

# Note on AR(1)-characterisation of stationary processes and model fitting

Marko Voutilainen<sup>a,\*</sup>, Lauri Viitasaari<sup>b</sup>, Pauliina Ilmonen<sup>a</sup>

<sup>a</sup>*Department of Mathematics and Systems Analysis, Aalto University School of Science, P.O. Box 11100, FI-00076 Aalto, Finland*

<sup>b</sup>*Department of Mathematics and Statistics, University of Helsinki, P.O. Box 68, FI-00014 University of Helsinki, Finland*

[marko.voutilainen@aalto.fi](mailto:marko.voutilainen@aalto.fi) (M. Voutilainen), [lauri.viitasaari@iki.fi](mailto:lauri.viitasaari@iki.fi) (L. Viitasaari),  
[pauliina.ilmonen@aalto.fi](mailto:pauliina.ilmonen@aalto.fi) (P. Ilmonen)

Received: 5 October 2018, Revised: 13 February 2019, Accepted: 13 February 2019,  
Published online: 8 March 2019

**Abstract** It was recently proved that any strictly stationary stochastic process can be viewed as an autoregressive process of order one with coloured noise. Furthermore, it was proved that, using this characterisation, one can define closed form estimators for the model parameter based on autocovariance estimators for several different lags. However, this estimation procedure may fail in some special cases. In this article, a detailed analysis of these special cases is provided. In particular, it is proved that these cases correspond to degenerate processes.

**Keywords** AR(1)-characterisation, stationary processes, covariance functions

**2010 MSC** 60G10, 62M10

## 1 Introduction

Stationary processes are an important tool in many practical applications of time series analysis, and the topic is extensively studied in the literature. Traditionally, stationary processes are modelled by using autoregressive moving average processes or linear processes (see monographs [2, 4] for details).

---

\*Corresponding author.

One of the most simple example of an autoregressive moving average process is an autoregressive process of order one. That is, a process  $(X_t)_{t \in \mathbb{Z}}$  defined by

$$X_t = \phi X_{t-1} + \varepsilon_t, \quad t \in \mathbb{Z}, \quad (1)$$

where  $\phi \in (-1, 1)$  and  $(\varepsilon_t)_{t \in \mathbb{Z}}$  is a sequence of independent and identically distributed square integrable random variables. The continuous time analogue of (1) is called the Ornstein–Uhlenbeck process, which can be defined as the stationary solution of the Langevin-type stochastic differential equation

$$dU_t = -\phi U_t dt + dW_t, \quad (2)$$

where  $\phi > 0$  and  $(W_t)_{t \in \mathbb{R}}$  is a two-sided Brownian motion. Such equations have also applications in mathematical physics.

Statistical inference for AR(1)-process or Ornstein–Uhlenbeck process is well-established in the literature. Furthermore, recently a generalised continuous time Langevin equation, where the Brownian motion  $W$  in (2) is replaced with a more general driving force  $G$ , have been a subject of an active study. Especially, the so-called fractional Ornstein–Uhlenbeck processes introduced by [3] have been studied extensively. For the parameter estimation in such models, we mention a recent monograph [6] dedicated to the subject, and the references therein.

When the model becomes more complicated, the number of parameters increases and the estimation may become a challenging task. For example, it may happen that standard maximum likelihood estimators cannot be expressed in closed form [2]. Even worse, it may happen that classical estimators such as maximum likelihood or least squares estimators are biased and not consistent (cf. [1] for discussions on the generalised ARCH-model with liquidity given by fractional Brownian motion). One way to tackle such problems is to consider one parameter model, and to replace the white noise in (1) with some other stationary noise. It was proved in [8] that each discrete time strictly stationary process can be characterised by

$$X_t = \phi X_{t-1} + Z_t, \quad (3)$$

where  $\phi \in (0, 1)$ . This representation can be viewed as a discrete time analogue of the fact that Langevin-type equation characterises strictly stationary processes in continuous time [7].

The authors in [8] applied characterisation (3) to model fitting and parameter estimation. The presented estimation procedure is straightforward to apply with the exception of certain special cases. The purpose of this paper is to provide a comprehensive analysis of these special cases. In particular, we show that such cases do not provide very useful models. This highlights the wide applicability of characterisation (3) and the corresponding estimation procedure.

The rest of the paper is organised as follows. In Section 2 we briefly discuss the motivating estimation procedure of [8]. We also present and discuss our main results together with some illustrative figures. All the proofs and technical lemmas are postponed to Section 3. Section 4 provides a small simulation study comparing an estimator of quadratic type arising out of (3) with the classical Yule–Walker estimator in the case of an AR(1)-process. We end the paper with discussion.

## 2 Motivation and formulation of the main results

Let  $X = (X_t)_{t \in \mathbb{Z}}$  be a stationary process. It was shown in [8] that equation

$$X_t = \phi X_{t-1} + Z_t, \tag{4}$$

where  $\phi \in (0, 1)$  and  $Z_t$  is another stationary process, characterises all discrete time (strictly) stationary processes. Throughout this paper we suppose that  $X$  and  $Z$  are square integrable processes with autocovariance functions  $\gamma(\cdot)$  and  $r(\cdot)$ , respectively. Using Equation (4), one can derive the quadratic equations of the Yule–Walker type for the parameter  $\phi$ , which can be solved in an explicit form. Namely, for any  $m \in \mathbb{Z}$  such that  $\gamma(m) \neq 0$  we have

$$\phi = \frac{\gamma(m+1) + \gamma(m-1) \pm \sqrt{(\gamma(m+1) + \gamma(m-1))^2 - 4\gamma(m)(\gamma(m) - r(m))}}{2\gamma(m)}. \tag{5}$$

The estimation of the parameter  $\phi$  is obvious from (5) provided that one can determine which sign, plus or minus, should be chosen. In practice, this can be done by choosing different lags  $m$  for which the covariance function  $\gamma(m)$  is estimated. Then one can determine the correct value  $\phi$  by comparing different signs in (5) for different lags  $m$  (We refer to [8, p. 387] for detailed discussion). However, this approach fails, i.e. one cannot find suitable lags leading to the correct choice of the sign and only one value  $\phi$ , if, for  $m \in \mathbb{Z}$  such that  $\gamma(m) = 0$  we also have  $r(m) = 0$ , and for any  $m \in \mathbb{Z}$  such that  $\gamma(m) \neq 0$ , the ratio

$$a_m = \frac{r(m)}{\gamma(m)} = a \tag{6}$$

for some constant  $a \in (0, 1)$ . The latter is equivalent [8, p. 387] to the fact that

$$\frac{\gamma(m+1) + \gamma(m-1)}{\gamma(m)} = b$$

for some constant  $b$  with  $\phi < b < \phi + \phi^{-1}$ . This leads to

$$\gamma(m+1) = b\gamma(m) - \gamma(m-1). \tag{7}$$

Moreover, if  $\gamma(m) = r(m) = 0$  for some  $m$ , it is straightforward to verify that (7) holds in this case as well. Thus (7) holds for all  $m \in \mathbb{Z}$ . Since covariance functions are necessarily symmetric, we obtain an “initial” condition  $\gamma(1) = \frac{b}{2}\gamma(0)$ . Thus (7) admits a unique symmetric solution.

By the Cauchy–Schwarz inequality and equality  $\gamma(1) = \frac{b}{2}\gamma(0)$ , it is clear that (7) does not define covariance function for  $b > 2$ . Furthermore, since  $\phi > 0$ , we conduct a comprehensive analysis of the special cases by studying the functions given by (7) with  $b \in [0, 2]$  (we include the trivial case  $b = 0$ ). For  $b = 2$  Equation (7) corresponds to the case  $X_t = X_0$  for all  $t \in \mathbb{Z}$  which is hardly interesting. Similarly, the case  $b = 0$  leads to a process  $(\dots, X_0, X_1, -X_0, -X_1, X_0, X_1, \dots)$  which again does not provide a practical model. On the other hand, it is not clear whether for some

other values  $b \in (0, 2)$  Equation (7) can lead to some non-trivial model in which the estimation procedure explained above cannot be applied. By our first main theorem, for any  $b \in [0, 2]$ , Equation (7) defines a covariance function. On the other hand, the resulting covariance function, denoted by  $\gamma_b$ , leads to a model that is not very interesting either.

**Theorem 2.1.** *Let  $b \in (0, 2)$  and  $\gamma_b$  be the (unique) symmetric function satisfying (7). Then*

1. *Let  $b = 2 \sin(\frac{k}{l} \frac{\pi}{2})$ , where  $k$  and  $l$  are strictly positive integers such that  $\frac{k}{l} \in (0, 1)$ . Then  $\gamma_b(m)$  is periodic.*
2. *Let  $b = 2 \sin(r \frac{\pi}{2})$ , where  $r \in (0, 1) \setminus \mathbb{Q}$ . Then for any  $M \geq 0$ , the set  $\{\gamma_b(M + m) : m \geq 0\}$  is dense in  $[-\gamma(0), \gamma(0)]$ .*
3. *For any  $b \in [0, 2]$ ,  $\gamma_b(\cdot)$  is a covariance function.*

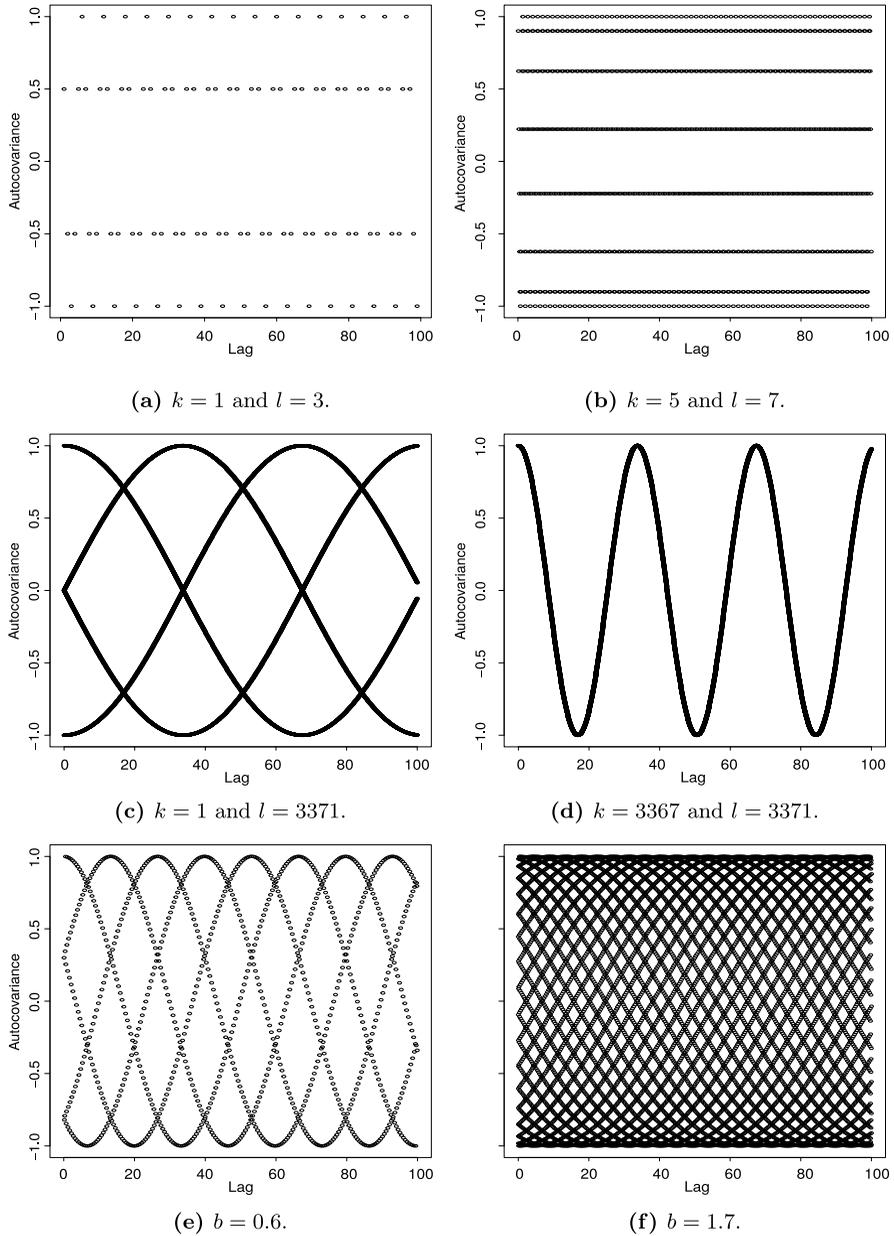
In many applications of stationary processes, it is assumed that the covariance function  $\gamma(\cdot)$  vanishes at infinity, or that  $\gamma(\cdot)$  is periodic. Note that the latter case corresponds simply to the analysis of finite-dimensional random vectors with identically distributed components. Indeed,  $\gamma(m) = \gamma(0)$  implies  $X_n = X_0$  almost surely, so periodicity of  $\gamma(\cdot)$  with period  $N$  implies that there exists at most  $N$  random variables as the source of randomness. By items (2) and (3) of Theorem 2.1, we observe that, for suitable values of  $b$ , (7) can be used to construct covariance functions that are neither periodic nor vanishing at infinity. On the other hand, in this case there are arbitrary large lags  $m$  such that  $\gamma_b(m)$  is arbitrary close to  $\gamma_b(0)$ . Consequently, due to the strong dependency structure, it is expected that different estimation procedures will fail. Indeed, even the standard covariance estimators are not consistent. A consequence of Theorem 2.1 is that only a little structure in the noise  $Z$  is needed in order to apply the estimation procedure of the parameter  $\phi$  introduced in [8], provided that one has consistent estimators for the covariances of  $X$ . The following is a precise mathematical formulation of this observation.

**Theorem 2.2.** *Let  $X$  be given by (4) for some  $\phi \in (0, 1)$  and noise  $Z$ . Assume that there exists  $\epsilon > 0$  and  $M \in \mathbb{N}$  such that  $r(m) \leq r(0)(1 - \epsilon)$  or  $r(m) \geq -r(0)(1 - \epsilon)$  for all  $m \geq M$ . Then the covariance function  $\gamma$  of  $X$  does not satisfy (7) for any  $b \in [0, 2]$ .*

In most situations, a natural assumption regarding the covariance of the noise is  $r(m) \rightarrow 0$  as  $m \rightarrow \infty$ . In this case, Theorem 2.2 gets obviously satisfied. We end this section by visual illustrations of the covariance functions defined by (7). In these examples we have set  $\gamma_b(0) = 1$ . In Figures 1a and 1b we have illustrated the case of item (1) of Theorem 2.1. Note that in Figure 1a we have  $b = 2 \sin(\frac{1}{3} \frac{\pi}{2}) = 1$ . Figures 1c and 1d demonstrate how  $k$  can affect the shape of the covariance function. Finally, Figures 1e and 1f illustrate the case of item (2) of Theorem 2.1.

### 3 Proofs

Throughout this section, without loss of generality, we assume  $\gamma_b(0) = 1$ . We also drop the subscript and simply write  $\gamma$  for  $\gamma_b$ . The following first result gives explicit formula for the solution to (7).



**Fig. 1.** Examples of covariance functions corresponding to  $b = 2 \sin(\frac{k}{l} \frac{\pi}{2})$  and  $b = 2 \sin(r \frac{\pi}{2})$

**Proposition 3.1.** *The unique symmetric solution to (7) is given by*

$$\gamma(m) = \begin{cases} (-1)^{\frac{m}{2}} \cos(m \arcsin(\frac{b}{2})), & \text{for } m \text{ is even} \\ (-1)^{\frac{(m-1)}{2}} \sin(m \arcsin(\frac{b}{2})), & \text{for } m \text{ is odd.} \end{cases} \quad (8)$$

**Proof.** Clearly,  $\gamma(m)$  given by (8) is symmetric, and thus it suffices to consider  $m \geq 0$ . Moreover  $\gamma(0) = 1$  and  $\gamma(1) = \frac{b}{2}$ . We use the short notation  $A = \arcsin(\frac{b}{2})$ , so that  $\sin A = \frac{b}{2}$ . Assume first  $m + 2 \equiv 2 \pmod{4}$ . Then

$$\begin{aligned} \gamma(m + 2) &= -\cos((m + 2)A) = -\cos(mA)\cos(2A) + \sin(mA)\sin(2A) \\ &= -\cos(mA)(1 - 2\sin^2 A) + 2\sin(mA)\sin A \cos A \\ &= -\cos(mA)(1 - b\sin A) + b\sin(mA)\cos A \\ &= b(\cos(mA)\sin A + \sin(mA)\cos A) - \cos(mA) \\ &= b\sin((m + 1)A) - \cos(mA) = b\gamma(m + 1) - \gamma(m). \end{aligned}$$

Treating the cases  $m + 2 \equiv 3 \pmod{4}$ ,  $m + 2 \equiv 0 \pmod{4}$  and  $m + 2 \equiv 1 \pmod{4}$  similarly, we deduce that (8) satisfies (7).  $\square$

**Remark 3.2.** Using (7) directly, we observe, for even  $m \geq 1$ , that

$$\gamma(m) = b^m + \sum_{n=\frac{m}{2}}^{m-1} (-1)^{m-n} \left( \binom{n}{m-n} b^{2n-m} + \binom{n}{m-n-1} \frac{b^{2n-m+2}}{2} \right). \tag{9}$$

Similarly, for odd  $m \geq 1$ , we obtain

$$\gamma(m) = \sum_{n=\frac{m+1}{2}}^m (-1)^{m-n} \binom{n}{m-n} b^{2n-m} + \sum_{n=\frac{m-1}{2}}^{m-1} (-1)^{m-n} \binom{n}{m-n-1} \frac{b^{2n-m+2}}{2}. \tag{10}$$

These formulas are finite polynomial expansions, in variable  $b$ , of the functions presented in (8) which could have been deduced also by using some well-known trigonometric identities.

Before proving our main theorems we need several technical lemmas.

**Definition 3.3.** We denote with  $Q$  a subset of rationals defined by

$$Q := \left\{ \frac{k}{l} : k, l \in \mathbb{N}, \frac{k}{l} \in (0, 1), k - l \equiv 1 \pmod{2} \right\}.$$

**Remark 3.4.** The modulo condition above means only that either  $k$  is even and  $l$  is odd, or vice versa.

**Lemma 3.5.** Let  $x = \frac{k}{l}\frac{\pi}{2}$ , where  $\frac{k}{l} \in Q$ . Then

$$\sum_{j=1}^{2l-1} \cos^2(jx)(-1)^j = -1.$$

**Proof.** We write

$$\sum_{j=1}^{2l-1} \cos^2(jx)(-1)^j = \cos^2(lx)(-1)^l + \sum_{j=1}^{l-1} \cos^2(jx)(-1)^j + \sum_{j=l+1}^{2l-1} \cos^2(jx)(-1)^j.$$

Change of variable  $t = j - l$  gives

$$\begin{aligned} \sum_{j=l+1}^{2l-1} \cos^2(jx)(-1)^j &= \sum_{t=1}^{l-1} \cos^2((t+l)x)(-1)^{t+l} \\ &= \sum_{t=1}^{l-1} \cos^2\left(tx + k\frac{\pi}{2}\right)(-1)^{t+l} = \begin{cases} \sum_{t=1}^{l-1} \cos^2(tx)(-1)^{t+l}, & k \text{ even} \\ \sum_{t=1}^{l-1} \sin^2(tx)(-1)^{t+l}, & k \text{ odd.} \end{cases} \end{aligned}$$

Consequently, for even  $k$  and odd  $l$  we have

$$\sum_{j=1}^{2l-1} \cos^2(jx)(-1)^j = -\cos^2\left(k\frac{\pi}{2}\right) = -1.$$

Similarly, for odd  $k$  and even  $l$ ,

$$\sum_{j=1}^{2l-1} \cos^2(jx)(-1)^j = \cos^2\left(k\frac{\pi}{2}\right) + \sum_{j=1}^{l-1} (-1)^j = -1.$$

□

**Lemma 3.6.** Let  $\gamma(\cdot)$  be given by (8) with  $b = 2 \sin\left(\frac{k}{l}\frac{\pi}{2}\right)$  for some  $\frac{k}{l} \in \mathcal{Q}$ . Then the non-zero eigenvalues of the matrix

$$C := \begin{bmatrix} \gamma(0) & \gamma(1) & \cdots & \gamma(4l-1) \\ \gamma(1) & \gamma(0) & \cdots & \gamma(4l-2) \\ \vdots & \vdots & \ddots & \vdots \\ \gamma(4l-1) & \gamma(4l-2) & \cdots & \gamma(0) \end{bmatrix} \quad (11)$$

are either  $2l$  of multiplicity 2 or  $4l$  of multiplicity 1.

**Proof.** Let  $c_i$  denote the  $i$ th column of  $C$ . Then, by the defining equation (7),  $c_i = bc_{i-1} - c_{i-2}$  for any  $i \geq 3$ . Consequently, there exists at most two linearly independent columns. Thus  $\text{rank}(C) \leq 2$ , which in turn implies that there exists at most two non-zero eigenvalues  $\lambda_1$  and  $\lambda_2$ . In order to compute  $\lambda_1$  and  $\lambda_2$ , we recall the following identities:

$$\text{tr}(C) = \lambda_1 + \lambda_2 = 4l, \quad (12)$$

$$\text{tr}(C^2) = \lambda_1^2 + \lambda_2^2 = \|C\|_F^2, \quad (13)$$

where  $\|\cdot\|_F$  is the Frobenius norm. If  $\text{rank}(C) = 1$ , then  $\lambda_2 = 0$ , implying the second part of the claim. Suppose then  $\text{rank}(C) = 2$ . Observing that the squared sum of the diagonals is  $4l$  and, for  $j = 1, 2, \dots, 4l-1$ , a term  $\gamma(j)$  appears in  $C$  exactly  $2(4l-j)$  times, we obtain

$$\|C\|_F^2 = 4l + 2 \sum_{j=1}^{4l-1} (4l-j)\gamma(j)^2.$$

Dividing the sum into two parts and using  $\sin^2(x) = 1 - \cos^2(x)$  we have

$$\begin{aligned}
 \|C\|_F^2 &= 4l + 2 \sum_{j=0}^{2l-1} (4l - (2j + 1)) \gamma(2j + 1)^2 + 2 \sum_{j=1}^{2l-1} (4l - 2j) \gamma(2j)^2 \\
 &= 4l + 2 \sum_{j=0}^{2l-1} (4l - (2j + 1)) \sin^2((2j + 1)x) + 2 \sum_{j=1}^{2l-1} (4l - 2j) \cos^2(2jx) \\
 &= 4l + 2 \sum_{j=0}^{2l-1} (4l - (2j + 1)) + 2 \sum_{j=1}^{4l-1} (4l - j) \cos^2(jx) (-1)^j \\
 &= 8l^2 + 4l + 2 \sum_{j=1}^{4l-1} (4l - j) \cos^2(jx) (-1)^j,
 \end{aligned}$$

where in the last equality we have used

$$\sum_{j=0}^{2l-1} (4l - (2j + 1)) = \sum_{j=0}^{2l-1} (4l - 1) - 2 \sum_{j=0}^{2l-1} j = 2l(4l - 1) + 2l(2l - 1) = 4l^2.$$

Now

$$\begin{aligned}
 \sum_{j=1}^{4l-1} (4l - j) \cos^2(jx) (-1)^j &= 2l + \sum_{j=1}^{2l-1} (4l - j) \cos^2(jx) (-1)^j \\
 &\quad + \sum_{j=2l+1}^{4l-1} (4l - j) \cos^2(jx) (-1)^j, \quad (14)
 \end{aligned}$$

where substitution  $j = 4l - t$  yields

$$\sum_{j=2l+1}^{4l-1} (4l - j) \cos^2(jx) (-1)^j = \sum_{t=1}^{2l-1} t \cos^2(tx) (-1)^t. \quad (15)$$

Now (14), (15), and Lemma 3.5 imply

$$\|C\|_F^2 = 8l^2 + 4l + 2 \left( 2l + 4l \sum_{j=1}^{2l-1} \cos^2(jx) (-1)^j \right) = 8l^2.$$

Finally, using (12) and (13) together with  $\|C\|_F^2 = 8l^2$ , we obtain

$$\lambda_1^2 + (4l - \lambda_1)^2 - 8l^2 = 2\lambda_1^2 - 8l\lambda_1 + 8l^2 = (\sqrt{2}\lambda_1 - \sqrt{8}l)^2 = 0.$$

Hence  $\lambda_1 = \lambda_2 = 2l$ . □

We are now ready to prove Theorem 2.1 and Theorem 2.2.

**Proof the Theorem 2.1.** Throughout the proof we denote  $a_2 \equiv a_1 \pmod{2\pi}$  if  $a_2 = a_1 + 2k\pi$  for some  $k \in \mathbb{Z}$ . That is,  $a_1$  and  $a_2$  are identifiable when regarding them as points on the unit circle. By  $a_3 \in (a_1, a_2) \pmod{2\pi}$  we mean that  $a_3 \equiv a \pmod{2\pi}$  for some  $a \in (a_1, a_2)$ .

1. Since  $\arcsin\left(\frac{b}{2}\right) = \frac{k}{l}\frac{\pi}{2}$ , the first claim follows from Proposition 3.1 together with the fact that functions  $\sin(\cdot)$  and  $\cos(\cdot)$  are periodic. In particular, we have  $\gamma(4l + m) = \gamma(m)$  for every  $m \in \mathbb{Z}$ .
2. Denote  $A = \arcsin\left(\frac{b}{2}\right) = r\frac{\pi}{2}$ . By Proposition 3.1,  $mA$  is the corresponding angle for  $\gamma(m)$  on the unit circle. Note first that, due the periodic nature of  $\cos$  and  $\sin$  functions, it suffices to prove the claim only in the case  $M = 0$ . In what follows, we assume that  $m \geq 0$ . We show that the function  $\gamma(m)$ ,  $m \equiv 0 \pmod{4}$  is dense in  $[-1, 1]$ , while a similar argument could be used for other equivalence classes as well. That is, we show that the function  $\cos(mA)$ ,  $m \equiv 0 \pmod{4}$  is dense in  $[-1, 1]$ . Essentially this follows from the observation that, as  $r \notin \mathbb{Q}$ , the function  $m \mapsto \cos(mA)$  is injective. Indeed, if  $\cos(\tilde{m}A) = \cos(mA)$  for some  $\tilde{m}, m \geq 0$ ,  $\tilde{m} \neq m$ , it follows that

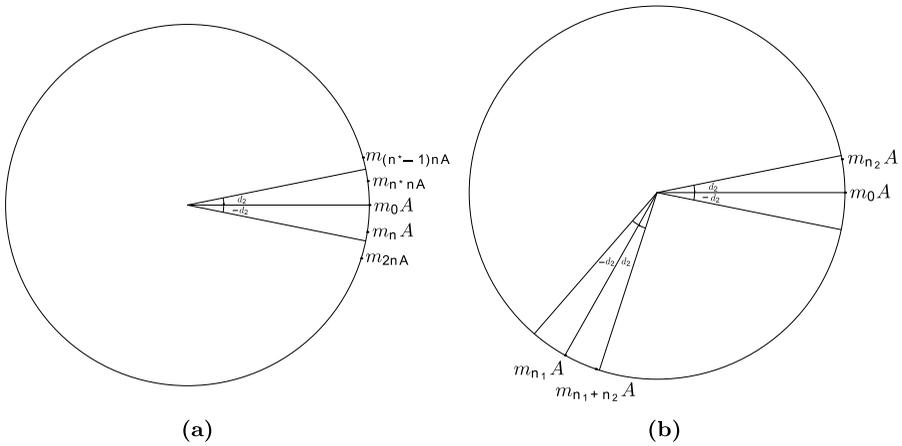
$$\tilde{m}A = \tilde{m}r\frac{\pi}{2} = \pm mr\frac{\pi}{2} + k2\pi = \pm mA + k2\pi \quad \text{for some } k \in \mathbb{Z}.$$

This implies

$$r = \frac{4k}{\tilde{m} \pm m},$$

which contradicts  $r \notin \mathbb{Q}$ . Since  $\cos(mA)$  is injective, it is intuitively clear that  $\cos(mA)$ ,  $m \equiv 0 \pmod{4}$  is dense in  $[-1, 1]$ . For a precise argument, we argue by contradiction and assume there exists an interval  $(c_1, d_1) \subset [-1, 1]$  such that  $\cos(mA) \notin (c_1, d_1)$  for any  $m \equiv 0 \pmod{4}$ . This implies that there exists an interval  $(c_2, d_2) \subset [0, 2\pi]$  such that for every  $m \equiv 0 \pmod{4}$  it holds that  $mA \notin (c_2, d_2) \pmod{2\pi}$ . Without loss of generality, we can assume  $c_2 = 0$  and that for some  $m_0 \equiv 0 \pmod{4}$  we have  $m_0A \equiv 0 \pmod{2\pi}$ . Let  $m_n = m_0 + 4n$  with  $n \in \mathbb{N}$  and denote by  $\lfloor \cdot \rfloor$  the standard floor function. Suppose that for some  $n \in \mathbb{N}$  and  $p_n \in (-d_2, 0)$  we have  $m_nA \equiv p_n \pmod{2\pi}$ . Since by injectivity  $\frac{2\pi}{|p_n|} \notin \mathbb{N}$ , we get  $m_n \lfloor \frac{2\pi}{|p_n|} \rfloor A \in (0, d_2) \pmod{2\pi}$  leading to a contradiction. This implies that for every  $n \in \mathbb{N}$  we have  $m_nA \notin (-d_2, d_2) \pmod{2\pi}$  (for a visual illustration, see Figure 2a). Similarly, assume next that  $m_{n_1}A \equiv p_{n_1} \pmod{2\pi}$  and  $m_{n_1+n_2}A - m_{n_1}A \in (-d_2, d_2) \pmod{2\pi}$ . Then  $m_{n_2}A \in (-d_2, d_2) \pmod{2\pi}$  which again leads to a contradiction (see Figure 2b). This means that for an arbitrary point  $p_n$  on the unit circle such that  $m_nA \equiv p_n \pmod{2\pi}$ , we get an interval  $(p_n - d_2, p_n + d_2)$  (understood as an angle on the unit circle) such that this interval cannot be visited later. As the whole unit circle is covered eventually, we obtain the expected contradiction.

3. Consider first the case  $b = 2 \sin\left(\frac{k}{l}\frac{\pi}{2}\right)$ , where  $\frac{k}{l} \in \mathbb{Q}$ . By Lemma 3.6, the symmetric matrix  $C$  defined by (11) has non-negative eigenvalues, and thus  $C$  is a covariance matrix of some random vector  $(X_0, X_1, \dots, X_{4l-1})$ . Now it suffices to extend this vector to a process  $X = (X_t)_{t \in \mathbb{Z}}$  by the relation  $X_{4l+t} =$



**Fig. 2.** Examples of excluded intervals. In part (a) we have set  $n^* = \lfloor \frac{2\pi}{|p_n|} \rfloor$  and visualized the points on the unit circle corresponding to the steps  $0, n, 2n, (n^* - 1)n$  and  $n^*n$ . In part (b) we have visualised excluded intervals around zero and an angle  $m_{n_1}A$

$X_t$ . Indeed, it is straightforward to verify that  $X$  has the covariance function  $\gamma$ . Assume next  $b = 2 \sin(r \frac{\pi}{2})$ , where  $r \in (0, 1) \setminus Q$ . We argue by contradiction and assume that there exists  $k \in \mathbb{N}$ , and vectors  $t = (t_1, t_2, \dots, t_k)^T \in \mathbb{Z}^k$  and  $a = (a_1, a_2, \dots, a_k)^T \in \mathbb{R}^k$  such that

$$\sum_{i,j=1}^k a_i \gamma(t_i - t_j) a_j = -\epsilon \quad \text{for some } \epsilon > 0,$$

where  $\gamma(\cdot)$  is the covariance function corresponding to the value  $b$ . Since  $Q$  is dense in  $[0, 1]$ , it follows that there exists  $(q_n)_{n \in \mathbb{N}} \subset Q$  such that  $q_n \rightarrow r$ . Denote the corresponding sequence of covariance functions with  $(\gamma_n(\cdot))_{n \in \mathbb{N}}$ . By definition,

$$\sum_{i,j=1}^k a_i \gamma_n(t_i - t_j) a_j \geq 0 \quad \text{for every } n.$$

On the other hand, continuity implies  $\gamma_n(m) \rightarrow \gamma(m)$  for every  $m$ . This leads to

$$\lim_{n \rightarrow \infty} \sum_{i,j=1}^k a_i \gamma_n(t_i - t_j) a_j = \sum_{i,j=1}^k a_i \gamma(t_i - t_j) a_j = -\epsilon$$

giving the expected contradiction. □

**Remark 3.7.** Note that in the periodic case the covariance matrix  $C$  defined by (11) satisfies  $\text{rank}(C) \leq 2$ . Thus, in this case, the process  $(X_t)_{t \in \mathbb{Z}}$  is driven linearly by only two random variables  $Y_1$  and  $Y_2$ . In other words, we have

$$X_t = a_1(t)Y_1 + a_2(t)Y_2, \quad t \in \mathbb{Z},$$

for some deterministic coefficients  $a_1(t)$  and  $a_2(t)$ .

**Proof of Theorem 2.2.** Suppose  $\gamma$  satisfies (7) and  $r(m) \leq r(0)(1 - \epsilon)$  for all  $m \geq M$ . By Theorem 2.1, there exists  $m^* \geq M$  such that

$$\gamma(m^*) \geq \gamma(0) \left(1 - \frac{\epsilon}{2}\right).$$

Furthermore, (7) implies (6) for every  $m$  such that  $\gamma(m) \neq 0$ . Now

$$a_{m^*} = \frac{r(m^*)}{\gamma(m^*)} \leq \frac{r(0)(1 - \epsilon)}{\gamma(0)(1 - \frac{\epsilon}{2})} < \frac{r(0)}{\gamma(0)} = a_0$$

leads to a contradiction. Treating the case  $r(m) \geq -r(0)(1 - \epsilon)$  for all  $m \geq M$  similarly concludes the proof.  $\square$

#### 4 Simulations

In this section we present a simulation study in order to compare the classical Yule–Walker estimator with our quadratic type estimator in the case of an AR(1)-process.

If  $(Z_t)_{t \in \mathbb{Z}}$  is chosen to be white noise in the characterisation (4), the process is an AR(1)-process with  $\phi > 0$  and equations (5) provide natural estimators for the AR(1)-parameter. In this case, it can be verified that the minus sign in (5) is the correct choice whenever  $m \neq 0$  (see the discussion about determining the correct sign based on the ratio  $\frac{r(m)}{\gamma(m)}$  in [8]). If  $m = 0$ , then the discriminant of (5) equals to zero yielding

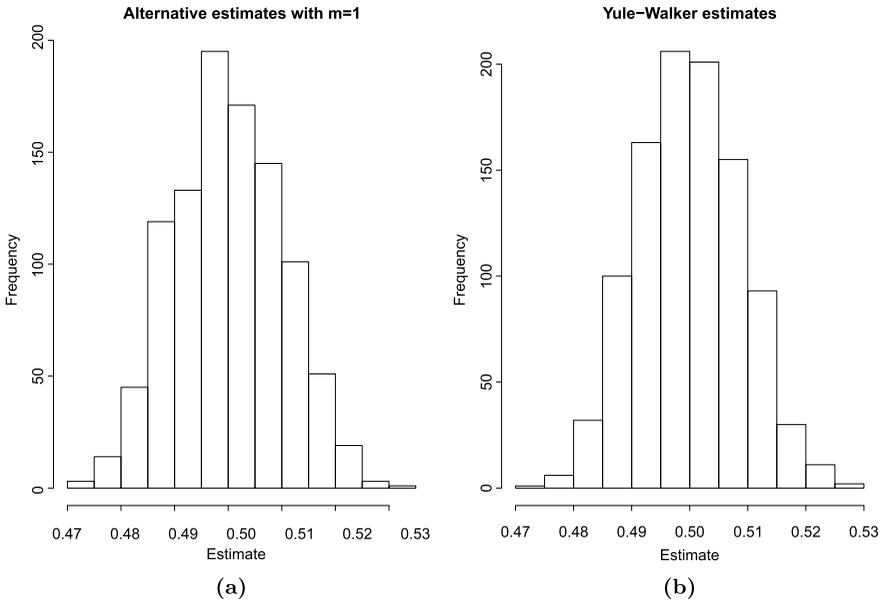
$$\phi = \frac{\gamma(1)}{\gamma(0)}, \tag{16}$$

which is the classical Yule–Walker equation for the model parameter. The same equation is also given by Theorem 2 of [8]. We would also like to point out that, when  $k = 0$  is chosen, the other Yule–Walker equation

$$\gamma(0) = \phi\gamma(1) + r(0)$$

related to AR(1)-processes is given by Equation 7 of [8].

Figure 3 displays histograms comparing efficiencies of the Yule–Walker estimator given by (16) and the alternative estimator given by (5) with  $m = 1$ . We simulated data from an AR(1)-process with  $\phi = 0.5$ . The used sample size and number of iterations were 10000 and 1000, respectively. The sample mean and sample variance of the alternative estimates (3a) are 0.5000318 and 0.0001030415. For classical Yule–Walker estimates (3b), the corresponding sample statistics are 0.4998632 and 7.771528e-05. Thereby it seems that, in this setting, the variances of the two estimators are of the same order. Moreover, the slightly better performance of the Yule–Walker estimator is something that could have been expected. Indeed, the autocovariance function of an AR(1)-process is exponentially decreasing and consequently, the denominator  $\gamma(m)$  acts as an variance increasing factor in the estimators as  $m$  grows. For asymptotic distributions of estimators given by (5) and a more extensive simulation study, we refer to [8].



**Fig. 3.** Classical Yule–Walker estimates of an AR(1)-process and alternative estimates corresponding to the lag  $m = 1$

## 5 Discussion

We have shown (Theorem 2.2) that the estimation method arising out of the AR(1)-characterisation introduced in [8] is applicable except in a very special class of processes. These processes are highly degenerate, namely, they are either driven by two random variables only (Remark 3.7) or their covariance functions can be approximated with covariance functions of such processes (proof of Theorem 2.1 item 3).

The discussed estimation procedure has recently been applied in practice in [5], where we considered a generalized ARCH model with stationary liquidity. As mentioned in [1], the usual approaches of LS and ML methods fail even in the case of liquidity given by the squared increments of a fractional Brownian motion with  $H \neq \frac{1}{2}$ . However, with our method, we were able to derive estimators for the model parameters in the case of liquidity given by a general class of stationary processes. In a more general context, it could be argued that whenever it is possible to derive the maximization problem related, e.g., to ML and QL methods in an adequate way, then these methods provide more efficient estimators. However, deriving ML or QL estimators may turn out to be a difficult task to accomplish. Moreover, unlike our method, these approaches require that the practitioner knows the underlying distribution up to the parameters of interest.

In addition to the above mentioned generalized ARCH model, it would be interesting to study whether our method can be applied in modeling and estimation of different temporal (stationary) models. For example, one could consider GARCH-X models or even integer valued processes such as INAR(1). However, caution should

be taken when comparing different methods and interpreting the results since, in general, the parameters arising out of the AR(1)-characterisation do not coincide with the parameters of the original model.

## References

- [1] Bahamonde, M., Torres, S., Tudor, C.A.: ARCH model with fractional Brownian motion. *Stat. Probab. Lett.* **134**, 70–78 (2018). [MR3758583](#). <https://doi.org/10.1016/j.spl.2017.10.003>
- [2] Brockwell, P.J., Davis, R.A.: *Time Series: Theory and Methods*. Springer (2013). [MR2839251](#)
- [3] Cheridito, P., Kawaguchi, H., Maejima, M.: Fractional Ornstein-Uhlenbeck processes. *Electron. J. Probab.* **8**, 3–14 (2003). [MR1961165](#). <https://doi.org/10.1214/EJP.v8-125>
- [4] Hamilton, J.D.: *Time Series Analysis vol. 2*. Princeton University Press (1994). [MR1278033](#)
- [5] Ilmonen, P., Torres, S., Tudor, C., Viitasaari, L., Voutilainen, M.: On generalized ARCH model with stationary liquidity. arXiv preprint [arXiv:1806.08608](#) (2018)
- [6] Kubilius, K., Mishura, Y., Ralchenko, K.: *Parameter Estimation in Fractional Diffusion Models*. Springer (2018). [MR3752152](#). <https://doi.org/10.1007/978-3-319-71030-3>
- [7] Viitasaari, L.: Representation of stationary and stationary increment processes via Langevin equation and self-similar processes. *Stat. Probab. Lett.* **115**, 45–53 (2016). [MR3498367](#). <https://doi.org/10.1016/j.spl.2016.03.020>
- [8] Voutilainen, M., Viitasaari, L., Ilmonen, P.: On model fitting and estimation of strictly stationary processes. *Mod. Stoch.: Theory Appl.* **4**(4), 381–406 (2017). [MR3739015](#). <https://doi.org/10.15559/17-vmsta91>