3-2017

# Stationary Gaussian Markov Processes as Limits of Stationary Autoregressive Time Series

Philip A. Ernst

Lawrence D. Brown
*University of Pennsylvania*

Larry Shepp

Robert L. Wolpert



## Recommended Citation

Ernst, P. A., Brown, L. D., Shepp, L., & Wolpert, R. L. (2017). Stationary Gaussian Markov Processes as Limits of Stationary Autoregressive Time Series. *Journal of Multivariate Analysis, 155* 180-186. http://dx.doi.org/10.1016/j.jmva.2016.12.008

# Stationary Gaussian Markov Processes as Limits of Stationary Autoregressive Time Series

**Abstract**

We consider the class, $\mathbb{C}_p$, of all zero mean stationary Gaussian processes, $\{Y_t : t \in (-\infty, \infty)\}$ with $p$ derivatives, for which the vector valued process $\{(Y_t^{(0)}, ..., Y_t^{(p)}) : t \geq 0\}$ 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.

**Keywords**

continuous autoregressive processes, stationary Gaussian Markov processes, stochastic differential equations

**Disciplines**

Business | Mathematics | Partial Differential Equations | Statistics and Probability

# Stationary Gaussian Markov Processes As Limits of Stationary Autoregressive Time Series

Lawrence D. Brown, Philip A. Ernst, Larry Shepp, and Robert Wolpert

August 27, 2015

**Abstract**

We consider the class, $\boldsymbol{C}_p$, of all zero mean stationary Gaussian processes, $Y_t, t \in (-\infty, \infty)$ with $p$ derivatives, for which the vector valued process

$$\left(Y_t^{(0)}, Y_t^{(1)}, \ldots, Y_t^{(p)}\right), t \geq 0$$

is a $p+1$-vector Markov process, where $\left(Y_t^{(0)} = Y(t)\right)$. We provide a rigorous description and treatment of these stationary Gaussian processes as limits of stationary AR($p$) time series.

MSC 2010 Primary: 60G10, Secondary: 60G15

Keywords: Continuous autoregressive processes, stationary Gaussian Markovian processes, stochastic differential. equations

## 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) has been well studied by many mathematicians, statisticians, and economists. See, for example, the works of [3], [4], [2], and [5].

A special class of autoregressive processes are the discrete-time zero-mean stationary Gaussian Markovian processes on the line $(-\infty, \infty)$. The continuous time analogs of these processes are documented in [9] (ch.10) and [10] (pp. 207-212). 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 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 processes 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.

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. This construction is, to the best of our knowledge, novel.

# 2 Equivalent Descriptions of the Class $C_p$

We give here three distinct descriptions of processes comprising the class $\boldsymbol{C}_p$, which are documented in [10] (p. 212), but in different notation. [10] prove (p.211-212) 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 $Y = \{Y(t)\}$ symbolizes a zero-mean Gaussian process on $t \in [0, \infty)$.

## 2.1 Three Definitions

(I) $Y$ is 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). We denote the derivatives at $t$ by $Y^{(i)}(t), i = 1, ..., p-1$. At any $t_0 \in (0, \infty)$, the conditional evolution of the process given $Y(t), t \in [0, t_0]$ depends only on the values of $Y^{(i)}(t_0), i = 0, ..., p-1$. The above can be formalized as follows: let $\left(Y_t^{(0)}, Y_t^{(1)}, \ldots, Y_t^{(p-1)}\right), t \geq 0$ denote the values of a mean zero Itô vector diffusion process defined by the system of equations

$$\begin{aligned} dY_t^{(i-1)} &= Y_t^{(i)} dt, \quad t > 0, \quad i = 1, 2, ..., p-1 \\ dY_t^{(p-1)} &= \sum_{i=0}^{p-1} a_{i+1} Y_t^{(i)} dt + \sigma dW_t \end{aligned} \tag{2.1}$$

for all $t > 0$, where $\sigma > 0$ and the coefficients $\{a_j\}$ satisfy conditions (2.3) and (2.4), below. Then let $Y(t) = Y_t^{(0)}$.

(II) Each $Y \in \boldsymbol{C}_p$ is given uniquely by a certain polynomial $P(z)$ via the covariance of any such $Y$ as follows:

$$\begin{aligned} \mathbb{E}\left[Y_s Y_t\right] &= R(s,t) \\ &= r(t-s) \\ &= \int_{-\infty}^{\infty} \frac{e^{i(t-s)z}}{|P(z)|^2} dz. \end{aligned} \tag{2.2}$$

$P(z)$ is a complex polynomial of degree $p+1$. It has positive leading coefficients and all complex roots $\zeta_j = \rho_j + i\sigma_j, j = 0, \ldots, p$ with $\sigma_j > 0$, and $\rho_j$ real. We impose the following constraint on the roots of $P(z)$: whenever $\rho_j \neq 0$, then there is another $\zeta_j' = -\zeta_j^*$ which is the negative conjugate of $\zeta_j$. This definition of $P(z)$ ensures that $|P(z)|^2$ is an even function. Finally, it can easily be shown that, for all $t$, $r(t)$ automatically has $2p$ derivatives. The conditions in equations (2.3) and (2.4) below link equations (2.2) and (2.1) and characterize which processes satisfying (2.1) are stationary.

The coefficients $a_i$ are the unique solution of the equations:

$$r^{(p+i+1)}(0^+) = \sum_{j=0}^{p} a_j r^{(i+j)}(0), i = 0, 1, \dots, p-1. \tag{2.3}$$

Note that the left and right derivatives of $r^{(j)}$ are equal except for $j = 2p$.

The diffusion coefficient, $\sigma$, is given by

$$\sigma^2 = \sum_{j=0}^{p} a_j r^{(j+p)}(0)(-1)^{j+1} + (-1)^p r^{(2p+1)}(0^-). \tag{2.4}$$

The stationarity of the process in (2.1) can also be determined via the characterization in Section 2.2.

(III) Equivalently, it is necessary and sufficient that $Y \in \boldsymbol{C}_p$ has the representation via Wiener integrals with a standard Brownian motion, $W$,

$$Y_t = \int_{-\infty}^{\infty} g(t-u)dW(u), \tag{2.5}$$

where $g$ has the $L^2$ Fourier transform

$$\hat{g}(z) = \frac{1}{|P(z)|}. \tag{2.6}$$

By well-known results from both Fourier analysis and stochastic integration, a full treatment of which is given in [6], (2.5) and (2.6) jointly are equivalent to construction (II). Further, by [6], an equivalent construction to that given jointly by equations (2.5) and (2.6) is the following: given a pair of independent standard Brownian motions, $W_1$ and $W_2$, $Y$ has the following spectral representation:

$$Y_t = \int_{-\infty}^{\infty} \cos tz\hat{g}\, dW_1(z) + \int_{-\infty}^{\infty} \sin tz\hat{g}\, dW_2(z). \tag{2.7}$$

With regard to initial conditions for (2.1), we note that the process in (2.2) has a stationary distribution. If we use that distribution as the initial distribution for (2.1), and check that equations (2.3) and (2.4) hold, we arrive at a stationary process.

## 2.2 Characterization of Stationarity via (2.1)

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

$$\lambda^p - a_p\lambda^{p-1} - \ldots - a_2\lambda - a_1 = 0. \tag{2.8}$$

The process is stationary if and only if all the roots of equation (2.8) 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 equation (2.8) 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 $a_i < 0, i = 1, 2$. The covariance functions for $\boldsymbol{C}_2$ can be found in [9] (p. 326). For higher order order processes the conditions for stationarity are not so easily described. Indeed, for $\boldsymbol{C}_3$ it is necessary that $a_i < 0, i = 1, 2, 3$, 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 equation (2.8).

# 3 Weak Convergence of the $h$-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 domain $\{0, h, 2h, \ldots\}$ via

$$\begin{aligned} &X_t = b_1^h X_{t-h} + b_2^h X_{t-2h} + \varsigma^h Z_t, \\ &Z_t \sim \text{IID } N(0,1),\ t = 2h, 3h, \ldots \end{aligned} \tag{3.1}$$

The goal is to establish conditions on the coefficients $b_1^h$, $b_2^h$ and $\varsigma^h$ 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 $\{\Delta_{0;t}^h, \Delta_{1;t}^h\}$ defined via

$$\begin{aligned} \Delta_{0;t}^h - \Delta_{0;t-h}^h &= h\Delta_{1;t}^h, \\ \Delta_{1;t}^h - \Delta_{1;t-h}^h &= \left[c_1^h \Delta_{0;t-h}^h + c_2^h \Delta_{1;t-h}^h\right] h + \xi^h Z_t. \\ Z_t \sim \text{IID } N(0,1),& \quad t = h, 2h, ... \end{aligned} \tag{3.2}$$

From (3.2) we see that

$$\begin{aligned} \Delta_{0;t}^h &= \Delta_{0;t-h}^h + h\Delta_{1;t}^h \\ &= \Delta_{0;t-h}^h + h\Delta_{1;t-h}^h + \left[c_1^h \Delta_{0;t-h}^h + c_2^h \Delta_{1;t-h}^h\right] h^2 + \xi^h h Z_t \\ &= \left[2 + c_1^h h^2 + c_2^h h\right] \Delta_{0;t-h}^h - \left[1 + c_2^h h\right] \Delta_{0;t-2h}^h + \xi^h h Z_t. \end{aligned}$$

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

$$b_1^h \triangleq c_1^h h^2 + c_2^h h + 2, \qquad b_2^h \triangleq -c_2^h h - 1, \qquad \text{and} \qquad \varsigma^h \triangleq \xi^h h, \tag{3.3}$$

or, equivalently,

$$c_1^h \triangleq h^{-2}\left[b_1^h + b_2^h - 1\right], \qquad c_2^h \triangleq h^{-1}\left[-1 - b_2^h\right], \qquad \text{and} \qquad \xi^h \triangleq h^{-1}\varsigma^h. \tag{3.4}$$

From above, the $h$-AR(2) process of (3.1) with coefficients given by (3.3) is equivalent to the $h$-VAR(2) in equations (3.2).

**Theorem 3.1.** *Consider a sequence of $h$-AR(2) processes of (3.1) with coefficients given by (3.3), where $c_j^h \to a_j$, $j = 1, 2$, and $\xi^h/\sqrt{h} \to \sigma$ as $h \downarrow 0$. This sequence converges in distribution to the CAR(2) process of (2.1).*

*Proof.* It suffices to show that the $h$-VAR(2) process of (3.2) converges to the SDE system of

$$\begin{aligned} dY_t &= \dot{Y}_t dt, \quad t > 0 \\ d\dot{Y}_t &= \left[a_1 Y_t + a_2 \dot{Y}_t\right] dt + \sigma dW_t, \quad t > 0. \end{aligned} \tag{3.5}$$

We employ the framework of Theorems 2.1 and 2.2 of [8]. Let $M_t$ be the $\sigma$-algebra generated by

$$\Delta^h_{0;0}, \Delta^h_{0;h}, \Delta^h_{0;2h}, ..., \Delta^h_{0;t-h}$$

and

$$\Delta^h_{1;0}, \Delta^h_{1;h}, \Delta^h_{1;2h}, ..., \Delta^h_{1;t}$$

for $t = h, 2h, \ldots$. The $h$-VAR(2) process of (3.2) is clearly Markovian of order 1, since to construct $\{\Delta^h_{0;t}, \Delta^h_{1;t}\}$ from $\{\Delta^h_{0;t-h}, \Delta^h_{1;t-h}\}$ one needs to use the second equation of (3.2) to construct $\Delta^h_{1;t}$ and then use the first equation to construct $\Delta^h_{0;t}$ as well. This establishes that $\Delta^h_{0;t}$ is $M_t$ adapted. Thus, the corresponding drifts per unit of time conditioned on information at time $t$ are given by:

$$\mathbb{E}\left[\frac{\Delta^h_{0;t} - \Delta^h_{0;t-h}}{h}\middle|M_t\right] = \mathbb{E}\left[\frac{\Delta^h_{0;t-h} + h\,\Delta^h_{1;t} - \Delta^h_{0;t-h}}{h}\middle|M_t\right] = \Delta^h_{1;t} \tag{3.6}$$

and

$$\begin{aligned}\mathbb{E}\left[\frac{\Delta^h_{1;t+h} - \Delta^h_{1;t}}{h}\middle|M_t\right] &= \mathbb{E}\left[\frac{\left(c^h_1\Delta^h_{0;t} + c^h_2\Delta^h_{1;t}\right)h + \xi^h Z_{t+h}}{h}\middle|M_t\right] \\ &= c^h_1\Delta^h_{0;t} + c^h_2\Delta^h_{1;t}.\end{aligned} \tag{3.7}$$

Furthermore, the variances and covariances per unit of time are given by

$$\mathbb{E}\left[\frac{\left(\Delta^h_{0;t} - \Delta^h_{0;t-h}\right)^2}{h}\middle|M_t\right] = \mathbb{E}\left[\frac{\left(h\,\Delta^h_{1;t}\right)^2}{h}\middle|M_t\right] = h\left(\Delta^h_{1;t}\right)^2 \tag{3.8}$$

and

$$\begin{aligned}\mathbb{E}\left[\frac{\left(\Delta^h_{1;t+h} - \Delta^h_{1;t}\right)^2}{h}\middle|M_t\right] &= \mathbb{E}\left[\frac{\left[\left(c^h_1\Delta^h_{0;t} + c^h_2\Delta^h_{1;t}\right)h + \xi^h Z_{t+h}\right]^2}{h}\middle|M_t\right] \\ = \left[c^h_1\Delta^h_{0;t} + c^h_2\Delta^h_{1;t}\right]^2 h &+ 2\left[c^h_1\Delta^h_{0;t} + c^h_2\Delta^h_{1;t}\right]\xi^h\mathbb{E}\left[Z_{t+h}\right] + \frac{\left(\xi^h\right)^2}{h}\mathbb{E}\left[Z^2_{t+h}\right] \\ = \left[c^h_1\Delta^h_{0;t} + c^h_2\Delta^h_{1;t}\right]^2 h &+ \frac{\left(\xi^h\right)^2}{h},\end{aligned} \tag{3.9}$$

where the last equality assumes that $\epsilon_{t+h} \sim$ IID $N(0,1)$. By the same logic,

$$\begin{aligned}
&\mathbb{E}\left[\frac{\left(\Delta^h_{0;t}-\Delta^h_{0;t-h}\right)\left(\Delta^h_{1;t+h}-\Delta^h_{1;t}\right)}{h}\middle| M_t\right] \\
&= \mathbb{E}\left[\frac{h\Delta^h_{1;t}\left[\left(c^h_1\Delta^h_{0;t}+c^h_2\Delta^h_{1;t}\right)h+\xi^h Z_{t+h}\right]}{h}\middle| M_t\right] \qquad (3.10)\\
&= h\Delta^h_{1;t}\left[c^h_1\Delta^h_{0;t}+c^h_2\Delta^h_{1;t}\right].
\end{aligned}$$

The relationships in (3.8) - (3.10) become

$$\mathbb{E}\left[\frac{\left(\Delta^h_{0;t}-\Delta^h_{0;t-h}\right)^2}{h}\middle| M_t\right] = \mathrm{o}(1), \qquad (3.11)$$

$$\mathbb{E}\left[\frac{\left(\Delta^h_{1;t+h}-\Delta^h_{1;t}\right)^2}{h}\middle| M_t\right] = \frac{\left(\xi^h\right)^2}{h} + \mathrm{o}(1), \qquad (3.12)$$

and

$$\mathbb{E}\left[\frac{\left(\Delta^h_{0;t}-\Delta^h_{0;t-h}\right)\left(\Delta^h_{1;t+h}-\Delta^h_{1;t}\right)}{h}\middle| M_t\right] = \mathrm{o}(1). \qquad (3.13)$$

The o(1) terms vanish uniformly on compact sets. We may additionally show by brute force that the limits of

$$\mathbb{E}\left[\frac{\left(\Delta^h_{0;t}-\Delta^h_{0;t-h}\right)^4}{h}\middle| M_t\right]$$

and

$$\mathbb{E}\left[\frac{\left(\Delta^h_{1;t+h}-\Delta^h_{1;t}\right)^4}{h}\middle| M_t\right]$$

exist and converge to zero as $h \downarrow 0$.

We proceed to define the continuous time version of the $h$-VAR(2) process of (3.2) by

$$\Delta_{0;t}^h \triangleq \Delta_{0;kh}^h \qquad \text{and} \qquad \Delta_{1;t}^h \triangleq \Delta_{1;kh}^h$$

for $kh \leq t < (k+1)h$. Then, according to the Theorem 2.2 in [8], the relationships (3.6), (3.7) and (3.11) - (3.13) provide the weak (in distribution) limit diffusion

$$d \begin{bmatrix} Y_t \\ \dot{Y}_t \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ a_1 & a_2 \end{bmatrix} \begin{bmatrix} Y_t \\ \dot{Y}_t \end{bmatrix} dt + \begin{bmatrix} 0 & 0 \\ 0 & \sigma \end{bmatrix} d \begin{bmatrix} 0 \\ W_t \end{bmatrix},$$

where $W_t,\ t \geq 0$, is a Brownian motion. This is the linear SDE system of (3.5) and it has a unique solution. □

*Remark* 3.1. [8] Theorems 2.1 and 2.2 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 [8] considers 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 [8] apply.

## 3.2 Stationarity of AR(2)-VAR(2) process

In this section we present necessary and sufficient conditions for stationarity of the AR(2) process and its equivalent VAR(2) process and connect this to the stationary condition for CAR(2). First, we invoke the following proposition, which is easily proved using well-known results from the time series literature (see [1]).

**Proposition 3.1.** *The AR(2) process*

$$X_t = b_1 X_{t-1} + b_2 X_{t-2} + \varsigma Z_t, \qquad Z_t \sim \mathit{IID}\ N(0,1). \tag{3.14}$$

*is stationary if and only if*

$$-2 < b_1 < 2, \qquad b_2 - b_1 < 1 \quad \text{and} \quad b_1 + b_2 < 1, \tag{3.15}$$

*which require* $(b_1, b_2)$ *to lie in the interior of the triangle with vertices (-2,-1), (0,1) and (2,-1). Its stationary variance is given by*

$$\gamma_0 = \frac{\varsigma^2}{1 - \frac{b_1^2}{1-b_2} - b_2\left(\frac{b_1^2}{1-b_2} + b_2\right)}. \tag{3.16}$$

Because of the relations in (3.3) and (3.4), it follows that we have the following corollary to Proposition 3.1.

**Corollary 3.1.** *The VAR(2) process*

$$\begin{aligned}\Delta_{0;t} - \Delta_{0;t-1} &= \Delta_{1;t},\\ \Delta_{1;t} - \Delta_{1;t-1} &= c_1\Delta_{0;t-1} + c_2\Delta_{1;t-1} + \xi Z_t,\\ Z_t &\sim IID\ N(0,1), \quad t = 1, 2, ...,\end{aligned}$$

*is equivalent to the AR(2) process of (3.14) for*

$$b_1 = c_1 + c_2 + 2 \qquad b_2 = -c_2 - 1 \qquad \textit{and} \qquad \varsigma = \xi.$$

*Furthermore, the VAR(2) process is stationary if and only if*

$$-4 < c_1 < 0 \qquad \textit{and} \qquad -2 - \frac{c_1}{2} < c_2 < 0. \tag{3.17}$$

*(3.17) is equivalent to the condition that* $(c_1, c_2)$ *lies in the interior of the triangle with vertices (-4,0), (0,0) and (0,-2).*

## 3.3 Stationary Variance of $h$-VAR(2) Process

In this section we investigate the stationarity of a special version of the $h$-VAR(2) process in Theorem 5.1.1, which converges to the CAR(2) process of (3.5) as $h \downarrow 0$, through the stationarity of its equivalent $h$-AR(2) process.

**Proposition 3.2.** *The h-VAR(2) process of (3.2) for* $c_1^h = a_1$, $c_2^h = a_2$ *and* $\xi^h = \sigma\sqrt{h}$ *is stationary as* $h \downarrow 0$ *if and only if* $a_j < 0$, $j = 1, 2$. *Then its stationary variance satisfies*

$$\lim_{h\downarrow 0} \gamma_0^h = \frac{\sigma^2}{2a_1a_2}. \tag{3.18}$$

*Proof.* From (3.3), the $h$-VAR(2) process of the hypothesis is equivalent to the $h$-AR(2) process of equation (3.1) with coefficients given by

$$b_1^h = a_1h^2 + a_2h + 2, \qquad b_2^h = -a_2h - 1 \qquad \text{and} \qquad \varsigma^h = \sigma\big(\sqrt{h}\big)^3. \tag{3.19}$$

From condition (3.15) of Proposition 3.1, this $h$-AR(2) process is stationary if and only if the following conditions hold:

- $b_1^h + b_2^h < 1 \;\Leftrightarrow\; a_1h^2 + a_2h + 2 - a_2h - 1 < 1 \;\Leftrightarrow\; \quad a_1 < 0,$

- $b_1^h < 2 \;\Leftrightarrow\; a_1h^2 + a_2h + 2 < 2 \quad \Leftrightarrow\; a_2 < -a_1h \overset{h\downarrow 0}{\Longleftrightarrow} \quad a_2 < 0,$

- $-2 < b_1^h \;\Leftrightarrow\; -2 < a_1h^2 + a_2h + 2 \;\Leftrightarrow\; a_1h^2 + a_2h + 4 > 0 \;\Leftrightarrow\; 0 < h < \frac{-a_2 - \sqrt{a_2^2 - 16a_1}}{2a_1},$

- $b_2^h - b_1^h < 1 \;\Leftrightarrow\; -a_2h - 1 - a_1h^2 - a_2h - 2 < 1 \;\Leftrightarrow\; a_1h^2 + 2a_2h + 4 > 0 \;\Leftrightarrow\; 0 < h < \frac{-a_2 - \sqrt{a_2^2 - 4a_1}}{a_1},$

where the last two conditions hold as $h \downarrow 0$.

Finally, from equation (3.16) and (3.19) we can compute the stationary variance as follows:

$$
\begin{aligned}
\gamma_0^h &= \frac{\left(\varsigma^h\right)^2}{1 - \frac{\left(b_1^h\right)^2}{1-b_2^h} - b_2^h \left[\frac{\left(b_1^h\right)^2}{1-b_2^h} + b_2^h\right]}, \\
&= \frac{\sigma^2(2 + a_2h)}{a_1a_2(4 + a_1h^2 + 2a_2h)} \quad \overset{h\downarrow 0}{\longrightarrow} \quad = \frac{\sigma^2}{2a_1a_2}.
\end{aligned}
$$

This concludes the proof. $\square$

## 3.4 Stationarity of CAR(2) Process

This section provides a new derivation of the necessary and sufficient conditions for the stationarity of a CAR(2) process. It also gives the stationary covariance function of the vector $\left(Y_t^{(0)}, Y_t^{(1)}\right)$.

**Theorem 3.2.** *The CAR(2) process given by the SDEs system*

$$
d\begin{bmatrix} Y_t \\ \dot{Y}_t \end{bmatrix} = \Lambda \begin{bmatrix} Y_t \\ \dot{Y}_t \end{bmatrix} dt + \Sigma\, d\begin{bmatrix} 0 \\ W_t \end{bmatrix}, \qquad t > 0, \tag{3.20}
$$

*where*

$$
\Lambda = \begin{bmatrix} 0 & 1 \\ a_1 & a_2 \end{bmatrix} \qquad \textit{and} \qquad \Sigma = \begin{bmatrix} 0 & 0 \\ 0 & \sigma \end{bmatrix},
$$

*is stationary if and only if $a_j < 0$, $j = 1, 2$. Then its solution $[Y_t, \dot{Y}_t]^\top$, $t \geq 0$, is a zero-mean 2-dimensional Gaussian process with covariance*

$$V \triangleq \int_0^\infty e^{t\Lambda}\Sigma\Sigma^\top e^{t\Lambda^\top}\, dt \tag{3.21}$$

*and covariance function*

$$\begin{aligned} \rho(s,t) &\triangleq E\left( \begin{bmatrix} Y_s \\ \dot{Y}_s \end{bmatrix} \begin{bmatrix} Y_t, \dot{Y}_t \end{bmatrix} \right) \\ &= \begin{cases} e^{(s-t)\Lambda}V, & 0 \leq t \leq s < \infty; \\ Ve^{(t-s)\Lambda^\top}, & 0 \leq s \leq t < \infty. \end{cases} \end{aligned}$$

*Proof.* According to Theorem 6.7 in Chapter 5 of [7], the assertion of the theorem holds if all the eigenvalues of matrix $\Lambda$ have negative real parts. Hence, we compute the eigenvalues of matrix $\Lambda$. We calculate the characteristic polynomial as

$$\phi(\lambda) = |\Lambda - \lambda I| = \begin{vmatrix} -\lambda & 1 \\ a_1 & a_2 - \lambda \end{vmatrix} = \lambda^2 - a_2\lambda - a_1,$$

which is of quadratic order with a discriminant equal to $D = a_2^2 + 4a_1$. We then consider the following cases:

(i) If $D = 0 \Leftrightarrow \frac{a_2^2}{4} = -a_1$, the characteristic polynomial has the double root

$$\lambda_{1,2} = \frac{a_2}{2}.$$

(ii) If $D > 0 \Leftrightarrow \frac{a_2^2}{4} > -a_1$, the characteristic polynomial has the two real roots

$$\lambda_{1,2} = \frac{a_2 \pm \sqrt{a_2^2 + 4a_1}}{2}.$$

(iii) If $D < 0 \Leftrightarrow \frac{a_2^2}{4} < -a_1$, the characteristic polynomial has the two complex roots

$$\lambda_{1,2} = \frac{a_2 \pm i\sqrt{4a_1 + a_2^2}}{2}.$$

In every case we need to impose conditions on the coefficients of the characteristic polynomial so as the real part of all eigenvalues is negative.

Indeed, in case (i) the double real root of the characteristic polynomial is negative if and only if $a_2 < 0$, which through the discriminant condition implies also that $a_1 < 0$. In case (ii) we need to impose that both real eigenvalues are negative; i.e.,

$$\lambda_1 = \frac{a_2 - \sqrt{a_2^2 + 4a_1}}{2} < 0,$$

$$\begin{aligned} \lambda_2 = \frac{a_2 + \sqrt{a_2^2 + 4a_1}}{2} &< 0 \\ \iff \quad \sqrt{a_2^2 + 4a_1} &< -a_2 \\ \overset{a_2<0}{\iff} \quad a_2^2 + 4a_1 &< a_2^2 \\ \iff \quad a_1 &< 0. \end{aligned}$$

Note that the latter condition implies both that $a_j < 0$ for $j = 1, 2$. Then the former condition holds as well. In case (iii) the common real part of the two complex eigenvalues is negative if and only if $a_2 < 0$, which through the discriminant condition also implies that $a_1 < 0$. Consequently, in all cases the real part of both eigenvalues of matrix $\Lambda$ is negative if and only if $a_j < 0$, $j = 1, 2$. $\square$

We now compute the stationary variance $V$ of the CAR(2) process of (3.20), as given in (3.21), beginning with the computation of the matrix $e^{t\Lambda}$, $t \geq 0$. In particular, we are looking for $f(\Lambda)$, where $f(\lambda) = e^{\lambda t}$. From standard matrix theory, this can be computed via a polynomial expression of order 1, and thus $f(\Lambda) = \delta_0 I + \delta_1 \Lambda$. Hence, it suffices to set $g(\lambda) = \delta_0 + \delta_1 \lambda$ and to demand that $f(\lambda)$ and $g(\lambda)$ to be equal on the spectrum of $\Lambda$. Then we will have that $f(\Lambda) = g(\Lambda)$.

The roots (one double or two real/complex) $\lambda_1, \lambda_2$ of the characteristic polynomial $\phi(\lambda)$ of $\Lambda$ satisfy the relationships:

$$\lambda_1 + \lambda_2 = a_2 \qquad \text{and} \qquad \lambda_1\lambda_2 = -a_1. \tag{3.22}$$

Since the polynomials $f(\lambda) = e^{\lambda t}$ and $g(\lambda) = \delta_0 + \delta_1 \lambda$ must be equal on the spectrum of $\Lambda$, we have that

$$f(\lambda_1) = g(\lambda_1) \quad \Leftrightarrow \quad e^{\lambda_1 t} = \delta_0 + \delta_1\lambda_1,$$

$$f(\lambda_2) = g(\lambda_2) \quad \Leftrightarrow \quad e^{\lambda_2 t} = \delta_0 + \delta_1\lambda_2.$$

Then,

$$e^{\Lambda t} = f(\Lambda) = g(\Lambda) = \delta_0\, I + \delta_1\, \Lambda = \begin{bmatrix} \delta_0 & \delta_1 \\ & \\ a_1\delta_1 & \delta_0 + a_2\delta_1 \end{bmatrix}.$$

From (3.21), we compute the stationary variance

$$V = \int_0^\infty e^{t\Lambda}\Sigma\Sigma^\top e^{t\Lambda^\top} dt = \sigma^2 \int_0^\infty e^{t\Lambda} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \left(e^{t\Lambda}\right)^\top dt.$$

For any $a < 0$ we have that

$$\int_0^\infty e^{(a+bi)t} dt = \frac{e^{(a+bi)t}}{(a+bi)}\bigg|_0^\infty = -\frac{1}{a+bi}.$$

Using the equalities in (3.22), we can rewrite the covariance function as:

$$= \begin{bmatrix} \sigma^2 \frac{\lambda_2 e^{\lambda_1 (s-t)} - \lambda_1 e^{\lambda_2 (s-t)}}{2a_1 a_2(\lambda_2-\lambda_1)} & -\sigma^2 \frac{e^{\lambda_2 (s-t)} - e^{\lambda_1 (s-t)}}{2a_2(\lambda_2-\lambda_1)} \\ & \\ \sigma^2 \frac{e^{\lambda_2 (s-t)} - e^{\lambda_1 (s-t)}}{2a_2(\lambda_2-\lambda_1)} & -\sigma^2 \frac{\lambda_2 e^{\lambda_2 (s-t)} - \lambda_1 e^{\lambda_1 (s-t)}}{2a_2(\lambda_2-\lambda_1)} \end{bmatrix}, \quad 0 \le t \le s < \infty.$$

## 3.5 Weak Convergence of $h$-AR($p$) process to CAR($p$) process

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

$$\begin{aligned} &X_t = b_1^h X_{t-h} + b_2^h X_{t-2h} + \cdots + b_i^h X_{t-ih} + \cdots + b_p^h X_{t-ph} + \varsigma^h Z_t, \qquad (3.23) \\ &Z_t \sim \text{IID } N(0,1),\ \ t = ph, (p+1)h, ..., \end{aligned}$$

and show that subject to suitable conditions on the coefficients $b_1^h$, $b_2^h$, ... $b_p^h$ and $\varsigma^h$, this converges as $h \downarrow 0$ to its continuous time CAR($p$) process of the form

$$Y_t^{(p)} = \sum_{i=0}^{p-1} a_{i+1} Y_t^{(i)} + \sigma W_t, \quad t > 0, \qquad (3.24)$$

for $a_j \neq 0$, $j = 1, 2, ..., p$, and $\sigma^2 > 0$.

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

$$\begin{aligned} b_i^h &\triangleq (-1)^{i-1} \left\{ \binom{p}{i} + \sum_{k=i}^{p} \binom{k-1}{i-1} h^{p-k+1} c_k^h \right\}, \qquad \text{and} \\ \varsigma^h &\triangleq h^{p-1} \xi^h. \end{aligned} \tag{3.25}$$

The following theorem, which is relegated to the Appendix, can be proven:

**Theorem 3.3.** *The h-AR(p) process of (3.23) with coefficients given by (3.25), where $c_j^h \to a_j$, $j = 1, 2, ..., p$, and $\xi^h/\sqrt{h} \to \sigma$ as $h \downarrow 0$, converges in distribution to the CAR(p) process of (3.24).*

It is of interest to note the scaling for the Gaussian variable $Z_t$ in (3.23). In order to have the desired convergence, one must have $\zeta^h \to \sigma\sqrt{h}$ and via (3.25 this entails $\zeta^h \to \sigma h^{p-1/2}$.

## APPENDIX

The Appendix proves Theorem 5.3. To do so, we first study the similarity of the $h$-AR($p$) process in (3.23) with the CAR($p$) process (3.24). We begin by introducing the corresponding $h$-VAR($p$) process

$$\begin{aligned} \Delta_{0;t}^h - \Delta_{0;t-h}^h &= h\Delta_{1;t}^h, \\ \Delta_{1;t}^h - \Delta_{1;t-h}^h &= h\Delta_{2;t}^h, \\ &\vdots \\ \Delta_{i-1;t}^h - \Delta_{i-1;t-h}^h &= h\Delta_{i;t}^h, \\ &\vdots \\ \Delta_{p-2;t}^h - \Delta_{p-2;t-h}^h &= h\Delta_{p-1;t}^h, \\ \Delta_{p-1;t}^h - \Delta_{p-1;t-h}^h &= h\sum_{i=0}^{p-1} c_{i+1}^h \Delta_{i;t-h}^h + \xi^h Z_t, \\ Z_t &\sim \text{IID } N(0,1), \quad t = h, 2h, .... \end{aligned} \tag{A.1}$$

The process of (A.1) immediately yields:

$$\overset{i=1}{\Longrightarrow} \quad h\Delta_{1;t-h}^h = \Delta_{0;t-h}^h - \Delta_{0;t-2h}^h$$

$$= \binom{1}{0}(-1)^0\Delta^h_{0;t-h} + \binom{1}{1}(-1)^1\Delta^h_{0;t-2h},$$

$$\begin{aligned}
\overset{i=2}{\Longrightarrow} \quad h^2\Delta^h_{2;t-h} &= h\left[\Delta^h_{1;t-h} - \Delta^h_{1;t-2h}\right] \\
&= \left[\Delta^h_{0;t-h} - \Delta^h_{0;t-2h}\right] - \left[\Delta^h_{0;t-2h} - \Delta^h_{1;t-3h}\right] \\
&= \Delta^h_{0;t-h} - 2\Delta^h_{0;t-2h} + \Delta^h_{0;t-3h} \\
&= \binom{2}{0}(-1)^0\Delta^h_{0;t-h} + \binom{2}{1}(-1)^1\Delta^h_{0;t-2h} + \binom{2}{2}(-1)^2\Delta^h_{0;t-3h}.
\end{aligned}$$

The process of (A.1) generalizes as follows:

$$h^i\Delta^h_{i;t-h} = \sum_{k=1}^{i+1}\binom{i}{k-1}(-1)^{k-1}\Delta^h_{0;t-kh}, \tag{A.2}$$

for $i = 1, 2, ..., p-1$.
We prove (A.2) via mathematical induction. (i) For $i = 1$ the relationship (A.2) holds trivially. (ii) Let (A.2) hold for $i = m$. (iii) We shall show that (A.2) also holds for $i = m+1$.

$$\begin{aligned}
h^{m+1}\Delta^h_{m+1;t-h} &\overset{\text{(A.1) for } i=m+1}{=\!=\!=} h^m\Delta^h_{m;t-h} - h^m\Delta^h_{m;t-2h} \\
&\overset{(ii)}{=\!=} \sum_{k=1}^{m+1}\binom{m}{k-1}(-1)^{k-1}\Delta^h_{0;t-kh} - \\
&\quad \sum_{k=1}^{m+1}\binom{m}{k-1}(-1)^{k-1}\Delta^h_{0;t-(k+1)h} \\
&\overset{\text{2nd sum: } l=k+1}{=\!=\!=} \sum_{k=1}^{m+1}\binom{m}{k-1}(-1)^{k-1}\Delta^h_{0;t-kh} + \\
&\quad \sum_{l=2}^{m+2}\binom{m}{l-2}(-1)^{l-1}\Delta^h_{0;t-lh} \\
&= \sum_{k=1}^{m+2}\binom{m+1}{k-1}(-1)^{k-1}\Delta^h_{0;t-kh}.
\end{aligned}$$

Furthermore, we have that

$$
\begin{aligned}
\Delta^h_{0;t} &\xlongequal{\text{(A.1) for } i=1} \Delta^h_{0;t-h} + h\Delta^h_{1;t} \\
&\xlongequal{\text{(A.1) for } i=2} \Delta^h_{0;t-h} + h\left[\Delta^h_{1;t-h} + h\Delta^h_{2;t}\right] = \Delta^h_{0;t-h} + h\Delta^h_{1;t-h} + h^2\Delta^h_{2;t} \\
&= \cdots = \sum_{i=0}^{p-2} h^i \Delta^h_{i;t-h} + h^{p-1}\Delta^h_{p-1;t} \\
&\xlongequal{\text{by the last equation of (A.1)}} \sum_{i=0}^{p-2} h^i \Delta^h_{i;t-h} + h^{p-1}\Big[\Delta^h_{p-1;t-h} + \\
&\qquad h\sum_{i=0}^{p-1} c^h_{i+1}\Delta^h_{i;t-h} + \xi^h Z_t\Big] \\
&\xlongequal{\text{(A.2)}} \sum_{i=0}^{p-1}\left[1 + h^{p-i}c^h_{i+1}\right]\sum_{k=1}^{i+1}\binom{i}{k-1}(-1)^{k-1}\Delta^h_{0;t-kh} + h^{p-1}\xi^h Z_t \\
&\xlongequal{\text{by interchanging the sums}} \sum_{k=1}^{p}(-1)^{k-1}\left\{\sum_{i=k-1}^{p-1}\binom{i}{k-1}\left[1 + h^{p-i}c^h_{i+1}\right]\right\} \\
&\qquad \Delta^h_{0;t-kh} + h^{p-1}\xi^h Z_t \\
&\xlongequal{\text{telescopic sum}} \sum_{k=1}^{p}(-1)^{k-1}\left\{\binom{p}{k} + \sum_{i=k}^{p}\binom{i-1}{k-1}h^{p-i+1}c^h_i\right\}\Delta^h_{0;t-kh} + \\
&\qquad h^{p-1}\xi^h Z_t,
\end{aligned}
$$

which, when compared with the $h$-AR(p) process of (3.23), yields the relationships

$$
\begin{aligned}
b^h_i &\triangleq (-1)^{i-1}\left\{\binom{p}{i} + \sum_{k=i}^{p}\binom{k-1}{i-1}h^{p-k+1}c^h_k\right\}, \qquad \text{and} \qquad \text{(A.3)} \\
\varsigma^h &\triangleq h^{p-1}\xi^h
\end{aligned}
$$

for all $i = 1, 2..., p$. From above, the $h$-AR(p) process of (3.23) with coefficients given by (3.25) is equivalent to the $h$-VAR(p) of (A.1). We conclude with a statement of Theorem 5.3, the proof and supporting details of which are in the Appendix.

We shall next find the coefficients $c_i^h$, $i = 1, 2, ..., p$, in terms of the coefficients $b_i^h$, $i = 1, 2, ..., p$. In particular, we have that, from (3.25),

$$\begin{aligned} &\overset{i=p}{\Longrightarrow} \quad b_p^h = (-1)^{p-1}\left\{\binom{p}{p} + \binom{p-1}{p-1} h\, c_p^h\right\} \\ &\Rightarrow \quad c_p^h = h^{-1}\left[(-1)^{p-1}\binom{p-1}{p-1} b_p^h - 1\right] \end{aligned}$$

and

$$\begin{aligned} &\overset{i=p-1}{\Longrightarrow} \quad b_{p-1}^h = (-1)^{p-2}\left\{\binom{p}{p-1} + \sum_{k=p-1}^{p}\binom{k-1}{p-2} h^{p-k+1} c_k^h\right\} \\ &\Rightarrow \quad c_{p-1}^h = h^{-2}\left\{(-1)^{p-2}\left[\binom{p-2}{p-2} b_{p-1}^h + \binom{p-1}{p-2} b_p^h\right] - 1\right\}. \end{aligned}$$

**Proposition A.1.** *This gives the general formula*

$$c_i^h = h^{-p+i-1}\left[(-1)^{i-1}\sum_{k=i}^{p}\binom{k-1}{i-1} b_k^h - 1\right], \quad i = 1, 2, ..., p. \tag{A.4}$$

*Proof.* We prove Proposition A.1 by backward induction. (i) For $i = p$ the relationship (A.4) holds trivially. (ii) Let (A.4) hold for every $i = p-1, p-2, ..., m+1$. (iii) We shall show that (A.4) also holds for $i = m$.

$$\overset{i=m}{\Longrightarrow} \quad b_m^h = (-1)^{m-1}\left\{\binom{p}{m} + \sum_{i=m}^{p}\binom{i-1}{m-1} h^{p-i+1} c_i^h\right\}$$

$$\begin{aligned} \overset{(i),\ (ii)}{\Longrightarrow} \quad (-1)^{m-1} b_m^h = &\binom{p}{m} + h^{p-m+1} c_m^h + \\ &\sum_{i=m+1}^{p}\binom{i-1}{m-1}(-1)^{i-1}\sum_{k=i}^{p}\binom{k-1}{i-1} b_k^h - \sum_{i=m+1}^{p}\binom{i-1}{m-1}. \end{aligned}$$

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

$$\sum_{i=m+1}^{p} \binom{i-1}{m-1} (-1)^{i-1} \sum_{k=i}^{p} \binom{k-1}{i-1} b_k^h =$$
$$= \sum_{k=m+1}^{p} b_k^h \sum_{i=m+1}^{k} (-1)^{i-1} \binom{i-1}{m-1} \binom{k-1}{i-1}$$
$$\overset{j=i-m}{=\!=\!=} \sum_{k=m+1}^{p} b_k^h \binom{k-1}{m-1} \sum_{j=1}^{k-m} (-1)^{j+m-1} \binom{k-m}{j}$$
$$= (-1)^{m-1} \sum_{k=m+1}^{p} b_k^h \binom{k-1}{m-1} \left[ \sum_{j=0}^{k-m} (-1)^j \binom{k-m}{j} - 1 \right]$$
$$\overset{\text{Newton}}{=\!=\!=} (-1)^{m-1} \sum_{k=m+1}^{p} b_k^h \binom{k-1}{m-1} \left[ (-1+1)^{k-m} - 1 \right]$$
$$= (-1)^m \sum_{k=m+1}^{p} b_k^h \binom{k-1}{m-1},$$

as well as

$$\sum_{i=m+1}^{p} \binom{i-1}{m-1} = \sum_{i=m+1}^{p} \left[ \binom{i}{m} - \binom{i-1}{m} \right] \overset{\text{telescopic sum}}{=\!=\!=\!=\!=} \binom{p}{m} - 1.$$

The last relationship yields

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

concluding part (iii) and thus the proof. □

Finally, for $t > 0$ we know that:

$$\begin{aligned}
dY_t &= Y_t^{(1)} dt, \\
dY_t^{(1)} &= Y_t^{(2)} dt, \\
&\vdots \\
dY_t^{(i-1)} &= Y_t^{(i)} dt, \\
&\vdots \\
dY_t^{(p-2)} &= Y_t^{(p-1)} dt,
\end{aligned} \tag{A.5}$$

and from (3.24) we have also that

$$\begin{aligned}
dY_t^{(p-1)} = \Big[ & a_1 Y_t + a_2 Y_t^{(1)} + \cdots + a_i Y_t^{(i-1)} + \cdots \\
& + a_{p-1} Y_t^{(p-2)} + a_p Y_t^{(p-1)} \Big] dt + \sigma dW_t.
\end{aligned} \tag{A.6}$$

Thus the CAR($p$) process of (3.24) is equivalent from (A.5) and (A.6) to the system of stochastic differential equations in (2.1).

**Theorem 5.3** The $h$-AR(p) process of (3.23) with coefficients given by (3.25), where $c_j^h \to a_j$, $j = 1, 2, ..., p$, and $\xi^h/\sqrt{h} \to \sigma$ as $h \downarrow 0$, converges in distribution to the CAR($p$) process of (3.24).

**Proof.** We generalize Theorem 3.3 by proving that it suffices to show that the $h$-VAR(p) process of (A.1) converges to the SDEs system of (2.1). As in Theorem 5.1.1, we employ the framework of Theorems 2.1 and 2.2 of [8]. Let $M_t$ be the $\sigma$-algebra generated by $\Delta_{i;0}^h, \Delta_{i;h}^h, \Delta_{i;2h}^h, ..., \Delta_{i;t-h}^h$, $i = 0, 1, ..., p-2$, and $\Delta_{p-1;0}^h, \Delta_{p-1;h}^h, \Delta_{p-1;2h}^h, ..., \Delta_{p-1;t}^h$ for $t = h, 2h, ....$ The $h$-VAR(p) process of (A.1) is clearly Markovian of order 1, since we may construct $\Delta_{0;t}^h, \Delta_{1;t}^h, \Delta_{2;t}^h, ..., \Delta_{p-1;t}^h$ from $\Delta_{0;t-h}^h, \Delta_{1;t-h}^h, \Delta_{2;t-h}^h, ..., \Delta_{p-1;t-h}^h$ by constructing first $\Delta_{p-1;t}^h$ from the last equation of (A.1), $\Delta_{p-2;t}^h$ from (A.1) for $i = p-1$, and so forth, and then finally $\Delta_{0;t}^h$ from (A.1) for $i = 1$. This also establishes that $\Delta_{i;t}^h$, $i = 0, 1, ..., p-2$ is $M_t$ adapted. Thus the corresponding drifts per unit of time conditioned on information at time $t$ are given by:

$$E\left[\left.\frac{\Delta_{i-1;t}^h - \Delta_{i-1;t-h}^h}{h}\right| M_t\right] \xlongequal{(A.1)} E\left[\left.\frac{\Delta_{i-1;t-h}^h + h\,\Delta_{i;t}^h - \Delta_{i-1;t-h}^h}{h}\right| M_t\right]$$

$$= \Delta^h_{i;t}, \qquad i = 1, 2, ..., p-1, \tag{A.7}$$

and

$$\begin{aligned} & E\left[\frac{\Delta^h_{p-1;t+h} - \Delta^h_{p-1;t}}{h}\bigg| M_t\right] \\ \xlongequal{\text{by the last equation of (A.1)}} \; & E\left[\frac{h\sum_{i=0}^{p-1} c^h_{i+1}\Delta^h_{i;t} + \xi^h Z_{t+h}}{h}\bigg| M_t\right] \\ = \; & c^h_1\Delta^h_{0;t} + c^h_2\Delta^h_{1;t} + \cdots + c^h_p\Delta^h_{p-1;t}\ . \end{aligned} \tag{A.8}$$

Furthermore, the variances and covariances per unit of time are given by

$$\begin{aligned} \mathbb{E}\left[\frac{\left(\Delta^h_{i-1;t} - \Delta^h_{i-1;t-h}\right)^2}{h}\bigg| M_t\right] \xlongequal{\text{(A.1)}} \; & \mathbb{E}\left[\frac{\left(h\,\Delta^h_{i;t}\right)^2}{h}\bigg| M_t\right] \\ = \; & h\left(\Delta^h_{i;t}\right)^2, \qquad i = 1, 2, ..., p-1, \end{aligned} \tag{A.9}$$

and

$$\begin{aligned} & \mathbb{E}\left[\frac{\left(\Delta^h_{p-1;t+h} - \Delta^h_{p-1;t}\right)^2}{h}\bigg| M_t\right] \\ \xlongequal{\text{by the last equation of (A.1)}} \; & \mathbb{E}\left[\frac{\left[h\sum_{i=0}^{p-1} c^h_{i+1}\Delta^h_{i;t})h + \xi^h Z_{t+h}\right]^2}{h}\bigg| M_t\right] \end{aligned}$$

$$= \left[c^h_1\Delta^h_{0;t} + c^h_2\Delta^h_{1;t} + \cdots + c^h_p\Delta^h_{p-1;t}\right]^2 h + \frac{\left(\xi^h\right)^2}{h}, \tag{A.10}$$

where the last equality assumes that $Z_{t+h} \sim$ IID $N(0,1)$. By the same logic:

$$\mathbb{E}\left[\frac{\left(\Delta^h_{i-1;t} - \Delta^h_{i-1;t-h}\right)\left(\Delta^h_{j-1;t} - \Delta^h_{j-1;t-h}\right)}{h}\bigg| M_t\right]$$

$$\stackrel{(A.1)}{=\!=} \mathbb{E}\left[\frac{\left(h\,\Delta^h_{i;t}\right)\left(h\,\Delta^h_{j;t}\right)}{h}\middle| M_t\right]$$

$$= h\Delta^h_{i;t}\,\Delta^h_{j;t},\;\; i,j = 1,2,...,p-1,\;\; i \neq j, \tag{A.11}$$

and

$$\mathbb{E}\left[\frac{\left(\Delta^h_{i-1;t} - \Delta^h_{i-1;t-h}\right)\left(\Delta^h_{p-1;t+h} - \Delta^h_{p-1;t}\right)}{h}\middle| M_t\right]$$

$$\stackrel{(A.1)}{=\!=} \mathbb{E}\left[\frac{h\Delta^h_{i;t}\Big[\big(c^h_1\Delta^h_{0;t} + c^h_2\Delta^h_{1;t} + \cdots + c^h_p\Delta^h_{p-1;t}\big)h + \xi^h Z_{t+h}\Big]}{h}\middle| M_t\right]$$

$$= h\Delta^h_{i;t}\Big[c^h_1\Delta^h_{0;t} + c^h_2\Delta^h_{1;t} + \cdots + c^h_p\Delta^h_{p-1;t}\Big], \qquad i = 1,2,...,p-1. \tag{A.12}$$

Therefore, the relationships of (A.9) - (A.12) become

$$\mathbb{E}\left[\frac{\left(\Delta^h_{i-1;t} - \Delta^h_{i-1;t-h}\right)^2}{h}\middle| M_t\right] = \mathrm{o}(1), \qquad i = 1,2,...,p-1, \tag{A.13}$$

$$\mathbb{E}\left[\frac{\left(\Delta^h_{p-1;t+h} - \Delta^h_{p-1;t}\right)^2}{h}\middle| M_t\right] = \frac{\left(\xi^h\right)^2}{h} + \mathrm{o}(1), \tag{A.14}$$

$$\mathbb{E}\left[\frac{\left(\Delta^h_{i-1;t}-\Delta^h_{i-1;t-h}\right)\left(\Delta^h_{j-1;t}-\Delta^h_{j-1;t-h}\right)}{h}\middle| M_t\right] = \mathrm{o}(1) \tag{A.15}$$

$$i,j = 1,2,...,p-1,\; i \neq j$$

and

$$\mathbb{E}\left[\frac{\left(\Delta^h_{i-1;t}-\Delta^h_{i-1;t-h}\right)\left(\Delta^h_{p-1;t+h}-\Delta^h_{p-1;t}\right)}{h}\middle| M_t\right] = \mathrm{o}(1), \tag{A.16}$$

$$i = 1, 2, ..., p - 1,$$

where the o(1) terms vanish uniformly on compact sets.

We may also show by brute force that the limits of

$$\mathbb{E}\left[\frac{\left(\Delta^h_{i-1;t} - \Delta^h_{i-1;t-h}\right)^4}{h}\middle| M_t\right], \qquad i = 1, 2, ..., p - 1$$

and

$$\mathbb{E}\left[\frac{\left(\Delta^h_{p-1;t+h} - \Delta^h_{p-1;t}\right)^4}{h}\middle| M_t\right]$$

exist and converge to zero as $h \downarrow 0$. 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 $kh \leq t < (k+1)h$ and $i = 0, 1, ..., p-1$. Thus, according to Theorem 2.2 in [8], the relationships (A.7), (A.8) and (A.13) - (A.16) provide the weak (in distribution) limit diffusion. This is precisely the linear SDE system of (2.1) and it has a unique solution.

# References


[1] Anderson, T. *The Statistical Analysis of Time Series.* Wiley-Interscience, 1994.

[2] Brockwell, P. J., Ferrazzano, V., and Klüppelberg, C. High-frequency sampling and kernel estimation for continuous-time moving average processes. *J. Time Series Anal. 34*, 3 (2013), 385–404.

[3] Brockwell, P. J., and Lindner, A. Existence and uniqueness of stationary Lévy-driven CARMA processes. *Stochastic Process. Appl. 119*, 8 (2009), 2660–2681.

[4] Brockwell, P. J., and Lindner, A. Strictly stationary solutions of autoregressive moving average equations. *Biometrika 97*, 3 (2010), 765–772.

[5] Brockwell, P. Davis, R., and Yang, Y. Continuous-time Gaussian autoregression. *Statistica Sinica 17*, 1 (2007), 63.

[6] Dym, H., and McKean, H. *Gaussian Processes, Function Theory, an the Inverse Spectral Problem*. Academic Press, 1976.

[7] Karatzas, I., and Shreve, S. *Brownian Motion and Stochastic Calculus*. Springer, 1991.

[8] Nelson, D. ARCH models as diffusion approximations. *Journal of Econometrics 45*, 1 (1990), 7–38.

[9] Papoulis, A. *Probability, Random Variables, Stochastic Processes*. McGraw Hill, 1991.

[10] Rasmussen, C., and Williams, C. *The Statistical Analysis of Time Series*. MIT Press, 2006.