



Modeling and Estimation of Multivariate Discrete and Continuous Time Stationary Processes

Marko Voutilainen*

Department of Mathematics and Systems Analysis, Aalto University School of Science, Espoo, Finland

In this paper, we give an autoregressive model of order 1 type of characterization covering all multivariate strictly stationary processes indexed by the set of integers. Consequently, under square integrability, we derive continuous time algebraic Riccati equations for the parameter matrix of the characterization. This provides us with a natural way to define the corresponding estimator. In addition, we show that the estimator inherits consistency from autocovariances of the stationary process. Furthermore, the limiting distribution is given by a linear function of the limiting distribution of the autocovariances. We also present the corresponding existing results of the continuous time setting paralleling them to the discrete case treated in this paper.

OPEN ACCESS

Edited by:

Tommi Sottinen,
University of Vaasa, Finland

Reviewed by:

Foad Shokrollahi,
University of Vaasa, Finland
Ehsan Azmoodeh,
Ruhr University Bochum, Germany

*Correspondence:

Marko Voutilainen
marko.voutilainen@aalto.fi

Specialty section:

This article was submitted to
Mathematical Finance,
a section of the journal
Frontiers in Applied Mathematics and
Statistics

Received: 04 November 2019

Accepted: 03 August 2020

Published: 17 September 2020

Citation:

Voutilainen M (2020) Modeling and
Estimation of Multivariate Discrete and
Continuous Time Stationary
Processes.
Front. Appl. Math. Stat. 6:43.
doi: 10.3389/fams.2020.00043

Keywords: time-series analysis, stationary processes, characterization, multivariate Ornstein-Uhlenbeck processes, generalized Langevin equation, algebraic Riccati equations, estimation, consistency

1. INTRODUCTION

Stationary stochastic processes provide a significant instrument for modeling numerous temporal phenomena related to different fields of science. In particular, due to the evidence of long dependence structures in the real financial data, stationary processes possessing long-memory have been widely applied in mathematical finance.

When discrete time is considered, stationary data is typically modeled by applying autoregressive-moving-average (ARMA) processes or their extensions. One focal reason for popularity of ARMA processes is that for every stationary process with a vanishing autocovariance $\gamma(\cdot)$ and for every $n \in \mathbb{N}$ there exists an ARMA process X such that $\gamma_X(k) = \gamma(k)$ for $|k| \leq n$. For a comprehensive overview of ARMA processes we mention [1–3]. The immense ARMA family include for example SARIMA (seasonal autoregressive integrated moving-average) models, where a seasonal ARMA process is obtained by differencing the original data. We mention also different GARCH (generalized autoregressive conditional heteroskedasticity) models originating from [4] and [5] that are commonly used in financial modeling taking account of the time-dependent volatility. ARMA processes, their extensions and estimation in these models have been concerned e.g., in [6–16], to name but a few. Moreover, in [17] we showed that all univariate strictly stationary processes indexed by the integers are characterized by the AR(1) (autoregressive model of order 1) equation

$$X_t - \phi X_{t-1} = Z_t, \quad t \in \mathbb{Z}.$$

However, contrary to the classical AR(1), the noise Z belonging to a certain class of stationary processes is not necessarily white. Established on the characterization, we proposed an estimation method for ϕ in the case of a square integrable stationary process. This method has several

advantages over conventional ones such as maximum likelihood and least squares fitting of ARMA models. Furthermore, in [18] we applied our method in estimation of a generalization of the ARCH model involving a covariate process that can be interpreted as the liquidity of an asset. We would like to emphasize that the proposed method is applicable in estimation of essentially any square integrable one-dimensional stationary process. Hence, it also covers stationary solutions, which are often of central interest, of models of the vast ARMA family.

In the case of continuous time, the Ornstein-Uhlenbeck process X given by the Langevin equation

$$dX_t = -\theta X_t dt + dB_t, \quad t \in \mathbb{Z}, \tag{1}$$

where $\theta > 0$ and B is a two-sided Brownian motion, can be seen as the analog of the discrete time AR(1) process. By posing a suitable initial condition, (1) yields a stationary solution. The foregoing can be generalized, for example, by replacing Brownian motion with other stationary increment processes satisfying certain integrability conditions. One popular option as the replacement is fractional Brownian motion recovering the fractional Ornstein-Uhlenbeck process introduced in [19]. This kind of generalized Ornstein-Uhlenbeck processes are applied e.g., in mathematical finance to describe mean-reverting systems under the influence of shocks, and they are a highly active topic of research. Equations of type (1) with varying driving forces, and estimation in such models have been concerned e.g., in [20–32], to mention but a few. Furthermore, in [33] we showed that a generalized version of (1) characterizes all multivariate strictly stationary processes with continuous paths. Consequently, we proposed an estimation method for the parameter matrix of (1). The method is based on continuous time algebraic matrix Riccati equations (CAREs) and it is applicable in estimation of essentially any square integrable multivariate stationary process.

Algebraic Riccati equations, occurring naturally e.g., in optimal control and filtering theory, is an intensively studied topic in the literature on its own. In many applications, real-valued CAREs often take the symmetric form

$$B^T A + AB - ACA + D = 0, \tag{2}$$

where C and D are symmetric, and symmetric solutions A are to be found. For a general approach to algebraic Riccati equations the reader may consult for example [34]. The existence and uniqueness of a solution to (2) is a well-studied topic, especially when C and D are positive semidefinite [see e.g., [35–37]].

The main contribution of this paper is to extend the characterization and the consequent estimation method of discrete stationary processes of [17] to multivariate settings. Serving as an instrument for the characterization, we also define a multivariate discrete Lamperti transform giving a one-to-one correspondence between stationary and H -self-similar processes with a positive definite matrix H . The characterization leads to covariance based symmetric CAREs for the related parameter matrix providing us with a novel estimation method of multivariate discrete stationary processes. In outline, this results in the following correspondence between the continuous

time case. When the concepts of noises are extended from the conventional ones, the reputed analogous AR(1) and Langevin equations characterize discrete and continuous time stationary processes, respectively. The characterizations provide us with models of stationary processes yielding symmetric CAREs for the corresponding parameter matrices. Furthermore, these equations can be applied in a similar manner in estimation in both cases. For readers' interest and to highlight the obtained connection between discrete and continuous time, we also give the key results of [33].

The rest of the paper is organized as follows. In section 2.1, we first give an AR(1) type of characterization covering all multivariate strictly stationary processes indexed by the integers. Consequently, under the assumption of square integrability, we obtain a set of symmetric CAREs for the model parameter matrix serving as a basis for estimation. Finally, we state theorems for consistency and asymptotic distribution of the parameter (matrix) estimator. In section 2.2, we present the main results of [33], while at the same time comparing them to the results obtained in discrete time. For the reader's convenience, all technical proofs are postponed to section 3.

2. MAIN RESULTS

We begin with some preliminaries and a short notational introduction. The processes we consider in this paper are n -dimensional, real-valued and indexed by $I \in \{\mathbb{Z}, \mathbb{R}\}$. For such a process Y we write $Y = (Y_t)_{t \in I}$, where the i th component of the random vector Y_t is denoted by $Y_t^{(i)}$. Equality of the distributions of two random vectors Y_t and Z_t is denoted by $Y_t \stackrel{\text{law}}{=} Z_t$. Similarly, equality of two processes Y and Z in the sense of finite dimensional distributions is denoted by $Y = (Y_t)_{t \in I} \stackrel{\text{law}}{=} (Z_t)_{t \in I} = Z$. Throughout the paper, we investigate strictly stationary processes meaning that $(X_{t+s})_{t \in I} \stackrel{\text{law}}{=} (X_t)_{t \in I}$ for every $s \in I$. Consequently, we omit the word “strictly” and simply say that X is stationary. By writing $A \geq 0$ or $A > 0$ we mean that the matrix A is positive semidefinite or positive definite, respectively. We denote an eigendecomposition of a symmetric matrix by $A = Q\Lambda Q^T$, where $\Lambda = \text{diag}(\lambda_i)$. Furthermore, the L^2 vector norm and the corresponding induced matrix norm is denoted by $\|\cdot\|$.

In sections 2.1 and 2.2, we introduce models for discrete and continuous time stationary processes, respectively. Consequently, this leads to symmetric CAREs for the model parameter matrix of the form

$$B^T A + AB - ACA + D = 0, \tag{3}$$

where $C, D \geq 0$, and we are solving the equation for a positive definite A . There exists a vast amount of literature on existence and uniqueness of a solution [see e.g., [35] or [36]] in the described setting. In particular, if $C, D > 0$, then there exists a unique positive semidefinite solution to (3). Furthermore, there exists several numerical methods for finding the positive semidefinite solution of (3) [see e.g., [38, 39] or monograph [40]].

Hence, provided that the solution to (3) is unique, a prospective estimation method of A can be based on the equation.

2.1. Discrete Time

In this subsection, we extend the characterization of discrete ($I = \mathbb{Z}$) stationary processes of [17] to multivariate settings. Consequently, we derive quadratic equations for the corresponding model parameter matrix providing us with a natural way to define an estimator for the parameter. Finally, we pose theorems for consistency and asymptotic distribution. A strong analog with the continuous time case $I = \mathbb{R}$ covered in [33] is obtained. We start by providing some definitions.

DEFINITION 2.1. Let $G = (G_t)_{t \in \mathbb{Z}}$ be an n -dimensional stationary increment process. We define a stationary process $\Delta G = (\Delta_t G)_{t \in \mathbb{Z}}$ by

$$\Delta_t G = G_t - G_{t-1}.$$

Next, we define a class of stationary increment processes having sub-exponentially deviating sample paths. These processes serve as the noise in the subsequent AR(1) type of characterization of stationary processes.

DEFINITION 2.2. Let $H > 0$ be a positive definite $n \times n$ matrix, and let $G = (G_t)_{t \in \mathbb{Z}}$ be an n -dimensional stochastic process with stationary increments and $G_0 = 0$. If

$$\lim_{l \rightarrow -\infty} \sum_{k=l}^0 e^{kH} \Delta_k G \tag{4}$$

exists in probability and defines an almost surely finite random variable, we denote $G \in \mathcal{G}_H$.

REMARK 2.3. Lemma 3.1 shows that existence of a logarithmic moment is sufficient for $G \in \mathcal{G}_H$ for all $H > 0$. Particularly, this is the case if G is square integrable. On the other hand, an example of an one-dimensional stationary increment process G with $G_0 = 0$, but $G \notin \mathcal{G}_H$ for any $H > 0$ was provided in [41]. Moreover, Definition 2.2 could also be stated without the assumption of positive definiteness. However, in the one-dimensional case with $H \leq 0$, the convergence of (4) would imply $G \equiv 0$. As expected, we encounter a similar kind of dimensional degeneracy when considering e.g., symmetric matrices with non-positive eigenvalues. Hence, the assumption $H > 0$ may be regarded as natural. See also Remark 3.3.

The next theorem characterizes all multivariate stationary processes, including processes possessing long-memory.

THEOREM 2.4. Let $H > 0$ be a positive definite $n \times n$ matrix, and let $X = (X_t)_{t \in \mathbb{Z}}$ be an n -dimensional stochastic process. Then X is stationary if and only if $\lim_{t \rightarrow -\infty} e^{tH} X_t \stackrel{\mathbb{P}}{=} 0$ and

$$\Delta_t X = (e^{-H} - I)X_{t-1} + \Delta_t G \tag{5}$$

for $G \in \mathcal{G}_H$ and $t \in \mathbb{Z}$. Moreover, the process $G \in \mathcal{G}_H$ is unique.

COROLLARY 2.5. Let $H > 0$ be a positive definite $n \times n$ matrix, and let X be stationary. Then X admits an AR(1) type of representation

$$X_t - \Phi X_{t-1} = \Delta_t G, \tag{6}$$

where $\Phi = e^{-H}$ and $G \in \mathcal{G}_H$.

By using (6) and the expression (14) from the proof of Theorem 2.4, it is straightforward to show that ΔG is centered and square integrable if and only if X is centered and square integrable, respectively. In what follows, we assume these two attributes and write $\gamma(t) = \mathbb{E}(X_t X_0^\top)$ and $r(t) = \mathbb{E}[(\Delta_t G)(\Delta_0 G)^\top]$. Furthermore, since

$$G_t = \sum_{k=1}^t \Delta_k G, \quad t \geq 1,$$

in this case also G is centered and square integrable, and we denote $v(t) = \text{cov}(G_t) = \mathbb{E}(G_t G_t^\top)$. We would like to point out that centeredness can be assumed without loss of generality (see Remark 2.11).

Under the discussed assumptions, we obtain an expression for $\gamma(t)$ in terms of the noise process.

REMARK 2.6. The autocovariance function $\gamma(t)$ is given by

$$\gamma(t) = e^{-tH} \sum_{k=-\infty}^t \sum_{j=-\infty}^0 e^{kH} r(k-j) e^{jH}.$$

Furthermore, if G has independent components, we obtain

$$\gamma(t) = \frac{e^{-tH}}{2} \sum_{k=-\infty}^t \sum_{j=-\infty}^0 e^{kH} (v(k-j+1) + v(k-j-1) - 2v(k-j)) e^{jH}.$$

The following lemma writes the quadratic equations for the model parameter $\Phi = e^{-H}$ presented in [17] in our multivariate setting.

LEMMA 2.7. Let $H > 0$ be a positive definite $n \times n$ matrix, and let X be stationary of the form (6). Then

$$r(t) = \Phi \gamma(t) \Phi - \gamma(t+1) \Phi - \Phi \gamma(t-1) + \gamma(t) \tag{7}$$

for every $t \in \mathbb{Z}$.

REMARK 2.8. In the proof of the lemma, we utilize the increment process ΔG similarly as in [17] yielding quadratic equations for Φ in terms of r and γ . One could consider an estimation method for Φ based on the above equations. However, for a general stationary X , (7) is a symmetric CARE only if $t = 0$, and even in this case, existence of a unique positive semidefinite solution is not guaranteed.

By adapting the approach of [33], we obtain a set of symmetric CAREs on which we construct an estimator for the model parameter $\Phi = e^{-H}$. For this, we define the following matrix coefficients.

DEFINITION 2.9. We set

$$B_t = \sum_{k=1}^t \gamma(k-1) - \gamma(k)^\top$$

$$C_t = \sum_{k=1}^t \sum_{j=1}^t \gamma(k-j)$$

$$D_t = v(t) - 2\gamma(0) + \gamma(t) + \gamma(t)^\top$$

for every $t \in \mathbb{N}$.

THEOREM 2.10. Let $H > 0$ be a positive definite $n \times n$ matrix, and set $\Theta = I - e^{-H}$. Let $X = (X_t)_{t \in \mathbb{Z}}$ be stationary of the form (5). Then the CARE

$$B_t^\top \Theta + \Theta B_t - \Theta C_t \Theta + D_t = 0 \tag{8}$$

is satisfied for every $t \in \mathbb{N}$.

REMARK 2.11. Equations (7) and (8) are covariance based. Consequently, they hold also when X and G in Theorem 2.4 are not centered.

REMARK 2.12. It is worth to emphasize the interplay between the parameter matrix H and the noise G in the context of a fixed stationary X . The noise in Theorem 2.4 and Corollary 2.5 is unique only after $H > 0$ is fixed. On the other hand, if we fix some covariances of the noise process such that the equations of Theorem 2.10 (or Lemma 2.7) yield a positive definite solution, then we may set H according to this solution. Furthermore, the corresponding noise is now given by Corollary 2.5 and its covariance satisfies the set prerequisites. In practice, the parameter matrix H is estimated by assuming some information on the noise, and by estimating the autocovariance of the observed stationary process X . The estimation will be discussed in more detail at the end of this subsection. See also the following examples illustrating how to obtain familiar noise terms in the case of two stationary ARMA type of processes.

We give a couple of examples on how some basic multivariate processes of ARMA type can be presented in the form (6), and how to derive the corresponding noise G together with its covariance function v .

EXAMPLE 2.13. Let X be an n -dimensional stationary AR(1) type of process given by

$$X_t - \phi X_{t-1} = \epsilon_t,$$

with $0 < \phi = Q\Lambda Q^\top$, $\|\phi\| < 1$ and $\epsilon \sim iid(0, \Sigma)$. Then, we may set $H = -Q \text{diag}(\log \lambda_i) Q^\top$ giving $\Phi = \phi$. Now $\Delta G = \epsilon$ and $G_t = \sum_{k=1}^t \epsilon_k$. Furthermore, $v(t) = \sum_{k=1}^t \text{cov}(\epsilon_k) = t\Sigma$ for $t \geq 1$.

EXAMPLE 2.14. Let X be an n -dimensional stationary ARMA(1, q) type of process given by

$$X_t - \phi X_{t-1} = \epsilon_t + \theta_1 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-q},$$

with $0 < \phi$, $\|\phi\| < 1$ and $\epsilon \sim iid(0, \Sigma)$. Similarly as above, we may set $\Phi = \phi$ and now ΔG equals to the MA(q) process on the right. Consequently, for $t \geq 1$,

$$G_t = \sum_{k=1}^t \epsilon_k + \theta_1 \epsilon_{k-1} + \dots + \theta_q \epsilon_{k-q}$$

and

$$v(t) = \sum_{i,j=0}^q \max(0, t - |i - j|) \theta_i \Sigma \theta_j^\top,$$

where $\theta_0 = I$.

In [17], we proposed an estimation method of one-dimensional stationary processes based on equations (7). In particular, we showed that the method is applicable except in some special class of stationary processes. In [42], we provided a comprehensive analysis of the class, and proved that it consists of highly degenerate processes. On the other hand, due to the strong dependence structure, the failure of different estimation methods is expected. Fundamentally, a stationary process X belongs to the class if there exists two values H and \tilde{H} such that the corresponding processes ΔG and $\tilde{\Delta G}$ in (6) have identical autocovariance functions. Next, we state a lemma showing that these degenerate processes have a special characteristic also under the new set of Equations (8).

LEMMA 2.15. Let X be a one-dimensional stationary process and let $H > 0$ be fixed. Set $\Phi = e^{-H}$ and

$$X_t - \Phi X_{t-1} = \Delta_t G, \quad t \in \mathbb{Z}.$$

If the equation

$$\gamma(t)\Phi^2 - (\gamma(t+1) - \gamma(t-1))\Phi + \gamma(t) - r(t) = 0 \tag{9}$$

yields the same two solutions $\Phi, \tilde{\Phi} > 0$ for every $t \in \mathbb{Z}$, then also the equation

$$C_t \Theta^2 - 2B_t \Theta - D_t = 0 \tag{10}$$

yields the same two solutions $\Theta = 1 - \Phi$ and $\tilde{\Theta} = 1 - \tilde{\Phi}$ for every $t \in \mathbb{N}$.

The following remark illuminates the connection between the coefficient matrices of Definition 2.9 and uniqueness of the solution to (8), which we discussed in the beginning of section 2.

REMARK 2.16. By Lemma 3.8, the matrix Θ is positive definite. Since

$$C_t = \mathbb{E} \left[\sum_{k=1}^t X_{k-1} \left(\sum_{k=1}^t X_{k-1} \right)^\top \right] = \text{cov} \left(\sum_{k=1}^t X_{k-1} \right),$$

the matrix C_t is positive semidefinite. Furthermore, if the smallest eigenvalue of $v(t)$ grows enough in time, D_t becomes positive definite [see [33]]. This is the case e.g., when the noise G has independent components with growing variances.

For estimation, it is desirable that (8) admits a unique positive semidefinite solution. In this case, the solution is the correct parameter matrix Θ by the construction. Moreover, it guarantees that a convergent numerical scheme has the desired limit Θ . In the sequel, we simply assume that t is chosen in such a way that $C_t, D_t > 0$ ensuring the uniqueness of the solution, and we omit the subindex t from Equation (8). We have justified the assumption of positive definiteness in detail in continuous time [see section 2.1 and Remark 2.11 in [33]]. Furthermore, we assume that $v(t)$ is known and the stationary process X is observed up to the time $T > t$, and the coefficient matrices B, C , and D are estimated from these observations by replacing the autocovariances $\gamma(\cdot)$ with some estimators $\hat{\gamma}_T(\cdot)$. The coefficient estimators are denoted by \hat{B}_T, \hat{C}_T , and \hat{D}_T , and we set

$$\Delta_T B = \hat{B}_T - B, \quad \Delta_T C = \hat{C}_T - C, \quad \Delta_T D = \hat{D}_T - D.$$

Next, we define an estimator $\hat{\Theta}_T$ for the matrix $\Theta = I - \Phi = I - e^{-H}$. The proofs of the related asymptotic results allow a certain amount of flexibility in the definition. Thus, we give a definition that probably is the most convenient from the practical point of view. Consistency and the rate of convergence of $\hat{\Theta}_T$ are inherited from autocovariance estimators $\hat{\gamma}_T(\cdot)$ of the observed stationary process. In addition, the limiting distribution is obtained as a linear function of the limiting distribution of the autocovariance estimators.

DEFINITION 2.17. *The estimator $\hat{\Theta}_T$ is defined as the unique positive semidefinite solution to the perturbed CARE*

$$\hat{B}_T^\top \hat{\Theta}_T + \hat{\Theta}_T \hat{B}_T - \hat{\Theta}_T \hat{C}_T \hat{\Theta}_T + \hat{D}_T = 0$$

whenever $\hat{C}_T, \hat{D}_T > 0$. Otherwise, we set $\hat{\Theta}_T = 0$.

THEOREM 2.18. *Let $C, D > 0$. Assume that*

$$\max_{s \in \{0, 1, \dots, t\}} \|\hat{\gamma}_T(s) - \gamma(s)\| \xrightarrow{\mathbb{P}} 0.$$

Then

$$\|\hat{\Theta}_T - \Theta\| \xrightarrow{\mathbb{P}} 0,$$

where $\hat{\Theta}_T$ is given by Definition 2.17.

THEOREM 2.19. *Let $l(T)$ be a rate function. If*

$$l(T) \begin{bmatrix} \text{vec}(\hat{\gamma}_T(0) - \gamma(0)) \\ \text{vec}(\hat{\gamma}_T(1) - \gamma(1)) \\ \vdots \\ \text{vec}(\hat{\gamma}_T(t) - \gamma(t)) \end{bmatrix} \xrightarrow{\text{law}} Z,$$

where Z is a $(t + 1)n^2$ -dimensional random vector, then:

(1) Let \tilde{Z} be the permutation of elements of Z corresponding to the order of elements of

$$\begin{bmatrix} \text{vec}((\hat{\gamma}_T(0) - \gamma(0))^\top) \\ \text{vec}((\hat{\gamma}_T(1) - \gamma(1))^\top) \\ \vdots \\ \text{vec}((\hat{\gamma}_T(t) - \gamma(t))^\top) \end{bmatrix}.$$

Define a linear mapping $L_1 : \mathbb{R}^{(t+1)n^2} \rightarrow \mathbb{R}^{3n^2}$ by

$$L_1(Z) = \begin{bmatrix} \sum_{k=0}^{t-1} (t-k) \begin{bmatrix} Z^{(kn^2+1)} \\ \vdots \\ Z^{((k+1)n^2)} \end{bmatrix} + \sum_{k=1}^{t-1} (t-k) \begin{bmatrix} \tilde{Z}^{(kn^2+1)} \\ \vdots \\ \tilde{Z}^{((k+1)n^2)} \end{bmatrix} \\ \sum_{k=1}^t \begin{bmatrix} Z^{((k-1)n^2+1)} - \tilde{Z}^{(kn^2+1)} \\ \vdots \\ Z^{(kn^2)} - \tilde{Z}^{((k+1)n^2)} \end{bmatrix} \\ \begin{bmatrix} Z^{(tn^2+1)} + \tilde{Z}^{(tn^2+1)} \\ \vdots \\ Z^{((t+1)n^2)} + \tilde{Z}^{((t+1)n^2)} \end{bmatrix} - 2 \begin{bmatrix} Z^{(1)} \\ \vdots \\ Z^{(n^2)} \end{bmatrix} \end{bmatrix},$$

where \sum_1^0 is an empty sum. Then

$$l(T) \text{vec}(\Delta_T C, \Delta_T B, \Delta_T D) \xrightarrow{\text{law}} L_1(Z).$$

(2) If $D, C > 0$ and $\hat{\Theta}_T$ is given by Definition 2.17, then

$$l(T) \text{vec}(\hat{\Theta}_T - \Theta) \xrightarrow{\text{law}} L_2(L_1(Z)),$$

where $L_2 : \mathbb{R}^{3n^2} \rightarrow \mathbb{R}^{n^2}$ is a linear mapping expressible in terms of Θ, t and r .

2.2. Continuous Time

We have collected the main results (Theorems 2.21, 2.28, 2.32 and 2.33) of [33] concerning continuous time multivariate stationary processes into this subsection. In addition, in order to complete the analog between discrete and continuous time, we derive quadratic equations for the model parameter (Proposition 2.24) that correspond to the equations of Lemma 2.7. Throughout the subsection, we assume that the considered processes have continuous paths almost surely and hence, the related stochastic integrals can be interpreted as pathwise Riemann-Stieltjes integrals. Again, we start by defining the class \mathcal{G}_H of stationary increment processes for $H > 0$.

DEFINITION 2.20. *Let $H > 0$ be a positive definite $n \times n$ matrix, and let $G = (G_t)_{t \in \mathbb{R}}$ be an n -dimensional stochastic process with stationary increments and $G_0 = 0$. If*

$$\lim_{s \rightarrow -\infty} \int_s^0 e^{Hu} dG_u$$

exists in probability and defines an almost surely finite random variable, we denote $G \in \mathcal{G}_H$.

As in discrete time, it can be shown that existence of some logarithmic moments ensure that $G \in \mathcal{G}_H$ for all $H > 0$. In particular, square integrability of G suffices, which is the case in our second moment based estimation method.

The next theorem is the continuous time counterpart of Theorem 2.4 showing that all stationary processes are characterized by the Langevin equation, whereas in discrete time, the characterization was given by an AR(1) type of equation.

THEOREM 2.21. *Let $H > 0$ be a positive definite $n \times n$ matrix, and let $X = (X_t)_{t \in \mathbb{R}}$ be an n -dimensional stochastic process. Then X is stationary if and only if*

$$X_0 = \int_{-\infty}^0 e^{Hu} dG_u$$

and

$$dX_t = -HX_t dt + dG_t, \tag{11}$$

for $G \in \mathcal{G}_H$ and $t \in \mathbb{R}$. Moreover, the process $G \in \mathcal{G}_H$ is unique.

COROLLARY 2.22. *From Theorem 2.21 it follows that X is the unique stationary solution*

$$X_t = e^{-Ht} \int_{-\infty}^t e^{Hu} dG_u \tag{12}$$

to (11).

In order to apply Theorem 2.21 in estimation, we pose the assumption

$$\sup_{s \in [0,1]} \mathbb{E} (\|G_s\|^2) < \infty.$$

This guarantees that $G \in \mathcal{G}_H$ for all $H > 0$, and square integrability of X and G . On the other hand, if X is square integrable, then G is also. In addition and without loss of generality, we assume that the processes are centered. Again, we write $\gamma(t) = \mathbb{E}(X_t X_0^\top)$ and $\nu(t) = \mathbb{E}(G_t G_t^\top)$. Now, the autocovariance function of the following stationary process is well-defined.

DEFINITION 2.23. *Let $G = (G_t)_{t \in \mathbb{R}}$ be a centered square integrable stationary increment process and let $\delta > 0$. We define a stationary process $\Delta^\delta G = (\Delta_t^\delta G)_{t \in \mathbb{R}}$ by*

$$\Delta_t^\delta G = G_t - G_{t-\delta}$$

and the corresponding autocovariance function r_δ by

$$r_\delta(t) = \mathbb{E} \left[(\Delta_t^\delta G) (\Delta_0^\delta G)^\top \right].$$

As in discrete time (Lemma 2.7), we obtain quadratic equations for the model parameter H in terms of r_δ and γ . The equations could potentially be used to construct an estimator for H and they might be also of independent interest.

PROPOSITION 2.24. *Let $H > 0$ be a positive definite $n \times n$ matrix, and let X be of the form (12). Then*

$$\begin{aligned} r_\delta(t) = & 2\gamma(t) - \gamma(t+\delta) - \gamma(t-\delta) + \left(\int_t^{t+\delta} \gamma(s) ds - \int_{t-\delta}^t \gamma(s) ds \right) H \\ & + H \left(\int_{t-\delta}^t \gamma(s) ds - \int_t^{t+\delta} \gamma(s) ds \right) + H \left(\int_{t-\delta}^t (s-t+\delta)\gamma(s) ds \right. \\ & \left. + \int_t^{t+\delta} (t-s+\delta)\gamma(s) ds \right) H \end{aligned}$$

for every $t \in \mathbb{R}$.

REMARK 2.25. *The advantage of the equations above is that we have to consider $\gamma(s)$ only for $s \in [t-\delta, t+\delta]$, but as in the discrete case, for a general stationary X we obtain a symmetric CARE only when $t = 0$. In addition, similarly as above, we could set in discrete time $\Delta_t^k G := G_t - G_{t-k}$, $k \in \mathbb{N}$. However, this would lead to more complicated equations in Lemma 2.7.*

A significant difference compared to the discrete time Equations (7) occurs in the univariate case. Namely, the first order term with respect to H vanishes.

COROLLARY 2.26. *The univariate case yields*

$$\begin{aligned} r_\delta(t) = & 2\gamma(t) - \gamma(t+\delta) - \gamma(t-\delta) + H^2 \left(\int_{t-\delta}^t (s-t+\delta)\gamma(s) ds \right. \\ & \left. + \int_t^{t+\delta} (t-s+\delta)\gamma(s) ds \right) \end{aligned}$$

for every $t \in \mathbb{R}$.

One could base a univariate estimation method on the above equations without the concern of existence of a unique positive solution. However, since we wish to treat also multivariate settings, we present the most central results of [33] that are obtained from Theorem 2.21 by considering the noise G directly. First, we define matrix coefficients corresponding to Definition 2.9. Consequently, we write symmetric CAREs for the parameter H that are similar to the CAREs (8) for the discrete time parameter Θ .

DEFINITION 2.27. *We set*

$$\begin{aligned} B_t &= \int_0^t \gamma(s) - \gamma(s)^\top ds \\ C_t &= \int_0^t \int_0^t \gamma(s-u) du ds \\ D_t &= \nu(t) - 2\gamma(0) + \gamma(t) + \gamma(t)^\top \end{aligned}$$

for every $t \geq 0$.

THEOREM 2.28. *Let $H > 0$ be a positive definite $n \times n$ matrix, and let $X = (X_t)_{t \in \mathbb{R}}$ be stationary of the form (12). Then the CARE*

$$B_t^\top H + HB_t - HC_t H + D_t = 0 \tag{13}$$

is satisfied for every $t \geq 0$.

REMARK 2.29. As in discrete time, the Equations (13) and (14) are covariance based and hence, they hold also for a non-centered stationary X .

REMARK 2.30. Contrary to the discrete time Equations (8), the first order term vanishes in the univariate setting as in (13).

Again, we assume that t is chosen in such a way that $C_t, D_t > 0$ ensuring the existence of a unique positive semidefinite solution. We have discussed this assumption in detail in [33]. We define an estimator \hat{H}_T for the model parameter matrix H identically to the discrete time by replacing the autocovariances $\gamma(\cdot)$ in the matrix coefficients with their estimators $\hat{\gamma}_T(\cdot)$. The below given definition differs slightly from the definition in [33], but the same asymptotic results still apply.

DEFINITION 2.31. The estimator \hat{H}_T is defined as the unique positive semidefinite solution to the perturbed CARE

$$\hat{B}_T^\top \hat{H}_T + \hat{H}_T \hat{B}_T - \hat{H}_T \hat{C}_T \hat{H}_T + \hat{D}_T = 0$$

whenever $\hat{C}_T, \hat{D}_T > 0$. Otherwise, we set $\hat{H}_T = 0$.

As in discrete time, asymptotic properties of \hat{H}_T are inherited from the autocovariance estimators. However, due to the continuous time setting, instead of pointwise convergence, we have to consider functional form of convergence of $\hat{\gamma}_T(\cdot)$. In [33], we have provided sufficient conditions in the case of Gaussian noise G with independent components, under which the assumptions of the following theorems are satisfied. In particular, the results are valid for fractional Brownian motion that is widely applied in the field of mathematical finance.

THEOREM 2.32. Let $C, D > 0$. Assume that

$$\sup_{s \in [0, t]} \|\hat{\gamma}_T(s) - \gamma(s)\| \xrightarrow{\mathbb{P}} 0.$$

Then

$$\|\hat{H}_T - H\| \xrightarrow{\mathbb{P}} 0,$$

where \hat{H}_T is given by Definition 2.31.

THEOREM 2.33. Let $Y = (Y_s)_{s \in [0, t]}$ be an n^2 -dimensional stochastic process with continuous paths almost surely and let $l(T)$ be a rate function. If

$$l(T) \text{vec}(\hat{\gamma}_T(s) - \gamma(s)) \xrightarrow{\text{law}} Y_s$$

in the uniform topology of continuous functions, then:

(1) Let \tilde{Y}_s be the permutation of elements of Y_s that corresponds to the order of elements of $\text{vec}((\hat{\gamma}_T(s) - \gamma(s))^\top)$. Then

$$l(T) \text{vec}(\Delta_T C, \Delta_T B, \Delta_T D) \xrightarrow{\text{law}} \begin{bmatrix} \int_0^t (t-s)(Y_s + \tilde{Y}_s) ds \\ \int_0^t (Y_s - \tilde{Y}_s) ds \\ Y_t + \tilde{Y}_t - 2Y_0 \end{bmatrix} =: L_1(Y).$$

(2) If $C, D > 0$ and \hat{H}_T is given by Definition 2.31, then

$$l(T) \text{vec}(\hat{H}_T - H) \xrightarrow{\text{law}} L_2(L_1(Y)),$$

where $L_2: \mathbb{R}^{3n^2} \rightarrow \mathbb{R}^{n^2}$ is a linear mapping expressible in terms of H, t and the covariance function of G .

3. PROOFS

In the following, we denote the smallest eigenvalue of $H > 0$ by λ_{\min} . Consequently $\|e^{kH}\| = \|\text{Qdiag}(e^{\lambda_{\min} k})\text{Q}^\top\| = e^{\lambda_{\min} k}$ for a negative k .

3.1. Discrete Time

The proof of the next lemma follows the lines of the proof of Theorem 2.2. in [41] that concerns the one-dimensional continuous time case. However, in our setting, we obtain a weaker sufficient condition for $G \in \mathcal{G}_H$ for all $H > 0$.

LEMMA 3.1. Let $G = (G_t)_{t \in \mathbb{Z}}$ be an n -dimensional stationary increment process with $G_0 = 0$. Assume that

$$\mathbb{E} \left(\left| \log \|G_1\| \mathbb{1}_{\{\|G_1\| > 1\}} \right|^{1+\delta} \right) < \infty$$

for some $\delta > 0$. Then $G \in \mathcal{G}_H$ for all positive definite $n \times n$ matrices H .

PROOF. Let $H > 0$. We apply the Borel-Cantelli lemma together with Markov's inequality to show that $\|e^{kH} \Delta_k G\| \rightarrow 0$ almost surely as $k \rightarrow -\infty$. Let $\epsilon > 0$ be fixed below.

$$\begin{aligned} \mathbb{P} \left(\|e^{kH} \Delta_k G\| > \epsilon \right) &\leq \mathbb{P} \left(e^{\lambda_{\min} k} \|\Delta_k G\| > \epsilon \right) \\ &= \mathbb{P} \left(e^{\lambda_{\min} k} \|G_1\| > \epsilon \right) = \mathbb{P} \left(\|G_1\| > \frac{\epsilon}{e^{\lambda_{\min} k}} \right) \\ &= \mathbb{P} \left(\log \|G_1\| > \log \epsilon - \lambda_{\min} k \right), \end{aligned}$$

since ΔG is stationary and $G_0 = 0$. Furthermore, $k(\frac{\log \epsilon}{\lambda_{\min}} - \lambda_{\min}) \geq -Ck$ for some $C > 0$ and $k \leq k_\epsilon$. Thus, for $k \leq k_\epsilon$,

$$\begin{aligned} \mathbb{P} \left(\|e^{kH} \Delta_k G\| > \epsilon \right) &\leq \mathbb{P} \left(\log \|G_1\| \geq -Ck \right) \\ &= \mathbb{P} \left(\log \|G_1\| \mathbb{1}_{\{\|G_1\| > 1\}} \geq -Ck \right) \\ &\leq \frac{\mathbb{E} \left| \log \|G_1\| \mathbb{1}_{\{\|G_1\| > 1\}} \right|^{1+\delta}}{(-Ck)^{1+\delta}} \leq c \frac{1}{(-k)^{1+\delta}} \end{aligned}$$

giving the wanted result. We conclude the proof by noting that

$$\begin{aligned} \sum_{k=-\infty}^0 \|e^{kH} \Delta_k G\| &\leq \sum_{k=-\infty}^0 \|e^{\frac{1}{2}kH}\| \|e^{\frac{1}{2}kH} \Delta_k G\| \leq \\ \sup_k \|e^{\frac{1}{2}kH} \Delta_k G\| &\sum_{k=-\infty}^0 e^{\frac{1}{2}\lambda_{\min} k} < \infty \end{aligned}$$

almost surely.

Next, we extend the concept of self-similarity to discrete time multivariate processes. While the following two definitions are natural, to the best of our knowledge they are not widely acknowledged in the literature.

DEFINITION 3.2. Let $H > 0$ be a positive definite $n \times n$ matrix, and let $Y = (Y_{e^t})_{t \in \mathbb{Z}}$ be an n -dimensional stochastic process. Then Y is H -self-similar if

$$(Y_{e^{t+s}})_{t \in \mathbb{Z}} \stackrel{\text{law}}{=} (e^{sH} Y_{e^t})_{t \in \mathbb{Z}}$$

for every $s \in \mathbb{Z}$.

REMARK 3.3. Similarly as in Definition 2.2, positive definite matrices serve as natural counterparts of the conventional positive scalar valued exponents of self-similarity. In addition, $\lim_{l \rightarrow -\infty} Y_{e^l}$ does not converge (see the proof of the auxiliary Lemma 3.7) in the case of a general H -self-similar Y with non-positive definite H . It seems intuitive to expect the term to be convergent. However, posing such requirement leads to degeneracy of Y e.g., when H is symmetric with negative eigenvalues [cf. Remark 3.2 in [33]]. Furthermore, this indicates a possibility of dimensional reduction. For details on continuous time self-similar processes, we refer to [43]. See also the discussion on non-positive exponents of self-similarity in [44].

The following transform and the corresponding theorem giving one-to-one correspondence between self-similar and stationary processes were originally introduced by Lamperti in the univariate continuous time setting [45].

DEFINITION 3.4. Let $H > 0$ be a positive definite $n \times n$ matrix, and let $X = (X_t)_{t \in \mathbb{Z}}$ and $Y = (Y_{e^t})_{t \in \mathbb{Z}}$ be n -dimensional stochastic processes. We define

$$(\mathcal{L}_H X)_{e^t} = e^{tH} X_t$$

and

$$(\mathcal{L}_H^{-1} Y)_t = e^{-tH} Y_{e^t}.$$

THEOREM 3.5. The operator \mathcal{L}_H together with its inverse \mathcal{L}_H^{-1} define a bijection between n -dimensional stationary processes and n -dimensional H -self-similar processes.

PROOF. First, let X be stationary and set $Z_{e^t} = (\mathcal{L}_H X)_{e^t}$. Then

$$\begin{bmatrix} Z_{e^{t_1+s}} \\ Z_{e^{t_2+s}} \\ \vdots \\ Z_{e^{t_m+s}} \end{bmatrix} = \begin{bmatrix} e^{(t_1+s)H} X_{t_1+s} \\ e^{(t_2+s)H} X_{t_2+s} \\ \vdots \\ e^{(t_m+s)H} X_{t_m+s} \end{bmatrix} \stackrel{\text{law}}{=} \begin{bmatrix} e^{(t_1+s)H} X_{t_1} \\ e^{(t_2+s)H} X_{t_2} \\ \vdots \\ e^{(t_m+s)H} X_{t_m} \end{bmatrix} = \begin{bmatrix} e^{sH} Z_{e^{t_1}} \\ e^{sH} Z_{e^{t_2}} \\ \vdots \\ e^{sH} Z_{e^{t_m}} \end{bmatrix}$$

for every $m \in \mathbb{N}$, $t \in \mathbb{Z}^m$ and $s \in \mathbb{Z}$. Hence, Z is H -self-similar.

Now, let Y be H -self-similar and set $Z_t = (\mathcal{L}_H^{-1} Y)_t$. Then

$$\begin{bmatrix} Z_{t_1+s} \\ Z_{t_2+s} \\ \vdots \\ Z_{t_m+s} \end{bmatrix} = \begin{bmatrix} e^{-(t_1+s)H} Y_{e^{t_1+s}} \\ e^{-(t_2+s)H} Y_{e^{t_2+s}} \\ \vdots \\ e^{-(t_m+s)H} Y_{e^{t_m+s}} \end{bmatrix} \stackrel{\text{law}}{=} \begin{bmatrix} e^{-t_1H} Y_{e^{t_1}} \\ e^{-t_2H} Y_{e^{t_2}} \\ \vdots \\ e^{-t_mH} Y_{e^{t_m}} \end{bmatrix} = \begin{bmatrix} Z_{t_1} \\ Z_{t_2} \\ \vdots \\ Z_{t_m} \end{bmatrix}$$

for every $m \in \mathbb{N}$, $t \in \mathbb{Z}^m$ and $s \in \mathbb{Z}$. Hence, Z is stationary completing the proof.

REMARK 3.6. The form of the conventional continuous time Lamperti transform is not directly applicable in the discrete time setting due to the scaling of time. Hence, in order to stay within the given discrete parameter sets, we use exponential clocks with self-similar processes in Definitions 3.2 and 3.4.

Before the proof of Theorem 2.4 we state an auxiliary lemma.

LEMMA 3.7. Let $H > 0$ be a positive definite $n \times n$ matrix, and let $(Y_{e^t})_{t \in \mathbb{Z}}$ be an n -dimensional H -self-similar process. We define a process $G = (G_t)_{t \in \mathbb{Z}}$ by

$$G_t = \begin{cases} \sum_{k=1}^t e^{-kH} \Delta_k Y_{e^k}, & t \geq 1 \\ 0, & t = 0 \\ -\sum_{k=t+1}^0 e^{-kH} \Delta_k Y_{e^k}, & t \leq -1. \end{cases}$$

Then $G \in \mathcal{G}_H$.

PROOF. It is straightforward to verify that $\Delta_t G = e^{-tH} \Delta_t Y_{e^t}$ for every $t \in \mathbb{Z}$. In addition

$$\lim_{l \rightarrow -\infty} \sum_{k=l}^0 e^{kH} \Delta_k G = \lim_{l \rightarrow -\infty} \sum_{k=l}^0 \Delta_k Y_{e^k} = Y_1 - \lim_{l \rightarrow -\infty} Y_{e^{l-1}},$$

where by self-similarity of Y

$$\mathbb{P}(\|Y_{e^{l-1}}\| \geq \epsilon) = \mathbb{P}(\|e^{H(l-1)} Y_1\| \geq \epsilon) \leq \mathbb{P}(e^{\lambda_{\min}(l-1)} \|Y_1\| \geq \epsilon) \rightarrow 0.$$

Hence, we set

$$\lim_{l \rightarrow -\infty} \sum_{k=l}^0 e^{kH} \Delta_k G = Y_1.$$

PROOF OF THEOREM 2.4. Assume that $\lim_{t \rightarrow -\infty} e^{tH} X_t \stackrel{\mathbb{P}}{=} 0$ and (5) holds for $G \in \mathcal{G}_H$. Then, by using (5) repeatedly

$$\begin{aligned} X_t &= e^{-H} X_{t-1} + \Delta_t G = e^{-(n+1)H} X_{t-n-1} + \sum_{j=0}^n e^{-jH} \Delta_{t-j} G \\ &= e^{-(n+1)H} X_{t-n-1} + e^{-tH} \sum_{k=t-n}^t e^{kH} \Delta_k G = e^{-tH} \left(e^{(t-n-1)H} X_{t-n-1} + \sum_{k=t-n}^t e^{kH} \Delta_k G \right) \end{aligned}$$

for every $n \in \mathbb{N}$. Since, as $n \rightarrow \infty$, the limit of the sum above is well-defined, and $\lim_{n \rightarrow \infty} e^{(t-n-1)H} X_{t-n-1} \stackrel{\mathbb{P}}{=} 0$, we obtain that

$$X_t = e^{-tH} \sum_{k=-\infty}^t e^{kH} \Delta_k G. \tag{14}$$

Let $m \in \mathbb{N}, t \in \mathbb{Z}^m$, and $s \in \mathbb{Z}$. Then, by stationary increments of G , we have

$$\begin{bmatrix} e^{-(t_1+s)H} \sum_{j=-M}^{t_1} e^{(j+s)H} \Delta_{j+s} G \\ \vdots \\ e^{-(t_m+s)H} \sum_{j=-M}^{t_m} e^{(j+s)H} \Delta_{j+s} G \end{bmatrix} \stackrel{\text{law}}{=} \begin{bmatrix} e^{-t_1 H} \sum_{j=-M}^{t_1} e^{jH} \Delta_j G \\ \vdots \\ e^{-t_m H} \sum_{j=-M}^{t_m} e^{jH} \Delta_j G \end{bmatrix}$$

for every $-M < \min\{t_i\}$. Since the random vectors above converge in probability as $M \rightarrow \infty$, we obtain that

$$\begin{bmatrix} X_{t_1+s} \\ \vdots \\ X_{t_m+s} \end{bmatrix} = \begin{bmatrix} e^{-(t_1+s)H} \sum_{j=-\infty}^{t_1} e^{(j+s)H} \Delta_{j+s} G \\ \vdots \\ e^{-(t_m+s)H} \sum_{j=-\infty}^{t_m} e^{(j+s)H} \Delta_{j+s} G \end{bmatrix} \stackrel{\text{law}}{=} \begin{bmatrix} e^{-t_1 H} \sum_{j=-\infty}^{t_1} e^{jH} \Delta_j G \\ \vdots \\ e^{-t_m H} \sum_{j=-\infty}^{t_m} e^{jH} \Delta_j G \end{bmatrix} = \begin{bmatrix} X_{t_1} \\ \vdots \\ X_{t_m} \end{bmatrix}$$

and hence, X is stationary.

Next, assume that X is stationary. Then, by Theorem 3.5 there exists a H -self-similar Y such that

$$\Delta_t X = e^{-tH} Y_{e^t} - e^{-(t-1)H} Y_{e^{t-1}} = (e^{-H} - I)X_{t-1} + e^{-tH} \Delta_t Y_{e^t}.$$

Defining G as in Lemma 3.7 completes the proof of the other direction.

To prove uniqueness, we use (15). Assume that, for $G, \tilde{G} \in \mathcal{G}_H$,

$$e^{tH} X_t = \sum_{k=-\infty}^t e^{kH} \Delta_k G = \sum_{k=-\infty}^t e^{kH} \Delta_k \tilde{G}$$

for every $t \in \mathbb{Z}$. Then

$$e^{tH} X_t - e^{(t-1)H} X_{t-1} = e^{tH} \Delta_t G = e^{tH} \Delta_t \tilde{G}.$$

Since e^{tH} is invertible and both processes start from zero, we conclude that $G = \tilde{G}$.

PROOF OF LEMMA 2.7. We have that

$$\Delta_t G (\Delta_0 G)^\top = (X_t - \Phi X_{t-1})(X_0^\top - X_{-1}^\top \Phi).$$

Taking expectations yields

$$r(t) = \Phi \gamma(t) \Phi - \gamma(t+1) \Phi - \Phi \gamma(t-1) + \gamma(t).$$

PROOF OF THEOREM 2.10. Let

$$\Delta_t X = -\Theta X_{t-1} + \Delta_t G.$$

Then for $t \in \mathbb{N}$ we have

$$G_t = \sum_{k=1}^t \Delta_k G = \sum_{k=1}^t \Delta_k X + \Theta \sum_{k=1}^t X_{k-1} = X_t - X_0 + \Theta \sum_{k=1}^t X_{k-1}.$$

Hence

$$\begin{aligned} \text{cov}(G_t) &= \text{cov}(X_t - X_0) + \mathbb{E} \left[(X_t - X_0) \sum_{k=1}^t X_{k-1}^\top \right] \Theta \\ &+ \Theta \mathbb{E} \left[\sum_{k=1}^t X_{k-1} (X_t - X_0)^\top \right] + \Theta \mathbb{E} \left[\sum_{k=1}^t X_{k-1} \left(\sum_{k=1}^t X_{k-1} \right)^\top \right] \Theta \end{aligned}$$

giving (8) since

$$\begin{aligned} \mathbb{E} \left[(X_t - X_0) \sum_{k=1}^t X_{k-1}^\top \right] &= \sum_{k=1}^t \gamma(t-k+1) - \gamma(-k+1) \\ &= \sum_{k=1}^t \gamma(k) - \gamma(k-1)^\top \end{aligned}$$

and

$$\begin{aligned} \text{cov}(X_t - X_0) &= \mathbb{E} \left[(X_t - X_0)(X_t - X_0)^\top \right] = 2\gamma(0) - \gamma(t) \\ &- \gamma(-t) = 2\gamma(0) - \gamma(t) - \gamma(t)^\top \end{aligned}$$

PROOF OF LEMMA 2.15. Assume that $\tilde{\Phi} = e^{-\tilde{H}}$ satisfies (9) for every $t \in \mathbb{Z}$ and set

$$X_t - \tilde{\Phi} X_{t-1} = \Delta_t \tilde{G}, \quad t \in \mathbb{Z}.$$

Consequently

$$\tilde{r}(t) = \gamma(t) \tilde{\Phi}^2 - (\gamma(t+1) - \gamma(t-1)) \tilde{\Phi} + \gamma(t) = r(t), \quad t \in \mathbb{Z},$$

where $\tilde{r}(t)$ is the autocovariance function of $(\Delta_t \tilde{G})_{t \in \mathbb{Z}}$. Now, since $G_0 = \tilde{G}_0 = 0$, we obtain that

$$\begin{aligned} \text{var}(G_t) &= \text{var} \left(\sum_{k=1}^t \Delta_k G \right) = \sum_{k,j=1}^t \text{cov}(\Delta_k G, \Delta_j G) = \sum_{k,j=1}^t r(k-j) \\ &= \text{var} \left(\sum_{k=1}^t \Delta_k \tilde{G} \right) = \text{var}(\tilde{G}_t) \end{aligned}$$

for all $t \in \mathbb{N}$. Hence, both Θ and $\tilde{\Theta}$ are solutions to (10).

LEMMA 3.8. The matrix $\Theta = I - e^{-H}$ is positive definite.

PROOF. Let a be a real vector of length n , and let $H = \mathcal{Q} \Lambda \mathcal{Q}^\top$ be an eigendecomposition of H . Then

$$a^\top (I - e^{-H}) a = \|a\|^2 - a^\top e^{-H} a,$$

where

$$|a^\top e^{-H} a| \leq \|a\|^2 e^{-\lambda_{\min}} < \|a\|^2$$

completing the proof.

In order to show that $\hat{\Theta}_T$ is consistent, we simply need to find suitable bounds for $\Delta_T B, \Delta_T C$ and $\Delta_T D$ in terms of the autocovariance estimators. After that, the same strategy as in [33] can be applied.

LEMMA 3.9. Set

$$M_{t,T} = \max_{s \in \{0,1,\dots,t\}} \|\hat{\gamma}_T(s) - \gamma(s)\|.$$

Then the coefficients of the perturbed CARE satisfy

$$\begin{aligned} \|\Delta_T D\| &\leq 4M_{t,T} \\ \|\Delta_T C\| &\leq t^2 M_{t,T} \\ \|\Delta_T B\| &\leq 2tM_{t,T}, \end{aligned}$$

PROOF. First, we recall first that

$$\|\hat{\gamma}_T(-s) - \gamma(-s)\| = \|\hat{\gamma}_T(s)^\top - \gamma(s)^\top\| = \|\hat{\gamma}_T(s) - \gamma(s)\|.$$

Now, since $v(t)$ is known,

$$\|\Delta_T D\| \leq 2\|\hat{\gamma}_T(0) - \gamma(0)\| + \|\hat{\gamma}_T(t) - \gamma(t)\| + \|\hat{\gamma}_T(t)^\top - \gamma(t)^\top\| \leq 4M_{t,T}.$$

Moreover

$$\|\Delta_T C\| \leq \sum_{k=1}^t \sum_{j=1}^t \|\hat{\gamma}_T(k-j) - \gamma(k-j)\| \leq t^2 M_{t,T}.$$

Finally

$$\begin{aligned} \|\Delta_T B\| &\leq \sum_{k=1}^t \|\hat{\gamma}_T(k-1) - \gamma(k-1)\| + \|\gamma(k)^\top - \hat{\gamma}_T(k)^\top\| \\ &\leq 2tM_{t,T}. \end{aligned}$$

PROOF OF THEOREM 2.18. The result follows by replacing $\sup_{s \in [0,t]} \|\hat{\gamma}_T(s) - \gamma(s)\|$ with $M_{t,T}$ in Corollary 3.14 and in the proof of Theorem 2.9 of [33]. The details are left to the reader.

PROOF OF THEOREM 2.19. For the first part of the theorem, we notice that

$$\begin{aligned} C_t &= \sum_{k=1}^t \sum_{j=1}^t \gamma(k-j) = \sum_{k=1}^t \sum_{l=k-t}^{k-1} \gamma(l) = \sum_{l=1-t}^{-1} \sum_{k=1}^{t+l} \gamma(l) \\ &+ \sum_{l=0}^{t-1} \sum_{k=l+1}^t \gamma(l) = \sum_{l=1-t}^{-1} (t+l)\gamma(l) + \sum_{l=0}^{t-1} (t-l)\gamma(l) \\ &= \sum_{l=0}^{t-1} (t-l)\gamma(l) + \sum_{l=1}^{t-1} (t-l)\gamma(l)^\top, \end{aligned}$$

where \sum_0^{-1} and \sum_1^0 are interpreted as empty sums. Now we have that

$$\Delta_T C = \sum_{k=0}^{t-1} (t-k)(\hat{\gamma}_T(k) - \gamma(k))$$

$$\begin{aligned} &+ \sum_{k=1}^{t-1} (t-k) \left(\hat{\gamma}_T(k)^\top - \gamma(k)^\top \right) \\ \Delta_T B &= \sum_{k=1}^t \hat{\gamma}_T(k-1) - \gamma(k-1) - \hat{\gamma}_T(k)^\top + \gamma(k)^\top \\ \Delta_T D &= 2(\gamma(0) - \hat{\gamma}_T(0)) + \hat{\gamma}_T(t) - \gamma(t) + \hat{\gamma}_T(t)^\top - \gamma(t)^\top \end{aligned}$$

and furthermore

$$\begin{aligned} &I(T) \text{vec}(\Delta_T C, \Delta_T B, \Delta_T D) = I(T) \\ &\begin{bmatrix} \sum_{k=0}^{t-1} (t-k) \text{vec}(\hat{\gamma}_T(k) - \gamma(k)) + \sum_{k=1}^{t-1} (t-k) \text{vec} \left((\hat{\gamma}_T(k) - \gamma(k))^\top \right) \\ \sum_{k=1}^t \text{vec}(\hat{\gamma}_T(k-1) - \gamma(k-1)) - \text{vec} \left((\hat{\gamma}_T(k) - \gamma(k))^\top \right) \\ -2 \text{vec}(\hat{\gamma}_T(0) - \gamma(0)) + \text{vec}(\hat{\gamma}_T(t) - \gamma(t)) + \text{vec} \left((\hat{\gamma}_T(t) - \gamma(t))^\top \right) \end{bmatrix} \\ &= L_1 \begin{pmatrix} I(T) \begin{bmatrix} \text{vec}(\hat{\gamma}_T(0) - \gamma(0)) \\ \text{vec}(\hat{\gamma}_T(1) - \gamma(1)) \\ \vdots \\ \text{vec}(\hat{\gamma}_T(t) - \gamma(t)) \end{bmatrix} \end{pmatrix} \xrightarrow{\text{law}} L_1(Z) \end{aligned}$$

by the continuous mapping theorem. For the second part of the theorem, the proof of the continuous time case of [33] can be applied just by replacing $\sup_{s \in [0,t]} \|\hat{\gamma}_T(s) - \gamma(s)\|$ with $M_{t,T}$ in the definition of the set A_T .

3.2. Continuous Time

We only provide the proof of Proposition 2.24, while the other proofs can be found from [33]

PROOF OF PROPOSITION 2.24. Integrating (11) from 0 to t gives

$$G_t = X_t - X_0 + H \int_0^t X_s ds.$$

Hence

$$\begin{aligned} \Delta_t^\delta G (\Delta_0^\delta G)^\top &= \left(X_t - X_{t-\delta} + H \int_{t-\delta}^t X_s ds \right) \\ &\left(X_0^\top - X_{-\delta}^\top + \int_{-\delta}^0 X_s^\top ds H \right). \end{aligned}$$

Taking expectations yields

$$\begin{aligned} r_\delta(t) &= 2\gamma(t) - \gamma(t+\delta) - \gamma(t-\delta) + \int_{-\delta}^0 \gamma(t-s) - \gamma(t-\delta-s) ds H \\ &+ H \int_{t-\delta}^t \gamma(s) - \gamma(s+\delta) ds + H \int_{t-\delta}^t \int_{-\delta}^0 \gamma(s-u) dud s H, \end{aligned}$$

where the first order terms can be treated with a simple change of variables. For the second order term we obtain that

$$\begin{aligned} &\int_{t-\delta}^t \int_{-\delta}^0 \gamma(s-u) dud s = \int_{t-\delta}^t \int_s^{s+\delta} \gamma(x) dx ds \\ &= \int_{t-\delta}^t \int_{t-\delta}^x \gamma(x) ds dx + \int_t^{t+\delta} \int_{x-\delta}^t \gamma(x) ds dx \\ &= \int_{t-\delta}^t (x-t+\delta)\gamma(x) dx + \int_t^{t+\delta} (t-x+\delta)\gamma(x) dx. \end{aligned}$$

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

ACKNOWLEDGMENTS

I would like to thank Lauri Viitasaari for his comments and suggestions.

REFERENCES

- Brockwell PJ, Davis RA. *Time Series: Theory and Methods*. New York, NY: Springer Science & Business Media (1991). doi: 10.1007/978-1-4419-0320-4
- Hamilton JD. *Time Series Analysis*. Vol. 2. Princeton: Princeton University Press (1994).
- Neusser K. *Time Series Econometrics*. Springer (2016). doi: 10.1007/978-3-319-32862-1
- Engle RF. Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometr J Econometr Soc.* (1982) **50**:987–1007. doi: 10.2307/1912773
- Bollerslev T. Generalized autoregressive conditional heteroskedasticity. *J. Econometr.* (1986) **31**:307–27. doi: 10.1016/0304-4076(86)90063-1
- Hannan EJ. The asymptotic theory of linear time-series models. *J Appl Probabil.* (1973) **10**:130–45. doi: 10.1017/S0021900200042145
- Tiao GC, Tsay RS. Consistency properties of least squares estimates of autoregressive parameters in ARMA models. *Ann Stat.* (1983) **11**:856–71. doi: 10.1214/aos/1176346252
- Mikosch T, Gadrich T, Kluppelberg C, Adler RJ. Parameter estimation for ARMA models with infinite variance innovations. *Ann Stat.* (1995) **23**:305–26. doi: 10.1214/aos/1176324469
- Mauricio JA. Exact maximum likelihood estimation of stationary vector ARMA models. *J Am Stat Assoc.* (1995) **90**:282–91. doi: 10.1080/01621459.1995.10476511
- Ling S, McAleer M. Asymptotic theory for a vector ARMA-GARCH model. *Econometr Theory.* (2003) **19**:280–310. doi: 10.1017/S0266466603192092
- Franco C, Zakoian JM. Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli.* (2004) **10**:605–37. doi: 10.3150/bj/1093265632
- Bollerslev T. Glossary to ARCH (GARCH). *CREATES Res Pap.* (2008) **49**:1–44. doi: 10.2139/ssrn.1263250
- Han H, Kristensen D. Asymptotic theory for the QMLE in GARCH-X models with stationary and nonstationary covariates. *J Bus Econ Stat.* (2014) **32**:416–29. doi: 10.1080/07350015.2014.897954
- Baillie RT, Bollerslev T, Mikkelsen HO. Fractionally integrated generalized autoregressive conditional heteroskedasticity. *J Econometr.* (1996) **74**:3–10. doi: 10.1016/S0304-4076(95)01749-6
- Ling S, Li WK. On fractionally integrated autoregressive moving-average time series models with conditional heteroscedasticity. *J Am Stat Assoc.* (1997) **92**:1184–94. doi: 10.1080/01621459.1997.10474076
- Zhang MY, Russell JR, Tsay RS. A nonlinear autoregressive conditional duration model with applications to financial transaction data. *J Econometr.* (2001) **104**:179–207. doi: 10.1016/S0304-4076(01)00063-X
- Voutilainen M, Viitasaari L, Ilmonen P. On model fitting and estimation of strictly stationary processes. *Modern Stochast Theory Appl.* (2017) **4**:381–406. doi: 10.15559/17-VMSTA91
- Voutilainen M, Viitasaari L, Ilmonen P, Torres S, Tudor C. On the ARCH model with stationary liquidity. (2020). Available online at: <https://link.springer.com/content/pdf/10.1007/s00184-020-00779-x.pdf>
- Cheridito P, Kawaguchi H, Maejima M. Fractional Ornstein-Uhlenbeck processes. *Electron J Probabil.* (2003) **8**:1–44. doi: 10.1214/EJP.v8-125
- Hu Y, Nualart D. Parameter estimation for fractional Ornstein-Uhlenbeck processes. *Stat Probabil Lett.* (2010) **80**:1030–8. doi: 10.1016/j.spl.2010.02.018
- Kleptsyna ML, Le Breton A. Statistical analysis of the fractional Ornstein-Uhlenbeck type process. *Stat Inference Stochast Process.* (2002) **5**:229–48. doi: 10.1023/A:1021220818545
- Azmoodeh E, Viitasaari L. Parameter estimation based on discrete observations of fractional Ornstein-Uhlenbeck process of the second kind. *Stat Inference Stochast Process.* (2015) **18**:205–27. doi: 10.1007/s11203-014-9111-8
- Bajja S, Es-Sebaïy K, Viitasaari L. Least squares estimator of fractional Ornstein-Uhlenbeck processes with periodic mean. *J Korean Stat Soc.* (2017) **46**:608–22. doi: 10.1016/j.jkss.2017.06.002
- Balde MF, Es-Sebaïy K, Tudor CA. Ergodicity and drift parameter estimation for infinite-dimensional fractional Ornstein-Uhlenbeck process of the second kind. *Appl Math Opt.* (2020) **81**:785–814. doi: 10.1007/s00245-018-9519-4
- Brouste A, Iacus SM. Parameter estimation for the discretely observed fractional Ornstein-Uhlenbeck process and the Yuima R package. *Comput Stat.* (2013) **28**:1529–47. doi: 10.1007/s00180-012-0365-6
- Douissi S, Es-Sebaïy K, Tudor CA. Hermite Ornstein-Uhlenbeck processes mixed with a Gamma distribution. *Publicaciones Mathematicae Debrecen.* (2020) **96**:1–22. doi: 10.5486/PMD.2020.8443
- Hu Y, Nualart D, Zhou H. Parameter estimation for fractional Ornstein-Uhlenbeck processes of general Hurst parameter. *Stat Inference Stochast Process.* (2019) **22**:111–42. doi: 10.1007/s11203-017-9168-2
- Nourdin I, Tran TD. Statistical inference for Vasicek-type model driven by Hermite processes. *Stochast Process Appl.* (2019) **129**:3774–91. doi: 10.1016/j.spa.2018.10.005
- Sottinen T, Viitasaari L. Parameter estimation for the Langevin equation with stationary-increment Gaussian noise. *Stat Inference Stochast Process.* (2018) **21**:569–601. doi: 10.1007/s11203-017-9156-6
- Tanaka K. Maximum likelihood estimation for the non-ergodic fractional Ornstein-Uhlenbeck process. *Stat Inference Stochast Process.* (2015) **18**:315–32. doi: 10.1007/s11203-014-9110-9
- Applebaum D, et al. Infinite dimensional Ornstein-Uhlenbeck processes driven by Lévy processes. *Probabil Surv.* (2015) **12**:33–54. doi: 10.1214/14-PS249
- Magdziarz M. Fractional Ornstein-Uhlenbeck processes. Joseph effect in models with infinite variance. *Phys A Stat Mech Appl.* (2008) **387**:123–33. doi: 10.1016/j.physa.2007.08.016
- Voutilainen M, Viitasaari L, Ilmonen P, Torres S, Tudor C. Vector-valued generalised Ornstein-Uhlenbeck processes. *arXiv[Preprint].arXiv:1909.02376.* (2019).
- Lancaster P, Rodman L. *Algebraic Riccati Equations*. New York, NY: Oxford University Press (1995).
- Kucera V. A contribution to matrix quadratic equations. *IEEE Trans Automat Control.* (1972) **17**:344–7. doi: 10.1109/TAC.1972.1099983
- Wonham WM. On a matrix Riccati equation of stochastic control. *SIAM J Control.* (1968) **6**:681–97. doi: 10.1137/0306044
- Sun Jg. Perturbation theory for algebraic Riccati equations. *SIAM J Matrix Anal Appl.* (1998) **19**:39–65. doi: 10.1137/S0895479895291303
- Byers R. Solving the algebraic Riccati equation with the matrix sign function. *Linear Algebra Appl.* (1987) **85**:267–79. doi: 10.1016/0024-3795(87)90222-9
- Laub A. A Schur method for solving algebraic Riccati equations. *IEEE Trans Automat Control.* (1979) **24**:913–21. doi: 10.1109/TAC.1979.1102178
- Bini DA, Iannazzo B, Meini B. *Numerical Solution of Algebraic Riccati Equations*. Vol. 9. Philadelphia: SIAM (2012). doi: 10.1137/1.9781611972092
- Viitasaari L. Representation of stationary and stationary increment processes via Langevin equation and self-similar processes. *Stat Probabil Lett.* (2016) **115**:45–53. doi: 10.1016/j.spl.2016.03.020

42. Voutilainen M, Viitasaari L, Ilmonen P. Note on AR(1)-characterisation of stationary processes and model fitting. *Modern Stochast Theory Appl.* (2019) **6**:195–207. doi: 10.15559/19-VMSTA132
43. Embrechts P, Maejima M. *Selfsimilar Processes*. Princeton: Princeton University Press (2002). doi: 10.1515/9781400825103
44. Pipiras V, Taqqu MS. *Long-Range Dependence and Self-Similarity*. Vol. 45. Cambridge: Cambridge University Press (2017). doi: 10.1017/CBO9781139600347
45. Lamperti J. Semi-stable stochastic processes. *Trans Am Math Soc.* (1962) **104**:62–78. doi: 10.1090/S0002-9947-1962-0138128-7

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Voutilainen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.