

From physical linear systems to discrete-time series. A guide for analysis of the sampled experimental data

Jakub Ślęzak* and Aleksander Weron†

Hugo Steinhaus Center, Wrocław University of Technology, 50-370 Wrocław, Poland

(Received 5 February 2015; published 11 May 2015)

Modeling physical data with linear discrete-time series, namely, the autoregressive fractionally integrated moving average (ARFIMA) model, is a technique that has attracted attention in recent years. However, this model is used mainly as a statistical tool only, with weak emphasis on the physical background of the model. The main reason for this lack of attention is that the ARFIMA model describes discrete-time measurements, whereas physical models are formulated using continuous-time parameters. In order to eliminate this discrepancy, we show that time series of this type can be regarded as sampled trajectories of the coordinates governed by a system of linear stochastic differential equations with constant coefficients. The observed correspondence provides formulas linking ARFIMA parameters and the coefficients of the underlying physical stochastic system, thus providing a bridge between continuous-time linear dynamical systems and ARFIMA models.

DOI: [10.1103/PhysRevE.91.053302](https://doi.org/10.1103/PhysRevE.91.053302)

PACS number(s): 05.10.Gg, 05.45.Tp

I. INTRODUCTION

Discrete time-series methods based on the autoregressive fractionally integrated moving average (ARFIMA) model [1–3], provide powerful and flexible statistical tools that are successful in analyzing data in econometrics, finance, and engineering [4–6]. Let us note that acronym ARFIMA is used interchangeably with FARIMA in the scientific literature. It is a model that fully describes the behavior of time series using a small number of parameters, which can be estimated from the data using well-established techniques and widely available statistical packages [2,7]. Moreover, these techniques allow for control of the estimation's quality, checking the correctness of the model or even forecasting future values of the time series. In recent years, new physical [8,9], biological [10], and medical [11] applications of the ARFIMA model were found, allowing for empirical description of complex systems with long- (powerlike), short- (exponential), and finite-range dependences [7,12] (see Fig. 1). Autoregressive moving average processes were also studied as models of physical data governed by discrete-time Langevin equations [13,14].

The main physical interpretation of this model was based on the fact that the ARFIMA model approximates processes such as fractional Brownian motion, Lévy stable motion [7,15], and the corresponding noises, whereas its special case, the ARMA model, can model properties of various stationary processes with finite or exponentially decaying memory [2,3]. However, most of these continuous-time processes themselves reflect the behavior of the process rather than its internal physical dynamics. On the other hand, the mathematical theory proposed by Phillips [16], and in recent years developed further by Brockwell *et al.* [17], established a connection between the ARMA model and a class of continuous-time stochastic dynamical systems. Here we show that the physical importance of these results is significant and, after suitable refinement, this

connection establishes a reliable physical basis for the ARMA and ARFIMA models.

II. THE ARFIMA MODEL

The studied model ARFIMA(p, d, q) states that the considered time series X_n fulfills the recursive relation [7]

$$\Delta^d \left(X_n - \sum_{k=1}^p \phi_k X_{n-k} \right) = \xi_n + \sum_{j=1}^q \theta_j \xi_{n-j}, \quad (1)$$

where ϕ_k, θ_j, d are deterministic coefficients and ξ_n is white noise that generates the stochastic dynamics; it is Gaussian or non-Gaussian, e.g., α stable [7,18], which determines the distribution of X_n . Equation (1) is comprised of three basic building blocks: the autoregressive (AR) part (the term in parentheses on the left-hand side), the fractionally integrated (FI) part (the operator Δ^d), and the moving average (MA) part (the term on the right-hand side). Each of these blocks models a different type of memory and has a distinct interpretation. If no memory is present we deal with ARFIMA(0,0,0), which is a white noise: $X_n = \xi_n$.

The term in parentheses on the left-hand side of Eq. (1) is AR(p) (the AR part), in which coefficients ϕ_k determine how the present value of the time series X_n depends linearly on the past values X_{n-k} ; in other words, it models the internal dynamics of the system. Because this dynamics is linear, it describes the exponential components of the memory. The most basic process from this class is called AR(1) or ARMA(1,0), which is the simplest exponential memory process with the correlation function [2]

$$\rho_X(k) = \frac{\langle X_n X_{n+k} \rangle}{\sqrt{\langle X_n^2 \rangle \langle X_{n+k}^2 \rangle}} = e^{-\phi_1 k}. \quad (2)$$

More general AR(p) models have memory functions that are mixtures of exponential decays [2]. These processes have great importance for statistics because AR(p) are maximal entropy processes for the fixed first $p + 1$ values of the correlation function [19].

*jakub.slezak@pwr.edu.pl

†aleksander.weron@pwr.edu.pl

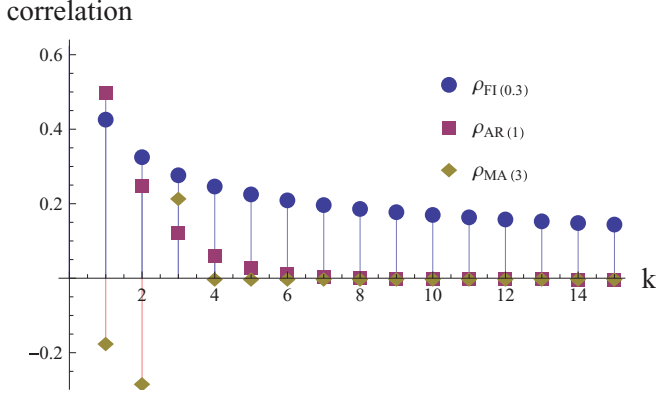


FIG. 1. (Color online) Power-law (circles), exponential (squares), and finite (diamonds) memory functions of the ARFIMA processes.

The right-hand side of Eq. (1) is the $MA(q)$ process, which determines the external dynamics through modification of the white noise. It models finite-range components of the memory that depend on coefficients θ_j . The actual value of the $MA(q)$ process depends only on the q last values of the generating noise ξ_n and because of that it does not contain any information about the history of the process older than $q\Delta t$. For example, the process called $MA(1)$ or $ARMA(0,1)$ is the simplest type of colored noise with ultrashort memory and the correlation function $\rho_{MA(1)}(1) = \theta_1$, $\rho_{MA(1)}(k) = 0$ for $k > 1$ [2,20].

The last part of Eq. (1), the operator Δ^d denoting the FI part, reflects both nonstationarity and fractional memory. The symbol Δ denotes the discrete difference operator $\Delta X_n = X_{n+1} - X_n$. When d is a natural number, the nonstationary process $ARFIMA(p,d,q)$, in this case also called $ARIMA(p,d,q)$, is understood as a process that after d differentiations is stationary $ARMA(p,q)$. A basic example is $ARFIMA(0,1,0)$, which is summed white noise. It is a sampled trajectory of Brownian motion.

In a situation when d is real, it can be decomposed into a natural-number part d_n and a fractional remainder d_f , $-1/2 \leq d_f \leq 1/2$, such that $\Delta^d = \Delta^{d_n} \Delta^{d_f}$. This remainder accounts for the power-law type of memory common, e.g., for anomalous diffusion [21]. This operator is understood as a series [7]

$$\Delta^{d_f} X_n = \sum_{k=0}^{\infty} \frac{d_f(d_f - 1) \cdots (d_f - k + 1)}{k!} (-1)^k X_{n-k}. \quad (3)$$

Applying Δ^{-d_f} to both sides of Eq. (1), it can be confirmed that $ARFIMA(p,d,q)$ can be interpreted as a modification of $ARFIMA(p,d_n,q)$ generated by noise $\Delta^{-d_f} \xi_n$. This noise, which is denoted by $FI(d_f)$ or $ARFIMA(0,d_f,0)$, is a stationary process with a power-law memory function that has a tail $\sim n^{2d_f-1}$; when ξ_n are Gaussian this time series is very similar to the fractional Brownian noise [15,22,23].

III. CONTINUOUS- VERSUS DISCRETE-TIME PROCESSES IN EXPERIMENTS

A continuous-time process $X(t)$ in a natural way contain much more information than its discrete-time counterpart

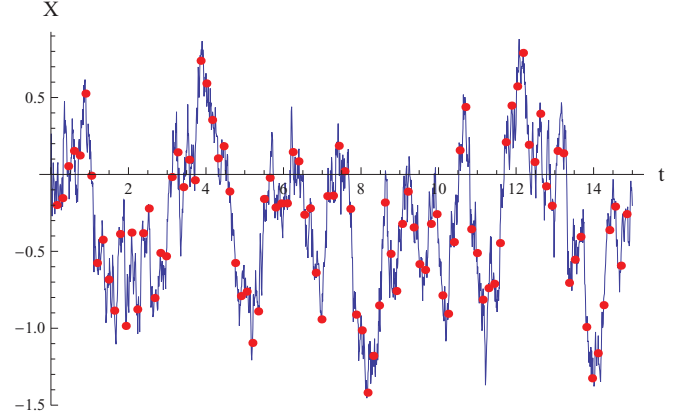


FIG. 2. (Color online) Continuous-time process (blue line) and sampled process (red dots) for the trajectory of the stochastic harmonic oscillator.

$X_n = X(n\Delta t)$ (see Fig. 2). During sampling we lose information, e.g., about the geometrical properties of trajectories. Memory functions become discrete and do not contain information about the dependence within intervals smaller than the sampling rate Δt . However, only in discrete case we can define some more refined memory functions, such as the partial autocorrelation function, which is a correlation of X_n and X_{n+k} with the influence of all in-between X_{n+j} removed [2,20]. In many cases it has a simple form that leads to greater usability.

Some state functions, such as power spectral density (PSD), differ considerably for discrete time and continuous time. These are memory functions of the considered process in the Fourier space. Continuous time PSD (CPSD) of the process X , f_X , is most easily defined as a Fourier transform of the covariance function

$$f_X(\omega) = \int_{-\infty}^{\infty} d\tau \langle X(t)X(t+\tau) \rangle e^{-i\omega\tau}. \quad (4)$$

However, in contrast to the PSD f_X , discrete-time power spectral density (DPSD) is a Fourier series of the sampled covariance function

$$f_X^{\Delta t}(\omega) = \frac{1}{\Delta t} \sum_{k=-\infty}^{\infty} \langle X_n X_{n+k} \rangle e^{-i\omega\Delta t k} \quad (5)$$

and is a periodic function that repeats after $2\pi/\Delta t$. The relation between these two functions is given by the Poisson summation formula [24,25]. It allows us to calculate DPSD numerically or analytically given the CPSD, but for the processes considered in this paper it is not a very practical tool. However, using the Poisson summation formula, one can prove that the DPSD converges to the CPSD as $\Delta t \rightarrow 0$ [20,24],

$$\lim_{\Delta t \rightarrow 0} f_X^{\Delta t} = f_X; \quad (6)$$

this fact can be interpreted as convergence of the time series X_n to the process $X(t)$ as $\Delta t \rightarrow 0$ (this limit is often called infill asymptotics). However, in realistic conditions, we often are far away from this limit and only the discrete-time model properly reflects the behavior of the observed system. Note also that all distortions of the data caused by the measurement equipment are changes of the sampled series X_n as this is

the object that is actually processed by the hardware. Thus, accounting for unwanted effects such as blur or measurement noise must be performed in the discrete-time setting [26]. In order to perform this procedure, a discrete-time model of the undistorted observations is needed.

IV. LINEAR DYNAMICAL SYSTEMS

To link the continuous- and discrete-time processes using the ARFIMA model we consider a linear stochastic system with the N -dimensional state vector $\mathbf{S}(t) = [S^1(t), S^2(t), \dots, S^N(t)]^T$, which evolves in continuous time according to the stochastic differential equation of the first order [27]

$$\frac{d}{dt}\mathbf{S}(t) = A\mathbf{S}(t) + \mathbf{F}(t), \tag{7}$$

where A is an $N \times N$ matrix with constant coefficients and $\mathbf{F}(t)$ is some stationary noise, acting as a random force. This general model describes a class of systems with a time-independent environment and additive stochastic disturbance. Note that if we study a state vector $\mathbf{S}(t)$ in which some coordinates $S^i(t)$ are described by differential equations of order bigger than one, we can complement the state vector $\mathbf{S}(t)$ by auxiliary coordinates $\{\frac{d}{dt}S^i(t), \frac{d^2}{dt^2}S^i(t), \dots\}$ and also reduce the problem to the form (7). One well-known property described by Eq. (7) is the position of a particle trapped in the harmonic potential within liquid [28]

$$m \frac{d^2}{dt^2}X(t) = -\kappa X(t) - \beta \frac{d}{dt}X(t) + F(t), \tag{8}$$

where m is the mass of particle, κ is the stiffness of the harmonic trap, β is the friction coefficient of the liquid, and $F(t)$ is white noise modeling the exchange of momenta with surrounding particles. A phase plot of this equation is shown in Fig. 3. Other examples include evolution of the charge $Q(t)$ in a linear RLC circuit disturbed by the noise electromotive

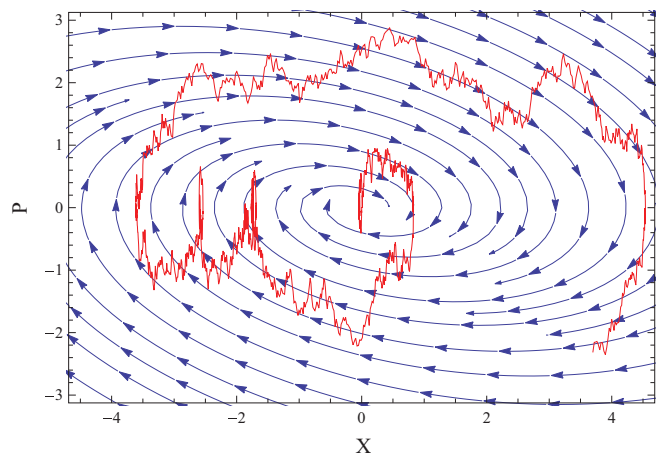


FIG. 3. (Color online) Phase plot of Eq. (8) (blue lines with arrows) for coordinates X and P and the stochastic solution of Eq. (8) (red line) for $m = 1$, $\kappa = 1/4$, and $\beta = 1/4$.

force $\mathcal{E}(t)$,

$$L \frac{d^2}{dt^2}Q(t) + R \frac{d}{dt}Q(t) + \frac{1}{C}Q(t) = \mathcal{E}(t), \tag{9}$$

as well as other types of linear disturbed circuits [29], harmonic heat bath models [30], and Brownian magnetic particles in a constant magnetic field [31].

V. TIME DISCRETIZATION PROCEDURE (SAMPLING)

In any of these cases the real system evolves in continuous time, but the experimental observations must be discrete and usually have the form of a time series \mathcal{S}_n sampled with constant sampling time Δt : $\mathcal{S}_n = \mathbf{S}(n\Delta t)$. The necessary condition for $\mathbf{S}(t)$ and \mathcal{S}_n to be stationary is for the matrix A to be negative definite, in other words, it needs to have eigenvalues with a negative real part. In such a case there exists a stationary solution of (7) given by convolution of the force $\mathbf{F}(t)$ with the matrix exponential e^{At} [27],

$$\mathbf{S}(t) = \int_{-\infty}^t ds e^{A(t-s)} \mathbf{F}(s). \tag{10}$$

From elementary properties of the matrix exponential and integration it follows that

$$\begin{aligned} \mathbf{S}(n\Delta t) &= e^{A\Delta t} \int_{-\infty}^{(n-1)\Delta t} ds e^{A(t-s)} \mathbf{F}(s) \\ &+ \int_{(n-1)\Delta t}^{n\Delta t} ds e^{A(n\Delta t-s)} \mathbf{F}(s), \end{aligned} \tag{11}$$

that is, the sampled process \mathcal{S}_n fulfills the equation

$$\mathcal{S}_n = E\mathcal{S}_{n-1} + \mathfrak{E}_n. \tag{12}$$

The obtained recursive formula is a vector counterpart of the process AR(1) and is called VAR(1) [2]. The VAR(1) process \mathcal{S}_n can be explicitly expressed in terms of the generating noise \mathfrak{E}_n as

$$\mathcal{S}_n = \sum_{k=0}^{\infty} E^k \mathfrak{E}_{n-k}. \tag{13}$$

This equation is a discrete counterpart of the convolution formula (10).

From Eq. (13) it follows that neither matrix A directly nor all values of $\mathbf{F}(t)$ affect the state \mathcal{S}_n ; they do this only through their discrete-time counterparts

$$E = e^{A\Delta t}, \quad \mathfrak{E}_n = \int_{(n-1)\Delta t}^{n\Delta t} ds e^{A(n\Delta t-s)} \mathbf{F}(s). \tag{14}$$

The first of Eqs. (14) defines the discretized evolution operator E , which is responsible for the deterministic part of the transition from state \mathcal{S}_{n-1} to \mathcal{S}_n . The stochastic deviation from the deterministic path is described by the second of Eqs. (14), which defines the packing operator $\mathbf{F}(t) \mapsto \mathfrak{E}_n$. The whole influence of the process $\mathbf{F}(t)$ on \mathcal{S}_n is fully determined by the packed force \mathfrak{E}_n . Each value of \mathfrak{E}_n gathers values of $\mathbf{F}(t)$ from the interval $((n-1)\Delta t, n\Delta t)$. If $\mathbf{F}(t)$ is stationary, also the discrete noise \mathfrak{E}_n is stationary. The packed force process inherits most of the properties of the underlying continuous-time $\mathbf{F}(t)$: the same type of distribution (Gaussian

$F(t)$ gives Gaussian Ξ_n and α -stable $F(t)$ gives α -stable Ξ_n) and the same type of memory (white noise, finite range, exponential, or power law).

Because of the above correspondence, the statistical methods available for the VAR(1) model [32] can be used in physical applications. Using estimators for the matrix E and taking matrix logarithm, we obtain estimates for the underlying matrix $A = \ln(E)/\Delta t$. The properties of force $F(t)$ can be studied through analysis of Ξ_n , which can be estimated as $\Xi_n = S_n - ES_{n-1}$.

VI. SINGLE-COMPONENT ANALYSIS

The above approach has a huge disadvantage: it requires that we can observe the whole sampled state vector S_n , which is unrealistic for more complex systems. To avoid this difficulty we want to analyze the behavior of one chosen component of S_n ; without loss of generality we will assume that it is S_n^1 . The evolution of this component is coupled with the evolution of the rest of the state vector by the action of nondiagonal components of the evolution matrix E .

In order to decouple the component S_n^1 , we use the Cayley-Hamilton theorem, which states that if an $N \times N$ matrix E has a characteristic equation $p(\lambda) = \det(E - \lambda I) = \lambda^N - \sum_{k=1}^N \phi_k \lambda^{N-k} = 0$, then the matrix E itself fulfills this equation, that is,

$$p(E) = E^N - \sum_{k=1}^N \phi_k E^{N-k} = 0. \quad (15)$$

This polynomial equation of order N has coefficients ϕ_k determined by the eigenvalues λ_i of the matrix E ,

$$\phi_k = (-1)^{k+1} \sum_{D_k} \prod_{i \in D_k} \lambda_i, \quad (16)$$

where D_k denotes the family of all k -element subsets of the set $\{1, 2, \dots, N\}$. The eigenvalues ν_i of matrix A from (7) and the eigenvalues of matrix E are related by formula $\lambda_i = e^{\nu_i \Delta t}$. Therefore, E fulfills (15) with the coefficients

$$\phi_k = (-1)^{k+1} \sum_{D_k} \exp\left(\Delta t \sum_{i \in D_k} \nu_i\right). \quad (17)$$

As a preparation for using the Cayley-Hamilton theorem, we recursively use (12) and express the variables S_{n-k} as functions of S_{n-N} ,

$$\begin{aligned} S_{n-N} &= S_{n-N}, \\ S_{n-N+1} &= ES_{n-N} + \Xi_{n-N+1}, \\ S_{n-N+2} &= E^2 S_{n-N} + E \Xi_{n-N+1} + \Xi_{n-N+2}, \\ S_{n-N+3} &= E^3 S_{n-N} + E^2 \Xi_{n-N+1} + E \Xi_{n-N+2} + \Xi_{n-N+3}, \\ &\vdots \\ S_n &= E^N S_{n-N} + \sum_{j=0}^{N-1} E^j \Xi_{n-j}, \end{aligned} \quad (18)$$

thus coupling them to only this one variable. After multiplying these equations by ϕ_k in order to obtain terms $\phi_k E^{N-k} S_{n-N}$

and subtracting them from the last one, we remove all action of the matrix E on the time series S_n using the equality $(E^N - \sum_{k=1}^N \phi_k E^{N-k}) S_{n-N} = 0$ obtained from the Cayley-Hamilton theorem. The cost of this decoupling is performing complicated transformations of the discretized force Ξ_n along the way. The equation that we obtain after this procedure has the form

$$S_n - \sum_{k=1}^N \phi_k S_{n-k} = \Xi_n + \sum_{k=1}^{N-1} R_k \Xi_{n-k}. \quad (19)$$

The left-hand side of Eq. (19) is the AR(N) part described by the coefficients ϕ_k , which depend only on the deterministic matrix A . The left-hand side acts as an effective noise η_n ,

$$\eta_n = \Xi_n + \sum_{k=1}^{N-1} R_k \Xi_{n-k}, \quad (20)$$

which generates the stochastic dynamics of the vector S_n . The behavior of this noise is determined by the matrices R_k ,

$$R_k = E^k - \sum_{j=1}^k \phi_j E^{k-j}, \quad (21)$$

composed of mixtures of time-shifting operators E^{k-j} . They affect the evolution by mixing different Ξ_{n-k} ; as a result, the first component S_n^1 fulfills the recurrence relation

$$S_n^1 - \sum_{k=1}^N \phi_k S_{n-k}^1 = \eta_n^1. \quad (22)$$

There is a deep connection between this formula and the classical Mori-Zwanzig theory [30,33]. Equation (22) can be written in a slightly different manner

$$\frac{\Delta S_n^1}{\Delta t} - \sum_{k=1}^N \phi'_k S_{n-k}^1 = \frac{\eta_n^1}{\Delta t}, \quad (23)$$

where $\phi'_1 = (\phi_1 + 1)/\Delta t$, $\phi'_k = \phi_k/\Delta t$, and $k > 1$. Now it becomes clear that this is the discrete-time analog of the generalized Langevin equation [23,30,33]. It is no coincidence, as derivations in both cases use the same ideas, decoupling most of the coordinates of freedom, at the same time introducing the memory kernel and effective noise. Despite many similarities, this analogy has its limitations, e.g., there is no clear discrete-time equivalent of the fluctuation-dissipation theorem.

VII. ANALYSIS OF THE STOCHASTIC FORCE

The effective noise η_n^1 is a mixture of the various components of the last $N - 1$ vectors Ξ_n . If the underlying force $F(t)$ is white Gaussian noise, then the resulting η_n^1 are a mixture of Gaussian white noises that forms a time series with an $N - 1$ finite-range dependence. When analyzing only one component, we can ignore its internal structure and represent it as the MA($N - 1$) process

$$\eta_n^1 = \xi_n^1 + \sum_{j=1}^{N-1} \theta_j^1 \xi_{n-j}^1, \quad (24)$$

generated by a white noise ξ_n^1 , which is in fact the orthogonalized series η_n^1 [34]. Such orthogonalization can be always performed and the coefficients θ_j^1 may be obtained by solving the system of equations resulting from comparing the covariance function of the process defined by the left-hand side of the Eq. (24) and the effective noise η_n^1 from Eq. (22) [2].

Thus, we arrive at the conclusion that the coefficient S_n^1 is the ARMA($N, N - 1$) process with the coefficients ϕ_k and θ_j^1 . Similar statements hold for more complex models of the force $F(t)$.

If the force $F(t)$ has a finite range of memory smaller than $K\Delta t$, then η_n^1 by the same reasoning as above can be regarded as the MA($N - 1 + K$) process. In such a case S_n^1 is ARMA($N, N - 1 + K$).

If the force $F(t)$ has power-law memory tails $\sim t^{2d_f-1}$, then η_n^1 is a composition of finite-range mixing introduced by the operators R_k and the power-law behavior. The best approximation of such time series is FIMA(d_f, q), i.e., a process similar to (24), but where ξ_n^1 are FI(d_f) time series [7]. So the component S_n^1 is ARFIMA(N, d_f, q).

If the force $F(t)$ has exponential tails of memory, that is, if it can be represented in a form similar to (10), then we may treat it as a time-dependent state of the same class as $S(t)$, which confirms that η_n^1 is ARMA($L, L - 1 + N$) for some L ; added N results from the mixing of Ξ_{n-k} . The AR(L) part, understood as an operator, can be freely moved from acting on η_n^1 to acting on S_n^1 , leading to the ARMA($N + L, N + L - 1$) model. The operator AR($N + L$) can be easily calculated as the composition AR(L)AR(N).

In our considerations we assumed that $S(t)$ was stationary. However, we may observe the possibly nonstationary integral of the stationary coordinate $X(t) = \int_0^t d\tau S^1(\tau)$, $X(t)$ being position, charge, etc. In this case the process of differences $\Delta X_n = \int_{(n-1)\Delta t}^{n\Delta t} d\tau S^1(\tau)$ is ARFIMA($p, d_f, q + 1$), which follows from the fact that $S^1(\tau + n\Delta t)$ is some ARFIMA(p, d_f, q) for any τ , with p, d_f, q determined by the proper model from the description above. An increase by one order in the MA part accounts for additional short-time memory introduced by the integral $\int_{(n-1)\Delta t}^{n\Delta t} ds$ and so X_n is ARFIMA($p, d_f + 1, q + 1$).

An analogous statement holds for all other components S_n^i . The AR coefficients ϕ_k are identical for all of them and the MA coefficients θ_j^i and noises ξ_n^i from Eq. (24) vary. The equations governing the evolution of different S_n^i were decoupled; however, these components are dependent because ξ_n^i are mixtures of the Ξ_n components, so they are dependent set of variables with respect to i .

We stress that all modeling is performed at the level of the discretized stochastic force; the obtained ARFIMA model of the observed coordinate is by no means phenomenological as is often the case for discrete-time models, but derived from the theory of the continuous-time dynamical system. For all cases except the power-law memory, the correspondence is exact; in the latter case the FIMA approximation must be made for the discretized stochastic force and the AR part is still exact. As the FI part can reflect any type of power-law long-time memory asymptotics and the MA part can account for any finite-range deviations, such a model is most often sufficient [7,10].

In our reasoning we used the fact that the Gaussian process is fully determined by its covariance structure at the moment when we orthogonalized the effective noise series η_n^1 . For non-Gaussian processes it is no longer true, as they can have richer than linear memory structure [18]. Therefore, for non-Gaussian forces $F(t)$, the obtained ARFIMA model reflects only the linear aspect of the memory. In this case it is approximate, but has the same autocovariance and power spectral density as the original process.

VIII. PARTICLE IN A HARMONIC POTENTIAL

Let us return to Eq. (8), which is the second-order differential equation describing the particle trapped in a harmonic potential. An approximation in which the inertial term $m d^2X/dt^2$ is considered negligible simplifies analysis. In such conditions the state of the particle evolves according to the force-balance equation $\beta dX/dt = -\kappa X + \xi$. Its stationary solution

$$X(t) = \frac{1}{\beta} \int_{-\infty}^t ds e^{-(\kappa/\beta)(t-s)} \xi(s) \quad (25)$$

is called the Ornstein-Uhlenbeck process [35,36] and has a well-known Lorentzian continuous-time power spectral density [26,28]

$$f_{OU}(\omega) = \frac{1}{\beta^2} \frac{\sigma^2}{(\kappa/\beta)^2 + \omega^2}, \quad (26)$$

where σ^2 is the variance of the noise $F(t)$; when the fluctuation-dissipation theorem holds $\sigma^2 = 2k_B T \beta$ [30]. The sampled trajectory of (25) is the AR(1) process with a coefficient $\phi_1 = e^{-\Delta t \kappa/\beta}$. The discrete-time power spectral density can be calculated from the general formula for all ARFIMA processes [3,24], which for AR(1) yields [26]

$$f_{AR(1)}^{\Delta t}(\omega) = (\phi_1^{-2} - 1) \frac{1}{2\kappa\beta} \frac{\sigma^2 \Delta t}{1 + \phi_1^2 - 2\phi_1 \cos(\omega \Delta t)}. \quad (27)$$

We can see that the CPSD and DPSD functions differ when Δt is not considerably smaller than β/κ (see Fig. 4), which is often the case for mesoscopic objects observed in normal conditions

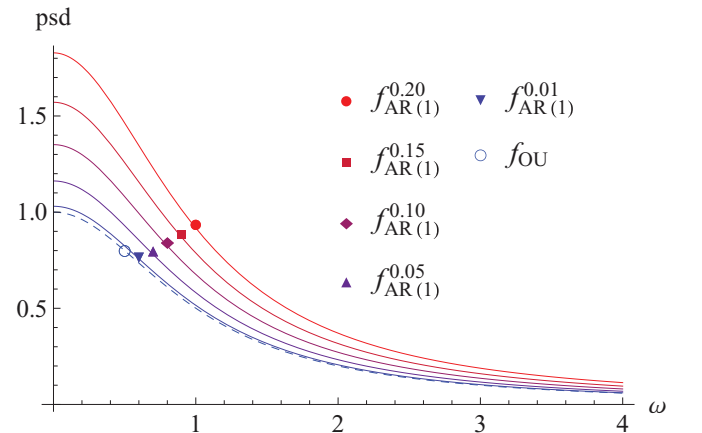


FIG. 4. (Color online) Comparison of the DPSD and CPSD of the process (25) for decreasing sampling times Δt in nondimensional units $\kappa = \beta = \sigma = 1$.

[26]. As we provide an exact formula for the observed spectral density, there is no reason to use approximate Eq. (26).

In the case when the mass is not negligible, the effects of sampling become more complex. The full state vector \mathbf{S} is then composed of position and momentum $\mathbf{S}(t) = [X(t), P(t)]^T$. The stochastic force affects the change of momenta $\mathbf{F}(t) = [0, F(t)]^T$. Note that because of the identical forms of Eqs. (8) and (9), all further results would follow also for the RLC circuit (9) after a simple change of letters. In this case the state vector would consist of charge and electric current.

Calculating the eigenvalues

$$v_{1,2} = -\frac{\beta}{2m} \pm \sqrt{\left(\frac{\beta}{2m}\right)^2 - \frac{\kappa}{m}} \quad (28)$$

of the evolution matrix $A = \begin{bmatrix} 0 & 1/m \\ -\kappa & -\beta/m \end{bmatrix}$, we obtain the AR coefficients of the sampled position process X_n , which, if $F(t)$ is a white noise, is ARMA(2,1) with AR(2) coefficients

$$\begin{aligned} \phi_1 &= 2 \exp\left(-\Delta t \frac{\beta}{2m}\right) \cosh\left(\Delta t \sqrt{\left(\frac{\beta}{2m}\right)^2 - \frac{\kappa}{m}}\right), \\ \phi_2 &= -\exp\left(-\Delta t \frac{\beta}{m}\right). \end{aligned} \quad (29)$$

The MA coefficient θ_1^1 is also determined by the calculated eigenvalues and is given by a complicated formula, but can

be easily calculated numerically. By estimating the AR(2) coefficients from the data, the ratios κ/m and β/m of the underlying process can be assessed and the parameter m can be subsequently estimated from the variance of the sampled process X_n . If $F(t)$ is not white noise, the MA part may differ and if $F(t)$ would have the power-law dependence it would be reflected in the FI part of the ARFIMA model.

IX. SUMMARY

In our work we tried to construct a bridge between continuous-time linear dynamical systems and the discrete-time ARMA model or, more generally, ARFIMA model. The studied correspondence for many cases might serve as a physical interpretation of the ARFIMA model and justification for its usage. Additionally, we have shown what order the physical ARFIMA model should have for a given dynamical system and we have given explicit formulas for its AR coefficients, which allows for estimation of the dynamical system's parameters using standard statistical tools. Its MA and FI coefficients can also be calculated, but they depend on the assumed model of the stochastic force. The coefficients of the ARFIMA model determine its characteristics, such as power spectral density, linking the basic dynamical system model with functions that can be estimated from the sampled data measured during experiment.

ACKNOWLEDGMENT

This research was supported by NCN Maestro Grant No. 2012/06/A/ST1/00258.

-
- [1] C. W. J. Granger and R. Joyeux, *J. Time Ser. Anal.* **1**, 15 (1980).
 - [2] P. Brockwell and R. Davis, *Time Series: Theory and Methods* (Springer, New York, 2006).
 - [3] J. Beran, *Statistics for Long-Memory Processes* (Chapman and Hall, New York, 1994).
 - [4] N. Crato and P. Rothman, *Econ. Lett.* **45**, 287 (1994).
 - [5] L. A. Gil-Alana, *International Review of Financial Analysis* **13**, 265 (2004).
 - [6] G. Fouskitakis and S. Fassois, *IEEE Trans. Signal Process.* **47**, 3365 (1999).
 - [7] K. Burnecki, G. Sikora, and A. Weron, *Phys. Rev. E* **86**, 041912 (2012).
 - [8] K. Burnecki, J. Klafter, M. Magdziarz, and A. Weron, *Physica A* **387**, 1077 (2008).
 - [9] F. D. Di Tolla and M. Ronchetti, *Phys. Rev. E* **48**, 1726 (1993).
 - [10] K. Burnecki, *J. Stat. Mech.* (2012) P05015.
 - [11] D. T. J. Liley, P. J. Cadusch, M. Gray, and P. J. Nathan, *Phys. Rev. E* **68**, 051906 (2003).
 - [12] T. Nakamura, Y. Hirata, and M. Small, *Phys. Rev. E* **74**, 041114 (2006).
 - [13] I. Horenko, F. Noe, C. Hartmann, and C. Schütte, *Phys. Rev. E* **76**, 016706 (2007).
 - [14] M. Niemann, T. Laubrich, E. Olbrich, and H. Kantz, *Phys. Rev. E* **77**, 011117 (2008).
 - [15] B. Sikdar and K. S. Vastola, *Proceedings of the 34th Conference on Information Sciences and Systems* (Princeton University Press, Princeton, 2000).
 - [16] A. W. Phillips, *Biometrika* **46**, 67 (1959).
 - [17] P. J. Brockwell, R. A. Davis, and Y. Yang, *Stat. Sin.* **17**, 63 (2007).
 - [18] P. Kokoszka and M. Taqqu, *J. Time Ser. Anal.* **15**, 203 (1994).
 - [19] J. Franke, *Adv. Appl. Probab.* **17**, 810 (1985).
 - [20] G. Box, G. Jenkins, and G. Reinsel, *Time Series Analysis: Forecasting and Control* (Prentice-Hall, Englewood Cliffs, 1994).
 - [21] R. Morgado, F. A. Oliveira, G. G. Batrouni, and A. Hansen, *Phys. Rev. Lett.* **89**, 100601 (2002).
 - [22] S. Burov and E. Barkai, *Phys. Rev. Lett.* **100**, 070601 (2008).
 - [23] S. C. Kou and X. S. Xie, *Ann. Appl. Stat.* **2**, 501 (2008).
 - [24] D. B. Percival and A. T. Walden, *Spectral Analysis for Physical Applications* (Cambridge University Press, Cambridge, 1993).
 - [25] M. Unser, *Proc. IEEE* **88**, 569 (2000).
 - [26] J. Ślęzak, S. Drobczyński, K. Weron, and J. Masajada, *Appl. Opt.* **53**, B254 (2014).
 - [27] K. Sobczyk, *Stochastic Differential Equations* (Kluwer Academic, Dordrecht, 1991).
 - [28] W. Coffey, *Adv. Chem. Phys.* **63**, 69 (1985).
 - [29] T. K. Rawat and H. Parthasarathy, *Int. J. Circ. Theor. Appl.* **38**, 259 (2010).
 - [30] R. Zwanzig, *J. Stat. Phys.* **9**, 215 (1973).
 - [31] B. Kursunoglu, *Ann. Phys. (NY)* **17**, 259 (1962).

- [32] B. Pfaff, *J. Stat. Softw.* **27**, 4 (2008).
- [33] H. Mori, *Prog. Theor. Phys.* **33**, 423 (1965).
- [34] F. Ansley, W. A. Spivey, and W. J. Wroblewski, *J. Econ.* **6**, 121 (1977).
- [35] G. E. Uhlenbeck and L. S. Ornstein, *Phys. Rev.* **36**, 823 (1930).
- [36] P. Protter, *Stochastic Integration and Differential Equations* (Springer, Berlin, 1990).