# Asymptotic behavior of CLS estimators
# for unstable INAR(2) models

Mátyás Barczy*,◇, Márton Ispány*, Gyula Pap⋆

* Faculty of Informatics, University of Debrecen, Pf. 12, H–4010 Debrecen, Hungary.

⋆ Bolyai Institute, University of Szeged, Aradi vértanúk tere 1, H–6720 Szeged, Hungary.

e–mails: barczy.matyas@inf.unideb.hu (M. Barczy), ispany.marton@inf.unideb.hu (M. Ispány), papgy@math.u-szeged.hu (G. Pap).

◇ Corresponding author.

1 2 3

## Abstract

In this paper the asymptotic behavior of the conditional least squares estimators of the autoregressive parameters $(\alpha, \beta)$, of the stability parameter $\varrho := \alpha + \beta$, and of the mean $\mu$ of the innovation $\varepsilon_k$, $k \in \mathbb{N}$, for an unstable integer-valued autoregressive process $X_k = \alpha \circ X_{k-1} + \beta \circ X_{k-2} + \varepsilon_k$, $k \in \mathbb{N}$, is described. The limit distributions and the scaling factors are different according to the following three cases: (i) decomposable, (ii) indecomposable but not positively regular, and (iii) positively regular models.

# 1 Introduction

The theory and practice of statistical inference for integer-valued time series models are rapidly developing and important topics of the modern theory of statistics. A number of results are now available in specialized monographs and review papers, to name a few, see, e.g., Steutel and van Harn [33] and Weiß [36]. Among the most successful integer-valued time series models proposed in the literature we mention the INteger-valued AutoRegressive model of order $p$ (INAR($p$)). This model was first introduced by McKenzie [27] and Al-Osh and Alzaid [1] for the case $p = 1$. The INAR(1) model has been investigated by several authors. The more general INAR($p$) processes were first introduced by Alzaid and Al-Osh [2]. In their setup the autocorrelation structure of the process corresponds to that of an ARMA($p, p - 1$) process. Another definition of an INAR($p$) process was proposed independently by Du and Li [10] and

by Gauthier and Latour [13] and Latour [26], and is different from that of Alzaid and Al-Osh [2]. In Du and Li's setup the autocorrelation structure of an INAR($p$) process is the same as that of an AR($p$) process. The setup of Du and Li [10] has been followed by most of the authors, and our approach will also be the same. In Barczy et al. [5] we investigated the asymptotic behavior of unstable INAR($p$) processes, i.e., when the characteristic polynomial has a unit root. Under some natural assumptions we proved that the sequence of appropriately scaled random step functions formed from an unstable INAR($p$) process converges weakly towards a squared Bessel process. This limit process is a continuous time branching process with immigration also known as the square-root process or the Cox–Ingersoll–Ross process.

Parameter estimation for INAR($p$) models has a long history. Franke and Seligmann [12] analyzed conditional maximum likelihood estimator of some parameters (including the autoregressive parameter) for stable INAR(1) models with Poisson innovations. Du and Li [10, Theorem 4.2] proved asymptotic normality of the conditional least squares (CLS) estimator of the autoregressive parameters for stable INAR($p$) models (see also Latour [26, Proposition 6.1]), Brännäs and Hellström [7] considered generalized method of moment estimation. Silva and Oliveira [31] proposed a frequency domain based estimator of the autoregressive parameters for stable INAR($p$) models with Poisson innovations. Ispány et al. [17], [18] derived asymptotic inference for nearly unstable INAR(1) models which has been refined by Drost et al. [9] later. In [17] the mean of the innovation was supposed to be known, while in [18] both the autoregressive parameter and the mean of the innovation have been estimated jointly. Drost et al. [8] studied asymptotically efficient estimation of the parameters for stable INAR($p$) models. The stability parameter $\varrho := \alpha_1 + \cdots + \alpha_p$ of an INAR($p$) model with autoregressive parameters $(\alpha_1, \ldots, \alpha_p)$ has not been treated yet, but this stability parameter is well investigated in case of unstable AR($p$) processes, see the unit root tests, e.g., in Hamilton [15, Section 17, Table 17.3, Case 1]. Namely, for the simplicity in case of $p = 1$, if $(Y_k)_{k \geqslant 0}$ is an AR(1) process, i.e., $Y_k = \varrho Y_{k-1} + \zeta_k$, $k \geqslant 1$, with $Y_0 := 0$ and an i.i.d. sequence $(\zeta_k)_{k \geqslant 1}$ having mean $0$ and positive variance, then the ordinary least squares estimator of the stability parameter $\varrho$ based on the sample $\boldsymbol{Y}_n := (Y_1, \ldots, Y_n)$ takes the form

$$\widehat{\varrho}_n(\boldsymbol{Y}_n) = \frac{\sum_{k=1}^n Y_{k-1} Y_k}{\sum_{k=1}^n Y_k^2}, \qquad n \geqslant 1,$$

see, e.g., Hamilton [15, 17.4.2], and, by Hamilton [15, 17.4.7], in the unstable case, i.e., when $\varrho = 1$,

$$n(\widehat{\varrho}_n(\boldsymbol{Y}_n) - 1) \xrightarrow{\mathcal{L}} \frac{\int_0^1 \mathcal{W}_t \, d\mathcal{W}_t}{\int_0^1 \mathcal{W}_t^2 \, dt} \qquad \text{as } n \to \infty,$$

where $(\mathcal{W}_t)_{t \geqslant 0}$ is a standard Wiener process and $\xrightarrow{\mathcal{L}}$ denotes convergence in distribution. Here $n(\widehat{\varrho}_n(\boldsymbol{Y}_n) - 1)$ is known as the Dickey–Fuller statistics. In this paper the asymptotic behavior of the CLS estimators of the autoregressive and the stability parameters and of the mean of the innovation for unstable INAR(2) models is described (see our main results in Section 2) which can be considered as a first step of examining this question for general unstable INAR(p) processes and more generally for critical multitype branching processes. We call the attention

2

that in case of unstable INAR(2) processes new types of limit distribution occur (see Theorem 2.1) compared to those of unstable AR($p$) processes.

First we recall INAR(2) models. Let $\mathbb{Z}_+$, $\mathbb{N}$, $\mathbb{R}$ and $\mathbb{R}_+$ denote the set of non-negative integers, positive integers, real numbers and non-negative real numbers, respectively. Every random variable will be defined on a fixed probability space $(\Omega, \mathcal{A}, \mathbb{P})$.

**1.1 Definition.** *Let $(\varepsilon_k)_{k \in \mathbb{N}}$ be an independent and identically distributed (i.i.d.) sequence of non-negative integer-valued random variables, and let $(\alpha, \beta) \in [0,1]^2$. An INAR(2) time series model with autoregressive parameters $(\alpha, \beta)$ and innovations $(\varepsilon_k)_{k \in \mathbb{N}}$ is a stochastic process $(X_k)_{k \geqslant -1}$ given by*

$$(1.1) \qquad X_k = \sum_{j=1}^{X_{k-1}} \xi_{k,j} + \sum_{j=1}^{X_{k-2}} \eta_{k,j} + \varepsilon_k, \qquad k \in \mathbb{N},$$

*where for all $k \in \mathbb{N}$, $(\xi_{k,j})_{j \in \mathbb{N}}$ and $(\eta_{k,j})_{j \in \mathbb{N}}$ are sequences of i.i.d. Bernoulli random variables with mean $\alpha$ and $\beta$, respectively, such that these sequences are mutually independent and independent of the sequence $(\varepsilon_k)_{k \in \mathbb{N}}$, and $X_0$ and $X_{-1}$ are non-negative integer-valued random variables independent of the sequences $(\xi_{k,j})_{j \in \mathbb{N}}$, $(\eta_{k,j})_{j \in \mathbb{N}}$, $k \in \mathbb{N}$, and $(\varepsilon_k)_{k \in \mathbb{N}}$.*

The INAR(2) model (1.1) can be written in another way using the binomial thinning operator $\circ$ (due to Steutel and van Harn [33]) which we recall now. Let $X$ be a non-negative integer-valued random variable. Let $(\xi_j)_{j \in \mathbb{N}}$ be a sequence of i.i.d. Bernoulli random variables with mean $\alpha \in [0,1]$. We assume that the sequence $(\xi_j)_{j \in \mathbb{N}}$ is independent of $X$. The non-negative integer-valued random variable $\alpha \circ X$ is defined by

$$\alpha \circ X := \begin{cases} \sum_{j=1}^{X} \xi_j, & \text{if } X > 0, \\ 0, & \text{if } X = 0. \end{cases}$$

The sequence $(\xi_j)_{j \in \mathbb{N}}$ is called a counting sequence. Then the INAR(2) model (1.1) takes the form

$$X_k = \alpha \circ X_{k-1} + \beta \circ X_{k-2} + \varepsilon_k, \qquad k \in \mathbb{N}.$$

Note that the above form of the INAR(2) model is quite analogous with a usual AR(2) process (another slight link between them is the similarity of some conditional expectations, see (3.1)).

For the sake of simplicity we consider a zero start INAR(2) process, that is we suppose $X_0 = X_{-1} = 0$. The general case of nonzero initial values may be handled in a similar way, but we renounce to consider it.

In the sequel we always assume $\mathbb{E}(\varepsilon_1^2) < \infty$. Let us denote the mean and variance of $\varepsilon_1$ by $\mu$ and $\sigma^2$. Further, we assume $\mu > 0$, otherwise $X_k = 0$ for all $k \in \mathbb{N}$.

Based on the asymptotic behavior of $\mathbb{E}(X_k)$ as $k \to \infty$ described in Barczy et al. [5, Proposition 2.6], we distinguish three types of INAR(2) models. The asymptotic behavior of

$\mathbb{E}(X_k)$ as $k \to \infty$ is determined by the spectral radius $r$ of the matrix

$$(1.2) \qquad A := \begin{bmatrix} \alpha & \beta \\ 1 & 0 \end{bmatrix},$$

i.e., by the maximum of the modulus of the eigenvalues of $A$. The case $r < 1$, when $\mathbb{E}(X_k)$ converges to a finite limit as $k \to \infty$, is called *stable* or *asymptotically stationary*, whereas the cases $r = 1$, when $\mathbb{E}(X_k)$ tends linearly to $\infty$, and $r > 1$, when $\mathbb{E}(X_k)$ converges to $\infty$ with an exponential rate, are called *unstable* and *explosive*, respectively. It is easy to check that $r < 1$, $r = 1$, and $r > 1$ are equivalent with $\varrho < 1$, $\varrho = 1$, and $\varrho > 1$, respectively, where $\varrho := \alpha + \beta$ is called the *stability parameter*, see Barczy et al. [5, Proposition 2.2].

We also note that an INAR(2) process can be considered as a special 2-type branching process with immigration. Namely, by (1.1),

$$\begin{bmatrix} X_k \\ X_{k-1} \end{bmatrix} = \sum_{j=1}^{X_{k-1}} \begin{bmatrix} \xi_{k,j} \\ 1 \end{bmatrix} + \sum_{j=1}^{X_{k-2}} \begin{bmatrix} \eta_{k,j} \\ 0 \end{bmatrix} + \begin{bmatrix} \varepsilon_k \\ 0 \end{bmatrix}, \qquad k \in \mathbb{N},$$

and hence the so-called mean matrix of an INAR(2) process with autoregressive parameters $(\alpha, \beta)$ (considered as a 2-type branching process) is nothing else but $A$. This process is called *positively regular* if there is a positive integer $k \in \mathbb{N}$ such that the entries of $A^k$ are positive (see Kesten and Stigum [23]), which is equivalent with $\alpha > 0$ and $\beta > 0$. The model is called *decomposable* if the matrix $A$ is decomposable (see Kesten and Stigum [25]), which is equivalent with $\beta = 0$. If $\alpha = 0$ and $\beta > 0$, then the process is *indecomposable but not positively regular* (see Kesten and Stigum [24]). If $\alpha > 0$ and $\beta = 0$, then the decomposable process $(X_k)_{k \geqslant -1}$ is an INAR(1) process with autoregressive parameter $\alpha$. If $\alpha = 0$ and $\beta > 0$, then the indecomposable process $(X_k)_{k \geqslant -1}$ takes the form

$$X_k = \beta \circ X_{k-2} + \varepsilon_k, \qquad k \in \mathbb{N},$$

and hence the subsequences $(X_{2k-j})_{k \geqslant 0}$, $j \in \{0, 1\}$, form independent positively regular INAR(1) processes with autoregressive parameter $\beta$ such that $X_{-j} = 0$, $j \in \{0, 1\}$. For more details of this classification of INAR(2) processes, see Appendix A.

Next we give an overview of the structure of the paper. Section 2 contains our main results, see Theorem 2.1 for unstable and positively regular INAR(2) processes, Theorem 2.2 for unstable and decomposable INAR(2) processes, and Theorem 2.3 for unstable, indecomposable but not positively regular ones. In order to highlight our main results, the preliminaires and (technical) details on CLS estimators are presented only after our main results, see Section 3. In Theorems 4.1, 4.2 and 4.3 of Section 4 we present joint asymptotic behaviours of the building blocks of the CLS estimators (according to the above mentioned three cases), and by applying a version of the continuous mapping theorem (which is formulated for completeness in Appendix B) we show how one can derive Theorems 2.1, 2.2 and 2.3 using these theorems. Section 5 is devoted to the proof of Theorem 4.1 which is based on Lemma 5.1 and Theorem 5.1. Due to its length, the proof of Theorem 5.1 is given separately in Section 6. Sections 7 and 8 are

devoted to the proofs of Theorem 4.2 and Theorem 4.3, respectively. In Section 9 we present estimates for the moments of the processes involved, these estimates are used throughout the paper. In Appendix C we recall a result about convergence of random step processes noting that the proof of Theorem 5.1 is based on this result.

# 2   Main results

In what follows we always assume $\varrho = \alpha + \beta = 1$, that is, the process $(X_k)_{k \geqslant -1}$ is unstable.

For each $n \in \mathbb{N}$, any CLS estimator $(\widehat{\alpha}_n(\boldsymbol{X}_n), \widehat{\beta}_n(\boldsymbol{X}_n), \widehat{\mu}_n(\boldsymbol{X}_n))$ of the autoregressive parameters $(\alpha, \beta)$ and of the mean $\mu$ of the innovation based on a sample $\boldsymbol{X}_n := (X_1, \ldots, X_n)$ has the form

$$
\begin{bmatrix} \widehat{\alpha}_n(\boldsymbol{X}_n) \\ \widehat{\beta}_n(\boldsymbol{X}_n) \\ \widehat{\mu}_n(\boldsymbol{X}_n) \end{bmatrix} = \left( \sum_{k=1}^{n} \begin{bmatrix} X_{k-1}^2 & X_{k-1}X_{k-2} & X_{k-1} \\ X_{k-1}X_{k-2} & X_{k-2}^2 & X_{k-2} \\ X_{k-1} & X_{k-2} & 1 \end{bmatrix} \right)^{-1} \sum_{k=1}^{n} \begin{bmatrix} X_k X_{k-1} \\ X_k X_{k-2} \\ X_k \end{bmatrix}
$$

on the set $\{\omega \in \Omega : \sum_{k=1}^{n} X_{k-2}(\omega)^2 > 0\}$ with $\lim_{n \to \infty} \mathbb{P}\left(\sum_{k=1}^{n} X_{k-2}^2 > 0\right) = 1$, see Proposition 3.1. Moreover, for each $n \in \mathbb{N}$, any CLS estimator of the stability parameter $\varrho$ takes the form

$$
\widehat{\varrho}_n(\boldsymbol{X}_n) = \widehat{\alpha}_n(\boldsymbol{X}_n) + \widehat{\beta}_n(\boldsymbol{X}_n)
$$

on the set $\{\omega \in \Omega : \sum_{k=1}^{n} X_{k-2}(\omega)^2 > 0\}$, see Section 3.

**2.1 Theorem.** *Let* $(X_k)_{k \geqslant -1}$ *be an* INAR(2) *process with autoregressive parameters* $(\alpha, \beta) \in (0,1)^2$ *such that* $\alpha + \beta = 1$ *(hence it is unstable and positively regular). Suppose that* $X_0 = X_{-1} = 0$, $\mathbb{E}(\varepsilon_1^8) < \infty$ *and* $\mu > 0$. *Then*

$$(2.1) \qquad n\big(\widehat{\varrho}_n(\boldsymbol{X}_n) - 1\big) \xrightarrow{\mathcal{L}} \frac{\sqrt{2\alpha\beta} \int_0^1 \mathcal{X}_t^{3/2}\, d\mathcal{W}_t - [(1+\beta)\mathcal{X}_1 - \mu] \int_0^1 \mathcal{X}_t\, dt}{\int_0^1 \mathcal{X}_t^2\, dt - \left(\int_0^1 \mathcal{X}_t\, dt\right)^2}$$

$$(2.2) \qquad \begin{bmatrix} n^{1/2}(\widehat{\alpha}_n(\boldsymbol{X}_n) - \alpha) \\ n^{1/2}(\widehat{\beta}_n(\boldsymbol{X}_n) - \beta) \end{bmatrix} \xrightarrow{\mathcal{L}} \sqrt{\alpha(1+\beta)} \frac{\int_0^1 \mathcal{X}_t\, d\widetilde{\mathcal{W}}_t}{\int_0^1 \mathcal{X}_t\, dt} \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

*and*

$$(2.3) \qquad \widehat{\mu}_n(\boldsymbol{X}_n) - \mu \xrightarrow{\mathcal{L}} \frac{-\sqrt{2\alpha\beta} \int_0^1 \mathcal{X}_t\, dt \int_0^1 \mathcal{X}_t^{3/2}\, d\mathcal{W}_t + [(1+\beta)\mathcal{X}_1 - \mu] \int_0^1 \mathcal{X}_t^2\, dt}{\int_0^1 \mathcal{X}_t^2\, dt - \left(\int_0^1 \mathcal{X}_t\, dt\right)^2}$$

*as* $n \to \infty$, *where* $(\mathcal{X}_t)_{t \in \mathbb{R}_+}$ *is the unique strong solution of the stochastic differential equation (SDE)*

$$(2.4) \qquad d\mathcal{X}_t = \frac{1}{1+\beta}\Big(\mu\, dt + \sqrt{2\alpha\beta\mathcal{X}_t^+}\, d\mathcal{W}_t\Big), \qquad t \in \mathbb{R}_+,$$

*with initial value* $\mathcal{X}_0 = 0$, *where* $(\mathcal{W}_t)_{t \in \mathbb{R}_+}$, $(\widetilde{\mathcal{W}}_t)_{t \in \mathbb{R}_+}$ *are independent standard Wiener processes, and* $x^+$ *denotes the positive part of* $x \in \mathbb{R}$.

5

**2.1 Remark.** The moment condition $\mathbb{E}(\varepsilon_1^8) < \infty$ in Theorem 2.1 seems to be too strong, but we call the attention that the process $(X_k)_{k \geqslant -1}$ can be considered as a heteroscedastic time series. Indeed, $X_k = \alpha X_{k-1} + \beta X_{k-2} + M_k + \mu$, see (3.3), and by (9.1), $\mathbb{E}(M_k^2 \,|\, \mathcal{F}_{k-1}) = \alpha(1-\alpha)X_{k-1} + \beta(1-\beta)X_{k-2} + \sigma^2$, $k \in \mathbb{N}$. That is why we think that the behavior of the process $(X_k)_{k \geqslant -1}$ is similar to GARCH models, where, even in the stable case, high moment conditions are needed for convergence of estimators such as the quasi-maximum likelihood estimator in Hall and Yao [14] or the Whittle estimator in Mikosch and Straumann [28]. □

**2.2 Remark.** The SDE (2.4) has a unique strong solution $(\mathcal{X}_t^{(x)})_{t \geqslant 0}$ for all initial values $\mathcal{X}_0^{(x)} = x \in \mathbb{R}$. Indeed, since $|\sqrt{x} - \sqrt{y}| \leqslant \sqrt{|x-y|}$ for $x, y \geqslant 0$, the coefficient functions $\mathbb{R} \ni x \mapsto \mu/(1+\beta)$ and $\mathbb{R} \ni x \mapsto \sqrt{2\alpha\beta x^+}/(1+\beta)$ satisfy conditions of part (ii) of Theorem 3.5 in Chapter IX in Revuz and Yor [30] or the conditions of Proposition 5.2.13 in Karatzas and Shreve [22]. Further, by the comparison theorem (see, e.g., Revuz and Yor [30, Theorem 3.7, Chapter IX]), if the initial value $\mathcal{X}_0^{(x)} = x$ is nonnegative, then $\mathcal{X}_t^{(x)}$ is nonnegative for all $t \in \mathbb{R}_+$ with probability one. Hence $\mathcal{X}_t^+$ may be replaced by $\mathcal{X}_t$ under the square root in (2.4). The unique strong solution of the SDE (2.4) is known as a squared Bessel process, a squared-root process or a Cox–Ingersoll–Ross (CIR) process. □

**2.3 Remark.** By Itô's formula and Remark 2.2, $\mathcal{M}_t := (1+\beta)\mathcal{X}_t - \mu t$, $t \in \mathbb{R}_+$, is the unique strong solution of the SDE

$$(2.5) \qquad \mathrm{d}\mathcal{M}_t = \sqrt{\frac{2\alpha\beta}{1+\beta}(\mathcal{M}_t + \mu t)^+} \, \mathrm{d}\mathcal{W}_t, \qquad t \in \mathbb{R}_+,$$

with initial value $\mathcal{M}_0 = 0$, and $(\mathcal{M}_t + \mu t)^+$ may be replaced by $\mathcal{M}_t + \mu t$ under the square root in (2.5). Hence $\mathrm{d}\mathcal{M}_t = \sqrt{2\alpha\beta\mathcal{X}_t} \, \mathrm{d}\mathcal{W}_t$, and the convergences (2.1) and (2.3) can also be formulated as

$$(2.6) \qquad n(\widehat{\varrho}_n(\boldsymbol{X}_n) - 1) \xrightarrow{\mathcal{L}} \frac{\int_0^1 \mathcal{X}_t \, \mathrm{d}\mathcal{M}_t - \mathcal{M}_1 \int_0^1 \mathcal{X}_t \, \mathrm{d}t}{\int_0^1 \mathcal{X}_t^2 \, \mathrm{d}t - \left(\int_0^1 \mathcal{X}_t \, \mathrm{d}t\right)^2} \qquad \text{as } n \to \infty,$$

$$(2.7) \qquad \widehat{\mu}_n(\boldsymbol{X}_n) - \mu \xrightarrow{\mathcal{L}} \frac{-\int_0^1 \mathcal{X}_t \, \mathrm{d}t \int_0^1 \mathcal{X}_t \, \mathrm{d}\mathcal{M}_t + \mathcal{M}_1 \int_0^1 \mathcal{X}_t^2 \, \mathrm{d}t}{\int_0^1 \mathcal{X}_t^2 \, \mathrm{d}t - \left(\int_0^1 \mathcal{X}_t \, \mathrm{d}t\right)^2} \qquad \text{as } n \to \infty.$$
□

The next theorem contains our result for decomposable unstable INAR(2) processes.

**2.2 Theorem.** *Let* $(X_k)_{k \geqslant -1}$ *be an* INAR(2) *process with autoregressive parameters* $(1, 0)$ *(hence it is unstable and decomposable). Suppose that* $X_0 = X_{-1} = 0$, $\mathbb{E}(\varepsilon_1^4) < \infty$ *and* $\mu > 0$. *Then*

$$(2.8) \qquad n^{3/2}\left(\widehat{\varrho}_n(\boldsymbol{X}_n) - 1\right) \xrightarrow{\mathcal{L}} \mathcal{N}_1\left(0, \frac{12\sigma^2}{\mu^2}\right) \qquad \text{as } n \to \infty,$$

$$(2.9) \qquad \begin{bmatrix} n^{1/2}(\widehat{\alpha}_n(\boldsymbol{X}_n) - 1) \\ n^{1/2}\widehat{\beta}_n(\boldsymbol{X}_n) \end{bmatrix} \xrightarrow{\mathcal{L}} Z \begin{bmatrix} -1 \\ 1 \end{bmatrix} \qquad as \ \ n \to \infty,$$

*and*

$$(2.10) \qquad n^{1/2}\big(\widehat{\mu}_n(\boldsymbol{X}_n) - \mu\big) \xrightarrow{\mathcal{L}} \mathcal{N}_1(0, \mu^2 + 4\sigma^2) \qquad as \ \ n \to \infty,$$

*where $Z$ is a standard normally distributed random variable.*

**2.4 Remark.** Note that an unstable and decomposable INAR(2) process has autoregressive parameters $(1, 0)$, i.e., it is actually an unstable INAR(1) process. However, we call the attention that the asymptotic behaviour of the estimators $\widehat{\varrho}_n(\boldsymbol{X}_n)$, $(\widehat{\alpha}_n(\boldsymbol{X}_n), \widehat{\beta}_n(\boldsymbol{X}_n))$ and $\widehat{\mu}_n(\boldsymbol{X}_n)$ as $n \to \infty$ in Theorem 2.2 can not be derived from the corresponding results for an unstable INAR(1) process, since the CLS estimator of the coefficient (which can also be considered as the stability parameter) of an INAR(1) process is different from $\widehat{\varrho}_n(\boldsymbol{X}_n)$, see, e.g., Ispány et al. [18]. Remark also that the CLS estimator of the coefficient of an unstable INAR(1) process is also asymptotically normal with the same scaling $n^{3/2}$, but the asymptotic variance $3\sigma^2/\mu^2$ is different from the corresponding one $12\sigma^2/\mu^2$ for an unstable and decomposable INAR(2) process, see Ispány et al. [18, Theorem 2.1]. $\qquad\qquad \square$

The last theorem contains our result for unstable, indecomposable but not positively regular INAR(2) processes.

**2.3 Theorem.** *Let $(X_k)_{k \geqslant -1}$ be an INAR(2) process with autoregressive parameters $(0, 1)$ (hence it is unstable, indecomposable but not positively regular). Suppose that $X_0 = X_{-1} = 0$, $\mathbb{E}(\varepsilon_1^2) < \infty$ and $\mu > 0$. Then*

$$(2.11) \qquad n^{3/2}\big(\widehat{\varrho}_n(\boldsymbol{X}_n) - 1\big) \xrightarrow{\mathcal{L}} \mathcal{N}_1\left(0, \frac{48\sigma^2}{\mu^2}\right) \qquad as \ \ n \to \infty,$$

$$(2.12) \qquad \begin{bmatrix} n\widehat{\alpha}_n(\boldsymbol{X}_n) \\ n(\widehat{\beta}_n(\boldsymbol{X}_n) - 1) \end{bmatrix} \xrightarrow{\mathcal{L}} \frac{\int_0^1 \mathcal{W}_t \, d\mathcal{W}_t}{\int_0^1 \mathcal{W}_t^2 \, dt} \begin{bmatrix} -1 \\ 1 \end{bmatrix} \qquad as \ \ n \to \infty,$$

*and*

$$(2.13) \qquad n^{1/2}\big(\widehat{\mu}_n(\boldsymbol{X}_n) - \mu\big) \xrightarrow{\mathcal{L}} \mathcal{N}_1\left(0, 4\sigma^2\right) \qquad as \ \ n \to \infty,$$

*where $(\mathcal{W}_t)_{t \in \mathbb{R}_+}$ is a standard Wiener process.*

**2.5 Remark.** We note that in all unstable cases the limit distributions for the estimators of the autoregressive parameters are concentrated on the same line $\{(x, y) \in \mathbb{R}^2 : x + y = 0\}$. However, these limit distributions are pairwise different. Surprisingly, both in the unstable positively regular case and in the unstable decomposable case the scaling factor is $\sqrt{n}$, while in the unstable, indecomposable but not positively regular case it is $n$. In the stable case

7

this factor is again $\sqrt{n}$ (see Du and Li [10, Theorem 4.2] or Latour [26, Proposition 6.1]). The reason of this strange phenomena can be understood from the asymptotic behavior of the sequence $(\boldsymbol{A}_n, \boldsymbol{d}_n)_{n \in \mathbb{N}}$ of random vectors defined and analyzed in Sections 3, 4, 5, 7 and 8. Namely, the scaling factor for the entries of the matrices $(\boldsymbol{A}_n)_{n \in \mathbb{N}}$ as well as for the entries of the vectors $(\boldsymbol{d}_n)_{n \in \mathbb{N}}$ are different. In order to get over these difficulties, we use the canonical form of the process $(X_k)_{k \in \mathbb{N}}$ due to Sims, Stock and Watson [32]. Further, one of the decisive tools in deriving the needed asymptotic behavior is a good bound for the moments of the involved processes, see Corollary 9.1. $\qquad\square$

**2.6 Remark.** We recall that the distribution of $\int_0^1 \mathcal{W}_t \, \mathrm{d}\mathcal{W}_t / \int_0^1 \mathcal{W}_t^2 \, \mathrm{d}t$ in Theorem 2.3 agrees with the limit distribution of the Dickey–Fuller statistics for unit root test of AR(1) time series, see, e.g., Hamilton [15, 17.4.2 and 17.4.7] or Tanaka [34, (7.14) and Theorem 9.5.1]. The limit distribution in (2.2) is also a fraction of two stochastic integrals, but it contains two independent standard Wiener processes. This phenomena is very similar to the appearing of two independent standard Wiener processes in limit theorems for CLS estimators of the variance of the offspring and immigration distributions for critical branching processes with immigration in Winnicki [37, Theorems 3.5 and 3.8]. Finally, note that the limit distribution of the CLS estimator of the autoregressive parameters $(\alpha, \beta)$ is symmetric in Theorems 2.1 and 2.2, and non-symmetric in Theorem 2.3. Indeed, since $(\mathcal{W}_t)_{t \in \mathbb{R}_+}$ and $(\widetilde{\mathcal{W}}_t)_{t \in \mathbb{R}_+}$ are independent, by the SDE (2.4), the processes $(\mathcal{X}_t)_{t \in \mathbb{R}_+}$ and $(\widetilde{\mathcal{W}}_t)_{t \in \mathbb{R}_+}$ are also independent, which yields that the limit distribution of the CLS estimator of $(\alpha, \beta)$ is symmetric in Theorem 2.1. $\qquad\square$

**2.7 Remark.** We note that the CLS estimator $\widehat{\varrho}_n(\boldsymbol{X}_n)$ of $\varrho$, and the CLS estimator $(\widehat{\alpha}_n(\boldsymbol{X}_n), \widehat{\beta}_n(\boldsymbol{X}_n))$ of $(\alpha, \beta)$ are asymptotically weakly consistent as $n \to \infty$ in Theorems 2.1, 2.2 and 2.3. The CLS estimator $\widehat{\mu}_n(\boldsymbol{X}_n)$ of $\mu$ in Theorems 2.2 and 2.3 is also asymptotically weakly consistent as $n \to \infty$, however in Theorem 2.1 it is not asymptotically weakly consistent. Note that in the case of an unstable INAR(1) model the CLS estimator of the mean of the innovation is asymptotically weakly consistent, see Ispány et al. [18]. Further, we remark that in Theorem 2.1 the variance $\sigma^2$ of the innovation does not show up in the limit distributions, while in Theorems 2.2 and 2.3 it appears. Finally, in Theorems 2.1, 2.2 and 2.3 one could prove joint convergence as well. $\qquad\square$

## 3 CLS estimators

For all $k \in \mathbb{Z}_+$, let us denote by $\mathcal{F}_k$ the $\sigma$-algebra generated by the random variables $X_{-1}, X_0, X_1, \ldots, X_k$. (Note that $\mathcal{F}_0 = \{\Omega, \emptyset\}$, since $X_0 = X_{-1} = 0$.) By (1.1),

$$(3.1) \qquad \mathbb{E}(X_k \,|\, \mathcal{F}_{k-1}) = \alpha X_{k-1} + \beta X_{k-2} + \mu, \qquad k \in \mathbb{N}.$$

Let us introduce the sequence

$$(3.2) \qquad M_k := X_k - \mathbb{E}(X_k \,|\, \mathcal{F}_{k-1}) = X_k - \alpha X_{k-1} - \beta X_{k-2} - \mu, \qquad k \in \mathbb{N},$$

of martingale differences with respect to the filtration $(\mathcal{F}_k)_{k\in\mathbb{Z}_+}$. The process $(X_k)_{k\geqslant-1}$ satisfies the recursion

$$(3.3) \qquad X_k = \alpha X_{k-1} + \beta X_{k-2} + M_k + \mu, \qquad k \in \mathbb{N}.$$

For each $n \in \mathbb{N}$, a CLS estimator $(\widehat{\alpha}_n(\boldsymbol{X}_n), \widehat{\beta}_n(\boldsymbol{X}_n), \widehat{\mu}_n(\boldsymbol{X}_n))$ of the parameters $(\alpha, \beta, \mu)$ based on a sample $\boldsymbol{X}_n = (X_1, \ldots, X_n)$ can be obtained by minimizing the sum of squares

$$(3.4) \qquad \sum_{k=1}^{n} \left( X_k - \mathbb{E}(X_k \,|\, \mathcal{F}_{k-1}) \right)^2 = \sum_{k=1}^{n} (X_k - \alpha X_{k-1} - \beta X_{k-2} - \mu)^2$$

with respect to $(\alpha, \beta, \mu)$ over $\mathbb{R}^3$. For all $n \in \mathbb{N}$ and $x_1, \ldots, x_n \in \mathbb{R}$, let us put

$$\boldsymbol{x}_n := (x_1, \ldots, x_n),$$

and in what follows we use the convention

$$x_{-1} := x_0 := 0.$$

For all $n \in \mathbb{N}$, we define the function $Q_n : \mathbb{R}^n \times \mathbb{R}^3 \to \mathbb{R}$ by

$$Q_n(\boldsymbol{x}_n; \alpha', \beta', \mu') := \sum_{k=1}^{n} (x_k - \alpha' x_{k-1} - \beta' x_{k-2} - \mu')^2$$

for all $\alpha', \beta', \mu' \in \mathbb{R}$ and $\boldsymbol{x}_n \in \mathbb{R}^n$. By definition, for all $n \in \mathbb{N}$, a CLS estimator of the parameters $(\alpha, \beta, \mu)$ is a measurable function $(\widehat{\alpha}_n, \widehat{\beta}_n, \widehat{\mu}_n) : \mathbb{R}^n \to \mathbb{R}^3$ such that

$$Q_n(\boldsymbol{x}_n; \widehat{\alpha}_n(\boldsymbol{x}_n), \widehat{\beta}_n(\boldsymbol{x}_n), \widehat{\mu}_n(\boldsymbol{x}_n)) = \inf_{(\alpha', \beta', \mu') \in \mathbb{R}^3} Q_n(\boldsymbol{x}_n; \alpha', \beta', \mu') \qquad \forall \, \boldsymbol{x}_n \in \mathbb{R}^n.$$

Since the variance $\sigma^2$ of the innovation does not appear in the conditional expectation $\mathbb{E}(X_k \,|\, \mathcal{F}_{k-1})$ given in (3.1), and hence, in the definition of $Q_n$, we do not need to know the value of $\sigma^2$ for the calculation of the CLS estimator of the parameters $(\alpha, \beta, \mu)$.

Next we give the solutions of this extremum problem.

**3.1 Lemma.** *For each $n \geqslant 2$, $n \in \mathbb{N}$, any CLS estimator of the parameters $(\alpha, \beta, \mu)$ is a measurable function $(\widehat{\alpha}_n, \widehat{\beta}_n, \widehat{\mu}_n) : \mathbb{R}^n \to \mathbb{R}^3$ for which*

$$(3.5) \qquad \begin{bmatrix} \widehat{\alpha}_n(\boldsymbol{x}_n) \\ \widehat{\beta}_n(\boldsymbol{x}_n) \\ \widehat{\mu}_n(\boldsymbol{x}_n) \end{bmatrix} = F_n(\boldsymbol{x}_n)^{-1} g_n(\boldsymbol{x}_n)$$

*if $\sum_{k=1}^{n} x_{k-2}^2 > 0$, where*

$$F_n(\boldsymbol{x}_n) := \sum_{k=1}^{n} \begin{bmatrix} x_{k-1} \\ x_{k-2} \\ 1 \end{bmatrix} \begin{bmatrix} x_{k-1} \\ x_{k-2} \\ 1 \end{bmatrix}^{\top}, \qquad g_n(\boldsymbol{x}_n) := \sum_{k=1}^{n} x_k \begin{bmatrix} x_{k-1} \\ x_{k-2} \\ 1 \end{bmatrix},$$

9

$$(3.6) \qquad \widehat{\alpha}_n(\boldsymbol{x}_n) = \frac{x_n}{x_{n-1}} - \frac{1}{n-1}, \qquad \widehat{\mu}_n(\boldsymbol{x}_n) = \frac{x_{n-1}}{n-1}$$

if $x_1 = \cdots = x_{n-2} = 0$ and $x_{n-1} \neq 0$, and

$$(3.7) \qquad \widehat{\mu}_n(\boldsymbol{x}_n) = \frac{x_n}{n}$$

if $x_1 = \cdots = x_{n-1} = 0$.

Note that $(\widehat{\alpha}_n, \widehat{\beta}_n, \widehat{\mu}_n)$ is not defined uniquely on the set $\{\boldsymbol{x}_n \in \mathbb{R}^n : x_1 = \cdots = x_{n-2} = 0\}$. Namely, if $x_1 = \cdots = x_{n-2} = 0$ and $x_{n-1} \neq 0$, then $\widehat{\beta}_n$ can be chosen as an arbitrary measurable function, while if $x_1 = \cdots = x_{n-1} = 0$, then the same holds for $(\widehat{\alpha}_n, \widehat{\beta}_n)$. We call the attention that Lemma 3.1 holds for all types of INAR(2) processes, i.e., it covers the stable, unstable and explosive cases as well.

**Proof of Lemma 3.1.** For any fixed $\boldsymbol{x}_n \in \mathbb{R}^n$ with $\sum_{k=1}^n x_{k-2}^2 > 0$, the quadratic function $\mathbb{R}^3 \ni (\alpha', \beta', \mu') \mapsto Q_n(\boldsymbol{x}_n; \alpha', \beta', \mu')$ can be written in the form

$$Q_n(\boldsymbol{x}_n; \alpha', \beta', \mu') = \left( \begin{bmatrix} \alpha' \\ \beta' \\ \mu' \end{bmatrix} - F_n(\boldsymbol{x}_n)^{-1} g_n(\boldsymbol{x}_n) \right)^{\top} F_n(\boldsymbol{x}_n) \left( \begin{bmatrix} \alpha' \\ \beta' \\ \mu' \end{bmatrix} - F_n(\boldsymbol{x}_n)^{-1} g_n(\boldsymbol{x}_n) \right) + \widetilde{Q}_n(\boldsymbol{x}_n),$$

where

$$\widetilde{Q}_n(\boldsymbol{x}_n) := \sum_{k=1}^n x_k^2 - g_n(\boldsymbol{x}_n)^{\top} F_n(\boldsymbol{x}_n)^{-1} g_n(\boldsymbol{x}_n).$$

We check that the matrix $F_n(\boldsymbol{x}_n)$ is strictly positive definite. For this, it is enough to show that $\sum_{k=1}^n x_{k-2}^2 > 0$ implies that the rank of the system of vectors

$$(3.8) \qquad \begin{bmatrix} x_{k-1} \\ x_{k-2} \\ 1 \end{bmatrix}, \qquad k \in \{1, \ldots, n\},$$

equals 3. Indeed, if $a_i \in \mathbb{R}^3$, $i \in \{1, \ldots, n\}$, and the rank of the system $\{a_1, \ldots, a_n\}$ is 3, then $A := \sum_{i=1}^n a_i a_i^{\top}$ is strictly positive definite which can be checked as follows. For any $z \in \mathbb{R}^3$,

$$\langle Az, z \rangle = \sum_{i=1}^n \langle a_i a_i^{\top} z, z \rangle = \sum_{i=1}^n \langle a_i^{\top} z, a_i^{\top} z \rangle \geqslant 0,$$

and $\langle Az, z \rangle = 0$ holds if and only if $a_i^{\top} z = 0$, $i = 1, \ldots, n$. Since the rank of the system $\{a_1, \ldots, a_n\}$ is 3, we have $z = 0$.

The rank of the system of vectors in (3.8) is 3, since the rank of the matrix

$$\begin{bmatrix} x_{n-1} & x_{n-2} & \cdots & x_2 & x_1 & 0 \\ x_{n-2} & x_{n-3} & \cdots & x_1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

10

equals 3. Indeed, there exists some $i \in \{1, \ldots, n-2\}$ such that $x_i \neq 0$ and $x_{i-1} = 0$, and hence there exists a submatrix with negative determinant

$$\begin{vmatrix} x_{i+1} & x_i & 0 \\ x_i & x_{i-1} & 0 \\ 1 & 1 & 1 \end{vmatrix} = x_{i-1}x_{i+1} - x_i^2 = -x_i^2 < 0.$$

Hence we obtain (3.5).

For any fixed $\boldsymbol{x}_n \in \mathbb{R}^n$ with $x_1 = \cdots = x_{n-2} = 0$ and $x_{n-1} \neq 0$, the quadratic function $\mathbb{R}^3 \ni (\alpha', \beta', \mu') \mapsto Q_n(\boldsymbol{x}_n; \alpha', \beta', \mu')$ can be written in the form

$$Q_n(\boldsymbol{x}_n; \alpha', \beta', \mu') = (x_n - \alpha' x_{n-1} - \mu')^2 + (x_{n-1} - \mu')^2 + (n-2)(\mu')^2, \qquad (\alpha', \beta', \mu') \in \mathbb{R}^3.$$

The system of equation consisting of the first order partial derivates of $Q_n$ with respect to $\alpha'$ and $\mu'$ takes the form

$$x_n - \alpha' x_{n-1} - \mu' = 0,$$

$$x_n - \alpha' x_{n-1} - \mu' + x_{n-1} - \mu' - (n-2)\mu' = 0.$$

Using that $n \geqslant 2$, by an easy computation, we conclude (3.6).

If $\boldsymbol{x}_n \in \mathbb{R}^n$ with $x_1 = \cdots = x_{n-1} = 0$, then $Q_n(\boldsymbol{x}_n; \alpha', \beta', \mu') = (x_n - \mu')^2 + (n-1)(\mu')^2$, which implies (3.7). $\qquad \square$

We note that one could give a different proof of Lemma 3.1 as in Barczy et al. [4, Lemma 2.1].

Next we present a result about the existence and uniqueness of $(\widehat{\alpha}_n(\boldsymbol{X}_n), \widehat{\beta}_n(\boldsymbol{X}_n), \widehat{\mu}_n(\boldsymbol{X}_n))$.

**3.1 Proposition.** *Let $(X_k)_{k \geqslant -1}$ be an INAR(2) process with autoregressive parameters $(\alpha, \beta) \in [0, 1]^2$ such that $\alpha + \beta = 1$ (hence it is unstable). Suppose that $X_0 = X_{-1} = 0$, $\mathbb{E}(\varepsilon_1^2) < \infty$ and $\mu > 0$. Then*

$$\lim_{n \to \infty} \mathbb{P}\left( \sum_{k=1}^{n} X_{k-2}^2 > 0 \right) = 1,$$

*and hence the probability of the existence of a unique CLS estimator $(\widehat{\alpha}_n(\boldsymbol{X}_n), \widehat{\beta}_n(\boldsymbol{X}_n), \widehat{\mu}_n(\boldsymbol{X}_n))$ converges to 1 as $n \to \infty$, and this CLS estimator has the form*

(3.9)
$$\begin{bmatrix} \widehat{\alpha}_n(\boldsymbol{X}_n) \\ \widehat{\beta}_n(\boldsymbol{X}_n) \\ \widehat{\mu}_n(\boldsymbol{X}_n) \end{bmatrix} = \boldsymbol{F}_n^{-1} \boldsymbol{g}_n$$

*on the set $\{\omega \in \Omega : \sum_{k=1}^{n} X_{k-2}(\omega)^2 > 0\}$, where*

$$\boldsymbol{F}_n := F_n(\boldsymbol{X}_n) = \sum_{k=1}^{n} \begin{bmatrix} X_{k-1}^2 & X_{k-1}X_{k-2} & X_{k-1} \\ X_{k-1}X_{k-2} & X_{k-2}^2 & X_{k-2} \\ X_{k-1} & X_{k-2} & 1 \end{bmatrix}, \qquad \boldsymbol{g}_n := g_n(\boldsymbol{X}_n) = \sum_{k=1}^{n} \begin{bmatrix} X_k X_{k-1} \\ X_k X_{k-2} \\ X_k \end{bmatrix}.$$

11

**Proof.** First we prove the statements for $(\alpha, \beta) \in (0,1)^2$. For each $n \in \mathbb{N}$, consider the random step process

$$\mathcal{X}_t^{(n)} := n^{-1} X_{\lfloor nt \rfloor}, \qquad t \in \mathbb{R}_+,$$

where $\lfloor x \rfloor$ denotes the integer part of a real number $x \in \mathbb{R}$. By Barczy et al. [5, Theorem 3.1] we have

$$(3.10) \qquad\qquad \mathcal{X}^{(n)} \xrightarrow{\mathcal{L}} \mathcal{X} \qquad \text{as } n \to \infty,$$

where the process $(\mathcal{X}_t)_{t \in \mathbb{R}_+}$ is the unique strong solution of the SDE (2.4) with initial value $\mathcal{X}_0 = 0$. Next we show that

$$(3.11) \qquad\qquad \frac{1}{n^3} \sum_{k=1}^{n} X_{k-2}^2 \xrightarrow{\mathcal{L}} \int_0^1 \mathcal{X}_t^2 \, dt \qquad \text{as } n \to \infty.$$

Let us apply Lemmas B.2 and B.3 with the special choices $d := p := q := 1$, $h : \mathbb{R} \to \mathbb{R}$, $h(x) := x$, $x \in \mathbb{R}$, $K : [0,1] \times \mathbb{R}^2 \to \mathbb{R}$,

$$K(s, x_1, x_2) := x_1^2, \qquad (s, x_1, x_2) \in [0,1] \times \mathbb{R}^2,$$

and $\mathcal{U} := \mathcal{X}$, $\mathcal{U}^{(n)} := \mathcal{X}^{(n)}$, $n \in \mathbb{N}$. Then

$$|K(s, x_1, x_2) - K(t, y_1, y_2)| = |x_1^2 - y_1^2| \leqslant (|x_1| + |y_1|)|x_1 - y_1| \leqslant 2R(|t - s| + |x_1 - y_1|)$$

$$\leqslant 2R\big(|t - s| + \|(x_1, x_2) - (y_1, y_2)\|\big)$$

for all $s, t \in [0,1]$ and $(x_1, x_2), (y_1, y_2) \in \mathbb{R}^2$ with $\|(x_1, x_2)\| \leqslant R$ and $\|(y_1, y_2)\| \leqslant R$, where $R > 0$. Further, using the definition of $\Phi$ and $\Phi_n$, $n \in \mathbb{N}$, given in Lemma B.3,

$$\Phi_n(\mathcal{X}^{(n)}) = \left( \mathcal{X}_1^{(n)}, \frac{1}{n} \sum_{k=1}^{n} (\mathcal{X}_{k/n}^{(n)})^2 \right) = \left( \frac{1}{n} X_n, \frac{1}{n^3} \sum_{k=1}^{n} X_k^2 \right),$$

$$\Phi(\mathcal{X}) = \left( \mathcal{X}_1, \int_0^1 \mathcal{X}_u^2 \, du \right).$$

Since the process $(\mathcal{X}_t)_{t \in \mathbb{R}_+}$ admits continuous paths with probability one, Lemma B.2 (with the choice $C := \mathsf{C}(\mathbb{R}_+, \mathbb{R})$) and Lemma B.3 yield (3.11). Since $\mu > 0$, by the SDE (2.4), we have $\mathbb{P}(\mathcal{X}_t = 0, t \in [0,1]) = 0$, which implies that $\mathbb{P}(\int_0^1 \mathcal{X}_t^2 \, dt > 0) = 1$. Consequently, the distribution function of $\int_0^1 \mathcal{X}_t^2 \, dt$ is continuous at 0, and hence, by (3.11),

$$\mathbb{P}\left( \sum_{k=1}^{n} X_{k-2}^2 > 0 \right) = \mathbb{P}\left( \frac{1}{n^3} \sum_{k=1}^{n} X_{k-2}^2 > 0 \right) \to \mathbb{P}\left( \int_0^1 \mathcal{X}_t^2 \, dt > 0 \right) = 1 \qquad \text{as } n \to \infty.$$

Clearly, (3.9) also holds, hence we obtain the statement in the case of $(\alpha, \beta) \in (0,1)^2$.

Next we consider the case of $(\alpha, \beta) = (1, 0)$. In this case equation (1.1) has the form $X_k = X_{k-1} + \varepsilon_k$, $k \in \mathbb{N}$, and hence $X_n = \sum_{k=1}^{n} \varepsilon_k$, $n \in \mathbb{N}$. By the strong law of large numbers we have

$$(3.12) \qquad\qquad n^{-1} X_n \xrightarrow{\text{a.s.}} \mu,$$

12

and hence
$$n^{-2}X_n^2 \xrightarrow{\text{a.s.}} \mu^2,$$

where $\xrightarrow{\text{a.s.}}$ denotes almost sure convergence. Then, by Toeplitz theorem, we conclude

$$(3.13) \qquad n^{-3}\sum_{k=1}^n X_k^2 = \sum_{k=1}^n \frac{k^2}{n^3}k^{-2}X_k^2 \xrightarrow{\text{a.s.}} \frac{1}{3}\mu^2,$$

where we used that

$$\lim_{n\to\infty}\sum_{k=1}^n \frac{k^2}{n^3} = \lim_{n\to\infty}\frac{n(n+1)(2n+1)}{6n^3} = \frac{1}{3}.$$

Since $\mu > 0$, this implies the existence of an event $\Omega_0 \in \mathcal{A}$ such that $\mathbb{P}(\Omega_0) = 1$, and for all $\omega \in \Omega_0$ there exists an $n_0(\omega) \in \mathbb{N}$ such that $\sum_{k=1}^n X_{k-2}(\omega)^2 > 0$ for $n \geqslant n_0(\omega)$. This is equivalent with $\mathbb{P}(\bigcup_{n=1}^\infty \{\sum_{k=1}^n X_{k-2}^2 > 0\}) = 1$, and, by continuity of probability, is also equivalent with $\lim_{n\to\infty}\mathbb{P}(\{\sum_{k=1}^n X_{k-2}^2 > 0\}) = 1$. Clearly (3.9) also holds, hence we obtain the statement in case $(\alpha, \beta) = (1, 0)$.

Finally, we consider the case $(\alpha, \beta) = (0, 1)$. In this case equation (1.1) has the form $X_k = X_{k-2} + \varepsilon_k$, $k \in \mathbb{N}$, and hence $X_{2n} = \sum_{k=1}^n \varepsilon_{2k}$, $X_{2n-1} = \sum_{k=1}^n \varepsilon_{2k-1}$, $n \in \mathbb{N}$. By the strong law of large numbers we have

$$n^{-1}X_{2n} \xrightarrow{\text{a.s.}} \mu, \qquad n^{-1}X_{2n-1} \xrightarrow{\text{a.s.}} \mu,$$

which yield that

$$(3.14) \qquad n^{-1}X_n \xrightarrow{\text{a.s.}} \frac{1}{2}\mu.$$

Using Toeplitz theorem, as in case $(\alpha, \beta) = (1, 0)$, we get

$$(3.15) \qquad n^{-3}\sum_{k=1}^n X_k^2 \xrightarrow{\text{a.s.}} \frac{1}{12}\mu^2.$$

One can finish the proof as in case $(\alpha, \beta) = (1, 0)$. $\qquad\square$

The recursion (3.3) can also be written in the form

$$(3.16) \qquad X_k = \varrho X_{k-1} - \beta(X_{k-1} - X_{k-2}) + M_k + \mu, \qquad k \in \mathbb{N}.$$

The representation (3.16) is called the canonical form of Sims, Stock and Watson [32], see also Hamilton [15, 17.7.6]. A natural CLS estimator of the stability parameter $\varrho$ takes the form $\widehat{\varrho}_n(\boldsymbol{X}_n) = \widehat{\alpha}_n(\boldsymbol{X}_n) + \widehat{\beta}_n(\boldsymbol{X}_n)$, since, for each $n \in \mathbb{N}$, a CLS estimator $(\widehat{\varrho}_n(\boldsymbol{X}_n), \widehat{\beta}_n(\boldsymbol{X}_n), \widehat{\mu}_n(\boldsymbol{X}_n))$ of $(\varrho, \beta, \mu)$ based on a sample $\boldsymbol{X}_n = (X_1, \ldots, X_n)$ can be obtained by minimizing the sum of squares

$$(3.17) \qquad \sum_{k=1}^n \left(X_k - \mathbb{E}(X_k \mid \mathcal{F}_{k-1})\right)^2 = \sum_{k=1}^n \left(X_k - \varrho X_{k-1} + \beta(X_{k-1} - X_{k-2}) - \mu\right)^2$$

13

with respect to $(\varrho, \beta, \mu)$ over $\mathbb{R}^3$. One can easily argue that any CLS estimator $(\widehat{\varrho}_n, \widehat{\beta}_n, \widehat{\mu}_n)$ : $\mathbb{R}^n \to \mathbb{R}^3$ of $(\varrho, \beta, \mu)$ is of the form

(3.18)
$$
\begin{bmatrix} \widehat{\varrho}_n(\boldsymbol{x}_n) \\ \widehat{\beta}_n(\boldsymbol{x}_n) \\ \widehat{\mu}_n(\boldsymbol{x}_n) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \widehat{\alpha}_n(\boldsymbol{x}_n) \\ \widehat{\beta}_n(\boldsymbol{x}_n) \\ \widehat{\mu}_n(\boldsymbol{x}_n) \end{bmatrix}, \qquad \boldsymbol{x}_n \in \mathbb{R}^n,
$$

where $(\widehat{\alpha}_n, \widehat{\beta}_n, \widehat{\mu}_n)$ is a CLS estimator of $(\alpha, \beta, \mu)$. Namely, if $\psi : \mathbb{R}^3 \to \mathbb{R}^3$ is a bijective measurable function such that

$$
\mathbb{R}^3 \ni (\alpha', \beta', \mu') \mapsto \psi(\alpha', \beta', \mu') := \begin{bmatrix} \alpha' + \beta' \\ h(\alpha', \beta', \mu') \end{bmatrix} =: \begin{bmatrix} \varrho' \\ \gamma' \\ \delta' \end{bmatrix}
$$

with some function $h : \mathbb{R}^3 \to \mathbb{R}^2$, then there is a bijection between the set of CLS estimators of the parameters $(\alpha, \beta, \mu)$ and the set of CLS estimators of the parameters $\psi(\alpha, \beta, \mu)$. Indeed, for all $n \in \mathbb{N}$, $(x_1, \ldots, x_n) \in \mathbb{R}^n$ and $(\alpha', \beta', \mu') \in \mathbb{R}^3$,

$$
\sum_{k=1}^{n} (x_k - \alpha' x_{k-1} - \beta' x_{k-2} - \mu')^2 = \sum_{k=1}^{n} \left( x_k - \begin{bmatrix} \alpha' \\ \beta' \\ \mu' \end{bmatrix}^{\top} \begin{bmatrix} x_{k-1} \\ x_{k-2} \\ 1 \end{bmatrix} \right)^2
$$

$$
= \sum_{k=1}^{n} \left( x_k - \left( \psi^{-1}(\varrho', \gamma', \delta') \right)^{\top} \begin{bmatrix} x_{k-1} \\ x_{k-2} \\ 1 \end{bmatrix} \right)^2,
$$

hence $(\widehat{\alpha}_n, \widehat{\beta}_n, \widehat{\mu}_n) : \mathbb{R}^n \to \mathbb{R}^3$ is a CLS estimator of $(\alpha, \beta, \mu)$ if and only if $\psi(\widehat{\alpha}_n, \widehat{\beta}_n, \widehat{\mu}_n)$ is a CLS estimator of $\psi(\alpha, \beta, \mu)$. With the special choice $h : \mathbb{R}^3 \to \mathbb{R}^2$, $h(\alpha, \beta, \mu) := (\beta, \mu)$, $(\alpha, \beta, \mu) \in \mathbb{R}^3$, we get (3.18). In what follows, by speaking about the CLS estimator $\widehat{\varrho}_n$ of $\varrho$ we mean the first coordinate of $\psi(\widehat{\alpha}_n, \widehat{\beta}_n)$. Hence, by Proposition 3.1, the probability of the existence of a unique CLS estimator $(\widehat{\varrho}_n(\boldsymbol{X}_n), \widehat{\beta}_n(\boldsymbol{X}_n), \widehat{\mu}_n(\boldsymbol{X}_n))$ converges to 1 as $n \to \infty$, and this CLS estimator has the form

(3.19)
$$
\begin{bmatrix} \widehat{\varrho}_n(\boldsymbol{X}_n) \\ \widehat{\beta}_n(\boldsymbol{X}_n) \\ \widehat{\mu}_n(\boldsymbol{X}_n) \end{bmatrix} = \boldsymbol{A}_n^{-1} \boldsymbol{b}_n
$$

on the set $\{\omega \in \Omega : \sum_{k=1}^{n} X_{k-2}(\omega)^2 > 0\}$, where

$$
\boldsymbol{A}_n := \sum_{k=1}^{n} \begin{bmatrix} X_{k-1}^2 & -X_{k-1}V_{k-1} & X_{k-1} \\ -X_{k-1}V_{k-1} & V_{k-1}^2 & -V_{k-1} \\ X_{k-1} & -V_{k-1} & 1 \end{bmatrix}, \qquad \boldsymbol{b}_n := \sum_{k=1}^{n} \begin{bmatrix} X_k X_{k-1} \\ -X_k V_{k-1} \\ X_k \end{bmatrix}
$$

14

with

$$V_{k-1} := X_{k-1} - X_{k-2}, \qquad k \in \mathbb{N}.$$

(In Appendix A, in Remark 9.2 one can find a detailed motivation of the definition of $V_k$, $k \in \mathbb{N}$.) Indeed, by (3.9),

$$\begin{bmatrix} \widehat{\varrho}_n(\boldsymbol{X}_n) \\ \widehat{\beta}_n(\boldsymbol{X}_n) \\ \widehat{\mu}_n(\boldsymbol{X}_n) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \boldsymbol{F}_n^{-1} \boldsymbol{g}_n = \left( \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \boldsymbol{F}_n \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}^{-1} \right)^{-1} \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \boldsymbol{g}_n = \boldsymbol{A}_n^{-1} \boldsymbol{b}_n,$$

which also shows that $\boldsymbol{A}_n^{-1}$ exists on the set $\{\omega \in \Omega : \sum_{k=1}^n X_{k-2}(\omega)^2 > 0\}$.

Alternatively, the CLS estimator $\widehat{\varrho}_n(\boldsymbol{X}_n)$ of the stability parameter $\varrho$ could also be obtained via a CLS estimator $(\widehat{\alpha}_n(\boldsymbol{X}_n), \widehat{\varrho}_n(\boldsymbol{X}_n), \widehat{\mu}_n(\boldsymbol{X}_n))$ of $(\alpha, \varrho, \mu)$.

Note also that in case of an unstable INAR(2) process, i.e., when $\varrho = 1$, we have

$$(3.20) \qquad V_k = -\beta V_{k-1} + M_k + \mu, \qquad k \in \mathbb{N},$$

hence $(V_k)_{k \in \mathbb{N}}$ is a stable AR(1) process with heteroscedastic innovations $(M_k)_{k \in \mathbb{N}}$ and with positive drift $\mu$ whenever $0 \leqslant \beta < 1$.

# 4 Proof of the main results

In case of an unstable INAR(2) process, i.e., when $\varrho = \alpha + \beta = 1$, by (3.19), we have

$$(4.1) \qquad \begin{bmatrix} \widehat{\varrho}_n(\boldsymbol{X}_n) - 1 \\ \widehat{\beta}_n(\boldsymbol{X}_n) - \beta \\ \widehat{\mu}_n(\boldsymbol{X}_n) - \mu \end{bmatrix} = \boldsymbol{A}_n^{-1} \boldsymbol{d}_n, \qquad n \in \mathbb{N},$$

on the set $\{\omega \in \Omega : \sum_{k=1}^n X_{k-2}(\omega)^2 > 0\}$, where

$$\boldsymbol{d}_n := \sum_{k=1}^n \begin{bmatrix} M_k X_{k-1} \\ -M_k V_{k-1} \\ M_k \end{bmatrix}, \qquad n \in \mathbb{N}.$$

Theorems 2.1, 2.2, and 2.3 will follow from Theorems 4.1, 4.2, and 4.3, respectively (see the details below).

**4.1 Theorem.** *Under the assumptions of Theorem 2.1 we have* $(\widetilde{\boldsymbol{A}}_n, \widetilde{\boldsymbol{d}}_n) \xrightarrow{\mathcal{L}} (\widetilde{\boldsymbol{A}}, \widetilde{\boldsymbol{d}})$ *as* $n \to \infty$, *where*

$$\widetilde{\boldsymbol{A}}_n := \begin{bmatrix} n^{-3/2} & 0 & 0 \\ 0 & n^{-1} & 0 \\ 0 & 0 & n^{-1/2} \end{bmatrix} \boldsymbol{A}_n \begin{bmatrix} n^{-3/2} & 0 & 0 \\ 0 & n^{-1} & 0 \\ 0 & 0 & n^{-1/2} \end{bmatrix}, \qquad \widetilde{\boldsymbol{d}}_n := \begin{bmatrix} n^{-2} & 0 & 0 \\ 0 & n^{-3/2} & 0 \\ 0 & 0 & n^{-1} \end{bmatrix} \boldsymbol{d}_n,$$

15

$$\widetilde{A} := \begin{bmatrix} \int_0^1 \mathcal{X}_t^2 \, dt & 0 & \int_0^1 \mathcal{X}_t \, dt \\ 0 & \frac{2\beta}{1+\beta} \int_0^1 \mathcal{X}_t \, dt & 0 \\ \int_0^1 \mathcal{X}_t \, dt & 0 & 1 \end{bmatrix}, \qquad \widetilde{d} := \begin{bmatrix} \sqrt{2\alpha\beta} \int_0^1 \mathcal{X}_t^{3/2} \, d\mathcal{W}_t \\ -\frac{2\beta\sqrt{\alpha}}{\sqrt{1+\beta}} \int_0^1 \mathcal{X}_t \, d\widetilde{\mathcal{W}}_t \\ (1+\beta)\mathcal{X}_1 - \mu \end{bmatrix},$$

where $(\mathcal{W}_t)_{t \in \mathbb{R}_+}$ and $(\widetilde{\mathcal{W}}_t)_{t \in \mathbb{R}_+}$ are independent standard Wiener processes.

**4.2 Theorem.** *Under the assumptions of Theorem 2.2 we have* $(\widetilde{A}_n, \widetilde{d}_n) \xrightarrow{\mathcal{L}} (\widetilde{A}, \widetilde{d})$ *as* $n \to \infty$, *where*

$$\widetilde{A}_n := \begin{bmatrix} n^{-3/2} & 0 & 0 \\ 0 & n^{-1/2} & 0 \\ 0 & 0 & n^{-1/2} \end{bmatrix} A_n \begin{bmatrix} n^{-3/2} & 0 & 0 \\ 0 & n^{-1/2} & 0 \\ 0 & 0 & n^{-1/2} \end{bmatrix}, \quad \widetilde{d}_n := \begin{bmatrix} n^{-3/2} & 0 & 0 \\ 0 & n^{-1/2} & 0 \\ 0 & 0 & n^{-1/2} \end{bmatrix} d_n,$$

$$\widetilde{A} := \begin{bmatrix} \frac{1}{3}\mu^2 & -\frac{1}{2}\mu^2 & \frac{1}{2}\mu \\ -\frac{1}{2}\mu^2 & \mu^2 + \sigma^2 & -\mu \\ \frac{1}{2}\mu & -\mu & 1 \end{bmatrix}, \qquad \widetilde{d} \overset{\mathcal{L}}{=} \mathcal{N}_3(\mathbf{0}, \sigma^2 \widetilde{A}).$$

**4.3 Theorem.** *Under the assumptions of Theorem 2.3 we have* $(\widetilde{A}_n, \widetilde{d}_n) \xrightarrow{\mathcal{L}} (\widetilde{A}, \widetilde{d})$ *as* $n \to \infty$, *where*

$$\widetilde{A}_n := \begin{bmatrix} n^{-3/2} & 0 & 0 \\ 0 & n^{-1} & 0 \\ 0 & 0 & n^{-1/2} \end{bmatrix} A_n \begin{bmatrix} n^{-3/2} & 0 & 0 \\ 0 & n^{-1} & 0 \\ 0 & 0 & n^{-1/2} \end{bmatrix}, \qquad \widetilde{d}_n := \begin{bmatrix} n^{-3/2} & 0 & 0 \\ 0 & n^{-1} & 0 \\ 0 & 0 & n^{-1/2} \end{bmatrix} d_n,$$

$$\widetilde{A} := \begin{bmatrix} \frac{1}{12}\mu^2 & 0 & \frac{1}{4}\mu \\ 0 & \sigma^2 \int_0^1 \mathcal{W}_t^2 \, dt & 0 \\ \frac{1}{4}\mu & 0 & 1 \end{bmatrix}, \qquad \widetilde{d} := \begin{bmatrix} \frac{1}{2}\mu\sigma \int_0^1 t \, d\widetilde{\mathcal{W}}_t \\ \sigma^2 \int_0^1 \mathcal{W}_t \, d\mathcal{W}_t \\ \sigma\widetilde{\mathcal{W}}_1 \end{bmatrix},$$

where $(\mathcal{W}_t)_{t \in \mathbb{R}_+}$ and $(\widetilde{\mathcal{W}}_t)_{t \in \mathbb{R}_+}$ are independent standard Wiener processes.

Now we briefly summarize how Theorem 4.1 yields Theorem 2.1. The function $g : \mathbb{R}^{3\times3} \times \mathbb{R}^{3\times1} \to \mathbb{R}^{3\times1}$, defined by

(4.2) $$g(\mathbf{X}, \mathbf{y}) := \begin{cases} \mathbf{X}^{-1}\mathbf{y}, & \text{if } \exists \, \mathbf{X}^{-1}, \\ \mathbf{0}, & \text{otherwise,} \end{cases}$$

is continuous on the set $\{\mathbf{X} \in \mathbb{R}^{3\times3} : \exists \, \mathbf{X}^{-1}\} \times \mathbb{R}^{3\times1}$, and the limit distribution in Theorem 4.1 is concentrated on this set, since,

$$\det(\widetilde{A}) = \frac{2\beta}{1+\beta} \int_0^1 \mathcal{X}_t \, dt \left( \int_0^1 \mathcal{X}_t^2 \, dt - \left( \int_0^1 \mathcal{X}_t \, dt \right)^2 \right),$$

16

and, by Remark 2.2 and the proof of Proposition 3.1,

$$\mathbb{P}\left(\int_0^1 \mathcal{X}_t \, dt > 0\right) = \mathbb{P}\left(\int_0^1 \mathcal{X}_t^2 \, dt > 0\right) = 1,$$

and, by Lemma 4.3 in Barczy et al. [3],

$$\mathbb{P}\left(\int_0^1 \mathcal{X}_t^2 \, dt - \left(\int_0^1 \mathcal{X}_t \, dt\right)^2 > 0\right) = 1.$$

Hence the continuous mapping theorem (see, e.g., Theorem 2.3 in van der Vaart [35]) yields that

$$g(\widetilde{\boldsymbol{A}}_n, \widetilde{\boldsymbol{d}}_n) \xrightarrow{\mathcal{L}} g(\widetilde{\boldsymbol{A}}, \widetilde{\boldsymbol{d}})$$

as $n \to \infty$. Under the conditions of Proposition 3.1, by (3.11) and (4.1), we have

$$\mathbb{P}\left(\begin{bmatrix} n & 0 & 0 \\ 0 & n^{1/2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \widehat{\varrho}_n(\boldsymbol{X}_n) - 1 \\ \widehat{\beta}_n(\boldsymbol{X}_n) - \beta \\ \widehat{\mu}_n(\boldsymbol{X}_n) - \mu \end{bmatrix} = g(\widetilde{\boldsymbol{A}}_n, \widetilde{\boldsymbol{d}}_n)\right) \geqslant \mathbb{P}\left(\exists \, \widetilde{\boldsymbol{A}}_n^{-1}\right) = \mathbb{P}\left(\exists \, \boldsymbol{A}_n^{-1}\right)$$

$$\geqslant \mathbb{P}\left(\sum_{k=1}^n X_{k-2}^2 > 0\right) = \mathbb{P}\left(\frac{1}{n^3}\sum_{k=1}^n X_{k-2}^2 > 0\right) \to \mathbb{P}\left(\int_0^1 \mathcal{X}_t^2 \, dt > 0\right) = 1$$

as $n \to \infty$. Clearly, if $\xi_n$, $\eta_n$, $n \in \mathbb{N}$, and $\xi$ are random variables such that $\xi_n \xrightarrow{\mathcal{L}} \xi$ as $n \to \infty$ and $\lim_{n\to\infty} \mathbb{P}(\xi_n = \eta_n) = 1$, then $\eta_n \xrightarrow{\mathcal{L}} \xi$ as $n \to \infty$, see, e.g., Barczy et al. [4, Lemma 3.1]. Consequently, under the conditions of Theorem 2.1, Theorem 4.1 yields that

$$\begin{bmatrix} n(\widehat{\varrho}_n(\boldsymbol{X}_n) - 1) \\ n^{1/2}(\widehat{\beta}_n(\boldsymbol{X}_n) - \beta) \\ \widehat{\mu}_n(\boldsymbol{X}_n) - \mu \end{bmatrix} \xrightarrow{\mathcal{L}} g(\widetilde{\boldsymbol{A}}, \widetilde{\boldsymbol{d}}) \qquad \text{as } n \to \infty,$$

where

$$g(\widetilde{\boldsymbol{A}}, \widetilde{\boldsymbol{d}}) = \widetilde{\boldsymbol{A}}^{-1}\widetilde{\boldsymbol{d}}$$

$$= \frac{1}{\int_0^1 \mathcal{X}_t^2 \, dt - \left(\int_0^1 \mathcal{X}_t \, dt\right)^2} \begin{bmatrix} 1 & 0 & -\int_0^1 \mathcal{X}_t \, dt \\ 0 & \frac{1+\beta}{2\beta}\frac{\int_0^1 \mathcal{X}_t^2 \, dt - \left(\int_0^1 \mathcal{X}_t \, dt\right)^2}{\int_0^1 \mathcal{X}_t \, dt} & 0 \\ -\int_0^1 \mathcal{X}_t \, dt & 0 & \int_0^1 \mathcal{X}_t^2 \, dt \end{bmatrix} \begin{bmatrix} \sqrt{2\alpha\beta}\int_0^1 \mathcal{X}_t^{3/2} \, d\mathcal{W}_t \\ -\frac{2\beta\sqrt{\alpha}}{\sqrt{1+\beta}}\int_0^1 \mathcal{X}_t \, d\widetilde{\mathcal{W}}_t \\ (1+\beta)\mathcal{X}_1 - \mu \end{bmatrix}$$

$$= \begin{bmatrix} \dfrac{\sqrt{2\alpha\beta}\int_0^1 \mathcal{X}_t^{3/2} \, d\mathcal{W}_t - [(1+\beta)\mathcal{X}_1 - \mu]\int_0^1 \mathcal{X}_t \, dt}{\int_0^1 \mathcal{X}_t^2 \, dt - \left(\int_0^1 \mathcal{X}_t \, dt\right)^2} \\[2em] -\dfrac{\sqrt{\alpha(1+\beta)}\int_0^1 \mathcal{X}_t \, d\widetilde{\mathcal{W}}_t}{\int_0^1 \mathcal{X}_t \, dt} \\[2em] \dfrac{-\sqrt{2\alpha\beta}\int_0^1 \mathcal{X}_t \, dt \int_0^1 \mathcal{X}_t^{3/2} \, d\mathcal{W}_t + [(1+\beta)\mathcal{X}_1 - \mu]\int_0^1 \mathcal{X}_t^2 \, dt}{\int_0^1 \mathcal{X}_t^2 \, dt - \left(\int_0^1 \mathcal{X}_t \, dt\right)^2} \end{bmatrix}.$$

17

Hence we obtain (2.1) and (2.3), and, using that $\int_0^1 \mathcal{X}_t \, d\widetilde{\mathcal{W}}_t$ is symmetric, also the convergence of the second coordinate in (2.2). By Slutsky's lemma, convergence (2.1) implies $n^{1/2}(\widehat{\varrho}_n(\boldsymbol{X}_n) - 1) \xrightarrow{\mathbb{P}} 0$ as well, where $\xrightarrow{\mathbb{P}}$ denotes convergence in probability, hence

$$\begin{bmatrix} n^{1/2}(\widehat{\alpha}_n(\boldsymbol{X}_n) - \alpha) \\ n^{1/2}(\widehat{\beta}_n(\boldsymbol{X}_n) - \beta) \end{bmatrix} = n^{1/2}(\widehat{\beta}_n(\boldsymbol{X}_n) - \beta) \begin{bmatrix} -1 \\ 1 \end{bmatrix} + n^{1/2}(\widehat{\varrho}_n(\boldsymbol{X}_n) - 1) \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

yields (2.2).

Next we briefly summarize how Theorem 4.2 yields Theorem 2.2. Similarly to the previous case, under the conditions of Proposition 3.1, by (3.11) and (4.1), we have

$$\mathbb{P}\left( \begin{bmatrix} n^{3/2} & 0 & 0 \\ 0 & n^{1/2} & 0 \\ 0 & 0 & n^{1/2} \end{bmatrix} \begin{bmatrix} \widehat{\varrho}_n(\boldsymbol{X}_n) - 1 \\ \widehat{\beta}_n(\boldsymbol{X}_n) \\ \widehat{\mu}_n(\boldsymbol{X}_n) - \mu \end{bmatrix} = g(\widetilde{\boldsymbol{A}}_n, \widetilde{\boldsymbol{d}}_n) \right) \geqslant \mathbb{P}\left( \sum_{k=1}^n X_{k-2}^2 > 0 \right) \to 1$$

as $n \to \infty$. Consequently, under the conditions of Theorem 2.2, Theorem 4.2 yields that

$$\begin{bmatrix} n^{3/2}(\widehat{\varrho}_n(\boldsymbol{X}_n) - 1) \\ n^{1/2}\widehat{\beta}_n(\boldsymbol{X}_n) \\ n^{1/2}(\widehat{\mu}_n(\boldsymbol{X}_n) - \mu) \end{bmatrix} \xrightarrow{\mathcal{L}} g(\widetilde{\boldsymbol{A}}, \widetilde{\boldsymbol{d}}) \qquad \text{as } n \to \infty,$$

where

$$g(\widetilde{\boldsymbol{A}}, \widetilde{\boldsymbol{d}}) = \widetilde{\boldsymbol{A}}^{-1}\widetilde{\boldsymbol{d}} \overset{\mathcal{L}}{=} \widetilde{\boldsymbol{A}}^{-1}\mathcal{N}_3(\boldsymbol{0}, \sigma^2\widetilde{\boldsymbol{A}}) \overset{\mathcal{L}}{=} \mathcal{N}_3(\boldsymbol{0}, \sigma^2\widetilde{\boldsymbol{A}}^{-1})$$

$$\overset{\mathcal{L}}{=} \mathcal{N}_3\left( \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \frac{1}{\mu^2} \begin{bmatrix} 12\sigma^2 & 0 & -6\mu\sigma^2 \\ 0 & \mu^2 & \mu^3 \\ -6\mu\sigma^2 & \mu^3 & \mu^2(\mu^2 + 4\sigma^2) \end{bmatrix} \right).$$

Hence we obtain (2.8), (2.10), and convergence of the second coordinate in (2.9). By Slutsky's lemma, convergence (2.8) implies $n^{1/2}(\widehat{\varrho}_n(\boldsymbol{X}_n) - 1) \xrightarrow{\mathbb{P}} 0$ as well, hence

$$\begin{bmatrix} n^{1/2}(\widehat{\alpha}_n(\boldsymbol{X}_n) - 1) \\ n^{1/2}\widehat{\beta}_n(\boldsymbol{X}_n) \end{bmatrix} = n^{1/2}\widehat{\beta}_n(\boldsymbol{X}_n) \begin{bmatrix} -1 \\ 1 \end{bmatrix} + n^{1/2}(\widehat{\varrho}_n(\boldsymbol{X}_n) - 1) \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

yields (2.9).

Finally, we briefly summarize how Theorem 4.3 yields Theorem 2.3. Similarly as above, under the conditions of Proposition 3.1, by (3.11) and (4.1), we have

$$\mathbb{P}\left( \begin{bmatrix} n^{3/2} & 0 & 0 \\ 0 & n & 0 \\ 0 & 0 & n^{1/2} \end{bmatrix} \begin{bmatrix} \widehat{\varrho}_n(\boldsymbol{X}_n) - 1 \\ \widehat{\beta}_n(\boldsymbol{X}_n) - 1 \\ \widehat{\mu}_n(\boldsymbol{X}_n) - \mu \end{bmatrix} = g(\widetilde{\boldsymbol{A}}_n, \widetilde{\boldsymbol{d}}_n) \right) \geqslant \mathbb{P}\left( \sum_{k=1}^n X_{k-2}^2 > 0 \right) \to 1$$

18

as $n \to \infty$. Consequently, under the conditions of Theorem 2.3, Theorem 4.3 yields that

$$
\begin{bmatrix} n^{3/2}(\widehat{\varrho}_n(\boldsymbol{X}_n) - 1) \\ n(\widehat{\beta}_n(\boldsymbol{X}_n) - 1) \\ n^{1/2}(\widehat{\mu}_n(\boldsymbol{X}_n) - \mu) \end{bmatrix} \xrightarrow{\mathcal{L}} g(\widetilde{\boldsymbol{A}}, \widetilde{\boldsymbol{d}}) \qquad \text{as } n \to \infty,
$$

where

$$
g(\widetilde{\boldsymbol{A}}, \widetilde{\boldsymbol{d}}) = \widetilde{\boldsymbol{A}}^{-1}\widetilde{\boldsymbol{d}} = \begin{bmatrix} \frac{48}{\mu^2} & 0 & -\frac{12}{\mu} \\ 0 & \frac{1}{\sigma^2 \int_0^1 \mathcal{W}_t^2 \, \mathrm{d}t} & 0 \\ -\frac{12}{\mu} & 0 & 4 \end{bmatrix} \begin{bmatrix} \frac{1}{2}\mu\sigma \int_0^1 t \, \mathrm{d}\widetilde{\mathcal{W}}_t \\ \sigma^2 \int_0^1 \mathcal{W}_t \, \mathrm{d}\mathcal{W}_t \\ \sigma\widetilde{\mathcal{W}}_1 \end{bmatrix}
$$

$$
= \begin{bmatrix} \frac{24\sigma}{\mu} \int_0^1 t \, \mathrm{d}\widetilde{\mathcal{W}}_t - \frac{12\sigma}{\mu}\widetilde{\mathcal{W}}_1 \\ \frac{\int_0^1 \mathcal{W}_t \, \mathrm{d}\mathcal{W}_t}{\int_0^1 \mathcal{W}_t^2 \, \mathrm{d}t} \\ -6\sigma \int_0^1 t \, \mathrm{d}\widetilde{\mathcal{W}}_t + 4\sigma\widetilde{\mathcal{W}}_1 \end{bmatrix} = \begin{bmatrix} \frac{12\sigma}{\mu} \int_0^1 (2t - 1) \, \mathrm{d}\widetilde{\mathcal{W}}_t \\ \frac{\int_0^1 \mathcal{W}_t \, \mathrm{d}\mathcal{W}_t}{\int_0^1 \mathcal{W}_t^2 \, \mathrm{d}t} \\ 2\sigma \int_0^1 (2 - 3t) \, \mathrm{d}\widetilde{\mathcal{W}}_t \end{bmatrix}.
$$

Since $\int_0^1 (2t - 1) \, \mathrm{d}\widetilde{\mathcal{W}}_t$ and $\int_0^1 (2 - 3t) \, \mathrm{d}\widetilde{\mathcal{W}}_t$ are normally distributed random variables with mean $0$ and with variance

$$
\int_0^1 (2t - 1)^2 \, \mathrm{d}t = \frac{1}{3}, \qquad \int_0^1 (2 - 3t)^2 \, \mathrm{d}t = 1,
$$

respectively, we obtain (2.11), (2.13), and convergence of the second coordinate in (2.12). By Slutsky's lemma, convergence (2.11) implies $n(\widehat{\varrho}_n(\boldsymbol{X}_n) - 1) \xrightarrow{\mathbb{P}} 0$ as well, hence

$$
\begin{bmatrix} n\widehat{\alpha}_n(\boldsymbol{X}_n) \\ n(\widehat{\beta}_n(\boldsymbol{X}_n) - 1) \end{bmatrix} = n(\widehat{\beta}_n(\boldsymbol{X}_n) - 1) \begin{bmatrix} -1 \\ 1 \end{bmatrix} + n(\widehat{\varrho}_n(\boldsymbol{X}_n) - 1) \begin{bmatrix} 1 \\ 0 \end{bmatrix}
$$

yields (2.12).

# 5 Proof of Theorem 4.1

We have

$$
\widetilde{\boldsymbol{A}}_n = \sum_{k=1}^n \begin{bmatrix} n^{-3}X_{k-1}^2 & -n^{-5/2}X_{k-1}V_{k-1} & n^{-2}X_{k-1} \\ -n^{-5/2}X_{k-1}V_{k-1} & n^{-2}V_{k-1}^2 & -n^{-3/2}V_{k-1} \\ n^{-2}X_{k-1} & -n^{-3/2}V_{k-1} & n^{-1} \end{bmatrix},
$$

(5.1)

$$
\widetilde{\boldsymbol{d}}_n = \sum_{k=1}^n \begin{bmatrix} n^{-2}M_kX_{k-1} \\ -n^{-3/2}M_kV_{k-1} \\ n^{-1}M_k \end{bmatrix}.
$$

19

**5.1 Lemma.** *Under the assumptions of Theorem 2.1 we have*

$$(5.2) \qquad n^{-3/2} \sum_{k=1}^{n} V_k \xrightarrow{\mathbb{P}} 0 \qquad as \ n \to \infty,$$

$$(5.3) \qquad n^{-5/2} \sum_{k=1}^{n} X_k V_k \xrightarrow{\mathbb{P}} 0 \qquad as \ n \to \infty,$$

$$(5.4) \qquad n^{-2} \left( \sum_{k=1}^{n} V_k^2 - \frac{2\beta}{1+\beta} \sum_{k=1}^{n} X_{k-1} \right) \xrightarrow{\mathbb{P}} 0 \qquad as \ n \to \infty.$$

**Proof.** We have $\sum_{k=1}^{n} V_k = X_n \geqslant 0$, $n \in \mathbb{N}$, and, by Corollary 9.1, $\mathbb{E}(X_n) = \mathrm{O}(n)$, hence we conclude (5.2). We have

$$\sum_{k=1}^{n} (X_k - X_{k-1})^2 = \sum_{k=1}^{n} X_k^2 - 2 \sum_{k=1}^{n} X_k X_{k-1} + \sum_{k=1}^{n} X_{k-1}^2 = 2 \sum_{k=1}^{n} X_k^2 - 2 \sum_{k=1}^{n} X_k X_{k-1} - X_n^2$$

$$= 2 \sum_{k=1}^{n} X_k (X_k - X_{k-1}) - X_n^2,$$

thus

$$(5.5) \qquad \sum_{k=1}^{n} X_k V_k = \frac{1}{2} X_n^2 + \frac{1}{2} \sum_{k=1}^{n} V_k^2 \geqslant 0.$$

Corollary 9.1 implies

$$\mathbb{E} \left( \sum_{k=1}^{n} X_k V_k \right) = \frac{1}{2} \mathbb{E}(X_n^2) + \frac{1}{2} \sum_{k=1}^{n} \mathbb{E}(V_k^2) = \mathrm{O}(n^2),$$

hence we obtain (5.3).

In order to prove (5.4) we derive a decomposition of $\sum_{k=1}^{n} V_k^2$ as a sum of a martingale and some negligible terms. Using recursion (3.20) and Lemma 9.1, we obtain

$$\mathbb{E}(V_k^2 \mid \mathcal{F}_{k-1}) = \mathbb{E} \left( (-\beta V_{k-1} + M_k + \mu)^2 \mid \mathcal{F}_{k-1} \right)$$

$$= \beta^2 V_{k-1}^2 - 2\beta\mu V_{k-1} + \mu^2 + \mathbb{E}(M_k^2 \mid \mathcal{F}_{k-1})$$

$$= \beta^2 V_{k-1}^2 - 2\beta\mu V_{k-1} + \mu^2 + \alpha\beta(X_{k-1} + X_{k-2}) + \sigma^2$$

$$= \beta^2 V_{k-1}^2 + 2\alpha\beta X_{k-1} + \mu^2 + \sigma^2 - (2\beta\mu + \alpha\beta)V_{k-1}$$

$$= \beta^2 V_{k-1}^2 + 2\alpha\beta X_{k-1} + \mathrm{constant} + \mathrm{constant} \times V_{k-1},$$

where we used that $X_{k-1} + X_{k-2} = 2X_{k-1} - V_{k-1}$, $k \in \mathbb{N}$. Thus

$$\sum_{k=1}^{n} V_k^2 = \sum_{k=1}^{n} \left[ V_k^2 - \mathbb{E}(V_k^2 \mid \mathcal{F}_{k-1}) \right] + \sum_{k=1}^{n} \mathbb{E}(V_k^2 \mid \mathcal{F}_{k-1})$$

$$= \sum_{k=1}^{n} \left[ V_k^2 - \mathbb{E}(V_k^2 \mid \mathcal{F}_{k-1}) \right] + \beta^2 \sum_{k=1}^{n} V_{k-1}^2 + 2\alpha\beta \sum_{k=1}^{n} X_{k-1} + \mathrm{O}(n) + \mathrm{constant} \times \sum_{k=1}^{n} V_{k-1}.$$

20

Consequently,

$$\sum_{k=1}^{n} V_k^2 = \frac{1}{1-\beta^2} \sum_{k=1}^{n} \left[ V_k^2 - \mathbb{E}(V_k^2 \,|\, \mathcal{F}_{k-1}) \right] + \frac{2\beta}{1+\beta} \sum_{k=1}^{n} X_{k-1}$$

$$- \frac{\beta^2}{1-\beta^2} V_n^2 + \mathrm{O}(n) + \text{constant} \times \sum_{k=1}^{n} V_{k-1}.$$

By the tower rule of conditional expectation, $V_k^2 - \mathbb{E}(V_k^2 \,|\, \mathcal{F}_{k-1})$ and $V_\ell^2 - \mathbb{E}(V_\ell^2 \,|\, \mathcal{F}_{\ell-1})$ are uncorrelated if $k \neq \ell$, so

$$\mathbb{E}\left( \left( \sum_{k=1}^{n} \left[ V_k^2 - \mathbb{E}(V_k^2 \,|\, \mathcal{F}_{k-1}) \right] \right)^2 \right) = \sum_{k=1}^{n} \mathbb{E}\left( \left[ V_k^2 - \mathbb{E}(V_k^2 \,|\, \mathcal{F}_{k-1}) \right]^2 \right) \leqslant \sum_{k=1}^{n} \mathbb{E}(V_k^4) = \mathrm{O}(n^3),$$

where we also used Corollary 9.1 and

(5.7) $$\mathbb{E}\left( \left[ \xi - \mathbb{E}(\xi \,|\, \mathcal{F}) \right]^2 \right) = \mathbb{E}(\xi^2) - \mathbb{E}\left( \mathbb{E}(\xi \,|\, \mathcal{F})^2 \right) \leqslant \mathbb{E}(\xi^2)$$

for an arbitrary random variable $\xi$ with $\mathbb{E}(\xi^2) < \infty$ and $\sigma$-algebra $\mathcal{F} \subset \mathcal{A}$. Hence

$$\frac{1}{n^2} \sum_{k=1}^{n} \left[ V_k^2 - \mathbb{E}(V_k^2 \,|\, \mathcal{F}_{k-1}) \right] \xrightarrow{\mathbb{P}} 0 \qquad \text{as } n \to \infty.$$

We note that this convergence follows also by (9.16) with the choice $(\ell, i, j) = (8, 0, 2)$. Again, by Corollary 9.1, we obtain $\mathbb{E}(V_n^2) = \mathrm{O}(n)$ and $\mathbb{E}(X_{n-1}^2) = \mathrm{O}(n^2)$, and since $\sum_{k=1}^{n} V_{k-1} = X_{n-1}$, $n \in \mathbb{N}$, we get $n^{-2} V_n^2 \xrightarrow{\mathbb{P}} 0$ and $n^{-2} \sum_{k=1}^{n} V_{k-1} \xrightarrow{\mathbb{P}} 0$ as $n \to \infty$ (we note that the second convergence follows also by (9.14) with the choice $(\ell, i, j) = (8, 0, 1)$). Consequently, by (5.6), we obtain (5.4). $\qquad\square$

Now let

$$U_k := X_k + \beta X_{k-1}, \qquad k \in \mathbb{Z}_+,$$

with the convention $U_{-1} := U_0 := 0$. In Appendix A, in Remark 9.2 one can find a detailed motivation of the definition of $U_k$, $k \in \mathbb{N}$. One can observe that $U_k \geqslant 0$ for all $k \in \mathbb{Z}_+$, and, by $\alpha + \beta = 1$,

(5.8) $$U_k = U_{k-1} + M_k + \mu, \qquad k \in \mathbb{Z}_+,$$

hence $(U_k)_{k \in \mathbb{Z}_+}$ is a nonnegative unstable AR(1) process with positive drift $\mu$ sharing the innovation $(M_k)_{k \in \mathbb{N}}$ with the stable AR(1) process $(V_k)_{k \in \mathbb{Z}_+}$.

Consider the sequence of stochastic processes

$$\boldsymbol{\mathcal{Z}}_t^{(n)} := \begin{bmatrix} \mathcal{M}_t^{(n)} \\ \mathcal{N}_t^{(n)} \\ \mathcal{P}_t^{(n)} \end{bmatrix} := \sum_{k=1}^{\lfloor nt \rfloor} \boldsymbol{Z}_k^{(n)} \quad \text{with} \quad \boldsymbol{Z}_k^{(n)} := \begin{bmatrix} n^{-1} M_k \\ n^{-2} M_k U_{k-1} \\ n^{-3/2} M_k V_{k-1} \end{bmatrix}, \qquad t \in \mathbb{R}_+, \quad k, n \in \mathbb{N}.$$

Theorem 4.1 will follow from Lemma 5.1 and the following theorem (which will be detailed after Theorem 5.1).

**5.1 Theorem.** *Under the assumptions of Theorem 2.1 we have*

(5.9)
$$\mathbf{Z}^{(n)} \xrightarrow{\mathcal{L}} \mathbf{Z} \qquad as \ \ n \to \infty,$$

*where the process* $(\mathbf{Z}_t)_{t \in \mathbb{R}_+}$ *with values in* $\mathbb{R}^3$ *is the unique strong solution of the SDE*

(5.10)
$$\mathrm{d}\mathbf{Z}_t = \gamma(t, \mathbf{Z}_t) \, \mathrm{d}\mathbf{W}_t, \qquad t \in \mathbb{R}_+,$$

*with initial value* $\mathbf{Z}_0 = \mathbf{0}$, *where* $\mathbf{W}_t := \begin{bmatrix} \mathcal{W}_t & \widetilde{\mathcal{W}}_t \end{bmatrix}^\top$, $t \in \mathbb{R}_+$, *being a 2-dimensional standard Wiener process, and* $\gamma : \mathbb{R}_+ \times \mathbb{R}^3 \to \mathbb{R}^{3 \times 2}$ *is defined by*

$$\gamma(t, \mathbf{x}) := \begin{bmatrix} \sqrt{\frac{2\alpha\beta}{1+\beta}}[(x_1 + \mu t)^+]^{1/2} & 0 \\ \sqrt{\frac{2\alpha\beta}{1+\beta}}[(x_1 + \mu t)^+]^{3/2} & 0 \\ 0 & \frac{2\beta\sqrt{\alpha}}{(1+\beta)^{3/2}}(x_1 + \mu t) \end{bmatrix}$$

*for* $t \in \mathbb{R}_+$ *and* $\mathbf{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$.

Indeed, the unique strong solution of (5.10) with initial value $\mathbf{Z}_0 = \mathbf{0}$ can be written in form

$$\mathbf{Z}_t := \begin{bmatrix} \mathcal{M}_t \\ \mathcal{N}_t \\ \mathcal{P}_t \end{bmatrix} := \begin{bmatrix} (1+\beta)\mathcal{X}_t - \mu t \\ (1+\beta)\sqrt{2\alpha\beta} \int_0^t \mathcal{X}_s^{3/2} \, \mathrm{d}\mathcal{W}_s \\ \frac{2\beta\sqrt{\alpha}}{\sqrt{1+\beta}} \int_0^t \mathcal{X}_s \, \mathrm{d}\widetilde{\mathcal{W}}_s \end{bmatrix}, \qquad t \in \mathbb{R}_+,$$

since, by Remark 2.3,

$$\mathrm{d}\mathbf{Z}_t = \begin{bmatrix} \mathrm{d}\mathcal{M}_t \\ \mathrm{d}\mathcal{N}_t \\ \mathrm{d}\mathcal{P}_t \end{bmatrix} = \begin{bmatrix} \sqrt{\frac{2\alpha\beta}{1+\beta}}[(\mathcal{M}_t + \mu t)^+]^{1/2} \, \mathrm{d}\mathcal{W}_t \\ \sqrt{\frac{2\alpha\beta}{1+\beta}}[(\mathcal{M}_t + \mu t)^+]^{3/2} \, \mathrm{d}\mathcal{W}_t \\ \frac{2\beta\sqrt{\alpha}}{(1+\beta)^{3/2}}(\mathcal{M}_t + \mu t) \, \mathrm{d}\widetilde{\mathcal{W}}_t \end{bmatrix} = \begin{bmatrix} \sqrt{2\alpha\beta}\mathcal{X}_t^{1/2} \, \mathrm{d}\mathcal{W}_t \\ (1+\beta)\sqrt{2\alpha\beta}\mathcal{X}_t^{3/2} \, \mathrm{d}\mathcal{W}_t \\ \frac{2\beta\sqrt{\alpha}}{\sqrt{1+\beta}}\mathcal{X}_t \, \mathrm{d}\widetilde{\mathcal{W}}_t \end{bmatrix}, \qquad t \in \mathbb{R}_+.$$

By the method of the proof of $\mathcal{X}^{(n)} \xrightarrow{\mathcal{L}} \mathcal{X}$ in Theorem 3.1 in Barczy et al. [5] one can easily derive

(5.11)
$$\begin{bmatrix} \mathcal{X}^{(n)} \\ \mathbf{Z}^{(n)} \end{bmatrix} \xrightarrow{\mathcal{L}} \begin{bmatrix} \mathcal{X} \\ \mathbf{Z} \end{bmatrix} \qquad as \ \ n \to \infty.$$

More precisely, using that

$$X_k = \sum_{j=1}^{k} (M_j + \mu) \mathbf{e}_1^\top A^{k-j} \mathbf{e}_1, \qquad k \in \mathbb{N}, \qquad where \ \ \mathbf{e}_1 := \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

see, e.g., Barczy et al. [5, (3.11)], we have

$$\begin{bmatrix} \mathcal{X}^{(n)} \\ \mathbf{Z}^{(n)} \end{bmatrix} = \psi_n(\mathbf{Z}^{(n)}), \qquad n \in \mathbb{N},$$

22

where the mapping $\psi_n : \mathsf{D}(\mathbb{R}_+, \mathbb{R}^3) \to \mathsf{D}(\mathbb{R}_+, \mathbb{R}^4)$ is given by

$$\psi_n(f_1, f_2, f_3)(t) := \begin{bmatrix} \sum_{j=1}^{\lfloor nt \rfloor} \left( f_1\left(\frac{j}{n}\right) - f_1\left(\frac{j-1}{n}\right) + \frac{\mu}{n} \right) \boldsymbol{e}_1^\top A^{\lfloor nt \rfloor - j} \boldsymbol{e}_1 \\ f_1(t) \\ f_2(t) \\ f_3(t) \end{bmatrix}$$

for $f_1, f_2, f_3 \in \mathsf{D}(\mathbb{R}_+, \mathbb{R})$, $t \in \mathbb{R}_+$, $n \in \mathbb{N}$. Further, using that, by Remark 2.3,

$$\mathcal{X}_t = \frac{1}{1 + \beta}(\mathcal{M}_t + \mu t), \qquad t \in \mathbb{R}_+,$$

we have

$$\begin{bmatrix} \mathcal{X} \\ \mathcal{Z} \end{bmatrix} = \psi(\mathcal{Z}),$$

where the mapping $\psi : \mathsf{D}(\mathbb{R}_+, \mathbb{R}^3) \to \mathsf{D}(\mathbb{R}_+, \mathbb{R}^4)$ is given by

$$\psi(f_1, f_2, f_3)(t) := \begin{bmatrix} \frac{1}{1+\beta}(f_1(t) + \mu t) \\ f_1(t) \\ f_2(t) \\ f_3(t) \end{bmatrix}$$

for $f_1, f_2, f_3 \in \mathsf{D}(\mathbb{R}_+, \mathbb{R})$ and $t \in \mathbb{R}_+$. By page 603 in Barczy et al. [5], the mappings $\psi_n$, $n \in \mathbb{N}$, and $\psi$ are measurable (the latter one is continuous too), since the coordinate functions are measurable. Using page 604 in Barczy et al. [5], we get the set

$$C := \left\{ f \in \mathsf{C}(\mathbb{R}_+, \mathbb{R}^3) : f(0) = \boldsymbol{0} \in \mathbb{R}^3 \right\}$$

has the properties $C \subseteq C_{\psi,(\psi_n)_{n \in \mathbb{N}}}$ with $C \in \mathcal{B}(\mathsf{D}(\mathbb{R}_+, \mathbb{R}^3))$ and $\mathbb{P}(\mathcal{Z} \in C) = 1$, where $C_{\psi,(\psi_n)_{n \in \mathbb{N}}}$ is defined in Appendix B. Hence, by (5.9) and Lemma B.2, we have

$$\begin{bmatrix} \mathcal{X}^{(n)} \\ \mathcal{Z}^{(n)} \end{bmatrix} = \psi_n(\mathcal{Z}^{(n)}) \xrightarrow{\mathcal{L}} \psi(\mathcal{Z}) = \begin{bmatrix} \mathcal{X} \\ \mathcal{Z} \end{bmatrix} \qquad \text{as } n \to \infty,$$

as desired. Next, similarly to the proof of (3.11), by Lemmas B.2 and B.3, convergence (5.11) implies

$$(5.12) \qquad \sum_{k=1}^{n} \begin{bmatrix} n^{-1} M_k \\ n^{-3} X_{k-1}^2 \\ n^{-2} X_{k-1} \\ n^{-2} M_k U_{k-1} \\ n^{-3/2} M_k V_{k-1} \end{bmatrix} \xrightarrow{\mathcal{L}} \begin{bmatrix} (1+\beta)\mathcal{X}_1 - \mu \\ \int_0^1 \mathcal{X}_t^2 \, dt \\ \int_0^1 \mathcal{X}_t \, dt \\ (1+\beta)\sqrt{2\alpha\beta} \int_0^1 \mathcal{X}_t^{3/2} \, d\mathcal{W}_t \\ \frac{2\beta\sqrt{\alpha}}{\sqrt{1+\beta}} \int_0^1 \mathcal{X}_t \, d\widetilde{\mathcal{W}}_t \end{bmatrix} \qquad \text{as } n \to \infty.$$

23

Namely,

$$
\sum_{k=1}^{\lfloor nt \rfloor}
\begin{bmatrix}
n^{-1} M_k \\
n^{-3} X_{k-1}^2 \\
n^{-2} X_{k-1} \\
n^{-2} M_k U_{k-1} \\
n^{-3/2} M_k V_{k-1}
\end{bmatrix}
= \widetilde{\psi}_n \left( \begin{bmatrix} \mathcal{X}^{(n)} \\ \mathcal{Z}^{(n)} \end{bmatrix} \right)(t), \qquad t \in \mathbb{R}_+, \quad n \in \mathbb{N},
$$

where $\widetilde{\psi}_n : \mathsf{D}(\mathbb{R}_+, \mathbb{R}^4) \to \mathsf{D}(\mathbb{R}_+, \mathbb{R}^5)$ is given by

$$
\widetilde{\psi}_n(f_1, f_2, f_3, f_4)(t) :=
\begin{bmatrix}
f_2(t) \\
n^{-1} \sum_{k=1}^{\lfloor nt \rfloor} \left( f_1 \left( \frac{k-1}{n} \right) \right)^2 \\
n^{-1} \sum_{k=1}^{\lfloor nt \rfloor} f_1 \left( \frac{k-1}{n} \right) \\
f_3(t) \\
f_4(t)
\end{bmatrix}
$$

for $f_1, f_2, f_3, f_4 \in \mathsf{D}(\mathbb{R}_+, \mathbb{R})$, $t \in \mathbb{R}_+$, $n \in \mathbb{N}$. Further,

$$
\begin{bmatrix}
(1 + \beta)\mathcal{X}_t - \mu \\
\int_0^t \mathcal{X}_s^2 \, \mathrm{d}s \\
\int_0^t \mathcal{X}_s \, \mathrm{d}s \\
(1 + \beta)\sqrt{2\alpha\beta} \int_0^t \mathcal{X}_s^{3/2} \, \mathrm{d}\mathcal{W}_s \\
\frac{2\beta\sqrt{\alpha}}{\sqrt{1+\beta}} \int_0^t \mathcal{X}_s \, \mathrm{d}\widetilde{\mathcal{W}}_s
\end{bmatrix}
= \widetilde{\psi} \left( \begin{bmatrix} \mathcal{X} \\ \mathcal{Z} \end{bmatrix} \right)(t), \qquad t \in \mathbb{R}_+,
$$

where $\widetilde{\psi} : \mathsf{D}(\mathbb{R}_+, \mathbb{R}^4) \to \mathsf{D}(\mathbb{R}_+, \mathbb{R}^5)$ is given by

$$
\widetilde{\psi}(f_1, f_2, f_3, f_4)(t) :=
\begin{bmatrix}
f_2(t) \\
\int_0^1 (f_1(s))^2 \, \mathrm{d}s \\
\int_0^1 f_1(s) \, \mathrm{d}s \\
f_3(t) \\
f_4(t)
\end{bmatrix}
$$

for $f_1, f_2, f_3, f_4 \in \mathsf{D}(\mathbb{R}_+, \mathbb{R})$, $t \in \mathbb{R}_+$. As in the proof of Lemma B.3, one can check that the set

$$
\widetilde{C} := \left\{ f \in \mathsf{C}(\mathbb{R}_+, \mathbb{R}^4) : f(0) = \mathbf{0} \in \mathbb{R}^4 \right\}
$$

has the properties $\widetilde{C} \subseteq C_{\widetilde{\psi}, (\widetilde{\psi}_n)_{n \in \mathbb{N}}}$ with $\widetilde{C} \in \mathcal{B}(\mathsf{D}(\mathbb{R}_+, \mathbb{R}^4))$ and

$$
\mathbb{P} \left( \begin{bmatrix} \mathcal{X} \\ \mathcal{Z} \end{bmatrix} \in \widetilde{C} \right) = 1.
$$

Hence, by (5.11) and Lemma B.2, we have

$$\widetilde{\psi}_n\left(\begin{bmatrix}\mathcal{X}^{(n)} \\ \mathbf{Z}^{(n)}\end{bmatrix}\right) \xrightarrow{\mathcal{L}} \widetilde{\psi}\left(\begin{bmatrix}\mathcal{X} \\ \mathbf{Z}\end{bmatrix}\right) \quad \text{as } n \to \infty,$$

which yields (5.12).

Using $U_{k-1} = (1+\beta)X_{k-1} - \beta V_{k-1}$ and convergence of the third coordinates in $\mathbf{Z}^{(n)} \xrightarrow{\mathcal{L}} \mathbf{Z}$ as $n \to \infty$ we obtain

$$n^{-2}\left(\sum_{k=1}^{n} M_k X_{k-1} - \frac{1}{1+\beta}\sum_{k=1}^{n} M_k U_{k-1}\right) = \frac{\beta}{(1+\beta)n^2}\sum_{k=1}^{n} M_k V_{k-1} \xrightarrow{\mathbb{P}} 0 \quad \text{as } n \to \infty.$$

Using (5.1), the above two convergences and Lemma 5.1 we obtain Theorem 4.1 by Slutsky's lemma.

# 6 Proof of Theorem 5.1

In order to show convergence $\mathbf{Z}^{(n)} \xrightarrow{\mathcal{L}} \mathbf{Z}$, we apply Theorem C.1 with the special choices $\mathcal{U} := \mathbf{Z}$, $\mathbf{U}_k^{(n)} := \mathbf{Z}_k^{(n)}$, $n, k \in \mathbb{N}$, $(\mathcal{F}_k^{(n)})_{k \in \mathbb{Z}_+} := (\mathcal{F}_k)_{k \in \mathbb{Z}_+}$ and the function $\gamma$ which is defined in Theorem 5.1. Note that the arguments in Section 5 and Remark 2.2 show that the SDE (5.10) admits a unique strong solution $(\mathbf{Z}_t^z)_{t \in \mathbb{R}_+}$ for all initial values $\mathbf{Z}_0^z = z \in \mathbb{R}^3$.

Now we show that conditions (i) and (ii) of Theorem C.1 hold. The conditional variances have the form

$$\mathbb{E}\big(\mathbf{Z}_k^{(n)}(\mathbf{Z}_k^{(n)})^\top \,|\, \mathcal{F}_{k-1}\big) = \mathbb{E}(M_k^2 \,|\, \mathcal{F}_{k-1})\begin{bmatrix} n^{-2} & n^{-3}U_{k-1} & n^{-5/2}V_{k-1} \\ n^{-3}U_{k-1} & n^{-4}U_{k-1}^2 & n^{-7/2}U_{k-1}V_{k-1} \\ n^{-5/2}V_{k-1} & n^{-7/2}U_{k-1}V_{k-1} & n^{-3}V_{k-1}^2 \end{bmatrix}$$

for $n \in \mathbb{N}$, $k \in \{1, \dots, n\}$, and

$$\gamma(s, \mathbf{Z}_s^{(n)})\gamma(s, \mathbf{Z}_s^{(n)})^\top = \begin{bmatrix} \frac{2\alpha\beta}{1+\beta}(\mathcal{M}_s^{(n)} + \mu s) & \frac{2\alpha\beta}{1+\beta}(\mathcal{M}_s^{(n)} + \mu s)^2 & 0 \\ \frac{2\alpha\beta}{1+\beta}(\mathcal{M}_s^{(n)} + \mu s)^2 & \frac{2\alpha\beta}{1+\beta}(\mathcal{M}_s^{(n)} + \mu s)^3 & 0 \\ 0 & 0 & \frac{4\alpha\beta^2}{(1+\beta)^3}(\mathcal{M}_s^{(n)} + \mu s)^2 \end{bmatrix}$$

for $s \in \mathbb{R}_+$, where we used that $(\mathcal{M}_s^{(n)} + \mu s)^+ = \mathcal{M}_s^{(n)} + \mu s$, $s \in \mathbb{R}_+$, $n \in \mathbb{N}$, see Barczy et al. [5, page 598] or (6.7) later on. In order to check condition (i) of Theorem C.1, we need to

25

prove that for each $T > 0$,

$$(6.1) \qquad \sup_{t \in [0,T]} \left| \frac{1}{n^2} \sum_{k=1}^{\lfloor nt \rfloor} \mathbb{E}(M_k^2 \mid \mathcal{F}_{k-1}) - \frac{2\alpha\beta}{1+\beta} \int_0^t (\mathcal{M}_s^{(n)} + \mu s)\, \mathrm{d}s \right| \xrightarrow{\mathbb{P}} 0,$$

$$(6.2) \qquad \sup_{t \in [0,T]} \left| \frac{1}{n^3} \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} \mathbb{E}(M_k^2 \mid \mathcal{F}_{k-1}) - \frac{2\alpha\beta}{1+\beta} \int_0^t (\mathcal{M}_s^{(n)} + \mu s)^2\, \mathrm{d}s \right| \xrightarrow{\mathbb{P}} 0,$$

$$(6.3) \qquad \sup_{t \in [0,T]} \left| \frac{1}{n^4} \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^2 \mathbb{E}(M_k^2 \mid \mathcal{F}_{k-1}) - \frac{2\alpha\beta}{1+\beta} \int_0^t (\mathcal{M}_s^{(n)} + \mu s)^3\, \mathrm{d}s \right| \xrightarrow{\mathbb{P}} 0,$$

$$(6.4) \qquad \sup_{t \in [0,T]} \left| \frac{1}{n^3} \sum_{k=1}^{\lfloor nt \rfloor} V_{k-1}^2 \mathbb{E}(M_k^2 \mid \mathcal{F}_{k-1}) - \frac{4\alpha\beta^2}{(1+\beta)^3} \int_0^t (\mathcal{M}_s^{(n)} + \mu s)^2\, \mathrm{d}s \right| \xrightarrow{\mathbb{P}} 0,$$

$$(6.5) \qquad \sup_{t \in [0,T]} \left| \frac{1}{n^{5/2}} \sum_{k=1}^{\lfloor nt \rfloor} V_{k-1} \mathbb{E}(M_k^2 \mid \mathcal{F}_{k-1}) \right| \xrightarrow{\mathbb{P}} 0,$$

$$(6.6) \qquad \sup_{t \in [0,T]} \left| \frac{1}{n^{7/2}} \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} V_{k-1} \mathbb{E}(M_k^2 \mid \mathcal{F}_{k-1}) \right| \xrightarrow{\mathbb{P}} 0$$

as $n \to \infty$. Covergence (6.1) follows from (5.1) in Barczy et al. [5] with the special choices $p = 2$, $\alpha_1 = \alpha$ and $\alpha_2 = \beta$.

Next we turn to prove (6.2). Since $\alpha + \beta = 1$, by (3.2), we get

$$(6.7) \qquad \begin{aligned} \mathcal{M}_s^{(n)} + \mu s &= \frac{1}{n} \sum_{k=1}^{\lfloor ns \rfloor} \left( X_k - \alpha X_{k-1} - \beta X_{k-2} - \mu \right) + \mu s \\ &= \frac{1}{n} \left( X_{\lfloor ns \rfloor} + \beta X_{\lfloor ns \rfloor - 1} \right) + \frac{ns - \lfloor ns \rfloor}{n} \mu = \frac{1}{n} U_{\lfloor ns \rfloor} + \frac{ns - \lfloor ns \rfloor}{n} \mu \end{aligned}$$

for $s \in \mathbb{R}_+$, $n \in \mathbb{N}$. Thus

$$\begin{aligned} \int_0^t (\mathcal{M}_s^{(n)} + \mu s)^2\, \mathrm{d}s &= \frac{1}{n^3} \sum_{k=1}^{\lfloor nt \rfloor - 1} U_k^2 + \frac{\mu}{n^3} \sum_{k=1}^{\lfloor nt \rfloor - 1} U_k + \frac{nt - \lfloor nt \rfloor}{n^3} U_{\lfloor nt \rfloor}^2 \\ &\quad + \frac{\mu(nt - \lfloor nt \rfloor)^2}{n^3} U_{\lfloor nt \rfloor} + \frac{\lfloor nt \rfloor + (nt - \lfloor nt \rfloor)^3}{3n^3} \mu^2. \end{aligned}$$

Since

$$(6.8) \qquad X_{k-1} = \frac{1}{1+\beta}(U_k - V_k), \qquad X_k = \frac{1}{1+\beta}(U_k + \beta V_k), \qquad k \in \mathbb{N},$$

26

using Lemma 9.1, we obtain

$$\sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}\,\mathbb{E}(M_k^2\,|\,\mathcal{F}_{k-1}) = \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}\big[\alpha\beta(X_{k-1}+X_{k-2})+\sigma^2\big]$$

(6.9)
$$= \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}\left[\frac{\alpha\beta}{1+\beta}(2U_{k-1}-\alpha V_{k-1})+\sigma^2\right]$$

$$= \frac{2\alpha\beta}{1+\beta}\sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^2 - \frac{\alpha^2\beta}{1+\beta}\sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}V_{k-1} + \sigma^2\sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}.$$

Thus, in order to show (6.2), it suffices to prove

(6.10)
$$n^{-3}\sum_{k=1}^{\lfloor nT \rfloor} |U_k V_k| \xrightarrow{\mathbb{P}} 0,$$

(6.11)
$$n^{-3}\sum_{k=1}^{\lfloor nT \rfloor} U_k \xrightarrow{\mathbb{P}} 0,$$

(6.12)
$$n^{-3/2}\sup_{t\in[0,T]} U_{\lfloor nt \rfloor} \xrightarrow{\mathbb{P}} 0,$$

(6.13)
$$n^{-3}\sup_{t\in[0,T]} \big[\lfloor nt \rfloor + (nt-\lfloor nt \rfloor)^3\big] \to 0$$

as $n\to\infty$. Using (9.14) with $(\ell,i,j)=(8,1,1)$ and $(\ell,i,j)=(8,1,0)$, we have (6.10) and (6.11), respectively. Using (9.15) with $(\ell,i,j)=(8,1,0)$, we have (6.12). Clearly, (6.13) follows from $|nt-\lfloor nt \rfloor|\leqslant 1$, $n\in\mathbb{N}$, $t\in\mathbb{R}_+$, thus we conclude (6.2).

Now we turn to check (6.3). Again by (6.7), we have

$$\int_0^t (\mathcal{M}_s^{(n)}+\mu s)^3\,\mathrm{d}s = \frac{1}{n^4}\sum_{k=1}^{\lfloor nt \rfloor-1} U_k^3 + \frac{3\mu}{2n^4}\sum_{k=1}^{\lfloor nt \rfloor-1} U_k^2 + \frac{\mu^2}{n^4}\sum_{k=1}^{\lfloor nt \rfloor-1} U_k + \frac{nt-\lfloor nt \rfloor}{n^4}U_{\lfloor nt \rfloor}^3$$

$$+ \frac{3\mu(nt-\lfloor nt \rfloor)^2}{2n^4}U_{\lfloor nt \rfloor}^2 + \frac{\mu^2(nt-\lfloor nt \rfloor)^3}{n^4}U_{\lfloor nt \rfloor} + \frac{\lfloor nt \rfloor+(nt-\lfloor nt \rfloor)^4}{4n^4}\mu^3.$$

Using Lemma 9.1, we obtain

$$\sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^2\,\mathbb{E}(M_k^2\,|\,\mathcal{F}_{k-1}) = \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^2\big[\alpha\beta(X_{k-1}+X_{k-2})+\sigma^2\big]$$

$$= \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^2\left[\frac{\alpha\beta}{1+\beta}(2U_{k-1}-\alpha V_{k-1})+\sigma^2\right]$$

$$= \frac{2\alpha\beta}{1+\beta}\sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^3 - \frac{\alpha^2\beta}{1+\beta}\sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^2 V_{k-1} + \sigma^2\sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^2.$$

Thus, in order to show (6.3), it suffices to prove

(6.14)
$$n^{-4} \sum_{k=1}^{\lfloor nT \rfloor} |U_k^2 V_k| \xrightarrow{\mathbb{P}} 0,$$

(6.15)
$$n^{-4} \sum_{k=1}^{\lfloor nT \rfloor} U_k^2 \xrightarrow{\mathbb{P}} 0,$$

(6.16)
$$n^{-4} \sum_{k=1}^{\lfloor nT \rfloor} U_k \xrightarrow{\mathbb{P}} 0,$$

(6.17)
$$n^{-4/3} \sup_{t \in [0,T]} U_{\lfloor nt \rfloor} \xrightarrow{\mathbb{P}} 0,$$

(6.18)
$$n^{-4} \sup_{t \in [0,T]} \left[ \lfloor nt \rfloor + (nt - \lfloor nt \rfloor)^4 \right] \to 0$$

as $n \to \infty$. Using (9.14) with $(\ell, i, j) = (8, 2, 1)$, $(\ell, i, j) = (8, 2, 0)$ and $(\ell, i, j) = (8, 1, 0)$, we have (6.14), (6.15) and (6.16), respectively. Using (9.15) with $(\ell, i, j) = (8, 1, 0)$, we have (6.17). Clearly, (6.18) follows again from $|nt - \lfloor nt \rfloor| \leqslant 1$, $n \in \mathbb{N}$, $t \in \mathbb{R}_+$, thus we conclude (6.3).

Next we turn to prove (6.4). By (6.9), (6.10) and (6.11) we get

(6.19)
$$n^{-3} \sup_{t \in [0,T]} \left| \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} \mathbb{E}(M_k^2 \mid \mathcal{F}_{k-1}) - \frac{2\alpha\beta}{1+\beta} \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^2 \right| \xrightarrow{\mathbb{P}} 0 \qquad \text{as } n \to \infty$$

for all $T > 0$. Using (6.2), in order to prove (6.4), it is sufficient to show that

(6.20)
$$n^{-3} \sup_{t \in [0,T]} \left| \sum_{k=1}^{\lfloor nt \rfloor} V_{k-1}^2 \mathbb{E}(M_k^2 \mid \mathcal{F}_{k-1}) - \frac{4\alpha\beta^2}{(1+\beta)^3} \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^2 \right| \xrightarrow{\mathbb{P}} 0 \qquad \text{as } n \to \infty$$

for all $T > 0$. As in the previous case, using Lemma 9.1 and (6.8), we obtain

$$\sum_{k=1}^{\lfloor nt \rfloor} V_{k-1}^2 \mathbb{E}(M_k^2 \mid \mathcal{F}_{k-1}) = \sum_{k=1}^{\lfloor nt \rfloor} V_{k-1}^2 [\alpha\beta(X_{k-1} + X_{k-2}) + \sigma^2]$$

$$= \sum_{k=1}^{\lfloor nt \rfloor} V_{k-1}^2 \left[ \frac{\alpha\beta}{1+\beta}(2U_{k-1} - \alpha V_{k-1}) + \sigma^2 \right]$$

$$= \frac{2\alpha\beta}{1+\beta} \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} V_{k-1}^2 - \frac{\alpha^2\beta}{1+\beta} \sum_{k=1}^{\lfloor nt \rfloor} V_{k-1}^3 + \sigma^2 \sum_{k=1}^{\lfloor nt \rfloor} V_{k-1}^2.$$

Using (9.14) with $(\ell, i, j) = (8, 0, 3)$ and $(\ell, i, j) = (8, 0, 2)$, we have

$$n^{-3} \sum_{k=1}^{\lfloor nT \rfloor} |V_k|^3 \xrightarrow{\mathbb{P}} 0, \qquad n^{-3} \sum_{k=1}^{\lfloor nT \rfloor} V_k^2 \xrightarrow{\mathbb{P}} 0 \qquad \text{as } n \to \infty,$$

28

hence (6.20) will follow from

$$(6.21) \qquad n^{-3} \sup_{t\in[0,T]} \left| \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} V_{k-1}^2 - \frac{2\beta}{(1+\beta)^2} \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^2 \right| \xrightarrow{\mathbb{P}} 0 \qquad \text{as } n \to \infty$$

for all $T > 0$.

The aim of the following discussion is to decompose $\sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} V_{k-1}^2 - 2\beta(1+\beta)^{-2} \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^2$ as a sum of a martingale and some negligible terms. Using recursions (3.20), (5.8) and Lemma 9.1 (formula (9.1)), we obtain

$$\mathbb{E}(U_{k-1} V_{k-1}^2 \,|\, \mathcal{F}_{k-2}) = \mathbb{E}\left( (U_{k-2} + M_{k-1} + \mu)(-\beta V_{k-2} + M_{k-1} + \mu)^2 \,|\, \mathcal{F}_{k-2} \right)$$

$$= \beta^2 U_{k-2} V_{k-2}^2 + \alpha\beta(X_{k-2} + X_{k-3})(U_{k-2} - 2\beta V_{k-2} + 3\mu) + \mathbb{E}(M_{k-1}^3 \,|\, \mathcal{F}_{k-2})$$

$$+ \text{constant} + \text{linear combination of } U_{k-2} V_{k-2}, \ V_{k-2}^2, \ U_{k-2} \text{ and } V_{k-2}.$$

Using again Lemma 9.1 (formula (9.3)) and (6.8), we get

$$\mathbb{E}(U_{k-1} V_{k-1}^2 \,|\, \mathcal{F}_{k-2})$$

$$= \beta^2 U_{k-2} V_{k-2}^2 + \frac{\alpha\beta}{1+\beta}(2U_{k-2} - \alpha V_{k-2})(U_{k-2} - 2\beta V_{k-2} + 3\mu) + \mathbb{E}(M_{k-1}^3 \,|\, \mathcal{F}_{k-2})$$

$$(6.22) \qquad + \text{constant} + \text{linear combination of } U_{k-2} V_{k-2}, \ V_{k-2}^2, \ U_{k-2} \text{ and } V_{k-2}$$

$$= \beta^2 U_{k-2} V_{k-2}^2 + \frac{2\alpha\beta}{1+\beta} U_{k-2}^2 + \text{constant}$$

$$+ \text{linear combination of } U_{k-2} V_{k-2}, \ V_{k-2}^2, \ U_{k-2} \text{ and } V_{k-2}.$$

Thus

$$\sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} V_{k-1}^2 = \sum_{k=2}^{\lfloor nt \rfloor} \left[ U_{k-1} V_{k-1}^2 - \mathbb{E}(U_{k-1} V_{k-1}^2 \,|\, \mathcal{F}_{k-2}) \right] + \sum_{k=2}^{\lfloor nt \rfloor} \mathbb{E}(U_{k-1} V_{k-1}^2 \,|\, \mathcal{F}_{k-2})$$

$$= \sum_{k=2}^{\lfloor nt \rfloor} \left[ U_{k-1} V_{k-1}^2 - \mathbb{E}(U_{k-1} V_{k-1}^2 \,|\, \mathcal{F}_{k-2}) \right] + \beta^2 \sum_{k=2}^{\lfloor nt \rfloor} U_{k-2} V_{k-2}^2 + \frac{2\alpha\beta}{1+\beta} \sum_{k=2}^{\lfloor nt \rfloor} U_{k-2}^2$$

$$+ \mathrm{O}(n) + \text{linear combination of } \sum_{k=2}^{\lfloor nt \rfloor} U_{k-2} V_{k-2}, \ \sum_{k=2}^{\lfloor nt \rfloor} V_{k-2}^2, \ \sum_{k=2}^{\lfloor nt \rfloor} U_{k-2} \text{ and } \sum_{k=2}^{\lfloor nt \rfloor} V_{k-2}.$$

Consequently,

$$\sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}V_{k-1}^2 = \frac{1}{1-\beta^2} \sum_{k=2}^{\lfloor nt \rfloor} \left[ U_{k-1}V_{k-1}^2 - \mathbb{E}(U_{k-1}V_{k-1}^2 \mid \mathcal{F}_{k-2}) \right] + \frac{2\alpha\beta}{(1+\beta)(1-\beta^2)} \sum_{k=2}^{\lfloor nt \rfloor} U_{k-2}^2$$

$$- \frac{\beta^2}{1-\beta^2} U_{\lfloor nt \rfloor -1} V_{\lfloor nt \rfloor -1}^2 + \mathrm{O}(n)$$

$$+ \text{ linear combination of } \sum_{k=2}^{\lfloor nt \rfloor} U_{k-2}V_{k-2}, \ \sum_{k=2}^{\lfloor nt \rfloor} V_{k-2}^2, \ \sum_{k=2}^{\lfloor nt \rfloor} U_{k-2} \ \text{ and } \ \sum_{k=2}^{\lfloor nt \rfloor} V_{k-2}.$$

Using (9.16) with $(\ell, i, j) = (8, 1, 2)$ we have

$$n^{-3} \sup_{t\in[0,T]} \left| \sum_{k=2}^{\lfloor nt \rfloor} \left[ U_{k-1}V_{k-1}^2 - \mathbb{E}(U_{k-1}V_{k-1}^2 \mid \mathcal{F}_{k-2}) \right] \right| \xrightarrow{\mathbb{P}} 0 \qquad \text{as } n \to \infty.$$

Thus, in order to show (6.21), it suffices to prove

(6.23) $$n^{-3} \sum_{k=1}^{\lfloor nT \rfloor} |U_k V_k| \xrightarrow{\mathbb{P}} 0,$$

(6.24) $$n^{-3} \sum_{k=1}^{\lfloor nT \rfloor} V_k^2 \xrightarrow{\mathbb{P}} 0,$$

(6.25) $$n^{-3} \sum_{k=1}^{\lfloor nT \rfloor} U_k \xrightarrow{\mathbb{P}} 0,$$

(6.26) $$n^{-3} \sum_{k=1}^{\lfloor nT \rfloor} |V_k| \xrightarrow{\mathbb{P}} 0,$$

(6.27) $$n^{-3} \sup_{t\in[0,T]} U_{\lfloor nt \rfloor} V_{\lfloor nt \rfloor}^2 \xrightarrow{\mathbb{P}} 0,$$

(6.28) $$n^{-3} \sup_{t\in[0,T]} U_{\lfloor nt \rfloor}^2 \xrightarrow{\mathbb{P}} 0$$

as $n \to \infty$. Using (9.14) with $(\ell, i, j) = (8, 1, 1)$, $(\ell, i, j) = (8, 0, 2)$, $(\ell, i, j) = (8, 1, 0)$, and $(\ell, i, j) = (8, 0, 1)$, we have (6.23), (6.24), (6.25) and (6.26). Using (9.15) with $(\ell, i, j) = (8, 1, 2)$ and $(\ell, i, j) = (8, 2, 0)$, we have (6.27) and (6.28). Thus we conclude (6.4).

30

For (6.5), consider

$$\sum_{k=1}^{\lfloor nt \rfloor} V_{k-1}\, \mathbb{E}(M_k^2 \mid \mathcal{F}_{k-1}) = \sum_{k=1}^{\lfloor nt \rfloor} V_{k-1}\big(\alpha\beta(X_{k-1} + X_{k-2}) + \sigma^2\big)$$

$$= \sum_{k=1}^{\lfloor nt \rfloor} V_{k-1}\left(\frac{\alpha\beta}{1+\beta}(2U_{k-1} - \alpha V_{k-1}) + \sigma^2\right)$$

$$= \frac{2\alpha\beta}{1+\beta}\sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}V_{k-1} - \frac{\alpha^2\beta}{1+\beta}\sum_{k=1}^{\lfloor nt \rfloor} V_{k-1}^2 + \sigma^2\sum_{k=1}^{\lfloor nt \rfloor} V_{k-1},$$

where we used Lemma 9.1 and (6.8). Using (9.14) with $(\ell, i, j) = (8, 0, 2)$, and $(\ell, i, j) = (8, 0, 1)$, we have

$$n^{-5/2}\sum_{k=1}^{\lfloor nT \rfloor} V_k^2 \xrightarrow{\mathbb{P}} 0, \qquad n^{-5/2}\sum_{k=1}^{\lfloor nT \rfloor} |V_k| \xrightarrow{\mathbb{P}} 0 \qquad \text{as } n \to \infty,$$

hence (6.5) will follow from

(6.29) 
$$n^{-5/2}\sup_{t\in[0,T]}\left|\sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}V_{k-1}\right| \xrightarrow{\mathbb{P}} 0.$$

The aim of the following discussion is to decompose $\sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}V_{k-1}$ as a sum of a martingale and some negligible terms. Using the recursions (3.20), (5.8) and Lemma 9.1, we obtain

$$\mathbb{E}(U_{k-1}V_{k-1} \mid \mathcal{F}_{k-2}) = \mathbb{E}\big((U_{k-2} + M_{k-1} + \mu)(-\beta V_{k-2} + M_{k-1} + \mu) \mid \mathcal{F}_{k-2}\big)$$

$$= -\beta U_{k-2}V_{k-2} + \mu U_{k-2} - \beta\mu V_{k-2} + \mu^2 + \mathbb{E}(M_{k-1}^2 \mid \mathcal{F}_{k-2})$$

$$= -\beta U_{k-2}V_{k-2} + \text{constant} + \text{linear combination of } U_{k-2} \text{ and } V_{k-2}.$$

Thus

$$\sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}V_{k-1} = \sum_{k=2}^{\lfloor nt \rfloor} \big[U_{k-1}V_{k-1} - \mathbb{E}(U_{k-1}V_{k-1} \mid \mathcal{F}_{k-2})\big] + \sum_{k=2}^{\lfloor nt \rfloor} \mathbb{E}(U_{k-1}V_{k-1} \mid \mathcal{F}_{k-2})$$

$$= \sum_{k=2}^{\lfloor nt \rfloor} \big[U_{k-1}V_{k-1} - \mathbb{E}(U_{k-1}V_{k-1} \mid \mathcal{F}_{k-2})\big] - \beta\sum_{k=2}^{\lfloor nt \rfloor} U_{k-2}V_{k-2}$$

$$+ O(n) + \text{linear combination of } \sum_{k=2}^{\lfloor nt \rfloor} U_{k-2} \text{ and } \sum_{k=2}^{\lfloor nt \rfloor} V_{k-2}.$$

31

Consequently

$$\sum_{k=2}^{\lfloor nt \rfloor} U_{k-1} V_{k-1} = \frac{1}{1+\beta} \sum_{k=2}^{\lfloor nt \rfloor} \left[ U_{k-1} V_{k-1} - \mathbb{E}(U_{k-1} V_{k-1} \mid \mathcal{F}_{k-2}) \right] + \frac{\beta}{1+\beta} U_{\lfloor nt \rfloor - 1} V_{\lfloor nt \rfloor - 1}$$

$$+ \mathrm{O}(n) + \text{linear combination of } \sum_{k=2}^{\lfloor nt \rfloor} U_{k-2} \text{ and } \sum_{k=2}^{\lfloor nt \rfloor} V_{k-2}.$$

Using (9.16) with $(\ell, i, j) = (8, 1, 1)$ we have

$$n^{-5/2} \sup_{t \in [0,T]} \left| \sum_{k=2}^{\lfloor nt \rfloor} \left[ U_{k-1} V_{k-1} - \mathbb{E}(U_{k-1} V_{k-1} \mid \mathcal{F}_{k-2}) \right] \right| \xrightarrow{\mathbb{P}} 0 \qquad \text{as } n \to \infty.$$

Thus, in order to show (6.29), it suffices to prove

$$(6.30) \qquad\qquad n^{-5/2} \sum_{k=1}^{\lfloor nT \rfloor} U_k \xrightarrow{\mathbb{P}} 0,$$

$$(6.31) \qquad\qquad n^{-5/2} \sum_{k=1}^{\lfloor nT \rfloor} |V_k| \xrightarrow{\mathbb{P}} 0,$$

$$(6.32) \qquad\qquad n^{-5/2} \sup_{t \in [0,T]} |U_{\lfloor nt \rfloor} V_{\lfloor nt \rfloor}| \xrightarrow{\mathbb{P}} 0$$

as $n \to \infty$. Using (9.14) with $(\ell, i, j) = (8, 1, 0)$, and $(\ell, i, j) = (8, 0, 1)$, we have (6.30) and (6.31). Using (9.15) with $(\ell, i, j) = (8, 1, 1)$ we have (6.32), thus we conclude (6.5).

Convergence (6.6) can be handled in the same way as (6.5). For completeness we present all of the details. By Lemma 9.1 and (6.8), we have

$$\sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} V_{k-1} \mathbb{E}(M_k^2 \mid \mathcal{F}_{k-1}) = \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} V_{k-1} \left( \alpha \beta (X_{k-1} + X_{k-2}) + \sigma^2 \right)$$

$$= \frac{2\alpha\beta}{1+\beta} \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^2 V_{k-1} - \frac{\alpha^2 \beta}{1+\beta} \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} V_{k-1}^2 + \sigma^2 \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} V_{k-1}.$$

Using (9.14) with $(\ell, i, j) = (8, 1, 2)$, and $(\ell, i, j) = (8, 1, 1)$, we have

$$n^{-7/2} \sum_{k=1}^{\lfloor nT \rfloor} U_{k-1} V_{k-1}^2 \xrightarrow{\mathbb{P}} 0, \qquad n^{-7/2} \sum_{k=1}^{\lfloor nT \rfloor} U_{k-1} |V_{k-1}| \xrightarrow{\mathbb{P}} 0 \qquad \text{as } n \to \infty,$$

hence (6.6) will follow from

$$(6.33) \qquad\qquad n^{-7/2} \sup_{t \in [0,T]} \left| \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^2 V_{k-1} \right| \xrightarrow{\mathbb{P}} 0 \qquad \text{as } n \to \infty.$$

32

The aim of the following discussion is to decompose $\sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^2 V_{k-1}$ as a sum of a martingale and some negligible terms. Using the recursions (3.20) and (5.8), we obtain

$$\mathbb{E}(U_{k-1}^2 V_{k-1} \mid \mathcal{F}_{k-2}) = \mathbb{E}\left((U_{k-2} + M_{k-1} + \mu)^2(-\beta V_{k-2} + M_{k-1} + \mu) \mid \mathcal{F}_{k-2}\right)$$

$$= -\beta U_{k-2}^2 V_{k-2} + \mu U_{k-2}^2 - \beta \mu^2 V_{k-2} - 2\beta \mu U_{k-2} V_{k-2} + 2\mu^2 U_{k-2}$$

$$+ (2U_{k-2} - \beta V_{k-2} + 3\mu)\,\mathbb{E}(M_{k-1}^2 \mid \mathcal{F}_{k-2}) + \mathbb{E}(M_{k-1}^3 \mid \mathcal{F}_{k-2}) + \mu^3.$$

Hence, by Lemma 9.1 and (6.8),

$$\mathbb{E}(U_{k-1}^2 V_{k-1} \mid \mathcal{F}_{k-2}) = -\beta U_{k-2}^2 V_{k-2} + \text{constant}$$

$$+ \text{linear combination of } U_{k-2},\, V_{k-2},\, U_{k-2}^2,\, V_{k-2}^2 \text{ and } U_{k-2} V_{k-2}.$$

Thus

$$\sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^2 V_{k-1} = \sum_{k=2}^{\lfloor nt \rfloor} \left[U_{k-1}^2 V_{k-1} - \mathbb{E}(U_{k-1}^2 V_{k-1} \mid \mathcal{F}_{k-2})\right] + \sum_{k=2}^{\lfloor nt \rfloor} \mathbb{E}(U_{k-1}^2 V_{k-1} \mid \mathcal{F}_{k-2})$$

$$= \sum_{k=2}^{\lfloor nt \rfloor} \left[U_{k-1}^2 V_{k-1} - \mathbb{E}(U_{k-1}^2 V_{k-1} \mid \mathcal{F}_{k-2})\right] - \beta \sum_{k=2}^{\lfloor nt \rfloor} U_{k-2}^2 V_{k-2} + \mathrm{O}(n)$$

$$+ \text{linear combination of } \sum_{k=1}^{\lfloor nt \rfloor} U_{k-2},\, \sum_{k=1}^{\lfloor nt \rfloor} V_{k-2},\, \sum_{k=1}^{\lfloor nt \rfloor} U_{k-2}^2,\, \sum_{k=1}^{\lfloor nt \rfloor} V_{k-2}^2 \text{ and } \sum_{k=1}^{\lfloor nt \rfloor} U_{k-2} V_{k-2}.$$

Consequently

$$\sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^2 V_{k-1} = \frac{1}{1+\beta} \sum_{k=2}^{\lfloor nt \rfloor} \left[U_{k-1}^2 V_{k-1} - \mathbb{E}(U_{k-1}^2 V_{k-1} \mid \mathcal{F}_{k-2})\right] + \frac{\beta}{1+\beta} U_{\lfloor nt \rfloor -1}^2 V_{\lfloor nt \rfloor -1}$$

$$+ \mathrm{O}(n) + \text{linear combination of } \sum_{k=1}^{\lfloor nt \rfloor} U_{k-2},\, \sum_{k=1}^{\lfloor nt \rfloor} V_{k-2},\, \sum_{k=1}^{\lfloor nt \rfloor} U_{k-2}^2,\, \sum_{k=1}^{\lfloor nt \rfloor} V_{k-2}^2 \text{ and } \sum_{k=1}^{\lfloor nt \rfloor} U_{k-2} V_{k-2}.$$

Using (9.16) with $(\ell, i, j) = (8, 2, 1)$ we have

$$n^{-7/2} \sup_{t \in [0,T]} \left| \sum_{k=2}^{\lfloor nt \rfloor} \left[U_{k-1}^2 V_{k-1} - \mathbb{E}(U_{k-1}^2 V_{k-1} \mid \mathcal{F}_{k-2})\right] \right| \xrightarrow{\mathbb{P}} 0 \qquad \text{as } n \to \infty.$$

Thus, in order to show (6.33), it suffices to prove

(6.34)
$$n^{-7/2} \sum_{k=1}^{\lfloor nT \rfloor} U_k \xrightarrow{\mathbb{P}} 0,$$

(6.35)
$$n^{-7/2} \sum_{k=1}^{\lfloor nT \rfloor} U_k^2 \xrightarrow{\mathbb{P}} 0,$$

33

$$(6.36) \qquad n^{-7/2} \sum_{k=1}^{\lfloor nT \rfloor} |V_k| \xrightarrow{\mathbb{P}} 0,$$

$$(6.37) \qquad n^{-7/2} \sum_{k=1}^{\lfloor nT \rfloor} V_k^2 \xrightarrow{\mathbb{P}} 0,$$

$$(6.38) \qquad n^{-7/2} \sum_{k=1}^{\lfloor nT \rfloor} |U_k V_k| \xrightarrow{\mathbb{P}} 0,$$

$$(6.39) \qquad n^{-7/2} \sup_{t \in [0,T]} |U_{\lfloor nt \rfloor}^2 V_{\lfloor nt \rfloor}| \xrightarrow{\mathbb{P}} 0$$

as $n \to \infty$. Here (6.34), (6.35), (6.36), (6.37) and (6.38) follow by (9.14), and (6.39) by (9.15), thus we conclude (6.6).

Finally, we check condition (ii) of Theorem C.1, i.e., the conditional Lindeberg condition

$$(6.40) \qquad \sum_{k=1}^{\lfloor nT \rfloor} \mathbb{E} \left( \|\boldsymbol{Z}_k^{(n)}\|^2 \mathbb{1}_{\{\|\boldsymbol{Z}_k^{(n)}\| > \theta\}} \,\big|\, \mathcal{F}_{k-1} \right) \xrightarrow{\mathbb{P}} 0 \qquad \text{as } n \to \infty \text{ for all } \theta > 0 \text{ and } T > 0.$$

We have $\mathbb{E} \left( \|\boldsymbol{Z}_k^{(n)}\|^2 \mathbb{1}_{\{\|\boldsymbol{Z}_k^{(n)}\| > \theta\}} \,\big|\, \mathcal{F}_{k-1} \right) \leqslant \theta^{-2} \mathbb{E} \left( \|\boldsymbol{Z}_k^{(n)}\|^4 \,\big|\, \mathcal{F}_{k-1} \right)$ and

$$\|\boldsymbol{Z}_k^{(n)}\|^4 \leqslant 3 \left( n^{-4} M_k^4 + n^{-8} M_k^4 U_{k-1}^4 + n^{-6} M_k^4 V_{k-1}^4 \right).$$

Hence

$$\sum_{k=1}^{\lfloor nT \rfloor} \mathbb{E} \left( \|\boldsymbol{Z}_k^{(n)}\|^2 \mathbb{1}_{\{\|\boldsymbol{Z}_k^{(n)}\| > \theta\}} \right) \to 0 \qquad \text{as } n \to \infty \text{ for all } \theta > 0 \text{ and } T > 0,$$

since $\mathbb{E}(M_k^4) = \mathrm{O}(k^2)$, $\mathbb{E}(M_k^4 U_{k-1}^4) \leqslant \sqrt{\mathbb{E}(M_k^8) \mathbb{E}(U_{k-1}^8)} = \mathrm{O}(k^6)$ and $\mathbb{E}(M_k^4 V_{k-1}^4) \leqslant \sqrt{\mathbb{E}(M_k^8) \mathbb{E}(V_{k-1}^8)} = \mathrm{O}(k^4)$ by Corollary 9.1. Here we call the attention that our eight order moment condition $\mathbb{E}(\varepsilon_1^8) < \infty$ is used for applying Corollary 9.1. This yields (6.40).

# 7 Proof of Theorem 4.2

We have

$$(\widetilde{\boldsymbol{A}}_n, \widetilde{\boldsymbol{d}}_n) = \left( \sum_{k=1}^{n} \begin{bmatrix} n^{-3} X_{k-1}^2 & -n^{-2} X_{k-1} V_{k-1} & n^{-2} X_{k-1} \\ -n^{-2} X_{k-1} V_{k-1} & n^{-1} V_{k-1}^2 & -n^{-1} V_{k-1} \\ n^{-2} X_{k-1} & -n^{-1} V_{k-1} & n^{-1} \end{bmatrix}, \sum_{k=1}^{n} \begin{bmatrix} n^{-3/2} M_k X_{k-1} \\ -n^{-1/2} M_k V_{k-1} \\ n^{-1/2} M_k \end{bmatrix} \right).$$

Theorem 4.2 will follow from the following statement (using also Slutsky's lemma).

**7.1 Theorem.** *Under the assumptions of Theorem 2.2 we have*

$$n^{-2} \sum_{k=1}^{n} X_{k-1} \xrightarrow{\text{a.s.}} \frac{\mu}{2}, \qquad n^{-1} \sum_{k=1}^{n} V_{k-1} \xrightarrow{\text{a.s.}} \mu, \qquad n^{-3} \sum_{k=1}^{n} X_{k-1}^2 \xrightarrow{\text{a.s.}} \frac{\mu^2}{3},$$

$$n^{-2} \sum_{k=1}^{n} X_{k-1} V_{k-1} \xrightarrow{\text{a.s.}} \frac{\mu^2}{2}, \qquad n^{-1} \sum_{k=1}^{n} V_{k-1}^2 \xrightarrow{\text{a.s.}} \mu^2 + \sigma^2,$$

*and*

$$\sum_{k=1}^{n} \begin{bmatrix} n^{-3/2} M_k X_{k-1} \\ -n^{-1/2} M_k V_{k-1} \\ n^{-1/2} M_k \end{bmatrix} \xrightarrow{\mathcal{L}} \mathcal{N}_3 \left( \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \sigma^2 \begin{bmatrix} \frac{1}{3}\mu^2 & -\frac{1}{2}\mu^2 & \frac{1}{2}\mu \\ -\frac{1}{2}\mu^2 & \mu^2 + \sigma^2 & -\mu \\ \frac{1}{2}\mu & -\mu & 1 \end{bmatrix} \right).$$

**Proof.** In this case equation (1.1) has the form $X_k = X_{k-1} + \varepsilon_k$, $k \in \mathbb{N}$, and hence $X_k = \varepsilon_1 + \cdots + \varepsilon_k$, $M_k = X_k - X_{k-1} - \mu = \varepsilon_k - \mu$ and $V_k = X_k - X_{k-1} = \varepsilon_k$, $k \in \mathbb{N}$.

The first statement follows from (3.12) by Toeplitz theorem, where we used that

$$\lim_{n \to \infty} \sum_{k=1}^{n} \frac{k}{n^2} = \frac{1}{2}.$$

Again by (3.12),

$$n^{-1} \sum_{k=1}^{n} V_k = n^{-1} X_k \xrightarrow{\text{a.s.}} \mu \qquad \text{as } n \to \infty.$$

We have already shown the third statement, see (3.13). By the strong law of large numbers we have

(7.1) $$n^{-1} \sum_{k=1}^{n} V_k^2 = n^{-1} \sum_{k=1}^{n} \varepsilon_k^2 \xrightarrow{\text{a.s.}} \sigma^2 + \mu^2 \qquad \text{as } n \to \infty.$$

Moreover,

$$\sum_{k=1}^{n} X_{k-1} V_{k-1} = \sum_{k=1}^{n} X_{k-1} \varepsilon_{k-1} = \sum_{k=1}^{n} \varepsilon_{k-1}^2 + \sum_{k=1}^{n} X_{k-2} \varepsilon_{k-1} = \sum_{k=1}^{n} \varepsilon_{k-1}^2 + \sum_{k=1}^{n} \varepsilon_{k-1} \sum_{i=1}^{k-2} \varepsilon_i$$

$$= \sum_{k=1}^{n} \varepsilon_{k-1}^2 + \sum_{1 \leqslant i < j \leqslant n-1} \varepsilon_i \varepsilon_j = \frac{1}{2} \left( \left( \sum_{k=1}^{n} \varepsilon_{k-1} \right)^2 + \sum_{k=1}^{n} \varepsilon_{k-1}^2 \right)$$

with $\varepsilon_0 := 0$, and hence by (3.12) and (7.1),

(7.2) $$n^{-2} \sum_{k=1}^{n} X_{k-1} V_{k-1} \xrightarrow{\text{a.s.}} \frac{\mu^2}{2}.$$

The last statement can be proved by the multidimensional martingale central limit theorem (see, e.g., Jacod and Shiryaev [20, Chapter VIII, Theorem 3.33]) for the sequence $(\boldsymbol{Y}_k^{(n)}, \mathcal{F}_k)_{k \in \mathbb{N}}$,

35

$n \in \mathbb{N}$, of square-integrable martingale differences given by

$$
\boldsymbol{Y}_k^{(n)} := \begin{bmatrix} n^{-3/2} M_k X_{k-1} \\ -n^{-1/2} M_k V_{k-1} \\ n^{-1/2} M_k \end{bmatrix} = \begin{bmatrix} n^{-3/2} (\varepsilon_k - \mu) X_{k-1} \\ -n^{-1/2} (\varepsilon_k - \mu) \varepsilon_{k-1} \\ n^{-1/2} (\varepsilon_k - \mu) \end{bmatrix}, \qquad n, k \in \mathbb{N}.
$$

We have

$$
\mathbb{E}(\boldsymbol{Y}_k^{(n)} (\boldsymbol{Y}_k^{(n)})^\top \mid \mathcal{F}_{k-1}) = \sigma^2 \begin{bmatrix} n^{-3} X_{k-1}^2 & -n^{-2} X_{k-1}\varepsilon_{k-1} & n^{-2} X_{k-1} \\ -n^{-2} X_{k-1}\varepsilon_{k-1} & n^{-1}\varepsilon_{k-1}^2 & -n^{-1}\varepsilon_{k-1} \\ n^{-2} X_{k-1} & -n^{-1}\varepsilon_{k-1} & n^{-1} \end{bmatrix}, \qquad n, k \in \mathbb{N},
$$

hence by (3.13), (7.1), and (7.2), we have the asymptotic covariance matrix

$$
\sum_{k=1}^n \mathbb{E}(\boldsymbol{Y}_k^{(n)} (\boldsymbol{Y}_k^{(n)})^\top \mid \mathcal{F}_{k-1}) \xrightarrow{\text{a.s.}} \sigma^2 \begin{bmatrix} \frac{\mu^2}{3} & -\frac{\mu^2}{2} & \frac{\mu}{2} \\ -\frac{\mu^2}{2} & \sigma^2 + \mu^2 & -\mu \\ \frac{\mu}{2} & -\mu & 1 \end{bmatrix} \qquad \text{as } n \to \infty.
$$

The conditional Lindeberg condition

$$
\sum_{k=1}^n \mathbb{E}(\|\boldsymbol{Y}_k^{(n)}\|^2 \mathbb{1}_{\{\|\boldsymbol{Y}_k^{(n)}\| > \theta\}} \mid \mathcal{F}_{k-1}) \xrightarrow{\mathbb{P}} 0 \qquad \text{as } n \to \infty
$$

is satisfied for all $\theta > 0$, since using that $\mathbb{E}(\varepsilon_1^4) < \infty$,

$$
\sum_{k=1}^n \mathbb{E}(\|\boldsymbol{Y}_k^{(n)}\|^2 \mathbb{1}_{\{\|\boldsymbol{Y}_k^{(n)}\| > \theta\}} \mid \mathcal{F}_{k-1}) \leqslant \frac{1}{\theta^2} \sum_{k=1}^n \mathbb{E}(\|\boldsymbol{Y}_k^{(n)}\|^4 \mid \mathcal{F}_{k-1})
$$

$$
\leqslant \frac{3}{\theta^2} \sum_{k=1}^n \mathbb{E}\big(n^{-6}(\varepsilon_k - \mu)^4 X_{k-1}^4 + n^{-2}(\varepsilon_k - \mu)^4 (\varepsilon_{k-1}^4 + 1) \mid \mathcal{F}_{k-1}\big)
$$

$$
\leqslant \frac{3\,\mathbb{E}\big((\varepsilon_1 - \mu)^4\big)}{\theta^2} \sum_{k=1}^n \big(n^{-6} X_{k-1}^4 + n^{-2}(\varepsilon_{k-1}^4 + 1)\big) \xrightarrow{\mathbb{P}} 0
$$

as $n \to \infty$, where the last but one step follows by that $\varepsilon_k$ and $\mathcal{F}_{k-1}$ are independent, $\varepsilon_{k-1}$ is measurable with respect to the $\sigma$-algebra $\mathcal{F}_{k-1}$ (since $\varepsilon_{k-1} = X_{k-1} - X_{k-2}$), and the last step follows by $\mathbb{E}(X_k^4) = \mathrm{O}(k^4)$ (see Corollary 9.1). $\qquad \square$

## 8 Proof of Theorem 4.3

We have

$$
\widetilde{\boldsymbol{A}}_n = \sum_{k=1}^n \begin{bmatrix} n^{-3} X_{k-1}^2 & -n^{-5/2} X_{k-1} V_{k-1} & n^{-2} X_{k-1} \\ -n^{-5/2} X_{k-1} V_{k-1} & n^{-2} V_{k-1}^2 & -n^{-3/2} V_{k-1} \\ n^{-2} X_{k-1} & -n^{-3/2} V_{k-1} & n^{-1} \end{bmatrix}, \qquad \widetilde{\boldsymbol{d}}_n = \sum_{k=1}^n \begin{bmatrix} n^{-3/2} M_k X_{k-1} \\ -n^{-1} M_k V_{k-1} \\ n^{-1/2} M_k \end{bmatrix}.
$$

**8.1 Lemma.** *Under the assumptions of Theorem 2.3, as $n \to \infty$, we have*

$$\text{(8.1)} \qquad n^{-2} \sum_{k=1}^{n} X_{k-1} \xrightarrow{\text{a.s.}} \frac{\mu}{4},$$

$$\text{(8.2)} \qquad n^{-1} \sum_{k=1}^{n} V_{k-1} \xrightarrow{\text{a.s.}} \frac{\mu}{2},$$

$$\text{(8.3)} \qquad n^{-3} \sum_{k=1}^{n} X_{k-1}^{2} \xrightarrow{\text{a.s.}} \frac{\mu^{2}}{12},$$

$$\text{(8.4)} \qquad n^{-5/2} \sum_{k=1}^{n} X_{k-1} V_{k-1} \xrightarrow{\mathbb{P}} 0,$$

$$\text{(8.5)} \qquad n^{-2} \sum_{k=1}^{n} \left[ \mathbb{E}(V_{k-1}) \right]^{2} \to 0,$$

$$\text{(8.6)} \qquad n^{-2} \sum_{k=1}^{n} (V_{k-1} - \mathbb{E}(V_{k-1})) \, \mathbb{E}(V_{k-1}) \xrightarrow{\mathbb{P}} 0,$$

$$\text{(8.7)} \qquad n^{-3/2} \sum_{k=1}^{n} M_{k} \left( X_{k-1} - \mathbb{E}(X_{k-1}) \right) \xrightarrow{\mathbb{P}} 0,$$

$$\text{(8.8)} \qquad n^{-1} \sum_{k=1}^{n} M_{k} \, \mathbb{E}(V_{k-1}) \xrightarrow{\mathbb{P}} 0.$$

**Proof.** In this case equation (1.1) has the form $X_k = X_{k-2} + \varepsilon_k, \quad k \in \mathbb{N}, \quad$ and hence $X_{2k} = \varepsilon_2 + \varepsilon_4 + \cdots + \varepsilon_{2k}, \quad X_{2k-1} = \varepsilon_1 + \varepsilon_3 + \cdots + \varepsilon_{2k-1}, \quad M_k = X_k - X_{k-2} - \mu = \varepsilon_k - \mu \quad$ and $V_{2k} = X_{2k} - X_{2k-1} = (\varepsilon_2 - \varepsilon_1) + \cdots + (\varepsilon_{2k} - \varepsilon_{2k-1}), \quad V_{2k-1} = X_{2k-1} - X_{2k-2} = (\varepsilon_1 - \varepsilon_2) + \cdots + (\varepsilon_{2k-3} - \varepsilon_{2k-2}) + \varepsilon_{2k-1}, \quad k \in \mathbb{N}$.

Convergence (8.1) follows from (3.14) by Toeplitz theorem. Again by (3.14), we obtain

$$n^{-1} \sum_{k=1}^{n} V_{k-1} = n^{-1} X_{n-1} \xrightarrow{\text{a.s.}} \frac{\mu}{2},$$

yielding (8.2). We have already shown (8.3), see (3.15).

In order to show (8.4), we use (5.5). Clearly we have

$$\mathbb{E}\left( \sum_{k=1}^{n} X_k V_k \right) = \frac{1}{2} \mathbb{E}(X_n^2) + \frac{1}{2} \sum_{k=1}^{n} \mathbb{E}(V_k^2) = \mathrm{O}(n^2),$$

since, by Corollary 9.1, $\mathbb{E}(X_n^2) = \mathrm{O}(n^2), \quad n \in \mathbb{N}, \quad$ and $\mathbb{E}(V_k^2) = \mathrm{O}(k), \quad k \in \mathbb{N}, \quad$ and hence we obtain (8.4). For each $k \in \mathbb{N}$, we have $\mathbb{E}(V_{2k}) = 0 \quad$ and $\mathbb{E}(V_{2k-1}) = \mu, \quad$ hence we conclude

(8.5), and

$$\mathbb{E}\left(\left|\sum_{k=1}^{n}(V_{k-1} - \mathbb{E}(V_{k-1}))\,\mathbb{E}(V_{k-1})\right|\right) \leqslant \sum_{k=1}^{n} \mu\,\mathbb{E}\big(|V_{k-1} - \mathbb{E}(V_{k-1})|\big)$$

$$\leqslant \sum_{k=1}^{n} \mu\sqrt{\mathbb{E}\big((V_{k-1} - \mathbb{E}(V_{k-1}))^2\big)} = \mathrm{O}(n^{3/2}),$$

since $\mathbb{E}(V_k - \mathbb{E}(V_k))^2 \leqslant \mathbb{E}(V_k^2) = \mathrm{O}(k)$, $k \in \mathbb{N}$ (by Corollary 9.1), which implies (8.6).

Moreover, using that $M_k(X_{k-1} - \mathbb{E}(X_{k-1}))$, $k \in \{1, \dots, n\}$, are uncorrelated,

$$\mathbb{E}\left(\left(\sum_{k=1}^{n} M_k\big(X_{k-1} - \mathbb{E}(X_{k-1})\big)\right)^2\right) = \sum_{k=1}^{n} \mathbb{E}\left((\varepsilon_k - \mu)^2\big(X_{k-1} - \mathbb{E}(X_{k-1})\big)^2\right)$$

$$= \sigma^2 \sum_{k=1}^{n} \mathbb{E}\left(\big(X_{k-1} - \mathbb{E}(X_{k-1})\big)^2\right) = \mathrm{O}(n^2),$$

since $\mathbb{E}(X_{k-1} - \mathbb{E}(X_{k-1}))^2 \leqslant \lfloor k/2 \rfloor \sigma^2$, $k \in \mathbb{N}$ (by Corollary 9.1), thus we get (8.7).

Since $\mathbb{E}(V_{2k}) = 0$ and $\mathbb{E}(V_{2k-1}) = \mu$, $k \in \mathbb{N}$, we have

$$\mathbb{E}\left(\left(\sum_{k=1}^{n} M_k\,\mathbb{E}(V_{k-1})\right)^2\right) = \sum_{k=1}^{n} \big[\mathbb{E}(V_{k-1})\big]^2\,\mathbb{E}\big((\varepsilon_k - \mu)^2\big) = \sigma^2 \sum_{k=1}^{n} \big[\mathbb{E}(V_{k-1})\big]^2 = \mathrm{O}(n),$$

which implies (8.8). $\qquad\square$

Theorem 4.3 will follow from Lemma 8.1 and the following statement (using Slutsky's lemma).

**8.1 Theorem.** *Under the assumptions of Theorem 2.3 we have*

$$\sum_{k=1}^{n} \begin{bmatrix} n^{-2}\big(V_{k-1} - \mathbb{E}(V_{k-1})\big)^2 \\ n^{-3/2} M_k\,\mathbb{E}(X_{k-1}) \\ -n^{-1} M_k\big(V_{k-1} - \mathbb{E}(V_{k-1})\big) \\ n^{-1/2} M_k \end{bmatrix} \xrightarrow{\mathcal{L}} \begin{bmatrix} \sigma^2 \int_0^1 \mathcal{W}_t^2\,\mathrm{d}t \\ \tfrac{1}{2}\mu\sigma \int_0^1 t\,\mathrm{d}\widetilde{\mathcal{W}}_t \\ \sigma^2 \int_0^1 \mathcal{W}_t\,\mathrm{d}\mathcal{W}_t \\ \sigma\widetilde{\mathcal{W}}_1 \end{bmatrix} \qquad as \;\; n \to \infty,$$

*where* $(\mathcal{W}_t)_{t\in\mathbb{R}_+}$ *and* $(\widetilde{\mathcal{W}}_t)_{t\in\mathbb{R}_+}$ *are independent standard Wiener processes.*

**Proof.** Consider the sequence

$$\begin{bmatrix} \mathcal{S}_t^{(n)} \\ \mathcal{T}_t^{(n)} \end{bmatrix} := \begin{bmatrix} n^{-1/2}\big(X_{2\lfloor nt \rfloor} - \mathbb{E}(X_{2\lfloor nt \rfloor})\big) \\ n^{-1/2}\big(X_{2\lfloor nt \rfloor - 1} - \mathbb{E}(X_{2\lfloor nt \rfloor - 1})\big) \end{bmatrix}, \qquad t \in \mathbb{R}_+, \quad n \in \mathbb{N},$$

of stochastic processes. Then, by the multidimensional martingale central limit theorem,

(8.9) $$\begin{bmatrix} \mathcal{S}^{(n)} \\ \mathcal{T}^{(n)} \end{bmatrix} \xrightarrow{\mathcal{L}} \sigma \begin{bmatrix} \mathcal{B} \\ \widetilde{\mathcal{B}} \end{bmatrix} \qquad as \;\; n \to \infty,$$

where $(\mathcal{B}_t)_{t\in\mathbb{R}_+}$ and $(\widetilde{\mathcal{B}}_t)_{t\in\mathbb{R}_+}$ are independent standard Wiener processes. Indeed, with the notation

$$\boldsymbol{Y}_k^{(n)} := \begin{bmatrix} n^{-1/2}(\varepsilon_{2k} - \mu) \\ n^{-1/2}(\varepsilon_{2k-1} - \mu) \end{bmatrix}, \qquad n, k \in \mathbb{N},$$

we have that $(\boldsymbol{Y}_k^{(n)}, \mathcal{F}_{2k})_{k\in\mathbb{N}}$, $n \in \mathbb{N}$, is a sequence of square integrable martingale differences such that

$$\sum_{k=1}^{\lfloor nt \rfloor} \boldsymbol{Y}_k^{(n)} = \begin{bmatrix} \mathcal{S}_t^{(n)} \\ \mathcal{T}_t^{(n)} \end{bmatrix}, \qquad n \in \mathbb{N}, \quad t \in \mathbb{R}_+,$$

$\mathbb{E}\left(\boldsymbol{Y}_k^{(n)} \mid \mathcal{F}_{2(k-1)}\right) = \boldsymbol{0} \in \mathbb{R}^2$ and

$$\mathbb{E}\left(\boldsymbol{Y}_k^{(n)}(\boldsymbol{Y}_k^{(n)})^\top \mid \mathcal{F}_{2(k-1)}\right) = \sigma^2 n^{-1} \boldsymbol{I}_2, \qquad n, k \in \mathbb{N},$$

where $\boldsymbol{I}_2$ denotes the $2 \times 2$ identity matrix. Then the asymptotic covariance matrix

$$\sum_{k=1}^{\lfloor nt \rfloor} \mathbb{E}\left(\boldsymbol{Y}_k^{(n)}(\boldsymbol{Y}_k^{(n)})^\top \mid \mathcal{F}_{2(k-1)}\right) \xrightarrow{\text{a.s.}} \sigma^2 t \boldsymbol{I}_2 \qquad \text{as } n \to \infty \text{ for } t \in \mathbb{R}_+.$$

The conditional Lindeberg condition

$$(8.10) \qquad \sum_{k=1}^{\lfloor nt \rfloor} \mathbb{E}\left(\|\boldsymbol{Y}_k^{(n)}\|^2 \mathbb{1}_{\{\|\boldsymbol{Y}_k^{(n)}\|>\theta\}} \mid \mathcal{F}_{2(k-1)}\right) \xrightarrow{\mathbb{P}} 0 \qquad \text{as } n \to \infty$$

is satisfied for all $t \in \mathbb{R}_+$ and $\theta > 0$. Indeed, we have

$$\sum_{k=1}^{\lfloor nt \rfloor} \mathbb{E}\left(\|\boldsymbol{Y}_k^{(n)}\|^2 \mathbb{1}_{\{\|\boldsymbol{Y}_k^{(n)}\|>\theta\}}\right)$$

$$= \frac{1}{n} \sum_{k=1}^{\lfloor nt \rfloor} \mathbb{E}\left(\left((\varepsilon_{2k} - \mu)^2 + (\varepsilon_{2k-1} - \mu)^2\right) \mathbb{1}_{\{(\varepsilon_{2k}-\mu)^2+(\varepsilon_{2k-1}-\mu)^2>n\theta^2\}}\right)$$

$$= \frac{\lfloor nt \rfloor}{n} \mathbb{E}\left(\left((\varepsilon_2 - \mu)^2 + (\varepsilon_1 - \mu)^2\right) \mathbb{1}_{\{(\varepsilon_2-\mu)^2+(\varepsilon_1-\mu)^2>n\theta^2\}}\right) \to 0,$$

by dominated convergence theorem. This yields that the convergence in (8.10) holds in fact in $L_1$-sense. Thus we obtain (8.9). We are going to prove that convergence (8.9) implies

$$(8.11) \qquad \sum_{k=1}^{n} \begin{bmatrix} n^{-2}\left(V_{k-1} - \mathbb{E}(V_{k-1})\right)^2 \\ n^{-3/2} M_k \mathbb{E}(X_{k-1}) \\ -n^{-1} M_k\left(V_{k-1} - \mathbb{E}(V_{k-1})\right) \\ n^{-1/2} M_k \end{bmatrix} \xrightarrow{\mathcal{L}} \begin{bmatrix} \frac{1}{2}\sigma^2 \int_0^1 (\mathcal{B}_t - \widetilde{\mathcal{B}}_t)^2 \, dt \\ \frac{1}{2^{3/2}}\mu\sigma\left(\mathcal{B}_1 + \widetilde{\mathcal{B}}_1 - \int_0^1 (\mathcal{B}_t + \widetilde{\mathcal{B}}_t) \, dt\right) \\ \frac{1}{4}\sigma^2\left[(\mathcal{B}_1 - \widetilde{\mathcal{B}}_1)^2 - 2\right] \\ \frac{1}{2^{1/2}}\sigma(\mathcal{B}_1 + \widetilde{\mathcal{B}}_1) \end{bmatrix}$$

as $n \to \infty$, which yields the statement. Indeed, $\left(2^{-1/2}(\mathcal{B}_t + \widetilde{\mathcal{B}}_t)\right)_{t\in\mathbb{R}_+}$ and $\left(2^{-1/2}(\mathcal{B}_t - \widetilde{\mathcal{B}}_t)\right)_{t\in\mathbb{R}_+}$ are independent standard Wiener processes, and by Itô's formula, $\int_0^1 t \, d\widetilde{\mathcal{W}}_t = \widetilde{\mathcal{W}}_1 - \int_0^1 \widetilde{\mathcal{W}}_t \, dt$

39

and $\int_0^1 \mathcal{W}_t \, d\mathcal{W}_t = 2^{-1}(\mathcal{W}_1^2 - 1)$, which yield the statement with the choices $\widetilde{\mathcal{W}}_t := 2^{-1/2}(\mathcal{B}_t + \widetilde{\mathcal{B}}_t)$, $t \geqslant 0$, and $\mathcal{W}_t := 2^{-1/2}(\mathcal{B}_t - \widetilde{\mathcal{B}}_t)$, $t \geqslant 0$.

Applying Lemmas B.2 and B.3 as in the proof of Proposition 3.1 and using Slutsky's lemma, (8.11) will follow from

$$(8.12) \qquad \frac{1}{n^2} \sum_{k=1}^n \left(V_{k-1} - \mathbb{E}(V_{k-1})\right)^2 - \frac{1}{n} \sum_{k=1}^{\lfloor n/2 \rfloor} \left(\mathcal{S}_{2k/n}^{(\lfloor n/2 \rfloor)} - \mathcal{T}_{2k/n}^{(\lfloor n/2 \rfloor)}\right)^2 \overset{\mathbb{P}}{\longrightarrow} 0,$$

$$(8.13) \quad \frac{1}{n^{3/2}} \sum_{k=1}^n M_k \, \mathbb{E}(X_{k-1}) - \frac{\mu}{2^{3/2}} \left(\mathcal{S}_1^{(\lfloor n/2 \rfloor)} + \mathcal{T}_1^{(\lfloor n/2 \rfloor)} - \frac{2}{n} \sum_{k=1}^{\lfloor n/2 \rfloor} \left(\mathcal{S}_{2k/n}^{(\lfloor n/2 \rfloor)} + \mathcal{T}_{2k/n}^{(\lfloor n/2 \rfloor)}\right)\right) \overset{\mathbb{P}}{\longrightarrow} 0,$$

$$(8.14) \qquad \frac{1}{n} \sum_{k=1}^n M_k \left(V_{k-1} - \mathbb{E}(V_{k-1})\right) + \frac{1}{4}\left[(\mathcal{S}_1^{(\lfloor n/2 \rfloor)} - \mathcal{T}_1^{(\lfloor n/2 \rfloor)})^2 - 2\sigma^2\right] \overset{\mathbb{P}}{\longrightarrow} 0,$$

$$(8.15) \qquad \frac{1}{n^{1/2}} \sum_{k=1}^n M_k - \frac{1}{2^{1/2}}(\mathcal{S}_1^{(\lfloor n/2 \rfloor)} + \mathcal{T}_1^{(\lfloor n/2 \rfloor)}) \overset{\mathbb{P}}{\longrightarrow} 0.$$

Indeed, first considering the subsequence $(2n)_{n \in \mathbb{N}}$, let us apply Lemmas B.2 and B.3 with the special choices $d := 2$, $p := 2$, $q := 2$, $h : \mathbb{R}^2 \to \mathbb{R}^2$,

$$h(x_1, x_2) := \left(x_1 + x_2, \frac{1}{4}(x_1 - x_2)^2 - \frac{\sigma^2}{2}\right), \qquad (x_1, x_2) \in \mathbb{R}^2,$$

$K : [0, 1] \times \mathbb{R}^4 \to \mathbb{R}^2$,

$$K(s, x_1, x_2, x_3, x_4) := \left(\frac{1}{2}(x_1 - x_2)^2, x_1 + x_2\right), \qquad (s, x_1, x_2, x_3, x_4) \in [0, 1] \times \mathbb{R}^4,$$

and

$$\mathcal{U} := \sigma \begin{bmatrix} \mathcal{B} \\ \widetilde{\mathcal{B}} \end{bmatrix}, \qquad \mathcal{U}^{(n)} := \begin{bmatrix} \mathcal{S}^{(n)} \\ \mathcal{T}^{(n)} \end{bmatrix}, \qquad n \in \mathbb{N}.$$

Then

$$\|K(s, x_1, x_2, x_3, x_4) - K(t, y_1, y_2, y_3, y_4)\|$$

$$= \left(\frac{1}{4}\left((x_1 - x_2)^2 - (y_1 - y_2)^2\right)^2 + (x_1 - y_1 + x_2 - y_2)^2\right)^{1/2}$$

$$= \left(\frac{1}{4}(x_1 - y_1 + y_2 - x_2)^2(x_1 - x_2 + y_1 - y_2)^2 + (x_1 - y_1 + x_2 - y_2)^2\right)^{1/2}$$

$$\leqslant 2\left(\left((x_1 - y_1)^2 + (y_2 - x_2)^2\right)\left((x_1 - x_2)^2 + (y_1 - y_2)^2\right) + \left((x_1 - y_1)^2 + (x_2 - y_2)^2\right)\right)^{1/2}$$

$$\leqslant 8R \,\|(x_1, x_2, x_3, x_4) - (y_1, y_2, y_3, y_4)\|$$

40

for all $s, t \in [0, 1]$ and $(x_1, x_2, x_3, x_4), (y_1, y_2, y_3, y_4) \in \mathbb{R}^4$ with $\|(x_1, x_2, x_3, x_4)\| \leqslant R$ and $\|(y_1, y_2, y_3, y_4)\| \leqslant R,$ where $R > 0,$ since

$$(x_1 - x_2)^2 + (y_1 - y_2)^2 \leqslant 2(x_1^2 + x_2^2) + 2(y_1^2 + y_2^2) \leqslant 8R^2.$$

Further, using the definitions of $\Phi$ and $\Phi_n,$ $n \in \mathbb{N},$ given in Lemma B.3, we have

$$\Phi_n\left(\begin{bmatrix} \mathcal{S}^{(n)} \\ \mathcal{T}^{(n)} \end{bmatrix}\right)$$

$$= \left( \mathcal{S}_1^{(n)} + \mathcal{T}_1^{(n)}, \frac{1}{4}(\mathcal{S}_1^{(n)} - \mathcal{T}_1^{(n)})^2 - \frac{\sigma^2}{2}, \frac{1}{n}\sum_{k=1}^n \frac{1}{2}(\mathcal{S}_{k/n}^{(n)} - \mathcal{T}_{k/n}^{(n)})^2, \frac{1}{n}\sum_{k=1}^n (\mathcal{S}_{k/n}^{(n)} + \mathcal{T}_{k/n}^{(n)}) \right)$$

and

$$\Phi\left( \sigma \begin{bmatrix} \mathcal{B} \\ \widetilde{\mathcal{B}} \end{bmatrix}\right) = \left( \sigma(\mathcal{B}_1 + \widetilde{\mathcal{B}}_1), \frac{\sigma^2}{4}(\mathcal{B}_1 - \widetilde{\mathcal{B}}_1)^2 - \frac{\sigma^2}{2}, \int_0^1 \frac{\sigma^2}{2}(\mathcal{B}_u - \widetilde{\mathcal{B}}_u)^2 \, du, \int_0^1 \sigma(\mathcal{B}_u + \widetilde{\mathcal{B}}_u) \, du \right).$$

Since the process $\sigma[\mathcal{B}_t \ \widetilde{\mathcal{B}}_t]_{t\in\mathbb{R}_+}^\top$ admits continuous paths with probability one, (8.9), Lemma B.2 (with the choice $C := \mathsf{C}(\mathbb{R}_+, \mathbb{R}^2)),$ and Lemma B.3 yield that

$$\Phi_n\left(\begin{bmatrix} \mathcal{S}^{(n)} \\ \mathcal{T}^{(n)} \end{bmatrix}\right) \xrightarrow{\mathcal{L}} \Phi\left( \sigma \begin{bmatrix} \mathcal{B} \\ \widetilde{\mathcal{B}} \end{bmatrix}\right) \qquad \text{as } n \to \infty.$$

By another easy application of continuous mapping theorem (one can again apply Lemmas B.2 and B.3) we have

$$\begin{bmatrix} \frac{1}{2n}\sum_{k=1}^n (\mathcal{S}_{k/n}^{(n)} - \mathcal{T}_{k/n}^{(n)})^2 \\ \mathcal{S}_1^{(n)} + \mathcal{T}_1^{(n)} - \frac{1}{n}\sum_{k=1}^n (\mathcal{S}_{k/n}^{(n)} + \mathcal{T}_{k/n}^{(n)}) \\ \frac{1}{4}(\mathcal{S}_1^{(n)} - \mathcal{T}_1^{(n)})^2 - \frac{\sigma^2}{2} \\ \mathcal{S}_1^{(n)} + \mathcal{T}_1^{(n)} \end{bmatrix} \xrightarrow{\mathcal{L}} \begin{bmatrix} \frac{\sigma^2}{2}\int_0^1 (\mathcal{B}_u - \widetilde{\mathcal{B}}_u)^2 \, du \\ \sigma\left(\mathcal{B}_1 + \widetilde{\mathcal{B}}_1 - \int_0^1 (\mathcal{B}_u + \widetilde{\mathcal{B}}_u) \, du\right) \\ \frac{\sigma^2}{4}\left((\mathcal{B}_1 - \widetilde{\mathcal{B}}_1)^2 - 2\right) \\ \sigma(\mathcal{B}_1 + \widetilde{\mathcal{B}}_1) \end{bmatrix} \qquad \text{as } n \to \infty.$$

Hence, using (8.12), (8.13), (8.14), and Slutsky's lemma, we have (8.11) for the subsequence $(2n)_{n\in\mathbb{N}}.$ To prove (8.11) for the subsequence $(2n-1)_{n\in\mathbb{N}},$ by Slutsky's lemma, it is enough to check that as $n \to \infty,$

$$(8.16) \qquad \frac{1}{n^2}(V_n - \mathbb{E}(V_n))^2 \xrightarrow{\mathbb{P}} 0,$$

$$(8.17) \qquad \frac{1}{n^{3/2}}M_n \, \mathbb{E}(X_{n-1}) \xrightarrow{\mathbb{P}} 0,$$

$$(8.18) \qquad \frac{1}{n}M_n(V_{n-1} - \mathbb{E}(V_{n-1})) \xrightarrow{\mathbb{P}} 0$$

$$(8.19) \qquad \frac{1}{n^{1/2}}M_n \xrightarrow{\mathbb{P}} 0.$$

By Corollary 9.1, $\mathbb{E}((V_n - \mathbb{E}(V_n))^2) = O(n)$,

$$\mathbb{E}(|M_n \mathbb{E}(X_{n-1})|) = \mathbb{E}(|M_n|) \mathbb{E}(X_{n-1}) = \mathbb{E}(|\varepsilon_1 - \mu|) \mathbb{E}(X_{n-1}) = O(n),$$

$$\mathbb{E}(M_n^2 (V_{n-1} - \mathbb{E}(V_{n-1}))^2) = \mathbb{E}\left((V_{n-1} - \mathbb{E}(V_{n-1}))^2 \mathbb{E}(M_n^2 \,|\, \mathcal{F}_{n-1})\right)$$

$$= \sigma^2 \, \mathbb{E}((V_{n-1} - \mathbb{E}(V_{n-1}))^2) = O(n),$$

and

$$\mathbb{E}(M_n^2) = \mathbb{E}((\varepsilon_n - \mu)^2) = \sigma^2, \qquad n \in \mathbb{N},$$

thus we obtain (8.16), (8.17), (8.18), and (8.19).

First we will prove (8.12), (8.13), (8.14) and (8.15) for the subsequence $(2n)_{n\in\mathbb{N}}$ and then the subsequence $(2n-1)_{n\in\mathbb{N}}$. In order to prove (8.12) first observe that, for all $k \in \mathbb{N}$,

$$V_{2k} - \mathbb{E}(V_{2k}) = (X_{2k} - \mathbb{E}(X_{2k})) - (X_{2k-1} - \mathbb{E}(X_{2k-1})) = n^{1/2}(\mathcal{S}_{k/n}^{(n)} - \mathcal{T}_{k/n}^{(n)}),$$

$$V_{2k-1} - \mathbb{E}(V_{2k-1}) = (\varepsilon_{2k} - \mathbb{E}(\varepsilon_{2k})) - (V_{2k} - \mathbb{E}(V_{2k})) = (\varepsilon_{2k} - \mu) - n^{1/2}(\mathcal{S}_{k/n}^{(n)} - \mathcal{T}_{k/n}^{(n)}).$$

Then

$$\frac{1}{(2n)^2} \sum_{k=1}^{2n} \left(V_{k-1} - \mathbb{E}(V_{k-1})\right)^2 = \frac{1}{4n^2} \sum_{k=1}^{n-1} \left(V_{2k} - \mathbb{E}(V_{2k})\right)^2 + \frac{1}{4n^2} \sum_{k=1}^{n} \left(V_{2k-1} - \mathbb{E}(V_{2k-1})\right)^2$$

$$= \frac{1}{4n} \sum_{k=1}^{n-1} (\mathcal{S}_{k/n}^{(n)} - \mathcal{T}_{k/n}^{(n)})^2 + \frac{1}{4n^2} \sum_{k=1}^{n} \left[(\varepsilon_{2k} - \mu) - n^{1/2}(\mathcal{S}_{k/n}^{(n)} - \mathcal{T}_{k/n}^{(n)})\right]^2$$

$$= \frac{1}{2n} \sum_{k=1}^{n} (\mathcal{S}_{k/n}^{(n)} - \mathcal{T}_{k/n}^{(n)})^2 - \frac{1}{4n} (\mathcal{S}_1^{(n)} - \mathcal{T}_1^{(n)})^2$$

$$- \frac{1}{2n^{3/2}} \sum_{k=1}^{n} (\varepsilon_{2k} - \mu)(\mathcal{S}_{k/n}^{(n)} - \mathcal{T}_{k/n}^{(n)}) + \frac{1}{4n^2} \sum_{k=1}^{n} (\varepsilon_{2k} - \mu)^2$$

$$= \frac{1}{2n} \sum_{k=1}^{n} (\mathcal{S}_{k/n}^{(n)} - \mathcal{T}_{k/n}^{(n)})^2 - \frac{1}{4n^2} (V_{2n} - \mathbb{E}(V_{2n}))^2$$

$$- \frac{1}{2n^2} \sum_{k=1}^{n} (\varepsilon_{2k} - \mu)(V_{2k} - \mathbb{E}(V_{2k})) + \frac{1}{4n^2} \sum_{k=1}^{n} (\varepsilon_{2k} - \mu)^2.$$

Thus, in order to prove (8.12) for the subsequence $(2n)_{n\in\mathbb{N}}$, it suffices to prove

$$(8.20) \qquad \frac{1}{n^2}(V_{2n} - \mathbb{E}(V_{2n}))^2 \xrightarrow{\mathbb{P}} 0,$$

$$(8.21) \qquad \frac{1}{n^2} \sum_{k=1}^{n} (\varepsilon_{2k} - \mu)(V_{2k} - \mathbb{E}(V_{2k})) \xrightarrow{\mathbb{P}} 0,$$

$$(8.22) \qquad \frac{1}{n^2} \sum_{k=1}^{n} (\varepsilon_{2k} - \mu)^2 \xrightarrow{\mathbb{P}} 0$$

42

as $n \to \infty$. By Corollary 9.1, we have $\mathbb{E}\big((V_{2n} - \mathbb{E}(V_{2n}))^2\big) = \mathrm{O}(n)$ and $\mathbb{E}\big((\varepsilon_{2k} - \mu)^2\big) = \sigma^2$, thus we obtain (8.20) and (8.22). Further, $V_{2k} - \mathbb{E}(V_{2k}) = (\varepsilon_{2k} - \mu) - (V_{2k-1} - \mathbb{E}(V_{2k-1}))$, hence (8.21) follows from (8.22) and from

$$
\mathbb{E}\left(\left(\sum_{k=1}^{n}(\varepsilon_{2k} - \mu)(V_{2k-1} - \mathbb{E}(V_{2k-1}))\right)^2\right) = \sum_{k=1}^{n}\mathbb{E}\left((\varepsilon_{2k} - \mu)^2(V_{2k-1} - \mathbb{E}(V_{2k-1}))^2\right)
$$

$$
= \sigma^2 \sum_{k=1}^{n}\mathbb{E}\left((V_{2k-1} - \mathbb{E}(V_{2k-1}))^2\right) = \mathrm{O}(n^2),
$$

and we finish the proof of (8.12) for the subsequence $(2n)_{n\in\mathbb{N}}$.

Now we turn to prove (8.13) for the subsequence $(2n)_{n\in\mathbb{N}}$. First observe that

$$
\sum_{k=1}^{2n} M_k\,\mathbb{E}(X_{k-1}) = \mu \sum_{k=1}^{n-1}(\varepsilon_{2k+1} - \mu)k + \mu \sum_{k=1}^{n}(\varepsilon_{2k} - \mu)k.
$$

We have

$$
\sum_{k=1}^{n}(\varepsilon_{2k} - \mu)k = \sum_{k=1}^{n}\sum_{j=1}^{k}(\varepsilon_{2k} - \mu) = \sum_{j=1}^{n}\sum_{k=j}^{n}(\varepsilon_{2k} - \mu) = \sum_{j=1}^{n}\left(\sum_{k=1}^{n}(\varepsilon_{2k} - \mu) - \sum_{k=1}^{j-1}(\varepsilon_{2k} - \mu)\right)
$$

$$
= \sum_{j=1}^{n}\left[(X_{2n} - \mathbb{E}(X_{2n})) - (X_{2j-2} - \mathbb{E}(X_{2j-2}))\right] = n^{3/2}\mathcal{S}_1^{(n)} - n^{1/2}\sum_{j=1}^{n}\mathcal{S}_{(j-1)/n}^{(n)},
$$

and, in a similar way,

$$
\sum_{k=1}^{n-1}(\varepsilon_{2k+1} - \mu)k = \sum_{k=1}^{n-1}\sum_{j=1}^{k}(\varepsilon_{2k+1} - \mu) = \sum_{j=1}^{n-1}\sum_{k=j}^{n-1}(\varepsilon_{2k+1} - \mu)
$$

$$
= \sum_{j=1}^{n-1}\left(\sum_{k=1}^{n-1}(\varepsilon_{2k+1} - \mu) - \sum_{k=1}^{j-1}(\varepsilon_{2k+1} - \mu)\right)
$$

$$
= \sum_{j=1}^{n-1}\left(X_{2n-1} - \varepsilon_1 - (n-1)\mu - (X_{2j-1} - \varepsilon_1 - (j-1)\mu)\right)
$$

$$
= \sum_{j=1}^{n-1}\left((X_{2n-1} - \mathbb{E}(X_{2n-1}) - \varepsilon_1 + \mu) - (X_{2j-1} - \mathbb{E}(X_{2j-1}) - \varepsilon_1 + \mu)\right)
$$

$$
= \sum_{j=1}^{n-1}\left(X_{2n-1} - \mathbb{E}(X_{2n-1}) - (X_{2j-1} - \mathbb{E}(X_{2j-1}))\right)
$$

$$
= n^{1/2}(n-1)\mathcal{T}_1^{(n)} - n^{1/2}\sum_{j=1}^{n-1}\mathcal{T}_{j/n}^{(n)} = n^{3/2}\mathcal{T}_1^{(n)} - n^{1/2}\sum_{j=1}^{n}\mathcal{T}_{j/n}^{(n)}.
$$

43

Hence

$$\frac{1}{(2n)^{3/2}} \sum_{k=1}^{2n} M_k \, \mathbb{E}(X_{k-1}) = \frac{\mu}{2^{3/2}} \left( \mathcal{S}_1^{(n)} + \mathcal{T}_1^{(n)} - \frac{1}{n} \sum_{k=1}^{n} \left( \mathcal{S}_{k/n}^{(n)} + \mathcal{T}_{k/n}^{(n)} \right) \right) + \frac{\mu}{2^{3/2} n} \mathcal{S}_1^{(n)}.$$

Convergence (8.9) implies $\mathcal{S}_1^{(n)} \xrightarrow{\mathcal{L}} \sigma \mathcal{B}_1$ and hence $n^{-1} \mathcal{S}_1^{(n)} \xrightarrow{\mathbb{P}} 0$, thus we obtain (8.13) for the subsequence $(2n)_{n \in \mathbb{N}}$.

Now we turn to prove (8.14) for the subsequence $(2n)_{n \in \mathbb{N}}$. First observe that

$$\sum_{k=1}^{2n} M_k \big( V_{k-1} - \mathbb{E}(V_{k-1}) \big) = \sum_{k=1}^{n-1} (\varepsilon_{2k+1} - \mu) \big[ (X_{2k} - \mathbb{E}(X_{2k})) - (X_{2k-1} - \mathbb{E}(X_{2k-1})) \big]$$

$$+ \sum_{k=1}^{n} (\varepsilon_{2k} - \mu) \big[ (X_{2k-1} - \mathbb{E}(X_{2k-1})) - (X_{2k-2} - \mathbb{E}(X_{2k-2})) \big]$$

$$= \sum_{k=1}^{n-1} (\varepsilon_{2k+1} - \mu) \sum_{j=1}^{k} (\varepsilon_{2j} - \mu) - \sum_{k=1}^{n-1} (\varepsilon_{2k+1} - \mu) \sum_{j=1}^{k} (\varepsilon_{2j-1} - \mu)$$

$$+ \sum_{k=1}^{n} (\varepsilon_{2k} - \mu) \sum_{j=1}^{k} (\varepsilon_{2j-1} - \mu) - \sum_{k=1}^{n} (\varepsilon_{2k} - \mu) \sum_{j=1}^{k-1} (\varepsilon_{2j} - \mu).$$

Here the sum of the first and third summands is

$$\sum_{k=1}^{n-1} (\varepsilon_{2k+1} - \mu) \sum_{j=1}^{k} (\varepsilon_{2j} - \mu) + \sum_{k=1}^{n} (\varepsilon_{2k} - \mu) \sum_{j=1}^{k} (\varepsilon_{2j-1} - \mu)$$

$$= \sum_{k=2}^{n} \sum_{j=1}^{k-1} (\varepsilon_{2k-1} - \mu)(\varepsilon_{2j} - \mu) + \sum_{j=1}^{n} \sum_{k=1}^{j} (\varepsilon_{2j} - \mu)(\varepsilon_{2k-1} - \mu)$$

$$= \sum_{j=1}^{n} \sum_{k=j+1}^{n} (\varepsilon_{2k-1} - \mu)(\varepsilon_{2j} - \mu) + \sum_{j=1}^{n} \sum_{k=1}^{j} (\varepsilon_{2j} - \mu)(\varepsilon_{2k-1} - \mu)$$

$$= \sum_{j=1}^{n} \sum_{k=1}^{n} (\varepsilon_{2k-1} - \mu)(\varepsilon_{2j} - \mu)$$

$$= \sum_{j=1}^{n} (\varepsilon_{2j} - \mu) \sum_{k=1}^{n} (\varepsilon_{2k-1} - \mu) = n \mathcal{S}_1^{(n)} \mathcal{T}_1^{(n)},$$

the second summand is

$$\sum_{k=1}^{n-1} (\varepsilon_{2k+1} - \mu) \sum_{j=1}^{k} (\varepsilon_{2j-1} - \mu) = \sum_{1 \leqslant j < \ell \leqslant n} (\varepsilon_{2j-1} - \mu)(\varepsilon_{2\ell-1} - \mu)$$

$$= \frac{1}{2} \left[ \left( \sum_{k=1}^{n} (\varepsilon_{2k-1} - \mu) \right)^2 - \sum_{k=1}^{n} (\varepsilon_{2k-1} - \mu)^2 \right] = \frac{1}{2} \left[ n(\mathcal{T}_1^{(n)})^2 - \sum_{k=1}^{n} (\varepsilon_{2k-1} - \mu)^2 \right],$$

44

and similarly, the forth summand is

$$\sum_{k=1}^{n}(\varepsilon_{2k}-\mu)\sum_{j=1}^{k-1}(\varepsilon_{2j}-\mu) = \frac{1}{2}\left[n(\mathcal{S}_1^{(n)})^2 - \sum_{k=1}^{n}(\varepsilon_{2k}-\mu)^2\right].$$

Consequently,

$$\frac{1}{2n}\sum_{k=1}^{2n}M_k\big(V_{k-1}-\mathbb{E}(V_{k-1})\big) = -\frac{1}{4}\left[(\mathcal{S}_1^{(n)}-\mathcal{T}_1^{(n)})^2 - 2\sigma^2\right] + \frac{1}{4n}\sum_{k=1}^{2n}(\varepsilon_k-\mu)^2 - \frac{1}{2}\sigma^2.$$

By the strong law of large numbers $(2n)^{-1}\sum_{k=1}^{2n}(\varepsilon_k-\mu)^2 \xrightarrow{\text{a.s.}} \sigma^2$ as $n\to\infty$, hence we obtain (8.14) for the subsequence $(2n)_{n\in\mathbb{N}}$. Note also that the convergence in (8.14) holds almost surely, too.

Now we turn to prove (8.15) for the subsequence $(2n)_{n\in\mathbb{N}}$. First observe that

$$\sum_{k=1}^{2n}M_k = \sum_{k=1}^{2n}(\varepsilon_k-\mu) = \sum_{k=1}^{n}(\varepsilon_{2k}-\mu) + \sum_{k=1}^{n}(\varepsilon_{2k-1}-\mu).$$

Hence

$$\frac{1}{(2n)^{1/2}}\sum_{k=1}^{2n}M_k = \frac{1}{2^{1/2}}\left[\frac{1}{n^{1/2}}(X_{2k}-\mathbb{E}(X_{2k})) + \frac{1}{n^{1/2}}(X_{2k-1}-\mathbb{E}(X_{2k-1}))\right] = \frac{1}{2^{1/2}}\left(\mathcal{S}_1^{(n)}+\mathcal{T}_1^{(n)}\right),$$

thus we obtain (8.15) for the subsequence $(2n)_{n\in\mathbb{N}}$.

Finally, one can show (8.12), (8.13), (8.14), and (8.15) for the subsequence $(2n-1)_{n\in\mathbb{N}}$ in the same way. $\qquad\square$

# 9  Estimations of moments

In the proofs of Theorem 2.1, Theorem 2.2 and Theorem 2.3 good bounds for moments of the random variables $(M_k)_{k\in\mathbb{Z}_+}$, $(X_k)_{k\in\mathbb{Z}_+}$, $(U_k)_{k\in\mathbb{Z}_+}$ and $(V_k)_{k\in\mathbb{Z}_+}$ are extensively used. First note that, for all $k\in\mathbb{N}$, $\mathbb{E}(M_k\,|\,\mathcal{F}_{k-1})=0$ and $\mathbb{E}(M_k)=0$, since $M_k = X_k - \mathbb{E}(X_k\,|\,\mathcal{F}_{k-1})$.

**9.1 Lemma.** *Let* $(X_k)_{k\geqslant-1}$ *be an* INAR(2) *process. Suppose that* $X_0 = X_{-1} = 0$ *and*

$\mathbb{E}(\varepsilon_1^2) < \infty.$ *Then, for all* $k, \ell \in \mathbb{N},$

(9.1) $\qquad \mathbb{E}(M_k M_\ell \,|\, \mathcal{F}_{\max\{k,\ell\}-1}) = \begin{cases} \alpha(1-\alpha)X_{k-1} + \beta(1-\beta)X_{k-2} + \sigma^2 & \text{if } k = \ell, \\ 0 & \text{if } k \neq \ell, \end{cases}$

(9.2) $\qquad \mathbb{E}(M_k M_\ell) = \begin{cases} \alpha(1-\alpha)\,\mathbb{E}(X_{k-1}) + \beta(1-\beta)\,\mathbb{E}(X_{k-2}) + \sigma^2 & \text{if } k = \ell, \\ 0 & \text{if } k \neq \ell, \end{cases}$

(9.3)
$$\mathbb{E}(M_k^3 \,|\, \mathcal{F}_{k-1}) = X_{k-1}\,\mathbb{E}\left[(\xi_{1,1} - \mathbb{E}(\xi_{1,1}))^3\right] + X_{k-2}\,\mathbb{E}\left[(\eta_{1,1} - \mathbb{E}(\eta_{1,1}))^3\right] + \mathbb{E}\left[(\varepsilon_1 - \mathbb{E}(\varepsilon_1))^3\right],$$

(9.4)
$$\mathbb{E}(M_k^3) = \mathbb{E}\left[(\xi_{1,1} - \mathbb{E}(\xi_{1,1}))^3\right]\mathbb{E}(X_{k-1}) + \mathbb{E}\left[(\eta_{1,1} - \mathbb{E}(\eta_{1,1}))^3\right]\mathbb{E}(X_{k-2}) + \mathbb{E}\left[(\varepsilon_1 - \mathbb{E}(\varepsilon_1))^3\right].$$

**Proof.** By (1.1) and (3.2),

(9.5) $\qquad M_k = \sum_{j=1}^{X_{k-1}} \left(\xi_{k,j} - \mathbb{E}(\xi_{k,j})\right) + \sum_{j=1}^{X_{k-2}} \left(\eta_{k,j} - \mathbb{E}(\eta_{k,j})\right) + \left(\varepsilon_k - \mathbb{E}(\varepsilon_k)\right), \qquad k \in \mathbb{N}.$

For all $k \in \mathbb{N}$, the random variables $\{\xi_{k,j} - \mathbb{E}(\xi_{k,j}),\, \eta_{k,j} - \mathbb{E}(\eta_{k,j}),\, \varepsilon_k - \mathbb{E}(\varepsilon_k) : j \in \mathbb{N}\}$ are independent of each other, independent of $\mathcal{F}_{k-1}$, and have zero mean, thus in case $k = \ell$ we conclude (9.1) and hence (9.2). If $k < \ell$, then $\mathbb{E}(M_k M_\ell \,|\, \mathcal{F}_{\ell-1}) = M_k\,\mathbb{E}(M_\ell \,|\, \mathcal{F}_{\ell-1}) = 0$. Thus we obtain (9.1) and (9.2) in case $k \neq \ell$. Shedding more light we give more details for deriving (9.3) and (9.4). Namely, using multinomial theorem the above mentioned properties of the random variables $\{\xi_{k,j} - \mathbb{E}(\xi_{k,j}),\, \eta_{k,j} - \mathbb{E}(\eta_{k,j}),\, \varepsilon_k - \mathbb{E}(\varepsilon_k) : j \in \mathbb{N}\}$ yield that

$$\mathbb{E}(M_k^3 \,|\, \mathcal{F}_{k-1}) = \mathbb{E}\left( \sum_{j=1}^{X_{k-1}} (\xi_{k,j} - \mathbb{E}(\xi_{k,j}))^3 + \sum_{j=1}^{X_{k-2}} (\eta_{k,j} - \mathbb{E}(\eta_{k,j}))^3 + (\varepsilon_k - \mathbb{E}(\varepsilon_k))^3 \,\Big|\, \mathcal{F}_{k-1} \right)$$

$$= X_{k-1}\,\mathbb{E}[(\xi_{1,1} - \mathbb{E}(\xi_{1,1}))^3] + X_{k-2}\,\mathbb{E}[(\eta_{1,1} - \mathbb{E}(\eta_{1,1}))^3] + \mathbb{E}[(\varepsilon_1 - \mathbb{E}(\varepsilon_1))^3].$$

This readily implies (9.4). $\qquad \qquad \square$

**9.2 Lemma.** *Let* $(\zeta_k)_{k\in\mathbb{N}}$ *be independent and identically distributed random variables such that* $\mathbb{E}(|\zeta_1|^\ell) < \infty$ *for some* $\ell \in \mathbb{N}$.

(i) *If* $\mathbb{E}(\zeta_1) \neq 0$, *then there exists a polynomial* $Q_\ell$ *of degree* $\ell$ *such that its leading coefficient is* $[\mathbb{E}(\zeta_1)]^\ell$ *and*

$$\mathbb{E}\left((\zeta_1 + \cdots + \zeta_N)^\ell\right) = Q_\ell(N), \qquad N \in \mathbb{N}.$$

(ii) *If* $\mathbb{E}(\zeta_1) = 0$, *then there exists a polynomial* $R_\ell$ *of degree at most* $\ell/2$ *such that*

$$\mathbb{E}\left((\zeta_1 + \cdots + \zeta_N)^\ell\right) = R_\ell(N), \qquad N \in \mathbb{N}.$$

*The coefficients of the polynomials in question depend on the moments $\mathbb{E}(\zeta_1^j)$, $j \in \{1, \ldots, \ell\}$.*

**Proof.** (i) By multinomial theorem,

$$\mathbb{E}\big((\zeta_1 + \cdots + \zeta_N)^\ell\big) = \sum_{\substack{\ell_1 + \cdots + \ell_N = \ell, \\ \ell_1, \ldots, \ell_N \in \mathbb{Z}_+}} \frac{\ell!}{\ell_1! \cdots \ell_N!} \, \mathbb{E}(\zeta_1^{\ell_1} \cdots \zeta_N^{\ell_N})$$

$$= \sum_{\substack{\ell_1 + \cdots + \ell_N = \ell, \\ \ell_1, \ldots, \ell_N \in \mathbb{Z}_+}} \frac{\ell!}{\ell_1! \cdots \ell_N!} \, \mathbb{E}(\zeta_1^{\ell_1}) \cdots \mathbb{E}(\zeta_1^{\ell_N})$$

$$= \sum_{\substack{k_1 + 2k_2 + \cdots + sk_s = \ell, \\ k_1, \ldots, k_s \in \mathbb{Z}_+, \, 1 \leqslant s \leqslant \ell}} \binom{N}{k_1} \binom{N - k_1}{k_2} \cdots \binom{N - k_1 - \cdots - k_{s-1}}{k_s}$$

$$\times \frac{\ell!}{(2!)^{k_2}(3!)^{k_3} \cdots (s!)^{k_s}} \big[\mathbb{E}(\zeta_1)\big]^{k_1} \cdots \big[\mathbb{E}(\zeta_1^s)\big]^{k_s}.$$

Since

$$\binom{N}{k_1} \binom{N - k_1}{k_2} \cdots \binom{N - k_1 - \cdots - k_{s-1}}{k_s} = \frac{N(N-1) \cdots (N - k_1 - k_2 - \cdots - k_s + 1)}{k_1! k_2! \cdots k_s!}$$

is a polynomial of the variable $N$ having degree $k_1 + \cdots + k_s \leqslant \ell$, there is a polynomial $Q_\ell$ of degree at most $\ell$ such that $\mathbb{E}\big((\zeta_1 + \cdots + \zeta_N)^\ell\big) = Q_\ell(N)$, $N \in \mathbb{N}$. Note that a term of degree $\ell$ can occur only in the case $k_1 + \cdots + k_s = \ell$. Since $k_1 + 2k_2 + \cdots + sk_s = \ell$, we have $s = 1$ and $k_1 = \ell$, and the corresponding term of degree $\ell$ is $N(N-1) \cdots (N - \ell + 1)\big[\mathbb{E}(\zeta_1)\big]^\ell$. Hence $Q_\ell$ is polynomial of degree $\ell$ having leading coefficient $\big[\mathbb{E}(\zeta_1)\big]^\ell$.

(ii) Using again the multinomial theorem we have

$$\mathbb{E}\big((\zeta_1 + \cdots + \zeta_N)^\ell\big) = \sum_{\substack{\ell_1 + \cdots + \ell_N = \ell, \\ \ell_1, \ldots, \ell_N \in \mathbb{Z}_+}} \frac{\ell!}{\ell_1! \cdots \ell_N!} \, \mathbb{E}(\zeta_1^{\ell_1} \cdots \zeta_N^{\ell_N})$$

$$= \sum_{\substack{\ell_1 + \cdots + \ell_N = \ell, \\ \ell_1, \ldots, \ell_N \in \mathbb{Z}_+ \backslash \{1\}}} \frac{\ell!}{\ell_1! \cdots \ell_N!} \, \mathbb{E}(\zeta_1^{\ell_1}) \cdots \mathbb{E}(\zeta_1^{\ell_N})$$

$$= \sum_{\substack{2k_2 + 3k_3 + \cdots + sk_s = \ell, \\ k_2, \ldots, k_s \in \mathbb{Z}_+, \, 2 \leqslant s \leqslant \ell}} \binom{N}{k_2} \binom{N - k_2}{k_3} \cdots \binom{N - k_2 - \cdots - k_{s-1}}{k_s}$$

$$\times \frac{\ell!}{(2!)^{k_2}(3!)^{k_3} \cdots (s!)^{k_s}} \big[\mathbb{E}(\zeta_1^2)\big]^{k_2} \cdots \big[\mathbb{E}(\zeta_1^s)\big]^{k_s}.$$

Here

$$\binom{N}{k_2} \binom{N - k_2}{k_3} \cdots \binom{N - k_2 - \cdots - k_{s-1}}{k_s} = \frac{N(N-1) \cdots (N - k_2 - k_3 - \cdots - k_s + 1)}{k_2! k_3! \cdots k_s!}$$

is a polynomial of the variable $N$ having degree $k_2 + \cdots + k_s$. Since

$$\ell = 2k_2 + 3k_3 + \cdots + sk_s \geqslant 2(k_2 + k_3 + \cdots + k_s),$$

we have $k_2 + \cdots + k_s \leqslant \ell/2$ yielding part (ii). Note that if $\ell$ is even and $\mathbb{E}(\zeta_1^2) \neq 0$, then the degree of $R_\ell$ is $\ell/2$; if $\ell$ is odd and $\mathbb{E}(\zeta_1^2) \neq 0$, $\mathbb{E}(\zeta_1^3) \neq 0$, then the degree of $R_\ell$ is also $\ell/2$. $\qquad\square$

**9.1 Remark.** In what follows using the proof of Lemma 9.2 we give a bit more explicit form of the polynomial $R_\ell$ in part (ii) of Lemma 9.2 for the special cases $\ell \in \{1, 2, 3, 4, 5, 6\}$. If $\ell = 1$, then $\mathbb{E}(\zeta_1 + \cdots + \zeta_N) = 0$ and $R_1 : \mathbb{R} \to \mathbb{R}$, $R_1(x) := 0$, $x \in \mathbb{R}$.

If $\ell = 2$, then

$$\mathbb{E}((\zeta_1 + \cdots + \zeta_N)^2) = N\,\mathbb{E}(\zeta_1^2),$$

and $R_2 : \mathbb{R} \to \mathbb{R}$, $R_2(x) := \mathbb{E}(\zeta_1^2)x$, $x \in \mathbb{R}$.

If $\ell = 3$, then

$$\mathbb{E}((\zeta_1 + \cdots + \zeta_N)^3) = N\,\mathbb{E}(\zeta_1^3),$$

and $R_3 : \mathbb{R} \to \mathbb{R}$, $R_3(x) := \mathbb{E}(\zeta_1^3)x$, $x \in \mathbb{R}$.

If $\ell = 4$, then

$$\mathbb{E}((\zeta_1 + \cdots + \zeta_N)^4) = N\,\mathbb{E}(\zeta_1^4) + \binom{N}{2}\frac{4!}{2!2!}(\mathbb{E}(\zeta_1^2))^2,$$

and $R_4 : \mathbb{R} \to \mathbb{R}$, $R_4(x) := \mathbb{E}(\zeta_1^4)x + 3(\mathbb{E}(\zeta_1^2))^2x(x-1)$, $x \in \mathbb{R}$.

If $\ell = 5$, then

$$\mathbb{E}((\zeta_1 + \cdots + \zeta_N)^5) = N\,\mathbb{E}(\zeta_1^5) + 2\binom{N}{2}\frac{5!}{2!3!}\,\mathbb{E}(\zeta_1^3)\,\mathbb{E}(\zeta_1^2),$$

and $R_5 : \mathbb{R} \to \mathbb{R}$, $R_5(x) := \mathbb{E}(\zeta_1^5)x + 10\,\mathbb{E}(\zeta_1^3)\,\mathbb{E}(\zeta_1^2)x(x-1)$, $x \in \mathbb{R}$.

If $\ell = 6$, then

$$\mathbb{E}((\zeta_1 + \cdots + \zeta_N)^6) = N\,\mathbb{E}(\zeta_1^6) + 2\binom{N}{2}\frac{6!}{2!4!}\,\mathbb{E}(\zeta_1^4)\,\mathbb{E}(\zeta_1^2) + \binom{N}{2}\frac{6!}{3!3!}(\mathbb{E}(\zeta_1^3))^2 + \binom{N}{3}\frac{6!}{2!2!2!}(\mathbb{E}(\zeta_1^2))^3,$$

and $R_6 : \mathbb{R} \to \mathbb{R}$,

$$R_6(x) := \mathbb{E}(\zeta_1^6)x + 15\,\mathbb{E}(\zeta_1^4)\,\mathbb{E}(\zeta_1^2)x(x-1) + 10(\mathbb{E}(\zeta_1^3))^2x(x-1) + 15(\mathbb{E}(\zeta_1^2))^3x(x-1)(x-2), \quad x \in \mathbb{R}.$$

$\qquad\square$

**9.3 Lemma.** *If* $\alpha + \beta = 1$, *then the matrix* $A$ *defined in (1.2) has eigenvalues* $1$ *and* $\alpha - 1 = -\beta$, *and the powers of* $A$ *take the following form*

$$A^k = \frac{1}{1+\beta}\begin{bmatrix} 1 & \beta \\ 1 & \beta \end{bmatrix} + \frac{(-\beta)^k}{1+\beta}\begin{bmatrix} \beta & -\beta \\ -1 & 1 \end{bmatrix} = \boldsymbol{u}\widetilde{\boldsymbol{u}}^\top + (-\beta)^k\boldsymbol{v}\widetilde{\boldsymbol{v}}^\top, \qquad k \in \mathbb{Z}_+,$$

*with*

$$\boldsymbol{u} := \frac{1}{1+\beta}\begin{bmatrix}1\\1\end{bmatrix}, \quad \widetilde{\boldsymbol{u}} := \begin{bmatrix}1\\\beta\end{bmatrix}, \quad \boldsymbol{v} := \frac{1}{1+\beta}\begin{bmatrix}\beta\\-1\end{bmatrix}, \quad \widetilde{\boldsymbol{v}} := \begin{bmatrix}1\\-1\end{bmatrix}.$$

**Proof.** The formula for the powers of $A$ follows by the so-called Putzer's spectral formula, see, e.g., Putzer [29]. $\square$

**9.2 Remark.** Using Lemma 9.3 we obtain the decomposition

$$(9.6) \qquad \begin{bmatrix}X_k\\X_{k-1}\end{bmatrix} = U_k\boldsymbol{u} + V_k\boldsymbol{v} = \frac{1}{1+\beta}\begin{bmatrix}U_k+\beta V_k\\U_k-V_k\end{bmatrix} = \frac{1}{1+\beta}\begin{bmatrix}1&\beta\\1&-1\end{bmatrix}\begin{bmatrix}U_k\\V_k\end{bmatrix}, \qquad k\in\mathbb{N},$$

with

$$(9.7) \qquad\qquad U_k = X_k + \beta X_{k-1}, \qquad V_k = X_k - X_{k-1}, \qquad k\in\mathbb{N}.$$

Note that (9.6) is valid for $k=0$ with the convention $U_0 := 0$ and $V_0 := 0$. The decomposition (9.6) can be considered as a motivation for the definition of $U_k$ and $V_k$, $k\in\mathbb{N}$, given in Sections 3 and 5. $\square$

**9.4 Lemma.** *Let* $(X_k)_{k\geqslant-1}$ *be an* INAR(2) *process with autoregressive parameters* $(\alpha,\beta)\in [0,1]^2$ *such that* $\alpha+\beta=1$ *(hence it is unstable). Suppose that* $X_0 = X_{-1} = 0$ *and* $\mathbb{E}(\varepsilon_1^\ell) < \infty$ *with some* $\ell\in\mathbb{N}$. *Then there exists a constant* $c_\ell$ *such that* $\mathbb{E}(X_n^{\ell_1}X_{n-1}^{\ell_2}) \leqslant c_\ell n^\ell$, $n\in\mathbb{N}$, *for all* $\ell_1,\ell_2\in\mathbb{Z}_+$ *with* $\ell_1+\ell_2\leqslant\ell$.

**First proof.** Observe that the statement is equivalent with the following: for each polynomial $P$ of two variables having degree at most $\ell$, there exists a constant $c_P$ such that $\mathbb{E}\big(|P(X_n, X_{n-1})|\big) \leqslant c_P n^\ell$, $n\in\mathbb{N}$.

First let us suppose that $(\alpha,\beta)\in(0,1)^2$. If $\ell=1$, i.e., $(\ell_1,\ell_2)=(1,0)$ or $(\ell_1,\ell_2)=(0,1)$, then to conclude the statement we show that

$$(9.8) \qquad\qquad \mathbb{E}(X_n) = \frac{\mu}{1+\beta}n + \frac{\mu\beta}{(1+\beta)^2}(1-(-\beta)^n), \qquad n\in\mathbb{N}.$$

Since $\mathbb{E}(X_n\,|\,\mathcal{F}_{n-1}) = \alpha X_{n-1} + \beta X_{n-2} + \mu$, $n\in\mathbb{N}$, we have $\mathbb{E}(X_n) = \alpha\,\mathbb{E}(X_{n-1}) + \beta\,\mathbb{E}(X_{n-2}) + \mu$, $n\in\mathbb{N}$, yielding that

$$\begin{bmatrix}\mathbb{E}(X_n)\\\mathbb{E}(X_{n-1})\end{bmatrix} = \begin{bmatrix}\alpha&\beta\\1&0\end{bmatrix}\begin{bmatrix}\mathbb{E}(X_{n-1})\\\mathbb{E}(X_{n-2})\end{bmatrix} + \begin{bmatrix}\mu\\0\end{bmatrix} = A\begin{bmatrix}\mathbb{E}(X_{n-1})\\\mathbb{E}(X_{n-2})\end{bmatrix} + \begin{bmatrix}\mu\\0\end{bmatrix}, \qquad n\in\mathbb{N}.$$

By Lemma 9.3, we get

$$\begin{bmatrix}\mathbb{E}(X_n)\\\mathbb{E}(X_{n-1})\end{bmatrix} = \sum_{j=1}^n A^{n-j}\begin{bmatrix}\mu\\0\end{bmatrix} = \left(\frac{n}{1+\beta}\begin{bmatrix}1&\beta\\1&\beta\end{bmatrix} + \frac{1-(-\beta)^n}{(1+\beta)^2}\begin{bmatrix}\beta&-\beta\\-1&1\end{bmatrix}\right)\begin{bmatrix}\mu\\0\end{bmatrix}, \qquad n\in\mathbb{N},$$

49

which yields (9.8).

Let us suppose now that the statement holds for $1, \ldots, \ell - 1$. By multinomial theorem,

$$(9.9) \qquad X_n^k = \sum_{\substack{k_1+k_2+k_3=k, \\ k_1,k_2,k_3 \in \mathbb{Z}_+}} \frac{k!}{k_1!k_2!k_3!} \left( \sum_{j=1}^{X_{n-1}} \xi_{n,j} \right)^{k_1} \left( \sum_{j=1}^{X_{n-2}} \eta_{n,j} \right)^{k_2} \varepsilon_n^{k_3}, \qquad k \in \mathbb{N}.$$

Since for all $n \in \mathbb{N}$ the random variables $\{\xi_{n,j}, \eta_{n,j}, \varepsilon_n : j \in \mathbb{N}\}$ are independent of each other and of the $\sigma$-algebra $\mathcal{F}_{n-1}$, we have for all $\ell_1, \ell_2 \in \mathbb{Z}_+$ with $\ell_1 + \ell_2 = \ell$

$$\mathbb{E}(X_n^{\ell_1} X_{n-1}^{\ell_2} \,|\, \mathcal{F}_{n-1})$$

$$= X_{n-1}^{\ell_2} \sum_{\substack{k_1+k_2+k_3=\ell_1, \\ k_1,k_2,k_3 \in \mathbb{Z}_+}} \frac{\ell_1!}{k_1!k_2!k_3!} \, \mathbb{E} \left( \left( \sum_{j=1}^{M} \xi_{n,j} \right)^{k_1} \right) \Bigg|_{M=X_{n-1}} \mathbb{E} \left( \left( \sum_{j=1}^{N} \eta_{n,j} \right)^{k_2} \right) \Bigg|_{N=X_{n-2}} \mathbb{E}(\varepsilon_1^{k_3}).$$

Using part (i) of Lemma 9.2 and separating the terms having degree $\ell$ and less than $\ell$, we have

$$\mathbb{E}(X_n^{\ell_1} X_{n-1}^{\ell_2} \,|\, \mathcal{F}_{n-1}) = \sum_{\substack{k_1+k_2=\ell_1, \\ k_1,k_2 \in \mathbb{Z}_+}} \frac{\ell_1!}{k_1!k_2!} \alpha^{k_1} X_{n-1}^{\ell_2+k_1} \beta^{k_2} X_{n-2}^{k_2} + Q_{\ell_1,\ell_2}(X_{n-1}, X_{n-2}),$$

where $Q_{\ell_1,\ell_2}$ is a polynomial of two variables having degree at most $\ell - 1$. Hence

$$\mathbb{E}(X_n^{\ell_1} X_{n-1}^{\ell_2}) = \sum_{\substack{k_1+k_2=\ell_1, \\ k_1,k_2 \in \mathbb{Z}_+}} \frac{\ell_1!}{k_1!k_2!} \alpha^{k_1} \beta^{k_2} \, \mathbb{E} \left( X_{n-1}^{\ell_2+k_1} X_{n-2}^{k_2} \right) + \mathbb{E} \left( Q_{\ell_1,\ell_2}(X_{n-1}, X_{n-2}) \right).$$

By the induction hypothesis (used for polynomials, see the beginning of the proof), there exists a constant $c_{Q_{\ell_1,\ell_2}}$ such that $\mathbb{E} \left( |Q_{\ell_1,\ell_2}(X_n, X_{n-1})| \right) \leqslant c_{Q_{\ell_1,\ell_2}} n^{\ell-1}$, $n \in \mathbb{N}$. In fact, we have

$$(9.10) \qquad \mathbb{E} \left( |Q_{\ell_1,\ell_2}(X_n, X_{n-1})| \right) \leqslant \widetilde{c}_\ell n^{\ell-1}$$

for $n \in \mathbb{N}$ and $\ell_1, \ell_2 \in \mathbb{Z}_+$ with $\ell_1 + \ell_2 = \ell$, where $\widetilde{c}_\ell := \max_{0 \leqslant i \leqslant \ell} c_{Q_{i,\ell-i}}$. Consequently, we have

$$\mathbb{E}(X_n^{\ell_1} X_{n-1}^{\ell_2}) \leqslant \sum_{\substack{k_1+k_2=\ell_1, \\ k_1,k_2 \in \mathbb{Z}_+}} \frac{\ell_1!}{k_1!k_2!} \alpha^{k_1} \beta^{k_2} \, \mathbb{E} \left( X_{n-1}^{\ell_2+k_1} X_{n-2}^{k_2} \right) + \widetilde{c}_\ell (n-1)^{\ell-1}.$$

Similarly, for all $k_1, k_2 \in \mathbb{Z}_+$ with $k_1 + k_2 = \ell_1$, we have

$$\mathbb{E} \left( X_{n-1}^{\ell_2+k_1} X_{n-2}^{k_2} \right) = \sum_{\substack{j_1+j_2=\ell_2+k_1, \\ j_1,j_2 \in \mathbb{Z}_+}} \frac{(\ell_2+k_1)!}{j_1!j_2!} \alpha^{j_1} \beta^{j_2} \, \mathbb{E} \left( X_{n-2}^{k_2+j_1} X_{n-3}^{j_2} \right) + \mathbb{E} \left( Q_{\ell_2+k_1,k_2}(X_{n-2}, X_{n-3}) \right).$$

50

Hence we have

$$\mathbb{E}(X_n^{\ell_1} X_{n-1}^{\ell_2}) = \sum_{\substack{k_1+k_2=\ell_1, \\ k_1,k_2 \in \mathbb{Z}_+}} \frac{\ell_1!}{k_1!k_2!} \alpha^{k_1} \beta^{k_2} \sum_{\substack{j_1+j_2=\ell_2+k_1, \\ j_1,j_2 \in \mathbb{Z}_+}} \frac{(\ell_2+k_1)!}{j_1!j_2!} \alpha^{j_1} \beta^{j_2} \mathbb{E}\left(X_{n-2}^{k_2+j_1} X_{n-3}^{j_2}\right)$$

$$+ \sum_{\substack{k_1+k_2=\ell_1, \\ k_1,k_2 \in \mathbb{Z}_+}} \frac{\ell_1!}{k_1!k_2!} \alpha^{k_1} \beta^{k_2} \mathbb{E}\left(Q_{\ell_2+k_1,k_2}(X_{n-2}, X_{n-3})\right) + \mathbb{E}\left(Q_{\ell_1,\ell_2}(X_{n-1}, X_{n-2})\right).$$

Applying (9.10) and

$$\sum_{\substack{k_1+k_2=\ell_1, \\ k_1,k_2 \in \mathbb{Z}_+}} \frac{\ell_1!}{k_1!k_2!} \alpha^{k_1} \beta^{k_2} = (\alpha + \beta)^{\ell_1} = 1,$$

we conclude

$$\mathbb{E}(X_n^{\ell_1} X_{n-1}^{\ell_2}) \leqslant \sum_{\substack{k_1+k_2=\ell_1, \\ k_1,k_2 \in \mathbb{Z}_+}} \frac{\ell_1!}{k_1!k_2!} \alpha^{k_1} \beta^{k_2} \sum_{\substack{j_1+j_2=\ell_2+k_1, \\ j_1,j_2 \in \mathbb{Z}_+}} \frac{(\ell_2+k_1)!}{j_1!j_2!} \alpha^{j_1} \beta^{j_2} \mathbb{E}\left(X_{n-2}^{k_2+j_1} X_{n-3}^{j_2}\right)$$

$$+ \widetilde{c}_\ell (n-2)^{\ell-1} + \widetilde{c}_\ell (n-1)^{\ell-1}.$$

Using that $\mathbb{E}(X_1^r X_0^q) = 0$, $r,q \in \mathbb{Z}_+$ (since $X_0 = 0$), after $n-1$ steps, one can derive

$$\mathbb{E}(X_n^{\ell_1} X_{n-1}^{\ell_2}) \leqslant \widetilde{c}_\ell \sum_{i=1}^{n-1} i^{\ell-1} \leqslant \widetilde{c}_\ell n \cdot n^{\ell-1} = \mathrm{O}(n^\ell), \qquad n \in \mathbb{N},$$

that is, $\mathbb{E}(P(X_n, X_{n-1})) \leqslant \widetilde{c}_\ell n^\ell$ for all monomials $P(x,y) := x^{\ell_1} y^{\ell_2}$, $x,y \in \mathbb{R}$, with $\ell_1 + \ell_2 = \ell$, $\ell_1, \ell_2 \in \mathbb{Z}_+$. If $P$ has the form

$$P(x,y) := \sum_{i=0}^{\ell} p_i x^i y^{\ell-i} + Q(x,y), \qquad x,y \in \mathbb{R},$$

where $p_i \in \mathbb{R}$, $i \in \{0, \ldots, \ell\}$, and $Q$ is a polynomial of two variables having degree at most $\ell - 1$, then for all $n \in \mathbb{N}$,

$$\mathbb{E}(|P(X_n, X_{n-1})|) \leqslant \sum_{i=0}^{\ell} |p_i| \, \mathbb{E}(X_n^i X_{n-1}^{\ell-i}) + \mathbb{E}(Q(X_n, X_{n-1})) \leqslant \left(\sum_{i=0}^{\ell} |p_i| c_\ell\right) n^\ell + c_Q n^{\ell-1} \leqslant c_P n^\ell,$$

where $c_P := c_Q + c_\ell \sum_{i=0}^{\ell} |p_i|$, as desired.

Next let us suppose that $(\alpha, \beta) = (1, 0)$. Then $X_n = X_{n-1} + \varepsilon_n$, $n \in \mathbb{N}$, which implies that $X_n = \sum_{i=1}^{n} \varepsilon_i$, $n \in \mathbb{N}$. By part (i) of Lemma 9.2,

$$(9.11) \qquad\qquad \mathbb{E}(X_n^\ell) = Q_\ell(n), \qquad n \in \mathbb{N},$$

where $Q_\ell$ is a polynomial of degree $\ell$. If $\ell_1, \ell_2 \in \mathbb{Z}_+$ with $\ell_1 + \ell_2 \leqslant \ell$, then using independence of $X_{n-1}$ and $\varepsilon_n$ we have

$$\mathbb{E}(X_n^{\ell_1} X_{n-1}^{\ell_2}) = \mathbb{E}((X_{n-1} + \varepsilon_n)^{\ell_1} X_{n-1}^{\ell_2}) = \mathbb{E}\left(\sum_{j=0}^{\ell_1} \binom{\ell_1}{j} X_{n-1}^j \varepsilon_n^{\ell_1 - j} X_{n-1}^{\ell_2}\right)$$

$$= \sum_{j=0}^{\ell_1} \binom{\ell_1}{j} \mathbb{E}(X_{n-1}^{j+\ell_2}) \mathbb{E}(\varepsilon_n^{\ell_1 - j}), \qquad n \in \mathbb{N}.$$

Using (9.11),

$$\mathbb{E}(X_n^{\ell_1} X_{n-1}^{\ell_2}) = \sum_{j=0}^{\ell_1} \binom{\ell_1}{j} Q_{j+\ell_2}(n-1) \mathbb{E}(\varepsilon_1^{\ell_1 - j}) = \mathrm{O}(n^\ell), \qquad n \in \mathbb{N},$$

since for each $j \in \{0, \ldots, \ell_1\}$, the polynomial $Q_{j+\ell_2}$ is of degree $j + \ell_2 \leqslant \ell$, which yields the statement in case $(\alpha, \beta) = (1, 0)$.

Finally, let us suppose that $(\alpha, \beta) = (0, 1)$. Then $X_n = X_{n-2} + \varepsilon_n$, $n \in \mathbb{N}$, which implies that

$$X_{2n} = \sum_{i=1}^n \varepsilon_{2i}, \qquad X_{2n-1} = \sum_{i=1}^n \varepsilon_{2i-1}, \qquad n \in \mathbb{N}.$$

By part (i) of Lemma 9.2, we have

$$\mathbb{E}(X_{2n}^\ell) = Q_\ell(n), \qquad n \in \mathbb{N}, \qquad \mathbb{E}(X_{2n-1}^\ell) = Q_\ell(n), \qquad n \in \mathbb{N},$$

where $Q_\ell$ is a polynomial of degree $\ell$. Using the independence of $X_{2n}$ and $X_{2n-1}$, for $\ell_1 + \ell_2 \leqslant \ell$, $\ell_1, \ell_2 \in \mathbb{Z}_+$, we have

$$\mathbb{E}(X_{2n}^{\ell_1} X_{2n-1}^{\ell_2}) = \mathbb{E}(X_{2n}^{\ell_1}) \mathbb{E}(X_{2n-1}^{\ell_2}) = Q_{\ell_1}(n) Q_{\ell_2}(n) = \mathrm{O}(n^\ell), \qquad n \in \mathbb{N},$$

as desired. Similarly,

$$\mathbb{E}(X_{2n-1}^{\ell_1} X_{2n-2}^{\ell_2}) = \mathbb{E}(X_{2n-1}^{\ell_1}) \mathbb{E}(X_{2n-2}^{\ell_2}) = Q_{\ell_1}(n) Q_{\ell_2}(n-1) = \mathrm{O}(n^\ell), \qquad n \in \mathbb{N}.$$

Hence we have the assertion.

**Second proof.** It is enough to prove that there exists some $c_\ell \in \mathbb{R}_+$ such that $\mathbb{E}(X_n^{\ell_1} X_{n-1}^{\ell_2}) \leqslant c_\ell n^\ell$ for all $n \in \mathbb{N}$ and $\ell_1, \ell_2 \in \mathbb{Z}_+$ with $\ell_1 + \ell_2 = \ell$. Let us introduce the notation

$$\boldsymbol{X}_n^{(k)} := \left[X_n^k \quad X_n^{k-1} X_{n-1} \quad X_n^{k-2} X_{n-1}^2 \quad \cdots \quad X_n X_{n-1}^{k-1} \quad X_{n-1}^k\right]^\top \in \mathbb{R}_+^{k+1}, \qquad n, k \in \mathbb{N}.$$

First we check that

$$(9.12) \qquad \mathbb{E}(\boldsymbol{X}_n^{(k)} \mid \mathcal{F}_{n-1}) = A_k \boldsymbol{X}_{n-1}^{(k)} + \sum_{j=1}^{k-1} B_{k,j} \boldsymbol{X}_{n-1}^{(j)} + \boldsymbol{\mu}_k, \qquad n \in \mathbb{N}, \quad k \in \{1, \ldots, \ell\},$$

52

where

$$
A_k := \begin{bmatrix}
\alpha^k & \binom{k}{1}\alpha^{k-1}\beta & \cdots & \binom{k}{k-1}\alpha\beta^{k-1} & \beta^k \\
\alpha^{k-1} & \binom{k-1}{1}\alpha^{k-2}\beta & \cdots & \beta^{k-1} & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
\alpha & \beta & \cdots & 0 & 0 \\
1 & 0 & \cdots & 0 & 0
\end{bmatrix} \in \mathbb{R}_+^{(k+1)\times(k+1)}
$$

and $B_{k,j} \in \mathbb{R}_+^{(k+1)\times(j+1)}$ are appropriate matrices of which the entries are non-negative and depend only on $\alpha$ and the moments of $\varepsilon_1$ of order less than or equal to $(k-j)$ and

$$
\boldsymbol{\mu}_k := \begin{bmatrix} \mathbb{E}(\varepsilon_1^k) & 0 & \cdots & 0 \end{bmatrix}^\top \in \mathbb{R}_+^{k+1}.
$$

For a better understanding, first we give a proof for (9.12) in the case of $k = 1$ and $k = 2$. If $k = 1$, then

$$
\mathbb{E}(\boldsymbol{X}_n^{(1)} \,|\, \mathcal{F}_{n-1}) = \begin{bmatrix} \mathbb{E}(X_n \,|\, \mathcal{F}_{n-1}) \\ \mathbb{E}(X_{n-1} \,|\, \mathcal{F}_{n-1}) \end{bmatrix} = \begin{bmatrix} \alpha & \beta \\ 1 & 0 \end{bmatrix} \begin{bmatrix} X_{n-1} \\ X_{n-2} \end{bmatrix} + \begin{bmatrix} \mu \\ 0 \end{bmatrix} = A_1 \boldsymbol{X}_{n-1}^{(1)} + \boldsymbol{\mu}_1, \qquad n \in \mathbb{N}.
$$

If $k = 2$, then, by (9.9), we have

$$
\mathbb{E}(X_n^2 \,|\, \mathcal{F}_{n-1}) = \mathbb{E}\left( \sum_{\substack{k_1+k_2+k_3=2, \\ k_1,k_2,k_3 \in \mathbb{Z}_+}} \frac{2!}{k_1!k_2!k_3!} \left( \sum_{j=1}^{X_{n-1}} \xi_{n,j} \right)^{k_1} \left( \sum_{j=1}^{X_{n-2}} \eta_{n,j} \right)^{k_2} \varepsilon_n^{k_3} \;\middle|\; \mathcal{F}_{n-1} \right)
$$

$$
= \alpha X_{n-1} + \alpha^2(X_{n-1}^2 - X_{n-1}) + \beta X_{n-2} + \beta^2(X_{n-2}^2 - X_{n-2}) + 2\alpha\beta X_{n-1}X_{n-2}
$$

$$
+ 2\alpha X_{n-1}\mathbb{E}(\varepsilon_1) + 2\beta X_{n-2}\mathbb{E}(\varepsilon_1) + \mathbb{E}(\varepsilon_1^2), \qquad n \in \mathbb{N},
$$

and hence, using also that

$$
\mathbb{E}(X_n X_{n-1} \,|\, \mathcal{F}_{n-1}) = X_{n-1}\mathbb{E}(X_n \,|\, \mathcal{F}_{n-1}) = \alpha X_{n-1}^2 + \beta X_{n-1}X_{n-2} + X_{n-1}\mathbb{E}(\varepsilon_1), \qquad n \in \mathbb{N},
$$

we have

$$
\mathbb{E}(\boldsymbol{X}_n^{(2)} \,|\, \mathcal{F}_{n-1}) = \mathbb{E}\left( \begin{bmatrix} X_n^2 \\ X_n X_{n-1} \\ X_{n-1}^2 \end{bmatrix} \;\middle|\; \mathcal{F}_{n-1} \right) = A_2 \begin{bmatrix} X_{n-1}^2 \\ X_{n-1}X_{n-2} \\ X_{n-2}^2 \end{bmatrix} + B_{2,1} \begin{bmatrix} X_{n-1} \\ X_{n-2} \end{bmatrix} + \boldsymbol{\mu}_2, \qquad n \in \mathbb{N},
$$

where

$$
A_2 = \begin{bmatrix} \alpha^2 & 2\alpha\beta & \beta^2 \\ \alpha & \beta & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad \text{and} \quad B_{2,1} = \begin{bmatrix} \alpha\beta + 2\alpha\mathbb{E}(\varepsilon_1) & \alpha\beta + 2\beta\mathbb{E}(\varepsilon_1) \\ \mathbb{E}(\varepsilon_1) & 0 \\ 0 & 0 \end{bmatrix}
$$

53

as desired. In the general case using part (i) of Lemma 9.2 one can prove (9.12) (giving also explicit forms for the matrices $B_{k,j}$).

Taking expectation of (9.12), we have

$$(9.13) \qquad \mathbb{E}(\boldsymbol{X}_n^{(k)}) = A_k \, \mathbb{E}(\boldsymbol{X}_{n-1}^{(k)}) + \sum_{j=1}^{k-1} B_{k,j} \, \mathbb{E}(\boldsymbol{X}_{n-1}^{(j)}) + \boldsymbol{\mu}_k, \qquad n \in \mathbb{N}, \quad k \in \{1, \dots, \ell\}.$$

For a $d$-dimensional vector $\mathbf{v} = (v_i)_{i=1}^d \in \mathbb{R}^d$ and a $d \times d$ matrix $M = (m_{i,j})_{i,j=1}^d \in \mathbb{R}^{d \times d}$, let us introduce the notations

$$\|\mathbf{v}\|_\infty := \max_{1 \leqslant i \leqslant d} |v_i| \qquad \text{and} \qquad \|M\|_\infty := \max_{1 \leqslant i \leqslant d} \sum_{j=1}^d |m_{i,j}|.$$

By the binomial theorem one can easily have $\|A_k\|_\infty = 1$, $k \in \{1, \dots, \ell\}$. We prove the statement using a double induction with respect to $k \in \{1, \dots, \ell\}$ and $n \in \mathbb{N}$. First we show that the statement holds for $k = 1$ using induction with respect to $n$. Namely, we show that

$$\|\mathbb{E}(\boldsymbol{X}_n^{(1)})\|_\infty \leqslant c_1 n, \qquad n \in \mathbb{N},$$

where $c_1 := \|\boldsymbol{\mu}_1\|_\infty$. If $n = 1$, then

$$\mathbb{E}(\boldsymbol{X}_1^{(1)}) = \begin{bmatrix} \mathbb{E}(X_1) \\ \mathbb{E}(X_0) \end{bmatrix} = \begin{bmatrix} \mathbb{E}(\varepsilon_1) \\ 0 \end{bmatrix} = \boldsymbol{\mu}_1,$$

which implies that $\|\mathbb{E}(\boldsymbol{X}_1^{(1)})\|_\infty = c_1$. Let us suppose now that $\|\mathbb{E}(\boldsymbol{X}_m^{(1)})\|_\infty \leqslant c_1 m$ holds for $m \in \{1, \dots, n-1\}$ with $n \geqslant 2$. Then, (9.13),

$$\|\mathbb{E}(\boldsymbol{X}_n^{(1)})\|_\infty = \|A \, \mathbb{E}(\boldsymbol{X}_{n-1}^{(1)}) + \boldsymbol{\mu}_1\|_\infty \leqslant \|A \, \mathbb{E}(\boldsymbol{X}_{n-1}^{(1)})\|_\infty + \|\boldsymbol{\mu}_1\|_\infty$$

$$\leqslant \|A\|_\infty \|\mathbb{E}(\boldsymbol{X}_{n-1}^{(1)})\|_\infty + \|\boldsymbol{\mu}_1\|_\infty \leqslant c_1(n-1) + c_1 = c_1 n,$$

as desired.

Let us suppose now that the statement holds for $j = 1, \dots, \ell - 1$, i.e.,

$$\|\mathbb{E}(\boldsymbol{X}_n^{(j)})\|_\infty \leqslant c_j n^j, \qquad n \in \mathbb{N}, \quad j \in \{1, \dots, \ell - 1\}.$$

Next, using induction with respect to $n \in \mathbb{N}$ we prove that

$$\|\mathbb{E}(\boldsymbol{X}_n^{(\ell)})\|_\infty \leqslant c_\ell n^\ell, \qquad n \in \mathbb{N},$$

where

$$c_\ell := \sum_{j=1}^{\ell-1} c_j \|B_{\ell,j}\|_\infty + \|\boldsymbol{\mu}_\ell\|_\infty.$$

If $n = 1$, then, using that $X_0 = 0$ and $X_1 = \varepsilon_1$, we have

$$\mathbb{E}(\boldsymbol{X}_1^{(\ell)}) = \begin{bmatrix} \mathbb{E}(\varepsilon_1^\ell) & 0 & \cdots & 0 \end{bmatrix}^\top = \boldsymbol{\mu}_\ell,$$

54

which yields that $\|\mathbb{E}(\boldsymbol{X}_1^{(\ell)})\|_\infty = \|\boldsymbol{\mu}_\ell\|_\infty \leqslant c_\ell.$ Let us suppose now that

$$\|\mathbb{E}(\boldsymbol{X}_m^{(\ell)})\|_\infty \leqslant c_\ell m^\ell, \qquad m \in \{1, \dots, n-1\},$$

where $n \geqslant 2.$ Then, by (9.13),

$$\|\mathbb{E}(\boldsymbol{X}_n^{(\ell)})\|_\infty \leqslant \|A_\ell\|_\infty \|\mathbb{E}(\boldsymbol{X}_{n-1}^{(\ell)})\|_\infty + \sum_{j=1}^{\ell-1} \|B_{\ell,j}\|_\infty \|\mathbb{E}(\boldsymbol{X}_{n-1}^{(j)})\|_\infty + \|\boldsymbol{\mu}_\ell\|_\infty$$

$$\leqslant c_\ell(n-1)^\ell + \sum_{j=1}^{\ell-1} \|B_{\ell,j}\|_\infty c_j (n-1)^j + \|\boldsymbol{\mu}_\ell\|_\infty$$

$$\leqslant c_\ell(n-1)^\ell + \left( \sum_{j=1}^{\ell-1} c_j \|B_{\ell,j}\|_\infty + \|\boldsymbol{\mu}_\ell\|_\infty \right)(n-1)^{\ell-1}$$

$$= c_\ell(n-1)^{\ell-1}(n-1+1)$$

$$\leqslant c_\ell n^\ell,$$

as desired. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

**9.1 Corollary.** *Let* $(X_k)_{k \geqslant -1}$ *be an INAR(2) process with autoregressive parameters* $(\alpha, \beta) \in [0,1]^2$ *such that* $\alpha + \beta = 1$ *(hence it is unstable). Suppose that* $X_0 = X_{-1} = 0$ *and* $\mathbb{E}(\varepsilon_1^\ell) < \infty$ *with some* $\ell \in \mathbb{N}.$ *Then*

$$\mathbb{E}(X_k^i) = \mathrm{O}(k^i), \qquad \mathbb{E}(M_k^i) = \mathrm{O}(k^{\lfloor i/2 \rfloor}), \qquad \mathbb{E}(U_k^i) = \mathrm{O}(k^i), \qquad \mathbb{E}(V_k^{2j}) = \mathrm{O}(k^j), \qquad k \in \mathbb{N},$$

*for* $i, j \in \mathbb{Z}_+$ *with* $i \leqslant \ell$ *and* $2j \leqslant \ell.$

**Proof.** The estimate $\mathbb{E}(X_k^i) = \mathrm{O}(k^i)$ readily follows by Lemma 9.4. Next we turn to prove $\mathbb{E}(M_k^i) = \mathrm{O}(k^{\lfloor i/2 \rfloor}).$ Using (9.5) and that the random variables $\{\xi_{n,j}, \eta_{n,j}, \varepsilon_n : j \in \mathbb{N}\}$ are independent of each other and of the $\sigma$-algebra $\mathcal{F}_{n-1},$ we have for all $n \in \mathbb{N},$

$$\mathbb{E}(M_n^i \mid \mathcal{F}_{n-1}) = \sum_{\substack{i_1+i_2+i_3=i, \\ i_1,i_2,i_3 \in \mathbb{Z}_+}} \frac{i!}{i_1! i_2! i_3!} \mathbb{E}\left( \left( \sum_{j=1}^M (\xi_{n,j} - \mathbb{E}(\xi_{n,j})) \right)^{i_1} \right) \Bigg|_{M=X_{n-1}}$$

$$\times \mathbb{E}\left( \left( \sum_{j=1}^N (\eta_{n,j} - \mathbb{E}(\eta_{n,j})) \right)^{i_2} \right) \Bigg|_{N=X_{n-2}} \mathbb{E}\left( (\varepsilon_n - \mathbb{E}(\varepsilon_n))^{i_3} \right).$$

By part (ii) of Lemma 9.2, there exist polynomials $Q_{i_1},\ i_1 \in \mathbb{N},$ of degree at most $i_1/2,$ and $\widetilde{Q}_{i_2},\ i_2 \in \mathbb{N},$ of degree at most $i_2/2$ such that

$$\mathbb{E}(M_n^i \mid \mathcal{F}_{n-1}) = \sum_{\substack{i_1+i_2+i_3=i, \\ i_1,i_2,i_3 \in \mathbb{Z}_+}} \frac{i!}{i_1! i_2! i_3!} Q_{i_1}(X_{n-1}) \widetilde{Q}_{i_2}(X_{n-2}) \mathbb{E}\left( (\varepsilon_1 - \mathbb{E}(\varepsilon_1))^{i_3} \right).$$

55

Hence

$$\mathbb{E}(M_n^i) = \sum_{\substack{i_1+i_2+i_3=i, \\ i_1,i_2,i_3 \in \mathbb{Z}_+}} \frac{i!}{i_1!i_2!i_3!} \mathbb{E}\left(Q_{i_1}(X_{n-1})\widetilde{Q}_{i_2}(X_{n-2})\right) \mathbb{E}\left((\varepsilon_1 - \mathbb{E}(\varepsilon_1))^{i_3}\right), \qquad n \in \mathbb{N}.$$

Clearly, $Q_{i_1}(X_{k-1})\widetilde{Q}_{i_2}(X_{k-2}) = Q_{i_1+i_2}^*(X_{k-1}, X_{k-2})$, where $Q_{i_1+i_2}^*$ is a polynomial of two variables having degree at most $(i_1 + i_2)/2 \leqslant i/2$, and hence, at most $\lfloor i/2 \rfloor$. By Lemma 9.4, there exists a constant $c_{Q_{i_1+i_2}^*}$ such that $\mathbb{E}\left(|Q_{i_1+i_2}^*(X_{k-1}, X_{k-2})|\right) \leqslant c_{Q_{i_1+i_2}^*}(k-1)^{\lfloor i/2 \rfloor}$. Hence

$$|\mathbb{E}(M_k^i)| \leqslant (k-1)^{\lfloor i/2 \rfloor} \sum_{\substack{i_1+i_2+i_3=i, \\ i_1,i_2,i_3 \in \mathbb{Z}_+}} \frac{i!}{i_1!i_2!i_3!} c_{Q_{i_1+i_2}^*} \left|\mathbb{E}\left((\varepsilon_1 - \mathbb{E}(\varepsilon_1))^{i_3}\right)\right|$$

for all $k \in \mathbb{N}$, as desired.

Next we turn to prove $\mathbb{E}(U_k^i) = \mathrm{O}(k^i)$, $i, k \in \mathbb{N}$ with $i \leqslant \ell$. First note that, by power mean inequality, for all $i \in \mathbb{N}$,

$$\frac{a+b}{2} \leqslant \left(\frac{a^i + b^i}{2}\right)^{\frac{1}{i}}, \qquad a, b \geqslant 0,$$

yielding that $(a + b)^i \leqslant 2^{i-1}(a^i + b^i)$, $a, b \geqslant 0$. Hence, by Lemma 9.4,

$$\mathbb{E}(U_k^i) = \mathbb{E}((X_k + \beta X_{k-1})^i) \leqslant 2^{i-1}(\mathbb{E}(X_k^i) + \beta^i \mathbb{E}(X_{k-1}^i)) \leqslant 2^{i-1}(P_i(k) + \beta^i P_i(k-1)),$$

where $P_i$ is a polynomial of degree at most $i$, which yields that $\mathbb{E}(U_k^i) = \mathrm{O}(k^i)$.

Finally, for $2j \leqslant \ell$, $j \in \mathbb{Z}_+$, we prove $\mathbb{E}(V_k^{2j}) = \mathrm{O}(k^j)$, $k \in \mathbb{N}$, using induction in $k$. By the recursion $V_k = -\beta V_{k-1} + M_k + \mu$, $k \in \mathbb{N}$, we have $\mathbb{E}(V_k) = -\beta \mathbb{E}(V_{k-1}) + \mu$, $k \in \mathbb{N}$, with initial value $\mathbb{E}(V_0) = 0$, hence

$$\mathbb{E}(V_k) = \mu \sum_{i=0}^{k-1}(-\beta)^i, \qquad k \in \mathbb{N},$$

which yields that $|\mathbb{E}(V_k)| = \mathrm{O}(1)$. Indeed, for all $k \in \mathbb{N}$,

$$\left|\sum_{i=0}^{k-1}(-\beta)^i\right| \leqslant \begin{cases} \frac{1}{1-\beta} & \text{if } 0 \leqslant \beta < 1, \\ 1 & \text{if } \beta = 1, \end{cases}$$

where the inequality for the case $\beta = 1$ follows by that the sequence of partial sums in question is nothing else but the alternating one $1, 0, 1, 0, 1, 0, \ldots$. Let us introduce the notation $\widetilde{V}_k := V_k - \mathbb{E}(V_k)$, $k \in \mathbb{N}$. Since, by the triangular inequality for the $L_{2j}$-norm,

$$\left(\mathbb{E}(V_k^{2j})\right)^{\frac{1}{2j}} \leqslant \left(\mathbb{E}(\widetilde{V}_k^{2j})\right)^{\frac{1}{2j}} + \mathbb{E}(|V_k|),$$

and $|\mathbb{E}(V_k)| = \mathrm{O}(1)$, for proving $\mathbb{E}(V_k^{2j}) = \mathrm{O}(k^j)$, $k \in \mathbb{N}$, it is enough to show that $\mathbb{E}(\widetilde{V}_k^{2j}) = \mathrm{O}(k^j)$, $k \in \mathbb{N}$. Using again the recursion $V_k = -\beta V_{k-1} + M_k + \mu$, $k \in \mathbb{N}$, we get $\widetilde{V}_k = -\beta \widetilde{V}_{k-1} + M_k$, $k \in \mathbb{N}$. Hence

$$\left(\mathbb{E}(\widetilde{V}_k^{2j})\right)^{\frac{1}{2j}} \leqslant \beta \left(\mathbb{E}(\widetilde{V}_{k-1}^{2j})\right)^{\frac{1}{2j}} + \left(\mathbb{E}(M_k^{2j})\right)^{\frac{1}{2j}} = \left(\mathrm{O}((k-1)^j)\right)^{\frac{1}{2j}} + \left(\mathrm{O}(k^j)\right)^{\frac{1}{2j}} = \mathrm{O}(k^{1/2}),$$

where the first inequality follows by the triangular inequality for the $L_{2j}$-norm, and the second one by the induction hypothesis and that $\mathbb{E}(M_k^{2j}) = \mathrm{O}(k^j)$. Hence $\mathbb{E}(\widetilde{V}_k^{2j}) = \mathrm{O}(k^j)$, $k \in \mathbb{N}$, as desired. $\qquad\square$

**9.2 Corollary.** *Let* $(X_k)_{k \geqslant -1}$ *be an* INAR(2) *process with autoregressive parameters* $(\alpha, \beta) \in [0,1]^2$ *such that* $\alpha + \beta = 1$ *(hence it is unstable). Suppose that* $X_0 = X_{-1} = 0$ *and* $\mathbb{E}(\varepsilon_1^\ell) < \infty$ *with some* $\ell \in \mathbb{N}$. *Then*

(i) *for all* $i, j \in \mathbb{Z}_+$ *with* $\max\{i, j\} \leqslant \ell/2$, *and for all* $\kappa > i + \frac{j}{2} + 1$, *we have*

$$(9.14) \qquad n^{-\kappa} \sum_{k=1}^{n} |U_k^i V_k^j| \overset{\mathbb{P}}{\longrightarrow} 0 \qquad as \ n \to \infty,$$

(ii) *for all* $i, j \in \mathbb{Z}_+$ *with* $\max\{i, j\} \leqslant \ell$, *for all* $T > 0$, *and for all* $\kappa > i + \frac{j}{2} + \frac{i+j}{\ell}$, *we have*

$$(9.15) \qquad n^{-\kappa} \sup_{t \in [0,T]} |U_{\lfloor nt \rfloor}^i V_{\lfloor nt \rfloor}^j| \overset{\mathbb{P}}{\longrightarrow} 0 \qquad as \ n \to \infty,$$

(iii) *for all* $i, j \in \mathbb{Z}_+$ *with* $\max\{i, j\} \leqslant \ell/4$, *for all* $T > 0$, *and for all* $\kappa > i + \frac{j}{2} + \frac{1}{2}$, *we have*

$$(9.16) \qquad n^{-\kappa} \sup_{t \in [0,T]} \left| \sum_{k=1}^{\lfloor nt \rfloor} [U_k^i V_k^j - \mathbb{E}(U_k^i V_k^j \,|\, \mathcal{F}_{k-1})] \right| \overset{\mathbb{P}}{\longrightarrow} 0 \qquad as \ n \to \infty.$$

**Proof.** By Cauchy-Schwartz's inequality and Lemma 9.1, we have

$$\mathbb{E}\left(\sum_{k=1}^{n} |U_k^i V_k^j|\right) \leqslant \sum_{k=1}^{n} \sqrt{\mathbb{E}(U_k^{2i})\,\mathbb{E}(V_k^{2j})} = \sum_{k=1}^{n} \sqrt{\mathrm{O}(k^{2i})\,\mathrm{O}(k^j)} = \sum_{k=1}^{n} \mathrm{O}(k^{i+j/2}) = \mathrm{O}(n^{1+i+j/2}).$$

Using Slutsky's lemma this implies (9.14).

Now we turn to prove (9.15). First note that

$$(9.17) \qquad \sup_{t \in [0,T]} |U_{\lfloor nt \rfloor}^i V_{\lfloor nt \rfloor}^j| \leqslant \sup_{t \in [0,T]} |U_{\lfloor nt \rfloor}^i| \sup_{t \in [0,T]} |V_{\lfloor nt \rfloor}^j|,$$

and for all $\varepsilon > 0$ and $\delta > 0$, we have, by Markov's inequality,

$$\mathbb{P}\left(n^{-\varepsilon} \sup_{t\in[0,T]} |U^i_{\lfloor nt \rfloor}| > \delta\right) = \mathbb{P}\left(n^{-\ell\varepsilon/i} \sup_{t\in[0,T]} |U^\ell_{\lfloor nt \rfloor}| > \delta^{\ell/i}\right) \leqslant \sum_{k=1}^{\lfloor nT \rfloor} \mathbb{P}(U^\ell_k > \delta^{\ell/i} n^{\ell\varepsilon/i})$$

$$\leqslant \sum_{k=1}^{\lfloor nT \rfloor} \frac{\mathbb{E}(U^\ell_k)}{\delta^{\ell/i} n^{\ell\varepsilon/i}} = \sum_{k=1}^{\lfloor nT \rfloor} \frac{\mathrm{O}(k^\ell)}{\delta^{\ell/i} n^{\ell\varepsilon/i}} = \mathrm{O}(n^{\ell+1-\ell\varepsilon/i}), \qquad i \in \{1, 2, \ldots, \ell\},$$

and

$$\mathbb{P}\left(n^{-\varepsilon} \sup_{t\in[0,T]} |V^j_{\lfloor nt \rfloor}| > \delta\right) = \mathbb{P}\left(n^{-\ell\varepsilon/j} \sup_{t\in[0,T]} |V^\ell_{\lfloor nt \rfloor}| > \delta^{\ell/j}\right) \leqslant \sum_{k=1}^{\lfloor nT \rfloor} \mathbb{P}(|V^\ell_k| > \delta^{\ell/j} n^{\ell\varepsilon/j})$$

$$\leqslant \sum_{k=1}^{\lfloor nT \rfloor} \frac{\mathbb{E}(|V^\ell_k|)}{\delta^{\ell/j} n^{\ell\varepsilon/j}} \leqslant \sum_{k=1}^{\lfloor nT \rfloor} \frac{\sqrt{\mathbb{E}(V^{2\ell}_k)}}{\delta^{\ell/j} n^{\ell\varepsilon/j}} = \sum_{k=1}^{\lfloor nT \rfloor} \frac{\mathrm{O}(k^{\ell/2})}{\delta^{\ell/j} n^{\ell\varepsilon/j}} = \mathrm{O}(n^{\ell/2+1-\ell\varepsilon/j}), \quad j \in \{1, 2, \ldots, \ell\}.$$

Hence, if $\ell + 1 - \ell\varepsilon/i < 0$, i.e., $\varepsilon > \frac{\ell+1}{\ell} i$, then

$$n^{-\varepsilon} \sup_{t\in[0,T]} |U^i_{\lfloor nt \rfloor}| \xrightarrow{\mathbb{P}} 0 \qquad \text{as } n \to \infty,$$

and if $\ell/2 + 1 - \ell\varepsilon/j < 0$, i.e., $\varepsilon > \frac{\ell/2+1}{\ell} j$, then

$$n^{-\varepsilon} \sup_{t\in[0,T]} |V^j_{\lfloor nt \rfloor}| \xrightarrow{\mathbb{P}} 0 \qquad \text{as } n \to \infty.$$

By (9.17), we get (9.15).

Finally, we show (9.16). Applying Doob's maximal inequality (see, e.g., Revuz and Yor [30, Chapter II, Theorem 1.7]) for the martingale

$$\sum_{k=1}^{n} \left[U^i_k V^j_k - \mathbb{E}(U^i_k V^j_k \mid \mathcal{F}_{k-1})\right], \qquad n \in \mathbb{N},$$

(with the filtration $(\mathcal{F}_k)_{k\in\mathbb{N}}$) and then (5.7), we obtain

$$\mathbb{E}\left(\sup_{t\in[0,T]} \left(\sum_{k=1}^{\lfloor nt \rfloor} \left[U^i_k V^j_k - \mathbb{E}(U^i_k V^j_k \mid \mathcal{F}_{k-1})\right]\right)^2\right) \leqslant 4\,\mathbb{E}\left(\left(\sum_{k=1}^{\lfloor nT \rfloor} \left[U^i_k V^j_k - \mathbb{E}(U^i_k V^j_k \mid \mathcal{F}_{k-1})\right]\right)^2\right)$$

$$\leqslant 4 \sum_{k=1}^{\lfloor nT \rfloor} \mathbb{E}(U^{2i}_k V^{2j}_k) = \sum_{k=1}^{\lfloor nT \rfloor} \mathrm{O}(k^{2i+j}) = \mathrm{O}(n^{2i+j+1}),$$

since $\mathbb{E}(U^{2i}_k V^{2j}_k) \leqslant \sqrt{\mathbb{E}(U^{4i}_k)\,\mathbb{E}(V^{4j}_k)} = \mathrm{O}(k^{2i+j})$ by Corollary 9.1. $\qquad \square$

**9.3 Remark.** We note that in the special case $(\ell, i, j) = (2, 1, 0)$, we also get

(9.18)
$$n^{-\kappa} \sup_{t \in [0,T]} U_{\lfloor nt \rfloor} \xrightarrow{\mathbb{P}} 0 \qquad \text{as } n \to \infty \text{ for } \kappa > 1.$$

Indeed, by (5.8), we have

(9.19)
$$U_n = \sum_{k=1}^{n} (M_k + \mu), \qquad n \in \mathbb{N},$$

and hence convergence (9.18) will follow from

(9.20)
$$n^{-\kappa} \sup_{t \in [0,T]} \left| \sum_{k=1}^{\lfloor nt \rfloor} M_k \right| \xrightarrow{\mathbb{P}} 0 \qquad \text{as } n \to \infty \text{ for all } \kappa > 1.$$

Doob's maximal inequality (see, e.g., Revuz and Yor [30, Chapter II, Theorem 1.7]) for the martingale $\sum_{i=1}^{k} M_i$, $k \in \mathbb{N}$, (with the filtration $(\mathcal{F}_k)_{k \in \mathbb{N}}$) gives

$$\mathbb{E} \left( \sup_{t \in [0,T]} \left( \sum_{k=1}^{\lfloor nt \rfloor} M_k \right)^2 \right) \leqslant 4 \, \mathbb{E} \left( \left( \sum_{k=1}^{\lfloor nT \rfloor} M_k \right)^2 \right) = 4 \sum_{k=1}^{\lfloor nT \rfloor} \mathbb{E}(M_k^2) = \mathrm{O}(n^2),$$

since $\mathbb{E}(M_k^2) = \mathrm{O}(k)$ by Corollary 9.1. This implies (9.20), hence (9.18).

However, it turns out that we do not need this stronger statement. $\qquad \square$

# Appendicies

# A    Classification of INAR(2) processes

An INAR(2) process is called *positively regular* if there is a positive integer $k$ such that the entries of $A^k$ are positive (see Kesten and Stigum [23]). If $\alpha > 0$ and $\beta > 0$ then the INAR(2) process is positively regular, since

$$A = \begin{bmatrix} \alpha & \beta \\ 1 & 0 \end{bmatrix}, \qquad A^2 = \begin{bmatrix} \alpha^2 + \beta & \alpha\beta \\ \alpha & \beta \end{bmatrix}.$$

If $\alpha = 0$, then

$$A^{2k+1} = \beta^k A = \beta^k \begin{bmatrix} 0 & \beta \\ 1 & 0 \end{bmatrix}, \qquad A^{2k} = \beta^k \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad k \in \mathbb{Z}_+,$$

hence the process is not positively regular. If $\beta = 0$, then

$$A^k = \alpha^{k-1} A = \alpha^{k-1} \begin{bmatrix} \alpha & 0 \\ 1 & 0 \end{bmatrix}, \qquad k \in \mathbb{N},$$

hence the process is not positively regular. Consequently, an INAR(2) process is positively regular if and only if $\alpha > 0$ and $\beta > 0$.

An INAR(2) process is called *decomposable* if the matrix $A$ is decomposable (see Kesten and Stigum [25]). Note that an INAR(2) process is decomposable if and only if the matrix $A$ is reducible (see Horn and Johnson [16, Definition 6.2.21]), that is, there exists a permutation matrix $P \in \mathbb{R}^{2 \times 2}$ such that

$$P^\top A P = \begin{bmatrix} b & c \\ 0 & d \end{bmatrix},$$

where $b, c, d \in \mathbb{R}$. Since

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \alpha & \beta \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ \beta & \alpha \end{bmatrix},$$

we get an INAR(2) process is decomposable if and only if $\beta = 0$. Moreover, an INAR(2) process is indecomposable but not positively regular if and only if $\alpha = 0$ and $\beta > 0$.

Note that an INAR(2) process is positively regular if and only if the matrix $A$ is primitive (see Horn and Johnson [16, Definition 8.5.0 and Theorem 8.5.2]), so this case can also be called *primitive* (see Barczy et al. [5, Definition 2.4]). Further we remark that the not positively regular case is also called *non-primitive*.

# B   A version of the continuous mapping theorem

A function $f : \mathbb{R}_+ \to \mathbb{R}^d$ is called *càdlàg* if it is right continuous with left limits. Let $\mathsf{D}(\mathbb{R}_+, \mathbb{R}^d)$ and $\mathsf{C}(\mathbb{R}_+, \mathbb{R}^d)$ denote the space of all $\mathbb{R}^d$-valued càdlàg and continuous functions on $\mathbb{R}_+$, respectively. Let $\mathcal{B}(\mathsf{D}(\mathbb{R}_+, \mathbb{R}^d))$ denote the Borel $\sigma$-algebra on $\mathsf{D}(\mathbb{R}_+, \mathbb{R}^d)$ for the metric defined in Jacod and Shiryaev [20, Chapter VI, (1.26)] (with this metric $\mathsf{D}(\mathbb{R}_+, \mathbb{R}^d)$ is a complete and separable metric space and the topology induced by this metric is the so-called Skorokhod topology). For $\mathbb{R}^d$-valued stochastic processes $(\boldsymbol{\mathcal{Y}}_t)_{t \in \mathbb{R}_+}$ and $(\boldsymbol{\mathcal{Y}}_t^{(n)})_{t \in \mathbb{R}_+}$, $n \in \mathbb{N}$, with càdlàg paths we write $\boldsymbol{\mathcal{Y}}^{(n)} \xrightarrow{\mathcal{L}} \boldsymbol{\mathcal{Y}}$ if the distribution of $\boldsymbol{\mathcal{Y}}^{(n)}$ on the space $(\mathsf{D}(\mathbb{R}_+, \mathbb{R}), \mathcal{B}(\mathsf{D}(\mathbb{R}_+, \mathbb{R}^d)))$ converges weakly to the distribution of $\boldsymbol{\mathcal{Y}}$ on the space $(\mathsf{D}(\mathbb{R}_+, \mathbb{R}), \mathcal{B}(\mathsf{D}(\mathbb{R}_+, \mathbb{R}^d)))$ as $n \to \infty$. Concerning the notation $\xrightarrow{\mathcal{L}}$ we note that if $\xi$ and $\xi_n$, $n \in \mathbb{N}$, are random elements with values in a metric space $(E, d)$, then we also denote by $\xi_n \xrightarrow{\mathcal{L}} \xi$ the weak convergence of the distributions of $\xi_n$ on the space $(E, \mathcal{B}(E))$ towards the distribution of $\xi$ on the space $(E, \mathcal{B}(E))$ as $n \to \infty$, where $\mathcal{B}(E)$ denotes the Borel $\sigma$-algebra on $E$ induced by the given metric $d$.

The following version of continuous mapping theorem can be found for example in Kallenberg [21, Theorem 3.27].

**B.1 Lemma.** *Let $(S, d_S)$ and $(T, d_T)$ be metric spaces and $(\xi_n)_{n \in \mathbb{N}}$, $\xi$ be random elements with values in $S$ such that $\xi_n \xrightarrow{\mathcal{L}} \xi$ as $n \to \infty$. Let $f : S \to T$ and $f_n : S \to T$, $n \in \mathbb{N}$, be*

*measurable mappings and* $C \in \mathcal{B}(S)$ *such that* $\mathbb{P}(\xi \in C) = 1$ *and* $\lim_{n \to \infty} d_T(f_n(s_n), f(s)) = 0$ *if* $\lim_{n \to \infty} d_S(s_n, s) = 0$ *and* $s \in C$. *Then* $f_n(\xi_n) \xrightarrow{\mathcal{L}} f(\xi)$ *as* $n \to \infty$.

For the case $S := \mathsf{D}(\mathbb{R}_+, \mathbb{R}^d)$ and $T := \mathbb{R}^q$ ($T := \mathsf{D}(\mathbb{R}_+, \mathbb{R}^q)$), where $d, q \in \mathbb{N}$, we formulate a consequence of Lemma B.1.

For functions $f$ and $f_n$, $n \in \mathbb{N}$, in $\mathsf{D}(\mathbb{R}_+, \mathbb{R}^d)$, we write $f_n \xrightarrow{\mathrm{lu}} f$ if $(f_n)_{n \in \mathbb{N}}$ converges to $f$ locally uniformly, i.e., if $\sup_{t \in [0,T]} \|f_n(t) - f(t)\| \to 0$ as $n \to \infty$ for all $T > 0$. For measurable mappings $\Phi : \mathsf{D}(\mathbb{R}_+, \mathbb{R}^d) \to \mathbb{R}^q$ ($\mathsf{D}(\mathbb{R}_+, \mathbb{R}^q)$) and $\Phi_n : \mathsf{D}(\mathbb{R}_+, \mathbb{R}^d) \to \mathbb{R}^q$ ($\mathsf{D}(\mathbb{R}_+, \mathbb{R}^q)$), $n \in \mathbb{N}$, we will denote by $C_{\Phi, (\Phi_n)_{n \in \mathbb{N}}}$ the set of all functions $f \in \mathsf{C}(\mathbb{R}_+, \mathbb{R}^d)$ such that $\Phi_n(f_n) \to \Phi(f)$ ($\xrightarrow{\mathrm{lu}} \Phi(f)$) whenever $f_n \xrightarrow{\mathrm{lu}} f$ with $f_n \in \mathsf{D}(\mathbb{R}_+, \mathbb{R}^d)$, $n \in \mathbb{N}$.

We will use the following version of the continuous mapping theorem several times, see, e.g., Barczy et al. [4, Lemma 4.2] and Ispány and Pap [19, Lemma 3.1].

**B.2 Lemma.** *Let* $d, q \in \mathbb{N}$, *and* $(\mathcal{U}_t)_{t \in \mathbb{R}_+}$ *and* $(\mathcal{U}_t^{(n)})_{t \in \mathbb{R}_+}$, $n \in \mathbb{N}$, *be* $\mathbb{R}^d$-*valued stochastic processes with càdlàg paths such that* $\mathcal{U}^{(n)} \xrightarrow{\mathcal{L}} \mathcal{U}$. *Let* $\Phi : \mathsf{D}(\mathbb{R}_+, \mathbb{R}^d) \to \mathbb{R}^q$ ($\mathsf{D}(\mathbb{R}_+, \mathbb{R}^q)$) *and* $\Phi_n : \mathsf{D}(\mathbb{R}_+, \mathbb{R}^d) \to \mathbb{R}^q$ ($\mathsf{D}(\mathbb{R}_+, \mathbb{R}^q)$), $n \in \mathbb{N}$, *be measurable mappings such that there exists* $C \subset C_{\Phi, (\Phi_n)_{n \in \mathbb{N}}}$ *with* $C \in \mathcal{B}(\mathsf{D}(\mathbb{R}_+, \mathbb{R}^d))$ *and* $\mathbb{P}(\mathcal{U} \in C) = 1$. *Then* $\Phi_n(\mathcal{U}^{(n)}) \xrightarrow{\mathcal{L}} \Phi(\mathcal{U})$.

In order to apply Lemma B.2, we will use the following statement several times.

**B.3 Lemma.** *Let* $d, p, q \in \mathbb{N}$, $h : \mathbb{R}^d \to \mathbb{R}^q$ *be a continuous function and* $K : [0, 1] \times \mathbb{R}^{2d} \to \mathbb{R}^p$ *be a function such that for all* $R > 0$ *there exists* $C_R > 0$ *such that*

(B.1)
$$\|K(s, x) - K(t, y)\| \leqslant C_R (|t - s| + \|x - y\|)$$

*for all* $s, t \in [0, 1]$ *and* $x, y \in \mathbb{R}^{2d}$ *with* $\|x\| \leqslant R$ *and* $\|y\| \leqslant R$. *Moreover, let us define the mappings* $\Phi, \Phi_n : \mathsf{D}(\mathbb{R}_+, \mathbb{R}^d) \to \mathbb{R}^{q+p}$, $n \in \mathbb{N}$, *by*

$$\Phi_n(f) := \left( h(f(1)), \frac{1}{n} \sum_{k=1}^n K\left( \frac{k}{n}, f\left( \frac{k}{n} \right), f\left( \frac{k-1}{n} \right) \right) \right),$$

$$\Phi(f) := \left( h(f(1)), \int_0^1 K(u, f(u), f(u)) \, \mathrm{d}u \right)$$

*for all* $f \in \mathsf{D}(\mathbb{R}_+, \mathbb{R}^d)$. *Then the mappings* $\Phi$ *and* $\Phi_n$, $n \in \mathbb{N}$, *are measurable, and* $C_{\Phi, (\Phi_n)_{n \in \mathbb{N}}} = \mathsf{C}(\mathbb{R}_+, \mathbb{R}^d) \in \mathcal{B}(\mathsf{D}(\mathbb{R}_+, \mathbb{R}^d))$.

**Proof.** For an arbitrary Borel set $B \in \mathcal{B}(\mathbb{R}^{q+p})$ we have

$$\Phi_n^{-1}(B) = \pi_{0, \frac{1}{n}, \frac{2}{n}, \dots, 1}^{-1}(\widetilde{K}_n^{-1}(B)), \qquad n \in \mathbb{N},$$

where for all $n \in \mathbb{N}$ the mapping $\widetilde{K}_n : (\mathbb{R}^d)^{n+1} \to \mathbb{R}^{q+p}$ is defined by

$$\widetilde{K}_n(x_0, x_1, \dots, x_n) := \left( h(x_n), \frac{1}{n} \sum_{k=1}^n K\left( \frac{k}{n}, x_k, x_{k-1} \right) \right), \qquad x_0, x_1, \dots, x_n \in \mathbb{R}^d,$$

and the natural projections $\pi_{t_0,t_1,t_2,\ldots,t_n} : \mathsf{D}(\mathbb{R}_+, \mathbb{R}^d) \to (\mathbb{R}^d)^{n+1}$, $t_0, t_1, t_2, \ldots, t_n \in \mathbb{R}_+$, are given by $\pi_{t_0,t_1,t_2,\ldots,t_n}(f) := (f(t_0), f(t_1), f(t_2), \ldots, f(t_n))$, $f \in \mathsf{D}(\mathbb{R}_+, \mathbb{R}^d)$, $t_0, t_1, t_2, \ldots, t_n \in \mathbb{R}_+$. Since $h$ and $K$ are continuous, $\widetilde{K}_n$ is also continuous, and hence $\widetilde{K}_n^{-1}(B) \in \mathcal{B}((\mathbb{R}^d)^{n+1})$. It is known that $\pi_{t_0,t_1,t_2,\ldots,t_n}$, $t_0, t_1, t_2, \ldots, t_n \in \mathbb{R}_+$, are measurable mappings (see, e.g., Billingsley [6, Theorem 16.6 (ii)] or Ethier and Kurtz [11, Proposition 3.7.1]), and hence $\Phi_n = \widetilde{K}_n \circ \pi_{0,\frac{1}{n},\frac{2}{n},\ldots,1}$ is also measurable.

Next we show the measurability of $\Phi$. Since the natural projection $\mathsf{D}(\mathbb{R}_+, \mathbb{R}^d) \ni f \mapsto f(1) = \pi_1(f)$ is measurable, $h$ is continuous, it is enough to show that the mapping

$$\mathsf{D}(\mathbb{R}_+, \mathbb{R}^d) \ni f \mapsto \widetilde{\Phi}(f) := \int_0^1 K(t, f(t), f(t))\, \mathrm{d}t$$

is measurable. Namely, we show that $\widetilde{\Phi}$ is continuous. We have to check that $\widetilde{\Phi}(f_n) \to \widetilde{\Phi}(f)$ in $\mathbb{R}^p$ as $n \to \infty$ whenever $f_n \to f$ in $\mathsf{D}(\mathbb{R}_+, \mathbb{R}^d)$ as $n \to \infty$, where $f, f_n \in D(\mathbb{R}_+, \mathbb{R}^d)$, $n \in \mathbb{N}$. Due to Ethier and Kurtz [11, Proposition 3.5.3], for all $T > 0$ there exists a sequence $\lambda_n : \mathbb{R}_+ \to \mathbb{R}_+$, $n \in \mathbb{N}$, of strictly increasing continuous functions with $\lambda_n(0) = 0$ and $\lim_{t\to\infty} \lambda_n(t) = \infty$ such that

(B.2) $$\lim_{n\to\infty} \sup_{t\in[0,T]} |\lambda_n(t) - t| = 0, \qquad \lim_{n\to\infty} \sup_{t\in[0,T]} \|f_n(t) - f(\lambda_n(t))\| = 0.$$

We check that $\lim_{n\to\infty} f_n(t) = f(t)$ whenever $t \in \mathbb{R}_+$ is a continuity point of $f$. This readily follows by

$$\|f_n(t) - f(t)\| \leqslant \|f_n(t) - f(\lambda_n(t))\| + \|f(\lambda_n(t)) - f(t)\|, \qquad n \in \mathbb{N}, \quad t \in \mathbb{R}_+.$$

Using that $f$ has at most countably many discontinuities (see, e.g., Jacod and Shiryaev [20, page 326]), we have $\lim_{n\to\infty} f_n(t) = f(t)$ for all $t \in \mathbb{R}_+$ except a countable set having Lebesgue measure zero. In what follows we check that

$$\sup_{n\in\mathbb{N}} \sup_{t\in[0,1]} \|K(t, f_n(t), f_n(t))\| < \infty.$$

Since $K$ is continuous and hence it is bounded on a compact set, it is enough to verify that

$$\sup_{n\in\mathbb{N}} \sup_{t\in[0,1]} \|f_n(t)\| < \infty.$$

This follows by Jacod and Shiryaev [20, Chapter VI, Lemma 1.14 (b)], since $f_n \to f$ in $D(\mathbb{R}_+, \mathbb{R}^d)$ yields that $\{f_n : n \in \mathbb{N}\}$ is a relatively compact set (with respect to the Skorokhod topology). Then Lebesgue dominated convergence theorem yields the continuity of $\widetilde{\Phi}$.

In order to show $C_{\Phi,(\Phi_n)_{n\in\mathbb{N}}} = \mathsf{C}(\mathbb{R}_+, \mathbb{R}^d)$ we have to check that $\Phi_n(f_n) \to \Phi(f)$ whenever

$f_n \xrightarrow{\mathrm{lu}} f$ with $f \in \mathsf{C}(\mathbb{R}_+, \mathbb{R}^d)$ and $f_n \in \mathsf{D}(\mathbb{R}_+, \mathbb{R}^d)$, $n \in \mathbb{N}$. We have

$$\|\Phi_n(f_n) - \Phi(f)\| \leqslant \|h(f_n(1)) - h(f(1))\|$$

$$+ \frac{1}{n} \sum_{k=1}^{n} \left\| K\left(\frac{k}{n}, f_n\left(\frac{k}{n}\right), f_n\left(\frac{k-1}{n}\right)\right) - K\left(\frac{k}{n}, f\left(\frac{k}{n}\right), f\left(\frac{k-1}{n}\right)\right) \right\|$$

$$+ \sum_{k=1}^{n} \int_{(k-1)/n}^{k/n} \left\| K\left(\frac{k}{n}, f\left(\frac{k}{n}\right), f\left(\frac{k-1}{n}\right)\right) - K(t, f(t), f(t)) \right\| \, \mathrm{d}t$$

$$=: \|h(f_n(1)) - h(f(1))\| + A_n^{(1)} + A_n^{(2)}.$$

Since $f_n \xrightarrow{\mathrm{lu}} f$ implies that $f_n(1) \to f(1)$ as $n \to \infty$, using the continuity of $h$, we get

$$\|h(f_n(1)) - h(f(1))\| \to 0 \qquad \text{as} \quad n \to \infty.$$

Let us also observe that

$$\sup_{n \in \mathbb{N}} \sup_{t \in [0,1]} \|f_n(t)\| \leqslant \sup_{n \in \mathbb{N}} \sup_{t \in [0,1]} \|f_n(t) - f(t)\| + \sup_{t \in [0,1]} \|f(t)\| =: c < \infty,$$

hence

$$\left\| \left( f_n\left(\frac{k}{n}\right), f_n\left(\frac{k-1}{n}\right) \right) \right\| \leqslant \sqrt{2}c, \qquad n \in \mathbb{N}, \quad k \in \{1, \ldots, n\},$$

and then, by (B.1),

$$A_n^{(1)} \leqslant \sqrt{2} C_{\sqrt{2}c} \sup_{t \in [0,1]} \|f_n(t) - f(t)\| \to 0$$

as $n \to \infty$. Moreover,

$$A_n^{(2)} \leqslant C_{\sqrt{2}c} \sum_{k=1}^{n} \int_{(k-1)/n}^{k/n} \left( \left| \frac{k}{n} - t \right| + \left\| \left( f\left(\frac{k}{n}\right), f\left(\frac{k-1}{n}\right) \right) - (f(t), f(t)) \right\| \right) \, \mathrm{d}t$$

$$\leqslant \sqrt{2} C_{\sqrt{2}c} (n^{-1} + \omega_1(f, n^{-1})),$$

where

$$\omega_1(f, \varepsilon) := \sup_{t, s \in [0,1], |t-s| < \varepsilon} \| f(t) - f(s) \|, \qquad \varepsilon > 0,$$

denotes the modulus of continuity of $f$ on $[0, 1]$. Since $f$ is continuous, $\omega_1(f, n^{-1}) \to 0$ as $n \to \infty$ (see, e.g., Jacod and Shiryaev [20, Chapter VI, 1.6]), and we obtain $A_n^{(2)} \to 0$ as $n \to \infty$. Then $C_{\Phi, (\Phi_n)_{n \in \mathbb{N}}} = \mathsf{C}(\mathbb{R}_+, \mathbb{R}^d)$.

Finally, $\mathsf{C}(\mathbb{R}_+, \mathbb{R}^d) \in \mathcal{B}(\mathsf{D}(\mathbb{R}_+, \mathbb{R}^d))$ holds since $\mathsf{D}(\mathbb{R}_+, \mathbb{R}^d) \setminus \mathsf{C}(\mathbb{R}_+, \mathbb{R}^d)$ is open. Indeed, if $f \in \mathsf{D}(\mathbb{R}_+, \mathbb{R}^d) \setminus \mathsf{C}(\mathbb{R}_+, \mathbb{R}^d)$ then there exists $t \in \mathbb{R}_+$ such that $\varepsilon := \|f(t) - \lim_{s \uparrow t} f(s)\| > 0$, and then the open ball in $\mathsf{D}(\mathbb{R}_+, \mathbb{R}^d)$ with centre $f$ and radius $\varepsilon/2$ does not contain any continuous function. We note that for $\mathsf{C}(\mathbb{R}_+, \mathbb{R}^d) \in \mathcal{B}(\mathsf{D}(\mathbb{R}_+, \mathbb{R}^d))$ one can also simply refer to Ethier and Kurtz [11, Problem 3.11.25]. □

# C  Convergence of random step processes

We recall a result about convergence of random step processes towards a diffusion process, see Ispány and Pap [19]. This result is used for the proof of convergence (5.9).

**C.1 Theorem.** *Let* $\boldsymbol{\gamma} : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^{d \times r}$ *be a continuous function. Assume that uniqueness in the sense of probability law holds for the SDE*

$$(\text{C.1}) \qquad \mathrm{d}\boldsymbol{\mathcal{U}}_t = \gamma(t, \boldsymbol{\mathcal{U}}_t)\,\mathrm{d}\boldsymbol{\mathcal{W}}_t, \qquad t \in \mathbb{R}_+,$$

*with initial value* $\boldsymbol{\mathcal{U}}_0 = \boldsymbol{u}_0$ *for all* $\boldsymbol{u}_0 \in \mathbb{R}^d$, *where* $(\boldsymbol{\mathcal{W}}_t)_{t \in \mathbb{R}_+}$ *is an $r$-dimensional standard Wiener process. Let* $(\boldsymbol{\mathcal{U}}_t)_{t \in \mathbb{R}_+}$ *be a solution of* (C.1) *with initial value* $\boldsymbol{\mathcal{U}}_0 = \boldsymbol{0} \in \mathbb{R}^d$.

*For each* $n \in \mathbb{N}$, *let* $(\boldsymbol{U}_k^{(n)})_{k \in \mathbb{N}}$ *be a sequence of $d$-dimensional martingale differences with respect to a filtration* $(\mathcal{F}_k^{(n)})_{k \in \mathbb{Z}_+}$, *i.e.,* $\mathbb{E}(\boldsymbol{U}_k^{(n)} \,|\, \mathcal{F}_{k-1}^{(n)}) = \boldsymbol{0}$, $k, n \in \mathbb{N}$. *Let*

$$\boldsymbol{\mathcal{U}}_t^{(n)} := \sum_{k=1}^{\lfloor nt \rfloor} \boldsymbol{U}_k^{(n)}, \qquad t \in \mathbb{R}_+, \quad n \in \mathbb{N}.$$

*Suppose that* $\mathbb{E}\left(\|\boldsymbol{U}_k^{(n)}\|^2\right) < \infty$ *for all* $k, n \in \mathbb{N}$. *Suppose that for each* $T > 0$,

(i) $\displaystyle \sup_{t \in [0,T]} \left\| \sum_{k=1}^{\lfloor nt \rfloor} \mathbb{E}\left( \boldsymbol{U}_k^{(n)} (\boldsymbol{U}_k^{(n)})^\top \,|\, \mathcal{F}_{k-1}^{(n)} \right) - \int_0^t \boldsymbol{\gamma}(s, \boldsymbol{\mathcal{U}}_s^{(n)}) \boldsymbol{\gamma}(s, \boldsymbol{\mathcal{U}}_s^{(n)})^\top \mathrm{d}s \right\| \xrightarrow{\mathbb{P}} 0$,

(ii) $\displaystyle \sum_{k=1}^{\lfloor nT \rfloor} \mathbb{E}\left( \|\boldsymbol{U}_k^{(n)}\|^2 \mathbf{1}_{\{\|\boldsymbol{U}_k^{(n)}\| > \theta\}} \,\big|\, \mathcal{F}_{k-1}^{(n)} \right) \xrightarrow{\mathbb{P}} 0$ *for all* $\theta > 0$,

*where* $\xrightarrow{\mathbb{P}}$ *denotes convergence in probability. Then* $\boldsymbol{\mathcal{U}}^{(n)} \xrightarrow{\mathcal{L}} \boldsymbol{\mathcal{U}}$ *as* $n \to \infty$.

Note that in (i) of Theorem C.1, $\|\cdot\|$ denotes a matrix norm, while in (ii) it denotes a vector norm.

# References

[1] AL-OSH, M. A. and ALZAID, A. A. (1987). First order integer-valued autoregressive INAR(1) process. *Journal of Time Series Analysis* **8(3)** 261–275.

[2] ALZAID, A. A. and AL-OSH, M. A. (1990). An integer-valued $p$th-order autoregressive structure (INAR(p)) process. *Journal of Applied Probability* **27(2)** 314–324.

[3] BARCZY, M., DOERING, L., LI, Z. and PAP, G. (2013). Parameter estimation for an affine two factor model. Available on the ArXiv: `http://arxiv.org/abs/1302.3451`.

[4] BARCZY, M., ISPÁNY, M. and PAP, G. (2010). Asymptotic behavior of CLS estimators of autoregressive parameter for nonprimitive unstable INAR(2) models. Available on the ArXiv: `http://arxiv.org/abs/1006.4641`.

[5] BARCZY, M., ISPÁNY, M. and PAP, G. (2011). Asymptotic behavior of unstable INAR($p$) processes. *Stochastic Process. Appl.* **121(3)** 583–608.

[6] BILLINGSLEY, P. (1968). *Convergence of Probability Measures*, Wiley, 1968.

[7] BRÄNNÄS, K. and HELLSTRÖM, J. (2001). Generalized integer-valued autoregression. *Econometric Reviews* **20(4)** 425–443.

[8] DROST, F. C., VAN DEN AKKER, R. and WERKER, B. J. M. (2008). Local asymptotic normality and efficient estimation for INAR($p$) models. *Journal of Time Series Analysis* **29(5)** 783–801.

[9] DROST, F. C., VAN DEN AKKER, R. and WERKER, B. J. M. (2009). The asymptotic structure of nearly unstable non-negative integer-valued AR(1) models. *Bernoulli* **15(2)** 297–324.

[10] DU, J. G. and LI, Y. (1991). The integer valued autoregressive (INAR($p$)) model. *Journal of Time Series Analysis* **12(2)** 129–142.

[11] ETHIER, S. N. and KURTZ, T. G. (1986). *Markov Processes*, John Wiley & Sons, Inc., New York.

[12] FRANKE, J. and SELIGMANN, T. (1993). Conditional maximum-likelihood estimates for INAR(1) processes and their applications to modelling epileptic seizure counts. In: Subba Rao, T. (Ed.), *Developments in time series*, pp. 310–330. London: Chapman & Hall.

[13] GAUTHIER, G. and LATOUR, A. (1994). Convergence forte des estimateurs des paramètres d'un processus GENAR($p$). *Annales des Sciences Mathématiques du Québec* **18(1)** 49–71.

[14] HALL, P. and YAO, Q. (2003). Inference in ARCH and GARCH models with heavy-tailed errors. *Econometrica* **71(1)** 285–317.

[15] HAMILTON, J. D. (1994). *Time series analysis.* Princeton University Press, Princeton.

[16] HORN, R. A. and JOHNSON, CH. R. (1985). *Matrix Analysis.* Cambridge University Press, Cambridge.

[17] ISPÁNY, M., PAP, G. and VAN ZUIJLEN, M. C. A. (2003). Asymptotic inference for nearly unstable INAR(1) models. *Journal of Applied Probability* **40(3)** 750–765.

[18] ISPÁNY, M., PAP, G. and VAN ZUIJLEN, M. C. A. (2003). Asymptotic behaviour of estimators of the parameters of nearly unstable INAR(1) models. *Foundations of Statistical Inference (Shoresh, 2000),* pp. 193–204, Contrib. Statist., Physica, Heidelberg.

[19] ISPÁNY, M. and PAP, G. (2010). A note on weak convergence of random step processes. *Acta Mathematica Hungarica* **126(4)** 381–395.

[20] JACOD, J. and SHIRYAEV, A. N. (2003). *Limit Theorems for Stochastic Processes*, 2nd ed. Springer-Verlag, Berlin.

[21] KALLENBERG, O. (1997). *Foundations of Modern Probability.* Springer, New York, Berlin, Heidelberg.

[22] KARATZAS, I. and SHREVE, S. E. (1991). *Brownian Motion and Stochastic Calculus*, 2nd ed. Springer, Berlin.

[23] KESTEN, H. and STIGUM, B. P. (1966). A limit theorem for multidimensional Galton–Watson processes. *Ann. Math. Statist.* **37(5)** 1211–1223.

[24] KESTEN, H. and STIGUM, B. P. (1966). Additional limit theorems for indecomposable multidimensional Galton–Watson processes. *Ann. Math. Statist.* **37(6)** 1463–1481.

[25] KESTEN, H. and STIGUM, B. P. (1967). Limit theorems for decomposable multidimensional Galton–Watson processes. *J. Math. Anal. Appl.* **17(2)** 309–338.

[26] LATOUR, A. (1998). Existence and stochastic structure of a non-negative integer-valued autoregressive processes. *Journal of Time Series Analysis* **19(4)** 439–455.

[27] MCKENZIE, E. (1985). Some simple models for discrete variate time series. *Journal of the American Water Resources Association* **21(4)** 645–650.

[28] MIKOSCH, T. and STRAUMANN, D. (2002). Whittle estimation in a heavy-tailed GARCH(1,1) model. *Stochastic Processes and their Applications* **100** 187–222.

[29] PUTZER, E. J. (1966). Avoiding the Jordan canonical form in the discussion of linear systems with constant coefficients. *The American Mathematical Monthly* **73(1)** 2–7.

[30] REVUZ, D. and YOR, M. (2001). *Continuous Martingales and Brownian Motion*, 3rd ed., corrected 2nd printing. Springer-Verlag, Berlin.

[31] SILVA, M. and OLIVEIRA, V. (2005). Difference equations for the higher order moments and cumulants of the INAR($p$) model. *Journal of Time Series Analysis* **26(1)** 17–36.

[32] SIMS, C. A., STOCK, J. H. and WATSON, M. W. (1990). Inference in linear time series models with some unit roots. *Econometrica* **58(1)** 113–144.

[33] STEUTEL, F. and VAN HARN, K. (1979). Discrete analogues of self-decomposability and stability. *The Annals of Probability* **7(5)** 893–899.

[34] TANAKA, K. (1996). *Time Series Analysis, Nonstationary and Noninvertible Distribution Theory.* Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., New York.

[35] VAN DER VAART, A. W. (1998). *Asymptotic Statistics*. Cambridge University Press, Cambridge.

[36] WEISS, C. H. (2008). Thinning operations for modelling time series of counts—a survey. *AStA Advances in Statistical Analysis* **92(3)** 319–341.

[37] WINNICKI, J. (1991). Estimation of the variances in the branching process with immigration. *Probability Theory and Related Fields* **88(1)** 77–106.