Martingale approximation and the central limit theorem for random dynamical systems of affine transformations

M. Denker and Y. Iwata

REPORT No. 8, 2009/2010, spring
ISSN 1103-467X
ISRN IML-R-8-09/10-SE+spring
Martingale approximation and the central limit theorem for random dynamical systems of affine transformations

Manfred Denker∗ and Yukiko Iwata†

May 28, 2010

Abstract

Let $X_{n+1} = A_n X_n + B_n$ be a stochastic recurrence relation of real valued variables $X_n$ with stationary and ergodic coefficient process $(A_n, B_n)$. We show a central limit theorem for the partial sums of the variables $X_n$ when the initial variable $X_0$ is tempered. The problem is solved by writing the equation as a random dynamical system of affine transformations and proving a CLT for its stationary solution.

AMS Classification: 37H10, 37A50, 37A05, 60F05, 60G10, 62M10

Key Words : ergodic theorem, central limit theorem, random dynamical system, affine transformation, stochastic recurrence equation, autoregressive process.

1 Introduction

Let $(A_n, B_n)_{n \geq 0}$ be an $\mathbb{R}^{2d} \times \mathbb{R}^d$-valued strictly stationary sequence of $d \times d$ random matrices $A_n$ and $\mathbb{R}^d$-valued random vectors $B_n$. The stochastic recurrence equation

$$X_{n+1} = A_n X_n + B_n, \quad n \geq 0,$$

where $X_0$ is some initial random variable, appears frequently in the literature, almost exclusively when the process $(A_n, B_n)$ is independent and identically

∗Research in part supported by Institut Mittag-Leffler, Stockholm.
†Research supported by JSPS International Training Program (ITP) at Department of Mathematics, Hokkaido University.
distributed. In this note we study this equation under stationarity assumptions when \( d = 1 \).

The equation (1) has been investigated even before the seminal paper [11] by Furstenberg and Kesten on the behavior of products of random matrices. More recently and related to the present goal, the equation was studied in the work of Verwaat [25], Brandt [9], Bougerol and Picard [7], Goldie and Maller [12], Aue et al. [2] among others.

The equation (1) can be understood as an equality in distribution. In [25] and [9] it was shown that a stationary distribution exists when the process \((A_n, B_n)\) is independent and identically distributed (i.i.d.). In this case (1) is a special case of Röschler’s contraction method (see [20] for a survey of this method).

In this note we are considering strong solutions of equation (1) which means (strictly) stationary solutions of this equation. If \( A_n \) is non-random and \( B_n \) is i.i.d. the equation reduces to an autoregressive process (AR-process) of order 1, and if, moreover, \( A_n = 1 \) the process is a random walk on \( \mathbb{R}^d \) for some \( d \geq 1 \). Another special class is given by augmented GARCH processes (see [2]) where \( A_n \) is a 'function' of \( B_n \). Applications in other disciplines are numerous, we just mention [5] for an application in queueing theory and [6] or [23] for applications in econometrics. The existence of a stationary solution to equation (1) was shown in [7] for i.i.d. processes \((A_n, B_n)\) and later in [2] it was further strengthened, including GARCH processes.

The non i.i.d. case of \((A_n, B_n)\) has been considered in [22] and [21], where a distributional solution is obtained for some Markovian type coefficients \( A_n \) and \( B_n \). More importantly, it was noticed in [7] that the equation (1) can be reformulated in terms of random maps, for details see [1]. In fact, one may introduce the random linear transformation by the cocycle

\[
\varphi : \mathbb{Z}_+ \times \Omega \times \mathbb{R}^d \to \mathbb{R}^d
\]

defined by

\[
\begin{align*}
\varphi(1, \omega, x) & = a(\omega)x + b(\omega) \\
\varphi(n + m, \omega, x) & = \varphi(n, \theta^m \omega, \varphi(m, \omega, x)) \quad n, m \geq 0,
\end{align*}
\]

where \((\Omega, \mathcal{F}, \theta, P)\) denotes a \( P \)-probability preserving dynamical system with invariant transformation \( \theta : \Omega \to \Omega \), \( a \) is a measurable function with values in the space of \( d \times d \)-matrices and \( b \) is a measurable \( d \)-dimensional vector function. Setting \( A_n = a \circ \theta^n \) and \( B_n = b \circ \theta^n \) we rediscover equation (1).
stationary solution to (1) is then given by a measurable function $x^* : \Omega \to \mathbb{R}^d$ such that

$$x^* \circ \theta = \varphi(1, \cdot, x^*).$$

A general fixed point theorem for random transformations can be found in [24]. In the present context of equation (2) for $d = 1$ the solution is constructed in [23] based on [1], Proposition 4.1.3. For completeness, in Section 3, we provide a detailed proof of the existence of the stationary solution following the work in [7] and [23].

The main objective is to prove a central limit theorem (CLT) for the partial sums of the random variables $X_n$ appearing in equation (1) when the initial distribution of $X_0$ is tempered. Note that (1) is a contraction in the sense that every tempered distribution will force the partial sums to behave like the stationary solution. Thus we are led to study the CLT for the stationary solution, i.e. for the function $x^*$ under the action of $\theta$ in equation (3). The central limit theorem for processes generated by dynamical systems has gained dominant attention in the study of random phenomena in dynamics (see [10] for an introduction and basic concepts of the theory). The proof of limit theorems for dynamical systems is either based on the concept of probabilistic mixing (described in [15]) or on Gordin’s martingale approximation method ([13]). Here we follow both, giving the general result on the central limit theorem for square integrable functions in terms of Gordin’s approach for arrays of martingale differences (see the book by Hall and Heyde [14], for example). The example in Section 4, moreover, uses uniform mixing to assure the approximation by arrays of martingale differences. Note that under mixing conditions such a result was discussed in [4], however, their result, a weak invariance principle of the partial sums for i.i.d. coefficients, is proved by a different technique (also different from ours) and thus does not cover the present result. For completeness, we also mention the result in [19] where convergence to stable distributions is considered (under i.i.d. coefficients). The article [26] is closely related to our work in as much that Voliś considers similar martingale methods for proving central limit theorems for linear processes.

Section 4 contains the main results (which are in case $d = 1$). We assume the existence of an invariant family of sub-$\sigma$-algebras $\mathcal{F}_M$ such that $a$ and $b$ are $\mathcal{F}_0$-measurable (by [2] such a condition is natural). Under fairly general
conditions we are able to show (Theorem 4.2) that
\[ \frac{1}{\sqrt{n}} S_n (x^* - E(x^*|\mathcal{F}_{-M})) \]
is asymptotically normal. Here we just use an approximation by an array of martingale difference sequences. The centering can be replaced by a non-random centering if some additional conditions are satisfied. The condition given in Theorem 4.4 is weaker than those given in the literature, in particular weaker than the condition
\[ \sum_{k=1}^{\infty} V_k (x^* - E(x^*)) \text{ converges in } L^2(P), \]
which implies Gordin’s condition for the CLT for ergodic stationary sequences in [13]. Here \( V \) denotes the dual operator of \( U f = f \circ \theta \) defined on \( L^2(\mathcal{F}_0, P) \). It is known that certain mixing conditions of the filtration will imply the CLT. We show this in Example 4.9 for uniformly mixing filtrations \( \mathcal{F}_{-M}, M \geq 0 \). As a corollary of the main results we obtain the central limit theorem as well replacing the initial condition \( X_0 = x^* \) in the recursion equation (3) by any other variable \( f : \Omega \to \mathbb{R} \) as long as this variable is tempered. In particular, if a function is integrable, it is tempered, so the CLT holds.

Sections 5 and 6 contain the proofs of the results.

2 Preliminaries

We begin recalling the definition of a random dynamical system following the exposition in [1], and introducing the necessary notation.

Let \((\Omega, \mathcal{F}, P)\) be a probability space, where \( \mathcal{F} \) denotes a \( \sigma \)-field and \( P \) a probability measure. Let \( \theta : \Omega \to \Omega \) be an invertible probability preserving transformation. \( \theta \) defines a unitary operator \( U : L^2(\Omega, P) \to L^2(\Omega, P) \) by
\[ \int U f \cdot gdP = \int f \circ \theta \cdot gdP \quad g \in L^2(\Omega, P). \]
Throughout the paper, we fix a sub-\( \sigma \)-field \( \mathcal{F}_0 \subset \mathcal{F} \) such that
\[ \mathcal{F}_0 \subset \theta^{-1} \mathcal{F}_0 = \mathcal{F}_1 \]
and let
\[ F_j = \theta^{-j}F_0 \quad j \in \mathbb{Z}. \]
We shall call the system \((\Omega, \mathcal{F}, P, \theta, F_0)\) a filtered invariant measure preserving transformation. Let \(H_k = L^2(\Omega, \mathcal{F}_k, P)\) be the \(L^2\)-space of all square integrable functions \(f\) which are measurable with respect to \(\mathcal{F}_k\), i.e. \(U^{-k}f = f \circ \theta^{-k}\) is \(F_0\) measurable. More generally, for \(l, k \in \mathbb{Z}\) and an integrable \(\mathcal{F}_k\)-measurable function \(f\), \(U^{l-k}f\) is \(\mathcal{F}_l\)-measurable. Define
\[ Q = \bigcup_{-\infty < k < l < \infty} H_l \ominus H_k. \]
Note that \(Q\) is empty whenever \(\mathcal{F}_1 = \mathcal{F}_0\). We denote
\[ P_l(g) := E(g|\mathcal{F}_l) - E(g|\mathcal{F}_{l-1}) \]
for each \(l \in \mathbb{Z}\) and every random variable \(g\). Note that for any random variable \(g\) on \(\Omega\), we have that
\[ E(U^kg|\mathcal{F}_l) = U^kE(g|\mathcal{F}_{l-k}) \quad \text{for } l, k \in \mathbb{Z}, \quad (4) \]
\[ P_{-l}(g - E[g]) = P_{-l}(g) \quad \text{for } l \geq 1, \quad (5) \]
\[ U^{k+l}P_{-l} = P_kU^{l+k} \quad \text{for } k, l \in \mathbb{Z}. \quad (6) \]

Recall ([1]) that a random dynamical system on a topological space \(X\) with Borel field \(B(X)\) over a dynamical system \((\Omega, \mathcal{F}, P, \theta)\) with time \(\mathbb{Z}^+\) is a mapping \(\varphi: \mathbb{Z}^+ \times \Omega \times X \to X\) with the following properties:

1. \(\varphi\) is \((B(\mathbb{Z}^+) \otimes \mathcal{F} \otimes B(X), B(X))\)-measurable.

2. For all \(m, n \in \mathbb{Z}^+, \omega \in \Omega\) and \(x \in X\), we have
\[ \varphi(0, \omega, x) = x, \quad \text{and} \quad \varphi(n + m, \omega, x) = \varphi(n, \theta^m\omega, \varphi(m, \omega, x)). \]

Condition 2. is called the ”cocycle property”, and \(\varphi(1, \omega, \cdot)\) is the generator of random dynamical system \(\varphi\). If the map \(\varphi(1, \omega, \cdot): X \to X\) is invertible, then the random dynamical system \(\varphi\) can be extended in time to \(\mathbb{Z}\).

**Definition 2.1** A (strong) stationary solution of random dynamical system \(\varphi\) is a random variable \(x^* : \Omega \to X\) such that
\[ x^*(\theta\omega) = \varphi(1, \omega, x^*(\omega)) \quad \text{for almost all } \omega \in \Omega. \]
This equation implies that \( x^*(\theta^n \omega) = \varphi(n, \omega, x^*(\omega)) \) holds for all \( n \in \mathbb{Z}^+ \).
Hence the stationary solution generates a stationary process \( \{x^*(\theta^n \omega)\}_{n \geq 0} \).
As explained in the introduction we consider central limit theorem for a random process \( \{\varphi(n, \omega, g(\omega))\}_{n \geq 0} \) where \( g : \Omega \to X \) is a tempered function.

### 3 Random dynamical systems of affine transformations

**Definition 3.1** A random dynamical system \((\Omega, \mathcal{F}, P, \theta, \mathbb{R}, \varphi)\) is called a random dynamical system of affine transformations if for all \( x \in \mathbb{R} \), the generator is given by

\[
\varphi(1, \omega, x) := a(\omega) x + b(\omega)
\]  

(7)

where \( a, b : \Omega \to \mathbb{R} \) are random variables (on \((\Omega, \mathcal{F}, P)\)).

A filtered random dynamical system of affine transformations has the additional property that the probability space is filtered with respect to \( \mathcal{F}_0 \subset \mathcal{F} \) such that the tail field

\[
\mathcal{F}_{-\infty} = \bigcap_{k=0}^{\infty} \mathcal{F}_{-k}
\]

is trivial. The system will be denoted by \((\Omega, \mathcal{F}, P, \theta, \mathcal{F}_0, \mathbb{R}, \varphi)\).

Distributional solutions exist for random affine transformations under mild conditions. The result is due to Brandt in [9]. We give an existence proof for a stationary solution in the present context of random dynamics and extend it with respect to a filtration (cf. [23] and the introduction).

**Theorem 3.2** Let \((\Omega, \mathcal{F}, P, \theta, \mathcal{F}_0, \mathbb{R}, \varphi)\) be a filtered random dynamical system of affine transformations, where \( a, b \in L^2(\Omega, \mathcal{F}_0, P) \). If

\[
E[\log |a|] < 0 \quad \text{and} \quad E[\log^+ |b|] < \infty,
\]

(8)

where \( \log^+ |b| = \max\{0, \log |b|\} \). Then

\[
\lim_{n \to \infty} \sum_{i=1}^{n} \prod_{j=1}^{i} U^{-j} a U^{-i-1} b \quad \text{and} \quad \lim_{n \to \infty} \sum_{i=1}^{n} \prod_{j=0}^{i-1} U^j a
\]
exist $P$ a.e. Moreover,

$$x^* = U^{-1}b + \sum_{i=1}^{\infty} \prod_{j=1}^{i} U^{-j}aU^{-i-1}b$$

is a stationary solution for the random dynamical system of affine transformations.

Proof. Let $\lfloor x \rfloor$ denote the Gauss symbol of $x$ and let $\varepsilon = -E[\log |a(\omega)|] > 0$. If $N \leq \frac{2}{\varepsilon} \log^+ |b(\omega)| < N + 1$, we have that

$$1_{\left\{ \frac{2}{\varepsilon} \log^+ |b| \geq n \right\}}(\omega) = \begin{cases} 1 & \text{if } 1 \leq n \leq N, \\ 0 & \text{if } N < n. \end{cases}$$

whence

$$\sum_{n=1}^{\infty} 1_{\left\{ \frac{2}{\varepsilon} \log^+ |b| \geq n \right\}}(\omega) = \lfloor \frac{2}{\varepsilon} \log^+ |b(\omega)| \rfloor.$$

By assumption (8) and the monotone convergence theorem

$$\sum_{n=1}^{\infty} P\left( \frac{1}{n} \log^+ |U^{-n-1}b| \geq \frac{\varepsilon}{2} \right) \leq \frac{2}{\varepsilon} E[\log^+ |b|] < \infty.$$

The Borel-Cantelli lemma now yields

$$\limsup_{n \to \infty} \frac{1}{n} \log^+ |U^{-n-1}b| < \frac{\varepsilon}{2} \quad P\text{-a.e.}$$

Then by the ergodic theorem with respect to $\theta^{-1}$,

$$\limsup_{n \to \infty} \frac{1}{n} \log \left| \prod_{j=1}^{n} U^{-j}aU^{-n-1}b \right|$$

$$\leq \lim_{n \to \infty} \frac{1}{n} \sum_{j=1}^{n} \log |U^{-j}a| + \limsup_{n \to \infty} \frac{1}{n} \log^+ |U^{-n-1}b|$$

$$< -\varepsilon + \frac{\varepsilon}{2} = -\frac{\varepsilon}{2} \quad P\text{-a.e.}$$

If follows that

$$\sum_{n=1}^{\infty} \left| \prod_{j=1}^{n} U^{-j}aU^{-n-1}b \right| < \infty \quad P\text{-a.e.},$$

7
and therefore \( \sum_{n=1}^{\infty} \prod_{j=1}^{n} U^{-j} a U^{-n+1} b \) is a.s. pointwise convergent proving the a.s. existence of \( x^* \).

By the ergodicity with respect to \( \theta \), we have that
\[
\lim_{n \to \infty} \frac{1}{n} \log |\prod_{j=1}^{n} U^{j} a| = E[\log |a|] = -\varepsilon \quad P\text{-a.e.,}
\]
hence \( \lim \sup_{n \to \infty} |\prod_{j=1}^{n} U^{j} a|^{1/n} < 1 \) and
\[
\sum_{n=1}^{\infty} |\prod_{j=1}^{n} U^{j} a| < \infty \quad P\text{-a.e.}
\]

Therefore \( \sum_{n=1}^{\infty} \prod_{j=1}^{n} U^{j} a \) is well-defined \( P\)-a.e.

Finally observe that by definition, \( x^* \) is a stationary solution for the random dynamical system of affine transformations.

4 Main results

Let \( (\Omega, \mathcal{F}, P, \theta, \mathcal{F}_0, \mathbb{R}, \varphi) \) be a filtered random dynamical system of affine transformations. Define the formal expressions
\[
A^- := \lim_{n \to \infty} \sum_{i=1}^{n} \prod_{j=1}^{i} |U^{-j} a||U^{-i+1} b|
\]
\[
A^+ := \lim_{n \to \infty} \sum_{i=1}^{n} \prod_{j=0}^{i-1} |U^{j} a|
\]

We say that the random dynamical system of affine transformations is \textit{integrable} if the formal expressions exist and the following integrability conditions C1–C3 hold:

- C1: \( a, b \in L^2(\Omega, \mathcal{F}_0, P) \);
- C2: \( \log |a|, \log^+ |b| \in L^1(\Omega, \mathcal{F}_0, P) \);
- C3: \( A^+, A^- \in L^2(\Omega, \mathcal{F}_0, P) \);
Note that the proof of Theorem 3.2 shows that \( A^\pm \) exist a.s., but does not show any integrability. Thus C1 and C2 are sufficient to ensure the existence of the formal expressions.

**Remark 4.1** Denote by \( S_n f = f + f \circ \theta + \ldots + f \circ \theta^{n-1} \). Condition C3 can be obtained from the following condition:

- the free energy function for \( \log |a| \) exists in a region \( |t| \leq 4 \) and is negative, i.e.

  \[
  F(t) = \lim_{n \to \infty} \frac{1}{n} \log \int e^{tS_n \log |a|} dP < 0
  \]

  exist for \( |t| \leq 4 \).

- \( b \in \mathbb{L}^4(P) \).

The next theorem is the basic result of the paper.

**Theorem 4.2** Let \((\Omega, \mathcal{F}, P, \theta, \mathcal{F}_0, \mathbb{R}, \varphi)\) be a filtered and integrable random dynamical system of affine transformations. Then for \( M \geq 1 \)

\[
\frac{1}{\sqrt{n}} \sum_{k=1}^{n} U^k [x^* - E(x^*|\mathcal{F}_{-M})]
\]

converges in distribution to a normal law with zero mean and variance

\[
\sigma_M^2 := E \left( \sum_{l=1}^{\infty} U^l P_{-l}(x^* - E(x^*|\mathcal{F}_{-M})) \right)^2 < \infty.
\]

If \( \sigma_M^2 = 0 \), the statement means convergence is to 0 in probability. Suppose that C1 and C2 hold and that the conditions in Remark 4.1 hold. Then the central limit theorem with random centering as in Theorem 4.2 holds.

The random centering certainly can be replaced by the expectation of \( x^* \) if

\[
\lim_{n \to \infty} \frac{1}{\sqrt{n} \sigma_M} \sum_{k=1}^{n} (U^k E(x^*|\mathcal{F}_{-M}) - Ex^*) = 0
\]

in probability. A sufficient condition is contained in the next corollary.
Corollary 4.3 In addition to the assumptions for Theorem 4.2 assume that

1. \( \limsup_{M \to \infty} \sigma_M^2 = s^2 > 0 \).

2. \( \lim_{M \to \infty} \sum_{n=1}^{\infty} E \left[ U^n (E(x^* - Ex^* | \mathcal{F}_M)) E(x^* | \mathcal{F}_M) \right] = 0 \).

Then there exists a sequence \( \tau_n^2 \) such that

\[
\frac{1}{\sqrt{n} \tau_n^2} \sum_{k=1}^{n} U^k (x^* - Ex^*)
\]

converges weakly to the standard normal distribution.

The proof of Theorem 4.2 is given in Section 6, the proof of Corollary 4.3 is standard and therefore omitted. The theorem implies the following result by martingale approximation. Its proof is also postponed to Section 6. In fact, the condition of the last corollary is slightly weaker than the condition imposed in the next theorem, which also ensures the convergence of \( \sigma_M^2 \).

Theorem 4.4 Let \((\Omega, \mathcal{F}, P, \theta, \mathcal{F}_0, \mathbb{R}, \varphi)\) be a filtered and integrable random dynamical system of affine transformations. If for \( \bar{x}^* = x^* - E[x^*] \)

\[
\limsup_{M \to \infty} \sup_{K > M} \left| \sum_{k=M+1}^{K} E \left[ U^k \bar{x}^* \cdot E(U^K \bar{x}^* | \mathcal{F}_0) \right] \right| = 0, \tag{9}
\]

then

\[
\frac{1}{\sqrt{n}} \sum_{k=1}^{n} (U^k x^* - E[x^*])
\]

converges in distribution as \( n \to \infty \) to a normal distribution with zero mean and variance

\[
\sigma^2 := \lim_{M \to \infty} \sigma_M^2. \tag{10}
\]

Remark 4.5 We need to discuss condition (9). Since \( x^* \) is a square integrable random variable, according to Gordin’s theorem it is well known that the central limit theorem holds if \( \sum_{k=0}^{\infty} V_k \bar{x}^* \) converges in \( L_2(P) \), where \( V_k \)
denotes the dual operator of \( U^{-k} : L_2(\mathcal{F}_0) \to L_2(\mathcal{F}_{-k}) \). Note that the conditional expectation \( E(g | \mathcal{F}_{-k}) \) (\( g \in L_2(\mathcal{F}_0) \)) can be expressed by the operator \( V^k \):

\[
E(g | \mathcal{F}_{-k}) = U^{-k} V^k g.
\]

Assuming this condition and writing \( x^* = x^* - E x^* \),

\[
E [U^k \bar{x}^* E(U^M \bar{x}^* | \mathcal{F}_0)] = E [E(U^k \bar{x}^* | \mathcal{F}_0) E(U^M \bar{x}^* | \mathcal{F}_0)] = E [U^k E(\bar{x}^* | \mathcal{F}_{-k}) U^M E(\bar{x}^* | \mathcal{F}_{-M})] = E [U^k U^{-k} V^k \bar{x}^* U^M U^{-M} V^k \bar{x}^*].
\]

Therefore, by the Cauchy-Schwarz inequality

\[
\left| E \left[ \sum_{k=M+1}^{K} U^k \bar{x}^* E(U^M \bar{x}^* | \mathcal{F}_0) \right] \right| \leq \left\| \sum_{k=M+1}^{K} V^k \bar{x}^* \right\| \| \bar{x}^* \|_2.
\]

since the \( V^k \) are contractions. This shows that condition (9) is weaker than the usual condition for the CLT.

It is also easy to see that condition (9) is weaker than the corresponding condition (5.9) in [14], page 131.

**Remark 4.6** It is possible to translate the condition (9) into conditions on the functions \( a \) and \( b \). Assume that

\[
\limsup_{M \to \infty} \sup_{K > M} \sum_{k=M}^{K-1} E[E(U^K \bar{x}^* | \mathcal{F}_0) U^k b] = 0
\]

\[
\limsup_{M \to \infty} \sup_{K > M} \sum_{k=M}^{K-1} \| E(U^K \bar{x}^* | \mathcal{F}_0) U^k a \|_2 = 0
\]

\[
\sum_{i=1}^{\infty} \left\| \prod_{j=1}^{i-1} b \cdot U^j a \right\|_2 < \infty.
\]

The last condition is clearly satisfied under the condition of Remark 4.1.
Indeed, for fixed $M < K$ and letting $G = E(U^K x^* | F_0)$

\[
\begin{align*}
\sum_{k=M+1}^{K} E[G U^k x^*] &= \\
&= \sum_{k=M+1}^{K} E[G U^{k-1} b] + \sum_{i=1}^{\infty} \sum_{k=M+1}^{K} E \left[ G \cdot \prod_{j=1}^{i} U^{k-j} a U^{k-i-1} b \right] \\
&\leq \sum_{k=M+1}^{K} E[G U^{k-1} b] + \sum_{i=1}^{\infty} \sum_{k=M+1}^{K} \| G U^{k-1} a \|_2 \left\| \prod_{j=2}^{i} U^{k-j} a U^{k-i-1} b \right\|_2 \\
&= \sum_{k=M+1}^{K} E[G U^{k-1} b] + \sum_{k=M}^{K-1} \| G U^{k-1} a \|_2 \sum_{i=1}^{\infty} \left\| \prod_{j=1}^{i-1} U^{j} a \right\|_2 \\
&\leq \sum_{k=M}^{K} E[G U^{k-1} b] + \sum_{i=1}^{\infty} \left\| \prod_{j=1}^{i-1} U^{j} a \right\|_2 \\
&= \sum_{k=M}^{K} E[G U^{k-1} b] + \sum_{i=1}^{\infty} \left\| \prod_{j=1}^{i-1} U^{j} a \right\|_2 \\
&= \sum_{k=M}^{K} E[G U^{k-1} b] + \sum_{i=1}^{\infty} \left\| \prod_{j=1}^{i-1} U^{j} a \right\|_2 \\
&= \sum_{k=M}^{K} E[G U^{k-1} b] + \sum_{i=1}^{\infty} \left\| \prod_{j=1}^{i-1} U^{j} a \right\|_2 \\
&= \sum_{k=M}^{K} E[G U^{k-1} b] + \sum_{i=1}^{\infty} \left\| \prod_{j=1}^{i-1} U^{j} a \right\|_2 \\
\end{align*}
\]

**Definition 4.7** A random variable $f : \Omega \rightarrow (0, \infty)$ is called tempered with respect to the dynamical system $\theta$ if

\[
\lim_{n \to \pm \infty} \frac{1}{n} \log f(\theta^n \omega) = 0 \quad P\text{-a.e.}
\]

In general, the property “tempered” is a weaker condition than integrability (see Proposition 4.1.3, p.165 in [1]). By the result of [9], we have that for each tempered function $f$,

\[
\lim_{n \to \infty} |\phi(n, \omega, f(\omega)) - U^n x^*(\omega)| = 0 \quad P\text{-a.e.}
\]

In fact, if the random dynamical system of affine transformations satisfies the assumption (8) then $\lim_{n \to \infty} \prod_{j=1}^{n-1} |U^j a(\omega)| = 0$ for $P\text{-a.e. } \omega$, which follows from $\sum_{n=1}^{\infty} \prod_{j=1}^{n} |U^j a(\omega)| < \infty$ in the proof of Theorem 3.2. Moreover, $\{ \omega : f(\omega) = \infty \}$ is a null set, whence

\[
\lim_{n \to \infty} |\phi(n, \omega, f(\omega)) - U^n x^*(\omega)| = \lim_{n \to \infty} \prod_{j=1}^{n-1} |U^j a| \cdot |f(\omega) - x^*(\omega)| = 0 \quad P\text{-a.e.}
\]

12
In addition,

\[
\limsup_{n \to \infty} \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \left| \phi(k, \omega, f(\omega)) - U^k x(\omega) \right| \\
\leq \limsup_{n \to \infty} \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \prod_{j=1}^{k-1} |U^j a| \cdot |f(\omega) - x(\omega)| \\
\leq \limsup_{n \to \infty} \frac{1}{\sqrt{n}} \sum_{k=1}^{\infty} \prod_{j=1}^{k-1} |U^j a| \cdot |f(\omega) - x(\omega)| = 0.
\]

This estimate is needed to derive the central limit theorem for tempered functions.

**Corollary 4.8** If \( \sigma^2 > 0 \), then for any tempered function \( f \) (in particular any integrable function)

\[
\frac{1}{\sqrt{n}} \sum_{k=1}^{n} \left( \phi(k, \cdot, f) - E[\phi(k, \cdot, f)] \right)
\]

converges in distribution as \( n \to \infty \) to a normal distribution with zero mean and variance \( \sigma^2 \).

**Example 4.9** Let \( \theta : \Omega \to \Omega \) be bijective bimeasurable ergodic and \( P \)-preserving transformation of the probability space \( (\Omega, \mathcal{F}, P) \).

We assume that there exists a \( \sigma \)-field \( \mathcal{F}_0 \subset \mathcal{F} \) satisfying \( \mathcal{F}_0 \subset \theta^{-1} \mathcal{F}_0 \) and define the filtration \( (\mathcal{F}_i)_{i \in \mathbb{Z}} \) by \( \mathcal{F}_i = \theta^{-i} \mathcal{F}_0 \). Let

\[
\phi(k) = \sup_{A \in \mathcal{F}_k, B \in \mathcal{F}_0, P(A) > 0} |P(B|A) - P(A)|,
\]

\( b \) be a \( \mathcal{F}_0 \)-measurable \( L^2(\Omega, P) \)-function such that \( E[\log^+ |b|] < \infty \) and \( c \) be a constant in \((0,1)\). Consider the random dynamical system given by

\[
\varphi(1, \omega, x) = cx + b(\omega).
\]

If

\[
\sum_{k=1}^{\infty} \phi(k)^{1/2} < \infty,
\]
then, for all $f \in L^2(\Omega, P)$ (resp. tempered function $f$) and as $n \to \infty$,

$$
\frac{1}{\sqrt{n}} \sum_{k=1}^{n} [\phi(k, \omega, f(\omega)) - E[\phi(k, \omega, f(\omega))]]
$$

and

$$
\frac{1}{\sqrt{n}} \sum_{k=1}^{n} [x^*(\theta^k \omega) - E[x^*(\theta^k \omega)]]
$$

converge in distribution to a normal distribution with mean zero and variance $\sigma^2$, where $x^* = \sum_{i=0}^{\infty} c_i b \circ \theta^{-i-1}$.

In order to show the claim it suffices to show that the assumptions of Theorem 4.4 are satisfied.

Let $a(\omega) = c 1_{\Omega}(\omega)$. Since $c \in (0, 1)$, we have that $\sum_{j=1}^{\infty} j c^j < \infty$. Then we have the following:

- $E[\log |a(\omega)|] = \log c < 0$,
- $\| \sum_{i=1}^{\infty} \prod_{j=1}^{i} |U^{-j} a U^{-i-1} b| \|_2 = \sum_{i=1}^{\infty} c^i \| b \|_2 < \infty$,
- $\| \sum_{i=1}^{\infty} \prod_{j=0}^{i-1} |U^j a| \|_2 = \sum_{i=1}^{\infty} c^i < \infty$.

Therefore the assumptions C1-C3 are satisfied with respect to $a(\omega) = c 1_{\Omega}$. It follows that $x^* \in L^2(\Omega, P)$ since $\|x^*\|_2 = \sum_{i=0}^{\infty} c^i \| b \|_2 < \infty$. Thus we use Theorem A.6 of Appendix III in [14] and obtain

$$
\sum_{k=1}^{\infty} E \left| U^k \bar{x}^* E(U^M \bar{x}^* | F_0) \right| \leq 2 \sum_{k=1}^{\infty} \phi(k)^{1/2} \| x^* \|_2^2 < \infty,
$$

which implies the claim.

5 Auxiliary lemmas

In this section we prove three lemmas which contain the essential steps for the approximation by arrays of martingale difference sequences used in the proofs of Theorems 4.2 and 4.4.
Lemma 5.1 Let \( N, M \in \mathbb{N} \) be fixed and define

\[
(\text{I}) \quad g^*_{N,M} = \sum_{l=1}^{N} U^l P_{-l}(x^* - E(x^*|\mathcal{F}_M)),
\]

\[
(\text{II}) \quad z^*_{N,M} = \sum_{l=1}^{N} \sum_{m=0}^{l-1} U^m P_{-l}(x^* - E(x^*|\mathcal{F}_M)).
\]

Under the conditions C1–C3 in Section 4 the limits \( g^*_M := \lim_{N \to \infty} g^*_{N,M} \) and \( z^*_M := \lim_{N \to \infty} z^*_{N,M} \) exist in \( L^2(\Omega, \mathcal{F}, P) \) and \( P \text{-a.e.} \)

Proof. (I) Recall that \( \mathcal{F}_l \subset \mathcal{F}_{l+1} \) for each \( l \in \mathbb{Z} \). Therefore, if \( l \geq M \),

\( P_{-l}(x^* - E(x^*|\mathcal{F}_M)) = 0. \)

On the other hand, if \( 1 \leq l \leq M-1 \) then

\( P_{-l}(x^* - E(x^*|\mathcal{F}_M)) = P_{-l}(x^*). \)

Then, almost surely, \( g^*_M \) exists and equals

\[
g^*_M = \lim_{N \to \infty} \sum_{l=1}^{N} U^l P_{-l}(x^*) = \sum_{l=1}^{M-1} U^l P_{-l}(x^*).
\]

By condition C3, \( x^* \in L^2(\Omega, \mathcal{F}, P) \), hence, under the respective condition,

\[
\|g^*_M\|_2 = \left\| \sum_{l=1}^{M-1} U^l P_{-l}(x^*) \right\|_2 \leq \sum_{l=1}^{M-1} \|P_{-l}(x^*)\|_2 < \infty.
\]

Therefore \( g^*_{N,M} \) converges a.s. and in \( L^2 \).

(II) It has been shown in (I) that \( P_{-l}(x^* - E(x^*|\mathcal{F}_M)) \) equals \( P_{-l}(x^*) \) for \( 1 \leq l \leq M-1 \) and equals 0 for \( l \geq M \). Therefore, it follows that

\[
z^*_M = \lim_{N \to \infty} \sum_{l=1}^{N} \sum_{m=0}^{l-1} U^m P_{-l}(x^*) = \sum_{l=1}^{M-1} \sum_{m=0}^{l-1} U^m P_{-l}(x^*)
\]

exists a.s. Moreover, as before,

\[
\|z^*_M\|_2 = \left\| \sum_{l=1}^{M-1} \sum_{m=0}^{l-1} U^m P_{-l}(x^*) \right\|_2 < \infty.
\]

Therefore, \( z^*_{N,M} \) converges a.s. and in \( L^2 \) by condition C3.
Lemma 5.2 Let the conditions C1-C3 in Section 4 be satisfied.

Then the sequence \( \{U^k g_M^* : 1 \leq k \leq n, n \geq 1\} \) is a zero-mean, square-integrable, and stationary ergodic sequence of martingale difference sequences with respect to the filtrations \( \{\mathcal{F}_k\}_{1 \leq k \leq n}, n \geq 1 \).

Proof. For each \( k, U^k g_M^* \) is an \( L^2 \)-function by Lemma 5.1. Moreover, since \( E[g_M^*] = 0 \), the sequence \( \{U^k g_M^* : k \geq 1\} \) is a zero-mean square-integrable stationary ergodic sequence.

We prove that the sequence \( \{U^k g_M^* : 1 \leq k \leq n, n \geq 1\} \) is a sequence of martingale differences with respect to the filtration \( \{\mathcal{F}_k\}_{1 \leq k \leq n} \). By (11) and using (6), we have that for each \( 1 \leq k \leq n \) and \( n \geq 1 \),

\[
U^k g_M^* = \sum_{l=1}^{M-1} P_k(U^{l+k}x^*) = \sum_{l=1}^{M-1} [E(U^{l+k}x^*|\mathcal{F}_k) - E(U^{l+k}x^*|\mathcal{F}_{k-1})] \quad P\text{-a.e.}
\]

Hence for each \( k \geq 1 \), the random variable \( U^k g_M^* \) is \( \mathcal{F}_k \)-measurable. It also follows that for each \( k \geq 1 \)

\[
E(U^k g_M^*|\mathcal{F}_{k-1}) = \sum_{l=1}^{M-1} E\left(E(U^{l+k}x^*|\mathcal{F}_k) - E(U^{l+k}x^*|\mathcal{F}_{k-1})|\mathcal{F}_{k-1}\right) = 0.
\]

Lemma 5.3 Under conditions C1–C3 in Section 4, \( \frac{1}{\sqrt{n}} \sum_{k=1}^{n} U^k g_M^* \) converges in distribution to a random variable \( Z_M \) which has characteristic function \( E[e^{itZ_M^2}] \) as \( n \to \infty \), where \( \sigma^2_M = E[(g_M^*)^2] \).

Proof. We begin recalling Theorem 3.2 in [14]. If the array of martingale difference sequences \( \{U^k g_M^* : 1 \leq k \leq n\} \) with filtration \( \{\mathcal{F}_k\}_{1 \leq k \leq n} \) \( (n \geq 1) \) satisfies the following conditions:

(i) \( \frac{1}{\sqrt{n}} \max_{1 \leq k \leq n}|U^k g_M^*| \) converges to zero in probability as \( n \to \infty \),

(ii) \( \frac{1}{n} \sum_{k=1}^{n} U^k(g_M^*)^2 \) converges to a constant \( \sigma^2_M \) in probability as \( n \to \infty \),

(iii) \( \frac{1}{n} E \left[\max_{1 \leq k \leq n} |U^k(g_M^*)^2|\right] \) is bounded in \( n \),

then \( \frac{1}{\sqrt{n}} \sum_{k=1}^{n} U^k g_M^* \) converges in distribution to a normal random variable \( Z_M \) with expectation zero and variance \( \sigma^2_M \).
First, we consider the condition (i). Let $X_n = \frac{1}{\sqrt{n}} \max_{1 \leq k \leq n} |U^k g_M^*|$. By Chebychev’s inequality and square-integrability of $g_M^*$ (see Lemma 5.1), we obtain for any $\delta > 0$ that

$$P(X_n \geq \delta) = P(\max_{1 \leq k \leq n} |U^k g_M^*| \geq \delta \sqrt{n}) \\
\leq \sum_{k=1}^{n} P(|U^k g_M^*| \geq \delta \sqrt{n}) = n P(|g_M^*| \geq \delta \sqrt{n}) = O(1).$$

Condition (ii) holds by the individual ergodic theorem. Since $g_M^*$ is a $L^2$-function as shown in Lemma 5.1, the ergodic theorem yields

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} U^k (g_M^*)^2 = E[(g_M^*)^2] =: \sigma_M^2 \quad P\text{-a.e.}$$

Finally we consider the condition (iii). We have that

$$\frac{1}{n} \mathbb{E} \left[ \max_{1 \leq k \leq n} U^k (g_M^*)^2 \right] \leq \frac{1}{n} \sum_{k=1}^{n} \mathbb{E} [U^k (g_M^*)^2] = \mathbb{E}[(g_M^*)^2] < \infty.$$

Therefore the martingale difference sequences $\{ \frac{1}{\sqrt{n}} U^k g_M^* : k \geq 1 \}$ satisfies the conditions (i)-(iii), and we get that $\frac{1}{\sqrt{n}} \sum_{k=1}^{n} U^k g_M^*$ converges in distribution to a normal random variable $Z_M$ with expectation 0 and variance $\sigma_M^2 = E[(g_M^*)^2]$ as $n \to \infty$.

6 Proof of the theorems

Proof of Theorem 4.2. By Lemma 5.1 and 5.2, we have that

$$\left\| \sum_{l \in \mathbb{Z}} U^l P_{-l} (x^* - E(x^* | \mathcal{F}_{-M})) \right\|_2 = \left\| \sum_{l=1}^{M-1} U^l P_{-l} (x^*) \right\|_2 < \infty.$$

Therefore by Theorem 6 in [26]

$$\lim_{n \to \infty} \frac{1}{\sqrt{n}} \sum_{k=1}^{n} U^k [x^* - E(x^* | \mathcal{F}_{-M}) - g_M^*] = 0 \quad \text{in } L^2.$$
Since the sequence $\{U_{kgM^{\sqrt{n}}}\}_{k \geq 1}$ satisfies the central limit theorem by Lemma 5.3, this implies that the proof is completed.

Proof of Theorem 4.4.

Since $\bar{x} - (x^* - E(x^* | \mathcal{F}_{-M})) = E(\bar{x}^* | \mathcal{F}_{-M})$, by Gordin’s central limit theorem (cf. Theorem 5.1 in [14]) and Theorem 4.2, it is sufficient to show that for each arbitrary $\epsilon > 0$, there exists $M_0 \geq 1$ such that for all $M \geq M_0$

$$\limsup_{n \to \infty} E \left| \frac{1}{\sqrt{n}} \sum_{k=1}^{n} U_k [E(\bar{x}^* | \mathcal{F}_{-M})] \right|^2 \leq \epsilon.$$ 

This implies that $\inf_{f \in Q} \limsup_{n \to \infty} E \left| \frac{1}{\sqrt{n}} \sum_{k=1}^{n} U_k [\bar{x}^* - f] \right|^2 = 0$ since $x^* - E(x^* | \mathcal{F}_{-M}) \in H_{-1} \ominus H_{-M} \subset Q$.

Fix $\epsilon > 0$. Observe that

$$E \left[ \frac{1}{n} \left| \sum_{k=1}^{n} U_k [E(\bar{x}^* | \mathcal{F}_{-M})] \right|^2 \right] = E[E(\bar{x}^* | \mathcal{F}_{-M})^2] + \frac{2}{n} \sum_{k=1}^{n-1} (n - k) E \left[ U_k \bar{x}^* E(U^M \bar{x}^* | \mathcal{F}_0) \right].$$ 

For the first term on the right-hand side of (12) it follows from the triviality of the tail field $\mathcal{F}_{-\infty}$ that

$$E[E(\bar{x}^* | \mathcal{F}_{-M})^2] \leq \epsilon/2$$ 

for all $M$ large enough.

The second term on the right-hand side of (12) is equal to

$$\frac{2}{n} \sum_{k=1}^{n-1} (n - k) E \left[ U_k^{k+M} \bar{x}^* E(U^M \bar{x}^* | \mathcal{F}_0) \right]$$

$$= 2 \sum_{k=M+1}^{M+n-1} E \left[ U_k \bar{x}^* E(U^M \bar{x}^* | \mathcal{F}_0) \right] - \frac{2}{n} \sum_{k=1}^{n-1} k E \left[ U_k^{k+M} \bar{x}^* E(U^M \bar{x}^* | \mathcal{F}_0) \right].$$

By assumption (9)

$$\left| \sum_{k=M+1}^{M+n-1} E \left[ U_k \bar{x}^* E(U^M \bar{x}^* | \mathcal{F}_0) \right] \right| \leq \epsilon/4$$ 

(14)
for all \( n \geq 1 \) and for all \( M \) large enough. Also note that for fixed \( M \) by assumption (9) \( \sum_{k=1}^{\infty} E \left[ U^k \bar{x}^* E(U^M \bar{x}^* | \mathcal{F}_0) \right] \) converges. By Kronecker's lemma it follows that

\[
\lim_{n \to \infty} \frac{2}{n} \sum_{k=1}^{n-1} k E \left[ U^k \bar{x}^* E(U^M \bar{x}^* | \mathcal{F}_0) \right] = 0.
\]

Using this relation together with (13) and (14) we obtain

\[
\limsup_{n \to \infty} E \left[ \frac{1}{n} \left| \sum_{k=1}^{n} U^k \bar{x}^* E(U^M \bar{x}^* | \mathcal{F}_0) \right| \right] < \epsilon
\]

for all sufficiently large \( M \), say \( M \geq M_0 \).

It remains to show (10).

Recall that for integers \( i < j \) and \( M, M' \) sufficiently large as above

\[
E[U^i g_M^* U^j g_{M'}^*] = E [E(U^i g_M^* U^j g_{M'}^* | \mathcal{F}_i)] = E [U^i g_M^* E(U^j g_{M'}^* | \mathcal{F}_i)] = 0
\]

where \( g_M^* \) is defined in Lemma 5.1. This follows from Lemma 5.2, where it is shown that \( U^j g_{M'}^* \) (\( 1 \leq j \leq n \)) is a mean zero martingale difference sequence with respect to the filtration \( (\mathcal{F}_k) \). Therefore we have that for sufficiently large \( M, M' \geq M_0 \),

\[
(\sigma_M - \sigma_{M'})^2 = \left( E[(g_M^*)^2]^{1/2} - E[(g_{M'}^*)^2]^{1/2} \right)^2
\]

\[
\leq 2 \limsup_{n \to \infty} \frac{1}{n} E \left[ \left( \sum_{k=1}^{n} U^k [\bar{x}^* - g_M^*] \right)^2 \right]
\]

\[
+ 2 \limsup_{n \to \infty} \frac{1}{n} E \left[ \left( \sum_{k=1}^{n} U^k [\bar{x}^* - g_{M'}^*] \right)^2 \right]
\]

\[
\leq 4 \epsilon.
\]

Thus \( \sigma_M \) converges to some limit \( \sigma \) as \( M \to \infty \).

\textit{Proof of Corollary 4.8.}

By Theorem 4.4, it is sufficient to prove that

\[
\lim_{n \to \infty} \frac{1}{\sqrt{n}} \left| \sum_{k=1}^{n} (\phi(k, \cdot, f) - U^k x^* - E[\phi(k, \cdot, f) - U^k x^*]) \right| = 0
\]
in probability. By the condition C3, \( \left\| \sum_{k=1}^{\infty} |\prod_{j=0}^{k-1} U_j a| \right\|_2 < \infty \). Hence the family of random variables \( \{ \sum_{k=1}^{n} \prod_{j=0}^{k-1} |U_j a| \cdot |f - x^*| \}_{n \geq 1} \) is bounded by an integrable function and \( \lim_{n \to \infty} \sum_{k=1}^{n} \prod_{j=0}^{k-1} |U_j a| \cdot |f - x^*| = \sum_{k=1}^{\infty} \prod_{j=0}^{k-1} |U_j a| \cdot |f - x^*| \) P-a.e. Therefore, by dominated convergence, it follows that

\[
\frac{1}{\sqrt{n}} E \left[ \sum_{k=1}^{n} \prod_{j=0}^{k-1} |U_j a| \cdot |f - x^*| \right] \to 0 \quad \text{as } n \to \infty,
\]

whence

\[
\frac{1}{\sqrt{n}} \left| \sum_{k=1}^{n} \left( \varphi(k, \cdot, f) - U^k x^* - E[\varphi(k, \cdot, f) - U^k x^*] \right) \right|
\leq \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \prod_{j=0}^{k-1} |U_j a| \cdot |f - x^*| + \frac{1}{\sqrt{n}} E \left[ \sum_{k=1}^{n} \prod_{j=0}^{k-1} |U_j a| \cdot |f - x^*| \right]
\to 0 \quad \text{as } n \to \infty \quad \text{P-a.e.}
\]

Then, as \( n \to \infty \), \( \frac{1}{\sqrt{n}} \sum_{k=1}^{n} [\varphi(k, \cdot, f) - E[\varphi(k, \cdot, f)]] \) converges in distribution to a normal distribution with mean zero and variance \( \sigma^2 \), which completes the proof of Corollary 4.8.

### References


Authors’ addresses: