Identification of Volterra Kernels Using Interpolation

József G. Németh*, István Kollár*, and Johan Schoukens**

* Department of Measurement and Information Systems
Budapest University of Technology and Economics
Phone: +36 1 463 2057, fax: +36 1 463 4112, nemeth@mit.bme.hu
** Department ELEC – Fundamental Electricity and Instrumentation
Vrije Universiteit Brussel

Abstract – The paper presents a new method for the identification of frequency-domain Volterra kernels. Based on the assumption that frequency-domain kernels are locally smooth, the kernel surface can be approximated by interpolation techniques, thus reducing the complexity of the model. Similarly to the unreduced (Volterra) model, this smaller model is also (i) linear in the unknowns, (ii) only locally sensitive to its parameters and (iii) free of structural assumptions about the system. The parameter estimation boils down to solving a linear system of equations in the least-squares (LS) sense.

The design of the interpolation scheme is described, and the performance of the approximation is analyzed, and illustrated by simulation. The algorithm allows a significant saving in measurement time compared to other kernel estimation methods.

Keywords – nonlinear system, system identification, Volterra series, Volterra kernel, interpolation, random multisine, B-spline.

I. INTRODUCTION

Many nonlinear systems can be described by a Volterra series [1], and can be well approximated around an operating point by the first few terms of the series. In this paper the focus is on the identification of a second order kernel. The frequency domain transform of a second order Volterra series is shown in Eq. 3. It is linear in the unknowns ($G^{(1)}(k_1)$, $G^{(2)}(k_1, k_2)$), however, there are many more unknowns than equations obtained from a single measurement. The objective is to reduce the number of necessary measurements.

II. PRELIMINARIES

A. Normalized random multi-sine excitation signals

In this paper, a *normalized random multisine* excitation will be used. Multisine excitation and frequency domain methods

for the identification of linear systems are treated in [2]-[5]. A random multisine is a broadband, periodic signal:

$$u(t) = \sum_{k=-N}^{N} U(k) e^{j2\pi \frac{k}{N} f_{\max} t} , \qquad (1)$$

with $U(k) = \overline{U}(-k) = |U(k)| e^{j\varphi_k}$, where $f_{\max}$ is the maximum frequency of the of the excitation signal, N is the number of frequency components, and the phases φ_k are independent, uniformly distributed, random variables on $[0, 2\pi)$, such that $E\left\{ e^{j\varphi_k} \right\} = 0$. Different random phases result in different realizations of the random multisine.

Let $\hat{U}(f): [0, f_{\max}] \rightarrow \mathrm{R}^+$. If N is varied, then $u(t)$ can be normalized: $|U(k)| = \frac{1}{\sqrt{N}} \hat{U}\left(\frac{|k|}{N} \cdot f_{\max} \right)$, resulting in a normalized random multisine. Its time-domain amplitude distribution is asymptotically Gaussian $(as\ N \rightarrow \infty)$, its RMS is independent from N, whereas its spectral resolution and period length varies in proportion to N:

$$T_{\text{period}} = \frac{N}{f_{\max}} . \qquad (2)$$

B. Frequency-domain truncated Volterra model

A Volterra series [1] provides a description for dynamic systems in a similar way as the Taylor series does for static input/output relationships, i.e., the system output is split into linear, quadratic, cubic, etc. contributions (see *Fig. 1*) (N. B. a Volterra series cannot model some nonlinear behaviors, such as hysteresis and chaos but these are out of scope, here.)
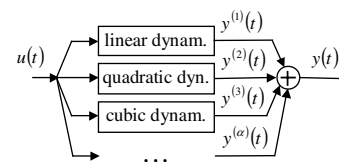

Figure 1. Linear, quadratic, etc., contributions in the Volterra series expansion.
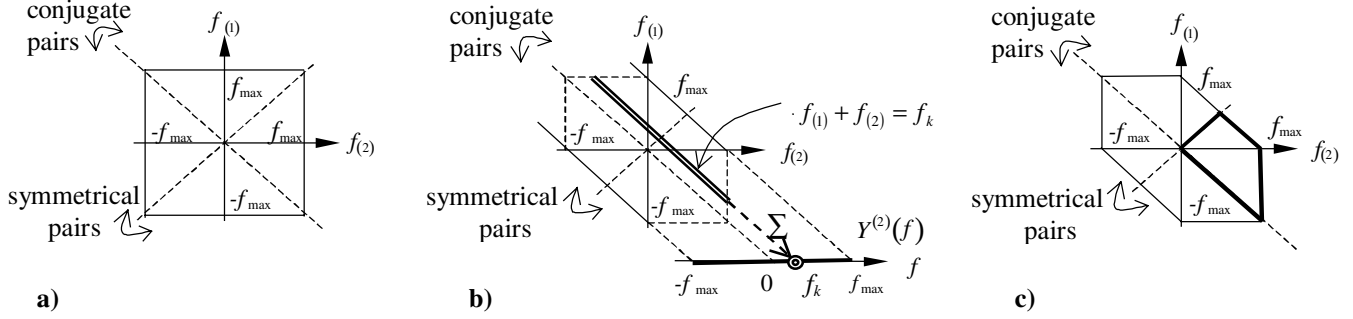
Figure 2. a) Domain on which the quadratic polyspectrum is different from zero. Symmetrical and complex conjugate pairs.

b) Summing the weighted polyspectrum along $f_{(1)} + f_{(2)} = \frac{k}{N} \cdot f_{max} = f_k$ yields $Y^{(2)}(f_k)$. Horizontal, bold segment: band of observed output.

c) Hexagon: domain on which the kernel can be identified. Trapezoid: selected, non-redundant part of the kernel.

In this paper the system is restricted to its 1st- and 2nd-order frequency-domain Volterra kernels. The response of this model to a multisine excitation (Eq. 1) is periodic and can be described by the following Fourier coefficients:

$$Y(k) = Y^{(1)}(k) + Y^{(2)}(k) =$$
$$= G^{(1)}(k) \cdot U(k) + \sum_{k_1=-N+k}^{N} G^{(2)}(k_1, k-k_1) \cdot U^{(2)}(k_1, k-k_1), \quad (3)$$
$$k = 1, ..., N$$

where $U^{(2)}$ is a second order *polyspectrum* (the tensor product of two spectra):

$$U^{(2)}(k_1, k_2) = U(k_1) \cdot U(k_2). \quad (4)$$

In these equations,

- $U(k_i)$ is the complex Fourier coefficient of the k_i^{th} harmonic of the input.

- $Y(k)$ is the complex Fourier coefficient of the kth harmonic of the output; $Y^{(1)}(k)$ and $Y^{(2)}(k)$ are the linear and the quadratic contributions, respectively. Only the excited band is observed, hence: $k <= N$.

- $G^{(1)}(k_1)$ and $G^{(2)}(k_1, k_2)$ are samples of the 1st and 2nd-order Volterra kernels, respectively; the index k_i being a substitute for the explicit frequency $\frac{k_i}{N} \cdot f_{max}$. The first-order kernel $G^{(1)}$ is a frequency response function (FRF).

The frequency band of the excitation determines the band in which the system characteristics can be identified. This is illustrated for the quadratic kernel in *Fig. 2*. The quadratic polyspectrum is zero outside the square shown in *Fig. 2.a*, and contains symmetrical and complex conjugate pairs according to Eq. 4. Equation 3 is a weighted sum of (poly)spectrum points, where $G^{(1)}(k_1)$ and $G^{(2)}(k_1, k_2)$ are

the weights; the quadratic contributions are illustrated in *Fig. 2.b*. The $G^{(2)}$ kernel can be made symmetrical, too, and the domain depicted in *Fig. 2.c* can be selected for estimation.

C. Measurement set-up and operating point

$U(k)$'s and $Y(k)$'s are obtained by executing band-limited experiments with the excitation described in Eq. 1, and averaging FFT's of whole periods in steady state. It is important to note that not only the observations are band-limited, but the input to the system, as well. For the sake of simplicity, at first, noise-free observations are assumed.

The DC component at the input and the output is considered as an operating point, thus all terms containing $U(0)$ must be omitted from Eq. 3 (where $Y(0)$ is excluded already). This involves that for a pure square-law device, for instance, the identified model will contain non-zero linear components ($G^{(1)}$) if the operating point is different from 0 ($U(0) \neq 0$).

III. PROBLEM SETTING & NEW APPROACH

Equation 3 is a linear problem vis-a-vis the unknowns $G^{(1)}(k_1), G^{(2)}(k_1, k_2)$. By executing a measurement with a single realization of the random multisine, N equations are obtained $(k = 1, ..., N)$. However, the quadratic contributions contain $O(N^2)$ complex unknowns. This means that $O(N)$ measurements must be accomplished (each with a different realization of the random multisine) to identify the kernels.

In this paper, $G^{(2)}(k_1, k_2)$'s will be considered linearly dependent on each other: the quadratic kernel is conceived as an interpolated surface, of which the parameters are linear. Thus the number of the linear unknowns can be reduced to $O(N)$, hence $O(N^0)$ realization of the measurement will be sufficient. This approach is introduced in Section III.B.

A. Fundamental problem setting ($O(N^2)$ unknowns)

The linear problem to be solved is different from Eq. 3, since:
1. according to Section II.C, all terms containing $U(0)$ must be omitted, and
2. the redundant unknowns (symmetrical pairs in $G^{(2)}$) must be eliminated in order to get a full-rank problem.

For each k, the resulting equations (which are not written out explicitly) contain a sum of products (similarly to Eq. 3), which can be written as a scalar product of vectors:

$$Y(k) = \mathbf{u}_k^T \cdot \mathbf{g}_k, \quad k = 1, \ldots, N \tag{5}$$

where $\mathbf{u}_k$'s and $\mathbf{g}_k$'s contain the samples of the two kernels and the (poly)spectra, respectively, (the superscript T denoting the transpose, not conjugated). The N equations, which are not coupled, can be assembled into a matrix form:

$$\mathbf{y} = \mathbf{U} \cdot \mathbf{g}, \tag{6}$$

where $\mathbf{y} = [Y(1), \ldots, Y(N)]^T \in C^{N \times 1}$, $\mathbf{g} \in C^{M \times 1}$ and $\mathbf{U} \in C^{N \times M}$, with $M = O(N^2)$ the number of complex unknowns[†]. $\mathbf{U}$ is a sparse matrix, in each column having a single nonzero entry.

By executing the measurement with more than one realizations of the excitation, the obtained $\mathbf{y}$'s and $\mathbf{U}$'s can be concatenated by rows, yielding:

$$\mathbf{y}_{(1 \text{ to } R)} = \mathbf{U}_{(1 \text{ to } R)} \cdot \mathbf{g}, \tag{7}$$

with

$$\mathbf{y}_{(1 \text{ to } R)} = [\mathbf{y}_1; \ldots; \mathbf{y}_i; \ldots; \mathbf{y}_R] \in C^{R \cdot N \times 1}, \tag{8}$$

and

$$\mathbf{U}_{(1 \text{ to } R)} = [\mathbf{U}_1; \ldots; \mathbf{U}_i; \ldots; \mathbf{U}_R] \in C^{R \cdot N \times M}, \tag{9}$$

where R denotes the number of different realizations, and the subscript i refers to the ith realization of the measurements.

By generating $O(N)$ different realizations of the measurements, enough rows could be obtained for solving Eq. 7. However, the measurement would take relatively long, unless N is reduced. A similar approach can be found in [6,7].

B. New approach: a reduced complexity model (only $O(N)$ unknowns)

In order to reduce the number of unknowns, an assumption is made. Namely, that the kernels are locally smooth, thus the

real and imaginary parts of the quadratic kernel can be sought for in the form of interpolated surfaces. For the sake of simplicity, we consider here the case when the same kind of interpolation is used for the real and the imaginary components.

Without considering the exact choice of interpolation, only its linear property is used:

$$\hat{\mathbf{g}}(\mathbf{a}) = \mathbf{P} \cdot \mathbf{a}, \tag{10}$$

where $\mathbf{P} \in R^{M \times L}$ is the interpolation matrix, $\mathbf{a} \in C^{L \times 1}$ is the parameter vector, with $L = O(N)$ the number of complex unknowns. Thus, the number of the unknowns may be significantly reduced. Eq. 10 implies

$$\text{Re}(\hat{\mathbf{g}}) = \mathbf{P} \cdot \text{Re}(\mathbf{a}), \text{ and } \text{Im}(\hat{\mathbf{g}}) = \mathbf{P} \cdot \text{Im}(\mathbf{a}), \tag{11}$$

i.e., the real and imaginary components of the kernels are approximated using the same basis vectors (columns of $\mathbf{P}$).

The output of the approximate model is:

$$\hat{\mathbf{y}}(\mathbf{a}) = \mathbf{U} \cdot \hat{\mathbf{g}}(\mathbf{a}). \tag{12}$$

The output error of the approximate model is:

$$\mathbf{h} = \hat{\mathbf{y}} - \mathbf{y}. \tag{13}$$

By adopting the notations of Eqs. 7-9 for the case when several realizations are observed,

$$\mathbf{h}_{(1 \text{ to } R)} = \hat{\mathbf{y}}_{(1 \text{ to } R)} - \mathbf{y}_{(1 \text{ to } R)}, \tag{14}$$

and

$$\left\| \mathbf{h}_{(1 \text{ to } R)} \right\|_2^2 = \sum_{i=1}^{R} \left\| \mathbf{h}_i \right\|_2^2. \tag{15}$$

The performance function will be the mean-square output error:

$$MSE(\mathbf{a}; \mathbf{P}) = E_\varphi \left\{ \left\| \mathbf{h} \right\|_2^2 \right\} \Big|_{R=1}, \tag{16}$$

where the statistical expectation is taken with respect to all realizations of φ (see Sect II.A). $MSE(\mathbf{a})$ is a quadratic function of $\mathbf{a}$, and is parameterized by $\mathbf{P}$. The optimal parameter setting $\mathbf{a} = \boldsymbol{\alpha}$ is the one yielding the least-mean-square output error (using a given $\mathbf{P}$):

$$\boldsymbol{\alpha} = \arg \min_{\mathbf{a}} MSE(\mathbf{a}). \tag{17}$$

[†] $M = \frac{3}{4}N^2 - \frac{1}{2}N - \frac{1}{4}$ if N is odd, $M = \frac{3}{4}N^2 + \frac{1}{2}N$ if N is even.

To approximate $\boldsymbol{\alpha}$, the following equation may be solved in LS sense:

$$\mathbf{y}_{(1\,\text{to}\,R)} = \left(\mathbf{U}_{(1\,\text{to}\,R)} \cdot \mathbf{P}\right) \cdot \mathbf{a} \quad . \tag{18}$$

The resulting $\mathbf{a} = \hat{\boldsymbol{\alpha}}$ minimizes $\frac{1}{R}\left\|\mathbf{h}_{(1\,\text{to}\,R)}\right\|_2^2$ for the realizations involved in the specific experiment. Since $\frac{1}{R}\left\|\mathbf{h}_{(1\,\text{to}\,R)}\right\|_2^2$ converges to *MSE* as $R \to \infty$, it can be proved that $\hat{\boldsymbol{\alpha}}$ is an asymptotically unbiased and weakly consistent estimate of $\boldsymbol{\alpha}$:

$$\forall \varepsilon > 0: \ \lim_{R\to\infty} \text{Prob}\left(\left|\hat{\boldsymbol{\alpha}}(R) - \boldsymbol{\alpha}\right| > \varepsilon\right) = 0 . \tag{19}$$

IV. ALGORITHM

A. Steps of the identification process

1) Select the operation point and the desired excitation class (RMS, $f_{\max}$, etc. ; see Sect. II.A).

2) Design $\mathbf{P}$. (See subsection B.)

3) Select N and R (i.e., the number of harmonics and the number of realizations), so that the equation to unknown ratio of Eq. 18 be $R \cdot N / L > 1$. The necessary ratio also depends on the noise level and the higher order nonlinear contributions. Then generate a multisine to start the measurements. The duration of the measurements is:

$$T_{\text{meas}} = p \cdot \frac{R \cdot N}{f_{\max}}, \tag{20}$$

where p is the number of periods measured with each realization. It is necessary to wait several periods to achieve steady state. Additional periods can be measured to filter the additive noise by averaging the FFTs. The higher order nonlinear contributions are periodic, hence, these are not filtered with the noise. On the other hand, these can be averaged by increasing R and N.

It is advised to carry out additional realizations of the measurements (e.g. $R_{\text{test}} = R/4$) to collect observations as test data, for the assessment of the later estimate.

4) When the data are available, solve Eq. 18 in LS sense to obtain $\hat{\boldsymbol{\alpha}}$, but do not use the test data for this estimation.

5) Estimate $MSE(\hat{\boldsymbol{\alpha}};\mathbf{P})$ by using the test data:
$\hat{MSE}(\hat{\boldsymbol{\alpha}};\mathbf{P},\mathbf{U}_{\text{test}}) = \frac{1}{R_{\text{test}}}\left\|\mathbf{h}(\hat{\boldsymbol{\alpha}};\mathbf{P},\mathbf{U}_{\text{test}})\right\|_2^2$. If $\hat{MSE}(\hat{\boldsymbol{\alpha}})$ is

satisfactory, then the procedure may be finished, or the result might be improved by including the 'test data' in the estimation of $\boldsymbol{\alpha}$. If $\hat{MSE}(\hat{\boldsymbol{\alpha}})$ is *not* satisfactory, then further iterations are necessary through steps 3-5 (keeping $\mathbf{P}$) or 2-5. (The recursive least-squares solution is out of the scope, here. See [7].)

B. Design of the interpolation scheme: $\mathbf{P}$

Interpolation by B-splines [8] is used because it is (*i*) locally smooth, (*ii*) only locally sensitive to the parameters, and (*iii*) it can be applied to arbitrary dimensions by tensor product extension. The two-dimensional extension is as follows:

$$\hat{f}(x,y) = \sum_{i,j=1}^{n,m} a_{i,j} B_{i,j}^{(2)}(x,y) = \sum_{l=1}^{n\cdot m} a_l B_l^{(2)}(x,y), \tag{21}$$

where $B_{i,j}^{(2)}(x,y) = B_i(x) \cdot B_j(y). \tag{22}$

$\{B_i(x)\}_1^n$ is a "horizontal" B-spline basis, whereas $\{B_j(y)\}_1^m$ is a "vertical" B-spline basis. If values of $\hat{f}(x,y)$ are arranged in a column vector $\hat{\mathbf{g}}$, and $\{B_l^{(2)}(x,y)\}_1^{m\cdot n}$ are arranged likewise into the columns of a matrix $\mathbf{P}$, then Eq. 10 is obtained formally. (The difference is that $\hat{\mathbf{g}}$ in Eq. 10 includes the linear kernel, too, which is not reduced by interpolation.)

Two alternatives are obvious in selecting the directions for the horizontal and the vertical axes:
1. parallel to the frequency axes of the quadratic kernel, or
2. parallel to the complex conjugation axis and the symmetry axis of the quadratic kernel, respectively.

Only the latter option is analyzed here. (See *Fig. 4*). To avoid unnecessary boundary effects, first, a basis must be designed over the redundant kernel area, then the redundant parameters must be eliminated by taking into account the symmetries of the quadratic kernel (see *Figs. 3-4*). Finally, the bases functions must be 'clipped' to the non-redundant kernel area.
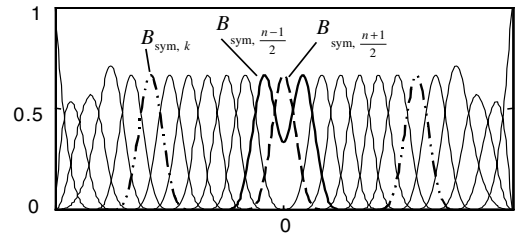

Figure 3. Symmetrical basis obtained from cubic B-splines (with uniform knot-placement and not-a-knot boundary condition):
$$B_{\text{sym},k} = B_k + B_{n-k}, \ k = 1,\ldots,\frac{n-1}{2}, \ \text{ and } \ B_{\text{sym},(n+1)/2} = B_{(n+1)/2}$$

Since the real part of the kernel is symmetrical to $f_{(1)} + f_{(2)} = 0$, whereas the imaginary part is anti-symmetrical to it, different basis vectors should be used for the two surfaces. If this choice is taken, then Eq. 11 is replaced by $\text{Re}(\hat{\mathbf{g}}) = \mathbf{P}_{\text{sym}} \cdot \mathbf{a}_{\text{Re}}$ and $\text{Im}(\hat{\mathbf{g}}) = \mathbf{P}_{\text{anti-sym}} \cdot \mathbf{a}_{\text{Im}}$, and the following equation must be solved in LS sense, instead of Eq. 18:

$$\begin{bmatrix} \text{Re}(\mathbf{y}) \\ \text{Im}(\mathbf{y}) \end{bmatrix}_{(1 \text{ to } R)} = \begin{bmatrix} \text{Re}(\mathbf{U}) & -\text{Im}(\mathbf{U}) \\ \text{Im}(\mathbf{U}) & \text{Re}(\mathbf{U}) \end{bmatrix}_{(1 \text{ to } R)} \cdot \begin{bmatrix} \mathbf{P}_{\text{sym}} & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_{\text{anti-sym}} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{a}_{\text{Re}} \\ \mathbf{a}_{\text{Im}} \end{bmatrix} \quad (23)$$

where all values are real. $\mathbf{a}_{\text{Re}}$ and $\mathbf{a}_{\text{Im}}$ may have different dimensions. The total number of real-valued parameters shall be denoted by L_{Re} (sum of lengths of $\mathbf{a}_{\text{Re}}$ and $\mathbf{a}_{\text{Im}}$).

At interpolation boundaries, the constraints are fewer. When the B-splines are constructed, this is taken into account, which solves the probem for the boundaries of the basis (see the rectangle in *Fig. 4*). However, these boundaries do not match the boundaries of the excited domain of the kernel at $f_{(1)} = f_{\text{max}}$ and $f_{(2)} = f_{\text{max}}$. Consequently, here the degree of liberty of the interpolation is too high. This will set back the performance of the estimation.

C. *Computational complexity*

The computational complexity is dominated by the solution of the least-squares problem. In the simulations, Gaussian elimination was used implying $\approx 2 \cdot R \cdot N \cdot L_{\text{Re}}^2$ flops.

V. ERROR SOURCES

Suppose that the physical system being identified can be described by a Volterra series, then some of the following effects may dominate the $MSE(\hat{\boldsymbol{\alpha}}; \mathbf{P})$, and call for remedy:

1. *Capability of the interpolation scheme to approximate the surface.* The model set is reduced by Eq. 10. Since in general $\mathbf{g} \notin \text{range}(\mathbf{P})$, an approximation error is necessarily present, which depends on $\mathbf{P}$ and the system.

2. *Convergence of the LS solution.* $MSE(\mathbf{a}; \mathbf{P})$ has one single minimum: $MSE(\boldsymbol{\alpha}; \mathbf{P})$; hence, $\hat{\boldsymbol{\alpha}} \neq \boldsymbol{\alpha}$ introduces an additional approximation error.

3. *Observation noise.* Zero-mean, uncorrelated, additive, output noise introduces no additional bias into $\boldsymbol{\alpha}$, but increases its variance. Therefore, if certain output frequencies are more noisy than others, then a weighted least-squares solution can be applied instead of Eq. 23, with weights inversely proportional to the estimated
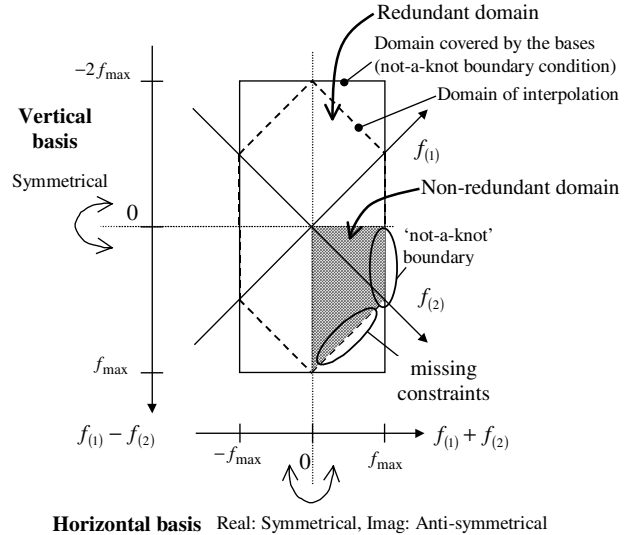

Figure 4. The bases (for the real and imaginary parts of the quadratic kernel) are defined by the tensor product of the horizontal and the vertical bases.

noise variance. Noise in the input observations always introduces a bias.

4. *Convergence of the Volterra series that describes the true system.* Higher order nonlinear contributions have similar effects to additive, possibly correlated, output noise. In [4,5] it is shown that even order nonlinear contributions do not bias the linear FRF estimate, when a linear model is being identified, whereas odd order contributions do. It also follows, that the output error that is due to the approximation of the quadratic kernel does not introduce a bias in the linear kernel estimate.

VI. SIMULATIONS

A. *The simulated system*

In this section, simulations are provided using the system:

$$\dot{y}(t) + a \cdot y(t) = b \cdot u(t) + c \cdot u^2(t), \quad (24)$$

where $a = b = c = 2\pi$. The operating point is $U(0) = 0$. The RMS of the excitation is 0.25 V, whereas its bandwidth is $f_{\text{max}} = 8.1$ Hz. Throughout the simulations $N = 81$ harmonics are applied and the magnitude spectrum is flat. The observations are noiseless, and the system does not produce higher order nonlinear contributions, thus points one and two of the previous section are illustrated in the followings.

B. *Illustration of the consistency*

By increasing the number of realizations (R), the *MSE* must decrease. In *Fig. 5* Y_{av} is the root-mean-square of the output

of the system, $Y_{av}^{(2)}$ is the quadratic component of Y_{av}, whereas the lower curves show the root-mean-square error at each frequency. The equation to unknown ratio in Eq. 23 is $2 \cdot R \cdot N / L_{Re}$, which yields 1.1, 1.3, 1.5 and 6.6, respectively, in the presented cases. The convergence is rapid. By using 30 realizations, the *MSE* is bound to hit a floor depending on the interpolation scheme. (See Section V.)

C. Illustration of the performance for several interpolation schemes

The performance of four interpolation schemes will be compared, and the notation adopted in Table 1 will be used. For the interpolation schemes that have fewer parameters, fewer observations are used at the parameter estimation.

Table 1. Interpolation schemes with the number of realizations used.

# of quadratic kernel param. $L_{Re} - 2 \cdot N$	cubic	5th order	# of realizations R	Equation to unknown ratio $2 \cdot R \cdot N / L_{Re}$
573	**P₁**	**P₂**	7	1.54
1179	**P₃**	**P₄**	12	1.45

In *Fig. 6* the lower curves are the mean-square error components for each frequency. Curves P1, P2, P3 and P4 must be compared to $Y_{av}^{(2)}$ since they mostly result from the error of the quadratic kernel. Near $f_{(1)} + f_{(2)} = 0$ the approximation performs worse due to the steep variation of the kernel surface in this region (*Fig. 7*). Near $f_{max} = 8.1\,\text{Hz}$ the error grows rapidly due to boundary effects (*Figs. 4,6*). The curves are in agreement with the expectations. Namely, that 5th order splines can better follow the variations of the kernel surface, but they are also more vulnerable to boundary effects, since the basis functions have larger support. These effects are smaller when the number of the parameters is high.

The quadratic kernel shown in *Fig. 7* has been identified using **P₁**. Compared to the solution of the non-parametric problem setting as described in Section III.A, the saving in measurement time is a ratio of $N / R = 13$.

REFERENCES

[1] Schetzen, M., "The Volterra and Wiener Theories of Nonlinear Systems", Wiley-Interscience, 1980.
[2] Schoukens, J. and R. Pintelon, "Identification of linear Systems: A Practical Guide to Accurate Modeling", Pergamon, 1991.
[3] Godfrey, K. R., "Perturbation Signals for System Identification", Prentice-Hall, 1993.
[4] Pintelon, R. and J. Schoukens, "System Identification: a Frequency Domain Approach", IEEE Press, to appear in 2001.
[5] Schoukens, J., T. Dobrowiecki and R. Pintelon, "Identification of linear systems in the presence of nonlinear distortions. A frequency domain approach.", IEEE Trans. on Automatic Control, vol.43, no 2, 1998.
[6] Kim K. I. and E. J. Powers, "A digital method of modeling quadratically nonlinear systems with a general random input", IEEE Trans. Acoust. Speech Signal Process., vol. 36, pp. 1758-1769, Nov. 1988.
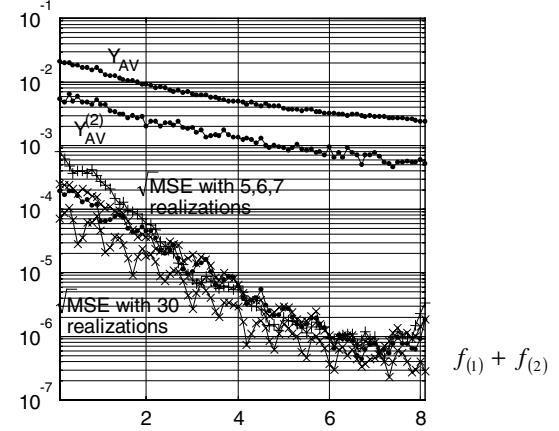
Figure 5. Convergence of the root-MSE components as *R* increases. Four cases: 5,6,7 and 30 realizations are used for the parameter estimation. In each case, the curves are estimated by using 30 independent realizations.
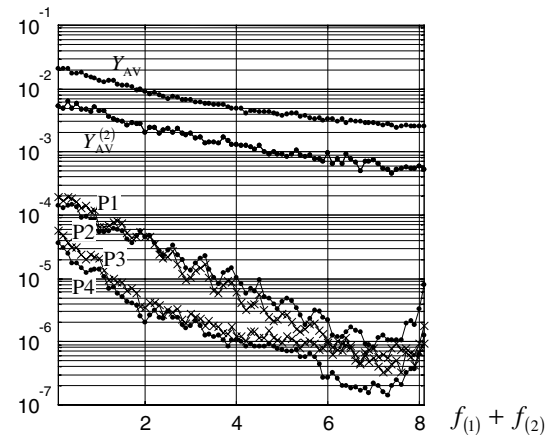

Figure 6. Root-MSE components estimated using 30 independent realizations. Cases P1,P3: cubic splines – P2,P4: 5th order splines P1,P2: 573 real parameters, 7 realizations P3,P4: 1179 params., 12 realiz.
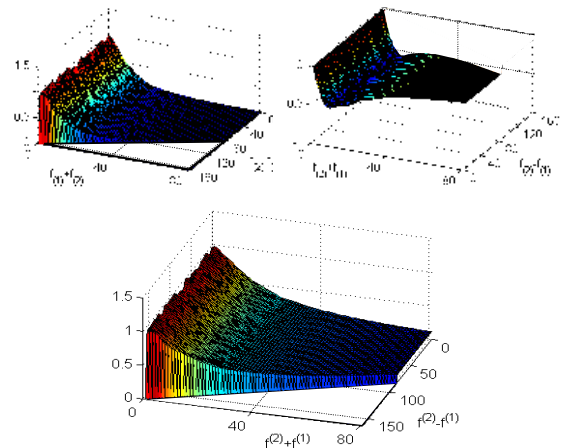

Figure 7. Identified quadratic kernel (*Re & Im* part & Magnitude) using **P₁**.

[7] Nam S. W. and E. J. Powers, "Application of higher order spectral analysis to cubically nonlinear system identification", IEEE Trans. Signal Proc., vol. 42, no. 7, pp. 1746-1765, July 1994.
[8] de Boor, C., "A Practical Guide to Splines", Springer, 1978.