Journal of Multivariate Analysis 155 (2017) 180–186

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Stationary Gaussian Markov processes as limits of stationary autoregressive time series

Philip A. Ernst [a,*], Lawrence D. Brown [b], Larry Shepp [†], Robert L. Wolpert [c]

[a] Department of Statistics, Rice University, Houston, TX 77005, USA
[b] Department of Statistics, University of Pennsylvania, Philadelphia, PA 19104, USA
[c] Department of Statistical Science, Duke University, Durham, NC 27710, USA

ARTICLE INFO

Article history:
Received 26 June 2016
Available online 3 January 2017

AMS 2010 subject classifications:
primary 60G10
secondary 60G15

Keywords:
Continuous autoregressive processes
Stationary Gaussian Markov processes
Stochastic differential equations

ABSTRACT

We consider the class, $\mathbf{C}_p$, of all zero mean stationary Gaussian processes, $\{Y_t : t \in (-\infty, \infty)\}$ with p derivatives, for which the vector valued process $\{(Y_t^{(0)}, \ldots, Y_t^{(p)}) : t \geq 0\}$ is a $p+1$-vector Markov process, where $Y_t^{(0)} = Y(t)$. We provide a rigorous description and treatment of these stationary Gaussian processes as limits of stationary AR(p) time series.

1. Introduction

In many data-driven applications in both the natural sciences and in finance, time series data are often discretized prior to analysis and are then formulated using autoregressive models. The theoretical and applied properties of the convergence of discrete autoregressive ("AR") processes to their continuous analogs (continuous autoregressive or "CAR" processes) have been well studied by many mathematicians, statisticians, and economists; see, e.g., [1,2,4,8,12]. For references on stochastic differential equations, which underlie the theory of CAR processes, we refer to [6,7,14,15].

A special class of autoregressive processes are the discrete-time zero-mean stationary Gaussian Markovian processes on $\mathbb{R}$. The continuous time analogs of these processes are documented in Chapter 10 of [11] and in pages 207–212 of [13]. For processes in this class, the sample paths possess $p-1$ derivatives at each value of t, and the evolution of the process following t depends in a linear way only on the values of these derivatives at t. Notationally, we term such a process as a member of the class $\boldsymbol{C}_p$. For convenience, we will use the notation CAR(p) $= \boldsymbol{C}_p$. The standard Ornstein–Uhlenbeck process is of course a member of $\boldsymbol{C}_1$, and hence CAR(p) processes can be described as a generalization of the Ornstein–Uhlenbeck process.

It is well understood that the Ornstein–Uhlenbeck process is related to the usual Gaussian AR(1) process on a discrete-time index, and that an Ornstein–Uhlenbeck process can be described as a limit of appropriately chosen AR(1) processes; see [7]. In an analogous fashion we show that processes in $\boldsymbol{C}_p$ are related to AR(p) processes and can be described as limits of an appropriately chosen sequence of AR(p) processes.

* Corresponding author.
E-mail addresses: philip.ernst@rice.edu (P.A. Ernst), lbrown@wharton.upenn.edu (L.D. Brown), shepp@wharton.upenn.edu (L. Shepp), wolpert@stat.duke.edu (R.L. Wolpert).
† Deceased April 23, 2013.

http://dx.doi.org/10.1016/j.jmva.2016.12.008

Of course, there is also extensive literature on the weak convergence of discrete-time time series processes, particularly that of ARMA and GARCH processes. For example, Duan [5] considers the diffusion limit of an augmented GARCH process and Lorenz [9] discusses (in Chapter 3) limits of ARMA processes. However, to the best of our knowledge, none of these references (nor simplifications of their results) discuss how to correctly approximate $\boldsymbol{C}_p$ by discrete AR(p) processes, and thus this is the goal of our paper.

Section 2 begins by reviewing the literature on CAR(p) processes, recalling three equivalent definitions of the processes in $\boldsymbol{C}_p$. Section 3 discusses how to correctly approximate $\boldsymbol{C}_p$ by discrete AR(p) processes. The Appendix contains the proof of our main result, Theorem 3.2.

2. Equivalent descriptions of the class $\boldsymbol{C}_p$

There are three distinct descriptions of processes comprising the class $\boldsymbol{C}_p$, which are documented on p. 212 of [13] but in different notation. On pp. 211–212, Rasmussen and Williams [13] prove that these descriptions are equivalent ways of describing the same class of processes. The first description matches the heuristic description given in the introduction. The remaining descriptions provide more explicit descriptions that can be useful in construction and interpretation of these processes. In all the descriptions $Y = \{Y(t) : t \in [0, \infty)\}$ symbolizes a zero-mean Gaussian process on $[0, \infty)$.

In the present paper we use the first of the three equivalent descriptions in [13], as follows. Let Y be stationary. The sample paths are continuous and are $p-1$ times differentiable, a.e., at each $t \in [0, \infty)$. (The derivatives at $t = 0$ are defined only from the right. At all other values of t, the derivatives can be computed from either the left or the right, and both right and left derivatives are equal.) For each $i \in \{1, \ldots, p-1\}$, we denote the ith derivative at t by $Y^{(i)}(t)$. At any $t_0 \in (0, \infty)$, the conditional evolution of the process $\{Y(t) : t \in [0, t_0]\}$ depends in a linear way only on the set of values $\{Y^{(i)}(t_0) : i \in \{0, \ldots, p-1\}\}$. The above can be formalized as follows: let $\{(Y_t^{(0)}, \ldots, Y_t^{(p-1)}) : t \geq 0\}$ denote the values of a mean zero Itô vector diffusion process defined by the system of equations

$$
\begin{aligned}
dY_t^{(i-1)} &= Y_t^{(i)} dt, \quad t > 0,\ i = 1, \ldots, p-1 \\
dY_t^{(p-1)} &= \sum_{i=0}^{p-1} a_{i+1} Y_t^{(i)} dt + \sigma dW_t
\end{aligned}
\tag{2.1}
$$

for all $t > 0$, where W_t is the Wiener process, $\sigma > 0$. Then let $Y(t) = Y_t^{(0)}$.

2.1. Characterization of stationarity via (2.1)

The system in (2.1) is linear. Stationarity of vector-valued processes described in such a way has been studied elsewhere; see in particular Theorem 5.6.7 on p. 357 in [7]. The coefficients in (2.1) that yield stationarity can be characterized via the characteristic polynomial of the matrix Λ, where $|\Lambda - \lambda I|$ is

$$\lambda^p - a_p \lambda^{p-1} - \cdots - a_2 \lambda - a_1 = 0. \tag{2.2}$$

The process is stationary if and only if all the roots of Eq. (2.2) have strictly negative real parts.

In order to discover whether the coefficients in (2.1) yield a stationary process it is thus necessary and sufficient to check whether all the roots of Eq. (2.2) have strictly negative real parts. In the case of $\boldsymbol{C}_2$ the condition for stationarity is quite simple, namely that a_1, a_2 should lie in the quadrant $a_1 < 0, a_2 < 0$. The covariance functions for $\boldsymbol{C}_2$ can be found in [11, p. 326]. For higher order processes the conditions for stationarity are not so easily described. Indeed, for $\boldsymbol{C}_3$ it is necessary that $a_1 < 0, a_2 < 0, a_3 < 0$ simultaneously, but the set of values for which stationarity holds is not the entire octant. For larger p one needs to study the solutions of the higher order polynomial in Eq. (2.2).

3. Weak convergence of the $\boldsymbol{h}$-AR(2) process to CAR(2) process

3.1. Discrete time analogs of the CAR processes

We now turn our focus to describing the discrete time analogs of the CAR processes and the expression of the CAR processes as limits of these discrete time processes. In this section, we discuss the situation for $p = 2$. Define the h-AR(2) processes on the discrete time domain $\{0, h, 2h, \ldots\}$ via

$$X_t = b_1^h X_{t-h} + b_2^h X_{t-2h} + \varsigma^h Z_t, \tag{3.1}$$

with $Z_t \sim$ IID $\mathcal{N}(0, 1)$ for all $t = 2h, 3h, \ldots$. The goal is to establish conditions on the coefficients b_1^h, b_2^h and ς^h so that these AR(2) processes converge to the continuous time CAR(2) process as in the system of equations given in (2.1). We then discuss some further features of these processes.

To see the similarity of the h-AR(2) process in (3.1) with the CAR(2) process of (2.1), we introduce the corresponding h-VAR(2) processes $\Delta^h_{0;t}$, $\Delta^h_{1;t}$ defined via

$$\begin{aligned} \Delta^h_{0;t} - \Delta^h_{0;t-h} &= h\Delta^h_{1;t}, \\ \Delta^h_{1;t} - \Delta^h_{1;t-h} &= (c^h_1\Delta^h_{0;t-h} + c^h_2\Delta^h_{1;t-h})h + \xi^h Z_t \end{aligned} \tag{3.2}$$

with $Z_t \sim$ IID $\mathcal{N}(0, 1)$ for all $t = h, 2h, \ldots$ and

$$\Pr\{(\Delta^h_{0;0}, \Delta^h_{1;0}) \in \Gamma\} = \nu^h_2(\Gamma) \quad \text{for all } \Gamma \in \mathcal{B}(\mathbb{R}^2), \tag{3.3}$$

where ν^h_2 is a probability measure on $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$ for the Borel σ-algebra $\mathcal{B}(\mathbb{R}^2)$. We assume also that as $h \downarrow 0$, $(\Delta^h_{0;0}, \Delta^h_{1;0})$ converges in distribution to a random variable pair $(\Delta^0_{0;0}, \Delta^0_{1;0})$ with probability measure ν^0_2 on $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$. The above assumptions for the initial distribution follow Assumption 3 of [10] as well as the discretized process defined by (2.22)–(2.24) therein.

From (3.2) we see that

$$\begin{aligned} \Delta^h_{0;t} &= \Delta^h_{0;t-h} + h\Delta^h_{1;t} \\ &= \Delta^h_{0;t-h} + h\Delta^h_{1;t-h} + (c^h_1\Delta^h_{0;t-h} + c^h_2\Delta^h_{1;t-h})h^2 + \xi^h hZ_t \\ &= (2 + c^h_1h^2 + c^h_2h)\Delta^h_{0;t-h} - (1 + c^h_2h)\Delta^h_{0;t-2h} + \xi^h hZ_t. \end{aligned}$$

This shows that the h-AR(2) process of (3.1) is equivalent to the h-VAR(2) in (3.2) with

$$b^h_1 \triangleq c^h_1h^2 + c^h_2h + 2, \qquad b^h_2 \triangleq -c^h_2h - 1, \qquad \varsigma^h \triangleq \xi^h h, \tag{3.4}$$

or, equivalently,

$$c^h_1 \triangleq h^{-2}(b^h_1 + b^h_2 - 1), \qquad c^h_2 \triangleq h^{-1}(-1 - b^h_2), \qquad \xi^h \triangleq h^{-1}\varsigma^h. \tag{3.5}$$

In Theorems 3.1 and 3.2, we consider weak convergence in the same sense as that of [10]; see footnote 7 on p. 13 therein.

Theorem 3.1. *Consider a sequence of h-AR (2) processes of (3.1) with coefficients given by (3.4), where $c^h_j \to a_j \neq 0$ for $j = 1, 2$ and $\xi^h/\sqrt{h} \to \sigma$ as $h \downarrow 0$. We also assume that as $h \downarrow 0$, $(\Delta^h_{0;0}, \Delta^h_{1;0})$ converges in distribution to a random variable pair $(\Delta^0_{0;0}, \Delta^0_{1;0})$ with probability measure ν^0_2 on $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$. Then the sequence of h-AR (2) processes of (3.1) converges in distribution to the CAR (2) process of (2.1), where $\{W_t : t \geq 0\}$ is a one-dimensional Brownian motion, independent of $\left(Y_0, Y_0^{(1)}\right)$, and $\Pr\{(Y_0, Y_0^{(1)}) \in \Gamma\} = \nu^0_2(\Gamma)$ for any $\Gamma \in \mathcal{B}(\mathbb{R}^2)$.*

Proof. This proof is a special case of Theorem 3.2 and is thus omitted. □

Remark 3.1. Theorems 2.1 and 2.2 in [10] are explicitly stated for real-valued processes but apply to vector-valued processes as well. One only needs to explicitly allow the processes to be vector-valued, and to write the regularity conditions to allow for the full cross-covariance of the vector-valued observations, rather than just ordinary covariance functions. Our processes are much better behaved than the most general type of process considered in [10] since our error variance is constant (depending only on h) and our distributions are Gaussian, and hence very light tailed. Thus both of Theorems 2.1 and 2.2 in [10] apply.

Remark 3.2. Theorem 3.1 motivates the following question: what happens when a given CAR(2) process is sampled? This question is resolved in [3], which found that a sampled CAR(2) process gives rise to an ARMA (2,1) process.

3.2. Weak convergence of h-AR(p) process to CAR(p) process

We now consider the AR(p) process on the discrete time domain $\{0, h, 2h, \ldots\}$, given as

$$X_t = b^h_1X_{t-h} + \cdots + b^h_pX_{t-ph} + \varsigma^hZ_t, \tag{3.6}$$

where $Z_t \sim$ IID $\mathcal{N}(0, 1)$ for all $t = ph, (p+1)h$ and show that subject to suitable conditions on the coefficients $b^h_1, \ldots, b^h_p$ and ς^h, this converges as $h \downarrow 0$ to its continuous time CAR(p) process of the form

$$Y^{(p)}_t = \sum_{i=0}^{p-1} a_{i+1}Y^{(i)}_t + \sigma W_t, \quad t > 0 \tag{3.7}$$

where $a_j \neq 0$ for all $j \in \{1, \ldots, p\}$ and $\sigma^2 > 0$.

Define the coefficients $\{c_j^h : j = 1, \ldots, p\}$ and ζ^h through the equations

$$b_i^h \triangleq (-1)^{i-1} \left\{ \binom{p}{i} + \sum_{k=i}^{p} \binom{k-1}{i-1} h^{p-k+1} c_k^h \right\}, \qquad \varsigma^h \triangleq h^{p-1} \xi^h. \tag{3.8}$$

The following theorem is proven in the Appendix.

Theorem 3.2. *Consider the h-AR (p) process of (3.6) with coefficients given by (3.8), where* $c_j^h \to a_j \neq 0$ *for all* $j \in \{1, \ldots, p\}$ *and* $\xi^h/\sqrt{h} \to \sigma$ *as* $h \downarrow 0$, *we further assume that as* $h \downarrow 0$, $(\Delta_{0;0}^h, \ldots, \Delta_{p-1;0}^h)$ *converges in distribution to* $(\Delta_{0;0}^0, \ldots, \Delta_{p-1;0}^0)$ *with probability measure* ν_p^0 *on* $(\mathbb{R}^p, \mathcal{B}(\mathbb{R}^p))$. *Then the h-AR(p) process of (3.6) converges in distribution to the CAR(p) process of (3.7), where* $\{W_t : t \geq 0\}$ *is a one-dimensional Brownian motion, independent of* $(Y_0, Y_0^{(1)}, \ldots, Y_0^{(p-1)})$, *and* $\Pr\{(Y_0, Y_0^{(1)}, \ldots, Y_0^{(p-1)}) \in \Gamma\} = \nu_p^0(\Gamma)$ *for any* $\Gamma \in \mathcal{B}(\mathbb{R}^p)$.

It is of interest to note the scaling for the Gaussian variable Z_t in (3.6). In order to have the desired convergence, one must have $\zeta^h/(\sigma\sqrt{h}) \to 1$ and via (3.8) this entails $\zeta^h/(\sigma h^{p-1/2}) \to 1$.

Acknowledgments

We greatly thank two anonymous referees for their careful reading of the paper and for their constructive feedback, which greatly helped to improve the quality of the paper. We are also very grateful for both the very helpful input as well as the generosity of time from the Editor-in-Chief, Prof. Christian Genest.

Appendix

This Appendix proves Theorem 3.2. To do so, we first study the similarity of the h-AR(p) process in (3.6) with the CAR(p) process (3.7). We begin by introducing the corresponding h-VAR(p) process

$$\Delta_{0;t}^h - \Delta_{0;t-h}^h = h\Delta_{1;t}^h, \ldots, \Delta_{p-2;t}^h - \Delta_{p-2;t-h}^h = h\Delta_{p-1;t}^h, \tag{A.1}$$

from which one gets, through iteration,

$$\Delta_{p-1;t}^h - \Delta_{p-1;t-h}^h = h \sum_{i=0}^{p-1} c_{i+1}^h \Delta_{i;t-h}^h + \xi^h Z_t,$$

with $Z_t \sim$ IID $\mathcal{N}(0, 1)$ for all $t = h, 2h, \ldots$ and

$$\Pr\{(\Delta_{0;0}^h, \ldots, \Delta_{p-1;0}^h) \in \Gamma\} = \nu_p^h(\Gamma) \quad \text{for all } \Gamma \in \mathcal{B}(\mathbb{R}^p), \tag{A.2}$$

ν_p^h being a probability measure on $(\mathbb{R}^p, \mathcal{B}(\mathbb{R}^p))$ for the Borel σ-algebra $\mathcal{B}(\mathbb{R}^p)$. For $i = 1$, the process of (A.1) immediately yields

$$h\Delta_{1;t-h}^h = \Delta_{0;t-h}^h - \Delta_{0;t-2h}^h = \binom{1}{0}(-1)^0 \Delta_{0;t-h}^h + \binom{1}{1}(-1)^1 \Delta_{0;t-2h}^h.$$

For $i = 2$, the process of (A.1) immediately yields

$$\begin{aligned} h^2 \Delta_{2;t-h}^h &= h(\Delta_{1;t-h}^h - \Delta_{1;t-2h}^h) \\ &= \binom{2}{0}(-1)^0 \Delta_{0;t-h}^h + \binom{2}{1}(-1)^1 \Delta_{0;t-2h}^h + \binom{2}{2}(-1)^2 \Delta_{0;t-3h}^h. \end{aligned}$$

The process of (A.1) generalizes as follows: for all $i \in \{1, \ldots, p-1\}$,

$$h^i \Delta_{i;t-h}^h = \sum_{k=1}^{i+1} \binom{i}{k-1} (-1)^{k-1} \Delta_{0;t-kh}^h. \tag{A.3}$$

We now prove (A.3) via mathematical induction. For $i = 1$ the relationship (A.3) holds trivially. Let (A.3) hold for $i = m$. It is now straightforward to show that (A.3) also holds for $i = m + 1$. Furthermore, for $i = 1$, it follows from (A.1) that

$\Delta^h_{0;t} = \Delta^h_{0;t-h} + h\Delta^h_{1;t}$. For $i = 2$, it follows from (A.1) that

$$\begin{aligned}\Delta^h_{0;t} &= \Delta^h_{0;t-h} + h(\Delta^h_{1;t-h} + h\Delta^h_{2;t}) \\ &= \Delta^h_{0;t-h} + h\Delta^h_{1;t-h} + h^2\Delta^h_{2;t} = \cdots = \sum_{i=0}^{p-2} h^i \Delta^h_{i;t-h} + h^{p-1}\Delta^h_{p-1;t} \\ &= \sum_{i=0}^{p-2} h^i \Delta^h_{i;t-h} + h^{p-1}\left(\Delta^h_{p-1;t-h} + h\sum_{i=0}^{p-1} c^h_{i+1}\Delta^h_{i;t-h} + \xi^h Z_t\right) \\ &= \sum_{k=1}^{p} (-1)^{k-1}\left\{\binom{p}{k} + \sum_{i=k}^{p}\binom{i-1}{k-1} h^{p-i+1} c^h_i\right\}\Delta^h_{0;t-kh} + h^{p-1}\xi^h Z_t,\end{aligned} \tag{A.4}$$

which, when compared with the h-AR(p) process of (3.6), yields the relationships

$$b^h_i \triangleq (-1)^{i-1}\left\{\binom{p}{i} + \sum_{k=i}^{p}\binom{k-1}{i-1} h^{p-k+1} c^h_k\right\}, \qquad \varsigma^h \triangleq h^{p-1}\xi^h \tag{A.5}$$

for all $i \in \{1, \ldots, p\}$. From above, the h-AR(p) process of (3.6) with coefficients given by (3.8) is equivalent to the h-VAR(p) of (A.1).

We shall next find the coefficients $c^h_1, \ldots, c^h_p$ in terms of the coefficients $b^h_1, \ldots, b^h_p$. In particular, we have that, from (3.8), when $i = p$,

$$b^h_p = (-1)^{p-1}\left\{\binom{p}{p} + \binom{p-1}{p-1} h\, c^h_p\right\}, \quad c^h_p = h^{-1}\left\{(-1)^{p-1}\binom{p-1}{p-1} b^h_p - 1\right\}.$$

When $i = p - 1$,

$$\begin{aligned} b^h_{p-1} &= (-1)^{p-2}\left\{\binom{p}{p-1} + \sum_{k=p-1}^{p}\binom{k-1}{p-2} h^{p-k+1} c^h_k\right\}, \\ c^h_{p-1} &= h^{-2}\left[(-1)^{p-2}\left\{\binom{p-2}{p-2} b^h_{p-1} + \binom{p-1}{p-2} b^h_p\right\} - 1\right].\end{aligned}$$

This leads to the following general formula.

Proposition A.1. *For all $i \in \{1, \ldots, p\}$, one has*

$$c^h_i = h^{-p+i-1}\left\{(-1)^{i-1}\sum_{k=i}^{p}\binom{k-1}{i-1} b^h_k - 1\right\}. \tag{A.6}$$

Proof. We prove Proposition A.1 by backward induction. (i) For $i = p$ the relationship (A.6) holds trivially. (ii) Let (A.6) hold for every $i \in \{p-1, p-2, \ldots, m+1\}$. (iii) We shall show that (A.6) also holds for $i = m$. For $i = m$, one has

$$b^h_m = (-1)^{m-1}\left\{\binom{p}{m} + \sum_{i=m}^{p}\binom{i-1}{m-1} h^{p-i+1} c^h_i\right\}$$

and hence it follows from (i) and (ii) that

$$(-1)^{m-1} b^h_m = \binom{p}{m} + h^{p-m+1} c^h_m + (-1)^{m-1} b^h_m \sum_{i=m+1}^{p}\binom{i-1}{m-1}(-1)^{i-1}\sum_{k=i}^{p}\binom{k-1}{i-1} b^h_k - \sum_{i=m+1}^{p}\binom{i-1}{m-1}.$$

Interchanging the order of summation in the double sum, the former index bounds $i \le k \le p$ and $m + 1 \le i \le p$ have now become $m + 1 \le i \le k$ and $m + 1 \le k \le p$. We further have

$$\begin{aligned}\sum_{i=m+1}^{p}\binom{i-1}{m-1}(-1)^{i-1}\sum_{k=i}^{p}\binom{k-1}{i-1} b^h_k &= \sum_{k=m+1}^{p} b^h_k \binom{k-1}{m-1}\sum_{j=1}^{k-m}(-1)^{j+m-1}\binom{k-m}{j} \\ &= (-1)^m \sum_{k=m+1}^{p} b^h_k \binom{k-1}{m-1},\end{aligned}$$

as well as

$$\sum_{i=m+1}^{p}\binom{i-1}{m-1}=\sum_{i=m+1}^{p}\left\{\binom{i}{m}-\binom{i-1}{m}\right\}=\binom{p}{m}-1.$$

The last relationship yields

$$(-1)^{m-1}b_m^h=\binom{p}{m}+h^{p-m+1}c_m^h+(-1)^m\sum_{k=m+1}^{p}b_k^h\binom{k-1}{m-1}-\left\{\binom{p}{m}-1\right\},$$

$$h^{p-m+1}c_m^h=(-1)^{m-1}\sum_{k=m}^{p}b_k^h\binom{k-1}{m-1}-1,$$

concluding part (iii) and thus the proof. □

Finally, for $t>0$ we know that

$$dY_t=Y_t^{(1)}dt,\ dY_t^{(1)}=Y_t^{(2)}dt,\ldots,dY_t^{(p-2)}=Y_t^{(p-1)}dt, \tag{A.7}$$

and from (3.7) we have also that

$$dY_t^{(p-1)}=\{a_1Y_t+a_2Y_t^{(1)}+\cdots+a_pY_t^{(p-1)}\}dt+\sigma dW_t. \tag{A.8}$$

Thus the CAR(p) process of (3.7) is equivalent from (A.7) and (A.8) to the system of stochastic differential equations in (2.1).

We are now ready to prove Theorem 3.2.

Proof. We prove Theorem 3.2 by proving that it suffices to show that the h-VAR(p) process of (A.1) converges to the SDEs system of (2.1). We employ the framework of Theorems 2.1 and 2.2 of [10]. Let M_t be the σ-algebra generated by $\Delta^h_{i;0}, \Delta^h_{i;h}, \Delta^h_{i;2h}, \ldots, \Delta^h_{i;t-h}$, for $i=0,\ldots,p-2$, and $\Delta^h_{p-1;0}, \Delta^h_{p-1;h}, \Delta^h_{p-1;2h}, \ldots, \Delta^h_{p-1;t}$ for $t=h,2h,\ldots$ The h-VAR(p) process of (A.1) is clearly Markovian of order 1, since we may construct $\Delta^h_{0;t}, \Delta^h_{1;t}, \Delta^h_{2;t}, \ldots, \Delta^h_{p-1;t}$ from $\Delta^h_{0;t-h}, \Delta^h_{1;t-h}, \Delta^h_{2;t-h}, \ldots, \Delta^h_{p-1;t-h}$ by constructing first $\Delta^h_{p-1;t}$ from the last equation of (A.1), $\Delta^h_{p-2;t}$ from (A.1) for $i=p-1$, and so forth, and then finally $\Delta^h_{0;t}$ from (A.1) for $i=1$. This also establishes that the set $\{\Delta^h_{i;t}:i\in\{0,\ldots,p-2\}\}$ is M_t-adapted. Thus the corresponding drifts per unit of time conditioned on information at time t are given, for all $i\in\{1,\ldots,p-1\}$, by

$$\mathrm{E}\left\{\frac{\Delta^h_{i-1;t}-\Delta^h_{i-1;t-h}}{h}\middle|M_t\right\}=\mathrm{E}\left\{\frac{\Delta^h_{i-1;t-h}+h\,\Delta^h_{i;t}-\Delta^h_{i-1;t-h}}{h}\middle|M_t\right\}=\Delta^h_{i;t} \tag{A.9}$$

and

$$\mathrm{E}\left\{\frac{\Delta^h_{p-1;t+h}-\Delta^h_{p-1;t}}{h}\middle|M_t\right\}=c_1^h\Delta^h_{0;t}+\cdots+c_p^h\Delta^h_{p-1;t}, \tag{A.10}$$

where the second inequality holds by the last equation of (A.1).

Furthermore, the variances and covariances per unit of time are respectively given, for all $i\in\{1,\ldots,p-1\}$, by

$$\mathrm{E}\left\{\frac{(\Delta^h_{i-1;t}-\Delta^h_{i-1;t-h})^2}{h}\middle|M_t\right\}=h\big(\Delta^h_{i;t}\big)^2, \tag{A.11}$$

and

$$\mathrm{E}\left\{\frac{(\Delta^h_{p-1;t+h}-\Delta^h_{p-1;t})^2}{h}\middle|M_t\right\}=(c_1^h\Delta^h_{0;t}+\cdots+c_p^h\Delta^h_{p-1;t})^2h+\frac{(\xi^h)^2}{h}, \tag{A.12}$$

where the last equality assumes that $Z_{t+h}\sim$ IID $\mathcal{N}(0,1)$. By the same logic, one has, for all $i,j\in\{1,\ldots,p-1\}$ with $i\neq j$,

$$\mathrm{E}\left\{\frac{(\Delta^h_{i-1;t}-\Delta^h_{i-1;t-h})(\Delta^h_{j-1;t}-\Delta^h_{j-1;t-h})}{h}\middle|M_t\right\}=h\Delta^h_{i;t}\,\Delta^h_{j;t}, \tag{A.13}$$

and

$$\mathrm{E}\left\{\frac{(\Delta^h_{i-1;t}-\Delta^h_{i-1;t-h})(\Delta^h_{p-1;t+h}-\Delta^h_{p-1;t})}{h}\middle|M_t\right\}=h\Delta^h_{i;t}(c_1^h\Delta^h_{0;t}+\cdots+c_p^h\Delta^h_{p-1;t}). \tag{A.14}$$

Therefore, the relationships of (A.11)–(A.14) are such that, for all $i, j \in \{1, \ldots, p-1\}$ with $i \neq j$,

$$\mathrm{E}\left\{\frac{(\Delta^h_{i-1;t} - \Delta^h_{i-1;t-h})^2}{h}\bigg|M_t\right\} = \mathrm{o}(1), \tag{A.15}$$

$$\mathrm{E}\left\{\frac{(\Delta^h_{p-1;t+h} - \Delta^h_{p-1;t})^2}{h}\bigg|M_t\right\} = \frac{\left(\xi^h\right)^2}{h} + \mathrm{o}(1), \tag{A.16}$$

$$\mathrm{E}\left\{\frac{(\Delta^h_{i-1;t} - \Delta^h_{i-1;t-h})(\Delta^h_{j-1;t} - \Delta^h_{j-1;t-h})}{h}\bigg|M_t\right\} = \mathrm{o}(1) \tag{A.17}$$

and

$$\mathrm{E}\left\{\frac{(\Delta^h_{i-1;t} - \Delta^h_{i-1;t-h})(\Delta^h_{p-1;t+h} - \Delta^h_{p-1;t})}{h}\bigg|M_t\right\} = \mathrm{o}(1), \tag{A.18}$$

where the $\mathrm{o}(1)$ terms vanish uniformly on compact sets.

We can then define the continuous time version of the h-VAR(p) process of (A.1) by $\Delta^h_{i;t} \triangleq \Delta^h_{i;kh}$ for $kh \le t < (k+1)h$ and all $i \in \{0, \ldots, p-1\}$. Thus, according to Theorem 2.2 in [10], the relationships (A.9)–(A.10) and (A.15)–(A.18) provide the weak limit diffusion. This is precisely the linear SDE system of (2.1) and it has a unique solution with the asserted initial distribution.

References

[1] P.J. Brockwell, R. Davis, Y. Yang, Continuous-time Gaussian autoregression, Statist. Sinica 17 (2007) 63–80.
[2] P.J. Brockwell, V. Ferrazzano, C. Klüppelberg, High-frequency sampling and Kernel estimation for continuous-time moving average processes, J. Time Series Anal. 34 (2013) 385–404.
[3] P.J. Brockwell, A. Lindner, Existence and uniqueness of stationary Lévy-driven CARMA processes, Stochastic Process. Appl. 119 (2009) 2660–2681.
[4] P.J. Brockwell, A. Lindner, Strictly stationary solutions of autoregressive moving average equations, Biometrika 97 (2010) 765–772.
[5] J.-C. Duan, Augmented GARCH (p, q) process and its diffusion limit, J. Econometrics 79 (1997) 97–127.
[6] R. Durrett, Stochastic Calculus: A Practical Introduction, CRC Press, Boca Raton, FL, 1996.
[7] I. Karatzas, S. Shreve, Brownian Motion and Stochastic Calculus, Springer, New York, 2012.
[8] H. Kirshner, M. Unser, J.P. Ward, On the unique identification of continuous-time autoregressive models from sampled data, IEEE Trans. Signal Process. 62 (2014) 1361–1376.
[9] R. Lorenz, Weak approximation of stochastic delay: differential equations with bounded memory by discrete time series (Doctoral Dissertation), Humboldt Universität zu Berlin, Germany, 2006.
[10] D.B. Nelson, ARCH models as diffusion approximations, J. Econometrics 45 (1990) 7–38.
[11] A. Papoulis, Probability, Random Variables and Stochastic Processes, McGraw-Hill, New York, 1984.
[12] D.-T. Pham, Estimation of continuous-time autoregressive model from finely sampled data, IEEE Trans. Signal Process. 48 (2000) 2576–2584.
[13] C.E. Rasmussen, C. Williams, Gaussian Processes for Machine Learning, MIT Press, Cambridge, MA, 2006.
[14] S.I. Resnick, Adventures in Stochastic Processes, Springer, New York, 2013.
[15] L.C.G. Rogers, D. Williams, Diffusions, Markov Processes, and Martingales, Vol. I, Cambridge University Press, Cambridge, 2000.