Automatica 72 (2016) 217–229

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

# Transfer function and transient estimation by Gaussian process regression in the frequency domain[☆]

John Lataire[a], Tianshi Chen[b]

[a] Vrije Universiteit Brussel (VUB), Department ELEC, Pleinlaan 2, 1050 Brussels, Belgium
[b] School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen, China

ARTICLE INFO

*Article history:*
Received 2 August 2014
Received in revised form
23 October 2015
Accepted 26 May 2016

*Keywords:*
Transfer function estimation
Gaussian process regression
Frequency domain
System identification
Regularised least squares

ABSTRACT

Inspired by the recent promising developments of Bayesian learning techniques in the context of system identification, this paper proposes a Transfer Function estimator, based on Gaussian process regression. Contrary to existing kernel-based impulse response estimators, a frequency domain approach is adopted. This leads to a formulation and implementation which is seamlessly valid for both continuous- and discrete-time systems, and which conveniently enables the selection of the frequency band of interest. A pragmatic approach is proposed in an output error framework, from sampled input and output data. The transient is dealt with by estimating it simultaneously with the transfer function.

Modelling the transfer function and the transient as Gaussian processes allows for the incorporation of relevant prior knowledge on the system, in the form of suitably designed kernels. The SS (Stable Spline) and DC (Diagonal Correlated) kernels from the literature are translated to the frequency domain, and are proven to impose the stability of the estimated transfer function. Specifically, the estimates are shown to be stable rational functions in the frequency variable. The hyperparameters of the kernel are tuned via marginal likelihood maximisation.

The comparison between the proposed method and three existing methods from the literature – the regularised finite impulse response (RFIR) estimator, the Local Polynomial Method (LPM), and the Local Rational Method for Frequency Response Function estimation – is illustrated on simulations on multiple case studies.



## 1. Introduction

When considering the identification of Linear Time Invariant (LTI) systems, an important initial step is the non-parametric estimation of their Transfer Functions (TF) (Ljung, 1985), (Pintelon & Schoukens, 2012, Chapters 2–7). It provides the user with insight into the dynamic behaviour of the system, even before any attempt is made to determine a parametric model.

The estimation of a transfer function ought to consider that the available signals are confined to a finite time interval. This is typically handled by including initial conditions (time domain) or an additional transient (frequency domain) in the estimation process. The transient takes into account the fact that the input and output signals are not necessarily periodic, or that their periodicity does not match the length of the measurement window, as explained later on.

The Frequency Response Function (FRF) of a system is defined in Pintelon and Schoukens (2012, Chapter 2) as the evaluation of its TF – a continuous function – at a discrete set of frequencies. FRF estimation has been studied extensively in the past, starting with tools for spectral analysis (Antoni & Schoukens, 2007; Bendat & Piersol, 1993; Schoukens, Rolain, & Pintelon, 2006). These tools aim at suppressing the transient by applying carefully designed windows. Alternatively, Stenman, Gustafsson, Rivera, Ljung, and McKelvey (2000) applies a frequency dependent smoothing procedure to the Empirical Transfer Function Estimate (ETFE) to suppress the transient. More recent work makes use of an intrinsic property of the transient to estimate it simultaneously with the FRF, yielding much better results. Namely, both the FRF and the transient are known to be smooth functions of the frequency (Pintelon & Schoukens, 2012, Appendix 6.B). Therefore, their estimation can be performed via smoothers. By following this point of

☆ The material in this paper was not presented at any conference. Corresponding author: Tianshi Chen. Tel.: +86 0755 84273835. A part of Tianshi Chen's work was done when he was with the Division of Automatic Control, Linköping University, Linköping, Sweden. This paper was recommended for publication in revised form by Associate Editor Brett Ninness under the direction of Editor Torsten Söderström.

*E-mail addresses:* jlataire@vub.ac.be (J. Lataire), tschen@cuhk.edu.cn (T. Chen).

http://dx.doi.org/10.1016/j.automatica.2016.06.009

view, Refs. Pintelon, Schoukens, Vandersteen, and Barbé (2010a,b) and Schoukens, Vandersteen, Barbé, and Pintelon (2009) use a local polynomial smoother, and will be referred to as the Local Polynomial Method (LPM), while (McKelvey & Guérin, 2012) discusses the Local Rational Method (LRM), which uses a local rational function as a smoother. It is worth to note that both the LPM and the LRM provide a set of local models centred around the bins of the DFT (Discrete Fourier Transform), for which the interpolation in-between the DFT bins remains an open question. Consequently, the stability of the LPM and LRM estimates is also undefined.

Recently, new results for LTI system identification have been reported (Pillonetto & De Nicolao, 2010; Pillonetto, Chiuso, & De Nicolao, 2011; Chen, Ohlsson, & Ljung, 2012) on the estimation of impulse responses, which is the time domain equivalent of the TF. The impulse response estimation is formulated as a Gaussian process regression problem, which can also be interpreted as a regularisation method. More specifically, the impulse response is modelled as a real and zero mean Gaussian process with suitably chosen and tuned covariance (often called kernel) functions. The impulse response estimate is then given by the conditional mean of the Gaussian process conditioned on the given data. This method will be denoted by RFIR (Regularised least squares for estimating the Finite Impulse Response). It has two unique features. The first one is that the kernel function is designed to embed the prior knowledge, e.g., stability and smoothness of the impulse response, into the estimation problem. The second one is that the model complexity is tuned in a continuous way and handled by maximising the marginal likelihood of the hyper-parameters used to parameterise the kernel function. This approach is known to enable an automatic trade-off between the model fit and the model complexity (MacKay, 1998; Rasmussen & Williams, 2006), and as pointed out in Pillonetto, Dinuzzo, Chen, and De Nicolao (2014), is more reliable than existing complexity measures, such as the Akaike's criterium (AIC) or cross validation, especially for small data sets.

In this paper, Gaussian process regression is applied directly to the TF and the transient estimation. That is, the estimation is formulated in the frequency domain. The resulting estimate will be denoted by GPTF. A particular attention is deserved to the fact that the TF and the transient are complex valued functions, but that at some frequencies – at 0 Hz and at the Nyquist frequency for discrete time systems – they should be real valued. For that reason, the method will be developed in the context of mixed real/complex Gaussian processes. Next, properties of the associated frequency domain kernels, applicable to LTI systems, will be derived and a sufficient condition on the kernel will be formulated to impose the stability of the GPTF estimate. Then, it will be shown how the time domain kernels – Stable Spline and Diagonal Correlated – proposed in Pillonetto and De Nicolao (2010), Pillonetto et al. (2011) and Chen et al. (2012) are transformed to the frequency domain and satisfy the condition of stability.

It will be shown that the RFIR estimate in Pillonetto and De Nicolao (2010), Pillonetto et al. (2011), Chen et al. (2012), Pillonetto et al. (2014) and the GPTF in the frequency domain are dual to each other, under specific conditions. However, from a practical point of view, the frequency-domain formulation is shown to give a more appealing implementation than the RFIR when working with continuous-time systems. This is because the explicit computation of the convolution between the input and the impulse response is circumvented. A second advantage of the GPTF over the RFIR is that it allows the estimation to be performed in a limited frequency band. The main advantages of the GPTF over the LPM and the LRM are that the estimated transfer function is guaranteed to be stable, and that it is expressed as a continuous function of the frequency.

The remaining part of this paper is organised as follows. Gaussian processes for regression of mixed real and complex valued functions are developed in Section 2. The formulation of the TF estimation problem is given in Section 3, and is rewritten as a Bayesian regression problem in Section 4. The choice and construction of kernels in the frequency domain is discussed in Section 5, and the duality with regularised impulse response estimation is given in Section 6. The Gaussian process TF estimator is compared with the LPM, the LRM and the RFIR in Section 7 on case studies. Section 8 concludes this paper. The appendix provides the proofs of the lemma and the theorems.

## 2. Real/complex Gaussian distributions

Transfer Functions (TFs) will be modelled as Gaussian processes. Since a TF takes both real (at 0 Hz and at the Nyquist frequency for discrete time systems) and complex values, it cannot be modelled as either a real or a complex random variable. In particular, treating a real Gaussian random variable as a complex one leads to a singular covariance matrix, see Remark 1. This prompts us to introduce the so-called real/complex Gaussian (RCG) distribution to model the TF.

**Definition 1** (*RCG Distribution*)**.** A random vector

$$Z = \begin{bmatrix} Z_r^T & Z_c^T \end{bmatrix}^T, \quad Z_r \in \mathbb{R}^{n_r},\ Z_c \in \mathbb{C}^{n_c} \tag{1}$$

is said to be real/complex Gaussian distributed, if $\begin{bmatrix} Z_r^T & \Re Z_c^T & \Im Z_c^T \end{bmatrix}^T$ is Gaussian distributed, where $\Re$, $\Im$ denote the real and imaginary parts respectively, and the superscript $T$ denotes the transpose of a vector.

We introduce below its probability density function and derive its characteristic parameters. Define

$$Z_{re} = \begin{bmatrix} Z_r^T & \Re Z_c^T & \Im Z_c^T \end{bmatrix}^T, \qquad z_{re} = \begin{bmatrix} z_r^T & \Re z_c^T & \Im z_c^T \end{bmatrix}^T,$$
$$z_r \in \mathbb{R}^{n_r}, z_c \in \mathbb{C}^{n_c}. \tag{2}$$

The probability density function of $Z_{re}$ is described by

$$p(z_{re}) = \frac{1}{\sqrt{\det 2\pi \Gamma_{re}}} \ldots \times \exp\left(-\tfrac{1}{2}(z_{re} - m_{re})^H \Gamma_{re}^{-1} (z_{re} - m_{re})\right) \tag{3}$$

where $m_{re} = \mathbb{E}\{Z_{re}\}$ and $\Gamma_{re} = \mathbb{E}\left\{(Z_{re} - m_{re})(Z_{re} - m_{re})^T\right\}$, and the superscript $H$ denotes the Hermitian transpose of a vector. Define

$$\begin{bmatrix} m_r^T & m_c^T \end{bmatrix}^T = \mathbb{E}\{Z\} \tag{4a}$$
$$\Gamma_r = \mathbb{E}\left\{(Z_r - m_r)(Z_r - m_r)^T\right\} \tag{4b}$$
$$\Gamma_{rc} = \mathbb{E}\left\{(Z_r - m_r)(Z_c - m_c)^H\right\} \tag{4c}$$
$$\Gamma_c = \mathbb{E}\left\{(Z_c - m_c)(Z_c - m_c)^H\right\} \tag{4d}$$
$$C_c = \mathbb{E}\left\{(Z_c - m_c)(Z_c - m_c)^T\right\} \tag{4e}$$

then it is easy to verify that

$$m_{re} = \begin{bmatrix} m_r^T & \Re m_c^T & \Im m_c^T \end{bmatrix}^T, \tag{5}$$

$$\Gamma_{re} = \begin{bmatrix} \Gamma_r & \Re \Gamma_{rc} & -\Im \Gamma_{rc} \\ \Re \Gamma_{rc}^H & \Re(\Gamma_c + C_c)/2 & \Im(C_c - \Gamma_c)/2 \\ \Im \Gamma_{rc}^H & \Im(\Gamma_c + C_c)/2 & \Re(\Gamma_c - C_c)/2 \end{bmatrix}. \tag{6}$$

Now, define $\tilde{z} = \begin{bmatrix} z_r^T & z_c^T & z_c^H \end{bmatrix}^T$. It holds that

$$z_{re} = M\tilde{z}, \quad \text{with } M = \begin{bmatrix} I_{n_r} & 0 & 0 \\ 0 & I_{n_c}/2 & I_{n_c}/2 \\ 0 & -jI_{n_c}/2 & jI_{n_c}/2 \end{bmatrix}. \tag{7}$$