On Conditional Central Limit Theorems For Stationary Processes

Manfred Denker[1]
Universität Göttingen

and

Mikhail Gordin
V.A. Steklov Institute of Mathematics

Abstract

The central limit theorem for stationary processes arising from measure preserving dynamical systems has been reduced in [6] and [7] to the central limit theorem of martingale difference sequences. In the present note we discuss the same problem for conditional central limit theorems, in particular for Markov chains and immersed filtrations.

1 Introduction

Let $(\zeta_k)_{k\in\mathbb{Z}} = ((\xi_k, \eta_k))_{k\in\mathbb{Z}}$ be a two-component strictly stationary random process. Every measurable real-valued function f on the state space of the process defines another stationary sequence $(f(\zeta_k))_{k\in\mathbb{Z}}$. Various questions in stochastic control theory, modeling of random environment among many other applications lead to the study of conditional distributions of the sums $\sum_{k=0}^{n-1} f(\zeta_k)$ given $\eta_0, \ldots, \eta_{n-1}$. In particular, the asymptotic behaviour of these conditional distributions is of interest, including the case when the limit distribution is normal.

We shall prove conditional central limit theorems in the slightly more abstract situation of measure preserving dynamical systems $(X, \mathcal{F}, P, T)$, where $(X, \mathcal{F}, P)$ is a probability space and $T : X \to X$ is P-preserving.

Let f be a measurable function and $\mathcal{H}$ be a sub-σ-algebra. f is said to satisfy the conditional central limit theorem with respect to $\mathcal{H}$ (CCLT($\mathcal{H}$)), if P a.s. the conditional distributions of

$$\frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} f \circ T^k,$$

given $\mathcal{H}$, converge weakly to a normal distribution with some non-random variance $\sigma^2 \geq 0$.

This leads to the identification problem for $L_2(P)$-subspaces consisting of functions satisfying a CCLT. Following [6], an elegant way to describe such subclasses

[1]This paper is partially supported by the DFG-RFBR grant 99-01-04027. The second named author was also supported by the RFBR grants 00-15-96019 and 02-01-00265.

uses T-filtrations, i.e. increasing sequences of σ-fields $\mathcal{F}_n = T^{-1}\mathcal{F}_{n+1}$, $n \in \mathbb{Z}$. Here we need to consider a pair of T-filtrations $(\mathcal{F}_n)_{n\in\mathbb{Z}}$ and $(\mathcal{G}_n)_{n\in\mathbb{Z}}$ satisfying $\mathcal{G}_n \subset \mathcal{F}_n$ for every $n \in \mathbb{Z}$. For example, in case of a strictly stationary random process $(\xi_k)_{k\in\mathbb{Z}}$ as above the σ-field $\mathcal{F}_n$ (or $\mathcal{G}_n$) is generated by $(\zeta_k)_{k\leq n}$ (or $(\eta_k)_{k\leq n}$, respectively). First of all, the conditional distributions in CCLT($\mathcal{H}$) are determined by

$$\mathcal{H} = \bigvee_{k\in\mathbb{Z}} \mathcal{G}_k \vee \bigvee_{k\leq 0} \mathcal{F}_k.$$

Secondly, a general condition describing the class of functions f for which the CCLT($\mathcal{H}$) holds is given by the coboundary equation $f = h + g - g \circ T$ with a $(\mathcal{F}_n)_{n\in\mathbb{Z}}$-martingale difference sequence $h \circ T^k$ (i.e. h is $U_T\mathcal{H}$-measurable and $E^{\mathcal{H}} f := E(f|\mathcal{H}) = 0$).

The coboundary equation is implicitely also used in [10] and [9]. In [10], sufficient conditions for CCLT($\mathcal{H}$) are obtained, when $\mathcal{H}$ is replaced by $\widetilde{\mathcal{H}} = \bigvee_{k\in\mathbb{Z}} \mathcal{G}_k$, and our Proposition 3.1 contains this result as a special case. This proposition also specializes in case of skew products $T(x, y) = (\tau(x), T_x(y))$ as in [9], where $\mathcal{G}_n$ is a τ-filtration, and where $\mathcal{H}$ is also replaced by $\widetilde{\mathcal{H}}$.

It is hardly possible to verify this coboundary condition using properties of the σ-fields $(\mathcal{F}_n)_{n\in\mathbb{Z}}$ and $(\mathcal{G}_n)_{n\in\mathbb{Z}}$ without making assumptions about their interaction. It has been noticed in [5] that conditional independence plays a fundamental role when studying conditional measures and their properties in connection with thermodynamic formalism. This additional property of conditional independence has been called *immersion* in [1], and we shall adopt this terminology. It means that for every $n \in \mathbb{Z}$ the σ-fields $\mathcal{F}_n$ and $\mathcal{G}_{n+1}$ are conditionally independent given $\mathcal{G}_n$. The property of immersion is an essential simplification, although it seems to be rather strong. However, it looks quite natural in several situations (see e.g. [5]), in particular, when both $(\zeta_k)_{k\in\mathbb{Z}}$ and $(\eta_k)_{k\in\mathbb{Z}}$ are Markovian. Indeed, if the sequence $(\eta_k)_{k\in\mathbb{Z}}$ models the time evolution of a random environment influencing the process $(\xi_k)_{k\in\mathbb{Z}}$, the condition just means that there is no interaction between the process $(\xi_k)_{k\in\mathbb{Z}}$ and the environment $(\eta_k)_{k\in\mathbb{Z}}$. The same picture arises when $(\xi_k)_{k\in\mathbb{Z}}$ models the outcome of non-anticipating observations over the process $(\eta_k)_{k\in\mathbb{Z}}$, mixed with noise. If the sequence $(\zeta_k)_{k\in\mathbb{Z}}$ is a Markov chain, there is a natural assumption in terms of transition probabilities to guarantee that the corresponding filtrations are immersed (see Section 4).

The notion of immersed filtrations was first recognized as an important concept in connection with the classification problem of filtrations (see [1] and references therein). A closely related notion, *regular factors*, was introduced in [5]. The latter paper also contains some examples of regular factors originating in two-dimensional complex dynamics.

In more general situations (like in control theory) some form of the feed-back between the two processes may be present, and we cannot expect that the corresponding filtrations are immersed. In this case more general concepts and results (like Theorem 3.7 of the present paper) have to be developed.

In particular, we study the CCLT-problem for functions of Markov chains. We follow the ideas in [7] closely where a rather general and natural condition in terms of the transition operator was introduced for the CLT-problem. This condition means that the Poisson equation is solvable, and it avoids mixing assumptions and similar concepts (e.g. [9] contains results in this direction). There is a natural construction embedding the original Markov chain into another one, for which the Poisson equation has to be solved. We give some comments how this verification can be done, in particular, in the context of fibred dynamical systems [5]. However, we do not go into much of details. As a consequence we obtain the functional form of the CLT for fluctuations of a random sequence around the conditional mean.

Finally, we consider the case of immersed Markov chains. This property together with a solution of the Poisson equations for the original and extended Markov chains establishes an analogous result for conditional mean values of the original sequence, in addition to the CLT for fluctuations.

The present paper arose from an attempt to understand Bezhaeva's paper [2] from the viewpoint of martingales. Bezhaeva's article studies the same problem as in the present note in the special case of finite state Markov chains. We do not reproduce these results in detail and formulate the conclusions of our theorems in a way different from the viewpoint taken in [2]. However, we would like to sketch the differences in both approaches. There are two results on the CLT in [2]: Theorem 3 and Theorem 5 (the latter theorem seems to be the most important result of [2]). Our corresponding results are Theorem 3.7 and Theorem 4.4. Though, we do not verify here that the conditions of our Theorem 4.4 are satisfied for a class of Markov chains considered in [2] and arbitrary centered functions: this would be just a reproduction of a part of [2]. Its proof and the content of our Section 4 clearly show that even for finite state Markov chain we really deal with continuous state space when considering a conditional setup. In fact much more general chains than in Theorem 5 in [2] (for example, geometrically ergodic) can be considered on the basis of our Theorem 4.4. Our method of proving the CLT is quite different from that of [2] and, as was remarked above, is based on approximation by martingales.

We assume in this paper that all probability spaces and σ-fields satisfy the requirements of Rokhlin's theory of Lebesgue spaces and measurable partitions. This does not imply any restriction to the joint distributions of random sequences we are considering; hence we may freely use conditional probability distributions given a σ-field. An alternative approach would be to reformulate the results avoiding conditional distributions. However, we do not think that the advantages given by such an approach justifies the complexity of such a description.

2 Immersed Filtrations

Throughout this paper, let $(X, \mathcal{F}, P)$ and $T : X \to X$ be, respectively, a probability space and an automorphism of $(X, \mathcal{F}, P)$ (that is an invertible P-preserving measurable transformation). An increasing sequence of σ-subfields

$(\mathcal{F}_n)_{n\in\mathbb{Z}}$ of $\mathcal{F}$ will be called a *filtration* and a T-*filtration* if, in addition, $T^{-1}(\mathcal{F}_n) = \mathcal{F}_{n+1}$ for every $n \in \mathbb{Z}$. Any σ-field $\mathcal{E} \subseteq \mathcal{F}$ defines a natural T-filtration $(\mathcal{E}_n)_{n\in\mathbb{Z}} = (T^{-n}\mathcal{E})_{n\in\mathbb{Z}}$, whenever $T^{-1}\mathcal{E} \supseteq \mathcal{E}$. A filtration $(\mathcal{G}_n)_{n\in\mathbb{Z}}$ is said to be *subordinated* to a filtration $(\mathcal{F}_n)_{n\in\mathbb{Z}}$, if for every $n \in \mathbb{Z}$

$$\mathcal{G}_n \subseteq \mathcal{F}_n, \tag{2.1}$$

and it is called *immersed* into the filtration $(\mathcal{F}_n)_{n\in\mathbb{Z}}$, if $(\mathcal{G}_n)_{n\in\mathbb{Z}}$ is subordinated to $(\mathcal{F}_n)_{n\in\mathbb{Z}}$ and for every $n \in \mathbb{Z}$ the σ-fields $\mathcal{F}_n$ and $\mathcal{G}_{n+1}$ are conditionally independent given $\mathcal{G}_n$.

We shall always assume that

$$\mathcal{F} = \bigvee_{n\in\mathbb{Z}} \mathcal{F}_n \tag{2.2}$$

($\bigvee_{s\in S} \mathcal{E}_s$ denotes the smallest σ-field containing all σ-fields $\mathcal{E}_s$, $s \in S$). Setting $\mathcal{G} = \bigvee_{n\in\mathbb{Z}} \mathcal{G}_n$ it follows from the definition of a T-filtration that $\mathcal{G}$ is *completely invariant* with respect to T (that is $T^{-1}(\mathcal{G}) = \mathcal{G}$). Finally, define $\mathcal{F}_- = \bigcap_{k\in\mathbb{Z}} \mathcal{F}_k$, and similarly $\mathcal{G}_- = \bigcap_{k\in\mathbb{Z}} \mathcal{G}_k$.

Throughout this paper $(\mathcal{G}_n)_{n\in\mathbb{Z}}$ always denotes a T-filtration which is subordinated to the T-filtration $(\mathcal{F}_n)_{n\in\mathbb{Z}}$. We then set

$$\mathcal{H}_n = \mathcal{G} \vee \mathcal{F}_n.$$

The transformation T defines a unitary operator U_T on $L_2 = L_2(X, \mathcal{F}, P)$ by $U_T f = f \circ T$, $f \in L_2$. Given a sub-σ-field $\mathcal{H} \subset \mathcal{F}$, we denote its conditional expectation operator (on L_2) by $E^{\mathcal{H}}$ and its conditional probability by $P(\cdot|\mathcal{H})$. Let $\|\cdot\|_2$ denote the L_2-norm.

As mentioned above, the notion of immersed filtrations arises naturally in the context of Gibbs measures in the thermodynamic formalism (see [5]) and of Markov chains (see e.g. [2]). In order to simplify our conditions in the CCLT for these applications we need the following lemma for immersed filtrations.

Lemma 2.1. *The T-filtration $(\mathcal{G}_k)_{k\in\mathbb{Z}}$ is immersed into the T-filtration $(\mathcal{F}_k)_{k\in\mathbb{Z}}$, if for every $n \in \mathbb{Z}$*

$$E^{\mathcal{F}_n} \circ E^{\mathcal{G}_{n+1}} = E^{\mathcal{G}_n}. \tag{2.3}$$

or, equivalently,

$$E^{\mathcal{G}_{n+1}} \circ E^{\mathcal{F}_n} = E^{\mathcal{G}_n}. \tag{2.4}$$

Conversely, if $(\mathcal{G}_k)_{k\in\mathbb{Z}}$ is immersed into $(\mathcal{F}_k)_{k\in\mathbb{Z}}$, then the following equalities hold for every $n \in \mathbb{Z}$ and $m \geq 1$:

$$E^{\mathcal{F}_n} \circ E^{\mathcal{G}_{n+m}} = E^{\mathcal{G}_{n+m}} \circ E^{\mathcal{F}_n} = E^{\mathcal{G}_n}. \tag{2.5}$$

Proof. We first show that $(\mathcal{G}_k)_{k\in\mathbb{Z}}$ is immersed into $(\mathcal{F}_k)_{k\in\mathbb{Z}}$, if (2.3) holds. Let $n \in \mathbb{Z}$ be fixed and let ξ and η be bounded functions measurable with respect to $\mathcal{F}_n$ and $\mathcal{G}_{n+1}$, respectively. It follows from (2.3) that $E^{\mathcal{F}_n}\eta = E^{\mathcal{G}_n}\eta$. Therefore we have

$$\begin{aligned} E^{\mathcal{G}_n}(\xi\eta) &= E^{\mathcal{G}_n} E^{\mathcal{F}_n}(\xi\eta) = E^{\mathcal{G}_n}(\xi E^{\mathcal{F}_n}\eta) \\ &= E^{\mathcal{G}_n}(\xi E^{\mathcal{G}_n}\eta) = E^{\mathcal{G}_n}(\xi) E^{\mathcal{G}_n}(\eta), \end{aligned}$$

which implies the conditional independence of $\mathcal{F}_n$ and $\mathcal{G}_{n+1}$ given $\mathcal{G}_n$. In a similar way (replacing $\mathcal{F}_n$ by $\mathcal{G}_{n+1}$) one shows conditional independence assuming (2).

Conversely, we first show that conditional independence of $\mathcal{F}_n$ and $\mathcal{G}_{n+1}$ given $\mathcal{G}_n$ for some $n \in \mathbb{Z}$ implies (2.3). Indeed, it suffices to verify (2.3) for all bounded $\mathcal{F}_n \vee \mathcal{G}_{n+1}$-measurable functions of the form $\xi\eta$, where ξ and η are $\mathcal{F}_n$- and $\mathcal{G}_{n+1}$-measurable, respectively. By conditional independence, for a $\mathcal{G}_{n+1}$-measurable, bounded function h,

$$\int h \cdot E^{\mathcal{G}_{n+1}}\xi dP = \int E^{\mathcal{G}_n}(\xi h)dP = \int E^{\mathcal{G}_n}\xi E^{\mathcal{G}_n} h dP = \int h E^{\mathcal{G}_n}\xi dP,$$

whence $E^{\mathcal{G}_{n+1}}\xi = E^{\mathcal{G}_n}\xi$. Similarly one shows that $E^{\mathcal{F}_n}\eta = E^{\mathcal{G}_n}\eta$. It follows that

$$\begin{aligned} E^{\mathcal{F}_n} E^{\mathcal{G}_{n+1}}(\xi\eta) &= E^{\mathcal{F}_n}(\eta E^{\mathcal{G}_{n+1}}\xi) = E^{\mathcal{F}_n}(\eta E^{\mathcal{G}_n}\xi) \\ &= (E^{\mathcal{G}_n}\xi)(E^{\mathcal{F}_n}\eta) = (E^{\mathcal{G}_n}\xi)(E^{\mathcal{G}_n}\eta) \\ &= E^{\mathcal{G}_n}(\xi\eta). \end{aligned}$$

Since the equation (2.4) can be proved similarly, we obtain the equivalence of (2.3) and (2.4). Moreover, by induction one easily verifies (2.5). □

3 A Conditional Central Limit Theorem

Let $(\nu_k)_{k\geq 1}$ be a sequence of real-valued random variables. For every $n \in \mathbb{Z}_+$ define a random function with values in the Skorokhod space $D([0,1])$ ([3], [8]) in the standard way: it is piecewise constant, right continuous, equals 0 in the interval $[0, 1/n)$ and equals $n^{-1/2}\sum_{1\leq m\leq [nt]}\nu_m$ for a point $t \in [1/n, 1]$. This random function will be denoted by $R_n(\nu_1, \ldots, \nu_n)$ and has a distribution on $D([0,1])$, denoted by $P_n(\nu_1, \ldots, \nu_n)$. We write w_σ for the Brownian motion on $[0,1]$ with variance σ^2 of $w_\sigma(1)$ (we need not exclude $\sigma^2 = 0$ since w_0 is the process which identically vanishes). The distribution of w_σ in $C([0,1])$ will be denoted by W_σ.

Remark 3.1. In the sequel we deal with convergence in probability of a sequence of *random* probability distributions in $D([0,1])$ to a *non-random* probability distribution. It is assumed here that the set of all probability distributions in $D([0,1])$ is endowed with the *weak topology.* It is well known that the piecewise constant random functions (in $D([0,1])$) can be replaced by piecewise linear functions (in $C([0,1])$) without changing the essence of the results formulated below.

3.1 A general CCLT

As mentioned in the introduction the conditional central limit theorems in [9] and [10] are proved using some martingale approximation. There are different versions of a martingale central limit theorem which may be used in the present context. They all are versions and extensions of Brown's martingale central limit theorem. It has been used in [10] directly, and is used in [9] and here in a modified form. We apply a corollary of Theorem 8.3.33 in [8] to obtain the following CLT for arrays of martingale difference sequences.

Lemma 3.1. *For $n \in \mathbb{Z}_+$ let $(\Omega^n, \mathcal{F}^n, (\mathcal{F}_{k,n})_{k\geq 0}, P^n)$ be a probability space with filtration $\mathcal{F}_{k,n} \subset \mathcal{F}^n$ $(k \geq 0)$, and let $(\nu_{k,n})_{k\geq 1}$ be a square integrable martingale difference sequence with respect to $((\mathcal{F}_{k,n})_{k\geq 0}, P^n)$. If for every $\epsilon > 0$ and $t \geq 0$ we have*

$$\sum_{1\leq k\leq nt} E^{\mathcal{F}_{k-1,n}}(\nu_{k,n}^2 \mathbf{1}_{\{|\nu_{k,n}|>\epsilon\}}) \to 0 \tag{3.6}$$

and

$$\sum_{1\leq k\leq nt} E^{\mathcal{F}_{k-1,n}} \nu_{k,n}^2 \to t\sigma^2 \tag{3.7}$$

in probability as $n \to \infty$ then $\{P_n(\nu_1, \ldots, \nu_n) : n \geq 1\}$, converges weakly to W_{σ^2}.

The following proposition is the key result in the martingale approximation method for the CCLT. Implicitly it also appears in [10], and its proof is analogous to that for the central limit theorem in [6] or [7].

Proposition 3.1. *Let T be an ergodic automorphism and $(\mathcal{H}_n)_{n\in\mathbb{Z}}$ be a T-filtration. Assume that $g, h \in L_2$ and*

$$E^{\mathcal{H}_1} h = h, E^{\mathcal{H}_0} h = 0. \tag{3.8}$$

If f is defined by

$$f = h + g - U_T g, \tag{3.9}$$

then, with probability 1, the conditional distributions $P_n(f, U_T f, \ldots, U_T^{n-1} f|\mathcal{H}_0)$ given $\mathcal{H}_0$ of the random functions $R_n(f, U_T f, \ldots, U_T^{n-1} f)$ converge weakly to the (non-random) probability distribution W_σ, where $\sigma = \|h\|_2 \geq 0$.

Remark 3.2. The equations in (3.8) say that the sequence $(U_T^n h)_{n\in\mathbb{Z}}$ is a stationary martingale difference sequence with respect to the filtration $(\mathcal{H}_n)_{n\in\mathbb{Z}}$.

Remark 3.3. The conclusion of Proposition 3.1 remains true if the σ-field $\mathcal{H}_0$ in the statement is changed to any coarser one. This follows easily from the definition of weak convergence and the non-randomness of the limit distribution.

Proof of Proposition 3.1. By remark 3.2 the sequence of finite series $\nu_{k,n} = n^{-1/2} U_T^{k-1} h$, $(1 \leq k \leq n)$, form a martingale difference sequence with respect to the filtrations $(\mathcal{H}_k)_{0\leq k\leq n}$. Assume first that $\sigma > 0$. We show that the sequence $\{\nu_{k,n} | 1 \leq k \leq n,\ n \in \mathbb{Z}\}$ with probability 1 satisfies the conditions 3.6 and 3.7 of Lemma 3.1 with respect to the conditional distribution given $\mathcal{H}_0$. Relative to this conditional distribution with probability 1 the sequence $(U_T^n h)_{n\in\mathbb{Z}}$ is a (non-stationary) sequence of martingale differences with finite second moments. The ergodic theorem implies that with P-probability 1

$$\frac{1}{n}\sum_{k=0}^{n-1} E^{\mathcal{H}_k}(U_T^k h)^2 = \frac{1}{n}\sum_{k=0}^{n-1} U_T^k (E^{\mathcal{H}_0} h^2) \to \|h\|_2^2$$

as $n \to \infty$. It follows that with probability 1 the same relation holds almost surely with respect to the conditional probability given $\mathcal{H}_0$, establishing (3.7). We need to check (3.6). By the ergodic theorem again, for every $\epsilon > 0$ and $A > 0$ we have with P-probability 1

$$\begin{aligned}
& \limsup_{n\to\infty} \sum_{1\le k\le nt} E^{\mathcal{H}_{k-1}}(\nu_{k,n}^2 \mathbf{1}_{\{|\nu_{k,n}|>\epsilon\}}) \\
= & \limsup_{n\to\infty} n^{-1} \sum_{0\le k\le (n-1)t} E^{\mathcal{H}_k}((U_T^k h)^2 \mathbf{1}_{\{|U_T^k h|>\epsilon n^{1/2}\}}) \\
\le & \limsup_{n\to\infty} n^{-1} \sum_{0\le k\le (n-1)t} E^{\mathcal{H}_k}((U_T^k h)^2 \mathbf{1}_{\{|U_T^k h|>A\}}) \\
= & \limsup_{n\to\infty} n^{-1} \sum_{0\le k\le (n-1)t} E^{\mathcal{H}_k}(U_T^k h^2 U_T^k \mathbf{1}_{\{|h|>A\}}) \\
= & \limsup_{n\to\infty} n^{-1} \sum_{0\le k\le (n-1)t} U_T^k(E^{\mathcal{H}_0}(h^2 \mathbf{1}_{\{|h|>A\}})) \\
= & EE^{\mathcal{H}_0}(h^2\mathbf{1}_{\{|h|>A\}}) = E(h^2\mathbf{1}_{\{|h|>A\}}),
\end{aligned}$$

and, choosing A large enough, the latter expression can be made arbitrarily small. Thus for every $\epsilon > 0$ with P-probability 1

$$\sum_{1\le k\le nt} E^{\mathcal{H}_{k-1}}(\nu_{k,n}^2 \mathbf{1}_{\{|\nu_{k,n}|>\epsilon\}}) \to 0$$

as $n \to \infty$. This implies that with probability 1 the same expression tends to zero with respect to the conditional probability given $\mathcal{H}_0$, proving (3.6).

It follows from Lemma 3.1 that $P_n(h, \ldots, U_T^{n-1}h|\mathcal{H}_0)$ converges weakly to W_σ P-a.s. The same conclusion also holds if $\sigma = 0$ ($h = 0$ in this case).

Finally we need to show that the sequences $(U_T^n h)_{n\in\mathbb{Z}}$ and $(U_T^n f)_{n\in\mathbb{Z}}$ are stochastically equivalent. We have

$$\begin{aligned}
& R_n(n^{-1/2}f, \ldots, n^{-1/2}U_T^{n-1}f) - R_n(n^{-1/2}h, \ldots, n^{-1/2}U_T^{n-1}h) = \\
& R_n(n^{-1/2}(U_T g - g), n^{-1/2}(U_T^2 g - U_T g), \ldots, n^{-1/2}(U_T^n g - U_T^{n-1} g)).
\end{aligned}$$

It is easy to see that the maximum (over the interval $[0,1]$) of the modulus of the latter random function equals $n^{-1/2}\max_{1\le k\le n}|U_T^k g - g|$ and does not exceed $n^{-1/2}(|g| + \max_{1\le k\le n}|U_T^k g|)$. Since by the ergodic theorem $n^{-1}U_T^n g^2 \to 0$, this expression tends to zero P-a.s. Thus we see that P-a.s. the distance in $D([0,1])$ between $R_n(h, \ldots, U_T^{n-1}h)$ and $R_n(f, \ldots, U_T^{n-1}f)$ tends to zero as $n \to \infty$. This implies that, with probability 1, the conditional distributions $P_n(f, \ldots, U_T^{n-1}f)|\mathcal{H}_0)$ in $D([0,1])$ have the same weak limit as

$$P_n(h, \ldots, U_T^{n-1}h|\mathcal{H}_0).$$

$\square$

3.2 On Rubshtein's CCLT

Proposition 3.1 is in fact a general result which can be seen when compared to other theorems in the literature. We begin recalling Rubshtein's result in [10].

Theorem 3.4. *Let $(\xi_n, \eta_n)_{n\in\mathbb{Z}}$ be an ergodic stationary process with $\xi \in L_2$ and $E^{\mathcal{G}}\xi_0 = 0$. If*

$$\sup_{n\geq 1} E\left(E^{\mathcal{H}_0}\left(\sum_{k=1}^{n}\xi_k\right)\right)^2 < \infty, \tag{3.10}$$

then, with probability 1, the conditional distributions $P_n(\xi_1, \xi_2, \dots, \xi_n|\,\mathcal{G})$ of $R_n(\xi_1, \xi_2, \dots, \xi_n)$ converge weakly to the non-random probability W_σ, where

$$\lim_{n\to\infty}\frac{1}{n}E(\xi_1 + \dots + \xi_n)^2 = \sigma^2.$$

The proof of this result can be reduced to Proposition 3.1 observing that (3.10) implies a representation as in (3.9). The result in [9], Theorem 2.3 is of the same nature, but in the special situation of a skew product. Another special case of Proposition 3.1 is the following theorem, which is also a generalization of Theorem 2 in [6], when $p = 2$.

Theorem 3.5. *Let T be an ergodic automorphism and $(\mathcal{H}_n)_{n\in\mathbb{Z}}$ be a T-filtration. If $f \in L_2$ is a real-valued function satisfying*

$$\sum_{k=0}^{\infty}(\|f - E^{\mathcal{H}_k}f\|_2 + \|E^{\mathcal{H}_{-k}}f\|_2) < \infty, \tag{3.11}$$

then Proposition 3.1 applies to f. In particular, there exists $\sigma \geq 0$ such that with probability 1 the conditional distributions $P_n(f, U_T f, \dots, U_T^{n-1}f|\ \mathcal{H}_0)$ converge weakly to the probability distribution W_σ.

Proof. The following explicit formula defines a function g which permits a representation as in (3.9), where we set $h = f - g\circ T + g$:

$$g = \sum_{k=1}^{\infty} U_T^{-k}(f - E^{\mathcal{H}_k}f) - \sum_{k=0}^{\infty} U_T^{k}(E^{\mathcal{H}_{-k}}f)$$

(here and below the series are L_2-norm convergent due to the assumption

(3.11)). It follows that

$$\begin{aligned}
h &= f - U_T g + g \\
&= f - \sum_{k=1}^{\infty} U_T^{-k+1}(f - E^{\mathcal{H}_k} f) + \sum_{k=0}^{\infty} U_T^{k+1}(E^{\mathcal{H}_{-k}} f) \\
&\quad + \sum_{k=1}^{\infty} U_T^{-k}(f - E^{\mathcal{H}_k} f) - \sum_{k=0}^{\infty} U_T^{k}(E^{\mathcal{H}_{-k}} f) \\
&= f - \sum_{k=0}^{\infty} U_T^{-k}(f - E^{\mathcal{H}_{k+1}} f) + \sum_{k=1}^{\infty} U_T^{k}(E^{\mathcal{H}_{-k+1}} f) \\
&\quad + \sum_{k=1}^{\infty} U_T^{-k}(f - E^{\mathcal{H}_k} f) - \sum_{k=0}^{\infty} U_T^{k}(E^{\mathcal{H}_{-k}} f) \\
&= f - (f - E^{\mathcal{H}_1} f) + \sum_{k=1}^{\infty} U_T^{-k}(E^{\mathcal{H}_{k+1}} f - E^{\mathcal{H}_k} f) \\
&\quad - E^{\mathcal{H}_0} f + \sum_{k=1}^{\infty} U_T^{k}(E^{\mathcal{H}_{-k+1}} f - E^{\mathcal{H}_{-k}} f) \\
&= \sum_{k \in \mathbb{Z}} U_T^{k}(E^{\mathcal{H}_{-k+1}} - E^{\mathcal{H}_{-k}}) f \\
&= \lim_{n \to \infty} (E^{\mathcal{H}_1} - E^{\mathcal{H}_0}) \sum_{k=-n}^{n} U_T^{k} f.
\end{aligned}$$

This representation clearly shows that h satisfies (3.8) and the theorem follows from Proposition 3.1. □

3.3 The CCLT for subordinated filtrations

Let $(\mathcal{G}_n)_{n\in\mathbb{Z}}$ and $(\mathcal{F}_n)_{n\in\mathbb{Z}}$ be two subordinated T-filtrations as explained in section 2 on filtrations. We shall use Proposition 3.1 to obtain sufficient conditions that the CCLT holds together with the CLT for the conditional mean. We begin with the following reformulation of Proposition 3.1.

Proposition 3.2. *Let T be an ergodic automorphism, $(\mathcal{G}_n)_{n\in\mathbb{Z}}$ and $(\mathcal{F}_n)_{n\in\mathbb{Z}}$ be a pair of T-filtrations such that $(\mathcal{G}_n)_{n\in\mathbb{Z}}$ is subordinated to $(\mathcal{F}_n)_{n\in\mathbb{Z}}$. For $f \in L_2$ define $\widehat{f} = E^{\mathcal{G}} f$ and $\widetilde{f} = f - \widehat{f}$. Assume that $\widehat{f}$ and $\widetilde{f}$ admit representations*

$$\widehat{f} = \widehat{h} + \widehat{g} - U_T \widehat{g}, \tag{3.12}$$

and

$$\widetilde{f} = \widetilde{h} + \widetilde{g} - U_T \widetilde{g}, \tag{3.13}$$

where $\widehat{g}, \widetilde{g} \in L_2$,

$$E^{\mathcal{G}_1} \widehat{h} = \widehat{h}, \;\; E^{\mathcal{G}_0} \widehat{h} = 0$$

$$E^{\mathcal{H}_1} \widetilde{h} = \widetilde{h} \text{ and } E^{\mathcal{H}_0} \widetilde{h} = 0;$$

then

i) *the distributions* $P_n(\widehat{f}, U_T\widehat{f}, \ldots, U_T^{n-1}\widehat{f})$ *of the random functions* $R_n(\widehat{f}, U_T\widehat{f}, \ldots, U_T^{n-1}\widehat{f})$ *converge weakly to the probability distribution* $W_{\widehat{\sigma}}$, *where* $\widehat{\sigma} = \|\widehat{h}\|_2 \geq 0$.

ii) *with probability 1, the conditional distributions* $P_n(\widetilde{f}, U_T\widetilde{f}, \ldots, U_T^{n-1}\widetilde{f}|\mathcal{H}_0)$ *given* $\mathcal{H}_0$ *of the random functions* $R_n(\widetilde{f}, U_T\widetilde{f}, \ldots, U_T^{n-1}\widetilde{f})$ *converge weakly to the (non-random) probability distribution* $W_{\widetilde{\sigma}}$, *where* $\widetilde{\sigma} = \|\widetilde{h}\|_2 \geq 0$.

Remark 3.6. The same proof as for Proposition 3.2 shows that the joint distribution of the partial sums of $(\widehat{f}, \widetilde{f})$ converge to aGaussian law with covariance matrix (σ_{ij}), where $\sigma_1^2 = \|\widehat{h}\|_2^2$, $\sigma_2^2 = \|\widetilde{h}\|_2^2$ and $\sigma_{1,2} = \sigma_{2,1} = \int \widetilde{h}\widehat{h}d\mu$. One easily deduces from this that also f is asymptotically normal with variance $\|\widehat{h} + \widetilde{h}\|_2^2$.

Proof. The assertion ii) is a direct consequence of Proposition 3.1. The assertion i) also follows from the Proposition 3.1 (applied to the filtration $(\mathcal{G}_n)_{n\in\mathbb{Z}}$) and Remark 3.3. □

Corollary 3.1. *Under the assumptions of Proposition 3.2, with probability 1, the conditional distributions* $P_n(\widetilde{f}, U_T\widetilde{f}, \ldots, U_T^{n-1}\widetilde{f}|\widehat{f}, U_T\widehat{f}, \ldots, U_T^{n-1}\widehat{f})$ *converge weakly to* $W_{\widetilde{\sigma}}$, *where* $\widetilde{\sigma} = \|\widetilde{h}\|_2 \geq 0$.

Proof. This follows from Remark 3.3, because the functions $\widehat{f}, U_T\widehat{f}, \ldots, U_T^{n-1}\widehat{f}$ are $\mathcal{G}$-measurable and $\mathcal{G} \subseteq \mathcal{H}_0$. □

Theorem 3.7. *Let T be an ergodic automorphism, and let $(\mathcal{G}_n)_{n\in\mathbb{Z}}$ and $(\mathcal{F}_n)_{n\in\mathbb{Z}}$ be a pair of T-filtrations such that $(\mathcal{G}_n)_{n\in\mathbb{Z}}$ is subordinated to $(\mathcal{F}_n)_{n\in\mathbb{Z}}$. Let $f \in L_2$ be a real-valued function satisfying*

$$\sum_{k=0}^{\infty} \|f - E^{\mathcal{F}_k} f\|_2 < \infty, \tag{3.14}$$

$$\sum_{k=0}^{\infty} \|E^{\mathcal{G}} f - E^{\mathcal{H}_{-k}} f\|_2 < \infty \tag{3.15}$$

$$\sum_{k=0}^{\infty} \|E^{\mathcal{G}} f - E^{\mathcal{G}_k} f\|_2 < \infty \tag{3.16}$$

and

$$\sum_{k=0}^{\infty} \|E^{\mathcal{G}_{-k}} f\|_2 < \infty. \tag{3.17}$$

Setting

$$\widehat{f} = E^{\mathcal{G}} f \quad \textit{and} \quad \widetilde{f} = f - E^{\mathcal{G}} f,$$

then $\widehat{f}$ and $\widetilde{f}$ admit, respectively, the representations

$$\widehat{f} = \widehat{h} + \widehat{g} - U_T\widehat{g}$$

and

$$\widetilde{f} = \widetilde{h} + \widetilde{g} - U_T\widetilde{g},$$

where

$$E^{\mathcal{G}_1}\widehat{h} = \widehat{h}, E^{\mathcal{G}_0}\widehat{h} = 0, E^{\mathcal{H}_1}\widetilde{h} = \widetilde{h}, E^{\mathcal{H}_0}\widetilde{h} = 0, \widehat{g}, \widetilde{g} \in L_2, E^{\mathcal{G}}\widetilde{g} = \widetilde{g}.$$

Moreover,

i) *the distributions* $P_n(\widehat{f}, U_T\widehat{f}, \ldots, U_T^{n-1}\widehat{f})$ *of the random functions* $R_n(\widehat{f}, U_T\widehat{f}, \ldots, U_T^{n-1}\widehat{f})$ *converge weakly to the probability distribution* $W_{\widehat{\sigma}}$, *where* $\widehat{\sigma} = \|\widehat{h}\|_2 \geq 0$.

ii) *with probability* 1, *the conditional distributions* $P_n(\widetilde{f}, U_T\widetilde{f}, \ldots, U_T^{n-1}\widetilde{f}|\mathcal{H}_0)$ *given* $\mathcal{H}_0$ *of the random functions* $R_n(\widetilde{f}, U_T\widetilde{f}, \ldots, U_T^{n-1}\widetilde{f})$ *converge weakly to the (non-random) probability distribution* $W_{\widetilde{\sigma}}$, *where* $\widetilde{\sigma} = \|\widetilde{h}\|_2 \geq 0$.

Remark 3.8. (1) Instead of (3.17) it is sometimes more convenient to verify the stronger condition

$$\sum_{k=0}^{\infty} \|E^{\mathcal{F}_{-k}} f\|_2 < \infty.$$

(2) If

$$\mathcal{H}_- = \mathcal{G},$$

then the class of functions satisfying the assumptions of Theorem 3.7 is dense in the subspace of the functions $f \in L_2$ satisfying $E^{\mathcal{G}_-} f = 0$. A sufficient condition for this can be found in subsection 4.4.

Proof of Theorem 3.7. We apply Theorem 3.5 twice. Let us show first that $\widetilde{f}$ and $(\mathcal{H}_n)_{n\in\mathbb{Z}}$ satisfy the assumptions of Theorem 3.5. We have by (3.14) and (3.15)

$$\begin{aligned}
\sum_{k=0}^{\infty} \|\widetilde{f} - E^{\mathcal{H}_k}\widetilde{f}\|_2 &= \sum_{k=0}^{\infty} \|f - E^{\mathcal{G}}f - E^{\mathcal{H}_k}f + E^{\mathcal{H}_k}E^{\mathcal{G}}f\|_2 \\
&= \sum_{k=0}^{\infty} \|f - E^{\mathcal{H}_k}f\|_2 \\
&\leq \sum_{k=0}^{\infty} \|f - E^{\mathcal{F}_k}f\|_2 < \infty
\end{aligned}$$

and

$$\begin{aligned}
\sum_{k=0}^{\infty} \|E^{\mathcal{H}_{-k}}\widetilde{f}\|_2 &= \sum_{k=0}^{\infty} \|E^{\mathcal{H}_{-k}}f - E^{\mathcal{H}_{-k}}E^{\mathcal{G}}f\|_2 \\
&= \sum_{k=0}^{\infty} \|E^{\mathcal{G}}f - E^{\mathcal{H}_{-k}}f\|_2 < \infty.
\end{aligned}$$

By (3.16) and (3.17) we can also apply Theorem 3.5 to $\widehat{f}$ and $(\mathcal{G}_n)_{n\in\mathbb{Z}}$ (instead of f and $(\mathcal{H}_n)_{n\in\mathbb{Z}}$), since

$$\sum_{k=0}^{\infty} \|E^{\mathcal{G}_{-k}}\widehat{f}\|_2 = \sum_{k=0}^{\infty} \|E^{\mathcal{G}_{-k}}E^{\mathcal{G}}f\|_2 = \sum_{k=0}^{\infty} \|E^{\mathcal{G}_{-k}}f\|_2 < \infty$$

and

$$\sum_{k=0}^{\infty} \|\widehat{f} - E^{\mathcal{G}_k} \widehat{f}\|_2 = \sum_{k=0}^{\infty} \|E^{\mathcal{G}} f - E^{\mathcal{G}_k} f\|_2 < \infty.$$

□

Corollary 3.2. *Let T be an ergodic automorphism, $(\mathcal{G}_n)_{n\in\mathbb{Z}}$ and $(\mathcal{F}_n)_{n\in\mathbb{Z}}$ be a pair of T-filtrations such that $(\mathcal{G}_n)_{n\in\mathbb{Z}}$ is immersed into $(\mathcal{F}_n)_{n\in\mathbb{Z}}$. Assume that*

$$\mathcal{F}_- \subseteq \mathcal{G}, \tag{3.18}$$

and that $f \in L_2$ is a real-valued function satisfying

$$\sum_{k=0}^{\infty} \|f - E^{\mathcal{F}_k} f\|_2 < \infty, \tag{3.19}$$

$$\sum_{k=0}^{\infty} \|E^{\mathcal{G}} f - E^{\mathcal{H}_{-k}} f\|_2 < \infty$$

and

$$\sum_{k=0}^{\infty} \|E^{\mathcal{G}_{-k}} f\|_2 < \infty.$$

Set

$$\widehat{f} = E^{\mathcal{G}} f \quad \text{and} \quad \widetilde{f} = f - E^{\mathcal{G}} f,$$

then (3.14)-(3.17) *of Theorem 3.7 are satisfied and its conclusion applies to $\widehat{f}$ and $\widetilde{f}$. Moreover, the class of functions satisfying the assumptions* (3.14)-(3.17) *is dense in* $\{f \in L_2 : E^{\mathcal{G}_-} f = 0\}$.

Remark 3.9. In many applications we have $\mathcal{F}_- = \mathcal{N}$ where $\mathcal{N}$ is the trivial σ-subfield. This obviously implies (3.18).

Proof of Corollary 3.2. We only need to verify (3.16). This can be deduced from (2.5) in the statement of Lemma 2.1 as follows:

$$\begin{aligned} \|E^{\mathcal{G}} f - E^{\mathcal{G}_k} f\|_2 &= \|E^{\mathcal{G}} f - E^{\mathcal{G}}(E^{\mathcal{F}_k} f)\|_2 \\ &= \|E^{\mathcal{G}}(f - E^{\mathcal{F}_k} f)\|_2 \\ &\le \|f - E^{\mathcal{F}_k} f\|_2, \end{aligned}$$

and (3.16) follows from (3.19). By (3.18) Remark 3.8 (2) applies and the set of functions satisfying (3.14)-(3.17) is dense in $\{f \in L_2 : E^{\mathcal{G}_-} f = 0\}$. □

4 Markov chains

4.1 A general result

Let $(\varkappa_n)_{n\in\mathbb{Z}}$ be a stationary Markov chain with state space $(S_\varkappa, \mathcal{A}_\varkappa)$ (where $S_\varkappa$ is a non-empty set and $\mathcal{A}_\varkappa$ a σ-field in $S_\varkappa$), transition probability $Q_\varkappa : S_\varkappa \times$

$\mathcal{A}_\varkappa \to [0,1]$ and stationary probability measure $\mu_\varkappa$ on $\mathcal{A}_\varkappa$. We assume that the random sequence $(\varkappa_k)_{k\in\mathbb{Z}}$ is defined on some fixed sample space $(X, \mathcal{F}, P)$ where the probability measure P is the distribution of the Markov chain with initial distribution $\mu_\varkappa$, the stationary distribution. Then every $\varkappa_k$ maps $(X, \mathcal{F}, P)$ onto $(S_\varkappa, \mathcal{A}_\varkappa, \mu_\varkappa)$ in a measurable and measure preserving way. For every $n \in \mathbb{Z}$ denote by $\mathcal{K}_n$ the σ-field in X generated by $\varkappa_n$ and by $\mathcal{H}_n$ the σ-field generated by $\{\varkappa_k : k \leq n\}$, i.e. $\mathcal{H}_n = \bigvee_{k\leq n} \mathcal{K}_k$. The shift transformation T in X preserves P and operates on the Markov chain by $\varkappa_{n+1} = \varkappa_n \circ T$ for every $n \in \mathbb{Z}$.

We use the same notation for a transition probability Q and for the corresponding transition operator $Qf(x) = \int f(y)Q(x, dy)$ (f bounded, measurable). Recall that a transition operator Q with a specified stationary probability measure λ is *ergodic* if $QF = F$ for $F \in L_2(\lambda)$ implies that F is constant. For an ergodic Q the shift transformation in the path space of the corresponding stationary Markov chain is ergodic [4], Ch.4, Lemma 7.1.

The following result on the CLT for Markov chains is well known (see [4], in a weaker form also [7]).

Proposition 4.1. *Let $(\varkappa_n)_{n\in\mathbb{Z}}$ be an ergodic stationary Markov chain with stationary probability measure $\mu_\varkappa$ and transition operator $Q_\varkappa$. If $F \in L_2(\mu_\varkappa)$ has the representation*

$$F = G - Q_\varkappa G \tag{4.20}$$

for some $G \in L_2(\mu_\varkappa)$, then, with probability 1, the conditional distributions $P_n(F \circ \varkappa_0, F \circ \varkappa_1, \ldots, F \circ \varkappa_{n-1}|\mathcal{H}_0)$ given $\mathcal{H}_0$ of the random functions $R_n(F \circ \varkappa_0, F \circ \varkappa_1, \ldots, F \circ \varkappa_{n-1})$ converge weakly to the (non-random) probability distribution W_σ, where $\sigma^2 = \|G\|_2^2 - \|Q_\varkappa G\|_2^2 \geq 0$.

Proof. We apply Proposition 3.1 to $F \circ \varkappa_0$. Indeed, the representation (3.9) has now the form

$$\begin{aligned} F \circ \varkappa_0 &= (G \circ \varkappa_1 - (Q_\varkappa G) \circ \varkappa_0) - G \circ \varkappa_1 + G \circ \varkappa_0 \\ &= H + G \circ \varkappa_0 - U_T(G \circ \varkappa_0), \end{aligned}$$

where $H = G \circ \varkappa_1 - (Q_\varkappa G) \circ \varkappa_0$ satisfies (3.8). To complete the proof it is sufficient to notice that

$$\begin{aligned} &\|H\|_2^2 \\ = \;& E_P|G \circ \varkappa_1|^2 - 2E_P\big(((G \circ \varkappa_1) \cdot (Q_\varkappa G) \circ \varkappa_0)\big) + E_P|(Q_\varkappa G) \circ \varkappa_0|^2 \\ = \;& \|G\|_2^2 - \|Q_\varkappa G\|_2^2. \end{aligned}$$

□

4.2 Markov chains fibred over invertible transformations

We keep the notation as in the previous subsection. In addition, let $(S_\pi, \mathcal{A}_\pi)$ be a measurable space and $\psi : S_\varkappa \to S_\pi$ a measurable map. ψ defines a stationary sequence $\pi_n = \psi \circ \varkappa_n$ ($n \in \mathbb{Z}$) with one-dimensional marginal $\mu_\pi = \mu_\varkappa \circ \psi^{-1}$, the image of $\mu_\varkappa$ under ψ. We assume that there exists an invertible measurable transformation V of S_π onto itself such that

$$\psi(\varkappa_{n+1}) = V(\psi(\varkappa_n)), \quad n \in \mathbb{Z}. \tag{4.21}$$

Since $(\varkappa_n)_{n\in\mathbb{Z}}$ is a stationary sequence with one-dimensional distribution $\mu_\varkappa$, it follows from (4.21) that V preserves μ_π. Next, consider the following identity for the transition operator $Q_\varkappa$, for all bounded, $\mathcal{A}_\varkappa$-measurable functions F on $S_\varkappa$ and all bounded, $\mathcal{A}_\pi$-measurable functions G on S_π :

$$(Q_\varkappa((G\circ\psi)F))(\cdot) = G(V(\psi(\cdot)))(Q_\varkappa F)(\cdot). \tag{4.22}$$

If $S_{\varkappa,z} = \psi^{-1}(z)$ denotes the fibre over $z \in S_\pi$, then property (4.22) means that the transition probability for an initial point $x \in S_\varkappa$ is concentrated on the fibre $S_{\varkappa,V(\psi(x))}$. In this case the transition operator $Q_\varkappa$ is *fibred over the transformation* V, and $(S_\pi, \mathcal{A}_\pi, \mu_\pi)$ and V are called the *base probability space* and the *base transformation*, respectively.

Fix some $x \in S_\varkappa$. We are interested in the distribution of $(\varkappa_n)_{n\geq 0}$ conditioned by the constraints $\varkappa_0 = x, \psi(\varkappa_n) = V^n(\psi(x)), n \in \mathbb{Z}$. In order to describe this behaviour let $\mathcal{C}$ be a σ-field generated by some fixed random variables π_l. The following observation follows from Propsition 4.1 by passing from the σ-field $\mathcal{H}_0$ to the coarser σ-field $\mathcal{C}$.

Proposition 4.2. *Let $Q_\varkappa$ be an ergodic transition probability with stationary probability measure $\mu_\varkappa$, and assume that $Q_\varkappa$ is fibred over a transformation V with base probability space $(S_\pi, \mathcal{A}_\pi, \mu_\pi)$.*

If $F \in L_2(\mu_\varkappa)$ has a representation (4.20)

$$F = G - Q_\varkappa G$$

for some $G \in L_2(\mu_\varkappa)$, then, with probability 1, *the conditional distributions $P_n(F\circ\varkappa_0, F\circ\varkappa_1, \ldots, F\circ\varkappa_{n-1}|\mathcal{C})$ of the random functions $R_n(F\circ\varkappa_0, F\circ\varkappa_1, \ldots, F\circ\varkappa_{n-1})$ converge weakly to the (non-random) probability distribution W_σ, where $\sigma^2 = \|G\|_2^2 - \|Q_\varkappa G\|_2^2 \geq 0$.*

The same conclusion holds for $\widetilde{F} = F - E^{\mathcal{A}'_\pi}F$, where $\mathcal{A}'_\pi = \psi^{-1}(\mathcal{A}_\pi)$.

Proof. First note that the first claim follows from Proposition 4.1. By the assumptions we have the identity

$$Q_\varkappa E^{\mathcal{A}'_\pi} = E^{\mathcal{A}'_\pi} Q_\varkappa,$$

which implies that both functions $\widehat{F}$ and $\widetilde{F}$ defined by

$$\widehat{F} = E^{\mathcal{A}'_\pi}F,\ \widetilde{F} = F - \widehat{F} \tag{4.23}$$

also satisfy (4.20), because $E^{\mathcal{A}'_\pi}(G - Q_\varkappa G) = E^{\mathcal{A}'_\pi}G - Q_\varkappa E^{\mathcal{A}'_\pi}G$. ☐

Remark 4.1. Only $\widetilde{F}$ defines a stationary process $\widetilde{f}\circ\varkappa_n$ $(n \geq 0)$ with a possibly non-generate CLT, while $\widehat{F}$ has the form $\widehat{F} = \widehat{G}\circ V - \widehat{G}$, hence is a coboundary and defines a stationary process with a degenerate limit in the CLT. For a function F with decomposition (4.23), we can always assume that the function G in a representation (4.20) has a decomposition of the form (4.23) as well, i.e.

$$\widehat{F} = E^{\mathcal{A}'_\pi}G - Q\varkappa E^{\mathcal{A}'_\pi}G = \widehat{G} - Q_\varkappa\widehat{G}$$

and

$$\widetilde{F} = G - \widehat{G} - Q_\varkappa(G - \widehat{G}) = \widetilde{G} - Q_\varkappa \widetilde{G}.$$

Under this condition (4.20) admits at most one solution.

Functions satisfying (4.20) form a dense subset in $L_2(\mu_\varkappa)$. This follows from the fact that their orthogonal complement in $L_2(\mu_\varkappa)$ is the space of $Q^*_\varkappa$-invariant functions, whence are constant by ergodicity of $Q_\varkappa$. They are also dense in the subspace of functions F, satisfying

$$E^{\mathcal{A}'_\pi} F = 0. \tag{4.24}$$

Remark 4.2. There are different strategies to obtain (4.20) for a given function. If $Q_\varkappa$ is a normal operator (in the sense that it commutes with its conjugate), very precise conditions for (4.20) to hold can be given in terms of the spectral decomposition of F relative to $Q_\varkappa$ ([4]).

For a function F a solution G to the equation (4.20) can be written down as a formal power series:

$$G = \sum_{n=0}^{\infty} Q^n_\varkappa F. \tag{4.25}$$

In some cases this series converges with respect to an appropriate norm.

Remark 4.3. Fibration over the base space is of particular interest for fibred dynamical systems (see [5]). The fibres are given by $S_{\varkappa,z} = \psi^{-1}(z)$ and the measure $\mu_\varkappa$ has a disintegration into probability measures $\mu_{\varkappa,z}$ which are supported on the fibres $S_{\varkappa,z}$. Under (4.22) fibrewise transition probabilities are defined by

$$Q^{(z,n)}_\varkappa : S_{\varkappa,z} \times \mathcal{A}(V^n(z)) \to [0,1],$$

$$Q^{(z,n)}_\varkappa(x, A) = Q^n_\varkappa(x, A), \qquad x \in S_{\varkappa,z},\ A \in \mathcal{A}(V^n(z)),$$

where $z \in S_\pi$ and $\mathcal{A}(z)$ is the restriction of the σ-field $\mathcal{A}_\varkappa$ to the fibre $S_{\varkappa,z}$. The family $(Q^{(z,n)}_\varkappa)_{z\in S_\pi, n\geq 0}$ is measurable in z and satisfies the cocycle identity in n, i.e.

$$\int_{S_{\varkappa,V^k(z)}} Q^{(V^k(z),l)}_\varkappa(u, A) Q^{(z,k)}_\varkappa(x, du) = Q^{(z,k+l)}_\varkappa(x, A), \tag{4.26}$$

for $z \in S_\pi$, $x \in S_{\varkappa,z}$, $A \in \mathcal{A}(V^{k+l}(z))$, $k, l \geq 0$. The transition probability $Q^{(z,n)}_\varkappa$ transports the conditional measure $\mu_{(\varkappa,z)}$ on the fibre $S_{\varkappa,z}$ to the conditional measure $\mu_{(\varkappa,V^n(z))}$ on $S_{\varkappa,V^n(z)}$. The condition (4.24) means that F has vanishing integrals with respect to each fibre probability measure $\mu_{\varkappa,z}$, thus defining the family of function spaces on fibres $S_{\varkappa,z}$ given by functions of vanishing integral with respect to $\mu_{(\varkappa,z)}$. The family $Q^{(z,n)}_\varkappa$ also defines a family of operators between these function spaces with the cocycle property (4.26) (the operator $Q^{(z,n)}_\varkappa$ maps functions on the fibre $S_{\varkappa,V^n(z)}$ to those on $S_{\varkappa,z}$). They also preserve integrals with respect to the conditional measures, in particular, the set of function with integral 0 is invariant with respect to these operators. Various conditions are known in the literature ensuring that this family of operators, restricted to spaces of functions with vanishing integrals over all fibres, are contractions with respect to an appropriate norm (provided

n is sufficiently large). For example, in the case of immersed finite state Markov chains considered in Theorem 5 of [2] (we shall treat the immersed case in the next section avoiding such considerations) there are only finitely many types of finite fibres with finitely many types of transition probabilities between them. Under some additional assumptions the contraction property is ensured in the uniform norm. Alternatively, assuming that $S_\varkappa$ is a metric space, we can use Hölder norms to achieve the contraction property. This technique is often used in connection with thermodynamic formalism and its relativized version (see [5] and references therein). The transfer operator considered there is a generalization of the transition operator, because it does not need to preserve the space of constant functions; however, it is a specialization at the same time, because the "reversed process" is deterministic. Notice, that there is no need to apply the Hilbert projective norm technique because we assume the existence of a stationary probability measure (though this technique is very helpful in proving the existence of these measures).

4.3 Reduction of conditional Markov chains to chains with deterministic base

In this section we sketch the application of subsection 4.2 to the general problem mentioned in the introduction. Recall that we are interested in the asymptotic distribution of $\sum_{k=0}^{n-1} f(\zeta_k)$ given $\eta_0, ..., \eta_{n-1}$, where $\zeta_k = (\xi_k, \eta_k)$ is a two component strictly stationary homogeneous Markov chain.

Let $(\zeta_k)_{k\in\mathbb{Z}}$ be a stationary homogeneous Markov chain. Its state space is denoted by $(S_\zeta, \mathcal{A}_\zeta)$ (where S_ζ is a set and $\mathcal{A}_\zeta$ is a σ-field in S_ζ), its transition probability by $Q_\zeta : S_\zeta \times \mathcal{A}_\zeta \to [0,1]$ and its stationary probability measure by μ_ζ on $\mathcal{A}_\zeta$, i.e. $E(F(\zeta_{n+1})|\zeta_k, k \le n) = (Q_\zeta F)(\zeta_n)$. We assume that the random sequence $(\zeta_k)_{k\in\mathbb{Z}}$ is defined on some fixed probability space $(X, \mathcal{F}, P)$ where the probability measure P is derived from the stationary distribution μ_ζ (as in subsection 4.2). Then every ζ_k maps $(X, \mathcal{F}, P)$ onto $(S_\zeta, \mathcal{A}_\zeta, \mu_\zeta)$ in a measurable and measure preserving way. For every $n \in \mathbb{Z}$ denote by $\mathcal{A}_n$ the σ-field in X generated by ζ_n and by $\mathcal{F}_n$ the σ-field generated by $\{\zeta_k : k \le n\}$, i.e. $\mathcal{F}_n = \bigvee_{k\le n} \mathcal{A}_k$. The shift transformation $T : X \to X$ preserves P and $\zeta_{n+1} = \zeta_n \circ T$ for every $n \in \mathbb{Z}$.

The process $(\eta_n)_{n\in\mathbb{Z}}$ as above can be described by a measurable map φ from $(S_\zeta, \mathcal{A}_\zeta)$ to a measurable space $(S_\eta, \mathcal{A}_\eta)$. We set for every $n \in \mathbb{Z}$ $\eta_n = \varphi(\zeta_n)$, $\mathcal{B}_n = \zeta_n^{-1}\mathcal{A}' \ (= \eta_n^{-1}\mathcal{A}_\eta)$, $\mathcal{G}_n = \bigvee_{k\le n} \mathcal{B}_k$, $\mathcal{A}' = \varphi^{-1}\mathcal{A}_\eta$ and $\mathcal{G} = \bigvee_{k\in\mathbb{Z}} \mathcal{B}_k$.

The Markov chain has a representation as in subsection 4.2 as follows.

Take S_π to be the set $S_\eta^{\mathbb{Z}}$ formed by two sided inifinite sequences of elements of S_η. Denote by S_ζ^- the set of all left-infinite sequences of the elements in $X = S_\zeta^{\mathbb{Z}}$. We define the shift transformation V on $S_\eta^{\mathbb{Z}}$ by $V(\dots, z_{-1}, z_0, z_1, \dots) = (\dots, z_0, z_1, z_2, \dots)$. The set $S_\varkappa$ consists of those pairs $(\mathbf{x}, \mathbf{z}) \in S_\zeta^- \times S_\pi$ with $\mathbf{x} = (\dots, x_{-1}, x_0)$, $\mathbf{z} = (\dots, z_{-1}, z_0, z_1, \dots)$ which satisfy $\varphi(x_0) = z_0$. Then we set $\psi((\mathbf{x}, \mathbf{z})) = \mathbf{z}$. The random sequence $(\varkappa_n)_{n\in\mathbb{Z}}$ can be defined on the same probability space $(X, \mathcal{F}, P)$ as $(\zeta_n)_{n\in\mathbb{Z}}$ by setting $\varkappa_n = ((\ldots, \zeta_{n-1}, \zeta_n), (\ldots, \eta_{n-1}, \eta_n, \eta_{n+1}, \ldots)), n \in \mathbb{Z}$, where η_n in the second coordinate marks the position 0 in the infinite string. It is obvious that $\varkappa_n \circ T = \varkappa_{n+1}$ and that $(\varkappa_n)_{n\in\mathbb{Z}}$ generates $\mathcal{F}$. Therefore $(X, \mathcal{F}, P)$ can be also considered as the path space of $(\varkappa_n)_{n\in\mathbb{Z}}$. Note that $(\varkappa_n)_{n\in\mathbb{Z}}$ is a Markov chain, because $(\varkappa_n)_{n\in\mathbb{Z}}$ is a random sequence for which the past can be reconstructed from the present. Now we see that we are essentially in the situation of subsection 4.2. The operator $Q_\varkappa$ can be defined correctly at least as an operator on $L_2(S_\varkappa, \mathcal{A}_\varkappa, \mu_\varkappa)$, and Proposition 4.2 applies. Given a function F on S_ζ, the problem remains to check (4.20) for the function F' defined on $S_\varkappa$ by

$$F'((\ldots, x_{-1}, x_0), (\ldots, z_{-1}, z_0, z_1, \ldots)) = F(x_0).$$

First we need to subtract from F' the function $\mathbf{z} \mapsto \int F'(u)\mu_{\varkappa,\mathbf{z}}(du)$, the conditional expectation with respect to the base. Then we may prove, for example, convergence of the series (4.25) for the function $F' - \int F'(u)\mu_{\varkappa,\mathbf{z}}(du)$. As to the behavior of the random sequence $(E(F \circ (\zeta_n)|\{\eta_k\}_{k\in\mathbb{Z}})_{n\in\mathbb{Z}}$ related to the function $\mathbf{z} \mapsto \int F'(u)\mu_{\varkappa,\mathbf{z}}(du)$, it requires some estimates showing that $\mu_{\varkappa,\mathbf{z}}$ is mainly determined by the finite part of the sequence $\mathbf{z}$. In Bezhaeva [2] this is assured by condition (A).

4.4 Immersed Markov chains

We keep the notation of the previous subsection.

Let Q be a transition probability on $S \times \mathcal{A}$ and $\mathcal{A}'$ be a σ-subfield of $\mathcal{A}$. Then Q is said to be $\mathcal{A}'$*-compatible* if the transition probability $Q(\cdot, A)$ is a $\mathcal{A}'$-measurable function for every $A \in \mathcal{A}'$.

Let $(\zeta_n)_{n\in\mathbb{Z}}$ be a stationary Markov chain and $(\eta_n)_{n\in\mathbb{Z}}$ be a random sequence defined by $\eta_n = \varphi(\zeta_n), n \in \mathbb{Z}$. We say that $(\eta_n)_{n\in\mathbb{Z}}$ is *immersed into* $(\zeta_n)_{n\in\mathbb{Z}}$, if Q_ζ is $\varphi^{-1}(\mathcal{A}_\eta)$-compatible. Under this condition a straight forward calculation shows that the sequence $(\eta_n)_{n\in\mathbb{Z}}$ is a Markov chain, and that the filtration $\mathcal{G}_n = \bigvee_{k\le n} \mathcal{A}'_k$ is immersed into $\mathcal{F}_n = \bigvee_{k\le n} \mathcal{A}_k$. There are two main properties which specifically hold in the immersed case, but not in the general situation of the previous subsection:

i) the fibrewise transition probability $Q^{(\mathbf{z},1)}$ depends on z_0 only where $\mathbf{z} = (\ldots, z_{-1}, z_0, z_1, \ldots)$;

ii) the conditional measure $\mu_\mathbf{z}$ is a function of $z_0, z_{-1}, \ldots$ (here again $\mathbf{z} = (\ldots, z_{-1}, z_0, z_1, \ldots)$).

Recall that $\mathcal{G}$ is the σ-field generated by $(\eta_n)_{n\in\mathbb{Z}}$ and $\mathcal{A}'_\pi = \psi^{-1}(\mathcal{A}_\pi)$ is the σ-field on the state space of the Markov chain $(\varkappa_n)_{n\in\mathbb{Z}}$ generated by the map ψ. In other words it is generated by the map $(\mathbf{x}, \mathbf{z}) \mapsto \mathbf{z}$, where $\mathbf{x} = (\ldots, x_{-1}, x_0)$ and $\mathbf{z} = (\ldots, z_{-1}, z_0, z_1, \ldots)$. Let Λ be the map sending $(\mathbf{x}, \mathbf{z})$ to x_0. In the following theorem we use the notations introduced above.

Theorem 4.4. *Let $(\zeta_n)_{n\in\mathbb{Z}}$ be a Markov chain and $\eta_n = \varphi(\zeta_n), n \in \mathbb{Z}$. Assume, that $(\eta_n)_{n\in\mathbb{Z}}$ is immersed into $(\zeta_n)_{n\in\mathbb{Z}}$. Let $(\varkappa_n)_{n\in\mathbb{Z}}$ denote the Markov chain associated to $(\zeta_n)_{n\in\mathbb{Z}}$ as in subsection 4.3. For a function $F' = F \circ \Lambda$ on $S_\varkappa$, define*

$$\widehat{F}' = E^{\mathcal{A}'_\pi} F', \text{ and } \widetilde{F}' = F' - \widehat{F}'.$$

If the functions F and $\widetilde{F}'$ admit representations

$$F = G - Q_\zeta G$$

and

$$\widetilde{F}' = \widetilde{G}' - Q_\varkappa \widetilde{G}',$$

where $G \in L_2(\mu_\zeta)$ and $\widetilde{G}' \in L_2(\mu_\varkappa)$, then $\widetilde{f} = \widetilde{F}' \circ \varkappa_0$ and $\widehat{f} = \widehat{F}' \circ \varkappa_0$ satisfy the assumptions of Proposition 3.2. Thus,

i) *the distributions $P_n(\widehat{f}, U_T\widehat{f}, \ldots, U_T^{n-1}\widehat{f})$ of the random functions $R_n(\widehat{f}, U_T\widehat{f}, \ldots, U_T^{n-1}\widehat{f})$ converge weakly to the probability distribution $W_{\widehat{\sigma}}$, where $\widehat{\sigma}^2 = \|G\|_2^2 - \|Q_\zeta\|_2^2 \geq 0$;*

ii) *with probability 1, the conditional distributions $P_n(\widetilde{f}, U_T\widetilde{f}, \ldots, U_T^{n-1}\widetilde{f}|\mathcal{H}_0)$ given $\mathcal{H}_0$ of the random functions $R_n(\widetilde{f}, U_T\widetilde{f}, \ldots, U_T^{n-1}\widetilde{f})$ converge weakly to the (non-random) probability distribution $W_{\widetilde{\sigma}}$, where $\widetilde{\sigma}^2 = \|\widetilde{G}'\|_2^2 - \|Q_\varkappa\widetilde{G}'\|_2^2 \geq 0$.*

Proof. We apply Proposition 3.2 to the functions $\widetilde{f}$ and $\widehat{f}$.

It is clear from the proof of Proposition 4.1 that $\widetilde{f}$ satisfies the condition (3.13) of Proposition 3.2 with $\widetilde{g} = G' \circ \varkappa_0$.

Setting $\rho_n = (\ldots, \eta_{n-1}, \eta_n)$ we introduce a stationary Markov chain $(\rho_n)_{n\in\mathbb{Z}}$ with state space S_ρ, transition operator Q_ρ and stationary measure μ_ρ. Let $\chi : S_\varkappa \to S_\rho$ be the map sending $((\ldots, x_{-1}, x_0), (\ldots, z_{-1}, z_0, z_1, \ldots))$ to $(\ldots, z_{-1}, z_0)$ and by $\mathcal{A}''$ the σ-field in $S_\varkappa$ generated by χ. Then by immersion it follows that $E^{\mathcal{A}'_\pi}(G \circ \Lambda)$ is $\mathcal{A}''$-measurable. Therefore, it can be written in the form $\widetilde{G} \circ \Lambda$ with an appropriate function $\widetilde{G}$ on S_ρ, and, applying the immersion property again, we obtain

$$E^{\mathcal{A}'_\pi}(Q_\zeta G) \circ \Lambda) = (Q_\rho \widetilde{G}) \circ \chi.$$

This implies the representation

$$\widehat{F}'(\varkappa_n) = \widetilde{G}(\rho_n) - (Q_\rho \widetilde{G})(\rho_n),$$

whence (3.12) holds for $\widehat{f}$ with $\widehat{g} = \widetilde{G}(\rho_n)$.

It follows that Proposition 3.2 applies to the function $f = \widetilde{f} + \widehat{f}$. □

Bibliography

[1] Beghdadi-Sakrani M., Émery M.: On certain probabilities equivalent to coin-tossing, d'après Schachermayer. Sém. de Probab. XXXIII, Lect. Notes in Math. **1709** (1999), 240–256.

[2] Bezhaeva Z.I.: Limit theorems for conditional Markov chains. Theory of Probability and its Applications (translated from the russian original) **163** (1971), 428–437.

[3] Billingsley P.: Convergence of Probability Measures. John Wiley & Sons, Inc. New York-London-Sydney-Toronto, 1968.

[4] Borodin A.N., Ibragimov I.A.: Limit Theorems for Functionals of Random Walks. Proc. of Steklov Inst. of Math. **195**, Providence RI, 1995.

[5] Denker M., Gordin M., Heinemann S.-M.: On the relative variational principle for fibre expanding maps. Ergodic Theory and Dynamical Systems **22** (2002), 757–782.

[6] Gordin M.I.: The central limit theorem for stationary processes. Russian. Dokl. Akad. Nauk SSSR **188**9 (1969), 739–741.

[7] Gordin M. I., Lifšic B. A.: Central limit theorem for stationary Markov processes. Russian. Dokl. Akad. Nauk SSSR **239**4 (1978), 766–767.

[8] Jacod J., Shiryaev A. N.: Limit Theorems for Stochastic Processes. Grundlehren der math. Wiss. **288** Springer Verlag Berlin etc. 1987.

[9] Kifer, Y.: Limit theorems for random transformations and processes in random environments. Trans. Amer. Math. Soc. **350** (1998), 1481–1518.

[10] Rubshtein, B.-Z.: A central limit theorem for conditional distributions. Convergence in Ergodic Theory and Probability (Bergelson, March, Rosenblatt eds.), Walter de Gruyter, Berlin 1996, 373–380.