Preprints of the 20th World Congress
The International Federation of Automatic Control
Toulouse, France, July 9-14, 2017

Bayesian Frequency Domain Identification of LTI Systems with OBFs Kernels ⋆

Mohamed A. H. Darwish [*,**] **John Lataire** [***] **Roland Tóth** [*]

[*] *Control Systems Group, Department of Electrical Engineering, Eindhoven University of Technology, P.O. Box 513, 5600 MB, Eindhoven, The Netherlands (e-mail: m.a.h.darwish@tue.nl).*
[**] *Electrical Engineering Department, Faculty of Engineering, Assiut University, 71515 Assiut, Egypt.*
[***] *Dept. ELEC (Fundamental Electricity and Instrumentation), Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium.*

Abstract: Regularised *Frequency Response Function* (FRF) estimation based on Gaussian process regression formulated directly in the frequency-domain has been introduced recently. The underlying approach largely depends on the utilised kernel function, which encodes the relevant prior knowledge on the system under consideration. In this paper, we show how to construct a rich class of kernel functions, directly in the frequency-domain, based on Orthonormal Basis Functions (OBFs), which is capable of representing a wide range of dynamical properties, e.g., stability, resonance frequencies, damping, etc, in terms of the poles of the employed basis functions that are treated as hyperparameters to efficiently shape the model class, i.e., the prior in the corresponding Bayesian setting. This class of kernel functions also implicitly guarantees the stability of the estimated FRF. The generating poles of the OBFs are tuned along with other hyperparameters, e.g., noise variance, by maximising the marginal likelihood. Multiple case studies are considered to show the potential of the considered kernels.

Keywords: Gaussian process regression; Orthonormal basis functions; Regularisation, Frequency-domain; Kernel functions.

1. INTRODUCTION

Nonparametric estimation of *Transfer Functions* (TF) of *Linear Time-Invariant* (LTI) systems provides valuable information about the dynamics of the system under consideration that can be used further to obtain an accurate parametric model [Pintelon and Schoukens, 2012, Ljung, 1999]. The evaluation of the TF on the unit circle will be called the *Frequency Response Function* (FRF) and it has been studied extensively in the literature [Antoni and Schoukens, 2007, Pintelon and Schoukens, 2012], etc.

One main challenge in the data-driven estimation of FRF is the transient effect, which is due to the fact that the input and the output signals are not periodic or their periodicity does not match the length of the measurement window. As a result, most of the available approaches try suppressing the transient effect in different ways. More specifically, via spectral analysis as in Schoukens et al. [2006], or via a frequency-dependent smoothing procedure that is applied to the *Empirical Transfer Function Estimate* (ETFE) [Stenman et al., 2000]. More recent approaches are estimating both the FRF and the transient simultaneously [Pintelon and Schoukens, 2012], e.g., the *Local Polynomial Method* (LPM) [Pintelon et al., 2010a,b], which uses a local polynomial smoother, and the *Local Rational Method* (LRM) [McKelvey and Guérin, 2012], which uses a local rational function as a smoother. The drawbacks of these approaches, i.e., LPM and LRM, are: i) both methods provide a set of local models centered around the bins of the *Discrete Fourier Transform* (DFT), for which the interpolation between the DFT bins is still an open question; ii) the stability of the resulting estimates cannot be guaranteed.

Alternatively, inspired by new developments of nonparametric estimation of LTI impulse response models in the time-domain [Pillonetto et al., 2014, Chen et al., 2012], regularised frequency domain estimates of both the FRF and the transient effects within the *Gaussian Process Regression* (GPR) framework has been introduced in Lataire and Chen [2016]. More specifically, both the FRF and the transient are assumed to be a realisation of a zero-mean real/complex GP [Schreier and Scharf, 2010] with a certain covariance (kernel) function that encodes the relevant prior knowledge on the system, e.g., smoothness and stability, etc. The direct formulation of the estimation problem in the frequency domain offers many advantages: i) it allows the estimation to be performed in a limited frequency band; ii) it allows for an efficient implementation for continuous-time systems.

One critical aspect that is related to the aforementioned approach is the design of the kernel function. It has to be

⋆ This research has benefited from the financial support of the Student Mission, Ministry of Higher Education, Government of Egypt, and of the Belgian Federal Government (Interuniversity Attraction Poles programme, IAP VII-DYSCO)

Copyright by the International Federation of Automatic Control (IFAC)

flexible enough to describe a wide range of dynamical properties, e.g., stability, resonance behaviour, damping, etc., and at the same time parameterised by a low number of hyperparameters. In Lataire and Chen [2016], the kernels from the time-domain, e.g., *Diagonal/Correlated* (DC) kernel [Chen et al., 2012] and *Stable/Spline* (SS) kernel [Pillonetto and De Nicolao, 2010], have been formulated in the frequency-domain. Moreover, it has been shown that for both of these kernels the resulting estimates are stable, i.e., all poles of the estimated FRF lie inside the unit circle.

On the other hand, inspired by realisation theory of dynamical systems, a rich class of kernel functions for impulse response estimation based on *Orthonormal Basis Functions* (OBFs) has been introduced [Chen and Ljung, 2015, Darwish et al., 2015]. These OBFs are generated by a cascaded network of stable inner transfer functions, i.e., all-pass filters, completely determined, modulo the sign, by their poles. In the frequency-domain, these OBFs constitute a complete orthonormal basis for the function space $\mathcal{RH}_2$, which is the space of functions over $\mathbb{C}$, that are squared integrable on the unit circle, analytic outside of it and their inverse $z-$transform is a real sequence [Heuberger et al., 2005, 1995]. This space is a *Reproducing Kernel Hilbert Space* (RKHS) [Aronszajn, 1950] with a well-defined reproducing kernel function [Heuberger et al., 2005]. In terms of regularised estimate of FRF, this kernel function has been suggested in [Chen and Ljung, 2015, Equation 17]. However, the stability of the estimated FRF by using such kernels has not been discussed. Furthermore, the problem of choosing the proper number of basis functions to be used has not been addressed.

In this paper, the formulation of a stable OBFs-based kernels formulated directly in the frequency-domain is given. The stability of the estimated FRF is guaranteed by introducing a decay term that weighs the OBFs. This simultaneously circumvents the problem of selecting the number of basis functions to be introduced in the model. This class of kernel functions provides an efficient way to incorporate a wide range of prior knowledge, i.e., resonance behaviour, stability, damping, via the generating poles of the OBFs. Special cases of the presented kernels will be discussed in details, i.e., Laguerre, Kautz, *Generalised* OBFs (GOBFs) [Heuberger et al., 1995]. The generating poles are considered to be unknown hyperparameters and are tuned via maximising the *Marginal Likelihood* (ML) [Rasmussen and Williams, 2006, Pillonetto and Chiuso, 2015].

This paper is organised as follows. Section 2 presents the problem formulation, whereas Bayesian frequency domain identification is discussed in Section 3. Section 4 gives an overview of OBFs, their associated RKHS, reproducing kernel defined in the frequency-domain and how such kernel can be adopted for regularised FRF estimation. Extensive simulation studies are reported in Section 5. Finally, the paper is concluded in Section 6.

Notation

In the following: $\mathbb{C}$ denotes the complex plane, $\mathbb{D}$ is the interior of the unit disc, i.e., $\{z \in \mathbb{C} \mid |z| < 1\}$. $\mathcal{E}$ denotes the expectation operator. z^* denotes the complex conjugate of a complex number $z \in \mathbb{C}$, whereas the superscript H denotes the Hermitian (complex conjugate) transpose of a vector. $\mathbb{R}$ stands for real numbers, while $\mathbb{Z}$ denotes the set of the integer numbers and $\mathbb{N}$ is the set of all positive integers. $|A|$ is the determinant of a square matrix A.

2. PROBLEM FORMULATION

Consider a *Single-Input Single-Output* (SISO), finite-order, asymptotically stable and LTI discrete-time data-generating system described by

$$y(t) = (g * u)(t) + v(t), \tag{1}$$

where $t \in \mathbb{Z}$ is the discrete-time, $y : \mathbb{Z} \to \mathbb{R}$ is the output, $u : \mathbb{Z} \to \mathbb{R}$ is the input of the system, $v(t)$ is a zero-mean quasi-stationary noise process, independent of u, and $(g * u)(t)$ is the discrete convolution of the (im)pulse response $g(\cdot)$ and the input $u(\cdot)$ at time instant t and is defined as

$$(g * u)(t) = \sum_{\tau=0}^{\infty} u(t-\tau)g(\tau). \tag{2}$$

It is assumed that $y(t)$ is measured at $t = 0, 1, \ldots, N-1$.

Denote $\Omega = e^{j\omega}$ the frequency variable for $\omega \in \mathbb{R}$, then Ω_k, for $k \in \mathbb{R}$, is defined as

$$\Omega_k = e^{j\omega_k} = e^{\frac{j2\pi k}{N}}. \tag{3}$$

It is worth to mention that for $k \in \mathbb{Z}$, Ω_k corresponds to the k^{th} bin of an $N-$point DFT, where for the sampled signal $x(\tau), \tau = 0, \ldots, N-1$, the $N-$point DFT, at frequency bin k is given by

$$X(k) = \frac{1}{\sqrt{N}} \sum_{\tau=0}^{N-1} x(\tau) e^{\frac{-j2\pi k\tau}{N}}, \quad k \in \mathbb{Z}. \tag{4}$$

Accordingly, denote $U(k), Y(k), V(k)$ the $N-$point DFT at frequency bin k of $u(t), y(t), v(t)$, respectively.

The frequency-domain representation of (1) is

$$Y(k) = G(\Omega_k)U(k) + T(\Omega_k) + V(k), \tag{5}$$

where the FRF $G(\Omega)$ is computed as the *Discrete-Time Fourier Transform* of the impulse response $g(t)$, i.e., $G(\Omega) = \mathcal{F}\{g(t)\} = \sum_{\tau=0}^{\infty} g(\tau)e^{-j\omega\tau}$ and $T(\Omega)$ is the transient, which depends on the difference $u(t) - u(t+N)$, for $t < 0$, and on the impulse response of the system [Lataire and Chen, 2016, Lemma 1].

Given the exact input and measured output $N-$point DFT spectra $U(k)$ and $Y(k)$, we are aiming at estimating $G(\Omega)$ and $T(\Omega)$.

3. BAYESIAN FREQUENCY-DOMAIN IDENTIFICATION

In the Bayesian approach to system identification within the GPR framework, the unknown function to be estimated is assumed to be a realisation of a zero-mean GP with a certain covariance (kernel) function that encodes our priori knowledge about it. Given observed data of joint Gaussian processes and *a prior* mean and covariance, the goal is to obtain the *a posteriori* mean and covariance, which can be used for prediction of the unknown function at arbitrary input values.

For the LTI system (1), it holds true that the FRF takes both real (typically at 0 Hz and at the Nyquist frequency) and complex values. As a result, it is not possible to model it as a real or as a complex GP. Hence, both the FRF and the transient have to be defined as a *Real/Complex* GP (RCGP) [Lataire and Chen, 2016, Section 2]. More specifically, an RCGP $\eta(k)$ is defined as

$$\eta(k) \sim \mathcal{RCGP}(m, K, C) \mid \mathbb{K}_{\mathcal{R}}, \tag{6}$$

where m, K, C are the mean, covariance, and relation functions, respectively, and $\mathbb{K}_{\mathcal{R}}$ is a set of indices that indicates where $\eta(k)$ is real, i.e., $\mathbb{K}_{\mathcal{R}} = \{0, \pm N/2, \pm 2(N/2), \ldots\}$. Following the Bayesian approach within the GPR framework, the FRF $G(\Omega_k)$ and the transient $T(\Omega_k)$ are assumed to be independent of each other and are assumed to be RCGPs over $k \in \mathbb{R}$:

$$G(\Omega_k) \sim \mathcal{RCGP}(0, \alpha_{\mathrm{G}} K, \alpha_{\mathrm{G}} C) \mid \mathbb{K}_{\mathcal{R}} \tag{7}$$

$$T(\Omega_k) \sim \mathcal{RCGP}(0, \alpha_{\mathrm{T}} K, \alpha_{\mathrm{T}} C) \mid \mathbb{K}_{\mathcal{R}}, \tag{8}$$

where $\alpha_{\mathrm{G}} \geq 0, \alpha_{\mathrm{T}} \geq 0$, K, C are well-defined covariance and relation functions, respectively. Once K, C are defined, the *Maximum a Posteriori* (MAP) estimates $\hat{G}$ and $\hat{T}$ of the FRF and the transient, respectively, can be easily computed [Lataire and Chen, 2016].

3.1 Kernel functions in the frequency-domain

A natural way to construct the kernel function for FRF estimation is to utilise the duality between the FRF and impulse response function, i.e., $G(e^{j\omega}) = \sum_{\tau=0}^{\infty} g(\tau) e^{-j\omega\tau}$, and the linearity of the Fourier transform, to derive the corresponding covariance and relation functions in the frequency-domain. More specifically, if the impulse response function g is assumed to be a realisation of a zero-mean GP with covariance $\operatorname{cov}(g(t), g(s)) = \alpha_{\mathrm{G}} P(t, s),\ t, s = 0, 1, \ldots$, then, $G(e^{j\omega})$ is an RCGP with

$$\mathcal{E}\{G(e^{j\omega})\} = \mathcal{F}\{\mathcal{E}\{g(t)\}\} = 0, \tag{9}$$

$$\alpha_{\mathrm{G}} K(e^{j\omega_k}, e^{j\omega_l}) = \mathcal{E}\{G(e^{j\omega_k}) G^*(e^{j\omega_l})\}, \tag{10}$$

$$\alpha_{\mathrm{G}} C(e^{j\omega_k}, e^{j\omega_l}) = \mathcal{E}\{G(e^{j\omega_k}) G(e^{j\omega_l})\} = \alpha_{\mathrm{G}} K(e^{j\omega_k}, e^{-j\omega_l}), \tag{11}$$

$$K(e^{j\omega_k}, e^{j\omega_l}) = \sum_{\tau=0}^{\infty} \sum_{\tau'=0}^{\infty} P(\tau, \tau') e^{-j\omega_k \tau} e^{j\omega_l \tau'}. \tag{12}$$

In Lataire and Chen [2016], the authors make use of such approach to define kernel functions used in the time-domain based literature for FRF estimation, e.g., *Diagonal/Correlated* (DC) [Chen et al., 2012] and *Stable/Spline* (SS) kernel [Pillonetto and De Nicolao, 2010]. For the sake of space, we refer the reader to the formulation of these kernels in the frequency domain in [Lataire and Chen, 2016, Equations 55,56]. Furthermore, it has been proven that these kernels guarantee the stability of the resulting estimates. A sufficient condition on the kernel function to guarantee the stability of the estimated FRF is to satisfy the condition in [Lataire and Chen, 2016, Property 7] or equivalently, the corresponding impulse response of the estimated FRF must be absolutely summable, i.e., $g(t) = \mathcal{F}^{-1}\{\hat{G}(e^{j\omega})\}$, where $\mathcal{F}^{-1}$ denotes the inverse Fourier transform.

Remark 1. Regarding the kernel function for the transient $T(\Omega)$, it has been shown in [Lataire and Chen, 2016, Section 5.3] that a computational convenient way is to assume $G(\Omega)$ and $T(\Omega)$ have the same kind of covariance function, but with different scaling hyperparameters α_{G} and α_{T}.

The aforementioned kernels can describe stability and smoothness of the estimated FRF. However, as recommended in Lataire and Chen [2016], kernels that are able to describe other dynamical properties would be beneficial in the frequency-domain identification, e.g., resonance behaviour, damping, etc., but keeping a simple structure of the kernel function. In the next section, we show how to construct a class of kernel functions directly in the frequency-domain which have such advantages via the use of OBFs.

4. RATIONAL ORTHONORMAL BASIS FUNCTIONS

In this section, we introduce a general class of OBFs, namely, GOBFs, and its special case, i.e., *Laguerre* and *Kautz* basis, their properties, the corresponding RKHS and the associated reproducing kernel function.

4.1 An overview

Since we are interested in proper, stable and real rational transfer functions with real-valued impulse response, we introduce a particular set of orthonormal basis functions which constitute a complete basis for $\mathcal{RH}_2$. Let $G_{\mathrm{b}} \in \mathcal{RH}_{2-}$, i.e., the subspace of $\mathcal{RH}_2$ which is restricted to real, rational and proper functions, be an inner function, i.e., all-pass filter, which satisfy $G_{\mathrm{b}}(e^{j\omega}) G_{\mathrm{b}}^*(e^{j\omega}) = 1$ with McMillan degree $n_{\mathrm{g}} > 0$. Let (A, B, C, D) be a minimal balanced *State-Space* realisation of $G_{\mathrm{b}}(e^{j\omega})$. Note that G_{b} is completely determined, modulo the sign, by its poles $\Lambda_{n_{\mathrm{g}}} = [\xi_1 \cdots \xi_{n_{\mathrm{g}}}] \in \mathbb{D}^{n_{\mathrm{g}}}$:

$$G_{\mathrm{b}}(e^{j\omega}) = \pm \prod_{i=1}^{n_{\mathrm{g}}} \frac{1 - \xi_i^* e^{j\omega}}{e^{j\omega} - \xi_i}, \tag{13}$$

with $\Lambda_{n_{\mathrm{g}}}$ containing real poles and/or complex conjugate pole pairs. The class of GOBFs is obtained by cascading identical $n_{\mathrm{g}}^{\mathrm{th}}$ order all-pass filters and can be written in a vector form as [Heuberger et al., 2005]:

$$V_\tau(e^{j\omega}) = V_1(e^{j\omega}) G_{\mathrm{b}}^{\tau-1}(e^{j\omega}), \quad \text{for } \tau > 1, \tag{14}$$

where $V_1(e^{j\omega}) = (e^{j\omega} I - A)^{-1} B$. Let $\varphi_i = [V_1]_i$ denote the i^{th} element of the vector transfer function V_1. Then, a particular GOBF basis consists of the functions

$$\Psi = \{\psi_\tau\}_{\tau=1}^{\infty} = \left\{\varphi_i G_b^l\right\}_{i=1, l=0}^{n_{\mathrm{g}}, \infty}, \quad \text{with } \tau = l \cdot n_{\mathrm{g}} + i. \tag{15}$$

These functions, i.e., (15), constitute a complete orthonormal basis for $\mathcal{RH}_2$. As a result, any $G(e^{j\omega}) \in \mathcal{RH}_2$ can be decomposed as

$$G(e^{j\omega}) = \sum_{\tau=1}^{\infty} c_\tau \psi_\tau(e^{j\omega}), \tag{16}$$

which is the generalisation of *Trigonometric Basis Functions* (TBF), i.e., $\left\{e^{-j\omega\tau}\right\}_{\tau=1}^{\infty}$, where all the generating poles are assumed to be at 0 . It can be shown that the rate of convergence of this series expansion is bounded by $\rho = \max_\tau |G_{\mathrm{b}}(\varepsilon_\tau^{-1})|$, called the decay rate, where $\{\varepsilon_\tau\}$ are the poles of $G(e^{j\omega})$ [Oliveira e Silva, 1996]. In practice, only a finite number of terms $\{\psi_\tau\}_{\tau=1}^{n_{\mathrm{b}}}$, i.e., truncation of the expansion (16) is used. However, a model structure corresponding to a truncation of (16) can achieve an arbitrary low modeling error with a relatively small number of parameters due to the faster convergence of the series representation than in the TBF case, which in system identification results in decreased variance of the final model estimate [Tóth et al., 2009].

The two interesting special cases of GOBFs defined in (14) are detailed below, namely, the (2-*parameter Kautz* *functions*) for $n_{\mathrm{g}} = 2$, and as *Laguerre functions* for $n_{\mathrm{g}} = 1$ [Heuberger et al., 2005]. Laguerre basis are defined as

$$\psi_\tau(e^{j\omega}) = \frac{\sqrt{(1-\xi^2)}}{e^{j\omega}-\xi}\left(\frac{1-\xi e^{j\omega}}{e^{j\omega}-\xi}\right)^{\tau-1}, \quad \xi \in (-1,1), \quad (17)$$

where the parameter ξ is known as the Laguerre parameter or generating pole. The impulse response of Laguerre basis functions exhibit an exponential decay. However, Laguerre functions do not allow the use of complex poles, hence, they are less suitable, i.e., they offer a lower achievable decay rate in capturing systems with oscillatory response. In this case, the two-parameter Kautz basis functions result in a more appropriate structure. The two-parameter Kautz basis are the set of orthonormal functions

$$\psi_{2\tau-1}=\frac{\sqrt{1-c^2}(e^{j\omega}-b)}{e^{j2\omega}+b(c-1)e^{j\omega}-c}\left(\frac{-ce^{j2\omega}+b(c-1)e^{j\omega}+1}{e^{j2\omega}+b(c-1)e^{j\omega}-c}\right)^{\tau-1}$$
$$\psi_{2\tau}=\frac{\sqrt{(1-c^2)(1-b^2)}}{e^{j2\omega}+b(c-1)e^{j\omega}-c}\left(\frac{-ce^{j2\omega}+b(c-1)e^{j\omega}+1}{e^{j2\omega}+b(c-1)e^{j\omega}-c}\right)^{\tau-1}, \quad (18)$$

where $b, c \in (-1, 1)$. Note that (18) corresponds to a repeated complex pair $\xi, \xi^* \in \mathbb{D}$ [Wahlberg, 1994].

4.2 OBFs based kernels in the frequency domain

It is well-known that the space spanned by the OBFs, i.e., $\mathcal{RH}_2$, is a RKHS [Ninness et al., 1999] with the following reproducing kernel

$$K(e^{j\omega}, e^{j\omega'}) = \sum_{\tau=0}^{\infty} \psi_\tau(e^{j\omega})\psi_\tau^*(e^{j\omega'}), \quad (19)$$

and with the following well-defined inner product

$$\langle f_1, f_2\rangle_{\mathcal{RH}_2} = \frac{1}{2\pi}\int_{-\pi}^{\pi} f_1(e^{j\omega})f_2^*(e^{j\omega})dw, \quad (20)$$

for any $f_1, f_2 \in \mathcal{RH}_2$.

Now, let's investigate the stability of the estimated FRF. Similarly to the reasoning presented in Darwish et al. [2015], when the space spanned by the OBFs. i.e., $\mathcal{RH}_2$, where the functions in that space are not necessarily rational, is utilised as a hypothesis space, the corresponding impulse response function estimates will belong to $\mathcal{R}\ell_2(\mathbb{N})$, i.e., the space of squared summable real sequences, due to the isomorphism between $\mathcal{RH}_2$ and $\mathcal{R}\ell_2(\mathbb{N})$ [Oliveira e Silva, 1996]. The stability condition for finite-dimensional LTI systems (with a rational proper transfer function) is that the impulse response should be absolutely summable, i.e., $\hat{g}(t) \in \mathcal{R}\ell_1(\mathbb{N})$, where $\mathcal{R}\ell_1(\mathbb{N})$ is the space of absolutely summable real sequences. However, $\mathcal{R}\ell_2(\mathbb{N}) \not\subset \mathcal{R}\ell_1(\mathbb{N})$, which means that the hypothesis space should be restricted to a subspace that only contains real, rational and proper functions, i.e., $\mathcal{RH}_{2-}$. In order to guarantee the stability of the estimated FRF and to solve the problem of determining the right number of basis functions, in this paper, we include a decay term that weighs the OBFs

$$K(e^{j\omega}, e^{j\omega'}) = \sum_{\tau=0}^{\infty} d_\tau(\beta)\psi_\tau(e^{j\omega})\psi_\tau^*(e^{j\omega'}), \quad (21)$$

where the decay term $d_\tau(\beta) \to 0$ as $\tau \to \infty$. Possible choices are

$$d_\tau(\beta) = \tau^{-\beta},\ d_\tau(\beta) = \beta^\tau \text{ with } \beta \geq 0,\ 0 \leq \beta < 1, \quad (22)$$

respectively, where β is considered to be a hyperparameter that determines the decay rate of the expansion in (21) and the decay term, i.e., $d_\tau(\beta)$, with β tuned by marginal likelihood optimization acts as an automatic way to select the number of significant basis functions that are needed to construct the kernel. Note that the relation function C can be constructed accordingly via (21) and (10)-(11).

Remark 2. In absence of more sophisticated prior information, in many circumstances, monotonically decreasing weights (22) is effective and able to well guard against the ill-conditioning affecting the system identification problem. However, depending on the available knowledge, other parameters can be introduced in the decay term that describe more complicated shapes for the weights. Similarly, when a prior information is available, this can support the choice of the basis functions. This fits in the framework developed in the paper, e.g., if the number of resonance peaks is known, we can use such information to decide the number of real or complex pairs that should be considered for GOBFs.

4.3 Hyperparameters tuning

The kernel function defined above, i.e., the OBFs based kernel, depends on some unknown hyperparameters that need to be tuned from the observed data. These hyperparameters are the scaling parameters α_{G} and α_{T}, the noise variance σ^2, β the parameter that determines the decay rate of the expansion and ϑ a vector of the generating poles of the OBFs. Denote by θ the vector of the hyperparameters, i.e., $\theta = [\,\alpha_{\mathrm{G}}\ \alpha_{\mathrm{T}}\ \sigma^2\ \beta\ \vartheta^\top\,]^\top$. One popular approach to tune θ within the Bayesian framework is by maximising the log *Marginal Likelihood*, i.e., $\log p(Y(\mathcal{K}) \mid \theta)$ of the output spectrum [Rasmussen and Williams, 2006]

$$\log\left(Y(\mathcal{K} \mid \theta)\right) = -\frac{1}{2}Y(\mathcal{K})^H\Gamma_{\mathrm{Y}}^{-1}(\theta)Y(\mathcal{K}) - \frac{1}{2}\log|\Gamma_{\mathrm{Y}}(\theta)| - \frac{n_{\mathrm{r}}}{2}\log 2\pi - n_{\mathrm{c}}\log\pi, \quad (23)$$

where $\mathcal{K} = \{k_1, k_2, \ldots, k_{\mathrm{n}}\} \subset \{0, \ldots, N/2\}$ is the set of DFT-frequency indices that lie in the frequency band of interest, Γ_{Y} is the augmented covariance matrix and can be constructed from the covariance and relation functions, i.e., K and C (21) and (10)-(11), see [Lataire and Chen, 2016, Equation 36] for constructing Γ_{Y}, n_{r} is the number of frequencies where the FRF has real values and n_{c} is the number of frequencies where the FRF has complex values.

5. SIMULATION STUDIES

In this section, the presented OBFs based kernel function is tested and compared to the existing kernels, i.e., DC kernel, for FRF estimation. A challenging system is considered to show the capability of the presented kernel to model a wide range of dynamical properties, specifically resonance behaviour, with a simple kernel structure.

5.1 Considered system

We consider a randomly generated 20th order, LTI and discrete-time system G generated by the `drss` Matlab function. The sampling period T_{s} is $1\,s$. We make sure that there are two dominant complex conjugate pole pairs. These dominant poles are located at $0.95 \pm j0.25$, $-0.17 \pm j0.89$, see Figure 1 for the the impulse response and the pole/zero plots of the generated system.

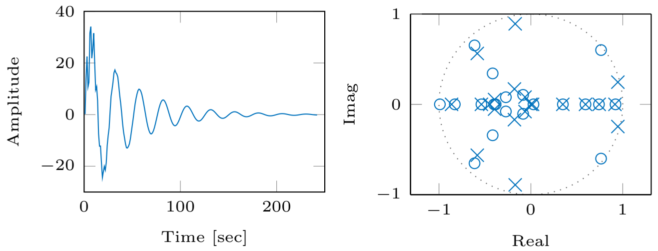

Fig. 1. Left: The impulse response. Right: Poles (×) and zeros (o), of the considered system.

5.2 Identification settings

Monte-Carlo (MC) simulations of 100 runs are performed, where at each run a new realisation of the input $u(t)$ and the noise $v(t)$ are utilised according to (1). The considered system is used to generate a data set of length $N = 512$ for each MC run using a white and periodic zero-mean input u and an additive white Gaussian noise v. The variance of v is chosen such that the *Signal-to-Noise Ratio* (SNR), which is defined as

$$\mathrm{SNR} = 10\log_{10}\left(\frac{\sum_{\tau=1}^{N}\tilde{y}^2(\tau)}{\sum_{\tau=1}^{N}v^2(\tau)}\right) \qquad (24)$$

where $\tilde{y}$ denotes the noise-free system output, is corresponding to two case for the SNR: 10 dB or 40 dB.

The considered estimators are:

- GPTF with DC kernel;
- GPTF with OBFs based kernel, specifically, with GOBFs based kernel where the poles of the inner function G_b are $\{\xi_1, \xi_1^*, \xi_2, \xi_2^*\}$, which are considered as hyperparameters;
- Parametric model identified with the Identification toolbox of Matlab (2016a), more specifically, an *Output-Error* (OE) model with the true order of the system, i.e., using the command `oe(20,20)`. We will call this estimator as an Oracle estimator, in the sense that it knows the true model structure and order.

For each MC run, the hyperparameters of the GPTF estimators, for both DC and OBFs kernels, are tuned by maximising the ML (23). The estimation is performed on a limited frequency band, i.e., from $\omega = 0.1$ rad/s to $\omega = 3$ rad/s. For the GPTF estimator, 241 frequency domain samples in the mentioned range were used, whereas the OE model was estimated based on the whole data record.

5.3 Results and discussion

The performance measure that is used to determine the quality of the estimated FRF with different estimators is the averaged *Mean Squared Error* (MSE) over all frequencies in the band of interest, i.e.,

$$\mathrm{MSE} = \frac{1}{100}\sum_{i=1}^{100}\left(\frac{1}{N}\sum_{k=1}^{N}|\hat{G}_i(\Omega_k) - G(\Omega_k)|^2\right), \qquad (25)$$

where $\hat{G}_i$ is the estimated FRF at the MC run i, which is calculated on a more dense frequency grid, i.e., 966 frequencies, but within the same frequency band as the training data set.

The averaged MSE for the estimates over all the frequencies in the considered frequency band is summarised in Table 1. It can be seen from the table that the GPTF estimators perform better than the parametric estimator, even though the latter makes use of more data points and more importantly it makes use of the true model structure. Moreover, the GPTF estimator with the GOBFs based kernel shows a significant improvement with respect to the GPTF estimate with the DC kernel. The main reason is that the complex conjugate poles included in the GOBFs based kernel are better at modeling the resonance behaviour and result in a smoother estimate. To visualise such results, the left parts of Figure 2, 3, show the estimated FRF and the true function at the validation set of frequencies of one MC run within the considered frequency band for both cases of SNR of 10 dB and 40 dB, respectively. The right parts of the figures show the error associated with the employed estimators in dB. From these figures, it can be easily seen that the OBFs based kernel performs well compared to the DC kernel estimator and can deliver an acceptable response even in the high frequency range.

Table 1.
Averaged MSE of all estimates (in dB) for different SNR scenarios.

Estimator	10 dB	40 dB
GPTF (DC)	46.45	-0.28
GPTF (GOBFs)	33.46	-7.05
OE (20th-order)	50.94	6.89

6. CONCLUSION

In this paper, we presented the formulation of the OBFs-based kernels directly in the frequency-domain. These kernels can be used for simultaneously obtaining a regularised estimate for the FRF and the transient in the Bayesian setting within the GPR framework. Such formulation results in a flexible class of kernel functions, which are capable of describing a wide range of dynamical properties, e.g., stability, resonance behaviour, damping, etc, directly via the generating poles of the OBFs. The generating poles are dealt with as unknown hyperparameters and are tuned by maximising the ML. Special cases of the presented class of kernel functions are considered and their capability to model the FRF and the transient is shown with extensive simulation studies.

REFERENCES

J. Antoni and J. Schoukens. A comprehensive study of the bias and variance of frequency-response-function measurements: Optimal window selection and overlapping strategies. *Automatica*, 43(10):17231736, 2007.

N. Aronszajn. Theory of reproducing kernels. *Trans. of the American mathematical society*, pages 337–404, 1950.

T. Chen and L. Ljung. Regularized system identification using orthonormal basis functions. In *Proc. of the European Control Conference*, pages 1291 – 1296, Linz, Austria, July 2015.

T. Chen, H. Ohlsson, and L. Ljung. On the estimation of transfer functions, regularizations and Gaussian processes–revisited. *Automatica*, 48(8):1525–1535, 2012.

M. A. H. Darwish, G. Pillonetto, and R. Tóth. Perspectives of orthonormal basis functions based kernels in Bayesian system identification. In *Proc. of the 54th IEEE Conference on Decision and Control*, pages 2713–2718, Osaka, Japan, Dec. 2015.

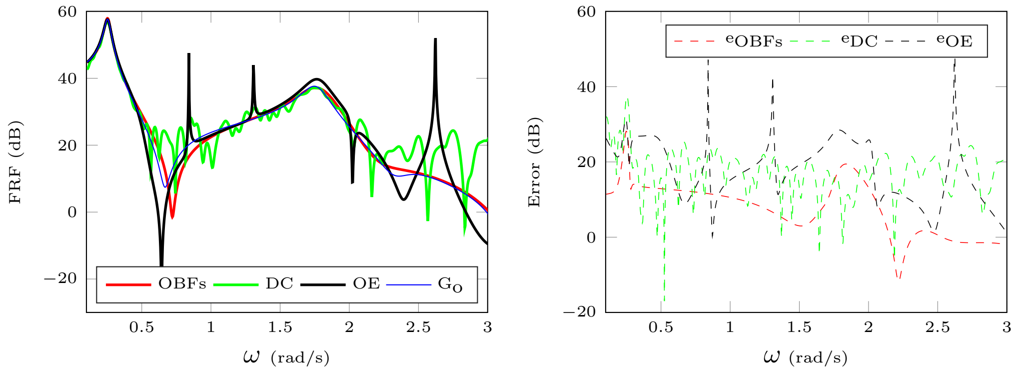

Fig. 2. Left: Plot of one MC realisation of the estimated FRF with different estimators and the true one in case of SNR=10dB. Right: The error (in dB) associated with the considered estimators.

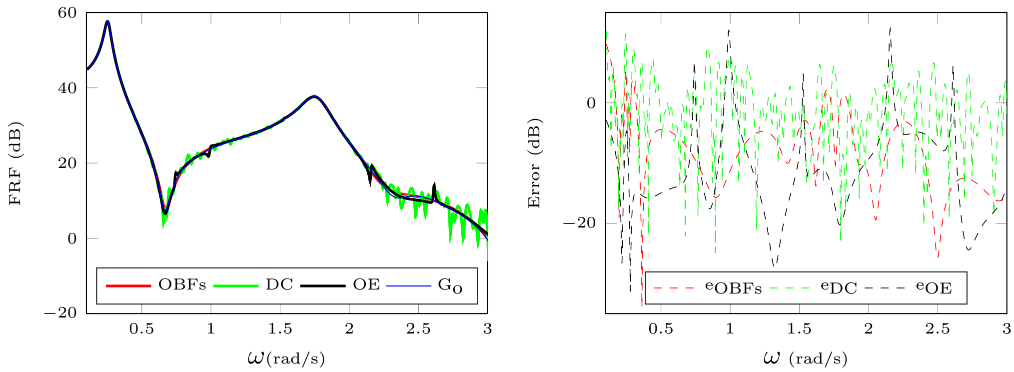

Fig. 3. Left: Plot of one MC realisation of the estimated FRF with different estimators and the true one in case of SNR=40dB. Right: The error (in dB) associated with the considered estimators.

P. S. C. Heuberger, P. M. J. Van den Hof, and O. H. Bosgra. A generalized orthonormal basis for linear dynamical systems. 40(3):451–465, 1995.

P. S. C. Heuberger, P. M. J. Van den Hof, and B. Wahlberg. *Modeling and Identification with Rational Orthonormal Basis Functions*. Springer-Verlag, 2005.

J. Lataire and T. Chen. Transfer function and transient estimation by Gaussian process regression in the frequency domain. *Automatica*, 72:217–229, 2016.

L. Ljung. *System identification - Theory for the User*. Prentice-Hall, 1999.

T. McKelvey and G. Guérin. Non-parametric frequency response estimation using a local rational model. In *Proc. of the 16th IFAC Symposium on System Identification*, pages 49 –54, Brussels, Belgium, July 2012.

B. Ninness, H. Hjalmarsson, and F. Gustafsson. Generalized Fourier and Toeplitz results for rational orthonormal bases. *SIAM Journal on Control and Optimization*, 37(2):429–460, 1999.

T. Oliveira e Silva. A n-width result for the generalized orthonormal basis function model. In *Proc. of the 13th IFAC World Congress*, pages 375–380, Sydney, Australia, July 1996.

G. Pillonetto and A. Chiuso. Tuning complexity in regularized kernel-based regression and linear system identification: The robustness of the marginal likelihood estimator. *Automatica*, 58:106–117, 2015.

G. Pillonetto and G. De Nicolao. A new kernel-based approach for linear system identification. *Automatica*, 46(1):81–93, 2010.

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*, 50(3):657–682, 2014.

R. Pintelon and J. Schoukens. *System identification, a frequency domain approach*. IEEE Press, Wiley, 2012.

R. Pintelon, J. Schoukens, G. Vandersteen, and K. Barbé. Estimation of nonparametric noise and frf models for multivariable systemspart I: Theory. *Mechanical Systems and Signal Processing*, 24(3):573–595, 2010a.

R. Pintelon, J. Schoukens, G. Vandersteen, and K. Barbé. Estimation of nonparametric noise and frf models for multivariable systemspart II: Extensions, applications. *Mechanical Systems and Signal Processing*, 24(3):596–616, 2010b.

C. E. Rasmussen and C. K. I. Williams. *Gaussian processes for machine learning*. The MIT Press, 2006.

J. Schoukens, Y. Rolain, and R. Pintelon. Analysis of windowing/leakage effects in frequency response function measurements. *Automatica*, 42(1):27–38, 2006.

P. J. Schreier and L. L. Scharf. *Statistical signal processing of complex-valued data: the theory of improper and noncircular signals*. Cambridge University Press, 2010.

A. Stenman, F. Gustafsson, D. E. Rivera, L. Ljung, and T. McKelvey. On adaptive smoothing of empirical transfer function estimates. *Automatica*, 8(11):13091315, 2000.

R. Tóth, P. S. C. Heuberger, and P. M. J. Van den Hof. Asymptotically optimal orthonormal basis functions for LPV system identification. *Automatica*, 45(6):1359–1370, 2009.

B. Wahlberg. System identification using Kautz models. *IEEE Trans. on Automatic Control*, 39(6):1276–1282, 1994.