

Formula for Estimating the Frequency Response of LTI Systems From Noisy Finite-Length Datasets

Hamid R. Ossareh^{ID}, *Senior Member, IEEE*, and Florian Dörfler^{ID}, *Senior Member, IEEE*

Abstract—In this letter, we revisit the classical problem of estimating the frequency response of an LTI system from noisy, non-periodic input-output data. Existing solutions fall into two categories: indirect methods, which compute the frequency response using identified models, and direct methods, which estimate the response directly from data. Direct methods bypass system identification, but have challenges when applied to noisy, finite-length datasets with unknown initial conditions. This letter proposes a new direct method that addresses these challenges, and provides an explicit formula for computing the frequency response. To develop this method, this letter leverages ideas from behavioral system theory and poses the problem as an optimization problem, whose objective is to minimize the projection of the solution onto the nullspace of an input-output data matrix. This letter also offers an alternative derivation of the formula, based on identifying an ARX model, thereby bridging the gap with classical indirect approaches. The proposed method is applied to experimental data collected from a DC motor, where we show that the proposed method outperforms other direct approaches based on Fourier transforms and low-rank approximations, and performs equally as good as direct subspace identification methods, even though no model class is prescribed.

Index Terms—Frequency response estimation, non-parametric methods, subspace methods, behavioral system theory.

I. INTRODUCTION

THE ESTIMATION of frequency response from noisy and non-periodic input-output data is a well-known problem in the field of system identification [1] and in applications [2]. Existing solutions can be classified as either parametric or non-parametric. Parametric methods identify a model, such as a state-space or transfer function model, and then use the

identified model to indirectly compute the frequency response. As such, we refer to these methods as “indirect”. Non-parametric methods, on the other hand, are “direct” in the sense that they estimate the frequency response directly from data. Non-parametric methods bypass system identification and do not require explicit model order selection. As such, they simplify the overall identification process and provide an end-to-end solution to the practitioner. However, they face challenges when dealing with noisy, non-periodic, and finite-length data, and with non-zero initial conditions [3], [4], [5], [6]. Even in the absence of noise, the finite nature of the data results in the so-called “leakage errors” due to windowing, which distorts the estimated spectrum.

The recent work in [7] introduced a new direct method that addresses the leakage problem in the case of exact, noise-free data. The method was developed in the behavioral setting (see [8], [9], [10] for an overview), which thanks to the *fundamental lemma* [8], [11], [12], allows a harmonic output of an LTI system to be expressed as a function of offline, persistently exciting input-output data. In the case of inexact (i.e., noisy) data, [7] proposed a pre-processing heuristic to denoise the data matrix using low-rank approximation (LRA) to enforce the rank condition required by the fundamental lemma. We recently applied the method presented in [7] to experimental data collected from a real DC motor. The method performed poorly out of the box, which motivated further investigation that culminated in this letter.

As such, this letter introduces a new direct method, as well as an explicit formula, for computing the frequency response of LTI systems from noisy, non-periodic input-output datasets. The formula does not require any hyper-parameter tuning and, in the case of noise-free data, provides the exact frequency response, even when applied to finite-length datasets with non-zero, unknown initial conditions. We offer two different derivations of this formula. The first one builds upon the results in [7], where we take a representation-free and non-parametric perspective, but unlike the LRA heuristic in [7], we pose the estimation problem as an optimization problem, whose objective is to minimize the projection of the solution onto the nullspace of the input-output data matrix. This approach is inspired by the recent data-driven control literature (see [8], [13], [14], [15], [16], [17], specifically [14], [15], [16]). The second derivation is based on identifying an ARX model by minimizing a least-square prediction criterion, and computing the frequency response using this model. Thus, the two derivations provide a bridge

Manuscript received 8 September 2023; revised 11 November 2023; accepted 30 November 2023. Date of publication 8 December 2023; date of current version 27 December 2023. This work was supported in part by NASA under Grant VT-80NSSC20M0213; in part by NSF under Grant CMMI-2238424; and in part by SNF through the NCCR Automation. Recommended by Senior Editor S. Olaru. (Corresponding author: Hamid R. Ossareh.)

Hamid R. Ossareh is with the Electrical Engineering Program, University of Vermont, Burlington, VT 05405 USA (e-mail: hrossareh@uvm.edu).

Florian Dörfler is with the Department of Electrical Engineering and Information Technology, Swiss Federal Institute of Technology (ETH), 8092 Zürich, Switzerland (e-mail: dorfler@control.ee.ethz.ch).

Digital Object Identifier 10.1109/LCSYS.2023.3340997

2475-1456 © 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See <https://www.ieee.org/publications/rights/index.html> for more information.

between direct and indirect approaches. We apply the proposed formula to experimental data collected from a DC motor, where we show that the proposed method outperforms other direct approaches based on LRA (as proposed in [7]) and Fourier transforms, and performs equally as good as direct subspace identification methods, even though no model class is prescribed.

This letter is organized as follows. Preliminaries are provided in Section II. The main results are reported in Section III. Application to a DC motor is provided in Section IV. Conclusions and future work are reported in Section V.

The notation throughout this letter is as follows. The sets $\mathbb{R}$, $\mathbb{R}^n$, $\mathbb{R}^{n \times n}$, and $\mathbb{C}$ denote the set of real numbers, n -dimensional vectors of real numbers, $n \times n$ matrices with real entries, and complex numbers, respectively. The variable j denotes the imaginary unit: $j = \sqrt{-1}$. We use the integer variable t to denote the discrete time index. For a discrete-time LTI system with discrete transfer function $P(z)$, its frequency response is defined as $P(e^{j\omega})$, where $\omega \in \mathbb{R}$. For a matrix A , $A^\dagger$ denotes its Moore–Penrose pseudoinverse and $\|A\|_F$ its Frobenius norm. For a sequence of matrices $X_1, \dots, X_n$ with the same number of columns, we use the **col** operator to denote $\text{col}(X_1, \dots, X_n) = [X_1^\top, \dots, X_n^\top]^\top$.

II. PRELIMINARIES AND PROBLEM STATEMENT

A. Model- & Data-Based Representations of LTI Systems

Consider a discrete-time, SISO,¹ LTI, and bounded-input-bounded-output stable system with input $u(t) \in \mathbb{R}$, output $y(t) \in \mathbb{R}$, and represented by the disturbance-free ARX model

$$y(t) = \mathcal{G} \begin{bmatrix} u(t-n) \\ \vdots \\ u(t) \\ y(t-n) \\ \vdots \\ y(t-1) \end{bmatrix}, \quad (1)$$

where $\mathcal{G} \in \mathbb{R}^{1 \times (2n-1)}$ collects the ARX coefficients. For now, we assume that $\mathcal{G}$ is known, but will eventually estimate it from (noisy) data in Section III-B. Eq. (1) is often posed in an n -dimensional state-space form. Both are parametric models.

Suppose that the model parameters are unknown, but a single input-output trajectory from the system is available. Let the trajectory be denoted by $u^d(t)$ and $y^d(t)$, $t = 0, \dots, T_d - 1$, where $T_d \geq 1$ is the length of the data, and d in the subscript and superscript stands for “data”. Given a fixed $T \geq 1$, we define the “Hankel data matrix” of order T as follows:

$$\mathcal{H} = \begin{bmatrix} \mathcal{H}_u \\ \mathcal{H}_y \end{bmatrix} \in \mathbb{R}^{2T \times M}$$

where $M = T_d - T + 1$,

$$\mathcal{H}_u = \begin{bmatrix} u^d(0) & u^d(1) & \cdots & u^d(T_d - T) \\ u^d(1) & u^d(2) & \cdots & u^d(T_d - T + 1) \\ \vdots & \vdots & \ddots & \vdots \\ u^d(T-1) & u^d(T) & \cdots & u^d(T_d - 1) \end{bmatrix},$$

and $\mathcal{H}_y$ can be similarly defined by replacing all instances of u by y in the above. All of our subsequent developments

¹For the sake of clarity, this letter considers single-input, single-output (SISO) systems. The multi-input, multi-output case can be handled similarly by applying our methods to all permutations of the input and output channels.

also hold for a more general matrix data structure, where the columns of $\mathcal{H}_u$ and $\mathcal{H}_y$ may comprise of M different, possibly independent, trajectories, as discussed in [8]. We will not pursue the general case in this letter.

In the case of exact data, any combination of the columns of $\mathcal{H}$ will be a valid length- T trajectory of system (1). The converse result is true as well under certain conditions: the column span of the data matrix $\mathcal{H}$ spans the entire set of length- T trajectories if and only if the data matrix satisfies the *low-rank condition*²

$$\text{rank}(\mathcal{H}) = T + n, \quad (2)$$

for $T \geq n$. This result is known as the “generalized persistency of excitation” condition [18, Corollary 21]. Said differently, under condition (2), for any length- T input sequence u and corresponding output y , where $u = [u(0), u(1), \dots, u(T-1)]^\top$ and $y = [y(0), y(1), \dots, y(T-1)]^\top$, there exists a (non-unique) vector g such that

$$\mathcal{H}g = \begin{bmatrix} u \\ y \end{bmatrix} \quad (3)$$

For later use, we define Y_P as the first $T-1$ rows of $\mathcal{H}_y$ and Y_F as the last row of $\mathcal{H}_y$, so that

$$\mathcal{H} = \begin{bmatrix} \mathcal{H}_u \\ \mathcal{H}_y \end{bmatrix} = \begin{bmatrix} \mathcal{H}_u \\ Y_P \\ Y_F \end{bmatrix}. \quad (4)$$

In the case of exact data, a persistently exciting input, and for $T = n + 1$, the matrix $\text{col}(\mathcal{H}_u, Y_P)$ has full row rank [16, Remark 3]. In case of a persistently exciting input and noisy output measurements, this matrix also has full row rank almost surely; see [16, Lemma 3]. In either case, the matrix is right-invertible. We thus make the following assumption for the rest of this letter:

Assumption 1: The matrix $\text{col}(\mathcal{H}_u, Y_P)$ has full row rank.

B. Frequency Response Estimation Problem

Let the transfer function of (1) be denoted by $P(z)$. Recall that its frequency response for a frequency $\omega \in \mathbb{R}$ is defined as $P(e^{j\omega})$, which is a complex number. Its magnitude and phase can be used to construct the Bode plot, which famously underlies numerous analysis and synthesis concepts.³

If the model parameters are known, $P(e^{j\omega})$ can be readily computed. Without a model, $P(e^{j\omega})$ can be obtained “experimentally” (at least conceptually for a stable system) by applying a complex exponential input signal $u(t) = e^{j\omega t}$, which is a continuous and periodic function of time with frequency ω , and collecting the associated output signal. The output at steady-state (i.e., after all transients have decayed), is given by $y(t) = P(e^{j\omega})e^{j\omega t}$, another complex exponential of frequency ω .

In the same manner, the frequency response $P(e^{j\omega})$ can be obtained from a finite length- T input-output dataset. Namely, let $u = \mathbf{z} := [e^{j\omega} \ e^{j2\omega} \ \cdots \ e^{jT\omega}]^\top$ be a harmonic input signal of frequency ω and length $T > n$, where the notation u is consistent with (3). If the associated length- T output signal y were equal to $\mathbf{z}P_\omega$, i.e., harmonic with frequency ω and

²From an input-design perspective, the low-rank condition (2) is met if the system is controllable and the input data, u^d , is a single long *persistently exciting* time series. This is known as the *fundamental lemma* [11].

³We acknowledge that synthesis methods that leverage the frequency response are more widely used in continuous time or in discrete-time scenarios when the desired cross-over frequency is sufficiently below the Nyquist rate.

scaled by a complex $P_\omega \in \mathbb{C}$, then P_ω is the sought frequency response, i.e., $P_\omega = P(e^{j\omega})$ [7].

Define $\tilde{\mathbf{z}} := [e^{j\omega} \ e^{j2\omega} \ \dots \ e^{j(T-1)\omega}]^\top$, which has one less element than $\mathbf{z}$ defined earlier but is otherwise the same. In a model-based setting and for $T = n + 1$, we may readily insert the signals $u = \mathbf{z}$ and $y = \mathbf{z}P_\omega$ into (1), and leverage the notation $\tilde{\mathbf{z}}$, to obtain:

$$e^{jT\omega}P_\omega = \mathcal{G} \begin{bmatrix} \mathbf{z} \\ \tilde{\mathbf{z}}P_\omega \end{bmatrix} \quad (5)$$

and solve equation (5) for the sought frequency response P_ω .

In a data-driven setting, the above principle has been successfully used in [7] for arbitrary $T > n$. Specifically, suppose the data matrix $\mathcal{H}$ has been populated with the offline input-output data. Then, by substituting $u = \mathbf{z}$ and $y = \mathbf{z}P_\omega$ into (3) (parameterizing all length- T trajectories) we obtain

$$\mathcal{H}g = \begin{bmatrix} \mathbf{z} \\ \mathbf{z}P_\omega \end{bmatrix}, \quad (6)$$

where the unknowns are now $g \in \mathbb{C}^{M \times 1}$ and $P_\omega \in \mathbb{C}$. The system of equations (6) can be equivalently written in the following standard form:

$$\begin{bmatrix} \begin{bmatrix} 0 \\ -\mathbf{z} \end{bmatrix} & \mathcal{H} \end{bmatrix} \begin{bmatrix} P_\omega \\ g \end{bmatrix} = \begin{bmatrix} \mathbf{z} \\ 0 \end{bmatrix} \quad (7)$$

If $\mathcal{H}$ satisfies (2) and $T > n$, then (7) has a unique solution for P_ω (but not g) [7]. Of practical relevance is the case of inexact data, where the trajectories in $\mathcal{H}$ are corrupted by measurement noise, process disturbances, or nonlinearities. In this case, $\mathcal{H}$ will have full row rank (instead of rank $T + n$) and so P_ω is not unique. To address this, [7] proposed a heuristic based on unstructured low-rank approximation (LRA) to “denoise” the data matrix before using it in Eq. (7). This was achieved by computing the singular value decomposition of $\mathcal{H}$: $\mathcal{H} = U\Sigma V^\top$ and replacing $\mathcal{H}$ in (7) with the first $T + n$ columns of U , which guarantees that the rank condition in (2) is met.

However, LRA provides no guarantees of accuracy and, in fact, it may perform poorly in practice, as we show in Section IV. The main reason for this poor performance is that LRA merely enforces the rank condition as required by the fundamental lemma. In the process, it removes the Hankel structure of the matrix (which encodes time-invariance of the underlying system), and modifies both $\mathcal{H}_u$ and $\mathcal{H}_y$ parts of $\mathcal{H}$, even though only $\mathcal{H}_y$ is corrupted by noise. Furthermore, LRA disregards the interplay between g and P_ω .

The main goal of this letter is to propose a new direct data-driven approach to robustly (and uniquely) compute P_ω from noisy data.

III. MAIN RESULTS

In this section, we present our main result, namely an explicit formula to compute P_ω from noisy data. We offer two derivations of this formula: one in the non-parametric setting by leveraging (7), and one in the parametric setting by leveraging (5).

A. The Non-Parametric Perspective

To begin, note that (6) can be equivalently expressed as

$$\begin{bmatrix} \mathcal{H}_u \\ Y_P \\ Y_F \end{bmatrix} g = \begin{bmatrix} \mathbf{z} \\ \tilde{\mathbf{z}}P_\omega \\ e^{jT\omega}P_\omega \end{bmatrix}, \quad (8)$$

where Y_P and Y_F are defined in (4), and $\tilde{\mathbf{z}}$ is as defined above Eq. (5). In the noise-free case, the semantics of (8) are as follows: P_ω must be chosen so that the right-hand side of (8) is in the image of the matrix $\mathcal{H}$ on the left-hand side. According to (2), $\mathcal{H}$ is of low-rank and admits a null-space (of considerable size for large M) and thus a non-unique solution for g , but any of these solutions lead to the same frequency response P_ω . In case of inexact data, $\mathcal{H}$ is likely of full row rank and thus any frequency response P_ω is compatible with the data. This incompatibility with the ground-truth system is clearly undesired and can only be poorly mitigated by LRA, as will be shown in Section IV.

Instead, contemporary data-driven predictors assure a robust prediction by biasing the solution g , e.g., by searching for a g of small 2-norm $\|g\|_2$ or of small 1-norm $\|g\|_1$ (i.e., high sparsity), which serve as surrogates for noise robustness and low rank of $\mathcal{H}$, respectively; see [19] for a discussion and survey. One particular solution for g is

$$g \perp \text{kernel} \begin{bmatrix} \mathcal{H}_u \\ Y_P \end{bmatrix}, \quad (9)$$

i.e., g is orthogonal to the nullspace of the first two block-equations in (8). This choice of g is related to a least-square estimate of an ARX model, as shown in the next subsection.

The *unique* solution of (8) subject to the orthogonality (9) can be formulated by minimizing the projection of g onto the nullspace of the (first two blocks of the) data matrix $\mathcal{H}$, i.e.,

$$\min_{g, P_\omega} \|(I - \Pi)g\|_2^2 \quad (10)$$

subject to (8)

where $\Pi \in \mathbb{R}^{M \times M}$ is the projector assuring (9):

$$\Pi = \begin{bmatrix} \mathcal{H}_u \\ Y_P \end{bmatrix}^\dagger \begin{bmatrix} \mathcal{H}_u \\ Y_P \end{bmatrix}.$$

To solve (10) explicitly, note that (9) is equivalent to requiring that g belongs to the image of $\text{col}(\mathcal{H}_u, Y_P)^\top$, i.e., $g = \text{col}(\mathcal{H}_u, Y_P)^\top \alpha$ for some $\alpha \in \mathbb{C}^{2T-1}$. Substituting this into (8) yields the following two equations:

$$\begin{bmatrix} \mathcal{H}_u \\ Y_P \end{bmatrix} \begin{bmatrix} \mathcal{H}_u \\ Y_P \end{bmatrix}^\top \alpha = \begin{bmatrix} \mathbf{z} \\ \tilde{\mathbf{z}}P_\omega \end{bmatrix}$$

$$Y_F \begin{bmatrix} \mathcal{H}_u \\ Y_P \end{bmatrix}^\top \alpha = e^{jT\omega}P_\omega$$

or, by leveraging Assumption 1,

$$Y_F \begin{bmatrix} \mathcal{H}_u \\ Y_P \end{bmatrix}^\top \left(\begin{bmatrix} \mathcal{H}_u \\ Y_P \end{bmatrix} \begin{bmatrix} \mathcal{H}_u \\ Y_P \end{bmatrix}^\top \right)^{-1} \begin{bmatrix} \mathbf{z} \\ \tilde{\mathbf{z}}P_\omega \end{bmatrix} = e^{jT\omega}P_\omega \quad (11)$$

Now, define the matrix $\mathcal{X} \in \mathbb{R}^{1 \times (2T-1)}$ as follows:

$$\mathcal{X} = Y_F \begin{bmatrix} \mathcal{H}_u \\ Y_P \end{bmatrix}^\top \left(\begin{bmatrix} \mathcal{H}_u \\ Y_P \end{bmatrix} \begin{bmatrix} \mathcal{H}_u \\ Y_P \end{bmatrix}^\top \right)^{-1}$$

As we will show in the next subsection, $\mathcal{X}$ corresponds to the least-square estimate of the ARX model coefficients $\mathcal{G}$ in (1). Finally, we break up

$$\begin{bmatrix} \mathbf{z} \\ \tilde{\mathbf{z}}P_\omega \end{bmatrix} = \begin{bmatrix} \mathbf{z} \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \tilde{\mathbf{z}} \end{bmatrix} P_\omega$$

and solve (11) for P_ω , which yields the following formula for computing the frequency response directly from data (encoded in $\mathcal{X}$):

$$P_\omega = \frac{\mathcal{X} \begin{bmatrix} \mathbf{z} \\ 0 \end{bmatrix}}{e^{j\omega T} - \mathcal{X} \begin{bmatrix} 0 \\ \tilde{\mathbf{z}} \end{bmatrix}}. \quad (12)$$

Equation (12) is the main result of this letter. Specifically, if the data satisfies Assumption 1, we can use (12) to compute the frequency response of the system at frequency ω , without identifying a parametric model. To use this formula to generate the Bode plot, we compute P_ω using (12) on a frequency grid, one frequency at a time. Note that the matrix $\mathcal{X}$ needs to be computed only once.

Remark 1: Formula (12) can be further leveraged for downstream analysis and synthesis tasks, e.g., to evaluate the H_2 and H_∞ norms of the system directly from data. Specifically, the H_2 norm can be computed as the square root of $\frac{1}{\pi} \int_0^\pi |P_\omega|^2 d\omega$ through numerical integration, and the H_∞ norm can be computed as $\max_{\omega \in [0, \pi]} |P_\omega|$, which can be obtained by finding the analytical expression for $\frac{d|P_\omega|}{d\omega}$ and computing its roots. We will not pursue these further here.

Remark 2: Although condition (2) requires the order, n , of the system to be known, the method itself does not use it. Without prior knowledge of n , the parameter T can be chosen as the maximum value for which $\mathcal{H}$ has at least as many columns as rows.

B. Bridging the Gap With Parametric Approaches

In what follows, we show how to equivalently derive the frequency response formula (12) in an indirect setting, i.e., based on a parametric model identified from data. Consider the offline input-output data stored in $u^d(t)$ and $y^d(t)$. We know that for any $t \in \{n, \dots, T_d - 1\}$, the output, $y(t)$, must satisfy the ARX model (1). The coefficients $\mathcal{G}$ of the model in (1) can thus be estimated according to the following least-squares criterion

$$\mathcal{G}^* = \operatorname{argmin}_{\mathcal{G}} \sum_{t=n}^{T_d-1} \left\| y^d(t) - \mathcal{G} \begin{bmatrix} u^d(t-n) \\ \vdots \\ u^d(t) \\ y^d(t-n) \\ \vdots \\ y^d(t-1) \end{bmatrix} \right\|_2^2$$

By selecting $T = n + 1$ and forming the Hankel data matrix $\mathcal{H}$ with this T , the above optimization problem can be equivalently expressed in terms of $\mathcal{H}$ and explicitly solved, as follows:

$$\mathcal{G}^* = \operatorname{argmin}_{\mathcal{G}} \left\| Y_F - \mathcal{G} \begin{bmatrix} \mathcal{H}_u \\ \mathcal{Y}_P \end{bmatrix} \right\|_F^2 = Y_F \begin{bmatrix} \mathcal{H}_u \\ \mathcal{Y}_P \end{bmatrix}^\dagger, \quad (13)$$

where this is the unique solution, thanks to Assumption 1.

Next, following the certainty-equivalence paradigm, the above solution $\mathcal{G}^*$ can be used to calculate the frequency response P_ω , as in the model-based setting (5). Formally, this sequential procedure can be posed as the following bi-level problem:

$$\text{Find } P_\omega$$

$$\text{such that } e^{jT\omega} P_\omega = \mathcal{G}^* \begin{bmatrix} \mathbf{z} \\ \tilde{\mathbf{z}} P_\omega \end{bmatrix}$$

$$\text{where } \mathcal{G}^* = \operatorname{argmin}_{\mathcal{G}} \left\| Y_F - \mathcal{G} \begin{bmatrix} \mathcal{H}_u \\ \mathcal{Y}_P \end{bmatrix} \right\|_F^2. \quad (14)$$

It turns out that the certainty-equivalent model-based approach (14) also results in the previously derived frequency response formula (12). Namely, upon inserting the explicit least-squares solution (13) into (14), we arrive at the equation

$$e^{jT\omega} P_\omega = Y_F \begin{bmatrix} \mathcal{H}_u \\ \mathcal{Y}_P \end{bmatrix}^\dagger \begin{bmatrix} \mathbf{z} \\ \tilde{\mathbf{z}} P_\omega \end{bmatrix}$$

which coincides with (11) based on which the frequency response formula (12) was derived. The attentive reader may note that $\mathcal{X}$ in (12) equals the model estimate $\mathcal{G}^*$ in (13).

Remark 3 (Contextualization of Our Approach): We conclude with a few remarks on this model-based perspective on the frequency response formula (12). First, we relate our results to the subspace literature: the vector of optimal ARX coefficients $\mathcal{G}^*$ (or equivalently, $\mathcal{X}$) in (13) is also known as the (single-step) subspace predictive control model [20]. It is known that in the presence of noise, this predictor has accuracy $1/\sqrt{M}$, where M is the number of columns in the data matrix $\mathcal{H}$; see [16, Th. 2] for details. In the equivalent derivation in the previous subsection, the objective (10) is also known as the certainty-equivalence regularizer in data-driven control approaches [14], [15]. Last, an instrumental variable scheme can also be used to derive the predictor (13) (suitably adapted to the instrument) [21].

Now we take the perspective of prediction error methods minimizing the least square prediction loss. In the noisy case, the optimal estimate $\mathcal{G}^*$ in (13) is likely dense, in which case the estimated model has relative degree zero. Further, the derivations in this subsection rely on the data-window length $T = n + 1$. In general, n is unknown and thus one may opt for a sufficiently large T , i.e., a possibly over-parameterized model. As our numerical studies show (e.g., see next section), this class of (possibly over-parameterized and relative-degree zero) models performs well in experimental case studies.

IV. EXPERIMENTAL RESULTS

We use an experimental setup comprising a Quanser QUBE Servo 2 DC motor interfaced with a laptop computer running MATLAB/Simulink 2020b. The manipulated variable (input) is the voltage applied to the DC motor and the measured variable (output) is the motor speed. This is a single-input single-output, stable system.

First, a sequence of quasi-random steps are applied to the motor, as shown in the top subplot of Fig. 1. The resulting output, collected at $T_s = 2$ ms, is shown in the bottom subplot. This data is used to obtain a model of the system using subspace identification (n4sid command in MATLAB). Since the dynamics appear linear and first order, the order of the model is selected to be $n = 1$. The identified transfer function is $\frac{21.71}{s-0.9837}$, which will serve as benchmark comparison later on. The output of the model, driven by the same input as shown in the top subplot of Fig. 1, is superimposed in the bottom subplot (red curve). Clearly, the identified first-order model captures the dynamics accurately, though there is visible measurement noise and quantization errors, as well as nonlinear effects (e.g., around 3 seconds).

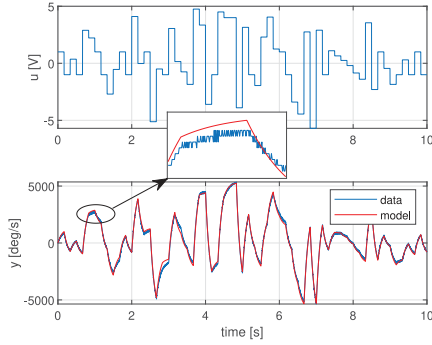


Fig. 1. Input-output data $u^d(t)$ and $y^d(t)$ collected from the DC motor experiment.

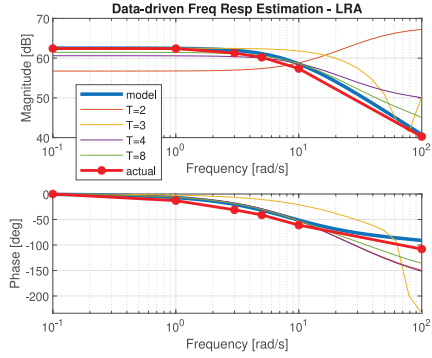


Fig. 2. The Bode plots obtained by applying LRA pre-processing and solving (7), with various values of horizon T , on the dataset shown in Fig. 1. The Bode plot generated using model-based estimation technique (`n4sid`) and the actual, empirical frequency response are also plotted.

The Bode plot of this model, generated using MATLAB's `bode` command, is shown in Fig. 2 (solid blue line). Note that in this figure, and in all the remaining Bode plots in this letter, we plot $P(e^{j\Omega T_s})$ as a function of Ω instead of $P(e^{j\omega})$, where T_s is the sampling period with which the input and output data are collected, and Ω is the continuous-time frequency in the units of radians per second. Recall that the continuous-time and discrete-time frequencies are related through $\omega = \Omega T_s$.

To determine the accuracy of the model-based Bode plot, a sequence of sinusoidal inputs $u(t) = \sin(\omega t)$, with $\omega \in \{0.1, 1, 3, 5, 10, 100\}$, are applied to the motor input and the output is collected. The magnitude and phase shift of the output sinusoid are then visually obtained and are used to empirically calculate the Bode plot. The results are superimposed in Fig. 2 (red curve). As can be seen, the model-based Bode plot is highly accurate.

Next, the same input-output data is used to generate the Bode plot using the LRA-based data-driven approach presented in [7] and reviewed in Section II. To this end, the prediction horizon T is varied from 2 to 8 ($T = 2$ is the theoretical minimum for a first order system). For each T , we construct the Hankel matrix $\mathcal{H}$, which has dimension $2T \times (5001 - T)$, and use the LRA heuristic to denoise it, as reviewed in Section II-B. Finally, we use (7) to compute the frequency response. The resulting “data-driven Bode plots” are superimposed over the model-based one in Fig. 2. As can be seen, the LRA-based data-driven approach cannot accurately estimate the frequency response. Interestingly, LRA exhibits a bias at low frequencies for lower values of T . This observation may have important ramifications in the context of certain

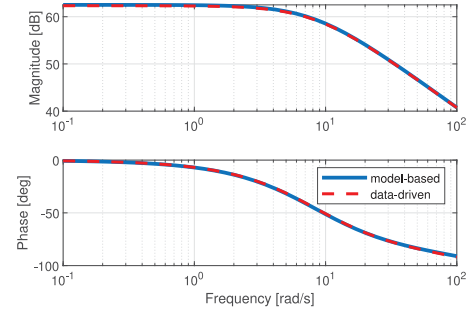


Fig. 3. Bode plot obtained using (12). The model-based Bode plot obtained using `n4sid` is also plotted.

data-driven control schemes such as the DeePC [13]. Indeed, if LRA-based data-driven predictors are used to compute optimal control commands to track constant references, the bias in the predictor at low frequencies may lead to poor tracking performance. Also interestingly, Fig. 2 shows that LRA performance improves with T . A plausible reasoning for this observation is that condition (6) – encoding that the response to a sinusoidal input is a sinusoid of the same frequency – is increasingly hard to be met (for larger T) for a system inconsistent with the data.

Next, we use the proposed data-driven approach to generate the Bode plot. To this end, we use the dataset shown in Fig. 1 to form the Hankel matrix with $T = 2$ without any LRA pre-processing, and compute P_ω on a logarithmic frequency grid using (12). The resulting Bode plot is shown in Fig. 3. As can be seen, the estimation performance has greatly improved. In fact, the accuracy is the same as the indirect method of obtaining a state-space model using subspace identification via the `n4sid` routine followed by evaluation of the frequency response using the identified model. Note that, unlike the indirect method, (12) gives the frequency response in one-shot and does not require explicit model order selection (recall Remark 2).

Finally, we study the effects of dataset size (i.e., windowing) and compare the performance of the proposed approach with common frequency response estimation methods. To this end, we window (i.e., truncate) the time series data in Fig. 1 with $T_s = 6\text{ms}$ from $t = 0$ to $t = L - 1$, where L is varied from 50 timesteps (i.e., 300ms) to 1600 (i.e., 9.6 s). For each L , we form the Hankel data matrix with $T = 2$, and compute P_ω on a frequency grid using the LRA-based approach and also using Eq. (12). Note that regardless of L , the resulting Hankel matrix has 4 rows, but the number of columns will be larger for larger values of L (i.e., the larger the L , the wider the matrix). As a point of comparison, we also use the following routines in MATLAB to obtain the Bode plots: `n4sid` (subspace method), `arx` (least-squares method), `spafdr` (spectral method), and `etfe` (also spectral). The first two are parametric and the latter two are non-parametric. In `n4sid`, as before, we assume system order of $n = 1$. In `arx`, we regularize the problem using the squared exponential kernel with hyperparameters that are tuned optimally using `arxRegul`, and assume $n = 2$. This combination of regularization and order resulted in the best estimation performance with this method. For the spectral methods, we do not apply any window functions to the data.

To quantify the performance, we compute the error at the six frequencies where the empirical frequency response

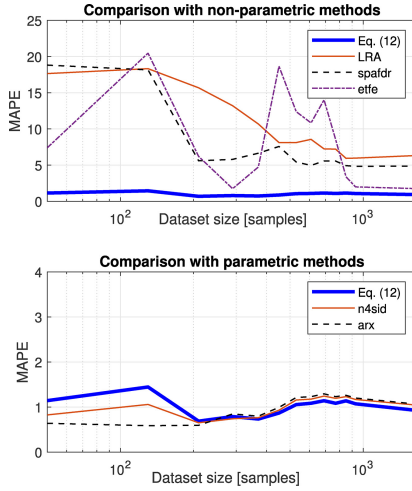


Fig. 4. Comparison of the proposed method (Eq. (12)) with *n4sid*, LRA, *spafdr*, *etfe*, and *arx* for varying window sizes (i.e., dataset size). Smaller MAPE is better.

(i.e., “the ground truth”) is available: 0.1, 1, 3, 5, 10, and 100 rad/s. For this quantification, we define the mean absolute percentage error (MAPE) for each Bode plot as: $\text{MAPE}_x = \frac{1}{6} \sum_{i=1}^6 \left| \frac{m_{\text{emp}}(\omega_i) - m_x(\omega_i)}{m_{\text{emp}}(\omega_i)} \right| \times 100$, where ω_i take values in $\{0.1, 1, 3, 5, 10, 100\}$, $m_{\text{emp}}(\omega_i)$ is the empirical Bode magnitude at frequency ω_i in dB, subscript x takes values from the set $\{\text{Eq. (12)}, \text{n4sid}, \text{LRA}, \text{spafdr}, \text{etfe}, \text{arx}\}$, and m_x is the Bode magnitude in dB, estimated using method x . The MAPE values as a function of dataset size are plotted in Fig. 4. The following conclusions can be made:

- The proposed method performs well, even on small datasets. In fact, it results in similar performance as subspace identification (*n4sid*) and least-squares (*arx*), independent of the window size and despite it being direct and not imposing any model parametrization.
- The LRA-based estimation performance is poor, though it improves with dataset size. Interestingly, the LRA method performs worse than *spafdr* and *etfe* on this dataset.
- The estimation accuracy of all methods reaches a plateau and remains flat. This can be attributed to bias errors resulting from quantization errors, sampling, and nonlinearities such as nonlinear friction, as well as the fact that the empirical frequency response itself may contain small errors.

As a final note, we emphasize that this letter does not claim superiority over either time-domain or frequency-domain identification methods. We merely claim that the proposed formula (12) can achieve similar performance as indirect *parametric* methods (in either domain) without explicitly assuming a model structure.

V. CONCLUSION

This letter provided an explicit formula to compute the frequency response of LTI systems directly from finite-length input-output, possibly-noisy time-domain data. The formula is derived in two settings: the behavioral setting, where the estimation problem is formulated as an optimization problem, whose objective is to minimize the projection of the solution onto the nullspace of the input-output data matrix; and the indirect model-based setting, where we show that the proposed

formula is equivalent to first identifying a model through a certainty-equivalent ARX predictor and then obtaining the frequency response from the model.

We applied the proposed method to experimental data collected from a DC motor and showed that the proposed method outperforms other direct approaches based on Fourier transforms and low-rank approximations, and performs similar to direct methods based on subspace identification, even though no model class is assumed. Future work will investigate reducing bias errors caused by nonlinearities and quantization effects, as well as applying the proposed method to other practical experiments. We will also compare the proposed method against other parametric and non-parametric identification techniques. Finally, we will study the statistical properties of the proposed estimator.

REFERENCES

- [1] L. Ljung, *System Identification*. Berlin, Germany, Springer, 1998.
- [2] V. Häberle, L. Huang, X. He, E. Prieto-Araujo, R. S. Smith, and F. Dörfler, “MIMO grid impedance identification of three-phase power systems: Parametric vs. nonparametric approaches,” 2023, *arXiv:2305.00192*.
- [3] S. Kay, “Spectral estimation,” in *Advanced Topics in Signal Processing*. Hoboken, NJ, USA: Prentice-Hall, 1988, pp. 58–122.
- [4] P. Stoica, and R. L. Moses, *Spectral Analysis of Signals*, vol. 452, Upper Saddle River, NJ, USA: Pearson Prentice Hall, 2005.
- [5] Y. Wang, J. Li, and P. Stoica, *Spectral Analysis of Signals: The Missing Data Case* (Synthesis Lectures Signal Process), vol. 1, Cham, Switzerland, Springer, 2006, pp. 1–102.
- [6] M. Gevers, R. Pintelon, and J. Schoukens, “The local polynomial method for nonparametric system identification: Improvements and experimentation,” in *Proc. 50th IEEE Conf. Decis. Control Eur. Control Conf.*, 2011, pp. 4302–4307.
- [7] I. Markovsky and H. Ossareh, “Finite-data nonparametric frequency response evaluation without leakage,” *Automatica*, vol. 159, Jan. 2024, Art. no. 111351.
- [8] I. Markovsky and F. Dörfler, “Behavioral systems theory in data-driven analysis, signal processing, and control,” *Annu. Rev. Control*, vol. 52, pp. 42–64, Jan. 2021.
- [9] C. De Persis and P. Tesi, “Formulas for data-driven control: Stabilization, optimality, and robustness,” *IEEE Trans. Autom. Control*, vol. 65, no. 3, pp. 909–924, Mar. 2020.
- [10] J. Coulson, H. J. van Waarde, J. Lygeros, and F. Dörfler, “A quantitative notion of persistency of excitation and the robust fundamental lemma,” *IEEE Control Syst. Lett.*, vol. 7, pp. 1243–1248, 2023.
- [11] J. C. Willems, P. Rapisarda, I. Markovsky, and B. L. De Moor, “A note on persistency of excitation,” *Syst. Control Lett.*, vol. 54, no. 4, pp. 325–329, 2005.
- [12] I. Markovsky, E. Prieto-Araujo, and F. Dörfler, “On the persistency of excitation,” *Automatica*, vol. 147, Jan. 2023, Art. no. 110657.
- [13] J. Coulson, J. Lygeros, and F. Dörfler, “Data-enabled predictive control: In the shallows of the DeePC,” in *Proc. 18th Eur. Control Conf. (ECC)*, 2019, pp. 307–312.
- [14] F. Dörfler, J. Coulson, and I. Markovsky, “Bridging direct and indirect data-driven control formulations via regularizations and relaxations,” *IEEE Trans. Autom. Control*, vol. 68, no. 2, pp. 883–897, Feb. 2023.
- [15] F. Dörfler, P. Tesi, and C. De Persis, “On the certainty-equivalence approach to direct data-driven LQR design,” *IEEE Trans. Autom. Control*, vol. 68, no. 12, pp. 7989–7996, Dec. 2023.
- [16] V. Breschi, A. Chiuso, and S. Formentin, “Data-driven predictive control in a stochastic setting: A unified framework,” *Automatica*, vol. 152, Jun. 2023, Art. no. 110961.
- [17] L. Huang, J. Zhen, J. Lygeros, and F. Dörfler, “Robust data-enabled predictive control: Tractable formulations and performance guarantees,” *IEEE Trans. Autom. Control*, vol. 68, no. 5, pp. 3163–3170, May 2023.
- [18] I. Markovsky and F. Dörfler, “Identifiability in the behavioral setting,” *IEEE Trans. Autom. Control*, vol. 68, no. 3, pp. 1667–1677, Mar. 2023.
- [19] I. Markovsky, L. Huang, and F. Dörfler, “Data-driven control based on the behavioral approach: From theory to applications in power systems,” *IEEE Control Syst. Mag.*, vol. 43, no. 5, pp. 28–68, Oct. 2023.
- [20] W. Favoreel, B. De Moor, and M. Gevers, “SPC: Subspace predictive control,” *IFAC Proc. Volumes*, vol. 32, no. 2, pp. 4004–4009, 1999.
- [21] J.-W. Van Wingerden, S. P. Mulders, R. Dinkla, T. Oomen, and M. Verhaegen, “Data-enabled predictive control with instrumental variables: The direct equivalence with subspace predictive control,” in *Proc. IEEE 61st Conf. Decis. Control*, 2022, pp. 2111–2116.