Automatica 159 (2024) 111351

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Technical communique

Finite-data nonparametric frequency response evaluation without leakage[☆]

Ivan Markovsky[a,b,*], Hamid Ossareh[c]

[a] Catalan Institution for Research and Advanced Studies (ICREA), Pg. Lluis Companys 23, Barcelona, Spain
[b] International Centre for Numerical Methods in Engineering (CIMNE), Gran Capitàn, Barcelona, Spain
[c] Department of Electrical and Biomedical Engineering University of Vermont, 33 Colchester Ave, Burlington, USA

ARTICLE INFO

Article history:
Received 21 September 2022
Received in revised form 14 July 2023
Accepted 11 September 2023
Available online 3 November 2023

Keywords:
Frequency response estimation
Leakage
Direct data-driven methods
Subspace methods
Behavioral approach

ABSTRACT

The existing nonparametric frequency response estimation methods suffer from leakage. Because of this, these methods do not yield the correct result in case of exact data of a linear time-invariant system. Our main contribution is a time-domain direct data-driven nonparameteric frequency response estimation method that, in case of exact data satisfying standard persistency of excitation condition, eliminates the leakage and has infinite frequency resolution. The method is derived in the behavioral setting. It requires solving a system of linear equations and has no hyper-parameters. In case of noisy data, a modification of the method with preprocessing with low-rank approximation results in a subspace-type method for frequency response estimation.

1. Introduction

Nonparametric frequency response estimation is a classical system identification problem, see Ljung (1999, Chapter 6). The basic solution, referred to as the *empirical transfer-function estimate* computes the discrete-time Fourier transforms (DTFT) of the given time-domain input/output data sequences and divides per frequency the Fourier transform of the output by the Fourier transform of the input. The resulting method is frequency-domain direct data-driven. It is conceptually simple, computationally efficient, thanks to the fast Fourier transform (FFT), and easy to use. Even with exact finite-length data, however, the empirical transfer-function estimate is not guaranteed to deliver the exact frequency response because of *leakage errors*.

From a system-theoretic perspective, the leakage error is the effect of the ignored initial conditions and the resulting transient response (Pintelon & Schoukens, 2012, Section 6.3.2). From a signal processing perspective, the leakage error is the effect of the windowing of the data. Numerous modifications of the basic empirical transfer-function method aim to reduce the errors due to the leakage, see, *e.g.*, Gevers, Pintelon, and Schoukens (2011), Kay (1988) and Stoica and Moses (2005). The state-of-the-art methods are based on pre-processing of the data by filtering, which often involve hyper-parameters whose tuning requires user expertise and prior knowledge about the system. Thus, although nonparametric frequency response estimation methods have been developed for many years, the fundamental problem of leakage remains open. It should be noted that this problem can be resolved in the parametric setting using classical identification methods. The parametric approach, however, leads to *indirect* data-driven methods. Moreover, the parametric methods use prior knowledge about the model structure, which the nonparametric methods do not use.

Our objective is to develop a direct data-driven nonparametric method that resolves the leakage problems in the case of exact (noise free) data. The basic problem, referred to as *frequency response evaluation*, is: Given an exact finite-length time-domain input/output trajectory of a finite-order deterministic linear time-invariant system and a set of frequencies, find the frequency response of the system at the given frequencies. Contrary to the existing nonparameteric methods that convert the data to the frequency-domain, the method proposed here operates directly on the time-domain data. The method has no hyper-parameters and is provably correct under a standard persistency of excitation condition on the data. In case of inexact (noisy) data, we propose a modification of the method based on low-rank approximation, which has as a hyper-parameter the order of the system. The resulting subspace-type estimation method is computationally cheap but suboptimal. Currently, it lacks theoretical performance

☆ The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor A. Pedro Aguiar under the direction of Editor André Tits.

* Corresponding author at: Catalan Institution for Research and Advanced Studies (ICREA), Pg. Lluis Companys 23, Barcelona, Spain.

E-mail addresses: imarkovsky@cimne.upc.edu (I. Markovsky), Hamid.Ossareh@uvm.edu (H. Ossareh).

https://doi.org/10.1016/j.automatica.2023.111351
0005-1098/

guarantees, however, empirical results show that it works well in practice.

The derivation of the method is done in the behavioral setting, which views a dynamical system as a set of trajectories. The key tool at the heart of the data-driven method proposed in the paper is a time-domain nonparametric representation of the finite horizon behavior of the system. The result that gives conditions for validity of the data-driven representation became known as the fundamental lemma (Willems, Rapisarda, Markovsky, & De Moor, 2005). The data-driven representation was effectively used in Markovsky and Rapisarda (2008) for solving simulation and tracking control problems. The methods were originally developed for exact data but were subsequently generalized to noisy data and some classes of nonlinear time-varying systems, see the overviews (Markovsky & Dörfler, 2021; Markovsky, Huang, & Dörfler, 2023). The necessary background is given in Section 2. Section 3.1 presents to solution in case of exact data and Section 3.2 presents the modification of the method for the case of noisy data. The method is illustrated and validated in Section 4.

2. Preliminaries and problem statement

Let $(\mathbb{C}^q)^{\mathbb{N}}$ be the set of q-variate complex-valued signals with time-axis the natural numbers $\mathbb{N}$ (*i.e.*, vector sequences). In the *behavioral approach* to systems theory (Polderman & Willems, 1998; Willems, 1986, 1997), a dynamical system is defined as a set of trajectories $\mathscr{B} \subset (\mathbb{C}^q)^{\mathbb{N}}$. The important difference from the classical approach is the distinction of the system (set of trajectories) from its representations (algebraic/difference/differential equations).

Note 1 (*Complex-valued Signals and Systems*)**.** In the system identification and data-driven signal processing and control literature using the behavioral approach, trajectories and behaviors have traditionally been real-valued. In this paper, we extend the methods to complex-valued signals w and correspondingly define the system $\mathscr{B}$ as a subset of $(\mathbb{C}^q)^{\mathbb{N}}$. In practice, however, the response to a real-valued input and real-valued initial condition is a real-valued signal. With some abuse of notation, we call systems with such property real-valued. The response $y = y_{\text{real}} + \mathbf{i}y_{\text{imag}}$ of a real-valued system $\mathscr{B}$ to a complex-valued input $u = u_{\text{real}} + \mathbf{i}u_{\text{imag}}$ and/or complex-valued initial condition is complex-valued. However, $w_{\text{real}} = \left[\begin{smallmatrix} u_{\text{real}} \\ y_{\text{real}} \end{smallmatrix}\right]$ and $w_{\text{imag}} = \left[\begin{smallmatrix} u_{\text{imag}} \\ y_{\text{imag}} \end{smallmatrix}\right]$ are decoupled, *i.e.*, two independent real-valued trajectories are formally put together in one complex-valued trajectory. This formalism is used in Section 3, where we consider complex exponentials instead of sine and cosine trajectories. □

We use the behavioral approach because of its relevance to the data-driven methods in signal processing and control. For example, it allows us to use the short-hand notation $w \in \mathscr{B}$ for "the signal w is a trajectory of the system $\mathscr{B}$". Since we consider finite signals we use the following notation for restricting the time axis to a finite interval: $w|_T := \big(w(1), \dots, w(T)\big)$ is the restriction of w to the interval $[1, T]$ and $\mathscr{B}|_T \subset (\mathbb{R}^q)^T$ is the restriction of $\mathscr{B}$ to $[1, T]$.

A linear time-invariant system $\mathscr{B}$ is a shift-invariant subspace of the space of trajectories $(\mathbb{R}^q)^{\mathbb{N}}$. The number of inputs m, lag ℓ, and order n of $\mathscr{B}$ are invariant of its representations and are therefore properties of $\mathscr{B}$ (Willems, 1986). The restricted behavior $\mathscr{B}|_T$ for $T \geq \ell$ admits a representation

$$\mathscr{B}|_T = \text{image} \underbrace{\begin{bmatrix} w_{\text{d}}(1) & w_{\text{d}}(2) & \cdots & w_{\text{d}}(T_{\text{d}} - T + 1) \\ w_{\text{d}}(2) & w_{\text{d}}(3) & \cdots & w_{\text{d}}(T_{\text{d}} - T + 2) \\ \vdots & \vdots & & \vdots \\ w_{\text{d}}(T) & w_{\text{d}}(T+1) & \cdots & w_{\text{d}}(T_{\text{d}}) \end{bmatrix}}_{\mathscr{H}_T(w_{\text{d}})}, \quad (1)$$

by a trajectory $w_{\text{d}} \in \mathscr{B}|_{T_{\text{d}}}$ of length T_{d} that satisfies the *generalized persistency of excitation* condition (Markovsky & Dörfler, 2023)

$$\operatorname{rank} \mathscr{H}_T(w_{\text{d}}) = mT + n. \quad (2)$$

The matrix $\mathscr{H}_T(w_{\text{d}}) \in \mathbb{R}^{qT \times (T_{\text{d}} - T + 1)}$ is the Hankel matrix with depth T constructed from the data w_{d}. Note that (2) is verifiable from the data w_{d} and the prior knowledge of the system's number of inputs and order.

Note 2. The data-driven representation (1) and the generalized persistency of excitation condition (2) are derived for real-valued systems and signals. The derivation, however, is also valid in the complex-valued case. If $\mathscr{B}$ is real-valued as defined in Note 1, it is sufficient to use a real-valued trajectory w_{d} that satisfies (2). Complex-valued trajectories are represented then by linear combination of the columns of $\mathscr{H}_T(w_{\text{d}})$ with complex-valued coefficients. □

Consider a linear time-invariant system $\mathscr{B}$ with an input/output partitioning of the variables $w = \left[\begin{smallmatrix} u \\ y \end{smallmatrix}\right]$. Let $\mathscr{B}(H)$ be the transfer function representation of the controllable part of $\mathscr{B}$, corresponding to the input/output partitioning $w = \left[\begin{smallmatrix} u \\ y \end{smallmatrix}\right]$. The data-driven frequency response evaluation problem considered is defined as follows.

Problem 1 (*Data-driven Frequency Response evaluation*)**.** Given a finite input/output trajectory $(u_{\text{d}}, y_{\text{d}})$ of a linear time-invariant system $\mathscr{B}$ and a frequency $\omega \in [0, \pi)$, find the frequency response $H(e^{\mathbf{i}\omega})$ of $\mathscr{B}$ at the frequency ω.

Trivial generalizations of the problem are to have as data multiple trajectories $\{ w_{\text{d}}^1, \dots, w_{\text{d}}^N \}$, where $w_{\text{d}}^i \in (\mathbb{R}^q)^{T_i}$ (this is achieved by using a mosaic Hankel matrix Markovsky, 2014) and to aim at evaluation of the frequency response at multiple frequencies $\Omega := \{ \omega_1, \dots, \omega_K \}$. Nontrivial generalizations are to consider noisy data in the errors-in-variables and output error setups as well as nonlinear systems. We will address this extension in Section 3.2.

3. The proposed method

Section 3.1 presents a solution of the data-driven frequency response evaluation problem (*i.e.*, assuming exact data w_{d}) that uses the data-driven representation (1) as the main tool. Section 3.2 presents a modification of the solution for inexact case, *i.e.*, noisy data and/or data from a nonlinear system, based on preprocessing of the data matrix $\mathscr{H}_L(w_{\text{d}})$ with low-rank approximation (Markovsky, 2019).

3.1. Solution of Problem 1

We consider general multivariable linear time-invariant system $\mathscr{B}$ with an input/output partitioning $w = (u, y)$ and a corresponding transfer function H. Our main result is stated in the following theorem.

Theorem 2. *For exact data $w_{\text{d}} = (u_{\text{d}}, y_{\text{d}}) \in \mathscr{B}|_{T_{\text{d}}}$ satisfying (2), $T \geq \ell + 1$, and $z \in \mathbb{C}$ that is not a pole of H, the system of linear equations*

$$\begin{bmatrix} 0_{mT \times p} & \mathscr{H}_T(u_{\text{d}}) \\ -\mathbf{z} \otimes I_p & \mathscr{H}_T(y_{\text{d}}) \end{bmatrix} \begin{bmatrix} H_z \\ G \end{bmatrix} = \begin{bmatrix} \mathbf{z} \otimes I_m \\ 0_{pT \times m} \end{bmatrix}, \quad (3)$$

where $\otimes$ is the Kronecker product and $\mathbf{z} := [z^1 \ \dots \ z^T]^\top$ has a unique solution for H_z, such that $H_z = H(z)$.

Proof. Using the data-driven representation, first, we describe the m complex exponential responses of the system to complex exponential inputs applied separately on the m input channels. Then, we show that the m vector-valued trajectories are compactly described as a matrix valued trajectory. Let $e_i \in \mathbb{R}^m$ be the ith unit vector (ith column of the $m \times m$ identity matrix I_m) and $\exp_z(t) := z^t$ be the complex exponential function with $z \in \mathbb{C}$. From the behavioral point of view, $H(z)$ describes the subbehavior of $\mathscr{B}$ spanned by

$$w^i = \big(e_i \exp_z,\ h_{z,i} \exp_z\big), \qquad \text{for } i = 1, \dots, m \quad \text{and } z \in \mathbb{C}\ , z \notin \lambda(\mathscr{B}), \tag{4}$$

where $\lambda(\mathscr{B})$ denotes the set of poles of $\mathscr{B}$. The input of w_i is the complex exponential $e_i \exp_z$ and the output is a scaled version of the input $h_{z,i} \exp_z$, where the scaling factor $h_{z,i}$ is the ith column of $H_z := H(z)$. The $w^1, \dots, w^m$ can be written as a matrix-valued trajectory

$$W = \big(I_m \exp_z,\ H_z \exp_z\big), \quad \text{for } z \in \mathbb{C}, z \notin \lambda(\mathscr{B}). \tag{5}$$

Using the data-driven representation (1) for (5) with length $T \geq \ell + 1$, we obtain the following system of equations

$$\begin{bmatrix} \mathscr{H}_T(u_\mathrm{d}) \\ \mathscr{H}_T(y_\mathrm{d}) \end{bmatrix} G = \begin{bmatrix} \mathbf{z} \otimes I_m \\ \mathbf{z} \otimes H_z \end{bmatrix},$$

where $G \in \mathbb{R}^{(T_\mathrm{d}-T+1)\times m}$. Rewritten in the standard form of a system of linear equations, this gives us Eq. (3). □

The result stated in Theorem 2 is constructive and leads to a method for direct data-driven frequency response estimation. Indeed, the parameter of interest H_z can be computed by solving (3) for the unknowns H_z and G. Note that the solution based on (3) allows us to evaluate the transfer function $H(z)$ at any complex number z, not only at z on the unit circle $e^{\mathbf{i}\omega}$.

Note 3 (*No Hyper-parameters*)**.** Although the conditions of Theorem 2 require the lag ℓ and the order n of the system to be known, the method itself does not use them. Without prior knowledge of ℓ, the parameter T should be chosen as the maximum value for which $\mathscr{H}_T(w_\mathrm{d})$ has at least as many columns as rows, *i.e.*, $T = T_\mathrm{max} = \lfloor (T_\mathrm{d}+1)/(q+1) \rfloor$.

3.2. Modification for inexact data

One way of modifying the method presented in Section 3.1 for the case of noisy data is to first preprocess the data, aiming to approximate the exact noise-free data, and then apply the method on the preprocessed data. A popular preprocessing heuristic is unstructured low-rank approximation of the data matrix $\mathscr{H}_T(w_\mathrm{d})$ enforcing the prior knowledge that rank $\mathscr{H}_T(w_\mathrm{d}) = mT + n$. If the model order is known, the rank $mT + n$ approximation can be obtained by truncation of the singular value decomposition of $\mathscr{H}_T(w_\mathrm{d})$, with $T = n + 1$ (since $n \geq \ell$). If the model order is not known, it can be estimated from the decay of the singular values, by visual inspection or by a range of rank estimation heuristics.

The method with the low-rank approximation preprocessing is summarized in Algorithm 1. Its Matlab implementation, available from https://imarkovs.github.io/frest, is essentially five lines of code. Moreover, it applies to general multivariable systems and can use data from multiple trajectories $\{ w_\mathrm{d}^1, \dots, w_\mathrm{d}^N \}$ as well as estimate the transfer function at multiple points in the complex plane. In the next section, the implementation `dd_frest` of Algorithm 1 is tested on simulated data and is compared with alternative direct and indirect frequency response estimation methods.

Algorithm 1 Data-driven frequency response estimation.

Input: Trajectory $(u_\mathrm{d}, y_\mathrm{d})$, $z \in \mathbb{C}$, and order n.

1: Let $T := n + 1$.

2: Compute the singular value decomposition

$$\begin{bmatrix} \mathscr{H}_T(u_\mathrm{d}) \\ \mathscr{H}_T(y_\mathrm{d}) \end{bmatrix} = U \Sigma V^\top \quad \text{and let} \quad P := U(:, 1 : mT + n).$$

3: Solve the system

$$\left[\begin{bmatrix} 0_{mT\times p} \\ -\mathbf{z} \otimes I_p \end{bmatrix} \quad P \right] \begin{bmatrix} H_z \\ G \end{bmatrix} = \begin{bmatrix} \mathbf{z} \otimes I_m \\ 0_{pT\times m} \end{bmatrix}.$$

Output: $H_z = H(z)$

4. Numerical examples

This section illustrates and empirically validates the proposed method—Algorithm 1, implemented in the function `dd_frest`—in case of exact and noisy data obtained in the errors-in-variables setup. An alternative nonparameteric frequency response method, used for comparison with `dd_frest` is the spectral analysis estimator `spa` with the Welch method to calculate spectral densities (Stoica & Moses, 2005). Both non-parametric methods—`dd_frest` and `spa`—are referenced against the theoretical optimal performance achieved by the maximum-likelihood estimator `ident` (Markovsky, 2013) in the errors-in-variables setup. Note that `ident` implements an indirect parametric method based on local optimization. In the case of noisy data both `dd_frest` and `ident` use the correct order n of the data-generating system as prior knowledge.

In the simulation, we use the benchmark of Landau, Rey, Karimi, Voda, and Franco (1995), which is a 4th order single-input single-output system $\mathscr{B}$ defined by the transfer function

$$H(z) = \frac{0.2826z + 0.5067z^2}{1 - 1.4183z + 1.5894z^2 - 1.3161z^3 + 0.8864z^4}.$$

The data is obtained in the errors-in-variables setting (Söderström, 2007), *i.e.*, $w_\mathrm{d} = \overline{w}_\mathrm{d} + \widetilde{w}_\mathrm{d}$, where $\overline{w}_\mathrm{d} \in \mathscr{B}|_{T_\mathrm{d}}$ is the true value and $\widetilde{w}_\mathrm{d}$ is a zero mean white Gaussian noise with variance s^2.

Fig. 1 shows the frequency response estimates by the three methods—`dd_frest`, `ident`, and `spa`—in an experiment with $T_\mathrm{d} = 1000$ samples and noise level $s = 10\%$ (*i.e.*, signal-to-noise ratio 10 dB). On the resolution of the figure, the amplitude estimates of `dd_frest` and `ident` are indistinguishable and are close to the true value. The `spa` method is less accurate.

Next, we show the relative percentage estimation errors

$$e_a := 100\% \frac{||\overline{H}_z| - |\widehat{H}_z||}{|\overline{H}_z|} \quad \text{and} \quad e_p := 100\% \frac{|\angle\overline{H}_z - \angle\widehat{H}_z|}{\angle\overline{H}_z},$$

where $\widehat{H}_z$ is the estimated frequency response and $\overline{H}_z$ the true frequency response, averaged over 100 trials of the simulation experiment with different realization of the measurement noise. The frequency response is estimated at $\omega = \pi/4$. Fig. 2 shows the relative averaged estimation errors e_a and e_p as a function of the noise level in an experiment with $T_\mathrm{d} = 500$ data samples and Fig. 3 shows the relative averaged estimation errors e_a and e_p as a function of the number of samples T_d for noise level $s = 5\%$.

For exact data (zero noise level) the proposed method and the maximum-likelihood method give exact result (zero errors) and for increasing noise levels the increase of the errors. The fact that the proposed method given the correct result for exact data is an empirical confirmation of Theorem 2. The gap between the

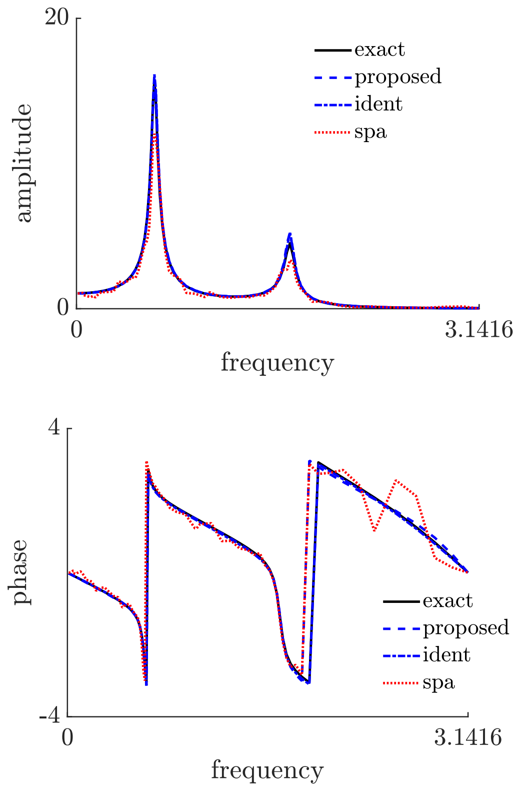

Fig. 1. Frequency response estimates in an experiment with 1000 samples and 10% noise level. The amplitude estimates of the proposed method are nearly identical to the one of the maximum-likelihood method. The classical nonparametric method is less accurate.

error of the proposed method and the errors of the maximum-likelihood method quantifies the sub-optimality of the proposed method due to the low-rank approximation heuristic. The errors of the classical nonparametric method are much higher.

Note 4 (*Indirect Time-domain Nonparametric Approach*)**.** An alternative indirect time-domain nonparametric approach is to (1) compute the impulse response and (2) convert it to the frequency response. A nonparametric impulse response estimation method that for exact data delivers the exact impulse response is proposed in Markovsky and Rapisarda (2008, Section 4.5). The conversion from time-domain to frequency-domain on step 2 can be done by the discrete Fourier transform. This computation, however, involves an error due to the finite number of samples of the impulse response. In addition, classical nonparametric impulse response estimation methods, such as the one implemented in the function `impulseest` from the System Identification Toolbox of Matlab, involve an additional error in the computation of the impulse response. We illustrate this point on a numerical example:

```
hh = freqresp(impulseest(iddata(yd0, ud0)), ...
          Omega);
e = norm(h0(:) - hh(:)) / norm(h0(:)) % -> 0.1918
```

Although the data is exact, the time-domain nonparameteric indirect frequency response method, based on `impulseest`, results in 19% error. Using the noisy data in the errors-in-variables setup with 10% noise level

```
hh = freqresp(impulseest(iddata(yd, ud)), ...
          Omega);
e = norm(h0(:) - hh(:)) / norm(h0(:)) % -> 0.2150
```

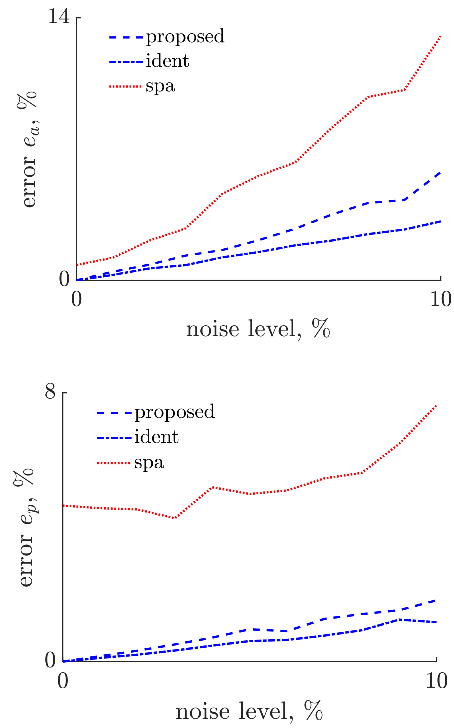

Fig. 2. The relative averaged estimation errors e_a and e_p as a function of the noise level show that for exact data (zero noise level) the proposed method and the maximum-likelihood method give exact result (zero errors) and for increasing noise levels the increase of the errors. The gap between the error of the proposed method and the errors of the maximum-likelihood method quantifies the lack of efficiency of the proposed method due to the low-rank approximation heuristic. The errors of the classical nonparameteric method are much higher.

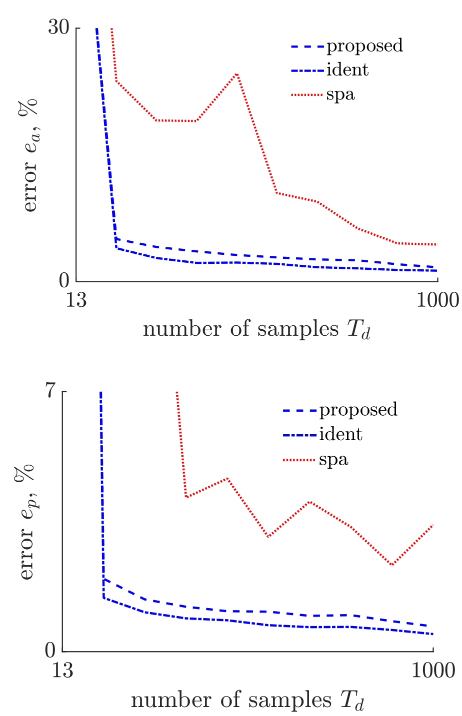

Fig. 3. Relative averaged estimation errors e_a and e_p as a function of the number of samples T_d. The errors of both the proposed method and the maximum-likelihood method show the typical $1/\sqrt{T_d}$ convergence rate. Again the gap between the error of the proposed method and the errors of the maximum-likelihood method quantifies the lack of efficiency of the proposed method due to the low-rank approximation heuristic. The errors of the classical nonparameteric method are much higher.

the estimation error is 21%. Using the method proposed in the paper with the low-rank approximation modification for dealing with the noise

```
hh = dd_frest(ud, yd, exp(i * Omega), n);
e = norm(h0(:) - hh(:)) / norm(h0(:)) % -> 0.0823
```

the estimation error is 8%.

5. Conclusions and perspectives

We proposed a direct data-driven method for frequency response evaluation and estimation that does not suffer from leakage errors and has unlimited frequency resolution. The assumption for exact evaluation in case of noise free data is standard persistency of excitation condition on the data. The resulting algorithm has no hyper-parameters and requires solution of a system of linear equations only. In the noise-free setting, the proposed method can be viewed akin to the two-step, indirect frequency response evaluation method consisting of the identification of an overparametrized ARX model as the first step and the evaluation of the frequency response using this model as the second step.

In case of noisy data obtained in the errors-in-variables setting, we propose a modification of the method that has a preprocessing step by low-rank approximation, using knowledge of the model order. Empirical results show that the modified method is more accurate than state-of-the-art nonparameteric frequency response evaluation methods. The advantages are particularly pronounced for short data records, lightly damped systems, and low noise level. In particular, the method is applicable to marginally stable and unstable systems. Future work includes uncertainty quantification of the method as well as other modifications for noisy data.

Acknowledgments

The authors wish to thank Florian Dörfler, Petre Stoica, Torsten Söderström, Alessandro Chiuso, Maarten Schoukens, Keith Godfrey, and Koen Tiels for feedback and discussions on the results in the paper. The research leading to these results has received funding from the Catalan Institution for Research and Advanced Studies (ICREA), Spain, the Fond for Scientific Research Vlaanderen (FWO) projects G090117N and G033822N, the Fonds de la Recherche Scientifique FNRS–FWO EOS Project 30468160, Belgium, NASA under cooperative agreement VT-80NSSC20M0213, United States, and the US National Science Foundation under award CMMI2238424.

References

Gevers, M., Pintelon, R., & Schoukens, J. (2011). The local polynomial method for nonparametric system identification: Improvements and experimentation. In *IEEE conf. on decision and control* (pp. 4302–4307).

Kay, S. (1988). *Modern spectral estimation: theory and applications*. Prentice-Hall.

Landau, I., Rey, D., Karimi, A., Voda, A., & Franco, A. (1995). A flexible transmission system as a benchmark for robust digital control. *European Journal of Control*, *1*(2), 77–96.

Ljung, L. (1999). *System identification: theory for the user*. Prentice-Hall.

Markovsky, I. (2013). A software package for system identification in the behavioral setting. *Control Engineering Practice*, *21*, 1422–1436.

Markovsky, I. (2014). Recent progress on variable projection methods for structured low-rank approximation. *Signal Processing*, *96PB*, 406–419.

Markovsky, I. (2019). *Low-rank approximation: algorithms, implementation, applications* (2nd ed.). Springer.

Markovsky, I., & Dörfler, F. (2021). Behavioral systems theory in data-driven analysis, signal processing, and control. *Annual Reviews in Control*, *52*, 42–64.

Markovsky, I., & Dörfler, F. (2023). Identifiability in the behavioral setting. *IEEE Transactions on Automatic Control*, *68*, 1667–1677.

Markovsky, I., Huang, L., & Dörfler, F. (2023). Data-driven control based on behavioral approach: From theory to applications in power systems. *IEEE Control Systems Magazine*, *43*(5), 28–68. http://dx.doi.org/10.1109/MCS.2023.3291638.

Markovsky, I., & Rapisarda, P. (2008). Data-driven simulation and control. *International Journal of Control*, *81*(12), 1946–1959.

Pintelon, R., & Schoukens, J. (2012). *System identification: a frequency domain approach* (2nd ed.). Piscataway, NJ: IEEE Press.

Polderman, J., & Willems, J. C. (1998). *Introduction to mathematical systems theory*. Springer-Verlag.

Söderström, T. (2007). Errors-in-variables methods in system identification. *Automatica*, *43*, 939–958.

Stoica, P., & Moses, R. (2005). *Spectral analysis of signals*. Prentice Hall.

Willems, J. C. (1986). From time series to linear system—Part I. Finite dimensional linear time invariant systems, part II. Exact modelling, part III. Approximate modelling. *Automatica*, *22*(23), 561–580, 87–115, 1987 675–694.

Willems, J. C. (1997). On interconnections, control, and feedback. *IEEE Transactions on Automatic Control*, *42*, 326–339.

Willems, J. C., Rapisarda, P., Markovsky, I., & De Moor, B. (2005). A note on persistency of excitation. *Control Letters*, *54*(4), 325–329.