The Local Polynomial Method for nonparametric system identification: improvements and experimentation

Michel Gevers, Rik Pintelon and Johan Schoukens

Abstract— The Local Polynomial Method (LPM) is a recently developed procedure for nonparametric estimation of the Frequency Response Function (FRF) of a linear system. Compared with other nonparametric FRF estimates based on windowing techniques, it has proved to be remarkably efficient in reducing the leakage errors caused by the application of Fourier transform techniques to non periodic data. In this paper we propose a modification of the LPM that takes account explicitly of constraints between the coefficients of the polynomials at neighbouring frequencies. This new variant contributes a new and significant reduction in the Mean Square Error of the FRF estimates. We also discuss the effects of the various design parameters on the accuracy of the estimates.

I. INTRODUCTION

This paper addresses the nonparametric estimation of the Frequency Response Function (FRF) of a linear dynamic system from input-output measurements and it proposes a new method for the reduction of *leakage errors* that are inherent in the computation of frequency response estimates. The inputs are assumed known but not necessarily periodic, while the outputs are perturbed by additive quasistationary noise. There are many good reasons for the estimation of these nonparametric quantities. The obtention of a high quality estimate of the FRF can be of independent interest, yielding a completely nonparametric approach to the identification problem. The availability of this FRF estimate can give a preliminary idea of the complexity of the system, and can be used as a benchmark to test and validate parametric model estimates. The advantage of such nonparametric approach is that it avoids the difficult problem of structure selection, which is really the hardest part of the identification problem. The disadvantage is that for many applications (e.g. control design, prediction, etc) a finite dimensional parametric model is much more practical, if not essential.

Even if the identification of a parametric model is the final goal, it has been shown that the estimation of a nonparametric model can be a very useful first step in a parametric estimation procedure because it allows one to compute a prior nonparametric estimate of the noise spectrum, which can significantly improve the quality of the ensuing parametric input-output model estimate, as well as avoid the possibility of getting trapped in local minima during the minimization procedure of the parametric identification criterion [1].

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. These leakage errors can be significantly reduced by the application of periodic input signals; however, this is not always practically possible. As a result, 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. The main contribution of this paper is a novel technique that leads to a signficant reduction of these leakage errors.

Until the 80's, leakage errors on FRF-measurements were studied at the input and output signal level, without considering the linear system relation between the input and output [2], [4]. In FRF-measurements, the leakage errors are due to unknown past inputs and missing future outputs. Both effects are highly structured, and as a result the leakage errors can be represented in the frequency domain as rational functions added to the output [5], [6], [7], [8], [9], [10]. This key observation that led to the adoption of windowing techniques that are based on a differentiation of the input and output signals around a central frequency, thereby reducing the effect of the smooth leakage term [11]. One of the drawbacks of these windowing techniques is that by reducing the leakage errors the window introduces an interpolation error.

Recently a new method, called *the Local Polynomial Method (LPM)*, has been introduced to estimate the FRF and the power spectrum of the disturbing noise. Using a Taylor series expansion, the transfer function and the leakage term are expressed in a narrow window around some central frequency as two local polynomial models. The coefficients of this local polynomial are estimated by Least Squares using the input and output data over the narrow window around the frequency of interest [12], [13], [14]. The least squares estimate of the local polynomial coefficients delivers an estimate of the FRF at the central frequency from which the leakage errors and the transient errors have been substantially reduced. The LPM outperforms the classical methods: the leakage errors are reduced with several orders of magnitude, depending upon the system and the record length, at the expense of an increase in the computation time.

In the LPM, the LS estimation of the polynomial coef-

M. Gevers is with CESAME, Université catholique de Louvain, B1348 Louvain la Neuve, Belgium, and with Dept ELEC, Vrije Universiteit Brussel, Pleinlaan 2, B1050 Brussels, Belgium `Michel.Gevers@uclouvain.be`

R. Pintelon and J. Schoukens are with Dept ELEC, Vrije Universiteit Brussel, Pleinlaan 2, B1050 Brussels, Belgium `Rik.Pintelon@vub.ac.be` and `Johan.Schoukens@vub.ac.be`

ficients is applied locally at every frequency using data in narrow windows around each frequency. Since neighbouring intervals overlap, some of the estimated polynomial coefficients appear in common in the LS problems formulated over neighbouring intervals. This is not taken into account in the solution of the standard Local Polynomial Method, since the LS estimation problem solved over one window of input-output data is solved independently of the LS problem solved over neighbouring windows.

The contribution of this paper is to propose a modification of the LPM that takes account of the appearance of identical polynomial coefficients in neighbouring windows. We call this new method the *Local Polynomial Method with Constraints (LPMC)*. Introducing these constraints reduces the variance error of the estimated coefficients since more information is used for the estimation of each coefficient. However, the bias error is increased since the coefficients estimated over one frequency interval influence those estimated on neighbouring intervals. Thus, a proper trade-off is required, which can be achieved by adding the error on the constraints to the LS criterion of LPM, with a proper scaling between the two terms of the modified LS criterion. We shall present the new LPMC, illustrate the benefits in terms of accuracy of the estimated FRF, and discuss the influence of its design parameters through some simulated examples.

In Section II we present the "classical LPM". In Section III we show how the constraints between neighbouring parameter vectors can be introduced, while in Section IV we compare the classical LPM with the new constrained version on some simulated examples, which will illustrate the trade-offs mentioned above. In Section V we explain how these first results on the LPMC pave the way for further possible improvements of the Local Polynomial Method.

II. THE LOCAL POLYNOMIAL METHOD

We start by presenting the 'classical' Local Polynomial Method, first published in [14]. A complete analysis of the LPM for the multiple-input multiple-output (MIMO) case can be found in [12], [13]. Here we focus on the SISO case for pedagogical reasons; the extension to MIMO systems is straightforward but tedious.

Thus, consider a linear discrete time single-input single-output (SISO) system $G_0(q)$ 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 [3] so that asymptotic analysis can be used for the computation of the Mean Square Error. In particular, $v(t)$ can be modeled as the output of a white-noise driven filter. Thus the input-output system can be represented as

$$y(t) = G_0(q)u(t) + v(t) = G_0(q)u(t) + H_0(q)e(t) \quad (1)$$

where q^{-1} is the backward shift operator, $G_0(q)$ and $H_0(q)$ are causal rational functions of q, and $e(t)$ is zero mean white noise with variance σ_e^2. This input-output representation assumes an infinite data record of input and output signals, for $t = -\infty, \ldots, N-1$. For a finite record $t = 0, \ldots, N-1$

- as it is in practical applications - this equation has to be modified to take account of the initial condition (or transient) terms t_G and t_H due to the action of the transfer function G_0 and the noise model H_0, leading to:

$$y(t) = G_0(q)u(t) + t_G(t) + H_0(q)e(t) + t_H(t). \quad (2)$$

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 (2) is obtained:

$$Y(k) = G_0(\Omega_k)U(k) + T_G(\Omega_k) + H_0(\Omega_k)E(k) + T_H(\Omega_k) \quad (3)$$

where the index k points to the frequency kf_s/N with f_s the sampling frequency, and $\Omega_k = e^{j2\pi k/N}$. The contributions U, E, Y in (3) are an $O(N^0)$, while the transient terms T_G and T_H are an $O(N^{-1/2})$, where $X = O(N^p)$ means that $\lim_{N^p \to 0} |\frac{X}{N^p}| < \infty$.

It is important to understand that (3) is an exact relation [8], [9], [15], [16]. The transient terms $t_G(t)$ and $t_H(t)$ are rational forms in q^{-1} applied to a delta-input, while the leakage terms T_G and T_H are rational forms in z^{-1}, and hence smooth functions of the frequency. For simplicity of notation we shall from now on rewrite (3) as

$$Y_k = G_kU_k + T_k + V_k, \quad (4)$$

where T_k denotes the sum of the plant and noise leakage errors and $V_k = H_0(q)E_k$. The basic idea of the LPM, based on the smoothness of the transfer function G_0 and of the transient term T as functions of frequency, is to approximate these functions in a narrow frequency band around a central frequency Ω_k by a complex polynomial. The complex polynomial parameters are estimated from the experimental data collected in this frequency band. Next G_k, at the central frequency Ω_k, is retrieved from this local polynomial model as the estimate of the FRF at that frequency Ω_k.

By the smoothness of G_0 and T, the following polynomial representation holds for the frequency lines $k+r$, with $r = 0, \pm 1, \ldots, \pm n$:

$$G_{k+r} = G_k + \sum_{s=1}^{R} g_s(k)r^s + O\left(\left(\frac{r}{N}\right)^{(R+1)}\right) \quad (5)$$

$$T_{k+r} = T_k + \sum_{s=1}^{R} t_s(k)r^s + N^{-\frac{1}{2}}O\left(\left(\frac{r}{N}\right)^{(R+1)}\right) \quad (6)$$

We can now collect G_k, T_k and all polynomial coefficients into a $2(R+1)$-vector of unknown complex coefficients defined as

$$\theta_k \triangleq [G_k\ g_1(k) \ldots g_R(k); T_k\ t_1(k) \ldots t_R(k)]^T, \quad (7)$$

where A^T denotes the transpose of A. Rewriting (4) at frequency Ω_{k+r} and substituting G_{k+r} and T_{k+r} by their expressions (5)-(6) while neglecting the remainder terms allows one to re-express Y_{k+r} as follows

$$Y_{k+r} = K(R, k+r)\theta_k + V_{k+r}, \quad \text{for } r = 0, \pm 1, \ldots, \pm n \quad (8)$$

where $K(R, k + r)$ is a $2(R + 1)$ row-vector that contains both structural information (the powers of r in the polynomial expansions (5)-(6)) and input signal information. We now collect the $2n + 1$ equations (8) obtained for $r = 0, \pm 1, \ldots, \pm n$ into one matrix equation by defining the $(2n + 1)$-vectors $\bar{Y}_{k,n}$ and $\bar{V}_{k,n}$ as follows:

$$\bar{Y}_{k,n} \triangleq [Y_{k-n}\ Y_{k-n+1}\ \ldots Y_k\ \ldots Y_{k+n-1}\ Y_{k+n}]^T$$
$$\bar{V}_{k,n} \triangleq [V_{k-n}\ V_{k-n+1}\ \ldots V_k\ \ldots V_{k+n-1}\ V_{k+n}]^T$$

This then leads to the following matrix version of (8):

$$\bar{Y}_{k,n} = K_{k,n}(R, \bar{U}_{k,n})\theta_k + \bar{V}_{k,n} \tag{9}$$

where the $2(n + 1)$-vector $\bar{U}_{k,n}$ is defined in the same way as $\bar{Y}_{k,n}$ and $\bar{V}_{k,n}$. The matrix $K_{k,n}(R, \bar{U}_{k,n})$ is a $2(n+1) \times 2(R + 1)$ matrix whose structure is entirely determined by the indices n and R and which contains the input signals U_{k+r} that appear in the input vector $\bar{U}_{k,n}$. In the standard LPM, the parameter estimate $\hat{\theta}_k$ is obtained by solving the following LS problem:

$$\min_{\theta_k} \left[\bar{Y}_{k,n} - K_{k,n}(R, \bar{U}_{k,n})\theta_k\right]^H \left[\bar{Y}_{k,n} - K_{k,n}(R, \bar{U}_{k,n})\theta_k\right] \tag{10}$$

where for any complex vector or matrix A, A^H denotes its Hermitian conjugate transpose. It follows from (7) that an estimate of the FRF at the frequency Ω_k is then obtained as the first component of the estimate $\hat{\theta}_k$: $\widehat{G}(\Omega_k) = \hat{\theta}_k(1)$. In order to get a full column rank matrix $K_{k,n}(R, \bar{U}_{k,n})$ the following condition is required between the number of spectral lines in the frequency window around Ω_k and the order of the polynomial approximation: $n \geq R + 1$. Taking a larger number of frequencies in the frequency window reduces the variance of the parameter estimate because the noise is averaged over a larger number of data, and the leakage error decreases with increasing R. On the other hand, the larger the window, the larger the bias error (or interpolation error) caused by the fact that the transfer function varies over the interval. The smallest interpolation error is obtained for $n = R + 1$. A detailed error analysis of the LPM is presented in [12] where this bias-variance trade-off is discussed. In practice, the LPM is mostly used with polynomials of degree two only, i.e. $R = 1$ or 2, which offers a good compromise between leakage error and interpolation error.

III. LPM WITH CONSTRAINTS

In the LPM described above each parameter vector θ_k is estimated using local data U_{k+r}, Y_{k+r} in a frequency window defined by $r = 0, \pm 1, \ldots, \pm n$. As a result, for $r \leq n$, the estimates, θ_k and θ_{k+r} are computed by solving two separate Least Squares problems that use data which partly overlap. This means that these estimates are correlated because the data that are used in the two LS problems are correlated. But in addition, for $|r| \leq n$, the parameters in θ_k and θ_{k+r} are not independent, since they are related by the polynomial constraints (5)-(6). Indeed, it follows from (5)-(6) that up to the remainder terms appearing in these expressions, the following relationships exist between θ_{k+r} and θ_k, for $r = 0, \pm 1, \ldots, \pm n$:

$$G_{k+r} = \theta_{k+r}(1) = \theta_k(1) + \sum_{s=1}^{R} \theta_k(s + 1)r^s \tag{11}$$

$$T_{k+r} = \theta_{k+r}(R + 2) = \theta_k(R + 2) + \sum_{s=1}^{R} \theta_k(s + R + 2)r^s$$

In the standard LPM these relationships have not been exploited. The contribution of this paper is to explore ways in which these constraints can be exploited to decrease the Mean Square Error (MSE) in the estimates of the parameters θ_k, $k = 1, \ldots, N$, and in particular the MSE of the FRF estimates $\widehat{G}(\Omega_k)$, which are the first component of these vectors θ_k. Observe that (11) represents $4n$ constraints on the $2(R + 1)$-parameter vector θ_k, with $n \geq R + 1$. If the θ_{k+r}, $r = \pm 1, \ldots, \pm n$, were considered as known data in the estimation problem of θ_k, then θ_k would be entirely determined by this set of equations. Thus adding the constraints (11) to the LS problem (10) would lead to an overdetermined set of constraints on the solution θ_k. However, the θ_{k+r} are themselves the solution of a LS problem (10) that depends on the data $\bar{U}_{k+r}, \bar{Y}_{k+r}$. Thus, in the formulation of a modified LPM that takes these constraints into account one needs to find a compromise between "letting the data speak", and "letting the constraints speak".

A first idea would be to formulate one global optimization problem for $\{\theta_0, \ldots, \theta_{N-1}\}$ using all data $\{U_k, Y_k, k = 0, \ldots, N - 1\}$ and taking the constraints into account. This would significantly increase the computational load and the attractivity of the *local* polynomial approach would be lost.

The alternative proposed in this paper is to keep the advantage of the local computation of θ_k based on data $\bar{U}_{k,n}$ and $\bar{Y}_{k,n}$ in a narrow frequency band around Ω_k, but to turn the local LS criterion (10) into a multiobjective LS criterion by adding a penalty on the mismatch between left and right hand side of the constraints (11). In order to arrive at a feasible implementation of this idea, we first analyze the constraints.

A. Analysis of the constraints

We first rewrite the constraints (11) in matrix form. In order to do so, we introduce the following matrices, for positive integers R and n.

$$M(R, -n) \triangleq \begin{pmatrix} 1 & -n & (-n)^2 & \ldots & (-n)^R \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & -2 & (-2)^2 & \ldots & (-2)^R \\ 1 & -1 & (-1)^2 & \ldots & (-1)^R \end{pmatrix}$$

$$M(R, n) \triangleq \begin{pmatrix} 1 & 1 & 1 & \ldots & 1 \\ 1 & 2 & 2^2 & \ldots & 2^R \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & n & n^2 & \ldots & n^R \end{pmatrix} \tag{12}$$

Notice that the matrices $M(R, -n)$ and $M(R, n)$ are Vandermonde matrices; therefore, for $n \geq R+1$, these matrices have rank $R+1$. The constraints (11) can be written in matrix form as follows.

$$\begin{bmatrix} M(R,n) & 0 \\ 0 & M(R,n) \\ M(R,-n) & 0 \\ 0 & M(R,-n) \end{bmatrix} \theta_k = \begin{bmatrix} \Psi_k \\ \dots \\ \Phi_k \end{bmatrix}, \qquad (13)$$

or, equivalently, as

$$\bar{M}\theta_k = \Xi_k \qquad (14)$$

where
$\Psi_k \triangleq [\theta_{k+1}(1) \dots \theta_{k+n}(1); \theta_{k+1}(R+2) \dots \theta_{k+n}(R+2)]^T$,
$\Phi_k \triangleq [\theta_{k-n}(1) \dots \theta_{k-1}(1); \theta_{k-n}(R+2) \dots \theta_{k-1}(R+2)]^T$,
$\Xi_k = [\Psi_k^T \ \Phi_k^T]^T$ and $\bar{M}$ is the block matrix on the left hand side of (13).

The constraints (13) split up into two subsets of constraints. The top half relates the parameter vector θ_k to parameter vectors at higher frequencies in the frequency window, while the bottom half relates θ_k to parameters at lower frequencies. In addition, for $n \geq R+1$ each of these two sets of equations is an overdetermined set in that it contains $2n$ equations for $R+1$ unknowns. $M(R,-n)$ and $M(R,n)$ have full column rank and hence the top or the bottom half of (14) is sufficient to fully determine the vector θ_k. The difficulty is that the parameters on the right hand side of (13) are unknown. One can think of two ways to overcome this difficulty.

The first is a *recursive* (in k) solution to the local LS problem, starting from $k = 0$ and going up in frequencies, say, and applying one-sided constraints only, i.e. the solution for θ_k is obtained using only the bottom half of the constraints, namely those containing $\theta_{k-r}(1)$ and $\theta_{k-r}(R + 2)$ for $r = 1, \dots, n$. In order to initialize the recursions, the first n vectors, $\theta_0, \dots, \theta_{n-1}$, can be set at the estimates obtained by the standard LPM. We have applied this recursive procedure to a range of systems and found that, in each case, it performed worse than the two-step procedure that we now describe.

B. Two-step implementation of the LPM with constraints
In the first step, θ_k is estimated for all $k = 0, \dots, N-1$ using the standard LPM, yielding estimates that we denote $\hat{\theta}_k^{(LPM)}, k = 0, \dots, N-1$. These estimates are then used in the right hand side of the constraints (13), alternatively (14), yielding the constraints

$$\bar{M}\theta_k = \widehat{\Xi}_k, \qquad (15)$$

which are now entirely feasible. As explained above, solving the LS problem (10) subject to (15) would yield an estimate θ_k that is entirely determined by the constraints, leaving no degrees of freedom for the minimization of the LS criterion. Instead, a penalty is added to the LS criterion (10) which represents the 2-norm of the error on the constraints (15). The modified LPM is thus obtained by solving, for $k =$

$0, \dots, N-1$, the following multiobjective LS problem

$$\min_{\theta_k} \left\{ \left[\bar{Y}_{k,n} - K_{k,n}(R, \bar{U}_{k,n})\theta_k \right]^H \left[\bar{Y}_{k,n} - K_{k,n}(R, \bar{U}_{k,n})\theta_k \right] \right.$$
$$\left. + \lambda\Phi_u(\Omega_k) \left(\bar{M}\theta_k - \widehat{\Xi}_k \right)^H \left(\bar{M}\theta_k - \widehat{\Xi}_k \right) \right\}. \qquad (16)$$

The weighting factor λ allows one to tune the relative importance of the constraint mismatch versus the error fit to the measured data. Increasing λ will impose more smoothness in the estimated FRF, thus decreasing the variance error at the expense of an increased bias error. The scaling by the spectrum of the input, $\Phi_u(\Omega_k)$, ensures that the relative importance of the two terms of (16) are independent of the power of the input signal since the first term is proportional to $\Phi_u(\Omega_k)$. The automatic tuning of λ is an important issue that is out of the scope of this paper.

IV. LPM WITH CONSTRAINTS AT WORK

In this section we illustrate the behaviour of the LPM with constraints, denoted LPMC, by presenting the results of Monte-Carlo simulations obtained on two different input-output systems and noise models. In particular we examine the role of the three design parameters: the degree R of the polynomial approximation, the width $2n$ of the frequency window over which the local estimates are computed, and the weighting λ that accounts for the tradeoff between data information and structural information, i.e. the knowledge that the coefficients of neighbouring parameter vectors are related by the polynomial constraints.

Example 1
We first consider a system with the following Box-Jenkins (BJ) structure:

$$\begin{aligned} y(t) &= 0.1943\frac{(q+1)^2}{q^2 + 0.7125q + 0.7449}u(t) \\ &+ 0.1084\frac{(q+1)^2}{q^2 - 0.8773q + 0.3111}e(t) \end{aligned}$$

where $e(t)$ is zero mean white noise with standard deviation σ_e and where the input signal $u(t)$ is a colored noise generated by

$$u(t) = 0.5276\frac{(q+1)^3}{q^3 + 1.7600q^2 + 1.1829q + 0.2781}w(t) \qquad (17)$$

with $w(t)$ white noise of standard deviation $\sigma_w = 1$.

200 Monte-Carlo runs are used to produce 200 sets of input-output data, each of length 8192, from which the first 1024 are eliminated in order to remove transient effects of the simulation; thus, each data set contains 7168 input-output data. The LPMC is applied on each of these 200 runs to estimate the FRF $G_0(\Omega_k)$ for $k = 0, \dots, 7168$, obtained with the standard LPM, denoted $\widehat{G}^{(LPM)}(\Omega_k)$, and with the constrained estimate, denoted $\widehat{G}^{(LPMC)}(\Omega_k)$. The Mean Square Errors between these two estimates and the exact $G_0(\Omega_k)$ are computed and plotted as a function of frequency, in a log-scale. Finally, the average of these mean square errors over all frequencies are computed, because
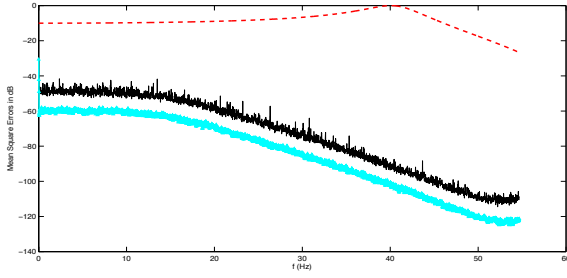
Fig. 1. BJ model. Top line (red dotted) = true FRF, middle line (black) = MSE on $\widehat{G}^{(LPM)}(\Omega_k)$, bottom line (cyan) = MSE on $\widehat{G}^{(LPMC)}(\Omega_k)$ as a function of frequency; all plots in dB. $R = 2, 2n = 6, \lambda = 0.05, \sigma_e = 0.05$

Fig. 3. BJ model. Top line (red dotted) = true FRF, middle line (black) = MSE on $\widehat{G}^{(LPM)}(\Omega_k)$, bottom line (cyan) = MSE on $\widehat{G}^{(LPMC)}(\Omega_k)$ as a function of frequency; all plots in dB. $R = 2, 2n = 6, \lambda = 0.05, \sigma_e = 0.15$

these numbers give a global indication of the quality of each of the two estimates.

Figure 1 shows the Mean Square Error obtained using the LPM and LPMC estimates on the BJ system described above, with a noise e with standard deviation $\sigma_e = 0.05$, for the following design choices: polynomial degree $R = 2$, local bandwidth $2n = 6$ and weighting $\lambda = 0.05$. The average MSE, over all frequencies, of the FRF estimates are as follows: for $\widehat{G}^{(LPM)}$: 0.00123, for $\widehat{G}^{(LPMC)}$: 0.00035. The signal to noise ratio for this first experiment, expressed as 10 times the logarithm of the input contribution to the output power spectrum $\Phi_y^{(u)}(\Omega_k)$ divided by the noise contribution $\Phi_v(\Omega_k)$, is presented as the top line in Figure 2; the other two lines represent, respectively, $\Phi_y^{(u)}(\Omega_k)$ and $\Phi_v(\Omega_k)$.
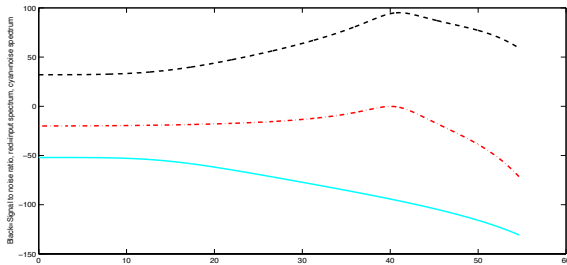

Fig. 2. BJ model. Top line (black dotted) = signal to noise ratio, middle line (red dash-dot) = output power due to input signal, bottom line (cyan full) = noise power on output as a function of frequency; all plots in dB. $R = 2, 2n = 6, \lambda = 0.05, \sigma_e = 0.05$

Figure 3 provides the same information as Figure 1 for the same system and with the same design parameters, but with an increased value of the white noise level e, i.e. $\sigma_e = 0.15$. The average MSE over all frequencies are, respectively, 0.0111 for $\widehat{G}^{(LPM)}$ and 0.0024 for $\widehat{G}^{(LPMC)}$. We observe that the superiority of the constrained LPM estimate over the classical one is even higher when the noise level is higher. The addition of constraints has a smoothing effect on the estimate, whose contribution is all the more important when the data are more noisy.

We now examine the effect of the polynomial degree and of the bandwidth. The third simulation is performed with the
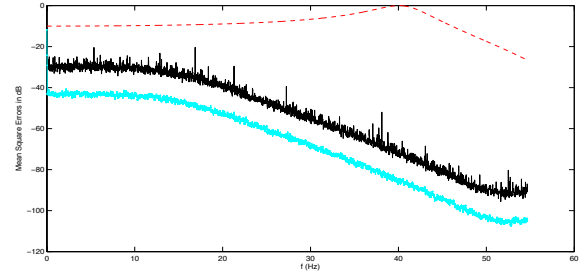
same system and design variables as the first, except that $R = 1$ and $2n = 4$. The MSE of the two estimates are presented in Figure 4. The figure shows that the superiority of the LPMC estimate over the LPM estimate is significantly reduced. The reason is that with lower R and n, the number of constraints and hence their impact is significantly reduced. To confirm this interpretation, we have multiplied the weighting factor λ by 10, i.e. $\lambda = 0.5$ in order to give more weight to the constraints. The results are shown in Figure 5, which shows that this increased penalty on the constraint mismatch leads to a much smaller MSE for $\widehat{G}^{(LPMC)}(\Omega_k)$.
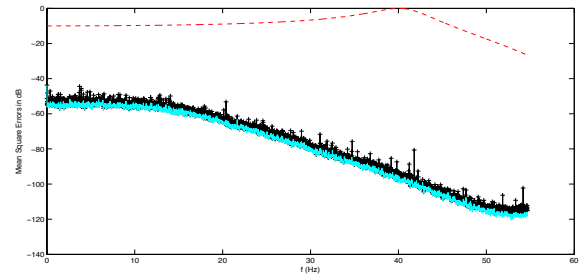

Fig. 4. BJ model. Top line (red dotted) = true FRF, middle line (black with crosses) = MSE on $\widehat{G}^{(LPM)}(\Omega_k)$, bottom line (cyan) = MSE on $\widehat{G}^{(LPMC)}(\Omega_k)$ as a function of frequency; all plots in dB. $R = 1, 2n = 4, \lambda = 0.05, \sigma_e = 0.05$
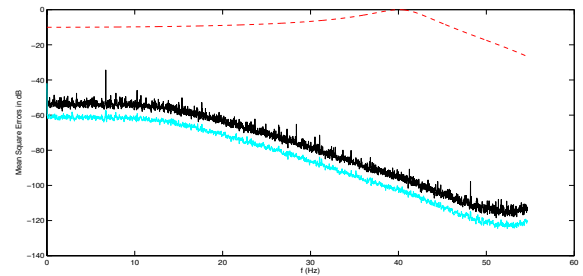

Fig. 5. BJ model. Top line (red dotted) = true FRF, middle line (black)= MSE on $\widehat{G}^{(LPM)}(\Omega_k)$, bottom line (cyan) = MSE on $\widehat{G}^{(LPMC)}(\Omega_k)$ as a function of frequency; all plots in dB. $R = 1, 2n = 4, \lambda = 0.5, \sigma_e = 0.05$

Example 2
We now consider an ARX system with the same input-output model as in example 1:

$$y(t) = 0.1943 \frac{(q+1)^2}{q^2 + 0.7125q + 0.7449} u(t)$$
$$+ \frac{1}{q^2 + 0.7125q + 0.7449} e(t)$$

where $e(t)$ is white noise with standard deviation σ_e and where the input signal $u(t)$ is now a white noise sequence with standard deviation $\sigma_u = 1$. We perform 200 Monte Carlo simulations as before, computing again the Mean Square Error between the true FRF and the estimates $\widehat{G}^{(LPM)}(\Omega_k)$ and $\widehat{G}^{(LPMC)}(\Omega_k)$, with the following design variables: $R = 2, n = 3, \lambda = 1$ and $\sigma_e = 0.05$. The average over all frequencies of the mean square errors on $\widehat{G}^{(LPM)}(\Omega_k)$ and on $\widehat{G}^{(LPMC)}(\Omega_k)$ are, respectively, 0.0111 and 0.0026, a ratio of improvement of more than 4 in favour of the new constrained LPM. The results are shown in Figure 6, and the signal to noise ratio is represented in Figure 7.
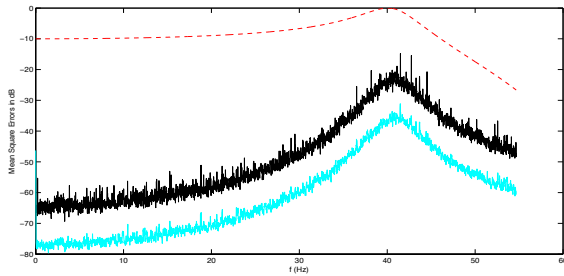

Fig. 6. ARX model. Top line (red dotted) = true FRF, middle line (black)= MSE on $\widehat{G}^{(LPM)}(\Omega_k)$, bottom line (cyan) = MSE on $\widehat{G}^{(LPMC)}(\Omega_k)$ as a function of frequency; all plots in dB. $R = 1, 2n = 4, \lambda = 0.5, \sigma_e = 0.05$
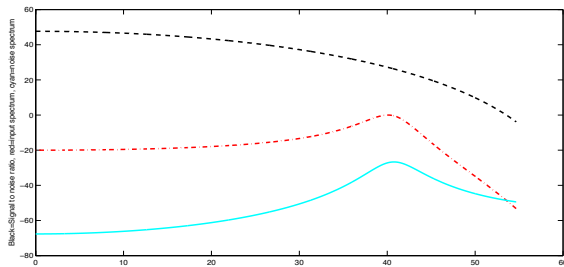

Fig. 7. ARX model. Top line (black dotted) = signal to noise ratio, middle line (red dash-dot) = output power due to input signal, bottom line (cyan full) = noise power on output as a function of frequency; all plots in dB. $R = 2, 2n = 6, \lambda = 0.05, \sigma_e = 0.05$

V. CONCLUSIONS AND FUTURE WORK

We have proposed a modification to the Local Polynomial Method for the computation of a non parametric estimate of the FRF of a linear time-invariant system. The modification consists of applying to the estimated parameter vectors

the constraints that exist between estimates at neighbouring frequencies; the classical LPM was treating these parameter vectors as independent. The constrained estimates have been shown to yield estimates with significantly smaller mean square errors. The gain in accuracy that can be made depends on the choice of a small number of design parameters, whose impact we have exhibited. Our next goal is to provide an almost automatic procedure for the selection of these design parameters, based on the collected data. In addition, we plan to compare the performance of this new Constrained LPM with a one-step procedure, in which the full set of parameters θ_k over the whole frequency range would be computed as the solution of one large LS problem subject to a 2-norm penalty on the constraint errors.

REFERENCES

[1] J. Schoukens, Y. Rolain, G. Vandersteen and R. Pintelon, *User friendly Box-Jenkins identification using nonparametric noise models.* Submitted for presentation at 50th IEEE Conference on Decision Control and 10-th European Control Conference, Orlando-Florida (USA), Dec. 12-15, 2011.
[2] J.S. Bendat and A.G. Piersol, *Engineering Applications of Correlations and Spectral Analysis.* Wiley, New York (1980).
[3] L. Ljung, *System Identification: Theory for the User, 2nd Edition.* Prentice-Hall,Englewood Cliffs, NJ (1999).
[4] P.E. Wellstead, *Non-parametric methods of system identification.* Automatica, Vol. 17, (1981), pp. 55-69.
[5] L. Rabiner and J. Allen, *On the implementation of a short-time spectral analysis method for system identification.* IEEE Trans. Acoustics, Speech, and Signal Processing ASSP, Vol. 28, (1980), pp. 69-78.
[6] J.L. Douce and L. Balmer, *Transient effects in spectrum estimation.* IEE Proc. Pt. D, Vol. 132, (1985), pp. 25-29.
[7] J.L. Douce and L. Balmer, *Statistics of frequency-response estimates.* IEE Proc. Pt. D, Vol. 137, (1990), pp. 290-296.
[8] R. Pintelon, J. Schoukens, and G. Vandersteen, *Frequency domain system identification using arbitrary signals.* IEEE Trans. Automatic Control, Vol. 42, (1997), pp. 1717-1720.
[9] T. McKelvey, *Frequency domain identification methods.* Circuits Systems Signal Processing, Vol. 21, (2002), pp. 39-55.
[10] J. Schoukens, Y. Rolain, and R. Pintelon, *Improved frequency response function measurements for random noise excitations*, IEEE Trans. Instrum. Meas. Vol. 47, (1998), pp. 322-326.
[11] J. Schoukens, Y. Rolain and R. Pintelon, *Analysis of window-ing/leakage effects in frequency response function measurements*, Automatica, Vol. 42, No 1, (2006), pp. 27-38.
[12] R. Pintelon, J. Schoukens, and G. Vandersteen, *Estimation of non-parametric noise and FRF models for multivariable systems—Part I: Theory. Mechanical Systems and Signal Processing*, 2010; 24, pp. 573-595.
[13] R. Pintelon, J. Schoukens, and G. Vandersteen, *Estimation of non-parametric noise and FRF models for multivariable systems— Part II: applications and extensions.* Mechanical Systems and Signal Processing, Vol. 24, (2010), pp. 596-616.
[14] J. Schoukens, G. Vandersteen, K. Barbe, and R. Pintelon, *Nonparametric Preprocessing in System Identification: a Powerful Tool.* European Journal of Control, Vol. 15, (2009), 260-274.
[15] R. Pintelon and J. Schoukens, *System Identification. A Frequency Domain Approach.* IEEE-press, Piscataway, (2001).
[16] J.C. Agüero, J.I. Yuz, G.C. Goodwin, and R.A. Delgado, *On the equivalence of time and frequency domain maximum likelihood estimation.* Automatica, Vol. 46, (2010), pp. 260-270.