The Transient Impulse Response Modeling Method and the Local Polynomial Method for nonparametric system identification

Michel Gevers [*] **Per Hägg** [**] **Håkan Hjalmarsson** [**]
Rik Pintelon [***] **Johan Schoukens** [***]

[*] *CESAME, Université catholique de Louvain, B1348 Louvain la Neuve, Belgium, and Dept ELEC, Vrije Universiteit Brussel, Pleinlaan 2, B1050 Brussels, Belgium (e-mail: Michel.Gevers@uclouvain.be)*
[**] *Automatic Control Lab and ACCESS, School of Electrical Engineering, KTH, SE-100 44 Stockholm, Sweden (e-mail: per.hagg,hakan.hjalmarsson@ee.kth.se)*
[***] *Dept ELEC, Vrije Universiteit Brussel, Pleinlaan 2, B1050 Brussels, Belgium (e-mail: Rik.Pintelon, Johan.Schoukens@vub.ac.be)*

Abstract: This paper analyzes two recent methods for the nonparametric estimation of the Frequency Response Function (FRF) from input-output data using Prediction Error identification. Such FRF estimate can be the main goal of the identification exercise, or it can be a tool for the computation of a nonparametric estimate of the noise spectrum. We show that the choice of the method depends on the signal to noise ratio and on the objective. The method that delivers the best FRF estimate may not deliver the best estimate of the noise spectrum. Our theoretical analysis is illustrated by simulations.

Keywords: System identification, nonparametric identification, noise spectrum estimation

1. INTRODUCTION

This paper deals with the nonparametric estimation of the Frequency Response Function (FRF) and of the noise spectrum of a linear dynamic system from input-output data in a Prediction Error (PE) identification framework. The inputs are known but not necessarily periodic, while the outputs are perturbed by quasistationary noise.

There are good reasons for the estimation of these nonparametric quantities. The nonparametric FRF estimate can give a preliminary idea of the complexity of the system and guide the user for the selection of a model structure in a subsequent parametric estimation step. The nonparametric estimate of the noise spectrum can be used as frequency weighting for a subsequent estimation of a parametric input-output model, thereby simplifying the PE criterion and reducing the risk of local minima, as shown in Schoukens et al. (2011).

The estimation of the FRF of the input-output transfer function is obtained from Fourier transforms of finite sets of input and output data, and this introduces *leakage errors* which are the frequency domain equivalent of transient errors in time domain identification. Leakage errors have for a long time been a major deterrent against the use of nonparametric estimates of the FRF in the presence of random input signals. Recently, two new methods have been introduced that significantly reduce the leakage effects in the nonparametric estimation of the FRF. The first method, called *Local Polynomial Method (LPM)* was introduced in Schoukens et al. (2009) and Pintelon et al. (2010). It is based on an approximation of the FRF and the leakage term in a local bandwidth around each frequency, and the estimation of these local parameters as the solution of N local Least Squares (LS) problems, where N is the number of data. The second method was developed in Hägg et al. (2011) and is now called the *TRansient Impulse response Modeling Method (TRIMM)*. It is based on an approximation of the FRF and the leakage term as a function of global parameters which are estimated as the solution of one global LS problem.

In this paper we analyze these two methods and establish the relations between the approximation errors that are made in each one. We show that the two methods are essentially different, and we compare the bias and variance errors induced by each method for the estimation of the FRF. We show that the objective (estimation of the FRF, or estimation of the noise spectrum) and the signal to noise ratio should guide the choice of the method and its design parameters.

The paper is organized as follows. In section 2 we formulate the input-output description of the system and introduce the leakage terms. The LPM and TRIMM methods are presented in section 3, and the error analysis for both methods is presented in section 4. A detailed comparative analysis is performed in section 5. Section 6 deals with the

[⋆] This work is supported by the Fund for Scientific Research (FWO-Vlaanderen), the Flemish Government (Methusalem Fund METH1), and the Belgian Programme on Interuniversity Attraction Poles, initiated by the Belgian Federal Science Policy Office.

estimation of the noise spectrum. In section 7 we illustrate the theoretical results with simulations.

2. THE INPUT OUTPUT DESCRIPTION

Consider a linear discrete-time single-input single-output (SISO) system $G(q)$ that is excited with a known random input signal $u(t)$, and whose output is the sum of the input contribution and of a disturbance term $v(t)$. It is assumed that $u(t)$ and $v(t)$ are quasistationary (Ljung (1999)); thus $v(t)$ can be modeled as the output of a white-noise driven filter. The input-output system can be represented as

$$y(t) = G(q)u(t) + v(t) = G(q)u(t) + H(q)e(t) \quad (1)$$

where q is the forward shift operator, $G(q)$ and $H(q)$ are causal rational functions of q, and $e(t)$ is zero mean white noise with variance σ_e^2. It can also be represented in state-space form as:

$$x(t+1) = Ax(t)+Bu(t)+Ke(t), \quad y(t) = Cx(t)+e(t). \quad (2)$$

The input-output model (1) assumes an infinite record of input and output signals. For a finite record $t = 0,\dots,N-1$ this equation has to be modified to take account of the initial condition (or transient) terms τ_G and τ_H due to the action of the transfer functions G and H, leading to:

$$y(t) = G(q)u(t) + \tau_G(t) + H(q)e(t) + \tau_H(t). \quad (3)$$

Using the discrete Fourier transform (DFT)

$$X(k) = \frac{1}{\sqrt{N}} \sum_{t=0}^{N-1} x(t)e^{-j2\pi kt/N},$$

an exact frequency domain formulation of (3) is obtained:

$$Y(k) = G(\Omega_k)U(k)+T_G(\Omega_k)+H(\Omega_k)E(k)+T_H(\Omega_k) \quad (4)$$

where the index k points to the frequency kf_s/N with f_s the sampling frequency, and $\Omega_k = e^{-j2\pi k/N}$. We also introduce the notation $\omega_k \triangleq \frac{2\pi k}{N}$; hence we can also write $\Omega_k = e^{-j\omega_k}$. The contributions U, E, Y in (4) are an $O(N^0)$ signal, while the transient terms T_G and T_H are $O(N^{-1/2})$, where $X = O(N^p)$ means that $\lim_{N\to 0} |\frac{X}{N^p}| < \infty$.

It is important to understand that (3) and (4) are exact relations (Pintelon et al. (1997); McKelvey (2002); Agüero et al. (2008)). The leakage terms T_G and T_H are rational forms in z^{-1}, and hence smooth functions of the frequency. We shall lump the transient terms together, and define $\tau(t) \triangleq \tau_G(t) + \tau_H(t)$ and, similarly, $T(\Omega_k) \triangleq T_G(\Omega_k) + T_H(\Omega_k)$. For simplicity of notation we shall from now on rewrite the frequency domain expression of (4) as

$$Y_k = G_kU_k + T_k + V_k, \quad k = 1,\dots,N \quad (5)$$

where $G_k = G(\Omega_k), T_k = T(\Omega_k)$ and $V_k = H(\Omega_k)E(k)$. Consider the system (2) with initial condition x_0 and let $x(N)$ be its response at time N. It can then be shown (Pintelon and Schoukens (2001)) that the transient terms can be expressed as a function of the state-space model as follows in the time domain and in the frequency domain:

$$\tau(t) = CA^t(x_0 - x_N) \quad (6)$$

$$T_k = \frac{1}{\sqrt{N}} \sum_{t=0}^{N-1} \tau(t)e^{-j\omega_k t} = \frac{1}{\sqrt{N}} C(I - e^{-j\omega_k}A)^{-1}(x_0 - x_N). \quad (7)$$

It follows that T and G are both smooth functions of frequency. This is the basis for the joint nonparametric estimation of the Frequency Response Functions $G(\Omega)$ and $T(\Omega)$ developed in the two methods that we analyze and compare in this paper: the Local Polynomial Method (LPM) first developed in (Schoukens et al. (2009)) and thoroughly analyzed in (Pintelon et al. (2010)), and the TRIMM method developed in (Hägg et al. (2011)).

3. THE LPM AND THE TRIMM METHOD

The estimation of $\{G_k, T_k, k = 1,\dots,N\}$ from (5) is impossible because there are only N equations for $2N$ unknown parameters. The main idea of the methods discussed in this paper is to generate additional equations by approximating G_k and T_k in a window around each frequency Ω_k - taking account of the smoothness of these functions - and thereby generating a sufficient number of equations. With these additional equations, Least Squares (LS) problems can be set up involving the unknown parameters G_k as well as additional parameters that account for the estimation of T_k and the approximation of G_k in this window. The LPM and TRIMM methods vary essentially in the way the approximations are performed and in the way the LS problems are set up. The LPM sets up N local LS problems involving a vector of local parameters for each frequency window, while the KTH method solves one global LS problem involving global parameters for the approximation of $G(\Omega)$ and $T(\Omega)$. We now develop the expressions that lead to the LS problems.

Suppose that we want to estimate G_k and T_k at frequency ω_k, and consider the data Y_{k+r}, U_{k+r} for $r = -L,\dots,0,\dots,L$ in a window around the frequency ω_k, with window size $2L+1$. We can then write:

$$\begin{aligned} Y_{k+r} &= G_{k+r}U_{k+r} + T_{k+r} + V_{k+r} \quad r = -L,\dots,0,\dots,L \\ &= G_kU_{k+r} + [G_{k+r} - G_k]U_{k+r} + T_{k+r} + V_{k+r}. \quad (8) \end{aligned}$$

The LPM method

With LPM a Taylor series expansion is used for $G_{k+r} - G_k$ and $T_{k+r} - T_k$:

$$G_{k+r} = G_k + \sum_{s=1}^{\infty} g_s(k)r^s, \quad (9)$$

$$T_{k+r} = T_k + \frac{1}{\sqrt{N}} \sum_{s=1}^{\infty} t_s(k)r^s \quad (10)$$

Substituting in (8) yields the following expressions, which are then truncated to produce a well-posed LS problem:

$$\begin{aligned} Y_{k+r} =& \left[G_k + \sum_{s=1}^{R} g_s(k)r^s\right] U_{k+r} + T_k + \frac{1}{\sqrt{N}} \sum_{s=1}^{R} t_s(k)r^s + V_{k+r} \\ &+ \left[\sum_{s=R+1}^{\infty} g_s(k)r^s\right] U_{k+r} + \frac{1}{\sqrt{N}} \sum_{s=R+1}^{\infty} t_s(k)r^s \quad (11) \\ \approx& \left[G_k + \sum_{s=1}^{R} g_s(k)r^s\right] U_{k+r} + T_k + \frac{1}{\sqrt{N}} \sum_{s=1}^{R} t_s(k)r^s + V_{k+r}, \\ &\text{for} \quad r = -L,\dots,0,\dots,L. \quad (12) \end{aligned}$$

Collecting the $2(R+1)$ unknown complex parameters into a local parameter vector

$$\theta_k \triangleq [G_k \;\; g_1(k)\dots g_R(k); T_k \;\; t_1(k)\dots t_R(k)]^T \quad (13)$$

and defining the $(2L+1)$-vectors $\bar{Y}_{k,L}$ and $\bar{U}_{k,L}$ as

$$\bar{Y}_{k,L} \triangleq [Y_{k-L} \; Y_{k-L+1} \; \dots Y_k \; \dots Y_{k+L-1} \; Y_{k+L}]^T$$

$$\bar{U}_{k,L} \triangleq [U_{k-L} \; U_{k-L+1} \; \dots U_k \; \dots U_{k+L-1} \; U_{k+L}]^T$$

then leads to the following matrix version of (12):

$$\bar{Y}_{k,L} = K_{k,L}(R, \bar{U}_{k,L})\theta_k. \quad (14)$$

This constitutes an overdetermined set of $2L+1$ equations in the $2(R+1)$ unknowns θ_k, based on the local data $\{U_{k+r}, Y_{k+r}; r = -L, \dots, 0, \dots, L\}$, in which $K_{k,L}(R, \bar{U}_{k,L})$ is entirely known. This can be solved by Least Squares to estimate θ_k, of which G_k is the first component. The LPM thus consists of the solution of N such local LS problems, one for each frequency.

The Constrained LPM method (CLPM)

The expressions $G_{k+r} \approx G_k + \sum_{s=1}^{R} g_s(k) r^s$ and $T_{k+r} \approx T_k + \frac{1}{\sqrt{N}} \sum_{s=1}^{R} t_s(k) r^s$ induce constraints between the parameter vectors θ_k and $\theta_{k+r}, r = -R, \dots, R$. In Gevers et al. (2011) a variation of the LPM method (called CLPM) was proposed that takes account of these constraints, typically yielding a better estimate of the FRF compared to the standard LPM, but a worse estimate of the noise spectrum, given that the CLPM estimate has a larger bias error. For reasons of space limitations, we shall not analyze CLPM in this paper.

The TRIMM method

In the TRIMM method, the term $G_{k+r} - G_k$ of (8) is expressed as

$$G_{k+r} - G_k = \sum_{t=1}^{\infty} g(t)[e^{-j\omega_{k+r}t} - e^{-j\omega_k t}] \quad (15)$$
$$= \sum_{t=1}^{\infty} g(t)[\Omega_{k+r}^t - \Omega_k^t] = \sum_{t=1}^{\infty} g(t)\phi_t(\omega_{k+r}, \omega_k)$$

where $g(t) \triangleq CA^{t-1}B$ and where $\phi_t(\omega_{k+r}, \omega_k) \triangleq e^{-j\omega_{k+r}t} - e^{-j\omega_k t}$ are known functions of frequency. The frequency function T_{k+r} in (8) can be written as (see (7)):

$$T_{k+r} = \frac{1}{\sqrt{N}} \sum_{t=0}^{N-1} \tau(t) e^{-j\omega_{k+r}t}. \quad (16)$$

We now substitute these two expressions in (8) and we approximate by Finite Impulse Response models :

$$\begin{aligned} Y_{k+r} &= G_k U_{k+r} + [\sum_{t=1}^{n_1} g(t)\phi_t(\omega_{k+r}, \omega_k)] U_{k+r} \\ &\quad + \frac{1}{\sqrt{N}} \sum_{t=0}^{n_2-1} \tau(t) e^{-j\omega_{k+r}t} + V_{k+r} \quad (17) \\ &\quad + \left[\sum_{t=n_1+1}^{\infty} g(t)\phi_t(\omega_{k+r}, \omega_k)\right] U_{k+r} + \frac{1}{\sqrt{N}} \sum_{t=n_2}^{N-1} \tau(t) e^{-j\omega_{k+r}t} \\ &\approx G_k U_{k+r} + \left[\sum_{t=1}^{n_1} g(t)\phi_t(\omega_{k+r}, \omega_k)\right] U_{k+r} + V_{k+r} \\ &\quad + \frac{1}{\sqrt{N}} \sum_{t=0}^{n_2-1} \tau(t) e^{-j\omega_{k+r}t}, \quad r = -L, \dots, L \quad (18) \end{aligned}$$

Equating real part and imaginary part, we note that for each frequency ω_k, the equations (18) constitute a set of $2(2L+1)$ equations for the $2 + n_1 + n_2$ unknowns $\{G_k; g(1), \dots, g(n_1); \; \tau(0), \dots, \tau(n_2-1)\}$. The TRIMM method consists in setting up a large LS problem obtained by assembling the equations (18) for the N DFT frequencies $\omega_1, \dots, \omega_N$ and solving these by Least Squares for the global vector of $2N + n_1 + n_2$ parameters

$$\theta \triangleq [G_1, \dots, G_N; g(1), \dots, g(n_1); \tau(0), \dots, \tau(n_2-1)]^T, \quad (19)$$

where G_k is counted for two unknowns (real and imaginary part). This LS problem consists of $2N(2L+1)$ equations for $2N + n_1 + n_2$ real unknowns. It can be written as:

$$Y = \Phi\theta + E + V \quad (20)$$

where the vector E accounts for the approximation errors. For small values of L the matrix Φ can be poorly conditioned; hence a regularization term can be added leading to the following LS solution:

$$\hat{\theta} = [\Phi^H\Phi + \beta\gamma(\Phi)I]^{-1}\Phi^H Y \quad (21)$$

where Φ^H denotes the complex conjugate transpose of Φ, while $\gamma(\Phi)$ is the Frobenius norm of Φ and β will be taken as 10^{-3} in the simulations of Section 7.

4. ERROR ANALYSIS

In this section we analyze the bias and variance errors of the LPM and TRIMM methods. We assume that A is a stability matrix and we denote by λ_1 the absolute value of its largest eigenvalue, so that for any row vector C and column vector B, we have $|CA^tB| \leq \alpha\lambda_1^t$ for some real positive number α. Thus $|g(t)| \leq \alpha_1\lambda_1^t$ and $|\tau(t)| \leq \alpha_2\lambda_1^t$.

Bias errors

Observe from (15) that $G_{k+r} - G_k$ can be written as

$$G_{k+r} - G_k = \sum_{t=1}^{\infty} g(t)\Omega_k^t[\Omega_r^t - 1] \quad (22)$$

Expanding $\Omega_r^t = e^{-j\omega_r t}$ around $r = 0$, we can write

$$\Omega_r^t - 1 = -j\omega_t r + \frac{1}{2}(-j\omega_t)^2 r^2 + \frac{1}{3!}(-j\omega_t)^3 r^3 + \dots \quad (23)$$

Inserting this expression into (22) leads to (9) where

$$g_s(k) \triangleq \sum_{t=1}^{\infty} \frac{1}{s!} g(t)\Omega_k^t(-j\omega_t)^s. \quad (24)$$

Similarly it follows from (10) and (16) that

$$t_s(k) \triangleq \sum_{t=1}^{N-1} \frac{1}{s!} \tau(t)\Omega_k^t(-j\omega_t)^s. \quad (25)$$

The approximation error for the LS problem in the LPM is due to the terms neglected in the second line of (11), which we call E_k^{LPM}. Using (24) and (25) it can be expressed as:

$$E_k^{LPM} = \left[\sum_{s=R+1}^{\infty} \frac{1}{s!} \left(\sum_{t=1}^{\infty} g(t)\Omega_k^t(-j\omega_t)^s\right) r^s\right] U_{k+r} + \frac{1}{\sqrt{N}} \sum_{s=R+1}^{\infty} t_s(k) r^s \quad (26)$$

The approximation error in the LS problem of TRIMM is due to the terms in the last line of (17) which we denote by E_k^{TRIMM}. Replacing $\phi_t(\omega_{k+r}, \omega_k)$ by $\Omega_k^t[\Omega_r^t - 1]$ and using (23), it can be expressed as

$$E_k^{TRIMM} = \left[\sum_{s=1}^{\infty} \frac{1}{s!} \left(\sum_{t=n_1}^{\infty} g(t)\Omega_k^t (-j\omega_t)^s \right) r^s \right] U_{k+r} + \frac{1}{\sqrt{N}} \sum_{t=n_2}^{N-1} \tau(t) e^{-j\omega_{k+r}t} \quad (27)$$

It is interesting to compare the approximation errors for LPM and TRIMM. The sum in the first term of (26) starts with the term $s = R+1$ while it starts with $s = 1$ in (27); on the other hand the coefficients of r^s in (27) start at $t = n_2$ instead of $t = 1$ in (26). Taking into account that the size of the window $[0, r]$ is expressed in the frequency domain as $\frac{r}{N} f_s$, these errors can be approximated by

$$E_k^{LPM} = G_k^{(R+1)} O((\frac{r}{N})^{(R+1)}) + T_k^{(R+1)} O((\frac{r}{N})^{(R+3/2)})$$
$$E_k^{TRIMM} = \lambda_1^{n_1} O(\frac{r}{N}) + \lambda_1^{n_2} (\frac{1-\lambda_1^{N-n_2}}{1-\lambda_1}) O(N^{-1/2}) \quad (28)$$

where λ_1 is the absolute value of the largest eigenvalue of A, $G_k^{(R+1)}$ is the $(R+1)$-th order derivative of G_k, and similarly for $T_k^{(R+1)}$. It can be shown using the methods of (Pintelon et al. (2010)) that these approximation errors induce the following bias errors:

$$Bias(\hat{\theta}_k^{LPM}) = G_k^{(R+1)} O((\frac{L}{N})^{(R+1)}) + T_k^{(R+1)} O((\frac{L}{N})^{(R+2)})$$
$$Bias(\hat{\theta}_k^{TRIMM}) = \lambda_1^{n_1} O(\frac{L}{N}) + \lambda_1^{n_2} (\frac{1-\lambda_1^{N-n_2}}{1-\lambda_1}) O(N^{-1}) \quad (29)$$

Variance errors

In the LPM, the parameter vector θ_k of size $2(R+1)$ is estimated in the local windows using $2L+1$ data points, while in the TRIMM method the parameter vector θ of size $N + n_1 + n_2$ is estimated using all N data. As a result the variances for the two methods due to the noise V can be approximated as follows:

$$Var(\hat{\theta}_k^{LPM}) = O(\frac{2(R+1)}{2L+1}) \sigma_V^2(k)$$
$$Var(\hat{\theta}^{TRIMM}) = O(\frac{N+n_1+n_2}{(2L+1)N}) \sigma_V^2 \quad (30)$$

where $\sigma_V^2(k)$ is the noise spectrum at frequency Ω_k and σ_V^2 is the variance of the noise $v(t)$. Here the interpolation error due to the variation of the noise variance in the local window $[k-L, k+L]$ has been neglected.

5. COMPARATIVE ANALYSIS

We make a number of observations concerning the LPM and TRIMM methods, and we discuss the role of their respective design parameters. These observations are based on the theoretical analysis of the previous section but also on extensive Monte Carlo simulations in which the two methods were applied. We shall present some simulations in the next section to support these observations.

Observations on the LPM method

- The design parameters are the width $2L$ of the local window and the degree R of the polynomials.
- In the LPM, the parameters are locally defined and they are estimated in each of the local frequency windows. The bias increases with the width $2L+1$ of the window and it decreases with the degree R of the local polynomials. However, these two parameters cannot be chosen independently: in order to have a full rank set of equations for the LS problem $L \geq R+1$ is required. The smallest interpolation error is obtained for $L = R+1$.
- The variance error is inversely proportional to the window width L since $2L+1$ data are available for each local LS problem. Thus, if the noise dominates, one should choose a larger window width. However, the bias in the estimation of the noise spectrum increases with the window width: see Section 6. Hence, if the estimate of the noise spectrum is the main goal, one should refrain from taking L too large.
- The speed of computation is essentially determined by the size of the LS problem; since the LPM solves LS problems of size $2L+1$, the method is fast.

Observations on the TRIMM method

- The design parameters are the size $2L+1$ of the window used for interpolation, the length n_1 of the Finite Impulse Response (FIR) used for the approximation of $G_{k+r} - G_k$, and the length n_2 of the FIR used for the approximation of T_k.
- The parameters estimated are the desired FRF G_k as well as a set of parameters $g(1), \ldots, g(n_1)$; $\tau(0), \ldots, \tau(n_2-1)$ that are common to all equations of the LS problem.
- Without any windowing there would be fewer equations than unknowns. Assuming that $N > \frac{1}{4}(n_1+n_2)$, we observe that choosing $L = 1$ already gives more equations than unknowns. Increasing L above this value decreases the variance, as shown in (30), but it increases the bias as shown in (29). In addition the computation time is directly proportional to L.
- When n_1 and n_2 increase, the bias error decreases but the variance error increases. Thus, n_1 and n_2 should be chosen larger when the noise variance is small, but smaller when the noise variance is large and the variance errors dominate the bias errors.
- The parameters $g(1), \ldots, g(n_1)$ are only needed to produce approximations of $G_{k+r} - G_k$ in the window around Ω_k. If this window is chosen narrow (i.e. L small) then G_{k+r} is close to G_k and the approximation need not be as precise as when L is large, so that n_1 can also be chosen small. Simulations show that for a system $G(q)$ with a slowly decaying impulse response $g(t)$ the TRIMM method performs well with a small L and a choice of n_1 that is much smaller than the length of the impulse response.
- Computing $\{G_k, k = 1, \ldots, N\}$ requires the solution of one large LS problem with $2N(2L+1)$ equations and $2N + n_1 + n_2$ real unknowns; the computational time may be quite large. Since the computation time is proportional to L, this favors the choice of a small L when that is possible.

- Unlike the LPM which is designed to work with a small local window, the TRIMM method can also be used with a window of length $L = N/2$. This special case has been analyzed in Hägg and Hjalmarsson (2012), where it is also shown that the other extreme choice, $L = 0$, corresponds to the well-known Empirical Transfer Function Estimate Ljung (1999).

Comparative observations

- For some applications, the estimated FRF is mainly used as a tool for the estimation of related quantities, such as the estimate of the noise variance (see Section 6 for details). In such case, it is important that the estimate of the FRF has the smallest possible bias, and the LPM estimate may be preferred.
- On the other hand, if the goal is to get an estimate of the FRF with the smallest Mean Square Error, then the TRIMM method will typically perform better as soon as the noise variance becomes significant. For large N the ratio between the variance of the estimates is $Var(\hat{\theta}_k^{LPM}) \approx 2(R+1)Var(\hat{\theta}^{TRIMM})$.

6. ESTIMATION OF THE NOISE SPECTRUM

The estimation of the FRF $G(e^{j\omega})$ is often only an intermediate step that allows one to compute other quantities of interest. One such important application is the estimation of the noise spectrum $\sigma_V^2(k)$ as a function of frequency which can serve as a frequency weighting in a subsequent parametric estimate of the input-output transfer function $G(q)$. This application has been extensively discussed in (Schoukens et al. (2011, 2012).

For both the LPM and TRIMM methods, the estimate of the noise spectrum is obtained by computing the residuals $\hat{V}_{k+r}$ in a local frequency window of size $2M+1$ around Ω_k. They are computed by inverting the second relation in (8) in which G_k, $G_{k+r} - G_k$ and T_{k+r} are replaced by their estimates. The power spectrum at frequency Ω_k can then be obtained as

$$\hat{\sigma}_V^2(k) = \frac{1}{q} \sum_{r=-M}^{M} |\hat{V}_{k+r}|^2 \tag{31}$$

where q depends on the method used. The following observations can be made about this estimate.

- The noise power estimate at frequency Ω_k is contructed for both methods as the mean of the square of the residuals at the neighbouring frequencies, i.e. in the local window $[k-M, k+M]$. This estimate will have a bias that is caused by the variation of $|H_0(\Omega)|^2$ over this interval as well as the bias on the estimated parameters, which have been studied in Section 4.
- In the LPM the estimate of $\sigma_V^2(k)$ takes the form (31) with $M = L$, where $2L$ is the width of the local bandwidth in which the parameters of G_k, $G_{k+r} - G_k$ and T_{k+r} have been estimated.
- In the TRIMM these parameters have been estimated using global parameters and on the basis of the whole data record of N data. Since the design parameter L can sometimes be taken small, as argued in Section 5, the mean value should be computed over a frequency window that may be larger than the window $[k-L, k+L]$ used for the LS problem. In the simulations of the next section we have taken M equal to the size L of the local window used in the LPM in order to have a fair comparison of these two estimates.

7. SIMULATIONS

In this section we illustrate the behaviour of the LPM and TRIMM methods by Monte-Carlo simulations. We consider the following "true" system:

$$y(t) = \frac{0.1943 + 0.3885q^{-1} + 0.1943q^{-2}}{1 + 0.7125q^{-1} + 0.7449q^{-2}} u(t) \quad + \\ \frac{0.0389+0.0837q^{-1}+0.1016q^{-2}+0.0597q^{-3}+0.0187q^{-4}}{1+1.5281q^{-1}+2.2864q^{-2}+1.2918q^{-3}+0.7154q^{-4}} e(t)$$

The input signal is a coloured noise generated by

$$u(t) = \frac{0.5276q^3 + 1.583q^2 + 1.583q + 0.5276}{q^3 + 1.76q^2 + 1.183q + 0.2781} w(t),$$

where $w(t)$ is a white noise excitation signal with unit variance. This system and noise model have been chosen because the signal to noise ratio varies greatly around the resonance peak and beyond, which will illustrate the dependence of both methods on the signal to noise ratio.

We have performed 100 Monte Carlo simulations with 2500 data each from which the first 500 data have been removed, leaving 2000 useful I/O data for the estimation, for a range of choices of the design parameters. Figure 1 shows the true $G(\Omega_k)$ as well as the RMS error on the FRF estimate of G for the LPM and TRIMM estimates obtained with $\sigma_e^2 = 0.04$ and with the following design parameters:
- for LPM: $R = 3$ and $L = 6$
- for TRIMM: $n_1 = n_2 = 10, L = 5, \beta = 0.01$

Figure 2 shows the absolute value of the error between the true noise spectrum and the estimate obtained by the two methods under the same conditions and with the same design parameters. The TRIMM method gives a better

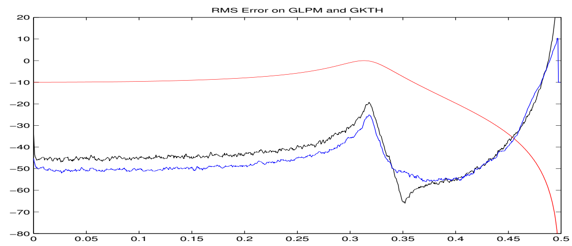

Fig. 1. FRF of G (red), RMS error on $\hat{G}^{LPM}$ (black) and $\hat{G}^{TRIMM}$ (blue)

estimate of G over much of the frequency range; the LPM estimate is better only where the signal to noise ratio is large. On the other hand, the LPM gives a better estimate of the noise spectrum due to its smaller bias error.

The next table gives the mean value over the whole frequency range of the MSE of $\hat{G}$ and of the error on $\hat{\sigma}_V^2(k)$ for a range of design values, with $\sigma_e^2 = 0.0001$ and with $\beta = 0.001$. For TRIMM, the design parameters n_1 and n_2 were taken identical and are denoted n in the tables below.

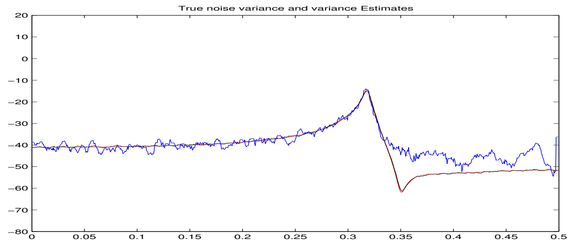

Fig. 2. True noise spectrum (red) and $\hat{\sigma}_V^2(k)$ for LPM (black) and TRIMM (blue)

	L	MSE $\hat{G}$	Mean Err $\hat{\sigma}_V^2(k)$
LPM			
$R = 1$	5	0.00049562	4.2004e-07
$R = 1$	10	0.00121750	5.6839e-06
$R = 2$	5	0.00052861	1.1893e-07
$R = 2$	10	0.00032632	3.6241e-07
$R = 3$	5	0.00064776	1.2098e-07
$R = 3$	10	0.00034635	3.4201e-07
TRIMM			
$n = 4$	2	0.00320860	0.00018504
$n = 4$	10	0.00415630	0.00042724
$n = 50$	2	0.00085536	3.3143e-06
$n = 50$	10	0.00051315	8.3850e-06
$n = 80$	2	0.00097818	1.8631e-05
$n = 80$	10	0.00067632	2.8348e-06

The next table gives the results for the same design variables, but now with $\sigma_e^2 = 0.01$.

	L	MSE $\hat{G}$	Mean Err $\hat{\sigma}_V^2(k)$
LPM			
$R = 1$	5	0.0032674	1.2814e-05
$R = 1$	10	0.0026382	3.9703e-05
$R = 2$	5	0.0053735	8.665e-06
$R = 2$	10	0.0032151	3.3414e-05
$R = 3$	5	0.0064713	1.4554e-05
$R = 3$	10	0.0034249	3.1455e-05
TRIMM			
$n = 4$	2	0.0057070	0.00019851
$n = 4$	10	0.0046941	0.00041201
$n = 50$	2	0.0050661	8.1015e-05
$n = 50$	10	0.0029149	8.7244e-05
$n = 80$	2	0.0058321	9.9308e-05
$n = 80$	10	0.0037225	6.5780e-05

For LPM, except for the choice $R = 1, \sigma_e^2 = 0.0001$, the larger window gives better results than the smaller window for the estimation of G; this probably means that the variance error for $\hat{G}^{LPM}$ dominates the bias error in all cases. For TRIMM, with low noise variance the best results for $\hat{G}$ are obtained with $n = 50$; for $n = 80$ the reduction in bias error is offset by an increase in variance error. With $\sigma_e^2 = 0.01$ the best results for $\hat{G}^{TRIMM}$ are again obtained with $n = 50$ but the differences are small between the different cases; however in call cases the larger window gives better results because the variance error dominates. Finally, for the estimation of the noise spectrum, LPM typically gives better results. We conjecture that this is because $\hat{G}^{LPM}$ has a smaller bias in the cases studied in these tables; this is the object of our present investigations.

8. CONCLUSIONS

We have compared two recent methods for the nonparametric estimation of the FRF from stationary input signals. Even though TRIMM was inspired by LPM, we have shown that they are based on the solution of two very different LS problems: LPM estimates parameters in a local frequency band, while TRIMM estimates global parameters using all data. As a result, the statistical properties of these two methods are quite different, and we have shown their relationships. The performance of both methods rely heavily on an adequate choice of their design parameters: with low noise, the bias errors dominate and the design parameters must be chosen to minimize these bias errors; with high noise, the design parameters must be chosen to minimize the dominating variance errors. Future work should lead to an automated selection of the design parameters.

REFERENCES

Agüero, J.C., Yuz, J., Goodwin, G., and Delgado, R. (2008). On the equivalence of time and frequency domain maximum likelihood estimation. *Automatica*, 46(2), 260–270.

Gevers, M., Pintelon, R., and Schoukens, J. (2011). The Local Polynomial Method for nonparametric system identification: improvements and experimentation. In *accepted for presentation at CDC-ECC 2011*. Orlando, USA.

Hägg, P. and Hjalmarsson, H. (2012). Non-parametric fequency function estimation using Transient Impulse Response Modelling. In *submitted to IFAC Symp SYSID 2012*. Brussels, Belgium.

Hägg, P., Hjalmarsson, H., and Wahlberg, B. (2011). A least squares approach to direct frequency response estimation. In *accepted for presentation at CDC-ECC 2011 CDC-ECC 2011*. Orlando, Florida, USA.

Ljung, L. (1999). *System Identification: Theory for the User, 2nd Edition*. Prentice-Hall, Englewood Cliffs, NJ.

McKelvey, T. (2002). Frequency domain identification methods. *Circuits Systems Signal Processing*, 21, 39–55.

Pintelon, R. and Schoukens, J. (2001). *System Identification - A Frequency Domain Approach*. IEEE Press, Piscataway, New Jersey.

Pintelon, R., Schoukens, J., and Vandersteen, G. (1997). Frequency domain system identification using arbitrary signals. *IEEE Trans. Automatic Control*, 42, 1717–1720.

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

Schoukens, J., Rolain, Y., Vandersteen, G., and Pintelon, R. (2011). User-friendly Box-Jenkins identification using nonparametric noise models. In *accepted for presentation IEEE CDC and ECC 2011*. Orlando, USA.

Schoukens, J., Vandersteen, G., Barbé, K., and Pintelon, R. (2009). Nonparametric preprocessing in system identification: a powerful tool. *European Journal of Control*, 15(3-4), 260–274.

Schoukens, J., Vandersteen, G., Gevers, M., Pintelon, R., and Rolain, Y. (2012). User choices for nonparametric preprocessing in system identification. In *submitted to IFAC Symp SYSID 2012*.