Frequency Response Function Measurements via Local Rational Modeling, Revisited

Pintelon, Rik; Peumans, Dries; Vandersteen, Gerd; Lataire, John

Published in:
IEEE Transactions on Instrumentation and Measurement

DOI:
10.1109/TIM.2020.3020601

Publication date:
2021

Document Version:
Accepted author manuscript

Frequency Response Function Measurements via Local Rational Modeling, Revisited

Rik Pintelon, *Fellow, IEEE,* Dries Peumans, *Member, IEEE,* Gerd Vandersteen, *Senior Member, IEEE,* and John Lataire, *Member, IEEE*

***Abstract*—A finite measurement time is at the origin of transient (sometimes called leakage) errors in nonparametric frequency response function (FRF) estimation. If the FRF varies significantly over the frequency resolution of the experiment (= reciprocal of the measurement time), then these transients (leakage) errors cause important bias and variance errors in the FRF estimate. To decrease these errors, several local modeling techniques have been proposed in the literature. This paper presents an overview of the existing methods and gives an in-depth bias and variance analysis of the FRF and disturbing noise variance estimates. In addition, a new local modeling approach is described that combines the small bias error of the local rational approximation with the low noise sensitivity of the local polynomial approximation. It is based on an automatic local model order selection procedure applied to a specific subclass of rational functions.**

***Index Terms*—frequency response function, nonparametric estimation, local rational modeling, local polynomial modeling, transient, leakage.**

I. Introduction

MEASURING frequency response functions (FRF) is a first step towards the (physical) understanding of the dynamic behavior of real-life systems. Practical applications can be found in many disciplines of engineering and (bio-medical) sciences such as, vibration analysis of mechanical and civil engineering structures [1], [2], state-of-health monitoring of lithium-ion batteries [3], hemodialysis measurements [4], characterization of vegetable tissues [5], ... An important difficulty in retrieving accurate and precise FRF measurements are the transient (sometimes also called leakage) errors, which can become significant whenever the FRF varies 'largely' over the frequency resolution (inverse of the experiment time). Confining these errors within a user defined value imposes a lower bound on the experiment time needed. This is especially true for lightly damped systems, such as, vibrating mechanical and civil engineering structures.

Spectral analysis via averaging of (overlapping) windowed subrecords is a classical engineering tool for reducing the impact of transient errors in FRF measurements [6], [7]. A drawback of the averaging over P non-overlapping subrecords is the loss in frequency resolution of a factor P. Via a local polynomial approximation of the FRF and the transient error, a smaller bias and variance error of the FRF estimate is obtained, and this at the full frequency resolution of the experiment [8], [9]. A local rational approximation of the FRF and the transient error combined with a linear least squares estimator (called local Levy method in this paper) has been introduced in [10], and a detailed comparison with the local polynomial approach can be found in [11], [12]. Compared with the local polynomial (LP) method, the local Levy (LL) estimator reduces significantly the transient error in the neighborhood of lightly damped resonances. Due to the disturbing noise in the measurements, and in contrast to the LP technique, the LL estimates of the FRF and the noise variance are biased [11], [13].

To cope with the bias in the LL method originating from the disturbing noise, a local output error (LOE) approach was proposed in [11], [14]. Beside requiring an iterative minimization procedure, LOE has the drawback of an increased noise sensitivity compared with LL [11], [12]. However, in this paper it is shown that LOE has similar or better noise sensitivity than LL, when combined with an appropriate model selection procedure. Although this property of LOE has already been predicted in [11], the key to the success of LOE for FRF estimation is that the model selection is performed on a specific subset of rational functions.

Summarized, the main contributions of this paper are:

- Rationale explaining why and when a particular subclass of local rational models performs better than a local polynomial approximation.
- Overview and comparison of local parametric modeling techniques for nonparametric FRF estimation.
- Noise variance estimation for all local methods – and the analysis of the bias.
- Determination of the asymptotic distribution of the FRF estimates.
- Automatic procedure for selecting the local model order of a specific subclass of rational functions.

The paper is organized as follows. First, we define the problem to be solved (Section II). Next, we describe the basic idea of the local modeling methods (Section III). Further, an in-depth study of the properties of the existing local parametric modeling techniques is made (Section IV), resulting in a novel local rational modeling approach (Section V). The theory is illustrated on a challenging simulation (Section VI) and measurement (Section VII) example. Finally, a summary of the findings can be found in Section VIII.

This work was supported in part by the Research Council of the VUB (SRP19), the Research Foundation Flanders (FWO-Vlaanderen), and by the Flemish Government (Methusalem Fund METH1).

R. Pintelon, D. Peumans, G. Vandersteen and J. Lataire are with the Department ELEC of the Vrije Universiteit Brussel, 1050 Brussel, Belgium (email: Rik.Pintelon@vub.ac.be and John.Lataire@vub.ac.be).

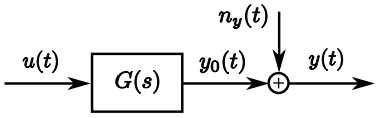

Fig. 1. Known input $u(t)$, noisy output $y(t)$ measurement of a linear time-invariant continuous-time system with transfer function $G(s)$. The system operates in open loop and the output measurement noise $n_y(t)$ is independently distributed of the input $u(t)$.

II. Problem Statement

Consider the problem of estimating nonparametrically the frequency response function (FRF) of a linear time-invariant system operating in open loop. We assume that the input $u(t)$ is known and that the output $y(t) = y_0(t)+n_y(t)$ is noisy, with $y_0(t)$ the true value and $n_y(t)$ the disturbing noise [see Fig. 1]. Define furthermore the scaled discrete Fourier transform (DFT) of N samples of a signal $x(t)$ as,

$$X(k) = \text{DFT}\{x(t)\} = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x(nT_s)e^{-j2\pi \frac{kn}{N}} \quad (1)$$

with T_s the sampling period. The relationship between $U(k) = \text{DFT}\{u(t)\}$ and $Y_0(k) = \text{DFT}\{y_0(t)\}$ is then given by

$$Y_0(k) = G(j\omega_k)U(k) + T_G(j\omega_k) \quad (2)$$

where $G(j\omega)$ is the plant frequency response function and $T_G(j\omega)$ the plant transient term. The latter is a smooth function of $j\omega$ that decreases to zero as an $O(N^{-1/2})$ for $N \to \infty$, and it accounts for the difference between the initial and final conditions of the experiment and a residual alias error [see [9], Section 6.3, pp. 184–188]. This is in the literature also know as leakage introduced by a rectangular window. From (1) and (2), it follows that the frequency resolution of the measurement equals f_s/N, which is the reciprocal of the experiment duration $T = NT_s$.

Assuming that the output measurement noise $n_y(t)$ can be written as continuous-time filtered band-limited white Gaussian noise $e(t)$, the DFT (1) $N_Y(k)$ of N samples of $n_y(t)$ can be written as

$$N_Y(k) = H(j\omega_k)E(k) + T_H(j\omega_k) \quad (3)$$

where $E(k)$ is the DFT of N samples of the unobserved $e(t)$. $H(j\omega)$ is the noise frequency response function, $T_H(j\omega)$ is the noise transient term that has the same properties as $T_G(j\omega)$ in (2), and $E(k)$ is white and circular complex normally distributed [see [9], Section 6.7, pp. 195–200].

Combining (2) and (3) we find an expression for the DFT $Y(k)$ of the measured output $y(t)$

$$Y(k) = G(j\omega_k)U(k) + T(j\omega_k) + V(k) \quad (4a)$$
$$T(j\omega) = T_G(j\omega) + T_H(j\omega) \quad (4b)$$
$$V(k) = H(j\omega_k)E(k) \quad (4c)$$

The properties of $T_G(j\omega)$ and $T_H(j\omega)$ in (4b) differ in two important ways:

1) The contribution of the plant transient term $T_G(j\omega)$ to (4a) can be put to zero via coherent sampling of the steady state response to a periodic excitation [see [9], Section 2.4, pp. 44–49], while the noise transient term $T_H(j\omega)$ will always remain different from zero for colored noise.
2) Consider the spectral analysis estimate based on P consecutive subrecords of N/P samples each [6]

$$\hat{G}(j\omega_k) = \frac{\frac{1}{P}\sum_{p=1}^{P} Y^{[p]}(k)\overline{U^{[p]}(k)}}{\frac{1}{P}\sum_{p=1}^{P} |U^{[p]}(k)|^2} \quad (5)$$

where $U^{[p]}(k)$ and $Y^{[p]}(k)$ are the DFTs of the p-th subrecord of $u(t)$ and $y(t)$. Substituting (4) in (5), gives

$$\hat{G}(j\omega_k) = G(j\omega_k) + \frac{\frac{1}{P}\sum_{p=1}^{P} T_G^{[p]}(k)\overline{U^{[p]}(k)}}{\frac{1}{P}\sum_{p=1}^{P} |U^{[p]}(k)|^2} + \frac{\frac{1}{P}\sum_{p=1}^{P} T_H^{[p]}(k)\overline{U^{[p]}(k)}}{\frac{1}{P}\sum_{p=1}^{P} |U^{[p]}(k)|^2} + \frac{\frac{1}{P}\sum_{p=1}^{P} V^{[p]}(k)\overline{U^{[p]}(k)}}{\frac{1}{P}\sum_{p=1}^{P} |U^{[p]}(k)|^2}$$

It shows that $T_G(j\omega)$ increases the variability of (5), and introduces a bias error as $T_G^{[p]}(j\omega_k)$ is correlated with $U^{[p]}(k)$ [see [9], Theorem 6.17 on p. 202], while $T_H(j\omega)$ only increases the variability of (5) as $T_H^{[p]}(j\omega_k)$ is independent of $U^{[p]}(k)$.

Especially for lightly damped systems, the variance and bias errors introduced by the transient term $T(j\omega)$ in (4) are an issue in nonparametric FRF estimation [7].

III. Basic Idea Local Modeling Methods

In classical spectral analysis estimation (5), the transient (leakage) problem is handled via time-domain windowing of the subrecords. Windows often used are the half-sine and Hanning windows that perform, respectively, a scaled differencing and a scaled double differencing of the DFT spectra [7], [9]. Hence, by combining two (half-sine) or three (Hanning) neighboring frequencies, the impact of the transient term in (5) is diminished. Based on this observation, the local modeling techniques use multiple neighboring frequencies to suppress furthermore the transient error [see Section III-A].

This section is organized as follows. First, we define the type of local models considered [Section III-A] and, next, we explain intuitively why and when local rational models perform better than local polynomial approximations [Section III-B].

A. Definition of Local Models

To decrease the bias and variance error of the FRF estimate due to the transient term $T(j\omega)$ in (4), the local modeling methods [8], [10], [13], [14] approximate the FRF $G(j\omega)$ and the transient $T(j\omega)$ in a small frequency band centered around an excited frequency $f_k = kf_s/N$

$$[f_{k-n}, f_{k-n+1}, \ldots, f_{k-1}, f_k, f_{k+1}, \ldots, f_{k+n-1}, f_{k+n}] \quad (6)$$

by a low order parametric model [see Fig. 2]. This parametric model is either a complex polynomial or a complex rational function of $j(\omega_{k+r} - \omega_k) = j\omega_r$,

$$G(j\omega_{k+r}) = \frac{B_k^c(j\omega_r)}{A_k^c(j\omega_r)} + O\Big(\Big(\frac{r}{N}\Big)^{g_k}\Big)G^{(g_k)}(j\omega_k) \quad (7a)$$

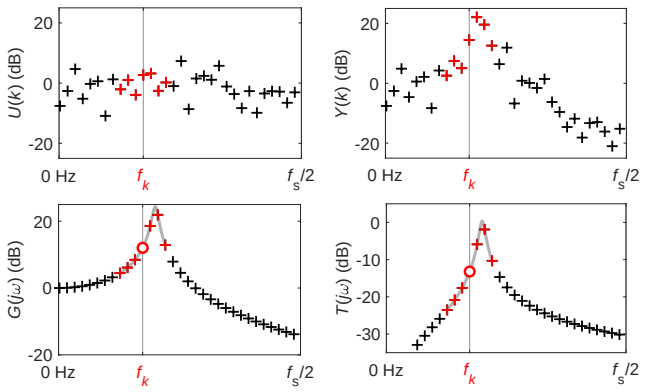

Fig. 2. Basic idea of the local modeling approach. Using the known input and noisy output DFTs $U(k)$ and $Y(k)$ [top row] in a small frequency band [red '+'] centered around an excited frequency $f_k = kf_s/N$, the frequency response function $G(j\omega)$ [bottom left] and the transient term $T(j\omega)$ [bottom right] in (4) are locally approximated by a complex polynomial or a complex rational function of $j(\omega-\omega_k)$ [gray lines]. The '+' and 'o' in the bottom row indicate the true values inside (red) and outside (black) the local frequency band.

$$T(j\omega_{k+r}) = \frac{I_k^c(j\omega_r)}{A_k^c(j\omega_r)} + O\Big(\Big(\frac{r}{N}\Big)^{t_k+0.5}\Big)T^{(t_k)}(j\omega_k) \quad (7b)$$

with $\omega = 2\pi f$, $g_k = n_{a,k} + n_{b,k} + 1$, $t_k = n_{a,k} + n_{i,k} + 1$, $r = -n, -n+1, \ldots, -1, 0, 1, \ldots, n$, $X^{(n)}(j\omega)$ the n-th order derivative w.r.t. $j\omega$, and where $X_k^c(j\omega)$, $X = B, A, I$, are complex polynomials of order $n_{x,k}$

$$X_k^c(j\omega) = \sum_{n=0}^{n_{x,k}} x_{n,k}(j\omega)^n \quad \text{with} \quad x_{n,k} \in \mathbb{C} \quad (8)$$

The remainders in (7) originate from a Taylor series expansion at $j\omega_k$ [see, [15] for the local polynomial model, $n_{a,k} = 0$, and [16] for the local rational model, $n_{a,k} > 0$], and the extra 0.5 in the power of the remainder of the transient term (7b) is a consequence of the fact that $T(j\omega) = O(N^{-1/2})$. Note that the index k in the polynomial order $n_{x,k}$ has been added to indicate that the local model order is frequency dependent.

Combining (7) with (8), we find the following relationship between the complex polynomial coefficients and the FRF and the transient term at frequency f_k

$$G(j\omega_k) = \frac{b_{0,k}}{a_{0,k}} \quad \text{and} \quad T(j\omega_k) = \frac{i_{0,k}}{a_{0,k}} \quad (9)$$

where $a_{0,k} = 1$ for a local polynomial approximation. Repeating the local modeling procedure at all frequencies f_k, $k = 1, 2, \ldots, N/2-1$ eventually gives the nonparametric FRF estimate. Hence, in total, much more parameters are estimated than the amount of data available. This apparent contradiction is resolved at the end of Section V-B.

In (7), it is silently assumed that the poles of the noise and plant transient terms $T_H(s)$ and $T_G(s)$ are the same. However, this condition is fulfilled for ARX and ARMAX model structures only [17]. To deal with this issue, the order $n_{i,k}$ of the numerator of $I_k^c(s)/A_k^c(s)$ in (7b) is chosen to be larger than the order $n_{a,k}$ of the denominator. With this choice, $I_k^c(s)/A_k^c(s)$ can be written as the sum of a rational function of order $n_{a,k}$ and a polynomial of order $n_{i,k} - n_{a,k}$, where the rational function models $T_G(s)$ and the polynomial $T_H(s)$.

There are two reasons why the poles of $T_G(s)$ are modeled instead of those of $T_H(s)$:

1) $T_G(s)$ introduces a variance and a bias error in the FRF estimate, while $T_H(s)$ only increases the variance.
2) Parametrization (7) simplifies the generation of starting values for the local rational modeling [see Section IV-B].

Notice that the transient term $T(j\omega)$ is most important at the location of the poles of the system $G(s)$ and noise $H(s)$ transfer functions [9]. Since these pole locations are unknown beforehand, $T(j\omega)$ is accounted for over the whole frequency band.

B. Rational versus Polynomial Approximation

A real rational transfer function $G(s)$ of order n_b/n_a can always be expanded in partial fractions. Assuming simple poles $p_n \in \mathbb{C}$ we get

$$G(s) = \sum_{n=1}^{n_a} \frac{R_n}{s - p_n} + P(s) \quad (10)$$

where $R_n \in \mathbb{C}$ is the residue of the pole p_n, and with $P(s)$ a real polynomial of order $n_b - n_a$, if $n_b \geqslant n_a$, and zero otherwise. Subtracting the term $R_1/(s - p_1)$ from $G(s)$ in (10) results in a transfer function $G_{\text{pr}}(s)$

$$G_{\text{pr}}(s) = G(s) - \frac{R_1}{s - p_1} \quad (11)$$

whose amplitude and phase vary much less than those of $G(s)$ in the neighborhood of the resonance linked to the pole p_1. Since the remainder in the polynomial approximation (7a), with $n_{a,k} = 0$, depends on $(n_{b,k}+1)$-th derivative of $G(j\omega_k)$ [see [15]], it follows that, after removal of the pole p_1, the polynomial approximation error in the neighborhood of the corresponding resonance will be much smaller than that of the original FRF. This is illustrated in Fig. 3 for the case

$$G(s) = \frac{1}{\frac{s^2}{\omega_0^2} + 2\zeta\frac{s}{\omega_0} + 1} \quad (12)$$

with $\zeta = 0.03$, $f_0 = 1\,\text{kHz}$ and $p_1 = \omega_0(-\zeta + j\sqrt{1-\zeta^2})$. When comparing the gray line with the red circle in the top row, it can be seen that the FRF cannot be approximated well by a second order polynomial, resulting in a relative bias error of -6 dB [50% error] at frequency f_k. After removal of the pole, the relative bias error of the second order polynomial approximation is reduced to -90 dB [3×10^{-3}% error] at frequency f_k, as shown in the bottom row.

The smoothing effect of the pole removal motivates the following local rational modeling strategy. In a 'close' neighborhood of a resonance we choose $n_{a,k} \geqslant 1$ and $n_{b,k} > n_{a,k}$ (rational approximation), while elsewhere we select $n_{a,k} = 0$ (polynomial approximation) to avoid nearly coinciding poles and zeroes in the estimated FRF. How to detect automatically the 'close' neighborhood of a resonance described by a pole $s = p_1$ is explained in Section V-A. For example, with the choice $n_{a,k} = 1$ and $n_{b,k} > n_{a,k}$, the complex rational function $B_k^c(j\omega_r)/A_k^c(j\omega_r)$ in (7a) can be split as

$$\frac{B_k^c(j\omega_r)}{A_k^c(j\omega_r)} = \frac{R_{1,k}}{j\omega_r - p_{1,k}} + P_k(j\omega_r) \quad (13)$$

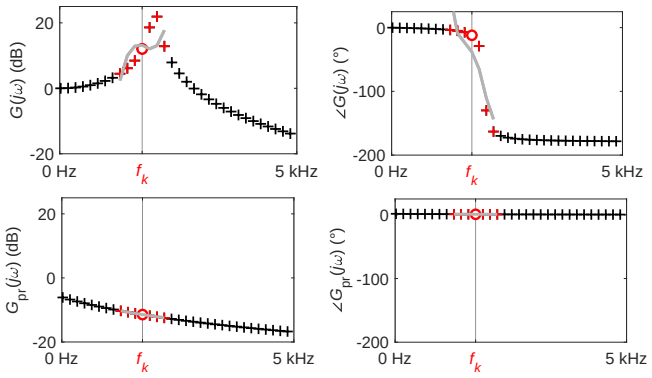

Fig. 3. Second order local polynomial modeling [gray lines] in a frequency band around f_k [red '+'] before [top row] and after [bottom row] removal of the pole $s = \omega_0(-\zeta + j\sqrt{1-\zeta^2})$ in the transfer function $G(s)$ (12) [$f_s = 10\,\text{kHz}$]. Left column: amplitude; and right column: phase. The '+' and 'o' indicate the true values inside (red) and outside (black) the local frequency band.

with $R_{1,k} \in \mathbb{C}$ and $P_k(j\omega_r)$ a complex polynomial of order $n_{b,k} - 1$. If $p_{1,k} + j\omega_k$ and $R_{1,k}$ in (13) are close to, respectively, the pole p_1 and its corresponding residue R_1, then

$$\frac{R_{1,k}}{j\omega_r - p_{1,k}} \approx \frac{R_1}{j\omega_{k+r} - p_1} \tag{14}$$

and the polynomial part $P_k(j\omega_r)$ can locally approximate very well the remaining dynamics $G(j\omega_{k+r}) - R_{1,k}/(j\omega_r - p_{1,k})$. If $n_{a,k} > 1$, then the same reasoning is applied to the nearest $n_{a,k}$ poles.

Summarized, in the 'close' neighborhood of a resonance caused by a pole p_1, the remainders of the local polynomial [(7a) with $n_{a,k} = 0$] and local rational [(7a) with $n_{a,k} > 0$] models can be approximated as, respectively,

$$G^{(m_k)}(j\omega_k) O\Big(\Big(\frac{r}{N}\Big)^{m_k}\Big) \tag{15a}$$

$$\Big(G(j\omega) - \frac{R_{1,k}}{j\omega - (p_{1,k} + j\omega_k)}\Big)^{(m_k)}\Big|_{\omega=\omega_k} O\Big(\Big(\frac{r}{N}\Big)^{m_k}\Big) \tag{15b}$$

where $m_k = n_{b,k} + 1$. To construct (15b), we assumed that $R_{1,k}$ and $p_{1,k}$ in (13) are fixed. In practice, $R_{1,k}$ and $p_{1,k}$ are estimated simultaneously with the polynomial part in (13). This will lower the approximation error because more parameters are estimated. Therefore, the magnitude of (15b) is an upper bound for the rational approximation error. The remainder (15b) of the rational approximation is smaller than the remainder (15a) of the polynomial approximation if

$$\left|\Big(G(j\omega) - \frac{R_{1,k}}{j\omega - (p_{1,k} + j\omega_k)}\Big)^{(m_k)}\Big|_{\omega=\omega_k}\right| < \left|G^{(m_k)}(j\omega_k)\right| \tag{16}$$

Note that (16) puts a condition on the smoothing effectiveness of the subtraction $G(j\omega) - R_{1,k}/(j\omega - (p_{1,k} + j\omega_k))$ [the error of (14) should be sufficiently small] and the location of the frequency f_k [f_k should be sufficiently close to the resonance]. If both conditions are not fulfilled, then a polynomial approximation will perform better.

IV. Local Modeling Methods

First, we state the conditions on the input and the disturbing noise under which the FRF can be estimated from (4) [Section IV-A]. Next, we give an overview of the existing local methods for estimating the FRF [Section IV-B]. Their pros and cons are discussed and a detailed bias [Section IV-B] and variance [Section IV-D] analysis is made. Finally, a local noise variance estimator is proposed [Section IV-C] which is used for estimating the variance of the FRF [Section IV-D].

A. Basic Assumptions

Due to the smoothness of the FRF $G(j\omega)$ and the transient $T(j\omega)$, the terms $G(j\omega_k)U(k)$ and $T(j\omega_k)$ in (4) can be distinguished only if the input DFT $U(k)$ varies in a non-smooth way over the frequency. This leads to the following identifiability condition.

Assumption 1 (Roughness Input DFT)**.** *At the excited frequencies $f_k = kf_s/N$, the spectral difference*

$$|U(k+1) - U(k)| = O(N^0)$$

with $U(k)$ the scaled DFT (1)*, does not vanish to zero as $N \to \infty$.*

Examples of 'rough' signals are random noise, periodic noise, random phase multisines, and (pseudo)random binary sequences [see [9], Chapter 5]. To avoid any bias due to a correlation between the true input and the output noise it is assumed that the system operates in open loop.

Assumption 2 (Known input, Noisy Output Observation of a System Operating in Open Loop)**.** *The system operates in open loop and the true input $u(t)$ is independent of the noise $n_y(t)$ at the output. $n_y(t)$ can be written as continuous-time filtered band-limited white Gaussian noise $e(t)$.*

B. Overview of Local Methods for FRF Estimation

According to the model structure used, (7) with $n_{a,k} = 0$ or (7) with $n_{a,k} > 0$, the local methods can be classified in the local polynomial and the local rational methods. The *local polynomial (LP)* method minimizes

$$V_{\text{LP}}(\theta_k, Z_k) = \sum_{r=-n}^{n} |e_{\text{LP}}(j\omega_r, \theta_k, Z_k)|^2 \tag{17a}$$

$$e_{\text{LP}}(j\omega_r, \theta_k, Z_k) = Y(k+r) - B_k^{\text{c}}(j\omega_r, \theta_k)U(k+r) - I_k^{\text{c}}(j\omega_r, \theta_k) \tag{17b}$$

$$\theta_k = [b_{0,k}, \ldots, b_{n_{b,k},k}, i_{0,k}, \ldots, i_{n_{i,k},k}]^T \tag{17c}$$

w.r.t. $\theta_k \in \mathbb{C}^{n_{b,k}+n_{i,k}+2}$, where Z_k is a vector containing the known input and noisy output DFT spectra in the band $[k-n, \ldots, k+n]$ (see [8]). The *local rational* methods aim to minimize the *local output error (LOE)* cost function

$$V_{\text{LOE}}(\theta_k, Z_k) = \sum_{r=-n}^{n} |e_{\text{LOE}}(j\omega_r, \theta_k, Z_k)|^2 \tag{18a}$$

$$e_{\text{LOE}}(j\omega_r, \theta_k, Z_k) = Y(k+r) - \frac{B_k^{\text{c}}(j\omega_r, \theta_k)}{A_k^{\text{c}}(j\omega_r, \theta_k)}U(k+r) - \frac{I_k^{\text{c}}(j\omega_r, \theta_k)}{A_k^{\text{c}}(j\omega_r, \theta_k)} \tag{18b}$$

$$\theta_k = [a_{0,k}, \ldots, a_{n_{a,k},k}, b_{0,k}, \ldots, b_{n_{b,k},k}, i_{0,k}, \ldots, i_{n_{i,k},k}]^T \tag{18c}$$

w.r.t. $\theta_k \in \mathbb{C}^{n_{a,k}+n_{b,k}+n_{i,k}+3}$ and subject to a constraint, for example, $a_{n_{a,k},k} = 1$ or $\|\theta_k\|_2 = 1$. While (17) is a linear least squares problem, (18) is a nonlinear least squares problem requiring an iterative optimization method and starting values.

In [10] the nonlinear minimization problem (18a) is avoided by calculating the Levy solution [18] of the rational approximation problem. The *local Levy (LL)* method minimizes

$$V_{\mathrm{LL}}(\theta_k, Z_k) = \sum_{r=-n}^{n} |e_{\mathrm{LL}}(j\omega_r, \theta_k, Z_k)|^2 \tag{19a}$$

$$e_{\mathrm{LL}}(j\omega_r, \theta_k, Z_k) = A_k^{\mathrm{c}}(j\omega_r, \theta_k)Y(k+r) - B_k^{\mathrm{c}}(j\omega_r, \theta_k)U(k+r) - I_k^{\mathrm{c}}(j\omega_r, \theta_k) \tag{19b}$$

w.r.t. θ_k defined in (18c). Due to the weighting of the output noise $V(k+r)$ (4c) with $A_k^{\mathrm{c}}(j\omega_r, \theta_k)$ in the residual (19b), the noise at the borders of the local frequency band are overemphasized and the minimizer of (19a) is biased [see [9], Section 9.8.2, pp. 301–303], even if the disturbing noise on the equation error (19b) is white [17].

Dividing the residual in (19b) by an initial guess of the denominator polynomial $A_k^{\mathrm{c}}(j\omega_r, \theta_k^{[0]})$, where $\theta_k^{[0]}$ is the minimizer of (19a), reduces the amplification of the output noise at the borders of the local frequency band. The obtained weighted linear least squares estimate $\theta_k^{[1]}$ can be used to calculate a (hopefully) improved estimate of the denominator polynomial $A_k^{\mathrm{c}}(j\omega_r, \theta_k^{[1]})$, giving a (hopefully) better estimate $\theta^{[2]}$. Recursive application of this procedure results in the Sanathanan and Koerner solution [19] of the local rational approximation problem [14]. The i-th step of the *local Sanathanan and Koerner (LSK)* method minimizes

$$V_{\mathrm{LSK}}(\theta_k^{[i]}, Z_k) = \sum_{r=-n}^{n} \left| \frac{e_{\mathrm{LL}}(j\omega_r, \theta_k^{[i]}, Z_k)}{A_k^{\mathrm{c}}(j\omega_r, \theta_k^{[i-1]})} \right|^2 \tag{20}$$

w.r.t. $\theta_k^{[i]}$ and with $e_{\mathrm{LL}}(j\omega_r, \theta_k, Z_k)$ defined in (19b). Similar to the local Levy method (19) the local Sanathanan and Koerner method is biased due to the weighting of the output noise with $A_k^{\mathrm{c}}(j\omega_r, \theta_k^{[i]})$ [see [9], Section 9.8.3, pp. 303–305]. Applying the Sanathanan and Koerner iteration to a partial fraction parametrization of the rational functions results in the local vector fitting method [13].

Solving the full nonlinear minimization problem (18) reduces the bias caused by the output noise in the local Levy and local Sanathanan and Koerner estimates [14]. To avoid almost coinciding estimated pole-zero pairs in local frequency bands that are 'far away' from a resonance, a model selection procedure should be added [see [11] and Section V-A].

An approach that combines the bias reduction property of the LOE method (18) with the global minimization property of the LL (19) and LSK (20) methods is the *local generalized total least squares (LGTLS)* estimator [13], based on the GTLS estimator of [20]. Approximating the noise variance $\sigma_V^2(k+r) = \mathrm{var}(V(k+r)) \approx \sigma_V^2(k)$, for $r = -n, -n+1, \ldots, n$, the LGTLS estimator minimizes

$$V_{\mathrm{LGTLS}}(\theta_k, Z_k) = \frac{\sum_{r=-n}^{n} |e_{\mathrm{LL}}(j\omega_r, \theta_k, Z_k)|^2}{\sum_{r=-n}^{n} \sigma_V^2(k) |A_k^{\mathrm{c}}(j\omega_r, \theta_k)|^2} \tag{21}$$

TABLE I
VALUES OF α_k, β AND γ IN THE BIAS EXPRESSION (23) OF THE LOCAL METHODS.

local method	α_k	β	γ
LP (17)	$n_{b,k}+1$	∞	∞
LOE (18)	$n_{a,k}+n_{b,k}+1$	1	∞
LL (19), LSK (20)	$n_{a,k}+n_{b,k}+1$	0	∞
LGTLS (21), LBTLS (22)	$n_{a,k}+n_{b,k}+1$	1	1

w.r.t. θ_k and subject to the constraint $\|\theta_k\|_2 = 1$. Similar to the local Levy method (19), the noise at the borders of the local frequency band is overemphasized in (21).

Following the same lines of the Sanathanan and Koerner iteration (20), the noise amplification in (21) can be suppressed by dividing the residual (19b) with an initial guess of the denominator polynomial. The i-th step of the resulting *local bootstrapped total least squares (LBTLS)* estimator[13], based on the BTLS estimator of [20], minimizes

$$V_{\mathrm{LBTLS}}(\theta_k^{[i]}, Z_k) = \frac{\sum_{r=-n}^{n} \frac{|e_{\mathrm{LL}}(j\omega_r, \theta_k^{[i]}, Z_k)|^2}{\sigma_V^2(k)|A_k^{\mathrm{c}}(j\omega_r, \theta_k^{[i-1]})|^2}}{\sum_{r=-n}^{n} \frac{|A_k^{\mathrm{c}}(j\omega_r, \theta_k^{[i]})|^2}{|A_k^{\mathrm{c}}(j\omega_r, \theta_k^{[i-1]})|^2}} \tag{22}$$

w.r.t. $\theta_k^{[i]}$ and subject to the constraint $\left\|\theta_k^{[i]}\right\|_2 = 1$. As initial value $\theta_k^{[0]}$ we take the LGTLS solution (21). Similar to the LOE method (18), near pole-zero cancellations in local frequency bands 'far away' from a resonance are avoided via a model selection procedure [see [13] and Section V-A].

Note that the global minimizers of the nonlinear least squares cost functions (21) and (22), (i) are independent of $\sigma_V^2(k)$, and (ii) are calculated via a generalized singular value decomposition of a matrix pair [see [20]]. In the remainder of this section we analyze the bias of the local methods and the impact of the approximation $\sigma_V^2(k+r) \approx \sigma_V^2(k)$ in (21) and (22).

All local methods are subject to the bias errors of the polynomial or rational approximations (7). In addition, all methods – except LP (17) – also have (for finite n and N) a bias contribution due to the disturbing output noise. The following theorem quantifies these two bias contributions as a function of the number of time domain samples N and the number of frequencies $2n+1$ in the local frequency band.

Theorem 1 (Bias Local FRF Estimate)**.** *Under Assumptions 1 and 2, the bias of the local FRF estimates equals*

$$\mathbb{E}\{\hat{G}(j\omega_k)\} - G(j\omega_k) = G^{(\alpha_k)}(j\omega_k) O\Big(\big(\frac{n}{N}\big)^{\alpha_k}\Big) + \frac{\sigma_V^2(k)}{\mathbb{E}\{\|Y_0(k)\|^2\}} \Big[O\Big(\frac{1}{(2n+1)^{\beta}}\Big) + \frac{(\sigma_V^2(k))'}{\sigma_V^2(k)} O\Big(\big(\frac{n}{N}\big)^{\gamma}\Big)\Big] \tag{23}$$

with $()'$ the derivative w.r.t. k, N the number of time domain samples used for calculating the DTF, and $2n+1$ the number of frequencies in the local frequency band. The values α_k, β and γ in (23) *are given by Table I and correspond to, respectively, the polynomial or rational approximation error, the bias caused by the disturbing noise, and the bias due to the approximation $\sigma_V^2(k+r) \approx \sigma_V^2(k)$ for $r = -n, -n+1 \ldots, n$.*

Proof. See Appendix A.

From (23) it can be seen that for the LOE, LGTLS and LBTLS estimators the approximation errors [first and third term] vanish as $N \to \infty$ for n fixed, while the bias error due to the noise tends to zero as $n \to \infty$. To reconcile these conflicting demands, n should increase slower to infinity than N, for example, $n = O(N^\delta)$ with $0 < \delta < 1$. Proceeding in this way the width $(2n+1)f_s/N$ in Hz of the local frequency band still decreases to zero while the number of frequencies $2n+1$ in that band increases to infinity.

C. Local Estimation of the Noise Variance

Assuming that the disturbing noise is white in the local frequency band $[k-n, k+n]$, a noise variance estimate is obtained from the difference between the measured output $Y(k)$ and the output $\hat{Y}(k)$ predicted by the local models

$$\hat{\sigma}_V^2(k) = \frac{1}{2n+1-n_{\theta_k}} \sum_{r=-n}^{n} \left|Y(k+r) - \hat{Y}(k+r)\right|^2 \quad (24a)$$

$$\hat{Y}(k+r) = \hat{G}_k^{\mathrm{c}}(j\omega_r)U(k+r) + \hat{T}_k^{\mathrm{c}}(j\omega_r) \quad (24b)$$

where $\hat{G}_k^{\mathrm{c}}$ and $\hat{T}_k^{\mathrm{c}}$ are the local estimates of the FRF and the transient term

$$\hat{G}_k^{\mathrm{c}}(j\omega_r) = \frac{B_k^{\mathrm{c}}(j\omega_r, \hat{\theta}_k)}{A_k^{\mathrm{c}}(j\omega_r, \hat{\theta}_k)}, \; \hat{T}_k^{\mathrm{c}}(j\omega_r) = \frac{I_k^{\mathrm{c}}(j\omega_r, \hat{\theta}_k)}{A_k^{\mathrm{c}}(j\omega_r, \hat{\theta}_k)} \quad (24c)$$

and with $n_{\theta,k}$ the number of free local parameters (= total number minus one for the rational models)

$$n_{\theta,k} = n_{a,k} + n_{b,k} + n_{i,k} + 2 \quad (24d)$$

The difference between the number of local equations $2n+1$ and the number of local parameters $n_{\theta,k}$ (24d) is called the degrees of freedom dof_k

$$\mathrm{dof}_k = 2n + 1 - n_{\theta,k} \quad (25)$$

It accounts for the decrease in degrees of freedom of the output residual by the $n_{\theta,k}$ estimated model parameters, and quantifies the uncertainty of the noise variance estimate (24a) via a χ^2-distribution with $2\mathrm{dof}_k$ degrees of freedom.

For all local methods the noise variance estimate (24a) is biased due to the white noise approximation and the bias error of the estimated local models. In addition, all methods – except LP (17) – also have (for finite n and N) a bias contribution due to the correlation between the disturbing noise $V(k+r)$ (4c) and the estimated local models $\hat{G}_k^{\mathrm{c}}(j\omega_r)$ and $\hat{T}_k^{\mathrm{c}}(j\omega_r)$. The following theorem quantifies these three bias contributions as a function of the number of time domain samples N and the number of frequencies $2n+1$ in the local frequency band.

Theorem 2 (Bias Local Noise Variance Estimate)**.** *Under Assumptions 1 and 2 the bias of the local variance estimate* (24) *equals*

$$\mathbb{E}\{\hat{\sigma}_V^2(k)\} - \sigma_V^2(k) = (\sigma_V^2(k))' O\left(\frac{n}{N}\right) + \sigma_V^2(k) O\left(\frac{1}{n^\alpha}\right) + O(b_G^2) \quad (26)$$

with $()'$ the derivative w.r.t. k, N the number of time domain samples used for calculating the DTF, $2n+1$ the number of frequencies in the local frequency band, and b_G the right hand side of (23)*. $\alpha = \infty$ for LP* (17) *and $\alpha = 1$ for LOE* (18)*, LL* (19)*, LSK* (20)*, LGTLS* (21)*, and LBTLS* (22)*. The terms in the right hand side of* (26) *correspond to, respectively, the approximation $\sigma_V^2(k+r) \approx \sigma_V^2(k)$ in the local frequency band, the correlation between the disturbing noise and the estimated local models, and the bias error of the local models.*

Proof. See Appendix B.

D. Variance Nonparametric FRF Estimate

First, we show that the local FRF estimate (9) is asymptotically ($n \to \infty$) (circular) complex normally distributed. Next, we handle the practical calculation of the variance of the FRF estimate and the construction of a $100 \times p\%$ confidence bound.

Theorem 3 (Asymptotic Distribution Local FRF Estimate)**.** *Under Assumptions 1 and 2 the FRF estimate* (9) *has – in the absence of local modeling errors – the following (asymptotic) distribution:*

1) *LP* (17) *method: circular complex normally distributed.*
2) *LOE* (18) *method: asymptotically ($n \to \infty$) circular complex normally distributed.*
3) *LL* (19)*, LSK* (20)*, LGTLS* (21)*, and* (22) *are asymptotically ($n \to \infty$) complex normally distributed.*

Proof. See Appendix C.

A practical consequence of Theorem 3 is that the (asymptotic) distribution of the LP and LOE FRF estimates is solely determined by the (asymptotic) variance

$$\mathrm{var}(\hat{G}(j\omega_k)) = \mathbb{E}\{|\hat{G}(j\omega_k) - \mathbb{E}\{\hat{G}(j\omega_k)\}|^2\} \quad (27)$$

while that of the LL, LSK, LGTLS and LBTLS estimates also requires the knowledge of

$$\mathbb{E}\{(\hat{G}(j\omega_k) - \mathbb{E}\{\hat{G}(j\omega_k)\})^2\} \quad (28)$$

Hence, the most compact uncertainty regions of the LP and LOE FRF estimates are circles [see [9], Section 2.4.2.3 on pp. 48–49], while those of the LL, LSK, LGTLS and LBTLS FRF estimates are – in general – ellipses.

In the remainder of this section we discuss the practical calculation of (27) and the construction of a $100 \times p\%$ confidence bound. If the disturbing noise variance is constant in the local frequency band $[k-n, k+n]$ then, under Assumption 2, the LOE estimator (18) is the maximum likelihood solution which is asymptotically ($n \to \infty$) efficient [see [9], Theorem 9.21 on pp. 298–299]. Hence, within the class of consistent estimators, the covariance matrix of the LOE estimate $\hat{\theta}_k$ is asymptotically ($n \to \infty$) the smallest, and can be approximated by the Cramér-Rao lower bound (CRLB)

$$\mathrm{CRLB}(\theta_{k,0}) = \sigma_V^2 \left[J_{\mathrm{LOE}}^H(\theta_{k,0}) J_{\mathrm{LOE}}(\theta_{k,0})\right]^{-1} \quad (29)$$

with $J_{\mathrm{LOE}}(\theta_k) = \partial e_{\mathrm{LOE}}/\partial \theta_k$, $e_{\mathrm{LOE}} \in \mathbb{C}^{(2n+1)\times 1}$ the vector of the local output error residuals (18b), and $\theta_{k,0}$ the vector of the true local model parameters. In practice σ_V^2 and $\theta_{k,0}$ in

(29) are unknown and replaced by, respectively, $\hat{\sigma}_V^2(k)$ (24a) and $\hat{\theta}_k$, resulting in the following estimate

$$\widehat{\mathrm{Cov}}(\hat{\theta}_k) = \hat{\sigma}_V^2(k)\big[J_{\mathrm{LOE}}^H(\hat{\theta}_k)J_{\mathrm{LOE}}(\hat{\theta}_k)\big]^{-1} \tag{30}$$

Note that (30) is also valid for the LP method (17). Since the BTLS estimate is mostly (very) close to the ML solution, the covariance of the LBTLS estimate (22) can be approximated well by (30), where θ_k is the minimizer of (22). For the other local estimators (LL, LSK, and LGTLS) the approximation error of (30) is bigger.

Finally, using (30), a variance estimate $\hat{\sigma}_{\hat{G}}^2(j\omega_k)$ of the FRF (9) is obtained as, respectively,

$$\hat{\sigma}_{\hat{G}}^2(j\omega_k) = \widehat{\mathrm{var}}(\hat{b}_{0,k}) \tag{31a}$$

$$\hat{\sigma}_{\hat{G}}^2(j\omega_k) = |\hat{G}(j\omega_k)|^2\Big(\frac{\widehat{\mathrm{var}}(\hat{a}_{0,k})}{|\hat{a}_{0,k}|^2} + \frac{\widehat{\mathrm{var}}(\hat{b}_{0,k})}{|\hat{b}_{0,k}|^2} - 2\mathrm{Re}\Big(\frac{\widehat{\mathrm{covar}}(\hat{a}_{0,k},\hat{b}_{0,k})}{\hat{a}_{0,k}\overline{\hat{b}_{0,k}}}\Big)\Big) \tag{31b}$$

where the second equation is the result of a first order Taylor series expansion of (9) [proof: use Eq. (79) of Appendix C]. Note that the degrees of freedom of (31) determines the uncertainty of the noise variance estimate (24a).

Taking into account the uncertainty (degrees of freedom dof_k) of the FRF variance estimate (31), a $100\times p\%$ confidence bound for $\hat{G}(j\omega_k)$ can be constructed as a circle with center $\hat{G}(j\omega_k)$ and radius $\hat{\sigma}_{\hat{G}}(j\omega_k)\sqrt{F_p(2,2\mathrm{dof}_k)}$,

$$\mathrm{Prob}\Big(|\hat{G}(j\omega_k) - G(j\omega_k)| \leqslant \hat{\sigma}_{\hat{G}}(j\omega_k)\sqrt{F_p(2,2\mathrm{dof}_k)}\Big) = p \tag{32}$$

where $F_p(2,2\mathrm{dof}_k)$ is the $100\times p\%$ percentile of an $F(2,2\mathrm{dof}_k)$-distributed random variable with 2 and $2\mathrm{dof}_k$ degrees of freedom [see [9], Eq. (2-40), p. 51].

V. Automatic Procedure for Selecting the Local Modal Order of a Specific Subclass of Rational Functions

Based on the insights provided by Sections III and IV, an automatic local model order selection procedure is proposed in Section V-A. Application of this procedure on a specific subclass of rational functions, defines a new local rational modeling approach [Section V-B].

While in the previous sections it is explicitly assumed that the input is known, and that the system is linear and operates in open loop, an extension of the novel local rational modeling method to noisy input, noisy output measurements of a certain class of nonlinear systems operating in feedback is presented in Section V-C.

A. Model Order Selection

From Figs. 2 and 3 it can be deduced that the complexity of the local polynomial and local rational models (7) depends on the local frequency band. Therefore, an appropriate procedure for selecting the polynomial orders in these models is needed to avoid undermodeling (the bias error is larger than the noise error) and overmodeling (overfitting) such as, for example, an estimated pole-zero pair that is almost coinciding because it fits a noise peak instead of the plant dynamics. This is valid for each of the local modeling techniques discussed in Section IV-B. Since the rational model (7) with $n_{a,k} = 0$ simplifies to the polynomial model, we will concentrate the discussion to the selection of the polynomial orders $n_{a,k}$, $n_{b,k}$ and $n_{i,k}$ in (7).

A difficulty in selecting the appropriate model order (values of $n_{a,k}$, $n_{b,k}$ and $n_{i,k}$), is that the amount of local data $2n+1$ is marginally larger than the number of free local model parameters $n_{\theta,k}$ (24d). Indeed, as a rule of thumb, the amount of data should be at least ten times the number of model parameters for the large sample results ($(2n+1)/n_{\theta,k} \gg 1$) to be useful for finite sample sizes [21]. To cope with small sample sizes ($2n+1 \sim n_{\theta,k}$), a modified Akaike information criterion (AIC) has been proposed in [21]. Here, following the same lines of [22], we use the equivalently modified minimum description length (MDL) criterion because it rejects better complex models than AIC. Since the ultimate goal of all local modeling techniques is to minimize the output error cost function (18) without overfitting, the modified MDL criterion to be minimized over $n_{a,k}$, $n_{b,k}$ and $n_{i,k}$ takes the form

$$\frac{V_{\mathrm{LOE}}(\hat{\theta}_k, Z_k)}{\mathrm{dof}_k} e^{\log(4n+2)\frac{n_{\theta,k}}{\mathrm{dof}_k - 2}} \tag{33}$$

where $n_{\theta,k}$ and dof_k are defined in (24d) and (25) respectively, and with $\hat{\theta}_k$ the minimizer of LP (17), LOE (18), LL (19), LSK (20), LGTLS (21) or LBTLS (22) cost function. The modified MDL criterion (33) predicts what would happen if the estimated local model is evaluated on a validation data set [23]. It avoids splitting the data record in an identification and a validation data set of equal length, which would result in a loss of frequency resolution of a factor two.

Notice that the minimization of (33) in each local frequency band $[k-n, k+n]$ requires the estimation of multiple local models over a three dimensional space of integer numbers $n_{a,k}, n_{b,k}, n_{i,k} \in \mathbb{N}$. Some practical guidelines for limiting the search space and, hence, the computation time, are given in Section V-B.

B. Local Rational Modeling Procedure

First, note that the choice $n_{b,k} > n_{a,k}$ is beneficial for modeling the residual plant dynamics after a pole removal [see Section III-B for a numerical example]. Next, practice showed that the local polynomial approximations of odd degree $2R+1$ do not perform better than the even degree $2R$ ones [see Section III.B of [24] for a rationale]. Therefore, the difference between the numerator $n_{b,k}$ and denominator $n_{a,k}$ orders of the plant transfer function (7a) will be limited to even numbers.

In order to model the noise transient in (4), the numerator order $n_{i,k}$ of the transient term (7b) is chosen to be larger than the denominator order $n_{a,k}$ of the plant transfer function [see Section III-A for a motivation]. Finally, the local (rational) models should be kept as simple as possible, and poles should only be introduced when needed.

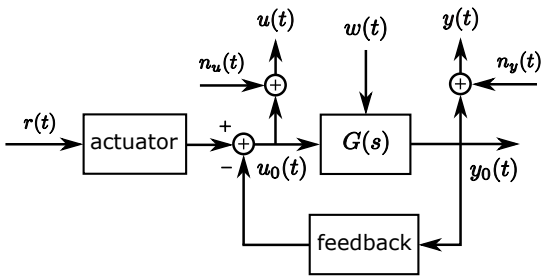

Fig. 4. Noisy input $u(t)$, noisy output $y(t)$ measurement of a system $G(s)$ subject to process noise $w(t)$ and operating in feedback. The input-output measurement noise $n_u(t)$, $n_y(t)$ and the process noise $w(t)$ are independently distributed of the reference signal $r(t)$.

Based on these observations we impose the following constraints on $n_{a,k}$, $n_{b,k}$ and $n_{i,k}$ during the minimization of (33)

$$n_{a,k} = 0,\ 1,\ \ldots,\ n_{a,\max} \leqslant 3 \tag{34a}$$
$$n_{b,k} = n_{a,k} + 2R \quad \text{with} \quad R = 1, 2, \ldots, R_{\max} \leqslant 4 \tag{34b}$$
$$n_{i,k} = n_{b,k} \tag{34c}$$

Hence, in each local frequency band $[k-n, k+n]$, at most 16 different local models are estimated, and the one minimizing (33) is selected.

Note that the model selection can be performed at the level of the initial [LL (19) and LGTLS (21) methods] or the final [LSK (20), LBTLS (22) and LOE (18) methods] estimates. While the second option is, from a statistical point of view, better than the first option, it takes (much) more computation time. Practice shows that the model selection based on the starting values performs well and, therefore, this solution is preferred.

Finally, the width $2n+1$ of the local frequency band $[k-n, k+n]$ is chosen to be the same for all frequencies k. It is determined by the minimal degrees of freedom $\mathrm{dof}_{\min}$ set by the user. From (25), we derive

$$n = \lceil 0.5\big(\mathrm{dof}_{\min} + (n_{\theta,k})_{\max} - 1\big)\rceil \tag{35}$$

with $\lceil x \rceil$ the smallest integer larger than or equal to x, and where $(n_{\theta,k})_{\max}$ is the number of model parameters of the most complex model in (34) and equals $3n_{a,max} + 4R_{max} + 2$ [(24d) with $n_{a,k} = n_{a,\max}$ and $n_{b,k} = n_{i,k} = n_{a,\max} + 2R_{\max}$]. Due to the overlap of the local frequency bands, the local FRF and noise variance estimates are each correlated over the frequency. With the choice (35), the correlation length over the frequency k of these nonparametric estimates equals $\pm 2n$ at all frequencies. It explains the apparent contradiction that the total number of estimated parameters $\sum_{k=1}^{N/2-1} n_{\theta,k}$, with $n_{\theta,k}$ defined in (24d), is much larger than the total amount of data $N/2-1$.

C. Extensions

Provided that a known reference signal $r(t)$ is available, noisy input $u(t)$, noisy output $y(t)$ measurements of a system $G(s)$ subject to process noise $w(t)$ and operating in feedback [see Fig. 4] can be handled using the procedure of Section V-B. Such a reference signal is typically the signal stored in an arbitrary waveform generator. In the first step, the FRFs from the known reference $r(t)$ to the noisy input $u(t)$ and the noisy output $y(t)$ are estimated. These are open loop problems satisfying the assumptions of Section IV-A. In the second step, the FRF from input to output is obtained as the ratio of the FRF from reference to output to the FRF from reference to input.

If the system behaves nonlinearly, then the proposed procedure estimates the best linear approximation [see [25] for the definition of the best linear approximation and the class of nonlinear systems considered]. Its variance depends on the input-output measurement noise, the process noise, and the nonlinear distortions due to the nonlinear input-output and nonlinear input-process noise interactions [see [25] for the details].

VI. Simulation Example

A simulation example was chosen to illustrate the above theoretical results. Therefore, note that the bias of the LSK and LBTLS methods is larger than that of the LOE estimator [see Theorems 1 and 2]. Next, LL requires (much) less calculation time than LGTLS, which is important for the model order selection procedure [see Sections V-A and V-B]. Finally, we want to compare the performance of the local polynomial and the mixed local polynomial-local rational modeling approaches. Therefore, the LP, LL and LOE estimates are calculated for a lightly damped sixteenth order continuous-time system. Note that the LL estimates are used for initializing the minimization of the LOE cost function (18).

Each pole of the system has a damping ratio $\zeta = 1.25 \times 10^{-3}$, and the eight resonance peaks have frequencies f_m^{res}, $m = 1, 2, \ldots, 8$, and corresponding 3 dB bandwidth $\Delta_m^{3\mathrm{dB}} \approx 2\zeta f_m^{\mathrm{res}}$ [see Table II]. The simulated transfer function $G(s)$

TABLE II
Resonance Frequencies f_m^{res} and Corresponding 3 dB Bandwidth $\Delta_m^{3\mathrm{dB}}$.

m	$f_m^{\mathrm{res}}(Hz)$	$\Delta_m^{3\mathrm{dB}}(Hz)$
1	104.4	0.26
2	203.9	0.51
3	303.3	0.76
4	402.8	1.0
5	502.2	1.3
6	601.7	1.5
7	701.1	1.8
8	800.6	2.0

has the form

$$G(s) = \sum_{m=-8,\, m\neq 0}^{8} \frac{-10 s_m^2 R_m}{s - s_m} \tag{36}$$

[see Fig. 5], with $s_m = 2\pi f_m^{\mathrm{res}}(-\zeta + j\sqrt{1-\zeta^2})$, $s_{-m} = \overline{s_m}$, $R_{-m} = \overline{R_m}$, and where R_m is a uniformly $[-1, 1]$ distributed random variable [the Matlab random generator is started from its default value, rng('default'), which is the Mersenne Twister with seed 0].

The noisy transient ~~transient~~ response $y(t)$ of the continuous-time system $G(s)$ (36) to a band-limited (zero

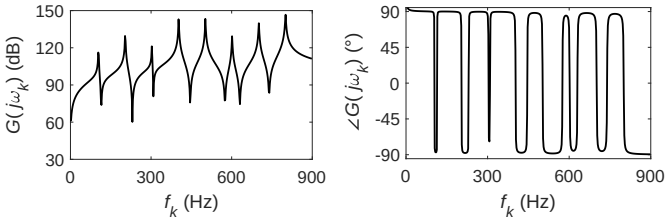

Fig. 5. True amplitude [left] and phase [right] at the frequency resolution $f_s/N \approx 1.95\,\text{Hz}$ of the simulation – High system order, white disturbing noise example.

power spectral density for $|f| > 900\,\text{Hz}$), continuous-time white Gaussian noise excitation $u(t)$, with zero mean value and standard deviation 0.0210, is obtained as follows. First, one period of 2048 samples of the steady state response of $G(s)$ to periodic white Gaussian noise with a bandwidth of 900 Hz is calculated at the rate $f_s = 2\,\text{kHz}$. Next, zero mean, discrete-time filtered white Gaussian noise $n_y(t)$ is added at the sampling instances. The noise transfer function is a second order Chebyshev filter with a passband ripple of 10 dB and a cutoff frequency of $0.8 f_s/2$. The discrete-time white noise standard deviation σ is chosen such that the output signal-to-noise ratio (SNR) is either 30 dB (large noise level: $\sigma = 2\sqrt{210^3}$) or 50 dB (low noise level: $\sigma = 2\sqrt{210^2}$). Finally, starting from the first $N = 1024$ samples of the known input $u(t)$ and the noisy output $y(t)$, the local polynomial (LP), the local Levy (LL) and local output error (LOE) estimates of the FRF and its variance are calculated for the $F = 460$ excited frequencies in the band $[2\,\text{Hz}, 900\,\text{Hz}]$ with the following parameters in equations (34) and (35)

$$\text{LP}: n_{a,\max} = 0,\ R_{\max} = 3 \text{ and } \text{dof}_{\min} = 10 \quad (37a)$$

$$\text{LL, LOE}: n_{a,\max} = 3,\ R_{\max} = 2 \text{ and } \text{dof}_{\min} = 10 \quad (37b)$$

The corresponding local bandwidths are

$$\text{LP}: 2n+1 = 25 \ \Rightarrow\ \Delta f = 48.8\,\text{Hz} \quad (38a)$$

$$\text{LL, LOE}: 2n+1 = 29 \ \Rightarrow\ \Delta f = 56.6\,\text{Hz} \quad (38b)$$

which is obtained by combining (24d), (34), (35) and (37). The top row of Fig. 6 shows the orders of the local rational models selected by the procedure of Section V-A for one random realization of the input and the disturbing noise. One could be surprised that the most complex models (two poles) are not located at the resonances but in between. It just indicates that in between resonances, more than one peak significantly contributes to the FRF. Note also that the LL method selects a polynomial model in local frequency bands where the variation of the FRF is limited. As such, the LL model selection procedure combines the low bias of the rational approximation with the low noise sensitivity of the polynomial model. The whole procedure is repeated for 1000 independent random realizations of the excitation $u(t)$ and the disturbing noise $n_y(t)$, and this for an output SNR of 30 dB and 50 dB.

Fig. 6 compares the sample mean of the 1000 LP and LOE estimates to the true value for an output SNR of 30 dB. It can be seen that the LP method completely fails in the neighborhood of the resonance peaks. This poor performance

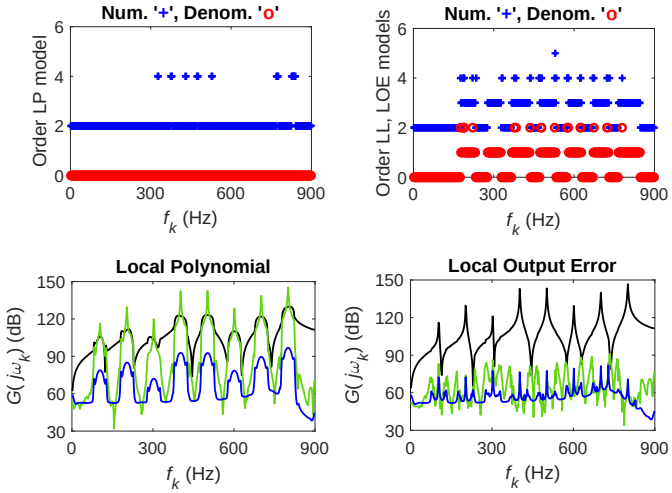

Fig. 6. Local polynomial [LP: left column] and local output error [LOE: right column] estimates of the sixteenth order continuous-time system (36) for an output SNR of 30 dB. Top row: model complexity selected by LP [left] and local Levy [LL: right] methods for one realization of the 1000 Monte-Carlo runs. Bottom row: sample mean local FRF estimates [black] over 1000 Monte-Carlo runs. Blue: sample variance of the sample mean. Green: magnitude of the complex bias error of the sample mean.

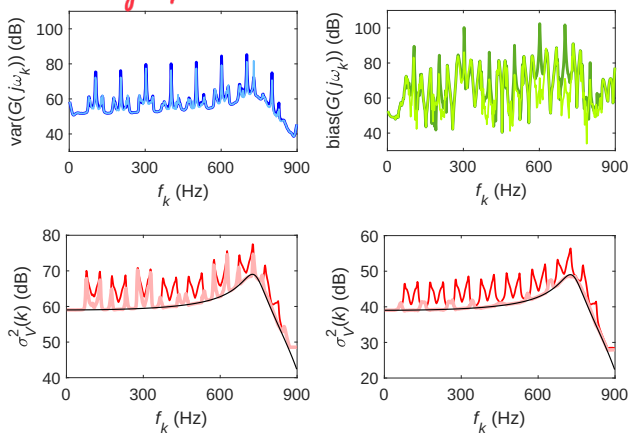

Fig. 7. Comparison of the sample variances [blue] and biases [green] of the mean FRF estimates and of the mean output variance estimates [red]. Dark color: local Levy estimate. Light color: local output error estimate. Black: true value. Top row and bottom left: SNR of 30 dB. Bottom right: SNR of 50 dB.

is due to the fact that no or just one DFT frequency lies within the 3 dB bandwidths of the resonances [compare the frequency resolution $f_s/N \approx 1.95\,\text{Hz}$ with the 3 dB bandwidths $\Delta_i^{3\text{dB}}$ in Table II], while at least seven frequencies are needed [24]. Despite the very small frequency resolution, the LOE estimates perform quite well. Although in between the resonances the LOE bias error of the sample mean [green line] is almost everywhere larger than the variance error [blue line], for one realization, the variance is the dominant error source [$10\log_{10}(1000) = 30\,\text{dB}$ larger than the blue line].

A detailed comparison of the local Levy (LL) and local output error (LOE) estimates is shown in Fig. 7 for an SNR of 30 dB [top row and bottom left] and 50 dB [bottom right]. At the resonance frequencies, the sample variance [top left plot] and bias [top right plot] of the mean LOE estimate is

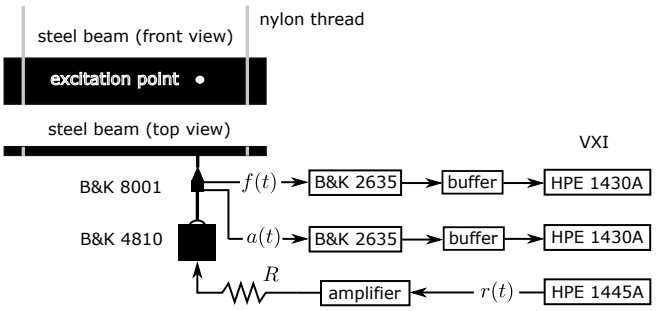

Fig. 8. Experimental setup for measuring the flexural vibrations of a steel beam. The input of the mini-shaker [B&K 4810] is controlled by the output $r(t)$ of a VXI arbitrary waveform generator [HPE 1445A, $Z_{\text{out}} = 50\,\Omega$] via an amplifier and a series resistance $R = 20\,\Omega/5\,\text{W}$. The force $f(t)$ and acceleration $a(t)$ signals, measured with an impedance head [B&K 8001], are amplified [B&K 2635] and buffered [$Z_{\text{in}} > 5\,M\Omega$, $Z_{\text{out}} = 50\,\Omega$], before being applied to the VXI data acquisition channels [HP 1430A, $Z_{\text{in}} = 50\,\Omega$].

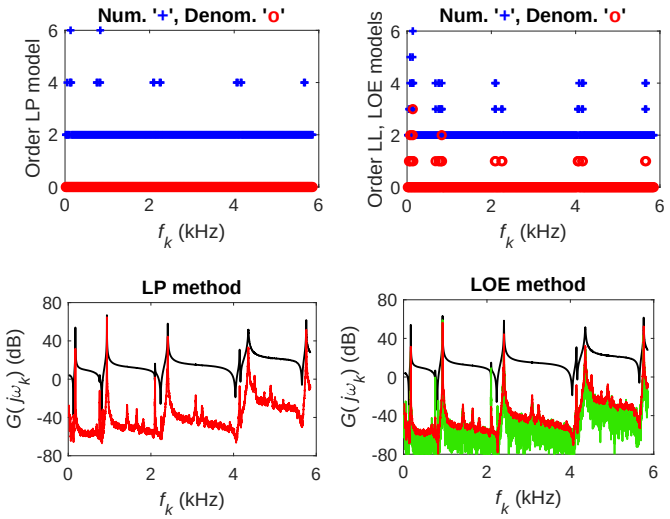

Fig. 9. Local polynomial [LP: left column] and local output error [LOE: right column] modeling of the full steal beam data [$f_k = kf_{\text{s}}/N_{\text{full}}$ with $f_{\text{s}} = 19.531\,\text{kHz}$ and $N_{\text{full}} = 50 \times 1024$]. Top row: model complexity selected by the LP [left] and the local Levy [LL: right] methods. Bottom row: estimate of the FRF [black] and its variance [red], and magnitude of the complex difference $|\hat{G}^{\text{full}}_{\text{LP}}(j\omega_k) - \hat{G}^{\text{full}}_{\text{LOE}}(j\omega_k)|$ between the LP and LOE estimates [green].

(significantly) smaller than that of the LL method, while it is the other way around at a few other frequencies. Hence, the LOE estimate of the FRF is not uniformly better than that of the LL method. The bottom row of Fig. 7 shows the LOE and LL output noise variance estimates $\hat{\sigma}^2_V(k)$ for an output SNR of 30 dB [left] and 50 dB [right]. From the left plot, a (significant) bias error can be observed in the neighborhood of the resonances for both estimates. However, the bias of the LOE estimate is everywhere (much) smaller than that of the LL method. Comparing the bottom plots of Figure 7, it can be seen that the relative bias error $(\hat{\sigma}^2_V(k) - \sigma^2_V(k))/\sigma^2_V(k)$ of the LOE estimate decreases with increasing SNR, which is not the case for the LL method. It indicates that the second term in (23) dominates the bias error (26) for the LOE estimate, which is not the case for the LL method.

Note that the non-zero bias errors of the LOE estimates in Figures 6 and 7 are due to the non-zero ratio n/N and the finite number of frequencies $2n+1$ in the local frequency bands, while those of the LL method remain non-zero, even if $n \to \infty$ and $n/N \to 0$ [see Theorems 1 and 2 on, respectively, pages 5 and 6].

VII. Measurement Example

Fig. 8 shows the experimental setup for measuring the flexural vibrations of a steel beam. The steel beam of density $7800\,\text{kg/m}^3$ and size $61\,\text{cm} \times 2.17\,\text{cm} \times 4.93\,\text{mm}$, is hung by two nylon threads to the ceiling. Proceeding in this way the impact of external effects on the setup are minimized [free-free boundary conditions]. A transverse force is applied at 10 cm from the end of the beam by means of a mini-shaker [B&K 4810]. The output signal of an arbitrary waveform generator is first amplified before connecting it resistively to the input of the mini-shaker. This series resistor is inserted to avoid the adverse effect of the inductive loading of the amplifier by the mini-shaker. Finally, the force (input) and acceleration (output) signals, measured with an impedance head [B&K 8001], are amplified and buffered before they are sampled at the rate $f_{\text{s}} = 10\,\text{MHz}/2^9 \approx 19.531\,\text{kHz}$ by the alias protected data acquisition channels. Synchronized sampling of the signals is guaranteed by an internal 10 MHz master clock that synchronizes the generator and acquisition boards.

A random binary sequence $r(t)$ of length $N_{\text{full}} = 50 \times 1024$ is applied to the mini-shaker resulting in input (force) and output (acceleration) signals of, respectively, $u = 242\,\text{mV rms}$ and $y = 256\,\text{mV rms}$. Starting from this full data set, the local polynomial (LP) and local output error (LOE) estimates of the FRF and its variance are calculated with the following parameters in (34) and (35)

$$\text{LP}: n_{a,\max} = 0,\ R_{\max} = 3 \text{ and } \text{dof}_{\min} = 10 \tag{39a}$$
$$\text{LOE}: n_{a,\max} = 2,\ R_{\max} = 2 \text{ and } \text{dof}_{\min} = 10 \tag{39b}$$

Since the modeling starts from the known reference signal $r(t)$ and the noisy input $u(t)$, noisy output $y(t)$ observations, the procedure of Section V-C is followed. Therefore, the local bandwidths corresponding to (39) are

$$\text{LP}: 2n+1 = 25 \ \Rightarrow\ \Delta f = 9.5\,\text{Hz} \tag{40a}$$
$$\text{LOE}: 2n+1 = 29 \ \Rightarrow\ \Delta f = 11.1\,\text{Hz} \tag{40b}$$

This is obtained by combining (35) and (39) with $n_{\theta,k} = n_{a,k} + 2(n_{b,k} + n_{i,k} + 2)$ and $n = \lceil 0.5\big(\text{dof}_{\min} + (n_{\theta,k})_{\max}/2 - 1\big)\rceil$.

The local Levy (LL) estimates are used for initializing the local output error (LOE) method and for selecting the local rational model order. Fig. 9 shows the LP and LOE estimates and the selected local polynomial and local rational model orders. Eight resonances at f^{res}_i, $i = 1, 2, \ldots, 8$, can be distinguished:

$$f^{\text{res}}_1 = 17.3\,\text{Hz},\ f^{\text{res}}_2 = 760\,\text{Hz},\ f^{\text{res}}_3 = 936\,\text{Hz}, \tag{41a}$$
$$f^{\text{res}}_4 = 2.10\,\text{kHz},\ f^{\text{res}}_5 = 2.41\,\text{kHz},\ f^{\text{res}}_6 = 4.15\,\text{kHz}, \tag{41b}$$
$$f^{\text{res}}_7 = 4.36\,\text{kHz},\ f^{\text{res}}_8 = 5.76\,\text{kHz} \tag{41c}$$

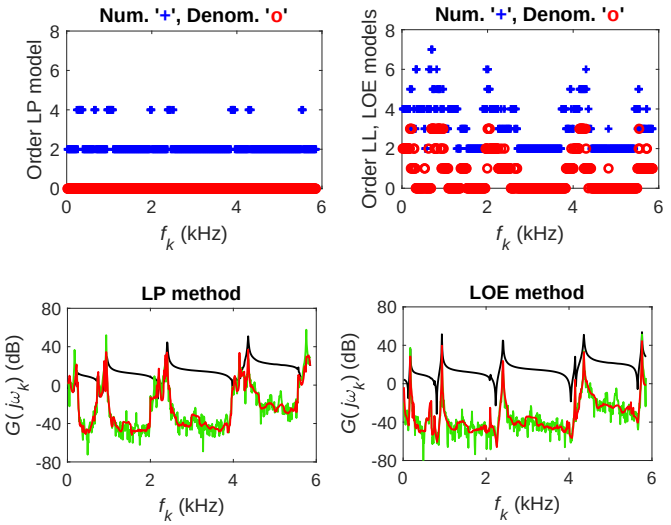

Fig. 10. Local polynomial [LP: left column] and local output error [LOE: right column] modeling of the short steal beam data [$f_k = kf_s/N_{\text{short}}$ with $f_s = 19.531\,\text{kHz}$ and $N_{\text{short}} = 2 \times 1024$]. Top row: model complexity selected by the LP [left] and the local Levy [LL: right] methods. Bottom row: LP $\hat{G}_{\text{LP}}^{\text{short}}(j\omega_k)$ and LOE $\hat{G}_{\text{LOE}}^{\text{short}}(j\omega_k)$ estimates [black] and their variance [red], and the magnitude of the complex difference [green] between the full and short data estimates $|\hat{G}_{\text{LP}}^{\text{short}}(j\omega_k) - \hat{G}_{\text{LOE}}^{\text{full}}(j\omega_k)|$ [left] and $|\hat{G}_{\text{LOE}}^{\text{short}}(j\omega_k) - \hat{G}_{\text{LOE}}^{\text{full}}(j\omega_k)|$ [right].

From the bottom right plot of Fig. 9 it can be seen that the LP estimates perform equally well as the LOE estimates [the green line is below the red line], except in the neighborhood of the anti-resonances and the resonances f_1^{res}, f_2^{res} and f_3^{res}. Although mostly a polynomial of order two is selected by the LP and LL methods [see the top row], the LP and LOE estimates are not exactly the same. This is due to the fact that the local bandwidths (40) are different. Similar to the simulation example, the most complex models appear in the neighborhood of, but not exactly at, the resonances, and LL selects polynomial models in local frequency bands with a restricted variation of the FRF. The high variability of the FRF estimates at the resonance frequencies originates from the poor input signal-to-noise ratio, since at those frequencies almost no energy can be injected into the system.

To verify the performance of the local modeling techniques on short data records, the last $N_{\text{short}} = 2 \times 1024$ samples of the response to the random binary excitation are used for estimating the FRF and its variance, and the results are compared to those on the full data record [$N_{\text{full}} = 50 \times 1024$ samples]. Starting from the short data set, the local polynomial (LP) and local output error (LOE) estimates of the FRF and its variance are calculated with the following parameters in (34) and (35)

$$\text{LP} : n_{a,\max} = 0,\ R_{\max} = 3 \text{ and } \text{dof}_{\min} = 6 \tag{42a}$$
$$\text{LOE} : n_{a,\max} = 3,\ R_{\max} = 3 \text{ and } \text{dof}_{\min} = 6 \tag{42b}$$

and the corresponding local bandwidths are

$$\text{LP} : 2n+1 = 21 \ \Rightarrow\ \Delta f = 200\,\text{Hz} \tag{43a}$$
$$\text{LOE} : 2n+1 = 29 \ \Rightarrow\ \Delta f = 277\,\text{Hz} \tag{43b}$$

Fig. 10 shows the LP and LOE estimates and the selected local polynomial and local rational model orders. From the bottom row it can be seen that the LP estimates fail completely in the neighborhood of almost all resonances, all except the fifth and seventh, and all anti-resonances. The LOE correctly estimates six out of the eight resonances (41) and all anti-resonances. However, since the bias errors of the LOE estimates at the first and last resonances are (almost) as large as the estimates themselves, we conclude that the measurement time of the short data record is too small for getting reliable results at these frequencies.

Note that the estimated variances (red lines) predict quite well the difference between the short and full data FRF estimates (green lines). Hence, the user gets automatically a warning in which frequency band(s) the results are unreliable. However, since the fraction outside the 95% confidence bound (32) of the LP and LOE FRF estimates equals, respectively, 22% and 16% instead of 5%, the error bounds generated for the short data set are somewhat too small in this measurement example.

From the top right plots of Figs. 9 and 10, it can be concluded that the LOE estimates of the short data record require much more complex local rational models than those of the full data record. This is due to the much larger local frequency band of the short data record (43b) compared with that of the full record (40b).

Finally, it has been observed that the local Levy estimates (not shown here) of the full and short data records are quite similar to the LOE estimates with the difference of an overestimated noise variance when $n_{a,k} > 0$.

VIII. Conclusion

As already mentioned in [11] for the local output error (LOE) method and in [13] for the local bootstrapped total least squares, an appropriate model order selection procedure is a key element for the success of all local rational modeling methods. Another key element is the choice of the model structure (34), where the numerator order of the local rational function is larger than the denominator order. All conclusions concerning the local modeling techniques strongly rely on these two elements.

The LOE method combined with the automatic model selection procedure of Section V-B is a valid alternative for the local polynomial (LP) approach, because it optimally integrates the small bias error of the rational approximation with the low noise sensitivity of the polynomial approximation. If the frequency resolution of the measurement [= inverse of the experiment duration] is such that less than seven DFT frequencies lie within the 3 dB bandwidths of the resonances, then the bias reduction of LOE compared with LP is significant. Hence, to achieve in these cases a user specified accuracy, the required LOE measurement time will be much smaller than that of the LP method.

Initializing LOE with the local Levy (LL) estimate, and selecting the local rational model order via the LL method, works well and is a good compromise between calculation time and statistical performance. In most cases the LL and

LOE FRF estimators perform equally well. Only at the resonance peaks of lowly damped poles, the root mean squared error of LL can be significantly larger. However, the major difference between both methods is the (much) larger bias of the LL noise variance estimate. Since a reliable uncertainty bound is as important as the FRF estimate itself, and since an (asymptotically) unbiased estimate of the noise variance is essential for parametric transfer function modeling [see [9], Chapters 10 and 12], LOE is preferred over LL.

APPENDIX A
PROOF OF THEOREM 1

The first term in the right hand side of (23) originates from the polynomial and rational approximation errors in (7).

The second term in the right hand side of (23) reflects the noise contribution to the bias error. Under Assumptions 1 and 2, this noise bias error is zero ($\beta = \infty$) for the LP estimator (17) because it is a linear least squares problem with noiseless regression matrix. For all the other local methods the bias is a consequence of their stochastic behavior for increasing values of the number of frequencies $2n+1$ in the local frequency band. Under Assumptions 1 and 2:

1) LL (19) and LSK (20) are inconsistent estimators [see [9], Section 9.8, pp. 301–305] and, therefore, are biased. Hence, $\beta = 0$ in (23).
2) LOE (18) is a consistent estimator. If in addition the disturbing noise variance is constant in the local frequency band, then also the LGTLS (21) and LBTLS (22) estimators are consistent. Hence, the biases of the LOE, LGTLS and LBTLS estimates decrease to zero as an $O(1/(2n+1))$ [see [9], Theorem 9.21, pp. 298–299], and $\beta = 1$ in (23).

Assuming that the rational approximation errors can be neglected, we show in the remainder of this appendix that the noise bias term is proportional to the noise-to-signal ratio and prove the bias contribution of the approximation $\sigma_V^2(k+r) \approx \sigma_V^2(k)$ for the LGTLS and LBTLS estimators.

The (noise) bias terms in Sections A-A to A-C are calculated for $2n+1 \to \infty$. Given the relationship between the FRF and the rational model, $G(j\omega_k) = b_{0,k}/a_{0,k}$ (9), and the fact that the transient parameters are asymptotically ($2n+1 \to \infty$) zero, it is sufficient to perform the analysis for the a- and b-coefficients only.

In this appendix, $f'(\theta_k)$ denotes the derivative of $f(\theta_k)$ w.r.t. the local parameter vector $(\theta_k)_{\mathrm{re}}$, where $(X)_{\mathrm{re}}$ stacks the real and imaginary parts of X on top of each other

$$(X)_{\mathrm{re}} = \begin{pmatrix} \mathrm{Re}(X) \\ \mathrm{Im}(X) \end{pmatrix}. \tag{44}$$

By extension, $f''(\theta_k)$ is the second order derivative of $f(\theta_k)$ w.r.t. $(\theta_k)_{\mathrm{re}}$.

A. Noise Bias LL and LSK

We calculate the noise bias term of the LL method (19). The calculation for the LSK estimator (20) follows exactly the same lines.

Under Assumptions 1 and 2, the expected value of the LL cost (19a) equals

$$\begin{aligned} V_{\mathrm{LL}}(\theta_k) &= \mathbb{E}\{V_{\mathrm{LL}}(\theta_k, Z_k)\} \\ &= \mathbb{E}\{V_{\mathrm{LL}}(\theta_k, Z_{k,0})\} \\ &\quad + \sum_{r=-n}^{n} \sigma_V^2(k+r)|A_k^{\mathrm{c}}(j\omega_r, \theta_k)|^2 \end{aligned} \tag{45}$$

where $Z_{k,0} = \mathbb{E}\{Z_k\}$. A first order Taylor series expansion at the true parameter value $\theta_{k,0}$ of $V'_{\mathrm{LL}}(\theta_k)$, the derivative of (45) w.r.t. $(\theta_k)_{\mathrm{re}}$, gives

$$V'^T_{\mathrm{LL}}(\theta_k) \approx V'^T_{\mathrm{LL}}(\theta_{k,0}) + V''_{\mathrm{LL}}(\theta_{k,0})(\theta_k - \theta_{k,0})_{\mathrm{re}} \tag{46}$$

where, using $\mathbb{E}\{V'_{\mathrm{LL}}(\theta_{k,0}, Z_{k,0})\} = 0$,

$$\begin{aligned} V'^T_{\mathrm{LL}}(\theta_{k,0}) &= \sum_{r=-n}^{n} \sigma_V^2(k+r)\big(|A_k^{\mathrm{c}}(j\omega_r, \theta_{k,0})|^2\big)'^T \\ &= \sigma_V^2(k) O(2n+1) \end{aligned} \tag{47}$$

$$\begin{aligned} V''_{\mathrm{LL}}(\theta_{k,0}) &\approx \mathbb{E}\{V''_{\mathrm{LL}}(\theta_{k,0}, Z_{k,0})\} \\ &\approx \mathbb{E}\{|Y_0(k)|^2\} O(2n+1) \end{aligned} \tag{48}$$

Evaluating (46) at $\tilde{\theta}_k$, the minimizer of (45), and using $V'^T_{\mathrm{LL}}(\tilde{\theta}_k) = 0$, we get an expression for the bias $\tilde{\theta}_k - \theta_{k,0}$ of the parameter estimate

$$(\tilde{\theta}_k - \theta_{k,0})_{\mathrm{re}} \approx -V''^{-1}_{\mathrm{LL}}(\theta_{k,0}) V'^T_{\mathrm{LL}}(\theta_{k,0}) \tag{49}$$

Combining (9) with (47)–(49) proves the noise bias term in (23).

B. Noise Bias LOE, LGTLS and LBTLS

We derive a bias expression for the LOE estimator (18) under Assumptions 1 and 2. The derivation for the LGTLS (21) and LBTLS (22) estimators, under the additional condition that the disturbing noise variance is white in the local frequency band $[k-n, k+n]$, follows exactly the same lines.

A first order Taylor series expansion with remainder of the derivative $V'_{\mathrm{LOE}}(\theta_k, Z_k)$ of the LOE cost (18) at the true parameter value $\theta_{k,0}$, gives

$$V'^T_{\mathrm{LOE}}(\theta_k, Z_k) = V'^T_{\mathrm{LOE}}(\theta_{k,0}, Z_k) + V''_{\mathrm{LOE}}(\widehat{\theta}_k, Z_k)(\theta_k - \theta_{k,0})_{\mathrm{re}} \tag{50}$$

where $\widehat{\theta}_k = t\theta_k + (1-t)\theta_{k,0}$ with $t \in [0,1]$. Evaluating (50) at $\hat{\theta}_k$, the minimizer of (18), and using $V'_{\mathrm{LOE}}(\hat{\theta}_k, Z_k) = 0$, we get an exact expression for the difference $\hat{\theta}_k - \theta_{k,0}$

$$(\hat{\theta}_k - \theta_{k,0})_{\mathrm{re}} = -V''^{-1}_{\mathrm{LOE}}(\widehat{\theta}_k, Z_k) V'^T_{\mathrm{LOE}}(\theta_{k,0}, Z_k) \tag{51}$$

Using the strong law of large numbers for $2n+1 \to \infty$, the right hand side of (51) can be elaborated as

$$(\hat{\theta}_k - \theta_{k,0})_{\mathrm{re}} = \delta_{\theta_k}(Z_k) + b_{\theta_k}(Z_k) \tag{52a}$$

$$\begin{aligned} \delta_{\theta_k}(Z_k) &= -V''^{-1}_{\mathrm{LOE}}(\theta_{k,0}) V'^T_{\mathrm{LOE}}(Z_k, \theta_{k,0}) \\ &= O((2n+1)^{-1/2}) \end{aligned} \tag{52b}$$

$$\begin{aligned} b_{\theta_k}(Z_k) &= \big[V''^{-1}_{\mathrm{LOE}}(\theta_{k,0}) - V''^{-1}_{\mathrm{LOE}}(\widehat{\theta}_k, Z_k)\big] V'^T_{\mathrm{LOE}}(\theta_{k,0}, Z_k) \\ &= O((2n+1)^{-1}) \end{aligned} \tag{52c}$$

with $V_{\mathrm{LOE}}(\theta_k) = \mathbb{E}\{V_{\mathrm{LOE}}(\theta_k, Z_k)\}$, $\delta_{\theta_k}(Z_k)$ the zero-mean random contribution [$\mathbb{E}\{\delta_{\theta_k}(Z_k)\} = 0$], and with $b_{\theta_k}(Z_k)$ the

non-zero mean random part [proof: see [9], Theorem 17.21 and Corollary 17.25 on pp. 634 and 635, respectively].

Via a first order Taylor series expansion of $V''_{\mathrm{LOE}}(\widehat{\theta}_k, Z_k)$ at $\theta_{k,0}$, the expression of $b_{\theta_k}(Z_k)$ (52c) can be refined. After some calculations we find that

$$\mathbb{E}\{b_{\theta_k}(Z_k)\} = \frac{\sigma_V^2(k)}{\mathbb{E}\{|Y_0(k)|^2\}} O\big(\frac{1}{2n+1}\big) \tag{53}$$

[proof: see [9], Appendix 19.D, part (ii), on pp. 689–690]. Combining (9) with (53) proves the noise bias term in (23).

C. Bias LGTLS and LBTLS Caused by the Local White Noise Approximation

We derive an expression for the bias of the LGTLS estimator (21) due to the approximation $\sigma_V^2(k+r) \approx \sigma_V^2(k)$ in the local frequency band $[k-n, k+n]$. The bias term of the LBTLS method (22) is obtained in a similar way.

Under Assumptions 1 and 2, the expected value of the LGTLS cost (21) equals

$$\begin{aligned} V_{\mathrm{LGTLS}}(\theta_k) &= \mathbb{E}\{V_{\mathrm{LGTLS}}(\theta_k, Z_k)\} \\ &= \mathbb{E}\{V_{\mathrm{LGTLS}}(\theta_k, Z_{k,0})\} \\ &\quad + \frac{\sum_{r=-n}^{n} \sigma_V^2(k+r)|A_k^{\mathrm{c}}(j\omega_r, \theta_k)|^2}{\sum_{r=-n}^{n} \sigma_V^2(k)|A_k^{\mathrm{c}}(j\omega_r, \theta_k)|^2} \end{aligned} \tag{54}$$

with $Z_{k,0} = \mathbb{E}\{Z_k\}$. Under Assumption 2, the noise variance $\sigma_V^2(k)$ is a smooth function of the frequency. Hence, in the local frequency band $[k-n, k+n]$, $\sigma_V^2(k+r)$ can be approximated as

$$\sigma_V^2(k+r) \approx \sigma_V^2(k) + (\sigma_V^2(k))' \frac{r}{N} \tag{55}$$

where $()'$ is the derivative w.r.t. k. Combining (54) and (55), gives

$$\begin{aligned} V_{\mathrm{LGTLS}}(\theta_k) &\approx \mathbb{E}\{V_{\mathrm{LGTLS}}(\theta_k, Z_k)\} + 1 \\ &\quad + \frac{(\sigma_V^2(k))'}{N\sigma_V^2(k)} \frac{\sum_{r=-n}^{n} r|A_k^{\mathrm{c}}(j\omega_r, \theta_k)|^2}{\sum_{r=-n}^{n} |A_k^{\mathrm{c}}(j\omega_r, \theta_k)|^2} \end{aligned} \tag{56}$$

Expanding $V'_{\mathrm{LGTLS}}(\theta_k)$, the derivative of (56) w.r.t. $(\theta_k)_{\mathrm{re}}$, in a first order Taylor series at the true parameter value $\theta_{k,0}$, we get

$$V'^T_{\mathrm{LGTLS}}(\theta_k) \approx V'^T_{\mathrm{LGTLS}}(\theta_{k,0}) + V''_{\mathrm{LGTLS}}(\theta_{k,0})(\theta_k - \theta_{k,0})_{\mathrm{re}} \tag{57}$$

where, using $\mathbb{E}\{V'_{\mathrm{LGTLS}}(\theta_{k,0}, Z_{k,0})\} = 0$,

$$V'^T_{\mathrm{LGTLS}}(\theta_{k,0}) \approx \frac{(\sigma_V^2(k))'}{\sigma_V^2(k)} O\Big(\frac{n}{N}\Big) \tag{58}$$

$$\begin{aligned} V''_{\mathrm{LGTLS}}(\theta_{k,0}) &\approx \mathbb{E}\{V''_{\mathrm{LGTLS}}(\theta_{k,0}, Z_{k,0})\} \\ &\approx \mathbb{E}\{|Y_0(k)|^2\} O(2n+1) \end{aligned} \tag{59}$$

Evaluating (57) at $\tilde{\theta}_k$, the minimizer of (54), and using $V'^T_{\mathrm{LGTLS}}(\tilde{\theta}_k) = 0$, we get an expression for the bias $\tilde{\theta}_k - \theta_{k,0}$ of the parameter estimate

$$(\tilde{\theta}_k - \theta_{k,0})_{\mathrm{re}} \approx -V''^{-1}_{\mathrm{LGTLS}}(\theta_{k,0}) V'^T_{\mathrm{LGTLS}}(\theta_{k,0}) \tag{60}$$

Combining (9) with (58)–(60) proves the third bias term in (23).

APPENDIX B
PROOF OF THEOREM 2

The proof of (26) is given in [9], Appendix 7.D, pp. 271–272, for the LP method. In this appendix we prove (26) for the LOE, LL, LSK, LGTLS and LBTLS estimators. First, we derive an approximate expression for the output residual $Y(k+r) - \hat{Y}(k+r)$ [Section B-A]. Next, using this expression, the expected value of the local noise variance estimate is analyzed [Section B-B].

A. Approximate Expression for the Output Residual

For all local rational methods, the estimated polynomials $X_k^{\mathrm{c}}(j\omega_r, \hat{\theta}_k)$, $X = A$, B, and I, can be split into three contributions: the true value $X_{k,0}^{\mathrm{c}}(j\omega_r)$, the bias error $X_{k,\mathrm{bias}}^{\mathrm{c}}(j\omega_r)$, and the zero-mean noise contribution $X_{k,\mathrm{noise}}^{\mathrm{c}}(j\omega_r) = O((2n+1)^{-1/2})$

$$X_k^{\mathrm{c}}(j\omega_r, \hat{\theta}_k) = X_{k,0}^{\mathrm{c}}(j\omega_r) + X_{k,\mathrm{bias}}^{\mathrm{c}}(j\omega_r) + X_{k,\mathrm{noise}}^{\mathrm{c}}(j\omega_r) \tag{61}$$

Using $\frac{Y+y}{X+x} \approx \frac{Y}{X}(1 + \frac{y}{Y} - \frac{x}{X})$ for the rational functions, the predicted output (24b) can be approximated as

$$\hat{Y}(k+r) \approx Y_0(k+r) + Y_{\mathrm{bias}}(k+r) + Y_{\mathrm{noise}}(k+r) \tag{62a}$$

with $Y_0(k+r)$ the true value, $Y_{\mathrm{bias}}(k+r)$ the bias due to the noise and the local rational approximations

$$\begin{aligned} Y_{\mathrm{bias}}(k+r) = \\ G_{k,0}^{\mathrm{c}}(j\omega_r)U(k+r)\big(\frac{B_{k,\mathrm{bias}}^{\mathrm{c}}(j\omega_r)}{B_{k,0}^{\mathrm{c}}(j\omega_r)} - \frac{A_{k,\mathrm{bias}}^{\mathrm{c}}(j\omega_r)}{A_{k,0}^{\mathrm{c}}(j\omega_r)}\big) \\ + T_{k,0}^{\mathrm{c}}(j\omega_r)\big(\frac{I_{k,\mathrm{bias}}^{\mathrm{c}}(j\omega_r)}{I_{k,0}^{\mathrm{c}}(j\omega_r)} - \frac{A_{k,\mathrm{bias}}^{\mathrm{c}}(j\omega_r)}{A_{k,0}^{\mathrm{c}}(j\omega_r)}\big) \end{aligned} \tag{62b}$$

and $Y_{\mathrm{noise}}(k+r) = O((2n+1)^{-1/2})$ the zero-mean noise contribution

$$\begin{aligned} Y_{\mathrm{noise}}(k+r) = \\ G_{k,0}^{\mathrm{c}}(j\omega_r)U(k+r)\big(\frac{B_{k,\mathrm{noise}}^{\mathrm{c}}(j\omega_r)}{B_{k,0}^{\mathrm{c}}(j\omega_r)} - \frac{A_{k,\mathrm{noise}}^{\mathrm{c}}(j\omega_r)}{A_{k,0}^{\mathrm{c}}(j\omega_r)}\big) \\ + T_{k,0}^{\mathrm{c}}(j\omega_r)\big(\frac{I_{k,\mathrm{noise}}^{\mathrm{c}}(j\omega_r)}{I_{k,0}^{\mathrm{c}}(j\omega_r)} - \frac{A_{k,\mathrm{noise}}^{\mathrm{c}}(j\omega_r)}{A_{k,0}^{\mathrm{c}}(j\omega_r)}\big) \end{aligned} \tag{62c}$$

Taking into account (62), the output residual $\hat{V}(k+r) = Y(k+r) - \hat{Y}(k+r)$ can be written as

$$\hat{V}(k+r) = V(k+r) + Y_{\mathrm{bias}}(k+r) + Y_{\mathrm{noise}}(k+r) \tag{63}$$

Hence, a natural estimate of the noise variance is given by

$$\hat{\sigma}_V^2(k) = \frac{1}{2n+1} \sum_{r=-n}^{n} |\hat{V}(k+r)|^2 \tag{64}$$

In the next section of this appendix we analyze the expected value of (64).

B. Expected Value Noise Variance Estimate

Since $\mathbb{E}\{Y_{\rm bias}(k+r)\overline{V(k+r)}\} = 0$ and $\mathbb{E}\{Y_{\rm bias}(k+r)\overline{Y_{\rm noise}(k+r)}\} = 0$, the expected value of (64) has four non-zero contributions

$$\mathbb{E}\{\hat{\sigma}_V^2(k)\} = \frac{1}{2n+1}\Big[\sum_{r=-n}^{n}\sigma_V^2(k+r) + \sum_{r=-n}^{n}|Y_{\rm bias}(k+r)|^2 + \sum_{r=-n}^{n}\mathbb{E}\big\{|Y_{\rm noise}(k+r)|^2\big\} + 2{\rm Re}\Big(\sum_{r=-n}^{n}\mathbb{E}\big\{Y_{\rm noise}(k+r)\overline{V(k+r)}\big\}\Big)\Big] \quad (65)$$

where ${\rm Re}(x)$ denotes the real part of x. Using the approximation (55), the first sum of (65) becomes

$$\frac{1}{2n+1}\sum_{r=-n}^{n}\sigma_V^2(k+r) = \sigma_V^2(k) + (\sigma_V^2(k))'O\big(\frac{n}{N}\big) \quad (66)$$

In the second sum of (65), $Y_{\rm bias}(k+r)$ depends on the bias of the estimated local model parameters θ_k. Since the latter is described by b_G, the right hand side of (23), we get

$$\frac{1}{2n+1}\sum_{r=-n}^{n}|Y_{\rm bias}(k+r)|^2 = O(b_G^2) \quad (67)$$

Using $Y_{\rm noise}(k+r) = O((2n+1)^{-1/2})$, the third sum of (65) equals

$$\frac{1}{2n+1}\sum_{r=-n}^{n}\mathbb{E}\big\{|Y_{\rm noise}(k+r)|^2\big\} = O\big(\frac{1}{2n+1}\big) \quad (68)$$

In Section B-C of this appendix it is shown that $\mathbb{E}\{Y_{\rm noise}(k+r)\overline{V(k+r)}\} = O(n^{-1})$. Therefore, the contribution of the fourth sum of (65) is

$$2{\rm Re}\Big(\sum_{r=-n}^{n}\frac{\mathbb{E}\big\{Y_{\rm noise}(k+r)\overline{V(k+r)}\big\}}{2n+1}\Big) = O\big(\frac{1}{2n+1}\big) \quad (69)$$

Collecting (65)–(69) proves the right hand side of (26).

Note that the estimated local model parameters θ_k introduce, via $Y_{\rm noise}(k+r)$, a correlation over the frequency r of the output residual $\hat{V}(k+r)$ (63). For the LP method the corresponding decrease in degrees of freedom of $\hat{V}(k+r)$ is exactly n_{θ_k} [proof: see [9], Appendix 7.B, pp. 269–270]. Although this is an approximation for the other methods, it motivates why $2n+1$ in (64) is replaced by $2n+1-n_{\theta_k}$ giving (24a). The effect of this modification is a reduction of the $O(n^{-1})$ term in (26).

C. Correlation Local Estimates – Noise

We prove (69) for the LL estimator (19). The proof for the other local methods follows the same lines.

The linear least squares problem (19) can be written as an overdetermined set of equations

$$Y_k = H(V_k)\theta_k + V_k + R_k \quad (70a)$$

with Y_k, V_k and R_k $(2n+1)\times 1$ vectors of, respectively, the output $Y(k)$, the disturbing noise $V(k)$ and the rational approximation errors $R(k)$ in the local frequency band

$$X_k = [X(k-n),\ldots,X(k),\ldots,X(k+n)]^T \quad (70b)$$

with $X = Y$, V and R, and where $H(V_k)$ is a $(2n+1)\times n_{\theta_k}$ regression matrix depending on the disturbing noise V_k. The linear least squares estimate

$$\hat{\theta}_k = (H^H(V_k)H(V_k))^{-1}H^H(V_k)Y_k \quad (71)$$

can be split into three contributions

$$\hat{\theta}_k = \theta_{k,0} + \theta_{k,\rm bias} + \theta_{k,\rm noise} \quad (72)$$

where $\theta_{k,0}$ is the true value, $\theta_{k,\rm bias}$ the bias due to the disturbing noise V_k and the rational approximation errors R_k, and $\theta_{k,\rm noise}$ the zero-mean random contribution. Note that $Y_0(k+r)$, $Y_{\rm bias}(k+r)$ and $Y_{\rm noise}$ in (62a) depend on, respectively, $\theta_{k,0}$, $\theta_{k,\rm bias}$ and $\theta_{k,\rm noise}$ [see (61)].

Via a first order approximation of $\theta_{k,\rm noise}$

$$\theta_{k,\rm noise} \approx (H^H(0)H(0))^{-1}H^H(0)V_k \quad (73)$$

the correlation between $\theta_{k,\rm noise}$ and $V(k+r)$ can be calculated

$$\mathbb{E}\{\theta_{k,\rm noise}\overline{V(k+r)}\} = (H^H(0)H(0))^{-1}H^H(0)\mathbb{E}\{V_k\overline{V(k+r)}\} = O\big(\frac{\sigma_V^2(k)}{2n+1}\big) \quad (74)$$

where the second equality uses $H^H(0)H(0) = O(2n+1)$ and the fact that $\mathbb{E}\{V_k\overline{V(k+r)}\}$ has only one non-zero entry, namely $\sigma_V^2(k+r)$. Combining (62c) with (74), it follows that $\mathbb{E}\{Y_{\rm noise}(k+r)\overline{V(k+r)}\} = \sigma_V^2(k)O((2n+1)^{-1})$, which proves (69).

Appendix C
Proof of Theorem 3

Under Assumption 2, the disturbing noise $V(k)$ in (4) is circular complex normally distributed. Hence, applying the results of Chapter 9 of [9] it follows immediately that the LP, LL, LSK, LGTLS, LBTLS and LOE estimates of the local model parameters θ_k are asymptotically $(n\to\infty)$ complex normally distributed.

In the remainder of this appendix we will show that the (asymptotic) distribution of the LP and LOE estimates is *circular* complex. It means that

$$\mathbb{E}\{(\hat{\theta}_k - \mathbb{E}\{\hat{\theta}_k\})(\hat{\theta}_k - \mathbb{E}\{\hat{\theta}_k\})^T\} = 0 \quad (75)$$

so that the second order moments of $\hat{\theta}_k$ are solely determined by

$${\rm Cov}(\hat{\theta}_k) = \mathbb{E}\{(\hat{\theta}_k - \mathbb{E}\{\hat{\theta}_k\})(\hat{\theta}_k - \mathbb{E}\{\hat{\theta}_k\})^H\} \quad (76)$$

Note that (75) is not zero for the LL, LSK, LGTLS and LBTLS estimates and, hence, their complex normal distributions are fully characterized by both second order moments (75) and (76).

The minimizer of the LP cost function (17) can be written as

$$\hat{\theta}_k = -\big(J_{\rm LP}^H J_{\rm LP}\big)^{-1}J_{\rm LP}^H Y_k \quad (77)$$

with $J_{\rm LP} = \partial e_{\rm LP}/\partial\theta_k$, $e_{\rm LP}\in\mathbb{C}^{(2n+1)\times 1}$ the local polynomial residuals (17b) and Y_k defined in (70b). Since θ_k (77) is via Y_k (4) an affine analytic function of the disturbing noise V_k defined in (70b) [$J_{\rm LP}$ is independent of V_k], it is also circular complex distributed for finite values of n [see [26], Chapter 4]. Given relationship (9), with $a_{0,k} = 1$, between the FRF

and the LP estimate, Property 1 of Theorem 3 immediately follows.

Using (52), and in the absence of local modeling errors, the minimizer of the LOE cost function (18) can asymptotically ($n \to \infty$) be written as

$$\hat{\theta}_k - \theta_{k,0} = -\big(J_{\mathrm{LOE}}^H(\theta_{k,0}) J_{\mathrm{LOE}}(\theta_{k,0})\big)^{-1} \times J_{\mathrm{LOE}}^H(\theta_{k,0}) e_{\mathrm{LOE}}(\theta_{k,0}) \tag{78}$$

with $J_{\mathrm{LOE}}(\theta_k) = \partial e_{\mathrm{LOE}}(\theta_k)/\partial\theta_k$ and $e_{\mathrm{LOE}}(\theta_k) \in \mathbb{C}^{(2n+1)\times 1}$ the vector of the local output error residuals (18b) [proof: see Appendix D]. Since $e_{\mathrm{LOE}}(\theta_{k,0})$ is a linear analytic function of V_k and since $J_{\mathrm{LOE}}(\theta_{k,0})$ is independent of V_k, the right hand side of (78) is circular complex distributed. It proves that $\hat{\theta}_k$ is asymptotically ($n \to \infty$) circular complex normally distributed. A first order Taylor series expansion of (9)

$$\hat{G}(j\omega_k) \approx G(j\omega_k) + \frac{\partial G(j\omega_k)}{\partial \theta_{k,0}} (\hat{\theta}_k - \theta_{k,0}) \tag{79}$$

shows that $\hat{G}(j\omega_k)$ is asymptotically an affine function of $\hat{\theta}_k$, which concludes the proof of Property 2 of Theorem 3.

Note that the LL, LSK, LGTLS, and LBTLS estimates are not circular complex distributed because they are not an analytic function of the disturbing noise V_k. See, for example, the factor $H^H(V_k)H(V_k)$ in the LL estimate (71). Using (79) the asymptotic ($n \to \infty$) complex normal distribution of the FRF estimate (9) follows from that of $\hat{\theta}_k$.

APPENDIX D
PROOF OF EQ. (78)

In the absence of modeling errors, (52) is asymptotically ($n \to \infty$) equal to

$$(\hat{\theta}_k - \theta_{k,0})_{\mathrm{re}} = -V_{\mathrm{LOE}}''^{-1}(\theta_0) V_{\mathrm{LOE}}'^T(\theta_0, Z_k) \tag{80}$$

where

$$V_{\mathrm{LOE}}''(\theta_0) = 2\mathrm{Re}\big(J^H(\theta_{k,0}) J(\theta_{k,0})\big) \tag{81a}$$

$$V_{\mathrm{LOE}}'^T(\theta_0, Z_k) = 2\mathrm{Re}\big(J^H(\theta_{k,0}) e_{\mathrm{LOE}}(\theta_{k,0})\big) \tag{81b}$$

with $J(\theta_k) = \partial e_{\mathrm{LOE}}(\theta_k)/\partial(\theta_k)_{\mathrm{re}}$. Using the chain rule of the partial derivative, taking into account that $e_{\mathrm{LOE}}(\theta_k)$ is an analytic function of θ_k, we find

$$\begin{aligned} J(\theta_k) &= \left[\begin{array}{cc} \frac{\partial e_{\mathrm{LOE}}(\theta_k)}{\partial \mathrm{Re}(\theta_k)} & \frac{\partial e_{\mathrm{LOE}}(\theta_k)}{\partial \mathrm{Im}(\theta_k)} \end{array} \right] \\ &= \left[\begin{array}{cc} J_{\mathrm{LOE}}(\theta_k) & jJ_{\mathrm{LOE}}(\theta_k) \end{array} \right] \end{aligned} \tag{82}$$

with $J_{\mathrm{LOE}}(\theta_k) = \partial e_{\mathrm{LOE}}(\theta_k)/\partial\theta_k$. Combining (82) with (81), gives

$$\begin{aligned} \mathrm{Re}\big(J^H(\theta_{k,0}) J(\theta_{k,0})\big) &= \left[\begin{array}{cc} \mathrm{Re}\big(M(\theta_{k,0})\big) & -\mathrm{Im}\big(M(\theta_{k,0})\big) \\ \mathrm{Im}\big(M(\theta_{k,0})\big) & \mathrm{Re}\big(M(\theta_{k,0})\big) \end{array} \right] \\ &= \big(M(\theta_{k,0})\big)_{\mathrm{Re}} \end{aligned} \tag{83a}$$

$$\begin{aligned} \mathrm{Re}\big(J^H(\theta_{k,0}) e_{\mathrm{LOE}}(\theta_{k,0})\big) &= \left[\begin{array}{c} \mathrm{Re}\big(X(\theta_{k,0})\big) \\ \mathrm{Im}\big(X(\theta_{k,0})\big) \end{array} \right] \\ &= \big(X(\theta_{k,0})\big)_{\mathrm{re}} \end{aligned} \tag{83b}$$

with $M(\theta_k) = J_{\mathrm{LOE}}^H(\theta_k) J_{\mathrm{LOE}}(\theta_k)$ and $X(\theta_k) = J_{\mathrm{LOE}}^H(\theta_k) e_{\mathrm{LOE}}(\theta_k)$, and where the second equality in (83a) and (83b) define the matrix operators $()_{\mathrm{Re}}$ and $()_{\mathrm{re}}$ respectively. Collecting (80), (81) and (83) we finally obtain

$$\begin{aligned} (\hat{\theta}_k - \theta_{k,0})_{\mathrm{re}} &= -\Big(\big(M(\theta_{k,0})\big)_{\mathrm{Re}}\Big)^{-1} \big(X(\theta_{k,0})\big)_{\mathrm{re}} \\ &= -\big(M^{-1}(\theta_{k,0})\big)_{\mathrm{Re}} \big(X(\theta_{k,0})\big)_{\mathrm{re}} \\ &= -\big(M^{-1}(\theta_{k,0}) X(\theta_{k,0})\big)_{\mathrm{re}} \end{aligned} \tag{84}$$

where the second and third equality use, respectively, Lemmas 15.3 and 15.4 on pp. 553 and 554 of [9]. From (84) we deduce immediately that $\hat{\theta}_k - \theta_{k,0} = -M^{-1}(\theta_{k,0}))X(\theta_{k,0})$, which proves (78).

REFERENCES

[1] D. J. Ewins, *Modal Testing: Theory and Practice*. New York: Wiley, 1991.

[2] K. Deckers, P. Guillaume, C. Vuye and D. Lefeber, "Implementation of the scanning laser Doppler vibrometer combined with a light-weight pneumatic artificial muscle actuator for the modal analysis of a civil structure," *Shock and Vibration*, vol. 19, no. 3, pp. 421–431, 2012.

[3] A. Guha and A. Patra, "Online estimation of the electrochemical impedance spectrum and remaining useful life of lithium-ion batteries," *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 8, pp. 1836–1849, 2018.

[4] E. Ravagli, M. Crescentini, P. Rovatti and S. Severi, "Noninvasive estimation of plasma sodium concentration during hemodialysis via capacitively coupled electrical impedance spectroscopy," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 4, pp. 1673–1681, 2020.

[5] J. J. Cabrera-López and J. Velasco-Medina, "Structured approach and impedance spectroscopy microsystem for fractional-order electrical characterization of vegetable tissues," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 2, pp. 469–478, 2020.

[6] D. R. Brillinger, *Time Series: Data Analysis and Theory*. New York: McGraw-Hill, 1981.

[7] J. Antoni and J. Schoukens, "Optimal settings for measuring frequency response functions with weighted overlapped segment averaging,", *IEEE Transactions on Instrumentation and Measurement*, vol. 58, no. 9, pp. 3276–3287, 2009.

[8] R. Pintelon, J. Schoukens, G. Vandersteen and K. Barbé, "Estimation of nonparametric noise and FRF models for multivariate systems – Part I: theory," *Mechanical Systems and Signal Processing*, vol. 24, no. 3, pp. 573–595, 2010.

[9] R. Pintelon and J. Schoukens, *System Identification: A Frequency Domain Approach, 2nd ed.* Hoboken, New Jersey: Wiley–IEEE Press, 2012.

[10] T. McKelvey and G. Guérin, "Non-parametric frequency response estimation using a local rational model," In *IFAC Proceedings Volumes*, vol. 45, Brussels (Belgium), July 2012, pp. 49–54.

[11] E. Geerardyn, *Development of User-Friendly System Identification Techniques*. PhD thesis, Vrije Universiteit Brussel, Faculty of Engineering, dept. ELEC, August 2016.

[12] D. Verbeke and J. Schoukens, "Frequency response measurements with local parametric modeling," *IEEE Transactions on Instrumentation and Measurement*,vol. 69, no. 6, pp. 3249–3261, 2020.

[13] D. Peumans, C. Busschots, G. Vandersteen, and R. Pintelon, "Improved FRF measurement of lightly damped systems using local rational models," *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 7, pp. 1749–1759, 2018.

[14] R. Voorhoeve, A. van der Maas and T. Oomen, "Non-parametric identification of multivariable systems: A local rational modeling approach with application to a vibration isolation benchmark," *Mechanical Systems and Signal Processing*, vol. 105, pp. 129–152, 2018.

[15] P. Henrici, *Applied and Computational Complex Analysis. Vol. 1: Power Series–Integration–Conformal Mapping–Location of Zeroes*, New York: John Wiley & Sons, 1974.

[16] W. Van Assche, "Padé and Hermite-Padé approximation and orthogonality," *Surveys in Approximation Theory*, vol. 2, pp. pp. 61–91, 2006.

[17] R. Pintelon and J. Schoukens, "Box-Jenkins identification revisited – Part I: Theory," *Automatica*, vol. 42, no. 1, pp. 63–75, 2006.

[18] E. C. Levy, "Complex curve fitting," *IRE Transactions on Automatic Control*, vol. 4, no. 1, pp. 37–43, 1959.

[19] C. K. Sanathanan and J. Koerner, "Transfer function synthesis as a ratio of two complex polynomials," *IEEE Transactions on Automatic Control*, vol. 8, no. 1, pp. 56–58, 1963.
[20] R. Pintelon, P. Guillaume, G. Vandersteen and Y. Rolain, "Analyses, development and applications of TLS algorithms in frequency-domain system identification," *SIAM journal on Matrix Analysis and Applications*, vol. 19, no. 4, pp. 983–1004, 1998.
[21] A. McQuarrie, R. Shumway and C. L. Tsai, "The model selection criterion AIC$_u$," *Statistics & Probability Letters*, vol. 34, no. 3, pp. 285–292, 1997.
[22] F. De Ridder, R. Pintelon, J. Schoukens and D. P. Gillikin, "Modified AIC and MDL model selection criteria for short data records," *IEEE Transactions on Instrumentation and Measurement*, vol. 54, no. 1, pp. 144–150, 2005.
[23] L. Ljung, *System Identification: Theory for the User, 2nd ed.* Upper Saddle River, NJ: Prentice-Hall, 1999.
[24] J. Schoukens, G. Vandersteen, R. Pintelon, Z. Emedi and Y. Rolain, "Bounding the polynomial approximation errors of frequency response functions," *IEEE Transactions on Instrumentation and Measurement*, vol. 62, no. 5, pp. 1346–1353, 2013.
[25] R. Pintelon, M. Schoukens and J. Lataire, "Best linear approximation of nonlinear continuous-time systems subject to process noise and operating in feedback," *IEEE Transactions on Instrumentation and Measurement*, doi 10.1109/TIM.2020.2987476.
[26] B. Picinbono, *Random Signals and Systems*, Englewood Cliffs: Prentice-Hall, 1993.