A Greedy Data Collection Scheme For Linear Dynamical Systems

Karim Cherifi[1]   Pawan Goyal[2]   Peter Benner[3]

[1]*Institut für Mathematik MA 4-5, TU Berlin, Straße des 17. Juni 136, D-10623 Berlin, Germany**
Email: cherifi@math.tu-berlin.de, ORCID: 0000-0003-1294-9291
[2]*Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstraße 1, 39106 Magdeburg,
Germany*
Email: goyalp@mpi-magdeburg.mpg.de, ORCID: 0000-0003-3072-7780
[3]*Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstraße 1, 39106 Magdeburg,
Germany*
Email: benner@mpi-magdeburg.mpg.de, ORCID: 0000-0003-3362-4103

Mathematical models are essential to analyze and understand the dynamics of complex systems. Recently, data-driven methodologies have got a lot of attention which is leveraged by advancements in sensor technology. However, the quality of obtained data plays a vital role in learning a good and reliable model. Therefore, in this paper, we propose an efficient heuristic methodology to collect data both in the frequency domain and time-domain, aiming at the best possible information gain from limited experimental data. The efficiency of the proposed methodology is illustrated by means of several examples, and also, its robustness in the presence of noisy data is shown.

1 Introduction

In this paper, our focus is on inferring linear time-invariant (LTI) systems of the form

$$
\begin{aligned}
\mathbf{E}\dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t), \\
\mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t)
\end{aligned}
\tag{1}
$$

from data. Using the Laplace transform, the transfer function $\mathbf{H}(s)$ of the system (1) can be obtained and is given as

$$
\mathbf{H}(s) = \mathbf{C}(s\mathbf{E} - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D},
\tag{2}
$$

where s represents the frequency. The problem of learning dynamical systems of the form (1) is widely studied and well understood, see, e.g., [1–8]. A key ingredient to data-driven methodologies is data, and it is necessary to provide a good quality of it, ensuring a good quality of the inferred model. Typically, there are two possible ways to collect data, namely in frequency-domain and time-domain. In the frequency-domain, samples of transfer functions are given, whereas, in the time-domain, the output $\mathbf{y}$ is measured for an input $\mathbf{u}$. Moreover, these data may be subject to noise. Depending on an experimental set-up, we may collect data in either of these domains.

Among several widely used approaches, Loewner-based approaches [1,5] to learn dynamical systems have gained increasing popularity. The fundamental idea of the Loewner-based approach is to construct

---

*This work was done when the author was at the Max Planck Institute for Dynamics of Complex Technical Systems, Magdeburg, Germany.

a model that interpolates the given data points in the frequency domain. It has also been extended to time-domain data [4]. Furthermore, the authors in [5] have shown that if the minimum order of a system that realizes the system is n, then using $2n$ data points, one can construct the realization using the Loewner approach. However, the minimal order n of a system is not known a priori; therefore, in practice, as many data as possible are collected, and the minimal order can be found using a truncation procedure, which is then followed by the construction of a minimal order realization. We refer the reader to [1] for more details. However, this approach can impose practical constraints if collecting data is expensive. In this case, we aim to gather data points carefully and focus on collecting the data points to obtain as much information as possible about the underlying dynamics. In this direction, one may think of methods proposed in, e.g., [9–12], where interpolation points (or data) are chosen adaptively to obtain a good low-order or minimal realization of a high-fidelity model. However, these methods are not fully data-driven in the sense that they still require the high-fidelity model, which we here assume to be unknown. Towards fully data-driven approaches, the method proposed in [13] determines the choice of measurement data in an optimal way (e.g., $\mathcal{H}_2$-optimal). But the method contains a few drawbacks – these are: the method takes the order of the realization as an input, which is a hyper-parameter and needs to be chosen carefully. Secondly, it may require measurements of the transfer function in the complex domain, but in an experimental set-up, we usually can only obtain measurements on the $\jmath\omega$-axis.

In this paper, we focus on inferring minimal linear dynamical systems using minimum possible measurement data, but without knowing the minimal order of the realization a priori and its realization in any form. For this, we discuss a heuristic approach, allowing us to select adaptively frequency points that provide the maximum information about the dynamics of the system. We also discuss the extension of the approach to time-domain data.

The paper is structured as follows. In Section 2, the Loewner framework to learn linear dynamical systems is briefly presented. In Section 3, we present the heuristic approach to select frequency points (or transfer function measurements) to extract the maximum information about the system. Then, in Section 4, we discuss an extension of the approach when input-output data are collected in the time-domain. In Section 5, we illustrate the efficiency of the proposed greedy methodology and compare it with models that are obtained by taking measurements on a uniform grid. Finally, we conclude the paper with a summary and future research in Section 6.

2 Data-Driven Modeling Based on the Loewner Framework

In this section, we briefly recap the Loewner approach [1] to construct a realization from given transfer function measurements. To that end, for simplicity, we first write down the underlying realization problem for single-input single-output systems – that is as follows:

Problem 2.1 *Given a set of interpolation points $Z := \{\sigma_1, \ldots, \sigma_{2N}\}$ and corresponding transfer function measurements $\mathbf{H}(\sigma_i)$, the goal is to identify a minimal realization whose rational transfer function is denoted by $\widehat{\mathbf{H}}(s)$, satisfying the following interpolation conditions:*

$$\mathbf{H}(\sigma_i) = \widehat{\mathbf{H}}(\sigma_i), \quad i \in \{1, \ldots, 2N\}, \tag{3}$$

and the transfer function is given as:

$$\widehat{\mathbf{H}}(s) = \widehat{\mathbf{C}}(s\widehat{\mathbf{E}} - \widehat{\mathbf{A}})^{-1}\widehat{\mathbf{B}} + \widehat{\mathbf{D}}, \tag{4}$$

where $\widehat{\mathbf{E}}, \widehat{\mathbf{A}} \in \mathbb{R}^{r \times r}$, $\widehat{\mathbf{B}} \in \mathbb{R}^r$, $\widehat{\mathbf{C}} \in \mathbb{R}^r$ and $\widehat{\mathbf{D}} \in \mathbb{R}$ with r being the order of a minimal realization.

This realization problem can be solved using the Loewner approach. For this, we need to split the measurement data into left and right measurements. Let us define the left and right interpolation points as λ_j and μ_j, where $j \in \{1, \ldots, N\}$, corresponding transfer function measurements are denoted by $\mathbf{H}(\lambda_j)$ and $\mathbf{H}(\mu_j)$, and $Z = \{\lambda_1, \ldots, \lambda_N\} \cup \{\mu_1, \ldots, \mu_N\}$. Having this data, in the following, we define the Loewner and shifted Loewner matrices.

Definition 2.1 *[1] Given left data* $(\lambda_j, \mathbf{H}(\lambda_j))$ *and right data* $(\mu_j, \mathbf{H}(\mu_j))$, *the Loewner* $\mathbb{L}$ *and shifted Loewner* $\mathbb{L}_\mathrm{s}$ *matrices are defined as follows:*

$$
\mathbb{L} = \begin{bmatrix} \frac{\mathbf{H}(\lambda_1)-\mathbf{H}(\mu_1)}{\lambda_1-\mu_1} & \cdots & \frac{\mathbf{H}(\lambda_n)-\mathbf{H}(\mu_1)}{\lambda_n-\mu_1} \\ \vdots & \ddots & \vdots \\ \frac{\mathbf{H}(\lambda_1)-\mathbf{H}(\mu_n)}{\lambda_1-\mu_n} & \cdots & \frac{\mathbf{H}(\lambda_n)-\mathbf{H}(\mu_n)}{\lambda_n-\mu_n} \end{bmatrix}, \quad \mathbb{L}_\mathrm{s} = \begin{bmatrix} \frac{\lambda_1\mathbf{H}(\lambda_1)-\mu_1\mathbf{H}(\mu_1)}{\lambda_1-\mu_1} & \cdots & \frac{\lambda_n\mathbf{H}(\lambda_n)-\mu_1\mathbf{H}(\mu_1)}{\lambda_n-\mu_1} \\ \vdots & \ddots & \vdots \\ \frac{\lambda_1\mathbf{H}(\lambda_1)-\mu_n\mathbf{H}(\mu_n)}{\lambda_1-\mu_n} & \cdots & \frac{\lambda_n\mathbf{H}(\lambda_n)-\mu_n\mathbf{H}(\mu_n)}{\lambda_n-\mu_n} \end{bmatrix}.
$$

It was proven in [1] that based on the Loewner and shifted Loewner matrices, a realization can be constructed that solves Problem 2.1. This result is presented in the following theorem:

Theroem 1 *([1]) Given left data* $(\lambda_j, \mathbf{H}(\lambda_j))$ *and right data* $(\mu_j, \mathbf{H}(\mu_j))$, *consider the Loewner and shifted Loewner matrices as defined in Definition 2.1. Moreover, assume that the feed-through term* $\mathbf{D}$ *is known. Then, an interpolating realization can be constructed as follows:*

$$
\begin{aligned}
\widehat{\mathbf{E}} &= \mathbb{L}, \quad \widehat{\mathbf{A}} = \mathbb{L}_s - \mathbb{1}\mathbf{D}\mathbb{1}^\top, \quad \widehat{\mathbf{B}} = \mathbf{V} - \mathbb{1}\mathbf{D}, \\
\widehat{\mathbf{C}} &= \mathbf{W} - \mathbf{D}\mathbb{1}^\top, \quad \widehat{\mathbf{D}} = \mathbf{D}.
\end{aligned}
\tag{5}
$$

where $\mathbb{1} \in \mathbb{R}^N$ *is the vector of ones. Furthermore, the realization* (5) *is minimal, assuming* $\mathbb{L}$ *is of full-rank.*

Furthermore, in case the Loewner matrix $\mathbb{L}$ is singular, there exists a lower order realization of order $\widehat{n} < n$, where $\widehat{n} = \mathrm{rank}\,(\mathbb{L})$, satisfying the interpolation conditions. To remove the redundant information and obtain the minimal realization, a compression step based on the SVD of $[\mathbb{L}, \mathbb{L}_\mathrm{s}]$ can be employed, see, e.g., [1,5]. For simplicity, we have presented the realization theory using the Loewner approach for SISO systems; however, it can be extended to multi-input multi-output using tangential interpolation. We refer to [1,5] for more details. We also remark that the realization (5) can be complex if data are complex, but if the interpolation points are closed under conjugation, then there exists an orthogonal transformation that allows to determine the real realization using orthogonal transformation.

The success of the Loewner approach has been shown in various applications, see, e.g., [14–16]. Its success lies in the quality of data and assumes that there are enough data for a wide range of frequencies. However, suppose the data collection process is expensive. In this case, it is essential to make a smart choice to measurements to obtain maximum information about the system with limited data. With this aim, in the following section, we discuss an adaptive scheme to collect measurements to extract maximum information about the system.

3 A Greedy-based Data Collection Scheme

Naturally, the quality of collected data is a key ingredient to learning a good and reliable model, describing the underlying dynamics of a process. Obviously, one can collect as much data as possible if feasible. It can then be followed by inferring a system realization by using, e.g., the Loewner approach [1]. The approach involves a compression step that squeezes the information lying in all the given data. However, a mass collection could be expensive in many scenarios. Therefore, in this section, we discuss a greedy scheme that can guide us to collect the data so that we can expect to extract as much new information about the dynamical system from every measurement as possible. In the following, we first note down a corresponding problem.

Problem 3.1 *Identify the underlying state-space dynamical system of a process using minimal number of transfer function evaluations.*

Problem 3.1 has been discussed in the literature in different ways. Towards this, it has been shown in [5] that the minimum number of transfer function measurements needed to determine a minimal state-space realization of order n is $2n$. And such a realization can be determined by employing the

Loewner approach [1]. However, if only data is given, we cannot know a priori the minimal order of state-space realization that captures the dynamics of the process. Therefore, typically, one aims at collecting $N \gg 2n$ data points, which is then followed by a compression step to determine the minimal order of the state-space model that captures the dynamics of the process. Furthermore, a step towards collecting good data, the authors in [13], presented an approach inspired by the IRKA algorithm [9] that focused on using interpolation points that facilitate obtaining a model, minimizing the $\mathcal{H}_2$ error between the learned model and the unknown ground-truth model. However, again, the approach needs the order of the state-space realization as input, which may not be known in advance. Furthermore, we mention that there also exist other approaches that aim at determining or collecting measurements greedily but require a ground-truth high-fidelity model, see, e.g., [10–12, 17].

On the other hand, our goal is to construct a minimal realization of a process using as a minimum number of measurement data as possible, where we neither know the order of a minimal realization in advance, nor have access to any high-fidelity model. Furthermore, one can theoretically consider taking the transfer function measurements on the complex domain, and often, it is known that good, in fact optimal, measurement points are complex as shown in [9,13]. However, from an experimental viewpoint and physical interpretation of a transfer function, we assume that transfer function measurements are taken anywhere on the imaginary $\jmath\omega$-axis. The motivation of considering measurements on the $\jmath\omega$-axis are:

(a) A transfer function at $\jmath\omega$ can be estimated by exciting the system using an input containing sine and cosine functions of the frequency ω. It can be extended to multiple frequencies as well.

(b) From the control theory perspective, a transfer function can be better interpreted on the $\jmath\omega$-axis, for example, using the Bode plot. More details can be found in, e.g., [18].

To that end, we first assume that we have an initial rough construct of a realization that still can be far from satisfactory. An initial rough model can be constructed using only a few measurement points. We denote the transfer function of the initial realization by $\mathbf{H}_{\mathrm{init}}(s)$. Furthermore, let us denote the transfer function of the unknown ground truth realization by $\mathbf{H}_{\mathrm{true}}(s)$. In this case, we ideally would like to update the model using the point on the $j\omega$-axis, where the maximum error between the true and initial realized systems occurs; hence, we select the interpolation point that solves the following optimization problem:

$$\sigma_{\mathtt{new}} := \arg\max_{\sigma \in \jmath\omega} \|\mathbf{H}_{\mathtt{true}}(\sigma) - \mathbf{H}_{\mathtt{init}}(\sigma)\|, \tag{6}$$

where $\omega \in \mathbb{R}$. However, to solve the optimization problem (6), we require the transfer function of the ground-truth realization, which is not available. Therefore, we seek for an alternative approach, allowing us to estimate good interpolation points in an iterative procedure. For this, let us assume that we have $2k$ measurement points, to begin with, and denote these measurements by tuples $(\sigma_i, \mathbf{H}(\sigma_i))$, $i \in \{1, \ldots, 2k\}$, where σ_i and $\mathbf{H}(\sigma_i)$ are the interpolation points and the transfer function measurements at these points. Before we proceed further, we define the following notation:

- $\Sigma_i := \{\sigma_1, \ldots, \sigma_{2i}\}$,

- The transfer function of the realization constructed using interpolation points Σ_m is denoted by $\mathbf{H}_m(s)$, and the order of the constructed realization is m.

Next, we construct realizations using interpolation points $\Sigma_{k-1} \subset \Sigma_k$ and Σ_k, which are respectively denoted by $\mathbf{H}_{k-1}(s)$ and $\mathbf{H}_k(s)$. As discussed earlier, we ideally would add a new measurement at the frequency that solves the problem (6). Since we do not know $\mathbf{H}_{\mathrm{true}}$, in the following, we discuss a heuristic approach to relax the optimization problem. First note that

$$\|\mathbf{H}_{\mathrm{true}}(\sigma) - \mathbf{H}_k(\sigma)\| = \|\mathbf{H}_{\mathrm{true}}(\sigma) - \mathbf{H}_{k-1}(\sigma) - (\mathbf{H}_k(\sigma) - \mathbf{H}_{k-1}(\sigma))\|$$
$$\leq \|\mathbf{H}_{\mathrm{true}}(\sigma) - \mathbf{H}_{k-1}(\sigma)\| + \|\mathbf{H}_k(\sigma) - \mathbf{H}_{k-1}(\sigma)\|. \tag{7}$$
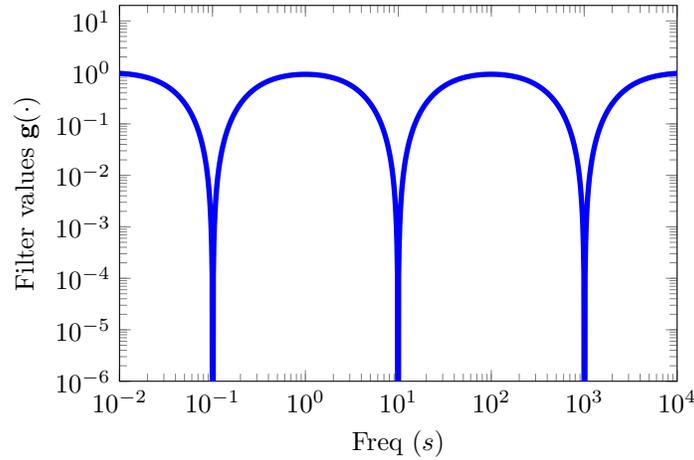
Figure 1: A visual illustration of the filter when the measurements are already taken at the frequencies $\sigma_i = \{10^{-1}, 10^1, 10^3\}$.

With unknown $\mathbf{H}_{\text{true}}(\sigma)$, we rather focus on defining our next measurements data based on the second part of (7) with a constraint; that is – we should exclude the regime of already taken measurements points. Hence, we solve an optimization problem to obtain a frequency point at which transfer function measurement needs to be taken to obtain more information about the underlying dynamics as follows:

$$\max_{\sigma \in \jmath \omega} \mathbf{g}(\sigma) \|\mathbf{H}_k(\sigma) - \mathbf{H}_{k-1}(\sigma)\|, \tag{8}$$

where the function $\mathbf{g}(\sigma)$ can be thought of as a mask that aims at excluding the regime of points at which measurements have already been collected. If $\mathbf{g}(\sigma)$ is known, then we can solve (8) to obtain our new measurement points which possibly bring the most information about the system as they would occur at the maximum change in the transfer function from the previous to the next step.

Naturally, the choice of the function $\mathbf{g}(\cdot)$ plays an important role in determining next measurement points. So, in the following, we discuss a choice of the mask function $\mathbf{g}(\sigma)$. As discussed, a choice of the mask function should be such that it excludes the regime of already considered data points. Among many possible choices, here, we propose the following choice:

$$\mathbf{g}(\sigma) = \prod_{i=1}^{2k} \tilde{\mathbf{g}}(\sigma, \sigma_i),$$

where $\tilde{\mathbf{g}}(\sigma, \sigma_i)$ is defined as:

$$\tilde{\mathbf{g}}(\sigma, \sigma_i) = 1 - \mathrm{e}^{(-\beta(\log(|\sigma|+\epsilon) - \log(|\sigma_i|+\epsilon))^2)}, \tag{9}$$

with $\sigma_i, i \in \{1, \ldots, 2k\}$, being already considered measurement points, and ϵ and β are positive constants and hyper-parameters. It can be noticed that $\mathbf{g}(\sigma)$ is zero at all the considered points with small values in their neighborhood in order to favor the exploration over the already chosen interpolation points. The mask function can also be seen as a notch filter at multiple frequencies and β can be seen as the bandwidth of the filter. To illustrate the mask function, we plot the function for $\beta = 0.6$, $\epsilon = 10^{-15}$ and the notch frequencies $\sigma_i = \{10^{-1}, 10^1, 10^3\}$ in Figure 1. It clearly shows that the mask takes smaller values near the considered frequency points.

Next, we discuss how to further simplify the choice of the function $\mathbf{g}(\sigma)$. First, note that $\mathbf{H}_k(\sigma_i) = \mathbf{H}_{k-1}(\sigma_i)$ for $i \in \{1, \ldots, 2k-2\}$ assuming that the realization is determined by the Loewner approach, and thus, the interpolation at the given measurement points is exact. As a result, we can also consider

---

Algorithm 1: A greedy selection of measurement points in the frequency-domain.

  Input: A parameter β, and initial measurement points $\Sigma_k := \{\sigma_1, \ldots, \sigma_{2k}\}$.

1 Determine realizations $\mathbf{H}_{k-1}$ and $\mathbf{H}_k$ by using measurement points $\Sigma_{k-1} \subset \Sigma_k$ and Σ_k,
  respectively and by employing the Loewner approach.

2 Compute $\texttt{Err} := \max\limits_{s \in j\omega} \|\mathbf{H}_k(s) - \mathbf{H}_{k-1}(s)\|$.

3 while *Err* > *tol* **do**

4   Compute the new measurement point σ_{2k+1} according to the optimization problem (10).

5   Deduce σ_{2k+2} according to (11) .

6   Estimate the transfer function at the measurement points σ_{2k+1} and σ_{2k+2}.

7   Determine a realization using measurement points at $\Sigma_k \cup \{\sigma_{2k+1}, \sigma_{2k+2}\}$.

8 $k \leftarrow k + 1$.

9   Compute $\texttt{Err} := \max\limits_{s \in j\omega} \|\mathbf{H}_k(s) - \mathbf{H}_{k-1}(s)\|$.

  Output: A learned model, whose transfer function is given by $\mathbf{H}_k$.

---

the optimization problem to determine our next measurement point as follows:

$$\sigma_{2k+1} := \arg\max_{\sigma \in j\omega} \prod_{i=2k-1}^{2k} \tilde{\mathbf{g}}(\sigma, \sigma_i) \|\mathbf{H}_k(\sigma) - \mathbf{H}_{k-1}(\sigma)\|. \tag{10}$$

In this paper, we utilize the Loewner approach to construct the realization. Therefore, it is preferred to have an even number of measurement points to avoid the construction of a rectangular realization. Hence, it is desired to include two new additional points at each step. Consequently, in order to choose one additional point, we solve the following optimization problem once σ_{2k+1} is obtained:

$$\sigma_{2k+2} := \arg\max_{\sigma \in j\omega} \prod_{i=2k-1}^{2k+1} \tilde{\mathbf{g}}(\sigma, \sigma_i) \|\mathbf{H}_k(\sigma) - \mathbf{H}_{k-1}(\sigma)\|. \tag{11}$$

As a result, we have two new frequency points σ_{2k+1} and σ_{2k+2} at which we take transfer function measurements. Thus, we update our realization $\mathbf{H}_{k+1}$ by using measurements at σ_i, $i \in \{1, \ldots, 2k+2\}$. In the SISO case, the norm between the transfer function measurements just becomes the absolute value.

We repeat the procedure until:

$$\max_{s \in j\omega} \|\mathbf{H}_k(s) - \mathbf{H}_{k-1}(s)\| \leq \texttt{tol}. \tag{12}$$

We sketch the whole procedure in Algorithm 1. When the condition (12) is met, the algorithm converges as we have obtained the minimal realization of the system and no new information can be added to model.

The algorithm is not sensitive to small changes in the input hyperparameters. For example, small changes of β in the mask function will not have a big impact on the convergence of the algorithm as long as the mask function keeps the properties described above. This is shown in the numerical section. The other hyperparameter $\texttt{tol}$ needs to be set by the user as a trade off between accuracy of the resulting model and the number of interpolation points used.

4 Extension to Time Domain Data

In the previous section, we have discussed a greedy approach for frequency domain measurements. Here, we suppose that we have access to the time-domain data instead of the frequency-domain data.

In this case, we need to design an input to collect time domain data so that we extract the maximum information about the process. In this case, we can also employ Algorithm 1 with a slight modification. First, note that the authors in [19] have proposed a methodology to realize a state-space model using time-domain data, where frequency-domain measurements are estimated using time-domain data. This is followed by obtaining a realization using the Loewner framework described in Section 2. However, the methodology heavily depends on the choice of input, and the choice should be made in such a way that it allows us to describe the dynamics of the system completely. Hence, in this section, we discuss a suitable choice of inputs. Using an extension of Algorithm 1, one can adaptively choose frequency points composing an input. In practice, time-domain measurements are typically collected at a regular interval; hence, the measurements are discrete. Therefore, the proposed greedy procedure needs to be adapted in a discrete setting. In what follows, we discuss the adaptation that needs to be made to allow us to design an input to extract as much information as possible. Note that the interpolation points are no longer on the $s-$plane but are rather on the unit circle of the $z-$plane. Thus, the frequency range on the $\jmath\omega$-axis needs to be mapped on the unit circle using an appropriate discretization method. In order to obtain time-domain measurements, the system is excited with an input spanning a set of interpolation points and the system response is measured. We design an input $\mathbf{u}_p^{(k)}$ using a sum of sine and cosine functions, i.e., the input in the kth step of the algorithm at time $\mathbf{T}_s p$, where $\mathbf{T}_s$ is a sampling time and p is a non-negative integer, can be given as:

$$\mathbf{u}_p^{(k)} = \frac{1}{K} \sum_{l=1}^{2} (1 + \jmath) \left(\cos\left(\sigma_{2k+l}\right) + \jmath \sin\left(\sigma_{2k+l}\right) \right), \tag{13}$$

where $\jmath := \sqrt{-1}$, $p \in \{0, \ldots, K-1\}$, K is the number of time-domain measurements.

Following the same strategy as in Section 3, the goal is to estimate frequency-domain data, construct a realization using the interpolation points and corresponding transfer function estimates. However, since only time domain data are available, one has first to use this data to estimate the frequency measurements.

The complete procedure is as follow. Given an initial realization $\mathbf{H}_{\text{init}}(s)$, the next two interpolation points σ_{2k+1} and σ_{2k+2} are computed as in (10) and (11). Then, we construct the input $\mathbf{u}_p^{(k)}$ (at step k of the algorithm and time step p) as in (13) with the two interpolation points σ_{2k+1} and σ_{2k+2}. We simulate the system to obtain an output having a value of $\mathbf{y}_p^{(k)}$ at the step k of the algorithm and at time step p. Assembling this time-domain data, one can then compute estimates for the transfer function $\mathbf{H}_{\text{true}}(\sigma_{2k+1})$ and $\mathbf{H}_{\text{true}}(\sigma_{2k+2})$. These values are computed using a least-squares problem [19] as follows:

$$\underset{\widehat{\mathbf{H}} \in \mathbb{C}^r}{\arg\min} \left\| \mathbf{F}\widehat{\mathbf{H}} - \bar{\mathbf{y}} \right\|_2^2, \tag{14}$$

where $\mathbf{F} \in \mathbb{C}^{(\mathbf{K}-k_{\min}) \times 2}$ and $\widehat{\mathbf{H}}$ are as follows :

$$\mathbf{F} = \begin{bmatrix} \mathbf{U}_1 e^{\jmath\sigma_{2k+1}k_{\min}} & \mathbf{U}_2 e^{\jmath\sigma_{2k+2}k_{\min}} \\ \vdots & \vdots \\ \mathbf{U}_1 e^{\jmath\sigma_{2k+1}(K-1)} & \mathbf{U}_2 e^{\jmath\sigma_{2k+2}(K-1)} \end{bmatrix}, \quad \text{and} \quad \widehat{\mathbf{H}} = \begin{bmatrix} \widetilde{\mathbf{H}}(\sigma_{2k+1}) \\ \widetilde{\mathbf{H}}(\sigma_{2k+2}) \end{bmatrix}, \tag{15}$$

where $\widetilde{\mathbf{H}}(\sigma)$ is an estimate of $\mathbf{H}_{\text{true}}(\sigma)$, and the output vector is defined as $\bar{\mathbf{y}} := \left[\mathbf{y}_{k_{\min}}^{(k)}, \ldots, \mathbf{y}_{K-1}^{(k)} \right]^\top$ such that $\mathbf{y}_p^{(k)}$ is the measurement of the output at time step p and the step k of the algorithm. $\mathbf{U}_1$ and $\mathbf{U}_2$ are the nonzero (discrete) Fourier transform components of the input $\mathbf{u}$, given in (13) which will be, in fact, at frequencies σ_{2k+1} and σ_{2k+2}, respectively. Moreover, $k_{\min}$ is chosen such that the system reaches a steady state (approximately) after $k_{\min}$ time steps. For a proof of the derivation of the least-squares problem (14), we refer to the discussion in [19].

The next steps are then similar to the steps of Algorithm 1. A realization is constructed using all the estimated points at $\Sigma_k \cup \{\sigma_{2k+1}, \sigma_{2k+2}\}$. The new interpolation points are then computed as in

(10) and (11). The whole procedure is repeated until the error between two iterations is small enough, i.e.,

$$\max_{\|z\|=1} \|\mathbf{H}_k(z) - \mathbf{H}_{k-1}(z)\| \leq \texttt{tol}.$$

Remark 1 *Although the realization is constructed using all collected interpolation points, the input is constructed at each step using only two interpolation points. It results in a considerable reduction in the computation time of the least-squares problem* (14) *compared to the case where all interpolation points have to be inferred at once.*

5 Numerical experiments

In this section, we illustrate the proposed methodology by means of three examples. The first two examples, namely, the penzl [20] and the beam [21] examples, consider frequency-domain measurements. In the last example, time-domain data is used to construct a realization for an RLC circuit [22] with an adaptive choice of measurements. We compare our proposed approach, where we make a careful choice of measurement points, with the Loewner approach, where logarithmically equidistant measurement points are considered. In both cases we choose the best splitting method of right and left interpolation points as discussed in [23]. In the following, we also note some details. Typically, we are interested in a frequency range $[\omega_{min}, \omega_{max}]$, and we consider M points, on a logarithmic scale, from that given range, denoted by $\mathbf{Q} := [q_1, \ldots, q_M]$. Whenever a measurement point is taken, we assume that it is taken for the set $\mathbf{Q}$. Moreover, the Loewner framework [1] is utilized to construct a realization using measurement points and transfer function evaluations. The initial realization is constructed using 6 approximately logarithmically equidistant interpolation points taken from the set $\mathbf{Q}$. This is done by computing first 6 logarithmically equidistant interpolation points $[\widehat{q}_1, \ldots, \widehat{q}_6]$ in the frequency range $[\omega_{min}, \omega_{max}]$ and then find the interpolation points $[\tilde{q}_1, \ldots, \tilde{q}_6]$ that are used in the algorithm by solving the following optimization problem:

$$\arg\min_{\tilde{q}_i \in \mathbf{Q}} \|\tilde{q}_i - \widehat{q}_i\|. \tag{16}$$

In each step of the algorithm, the interpolation points are organized into an interlacing manner []. The filter parameters in (9) are set to $\beta = 0.6$ and $\epsilon = 10^{-15}$. The tolerance in Algorithm 1 is set to $\texttt{tol} = 10^{-8}$. All the numerical experiments are done on an AMD Ryzen 7 PRO 4750U processor CPU@1.7GHz, up to 8MB cache, 16 GB RAM, Ubuntu 20.04 LTS, MATLAB Version 9.8.0.1323502(R2020a) 64-bit(glnxa64).

5.1 Penzl example

As a first example, we consider the penzl example [20] of order $N = 1\,006$, which is referred to as the FOM example in the `SLICOT` model reduction benchmark [21]. We are interested in the frequency range $[10^{-1}, 10^3]$. Using Algorithm 1, we obtain a realization that on termination uses 24 interpolation points. For comparison, we also construct a realization using 24 logarithmically equidistant interpolation points in the given frequency range. We compare these two learned realizations in Figure 2, where we observe that our greedy scheme to collect measurements clearly outperforms the scheme when measurements are taken equidistantly. Algorithm 1 focuses on the region where the transfer function is involved, and it automatically trends to add more points around the peaks of the transfer function. On the other hand, the realization, constructed using the logarithmically equidistant points, has a larger error as it fails to capture these peaks accurately.

Furthermore, in Figure 3, we compare the frequency-limited $\mathcal{H}_2$-norm [24] of the learned and ground-truth systems in terms of the number of interpolation points used. We use the `MORLAB` software package for the computation of the frequency-limited norm [25]. The figure shows that the frequency-limited $\mathcal{H}_2$ error decays faster when the measurements are collected adaptively (Algorithm 1). Finally, we
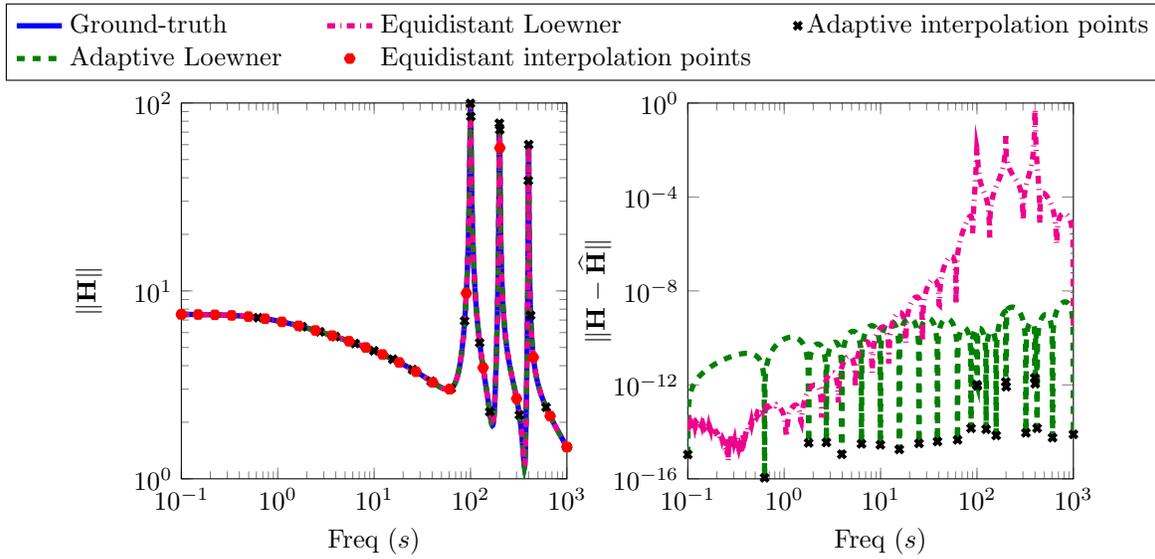
Figure 2: Penzl example: The Bode plot of the ground-truth, adaptively generated system and a realization with equidistant points and the corresponding error between them.

Method	Standard deviation of noise (σ)			
	0	10^{-6}	10^{-5}	10^{-4}
Adaptive	$3.287 \cdot 10^{-9}$	$1.544 \cdot 10^{-7}$	$1.788 \cdot 10^{-5}$	$5.886 \cdot 10^{-4}$
Equidistant	$7.737 \cdot 10^{-6}$	$2.400 \cdot 10^{-3}$	$1.400 \cdot 10^{-2}$	$3.259 \cdot 10^{-1}$

Table 1: Penzl example: A comparison of the frequency-limited $\mathcal{H}_2$-norm of the error between the ground-truth and realized systems under different levels of noise in the measurement data.

study the robustness of Algorithm 1 with respect to the hyper-parameter β. We report the quality of the learned realization with respect to the parameter β in Figure 4. We note that Algorithm 1 is quite robust to the parameter, and all learned realizations outperform the one obtained using the logarithmically equidistant measurement points.

Measurement noise

Many studies have been made about the robustness of the Loewner framework to noise, see, e.g., [26–28]. We do a preliminary study and observe the performance of our approach under noisy measurements. We conduct three experiments, where we corrupt the transfer function measurements with Gaussian white noise of standard deviation $\sigma = [10^{-4}, 10^{-5}, 10^{-6}]$, respectively. We compare our approach with a realization based on logarithmically equidistant interpolation points and measurements corrupted with the same noise level as for the adaptive case. The $\mathcal{H}_2$-norms of learned and the ground-truth models are compared in Table 1. The table shows that even though the $\mathcal{H}_2$ error for the adaptive algorithm increases with the level of noise, it yields much better models as compared to the one obtained using the logarithmically equidistant points. This shows that the adaptive method is superior even in the presence of noise in the measurements. It is important to note that the Loewner framework was used to illustrate the application of our method. In the noisy data case, there maybe better methods where our method can be used.
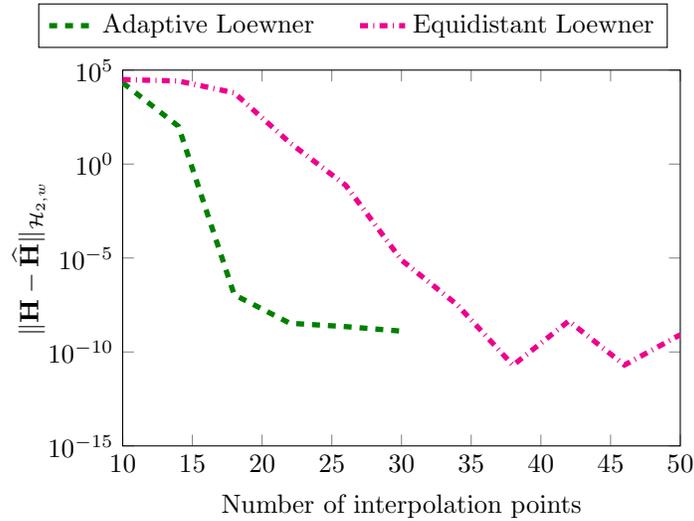
Figure 3: Penzl example: A comparison between the frequency-limited $\mathcal{H}_2$-norm error for the adaptively chosen points and the equidistant points.

5.2 Beam example

In this example, we consider the beam model from the SLICOT model reduction library [21]. It comes from the discretization of a PDE [29]. The input is the force applied at one end, and the output is the displacement resulting from the applied force at another end. We consider the frequency range of operation $[10^{-1}, 5]$. Algorithm 1 is applied, and a realization is obtained using 30 interpolation points as shown in Figure 5. As in the first example, we also construct a realization using 30 logarithmically equidistant logarithmic scale interpolation points distributed over the chosen range of frequencies. Here again, we notice that most of the interpolation points using the adaptive method are chosen around the areas where most changes are happening. In contrast, measurements taken at logarithmically equidistant points yield a realization having a larger error. In Figure 6, we observe that adaptively chosen interpolation points decrease the $\mathcal{H}_2$-norm error faster as compared to the equidistant ones.

5.3 RLC circuit

Finally, we consider an RLC circuit with 100 resistors, capacitors and inductors [22, 30]. The frequencies are taken within the range of frequencies $[10^{-2}, 10^3]$. In order to obtain input/output time-domain measurements, a system simulation using the input (13) is performed to obtain the output in each step. Applying the procedure described in Section 4 directly realizes a model of order 12. As in the first example, we consider the realization of a model using 12 logarithmically equidistant interpolation points and their complex conjugates on the unit circle in order to ensure that the resulting system is real, as discussed in [19]. A comparison of the Bode plots is shown in Figure 7 and a comparison between the output response of the ground-truth system and its realization is shown in Figure 8 using the input

$$\mathbf{u}(t) = \sin(2t) + \sin(20t).$$

Our algorithm recovers the model with a relatively low error, and the model captures the time-domain response of the ground-truth model. It shows that our method provides a promising direction for the long-standing subject of input choice for time-domain system identification.
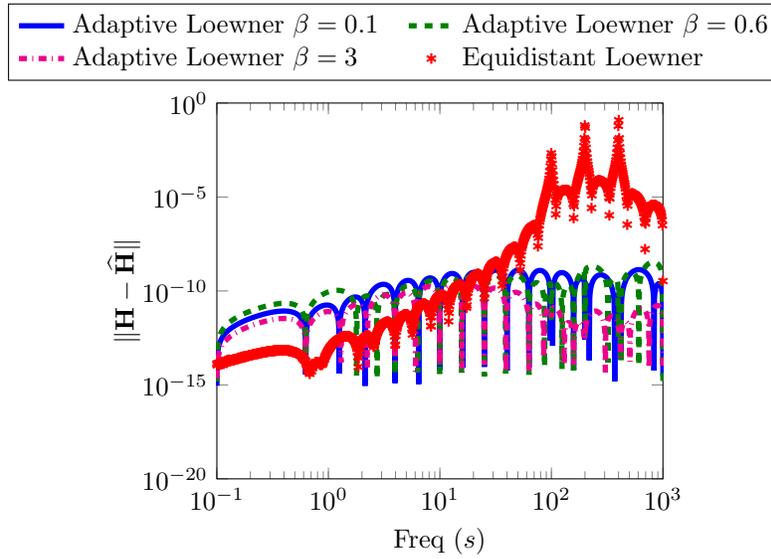
Figure 4: Penzl example: A comparison between the resulting bode plot for different values of β used to defined the filter $\mathbf{g}(\cdot)$ in (9).
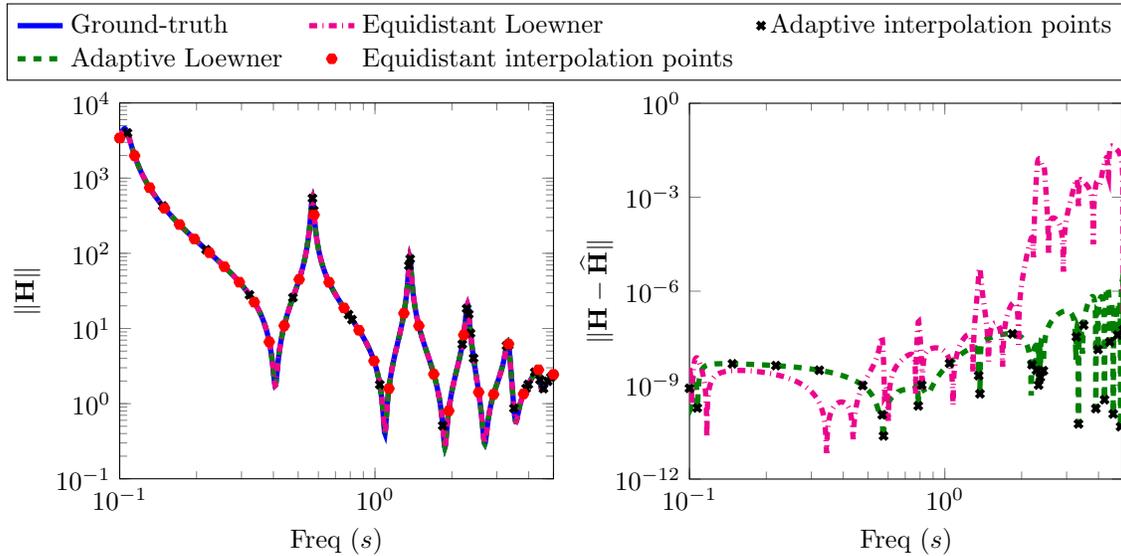

Figure 5: Beam Example: The Bode plot of the ground-truth system, adaptively generated system, and a realized system with equidistant points. The right figure shows the corresponding error between the ground-truth and realized systems.

6 Conclusion

In this paper, we have presented a purely data-driven realization method that greedily chooses measurement points to extract maximum information about the dynamics of a system. We have discussed both cases where data are taken in frequency and time-domain. We have illustrated the efficiency of the proposed methodology by means of three benchmark examples. It was shown that our method outperforms the case where the measurement data are taken at logarithmically equidistant points in
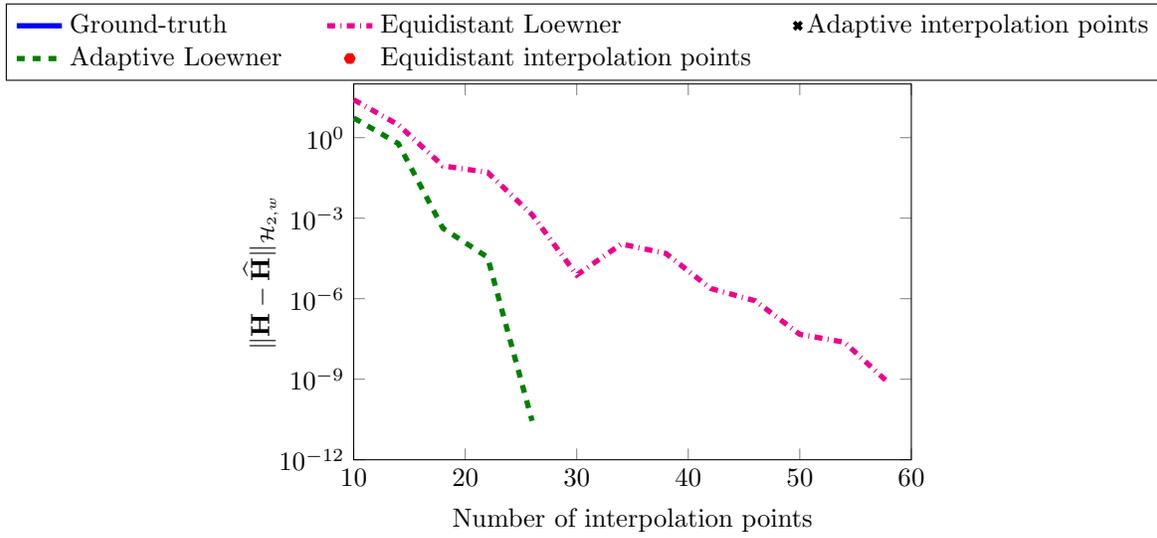
Figure 6: Beam Example: A comparison between the frequency-limited $\mathcal{H}_2$ norm error for the adaptively chosen points and the equidistant points.
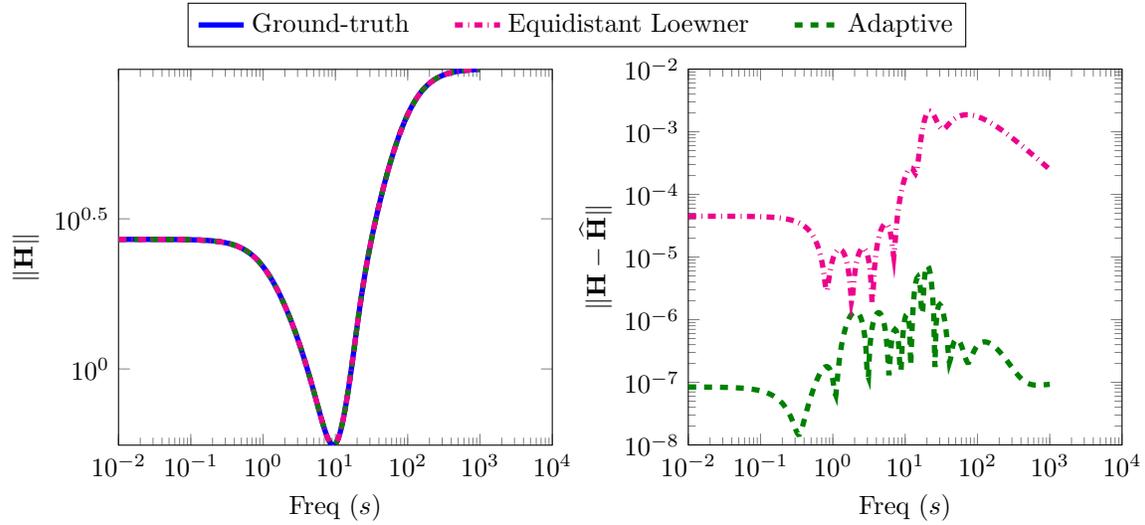

Figure 7: RLC example: A comparison for the Bode plot of the ground-truth and the identified models.

a given frequency range.

References

[1] A. J. Mayo and A. C. Antoulas. A framework for the solution of the generalized realization problem. *Linear Algebra Appl.*, 425(2-3):634–662, 2007.

[2] Z. Drmač, S. Gugercin, and C. Beattie. Quadrature-based vector fitting for discretized $\mathcal{H}_2$ approximation. *SIAM J. Sci. Comput.*, 37(2):A625–A652, 2015.
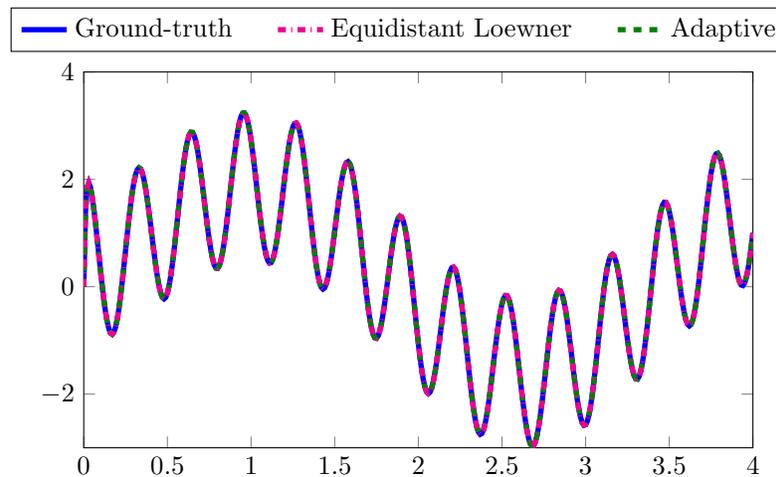
Figure 8: RLC example: A comparison between the output response of the ground-truth and the identified models.

[3] Z. Drmač, S. Gugercin, and C. Beattie. Vector fitting for matrix-valued rational approximation. *SIAM J. Sci. Comput.*, 37(5):A2346–A2379, 2015.

[4] B. Peherstorfer and K. Willcox. Data-driven operator inference for non-intrusive projection-based model reduction. *Comp. Meth. Appl. Mech. Eng.*, 306:196–215, 2016.

[5] A. C. Antoulas, S. Lefteriu, and A.C. Ionita. A tutorial introduction to the Loewner framework for model reduction. In P. Benner, A. Cohen, M. Ohlberger, and K. Willcox, editors, *Model Reduction and Approximation: Theory and Algorithms*, pages 335–376. SIAM, 2017.

[6] Y. Nakatsukasa, O. Sète, and L. N. Trefethen. The AAA algorithm for rational approximation. *SIAM J. Sci. Comput.*, 40:A1494–A1522, 2018.

[7] J. Yu, C. Yan, and M. Guo. Non-intrusive reduced-order modeling for fluid problems: A brief review. *Proc. the Institution of Mechanical Engineers, Part G: J. Aerospace Eng.*, 233(16):5896–5912, 2019.

[8] S. L. Brunton and J. N. Kutz. *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control.* Cambridge University Press, 2019.

[9] S. Gugercin, A. C. Antoulas, and C. Beattie. $\mathcal{H}_2$ model reduction for large-scale linear dynamical systems. *SIAM J. Matrix Anal. Appl.*, 30(2):609–638, 2008.

[10] A. Paul-Dubois-Taine and D. Amsallem. An adaptive and efficient greedy procedure for the optimal training of parametric reduced-order models. *Internat. J. Numer. Methods Engrg.*, 102(5):1262–1292, 2015.

[11] S. Chellappa, L. Feng, and P. Benner. Adaptive basis construction and improved error estimation for parametric nonlinear dynamical systems. *Internat. J. Numer. Methods Engrg.*, 121:5320– 5349, 2020.

[12] S. Chellappa, L. Feng, V. de la Rubia, and P. Benner. Adaptive interpolatory MOR by learning the error estimator in the parameter domain. e-prints 2003.02569, arXiv, 2020.

[13] C. A. Beattie and S. Gugercin. Realization-independent $\mathcal{H}_2$-approximation. In *51st IEEE Conference on Decision and Control (CDC)*, pages 4953–4958, 2012.

[14] A.C. Ionita and A.C. Antoulas. Case study: Parametrized reduction using reduced-basis and the loewner framework. In Quarteroni A. and Rozza G., editors, *Reduced Order Methods for Modeling and Computational Reduction*, volume 9, pages 51–66. Springer-Verlag, 2014.

[15] A. C. Ionita and A. C. Antoulas. Data-driven parametrized model reduction in the Loewner framework. *SIAM J. Sci. Comput.*, 36(3):A984–A1007, 2014.

[16] C. Poussot-Vassal, P. Kergus, F. Kerhervé, D. Sipp, and L. Cordier. Interpolatory-based data-driven pulsed fluidic actuator control design and experimental validation. e-prints 2012.01061, arXiv, 2020.

[17] R. S. Beddig, P. Benner, I. Dorschky, T. Reis, P. Schwerdtner, M. Voigt, and S. W. R. Werner. Structure-preserving model reduction for dissipative mechanical systems. e-prints 2010.06331, arXiv, 2020.

[18] K. Ogata. *Modern Control Engineering*. Prentice Hall, 2010.

[19] B. Peherstorfer, S. Gugercin, and K. Willcox. Data-driven reduced model construction with time-domain Loewner models. *SIAM J. Sci. Comput.*, 39:A2152–A2178, 2017.

[20] T. Penzl. Algorithms for model reduction of large dynamical systems. *Linear Algebra Appl.*, 415:322–343, 2006.

[21] Y. Chahlaoui and P. Van Dooren. A collection of benchmark examples for model reduction of linear time invariant dynamical systems. Technical Report 2002–2, SLICOT Working Note, 2002. Available from www.slicot.org.

[22] S. Gugercin and A. C. Antoulas. A survey of balancing methods for model reduction. In *Proc. European Control Conf. ECC 2003, Cambridge, UK*, 2003. CD Rom.

[23] D. S. Karachalios, I. V. Gosea, and A. C. Antoulas. The Loewner framework for system identification and reduction. In P. Benner, S. Grivet-Talocia, A. Quarteroni, G. Rozza, W. H. A. Schilders, and L. M. Silveira, editors, *Methods and Algorithms*, volume 1 of *Handbook on Model Reduction*. De Gruyter, 2020.

[24] D. A. Wilson. Optimum solution of model-reduction problem. *Proceedings of the Institution of Electrical Engineers*, 117(6):1161–1165, 1970.

[25] P. Benner and S. W. R. Werner. Limited balanced truncation for large-scale sparse second-order systems (version 2.0), 2020. Available from http://doi.org/10.5281/zenodo.3331592.

[26] S. Lefteriu, A. C. Ionita, and A. C. Antoulas. Modeling systems based on noisy frequency and time domain measurements. In Willems J.C., Hara S., Ohta Y., and Fujioka H., editors, *Perspectives in Mathematical System Theory, Control, and Signal Processing*, volume 398. Springer, Berlin, Heidelberg, 2010.

[27] Z. Drmač and B. Peherstorfer. Learning low-dimensional dynamical-system models from noisy frequency-response data with Loewner rational interpolation. e-prints 1910.00110, arXiv, 2019.

[28] P. Kergus, S. Formentin, C. Poussot-Vassal, and F. Demourant. Data-driven control design in the Loewner framework: Dealing with stability and noise. In *European Control Conference (ECC)*, pages 1704–1709. IEEE, 2018.

[29] A. C. Antoulas, D. C. Sorensen, and S. Gugercin. A survey of model reduction methods for large-scale systems. *Contemp. Math.*, 280:193–219, 2001.

[30] P. Benner, P. Goyal, and P. Van Dooren. Identification of port-Hamiltonian systems from frequency response data. *Syst. Control Lett.*, 143:104741, 2020.