



Universiteit
Leiden
The Netherlands

Approach to Markov operators on spaces of measures by means of equicontinuity

Ziemlańska, M.A.

Citation

Ziemlańska, M. A. (2021, February 10). *Approach to Markov operators on spaces of measures by means of equicontinuity*. Retrieved from <https://hdl.handle.net/1887/3135034>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/3135034>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <https://hdl.handle.net/1887/3135034> holds various files of this Leiden University dissertation.

Author: Ziemiańska, M.A.

Title: Approach to Markov Operators on spaces of measures by means of equicontinuity

Issue Date: 2021-02-10

Chapter 5

Central Limit Theorem for some non-stationary Markov chains

This chapter is based on:

Jacek Gulgowski, Sander C. Hille, Tomasz Szarek, Maria A. Ziemlańska. Central Limit Theorem for some non-stationary Markov chains. *Studia Mathematica*, Number 246 (2019), Pages 109-131.

Abstract:

Using the classical Central Limit Theorem for stationary Markov chains proved by M. I. Gordin and B. A. Lifšic in [GL78] we show that it also holds for non-stationary Markov chains provided the transition probabilities satisfy the spectral gap property in the Kantorovich–Rubinstein norm.

5.1 Introduction

In this chapter we aim at showing that the Central Limit Theorem (**CLT**) obtained for stationary Markov chains by Gordin and Lifšic in [GL78] (see also its extension due to M. Maxwell and M. Woodroffe in [MW00]) may be extended to non-stationary ones provided that their transition probabilities satisfy the Spectral Gap Property in the Kantorovich–Rubinstein norm (see condition **(A2)** below). This condition implies that the iteration of the Markov operator associated to these transition probabilities constitute an equicontinuous family of continuous maps on probability measures, equipped with the Dudley metric, see Chapter 1, Theorem 2.8.4. In this way we obtain the result stronger than the almost sure result known in the literature as a quenched Central Limit Theorems (see [Pel15, BI95]).

Recently the **CLT** was proved for various non-stationary Markov processes (see [KW12, Kuk02]). For more details we refer readers to the book by T. Komorowski *et al.* [KLO12], where a more detailed description of recent results on central limit theorems is provided. Our result is in the same spirit as the main theorem in [KW12]. However some delicate approximation allow us to obtain the **CLT** for initial distributions with 2-nd moment finite instead of $2 + \delta$.

In this chapter we introduce notation that deviates from previous chapters, mainly because those used here are more common in the field of probability theory, while the topic considered is specially targeted to an audience from this field.

Suppose that (X, ρ) is a Polish space. By $\mathcal{B}(X)$ we denote the family of all Borel sets in X . Denote by $B_b(X)$ the set of all bounded Borel measurable functions equipped with the supremum norm and let $C_b(X)$ be its subset consisting of all bounded continuous functions.

By $\mathcal{M}_1$ and $\mathcal{M}$ we denote the spaces of all probability Borel measures and of all Borel measures on X , respectively. Let $\pi : X \times \mathcal{B}(X) \rightarrow [0, 1]$ be a transition probability on X and let $U : B_b(X) \rightarrow B_b(X)$ be defined by $Uf(x) = \int_X f(y)\pi(x, dy)$ for every $f \in B_b(X)$. The operator U is dual to the Markov operator P defined on $\mathcal{M}$ and given by the formula $P\mu(\cdot) = \int_X \pi(x, \cdot)\mu(dx)$ for $\mu \in \mathcal{M}$, i.e.

$$\int_X f(x)P\mu(dx) = \int_X Uf(x)\mu(dx) \quad \text{for any } f \in B_b(X) \text{ and } \mu \in \mathcal{M}.$$

In particular, we have $P\delta_x(\cdot) = \pi(x, \cdot)$ for $x \in X$.

Suppose that (X_n) is an X -valued Markov chain given over some probability space $(\Omega, \mathcal{F}, P)$,

whose transition probability is π . If the distribution of X_0 is μ_0 , then the distribution of X_n equals $P^n\mu_0$, $n \geq 1$. By $(\mathcal{F}_n)$ we denote the natural filtration of the chain, i.e. the increasing family of σ -algebras $\mathcal{F}_n := \sigma(X_i : i \leq n)$.

For any $x \in X$ and $n \geq 1$ we define the measure $\mathbf{P}_x^n$ on $(X^n, \mathcal{B}(X^n))$, where

$$\mathcal{B}(X^n) = \sigma(\{A_1 \times A_2 \times \cdots \times A_n : A_i \in \mathcal{B}(X), i = 1, \dots, n\})$$

by the formula

$$\begin{aligned} \mathbf{P}_x^n(A_1 \times \cdots \times A_n) &= \int_{X^n} \chi_{A_1 \times \cdots \times A_n}(x_1, \dots, x_n) \pi(x_{n-1}, dx_n) \cdots \pi(x_1, dx_2) \delta_x(dx_1) \\ &= \int_{X^n} \chi_{A_1 \times \cdots \times A_n}(x_1, \dots, x_n) P\delta_{x_{n-1}}(dx_n) \cdots P\delta_{x_1}(dx_2) P\delta_x(dx_1). \end{aligned}$$

Here $\chi_{A_1 \times \cdots \times A_n}$ denotes the characteristic function of $A_1 \times \cdots \times A_n$. Obviously, the distribution of the random vector $(X_1, \dots, X_n)$ if $X_0 = x$ is given by $\mathbf{P}_x^n$. On the other hand, if the distribution of X_0 equals μ_0 , then the distribution of the random vector $(X_1, \dots, X_n)$ is equal to $\mathbf{P}_{\mu_0}^n(\cdot) = \int_X \mathbf{P}_x^n(\cdot) \mu_0(dx)$.

By $\mathbf{P}_x$ we will denote the probability measure on the Borel σ -algebra of the trajectory space X^∞ associated with (X_n) with $X_0 = x$. Further, $\mathbb{E}_x$ denotes the expected value with respect to $\mathbf{P}_x$. Analogously, if X_0 is distributed with μ_0 , then $\mathbf{P}_{\mu_0}(\cdot) = \int_X \mathbf{P}_x(\cdot) \mu_0(dx)$. Similarly, we have then $\mathbb{E}_\mu(\cdot) = \int_X \mathbb{E}_x(\cdot) \mu(dx)$.

5.2 Assumptions

Set

$$\mathcal{M}_1^p := \left\{ \mu \in \mathcal{M}_1 : \int_X [\rho(x, x_0)]^p \mu(dx) < \infty \right\} \quad \text{for } p \geq 1.$$

We equip the space $\mathcal{M}_1^1$ with the Kantorovich-Rubinstein distance

$$\|\mu - \nu\| := \sup \left| \int_X f d\mu - \int_X f d\nu \right| \quad \text{for } \mu, \nu \in \mathcal{M}_1^1,$$

where the supremum is taken over all Lipschitz functions $f : X \rightarrow \mathbb{R}$ with the Lipschitz constant bounded by 1. If μ and ν are two Borel probability measures on some Polish spaces W and Z respectively, then by $\mathcal{C}(\mu, \nu)$ we denote the set of all joint Borel probability measures on $W \times Z$ whose marginals are μ and ν (the so-called *coupling*). If $\mu, \nu \in \mathcal{M}_1^1$, then according to the following Kantorovich-Rubinstein Theorem

Theorem 5.2.1. *[Kantorovich-Rubinstein Theorem [Bog07a], Theorem 8.10.45] The Kantorovich-Rubinstein distance $\|\mu - \nu\|$ between Radon probability measures $\mu, \nu \in \mathcal{M}_1^1$ can be represented in the form*

$$\|\mu - \nu\| = \inf_{\gamma \in \mathcal{C}(\mu, \nu)} \int_{X \times X} \rho(x, y) \gamma(\mathrm{d}x, \mathrm{d}y).$$

Moreover, there exists a measure $\lambda_0 \in C(\mu, \nu)$ at which the value $\|\mu - \nu\|$ is attained.

This immediately implies

$$\|\delta_x - \delta_y\| = \rho(x, y) \quad \text{for } x, y \in X.$$

Assumptions:

(A1) We assume that P leaves $\mathcal{M}_1^1$ invariant;

(A2) there exist $C > 0$ and $0 < q < 1$ such that

$$\|P^n \mu - P^n \nu\| \leq C q^n \|\mu - \nu\| \quad (5.1)$$

for all $\mu, \nu \in \mathcal{M}_1^1$ and $n \in \mathbb{N}$;

(A3) we assume that there exists $\mu \in \mathcal{M}_1^2$ such that

$$\sup_{n \geq 1} \int_X [\rho(x, x_0)]^2 P^n \mu(\mathrm{d}x) < +\infty \quad \text{for some (thus all) } x_0 \in X. \quad (5.2)$$

The following theorem is standard but we provide its proof for completeness of our presentation.

Proposition 5.2.1. *If (A1) and (A2) hold, then P has a unique invariant measure $\mu_* \in \mathcal{M}_1^1$ and for any $\mu \in \mathcal{M}_1^1$ we have $\|P^n \mu - \mu_*\| \rightarrow 0$ as $n \rightarrow \infty$. Moreover, if (A3) holds, then $\mu_* \in \mathcal{M}_1^2$.*

Proof. For any $\mu \in \mathcal{M}_1^1(X)$ and any $m, n \in \mathbb{N}$, $m \geq n$ one obtains from (5.1) that

$$\|P^n \mu - P^m \mu\| \leq \sum_{k=0}^{m-n-1} \|P^{n+k} \mu - P^{n+k+1} \mu\| \leq \frac{C q^n}{1-q} \|\mu - P\mu\|.$$

So $(P^n \mu)$ is a Cauchy sequence in the $\|\cdot\|$ -complete space $\mathcal{M}_1^1$ ([Vil08], Theorem 6.18). Hence it converges to some invariant measure $\mu_* \in \mathcal{M}_1^1(X)$. Then again by (5.1) for any

$$\mu \in \mathcal{M}_1^1(X),$$

$$\|P^n \mu - \mu_*\| = \|P^n \mu - P^n \mu_*\| \leq C q^n \|\mu - \mu_*\| \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

So μ_* must be unique. Now suppose that **(A3)** holds and let $\mu \in \mathcal{M}_1^2$ be such that (5.2) is satisfied. Let $x_0 \in X$. For any $K > 0$, one has

$$\sup_{n \geq 1} \int_X [\rho(x, x_0) \wedge K]^2 P^n \mu(dx) \leq \sup_{n \geq 1} \int_X \rho(x, x_0)^2 P^n \mu(dx) =: M_0 < \infty.$$

Since $x \mapsto [\rho(x, x_0) \wedge K]^2$ is a Lipschitz function on X and $\|P^n \mu - \mu_*\| \rightarrow 0$, we have

$$\int_X [\rho(x, x_0) \wedge K]^2 P^n \mu(dx) \rightarrow \int_X [\rho(x, x_0) \wedge K]^2 \mu_*(dx).$$

So $\int_X [\rho(x, x_0) \wedge K]^2 \mu_*(dx) \leq M_0$ for every K . By the Monotone Convergence Theorem we obtain that $\int_X [\rho(x, x_0)]^2 \mu_*(dx) \leq M_0$ too. $\square$

5.3 Gordin–Lifšic results for stationary case

We start with the following simple consequences of the Gordin and Lifšic result on the CLT for stationary Markov chains:

Proposition 5.3.1. *Let P be a Markov operator that satisfies **(A1)** - **(A3)** and let (X_n) be a stationary Markov chain corresponding to P . Then for any bounded Lipschitz function $g : X \rightarrow \mathbb{R}$ such that $\int_X g d\mu_* = 0$, where μ_* is the unique invariant distribution for P , the limit*

$$\sigma^2 := \lim_{n \rightarrow \infty} \mathbb{E}_{\mu_*} \left(\frac{g(X_0) + g(X_1) + \cdots + g(X_n)}{\sqrt{n}} \right)^2$$

exists and is finite. Moreover, if $\sigma > 0$, then

$$\lim_{n \rightarrow \infty} P \left(\frac{g(X_0) + g(X_1) + \cdots + g(X_n)}{\sqrt{n}} < a \right) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^a e^{-\frac{y^2}{2\sigma^2}} dy \quad \text{for all } a \in \mathbb{R}.$$

Otherwise, if $\sigma = 0$, then the sequence

$$\frac{g(X_0) + g(X_1) + \cdots + g(X_n)}{\sqrt{n}} \quad \text{for } n \geq 1$$

converges in distribution to 0.

Proof. Let a bounded Lipschitz function $g : X \rightarrow \mathbb{R}$ such that $\int_X g d\mu_* = 0$ be given. With no loss of generality we may assume that the Lipschitz constant of g equals 1. From Gordin and Lifšic [GL78] it follows that to finish the proof it is enough to show that

$$\sum_{n=0}^{\infty} \|U^n g\|_{L^2(X, \mu_*)} < \infty,$$

where U is dual to P . In fact, since $\int_X U^i g(x) \mu_*(dx) = \int_X g(x) \mu_*(dx) = 0$ for all $i \geq 0$, we have

$$\begin{aligned} \sum_{n=1}^{\infty} \|U^n g\|_{L^2(X, \mu_*)} &= \sum_{n=1}^{\infty} \left(\int_X (U^n g(x))^2 \mu_*(dx) \right)^{1/2} \\ &= \sum_{n=1}^{\infty} \left(\int_X (U^n g(x) - \int_X U^n g(y) \mu_*(dy))^2 \mu_*(dx) \right)^{1/2} \\ &\leq \sum_{n=1}^{\infty} \left(\int_X \left(\int_X |U^n g(x) - U^n g(y)| \mu_*(dy) \right)^2 \mu_*(dx) \right)^{1/2} \\ &\leq \sum_{n=1}^{\infty} \left(\int_X \left(\int_X Cq^n \rho(x, y) \mu_*(dy) \right)^2 \mu_*(dx) \right)^{1/2} \\ &\leq C \sum_{n=1}^{\infty} q^n \left(\int_X \left(\int_X \rho(x, y) \mu_*(dy) \right)^2 \mu_*(dx) \right)^{1/2} \\ &\leq C \sum_{n=1}^{\infty} q^n \left(\int_X \left(\rho(x, x_0) + \int_X \rho(x_0, y) \mu_*(dy) \right)^2 \mu_*(dx) \right)^{1/2} \\ &\leq C(1-q)^{-1} \left(2 \int_X [\rho(x, x_0)]^2 \mu_*(dx) + 2 \left(\int_X \rho(x_0, y) \mu_*(dy) \right)^2 \right)^{1/2} < \infty, \end{aligned}$$

by the fact that $\mu_* \in \mathcal{M}_1^2$. The proof is complete. $\square$

5.4 Auxiliary lemmas

We start with a theorem on the existence of a suitable coupling for the trajectories of a given Markov chain.

Theorem 5.4.1. *Assume that a Markov operator P satisfies (A1) and (A2). Let l_0 be a positive integer such that $Cq^{l_0} < 1$, where the constants C, q are given by (5.1). Then there exists a constant $\kappa > 0$ such that for every integers $l \geq l_0$ and $m \geq 1$ and every two points $x, y \in X$ we have a measure $\mathbf{P}_{x,y}^{ml} \in \mathcal{C}(\mathbf{P}_x^{ml}, \mathbf{P}_y^{ml})$ satisfying*

$$\sum_{i=1}^m \int_{X^{lm} \times X^{lm}} \rho(x_{il}, y_{il}) \mathbf{P}_{x,y}^{ml}(dx_1, \dots, dx_{ml}, dy_1, \dots, dy_{ml}) \leq \kappa \rho(x, y). \quad (5.3)$$

Proof. Choose $\tilde{q} \in (Cq^{l_0}, 1)$. Fix $l \geq l_0$. By induction on m for every two points $x, y \in X$ we construct $\mathbf{P}_{x,y}^{ml} \in \mathcal{C}(\mathbf{P}_x^{ml}, \mathbf{P}_y^{ml})$ and prove that

$$\int_{X^{lm} \times X^{lm}} \left(\sum_{i=1}^m \rho(x_{il}, y_{il}) \right) \mathbf{P}_{x,y}^{ml}(\mathrm{d}x_1, \dots, \mathrm{d}x_{ml}, \mathrm{d}y_1, \dots, \mathrm{d}y_{ml}) \leq \sum_{j=1}^m \tilde{q}^j \rho(x, y). \quad (5.4)$$

Then the hypothesis will hold with $\kappa = \tilde{q}(1 - \tilde{q})^{-1}$. Fix $x, y \in X$ and let $m = 1$. Since

$$\|P^l \delta_x - P^l \delta_y\| \leq Cq^{l_0} \rho(x, y),$$

by the Kantorovich–Rubinstein theorem there exists $\mu_{x,y}^1 \in \mathcal{C}(P^l \delta_x, P^l \delta_y)$ such that

$$\|P^l \delta_x - P^l \delta_y\| \leq \int_{X \times X} \rho(u, v) \mu_{x,y}^1(\mathrm{d}u, \mathrm{d}v) \leq \tilde{q} \rho(x, y).$$

From the proof of the Kantorovich–Rubinstein theorem it follows that the function $X \times X \ni (x, y) \rightarrow \mu_{x,y}^1 \in \mathcal{C}(P^l \delta_x, P^l \delta_y)$ is measurable if the space $\mathcal{C}(P^l \delta_x, P^l \delta_y)$ is endowed with some metric of weak topology (see Theorem 11.8.2 in [Dud02]). We may assume that the measure $\mu_{x,y}^1$ is absolutely continuous with respect to the product measure $P^l \delta_x \otimes P^l \delta_y$. Let $g_{x,y} : X \times X \rightarrow [0, +\infty)$ be the Radon–Nikodem density, i.e.

$$g_{x,y}(u, v) = \frac{\mathrm{d}\mu_{x,y}^1}{\mathrm{d}(P^l \delta_x \otimes P^l \delta_y)}(u, v) \quad u, v \in X.$$

Define

$$\begin{aligned} \mathbf{P}_{x,y}^l(A) &= \int_{X^l} \cdots \int_{X^l} \chi_A(x_1, \dots, x_l, y_1, \dots, y_l) g_{x,y}(x_l, y_l) \pi(x_{l-1}, \mathrm{d}x_l) \pi(y_{l-1}, \mathrm{d}y_l) \\ &\quad \cdots \pi(x_1, \mathrm{d}x_2) \pi(y_1, \mathrm{d}y_2) \delta_x(\mathrm{d}x_1) \delta_y(\mathrm{d}y_1). \end{aligned}$$

Let $A = [A_1 \times \cdots \times A_l] \times X^l$ for some Borel sets $A_1, \dots, A_l \subset X$. We have

$$\begin{aligned} \mathbf{P}_{x,y}^l(A) &= \int_{X^l} \chi_{A_1 \times \cdots \times A_l}(x_1, \dots, x_l) \left(\int_{X^l} g_{x,y}(x_l, y_l) \pi(y_{l-1}, \mathrm{d}y_l) \cdots \pi(y_1, \mathrm{d}y_2) \delta_y(\mathrm{d}y_1) \right) \\ &\quad \times \pi(x_{l-1}, \mathrm{d}x_l) \cdots \pi(x_1, \mathrm{d}x_2) \delta_x(\mathrm{d}x_1). \end{aligned}$$

Moreover, the measure

$$\int_{X^{l-1}} \chi_{A_1 \times \cdots \times A_{l-1}}(x_1, \dots, x_{l-1}) \pi(x_{l-1}, \cdot) \cdots \pi(x_1, \mathrm{d}x_2) \delta_x(\mathrm{d}x_1)$$

is absolutely continuous with respect to $P^l \delta_x$. Obviously,

$$\int_{X^l} g_{x,y}(x_l, y_l) \pi(y_{l-1}, dy_l) \cdots \pi(y_1, dy_2) \delta_y(dy_1) = \int_X g_{x,y}(x_l, y_l) P^l \delta_y(dy_l).$$

If we show that $\int_X g_{x,y}(\cdot, y_l) P^l \delta_y(dy_l) = 1$, $P^l \delta_x$ -a.s., then

$$\begin{aligned} \mathbf{P}_{x,y}^l(A) &= \int_{X^l} \chi_{A_1 \times \cdots \times A_l}(x_1, \dots, x_l) \pi(x_{l-1}, dx_l) \cdots \pi(x_1, dx_2) \delta_x(dx_1) \\ &= \mathbf{P}_x^l(A_1 \times \cdots \times A_l). \end{aligned}$$

To do this set $h(x_l) := \int_X g_{x,y}(x_l, y_l) P^l \delta_y(dy_l)$. From the definition of the coupling measure $\mu_{x,y}^1$ for any Borel set $B \subset X$ we have

$$\begin{aligned} P^l \delta_x(B) &= \mu_{x,y}^1(B \times X) = \int_B \left(\int_X g_{x,y}(x_l, y_l) P^l \delta_y(dy_l) \right) P^l \delta_x(dx_l) \\ &= \int_B h(x_l) P^l \delta_x(dx_l). \end{aligned}$$

Since the above equality holds for an arbitrary Borel set $B \subset X$, we obtain that $h(x_l) = 1$, $P^l \delta_x$ -a.s., and consequently $\mathbf{P}_{x,y}^l([A_1 \times \cdots \times A_l] \times X^l) = \mathbf{P}_x^l(A_1 \times \cdots \times A_l)$. In the same way we show that $\mathbf{P}_{x,y}^l(X^l \times [B_1 \times \cdots \times B_l]) = \mathbf{P}_y^l(B_1 \times \cdots \times B_l)$ for Borel sets $B_1, \dots, B_l \subset X$. Hence $\mathbf{P}_{x,y}^l \in \mathcal{C}(\mathbf{P}_x^l, \mathbf{P}_y^l)$. Finally, we have

$$\begin{aligned} &\int_{X^l \times X^l} \rho(x_l, y_l) \mathbf{P}_{x,y}^l(dx_1, \dots, dx_l, dy_1, \dots, dy_l) \\ &= \int_{X^l \times X^l} \rho(x_l, y_l) g_{x,y}(x_l, y_l) \pi(x_{l-1}, dx_l) \pi(y_{l-1}, dy_l) \cdots \pi(x_1, dx_2) \pi(y_1, dy_2) \delta_x(dx_1) \delta_y(dy_1) \\ &= \int_{X \times X} \rho(x_l, y_l) g_{x,y}(x_l, y_l) P^l \delta_x(dx_l) P^l \delta_y(dy_l) = \int_{X \times X} \rho(x_l, y_l) \mu^1(dx_l, dy_l) \leq \tilde{q} \rho(x, y) \end{aligned}$$

and the first step of the proof is finished.

Assume now that for $j = 1, \dots, m$ and arbitrary $x, y \in X$ there exists $\mathbf{P}_{x,y}^{jl} \in \mathcal{C}(\mathbf{P}_x^{jl}, \mathbf{P}_y^{jl})$ such that condition (5.4) holds with m replaced by j .

Define the measure $\mathbf{P}_{x,y}^{(m+1)l}$ on $X^{(m+1)l} \times X^{(m+1)l}$ by the formula

$$\begin{aligned} \mathbf{P}_{x,y}^{(m+1)l}(A) &= \int_{X^{(m+1)l} \times X^{(m+1)l}} \chi_A(x_1, \dots, x_{(m+1)l}, y_1, \dots, y_{(m+1)l}) \\ &\quad \mathbf{P}_{x_l, y_l}^{ml}(dx_{l+1}, \dots, dx_{(m+1)l}, dy_{l+1}, \dots, dy_{(m+1)l}) \mathbf{P}_{x,y}^l(dx_1, \dots, dx_l, dy_1, \dots, dy_l) \end{aligned}$$

for any Borel set $A \subset X^{(m+1)l} \times X^{(m+1)l}$. From the Markov property it easily follows that $\mathbf{P}_{x,y}^{(m+1)l} \in \mathcal{C}(\mathbf{P}_x^{(m+1)l}, \mathbf{P}_y^{(m+1)l})$. We also have

$$\begin{aligned}
 & \int_{X^{(m+1)l} \times X^{(m+1)l}} \sum_{j=1}^{m+1} \rho(x_{jl}, y_{jl}) \mathbf{P}_{x,y}^{(m+1)l} (dx_1, \dots, dx_{(m+1)l}, dy_1, \dots, dy_{(m+1)l}) \\
 & \leq \int_{X^l \times X^l} \left(\int_{X^{ml} \times X^{ml}} \sum_{j=2}^{m+1} \rho(x_{jl}, y_{jl}) \mathbf{P}_{x_l, y_l}^{ml} (dx_{l+1}, \dots, dx_{(m+1)l}, dy_{l+1}, \dots, dy_{(m+1)l}) \right) \\
 & \quad \mathbf{P}_{x,y}^l (dx_1, \dots, dx_l, dy_1, \dots, dy_l) + \int_{X^l \times X^l} \rho(x_l, y_l) \mathbf{P}_{x,y}^l (dx_1, \dots, dx_l, dy_1, \dots, dy_l) \\
 & \leq \sum_{j=1}^m \tilde{q}^j \int_{X^l \times X^l} \rho(x_l, y_l) \mathbf{P}_{x,y}^l (dx_1, \dots, dx_l, dy_1, \dots, dy_l) + \tilde{q} \rho(x, y) \\
 & \leq \sum_{j=1}^m \tilde{q}^j \tilde{q} \rho(x, y) + \tilde{q} \rho(x, y) = \sum_{j=1}^{m+1} \tilde{q}^j \rho(x, y)
 \end{aligned}$$

by the inductive hypothesis for 1 and m . This completes the proof. $\square$

As a consequence of Proposition 5.3.1 and Theorem 5.4.1 we have the following:

Proposition 5.4.1. *Let P be a Markov operator that satisfies (A1) - (A3) and let (X_n) be a Markov chain corresponding to P . Let l_0 be a positive integer such that $Cq^{l_0} < 1$, where the constants C, q are given by (5.1) and let $l \geq l_0$ be given. Set $Y_n = X_{nl}$ for $n \geq 0$ and assume that $X_0 = Y_0 = x$ for $x \in X$. Then for any bounded Lipschitz function $g : X \rightarrow \mathbb{R}$ such that $\int_X g d\mu_* = 0$, where μ_* is the unique invariant distribution for P , the limit*

$$\tilde{\sigma}^2 = \lim_{n \rightarrow \infty} \mathbb{E}_{\mu_*} \left(\frac{g(Y_0) + g(Y_1) + \dots + g(Y_n)}{\sqrt{n}} \right)^2$$

exists and is finite. Moreover, if $\tilde{\sigma}^2 > 0$, then the sequence of random vectors (Y_n) satisfies

$$\lim_{n \rightarrow \infty} P \left(\frac{g(Y_0) + g(Y_1) + \dots + g(Y_n)}{\sqrt{n}} < a \right) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^a e^{-\frac{y^2}{2\sigma^2}} dy$$

for all $a \in \mathbb{R}$.

Proof. Without loss of generality we may assume that $g : X \rightarrow \mathbb{R}$ is bounded by 1 and its Lipschitz constant is also bounded by 1. By φ_n^x , $x \in X$, we denote the characteristic function of the random variable

$$\frac{g(Y_0) + g(Y_1) + \dots + g(Y_n)}{\sqrt{n}},$$

where $Y_0 = x$. When the distribution of Y_0 is equal to μ_* the characteristic function is denoted by $\varphi_n^{\mu_*}$. Obviously,

$$\varphi_n^{\mu_*}(t) = \int_X \varphi_n^y(t) \mu_*(dy) \quad \text{for } t \in \mathbb{R}.$$

Moreover, from Proposition 5.3.1 it follows that

$$\lim_{n \rightarrow \infty} \varphi_n^{\mu_*}(t) = e^{-\frac{\tilde{\sigma}^2 t^2}{2}} \quad \text{for } t \in \mathbb{R}.$$

Theorem 5.4.1 allows us to evaluate the difference of characteristic functions $\varphi_n^x(t)$ and $\varphi_n^y(t)$ for $x, y \in X$ and $t \in \mathbb{R}$. Fix $x, y \in X$. We have

$$\begin{aligned} |\varphi_n^x(t) - \varphi_n^y(t)| &= \left| \int_{X^{ln}} \exp\left(it \frac{g(x_0) + g(x_l) \dots + g(x_{nl})}{\sqrt{n}}\right) \mathbf{P}_x^{nl}(dx_1, dx_2, \dots, dx_{nl}) \right. \\ &\quad \left. - \int_{X^{ln}} \exp\left(it \frac{g(y_0) + g(y_l) \dots + g(y_{nl})}{\sqrt{n}}\right) \mathbf{P}_y^{nl}(dy_1, dy_2, \dots, dy_{nl}) \right| \\ &\leq \frac{|t|}{\sqrt{n}} \int_{X^{ln} \times X^{ln}} [\rho(x, y) + \rho(x_l, x_l) + \dots + \rho(x_{nl}, y_{nl})] \mathbf{P}_{x,y}^{nl}(dx_1, \dots, dx_{nl}, dy_1, \dots, dy_{nl}) \\ &\leq \frac{|t|}{\sqrt{n}} (1 + \kappa) \rho(x, y) \quad \text{for } t \in \mathbb{R}, \end{aligned}$$

where $\mathbf{P}_{x,y}^{nl} \in \mathcal{C}(\mathbf{P}_x^{nl}, \mathbf{P}_y^{nl})$ is given by Theorem 5.4.1. Consequently, for $t \in \mathbb{R}$ we obtain

$$\begin{aligned} |\varphi_n^x(t) - \varphi_n^{\mu_*}(t)| &= |\varphi_n^x(t) - \int_X \varphi_n^y(t) \mu_*(dy)| \\ &\leq \lim_{n \rightarrow \infty} \frac{|t|(1 + \kappa)}{\sqrt{n}} \int_X \rho(x, y) \mu_*(dy) \rightarrow 0 \quad \text{as } n \rightarrow \infty. \end{aligned}$$

Since $\lim_{n \rightarrow \infty} \varphi_n^{\mu_*}(t) = e^{-t^2 \tilde{\sigma}^2 / 2}$, we obtain that $\lim_{n \rightarrow \infty} \varphi_n^x(t) = e^{-t^2 \tilde{\sigma}^2 / 2}$ and the proof is complete. $\square$

For any Markov chain (Z_n) with values in the space X and an arbitrary Lipschitz function $g : X \rightarrow \mathbb{R}$ we shall denote

$$S_m^n(g(Z_i)) := \sum_{i=m}^n g(Z_i) \quad \text{for } n \geq m \geq 0.$$

The key point in proving the main theorem is the following lemma:

Lemma 5.4.2. *Let P be a Markov operator that satisfies (A1) - (A2) and let (X_n) be a*

Markov chain corresponding to P . Assume that $g : X \rightarrow \mathbb{R}$ is a bounded Lipschitz function. Then there exists a constant $M > 0$ such that for any $n \in \mathbb{N}$ the function

$$x \rightarrow \mathbb{E}_x \left(\frac{S_0^n(g(X_i))}{\sqrt{n}} \right)^2$$

has Lipschitz constant $\leq M$.

Proof. Again assume that $g : X \rightarrow \mathbb{R}$ is a 1-Lipschitz function bounded by 1. Let l_0 be a positive integer such that $Cq^{l_0} < 1$ with the constants C, q given by (5.1). We have

$$\begin{aligned} \mathbb{E}_x \left(\frac{S_0^n(g(X_i))}{\sqrt{n}} \right)^2 &= \frac{1}{n} \sum_{i=0}^n \mathbb{E}_x (g(X_i))^2 + \frac{2}{n} \sum_{i=0}^{n-1} \sum_{j: 0 < j-i < l_0, j \leq n} \mathbb{E}_x (g(X_i)g(X_j)) \\ &\quad + \frac{2}{n} \sum_{i=0}^n \sum_{j: 0 \leq j < l_0, l_0+j+i \leq n} \mathbb{E}_x [g(X_i) \left(\sum_{m=1}^{k_{i,j}} g(X_{ml_0+j+i}) \right)], \end{aligned}$$

where $k_{i,j}$ are maximal positive integers such that $k_{i,j}l_0 + j + i \leq n$. We show that all three terms are Lipschitzian with a Lipschitz constant independent of n . To do this fix $x, y \in X$. Since the function $(g(\cdot))^2$ is a ≤ 2 -Lipschitz function, by (5.1) we have

$$|\mathbb{E}_x(g(X_i))^2 - \mathbb{E}_y(g(X_i))^2| \leq 2\|P^i\delta_x - P^i\delta_y\| \leq 2C\rho(x, y) \quad \text{for } i = 1, \dots, n$$

and the first term

$$\frac{1}{n} \sum_{i=1}^n \mathbb{E}_x (g(X_i))^2$$

is a $\leq 2C$ -Lipschitz function.

Further, for $j > i$ we have

$$\mathbb{E}_z g(X_i)g(X_j) = \mathbb{E}_z (g(X_i)\mathbb{E}_{|X_i} g(X_j)) \quad \text{for all } z \in X,$$

where $\mathbb{E}_{|X_i} g(X_j) := \mathbb{E}[g(X_j)|\mathcal{F}_i]$. Since $\mathbb{E}_{|X_i} g(X_j) = h(X_i)$, where h is a bounded by 1 a $\leq C$ -Lipschitz function, we obtain that

$$\begin{aligned} |\mathbb{E}_x g(X_i)g(X_j) - \mathbb{E}_y g(X_i)g(X_j)| &= |\mathbb{E}_x (g(X_i)h(X_i)) - \mathbb{E}_y (g(X_i)h(X_i))| \\ &\leq 2C\rho(x, y) \end{aligned}$$

for $i = 1, \dots, n$ and consequently the Lipschitz constant of the second term

$$\frac{2}{n} \sum_{i=1}^{n-1} \sum_{j: 0 < j-i < l_0, j \leq n} \mathbb{E}_x g(X_i) g(X_j)$$

is bounded from above by $4l_0 C$.

Finally, we have

$$\mathbb{E}_x \left(g(X_i) \left(\sum_{m=1}^{k_{i,j}} g(X_{ml_0+j+i}) \right) \right) = \mathbb{E}_x \left(g(X_i) \mathbb{E}_{|X_i} \left(\sum_{m=1}^{k_{i,j}} g(X_{ml_0+j}) \right) \right).$$

Using Theorem 5.4.1 we show that the function

$$X \ni z \rightarrow \mathbb{E}_z \left(\sum_{m=1}^{k_{i,j}} g(X_{ml_0+j}) \right) \quad \text{for all } i, j \geq 0$$

has a Lipschitz constant $\leq C\kappa q^j$, where κ is given by Theorem 5.4.1. Indeed, for any $i, j \geq 0$ we have

$$\mathbb{E}_z \left(\sum_{m=1}^{k_{i,j}} g(X_{ml_0+j}) \right) = \mathbb{E}_z \mathbb{E}_{|X_j} \left(\sum_{m=1}^{k_{i,j}} g(X_{ml_0+j}) \right).$$

In turn, from Theorem 5.4.1 we obtain that the function

$$X \ni u \rightarrow r(u) = \mathbb{E}_u \left(\sum_{m=1}^{k_{i,j}} g(X_{ml_0}) \right)$$

has a Lipschitz constant $\leq \kappa$. Indeed, fix $u, v \in X$. We have

$$\begin{aligned} |r(u) - r(v)| &= \left| \mathbb{E}_u \left(\sum_{m=1}^{k_{i,j}} g(x_{ml_0}) \right) - \mathbb{E}_v \left(\sum_{m=1}^{k_{i,j}} g(x_{ml_0}) \right) \right| \\ &\leq \left| \int_{(X^{k_{i,j}l_0})^2} \sum_{m=1}^{k_{i,j}} (g(x_{ml_0}) - g(y_{ml_0})) \mathbf{P}_{u,v}^{k_{i,j}l_0}(dx_1, \dots, dx_{k_{i,j}l_0}; dy_1, \dots, dy_{k_{i,j}l_0}) \right| \\ &\leq \int_{(X^{k_{i,j}l_0})^2} \sum_{m=1}^{k_{i,j}} \rho(x_{ml_0}, y_{ml_0}) \mathbf{P}_{u,v}^{k_{i,j}l_0}(dx_1, \dots, dx_{k_{i,j}l_0}; dy_1, \dots, dy_{k_{i,j}l_0}) \\ &\leq \kappa \rho(u, v), \end{aligned}$$

by Theorem 5.4.1. Since

$$\mathbb{E}_z \left(\sum_{m=1}^{k_{i,j}} g(X_{ml_0+j}) \right) = \int_X r(u) P^j \delta_z(du),$$

the function

$$X \ni z \rightarrow \mathbb{E}_z \left(\sum_{m=1}^{k_{i,j}} g(X_{ml_0+j}) \right)$$

has a Lipschitz constant $\leq C\kappa q^j$, by (5.1). Moreover its supremum norm is bounded by n . Since the Lipschitz constant of the function $z \rightarrow g(z)r(z)$ is bounded by $n + C\kappa q^j$ we obtain that the function

$$X \ni x \rightarrow \mathbb{E}_x \left(g(X_i) \left(\sum_{m=1}^{k_{i,j}} g(X_{ml_0+j+i}) \right) \right)$$

has a Lipschitz constant $\leq Cq^i(n + C\kappa q^j)$. Thus the third term

$$\frac{2}{n} \sum_{i=0}^n \sum_{j: 0 \leq j < l_0, l_0+j+i \leq n} \mathbb{E}_x [g(X_i) \left(\sum_{m=1}^{k_{i,j}} g(X_{ml_0+j+i}) \right)]$$

has Lipschitz constant bounded by

$$\frac{2}{n} \sum_{i=0}^n \sum_{j=0}^{l_0-1} Cq^i(n + C\kappa q^j) \leq \frac{2C}{n} \sum_{i=0}^n l_0(n + C\kappa) q^i \leq \frac{2l_0 C(1 + C\kappa)}{1 - q}.$$

Thus the function

$$x \rightarrow \mathbb{E}_x \left(\frac{S_0^n(g(X_i))}{\sqrt{n}} \right)^2$$

has Lipschitz constant $\leq M$, where

$$M = 2C + 4l_0 C + \frac{2l_0 C(1 + C\kappa)}{1 - q}.$$

This completes the proof. $\square$

Lemma 5.4.3. *Let P be a Markov operator that satisfies **(A1)** - **(A3)** and let (X_n) be a Markov chain corresponding to P . Assume that $g : X \rightarrow \mathbb{R}$ is a bounded Lipschitz function. Let $x \in X$ and $\varepsilon > 0$. Then there exists $K > 0$ and $N_0 \in \mathbb{N}$ such that*

$$\mathbb{E}_{P^n \delta_x} \left[\left(\frac{S_0^n(g(X_i))}{\sqrt{m}} \right)^2 \chi_{[K, +\infty)} \left(\left| \frac{S_0^n(g(X_i))}{\sqrt{m}} \right| \right) \right] \leq \varepsilon \quad (5.5)$$

for all $m \in \mathbb{N}$ and $n \geq N_0$.

Proof. Fix $x \in X$ and let $\varepsilon > 0$. Let g be given as above. Let l_0 be a positive integer

such that $Cq^{l_0} < 1$, where the constants C, q are given by (5.1) and let $l \geq l_0$ be given. Set $Y_n = X_{nl}$ for $n \geq 0$. We first show that the hypothesis holds if we replace X_n with Y_n . Set

$$S_m := \frac{g(Y_0) + g(Y_1) + \dots + g(Y_m)}{\sqrt{m}} \quad \text{for } m \geq 0.$$

The Markov chain (Y_n) corresponds to the operator P^l and the assumptions of Lemma 5.4.2 are satisfied with P and (X_n) replaced by P^l and (Y_n) , respectively. Thus the functions $y \rightarrow \mathbb{E}_y(S_m)^2$ are Lipschitzian with the Lipschitz constant bounded by, say $\tilde{M}$, independent of m . The Markov operator P satisfies **(A2)**, so that we may find $N_0 \in \mathbb{N}$ such that

$$|\mathbb{E}_{P^n \delta_x} S_m^2 - \mathbb{E}_{\mu_*} S_m^2| \leq \tilde{M} \|P^n \delta_x - \mu_*\| \leq \varepsilon / (9l^3) \quad \text{for } n \geq N_0 \text{ and } m \in \mathbb{N}. \quad (5.6)$$

Moreover, from Theorem 5.4.1 it follows that for any Lipschitz function $\Theta : \mathbb{R} \rightarrow \mathbb{R}$ with the Lipschitz constant $\hat{M} > 0$ such that we have for all $m \geq 1$

$$|\mathbb{E}_y \Theta(S_m) - \mathbb{E}_z \Theta(S_m)| \leq \frac{\kappa \hat{M}}{\sqrt{m}} \rho(y, z) \quad \text{for } y, z \in X.$$

Consequently, for some $N_1 \in \mathbb{N}$ we have

$$|\mathbb{E}_{P^n \delta_x} \Theta(S_m) - \mathbb{E}_{\mu_*} \Theta(S_m)| \leq \varepsilon / (9l^3) \quad \text{for } n \geq N_1 \text{ and } m \in \mathbb{N}. \quad (5.7)$$

Let $\varphi_K : \mathbb{R} \rightarrow \mathbb{R}$, $K \geq 1$ be a Lipschitz function such that

$$\chi_{[0, K-1]}(y) \leq \varphi_K(y) \leq \chi_{[0, K]}(y) \quad \text{for all } x \in \mathbb{R}.$$

We are going to show that there exists $K \geq 1$ such that

$$\mathbb{E}_{\mu_*}(S_m^2(1 - \varphi_K(|S_m|))) < \varepsilon / (9l^3) \quad \text{for all } m \geq 1.$$

By Proposition 5.3.1 we have

$$\mathbb{E}_{\mu_*} S_m^2 \rightarrow \tilde{\sigma}^2 \quad \text{as } m \rightarrow \infty$$

and

$$\mathbb{E}_{\mu_*}(S_m^2 \varphi_K(|S_m|)) \rightarrow \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} x^2 \varphi_K(|y|) e^{-y^2/(2\tilde{\sigma}^2)} dy \quad \text{as } m \rightarrow \infty.$$

In the case when $\tilde{\sigma} = 0$ we set $e^{-y^2/(2\tilde{\sigma}^2)} = 0$.

Since the integral on the right hand side converges to σ^2 as $K \rightarrow \infty$, we obtain that there exist constants m_0 and k_0 such that

$$\mathbb{E}_{\mu_*}(S_m^2(1 - \varphi_{K_0}(|S_m|))) = \mathbb{E}_{\mu_*}S_m^2 - \mathbb{E}_{\mu_*}(S_m^2\varphi_{K_0}(|S_m|)) < \varepsilon/(9l^3) \quad \text{for all } m \geq m_0.$$

Enlarging K_0 if necessary we obtain

$$\mathbb{E}_{\mu_*}(S_m^2(1 - \varphi_{K_0}(|S_m|))) < \varepsilon/(9l^3) \quad \text{for all } m \geq 1. \quad (5.8)$$

Observe that the function $y \rightarrow \Theta(y)$ given by the formula $\Theta(y) := y^2\varphi_{K_0}(|y|)$ is a Lipschitz function. Combining (5.6) - (5.8), we obtain that for $n \geq \max\{N_0, N_1\}$ we have

$$\begin{aligned} \mathbb{E}_{P^n\delta_x}[S_m^2\chi_{[K_0,+\infty)}(|S_m|)] &\leq \mathbb{E}_{P^n\delta_x}S_m^2(1 - \varphi_{K_0}(|S_m|)) \\ &\leq |\mathbb{E}_{P^n\delta_x}S_m^2 - \mathbb{E}_{\mu_*}S_m^2| + |\mathbb{E}_{P^n\delta_x}\Theta(S_m) - \mathbb{E}_{\mu_*}\Theta(S_m)| + \mathbb{E}_{\mu_*}(S_m^2(1 - \varphi_{K_0}(|S_m|))) \\ &\leq \varepsilon/(9l^3) + \varepsilon/(9l^3) + \varepsilon/(9l^3) = \varepsilon/(3l^3). \end{aligned}$$

Having this we may show that (5.5) holds with $K = lK_0$. Set $m_i = \max\{j : i + jl \leq m\}$ for $i = 0, \dots, l-1$. We have

$$\begin{aligned} &\mathbb{E}_{P^n\delta_x} \left[\left(\frac{g(X_0) + \dots + g(X_m)}{\sqrt{m}} \right)^2 \chi_{[K,+\infty)} \left(\left| \frac{g(X_0) + \dots + g(X_m)}{\sqrt{m}} \right| \right) \right] \\ &\leq l^2 \mathbb{E}_{P^n\delta_x} \left[\max_{0 \leq i \leq l-1} \left(\frac{g(X_i) + g(X_{i+l}) + \dots + g(X_{i+m_i l})}{\sqrt{m_i}} \right)^2 \chi_{[K,+\infty)} \left(\left| \frac{g(X_0) + \dots + g(X_m)}{\sqrt{m}} \right| \right) \right] \\ &\leq l^2 \mathbb{E}_{P^n\delta_x} \max_{0 \leq i \leq l-1} \left[\left(\frac{g(X_i) + g(X_{i+l}) + \dots + g(X_{i+m_i l})}{\sqrt{m_i}} \right)^2 \chi_{[K_0,+\infty)} \left(\left| \frac{g(X_i) + \dots + g(X_{i+m_i l})}{\sqrt{m_i}} \right| \right) \right] \\ &\leq l^2 \mathbb{E}_{P^n\delta_x} \left[\left(\frac{g(X_0) + g(X_l) + \dots + g(X_{m_0 l})}{\sqrt{m_1}} \right)^2 \chi_{[K_0,+\infty)} \left(\left| \frac{g(X_0) + g(X_l) + \dots + g(X_{m_0 l})}{\sqrt{m_0}} \right| \right) \right] \\ &\quad + l^2 \mathbb{E}_{P^n\delta_x} \left[\left(\frac{g(X_1) + g(X_{1+l}) + \dots + g(X_{1+m_1 l})}{\sqrt{m_1}} \right)^2 \chi_{[K_0,+\infty)} \left(\left| \frac{g(X_1) + g(X_{1+l}) + \dots + g(X_{1+m_1 l})}{\sqrt{m_1}} \right| \right) \right] \\ &\quad + \dots + l^2 \mathbb{E}_{P^n\delta_x} \left[\left(\frac{g(X_{l-1}) + g(X_{2l-1}) + \dots + g(X_{l-1+m_{l-1} l})}{\sqrt{m_{l-1}}} \right)^2 \right. \\ &\quad \left. \times \chi_{[K_0,+\infty)} \left(\left| \frac{g(X_{l-1}) + g(X_{2l-1}) + \dots + g(X_{l-1+m_{l-1} l})}{\sqrt{m_{l-1}}} \right| \right) \right] \leq l^2 l \varepsilon / l^3 = \varepsilon. \end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, the proof is complete. $\square$

We are now in a position to show the following:

Proposition 5.4.2. *Let P be a Markov operator that satisfies (A1) - (A3) and let (X_n) be a Markov chain corresponding to P . Let l_0 be a positive integer such that $Cq^{l_0} < 1$, where the constants C, q are given by (5.1) and let $l \geq l_0$ be given. Assume that $Y_n = X_{nl}$ for $n \geq 0$. If $g : X \rightarrow \mathbb{R}$ is a bounded Lipschitz function, then*

$$\lim_{m \rightarrow \infty} \mathbb{E}_x \left(\frac{g(Y_0) + g(Y_1) + \dots + g(Y_m)}{\sqrt{m}} \right)^2 = \tilde{\sigma}^2,$$

where

$$\tilde{\sigma}^2 = \lim_{m \rightarrow \infty} \mathbb{E}_{\mu_*} \left(\frac{g(X_0) + g(X_1) + \dots + g(X_n)}{\sqrt{n}} \right)^2.$$

Proof. First observe that for any $k \geq 1$ we have

$$\left| \mathbb{E}_x \left(\frac{g(Y_0) + g(Y_1) + \dots + g(Y_n)}{\sqrt{n}} \right)^2 - \mathbb{E}_x \left(\frac{g(Y_k) + g(Y_{k+1}) + \dots + g(Y_{n+k})}{\sqrt{n}} \right)^2 \right| \rightarrow 0$$

as $n \rightarrow \infty$. Further we have

$$\mathbb{E}_x \left(\frac{g(Y_k) + g(Y_{k+1}) + \dots + g(Y_{n+k})}{\sqrt{n}} \right)^2 = \mathbb{E}_{P^k \delta_x} \left(\frac{g(Y_0) + g(Y_1) + \dots + g(Y_n)}{\sqrt{n}} \right)^2,$$

by the Markov property. On the other hand, the Markov chain (Y_n) corresponds to the operator P^l and the assumptions of Lemma 5.4.2 are satisfied with P and (X_n) replaced with P^l and (Y_n) , respectively. Thus the functions

$$X \ni x \rightarrow \mathbb{E}_x \left(\frac{g(Y_0) + g(Y_1) + \dots + g(Y_n)}{\sqrt{n}} \right)^2 \quad \text{for } n \geq 0$$

are Lipschitzian with the Lipschitz constant bounded by, say $\tilde{M}$, independent of n . Consequently, we obtain

$$\begin{aligned} & \left| \mathbb{E}_{P^k \delta_x} \left(\frac{g(Y_0) + g(Y_1) + \dots + g(Y_n)}{\sqrt{n}} \right)^2 - \mathbb{E}_{\mu_*} \left(\frac{g(Y_0) + g(Y_1) + \dots + g(Y_n)}{\sqrt{n}} \right)^2 \right| \\ & \leq \tilde{M} \|P^k \delta_x - \mu_*\| \leq \tilde{M} C q^k \|\delta_x - \mu_*\|. \end{aligned}$$

Hence we finally have the desired convergence. □

5.5 The Central Limit Theorem

Now we are in a position to formulate and prove the main theorem of this paper.

Theorem 5.5.1. (*Central Limit Theorem*) Let P be a Markov operator that satisfies **(A1)** - **(A3)** and let (X_n) be a Markov chain corresponding to P . Assume that $X_0 = x$ for $x \in X$. Then for any bounded Lipschitz function $g : X \rightarrow \mathbb{R}$ such that $\int_X g d\mu_* = 0$, where μ_* is the unique invariant distribution for P , the sequence of random vectors (X_n) satisfies

$$\lim_{n \rightarrow \infty} P \left(\frac{g(X_0) + g(X_1) + \dots + g(X_n)}{\sqrt{n}} < a \right) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^a e^{-\frac{y^2}{2\sigma^2}} dy \quad \text{for all } a \in \mathbb{R},$$

if

$$\sigma^2 = \lim_{n \rightarrow \infty} \mathbb{E}_{\mu_*} \left(\frac{g(X_0) + g(X_1) + \dots + g(X_n)}{\sqrt{n}} \right)^2 > 0.$$

Otherwise, if $\sigma = 0$, the sequence

$$\frac{g(X_0) + g(X_1) + \dots + g(X_n)}{\sqrt{n}} \quad \text{for } n \geq 1$$

converges in distribution to 0.

Proof. To prove the theorem we show that for any $x \in X$

$$\lim_{n \rightarrow \infty} \mathbb{E}_x \exp \left[it \left(\frac{g(X_0) + g(X_1) + \dots + g(X_n)}{\sqrt{n}} \right) \right] = e^{-\sigma^2 t^2 / 2} \quad \text{for } t \in \mathbb{R}.$$

Without loss of generality we may assume that g is bounded by 1 and its Lipschitz constant is also bounded by 1. We will make use of the following formula:

$$e^{ia} = 1 + ia - a^2/2 - R(a)a^2,$$

where $|R(a)| \leq 1$ and $\lim_{a \rightarrow 0} R(a) = R(0) = 0$.

Fix $x \in X$, $t \in \mathbb{R} \setminus \{0\}$ and $\varepsilon > 0$. Let $k_0 \in \mathbb{N}$ and $\eta > 0$ be such that

$$|(1 - \sigma_0^2 t^2 / 2k)^k - e^{-\sigma_0^2 t^2 / 2}| < \varepsilon \quad \text{for } |\sigma_0 - \sigma| < \eta \text{ and } k \geq k_0.$$

Set $D := \sup_{n \geq 1} \int_X \int_X \rho(u, z) P^n \delta_x(du) \mu_*(dz)$ and observe that for any x_0 we have according

to Proposition 5.2.1

$$D \leq \sup_{n \geq 1} \int_X \rho(u, x_0) P^n \delta_x(du) + \int_X \rho(z, x_0) \mu_*(dz) < +\infty.$$

For given x, t, ε , by Lemma 5.4.3, we choose $K > 0$ and $N_0 \in \mathbb{N}$ such that

$$\mathbb{E}_{P^n \delta_x} \left[\left(\frac{g(X_0) + \dots + g(X_m)}{\sqrt{m}} \right)^2 \chi_{[K, +\infty)} \left(\left| \frac{g(X_0) + \dots + g(X_m)}{\sqrt{m}} \right| \right) \right] \leq \varepsilon / (2t^2) \quad (5.9)$$

for $n \geq N_0$ and $m \in \mathbb{N}$.

Fix $l \geq N_0$ such that

$$q^l C D |t| ((1 - q)^{-1} + M |t|) < \varepsilon \quad \text{and} \quad q^l C < 1, \quad (5.10)$$

where the constants q, C are given by condition (5.1) and the constant M is given by Lemma 5.4.2. Since

$$\sup_{m, n \geq 1} \mathbb{E}_{P^n \delta_x} \left[\left(\frac{g(X_0) + \dots + g(X_m)}{\sqrt{m}} \right)^2 \right] < \infty$$

one may choose $k \geq k_0$ such that

$$\begin{aligned} & \mathbb{E}_{P^n \delta_x} \left(\left(\frac{g(X_0) + \dots + g(X_m)}{\sqrt{m}} \right)^2 \left| R \left(\frac{\theta t}{\sqrt{k}} \frac{g(X_0) + \dots + g(X_m)}{\sqrt{m}} \right) \right| \right. \\ & \times \chi_{[0, K]} \left(\left| \frac{g(X_0) + \dots + g(X_m)}{\sqrt{m}} \right| \right) \Bigg) \leq \varepsilon / (2t^2) \quad \text{for all } m, n \geq 1 \text{ and } \theta \in [0, 1], \end{aligned} \quad (5.11)$$

by the fact that $|R(a)| \rightarrow 0$ as $a \rightarrow 0$ and $|\sigma_m - \sigma| < \eta$ for $m \geq k$, where

$$\sigma_m^2 = \mathbb{E}_{\mu_*} \left(\frac{g(X_0) + \dots + g(X_{m-l})}{\sqrt{m}} \right)^2 = \mathbb{E}_{\mu_*} \left(\frac{g(X_l) + \dots + g(X_m)}{\sqrt{m}} \right)^2.$$

This can be done due to the fact that σ_m tends to σ as $m \rightarrow +\infty$. For positive integers u, v and $w, v \geq l$, where l is such that condition (5.10) holds, we set

$$I_{u,v,w} = \mathbb{E}_x \exp \left[it \left(\frac{g(X_l) + \dots + g(X_w) + \dots + g(X_{(u-1)w+l}) + \dots + g(X_{uw})}{\sqrt{uw}} \right) \right]$$

and

$$I_w = \mathbb{E}_x \exp \left[it \left(\frac{g(X_0) + \dots + g(X_w)}{\sqrt{w}} \right) \right].$$

Note that $\lim_{n \rightarrow \infty} |I_n - I_{m,m,[n/m]}| = 0$ for every $m \geq 1$. So let us consider in further detail the terms $I_{j,k,m}$ for given $k, j \leq k$ and $m \geq k$. In the following we will use the notation

$$\mathbb{E}_{|X_{(j-1)m}}[\cdot] = \mathbb{E}[\cdot | \mathcal{F}_{(j-1)m}].$$

We have

$$\begin{aligned} I_{j,k,m} &= \mathbb{E}_x \exp \left[it \left(\frac{g(X_l) + \dots + g(X_m) + \dots + g(X_{(j-1)m+l}) + \dots + g(X_{jm})}{\sqrt{km}} \right) \right] \\ &= \mathbb{E}_x \left(\exp it \left[\left(\frac{g(X_l) + \dots + g(X_m) + \dots + g(X_{(j-2)m+l}) + \dots + g(X_{(j-1)m})}{\sqrt{km}} \right) \right] \right. \\ &\quad \times \mathbb{E}_{|X_{(j-1)m}} \exp it \left(\frac{g(X_{(j-1)m+l}) + \dots + g(X_{jm})}{\sqrt{km}} \right) \left. \right) \\ &= \mathbb{E}_x \left(\exp it \left[\left(\frac{g(X_l) + \dots + g(X_m) + \dots + g(X_{(j-2)m+l}) + \dots + g(X_{(j-1)m})}{\sqrt{km}} \right) \right] \right. \\ &\quad \left(1 + \frac{it}{\sqrt{km}} \mathbb{E}_{|X_{(j-1)m}} (g(X_{(j-1)m+l}) + \dots + g(X_{jm})) - \frac{t^2}{2k} \mathbb{E}_{|X_{(j-1)m}} \left(\frac{g(X_{(j-1)m+l}) + \dots + g(X_{jm})}{\sqrt{m}} \right)^2 \right. \\ &\quad \left. \left. - \frac{t^2}{k} \mathbb{E}_{|X_{(j-1)m}} \left[\left(\frac{g(X_{(j-1)m+l}) + \dots + g(X_{jm})}{\sqrt{m}} \right)^2 \right] R \left(\frac{t}{\sqrt{km}} (g(X_{(j-1)m+l}) + \dots + g(X_{jm})) \right) \right] \right) \right). \end{aligned} \quad (5.12)$$

Since $\mathbb{E}_{\mu_*} g(X_i) = 0$, by **(A2)** we obtain

$$\begin{aligned} |\mathbb{E}_u (g(X_l) + \dots + g(X_m))| &= |\mathbb{E}_{P^l \delta_u} (g(X_0) + \dots + g(X_{m-l})) - \mathbb{E}_{\mu_*} (g(X_0) + \dots + g(X_{m-l}))| \\ &\leq \sum_{i=l}^m |\mathbb{E}_{P^i \delta_u} (g(X_0)) - \mathbb{E}_{\mu_*} (g(X_0))| \leq \sum_{i=l}^m \|P^i \delta_u - \mu_*\| \\ &\leq q^l C(1-q)^{-1} \|\delta_u - \mu_*\| \leq q^l C(1-q)^{-1} \int_X \rho(u, z) \mu_*(dz) \end{aligned}$$

and hence we have

$$\begin{aligned} \mathbb{E}_x \left| \frac{it}{\sqrt{km}} \mathbb{E}_{|X_{(j-1)m}} (g(X_{(j-1)m+l}) + \dots + g(X_{jm})) \right| &\leq q^l C(1-q)^{-1} \frac{|t|}{\sqrt{km}} \mathbb{E}_x \left[\int_X \rho(X_{(j-1)m}, z) \mu_*(dz) \right] \\ &\leq q^l C(1-q)^{-1} \frac{|t|}{\sqrt{km}} \int_X \int_X \rho(u, z) \mu_*(dz) P^{(j-1)m} \delta_x(du) \leq q^l CD(1-q)^{-1} \frac{|t|}{\sqrt{km}}. \end{aligned}$$

On the other hand, by Lemma 5.4.2 and **(A2)** we have

$$\begin{aligned}
 & \left| \mathbb{E}_u \left(\frac{g(X_l) + \dots + g(X_m)}{\sqrt{m}} \right)^2 - \mathbb{E}_{\mu_*} \left(\frac{g(X_l) + \dots + g(X_m)}{\sqrt{m}} \right)^2 \right| \\
 &= \left| \mathbb{E}_{P^l \delta_u} \left(\frac{g(X_0) + \dots + g(X_{m-l})}{\sqrt{m}} \right)^2 - \mathbb{E}_{\mu_*} \left(\frac{g(X_0) + \dots + g(X_{m-l})}{\sqrt{m}} \right)^2 \right| \\
 &\leq q^l C M \frac{m-l}{m} \|\delta_u - \mu_*\| \leq q^l C M \int_X \rho(u, z) \mu_*(dz),
 \end{aligned}$$

where M is given in Lemma 5.4.2 and consequently we obtain

$$\begin{aligned}
 & \mathbb{E}_x \left| \frac{t^2}{k} \mathbb{E}_{|X_{(j-1)m}} \left(\frac{g(X_{(j-1)m+l}) + \dots + g(X_{jm})}{\sqrt{m}} \right)^2 - \frac{t^2}{k} \mathbb{E}_{\mu_*} \left(\frac{g(X_l) + \dots + g(X_m)}{\sqrt{m}} \right)^2 \right| \\
 &\leq q^l M \frac{t^2}{k} \mathbb{E}_x \left[\int_X \rho(X_{(j-1)m}, z) \mu_*(dz) \right] \leq q^l C M \frac{t^2}{k} \int_X \int_X \rho(u, z) \mu_*(dz) P^{(j-1)m} \delta_x(du) \\
 &\leq q^l C M D \frac{t^2}{k}.
 \end{aligned}$$

Finally conditions (5.9) and (5.11) will allow us to evaluate the term

$$\mathbb{E}_x \left| \frac{t^2}{k} \mathbb{E}_{|X_{(j-1)m}} \left[\left(\frac{g(X_{(j-1)m+l}) + \dots + g(X_{jm})}{\sqrt{m}} \right)^2 R \left(\frac{t}{\sqrt{km}} (g(X_{(j-1)m+l}) + \dots + g(X_{jm})) \right) \right] \right|.$$

Indeed, we have

$$\begin{aligned}
 & \mathbb{E}_x \left| \frac{t^2}{k} \mathbb{E}_{|X_{(j-1)m}} \left[\left(\frac{g(X_{(j-1)m+l}) + \dots + g(X_{jm})}{\sqrt{m}} \right)^2 R \left(\frac{t}{\sqrt{km}} (g(X_{(j-1)m+l}) + \dots + g(X_{jm})) \right) \right] \right| \\
 &\leq \frac{t^2}{k} \mathbb{E}_x \left[\mathbb{E}_{|X_{(j-1)m}} \left| \left(\frac{g(X_{(j-1)m+l}) + \dots + g(X_{jm})}{\sqrt{m}} \right)^2 R \left(\frac{t}{\sqrt{km}} (g(X_{(j-1)m+l}) + \dots + g(X_{jm})) \right) \right| \right] \\
 &= \frac{t^2}{k} \mathbb{E}_x \left[\left| \left(\frac{g(X_{(j-1)m+l}) + \dots + g(X_{jm})}{\sqrt{m}} \right)^2 R \left(\frac{t}{\sqrt{km}} (g(X_{(j-1)m+l}) + \dots + g(X_{jm})) \right) \right| \right]
 \end{aligned}$$

$$\begin{aligned}
&= \frac{t^2}{k} \mathbb{E}_{P^{(j-1)m+l}\delta_x} \left[\left(\frac{g(X_0) + \dots + g(X_{m-l})}{\sqrt{m}} \right)^2 \left| R \left(\frac{t}{\sqrt{km}} (g(X_0) + \dots + g(X_{m-l})) \right) \right| \right] \\
&\leq \frac{t^2}{k} \mathbb{E}_{P^{(j-1)m+l}\delta_x} \left[\left(\frac{g(X_0) + \dots + g(X_{m-l})}{\sqrt{m-l}} \right)^2 \chi_{[K,+\infty)} \left(\left| \frac{g(X_0) + \dots + g(X_{m-l})}{\sqrt{m-l}} \right| \right) \right] \\
&+ \frac{t^2}{k} \mathbb{E}_{P^{(j-1)m+l}\delta_x} \left(\left(\frac{g(X_0) + \dots + g(X_{m-l})}{\sqrt{m-l}} \right)^2 \left| R \left(\sqrt{\frac{m-l}{m}} \frac{t}{\sqrt{k}} \frac{g(X_0) + \dots + g(X_{m-l})}{\sqrt{m-l}} \right) \right| \right) \\
&\times \chi_{[0,K]} \left(\left| \frac{g(X_0) + \dots + g(X_{m-l})}{\sqrt{m-l}} \right| \right) \stackrel{\text{by (5.9) and (5.11)}}{\leq} \frac{t^2}{k} \varepsilon / (2t^2) + \frac{t^2}{k} \varepsilon / (2t^2) = \varepsilon / k.
\end{aligned}$$

Now from (5.12) it follows that

$$|I_{j,k,m} - I_{j-1,k,m}(1 - \sigma_m^2 t^2 / 2k)| \leq q^l C D (1 - q)^{-1} \frac{|t|}{\sqrt{km}} + q^l C M D \frac{t^2}{k} + \frac{\varepsilon}{k} \leq 2\varepsilon / k \quad \text{for } j = 1, \dots, k,$$

by condition (5.10) and the fact that $m \geq k$. Iterating this formula k -times we obtain

$$|I_{k,k,m} - (1 - \sigma_m^2 t^2 / 2k)^k| \leq 2\varepsilon$$

and consequently

$$|I_{k,k,m} - e^{-\sigma^2 t^2 / 2}| \leq 3\varepsilon \quad \text{for all } m \text{ sufficiently large.}$$

Since $\varepsilon > 0$ was arbitrary, we obtain that $\lim_{n \rightarrow \infty} |I_{k,k,[n/k]} - I_n| = 0$, which completes the proof. $\square$

5.6 Example

Let (X, ρ) be a Polish space and let $(T, \mathcal{A})$ be a measurable space. Let $\nu : \mathcal{A} \rightarrow [0, \infty)$ be some measure on T . Let $p : T \times X \rightarrow [0, \infty)$ be a measurable function such that

$$\int_T p(t, x) \nu(dt) = 1 \quad \text{for } x \in X.$$

We shall assume that $p(t, \cdot) : X \rightarrow [0, \infty)$ for $t \in T$ is a Lipschitz function. Denote its Lipschitz constant by $k(t)$.

Let $\pi_t : X \times \mathcal{B}(X) \rightarrow [0, 1]$, $t \in T$, be a transition probability and let P_t and U_t denote the corresponding Markov operator and its dual, respectively. We shall consider the Markov chain (X_n) on some space $(\Omega, \mathcal{F}, \mathbb{P})$ corresponding to the action of randomly chosen tran-

sition probabilities with probability depending on position, i.e. the Markov chain (X_n) given by the formula:

$$\mathbb{P}[X_{n+1} \in \cdot | X_n = x] = \pi(x, \cdot) = \int_T p(t, x) \pi_t(x, \cdot) \nu(dt) \quad \text{for } x \in X \text{ and } n \geq 1. \quad (5.13)$$

This model generalizes some random dynamical systems studied in [HSl16].

If the distribution of X_0 is equal to μ , then the distribution of X_n is given by $P^n \mu$, where $P : \mathcal{M} \rightarrow \mathcal{M}$ is of the form

$$P\mu(A) = \int_T \int_X p(t, x) \pi_t(x, A) \mu(dx) \nu(dt) \quad \text{for } A \in \mathcal{B}(X). \quad (5.14)$$

Theorem 5.6.1. *Assume that π_t for $t \in T$ are transition probabilities and there exists $q < 1$ such that for the Markov operators P_t corresponding to π_t we have*

$$\|P_t \mu_1 - P_t \mu_2\| \leq q \|\mu_1 - \mu_2\| \quad \text{for } \mu_1, \mu_2 \in \mathcal{M}_1^1.$$

Moreover, for some $x_0 \in X$ and the operator U_t dual to P_t , $t \in T$, we have

$$U_t[\rho(\cdot, x_0)]^2(x) \leq a[\rho(x, x_0)]^2 + b$$

for some $a < 1$ and $b > 0$. Finally set

$$\gamma(t) := \sup_{x \in X} \int_X \rho(z, x_0) \pi_t(x, dz)$$

for some $x_0 \in X$ and assume that

$$\int_T \gamma(t) k(t) \nu(dt) < 1 - q$$

and

$$\int_T p(t, x) \gamma(t) \nu(dt) < \infty \quad \text{for } x \in X.$$

Let (X_n) be a Markov chain given by (5.13). Assume that $X_0 = x$ for $x \in X$. Then for any bounded Lipschitz function $g : X \rightarrow \mathbb{R}$ such that $\int_X g d\mu_* = 0$, where μ_* is the unique invariant distribution for P , the sequence of random vectors (X_n) satisfies

$$\lim_{n \rightarrow \infty} P \left(\frac{g(X_0) + g(X_1) + \dots + g(X_n)}{\sqrt{n}} < v \right) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^v e^{-\frac{y^2}{2\sigma^2}} dy \quad \text{for } v \in \mathbb{R},$$

if

$$\sigma^2 = \lim_{n \rightarrow \infty} \mathbb{E}_{\mu_*} \left(\frac{g(X_0) + g(X_1) + \dots + g(X_n)}{\sqrt{n}} \right)^2 > 0.$$

Otherwise, if $\sigma = 0$, the sequence

$$\frac{g(X_0) + g(X_1) + \dots + g(X_n)}{\sqrt{n}} \quad \text{for } n \geq 1$$

converges in distribution to 0.

Proof. From Theorem 5.5.1 it follows that to finish the proof it is enough to show that the Markov operator P given by (5.14) satisfies conditions **(A1)** - **(A3)**.

Observe that P is a Feller operator and its dual is of the form

$$Uf(x) = \int_T \int_X p(t, x) f(z) \pi_t(x, dz) \nu(dt) \quad \text{for } f \in C_b(X). \quad (5.15)$$

The operator U may be extended to an arbitrary positive unbounded function f and then it is also given by formula (5.15). The operator U can be, in fact, extended to the space of all Lipschitz functions. To show this, observe that for $x_0 \in X$ we have

$$\begin{aligned} U\rho(\cdot, x_0)(x) &= \int_T p(t, x) \int_X \rho(z, x_0) \pi_t(x, dz) \nu(dt) \leq \int_T p(t, x) \gamma(t) \nu(dt) \\ &\leq \int_T |p(t, x) - p(t, x_0)| \gamma(t) \nu(dt) + \int_T p(t, x_0) \gamma(t) \nu(dt) \\ &\leq \int_T k(t) \gamma(t) \nu(dt) + \int_T p(t, x_0) \gamma(t) \nu(dt) < \infty. \end{aligned}$$

Therefore for any Lipschitz function $f : X \rightarrow \mathbb{R}$ with Lipschitz constant L and $x \in X$ we have

$$|Uf(x)| \leq U|f|(x) \leq |f(x_0)| + LU\rho(\cdot, x_0)(x) < \infty,$$

by the fact that the Lipschitz constant of the function $|f|$ is bounded by L .

Set $\hat{q} := \int_T \gamma(t) k(t) \nu(dt) + q < 1$ and let $f : X \rightarrow \mathbb{R}$ be an arbitrary Lipschitz function with Lipschitz constant bounded by 1 and such that $f(x_0) = 0$. We show that then Uf is a

Lipschitz function with Lipschitz constant bounded by $\hat{q}$. Indeed, for any $x, y \in X$ we have

$$\begin{aligned}
 |Uf(x) - Uf(y)| &= \left| \int_T \int_X p(t, x) f(z) \pi_t(x, dz) \nu(dt) - \int_T \int_X p(t, y) f(z) \pi_t(y, dz) \nu(dt) \right| \\
 &\leq \int_T \int_X |p(t, x) - p(t, y)| |f(z)| \pi_t(x, dz) \nu(dt) \\
 &\quad + \int_T p(t, y) \left| \int_X f(z) \pi_t(x, dz) - \int_X f(z) \pi_t(y, dz) \right| \nu(dt) \\
 &\leq \int_T k(t) \left(\int_X \rho(z, x_0) \pi_t(x, dz) \right) \nu(dt) \cdot \rho(x, y) + q \int_T p(t, y) \nu(dt) \cdot \rho(x, y) \\
 &\leq \left(\int_T k(t) \gamma(t) \nu(dt) + q \right) \rho(x, y) = \hat{q} \rho(x, y).
 \end{aligned}$$

To verify **(A1)** we compute

$$\begin{aligned}
 \int_X \rho(x, x_0) P\mu(dx) &= \int_X U\rho(\cdot, x_0)(x) \mu(dx) \\
 &= \int_X (U\rho(\cdot, x_0)(x) - U\rho(\cdot, x_0)(x_0)) \mu(dx) + U\rho(\cdot, x_0)(x_0) \\
 &\leq \hat{q} \int_X \rho(x, x_0) \mu(dx) + U\rho(\cdot, x_0)(x_0) \leq \int_X \rho(x, x_0) \mu(dx) + U\rho(\cdot, x_0)(x_0) < \infty
 \end{aligned}$$

for $\mu \in \mathcal{M}_1^1$.

Observe that for any $\mu_1, \mu_2 \in \mathcal{M}_1^1$ we have

$$\|\mu_1 - \mu_2\| = \sup \left| \int_X f(x) \mu_1(dx) - \int_X f(x) \mu_2(dx) \right|,$$

where the supremum is taken over all Lipschitz functions $f : X \rightarrow \mathbb{R}$ with Lipschitz constant bounded by 1 and such that $f(x_0) = 0$. Indeed for any $f : X \rightarrow \mathbb{R}$ with Lipschitz constant bounded by 1, we have

$$\left| \int_X f(x) \mu_1(dx) - \int_X f(x) \mu_2(dx) \right| = \left| \int_X (f(x) - f(x_0)) \mu_1(dx) - \int_X (f(x) - f(x_0)) \mu_2(dx) \right|.$$

To show that **(A2)** holds fix $\mu_1, \mu_2 \in \mathcal{M}_1^1$. Then we have

$$\begin{aligned}
 \|P\mu_1 - P\mu_2\| &= \sup \left| \int_X f(x) P\mu_1(dx) - \int_X f(x) P\mu_2(dx) \right| \\
 &= \hat{q} \sup \left| \int_X \hat{q}^{-1} Uf(x) \mu_1(dx) - \int_X \hat{q}^{-1} Uf(x) \mu_2(dx) \right| \leq \hat{q} \sup \left| \int_X f(x) \mu_1(dx) - \int_X f(x) \mu_2(dx) \right| \\
 &= \hat{q} \|\mu_1 - \mu_2\|
 \end{aligned}$$

Here the supremum is taken over all Lipschitz functions $f : X \rightarrow \mathbb{R}$ with Lipschitz constant bounded by 1 and such that $f(x_0) = 0$, by the fact that $\hat{q}^{-1}Uf$ is a Lipschitz function with Lipschitz constant bounded by 1. Hence **(A2)** holds.

Finally, we show that **(A3)** holds with $\mu = \delta_{x_0}$. To do this it is enough to prove that

$$U[\rho(\cdot, x_0)(x)]^2 \leq a[\rho(x, x_0)]^2 + b \quad (5.16)$$

for some $a < 1$ and $b > 0$. Indeed, iterating the above formula we have

$$U^n[\rho(\cdot, x_0)]^2(x) \leq a^n[\rho(x, x_0)]^2 + b(1 - a)^{-1} \quad \text{for } n \geq 1 \text{ and } x \in X,$$

consequently,

$$\begin{aligned} & \sup_{n \geq 1} \int_X [\rho(x, x_0)]^2 P^n \delta_{x_0}(\mathrm{d}x) \\ & \leq \sup_{n \geq 1} \left(a^n \int_X [\rho(x, x_0)]^2 \delta_{x_0}(\mathrm{d}x) + b(1 - a)^{-1} \right) = b(1 - a)^{-1} < \infty. \end{aligned}$$

and **(A3)** will hold. To verify condition (5.16) we compute

$$\begin{aligned} U[\rho(\cdot, x_0)(x)]^2 &= \int_T p(t, x_0) \int_X [\rho(z, x_0)]^2 \pi_t(x_0, \mathrm{d}z) \nu(\mathrm{d}t) = \int_T p(t, x_0) U_t[\rho(\cdot, x_0)]^2(x) \nu(\mathrm{d}t) \\ &\leq \int_T p(t, x_0) (a[\rho(x, x_0)]^2 + b) \nu(\mathrm{d}t) = a[\rho(x, x_0)]^2 + b. \end{aligned}$$

This completes the proof. □

