A Perturbation Theory for Ergodic Properties of Markov Chains

T. Shardlow[1] [2] and A. M. Stuart[1]

Abstract

Perturbations to Markov chains and Markov processes are considered. The unperturbed problem is assumed to be geometrically ergodic in the sense usually established through use of Foster-Lyapunov drift conditions. The perturbations are assumed to be uniform, in a weak sense, on bounded time intervals. The long-time behaviour of the perturbed chain is studied. Applications are given to numerical approximations of a randomly impulsed ODE, an Itô SDE and a parabolic SPDE subject to space-time Brownian noise. Existing perturbation theories for geometrically ergodic Markov chains are not readily applicable to these situations since they require very stringent hypotheses on the perturbations.

Key Words: Markov Chains, Ergodic Theory, Numerical Approximation, Random Impulses, Stochastic Differential Equations, Stochastic Partial Differential Equations.

AMS Subject Classifications: 60J10, 60J27, 65U, 60H10, 60H15, 34A37

[1]Scientific Computing and Computational Mathematics Program, Durand 257, Stanford University, Stanford CA94305-4040, USA. Supported by the National Science Foundation under grant DMS-95-04879.

[2]Current Address: IMA, University of Minnesota, 514 Vincent Hall, Minneapolis MN 55455–0436

1 Introduction

It is frequently of interest to understand how ergodic properties of Markov chains persist under various kinds of perturbations. Here perturbations to (discrete time) Markov chains and (continuous time) Markov processes evolving in a Banach space are considered. In both cases the perturbation is assumed to be a discrete time Markov chain and our primary motivation is to understand the *numerical* approximation of Markov chains and processes. The unperturbed Markov chain is assumed to be *geometrically ergodic* implying exponential convergence of expectations of functions from a certain class; the general framework of geometric ergodicity within which we operate is taken from the work of Meyn and Tweedie [20, 21] based on Foster-Lyapunov drift conditions. The perturbed Markov chains are assumed to be close to the unperturbed problem in a *weak sense*: the error in expectations of functions is small, uniformly on compact time-intervals disjoint from the origin, for functions in the same class. Perturbation theories for geometrically ergodic Markov chains do exist already, but it turns out that the class of perturbations considered there is typically too restrictive to admit application to the numerical methods considered here – at least for the finite time approximation results that we are currently able to obtain for these numerical methods. At the end of section 3 we will relate our perturbation theory to an existing perturbation theory due to Kartashov [13, 14, 15]. Properties of ergodic Markov chains under perturbation have been studied in many other contexts; for example in SDEs the idea of approximating white noise by a broad-band Gaussian noise process is of interest and this is studied in [1] and further in [17].

In section 2 our notation and general framework is established. A general theory is developed in section 3 and then, in section 4, applications described to three problems where, in all cases, the perturbation arises from numerical approximation. The first concerns an ODE subject to random impulses, the second an Itô SDE, and the third a parabolic SPDE subject to space-time Brownian noise.

The numerical simulation of ergodic stochastic processes in the context of finite dimensional SDEs has been studied by Talay [26] in the case where the generator of the process is uniformly parabolic; see also [27, 10] and [16]. However in many applications this uniform parabolicity does not hold; our theory encompasses problems for which the generator is not uniformly parabolic, albeit at a reduced rate of convergence when compared with the estimates in [26]. For stochastic parabolic PDEs, much is known about approximation properties on finite time intervals [4, 8, 11, 23], but the theory presented here enables us to prove long-time weak convergence properties in the geometrically ergodic case; such results have not been obtained before to the best of our knowledge.

2 Preliminaries

In the following $\mathbb{N} = \{1, 2, 3, \ldots\}$ and $\mathbb{Z}^+ = \{0, 1, 2, \ldots\}$. Let $\mathbb{S}$ be a Banach space and $\mathbb{B}(\mathbb{S})$ be the corresponding Borel σ–algebra. We consider Markov chains

$$\{u_n, n \in \mathbb{Z}^+\}, \tag{2.1}$$

of random variables from a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ to $(\mathbb{S}, \mathbb{B}(\mathbb{S}))$. For an initial distribution μ, let the probability triple $(\Omega, \mathcal{F}, \mathbb{P}^\mu)$ generate the chain (2.1). We use the notation δ^x to denote a point mass at $x \in \mathbb{S}$. Expectation with respect to $\mathbb{P}^\mu$ will be denoted by $\mathbb{E}^\mu$. Finally, we define measures μ_n on $(\mathbb{S}, \mathbb{B}(\mathbb{S}))$, parameterized by $\mu_0 = \mu$, according to

$$\mu_n(B) = \mathbb{P}^\mu\{\omega \in \Omega \colon u_n \in B\}, \quad B \in \mathbb{B}(\mathbb{S}).$$

We approximate the Markov chain (2.1) by a Markov chain

$$\{u_n^\varepsilon, n \in \mathbb{Z}^+\}, \tag{2.2}$$

of random variables from a probability space $(\Omega_\varepsilon, \mathcal{F}_\varepsilon, \mathbb{P}_\varepsilon)$ to $(\mathbb{S}, \mathbb{B}(\mathbb{S}))$. For an initial distribution μ, the probability triple $(\Omega_\varepsilon, \mathcal{F}_\varepsilon, \mathbb{P}_\varepsilon^\mu)$ is assumed to generate the chain (2.2). Expectation with respect to $\mathbb{P}_\varepsilon^\mu$ will be denoted by $\mathbb{E}^\mu$; this should cause no confusion as it will always be clear from the context which underlying probability space is giving rise to the expectation. In many applications the underlying probability spaces for (2.1) and (2.2) will be the same but this is not necessarily the case, for example when weak approximations of stochastic differential equations are studied. We define measures μ_n^ε on $(\mathbb{S}, \mathbb{B}(\mathbb{S}))$, parameterized by $\mu_0^\varepsilon = \mu^\varepsilon$, according to

$$\mu_n^\varepsilon(B) = \mathbb{P}_\varepsilon^\mu\{\omega \in \Omega_\varepsilon \colon u_n^\varepsilon \in B\}, \quad B \in \mathbb{B}(\mathbb{S}).$$

We will also consider the approximation of time-continuous Markov processes by time-discrete Markov chains of the form (2.2). Specifically we consider a stochastic process $\{u(t), t \geq 0\}$ of random variables from $(\Omega, \mathcal{F}, \mathbb{P})$ to $(\mathbb{S}, \mathbb{B}(\mathbb{S}))$. For an initial distribtuon μ, the probabilty triple $(\Omega, \mathcal{F}, \mathbb{P}^\mu)$ generates this process. For each fixed $\omega \in \Omega$ we thus have a path $u(\cdot) := u(\cdot\,; \omega)$. We define measures $\mu(t)$ on $(\mathbb{S}, \mathbb{B}(\mathbb{S}))$, parameterized by $\mu(0) = \mu$, according to

$$\mu(t)(B) = \mathbb{P}^\mu\{\omega \in \Omega : u(t) \in B\}, \quad B \in \mathbb{B}(\mathbb{S}).$$

If time is discretized with $t_n = n\varepsilon$ then we approximate the sampled chain $\{u(n\varepsilon), n \in \mathbb{Z}^+\}$ by a time-discrete Markov chain $\{u_n^\varepsilon, n \in \mathbb{Z}^+\}$ assumed to be of the form (2.2).

The notation $\lfloor x \rfloor$ will be used to denote the largest integer no larger than x. Throughout the paper C denotes a constant whose actual value may change between instances.

3 Basic Theory

The following two assumptions will be used in proving the basic results in this paper concerning approximation of (2.1) by (2.2):

- **(Ai)** The Markov chain (2.1) is geometrically ergodic. Specifically let $\bar{G} : \mathbb{S} \to [1, \infty)$ and define $\mathcal{G} = \{G : \mathbb{S} \to \mathbb{R}, |G| \leq \bar{G}\}$. Then $\bar{G} \in \mathcal{L}^1(\mathbb{S}, \mathbb{B}(\mathbb{S}), \mu_n)\ \forall n \geq 0$ and for some $R_1 > 0, r > 1$ and set $\mathcal{G}_0 \subseteq \mathcal{G}$ containing $\bar{G}$

$$\sup_{G \in \mathcal{G}_0} |\mathbb{E}^{\delta_x} G(u_n) - \pi(G)| \leq R_1 r^{-n} \bar{G}(x) \quad \forall n \geq 0$$

  with π the unique invariant measure on $(\mathbb{S}, \mathbb{B}(\mathbb{S}))$ generated by (2.1).

- **(Aii)** The Markov chains (2.1) and (2.2) satisfy, for some $\lambda > 1$ and $R_2 > 0$,

$$\sup_{G \in \mathcal{G}_0} |\mathbb{E}^{\delta_x} G(u_n) - \mathbb{E}^{\delta_x} G(u_n^\varepsilon)| \leq R_2 \bar{G}(x) \lambda^n \varepsilon \quad \forall n \geq 0.$$

The first assumption characterises the dependence of the rate of convergence of expectations under the chain to their limiting value, in terms of the deterministic initial value x. Typically such results are proved by deriving a Foster-Lyapunov type drift condition involving $\bar{G}$, together with construction of an irreducibility measure; see [20], Chapter 16 and [21]. Indeed for the randomly impulsed ODEs considered in this paper the condition in **(Ai)** is typically satisfied for any $G \in \mathcal{G}$. The second assumption states that expectations under (2.1) or (2.2) remain close, uniformly over a bounded number of transitions, and the dependence on deterministic initial data is again specified. Note however that the class of functions $\mathcal{G}_0$ for which **(Aii)** holds is typically smaller than $\mathcal{G}$ itself, at least for the currently available estimates we use for numerical approximations of Markov chains and processes.

We will prove the following theorems using these assumptions. For the second, we require two definitions: a function $\bar{G} : \mathbb{S} \to [1, \infty)$ is *norm-like* if

$$\lim_{n\to\infty} \left(\inf_{x \in \mathbb{S} - C_n} \bar{G}(x) \right) = \infty$$

for a sequence $C_n \subset \mathbb{S}$ of compact sets with $C_n \uparrow X$. A Markov chain is *weak Feller* if the transition kernel maps the space of bounded continuous functions on $\mathbb{S}$ into itself.

Theorem 3.1 *Under Assumptions* **(Ai)** *and* **(Aii)**, *there are* $K > 0$, $\varepsilon_c > 0$ *such that, for all* $0 < \varepsilon < \varepsilon_c$ *there is* $N = N(\varepsilon)$:

$$\sup_{G \in \mathcal{G}_0} |\mathbb{E}^{\delta_x} G(u_n^\varepsilon) - \pi(G)| \leq 2K \max\{\pi(\bar{G}), \bar{G}(x)\} \varepsilon^\gamma \quad \forall n \geq N \tag{3.1}$$

where $\gamma = \log r / \log(\lambda^2 r) \in (0,1)$. *If the Markov chain* (2.2) *has an invariant probability measure* π^ε, *then*

$$\sup_{G \in \mathcal{G}_0} |\pi^\varepsilon(G) - \pi(G)| \leq 6K\pi(\bar{G})\varepsilon^\gamma.$$

Corollary 3.2 *Let* $\mathbb{S}$ *be locally compact and separable and the functional* $\bar{G}$ *be norm-like and let assumptions* **(Ai)** *and* **(Aii)** *hold. If the perturbed chain* (2.2) *is weak Feller, then it has at least one invariant probability measure* π^ε.

The first theorem confines stabilization of expected values under (2.2) to a small neighbourhood of their limit under (2.1). The second theorem shows that, under some regularity conditions, the perturbed chain has an invariant measure, which in general is not unique. The hypothesis on the space $\mathbb{S}$ in the second theorem apply to finite dimensional vector spaces; those interested in existence of invariant measures on infinite dimensional spaces should look in [3].

Note that, if the perturbed chain is ergodic, then (3.1) implies that time averages of G under the perturbed chain will converge to a limit $\pi^{\varepsilon}(G)$ which is $\mathcal{O}(\varepsilon^{\gamma})$ close to the time average $\pi(G)$ under the original chain. It remains an open and interesting question to study when indeed ergodicity is inherited by the perturbed chain under the type of assumptions made in this paper.

Proof of Theorem 3.1 Choose $N \in \mathbb{R}$ such that $\varepsilon\lambda^{2N} = r^{-N}$ so that $r^{-N} = \varepsilon^{\gamma}$. By choosing ε_c sufficiently small we can ensure that $N \geq 2$ for all $\varepsilon \leq \varepsilon_c$. With $N \geq 2$ we have, by Assumptions **(Ai)** and **(Aii)**,

$$\begin{array}{rcl} \sup_{G\in\mathcal{G}_0} |\mathbb{E}^{\delta_x} G(u_n) - \pi(G)| & \leq & rR_1\varepsilon^{\gamma}\bar{G}(x) \quad \forall n \geq N-1, \\ \sup_{G\in\mathcal{G}_0} |\mathbb{E}^{\delta_x} G(u_n^{\varepsilon}) - \mathbb{E}^{\delta_x} G(u_n)| & \leq & R_2\varepsilon^{\gamma}\bar{G}(x) \qquad \forall n \leq 2N. \end{array}$$

Thus

$$\sup_{G\in\mathcal{G}_0} |\mathbb{E}^{\delta_x} G(u_n^{\varepsilon}) - \pi(G)| \leq (rR_1 + R_2)\varepsilon^{\gamma}\bar{G}(x) \quad \forall N-1 \leq n \leq 2N. \tag{3.2}$$

Let $K = rR_1 + R_2$ and assume, for induction, that the following holds for some $M \geq 1$:

$$\begin{array}{c} \sup_{G\in\mathcal{G}_0} |\mathbb{E}^{\delta_x} G(u_n^{\varepsilon}) - \pi(G)| \leq \pi(\bar{G})\sum_{j=1}^{M-1}(K\varepsilon^{\gamma})^j + \bar{G}(x)(K\varepsilon^{\gamma})^M \\ \forall n : NM \leq n < N(M+1). \end{array} \tag{3.3}$$

Note that (3.3) holds for $M = 1$ by (3.2).

Let p be any integer: $N(M+1) \leq p < N(M+2)$. Let n be the closest integer to $(p-N)$ subject to the constraint that n lies in $[NM, N(M+1))$. Then, since $N \geq 2$, $p = n + L$ where $N-1 \leq L \leq 2N$. Now, by conditional expectation and the Markov property, for any $G \in \mathcal{G}_0$,

$$\begin{array}{rcl} |\mathbb{E}^{\delta_x} G(u_p^{\varepsilon}) - \pi(G)| & = & |\mathbb{E}^{\delta_x} G(u_{n+L}^{\varepsilon}) - \pi(G)| \\ & = & |\mathbb{E}^{\mu_L^{\varepsilon}} G(u_n^{\varepsilon}) - \pi(G)| \\ & = & \left| \int_{\mathbb{S}} [\mathbb{E}^{\delta_x} G(u_n^{\varepsilon}) - \pi(G)]\mu_L^{\varepsilon}(dx) \right| \\ & \leq & \int_{\mathbb{S}} |\mathbb{E}^{\delta_x} G(u_n^{\varepsilon}) - \pi(G)|\mu_L^{\varepsilon}(dx). \end{array}$$

Thus, by the inductive hypothesis (3.3),

$$\begin{array}{rcl} |\mathbb{E}^{\delta_x} G(u_p^\varepsilon) - \pi(G)| & \leq & \mu_L^\varepsilon[\pi(\bar{G}) \sum_{j=1}^{M-1} (K\varepsilon^\gamma)^j + \bar{G}(\cdot)(K\varepsilon^\gamma)^M] \\ & = & \pi(\bar{G}) \sum_{j=1}^{M-1} (K\varepsilon^\gamma)^j + (K\varepsilon^\gamma)^M \mu_L^\varepsilon(\bar{G}). \end{array} \tag{3.4}$$

But $\mu_L^\varepsilon(\bar{G}) = \mathbb{E}^{\delta_x} \bar{G}(u_L^\varepsilon)$ and so, by (3.2) since $\bar{G} \in \mathcal{G}_0$,

$$\begin{array}{rcl} \mu_L^\varepsilon(\bar{G}) & = & \pi(\bar{G}) + [\mathbb{E}^{\delta_x} \bar{G}(u_L^\varepsilon) - \pi(\bar{G})] \\ & \leq & \pi(\bar{G}) + K\varepsilon^\gamma \bar{G}(x) \end{array} \tag{3.5}$$

and combining (3.4) and (3.5) gives

$$\begin{array}{rcl} |\mathbb{E}^{\delta_x} G(u_p^\varepsilon) - \pi(G)| & \leq & \pi(\bar{G}) \sum_{j=1}^{M-1} (K\varepsilon^\gamma)^j + (K\varepsilon^\gamma)^M \pi(\bar{G}) + (K\varepsilon^\gamma)^{M+1} \bar{G}(x) \\ & = & \pi(\bar{G}) \sum_{j=1}^{M} (K\varepsilon^\gamma)^j + (K\varepsilon^\gamma)^{M+1} \bar{G}(x). \end{array}$$

Since $G \in \mathcal{G}_0$ is arbitrary, (3.3) is established for all $M \geq 1$ by induction. Now reduce ε_c so that $K\varepsilon^\gamma \leq \frac{1}{2}$ for $0 < \varepsilon < \varepsilon_c$. Then (3.3) implies, for $0 < \varepsilon < \varepsilon_c$,

$$\begin{array}{rcl} \sup_{G \in \mathcal{G}_0} |\mathbb{E}^{\delta_x} G(u_n^\varepsilon) - \pi(G)| & \leq & \max\{\pi(\bar{G}), \bar{G}(x)\} \sum_{j=1}^{M} (K\varepsilon^\gamma)^j \\ & \leq & 2 \max\{\pi(\bar{G}), \bar{G}(x)\} K\varepsilon^\gamma, \end{array}$$

giving (3.1). Average (3.1) over data distributed according to π^ε to obtain

$$\sup_{G \in \mathcal{G}_0} |\pi^\varepsilon(G) - \pi(G)| \leq 2K\{\pi(\bar{G}) + \pi^\varepsilon(\bar{G})\}\varepsilon^\gamma.$$

Thus, for ε sufficiently small,

$$\pi^\varepsilon(\bar{G}) \leq 2\pi(\bar{G})$$

and the required result follows. □

Proof of Corollary 3.2 Under the stated hypotheses, (3.1) holds and so Proposition 12.1.4 in [20] applies,giving the existence of an invariant probability measure. □

We also consider the approximation of the time-continuous Markov process $\{u(t)\}_{t \geq 0}$ by (2.2). For this we will use the following modifications of **(Ai)**–**(Aii)**.

- **(ACi)** The Markov process $\{u(t)\}_{t\geq 0}$ is geometrically ergodic. Specifically let $\bar{G}:\mathbb{S}\to[1,\infty)$ and define $\mathcal{G}=\{G:\mathbb{S}\to\mathbb{R},|G|\leq\bar{G}\}$. Then $\bar{G}\in\mathcal{L}^1(\mathbb{S},\mathbb{B}(\mathbb{S}),\mu(t))\ \forall t\geq 0$ and for some $R_1>0,\omega>0$ and set $\mathcal{G}_0\subseteq\mathcal{G}$ containing $\bar{G}$

$$\sup_{G\in\mathcal{G}_0}|\mathbb{E}^{\delta_x}G(u(t))-\pi(G)|\leq R_1e^{-\omega t}\bar{G}(x)\quad\forall t\geq 0$$

  with π the unique invariant measure on $(\mathbb{S},\mathbb{B}(\mathbb{S}))$.

- **(ACii)** The Markov chain $\{u_n^\varepsilon\}_{n\geq 0}$ approximates the time-continuous Markov chain $\{u(t)\}_{t\geq 0}$ in the following way: for some $\alpha>0,R_2>0$ and $s>0$,

$$\sup_{G\in\mathcal{G}_0}|\mathbb{E}^{\delta_x}G(u(n\varepsilon))-\mathbb{E}^{\delta_x}G(u_n^\varepsilon)|\leq R_2\bar{G}(x)e^{\alpha n\varepsilon}\varepsilon^s\quad\forall n:n\varepsilon\geq 1.$$

Note that for stochastic differential equations the condition in **(Ai)** is typically satisfied for any $G\in\mathcal{G}$. However, for the stochastic PDEs considered here the condition is currently only verified for a proper subset of $\mathcal{G}$ in which a Lipschitz condition is satisfied. Hence the formulation of **(ACi)** in terms of $\mathcal{G}_0$. In any case for our current methods of analysis we can only verify **(ACii)** on a proper subset of G for the applications to SDEs and SPDEs considered here.

The following theorem may be proved similarly to Theorems 3.1—the necessary modifications to the proofs are outlined below.

Theorem 3.3 *Under Assumptions* **(ACi)** *and* **(ACii)** *there is* $K>0$, $\varepsilon_c>0$ *such that for all* $0<\varepsilon<\varepsilon_c$ *there is* $T=T(\varepsilon)$*:*

$$\sup_{G\in\mathcal{G}_0}|\mathbb{E}^{\delta_x}G(u_n^\varepsilon)-\pi(G)|\leq 2K\max\{\pi(\bar{G}),\bar{G}(x)\}\varepsilon^\gamma\quad\forall n:n\varepsilon\geq T \tag{3.6}$$

where $\gamma=\frac{\omega}{\omega+2\alpha}\in(0,1)$*. If the Markov chain* (2.2) *has an invariant probability measure* π^ε*, then*

$$\sup_{G\in\mathcal{G}_0}|\pi^\varepsilon(G)-\pi(G)|\leq 6K\pi(\bar{G})\varepsilon^\gamma.$$

Corollary 3.4 *Let $\mathbb{S}$ be locally compact and separable and the functional $\bar{G}$ be norm-like and let assumptions* **(ACi)** *and* **(ACii)** *hold. If the perturbed chain* (2.2) *is weak Feller, then it has at least one invariant probability measure π^{ε}.*

Proofs of Theorem 3.3 *and Corollary* 3.4 To prove Theorem 3.3 choose $\varepsilon_c < 1$ sufficiently small so that there is $T > 1$ solving

$$e^{-\omega T} = e^{2\alpha T}\varepsilon^{s}, \qquad 0 < \varepsilon < \varepsilon_c.$$

Then $e^{-\omega T} = \varepsilon^{\gamma s}$. By combining **(ACi)** and **(ACii)** we deduce that

$$\sup_{G \in \mathcal{G}_0} |\mathbb{E}^{\delta_x} G(u_n^{\varepsilon}) - \pi(G)| \le K\varepsilon^{s\gamma}\bar{G}(x), \quad T - \varepsilon_c \le n\varepsilon \le 2T \tag{3.7}$$

where $K = R_1 \varepsilon^{\varepsilon_c \omega} + R_2$. Now assume for induction that

$$\begin{gathered}\sup_{G \in \mathcal{G}_0} |\mathbb{E}^{\delta_x} G(u_n^{\varepsilon}) - \pi(G)| \le \pi(\bar{G}) \sum_{j=1}^{M-1} (K\varepsilon^{s\gamma})^j + \bar{G}(x)(K\varepsilon^{s\gamma})^M \\ \forall n : MT \le n\varepsilon \le (M+1)T.\end{gathered} \tag{3.8}$$

The induction now proceeds similarly to the proof of Theorem 3.1: choose p such that $(M+1)T \le p\varepsilon \le (M+2)T$ and define n to be the unique integer closest to $(p\varepsilon - T)/\varepsilon$ and such that $MT \le n\varepsilon \le (M+1)T$. Then define $N := p - n$. It follows that $T - \varepsilon \le N\varepsilon \le T + \varepsilon$. In particular $T - \varepsilon_c \le N\varepsilon \le 2T$ and (3.7) holds. The induction step can again be made by using (3.7); the details are omitted. Averaging as in Corollary 3.2 gives the Corollary 3.4. $\square$

Remark Let Q (resp. Q^{ε}) denote the transition kernel for the chain (2.1) (resp. (2.2)) and Π the lift of π to a transition kernel. The invariant measure is a solution of the eigenvalue problem

$$\pi Q = \pi, \quad \pi(\mathbb{S}) = 1$$

and in essence our interest is focused on perturbation theory for this eigenvalue problem. Such problems have, of course, been studied before. If, for given $v : \mathbb{S} \to [1, \infty)$, we define a norm on measures on $\mathbb{S}$ by

$$\|\mu\|_v = \int_{\mathbb{S}} v(x)|\mu|(dx)$$

then the induced operator norm for transition kernels P is

$$\|P\|_v = \sup_{x\in\mathbb{S}} \frac{\int_S v(y)|P(x,dy)|}{v(x)}.$$

This norm is discussed in some detail in [20], Chapter 16 and also in [13, 14]. (Note that if $v \equiv 1$ then $\|\cdot\|_v$ reduces to the total variation norm). It is straightforward to show that

$$\|\mu\|_v \le \varepsilon \Longleftrightarrow |\int_{\mathbb{S}} h(x)\mu(dx)| \le \varepsilon \quad \forall h : |h| \le v$$

and that

$$\|P\|_v \le \varepsilon \Longleftrightarrow |\int_{\mathbb{S}} h(y)P(x,dy)| \le \varepsilon v(x) \quad \forall h : |h| \le v.$$

In the case where $\mathcal{G}_0 \equiv \mathcal{G}$ assumption **(Ai)** states that

$$\|Q^n - \Pi\|_{\bar{G}} \le \frac{R_1}{r^n} \quad \forall n \ge 0$$

and **(Aii)** states that

$$\|Q^n - (Q^\varepsilon)^n\|_{\bar{G}} \le R_2 \lambda^n \varepsilon \quad \forall n \ge 0.$$

Under these two conditions Kartashov shows in [14] that Q^ε has an invariant probability measure $\mathcal{O}(\varepsilon)$ close to π, a stronger result than we are able to prove. (This is essentially a differentiability result for the invariant measures; related issues are discussed in [7].) However, whilst **(Ai)** can often be extended to the whole of $\mathcal{G}$, our current analysis of numerical methods does not enable us to extend **(Aii)** in this way for the problems considered here. Hence we are unable to apply the Kartashov theory. This motivates the development of the present theory.

4 Applications

We present applications of the foregoing theory: to randomly impulsed ordinary differential equations, to Itô stochastic differential equations and to stochastic parabolic PDEs subject to space-time Brownian noise.

4.1 Randomly Impulsed Differential Equations

Given a vector $\nu \in \mathbb{R}^m$ (the impulse), an initial state u_0 and a vector field $f \in C^\infty(\mathbb{R}^m, \mathbb{R}^m)$, we consider the equation

$$\frac{du}{dt} = f(u) + \sum_{n=1}^{\infty} \theta_n \nu \delta(t - \tau_n), \quad u(0^+) = x + \theta_0 \nu. \tag{4.1}$$

Here $\delta(\cdot)$ denotes a unit point mass at the origin, the $\{\theta_n\}_{n=0}^\infty$ are independent, identically distributed (IID) random variables with probability of $\theta_n = \pm 1$ being $\frac{1}{2}$, and the waiting times $t_n = \tau_{n+1} - \tau_n$, $n \in \mathbb{Z}^+$, $\tau_0 = 0$, are IID random variables exponentially distributed with parameter λ. Moreover we assume that the sequences $\{\theta_n\}_{n=0}^\infty$ and $\{t_n\}_{n=0}^\infty$ are independent. Note that $\tau_n \to \infty$, almost surely.

We denote by $S \in C^\infty(\mathbb{R}^m \times \mathbb{R}^+, \mathbb{R}^m)$ the semigroup assumed to be generated by the problem

$$\frac{du}{dt} = f(u), \quad u(0) = U; \tag{4.2}$$

thus the solution of (4.2) is $u(t) = S(U, t)$ for $t \geq 0$. If we define $u_n = u(\tau_n^-)$, $w_n = (\theta_n, t_n)^T$ and use the delta function to derive jump conditions on $u(t)$ across $t = t_n$, we obtain the Markov chain

$$u_{n+1} = H(u_n, w_n) := S(u_n + \theta_n \nu, t_n), \quad n \in \mathbb{Z}^+, \quad u_0 = x; \tag{4.3}$$

see [22]. (Note also that, when viewed in continuous time, (4.1) is an example of a piecewise deterministic Markov process in the general framework of Davis [5].) Let $\mathcal{F}_n$ denote the σ algebra generated by the random variables $\{u_i : i \leq n\}$.

We assume that f satisfies the following conditions:

- **(R1)** $\exists C > 0 : \|f(u)\| \leq C[1 + \|u\|] \quad \forall u \in \mathbb{R}^m$.
- **(R2)** all derivatives of f are uniformly bounded on $\mathbb{R}^m$ and the Lipschitz constant K for f satisfies $K < \lambda$.
- **(R3)** $\exists \alpha, \beta > 0 : \langle f(u), u \rangle \leq \alpha - \beta \|u\|^2 \quad \forall u \in \mathbb{R}^m$.

Under these conditions, the following lemma holds as proved in [22]:

Lemma 4.1 *Consider the Markov chain* (4.3) *with x distributed according to a measure μ on* $(\mathbb{R}^m, \mathbb{B}(\mathbb{R}^m))$ *satisfying* $\mu(\|x\|^2) < \infty$. *If* **(R1)–(R3)** *hold then* $\{u_n\}_{n=0}^{\infty}$ *exists* $\mathbb{P}^\mu - a.s.$ *and*

$$\|u_{n+1}\|^2 \leq \frac{\alpha}{\beta} + e^{-2\beta t_n}(\|u_n\|^2 + 2\theta_n\langle u_n, \nu\rangle + \|\nu\|^2 - \frac{\alpha}{\beta})$$

from which it follows that

$$\sup_{n\geq 0} \mathbb{E}^\mu \|u_n\|^2 \leq \max\{\mu(\|x\|^2), \frac{\alpha}{\beta} + \frac{1}{2\beta}\lambda\|\nu\|^2\} \tag{4.4}$$

and

$$\mathbb{E}^\mu\{\|u_{n+1}\| - \|u_n\| | \mathcal{F}_n\} \leq \left(\sqrt{\frac{\lambda}{\lambda + 2\beta}} - 1\right)\|u_n\| + \sqrt{\frac{\alpha}{\beta}} + \sqrt{\frac{\lambda}{\lambda+2\beta}}\|\nu\|. \tag{4.5}$$

By use of (4.5) it is possible to show that the Markov chain (4.3) is geometrically ergodic with respect to the function $\bar{G}(u) := 1 + \|u\|$, provided an irreducibility measure, satisfying certain continuity properties, exists – see [20] for the general theory and [22] for applications to (4.1). Thus we will assume later in this section:

- **(R4)** The Markov chain (4.3) is geometrically ergodic: for $\bar{G}(u) := 1 + \|u\|$, $\exists R > 0, r > 1$ such that, for

  $$\mathcal{G} = \{G : \mathbb{R}^m \to [1, \infty), |G| \leq \bar{G}\},$$

  we have

  $$\sup_{G\in\mathcal{G}} |\mathbb{E}^{\delta_x} G(u_n) - \pi(G)| \leq R\bar{G}(x)r^{-n} \quad \forall n \geq 0$$

  where π is the unique invariant measure on $(\mathbb{R}^m, \mathbb{B}(\mathbb{R}^m))$ generated by (4.3).

We assume that the deterministic part of the equation (4.1) is approximated by a one-step numerical method with map $S^1_{\Delta t} : \mathbb{R}^m \to \mathbb{R}^m$ denoting one-step of the method with time-step Δt (and $S^N_{\Delta t}$ its $N-$fold composition; we assume that this map is defined and continuous on $\mathbb{R}^m$ for all $\Delta t \in [0, \Delta t_c)$, which is true for all Runge-Kutta methods under **(R2)**. Given any $t \in \mathbb{R}^+$ we let $N = N(t)$ and $\Delta = \Delta(t)$ be defined by

$$N = \left\lfloor \frac{t}{\Delta t} \right\rfloor, \quad \Delta \in [0, \Delta t) : t = N\Delta t + \Delta.$$

We may thus define, for $w = (\theta, t)^T$,

$$H_{\Delta t}(u, w) := S^1_{\Delta}(S^N_{\Delta t}(u + \theta\nu))$$

and then the Markov chain

$$u^{\Delta t}_{n+1} = H_{\Delta t}(u_n, w_n), \quad u_0 = x \tag{4.6}$$

is our numerical approximation to the chain (4.3). This too is defined for all $\Delta t \in [0, \Delta t_c)$. Let $\mathcal{F}^{\Delta t}_n$ denote the $\sigma-$algebra generated by the random variables $\{u^{\Delta t}_i : i \leq n\}$.

We assume that, under **(R2)**, the one-step map satisfies the standard approximation property that $\exists C > 0$ so that, for all $\Delta t \in [0, \Delta t_c)$,

$$\|S^1_{\Delta t} u - S(\Delta t)u\| \leq C[1 + \|u\|]\Delta t^{r+1} \quad \forall u \in \mathbb{R}^m \tag{4.7}$$

for some integer $r \geq 1$; this holds for certain low order Runge-Kutta methods under **(R1)**, **(R2)**.

The following lemma may be proved by the techniques of Theorem 7.3.1 in [25] using (4.7).

Lemma 4.2 *Consider the Markov chain* (4.6) *with x distributed according to a measure μ on* $(\mathbb{R}^m, \mathbb{B}(\mathbb{R}^m))$ *satisfying $\mu(\|x\|^2) < \infty$. If* **(R1)–(R3)** *hold then $\{u_n\}_{n=0}^\infty$ exists $\mathbb{P}^\mu -$ a.s. and*

$$\|u^{\Delta t}_{n+1}\|^2 \leq \gamma + e^{-2\bar{\beta} t_n}(\|u^{\Delta t}_n\|^2 + 2\theta_n\langle u^{\Delta t}_n, \nu\rangle + \|\nu\|^2 - \gamma)$$

where $\gamma = \alpha/\beta + \mathcal{O}(\Delta t)$ and $\bar{\beta} < \beta$ satisfies $\bar{\beta} = \beta + \mathcal{O}(\Delta t)$. From this it follows that

$$\sup_{n \geq 0} \mathbb{E}^\mu \|u^{\Delta t}_n\|^2 \leq \max\{\mu(\|x\|^2), \gamma + \frac{1}{2\bar{\beta}}\lambda\|\nu\|^2\}$$

and

$$\mathbb{E}^{\mu}\{\|u_{n+1}^{\Delta t}\| - \|u_n^{\Delta t}\| | \mathcal{F}_n^{\Delta t}\} \le \left(\sqrt{\frac{\lambda}{\lambda+2\bar{\beta}}} - 1\right)\|u_n^{\Delta t}\| + \sqrt{\gamma} + \sqrt{\frac{\lambda}{\lambda+2\bar{\beta}}}\|\nu\|. \tag{4.8}$$

The next result addresses the pathwise approximation of (4.3) by (4.6).

Result 4.3 *Consider the approximation of the Markov chain* (4.3) *by the Markov chain* (4.6)*, under* **(R1)–(R3)**. *Then for any measure* μ *on* $\mathbb{R}^m$ *with* $\mu(\|x\|^2) < \infty$ *there is* $C > 0$:

$$\mathbb{E}^{\mu}\|u_n - u_n^{\Delta t}\| \le C[1 + \mu(\|x\|)]\left(\frac{\lambda}{\lambda - K}\right)^n \Delta t^r \quad \forall n \ge 0.$$

Proof By use of **(R2)** it follows that

$$\|H(u_1, w) - H(u_2, w)\| \le e^{Kt}\|u_1 - u_2\| \quad \forall u_1, u_2 \in \mathbb{R}^m \tag{4.9}$$

where $w = (\theta, t)^T$. Similarly (4.7) implies that

$$\|H(u, w) - H_{\Delta t}(u, w)\| \le Le^{Kt}[1 + \|u\|]\Delta t^r \quad \forall u \in \mathbb{R}^m. \tag{4.10}$$

The bound (4.10) is simply the standard convergence result for one-step methods modified by the uniform truncation error bound (4.7) and the global bound on solutions of (4.2), and its numerical approximations, induced by **(R3)**.

Now

$$\|u_{n+1} - u_{n+1}^{\Delta t}\| \le \|H(u_n, w_n) - H(u_n^{\Delta t}, w_n)\| + \|H(u_n^{\Delta t}, w_n) - H_{\Delta t}(u_n^{\Delta t}, w_n)\|. \tag{4.11}$$

Using the properties of the exponential distribution we have, from (4.9),

$$\mathbb{E}^{\mu}(\|H(u_n, w_n) - H(u_n^{\Delta t}, w_n)\| \,|\sigma(\mathcal{F}_n, \mathcal{F}_n^{\Delta t})) \le \left(\frac{\lambda}{\lambda - K}\right)\|u_n - u_n^{\Delta t}\|$$

(where $\sigma(\mathcal{F}_n, \mathcal{F}_n^{\Delta t})$ is the smallest σ–algebra containing both $\mathcal{F}_n$ and $\mathcal{F}_n^{\Delta t}$) and from (4.10)

$$\mathbb{E}^{\mu}(\|H(u_n^{\Delta t}, w_n) - H_{\Delta t}(u_n^{\Delta t}, w_n)\| \,|\mathcal{F}_n^{\Delta t}) \le C_1[1 + \|u_n^{\Delta t}\|]\Delta t^r.$$

Hence, by (4.11),

$$\mathbb{E}^\mu \|u_{n+1} - u_{n+1}^{\Delta t}\| \leq \left(\frac{\lambda}{\lambda - K}\right) \mathbb{E}^\mu \|u_n - u_n^{\Delta t}\| + C_1 \Delta t^r \mathbb{E}^\mu (1 + \|u_n^{\Delta t}\|). \quad (4.12)$$

By (4.8) there exists $q \in (0,1)$ and $r > 0$ such that

$$\mathbb{E}^\mu \|u_{n+1}^{\Delta t}\| \leq r + q\mathbb{E}^\mu \|u_n^{\Delta t}\|$$

so that

$$\mathbb{E}^\mu \|u_n^{\Delta t}\| \leq \left(\frac{r}{1-q}\right) [1 - q^n] + q^n \mathbb{E}^\mu \|u_0^{\Delta t}\|.$$

Thus (4.12) yields, for $e_n := \mathbb{E}^\mu \|u_n^{\Delta t} - u_n\|$,

$$e_n \leq C_2 \left[\left(\frac{\lambda}{\lambda - K} \right)^n - 1 \right] \Delta t^r + C_3 [1 + \mathbb{E}^\mu \|u_0^{\Delta t}\|][1 - q^n] \Delta t^r$$

giving the desired result.

In the following let $\mathcal{G}$ be as defined in **(R4)** and, given a Lipschitz constant C_G, define the following subset of $\mathcal{G}$:

$$\mathcal{G}_0 := \{G \in \mathcal{G} \colon |G(x) - G(y)| \leq C_G \|x - y\|\}.$$

Using the pathwise approximation result proved in Result 4.3, together with geometric ergodicity of (4.3), we prove:

Theorem 4.4 *Consider the approximation of the Markov chain* (4.3) *by the Markov chain* (4.6)*, under* **(R1)**–**(R4)**. *Then, for any* $G \in \mathcal{G}_0$ *there is* $C > 0$, $\delta \in (0,1)$ *and for any* $\Delta t \in [0, \Delta t_c)$ *an integer* $N = N(\Delta t)$ *such that*

$$|\mathbb{E}^{\delta_x} G(u_n^{\Delta t}) - \pi(G)| \leq 2C \max\{\pi(\bar{G}), \bar{G}(x)\} \Delta t^{r\delta} \quad \forall n \geq N. \qquad (4.13)$$

Furthermore the Markov chain (4.6) *has an invariant probability measure* $\mu^{\Delta t}$ *and*

$$|\pi^{\Delta t}(G) - \pi(G)| \leq 6C\pi(\bar{G})\Delta t^{r\delta}. \qquad (4.14)$$

Proof By Result 4.3 we have, for all $G \in \mathcal{G}_0$,

$$\begin{aligned} |\mathbb{E}^{\delta_x} G(u_n) - \mathbb{E}^{\delta_x} G(u_n^{\Delta t})| &\leq \mathbb{E}^{\delta_x} |G(u_n) - G(u_n^{\Delta t})| \\ &\leq C_G \mathbb{E}^{\delta_x} \|u_n - u_n^{\Delta t}\| \\ &\leq CC_G[1 + \|x\|] \left(\tfrac{\lambda}{\lambda - K}\right)^n \Delta t^r. \end{aligned}$$

Thus **(Aii)** holds with $\varepsilon = \Delta t^r$. Now **(Ai)** holds by **(R4)**. Thus Theorem 3.1 gives (4.13). Solutions of (4.6) are weak Feller by the continuity properties of $S^1_{\Delta t}$. Therefore, because we are working in finite dimensions, existence of an invariant measure comes from Corollary 3.2. □

4.2 Itô Stochastic Differential Equations

Consider the Itô stochastic differential equation

$$du = f(u)dt + \sigma(u)dW, \quad u(0) = x \tag{4.15}$$

where $u \in \mathbb{R}^m, f : \mathbb{R}^m \to \mathbb{R}^m, \sigma : \mathbb{R}^m \mapsto \mathbb{R}^{m \times s}$ and W is an s dimensional Brownian motion. We make the following assumptions about f and σ :

- **(I1)** $\|f(u)\| \leq C[1 + \|u\|]$ & $\|f(u) - f(v)\| \leq C\|u - v\| \quad \forall u, v \in \mathbb{R}^m$.
- **(I2)** $\|\sigma(u)\| \leq C$ & $\|\sigma(u) - \sigma(v)\| \leq C\|u - v\| \quad \forall u, v \in \mathbb{R}^m$.
- **(I3)** $\exists \alpha, \beta > 0 : \langle f(u), u \rangle \leq \alpha - \beta\|u\|^2 \quad \forall u \in \mathbb{R}^m$.

Conditions **(I1)** and **(I2)** ensure the existence and uniqueness of solutions $\mathbb{P}^{\delta_x} - a.s.$ with respect to the product measure induced by the Brownian motion and a point mass at x as a measure on the initial data. Furthermore, using **(I2)** and **(I3)**, we can prove

$$\sup_{t \geq 0} \mathbb{E}^{\delta_x} \|u(t)\|^2 \leq C[1 + \|x\|^2]. \tag{4.16}$$

We define the $m \times m$ matrix A with entries

$$A_{ij}(u) = \sum_{r=1}^{s} \sigma_{ir}(u)\sigma_{jr}(u)$$

and the second order differential operator $\mathcal{L}$ (the adjoint of the generator) defined by, for $u = (u_1, \ldots, u_m)$,

$$(\mathcal{L}\phi)(u) = \sum_{i=1}^{m} f_i(u) \frac{\partial \phi}{\partial u_i}(u) + \frac{1}{2} \sum_{i,j=1}^{m} A_{ij}(u) \frac{\partial^2 \phi(u)}{\partial u_i \partial u_j}.$$

It is well-known that, if the symmetric non-negative matrix A is positive definite in a compact set containing $B(0; \sqrt{\alpha/\beta})$ in its interior then, under **(I3)**, the Markov process generated by (4.15) is ergodic and the strong law of large numbers holds [12]. Under this type of strong ellipticity condition on $\mathcal{L}$ it is possible to show that a variety of numerical methods reproduce large-time expectations and time-averages accurately [26]. However many problems of interest do not satisfy the strong ellipticity condition but are still ergodic; we show how the theory of section 3 may be used to verify that numerical methods still reproduce large-time expectations and time averages in such cases.

As a simple example consider the Langevin equation

$$\frac{d^2x}{dt^2} + \gamma \frac{dx}{dt} + V'(x) = \frac{dW}{dt}, \quad x(0) = x_0, \frac{dx}{dt}(0) = x_1 \tag{4.17}$$

where W is a one-dimensional Brownian motion. This may be re-cast in the form (4.15) with $u = (u_1, u_2)^T := (x, \dot{x})^T$, $m = 2$ and $s = 1$. Then

$$f(u) = \begin{pmatrix} u_2 \\ -V'(u_1) - \gamma u_2 \end{pmatrix}, \quad \sigma(u) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad A = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Clearly A is not positive-definite so that the work in [26] cannot be applied to study long-time weak approximation properties of numerical methods. However the equation may be shown to be ergodic for $\gamma > 0$ (see [28] and [12], section 4.8) and certain potentials V. In [9] and [19] the effect of numerical approximation on this equation is studied when V is quadratic in x so that the equation is linear; however the results are based on explicit formulae and cannot be generalized to other potentials. The techniques we describe can be.

We assume that the SDE (4.15) is geometrically ergodic with respect to a quadratic function $\bar{G} = 1 + \|u\|^2$ on the phase space of the problem. To motivate this choice of Lyapunov function $\bar{G}$ note that

$$\mathcal{L}\bar{G} \leq 2(\alpha + \beta) + tr(A) - \beta\bar{G}.$$

But $tr(A)$ is equal to the square of the Frobenius norm of σ and is hence bounded by **(I2)**. Hence, provided all compact sets are petite for some skeleton chain (which essentially requires construction of an irreducibility measure with certain continuity properties) Theorem 6.1 of [21], part III proves geometric ergodicity with respect to $\bar{G}$. Hence we assume, for $\mathcal{G}$ as defined in **(Ai)**,

- **(I4)** The continuous time Markov process generated by (4.15) is geometrically ergodic: for $\bar{G}(u) := 1 + \|u\|^2 \ \exists R > 0, \gamma > 0$:

$$\sup_{G \in \mathcal{G}} |\mathbb{E}^{\delta_x} G(u(t)) - \pi(G)| \leq R\bar{G}(x)e^{-\gamma t} \quad \forall t \geq 0$$

  where π is the unique invariant measure on $(\mathbb{R}^m, \mathbb{B}(\mathbb{R}^m))$.

Now consider the approximation of (4.15) by the Euler-Maruyama scheme

$$U_{j+1} - U_j = f(U_j)\Delta t + \sigma(U_j)\Delta W_j, \quad U_0 = x \tag{4.18}$$

where $\Delta W_j = W((j+1)\Delta t) - W(j\Delta t)$. By the techniques described in [16] it may be shown that, for some $\alpha > 0$,

$$\mathbb{E}^{\delta_x}(\|U_J - u(J\Delta t)\|^2) \leq C_3 e^{2\alpha J\Delta t}[1 + \|x\|^2]\Delta t, \tag{4.19}$$

using **(I3)** to control all necessary constants appearing in terms of $\bar{G}(x)$. In the following we use $\mu_n^{\Delta t}$ to denote the measure on $(\mathbb{R}^m, \mathbb{B}(\mathbb{R}^m))$ induced by the Markov chain (4.18).

We now derive a uniform mean square bound on solutions of (4.18). Rearranging (4.18), squaring and taking expectations yields

$$\begin{aligned}
\mathbb{E}^{\delta_x}\|U_{j+1}\|^2 &\leq \mathbb{E}^{\delta_x}\|U_j\|^2 + 2\Delta t\mathbb{E}^{\delta_x}\langle U_j, f(U_j)\rangle + \mathbb{E}^{\delta_x}\|f(U_j)\|^2\Delta t^2 + C^2\Delta t \\
&\leq \mathbb{E}^{\delta_x}\|U_j\|^2 + 2\Delta t\mathbb{E}^{\delta_x}(\alpha - \beta\|U_j\|^2) \\
&\quad + C^2\mathbb{E}^{\delta_x}(1 + \|U_j\|^2)\Delta t^2 + C^2\Delta t.
\end{aligned}$$

(by using **(I1)**–**(I3)**, and the independence between ΔW_j and U_j.) Iterating this inequality gives a uniform in n bound for Δt sufficiently small:

$$\sup_{n \geq 0} \mathbb{E}^{\delta_x}\|U_n\|^2 < C[1 + \|x\|^2]. \tag{4.20}$$

We have $\bar{G}(x) = 1 + \|x\|^2$ and define

$$\mathcal{G}_0 = \{G : |G| \leq \bar{G} \;\&\; |G(x) - G(y)| \leq C_G[1 + \|x\| + \|y\|]\|x - y\|\}.$$

The previous estimate implies that $G \in \mathcal{G}_0$ are integrable with respect to $\mu_n^{\Delta t}$.

Theorem 4.5 *Consider the approximation of the continuous time Markov process* (4.15) *by the discrete time Markov chain* (4.18)*, under* **(I1)–(I4)***. Then, for any function $G \in \mathcal{G}_0$, there is $K > 0$, $\gamma \in (0,1)$ and for any $\Delta t \in (0, \Delta t_c)$ an integer $N = N(\Delta t)$ such that*

$$|\mathbb{E}^{\delta_x} G(U_n) - \pi(G)| \leq 2K \max\{\pi(\bar{G}), \bar{G}(x)\}\Delta t^{\gamma/2} \quad \forall n \geq N. \tag{4.21}$$

Furthermore the Markov chain (4.18) *has an invariant probability measure $\pi^{\Delta t}$ and*

$$|\pi^{\Delta t}(G) - \pi(G)| \leq 6K\pi(\bar{G})\Delta t^{\gamma/2}. \tag{4.22}$$

Proof By **(I4)**, **(ACi)** holds for the given definition of $\mathcal{G}_0$; indeed it holds for all $G \in \mathcal{G}$, which is strictly bigger than $\mathcal{G}_0$. Now let $G \in \mathcal{G}_0$. Then

$$\begin{array}{rc} & |\mathbb{E}^{\delta_x} G(u(n\Delta t)) - \mathbb{E}^{\delta_x} G(U_n)| \\ \leq & \mathbb{E}^{\delta_x}\{C_G[1 + \|u(n\Delta t)\| + \|U_n\|]\|u(n\Delta t) - U_n\|\} \\ \leq & C_G\{\mathbb{E}^{\delta_x}[1 + \|u(n\Delta t)\| + \|U_n\|]^2 \mathbb{E}^{\delta_x}\|u(n\Delta t) - U_n\|^2\}^{\frac{1}{2}}. \end{array}$$

By (4.16), (4.19) and (4.20) we deduce that

$$\sup_{G \in \mathcal{G}_0} |\mathbb{E}^{\delta_x} G(u(n\Delta t)) - \mathbb{E}^{\delta_x} G(U_n)| \leq C\bar{G}(x)\Delta t^{\frac{1}{2}} e^{\alpha n \Delta t}, \qquad \forall n\Delta t \geq 0.$$

Thus **(ACii)** also holds, yielding (4.21) by Theorem 3.3. Corollary 3.4 applies, because we are working in finite dimensions and because the weak Feller property holds for (4.18). This theorem completes the proof. $\square$

4.3 Parabolic Stochastic Partial Differential Equations

Consider the following stochastic PDE on $L_2(0,1)$:

$$du = \Big[- Au + f(u)\Big] dt + dW(t), \quad u(0) = x, \tag{4.23}$$

where $A = -\Delta$, the Laplacian with domain $H^2(0,1) \cap H_0^1(0,1)$, f is globally Lipschitz from $L_2(0,1)$ to itself, $W(\cdot)$ is an $L_2(0,1)$ valued Wiener process with covariance Q, and dW is interpreted as an Itô integral. If Q has a complete set of eigenvectors e_i, corresponding eigenvalues λ_i, and $\beta_i(\cdot)$ is a set of IID Brownian motions, then the process $W(t)$ may be written

$$W(t) = \sum_{i=1}^{\infty} \lambda_i^{1/2}\, e_i\, \beta_i(t). \tag{4.24}$$

Questions of existence and regularity of this equation are discussed by Da Prato–Zabczyk [2]. The hypotheses gathered here assure existence of a unique mild solution for bounded Q.

It is natural to ask when (4.23) is geometrically ergodic. The only published work on geometric ergodicity for stochastic PDEs known to the authors is [3] (although some related work may be found in [6]), in which **(ACi)** is obtained for bounded Lipschitz functions $\bar{G}$ with given Lipschitz constant and subject to the contractivity assumption: $\exists \omega > 0$ such that

$$\langle -A(u-v) + f(u) - f(v), u - v\rangle_{L_2(0,1)} \le -\omega \|u-v\|^2_{L_2(0,1)}, \quad u, v \in L_2(0,1). \tag{4.25}$$

This geometric ergodicity can be easily extended to globally Lipschitz test functionals when $\mathbb{E}^{\delta_x} \|u(t)\|^2_{L_2(0,1)} < \infty$ – see [23]. However the contractivity assumption is highly restrictive. The work of Shardlow [24] proves geometrically ergodic for space of test functionals dominated by $1 + \|\cdot\|_{L_2(0,1)}$ subject to the covariance operator Q being non-singular and the condition

$$\exists \alpha, \beta > 0 : \quad \langle -Au + f(u), u\rangle \le \alpha - \beta \|u\|^2_{L_2(0,1)}.$$

Motivated by the geometric ergodicity proved in [3] and extended in [23]–[24], we employ the following assumption:

- **(Si)** The continuous time Markov process generated by (4.23) is geometrically ergodic: given $\bar{G}(u) := 1 + \|u\|_{L_2(0,1)}$ define $\mathcal{G} = \{G : L^2(0,1) \to \mathbb{R}^+, |G| \le \bar{G}\}$ and $\mathcal{G}_0 = \{G \in \mathcal{G} : Lip\{G\} \le C_G\}$. Then $\exists R > 0, \gamma > 0 :$

$$\sup_{G \in \mathcal{G}_0} |\mathbb{E}^{\delta_x} G(u(t)) - \pi(G)| \le R\bar{G}(x) e^{-\gamma t} \quad \forall t \ge 0,$$

where π is the invariant measure on $(L_2(0,1), \mathbb{B}(L_2(0,1)))$ of (4.23).

We consider the numerical approximation of (4.23) by finite differences, in particular by the θ method in time, the standard 3 point approximation in space, and a spectral approximation to W (truncation of (4.24)). For fixed $J \in \mathbb{N}$, this method yields grid functions U_{nj} $j = 1, \ldots, J$ and $n \in \mathbb{N}$ where U_{nj} approximates $u(n\Delta t, j\Delta x)$ ($\Delta x = 1/(J+1)$ is the grid size, Δt the time step). We prefer to work with continuous interpolants u_n in $L_2(0,1)$ of U_{nj}. To do this, define $A_{\Delta x}: L_2(0,1) \to L_2(0,1)$ by

$$A_{\Delta x} \sin(k\pi\cdot) = \lambda_k \sin(k\pi\cdot), \qquad k = 1, 2, \ldots$$

where $\lambda_{k+nJ} = 4\sin^2(k\pi\Delta x/2)/\Delta x^2$ ($k = 1, \ldots, J$, $n = 1, \ldots$). This operator acts exactly as the standard three point approximation to the Laplacian on the grid $\{j\Delta x: j = 1, \ldots, J\}$. We study the following Markov chain

$$\begin{aligned} u_{n+1} - u_n = \quad & \Big[(1-\theta)A_{\Delta x}u_n + \theta A_{\Delta x}u_{n+1} \\ & +(1-\theta)f(u_n) + \theta f(u_{n+1})\Big]\Delta t + dW_n, \end{aligned} \tag{4.26}$$

where

$$dW_n = \int_{n\Delta t}^{(n+1)\Delta t} \mathbb{P}_J dW(t);$$

$\mathbb{P}_J$ is the projection of $L_2(0,1)$ onto the first J eigenfunctions of A. This Markov chain, which agrees with the finite difference scheme outlined above on the grid $\{j\Delta x : j = 1, \ldots, J\}$, is discussed in [23].

Other finite difference approximations to (4.23) are studied in the literature. For example, Gyöngy [11] uses an approximation to the white noise based on integrating over boxes of size $\Delta x \times \Delta t$. He works in a more general setting than we do here (with multiplicative noise and very mild measurability hypotheses on f) and obtains similar results to those below in Lemma 4.6; however the results in [11] are not directly relevant to the present perturbation theory for invariant measures because the dependence on the L_2 norm of the initial data and the growth of his estimate in large time are not described. The estimates of [23] given below for the method (4.26) depend on stronger hypothesis on f, but are L_2 estimates for L_2 initial data, which, with the geometric ergodicity results available for (4.23), enable us to apply Theorems 3.3.

The assumptions we make about f are as follows:

- **(Sii)** There exist constants K_1, K_2 such that

$$\begin{aligned}
\|f(x)-f(y)\|_{L_2(0,1)} &\le K_1\|x-y\|_{L_2(0,1)}, \quad x,y\in L_2(0,1);\\
\|f(x)\|_{L_2(0,1)} &\le K_2+K_1\|x\|_{L_2(0,1)}, \quad x\in L_2(0,1)\\
\|A^{1/4}f(x)\|_{L_2(0,1)} &\le K_2+K_1\|A^{1/4}x\|_{L_2(0,1)}, \quad x\in L_2(0,1).
\end{aligned}$$

Using **(Sii)**, **(ACii)** may be established for certain approximations of stochastic PDEs. The following is proved by Shardlow [23].

Lemma 4.6 *Let hypothesis* **(Sii)** *hold and consider approximation of* (4.23) *by* (4.26)*. Consider initial data $x \in L_2(0,1)$ for the problem* (4.23) *and let $(\Delta x, \Delta t) \to 0$ subject to the stability condition*

$$\frac{\Delta t}{\Delta x^2}4(1-\theta) \le 1. \tag{4.27}$$

For $0 < \delta < 1$, there exists $C, \alpha, \Delta t_c > 0$ such that, for $n\Delta t > 0$ and $0 < \Delta t < \Delta t_c$,

$$\left(\mathbb{E}^{\delta_x}\|u(n\Delta t)-u_n\|^2_{L_2(0,1)}\right)^{1/2} \le C\Delta x^{(1-2\delta)/2}e^{\alpha n\Delta t}(1+\|x\|_{L_2(0,1)})(1+(n\Delta t)^{(\delta-1)/4}).$$

In particular, for any $G \in \mathcal{G}_0$ the following holds for $n\Delta t > 1$,

$$|\mathbb{E}^{\delta_x}G(u(n\Delta t)) - \mathbb{E}^{\delta_x}G(u_n)| \le 2CC_G\Delta x^{(1-2\delta)/2}e^{\alpha n\Delta t}(1+\|x\|_{L_2(0,1)}).$$

The following result is now readily proved by application of Theorem 3.3. The result is weaker than that proved in §4.1–4.2, as we are unable to establish existence of an invariant measure for the numerical approximation. The authors expect this result would come from a uniform in time bound on the numerical solution in a stronger norm ($H^1_0(0,1)$, for example). However, we do not investigate this issue further.

Theorem 4.7 *Consider the approximation of the Markov process* (4.23) *by the Markov chain* (4.26)*, under* **(Si)**–**(Sii)***. Then, for any $G \in \mathcal{G}_0$, there is $C > 0$, $\gamma \in (0,1)$ and for any $\Delta t \in [0, \Delta t_c)$ an integer $N = N(\Delta t)$ such that*

$$|\mathbb{E}^{\delta_x}G(u_n) - \pi(G)| \le 2C\max\{\pi(\bar{G}), \bar{G}(x)\}\Delta x^{\gamma/2} \quad \forall n \ge N. \tag{4.28}$$

Furthermore, if the Markov chain (4.26) *has an invariant probability measure* $\pi^{\Delta x}$*, then*

$$|\pi^{\Delta x}(G) - \pi(G)| \leq 6C\pi(\bar{G})\Delta x^{\gamma/2}. \tag{4.29}$$

Acknowledgements. We thank Peter Baxendale and Peter Glynn for helpful discussions.

References

[1] G. Blankenship and G.C. Papanicolaou, *Stability and control of stochastic systems with wide-band noise disturbances. I.* SIAM J. Appl. Math. **34**(1978), 437–476.

[2] G. Da Prato and J. Zabczyk, *Stochastic Equations in Infinite Dimensions.*(1992) Cambridge University Press.

[3] G. Da Prato and J. Zabczyk, *Ergodicity for Infinite Dimensional Systems.*(1996) Cambridge University Press.

[4] A. M. Davie and J. G. Gaines, *Convergence of implicit schemes for numerical solution of parabolic stochastic partial differential equations*, In Progress.

[5] M.H.A. Davis, *Markov Models and Optimization*, Chapman and Hall, 1993.

[6] W. E, K. Khanin, A. Mazel and Y. Sinai, *Probability distribution functions for the Random forced Burgers equation.* Phys. Rev. Lett. **78**(1997), 1904–1907.

[7] P.W. Glynn and P. L'Ecuyer, *Likelihood ratio gradient estimation for stochastic recursions.* Adv. Appl. Prob. **27**(1995), 1019–1053.

[8] W. Grecksch and P. Kloeden, *Time–Discretized Galerkin Approximations of Parabolic Stochastic PDEs*, Bulletin of Australian Mathematics Society, (1996) 54. pp. 79–85.

[9] N. Gronbech-Jensen and S. Doniach, *Long-time overdamped dynamics of molecular chains.* J. Comp. Chem. **15**(1994), 997–1012.

[10] A. Grorud and D. Talay, *Approximation of Lyapunov exponents of nonlinear stochastic differential equations.* SIAM J. Appl. Math. **56**(1996), 627–650.

[11] I. Gyöngy, Lattice Approximations for stochastic quasi–linear parabolic partial differential equations driven by space–time white noise II, preprint no. 32/1995, University Budapest.

[12] R.Z. Hasminskii, *Stochastic stability of differential equations.* Sijthoff and Noordhoff, Rockville, 1980.

[13] N.V. Kartashov, *Criteria for uniform ergodicity and strong stability of Markov chains with a common phase space.* Theor. Prob. and Math. Stat. **30**(1985), 71–89.

[14] N.V. Kartashov, *Inequalities in theorems of ergodicity and stability for Markov chains with a common phase space, parts I and II.* Theo. Prob. and its Appl. **30**(1986), 247–259 and 505–515.

[15] N.V. Kartashov, *Strong Stable Markov Chains.* VSP, Utrecht, Netherlands, 1996.

[16] P.E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations.* Springer-Verlag, New York, 1991.

[17] H.J. Kushner, *Approximation and weak convergence methods for random processes.* MIT Press, Cambridge, 1984.

[18] T.M. Liggett, *Interacting Particle Systems.* Springer-Verlag, New York, 1985.

[19] B. Mishra and T. Schlick, *The notion of error in Langevin dynamics.* Preprint, 1996.

[20] S.P. Meyn and R.L. Tweedie, *Markov Chains and Stochastic Stability.* Second Edition, Springer-Verlag, London, 1996.

[21] S.P. Meyn and R.L. Tweedie, *Stability of Markovian Processes, I, II and III* Adv. Appl. Prob. **24**(1992), 542–574, **25**(1993), 487–517 and **25**(1933), 518–548.

[22] J.M. Sanz-Serna and A.M. Stuart, *Ergodic Properties of Dissipative Differential Equations Subject to Random Impulses*, (1997), submitted.

[23] T. Shardlow, *Numerical Methods for Parabolic Dissipative SPDEs*, In Progress. *Topics in Dissipative Evolution Equations*, (1997), Ph.D. Thesis, Stanford University.

[24] T. Shardlow, *Geometric Ergodicity for stochastic PDEs*, In Progress. (1998).

[25] A.M. Stuart and A.R. Humphries, *Dynamical Systems and Numerical Analysis*, (1996), Cambridge University Press.

[26] D. Talay, *Second-order discretization schemes for stochastic differential systems for the computation of the invariant law.* Stochastics and Stochastics Reports **29**(1990), 13–36

[27] D. Talay, *Approximation of upper Lyapunov exponents of bilinear stochastic differential systems.* SIAM J. Num. Anal. **28**(1991), 1141–1164.

[28] M.M. Tropper, *Ergodic properties and quasideterministic properties of finite-dimensional stochastic systems.* J. Stat. Phys. **17**(1977), 491–509.