A Lyapunov Bound for Solutions of Poisson's Equation

Peter W. Glynn
Department of Operations Research
Stanford University,
Stanford, CA 94305-4022

Sean P. Meyn
University of Illinois
and the Coordinated Science Laboratory
1308 W. Main St., Urbana, IL 61801

Abstract

In this paper we consider ψ-irreducible Markov processes evolving in discrete or continuous time, on a general state space. We develop a Lyapunov function criterion that permits one to obtain explicit bounds on the solution to Poisson's equation and, in particular, obtain conditions under which the solution is square integrable.

These results are applied to obtain sufficient conditions that guarantee the validity of a functional central limit theorem for the Markov process. As a second consequence of the bounds obtained, a perturbation theory for Markov processes is developed which gives conditions under which both the solution to Poisson's equation and the invariant probability for the process are continuous functions of its transition kernel. The techniques are illustrated with applications to queueing theory and autoregressive processes.

AMS subject classifications: 68M20, 60J10
Running head: Poisson's Equation
Keywords: Markov chain, Markov process, Poisson's equation, Lyapunov function, Foster's criterion, functional central limit theorem, perturbation theory.

1 Introduction

In this paper we develop a "Foster's criterion", or "Lyapunov function" approach to obtaining finite-valued solutions $\hat{g}$ to *Poisson's equation*, which in discrete time may be written

$$\overline{g} = \hat{g} - P\hat{g}, \tag{1}$$

where g is a given real-valued function on the state space, $\overline{g} = g - \pi(g)$, and π is an invariant probability. In the special case where $g \equiv 0$, solutions to Poisson's equation are precisely *harmonic functions.* In general, if $\hat{g}_1$ and $\hat{g}_2$ are two solutions to Poisson's equation then the difference $\hat{g}_1 - \hat{g}_2$ is harmonic.

For continuous time processes, Poisson's equation becomes

$$\overline{g} = -\widetilde{\mathcal{A}}\hat{g} \tag{2}$$

where $\widetilde{\mathcal{A}}$ is the extended generator of the Markov process $\boldsymbol{\Phi}$, formally defined in equations (11) and (12) below.

The Poisson equation and the general potential theory of positive kernels is developed in the seminal work of Neveu [28], Revuz [32] and Constantinescu and Cornea [7]. The reader is referred to Nummelin [30] for some of the most current results on Poisson's equation, to whom we owe much.

The solution $\hat{g}$ to Poisson's equation (1) is fundamental to the analysis of the additive functional

$$S_n = \sum_{k=0}^{n-1} \overline{g}(\Phi_k); \qquad S_0 = 0, \tag{3}$$

and it is equally valuable in studying the analogous additive functional in continuous time. The principal observation that underlies the analysis of S_n is that the behavior of such an additive functional is closely related to that of a certain martingale. Specifically, let

$$M_n = \hat{g}(\Phi_n) + S_n, \qquad n \in \mathbb{Z}_+. \tag{4}$$

We note that since $\hat{g}$ solves (1),

$$\begin{aligned} M_n &= \hat{g}(x) + \sum_{k=1}^{n} [\hat{g}(\Phi_k) - P\hat{g}\,(\Phi_{k-1})] \\ &= \hat{g}(x) + \sum_{k=1}^{n} [\hat{g}(\Phi_k) - \mathsf{E}(\hat{g}(\Phi_k) \mid \mathcal{F}_{k-1})] \end{aligned}$$

where $\mathcal{F}_n = \sigma(\Phi_0, \ldots, \Phi_n)$. Hence, under suitable integrability conditions on $\hat{g}$, the adapted process $\mathbf{M} = \{(M_n, \mathcal{F}_n) : n \geq 1\}$ is a square integrable martingale.

It is evident that the asymptotic behavior of $(S_n : n \geq 0)$ will typically mimic that of the martingale $\mathbf{M}$. In particular, laws of large numbers, central limit theorems, and laws of the iterated logarithm can often be derived for $(S_n : n \geq 0)$ by applying appropriate martingale theorems. This approach is taken in Maigret [17], and in the text Duflo [10] to obtain a functional central limit theorem (FCLT) for Markov chains. In related work, Kurtz [15] considers chains arising in models found in polymer chemistry. Bhattacharaya [5] considers Poisson's equation in continuous time to derive a FCLT and functional law of the iterated logarithm for ergodic Markov processes.

Using a more classical approach based upon the existence of atoms for a split chain, Meyn and Tweedie [23] derive a central limit theorem and law of the iterated logorithm for geometrically ergodic chains. The Lyapunov function approach of [23] is similar to that of the present paper, but the results reported here require far milder assumptions than geometric ergodicity.

In order to successfully apply the Poisson equation to obtain a FCLT for a given chain, one must basically show that the martingale difference sequence $\Delta_n = \hat{g}(\Phi_n) - (P\hat{g})(\Phi_{n-1})$ is appropriately square integrable. The Lyapunov function criterion developed in this paper permits one to verify that the solution $\hat{g} \in L^2(\pi)$, where π is the (unique) invariant measure of $\boldsymbol{\Phi}$. This guarantees that the martingale differences are well-behaved, so that the above mentioned FCLT's apply.

Before proceeding, we note that the solution $\hat{g}$ is often unique, in a certain sense. The reader is referred to Shwartz and Makowski [36] for further results in a discrete time – discrete space setting.

Proposition 1.1 *Suppose that $\boldsymbol{\Phi}$ is an ergodic Markov chain with unique invariant probability π, with discrete or continuous time parameter, and suppose that $\hat{g}$ and $\hat{g}_\bullet$ are two solutions to Poisson's equation with $\pi(|\hat{g}| + |\hat{g}_\bullet|) < \infty$. Then for some constant c, $\hat{g}(x) = c + \hat{g}_\bullet(x)$ for a.e. $x \in \mathsf{X}$ $[\pi]$.*

Proof The proofs are similar, so we will consider only the continuous time case. We have already remarked that $h := \hat{g} - \hat{g}_\bullet$ is harmonic.

We apply the ergodic theorem for Markov chains: to do so we consider the skeleton chain Φ_n, and the function $H = \int_0^1 P^s h\, ds$. From Poisson's equation we have that $P^t h = h$ for all t. From the Ergodic Theorem

$$n^{-1} \int_0^n P^t h\, dt = n^{-1} \sum_{k=0}^{n-1} (P^k H)(x) \to \pi(h),$$

a.e. $[\pi]$ (cf. [11]). Hence $h = \pi(h)$ a.e. $[\pi]$. □

2 Discrete Time Processes

We consider in this paper a Markov process $\boldsymbol{\Phi} = \{\Phi_t : t \in T\}$ where $T = \mathbb{R}_+$ or $\mathbb{Z}_+$, evolving on a locally compact separable metric space X, whose Borel σ-algebra shall be denoted $\mathcal{B}(\mathsf{X})$. We use P_μ and E_μ to denote probabilities and expectations conditional on Φ_0 having distribution μ, and P_x and E_x when μ is concentrated at x. In the discrete time setting the conditions on the state space can be relaxed somewhat, but it is convenient here to have the same set of assumptions for continuous and discrete time processes.

In this section we will develop our main results for discrete time processes. in Section 3 we show how the results may be extended to the continuous time case.

A set $\boldsymbol{\alpha} \in \mathcal{B}^+(\mathsf{X})$ is called an *atom* if transitions from distinct points in $\boldsymbol{\alpha}$ are identical: $P(x,\,\cdot\,) = P(y,\,\cdot\,)$, $x, y \in \boldsymbol{\alpha}$. When an atom exists, a solution to Poisson's equation (1) can easily be found: Define the function $\hat{g}$ by

$$\hat{g}(x) = \mathsf{E}_x\Big[\sum_{k=0}^{\sigma_\alpha} \overline{g}(\Phi_k)\Big] = \sum_{k=0}^{\infty} (I_{\alpha^c} P)^k \overline{g} \qquad (5)$$

when this is well defined. When $\boldsymbol{\Phi}$ is positive Harris [25] and $\pi(\boldsymbol{\alpha}) > 0$, we have

$$\int_\alpha \pi(dx) \mathsf{E}_x\Big[\sum_{k=1}^{\tau_\alpha} |g(\Phi_k)|\Big] = \int_\alpha \pi(dx) \mathsf{E}_x\Big[\sum_{k=0}^{\tau_\alpha - 1} |g(\Phi_k)|\Big] = \pi(|g|),$$

and hence when $\pi(|g|) < \infty$, the expression (5) is well defined on $\boldsymbol{\alpha}$, and in fact $\hat{g}$ is defined and finite a.e. $[\pi]$. It follows from the second equality in (5) that $\hat{g}$ solves Poisson's equation (1).

Even when an atom does not exist, by considering a *split chain*, one can construct an atom $\check{\alpha}$ on the split state space. Define then the kernel $G_{s,\nu}$ for functions f and states x by

$$G_{s,\nu}(x,f) = \check{\mathsf{E}}_{\delta_x^*}\Big[\sum_{k=0}^{\sigma_{\check{\alpha}}} f(\check{\Phi}_k)\Big]. \tag{6}$$

Then the function $\hat{g}(x) = G_{s,\nu}(x,\overline{g})$ solves Poisson's equation, and is finite a.e. if $\pi(|g|) < \infty$ (see [29, 30]).

The split chain is defined through a generalization of atoms known as *petite sets*. Let a be a probability distribution on $\mathbb{Z}_+$, and let K_a denote the Markov transition function $K_a = \sum_{i=0}^\infty a(i)P^i$. A set $C \subset \mathsf{X}$ is called ν_a-petite, where ν_a is a non-trivial measure on $\mathcal{B}(\mathsf{X})$, if for the distribution a on $\mathbb{Z}_+$,

$$K_a(x,A) \geq \nu_a(A), \qquad x \in C,\ A \in \mathcal{B}(\mathsf{X}).$$

The distribution a is called the *sampling distribution* for the petite set C. If the particular measure ν_a is unimportant, the prefix will be omitted so that C is simply called *petite*.

Much of the development to follow is concerned with verifying a recurrence condition for a Markov chain known as f-regularity. Let $f \geq 1$ be a real-valued function on X, and suppose that for some finite-valued function V_0, and any $B \in \mathcal{B}^+(\mathsf{X})$, there exists $c(B) < \infty$ such that

$$\mathsf{E}_x\Big[\sum_{k=0}^{\tau_B - 1} f(\Phi_k)\Big] \leq V_0(x) + c(B), \qquad x \in \mathsf{X}. \tag{7}$$

Then the chain is called *f-regular (with bounding function V_0)*. This definition is apparently stronger than similar notions of f-regularity given in [29] or [25], but we will find that all of these definitions are essentially equivalent.

This condition may be verified by establishing a drift property for the chain towards a single petite set.

For a function $f\colon \mathsf{X} \to [1,\infty)$, a petite set $C \in \mathcal{B}(\mathsf{X})$, a constant $b < \infty$, and a function $V\colon \mathsf{X} \to [0,\infty)$

$$PV(x) \leq V(x) - f(x) + b\mathbb{1}_C(x), \qquad x \in \mathsf{X}. \tag{8}$$

The power of (8) largely comes from the following

Theorem 2.1 (Comparison Theorem) *Suppose that the bound*

$$PV(x) \leq V(x) - f(x) + s(x) \qquad x \in \mathsf{X}$$

is satisfied. Then for each $x \in \mathsf{X}$, $N \in \mathbb{Z}_+$, and any stopping time τ we have

$$\begin{aligned} \sum_{k=0}^{N} \mathsf{E}_x[f(\Phi_k)] &\leq V(x) + \sum_{k=0}^{N} \mathsf{E}_x[s(\Phi_k)] \\ \mathsf{E}_x\Big[\sum_{k=0}^{\tau-1} f(\Phi_k)\Big] &\leq V(x) + \mathsf{E}_x\Big[\sum_{k=0}^{\tau-1} s(\Phi_k)\Big]. \end{aligned}$$

Proof By Dynkin's Formula (cf [23]) and the bound on PV we have

$$0 \le \mathsf{E}_x[V(\Phi_{\tau^n})] \le V(x) + \mathsf{E}_x\Big[\sum_{i=1}^{\tau^n}(s_{i-1}(\Phi_{i-1}) - [f_{i-1}(\Phi_{i-1}) \wedge N])\Big]$$

where $\tau^n = \min(\tau, n, \min(k : V(\Phi_k) \ge n))$. Hence by adding the finite term

$$\mathsf{E}_x\Big[\sum_{k=1}^{\tau^n}[f(\Phi_{k-1}) \wedge N]\Big]$$

to each side we get

$$\mathsf{E}_x\Big[\sum_{k=1}^{\tau^n}[f(\Phi_{k-1}) \wedge N]\Big] \le \mathsf{E}_x\Big[\sum_{k=1}^{\tau^n} s(\Phi_{k-1})\Big] \le \mathsf{E}_x\Big[\sum_{k=1}^{\tau} s(\Phi_{k-1})\Big].$$

Letting $n \to \infty$ and then $N \to \infty$ gives the result by the Monotone Convergence Theorem. □

We now characterize f-regularity using (8).

Theorem 2.2 *If (8) holds then*

(i) *The Markov chain $\boldsymbol{\Phi}$ is positive Harris recurrent with invariant probability π;*

(ii) *$\pi(f) < \infty$;*

(iii) *For any $B \in \mathcal{B}^+(\mathsf{X})$ there exists $c(B) < \infty$ such that*

$$\sum_{k=0}^{\infty}(PI_{B^c})^k f\,(x) = \mathsf{E}_x\Big[\sum_{k=0}^{\tau_B-1} f(\Phi_k)\Big] \le V(x) + c(B),$$

so that $\boldsymbol{\Phi}$ is f-regular with bounding function V.

Proof Results (i) and (ii) follow easily from (iii) and the structure of π in terms of mean occupancy times: See Theorem 10.0.1 of [25].

To prove (iii), suppose that (8) holds. By the Comparison Theorem 2.1, the strong Markov property, and the bound

$$\mathbb{1}_C(x) \le \psi_a(B)^{-1}K_a(x, B),$$

which follows from the fact that C is ψ_a-petite for some ψ_a, we have for any $B \in \mathcal{B}^+(\mathsf{X})$, $x \in \mathsf{X}$,

$$\begin{aligned}
\mathsf{E}_x\Big[\sum_{k=0}^{\tau_B-1} f(\Phi_k)\Big] &\le V(x) + b\mathsf{E}_x\Big[\sum_{k=0}^{\tau_B-1} \mathbb{1}_C(\Phi_k)\Big] \\
&\le V(x) + b\mathsf{E}_x\Big[\sum_{k=0}^{\tau_B-1} \psi_a(B)^{-1}K_a(\Phi_k, B)\Big] \\
&= V(x) + b\psi_a(B)^{-1}\sum_{i=0}^{\infty} a_i \mathsf{E}_x\Big[\sum_{k=0}^{\tau_B-1} P^i(\Phi_k, B)\Big] \\
&= V(x) + b\psi_a(B)^{-1}\sum_{i=0}^{\infty} a_i \mathsf{E}_x\Big[\sum_{k=0}^{\tau_B-1} \mathbb{1}_B(\Phi_{k+i})\Big] \\
&\le V(x) + b\psi_a(B)^{-1}\sum_{i=0}^{\infty} i a_i.
\end{aligned}$$

Since we can choose a so that $m_a = \sum_{i=0}^{\infty} i a_i < \infty$, result (iii) follows with $c(B) = b\psi_a(B)^{-1} m_a$. □

We may now present our main result which gives a specific bound on the *fundamental kernel* $Z := (I - P + \Pi)^{-1}$, where the kernel Π is defined as $\Pi(x, \cdot\,) \equiv \pi(\,\cdot\,)$. For any function g on X, the function $\hat{g} := Z\overline{g}$ is a solution to Poisson's equation whenever the inverse is well defined.

Let $h \geq 1$ be a finite-valued function on X, and let L_h^∞ denote the vector space of all measurable functions g on X such that $|g(x)|/h(x)$ is bounded in x. This vector space is a Banach space with the associated norm

$$|g|_h := \sup_{x \in \mathsf{X}} \frac{|g(x)|}{h(x)}.$$

We now give uniform bounds on solutions to Poisson's equation whenever a solution to (8) exists.

Theorem 2.3 *Suppose that $\boldsymbol{\Phi}$ is f-regular, so that (8) holds with V everywhere finite, $f \geq 1$, and C petite. Then the fundamental kernel Z is a bounded linear transformation from L_f^∞ to L_h^∞, with $h = V + 1$. That is, for some $c_0 < \infty$ and any $|g| \leq f$, Poisson's equation (1) admits a solution $\hat{g}$ satisfying the bound $|\hat{g}| \leq c_0(V + 1)$.*

Proof First consider the strongly aperiodic case. We can assume without loss of generality that the function V is bounded on the set S used in the definition of strong aperiodicity (for details, see [25, page 118].)

For any function h on X we can define a "split" function also denoted h on the split state space $\check{\mathsf{X}}$ which is identical to h on the two copies X_0 and X_1 of X. Similarly, we let $A \subset \check{\mathsf{X}}$ denote the "split set" $A = A_0 \cup A_1$.

Since V is bounded on S, it is straightforward to check that when (8) holds, we have the following bound for the split chain:

$$\check{P}V\,(x_i) \leq V(x_i) - f(x_i) + d\mathbb{1}_{C \cup S}(x_i), \qquad x_i \in \check{\mathsf{X}},$$

where d is a finite constant, and C is the petite set used in (8).

Hence we can apply Theorem 2.2 (iii) to the split chain to find that for some constant c

$$\check{\mathsf{E}}_{x_i}\Big[\sum_{k=0}^{\tau_{\check{\alpha}}-1} f(\check{\Phi}_k)\Big] \leq V(x_i) + c, \qquad x_i \in \check{\mathsf{X}}.$$

We then have the desired solution to Poisson's equation: For any $|g| \leq f$, we let $\hat{g}(x) = G_{s,\nu}(x, \overline{g})$, which is defined in (6). Then $\hat{g}$ solves Poisson's equation with $\hat{g}(x) \leq c(V(x) + 1)$, for a possibly larger constant c.

In the general case it is convenient to consider the K_{a_ε}-chain, which is always strongly aperiodic when $\boldsymbol{\Phi}$ is ψ-irreducible. We first show that the K_{a_ε}-chain satisfies a version of (8) with the same function f and a scaled version of the function V used in the theorem. We will on two occasions apply the identity

$$K_{a_\varepsilon} = \varepsilon K_{a_\varepsilon} P + (1 - \varepsilon) I, \tag{9}$$

whose derivation is straightforward. Hence by (8) for the kernel P,

$$K_{a_\varepsilon} V \leq \varepsilon K_{a_\varepsilon}(V - f + b\mathbb{1}_C) + (1 - \varepsilon)V.$$

Since $f \leq (1 - \varepsilon)^{-1} K_{a_\varepsilon} f$ it follows that with V_ε equal to a suitable constant multiple of V we have for some b',

$$K_{a_\varepsilon} V_\varepsilon \le V_\varepsilon - f + b' K_{a_\varepsilon} \mathbb{1}_C.$$

Since C is petite for $\boldsymbol{\Phi}$ and hence also for the K_{a_ε}-chain, the set $C_n := \{x : K_{a_\varepsilon}(x, C) \ge 1/n\}$ is petite for the K_{a_ε}-chain for all n. Note that $C \subseteq C_n$ for n sufficiently large. Scaling V_ε as necessary, we may choose n and b_ε so large that

$$K_{a_\varepsilon} V_\varepsilon \le V_\varepsilon - f + b_\varepsilon \mathbb{1}_{C_n}.$$

Thus the K_{a_ε}-chain is f-regular. By strong aperiodicity there exists a constant $c_\varepsilon < \infty$ such that for any $|g| \le f$, we have a solution $\hat{g}_\varepsilon$ to Poisson's equation

$$K_{a_\varepsilon} \hat{g}_\varepsilon = \hat{g}_\varepsilon - \overline{g}$$

satisfying $|\hat{g}_\varepsilon| \le V + c_\varepsilon$.

To complete the proof let $\hat{g} := \frac{\varepsilon}{1-\varepsilon} K_{a_\varepsilon} \hat{g}_\varepsilon = \frac{\varepsilon}{1-\varepsilon}(\hat{g}_\varepsilon - \overline{g})$. Writing (9) in the form

$$\frac{\varepsilon}{1-\varepsilon} P K_{a_\varepsilon} = \frac{1}{1-\varepsilon} K_{a_\varepsilon} - I,$$

we have by applying both sides to $\hat{g}_\varepsilon$,

$$P\hat{g} = \varepsilon^{-1}\hat{g} - \hat{g}_\varepsilon = \varepsilon^{-1}\hat{g} - (\varepsilon^{-1} - 1)\hat{g} - \overline{g} = \hat{g} - \overline{g}$$

so that Poisson's equation is satisfied. □

An important special case occurs when V is a constant multiple of f. In this case (8) may be written

$$PV \le \lambda V + b\mathbb{1}_C, \tag{10}$$

where $\lambda < 1$. Aperiodic chains for which (10) hold are called *V-uniformly ergodic* in [25]. When V is bounded from above and below, the inequality (10) is equivalent to uniform ergodicity as it is usually defined (see [32, 29, 25]).

By Theorem 2.3 we see that when (10) holds, the fundamental kernel $Z = (I - P + \Pi)^{-1}$ is a bounded linear transformation from L_V^∞ to itself. This provides another important consequence of V-uniform ergodicity which is especially valuable in analyzing perturbations of the chain. We will return to this after we consider the continuous time case.

3 Continuous Time Processes

We now show how the discrete time results developed thus far may be "lifted" through the resolvent chain to obtain analogous results for continuous time stochastic processes. We assume that $\boldsymbol{\Phi}$ is a Borel right process (cf. Sharpe [35]) so that, in particular, $\boldsymbol{\Phi}$ has the strong Markov property. We also assume that the escape time for the process is infinite. The reader is refered to [26, 27] for the relevant theory of Harris recurrence in continuous time.

For a measurable set A we let

$$\tau_A = \inf\{t \ge 0 : \Phi_t \in A\}, \qquad \eta_A = \int_0^\infty \mathbb{1}\{\Phi_t \in A\}\, dt.$$

The kernel K_a is defined exactly as in discrete time, where now a is a probability on $\mathbb{R}_+$. When a is an exponential distribution with unit mean, we let R denote the kernel K_a: this is the usual *resolvent* for the process.

We denote by $D(\widetilde{\mathcal{A}})$ the set of all functions $V\colon \mathsf{X}\times \mathbb{R}_+ \to \mathbb{R}$ for which there exists a measurable function $U\colon \mathsf{X}\to\mathbb{R}$ such that for each $x\in\mathsf{X}$, $t>0$,

$$P^t V\,(x) = V(x) + \int_0^t P^s U\,(x)\,ds \tag{11}$$

$$\int_0^t P^s|U|\,(x)\,ds < \infty. \tag{12}$$

We write $\widetilde{\mathcal{A}}V := U$ and call $\widetilde{\mathcal{A}}$ the *extended generator* of the process $\boldsymbol{\Phi}$. The function U is essentially unique (see Davis [8, p. 32] for a discussion).

Davis [8] requires only a local martingale property in his definition of the extended generator. If this weaker condition holds, then the results below are still valid, although some additional steps are required in the proofs. This more complicated situation may be treated as in [27].

We say that a function $h:\mathsf{X}\to\mathbb{R}$ is in the *domain of* R if $R(x,|h|)$ is finite for all $x\in\mathsf{X}$. From the definition (11) it is immediate that we have the identity

$$R\widetilde{\mathcal{A}}h = (R-I)h \tag{13}$$

for any h in the domain of R. The following result, which together with (13) states that R and $\widetilde{\mathcal{A}}$ commute, is central to obtaining solutions to the Poisson equation. For a proof see Down Meyn and Tweedie [9].

Lemma 3.1 *The extended generator satisfies the following identity for any h in the domain of R:*

$$\widetilde{\mathcal{A}}Rh = (R-I)h. \tag{14}$$

□

Equation (14) is a continuous time analogue of (9), which may be written

$$(P-I)\frac{\varepsilon}{1-\varepsilon}K_{a_\varepsilon} = K_{a_\varepsilon} - I,$$

and the identity (14) will be applied exactly as in the proof of Theorem 2.3 to obtain a solution to Poisson's equation in this continuous time context.

The following Foster-Lyapunov drift condition is taken from [27]. It is entirely analogous to (8), and will yield analogous results.

> *For a function $f\colon\mathsf{X}\to[1,\infty)$, a petite set $C\in\mathcal{B}(\mathsf{X})$, a constant $b<\infty$, and a function $V\colon\mathsf{X}\to[0,\infty)$*
>
> $$\widetilde{\mathcal{A}}V\,(x) \le -f(x) + b\mathbb{1}_C(x), \qquad x\in\mathsf{X}. \tag{15}$$

Theorem 3.2 *Suppose that $\boldsymbol{\Phi}$ is ψ-irreducible, and that (15) holds with V everywhere finite, $f\ge 1$, and C petite. Then $\boldsymbol{\Phi}$ is positive Harris recurrent with $\pi(f)<\infty$. For some $c_0<\infty$ and any $|g|\le f$, Poisson's equation (2) admits a solution $\hat g$ satisfying the bound $|\hat g| \le c_0(V+1)$.*

PROOF From (13) and the bound (15),

$$RV \le V - Rf + bR\mathbb{1}_C. \tag{16}$$

Since C is petite, the set $C_n = \{x : R\mathbb{1}_C(x) > 1/n\}$ is also petite, for any n. Hence, as in the proof of Theorem 2.3, we see that for a constant multiple V_0 of V, and some n, b' sufficiently large,

$$RV_0 \le V_0 - Rf + b'\mathbb{1}_{C_n}$$

Applying Theorem 2.2 (iii) we obtain the bound, for any $B \in \mathcal{B}^+(\mathsf{X})$,

$$\sum_{k=0}^{\infty}(RI_{C^c})^k Rf \le V_0 + c(B),$$

and adding f to both sides we obtain

$$\sum_{k=0}^{\infty}(RI_{C^c})^k f \le V_0 + f + c(B).$$

Hence the R-chain is f-regular with bounding function $V_0 + f$, and consequently $\pi(f) < \infty$.

From f-regularity of the resolvent chain, we may apply the discrete time result Theorem 2.3 to conclude that for some $c_1 < \infty$ and any $g \le f$, there exists a function $\hat{g}_1$ satisfying the bound $|\hat{g}_1| \le c_1(V + f + 1)$ such that

$$\hat{g}_1 - R\hat{g}_1 = \overline{g}.$$

We now let $\hat{g} = R\hat{g}_1$, and apply (16) to obtain the bound

$$|\hat{g}| \le R|\hat{g}_1| \le c_1(V + V + b + 1) \le c_0(V + 1)$$

where $c_0 = c_1(3 + b)$. Applying (14) we have that Poisson's equation is satisfied:

$$\widetilde{\mathcal{A}}\hat{g} = \widetilde{\mathcal{A}}R\hat{g}_1 = (R - I)\hat{g}_1 = -\overline{g},$$

which is the desired equality. □

4 Applications

In the remainder of the paper we describe several applications of our main results.

4.1 The Functional Central Limit Theorem

We first present several versions of the Functional Central Limit Theorem (FCLT) for Markov chains and processes; we begin with the discrete time case. For such models, the FCLT concerns the process obtained by interpolating the values of $\overline{g}(\Phi_k)$:

$$Z_n(t) = \frac{1}{\sqrt{n}}\Big(\sum_{k=0}^{[nt]}\overline{g}(\Phi_k)\Big), \qquad t \in \mathbb{R}_+.$$

Typically, in applications, the moment condition on V used in Theorem 4.1 will be obtained through an application of Theorem 2.2, which gives necessary and sufficient conditions for V^2-regularity, and hence a sufficient condition for finiteness of $\pi(V^2)$.

Theorem 4.1 *If the chain is f-regular with bounding function V, and if $\pi(V^2) < \infty$, then for any $|g| \le f$ there exists a constant $0 \le \gamma_g < \infty$ such that $Z_n \Rightarrow \gamma_g B$, P_μ-weakly, as $n \to \infty$ in $D[0,1]$ for any initial distribution μ.*

Furthermore, the constant γ_g^2 can be defined as $\gamma_g^2 = \pi(\hat{g}^2 - \{P\hat{g}\}^2)$, where $\hat{g}$ is the solution to Poisson's equation given in Theorem 2.3.

Proof This basically follows from Maigret [17]. Our theorem actually requires a slightly strengthened version of her result, because Theorem 4.1 asserts that the weak convergence holds for any initial distribution μ.

This extension is given in the continuous time case, in the proof of Theorem 4.3 below. To avoid serious repetition, we omit the proof of the discrete time result. □

The generalization of the FCLT to arbitrary initial conditions can also be carried out by applying the shift coupling property of positive Harris recurrent Markov processes. This is established for discrete time processes in Aldous and Thorrison [1].

An important special case occurs when (10) holds, in which case we obtain a simpler criterion for the FCLT. Taking square roots in (10) we obtain by Jensen's inequality

$$PV^{1/2} \le \lambda^{1/2} V^{1/2} + b^{1/2} \mathbb{1}_C .$$

Since by (10) we have that $\pi(V) < \infty$, the following corollary to Theorem 4.1 is immediate.

Theorem 4.2 *If (10) holds for some $V \ge 1$, $\lambda < 1$ and some petite set C then for any $g^2 \le V$ there exists a constant $0 \le \gamma_g < \infty$ such that $Z_n \Rightarrow \gamma_g B$, P_μ-weakly, as $n \to \infty$ in $D[0,1]$ for any initial distribution μ.*

The constant γ_g^2 can be defined as $\gamma_g^2 = \pi(\hat{g}^2 - \{P\hat{g}\}^2)$, where $\hat{g}$ is the solution to Poisson's equation given in Theorem 2.3. □

The analogous continuous time results may be obtained by applying the main result of Bhattacharaya [5]. We now consider the sequence of stochastic processes $\{\mathbf{Z}_n\}$ defined for each n by

$$Z_n(t) = \frac{1}{\sqrt{n}} \Big(\int_0^{nt} \overline{g}(\Phi_s)\, ds \Big), \qquad t \in \mathbb{R}_+ .$$

Theorem 4.3 *If (15) holds and if $\pi(V^2) < \infty$ then for any $|g| \le f$ there exists a constant $0 \le \gamma_g < \infty$ such that $Z_n \Rightarrow \gamma_g B$, P_μ-weakly, as $n \to \infty$ in $D[0,1]$ for any initial distribution μ.*

Furthermore, the constant γ_g^2 can be defined as $\gamma_g^2 = 2 \int \hat{g}(x)\overline{g}(x)\, \pi(dx)$, where $\hat{g}$ is the solution to Poisson's equation given in Theorem 2.3.

Proof In the special case where $\boldsymbol{\Phi}$ is stationary, or when a skeleton of the process is Harris ergodic, this follows from Theorem 2.1 and Theorem 2.6 of [5] together with Theorem 3.2. Note that we do not know if $\pi(g^2)$ is finite, a condition of Theorem 2.1 of [5]. However, by (16) and square integrability of V we do know that $\pi((R|g|)^2) < \infty$, and this is in fact enough to obtain the FCLT using the proof of his Theorem 2.1.

We now show how this result can be generalized to arbitrary initial conditions $\Phi_0 = x \in \mathsf{X}$; the result for an arbitrary initial distribution μ follows upon integrating over all initial conditions with respect to μ.

Recall that from Theorem 3.2, the process $\boldsymbol{\Phi}$ is positive Harris recurrent under the assumptions of Theorem 4.3. To show that the FCLT holds for arbitrary initial conditions under the assumptions of Theorem 4.3 we will apply the following two well known consequences of Harris recurrence: First of all, the Law of Large Numbers holds for any initial condition $x \in \mathsf{X}$, and any positive random variable H on sample space,

$$\frac{1}{T}\int_0^T \theta^s H\, ds \to \mathsf{E}_\pi[H] \qquad \text{a.s. } [\mathsf{P}_x]; \tag{17}$$

Secondly, the resolvent is Harris ergodic:

$$\|R^n(x,\,\cdot\,) - \pi\| \to 0, \qquad x \in \mathsf{X}. \tag{18}$$

This is a consequence of the fact that not only is the resolvent chain with transition function $R = \int e^{-t}P^t\,dt$ positive Harris recurrent when $\boldsymbol{\Phi}$ has this property, but the resolvent chain is necessarily also aperiodic since the measures $R(x,\,\cdot\,)$ and $\sum_{k=1}^\infty 2^{-k}R^k(x,\,\cdot\,)$ are equivalent for each x.

For any r and any n we let

$$Z_{n,r}(t) := \theta^r Z_n(t) = \frac{1}{\sqrt{n}}\Big(\int_r^{nt+r} \overline{g}(\Phi_s)\,ds\Big), \qquad t \in \mathbb{R}_+.$$

Using the Law of Large Numbers (17) one may show that

$$\sup_{0\le t\le 1} |Z_n(t) - Z_{n,r}(t)| \to 0, \qquad n \to \infty \text{ a.s.} \tag{19}$$

Letting ϕ be a bounded, continuous linear functional on $C[0,1]$, it follows from (19) that we have

$$|\mathsf{E}_x[\phi(Z_n(t))] - \mathsf{E}_x[\phi(Z_{n,r}(t))]| \to 0 \qquad n \to \infty.$$

By the Markov property this limit may be expressed

$$|\mathsf{E}_x[\phi(Z_n(t))] - \int P^r(x,dy)\mathsf{E}_y[\phi(Z_n(t))]| \to 0 \qquad n \to \infty.$$

Integrating over $r \ge 0$, it follows by dominated convergence that for any $m \ge 1$,

$$|\mathsf{E}_x[\phi(Z_n(t))] - \int R^m(x,dy)\mathsf{E}_y[\phi(Z_n(t))]| \to 0 \qquad n \to \infty,$$

where R denotes the resolvent. We can now apply (18): Let $\varepsilon > 0$, and choose m so large that $\|R^m - \pi\| < \varepsilon$. Then the limit above with this m implies that

$$\limsup_{n\to\infty} |\mathsf{E}_x[\phi(Z_n(t))] - \mathsf{E}_\pi[\phi(Z_n(t))]| \le \varepsilon|\phi|_\infty.$$

Since by Theorem 3.2 and Theorem 2.1 of [5] the FCLT holds when $\Phi \sim \pi$, this establishes the FCLT when $\Phi_0 = x$. □

When the exponential drift

$$\widetilde{\mathcal{A}}V \le -cV + b\mathbb{1}_C \tag{20}$$

holds for some $V \ge 1$, $c > 0$, then we obtain a simpler criterion for the FCLT, just as in the discrete time case.

We can use (16) to obtain, for some $\lambda < 1$, $b < \infty$ and a petite set C, the bound

$$RV \le \lambda V + b\mathbb{1}_C.$$

We then have by Jensen's inequality, $RV^{1/2} \le \lambda^{1/2}V^{1/2} + b^{1/2}\mathbb{1}_C$, and then following the proof of Theorem 3.2 we have, for some constant c_0, that Poisson's equation (2) can be solved for any g satisfying $g^2 \le V$, and the solution satisfies the bound $\hat{g}^2 \le c_0 V$. We immediately obtain from Theorem 4.3

Theorem 4.4 *If (20) holds then for any $g^2 \le V$ there exists a constant $0 \le \gamma_g < \infty$ such that $Z_n \Rightarrow \gamma_g B$, P_μ-weakly, as $n \to \infty$ in $D[0,1]$ for any initial distribution μ.*

The constant γ_g^2 can be defined as $\gamma_g^2 = 2\int \hat{g}(x)\overline{g}(x)\,\pi(dx)$, where $\hat{g}$ is the solution to Poisson's equation. □

In Section 4.3 we consider several models where moment conditions on the disturbance process may be given explicitly to ensure that (8) or (20) holds so that we can establish the FCLT.

4.2 Perturbations of Markov Processes

Smoothness of solutions to Poisson's equation is frequently assumed in applications to averaging and diffusion approximations, and in particular to establish the convergence of adaptive estimation algorithms of the stochastic approximation type (see [4, 19]), and well behaved solutions are required in the theory of Markov decision processes [18, 33]. An extensive bibliography of applications may be found in [36].

Suppose that $\{P_\theta : \theta \in \Theta\}$ is a family of Markov transition functions, where Θ denotes some open subset of Euclidean space. Assume that each of the corresponding Markov chains is ψ_θ-irreducible, and that for some $\theta_0 \in \Theta$, the chain with transition function P_{θ_0} satisfies the drift criterion (10). When the P_{θ_0}-chain is aperiodic, this means that the chain is V-uniformly ergodic.

We assume that $P_\theta \to P_{\theta_0}$ as $\theta \to \theta_0$ in the induced operator norm $\|\!| \cdot |\!\|_V$, defined as

$$\|\!| P_{\theta_0} - \Pi |\!\|_V := \sup_{\substack{h \in L_V^\infty \\ |h|_V = 1}} |(P_{\theta_0} - P_\theta)h|_V$$

Since each of the kernels is ψ_θ-irreducible, it may be shown that the drift criterion (10) holds, and that the set C used in the drift criterion is petite for the kernel P_θ for each θ in some open ball containing θ_0. Assume that this is the case for all $\theta \in \Theta$, and let $\{Z_\theta\}$ denote the corresponding collection of fundamental kernels and $\{\pi_\theta\}$ the associated invariant probabilities. From Theorem 2.3 we know that each of the kernels $\{Z_\theta, P_\theta, \Pi_\theta : \theta \in \Theta\}$ is a bounded linear transformation from L_V^∞ to itself.

Following [34] we define

$$U_{\theta_0,\theta} = [P_\theta - P_{\theta_0}]Z_{\theta_0} \tag{21}$$

$$H_{\theta_0,\theta} = [I - U_{\theta_0,\theta}]^{-1} \tag{22}$$

The first kernel is well defined since each of the kernels on the right hand side of the defining equation map L_V^∞ to itself. The inverse in the definition of $H_{\theta_0,\theta}$ is well defined for θ sufficiently close to θ_0.

A straightforward generalization of Theorem 2 of [34] shows that

$$\pi_\theta = \pi_{\theta_0} H_{\theta_0,\theta} \tag{23}$$

$$Z_\theta = Z_{\theta_0} H_{\theta_0,\theta} - \Pi_{\theta_0} H_{\theta_0,\theta} U_{\theta_0,\theta} Z_{\theta_0} H_{\theta_0,\theta} \tag{24}$$

where Π_{θ_0} is the Markov transition function defined as $\Pi_{\theta_0}(x, A) = \pi_{\theta_0}(A)$, $x \in \mathsf{X}$, $A \in \mathcal{B}(\mathsf{X})$. Hence the invariant probabilities converge in the V-total variation norm, and whenever $g \in L_V^\infty$, the solutions $\hat{g}_\theta = Z_\theta g$ to Poisson's equation converge in norm in L_V^∞ as $\theta \to \theta_0$.

4.3 Specific Models

Random walks and queues Consider the random walk on a half line given by $\Phi_n = [\Phi_{n-1} + W_n]^+$. We will assume also that the increment distribution Γ has a finite fifth moment. The chain $\boldsymbol{\Phi}$ can be viewed as the waiting-time sequence of a single-server queueing systems (see Asmussen [2]). It is well known that if $\mathsf{E}[|W_1|] < \infty$, then it is necessary and sufficient that $\mathsf{E}[W_1] < 0$ in order that $\boldsymbol{\Phi}$ be a positive recurrent Harris chain. Furthermore, since $K_{a_\varepsilon}(x, \{0\})$ is positive everywhere and bounded from below on compacta, all compact sets are petite, so that $\boldsymbol{\Phi}$ is a T-chain (cf [23]).

Let $f_p(x) = x^p + 1$ and $V_p(x) = cx^{p+1}$, with $c > 0$. We have that (8) holds for some c, and $p = 1, 4$: By Theorem 2.2 the chain is f_4-regular, and hence by definition the chain is simultaneously f_1-regular and V_1^2-regular. We see from Theorem 4.1 that the FCLT holds for any g satisfying $|g| \leq f_1$. In particular, on setting $g(x) = x$ we see that the FCLT holds for $\boldsymbol{\Phi}$ itself. For further discussion see Glynn [12].

These results also hold for some network models. See for example Meyn and Down [20] and Kumar and Meyn [14], where V-uniform ergodicity of generalized Jackson networks and certain re-entrant lines is established. In these papers, the Lyapunov function V may be taken as the exponential of a norm.

Linear state space models Consider the linear state space model which is defined by an $n \times n$ matrix F and an $n \times p$ matrix G such that for each $k \in \mathbb{Z}_+$, the random variables X_k and W_k take values in $\mathbb{R}^n$ and $\mathbb{R}^p$, respectively, and satisfy inductively for $k \in \mathbb{Z}_+$,

$$X_{k+1} = FX_k + GW_{k+1},$$

where $X_0 \in \mathbb{R}^n$ is arbitrary. The random variables $\{W_k\}$ are independent and identically distributed (i.i.d), and are independent of X_0, with common distribution $\Gamma(A) = \mathsf{P}(W_j \in A)$ having finite zero mean and finite covariance $\Sigma_W = \mathsf{E}[WW^\top]$.

We assume that Γ is nonsingular with respect to Lebesgue measure, so that in particular $\Sigma_W > 0$, and we assume that the controllability matrix $[F^{n-1}G \mid \cdots \mid FG \mid G]$ has rank n. In addition, we assume that the eigenvalues of F lie in the open unit disk in $\mathbb{C}$. Under these conditions, it follows as in [3] that every compact set is petite, and that the chain $\mathbf{X}$ is positive Harris.

To construct a Lyapunov function for the process, let M denote the solution to the *Lyapunov equation*

$$F^\top MF = M - I.$$

This is possible because of the eigenvalue condition imposed on F (cf Caines [6] and Duflo [10]). Letting $V_0(x) = x^\top Mx$, we have the bound

$$PV_0(x) = V_0(x) - x^\top x + \operatorname{trace}(M^{1/2}G\Sigma_W G^\top M^{1/2}) \leq \lambda V_0(x) + L$$

for some $\lambda < 1$ and $L < \infty$. It then follows that (10) holds for the chain with V a constant multiple of $\sqrt{V_0}$, so that the process is V-uniformly ergodic. Since $\pi(V_0)$, and hence also $\pi(V^2)$ is finite, the FCLT holds for any $|g(x)| \leq c(|x| + 1)$.

Clearly the approach used here may be extended to other models with a basically linear structure. A class of bilinear models are shown to be V-uniformly ergodic in [22], and adaptive control models are treated in [21]. Hence the FCLT holds for the processes of interest in each of these models.

Acknowledgement This paper was begun while the first author was visiting the Department of Mathematics and Statistics at Carleton University, Ottawa, Canada. He wishes to thank the Department for its hospitality.

We are grateful to Adam Shwartz for numerous useful comments on an earlier draft of the manuscript.

Research by the first author supported in part by the U.S. Army Research Office under Contract DAAL-03-kk-K-0063, and by the Natural Sciences of Engineering Research Council of Canada grant.

Research by the second author supported in part by NSF grants ECS 89-10088 and ECS 92-16487, and by University of Illinois Research Board grant 1-2-69637.

REFERENCES

[1] D. J. Aldous and H. Thorisson. Shift-coupling. *Stoch. Proc. Applns.*, 44:1–4, 1993.

[2] S. Asmussen. *Applied Probability and Queues.* John Wiley & Sons, New York, 1987.

[3] K. B. Athreya and S. G. Pantula. Mixing properties of Harris chains and autoregressive processes. *J. Appl. Probab.*, 23:880–892, 1986.

[4] A. Benveniste, M. Metivier, and P. Priouret. *Adaptive algorithms and stochastic approximations.* Applications of Mathematics, 22. Springer-Verlag, Berlin; New York, 1990. Translated from the French by Stephen S. Wilson.

[5] R. N. Bhattacharaya. On the functional central limit theorem and the law of the iterated logarithm for Markov processes. *Z. Wahrscheinlichkeitstheorie und Verw. Geb.*, 60:185–201, 1982.

[6] P. Caines. *Linear Stochastic Systems.* John Wiley & Sons, New York, 1988.

[7] C. Constantinescu and A. Cornea. *Potential Theory on Harmonic Spaces.* Springer-Verlag, Berlin, 1972.

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

[9] D. Down, S. P. Meyn, and R. L. Tweedie. Geometric and uniform ergodicity of Markov processes. submitted, 1993.

[10] M. Duflo. *Méthodes Récursives Aléatoires.* Masson, 1990.

[11] S. R. Foguel. *The Ergodic Theory of Markov Processes.* Van Nostrand Reinhold, New York, 1969.

[12] P. W. Glynn. Poisson's equation for the recurrent M/G/1 queue. to appear, *Advances in Applied Probability*, 1994.

[13] P. Hall and C. C. Heyde. *Martingale Limit Theory and Its Application.* Academic Press, New York, 1980.

[14] P. R. Kumar and S. P. Meyn, "Stability of queueing networks and scheduling policies," *IEEE Transactions on Automatic Control*, vol. 40, pp. 251–260, February 1995.

[15] T. G. Kurtz. The central limit theorem for Markov chains. *Ann. Probab.*, 9:557–560, 1981.

[16] H. J. Kushner. *Stochastic Stability and Control.* Academic Press, New York, 1967.

[17] N. Maigret. Théorème de limite centrale pour une chaine de Markov récurrente Harris positive. *Ann. Inst. Henri Poincaré Ser B*, 14:425–440, 1978.

[18] A. Makowski and A. Shwartz. Stochastic approximations and adaptive control of a discrete–time single-server network with random routing. *SIAM J. Control Optim.*, 30, 1992.

[19] M. Metivier and P. Priouret. Theoremes de convergence presque sure pour une classe d'algorithmes stochastiques a pas decroissants. *Prob. Theory Related Fields*, 74:403–428, 1987.

[20] S. P. Meyn and D. Down. Stability of generalized Jackson networks. *Ann. Appl. Probab.*, 4:124–148, 1994.

[21] S. P. Meyn and L. Guo. Stability, convergence, and performance of an adaptive control algorithm applied to a randomly varying system. *IEEE Trans. Automat. Control*, AC-37:535–540, 1992.

[22] S. P. Meyn and L. Guo. Geometric ergodicity of a bilinear time series model. *Journal of Time Series Analysis*, 14(1):93–108, 1993.

[23] S. P. Meyn and R. L. Tweedie. Stability of Markovian processes I: Discrete time chains. *Adv. Appl. Probab.*, 24:542–574, 1992.

[24] S. P. Meyn and R. L. Tweedie. Generalized resolvents and Harris recurrence of Markov processes. *Contemporary Mathematics*, 149:227–250, 1993.

[25] S. P. Meyn and R. L. Tweedie. *Markov Chains and Stochastic Stability.* Springer-Verlag, London, 1993.

[26] S. P. Meyn and R. L. Tweedie. Stability of Markovian processes II: Continuous time processes and sampled chains. *Adv. Appl. Probab.*, 25:487–517, 1993.

[27] S. P. Meyn and R. L. Tweedie. Stability of Markovian processes III: Foster-Lyapunov criteria for continuous time processes. *Adv. Appl. Probab.*, 25:518–548, 1993.

[28] J. Neveu. Potentiel Markovien récurrent des chaînes de Harris. *Ann. Inst. Fourier, Grenoble*, 22:7–130, 1972.

[29] E. Nummelin. *General Irreducible Markov Chains and Non-Negative Operators.* Cambridge University Press, Cambridge, 1984.

[30] E. Nummelin. On the Poisson equation in the potential theory of a single kernel. *Math. Scand.*, 68:59–82, 1991.

[31] E. Nummelin and P. Tuominen. Geometric ergodicity of Harris recurrent Markov chains with applications to renewal theory. *Stoch. Proc. Applns.*, 12:187–202, 1982.

[32] D. Revuz. *Markov Chains.* North-Holland, Amsterdam, 2nd edition, 1984.

[33] S.M. Ross. *Introduction to Stochastic Dynamic Programming.* Academic Press, New York, NY, 1984.

[34] P. J. Schweitzer. Perturbation theory and finite Markov chains. *J. Appl. Prob.*, 5:401–403, 1968.

[35] M. Sharpe. *General Theory of Markov Processes.* Academic Press, New York, 1988.

[36] A. Shwartz and A. Makowski. On the Poisson equation for Markov chains: existence of solutions and parameter dependence. Technical Report, Technion—Israel Institute of Technology, Haifa 32000, Israel., 1991.