10.24425/acs.2025.153957

Archives of Control Sciences
Volume 35(LXXI), 2025
No. 1*, pages 19–46*

Random time-series model identification from binary-valued observations and quantized measurements

Jarosław FIGWER

In the paper, two algorithms that allow identification of a parametric models of random time-series from binary-valued observations of their realizations, as well as from quantized measurements of their values, are proposed. The proposed algorithms are based on the idea of time-series decomposition either on a direct power spectral density or autocorrelation function approximation. They use the concepts of randomized search algorithms to recover the corresponding parametric models from calculated estimates of power spectral density or autocorrelation function. The considerations presented in the paper are illustrated with simulated identification examples in which linear and nonlinear block-oriented dynamic models of time-series are identified from the binary-valued observations and quantized measurements.

Key words: parametric time-series identification, spectral factorization, time-series decomposition, linear time-series models, nonlinear block-oriented time-series models, binary-valued observations, quantized measurements

1. Introduction

Spectral analysis is a well-known tool used by engineers and researchers mainly to analyze the spectral contents of observed time-series. In the real world, this analysis is based on power spectral densities identified using time-series observations acquired from their finite-length realizations or the corresponding quantized measurements. Nowadays, due to the fast development of computational hardware, the functionality of spectral analysis was supplemented with

---

Author (e-mail: Jaroslaw.Figwer@polsl.pl) is with Department of Measurements and Control, Silesian University of Technology, Akademicka 16, 44-100 Gliwice, Poland.

The partial financial support of this research by The Polish Ministry of Science and Higher Education is acknowledged.

Received 7.11.2024. Revised 14.01.2025.

a new function used to recover linear and nonlinear block-oriented dynamic parametric models of time-series from the identified power spectral densities or autocorrelation functions [17]. From an algorithm implementation point of view, such a function is only a piece of software running on specialized hardware. The paper describes how this piece of software may be applied to the identification of linear and nonlinear block-oriented dynamic parametric models of time-series from the corresponding binary-valued observations or quantized measurements.

The binary-valued observations are common in modern digital world. Large datasets often contain pieces of information that can be interpreted as binary-valued observations of the corresponding time-series realizations. These observations can be used for identification of dynamic parametric time-series models. Extensive literature research has resulted in a few publications [1–3, 5, 6, 21–23, 25, 33] discussing this identification problem. These publications are devoted solely to the case of AR time-series parametric model identification from binary-valued observations without considering the nonlinearity of the data acquisition system implied by processing time-series realization values to binary-valued observations.

The existence of nonlinearity of data acquisition systems should also be taken into account when time-series are represented by measurements of their real-world values. Quantizers present in data acquisition systems process time-series values into measurement results in a nonlinear manner. In well-known and highly cited bibliography devoted to the identification of time-series models (see e.g. [4, 9, 20, 24, 26, 28, 30, 32]), existence of this nonlinearity is omitted. An extensive literature search concerning the identification of time-series models based on quantized measurements resulted in only three items of literature [1, 2, 14] discussing the identification of only AR time-series models.

In the case where the above-mentioned nonlinearities of data acquisition system are taken into account dynamic models of time-series may be obtained using two mathematical tools. The first tool is an algorithm of random time-series decomposition into uncorrelated components [16, 17]. This algorithm is intended only for the identification of linear dynamic parametric models of time-series. The second tool allows for the recovery of both linear and nonlinear dynamic parametric models of time-series directly from estimates of power spectral densities or autocorrelation functions [17]. Both tools use optimization algorithms that utilize ideas of randomized search [8, 10, 19, 27, 31] to estimate parameters of the considered time-series models. They allow for the identification of more general classes of dynamic parametric models of time-series than those considered in the above-mentioned literature.

The paper is organized as follows: (1) the problem of identifying linear and nonlinear dynamic parametric models of time-series from binary-valued observations or quantized measurements is stated; (2) two algorithms used for this identification are described; (3) the properties of the described algorithms are illustrated with simulated identification examples in which dynamic parametric linear and nonlinear block-oriented time-series models are recovered from the calculated periodograms or autocorrelation function estimates.

2. Problem statement

Let $y_B(i)$ be a binary-valued observation at the discrete-time instant i of a realization of a weak ergodic zero-mean real-valued random time-series $y_p(i)$ with the power spectral density $S_{y_p y_p}(\omega T)$ satisfying the condition $0 \leqslant S_{y_p y_p}(\omega T) < \infty$ for relative frequencies ωT in the range $[0, 2\pi)$. For the set $\{y_B(0), y_B(1), \ldots, y_B(N-1)\}$ of the binary-valued observations acquired, the corresponding estimate of the power spectral density is calculated, resulting in a set of values $\hat{S}_{y_B y_B}(\omega T)$ given for N relative frequencies $\omega T = \Omega n$, where and $n = 0, 1, \ldots, N-1$ and $\Omega = \dfrac{2\pi}{N}$. Subsequently, this set of values is used to recover a dynamic parametric model of the time-series $y_p(i)$ assuming that:

a) realizations of the time-series $y_p(i)$ may be obtained as a filtration of a random process $e_p(i)$, which cannot be measured, through an asymptotically stable linear or nonlinear block-oriented dynamic model;

b) for the random process $e_p(i)$, its power spectral density or autocorrelation function is known and can be used to generate its realizations;

c) the mentioned linear or nonlinear block-oriented dynamic model is parameterized by a vector Θ_p of unknown parameters whose values will be recovered from $\hat{S}_{y_B y_B}(\omega T)$;

d) having the values of vector Θ_p, the corresponding output values $y_m(\Theta_p, i)$ of the linear and nonlinear dynamic time-series model can be calculated for the simulated realizations of $e_p(i)$, and then these values can be transformed into the corresponding binary-valued observations.

The transformation of time-series values $y_p(i)$ or $y_m(\Theta, i)$ into the corresponding binary-valued observations is a nonlinear transformation defined as:

$$y_B(i) = \mathcal{F}(y_p(i)) = \begin{cases} up; & y_p(i) \geqslant 0, \\ lo; & y_p(i) < 0, \end{cases} \tag{1}$$

where up and lo are real-valued numbers satisfying the relation $up > lo$. This transformation implies that the acquired set of binary-valued observations of time-series represents infinitely many different linear and nonlinear block-oriented dynamic time-series models differing only in their steady-state gains the output [18]. The acquired binary-valued observations of the time-series $y_p(i)$ are the original time-series values distorted by the transformation $\mathcal{F}(\cdot)$. The time-series $y_B(i)$ after this transformation may be represented as a sum of the two components: one having values proportional to the values of time-series under transformation, and the other being a random time-series that is uncorrelated with the time-series under transformation [11–13, 16, 17]. This implies that the above-formulated identification problem is a time-series identification problem in which the corresponding acquired binary-valued observations are interpreted as disturbed measurements of a time-series having values proportional to the values of the original time-series. The level of these disturbances depends on the choice of up and lo values. It can be expressed as a signal–to–noise ratio defined as the ratio of the variance of the time-series under transformation to the corresponding variance of disturbances. Thus, models of time-series are identified in the case of a very low signal–to–noise ratio [18].

Quantized measurements $y_Q(i)$, at the discrete-time instant i, of a realization of the weak ergodic zero-mean real-valued random time-series $y_p(i)$ or $y_m(\Theta, i)$ are acquired using a data acquisition system equipped with a quantizer having saturation with limits: lower $Q_{\min}$ and upper $Q_{\max}$, i.e.:

- values of the time-series $y_p(i)$ or $y_m(\Theta, i)$ from the range $[Q_{\min}, Q_{\max}]$ after being processed by the quantizer result in values from a set having a finite number of elements;
- values of the time-series $y_p(i)$ or $y_m(\Theta, i)$ less than $Q_{\min}$ or greater than $Q_{\max}$ are replaced by $Q_{\min}$ or $Q_{\max}$, respectively.

The above remarks imply that the transformation $\mathcal{Q}(\cdot)$ of time-series values $y_p(i)$ or $y_m(\Theta, i)$ into the corresponding quantized measurements is a nonlinear transformation having similar properties to the transformation $\mathcal{F}(\cdot)$. Nonlinear distortions implied by the transformation $\mathcal{Q}(\cdot)$ may be modelled by a random time-series that is uncorrelated with the time-series under transformation, with properties depending on the power of time-series processed by the quantizer. The following three cases of the power of time-series under considerations are taken into account [14, 16, 17]:

- ultra low-power case – values $y_p(i)$ or $y_m(\Theta, i)$ are such that they are less than or only occasionally reach the plus or minus value of the data acquisition system quant. In this case, data acquisition is added by randomized quantization [7] – prior to the quantization, independent realizations of a random variable uniformly distributed in a range covering the data acquisition system quant are added to processed values $y_p(i)$ or $y_m(\Theta, i)$;

- normal-power case – values of $y_p(i)$ or $y_m(\Theta, i)$ are in the range $[Q_{\min}, Q_{\max}]$. In this case, the influence of the nonlinearity of the data acquisition system on processed values $y_p(i)$ or $y_m(\Theta, i)$ decreases with the increase in the power of the processed time-series;
- ultra high-power case – values $y_p(i)$ or $y_m(\Theta, i)$ are less than $Q_{\min}$ or greater than $Q_{\max}$ for long time slots. Rarely are they in the range $[Q_{\min}, Q_{\max}]$.

The nonlinear distortions implied by the data acquisition system in the above-enumerated cases may be modelled as a white (the ultra low- and normal-power case) or colored random process (the ultra high-power case).

The two algorithms described in the sequel for recovering a parametric linear or nonlinear block-oriented dynamic model of the time-series $y_p(i)$ from an estimate of the power spectral density of the corresponding binary-valued observations or quantized measurements are based on the following property: the binary-valued observations $y_B(i)$ or quantized measurements $y_Q(i)$ are a nonlinear dynamic transformation of $e_p(i)$ that may be decomposed into two uncorrelated components – the first component having values proportional to these resulting from a linear or nonlinear filtration of $e_p(i)$, and the second component representing the nonlinear transformation defining the acquisition of binary-valued observations or quantized measurements [11–13, 16, 17]. It follows from this property that the power spectral density estimate $\hat{S}_{y_B y_B}(\omega T)$ or $\hat{S}_{y_Q y_Q}(\omega T)$ contains a component implied by the linear or nonlinear block-oriented dynamic model of the observed or measured time-series and may be used to recover the values of the vector Θ_p. It is worth emphasizing that for this purpose, the corresponding estimate of autocorrelation function may also be applied.

3. Model recovery using a random time-series decomposition

The first algorithm described in the paper is devoted to the identification of only linear dynamic models (AR, MA or ARMA) of time-series. Additionally, it is assumed that nonlinear distortions implied by the transformation $\mathcal{F}(\cdot)$ or $Q(\cdot)$ are sufficiently precisely modelled by a linear dynamic model of a time-series. This implies that the time-series $y_B(i)$ or $y_Q(i)$ may be regarded as a mixture of two components being ARMA time-series $y_1(i)$ and $y_2(i)$, i.e.:

$$y_B(i) = y_1(i) + y_2(i) \tag{2}$$

or

$$y_Q(i) = y_1(i) + y_2(i), \tag{3}$$

where $y_1(i)$ is a time-series having values proportional to those of $y_p(i)$ and $y_2(i)$ represents the aforementioned nonlinear distortions. The r-th ($r = 1, 2$)

component $y_r(i)$ of the mixture $y_B(i)$ or $y_Q(i)$ is defined in the time-domain as:

$$y_r(i) = \frac{\sum_{n=0}^{dB_r} b_{r,n} z^{-n}}{\sum_{n=0}^{dA_r} a_{r,n} z^{-n}} e_r(i), \tag{4}$$

where z^{-1} is the one-step backward shift operator, $e_r(i)$ is a white noise random process with zero mean and unit variance, dA_r and dB_r are the r-th ARMA component structural numbers being non-negative integers and $\Theta_r = [a_{r,0}, a_{r,1}, \ldots, a_{r,dA}, b_{r,0}, b_{r,1}, \ldots, b_{r,dB}]$ is a vector of the corresponding ARMA time-series parameters. It is assumed that white noises $e_1(i)$ and $e_2(i)$ are uncorrelated and for $r = 1, 2$:

- the parameter $a_{r,0}$ is greater than 0;
- the structural numbers dA_r and dB_r are known;
- the polynomials

  $$z^{dA_r} \sum_{n=0}^{dA_r} a_{r,n} z^{-n} \tag{5}$$

  and

  $$z^{dB_r} \sum_{n=0}^{dB_r} b_{r,n} z^{-n} \tag{6}$$

  have all roots inside the unit circle.

The decomposition of the mixture $y_B(i)$ or $y_Q(i)$ involves finding the values of the parameter vector $\boldsymbol{\Theta} = [\Theta_1,\ \Theta_2]$ by minimizing the following objective function:

$$\mathcal{S}_1(\Theta) = \sum_{n=0}^{N-1} \left(\hat{S}_1(\Omega n) - \widetilde{S}_1(\Theta, \Omega n) \right)^2 \tag{7}$$

with constraints on the placement the roots of all polynomials (5) and (6) inside the unit circle. In the above objective function $\hat{S}_1(\Omega n)$ is equal to $\hat{S}_{y_B y_B}(\omega T)$ or $\hat{S}_{y_Q y_Q}(\omega T)$ and the values $\widetilde{S}_1(\Theta, \Omega n)$ $(n = 0, 1, \ldots, N-1)$ denote estimates of the power spectral density of the mixture $y_B(i)$ or $y_Q(i)$ calculated using the corresponding model presented in equation (4). These estimates are calculated using the definition of the power spectral density of ARMA time-series and FFT algorithm, i.e.:

$$\widetilde{S}_1(\Theta,\Omega n) = \sum_{r=1}^{2} \left| \frac{\sum\limits_{n=0}^{dB_r} \hat{b}_{r,n} e^{-j\Omega n}}{\sum\limits_{n=0}^{dA_r} \hat{a}_{r,n} e^{-j\Omega n}} \right|^2, \tag{8}$$

where $\hat{a}_{r,0}, \hat{a}_{r,1}, \ldots, \hat{a}_{r,dA}, \hat{b}_{r,0}, \hat{b}_{r,1}, \ldots, \hat{b}_{r,dB}$ for $r = 1, 2$ denote estimates of the corresponding parameters of ARMA time-series models. To find the global minimum of $\mathcal{S}_1(\Theta)$, it is advisable to use randomized search optimization algorithms.

4. Model recovery based on a direct power spectral density approximation

In the second proposed algorithm, the nonlinear distortions implied by the transformation $\mathcal{F}(\cdot)$ or $\mathcal{Q}(\cdot)$ are incorporated into the time-series model identification at the stage of the corresponding simulation of realizations using a time-series model.

The vector Θ_1 of unknown parameters of a linear and nonlinear block-oriented dynamic time-series model may be recovered from a calculated power spectral density estimate by minimizing the following objective function:

$$\mathcal{S}_2(\Theta_1) = \sum_{n=0}^{N-1} \left(\hat{S}_1(\Omega n) - \widetilde{S}_{1m}(\Theta_1, \Omega n) \right)^2, \tag{9}$$

with constraints on the stability of linear or nonlinear block-oriented dynamic time-series models taken into account. In the above objective function $\hat{S}_1(\Omega n)$ is equal to $\hat{S}_{y_B y_B}(\omega T)$ or $\hat{S}_{y_Q y_Q}(\omega T)$ and the values $\widetilde{S}_{1m}(\Theta, \Omega n)$ $(n = 0, 1, \ldots, N-1)$ denote values of the power spectral density estimate calculated for the chosen Θ_1 based on a set of binary-valued observations or quantized measurements of a simulated output $y_m(\Theta_1, i)$ of the corresponding linear or nonlinear block-oriented dynamic time-series model. This simulated output is obtained by exciting the linear or nonlinear block-oriented dynamic time-series model with a simulated realization of $e_1(i)$. It is obvious that the objective function $\mathcal{S}_2(\Theta_1)$ may have many local minima. It follows from the author's experience that to obtain the global minimum of $\mathcal{S}_2(\Theta_1)$, it is worth using optimization algorithms from a family of randomized search algorithms. It should also be mentioned here that in the same way, values of the vector Θ_1 elements may be obtained when, in equation (9) instead of the power spectral density estimates $\hat{S}_1(\Omega n)$ and $\widetilde{S}_{1m}(\Theta_1, \Omega n)$, the corresponding estimates $\hat{R}_1(\tau)$ and $\widetilde{R}_{1m}(\Theta_1, \tau)$ $(\tau = 0, 1, \ldots, N-1)$ of autocorrelation functions are used. This defines a model recovery algorithm based on a direct autocorrelation function approximation. The above-described two algorithms of model recovery based on direct power spectral density or autocorrelation function approximation are a spectral factorization in the case of nonlinear dynamic time-series models [17].

5. Simulated identification examples

In the simulated identification examples presented below, ARMA models as well as Wiener and Wiener–Hammerstein block-oriented dynamic systems were chosen as dynamic time-series models. Minimizations of the objective function $\mathcal{S}_1(\Theta)$ were done using an optimization algorithm [15] based on random search aided by a local optimization made using MATLAB *lsqcurvefit* algorithm [29]. Minimizations of the objective function $\mathcal{S}_2(\Theta_1)$ were performed using a randomized optimization algorithm based on a random search for the minimum in a hypercube surrounding the origin of the coordinate system. The length of the hypercube edge was chosen randomly. This search was aided by additional local optimizations, which were also random searches but done in shrinking hypercubes surrounding potential solutions of the optimization problem. Objective functions $\mathcal{S}_1(\Theta)$ and $\mathcal{S}_2(\Theta_1)$ have infinitely many minima in each hypercube surrounding the origin of the coordinate system when all coefficients of polynomials being numerators and denominators of transfer functions defining linear dynamic models of time-series or linear components of nonlinear block-oriented dynamic models of time-series are recovered. Such parameterization of the mentioned transfer functions accelerates the minimization of the objective functions $\mathcal{S}_1(\Theta)$ and $\mathcal{S}_2(\Theta_1)$ and was used in the simulated identification examples described below. Additionally, it was assumed that structures of time-series models recovered in the simulated identification experiments are the same as the structures of simulated time-series.

Power spectral density estimates $\hat{S}_1(\Omega n)$, $\widetilde{S}_1(\Theta, \Omega n)$ and $\widetilde{S}_{1m}(\Theta, \Omega n)$ were calculated using MN values of the corresponding time-series realizations. As an estimator of the power spectral density an averaged periodogram (see e.g. [20]) was used, i.e. MN-sample data segment was divided into M non-overlapping N-sample data segments, and for the each N-sample data segment, the periodogram was calculated. Finally, the obtained periodograms were averaged for the relative frequencies Ωn ($n = 0, 1, \ldots, N - 1$), resulting in $\hat{S}_1(\Omega n)$, $\widetilde{S}_1(\Theta, \Omega n)$ and $\widetilde{S}_{1m}(\Theta, \Omega n)$. Patterns of power spectral densities for linear dynamic time-series models were calculated using the component number 1 from equation (4). As a pattern of the power spectral density of time-series generated using nonlinear block-oriented dynamic model averaged periodograms of $y_m(i)$ calculated for $2000M$ non-overlapping N-sample data segments were used. Calculation of all above-mentioned periodograms was done assuming a sampling interval value equal to 1.0000 [s]. During model identification using random time-series decomposition, uncorrelated white noises with variances equal to 1.0000 were applied as the random processes $e_1(i)$ and $e_2(i)$. Realizations of $e_1(i)$ being Gaussian white noise with a variance equal to 1.0000, used in the direct power spectral density or autocorrelation function approximation, were generated using a random number generator implemented in the MATLAB environment.

5.1. ARMA model identification using random time-series decomposition – binary-valued observations

The first simulated identification example was devoted to the model identification of an ARMA time-series defined by the following relation:

$$y_p(i) = \frac{1.0000 - 0.8000z^{-1}}{1.0000 - 1.5000z^{-1} + 0.7000z^{-2}} e_p(i), \tag{10}$$

where $e_p(i)$ was Gaussian white noise with a variance $\sigma^2_{e_p e_p}$ equal to 36.0000. Using this relation, 100 realizations of $y_p(i)$ were generated assuming $M = 2000$ and $N = 512$. The values of $y_p(i)$ were transformed to the corresponding binary-valued observations $y_B(i)$ by the transformation $\mathcal{F}(\cdot)$ with $up = 5.0000$ and $lo = -5.0000$. For each acquired realization $y_B(i)$ $(i = 1, 2, \ldots, MN - 1)$, the model identification was performed using the random time-series decomposition, assuming that the nonlinear distortions implied by the transformation $\mathcal{F}(\cdot)$ are modelled by an ARMA time-series $y_2(i)$ with structural numbers $dA_2 = dB_2 = 1$. As a search space during ARMA model recovery, a hypercube of dimension 9 surrounding the origin of the coordinate system was chosen. This means that for each MN-sample realization of $y_B(i)$, estimates of 9 parameters of the two ARMA models $y_1(i)$ and $y_2(i)$ were calculated. The obtained mean value of the minimum of the objective function $\mathcal{S}_1(\Theta)$ was equal to 275.6. Using 100 estimates of parameters of the ARMA time-series $y_1(i)$, the corresponding estimates of the power spectral density were calculated and compared with the pattern in Fig. 1. This pattern was calculated using the model presented in equation (10). In Fig. 2 the corresponding normalized, by maximum values, power spectral density of the pattern and of the mean value of calculated power spectral density estimates are presented. In Figs. 3 and 4, normalized in the same way, calculated power spectral density estimates, power spectral density of the pattern, and the mean value of calculated normalized power spectral density estimates are presented.

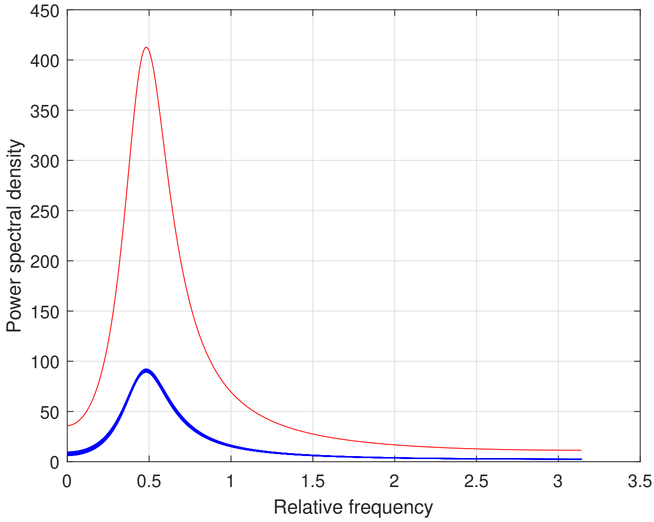

Figure 1: Power spectral densities calculated from 100 identified models of ARMA time-series represented by binary-valued observations (blue lines) and the pattern (red line) – random time-series decomposition, $M = 2000$ and $N = 512$

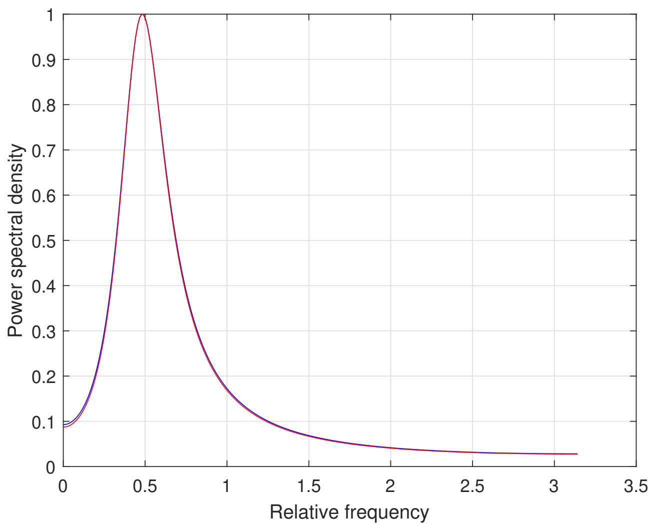

Figure 2: Normalized mean value of 100 power spectral densities calculated from identified models of ARMA time-series represented by binary-valued observations (blue line) and the normalized pattern (red line) – random time-series decomposition, $M = 2000$ and $N = 512$

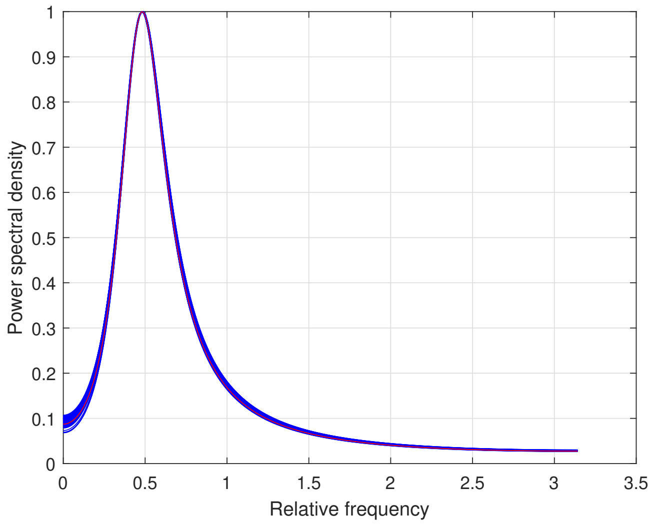

Figure 3: Normalized power spectral densities calculated from 100 identified models of ARMA time-series represented by binary-valued observations (blue lines) and the normalized pattern (red line) – random time-series decomposition, $M = 2000$ and $N = 512$

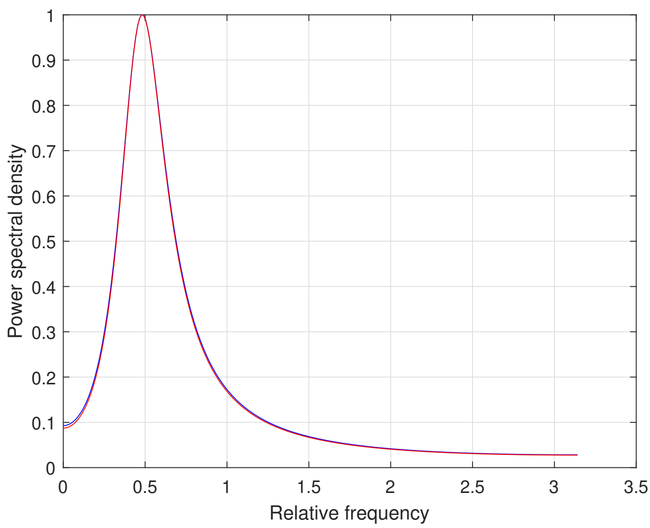

Figure 4: Mean value of 100 normalized power spectral densities calculated from identified models of ARMA time-series represented by binary-valued observations (blue line) and the normalized pattern (red line) – random time-series decomposition, $M = 2000$ and $N = 512$

5.2. ARMA model identification using direct power spectral density approximation – binary-valued observations

In the second simulated identification example, the above-described experiment was repeated with some changes. Now the variance $\sigma^2_{e_p e_p}$ was equal to 1.0000, up was equal to 1.0000, lo was equal to −1.0000, and N was equal to 128. Using the direct power spectral density approximation, 100 estimates of parameters of ARMA time-series $y_1(i)$ were recovered and used to calculate the corresponding normalized estimates of the power spectral density. In Fig. 5, the results of these calculations are compared with the normalized pattern. In Fig. 6, the normalized power spectral density of the pattern and the mean value of the calculated normalized power spectral density estimates are presented. As a search space during ARMA model recovery, a hypercube of dimension 5 surrounding the origin of the coordinate system was chosen – for each MN-sample realization of time-series $y_B(i)$, estimates of 5 coefficients of of ARMA model $y_1(i)$ were calculated. The obtained mean value of the minimum of the objective function $\mathcal{S}_2(\Theta_1)$ was equal to 0.16.

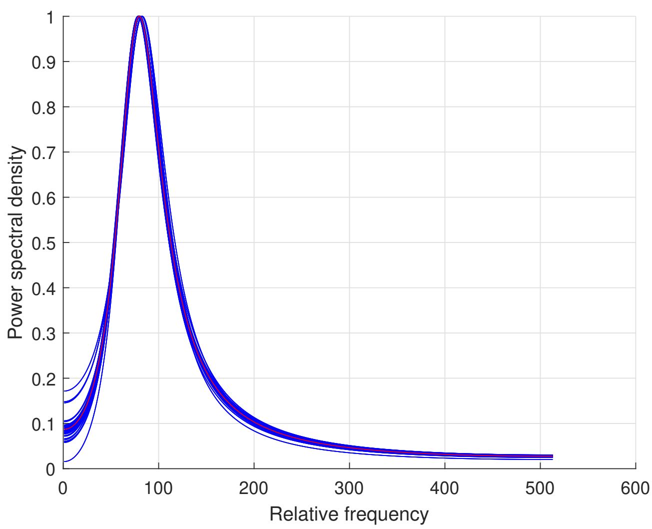

Figure 5: Normalized power spectral densities calculated from 100 identified models of ARMA time-series represented by binary-valued observations (blue lines) and the normalized pattern (red line) – direct power spectral density approximation, $M = 2000$ and $N = 128$

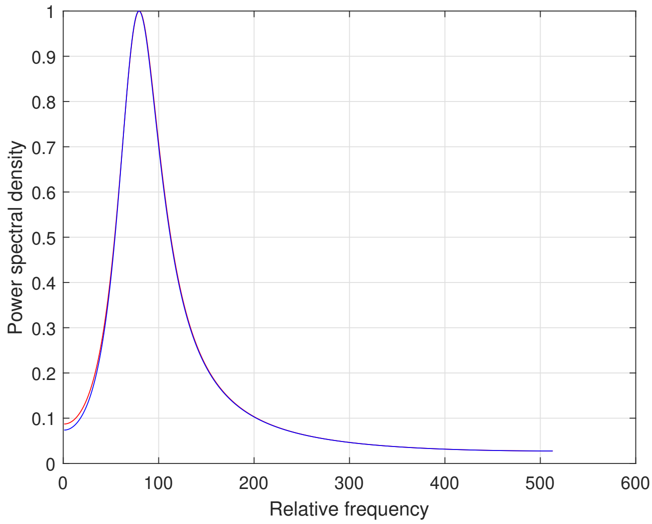

Figure 6: Mean value of 100 normalized power spectral densities calculated from identified models of ARMA time-series represented by binary-valued observations (blue line) and the normalized pattern (red line) – direct power spectral density approximation, $M = 2000$ and $N = 128$

5.3. ARMA model identification using direct autocorrelation function approximation – binary-valued observations

In the third simulated identification example, the above-described experiment, in which ARMA models were identified using direct power spectral density approximation, was repeated, but now ARMA models were identified using direct autocorrelation function approximation. Estimates of autocorrelation functions necessary to perform calculations were obtained using the biased autocorrelation estimator. These estimates were calculated for lags $0, 1, , \ldots, N-1$. The corresponding recovered 100 estimates of parameters of ARMA time-series $y_1(i)$ were used to calculate the corresponding estimates of the normalized power spectral density. In Fig. 7, the results of these calculations are compared with the normalized pattern. In Fig. 8, the normalized power spectral density of the pattern and the mean value of calculated normalized power spectral density estimates are presented. As the search space during ARMA model recovery, a hypercube of dimension 5 surrounding the origin of the coordinate system was chosen – once more for each MN-sample realization of the time-series $y_p(i)$ estimates of 5 coefficients of the ARMA model $y_1(i)$ were calculated. The obtained mean value of the minimum of the objective function $\mathcal{S}_2(\Theta_1)$, calculated for the estimates $\hat{R}_1(\tau)$ and $\widetilde{R}_{1m}(\Theta_1, \tau)$ $(\tau = 0, 1, \ldots, N-1)$, was equal to $4.7 \cdot 10^{-4}$.

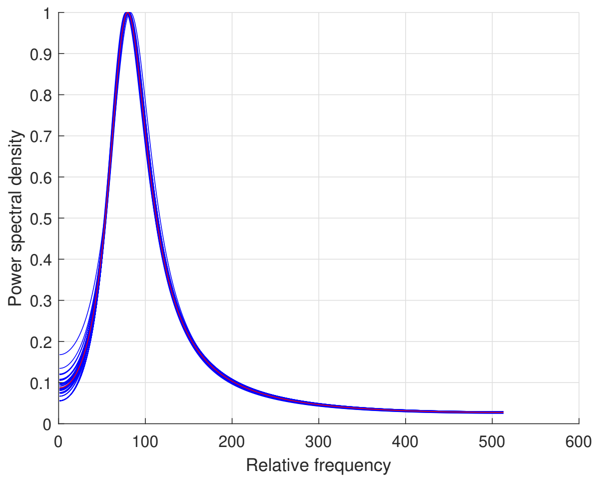

Figure 7: Normalized power spectral densities calculated from 100 identified models of ARMA time-series represented by binary-valued observations (blue lines) and the normalized pattern (red line) – direct autocorrelation function approximation, $M = 2000$ and $N = 128$

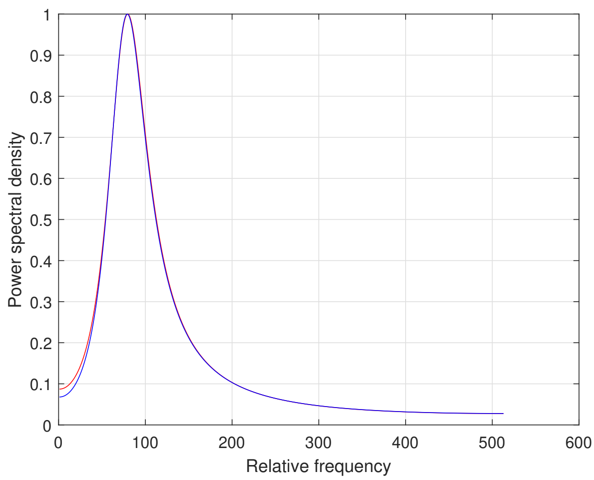

Figure 8: Normalized mean value of 100 power spectral densities calculated from identified models of ARMA time-series represented by binary-valued observations (blue line) and the normalized pattern (red line) – direct autocorrelation function approximation, $M = 2000$ and $N = 128$

5.4. Blind Wiener model identification using direct power spectral density approximation – binary-valued observations

In the next simulated identification example, parameters of a Wiener model used to simulate realizations of time-series were recovered from estimates of the power spectral density of binary-valued observations. This model had the following components:

- an input discrete-time linear dynamic subsystem described by the following difference equation:

$$y_W(i) = \frac{1.0000}{1.0000 - 1.5000z^{-1} + 0.7000z^{-2}} e_p(i); \tag{11}$$

- the output nonlinearity:

$$y_p(i) = y_W(i) + 0.2000\,(y_W(i))^3\,. \tag{12}$$

The corresponding 10 realizations of time-series $y_B(i)$ were simulated assuming $\sigma^2_{e_p e_p} = 1.000$, $up = -lo = 1.0000$, $M = 2000$ and $N = 512$. In this case of model recovery, the search space was of dimension 6, and parameters of Wiener models were recovered using the direct power spectral density approximation. In Fig. 9, results of the normalized power spectral density calculations using the identified models are compared with the normalized pattern. In Fig. 10, the normalized power spectral density of the pattern and the mean value of the calculated normalized power spectral densities are presented. The obtained mean value of the minimum of the objective function $\mathcal{S}_2(\Theta_1)$ was equal to 0.80.

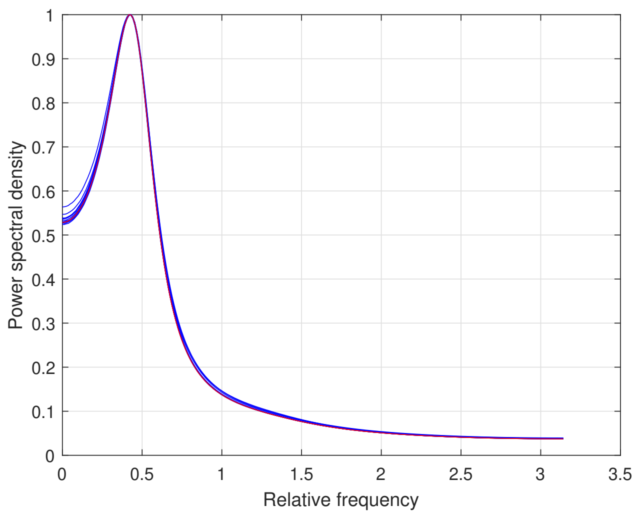

Figure 9: Normalized power spectral densities calculated using 10 identified Wiener models for the corresponding time-series represented by binary-valued observations (blue lines) and the normalized pattern (red line) – direct power spectral density approximation, $M = 2000$ and $N = 512$

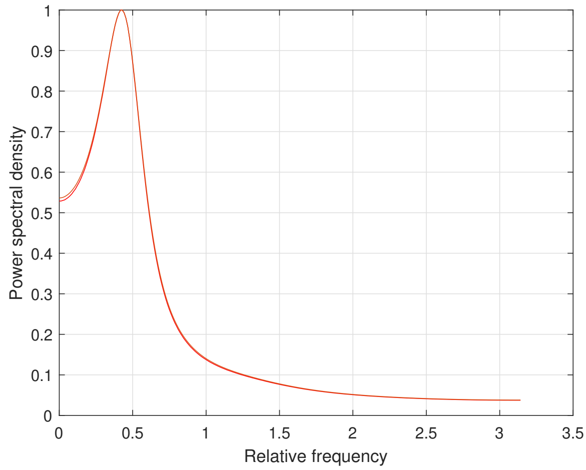

Figure 10: Normalized mean value of 10 power spectral densities calculated from identified Wiener models for the corresponding simulated time-series represented by binary-valued observations (blue line) and the normalized pattern (red line) – direct power spectral density approximation, $M = 2000$ and $N = 128$

5.5. Blind Wiener–Hammerstein model identification using direct power spectral density approximation – binary-valued observations

In the last simulated identification example devoted to model identification using binary-valued observations, parameters of a Wiener–Hammerstein model used to simulate realizations of time-series were recovered from estimates of the corresponding power spectral densities. The Wiener–Hammerstein model had the following components:

- an input discrete-time linear dynamic subsystem described by the following difference equation:

$$y_{WH}(i) = \frac{1.0000}{1.0000 - 1.5000z^{-1} + 0.7000z^{-2}} e_p(i); \tag{13}$$

- the internal nonlinearity:

$$u_{WH}(i) = y_{WH}(i) + 0.2000\,(y_{WH}(i))^3\,; \tag{14}$$

- an output discrete-time linear dynamic subsystem described by the following difference equation:

$$y_p(i) = \frac{1.0000}{1.0000 - 0.8000z^{-1}} u_{WH}(i). \tag{15}$$

A single realization of the corresponding time-series $y_B(i)$ was simulated using the above model, assuming the same values of $\sigma^2_{e_p e_p}$, up, lo, M and N as in the simulated blind Wiener model identification example. In the case of

Wiener–Hammerstein model recovery, the search space had dimension of 9, and parameters of Wiener–Hammerstein model were recovered using the direct power spectral density approximation. Model recovery gave the model consisting of:

- an input discrete-time linear dynamic subsystem described by the following difference equation:

$$y_{WHm}(i) = \frac{-1.6814}{1.1445 - 1.5052z^{-1} + 0.6604z^{-2}} e_1(i); \tag{16}$$

- the internal nonlinearity:

$$u_{WHm}(i) = -0.9521 y_{WHm}(i) - 0.6330\,(y_{WH}(i))^3\,; \tag{17}$$

- an output discrete-time linear dynamic subsystem described by the following difference equation:

$$y_m(i) = \frac{1.3756}{-1.0410 - 0.7984z^{-1}} u_{WHm}(i). \tag{18}$$

The corresponding value of the objective function $\mathcal{S}_2(\Theta_1)$ was equal to 3.2. In Fig. 11, normalized averaged periodograms calculated using 10000 realizations of binary-valued observations obtained using the simulated Wiener–Hammerstein model are compared with the normalized averaged periodogram of a single realization of binary-valued observations obtained using the recovered Wiener–Hammerstein model.

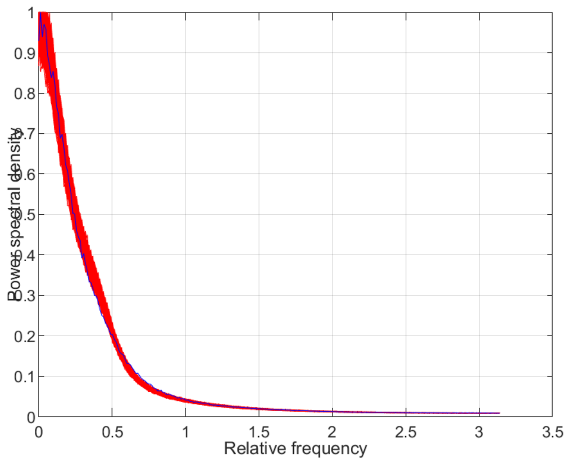

Figure 11: Normalized averaged periodograms calculated using 10000 realizations of binary-valued observations obtained using the simulated Wiener–Hammerstein model (red lines) and the normalized averaged periodogram of a single realization of binary-valued observations obtained using the recovered Wiener–Hammerstein model (blue line) – direct power spectral density approximation, $M = 2000$ and $N = 512$

In the next step, 50 realizations of the time-series $y_B(i)$ were simulated, and the corresponding Wiener–Hammerstein model recovery was repeated using the direct power spectral density approximation. In Fig. 12, the results of the normalized power spectral density calculations using the identified models are compared with the normalized pattern. In Fig. 13, the normalized power spectral

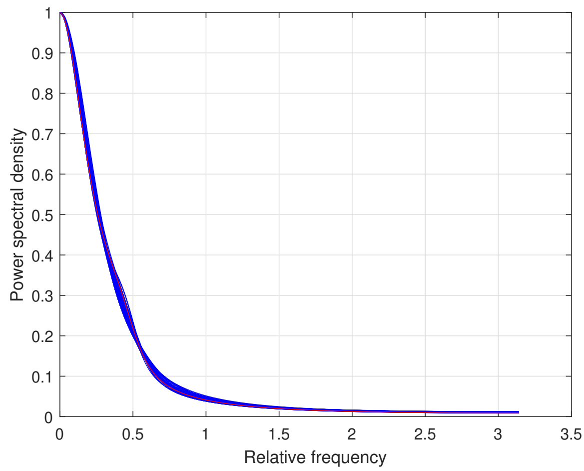

Figure 12: Normalized power spectral densities calculated using 50 identified Wiener–Hammerstein models for the corresponding time-series represented by binary-valued observations (blue lines) and the normalized pattern (red line) – direct power spectral density approximation, $M = 2000$ and $N = 512$

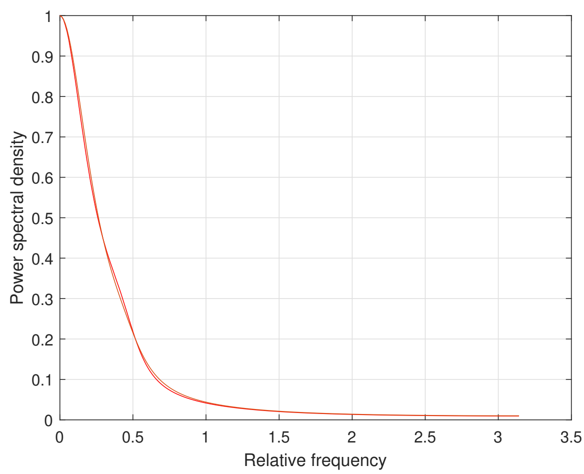

Figure 13: Normalized mean value of 50 power spectral densities calculated from the identified Wiener–Hammerstein models using the corresponding simulated time-series represented by binary-valued observations (blue line) and the normalized pattern (red line) – direct power spectral density approximation, $M = 2000$ and $N = 512$

density of the pattern and the mean value of the calculated normalized power spectral densities are presented. The obtained mean value of the minimum of the objective function $\mathcal{S}_2(\Theta_1)$ was equal to 3.29.

5.6. Ultra low-power ARMA model identification using random time-series decomposition – quantized measurements

In the case of binary-valued observations, time-series are represented by values taken from a set containing two numbers. The next possible time-series representation are quantized measurements. In the case of ultra low-power time-series, the corresponding quantized measurements take values from a set containing only a few (mainly 3) numbers. In the simulated identification example devoted to such time-series, case a model of ARMA time-series defined by the following relation:

$$y_p(i) = \frac{1.0000 - 0.4000z^{-1}}{1.0000 - 1.3000z^{-1} + 0.5000z^{-2}} e_p(i), \tag{19}$$

was identified, where $e_p(i)$ was a Gaussian white noise having the variance $\sigma^2_{e_p e_p}$ equal to $4 \cdot 10^{-6}$. Using the above model, 75 realizations of $y_p(i)$ were generated assuming $M = 1000$ and $N = 256$. Values $y_p(i)$ were transformed to the corresponding quantized measurements $y_Q(i)$ by the 8-bit midtread uniform quantizer having saturation. Its range was limited by saturation levels equal to 5.0000 and -5.0000, respectively. In Fig. 14, an exemplary 512-sample of $y_Q(i)$ is presented. For each acquired realization $y_Q(i)$ $(i = 1, 2, \ldots, MN - 1)$, the model identification was performed using the random time-series decomposition, assuming that

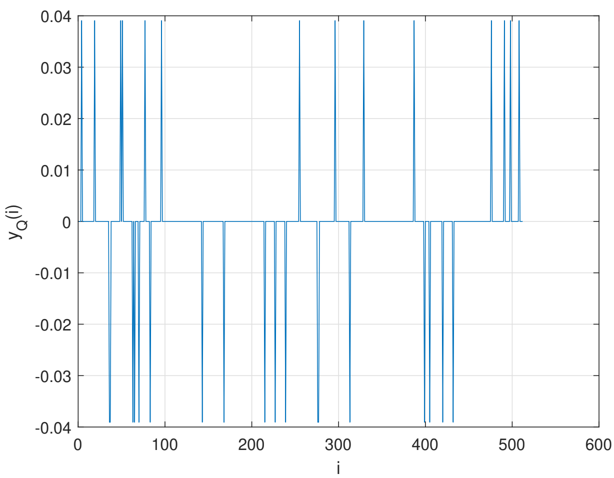

Figure 14: Exemplary 512-sample of quantized measurements acquired from an ultra low-power ARMA time-series realization

nonlinear distortions implied by the transformation $Q(\cdot)$ are modelled by a white noise $y_2(i)$ (ARMA time-series having the structural numbers $dA_2 = dB_2 = 0$). As a search space during ARMA model recovery, a hypercube of dimension 6 surrounding the origin of the coordinate system was chosen. For each MN-sample realization of $y_Q(i)$ estimates of 6 parameters of the two ARMA $y_1(i)$ and $y_2(i)$ time-series models were calculated. The obtained mean value of the minimum of the objective function $S_1(\Theta)$ was equal to 0.16. Using 75 estimates of parameters of ultra low-power ARMA time-series $y_1(i)$, the corresponding estimates of the power spectral density were calculated and compared with the pattern in Fig. 15. In Fig. 16, the corresponding power spectral densities of the pattern and the mean value of calculated power spectral density estimates are presented.

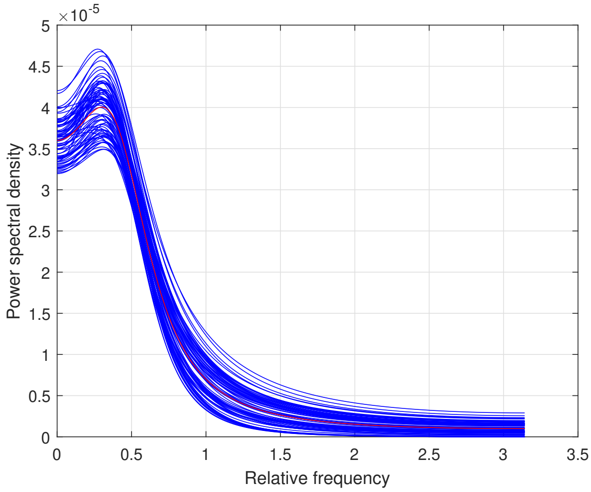

Figure 15: Power spectral densities calculated using 75 identified models of ultra low-power ARMA time-series represented by quantized measurements (blue lines) and the pattern (red line) – random time-series decomposition, $M = 1000$ and $N = 256$

5.7. Normal-power blind Wiener model identification using direct power spectral density approximation – quantized measurements

Next simulated identification example is devoted to normal-power time-series model identification using quantized measurements. The simulated identification experiment described in subsection V.D was repeated. Now, 20 realizations of Wiener time-series $y_p(i)$ defind by equations (11) and (12) were simulated assuming $\sigma^2_{e_p e_p} = 4 \cdot 10^{-4}$, $M = 2000$ and $N = 512$. Realizations of $y_p(i)$ were transformed to the corresponding quantized measurements $y_Q(i)$ by the 8-bit midtread uniform quantizer defined in the previous simulated identification ex-

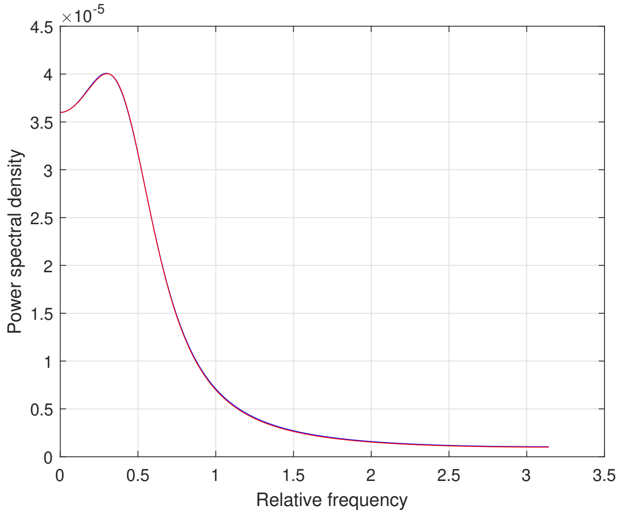

Figure 16: Mean value of 75 power spectral densities calculated from identified models of ultra low-power ARMA time-series represented by quantized measurements (blue line) and the pattern (red line) – random time-series decomposition, $M = 1000$ and $N = 256$

ample. In Fig. 17, an exemplary 512-sample of $y_Q(i)$ is presented. It can be noticed that after quantization, the processed time-series takes the values of only a few quantizer levels around the zero value. During model recovery, the search space was dimension of 6 and parameters of Wiener models were recovered using

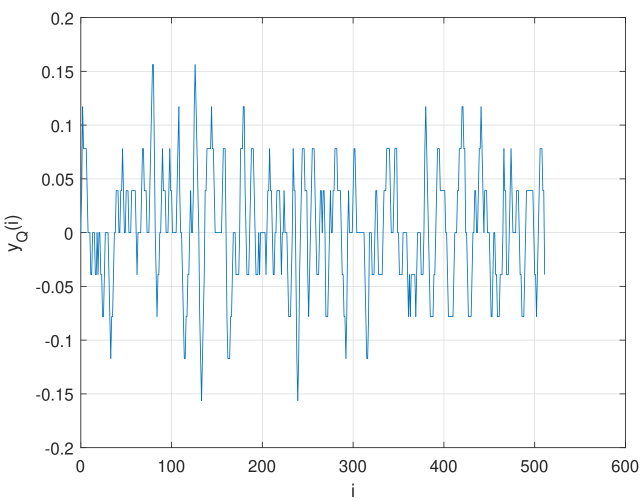

Figure 17: Exemplary 512-sample of quantized measurements of a normal-power Wiener time-series realization

the direct power spectral density approximation. In Fig. 18, results of the power spectral density calculations using the identified models are compared with the pattern. In Fig. 19, the power spectral density of the pattern and the mean value of calculated power spectral densities are presented. The obtained mean value of the minimum of the objective function $\mathcal{S}_2(\Theta_1)$ was equal to $1.99 \cdot 10^{-5}$.

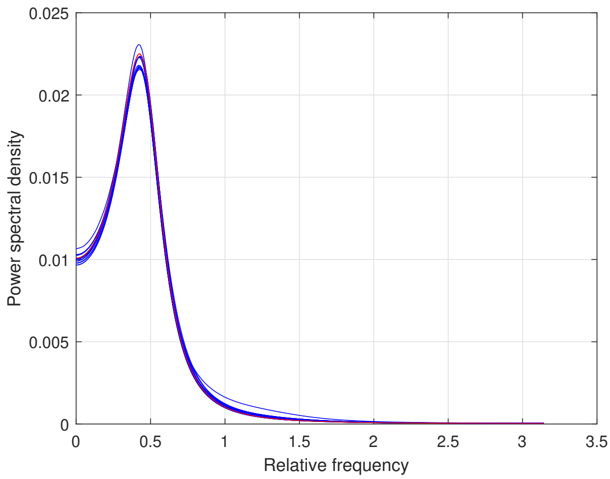

Figure 18: Power spectral densities calculated using 20 identified models of a normal-power Wiener time-series represented by quantized measurements (blue lines) and the pattern (red line) – power spectral density approximation, $M = 2000$ and $N = 512$

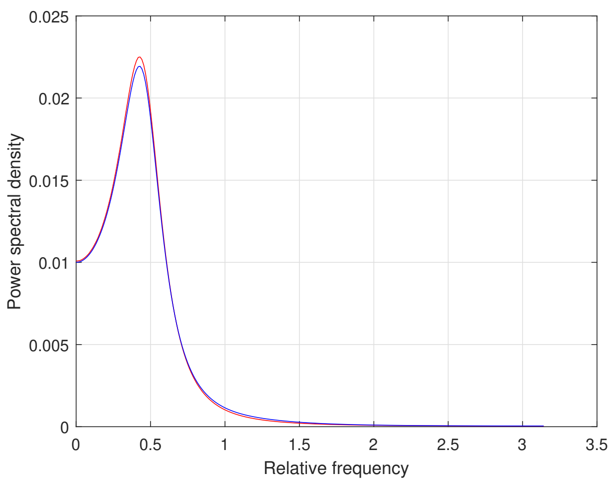

Figure 19: Mean value of 20 power spectral densities calculated from identified models of a normal-power Wiener time-series represented by quantized measurements (blue line) and the pattern (red line) – power spectral density approximation, $M = 2000$ and $N = 512$

5.8. Ultra high-power ARMA model identification using random time-series decomposition – quantized measurements

In the last simulated identification example, ARMA time-series model (19) was excited by $e_p(i)$ which is a Gaussian white noise with a variance $\sigma^2_{e_p e_p}$ equal to 36.0000. Using this relation, 100 realizations of $y_p(i)$ were generated for M equal to 10000 as well as 1000000. N was equal to 1024. Values $y_p(i)$ were transformed to the corresponding quantized measurements $y_Q(i)$ in the same way as in the previous simulated identification example. In Fig. 20, an exemplary 1024-sample of $y_Q(i)$ is presented. For each acquired realizations $y_Q(i)$ ($i = 1, 2, \ldots, MN - 1$), the model identification was performed using the random time-series decomposition, assuming that the nonlinear distortions implied by the transformation $Q(\cdot)$ are modelled by an ARMA time-series $y_2(i)$ having the structural numbers $dA_2 = dB_2 = 1$. As a search space during ARMA model recovery, a hypercube of dimension 9 surrounding the origin of the coordinate system was chosen,. For each MN-sample realization of $y_Q(i)$, estimates of 9 parameters of the two models of ARMA $y_1(i)$ and $y_2(i)$ time-series were calculated. The obtained mean value of the minimum of the objective function $\mathcal{S}_1(\Theta)$ for $M = 10000$ was equal to 36.6. Result of the mean value of $\mathcal{S}_1(\Theta)$ calculation for $M = 1000000$ was equal to 0.52. Using 100 estimates of parameters of ARMA time-series $y_1(i)$, the corresponding estimates of the power spectral density were calculated and, after normalization by maximum values, are compared in Figs. 21 and 22 with the normalized pattern. The power spectral density of the pattern

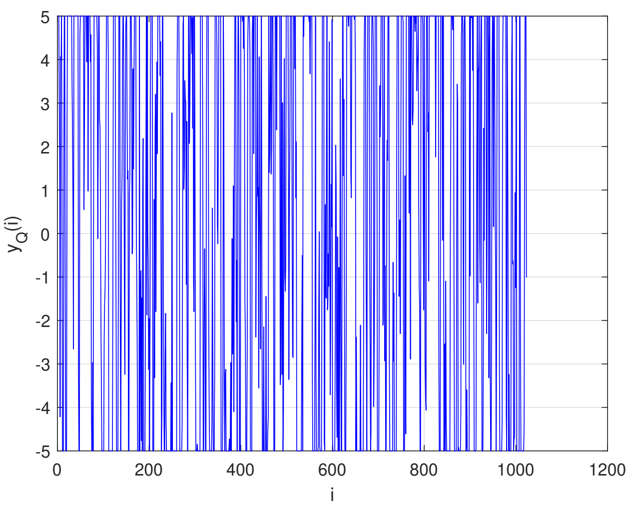

Figure 20: Exemplary 1024-sample of quantized measurements of a high-power ARMA time-series realization

was calculated using the model presented in equation (19). In Fig. 23, this power spectral density of the pattern and the mean value of calculated power spectral densities for $M = 10000$ are compared.

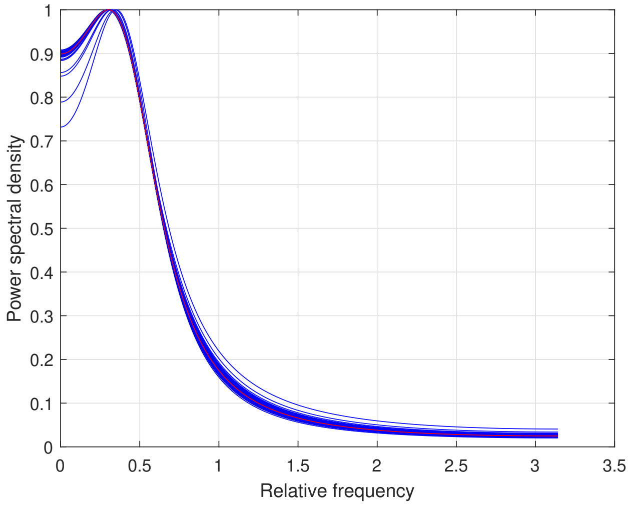

Figure 21: Power spectral densities calculated using 100 identified models of a high-power ARMA time-series represented by quantized measurements (blue lines) and the pattern (red line) – random time-series decomposition, $M = 10000$ and $N = 1024$

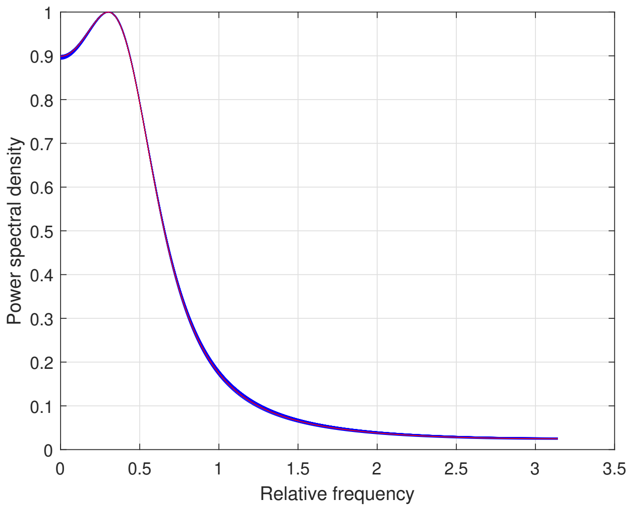

Figure 22: Power spectral densities calculated using 100 identified models of a high-power ARMA time-series represented by quantized measurements (blue lines) and the pattern (red line) – random time-series decomposition, $M = 1000000$ and $N = 1024$

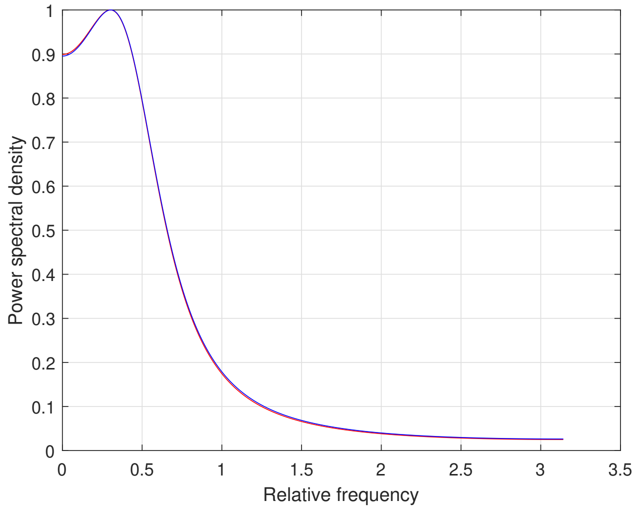

Figure 23: Mean value of 100 power spectral densities calculated from identified models of a high-power ARMA time-series represented by quantized measurements (blue line) and the pattern (red line) – random time-series decomposition, $M = 10000$ and $N = 1024$

6. Short summary

In the above-presented simulated identification examples, the results of identification are reported without discussion of their statistical properties. Now, they are summarized by interpreting identification algorithms as estimators. The identification algorithms discussed in the paper gave:

- mean values of normalized power spectral densities calculated from identified dynamic parametric models of time-series that are very close to the corresponding normalized patterns for models identified using binary-valued observations acquired or models identified for high-power time-series based on quantized measurements;
- mean values of power spectral densities calculated from identified dynamic parametric models of time-series that are very close to the corresponding patterns for models identified for ultra low- and normal-power time-series based on quantized measurements.

Additionally, the variance of the obtained identification results declines with the increase in the number of processed data (binary-valued observations or quantized measurements). This was not shown in the paper in a systematic way (apart from the last simulated identification example), but it follows from considerations presented in the monograph [17]. It should be emphasized that this is true only when the ratio of the number of parameters identified to the number of processed data declines to 0 with the increase in the number of processed data.

7. Conclusions

In the paper, an approach to the identification of linear and nonlinear dynamic parametric models of time-series from the corresponding binary-valued observations of their realizations, as well as quantized measurements of their values, is proposed. The approach proposed is based on randomized search optimization algorithms used to recover the corresponding parametric models from the calculated estimates of the power spectral density or autocorrelation function. The discussion presented is illustrated with a set of simulated identification examples showing the effectiveness of the proposed approach. Although the ideas introduced by the paper are illustrated with examples of linear and nonlinear block-oriented dynamic models of time-series identification, it is worth emphasizing that other nonlinear dynamic parametric models of time-series, e.g. bilinear or Volterra models, may be identified in the same way.

The ideas presented in the paper complement the considerations presented in the monograph [17] and result directly from them.

References

[1] R. Auber, M. Pouliquen, E. Pigeon, M. M'Saad, O. Gehan, P.A. Chapon, and S. Moussay: Estimation of auto-regressive models for time series using binary or quantized data. *IFAC-PapersOnLine*, **51**(15), (2018), 581–586. DOI: 10.1016/j.ifacol.2018.09.221

[2] R. Auber, M. Pouliquen, E. Pigeon, M. M'Saad, O. Gehan, P.A. Chapon, and S. Moussay: *Estimation of auto-regressive models for time series using binary or quantized data*. SYSID, Stockholm, 2018.

[3] R. Auber, M. Pouliquen, E. Pigeon, O. Gehan, M. M'Saad, P.A. Chapon, and S. Moussay: Identification of AR time-series based on binary data. *IET Signal Processing*, **14**(1), (2020), 24–31. DOI: 10.1049/iet-spr.2019.0152

[4] J.S. Bendat and A.G. Piersol: *Random Data Analysis and Measurement Procedures*. John Wiley & Sons, Inc., New York, 1986.

[5] A. Bagnall, G. Janacek: *Clustering time series from arma models with clipped data*. 10th ACM SIGKDD International Conference ed data. 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Seattle, 2004. DOI: 10.1145/1014052.1014061

[6] A. Bagnall and G. Janacek: Clustering time series with clipped data. *Machine Learning*, **58** (2005), 151–178. DOI: 10.1007/s10994-005-5825-6

[7] I. Bilinsks: *Digital Alias-free Signal Processing*, John Wiley & Sons, 2007.

[8] I. Boussaid, J. Lepagnot, and P. Siarry: A survey of optimization metaheuristics. *Information Science*, **237** (2013). DOI: 10.1016/j.ins.2013.02.041

[9] G.E.P. Box and G.M. Jenkins: *Time Series Analysis, Forecasting and Control*. Holden–Day, San Francisco 1976.

[10] J. Brownlee: *Clever Algorithms: Nature–Inspired Programming Recipes*. Jason Brownlee, LuLu 2012.

[11] J. Figwer: *Nonlinear System Modelling* (in Polish). in Inżynieria Wiedzy i Systemy Ekspertowe, A. Grzech, K. Juszczyszyn, H. Kwaśnicka, N.T. Nguyen (eds.), Akademicka Oficyna Wydawnicza EXIT, Warszawa, pp. 89–100, 2009.

[12] J. Figwer: Continuous-time dynamic system identification with multisine random excitation revisited. *Archives of Control Sciences* (2010), 123–139.

[13] J. Figwer: *Multisine Random processes. Theory and Applications* (in Polish). Akademicka Oficyna Wydawnicza EXIT, Warszawa, 2012.

[14] J. Figwer: *AR Time-Series Identification using Quantized Observations*. Proceedings of The 19th International Conference on Methods and Models in Automation and Robotics, pp. 235–239, 2014. DOI: 10.1109/MMAR.2014.6957357

[15] J. Figwer: *Continuous–time input–output linear dynamic system identification using sampled data*. Proceedings of The 20th International Conference on Methods and Models in Automation and Robotics, pp. 712–717, 2015. DOI: 10.1109/MMAR.2015.7283962

[16] J. Figwer: *A Random Time–Series Decomposition and Its Applications in Time–Series Identification*. Proceedings of The 21st International Conference on Methods and Models in Automation and Robotics, pp. 206–211, 2016. DOI: 10.1109/MMAR.2016.7575134

[17] J. Figwer: *Process Identification by Means of Random Search Methods Illustrated with Simulation Examples* (in Polish). Wydawnictwo Politechniki Slaskiej, Gliwice, Poland (2022)

[18] J. Figwer: Continuous-Time Dynamic Model Identification Using Binary–Valued Observations of Input and Output Signals. In: M. Pawelczyk, D. Bismor, Z. Ogonowski, J. Kacprzyk (eds.) Advanced, Contemporary Control. PCC 2023. *Lecture Notes in Networks and Systems*, vol. 708. Springer, Cham, 2023. DOI: 10.1007/978-3-031-35170-9_5

[19] K.A. De Jong: *Evolutionary Computation: A Unified Approach*. The MIT Press, 2006.

[20] S.M. Kay: *Modern Spectral Estimation: Theory and Application*. Prentice Hall, 1988.

[21] B. Kedem: Estimation of the parameters in stationary autoregressive processes after hard limiting, *Journal of The American Statistical Association*, **75** (1980), 146–153.

[22] V. Krishnamurthy and I. Mareels: Estimation of noisy quantized Gaussian AR time-series with randomly varying observation coefficient. *IEEE Transactions on Signal Processing*, **43** (1995), 1285–290.

[23] V. Krishnamurthy and H.V. Poor: Asymptotic Analysis of an Algorithm for Parameter Estimation and Identification of 1–b Quantized AR Time Series. *IEEE Transactions on Signal Processing*, **44** (1996), 62–73.

[24] L. Ljung:, *System Identification – Theory for the User*: Prentice Hall, New Jersey, 1999.

[25] A. Logothetis, V. Krishnamurthy, and H.V. Poor: Estimation of 1-bit quantized time-series with Markov regime. *Signal Processing*, **58** (1997), 273–292.

[26] H. Lütkepohl: *Introduction to Multiple Time Series Analysis*. Springer–Verlag, New York 1993.

[27] S. Mahdavi, M.E. Shiri, and S. Rahnamayan: Metaheuristics in large-scale global continues optimization: A survey. *Information Sciences*, **295** (2015), 407–428, DOI: 10.1016/j.ins.2014.10.042

[28] S.L. Marple: *Digital Spectral Analysis with Applications*. Prentice Hall, Englewood Cliffs, N.J. 1987.

[29] Optimization Toolbox: User's Guide Version 3, The MathWorks.

[30] M.B. Priestley: *Spectral analysis and time series*. Academic Press, New York 1981.

[31] J.C. Spall, *Introduction to Stochastic Search and Optimization*. Wiley, New York 2003.

[32] P. Stoica and R.L. Moses: *Spectral analysis of signals*. Pearson Prentice Hall, 2005.

[33] X. Zhen and I. Basawa: Estimation for binary models generated by gaussian autoregressive processes. *Journal of Statistical Computation and Simulation*, **80** (2010), 1041–1051. DOI: 10.1080/00949650902922711