MEASURING LINEAR APPROXIMATION TO WEAKLY NONLINEAR MIMO SYSTEMS

Tadeusz Dobrowiecki*, Johan Schoukens**

Dept. of Measurement and Information Systems, Budapest University of Technology and Economics, Budapest H-1117, Magyar tudósok krt-ja 2.,
***Dienst ELEC, Vrije Universiteit Brussel, Brussel, B-1050, Pleinlaan 2*

Abstract: The choice of the input signals has impact on the nonparametric frequency response function (FRF) measurements of a nonlinear MIMO system. It is shown that Gaussian noise, periodic noise, and random multisines are equivalent, yielding in the limit the same linear approximation to a nonlinear MIMO system. Even in the noiseless case, variability of the FRF is observed due to the presence of the nonlinearity and the randomness of the excitation. Based upon these effects a new class of equivalent input signals is proposed, decreasing essentially the variance of the nonlinear FRF measurements, while the same linear approximation is retrieved.

Keywords: MIMO systems, nonparametric, frequency-response, excitation, variance.

1. INTRODUCTION·

A SISO Volterra system, driven with a random multisine, i.e. a:

$$u(t) = F^{-1/2} \sum_{k=-M/2, k\neq 0}^{M/2} \hat{U}(kf_s / M) e^{j(2\pi k f_s t / M + \varphi_k)} \quad (1)$$

where $\varphi_{-k} = -\varphi_k$, the phases φ_k are iid random variables uniformly distributed on [0, 2π[, the real amplitudes obey $\hat{U}(f) \geq 0$, $f \leq f_{\max} = Ff_s / M$ and $\hat{U}(f = 0) = 0$, can be modeled as a linear approximation followed by a nonlinear noise source. The nonlinearity is translated into systematic 'bias' and noise like 'stochastic' contributions (Schoukens, *et al.*, 1998). This model has been extended to periodic noise (amplitudes $\hat{U}(f)$ in (1) are random) and Gaussian noise. FRFs estimated with any of those signals tend in the limit ($M \to \infty$) to the same linear approximation of the SISO system, provided that the signal spectra are comparable (Pintelon and Schoukens, 2002). Although the bias on the FRF is the same, it is not so with the nonlinear noise variance. Measuring with Gaussian noise introduces leakage. Periodic noise is leakage free, but its spectrum fluctuates from one realization to the other. Comparably, random multisines introduce least variance, consequently they are proposed as signals of choice in weakly nonlinear SISO measurements (Pintelon and Schoukens, 2001).

This theory will be extended to MIMO Volterra systems with no feedback, i.e. where outputs can be investigated separately as MISO systems. In the time domain such system can be described as:

$$y(t) = \sum_{\alpha=1}^{\infty} y^{\alpha}(t) = \sum_{\alpha=1}^{\infty} \sum_{j_1 j_2 \cdots j_\alpha} y^{j_1 j_2 \cdots j_\alpha}(t) \quad (2)$$

where the 2nd sum runs over all pure and mixed αth order kernels. A particular αth order kernel excited by input signals of indices $j_1, j_2, \ldots, j_\alpha$ is:

$$y^{j_1 j_2 \cdots j_\alpha}(t) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g^{j_1 j_2 \cdots j_\alpha}(\tau_1, \ldots, \tau_\alpha) \times u_{j_1}(t-\tau_1) u_{j_2}(t-\tau_2) \ldots u_{j_\alpha}(t-\tau_\alpha) d\tau_1 \ldots d\tau_\alpha \quad (3)$$

In the frequency domain this model becomes:

$$Y(l) = \sum_{\alpha=1}^{\infty} Y^{\alpha}(l) = \sum_{\alpha=1}^{\infty} \sum_{j_1 j_2 \cdots j_\alpha} Y^{j_1 j_2 \cdots j_\alpha}(l) \quad (4)$$

$$Y^{j_1 j_2 \cdots j_\alpha}(l) = N^{-\alpha/2} \times \sum_{k_1, \ldots, k_{\alpha-1} = -M/2}^{M/2} G^{j_1 j_2 \cdots j_\alpha}(k_1, k_2, \ldots, k_{\alpha-1}, k_\alpha) \times U_{j_1}(k_1) U_{j_2}(k_2) \ldots U_{j_\alpha}(k_\alpha) \quad (5)$$

· This work was supported by the Hungarian-Flemish Bilateral R&D Agreement of the Research & Technology Innovation Fund (Hungary) and the Ministry of the Flemish Community, by the Flemish Government (GOA-ILiNos), the FWO (Onderzoeks-gemeenschap ICCoS, research grants G.0211.05 and G.0528.04) and the Belgian government as a part of the Belgian program on Interuniversity Poles of Attraction (IUAP V/22).

where $U_k(l) = (M/F)^{1/2}\hat{U}_k(lf_s/M)e^{j\varphi_k}$, l is the discrete frequency, $l = \Sigma k_i$, i=1…α, (Schoukens, *et al.*, 1998). In Section 2 we examine the bias and the variance of the FRF approximation. In Section 3 MIMO equivalence of the signals mentioned above is proven. Section 4 introduces a new class of excitations with superior properties. Finally Section 5 provides some illustrative simulations and conclusions.

2. SOME REMARKS ON BIAS AND VARIANCE

Further on we omit the index of the output. Index k of the measured signal path is called *reference input* index. In practice all signal paths in a MISO system are computed simultaneously via a set of equations. For this purpose J experiments are made, cutting (after the transients settle) the successive records from the input and output signals. Signal amplitudes at frequency l are then arranged into:

$$\mathbf{Y}(l) = \mathbf{U}(l)\mathbf{G}(l) = \begin{bmatrix} Y^{(1)}(l) \\ \dots \\ Y^{(J)}(l) \end{bmatrix} = \begin{bmatrix} U_1^{(1)}(l) & \dots & U_N^{(1)}(l) \\ \dots & U_j^{(i)}(l) & \dots \\ U_1^{(J)}(l) & \dots & U_N^{(J)}(l) \end{bmatrix} \begin{bmatrix} G^1(l) \\ \dots \\ G^N(l) \end{bmatrix}, \quad (6)$$

(parentheses indicate the serial numbers of the experiments). The required FRF estimates can be computed as:

$$\hat{\mathbf{G}}(l) = \left[\hat{G}^i(l)\right] = (\mathbf{U}^H(l)\,\mathbf{U}(l))^{-1}\mathbf{U}^H(l)\,\mathbf{Y}(l) = \hat{\mathbf{S}}_{UU}^{-1}(l)\,\hat{\mathbf{S}}_{YU}(l), \quad (7)$$

where $(\,)^H$ means conjugate transpose.

With the equivalence proven in the next Section it is actually also shown that in the limit ($M \to \infty$) the output of the Volterra MIMO system, excited by the above mentioned input signals, can be written as:

$$\begin{aligned}\mathbf{Y}(l) &= \mathbf{G}_\mathbf{R}(l)\mathbf{U}(l) + \mathbf{Y}_\mathbf{S}(l) = \left[G_{R,l}^k(l)\right]\left[U_k(l)\right] + \left[Y_{S,l}(l)\right] \\ &= \left[G_l^k(l)\right]\left[U_k(l)\right] + \left[G_{B,l}^k(l)\right]\left[U_k(l)\right] + \left[Y_{S,l}(l)\right]\end{aligned} \quad (8)$$

where the so called related dynamic systems $G_{R,l}^k = E\{Y_l/U_k\} = G_l^k + G_{B,l}^k$ are linear approximations to the multidimensional Volterra series in the signal path $Y_l - U_k$, and $G_{B,l}^k$ are biases introduced by the nonlinearity. The equivalent noise sources $Y_{S,l}(l)$ capture all the nonsystematic nonlinear effects. (6) yields the MIMO additive nonlinear noise source model, a straightforward extension of the SISO and two-input two-output (TITO) cases (Dobrowiecki and Schoukens, 2004a).

Even in the absence of measurement noise, FRF measurements will vary from one realization to the other for two reasons. One is the fluctuation of the inverse matrix in (7) due to the experiment design. Note that $(\mathbf{U}^H\mathbf{U})^{-1}$ fluctuates even when random multisines with constant amplitude spectrum are used, contrary to the case of SISO systems (Pintelon and Schoukens, 2001). Other problem is the nonlinear noise source $Y_{S,l}(l)$ in (8), a zero mean stochastic contributions generated by the nonlinearity. Generally the variance on the measured FRF will depend on the used input signals (with the expected ranking of: $\mathrm{Var}_{\text{Gaussian noise}} > \mathrm{Var}_{\text{periodic noise}} > \mathrm{Var}_{\text{random multisine}}$), the order and the particular composition of the kernels.

The aim is then to design experiments suitable for the nonparametric FRF identification of MIMO Volterra systems. They should yield low measurement variance (by reducing the random fluctuations of the inverse in (7) and the influence of $Y_{S,l}(l)$), while the computed FRF estimates $\mathbf{G_R}$ in (8) should be the same for all kinds of excitations.

3. BIAS EQUIVALENT INPUT SIGNALS

We turn now to the equivalence of the measured MIMO FRF for signals evaluated in (Pintelon and Schoukens, 2002), i.e. random multisines (1) with $\hat{U}_k^{\,2}(f) = S_{\hat{U}\hat{U}}(f)$, periodic noise with random spectral amplitudes $E\{\hat{U}_k^2(f)\} = S_{\hat{U}\hat{U}}(f)$, and Gaussian noise with power spectrum $S_{UU}(f) = S_{\hat{U}\hat{U}}(f)/f_{\max}$. For more technical details on these signals, see (Pintelon and Schoukens, 2002). The results can be collected in the following corollary:

Assumption 1: Signals at different inputs are independent, their phases and (in case of periodic noise) spectral amplitudes are independent over the frequency. The signals have comparable spectral powers and are defined on the same frequency grid.

Assumption 2: The MIMO system can be of arbitrary input dimension N and order of the nonlinearity (assuming that the sums in (2-5) converge).

Corollary 1: Under Assumption 1 and 2, and for $M\to\infty$, all of the mentioned signals yield exactly the same linear approximation to a nonlinear MIMO system modeled by series (3-5). The contributing kernels with nonzero expected values (with respect to the random parameters of the excitations) are only those odd order kernels, which contain the reference input an odd number of times, and any other input an even number of times, including 0. (For the sketch of the proof see *App. A*).

Note that the SISO case is a special case with only reference input present).

4. ORTHOGONAL RANDOM MULTISINES

In the nonlinear measurements, only Gaussian signals in general, or multisines are assumed (note that all of the mentioned signals are Gaussian in the limit), see e.g. (Zi-Qiang and Billing, 2000). For linear MIMO systems (Briggs and Godfrey, 1966), later (Guillaume, *et al.*, 1996) introduced orthogonal inputs based on the Hadamard matrix (9) to minimize the influence of the measurement noise by the maximization the determinant of the input matrix. In nonlinear context it was criticized by (Dedene, *et al.*, 2002), as affecting the operating point of the system.

$$\mathbf{U}(l) = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} U(l) \qquad (9)$$

Considering that a linear approximation followed by a nonlinear output noise model had been extended to MIMO systems (8), Hadamard matrix inputs were investigated for cubic Volterra TITO system and shown to be bias equivalent to the random multisines with much less of the measurement variance (Dobrowiecki and Schoukens, 2004a). Further direct generalization to higher dimensions and orders failed (probably due to over constraining, as foreseen in (Dedene, *et al.*, 2002)). A drawback of the Hadamard design is also the restriction of the input dimension to $N = 4k$. For other dimensions truncated Hadamard matrix are used as an ad hoc solution.

Our aim is to design inputs suitable for an arbitrary MIMO Volterra system, bias equivalent to other, traditionally used excitations, and yielding even less measurement variance. Maximum determinant problem has solution in complex numbers for arbitrary dimensions (DFT matrix, Brenner and Cummings, 1972). However substituting one orthogonal matrix for another is not enough to obtain bias equivalence (i.e. keeping intact the operating point of the system).

Let us assume that $J = N_B \cdot N$ experiments are made and let partition the $N \times J$ input matrix $\mathbf{U}$ into N_B rectangular blocks as: $\mathbf{U}^H = \begin{bmatrix} \mathbf{U}_N^H & \mathbf{U}_N^H & ... & \mathbf{U}_N^H \end{bmatrix}$ and $\mathbf{Y}$ in a similar way. Instead of choosing:

$$\mathbf{U}_N(l) = \begin{bmatrix} U_1^{(1)}(l) & U_2^{(1)}(l) & ... & U_N^{(1)}(l) \\ U_1^{(2)}(l) & U_2^{(2)}(l) & ... & U_N^{(2)}(l) \\ ... & ... & ... & ... \\ U_1^{(N)}(l) & U_2^{(N)}(l) & ... & U_N^{(N)}(l) \end{bmatrix}, \qquad (10)$$

which requires independent excitations for every input and every experiment, we propose (11) where w_{ij} are entries of an arbitrary, deterministic, orthogonal (unitary) matrix: $\mathbf{W}^H\mathbf{W} = \mathbf{W}\mathbf{W}^H = N\,\mathbf{I}$.

$$\mathbf{U}_N(l) = \begin{bmatrix} w_{11}U_1^{(1)}(l) & w_{12}U_2^{(1)}(l) & ... & w_{1N}U_N^{(1)}(l) \\ w_{21}U_1^{(1)}(l) & w_{22}U_2^{(1)}(l) & ... & w_{2N}U_N^{(1)}(l) \\ ... & ... & ... & ... \\ w_{N1}U_1^{(1)}(l) & w_{N2}U_2^{(1)}(l) & ... & w_{NN}U_N^{(1)}(l) \end{bmatrix} \qquad (11)$$

For $\mathbf{W}$ a DFT matrix can be chosen, i.e. the Vandermode matrix of the roots of unity $[\mathbf{W}]_{kn} = \omega^{(k-1)(n-1)}, \omega = \exp^{2\pi j/N}$. Consider e.g. a 3-dim system. The basic input matrix could be:

$$\mathbf{U}_3(l) = \begin{bmatrix} U_1(l) & U_2(l) & U_3(l) \\ U_1(l) & e^{j\frac{2\pi}{3}}U_2(l) & e^{-j\frac{2\pi}{3}}U_3(l) \\ U_1(l) & e^{-j\frac{2\pi}{3}}U_2(l) & e^{j\frac{2\pi}{3}}U_3(l) \end{bmatrix}$$

For $N = 4k$ a simpler choice is the Hadamard matrix, composed solely from ± 1 entries.

The procedure is thus to generate random excitations for the first experiment in the block of N experiments and to shift them orthogonally for the next N-1 experiments, then start with another random choice for the next block. Due to the orthogonal properties of $\mathbf{W}$ (we omit the experiment index for clarity):

$$\begin{aligned} \mathbf{U}_N^H(l)\mathbf{U}_N(l) &= \left[\overline{U}_k(l)U_j(l)\sum_{i=1}^{N}\overline{w}_{ik}w_{ij} \right] \\ &= N\left[\overline{U}_k(l)U_j(l)\delta_{kj} \right] = N\,diag\{|U_k|^2\} \end{aligned}, \qquad (12)$$

is a simple amplitude scaling. Furthermore from (7): $(\mathbf{U}^H\mathbf{U})^{-1} = (N_B\mathbf{U}_N^H\mathbf{U}_N)^{-1}$ and finally:

$$\hat{\mathbf{G}}(l) = \frac{1}{N_B}\sum_{i=1}^{N_B}\hat{\mathbf{G}}_{N,i}(l), \text{ with}$$

$$\begin{aligned} \hat{\mathbf{G}}_{N,i}(l) &= (\mathbf{U}_N^H(l)\mathbf{U}_N(l))^{-1}\mathbf{U}_N^H(l)\,\mathbf{Y}_N(l) \\ &= \frac{1}{N}diag\{\frac{1}{|U_k|^2}\}\mathbf{U}_N^H(l)\,\mathbf{Y}_N \end{aligned} \qquad (13)$$

computed without taking inverse from one block of N equations. Because (12) gets rid of random fluctuations, a reasonable drop in variance should be expected. Moreover orthogonal random multisines lead to exactly the same $\mathbf{G_R}$ as for other input signals.

Corollary 2: Under Assumptions 1 and 2 the orthogonal random multisines (11) are equivalent to the signals specified in the *Corollary* 1. When normalized to the same spectral behavior and in the limit $M \to \infty$, they yield exactly the same $\mathbf{G_R}$. (For the sketch of the proof see *App. B*).

Kernels with nonzero expected values contribute to the bias and those with zero expected values to the variance. A thorough analysis of the nonlinear kernels reveals that due to the orthogonal $\mathbf{W}$ a number of zero expected value kernels will become zero, depressing the level of the variance further.

Corollary 3: The orthogonal random multisines (11) generate least variance comparing to the signals mentioned in the Corollary 1. A number of zero mean kernels present usually in the nonlinear variance drops out due to the properties of the entries of the orthogonal matrix $\mathbf{W}$. (For the sketch of the proof see *App. C*).

5. SIMULATIONS

Nonlinear variance levels were measured on $\mathbf{G_R}$ of a 3-dim MISO system, excited with Gaussian noise, periodic noise, random multisines, and orthogonal random multisines. For $\mathbf{W}$ in (11) a DFT matrix is used. All signals were scaled to unit power and the number of experiments was with N_B = 1, 2, 5, 10 number of blocks respectively ($J = 3N_B$ now). The

MISO system has a Wiener-Hammerstein structure with linear systems shown in Fig.1. The variance of the FRF of such systems follows their output dynamics (Dobrowiecki and Schoukens, 2002), consequently the comparison was based on the approximate level of the variance in the lower pass band of the system, and on the relative variance of the measurements with the signal *S*, with respect to that obtained with the orthogonal random multisines (*OMS*):

$$Vrel_S^k(l) = \frac{Var_S\{\hat{G}^k(l)\}}{(Var_S\{\hat{G}^k(l)\} + Var_{OMS}\{\hat{G}^k(l)\})/2} \qquad (14)$$

$$VR = \frac{\sum_{l,k} Vrel_S^k(l)}{\sum_{l,k} Vrel_{OMS}^k(l)}$$

where the relative variances are averaged over all input paths and the frequencies.

The static nonlinearity contains first all mixed powers up to the 5th order ($\varepsilon << 1$):

$$NL = x_1 + x_2 + x_2 + \varepsilon \sum_{0 \le i+j+k \le 5} x_1^i x_2^j x_3^k, \qquad (15)$$

It was designed to show the general situation of a weakly nonlinear system. Variances measured for this system are shown in Figs. 2., 3a., and 4a. Excitations randomizing the inverse matrix in (7) show the expected rapid drop in variance for small *J*. Variances measured for the orthogonal multisines follow the 1/*J* behavior. For a higher number of data all signals tend to the same limit. Variances produced with the orthogonal random multisines show a small shift, which does not disappear with increased number of data. This is due to the drop-out effect of some kernels. This effect is seen amplified in Figs. 3b. and 4b., which presents variances measured for:

$$NL = x_1 + x_2 + x_3 + x_1 x_2 + \varepsilon \underbrace{\sum_{i,\, j=0..5} x_1^i x_2^j x_3^k,}_{\text{other terms up to 5th order}} \qquad (16)$$

The strong nonlinear kernel $x_1 x_2$ drops out from the G^1 measurements (App. 3, case a.), appears fully in G^2 measurements (App 3, case b.), and only partly in the G^3 measurements (App 3, case c.). Please note that the drop-out effect is not influenced by the number of processed data, only by the structure of the kernels of the measured MIMO system.

6. CONCLUSIONS

Various practically important excitations yield the same linear approximation for the Volterra SISO systems. This equivalence is extended to the MIMO Volterra series. All excitations introduce however random fluctuations and an increased level of variance. With the new class of excitations the FRF measurements of the nonlinear MIMO Volterra system can be significantly improved in a sense that the equivalent linear approximations are measured with significantly less variance. Measurement procedure is easy, reducing further the computational burden.

REFERENCES

Brenner, J., & Cummings, L. (1972) The Hadamard Maximum Determinant Problem. *Amer. Math. Monthly*, **79**, 626-630.

Briggs, P.A.N. & Godfrey, K.R. (1966) Pseudorandom signals for the dynamic analysis of multivariable systems. *Proc. of the IEE* **113**(7), 1259-1267.

Dedene, N., Pintelon. R., & Lataire, Ph. (2002) Measurement of Multivariable Frequency Response Functions in the Presence of Nonlinear Distortions – Some Practical Aspects. *IEEE Trans. Instr. & Meas*., **51**(4), 577-582

Dobrowiecki, T., & Schoukens, J. (2002). Cascading Wiener-Hammerstein Systems. *Proc. of the 19th IEEE Instrumentation and Measurement Tech. Conference*, Anchorage, USA (pp. 881-886).

Dobrowiecki, T., & Schoukens, J. (2004a) Linear Approximation of Weakly Nonlinear MIMO Systems. *Proc. of the IMTC' 2004*, Como, Italy (pp. 1607-1612).

Dobrowiecki, T., & Schoukens, J. (2004b) Internal Report, www.mit.bme.hu/~tade/tmp/OrthMS.pdf

Guillaume, P., Pintelon, R., & Schoukens, J. (1996). Accurate estimation of multivariable frequency response functions. *Proc. of the 13th IFAC Triennial World Conference*, San Francisco, USA, (pp. 423-428).

Pintelon, R., & Schoukens, J. (2001). *System identification. A frequency domain approach*. Piscataway: IEEE Press.

Pintelon, R., & Schoukens, J. (2002). Measurement and Modeling of Linear Systems in the Presence of Non-linear Distortions. *Mechanical Systems and Signal Processing*, **16**(5), 785-801.

Schetzen, M. (1980). *The Volterra and Wiener Theories of Nonlinear Systems*. John Wiley & Sons, New York.

Schoukens, J., Dobrowiecki, T., Pintelon, R. (1998). Parametric and Non-Parametric Identification of Linear Systems in the Presence of Nonlinear Distortions. A Frequency Domain Approach. *IEEE Trans. Autom. Contr.*, **43**(2), 176-190.

Zi-Qiang, L. & Billing, S.A. (2000). Evaluation of output frequency responses of nonlinear systems under multiple inputs. *IEEE Trans. on Circuits and Systems II*, **47**(1), 28-38.

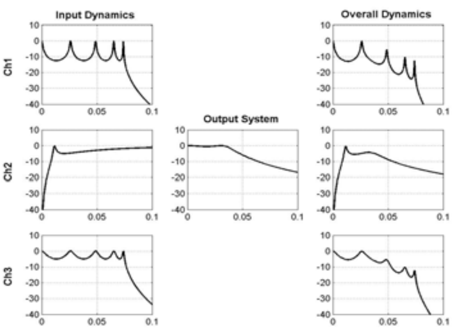

Fig. 1. Linear dynamics of 3-dim Wiener-Hammerstein system used in the simulations.

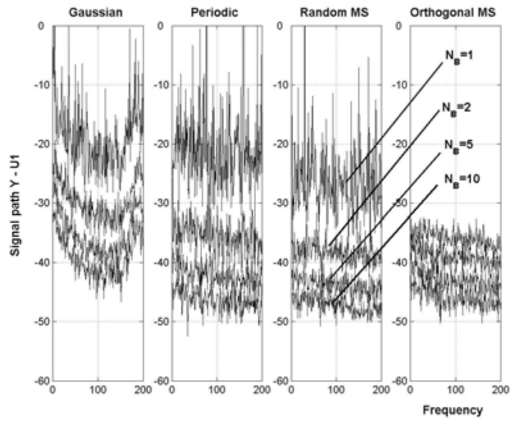

Fig. 2. Variances measured from 10 averages of the FRF in signal path Y - U_1, for N_B=1, 2, 5, 10.

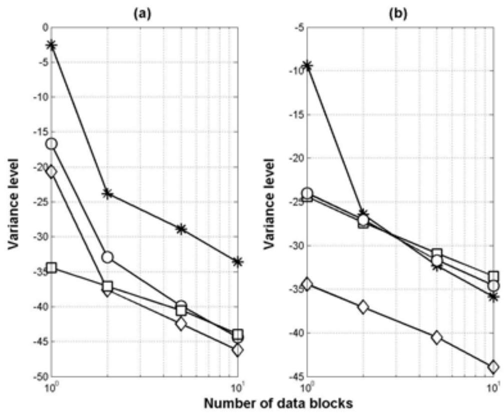

Fig. 3. Approximate variance at the lower 10% of the pass-band. (a) Path Y-U_1 and system (15): Gaussian (∗) and periodic noise (O), random (◊) and orthogonal multisine (□). (b) System (16): random multisine, all paths (∗), orthogonal multisine, path Y-U_2 (O), path Y-U_3 (□), path Y-U_1 (◊).

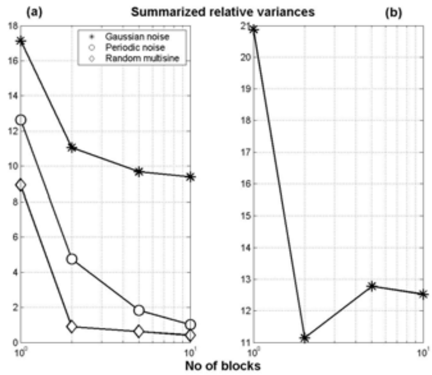

Fig. 4. (a) Relative variance (13) for system (15): Gaussian (∗) and periodic noise (O), random multisine (◊). Note the loss for small number of data and the convergence to 0 for higher number of data. (b) Relative variance for the signal path Y-U_1, system (16). Note no convergence to 0 for higher number of data. Compare with Fig. 3a. and 3b.

APPENDICES

App. A: Bias equivalence of the input signals.

It is enough to investigate a single kernel in (2)-(5). Bias on the measured FRF is the sum of all systematic contributions with nonzero expected values. The nonlinear noise variance comes from all other zero expected value stochastic contributions. The detailed analysis is presented in (Dobrowiecki and Schoukens, 2004b), only general ideas are explained here. The aim is to show that in every case kernels of exactly the same order and combination of inputs contribute to the bias. Scale factors based on the symmetry of the Volterra kernels and the frequency dependence of the kernels based on equivalent signal spectra lead in the limit $M \to \infty$ to exactly the same bias expressions. Signals in the kernel will be grouped as: $j_1, \dots, j_2, \dots, j_K$, where input j_1 appears in the kernel M_1 times, input j_2 M_2 times, etc., and $j_1 < j_2 < \dots < j_K$. Reference input index k will be identified usually with input j_1. The FRF measured with Gaussian noise is:

$$G_R^k(j\omega) = G^k(j\omega) + G_B^k(j\omega) = \frac{S_{YU_k}(j\omega)}{S_{U_kU_k}(j\omega)} \tag{17}$$

To compute: $$G_B^{j_1 \dots j_\alpha}(j\omega) = \frac{S_{Y^{j_1 \dots j_\alpha}U_k}(j\omega)}{S_{U_kU_k}(j\omega)} \tag{18}$$

we need the correlation:

$$\begin{aligned} &R_{y^{j_1 \dots j_\alpha}u_k}(\tau_0) = E\left\{y^{j_1 \dots j_\alpha}(t)u_k(t-\tau_0)\right\} = \\ &\int_{-\infty}^{\infty}\!\!\dots\!\int_{-\infty}^{\infty} \begin{matrix} g^{j_1 j_2 j \dots j_\alpha}(\tau_1, \dots, \tau_\alpha) \times \\ E\{u_k(t-\tau_0)u_{j_1}(t-\tau_1)\dots u_{j_\alpha}(t-\tau_\alpha)\} \end{matrix} d\tau_1 \dots d\tau_\alpha \end{aligned} \tag{19}$$

If the reference input k is not in $j_1, \dots, j_K$, the expected value is zero, so let $k = j_1$. Due to the independent inputs the expected value can be written as (simplifying the notation):

$$\begin{aligned} &E\{u_{j_1}u_{j_1} \dots u_{j_\alpha}\} \\ &= E\{\underbrace{u_{j_1}u_{j_1} \dots u_{j_1}}_{M_1+1}\}E\{\underbrace{u_{j_2}u_{j_2} \dots u_{j_2}}_{M_2}\} \dots E\{\underbrace{u_{j_K}u_{j_K} \dots u_{j_K}}_{M_K}\} \end{aligned} \tag{20}$$

Each of the expected values in (20) is zero for odd number of terms. For even number of terms they can be decomposed into sums of combinations of pair wise correlations $\Sigma\Pi R_{u_n u_m}(\tau_i - \tau_j)$ (Schetzen, 1980; Pintelon and Schoukens, 2001). Consequently the order M_1 of the reference signal in the kernel must be odd, and the orders M_n of other input signals even (e.g. in signal path with the reference index 1 kernels $G^{111}, G^{122}, G^{12233}, \dots$ etc. will contribute to the bias, but kernels $G^{12}, G^{233}, G^{1234}, \dots$ etc. not). For the final form of (18) we must take into account that by the symmetry of the kernels every combination of pair wise correlations leads the same bias term. The number of possible combinations for the left side of (20) is the product of the numbers of combinations at the right side, and we must introduce the frequency power spectrum via the Fourier transform, like in the SISO case. The final result is:

$$\begin{aligned} G_B^{j_1 \dots j_K}(j\omega) = &\frac{C_{\mathrm{ker}nel}}{(f_{\max})^{\beta-1}} \int_{f_2=0}^{f_{\max}} \!\!\dots\! \int G^{j_1 \dots j_K}(f, -f_2, f_2, \dots) \\ &\times S_{\hat{U}_{j_1}\hat{U}_{j_1}}(j\omega_2) \dots S_{\hat{U}_{j_K}\hat{U}_{j_K}}(j\omega_\beta) df_2 \dots df_\beta \end{aligned} \tag{21}$$

$$C_{\mathrm{ker}nel} = 2^{\beta-1} M_1!! \prod_{l=1}^{K} (M_l - 1)!! \tag{22}$$

where 2β-1 = $\alpha = \Sigma M_i$+1. Multidimensional Volterra frequency kernel in (21) is defined as the Fourier transform of the time domain kernel from (2-3). For more details see (Pintelon and Schoukens, 2001).

Periodic noise and random multisines will be analyzed with the FRF measured first as:

$$G_R^k(l) = E\{\hat{G}^k(l)\} = G^k(l) + G_B^k(l) = E\{\frac{Y(l)}{U_k(l)}\} \quad (23)$$

The expectation of a particular kernel is:

$$\begin{aligned} G_B^{j_1\cdots j_K}(l) &= E\{\frac{Y^{j_1\cdots j_K}(l)}{U_k(l)}\} = \\ N^{-\alpha/2}\sum_{k_1,\dots,k_{\alpha-1}=-M/2}^{M/2} & G^{j_1 j_2\cdots j_\alpha}(k_1,\dots,k_\alpha)\times E\{U_{j_1}(k_1)\dots U_{j_\alpha}(k_\alpha)\frac{1}{U_k(l)}\} \end{aligned} \quad (24)$$

For random multisines the expectation in (24) applies only to the random phases. To yield nonzero expected value the reference input k must be present among inputs $j_1,\dots, j_K$, then it must be possible to pair the remaining inputs (phases of different inputs are independent). The condition on the kernel is consequently the same and with it the bias (23) is:

$$\begin{aligned} G_B^{j_1\cdots j_K}(l) &= \frac{C_{\ker nel}}{(N)^{\beta-1}}\sum_{k_1,\dots,k_{\beta-1}=0}^{M} G^{j_1\cdots j_K}(l,-k_1,k_1,\dots)\times \\ &(\frac{N}{F})^{\beta-1} S_{\hat{U}_{j_1}\hat{U}_{j_1}}(k_1)\dots S_{\hat{U}_{j_K}\hat{U}_{j_K}}(k_{\beta-1}) \end{aligned} \quad (25)$$

For $M \to \infty$ the sum converges to the integral (21). For periodic noise the expectation in (24) applies to both amplitudes and phases:

$$\begin{aligned} & E\{U_{j_1}(k_1)\dots U_{j_\alpha}(k_\alpha)\frac{1}{U_k(l)}\} = \\ & E_{\hat{U}}\{\hat{U}_{j_1}(k_1)\dots\hat{U}_{j_\alpha}(k_\alpha)\frac{1}{\hat{U}_k(l)}\}E_\varphi\{e^{j\varphi_{j_1}(k_1)}\dots e^{-j\varphi_k(l)}\} \end{aligned} \quad (26)$$

Note here also that the reference input must be present (odd number of times) in the kernel and other inputs must appear in even numbers, otherwise the expectations are zero. In such case the denominator and phase expectation cancels and:

$$\begin{aligned} & E_{\hat{U}}\{\hat{U}_{j_1}(k_1)\dots\hat{U}_{j_\alpha}(k_\alpha)\} = \\ & \quad E_{\hat{U}}\{\hat{U}_{j_1}^2(k_1)\}\dots E_{\hat{U}}\{\hat{U}_{j_K}^2(k_{\beta-1})\} + O(M^{-1}) \end{aligned} \quad (27)$$

The asymptotically vanishing term contains higher even order moments. With this the bias equals (25).

When (7) is used as the measurement procedure:

$$\begin{aligned} & G_{B,k}^{j_1\cdots j_K}(l) = E\{\sum_{n=1}^{N} b_{kn}(l)\sum_{i=1}^{J}\overline{U}_n^{(i)}(l)Y^{j_1\cdots j_K(i)}(l)\} = \\ & \sum_{k_1,\dots,k_{\alpha-1}} G^{j_1\cdots j_K}\sum_{n=1}^{N}\sum_{i=1}^{J} E\{U_{j_1}^{(i)}(k_1)\dots U_{j_K}^{(i)}(k_\alpha)\,\overline{U}_n^{(i)}(l)b_{kn}(l)\} \end{aligned} \quad (28)$$

where: $[b_{kn}(l)] = \left[\sum_{i=1}^{J}\overline{U}_k^{(i)}(l)U_n^{(i)}(l)\right]^{-1}$. With the phases independent over frequency and with higher than 2nd order moments (pairing more than two frequencies together) leading to $O(M^{-1})$ order contributions:

$$\sum_{n=1}^{N}\sum_{i=1}^{J} E\{U_{j_1}^{(i)}(k_1)\dots U_{j_K}^{(i)}(k_\alpha)\,\overline{U}_n^{(i)}(l)b_{kn}(l)\} = \quad (29)$$

$$|U(k_1)|^2\dots|U(k_{\beta-1})|^2\, E\{\sum_{n=1}^{N}\sum_{i=1}^{J}\overline{U}_n^{(i)}(l)\dots U_j^{(i)}(l)b_{kn}(l)\}$$

where the term within the expectation is δ_{kn} as the entry of a matrix multiplied by its inverse. Putting (29) into (28) we again obtain (25). For periodic noise the expectation in (28) must be analyzed acc. to (26), and here the entries of the inverse pose problems:

$$\begin{aligned} & \sum_{n=1}^{N}\sum_{i=1}^{J} E_{\hat{U},\varphi}\{U_{j_1}^{(i)}(k_1)\dots U_{j_K}^{(i)}(k_\alpha)\,\overline{U}_n^{(i)}(l)b_{kn}(l)\} = \\ & \sum_{n=1}^{N}\sum_{i=1}^{J} E_{\hat{U}}\{\hat{U}_{j_1}^{(i)}(k_1)\dots\hat{U}_{j_K}^{(i)}(k_\alpha)\,\hat{U}_n^{(i)}(l)\times \\ & \qquad E_\varphi\{e^{j\varphi_{j_1}^{(i)}(k_1)}\dots e^{-j\varphi_k^{(i)}(l)}\,b_{kn}(l)\}\} \end{aligned}$$

Analysis of possible pairings toward nonzero expected value leads to the interim expression of:

$$E_{\hat{U}}\{\hat{U}_{j_1}^2(k_1)\dots\hat{U}_{j_K}^2(k_{\beta-1})E_\varphi\{\sum_{n=1}^{N}\sum_{i=1}^{J}\overline{U}_n^{(i)}(l)\dots U_j^{(i)}(l)\,b_{kn}(l)\}\}$$

which with the comments made to (27) and (29) yields exactly the same expression of the bias as (25).

App. B: Bias of the orthogonal random multisines.
From (12-13) it is enough to show the equivalence of the FRF measured for a single block of data (J=N):

$$\begin{aligned} & G_{B,k}^{j_1\cdots j_K}(l) = E\{\sum_{n=1}^{N} b_{kn}(l)\,Y^{j_1\cdots j_K(n)}(l)\} = \\ & \sum_{k_1,\dots,k_{\alpha-1}} G^{j_1\cdots j_K}E\{\sum_{n=1}^{N}U_{j_1}^{(n)}(k_1)\dots U_{j_K}^{(n)}(k_\alpha)\,b_{kn}(l)\} \end{aligned} \quad (30)$$

which with: $U_j^{(n)}(l) = w_{nj}U_j(l)$ and $b_{kn}(l) = \frac{\overline{w}_{nk}\overline{U}_k(l)}{N|U_k(l)|^2}$

can be written as:

$$\begin{aligned} & G_{B,k}^{j_1\cdots j_K}(l) = \\ & \sum_{k_1,\dots,k_{\alpha-1}} G^{j_1\cdots j_K}(k_1,\dots,k_\alpha)\times A\times E\{U_{j_1}(k_1)\dots U_{j_K}(k_\alpha)\,\overline{U}_k(l)\} \end{aligned}$$

For the expected value to be nonzero exactly the same conditions on the inputs are required as before. A represents dependency on the orthogonal matrix $\mathbf{W}$:

$$A = \frac{1}{N}\sum_{n=1}^{N} v_{nj_1}v_{nj_2}\dots v_{nj_\alpha}\overline{w}_{nk},\ v_{nj_m} = \begin{cases} w_{nj_m},\ k_m > 0 \\ \overline{w}_{nj_m},\ k_m < 0 \end{cases} \quad (31)$$

For frequency pairings leading to the nonzero expected value, due to $|w_{nk}|^2 = 1$, $A = 1$ and the bias expression equals (25).

App. C: Vanishing zero mean contributions.
For a particular kernel, A in (31) depends on the orthogonal matrix $\mathbf{W}$, on the indices of the input and the reference signals and on the frequency pairing introducing complex conjugate for the negative frequencies. Three cases can be distinguished depending whether the term under the sum in (31) simplifies to $w_{np} = 1$, or to a $w_{np} \neq 1$:

a. $A = 1$ for all frequencies. Such kernel contributes fully to the nonlinear variance on the FRF.

b. $A = 0$ for all frequencies. Such kernel drops out from the variance.

c. $A = 0$ only for particular frequencies, when suitable pairing is possible. Such kernel contributes to the variance at those frequencies only.