Martin Jacobsen

# WEAK CONVERGENCE OF AUTOREGRESSIVE PROCESSES

1
Preprint
January
1994

IMS

Institute of Mathematical Statistics
University of Copenhagen

ISSN 0902-8846

Martin Jacobsen

# WEAK CONVERGENCE OF AUTOREGRESSIVE PROCESSES

Preprint 1994 No. 1

INSTITUTE OF MATHEMATICAL STATISTICS
UNIVERSITY OF COPENHAGEN

January 1994

# WEAK CONVERGENCE OF AUTOREGRESSIVE PROCESSES

Martin Jacobsen
Institute of Mathematical Statistics
University of Copenhagen
5 Universitetsparken
2100 Copenhagen Ø
Denmark

**Abstract**

A result is presented on weak convergence of sequences of multivariate, higher order autoregressive processes in discrete time, to certain limit processes in continuous time. The emphasis is on performing the limit in such a way, that the order of the sequence is preserved in the limit. The simplest, and well known case, is that of first order autoregressive processes converging to a (Gaussian) diffusion limit. For sequences of higher order, the class of limits include processes with sample paths, that are differentiable a different number of times in different directions. The main limit result is formulated as an invariance principle.

*Keywords and phrases.* Multivariate time series, higher order Markov processes, processes with differentiable sample paths.

*AMS* 1991 *subject classification.* 60F17, 60J60, 60J10, 62M10.

# 1 INTRODUCTION

Consider a sequence $(Y_n)_{n\geq 1}$ of autoregressive processes of order $r+1$, given by

$$Y_{n,k} = \sum_{l=0}^{r} A_{n,l} Y_{n,k-l-1} + \epsilon_{n,k} \qquad (k \geq r+1), \tag{1}$$

where each $Y_{n,k} \in \mathbf{R}^{d\times 1}$ is viewed as a $d-$dimensional column vector, and where the $A_{n,l} \in \mathbf{R}^{d\times d}$ are non-random coefficient matrices, that do not depend on $k$. It will be convenient, but by no means necessary, to think of the error variables $(\epsilon_{n,k})_{k\geq r+1}$ as independent and identically distributed, $\epsilon_{n,k} \sim N(0,\Gamma_n)$, where $N(0,\Gamma_n)$ denotes the Gaussian law on $\mathbf{R}^{d\times 1}$ with mean vector 0 and covariance matrix $\Gamma_n$.

The $n'$th process $Y_n$ is completely specified by describing in addition the initial values $(Y_{n,k})_{0\leq k\leq r}$. Just now it will be convenient to keep in mind the simplest case with $Y_{n,0} \equiv \cdots \equiv Y_{n,r} \equiv 0$.

From the sequence $(Y_n)_{n\geq 1}$ in discrete time, create a new sequence of processes $(X_n)_{n\geq 1}$ in continuous time by

$$X_n(t) = Y_{n,[nt]} \qquad (t \geq 0). \tag{2}$$

The problem is now to find conditions on the coefficients $A_{n,l}$ and the errors $\epsilon_{n,k}$, such that $X_n \xrightarrow{\mathcal{D}} X$, with $X$ some continuous process in continuous time. Here $\xrightarrow{\mathcal{D}}$ refers to convergence in distribution on the Skorohod space $D_{\mathbf{R}^{d\times 1}}([0,\infty))$. It is particularly important that the limit process should have properties, that identify the order $r+1$ of the autoregressive sequence $(Y_n)$.

The case $r=0$ ($1'$st order autoregressive processes) is standard: consider for $B \in \mathbf{R}^{d\times d}$, writing $I_d$ for the $d\times d$ identity matrix,

$$Y_{n,k} = \left(I_d + \frac{1}{n}B\right) Y_{n,k-1} + \epsilon_{n,k} \tag{3}$$

and assume that $\epsilon_{n,k} \sim N\left(0, \frac{1}{n}\Gamma\right)$ where $\Gamma$ is a covariance matrix. Writing $\Delta_h$ for the difference operator

$$\Delta_h f(t) = f(t) - f(t-h) \tag{4}$$

(3) may be written as

$$\Delta_{\frac{1}{n}} X_n(t) = \frac{1}{n} B X_n\left(t - \tfrac{1}{n}\right) + \epsilon_{n,[nt]},$$

which in the limit resembles the SDE

$$dX(t) = BX(t)\,dt + \Gamma^{\frac{1}{2}}\,dW(t),$$

where $W$ is standard $d-$dimensional Brownian motion. In fact we have the following well known result:

**Proposition 1.1** *Suppose that the sequence $(\epsilon_n)$ of error processes satisfy that*

$$\left(\sum_{k=1}^{[nt]} \epsilon_{n,k}\right)_{t\geq 0} \xrightarrow{\mathcal{D}} U \tag{5}$$

*with $U$ a continuous semimartingale. Then for $X_n$ given by (2) with $Y_n$ as in (3) and $Y_{n,0} \equiv 0$ for all $n$, it holds that $X_n \xrightarrow{\mathcal{D}} X$, where $X$ is the unique solution to the SDE*

$$dX(t) = BX(t)\,dt + dU(t),$$

*subject to the boundary condition $X(0) \equiv 0$.*

For a proof in the univariate case, see [1, Lemma 2.1]. (Using a martingale functional central limit theorem, they show convergence of a certain sequence $(X_n)$ of processes to a Gaussian limit $X$. But $\widetilde{X}_n(t) := \left(1 + \frac{B}{n}\right)^{[nt]-n} X_n(t)$ is then the same as our $X_n$, converging to $\widetilde{X}(t) := e^{B(t-1)}X(t)$, which is the Ornstein-Uhlenbeck process). Using the techniques from [3], a different proof is given in the multivariate case in [9].

Note that if for each $n$, the $\epsilon_{n,k}$ are i.i.d $N\left(0, \frac{1}{n}\Gamma\right)$, then (5) holds with $U \stackrel{\mathcal{D}}{=} \Gamma^{\frac{1}{2}}W$. In this case $X$ is a Gaussian, homogeneous diffusion, see e.g the surveys [4], [5].

Already if $r = 1$, the situation is much more complex. For instance, modifying (3) to

$$Y_{n,k} = \left(I_d + \frac{1}{n}B\right) Y_{n,k-1} + \frac{1}{n^2} C Y_{n,k-2} + \epsilon_{n,k},$$

we would still obtain the same limit as in (3) (the term involving $Y_{n,k-2}$ is immaterial), while if

$$Y_{n,k} = \left(2I_d + \frac{1}{n}B\right) Y_{n,k-1} - \left(I_d + \frac{1}{n}B + \frac{1}{n^2}C\right) Y_{n,k-2} + \epsilon_{n,k}$$

one will obtain a Gaussian limit process with differentiable sample paths, *provided* the order of magnitude of the errors is changed with e.g $(\epsilon_{n,k})$ an i.i.d sequence with $\epsilon_{n,k} \sim N\left(0, \frac{1}{n^3}\Gamma\right)$. Also, the distribution of the limit process will depend on both matrices $B$ and $C$.

To obtain results about converging schemes of autoregressive processes of arbitrary order $r+1$, we shall first in Section 2 discuss a certain class of continuous processes in continuous time, and then show in Section 3 how they arise as limits. It will turn out that each of these processes will be the limit of a converging scheme of the form (for $r \geq 1$)

$$Y_{n,k} = \sum_{l=0}^{r} \left( \sum_{s=1-r}^{r+1} \frac{1}{n^s} \beta_{s,l} \right) Y_{n,k-l-1} + \epsilon_{n,k}, \tag{6}$$

with each $\beta_{s,l}$ a $d \times d-$matrix not depending on $n$. In the final Section 4, which contains our main result, we state an invariance principle with sufficient conditions for the sequence $(X_n)$ determined by (2) from (6), to converge in distribution to one of the limit processes.

As limits in continuous time of higher order AR$-$processes, it would be natural to call the processes that we shall discuss in Section 2, higher order CAR-processes (autoregressive processes in continuous time). The standard definition however [7, Section 3.7.5] requires CAR$-$processes of order $k$, to be $k$ times differentiable in time, and so, our class of processes is much larger: the standard definition would correspond to taking $V^{(0)} = V^{(1)} = \cdots = V^{(r)} = \mathbf{R}^{1\times d}$ in (8) below.

It may be noted already that for $r \geq 2$, strictly negative powers of $s$ are allowed in (6). Thus, for $r = 2$ it is for instance possible to obtain convergence from schemes of the following form,

$$Y_{n,k} = \sum_{l=0}^{2} \left( n\beta_{-1,l} + \beta_{0,l} + \frac{1}{n^3}\beta_{3,l} \right) Y_{n,k-l-1} + \epsilon_{n,k}$$

and with all three $\beta-$matrices affecting the limit distribution!

# 2 THE CLASS OF LIMIT PROCESSES

Let $X = (X(t))_{t\geq 0}$ be a $\mathbf{R}^{d\times 1}$–valued continuous process. If $v \in \mathbf{R}^{1\times d}$ is a row vector, we shall say that $X$ is *$p$ times differentiable in the direction $v$*, provided

$$t \to vX(t,\omega) \tag{7}$$

is, for almost all $\omega$, at least $p$ times continuously differentiable for all $t \in \mathbf{R}_0 := [0,\infty)$. Also, for $V \subset \mathbf{R}^{1\times d}$ a linear subspace, we call $X$ *$p$ times differentiable in the direction $V$*, if for almost all $\omega$, (7) is at least $p$ times continuously differentiable, simultaneously for all $v \in V$.

Now define $V^{(0)} = \mathbf{R}^{1\times d}$, and, for $p \geq 1$,

$$V^{(p)} = \left\{ v \in \mathbf{R}^{1\times d} : X \text{ is } p \text{ times differentiable in the direction } v \right\}.$$

Clearly $V^{(p)}$ is a linear subspace and

$$V^{(0)} \supset V^{(1)} \supset \cdots.$$

Note that for almost all $\omega$, the paths (7) are $p$ times differentiable in the direction $V^{(p)}$, as is seen by representing any $v \in V^{(p)}$ as a linear combination of a given finite set of base vectors.

Apart from the process $X$, we shall assume given a non-negative integer $r$, and shall then describe $X$ itself, using only the subspaces

$$V^{(0)} \supset V^{(1)} \supset \cdots \supset V^{(r)}. \tag{8}$$

It is natural, but not necessary to impose the condition $\dim\left(V^{(r)}\right) \geq 1$, i.e that $V^{(r)}$ must not be the null space $\{0\}$. Note that the equality $V^{(p)} = V^{(p+1)}$ is allowed.

If $\dim\left(V^{(r)}\right) \geq 1$ and $V^{(r+1)} = \{0\}$, we call $r+1$ the *order of the process* $X$. Thus the order is the maximal number plus one of differentiations possible in any direction. Note that in the description of processes $X$ to be given now, $r+1$ need not be the order, not even if $\dim\left(V^{(r)}\right) \geq 1$.

With $r$ and the nested subspaces (8) given, consider any decomposition

$$\mathbf{R}^{1\times d} = H^{(0)} \oplus \cdots \oplus H^{(r)} \tag{9}$$

of the space of $d$-dimensional row vectors, compatible with the structure (8):

$$V^{(p)} = H^{(p)} \oplus \cdots \oplus H^{(r)} \tag{10}$$

for $0 \leq p \leq r$.

Such a decomposition is in general not unique, but it always holds that $H^{(r)} = V^{(r)}$. Note that since $V^{(p)} = V^{(p+1)}$ may occur, $H^{(p)} = \{0\}$ is possible.

Corresponding to the decomposition (9), there are uniquely determined linear maps $\pi^{(p)}$ acting on $\mathbf{R}^{1\times d}$, such that for any $v \in \mathbf{R}^{1\times d}$,

$$v = v\pi^{(0)} + \cdots + v\pi^{(r)}$$

with $v\pi^{(p)} \in H^{(p)}$, $0 \leq p \leq r$. We shall identify each $\pi^{(p)}$ with a $d\times d$-matrix, and then note that the $\pi^{(p)}$ yield a decomposition of unity

$$I_d = \pi^{(0)} + \cdots + \pi^{(r)}$$

with

$$\sum_{p=0}^{r} \text{rank}\left(\pi^{(p)}\right) = d.$$

(If $H^{(p)} = \{0\}$, $\pi^{(p)} = 0_d$, the $d \times d$ null matrix). Note also that

$$\begin{array}{rl} \pi^{(p)}\pi^{(p)} = \pi^{(p)} & \qquad (0 \leq p \leq r) \\ \pi^{(p)}\pi^{(q)} = 0_d & \qquad (0 \leq p \neq q \leq r). \end{array} \tag{11}$$

The $\pi^{(p)}$ were derived from the decomposition (9) of row vectors, but in turn define a decomposition of the space of column vectors,

$$\mathbf{R}^{d\times 1} = K^{(0)} \oplus \cdots \oplus K^{(r)}$$

with $K^{(p)}$ the subspace spanned by the columns of $\pi^{(p)}$ and such that for any $x \in \mathbf{R}^{d\times 1}$,

$$x = \pi^{(0)}x + \cdots + \pi^{(r)}x$$

with $\pi^{(p)}x \in K^{(p)}$, $0 \leq p \leq r$. For further reference, introduce also

$$L^{(p)} = K^{(p)} \oplus \cdots \oplus K^{(r)}.$$

In the description of the process $X$ we shall now present, the basic ingredients are the (column) components $\pi^{(p)}X$. Of course here, the sample paths

$t \to \pi^{(p)} X(t, \omega)$ are at least $p$ times continuously differentiable, and the idea is now to consider the collection

$$Z = \left( \left( \pi^{(p)} X \right)^{(q)} \right)_{0 \le q \le p \le r} \tag{12}$$

of all derivatives under consideration, and then set up a simple SDE for this *big process* $Z$.

**Notation.** The upper index $q$ in (12) refers to the $q$'th time derivative,

$$\left( \pi^{(p)} X \right)^{(q)}(t) = \frac{d^q}{dt^q} \left( \pi^{(p)} X \right)(t). \tag{13}$$

Note that an expression like $(\rho X)^{(q)}$, where $\rho$ is some matrix, makes sense only if the rows of $\rho$ belong to $V^{(q)}$, and that it is not allowed to write e.g $\pi^{(p)} \left( X^{(q)} \right)$ for the left hand side of (13): $X$ may not have $q$'th order derivatives in all directions.

We shall now present the SDE for the big process $Z$ given by (12). Of course

$$d \left( \pi^{(p)} X \right)^{(q)} = \left( \pi^{(p)} X \right)^{(q+1)} dt \qquad (0 \le q < p \le r), \tag{14}$$

so it only remains to specify the SDE's for the maximal number $p$ of derivatives in each direction $\pi^{(p)}$, and here we shall impose the following structure: let $U$ be a continuous semimartingale, defined on a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ satisfying the usual conditions. Then we assume that for $0 \le p \le r$,

$$d \left( \pi^{(p)} X \right)^{(p)} = \sum_{m=0}^{r} \sum_{q=0}^{m} M^{(p,m,q)} \left( \pi^{(m)} X \right)^{(q)} dt + \pi^{(p)} \, dU \tag{15}$$

where each $M^{(p,m,q)} \in \mathbf{R}^{d \times d}$ is non-random and does not depend on time, and, viewed as a linear map acting on the space of column vectors, $M^{(p,m,q)} : K^{(m)} \to K^{(p)}$.

It is possible to write (15) in a more compact form. Define

$$\psi^{(p)} = \pi^{(p)} + \cdots + \pi^{(r)}$$

(so $v\psi^{(p)} = v$ for $v \in V^{(p)}$, $\psi^{(p)} x = x$ for $x \in L^{(p)}$). Then

$$d \left( \pi^{(p)} X \right)^{(p)} = \sum_{q=0}^{r} B^{(p,q)} \left( \psi^{(q)} X \right)^{(q)} dt + \pi^{(p)} \, dU, \tag{16}$$

with $B^{(p,q)} : L^{(q)} \to K^{(p)}$ given by

$$B^{(p,q)} = \sum_{m=q}^{r} M^{(p,m,q)} \pi^{(m)},$$

(recall that for $m \geq q$, $\pi^{(m)} = \pi^{(m)} \psi^{(q)}$). Obviously, any system of the form (16) is also of the form (15).

With (15) or (16) satisfied, it is clear that $Z$ itself satisfies a SDE of the form

$$dZ = \tilde{B} Z \, dt + d\tilde{U} \tag{17}$$

where $\tilde{B}$ is non-random and does not depend on time, and where $\tilde{U}$ is a continuous semimartingale.

Given a $\mathcal{F}_0$-measurable random vector $Z_0$, the SDE (17) has a unique strong solution with $Z(0) = Z_0$. With this and the preceding discussion in mind, the following result is easy to show.

**Proposition 2.1** (i) *With $U$ a $\mathbf{R}^{d\times 1}$-valued, continuous semimartingale, and $x^{(p,q)} \in K^{(p)}$ given non-random vectors $(0 \leq q \leq p \leq r)$, there is a unique process $X$, which solves* (16) *for $0 \leq p \leq r$, which is $p$ times differentiable in the direction $V^{(p)}$ for each $p$, and which satisfies the initial conditions*

$$\left(\pi^{(p)} X\right)^{(q)}(0) \equiv x^{(p,q)} \qquad (0 \leq q \leq p \leq r). \tag{18}$$

(ii) *If $U$ is a Brownian motion, the solution $X$ is Gaussian, and the big process $Z$ determined by $X$ is a Gaussian, homogeneous diffusion.*

**Remarks.**

- We call $U$ a Brownian motion (defined on $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ if $U$ is adapted, continuous, $U(0) = 0$ a.s. and for some drift vector $\xi \in \mathbf{R}^{d\times 1}$ and some $d \times d$ covariance matrix $\Gamma$, $U(t) - U(s)$ is for $0 \leq s \leq t$ independent of $\mathcal{F}_s$ and $N\left((t-s)\,\xi, (t-s)\,\Gamma\right)$-distributed.

- If $U$ is a Brownian motion with non-singular covariance $\Gamma$, then the process $X$ obtained by Proposition 2.1 has order $r+1$. In that case, if also $d = 1$, $X$ is an example of a (Gaussian) $r'$th order Markov process in the sense of [2, p. 272].

- With $U$ still Brownian, the big process $Z$ is an example of what [5] called a Gaussian homogeneous diffusion with smooth components. Any such process in $\tilde{d}$ dimensions satisfies an SDE of the form

$$dZ = \left(\tilde{A} + \tilde{B}Z\right)\, dt + \widetilde{D}\, d\widetilde{W} \tag{19}$$

with $\widetilde{W}$ a $\tilde{d}$–dimensional standard Brownian motion and the matrix $\widetilde{D}$ of reduced rank $< \tilde{d}$.

- It should now be clear, that the idea behind the system of equations (14) and (16) is to set up, in the case where $U$ is a Brownian motion, the most general class of processes $X$, such that the corresponding big processes $Z$ are Gaussian, homogeneous diffusions. Note that this is achieved without incorporating constant drift terms (such as $\tilde{A}$ in (19)) in (16) or (17): these terms are included in the driving Brownian motion $U$.

It is somewhat unsatisfactory that Proposition 2.1 depends on the arbitrary subspace decomposition (9). However, it is possible to describe $X$ as the solution to a system of SDE's that are intrinsic in the sense that the coefficients in the finite variation terms depend only on the fundamental spaces $V^{(p)}$, not on the arbitrary $H^{(p)}$'s.

**Theorem 2.2** (i) *Let the semimartingale $U$ and the vectors $x^{(p,q)}$ be as in Proposition* 2.1. *If $X$ denotes the unique solution to* (16) *with $\left(\pi^{(p)}X\right)^{(q)}(0) = x^{(p,q)}$, then there exist linear maps (not depending on the $x^{(p,q)}$)*

$$\alpha^{(p,q)} : V^{(p)} \to V^{(q)}$$

*for $0 \leq p, q \leq r$ and*

$$\gamma^{(p)} : V^{(p)} \to \mathbf{R}^{1\times d}$$

*for $0 \leq p \leq r$, such that*

$$d\,(vX)^{(p)} = \sum_{q=0}^{r} \left(v\alpha^{(p,q)}X\right)^{(q)} dt + v\gamma^{(p)}\, dU \tag{20}$$

*for $0 \leq p \leq r$ and all $v \in V^{(p)}$.*

*The $\alpha^{(p,q)}$ and the $\gamma^{(p)}$ satisfy the consistency conditions that for $0 \le p \le r-1$,*

$$v\alpha^{(p,q)} = \begin{cases} 0 & \left(q \neq p+1,\ v \in V^{(p+1)}\right) \\ v & \left(q = p+1,\ v \in V^{(p+1)}\right), \end{cases} \tag{21}$$

$$v\gamma^{(p)} = 0 \qquad \left(v \in V^{(p+1)}\right). \tag{22}$$

(ii) *Suppose conversely that*

$$\mathbf{R}^{1\times d} = V^{(0)} \supset \cdots \supset V^{(r)}$$

*are given subspaces, and let $\alpha^{(p,q)} : V^{(p)} \to V^{(q)}$ and $\gamma^{(p)} : V^{(p)} \to \mathbf{R}^{1\times d}$ be linear maps that satisfy* (21) *and* (22) *respectively. Let furthermore $f^{(p,q)} : V^{(p)} \to \mathbf{R}$ be linear functionals such that for $0 \le q \le p \le r-1$*

$$vf^{(p,q)} = vf^{(p+1,q)} \qquad \left(v \in V^{(p+1)}\right). \tag{23}$$

*Then, with $U$ a given continuous semimartingale, there is a unique process $X$, $p$ times differentiable in the direction $V^{(p)}$ for all $p$, that satisfies* (20) *for all $0 \le p \le r$, $v \in V^{(p)}$ subject to the initial conditions*

$$(vX)^{(q)}(0) = vf^{(p,q)} \qquad \left(0 \le q \le p \le r,\ v \in V^{(p)}\right). \tag{24}$$

**Proof.** (i) Let $v \in V^{(p)}$. From (16) and (14), since $v = \sum_{m=p}^{r} v\pi^{(m)}$ for $v \in V^{(p)}$, it follows that

$$d(vX)^{(p)} = \sum_{q=0}^{r} vB^{(p,q)}\left(\psi^{(q)}X\right)^{(q)} dt + \sum_{m=p+1}^{r} v\left(\pi^{(m)}X\right)^{(p+1)} dt + v\pi^{(p)}\, dU,$$

which is of the form (20) with

$$\alpha^{(p,q)} = \begin{cases} B^{(p,q)}\psi^{(q)} & (q \neq p+1) \\ \left(B^{(p,p+1)} + I_d\right)\psi^{(p+1)} & (q = p+1), \end{cases} \tag{25}$$

$$\gamma^{(p)} = \pi^{(p)}.$$

It is obvious that $v\alpha^{(p,q)} \in V^{(q)}$ if $v \in V^{(p)}$. Since $v\pi^{(p)} = 0$ for $v \in V^{(p+1)}$, (22) holds. (Here and in the sequel, (11) is used without comment). Since $B^{(p,q)} : L^{(q)} \to K^{(p)}$, we have $B^{(p,q)}\psi^{(q)} = \pi^{(p)}B^{(p,q)}\psi^{(q)}$ and hence $vB^{(p,q)}\psi^{(q)} = 0$ if $v \in V^{(p+1)}$. Since also then $v\psi^{(p+1)} = v$, (21) follows.

(ii) Suppose there is a process $X$ satisfying the requirements in part (ii) of the theorem. Consider a decomposition (9) satisfying (10) and define $\pi^{(p)}$ in the usual way. Then (20) implies

$$d\left(\pi^{(p)}X\right)^{(p)} = \sum_{q=0}^{r} \left(\pi^{(p)}\alpha^{(p,q)}X\right)^{(q)} dt + \pi^{(p)}\gamma^{(p)}\, dU.$$

Here $\pi^{(p)}\alpha^{(p,q)} = \pi^{(p)}\alpha^{(p,q)}\psi^{(q)}$ because $\alpha^{(p,q)} : V^{(p)} \to V^{(q)}$, and it is seen that $X$ satisfies (16) with

$$B^{(p,q)} = \pi^{(p)}\alpha^{(p,q)} \tag{26}$$

and the $U$ in (16) replaced by $\tilde{U} := \gamma^{(p)}U$. Furthermore $x^{(p,q)} := \left(\pi^{(p)}X\right)^{(q)}(0)$ for $0 \le q \le p \le r$ is given by

$$x^{(p,q)} = \pi^{(p)}f^{(p,q)}. \tag{27}$$

Thus $X$ satisfies the conditions of Proposition 2.1, and hence, by that proposition, $X$ is at most unique. To show that there is an $X$ obeying the requirements of the theorem, consider *the* solution $\widetilde{X}$ to (16), where $U$ is replaced by $\tilde{U}$, $B^{(p,q)}$ is given by (26), and $x^{(p,q)}$ by (27). We must then show that $\widetilde{X}$ satisfies (20) with (see (24)) $\left(v\widetilde{X}\right)^{(q)}(0) = vf^{(p,q)}$ for $0 \le q \le p \le r$, $v \in V^{(p)}$.

Since $\widetilde{X}$ satisfies (16), part (i) of Theorem 2.2 and its proof applies to $\widetilde{X}$, in particular it follows that $\widetilde{X}$ satisfies an equation of the form

$$d\left(v\widetilde{X}\right)^{(p)} = \sum_{q=0}^{r} \left(v\tilde{\alpha}^{(p,q)}\widetilde{X}\right)^{(q)} dt + v\tilde{\gamma}^{(p)}\, d\tilde{U},$$

where $\tilde{\alpha}^{(p,q)} = \pi^{(p)}\alpha^{(p,q)}\psi^{(q)}$ if $q \neq p+1$ and $= \left(\pi^{(p+1)}\alpha^{(p,p+1)} + I_d\right)\psi^{(p+1)}$ if $q = p+1$, is given by (25), using the $B^{(p,q)}$ from (26), and where $\tilde{\gamma}^{(p)} = \pi^{(p)}$. Thus, to show that $\widetilde{X}$ satisfies (20) itself, it suffices show that for $v \in V^{(p)}$,

$$v\tilde{\alpha}^{(p,q)} = v\alpha^{(p,q)}, \qquad v\pi^{(p)}\gamma^{(p)} = v\gamma^{(p)}. \tag{28}$$

But for $v \in V^{(p)}$, if $q \neq p+1$,

$$\begin{aligned} v\tilde{\alpha}^{(p,q)} &= v\pi^{(p)}\alpha^{(p,q)}\psi^{(q)} \\ &= v\pi^{(p)}\alpha^{(p,q)} \\ &= v\left(\pi^{(p)} + \psi^{(p+1)}\right)\alpha^{(p,q)} \\ &= v\alpha^{(p,q)}, \end{aligned}$$

where we have first used that $v\pi^{(p)}\alpha^{(p,q)} \in V^{(q)}$ and then that $v\psi^{(p+1)}\alpha^{(p,q)} = 0$ by (21). And if $q = p+1$,

$$\begin{aligned} v\widetilde{\alpha}^{(p,p+1)} &= v\left(\pi^{(p)}\alpha^{(p,p+1)} + I_d\right)\psi^{(p+1)} \\ &= v\pi^{(p)}\alpha^{(p,p+1)} + v\psi^{(p+1)} \\ &= v\pi^{(p)}\alpha^{(p,p+1)} + v\psi^{(p+1)}\alpha^{(p,p+1)} \\ &= v\alpha^{(p,p+1)}, \end{aligned}$$

using (21) for the third equality.

This proves the first identity in (28). The second follows from (22) since for $v \in V^{(p)}$,

$$v\gamma^{(p)} = v\left(\pi^{(p)} + \psi^{(p+1)}\right)\gamma^{(p)} = v\pi^{(p)}\gamma^{(p)}.$$

It remains to check that $\widetilde{X}$ satisfies the proper initial conditions,

$$\left(v\widetilde{X}\right)^{(q)}(0) = vf^{(p,q)}$$

for $0 \leq q \leq p \leq r$ and $v \in V^{(p)}$. But (see (27))

$$\left(v\widetilde{X}\right)^{(q)}(0) = \sum_{m=p}^{r}\left(v\pi^{(m)}\widetilde{X}\right)^{(q)}(0) = \sum_{m=p}^{r} v\pi^{(m)}f^{(m,q)}$$

and since by (23), $v\pi^{(m)}f^{(m,q)} = v\pi^{(m)}f^{(p,q)}$ for $m \geq p$, it follows that

$$\left(v\widetilde{X}\right)^{(q)}(0) = \sum_{m=p}^{r} v\pi^{(m)}f^{(p,q)} = vf^{(p,q)}$$

holds. □

# 3 SEQUENCES OF AR-PROCESSES CONVERGING TO A GIVEN LIMIT

As a first step towards our main result, we shall in this section show how any limit process $X$ of the type discussed in Section 2, may be obtained as the limit of sequences af autoregressive processes.

It will be convenient at this stage to use Proposition 2.1 and describe $X$ as the solution to (16) with initial conditions (18).

For $f : \mathbf{R}_0 \to \mathbf{R}^{d\times 1}$ a function, and $h > 0$ we define (cf (4)) $\Delta_h f$ as the function

$$\Delta_h f\,(t) = f(t) - f(t-h) \qquad (t \geq h)\,.$$

Repeated use of this difference operator gives

$$\Delta_h^p f(t) = \sum_{m=0}^{p} \binom{p}{m} (-1)^m f(t-mh) \qquad (t \geq ph) \tag{29}$$

for $p \in \mathbf{N}$, and for $p = 0$ we define $\Delta_h^0 f = f$.

The following observation forms the basis for the proofs in this section.

**Lemma 3.1** *Let $f : \mathbf{R}_0 \to \mathbf{R}^{d\times 1}$ be $q \geq 1$ times continuously differentiable. Then, for any $p \geq q$,*

$$\Delta_h^p f(t) = \int_{M_h^q(t)} du_1 \cdots du_q \, \Delta_h^{p-q} f^{(q)}(u_q) \qquad (t \geq ph), \tag{30}$$

*where, writing $u_0 = t$,*

$$M_h^q(t) = \{(u_1, \ldots, u_q) \in \mathbf{R}^q \, : \, u_{i-1} - h \leq u_i \leq u_{i-1}, \; 1 \leq i \leq q\}\,.$$

$\square$

The lemma is obviously true for $p = q = 1$ and is easily proved in general, using induction on $p$.

Let now $X$ be the unique solution to (16) with initial conditions (18). We shall use Lemma 3.1 to obtain approximate difference equations for $X$, involving the process at the lattice time points $\frac{k}{n}$, where $n \in \mathbf{N}$, $k \in \mathbf{N}_0$.

Using (30) with $p$ replaced by $p+1$ and $q = p$, it follows from (16) that for $n \in \mathbf{N}$, $k \geq p+1$,

$$\Delta_{\frac{1}{n}}^{p+1} \left(\pi^{(p)} X\right) \left(\tfrac{k}{n}\right) = \int_{M_{\frac{1}{n}}^{p}\left(\frac{k}{n}\right)} du_1 \cdots du_p \left( \left(\pi^{(p)} X\right)^{(p)} (u_p) - \left(\pi^{(p)} X\right)^{(p)} \left(u_p - \tfrac{1}{n}\right)\right)$$

$$= \int_{M_{\frac{1}{n}}^{p+1}\left(\frac{k}{n}\right)} du_1 \cdots du_{p+1} \sum_{q=0}^{r} B^{(p,q)} \left(\psi^{(q)} X\right)^{(q)} (u_{p+1}) + \tilde{\epsilon}_{n,k}^{(p)} \tag{31}$$

with errors

$$\widetilde{\epsilon}_{n,k}^{(p)} = \int\limits_{M_{\frac{1}{n}}^{p}\left(\frac{k}{n}\right)} du_1 \cdots du_p \left(\pi^{(p)} U\left(u_p\right) - \pi^{(p)} U\left(u_p - \tfrac{1}{n}\right)\right). \tag{32}$$

In what follows, consider only $k$ such that $\frac{k}{n} \leq T$, where $T > 0$ is an arbitrary constant.

Now, in the integral in (31), $\left|u_{p+1} - \frac{k}{n}\right| \leq \frac{p+1}{n}$, and since the integrand is uniformly continuous on $[0,T]$, we may therefore write

$$\left(\psi^{(q)} X\right)^{(q)}\left(u_{p+1}\right) = \left(n\Delta_{\frac{1}{n}}\right)^q \left(\psi^{(q)} X\right)\left(\tfrac{k-1}{n}\right) + o(1), \tag{33}$$

where $o(1)$ is a random vector, converging to 0 a.s as $n \to \infty$, uniformly in $k, u_{p+1}$ with $\frac{k}{n} \leq T$, $u_{p+1} \leq T$ and $\left|u_{p+1} - \frac{k}{n}\right| \leq \frac{p+1}{n}$. Since the Lebesgue measure of $M_{\frac{1}{n}}^{p+1}\left(\frac{k}{n}\right)$ is $n^{-p-1}$, it follows that

$$\Delta_{\frac{1}{n}}^{p+1}\left(\pi^{(p)} X\right)\left(\tfrac{k}{n}\right) = \sum_{q=0}^{r} \frac{1}{n^{p+1-q}} B^{(p,q)} \Delta_{\frac{1}{n}}^{q}\left(\psi^{(q)} X\right)\left(\tfrac{k-1}{n}\right) + \widetilde{\epsilon}_{n,k}^{(p)} + o_{k,p}\left(\tfrac{1}{n^{p+1}}\right) \tag{34}$$

for $r+1 \leq k \leq [nT]$, where $o_{k,p}\left(n^{-p-1}\right)$ is a random vector such that

$$\max_{k: r+1 \leq k \leq [nT]} n^{p+1} \left\| o_{k,p}\left(\tfrac{1}{n^{p+1}}\right)\right\| \overset{\text{a.s}}{\to} 0 \tag{35}$$

as $n \to \infty$. Here $\|\cdot\|$ denotes the Euclidean, or any other equivalent norm on $\mathbf{R}^{d\times 1}$.

**Remark.** For (33) we just needed some good approximation to the $q$'th derivative of $\psi^{(q)} X$ in terms of $X$ evaluated at the lattice points $\frac{j}{n}$ for $j \leq k-1$. We could equally well have replaced $\frac{k-1}{n}$ on the right of (33) by, say, $\frac{k-2}{n}$, and even used an approximant other than $\left(n\Delta_{\frac{1}{n}}\right)^q$ to the differentiation operator $\frac{d^q}{dt^q}$. Thus the converging scheme presented below, is by no means the only one with $X$ as limit.

**Proposition 3.2** *Let $(Y_n)_{n\geq 1}$ denote the sequence of autoregressive processes of order $r+1$ given by (1), where*

$$A_{n,l} = \sum_{p=0}^{r}\left[1_{(l\leq p)}\binom{p+1}{l+1}(-1)^l \pi^{(p)} + \sum_{q=l}^{r} \frac{1}{n^{p+1-q}}(-1)^l \binom{q}{l} B^{(p,q)} \psi^{(q)}\right], \tag{36}$$

$$\epsilon_{n,k} = \sum_{p=0}^{r} \frac{1}{n^p} \left( \pi^{(p)} U \left( \tfrac{k}{n} \right) - \pi^{(p)} U \left( \tfrac{k-1}{n} \right) \right) \tag{37}$$

*and with initial values $(Y_{n,j})_{0 \le j \le n}$ given by*

$$Y_{n,j} = \sum_{p=0}^{r} \pi^{(p)} Y_{n,j}$$

*where, for each $p$, $\pi^{(p)} Y_{n,j}$ is determined recursively by*

$$\pi^{(p)} Y_{n,j} = \begin{cases} \frac{1}{n^j} x^{(p,j)} - \sum_{i=1}^{j} \binom{j}{i} (-1)^i \pi^{(p)} Y_{n,j-i} & (0 \le j \le p) \\ \frac{1}{n^p} x^{(p,p)} - \sum_{i=1}^{p} \binom{p}{i} (-1)^i \pi^{(p)} Y_{n,j-i} & (p < j \le r). \end{cases} \tag{38}$$

*Defining $X_n$ by $X_n(t) = Y_{n,[nt]}$, it then holds that*

$$X_n \overset{\text{a.s.u.c}}{\to} X.$$

**Remark.** (38) may be written

$$\begin{array}{rcll} n^j \Delta^j_{\frac{1}{n}} \left( \pi^{(p)} X_n \right) \left( \frac{j}{n} \right) & = & x^{(p,j)} & (0 \le j \le p) \\ n^p \Delta^p_{\frac{1}{n}} \left( \pi^{(p)} X_n \right) \left( \frac{j}{n} \right) & = & x^{(p,p)} & (p < j \le r). \end{array} \tag{39}$$

**Notation.** a.s.u.c is short for 'convergence almost surely, uniformly on compact intervals of time'.

**Proof.** In (34), sum once over $k$ to obtain (since the number of terms in the sum is $\le nT$)

$$\begin{aligned} \Delta^p_{\frac{1}{n}} \left( \pi^{(p)} X \right) \left( \tfrac{k}{n} \right) = {} & \Delta^p_{\frac{1}{n}} \left( \pi^{(p)} X \right) \left( \tfrac{r}{n} \right) + \sum_{j=r+1}^{k} \sum_{q=0}^{r} \tfrac{1}{n^{p+1-q}} B^{(p,q)} \Delta^q_{\frac{1}{n}} \left( \psi^{(q)} X \right) \left( \tfrac{j-1}{n} \right) \\ & + \sum_{j=r+1}^{k} \tilde{\epsilon}^{(p)}_{n,j} + o\left( \tfrac{1}{n^p} \right) \end{aligned}$$

where $o\left(n^{-p}\right) = \sum_{j=r+1}^{k} o_j \left(n^{-p-1}\right)$. Define

$$\epsilon^{(p)}_{n,k} = \frac{1}{n^p} \left( \pi^{(p)} U \left( \tfrac{k}{n} \right) - \pi^{(p)} U \left( \tfrac{k-1}{n} \right) \right).$$

By (32)

$$\tilde{\epsilon}^{(p)}_{n,k} = \int\limits_{M^p_{\frac{1}{n}}\left(\frac{k}{n}\right)} du_1 \dots du_p\, \pi^{(p)} U\left(u_p\right) \quad - \int\limits_{M^p_{\frac{1}{n}}\left(\frac{k-1}{n}\right)} du_1 \dots du_p\, \pi^{(p)} U\left(u_p\right)$$

and hence

$$\begin{aligned}\sum_{j=r+1}^{k} \left(\tilde{\epsilon}^{(p)}_{n,j} - \epsilon^{(p)}_{n,j}\right) &= \int\limits_{M^p_{\frac{1}{n}}\left(\frac{k}{n}\right)} du_1 \dots du_p \left(\pi^{(p)} U\left(u_p\right) - \pi^{(p)} U\left(\frac{k}{n}\right)\right) \qquad (40)\\ &- \int\limits_{M^p_{\frac{1}{n}}\left(\frac{r}{n}\right)} du_1 \dots du_p \left(\pi^{(p)} U\left(u_p\right) - \pi^{(p)} U\left(\frac{r}{n}\right)\right)\\ &= o\left(\tfrac{1}{n^p}\right)\end{aligned}$$

uniformly in $k$, $r+1 \leq k \leq [nT]$, since $U$ is continuous.

Now define $X_n$ by $X_n(t) = X_n\left(\frac{[nt]}{n}\right)$ and, c.f (34), for $k \geq r+1$,

$$\Delta^{p+1}_{\frac{1}{n}} \left(\pi^{(p)} X_n\right) \left(\tfrac{k}{n}\right) = \sum_{q=0}^{r} \frac{1}{n^{p+1-q}} F^{(p,q)} \left(\Delta^q_{\frac{1}{n}} X_n\right) \left(\tfrac{k-1}{n}\right) + \epsilon^{(p)}_{n,k}, \qquad (41)$$

where $F^{(p,q)} = B^{(p,q)} \psi^{(q)}$.

Since the leading term on the left is $\pi^{(p)} X_n\left(\frac{k}{n}\right)$ and all other terms involve $X_n$ at points $\frac{j}{n}$, where $j < k$, this is a recursive system determining $X_n = \sum_p \pi^{(p)} X_n$ once the initial values

$$\left(X_n\left(\tfrac{j}{n}\right)\right)_{0 \leq j \leq r}$$

are fixed.

In the remainder of the proof, write $\Delta = \Delta_{\frac{1}{n}}$ and introduce

$$D^p_n = \pi^{(p)} X - \pi^{(p)} X_n,$$

$$D_n = X - X_n$$

and

$$\kappa_n = \max\left\{n^l \left\|\Delta^l D^p_n\left(\tfrac{j}{n}\right)\right\| \;:\; l, p, j \text{ with } 0 \leq l \leq p \leq r,\; l \leq j \leq r\right\},$$

$$\lambda_n = \max_{p,k:0\le p\le r,\, r+1\le k\le [nT]} \left\{ n^p \left\| \sum_{j=r+1}^{k} \left[ \left( \tilde{\epsilon}^{(p)}_{n,j} - \epsilon^{(p)}_{n,j} \right) + o_{j,p}\left(\tfrac{1}{n^{p+1}}\right) \right] \right\| \right\}$$

with $o_{j,p}\left(n^{-p-1}\right)$ as in (34).

The random variable $\kappa_n$ refers to the initial difference (before time $\frac{r}{n}$) between $X$ and $X_n$, while $\lambda_n$ describes the discrepancy between the error terms in (34) and those used in the definition of $X_n$, see (41). From (35) and (40) it follows, that for any $T > 0$,

$$\lambda_n \overset{\text{a.s}}{\to} 0. \tag{42}$$

We now claim that for $T > 0$ sufficiently small, there is a constant $K > 0$ such that for $0 \le m \le p \le r$,

$$\max_{j:m\le j\le [nT]} \left\{ \left\| \Delta^m D^p_n \left(\tfrac{j}{n}\right) \right\| \right\} \le \frac{K}{n^m}\left(\kappa_n + \lambda_n\right). \tag{43}$$

This claim is true for any $p, m$ and $j$ with $m \le p$, $m \le j \le r$ by the definition of $\kappa_n$, provided $K \ge 1$. Next, sum once over $k$ in (34) and (41) to obtain, for $r+1 \le k \le [nT]$,

$$\begin{aligned} \Delta^p D^p_n\left(\tfrac{k}{n}\right) = \Delta^p D^p_n\left(\tfrac{r}{n}\right) &+ \sum_{j=r+1}^{k} \sum_{q=0}^{r} \tfrac{1}{n^{p+1-q}} F^{(p,q)} \left(\Delta^q D_n\right)\left(\tfrac{j-1}{n}\right) \\ &+ \sum_{j=r+1}^{k} \left[ \tilde{\epsilon}^{(p)}_{n,j} - \epsilon^{(p)}_{n,j} + o_{j,p}\left(\tfrac{1}{n^{p+1}}\right) \right]. \end{aligned} \tag{44}$$

We can now show by induction on $j$ that

$$\left\| \Delta^m D^p_n\left(\tfrac{j}{n}\right) \right\| \le \frac{K}{n^m}\left(\kappa_n + \lambda_n\right) \tag{45}$$

for any $m \le p$ and any $j$ with m$\le j \le [nT]$ : assuming that (45) is true for all m$\le p$ and $1 \le j \le k-1$, it follows from (44) that

$$\begin{aligned} \left\| \Delta^p D^p_n\left(\tfrac{k}{n}\right) \right\| &\le \tfrac{\kappa_n}{n^p} + \sum_{j=r+1}^{k} \sum_{q=0}^{r} \tfrac{1}{n^{p+1}} \left\| F^{(p,q)} \right\| (r+1) K \left(\kappa_n + \lambda_n\right) + \tfrac{\lambda_n}{n^p} \\ &\le \tfrac{1}{n^p}\left(\kappa_n + \lambda_n\right) \left( 1 + (r+1) K T \sum_{q=0}^{r} \left\| F^{(p,q)} \right\| \right). \end{aligned} \tag{46}$$

(For $M$ a square matrix, $\|M\|$ denotes the operator norm corresponding to the vector norm $\|\cdot\|$ already in use). Thus (45) holds also for $m = p$ and $j = k$ provided $T > 0$ is first chosen so small that (for all $p$)

$$(r+1)T\sum_{q=0}^{r}\left\|F^{(p,q)}\right\| < 1,$$

and then $K \geq 1$ is chosen so large that

$$1 + (r+1)KT\sum_{q=0}^{r}\left\|F^{(p,q)}\right\| \leq K.$$

With these choices of $T$ and $K$, (43) has been proved for $m = p$, at time $\frac{k}{n}$. For other values of $m$, use induction backwards on $m$. So if (43) is true for $m$, use

$$\Delta^{m-1}D_n^p\left(\tfrac{k}{n}\right) = \sum_{j=r+1}^{k}\Delta^m D_n^p\left(\tfrac{j}{n}\right) + \Delta^{m-1}D_n^p\left(\tfrac{r}{n}\right) \tag{47}$$

to obtain

$$\left\|\Delta^{m-1}D_n^p\left(\tfrac{k}{n}\right)\right\| \leq \frac{1}{n^{m-1}}\left(\kappa_n + \lambda_n\right)(1+KT),$$

which shows (43) for $m-1$ at time $\frac{k}{n}$ provided $T < 1$ and $K$ is sufficiently large.

With the claim (43) established, take $m = 0$ and sum over $p$ to obtain

$$\max_{j:0\leq j\leq[nT]}\left\{\left\|X\left(\tfrac{j}{n}\right) - X_n\left(\tfrac{j}{n}\right)\right\|\right\} \leq (r+1)K\left(\kappa_n + \lambda_n\right). \tag{48}$$

Because of (42) it will follow that $X_n \overset{\text{a.s}}{\to} X$, uniformly on $[0,T]$ provided

$$\kappa_n \overset{\text{a.s}}{\to} 0. \tag{49}$$

Assuming this for the moment and repeating the argument leading to (48), one finds that $X_n \overset{\text{a.s}}{\to} X$, uniformly on an interval starting fractionally to the left of $T$ (to get the new 'initial values' in place before time $T$), at $(1-\delta)T$ say, and of length $T$. Thus $X_n \overset{\text{a.s}}{\to} X$ uniformly on $[0,(2-\delta)T]$ and continuing, the assertion $X_n \overset{\text{a.s.u.c}}{\to} X$ follows.

It remains to prove (49) and to read off the coefficients (36). The latter is easy, just use (41) and (29), isolate $\pi^{(p)}X_n\left(\frac{k}{n}\right)$ on the left, and sum over $p$

to obtain (c.f. (1))

$$X_n\left(\tfrac{k}{n}\right)=\sum_{l=0}^{r}A_{n,l}X_n\left(\tfrac{k-l-1}{n}\right)+\epsilon_{n,k}$$

with the $A_{n,l}$ as in (36) and $\epsilon_{n,k}$ as in (37).

Finally, to prove (49), note that from (30) and the initial condition $\left(\pi^{(p)}X\right)^{(q)}(0)=x^{(p,q)}$, it follows that for all $p$ and all $m,j$ with $0\le m\le p$, $m\le j\le r$,

$$n^m\Delta^m\pi^{(p)}X\left(\tfrac{j}{n}\right)=x^{(p,m)}+o(1)$$

(with $o(1)$ random, $\overset{\text{a.s}}{\to}0$). It is then clear that (49) follows if we show that (39) implies, for $p,m,j$ as above

$$n^m\Delta^m\pi^{(p)}X_n\left(\tfrac{j}{n}\right)=x^{(p,m)}+o(1) \tag{50}$$

(with $o(1)$ non-random since the $X_n\left(\frac{j}{n}\right)$ are non-random for $j\le r$). But by (39), (50) is true for $j=m$ where $0\le m\le p$, and for $m=p$ and $p<j\le r$. The simple observation (use (39))

$$n^m\Delta^m\pi^{(p)}X_n\left(\tfrac{j}{n}\right)=n^m\sum_{i=l+1}^{j}\Delta^{m+1}\pi^{(p)}X_n\left(\tfrac{i}{n}\right)+x^{(p,m)} \tag{51}$$

now permits a proof of (50) using backwards induction in $m$, starting from $m=p$: by the induction hypothesis each term in the sum in (51) is $O\left(n^{-m-1}\right)$ and the right hand side of (50) becomes $x^{(p,m)}+O\left(n^{-1}\right)$. □

The expression (36) is not particularly useful and not very illuminating! It is clear that the geometric structure of the limit process $X$ has been well and truly obscured by the autoregressive representation. It appears impossible to use (36) to recognize converging systems of autoregressive processes, so when addressing this problem in the next section, we shall rely on a different approach. However, the technique used in the proof of Proposition 3.2, will prove useful in Section 4 also.

Instead of using the method described above for determining AR-processes $X_n$ converging to $X$, one could have started with the big process $Z$ (an Itô-process) determined by $X$, and considered the standard Euler scheme [6, Chapter 9] converging to $Z$. Converting this into AR-processes converging to

$X$ is however cumbersome (the expression (36) is at least explicit), and in any case we find it more informative to work directly with $X$ as it was done above.

# 4 SUFFICIENT CONDITIONS FOR WEAK CONVERGENCE OF SEQUENCES OF AR-PROCESSES

In the previous section a concrete sequence of AR$(r+1)-$processes was exhibited, converging a.s.u.c to a given limit $X$ of order $r$. The sequence in question is of the form (1) with coefficients given by expressions of the following type (see (36))

$$A_{n,l} = \sum_{s \in R_r} \frac{1}{n^s} \beta_{s,l} \tag{52}$$

where each $\beta_{s,l} \in \mathbf{R}^{d \times d}$ and the range $R_r$ of summation always includes $s = 0$ and further consists of all values $s = p + 1 - q$ obtainable when $0 \leq p, q \leq r$. Thus

$$R_r = \begin{cases} \{0, 1\} & \text{if } r = 0 \\ \{-r+1, -r+2, \ldots, r, r+1\} & \text{if } r \geq 1. \end{cases}$$

Note that if $r \geq 2$, strictly negative values of $s$ may occur, cf. the closing remarks of Section 1.

We shall now consider arbitrary sequences $(Y_n)$ of AR$(r+1)-$processes of the form (1) with the $A_{n,l}$ as in (52), and shall then find sufficient conditions on the $\beta_{s,l}$, the errors $\epsilon_{n,k}$ and the initial conditions for each $Y_n$, that ensures that the sequence $(X_n)$ of continuous time processes derived from $(Y_n)$ as in (2) converges in distribution to a limit process $X$ of the type discussed in Section 2.

Initially we shall fix the limit $X$, and use the description (20), (24) provided by Theorem 2.2. From Lemma 3.1 and (20) it follows that for any $p$

with $0 \le p \le r$ and any $v \in V^{(p)}$, writing $\Delta = \Delta_{\frac{1}{n}}$,

$$\Delta^{p+1} v X\left(\tfrac{k}{n}\right) = \int_{M^{p+1}_{\frac{1}{n}}\left(\frac{k}{n}\right)} du_1 \cdots du_{p+1} \sum_{q=0}^{r} \left(v\alpha^{(p,q)} X\right)^{(q)} (u_{p+1}) \tag{53}$$

$$+ \int_{M^{p}_{\frac{1}{n}}\left(\frac{k}{n}\right)} du_1 \cdots du_p \, v\gamma^{(p)} \left(U\left(u_p\right) - U\left(u_p - \tfrac{1}{n}\right)\right),$$

while for $X_n$, (see (1) and (52)),

$$\Delta^{p+1} v X_n\left(\tfrac{k}{n}\right) = \sum_{l=0}^{r} \sum_{s \in R_r} \frac{1}{n^s} v\beta^{(p)}_{s,l} X_n\left(\tfrac{k-l-1}{n}\right) + v\epsilon_{n,k} \tag{54}$$

where

$$\beta^{(p)}_{s,l} = \begin{cases} \beta_{s,l} & \text{if } s \neq 0 \\ \beta_{s,l} - (-1)^l 1_{(l \le p)} \binom{p+1}{l+1} I_d & \text{if } s = 0. \end{cases} \tag{55}$$

The main problem is to find out how the geometric structure of the limit $X$, i.e the subspaces $V^{(0)} \supset \cdots \supset V^{(r)}$, is reflected in the matrices $\beta_{s,l}$. The idea we shall use is the following: the first term on the right of (53) is

$$\frac{1}{n^{p+1}} \sum_{q=0}^{r} \left(v\alpha^{(p,q)} X\right)^{(q)} \left(\tfrac{k-1}{n}\right) + o\left(\tfrac{1}{n^{p+1}}\right) \tag{56}$$

and we then aim for the sum on the right of (54), where $X_n$ is replaced by $X$, to approximate (56) with sufficient accuracy, i.e to within order $o\left(n^{-p-1}\right)$. Thus we want

$$\frac{1}{n^{p+1}} \sum_{q=0}^{r} \left(v\alpha^{(p,q)} X\right)^{(q)} \left(\tfrac{k-1}{n}\right) = \sum_{l=0}^{r} \sum_{s \in R_r} \frac{1}{n^s} v\beta^{(p)}_{s,l} X\left(\tfrac{k-l-1}{n}\right) + o\left(\tfrac{1}{n^{p+1}}\right). \tag{57}$$

For this to be possible, one must be able to Taylor expand, for an arbitrary sample path for $X$, the sum on the right of (57) around time $\frac{k-1}{n}$ with sufficient accuracy, which forces each $v\beta^{(p)}_{s,l}$ to belong to an appropriate subspace $V^{(m)}$—we must demand that $\beta^{(p)}_{s,l}$ maps $V^{(p)}$ into some $V^{(m)}$ or $\{0\}$.

We have indicated some of the conditions on the $\beta_{s,l}$, that will appear later. But the first task is to argue that the idea just sketched can be made to work.

Suppose that $X, X_n$ are as above, *but defined on the same probability space.* Define $D_n = X - X_n$.

**Lemma 4.1** *Suppose that there is a constant $T > 0$ such that for all $p$, $0 \leq p \leq r$ and $v \in V^{(p)}$,*

$$\kappa_n^p(v) \overset{\text{a.s}}{\to} 0, \quad \lambda_n^p(v) \overset{\text{a.s}}{\to} 0, \quad \rho_n^p(v) \overset{\text{a.s}}{\to} 0, \tag{58}$$

*where*

$$\begin{aligned} \kappa_n^p(v) &:= \max\left\{n^m \left|\Delta^m (vD_n)\left(\tfrac{j}{n}\right)\right| : m, j \text{ with } 0 \leq m \leq p,\ m \leq j \leq r\right\} \\ \lambda_n^p(v) &:= \max_{k:\, r+1 \leq k \leq [nT]} \left\{n^p \left|\sum_{j=r+1}^{k} v\epsilon_{n,j} - \tfrac{1}{n^p} v\gamma^{(p)} \left(U\left(\tfrac{k}{n}\right) - U(0)\right)\right|\right\} \\ \rho_n^p(v) &:= \max_{k:\, r+1 \leq k \leq [nT]} \left\{\left|\sum_{q=0}^{r} \left(v\alpha^{(p,q)} X\right)^{(q)} \left(\tfrac{k}{n}\right) - n^{p+1} \sum_{l=0}^{r} \sum_{s \in R_r} \tfrac{1}{n^s} v\beta_{s,l}^{(p)} X\left(\tfrac{k-l-1}{n}\right)\right|\right\}. \end{aligned}$$

*Assume also that for all $p$, $v \in V^{(p)}$ and $s \leq p$,*

$$\sum_{l=0}^{r} v\beta_{s,l}^{(p)} = 0. \tag{59}$$

*Then $X_n \overset{\text{a.s.u.c}}{\to} X$.*

**Remarks.** Note that (58) holds for all $v \in V^{(p)}$ iff it holds for finitely many vectors forming a base. The first condition in (58) ensures that $X_n$ is close to $X$ initially, the second controls the errors, and the third is a precise formulation of (57).

**Proof.** The proof resembles that of Proposition 3.2. For each $p$, let $E^{(p)}$ be a finite set of base vectors for $V^{(p)}$. We want first to show that for $T > 0$ sufficiently small, there is a constant $K > 0$ such that for all $p$ and $m$ with $0 \leq m \leq p \leq r$ and $v \in E^{(p)}$,

$$\max_{j:\, m \leq j \leq [nT]} \left\{\left|\Delta^m vD_n\left(\tfrac{j}{n}\right)\right|\right\} \leq \frac{K}{n^m}\zeta_n, \tag{60}$$

where

$$\zeta_n = \max\left\{\kappa_n^p(v) + \lambda_n^p(v) + T\rho_n^p(v) + vo(1) : p, v \text{ with } 0 \leq p \leq r,\ v \in E^{(p)}\right\}.$$

Since $\zeta_n \overset{\text{a.s}}{\to} 0$, it will then follow, arguing as in the proof of Proposition 3.2 (using $m = 0$ in (60)), firstly that $X_n \overset{\text{a.s}}{\to} X$, uniformly on $[0, T]$ and subsequently that $X_n \overset{\text{a.s.u.c}}{\to} X$.

In establishing (60), the important case is $m = p$. For $m < p$, argue as in the proof of Proposition 3.2, the paragraph containing (47).

Let $v \in V^{(p)}$. To show (60) for $m = p$, sum once on $k$ in (53) and (54) to obtain (cf. (44)) for $k \geq r+1$

$$\begin{aligned}
\Delta^p v D_n\left(\tfrac{k}{n}\right) = \; & \Delta^p v D_n\left(\tfrac{r}{n}\right) \\
& + \sum_{j=r+1}^{k} \left( \tfrac{1}{n^{p+1}} \left(v\alpha^{(p,q)} X\right)^{(q)} \left(\tfrac{j-1}{n}\right) - \sum_{l=0}^{r} \sum_{s \in R_r} \tfrac{1}{n^s} v\beta_{s,l}^{(p)} X\left(\tfrac{j-l-1}{n}\right)\right) \\
& + \sum_{j=r+1}^{k} \sum_{l=0}^{r} \sum_{s \in R_r} \tfrac{1}{n^s} v\beta_{s,l}^{(p)} D_n\left(\tfrac{j-l-1}{n}\right) \\
& + \tfrac{1}{n^p} v\gamma^{(p)} \left(U\left(\tfrac{k}{n}\right) - U(0)\right) + v o_k\left(\tfrac{1}{n^p}\right) - \sum_{j=r+1}^{k} v\epsilon_{n,j}
\end{aligned}$$

where $\max_{k:r+1 \leq k \leq [nT]} n^p \left|v o_k\left(n^{-p}\right)\right| \overset{\text{a.s}}{\to} 0$. Thus, with $o(1) \overset{\text{a.s}}{\to} 0$ not depending on $k$,

$$\begin{aligned}
n^p \left|\Delta^p v D_n\left(\tfrac{k}{n}\right)\right| \leq \; & \kappa_n^p(v) + \lambda_n^p(v) + T\rho_n^p(v) + vo(1) \\
& + n^p \sum_{j=r+1}^{k} \sum_{l=0}^{r} \sum_{s \in R_r,\, s \geq p+1} \tfrac{1}{n^s} \left|v\beta_{s,l}^{(p)} D_n\left(\tfrac{j-l-1}{n}\right)\right|,
\end{aligned} \tag{61}$$

where we have used assumption (59). Now, to show by induction in $k$, that this is $\leq$ the expression on the right of (60), one must be able to control the triple sum in (61). By the induction hypothesis, estimates of $\left|\Delta^m \tilde{v} D_n\left(\frac{j-l-1}{n}\right)\right|$ of the form (60) are available for all $\tilde{p}$, $\tilde{v} \in E^{(\tilde{p})}$, in particular, for $m = 0$, $\tilde{v} \in E^{(0)}$

$$\left|\tilde{v} D_n\left(\tfrac{j-l-1}{n}\right)\right| \leq K\zeta_n.$$

For the given $v \in E^{(p)}$, write $v\beta_{s,l}^{(p)}$ as a linear combination of the vectors in $E^{(0)}$ (which span $\mathbf{R}^{1\times d}$). It is then clear that there is a constant $C \geq 0$ such that for all $p, s, l$

$$\left|v\beta_{s,l}^{(p)} D_n\left(\tfrac{j-l-1}{n}\right)\right| \leq CK\zeta_n$$

and hence

$$\begin{aligned}
\sum_{l=0}^{r} \sum_{s \in R_r,\, s \geq p+1} \tfrac{1}{n^s} \left|v\beta_{s,l}^{(p)} D_n\left(\tfrac{j-l-1}{n}\right)\right| & \leq \sum_{l=0}^{r} \sum_{s \in R_r,\, s \geq p+1} \tfrac{1}{n^s} CK\zeta_n \\
& \leq \tfrac{1}{n^{p+1}} C_1 K\zeta_n
\end{aligned}$$

with $C_1$ a constant not depending on $T$ or $p$. With the sum over $j$ in (61) containing $\leq nT$ terms we may now proceed as in the proof of Proposition 3.2, (45), and obtain (60) at time $\frac{k}{n}$ for $T$ sufficiently small and $K$ sufficiently large. $\square$

We need one further result, the following curious combinatorial identity.

**Lemma 4.2** *For non-negative integers $i, p$ with $0 \leq i \leq p+1$, define*

$$S_{p,i} = (-1)^i \sum_{l=0}^{p} (-1)^l l^i \binom{p+1}{l+1}.$$

*Then*

$$\begin{aligned} S_{p,i} &= 1 && (0 \leq i \leq p) \\ S_{p,p+1} &= 1-(p+1)! \end{aligned}$$

**Note.** Use $0^0 = 1$ in the definition of $S_{p,0}$.
**Sketch of proof.** Because

$$(p+1)\binom{p}{l+1} - p\binom{p+1}{l+1} = -l\binom{p+1}{l+1}$$

one immediately verifies the recursion formula

$$S_{p,i+1} = (p+1)\, S_{p-1,i} - pS_{p,i}$$

for $i \geq 0,\ p \geq 1$. Now proceed by induction on $p$ and $i$. $\square$

Let now $(Y_n)_{n\geq 1}$ be a sequence of AR$(r+1)-$processes of the form (1) with coefficients of the form (52), and define $X_n(t) = Y_{n,[nt]}$ as usual. Recall the definition (55) of $\beta_{s,l}^{(p)}$.

**Theorem 4.3** *Suppose there exists subspaces $\mathbf{R}^{1\times d} = V^{(0)} \supset V^{(1)} \supset \cdots \supset V^{(r)}$ (with $V^{(r)} = \{0\}$ allowed), linear maps $\gamma^{(p)} : V^{(p)} \to \mathbf{R}^{1\times d}$ satisfying* (22) *and linear functionals $f^{(p,q)} : V^{(p)} \to \mathbf{R}$ satisfying* (23)*, such that the following three conditions hold:*

(i) (Initial conditions) *For all $p,\ v \in V^{(p)}$,*

$$n^j \Delta^j_{\frac{1}{n}} (vX_n)\left(\tfrac{j}{n}\right) = vf^{(p,j)} \qquad (0 \leq j \leq p)$$

$$n^j \Delta^p_{\frac{1}{n}} (vX_n) \left(\tfrac{j}{n}\right) = v f^{(p,p)} \qquad (p < j \leq r).$$

(ii) (Error convergence) *There is a continuous $\mathbf{R}^{d\times 1}-$valued semimartingale $U$, defined on some filtered space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ with $U(0) = 0$ a.s, such that if for any $p$, $v \in V^{(p)}$ one defines the scalar valued processes*

$$U_n^{(p)}(v) := n^p \sum_{k=r+1}^{[n\cdot]} v\epsilon_{n,k}$$

*it holds for all $N \in \mathbf{N}$, $0 \leq p_1, p_2, \ldots, p_N \leq r$, $v_1 \in V^{(1)}, v_2 \in V^{(2)}, \ldots, v_N \in V^{(N)}$ that*

$$\left(U_n^{(p_1)}(v_1), U_n^{(p_2)}(v_2) \ldots, U_n^{(p_N)}(v_N)\right) \xrightarrow{\mathcal{D}} \left(v_1\gamma^{(p_1)}U, v_2\gamma^{(p_2)}U, \ldots, v_N\gamma^{(p_N)}U\right). \tag{62}$$

(iii) (Subspace conditions). *For all $p, l$ with $0 \leq p, l \leq r$*

$$\beta_{s,l}^{(p)} : V^{(p)} \to V^{(p+1-s)} \qquad (p + 1 - r \leq s \leq p). \tag{63}$$

*Also, for all $0 \leq l \leq r$,*

$$\beta_{s,l} : V^{(r+s)} \to \{0\} \qquad (\min R_r \leq s < 0), \tag{64}$$

$$\left(\beta_{0,l} - (-1)^l \tbinom{r+1}{l+1} I_d\right) : V^{(r)} \to \{0\}. \tag{65}$$

*Finally, for $\min R_r \leq s < 0$*

$$\begin{aligned} \sum_{l=0}^{r} l^q \beta_{s,l} : V^{(0)} \to \{0\} \qquad & (0 \leq q \leq -s - 1) \\ \sum_{l=0}^{r} l^q \beta_{s,l} : V^{(q+s)} \to \{0\} \qquad & (-s \leq q \leq r - 1) \end{aligned} \tag{66}$$

*while for $s = 0$*

$$\sum_{l=0}^{r} l^q \beta_{0,l} - (-1)^q I_d : V^{(q)} \to \{0\} \qquad (0 \leq q \leq r - 1) \tag{67}$$

*and for $1 \leq s \leq r$*

$$\sum_{l=0}^{r} l^q \beta_{s,l} : V^{(q+s)} \to \{0\} \qquad (0 \leq q \leq r - s). \tag{68}$$

*Then it holds that*

$$X_n \xrightarrow{\mathcal{D}} X,$$

*where $X$ is given as the unique solution to* (20) *with initial conditions* (24), *and where the linear maps $\alpha^{(p,q)}$ from* (20) *are determined by*

$$\alpha^{(p,q)} = \frac{1}{q!}\sum_{l=0}^{r}(-l)^q \beta^{(p)}_{p+1-q,l} \tag{69}$$

*and satisfy* (21).

**Remarks.**

- Condition (i) should be compared with (39). Because of (23), it should be clear that the equations in (i) are consistent and define $X_n\left(\frac{j}{n}\right)$ for $0 \le j \le r$.

- Condition (ii) is a condition of process convergence of the sequence of errors from (1). It may be given in a different form, more useful for applications. Consider a decomposition (9) of $\mathbf{R}^{1\times d}$, which satisfies (10), and let $\pi^{(p)}$ denote the corresponding 'projections'. Now suppose that for each $p$ there is a $H^{(p)}$−valued, continuous semimartingale $U^{(p)}$ such that if

$$U_n^{(p)} := n^p \sum_{k=r+1}^{[n\cdot]} \pi^{(p)} \epsilon_{n,k}$$

it holds that

$$\left(U_n^{(0)}, U_n^{(1)}, \ldots, U_n^{(r)}\right) \xrightarrow{\mathcal{D}} \left(U^{(0)}, U^{(1)}, \ldots, U^{(r)}\right), \tag{70}$$

(Skorohod convergence on $\left(\mathbf{R}^{d\times 1}\right)^{r+1}$), then (ii) is satisfied: for $v \in V^{(p)}$

$$\begin{aligned} n^p \sum_{k=r+1}^{[n\cdot]} v\epsilon_{n,k} &= \sum_{q=p}^{r} v\left(n^p \sum_{k=r+1}^{[n\cdot]} \pi^{(q)}\epsilon_{n,k}\right) \\ &\xrightarrow{\mathcal{D}} vU^{(p)}, \end{aligned}$$

and similarly for the joint convergence in (62). Thus, defining $U = \sum U^{(p)}$, this limit is $v\pi^{(p)}U$, so (ii) holds with $\gamma^{(p)} = \pi^{(p)}$ and these $\gamma^{(p)}$ of course satisfy (22). It is clear that (ii) implies (70), so the two conditions are equivalent.

- Suppose that for each $n$, the $\epsilon_{n,k}$ for $k \geq r+1$ are i.i.d with $E\epsilon_{n,k} = 0$, Cov $\epsilon_{n,k} = \Gamma_n$. Then, using the preceding remark, it is seen that condition (ii) is satisfied provided for all $p, q$ there is $\Xi^{(p,q)} \in \mathbf{R}^{d\times d}$ with ${\Xi^{(p,q)}}^T = \Xi^{(q,p)}$ such that

$$\pi^{(p)}\,\Gamma_n\,{\pi^{(q)}}^T = \frac{1}{n^{p+q+1}}\pi^{(p)}\,\Xi^{(p,q)}\,{\pi^{(q)}}^T. \tag{71}$$

In that case, $U$ is a Brownian motion with 0 drift and covariance matrix $\Gamma$ (see the first remark after Proposition 2.1),

$$\Gamma = \sum_{p=0}^{r}\sum_{q=0}^{r}\pi^{(p)}\,\Xi^{(p,q)}\,{\pi^{(q)}}^T.$$

($^T$ stands for 'transpose'). (71) shows clearly how the errors $\epsilon_{n,k}$ must be normalized differently in different directions.

**Proof of Theorem** 4.3. We shall use Lemma 4.1. In order to do this, first note that by the Skorohod embedding theorem, [8] or e.g [3, Theorem 3.1.8], we may assume that for any $p$, $v \in V^{(p)}$,

$$U_n^{(p)}(v) \overset{\text{a.s.u.c}}{\to} v\gamma^{(p)}U. \tag{72}$$

Indeed, since (70) holds, the Skorohod embedding allows us to assume that all $U_n^{(p)}$ and $U^{(p)}$ are defined on the same probability space with

$$\left(U_n^{(0)}, U_n^{(1)}, \ldots, U_n^{(r)}\right) \overset{\text{a.s}}{\to} \left(U^{(0)}, U^{(1)}, \ldots, U^{(r)}\right). \tag{73}$$

Since the limit process is continuous, the a.s convergence of the sequence $\left(U_n^{(p)}\right)_{n,p}$ of random variables with values in the Skorohod space $D^{r+1}_{\mathbf{R}^{d\times 1}}(\mathbf{R}_0) \sim D_{\mathbf{R}^{d\times(r+1)}}(\mathbf{R}_0)$ translates into a.s.u.c convergence, and (72) follows.

Assuming, as we now may, that (72) holds, the proof will be complete if we show that $X_n \overset{\text{a.s.u.c}}{\to} X$, and this follows from Lemma 4.1 if we verify conditions (58) and (59). But the first assertion in (58) follows from (i) in the Theorem, copying the argument for (49) in the proof of Proposition 3.2. The second assertion in (58) follows directly from (72). To obtain the last assertion in (58), it suffices to show that for any continuous path $w : \mathbf{R}_0 \to \mathbf{R}^{d\times 1}$ with

$t \to vw(t)$, $p$ times continuously differentiable for any $p$, $v \in V^{(p)}$ (so $w$ is a typical sample path for $X$),

$$\lim_{n\to\infty} \left| \sum_{q=0}^{r} \left(v\alpha^{(p,q)}w\right)^{(q)}(t) - \sum_{l=0}^{r}\sum_{s\in R_r} n^{p+1-s}\left(v\beta_{s,l}^{(p)}w\right)\left(t-\tfrac{l}{n}\right)\right| = 0 \qquad (74)$$

uniformly in $t \in [0,T]$ for any $T > 0$.

For this to work, we first demand that the $(s,l)'$th term in the double sum, can be Taylor expanded around $t$ to within order $o(1)$, which is a problem only if $s \le p$, i.e we need that

$$v\beta_{s,l}^{(p)} \in V^{(p+1-s)} \qquad \left(v \in V^{(p)},\ 1 \le l \le r,\ p+1-r \le s \le p\right) \qquad (75)$$

$$v\beta_{s,l}^{(p)} = 0 \qquad \left(v \in V^{(p)},\ 1 \le l \le r,\ \min R_r \le s \le p-r\right). \qquad (76)$$

(Recall that $\min R_r$ is 0 if $r = 0$ and $-r+1$ if $r \ge 1$). Next, performing the expansion, the double sum becomes

$$\sum_{l=0}^{r}\sum_{s=p+1-r}^{p+1} n^{p+1-s}\sum_{q=0}^{p+1-s}\frac{1}{q!}\left(-\frac{l}{n}\right)^q\left(v\beta_{s,l}^{(p)}w\right)^{(q)}(t) + o(1) \qquad (77)$$

which for (74) to hold should equal

$$\sum_{q=0}^{r}\left(v\alpha^{(p,q)}w\right)^{(q)}(t) + o(1). \qquad (78)$$

View both (77) and (78) as expansions in powers of $n$ and the order $q$ of differentiation. Then match the two expressions term by term to obtain the conditions

$$v\left(\sum_{l=0}^{r} l^q\beta_{s,l}^{(p)}\right) = 0 \qquad \left(v \in V^{(p)},\ p+1-r \le s \le p,\ 0 \le q \le p-s\right), \quad (79)$$

$$v\left(\sum_{l=0}^{r}\frac{1}{q!}(-l)^q\beta_{p+1-q,l}^{(p)}\right) = v\alpha^{(p,q)} \qquad \left(v \in V^{(p)},\ 0 \le q \le r\right). \qquad (80)$$

We have now shown that (74) holds if (75), (76), (79) and (80) hold. But (75) is just (63), and (76) is the same as (65) in the case $s = 0$ and follows from (64) in the case $s < 0$ since $v \in V^{(p)} \subset V^{(r+s)}$. Further, (80) agrees with the definition (69) of the $\alpha^{(p,q)}$ and (79) follows from (66), (67) and (68) as we shall now see: if $s \neq 0$, (79) just reads $v\left(\sum_l l^q \beta_{s,l}\right) = 0$ for $v \in V^{(p)}$, an assertion which is strongest for $p$ as small as possible in agreement with (66) if $s < 0$ and with (68) if $s > 0$. For $s = 0$ the left hand side of (79) is the same as

$$\begin{aligned} & v\left(\sum_{l=0}^{r} l^q \beta_{0,l} - \sum_{l=0}^{p} l^q (-1)^l \binom{p+1}{l+1} I_d\right) \\ = \; & v\sum_{l=0}^{r} l^q \beta_{0,l} - (-1)^q S_{p,q} v \\ = \; & v\sum_{l=0}^{r} l^q \beta_{0,l} - (-1)^q v \end{aligned}$$

for $v \in V^{(p)}$ by Lemma 4.2. The smallest value of $p$ allowed is $p = q$, and we see that (79) for $s = 0$ follows from (67).

In order for us to apply Lemma 4.1, we still have to verify (59), but this is just (79) for $q = 0$. Of the assertions from the theorem it now only remains to prove, that the $\alpha^{(p,q)}$ map $V^{(p)}$ into $V^{(q)}$, which follows from (63) with $s = p+1-q$, and that they satisfy (21). But if $v \in V^{(p+1)}$, for $q \neq p+1$ we get

$$v\alpha^{(p,q)} = \frac{1}{q!} v \left(\sum_{l=0}^{r} (-l)^q \beta_{p+1-q,l}\right) = 0$$

in the case $q < r$ by (79) with $p$ replaced by $p+1$, $s = p+1-q$, and in the case $q = r$ by (76) with $p$ replaced by $p+1$, $s = p+1-r$, while if $q = p+1$

$$v\alpha^{(p,p+1)} = v \sum_{l=0}^{r} \tfrac{1}{(p+1)!}(-l)^{p+1} \left(\beta_{0,l} - (-1)^l 1_{(l \leq p)} \binom{p+1}{l+1} I_d\right).$$

By (79) with $p$ replaced by $p+1$, $s = 0$, $q = p+1$,

$$v \sum_{l=0}^{r} l^{p+1} \beta_{0,l}^{(p+1)} = 0$$

for $v \in V^{(p+1)}$. Thus

$$v\alpha^{(p,p+1)} = v \sum_{l=0}^{r} \tfrac{1}{(p+1)!}(-l)^{p+1} \left((-1)^l 1_{(l \leq p+1)} \binom{p+2}{l+1} - (-1)^l 1_{(l \leq p)} \binom{p+1}{l+1}\right)$$

and with the notation and assertion of Lemma 4.2, this

$$\begin{aligned} &= v \tfrac{1}{(p+1)!} \left(S_{p+1,p+1} - S_{p,p+1}\right) \\ &= v \end{aligned}$$

as desired. $\square$

Of the many identities appearing in the subspace conditions of the Theorem we emphasize one: taking $q = 0$ in (67) gives

$$\sum_{l=0}^{r} \beta_{0,l} = I_d.$$

Referring back to (52) and (1), this shows that approximately $X_n\left(\frac{k}{n}\right)$ is represented as an average of the $X_n\left(\frac{k-l-1}{n}\right)$ for $0 \le l \le r$.

We shall conclude with two examples that illustrate the effect of the subspace conditions in Theorem 4.3.

**Example.** Consider the case $V^{(1)} = \{0\}$ with $r$ arbitrary (so $r = 0$ gives the standard case from Proposition 1.1). One finds that (63)–(68) are equivalent to

$$\beta_{0,l} = \begin{cases} I_d & (l = 0), \\ 0_d & (l \ge 1), \end{cases} \qquad \beta_{s,l} = 0_d \quad (s < 0, \text{ all } l).$$

Thus, if the initial conditions and the errors behave, there will be convergence of the systems

$$X_n\left(\tfrac{k}{n}\right) = X_n\left(\tfrac{k-1}{n}\right) + \sum_{l=0}^{r}\sum_{s=1}^{r+1} \frac{1}{n^s}\beta_{s,l} X_n\left(\tfrac{k-l-1}{n}\right) + \epsilon_{n,k}. \tag{81}$$

But very few of the coefficients will affect the limit: by (69)

$$\alpha^{(0,0)} = \sum_{l=0}^{r} \beta_{1,l}, \qquad \alpha^{(0,q)} = 0 \quad (q \ge 1),$$

and all other $\alpha^{(p,q)}$ are required to map $\{0\}$ into $\{0\}$, which is automatic. So all terms with $s \ge 2$ in (81) are immaterial, and for $s = 1$ only the sum $\sum \beta_{1,l}$ is relevant for the structure of the limit process. $\square$

**Example.** Take $r = 1$ and assume that $d_1 := \dim V^{(1)} \ge 1$. The subspace conditions involve $\beta_{0,l}$ and $\beta_{1,l}$ only, not the $\beta_{2,l}$. One finds that the $\beta_{0,l}$ must satisfy

$$(\beta_{0,0} - I_d) : V^{(0)} \to V^{(1)}, \qquad \beta_{0,1} : V^{(0)} \to V^{(1)}, \tag{82}$$

$$(\beta_{0,0} - 2I_d) : V^{(1)} \to \{0\}, \qquad (\beta_{0,1} + I_d) : V^{(1)} \to \{0\}, \tag{83}$$

$$\beta_{0,0} + \beta_{0,1} = I_d, \tag{84}$$

while for the $\beta_{1,l}$ it holds that

$$\beta_{1,l} : V^{(1)} \to V^{(1)}, \qquad (\beta_{1,0} + \beta_{1,1}) : V^{(1)} \to \{0\}\,.$$

We can learn a little more about the structure of the $\beta_{0,l}$. Let $R^{1\times d} = H^{(0)} \oplus V^{(1)}$ be a decomposition as in (10), with 'projections' $\pi^{(0)}, \pi^{(1)}$. Then from (82) we get

$$\beta_{0,1} = \gamma\pi^{(1)} \tag{85}$$

for some matrix $\gamma$. In particular, $\operatorname{rank}\beta_{0,1} \leq d_1$ and since by (83), $\pi^{(1)}\left(\beta_{0,1} + I_d\right) = 0_d$ it follows that

$$\operatorname{rank}\beta_{0,1} = d_1.$$

Also by (83), $\beta_{0,1} + I_d = \pi^{(0)}\left(\beta_{0,1} + I_d\right)$, so $\operatorname{rank}\left(\beta_{0,1} + I_d\right) \leq d - d_1$. With $\left(\beta_{0,1} + I_d\right)\pi^{(0)} = \beta_{0,1}\left(I_d - \pi^{(1)}\right) + \pi^{(0)} = \pi^{(0)}$ as follows from (85), we find

$$\operatorname{rank}\left(\beta_{0,1} + I_d\right) = d - d_1 \tag{86}$$

and we have shown that

$$I_d = \left(\beta_{0,1} + I_d\right) - \beta_{0,1}$$

is a decomposition of unity with

$$\operatorname{rowspan}\beta_{0,1} = V^{(1)}. \tag{87}$$

If conversely $\beta_{0,1}$ is such that (87) and (86) hold, then $\beta_{0,1}\left(\beta_{0,1} + I_d\right) = 0$, and so, defining $\beta_{0,0}$ by (84), (82) and (83) hold.

The maps $\alpha^{(p,q)}$ are given as follows

$$\begin{aligned}
\alpha^{(0,0)} &= \beta_{1,0} + \beta_{1,1} \\
\alpha^{(0,1)} &= -\beta_{0,1} \\
\alpha^{(1,0)} &= \beta_{2,0} + \beta_{2,1} \\
\alpha^{(1,1)} &= -\beta_{1,1}.
\end{aligned}$$

In particular, for the $\beta_{2,l}$ only the sum $\beta_{2,0} + \beta_{2,1}$ and how it acts on $V^{(1)}$ is of interest. □

# References


[1] Chan, N.H. and Wei, C.Z. (1987). Asymptotic Inference for Nearly Nonstationary AR(1) Processes. *Ann. Statist.* **15**, 1050-1063.

[2] Doob, J.L. (1944). The Elementary Gaussian Processes. *Ann. Math. Statist.* **15**, 229-282.

[3] Ethier, S.N. and Kurtz, T.G. (1986). *Markov Processes: Characterization and Convergence.* Wiley, New York.

[4] Jacobsen, M. (1991). *Homogeneous Gaussian Diffusions in Finite Dimensions.* Preprint 3, Institute of Mathematical Statistics, University of Copenhagen.

[5] Jacobsen, M. (1993). A Brief Account of the Theory of Homogeneous Gaussian Diffusions in Finite Dimensions. In: *Frontiers in Pure and Applied Probability. Proceedings of the Third Finnish-Soviet Symposium on Probability and Mathematical Statistics.* H. Niemi, G. Högnäs, A.N. Shiryayev and A.V. Melnikov (Eds). VSP, Utrecht and TVP Science Publishers, Moscow, pp. 86-94.

[6] Kloeden, P.E. and Platen, E. (1992). *Numerical Solution of Stochastic Differential Equations.* Applications of Mathematics 23, Springer, Berlin.

[7] Priestley, M.B. (1981). *Spectral Analysis and Time Series, Vol. 1.* Academic Press, London.

[8] Skorohod, A.V. (1956). Limit Theorems for Stochastic Processes. *Theory Probab. Appl.* **1**, 261-290.

[9] Stockmarr, A. and Jacobsen, M. (1994). Gaussian Diffusions and Autoregressive Processes: Weak Convergence and Statistical Inference. *Scand. J. Statist.* **21**, (to appear).

## Preprints 1993

COPIES OF PREPRINTS ARE OBTAINABLE FROM THE AUTHOR OR FROM THE INSTITUTE OF MATHEMATICAL STATISTICS, UNIVERSITETSPARKEN 5, DK-2100 COPENHAGEN Ø, DENMARK. TELEPHONE +45 35 32 08 99.

No. 1 Hansen, Henrik and Johansen, Søren: Recursive Estimation in Cointegration VAR-Models.

No. 2 Stockmarr, A. and Jacobsen, M.: Gaussian Diffusions and Autoregressive Processes: Weak Convergence and Statistical Inference.

No. 3 Nishio, Atsushi: Testing for a Unit Root against Local Alternatives

No. 4 Tjur, Tue: StatUnit - An Alternative to Statistical Packages?

No. 5 Johansen, Søren: Likelihood Based Inference for Cointegration of Non-Stationary Time Series.

## Preprints 1994

COPIES OF PREPRINTS ARE OBTAINABLE FROM THE AUTHOR OR FROM THE INSTITUTE OF MATHEMATICAL STATISTICS, UNIVERSITETSPARKEN 5, DK-2100 COPENHAGEN Ø, DENMARK. TELEPHONE 45 35 32 08 99, FAX 45 35 32 07 72.

No. 1 Jacobsen, Martin: Weak Convergence of Autoregressive Processes.