The Stochastic Equation $Y_{t+1} = A_t Y_t + B_t$ with Non-Stationary Coefficients

Ulrich Horst *
Institut für Mathematik
Humboldt-Universität zu Berlin
Unter den Linden 6
D-10099 Berlin

March 30, 2001

Abstract

In this paper, we consider the stochastic sequence $\{Y_t\}_{t \in \mathbb{N}}$ defined recursively by the linear relation $Y_{t+1} = A_t Y_t + B_t$ in a random environment which is described by the non-stationary process $\{(A_t, B_t)\}_{t \in \mathbb{N}}$. We formulate sufficient conditions on the environment which ensure that the finite-dimensional distributions of $\{Y_t\}_{t \in \mathbb{N}}$ converge weakly to the finite-dimensional distribution of a unique stationary process. If the driving sequence $\{(A_t, B_t)\}_{t \in \mathbb{N}}$ becomes stationary in the long run, then we can establish a global convergence result. This extends results of Brandt (1986) and Borovkov (1998) from the stationary to the non-stationary case.

Key Words: Stochastic difference equation, stochastic stability, ergodicity

AMS 1991 subject classification: 60G35, 93E15, 60F99

*Support of Deutsche Forschungsgemeinschaft (SFB 373, "Quantification and Simulation of Economic Processes", Humboldt-Universität zu Berlin) is gratefully acknowledged. I would like to thank Peter Bank, Hans Föllmer and Alexander Schied for valuable discussions and an anonymous referee for a number of comments which helped to improve the presentation of the results.

1 Introduction

In this paper, we consider the stochastic sequence $\{Y_t\}_{t\in\mathbb{N}}$ defined recursively by the linear relation

$$Y_{t+1} = A_t Y_t + B_t \qquad (t \in \mathbb{N}) \tag{1}$$

in a random environment. The environment is described by the stochastic process $\{(A_t, B_t)\}_{t\in\mathbb{N}}$ defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We formulate sufficient conditions on the driving sequence $\{(A_t, B_t)\}_{t\in\mathbb{N}}$ which guarantee that the solution of (1) converges in distribution as $t \to \infty$.

Stochastic sequences of the form (1) have been extensively investigated under a mean contraction condition and under the assumption that the driving sequence $\{(A_t, B_t)\}_{t\in\mathbb{N}}$ is *stationary*. For example, Vervaat (1979) considers the case where the environment consists of i.i.d. random variables. Brandt (1986) assumes that the driving sequence is stationary and ergodic under the law $\mathbb{P}$; see also Borovkov (1998). However, in view of many applications it seems natural to consider the case where the environment is given by a *non-stationary* sequence $\{(A_t, B_t)\}_{t\in\mathbb{N}}$. For example, the process $\{Y_t\}_{t\in\mathbb{N}}$ could be a sequence of temporary equilibrium prices of a risky asset generated by the microeconomic interaction of many agents who are active on a financial market; see, e.g., Horst (2000). In such a model, the sequence $\{(A_t, B_t)\}_{t\in\mathbb{N}}$ is generated by the evolution of the "mood" of the market. Given that the "mood" is out of equilibrium but settles down in the long run, it is desirable to have sufficient conditions which ensure that the price process is driven into equilibrium.

We are going to formulate conditions on the non-stationary sequence $\{(A_t, B_t)\}_{t\in\mathbb{N}}$ which guarantee that the solution of (1) converges in law as $t \to \infty$. In a first step, we will analyse the dynamics of the process $\{Y_t\}_{t\in\mathbb{N}}$ governed by (1) under the assumption that the driving sequence has a "nice" tail structure. More precisely, we shall assume that there exists a probability measure $\mathbb{P}^*$ on $(\Omega, \mathcal{F})$ such that the driving sequence is stationary and ergodic under $\mathbb{P}^*$ and which coincides with the original measure $\mathbb{P}$ on the tail-field $\mathcal{T}$ generated by the process $\{(A_t, B_t)\}_{t\in\mathbb{N}}$. In this case, we establish a global convergence result under the measure $\mathbb{P}$, namely convergence in law of the shifted sequence $\{Y_{t+T}\}_{t\in\mathbb{N}}$ to the uniquely determined stationary solution of (1) under the measure $\mathbb{P}^*$ as $T \to \infty$. The results of Brandt (1986) correspond to the case $\mathbb{P} = \mathbb{P}^*$, and in this case one can prove almost sure convergence. Borovkov (1998) considers a situation where the environment converges in the sense of a coupling to a stationary process. This case is also covered by our method.

In a second step, we will weaken the regularity condition that the measures $\mathbb{P}^*$ and $\mathbb{P}$ coincide on the tail-field $\mathcal{T}$. Instead, we shall assume that the environment can be approximated in law by a sequence of "nice" processes. In this case, we establish a local version of the convergence result, i.e., weak convergence of the finite-dimensional distributions of the process $\{Y_{t+T}\}_{t\in\mathbb{N}}$ under $\mathbb{P}$ to those of the uniquely determined stationary solution of (1) under $\mathbb{P}^*$ as $T \to \infty$.

The paper is organized as follows. In Section 2, we formulate our main results. In Section 3, we prove our convergence result for a non-stationary but "nice" sequences $\{(A_t, B_t)\}_{t\in\mathbb{N}}$. In Section 4, we study the case where the environment can be approximated in law by "nice" processes. Section 5 is devoted to a Markovian case study where the assumption that the environment can be approximated by "nice" processes can indeed be verified.

2 Assumptions and the Main Results

Let $\psi := \{(A_t, B_t)\}_{t\in\mathbb{N}}$ be a sequence of $\mathbb{R}^2$-valued random variables defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and let $\{Y_t\}_{t\in\mathbb{N}}$ be the sequence in (1) driven by the "input" ψ. In this section, we formulate conditions which ensure that the finite-dimensional distributions of the shifted process $\{Y_{t+T}\}_{t\in\mathbb{N}}$ converge weakly to the finite-dimensional distributions of a uniquely determined stationary process as $T \to \infty$. In a first step, we shall assume that the environment ψ is stationary and ergodic under some law $\mathbb{P}^*$ and that the asymptotic behaviour of the driving sequence ψ is the same under the original measure $\mathbb{P}$ and under the law $\mathbb{P}^*$. In order to make this more precise, we introduce the σ-fields

$$\hat{\mathcal{F}}_{t,l} := \sigma\left(\{(A_t, B_t)\}_{t\le s\le l}\right) \qquad \text{and} \qquad \hat{\mathcal{F}}_t := \sigma\left(\{(A_t, B_t)\}_{s\ge t}\right), \tag{2}$$

where $0 \le t \le l \in \mathbb{N}$ and denote by

$$\mathcal{T}_\psi := \bigcap_{t\in\mathbb{N}} \hat{\mathcal{F}}_t \tag{3}$$

the tail-σ-algebra generated by the sequence ψ. We let $\mathcal{B}(\mathbb{R})$ be the Borel-σ-field on $\mathbb{R}$ and denote by $\mathbb{E}$ and $\mathbb{E}^*$ the expectation with respect to $\mathbb{P}$ and $\mathbb{P}^*$, respectively. Moreover, it will be convenient to denote by $\mathrm{Law}(X, \mathbb{P})$ the law of a random variable X defined on $(\Omega, \mathcal{F}, \mathbb{P})$.

Definition 2.1 *A driving sequence ψ is called "nice" if there exists a probability measure $\mathbb{P}^*$ on $(\Omega, \mathcal{F})$ with the following properties:*

(i) ψ is stationary and ergodic under $\mathbb{P}^*$ and satisfies one of the conditions

$$-\infty \le \mathbb{E}^* \ln|A_0| < 0 \quad \textit{and} \quad \mathbb{E}^*(\ln|B_0|)^+ < \infty \tag{4}$$

$$\mathbb{P}^*[A_0 = 0] > 0. \tag{5}$$

(ii) The asymptotic behaviour of ψ is the same under $\mathbb{P}$ and $\mathbb{P}^*$, i.e.,

$$\mathbb{P} = \mathbb{P}^* \quad \textit{on} \quad \mathcal{T}_\psi. \tag{6}$$

Remark 2.2 *(i) Our first integrability condition in (4) may be viewed as a mean contraction condition for the dynamics defined by (1). In the case $\mathbb{P} = \mathbb{P}^*$, such a condition has also been imposed by, e.g., Vervaat (1979) and Brandt (1986).*

(ii) We denote by $\|\cdot\|_{\mathcal{E}}$ the total variation of a signed measure on a measurable space $(E, \mathcal{E})$. Since

$$\lim_{t\to\infty} \|\mathbb{P} - \mathbb{P}^*\|_{\hat{\mathcal{F}}_t} = \|\mathbb{P} - \mathbb{P}^*\|_{\mathcal{T}_\psi}, \tag{7}$$

a driving sequence ψ satisfies (6) if and only if it becomes stationary in the long run. This is equivalent to the existence of a sequence $\{c_t\}_{t\in\mathbb{N}}$ satisfying $\lim_{t\to\infty} c_t = 0$ and

$$\sup_{l\ge t} ||\mathbb{P} - \mathbb{P}^*||_{\hat{\mathcal{F}}_{t,l}} \le c_t. \tag{8}$$

Here, both (7) and (8) follow from the continuity of the total variation distance along increasing and decreasing σ-algebras; a simple martingale proof can be found in, e.g., Föllmer (1979), Remark 2.1.

Before we formulate convergence results for the sequence $\{Y_t\}_{t\in\mathbb{N}}$ driven by a non-stationary input ψ, let us first consider an example where the assumption that the driving sequence is "nice" can indeed be verified.

Example 2.3 *Let $(\{\xi_t\}_{t\in\mathbb{N}}, (\mathbb{P}_\xi)_{\xi\in M})$ be a Markov chain on $(\Omega, \mathcal{F})$ with state space M and with transition operator Π. We assume that the sequence $\{\text{Law}(\xi_t, \mathbb{P}_\xi)\}_{t\in\mathbb{N}}$ converges in the total variation norm to a unique stationary measure μ^* as $t \to \infty$ and consider an environment of the form*

$$\psi := \{(f(\xi_t), g(\xi_t))\}_{t\in\mathbb{N}},$$

where $f, g : M \to \mathbb{R}$ are measurable functions. Due to Theorem 7.16 in Breiman (1968), uniqueness of the stationary probability measure μ^* implies that the sequence ψ is stationary and ergodic under

$$\mathbb{P}^*(\cdot) := \int_M \mathbb{P}_\xi(\cdot)\mu^*(d\xi).$$

Since the mapping $\xi \mapsto \mathbb{P}_\xi[\{\psi_1, \ldots, \psi_l\} \in B]$ $(l \in \mathbb{N}, B \in \mathcal{B}(\mathbb{R}^{2l}))$ is measurable and bounded above by 1 and because the sequence $\{\text{Law}(\xi_t, \mathbb{P}_\xi)\}_{t\in\mathbb{N}}$ converges in the total variation norm to μ^*, we obtain

$$\sup_{\xi,l} \|\mathbb{P}_\xi - \mathbb{P}^*\|_{\hat{\mathcal{F}}_{t,l}} \quad \leq \quad \sup_{|F|_\infty \leq 1} |\Pi^t F - \mu^*(F)|_\infty \to 0 \qquad (t \to \infty).$$

Thus, condition (8) holds for any initial distribution μ of the chain $\{\xi_t\}_{t\in\mathbb{N}}$, and so the environment $\psi = \{(f(\xi_t), g(\xi_t))\}_{t\in\mathbb{N}}$ defined on $(\Omega, \mathcal{F}, \mathbb{P}_\mu)$ is "nice" if, for example, the integrability conditions

$$-\infty \leq \int_M \ln|f| d\mu^* < 0 \quad \text{and} \quad \int_M (\ln|g|)^+ d\mu^* < 0$$

are satisfied.

Let us now turn to the solution $\{Y_t\}_{t\in\mathbb{N}}$ of (1). For a fixed environment ψ and for any initial value $Y_0 = y \in \mathbb{R}$, we have the explicit representation

$$Y_t = y_t(y, \psi) := \sum_{j=0}^{t-1} \left(\prod_{i=t-j}^{t-1} A_i \right) B_{t-j-1} + \left(\prod_{i=0}^{t-1} A_i \right) y \qquad (t \in \mathbb{N}). \tag{9}$$

If the driving process ψ is already in equilibrium, i.e., if $\mathbb{P} = \mathbb{P}^*$, then we find ourselves in the setting analysed in Brandt (1986). In such a situation, we may as well assume that the environment is defined for all $t \in \mathbb{Z}$, due to Kolmogorov's extension theorem. For notational convenience, we shall assume that the probability space $(\Omega, \mathcal{F}, \mathbb{P}^*)$ is rich enough to carry the sequence $\{(A_t, B_t)\}_{t\in\mathbb{Z}}$.

In the stationary setting, our assumptions in Definition 2.1 (i) coincide with conditions (0.4) and (0.5), respectively, in Brandt (1986). In this case, there exists a unique stationary process $\{Y_t^*\}_{t\in\mathbb{Z}}$ which satisfies $\mathbb{P}^*$-a.s. the recursive relation (1) for all $t \in \mathbb{Z}$, due to Theorem 1 in Brandt (1986). The random variable Y_t^* is $\mathbb{P}^*$-a.s. finite and takes the form

$$Y_t^* = \sum_{j=0}^{\infty} \left(\prod_{i=t-j}^{t-1} A_i \right) B_{t-j-1} \qquad (t \in \mathbb{Z}). \tag{10}$$

In the sequel we will call the process $\{Y_t^*\}_{t\in\mathbb{N}}$ the unique stationary solution of (1) under $\mathbb{P}^*$ driven by ψ. For any initial value $y \in \mathbb{R}$, the solution $\{y_t(y,\psi)\}_{t\in\mathbb{N}}$ of (1) converges almost surely to the stationary solution in the sense that

$$\lim_{t\to\infty} |y_t(y,\psi) - Y_t^*| = 0 \qquad \mathbb{P}^*\text{-a.s.}$$

We denote by ν the law of the process $\{Y_t^*\}_{t\in\mathbb{N}}$ under $\mathbb{P}^*$, and for any $0 \le t_1 < t_2 < \cdots < t_m \in \mathbb{N}$, we let $\nu_{t_1,\ldots,t_m}$ be the distribution of the vector $(Y_{t_1}^*, \ldots, Y_{t_m}^*)$, i.e.,

$$\nu := \mathrm{Law}(\{Y_t^*\}_{t\in\mathbb{N}}, \mathbb{P}^*) \quad \text{and} \quad \nu_{t_1,\ldots,t_m} := \mathrm{Law}((Y_{t_1}^*, \ldots, Y_{t_m}^*), \mathbb{P}^*). \tag{11}$$

We are now going to formulate a convergence result for the process $\{Y_t\}_{t\in\mathbb{N}}$ in a situation where the driving sequence ψ is out of equilibrium but "nice" in the sense of Definition 2.1. The following theorem will be proved in Section 3.

Theorem 2.4 *Suppose that ψ is "nice" in the sense of Definition 2.1. Then the process $\{y_t(y,\psi)\}_{t\in\mathbb{N}}$ converges in law to the uniquely determined stationary solution $\{Y_t^*\}_{t\in\mathbb{N}}$ of (1) defined on $(\Omega, \mathcal{F}, \mathbb{P}^*)$, i.e., for any initial value $y \in \mathbb{R}$, we have*

$$\mathrm{Law}(\{y_{t+T}(y,\psi)\}_{t\in\mathbb{N}}, \mathbb{P}) \stackrel{w}{\longrightarrow} \nu(\cdot) \qquad (T \to \infty). \tag{12}$$

Here, "$\stackrel{w}{\longrightarrow}$" denotes weak convergence of probability measures.

Let us now consider the case where the environment ψ does not have a "nice" tail structure. Theorem 2.6 below states that the finite-dimensional distributions of the process $\{y_t(y,\psi)\}_{t\in\mathbb{N}}$ under the law $\mathbb{P}$ converge weakly to the finite-dimensional distributions of the unique stationary solution of (1) under $\mathbb{P}^*$ as soon as ψ can be approximated in law by a suitable sequence of "nice" processes $\{\psi^n\}_{n\in\mathbb{N}}$. The approximating sequence $\{\psi^n\}_{n\in\mathbb{N}}$, $\psi^n := \{(A_t^n, B_t^n)\}_{t\in\mathbb{N}}$, is defined as follows. Let $\{\epsilon_t\}_{t\in\mathbb{N}}$ and $\{\eta_t\}_{t\in\mathbb{N}}$ be sequences of bounded random variables defined on $(\Omega, \mathcal{F})$. We put

$$A_t^n := A_t + \sigma_n \epsilon_t \quad \text{and} \quad B_t^n := B_t + \sigma_n \eta_t, \tag{13}$$

where σ_n is of the form $\sigma_n = \frac{c}{n}$ for some $c > 0$ which will be specified in Remark 2.5 (i) below. Finally, we introduce a driving sequence $\widetilde{\psi} := \{\widetilde{\psi}_t\}_{t\in\mathbb{N}}$ by

$$\widetilde{\psi}_t := (|A_t| + c|\epsilon_t|, |B_t| + c|\eta_t|). \tag{14}$$

Assumption 1 *The environment ψ is stationary and ergodic under $\mathbb{P}^*$ and satisfies*

$$-\infty < \mathbb{E}^* \ln |A_0| < 0 \quad \textit{and} \quad \mathbb{E}^* (\ln |B_0|)^+ < \infty.$$

The sequences $\widetilde{\psi}, \psi^1, \psi^2, \ldots$ are "nice" in the sense of Definition 2.1 and satisfy

$$\mathbb{E}^* (\ln |A_0^n|)^+ \to \mathbb{E}^* (\ln |A_0|)^+, \quad \mathbb{E}^* (\ln |B_0^n|)^+ \to \mathbb{E}^* (\ln |B_0|)^+ \quad (n \to \infty) \tag{15}$$

$$-\infty < \mathbb{E}^* \ln |A_0^n| \to \mathbb{E}^* \ln |A_0|, \quad \mathbb{E}^* \ln |B_0^n| \to \mathbb{E}^* \ln |B_0| \quad (n \to \infty). \tag{16}$$

In Section 5, we will consider a Markovian model where this assumption can indeed be verified.

Remark 2.5 *(i) Under our Assumption 1 we can always choose $c > 0$ such that*

$$-\infty < \mathbb{E}^* \ln[|A_0| + c|\epsilon_0|] < 0 \quad \text{and} \quad \mathbb{E}^* [\ln(|B_0| + c|\eta_0|)]^+ < \infty.$$

In this case, it follows from (16) that the sequences $\widetilde{\psi}, \psi^1, \psi^2, \dots$ satisfy (4).

(ii) In the stationary situation, i.e., for $\mathbb{P} = \mathbb{P}^*$, our assumptions (15) and (16) coincide with conditions (0.11) and (0.10), respectively, in Brandt (1986).

Let us now state our main result which will be proved in Section 4 below.

Theorem 2.6 *Suppose that Assumption 1 is satisfied. Then the finite dimensional distributions of the process $\{y_t(y,\psi)\}_{t\in\mathbb{N}}$ ($y \in \mathbb{R}$) under the law $\mathbb{P}$ converge weakly to the finite dimensional distributions of the unique stationary solution $\{Y_t^*\}_{t\in\mathbb{N}}$ of (1) under $\mathbb{P}^*$. More precisely, for any $m \in \mathbb{N}$ and for all $t_1 < \ldots < t_m$, we have*

$$\mathrm{Law}((y_{t_1+T}(y,\psi), \dots, y_{t_m+T}(y,\psi)), \mathbb{P}) \xrightarrow{w} \nu_{t_1,\dots,t_m}(\cdot) \qquad (T \to \infty)$$

3 Stochastic Sequences Driven by "Nice" Processes

This section is devoted to the proof of Theorem 2.4, and so we assume that the driving sequence ψ is "nice" in the sense of Definition 2.1. Let us start with the following result.

Lemma 3.1 *For any $l \in \mathbb{N}$, we have that*

$$\lim_{t\to\infty} \prod_{i=l}^{t} |A_i| = 0 \qquad \mathbb{P}\text{- and } \mathbb{P}^*\text{- a.s.}$$

Proof: Observe that

$$\prod_{i=l}^{t} |A_i| = \left\{ \exp\left(\frac{1}{t-l} \sum_{i=l}^{t-1} \ln|A_i| \right) \right\}^{t-l}.$$

Thus, we have

$$\left\{ \lim_{t\to\infty} \prod_{i=l}^{t} |A_i| = 0 \right\} \supseteqq \bigcup_{i=l}^{\infty} \{A_i = 0\} \cup C,$$

where

$$C := \bigcap_n \bigcup_{m\geq n} \left\{ \limsup_{t\to\infty} \frac{1}{t-m} \sum_{i=m}^{t-1} \ln|A_i| < 0 \right\}.$$

It is therefore enough to verify

$$\mathbb{P}\left[\bigcup_{i=l}^{\infty} \{A_i = 0\} \cup C \right] = \mathbb{P}^* \left[\bigcup_{i=l}^{\infty} \{A_i = 0\} \cup C \right] = 1.$$

Let us first assume that $\mathbb{P}^*[A_0 = 0] > 0$. Since ψ is stationary under $\mathbb{P}^*$ and because the event $\{A_t = 0 \text{ infinitely often}\}$ is a tail-event, it follows from (6) and from Theorem A.1.1.2 in Brandt, Franken, and Lisek (1990) that

$$\mathbb{P}[A_t = 0 \text{ infinitely often}] = \mathbb{P}^*[A_t = 0 \text{ infinitely often}] = 1.$$

This yields $\mathbb{P}\left[\bigcup_{i=l}^{\infty}\{A_i = 0\}\right] = \mathbb{P}^*\left[\bigcup_{i=l}^{\infty}\{A_i = 0\}\right] = 1$.

Suppose now that $\mathbb{P}^*[A_0 = 0] = 0$. In this case, we deduce form (4) and (6) that $\mathbb{P}[C] = \mathbb{P}^*[C] = 1$ because $C \in \mathcal{T}_\psi$ and because ψ is stationary and ergodic under $\mathbb{P}^*$. □

We are now ready to verify weak convergence of the one-dimensional distributions of the process $\{y_t(y,\psi)\}_{t\in\mathbb{N}}$ defined in (9) to the measure ν_0 defined in (11).

Proposition 3.2 *For each initial value $y \in \mathbb{R}$ we have that*

$$\text{Law}(y_t(y,\psi), \mathbb{P}) \xrightarrow{w} \nu_0(\cdot) \qquad (t \to \infty). \tag{17}$$

Proof: In order to verify (17) it is enough to show that

$$\lim_{t\to\infty} \int F(y_t(y,\psi))d\mathbb{P} = \int F d\nu_0$$

for any Lipschitz continuous function $F : \mathbb{R} \to \mathbb{R}$ with compact support. To this end, we will first approximate $y_t(y,\psi)$ by an $\hat{\mathcal{F}}_l$-measurable random variable $y_t^l(y,\psi)$ and show that

$$\left| \int F(y_t^l(y,\psi))d\mathbb{P} - \int F d\nu_0 \right| < \epsilon$$

whenever $l, t \in \mathbb{N}$ are large enough. This estimate will then be extended to the case $l = 0$.

1. For each $l \in \mathbb{N}$ and for all $t \geq l$, let us consider the random variable $y_t^l(y,\psi)$ given by

$$y_t^l(y,\psi) := \sum_{j=0}^{t-1-l} \left(\prod_{i=t-j}^{t-1} A_i \right) B_{t-j-1} + \left(\prod_{i=l}^{t-1} A_i \right) y. \tag{18}$$

   Note that $y_t^l(y,\psi)$ is $\hat{\mathcal{F}}_l$-measurable, where the σ-field $\hat{\mathcal{F}}_l$ is defined in (2). We have

$$\begin{aligned} |y_t(y,\psi) - y_t^l(y,\psi)| &\leq \left| \sum_{j=t-l}^{t-1} \left(\prod_{i=t-j}^{t-1} A_i \right) B_{t-j-1} + y \prod_{i=0}^{t-1} A_i - y \prod_{i=l}^{t-1} A_i \right| \\ &\leq \sum_{j=1}^{l} \left(\prod_{i=j}^{t-1} |A_i| \right) |B_{j-1}| + |y| \prod_{i=0}^{t-1} |A_i| + |y| \prod_{i=l}^{t-1} |A_i|, \end{aligned}$$

   and so it follows from Lemma 3.1 that

$$\lim_{t\to\infty} \left| y_t(y,\psi) - y_t^l(y,\psi) \right| = 0 \qquad \mathbb{P}\text{- and } \mathbb{P}^*\text{-a.s.} \tag{19}$$

2. Let us now fix $\epsilon > 0$. It follows from Theorem 1 in Brandt (1986) that there exists a constant $T_1 = T_1(\epsilon) \in \mathbb{N}$ such that

$$\left| \int F(y_t(y,\psi))d\mathbb{P}^* - \int F d\nu_0 \right| < \frac{\epsilon}{4} \qquad (t \geq T_1). \tag{20}$$

   Our aim is now to show that (20) holds true with $\mathbb{P}^*$ replaced by $\mathbb{P}$.

3. Since the environment ψ is "nice" and because the random variable $y_t^l(y,\psi)$ is $\hat{\mathcal{F}}_l$-measurable, we deduce from (7) that there exist a sequence $\{c_l\}_{l\in\mathbb{N}}$ satisfying $\lim_{l\to\infty} c_l = 0$ and

$$\sup_t \left| \int F(y_t^l(y,\psi))\,(d\mathbb{P} - d\mathbb{P}^*) \right| \quad \leq \quad c_l.$$

For the rest of the proof we fix $l \in \mathbb{N}$ such that $c_l \leq \frac{\epsilon}{4}$. Since the mapping $F : \mathbb{R} \to \mathbb{R}$ is Lipschitz continuous and has compact support it follows from (19) that there exists a constant $T_2 = T_2(l) \in \mathbb{N}$ such that

$$\left| \int \{F(y_t^l(y,\psi)) - F(y_t(y,\psi))\}\, d\mathbb{P} \right| \quad \leq \quad \frac{\epsilon}{4}$$

and

$$\left| \int \{F(y_t^l(y,\psi)) - F(y_t(y,\psi))\}\, d\mathbb{P}^* \right| \quad \leq \quad \frac{\epsilon}{4}$$

for all $t \geq T_2$. Thus, for all $t \geq \max\{T_1, T_2\}$ we obtain

$$\left| \int F(y_t(y,\psi))d\mathbb{P} - \int F d\nu_0 \right| \quad \leq \quad \epsilon.$$

This implies weak convergence of the sequence $\{\text{Law}(y_t(y,\psi),\mathbb{P})\}_{t\in\mathbb{N}}$ to ν_0.

□

We are now going to prove our global convergence result.

Proof of Theorem 2.4: From the proof of Proposition 3.2 we can easily deduce that the finite dimensional distributions of the shifted process $\{y_{t+T}(y,\psi)\}_{t\in\mathbb{N}}$ under the measure $\mathbb{P}$ converge weakly to the finite dimensional distributions of the unique stationary solution $\{Y_t^*\}_{t\in\mathbb{N}}$ of (1) under the law $\mathbb{P}^*$ as $T \to \infty$. Indeed, let $m \in \mathbb{N}$ and $l \leq t_1 < t_2 < \cdots < t_m \in \mathbb{N}$ be given. We put

$$y_{t_1,\cdots,t_m}(y,\psi) := (y_{t_1}(y,\psi),\ldots,y_{t_m}(y,\psi)) \quad \text{and} \quad y^l_{t_1,\cdots,t_m}(y,\psi) := (y^l_{t_1}(y,\psi),\ldots,y^l_{t_m}(y,\psi)).$$

As in the proof of Theorem 3.2 we have that

$$\lim_{t_1\to\infty} \left| y_{t_1,\cdots,t_m}(y,\psi) - y^l_{t_1,\cdots,t_m}(y,\psi) \right| = 0 \qquad \mathbb{P}\text{- and } \mathbb{P}^*\text{-a.s.},$$

where $|\cdot|$ denotes the Euclidean distance. Since, for any Borel set $B \subset \mathbb{R}^m$, the event $\{y^l_{t_1,\cdots,t_m}(y,\psi) \in B\}$ belongs to the σ-algebra $\hat{\mathcal{F}}_l$, we can use the same arguments as in the proof of Proposition 3.2 in order to verify weak convergence of the finite-dimensional distributions.

Thus, it remains to show that the family of random variables $(\{y_{t+T}(y,\psi)\}_{t\in\mathbb{N}})_{T\in\mathbb{N}}$ – viewed as a family of random variables taking values in the space $\mathbb{R}^{\mathbb{N}}$ – is tight. For this, we have to prove that, for any $\epsilon > 0$ and for all $t \in \mathbb{N}$, there exists a compact set $K_t \subset \mathbb{R}$ such that

$$\sup_T \mathbb{P}[y_{t+T}(y,\psi) \in K_t] \geq 1 - \epsilon;$$

see, e.g., Ethier and Kurtz (1986), Theorem 3.7.2.

To this end, we fix $\epsilon > 0$. Since the random variable Y_0^* is $\mathbb{P}^*$-a.s. finite there exists $k \in \mathbb{R}$ satisfying

$$\mathbb{P}^*[|Y_0^*| \leq k] > 1 - \frac{\epsilon}{4}.$$

Let us now fix $t \in \mathbb{N}$. Since the driving sequence ψ is "nice", we can choose $l \in \mathbb{N}$ such that

$$\sup_{B,T} \left| \mathbb{P}[y^l_{t+T}(y,\psi) \in B] - \mathbb{P}^*[y^l_{t+T}(y,\psi) \in B] \right| < \frac{\epsilon}{4},$$

Due to (19), there exists $T_0 = T_0(\epsilon, k)$ such that, for all $T \geq T_0$, we obtain

$$\mathbb{P}[|y_{t+T}(y,\psi)| \leq k + 2\epsilon] \geq \mathbb{P}[|y^l_{t+T}(y,\psi)| \leq k + \epsilon] - \frac{\epsilon}{4}.$$

Moreover, the process $\{y^l_{t+T}(y,\psi)\}_{T\in\mathbb{N}}$ converges $\mathbb{P}^*$-a.s. to the stationary sequence $\{Y^*_{t+T}\}_{T\in\mathbb{N}}$, due to Theorem 1 in Brandt (1986) and (19). This yields

$$\mathbb{P}^*[|y^l_{t+T}(y,\psi)| \leq k + \epsilon] \geq \mathbb{P}^*[|Y^*_{t+T}| \leq k] - \frac{\epsilon}{4} = \mathbb{P}^*[|Y^*_0| \leq k] - \frac{\epsilon}{4} \tag{21}$$

for all $T \geq T_0 = T_0(\epsilon, k)$, and so we obtain

$$\begin{aligned} \sup_{t \geq T_0} \mathbb{P}[|y_{t+T}(y,\psi)| \leq k + 2\epsilon] &\geq \sup_{t \geq T_0} \mathbb{P}[|y^l_{t+T}(y,\psi)| \leq k + \epsilon] - \frac{\epsilon}{4} \\ &\geq \sup_{t \geq T_0} \mathbb{P}^*[|y^l_{t+T}(y,\psi)| \leq k + \epsilon] - \frac{\epsilon}{2} \\ &\geq 1 - \epsilon. \end{aligned}$$

Thus, we can choose a compact set $K_t = K_t(\epsilon) \subset \mathbb{R}$ which satisfies $\sup_T \mathbb{P}[y_{t+T}(y,\psi) \in K_t] \geq 1 - 2\epsilon$. This proves the theorem. □

Our techniques may also be used in order to analyse the following situation which is studied in, e.g., Borovkov (1998), Chapter 3, Section 11: Let $\{(A_t, B_t)\}_{t\in\mathbb{N}}$ and $\{(\hat{A}_t, \hat{B}_t)\}_{t\in\mathbb{N}}$ be sequences of $\mathbb{R}^2$-valued random variables defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P}^*)$ and assume that the process $\{(\hat{A}_t, \hat{B}_t)\}_{t\in\mathbb{N}}$ is stationary and ergodic under $\mathbb{P}^*$. Following Borovkov (1998), we say that the sequence $\{(A_t, B_t)\}_{t\in\mathbb{N}}$ is coupling convergent to $\{(\hat{A}_t, \hat{B}_t)\}_{t\in\mathbb{N}}$ if

$$\lim_{T\to\infty} \mathbb{P}^*[(A_t, B_t) = (\hat{A}_t, \hat{B}_t) \text{ for all } t \geq T] = 1.$$

In such a situation, the process $\{(A_t, B_t)\}_{t\in\mathbb{N}}$ is not necessarily "nice" in the sense of our Definition 2.1. However, it satisfies

$$\lim_{T\to\infty} \sup_{B\in\mathcal{B}(\mathbb{R}^{\mathbb{N}})} \left| \mathbb{P}^*[\{(A_t, B_t)\}_{t\geq T} \in B] - \mathbb{P}^*[\{(\hat{A}_t, \hat{B}_t)\}_{t\geq T} \in B] \right| = 0. \tag{22}$$

It is easy to see that (22) is sufficient to carry out the arguments in the proofs of Lemma 3.1, Proposition 3.2 and Theorem 2.4. Thus, we have the following variant of Theorem 2.4.

Theorem 3.3 *Let $\psi = \{(A_t, B_t)\}_{t\in\mathbb{N}}$ and $\hat{\psi} = \{(\hat{A}_t, \hat{B}_t)\}_{t\in\mathbb{N}}$ be sequences of $\mathbb{R}^2$-valued random variables defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P}^*)$ such that the environment $\hat{\psi}$ is stationary and ergodic under $\mathbb{P}^*$. If (22) and one of the two conditions*

$$-\infty \leq \mathbb{E}^* \ln|\hat{A}_0| < 0 \quad \textit{and} \quad \mathbb{E}^* (\ln|\hat{B}_0|)^+ < \infty$$

$$\mathbb{P}^*[\hat{A}_0 = 0] > 0$$

are satisfied, then we have that

$$\text{Law}(\{y_{t+T}(y,\psi)\}_{t\in\mathbb{N}}, \mathbb{P}^*) \xrightarrow{w} \text{Law}(\{\hat{Y}_t\}_{t\in\mathbb{N}}, \mathbb{P}^*) \qquad (T \to \infty).$$

Here, $\{\hat{Y}_t\}_{t\in\mathbb{N}}$ denotes the unique stationary solution of (1) driven by the input $\hat{\psi}$.

4 Stochastic Sequences Driven by "Almost Nice" Processes

This subsection is devoted to the proof of Theorem 2.6, and so we will assume that our Assumption 1 is satisfied. In (13) we introduced a sequence of "nice" environments $\{\psi^n\}_{n\in\mathbb{N}}$ defined on $(\Omega,\mathcal{F})$ such that $\mathrm{Law}(\psi^n,\mathbb{P}) \xrightarrow{w} \mathrm{Law}(\psi,\mathbb{P})$ as $n\to\infty$. Let $\{Y_t^n\}_{t\in\mathbb{N}}$ be the unique stationary solution of (1) under $\mathbb{P}^*$ driven by ψ^n and put $\nu_0^n := \mathrm{Law}(Y_0^n,\mathbb{P}^*)$. We have

$$\mathrm{Law}(y_t(y,\psi^n),\mathbb{P}) \xrightarrow{w} \nu_0^n \qquad (t\to\infty), \tag{23}$$

due to Theorem 2.4. Our aim is now to prove that the process $\{y_t(y,\psi)\}_{t\in\mathbb{N}}$ converges in law to the unique weak limit ν_0 of the sequence $\{\nu_0^n\}_{n\in\mathbb{N}}$. To this end, we shall apply a continuity result which follows immediately from Theorem 2 in Brandt (1986).

Lemma 4.1 *Under the Assumptions of Theorem 2.6 we have that*

$$\nu_0^n \xrightarrow{w} \nu_0 \qquad (n\to\infty).$$

We are now ready to prove a uniform approximation result which will turn out to be the key for the proof of Theorem 2.6.

Proposition 4.2 *For any Lipschitz continuous function $F:\mathbb{R}\to\mathbb{R}$ with compact support we have*

$$\lim_{n\to\infty}\sup_t \mathbb{E}\,|F(y_t(y,\psi)) - F(y_t(y,\psi^n))| = 0.$$

Proof: Let F be a Lipschitz continuous function on $\mathbb{R}$ with compact support. With no loss of generality we may assume that F is Lipschitz with constant 1 and that diam(supp F)=1. In this case, we obtain $|F(x)-F(y)| \le G(|x-y|)$, where $G(x) := |x|\wedge 1$. It is therefore enough to show that

$$\lim_{n\to\infty}\sup_t \mathbb{E}G(|y_t(y,\psi) - y_t(y,\psi^n)|) = 0. \tag{24}$$

Let us put $\mathcal{I}_j := \{t-j,\ldots,t-1\}$ and $\mathcal{I}_0 := \emptyset$. The explicit representation (9) of $y_t(y,\psi)$ yields

$$\begin{aligned}
|y_t(y,\psi) - y_t(y,\psi^n)| &\le \sum_{j=1}^{t-1}\sum_{\mathcal{I}\subsetneq\mathcal{I}_j}\left(\prod_{i\in\mathcal{I}}|A_i|\prod_{i\in\mathcal{I}\setminus\mathcal{I}_j}\frac{c}{n}|\epsilon_i|\right)|B_{t-j-1}| \\
&\quad + \sum_{j=0}^{t-1}\sum_{\mathcal{I}\subset\mathcal{I}_j}\left(\prod_{i\in\mathcal{I}}|A_i|\prod_{i\in\mathcal{I}\setminus\mathcal{I}_j}\frac{c}{n}|\epsilon_i|\right)\frac{c}{n}|\eta_{t-j-1}| \\
&\quad + \sum_{\mathcal{I}\subsetneq\mathcal{I}_t}\left(\prod_{i\in\mathcal{I}}|A_i|\prod_{i\in\mathcal{I}_t\setminus\mathcal{I}}\frac{c}{n}|\epsilon_i|\right)|y| \\
&\le \frac{1}{n}\left\{\sum_{j=0}^{t-1}\sum_{\mathcal{I}\subset\mathcal{I}_j}\left(\prod_{i\in\mathcal{I}}|A_i|\prod_{i\in\mathcal{I}\setminus\mathcal{I}_j}c|\epsilon_i|\right)(|B_{t-j-1}|+c|\eta_{t-j-1}|)\right. \\
&\quad \left. + \sum_{\mathcal{I}\subset\mathcal{I}_t}\left(\prod_{i\in\mathcal{I}}|A_i|\prod_{i\in\mathcal{I}_t\setminus\mathcal{I}}c|\epsilon_i|\right)|y|\right\} \\
&= \frac{1}{n}y_t(|y|,\widetilde{\psi}),
\end{aligned}$$

where $\widetilde{\psi}$ is defined in (14). Thus, it suffices to show that

$$\limsup_{n\to\infty} \sup_t \mathbb{E}\left[G\left(\frac{1}{n}y_t(|y|,\widetilde{\psi})\right)\right] = 0. \tag{25}$$

To this end, we proceed in several steps.

1. Let us first show that (25) holds true for $\mathbb{P} = \mathbb{P}^*$. The unique stationary solution $\{\widetilde{Y}_t\}_{t\in\mathbb{N}}$ of (1) under $\mathbb{P}^*$ driven by the "nice" environment $\widetilde{\psi}$ satisfies

$$\lim_{n\to\infty} \sup_t \mathbb{E}^*\left[G\left(\frac{1}{n}\widetilde{Y}_t\right)\right] = \lim_{n\to\infty} \mathbb{E}^*\left[G\left(\frac{1}{n}\widetilde{Y}_0\right)\right] = G(0) = 0 \tag{26}$$

by Lebesgue's theorem. This yields

$$\limsup_{n\to\infty} \sup_t \mathbb{E}^* G\left(\frac{1}{n}y_t(|y|,\widetilde{\psi})\right) \le \limsup_{n\to\infty} \sup_t \mathbb{E}^*\left|G\left(\frac{1}{n}y_t(|y|,\widetilde{\psi})\right) - G\left(\frac{1}{n}\widetilde{Y}_t\right)\right|.$$

Since the driving sequence $\widetilde{\psi}$ is "nice" in the sense of Definition 2.1, it follows from Theorem 1 in Brandt (1986) that

$$\lim_{t\to\infty}\left|y_t(|y|,\widetilde{\psi}) - \widetilde{Y}_t\right| = 0 \qquad \mathbb{P}^*\text{-a.s.}$$

Thus, for any $\epsilon > 0$, there exists $T \in \mathbb{N}$ such that

$$\mathbb{E}^*\left|G\left(\frac{1}{n}y_t(|y|,\widetilde{\psi})\right) - G\left(\frac{1}{n}\widetilde{Y}_t\right)\right| \le \mathbb{E}^*\left[\left|\frac{1}{n}y_t(|y|,\widetilde{\psi}) - \frac{1}{n}\widetilde{Y}_t\right| \wedge 1\right] \le \epsilon$$

for all $t \ge T$ and all $n \in \mathbb{N}$. Hence, it follows from $G(0) = 0$ and from Lebesgue's Theorem that

$$\begin{aligned}
& \limsup_{n\to\infty} \sup_t \mathbb{E}^*\left|G\left(\frac{1}{n}y_t(|y|,\widetilde{\psi})\right) - G\left(\frac{1}{n}\widetilde{Y}_t\right)\right| \\
\le\ & \limsup_{n\to\infty} \sup_{t\ge T} \mathbb{E}^*\left|G\left(\frac{1}{n}y_t(|y|,\widetilde{\psi})\right) - G\left(\frac{1}{n}\widetilde{Y}_t\right)\right| \\
& + \limsup_{n\to\infty} \sum_{t=1}^{T}\left\{\mathbb{E}^* G\left(\frac{1}{n}y_t(|y|,\widetilde{\psi})\right) + \mathbb{E}^* G\left(\frac{1}{n}\widetilde{Y}_t\right)\right\} \\
\le\ & \epsilon.
\end{aligned}$$

Since $\epsilon > 0$ is arbitrary, we obtain

$$\lim_{n\to\infty} \sup_t \mathbb{E}^* G\left(\frac{1}{n}y_t(|y|,\widetilde{\psi})\right) = 0. \tag{27}$$

Our aim is now to prove that (27) holds true not only for $\mathbb{P}^*$ but also for $\mathbb{P}$.

2. To this end, we define a random variable $y_t^l(|y|,\widetilde{\psi})$ by analogy with (18) and shall first show that there exists $l \in \mathbb{N}$ such that the following estimate holds uniformly in $n \in \mathbb{N}$:

$$\sup_t \mathbb{E}G\left(\frac{1}{n}y_t^l(|y|,\widetilde{\psi})\right) \le \sup_t \mathbb{E}^* G\left(\frac{1}{n}y_t^l(|y|,\widetilde{\psi})\right) + \epsilon. \tag{28}$$

For any Borel set $B \subset \mathbb{R}$, the event $\{y_t^l(|y|,\widetilde{\psi}) \in B\}$ belongs to the σ-field $\sigma(\{\widetilde{\psi}_s\}_{s\ge l})$. Since $\widetilde{\psi}$ is "nice" by Assumption 1, it follows from Remark 2.2 (ii) that there exists a sequence $\{c_l\}_{l\in\mathbb{N}}$ satisfying $\lim_{l\to\infty} c_l = 0$ and

$$\int G\left(\frac{1}{n}y_t^l(|y|,\widetilde{\psi})\right)(d\mathbb{P} - d\mathbb{P}^*) \le \sup_{B\in\mathcal{B}(\mathbb{R}),t}\left|\mathbb{P}[y_t^l(|y|,\widetilde{\psi}) \in B] - \mathbb{P}^*[y_t^l(|y|,\widetilde{\psi}) \in B]\right| \le c_l.$$

This yields

$$\sup_t \mathbb{E} G\left(\frac{1}{n} y_t^l(|y|, \widetilde{\psi})\right) \leq \sup_t \mathbb{E}^* G\left(\frac{1}{n} y_t^l(|y|, \widetilde{\psi})\right) + c_l, \tag{29}$$

and so (28) holds true whenever l is large enough.

We can use the same arguments as in the proof of Theorem 2.4 in order to show that

$$\lim_{t\to\infty} \left| y_t^l(|y|, \widetilde{\psi}) - y_t(|y|, \widetilde{\psi}) \right| = 0 \qquad \mathbb{P}\text{- and } \mathbb{P}^*\text{-a.s.} \qquad (l \in \mathbb{N}). \tag{30}$$

In particular, for any fixed $l \in \mathbb{N}$, we have that

$$\lim_{t\to\infty} \left| y_t^l(|y|, \widetilde{\psi}) - \widetilde{Y}_t \right| = 0 \qquad \mathbb{P}^*\text{-a.s.}$$

Thus, it follows from (29) and by analogy with Step 1 that there exists $l \in \mathbb{N}$ such that

$$\limsup_{n\to\infty} \sup_t \mathbb{E} G\left(\frac{1}{n} y_t^l(|y|, \widetilde{\psi})\right) \leq \epsilon + \limsup_{n\to\infty} \sup_t \mathbb{E}^* G\left(\frac{1}{n} y_t^l(|y|, \widetilde{\psi})\right) = \epsilon. \tag{31}$$

We will now extend this estimate to the case $l = 0$, and this yields (25).

3. Let us choose $l \in \mathbb{N}$ such that (31) holds true. Using $G(0) = 0$ and (31), we obtain

$$\begin{aligned} & \limsup_{n\to\infty} \sup_t \mathbb{E} G\left(\frac{1}{n} y_t(|y|, \widetilde{\psi})\right) \\ \leq\ & \limsup_{n\to\infty} \sup_t \left| \mathbb{E}\left\{ G\left(\frac{1}{n} y_t(|y|, \widetilde{\psi})\right) - G\left(\frac{1}{n} y_t^l(|y|, \widetilde{\psi})\right) \right\} \right| + \epsilon. \end{aligned}$$

Due to (30), we can choose $T = T(\epsilon, l)$ such that

$$\sup_{t \geq T} \mathbb{E} \left| G\left(\frac{1}{n} y_t(|y|, \widetilde{\psi})\right) - G\left(\frac{1}{n} y_t^l(|y|, \widetilde{\psi})\right) \right| < \epsilon,$$

and so Lebesgue's theorem yields

$$\limsup_{n\to\infty} \sup_t \mathbb{E} G\left(\frac{1}{n} y_t(|y|, \widetilde{\psi})\right) < 2\epsilon.$$

Thus, we have shown (25). □

The next result follows immediately from Proposition 4.2.

Corollary 4.3 *Let $m \in \mathbb{N}$ be given. For any real-valued Lipschitz continuous function F on $\mathbb{R}^m$ with compact support we have that*

$$\lim_{n\to\infty} \sup_{t_1,\ldots,t_m} \mathbb{E}\left| F(y_{t_1}(y, \psi^n), \ldots, y_{t_m}(y, \psi^n)) - F(y_{t_1}(y, \psi), \ldots, y_{t_m}(y, \psi)) \right| = 0.$$

Using Proposition 4.2 and Theorem 3.2, we can now verify our local convergence result stated in Theorem 2.6.

Proof of Theorem 2.6: Let us first establish convergence of the one-dimensional distributions. To this end, it is enough to prove that

$$\lim_{t\to\infty} \int F(y_t(y, \psi) d\mathbb{P} = \int F d\nu_0$$

for any real-valued Lipschitz continuous function F on $\mathbb{R}$ with compact support. We have that

$$\begin{aligned} & \left| \int F(y_t(y,\psi)) d\mathbb{P} - \int F d\nu_0 \right| \\ \leq \; & \left| \int \left\{ F(y_t(y,\psi)) - F(y_t(y,\psi^n)) \right\} d\mathbb{P} \right| + \left| \int F(y_t(y,\psi^n)) d\mathbb{P} - \int F d\nu_0^n \right| + \left| \int F d\nu_0^n - \int F d\nu_0 \right|. \end{aligned}$$

According to Proposition 4.2 there exists a constant $N_1 = N_1(\epsilon) \in \mathbb{N}$ such that

$$\left| \int \left\{ F(y_t(y,\psi)) - F(y_t(y,\psi^n)) \right\} d\mathbb{P} \right| \leq \frac{\epsilon}{3}$$

for all $n \geq N_1$ and for all $t \in \mathbb{N}$. Moreover, we can choose $N_2 = N_2(\epsilon) \in \mathbb{N}$ satisfying

$$\left| \int F d\nu_0^n - \int F d\nu_0 \right| \leq \frac{\epsilon}{3} \qquad (n \geq N_2),$$

due to Lemma 4.1.

Let us now fix $n_0 \geq \max\{N_1, N_2\}$. Since the environment ψ^{n_0} is "nice", it follows from Theorem 2.4 that there exists a constant $T(n_0)$ such that

$$\left| \int F(y_t(y,\psi^{n_0})) d\mathbb{P} - \int F d\nu_0^{n_0} \right| < \frac{\epsilon}{3}$$

for all $t \geq T(n_0)$. In particular, we get

$$\left| \int F(y_t(y,\psi)) d\mathbb{P} - \int F d\nu_0 \right| \leq \epsilon$$

for all $t \geq T(n_0)$. This shows weak convergence of the sequence $\{\text{Law}(y_t(y,\psi), \mathbb{P})\}_{t \in \mathbb{N}}$ to the probability measure ν_0. Convergence of the finite-dimensional distributions follows by analogy with the proof of Theorem 2.4. □

5 A Markovian Case Study

In this section, we study a class of Markovian models where our basic Assumption 1 can indeed be verified. To this end, we recall the notion of a "random system with complete connections".

Let (M_1, d) be a compact metric space and $(M_2, \mathcal{M}_2)$ be a measurable space. Let Z denote a stochastic kernel from M_1 to $\mathcal{M}_2$ and let $v : M_1 \times M_2 \to M_1$ be a measurable mapping.

Definition 5.1 *Following Iosefescu and Theodorescu (1968), we call the quadruple*

$$\Sigma := ((M_1, d), (M_2, \mathcal{M}_2), Z, v) \tag{32}$$

a homogeneous random system with complete connections *(RSCC). Given an initial value $\xi \in M_1$, a RSCC induces two stochastic processes $\{\xi_t\}_{t \in \mathbb{N}}$ and $\{\zeta_t\}_{t \in \mathbb{N}}$ on the canonical probability space $(\Omega, \mathcal{F}, \mathbb{P}_\xi)$ taking values in M_1 and in M_2, respectively, by*

$$\xi_{t+1} = v(\xi_t, \zeta_t) \quad \textit{and} \quad \mathbb{P}_\xi(\zeta_t \in \cdot | \xi_t, \zeta_{t-1}, \xi_{t-1}, \zeta_{t-2}, \ldots) = Z(\xi_t; \cdot) \qquad (t \in \mathbb{N}).$$

Here, $\xi_0 = \xi$ $\mathbb{P}_\xi$-a.s. These processes are called the ***associated Markov chain*** *and the* ***signal sequence****, respectively.*

We assume that the following conditions are satisfied.

Assumption 2 *(i) For any starting point $\xi \in M_1$, the Markov chain $\{\xi_t\}_{t\in\mathbb{N}}$ converges in distribution to a unique stationary probability measure μ^* on M_1.*

(ii) There exists a constant $C < \infty$ such that

$$\sup_{B\in\mathcal{M}_2} \sup_{\xi\neq\hat{\xi}} \frac{|Z(\xi;B) - Z(\hat{\xi};B)|}{d(\xi,\hat{\xi})} \leq C.$$

(iii) The mapping $v : M_1 \times M_2 \to M_1$ satisfies the contraction condition

$$d(v(\xi,\zeta), v(\hat{\xi},\zeta)) \leq \theta d(\xi,\hat{\xi}) \qquad (\theta < 1).$$

Uniqueness of the invariant measure implies that the Markov chain $\{\xi_t\}_{t\in\mathbb{N}}$ is stationary and ergodic under the law

$$\mathbb{P}^*(\cdot) := \int_{M_1} \mathbb{P}_\xi(\cdot)\mu^*(d\xi). \tag{33}$$

However, as we assume that the sequence $\{\text{Law}(\xi_t, \mathbb{P}_\xi)\}_{t\in\mathbb{N}}$ converges weakly but not necessarily in the total variation norm to μ^*, the process $\{\xi_t\}_{t\in\mathbb{N}}$ typically does not have a “nice” asymptotic behaviour.

Let us first consider a driving sequence ψ which is generated by the signal process $\{\zeta_t\}_{t\in\mathbb{N}}$. More precisely, we assume that ψ is of the form

$$\psi = \{(f(\zeta_t), g(\zeta_t))\}_{t\in\mathbb{N}},$$

where $f, g : M_2 \to \mathbb{R}$ are measurable functions satisfying

$$-\infty \leq \int_{M_1}\int_{M_2} \ln|f(\zeta)| Z(\xi; d\zeta)\mu^*(d\xi) < 0 \quad \text{and} \quad \int_{M_1}\int_{M_2} (\ln|g(\zeta)|)^+ Z(\xi; d\zeta)\mu^*(d\xi) < \infty.$$

The environment ψ is stationary and ergodic under the measure $\mathbb{P}^*$ defined in (33) and satisfies (6); see Theorems 2.1.57 and 2.2.10 in Iosefescu and Theodorescu (1968). Thus ψ is “nice”, and so Theorem 2.4 yields the following result.

Proposition 5.2 *Under our above assumptions, the environment ψ defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P}_\xi)$ is “nice”. Thus, the stochastic sequence $\{Y_t\}_{t\in\mathbb{N}}$ defined by the linear recursive relation*

$$Y_{t+1} = f(\zeta_t)Y_t + g(\zeta_t)$$

converges in distribution to a unique stationary process in the sense specified in Theorem 2.4.

We are now going to consider a driving sequence which is defined in terms of the Markov chain $\{\xi_t\}_{t\in\mathbb{N}}$ instead of the signal process $\{\zeta_t\}_{t\in\mathbb{N}}$. More precisely, we consider an environment of the form

$$\psi = \{(f(\xi_t), g(\xi_t))\}_{t\in\mathbb{N}},$$

where $f, g : M_1 \to \mathbb{R}$ are Lipschitz continuous functions which are bounded away from zero. The environment ψ is stationary and ergodic under the law $\mathbb{P}^*$ defined by (33) but ψ is not necessarily "nice" with respect to the measure $\mathbb{P} = \mathbb{P}_\xi$. However, the driving sequence can be approximated in law by a sequence of "nice" processes $\{\psi^n\}_{n\in\mathbb{N}}$, and the approximating sequence $\{\psi^n\}_{n\in\mathbb{N}}$, $\psi^n = \{(A_t^n, B_t^n)\}_{t\in\mathbb{N}}$, may be constructed as follows:

Let $\{\epsilon_t\}_{t\in\mathbb{N}}$ and $\{\eta_t\}_{t\in\mathbb{N}}$ be sequences of non-negative random variables defined on $(\Omega, \mathcal{F})$. For each $\xi \in M_1$, these random variables are independent under the law $\mathbb{P}_\xi$, they are $\mathbb{P}_\xi$-independent of all ξ_t $(t \in \mathbb{N})$ and satisfy $\text{Law}((\epsilon_0, \eta_0), \mathbb{P}_\xi) = \text{Law}((\epsilon_0, \eta_0), \mathbb{P}_{\hat{\xi}})$ for all $\xi, \hat{\xi} \in M_1$. The probability measure $\text{Law}((\epsilon_0, \eta_0), \mathbb{P}_\xi)$ is assumed to have a bounded Lipschitz continuous density with respect to the 2-dimensional Lebesgue measure on $\mathbb{R}^2$.

Proposition 5.3 *Suppose that our Assumption 2 and the integrability conditions*

$$-\infty < \int_{M_1} \ln|f| d\mu^* < 0 \quad \textit{and} \quad \int_{M_1} (\ln|g|)^+ d\mu^* < \infty$$

are satisfied. Then the finite dimensional distributions of the process $\{Y_t\}_{t\in\mathbb{N}}$ *defined recursively by the linear relation*

$$Y_{t+1} = f(\xi_t) Y_t + g(\xi_t)$$

converge weakly to the finite dimensional distributions of a uniquely determined stationary process in the sense specified in Theorem 2.6.

Proof: The sequences $\widetilde{\psi}, \psi^1, \psi^2, \ldots$ defined on $(\Omega, \mathcal{F}, \mathbb{P}_\xi)$ are stationary and ergodic under $\mathbb{P}^*$. They satisfy (6), due to Proposition 4.30 in Horst (2000). Since the mapping $f, g : M_1 \to \mathbb{R}$ are bounded away from zero, it follows from a monotone convergence argument that ψ^n $(n \in \mathbb{N})$ is "nice" and that the regularity conditions (15) and (16) are satisfied. Thus, our Assumption 1 is met, and so the assertion follows from Theorem 2.6. □

References

Borovkov, A. A. (1998): *Ergodicity and stability of stochastic processes.* John Wiley & Sons, New York.

Brandt, A. (1986): "The Stochastic Equation $Y_{n+1} = A_n Y_n + B_n$ with Stationary Coefficients," *Adv. Appl. Probab.*, 18, 211–220.

Brandt, A., L. Franken, and B. Lisek (1990): *Stationary Stochastic Models.* John Wiley & Sons, New York.

Breiman, L. (1968): *Probability.* Addison-Wesley, Reading.

Ethier, S. N., and T. G. Kurtz (1986): *Markov Processes Characterisation and Convergence.* John Wiley & Sons, New York.

Föllmer, H. (1979): "Tail Structure of Markov Chains on Infinite Product Spaces," *Z. Wahrscheinlichkeitstheorie Verw. Geb.*, 50, 273–285.

Horst, U. (2000): "Asymptotics of Locally and Globally Interacting Markoc Processes Arising in Microstructure Models of Financial Markets," Dissertation, Humboldt Universität zu Berlin.

IOSEFESCU, M., AND M. THEODORESCU (1968): *Random Processes and Learning.* Springer-Verlag, Berlin.

VERVAAT, W. (1979): "On a stochastic difference equation and a representation of non-negative infinitely divisible random variables," *Adv. Appl. Probab.*, 11, 750–783.