2023 62nd IEEE Conference on Decision and Control (CDC)
December 13-15, 2023. Marina Bay Sands, Singapore

Kernel-based continuous-time system identification: A parametric approximation

Matteo Scandella, Alessio Moreschini, and Thomas Parisini

Abstract— **In this paper, we discuss the non-parametric estimate problem using kernel-based LTI system identification techniques by constructing a Loewner-based interpolant of the estimated model. Through this framework, we have been able to retrieve a finite-dimensional approximation of the infinite-dimensional estimate obtained using the classical kernel-based methodologies. The employment of the Loewner framework constitutes an enhancement of recent results which propose to use a Padé approximant to obtain a rational transfer function from an irrational transfer function corresponding to the identified impulse response. The enhancement has been illustrated for the identification of the Rao-Garnier benchmark.**

Index Terms— **System Identification, Kernel-based learning, Loewner framework, Continuous-time system identification, LTI system identification,**

I. Introduction

A. Motivations and state of the art

Modeling dynamical systems from data leans naturally toward the identification of discrete-time models due to how data are naturally collected. For this reason, the research effort in system identification is mainly focused on discrete-time modeling [1], [2], [3]. However, in many contexts, the use of continuous-time models is more appropriate because, in the physical world, dynamical systems are usually described in the continuous-time domain [4]. Thus, system identification for continuous-time models is a consolidated research area with many results [5], [6]. Continuous-time models are commonly identified relying on a two steps procedure that first uses the data to obtain a discrete-time model and then the identified model is converted into a continuous-time one [7], [8], [9]. However, these two-step approaches have multiple drawbacks as explained in detail in [4], [10], [11]. Firstly, they cannot handle non-uniformly sampled data, or time-delay that are not multiple of the sampling time. Furthermore, they struggle with stiff systems. For these reasons, several algorithms or optimization schemes that directly identify a continuous-time model were developed in recent years [6], [10], [12], [13], [14], [15], [16], and multiple Matlab toolboxes are available [17], [18].

Most of the aforecited methods are parametric and rely on the prior knowledge of the system order. When this information is unavailable, these methods are equipped with criteria that assess the more suitable order given the data. Among others, it is possible to use the Akaike Information Criterion [19], or the Young heuristic Information Criterion [20]. On the other hand, these methods select the model among a discrete set of possible structures. For this reason, recently, non-parametric techniques that employ regularization to tune continuously the model complexity were developed. These approaches, commonly known as kernel methods or regularization networks, were originally introduced for the identification of Linear Time-Invariant (LTI) systems in [21]. Since then, they gained popularity for their ability to easily embed specific desired properties into the identified model [22], [23], [24] such as, among others, Bounded Input Bounded Output (BIBO) stability, causality, or smoothness. Additionally, they can be easily equipped with various methods, such as Empirical Bayes [25], for automatically tuning the complexity of the identified model. For continuous-time models, unfortunately, these techniques provide identified models that are difficult to employ in practical applications because the identified model is defined by its impulse response.

Recently, [14] proposed a solution to this problem that directly identifies an approximated transfer function of the identified model. In this approach, the authors derive the irrational transfer function corresponding to the identified impulse response, and they propose to use a Padé approximant [26] to obtain a rational transfer function. This approach suffers from two drawbacks: **(D1)** the Padé approximation is designed to represent the irrational function only around a specific point in the domain, and **(D2)** it is not guaranteed that the approximation preserves the stability properties that the kernel methods guarantee.

B. Contributions

To cope with issues **(D1)** and **(D2)** we employ the Loewner-based interpolation method to construct a (finite-dimensional) rational approximant of the (infinite-dimensional) irrational estimated transfer function obtained in [14], [27], [28]. The Loewner framework [29], [30], [31], [32], originally developed for rational interpolation,

This work has been partially supported by the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No. 739551 (KIOS CoE); by the Italian Ministry for Research in the framework of the 2017 Program for Research Projects of National Interest (PRIN), Grant No. 2017YKXYXJ; and by the EPSRC grant "Model Reduction from Data", Grant No. EP/W005557.

M. Scandella, A. Moreschini are with the Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, UK (e-mail: m.scandella@imperial.ac.uk, a.moreschini@imperial.ac.uk)

T. Parisini is with the Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, UK, with the Department of Engineering and Architecture, University of Trieste, 34127 Trieste, Italy, and also with the KIOS Research and Innovation Center of Excellence, University of Cyprus, CY-1678 Nicosia, Cyprus (e-mail: t.parisini@imperial.ac.uk)

is a model reduction technique yielding a reduced order model that interpolates the transfer function of the high-order model at arbitrarily selected frequencies and directions. In this respect, since the reduced model is solely obtained by interpolating a set of points in multiple parts of the complex plane, the Loewner framework can be employed for solving **(D1)**. Additionally, the Loewner framework is well studied in the literature and various techniques that guarantee the stability of the reduced model are available [33], [32]. For this reason, it is also possible to address **(D2)**.

C. Organization

The paper is organized as follows. Section II briefly recalls the kernel-based methodology for the identification of LTI models. Then, Section III illustrates how the Loewner framework is employed to select a finite-dimensional approximation of the identified model. The proposed methodology is validated through numerical simulations in Section IV. Finally, V concludes the manuscript with some concluding remarks.

D. Notation

We denote by $\mathbb{C}$, $\mathbb{R}$ and $\mathbb{N}$ the set of complex, real and natural numbers, respectively ($0 \in \mathbb{N}$). ι is the imaginary unit. Given a complex number $z \in \mathbb{C}$, $\mathrm{Re}(z)$, $\mathrm{Im}(z)$, and z^* are the real part, the imaginary part, and the complex conjugate of z, respectively. Given $n \in \mathbb{N}$, $\mathbb{R}^n$ and $\mathbb{C}^n$ are the set of real and complex column vectors of dimension n, respectively. Given $n, m \in \mathbb{N}$, $\mathbb{R}^{n\times m}$ and $\mathbb{C}^{n\times m}$ are the set of real and complex matrices of dimension $n \times m$, respectively. With a slight abuse of notation, we use tuples of numbers and column vectors interchangeably. Given $n \in \mathbb{N}$, $I_n \in \mathbb{R}^{n\times n}$ is the identity matrix. Given a function $x : [0, \infty) \to \mathbb{R}$, $\mathcal{L}[x] : \mathbb{C} \to \mathbb{C}$ denotes the Laplace transform of x.

II. The system identification technique

Consider the continuous-time causal Single-Input Single-Output (SISO) LTI system whose output $y : [0, \infty) \to \mathbb{R}$ corresponding to the input $u : [0, \infty) \to \mathbb{R}$ is

$$y(t) = [g \star u](t) \coloneqq \int_0^\infty g(\psi)u(t-\psi)\,d\psi\,,$$

where $\star$ denotes the convolution operator and $g : [0, \infty) \to \mathbb{R}$ is the impulse response of the system. Suppose to have collected the dataset $\mathcal{D} = \{(t_i, y_i)\}_{i=1}^n \subseteq \mathbb{R}^2$ from an experiment on the system. In particular, we suppose that, for all $i \in \{1, \dots, n\}$, $y_i = [g\star u](t_i)+e_i$ where $u : [0, \infty) \to \mathbb{R}$ is a known input excitation signal and e is a Gaussian white noise. In particular, given $i, j \in \{1, \dots, n\}$ with $i \neq j$, e_i and e_j are independent and $e_i \sim \mathcal{N}(0, \beta^2)$ where $\mathcal{N}(\mu, \sigma^2)$ is the normal distribution with mean $\mu \in \mathbb{R}$ and variance $\sigma^2 \in (0, \infty)$. Since the data are collected from a causal system, for the remainder of the manuscript, we assume that, for every $i \in \{0, \dots, n\}$, $t_i > 0$.

The identification procedure for continuous-time LTI models is explained in detail in [23]. In particular, the impulse response of the identified model is obtained by solving the optimization problem

$$\underset{\hat{g}\in\mathcal{H}(k)}{\operatorname{argmin}} \quad \sum_{i=1}^{n} \big(y_i - [\hat{g} \star u](t_i)\big)^2 + \tau\big|\hat{g}\big|^2_{\mathcal{H}(k)}\,, \qquad (1)$$

where $\hat{g}$ is the impulse response of the identified model, $k : [0, \infty) \times [0, \infty) \to \mathbb{R}$ is a valid kernel function, *i.e.* it is a symmetric and positive semi-definite function, $\mathcal{H}(k)$ is the Reproducing Kernel Hilbert Space (RKHS) [34] with reproducing kernel k, $|\cdot|_{\mathcal{H}(k)}$ is its norm, and $\tau \in [0, \infty)$ is the parameter that tunes the complexity of the model. In particular, increasing τ decreases the complexity of the estimated model. Using the Representer Theorem [35], the solution of the optimization problem (1) is given by

$$\hat{g}(t) = \sum_{i=1}^{n} c_i \int_0^\infty u(t_i - \psi)k(t, \psi)\,d\psi, \quad \forall t \in [0, \infty) \quad (2)$$

where $c = (c_i)_{i=1}^n \in \mathbb{R}^n$ is the solution of the linear system

$$O(O + \tau I_n)c = O\bar{y}$$

with $\bar{y} = (y_i)_{i=1}^n \in \mathbb{R}^n$ and $O \in \mathbb{R}^{n\times n}$ being a symmetric positive semi-definite matrix whose (i, j)-th element is $o^u(t_i, t_j)$ that reads as

$$o^u(t_i, t_j) \coloneqq \int_0^\infty \int_0^\infty u(t_i-\xi)u(t_j-\psi)k(\xi, \psi)\,d\xi\,d\psi. \quad (3)$$

The properties of the estimated function $\hat{g}$ strongly depend on the selection of the kernel function k. Additionally, the Representer Theorem, given any bounded input, is guaranteed to hold only for absolutely integrable kernels [36]. In this early work, we focus exclusively on the stable spline kernel introduced in [21] because **(i)** it is widely used in the literature, *i.a.* [37], [38], [39], **(ii)** it is a stable kernel [23], [40], *i.e.* it guarantees that the estimated model is BIBO and the Representer Theorem is valid, **(iii)** they are strongly related to other useful stable kernels such as the DC and TC kernels [41].

The stable spline kernel can be defined as [14, Prop. 2.1]

$$k(a, b) = \lambda \sum_{h=0}^{q-1} \gamma_{q,h} \begin{cases} e^{-\beta[(2q-h-1)a+hb]} & \text{If } a \geq b \\ e^{-\beta[(2q-h-1)b+ha]} & \text{If } a < b \end{cases}, \quad (4)$$

where $\lambda, \beta \in [0, \infty)$, $q \in \mathbb{N} \setminus \{0\}$ are kernel parameters that need to be tuned, and

$$\gamma_{q,h} = \frac{(-1)^{q+h+1}}{h!(2q-h-1)!}\,.$$

Using kernel (4) and estimate (2), the method requires the tuning of the hyperparameters $\zeta = (\tau, \lambda, \beta, q) \in [0, \infty)^3 \times \mathbb{N} \setminus \{0\}$. This problem is usually addressed by solving an optimization problem capable of dealing with the trade-off between complexity and fitting, therefore enabling the avoidance of over- and under-fitting. For this type of estimation technique, the most popular method is based on the empirical Bayes theory [25], and it translates into the maximization of the likelihood of the available data given a certain ζ. More

precisely, we select ζ as the solution of the optimization problem

$$\underset{\zeta\in[0,\infty)^3\times\mathbb{N}\setminus\{0\}}{\operatorname{argmin}} \quad \bar{y}^\top(O+\tau I_n)^{-1}\bar{y}+\log\det\big(O+\tau I_n\big).$$

The estimated impulse response $\hat{g}$ described in (2) is a linear combination of convolutions between the kernel function k and functions that are defined using the excitation input signal. Handling this function in practical applications can be daunting, and it is usually approximated numerically in a finite set of time instants, *i.e.* by discretizing the model. To avoid this issue, we follow the rationale of [14] where they propose computing a rational approximation of the transfer function $\hat{G}=\mathcal{L}[\hat{g}]$. In particular, they first define the irrational transfer function of the identified model, and they employ it to define a rational approximant. The irrational transfer function of the identified model using kernel (4) is given by [14, Prop. 3.2]. In detail, for each $s\in\mathbb{C}$, we have

$$\hat{G}(s)=\lambda\left[\sum_{h=0}^{q-1}Q^u_{q,h}(s)+H^u_q(s)\right],$$

where

$$Q^u_{q,h}(s)=\frac{\gamma_{q,h}}{s+\beta h}\left(\sum_{i=1}^{n}c_iA^u_i\big(\beta(2q-h-1)\big)\right),$$

$$H^u_q(s)=\frac{(-1)^q\beta^{2q-1}}{\prod_{i=0}^{2q-1}(s+\beta i)}\left(\sum_{i=1}^{n}c_iA^u_i\big(s+\beta(2q-1)\big)\right),$$

and $A^u_i\coloneqq\mathcal{L}[u_i]$ with $u_i(t)=u(t_i-t)$, for every $t\in[0,\infty)$. By analyzing the term $H^u_q(s)$, the transfer function $\hat{G}$ is irrational for many input signals u. In the next Section, we propose a method to find a rational approximation of $\hat{G}$ using the Loewner framework.

III. Approximated finite-dimensional model

As anticipated in the Introduction, the Loewner-based order reduction approach yield interpolants of the transfer function at a finite number of points in its domain. For this reason, we consider $2m\in\mathbb{N}$ interpolation points that, for convenience, are divided into two disjointed sets. Specifically:

$$\{(\lambda_i,r_i,w_i)\}_{i=1}^m\subseteq\mathbb{C}^3,\qquad\{(\mu_i,\ell_i,v_i)\}_{i=1}^m\subseteq\mathbb{C}^3,$$

where, for every $i\in\{1,\dots,m\}$, we have

$$\hat{G}(\lambda_i)r_i=w_i,\qquad\qquad\ell_i\hat{G}(\mu_i)=v_i.$$

Furthermore, we assume that $\{\lambda_i\}_{i=1}^m\cap\{\mu_i\}_{i=1}^m=\varnothing$. The objective is to define an LTI model of order m whose transfer function $\hat{G}_{\mathrm{r}}$ interpolates all the considered interpolation points. Therefore, for every $i\in\{1,\dots,m\}$, we get the following interpolation conditions:

$$\hat{G}_{\mathrm{r}}(\lambda_i)r_i=w_i=\hat{G}(\lambda_i)r_i,\quad\ell_i\hat{G}_{\mathrm{r}}(\mu_i)=v_i=\ell_i\hat{G}(\mu_i).$$

Carrying out the procedure in [30], the descriptor of the reduced model is given by the equations

$$\mathbb{L}\dot{x}(t)=\mathbb{L}_\sigma x(t)-Vu(t),\qquad y(t)=Wx(t),\qquad(5)$$

where $V\coloneqq[v_1,\dots,v_m]^\top\in\mathbb{C}^{m\times1}$, $W\coloneqq[w_1,\dots,w_m]\in\mathbb{C}^{1\times m}$,

$$\mathbb{L}\coloneqq\begin{bmatrix}\frac{v_1r_1-\ell_1w_1}{\mu_1-\lambda_1}&\dots&\frac{v_1r_m-\ell_1w_m}{\mu_1-\lambda_m}\\\vdots&\ddots&\vdots\\\frac{v_mr_1-\ell_mw_1}{\mu_m-\lambda_1}&\dots&\frac{v_mr_m-\ell_mw_m}{\mu_m-\lambda_m}\end{bmatrix},$$

and

$$\mathbb{L}_\sigma\coloneqq\begin{bmatrix}\frac{\mu_1v_1r_1-\ell_1w_1\lambda_1}{\mu_1-\lambda_1}&\dots&\frac{\mu_1v_1r_m-\ell_1w_m\lambda_m}{\mu_1-\lambda_m}\\\vdots&\ddots&\vdots\\\frac{\mu_mv_mr_1-\ell_mw_1\lambda_1}{\mu_m-\lambda_1}&\dots&\frac{\mu_mv_mr_m-\ell_mw_m\lambda_m}{\mu_m-\lambda_m}\end{bmatrix}.$$

Therefore, the rational transfer function of the reduced model is $\hat{G}_{\mathrm{r}}(s)=-W(s\mathbb{L}-\mathbb{L}_\sigma)^{-1}V$, for every $s\in\mathbb{C}$. The matrices $\mathbb{L}\in\mathbb{C}^{n\times n}$ and $\mathbb{L}_\sigma\in\mathbb{C}^{n\times n}$ are commonly referred to as the Loewner Matrix and the shifted Loewner Matrix, respectively. Additionally, it is worth noting that these matrices are, respectively, the solution of the Sylvester equations

$$M\mathbb{L}-\mathbb{L}\Lambda=VR-LW,$$
$$M\mathbb{L}_\sigma-\mathbb{L}_\sigma\Lambda=MVR-LW\Lambda,$$

where $\Lambda\coloneqq\operatorname{diag}(\lambda_1,\dots,\lambda_m)\in\mathbb{C}^{m\times m}$, $M\coloneqq\operatorname{diag}(\mu_1,\dots,\mu_m)\in\mathbb{C}^{m\times m}$, $R\coloneqq[r_1,\dots,r_m]\in\mathbb{C}^{1\times m}$, and $L\coloneqq[\ell_1,\dots,\ell_m]^\top\in\mathbb{C}^{m\times1}$.

Unfortunately, the reduced model (5) cannot be applied *talis qualis* in this context for three reasons:

1) The reduced model takes on complex values *i.e.*, it maps complex functions to complex functions.
2) In general, for arbitrary interpolation points, the Loewner matrix might be nonsingular and thus the reduced model (5) might not be a minimum realization.
3) The reduced model can be an unstable interpolant, and its stability strongly depends on the selected interpolation points.

In the next three sections, problems **1)**-**3)** are tackled by suitably modifying the proposed interpolant.

A. Real reduced model

In this section, we propose to use a shift of coordinates in order to obtain a real reduced model, *i.e.* a model that maps real functions to real functions. In particular, following [30], we consider invertible matrices $J_\lambda,J_\mu\in\mathbb{C}^{m\times m}$ such that matrices

$$\bar{\Lambda}\coloneqq J_\lambda\Lambda J_\lambda^{-1}\qquad\bar{R}\coloneqq RJ_\lambda^{-1}\qquad\bar{W}\coloneqq WJ_\lambda^{-1}$$
$$\bar{M}\coloneqq J_\mu MJ_\mu^{-1}\qquad\bar{L}\coloneqq J_\mu L\qquad\bar{V}\coloneqq J_\mu V$$

are all real matrices. Then, the real Loewner matrix $\bar{\mathbb{L}}\in\mathbb{R}^{m\times m}$ and the real shifted Loewner matrix $\bar{\mathbb{L}}_\sigma\in\mathbb{R}^{m\times m}$ can be defined, respectively, as the solutions of the following Sylvester equations:

$$\bar{M}\bar{\mathbb{L}}-\bar{\mathbb{L}}\bar{\Lambda}=\bar{V}\bar{R}-\bar{L}\bar{W},$$
$$\bar{M}\bar{\mathbb{L}}_\sigma-\bar{\mathbb{L}}_\sigma\bar{\Lambda}=\bar{M}\bar{V}\bar{R}-\bar{L}\bar{W}\bar{\Lambda}.$$

Thus, the real reduced model is given by

$$\bar{\mathbb{L}}\dot{x}(t) = \bar{\mathbb{L}}_\sigma x(t) - \bar{V}u(t), \qquad y(t) = \bar{W}x(t), \tag{6}$$

with transfer function $\bar{G}_{\mathrm{r}}(s) = -\bar{W}(s\bar{\mathbb{L}} - \bar{\mathbb{L}}_\sigma)^{-1}\bar{V}$.

Notice that matrix $\bar{\Lambda}$ has the same spectrum of Λ. Therefore, since the spectrum of Λ is $\{\lambda_i\}_{i=1}^m$, J_λ exists if and only if for every $i \in \{1,\dots,m\}$, there exists $j \in \{1,\dots,m\}$ such that $\lambda_i = \lambda_j^*$. If this condition is satisfied, we can define J_λ as a block diagonal matrix with a 2×2 block, that is

$$\begin{bmatrix} 1 & -\iota \\ 1 & \iota \end{bmatrix} \in \mathbb{C}^{2\times 2},$$

for each complex conjugate pairs in $\{\lambda_i\}_{i=1}^m$ and a 1×1 block, that is I_1, for each real elements of $\{\lambda_i\}_{i=1}^m$. Analogously, J_μ is defined in the same way using the interpolation points in $\{\mu_i\}_{i=1}^m$.

B. Minimum realization of the reduced model

Usually, the sets of interpolation points $\{\lambda_i\}_{i=1}^m$ and $\{\mu_i\}_{i=1}^m$ are selected using the available knowledge of the model to reduce. However, in this work, we aim to find an interpolant for the transfer function $\hat{G}$ that is obtained through a black-box identification procedure. Therefore, we cannot select the interpolation points relying on the prior knowledge of the system. For this reason, it is useful to use as many interpolation points as possible to cover the relevant parts of the domain of $\hat{G}$. However, with the presence of many redundant interpolation points the reduced model can have more states than needed and/or a nonsingular Loewner matrix. In particular, this may occur when the matrix pencil $(\bar{\mathbb{L}}_\sigma, \bar{\mathbb{L}})$ is not regular, *i.e.* if, for every $s \in \{\lambda_i\}_{i=1}^m \cup \{\mu_i\}_{i=1}^m$,

$$\begin{aligned}\operatorname{rank}(s\bar{\mathbb{L}} - \bar{\mathbb{L}}_\sigma) &= \operatorname{rank}\begin{bmatrix} \bar{\mathbb{L}} \\ \bar{\mathbb{L}}_\sigma \end{bmatrix} \\ &= \operatorname{rank}\begin{bmatrix} \bar{\mathbb{L}} & \bar{\mathbb{L}}_\sigma \end{bmatrix} =: r < m.\end{aligned}$$

A procedure that solves this problem was proposed in [29]. Consider the two short Singular Value Decompositions (SVD)

$$\begin{bmatrix} \bar{\mathbb{L}} \\ \bar{\mathbb{L}}_\sigma \end{bmatrix} = U_v S_v Z_v^\top, \qquad \begin{bmatrix} \bar{\mathbb{L}} & \bar{\mathbb{L}}_\sigma \end{bmatrix} = U_h S_h Z_h^\top,$$

where $S_v, S_h \in \mathbb{R}^{r\times r}$, $U_v, Z_h \in \mathbb{R}^{2m\times r}$, and $Z_v, U_h \in \mathbb{R}^{m\times r}$. Then, the minimal realization of the reduced model can be described by the equations

$$U_h^\top \bar{\mathbb{L}} Z_v \; \dot{x}(t) = U_h^\top \bar{\mathbb{L}}_\sigma Z_v \; x(t) - U_h^\top \bar{V} \; u(t), \tag{7a}$$

$$y(t) = \bar{W} Z_v \; x(t). \tag{7b}$$

with transfer function

$$\breve{G}_{\mathrm{r}}(s) = -\bar{W} Z_v \Big(U_h^\top (s\bar{\mathbb{L}}_\sigma - \bar{\mathbb{L}}_\sigma) Z_v \Big)^{-1} U_h^\top \bar{V}.$$

C. Asymptotically stable reduced model

Using kernel (4), the identified model (2) is BIBO stable as shown in [21], [23]. However, it is not guaranteed that the approximated model (7) preserves this property. In the literature, there are various methods to enforce stability on the interpolated model. Here, we consider the method proposed in [33]. Specifically, let $r \in \mathbb{N}$, and define

$$\begin{aligned}\mathcal{S}_r &\coloneqq \bigcup_{i=1}^{r} \left[\mathbb{R}^{i\times i} \times \mathbb{R}^{i\times i} \times \mathbb{R}^{i\times 1} \times \mathbb{R}^{1\times i} \right], \\ \mathcal{S}_r^+ &\coloneqq \{(E,A,B,C) \in \mathcal{S}_r \,:\, \sigma(E,A) \subseteq \mathbb{C}_+\}, \\ \mathcal{S}_r^- &\coloneqq \{(E,A,B,C) \in \mathcal{S}_r \,:\, \sigma(E,A) \subseteq \mathbb{C}_-\},\end{aligned}$$

where $\sigma(E,A)$ is the set of eigenvalues of the matrix pencil (E,A), $\mathbb{C}_+ \coloneqq \{z \in C \,:\, \mathrm{Re}(z) \geq 0\}$ and $\mathbb{C}_- \coloneqq \{z \in C \,:\, \mathrm{Re}(z) < 0\}$. The set $\mathcal{S}_r$ can be interpreted as the set of the SISO models of order equal or smaller than r described using its descriptor form's matrices. Then, $\mathcal{S}_r^-$ is the set of BIBO models and $\mathcal{S}_r^+$ is the set of antistable models. Additionally, given $\Sigma = (E,A,B,C) \in \mathcal{S}_r$ we define the function $H_\Sigma : \mathbb{C} \to \mathbb{C}$ as the transfer function of the model associated to Σ, *i.e.* $H_\Sigma(s) = C(sE - A)^{-1}B$.

Using these definitions, [42] proposes to define the BIBO stable reduced model as the solution of the optimization problem

$$\Sigma_s = \underset{\Sigma \in \mathcal{S}_r^-}{\operatorname{argmin}} \quad \left| H_\Sigma - \breve{G}_{\mathrm{r}} \right|_2 \tag{8}$$

where $|\cdot|_2$ is the L^2 norm and $\breve{G}_{\mathrm{r}}$ is the transfer function of the reduced model defined in (7). To solve this optimization problem, let us recall from [33, Thm. 2] that for every $\Sigma \in \mathcal{S}_r$ there exist $\Sigma_+ \in \mathcal{S}_r^+$ and $\Sigma_- \in \mathcal{S}_r^-$ such that $H_\Sigma = H_{\Sigma_+} + H_{\Sigma_-}$. Then, the models Σ_- and Σ_+ can be constructed using the stable-unstable decomposition. [1] Then, there exist $Q_+ \in \mathcal{S}_+$ and $Q_- \in \mathcal{S}_-$ such that $\breve{G}_{\mathrm{r}} = H_{Q_+} + H_{Q_-}$ where $\breve{G}_{\mathrm{r}}$ is the transfer function of the interpolant defined in (7). Finally, the solution of the optimization problem (8) is $\Sigma_s = Q_-$ with transfer function $G_{\mathrm{r}} = H_{Q_-}$

IV. Simulations

To validate the proposed method numerically, we consider the Rao-Garnier benchmark [10] whose transfer function is

$$R(s) = \frac{-6400s + 1600}{s^4 + 5s^3 + 408s^2 + 416s + 1600}, \quad \forall s \in \mathbb{C}.$$

This benchmark is commonly used for continuous-time LTI system identification because it has some interesting properties. Firstly, the system has two oscillatory modes caused by its two complex conjugate pairs with low damping coefficients located at $2\,\mathrm{rad}/s$ and $20\,\mathrm{rad}/s$. Additionally, the model has a zero in the right half plane.

For the identification, we consider four different experimental settings:

A. Uniformly sampled measurement of the impulse response;

[1] In practice, it can be easily computed by means of various control systems toolbox, for example in MATLAB using the `stabsep` function.

B. Uniformly sampled measurement of the step response;
C. Irregularly sampled measurement of the impulse response;
D. Irregularly sampled measurement of the step response;

In all the experiments, we collect $n = 200$. For **A** and **C** the noise variance is $\sigma^2 = 1.732$, while for **B** and **D** is $\sigma^2 = 0.227$. For **A** and **B** the sampling time is $0.025\,\text{sec}$. Therefore, $t_i = (i-1)0.025\,\text{sec}$. Instead, for **C** and **D** the intersample times are distributed according to a uniform distribution in the interval $[0.0229, 0.0271]$. The identification procedure is carried out on all the experiments using the method proposed in [14], which uses a Padé approximant, and the proposed approach which uses the Loewner-based approach explained in Section III. Finally, the performances are evaluated on a validation dataset $\mathcal{D}_v = \{u_i^v, y_i^v\}_{i=1}^{50000}$ taken from a noiseless experiment on the system using a random Gaussian excitation band limited at $320Hz$ with a sampling time of $3200^{-1}\,\text{sec}$. To compare the identified models, we use the performance index

$$q(\mathcal{D}_v) := 1 - \sqrt{\frac{\sum_{t=1}^{50000} |y_i^v - \hat{y}^v|^2}{\sum_{t=1}^{50000} \left| y_i^v - \frac{1}{50000}\sum_{i=1}^{50000} y_i^v \right|^2}}$$

where $\hat{y}^v$ is the output of the estimated model using the same input of the validation dataset. To better assess the performance of the methods, we employ a Monte Carlo simulation with 125 runs.

For the evaluation of o_u, defined in (3), and of A_i^u, we refer to the equation reported in [14]. The stable spline order is set to $q = 1$. To reduce the computational complexity of the estimated model, we employ the procedure presented in [43] after selecting the hyperparameters. Furthermore, for the Loewner method, we use 1000 interpolation points, 200 reals and 400 complex conjugate pairs. In particular, let us define $\{a_i\}_{i=1}^{200}$ such that $a_1 = 0.2$, $a_{200} = 200$ and $\log(a_{i+1}) - \log(a_i) = \log(a_{j+1}) - \log(a_j)$ for every $i, j \in \{1, 199\}$. Then, the real interpolation points are $\{-a_i\}_{i=1}^{200}$ and the complex conjugate pairs are $\{(+\imath a_i, -\imath a_i)\}_{i=1}^{200}$ and $\{(a_i + \imath a_i, a_i - \imath a_i)\}_{i=1}^{200}$. The 600 interpolation points are split into two disjointed sets $\{\lambda_i\}_{i=1}^{150}$ and $\{\mu_i\}_{i=1}^{150}$ in such a way that the matrices J_λ and J_μ exist, as explained in Section III-A.

Figure 1 reports the box plots of the performance index obtained from the Monte Carlo simulation. Here, we can note that the proposed Loewner approximation outperforms the Padé approximant proposed in [14]. The main reason for this improvement can be noted in Figure 2 where the magnitude plot of the Bode diagram of the estimated models of experiment **A** are compared with the one of the true system. Since the Padé approximant is designed to approximate the irrational transfer function only around 0, the second oscillatory mode is lost during the approximation procedure. Instead, the Loewner-based approach finds an interpolant along a wide range inside the domain of the transfer function. Therefore, it is able to better approximate the effect of the second oscillatory mode of the identified model.

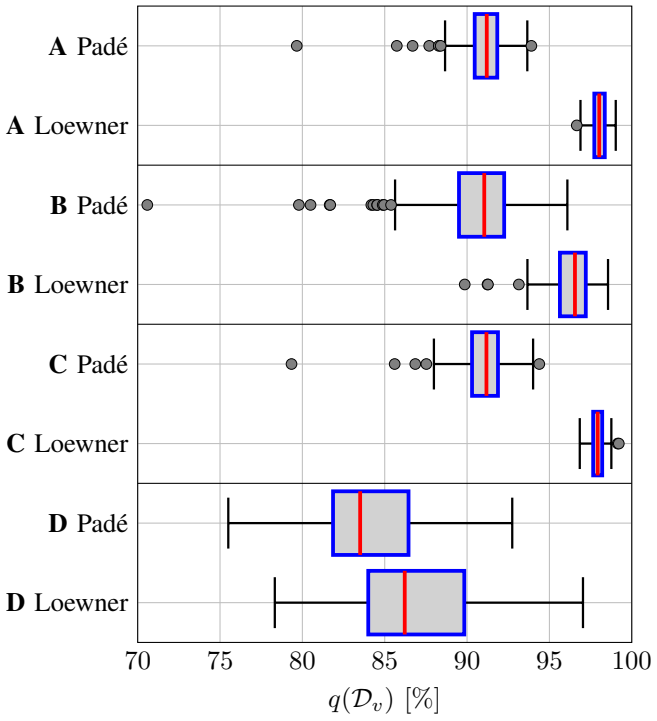

Fig. 1. Box plots of the performance index on the validation dataset in the various experiments.

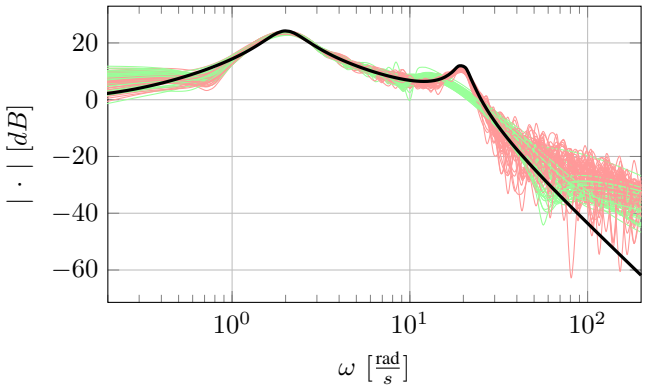

Fig. 2. Comparison between the magnitude of the true system (black line) with the ones identified using experiment **A**. In particular, the green and red lines are obtained using the Padé approximant and Loewner-based approximant, respectively.

V. Conclusions

In practical applications, the non-parametric estimate obtained using kernel-based LTI system identification techniques can be daunting, and it lacks interpretability. For this reason, in this work, we propose a method that tackles this problem by employing a Loewner-based interpolant of the estimated model. The proposed methodology is a refinement of the procedure introduced in [14] that shows significant improvements in the identification of the Rao-Garnier benchmark. With the proposed methodology, it is possible to retrieve a finite-dimensional approximation of the infinite-dimensional estimate obtained using the classical kernel-based methodologies. Additionally, the obtained

model preserves the BIBO stability guarantees of the original model, it is in minimal realization and the model order is automatically selected by the procedure.

In future developments, we aim to expand the identification procedure to more complicated input signals and to optimize the estimation procedure.

REFERENCES

[1] L. Ljung, *System Identification: Theory for the User*. Prentice Hall, 1999.

[2] S. Bittanti, *Model Identification and Data Analysis*. John Wiley and Sons Ltd, Mar. 2019.

[3] M. Milanese and A. Vicino, "Optimal estimation theory for dynamic systems with set membership uncertainty: An overview," *Automatica*, vol. 27, no. 6, pp. 997–1009, Nov. 1991.

[4] H. Garnier and P. C. Young, "What does continuous-time model identification have to offer?" in *Proceedings of the 16th IFAC Symposium on System Identification*, vol. 45, no. 16. Elsevier BV, Jul. 2012, pp. 810–815.

[5] H. Garnier and L. Wang, *Identification of Continuous-time Models from Sampled Data*. Springer London, Mar. 2008.

[6] R. Pintelon, P. Guillaume, Y. Rolain, J. Schoukens, and H. Van Hamme, "Parametric identification of transfer functions in the frequency domain - a survey," *IEEE Transactions on Automatic Control*, vol. 39, no. 11, pp. 2245–2260, Nov. 1994.

[7] R. A. González, C. R. Rojas, and J. S. Welsh, "An asymptotically optimal indirect approach to continuous-time system identification," in *Proceeding of the IEEE Conference on Decision and Control (CDC)*, Dec. 2018, pp. 638–643.

[8] Y. Y. Du, J. S. H. Tsai, H. Patil, L. S. Shieh, and Y. Chen, "Indirect identification of continuous-time delay systems from step responses," *Applied Mathematical Modelling*, vol. 35, no. 2, pp. 594–611, Feb. 2011.

[9] H. D. Unbehauen and G. Rao, "Continuous-time approaches to system identification – A survey," *Automatica*, vol. 26, no. 1, pp. 23–35, Jan. 1990.

[10] H. Garnier, "Direct continuous-time approaches to system identification. Overview and benefits for practical applications," *European Journal of Control*, vol. 24, pp. 50–62, Jul. 2015, sI: ECC15.

[11] G. P. Rao and H. Garnier, "Numerical illustrations of the relevance of direct continuous-time model indentification," in *Proceedings of the 15th IFAC World Congress*, vol. 35, no. 1. Elsevier BV, 2002, pp. 133–138.

[12] F. Chen, H. Garnier, and M. Gilson, "Robust identification of continuous-time models with arbitrary time-delay from irregularly sampled data," *Journal of Process Control*, vol. 25, pp. 19–27, 2015.

[13] F. Chen, J. C. Agüero, M. Gilson, H. Garnier, and T. Liu, "EM-based identification of continuous-time ARMA Models from irregularly sampled data," *Automatica*, vol. 77, pp. 293–301, Mar. 2017.

[14] M. Scandella, M. Mazzoleni, S. Formentin, and F. Previdi, "Kernel-based identification of asymptotically stable continuous-time linear dynamical systems," *International Journal of Control*, vol. 95, no. 6, pp. 1668–1681, 2022.

[15] A. Moreschini, M. Mattioni, S. Monaco, and D. Normand-Cyrot, "A gradient descent algorithm built on approximate discrete gradients," in *2022 26th International Conference on System Theory, Control and Computing (ICSTCC)*, 2022, pp. 343–348.

[16] P. Van Overschee and B. De Moor, "Continuous-time frequency domain subspace system identification," *Signal Processing*, vol. 52, no. 2, pp. 179–194, Jul. 1996, subspace Methods, Part II: System Identification.

[17] H. Garnier and M. Gilson, "CONTSID: a Matlab toolbox for standard and advanced identification of black-box continuous-time models," in *Proceedings of the 18th IFAC Symposium on System Identification SYSID 2018*, vol. 51, no. 15. Elsevier BV, 2018, pp. 688–693.

[18] I. Kollár, R. Pintelon, and J. Schoukens, "Frequency Domain System Identification Toolbox for MATLAB," in *Proceeding of the 9th IFAC/IFORS Symposium on Identification and System Parameter Estimation*, vol. 24, no. 3, Jul. 1991, pp. 1243–1247.

[19] J. E. Cavanaugh and A. A. Neath, "The Akaike information criterion: Background, derivation, properties, application, interpretation, and refinements," *WIREs Computational Statistics*, vol. 11, no. 3, p. 1460, Mar. 2019.

[20] P. C. Young, *Recursive Estimation and Time-Series Analysis*. Springer Berlin Heidelberg, 2011.

[21] G. Pillonetto and G. De Nicolao, "A new kernel-based approach for linear system identification," *Automatica*, vol. 46, no. 1, pp. 81–93, Jan. 2010.

[22] F. Dinuzzo, "Kernels for Linear Time Invariant System Identification," *SIAM Journal on Control and Optimization*, vol. 53, no. 5, pp. 3299–3317, 2015.

[23] G. Pillonetto, F. Dinuzzo, T. Chen, G. De Nicolao, and L. Ljung, "Kernel methods in system identification, machine learning and function estimation: A survey," *Automatica*, vol. 50, no. 3, pp. 657–682, Mar. 2014.

[24] T. Chen, "On kernel design for regularized LTI system identification," *Automatica*, vol. 90, pp. 109–122, 2018.

[25] D. J. C. MacKay, "Bayesian Interpolation," *Neural Computation*, vol. 4, no. 3, pp. 415–447, May 1992.

[26] G. A. Baker and P. Graves-Morris, *Padé approximants*. Cambridge University Press, 1996, vol. 59.

[27] M. Mazzoleni, M. Scandella, S. Formentin, and F. Previdi, "Black-box continuous-time transfer function estimation with stability guarantees: a kernel-based approach," in *Proceedings of the 2nd Conference on Learning for Dynamics and Control*, vol. 120. PMLR, 2020, pp. 267–276.

[28] ——, "Nonparametric continuous-time identification of linear systems: theory, implementation and experimental results," in *Proceedings of the Modeling, Estimation and Control Conference MECC 2021*, vol. 54, no. 20, 2021, pp. 699–704.

[29] A. J. Mayo and A. C. Antoulas, "A framework for the solution of the generalized realization problem," *Linear Algebra and its Applications*, vol. 425, no. 2, pp. 634–662, Sep. 2007, special Issue in honor of Paul Fuhrmann.

[30] A. C. Antoulas, S. Lefteriu, and A. C. Ionita, "A tutorial introduction to the Loewner framework for model reduction," *Model Reduction and Approximation: Theory and Algorithms*, vol. 15, p. 335, 2017.

[31] J. D. Simard, A. Moreschini, and A. Astolfi, "Parameterization of all moment matching interpolants," in *2023 European Control Conference (ECC)*, 2023.

[32] A. Moreschini, J. Simard, and A. Astolfi, "Model reduction for linear port-Hamiltonian systems in the Loewner framework," in *22st IFAC World Congress 2023*, 2023.

[33] I. V. Gosea and A. C. Antoulas, "Stability preserving post-processing methods applied in the Loewner framework," in *Proceeding of the IEEE 20th Workshop on Signal and Power Integrity (SPI)*. IEEE, May 2016, pp. 1–4.

[34] N. Aronszajn, "Theory of reproducing kernels," *Transactions of the American Mathematical Society*, vol. 68, no. 3, pp. 337–404, May 1950.

[35] A. Argyriou and F. Dinuzzo, "A Unifying View of Representer Theorems," in *Proceedings of the 31st International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, vol. 32, no. 2. Bejing, China: PMLR, 22–24 Jun 2014, pp. 748–756.

[36] M. Khosravi and R. S. Smith, "The existence and uniqueness of solutions for kernel-based system identification," *Automatica*, vol. 148, p. 110728, Feb. 2023.

[37] A. Y. Aravkin, J. V. Burke, and G. Pillonetto, "Generalized System Identification with Stable Spline Kernels," *SIAM Journal on Scientific Computing*, vol. 40, no. 5, pp. B1419–B1443, Jan. 2018.

[38] G. Pillonetto, "Identification of hybrid systems using stable spline kernels," in *Proceedings of the IEEE 25th International Workshop on Machine Learning for Signal Processing (MLSP)*, Sep. 2015, pp. 1–6.

[39] L. Blanken and T. Oomen, "Kernel-based identification of non-causal systems with application to inverse model control," *Automatica*, vol. 114, p. 108830, 2020.

[40] T. Chen and G. Pillonetto, "On the stability of reproducing kernel Hilbert spaces of discrete-time impulse responses," *Automatica*, vol. 95, pp. 529–533, Sep. 2018.

[41] T. Chen, "Continuous-Time DC Kernel–A Stable Generalized First-Order Spline Kernel," *IEEE Transactions on Automatic Control*, vol. 63, no. 12, pp. 4442–4447, Dec. 2018.

[42] M. Köhler, "On the closest stable descriptor system in the respective spaces RH_2 and RH_∞," *Linear Algebra and its Applications*, vol. 443, pp. 34–49, Feb. 2014.

[43] M. Scandella, M. Mazzoleni, S. Formentin, and F. Previdi, "A Note on the Numerical Solutions of Kernel-Based Learning Problems," *IEEE Transactions on Automatic Control*, vol. 66, no. 2, pp. 940–947, 2021.