *Stochastics*, 22:77–115, 1987.
- [16] N. Ikeda and S. Watanabe. *Stochastic differential equations and diffusion processes*, vol-

- ume 24 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam, second edition, 1989.
- [17] T. G. Kurtz and R. H. Stockbridge. *Martingale problems and linear programs for singular control*. 37th annual Allerton Conference on Communication Control and Computing (Monticello, Ill. 1999), 11-20, Univ. Illinois, Urbana-Champaign, Ill.
  - [18] H. J. Kushner. *Heavy traffic analysis of controlled queueing and communication networks*, volume 47 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, 2001. Stochastic Modelling and Applied Probability.
  - [19] H. J. Kushner and L. F. Martins. Limit theorems for pathwise average cost per unit time problems for controlled queues in heavy traffic. *Stochastics Stochastics Rep.*, 42(1):25–51, 1993.
  - [20] A. Mandelbaum and G. Pats. State-Dependent Stochastic Networks. Part I: Approximations and Applications with Continuous Diffusion Limits. *Ann. Appl. Probab.* 8 (2): 569–646, 1998.
  - [21] S. P. Meyn and R. L. Tweedie. *Markov chains and stochastic stability*. Springer-Verlag, London, 1993.
  - [22] W. Rudin. *Real and complex analysis*. McGraw-Hill Book Co., New York, third edition, 1987.
  - [23] D. W. Stroock and S. R. S. Varadhan. *Multidimensional diffusion processes*. Classics in Mathematics. Springer-Verlag, Berlin, 2006. Reprint of the 1997 edition.
  - [24] K. Yamada. Diffusion approximation for open state-dependent queueing networks in the heavy traffic situation. *Ann. Appl. Probab.*, 5(4):958–982, 1995.

A. BUDHIRAJA  
DEPARTMENT OF STATISTICS AND OPERATIONS RESEARCH  
UNIVERSITY OF NORTH CAROLINA  
CHAPEL HILL, NC 27599, USA

A. P. GHOSH  
DEPARTMENT OF STATISTICS  
IOWA STATE UNIVERSITY  
AMES, IA 50011-1210, USA

C. LEE  
DEPARTMENT OF STATISTICS  
COLORADO STATE UNIVERSITY  
FORT COLLINS, CO 80523-1877, USA