

# On the Unique Identification of Continuous-Time Autoregressive Models From Sampled Data

Hagai Kirshner, *Member, IEEE*, Michael Unser, *Fellow, IEEE*, and John Paul Ward

**Abstract**—In this work, we investigate the relationship between continuous-time autoregressive (AR) models and their sampled version. We consider uniform sampling and derive criteria for uniquely determining the continuous-time parameters from sampled data; the model order is assumed to be known. We achieve this by removing a set of measure zero from the collection of all AR models and by investigating the asymptotic behavior of the remaining set of autocorrelation functions. We provide necessary and sufficient conditions for uniqueness of general AR models, and we demonstrate the usefulness of this result by considering particular examples. We further exploit our theory and introduce an estimation algorithm that recovers continuous-time AR parameters from sampled data, regardless of the sampling interval. We demonstrate the usefulness of our algorithm for various Gaussian and non-Gaussian AR processes.

**Index Terms**—Sampling theory, approximation theory, stochastic processes.

## I. INTRODUCTION

CONTINUOUS-TIME autoregressive (AR) models are widely used in astronomy, geology, quantitative finance, control theory and signal processing, to name a few. In practice, the available data is discrete, and one is often required to estimate continuous-time parameters from sampled data. The intertwining relations between the continuous-time model and its sampled version play an important role in such estimation tasks, and studying these relations is the main focus of this work.

1) *Continuous-to-Discrete Mapping*: Within the context of state space representation, point-wise values of continuous-time state variables comply with a stochastic difference equation formulation ([2], Chapter 3.10). The exponential relation between the discrete- and the continuous-time poles can then be utilized for mapping a stochastic difference equation back to its corresponding stochastic differential equation [4], [7], [8], [32]. This relation is invariant to the addition of integer multiples of  $2\pi i$ , and the mapping of [32] avoids ambiguity by restricting the imaginary part of the poles to be in the interval  $[-\pi, \pi]$ . Such a

restriction also appears in [7], [8], which address the problem of embedding a discrete-time ARMA model in a continuous-time one. There is currently no closed-form expression for the zeros of the sampled model [14], [32]. Their asymptotic properties, however, are known for the limiting case of infinitesimally small sampling intervals [36]: the zeros of the sampled model converge to the exponential value of the continuous-domain zeros, and the sampling process introduces additional zeros that converge to the roots of an Euler-Frobenius polynomial. The rate of convergence depends on the model order, and this continuous-to-discrete mapping is useful in cases where the sampling interval is considerably smaller than the characteristic time constants of the model [23]. For larger sampling intervals, however, one can rely on the exponential spline property of the AR autocorrelation function and extract the zeros of the sampled model from the  $z$ -transform of the corresponding exponential B-spline [21].

2) *Algorithms*: The lack of a closed-form expression has led to the development of numerous estimation algorithms. Finite difference approximation of the derivative operator, for example, was shown in [3] to map a continuous-domain AR(2) model to a discrete-domain AR(2) model. The latter is then estimated from the data while including a multiplicative correcting term that accounts for the discretization error. A weighted-finite-difference approximation was considered for the general AR( $p$ ) case in [33]. The choice of weights is based on an optimization procedure that maximizes the rate of convergence; the model fitting is based on a least-squares criterion; and the imaginary parts of the complex poles are restricted to be in the interval  $[-\pi, \pi]$ . Finite difference approximation was also considered in [28] by extending an estimation algorithm that observes the data at an arbitrarily fine resolution, i.e., the estimation algorithm assumes a sufficiently high sampling rate. This method requires bias compensation, too, due to the discrete-domain zero that is introduced by the sampling process. Taking an indirect approach (cf. [22]), the algorithm of [27] fits available sample data of a continuous-time AR(2) process with a discrete-domain ARMA(2,1) model and inverse-transforms the discrete-domain parameters to their continuous-domain AR(2) counterparts. The continuous-domain zero is determined by solving a set of linear equations that stem from the partial fraction decomposition of the continuous-domain power spectrum; the imaginary parts of the poles are restricted to  $[-\pi, \pi]$ . Discrete-domain ARMA models of higher orders were considered in [32]; the estimation method there relies on partial-fraction decomposition and on restricting the imaginary part to  $[-\pi, \pi]$ , as well. Similarly, the algorithms of [13], [39] require a large number of sample values and a sufficiently high sampling rate.

Model ambiguity was also pointed out in [17] using a frequency-domain point of view. Power spectra which occupy high

Manuscript received June 03, 2013; revised October 04, 2013; accepted December 10, 2013. Date of publication January 02, 2014; date of current version February 21, 2014. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Maria Sabrina Greco. This work was supported by the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement no. 267439; and by the Swiss National Science Foundation under Grant 200020-144355.

The authors are with the Biomedical Imaging Group, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland (e-mail: hagai.kirshner@epfl.ch).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TSP.2013.2296879

frequencies may result in incorrect estimation, and the estimation approach of [17] designates a bandwidth of validity for carrying out the estimation task. While no restrictions were imposed on the AR model order  $p$ , the imaginary parts are still assumed to be in  $[-\pi, \pi]$ . The frequency-domain estimator of [16] approximates the autocorrelation function by means of a polynomial spline. While demonstrating improved approximation properties, the polynomial spline model cannot prefer one bandwidth over the other, and the reconstructed autocorrelation function has most of its energy in the baseband, i.e., the algorithm assumes a sufficiently high sampling rate. Non-quantitative restrictions were introduced in [19]. Having the input signal at one's disposal, the identification method is based on filtering both the input and the output signals with a set of predetermined filters. The continuous-domain filtering operations are then approximated by means of finite-differences, and the model parameters are fitted with the available input and output values using regression. Instead of limiting the imaginary part of the AR poles, the work in [19] assumes that there is no other set of continuous-domain poles that gives rise to the same input-output relation; this set, however, is not provided or characterized. The estimator of [34] formulates a likelihood function by means of the discrete-domain innovation process, and the optimization procedure is based on numerical approximation of autocorrelation values. While no restrictions are stated, simulation results describe baseband power spectra only; this implies that the algorithm assumes a sufficiently high sampling rate. The estimator of [4] formulates a likelihood function for general AR( $p$ ) processes, while relying on the state-space representation of the sampled model. The algorithm assumes identifiability (i.e., uniqueness) of the estimated model and convexity of the likelihood function; yet, no proofs are provided in this regard. A maximum likelihood estimator for a continuous-time ARMA(2,1) model was proposed in [8] while restricting the imaginary part of the poles to be in  $(0, \pi)$ .

3) *Evidence That Unique Identification From Sampled Data Is Feasible:* Model ambiguity can be overcome by a multi-rate sampling procedure [12], but the question we are raising in this work is whether there exists a set of continuous-domain AR( $p$ ) models for which the uniform sampling operator introduces no ambiguity. The AR(1) model, for example, has an exponentially decaying autocorrelation function that can be exactly reconstructed from its sampled version [29]. The two continuous-domain complex poles  $-1 \pm 20i$  are mapped upon unit-interval sampling to two discrete-domain poles and to a single discrete-domain zero of 0.1155; the poles  $-1 \pm (20 + 2\pi)i$  are mapped to yet another zero, 0.1201. As a matter of fact, there is no other pair of poles of the form  $-1 \pm (20 + 2\pi k)i$  where  $k = \pm 1, \pm 2, \dots, \pm 10$  that yields the same discrete-domain zero as  $k = 0$  does. Pole ambiguity in such a case can be resolved by shifting the identified continuous-domain poles by an integer multiple of the angular sampling frequency and by comparing the zeros of the sampled model with the zeros of the estimated discrete-domain model ([26], Remark 7). This, however, does not hold true in all cases, and ambiguity arises when continuous-time AR models are mapped to the same discrete-time model, e.g.,  $-1 \pm 2\pi k$ .

A related, yet different, problem is the embedding of discrete-time ARMA models [7], [8], [18]. The problem of em-

bedding is to determine, given a discrete-time process, if there is a continuous-time ARMA process that has the same autocorrelation values at the integers. It requires one to find a set of discrete-time models and a one-to-one mapping to the continuous-time domain that is inverted by the sampling operator. In our uniqueness problem, however, we seek to find a set of continuous-time models and a one-to-one mapping that inverts the sampling operator.

4) *Contribution and Organization of the Paper:* In this work, we revisit the ambiguity assumption of sampled continuous-time AR models by considering two subsets: 1) real poles, 2) real and complex poles with minimal restrictions. We show that the former introduces no ambiguity with respect to all possible AR models, and we provide an explicit criterion for the non-ambiguity property of the latter. In the AR(2) case, for example, uniqueness holds for a pair of complex poles when the imaginary part is an irrational multiple of  $\pi$ , regardless of the sampling interval. To this aim, we introduce two alternative AR representations and derive a closed-form expression for the autocorrelation function. We utilize its special structure—exponential spline and almost periodic properties—for extending our previous results on first and second order processes [37]. Specifically, we provide necessary and sufficient conditions for uniqueness of almost every AR( $p$ ) model. We then demonstrate the usefulness of this result by considering particular values of  $p$ . Furthermore, we exploit our theory and introduce an estimation algorithm that recovers continuous-time AR parameters from sampled data, while imposing no sampling rate constraints. We successfully apply it to Gaussian AR processes, as well as to non-Gaussian AR processes with Lévy-type innovation.

The paper is organized as follows. In Section II, we describe the estimation problem, and in Section III, we introduce alternative representations of AR processes. Section III contains a closed-form expression for the autocorrelation function too. In Section IV, we derive uniqueness results for general trigonometric polynomials, and in Section V, we apply them to AR( $p$ ) models. We then introduce our estimation algorithm in Section VI and demonstrate its performance in Section VII.

## II. THE PROBLEM

We consider the general continuous-time AR( $p$ ) model with parameters vector  $\theta$

$$x(t) = \int_0^\infty h(t - \tau; \theta) w(\tau) d\tau, \quad (1)$$

where  $w(t)$  is a Gaussian or non-Gaussian white noise process. An equivalent representation using a stochastic integral is

$$x(t) = \int_0^\infty h(t - \tau; \theta) dW(\tau), \quad (2)$$

where  $W(t)$  is a Brownian motion (in the Gaussian case) or a more general Lévy process (in the non-Gaussian case) [30]. The connection between both representations is

$$W(t) = \int_0^t w(\tau) d\tau \quad \text{or} \quad w(t) = \frac{dW}{dt}(t). \quad (3)$$

The white Lévy noise or innovation  $w(t)$  is a generalized stochastic process that corresponds to the derivative of  $W(t)$  in the

sense of distributions [15]. The shaping filter  $h(t; \theta)$  is given in the Fourier domain by

$$H(\omega; \theta) = \frac{1}{\prod_{n=1}^p (i\omega - s_n)}, \quad (4)$$

where

$$\theta = (s_1, s_2, \dots, s_p) \in \mathbb{C}^p. \quad (5)$$

The vector of parameters  $\theta$  consists of the poles of  $H(\omega; \theta)$ . The real part of each pole is strictly negative, and complex poles appear in conjugate pairs. Assuming that  $w(t)$  is white with finite variance  $\sigma^2$ , the autocorrelation function of  $x(t)$  is given in the Fourier domain by

$$\begin{aligned} \Phi(\omega; \theta, \sigma^2) &= \sigma^2 |H(\omega; \theta)|^2 \\ &= \frac{\sigma^2}{\prod_{n=1}^p (i\omega - s_n)(-i\omega - s_n)}, \end{aligned} \quad (6)$$

where we assume  $t \gg 0$  for approximate stationarity. In this work, we are interested in estimating  $\theta$  from a sampled version of  $x(t)$  while imposing no restrictions on the sampling interval. We first investigate whether two continuous-time AR( $p$ ) models can yield the same discrete-time model and then introduce an estimation algorithm. Without loss of generality, we assume a unit sampling interval and  $\sigma^2 = 1$ . For simplicity, we omit the notation  $\theta$  when possible.

### III. REVISITING AR REPRESENTATION

#### A. The Vector of Parameters

In this section, we introduce two new parameter vectors describing AR processes. The proposed representations are based on decay rates, modulation, and pole multiplicities. We use them to give an explicit formula for the autocorrelation functions of AR processes and to place a measure on the space of AR( $p$ ) processes. These representations are intended to make the properties mentioned above more accessible. We therefore use all three representations ( $\theta$  and the proposed  $\tilde{\theta}, \bar{\theta}$ ), as each has its particular use.

As a starting point, we write

$$\theta = (s_1, \dots, s_{2m}, s_{2m+1}, \dots, s_p), \quad (7)$$

where the first  $2m$  poles are complex, and conjugate pairs appear sequentially. Additionally, for a given complex pair, we require the one with positive imaginary part to be listed first (Table I). Our first alternative representation is based on decay rates and modulation values. It extends the AR representation of [37] in the following manner,

$$\tilde{\theta} = (a_1, b_1, a_2, b_2, \dots, a_m, b_m, a_{m+1}, \dots, a_{p-m}), \quad (8)$$

where  $a_1, \dots, a_{p-m}$  are the strictly negative real parts of the poles, and  $b_1, \dots, b_m$  are the strictly positive imaginary parts. The vector  $\tilde{\theta}$  is a point in  $\mathbb{R}^p$  and this identification can be made unique by imposing a dictionary-type ordering:

- $0 > a_1 \geq a_2 \geq \dots \geq a_m$ ;
- $0 > a_{m+1} \geq a_{m+2} \geq \dots \geq a_{p-m}$ ;
- if  $a_k = a_{k+1}$ , then  $b_{k+1} \geq b_k$ .

TABLE I  
THE RELATION BETWEEN  $\theta$  AND  $\tilde{\theta}$

Complex Poles		Real Poles	
$s_1$	$= a_1 + ib_1$	$s_{2m+1}$	$= a_{m+1}$
$s_2$	$= a_1 - ib_1$	$\vdots$	$\vdots$
$\vdots$	$\vdots$	$s_p$	$= a_{p-m}$
$s_{2m-1}$	$= a_m + ib_m$		
$s_{2m}$	$= a_m - ib_m$		

$a_1, \dots, a_{p-m}$  are strictly negative,  $b_1, \dots, b_m$  are strictly positive,  $m$  is the number of complex poles and  $p$  is the total number of poles.

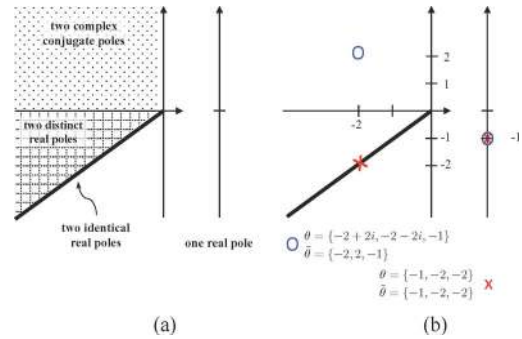


Fig. 1. Graphical representation of an AR(3) process: The representation is composed of a plane and a line, as shown in (a). The plane is used for describing two poles, either two reals or a complex conjugate pair, and the line axis describes a single real pole. A numerical example is given in (b). The representation  $\tilde{\theta}$  provides a means of measuring subsets of AR models.

The difference in sign between the  $a$ 's and  $b$ 's allows us to distinguish the two types of poles, so that there is no confusion. Our graphical representation of  $\tilde{\theta}$  is based on multiple copies of  $\mathbb{R}^2$  and  $\mathbb{R}$ . In the case  $p$  is even, we use  $p/2$  copies of  $\mathbb{R}^2$ . In the case  $p$  is odd, we use  $(p-1)/2$  copies of  $\mathbb{R}^2$  and one copy of  $\mathbb{R}$ .

*Example 1:* The AR(3) model has two possible constellations: (1) three real poles, e.g.,  $\tilde{\theta} = (-1, -2, -2)$ , and (2) a complex pair together with a real pole, e.g.,  $\tilde{\theta} = (-2, 2, -1)$  (Fig. 1).

The second new parameter vector  $\bar{\theta}$  indicates multiplicities of poles and will be used for obtaining an explicit formula of autocorrelation functions. The vector

$$\bar{\theta} = (\bar{s}_1, m_1, \bar{s}_2, m_2, \dots, \bar{s}_L, m_L) \quad (9)$$

will be used to denote an AR( $p$ ) process composed of the distinct poles  $\bar{s}_l$ . The multiplicity of a pole  $\bar{s}_l$  is represented by the number  $m_l$  following it in the parameter vector.

*Definition 1:* The collection of all parameter vectors  $\theta$  corresponding to AR( $p$ ) models is  $\Omega(p)$ . This is also the collection of all parameter vectors  $\tilde{\theta}$  or  $\bar{\theta}$ .

With the alternative representation  $\tilde{\theta}$  in mind, we view  $\Omega(p)$  as a subset of  $\mathbb{R}^p$ , and the Lebesgue measure on  $\mathbb{R}^p$  provides a means of measuring subsets of autocorrelation functions.

### B. The Autocorrelation Function

The rational form of  $H(\omega; \theta)$  is known to yield an autocorrelation function that is a sum of Hermitian symmetric exponentials, as the result of a decomposition in partial fractions. The explicit formula is obtained as follows.

*Proposition 1:* Let  $\bar{\theta} = (\bar{s}_1, m_1, \dots, \bar{s}_L, m_L)$  be a parameter vector in  $\Omega(p)$ , and let  $\lambda_l = \bar{s}_l^2$ . A formula for the autocorrelation function is given by

$$\varphi(t; \bar{\theta}) = (-1)^p \sum_{\ell=1}^L e^{-\lambda_\ell^{1/2}|t|} \sum_{n=1}^{m_\ell} \sum_{k=0}^{n-1} d_{\ell,n,k} |t|^{n-1-k}, \quad (10)$$

where

$$P(\xi) = \prod_{l=1}^L (\xi - \lambda_l)^{m_l} \quad (11)$$

$$c_{l,n} = \lim_{\xi \rightarrow \lambda_l} \frac{1}{(m_l - n)!} \frac{d^{m_l - n}}{d\xi^{m_l - n}} \left( \frac{(\xi - \lambda_l)^{m_l}}{P(\xi)} \right) \quad (12)$$

$$d_{l,n,k} = \frac{(-1)^n c_{l,n} (n-1+k)!}{(n-1)! k! (n-1-k)! (2\lambda_l^{1/2})^{n+k}}, \quad (13)$$

and  $\lambda_l^{1/2} \in \mathbb{C}$  denotes the principal square root of  $\lambda_l$ .

*Proof:* From (6), we see that

$$\Phi(\omega; \theta) = \frac{1}{\prod_{n=1}^p (\omega^2 + s_n^2)} \quad (14)$$

is a Green's function for an operator whose symbol is a polynomial in  $\omega^2$ . A general multi-dimensional formula for such functions is given in [6]. Here, we apply that more general result to (6) in order to obtain the representation above. The corresponding expression for  $\varphi(t; \bar{\theta})$  is

$$\begin{aligned} \varphi(t; \bar{\theta}) &= \frac{(-1)^p}{\pi^{\frac{1}{2}}} \\ &\times \sum_{l=1}^L \sum_{n=1}^{m_l} \frac{c_{l,n} \left( \frac{|t|}{2\lambda_l^{1/2}} \right)^{n-\frac{1}{2}} K_{\frac{1}{2}-n} \left( \lambda_l^{\frac{1}{2}} |t| \right)}{(-1)^n (n-1)!}. \end{aligned} \quad (15)$$

As the order of the Bessel function  $K$  is half of an odd integer, an explicit formula is known, and using properties of the Bessel function, we have

$$K_{1/2-n}(z) = \left( \frac{\pi}{2z} \right)^{1/2} e^{-z} \sum_{k=0}^{n-1} \frac{(n-1+k)!}{k!(n-1-k)!(2z)^k}, \quad (16)$$

cf. ([38], p. 80). Inserting (16) into (15), we obtain (10).  $\square$

*Example 2:* Let  $\bar{\theta} = (a, 3)$ , with  $a < 0$ . For this set of real poles, we have  $P(\xi) = (\xi - a^2)^3$ , so  $L = 1$ ,  $\lambda_1 = a^2$ , and  $m_1 = 3$ . This leads to

$$\varphi(t; \bar{\theta}) = \frac{-1}{16a^5} e^{a|t|} (a^2|t|^2 - 3a|t| + 3). \quad (17)$$

*Example 3:* Let  $\bar{\theta} = (-1+i, 1, -1-i, 1, -2, 1)$ . For this set of poles, we have  $\lambda_1 = -2i$ ,  $\lambda_2 = 2i$ , and  $\lambda_3 = 4$ . Therefore,

$$\varphi(t; \bar{\theta}) = \frac{1}{160} e^{-|t|} (2 \cos(t) + 6 \sin(|t|)) + \frac{1}{80} e^{-2|t|}. \quad (18)$$

*Example 4:* Let  $\bar{\theta} = (s_1, 1, s_2, 1, \dots, s_p, 1)$  consist of non-repeating poles. Then,

$$\varphi(t; \bar{\theta}) = (-1)^p \sum_{l=1}^p d_l e^{s_l |t|} \quad (19)$$

$$d_l = \frac{1}{2s_l \prod_{n=1, n \neq l}^p (s_l - s_n)(s_l + s_n)}, \quad (20)$$

### IV. ASYMPTOTIC PROPERTIES

We characterize the relationship between a continuous-time AR model and its sampled version in terms of the following uniqueness property. When we restrict to the collection of unique continuous-time models, there is a one-to-one mapping into the discrete domain.

*Definition 2:* Two parameter vectors  $\theta_1, \theta_2 \in \Omega(p)$  are equivalent if there exists  $\alpha < 0$  such that  $\varphi(n; \theta_1) + \alpha \cdot \varphi(n; \theta_2) = 0$  for all  $n \in \mathbb{Z}$ . If  $\theta_1$  is not equivalent to any distinct  $\theta_2$ , then it is unique.

The uniqueness property is related to linear combinations of autocorrelation functions. Closed form expressions for the autocorrelation functions are given in Proposition 1, and in the examples following the proposition, we see how the structure of the autocorrelation functions depends on the poles. In (10), the real parts of the parameters  $\lambda_l^{1/2}$  determine exponentially decaying terms, while the imaginary parts determine periods of trigonometric polynomials. In the case of multiple poles, a polynomial term multiplies the complex exponential. Linear combinations of these functions also have the same basic structure. Therefore, in this section we consider a general class of functions that contains such linear combinations as a subset, and we characterize the asymptotic behavior of this general collection of functions.

*Definition 3:* We denote by  $X$  the class of functions of the form

$$\sum_{l=1}^L \sum_{m=0}^{M_l} T_{l,m}(|t|) |t|^m e^{a_l |t|}, \quad (21)$$

where  $0 > a_1 > a_2 > \dots > a_L$  and each  $T_{l,m}$  is a trigonometric function. The latter is a finite sum of the form

$$T(t) = \gamma_0 + \sum_{k=1}^K \gamma_k \cos(b_k t) + \nu_k \sin(b_k t), \quad (22)$$

with real  $a$ 's,  $b$ 's,  $\gamma$ 's, and  $\nu$ 's.

In the previous definition,  $T(t)$  is an almost periodic function, which means that for any  $\epsilon > 0$  there are  $\epsilon$ -almost-periods  $\tau$  such that

$$\sup_t |f(t - \tau) - f(t)| < \epsilon, \quad (23)$$

cf. [20]. The space of real valued almost periodic functions on  $\mathbb{R}$  is denoted by  $\text{AP}(\mathbb{R}, \mathbb{R})$ , and the building blocks of this space are the trigonometric polynomials, where the frequencies are allowed to be arbitrary, as in  $T(t)$ . Of particular interest is the fact that uniform samples of almost periodic functions lie in the normed space of almost periodic sequences  $\text{AP}(\mathbb{Z}, \mathbb{R})$  (cf. ([9], Proposition 3.35) and Appendix A), and we exploit this fact to verify uniqueness.

*Definition 4:* ([9], Definition 3.8) The map  $f : \mathbb{Z} \rightarrow \mathbb{R}$  is called almost periodic if to any  $\epsilon > 0$  there corresponds a positive integer  $N(\epsilon)$  with the property that in any set of  $N(\epsilon)$  consecutive integers there exists an integer  $p$  such that

$$|f(n+p) - f(n)| < \epsilon, \quad n \in \mathbb{Z}. \quad (24)$$

*Definition 5:* ([9], pp. 94–95) For any integer  $n$ , the mean value of  $f \in \text{AP}(\mathbb{Z}, \mathbb{R})$  is

$$M(f) = \lim_{k \rightarrow \infty} \frac{f(n+1) + f(n+2) + \cdots + f(n+k)}{k}. \quad (25)$$

Note that, in this definition, we are free to choose any integer  $n$ , and the limit will be independent of this choice. A norm for  $\text{AP}(\mathbb{Z}, \mathbb{R})$  is given by

$$\|f\|_{\text{AP}(\mathbb{Z}, \mathbb{R})}^2 = M(|f|^2). \quad (26)$$

*Theorem 1:* If  $f \in X$  and  $f(n) = 0$  for all integers  $n$ , then the functions  $T_{l,m}(|t|)$  must also satisfy  $T_{l,m}(|n|) = 0$ .

*Proof:* Suppose that there are terms  $T_{l,m}(|n|) |n|^m e^{a_l |n|}$  of  $f(n)$  for which  $T_{l,m}(|n|)$  is not the zero sequence. Taking into account only the monomial and exponential portions of these terms, let  $|n|^{m_0} e^{a_{l_0} |n|}$  be the one with slowest decay as  $n \rightarrow \infty$ .

Since  $|n|^{m_0} e^{a_{l_0} |n|}$  decays the slowest and the functions  $|T_{l,m}(|n|)|$  are bounded above by a constant, the remaining terms must be asymptotically bounded above by a constant multiple of  $|n|^{m_0-1} e^{a_{l_0} |n|}$ . Therefore the sum of the remaining terms must satisfy the same bound.

As  $f(n) = 0$  for all integer values  $n$ , the slowest decaying term  $T_{l_0, m_0}(|n|) |n|^{m_0} e^{a_{l_0} |n|}$  must be canceled by the sum of the remaining terms. Therefore,

$$\left| T_{l_0, m_0}(|n|) |n|^{m_0} e^{a_{l_0} |n|} \right| \leq C \left| |n|^{m_0-1} e^{a_{l_0} |n|} \right| \quad (27)$$

$$\left| T_{l_0, m_0}(|n|) \right| \leq \frac{C}{|n|} \quad (28)$$

for some constant  $C > 0$ . It then follows that

$$\| \{ T_{l_0, m_0}(| \cdot |) \} \|_{\text{AP}(\mathbb{Z}, \mathbb{R})}^2 = \lim_{k \rightarrow \infty} \frac{\sum_{l=1}^k |T_{l_0, m_0}(|l|)|^2}{k} \quad (29)$$

$$\leq \lim_{k \rightarrow \infty} C^2 \frac{\sum_{l=1}^k l^{-2}}{k} \quad (30)$$

$$= 0 \quad (31)$$

Since  $\{T_{l_0, m_0}(|n|)\}$  is an almost periodic sequence with norm 0, it must be identically 0.  $\square$

The value of Theorem 1 is that it essentially allows us to compare functions from  $X$  in a segmented fashion, i.e., according to decay rates. For example, suppose  $\varphi$  is the autocorrelation function for an  $\text{AR}(p)$  process and assume it contains a term  $T(|t|) |t|^m e^{a|t|}$ , where  $T(|n|) |n|^m$  is not identically 0. Then  $\varphi$  cannot be equivalent to any autocorrelation function that lacks a term with similar decay.

## V. UNIQUENESS PROPERTIES

This section has two main results: 1) We prove uniqueness of AR models with real poles only. 2) We derive explicit uniqueness criteria for a large subset of AR models,  $\Omega(p)^* \subset \Omega(p)$ ,

whose complement has Lebesgue measure 0. Finally, as an example, we use the criteria to prove uniqueness of  $\text{AR}(3)$  models and to numerically validate uniqueness of some  $\text{AR}(4)$  models.

### A. Real Poles

*Lemma 1:* Consider two functions  $f_1$  and  $f_2$  of the form

$$\sum_{k=1}^K p_k(|t|) e^{a_k |t|} \quad (32)$$

where each  $p_k$  is a nonzero polynomial and  $a_k < 0$ . If there is a linear combination  $(\gamma_1 \gamma_2 \neq 0)$

$$\gamma_1 f_1 + \gamma_2 f_2 =: f \quad (33)$$

such that  $f(n) = 0$  for all integers  $n$ , then  $f$  is identically 0.

*Proof:* Any linear combination  $f$  must be of the form of (32), where each  $p_k$  is a (possibly zero) polynomial. By Theorem 1, each polynomial  $p_k$  must satisfy  $p_k(n) = 0$  for all integers  $n$ . Therefore all of the polynomials  $p_k$  are 0, and hence  $f$  is identically 0.  $\square$

*Theorem 2:* The elements of  $\Omega(p)$  that are composed entirely of real poles are unique.

*Proof:* First, suppose  $\tilde{\theta}_1, \tilde{\theta}_2 \in \Omega(p)$  are composed entirely of real poles and that they are equivalent. Then there is a linear combination  $(\gamma \neq 0)$

$$\varphi(\cdot; \tilde{\theta}_1) + \gamma \varphi(\cdot; \tilde{\theta}_2) =: f, \quad (34)$$

such that  $f(n) = 0$  for all integers  $n$ . Then Lemma 1 implies that  $f$  is identically 0, i.e.,  $\theta_1 = \theta_2$ .

Now, it suffices to show that any  $\tilde{\theta}_1$ , composed entirely of real poles, is not equivalent to any  $\tilde{\theta}_2 \in \Omega(p)$  that contains a complex pair of poles. Let us write

$$\tilde{\theta}_1 = (a_1, a_2, \dots, a_p) \quad (35)$$

$$\tilde{\theta}_2 = (\alpha_1, \beta_1, \dots, \alpha_\mu, \beta_\mu, \alpha_{\mu+1}, \dots, \alpha_{p-\mu}) \quad (36)$$

for  $1 \leq \mu \leq p/2$ . Furthermore, suppose that the set  $\{a_j, j = 1, \dots, p\}$  is equal to the set of distinct elements  $\{c_l : l = 1, \dots, L_1\}$  and that the set  $\{\alpha_j, j = 1, \dots, p - \mu\}$  is equal to the set of distinct elements  $\{d_l : l = 1, \dots, L_2\}$ . Then the autocorrelation functions are of the form

$$\varphi(t; \tilde{\theta}_1) = \sum_{l=1}^{L_1} f(t; c_l) e^{c_l |t|} \quad (37)$$

$$\varphi(t; \tilde{\theta}_2) = \sum_{l=1}^{L_2} g(t; d_l) e^{d_l |t|}. \quad (38)$$

For each decay rate, we determine the growth of the sequences  $f(n; c_l)$  and  $g(n; d_l)$ . In particular, for  $l = 1, \dots, L_1$ , let  $m(\tilde{\theta}_1; c_l) \geq 0$  denote the smallest integer such that

$$f(n; c_l) = \mathcal{O}\left(n^{m(\tilde{\theta}_1; c_l)-1}\right) \quad \text{as } n \rightarrow \infty. \quad (39)$$

Similarly, for  $l = 1, \dots, L_2$  let  $m(\tilde{\theta}_2; d_l) \geq 0$  denote the smallest integer such that

$$g(n; d_l) = \mathcal{O}\left(n^{m(\tilde{\theta}_2; d_l)-1}\right) \quad \text{as } n \rightarrow \infty. \quad (40)$$

Now, since  $\tilde{\theta}_1$  is composed of real poles

$$\sum_{l=1}^{L_1} m(\tilde{\theta}_1; c_l) = p, \quad (41)$$

and since  $\tilde{\theta}_2$  contains a complex pair of poles,

$$\sum_{l=1}^{L_2} m(\tilde{\theta}_2; d_l) < p. \quad (42)$$

Hence, there must be a decay rate  $a_{l_0}$  such that either

- $a_{l_0} \notin \{\alpha_l, l = 1 \dots p - m\}$
- or  $a_{l_0} \in \{\alpha_l, l = 1 \dots p - m\}$  and  $m(\tilde{\theta}_2; a_{l_0}) < m(\tilde{\theta}_1; a_{l_0})$ .

Suppose the first condition holds (i.e.,  $\tilde{\theta}_1$  has a decay rate that does not appear in  $\tilde{\theta}_2$ ), then Theorem 1 implies that  $\tilde{\theta}_1$  is not equivalent to  $\tilde{\theta}_2$ . If the second condition holds (e.g., the multiplicities of poles differ), then for the decay rate  $a_{l_0}$ ,  $f(n; a_{l_0})$  is growing faster than  $g(n; a_{l_0})$ , so again Theorem 1 implies  $\tilde{\theta}_1$  is not equivalent to  $\tilde{\theta}_2$ .  $\square$

### B. Real and Complex Poles With Minimal Restrictions

*Definition 6:* Let  $\Omega(p)^*$  be the collection of parameter vectors  $\tilde{\theta} \in \Omega(p)$  satisfying:

- $a_{k_1} \neq a_{k_2}$  for  $k_1 \neq k_2$ ;
- each  $b_k$  is an irrational multiple of  $\pi$ .

*Theorem 3:* As a subset of  $\mathbb{R}^p$ , the complement of  $\Omega(p)^*$  in  $\Omega(p)$  has Lebesgue measure 0.

*Proof:* Suppose there exists  $\tilde{\theta} \in \Omega(p)$  that is not in  $\Omega(p)^*$ .

Then there are two possibilities:

- $a_{k_1} = a_{k_2}$  for some  $k_1 \neq k_2$ ;
- or there exists  $1 \leq k_0 \leq p$  such that  $b_{k_0} \in \mathbb{Q}$ .

In the first case,  $\tilde{\theta}$  is contained in a finite union of hyperplanes in  $\mathbb{R}^p$

$$\bigcup_{j \neq k} \{y \in \mathbb{R}^p : y_j = y_k\}. \quad (43)$$

In the second case,  $\tilde{\theta}$  is contained in a countable union of hyperplanes in  $\mathbb{R}^p$

$$\bigcup_{1 \leq k \leq m, q \in \mathbb{Q}} \{y \in \mathbb{R}^p : y_{2k} = q\}. \quad (44)$$

In both cases, we arrive at the conclusion that the complement of  $\Omega(p)^*$  has measure 0. Furthermore, as we are excluding  $p - 1$  dimensional hyperplanes, the Hausdorff dimension of the complement is at most  $p - 1$ .  $\square$

*Lemma 2:* If an admissible vector  $\tilde{\theta}_1 \in \Omega(p)^*$  is equivalent to a vector  $\tilde{\theta}_2 \in \Omega(p)$ , then  $\tilde{\theta}_2$  must have the same number of complex pairs of poles as  $\tilde{\theta}_1$ . Furthermore, the complex pairs should exist at the same decay rates.

*Proof:* Suppose  $\tilde{\theta}_1 \in \Omega(p)^*$  is equivalent to  $\tilde{\theta}_2 \in \Omega(p)$

$$\tilde{\theta}_1 = (a_1, b_1, \dots, a_m, b_m, a_{m+1}, \dots, a_{p-m}) \quad (45)$$

$$\tilde{\theta}_2 = (\alpha_1, \beta_1, \dots, \alpha_\mu, \beta_\mu, \alpha_{\mu+1}, \dots, \alpha_{p-\mu}), \quad (46)$$

for  $\mu < m$ . Then there is a decay rate  $a_{l_0}$ , with  $l_0 \leq m$ , such that  $a_{l_0} \notin \{\alpha_l : l = 1, \dots, \mu\}$ . Due to the fact that  $b_{l_0}$  is

an irrational multiple of  $\pi$ , Theorem 1 implies that no linear combination of  $\varphi(n; \tilde{\theta}_1)$  and  $\varphi(n; \tilde{\theta}_2)$  can be identically 0. This is a contradiction.

Suppose  $\tilde{\theta}_1$  is equivalent to a  $\tilde{\theta}_2$  for  $\mu > m$ . Then  $\varphi(\cdot; \tilde{\theta}_1)$  consists of terms with non-zero samples at  $p - m$  distinct exponential decay rates, while  $\varphi(\cdot; \tilde{\theta}_2)$  can have terms with non-zero samples for at most  $p - m - 1$  distinct decay rates. By Theorem 1, this is a contradiction.  $\square$

*Lemma 3:* Suppose

$$\tilde{\theta}_1 = (a_1, b_1, \dots, a_m, b_m, a_{m+1}, \dots, a_{p-m}) \in \Omega(p)^* \quad (47)$$

is equivalent to

$$\tilde{\theta}_2 = (a_1, \beta_1, \dots, a_m, \beta_m, \alpha_{m+1}, \dots, \alpha_{p-m}) \in \Omega(p). \quad (48)$$

Then  $a_l = \alpha_l$  for all  $l$ .

*Proof:* Note that  $\varphi(\cdot; \tilde{\theta}_1)$  has the form

$$\varphi(t; \tilde{\theta}_1) = \sum_{l=1}^{p-m} f_l(t) e^{a_l |t|}, \quad (49)$$

and since each  $b_l$  is an irrational multiple of  $\pi$ , the sample values  $f_l(n)$  cannot be identically zero. Therefore Theorem 1 implies  $a_l = \alpha_l$  for  $l > m$ .  $\square$

*Proposition 2:* The autocorrelation function for every  $\theta \in \Omega(p)^*$  can be written as a finite sum of the form

$$\varphi(t; \theta) = \sum_l e^{a_l |t|} \left( \gamma_l e^{ib_l |t|} + \gamma_l^* e^{-ib_l |t|} \right), \quad (50)$$

for real numbers  $a_l, b_l$  and complex  $\gamma_l$ .

*Lemma 4:* Let  $f_1$  and  $f_2$  be functions of the form

$$f_1(t) = \gamma_1 e^{ib|t|} + \gamma_1^* e^{-ib|t|} \quad (51)$$

$$f_2(t) = \gamma_2 e^{i\beta|t|} + \gamma_2^* e^{-i\beta|t|}, \quad (52)$$

where  $\gamma_1, \gamma_2$  are non-zero complex numbers,  $b, \beta$  are positive real numbers, and  $b$  is an irrational multiple of  $\pi$ . If  $f_1(n) = f_2(n)$  for all non-negative integers  $n$ , then there is an integer  $k$  such that one of the following holds:

- $\gamma_1 = \gamma_2$  and  $\beta = b + 2\pi k$ ;
- $\gamma_1 = \gamma_2^*$  and  $\beta = -b + 2\pi k$ .

*Proof:* The proof of Lemma 4 follows the same lines as the proof of uniqueness of AR(2) processes, as discussed in ([37], Theorem 2: Case F).

We define

$$F(x, y) := (\gamma_1 e^{2\pi i x} + \gamma_1^* e^{-2\pi i x}) - (\gamma_2 e^{2\pi i y} + \gamma_2^* e^{-2\pi i y}), \quad (53)$$

which satisfies the conditions of Lemma 5. The lemma is then valid due to Lemma 5, Lemma 6, and Lemma 7 from Appendix B.  $\square$

Notice that the autocorrelation function corresponding to an element of  $\Omega(p)^*$  is relatively simple since the exponential decay rates are distinct. Lemma 3 implies that any process equivalent to an  $\Omega(p)^*$  process also has this simplified form.

**Proposition 3:** A process of the form (47) can be represented as

$$\begin{aligned}\varphi(t; \tilde{\theta}_1) &= \sum_{l=1}^m e^{a_l |t|} \left( \gamma_l e^{i b_l |t|} + \gamma_l^* e^{-i b_l |t|} \right) + \sum_{l=m+1}^{p-m} \gamma_l e^{a_l |t|} \\ &= \sum_{l=1}^m 2e^{a_l |t|} (\Re(\gamma_l) \cos(b_l |t|) - \Im(\gamma_l) \sin(b_l |t|)) \\ &\quad + \sum_{l=m+1}^{p-m} \gamma_l e^{a_l |t|}\end{aligned}$$

where for  $l \leq m$

$$\begin{aligned}\gamma_l^{-1} &:= -8i(a_l + i b_l) a_l b_l \prod_{l' > m} ((a_l + i b_l)^2 - a_{l'}^2) \\ &\quad \cdot \prod_{l' \neq l, l' \leq m} ((a_l + i b_l)^2 - (a_{l'} + i b_{l'})^2) \\ &\quad \times ((a_l + i b_l)^2 - (a_{l'} - i b_{l'})^2)\end{aligned}$$

and for  $l > m$

$$\begin{aligned}\gamma_l^{-1} &:= -2a_l \prod_{l' \neq l, l' > m} (a_l^2 - a_{l'}^2) \\ &\quad \cdot \prod_{l' \leq m} (a_l^2 - (a_{l'} + i b_{l'})^2) (a_l^2 - (a_{l'} - i b_{l'})^2).\end{aligned}$$

**Theorem 4:** Let  $\tilde{\theta}_1 \in \Omega(p)^*$  be defined by (47), and suppose that  $\theta_2 \in \Omega(p)$  is equivalent to  $\tilde{\theta}_1$ .

- i.  $\theta_2 \in \Omega(p)^*$
- ii. Let the autocorrelation functions  $\varphi(t; \tilde{\theta}_1)$  and  $\varphi(t; \tilde{\theta}_2)$  be defined by coefficients  $\gamma_l$  and  $\gamma'_l$ , respectively, as in Proposition 3. Then there is a positive number  $\sigma$  such that

$$\sigma^2 \gamma_l = \gamma'_l \quad \text{or} \quad \sigma^2 \gamma_l^* = \gamma'_l \quad (54)$$

for  $l = 1, \dots, p - m$ .

*Proof:* Since  $\theta_2$  is equivalent to  $\tilde{\theta}_1$ , it is defined as

$$(a_1, \beta_1, \dots, a_m, \beta_m, a_{m+1}, \dots, a_{p-m}) \in \Omega(p) \quad (55)$$

by Lemma 3. Furthermore, equivalence implies that there exists  $\sigma > 0$  such that  $\varphi(n; \tilde{\theta}_1) = \sigma^2 \varphi(n; \tilde{\theta}_2)$  for each non-negative integer  $n$ . From Theorem 1, we see that

$$\gamma_l e^{i b_l n} + \gamma_l^* e^{-i b_l n} = (\sigma^2 \gamma'_l) e^{i \beta_l n} + (\sigma^2 \gamma'_l)^* e^{-i \beta_l n} \quad (56)$$

for  $l \leq m$  and  $\gamma_l = \sigma^2 \gamma'_l$  for  $l > m$ . Lemma 4 then implies that  $\gamma_l = \sigma^2 \gamma'_l$  or  $\gamma_l^* = \sigma^2 \gamma'_l$  for every  $l$ .

Lemma 4 also guarantees that there are integers  $k_l$  such that  $\beta_l = \pm b_l + 2\pi k_l$ , so  $\tilde{\theta}_2 \in \Omega(p)^*$ .  $\square$

Theorem 4 provides criteria for uniqueness of general AR models in terms of systems of polynomial equations. In particular, in (54) there are two possibilities for each complex coefficient. Therefore, we obtain  $2^m$  systems of equations, which we denote  $E_{\mathbf{k}}$ ; the systems are indexed by a multi-integer  $\mathbf{k} \in \{0, 1\}^m$ . We use these systems of equations to give necessary and sufficient conditions for uniqueness of AR models in  $\Omega(p)^*$ .

**Definition 7:** Suppose  $p$  is a positive integer, and let  $m$  be a non-negative integer not greater than  $p/2$ . For  $\tilde{\theta} = (a_1, b_1, \dots, a_m, b_m, a_{m+1}, \dots, a_{p-m}) \in \mathbb{R}^p$  and  $\gamma_l$  defined as in Proposition 3, we define the polynomials

$$q_l^{\text{re}}(\tilde{\theta}) := \Re(\gamma_l^{-1}), \quad q_l^{\text{im}}(\tilde{\theta}) := \Im(\gamma_l^{-1}) \quad (57)$$

for  $l \leq m$ , and for  $l > m$  we define

$$q_l(\tilde{\theta}) := \gamma_l^{-1}. \quad (58)$$

**Definition 8:** Let  $\tilde{\theta}_1, \tilde{\theta}_2 \in \mathbb{R}^p$  be defined as

$$\tilde{\theta}_1 := (a_1, b_1, \dots, a_m, b_m, a_{m+1}, \dots, a_{p-m}) \quad (59)$$

$$\tilde{\theta}_2 := (a_1, \beta_1, \dots, a_m, \beta_m, a_{m+1}, \dots, a_{p-m}). \quad (60)$$

For each  $\mathbf{k} = (k_1, \dots, k_m) \in \{0, 1\}^m$ , we define the system of equations

$$\begin{aligned}E_{\mathbf{k}} &:= \left\{ q_l^{\text{re}}(\tilde{\theta}_1) - \sigma^2 q_l^{\text{re}}(\tilde{\theta}_2) = 0, \right. \\ &\quad \left. q_l^{\text{im}}(\tilde{\theta}_1) - (-1)^{k_l} \sigma^2 q_l^{\text{im}}(\tilde{\theta}_2) = 0 \right\}_{l=1}^m \\ &\quad \times \bigcup \left\{ q_l(\tilde{\theta}_1) - \sigma^2 q_l(\tilde{\theta}_2) = 0 \right\}_{l=m+1}^{p-m}\end{aligned}$$

in the variables  $\sigma, a_1, b_1, \beta_1, \dots, a_m, b_m, \beta_m, a_{m+1}, \dots, a_{p-m}$ .

**Corollary 1:** Let  $\tilde{\theta}_1 \in \Omega(p)^*$  be known, and let  $E_{\mathbf{k}}$  be the systems of equations, from Definition 8, in the variables

$$\sigma, \beta_1, \dots, \beta_m$$

under the constraints

$$\begin{aligned}\sigma &> 0, \\ \beta_l &> 0 \quad \text{and} \quad \frac{\beta_l - (-1)^{k_l} b_l}{2\pi} \in \mathbb{Z} \quad \text{for } l = 1, \dots, m.\end{aligned}$$

Then the AR model  $\tilde{\theta}_1$  is unique if and only if every solution of every system  $E_{\mathbf{k}}$  satisfies  $\beta_l = b_l$  for  $l = 1, \dots, m$ .

**Corollary 2:** Consider the equations  $E_{\mathbf{k}}$ , from Definition 8, in the variables

$$\sigma, a_1, b_1, \beta_1, \dots, a_m, b_m, \beta_m, a_{m+1}, \dots, a_{p-m},$$

under the constraints

$$\begin{aligned}\sigma &> 0 \\ \tilde{\theta}_1 &\in \Omega(p)^* \\ \beta_l &> 0 \quad \text{and} \quad \frac{\beta_l - (-1)^{k_l} b_l}{2\pi} \in \mathbb{Z} \quad \text{for } l = 1, \dots, m.\end{aligned}$$

Then the collection  $\Omega(p)^*$  is unique if and only if for every  $m = 1, \dots, \lfloor p/2 \rfloor$  the solutions of every system  $E_{\mathbf{k}}$  satisfy  $\beta_l = b_l$  for  $l = 1, \dots, m$ .

**Example 5:** Using the formulation of Corollary 2, we show, here, that every element of  $\Omega(2)^*$  is unique. The equations of  $E_{\mathbf{k}}$  are

$$\begin{aligned}0 &= 8a_1 b_1^2 - \sigma^2 8a_1 \beta_1^2 \\ 0 &= 8a_1^2 b_1 - (-1)^{k_1} \sigma^2 8a_1^2 \beta_1.\end{aligned}$$



As  $E_0$  and  $E_1$  contain non-linear polynomial equations, we use techniques from computational algebraic geometry to analyze their solutions. In particular, we use the *slimgb* algorithm of the software SINGULAR [11] to find reduced Gröbner bases [10] of the polynomial ideals corresponding to  $E_0$  and  $E_1$ . The bases are computed using the degree reverse lexicographic monomial order, with the variables ordered as:  $a_1, b_1, \beta_1, \sigma$ . The corresponding Gröbner basis systems are

$$0 = a_1^2 b_1 (b_1 - (-1)^{k_1} \beta_1) \quad (61)$$

$$0 = a_1 (\beta_1^2 \sigma^2 - b_1^2) \quad (62)$$

$$0 = a_1^2 (b_1 - (-1)^{k_1} \sigma^2 \beta_1). \quad (63)$$

Any solution of the system  $E_0$  must be a zero of the first polynomial. Due to the constraints on the variables, this polynomial can only be zero for  $b_1 = \beta_1$ . For  $E_1$ , the first polynomial has no solutions. Therefore, uniqueness holds.

*Example 6:* Applying the procedure of the previous example, we verify uniqueness for  $\Omega(3)^*$ . The systems of equations  $E_{\mathbf{k}}$ , described in Definition 8, are defined by the polynomials

$$\begin{aligned} q_1^{\text{re}}(a_1, b_1, a_2) &= 8a_1 b_1^2 (3a_1^2 - b_1^2 - a_2^2) \\ q_1^{\text{im}}(a_1, b_1, a_2) &= -8a_1^2 b_1 (a_1^2 - 3b_1^2 - a_2^2) \\ q_2(a_1, b_1, a_2) &= -2a_2 ((a_1 + a_2)^2 + b_1^2) ((a_1 - a_2)^2 + b_1^2). \end{aligned}$$

We compute Gröbner bases of the polynomial ideals corresponding to  $E_0$  and  $E_1$ . The bases are computed using the degree reverse lexicographic monomial order, with the variables ordered as:  $a_1, a_2, b_1, \beta_1, \sigma$ .

For  $E_0$ , the resulting Gröbner basis contains the polynomial

$$a_1^2 a_2 (a_2^2 + 4b_1^2) (4a_2^2 + (b_1 - \beta_1)^2) (b_1^2 - \beta_1^2) b_1^4, \quad (64)$$

and any solution of the system  $E_0$  must be a zero of this polynomial. Due to the constraints on the variables, this polynomial can only be zero for  $b_1 = \beta_1$ .

For  $E_1$ , the resulting Gröbner basis contains the polynomial

$$a_1^2 a_2 (a_2^2 + 4b_1^2) (4a_2^2 + (b_1 + \beta_1)^2) (b_1^2 - \beta_1^2) b_1^4, \quad (65)$$

and as in the previous case, this polynomial can only be zero for  $b_1 = \beta_1$ . Therefore, uniqueness holds.

The relationship between the coefficients of the autocorrelation functions in Theorem 4 is an essential component in the proof of known uniqueness results for AR( $p$ ) models. Since this theorem is valid for all values of  $p$ , we conjecture that uniqueness of  $\Omega(p)^*$  holds in general. In fact, we can apply the Gröbner basis approach of Example 6 for validating specific cases of the conjecture.

*Example 7:* The AR(4) model  $\tilde{\theta} = (-1/2, 1/2, -1, 1) \in \Omega(4)^*$  is unique. We verified this property by utilizing the Gröbner basis approach of Example 6. The model  $\tilde{\theta} = (-1/2, 1, -1, 2) \in \Omega(4)^*$  is unique, as well.

We note that computational complexity limits the use of Gröbner basis methods in general; however, there are other ways of exploring the polynomial ideals of Corollary 2 that could potentially be used for proving uniqueness for arbitrary AR model orders.

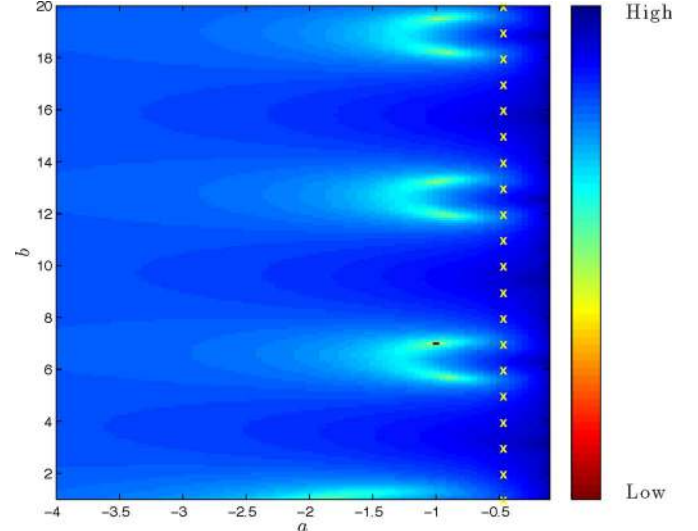


Fig. 2. Error map  $E(\tilde{\theta}_1, \tilde{\theta}_2)$ : Shown here are logarithmic values of the  $\ell_2$  error between uniform samples of  $\varphi(\cdot; \tilde{\theta}_1)$  and  $\varphi(\cdot; \tilde{\theta}_2)$  where  $\tilde{\theta}_1 = (-1, 7)$  and  $\tilde{\theta}_2 = (a, b)$ . The grid resolution is 0.05 for  $a$  and 0.1 for  $b$ .  $\tilde{\theta}_1$  in this case describes a unique process and the global minimum is the dark red pixel that appears at  $(a, b) = (-1, 7)$ . The sampling interval is  $T = 1$ . The yellow  $\times$  markers indicate the multiple initialization points that were used to minimize  $E(\tilde{\theta}_1, \tilde{\theta}_2)$  in Fig. 3 (only part of them are shown due to limited resolution).

### C. Interpolation Error Properties

Fig. 2 depicts the following error measure

$$E(\tilde{\theta}_1, \tilde{\theta}_2) := \log \left( \sum_{n=0}^{\infty} |\varphi(n; \tilde{\theta}_1) - \varphi(n; \tilde{\theta}_2)|^2 \right), \quad (66)$$

where  $\tilde{\theta}_1 = (-1, 7)$  and  $\tilde{\theta}_2 = (a, b)$  takes a range of values. The error map consists of several local minima. A single global minimum can be observed, however, at the correct position. This is in accordance with Theorem 4. The error map is essentially constant for small values of  $a$ , i.e., for large decay rates. This is due to the fact that the small values of  $a$  make the sampled version of the autocorrelation function almost zero, except at the origin.

In order to find the global minimum, we minimize the error measure using multiple initializations, as depicted in Fig. 3. The initialization parameters consist of  $a$  (the real part of the complex pole) and  $b$  (the imaginary part of the complex pole). We avoid the constant region of the error map by assigning a value to  $a$  that is larger than the oracle, e.g.,  $a = -0.5$ . Initial values for the imaginary part varied linearly between 1 and 20 in increments of 0.1. For each initial condition we minimized the error measure until reaching a local minimum. What stands out is the two clusters of local minima values. Each local minimum that belongs to the cluster of low values (less than  $-15$ ) corresponds to the oracle parameters. This is an important property from an algorithmic point of view. It means that the performance of the algorithm does not depend on a particular initialization point, and that by sweeping over the values of  $b$  we are guaranteed to find the correct AR parameters. This is not the case in the second example of the figure, though, which depicts estimation results for a non-unique AR model. There, we still observe two clusters; but the one with the low values is relatively wide, spanning ten orders of magnitude. Moreover, not all local minima there correspond to the oracle parameters. In the next section, we utilize this local minima structure of the error measure  $E(\cdot, \cdot)$  for



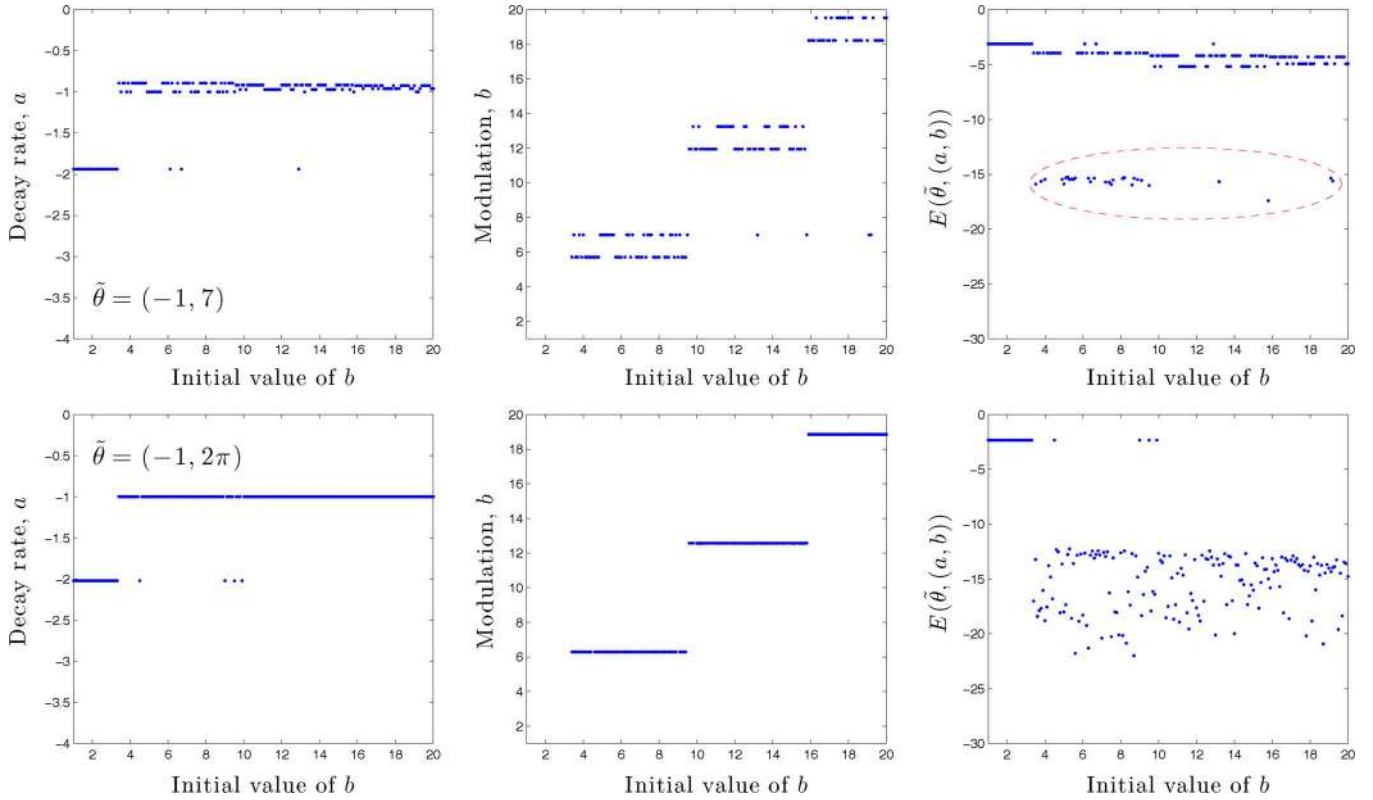


Fig. 3. Local minima in AR model estimation: The sampled version of  $\varphi(\cdot; \tilde{\theta})$  was used for estimating  $\tilde{\theta}$ . The sampling interval is  $T = 1$  and the number of samples was  $N = 30$ . The estimated parameters are  $\hat{\theta} = (a, b)$  (the two left columns). The estimation error is  $E(\tilde{\theta}, \hat{\theta})$  (the right column). Shown here are estimation results for various initial values of  $b$ ; the initial value of  $a$  was  $-0.5$ . When  $\tilde{\theta}$  is unique, we observe a cluster of local minima that are lower than  $-15$  (dashed red ellipse). Each local minimum in that cluster corresponds to the oracle parameters. In a non-unique case, as in the second row, we still observe two clusters. Nevertheless, the cluster having the low values spans more than ten orders of magnitude, and not all of the local minima there correspond to the oracle parameters.

estimating  $\tilde{\theta}$  from sampled data, while imposing no restriction on the sampling interval.

## VI. PARAMETERS ESTIMATION FROM SAMPLED DATA

Our algorithm estimates  $\theta$  from sample values of a continuous-time AR process. The input to the algorithm is: sample values, model order  $p$ , and the maximum allowed value of the imaginary part of the poles. The main stages of the algorithm are: decay rate estimation and multiple minimization.

*Slowest Decay Rate Estimation:* We estimate the autocorrelation sequence by

$$\hat{\varphi}(k) = \frac{1}{N-k} \sum_{n=0}^{N-k-1} (x(n) - \mu) \cdot (x(n+k) - \mu), \quad (67)$$

where  $\mu = \frac{1}{N} \sum_{n=0}^{N-1} x(n)$  and  $0 \leq k \leq K$ . We then fit the values of  $\log |\hat{\varphi}(k)|$  with a regression line; the slope of the line indicates the slowest decay rate of the autocorrelation function. The number of values  $K$  is chosen such that all values  $\hat{\varphi}(k)$ ,  $k > K$  are smaller than  $\hat{\varphi}(K)$  (Fig. 4). In order to accommodate for oscillating autocorrelation sequences, we restrict  $K$  to have a user-defined minimum value  $K_{\min}$ . We also restrict its maximum value by  $K_{\max}$ .

We assume that the order  $p$  of the AR model is known, and we construct the initialization points by taking all possible pole constellations into account. For an AR(1) model, we have a single

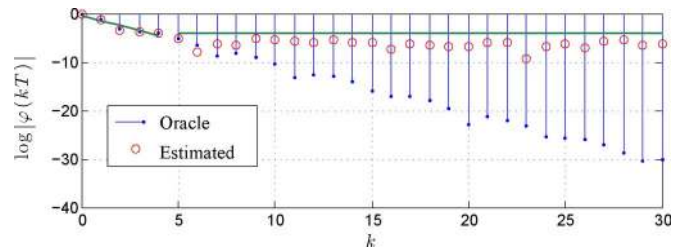


Fig. 4. Autocorrelation sequence of an AR(2) model with two complex poles  $s_{1,2} = -1 \pm 7i$  and a Gaussian process: The estimated autocorrelation sequence starts deviating from the model with increasing values of  $k$ . This sequence is used for determining the slowest decay rate of the model and the number of accurate samples  $K$  (six samples in this example). The first segment of the solid line is obtained by linear regression, and it provides an initial estimate of the decay rate. The second segment indicates an accuracy threshold for values that will not be used for estimation.

real pole. We therefore use a single initialization point and set it to be the negative value of the estimated decay rate, divided by two, namely  $\tilde{\theta} = \{-\frac{\alpha}{2}\}$ . For an AR(2) process, we may have two real poles or two complex conjugate poles. Denoting the estimated decay rate  $\alpha > 0$ , the initialization points are

$$\tilde{\theta}_0 = \underbrace{\left\{ \left( -\frac{\alpha}{2}, -\frac{\alpha}{3} \right) \right\}}_{\text{two real poles}} \cup \underbrace{\left\{ \left( -\frac{\alpha}{2}, \frac{m}{10T} \right) \right\}_{m=1}^{200}}_{\text{complex pair}}. \quad (68)$$

The number of initial complex pairs is defined by the user (we set it in our simulations to be 200). For an AR(3) model, we may have three real poles or a complex conjugate pair with a real pole. The initialization points in such a case are

$$\tilde{\theta}_0 = \underbrace{\left\{ \left( -\frac{\alpha}{2}, -\frac{\alpha}{3}, -\frac{\alpha}{4} \right) \right\}}_{\text{three real poles}} \cup \underbrace{\left\{ \left( -\frac{\alpha}{2}, \frac{m}{10T}, -\frac{\alpha}{3} \right) \right\}_{m=1}^{200}}_{\text{complex pair with a real pole}}. \quad (69)$$

For an AR(4) process we construct the initialization points by permuting two lists of AR(2) points with the necessary modifications. Specifically, we have

$$\tilde{\theta}_0 = \underbrace{\left\{ \left( -\frac{\alpha}{2}, -\frac{\alpha}{3}, -\frac{\alpha}{4}, -\frac{\alpha}{5} \right) \right\}}_{\text{four real poles}} \cup \underbrace{\left\{ \left( -\frac{\alpha}{2}, \frac{m}{10T}, -\frac{\alpha}{3}, -\frac{\alpha}{4} \right) \right\}_{m=1}^{200}}_{\text{complex pair \& two real poles}} \\ \times \cup \underbrace{\left\{ \left( -\frac{\alpha}{2}, \frac{m}{10T}, -\frac{\alpha}{3}, \frac{n}{10T} \right) \right\}_{m=1, n=1}^{200, 200}}_{\text{two complex pairs}}. \quad (70)$$

We follow the same multiple minimization principle for higher-order models as well. During the numerical minimization of the error criterion, we make sure that we reach a local minimum that corresponds to the same constellation of poles that we started with. This is done by restricting the elements of  $\tilde{\theta}_0$  to have the same sign. Practically, we restrict the negative initial elements to be in  $[-10, -0.1]$ , and the positive elements to be in  $[0.01, 20]$ .

Once we establish the set of initialization points, we numerically minimize the log value of the least-squares fit between the  $K$  points of the estimated autocorrelation sequence and the first  $K$  points of the AR model. We repeat this minimization for every point, and we choose the estimated parameters that yield the best fit.

Our algorithm is related to the estimation approach of [26]. The latter resolves aliasing ambiguity by shifting the identified poles by an integer multiple of the angular sampling frequency and by minimizing the difference between the coefficients vector and its projection onto a subspace of a matrix (Remark 7 in Section VII.B). However, our approach is based on multiple minimizations of the autocorrelation fit criterion. Our algorithm extends the results of [4], as well: it is computationally efficient; it characterizes the local minima of the cost function; it proposes a way of determining the multiple initial vectors; and it is based on a quantitative measure of uniqueness.

## VII. EXPERIMENTAL RESULTS

### A. Non-Gaussian Continuous-Time AR Generation

We consider four different Lévy processes  $W(t)$  as input for the computation of the stochastic integral (2): Gaussian, compound Poisson, variance-gamma and Student's t-distribution. These processes have the same power spectrum, but differ in their sparsity properties [35]. We generate their differences

$$y[n] = W(n\Delta) - W((n-1)\Delta) = \int_{(n-1)\Delta}^{n\Delta} w(\tau) d\tau, \quad (71)$$

and approximate (2) by the sum

$$x(nT) \cong \sum_{k=0}^{K'-1} h(k\Delta) \cdot y[nM - k]. \quad (72)$$

These differences are known to be independent and to follow an infinitely divisible distribution [31]. The sampling interval is  $T$  and the integration step size is  $\Delta \ll T$ . The parameter  $K'$  is determined by ensuring  $h(K'\Delta) \ll 1$  and  $M = T/\Delta$  is an integer number.

We simulate the Gaussian process by generating independent Gaussian increments of variance  $\sigma^2\Delta$ . The variance-gamma process is composed of unit-period increments that are conditionally normally distributed with respect to the variance parameter, which is a gamma random variable [25]. We set the variance random variable to be  $\sigma^2u(t)$ , where  $u(t)$  is a gamma process with a unit mean value and variance  $\nu$ . The variance-gamma process is equivalent to a difference of two independent gamma processes  $x(t) = y_1(t) - y_2(t)$ , while setting their mean to  $\Delta\sigma/(\sqrt{2\nu})$  and their variance to  $\Delta\sigma^2/2$  [24]. This is the way we generate the sample values of the process. The compound Poisson process is defined by integrating Dirac distributions that appear at arbitrary locations. We used the Poisson distribution to determine the total number of jumps. The mean value is  $\lambda T(N-1)$  where  $T$  is the sampling interval,  $N$  is the number of sampling points and  $\lambda$  is the average number of jumps per time-unit. The location of each jump is a uniform random variable in  $[0, (N-1)T]$ , and its amplitude is yet another uniform random variable in  $[-0.5, 0.5]$ . The Student's process is generated using independent increments with a Student's parameter  $\nu$ . The filter  $h(t)$  is analytically calculated from the continuous-time poles by means of a partial fraction decomposition of its transfer function.

### B. Estimation Results and Discussion

We validated the proposed algorithm on simulated Gaussian and non-Gaussian AR processes (Table II). Our Monte-Carlo experiments consist of 20 realizations for each set of poles. The error measure consists of two criteria: number of successful estimations, i.e., recall, and average error. A successful estimation is defined by  $\|\tilde{\theta} - \hat{\theta}_n\|_{\ell_2} < \pi/T$  where  $\hat{\theta}_n$  is the estimated vector of parameters at the  $n$ -th Monte Carlo simulation. This measure accounts for cases where the algorithm cannot overcome aliasing. The average error includes successful estimations only, and it is given by

$$\epsilon_m^2 = \frac{1}{N_s} \sum_{n=1}^{N_s} |\tilde{\theta}(m) - \hat{\theta}_n(m)|^2, \quad (73)$$

where  $\hat{\theta}_n(m)$  is the  $m$ -th parameter of the estimated vector of parameters at the  $n$ -th Monte Carlo simulation. The parameter  $N_s$  is the number of successful estimations. The user-defined parameters were:  $K_{\min} = 6$ ;  $K_{\max} = 30$ ; imaginary part values in  $[0.01, 20]$ , real part value in  $[-10, -0.1]$ ; step size along the search line was 0.1. We examined numerous AR models of varying orders and varying power-spectra. The models we examined consist of both real and imaginary poles. We focused on three decay rates:  $-0.5$ ,  $-1$  and  $-2$ . Higher decay rates make the sampled version of the autocorrelation function resemble the Kronecker delta, and the fitting process suffers from numerical limitations in such cases. The range of the modulation values that we examined is relatively large, corresponding to various constellations of power-spectra aliasing.

TABLE II  
ESTIMATION: RESULTS

Model $\tilde{\theta}$	Gaussian <sup>1,2</sup>		Compound Poisson <sup>1,3</sup>		Variance-Gamma <sup>1,4</sup>		Student's t-distribution <sup>1,5</sup>	
	error, $\epsilon$	recall	error, $\epsilon$	recall	error, $\epsilon$	recall	error, $\epsilon$	recall
(-0.5, 0.5)	(0.01, 0.02)	20/20	(0.007, 0.005)	20/20	(0.004, 0.004)	20/20	(0.01, 0.008)	20/20
(-0.5, 1)	(0.09, 0.03)	20/20	(0.008, 0.004)	20/20	(0.003, 0.002)	20/20	(0.01, 0.02)	20/20
(-0.5, 2)	(0.04, 0.01)	20/20	(0.004, 0.009)	20/20	(0.004, 0.003)	20/20	(0.01, 0.02)	20/20
(-0.5, 3)	(0.03, 0.09)	20/20	(0.03, 0.22)	18/20	(0.04, 0.20)	18/20	(0.04, 0.20)	19/20
(-0.5, 4)	(0.07, 0.005)	20/20	(0.005, 0.01)	20/20	(0.005, 0.003)	20/20	(0.01, 0.46)	14/20
(-0.5, 5)	(0.03, 0.004)	20/20	(0.005, 0.007)	19/20	(0.003, 0.003)	20/20	(0.003, 0.02)	20/20
(-0.5, 6)	(0.02, 0.21)	16/20	(0.02, 0.24)	17/20	(0.02, 0.0.29)	12/20	(0.02, 0.0.39)	16/20
(-0.5, 7)	(0.06, 0.01)	20/20	(0.005, 0.007)	17/20	(0.005, 0.004)	17/20	(0.02, 0.05)	16/20
(-0.5, 8)	(0.04, 0.01)	20/20	(0.003, 0.006)	17/20	(0.004, 0.004)	20/20	(0.002, 0.02)	15/20
(-1, 0.5)	(0.02, 0.03)	20/20	(0.02, 0.02)	20/20	(0.02, 0.03)	20/20	(0.04, 0.02)	20/20
(-1, 1)	(0.02, 0.01)	20/20	(0.02, 0.008)	20/20	(0.01, 0.005)	20/20	(0.03, 0.02)	20/20
(-1, 2)	(0.01, 0.01)	20/20	(0.008, 0.01)	20/20	(0.01, 0.006)	20/20	(0.03, 0.08)	20/20
(-1, 3)	(0.07, 0.20)	18/20	(0.07, 0.21)	17/20	(0.07, 0.18)	19/20	(0.08, 0.0.21)	19/20
(-1, 4)	(0.02, 0.01)	18/20	(0.01, 0.01)	19/20	(0.009, 0.01)	17/20	(0.12, 0.05)	18/20
(-1, 5)	(0.01, 0.01)	20/20	(0.007, 0.02)	19/20	(0.009, 0.006)	19/20	(0.008, 0.07)	18/20
(-1, 6)	(0.05, 0.32)	18/20	(0.05, 0.31)	19/20	(0.06, 0.38)	17/20	(0.05, 0.29)	17/20
(-1, 7)	(0.03, 0.33)	16/20	(0.011, 0.012)	14/20	(0.01, 0.01)	13/20	(0.11, 0.82)	13/20
(-1, 8)	(0.008, 0.005)	19/20	(0.008, 0.02)	19/20	(0.009, 0.006)	19/20	(0.01, 0.0.06)	15/20
(-0.5, 0.5, -1)	(0.03, 0.03, 0.15)	20/20	(0.009, 0.006, 0.04)	20/20	(0.013, 0.009, 0.05)	20/20	(0.01, 0.02, 0.09)	20/20
(-0.5, 1, -1)	(0.06, 0.16, 0.33)	20/20	(0.004, 0.007, 0.03)	20/20	(0.005, 0.007, 0.03)	20/20	(0.007, 0.008, 0.04)	20/20
(-0.5, 2, -1)	(0.05, 0.01, 0.08)	20/20	(0.005, 0.004, 0.01)	20/20	(0.005, 0.003, 0.009)	18/20	(0.008, 0.01, 0.02)	18/20
(-0.5, 4, -1)	(0.05, 0.03, 0.06)	20/20	(0.005, 0.007, 0.008)	20/20	(0.006, 0.006, 0.009)	20/20	(0.01, 0.05, 0.02)	18/20
(-0.5, 8, -1)	(0.06, 0.03, 0.10)	20/20	(0.004, 0.01, 0.01)	20/20	(0.005, 0.005, 0.01)	20/20	(0.03, 0.16, 0.09)	14/20
(-1, 0.5, -2)	(0.15, 0.11, 0.53)	20/20	(0.08, 0.08, 0.34)	20/20	(0.10, 0.09, 0.4)	19/20	(0.08, 0.08, 0.45)	20/20
(-1, 1, -2)	(0.03, 0.04, 0.19)	20/20	(0.02, 0.03, 0.17)	20/20	(0.03, 0.05, 0.29)	20/20	(0.03, 0.04, 0.32)	20/20
(-1, 2, -2)	(0.02, 0.02, 0.08)	20/20	(0.01, 0.01, 0.06)	20/20	(0.01, 0.01, 0.05)	20/20	(0.04, 0.03, 0.14)	20/20
(-1, 4, -2)	(0.07, 0.30, 0.06)	16/20	(0.02, 0.02, 0.04)	19/20	(0.02, 0.02, 0.05)	18/20	(0.06, 0.08, 0.08)	14/20
(-1, 8, -2)	(0.01, 0.02, 0.07)	16/20	(0.02, 0.02, 0.07)	15/20	(0.01, 0.02, 0.08)	14/20	(0.04, 0.04, 0.12)	13/20
(-0.5, 0.5, -1, 1)	(0.03, 0.15, 0.23, 0.21)	19/20	(0.02, 0.03, 0.21, 0.18)	19/20	(0.03, 0.03, 0.19, 0.12)	20/20	(0.02, 0.04, 0.12, 0.19)	19/20
(-0.5, 1, -1, 2)	(0.01, 0.01, 0.08, 0.04)	20/20	(0.03, 0.05, 0.2, 0.1)	20/20	(0.02, 0.04, 0.26, 0.01)	20/20	(0.003, 0.02, 0.09, 0.06)	20/20

<sup>1</sup> The sampling interval is  $T = 1$ ; integration was approximated by a step size of  $\Delta = 10^{-2}$

<sup>2</sup>  $\sigma = 1$  and the number of samples is  $N = 2 \cdot 10^5$

<sup>3</sup>  $\lambda = 0.01$ , the amplitude of the Dirac distribution are uniformly distributed random variables between  $[-0.5, 0.5]$ , and the number of samples is  $N = 2 \cdot 10^5$

<sup>4</sup>  $\sigma = 1, \nu = 1$

<sup>5</sup>  $\nu = 2$  and the number of samples is  $N = 3 \cdot 10^5$

Our estimation algorithm was able to correctly estimate the AR parameters in all cases, and with a high recall value. In cases where the success rate is less than 20/20, the estimated imaginary part was at least  $4\pi$  larger than the oracle value; the decay rate was properly estimated, though. A comparison between the different innovation models indicates that there is not much difference in terms of performance between the Gaussian and the non-Gaussian innovation. The heavy tail student's-t distribu-

tion is more challenging, as it requires more sampling values for achieving similar estimation results. The estimation error  $\epsilon$  is very small compared to the oracle values. In some cases, however, its elements reach values larger than 10%, e.g., in  $\tilde{\theta} = (-0.5, 0.5, -1, -1)$ . Two sources of this error are: (1) estimating the autocorrelation values from sample data, and (2) numerical limitations in simulating the data. Regarding the latter, the independent increments should, in principle, exhibit an auto-

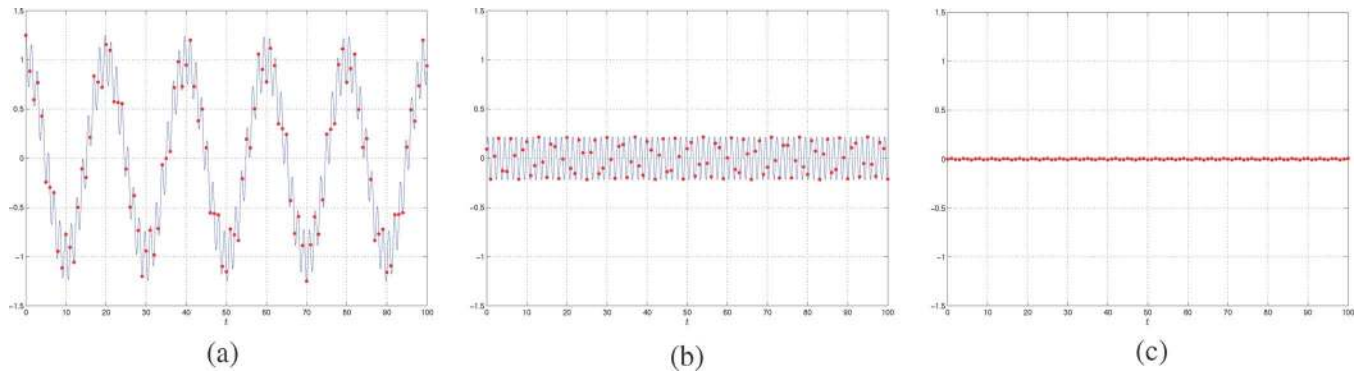


Fig. 5. An example of an almost periodic function: Shown here in (a) is the function  $y(t) = \cos(2\pi t/20) + \cos(2\pi t/\sqrt{2})$ . The function possesses an almost periodic property, which is demonstrated in (b) and in (c). The solid line indicates a sampling grid of 0.01 and the dots indicate a sampling grid of unit value. (a)  $y(t)$ , (b)  $y(t) - y(t - 20)$ , (c)  $y(t) - y(t - 140)$ .

correlation sequence that is the Kronecker delta. In practice, this sequence does not vanish, so the autocorrelation sequence of the sampled AR process does not exactly follow the model. This limitation can be observed in Fig. 4, where the estimated values (circles) deviate from autocorrelation model values (dots). It then follows that the algorithm can use only the first few values of the estimated autocorrelation sequence.

We observed a similar situation when estimating the poles of the Rao-Garnier ARMA(4,1) model. This model was used as a benchmark in [26], and it consists of four poles  $-0.5 \pm 1.937i, -2 \pm 19.9i$  and a single zero 0.25. Focusing on the AR part, we verified its uniqueness by means of Theorem 4. We then simulated sample data based on  $\tilde{\theta} = (-0.5, 1.937, -2, 19.9)$  and applied our algorithm in the same way we generated Table II. We used  $N = 10^6$  sample values and estimated the first ten autocorrelation values  $\hat{\varphi}(k)$  of (67). These values deviate from the oracle autocorrelation values  $\varphi(t, \tilde{\theta})$  of Proposition 1 by  $\|\varphi - \hat{\varphi}\|_{\ell_2} / \|\varphi\|_{\ell_2} = 1.75\%$ , which results in estimated parameters of  $(-0.5088, 1.9329, -0.2659, 17.0596)$ . Although different, they correspond to a continuous-time power spectrum that is very similar to the oracle. The main difference between the two is an additional side lobe that appears at  $\omega = 17.05$  [rad/time-unit]. The  $L_2$  difference between the two power spectra is 3.5%, and it reduces to 1.4% when excluding radial frequencies higher than 10 [rad/time-unit]. When, however, using the first ten values of the oracle autocorrelation function, the error in estimating  $\tilde{\theta}$  is, as expected, very small—less than 0.001 for each element.

The Gaussian maximum likelihood estimator of [21] can be interpreted as an autocorrelation fitting, and it performs as well as the proposed algorithm. We demonstrate its performance in Appendix C. The advantage of the proposed approach, however, is its computational complexity; it runs, on average, five times faster than the likelihood minimization algorithm.

## VIII. CONCLUSION

In this work, we investigated the intertwining relations between continuous-time AR models and their sampled versions. We introduced two new model representations that allowed us to derive a closed-form expression for the autocorrelation function and to associate a measure with subsets of AR models. We

then focused on two subsets: AR models with real poles, and AR models with minimal restrictions. The latter is defined by removing a set of measure zero from the collection of AR models. We characterized the models by a uniqueness property, which is the ability to map the parameters of the sampled model back to their continuous-time counterpart with no ambiguity. We then proved that an AR model with real poles is unique. We also provided uniqueness criteria for AR models with complex and real poles. To this aim, we exploited the decay rate property and the almost-periodic structure of the AR autocorrelation function. We utilized our theory for introducing an estimation algorithm that recovers continuous-time AR parameters from sampled data, regardless of the sampling interval. We then demonstrated its usefulness for general Gaussian and non-Gaussian AR processes. Our experimental results indicate that the uniqueness property of sampled continuous-time AR processes could provide a valuable tool for designing sampling-invariant algorithms for signal and image processing tasks.

## APPENDIX A

### NOTES ON ALMOST PERIODIC TRIGONOMETRIC SEQUENCES

*Example ([1], Appendix B):* Consider the trigonometric polynomial  $y(t) = \cos(2\pi t/20) + (1/4)\cos(2\pi t/\sqrt{2})$ . It is not periodic because the two periods  $T_1 = 20$  and  $T_2 = \sqrt{2}$  are incommensurable. However,  $y(t)$  looks similar to a periodic function (Fig. 5), and if we allow for a certain error  $\epsilon$  between  $y(t)$  and  $y(t + T)$ , then we can find a value of  $T$  satisfying

$$|y(t) - y(t + T)| < \epsilon. \quad (74)$$

The figure depicts the error for  $T = 20, 140$ .

Almost periodicity of sampled trigonometric polynomials is related to Kronecker's approximation theorem. In the following,  $\|t\|$  denotes the distance to the greatest integer smaller than or equal to  $t$ .

*Theorem (Kronecker's Approximation Theorem [5]):*

1. If the real numbers  $(x_1, \dots, x_n)$  are linearly independent over  $\mathbb{Q}$ , then for real  $\alpha_k$  and  $\epsilon > 0$  there are arbitrarily large real  $t$  for which

$$\|tx_k - \alpha_k\| < \epsilon \quad k = 1, \dots, n. \quad (75)$$

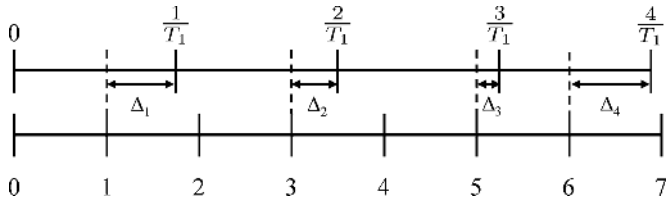


Fig. 6. An example of Kronecker's approximation theorem:  $T_1$  is the irrational period of a trigonometric function. Considering a sampling step of unit value, the sequence  $\{\Delta_n\}_n$  is dense in the interval  $[0, 1]$ .

2. If the numbers  $(1, x_1, \dots, x_n)$  are linearly independent over  $\mathbb{Q}$ , then in part 1,  $t$  can be taken to be an arbitrarily large integer.

Consider a single trigonometric function  $y(t) = \cos(2\pi t/T_1)$ :

- If  $1/T_1$  is rational, then the sequence  $\{y(n)\}_{n=0,1,2,\dots}$  takes a finite number of values in a periodic fashion.
- If  $1/T_1$  is irrational, then by Kronecker's theorem,  $p/T_1$  can be made arbitrarily close to an integer value for an arbitrarily large integer value of  $p$ , say  $p_0$ . This means that  $\Delta_n$ , in Fig. 6, can be made arbitrarily small. It then follows that  $y(n + p_0) \cong y(n)$  with an arbitrary precision due to the continuity of  $y(t)$ .

Consider now a sum of two trigonometric functions  $y(t) = \cos(2\pi t/T_1) + \cos(2\pi t/T_2)$ :

- If both  $1/T_1, 1/T_2$  are rational, then  $y(n)$  is periodic.
- If  $1/T_1$  is rational and  $1/T_2$  is irrational, then we denote the period of  $\cos(2\pi n/T_1)$  as  $N_1$ . It then follows by Kronecker's theorem that  $\tilde{p}N_1/T_2$  can be made arbitrarily close to an integer value, say by setting  $\tilde{p}$  to be  $\tilde{p}_0$ . Defining  $p_0 := N_1\tilde{p}_0$ , we then have that  $y(n + p_0) = y(n + \tilde{p}_0N_1) \cong y(n)$ .
- If both  $1/T_1, 1/T_2$  are irrational and
  - if  $\{1, 1/T_1, 1/T_2\}$  are independent over  $\mathbb{Q}$ , then  $n/T_1$  and  $n/T_2$  can be made arbitrarily close to an integer by Kronecker's theorem.
  - if  $\{1, 1/T_1, 1/T_2\}$  are dependent over  $\mathbb{Q}$ , then it must hold that  $n_1/T_1 + n_2/T_2 = n_3$  for some integer numbers  $n_1, n_2, n_3$ , which are not all zero. This, in turn, means that  $y(t)$  is periodic with an irrational period  $\max\{T_1, n_2/n_1T_1\}$ , and one can apply the same reasoning as was done for the single trigonometric function.

In a similar manner, Kronecker's approximation theorem can be used to show that the sum of more than two trigonometric functions yields an almost periodic sequence, as well.

APPENDIX B  
LEMMAS

*Lemma 5:* Define  $F : \mathbb{R}^2 \rightarrow \mathbb{R}$  by

$$F(x, y) := (\gamma_1 e^{2\pi i x} + \gamma_1^* e^{-2\pi i x}) - (\gamma_2 e^{2\pi i y} + \gamma_2^* e^{-2\pi i y}), \quad (76)$$

where  $\gamma_1, \gamma_2$  are non-zero complex numbers. Also, let  $b, \beta$  be positive real numbers, where  $b$  is an irrational multiple of  $\pi$ . If the collection of zeros of  $F$  includes

$$\left\{ \left( \frac{b}{2\pi}n, \frac{\beta}{2\pi}n \right) \right\}_{n \in \mathbb{Z}^+}, \quad (77)$$

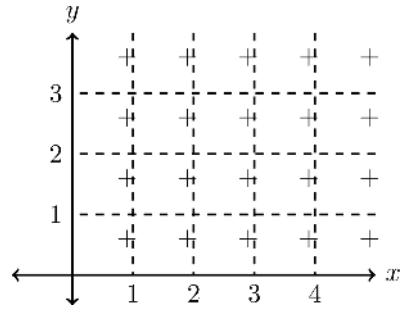


Fig. 7. Due to the periodicity of  $F$ , each zero of  $F$  corresponds to a lattice of zeros. The plus signs indicate zeros for a particular example.

then there are integers  $n_1, n_2, n_3$  with  $n_1 > 0$  and  $n_2 \neq 0$  such that

$$n_1 \frac{b}{2\pi} = n_2 \frac{\beta}{2\pi} + n_3. \quad (78)$$

*Proof:* First, suppose that  $1, b/2\pi, \beta/2\pi$  are linearly independent over the field of rational numbers, which will lead to a contradiction. Notice that the function  $F$  of (76) is 1-periodic in each variable. We therefore define its unit fundamental domain to be  $[0, 1) \times [0, 1)$ . Due to the periodicity, for each zero  $(nb/2\pi, n\beta/2\pi)$ , there correspond an infinite number of zeros lying on a lattice (see Fig. 7); i.e., for any  $k_1, k_2 \in \mathbb{Z}$ ,

$$\left( \frac{b}{2\pi}n + k_1, \frac{\beta}{2\pi}n + k_2 \right) \quad (79)$$

is also a zero of  $F$ . For each zero, we shall be primarily concerned with the associated lattice point lying in the unit fundamental domain. Specifically, for each zero of  $F$ ,  $((b/2\pi)n, (\beta/2\pi)n)$ , let  $k_1(n)$  and  $k_2(n)$  be integers such that

$$p(n) := \left( \frac{\beta}{2\pi}n - k_1(n), \frac{\beta}{2\pi}n - k_2(n) \right) \quad (80)$$

lies in  $[0, 1) \times [0, 1)$ . Now the linear independence assumption together with Kronecker's approximation theorem (cf. [5]) implies that  $\{p(n) : n \in \mathbb{Z}^+\}$  is dense in the unit fundamental domain. As  $F$  is continuous, the zero set of  $F$  contains the entire region  $[0, 1) \times [0, 1)$ , and the periodicity of  $F$  implies that  $F$  is everywhere 0. In particular, the zero set contains the line  $\{(x, 0) : x \in \mathbb{R}\}$ . This implies that

$$\gamma_1 e^{2\pi i x} + \gamma_1^* e^{-2\pi i x} = 2\Re(\gamma_2), \quad (81)$$

which is a contradiction. Hence,  $1, b/2\pi, \beta/2\pi$  must be linearly dependent over the field of rational numbers; i.e., there exist integers  $n_1, n_2, n_3$  (not all 0), with  $n_1 \geq 0$ , such that

$$n_1 \frac{b}{2\pi} = n_2 \frac{\beta}{2\pi} + n_3. \quad (82)$$

To finish the proof, we must show that  $n_1 \neq 0$ . For if this is indeed the case, then the irrationality of  $b/2\pi$  will imply that  $n_2 \neq 0$  and  $\beta/2\pi$  is irrational. To see that  $n_1 > 0$ , we assume  $n_1 = 0$ , and we shall derive a contradiction.

Under the assumption  $n_1 = 0$ , we must have that  $\beta/2\pi$  is a rational number. Now, define  $p(n)$  as in (80). As  $\beta/2\pi$  is rational,  $p_2(n) := n\beta/2\pi - k_2(n)$  takes only a finite number of

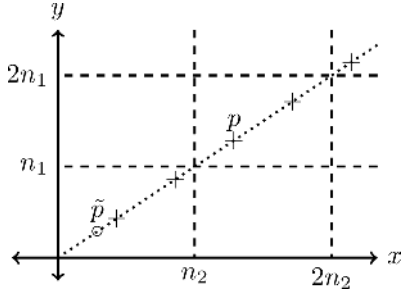


Fig. 8. Plus signs indicate zeros of the function  $F$  on a line of slope  $n_1/n_2$ . The point  $p$  is a zero of  $F$  lying outside of the fundamental domain, and  $\tilde{p}$  is a translated version of  $p$ , which is also a zero of  $F$ .

values. However,  $b/2\pi$  is irrational, so Kronecker's approximation theorem implies that  $\{p_1(n) := b/2\pi - k_1(n) : n \in \mathbb{Z}^+\}$  is dense in  $[0, 1)$ . As each  $p(n)$  is a zero of  $F$ , we have

$$\gamma_1 e^{2\pi i p_1(n)} + \gamma_1^* e^{-2\pi i p_1(n)} = \gamma_2 e^{2\pi i p_2(n)} + \gamma_2^* e^{-2\pi i p_2(n)}. \quad (83)$$

The left-hand side of (83) takes an infinite number of values, while the right-hand side takes a finite number of values. This is a contradiction.  $\square$

*Lemma 6:* Under the assumptions of Lemma 5, one of the following holds:

- $n_1 = n_2$  and  $\gamma_1 = \gamma_2$
- $n_1 = -n_2$  and  $\gamma_1 = \gamma_2^*$

*Proof:* A subset of the zeros from (77) is given by considering only those integers that are multiples of  $n_1$ :

$$\left\{ \left( \frac{b}{2\pi} n_1 n, \frac{\beta}{2\pi} n_1 n \right) \right\}_{n \in \mathbb{Z}^+}. \quad (84)$$

Note that this particular collection of zeros has the alternative representation

$$\left\{ \left( \frac{\beta}{2\pi} n_2 n + n_3 n, \frac{\beta}{2\pi} n_1 n \right) \right\}_{n \in \mathbb{Z}^+}, \quad (85)$$

due to (78). Then the periodicity of  $F$  implies that there are zeros at

$$\left\{ \left( \frac{\beta}{2\pi} n_2 n, \frac{\beta}{2\pi} n_1 n \right) \right\}_{n \in \mathbb{Z}^+}. \quad (86)$$

We can see that each of these points lies on the line  $y = (n_1/n_2)x$  (see Fig. 8). For the remainder of the proof, we shall be particularly concerned with the zeros lying on this line. Furthermore, the domain  $[0, n_2) \times [0, n_1)$  will be of primary importance, so we denote this as the fundamental domain of  $F$ . Note that, here, we are assuming  $n_2 > 0$ . The case  $n_2 < 0$  is similar, and it will be handled below.

Now, if a zero  $p$  of  $F$  lies outside of the fundamental domain, then there is an integer  $k$  such that  $\tilde{p} := p - (kn_2, kn_1)$  lies in the fundamental domain. The point  $\tilde{p}$  is also a zero of  $F$ , and it lies on the line  $y = (n_1/n_2)x$  (see Fig. 8). For each positive integer  $n$ , define the integer  $k(n)$  such that

$$p(n) := \left( \frac{\beta}{2\pi} n_2 n - k(n)n_2, \frac{\beta}{2\pi} n_1 n - k(n)n_1 \right) \quad (87)$$

TABLE III  
GAUSSIAN MAXIMUM LIKELIHOOD ESTIMATION OF AN AR MODEL WITH GAUSSIAN AND NON-GAUSSIAN INNOVATION

Model $\tilde{\theta}$	Gaussian Innovation <sup>1,2</sup>		Compound Poisson Innovation <sup>1,3</sup>	
	error, $\epsilon$	recall	error, $\epsilon$	recall
(-0.5, 0.5)	(0.003, 0.004)	20/20	(0.006, 0.004)	20/20
(-0.5, 1)	(0.002, 0.002)	20/20	(0.005, 0.003)	20/20
(-0.5, 2)	(0.004, 0.003)	20/20	(0.003, 0.005)	20/20
(-0.5, 3)	(0.03, 0.21)	18/20	(0.03, 0.15)	17/20
(-0.5, 4)	(0.003, 0.003)	20/20	(0.009, 0.008)	20/20
(-0.5, 5)	(0.003, 0.002)	20/20	(0.003, 0.009)	20/20
(-0.5, 6)	(0.02, 0.19)	9/20	(0.02, 0.26)	10/20
(-0.5, 7)	(0.002, 0.003)	20/20	(0.005, 0.003)	15/20
(-0.5, 8)	(0.003, 0.002)	20/20	(0.003, 0.01)	15/20
(-1, 0.5)	(0.004, 0.009)	20/20	(0.01, 0.01)	20/20
(-1, 1)	(0.005, 0.004)	20/20	(0.01, 0.005)	20/20
(-1, 2)	(0.006, 0.004)	20/20	(0.008, 0.01)	20/20
(-1, 3)	(0.08, 0.26)	18/20	(0.8, 0.26)	20/20
(-1, 4)	(0.009, 0.008)	20/20	(0.01, 0.01)	20/20
(-1, 5)	(0.01, 0.005)	19/20	(0.008, 0.01)	18/20
(-1, 6)	(0.05, 0.34)	18/20	(0.06, 0.4)	17/20
(-1, 7)	(0.07, 0.012)	15/20	(0.009, 0.012)	15/20
(-1, 8)	(0.006, 0.004)	16/20	(0.007, 0.02)	12/20
(-0.5, 0.5, -1)	(0.009, 0.005, 0.03)	20/20	(0.01, 0.006, 0.05)	20/20
(-0.5, 1, -1)	(0.004, 0.005, 0.02)	20/20	(0.004, 0.005, 0.02)	20/20
(-0.5, 2, -1)	(0.003, 0.002, 0.01)	20/20	(0.004, 0.003, 0.01)	20/20
(-0.5, 4, -1)	(0.006, 0.004, 0.009)	20/20	(0.007, 0.01, 0.008)	20/20
(-0.5, 8, -1)	(0.005, 0.05, 0.009)	18/20	(0.009, 0.01, 0.02)	11/20
(-1, 0.5, -2)	(0.03, 0.01, 0.99)	20/20	(0.05, 0.02, 0.2)	20/20
(-1, 1, -2)	(0.02, 0.02, 0.02)	20/20	(0.02, 0.01, 0.09)	20/20
(-1, 2, -2)	(0.01, 0.01, 0.06)	20/20	(0.02, 0.009, 0.02)	20/20
(-1, 4, -2)	(0.02, 0.02, 0.05)	17/20	(0.02, 0.02, 0.04)	18/20
(-1, 8, -2)	(0.02, 0.02, 0.06)	14/20	(0.02, 0.03, 0.08)	13/20
(-0.5, 0.5, -1, 1)	(0.01, 0.006, 0.01, 0.03)	20/20	(0.01, 0.007, 0.02, 0.03)	20/20
(-0.5, 1, -1, 2)	(0.005, 0.004, 0.01, 0.008)	20/20	(0.005, 0.006, 0.02, 0.007)	20/20

<sup>1</sup> the sampling interval is  $T = 1$ ; integration was approximated by a step size of  $\Delta = 10^{-2}$

<sup>2</sup>  $\sigma = 1$  and the number of samples is  $N = 2 \cdot 10^5$

<sup>3</sup>  $\lambda = 1$ , the amplitude of the Dirac distribution are uniformly distributed random variables between  $[-0.5, 0.5]$ , and the number of samples is  $N = 2 \cdot 10^5$

lies in the fundamental domain. Since  $(\beta/2\pi)n_2$  is irrational, Kronecker's approximation theorem implies that the set

$$\left\{ \left( \frac{\beta}{2\pi} n - k(n), n_2 \right) \right\}_{n \in \mathbb{Z}^+} \quad (88)$$

is dense in  $[0, n_2)$ . Therefore the set  $\{p(n) : n \in \mathbb{Z}^+\}$  is dense on the line  $y = (n_1/n_2)x$  within the fundamental domain. As  $F$  is a continuous function, the line segment of  $y = (n_1/n_2)x$  within the fundamental domain is a zero set for  $F$ . Periodicity of  $F$  then implies that the entire line  $y = (n_1/n_2)x$  is a zero set for  $F$ . This means

$$\gamma_1 e^{2\pi i x} + \gamma_1^* e^{-2\pi i x} = \gamma_2 e^{2\pi i x n_1/n_2} + \gamma_2^* e^{-2\pi i x n_1/n_2}. \quad (89)$$

Considering this equation in the Fourier domain, we see that we must have  $n_1/n_2 = 1$ , so

$$\gamma_1 e^{2\pi i x} + \gamma_1^* e^{-2\pi i x} = \gamma_2 e^{2\pi i x} + \gamma_2^* e^{-2\pi i x}. \quad (90)$$

Finally, by the uniqueness of Fourier series,  $\gamma_1 = \gamma_2$ .

For the case  $n_2 < 0$ , we can apply similar techniques. In this case, one can show that the line  $y = (n_1/n_2)x$  is a zero set for  $F$ , by translating the zeros of  $F$  to the fundamental domain  $[0, -n_2) \times [-n_1, 0)$ . As in the previous case, this implies (89), which can only hold if  $n_1/n_2 = -1$ . Therefore

$$\gamma_1 e^{2\pi i x} + \gamma_1^* e^{-2\pi i x} = \gamma_2 e^{-2\pi i x} + \gamma_2^* e^{2\pi i x}, \quad (91)$$

and uniqueness of Fourier series implies  $\gamma_1 = \gamma_2^*$ .  $\square$

*Lemma 7:* Under the assumptions of Lemma 5, we have that  $n_3$  is a multiple of  $n_1$ . That is, there exists an integer  $n_4$  such that

- $b = \beta + 2\pi n_4$  if  $n_1 = n_2$
- $b = -\beta + 2\pi n_4$  if  $n_1 = -n_2$ .

*Proof:* As both cases are similar, we prove only the first one. Combining the results of Lemma 5 and Lemma 6, we know that there are integers  $n_1, n_3$ , with  $n_1 > 0$ , such that

$$b = \beta + \frac{n_3}{n_1} 2\pi. \quad (92)$$

Since the function  $F$  of (76) is zero for points in (77), we have

$$0 = \gamma_1 e^{i b n} + \gamma_1^* e^{-i b n} - \gamma_1 e^{i \beta n} - \gamma_1^* e^{-i \beta n}. \quad (93)$$

We then use (92) to reduce this to

$$0 = \gamma_1 e^{i \beta n} (e^{i 2\pi n n_3 / n_1} - 1) + \gamma_1^* e^{-i \beta n} (e^{-i 2\pi n n_3 / n_1} - 1) \quad (94)$$

As this is true for any positive integer  $n$ , it must hold for  $n = k n_1 + 1$  for any positive integer  $k$ ; i.e.,

$$0 = (\gamma_1 e^{i \beta} (e^{i 2\pi n_3 / n_1} - 1)) e^{i \beta n_1 k} + (\gamma_1^* e^{i \beta} (e^{i 2\pi n_3 / n_1} - 1))^* e^{-i \beta n_1 k}, \quad (95)$$

which can be written as

$$0 = \eta e^{i \beta n_1 k} + \eta^* e^{-i \beta n_1 k}, \quad (96)$$

where

$$\eta := \gamma_1 e^{i \beta} (e^{i 2\pi n_3 / n_1} - 1). \quad (97)$$

Since  $\beta$  is an irrational multiple of  $\pi$ , Kronecker's approximation theorem implies that

$$\eta e^{i \beta n_1 k} + \eta^* e^{-i \beta n_1 k} \quad (98)$$

takes a non-zero value for some value of  $k$ , unless  $\eta = 0$ . Therefore  $\eta$  must be 0, and hence  $n_3$  is an integer multiple of  $n_1$ .  $\square$

## APPENDIX C

### COMPARISON WITH GAUSSIAN LIKELIHOOD ESTIMATOR

We provide estimation results of an estimator that minimizes a Gaussian likelihood function of the sampled data [21]. Table III describes estimation results for AR models that are driven by continuous-time Gaussian innovation. It also describes the minimization of the same Gaussian likelihood function for AR models that are driven by yet another continuous-time innovation process—the compound Poisson process.

## ACKNOWLEDGMENT

The authors thank the reviewers for their valuable comments and suggestions.

## REFERENCES

- [1] I. Amidror, *The Theory of the Moiré Phenomenon*, ser. Computational Imaging and Vision, 2nd ed. New York, NY, USA: Springer, 2009, vol. 1, Periodic Layers.
- [2] K. J. Åström, *Introduction to Stochastic Control Theory*. New York, NY, USA: Academic, 1970.
- [3] M. S. Bartlett, "On the theoretical specification and sampling properties of autocorrelated time-series," *J. Roy. Statist. Soc.*, vol. 8, no. 1, pp. 27–41, 1946.
- [4] A. R. Bergstrom, "Gaussian estimation of structural parameters in higher order continuous time dynamic models," *Econometrica* vol. 51, no. 1, pp. 117–152, 1983 [Online]. Available: <http://EconPapers.repec.org/RePEc:ecm:emtrp:v:51:y:1983:i:1:p:117-152>
- [5] H. Bohr and B. Jessen, "One more proof of Kronecker's theorem," *J. London Math. Soc.*, vol. 7, no. 4, pp. 274–275, 1932.
- [6] J. B. Boyling, "Green's functions for polynomials in the Laplacian," *Z. Angew. Math. Phys.* vol. 47, no. 3, pp. 485–492, 1996 [Online]. Available: <http://dx.doi.org/10.1007/BF00916651>
- [7] P. J. Brockwell, "A note on the embedding of discrete-time arma processes," *J. Time Ser. Anal.* vol. 16, no. 5, pp. 451–460, 1995 [Online]. Available: <http://dx.doi.org/10.1111/j.1467-9892.1995.tb00246.x>
- [8] M. J. Chambers and M. A. Thornton, "Discrete time representation of continuous time ARMA processes," *Econometr. Theory*, vol. 28, no. 01, pp. 219–238, 2012.
- [9] C. Corduneanu, *Almost Periodic Oscillations and Waves*. New York, NY, USA: Springer, 2009 [Online]. Available: <http://dx.doi.org/10.1007/978-0-387-09819-7>
- [10] D. A. Cox, J. Little, and D. O'Shea, *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra*. New York, NY, USA: Springer, 2007, vol. 10.
- [11] W. Decker, G.-M. Greuel, G. Pfister, and H. Schönemann, "Singular 3-1-6—A computer algebra system for polynomial computations," 2012 [Online]. Available: <http://www.singular.uni-kl.de>
- [12] F. Ding, L. Qiu, and T. Chen, "Reconstruction of continuous-time systems from their non-uniformly sampled discrete-time systems," *Automatica*, vol. 45, no. 2, pp. 324–332, 2009.
- [13] H. Fan, T. Söderström, M. Mossberg, B. Carlsson, and Y. Zou, "Estimation of continuous-time AR process parameters from discrete-time data," *IEEE Trans. Signal Process.*, vol. 47, no. 5, pp. 1232–1244, May 1999.
- [14] A. Feuer and G. Goodwin, *Sampling in Digital Signal Processing and Control*. Boston, MA, USA: Birkhäuser, 1996.
- [15] I. Gelfand and N. Y. Vilenkin, *Generalized Functions*, ser. Appl. Harmon. Anal.. New York, NY, USA: Academic, 1964, vol. 4.
- [16] J. Gillberg and L. Ljung, "Frequency-domain identification of continuous-time ARMA models from sampled data," *Automatica*, vol. 45, pp. 1371–1378, 2009.
- [17] G. C. Goodwin, J. I. Yuz, and H. Garnier, "Robustness issues in continuous-time system identification from sampled data," presented at the 16th IFAC World Congress, Prague, Czech Republic, Jul. 2005.
- [18] M. Huzii, "Embedding a Gaussian discrete-time autoregressive moving average process in a Gaussian continuous-time autoregressive moving average process," *J. Time Ser. Anal.* vol. 28, no. 4, pp. 498–520, 2007 [Online]. Available: <http://dx.doi.org/10.1111/j.1467-9892.2006.00520.x>
- [19] R. Johansson, "Identification of continuous-time models," *IEEE Trans. Signal Process.*, vol. 42, no. 4, pp. 887–897, Apr. 1994.
- [20] Y. Katznelson, *An Introduction to Harmonic Analysis*, ser. Cambridge Mathematical Library, 3rd ed. Cambridge, U.K.: Cambridge Univ. Press, 2004.
- [21] H. Kirshner, S. Maggio, and M. Unser, "A sampling theory approach to continuous ARMA identification," *IEEE Trans. Signal Process.*, vol. 59, no. 10, pp. 4620–4634, Oct. 2011.
- [22] E. K. Larsson, M. Mossberg, and T. Söderström, "An overview of important practical aspects of continuous-time ARMA system identification," *Circuits Syst. Signal Process.*, vol. 25, no. 1, pp. 17–46, May 2006.
- [23] E. K. Larsson, "Limiting sampling results for continuous-time ARMA systems," *Int. J. Control*, vol. 78, no. 7, pp. 461–473, May 2005.
- [24] D. B. Madan, P. Carr, and E. C. Chang, "The variance gamma process and option pricing," *Eur. Finance Rev.*, vol. 2, pp. 79–105, 1998.



- [25] D. B. Madanm and E. Seneta, "The variance gamma (v.g.) model for share market returns," *J. Business*, vol. 63, no. 4, pp. 511–524, 1990.
- [26] D. Marelli and M. Fu, "A continuous-time linear system identification method for slowly sampled data," *IEEE Trans. Signal Process.*, vol. 58, no. 5, pp. 2521–2533, 2010.
- [27] M. S. Phadke and S. M. Wu, "Modeling of continuous stochastic processes from discrete observations with application to sunspot data," *J. Amer. Statist. Assoc.*, vol. 69, no. 346, pp. 325–329, Jun. 1974.
- [28] D.-T. Pham, "Estimation of continuous-time autoregressive model from finely sampled data," *IEEE Trans. Signal Process.*, vol. 48, no. 9, pp. 2576–2584, Sep. 2000.
- [29] P. C. B. Phillips, "The problem of identification in finite parameter continuous time models," *J. Econometr.* vol. 1, no. 4, pp. 351–362, 1973 [Online]. Available: <http://EconPapers.repec.org/RePEc:eee:econom:v:1:y:1973:i:4:p:351-362>
- [30] P. Protter, *Stochastic Integration and Differential Equations*. New York, NY, USA: Springer, 2004.
- [31] K. Sato, *Lévy Processes and Infinitely Divisible Distributions*, ser. Cambridge Studies in Adv. Math.. Cambridge, U.K.: Cambridge Univ. Press, 1999.
- [32] T. Söderström, "Computing stochastic continuous-time models from ARMA models," *Int. J. Control*, vol. 53, no. 6, pp. 1311–1326, May 1991.
- [33] T. Söderström, H. Fan, B. Carlsson, and S. Bigi, "Least squares parameter estimation of continuous-time arx models from discrete-time data," *IEEE Trans. Autom. Control*, vol. 42, no. 5, pp. 659–673, 1997.
- [34] H. Tsai and K. S. Chan, "Maximum likelihood estimation of linear continuous time long memory processes with discrete time data," *J. Roy. Statist. Soc., Ser. B (Statist. Method.)*, vol. 67, no. 5, pp. 703–716, 2005.
- [35] M. Unser and P. Tafti, "Stochastic models for sparse and piecewise-smooth signals," *IEEE Trans. Signal Process.*, vol. 59, no. 3, pp. 989–1006, Mar. 2011.
- [36] B. Wahlberg, "Limit results for sampled systems," *Int. J. Control*, vol. 78, no. 3, pp. 1267–1283, 1988.
- [37] J. P. Ward, H. Kirshner, and M. Unser, "Is uniqueness lost for under-sampled continuous-time auto-regressive processes?," *IEEE Signal Process. Lett.*, vol. 19, no. 4, pp. 183–186, Apr. 2012.
- [38] G. N. Watson, *A Treatise on the Theory of Bessel Functions*. Cambridge, U.K.: Cambridge Univ. Press, 1944.
- [39] J. Yuz, J. Alfaro, J. Agüero, and G. Goodwin, "Identification of continuous-time state space models from nonuniform fast-sampled data," *IET Contr. Theory Appl.*, vol. 5, no. 7, pp. 842–855, 2011.



**Hagai Kirshner** (M'13) received the B.Sc. (*summa cum laude*), M.Sc., and Ph.D. degrees in electrical engineering from the Technion-Israel institute of Technology, Haifa, in 1997, 2005, and 2009, respectively.

He is currently a Postdoctoral Fellow at the Biomedical Imaging Group at the Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland. From 1997 to 2004, he was a System Engineer within the Technology Division, IDF. His research interests include sampling theory, stochastic image representation and processing, and computational methods in super resolution fluorescence microscopy.



**Michael Unser** (F'99) is Professor and Director of EPFL's Biomedical Imaging Group, Lausanne, Switzerland. His main research area is biomedical image processing. He has a strong interest in sampling theories, multiresolution algorithms, wavelets, the use of splines for image processing, and, more recently, stochastic processes. He has published about 250 journal papers on those topics. From 1985 to 1997, he was with the Biomedical Engineering and Instrumentation Program, National Institutes of Health, Bethesda, MD, conducting research on bioimaging and heading the Image Processing Group.

Dr. Unser is an EURASIP fellow (2009), and a member of the Swiss Academy of Engineering Sciences. He is the recipient of several international prizes including three IEEE-SPS Best Paper Awards and two Technical Achievement Awards from the IEEE (2008 SPS and EMBS 2010).



**John Paul Ward** received the B.S. degree in mathematics from the University of Georgia, Athens, and the Ph.D. degree in mathematics from Texas A&M University, College Station, in 2005 and 2010, respectively.

He is currently a postdoctoral researcher with the Biomedical Imaging Group, Swiss Federal Institute of Technology, Lausanne, Switzerland.