Proceedings of Machine Learning Research vol 120:1–10, 2020 2nd Annual Conference on Learning for Dynamics and Control

Black-box continuous-time transfer function estimation with stability guarantees: a kernel-based approach

Mirko Mazzoleni MIRKO.MAZZOLENI@UNIBG.IT
Department of Management, Information and Production Engineering, University of Bergamo, Via G. Marconi 5, 24044 Dalmine (BG), Italy.

Matteo Scandella MATTEO.SCANDELLA@UNIBG.IT
Department of Management, Information and Production Engineering, University of Bergamo, Via G. Marconi 5, 24044 Dalmine (BG), Italy.

Simone Formentin SIMONE.FORMENTIN@POLIMI.IT
Department of Electronics, Information and Bioengineering, Politecnico di Milano, Via G. Ponzio 34/5, 20133 Milano, Italy.

Fabio Previdi FABIO.PREVIDI@UNIBG.IT
Department of Management, Information and Production Engineering, University of Bergamo, Via G. Marconi 5, 24044 Dalmine (BG), Italy.

Editors: A. Bayen, A. Jadbabaie, G. J. Pappas, P. Parrilo, B. Recht, C. Tomlin, M.Zeilinger

Abstract

Continuous-time parametric models of dynamical systems are usually preferred given their physical interpretation. When there is a lack of prior physical knowledge, the user is faced with the model selection issue. In this paper, we propose a non-parametric approach to estimate a continuous-time stable linear model from data, while automatically selecting a proper structure of the transfer function and guaranteeing to preserve the system stability properties. Results show how the proposed approach outperforms the state of the art.

Keywords: Kernel methods; System identification; Linear identification; Continuous-time identification.

1. Introduction

System identification is the term used in the automatic control field for estimating dynamical models of systems, based on measurements of the system input and output signals. Given the discrete nature of sampled data, the community mostly focused on *discrete-time models*, developing methods either in time (Ljung and Glad (2016)) or frequency domain (Pintelon and Schoukens (2012)).

However, discrete-time models present the following shortcomings (Garnier and Young (2012)): (i) the model is valid only for a fixed sampling frequency; (ii) they must rely on uniformly sampled data; (iii) their performance degrade with stiff systems; (iv) physical insight is more difficult. Such issues can be tackled by *continuous-time models*, which are not defined by a specific sampling frequency. Most of the methods devised for this scope are parametric and require the prior knowledge of the system complexity, see Young (2015); Garnier (2015); Chen et al. (2015).

Nonparametric approaches such as kernel methods allow to select the bias-variance trade-off in a *continuous* way, due to the presence of various regularization terms (Formentin et al. (2019); Pillonetto et al. (2014); Mazzoleni et al. (2018a,b, 2019)). Their use in continuous-time Linear-

Time-Invariant (LTI) system identification is advocated in Pillonetto and Nicolao (2010) where Bounded-Input Bounded-Output (BIBO) stability of the identified model is guaranteed by the so-called *stable-spline* kernel. However, current approaches *need a second stage* to switch from a nonparametric (*discrete-time*) estimate to a parametric transfer function model representation (combining a model-reduction step and a conversion from discrete to continuous model), that is usually more suited for control applications (Pillonetto and Nicolao (2010); Mazzoleni et al. (2018c)).

This work presents a novel approach for *direct* nonparametric *continuous-time* identification of the (parametric) transfer function of asymptotically stable LTI systems. The method: (ii) automatically chooses model complexity; (ii) preserves the stability property of the system under study.

2. Setting and goal

Consider the continuous causal Single-Input Single-Output (SISO) LTI system $\breve{\mathscr{G}}$ with impulse response $\breve{g}:\mathbb{R}\to\mathbb{R}$. The input/output relation of $\breve{\mathscr{G}}$ is $y\left(t\right)=\left[\breve{g}\star u\right]\left(t\right)=\int_0^{+\infty}\breve{g}\left(\xi\right)u\left(t-\xi\right)\,d\xi$, where $u:\mathbb{R}_+\to\mathbb{R}$ and $y:\mathbb{R}_+\to\mathbb{R}$ are, respectively, the input and the output signals, and $\star$ denotes the convolution operator. In the Laplace domain, this relation becomes $Y\left(s\right)=\breve{G}\left(s\right)U\left(s\right)$, where, being $\mathcal{L}$ the Laplace operator, $U\left(s\right)=\mathcal{L}\left[u\right]\left(s\right)$, $Y\left(s\right)=\mathcal{L}\left[y\right]\left(s\right)$ and $\breve{G}\left(s\right)=\mathcal{L}\left[\breve{g}\right]\left(s\right)$ is the transfer function of the system $\breve{\mathscr{G}}$.

Suppose to have at disposal a dataset containing $n\in\mathbb{N}\backslash\{0\}$ noisy measurements, obtained with an experiment on the plant $\mathcal{D}=\left\{\left(t_i,y_i\right),1\leq i\leq n\right\}$, distributed according to the probabilistic model $y_i=\left[\breve{g}\star u\right]\left(t_i\right)+e_i$, $i=1,\ldots,n$ where $e_i\sim\mathcal{N}\left(0,\eta^2\right)$ are independent and identically distributed output-error Gaussian noises and $u:\mathbb{R}\to\mathbb{R}$ is the known input excitation used during the experiment. The excitation signal $u\left(t\right)$ is applied to the plant at the time instant $d\in\mathbb{R}$, i.e. $u\left(t\right)=0$, $\forall t<d$.

The aim is now to estimate the (continuous-time) impulse response $\breve{g}$ of the SISO LTI system $\breve{\mathcal{G}}$ using the noisy dataset $\mathcal{D}$ and the knowledge of the form of u. Following the rationale reported in Pillonetto et al. (2014); Pillonetto and Nicolao (2010), we can estimate $\breve{g}$ by

$$\hat{g}=\underset{g\in\mathcal{H}_k}{\arg\min}\left\{J\left(g\right)\right\},\quad J\left(g\right)=\sum_{i=1}^{n}\left(y_i-\left[g\star u\right]\left(t_i\right)\right)^2+\tau\left\|g\right\|_{\mathcal{H}}^2,\tag{1}$$

where $\mathcal{H}$ is a Reproducing Kernel Hilbert Space (RKHS) with kernel $k:\mathbb{R}_+\times\mathbb{R}_+\to\mathbb{R}$, $\tau>0$ controls the regularization strength and $\left\|\cdot\right\|_{\mathcal{H}}$ is the induced norm of the space $\mathcal{H}$. This estimator can be written as in Dinuzzo and Schölkopf (2012) $\hat{g}^u\left(t\right)=\sum_{i=1}^{n}c_i\hat{g}_i^u\left(t\right)$, where the dependency on the input u is highlighted and $\hat{g}_i^u\left(t\right)=\int_0^{\infty}u\left(t_i-\xi\right)k\left(t,\xi\right)\,d\xi$. The coefficients vector $\boldsymbol{c}=\left[c_1,\ldots,c_n\right]^\top\in\mathbb{R}^{n\times1}$ is found solving $\boldsymbol{O}\left(\boldsymbol{O}+\tau\boldsymbol{I}_n\right)\boldsymbol{c}=\boldsymbol{O}\boldsymbol{y}^\top$, where $\boldsymbol{y}=\left[y_1,\ldots,y_n\right]\in\mathbb{R}^{1\times n}$ and $\boldsymbol{O}\in\mathbb{R}^{n\times n}$ is a symmetric positive-definite matrix whose (i,j) element is $o^u\left(t_i,t_j\right)$,

$$o^u\left(t_i,t_j\right)=\int_0^{+\infty}\int_0^{+\infty}u\left(t_i-\psi\right)u\left(t_j-\xi\right)k\left(\psi,\xi\right)\,d\xi d\psi.\tag{2}$$

The tuning of the hyperparameters of the method $\boldsymbol{\zeta}=\left[\boldsymbol{\psi}^\top,\tau\right]^\top\in\mathbb{R}^{n_\zeta\times1}$ can be performed by resorting to its Bayesian interpretation (Pillonetto et al. (2014)).

A generic order stable-spline kernel can be represented as follows.

Proposition 1 (Spline kernels) *The stable-spline kernel $k_q : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}$ of order q and with λ, β strictly positive hyperparameters can be written as*

$$k_q(a,b) = \lambda \sum_{h=0}^{q-1} \gamma_{q,h} \begin{cases} e^{-\beta[(2q-h-1)a+hb]} & \text{if } a \geq b \\ e^{-\beta[(2q-h-1)b+ha]} & \text{if } a < b \end{cases}; \quad \gamma_{q,h} = \frac{(-1)^{q+h-1}}{h!\,(2q-h-1)!}, \tag{3}$$

Proof The proof is straightforward and omitted for the sake of brevity. ∎

The next section shows the proposed method to *directly* estimate the transfer function of the system without: (i) estimating a discrete-time impulse response $\hat{g}^u$, (ii) performing model reduction, (iii) converting the model from discrete-time to continuous time.

3. Asymptotically stable transfer function estimation from impulse input excitation

3.1. Continuous-time transfer function identification

Proposition 2 (TF expression) *Given the non-parametric estimator $\hat{g}^u$ of an LTI system, the corresponding transfer function $\hat{G}^u(s)$ is*

$$\hat{G}^u(s) = \sum_{i=1}^{n} c_i \hat{G}_i^u(s)\,; \;\; \hat{G}_i^u(s) = \int_d^{t_i} u(x)\, K(s; t_i - x)\; dx; \quad K(s;x) = \int_0^{\infty} k(t,x)\, e^{-s\tau}\, dt. \tag{4}$$

Proof The transfer function of an LTI system corresponds to the Laplace transform of its impulse response. For this reason, we have

$$\hat{G}^u(s) = \mathcal{L}[\hat{g}^u](s) = \int_0^{\infty} \hat{g}^u(t)\, e^{-st}\, dt = \int_0^{\infty} \left(\sum_{i=1}^{n} c_i \hat{g}_i^u(t) \right) e^{-st}\, dt = \sum_{i=1}^{n} c_i \hat{G}_i^u(s)\,,$$

where the term $\hat{G}_i^u(s) = \mathcal{L}[\hat{g}_i^u](s)$ reads as

$$\hat{G}_i^u(s) = \int_0^{\infty} \hat{g}_i^u(t)\, e^{-st}\, dt = \int_0^{\infty} \left[\int_0^{\infty} u(t_i - \xi)\, k(t,\xi)\; d\xi \right] e^{-st}\, dt = \int_0^{\infty} u(t_i - \xi)\, K(s;\xi)\; d\xi.$$

At last, since the integral can be limited to $t_i - d$, with a change of variable $x = t_i - \xi$, we obtain

$$\hat{G}_i^u(s) = \int_d^{t_i} u(x)\, K(s; t_i - x)\; dx.$$

∎

From the above result, we note that the estimated transfer function is composed by the convolution of two terms: the first one, $u(x)$, depends only on the shape of the excitation signal while the second one, $K(s; t_i - x)$, depends only on the kernel used. For the stable-spline kernel of order q, it is possible to compute a more informative formulation thanks to the following proposition.

Proposition 3 (Stable spline TF expression) *Let the kernel be a stable-spline k_q of order q with hyperparameter β, the term $\gamma_{q,h}$ as in Proposition 1 and the time instant d as defined in Section 2. The identified transfer function can be written as*

$$\hat{G}^u(s) = \lambda \left[\sum_{h=0}^{q-1} Q^u_{q,h}(s) + H^u_q(s) \right], \tag{5}$$

where

$$Q^u_{q,h}(s) = \frac{\gamma_{q,h}}{s+\beta h} \left(\sum_{i=1}^{n} c_i A^u_i\big(\beta(2q-h-1)\big) \right); \tag{6}$$

$$H^u_q(s) = \frac{(-1)^q \beta^{2q-1}}{\prod_{i=0}^{2q-1}(\beta i + s)} \left(\sum_{i=1}^{n} c_i A^u_i\big(s+\beta(2q-1)\big) \right); \quad A^u_i(x) = \int_d^{t_i} u(t)\, e^{x(t-t_i)}\, dt. \tag{7}$$

Proof Let us start by analyzing the term $K_q(s;x) = \int_0^\infty k_q(x,t)\, e^{-st}\, dt$, where the kernel k_q is a stable-spline of order q. It is useful to note that the parameter $x \in \mathbb{R}$ is always greater than 0 because in (4) this argument is always positive, thanks to the assumption that $t_i > d$. It is convenient to divide this integral in two parts:

$$K_q(s;x) = \int_0^x k_q(x,t)\, e^{-st}\, dt + \int_x^\infty k_q(x,t)\, e^{-st}\, dt.$$

Firstly, let us focus on the first integral:

$$\int_0^x k_q(x,t)\, e^{-st}\, dt = \int_0^x \lambda \sum_{h=0}^{q-1} \gamma_{q,h} e^{-\beta[(2q-h-1)x+ht]} e^{-st}\, dt = \lambda \sum_{h=0}^{q-1} \gamma_{q,h} e^{-\beta(2q-h-1)x} \int_0^x e^{-(s+\beta h)t}\, dt$$
$$= \lambda \sum_{h=0}^{q-1} \gamma_{q,h} \frac{e^{-\beta(2q-h-1)x}}{s+\beta h} \left(1 - e^{-(s+\beta h)x}\right) = \lambda \sum_{h=0}^{q-1} \gamma_{q,h} \left(\frac{e^{-\beta(2q-h-1)x}}{s+\beta h} - \frac{e^{-(s+\beta(2q-1))x}}{s+\beta h} \right).$$

Analogously, the second integral can be simplified as

$$\int_x^\infty k_q(x,t)\, e^{-st}\, dt = \int_x^\infty \lambda \sum_{h=0}^{q-1} \gamma_{q,h} e^{-\beta[(2q-h-1)t+hx]} e^{-st}\, dt = \lambda \sum_{h=0}^{q-1} \gamma_{q,h} e^{-\beta hx} \int_x^\infty e^{-(s+\beta(2q-h-1))t}\, dt$$
$$= \lambda \sum_{h=0}^{q-1} \gamma_{q,h} \frac{e^{-(s+\beta(2q-h-1))x}}{s+\beta(2q-h-1)}.$$

Thus, $K_q(s;x)$ can be reformulated as

$$K_q(s;x) = \lambda \sum_{h=0}^{q-1} \gamma_{q,h} \frac{e^{-\beta(2q-h-1)x}}{s+\beta h} + \lambda \sum_{h=0}^{q-1} \gamma_{q,h} \left[\frac{e^{-(s+\beta(2q-1))x}}{s+\beta(2q-h-1)} - \frac{e^{-(s+\beta(2q-1))x}}{s+\beta h} \right].$$

This can be further simplified by noting that

$$\sum_{h=0}^{q-1} \gamma_{q,h} \left(\frac{1}{s+\beta(2q-h-1)} - \frac{1}{s+\beta h} \right) = \frac{(-1)^q \beta^{2q-1}}{\prod_{i=0}^{2q-1}(s+\beta i)}.$$

obtaining

$$K_q(s;x) = \lambda \sum_{h=0}^{q-1} \gamma_{q,h} \frac{e^{-\beta(2q-h-1)x}}{s+\beta h} + \lambda e^{-(s+\beta(2q-1))x} \frac{(-1)^q \beta^{2q-1}}{\prod_{i=0}^{2q-1}(\beta i + s)}.$$

Now, it is possible to plug $K_q(s;a)$ in (4) to obtain $\hat{G}_i^u$ for the stable-spline kernel.

$$G_i^u(s) = \int_d^{t_i} u(y) \cdot K_q(s;t_i - y)\, dy = \lambda \sum_{h=0}^{q-1} \frac{\gamma_{q,h}}{s+\beta h} \underbrace{\int_d^{t_i} u(y)\, e^{-\beta(2q-h-1)(t_i-y)}\, dy}_{A_i^u\left(\beta(2q-h-1)\right)} +$$

$$+ \lambda \frac{(-1)^q \beta^{2q-1}}{\prod_{i=0}^{2q-1}(\beta i + s)} \underbrace{\int_d^{t_i} u(y)\, e^{-(s+\beta(2q-1))(t_i-y)}\, dy}_{A_i^u\left(s+\beta(2q-1)\right)} = \lambda \sum_{h=0}^{q-1} \frac{\gamma_{q,h}}{s+\beta h} A_i^u\big(\beta(2q-h-1)\big) +$$

$$+ \lambda \frac{(-1)^q \beta^{2q-1}}{\prod_{i=0}^{2q-1}(\beta i + s)} A_i^u\big(s+\beta(2q-1)\big).$$

The identified transfer function using the stable-spline kernel is then

$$\hat{G}^u(s) = \sum_{i=1}^{n} c_i \hat{G}_i^u(s) = \lambda \sum_{h=0}^{q-1} \underbrace{\frac{\gamma_{q,h}}{s+\beta h} \sum_{i=1}^{n} c_i A_i^u\left(\beta(2q-h-1)\right)}_{Q_{q,h}^u(s)} +$$

$$\lambda \underbrace{\frac{(-1)^q \beta^{2q-1}}{\prod_{i=0}^{2q-1}(s+\beta i)} \sum_{i=1}^{n} c_i A_i^u\big(s+\beta(2q-1)\big)}_{H_q^u(s)} = \lambda \left[\sum_{h=0}^{q-1} Q_{q,h}^u(s) + H_q^u(s)\right].$$

■

The expression (5) represents the estimated transfer function as a sum of $q+1$ transfer functions. The first q of them have one real pole located in a multiple of the $-\beta$ and a gain that depends on the coefficients $\boldsymbol{c}$, the hyper-parameters λ and β, the spline order and the shape of the input signal $u(t)$. The last one has $2q-1$ real poles that are multiple of $-\beta$ and, eventually, other poles that depend on the shape of the input $u(t)$. In particular, the transfer function $A_i^u\big(s+\beta(2q-1)\big)$ can have some poles or zeros that will be added to $\hat{G}^u$. For this reason, to evaluate the asymptotic stability of the identified system, we impose the following condition on the input signal.

Theorem 4 (Excitation for stability) *If the input signal $u(t)$ is such that the terms*

$$A_i^u\big(s+\beta(2q-1)\big), \qquad i = 1, \dots, n \tag{8}$$

are transfer functions whose poles have a negative real part, then $\hat{G}^u(s)$ is an asymptotically stable transfer function.

Proof Since the transfer function $\hat{G}^u$ is defined as the sum of $q+1$ transfer functions, we need to show that all these addends are asymptotically stable. First, let us consider the $q-1$ addends of the type

$$Q_{q,h}^u(s) = \lambda \frac{\gamma_{q,h}}{s+\beta h} \left(\sum_{i=1}^{n} c_i A_i^u\Big(\beta(2q-h-1)\Big) \right),$$

with $h > 0$. All these transfer functions have only one real pole in $-\beta h$ that it is strictly less than zero because $h > 0$ and $\beta > 0$. Therefore, these first $q-1$ transfer functions are asymptotically stable. The remainder of $\hat{G}^u$ is

$$\begin{aligned} R(s) &= \lambda Q^u_{q,0}(s) + \lambda H^u_q(s) \\ &= \frac{\lambda}{s} \cdot \gamma_{q,0} \sum_{i=1}^{n} c_i A^u_i\big(\beta(2q-1)\big) + \frac{\lambda}{s} \cdot \frac{(-1)^q \beta^{2q-1}}{\prod_{i=1}^{2q-1}(\beta i + s)} \sum_{i=1}^{n} c_i A^u_i\big(s + \beta(2q-1)\big). \end{aligned}$$

The poles of the transfer function $R(s)$ are $\{0, -\beta, -2\beta, \dots, -(2q-1)\beta\} \cup (\bigcup_{i=1}^{n} \mathcal{P}_i)$, where $\mathcal{P}_i$ are the poles, which real part is strictly negative (for the hypothesis of the Theorem), of the transfer function $A^u_i\big(s + \beta(2q-1)\big)$. Therefore, the only non-strictly negative pole is the one in 0 because $\beta > 0$. However, there is also a zero in the origin. To see this, consider the transfer function $\widetilde{R}(s)$ such that $R(s) = \frac{\lambda}{s} \cdot \tilde{R}(s)$. Then, the transfer function $R(s)$ has a zero in the origin if and only if $R(0) = 0$. This can be verified with some mathematical steps

$$\begin{aligned} \tilde{R}(0) &= \gamma_{q,0} \sum_{i=1}^{n} c_i A^u_i\big(\beta(2q-1)\big) + \frac{(-1)^q \beta^{2q-1}}{\prod_{i=1}^{2q-1} \beta i} \sum_{i=1}^{n} c_i A^u_i\big(\beta(2q-1)\big) \\ &= \left(\frac{(-1)^{q-1}}{(2q-1)!} + \frac{(-1)^q \cancel{\beta^{2q-1}}}{\cancel{\beta^{2q-1}}(2q-1)!} \right) \sum_{i=1}^{n} c_i A^u_i\big(\beta(2q-1)\big) = \frac{(-1)^{q-1} + (-1)^q}{(2q-1)!} \sum_{i=1}^{n} c_i A^u_i\big(\beta(2q-1)\big). \end{aligned}$$

Since $(-1)^{q-1}$ and $(-1)^q$ have opposite signs for every value of q, we have

$$\tilde{R}(0) = \frac{0}{(2q-1)!} \sum_{i=1}^{n} c_i A^u_i\big(\beta(2q-1)\big) = 0.$$

Therefore $R(s)$ has a zero in the origin that cancels out the pole in 0. Therefore, the identified system $\hat{G}^u$ is asymptotically stable. ■

From this Theorem, it is clear that the terms (8) have an important role in the identification procedure and on the stability of the identified model. Furthermore, note that the identified model is always at least BIBO stable because the stable-spline kernel is a stable kernel.

3.2. Identification using step response data

Consider now the case where a step input is applied at the time instant $d \in \mathbb{R}$, i.e.

$$u(t) = \text{step}(t) = \begin{cases} 1 & \text{if } t \geq d \\ 0 & \text{if } t < d \end{cases}. \tag{9}$$

In this case, since $t_i > d$, we have

$$A^{\text{step}}_i(x) = \int_d^{t_i} \text{step}(t)\, e^{x(t-t_i)}\, dt = e^{-xt_i} \int_d^{t_i} e^{xt}\, dt = e^{-xt_i} \frac{e^{xt_i} - e^{xd}}{x} = \frac{1 - e^{-x(t_i-d)}}{x} \tag{10}$$

Therefore, it is possible to check the condition of Theorem 4. In particular, we have:

$$A^{\text{step}}_i\big(s + \beta(2q-1)\big) = \frac{1 - e^{-\left(s+\beta(2q-1)\right)(t_i-d)}}{s + \beta(2q-1)} = \frac{1}{s + \beta(2q-1)} - e^{-s(t_i-d)} \frac{e^{-\beta(2q-1)(t_i-d)}}{s + \beta(2q-1)}.$$

This is a sum of two transfer functions (the second one with an input-ouput delay) that share the same pole in $p = -\beta(2q-1)$. Since $q \in \mathbb{N}$, $q \geq 1$ and $\beta > 0$, this pole is strictly negative for every value of the hyper-parameters and the theorem hypothesis is respected.

Applying Theorem 3 and using (10), we can to compute the identified transfer function $\hat{G}^{\text{step}}$

$$\hat{G}^{\text{step}}(s) = \lambda\left[\sum_{h=0}^{q-1} Q_{q,h}^{\text{step}}(s) + H_q^{\text{step}}(s)\right],$$

$$Q_{q,h}^{\text{step}}(s) = \frac{\gamma_{q,h}\sum_{i=1}^{n} c_i\left(1 - e^{-\beta(2q-h-1)(t_i-d)}\right)}{\beta(2q-h-1)(s+\beta h)}; \qquad H_{q,h}^{\text{step}}(s)\,. = \frac{(-1)^q\beta^{2q-1}}{\prod_{i=0}^{2q-1}(\beta i + s)}T_q(s)$$

Here, we can note that the transfer function $H_q^{\text{step}}(s)$ contains a non-rational term $T_q(s)$. This non-rational term can approximated using a specialized Padé approximant as explained in Remark 5. The derived kernel in this specific case reads as

$$o^{\text{step}}(t_i, t_j) = \sum_{h=0}^{q-1}\gamma_{q,h}\begin{cases} w_h(t_i, t_j) & \text{if } t_i \geq t_j \\ w_h(t_j, t_i) & \text{if } t_i < t_j, \end{cases}$$

where the term $w_h(t_i, t_j)$, when $h = 0$, is equal to

$$w_0(t_i, t_j) = 2\frac{1 - e^{-\beta(t_j-d)(2q-1)}}{\beta^2(2q-1)^2} - \frac{(t_j - d)\left(e^{-\beta(t_i-d)(2q-1)} + e^{-\beta(t_j-d)(2q-1)}\right)}{\beta(2q-1)}.$$

Instead, for $h > 0$, $w_h(t_i, t_j)$ is equal to:

$$w_h(t_i, t_j) = 2\frac{1 - e^{-\beta(t_j-d)(2q-1)}}{\beta^2(2q-h-1)(2q-1)} + \frac{e^{-\beta(t_j-d)(2q-1)} - e^{\beta(t_j-d)(2q-h-1)}}{\beta^2 h(2q-h-1)} + \\ - \frac{e^{-\beta(t_i-d)(2q-h-1)}\left(1 - e^{-\beta(t_j-d)h}\right)}{\beta^2 h(2q-h-1)}.$$

Remark 5 *In this work, a rational approximation $\widetilde{T}(s)$ of $T(s)$ is achieved by using a Padé approximant (Baker and Graves-Morris (1996)). In particular, a specialized approximant for $T_q(s)$ of order 25 was developed following the rationale described in Baker and Graves-Morris (1996).*

4. Simulation example

In the last decades, continuous-time system identification was studied in detail (Garnier (2015)). The most recent methods are implemented in the `CONTSID toolbox` (Garnier and Gilson (2018)). This section shows simulation results, where we compare the Simple Refined Instrumental Variable (SRIVC) method (Garnier (2015); Young (2011)) with the the proposed approach, using a step input.

The proposed method is tested on three different LTI systems

$$G_1(s) = -\frac{27}{20}\frac{2000s^3 + 3600s^2 + 2095s + 396}{1350s^4 + 7695s^3 + 12852s^2 + 7796s + 1520}$$

$$G_2(s) = 1600\frac{1 - 4s}{s^4 + 5s^3 + 408s^2 + 416s + 1600}; \quad G_3(s) = -\frac{1}{10}\frac{N_{G_3}(s)}{D_{G_3}(s)},$$

with $N_{G_3} = 1869s^4 + 17400s^3 + 68220s^2 + 72350s + 5075$, $D_{G_3} = 1000s^5 + 4419s^4 + 14160s^3 + 27180s^2 + 22220s + 5168$. The system $G_2(s)$ is the Rao-Garnier benchmark used in Garnier (2015); Rao and Garnier (2002). In all the simulations, we employ the stable-spline kernel with order q chosen in the set $\{1, 2, 3, 4, 5\}$. The `SRIVC` method requires the knowledge of the order of the system under analysis. In this comparison, the YIC (Young Information Criterion),see Young (2011), is used to select the best model order (the number of poles and zeros ranges from 1 to 5).

The output of the true model is compared with the estimated one on a test dataset, obtained using a random white Gaussian noise with 10 Hz of bandwidth as excitation signal. Both input and output are sampled for 1000 s. Then, the performance is computed according to the following fit index: Fit $= \left(1 - \frac{\sum_{t=1}^{n_v}(y_t - \hat{y}_t)^2}{\sum_{t=1}^{n_v}\left(y_t - \sum_{t=1}^{n_v} y_t\right)^2}\right) \cdot 100\%$, where n_v is the length of the obtained dataset, y_t and $\hat{y}_t$, with $t = 1, \ldots, n_v$, are, respectively, the samples of the true response and the estimated one.

The comparison with the `SRIVC` approach is performed in the following settings: (i) the input signal is a step; (ii) the dataset is composed by 250 output measurements, taken between 0 and $T = 4$s for $\mathcal{G}_1$, $T = 12$s for $\mathcal{G}_2$, $T = 15$s for $\mathcal{G}_3$; (iii) the dataset is sampled regularly, i.e. $t_i = \frac{i \cdot T}{250}$; (iv) the measurements noise has variance $\eta^2_{step} = 2.78 \cdot 10^{-2}$ for $\mathcal{G}_1$, $\eta^2_{step} = 6.80 \cdot 10^{-1}$ for $\mathcal{G}_2$, $\eta^2_{step} = 6.74 \cdot 10^{-3}$ for $\mathcal{G}_3$; (v) the pool of the possible number of poles and zeros for the `SRIVC` method is $\{1, 2, 3, 4, 5, 6\}$. The Signal-Noise-Ratio (SNR) is 5.

The results of a Monte Carlo simulation with 100 different noise values are reported in Figure 1. In the second example, the Rao-Garnier benchmark system(Garnier (2015); Rao and Garnier (2002)), is the one where the proposed kernel approach has more difficulties, but the median fit is still slightly better than the one obtained with the `CONTSID toolbox`.

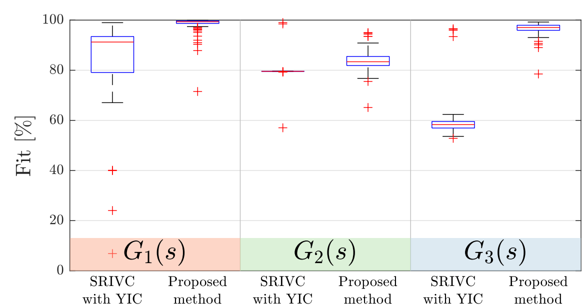

Figure 1: Comparison between the proposed method and `SRIVC` from the `CONTSID toolbox`.

5. Concluding remarks

This paper presented a novel black-box non-parametric continuous-time LTI identification technique. The proposed methodology directly identifies a parametric transfer function model, can work with non-regularly sampled data-points and preserves the stability properties of the system. The method showed very good performance when compared to the method proposed in Garnier (2015); Garnier and Gilson (2018) using a step input. Furthermore, a general parametrization of the stable-spline kernel is derived. Future research will be devoted to the development of the proposed method with more general excitation signals, along with optimal experiment design.

References

George A. Baker and Peter Graves-Morris. *Padé approximants*, volume 59. Cambridge University Press, 1996. ISBN 978-0521450072.

Fengwei Chen, Hugues Garnier, and Marion Gilson. Robust identification of continuous-time models with arbitrary time-delay from irregularly sampled data. *Journal of Process Control*, 25: 19–27, jan 2015. doi: 10.1016/j.jprocont.2014.10.003.

Tianshi Chen. On kernel design for regularized LTI system identification. *Automatica*, 90:109 – 122, apr 2018. ISSN 0005-1098. doi: 10.1016/j.automatica.2017.12.039.

Tianshi Chen, Henrik Ohlsson, and Lennart Ljung. On the estimation of transfer functions, regularizations and gaussian processes—revisited. *Automatica*, 48(8):1525–1535, aug 2012. ISSN 0005-1098. doi: 10.1016/j.automatica.2012.05.026.

Francesco Dinuzzo and Bernhard Schölkopf. The representer theorem for hilbert spaces: a necessary and sufficient condition. In F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 25*, pages 189–196. Curran Associates, Inc., 2012.

Simone Formentin, Mirko Mazzoleni, Matteo Scandella, and Fabio Previdi. Nonlinear system identification via data augmentation. *Systems & Control Letters*, 128:56 – 63, jun 2019. ISSN 0167-6911. doi: 10.1016/j.sysconle.2019.04.004.

H. Garnier and M. Gilson. CONTSID: a matlab toolbox for standard and advanced identification of black-box continuous-time models. *IFAC-PapersOnLine*, 51(15):688–693, 2018. ISSN 2405-8963. doi: 10.1016/j.ifacol.2018.09.203. 18th IFAC Symposium on System Identification SYSID 2018.

Hugues Garnier. Direct continuous-time approaches to system identification. overview and benefits for practical applications. *European Journal of Control*, 24:50 – 62, jul 2015. ISSN 0947-3580. doi: 10.1016/j.ejcon.2015.04.003. SI: ECC15.

Hugues Garnier and Liuping Wang, editors. *Identification of Continuous-time Models from Sampled Data*. Springer London, March 2008. ISBN 978-1-84800-160-2. doi: 10.1007/978-1-84800-161-9.

Hugues Garnier and Peter C. Young. What does continuous-time model identification have to offer? *IFAC Proceedings Volumes*, 45(16):810 – 815, jul 2012. ISSN 1474-6670. doi: https://doi.org/10.3182/20120711-3-BE-2027.00233. 16th IFAC Symposium on System Identification.

Lennart Ljung and Torkel Glad. *Modeling & Identification of Dynamic Systems*. Studentlitteratur AB, August 2016. ISBN 9144116888.

M. Mazzoleni, M. Scandella, S. Formentin, and F. Previdi. Identification of nonlinear dynamical system with synthetic data: a preliminary investigation. *IFAC-PapersOnLine*, 51(15):622 – 627, 2018a. ISSN 2405-8963. doi: https://doi.org/10.1016/j.ifacol.2018.09.227. 18th IFAC Symposium on System Identification SYSID 2018.

Mirko Mazzoleni, Simone Formentin, Matteo Scandella, and Fabio Previdi. Semi-supervised learning of dynamical systems: a preliminary study. In *2018 European Control Conference (ECC)*, pages 2824–2829. IEEE, June 2018b. doi: 10.23919/ECC.2018.8550550.

Mirko Mazzoleni, Matteo Scandella, Simone Formentin, and Fabio Previdi. Classification of light charged particles via learning-based system identification. In *2018 IEEE Conference on Decision and Control (CDC)*, pages 6053–6058. IEEE, dec 2018c. doi: 10.1109/cdc.2018.8618946.

Mirko Mazzoleni, Matteo Scandella, and Fabio Previdi. A comparison of manifold regularization approaches for kernel-based system identification. *Accepted for publication in Proceedings of the 13th IFAC Workshop on Adaptive and Learning Control Systems (ALCOS-2019), Guildhall Winchester, Winchester, United Kingdom*, 2019.

Valentin Pascu, Hugues Garnier, Lennart Ljung, and Alexandre Janot. Benchmark problems for continuous-time model identification: Design aspects, results and perspectives. *Automatica*, 107: 511 – 517, sep 2019. ISSN 0005-1098. doi: 10.1016/j.automatica.2019.06.011.

Gianluigi Pillonetto and Giuseppe De Nicolao. A new kernel-based approach for linear system identification. *Automatica*, 46(1):81–93, jan 2010. ISSN 0005-1098. doi: 10.1016/j.automatica.2009.10.031.

Gianluigi Pillonetto, Francesco Dinuzzo, Tianshi Chen, Giuseppe De Nicolao, and Lennart Ljung. Kernel methods in system identification, machine learning and function estimation: A survey. *Automatica*, 50(3):657–682, mar 2014. ISSN 0005-1098. doi: 10.1016/j.automatica.2014.01.001.

Rik Pintelon and Johan Schoukens. *System Identification*. John Wiley & Sons, April 2012.

Ganti Prasada Rao and Hugues Garnier. Numerical illustrations of the relevance of direct continuous-time model indentification. *IFAC Proceedings Volumes*, 35(1):133–138, 2002. ISSN 1474-6670. doi: https://doi.org/10.3182/20020721-6-ES-1901.01008. 15th IFAC World Congress.

A. Varga. Balancing free square-root algorithm for computing singular perturbation approximations. In *[1991] Proceedings of the 30th IEEE Conference on Decision and Control*, pages 1062–1065 vol.2. IEEE, December 1991. doi: 10.1109/CDC.1991.261486.

Peter C. Young. *Recursive Estimation and Time-Series Analysis*. Springer Berlin Heidelberg, 2011. doi: 10.1007/978-3-642-21981-8.

Peter C. Young. Refined instrumental variable estimation: Maximum likelihood optimization of a unified box-jenkins model. *Automatica*, 52:35 – 46, feb 2015. ISSN 0005-1098. doi: https://doi.org/10.1016/j.automatica.2014.10.126.