

DIFFUSION MODELS AND STEADY-STATE APPROXIMATIONS FOR EXPONENTIALLY ERGODIC MARKOVIAN QUEUES

BY ITAI GURVICH

Northwestern University

Motivated by queues with many servers, we study Brownian steady-state approximations for continuous time Markov chains (CTMCs). Our approximations are based on *diffusion models* (rather than a diffusion limit) whose steady-state, we prove, approximates that of the Markov chain with notable precision. Strong approximations provide such “limitless” approximations for process dynamics. Our focus here is on steady-state distributions, and the diffusion model that we propose is tractable relative to strong approximations.

Within an asymptotic framework, in which a scale parameter n is taken large, a uniform (in the scale parameter) Lyapunov condition imposed on the sequence of diffusion models guarantees that the gap between the steady-state moments of the diffusion and those of the properly centered and scaled CTMCs shrinks at a rate of \sqrt{n} .

Our proofs build on gradient estimates for solutions of the Poisson equations associated with the (sequence of) diffusion models and on elementary martingale arguments. As a by-product of our analysis, we explore connections between Lyapunov functions for the fluid model, the diffusion model and the CTMC.

1. Introduction. Fluid and diffusion limits for queuing systems have been applied successfully toward performance analysis and optimization of various queuing systems. We are concerned here with performance analysis in steady-state and, more specifically, with Brownian steady-state approximations for continuous time Markov chains (CTMCs).

The framework of diffusion limits begins with a sequence of CTMCs $\{X^n\}$, and properly scaled and centered versions

$$\widehat{X}^n = \frac{X^n - \bar{x}^n}{\sqrt{n}}$$

for some sequence $\{\bar{x}^n\}$ that arises from the specific structure of the model. With appropriate assumptions on the parameters of the CTMC, and on the sequence of initial conditions $\{\widehat{X}^n(0)\}$, one typically proceeds to establish process convergence

$$(1) \quad \widehat{X}^n \Rightarrow \widehat{X} \quad \text{as } n \rightarrow \infty,$$

Received April 2013; revised October 2013.

MSC2010 subject classifications. 60K25, 90B20, 90B36, 49L20, 60F17.

Key words and phrases. Markovian queues, steady-state, many servers, heavy-traffic, Halfin–Whitt regime, steady state approximations, strong approximations for queues.

in the appropriate function space where \widehat{X} is a diffusion process. If each of the $\{X^n\}$ as well as \widehat{X} are ergodic, and f is a continuous function such that $\{f(\widehat{X}^n(\infty))\}$ is uniformly integrable, one can subsequently conclude that

$$\mathbb{E}[f(\widehat{X}^n(\infty))] \rightarrow \mathbb{E}[f(\widehat{X}(\infty))] \quad \text{as } n \rightarrow \infty,$$

where $\widehat{X}^n(\infty)$ and $\widehat{X}(\infty)$ have, respectively, the steady-state distributions of \widehat{X}^n and \widehat{X} . A relatively general framework toward proving the required uniform integrability has been developed in [12] and applied there to generalized Jackson networks; see also [6]. It was subsequently applied successfully to other queueing systems. This so-called *interchange of limits* establishes that

$$(2) \quad \mathbb{E}[f(\widehat{X}^n(\infty))] = \mathbb{E}[f(\widehat{X}(\infty))] + o(1),$$

and supports using $\mathbb{E}[f(\widehat{X}(\infty))]$ as an approximation for $\mathbb{E}[f(\widehat{X}^n(\infty))]$.

A central benefit of the limit approach to approximations is the relative tractability of the diffusion \widehat{X} relative to the original CTMC. The convergence rate embedded in the $o(1)$ term is not, however, precisely captured by these convergence arguments. In this paper, we prove that an appropriately defined sequence of diffusion models, that are as tractable as the diffusion limit, provides accurate approximations for the steady-state of the CTMCs with an approximation gap that shrinks at a rate of \sqrt{n} . Our approach does not require process convergence as in (1).

We proceed to an informal exposition of the results and key ideas. The Markov chains that we consider have a semi-martingale representation

$$X^n(t) = X^n(0) + \int_0^t F^n(X^n(s)) ds + M^n(t),$$

where M^n is a local martingale with respect to a properly defined filtration. We define a fluid model by (heuristically) removing the martingale term, that is,

$$(FM) \quad \bar{x}^n(t) = \bar{x}^n(0) + \int_0^t F^n(\bar{x}^n(s)) ds.$$

If the FM has a unique stationary point \bar{x}_∞^n satisfying $F^n(\bar{x}_\infty^n) = 0$, it subsequently makes sense to center X^n around \bar{x}_∞^n and consider the centered and scaled process $\widehat{X}^n = (X^n - \bar{x}_\infty^n)/\sqrt{n}$. The process \widehat{X}^n satisfies the equation

$$\widehat{X}^n(t) = \widehat{X}^n(0) + \int_0^t \widehat{F}^n(\widehat{X}^n(s)) ds + M^n(t)/\sqrt{n},$$

where $\widehat{F}^n(y) = F^n(\sqrt{n}y + \bar{x}_\infty^n)/\sqrt{n}$, $y \in \mathbb{R}^d$. Under appropriate conditions, a strong approximation for \widehat{X}^n is given by the diffusion process

$$\widehat{S}^n(t) = \widehat{S}^n(0) + \int_0^t \widehat{F}^n(\widehat{S}^n(s)) ds + \int_0^t \sigma^n(\widehat{S}^n(s)) dB(s),$$

where B is a standard Brownian motion and σ^n arises naturally from the Markov-chain transition functions and is intimately related to the predictable quadratic variation of the martingale M^n . Strong-approximations theory predicts an approximation gap that is logarithmic in nT where T is the time horizon; see Remark 3.1.

A cruder approximation is obtained by replacing the (state dependent) diffusion coefficient with its value at the stationary point of the FM, \bar{x}_∞^n , to obtain the diffusion process specified by the equation

$$(DM) \quad \widehat{Y}^n(t) = \widehat{Y}^n(0) + \int_0^t \widehat{F}^n(\widehat{Y}^n(s)) ds + \sigma^n(\bar{x}_\infty^n)B(t).$$

Our main finding is that this straightforward heuristic derivation of the DM—building on a stationary point of the fluid model to construct a simplified diffusion model—may provide, insofar as steady-state analysis is concerned, an impressively accurate approximation.

More precisely, but still proceeding informally at this stage, we prove the following. Let \mathcal{A}^n be the generator of the diffusion \widehat{Y}^n . If there exists a function V together with finite positive constants b, δ and a compact set B (all not depending on n) such that

$$(UL) \quad \mathcal{A}^n V(x) \leq -\delta V(x) + b\mathbb{1}_B(x), \quad x \in \mathbb{R}^d,$$

then

$$\mathbb{E}[f(\widehat{Y}^n(\infty))] - \mathbb{E}[f(\widehat{X}^n(\infty))] = \mathcal{O}(1/\sqrt{n})$$

for all functions f with $|f| \leq V$. The uniform Lyapunov requirement UL must be proved on a case-by-case basis, and we illustrate this via two examples in Section 6. The requirement UL restricts the scope of our results to (sequences of) chains in which the corresponding DM is exponentially ergodic.

The sequence of Poisson equations (associated with the sequence of DMs) is central to our proofs. Let π^n be the steady-state distribution of the diffusion model and ν^n be that of the scaled CTMC. Let f be such that $\pi^n(f) = 0$. [The requirement that $\pi^n(f) = 0$ is not necessary and is imposed in this discussion for expositional purposes.] We will show that a solution $u_f^n \in \mathcal{C}^2(\mathbb{R}^d)$ exists for the DM’s Poisson equation

$$\mathcal{A}^n u = -f.$$

Based on Itô’s rule one expects that

$$\mathbb{E}_{\pi^n}[u_f^n(\widehat{Y}^n(t))] = \mathbb{E}_{\pi^n}[u_f^n(\widehat{Y}^n(0))] + \mathbb{E}_{\pi^n}\left[\int_0^t \mathcal{A}^n u_f^n(\widehat{Y}^n(s)) ds\right].$$

Since the DM has, by construction, a diffusion coefficient that does not depend on the state, the Poisson equation is (for each n) a linear PDE, and we are able to build on existing theory to identify gradient estimates that are uniform in the index n . These gradient estimates facilitate proving that

$$\mathbb{E}_{\nu^n}[u_f^n(\widehat{X}^n(t))] = \mathbb{E}_{\nu^n}[u_f^n(\widehat{X}^n(0))] + \mathbb{E}_{\nu^n}\left[\int_0^t \mathcal{A}^n u_f^n(\widehat{X}^n(s)) ds\right] + t\mathcal{O}(1/\sqrt{n}).$$

Informally speaking, this shows that u_f^n “almost solves” the Poisson equation for the CTMC.

Stationarity then allows us to conclude that

$$\mathbb{E}_{\nu^n} \left[\int_0^t \mathcal{A}^n u_f^n(\widehat{X}^n(s)) ds \right] = -t \mathbb{E}_{\nu^n} \left[\int_0^t f(\widehat{X}^n(s)) ds \right] = t \mathcal{O}(1/\sqrt{n}),$$

and, in particular, that

$$v^n(f) = \mathcal{O}(1/\sqrt{n}).$$

Recalling that $\pi^n(f) = 0$, it then follows that

$$v^n(f) - \pi^n(f) = \mathcal{O}(1/\sqrt{n}).$$

In the process of proving these results, we explore connections between the stability of the CTMC and that of the corresponding FM and DM.

Refined properties of the Poisson equation in the context of diffusion approximations for diffusions with a fast component are used in [21]. In the spirit of this paper, derivative bounds for certain Dirichlet problems are used in [15] to study *universal* approximations for the birth-and-death process underlying the so-called Erlang-A queue. The proofs there are based on the study of excursions but are closely related to ours; we revisit the Erlang-A queue in Section 6. The use of gradient estimates in conjunction with martingale arguments is also the theme in [1] where these are used to study optimality gaps in the control of a multi-class queue. The Poisson equation is replaced there with the PDE associated with the HJB equation.

Notation. Unless stated otherwise, all convergence statements are for $n \rightarrow \infty$. We use $|x|$ to denote the Euclidean norm of x in \mathbb{R}^d (the dimension d will be clear from the context). For two nonnegative sequences $\{a^n\}$ and $\{b^n\}$ we write $a^n = \mathcal{O}(b^n)$ if $\limsup_{n \rightarrow \infty} a^n/b^n < \infty$. Throughout we adopt the convention that $0/0 = 0$. We let

$$B_x(M) = \{y \in \mathbb{R}^d : |x - y| < M\},$$

and denote its closure by $\overline{B}_x(M)$. Following standard notation, we let $\mathcal{C}^j(\mathbb{R}^d)$ be the space of j -times continuously differentiable functions from \mathbb{R}^d to \mathbb{R} , and for $u \in \mathcal{C}^2(\mathbb{R}^d)$ we let Du and D^2u denote the gradient and the Hessian of u , respectively.

Given a Markov process $\Xi = (\Xi(t), t \geq 0)$ on a complete and separable metric space \mathcal{X} , we let \mathbb{P}_x be the probability distribution under which $\mathbb{P}\{\Xi(0) = x\} = 1$ for $x \in \mathcal{X}$ and $\mathbb{E}_x[\cdot] = \mathbb{E}[\cdot | \Xi(0) = x]$ be the expectation operator w.r.t. the probability distribution \mathbb{P}_x . Let \mathbb{P}_π denote the probability distribution under which $\Xi(0)$ is distributed according to π and put $\mathbb{E}_\pi[\cdot]$ to be the expectation operator w.r.t. this distribution. A probability distribution π defined on \mathcal{X} is said to be a stationary distribution if for every bounded continuous function f

$$\mathbb{E}_\pi[f(\Xi(t))] = \mathbb{E}_\pi[f(\Xi(0))] \quad \text{for all } t \geq 0.$$

It is said to be the steady-state distribution if for every such function and all $x \in \mathcal{X}$,

$$\mathbb{E}_x[f(\Xi(t))] \rightarrow \mathbb{E}_\pi[f(\Xi(0))] \quad \text{as } t \rightarrow \infty.$$

Given a probability distribution ν and a nonnegative function f , we define $\nu(f) = \int f(x) d\nu(x)$ (which can be infinite). For a general (not necessarily nonnegative) function, we define $\nu(f)$ as above whenever $\nu(|f|) < \infty$. Finally, whereas our results are not concerned with process-convergence, we will be making connections to the functional central limit theorem. All the processes that we study are assumed to be right continuous with left limits (RCLL), and \Rightarrow will be used for convergence in the space $\mathcal{D}^d[0, \infty)$ of such functions unless otherwise stated. For RCLL processes we use $x(t-) = \lim_{s \uparrow t} x(s)$ and let $\Delta x(t) = x(t) - x(t-)$.

2. A sequence of CTMCs. We consider a sequence $\{X^n, n \in \mathbb{N}\}$ of continuous time Markov chains (CTMCs). The chain X^n moves on a countable state space $E^n \subset \mathbb{R}^d$ according to transition rates $\beta_{y-x}^n(x) = q_{x,y}^n$ for $x, y \in E^n$. Given a nonrandom initial condition $X^n(0) \in E^n$, the dynamics of X^n are constructed as follows:

$$X^n(t) = X^n(0) + \sum_{\ell} \ell Y_{\ell} \left(\int_0^t \beta_{\ell}^n(X^n(s)) ds \right),$$

where $\ell \in \mathcal{L}^n = \{y - x : x, y \in E^n\}$ and $\{Y_{\ell}, \ell \in \mathcal{L}^n\}$ are independent unit-rate Poisson processes; see [10], Section 6.4. Letting $\tilde{Y}_{\ell}(t) = Y_{\ell}(t) - t$, we rewrite

$$X^n(t) = X^n(0) + \int_0^t F^n(X^n(s)) ds + \sum_{\ell} \ell \tilde{Y}_{\ell} \left(\int_0^t \beta_{\ell}^n(X^n(s)) ds \right),$$

where

$$(3) \quad F_i^n(x) = \sum_{\ell} \ell_i \beta_{\ell}^n(x).$$

Provided that X^n is nonexplosive,

$$M^n(t) = \sum_{\ell} \ell \tilde{Y}_{\ell} \left(\int_0^t \beta_{\ell}^n(X^n(s)) ds \right),$$

is a local martingale with respect to the filtration

$$(4) \quad \mathcal{F}_t^n = \sigma \left\{ X^n(0), \int_0^s \beta_{\ell}^n(X^n(u)) du, \tilde{Y}_{\ell} \left(\int_0^s \beta_{\ell}^n(X^n(u)) du \right); \ell \in \mathcal{L}^n, s \leq t \right\};$$

see [10], Theorem 6.4.1. The local (predictable) quadratic variation of M^n is given by

$$\langle M^n \rangle(t) = \int_0^t a^n(X^n(s)) ds,$$

where

$$(5) \quad a_{ij}^n(x) = \sum_{\ell} \ell_i \ell_j \beta_{\ell}^n(x).$$

In essence, F^n and a^n are defined only for values in E^n . We henceforth assume that they are extended to \mathbb{R}^d and, with some abuse of notation, denote by F^n and a^n these extensions. The requirements that we impose on these extensions will be clear in what follows.

Fluid models. Given x , we define the n th fluid model by

$$(FM) \quad \bar{x}^n(t) = x + \int_0^t F^n(\bar{x}^n(s)) ds,$$

or, in differential form,

$$\dot{\bar{x}}^n(t) = F^n(\bar{x}^n(t)), \quad \bar{x}^n(0) = x.$$

If F^n is Lipschitz continuous, the fluid model has a solution. We will assume that there exists a unique \bar{x}_{∞}^n satisfying

$$(6) \quad F^n(\bar{x}_{\infty}^n) = 0.$$

This requirement is intimately linked to our Lyapunov requirement; see Lemma 3.1.

Centered and scaled process. Define the processes

$$(7) \quad \widehat{X}^n = \frac{X^n - \bar{x}_{\infty}^n}{\sqrt{n}}, \quad \widehat{M}^n = \frac{M^n}{\sqrt{n}},$$

and denote by \widehat{E}^n the state space of \widehat{X}^n . Letting

$$\widehat{F}^n(x) = \frac{F^n(\bar{x}_{\infty}^n + \sqrt{n}x)}{\sqrt{n}}, \quad x \in \mathbb{R}^d,$$

we have

$$\widehat{X}^n(t) = \widehat{X}^n(0) + \int_0^t \widehat{F}^n(\widehat{X}^n(s)) ds + \widehat{M}^n(t).$$

The martingale \widehat{M}^n has the local predictable quadratic variation process

$$(8) \quad \langle \widehat{M}^n \rangle(t) = \int_0^t \bar{a}^n(\widehat{X}^n(s)) ds,$$

where

$$\bar{a}^n(x) = \frac{a^n(\bar{x}_{\infty}^n + x\sqrt{n})}{n}, \quad x \in \mathbb{R}^d.$$

Assumptions. We assume that the jump sizes are bounded uniformly in n :

$$(9) \quad \bar{\ell} = \sup_n \operatorname{argmax}\{|\ell| \in \mathcal{L}^n\} < \infty,$$

and that n is sufficiently large so that $\bar{\ell}/\sqrt{n} \leq 1$.

The sequence $\{\widehat{F}^n\}$ is assumed to be uniformly Lipschitz, and $\{\bar{a}^n\}$ is assumed to have linear growth around 0. Formally, there exist constants K_F, K_a such that, for all n ,

$$(10) \quad |\widehat{F}^n(x) - \widehat{F}^n(y)| \leq K_F|x - y|, \quad x, y \in \mathbb{R}^d$$

and

$$(11) \quad |\bar{a}^n(x) - \bar{a}^n(0)| \leq \frac{K_a}{\sqrt{n}}|x|, \quad x \in \mathbb{R}^d.$$

The requirements (10) and $\widehat{F}^n(0) = F^n(\bar{x}_n^\infty)/\sqrt{n} = 0$ guarantee, in particular, that $|\widehat{F}^n(x)| \leq 1 + K_F|x|$. Condition (11) is equivalently stated in terms of the (unscaled) a^n as

$$|a^n(x) - a^n(\bar{x}_\infty^n)| \leq K_a|x - \bar{x}_\infty^n|, \quad x \in \mathbb{R}^d.$$

We further assume that $\bar{a}^n(0)$ is positive definite for each n and that

$$(12) \quad \bar{a}^n(0) \rightarrow \bar{a},$$

where \bar{a} is itself positive definite. The matrix \bar{a} is not used in specifying the diffusion model in Section 3, but the assumption of convergence is used in our proofs, most notably in that of Theorem 3.1. In various settings, including our own examples in Section 6, $\bar{a}^n(0) \equiv \bar{a}$ in which case the convergence requirement is trivially satisfied.

The requirement that the continuous extension \widehat{F}^n satisfies the uniform Lipschitz requirement (10) is a restriction. It excludes, for example, single-server queueing systems; we revisit this point in Section 8.

ASSUMPTION 2.1. For each $n \in \mathbb{N}$, X^n is nonexplosive, irreducible, positive recurrent and satisfies (9)–(12).

Positive recurrence and irreducibility imply ergodicity of X^n and, in particular, the existence of a steady-state distribution (which is also the unique stationary distribution). In certain cases, positive recurrence of X^n need not be a priori assumed; see Theorem 3.3 and Remark 3.5.

Assumption 2.1 is imposed for the remainder of this paper.

3. A diffusion model. Recall that \bar{x}_∞^n is a stationary point for the fluid model

$$(FM) \quad \bar{x}^n(t) = \bar{x}^n(0) + \int_0^t F^n(\bar{x}^n(s)) ds,$$

and that $\bar{a}^n(0) = a^n(\bar{x}_\infty^n)/n$. Fix a probability space and a d -dimensional Brownian motion, and let \widehat{Y}^n be the strong solution to the SDE

$$(DM) \quad \widehat{Y}^n(t) = y + \int_0^t \widehat{F}^n(\widehat{Y}^n(s)) ds + \sqrt{\bar{a}^n(0)}B(t).$$

The existence and uniqueness of a strong solution follow from the Lipschitz continuity and linear growth of \widehat{F}^n and the constant diffusion coefficient; see, for example, [17], Theorems 5.2.5 and 5.2.9.

REMARK 3.1 (On strong approximations). The strong approximation for \widehat{X}^n is a diffusion obtained (heuristically at first) by taking the “density” $\bar{a}^n(x)$ of the quadratic variation in (8) as the diffusion coefficient, to define the process

$$\widehat{S}^n(t) = y + \int_0^t \widehat{F}^n(\widehat{S}^n(s)) ds + \int_0^t \sqrt{\bar{a}^n(\widehat{S}^n(s))} dB(s).$$

The process \widehat{S}^n provides a “good” approximation for the dynamics of the CTMC in the sense that

$$\sup_{0 \leq t \leq T} |\widehat{X}^n(t) - \widehat{S}^n(t)| \leq \Gamma_T^n \log(n),$$

where $\{\Gamma_T^n\}$ are random variables with exponential tails (uniformly in n); see, for example, [10], Chapters 7.5 and 11.3. Given the cruder (state independent) diffusion coefficient, the DM \widehat{Y}^n is not likely to be as precise, over finite horizons, as the strong approximation. In terms of tractability, however, the analysis of steady-state is simpler for the DM, insofar as its steady-state distribution (when it exists) involves linear PDEs; see, for example, [18], Chapter 4.9. Our main result, Theorem 3.2, shows that this increased tractability co-exists with an impressive steady-state-approximation accuracy.

REMARK 3.2 (On the diffusion model and the diffusion limit). Suppose that, in addition, Assumption 2.1

$$(13) \quad \frac{\beta_\ell^n(\bar{x}_\infty^n + \sqrt{n}x) - \beta_\ell^n(\bar{x}_\infty^n)}{\sqrt{n}} \rightarrow \widehat{\beta}_\ell(x),$$

uniformly on compact subsets of \mathbb{R}^d . If $\widehat{X}^n(0) \Rightarrow y$, then

$$\widehat{X}^n \Rightarrow \widehat{Y},$$

where \widehat{Y} is the strong solution to the SDE

$$\widehat{Y}(t) = y + \int_0^t \widehat{F}(\widehat{Y}(s)) ds + \sqrt{\bar{a}}B(t),$$

with $\widehat{F}(x) = \sum_\ell \ell \widehat{\beta}_\ell(x)$ and \bar{a} is as in (12); see [10], Theorem 6.5.4. Given (13), requirements (5.9) and (5.14) of that theorem are trivially satisfied here due to the bounded jumps. The final requirement in [10], Theorem 6.5.4, that $\tau_a = \inf\{t \geq 0 : |\widehat{Y}(t)| \geq a\}$ has $\tau_a \rightarrow \infty$ almost surely, follows immediately from the fact that \widehat{Y} is a strong solution. Further, it is easily proved that $\widehat{Y}^n \Rightarrow \widehat{Y}$. Thus, within a diffusion-limit framework, the DM is consistent with the diffusion limit in the sense that \widehat{Y}^n and \widehat{X}^n converge to the same limit.

For functions $f \in \mathcal{C}^2(\mathbb{R}^d)$, the generator of \widehat{Y}^n coincides with the second order differential operator \mathcal{A}^n defined, for such functions, by

$$(14) \quad \mathcal{A}^n f(x) = \sum_{i=1}^d \widehat{F}_i^n(x) \frac{\partial}{\partial x_i} f(x) + \frac{1}{2} \sum_{i,j} \bar{a}_{ij}^n(0) \frac{\partial^2}{\partial x_i \partial x_j} f(x);$$

see, for example, [17], Proposition 5.4.2.

We next state the uniform Lyapunov assumption. We say that $V \in \mathcal{C}^2(\mathbb{R}^d)$ is a *norm-like* function if $V(x) \rightarrow \infty$ as $|x| \rightarrow \infty$. A function $V \in \mathcal{C}^2(\mathbb{R}^d)$ is said to be sub-exponential if $V \geq 1$ and there exist constants c_1, c_2 and c_3 such that

$$(15) \quad |DV(x)| \vee |D^2V(x)| \leq c_1 e^{c_2|x|}, \quad x \in \mathbb{R}^d$$

and

$$(16) \quad \sup_{y:|y|\leq 1} \frac{V(x+y)}{V(x)} \leq c_3, \quad x \in \mathbb{R}^d.$$

ASSUMPTION 3.1. There exist a sub-exponential norm-like function $V \in \mathcal{C}^2(\mathbb{R}^d)$ and finite positive constants b, δ, K (not depending on n) such that

$$(UL) \quad \mathcal{A}^n V(x) \leq -\delta V(x) + b \mathbb{1}_{\bar{B}_0(K)}(x) \quad \text{for all } x \in \mathbb{R}^d,$$

and, for each n and all $x \in \widehat{E}^n$,

$$(17) \quad \mathbb{E}_x \left[\int_0^t ((1 + |\widehat{X}^n(s)|)^4 V(\widehat{X}^n(s)))^2 ds \right] < \infty, \quad t \geq 0.$$

Assumption 3.1 is imposed for the remainder of this paper. The requirement that $V \geq 1$ is made without loss of generality. If a norm-like function V satisfies UL, there exists re-defined constants b, δ and K such that $1 + V$ satisfies UL. All polynomials $V \geq 1$ satisfy (15) and (16)—the former is used only in the proof of Lemma 7.2, and the latter is used in the derivations of gradient bounds following the statement of Theorem 4.1. Requirement (17) is relatively unrestrictive as it is imposed on each individual n (rather than uniformly in n).

Lyapunov conditions are frequently used in the context of stability of continuous time Markov processes (corresponding to fixed n here); see [20]. The requirement of a uniform Lyapunov condition imposed on a family of Markov processes is less common (see [11] for a related example). In Section 6 we study two examples for which all the requirements of Assumption 3.1 are met.

With Assumption 3.1, the existence and uniqueness of a steady-state distribution, π^n , for \widehat{Y}^n follows from [20], Sections 4 and 6, as does the fact that \widehat{Y}^n is exponentially ergodic and that, for each $n, \pi^n(|f|) < \infty$ for all functions f with $|f| \leq V$; see [20], Theorem 4.2. For V that satisfies (15) we have, for all $t \geq 0$ and $x \in \mathbb{R}^d$, that

$$(18) \quad \mathbb{E}_x[V(\widehat{Y}^n(t))] = V(x) + \mathbb{E}_x \left[\int_0^t \mathcal{A}^n V(\widehat{Y}^n(s)) ds \right];$$

see, for example, [19], Theorem 6.3. UL then guarantees that

$$(19) \quad \mathbb{E}_x[V(\widehat{Y}^n(t))] \leq V(x) + \mathbb{E}_x \left[\int_0^t (-\delta V(\widehat{Y}^n(s)) + b) ds \right]$$

for all $t \geq 0$ and $x \in \mathbb{R}^d$ and, consequently, that

$$(20) \quad \limsup_{n \rightarrow \infty} \pi^n(|f|) \leq \frac{b}{\delta}$$

for all functions f with $|f| \leq V$; see also [14], Corollary 2.

Important for our analysis is the following consequence of Assumption 3.1.

THEOREM 3.1 (Uniform exponential ergodicity). *Let π^n be the steady-state distribution of \widehat{Y}^n . Then there exist finite positive constants \mathcal{M} and μ such that*

$$(21) \quad \sup_n \sup_{x \in \mathbb{R}^d} \sup_{|f| \leq V} \frac{1}{V(x)} |\mathbb{E}_x[f(\widehat{Y}^n(t))] - \pi^n(f)| \leq \mathcal{M}e^{-\mu t}, \quad t \geq 0.$$

Bounds on the convergence rate of exponentially ergodic Markov processes to their steady-state distribution have been studied extensively in recent literature. Our proof builds specifically on [2]. The constants \mathcal{M} and μ are related to a minorization condition for the discrete-time process $\{\widehat{Y}^n(m), m \in \mathbb{Z}_+\}$. In the standard application, these constants may depend on n . To obtain constants that can be used for all $n \in \mathbb{N}$ we must argue that a minorization condition is satisfied uniformly in n ; the proof of Theorem 3.1 is postponed to Section 7.

Theorem 3.1 has the following important implication: fixing a function f with $|f| \leq V$ and $\pi^n(f) = 0$, we have for all $x \in \mathbb{R}^d$, that

$$\sup_n |\mathbb{E}_x[f(\widehat{Y}^n(t))]| \leq \mathcal{M}V(x)e^{-\mu t}, \quad t \geq 0,$$

so that

$$\sup_n \int_0^\infty |\mathbb{E}_x[f(\widehat{Y}^n(s))]| ds \leq \mathcal{M}V(x) \int_0^\infty e^{-\mu s} ds = CV(x) < \infty$$

for all $x \in \mathbb{R}^d$, where the constant C does not depend on n or x . We conclude that

$$u_f^n(x) = \int_0^\infty \mathbb{E}_x[f(\widehat{Y}^n(s))] ds$$

is a well-defined function of $x \in \mathbb{R}^d$ and that, for all n ,

$$(22) \quad |u_f^n(x)| \leq CV(x), \quad x \in \mathbb{R}^d.$$

Also, for any fixed $M > 0$ and $n \in \mathbb{N}$,

$$(23) \quad \sup_{x \in \overline{B}_0(M)} \lim_{t \rightarrow \infty} \left| \int_0^t \mathbb{E}_x[f(\widehat{Y}^n(s))] ds - \int_0^\infty \mathbb{E}_x[f(\widehat{Y}^n(s))] ds \right| = 0.$$

Define

$$(24) \quad \mathcal{B}_x = B_x \left(\frac{1}{1 + |x|} \right), \quad x \in \mathbb{R}^d$$

and

$$(25) \quad \bar{f}(x) = \sup_{y \in \mathcal{B}_x} |f(y)| + \sup_{y, z \in \mathcal{B}_x} \frac{|f(y) - f(z)|}{|y - z|}.$$

The introduction of \bar{f} is motivated by the analysis of the (sequence of) Poisson equations, specifically by the gradient estimates that require bounds on local fluctuations of f ; see the derivations following Theorem 4.1.

Our main result, stated next, establishes that the steady-state distribution of the Markov chain and the DM are suitably close provided that moments of the former are uniformly bounded.

THEOREM 3.2. *Fix V that satisfies Assumption 3.1 and a function f such that $\pi^n(f) = 0$ and $\bar{f} \leq V$. Let ν^n and π^n be, respectively, the steady-state distributions of \hat{X}^n and \hat{Y}^n . If*

$$(26) \quad \limsup_{n \rightarrow \infty} \nu^n(V(\cdot)(1 + |\cdot|)^4) < \infty,$$

then

$$\nu^n(f) - \pi^n(f) = \mathcal{O}(1/\sqrt{n}).$$

Theorem 3.2 and the remaining results of this section are proved in Section 5.

REMARK 3.3. If f satisfies $\bar{f} \leq V$ but $\pi^n(f) \neq 0$, consider instead the function $\check{f}^n = f - \pi^n(f)$. Then $\pi^n(\check{f}^n) = 0$. By (20), $\limsup_{n \rightarrow \infty} \pi^n(|f|) \leq b/\delta < \infty$ and, in turn, $\limsup_{n \rightarrow \infty} \pi^n(|\check{f}^n|) \leq 2b/\delta < \infty$. Further, \check{f}^n satisfies that $\bar{\check{f}}^n \leq \bar{f} + \pi^n(|f|) \leq V + b/\delta$. Finally, if V satisfies Assumption 3.1, so does the function $\check{V} = V + b/\delta$. Thus the results that follow hold for functions f with $\bar{f} \leq V$ regardless of whether $\pi^n(f) = 0$ or not.

In general, proving requirement (26) (which implies, in particular, tightness of the sequence $\{\nu^n\}$ of steady-state distributions) is far from trivial. As we show next (26) can be argued in advance in our setting. One expects that, as n grows, the property (19) of the DM will be approximately valid for the CTMC allowing to draw an implication similar to (20) with \hat{Y}^n there replaced by \hat{X}^n . The next theorem shows that this intuition is valid provided that V satisfies additional simple properties.

Given a function $\Psi \in \mathcal{C}(\mathbb{R}^d)$, define for $x \in \mathbb{R}^d$,

$$(27) \quad [\Psi]_{2,1, B_x(\bar{\ell}/\sqrt{n})} = \sup_{y, z \in B_x(\bar{\ell}/\sqrt{n})} \frac{|D^2\Psi(y) - D^2\Psi(z)|}{|y - z|},$$

where the right-hand side may be infinite.

THEOREM 3.3 [From DM to CTMC Lyapunov]. *Let V be as in Assumption 3.1. Suppose, in addition, that there exists a finite positive constant C such that, for each n , and all $x \in \mathbb{R}^d$,*

$$(28) \quad (|DV(x)| + |D^2V(x)| + [V]_{2,1,B_x(\bar{\ell}/\sqrt{n})})(1 + |x|) \leq CV(x).$$

Then, for all sufficiently large n , and all $x \in \widehat{E}^n$,

$$(29) \quad \mathbb{E}_x[V(\widehat{X}^n(t))] \leq V(x) + \mathbb{E}_x \left[\int_0^t \left(-\frac{\delta}{2} V(\widehat{X}^n(s)) + b \right) ds \right], \quad t \geq 0,$$

where b is as in Assumption 3.1. Consequently, \widehat{X}^n is ergodic for all such n and, furthermore,

$$\limsup_{n \rightarrow \infty} v^n(V) \leq \frac{2b}{\delta}.$$

If $V \in \mathcal{C}^3(\mathbb{R}^d)$, condition (28) can be replaced with

$$(30) \quad (|DV(x)| + |D^2V(x)| + |D^3V(x)|)(1 + |x|) \leq CV(x).$$

Using Taylor’s theorem we have, for all $x \in \mathbb{R}^d$, that

$$\begin{aligned} (1 + |x|)[V]_{2,1,B_x(\bar{\ell}/\sqrt{n})} &\leq \sup_{\eta \in B_x(\bar{\ell}/\sqrt{n})} 2(1 + |\eta|)|D^3V(\eta)| \\ &\leq 2C \left(\sup_{\eta \in B_x(\bar{\ell}/\sqrt{n})} V(\eta) \right) \leq 2c_3 CV(x), \end{aligned}$$

where the last inequality follows from the sub-exponential property (16) of V and $\bar{\ell}/\sqrt{n} \leq 1$. Note that (30) is satisfied by any polynomial $V \geq 1$.

COROLLARY 3.4. *Fix V that satisfies Assumption 3.1. Suppose that there exists \bar{V} that, itself, satisfies Assumption 3.1 as well as (28) and*

$$V(\cdot)(1 + |\cdot|)^4 \leq \bar{V}(\cdot).$$

Then,

$$\limsup_{n \rightarrow \infty} v^n(\bar{V}) < \infty,$$

and, in particular, (26) holds for V .

REMARK 3.4 (A simple case). Suppose that $V \in \mathcal{C}^3(\mathbb{R}^d)$ and satisfies Assumption 3.1 and (30). If there exists $m \in \mathbb{N}$ such that $V_m(\cdot) = (V(\cdot))^m \geq V(\cdot)(1 +$

$|\cdot|)^4$ and V_m satisfies (17), then we can take $\bar{V} = V_m$ in Corollary 3.4. Indeed, for an integer $m \geq 2$,

$$\begin{aligned} \mathcal{A}^n V_m(x) &= mV_{m-1}(x)\mathcal{A}^n \Psi(x) + m(m-1)V_{m-2}(x) \frac{1}{2} \sum_{i,j}^d \bar{a}_{ij}^n(0) \frac{\partial}{\partial x_i} V(x) \\ &\leq -\delta m V_m(x) + bm V_{m-1}(x) + m(m-1)C V_{m-1}(x), \end{aligned}$$

with δ and b as in Assumption 3.1 and C as in (30). Thus if $V \in \mathcal{C}^3(\mathbb{R}^d)$ is sub-exponential and satisfies UL and (30), so does V_m .

REMARK 3.5 (A unified set of conditions). Combined, Theorem 3.2 and Corollary 3.4 establish the following: If there exist functions V and \bar{V} both satisfying Assumption 3.1 such that (28) holds for \bar{V} and $V(\cdot)(1 + |\cdot|)^4 \leq \bar{V}(\cdot)$, then we *simultaneously* have: (i) the positive recurrence of \bar{X}^n for sufficiently large n , (ii) the moment bound in (26) (which implies, in particular, the tightness of ν^n) and (iii) the $\mathcal{O}(1/\sqrt{n})$ approximation gap.

With the exception of the simple requirement (17), this reduces the requirements to properties of the DM.

We conclude this section with an observation pertaining to the connection between the stability of the FM and the DM. Suppose that there exist a norm-like function V and a constant η such that

$$(31) \quad V(x) > V(0) \quad \text{and} \quad \widehat{F}^n(x)' DV(x) \leq -\eta(V(x) - V(0)), \quad x \neq 0.$$

Letting $V^n(x) = V(\frac{x - \bar{x}_\infty^n}{\sqrt{n}}) - V(0)$ we have

$$F^n(x)' DV^n(x) \leq -\eta V^n(x), \quad x \neq \bar{x}_\infty^n,$$

so that the FM is stable in the sense that, for each n and any initial condition $\bar{x}^n(0) \in \mathbb{R}^d$, $\bar{x}^n(t) \rightarrow \bar{x}_\infty^n$ as $t \rightarrow \infty$. Moreover,

$$(32) \quad \begin{aligned} \mathcal{A}^n V(y) &\leq \widehat{F}^n(y)' DV(y) + |\bar{a}^n(0)| |D^2 V(y)| \\ &\leq -\eta(V(y) - V(0)) + |\bar{a}^n(0)| |D^2 V(y)|. \end{aligned}$$

The following is an immediate consequence.

LEMMA 3.1 [FM and DM stability]. *Let $V \in \mathcal{C}^2(\mathbb{R}^d)$ be a sub-exponential norm-like function satisfying (17) and (31). If*

$$\limsup_{|x| \rightarrow \infty} \frac{|D^2 V(x)|}{V(x)} = 0,$$

then V satisfies UL and, in turn, Assumption 3.1.

4. A sequence of Poisson equations. In what follows, fixing a set $\mathcal{B} \subseteq \mathbb{R}^d$, $\mathcal{C}^2(\mathcal{B})$ denotes the space of twice continuously differentiable functions from \mathcal{B} to \mathbb{R} . For $u \in \mathcal{C}^2(\mathcal{B})$, recall that Du and D^2u denote the gradient and the Hessian of u , respectively. The space $\mathcal{C}^{2,1}(\mathcal{B})$ is then the subspace of $\mathcal{C}^2(\mathcal{B})$ members of which have second derivatives that are Lipschitz continuous on \mathcal{B} . That is, a twice continuously differentiable function $u : \mathbb{R}^d \rightarrow \mathbb{R}$ is in $\mathcal{C}^{2,1}(\mathcal{B})$ if

$$[u]_{2,1,\mathcal{B}} = \sup_{x,y \in \mathcal{B}, x \neq y} \frac{|D^2u(x) - D^2u(y)|}{|x - y|} < \infty.$$

[In equation (27) the set \mathcal{B} is taken to be $B_x(\bar{\ell}/\sqrt{n})$.] We define $d_x = \text{dist}(x, \partial\mathcal{B}) = \inf\{|x - y|, y \in \partial\mathcal{B}\}$ where $\partial\mathcal{B}$ stands for the boundary of \mathcal{B} , and we let $d_{x,z} = \min\{d_x, d_z\}$. We define

$$(33) \quad |u|_{2,1,\mathcal{B}}^* = \sum_{j=0}^2 [u]_{j,\mathcal{B}}^* + \sup_{x,y \in \mathcal{B}, x \neq y} d_{x,y}^3 \frac{|D^2u(x) - D^2u(y)|}{|x - y|},$$

where $[u]_{j,\mathcal{B}}^* = \sup_{x \in \mathcal{B}} d_x^j |D^j u(x)|$ for $j = 0, 1, 2$. Above d_x^j (resp., $d_{x,y}^j$) denotes the j th power of d_x (resp., of $d_{x,y}$). We let $|u|_{0,\mathcal{B}} = [u]_{0,\mathcal{B}}^* = \sup_{x \in \mathcal{B}} |u(x)|$, and

$$|f|_{0,1,\mathcal{B}}^{(2)} = \sup_{x \in \mathcal{B}} d_x^2 |f(x)| + \sup_{x,y \in \mathcal{B}} d_{x,y}^3 \frac{|f(x) - f(y)|}{|x - y|}.$$

We say that the function is locally Lipschitz if $|f|_{0,1,\mathcal{B}_x}^{(2)} < \infty$ for all $x \in \mathbb{R}^d$, where \mathcal{B}_x is as in (24).

THEOREM 4.1. *Fix V that satisfies Assumption 3.1 and a locally Lipschitz function f with $|f| \leq V$ and $\pi^n(f) = 0$. Then, for each n , the Poisson equation*

$$(34) \quad \mathcal{A}^n u = -f$$

has a unique solution $u_f^n \in \mathcal{C}^2(\mathbb{R}^d)$ given by

$$(35) \quad u_f^n(x) = \int_0^\infty \mathbb{E}_x[f(\widehat{Y}^n(t))] dt.$$

Moreover, there exist a finite positive constant Θ (not depending on n) such that

$$|u_f^n|_{2,1,\mathcal{B}_x}^* \leq \Theta(|u_f^n|_{0,\mathcal{B}_x} + |f|_{0,1,\mathcal{B}_x}^{(2)}), \quad x \in \mathbb{R}^d.$$

Consequently, for all n and $x \in \mathbb{R}^d$,

$$(36) \quad |Du_f^n(x)| \leq 2\Theta(|u_f^n|_{0,\mathcal{B}_x} + |f|_{0,1,\mathcal{B}_x}^{(2)})(1 + |x|),$$

$$(37) \quad |D^2u_f^n(x)| \leq 4\Theta(|u_f^n|_{0,\mathcal{B}_x} + |f|_{0,1,\mathcal{B}_x}^{(2)})(1 + |x|)^2$$

and

$$(38) \quad [u_f^n]_{2,1,\mathcal{B}_x} \leq 8\Theta(|u_f^n|_{0,\mathcal{B}_x} + |f|_{0,1,\mathcal{B}_x}^{(2)})(1 + |x|)^3.$$

Several observations are useful for what follows: recall (22) that $|u_f^n(x)| \leq CV(x)$ for some constant C . By the assumed sub-exponentiality of V

$$|u_f^n|_{0, B_y} \leq \sup_{z \in B_y} CV(z) \leq c_3 CV(y)$$

for all $y \in \mathbb{R}^d$, where c_3 is as in (16). In turn,

$$\sup_{y \in B_x(\bar{\ell}/\sqrt{n})} |u_f^n|_{0, B_y} \leq \sup_{y \in B_x(\bar{\ell}/\sqrt{n})} c_3 CV(y) \leq c_3^2 CV(x).$$

For a function f with $\bar{f} \leq V$ [see (25)] and for all $y \in \mathbb{R}^d$,

$$|f|_{0,1, B_y}^{(2)} \leq \bar{f}(y) \leq V(y),$$

so that

$$\sup_{y \in B_x(\bar{\ell}/\sqrt{n})} |f|_{0,1, B_y}^{(2)} \leq c_3 V(x)$$

for all $x \in \mathbb{R}^d$. Defining

$$(39) \quad C_V(x) = 16\Theta(1 + c_3^2 C)V(x)(1 + |x|)^3, \quad x \in \mathbb{R}^d,$$

we have, by Theorem 4.1 (and assuming, without loss of generality that $c_3 \geq 1$), that for all $n \in \mathbb{N}$ and $x \in \mathbb{R}^d$,

$$(40) \quad \begin{aligned} |Du_f^n(x)| &\leq C_V(x)/(1 + |x|)^2, \\ |D^2u_f^n(x)| &\leq C_V(x)/(1 + |x|) \quad \text{and} \\ [u_f^n]_{2,1, B_x(\bar{\ell}/\sqrt{n})} &\leq C_V(x). \end{aligned}$$

PROOF OF THEOREM 4.1. We first prove that u_f^n in (35) solves the Poisson equation (34). Since f is fixed throughout we omit it from the notation.

Fixing M , let u_M^n be the solution to Dirichlet problem

$$\begin{aligned} \mathcal{A}^n u(x) &= -f(x), \quad x \in B_0(M); \\ u &= u^n, \quad x \in \partial B_0(M). \end{aligned}$$

In the boundary condition, u^n is as in (35). The existence and uniqueness of a solution $u_M^n \in \mathcal{C}^0(\bar{B}_0(M)) \cap \mathcal{C}^{2,1}(B_0(M))$ follows directly from [13], Theorem 6.13, recalling that \widehat{F}^n is Lipschitz continuous and $\bar{a}^n(0)$ is a constant matrix and hence trivially Lipschitz. Theorem 6.13 of [13] requires that u_M^n is continuous in x on $\partial B_0(M)$. This follows exactly as in part (c) of [21], Theorem 1, using (23). We omit the detailed argument.

It follows that

$$u_M^n(x) = \mathbb{E}_x \left[\int_0^{\tau_M^n} f(\widehat{Y}^n(s)) ds \right],$$

where $\tau_M^n = \inf\{t \geq 0 : \widehat{Y}^n(t) \notin B_0(M)\}$; see [17], Proposition 5.7.2 and Lemma 5.7.4. We have that

$$u_M^n(x) = u^n(x) \quad \text{for all } x \in B_0(M),$$

with $u^n(x)$ as in (35). This assertion is proved as in [21], Theorem 1, part (d). Since M is arbitrary we conclude that, $u^n(x)$ solves the Poisson equation (34).

To establish the gradient estimates observe that, since $\bar{a}^n(0)$ is bounded in n , there exists a constant C_a (not depending on n) such that (with the notation in [13], Theorem 6.2) $|\bar{a}^n(0)|_{0,1,\mathcal{B}_x}^{(0)} \leq C_a$. From the positive definiteness of $\bar{a}^n(0)$, and since $\bar{a}^n(0) \rightarrow \bar{a}$ for a positive definite \bar{a} , it follows that there exists a constant $\lambda > 0$ such that

$$(41) \quad \sum_{ij} \bar{a}_{ij}^n(0) \xi_i \xi_j \geq \lambda |\xi|^2$$

for all n and all $\xi \in \mathbb{R}^d$. Finally, following the notation in [13], Theorem 6.2,

$$\begin{aligned} |\widehat{F}^n|_{0,1,\mathcal{B}_x}^{(1)} &= |\widehat{F}^n|_{0,\mathcal{B}_x}^{(1)} + [\widehat{F}^n]_{0,1,\mathcal{B}_x}^{(1)} \\ &= [\widehat{F}^n]_{0,\mathcal{B}_x}^{(1)} + \sup_{y,z \in \mathcal{B}_x} d_{y,z}^2 \frac{|\widehat{F}^n(y) - \widehat{F}^n(z)|}{|y - z|} \\ &= \sup_{y \in \mathcal{B}_x} d_y |\widehat{F}^n(y)| + \sup_{y,z} d_{y,z}^2 \frac{|\widehat{F}^n(y) - \widehat{F}^n(z)|}{|y - z|} \\ &\leq 2K_F, \end{aligned}$$

where K_F is as in (10). In turn, by [13], Theorem 6.2, that

$$|u_f^n|_{2,1,\mathcal{B}_x}^* \leq \Theta(|u_f^n|_{0,\mathcal{B}_x} + |f|_{0,1,\mathcal{B}_x}^{(2)}),$$

where Θ depends only on K_F, C_a, d and the constant λ in (41) (for Λ there, we take $K_F \vee C_a$). Bounds (36)–(38) now follow from the definition of $|u_f^n|_{2,1,\mathcal{B}_x}^*$ applied to points in the subset $B_x(1/(2(1 + |x|)))$ of \mathcal{B}_x . Specifically, for each $y \in \mathcal{B}_x$,

$$d_y |Du_f^n(y)| \leq [u]_{1,\mathcal{B}_x}^* \leq |u_f^n|_{2,1,\mathcal{B}_x}^*.$$

Noting that $d_y \geq 1/(2(1 + |x|))$ for all $y \in B_x(1/(2(1 + |x|)))$ we have, for all such y (in particular for x itself), that

$$|Du_f^n(y)| \leq |u_f^n|_{2,1,\mathcal{B}_x}^* (1 + |x|).$$

Equations (37) and (38) are argued similarly. \square

5. Proofs of Theorems 3.2 and 3.3. The following simple lemma is proved in the Appendix. Given a function $\Psi \in C^2(\mathbb{R}^d)$ we write Ψ_i for the i th coordinate of $D\Psi$ and Ψ_{ij} for the ij th coordinate of $D^2\Psi$.

LEMMA 5.1. *Let $\Psi \in C^2(\mathbb{R}^d)$ be such that, for all $x \in \widehat{E}^n$ and $t \geq 0$,*

$$(42) \quad \mathbb{E}_x \left[\int_0^t (|D\Psi(\widehat{X}^n(s))| + |D^2\Psi(\widehat{X}^n(s))| + [\Psi]_{2,1,B_{\widehat{X}^n(s)}(\bar{\ell}/\sqrt{n})})(1 + |\widehat{X}^n(s)|) ds \right] < \infty.$$

Then, for all $x \in \widehat{E}^n$ and $t \geq 0$,

$$(43) \quad \mathbb{E}_x[\Psi(\widehat{X}^n(t))] = \Psi(x) + \mathbb{E}_x \left[\int_0^t \mathcal{A}^n \Psi(\widehat{X}^n(s)) ds \right] + A_{\Psi}^{n,x}(t) + D_{\Psi}^{n,x}(t),$$

where \mathcal{A}^n is as in (14) and, for all $x \in \widehat{E}^n$ and $t \geq 0$,

$$|A_{\Psi}^{n,x}(t)| \leq \frac{\bar{\ell}}{2\sqrt{n}} \mathbb{E}_x \left[\int_0^t [\Psi]_{2,1,B_{\widehat{X}^n(s)}(\bar{\ell}/\sqrt{n})} |\bar{a}^n(\widehat{X}^n(s))| ds \right],$$

$$D_{\Psi}^{n,x}(t) = \frac{1}{2} \mathbb{E}_x \left[\sum_{i,j} \int_0^t \Psi_{ij}(\widehat{X}^n(s)) (\bar{a}_{ij}^n(\widehat{X}^n(s)) - \bar{a}_{ij}^n(0)) ds \right].$$

Below \bar{f} is as in (25) and C_V as in (39).

COROLLARY 5.1. *Fix V that satisfies Assumption 3.1 and a function f such that $\bar{f} \leq V$. Then there exists a finite positive constant C (not depending on n), such that, for all $x \in \widehat{E}^n$ and $t \geq 0$,*

$$\left| \mathbb{E}_x[u_f^n(\widehat{X}^n(t))] - u_f^n(x) - \mathbb{E}_x \left[\int_0^t \mathcal{A}^n u_f^n(\widehat{X}^n(s)) ds \right] \right| \leq C \left(\mathbb{E}_x \left[\int_0^t \frac{C_V(\widehat{X}^n(s))}{\sqrt{n}} \left(1 + \frac{|\widehat{X}^n(s)|}{\sqrt{n}} \right) ds \right] \right).$$

PROOF. By (40) we have, for $x \in \mathbb{R}^d$, that

$$\begin{aligned} (|Du_f^n(x)| + |D^2u_f^n(x)| + [u_f^n]_{2,1,B_x(\bar{\ell}/\sqrt{n})})(1 + |x|) &\leq 3C_V(x)(1 + |x|) \\ &\leq \varepsilon(1 + |x|)^4 V(x) \end{aligned}$$

for some finite positive constant. By Assumption 3.1, specifically (17),

$$\mathbb{E}_x \left[\int_0^t (1 + |\widehat{X}^n(s)|)^4 (V(\widehat{X}^n(s)))^2 ds \right] < \infty,$$

so that V satisfies the requirements of Lemma 5.1, and we have that

$$\begin{aligned}
 |D_{u_f^n}^{n,x}(t)| &\leq \frac{1}{2} \mathbb{E}_x \left[\int_0^t |D^2 u_f^n(\widehat{X}^n(s))| |\bar{a}^n(\widehat{X}^n(s)) - \bar{a}^n(0)| ds \right] \\
 (44) \qquad &\leq \frac{K_a}{2\sqrt{n}} \mathbb{E}_x \left[\int_0^t |D^2 u_f^n(\widehat{X}^n(s))| |\widehat{X}^n(s)| ds \right] \\
 &\leq \frac{K_a}{2\sqrt{n}} \mathbb{E}_x \left[\int_0^t C_V(\widehat{X}^n(s)) ds \right].
 \end{aligned}$$

The second inequality follows from (11). The last inequality follows from (40).

Next,

$$\begin{aligned}
 |A_{u_f^n}^{n,x}(t)| &\leq \frac{\bar{\ell}}{2\sqrt{n}} \mathbb{E}_x \left[\int_0^t [u_f^n]_{2,1, B_{\widehat{X}^n(s)}(\bar{\ell}/\sqrt{n})} | \bar{a}^n(\widehat{X}^n(s)) | ds \right] \\
 (45) \qquad &\leq \frac{\bar{\ell}}{2\sqrt{n}} \mathbb{E}_x \left[\int_0^t [u_f^n]_{2,1, B_{\widehat{X}^n(s)}(\bar{\ell}/\sqrt{n})} | \bar{a}^n(0) | ds \right] \\
 &\quad + \frac{\bar{\ell}}{2\sqrt{n}} \mathbb{E}_x \left[\int_0^t [u_f^n]_{2,1, B_{\widehat{X}^n(s)}(\bar{\ell}/\sqrt{n})} | \bar{a}^n(\widehat{X}^n(s)) - \bar{a}^n(0) | ds \right].
 \end{aligned}$$

Using (11), (12) and (40) we conclude that

$$|A_{u_f^n}^{n,x}(t)| \leq \frac{\bar{\ell}}{2\sqrt{n}} \mathbb{E}_x \left[\int_0^t C_V(\widehat{X}^n(s)) (|\bar{a}^n(0)| + K_a |\widehat{X}^n(s)|/\sqrt{n}) ds \right],$$

which completes the proof. \square

We are ready to prove Theorem 3.2.

PROOF OF THEOREM 3.2. As ν^n is a stationary distribution we have, by (22) and (26), that

$$\mathbb{E}_{\nu^n} [u_f^n(\widehat{X}^n(t))] = \mathbb{E}_{\nu^n} [u_f^n(\widehat{X}^n(0))] \leq C \nu^n(V) < \infty$$

for all sufficiently large n and all $t \geq 0$. Recalling that $\mathcal{A}^n u_f^n = -f$, Corollary 5.1 guarantees the existence of a finite positive constant ϑ (not depending on n) such that

$$\begin{aligned}
 \left| \mathbb{E}_{\nu^n} \left[\int_0^t f(\widehat{X}^n(s)) ds \right] \right| &\leq \vartheta \mathbb{E}_{\nu^n} \left[\int_0^t \frac{C_V(\widehat{X}^n(s))}{\sqrt{n}} \left(1 + \frac{|\widehat{X}^n(s)|}{\sqrt{n}} \right) ds \right] \\
 (46) \qquad &= \vartheta t \mathbb{E}_{\nu^n} \left[\frac{C_V(\widehat{X}^n(0))}{\sqrt{n}} \left(1 + \frac{|\widehat{X}^n(0)|}{\sqrt{n}} \right) \right]
 \end{aligned}$$

for all $t \geq 0$, where the interchange of integral and expectation is justified by the nonnegativity of the integrands. Using again (26) and the nonnegativity of V we

have, for all $t \geq 0$, that

$$\mathbb{E}_{v^n} \left[\int_0^t |f(\widehat{X}^n(s))| ds \right] \leq \mathbb{E}_{v^n} \left[\int_0^t V(\widehat{X}^n(s)) ds \right] = tv^n(V) < \infty.$$

This justifies replacing integral and expectation in (46) to conclude that, with $t > 0$,

$$\begin{aligned} |v^n(f)| &= \frac{1}{t} \left| \mathbb{E}_{v^n} \left[\int_0^t f(\widehat{X}^n(s)) ds \right] \right| \leq \vartheta \mathbb{E}_{v^n} \left[\frac{C_V(\widehat{X}^n(0))}{\sqrt{n}} \left(1 + \frac{|\widehat{X}^n(0)|}{\sqrt{n}} \right) \right] \\ &= \mathcal{O}(1/\sqrt{n}) \end{aligned}$$

for a (re-defined) constant ϑ as required, where the last equality follows from (26) recalling the definition of C_V in (39). \square

PROOF OF THEOREM 3.3. Let V be as in Assumption 3.1. Applying Lemma 5.1 as in the proof of Corollary 5.1 we have that

$$\begin{aligned} |A_V^{n,x}(t)| &\leq \frac{\bar{\ell}}{2\sqrt{n}} \mathbb{E}_x \left[\int_0^t [V]_{2,1,B_{\widehat{X}^n(s)}(\bar{\ell}/\sqrt{n})} |\bar{a}^n(0)| ds \right] \\ &\quad + \frac{\bar{\ell}}{2\sqrt{n}} \mathbb{E}_x \left[\int_0^t [V]_{2,1,B_{\widehat{X}^n(s)}(\bar{\ell}/\sqrt{n})} |\bar{a}^n(\widehat{X}^n(s)) - \bar{a}^n(0)| ds \right] \\ &\leq \mathbb{E}_x \left[\int_0^t \frac{\delta}{4} V(\widehat{X}^n(s)) ds \right] \end{aligned}$$

for all sufficiently large n . The last inequality follows noting that, by (11), (12) and (28), there exists a finite positive constant C such that

$$[V]_{2,1,B_{\widehat{X}^n(s)}(\bar{\ell}/\sqrt{n})} |\bar{a}^n(0)| \leq CV(\widehat{X}^n(s))$$

and

$$[V]_{2,1,B_{\widehat{X}^n(s)}(\bar{\ell}/\sqrt{n})} |\bar{a}^n(\widehat{X}^n(s)) - \bar{a}^n(0)| \leq \frac{CK_a}{\sqrt{n}} V(\widehat{X}^n(s)),$$

where K_a is as in (11). Similarly one argues, using (11) and (28), that for all sufficiently large n ,

$$\begin{aligned} |D_V^{n,x}(t)| &\leq \frac{1}{2} \mathbb{E}_x \left[\int_0^t |D^2V(\widehat{X}^n(s))| |\bar{a}^n(\widehat{X}^n(s)) - \bar{a}^n(0)| ds \right] \\ &\leq \mathbb{E}_x \left[\int_0^t \frac{\delta}{4} V(\widehat{X}^n(s)) ds \right], \end{aligned}$$

to conclude from Assumption 3.1 and Lemma 5.1 that

$$\mathbb{E}_x [V(\widehat{X}^n(t))] \leq V(x) + \mathbb{E}_x \left[\int_0^t \left(-\frac{\delta}{2} V(\widehat{X}^n(s)) + b \right) ds \right].$$

In turn, (29) holds for all sufficiently large n .

This guarantees that \widehat{X}^n is ergodic for all such n ; see, for example, [23], Theorem 8.13. Using (29) and the nonnegativity of V , we have for all sufficiently large n and all $t > 0$ that

$$(47) \quad \frac{1}{t} \mathbb{E}_x \left[\int_0^t V(\widehat{X}^n(s)) ds \right] \leq \frac{1}{t} 2\delta^{-1}(V(x) + bt).$$

Letting ν^n be the steady-state distribution of \widehat{X}^n we have, for each M , that

$$\mathbb{E}_{\nu^n} [V(\widehat{X}^n(0)) \wedge M] = \lim_{t \rightarrow \infty} \frac{1}{t} \mathbb{E}_x \left[\int_0^t V(\widehat{X}^n(s)) \wedge M ds \right] \leq 2\delta^{-1}b.$$

The result now follows from the nonnegativity of V and the monotone convergence theorem. \square

6. Two examples. Lyapunov functions that satisfy Assumption 3.1 must be identified on a case-by-case basis. For the first example—the Erlang-A queue—this is a straightforward task. For the second example—a queue with many servers and phase-type service time distribution—this task is substantially more difficult, but recent work [9] provides us with the required function.

6.1. *The Erlang-A queue.* We consider a sequence of queues with a single pool of i.i.d. servers that serve one class of impatient i.i.d. customers. Arrivals follow a Poisson process (with rate n in the n th queue), service times are exponentially distributed with rate μ and customers’ patience times are exponentially distributed with rate θ . In the n th queue, there are N^n servers in the server pool. Let $X^n(t)$ be the total number of jobs in the n th queue (waiting or in service) at time t . Then $(X^n(t), t \geq 0)$ is a birth and death process with state space \mathbb{Z}_+ , birth rate n in all states and death rate $\mu(x \wedge N^n) + \theta(x - N^n)^+$ in state x where, for the remainder of the paper, we use $(x)^+ = \max\{0, x\}$, $(x)^- = \max\{0, -x\}$. We assume that $\theta > 0$ so that positive recurrence of X^n follows easily.

The drift F^n is then specified here by

$$F^n(x) = n - \mu(x \wedge N^n) - \theta(x - N^n)^+, \quad x \in \mathbb{Z}_+,$$

and is trivially extended here to the real line by allowing x to take real values (including negative values). The FM is then given by

$$(FM) \quad \bar{x}^n(t) = \bar{x}^n(0) + \int_0^t F^n(\bar{x}^n(s)) ds.$$

There exists a unique point \bar{x}_∞^n in which $F^n(\bar{x}_\infty^n) = 0$. At this point $n = \mu(\bar{x}_\infty^n \wedge N^n) + \theta(\bar{x}_\infty^n - N^n)^+$ so that

$$\bar{a}^n(0) = \frac{1}{n} (n + \mu(\bar{x}_\infty^n \wedge N^n) + \theta(\bar{x}_\infty^n - N^n)^+) \equiv 2.$$

The DM for the Erlang-A queue is subsequently given by

$$(DM) \quad \widehat{Y}^n(t) = \widehat{Y}^n(0) + \int_0^t \widehat{F}^n(\widehat{Y}^n(s)) ds + \sqrt{2}B(t),$$

where

$$\widehat{F}^n(x) = \mu((f^n(x))^- - (f^n(0))^-) - \theta((f^n(x))^+ - (f^n(0))^+),$$

and $f^n(x) = x + (\bar{x}_\infty^n - N^n)/\sqrt{n}$. It is easily verified that there exists $\eta > 0$ such that $\widehat{F}^n(x) \leq -\eta x$ when $x > 0$ and $\widehat{F}^n(x) \geq -\eta x$ if $x < 0$. Fixing $\varrho \geq 1$ and taking

$$V_m(x) = \varrho + x^{2m}, \quad x \in \mathbb{R}, m \in \mathbb{N},$$

we have that $V_m(x) > V_m(0)$ for all $x \neq 0$ and

$$DV_m(x)\widehat{F}^n(x) \leq -\eta(2m)(V_m(x) - V_m(0)) \quad \text{for all } x \neq 0.$$

Note that V_m is trivially sub-exponential. Further, for all sufficiently large $|x|$,

$$D^2V_m(x) = 2m(2m - 1)x^{2m-2} \leq \frac{\eta}{2}x^{2m},$$

so that the conditions of Lemma 3.1 are satisfied and, in turn, UL holds for the DM. Further, for each $t \geq 0$, $X^n(t) \leq X^n(0) + N^n + A^n(t)$ where $A^n(t)$ is the number of arrivals by time t . Condition (17) then follows from basic properties of the Poisson process. We have the following consequence.

LEMMA 6.1. *Fix $\varrho \geq 1$ and positive $m \in \mathbb{N}$. Then, $V_m(x) = \varrho + x^{2m}$ satisfies Assumption 3.1 for the DM of the Erlang-A queue.*

Fixing $m \in \mathbb{N}$ and choosing sufficiently large ϱ , we can take $\bar{V}_m = V_{4m}$ in Corollary 3.4; see Remark 3.4. The following is now a direct consequence of Theorem 3.2 and Corollary 3.4.

THEOREM 6.1 (Approximation gap for the Erlang-A queue in stationarity). *Consider a sequence of Erlang-A queues as above and let f be such that $\bar{f} \leq V_m$ for some $m \in \mathbb{N}$. Then*

$$\limsup_{n \rightarrow \infty} v^n(|f|) < \infty \quad \text{and} \quad v^n(f) - \pi^n(f) = \mathcal{O}(1/\sqrt{n}).$$

REMARK 6.1 (Universality and the connection to [15]). Above, we did not impose any restrictions on the way in which the number of servers, N^n , scales with n so that one may interpret our DM as a *universal* approximation for the Erlang-A queue. Universality for this queue (and its contrast with the assumption of a so-called operational regime) are discussed at length in [15]; see also the references therein. A similar result is proved there for the Erlang-A queue using an approach that, while having important similarities to the approach we take here, is based on approximating the excursions of the process X^n above and below \bar{x}_∞^n . In this one-dimensional Markov chain, the Poisson equation we use here is (informally) a “pasting” of the Dirichlet problems studied in [15].

In their greatest generality, the results of [15] are not a special case of Theorem 6.1 above. In [15] the authors allow the service rate μ to vary with n . This is facilitated by the excursion approach taken there but violates the assumptions required to apply our results, particularly, the uniform Lipschitz continuity of \widehat{F}^n . Moreover, the approach in [15] seems to be easily extendable to the case with $\theta = 0$ in which case the DM is not exponentially ergodic and Assumption 3.1 is not satisfied.

6.2. *A phase-type queue with many servers.* We next consider the single class $M/PH/n + M$ queue. This is a generalization of the Erlang-A queue where the exponential service time is replaced by a phase-type service-time; see [8] for a detailed construction. We repeat here only the essential details.

Let I be the number of service phases, and let $1/\nu_k$ be the average length of phase $k = 1, \dots, I$. We assume that $p = (1, \dots, 0)'$, corresponding to all customers commencing their service at phase 1 (the diffusion limits in [8] cover the general case where p is an arbitrary probability vector). Having completed phase i a job transitions into phase j with probability P_{ij} . The triplet (p, ν, P) defines the phase-type service-time distribution.

Let

$$R = (I - P') \text{diag}(\nu) \quad \text{and} \quad 1/\mu = e'R^{-1}p, \quad \gamma = \mu R^{-1}p.$$

Note that $\sum_k \gamma_k = 1$. As before, the patience rate is $\theta > 0$.

We consider a sequence of such queues indexed by the arrival rate $n \in \mathbb{Z}_+$. Let

$$\gamma^n = n\gamma, \quad n \in \mathbb{N}.$$

Let $X_1^n(t)$ be the number of customers in the first phase of their service and waiting in the queue at time t . For $i > 1$, let $X_i^n(t)$ be the number of customers in phase i of service at time t . The process

$$X^n(t) = (X_1^n(t), \dots, X_I^n(t)),$$

is then a CTMC.

For simplicity of exposition we assume here that $\sum_k \gamma_k^n$ is integer valued for each n and that the number of servers N^n satisfies $N^n = \sum_k \gamma_k^n$. This implies, trivially, that $N^n = \sum_k \gamma_k^n + \mathcal{O}(\sqrt{n})$ which corresponds to the so-called Halfin–Whitt many-server regime and allows us subsequently to build on the results of [8] and [9] that study diffusion limits in this regime. The analysis below is easily extended to the case $N^n = \sum_k \gamma_k^n + \beta\sqrt{n} + o(\sqrt{n})$ for some $\beta \neq 0$.

Define

$$\bar{x}_\infty^n = (\gamma_1^n, \dots, \gamma_I^n),$$

and the scaled and centered process \widehat{X}^n as in (7). Then,

$$(48) \quad \widehat{F}_i^n(x) = \begin{cases} -v_i x_i + \sum_{k \neq i, k \neq 1} P_{ki} \nu_k x_k + v_1 P_{1i} (x_1 - (e'x)^+), & \text{if } i \neq 1, \\ -v_1 (x_1 - (e'x)^+) - \theta (e'x)^+, & \text{if } i = 1. \end{cases}$$

This is written, in Matrix notation, as

$$(49) \quad \widehat{F}^n(x) = -Rx + (R - \theta I)p(e'x)^+.$$

$$(50) \quad n\bar{a}_{kk}^n(x) = \begin{cases} \sum_{i \neq k, i \neq 1} P_{ik}v_k(\gamma_k^n + \sqrt{n}x_k) + v_k(\gamma_k^n + \sqrt{n}x_k) \\ \quad + v_1P_{1k}(\gamma_1^n + \sqrt{n}x_1 - \sqrt{n}(e'x)^+), & \text{if } k \neq 1, \\ n + v_1(\gamma_1^n + \sqrt{n}x_1) + \theta\sqrt{n}(e'x)^+, & \text{if } k = 1, \end{cases}$$

and, for $k \neq j$,

$$(51) \quad n\bar{a}_{kj}^n(x) = \begin{cases} P_{kj}v_k(\gamma_k^n + \sqrt{n}x_k) + P_{jk}v_j(\gamma_j^n + \sqrt{n}x_j), & \text{if } k \neq 1, \\ P_{kj}v_k(\gamma_k^n + \sqrt{n}x_k - \sqrt{n}(e'x)^+) \\ \quad + P_{jk}v_j(\gamma_j^n + \sqrt{n}x_j), & \text{if } k = 1. \end{cases}$$

The functions \widehat{F}^n and \bar{a}^n satisfy (10) and (11). Assumption 2.1 holds in this example as the chain is trivially nonexplosive and irreducible. The positive recurrence follows immediately from the fact that $\theta > 0$.

The diffusion model is given by

$$(DM) \quad \widehat{Y}^n(t) = y + \int_0^t \widehat{F}^n(\widehat{Y}^n(s)) ds + \sqrt{\bar{a}^n(0)}B(t),$$

with \widehat{F}^n as in (49) and diffusion coefficient \bar{a}^n as in (50)–(51). Note (49)–(51) that \widehat{F}^n and $\bar{a}^n(0)$ do not, in fact, depend here on n . The existence of a quadratic Lyapunov function, V , for \widehat{Y}^n then follows from [9], Theorem 3—this function is specified in equation (5.24) there. (To extend this argument to the general case with $N^n = \sum_k \gamma_k^n + \beta\sqrt{n} + o(\sqrt{n})$, note that V in [9] is still a Lyapunov function for each n if we perturb \widehat{F}^n by a constant and $\bar{a}^n(0)$ by a term that shrinks proportional to $1/\sqrt{n}$.)

With a careful choice of the smoothing function ϕ there, the function $\Psi = \varrho + V$ (for any constant $\varrho \geq 1$) is also sub-exponential. Finally, (17) is argued as in the Erlang-A case using crude bounds on the Poisson arrivals.

The function $\Psi = \varrho + V$ thus satisfies Assumption 3.1. It is easily verified that $\Psi \in \mathcal{C}^3(\mathbb{R}^d)$ and satisfies (30) so that, as in Remark 3.4, $\Psi_m(x) = (\Psi(x))^m$ satisfies Assumption 3.1 with re-defined constants δ, b and K . Choosing sufficiently large ϱ guarantees that $\Psi_{4m}(\cdot) \geq \Psi_m(\cdot)(1 + |\cdot|)^4$. The following is then an immediate consequence of Theorem 3.2 and Corollary 3.4.

COROLLARY 6.2. *Consider the sequence of phase-type queues as above, and let f be such that $\bar{f} \leq \Psi_m$ for some $m \in \mathbb{N}$. Then*

$$\limsup_{n \rightarrow \infty} v^n(|f|) < \infty \quad \text{and} \quad v^n(f) - \pi^n(f) = \mathcal{O}(1/\sqrt{n}).$$

Thus, as in Remark 3.5, we have a Lyapunov function that allows us to establish simultaneously the stability of the Markov chain for each sufficiently large n , the uniform integrability of moments and the approximation gap. It is worth noting that the fact that $\limsup_{n \rightarrow \infty} v^n(|f|) < \infty$ was already established, by alternative means and for more general (multiclass) phase-type queues, in [7].

7. Proof of Theorem 3.1. The main step in this proof is a uniform minorization condition for a time-discretized version of \widehat{Y}^n . Once this is established (see Lemma 7.1 below), we build on [2] to complete the argument. The proofs of the lemmas that are stated in this section appear in the Appendix.

We first consider a linear transformation of \widehat{Y}^n . Specifically, let L_n be the unique square root of the matrix $\bar{a}^n(0)$; see [16], Theorem 7.2.6. In particular, $L_n(L_n)^T = \bar{a}^n(0)$. The matrix L_n is itself invertible and its inverse is the square root of the inverse of $\bar{a}^n(0)$; see [16], page 406. Let

$$(52) \quad \widehat{F}_L^n(x) = L_n^{-1} \widehat{F}^n(L_n x), \quad x \in \mathbb{R}^d,$$

and define

$$Z_L^n(t) = L_n^{-1} \widehat{Y}^n(t), \quad t \geq 0.$$

Then Z_L^n is a d -dimensional Brownian motion with drift \widehat{F}^n , that is,

$$Z_L^n(t) = z + \int_0^t \widehat{F}_L^n(Z_L^n(s)) ds + B(t),$$

where $z = L_n^{-1} \widehat{Y}^n(0)$.

We next consider the discrete-time analogues of both Z_L^n and \widehat{Y}^n . Let

$$\Phi_l^n = Z_L^n(l) \quad \text{and} \quad \psi_l^n = \widehat{Y}^n(l) \quad \text{for } l \in \mathbb{Z}_+.$$

Let $\mathbb{P}_{\Phi^n}(\cdot, \cdot)$ and $\mathbb{P}_{\psi^n}(\cdot, \cdot)$ be the corresponding one-step transition functions. Below $\mathcal{B}(\mathbb{R}^d)$ is the family of Borel sets in \mathbb{R}^d .

LEMMA 7.1. *Fixing $K > 0$, there exist a probability measure \mathcal{Q} with $\mathcal{Q}(B_0(K)) = 1$ and a constant $\epsilon < 1$ (both not depending on n) such that*

$$\mathbb{P}_{\Phi^n}(x, \mathcal{E}) \geq \epsilon \mathcal{Q}(L_n \mathcal{E}), \quad x \in L_n^{-1} B_0(K), \mathcal{E} \in \mathcal{B}(\mathbb{R}^d).$$

There consequently exists a constant $\tilde{\epsilon} < 1$ (not depending on n) such that

$$\mathbb{P}_{\psi^n}(x, \mathcal{E}) \geq \tilde{\epsilon} \mathcal{Q}(\mathcal{E}), \quad x \in B_0(K), \mathcal{E} \in \mathcal{B}(\mathbb{R}^d).$$

The following translates the Lyapunov property UL into the discrete time setting.

LEMMA 7.2. *Let V be as in Assumption 3.1. Then there exist finite positive constants $\gamma < 1$ and \bar{b} (not depending on n) such that for all $n \in \mathbb{N}$ and all $x \in \mathbb{R}^d$,*

$$(53) \quad \mathbb{E}_x[V(\widehat{Y}^n(1))] \leq (1 - \gamma)V(x) + \bar{b} \mathbb{1}_{\bar{B}_0(K)}(x).$$

Using the fact that $V(x) \rightarrow \infty$ as $|x| \rightarrow \infty$, (53) implies that there exist finite positive constants $K, \lambda < 1$ and M such that

$$(54) \quad \mathbb{E}_x[V(\widehat{Y}^n(1))] \leq \begin{cases} \lambda V(x), & \text{if } x \notin \overline{B}_0(K), \\ M, & \text{if } x \in \overline{B}_0(K). \end{cases}$$

The following is then a direct consequence of [2], Theorem 1.1. Assumptions (A1)–(A3) there hold by Lemmas 7.1, 7.2 and by (54).

COROLLARY 7.1. *There exist constants \mathbb{M} and μ (not depending on n) such that for each $m \in \mathbb{N}$,*

$$\sup_n \sup_{x \in \mathbb{R}^d} \sup_{|f| \leq V} \frac{1}{V(x)} |\mathbb{E}_x[f(\widehat{Y}^n(m))] - \pi^n(f)| \leq \mathbb{M}e^{-\mu m}.$$

With these we are ready for the proof of Theorem 3.1.

PROOF OF THEOREM 3.1. The proof of the theorem now follows as in [20], page 536. Specifically, let $s = t - \lfloor t \rfloor$

$$\begin{aligned} \sup_{|f| \leq V} |\mathbb{E}_x[f(\widehat{Y}^n(t))] - \pi^n(f)| &= \sup_{|f| \leq V} |\mathbb{E}_x[f(\widehat{Y}^n(\lfloor t \rfloor + s))] - \pi^n(f)| \\ &= \sup_{|f| \leq V} |\mathbb{P}_{\widehat{Y}^n}^s(x, dy)(\mathbb{E}_x[f(\widehat{Y}^n(\lfloor t \rfloor))] - \pi^n(f))| \\ &\leq \int_y \mathbb{P}_{\widehat{Y}^n}^s(x, dy) \sup_{|f| \leq V} |\mathbb{E}_y[f(\widehat{Y}^n(\lfloor t \rfloor))] - \pi^n(f)| \\ &\leq \mathbb{M}e^{-\mu \lfloor t \rfloor} \mathbb{E}_x[V(\widehat{Y}^n(s))] \\ &\leq \mathbb{M}e^\mu e^{-\mu t} (V(x) + b), \end{aligned}$$

where $\mathbb{P}_{\widehat{Y}^n}^s(x, A)$ is the transition probability function of \widehat{Y}^n in s time units. In the last inequality we used (19) and the fact that $s = t - \lfloor t \rfloor \leq 1$. Finally, since $V \geq 1$, the theorem holds with the constants $\mathcal{M} = \mathbb{M}e^{+1}(1 + b)$ and μ . \square

8. Concluding remarks. Diffusion models are useful in the approximation of Markov chains. We proved that, under a uniform Lyapunov condition, the steady-state of some multidimensional CTMCs can be approximated with impressive accuracy by the steady-state of a relatively tractable diffusion model.

The existence of a diffusion limit that satisfies the Lyapunov requirement—as is the case for the phase-type queue considered in Section 6.2—can facilitate the application of our results. The distinction between the diffusion model and diffusion limit is, however, important. A central motivation behind this work is to bypass the need for diffusion limits with the objective of providing steady-state diffusion

approximation whose precision does not depend on assumption with regards to limiting values of underlying parameters. That is, we ultimately seek to provide “limit-free” (or *universal*) approximations.

A uniform Lyapunov condition, as we require in Assumption 3.1, need not hold in general. Informally, one expects such a condition to hold if the scale parameter n has limited effect on the drift of the process around the FMs stationary point. Many-server queues with abandonment, as those we use to illustrate our results, seem to satisfy this characterizations: diffusion limits (regardless of the parameter regime, determining how the number of servers N^n scales with n) are generalizations of the OU process. It remains to identify the broadest characterization of Markov chains for which a uniform Lyapunov condition can be expected to hold.

In addition, the following extensions seem important:

State-space collapse. A fundamental phenomenon in diffusion limits for multi-class queueing systems is that of state-space collapse (SSC). With SSC, the diffusion limit “lives” on a state-space that is of lower dimension relative to the original CTMC: some coordinates of the CTMC become, asymptotically, deterministic functions of others. For example, if one allows for arbitrary initial-phase vectors p in the example of Section 6.2, the number of customers in queue with initial phase k is asymptotically equal to p_k ; see [8]. To exploit state-space collapse within the diffusion-model framework used in this paper, one must develop bounds (rather than convergence results) for state-space collapse.

Single server queues and reflection. A key challenge with single-server queueing systems is that of reflection. Such reflection may violate our assumptions on \widehat{F}^n . Consider, for example, the $M/M/1 + M$ queue—this is a single-server version of the Erlang-A queue discussed in Section 6. Suppose that the arrival rate and service rate in the n th queue satisfy $\lambda^n = n\lambda$, and $\mu^n = \lambda^n - \beta\sqrt{n}$ (for $\beta > 0$). Let $\theta > 0$ be the patience parameter. Then

$$\begin{aligned} F^n(x) &= \lambda^n - \mu^n \mathbb{1}\{x > 0\} - \theta x \\ &= \beta\sqrt{n} - \theta x + \mu^n \mathbb{1}\{x = 0\}, \end{aligned}$$

so that $\bar{x}_\infty^n = \beta\sqrt{n}/\theta$. Also, $\widehat{F}^n(-\beta/\theta) = F^n(0)/\sqrt{n} = \beta + \mu^n/\sqrt{n}$ and, in particular $|\widehat{F}^n(-\beta/\theta) - \widehat{F}^n(0)| = \beta + \mu^n/\sqrt{n} = \sqrt{n}\lambda \rightarrow \infty$ as $n \rightarrow \infty$. Clearly, (10) is violated.

It is fair to conjecture that similar results as ours can be proved in such settings provided that the reflection is explicitly captured in the DM. Extending our results to DMs with reflection seems to present a challenge insofar as the theory of PDEs that arise from the Poisson equation for such networks is less developed and poses a challenge in terms of the gradient bounds that are central to our analysis here; see, for example, [5], where the Poisson equation for constrained diffusion is discussed as well as, in the context of ergodic control, [3].

APPENDIX

PROOF OF LEMMA 5.1. Fix $x \in \widehat{E}^n$. By Itô's rule applied to the pure jump process $(\Psi(\widehat{X}^n(t)), t \geq 0)$ we have that

$$(55) \quad \begin{aligned} \Psi(\widehat{X}^n(t)) = & \Psi(x) + \sum_{s \leq t} \sum_{i=1}^d \Psi_i(\widehat{X}^n(s-)) \Delta \widehat{X}_i^n(s) \\ & + \sum_{s \leq t} \left[\Psi(\widehat{X}^n(s)) - \Psi(\widehat{X}^n(s-)) - \sum_{i=1}^d \Psi_i(\widehat{X}^n(s-)) \Delta \widehat{X}_i^n(s) \right]. \end{aligned}$$

From the linear growth of \widehat{F}^n and from (42), it then follows that

$$\mathbb{E}_x \left[\int_0^t |D\Psi(\widehat{X}^n(s))| |\widehat{F}^n(\widehat{X}^n(s))| ds \right] < \infty.$$

We can then apply Lévy's formula for CTMCs (see, e.g., [4], Exercise I.2.E2) to get that

$$\sum_{s \leq t} \sum_{i=1}^d \Psi_i(\widehat{X}^n(s-)) \Delta \widehat{X}_i^n(s) - \sum_{i=1}^d \int_0^t \Psi_i(\widehat{X}^n(s)) \widehat{F}_i^n(\widehat{X}^n(s)) ds$$

is a martingale with respect to the filtration in (4) and, in turn, for all $t \geq 0$,

$$\mathbb{E}_x \left[\sum_{s \leq t} \sum_{i=1}^d \Psi_i(\widehat{X}^n(s-)) \Delta \widehat{X}_i^n(s) \right] = \mathbb{E}_x \left[\sum_{i=1}^d \int_0^t \Psi_i(\widehat{X}^n(s)) \widehat{F}_i^n(\widehat{X}^n(s)) ds \right].$$

To treat the second line of (56), we decompose it into

$$(D) \quad \frac{1}{2} \sum_{s \leq t} \sum_{i,j}^d \Psi_{ij}(\widehat{X}^n(s-)) \Delta \widehat{X}_i^n(s) \Delta \widehat{X}_j^n(s)$$

and

$$(A) \quad \begin{aligned} \sum_{s \leq t} \left[\Psi(\widehat{X}^n(s)) - \Psi(\widehat{X}^n(s-)) - \sum_{i=1}^d \Psi_i(\widehat{X}^n(s-)) \Delta \widehat{X}_i^n(s) \right. \\ \left. - \frac{1}{2} \sum_{i,j}^d \Psi_{ij}(\widehat{X}^n(s-)) \Delta \widehat{X}_i^n(s) \Delta \widehat{X}_j^n(s) \right]. \end{aligned}$$

We treat (D) first. By (11), $|\bar{a}^n(x)| \leq |\bar{a}^n(0)| + K_a|x|/\sqrt{n}$ so that, by (42),

$$\mathbb{E}_x \left[\int_0^t |D^2\Psi(\widehat{X}^n(s))| |\bar{a}^n(\widehat{X}^n(s))| ds \right] < \infty, \quad t \geq 0, x \in \widehat{E}^n,$$

and applying Lévy’s formula once again, we obtain

$$\begin{aligned}
 & \frac{1}{2} \mathbb{E}_x \left[\sum_{s \leq t} \sum_{i,j}^d \Psi_{ij}(\widehat{X}^n(s-)) \Delta \widehat{X}_i^n(s) \Delta \widehat{X}_j^n(s) \right] \\
 &= \frac{1}{2} \mathbb{E}_x \left[\sum_{i,j}^d \sum_{\ell} \int_0^t \Psi_{ij}(\widehat{X}^n(s)) \ell_i \ell_j \frac{1}{n} \beta_{\ell}^n (\sqrt{n} \widehat{X}^n(s) + \bar{x}_{\infty}^n) ds \right] \\
 &= \frac{1}{2} \mathbb{E}_x \left[\sum_{i,j}^d \int_0^t \Psi_{ij}(\widehat{X}^n(s)) \bar{a}_{ij}^n(\widehat{X}^n(s)) ds \right] \\
 &= \frac{1}{2} \mathbb{E}_x \left[\sum_{i,j}^d \int_0^t \Psi_{ij}(\widehat{X}^n(s)) \bar{a}_{ij}^n(0) ds \right] \\
 &\quad + \frac{1}{2} \mathbb{E}_x \left[\sum_{i,j}^d \int_0^t \Psi_{ij}(\widehat{X}^n(s)) (\bar{a}^n(\widehat{X}^n(s)) - \bar{a}_{ij}^n(0)) ds \right].
 \end{aligned}$$

The second item in the last line is $D_{\Psi}^{n,x}(t)$ in the statement of the lemma. We have proven thus far that

$$\begin{aligned}
 & \mathbb{E}_x[\Psi(\widehat{X}^n(t))] \\
 &= \Psi(x) + \mathbb{E}_x \left[\sum_{i=1}^d \int_0^t \Psi_i(\widehat{X}^n(s)) \widehat{F}_i^n(\widehat{X}^n(s)) ds \right] \\
 &\quad + \frac{1}{2} \mathbb{E}_x \left[\sum_{i,j}^d \int_0^t \Psi_{ij}(\widehat{X}^n(s)) \bar{a}_{ij}^n(0) ds \right] + D_{\Psi}^{n,x}(t) + A_{\Psi}^{n,x}(t) \\
 &= \Psi(x) + \mathbb{E}_x \left[\int_0^t \mathcal{A}^n \Psi(\widehat{X}^n(s)) ds \right] + D_{\Psi}^{n,x}(t) + A_{\Psi}^{n,x}(t),
 \end{aligned}$$

where $D_{\Psi}^{n,x}$ is as in the statement of the lemma and $A_{\Psi}^{n,x}(t) = \mathbb{E}_x[\mathbf{A}]$ (we will prove below that this expectation is well defined). To bound $A_{\Psi}^{n,x}$ note that, by Taylor’s theorem,

$$\begin{aligned}
 & \Psi(\widehat{X}^n(s)) - \Psi(\widehat{X}^n(s-)) \\
 &= \sum_{i=1}^d \Psi_i(\widehat{X}^n(s-)) \Delta \widehat{X}_i^n(s) \\
 &\quad + \frac{1}{2} \sum_{i,j}^d \Psi_{ij}(\widehat{X}^n(s-) + \eta_{\widehat{X}^n(s-), \widehat{X}^n(s)}) \Delta \widehat{X}_i^n(s) \Delta \widehat{X}_j^n(s),
 \end{aligned}$$

where $\eta_{\widehat{X}^n(s-), \widehat{X}^n(s)} \in \prod_{i=1}^d [\widehat{X}_i^n(s-), \widehat{X}_i^n(s)]$. Thus

$$\mathbf{A} = \frac{1}{2} \sum_{s \leq t} \left(\sum_{i,j}^d (\Psi_{ij}(\widehat{X}^n(s-) + \eta_{\widehat{X}^n(s-)})) - \Psi_{ij}(\widehat{X}^n(s-)) \right) \Delta \widehat{X}_i^n(s) \Delta \widehat{X}_j^n(s).$$

Here note that $|\Delta \widehat{X}_i^n(s)| |\Delta \widehat{X}_j^n(s)| \leq \bar{\ell}^2/n$. Let $\tilde{\Psi}_{ij}(x, y) = \Psi_{ij}(x + \eta_{x,y}) - \Psi_{ij}(x)$. Note that $|\tilde{\Psi}_{ij}(x, y)| \leq \frac{\bar{\ell}}{\sqrt{n}} [\Psi]_{2,1, B_x(\bar{\ell}/\sqrt{n})}$ for $x, y \in \widehat{E}^n$ with $y \in B_x(\bar{\ell}/\sqrt{n})$. Since $\sum_{\ell} |\ell_i| |\ell_j| \beta_{\ell}^n(x) \leq \sum_{\ell} (|\ell_i|^2 + |\ell_j|^2) \beta_{\ell}^n(x) \leq |a^n(x)|$, we have that

$$\begin{aligned} & \frac{1}{2n} \mathbb{E}_x \left[\sum_{i,j}^d \int_0^t \sum_{\ell} |\tilde{\Psi}_{ij}(\widehat{X}^n(s), \widehat{X}^n(s) + \ell/\sqrt{n})| |\ell_i| |\ell_j| \beta_{\ell}^n(X^n(s)) ds \right] \\ & \leq \frac{\bar{\ell}}{\sqrt{n}} \frac{1}{2n} \mathbb{E}_x \left[\sum_{i,j}^d \int_0^t [\Psi]_{2,1, B_{\widehat{X}^n(s)}(\bar{\ell}/\sqrt{n})} \sum_{\ell} |\ell_i| |\ell_j| \beta_{\ell}^n(X^n(s)) ds \right] \\ & \leq \frac{\bar{\ell}}{\sqrt{n}} \frac{1}{2n} \mathbb{E}_x \left[\int_0^t [\Psi]_{2,1, B_{\widehat{X}^n(s)}(\bar{\ell}/\sqrt{n})} |a^n(X^n(s))| ds \right] \\ & = \frac{\bar{\ell}}{2\sqrt{n}} \mathbb{E}_x \left[\int_0^t [\Psi]_{2,1, B_{\widehat{X}^n(s)}(\bar{\ell}/\sqrt{n})} |\bar{a}^n(\widehat{X}^n(s))| ds \right] < \infty, \end{aligned}$$

where the finiteness follows from (11) and condition (42).

We can apply Lévy’s formula one final time to conclude that

$$\begin{aligned} |\mathbb{E}_x[\mathbf{A}]| & = \left| \frac{1}{2n} \sum_{i,j}^d \mathbb{E}_x \left[\int_0^t \sum_{\ell} \tilde{\Psi}_{ij}(\widehat{X}^n(s), \widehat{X}^n(s) + \ell/\sqrt{n}) \ell_i \ell_j \beta_{\ell}^n(X^n(s)) ds \right] \right| \\ & \leq \frac{\bar{\ell}}{2\sqrt{n}} \mathbb{E}_x \left[\int_0^t [\Psi]_{2,1, B_{\widehat{X}^n(s)}(\bar{\ell}/\sqrt{n})} |\bar{a}^n(\widehat{X}^n(s))| ds \right] \end{aligned}$$

as required. \square

Toward the proof of Lemma 7.1 we first prove that $\widehat{F}_L^n(x) = L_n^{-1} \widehat{F}^n(L_n x)$ inherits the Lipschitz continuity of \widehat{F}^n .

LEMMA A.1. *There exists a finite positive constant K (not depending on n) such that*

$$|\widehat{F}_L^n(x) - \widehat{F}_L^n(y)| \leq K|x - y|, \quad x, y \in \mathbb{R}^d.$$

PROOF. Since, for each n , $\bar{a}^n(0)$ is symmetric positive definite as is \bar{a} , these matrices have strictly positive eigenvalues; see, for example, [16], Theorem 7.2.1.

Also, the eigenvalues of the square-root matrix L_n are the square roots of the eigenvalues of $\bar{a}^n(0)$. Since $\bar{a}^n(0) \rightarrow \bar{a}$, the eigenvalues of L_n , $(\lambda_1^n, \dots, \lambda_d^n)$, converge to those of L , $(\lambda_1, \dots, \lambda_d)$. The eigenvalues of the inverses L_n^{-1} and L^{-1} are given by the reciprocals and, in turn, satisfy $(1/\lambda_1^n, \dots, 1/\lambda_d^n) \rightarrow (1/\lambda_1, \dots, 1/\lambda_d)$. In particular $\|L_n\|_2 \rightarrow \|L\|_2$ and $\|L_n^{-1}\|_2 \rightarrow \|L^{-1}\|_2$ (where, following common notation, $\|A\|_2$ is the spectral norm of A ; see [16], Section 5.1. Since the matrices are symmetric this norm is equal to the spectral radius of the matrix, that is, to its maximal eigenvalue). By definition of the matrix norm it then holds that

$$(56) \quad |L_n x - L_n y| \leq \|L_n\|_2 |x - y| \leq C_1 \|L\|_2 |x - y|, \quad x, y \in \mathbb{R}^d$$

for some finite positive constant C_1 where the last inequality follows from the fact $\|L_n\|_2 \rightarrow \|L\|_2$ argued above. Similarly,

$$(57) \quad |L_n^{-1} x - L_n^{-1} y| \leq C_2 \|L^{-1}\|_2 |x - y|, \quad x, y \in \mathbb{R}^d$$

for a finite positive constant C_2 . Finally, using (10) we have that

$$\begin{aligned} |L_n^{-1} \widehat{F}^n(L_n x) - L_n^{-1} \widehat{F}^n(L_n y)| &\leq \|L_n^{-1}\| \|\widehat{F}^n(L_n x) - \widehat{F}^n(L_n y)\| \\ &\leq C_2 K_F C_1 \|L^{-1}\|_2 \|L\|_2 |x - y|, \end{aligned}$$

which completes the proof. \square

PROOF OF LEMMA 7.1. We consider first the chain Φ^n . Fix K and let $\mathcal{K} = B_0(K)$. Let $\bar{\mathcal{K}}^n = L_n^{-1} \mathcal{K}$. By (57), there exists a constant \tilde{K} not depending on n such that

$$(58) \quad |x - y| \leq \tilde{K}, \quad x, y \in \bar{\mathcal{K}}^n.$$

By Lemma A.1 there exist ϵ and δ not depending on n such that $|\widehat{F}_L^n(x)| \leq \epsilon + \delta|x - y|$ for all $x \in \mathbb{R}^d$ and $y \in \bar{\mathcal{K}}^n$. Also, since $\widehat{F}_L^n(0) = L_n^{-1} \widehat{F}^n(L_n 0) = 0$ it satisfies also a linear growth condition uniformly in n . Using [22], Theorem 3.1 and (58) we have that

$$p(x, 1, y) \geq \check{\epsilon}, \quad x, y \in \bar{\mathcal{K}}^n$$

for some $\check{\epsilon} > 0$ where $p(x, t, y)$ is the transition density of Z_L^n from x to y in time t . In particular,

$$\mathbb{P}_{\Phi^n}(x, \mathcal{E}) \geq \int_{y \in \mathcal{E} \cap \bar{\mathcal{K}}^n} p(x, 1, y) dy \geq \check{\epsilon} \lambda(\bar{\mathcal{K}}^n) \mathcal{Q}^n(\mathcal{E}),$$

where λ is here the Lebesgue measure and

$$\mathcal{Q}^n(\cdot) = \frac{\lambda(\cdot \cap \bar{\mathcal{K}}^n)}{\lambda(\bar{\mathcal{K}}^n)}.$$

Using the invariance of Lebesgue measure under invertible linear transformations we have for any $\mathcal{E} \in \mathcal{B}(\mathbb{R}^d)$ that

$$\mathcal{Q}^n(L_n^{-1}\mathcal{E}) = \frac{\lambda(L_n^{-1}\mathcal{E} \cap L_n^{-1}\mathcal{K})}{\lambda(L_n^{-1}\mathcal{K})} = \frac{\det(L_n^{-1})\lambda(\mathcal{E} \cap \mathcal{K})}{\det(L_n^{-1})\lambda(\mathcal{K})},$$

where $\det(L_n^{-1}) > 0$ is here the determinant of the positive definite matrix L_n^{-1} , and we use the simple fact that $(L_n^{-1}\mathcal{E}) \cap (L_n^{-1}\mathcal{K}) = L_n^{-1}(\mathcal{E} \cap \mathcal{K})$. Since $L_n \rightarrow L$, it also holds that $\det(L_n^{-1}) = (\det(L_n))^{-1} \rightarrow (\det(L))^{-1} = \det(L^{-1}) > 0$ so that there exists $\varepsilon > 0$ (not depending on n) such that

$$\lambda(\bar{\mathcal{K}}^n) = \det(L_n^{-1})\lambda(\mathcal{K}) \geq \varepsilon.$$

Let $\epsilon = \tilde{\epsilon}\varepsilon$. Defining the measure

$$\mathcal{Q}(\cdot) = \frac{\lambda(\cdot \cap \mathcal{K})}{\lambda(\mathcal{K})},$$

we conclude that

$$\mathbb{P}_{\Phi^n}(x, \mathcal{E}) \geq \tilde{\epsilon}\lambda(\bar{\mathcal{K}}^n)\mathcal{Q}^n(\mathcal{E}) = \epsilon\mathcal{Q}(L_n\mathcal{E}), \quad x \in \bar{\mathcal{K}}^n, \mathcal{E} \in \mathcal{B}(\mathbb{R}^d).$$

The result for \mathbb{P}_{ψ^n} follows immediately from the above. Indeed,

$$\mathbb{P}_{\psi^n}(x, \mathcal{E}) = \mathbb{P}_{\Phi^n}(L_n^{-1}x, L_n^{-1}\mathcal{E}) \geq \epsilon\mathcal{Q}(\mathcal{E}), \quad x \in \mathcal{K}, \mathcal{E} \in \mathcal{B}(\mathbb{R}^d),$$

which completes the proof. \square

PROOF OF LEMMA 7.2. This argument is almost identical to the proof in [11], page 27. Under condition (15), Dynkin’s formula holds up to t , that is,

$$\mathbb{E}_y[V(\hat{Y}^n(t))] = V(y) + \mathbb{E}_y\left[\int_0^t \mathcal{A}^n V(\hat{Y}^n(s)) ds\right];$$

see, for example, [19], Theorem 6.3. Setting

$$g(t) = \mathbb{E}_y[V(\hat{Y}^n(t))] \quad \text{and} \quad h(t) = \mathbb{E}_y[\mathcal{A}^n V(\hat{Y}^n(t))] + \delta g(t),$$

we have that $h(t) \leq b\mathbb{1}_{\bar{B}_0(K)}(y)$ (b and δ as in Assumption 3.1) and

$$\dot{g}(t) = -\delta g(t) + h(t).$$

Solving this differential equation we get

$$\begin{aligned} g(t) &= g(0)e^{-\delta t} + \int_0^t e^{\delta(t-s)}h(s) ds \leq g(0)e^{-\delta t} + b\mathbb{1}_{\bar{B}_0(K)}(y)\frac{1 - e^{-\delta}}{\delta} \\ &= V(y)e^{-\delta t} + b\mathbb{1}_{\bar{B}_0(K)}(y)\frac{1 - e^{-\delta}}{\delta}. \end{aligned}$$

Setting $\gamma = 1 - e^{-\delta}$ and $\bar{b} = b\frac{1 - e^{-\delta}}{\delta}$ we have the statement of the lemma. \square

Acknowledgments. The author is grateful to Junfei Huang and to an anonymous referee for their careful reading of this paper and for their numerous insightful comments.

REFERENCES

- [1] ATA, B. and GURVICH, I. (2012). On optimality gaps in the Halfin–Whitt regime. *Ann. Appl. Probab.* **22** 407–455. [MR2932551](#)
- [2] BAXENDALE, P. H. (2005). Renewal theory and computable convergence rates for geometrically ergodic Markov chains. *Ann. Appl. Probab.* **15** 700–738. [MR2114987](#)
- [3] BORKAR, V. and BUDHIRAJA, A. (2004/05). Ergodic control for constrained diffusions: Characterization using HJB equations. *SIAM J. Control Optim.* **43** 1467–1492. [MR2124281](#)
- [4] BRÉMAUD, P. (1981). *Point Processes and Queues: Martingale Dynamics*. Springer, New York. [MR0636252](#)
- [5] BUDHIRAJA, A. and LEE, C. (2007). Long time asymptotics for constrained diffusions in polyhedral domains. *Stochastic Process. Appl.* **117** 1014–1036. [MR2340877](#)
- [6] BUDHIRAJA, A. and LEE, C. (2009). Stationary distribution convergence for generalized Jackson networks in heavy traffic. *Math. Oper. Res.* **34** 45–56. [MR2542988](#)
- [7] DAI, J., DIEKER, A. and GAO, X. (2012). Validity of heavy-traffic steady-state approximations in many-server queues with abandonment. Unpublished manuscript.
- [8] DAI, J. G., HE, S. and TEZCAN, T. (2010). Many-server diffusion limits for $G/Ph/n + GI$ queues. *Ann. Appl. Probab.* **20** 1854–1890. [MR2724423](#)
- [9] DIEKER, A. B. and GAO, X. (2013). Positive recurrence of piecewise Ornstein–Uhlenbeck processes and common quadratic Lyapunov functions. *Ann. Appl. Probab.* **23** 1291–1317. [MR3098433](#)
- [10] ETHIER, S. N. and KURTZ, T. G. (1986). *Markov Processes: Characterization and Convergence*. Wiley, New York. [MR0838085](#)
- [11] GALTCHOUK, L. and PERGAMENHTCHIKOV, S. (2012). Geometric ergodicity for families of homogeneous Markov chains. Available at [arXiv:1002.2341](#).
- [12] GAMARNIK, D. and ZEEVI, A. (2006). Validity of heavy traffic steady-state approximation in generalized Jackson networks. *Ann. Appl. Probab.* **16** 56–90. [MR2209336](#)
- [13] GILBARG, D. and TRUDINGER, N. S. (2001). *Elliptic Partial Differential Equations of Second Order*. Springer, Berlin. [MR1814364](#)
- [14] GLYNN, P. W. and ZEEVI, A. (2008). Bounding stationary expectations of Markov processes. In *Markov Processes and Related Topics: A Festschrift for Thomas G. Kurtz*. *Inst. Math. Stat. Collect.* **4** 195–214. IMS, Beachwood, OH. [MR2574232](#)
- [15] GURVICH, I., HUANG, J. and MANDELBAUM, A. (2014). Excursion-based universal approximations for the Erlang-A queue in steady-state. *Math. Oper. Res.* **39** 325–373.
- [16] HORN, R. A. and JOHNSON, C. R. (1994). *Matrix Analysis*. Cambridge Univ. Press, Cambridge.
- [17] KARATZAS, I. and SHREVE, S. E. (1991). *Brownian Motion and Stochastic Calculus*, 2nd ed. *Graduate Texts in Mathematics* **113**. Springer, New York. [MR1121940](#)
- [18] KHASHMINSKII, R. (2012). *Stochastic Stability of Differential Equations*, 2nd ed. *Stochastic Modelling and Applied Probability* **66**. Springer, Heidelberg. [MR2894052](#)
- [19] KLEBANER, F. C. (2005). *Introduction to Stochastic Calculus with Applications*, 2nd ed. Imperial College Press, London. [MR2160228](#)
- [20] MEYN, S. P. and TWEEDIE, R. L. (1993). Stability of Markovian processes. III. Foster–Lyapunov criteria for continuous-time processes. *Adv. in Appl. Probab.* **25** 518–548. [MR1234295](#)

- [21] PARDOUX, E. and VERETENNIKOV, A. Y. (2001). On the Poisson equation and diffusion approximation. I. *Ann. Probab.* **29** 1061–1085. [MR1872736](#)
- [22] QIAN, Z. and ZHENG, W. (2004). A representation formula for transition probability densities of diffusions and applications. *Stochastic Process. Appl.* **111** 57–76. [MR2049569](#)
- [23] ROBERT, P. (2003). *Stochastic Networks and Queues*, French ed. *Applications of Mathematics (New York)* **52**. Springer, Berlin. [MR1996883](#)

KELLOGG SCHOOL OF MANAGEMENT
NORTHWESTERN UNIVERSITY
EVANSTON, ILLINOIS 60208
USA
E-MAIL: i-gurvich@kellogg.northwestern.edu