# Multisine Measurement Technology of Linearly Approximated Weakly Non-linear Systems

Dissertation

**Tadeusz P. Dobrowiecki**
candidate of the technical sciences

Budapest, 2016

# Content

# Multisine Measurement Technology of Linearly Approximated Weakly Non-linear Systems

## 1. Introduction

One should always keep in mind that mathematical models of physical systems are necessarily better or worse approximations. A model is good, if the approximation errors do not jeopardize the model usage and the simplifications (meaning usually the choice of a more convenient model structure) are reasonable. A model is bad, if the rough approximation invalidates theories or artifacts built with the help of the model [155].

Building a good model means also solving a complex engineering problem. Whatever is the task, the model must be ready in time and must be simple enough to provide in time results pertinent to the task goals. Even the best approximating model can easily turn inconsequential, if the costs of the model building and model computations, in terms of equipment and time, cross reasonable limits.

The border between the linear and the non-linear behaviour is never very sharp, nor is it easy to handle. Linear systems exist as a pure abstraction, yet the linear system theory turned out to be one of the most fruitful practical engineering tools. If we conveniently forget that it is always an approximation we use, this theory provides us with a well developed methodology of linear analysis and synthesis [127].

Contrary to the linear theory based upon a single concept of a (linear) model, the non-linear system theory suffers from a multitude of possible non-linear models with widely varying functional properties [90-91, 155]. As a consequence non-linear models can be experimentally difficult to identify, to implement, and for the most part to evaluate. The so called semi-physical modelling helps a bit, since at least there is a fair chance for a good approximation, even if the model itself could be difficult to handle [128].

### 1.1 The problem of the non-linear distortions

In the linear system identification the cost function of the model fit is based upon the notion of the output and the measurement noises as the exclusive interfering agents[1]. Assuming that the phenomenon is linear and the model adequately fits it, the only difference between the both is the noise and the value of the cost function should reflect it.

In most linear measurement (identification) problems, though, the cost function based on the noise alone yields too large values, sending the message that the modeling errors are larger, than to be expected on the basis of the noise analysis alone, and that something (a non-linear

1 Beside a priori information introduced in the optional regularization terms.

distortion) lurks inside the system, which cannot be approximated well within the linear model.

Applying linear system theory without judging the importance and the potential consequences of the unexplainable discrepancies between the phenomena and their models is not an advisable engineering know-how. Furthermore the linear theory won't tell us how far, or how close are we to the validity limits of the model, or how robust it is[2]. The possible non-linear and other modelling problems can be taken into account only as the noise, where it may be impossible to identify and quantify them.

The foremost problem is that the linear system theory warrants that the obtained linear model will be valid for any kind of future experimental conditions (i.e. input signals), yet if the phenomenon is truly non-linear, the linear model is in principle valid solely for the input signal used in its identification. Driving the phenomenon and the model with different inputs can result in discrepancy much larger than experienced or foreseen by the identification process. Yet another problem is that in particular cases the non-linearly distorted system will produce an output "noise" deceiving the user, used to the noisiness of the measurements, as to the true nature of the system.

**Example 1.1.1: Hidden non-linearity and the level of the input signal/1.**

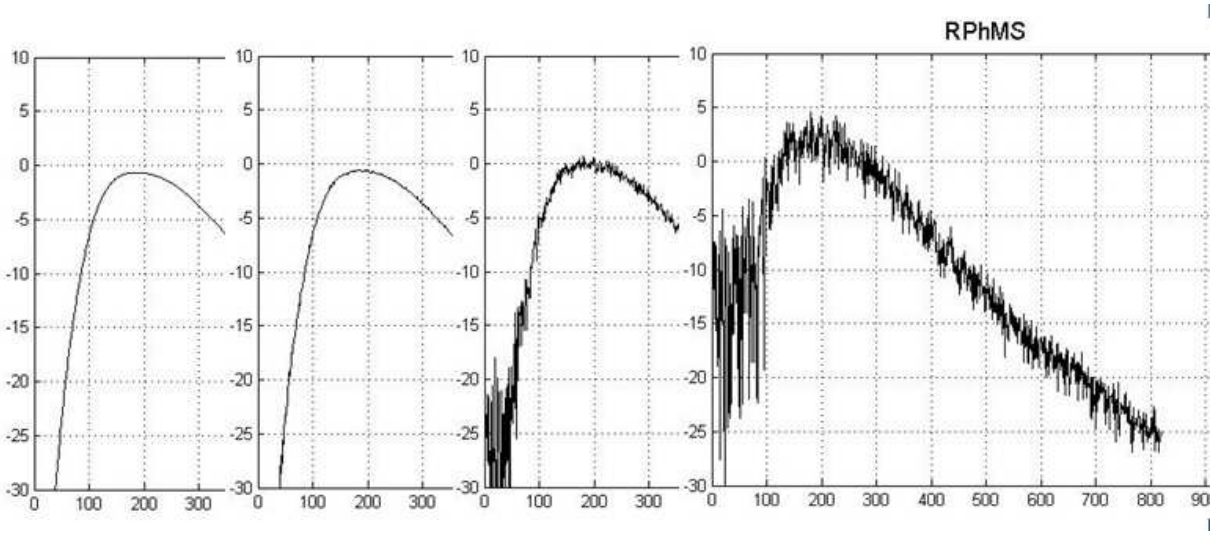


**Fig. 1.1.1** The system under study is composed from a 3rd order Butterworth high-pass filter (input linear system), a static nonlinearity: $y(t)=u(t)+.05u^2(t)+.1u^3(t)+.025u^4(t)+.01u^5(t)$ and a 3rd order Butterworth low-pass filter (output linear system), connected acc. to the Wiener-Hammerstein block-structure (see Section 2.4). The Empirical Transfer Function Estimate (ETFE, [125]) of this system is measured from a single application of the harmonic excitation (odd random phase multisine, see Section 3.3) containing 409 frequency

2 With the increasing number of the measurement data the linear (and time invariant) identification techniques always provide an unfalsified linear model with the uncertainty decreasing to 0, regardless of what is the true nature of the measured system. The reason lies in the fact that the linear techniques use only second order signal statistics and that they cannot make distinction between the system and its second order (least squares) linear approximation [127].

components with uniform amplitudes, under noiseless measurement conditions. The power level of the excitation was adjusted (from left to right) as $\sigma$ = 0.01, 0.1, 0.5, and 1. For small input levels the behavior of the system is convincingly linear, but for higher input levels the nonlinearity steps in more forcibly producing a seemingly random "noisy behaviour". An inexperienced user is here potentially at risk of misinterpreting this phenomenon from the linear theory point of view. Please note that an odd multisine excites only odd harmonics, i.e. the last of the 409 harmonics falls on the 818$^{th}$ frequency index. All the subfigures but the last to the right are printed overlapped for better visual comparison.

**General note on the figures in the dissertation**

The majority of the presented results are analytic and the figures serve only the purpose of illustration. Figures are based on discrete system Matlab(R2007b) simulations, using usually Wiener-Hammerstein system structure with various input and output linear systems, and polynomial static nonlinearities. The frequency band [1 … $f_{max}$] to excite the nonlinear system is always chosen to ensure that $d_{max} * f_{max} < f_s/2$, where $d_{max}$ is the maximum order of the nonlinearity and $f_s$ is the sampling frequency. The frequency axis of the figures usually shows the harmonic (frequency) index, and in some cases the relative frequency (as used in the parametrization of the Matlab functions).

**Example 1.1.2: Hidden non-linearity and the level of the input signal/2.**

A seemingly linear system $y(t) = u(t) + \varepsilon u^3(t), \quad \varepsilon = .01$ is measured in the noiseless measurement setup and is modeled as $y_M(t) = \alpha u(t)$. Its Empirical Transfer Function Estimate (ETFE) $\alpha$ is measured with zero mean Gaussian signal $u(t)$ of $\sigma$=1 as $\alpha = E\{y(t)u(t)\} / E\{u^2(t)\} = 1 + 3\varepsilon\sigma^2 = 1.03$. In the truly linear $\varepsilon = 0$ noiseless case the MSE = $E\{(y(t) - y_M(t))^2\}$ would be zero. Now the MSE = $6\varepsilon^2\sigma^6$ = 0.0006, which can be easily mistaken for some residual noise. The fraction of the non-linear power in the output seems also negligible:

$$100\% E\{(\varepsilon u^3(t))^2\} / E\{y^2(t)\} = 100\% \, 15\varepsilon^2 \sigma^6 / (\sigma^2 + 6\varepsilon\sigma^4 + 15\varepsilon^2 \sigma^6) = 0.14\%, \qquad (1.1.1)$$

and we may be satisfied with the linear identification results. In different experimental conditions, such model can be a source of trouble. Assume that the input is amplified 4 times ($\sigma$= 4). The fraction of the non-linear power is now 16.4%, and the MSE is 2.457 clearly indicating that the linear model is not serving its purpose.

Linear system theory is well developed and used efficiently, even if one may suspect or know that in the reality the system violates linearity assumptions. Linear system theory offers numerous advantages, like canonical model structures and the full theoretical equivalence between the identification problem posed in the time or the frequency domain, or as the input-output and the state-space models. In the following thus we keep on the linear measurement techniques we are however aware that we measure nonlinear systems. A typical situation will be where the non-linear part is negligible within the actual experimental conditions, the nonlinear behaviour contributes little to the measurement results and can be considered a kind of (nonlinear) distortion, but it will be in the interest to the user to know how serious such a distortion could be.

The state-of-the-art non-linear system identification, delivering a model describing well also the distortions, would of course yield answer to every question ([198, 188, 19, 104, 89-91, 54, 155]). In many cases it will be yet impractical at least for two reasons. First, the non-linearity is usually responsible for only a small fraction of the phenomenon, so why to pay the full cost of the parametric non-linear system identification? Second, we do not really want to know the

model of the non-linear distortion exactly, but only its influence on the originally identified linear model.

The literature shows that there is a considerable demand to embedd the nonlinear measurements in the linear measurement techniques, as this approach addresses a realistic situation. It also opens up possibilities for easier and faster measurements in the applications where the investigated system or behaviour is inherently nonlinear, but nobody wants or has the experience to tackle it with fully developed nonlinear models.

In the following we seek answers to these problems. We propose solution to the situation where the system under study shows non-linear behaviour and we show how to express its influence as errors to the measured linear Frequency Response Function (FRF), sidestepping thus the full parametric non-linear system identification. In designing the methods we keep in mind that the computing time is cheap, but the experiments are expensive (duration, or if sophisticated measurements are needed). The ideal solution would be to amend the linear FRF measurements in such a way, that the non-linear effects can be measured, or at least qualified in parallel with the main linear experiments, without extensive additional measurements or nonlinearity tests.

## 1.2 Overview of the literature

Modeling nonlinear distortions to linear systems (i.e. modeling weakly nonlinear systems) is a widely research field with continuous influx of the new theoretical and practical contributions. Linear approximation to the nonlinear behavior became very early the focus of interest. In the context of control [6] used Booton's decomposition of a static non-linearity into the best fitted (in the mean square sense) linear part and a "distortion factor" to analyze control systems with random inputs. J.L. Douce investigated the effect of the static non-linear distortions and their spectral behavior under random excitations [51, 53]; he even proposed a random-signal generator based upon the harmonic intermodulation due to non-linearity [52]. Distorting effect of the non-linearity on the input spectrum was analyzed in [263], and an interesting analysis of the static non-linear MIMO systems based on the separable signals was published in [264].

Probably the first serious attempt to deal with the non-linear distortions as a separate object of investigation, yet still within the linear system identification setting was [65]. Static or Volterra-like non-linear distortions of low order were tackled there from the measurement point of view, under harmonic excitations, within the deterministic setting. His aim was not to describe the non-linear distortions in general, but to get rid of them in concrete measurement and identification situations. To this purpose various kinds of harmonic signals were introduced [66-69], then a special kind of multisine signals was developed, to minimize the non-linear distortions for the assumed particular order of the non-linearity [70]. Later the investigations were extended to the concept of the best linear approximation to a Volterra system excited by the multisines with random harmonics (amplitudes and phases, or only phases), proposed by [30*]. Using Crest Factor minimization introduced in [85] in a series of

papers [72-73], [219-221, 223-225] presented heuristic comparison of modeling nonlinear errors if the Crest Factor minimization is also required.

Modeling non-linear systems with so-called output-error linear time-invariant second order equivalents (OE-LTI-SOE), within the Gaussian and the quasi-stationary framework, was introduced in [55-58], [125-127]. The analysis focused on the Non-linear Finite Impulse Response (NFIR) systems and the full characterization of stable causal OE-LTI-SOEs of such systems was given in [59-61]. The problem of the non-linear distortions was addressed via the notion of slightly non-linear system [58]. Seeking general conditions for the SOE of an NFIR system to be also a FIR system, separable signals were used as an extension to the notion of Gaussianity [59][3]. Bounds on the distance between the SOE and the linear part of the non-linear system were also studied.

Theoretically the most rigorous setting was the approach of Mäkilä and Partington, with the deterministic framework based upon the notion of the Generalized Harmonic Analysis (GHA) and quasi-stationary signals, drawing upon the normed space operator theory. In [134] the Frechet derivative was used to derive the best causal linear approximation of mildly non-linear systems. Beside the mean square error approximation, the absolute error approximation was also considered for smooth and non-smooth static non-linear systems in [135]. In [136-137] distribution theory of sequences was called in to refine the results obtained earlier for GHA and quasi-stationary signals. A very interesting notion of the nearly linear system and its LTI companion was introduced in [138], with the primary aim to investigate the controllability of a non-linear (NFIR) systems through the control of their LTI companions[4]. This work was extended in [139] to the linear approximations with the FIR and ARX parametrization, then in [140] to the notion of a non-linear companion system, providing also the state-space form for the linear companion.

## 1.3 Research focus

The reported work is based and extends the ideas put together in [27*, 28*, 30*], where the stochastic framework was proposed to deal with the Volterra-like non-linear distortions,

3 Although separable signals extend the properties of Gaussian signals with respect to the linear conditional expected value, their major disadvantage in the system identification is that the output of a linear system to a separable signal is not necessarily a separable signal. Separable signals are handy to identify systems with the non-linearity at the input (e.g. Hammerstein systems), they are not so useful where the non-linearity is hidden within the system, or at the system output (e.g. Wiener, or Wiener-Hammerstein systems) [148], [59-60].

4 Please note that the notion of a nearly linear system (and its linear companion) [138] is not comparable to the notion of a weakly non-linear Volterra series; see e.g. [15*]. The behavior of a nearly linear system, getting more and more linear as the input signal amplitude grows, does not reflect well what we observe in many practical measurement situations, where the signals are bounded, but the non-linearities (e.g. a cubic one, saturation, dead-zone) are not asymptotically linear. For small signals, the linear companion of such non-linear linear system can still yield large approximation errors, the related linear dynamic system of a weakly non-linear Volterra series, on the other hand, will always be closed to its linear component [162], [15*].

randomizing them through the randomization of the input signals, for which, due to the practical measurement reasons, random multisine excitations were used.

Contrary to the other contemporary approaches the aim was not only to produce the “best” linear approximation to a weakly[5] non-linear system, but to observe (and to influence) where and how the non-linear distortions manifest themselves in the measured linear non-parametric Frequency Response Function (FRF) (i.e. Empirical Transform Function Estimate – ETFE [125]). The theory was developed for input signals with a finite number of harmonics, and the asymptotic properties have been analysed when the number of harmonics tends to infinity[6].

The stochastic setting was essential to our purpose; otherwise there would be no possibility to qualify the error on the approximation. The stochastic setting makes it possible to design measurement procedures (here it was averaging) warranting the proper approximation of the theoretical limiting results (expected values) from the finite measurement data.[7]

Harmonic random signals are in the limit (in the number of harmonics) normally distributed, and even separable, the measured ETFE tends thus asymptotically to the FRF of the OE-LTI-SOE of a non-linear system described by a convergent Volterra series. Issues like stability, causality, memory length, impulse response structure, etc. are no investigated for the non-parametric FRF. However they may be pertinent to the parametric identification, which can be made after the analysis of the measured non-parametric FRF yields hints w.r.t. to the dynamics.

It is important to observe that the measurement of the non-parametric FRF precedes always, as the necessary introductory step, the parametric modeling in the frequency domain. The accurate judgement of the linear properties of the identified system, but also of the possible non-linear errors is essential for the successful further identification. The proposed methods

5 There are many definitions of being weakly non-linear or almost-linear (e.g. based on the coherent and incoherent output power, ratio of norms, ratio of coefficients, etc.), all sharing the notion that while nonlinear effects are essential and are observable, the linear behaviour dominates, and as a first approximation, the system can be considered linear.

6 Historically the starting point was the adoption in the 1980s by the Vrije Universiteit Brussel ELEC (Fundamental Electricity and Instrumentation) Department of the frequency-domain identification of the linear systems (contrary to the dominating then time-domain) and the following intensive research for the suitable deterministic excitation signals. From that time stems e.g. still state-of-the-art crest factor minimization technique [85]. Making research at the ELEC in the knowledge intensive signal processing methods ([24*-26*, 1*-*4*]) I joined the group which started to tackle the non-linear problems and became a member of the team working out the basic theory [27*-28*, 30*]. During multiple visiting periods at the ELEC I dealt with a number of open SISO (Single-Input Single-Output) problems and later I worked out the respective MIMO (Multi-Input Multi-Output) theory and solved some other open problems. From that time on this theory is intensively researched by a number of scientific and application groups.

7 Stochastic setting resembles the so called stochastic embedding of [129-130]. Stochastic embedding is a parametric framework where modeling errors are considered realizations of a random variable with a parametric distribution and the effort is to estimate these parameters. Our approach is nonparametric and the modeling errors will be classified into (nonparametric) systematic deterministic and stochastic components.

make it possible to measure the non-parametric FRF and to qualify the non-linear errors within the same experiments, minimizing the required measurement time[8].

In the following I review the principal assumptions underlying the research, and the basic theoretical concepts and results. Then I present my own research results, and finally I give a summary of the impact of these results on the applications.

Basic equations of the Volterra series in the time and the frequency domain under harmonics excitations can be found in [34, 198, 22, 9, 112]. The more specific results are referenced locally.

## 1.4 Research assumptions

The field of non-linear systems is too vast and too complicated to tackle with success any particular problem without carefully devised limiting assumptions. Their role is to bring the problem to the size and scope still valuable in practical modeling situations, yet admitting theoretical analysis and synthesis. The essential working assumptions were applied thus to the selection of the:

- non-linear system class,
- excitation signals,
- focus of the research.

The focus is the measurement methodology of the nonparametric linear frequency characteristics (Frequency Response Function - FRF), via its Empirical Transfer Function Estimate (ETFE [125]). Due to this reason the results were formulated in the frequency-domain for the input-output system descriptions. The information available in the time-domain and in the frequency-domain measurements is the same and the formulation of the measurement problem is in itself equivalent. Nevertheless the required information appears differently in the measurement data paving the way to the advantages stemming from different processing algorithms.

To the advantages of the frequency-domain belong the freedom in the selection of the frequencies where the model is matched to the measurements or freedom in restricting unwanted frequences, the possibility to model unstable systems, furthermore, if abiding to the periodic excitations (and periodic reference signals), no leakage bias on the ETFE, separation of the plant and the output noise modeling, possibility for the nonparametric noise modeling, possibility for modeling under close-loop conditions, finally (what is the topic of this work) the separation of the effect of the non-linearities and the output noise [204, 206, 125].

8 In recent developments the originally proposed harmonic random excitations are perturbed (power level, coloring) to verify how does it perturb the measured linear ETFE to gain indication about the possible structure of the nonlinear block model [63-64, 212-213, 114-116].

The non-linear system theory, as mentioned before, suffers from the multitude of possible models. Furthermore, the non-linear system identification is conditioned on the used excitation signals due to the presence of unavoidable model errors. Even if we intend to tackle practical situations when the level of the non-linear distortions is low, the choice of the class of the non-linear system models and the associated choice of the input signal class is important, as it is heavy in the consequences.

## On the system and signal models

In the present work convergent Volterra series were assumed as the model the measured systems. The particular usefulness of such model is dictated by the following (summarizing the wisdom of [23-25, 19, 21-22, 202, 211, 54, 59-60, 104, 137, 134, 162, 198, 188, 76]):

1. Natural (additive) way of how the linear and non-linear systems can be treated together, and the level of the non-linearity controlled;

2. A number of pragmatically important non-linear systems can be already modeled by finite, low order Volterra series;

3. Wide class of (even non-continuous) non-linear systems which can be approximated in the least square sense with the Volterra series;

4. Well developed frequency-domain representation;

5. Natural way of how the non-linear dynamics can be modeled;

6. Volterra models contain a number of practically important non-linear block models, i.e. Wiener-, Hammerstein-, and Wiener-Hammerstein models; and also Non-Linear Finite Response (NFIR) models;

7. Straightforward extension of the Single-Input Single-Output (SISO) models to the Multiple-Input Multiple-Output (MIMO) models;

8. Possibility to include a priori physical knowledge into the models (the number, the order, the symmetry, and the frequency bands of the Volterra kernels);

9. Volterra-series possess a unique steady state property and are Periodic-Input Same Periodic-Output (PISPO) systems, i.e. they do not generate subharmonics. Volterra series also yield almost periodic outputs to the almost periodic excitations.

Volterra models are not a universal tool and their expressiveness is limited. A number of interesting and important non-linear phenomena cannot be modeled well or at all with the Volterra series. As the Volterra series generalize the Taylor expansion, bifurcations, chaos, non-linear resonances, generation of sub-harmonics, etc. are out of reach for the Volterra models.

The second principal choice applied to the input excitation signals. Asymptotically Gaussian periodic signals were adopted (over non-periodic signals) due to:

1. Less problems in the ETFE measurements (no leakage due to transients);

2. It is easy to distinguish or to separate the input signal properties (periodic) from the noise properties (non-periodic);

3. An easy introduction of the randomness into the signal (via random phases and/or random spectral amplitudes);

4. A free hand in the construction of different signal characteristics by manipulating the spectral properties (coloring), the frequency grid, and the phases;

5. An easy realization of such signals in modern signal generators, meaning that the developed theory is straightforward enough to be widely used in practice;

6. Possibility to model approximately the non-periodic signals also by choosing high enough number of the harmonics in a bounded frequency band.

7. Considering that in many measurement application areas Gaussian (noise) signals are traditionally used, the proposed excitations signals provide portability of the new methods combined with additional advantages;

8. Gaussian signals are "non-linearity-friendly" (i.e. when applied to static non-linear systems).

9. Last, but not least harmonic signals can be analytically integrated and/or differentiated sparing error prone signal processing where different forms of the excitations are jointly needed (e.g. velocity, angular position, acceleration [227]).

It is important to mention, that the approach presented in the following and based on the characterization of the non-linear distortions as a noise and bias on the linear FRF, is valid for any convergent Volterra series, i.e. for any smooth enough non-linear dynamic system. If however the non-linear behaviour is strong, using such linear model does not make sense, as the excessive non-linear noise will make the measurements long and costly and then still the bias to the linear FRF will be too large to get a feeling of what the linear system dynamics is really like. Despite hence the universal validity of the results, their practical usage is confined to the situations where the level of non-linearity is small and/or the order of the non-linear system is low.

## Research methodology

The aim of the research was to separate in the measured FRF of a weakly non-linear system the systematic non-linear effects (resident and enduring in the measurement results) from other, noise-like non-linear effects that are removable with a suitable post-processing.

To this end it is not enough to excite the system (like a linear system) with a single deterministic signal. A manageable stochastic process is needed, to excite the system with its sample functions, one after the other. The non-linear system will answer to every input sample function differently, camouflaging the systematic error (with respect to the linear system) with "non-linear noise" of varying behavior (see Example 1.1.1).

Considering that the input to the investigated system is a stochastic process, the system output will be a stochastic process alike and in principle the systematic (non-zero expected value) and "noise-like" (zero expected value) error components can be grasped by computing higher order moments. Higher order moment estimates can be computed by assembly averages. In the measurement technical language it means that we should generate independent sample functions from the input stochastic process, apply them as excitations to the system, and then average the individual measurement results.

The used (usual) sample mean is appropriate because (1) sample averaging is a consistent estimate of the theoretical expected value. The practical bound limiting the number of averages is only the measurement duration (cost of the equipment, not met stationarity conditions, etc.); (2) for Gaussian signals the average is also a minimum variance estimate; (3) an average can be computed recursively without unnecessary data storage (important due to (1))[9].

At the beginning of the research the used excitation signal was the periodic Gauss noise (so called periodic noise). The break-through was however brought by the observation that if the phases of the multi-harmonic signals are random, independent, and uniformly distributed on the unit circle, then with the increasing number of the harmonics such signal - called random multisine - tends to a Gaussian signal. Furthermore the FRF measurements performed with random multisines tend asymptotically to the measurements performed with Gaussian signals, also in case of the non-linear systems.

## 1.5 Summary of the scientific results

The basic concepts of the underlying theory were established at the VUB ELEC Department, but due to the frequent research visits (from University of Glamorgan (UK), University of Warwick (UK), Linköping University (S), KTH Royal Institute of Technology (S), and last but not least the BME researchers) the developments were discussed continuously almost on the daily basis and were published in deliberately jointly authored publications (at the Department of Measurement and Information Systems, BME, these contacts took between 1997-2006 the organizational form of 4 successfully concluded Hungarian-Flemish Bilateral Research Grants).

Among the scientific results there are thus results, especially from the beginning of the years long research, where the identification of the individual responsibility of the authors is not possible, but there are also results, which I can responsibly call my own, despite the joint authorship of the publications. Such inseparable results merit mentioning because they reveal the research process and provide the context for presenting the strictly individual contributions. After summary review of the joint results, the individual contributions are structures into the research Theses, to make a clear distinction in the presentation.

9 In recent developments, due to the industrial technological demands, discrete level excitation signals were also considered leading to the “nonlinear noise” of different properties where the processing of the measured results was based on median filtering instead of simple averaging [210, 268-269].

## Jointly achieved research results

The starting point was the development of the theory of the nonparametric best linear approximation to weakly non-linear SISO (Single-Input Single-Output) systems. We gave the mathematical structure of the approximation and established the theoretical and practically verifiable properties of its components. This theoretical approach was extended, supported by analysis, simulations, and practical suggestions, to the measurement technology of the linear frequency characteristics ([27*-28*, 30*-31*, 33*-37*, 40*, 42*, 45*]). (Sections 2.2-2.3-2.4)

One of the important design parameters in the measurement technology was the frequency grid of the multisine excitations, frequencies where the excitation injects energy into the measured system. The selection of the frequency grid in the non-linear measurements strongly influences the properties of the measurable quantities. In our research many kinds of grids were considered pursuing inherent theoretical and practical possibilities ([32*, 247, 162, 170]) (Sections 3.2-3.3)

Expected values measured on the Volterra series with the harmonic signals with a large number of harmonics theoretically correspond to the Riemann integral sums. This made it possible to evaluate theoretically the robustness of the asymptotic properties of the systematic error (Best Linear Approximation - BLA) measurements. [17*]. (First part of Section 3.6)

An interesting and pragmatically important issue was the fact that in the designed (Best Linear Approximation) measurements the non-linear noise variance is directly measurable, but not so the non-linear bias (i.e. the systematic error on the FRF). Research attempted to clarify how much these two error components are interdependent, with the prospect to estimate the systematic FRF error from the measurements of the nonlinear noise variance. [47*-48*]. (Section 2.5, equ. 2.5.14)

Finally if the non-linearity is weak, linear system analysis may perhaps indicate that there are no problems with the stability in the close loop. The amplitude of the signals in the feed-back loop may nevertheless increase so much that the non-linear effects will appear with potentially unfavorable consequences. Some research was done to predict such situation well in advance ([38*, 41*, 43*]). (Section 6.3)

## Contributions to the SISO theory

The strictly joined research not only established the basic theory but also led to the formulation of a vast number of problems (not everyone solved yet) where the scientific contributions can be already attributed to the concrete individuals. In case of SISO systems, I have dealt in particular with various aspects of the multisine design and the systematic nonlinear error in more specific measurement situations. These particular research problems and results are formulated as independent research Theses.

**Thesis Group 1. Design methodology of multisine excitation signals – SISO systems**

Here I have collected results where the topic of the research was the flexibility of the multisine signals, shaping them to the application requirements and then analysing the related problem of how similar are the measurement results if the excitation signals differ in the design.

**Problem Topic 1.1** Before the theoretical and the practical well-founding of the multisine signals with a large number of harmonics the prevalent excitation signal in a number of measurement fields was the Gaussian noise. In theory both signals are asymptotically equivalent, but for the credibility the non-asymptotic behavior of the multisines had also to be examined.

**Thesis 1.1. Design considerations how to chose multisine excitation signals**

Based on the investigation I have proposed a methodology how to use multisine signals if nonasymptotic behavior is also important. In measuring the FRF of a weakly nonlinear systems I propose to use odd-odd (double odd: every second odd harmonic frequency) random phase multisine considering that:

- the measured FRF is the same as measured on the system with the Gaussian signal,

- the uncertainty of the measurement is largely reduced due to the drop-out of the effects of the even and in part of the odd nonlinearities (see Fig 3.5.1),

- it is possible to separately measure the even and odd nonlinearities.

It can be also stated that in case of the Gaussian noise excitation the required frequency band limitation and the amplitude censoring amplify the bias on the measured FRF, if the measured system does contain odd nonlinearities. Furthermore the usual amplitude censoring by $\pm 3\ \sigma$ is not enough if the nonlinearity is involved (I propose censoring by $\pm 4\ \sigma$). [49*-50*, 5*, 8*, 12*, 39*]. (Sections 3.2, 3.5, Th 3.2.1)

**Problem Topic 1.2** Frequency grid is the design parameter of the multisine signals. Frequency grids of various structures can be successfully used to solve special measurement problems. Important question is how consistent could be the measured FRF Best Linear Approximations, if the used multisine signals differ in the definition of the frequency grid?

**Thesis 1.2. Unifying asymptotic results for different frequency grids of the multisine excitation with the theory of the uniformly distributed sequences**

I have determined that if the frequency grids are are modeled as the uniformly distributed sequences with increasing resolution, then the error between the measurements obtained from different frequency grids gets smaller and is of the order of the magnitude of the grid resolution [20*, 46*]. (Section 3.6, from Def 3.6.1, Th 3.6.2)

**Problem Topic 1.3** In measuring non-linear systems an important issue is the control of the amplitude density of the excitation signals. It is a tool to put signal energy at the amplitudes relevant for the nonlinear behavior.

**Thesis 1.3. Designing application dependent wide band multisine excitation signals**

I have determined that the phases of the multisine, which are a design parameter, with suitable phase iterating algorithms, can be used to the user-demanded forming of the amplitude density of the multisine signal. Moreover, with some trade-off, the crest factor minimization can be also included [29*]. (Section 3.4, Algorithms 3.4.1-3.4.3)

**Thesis Group 2. Properties of the systematic nonlinear errors – SISO systems**

Here I have collected results of investigating in more detail the properties of the systematic error component and the mutual relation of the systematic and the stochastic nonlinear errors.

**Problem Topic 2.1** The systematic error on the FRF Best Linear Approximation measured in the presence of the non-linearity is in itself not measurable, however the variance of the non-linear noise is measurable. Important problem is to find out how the measurable (stochastic) error can be used to estimate the nonmeasurable (systematic) error.

**Thesis 2.1. Establishing bounds on the systematic error of the Best Linear Approximation**

I have determined that for the static polynomial non-linearity the cubic system can be considered the worst case instance and based on it I have formulated the worst-case estimate of the non-measurable error. I have extended the estimate heuristically to the Wiener-Hammerstein system model [9*]. (Section 2.5, Ths 2.5.3-2.5.4)

**Problem Topic 2.2** The coherence function is a well known tool in the recognition and examination of the non-linear systems (considered as black-box models). If the non-linear system admits the non-linear additive noise model, how this additional knowledge will influence the behavior of the coherence function?

**Thesis 2.2. Clarification of the relation of the Best Linear Approximation and the coherence function**

I have established that the coherence function can be expressed with the components of the Best Linear Approximation (non-linear bias and noise), as well as that the nonlinearity indicating properties of the measured coherence function are consistent with the behavior of the components of the Best Linear Approximation. To this end I have investigated general and also more specific non-linear system structures [10*-11*]. (Section 2.5, Ths 2.5.1-2.5.2)

**Problem Topic 2.3** The researched question was whether the product of the Best Linear Approximations to the superposed non-linear systems can be used to build an acceptable approximation of the whole system.

**Thesis 2.3. Analysis of the superposition of the SISO systems from the point of view of the Best Linear Approximation**

I have determined that in the frequency domain where both system components show high coherence, the product relation of the linear system theory (i.e. that the frequency

characteristic of the superposition is the product of the frequency characteristics of its components) remains valid [10*-11*]. (Section 6.1)

**Problem Topic 2.4** The Best Linear Approximation FRF is traditionally measured as a sample mean (for linear systems the minimum variance Maximum Likelihood estimate for Gaussian excitations). This minimum variance is the basis of the usual measurement time vs. measurement quality trade-off. However due to the non-linearity the sample mean is no more an optimal estimate; its variance is not an attainable theoretical minimum and can be improved. By using a limited a priori knowledge about the measured system this trade-off can be made sharper.

**Thesis 2.4. Reducing the measurement time of the Best Linear Approximation by Monte Carlo averaging methods**

I have developed an alternative method of measuring the Best Linear Approximation where a better trade-off between the measurement duration and the measurement accuracy can be achieved by using the Monte Carlo variance reduction techniques, without increasing the complexity of the measurement protocol [23*]. (Section 6.4)

## Contributions to the MIMO theory

In the real life problems there are usually more independent effects cumulated towards a common output. It is important thus to handle Multiple-Input (Multiple-Output) models also. It is also expected that the earlier SISO case should appear as a special case within the MIMO theory.

**Thesis Group 3. Properties of the systematic nonlinear errors – MISO systems**

Based on SISO Best Linear Approximation theory I have developed the ground results in the MIMO (MISO) Best Linear Approximation theory, focusing on the description of the systematic errors and the equivalence of the measurements for different kinds of excitations. I also have extended to the MISO case some of the more specific results developed for the SISO systems.

**Problem Topic 3.1** In case of multiple input systems separate Best Linear Approximations can be defined for every input-output signal channel. In computing such BLA system other inputs act as disturbances and complicate the computation of the non-zero expected values. Similarily to the SISO case some kernels do not contribute to the systematic errors, however the general picture is much more complicated.

**Thesis 3.1. Developing general MIMO BLA theory from the point of view of the systematic FRF error**

I have determined that using the random multisine BLA SISO measurement technique, the 2-dim MISO cubic system excited with independent random multisines defined on a common frequency grid, can be modeled (similarly to the SISO case) as a 2-dim linear FRF

characteristics and the nonlinear output noise [22*, 13*-14*, 44*]. (Sections 4.1-4.2, Ths 4.1.2, 4.1.3, 4.1.4, 4.1.5). As a full generalization I have determined that the multiple-input Volterra system excited with the random multisine excitations can be expressed as a linear BLA FRF system network, completed with the output non-linear noises. The earlier SISO and 2-dim results are the special cases of the general case [15*, 20*] (Sections 4.5-4.6, Ths 4.5.1, 4.5.2, 4.6.1)

**Problem Topic 3.2** Similarly to the SISO systems (see Thesis 2.3) an interesting question for the MIMO systems is how robust is the FRF BLA matrix from the point of view of further distortions superposed on the system.

**Thesis 3.2. Superposition of the MIMO systems from the point of view of the Best Linear Approximation**

I have established that the Best Linear Approximation to a MIMO system is robust when the excitations are nonideal and are modelled by the output of a nonlinear Volterra MIMO system. I have established that the nonlinear distorting effects cause larger FRF characteristics errors than the linear distorting effects of similar amplitude, the distorting effects originated in the cross input-output signal paths cause larger errors than the distorting effects in direct signal paths, as well as that the FRF phase errors increase faster than the FRF amplitude errors. [18*] (Section 6.2)

**Thesis Group 4. Design considerations about the excitation signals – MISO systems**

Multiple inputs introduce additional freedom into the excitation design. Not only we dispose the design parameters at a single channel, we can also decide how the excitations at different input channels and in different experiments could be related to yield better measurement results. The challenge lies in designing excitations that are better from the measurement technical point of view (i.e. warranting equivalent measurements at a lower cost). In case of multiple inputs a distinction must be made between the 2-dimensional and more-dimensional weakly nonlinear systems. Surprisingly, in low nonlinear order 2-dimensional system measurements the traditional (linear system theory) noise attenuating techniques are henceforward applicable, but this advantage is lost for higher order nonlinearity and/or higher input dimensions. There new noise attenuation techniques has to be developed.

**Problem Topic 4.1** In linear MIMO measurements an important part of the measurement methodology is the input design suitable for the noise cancellation. I have investigated whether such noise cancelling methods could be also used to the advantage in the BLA measurements to cancel the non-linear noise.

**Thesis 4.1. Optimizing excitations used in the 2-dim BLA theory**

I have determined that the linear noise attenuating technique is in the same way effective in case of the 2-dim cubic BLA measurements. I have also determined that this approach is not suitable in case of the system of a higher nonlinear order or the higher number of the inputs. [22*, 13*-14*, 44*] (Sections 4.3, 4.4)

**Problem Topic 4.2** The MIMO extension to the SISO BLA theory was formulated for the random multisine signals. An important question however is whether the excitation signals asymptotically equivalent in the SISO theory (multisine, periodic noise, Gaussian noise) are similarly equivalent in the MIMO case.

**Thesis 4.2. Equivalence of the excitations from the point of view of the systematic (MIMO) FRF error**

I have determined that the Gauss noise, the periodic noise, and the random phase multisine signal classes are asymptotically equivalent (if the number of harmonics is increasing and the spectral properties of the signals are comparable) also in the case of nonlinear MIMO Volterra systems in a sense that using these excitation signals the measured multidimensional FRF BLA systems tend in the limit to the same transfer characteristics matrix. [21*] (Section 5.4, Ths 5.4.1, 5.4.2)

**Problem Topic 4.3** In the detailed presentation of the results referred in Thesis 2.4 one can see that in the multidimensional case the inverse of the input matrix amplifies uncertainty, even in case of the multisine excitations (contrary to the SISO case). An important question is whether this situation can be improved or not?

**Thesis 4.3. Special orthogonal random multisines**

With the introduction of orthogonal random multisines I have developed a new efficient method of measuring the Best Linear Approximation of the MIMO Volterra systems. I have proved that the newly introduced excitation signals are equivalent in the sense of the Thesis 2.4 to other listed excitation signals, but result in an essentially lower level of the non-linear noise experienced on the measured FRF characteristics. [15*-17*] (Sections 5.1-5.2, Ths 5.1.1, 5.2.2, Lemma 5.2.1)

## 1.6 Review of the content

Section 2 introduces the general theory for the single-input single-output (SISO) systems. First a simple example is shown, without the in-depth formalism, providing the feeling of the problem and yet presenting every important issue and question, which later on will be elaborated in detail (Section 2.1). Section 2.2 presents the main result, i.e. the additive non-linear noise model to the non-linear system, and then the properties of the systematic non-linear distortion (Section 2.3) and the stochastic non-linear distortion (Section 2.4) are analyzed. As mentioned before, the Volterra models cover the usual non-linear block models. The developed theory is applied to them in Section 2.5. Finally the question of the mutual analysis of the distorting bias and variance is considered in Section 2.6.

The purpose of Section 3 is to show the enormous flexibility of the multisines as the excitation signals. Section 3.1 discusses tunable free parameters, i.e. the frequency grid, the amplitude spectrum, and the phases. Then various types of the multisines are presented (Section 3.2), discussing their design and the intended effects on the measured linear

approximation. The aim of Section 3.3 is to show, that the measurement results obtainable with the multisines are comparable for the increasing number of harmonics with the results obtained with traditional excitation signals (Gaussian noise and periodic noise).

Section 4 extends the results of Sect 2. to the multiple-input multiple-output (MIMO) Volterra models. The main result, the multidimensional additive non-linear noise model, is developed in Section 4.1 and the properties of the non-linear biases and non-linear noise variances are evaluated in Section 4.2 and 4.3.

In MIMO measurements excitations are applied to more than one input simultaneously. The excitation signals must be designed thus not only in themselves, but also in relation to the signals applied at other input points. Section 5 presents excitation design problem for the MIMO systems. Free design parameters are discussed in Section 5.1. Various excitation schemes and their effect on the model developed in Section 4 are analyzed in Section 5.2. Finally in a manner similar to Section 3.3 the equivalence of the measured results for various MIMO excitation schemes is considered.

Section 6 presents some applications of the introduced theory. In Section 6.1 simple measurement application are shown, based on the literature. Sections 6.2 and 6.3 elaborate on the problem of non-linear distortions in cascaded systems, for SISO and for MIMO systems respectively. Besides modeling practical measurement problems cascaded systems make it possible to study the robustness of the developed theoretical tools. In Section 6.4 an attempt is made to qualify stability problems in the non-linear feedback system based on the additive non-linear noise model.

In the development of the additive non-linear noise model, necessarily, plenty of problems remain still open. Section 7 discusses some of the more interesting and difficult open research issues.

Finally the Appendices contain the most lengthy proofs and examples.

# 2. General SISO theory

An unexpected non-linearity can deceive the user not familiar with the non-linear phenomena, or the user versed solely in the linear measurement or identification methods, making her/him thinking that the nature of the problem is quite different. Consider simulated measurements in the Example 1.1.2. The FRF of a linear system, otherwise smooth, can become scattered and acquires the "noisy look" if the excitation signal hits the hidden non-linear component. Besides scattering, the measured FRF will also be biased, this though is more difficult to discern, as the true frequency dependence of the characteristic is not known in advance. It is common knowledge, however, that without the noise, the FRF measurements should yield more or less smooth functions. The visible "noisiness" can thus easily be taken as the proof that the output noise is the real problem here (especially as it is usually present). The scattering is caused by the non-linear mechanism of summing various harmonic components in the input signal and shifting them to different places (frequencies) along the frequency axis.

**Example 2.1: Scattering of the frequencies due to a non-linearity.** Let the system $y(t) = u(t) + \varepsilon u^3(t)$ be excited by the input signal containing two harmonic components with frequencies $\omega_1$ = 1 and $\omega_2$ = 4. Then the output signal will have harmonics at frequencies ±1, ±2, ±3, ±4, ±6, ±7, ±9, ±12.

## 2.1 Problem introduction via simplified examples

In this Section the main points of the theory developed formally later will be shown in simplified measurement examples, without strict definitions and derivation.

A. **Measuring a linear system with a random signal**. Assume that for linear FRF measurements a periodic random $u(t) = u(t, \xi)$ stochastic process input signal is used. Let the collected measurement data come from noisy linear system, where the zero mean output noise is similarly a stochastic process bound to a random event $\zeta$, then:

$$y(t) = G_0(q)u(t,\xi) + n(t,\zeta), \text{ or} \tag{2.1.1}$$

in the frequency domain ($l$ is the discrete frequency)[10]: $Y(l) = G_0(l)U(l,\xi) + N(l,\zeta)$. (2.1.2)

The FRF is computed (from $k$ = 1… $M$ experiments) as:

$$\hat{G}(l) = \frac{\frac{1}{M}\sum_k Y^{(k)}(l)\overline{U}^{(k)}(l,\xi)}{\frac{1}{M}\sum_k |U^{(k)}(l,\xi)|^2} = G_0(l) + \frac{\frac{1}{M}\sum_k N^{(k)}(l,\zeta)\overline{U}^{(k)}(l,\xi)}{\frac{1}{M}\sum_k |U^{(k)}(l,\xi)|^2} \tag{2.1.3}$$

This estimate is unbiased, i.e.: $E\{\hat{G}(l)\} = G_0(l)$, (2.1.4)

because the output noise and the excitation are independent:

10 Considering that $Y$ is the output to the stochastic input and contains also random noise, formally we should write: $Y(l) = Y(l, \xi, \zeta)$ and similarly $G(l) = G(l, \xi, \zeta)$, however these arguments have been omitted for clarity.

$$E_{\xi\zeta}\{\frac{\frac{1}{M}\sum_k N^{(k)}(l,\zeta)\overline{U}^{(k)}(l,\xi)}{\frac{1}{M}\sum_k |U^{(k)}(l,\xi)|^2}\} = \frac{\frac{1}{M}\sum_k E_\zeta\{N^{(k)}(l,\zeta)\}E_\xi\{\overline{U}^{(k)}(l,\xi)\}}{\frac{1}{M}\sum_k E_\xi\{|U^{(k)}(l,\xi)|^2\}} = 0. \qquad (2.1.5)$$

and its variance depends on the degree of averaging introduced in (2.1.3).

**Note**: For the random multisine signals considered later: $E_\xi\{U(l,\xi)\overline{U}(l,\xi)\} = \hat{U}^2$.

B. **Measuring a non-linearly distorted system with a deterministic signal**. To measure the linear FRF it is enough to keep a single input realization (i.e. let $\xi = \xi_0$) and to average the results to get rid of the output noise $N(l,\zeta)$. Reassured that we can thus simplify the measurement we continue overlooking the point, that now the measurement data is coming from a noisy non-linear system:

$$Y(l) = G_0(l)U(l,\xi_0) + Y_{NL}(l,\xi_0) + N(l,\zeta), \qquad (2.1.6)$$

with $Y_{NL}(l,\xi_0) = V[U](l,\xi_0)$ the non-linear part of the system. The FRF estimate is now:

$$\hat{G}(l) = \frac{Y(l)}{U(l,\xi_0)} = \frac{G_0(l)U(l,\xi_0) + Y_{NL}(l,\xi_0) + N(l,\zeta)}{U(l,\xi_0)} = G_0(l) + \frac{Y_{NL}(l,\xi_0)}{U(l,\xi_0)} + \frac{N(l,\zeta)}{U(l,\xi_0)} \qquad (2.1.7)$$

Averaging gets rid of the output noise, but the non-linear term is not random:

$$E\{\hat{G}(l)\} = E\{\frac{Y(l)}{U(l,\xi_0)}\} = G_0(l) + \frac{Y_{NL}(l,\xi_0)}{U(l,\xi_0)}, \qquad (2.1.8)$$

and the measurement results are heavily distorted.

C. **Measuring a non-linearly distorted system with a random signal**. What will happen if we return to our original random input: $Y(l) = G_0(l)U(l,\xi) + Y_{NL}(l,\xi) + N(l,\zeta)$ ?

Now:

$$\hat{G}(l) = \frac{\frac{1}{M}\sum_k Y^{(k)}(l)\overline{U}^{(k)}(l,\xi)}{\frac{1}{M}\sum_k |U^{(k)}(l,\xi)|^2} = G_0(l) + \frac{\frac{1}{M}\sum_k Y_{NL}^{(k)}(l,\xi)\overline{U}^{(k)}(l,\xi)}{\frac{1}{M}\sum_k |U^{(k)}(l,\xi)|^2} + \frac{\frac{1}{M}\sum_k N^{(k)}(l,\zeta)\overline{U}^{(k)}(l,\xi)}{\frac{1}{M}\sum_k |U^{(k)}(l,\xi)|^2}$$

and:

$$E_{\zeta\xi}\{\hat{G}(l)\} = G_0(l) + E_\xi\{\frac{\frac{1}{M}\sum_k Y_{NL}^{(k)}(l,\xi)\overline{U}^{(k)}(l,\xi)}{\frac{1}{M}\sum_k |U^{(k)}(l,\xi)|^2}\} + E_{\zeta\xi}\{\frac{\frac{1}{M}\sum_k N^{(k)}(l,\zeta)\overline{U}^{(k)}(l,\xi)}{\frac{1}{M}\sum_k |U^{(k)}(l,\xi)|^2}\}. \qquad (2.1.9)$$

The second expected value is not a problem, as both random components are independent and zero mean. What about the non-linear expected value in the middle?

If the input signal has bounded realizations for every $\xi$ and the non-linear system V[.] is well behaving (e.g. continuous, BIBO stable, etc.), then the output of the non-linearity $Y_{NL}(l,\xi)$

will also have bounded realizations and the $\hat{G}_{NL}(l,\xi)=\dfrac{\frac{1}{M}\sum_k Y_{NL}^{(k)}(l,\xi)\overline{U}^{(k)}(l,\xi)}{\frac{1}{M}\sum_k |U^{(k)}(l,\xi)|^2}$ random variable for every frequency $l$ will have finite expected value and finite variance, which symbolically can be written as:

$$\begin{aligned}\hat{G}(l,\xi,\zeta) &= G_0(l)+E_\xi\{\hat{G}_{NL}(l,\xi)\}+(\hat{G}_{NL}(l,\xi)-E_\xi\{\hat{G}_{NL}(l,\xi)\})+G_N(l,\xi,\zeta)\\ &= G_0(l)+G_B(l)+G_S(l,\xi)+G_N(l,\xi,\zeta)\end{aligned} \qquad (2.1.10)$$

where $G_0(l)$ is the true linear FRF, $G_B(l)$ is systematic non-linear distortion, part of the measured FRF after averaging, $G_S(l,\xi)$ is a zero mean non-linear stochastic distortion, and $G_N(l,\xi,\zeta)$ is also a zero mean distortion caused by the output noise. Then:

$$E_{\xi\zeta}\{\hat{G}(l)\}=E_{\xi\zeta}\{\hat{G}(l,\xi,\zeta)\}=G_0(l)+G_B(l), \qquad (2.1.11)$$

and we will see later on that it is also possible to have:

$$Var[\hat{G}(l)]=Var[G_S(l,\xi)]+Var[G_N(l,\xi,\zeta)]. \qquad (2.1.12)$$

It is important to note that a part of the non-linear distortion observed in the transfer term $\dfrac{Y_{NL}(l,\xi)}{U(l,\xi)}$ has been eliminated by being pushed into the zero mean stochastic component, which can be got rid by averaging, already employed to get rid of the output noise, and this without any particular strict assumptions on the non-linear distortion.

D. **Measuring low order Volterra systems with random multisines**. Assume now that for the FRF measurements so called uniform random phase multisine signal is used, i.e. a multi-harmonic signal of period *N*, with *M* harmonic components distributed on a particular frequency grid, with independent random phases and equal spectral amplitudes, normalized to unit power, i.e.:

$$\begin{aligned}&u(t)=\sum_{k=1}^{M}\hat{U}(k)\cos(\omega_k t+\phi_k)=\sum_{k=-M,\,k\neq 0}^{M}U(k)\,e^{j2\pi k\,t/N}\\ &U(k)=\tfrac{1}{2}\hat{U}(k)e^{j\phi_k},\quad 2\sum_{k=1}^{M}|U(k)|^2=1,\quad |U(k)|^2=1/2M,\quad \hat{U}^2=2/M\end{aligned}, \qquad (2.1.13)$$

and assume, that the non-linear part of the system is a 3rd order Volterra system, which in the frequency domain can be written as ($k_1$, $k_2$, $k_3 \neq 0$):

$$\begin{aligned}Y_{\mathrm{NL}}(l) &= \sum_{k_1=-M,}^{M}\sum_{k_2=-M,}^{M}G_3(k_1,k_2,k_3)U(k_1)U(k_2)U(k_3)\\ &= (2M)^{-3/2}\sum_{k_1=-M,}^{M}\sum_{k_2=-M,}^{M}G_3(k_1,k_2,k_3)\,e^{j\phi(k_1)}\,e^{j\phi(k_2)}\,e^{j\phi(k_3)}\end{aligned}, \qquad (2.1.14)$$

where $G_3(.,.,.)$ is so called (symmetric) Volterra kernel of the 3rd order, and $k_1+k_2+k_3=l$.

Now:

$$E\{\frac{Y_{NL}(l,\xi)}{U(l,\xi)}\} = E_{\phi}\{\frac{Y_{NL}(l)}{U(l)}\}$$
$$= E_{\phi}\{\frac{(2M)^{-3/2}\sum_{k_1=-M,}^{M}\sum_{k_2=-M,}^{M} G_3(k_1,k_2,k_3)e^{j\phi(k_1)}e^{j\phi(k_2)}e^{j\phi(k_3)}}{U(l)}\} \qquad (2.1.15)$$
$$= (2M)^{-1}\sum_{k_1=-M,}^{M}\sum_{k_2=-M,}^{M} G_3(k_1,k_2,k_3)E_{\phi}\{e^{j\phi(k_1)}e^{j\phi(k_2)}e^{j\phi(k_3)}e^{-j\phi(l)}\}$$

Considering that at different frequencies phases are independently distributed on the unit circle, the expected value is nonzero only if the frequencies are suitably paired within the sum, e.g. $k_1 = l$, or $k_2 = l$, or $k_3 = l$. Taking into account that: $\varphi(-k) = -\varphi(k)$:

$$E_{\varphi}\{e^{j\varphi(k_1)}e^{j\varphi(k_2)}e^{j\varphi(l-k_1-k_2)}e^{-j\varphi(l)}\} = E_{\varphi}\{e^{j\varphi(l)}e^{j\varphi(k_2)}e^{j\varphi(-k_2)}e^{-j\varphi(l)}\} = 1 \qquad (2.1.16)$$

and with this: $E\{\frac{Y_{NL}(l,\xi)}{U(l,\xi)}\} = G_B(l) = 3(2M)^{-1}\sum_{k=-M,}^{M} G_3(l,k,-k) \approx 3\int_B G_3(l,f,-f)\,df$ (2.1.17)

for large $M$, in some frequency band $B$ (factor 3 comes from the ways frequency $l$ could be paired). After the averaging the FRF results thus in:

$$\hat{G}(l) = G_0(l) + 3(2M)^{-1}\sum_{k=-M,}^{M} G_3(l,k,-k) = G_{BLA}(l)\,. \qquad (2.1.18)$$

Taking together the points C and D we should note that:

(a) The obtained results permit to write down the original measurement problem (no output noise is assumed for clarity) as:

$$Y(l) = G_{BLA}(l)U(l) + Y_S(l)\,, \qquad (2.1.19)$$

i.e. under random multi-harmonic excitations the output of the Volterra system looks like the output of a linear system distorted by an additive noise;

(b) This additive noise is naturally a function of the input signal (as the problem is non-linear);

(c) Non-linear errors have been decomposed. A part of the non-linear effects went into the systematic distortion on the measured FRF, another part into the non-linear stochastic component, which can be handled with traditional tools; with this development a non-linear system becomes now easier to analyse;

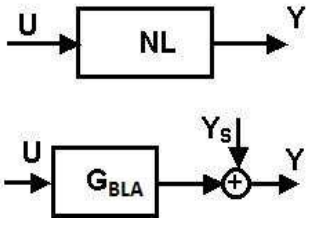


**Fig. 2.1.1:** Non-linear SISO system and its the additive non-linear noise source model.

(d) Numerous questions haven't been raised and addressing them requires a more formal setting, e.g.: what class of non-linear systems can we deal with, what class of excitations can

we use, what are the properties of the non-linear noise, what can we tell about the non-linear distortions if the system class or the excitation class is narrowed to well defined practical cases, etc.?

(e) As it was visible in the derivations, the number of the harmonics and their frequencies weren't particularly critical to the final result, especially interesting are thus the questions related to the frequency grid and the asymptotic behaviour of the excitations.

## 2.2 Best Linear Approximation of SISO Systems

### System and signal models

We will consider Volterra systems or systems being limits of convergent Volterra series. The reasons of this choice had been outlined in Section 1.3.

**Definition 2.2.1: Single-Input Single-Output (SISO) finite *K*th order Volterra system**. The output of such system can be written in time domain as [22, 198]:

$$y(t) = V^{(K)}[u](t) = \sum_{\alpha=1}^{K} y^{\alpha}(t) = \sum_{\alpha=1}^{K} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} g_{\alpha}(\tau_1, \dots, \tau_{\alpha}) \prod_{i=1}^{\alpha} u(t-\tau_i)\, d\tau_i \qquad (2.2.1)$$

where $g_\alpha$ is the time-domain Volterra kernel of order $\alpha$th. The primary domain of analysis will be the discrete frequency domain, where for the periodic inputs the so called fundamental frequency-domain formula [22, 25] is valid:

$$Y(l) = V^{(K)}[U](l) = \sum_{\alpha=1}^{K} Y^{\alpha}(l) = \sum_{\alpha=1}^{K} \sum_{k_1,\dots,k_{\alpha-1}=-M}^{M} G_{\alpha}(k_1, k_2, \dots, k_{\alpha-1}, k_{\alpha}) \prod_{i=1}^{\alpha} U(k_i) \qquad (2.2.2)$$

where $l = \Sigma\, k_i$, $i = 1 \dots \alpha$, is a discrete frequency, $G_\alpha$ is a symmetrized frequency-domain kernel of order $\alpha$ th, and $M$ is the number of harmonics present in the input signal. ■

**Definition 2.2.2: Single-Input Single-Output (SISO) Volterra series**. The Volterra series is defined by the convergent series:

$$y(t) = V[u](t) = \sum_{\alpha=1}^{\infty} y^{\alpha}(t) = \sum_{\alpha=1}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} g_{\alpha}(\tau_1, \dots, \tau_{\alpha}) \prod_{i=1}^{\alpha} u(t-\tau_i)\, d\tau_i \qquad (2.2.3)$$

or in the frequency domain by:

$$Y(l) = V[U](l) = \sum_{\alpha=1}^{\infty} Y^{\alpha}(l) = \sum_{\alpha=1}^{\infty} \sum_{k_1,\dots,k_{\alpha-1}=-M}^{M} G_{\alpha}(k_1, k_2, \dots, k_{\alpha-1}, k_{\alpha}) \prod_{i=1}^{\alpha} U(k_i) \qquad (2.2.4)$$

where the notation follows Def. 2.2.1. Kernels $G_\alpha$ are bounded by max $|G_\alpha| = M_\alpha$. The series is convergent for every $l$, if the input signal is normalized to unit power and has uniformly bounded spectral amplitudes $|U| \le M_U / \sqrt{2M}$ , furthermore if together [30*]:

$$\sum_{\alpha=1}^{\infty} M_{\alpha} M_U^{\alpha} < \infty \qquad (2.2.5)$$

■

**Note**: The above conditions on signals and kernels are equivalent to the conditions stated in [22, 25] for the validity of the fundamental frequency-domain formula (2.2.2), where it is required that the bounded ($\|u\|_\infty < \infty$) input signal should be of bounded variation over one period, otherwise the fundamental frequency-domain formula won't converge absolutely.

**Corollary 2.2.1**: **Random multisine signal (2.1.13) is of bounded variation over one period.**

**Proof**: For $u(t) = \sum_{k=1}^{M} a_k \cos(\omega_k t + \varphi_k)$ with period $T$, its variation is defined as:

$$\mathrm{var}[u(t)] = \lim_{n\to\infty} \sum_{n=1}^{N} |u(t_{n+1}) - u(t_n)| = \lim_{n\to\infty} \sum_{n=1}^{N} \left| \sum_{k=1}^{M} a_k \cos(\omega_k t_{n+1} + \varphi_k) - \sum_{k=1}^{M} a_k \cos(\omega_k t_n + \varphi_k) \right| \tag{2.2.6}$$

$$\mathrm{var}[u(t)] \le \lim_{n\to\infty} \sum_{n=1}^{N} \sum_{k=1}^{M} |a_k| |\cos(\omega_k t_{n+1} + \varphi_k) - \cos(\omega_k t_n + \varphi_k)| \tag{2.2.7}$$

where $t_n$ are timepoints within the period. Putting $t_{n+1} = t_n + \Delta_n$, where $\Delta_n \sim O(T/N)$, using $|a_k| \le M_U/\sqrt{M}$, and after some trigonometric manipulations we obtain:

$$\mathrm{var}[u(t)] \le \lim_{n\to\infty} \sum_{n=1}^{N} \sum_{k=1}^{M} \frac{M_U}{\sqrt{M}} \omega_k |\Delta_n| \le \frac{M_U}{\sqrt{M}} \sum_{k=1}^{M} \omega_k \lim_{n\to\infty} \sum_{n=1}^{N} \Delta_n = \frac{M_U}{\sqrt{M}} \sum_{k=1}^{M} \omega_k T < \infty . \tag{2.2.8}$$

■

Without the proof we state some essential theorems from [22, 25] about the behavior of the Volterra series. Everywhere we will assume bounded inputs within the convergence radius of the Volterra series.

**Theorem 2.2.1**: **Error bound for truncated Volterra series** is [22, 25]:

$$\| V[u](t) - V^{(K)}[u](t) \|_\infty \le \sum_{k=K+1}^{\infty} \| g_k \|_\infty (\| u \|_\infty)^k \tag{2.2.9}$$

where $\|.\|_\infty$ is sup norm. ■

**Theorem 2.2.2**: **Continuity of Volterra series** [22, 25]. Let $B_r$ be the ball of radius $r$ in $L^\infty$, and suppose $r < \rho$, where $\rho$ is the radius of convergence of the Volterra series (2.22-2.23), $\rho$ = Rad V = $(\limsup_{n\to\infty} |g_n|^{1/n})^{-1}$. Then: $V$: $B_r \to B_{f(r)}$ is Lipschitz continuous, $V$: $B_\rho \to L^\infty$ is continuous, where $f(x) = \Sigma_{n=1,..,\infty} \|g_n\|_\infty x^n$ is so called gain bound function. ■

**Theorem 2.2.3**: **Steady state theorem** [22, 25]. Let $u$, $u_s$ be any signals with $\|u\|_\infty$, $\|u_s\|_\infty < \rho$ = Rad V, and suppose that $u(t) \to u_s(t)$ as $t \to \infty$. Then $V[u](t) \to V[u_s](t)$ as $t \to \infty$. ■

**Theorem 2.2.4**: **Periodic steady state theorem** [22, 25]**.** If the input $u$ is periodic with period $T$ for $t \ge 0$ then the output $V[u]$ approaches a periodic steady state, with period $T$. ■

"Intuitively, an operator has *fading memory*, if two input signals which are *close* in the *recent past*, but *not necessarily close* in the *remote past* yield present outputs which are *close*" [23]. This intuitive definition will be enough for our purposes, for the formal definition see [22-23, 25]. Then:

**Theorem 2.2.5**: **Approximating fading memory systems** [22-23]. Finite Volterra system (operator) driven by bounded inputs has fading memory; Any time-invariant non-linear system with fading memory can be approximated by a finite Volterra system in $\|.\|_\infty$ sense. ■

**Note**: The concept of fading memory is not unique. Beside fading memory in the sense of [23] on the full time axis, there are also related concepts of fading memory on the positive time axis, approximately finite memory, uniformly fading memory, or myopic maps (see [192]). These concepts differ depending on the time axis involved, assumed causality of the operators, or the properties of the input signals. The concept of fading memory of [23] is perhaps the most natural (scalar amplitude continuous signals on the full time axis), but the differences are inconsequential considering that all kinds of fading memory systems can be approximated by the finite Volterra series.

**Definition 2.2.3**: **Non-linear system class of interest.** In the BLA modeling approach the class of systems of interest is restricted to those which are limits in least-square sense of the convergent Volterra series defined in Def. 2.2.2. If otherwise not specified, the term 'non-linear system' will be used in this context. ■

**Note**: The possible convergence schemes, the conditions, and the consequences essential in the system identification are discussed in more detail in [45*], and are summarized in the Table below:

| **System class** | Properties |
|---|---|
| Wiener system | Output converges in mean square sesnse. Point-wise convergence. Discontinuities and saturation allowed (bifurcations, chaos, sub harmonics, etc. excluded). Model valid for the Gaussian signals. |
| Fading memory system | Output converges uniformly. Saturation allowed. Model valid for bounded inputs. (bound set by the user) |
| Volterra system | Output converges uniformly. Derivation model converge uniformly. Saturation allowed. Model valid for bounded inputs. (bound cannot be set by the user) |

As mentioned earlier, periodic (multi-harmonic, so called multisine) excitations will be generally used and it's time to define them exactly.

Multisines will be defined on various, not necessarily uniform frequency grids. Beside some natural conditions (asymptotic Riemann-equivalence, see Section 3.6) the obtained general results do not depend on the particular frequency grid used. Let the period be *N*, the fundamental frequency $f_0 = 1/N$, and let the set of the integer indices corresponding to the full frequency grid be $S^+_{N,0} = \left[1 \quad 2 \quad \ldots \quad N/2 - 1\right]$, $S^-_{N,0} = -S^+_{N,0}$. Then an arbitrary permissible frequency grid of exactly *M* harmonics, of an *N*-periodic multisine is defined by the subset $S^+_M \subseteq S^+_{N,0}$, $\{1\} \in S^+_M$, $S^-_M = -S^+_M$, $\left|S^+_M\right| = M$ representing $k \in S^+_M$, $f_k = k\, f_0$ frequencies.

Asymptotic computations will be done for an increasing *N* and *M*, keeping $O(N) \sim O(M)$. The normalization is done by downscaling the otherwise $O(1)$ spectral amplitudes by $1/\sqrt{2M}$.

**Definition 2.2.4**: **Normalized random (phase) multisine**. **Normalized random multisines** (called also **periodic noise** excitations) are *N*-periodic signals with randomness introduced in the amplitudes and phases, defined as:

$$u(t) = (2M)^{-1/2} \sum_{k \in S_M^- \cap S_M^+} \hat{U}(k/N) e^{j(2\pi k\, t/N + \phi_k)} = (2M)^{-1/2} \sum_{k \in S_M^- \cap S_M^+} U(k/N) e^{j2\pi k\, t/N}\,, \tag{2.2.10}$$

$$U_k = U(k/N) = \hat{U}(k/N) e^{j\phi_k}\,. \tag{2.2.11}$$

The function $\hat{U}(f_k)$ takes nonnegative real values. $\hat{U}(f_k)$ and phases $\varphi_{-k} = -\varphi_k$ are the realizations of independent (jointly and over *k*) random processes satisfying the following conditions: phases $\varphi_k$ are iid. random variables uniformly distributed on [0, 2π), $\hat{U}(f_k)$ has bounded moments of any order, and $E\left\{\hat{U}^2(f_k)\right\} = S_u(f_k)$, where $S_u(f_k)$ is the input power spectrum defined for a continuous frequency argument.

Signal (2.2.10) is called **random phase multisine**, if only its phases are random, and the spectral amplitudes take real nonnegative values $\hat{U}(k) = \sqrt{S_u(f_k)} \geq 0$. Furthermore the spectral amplitudes are uniformly bounded by $S_u(f_k) \leq (M_U)^2 < \infty$, and have at most countable number of discontinuities in the considered band. The amplitudes of the sine waves in (2.2.10) decrease as $O(M^{-1/2})$, and the power:

$$\frac{1}{N}\sum_{n=0}^{N-1} u^2(n) = \frac{1}{M}\sum_{k \in S_M^+} |U_k|^2 = \frac{1}{M}\sum_{k \in S_M^+} \left(\hat{U}_k\right)^2 \tag{2.2.12}$$

is bounded by $(M_U)^2$ as there are exactly *M* nonzero harmonics in (2.2.10). ■

**Definition 2.2.5**: **Uniform vs. colored multisines**. If $|U_k|$= const, we speak about (normalized) uniform random phase multisines, otherwise we speak about (normalized) colored multisines. ■

When normalized random multisine signals (2.2.10) are applied to the system (2.2.4), the informal decomposition (2.1.19) can be stated formally as:

**Theorem 2.2.6**: N**on-linear additive noise model and the Best Linear Approximation**. Under random phase multisine excitations (2.2.10) the output of the system (2.2.4) can be written as:

$$Y(l) = (G_1(l) + G_{B,M}(l))U(l) + Y_{S,M}(l) = G_{BLA,M}(l)U(l) + Y_{S,M}(l) \tag{2.2.13}$$

where $G_{BLA,M}(l)$ is so called **Best Linear Approximation**, and is the solution of:

$$g_{BLA} = \arg\min_{g_{BLA}} E_u(y - g_{BLA} * u)^2 \qquad G_{BLA}(l) = \arg\min_{G_{BLA}} E_U \left|Y(l) - G_{BLA}(l)U(l)\right|^2 \tag{2.2.14}$$

$G_1(l)$ is the FRF of the true underlying linear system (if any exists), and the bias or systematic distortion term $G_{B,M}(l)$ is:

$$G_{B,M}(l) = \sum_{\alpha=2}^{\infty} G_{B,M}^{2\alpha-1}(l) + O(M^{-1}) \tag{2.2.15}$$

$$G_{B,M}^{2\alpha-1}(l) = \frac{c_\alpha}{M^{\alpha-1}} \sum_{k_1,\ldots,k_{\alpha-1} \in S_M^+} G_\alpha(l, k_1, -k_1, \ldots, k_{\alpha-1}, -k_{\alpha-1}) \prod_{i=1}^{\alpha-1} |U(k_i)|^2 \tag{2.2.16}$$

$$c_\alpha = 2^{\alpha-1}(2\alpha-1)!! \tag{2.2.17}$$

and the non-linear stochastic distortion $Y_{S,M}(l)$ is zero mean, and is asymptotically uncorrelated over the frequency $l$ and with input signal. Furthermore $Y_{S,M}(l)$ is asymptotically independent from U(k), for ∀ k, l; $Y_{S,M}(l)$ is asymptotically circular complex normally distributed and mixing of arbitrary order:

$$E\{Y_{S,M}(l)\} = 0 \tag{2.2.18}$$

$$E\{Y_{S,M}(l)\overline{U}(l)\} = 0 \tag{2.2.19}$$

The even moments do not disappear, but the odd moments converge to zero ($k \neq l$):

$$E\{M\, Y_{S,M}(l)\overline{Y}_{S,M}(k)\} = \begin{cases} \sigma^2_{Y_S,M}(l) = O(M^0), & k = l \\ O(M^{-1}), & k \neq l \end{cases} \tag{2.2.20}$$

$$E\{M\, Y_{S,M}(l) Y_{S,M}(k)\} = O(M^{-1}) \qquad \text{for } l \neq k \tag{2.2.21}$$

$$E\{M^2(|Y_{S,M}(k)|^2 - \sigma^2_{Y_S,M}(k))(|Y_{S,M}(l)|^2 - \sigma^2_{Y_S,M}(l))\} = \begin{cases} O(M^{-1}), & k \neq l \\ O(M^0), & k = l \end{cases} \tag{2.2.22}$$

$$E\{M^{3/2} Y_{S,M}(l) |Y_{S,M}(k)|^2\} = O(M^{-1}) \tag{2.2.23}$$

**Proof:** Original proof appeared in [30*], simplified and generalized later in [162-163, 170] to periodic and Gauss noises.

From those proofs we emphasize only the method of computing nonzero expected values for random multisine excitations, as it is an standard tool in the BLA related proofs.

Consider that we are interested in the calculation of the nonzero expected value of $E\{Y(l)\overline{U}(l)\}$ from (2.2.4). As the randomness is only in the amplitudes and the phases of the inputs, the computed expected value takes the form of $E\{U(k_1)U(k_2)\ldots U(k_n)\}$. The expected value will be different from zero, only if the number of terms is even and all the frequencies are paired like $(k_r, k_q = -k_r)$ leading to $U(k_r) = U(k_q) = \overline{U}(k_r)$. Consequently $E\{U(k_1)\ldots U(k_n)\} = E\{|U(l_1)|^2 \ldots |U(l_{n/2})|^2\} = E\{|U(l_1)|^2\}\ldots E\{|U(l_{n/2})|^2\}$, if all frequency indices are different. As the paired terms are of $O(M^{-1})$, after summation in (2.2.4) they yield $O(1)$ order contributions. In case when more indices coincide, like e.g. $l_q = l_r$, we obtain moments $E\{|U(l_p)|^4\}$ or higher. However the imposed additional constraint $l_q = l_r$ reduces the number of possible summable frequency combinations resulting for the higher order moments in the contribution of order

$O(M^{-1})$, which asymptotically disappears. When the number of indices (terms in the expected value) is odd, one of them cannot be paired resulting in zero expected value due to the circular distribution of the phase.

In case of the random phase multisine $E\left\{|U(l_1)|^2 ... |U(l_{n/2})|^2\right\} = |U(l_1)|^2 ... |U(l_{n/2})|^2$. ■

**Notes**:

(1) All the expected values are with respect to the random phases of the input signal.

(2) Measurements on a non-linear system depend on the input signal, consequently also on the number of harmonics in (2.2.10). This justifies the index notation in (2.2.15-2.2.23).

(3) The **Best Linear Approximation** system (**BLA**) $G_{BLA,M}(l)$ in (2.2.13), the measurable linear approximation to a non-linear system, was originally called **Related Linear Dynamic System** (**RLDS**) [27*-28*], as a system strongly "related" to the linear part of a weakly non-linear system. From (2.2.15-2.2.17) one can see that it depends on the properties of the input random multisine and on the odd non-linear distortions present in the measured system. Considering that the RLDS is anyhow the best linear approximation in the mean square sense, and that the recent literature on the linear approximation to non-linear systems strongly emphasizes this point, this component has been renamed for the better correspondence with the literature.

(4) From (2.2.13) and (2.2.19) we can see that the BLA can be measured as $H_1$-FRF:

$$G_{BLA,M}(l) = \frac{E\{Y(l)\overline{U}(l)\}}{E\{|U(l)|^2\}} \tag{2.2.24}$$

i.e. as the ratio of the cross-power spectrum by the auto-power spectrum which is also the Best Linear Approximation in the least-squares sense. For the random multisines (2.2.24) simplifies to:

$$G_{BLA,M}(l) = \frac{E\{Y(l)\overline{U}(l)\}}{E\{|U(l)|^2\}} = \frac{E\{Y(l)\overline{U}(l)\}}{|U(l)|^2} = E\{\frac{Y(l)\overline{U}(l)}{|U(l)|^2}\} = E\{\frac{Y(l)}{U(l)}\} \tag{2.2.25}$$

Dividing (2.2.13) by $U(l)$ we can write it down as:

$$\frac{Y(l)}{U(l)} = G_{BLA,M}(l) + \frac{Y_{S,M}(l)}{U(l)} = G(l) = G_{BLA,M}(l) + G_{S,M}(l) \tag{2.2.26}$$

where $G_{S,M}(l)$ represents the scattering visible on the measured non-linear FRF, see e.g. Example 2.1.1.

(5) The stochastic properties of the phases theoretically could be relaxed, because what is really required is to have $E\{\exp(j\varphi)\} = 0$, which could be achieved with other distributions also.

(6) In the research the true linear part of the weakly non-linear system was denoted interchangeably as $G_0$ or $G_1$.

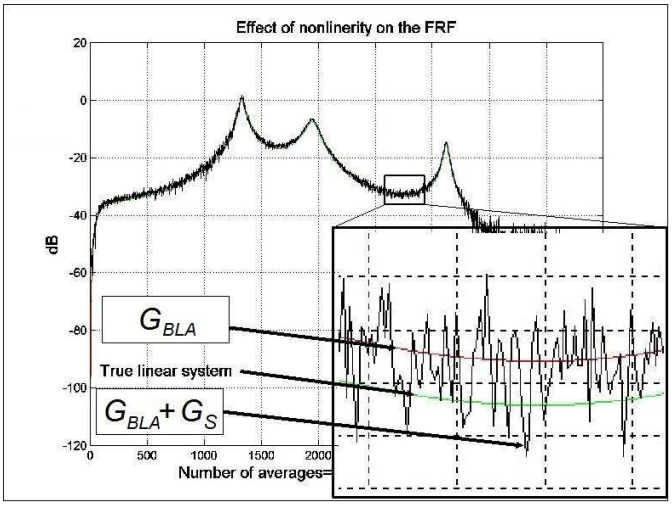


**Fig. 2.2.1:** The effect of the non-linear distortion on the FRF measurements in the light of the additive non-linear noise model. The measured FRF differs from the linear part of the system in level (systematic bias) and smoothness (stochastic non-linear noise). The measurements can be smoothed by averaging (obtaining the BLA approximation), however the results remain biased. Weakly nonlinear Wiener-Hammerstein system composed from two Chebyshev filters ($5^{th}$ order, 10 dB ripple, 0.08 relative cut-off frequency (input dynamics), and $3^{rd}$ order, 20 dB ripple, 0.035 relative cut-off frequency (output dynamics)) and a static polynomial nonlinearity (odd powers up to $11^{th}$ order) was measured with an uniform amplitude odd random phase multisine with 2185 harmonics. (The polynomial coefficients were set to the values corresponding to the nonlinear power content in the output signal being 10% of the overall output power)

## 2.3 Some properties of the BLA and the additive non-linear noise model

It is in the interest of the user leaning on the additive model (2.2.13) to know more about the properties of the systematic and the stochastic non-linear distortions. Three kinds of properties are investigated as being of interest to the practical applications:

- **asymptotic properties** as the number of harmonics in the input signal tends to ∞. Modern, memory based signal generators permit an easy design of dense multisine signals, consequently asymptotic properties are within reach in the otherwise finite measurements.

- **properties** of the systematic and stochastic distortions **as the function of frequency**.

- **robustness** of the systematic and stochastic distortions **to the free parameters of the measurement**, like e.g. the amplitude spectrum of the excitations, the overall level of the non-linearity in the measured system, or changes to the frequency grid.

### Theorem 2.3.1: BLA (2.2.13) is a bounded system, continuous in the level of the non-linear distortions and in the amplitudes of the input signal.

**Proof:** The boundedness of the BLA is the result of the boundedness of the Volterra system, of the input signal and of the BLA measurement procedure (2.2.24). For continuity in non-linear distortions let the Volterra system be decomposed as follows:

$$Y^{(a)}(l) = V[U](l) = V_1[U](l) + a\,V_2[U](l) = Y_1(l) + a\,Y_2(l) \qquad (2.3.1)$$

where $a$ is arbitrary real number which will represent the level of the non-linearities. If $V$ is well defined Volterra series, then $V_1$ and $V_2$ are also well defined (sub-series of a convergent series in (2.2.4)). By additive

decomposition (we omit the frequency and the index indicating finite numbers of harmonics):

$$Y^{(a)} = Y_1 + a\,Y_2 = G_{BLA}^{(a)}U + Y_S^{(a)} = (G_{BLA1}U + Y_{S1}) + a\,(G_{BLA2}U + Y_{S2}) \\ = (G_{BLA1} + a\,G_{BLA2})U + (Y_{S1} + a\,Y_{S2}) \tag{2.3.2}$$

Multiplying by $\overline{U}(l)$, taking expectation, and using (2.39) ($Y_s$-$U$ are uncorrelated) we can see that:

$$G_{BLA}^{(a)} = G_{BLA1} + a\,G_{BLA2}, \quad Y_S^{(a)} = Y_{S1} + a\,Y_{S2} \tag{2.3.3}$$

Let $a_k$ converge to some $a^*$, and let $|G_{BLA2}| \le M_2$ (BLA is bounded). Let choose arbitrary $\varepsilon > 0$. Then there exists $N_0$, that $|a_k - a^*| \le \varepsilon/M_2$, for k ≥ $N_0$, and:

$$|G_{BLA}^{(a_k)} - G_{BLA}^{(a^*)}| = |a_k\,G_{BLA2} - a^*\,G_{BLA2}| < |G_{BLA2}||a_k - a^*| < \frac{\varepsilon}{M_2}M_2 = \varepsilon \tag{2.3.4}$$

The proof of the continuity in input amplitudes is based on the product, ratio, expected value, superposition of continuous functions being also continuous functions.

$$\hat{U} \to U_k,\ \hat{U} \to u(t),\ u(t) \to y(t) = V[u(t)],\ y(t) \to Y_k,\ Y_k \to |Y_k|^2,\ U_k \to |U_k|^2,$$

$$U_k, Y_k \to Y_k/U_k,\ Y_k/U_k \to E\{Y_k/U_k\} = G_{BLA}. \tag{2.3.5}$$

■

**Theorem 2.3.2: Variance of the non-linear noise is continuous in the level of the non-linear distortions and in the amplitudes of the input signal.**

**Proof:** Consider the decomposition from Th. 2.3.1. For zero mean stochastic contributions we have similarly:

$$Var[Y_S^{(a)}] = E\{|Y_S^{(a)}|^2\} = E\{|Y_{S1}|^2\} + 2a\,\mathrm{Re}\,E\{Y_{S1}\overline{Y}_{S2}\} + a^2\,E\{|Y_{S2}|^2\} \tag{2.3.6}$$

$$|Var[Y_S^{(a_k)}] - Var[Y_S^{(a^*)}]| = (a_k - a^*)[2\,\mathrm{Re}\,E\{Y_{S1}\overline{Y}_{S2}\} + E\{|Y_{S2}|^2\}((a_k + a^*)] \tag{2.3.7}$$

The expected values are bounded, the term in the squared parentheses can be also bounded (say by $M_1$), and choosing $\varepsilon/M_1$ for the $a_k$ series, we have:

$$|Var[Y_S^{(a_k)}] - Var[Y_S^{(a^*)}]| < M_1|a_k - a^*| = \varepsilon \tag{2.3.8}$$

Let system $V_1$ be the whole linear part of system $V$, and system $V_2$ the whole non-linear part of system $V$, let further $a^* = 0$. Then the additive noise model (the BLA and the noise variance) tends, as the level of the non-linearity decreases, to the noiseless linear FRF measurement, i.e.:

$$G_{BLA}^{(a)} \to G_{BLA}^{(0)} = G_1, \text{ and } Var[Y_S^{(a)}] \to Var[Y_S^{(0)}] = 0 \tag{2.3.9}$$

For the input amplitudes consider that the product, ratio, expected value, superposition of continuous functions are also continuous functions.

$$\hat{U} \to U_k, \quad \hat{U} \to u(t), \quad u(t) \to y(t) = V[u(t)], \quad y(t) \to Y_k, \quad Y_k \to |Y_k|^2, \quad U_k \to |U_k|^2,$$
$$|Y_k|^2 \to E\{|Y_k|^2\}$$

finally from $|Y_k|^2, G_{BLA}, |U_k|^2 \to Var[Y_S] = E\{|Y_S|^2\} = E\{|Y_k|^2\} - |G_{BLA}|^2|U_k|^2$. (2.3.10)

■

The BLA model derived for the random multisines has been extended to the periodic noise (random amplitudes $\hat{U}(f)$) and to the Gaussian noise. For all these signal classes it has been shown that the measurement results are equivalent [162-163, 170]. The FRF of a SISO system

measured with the periodic signals and $Var[M^{-1/2}Y_{S,M}(l)]$ the non-linear noise variance tend in the least square sense (as the number of the harmonics tend to infinity, $M \to \infty$) to the best linear approximation and non-linear variance measured with the Gaussian signals, assuming that the spectral properties of the excitations are equivalent. For the finite number of the harmonics the respective results are comparable in order of $O(1/M)$ [30*, 162, 170].

From the point of view of the BLA measurements (2.2.24) the Gaussian noise causes leakage on the finite length records. Due to the small amplitudes in the amplitude spectrum of the Gaussian or periodic noise the FRF estimate (2.2.24) is noise sensitive and shows fluctuation and spikes difficult to be averaged. In comparison the amplitude spectrum of the random multisines does not fluctuate over sample functions, the denominator in (2.2.25) is constant and the FRF estimate is more stable. For this reason we advise to use random multisines in the measurements of weakly non-linear SISO systems [162, 170] (see Fig. 2.3.1).

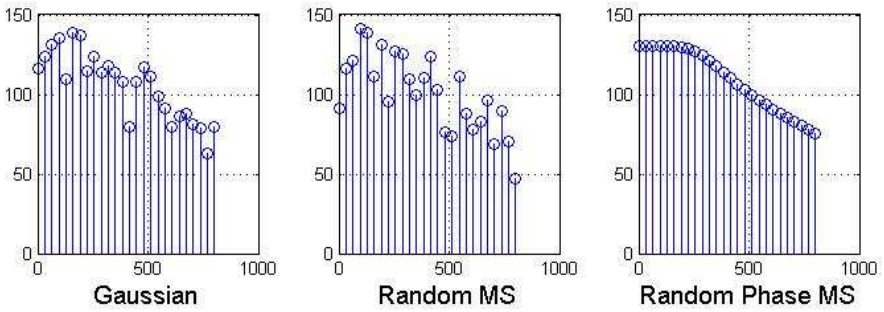


**Fig. 2.3.1:** Spectral behavior of investigated excitation signals. All signals are colored with a 3rd order Butterworth filter. For better readability the spectra are plotted at $M$=26 frequency points, which for periodic signals conincides with the number of the harmonics. Please take notice of the deterministic, nonoscillatory character of the random phase multisine spectra (contrary to the random multisine and the Gaussian noise) which ensures a better stability of the BLA measurements.

It is important to note that the above introduced characterization of the non-linear system as the biased and noisy linear FRF is valid for every convergent Volterra series. If however the non-linear effects are strong, such linear model is pointless, because its approximation errors will be high and the model won't convey any useful information (e.g. about the system dynamics). Despite the universal validity of the results their practical usage is limited to weakly non-linear (low order) non-linear systems.

**Note**: The basic theory is not suited for the close loop identification, because some essential assumptions are violated within the loop. Close loop BLA model (between the loop reference signal and the loop output signal) has been formulated later [171-172].

In measuring linear FRF BLA of nonlinear (Volterra) systems one must be ready for unexpected results. One such fortunate development is the following property:

**Theorem 2.3.3: Variance of the nonparametric BLA estimate**. Variance of the nonparametric BLA estimate $\hat{G}_{BLA}$ can be computed using the linear theory expression, i.e.:

$$\lim_{M\to\infty} M\sigma^2_{\hat{G}_{BLA}}(l) = \frac{S_{Y_S Y_S}(l)}{S_{UU}(l)} + O(N^{-1}) \qquad (2.3.11)$$

where $M$ is the number of averaged realizations of the input excitations and $N$ is the length of the excitation.

**Proof:** See Th. 1. in [208-209]. ■

**Note**: The surprise lies in the fact that although the nonlinear stochastic distortions are uncorrelated with the input (BLA is the solution to a least-squares problem), nevertheless they are mutually dependent which violates the classical hypothesis that the distortion (to be averaged out) should be independent from the input. The surprise stops here because e.g. in the parametric estimation $\hat{G}_{BLA}$ linear variance expression and the derived confidence intervals are no more applicable, a non-linear analysis is needed [205, 207-209].

## 2.4 Special case of block-models

Volterra model covers so called block models (Hammerstein, Wiener, and Wiener-Hammerstein models) widely used in the non-linear system modeling practice [91]. Block-models are built from linear dynamic and static non-linear blocks. Static non-linearity is also a special case of this model. Of theoretical interest is also the fact that general (fading-memory) non-linear systems can be approximated with relatively simple block structures composed from linear dynamic and non-linear static components [153-154, 19, 21, 24].

Special block structure of the block models (Fig. 2.4.1) is translated into special structure of the Volterra kernels and the principal question here is whether it may mean a new knowledge about the non-linear distortions. For the sake of brevity in the following we will use the notation $\langle R(f)\text{-}NL\text{-}S(f)\rangle$ to designate a Wiener-Hammerstein system with linear input dynamic $R(f)$, linear output dynamic $S(f)$, and the in-between static non-linearity $NL$.

Volterra kernels of order $\alpha$th are [198]:

(a) $$G_\alpha(k_1,k_2,\cdots,k_\alpha) = const\times\prod_{n=1}^{\alpha} R(k_n) \qquad (2.4.1)$$

(b) $$G_\alpha(k_1,k_2,\cdots,k_\alpha) = const\times S(k_1+k_2+\cdots k_\alpha) \qquad (2.4.2)$$

(c) $$G_\alpha(k_1,k_2,\cdots,k_\alpha) = const\times S(k_1+k_2+\cdots k_\alpha)\prod_{n=1}^{\alpha} R(k_n) \qquad (2.4.3)$$

(d) $$G_\alpha(k_1,k_2,\cdots,k_\alpha) = const \qquad (2.4.4)$$

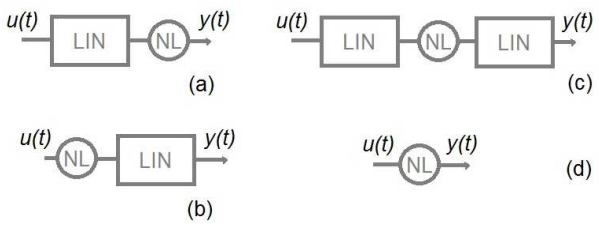


**Fig. 2.4.1** Some simpler non-linear block models: (a) Wiener $\langle R\rightarrow NL\rangle$, (b) Hammerstein $\langle NL\rightarrow S\rangle$, (c) Wiener-Hammerstein $\langle R\rightarrow NL\rightarrow S\rangle$, (d) static non-linearity $\langle NL\rangle$. From such basic blocks more involved model structures can be created, like e.g. parallel-branch Wiener, parallel-branch Wiener-Hammerstein, models with a feed-back, and also models where the static nonlinearity is substituted by a NFIR system.

The most unexpected result for the block-models is the behavior of the bias. Generally the bias is frequency dependent, but for Wiener-Hammerstein systems the bias on the measured

FRF is asymptotically (in the number of harmonics) proportional to the linear system (i.e. the relative bias is constant) (2.4.6). This proportionality called *Wiener-Hammerstein property* can be easily tested by exciting the system with a number of random multisines with a large number of harmonics and shifted power levels (Fig. 2.4.2).

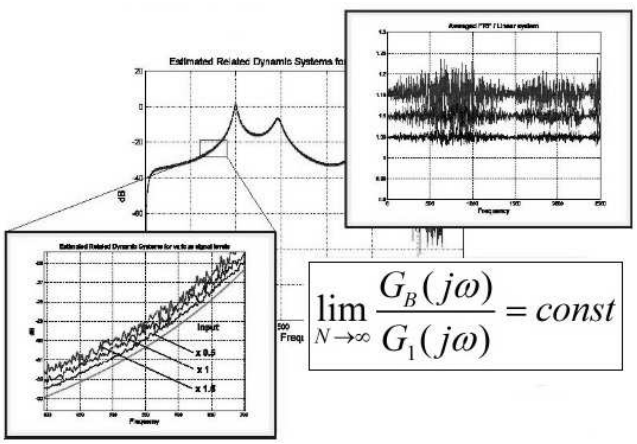


**Fig. 2.4.2**. Illustration of Wiener-Hammerstein property. The FRF characteristics of the nonlinear system from Fig. 2.2.1 (partly covered) is measured with different magnitude levels of the input signal ($\sigma$ = 0.5, 1, 1.5). A blow-up (left) shows varying levels (varying amount of the bias) of the BLA for different levels of the input, a blow-up (right) shows the measured relative bias for the same levels of the input signal.

**Theorem 2.4.1: For Wiener-Hammerstein systems the BLA is proportional to the product of the linear dynamics**. Assume that the ⟨*R*(*f*)-*NL*-*S*(*f*)⟩ Wiener-Hammerstein system is measured with normalized multisines with a large number of harmonics *M*. Then the bias is proportional to the linear part of the system, i.e. to the product *S(f) R(f)*:

$$G_{B,M}^{2\alpha-1} = C_\alpha \times S\,R + O(M^{-1}) \qquad (2.4.5)$$

Consequently: $G_{B,M} = \sum_{\alpha=1}^{\infty} G_{B,M}^{2\alpha-1} = (\sum_{\alpha=1}^{\infty} C_\alpha)\ S\,R + O(M^{-1}) = \varepsilon\, G_1 + O(M^{-1})$,

$$G_{BLA,M} = (1+\varepsilon)\ G_1 + O(M^{-1}) \qquad (2.4.6)$$

and $\varepsilon$ constant depends solely upon the excitation signal and the system dynamics.

**Proof:** Please recollect that a Volterra kernel of an odd ($2\alpha$-1) order adds to the bias as (2.2.15):

$$G_{B,M}^{2\alpha-1}(l) = \frac{2^{\alpha-1}(2\alpha-1)!!}{M^{\alpha-1}} \times \sum_{k_1 \in S_M^+} \cdots \sum_{k_{\alpha-1} \in S_M^+} G_{2\alpha-1}(l, k_1, -k_1, \cdots, k_{\alpha-1}, -k_{\alpha-1}) \prod_{n=1}^{\alpha-1} |U(k_n)|^2 + O(M^{-1}) \qquad (2.4.7)$$

Substituting (2.4.3) into (2.4.7) we get:

$$G_{B,M}^{2\alpha-1}(l) = \frac{2^{\alpha-1}(2\alpha-1)!!}{M^{\alpha-1}} S(l) R(l) \sum_{k_1 \in S_M^+} \cdots \sum_{k_{\alpha-1} \in S_M^+} \prod_{n=1}^{\alpha-1} |R(k_n)|^2 |U(k_n)|^2 + O(M^{-1})$$
$$= c_\alpha^* S(l) R(l) \prod_{n=1}^{\alpha-1} \sum_{k_n \in S_M^+} |R(k_n)|^2 |U(k_n)|^2 + O(M^{-1}) = c_\alpha S(l) R(l) + O(M^{-1}) \qquad (2.4.8)$$

where: $c_\alpha = c_\alpha^* \prod_{n=1}^{\alpha-1} \sum_{k_n \in S_M^+} |R(k_n)|^2 |U(k_n)|^2$ (2.4.9)

For normalized multisines of a large number of harmonics and for suitably smooth dynamics and input amplitude spectrum this constant can be written as:

$$c_\alpha = c_\alpha^* \prod_{n=1}^{\alpha-1} \sum_{k_n \in S_M^+} |R(k_n)|^2 |U(k_n)|^2 \approx c_\alpha^* \left(\int_0^{f_{\max}} |R(f)|^2 \hat{U}^2(f)\, df\right)^{\alpha-1} \tag{2.4.10}$$

and the (2.4.5) follows. ■

**Note**: Hammerstein and Wiener systems are special cases of the Th. 2.20 with $S(l) = 1$, or $R(l) = 1$. Static non-linearity is also a special case of Th. 2.4.1 with $S(l) = R(l) = 1$.

### Theorem 2.4.2 For Wiener-Hammerstein systems the non-linear noise variance is proportional to the output dynamics.

**Proof:** In case of an arbitrary Wiener-Hammerstein system, the variance of the zero mean stochastic component can be written as (we omit for brevity the frequency and the finite harmonics indices):

$$E\{|Y_S|^2\} = E\{\sum_{\alpha=2\ldots\infty} Y^\alpha \sum_{\beta=2\ldots\infty} \overline{Y}^\beta\} = \sum_{\alpha=2\ldots\infty} \sum_{\beta=2\ldots\infty} E\{Y^\alpha \overline{Y}^\beta\} = \sum_{\alpha=2\ldots\infty} E\{|Y^\alpha|^2\} + 2\,\mathrm{Re} \sum_{\substack{\alpha,\beta=2\ldots\infty \\ \alpha\neq\beta}} E\{Y^\alpha \overline{Y}^\beta\} \tag{2.4.11}$$

where $Y^\alpha$ are the outputs of the non-linear kernels of order $\alpha$th (2.2.2). Then:

$$E\{Y^\alpha \overline{Y}^\beta\} = E\{\sum_{k_1 \in S_M^- \cup S_M^+} \ldots \sum_{k_{\alpha-1} \in S_M^- \cup S_M^+} G^\alpha_{k_1,k_2,\ldots,k_{\alpha-1},L_k} U_{k_1} U_{k_2} \ldots U_{L_k} \times \sum_{z_1 \in S_M^- \cup S_M^+} \ldots \sum_{z_{\beta-1} \in S_M^- \cup S_M^+} \overline{G}^\beta_{z_1,z_2,\ldots,z_{\beta-1},L_z} \overline{U}_{z_1} \overline{U}_{z_2} \ldots \overline{U}_{L_z}\} \tag{2.4.12}$$

The expected value will be nonzero when both $\alpha$ and $\beta$ are odd or even. Substituting Wiener-Hammerstein kernels (2.4.3) into (2.4.12) we can observe that for:

$$E\{Y^\alpha \overline{Y}^\beta\} = a_\alpha a_\beta |S|^2 \times \sum_{k_1} \ldots \sum_{k_{\alpha-1}} \sum_{z_1} \ldots \sum_{z_{\beta-1}} R_{k_1} \ldots R_{L_k} \overline{R}_{z_1} \ldots \overline{R}_{L_z} E\{U_{k_1} \ldots U_{L_k} \overline{U}_{z_1} \ldots \overline{U}_{L_z}\} \tag{2.4.13}$$

We get nonzero expected value with frequency pairing $k_1 = z_1$, $k_2 = z_2$, etc. for $\alpha = \beta$, and pairing the remaining frequencies within themselves for $\alpha \neq \beta$. Other nonzero expected value pairings in the kernels went to the bias (BLA) terms or yields terms of lower order (i.e. $O(N^{-1-\alpha})$, $\alpha > 0$). The expression under the sums is a kind of a multiple (and thus smooth) convolution of the output linear system. Consequently:

$$E\{Y^\alpha \overline{Y}^\beta\} = \kappa_{\alpha\beta}\, a_\alpha a_\beta |S|^2 C^{\alpha\beta} |U|^2, \text{ and:} \tag{2.4.14}$$

$$E\{|Y_S|^2\} = \sum_{\alpha=2\ldots\infty} \sum_{\beta=2\ldots\infty} E\{Y^\alpha \overline{Y}^\beta\} = |S^2 \|U|^2 \sum_{\alpha=2\ldots\infty} \sum_{\beta=2\ldots\infty} \kappa_{\alpha\beta}\, a_\alpha a_\beta C_k^{\alpha\beta} = |S|^2 |U|^2\, C^* \tag{2.4.15}$$

where $C^*$ is an overall smooth function of $R$. ■

## 2.5 Non-linear bias and variance: the question of the mutual information

In the additive non-linear noise model the impact of the non-linear distortions appears in the measured FRF in two places, in its systematic, bias-like error, and in additional noise, blurring the FRF shape. Although in principle one can get rid of the stochastic component by averaging, still we have to deal with the bias, severity of which, without detailed a priori information about the non-linear system, is difficult to judge.

From the construction of the bias and the non-linear variance (2.2.15-2.2.17) we see that the same non-linear kernels (albeit being summed in different way) contribute to both. It is thus highly unlikely that the distorted FRF will be noisy, but not biased, or vice versa. On the other

hand we know already that both quantities are continuous in the level of non-linear distortions, and both grow or shrink, as the contribution of the non-linearity to the overall system increases or decreases.

The fact that we can get rid of the non-linear noise, but not of the non-linear bias, means also that the noise (its variance) is measurable, but the bias is not (only the BLA). In conclusion we can ask whether the bias could be roughly bounded by the measured variance. In the following we analyze the connection between the levels of the systematic and stochastic distortions.

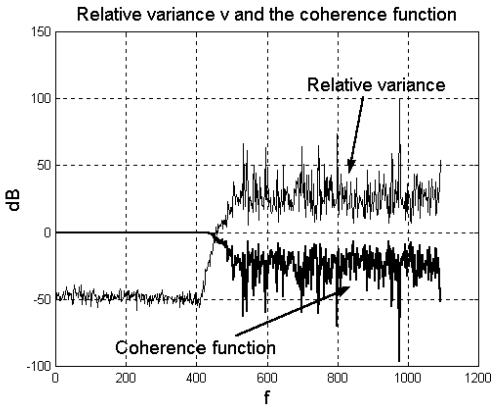


**Fig. 2.5.1** The inverse relation between the coherence function (gray) and the relative non-linear variance (dark gray) in case of a weak cubic Wiener-Hammerstein system composed from the 5th order 10 dB ripple Chebyshev high-pass input filter, a cubic static nonlinearity ($y = x + .1\ x^3$) and the 9th order 1 dB ripple low-pass Chebyshev output filter, measured with an odd random phase multisine of 546 harmonic components. N=10 measurements were averaged.

**Theorem 2.5.1 The additive BLA model and the coherence function**. The coherence of a non-linear system, excited with the uniform multisines, can be naturally expressed in terms of the BLA and the non-linear variance $V_S(l) = Var[G_S(l)] = E\{|Y_S(l)|^2\}/|U|(l)^2$ of the $G_S$ (2.2.26) (the finite harmonic index and the frequency argument are omitted for clarity):

$$\gamma^2 = \frac{|G_{BLA}|^2}{|G_{BLA}|^2 + V_S} = (1 + \frac{V_S}{|G_{BLA}|^2})^{-1} \tag{2.5.1}$$

**Proof:** Using the definitions of the coherence function [8], the BLA, and the properties of the multisine excitation, and noticing that $|U(l)| =$ constant, we obtain:

$$\gamma^2 = \frac{|E\{Y\bar{U}\}|^2}{E\{|Y|^2\}\,E\{|U|^2\}} = \frac{|U|^4|\,E\{Y/U\}|^2}{E\{|Y|^2\}\,|U|^2} = \frac{|U|^2|G_{BLA}|^2}{E\{|Y|^2\}} \tag{2.5.2}$$

with the expected value calculated over different realizations of the random multisines. The denominator can be written as:

$$\begin{aligned} E\{|Y|^2\} &= E\{|G_{BLA}U + Y_S|^2\} = |G_{BLA}|^2|U|^2 + 2\,\mathrm{Re}\,\bar{G}_{BLA}E\{Y_S\bar{U}\} + E\{|Y_S|^2\} \\ &= |G_{BLA}|^2|U|^2 + E\{|Y_S/U|^2\}\,|U|^2 = |G_{BLA}|^2|U|^2 + \mathrm{var}[G_S]\,|U|^2 = |G_{BLA}|^2|U|^2 + V_S\,|U|^2 \end{aligned} \tag{2.5.3}$$

due to the lack of correlation of the stochastic component and the input signal (2.2.19). Substituting (2.5.3) into (2.5.2), and dividing by $|U|^2$, yields (2.5.1). ■

**Theorem 2.5.2 Coherence for a Wiener-Hammerstein system.** For a Wiener-Hammerstein system $\langle R(f)\text{-NL-}S(f)\rangle$ driven with a uniform random multisines, the coherence function depends strongly on its input dynamics.

$$\gamma^2 \approx \frac{|R|^2}{|R|^2 + const} \tag{2.5.4}$$

**Proof:** In case of an arbitrary Wiener-Hammerstein system, its Best Linear Approximation and the variance of the stochastic component can be written as:

$$E\{|Y_S|^2\} = \sum_{\alpha=2\ldots\infty} \sum_{\beta=2\ldots\infty} E\{Y^{\alpha}\overline{Y}^{\beta}\} = |S^2|U|^2 \sum_{\alpha=2\ldots\infty} \sum_{\beta=2\ldots\infty} \kappa_{\alpha\beta}\ a_{\alpha}\ a_{\beta}\ C_k^{\alpha\beta} = |S|^2|U|^2\ C^* \tag{2.5.5}$$

where $C^*$ is an overall smooth function of $R$ (c.f. Th. 2.4.2 and (2.4.15)), and:

$$G_{BLA} = \sum_{\alpha} \nu_{\alpha} S\ R = \rho\ S\ R \tag{2.5.6}$$

Comparing (2.5.5-2.5.6) with (2.5.1) we see that indeed a behavior indicated in (2.5.4) is to be expected. ■

**Example 2.5.1: 3rd order Wiener-Hammerstein system.** For the Wiener-Hammerstein system possessing only linear and cubic terms ($NL = a_1 x + a_3 x^3$), $V_S$ at the frequency $k$ can be calculated as follows:

$$\begin{aligned} E\{|Y_S|^2\} = V_S|U|^2 &= E\{\sum_{k_1\in S_M^-\cup S_M^+} \sum_{k_2\in S_M^-\cup S_M^+} G^3_{k_1,k_2,L_k} U_{k_1} U_{k_2} U_{L_k} \sum_{z_1\in S_M^-\cup S_M^+} \sum_{z_2\in S_M^-\cup S_M^+} \overline{G}^3_{z_1,z_2,L_z} \overline{U}_{z_1} \overline{U}_{z_2} \overline{U}_{L_z}\} \\ &= \sum\sum |G^3_{k_1,k_2,L_k}|^2 |U_{k_1}|^2 |U_{k_2}|^2 |U_{L_k}|^2 = 3! a_3^2 |S_k|^2 \sum\sum |R_l|^2 |R_n|^2 |R_{k-l-n}|^2 |U_l|^2 |U_n|^2 |U_{k-l-n}|^2 \\ &= 3! a_3^2 2^{-2} |S_k|^2 \frac{1}{M^2} \sum_{l\in S_M^+} \sum_{n\in S_M^+} |R_l|^2 |R_n|^2 |R_{k-l-n}|^2\ |U_k|^2 = \frac{3}{2} a_3^2 |S_k|^2 C_k\ |U_k|^2 \end{aligned} \tag{2.5.7}$$

where $k_1 = z_1$, etc. frequency pairings yield nonzero expected value, $L_k = k-k_1-k_2$ and $L_z = k-z_1-z_2$ .The expression is simplified further by the subsequent substitution of the Wiener-Hammerstein kernels and normalized input amplitudes. Finally for a large number of multisine components the double sum can be treated as the approximation of the convolution $C_k$. Let us introduce the coefficient measuring the level of non-linearity as: $\chi = \sqrt{3/2}\,\varepsilon/(1+3\varepsilon r_1)$, with $\varepsilon = a_3/a_1$. Accordingly to (2.5.6) the Best Linear Approximation FRF can be written as: $G_{BLA} = (a_1 + 3a_3 r_1)\ S\ R$ , with $r_1 = M^{-1}\sum_{k\in S_M^+} |R_k|^2$ .

The coherence function can be now computed accordingly to (2.5.1), using (2.5.7). The coherence function shows strong dependence upon the input dynamics:

$$\gamma^2 = (1 + \frac{V_S}{|G_{BLA}|^2|U|^2})^{-1} = \frac{1}{1+\chi^2 C/|R|^2} = \frac{|R|^2}{|R|^2+\chi^2 C} \approx \frac{|R|^2}{|R|^2 + const} \tag{2.5.8}$$

where $\chi$ is proportional to the level of non-linearity, $C$ is smoothly behaving, and the constant is of order $\varepsilon^2$. Consequently the coherence function is close to 1 in the pass-band of $R$ and then follows the shape of $|R|^2$ when it drops, see Fig. 2.5.2. That way it can serve as an indicator of how the dynamics of the overall system are distributed between its input and output.

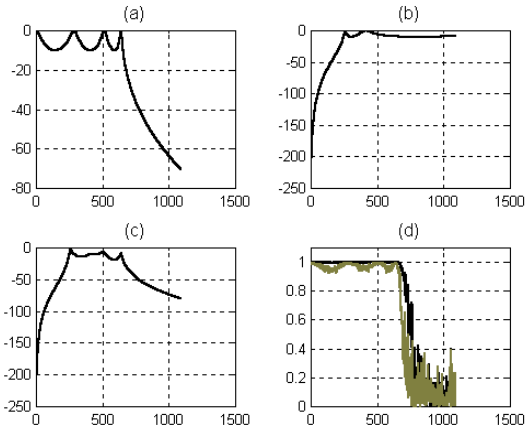


**Fig. 2.5.2**. Coherence function (d, black) of a cubic Wiener-Hammerstein system (Fig. 2.5.1), with input dynamics (a), output dynamics (b), and the overall dynamics (c). The coherence is dependent solely upon the input dynamics. For the comparison the coherence for the case of a general non-linear system (up to 9th order) is also shown (d, gray). $K$=10 measurements were averaged.

Now we switch over to a more involved situation. Even if no a priori information is available, the measurement still yields the Best (albeit distorted) Linear Approximation $G_{BLA}$ of $G_1$ with the observed level of the non-linear noise $Var[Y_S]$. An interesting question is whether this directly measurable quantity can be used to estimate the level of the systematic non-linear distortions.

We will investigate the worst-case situation, i.e. given the level of the measured non-linear noise, what order of a non-linearity may be assumed to yield systematic error bounds necessarily majoring the actual systematic error level. We show that for static monomial non-linearity the measurable non-linear variance contains enough information to compute the bounds on the FRF bias, even if the order of the non-linearity is not known. The result is based upon the fact that different powers contribute in different way to the bias and to the stochastic terms. We will compare the ratio of the non-linear variance to the bias as the relative variance:

$$v_B(l) = \frac{Var[Y_S(l)]}{|G_B(l)|^2 |U(l)|^2} = \frac{E\{|Y_S(l)|^2\}}{|G_B(l)|^2 |U(l)|^2} \qquad (2.5.9)$$

We will see that for a measured level of the variance the cubic system yields the largest bias, i.e. the cubic power is the safest (the worst-case) assumption with respect to the unknown bias.

**Theorem 2.5.3 For the pure $\alpha$th odd static monomial non-linear system the relative variance $v_B$ yields minimum for the cubic system $\alpha = 3$.**

**Proof:** Taking into account that the multisine excitations are asymptotically normally distributed, we will show the point using Gaussian signals, which generally yields O(1/$M$) ($M$ is the number of harmonics) approximation to the behavior of the random multisine. The exact derivation, counting harmonics, with a finite $M$ harmonics random multisine can be found in [6*, 9*].

Let the system be: $y(t) = c_\alpha u^\alpha(t)$, with $\alpha$ odd, excited with Gaussian $u(t)$ with unit $\sigma$. With no linear term: $G_{BLA} = G_B = E\{yu\}/E\{u^2\} = c_\alpha E\{u^{\alpha+1}\}/E\{u^2\} = c_\alpha\, \alpha!!\, \sigma^{\alpha+1}/\sigma^2 = c_\alpha\, \alpha!!\, \sigma^{\alpha-1}$. The variance of the stochastic component is:

$$E\{(y - G_{BLA}\, u)^2\} = E\{y^2\} - (G_{BLA})^2 E\{u^2\} = (c_\alpha)^2 E\{u^{2\alpha}\} - (G_{BLA})^2 E\{u^2\}$$
$$= (c_\alpha)^2 (2\alpha-1)!!\, \sigma^{2\alpha} - (c_\alpha\, \alpha!!\, \sigma^{\alpha-1})^2\, \sigma^2 = (c_\alpha)^2\, \sigma^{2\alpha} [(2\alpha-1)!! - (\alpha!!)^2] \qquad (2.5.10)$$

The relative variance $v_B$ can be derived now as:

$(c_\alpha)^2\ \sigma^{2\alpha}[(2\alpha-1)!! - (\alpha!!)^2] / (c_\alpha\,\alpha!!\,\sigma^{\alpha-1})^2\,\sigma^2 = [(2\alpha-1)!! - (\alpha!!)^2] / (\alpha!!)^2,$ (2.5.11)

which is an increasing function of the odd $\alpha$, see Fig.2.5.3. ■

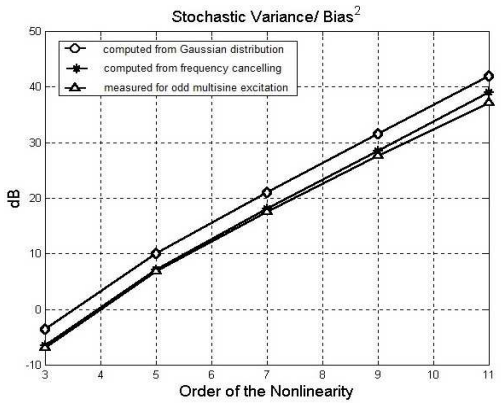


**Fig. 2.5.3** Behavior of the relative variance of a static monomial non-linear system, computed exactly for the random multisine with finite number of the harmonics, computed approximately via the Gaussian approximation, and measured with odd random multisines (2048 harmonic components, N=100 averages).

**Example 2.5.2 Cubic non-linearity is indeed the roughest**.

The heuristic explanation to the Th. 2.5.3 could be that the cubic non-linearity yields the least amount of frequency summations leading to the scattering (see [6*]). The higher the non-linearity, the more scattering, and due to the randomization through the input signal, the higher non-linear variance. Consider this phenomenon on a weakly non-linear system defined as: $y(t) = u(t) + \varepsilon u^\alpha(t)$, with $\alpha$ odd, excited with Gaussian $u(t)$ with some $\sigma$.

Now: $G_{BLA} = E\{yu\}/E\{u^2\} = E\{u^2 + \varepsilon u^{\alpha+1}\}/E\{u^2\} = 1+\varepsilon\,\alpha!!\,\sigma^{\alpha-1}$. The variance of the stochastic component is:

$E\{(y - G_{BLA}\,u)^2\} = E\{y^2\} - (G_{BLA})^2\,E\{u^2\} = E\{u^2+2\varepsilon u^{\alpha+1}+\varepsilon^2 u^{2\alpha}\} - (G_{BLA})^2\,E\{u^2\}$

$= \sigma^2 + 2\varepsilon\,\alpha!!\,\sigma^{\alpha+1} + \varepsilon^2(2\alpha-1)!!\,\sigma^{2\alpha}\} - (1+\varepsilon\,\alpha!!\,\sigma^{\alpha-1})^2\,\sigma^2 = \varepsilon^2\,\sigma^{2\alpha}[(2\alpha-1)!! - (\alpha!!)^2]$ (2.5.12)

The relative variance $v_B$ *is*: $\varepsilon^2\,\sigma^{2\alpha}[(2\alpha-1)!! - (\alpha!!)^2] / (1+\varepsilon\,\alpha!!\,\sigma^{\alpha-1})^2\,\sigma^2$.

Let us investigate its behaviour not only as a function of the order of the non-linearity, but also of the level of non-linearity $\varepsilon$. In Fig. 2.5.3 we can see that the cubic non-linearity is always the worst in the sense of the Th. 2.5.3.

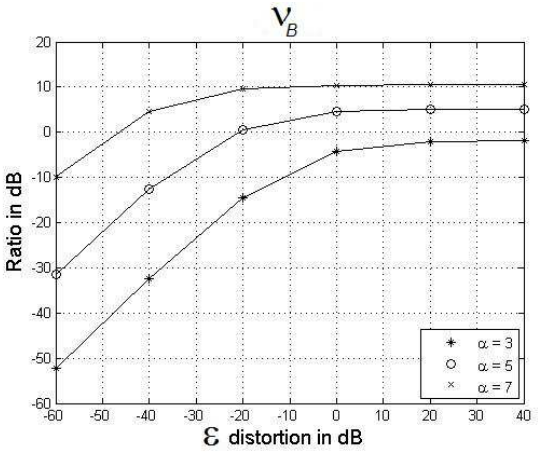


**Fig. 2.5.4** Ratio of the non-linear std to the BLA of the weakly non-linear system with static non-linear distortion of order α (i.e. $y = u + \varepsilon u^\alpha$) as a function of the level ε of the non-linear distortions. For large distortions (right side of the figure) the $G_{BLA} \approx G_B$.

The assumption of the pure αth order non-linearity is not realistic. Even if one of the powers in the non-linear system is dominant, the effect of the remaining powers should be tested,

especially when considering the worst-case without any a priori information. Although the general case of a dynamic non-linearity is too difficult to handle, some insight can be gained into the behavior of the static polynomial non-linearity, where it also turns out that in majority of practical cases the cubic order assumption can still serve as the worst case to judge the amount of the bias based upon the variance measurements.

An arbitrary static non-linear system poses a problem because the value of the relative variance will depend upon the unknown order of the system and the values of its coefficients. The way out is the worst-case derivation of the ratio of the variance to the bias (i.e. assuming the worst-case polynomial coefficients for a given level of the variance). Although $v_B$ cannot be measured, nor derived directly, $v_{BLA}$ a variance relative to the BLA approximation can be measured instead, serving as a useful empirical constraint in looking for the worst-case bias. The derivation shows that the situation does not change – the assumption of the cubic system is still the most conservative, yielding the largest possible bias for a given measured variance.

$$v_{BLA} = E\{|Y - G_{BLA}U|^2\}/|G_{BLA}U|^2 \qquad (2.5.13)$$

**Theorem 2.5.4 Bias bounds on a static polynomial non-linearity.** For an arbitrary polynomial static non-linear system, fulfilling the conditions of the proof, the relative variance $v_{BLA}$ yields minimum for the (lowest order) cubic system.

**Proof:** in Appendix A.1 ■

In the view of the Th. 2.14 the unknown $G_1$ linear part of the system can now be bounded under the worst-case assumption by the measured $G_{BLA}$: $G_{BLA}(1-\kappa) < G_1 < G_{BLA}(1+\kappa)$, with the bounding term $\kappa$ computed under the worst-case cubic assumption as: $\kappa = \sqrt{2}\sqrt{v_{BLA}}$, see (A.1.21-A.1.22).

**Note:** In this case the obtained bounds are not exactly the bias, because we defined them around the measured Best Linear Approximation, and not around the true linear system.

**Note:** In Appendix A.1 we make the assumption about the invertibility of the matrix (A.1.5). Such assumption can be violated in practice for certain combinations of the polynomial coefficients (typically for polynomial coefficients with alternating sign, leading to the mutual cancellation of the power terms in the bias or the non-linear variance expressions), however such non-linear systems are infrequent in practice. In practice the applicability of the Th. 2.5.4 requires the verification of the condition, based upon the measurements of the relative variance and the estimate of the highest non-linear order in the system.

The investigation of the systematic error bounds can be extended heuristically to the case of Wiener-Hammerstein systems, taking into account the constant relative bias property (2.4.5) and the frequency dependence of the non-linear noise (2.4.15) [6*, 9*].

It is important to note that due to the constant relative bias the smallest value of the relative variance measured somewhere in the band can be used to bound the bias in the whole band as well. Consequently the measurement should be done in the pass-band of the input linear system. This band can be easily judged from the measured $v_{BLA}$ itself.

This yields the following measurement strategy. After exciting the system with a broadband multisine excitation with a high number of harmonics *M*, the FRF should be measured to provide an insight into its dynamic behaviour. Next a rough estimate of the BLA system and

the non-linear noise variance should be computed by averaging the measured FRF over the neighbouring frequencies. Then the frequency band should be chosen where the amplitude of the measured relative variance is the smallest and the average ratio of the variance to the FRF (A.1.1) should be measured and used to compute the bounds (Fig. 2.5.5). This value can be used to bound the measured (averaged) BLA system acc. to (A.1.21).

The research of the interplay between the systematic non-measurable and stochastic measurable non-linear error components continued and was aimed next at the non-linear system without a true linear component [47*-48*]. At a particular excitation a polynomial system can expose a prominent linear behavior even if the linear term is in itself missing. In this case it is impossible to bound the bias from the variance and a new measure has to be designed for the characterization of the non-linear bias. The new measure (2.5.14) was analytically designed for the static polynomial systems.

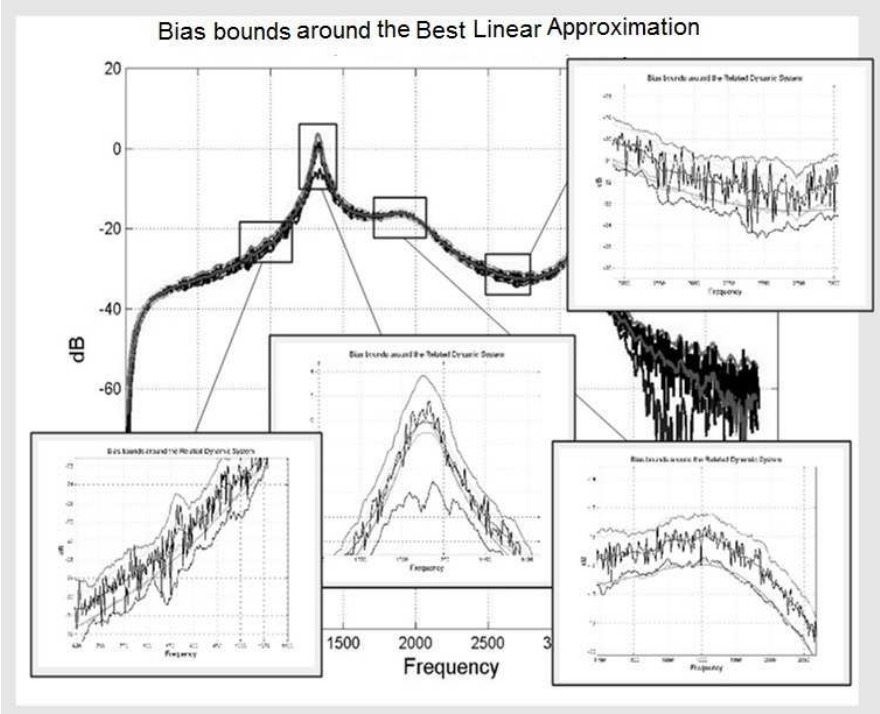


**Fig. 2.5.5** Heuristic systematic error bounds on the measured FRF of a weakly non-linear system (see Fig. 2.2.1), based on the measurements of the relative non-linear variance. Within the bounds one can see the measured 'noisy' FRF, its expected (averaged) value, i.e. the BLA, and below it the true linear component of the system. Odd uniform random phase multisine was used with 2185 harmonics.

The measure is based on a reference measurement with an excitation signal of a reference level $P_0$ ($\sigma = 1$), and on a new measurement at a different (lesser) level $P_1$ ($\tilde{\sigma} < \sigma$). Different levels yield different bias levels and $R(P_1)$ the ratio of the bias error (difference in bias) to the power of the reference stochastic contribution can be express in terms of moments ($\mu_p$) and polynomial coefficients ($a_p$):

$$R(P_1) = \frac{\tilde{\sigma}^2 \left( \sum_{p=2}^{n} a_p \mu_{p+1} (1 - \tilde{\sigma}^{p-1}) \right)^2}{\sum_{p=2}^{n} \sum_{q=2}^{n} a_p \, a_q (\mu_{p+q} - \mu_{p+1} \mu_{q+1})} \qquad (2.5.14)$$

and can be effectively bounded, yielding a hint of how the bias error evolves.

The worst-case behavior of the $R(P_1)$ for a particular power level $P_1$, non-linear order $n$, and a choice of the excitation signal was computed via numerical optimization and is visualized in contour plots in [48*]. These results were extended to the (generalized) Wiener-systems, but further steps toward more general non-linear systems were deemed practically infeasible.

# 3. Multisine excitations for SISO measurements

The used excitations – possibly colored normalized random multisines – are periodic signals. The reason of using periodic signals is that they generate no leakage in the FRF measurements. For the non-periodic noise input, the leakage error at the output of the linear system is of order $O(N^{-1/2})$, where $N$ is the number of the data points, and despite averaging it can scale up to a considerable bias error around sharp resonances in the measured FRF [125, 162]. Another reason is the easy control over their harmonic content, which can be impaired at most by the limits of the generating equipment (i.e. in high frequency microwave measurements). With time it turned out also that the primary tools of the multisine design are not the spectral amplitudes, but rather the phases and the frequency grid.

## 3.1 Multisine design – free parameters

Multisine signals are characterized by their **spectral amplitude**, **phase**, and **frequency grid** vectors, and altering them (with a little care if frequencies are also manipulated) won't affect the key asset – the periodicity of the signal. It means nonetheless that we are quite free in shaping the behavior of the multisines, which offers an enormous flexibility in the experiment design.

The normalized spectral amplitudes can be chosen to be uniform or colored, if more power is required in particular frequency bands to secure the persistence of the excitations. The coloring of the spectral amplitudes has no effect on the majority of the theoretical results presented here, albeit some of the results are much simpler for the uniform multisines. It may also happen that the coloring of the spectral amplitudes will influence the convergence rate of the iterative algorithms manipulating the phases.

Of special interest is the manipulation of the frequency grid. It can be full, i.e. every harmonics is present in the signal within the frequency band, or can be "hole-ridden", with particular harmonics purposefully left out. We speak in such case about the **excitation frequencies** (present in the signal) and **test frequencies** (omitted in the signal). The point is that (1) the non-linear scattering will generate output components also at the test frequencies, providing an "insight" into the character of the non-linearity; (2) a "hole-ridden" frequency grid may lead to less scattering on the excitation lines, distorting the linear FRF less. Both effects have been successfully explored.

**Example 3.1 "Seeing the non-linearity" at the test frequencies**.

A system containing both the 2nd and 3rd order non-linearities is measured with multisines containing harmonics up the 13th harmonic. Instead of using all harmonics (full) we may resort to the purely odd frequency grid, or retaining only every second of the odd harmonics, the so-called odd-odd grid. The 2nd and the 3rd order non-linearities will introduce frequency mixing and scattering, generating output power at the sums of two, or three excitation input frequencies. It is instructive to see how the choice of the grid makes it possible firstly to test for the presence, secondly to get rid of the 2nd order distortions (odd), or even 3rd order distortions (odd-odd).

In the Table below we see for both kinds of signals (odd, odd-odd) which harmonics are excited (•), and how many mixed spectral components (frequency combinations) appear at every excitation and test frequency (within the excited frequency band). Cubic non-linearity still can place non-linear power at the excitation frequencies,

distorting the measured linear FRF. It is not possible in the second odd-odd case, where the cubic power falls on test frequencies, and the linear FRF comes undistorted. Of course non-linearity of the 5th order would foul the design, or we would have to design even sparser odd harmonic signal.

| | Only odd harmonics | | | | Only odd-odd harmonics | | | |
|---|---|---|---|---|---|---|---|---|
| Full | Excited lines | Lin | Square | Cubic | Excited lines | Lin | Square | Cubic |
| 1 | • | 1 | | | • | 1 | | |
| 2 | | | 1 | | | | 1 | |
| 3 | • | 1 | | 1 | | | | 1 |
| 4 | | | 2 | | | | | |
| 5 | • | 1 | | 3 | • | 1 | | |
| 6 | | | 3 | | | | 2 | |
| 7 | • | 1 | | 6 | | | | 3 |
| 8 | | | 4 | | | | | |
| 9 | • | 1 | | 10 | • | 1 | | |
| 10 | | | 5 | | | | 3 | |
| 11 | • | 1 | | 15 | | | | 6 |
| 12 | | | 6 | | | | | |
| 13 | • | 1 | | 21 | • | 1 | | |

Theoretically phases shouldn't be considered a design parameter, as random phases of prescribed stochastic behavior are required in the very definition of the input signal. Besides, the input phases do not appear in spectral measurements and do not appear in the measured linear FRF. However phases do influence the time-domain behavior of the signal and with keeping the same spectral content phases can make it more advantageous in a particular situation, e.g. with respect to the measurement noise. We will see later how the algorithms developed to manipulate phases in the linear measurements can be effectively used when measuring non-linear distortions.

Multisine design, until recently, was addressed in the research literature as a tool:

- to provide persistency by shaping power injected into the system in different frequency bands (spectral amplitudes);

- to improve the SNR in the linear measurements by shaping crest-factor of the multisine signal (phases);

- to provide control over how much non-linearity is hit by the input power by shaping the amplitude spectrum of the multisine signal (phases);

- to minimize non-linear distortions of particular kind on the measured linear FRF (frequency grid);

- to qualify arbitrary non-linear distortions on the measured linear FRF (frequency grid, phases).

In the Section 3.3 some problems will be addressed in turn.

## 3.2 Multisines – asymptotic properties

In case of the finite number of harmonics, defined on different grids, with different amplitude and phase spectra, properties of the multisines are hopeless to analyze and compare effectively. Nonetheless the question of properties is important because otherwise measured results derived for different kinds of input signals wouldn't be comparable. The answer can be given in terms of the asymptotic properties, i.e. when the number of harmonics in the signal

grows beyond limit. We have seen already similar results in the additive non-linear noise model (Th. 2.2.6). With the technology allowing for an easy design of signals with $O(M) \sim 10^3$-$10^4$ we are close to the asymptotic properties already in the majority of practical measurement situations. The key asymptotic property is the Gaussianity, or the recently revived property of being separable [148, 59-61]. For the multisines defined earlier we have:

**Theorem 3.2.1: Random multisines are in the limit normally distributed and separable signals.**

**Proof: in Appendix A.2** ■

**Corollary 3.2.1**: **The uniform random multisine (i.e. $U_k = U$ = const) is separable.**

**Proof**: See [148], [59-62], but in the relation to the proof of the Th. 3.2.1 in App. 2, please note that that (A.2.22) for the uniform multisines yields:

$$D(\xi,\tau) = \frac{1}{2}U^3 J_1(U\xi) J_0^{M-1}(U\xi) \sum_{l\in S_M^+} \sum_{k\in S_M^+} (\cos(\omega_k\tau) - \cos(\omega_l\tau)) =$$
$$\frac{1}{2}U^3 J_1(U\xi) J_0^{M-1}(U\xi) (\sum_{l\in S_M^+} \sum_{k\in S_M^+} \cos(\omega_k\tau) - \sum_{l\in S_M^+} \sum_{k\in S_M^+} \cos(\omega_l\tau)) \qquad (3.2.1)$$
$$\frac{1}{2}U^3 J_1(U\xi) J_0^{M-1}(U\xi) M (\sum_{k\in S_M^+} \cos(\omega_k\tau) - \sum_{k\in S_M^+} \cos(\omega_l\tau)) = 0$$

consequently (A.2.6) is 0, and the signal by definition is separable. ■

**Note**: The concrete finite frequency grid does not play any role in the derivation, i.e. the results are valid for multisines defined on arbitrary grid.

## 3.3 Frequency grid families of multisine excitations

Shaping the spectral content of the signal is a task not particular to the multisines. On the other hand manipulating the frequencies and the phases is intimately related to the multisine structure. Modern instrumentation with PC-based computing power makes it easy to develop multisines with arbitrary frequency grids and phase properties. In the following we give a short review of these attempts, stating the design purpose and the effects they produce.

**Note**: Although making the frequency grid sparser serves always particular measurement purpose, in a finite frequency band of interest it easily leads to a contradictory situation, when making sparse grid dense enough pushes the first harmonic toward the zero frequency introducing thus extremely slow signals, which require considerable measurement time to settle the transients and to acquire the required amount of data.

## Regular grid based multisines

### A. **Full grid multisines**

| | |
|---|---|
| $f_0$ * [1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 … ] | $f_0 * k$, $k \in N$, natural numbers |

Full grid multisines contain all of the odd and even harmonics and have the best frequency resolution. They are signals of choice when no non-linear distortion is present. In the presence

of the non-linear distortions, measuring with the full grid multisines introduces non-linear bias and variance scattered all over the excited frequencies.

B. **Prime multisines**

| | |
|---|---|
| $f_0$ * [1 3 5 7 11 13 17 19 … ] | $f_0 * p$, $p \in P$, prime numbers |

Prime grid was used to get rid of the influence of the even non-linearities (odd frequency lines are not excited by the even non-linearities, if there are no even harmonics in the signal). The primary draw-back of the prime multisines was their sparse behavior for higher frequencies, consequently problems with an even frequency resolution and with the measurement time. [177, 66, 73]

C. **Odd-multisines**

| | |
|---|---|
| $f_0$ * [1 3 5 7 9 11 13 15 17 19 21 23 …] | $f_0 * (2k\text{-}1)$, $k \in N$, natural odd numbers |

Leaving out even harmonics serves more ends at the same time. Considering that the measured FRF is distorted by a smooth bias, the frequency resolution will still be sufficient, if we design sparser multisines with a number of frequencies left out. These left-out (test) frequencies can be used to estimate the level of the distorting non-linear noise, which can be used then to compensate FRF measurements at the remaining excited frequencies. If we leave out all the even harmonics from the excitation signal, then (a) the systematic distortion on the excitation lines will be smaller (less frequency combinations of nonzero expected value); (b) the non-linear noise caused by the even order non-linearity will be placed only on even (test) frequency lines; (c) even test lines can be used to detect whether even non-linearity is present in the system, and how strong it is, see Example 3.1. [68, 71, 73]

D. **Odd-odd multisines**

| | |
|---|---|
| $f_0$ * [1 5 9 13 17 21 25 …] | $f_0 * (4k\text{-}3)$, $k \in N$, every second natural odd number |

Accepting further limitation in the frequency resolution, even more opportunities open in handling the non-linear distortions. If we leave out e.g. every second line from the odd excitation lines, then beside every advantage listed above, the 3rd order non-linearity won't affect the measurements at the excitation lines, and will place its influence solely at the left-out odd test frequency lines (see Example 3.1.1). It means that if the non-linearly distorted system possesses non-linearities of only 2nd and 3rd order, it can be measured (albeit with sparser resolution) without non-linear errors at all, because the non-linearity will place its influence solely at the test frequency lines. In addition the noise variance measured at the test frequency lines can be used to detect and to judge the order and the severity of the non-linear distortions. The odd-odd multisine is not a cure for every problem, because already the non-linearity of 5th order will place its influence at the excitation lines, distorting the measurement. It will also mix up with the effects of the cubic non-linearity at the odd test lines, masking the problem. The effect of the 5th order non-linearity could be handled by an odd multisine with every second and third odd harmonic removed, such a signal would be however too sparse for practical applications, see Example 3.1.1. [66, 68, 72-73]

E. **Special-odd multisines**

| | |
|---|---|
| $f_0$ * [1 3 9 11 17 19 25 27…] | $f_0$* (8$k$ - 7) ∪ $f_0$* (8$k$ - 5), $k \in N$, every fourth natural odd number |

In case when the non-linearity is higher than the 3rd order (i.e. when the odd-odd multisines do not serve their purpose), a variation of the odd-odd grid has been tried, called special-odd multisines, where the odd excitation and the test frequencies are not coming in turn, but are grouped more closely together. The aim was to extrapolate the non-linear variance of the excitation lines from its measurements on the neighboring test lines. Published simulations have shown that special-odd multisines are better in this respect, than the odd-odd multisines. [49*-50*, 246-247]

F. **Log-tone multisines**

| | |
|---|---|
| an example: $f_0$ * [1 3 5 11 21 51 101 … ] | log($f_0$* $k$), $k \in N$, approximately uniformly spaced |

Log-tone multisines place the excitation at frequencies providing uniform resolution, when the measured FRF will be displayed on the logarithmic axis (Bode plot). The advantages of the log-tone multisines are however limited. They become sparser for higher frequencies, are difficult to suppress with Crest-Factor minimizing algorithms, increase the measurement time, and anyhow a dense enough odd multisine is easier to handle [81-82].

G. **No-interharmonic distortion (NID) multisines**

| | |
|---|---|
| an example: $f_0$ * [1 5 13 29 49 81 119 141 207 263 359 459 … ] | no close formula |

Interharmonic distortions or so called Type II contributions are non-linear stochastic contributions in the nomenclature of C. Evans. His idea was to devise multisine excitations with specially developed spacing between the harmonics to warrant that no Type II contribution will be generated at the excitation frequencies for a given order of the non-linearity. The example shows the grid designed for the cubic non-linearity. Beside the advantages warranted by the special design, the primary deficiency is the dependence on the assumed order of the non-linearity, a log-tone like sparseness for the higher frequencies, and the iterative, search-based procedure to procure the required exact harmonic numbers. [67, 69-70, 72]

## Random grid based multisines

H. **Random grid multisines**

| | |
|---|---|
| an example: $f_0$ * [1 3 5 9 13 15 17 19 23 …] | no close formula |

Random grid multisines are derivative of the special-odd multisines, where it was observed that the value of the non-linear noise variance computed from the observations at the test frequency lines does not entirely agree with those on the excitation lines, consequently only a rough extrapolation could be done.

The strict regularity of the grid was thought to be the culprit, and the idea was to destroy it in a random manner. The odd-multisines grid had been divided into consecutive blocks, where one frequency (but not the first in the block) was chosen at random to be left out (i.e. to serve as a test line). Extensive simulations show that the idea works and that the value of the non-linear distortions estimated at the test frequencies is in good agreement with the non-linear distortions on the excitation lines. Furthermore the grid is dense enough and approximately uniform, it does not suffer thus from the problems related to the more sparse grids. [49*-50*]

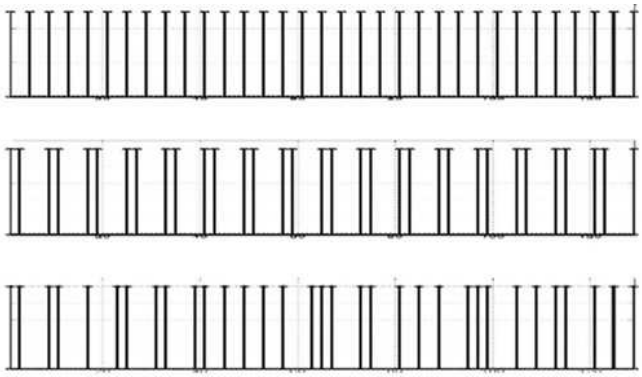

**Fig. 3.3.1** Examples of the multisine frequency grids: odd grid (upper), special grid (middle) and random grid (lower).

### I. **Randomized grid multisines**

| | |
|---|---|
| an example: $f_0$ * [1 2 3 • 5 6 • 8 • 10 11 12 13 • …] | no close formula |

Randomizing the frequency grid, i.e. changing the grid randomly from excitation to excitation, creates a nonstationary excitation which can be used to detect non-linearities in a fast way. The output of the system non-linearity, as a result of being a function of the nonstationary excitation, is therefore nonstationary too. Assuming that the measurement noise is stationary, the presence of non-linearities can be distinguished on that basis. To generate a randomized grid, groups of L (>2) consecutive lines are collected into blocks on a full frequency grid, from where one frequency line is dropped randomly, so as to form a random harmonic grid. [271]

In another approach used in the FRF measurements of the slowly time-varying linear systems a non-uniformly randomly spaced harmonics are used. For a multisine with $M$ frequency lines randomly selected (l-$q$)$M$, with $0 < q < 1$, harmonics are not excited (their amplitudes are set to zero), and the amplitudes of the $qM$ excited harmonics are all equal and chosen such that the rms value of the excitation equals 1. [173-174, 113]

## 3.4 Algorithms to work with the phases

### Crest Factor minimization

Phases in multisines were traditionally manipulated to keep their Crest-Factor low, i.e. to guarantee that:

$$CR(u(t)) = \frac{\max|u(t)|}{MSE(u(t))} = \frac{\|u(t)\|_\infty}{\|u(t)\|_2} = \min \tag{3.49}$$

Crest-Factor minimized signal has the smallest amplitude range for the same power level, which means that such signal can be amplified (injecting more power into the system and improving the SNR of the measurement) without fearing that the excessively large amplitudes will hit the hidden non-linearities and introduce non-linear distortions into the linear FRF measurements [199-201]. Crest Factor is minimal for the binary signal yielding value of 1. For any other signal this minimal value can be only approximated.

The required phases can be set algorithmically as so-called Schröder phases [214], [26-27], or by iterative minimizing algorithms: by suppressing the signal amplitudes by clipping [235], [189], or by minimizing $\|u(t)\|_{2p}$ norms with increasing $p$ (as $(\|u(t)\|_p)^p \rightarrow \|u(t)\|_\infty$ for $p \rightarrow \infty$) [85]. Recently it is considered that the second algorithm is a winner, yielding on the average better (lower) values of the Crest Factor in the general case and for large number of harmonics. It should be also noted that whatever the algorithm, log-tone like multisines are more difficult to compress and have higher Crest Factor values, than the more or less uniformly spaced multisines.

### $L_\infty$ multisines

The term was coined by [65] to denote the multisines with their Crest Factor minimized with the algorithm of [85].

**Note**: Manipulation of the phases seemingly destroys the randomness conditions imposed on the random multisines in (2.2.10-2.2.11), required in the proofs, so for a long time Crest Factor minimization was considered impossible in the non-linearly distorted FRF measurements. It turned out later (experimentally) that the situation isn't hopeless and that the $L_\infty$ multisines seem to retain their random properties required for the BLA measurements**.** The phenomenon is possibly related to the richness in local minima of the Crest Factor minimization surface, where randomly started minimum search stops at still randomly distributed places. The phenomenon however resists any kind of formal analysis[11].

### Shaping amplitude density

Crest Factor minimization improves the SNR conditions of the measurement, but it changes drastically the amplitude density towards that of a binary signal (the lowest Crest Factor = 1). Fig. 3.4.1 (left top) shows amplitude density of a Crest Factor optimized multisine. Such a design emphasizes strongly the extreme amplitudes; hence the linear approximation will be the best at the extreme amplitudes of the excitation. This unbalanced situation may be

11 Some experimental (not published) work of C. Evans, M. Solomou, J. Schoukens, and also my own.

undesirable for the general purpose excitation signal. An ideal signal would be a signal with e.g. a uniform (or normal) amplitude density. This requires an extended optimization method that would not only minimize the Crest Factor but would also impose the amplitude density with a specified power spectrum. The problem seems contradictory, because what improves Crest Factor, destroys good amplitude distribution and vice versa. The way out is the (heuristic) algorithm, which allows modifying the tails of the amplitude density function so that the user can balance between requirements regarding the distribution and the low Crest Factor [29*].

**Algorithm 3.4.1: Shaping amplitude density - basic algorithm**

Consider a single period of the multisine $u(t) = \sum_{k=1}^{M} U_k \cos(k\omega_0 t + \varphi_k)$, sampled at $t_k = kT_s$, with $k$ = 0, 1, …, $N$-1 and $\omega_0 = 2\pi/(NT_s)$. Consider the desired amplitude density function: $f_d(u)$ with $F_d(u) = \int_{v=-\infty}^{u} f_d(v)dv$, and define $Q$ as:

$$Q(l) = F_d^{-1}(P_l), \text{ with } P_l = \frac{1}{2N} + (l-1)\frac{1}{N}, \; l = 1, \ldots, N. \tag{3.4.2}$$

One iteration of the algorithm consists of four steps:

(1) Consider the set $U = \{u_L(t_k), k = 0, 1, \ldots, N\text{-}1\}$ in the $L$th iteration with:

$$u_L(t) = \sum_{k=1}^{M} U_k \cos(k\omega_0 t + \varphi_k^L); \tag{3.4.3}$$

(2) Sort $u_L(t_k)$ in the increasing order: $(Y_L, \tau) = \text{sort}(u_L(t_k))$, with $\tau$ the time instances of the sorted points;

(3) Create a new multisine $y_L(t_k)$ by replacing $Y_L$ with $Q$, such that $y_L = \text{sort}^{-1}(Q, \tau)$;

(4) Calculate the spectrum of $y_L$ (with the DFT applied to the samples $y_L(t_k)$), retain the phases and restore the original amplitude $U_k$ in this spectrum. The result is a new multisine:

$$u_{L+1}(t) = \sum_{k=1}^{M} U_k \cos(k\omega_0 t + \varphi_k^{L+1}). \tag{3.4.4}$$

This process is repeated until a suitable convergence criterion is met. ■

**Algorithm 3.4.2: Amplitude density with controlling the crest factor with a don't care zone**

If $f_d$ has long tails, the resulting multisine would have a very large Crest Factor. For this reason the amplitude domain must be partitioning into $D_{crest}$ and $D_f$.

Assuming symmetric distributions:

$$D_{crest} = ]-\infty, -a] \bigcup ]a, \infty], \text{ and } D_f = [-a, a], \tag{3.4.5}$$

where [-$a$, $a$] is the support of the desired amplitude density with $-a = F_d^{-1}(0), a = F_d^{-1}(1)$. Consequently after imposing $F_d$, all samples of the signal are concentrated in $D_f$, however after imposing $U_k$ some of the samples $x_{L+1}(t_k)$ will be smeared outside $D_f$, into $D_{crest}$, increasing the crest factor over the value of $a$.

The amplitude domain will be now partitioned in three parts by adding a don't care region. The amplitude distribution is only imposed in the given amplitude interval $D_f$ and left free

outside this interval:

$D_{crest} = ]-\infty,-a] \bigcup ]a,\infty]$, $D_{dc} = [-a, F_d^{-1}(\varepsilon)[ \bigcup ]F_d^{-1}(1-\varepsilon), a]$, and

$D_f = [F_d^{-1}(\varepsilon),\ F_d^{-1}(1-\varepsilon)], \qquad 0 \le \varepsilon \le 0.5$. (3.4.6)

To reduce the crest factor the samples belonging to $D_{crest}$ are clipped to the borders $a$ or $-a$ before a new iteration cycle is started. ■

**Note**: In practice $F_d$ is imposed on $N_\varepsilon = [(1\text{-}2\varepsilon)N]_{\text{ceil}}$ points belonging to $D_f$ and distributed following the Alg. 3.4.1 as: $Q(l) = F_d^{-1}(P_l)$, $P_l = \frac{1}{2} + \frac{(l - N_\varepsilon / 2)}{N}$, $l = 1, \ldots, N_\varepsilon$. (3.4.7)

During the design of the desired density function $f_d$ it is necessary to take care that no conflicting constraints are imposed. The power of the multisine is imposed by its amplitude spectrum by $\sum |U_k|^2$, on the other hand also the amplitude density function sets the power via its second moment $\int_{-\infty}^{\infty} u^2 f_d(u)\, du$. In case of desired densities of the infinite support, they must be truncated to a suitable finite interval (e.g. $\pm\alpha\sigma$) to get an acceptable Crest Factor. The second order moment is restored by selecting a proper scaling factor $S$ for $f_d$:

$$f_d(u) = \begin{cases} f(u)/S & -a \le u \le a \\ 0 & elsewhere \end{cases} \quad (3.4.8)$$

with $S$ the area under the truncated density.

**Algorithm 3.4.3: Amplitude density – optimizing the Crest Factor**

The algorithm resembles the Alg. 3.4.2., only the clipping algorithm is improved. Instead of clipping all the samples in $D_{crest}$ to the borders $a$ and $-a$, a varying clipping level $C_L \ge F_d^{-1}(\varepsilon)$ is chosen using the algorithm of [235] to compress the signal. All samples with amplitude larger than $C_L$ are clipped towards this level. That way even better Crest Factor can be achieved. ■

In the presentation of the algorithm the uniform frequency grid was used, however in theory the algorithm could work with any frequency grid. Experience shows that the irregular grids are much more difficult to obtain a fast convergence. Although no formal proof of the convergence is available, the tests indicate a good convergence of the algorithms in case of uniform grids. However theoretical handling of the convergence question seems hopeless, considering:

1. A particular solution is a fixed point (by construction). More solutions are possible for dense grids and finite error approximation.

2. The operator of Alg. 3.4.1 is a non-linear, discontinuous, and non-expansive operator.

3. The multisine in the algorithm is confined to a non-convex set (a shell of a fixed MSE value, as the amplitude sorting and similarly the imposing of the pdf and the amplitude spectrum is $\|.\|_\infty$, $\|.\|_2$ invariant).

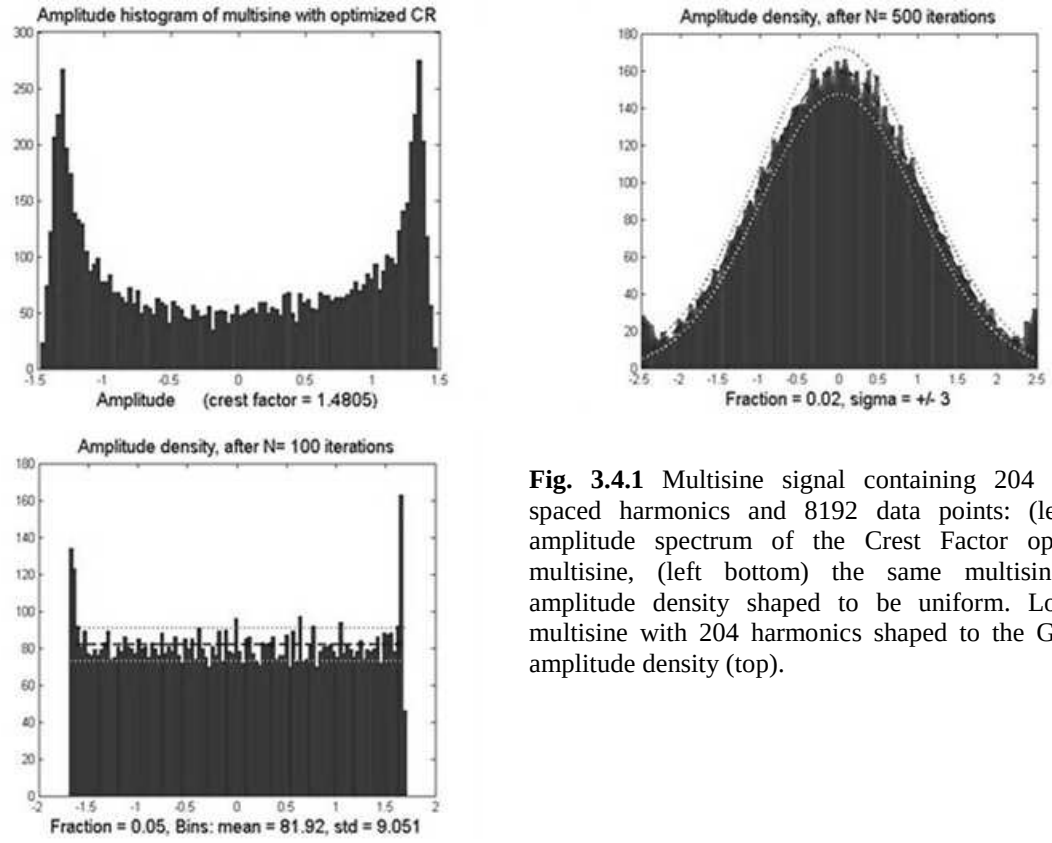


**Fig. 3.4.1** Multisine signal containing 204 linearly spaced harmonics and 8192 data points: (left top) amplitude spectrum of the Crest Factor optimized multisine, (left bottom) the same multisine with amplitude density shaped to be uniform. Log tone multisine with 204 harmonics shaped to the Gaussian amplitude density (top).

## 3.5 The question of choice

When multiple signals are available as the excitation, the well balanced choice is not always easy, and usually it is also a part of the trade-off, which may adversely affect the measurement results. What is then the situation with the multisine excitations?

Before the multisine signals of high harmonic content became theoretically and practically (instrumentation!) established, the prevailing excitation in many measurement fields was the white Gaussian noise. Although the developed theory stated the asymptotic equivalence of both kinds of excitations from the point of view of the non-linear distortions modelled by the Volterra series, practically important was to analyse the excitations within finite measurement conditions and to formulate pragmatic guidelines for the measurement design. The experimental comparison covered:

- ideal Gauss noise,

- Gauss noise filtered with 10th order low-pass Butterworth filter and clipped in the amplitude ("practical" Gauss noise),

- full frequency grid random phase multisine,

- random phase multisine defined on the odd frequency grid,

- random phase multisine defined on the odd-odd frequency grid,

- multisine defined on the odd frequency grid with the amplitude density shaped with the suitable choice of the phases.

Based on the theory and the simulations the practical advice for weakly non-linear FRF measurements is to use odd-odd random multisines of high harmonic content, considering that:

- the measured FRF is the same as the one measured with the Gaussian excitation,

- due to the drop-out of the effects of the even non-linearities the uncertainty of the measurement (the non-linear noise) is less,

- it is possible to detect and quantify odd and even nolinearities separately (even non-linearities at the even frequency lines, and odd non-linearities at the non-excited odd frequency lines).

It was also established that clipping the (amplitude of the) Gaussian noise elevates the bias on the FRF, assuming that the non-linear distortions do contain odd components. Furthermore it turned out that the usual ± 3 σ clipping is not enough in the weakly nonlinear measurements (it is advised to clip at ± 4 σ). [8*, 27*, 29*, 33*, 162]

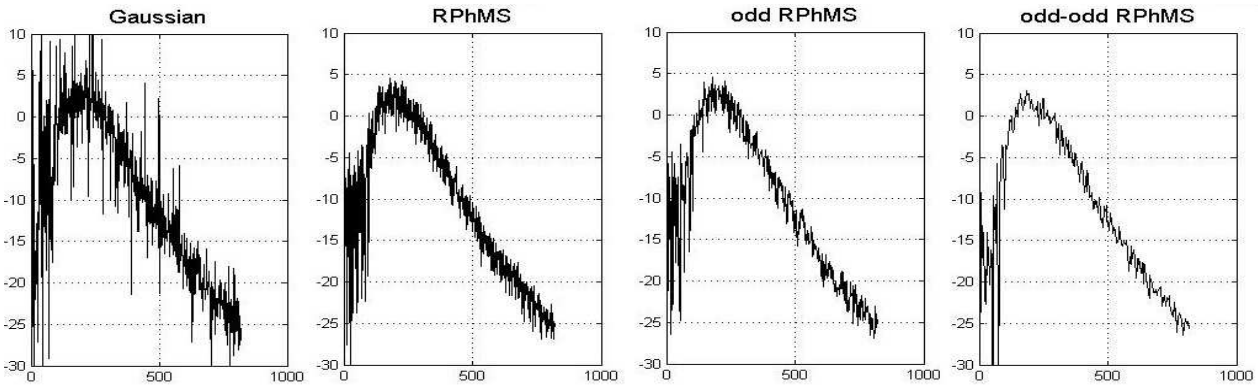


**Fig. 3.5.1.** A qualitative example: ETFE of a weakly nonlinear system from Fig. 1.1.1 measured with a single application of the Gaussian noise (left) and the random phase multisine defined on full, odd, and odd-odd frequency grids (all signals normalized to the unit power) (right). Using the random phase multisine yields lower level of the non-linear noise, consequently demands shorter measurement (averaging) time. This advantage can be further amplified by manipulating the frequency grid.

### Misusing random multisines

The extensive usage of the random multisine excitations, and the work on the Matlab Toolbox [106], where as the designing aim a “user-resistant” excitation was sought (which would provide acceptable results even if the user misuses the excitations in the measurement set-up), brought into light a very interesting phenomenon.

Random multisine (whatever the frequency grid) is a periodic signal and its proper usage requires processing of the integral number of periods in the measurement data. In a number of measurement applications however the measurement time (due to the chosen basic harmonic frequency and settling of the transients) can be so long, that the user will be tempted or pressed to stop the measurement before due. Are then the measurement results lost as not properly processable? Simulations indicate that the random multisines are very resistant to such misuse. Random multisine measured only in a part of its period is still very close in its

spectral properties to the ideal signal, meaning that the FRF measurement with such an input won't be much distorted. [215, 203]

**Compound experiment design**

The periodicity of the excitation not only means a tool against the leakage. It is a powerful property as it is (usually) not shared by the noise disturbances covering the measured data. The presence and the amount of the output or the measurement noise can be discovered on the fly by putting the periodicity at work [50].

The normal application of the random multisines would be to apply to the system input the series of the realizations for different independent phases: $\{u(t,\varphi_1),\ \ldots,\ u(t,\varphi_k),\ \ldots,\ u(t,\varphi_N)\}$.

Such measurement yields the Best Linear Approximation of the system. If the measurement noise is also present, it would be good idea to measure its variance to judge properly the amount of averaging needed to suppress it (or to introduce it as a weighting factor into the criterion function for the parametric estimates). Although the random multisine is a stochastic signal, its every single realization is deterministic, and the output of the non-linear system to such excitation will also be deterministic. Let $M = K*N$ and let stretch the input a bit more:

$$\{\underbrace{u(t,\phi_1),u(t,\phi_1),\cdots,u(t,\phi_1)}_{K\ times},\underbrace{u(t,\phi_2),u(t,\phi_2),\cdots,u(t,\phi_2)}_{K\ times},\cdots,\underbrace{u(t,\phi_N),u(t,\phi_N),\cdots,u(t,\phi_N)}_{K\ times}\} \qquad (3.5.1)$$

In case when $Y(l) = Y(l,\varphi,\eta) = V[U](l,\varphi) + N(l,\eta)$, i.e. the measurements are noisy, and the measured output depends not only on the input random parameter $\varphi$, but also on a random parameter $\eta$ characterizing the realization of the measurement noise, the usual computation of the Best Linear Approximation (averaging over all available records) will yield:

$$\hat{G}_{BLA}(l) \approx E_{\phi\eta}\{\frac{Y(l)}{U(l)}\} = E_{\phi\eta}\{\frac{Y(l,\phi,\eta)}{U(l,\phi)}\} = E_{\phi}\{\frac{E_{\eta}\{Y(l,\phi,\eta)\}}{U(l,\phi)}\}. \qquad (3.5.2)$$

As a result, within the same measurement set-up the measurement noise variance will be attenuated by $1/K$, and the non-linear noise variance by $1/N$. During the averaging both variances can be measured independently providing a fine insight into the distortion situation of the measurement (Fig. 3.5.2).

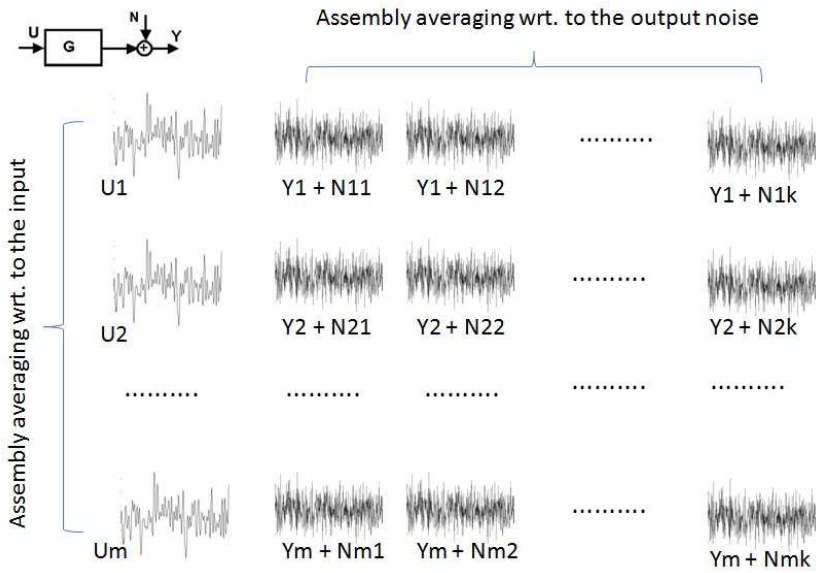


**Fig. 3.5.2.** Measurement setup to separate the ETFE BLA measurements from those of the output noise. When a single realization of a periodic input is applied and more periods measured from the output, the non-periodic output noise can be easily separated and its variance and nonparametric spectral properties measured. Applying independent inputs and averaging along them, on the other hand, gets rid of the non-linear noise and yields the FRF BLA.

## 3.6 Robustness in the SISO BLA measurements

Various signals can be used as the excitations for the FRF measurements. Some of them lead already in many measurement areas to the development of a specific measurement methodology (based upon these particular signals, e.g. the Gaussian noise). If new (better) excitations are proposed, the primary issue is that of portability and comparison – will we measure the same quantity or not?

If the measurements made by the older and the newer excitations are comparable, it makes it possible to switch over to the new methods and to extend the measurement technological toolbox, without the fear that the existing achievements will be uncomparable and jeopardized.

In the related research two questions were addressed. If the number of the harmonics tends to infinity, is the BLA asymptotically equivalent for different excitation signals? The second question dealt with the asymptotic properties of the BLA FRF measured with multisines defined on different frequency grids.

## SISO BLA equivalence of the multisine excitations

Luckily the first results were encouraging:

**Theorem 3.6.1: SISO model equivalence of random multisines, periodic noise, and Gaussian noise.** For the listed types of the excitation signals the SISO BLA and the non-linear variance retain all the properties proved for the random phase multisines and converge

(also in the moments) to the same limit values at the rate $O(M^{-1})$. The BLA is a continuous function of the frequency, with the continuous derivatives if the approximated Volterra system is continuous, with continuous derivatives.

**Proof:** See [162-163]. ■

## Robustness of the SISO BLA measurements

Further research developed a mature model of the asymptotic robustness of the SISO BLA measurements. It was recognized that if the Volterra-series is excited with a random multisine of a high harmonic content, then with the proper normalization, the BLA expression (2.2.16) is formally equivalent to a Riemann-(integral) sum. From this starting point, for a number of practically relevant excitations, we succeeded to formulate (Riemann) equivalence conditions, warranting for these excitations and for the non-linear systems approximable in the mean square sense by the Volterra-series the portability of the BLA measurements.

The BLA measured with the Riemann-equivalent excitation signals (taking into account the spectral equivalence from Fig. 3.6.1) can be written as:

$$G_{BLA}^{2\alpha-1}(f) = c_\alpha \int_0^{f_s/2} \dots \int_0^{f_s/2} G_{f,-f_1,f_1,\dots,f_{\alpha-1}}^{2\alpha-1} S_U(f_1) \dots S_U(f_{\alpha-1}) df_1 \dots df_{\alpha-1} \qquad (3.6.1)$$

and similar expression can be formulated for the non-linear noise variance (albeit with no so simple direct reference to the Volterra-kernels) (for the derivation, see [46*]).

SIGNALS OF THE EQUIVALENCE CLASS $E_{S_U}$ AT FREQUENCY $\Omega = 2\pi f_s k_\Omega / N$

| Signal | frequency grid | Power or amplitude spectrum |
|---|---|---|
| Gaussian noise | full | $S_U(\Omega)$ (user defined) |
| periodic noise | full | $E\{\lvert U(\lvert k_\Omega \rvert)\rvert^2\} = S_U(\Omega) f_s$ |
| multisine | full | $\lvert U(\lvert k_\Omega \rvert)\rvert^2 = S_U(\Omega) f_s$ |
| random grid multisine | full random with prob $p$ | $E\{\lvert U(\lvert k_\Omega \rvert)\rvert^2\} = S_U(\Omega) f_s / p$ |
| odd random multisine | odd | $\lvert U(\lvert k_\Omega \rvert)\rvert^2 = 2 S_U(\Omega) f_s$ if $k_\Omega$ is odd, and zero elsewhere |
| odd random grid multisine | odd random with prob $p$ | $E\{\lvert U(\lvert k_\Omega \rvert)\rvert^2\} = 2 S_U(\Omega) f_s / p$ if $k_\Omega$ is odd, and zero else where |

**Fig. 3.6.1.** Correspondence in the spectral content for various excitation signals.

The robustness of the BLA measurements for various families of the multisines can be addressed also from the frequency grid perspective. There are already many frequency grids defined in the measurement practice (see Section 3.3, but also other ideas found in the literature, e.g. [121]). Due to the fact that the grid, in the limit $M \to \infty$, is responsible in the bias and in the variance expressions for the convergence of the finite sums to the multiple frequency integrals (of arbitrarily high dimensions, see series (3.6.1)), this grid must fill the frequency band in a more and more uniformly dense manner, i.e. it must be a uniformly distributed sequence [144, 20*].

**Definition 3.6.1**: **Uniformly distributed multisine.** An uniformly distributed *N*-periodic random phase multisine with *M* harmonics is defined on the frequency grid $S_M^+$, which is a uniformly distributed sub-set of the reference frequency grid (i.e. the discrete full frequency grid of *N/2* harmonics): $S_M^+ = [k_1 \quad k_2 \quad ... \quad k_M] \subseteq S_{N,0}^+ = [1 \quad 2 \quad ... \quad (N/2)-1]$ ($k_1$=1, see Def. 2.2.4, the multiplication by the fundamental frequency $f_0$ is omitted for clarity), in a sense that for every *M* and *N* and for every frequency sub-band $I = [m_1, m_2]$, $1 < m_1 < m_2 < N/2$, the limit:

$$\lim_{M \to \infty} \frac{1}{M} \sum_{m \in I} 1 = \frac{m_2 - m_1}{2/N}. \tag{3.6.2}$$

exists, meaning that in the limit the frequencies of the $S_M^+$ grid uniformly and densely fill up every subband of the full frequency band of the signal . ■

The question now is to what extend the BLA measurements performed over different frequency grids are equivalent (and with the proper spectral equivalence, converge to the limit (3.6.1)). Based on the theory of the uniformly distributed sequences it can be stated that if the frequency grids used in the measurement design are characterized - as point sets - with the decreasing so called discrepancy, then the corresponding BLA measurements are asymptotically equivalent.

Let us start with a simple 3rd order SISO Volterra system. Let the frequency grid of the excitation be modeled as the point set:

$$S_M^+ = [k_1 \quad k_2 \quad ... \quad k_M] \subseteq S_{N,0}^+ = [1 \quad 2 \quad ... \quad (N/2)-1] \tag{3.6.3}$$

$$\text{Then: } Y(l) = G^1(l)U(l) + \sum_{k_1,k_2 \in S_M^+ \cup S_M^-} G^3(k_1,k_2,k_3) \prod_{i=1}^{3} U(k_i), \quad k_3 = l - k_1 - k_2 \tag{3.6.4}$$

$$\text{Denoting as: } K_{3,M} = \frac{1}{M} \sum_{k \in S_M^+} G^3(l,k,-k), \quad K_3 = \int_0^B G^3(l,f,-f)df \tag{3.6.5}$$

the kernels appearing in the BLA expression the question now is what is the bound on the:

$$\varepsilon_1(l) = \left| G_{BLA,M}(l) - G_{BLA}(l) \right| = 6 \left| K_{3,M} - K_3 \right| \leq ? \tag{3.6.6}$$

error, where $K_3$ is the kernel measured with the reference noise and $K_{3,M}$ is the kernel measured over the given frequency grid. In the general case, for the measurements performed over two different frequency grids the error is similarly:

$$\varepsilon_2(l) = \left|G_{BLA,M_1}(l) - G_{BLA,M_2}(l)\right| = 6\left|K_{3,M_1} - K_{3,M_2}\right| \leq ? \tag{3.6.7}$$

and can be bounded using so called Koksma-Hlawka inequality [144] as:

$$\varepsilon_2(l) \leq 6V(G^3(l, f, -f))(D^*_{M_1} + D^*_{M_2}) \tag{3.6.8}$$

where $V$ is the variation function measuring the smoothness of the kernel and $D^*_M$ is so called star discrepancy, the measure of the largest difference from the uniform distribution. Below one can see the discrepancy of some, already mentioned, characteristic frequency grids.

If the discrepancy of the frequency grids is of order $O(M^{-1})$, of similar order is the difference in the BLA measurements. The results can be generalized for the arbitrary order Volterra-system:

$$\varepsilon_1(l) = \left|G_{BLA,M}(l) - G_{BLA}(l)\right| \leq 6\sum_{\alpha=3,odd}^{\infty} K_\alpha \varepsilon_\alpha \tag{3.6.9}$$

where $K_\alpha$ depends on the kernel functions (cf. (3.6.5)) and $\varepsilon_\alpha$ can be estimated from the Koksma-Hlawka inequality [20*]:

$$\varepsilon_\alpha \leq D^*_M \sum_{k=0}^{\alpha} \sum_{1 \leq i_1 < i_2 < \ldots < i_k \leq \alpha} V^{(k)}(G^\alpha; i_1, \ldots, i_k) \tag{3.6.10}$$

where the term in the parentheses is the multidimensional variation of the frequency kernel $G^\alpha$ [144]. The results can be extended to the colored excitations with the additional condition on the bounded variation of the signal frequency spectrum (see [20*]). In consequence, based on the smoothness of the bias and the variance (Th. 3.6.1), furthermore on the standard properties of the Riemann sums (2.2.8) for (3.6.1) and on (3.6.10), the variations in the BLA measurements can be bounded by the variations in the frequency grid discrepancy leading to the following Corollary:

**Corollary 3.6.2: Frequency grid robustness of the BLA measurements.** The BLA and the non-linear variance are robust to the perturbations of the frequency grid. ■

More can be said in case of Wiener-Hammerstein systems, with more insight into the system structure and the place of the non-linear component. In such case [20*]:

$$\varepsilon_2(l) \leq C\,|G^1(l)|\,V(|R(f)|^2)(D^*_{M_1} + D^*_{M_2}) \tag{3.6.11}$$

The error thus is more dependent on the smoothness of the input linear system. [44*-45*].

$D_M^*$

| full | odd | odd-odd | special-odd | random grid |
|---|---|---|---|---|
| $1/M$ | $2/M$ | $4/M$ | $\approx 5/M$ | $\approx 4/M$ |

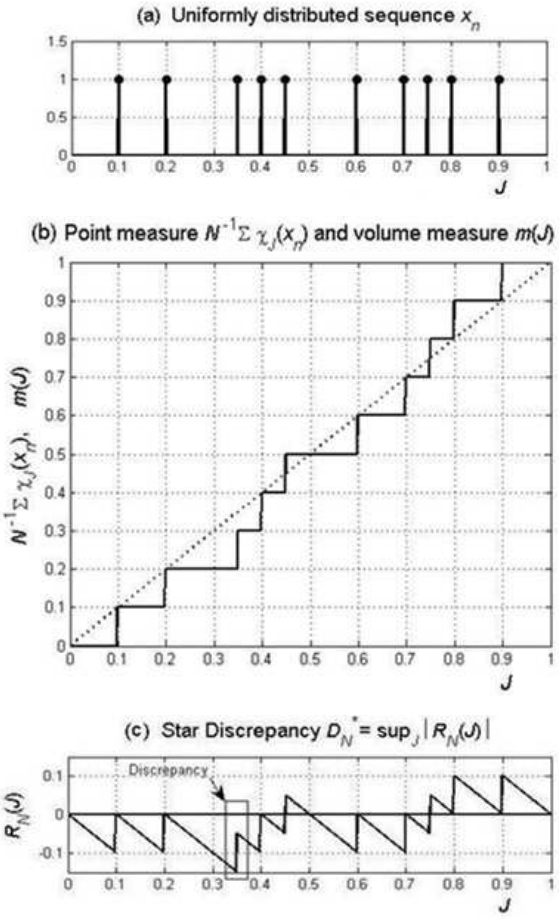


**Fig. 3.6.2** Quantities used to qualify properties of the uniformly distributed sequences: (a) mass distribution of the point set (point measure), (b) cumulative mass distribution (continuous) and the volume measure of the interval (point line), (c) star discrepancy.

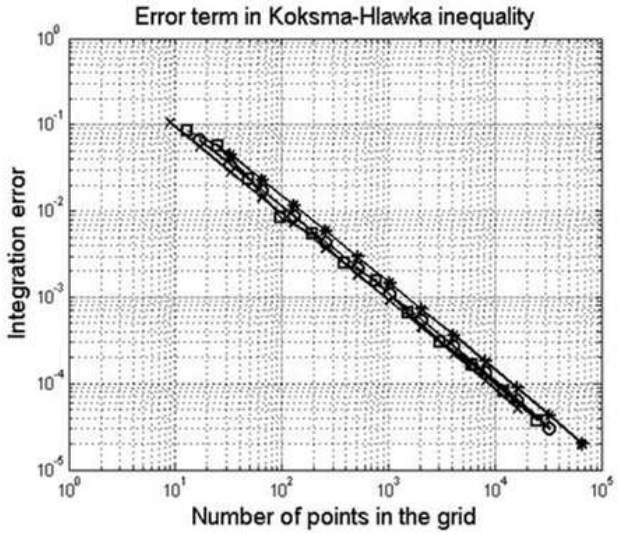


**Fig. 3.6.3 Behavior** of the error for the measurements with various random multisines. The measured system is a 3rd order Wiener-Hammerstein system with the 9th order Butterworth input system and the 9th order Tchebishev output system (10dB pass-band ripple). (*) full, (o) odd, (x) odd-odd, and (square) random frequency grid.

# 4. General MIMO BLA theory

## 4.1 MIMO Volterra systems

In numerous real life phenomena multiple independent or interdependent effects act together toward joint results. These can be modelled as so called Multiple-Input Multiple-Output (MIMO) systems. It is easy to extend SISO Volterra models to MIMO Volterra systems without feedback, i.e. when the outputs can be modelled separately as multiple-input single-output (MISO) systems with $N$ inputs.

A number of questions pops up naturally regarding the portability of the SISO BLA results and also about possible new phenomena specific to the higher dimensional systems. Once we define well behaving MIMO Volterra systems the primary issue will be the choice of the multiple input signals and the analysis how the measurements in one input-output channel influence or perturb the measurements in the other channels.

**General Assumptions**

The well developed SISO Volterra theory (Sect 2.2) lists numerous advantageous properties for well behaving kernels and input signals, resulting in further nice properties of the BLA approximation (Sect 2.3).

In the MIMO case there is more freedom. Feed-forward kernels and the cross-channel kernels can differ in properties, diverse input signals can be applied to various inputs, and in consequence the analysis of the MIMO Volterra systems can be difficult.

In the following we recall that the informal scope of the research is the nonparametric linear FRF measurement on a weakly non-linear system, i.e. the situation when little if any a priori information exists about the system under study and it is just the aim of the measurement to gain some information about the frequency band, the shape of dynamics, the level of non-linear distortions, etc. Consequently there is usually no measurement technical reason to specify essentially different excitation signals at the system inputs.

Considering the multisine signals applied to different inputs we may assume thus, that:

**Assumption 4.1.1**

(a) The frequency grids at the inputs coincide, whatever they are, or

(b) Though multisines at different $k$ inputs can be distributed on different frequency grids $S^+_{M,k}$ (see Def. 2.2.4), but these frequency grids are dense subsets of the same uniformly distributed frequency grid $S^+_{0,N}$, in a sense, that for $\forall k, S^+_{M,k} \subseteq S^+_{0,N}, \left|\bigcap_k S^+_{M,k}\right| = O(M)$ (i.e. they have plenty of common frequencies); that way all inputs will have the same common period. ■

The derivations in Sect 4.1 are made for the output of the MIMO Volterra system with bounded kernels, excited by normalized excitations, at the frequencies common to the all inputs (Case a.), as this situation constitutes the FRF measurement practice almost

exclusively. We will also assume that no other disturbing noise sources are present, focusing thus the analysis on the input signals and the non-linear effects. Although not investigated formally it can be nevertheless conjectured that a number of results will be valid also for the Case b.

**Definition 4.1.1. *N*-input *K*-output MIMO Volterra series.** An *N*-input *K*-output MIMO Volterra series can be described in the time domain as:

$$y_k(t) = V[u_1, u_2, \cdots, u_N](t) = \sum_{\alpha=1}^{\infty} y^{\alpha}(t) = \sum_{\alpha=1}^{\infty} \sum_{j_1 j_2 \ldots j_\alpha} y^{j_1 j_2 \ldots j_\alpha}(t) \tag{4.1.1}$$

where the second sum runs over the outputs of all pure and mixed $\alpha$th order kernels, and a particular $\alpha$th order kernel, excited by input signals of indices $j_1, j_2, \ldots, j_\alpha$, in time domain yields:

$$y^{j_1 j_2 \ldots j_\alpha}(t) = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} g^{j_1 j_2 j \ldots j_\alpha}(\tau_1, \ldots, \tau_\alpha) u_{j_1}(t-\tau_1) u_{j_2}(t-\tau_2) \ldots u_{j_\alpha}(t-\tau_\alpha) d\tau_1 \ldots d\tau_\alpha \tag{4.1.2}$$

and $k = 1, \ldots, K$, and every $j_m \in \{1, 2, \ldots, N\}$. In the following we will drop the output index, as redundant. Every output can be considered separately as the result of a different Volterra series.

In the frequency domain the system model is (c.f. (2.2.3-2.2.4) for SISO):

$$Y(l) = V[U_1, U_2, \cdots, U_N](l) = \sum_{\alpha=1}^{\infty} Y^{\alpha}(l) = \sum_{\alpha=1}^{\infty} \sum_{j_1 j_2 \ldots j_\alpha} Y^{j_1 j_2 \ldots j_\alpha}(l) \tag{4.1.3}$$

$$Y^{j_1 j_2 \ldots j_\alpha}(l) = M^{-\alpha/2} \sum_{k_1, \ldots, k_{\alpha-1} \in S_M^- \cup S_M^+} G^{j_1 j_2 \ldots j_\alpha}(k_1, k_2, \ldots, k_{\alpha-1}, k_\alpha) U_{j_1}(k_1) U_{j_2}(k_2) \ldots U_{j_\alpha}(k_\alpha) \tag{4.1.4}$$

where $l$ is the discrete frequency, $l = \Sigma\, k_i$, $i = 1 \ldots \alpha$, and:

$$U_j(k) = U_j(k/N) = \hat{U}_j(k/N) e^{j\phi_j(k)} \tag{4.1.5}$$

Similarly to the SISO case the kernels and the signals must be bounded. Kernels $G^{j_1 j_2 \ldots j_\alpha}$ are bounded by $\max |G^{j_1 j_2 \ldots j_\alpha}| = M_{j_1 j_2 \ldots j_\alpha}$. The series is convergent for every $l$, if the inputs are normalized to unit power and have uniformly bounded spectral amplitudes $|U_k| \le M_{U,k}/\sqrt{M} < \infty$, furthermore if together (c.f. (Def. 2.2.2)):

$$\sum_{\alpha=1}^{\infty} M_\alpha M_U^\alpha < \infty \tag{4.1.6}$$

where $M_\alpha = \max_{\text{all } \alpha \text{ order kernels}} |M_{j_1 j_2 \ldots j_\alpha}|$, $M_U = \max_{\text{all input signals}} |M_{U,k}|$. ■

**Definition 4.1.2**: **Non-linear system class of interest**. The class of systems of interest in the following is restricted to those systems which are the limits in the least-square sense of the convergent Volterra series defined in Def. 4.1.1. If otherwise not specified, the term 'non-linear system' will be used in this context. ■

**Note**: Special measurement situations, like e.g. measuring modulators with carrier inputs, or industrial installations with step-like signals, can be handled as a normal FRF measurement at a specific "working points" of the system. Furthermore the BLA theory was recently extended to (periodic) multilevel excitation signals [267-269].

We summarize now in the analogy to the SISO case some useful properties of the MIMO Volterra systems.

**Theorem 4.1.1**: **Error bound for the truncated MIMO Volterra series**.

$$\| V[u](t) - V^{(K)}[u](t) \|_\infty \leq \sum_{k=K+1}^{\infty} \| g_k \|_\infty (\| u \|_\infty)^k \qquad (4.1.7)$$

**Proof:** By the analogy to Th. 2.2.1, where $\|.\|_\infty$ is normal sup norm, maximized over all inputs, and the bounds are taken similarly as the maximum over all kernels of the same order, or over all inputs. ■

**Theorem 4.1.2**: **Boudedness of the MIMO Volterra series.** MIMO Volterra series is a Bounded-Input Bounded-Output for each of its input-output paths.

**Proof:** Inputs are bounded, so majorizing them with the worst-case input bound reduces the problem to the SISO case. ■

Within the measurement technical circumstances assumed in the dissertation the following observation may actually serve as theorems, however formally they can be stated only as the assumptions:

**Assumption 4.1.2**: **Continuity of the MIMO Volterra series.**

**Comment**: If the MIMO Volterra series is convergent (Def 4.1.1), due to (4.1.4)-(4.1.6) it can be stated that when all but one input are fixed in their properties (spectral content, frequency grid, phases) the remaining SISO Volterra system (characterized by $M_U^{\alpha-1} G^{j_1 j_2 \cdots j_\alpha}$ kernels) posses all the properties listed in Sect 2.2 for the bounded kernel, bounded input SISO Volterra systems, consequently is continuous. This is so called separate (component-like) continuity, which in itself does not imply the joint continuity of the multivariable function. However if an *N*-dim multivariable function is separately continuous in all of its variables, then on suitable domains (unit cube) the set of discontinuity points is of at most *N*-2 dimension, and with additional smoothness conditions, the set of discontinuity points is nowhere dense. The counter examples usually show discontinuity at the origin, which is without significance, as the origin means not applying the input signals at all. So for the practical purposes we assume that the MIMO Volterra system (if properly bounded) is also continuous. [103] ■

**Assumption 4.1.3**: **Steady state theorem for MIMO Volterra series.** For every input *k*, let $u_k$ and $u_{steady,k}$ be signals within the region of convergence of (4.1.2), and suppose that $u_k(t) \to u_{steady,k}(t)$ as $t \to \infty$ for all *k*. Then $V[u_1,u_2,\ldots,u_N](t) \to V[u_{steady,1},u_{steady,2},\ldots,u_{steady,N}](t)$ as $t \to \infty$.

**Comment:** By the assumed continuity of the MIMO Volterra system, see Ass. 4.1.2. ■

**Assumption 4.1.4**: **Periodic steady state theorem for MIMO Volterra series.** If the inputs $u_k$ are all periodic with the same period *T* for $t \geq 0$ then the output $V[u_1,u_2,\ldots,u_N]$ approaches a steady state, also periodic with period *T*.

**Comment:** If the inputs are of the same period, then their spectra can be jointly bounded and then SISO theorem Th. 2.2.3 applies. ■

**Assumption 4.1.5**: **Extended periodic steady state theorem for MIMO Volterra series**. If the inputs $u_k$ are all periodic with the periods $T_k$, possessing the least common multiple $T$ = lcm($T_1$, $T_2$, …, $T_N$), then the output V[$u_1$,$u_2$,…,$u_N$] approaches a steady state, also periodic with period $T$.

**Comment:** Since every input is also $T$-periodic by definition, then by Ass. 4.1.4 the output should be also $T$-periodic. There cannot be shorter periodicity in the output, because it would assume that all the inputs reached already a repetition, impossible by the definition of the least common multiple. ■

In a number of MIMO applications block models are also a useful modeling tool. The general unified definition of the MIMO Wiener-Hammerstein system and the related special cases does not exist. Here we present a definition used in the followings:

**Definition 4.1.3**: ***N*-input *K*-output MIMO Wiener-Hammerstein system** is built from $K$ independent $N$-input MISO Wiener-Hammerstein systems. Every $N$-input MISO Wiener-Hammerstein system has $N$ parallel different input dynamics, $N$-to-1 static non-linearity, and a common output dynamics (see Fig. 4.1.1 for 2-input MISO Wiener-Hammerstein system). ■

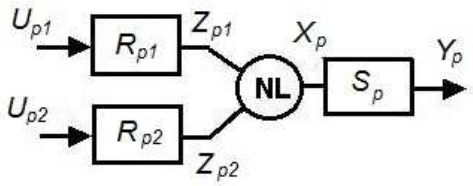


**Fig. 4.1.1**. Generic MISO Wiener-Hammerstein structure. $R_{p1}(l)$ and $R_{p2}(l)$ are the linear input dynamics, $S_p(l)$ is the linear output dynamics for the $p$th output, and $NL$ is the static non-linearity.

Volterra kernels of order α of $N$-input MISO Wiener-Hammerstein system are:

$$G^{j_1 j_2 \cdots j_\alpha}(k_1,k_2,\cdots,k_\alpha) = const \times S_p(k_1+k_2+\cdots k_\alpha)\prod_{k=1}^{\alpha} R_{p j_k}(k_n) \qquad (4.1.8)$$

where $j_1$, $j_2$, …, $j_\alpha$ are not necessarily all different (and similarly for Hammerstein, or Wiener systems).

## 4.2 MIMO FRF measurements

Multiple input setups were investigated extensively in the linear FRF measurements [199, 210, 125, 162]. As it was assumed in the SISO case in the following we consider the output error scheme, and the primary object of the investigation are the nonlinear effects coming from the system itself and not from the measurement noises. The MIMO system with $N$ inputs and $K$ outputs is measured as a set of MISO systems. To this purpose $J$ experiments are made, generating $J$ times all the input signals, then cutting (after the transients settle) the successive records from the inputs and the output.

The measured signal amplitudes at the frequency $l$ can be arranged as:

$$\mathbf{Y}(l)=\mathbf{G}(l)\mathbf{U}(l)+\mathbf{N}_Y(l)=\begin{bmatrix} Y_1^{(1)}(l) & Y_1^{(2)}(l) & \cdots & Y_1^{(J)}(l) \\ Y_2^{(1)}(l) & Y_2^{(2)}(l) & \cdots & Y_2^{(J)}(l) \\ \cdots & \cdots & \cdots & \cdots \\ Y_K^{(1)}(l) & Y_K^{(2)}(l) & \cdots & Y_K^{(J)}(l) \end{bmatrix}= \tag{4.2.1}$$

$$\begin{bmatrix} G_1^1(l) & G_1^2(l) & \cdots & G_1^N(l) \\ G_2^1(l) & G_2^2(l) & \cdots & G_2^N(l) \\ \cdots & \cdots & \cdots & \cdots \\ G_K^1(l) & G_K^2(l) & \cdots & G_K^N(l) \end{bmatrix}\begin{bmatrix} U_1^{(1)}(l) & U_1^{(2)}(l) & \cdots & U_1^{(J)}(l) \\ U_2^{(1)}(l) & U_2^{(2)}(l) & \cdots & U_2^{(J)}(l) \\ \cdots & \cdots & \cdots & \cdots \\ U_N^{(1)}(l) & U_N^{(2)}(l) & \cdots & U_N^{(J)}(l) \end{bmatrix}+\begin{bmatrix} N_{Y,1}^{(1)}(l) & N_{Y,1}^{(2)}(l) & \cdots & N_{Y,1}^{(J)}(l) \\ N_{Y,2}^{(1)}(l) & N_{Y,2}^{(2)}(l) & \cdots & N_{Y,2}^{(J)}(l) \\ \cdots & \cdots & \cdots & \cdots \\ N_{Y,K}^{(1)}(l) & N_{Y,K}^{(2)}(l) & \cdots & N_{Y,K}^{(J)}(l) \end{bmatrix}$$

The indices in parentheses are the serial numbers of the experiments. **U** is a $N \times J$ input matrix of complex input amplitudes (2.2.11), and will play important role in the followings:

$$\mathbf{U}(l)=\begin{bmatrix} U_1^{(1)}(l) & U_1^{(2)}(l) & \cdots & U_1^{(J)}(l) \\ U_2^{(1)}(l) & U_2^{(2)}(l) & \cdots & U_2^{(J)}(l) \\ \cdots & \cdots & \cdots & \cdots \\ U_N^{(1)}(l) & U_N^{(2)}(l) & \cdots & U_N^{(J)}(l) \end{bmatrix} \tag{4.2.2}$$

**G** is a $K \times N$ matrix of the true FRF values for a particular input-output channel; **Y** and $\mathbf{N}_Y$ are $K \times J$ matrices of the output amplitudes and the output measurement noise amplitudes accordingly. We will assume also in the following that $J = N_B \times N$, i.e. the number of experiments is an integral number of $N$ experiment blocks. In the future we will distinguish also a special $N \times N$ input matrix built from $J = N$ experiments:

$$\mathbf{U}_N(l)=\begin{bmatrix} U_1^{(1)}(l) & U_1^{(2)}(l) & \cdots & U_1^{(N)}(l) \\ U_2^{(1)}(l) & U_2^{(2)}(l) & \cdots & U_2^{(N)}(l) \\ \cdots & \cdots & \cdots & \cdots \\ U_N^{(1)}(l) & U_N^{(2)}(l) & \cdots & U_N^{(N)}(l) \end{bmatrix} \tag{4.2.3}$$

**Note**: The notation $G_k^m(l)$ calls for an explanation. To conform to the notation used in the Volterra kernels, the lower index is the output channel index, and the upper index always means the index of the input signal applied to the system.

As the MIMO system can be decoupled into $K$ MISO systems with the same inputs the general equation can be written for a single output channel (dropping also the redundant output index):

$$\mathbf{Y}(l)=\mathbf{G}(l)\mathbf{U}(l)+\mathbf{N}_Y(l)=\left[Y^{(1)}(l)\quad ...\quad Y^{(J)}(l)\right]=$$
$$\left[G^{1}(l)\quad ...\quad G^{N}(l)\right]\begin{bmatrix}U_1^{(1)}(l) & U_1^{(2)}(l) & \cdots & U_1^{(J)}(l)\\ U_2^{(1)}(l) & U_2^{(2)}(l) & \cdots & U_2^{(J)}(l)\\ \cdots & \cdots & \cdots & \cdots\\ U_N^{(1)}(l) & U_N^{(2)}(l) & \cdots & U_N^{(J)}(l)\end{bmatrix}+\left[N_Y^{(1)}(l)\quad ...\quad N_Y^{(J)}(l)\right] \qquad (4.2.4)$$

From now on **G** will be a 1×*N* matrix, and **Y**, $\mathbf{N}_Y$ are 1×*J* matrices accordingly. The FRF estimates $\hat{G}^i(l)$ can be computed from (4.2.4) as (from here on we will drop also the frequency argument when not explicitly required):

$$\hat{\mathbf{G}}=\mathbf{Y}\mathbf{U}^H(\mathbf{U}\mathbf{U}^H)^{-1}=\mathbf{G}+\mathbf{N}_Y\mathbf{U}^H(\mathbf{U}\mathbf{U}^H)^{-1}=\mathbf{G}+\mathbf{N}_Y\mathbf{B} \qquad (4.2.5)$$

where: $\mathbf{B}=\mathbf{U}^H(\mathbf{U}\mathbf{U}^H)^{-1}$, and $(\,)^H$ is the conjugate transpose. (4.2.6)

Taking $\tilde{\mathbf{G}}=\hat{\mathbf{G}}-\mathbf{G}$, (and assuming that $E\{\hat{\mathbf{G}}\}=E\{\mathbf{G}+\mathbf{N}_Y\mathbf{B}\}=\mathbf{G}$) the variances on the measured FRF-s can be computed from:

$$Cov\{\hat{\mathbf{G}}\}=E\{\tilde{\mathbf{G}}^H\tilde{\mathbf{G}}\}=E\{\mathbf{B}^H\mathbf{N}_Y^H\mathbf{N}_Y\mathbf{B}\} \qquad (4.2.7)$$

Matrix **B** will generally attenuate the distortion introduced by the output noise, its entries however could be complicated functions of the complex amplitudes of the input signals, and thus the noise smoothing effect of **B** will depend upon the particular choice of the inputs.

**Note**: Inverting the input matrix (direct, or pseudo inverse) is possible if all inputs are exciting at a given frequency (there are no zero rows in the input matrix). We have the following possibilities:

(a) every input is exciting at the frequency *l*;

(b) the measurement equation is considered only at the frequencies, where every input is exciting.

## 4.3 Problem of the input design – linear MIMO systems

The problem is now not only how to design the excitation at a particular input, but how to relate it to the other system inputs, i.e. instead of separately designing *U*(*l*) we should design a full matrix $\mathbf{U}=[U_k^{(n)}(l)]$ in (4.2.1), thinking globally about the all inputs and experiments. As a given column in **U** means the design of a single experiment, we should be free to change the input design from the experiment to the experiment.

The main problem can be best illustrated with the SISO FRF measurements, where the FRF estimate is (2.1.3) (2.2.24):

$$\hat{G}(l)=\frac{\sum_k Y_k(l)\overline{U_k(l)}}{\sum_k U_k(l)\overline{U_k(l)}}=\frac{\hat{S}_{YU}(l)}{\hat{S}_{UU}(l)} \qquad (4.3.1)$$

For the small number of averages the input auto spectrum estimate in the denominator can exhibit large variations and can take small values leading to the excessive variance of the FRF estimate. Such variations are present when the excitation signals are e.g. Gaussian noise or

periodic noise, and disappear when using random phase multisines, because then the denominator is deterministic:

$$\hat{S}_{UU}(l) = \sum_k U_k(l)\overline{U_k(l)} = \sum_k |U_k(l)|^2 = const \qquad (4.3.2)$$

In the MIMO case we can relate the inverse matrix in (4.2.5) to the denominator of (4.3.1). Intuitively if we use random signals as the excitations, the entries in (4.2.1) will be random, and in consequence the inverse in (4.2.5) will also fluctuate increasing the variance (4.2.7). Using random multisines won't entirely get rid of this problem in the MIMO case, because the randomness in the phases will still be visible in the off-diagonal entries in the matrix $\mathbf{U}\mathbf{U}^H$, consequently also in the inverse in (4.2.5). However the fluctuation will be less than for more randomized signals.

Linear MIMO FRF measurement theory developed input matrix design related to the minimization of the variances over the parametric FRF estimates [86, 87], [189-190]. Only when a suitable structure is enforced in the excitations applied to different inputs, i.e. into the input matrix **U**, we can expect a better behavior of the FRF estimates.

One of the possible approaches is through the uncertainty of the estimate (4.2.5), which can be characterized by an $N$-dim ellipsoid, which volume should be the smallest [125]. The smallest volume ellipsoid is related to the determinant of $\mathbf{U}\mathbf{U}^H$, and that to the determinant of **U** ($\det \mathbf{U}_N\mathbf{U}_N^H = (\det \mathbf{U}_N)^2$), which should be the largest and this brings into picture so called Hadamard maximum determinant problem.

**Note**: This is so called D-optimal design in the theory of linear measurements. There are more optimality criteria, like A-optimal design, etc. but the determinant based criterion is the most pertinent to the discussion [199, 125, 175].

**Definition 4.3.1**: **Hadamard maximum determinant problem**. The Hadamard maximum determinant problem [29] seeks complex $N \times N$ matrices **A** with entries in the unit disc satisfying the Hadamard bound $|\det \mathbf{A}| \leq N^{N/2}$. For dimensions $N = 2^K = 0 \ mod \ 4$ the real solutions are so called Hadamard matrices with entries ±1. For complex matrices the bound is always attained by the Vandermonde matrices of the $n$th roots of unity, i.e. the DFT matrices, which are defined for any dimension $N \neq 2^K$. ■

Hadamard matrices can be computed recursively via the Kronecker product as:

$$\mathbf{H}_{2^K} = \mathbf{H}_2 \otimes \mathbf{H}_{2^{K-1}}, \text{ where} \qquad (4.3.3)$$

$$\mathbf{H}_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \qquad (4.3.4)$$

The Hadamard matrix of e.g. order $N = 2^3 = 8$ is:

$$\mathbf{H}_8 = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \end{bmatrix} \qquad (4.3.5)$$

Hadamard matrices are also orthogonal, i.e.: $\mathbf{H}_N \mathbf{H}_N^H = \mathbf{H}_N^H \mathbf{H}_N = N\,\mathbf{I}_N$ (4.3.6)

where $\mathbf{I}_N$ is an $N \times N$ unit matrix.

Using Hadamard matrix to structure the MIMO excitation design for input dimension $N = 2^K = 0 \; mod \; 4$ would yield:

$\mathbf{U}_N(l) = \mathbf{H}_N \, U(l)$, and with this: (4.3.7)

$$\mathbf{B} = \mathbf{U}_N^H (\mathbf{U}_N \mathbf{U}_N^H)^{-1} = \frac{1}{N\overline{U}(l)} \mathbf{H}_N \,, \qquad (4.3.8)$$

$$\hat{\mathbf{G}} = \frac{1}{N\overline{U}(l)} \mathbf{Y}\mathbf{H}_N = \mathbf{G} + \frac{1}{N\overline{U}(l)} \mathbf{N}_Y \mathbf{H}_N \qquad (4.3.9)$$

assuming Gaussian white measurement noises:

$$Cov\{\hat{\mathbf{G}}\} = E\{\tilde{\mathbf{G}}^H \tilde{\mathbf{G}}\} = E\{\mathbf{B}^H \mathbf{N}_Y^H \mathbf{N}_Y \mathbf{B}\} = \frac{\sigma_Y^2}{N\,|U(l)|^2} \mathbf{I}_N \qquad (4.3.10)$$

**Note**: Input signal (matrix) designs in the linear and in the non-linear case essentially differ. Linear FRF estimates are input independent, only the output noise attenuation depends upon the inputs. In the non-linear case the measured FRF is only the Best Linear Approximation and is a kind of a least-square "linearization at the input point", i.e. theoretically input dependent.

**Note**: The principal drawback of using Hadamard input design is that Hadamard matrices are defined for input dimensions $N = 0$ mod 4. For other input dimensions a typical measurement practice is to approximate (4.3.7) with: $\mathbf{U}_M(l) = \hat{\mathbf{H}}_M \, U(l)$ , where for $M < 2^K$ , $\hat{\mathbf{H}}_M$ is the left upper $M \times M$ submatrix of $\mathbf{H}_{2^K}$ . The approximation yields of course correct solution for every $N = 2^K$. Note that due to the special structure in (4.3.7), the approximation will yield deterministic variance similar to (4.3.10) even for the approximated dimensions (with nonzero off-diagonal elements instead of diagonal $\mathbf{I}_N$) . To get a feeling of the phenomenon, consider:

$$\hat{\mathbf{H}}_3 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ -1 & -1 & 1 \end{bmatrix}, \qquad \hat{\mathbf{H}}_3 \hat{\mathbf{H}}_3^H = \begin{bmatrix} 3 & 1 & -1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{bmatrix} \qquad (4.3.11)$$

Comparison of the input designs based on the random multisines, the Hadamard matrices, the approximate Hadamard matrices, and the orthogonal matrices generalizing the idea of (4.3.3) designed for the non-linear systems (see Section 5.2.) can be found in [17*]. It was concluded, "… that for small system dimensions, in situations when no non-linear distortions are present and consequently when the Crest Factor minimization is allowed, the approximate orthogonal multisines yield the best results. For larger input dimensions, or when non-

linear distortions are present and the Crest Factor optimization won't be usually allowed, the ideal orthogonal multisines provide the best option."

## 4.4 Input design – non-linear effects in two-input two-output systems

The problem of an adequate input design to handle non-linear effects is introduced in case of simple Two-Input Two-Output (TITO) systems of low non-linear order, in noiseless measurement conditions, driven by the random phase multisines, defined on the same frequency grid. This particular choice is dictated by the theoretical importance of such models, e.g. in the microwaves, and also by the unexpected results, that in this case the linear FRF measurement technique is still fully functionable. The experimental setup is the special case of (4.2.4), for $J = N = 2$:

$$\mathbf{Y}(l) = \mathbf{G}(l)\mathbf{U}_2(l) = \left[Y^{(1)}(l) \quad Y^{(2)}(l)\right] = \left[G^1(l) \quad G^2(l)\right]\begin{bmatrix} U_1^{(1)}(l) & U_1^{(2)}(l) \\ U_2^{(1)}(l) & U_2^{(2)}(l) \end{bmatrix} \tag{4.4.1}$$

The required FRF estimates $\hat{G}^i(l)$ can be obtained simply as:

$$\hat{\mathbf{G}}(l) = \left[\hat{G}^1(l) \quad \hat{G}^2(l)\right] = \mathbf{Y}(l)\,\mathbf{U}^{-1}(l) \tag{4.4.2}$$

Now consider that the outputs **Y** belong to an up to 3rd order Volterra system:

$$\begin{aligned} Y(l) &= G^1(l)U_1(l) + G^2(l)U_2(l) + \\ &+ \sum_k G^{11}(k,L)U_1(k)U_1(L) + \sum_k G^{12}(k,L)U_1(k)U_2(L) + \sum_k G^{22}(k,L)U_2(k)U_2(L) + \\ &+ \sum_{k_1}\sum_{k_2} G^{111}(k_1,k_2,L)\,U_1(k_1)\,U_1(k_2)U_1(L) + \sum_{k_1}\sum_{k_2} G^{112}(k_1,k_2,L)\,U_1(k_1)\,U_1(k_2)U_2(L) + \\ &+ \sum_{k_1}\sum_{k_2} G^{122}(k_1,k_2,L)U_1(k_1)\,U_2(k_2)U_2(L) + \sum_{k_1}\sum_{k_2} G^{222}(k_1,k_2,L)U_2(k_1)U_2(k_2)U_2(L) \end{aligned} \tag{4.4.3}$$

where $L = l - k$ for the 2nd order, $L = l - k_1 - k_2$ for the 3rd order kernels. All sums run over the $S_M^- \cup S_M^+$ frequency grid, and $M$ is the number of harmonics in the input signal. In the general case the output of two experiments ($j = 1, 2$) is:

$$\begin{aligned} Y^{(j)}(l) &= G^1(l)U_1^{(j)}(l) + G^2(l)U_2^{(j)}(l) + \\ &\sum_k G^{11}(k,L)U_1^{(j)}(k)U_1^{(j)}(L) + \sum_k G^{12}(k,L)U_1^{(j)}(k)U_2^{(j)}(L) + \sum_k G^{22}(k,L)U_2^{(j)}(k)U_2^{(j)}(L) + \\ &\sum_{k_1}\sum_{k_2} G^{111}(k_1,k_2,L)U_1^{(j)}(k_1)U_1^{(j)}(k_2)U_1^{(j)}(L) + \sum_{k_1}\sum_{k_2} G^{112}(k_1,k_2,L)U_1^{(j)}(k_1)U_1^{(j)}(k_2)U_2^{(j)}(L) + \\ &\sum_{k_1}\sum_{k_2} G^{122}(k_1,k_2,L)U_1^{(j)}(k_1)U_2^{(j)}(k_2)U_2^{(j)}(L) + \sum_{k_1}\sum_{k_2} G^{222}(k_1,k_2,L)U_2^{(j)}(k_1)U_2^{(j)}(k_2)U_2^{(j)}(L) \end{aligned} \tag{4.4.4}$$

With: $$\mathbf{U}^{-1}(l) = \frac{1}{D(l)}\begin{bmatrix} U_2^{(2)}(l) & -U_1^{(2)}(l) \\ -U_2^{(1)}(l) & U_1^{(1)}(l) \end{bmatrix} \tag{4.4.5}$$

$$D(l) = \det\mathbf{U} = U_1^{(1)}(l)\,U_2^{(2)}(l) - U_1^{(2)}(l)\,U_2^{(1)}(l) \tag{4.4.6}$$

the FRF estimates from (4.4.2) are:

$$\hat{G}^1(l) = \frac{Y^{(1)}(l)U_2^{(2)}(l) - Y^{(2)}(l)U_2^{(1)}(l)}{D(l)} \tag{4.4.7}$$

$$\hat{G}^2(l) = \frac{-Y^{(1)}(l)U_1^{(2)}(l) + Y^{(2)}(l)U_1^{(1)}(l)}{D(l)} \tag{4.4.8}$$

Consider the $\hat{G}^1(l)$. Substituting the full expressions (4.4.4) into the estimate (4.4.7) yields:

$$\begin{aligned}
\hat{G}^1(l) = {} & G^1(l)\frac{U_1^{(1)}(l)\,U_2^{(2)}(l) - U_1^{(2)}(l)\,U_2^{(1)}(l)}{D(l)} + \\
& \sum_k G^{11}(k,L)\left(U_1^{(1)}(k)U_1^{(1)}(L)\frac{U_2^{(2)}(l)}{D(l)} - U_1^{(2)}(k)U_1^{(2)}(L)\frac{U_2^{(1)}(l)}{D(l)}\right) + \\
& \sum_k G^{12}(k,L)\left(U_1^{(1)}(k)U_2^{(1)}(L)\frac{U_2^{(2)}(l)}{D(l)} - U_1^{(2)}(k)U_2^{(2)}(L)\frac{U_2^{(1)}(l)}{D(l)}\right) + \\
& \sum_k G^{22}(k,L)\left(U_2^{(1)}(k)U_2^{(1)}(L)\frac{U_2^{(2)}(l)}{D(l)} - U_2^{(2)}(k)U_2^{(2)}(L)\frac{U_2^{(1)}(l)}{D(l)}\right) + \\
& \sum_{k_1}\sum_{k_2} G^{111}(k_1,k_2.L)\left(U_1^{(1)}(k_1)\,U_1^{(1)}(k_2)U_1^{(1)}(L)\frac{U_2^{(2)}(l)}{D(l)} - U_1^{(2)}(k_1)\,U_1^{(2)}(k_2)U_1^{(2)}(L)\frac{U_2^{(1)}(l)}{D(l)}\right) + \\
& \sum_{k_1}\sum_{k_2} G^{112}(k_1,k_2.L)\left(U_1^{(1)}(k_1)\,U_1^{(1)}(k_2)U_2^{(1)}(L)\frac{U_2^{(2)}(l)}{D(l)} - U_1^{(2)}(k_1)\,U_1^{(2)}(k_2)U_2^{(2)}(L)\frac{U_2^{(1)}(l)}{D(l)}\right) + \\
& \sum_{k_1}\sum_{k_2} G^{122}(k_1,k_2.L)\left(U_1^{(1)}(k_1)\,U_2^{(1)}(k_2)U_2^{(1)}(L)\frac{U_2^{(2)}(l)}{D(l)} - U_1^{(2)}(k_1)\,U_2^{(2)}(k_2)U_2^{(2)}(L)\frac{U_2^{(1)}(l)}{D(l)}\right) + \\
& \sum_{k_1}\sum_{k_2} G^{222}(k_1,k_2.L)\left(U_2^{(1)}(k_1)\,U_2^{(1)}(k_2)U_2^{(1)}(L)\frac{U_2^{(2)}(l)}{D(l)} - U_2^{(2)}(k_1)\,U_2^{(2)}(k_2)U_2^{(2)}(L)\frac{U_2^{(1)}(l)}{D(l)}\right)
\end{aligned} \tag{4.4.9}$$

The linear term cancels to $G^1(l)$. The measured FRF is the expected value of the $\hat{G}^1(l)$ (with respect to the random phases). The expected value of the terms in the parentheses will differ from 0 if suitable pairing of the frequencies can be found (see the Th. 2.2.1, consider also that different experiments and different input signals are statistically independent and that $|U_k^{(1)}|^2 = |U_k^{(2)}|^2$, $k = 1, 2$).

Consider e.g. the expression containing the kernel $G^{111}(k_1, k_2, L)$. Three frequency pairings $\{k = k_2 = -k_1, L = l\}$, $\{k = k_2 = -L_1, k_1 = l\}$, and $\{k = L = -k_1, k_2 = l\}$ are possible, yielding the nonzero expectation of:

$$E\left\{\sum_{k_1}\sum_{k_2} G^{111}(k_1,k_2,L)\left(U_1^{(1)}(k_1)U_1^{(1)}(k_2)U_1^{(1)}(L)\frac{U_2^{(2)}(l)}{D(l)} - U_1^{(2)}(k_1)U_1^{(2)}(k_2)U_1^{(2)}(L)\frac{U_2^{(1)}(l)}{D(l)}\right)\right\} =$$

$$3\sum_k G^{111}(l,k,-k)\left(U_1^{(1)}(k)U_1^{(1)}(-k)U_1^{(1)}(l)\frac{U_2^{(2)}(l)}{D(l)} - U_1^{(2)}(k)U_1^{(2)}(-k)U_1^{(2)}(l)\frac{U_2^{(1)}(l)}{D(l)}\right) =$$

$$3\sum_k G^{111}(l,k,-k)\left(|U_1^{(1)}(k)|^2\ U_1^{(1)}(l)\frac{U_2^{(2)}(l)}{D(l)} - |U_1^{(2)}(k)|^2\ U_1^{(2)}(l)\frac{U_2^{(1)}(l)}{D(l)}\right) =$$

$$3\sum_k G^{111}(l,k,-k)\,|U_1^{(1)}(k)|^2\,\frac{U_1^{(1)}(l)U_2^{(2)}(l) - U_1^{(2)}(l)U_2^{(1)}(l)}{D(l)} \quad + \quad O(M^{-1}) \tag{4.4.10}$$

Similar derivation but with less pairing is possible for $G^{122}$ kernel, and no nonzero pairing is possible at all for the other 3rd and 2nd order kernels:

$$\begin{aligned} E\{\hat{G}_1^1(l)\} &= G^1(l) + \\ &3\sum_k G_1^{111}(k,-k,l)\left|U_1^{(1)}(k)\right|^2\frac{U_1^{(1)}(l)U_2^{(2)}(l) - U_1^{(2)}(l)U_2^{(1)}(l)}{D(l)} \\ &+\sum_k G_1^{112}(k,-k,l)\left|U_1^{(1)}(k)\right|^2\frac{U_2^{(1)}(l)U_2^{(2)}(l) - U_2^{(2)}(l)U_2^{(1)}(l)}{D(l)} \\ &+\sum_k G_1^{122}(l,k,-k)\left|U_2^{(1)}(k)\right|^2\frac{U_1^{(1)}(l)U_2^{(2)}(l) - U_1^{(2)}(l)U_2^{(1)}(l)}{D(l)} \\ &+3\sum_k G_1^{222}(k,-k,l)\left|U_2^{(1)}(k)\right|^2\frac{U_2^{(1)}(l)U_2^{(2)}(l) - U_2^{(2)}(l)U_2^{(1)}(l)}{D(l)} \quad + O(M^{-1}) \end{aligned} \tag{4.4.11}$$

Within the 3rd order kernels the second and the fourth terms are zero, and the numerators in the first and the third terms equal to *D(l)*, leading finally to the nonzero expected value of (omitting for clarity the *M*-dependent asymptotic term):

$$\begin{aligned} &G_{BLA}^1(l) = E\{\hat{G}^1(l)\} = \\ &G^1(l) + 3\sum_k G^{111}(l,k,-k)\,|U_1(k|^2\ + \sum_k G^{122}(l,k,-k)\,|U_2(k|^2 = G^1(l) + G_B^1(l) \end{aligned} \tag{4.4.12}$$

The $\hat{G}^2(l)$ can be evaluated similarly as:

$$\begin{aligned} &G_{BLA}^2(l) = E\{\hat{G}^2(l)\} = \\ &G^2(l) + 3\sum_k G^{112}(l,k,-k)\,|U_1(k|^2\ + \sum_k G^{222}(l,k,-k)\,|U_2(k|^2 = G^2(l) + G_B^2(l) \end{aligned} \tag{4.4.13}$$

Zero mean terms (i.e. where the frequency pairing wasn't possible) belong to the non-linear noise $Y_S(l)$ - the noise observed over the measured FRF (see Th. 2.2.1 for more details). Considering that in (4.4.9) multiple terms (every kernel!) contribute to the noise and that the determinant (4.4.6) can be small, we should expect considerable non-linear noise and the lengthy averaging.

With this we have arrived to the linear representation of the TITO Volterra system as (introducing now the respective output indices):

$$\mathrm{Y}(l)=\mathrm{G}_{\mathbf{BLA}}(l)\mathrm{U}(l)+\mathrm{Y}_{\mathbf{S}}(l)=\begin{bmatrix}Y_1(l)\\ Y_2(l)\end{bmatrix}=\begin{bmatrix}G^1_{BLA,1} & G^2_{BLA,1}\\ G^1_{BLA,2} & G^2_{BLA,2}\end{bmatrix}\begin{bmatrix}U_1(l)\\ U_2(l)\end{bmatrix}+\begin{bmatrix}Y_{S,1}(l)\\ Y_{S,2}(l)\end{bmatrix} \tag{4.4.14}$$

where $\mathrm{G}_{\mathbf{BLA}}=\left[G^j_{BLA,i}\right]$ is the linear best approximation TITO system and $\boldsymbol{Y}_{\mathbf{S}}(l)$ is the non-linear noise. The particular investigated output can be written then as:

$$Y(l)=G^1_{BLA}(l)\,U_1(l)+G^2_{BLA}(l)\,U_2(l)+Y_S(l) \tag{4.4.15}$$

where the Best Linear Approximations: $G^k_{BLA}=E\{Y/U_k\}=G^k+G^k_B$, $k$=1, 2, are (as worked out in (2.2.13) for SISO systems) biased approximations to the non-linear relations described by Volterra series and $G^k_B$ are the bias terms introduced by the non-linearity. The equivalent noise source $Y_S(l)$ captures all other non-systematic effects.

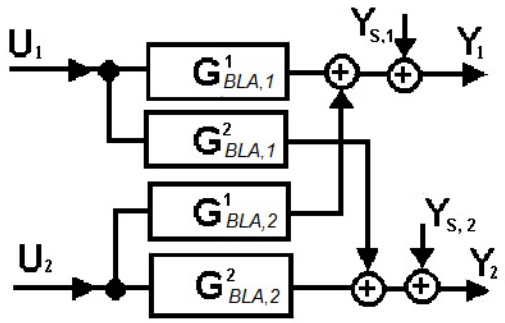


**Fig. 4.4.1**. Equivalent model of a non-linear TITO system. $G^k_{BLA,\,m}$ are the best linear approximation systems.

In the SISO theory we could see that the Best Linear Approximation takes on an especially simple form for the Wiener-Hammerstein systems. Assume then, that the computed 3rd order TITO Volterra model is really a Wiener-Hammerstein system.

**Example 4.4.1. Bias on the measured FRF of a 2-dim MISO Wiener-Hammerstein system**. The kernels (4.4.11) for the Wiener-Hammerstein model are particularly easy to compute:

$$G^{abc}(l,k,n)=\alpha_{122}R_{1a}(l)R_{1b}(k)R_{1c}(n)S_1(l+k+n) \tag{4.4.16}$$

With (4.4.16) the bias in (4.4.12-4.4.13) becomes:

$$\begin{aligned}G^1_{B,1}(l)&=3\sum_k G^{111}_1(l,k,-k)\left|U_1(k)\right|^2+\sum_k G^{122}_1(l,k,-k)\left|U_2(k)\right|^2=\\&=3\alpha_{111}R_{11}(l)S_1(l)\sum_k\left|R_{11}(k)\right|^2\left|U_1(k)\right|^2+\alpha_{122}R_{11}(l)\,S_1(l)\sum_k\left|R_{12}(k)\right|^2\left|U_2(k)\right|^2\\&=\alpha_{111}R_{11}(l)\,S_1(l)\,r_{11}+\alpha_{122}R_{11}(l)\,S_1(l)\,r_{12}=\left(\frac{3\alpha_{111}\,r_{11}}{\alpha_1}+\frac{\alpha_{122}\,r_{12}}{\alpha_1}\right)G^1_1(l)=K_1\,G^1_1(l)\end{aligned} \tag{4.4.17}$$

similarly for $G^2_{B,1}(l)$: $G^2_{B,1}(l)=\left(\frac{\alpha_{112}\,r_{11}}{\alpha_2}+\frac{3\alpha_{222}\,r_{12}}{\alpha_2}\right)G^2_1(l)=K_2\,G^2_1(l)$, (4.4.18)

because the input amplitudes are of order $O(N^{-1/2})$ and thus the sums $r_{11}$ and $r_{12}$ are of order $O(N^0)$. We can see that also in the MIMO case the relative bias remains constant for the Wiener-Hammerstein system (for general case see Section 4.6.) (the $M$-dependent asymptotic term has been omitted for clarity).

Let us turn to the orthogonal input design (4.3.7) proposed for linear MIMO FRF measurements. For dimension $N = 2$ the input matrix is:

$$\mathbf{U}(l) = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} U_1^{(1)}(l) \tag{4.4.19}$$

reapplying the same inputs in the second experiment, with sign reversed. The optimal choice (4.4.19) has been actually proposed for linear systems, yet it works extremely well also in this case. The common sense (and heuristically a powerful property of the Best Linear Approximation) is that the non-linear noise appears in the position of the output noise, thus techniques designed for the output noise should work also in this case. With (4.4.19):

$$\mathbf{U}^{-1}(l) = \frac{1}{2U_1^{(1)}(l)} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \tag{4.4.20}$$

and the FRF estimates are (c.f. with (4.4.7) and (4.4.8)):

$$\hat{G}^1(l) = \frac{Y_1^{(1)}(l) + Y_1^{(2)}(l)}{2U_1^{(1)}(l)}, \text{ and } \hat{G}^2(l) = \frac{Y_1^{(1)}(l) - Y_1^{(2)}(l)}{2U_1^{(1)}(l)} \tag{4.4.21}$$

Consider now the $\hat{G}^1(l)$. Substituting the full expressions of the outputs (4.4.4) into the equation of the FRF estimate (4.4.21) and taking into account that $U_1^{(1)}(l) = U_1^{(2)}(l) = U_2^{(1)}(l)$ and $U_2^{(2)}(l) = -U_1^{(1)}(l)$, yields:

$$\begin{aligned} \hat{G}^1(l) = G^1(l) + \sum_k G^{11} \frac{U_1^{(1)}(k)U_1^{(1)}(L)}{U_1^{(1)}(l)} + \sum_k G^{22} \frac{U_1^{(1)}(k)U_1^{(1)}(L)}{U_1^{(1)}(l)} + \\ \sum_{k_1}\sum_{k_2} G^{111} \frac{U_1^{(1)}(k_1)U_1^{(1)}(k_2)U_1^{(1)}(L)}{U_1^{(1)}(l)} + \sum_{k_1}\sum_{k_2} G^{122} \frac{U_1^{(1)}(k_1)U_1^{(1)}(k_2)U_1^{(1)}(L)}{U_1^{(1)}(l)} \end{aligned} \tag{4.4.22}$$

Terms with other kernels ($G^{12}, G^{112}, G^{222}$) are simply canceled out due to the change of sign in the input matrix. It is important to note, that:

- The expected value of (4.4.22) is exactly the same as that of (4.4.9) (nonzero mean frequency pairing is possible for 3rd order kernels, but not for 2nd order kernels).
- In the optimized case only four kernels visible in (4.4.22) will contribute to the non-linear variance, contrary to all(!) kernels contributing in the general case (c.f. (4.35)).

$\hat{G}^2(l)$ can be evaluated similarly.

**Example 4.4.2**. Non-linear FRF measurements using optimal linear input design.

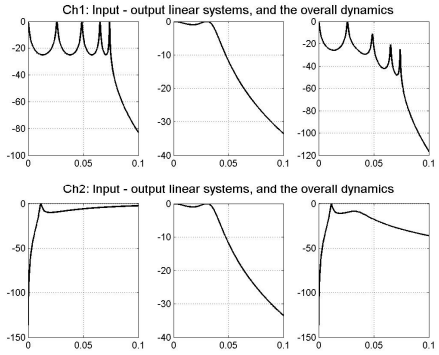


**Fig. 4.4.2.** The measured system is a 2-input MISO Wiener-Hammerstein system with the input dynamics, output dynamics, and overall dynamics in both channels shown in the figure. The static non-linearity *NL* is:

$x_p=z_1+z_1^2/2+z_1^3/5+z_2+z_2^2/2+z_2^3/5+z_1z_2/5+z_1^2z_2/2+z_1z_2^2/2.$

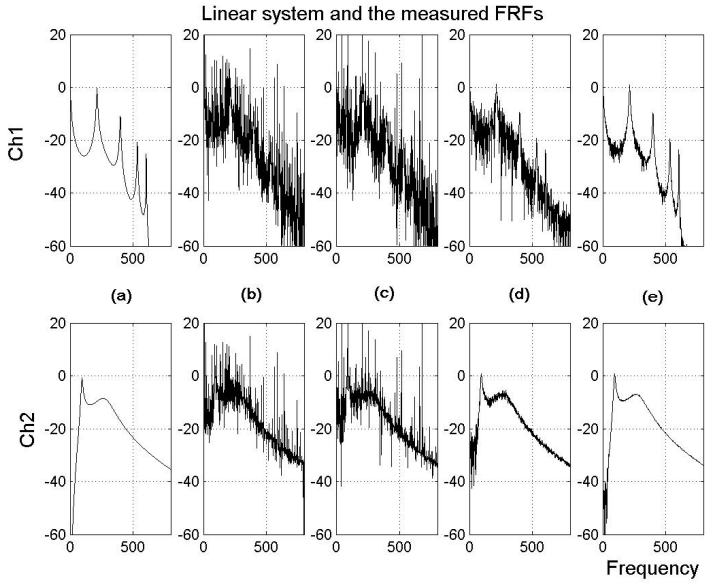


**Fig. 4.4.3.** The influence of various measurement setups on the measured FRF. (a) Linear systems for comparison. (b) FRF of a non-linearly distorted system measured with random multisines, without averaging, and (c) averaged from $M$ = 100 measurements, then (d) with orthogonal multisines, without averaging, and (e) averaged from $M$ = 100 measurements.

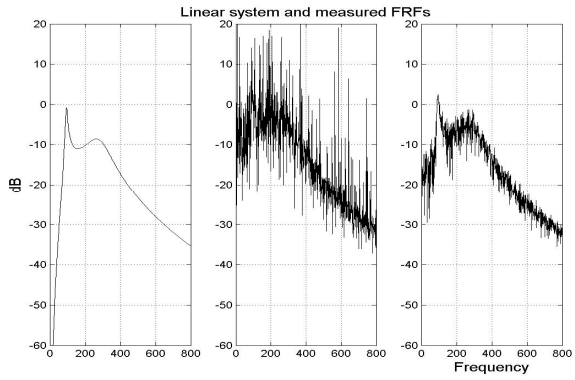


**Fig. 4.4.4.** The case of higher order non-linearity. All pure and mixed non-linearities up to the 5th order had been added to the static non-linearity *NL* in Fig. 4.4.2. Here we see the FRFs in Channel-2 (left), measured with general full grid multisines (middle) and with orthogonal multisines (right). The proposed method works well also in this case.

As expected from the derivations, optimizing the inputs (orthogonal multisines, (4.4.19)) yields a considerable gain in non-linear noise variance (almost 50 dB of difference, i.e. 300 times less averaging) with respect to the general case. We can conclude that when the Best Linear Approximation of a (weakly) non-linear Volterra TITO system is measured, it is profitable to used optimized orthogonal inputs (in addition to the odd frequency grid).

**Note**: The derivation shows that when (4.4.19) signal design is applied, the level of the variance is much lower, but the Best Linear Approximation to the cubic TITO Volterra system remains the same, which is an unexpected and positive result. In measuring non-linear systems one would normally expect that the measurement results depend strongly on the applied input signals. The same linear approximation with less noise makes the orthogonal inputs (4.4.19) a tool of choice for the TITO systems. We should add that when the output noise is

also present, the orthogonal input signals would tackle it as well. The question now is how much these results can be generalized to the MIMO Volterra series of arbitrary dimensions and arbitrary order of non-linearity?

## 4.5 Main results extended to MIMO (MISO) systems

We investigate now how the previous results can be extended to an arbitrary non-linear order Volterra MIMO system (4.1.1-4.1.2) (*N*-input MISO system).

In the MIMO measurements *N* experiments are made, using random multisines (2.2.10). As before the indices in parentheses indicate the serial number of the experiment. The upper index in the Volterra kernels lists the input signals belonging to the kernel. To prove the main theorem we will require a number of assumptions about the signals and the system:

**Assumption 4.5.1.** All multisines are defined on the same frequency grid. Signals at different inputs are independent from each other; their phases are independent over the frequency lines. ■

**Assumption 4.5.2.** MIMO system can be of arbitrary finite input dimension *N* and an arbitrary order of the non-linearity (assuming that the sums in (4.1.1)-(4.1.3) converge). ■

**Assumption 4.5.3.** Signals in different experiments are independent. ■

Then we have:

$$\mathbf{Y}(l)=\mathbf{G}(l)\mathbf{U}_N(l)=\left[Y^{(1)}(l)\quad ...\quad Y^{(N)}(l)\right]=\left[G^{1}(l)\quad ...\quad G^{N}(l)\right]\begin{bmatrix}U_1^{(1)}(l) & U_1^{(2)}(l) & \cdots & U_1^{(N)}(l)\\ U_2^{(1)}(l) & U_2^{(2)}(l) & \cdots & U_2^{(N)}(l)\\ \cdots & \cdots & \cdots & \cdots\\ U_N^{(1)}(l) & U_N^{(2)}(l) & \cdots & U_N^{(N)}(l)\end{bmatrix} \tag{4.5.1}$$

and the required FRF estimates $\hat{G}^{i}(l)$ can be computed as:

$$\hat{\mathbf{G}}(l)=\left[\hat{G}^{1}(l)\quad ...\quad \hat{G}^{N}(l)\right]=\mathbf{Y}(l)\mathbf{B}_N(l), \tag{4.5.2}$$

with $\mathbf{B}_N(l)=\left[b_{ij}\right]$ $N\times N$ any matrix fulfilling (e.g.: $\mathbf{B}_N(l)=\mathbf{U}_N^{-1}(l)$, or $\mathbf{B}_N(l)=\mathbf{U}_N^{H}(\mathbf{U}_N\mathbf{U}_N^{H})^{-1}$, etc.):

$$\mathbf{U}_N(l)\mathbf{B}_N(l)=\mathbf{I}_N \tag{4.5.3}$$

where $\mathbf{I}_N$ is the $N\times N$ unit matrix. In the following we will omit the frequency argument *l*, when unambiguous. First we will consider the general case, when: $\mathbf{B}_N(l)=\mathbf{U}_N^{-1}(l)$.

**Theorem 4.5.1: Bias on the general Volterra MIMO system.** When a Volterra MIMO system of arbitrary order is measured with random phase multisines (2.2.10) fulfilling the Assumptions 4.5.1 – 4.5.3, the bias on the Best Linear Approximation in the *kth* input signal channel is composed from the terms:

$$G_B^{j_1 j_2 \cdots j_\alpha}(l)=\sum_{k_1}\sum_{k_2}\cdots\sum_{k_{\alpha'}}G^{j_1 j_2 \cdots j_\alpha}(k_1,-k_1,...,l)\ \left|U_{j_{P1}}(k_1)\right|^2\cdots\left|U_{j_{P\alpha'}}(k_{\alpha'})\right|^2 \tag{4.5.4}$$

where $G^{j_1 j_2 \cdots j_\alpha}$ are all those odd order Volterra kernels, where the input signal belonging to the investigated channel appears an odd number of times (linear kernels are the special case of this). The overall bias is:

$$G_{\boldsymbol{B}}(l) = \sum_{\alpha=3,odd}^{\infty} G_{\boldsymbol{B}}^{\alpha}(l) = \sum_{\alpha=3,odd}^{\infty} \sum_{j_1 j_2 \cdots j_\alpha} G_B^{j_1 j_2 \cdots j_\alpha}(l) \tag{4.5.5}$$

**Proof: in Appendix A.3** ■

Please also note that the Theorem 2.2.1. and the bias expression for the SISO system (2.2.4) are special cases of the Theorem 4.5.1. (i.e. other inputs are not present, the input appears an odd number of times, consequently all odd and only odd non-linearities contribute to the bias in the SISO case).

$$\begin{aligned} &\sum_{k_1}\sum_{k_2}\cdots\sum_{k_{\alpha'}} G^{j_1 j_2 \cdots j_\alpha}(k_1,-k_1,\ldots,l)\ \left|U_{j_{P1}}(k_1)\right|^2 \cdots \left|U_{j_{P\alpha'}}(k_{\alpha'})\right| \Big|_{j_1=j_2=\cdots j_\alpha} = \\ &\quad \sum_{k_1}\sum_{k_2}\cdots\sum_{k_{\alpha'}} G^{\alpha}(k_1,-k_1,\ldots,l) \prod_{n=1}^{\alpha'} |U(k_n)|^2, \quad \alpha' = (\alpha-1)/2 \end{aligned} \tag{4.5.6}$$

**Note**: It is important to observe, that the proof of the Th. 4.5.1. did not require the assumption on the uniformity of the spectral amplitudes, consequently the Th. 4.5.1. and the bias expression is valid for the (normalized) multisine inputs independently colored over the input channels.

**Theorem 4.5.2: The Best Linear Approximation model of a MIMO Volterra system.** When the inputs of an arbitrary Volterra MIMO system are excited with multisine signals (2.2.10), the system outputs can be written as:

$$\begin{aligned} \mathbf{Y}(l) &= \left[Y_m(l)\right] = \mathbf{G}_{BLA}(l)\mathbf{U}(l) + \mathbf{Y_S}(l) = \left[G_{BLA,m}^{k}(l)\right]\left[U_k(l)\right] + \left[Y_{S,m}(l)\right] = \\ &= \left[G_m^k(l)\right]\left[U_k(l)\right] + \left[G_{B,m}^k(l)\right]\left[U_k(l)\right] + \left[Y_{S,m}(l)\right] \end{aligned} \tag{4.5.7}$$

where the systems: $G_{BLA,l}^{k} = E\{Y_l/U_k\} = G_l^k + G_{B,l}^k$ are biased Best Linear Approximations to the non-linear relations described by Volterra series and $G_{B,l}^k$ are the biases introduced by the non-linearity (note that the output index had been added for clarity). The equivalent noise sources $Y_{S,m}(l)$ capture other nonsystematic non-linear effects. This yields the additive non-linear noise model for the MIMO system, a straightforward extension to the SISO and TITO cases.

**Proof:** Zero mean terms in (A.3.8) (i.e. where frequency pairing wasn't possible) are the non-linear noises $Y_{S,k}(l)$, or $G_{S,k}(l) = Y_{S,k}(l)/U_k(l)$ the noises observed over the measured FRF. From this it is straightforward to get (4.5.7). Please note that multiple terms (every kernel!) contribute to the noise and that as the determinant det **U** can be closed to 0, we should expect considerable amount of the non-linear noise. ■

## 4.6 Properties of the BLA approximation

The construction (4.5.4), analogous to the SISO case (Th 2.2.6), the measurement friendly properties of the MIMO Volterra system under Assumptions 4.1.1-4.1.5 and 4.5.1-4.5.3 permit to conjecture that the MIMO BLA model has similarly nice properties.

**Conjecture 4.6.1: Properties of the MIMO BLA bias and non-linear variance.** Under similar conditions on the kernels as in Def 2.2.1-2.2.2, the MIMO BLA bias and non-linear variance are smooth functions in the frequency, are continuous in the input spectral amplitudes, and are robust with respect to the frequency grid of the input signals. ■

**Theorem 4.6.1: The Best Linear Approximation of the MIMO Wiener-Hammerstein system (Def 4.1.3) has constant relative bias.**

**Proof**: In the MIMO Volterra measurements (with random multisines) the bias on the Best Linear Approximation in the *kth* input signal channel is composed from terms like:

$$G_B^{j_1 j_2 \cdots j_\alpha}(l) = \sum_{k_1}\sum_{k_2}\cdots\sum_{k_{\alpha'}} G^{j_1 j_2 \cdots j_\alpha}(k_1, -k_1, \ldots, l)\ \left|U_{j_{P1}}(k_1)\right|^2 \cdots \left|U_{j_{P\alpha'}}(k_{\alpha'})\right|^2 \qquad (4.5.8)$$

$G^{j_1 j_2 \cdots j_\alpha}$-s are those odd order kernels, where the input signal in the investigated channel appears an odd number of times (linear kernels as special case). Kernels of order $\alpha$ of *N*-input MISO Wiener-Hammerstein system are:

$$G^{j_1 j_2 \cdots j_\alpha}(k_1, k_2, \cdots, k_\alpha) = const \times S_p(k_1 + k_2 + \cdots k_\alpha) \prod_{k=1}^{\alpha} R_{p j_k}(k_n) \qquad (4.5.9)$$

Consequently, after substitution and introducing canceling frequencies, the bias in the kth channel will be:

$$G_B^{j_1 j_2 \cdots j_\alpha}(l) = const \times \prod_{j=1}^{\alpha'} \sum_{k_j} \left|R_{p j}(k_j)\right|^2 \left|U_j(k_j)\right|^2 S_p(l)\, R_{pk}(l)$$

which with the normalization of the inputs yields: $G_B^k(l) = c_k S_p(l) R_{pk}(l) = c_k^* G^k(l)$ ■

The question we tackle now is could the input signals be designed in some special way, to make the BLA measurements even more advantageous? The seed of the idea is the successful application of the optimal input signal design for the 2-dim linear FRF measurements, applied to the 2-dim 3rd order Volterra systems. Is it a solution that could be generalized to the higher dimensions and nonlinear orders, or not? We seek now the answer to this problem.

# 5. Multisine excitations for MIMO measurements

MIMO measurement setup presents additional degrees of freedom to the design of the input signals. Besides focusing on the design of the amplitudes, phases, and frequencies of a particular signal at the particular input, we can ask questions of how these parameters should be related to the similar parameters set at other input channels. We could see already that applying independently the random multisines to system inputs yields extension to the MIMO case of the Best Linear Approximation devised for the SISO systems (Th. 2.2.1). We could also see that the introduction of a structure to the input matrix (4.3.7) can provide considerable gains not only when measuring linear FRF-s, but also in case of the non-linear distortions (4.4.21). Our question now is whether we can improve (Th.4.5.1–Th.4.5.2) in the following sense:

- the measured Best Linear Approximation should be the same (“equivalence”), but
- the non-linear noise variance should possibly be less, to make the measurements faster, and the collected data of better quality (“optimality”).

## 5.1 MIMO multisine design

Can the results obtained for TITO systems (4.4.22) be generalized to the full MIMO case? The problem is that the orthogonalization of the inputs and using only single random amplitude for all of the inputs and experiments (albeit with different weighting) can introduce constraints, which influence the bias. The answer is that it is indeed the case.

The traditional input matrix (4.2.2) built from random excitations is not a good choice. Even in the absence of disturbing output noise the FRF measurements will vary from one realization to the other. The reason is the fluctuation of the inverse matrix in (4.2.5), due to the randomness of the excitations. Note, that the matrix $(\mathbf{U}\mathbf{U}^H)^{-1}$ fluctuates even when random phase multisines are used, contrary to the measurements on SISO systems [162].

**Example 5.1.1**. Fluctuation of the input matrix for the random multisine measurements. We illustrate the problem with the simplest 2-dim case:

$$\mathbf{U}_2(l)=\begin{bmatrix} U_1^{(1)}(l) & U_1^{(2)}(l) \\ U_2^{(1)}(l) & U_2^{(2)}(l) \end{bmatrix} \tag{5.1.1}$$

Omitting for brevity the frequency index we have:

$$\mathbf{U}_2\mathbf{U}_2^H=\begin{bmatrix} D_1 & S \\ \overline{S} & D_2 \end{bmatrix} \tag{5.1.2}$$

where $D_\mathrm{k} = |U_\mathrm{k}^{(1)}|^2 + |U_\mathrm{k}^{(2)}|^2$ is deterministic, and off-diagonal $S = U_1^{(1)}\bar{U}_2^{(1)} + U_1^{(2)}\bar{U}_2^{(2)}$ is a zero mean random term. Consequently:

$$(\mathbf{U}_2\mathbf{U}_2^H)^{-1}=\frac{1}{D_1D_2-|S|^2}\begin{bmatrix} D_2 & -S \\ -\overline{S} & D_1 \end{bmatrix} \tag{5.1.3}$$

and the random off-diagonal terms introduce additional fluctuation in the FRF measurements. Note also the

expression of the determinant in the denominator. Its small values can further amplify the variance.

Generally the variance on the measured FRF is a function of the used input signals and the particular composition of the non-linear part of the system. In the SISO case the expected ranking is $Var\{\hat{G}\}\big|_{\substack{Gaussian\\noise}} > Var\{\hat{G}\}\big|_{\substack{periodic\\noise}} > Var\{\hat{G}\}\big|_{\substack{random\ phase\\multisine}}$, because $\hat{S}_{UU}(l) = \sum_{e=1}^{J} U^{(e)}(l)\overline{U}^{(e)}(l)$ in (2.2.24) (4.3.1) randomly fluctuates for the noises, but is deterministic for the (random phase) multisines.

In the MIMO case this wisdom won't be valid, because as mentioned earlier $\hat{\mathbf{S}}_{UU}^{-1}(l) = (\mathbf{U}_N \mathbf{U}_N^H)^{-1}$ in (4.2.5) will contain random components for all three considered inputs. For static non-linearities there is no leakage and variances will be comparable. For dynamic non-linear systems the leakage of the Gaussian noise will increase the variance comparing to the periodic input signals. In the comparison of the periodic noise and the random phase multisine periodic noise turns out a bit better, because its input matrix is better conditioned.

**Example 5.2.1. Condition number of the periodic noise and the random phase multisines**. Consider for illustration and simplicity the case $N = 2$ and let us assume that signals are independent over the channels, and that the amplitudes for the periodic noise are similarly distributed in different input channels, with $E\{A_{km}\} = A = 1$.

Let: $\mathbf{U}_2 = \begin{bmatrix} e^{j\varphi_{11}} & e^{j\varphi_{12}} \\ e^{j\varphi_{21}} & e^{j\varphi_{22}} \end{bmatrix}$ and $\mathbf{P}_2 = \begin{bmatrix} A_{11}e^{j\varphi_{11}} & A_{12}e^{j\varphi_{12}} \\ A_{21}e^{j\varphi_{21}} & A_{22}e^{j\varphi_{22}} \end{bmatrix}$ represent the input matrix $\begin{bmatrix} U_1^{(1)} & U_1^{(1)} \\ U_2^{(1)} & U_2^{(2)} \end{bmatrix}$ for the random phase multisines and the periodic noise respectively, where all the phases are uniformly distributed on the unit circle and the amplitudes $A_{km}$ are exponentially distributed with unit expected value. All the considered random variables are independent.

The condition number equals $\kappa(\mathbf{U}) = \|\mathbf{U}\|\,\|\mathbf{U}^{-1}\|$ [84], and let choose for the investigation the Frobenius norm, i.e.: $\|\mathbf{U}\|^2 = \Sigma_{km} |u_{km}|^2$. In this simple case the respective inverse matrices are:

$$\mathbf{U}_2^{-1} = \frac{1}{D(\mathbf{U}_2)} \begin{bmatrix} e^{j\varphi_{22}} & -e^{j\varphi_{12}} \\ -e^{j\varphi_{21}} & e^{j\varphi_{11}} \end{bmatrix} \text{ and } \mathbf{P}_2^{-1} = \frac{1}{D(\mathbf{P}_2)} \begin{bmatrix} A_{22}e^{j\varphi_{22}} & -A_{12}e^{j\varphi_{12}} \\ -A_{21}e^{j\varphi_{21}} & A_{11}e^{j\varphi_{11}} \end{bmatrix}, \tag{5.1.4}$$

with the determinants:

$$D(\mathbf{U}_2) = e^{j\varphi_{11}}e^{j\varphi_{22}} - e^{j\varphi_{12}}e^{j\varphi_{21}} \text{ and } D(\mathbf{P}_2) = A_{11}e^{j\varphi_{11}}A_{22}e^{j\varphi_{22}} - A_{12}e^{j\varphi_{12}}A_{21}e^{j\varphi_{21}}.$$

The condition numbers become then:

$$\kappa^2(\mathbf{U}_2) = \|\mathbf{U}_2\|^2 \|\mathbf{U}_2^{-1}\|^2 = 4 \Big/ \left| e^{j\varphi_{11}}e^{j\varphi_{22}} - e^{j\varphi_{12}}e^{j\varphi_{21}} \right| = 4 \Big/ \left| e^{j\xi_1} - e^{j\xi_2} \right|, \text{ and} \tag{5.1.5}$$

$$\kappa^2(\mathbf{P}_2) = \|\mathbf{P}_2\|^2 \|\mathbf{P}_2^{-1}\|^2 = \sum_{k,m} A_{km}^2 \Big/ \left| B_1 e^{j\xi_1} - B_2 e^{j\xi_2} \right| \tag{5.1.6}$$

where the phases $\xi_k$ are uniformly distributed and independent, and $B_k$-s are independent products of independent, exponentially distributed variables. Clearly for (5.1.6) to be singular not only the phazors must be colineated as for (5.1.5), but also the random amplitudes should match, which is an event of lower probability than the singularity of (5.1.5). On the other hand (5.1.6) can be excessively large without the colineation of the phases, simply when the amplitudes are small. Simulations show that the average condition number for the periodic noise is a bit better (simulations indicate a rough factor of 2, not really a difference).

If an input matrix built from random multisines does not work well, then what? Hadamard matrix makes the experiments the simplest; no computation is required to find the amplitudes for the new experiments. Using Hadamard matrix for the input matrix (4.2.3) makes it possible to Crest Factor optimize the input signals once for all, because the subsequent change of sign in an input channel does not influence the phases of the already optimized signal.

Another issue to consider is that although only Hadamard matrix is proposed in the literature to minimize the noise influence [88], it can be used with full impact only for the system with $N = 2^K$ inputs. Approximate design fares already not so well for linear systems, and is dubious when the Best Linear Approximation comes into question. One could use in (4.3.7) another orthogonal (or unitary) matrix, e.g. the Fourier matrix (DFT matrix), defined for any dimension (i.e. number of inputs). This choice however introduces already additional computation to the amplitudes.

First we will present a negative result telling that (4.4.20-4.4.22) cannot be generalized fully. This will lead to the introduction of a new input matrix design for which the equivalence and the optimality will be proved in the general case.

**Theorem 5.1.1: Input matrix design (4.3.7) does not generalize to the general MIMO Volterra system of arbitrary dimension and order**. When a Volterra MIMO system (Def. 4.1.1) of arbitrary order, with a number of inputs $N > 2$, is measured with random orthogonal inputs (2.2.10), with Hadamard or Fourier matrix, acc. to (4.3.7), the bias on the Best Linear Approximation (in the *kth* signal channel), besides terms mentioned in Theorem 4.5.1 will also contain some (not many) other terms as well.

**Proof: In Appendix A.4 ■**

**Example 5.1.1**: Some of the kernels adding to the bias, beside the "normal" bias terms (4.5.4) from Th. 4.5.1.:

$N = 4$, $\alpha = 3$, $k = 1$, kernel $(j_1 j_2 j_3) = (234)$ for both Fourier and Hadamard matrix,

$k = 2$, kernel $(j_1 j_2 j_3) = (123)$ for Fourier matrix,

$k = 2$, kernel $(j_1 j_2 j_3) = (134)$ for both Fourier and Hadamard matrix,

$k = 3$, kernel $(j_1 j_2 j_3) = (124)$ for both Fourier and Hadamard matrix,

$k = 4$, kernel $(j_1 j_2 j_3) = (233)$ for Fourier matrix, etc.

Generally it can be seen that Fourier matrix introduces more bias terms, than the Hadamard matrix. On the other hand Fourier matrix can be applied for an arbitrary (e.g. odd) number of inputs, but the Hadamard matrix cannot.

**Note:** Using orthogonal inputs amounts to a modified increased bias. Beside this all the evaluation leading to the additive non-linear noise source model remains valid; consequently in this case we also have the additive noise model, albeit with different Best Linear Approximation and non-linear variance.

**Note:** Although direct analytic comparison between the influence of the Fourier and the Hadamard cases (with respect to the level of the non-linear variance) does not seem realistic (different algebraic operations involved), low order calculations show that they eliminate roughly similar number of (similar) kernels, consequently, if the number of inputs permits, Hadamard matrix could be proposed as a simpler one for the calculations.

## 5.2 Orthogonal random multisines

Using arbitrary random phase multisines (i.e. independent phases for every input and experiment) yields too much of the variance. Restricting inputs to orthogonal combinations of the first channel-first experiment input (A.4.1) yields too much of the constraints. We would like then to be sure, as a designing principle, that **the bias is always the same and the variance is always less, than in the general input case**. To this end we should have a situation, where:

- $V$ in (A.3.8) has a form similar to (A.4.5), i.e. $V = A$ x $B$,
- $A = 0$ for a number of index combinations (non-linear kernels), to have less variance,
- $E\{B\}$ should be nonzero only where the investigated input appears an odd number of times, and other inputs an even number of times (the usual bias term, Th. 4.5.1).

Let us assume that $J = N_B$ x $N$ experiments are made ($N_B$ is the number of blocks in the experiments) and let partition the $N \times J$ input matrix $\mathbf{U}$ in (4.2.4) into $N_B$ rectangular blocks as: $\mathbf{U} = [\mathbf{U}_N\ \mathbf{U}_N \ldots \mathbf{U}_N]$ and $\mathbf{Y}$ in the similar way.

Instead of the general input design:

$$\mathbf{U}_N(l) = \begin{bmatrix} U_1^{(1)}(l) & U_1^{(2)}(l) & \ldots & U_1^{(N)}(l) \\ U_2^{(1)}(l) & U_2^{(2)}(l) & \ldots & U_2^{(N)}(l) \\ \ldots & \ldots & \ldots & \ldots \\ U_N^{(1)}(l) & U_N^{(2)}(l) & \ldots & U_N^{(N)}(l) \end{bmatrix}, \tag{5.2.1}$$

which requires independent excitations for every input and every experiment, I propose to use:

$$\mathbf{U}_N(l) = \begin{bmatrix} w_{11}U_1^{(1)}(l) & w_{12}U_1^{(1)}(l) & \ldots & w_{1N}U_1^{(1)}(l) \\ w_{21}U_2^{(1)}(l) & w_{22}U_2^{(1)}(l) & \ldots & w_{2N}U_2^{(1)}(l) \\ \ldots & \ldots & \ldots & \ldots \\ w_{N1}U_N^{(1)}(l) & w_{N2}U_N^{(1)}(l) & \ldots & w_{NN}U_N^{(1)}(l) \end{bmatrix} = diag\left\{U_k^{(1)}(l)\right\} \begin{bmatrix} w_{11} & \ldots & w_{1N} \\ \ldots & \ldots & \ldots \\ w_{N1} & \ldots & w_{NN} \end{bmatrix} = \mathbf{D}_U \mathbf{W} \tag{5.2.2}$$

where $w_{kj}$ are entries of an arbitrary, deterministic unitary (orthogonal) matrix:

$$\mathbf{W}^H \mathbf{W} = \mathbf{W}\mathbf{W}^H = N\,\mathbf{I} \tag{5.2.3}$$

e.g. the DFT matrix, with $[\mathbf{W}]_{kn} = e^{-j\,2\pi(k\text{-}1)\,(n\text{-}1)\,/\,N}$. Such signals will be called the **orthogonal multisines.**

The essence of this measurement design is that we use independent (phase) inputs for the first experiment and then for the next experiments we simply reuse them weighting with entries of a unitary coefficient matrix. Such input matrix is also unitary, if the amplitudes of various inputs are the same, i.e. if: $|U_k^{(1)}(l)|^2 = const,\ \forall\, k$.

**Example 5.4.** For $N$ = 4 and Hadamard coefficients the inputs are:

$$\mathbf{U}_N(l)=\begin{bmatrix} U_1(l) & U_2(l) & U_3(l) & U_4(l) \\ U_1(l) & -U_2(l) & U_3(l) & -U_4(l) \\ U_1(l) & U_2(l) & -U_3(l) & -U_4(l) \\ U_1(l) & -U_2(l) & -U_3(l) & U_4(l) \end{bmatrix} \tag{5.2.4}$$

For $N$ = 3 and Fourier coefficients the inputs are:

$$\mathbf{U}_N(l)=\begin{bmatrix} U_1(l) & U_2(l) & U_3(l) \\ U_1(l) & e^{j\frac{2\pi}{3}}U_2(l) & e^{-j\frac{2\pi}{3}}U_3(l) \\ U_1(l) & e^{-j\frac{2\pi}{3}}U_2(l) & e^{j\frac{2\pi}{3}}U_3(l) \end{bmatrix} \tag{5.2.5}$$

The measurement procedure is thus to generate random excitations for the first experiment in the block of the first $N$ experiments and to shift them orthogonally, accordingly to (5.2.2) for the next $N$-1 experiments, then to start with another random choice for the next block.

It is easy to see, that due to the unitary matrix **W** (the experiment index is omitted for clarity):

$$\mathbf{U}_N(l)\mathbf{U}_N^H(l)=\left[\overline{U}_k(l)U_j(l)\sum_{i=1}^{N}\overline{w}_{ik}w_{ij}\right]= \quad N\left[\overline{U}_k(l)U_j(l)\delta_{kj}\right]=N\,diag\{|U_k(l)|^2\}, \tag{5.2.6}$$

where $\delta_{kj}$ is the Kronecker Delta, is a simple amplitude scaling. Furthermore:

$$(\mathbf{U}\mathbf{U}^H)^{-1}=(N_B\mathbf{U}_N\mathbf{U}_N^H)^{-1} \text{ and finally:} \tag{5.2.7}$$

$$\hat{\mathbf{G}}(l)=\frac{1}{N_B}\sum_{i=1}^{N_B}\hat{\mathbf{G}}_{N,i}(l), \text{ with:} \tag{5.2.8}$$

$$\hat{\mathbf{G}}_{N,i}(l)=\mathbf{Y}_N(l)\mathbf{U}_N^H(l)(\mathbf{U}_N(l)\mathbf{U}_N^H(l))^{-1}=\frac{1}{N}diag\{|U_k|^{-2}\}\mathbf{Y}_N(l)\mathbf{U}_N^H(l) \tag{5.2.9}$$

computed without taking the inverse from one block of $N$ equations.

Because $\hat{\mathbf{S}}_{UU}^{-1}(l)$ gets rid of the random fluctuations (for advantageous condition number see below), a reasonable drop in variance should be expected, at least only for that reason.

**Lemma 5.2.1. Condition number of the orthogonal multisines.** Orthogonal multisines (5.2.2) with uniform amplitudes have relative condition number = 1, for any unitary matrix W built from the roots of unity and for equal input spectra at different inputs.

**Proof:** For simplicity we will omit the frequency and the experiment indices:

$$\mathbf{U}_N\mathbf{U}_N^H=\begin{bmatrix} w_{11}U_1 & w_{12}U_1 & \dots & w_{1N}U_1 \\ w_{21}U_2 & w_{22}U_2 & \dots & w_{2N}U_2 \\ \dots & \dots & \dots & \dots \\ w_{N1}U_N & w_{N2}U_N & \dots & w_{NN}U_N \end{bmatrix}\times\begin{bmatrix} \overline{w}_{11}\overline{U}_1 & \overline{w}_{21}\overline{U}_2 & \dots & \overline{w}_{N1}\overline{U}_N \\ \overline{w}_{12}U_1 & \overline{w}_{22}U_2 & \dots & \overline{w}_{N2}\overline{U}_N \\ \dots & \dots & \dots & \dots \\ \overline{w}_{1N}U_1 & \overline{w}_{2N}U_2 & \dots & \overline{w}_{NN}\overline{U}_N \end{bmatrix} \tag{5.2.10}$$

Then from (5.2.3):

$$[\mathbf{U}_N\mathbf{U}_N^H]_{kj} = \left[U_k\overline{U}_j\sum_{i=1}^{N} w_{ki}\overline{w}_{ji}\right]_{kj} = U_k\overline{U}_j\left[\mathbf{W}\mathbf{W}^H\right]_{kj} = N\,\delta_{kj}\,U_k\overline{U}_j \tag{5.2.11}$$

where $\delta_{kj}$ is Kronecker Delta.

Now: $\mathbf{U}_N\mathbf{U}_N^H = N\,\mathbf{D}$, and $\mathbf{D} = diag\{|U_k|^2\}$, and $\mathbf{U}_N^H = N\,\mathbf{U}_N^{-1}\mathbf{D}$, and $\mathbf{U}_N^{-1} = \frac{1}{N}\,\mathbf{U}_N^H\mathbf{D}^{-1}$ (5.2.12)

$$\mathbf{U}_N^{-1} = \frac{1}{N}\mathbf{U}_N^H \times diag\{\frac{1}{|U_k|^2}\} = \frac{1}{N}\begin{bmatrix} \overline{w}_{11}\frac{\overline{U}_1}{|U_1|^2} & \dots & \overline{w}_{N1}\frac{\overline{U}_N}{|U_N|^2} \\ \dots & \dots & \dots \\ \overline{w}_{1N}\frac{\overline{U}_1}{|U_1|^2} & \dots & \overline{w}_{NN}\frac{\overline{U}_N}{|U_N|^2} \end{bmatrix} \tag{5.2.13}$$

With the choice of the Frobenius matrix norm:

$$\|\,\mathbf{U}_N\,\|_F^2 = \sum_{k=1}^{N}\sum_{j=1}^{N}|\,u_{kj}\,|^2 = \sum_{k=1}^{N}|\,U_k\,|^2\sum_{j=1}^{N}|\,w_{kj}\,|^2 = N\sum_{k=1}^{N}|\,U_k\,|^2 \tag{5.2.14}$$

Similarly: $\|\,\mathbf{U}_N^{-1}\,\|_F^2 = \ldots = \frac{1}{N^2}\sum_{k=1}^{N}|\,U_k\,|^{-2}\sum_{j=1}^{N}|\,w_{kj}\,|^2 = \frac{1}{N}\sum_{k=1}^{N}|\,U_k\,|^{-2}$ (5.2.15)

The condition number becomes [84]:

$$\kappa(\mathbf{U}_N) = \|\,\mathbf{U}_N\,\|_F\|\,\mathbf{U}_N^{-1}\,\|_F = \sqrt{\sum_{k=1}^{N}|\,U_k\,|^2\sum_{k=1}^{N}|\,U_k\,|^{-2}}\,, \tag{5.2.16}$$

If the amplitudes are all equal, then $\kappa(\mathbf{U}_N) = N$, and $\kappa(\mathbf{U}_N)/N = 1$. The condition number becomes worse of course when the input amplitude levels are not equal. ■

**Example 5.2.1**: Determinant of $\mathbf{U}_N$.

Due to (5.2.2): $\det(\mathbf{U}_N\mathbf{U}_N^H) = \det(\mathbf{U}_N)\det(\mathbf{U}_N^H) = |\det(\mathbf{U}_H)|^2 = N\prod_{k=1}^{N}|\,U_k\,|^2$ (5.2.17)

Thus: $\det(\mathbf{U}_N) = e^{-j\theta}\sqrt{N}\prod_{k=1}^{N}|\,U_k\,|$, with some random $\theta \in [0, 2\pi[$. (5.2.18)

**Theorem 5.2.2: The Best Linear Approximation with orthogonal multisines**. When a Volterra MIMO system of arbitrary order, with inputs $N > 2$, is measured with random orthogonal multisines (5.2.2), the bias on the Best Linear Approximation (in the *kth* signal channel) equals to that from the Theorem 4.5.1, the additive non-linear noise source model from the Theorem 4.5.2 remains valid. The only difference is that the non-linear variance in case of the orthogonal inputs is lower, than in the general case, due to the deterministic inverse $(\mathbf{U}_N\mathbf{U}_N^H)^{-1}$ and the cancellation of a number of kernels.

**Proof:** The main part of the required results had been derived already in the proof of the Theorem 5.1.1 (see App. A.4). With matrix (5.2.2) the random component of the kernel (A.3.5-A.3.6) appears now as (cf. A.4.5):

$$V = \underbrace{\left(\sum_{i=1}^{N} w_{ij_1}w_{ij_2}\ldots w_{ij_\alpha}\,\overline{w}_{ik}\right)}_{A}\underbrace{\left(U_{j_1}(k_1)U_{j_2}(k_2)\ldots\overline{U}_k(l)\right)}_{B} \tag{5.2.19}$$

(the upper index of the first experiment is omitted for clarity). For such matrix coefficient *A* behaves exactly as before, and *B* also yields exactly the same bias, as for the general inputs, because pairing the frequencies in inputs with different indices still retains the randomness of the phases and yields zero expected value (viz. (A.4.12):

$$E\{U_j^{(i)}(k)U_k^{(i)}(-k)\} = \frac{1}{N\,|U|^2}E\{w_{ij}U_j^{(1)}(k)\overline{w}_{ik}\,\overline{U}_k^{(1)}(k)\} = \frac{1}{N\,|U|^2}w_{ij}\,\overline{w}_{ik}\,E\{U_j^{(1)}(k)\,\overline{U}_k^{(1)}(k)\} = 0 \quad , j \neq k \tag{5.2.20}$$

■

**Note:** As mentioned before, if for a particular choice of the orthogonal matrix and the combination of the input indices in the kernel, it turns out, that $A = \sum_{i=1}^{N} w_{ij_1} w_{ij_2} \ldots w_{ij_\alpha}\,\overline{w}_{ik} = 0$, this term will drop out and won't contribute to the variance at all. The exact quantification how much we can gain in the variance from the kernel drop out is possible only if *N*, a, and **W** are known.

**Example 5.2.2: Measurement on MIMO Wiener-Hammerstein system with general and orthogonal inputs.** Simulated measurements had been performed on a Wiener-Hammerstein MISO system with 4 inputs, containing pure and mixed non-linear kernels up to the 5th order (see Fig. 5.2.1) (Large ripples were chosen in simulation to provide resemblance to the FRF of highly resonant mechanical systems). Results for all channels are visible in Figs. 5.2.2 and 5.2.3.

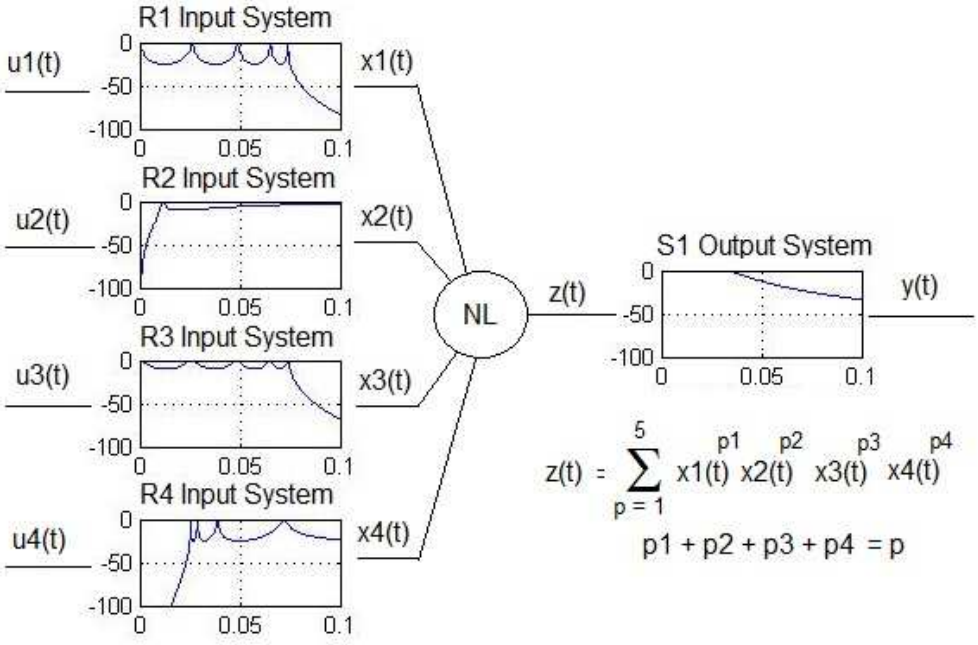


**Fig. 5.2.1.** MISO Wiener-Hammerstein system used in the simulations: R1 is 9th order low-pass Chebyshev filter with 25 db ripple, R2 is 3rd order high-pass Chebyshev filter with 10 dB ripple, R3 is 9th order low-pass Chebyshev filter with 10 dB ripple, R4 is 9th order high-pass Chebyshev filter with 25 dB ripple, and S1 is 3rd order low-pass Chebyshev filter with 1 dB ripple. Proportion of the non-linear power in the channels of order from -5dB … -10 dB.

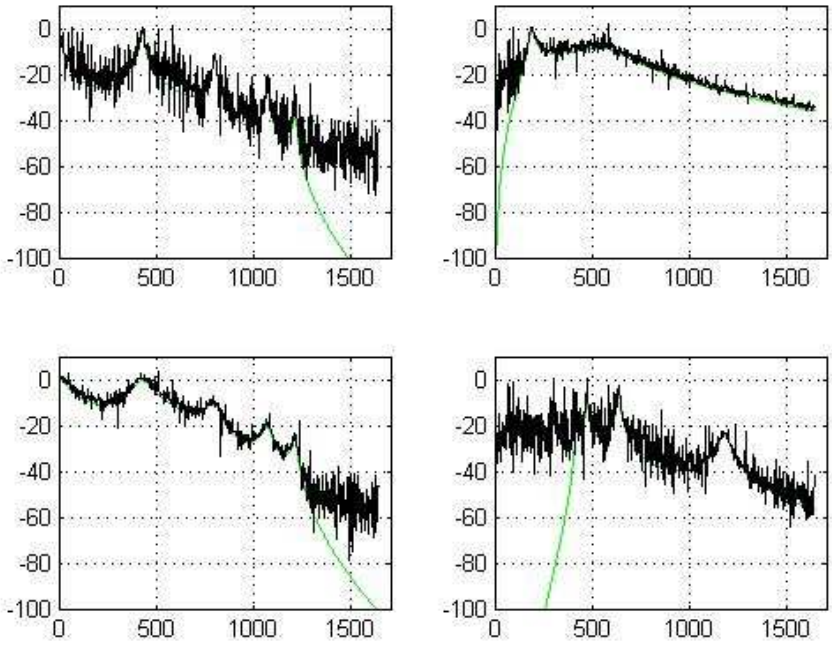


**Fig. 5.2.2**. FRF BLA measurements with a single application of the ordinary odd random phase multisine ($M$ = 819). Input-Output Channel 1: upper left, Channel 2: upper right, Channel 3: lower left, and Channel 4: lower right.

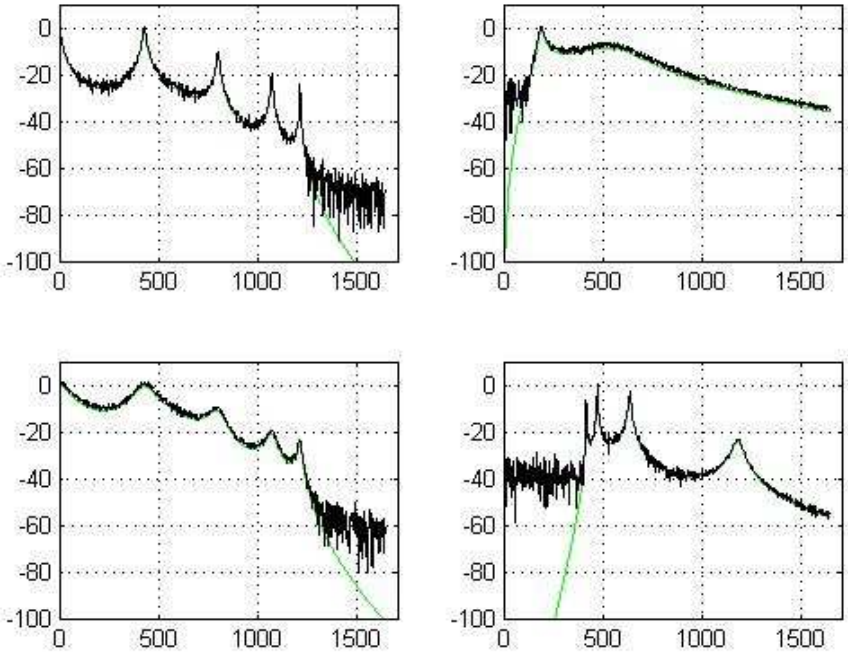


**Fig. 5.2.3**. FRF BLA measurements with a single application of the orthogonal odd random phase multisine ($M$ = 819 harmonics, 4-dim DFT matrix as the orthogonal coefficient matrix). Input-Output Channel 1: upper left, Channel 2: upper right, Channel 3: lower left, and Channel 4: lower right.

## 5.3 Other developments

Minimizing Crest Factor with time-domain constraints was used in the input signal design for linear SISO and MIMO measurements in [183-186]. The primary criterion was the Crest Factor minimization with a variant of the $L_\infty$ algorithm [85], with so called "plant-friendly" constraints on the maximum permissible time-domain amplitudes and moves (rates of change) in the inputs (or also jointly with the outputs). To account for the presence of the non-linearities Crest Factor minimization with dropped even harmonics was considered, to suppress even order non-linear effects. Accordingly to the advice in [85] "snowing effect" was intensively used [183-186] at both sides of the primary excitation band.

**Note**: Contrary to the claims, the design is not useful for the non-linear measurements as it is heavily optimized for a particular situation. If odd order non-linearities are present, then "snowing" the side bands "snows in" distortions into the primary band beside those interharmonic distortion already caused by the primary frequencies themselves. As the "snowing" is set by the optimizing algorithm and outside the reach of the user, the amount of the non-linear distortions on the primary excitations frequencies is difficult to quantify.

**Note**: This multisine design is oriented toward parametric identification, where the number of useful harmonics in the input needs to be higher (but not excessively higher), than the number of the parameters in the model [124]. Proper setting of the constraints requires plenty of information about the plant (like e.g. the desired closed-loop speed of response, estimated range of dominant time constant, order and structure of the model to be fit, acceptable signal length and amplitude, etc.) [28].

An extension to the constrained MIMO multisine design was the orthogonalization introduced by "zippering" the signal spectra [122], i.e. placing the spectrum of every input signal (otherwise a multisine optimized with constraints as before) on a different frequency grid, arranged alternately as Ch1, Ch2, Ch1, Ch2, … Signals with spectra having no common support are by definition mutually orthogonal, and the construction can be extended to an arbitrary number of inputs.

**Example 5.3.1**: If even harmonic are suppressed, then e.g. for the 4-dim systems the input signals will have the following grids in the primary excitation frequency band:

Ch1: 1 9 17 25 ….

Ch2: 3 11 19 27 …

Ch3: 5 13 21 29 ….

Ch4: 7 15 23 31 …. etc.

**Note**: Although signals at different channels are mutually orthogonal, as multisines (for higher input dimension systems) they are quite sparse. If the primary band is not densely filled up with the harmonics, the persistency at various inputs can be a problem. There is also the lost advantage of not having the basic harmonic in every frequency grid.

Pure orthogonal "zipper" multisines turned out not so useful, and to help the problem correlation was introduced into the inputs for large spectral amplitudes in the primary frequency band [119-122]. Such modified "zippered" multisines had frequency grids with mutually orthogonal sub-grids (smaller amplitude excitations) and a common grid (for larger excitations to introduce correlation into the signals).

So called directional multisine input design was done via manipulating amplitudes and phases to colineate the input signal with one of the SVD vectors of the gain matrix, selecting thus the low gain input direction [122]. Still another idea was to design signals at various inputs as delayed version of the signal (multisine) applied to the first channel [121].

**Note**: For random phase multisines the delay is not an issue. It does not destroy the properties of the random phases.

Another development in the constrained optimization of the multisine inputs is the optimization-based design of plant-friendly multisine signals using geometric discrepancy criteria for uniformly distributed sequences [142, 185]. The aim is to design excitations for which the output of the system will fill the state-space with its behaviour approximately uniformly, e.g. that e.g. the output state vector $[y(t)\ y(t-\tau)]$ is uniformly distributed in some 2-dim domain (the idea can be generalized to higher dimensions, with certain computational difficulties). To design such signals so called Weyl-criterion [143] is added to the criteria applied to the Crest Factor and the "plant-friendliness". As the Weyl-criterion can be verified only approximately for finite domains; also the overall optimum is sought up to a small error level.

## 5.4 MIMO equivalence of the random multisine excitations

For the portability of the theoretical results it is not enough to show that we measure exactly the same Best Linear Approximation for various multisine signals. Much more far reaching result is to show that these measurements are equivalent (in terms of the Best Linear Approximation) to the measurements made with the more traditional methods. We are able that way to present an (faster, cheaper, more precise, etc.) alternative to the measurement community, without instilling fear that the new results won't be compatible with the already gained experimental insight. In [163] it was shown that the random phase multisine measurements are in this sense equivalent to the periodic and Gaussian noise excitations (Th. 3.6.1). Here we prove the MIMO analogue of this theorem, extended also to the case of the newly introduced orthogonal multisines.

Assume that the signals have the following comparable spectral behavior:

- random phase multisines (2.2.10): $\hat{U}_k^{\,2}(f) = S_{\hat{U}\hat{U}}(f)$, (5.4.1)

- periodic noise with random spectral amplitudes: $E\left\{\hat{U}_k^2(f)\right\} = S_{\hat{U}\hat{U}}(f)$, (5.4.2)

- Gaussian noise with power spectrum: $S_{UU}(f) = S_{\hat{U}\hat{U}}(f) / f_{\max}$. (5.4.3)

Extending theory from SISO to MIMO systems requires assumptions how signals at different inputs are related to each other. We will require that (cf. Assumptions 4.5.1-4.5.3):

**Assumption 5.4.1.** Signals at different inputs are independent from each other, their phases and (in case of the periodic noise) spectral amplitudes are independent over the frequency. Signals have comparable spectral powers (5.4.1)-(5.4.3) and are defined on the same frequency grid. ■

**Assumption 5.4.2.** MIMO system can be of arbitrary input dimension *N* and an arbitrary order of the non-linearity (assuming that the sums in (4.1.1)-(4.1.3) converge). ■

**Assumption 5.4.3.** Signals in different experiments are independent. ■

As introduced before *J* independent experiments are made with independent realizations of the input signals $U^{(1)},\cdots U^{(J)}$. After the transients settle, the successive records to process are cut from the input and output signals. Signal amplitudes at frequency *l* are then arranged into:

$$\mathbf{Y}(l)=\mathbf{G}(l)\mathbf{U}(l)=\left[Y^{(1)}(l)\quad ...\quad Y^{(J)}(l)\right]=\left[G^{1}(l)\quad ...\quad G^{N}(l)\right]\begin{bmatrix}U_1^{(1)}(l) & ... & U_1^{(J)}(l)\\ ... & U_j^{(i)}(l) & ...\\ U_N^{(1)}(l) & ... & U_N^{(J)}(l)\end{bmatrix} \quad (5.4.4)$$

The required FRF estimates $\hat{G}^{i}(l)$ can be computed as:

$$\hat{\mathbf{G}}(l)=\left[\hat{G}^{i}(l)\right]=\mathbf{Y}(l)\mathbf{U}^{H}(l)(\mathbf{U}(l)\mathbf{U}^{H}(l))^{-1}=\hat{\mathbf{S}}_{YU}(l)\hat{\mathbf{S}}_{UU}^{-1}(l) \quad (5.4.5)$$

where $(\,)^{H}$ is the conjugate transpose.

By proving the Best Linear Approximation equivalence we actually also imply that in the limit ($M\rightarrow\infty$) the output of the Volterra MIMO system, excited by the above mentioned types of input signals, can be written for all these classes of the excitations as:

$$\begin{aligned}\mathbf{Y}(l)&=\mathbf{G}_{\mathbf{BLA}}(l)\mathbf{U}(l)+\mathbf{Y}_{\mathbf{S}}(l)=\mathbf{G}(l)\mathbf{U}(l)+\mathbf{G}_{\mathbf{B}}(l)\mathbf{U}(l)+\mathbf{Y}_{\mathbf{S}}(l)=\\ &\left[G_{BLA,m}^{k}(l)\right]\left[U_k(l)\right]+\left[Y_{S,m}(l)\right]=\left[G_m^{k}(l)\right]\left[U_k(l)\right]+\left[G_{B,m}^{k}(l)\right]\left[U_k(l)\right]+\left[Y_{S,m}(l)\right]\end{aligned} \quad (5.4.6)$$

where the measured quantities are the Best Linear Approximations $G_{BLA,m}^{k}=E\{Y_m/U_k\}=G_m^{k}+G_{B,m}^{k}$ (with expected value taken with respect to the random phases), to the non-linear relations described by the multidimensional Volterra series in the signal path $Y_m-U_k$, and $G_{B,m}^{k}$ are biases on the linear FRF introduced by the non-linearity. The equivalent noise sources $Y_{S,m}(l)$, $E\{Y_{S,m}(l)\}=0$ capture all the nonsystematic non-linear effects (Section 2.2, Th. 2.2.6), [30*].

Equation (5.4.6) yields thus the additive non-linear noise source model for MIMO systems, a straightforward extension to the SISO and Two-Input Two-Output (TITO) cases.

In the following the index of the output will be omitted, because we investigate essentially a MISO system. The index *k* of the measured signal path is called the '*reference input index*'.

The results on the equivalence can be collected in the following theorems:

**Theorem 5.4.1: Equivalence of the excitations I.** Under Assumptions 5.4.1.-5.4.3., with the input signals normalized to the same spectral behavior (5.4.1)-(5.4.3), all of the mentioned signal classes, i.e. the periodic noise, the random phase multisines in the limit $M\rightarrow\infty$, and the Gaussian noise, yield exactly the same linear approximation to a non-linear MIMO system, described by a multidimensional Volterra series (4.1.1)-(4.1.4). Kernels with nonzero expected value (with respect to the random parameters of the excitation signals) contributing to the bias

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.

**Proof: In Appendix A.5** ■

**Note**: For the illustration consider that e.g. in the signal path with the reference input of index '1' kernels:

$G^{111}, G^{122}, G^{12233}$, … etc. contribute to the bias, but kernels: $G^{12}, G^{233}, G^{1234}$, … etc. will not.

Considering that the orthogonal random phase multisines yield the same Best Linear Approximation as the normal random phase multisines, the Th. 5.4.1 can be immediately extended to:

**Theorem 5.4.2: Equivalence of excitations II.** Under Assumptions 5.4.1. – 5.4.2. the orthogonal random phase multisines defined by (5.2.2) are equivalent to all signals specified in the Th. 5.4.1. The orthogonal random phase multisines, when suitably normalized to the same spectral behavior (5.4.1) and in the limit $M \to \infty$, yield exactly the same Best Linear Approximation $G_{BLA}$. The presence of the orthogonal entries $w_{ij}$ combined within the kernels leads to three possible behaviors of the zero mean (stochastic) kernel contributions:

a. The cumulative effect of the entries $w_{ij}$ is nonzero and frequency independent. Such kernels contribute to the non-linear variance fully.

b. The cumulative effect of the entries $w_{ij}$ is zero and frequency independent. Such kernels drop out entirely from the non-linear variance.

c. The cumulative effect of the entries $w_{ij}$ is nonzero, but frequency dependent. Such kernels contribute to the non-linear variance in part and at particular frequencies only.

**Proof**: Due to (5.2.6-5.2.9) (and the expectation) it is enough to show the equivalence of the FRF measured for a single block of data (J = N). In that case:

$$G_{B,k}^{j_1 \dots j_K}(l) = E\{\sum_{n=1}^{N} b_{kn}(l)\, Y^{j_1 \dots j_K\,(n)}(l)\} = \sum_{k_1,\dots,k_{\alpha-1}} G^{j_1 \dots j_K} E\{\sum_{n=1}^{N} U_{j_1}^{(n)}(k_1) \dots U_{j_K}^{(n)}(k_\alpha)\; b_{kn}(l)\} \qquad (5.4.7)$$

which with (by definition, see also (5.2.2)): $U_j^{(n)}(l) = w_{nj} U_j(l)$ (5.4.8)

$$b_{kn}(l) = \frac{\overline{w}_{nk}\, \overline{U}_k(l)}{N\, |U_k(l)|^2} \qquad (5.4.9)$$

can be written as:

$$G_{B,k}^{j_1 \dots j_K}(l) = A \sum_{k_1,\dots,k_{\alpha-1}} G^{j_1 \dots j_K}(k_1,\dots,k_\alpha) \frac{1}{|U_k(l)|^2} E\{U_{j_1}(k_1) \dots U_{j_K}(k_\alpha)\; \overline{U}_k(l)\} \qquad (5.4.10)$$

For the expected value to be nonzero exactly the same conditions on the inputs are required as before (i.e. the reference input present odd number of times, other inputs present even number of times).

Coefficient *A* represents dependency of the bias on the choice of the particular unitary or orthogonal matrix **W**:

$$A = \frac{1}{N} \sum_{n=1}^{N} v_{nj_1} v_{nj_2} \dots v_{nj_\alpha} \overline{v}_{nk} \qquad (5.4.11)$$

$$v_{nk} = \begin{cases} w_{nk}, & k > 0 \\ \overline{w}_{nk}, & k < 0 \end{cases} \tag{5.4.12}$$

Considering, that: $U_j^{(n)}(-l) = \overline{U}_j^{(n)}(l) = \overline{w}_{nj}\overline{U}_j(l)$ (5.4.13)

consequently pairing the frequencies, which introduces complex conjugate to the signal amplitudes, will perform conjugation also on entries $w_{kn}$. For frequency pairings leading to the nonzero expected value in (5.4.10):

$$A = \frac{1}{N}\sum_{n=1}^{N} | w_{nj_1} |^2 \, | w_{nj_2} |^2 \ldots | w_{nk} |^2 = \frac{1}{N}\sum_{n=1}^{N} 1 = 1 \tag{5.4.14}$$

due to $|w_{nk}|^2 = 1$, and the bias again coincides with (A.5.17).

The value of (5.4.11) depends naturally on the choice of the orthogonal matrix **W**, on the indices of the inputs in the kernel, on the reference signal index of the measured signal path and on the frequency pairing introducing complex conjugate for the negative frequencies. Three cases can be distinguished in general for zero expected value kernels:

a. $A = 1$ for all frequencies (as in (5.4.14)). Such kernel contributes fully to the non-linear variance on the FRF.

b. $A = 0$ for all frequencies (if e.g. $A$ is reduced by the properties of orthogonal entries to $A = \frac{1}{N}\sum_{n=1}^{N} w_{np}$ for some $p \neq 1$. Such kernel drops out (does not contribute to) from the variance.

c. $A = 0$ only for some frequencies (when complex conjugate leads at a particular frequency to suitable reduction in the product of the entries in (5.4.11)). Such kernel contributes to the variance at those frequencies only. ■

**Note**: The orthogonal multisines will generate less variance because:

(1) due to (5.2.6) they do not introduce random fluctuations in the denominator of the estimate;

(2) due to (cases a., b., and c. in the proof) they eliminate some of the non-linear kernels from the non-linear stochastic component of (4.1.4).

**Example 5.4.1.:** A comparison is made of the non-linear variance levels measured in a 3-dim MISO system, excited with Gaussian noise, periodic noise, random phase multisines, and orthogonal random phase multisines accordingly. Matrix **W** is the DFT matrix. All input signals are scaled to unit power and in the measurement $N_B$ = 1, 2, 5, 10 number of blocks were used (note that now $J = 3N_B$). The MISO system has a Wiener-Hammerstein structure shown in Fig. 5.4.1. In the first simulation the static non-linearity contains all mixed powers up to the 5th order:

$$NL = x_1 + x_2 + x_2 + 10^{-2}\sum_{i,\,j,\,k=0..5} x_1^i\, x_2^j\, x_3^k, \qquad 1 < i + j + k \le 5 \tag{5.4.15}$$

and was designed to show the general situation with a weakly non-linear system. As mentioned before, no other noise sources are considered. The variance depends solely on the frequency and the signal channel.

Excitations that randomize the inverse in (4.2.5) show the expected rapid decrease in the variance for small $J$. For a higher number of data all signals tend to the same limit.

Variances produced with the orthogonal random phase multisines can be even lower due to the drop-out effect of some kernels. This effect can be seen amplified in Fig. 5.4.4., which presents variances measured in case of:

$$NL = x_1 + x_2 + x_2 + x_1 x_2 + 10^{-2}\sum_{i,\,j=0..5} x_1^i\, x_2^j\, x_3^k, \qquad 1 < i + j + k \le 5 \tag{5.4.16}$$

The strong non-linear kernel $x_1x_2$ drops out entirely from the $G^1$ measurements (Th. 5.4.2., case a.), appears fully in $G^2$ measurements (case b.), and partly in the $G^3$ measurements (case c.). Please note that the drop-out effect is not influenced by the number of data, only by the structure of the kernels of the measured MIMO system.

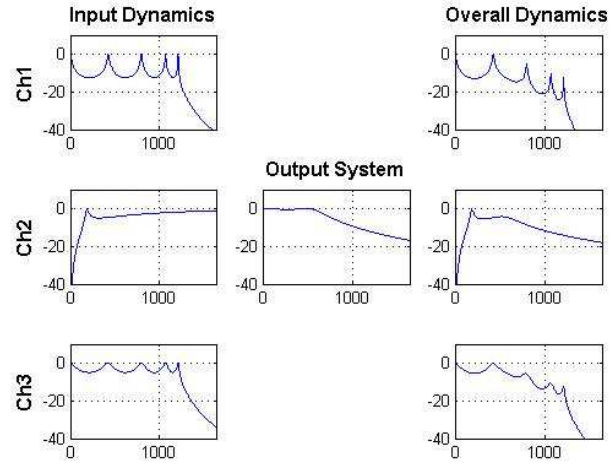


**Fig. 5.4.1** Linear dynamics of 3-dim Wiener-Hammerstein system used in the simulations.

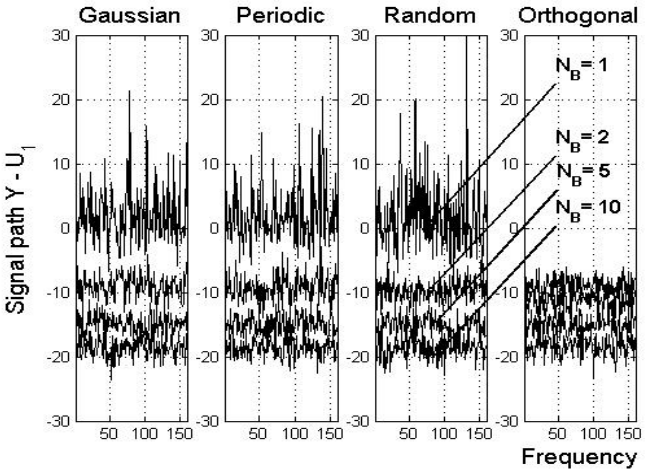


**Fig. 5.4.2** Variances (in dB) of the FRF measured in the signal path $Y$-$U_1$, for the static non-linear system (5.4.15), for $N_B$=1, 2, 5, 10. The decreasing levels of the variance are clearly visible for each kind of signals.

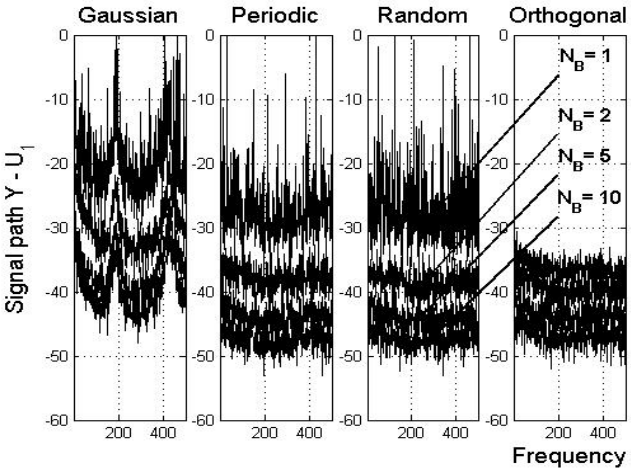


**Fig. 5.4.3** Variances (in dB) of the FRF measured in the signal path $Y$-$U_1$, for a Wiener-Hammerstein system composed from the dynamics in Fig. 5.4.1 and the static non-linearity (5.4.15), for $N_B$ =1, 2, 5, 10. The appearance of the leakage elevates the variance of the Gaussian measurements. Note also that the system dynamics influence the frequency behavior of the variance.

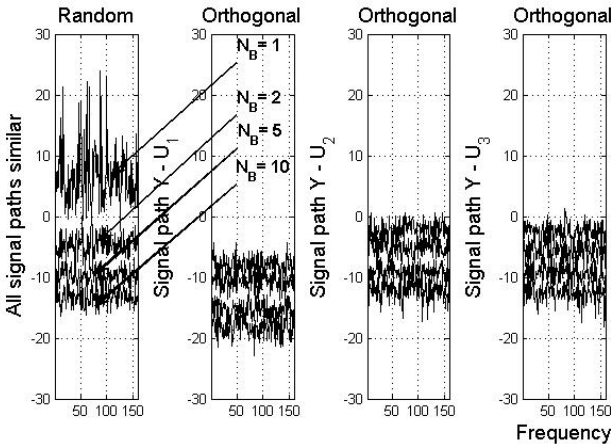


**Fig. 5.4.4** Variances (in dB) of the FRF measured for a Wiener-Hammerstein system composed from the dynamics in Fig. 5.4.1. and the static non-linearity (5.4.16), for $N_B$=1, 2, 5, 10, for random and orthogonal multisines. Due to interactions between non-linear kernels and the unitary matrix **W**, kernels usually generating the variance, when measured with random multisines, can drop-out in particular channels (here the kernel $x_1x_2$ does not affect the variance measured in the channel $Y–U_1$) when measured with orthogonal multisines.

# 6. Special applications

Modeling non-linear distortions with a bias on the linear FRF and the associated non-linear noise can be used alone as a well defined and flexible measurement technique, especially if the advantages of the periodic excitations will be amplified by the opportunities provided by the frequency grids. The approach can be used also as a part of methodology to tackle more involved modeling problems.

## 6.1 Non-linear distortion in cascaded SISO systems

The FRF $Q$ of two cascaded linear systems with the FRF $G^1$ and $H^1$ is given by the product $Q(l) = G^1(l)\, H^1(l)$. It is tempting to generalize this rule to the Best Linear Approximations of weakly non-linear systems, especially if the experiments indicate that passing random multisine excitation through a weakly non-linear Volterra system does not destroy much of their essential properties needed to measure the BLA [7*].

To get a feel the investigated problem will be limited to the uniform random phase multisines and to the SISO Wiener-Hammerstein systems, with cubic static non-linearity: $s = a_1 r + a_3 r^3$ ($a_2$ is set to 0, because even non-linearities do not contribute to the BLA measurements).

The general case is computationally too involved, and the cubic system constitutes an important special case. Furthermore the general case can be qualitatively extrapolated from this special case. We will consider weakly non-linear systems, i.e.: $\left\|a_3 r^3\right\|_2 / \left\|a_1 r\right\|_2 \leq \varepsilon$, with no other additive output measurement noises.

For the approximation of the cascaded Wiener-Hammerstein systems it will be shown that in those frequency bands where the coherence (2.5.2) is high, the cascaded systems can be approximated by the product of their Best Linear Approximations. In the other bands (where the linear input system attenuates the signal) such an approximation is useless because the stochastic output $Y_S(l)$ is much larger than the linear output $G_R(l)\, U(l)$ (see Figs. 6.1.1-6.1.2).

Let us investigate now the situation when two (cubic) Wiener-Hammerstein systems are cascaded, like in Fig. 6.1.1.

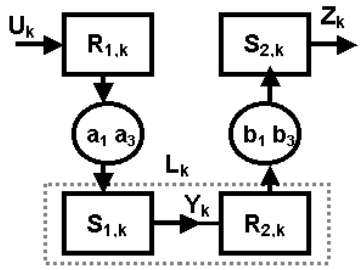


**Fig. 6.1.1** Cascade of two cubic Wiener-Hammerstein systems, with: $NL_1 = a_1 x + a_3 x^3$ and $NL_2 = b_1 x + b_3 x^3$.

Assuming small levels of non-linearity we will simplify the evaluation further, dropping out terms of order $\varepsilon^3$ and higher. The output of the cascade can be written (2.2.3-2.3.4) as:

$$Z(l) = Z^1(l) + Z^3(l) = G^1(l)Y(l) + \sum_{z_1=-M}^{M} \sum_{z_2=-M}^{M} G^{111}(z_1, z_2, L_z)\, Y(z_1)\, Y(z_2)\, Y(L_z) \qquad (6.1.1)$$

$$Y(l) = Y^1(l) + Y^3(l) = H^1(l)U(l) + \sum_{k_1=-M}^{M} \sum_{k_2=-M}^{M} H^{111}(k_1, k_2, L_k)U(k_1)U(k_2)U(L_k) \qquad (6.1.2)$$

Multiplication of three or more cubic terms means contribution of order higher than $\varepsilon^2$ and such terms will be omitted from further consideration. We have:

**Theorem 6.1.1**: **The Best Linear Approximation FRF of the cascade of two cubic Wiener-Hammerstein** systems is:

$$Q_{BLA}(l) = E\{Z(l)/U(l)\} = K\,(1 + \frac{C_1}{L(l)})\; G^1(l)\,H^1(l) \qquad (6.1.3)$$

where $L(l) = S_1(l)R_2(l)$. (6.1.4)

For small levels of $\varepsilon_1 = a_3/a_1$, $\varepsilon_2 = b_3/b_1$, $K \approx 1 + O(\varepsilon)$, $C_1 \approx O(\varepsilon^2)$, ($\varepsilon$ = max($\varepsilon_1$, $\varepsilon_2$)), and (6.1.3) yields: $Q_{BLA}(l) \approx G_{BLA}(l)\,H_{BLA}(l) \cong G^1(l)\,H^1(l)$, (6.1.5)

i.e. the usual product expression for the cascaded systems holds also in the case of weak non-linear systems in those frequency bands, where the coherence function is high (see Fig. 6.1.2). Observe however a small bias $K$ due to the presence of non-linearity ($\varepsilon_1$, $\varepsilon_2$ = 0.1). In those frequency bands using the best linear approximation is a sound modeling strategy, which provides the proper view of the dynamics of the cascade.

**Proof: In Appendix A.6** ■

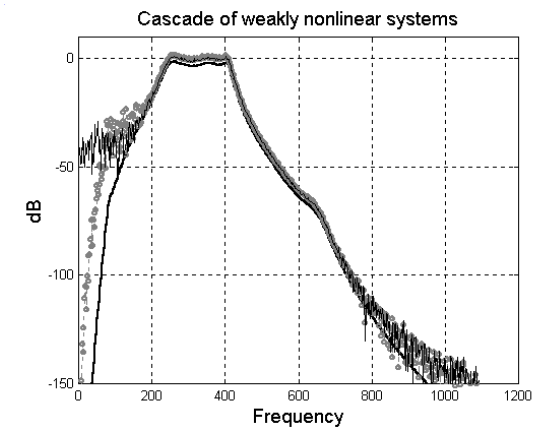


**Fig. 6.1.2** Cascading weakly non-linear systems: the FRF of the cascade of linear components, i.e. $R_1$ $S_1$ $R_2$ $S_2$ (solid black), the product $G_{BLA}$ $H_{BLA}$ (gray o), and the $Q_{BLA}$ averaged from $N$=10 (noisy solid black). In the frequency bands, where the coherence functions of the cascade components are high (see Figs. 2.5.1-2.5.2), the linear approximation is sufficient.

## 6.2 Non-linear distortion in cascaded MIMO systems

The simplified SISO cascade problem will be now investigated in the fully blown MIMO setting. Cascading is an elementary way to build complex systems and to model and solve practically important questions. Here we name only two:

(1) The excitation signals are applied through non-ideal actuators and/or are distorted by the non-linear loads. Should we accept the measurement results as they are or should we redesign the excitations to counter the effects of the actuators and the loads? [117-118]

(2) Linear FRF approximation to a non-linear system is theoretically valid only for the particular choice of input signals used during the measurements. This limits the usability of the linear approximations, because in different applications the inputs will generally be different. If however the new inputs do not differ much, the original linear approximation should hold, or shouldn't it? (see also Sect 3.6 for this problem)

These questions will be modeled as follows. What is the deterioration of the measurement quality, if the ideal excitation signal is distorted by passing through a non-linear system?

The measurement set-up is presented in Fig. 6.2.1. Ideally the designed reference signal ***R*** should be applied directly to the input of the system **V**. The Best Linear Approximation FRF $\hat{\mathbf{G}}_r = [G^m_{r,k}]$ of this system can be estimated from the ***Y*** - ***R*** measurements. We call this FRF the *reference estimate*. (Note that in the kernels $G^m_{\dots,k}$ the upper index, or indices, refer to the input signals and the multiple equals the order of the kernel, the lower index refers to the output signal and will be left out, if unambiguous).

In practice the reference signal undergoes distortions before reaching the excited input. Thus the FRF characteristics $\hat{\mathbf{G}} = [\hat{G}^m_k]$ are estimated from the ***Y*** - ***U*** data, where *U* is the output of the system **U** modeling the distortions. We assume that the distorted signals ***U*** can be directly measured. The question now is when is the estimated $\hat{\mathbf{G}}$ a fair approximation to the Best Linear Approximation $\hat{\mathbf{G}}_r$? The answer is not trivial because $\hat{\mathbf{G}}$ depends upon the actual excitation!

To tackle the problem we assume that both cascaded systems are MIMO Volterra systems limited to be at most 3rd order, to enumerate the distorting effects. For the input signals random phase multisines (2.2.17) and the orthogonal multisines will be used but results are valid also for other random excitations [15*, 16*, 21* 163].

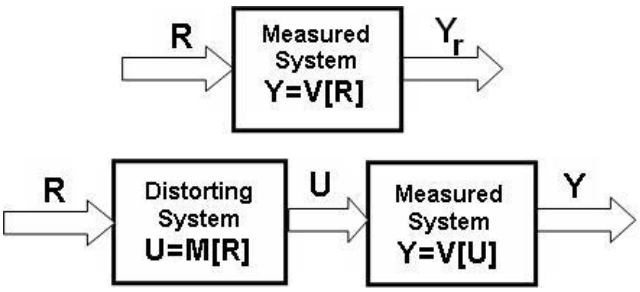


**Fig. 6.2.1** The ideal and the real measurement set-up.

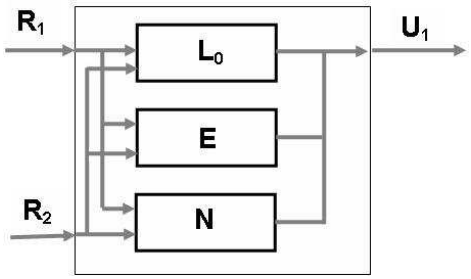


**Fig. 6.2.2** Channel distortions for $U_1$, for $N = 2$. Similar scheme is valid for signal $U_2$.

The distorting system **M** is the sum of an ideal system $\mathbf{L}_0$ (phase distortion only), a linear distortion system **E** (cross channel distortions) and a non-linear distortion system **N**, Fig. 6.2.2.

Generally the linear distortion between the reference inputs ***R*** and the distorted inputs ***U*** will be modeled as:

$$\mathbf{L}(l)=\begin{bmatrix}(1+\varepsilon_{11}(l))e^{j\varphi_{11}(l)} & \varepsilon_{12}(l)e^{j\varphi_{12}(l)} & \dots & \varepsilon_{1N}(l)e^{j\varphi_{1N}(l)}\\ \varepsilon_{21}(l)e^{j\varphi_{21}(l)} & (1+\varepsilon_{22}(l))e^{j\varphi_{22}(l)} & \dots & \varepsilon_{2N}(l)e^{j\varphi_{2N}(l)}\\ \dots & \dots & \dots & \dots\\ \varepsilon_{N1}(l)e^{j\varphi_{N1}(l)} & \varepsilon_{N2}(l)e^{j\varphi_{N2}(l)} & \dots & (1+\varepsilon_{NN}(l))e^{j\varphi_{NN}(l)}\end{bmatrix}=\mathbf{L}_0(l)+\mathbf{E}(l) \quad (6.2.1)$$

where $N$ is the input dimension of the cascade and $l$ is a discrete frequency.

The ideal system $\mathbf{L}_0$ can contain deterministic phase shifts. It will not influence the results because the random phases of the chosen excitation signals are random and uniformly distributed on the unit circle. The "ideal" channel is then:

$$\mathbf{L}_0(l)=\begin{bmatrix}e^{j\varphi_{11}(l)} & 0 & \dots & 0\\ 0 & e^{j\varphi_{22}(l)} & \dots & 0\\ \dots & \dots & \dots & \dots\\ 0 & 0 & \dots & e^{j\varphi_{NN}(l)}\end{bmatrix} \quad (6.2.2)$$

Observe also that: $\mathbf{L}_0\,\mathbf{L}_0^H=\mathbf{I}$ (6.2.3)

The linear distortion system models the phase and the amplitude distortion between any reference ***R*** and actual input ***U***:

$$\mathbf{E}(l)=\begin{bmatrix}\varepsilon_{11}(l)e^{j\varphi_{11}(l)} & \dots & \varepsilon_{1N}(l)e^{j\varphi_{1N}(l)}\\ \dots & \varepsilon_{km}(l)e^{j\varphi_{km}(l)} & \dots\\ \varepsilon_{N1}(l)e^{j\varphi_{N1}(l)} & \dots & \varepsilon_{NN}(l)e^{j\varphi_{11}(l)}\end{bmatrix}=\left[\varepsilon_{km}(l)e^{j\varphi_{km}(l)}\right]=[e_{km}(l)] \quad (6.2.4)$$

The non-linear distortions are modeled as a MIMO Volterra system of at most 3rd order (without linear terms, which are accounted for by **E**):

$$u_k(t)=\sum_{\alpha=2}^{3}u_k^{\alpha}(t)=\sum_{\alpha=2}^{3}\sum_{j_1 j_2\dots j_\alpha}u_k^{\,j_1 j_2\dots j_\alpha}(t) \quad (6.2.5)$$

$$u_k^{\,j_1 j_2\dots j_\alpha}(t)=\int_{-\infty}^{\infty}\dots\int_{-\infty}^{\infty}n_k^{\,j_1 j_2 j\dots j_\alpha}(\tau_1,\dots,\tau_\alpha)\,r_{j_1}(t-\tau_1)\dots r_{j_\alpha}(t-\tau_\alpha)\,d\tau_1\dots d\tau_\alpha \quad (6.2.6)$$

In the frequency domain the model is:

$$U_k(l) = \sum_{\alpha=2}^{3} U_k{}^{\alpha}(l) = \sum_{\alpha=2}^{3} \sum_{j_1 j_2 \dots j_\alpha} U_k{}^{j_1 j_2 \dots j_\alpha}(l) \tag{6.2.7}$$

$$U_k{}^{j_1 j_2 \dots j_\alpha}(l) = M^{-\alpha/2} \sum_{k_1,\dots,k_{\alpha-1} \in S_M^- \cup S_M^+} N_k{}^{j_1 j_2 \dots j_\alpha}(k_1, k_2, \dots, k_{\alpha-1}, k_\alpha)\, R_{j_1}(k_1)\, R_{j_2}(k_2) \dots R_{j_\alpha}(k_\alpha) \tag{6.2.8}$$

where $l$ is a discrete frequency, $l = \Sigma\, k_i$, $i = 1 \dots \alpha$, and $M$ is the number of the excited frequency lines. E.g. for $N = 2$, the model of a particular distorted excitation $U_k$ contains $N_k{}^{11}$, $N_k{}^{12}$, and $N_k{}^{22}$ 2nd order kernels, and $N_k{}^{111}$, $N_k{}^{112}$, $N_k{}^{122}$, and $N_k{}^{222}$ 3rd order kernels. The model is both general and simple. Effects of even and odd order non-linearities can be analyzed and it can be extended to higher order models, if required. We will also assume that the kernels can be written as:

$$N_k{}^{j_1 j_2 \cdots j_\alpha} = \delta_\alpha\, Q_k{}^{j_1 j_2 \cdots j_\alpha} \tag{6.2.9}$$

where $\delta_\alpha$ is the level of non-linear perturbations and $Q_k{}^{j_1 j_2 \cdots j_\alpha}$ is normalized as:

$\| Q_k{}^{j_1 j_2 \cdots j_\alpha} \|_\infty = \max | Q_k{}^{j_1 j_2 \cdots j_\alpha} | = 1$.

The measured system is a MIMO Volterra system of at most 3rd order (with linear terms):

$$y_k(t) = \sum_{\alpha=1}^{3} y_k^{\alpha}(t) = \sum_{\alpha=1}^{3} \sum_{j_1 j_2 \dots j_\alpha} y_k{}^{j_1 j_2 \dots j_\alpha}(t) \tag{6.2.10}$$

where a particular $\alpha$th order kernel, excited by input signals of indices $j_1, j_2, \dots, j_\alpha$, is:

$$y_k{}^{j_1 j_2 \dots j_\alpha}(t) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} g_k{}^{j_1 j_2 \dots j_\alpha}(\tau_1, \dots, \tau_\alpha)\, u_{j_1}(t - \tau_1) \dots u_{j_\alpha}(t - \tau_\alpha)\, d\tau_1 \dots d\tau_\alpha \tag{6.2.11}$$

In the frequency domain the system model is:

$$Y_k(l) = \sum_{\alpha=1}^{3} Y_k{}^{\alpha}(l) = \sum_{\alpha=1}^{3} \sum_{j_1 j_2 \dots j_\alpha} Y_k{}^{j_1 j_2 \dots j_\alpha}(l) \tag{6.2.12}$$

$$Y_k{}^{j_1 j_2 \dots j_\alpha}(l) = M^{-\alpha/2} \sum_{k_1,\dots,k_{\alpha-1} \in S_M^- \cup S_M^+} G_k{}^{j_1 j_2 \dots j_\alpha}(k_1, k_2, \dots, k_{\alpha-1}, k_\alpha)\, U_{j_1}(k_1) U_{j_2}(k_2) \dots U_{j_\alpha}(k_\alpha) \tag{6.2.13}$$

where $l$ is a discrete frequency, $l = \Sigma\, k_i$, $i = 1 \dots \alpha$, and $M$ is the number of the excited frequency lines. The kernels $G_k{}^{\cdots}$ are assumed to be of $O(1)$ order.

The main result can be stated then as:

**Theorem 6.2.1:** If the cascaded MIMO systems in Fig. 6.2.1-6.2.2 are weak non-linear systems in the sense that:

$$\varepsilon = \max \varepsilon_{km} << 1,\ \delta = \max \delta_k << 1,\ \xi = \max(\varepsilon, \delta) << 1. \tag{6.2.14}$$

then the 1st order perturbation (in the introduced distortions) of the measured Best Linear Approximation of the system $\mathbf{Y}=\mathbf{V}[\mathbf{U}]$ is:

$$\begin{aligned} \hat{\mathbf{G}}_{BLA} &= \mathbf{G}_{LIN} + \mathbf{G}_{BIAS}(\mathbf{I} - \mathbf{E}\mathbf{L}_0^{-1} - \mathbf{N}_{BLA}\mathbf{L}_0^{-1}) + \mathbf{H}_1 + \mathbf{H}_2 + O(\xi^2) \\ &= \mathbf{G}_{BLA} + \mathbf{G}_{BIAS}(\mathbf{I} - \mathbf{M}_{BLA}\mathbf{L}_0^{-1}) + \mathbf{H}_1 + \mathbf{H}_2 + O(\xi^2) = \mathbf{G}_{BLA} + O(\xi) \end{aligned} \tag{6.2.15}$$

where:

$\mathbf{G}_{BLA} = \mathbf{G}_{LIN} + \mathbf{G}_{BIAS} = [G^{m}_{BLA,k}] = O(1)$ is the true Best Linear Approximation (under zero distortions $\varepsilon$ and $\delta$) of the signal channel $Y_k$ - $U_m$; and the 1st order distortions are:

$\mathbf{G}_{BIAS}\mathbf{E}\mathbf{L}_0^{-1} = O(\varepsilon)$, due to the superposition of the linear distortions in the inputs;

$\mathbf{G}_{BIAS}\mathbf{N}_{BLA}\mathbf{L}_0^{-1} = O(\delta)$, where $\mathbf{N}_{BLA}$ is the Best Linear Approximation of the non-linear part of the distorting system, is due to the superposition of the non-linear distortions in the inputs;

$\mathbf{G}_{BIAS}(\mathbf{I} - \mathbf{M}_{BLA}\mathbf{L}_0^{-1}) = O(\xi)$, where $\mathbf{M}_{BLA} = \mathbf{L}_0 + \mathbf{E} + \mathbf{N}_{BLA}$;

$\mathbf{H}_1 = O(\varepsilon)$ comes from the distortion of the non-linear kernels in the measured system caused by the linear mixing of the excitation channels;

$\mathbf{H}_2 = O(\delta)$ is caused by the interaction of the non-linear kernels between the two (measured and distorting) systems.

**Proof: In Appendix A.7 ■**

The analysis of (6.2.15) shows that weak distortions do not cumulate and that the degradation in the measured FRF can be accounted for by 1st order distortions, if the overall distortion level is low. Furthermore, the effects of feed-forward and cross-channel distortions can be separated under such assumptions.

A weakly non-linear system does not make this much damage to the random multisines, and that they are still suitable to gain insight into the behavior of the measured system. But what about the orthogonal multisines?

If $\mathbf{R}_N(l)$ is an orthogonal random multisine:

$$\mathbf{R}_N(l) = \begin{bmatrix} w_{11}R_1^{(1)}(l) & w_{12}R_1^{(1)}(l) & \dots & w_{1N}R_1^{(1)}(l) \\ w_{21}R_2^{(1)}(l) & w_{22}R_2^{(1)}(l) & \dots & w_{2N}R_2^{(1)}(l) \\ \dots & \dots & \dots & \dots \\ w_{N1}R_N^{(1)}(l) & w_{N2}R_N^{(1)}(l) & \dots & w_{NN}R_N^{(1)}(l) \end{bmatrix} \qquad (6.2.16)$$

with $w_{km}$ entries of an orthogonal or unitary matrix:

$$[w_{km}][w_{km}]^H = N\,\mathbf{I}_N,\ \mathbf{R}_N(l)\,\mathbf{R}_N^H(l) = N\,\mathbf{I}_N. \qquad (6.2.17)$$

When such signal is distorted passing through the system $\mathbf{U}[\mathbf{R}_N]$, we have for the case of linear distortions:

$$\mathbf{U}_N(l) = [\mathbf{L}_0(l) + \mathbf{E}(l)]\mathbf{R}_N(l), \text{ and} \qquad (6.2.18)$$

$$\begin{aligned} \mathbf{U}_N(l)\mathbf{U}_N^H(l) &= [\mathbf{L}_0(l) + \mathbf{E}(l)]\,\mathbf{R}_N(l)\mathbf{R}_N^H(l)\,[\mathbf{L}_0^H(l) + \mathbf{E}^H(l)] \\ &= N\,[\mathbf{L}_0(l)\mathbf{L}_0^H(l) + \mathbf{L}_0(l)\mathbf{E}^H(l) + \mathbf{E}(l)\mathbf{L}_0^H(l) + \mathbf{E}(l)\mathbf{E}^H(l)] \end{aligned} \qquad (6.2.19)$$

Taking into account (6.2.3) and retaining only the first order terms, we have:

$$\mathbf{U}_N(l)\mathbf{U}_N^H(l) = N\,\mathbf{I}_N + 2N\,\mathrm{Re}\{\mathbf{E}(l)\mathbf{L}_0^H(l)\} \qquad (6.2.20)$$

i.e. the orthogonality of the excitations is perturbed up to the order of the linear distortions, but the deterministic character of $\mathbf{U}_N(l)\mathbf{U}_N^H(l)$ is not affected.

On the other hand when the non-linear distortions are also present in the system **U[R]**:

$$\mathbf{U}_N(l) = [\mathbf{L}_0(l) + \mathbf{E}(l)]\mathbf{R}_N(l) + \mathbf{N}[\mathbf{R}_N](l), \tag{6.2.21}$$

then retaining only the first order expressions:

$$\mathbf{U}_N(l)\mathbf{U}_N^H(l) = N\,\mathbf{I}_N + 2N\,\mathrm{Re}\{\mathbf{E}(l)\mathbf{L}_0^H(l)\} + 2\,\mathrm{Re}\{\mathbf{L}_0(l)\mathbf{R}_N(l)\mathbf{N}^H[\mathbf{R}_N](l)\} \tag{6.2.22}$$

the last term brings in also the randomness (when the frequencies in the non-linear kernels and the frequency $l$ are not paired [30*]), which means that the variance of the FRF estimate increases comparing to the ideal orthogonal case.

**Example 6.2.1:** (for the full experiment see [19*]) For the illustration 2-dim Wiener-Hammerstein systems were driven by unit power orthogonal multisines. Such systems permit an easy manipulation of the nonlinearities (contained between input and output dynamics). Furthermore the BLA to a Wiener-Hammerstein system is proportional to its linear dynamics (if $M >> 1$), which means that the expected influence of the distortion will mainly change the level of $G_{BLA}$ and less in its frequency behavior. The linear dynamics of the Wiener-Hammerstein systems are shown in Figs. 6.2.3-6.2.5, where every linear dynamics was normalized to the unit $\|.\|_\infty$ norm. The static non-linearity within systems $\mathbf{N}_1$ and $\mathbf{N}_2$ is:

$$u = \delta_2[r_1^2 + r_2^2 + r_1r_2] + \delta_3[r_1^3 + r_2^3 + r_1r_2^2 + r_1^2r_2], \tag{6.2.23}$$

and that in the system **V** is:

$$y = u_1 + u_2 + \alpha_2[u_1^2 + u_2^2 + u_1u_2] + \alpha_3[u_1^3 + u_2^3 + u_1u_2^2 + u_1^2u_2], \tag{6.2.24}$$

with suitably adjusted coefficients. The measurements are made on the output channel $Y_1$, thus the output index $k = 1$ is dropped. The comparison is based on the following measures of distortion, calculated in the pass band:

1. The $\alpha_k$* gain needed to scale up the measured FRF $\hat{\mathbf{G}}$ to match (in the LSE sense) the theoretical value of $\mathbf{G}_{BLA}$ (i.e. the FRF measured without any distortions):

$$\alpha_k^* = \arg\min_a \sum_{l=F_1}^{F_2} |\hat{G}^k(l) - a\,G_{BLA}^k(l)|^2 \tag{6.2.25}$$

The actual measure shown is the difference $\Delta_k = |1 - \alpha_k^*|$ in dB. The ideal gain is 1, consequently $\Delta_k$ grows steadily with the increasing distortion levels.

2. The rms value of the residual characteristic $\hat{\mathbf{G}} - \alpha_k^*\mathbf{G}_{BLA}$:

$$RMS_k = \sqrt{\frac{1}{F_2 - F_1}\sum_{l=F_1}^{F_2} |\hat{G}^k(l) - \alpha_k^* G_{BLA}^k(l)|^2} \tag{6.2.26}$$

3. The phase error of the measured $\hat{\mathbf{G}}$ with respect to that of the ideal $\mathbf{G}_{BLA}$: (a) its maximum value in the frequency band, and (b) the rms value of the phase residual.

In the first test only $e_{11}$, $e_{22}$ were set and the system **V** did not contain the 2nd order terms (Test A), then 2nd order terms have been added to **V** (Test B), then cross distortions were added to the system **E** (Test C), and finally non-linear distortions **N** were added (Test D) to the full system **M**, with linear and nonlinear distortions kept at the same level. The figures show results for a small (10) and a large number of averages (1000). For the x-axis in figures, the SNR of the overall distortion level was used, defined as:

$$Dist = \frac{\sum_l |U_1(l) - [\mathbf{L}_0]_{11}\,R_1(l)|^2}{\sum_l |R_1(l)|^2} \tag{6.2.27}$$

The figures show the expected slow increase in the distortion measures as the level of (13) increases. Robustness is attained until the distortion level becomes large (Fig. 6.2.7-6.2.9). Nonlinear distortions produce larger errors than those caused by linear distortions of similar magnitude (Fig. 6.2.7, Fig. 6.2.9). Similarly cross channel

distortions yield larger errors than feed-forward distortions. The phase errors grow faster comparing to the averaged amplitude errors. It means that the FRF becomes locally "twisted" rather, than globally distorted over larger frequency range. These effects are better visible in the results averaged from a larger number of measurements. For a low number of averages the effects are screened partly by the variance of the nonlinear noise on the FRF estimates.

In conclusion, the best linear approximation measurements are robust under perturbed excitations modeled by weakly nonlinear MIMO Volterra system (weak nonlinear amplitude and phase disturbances). Considered that finite order Volterra systems are smooth, these results are not so unexpected. Intuitively drastic changes for low level distortions could be difficult to explain. Figures show that distortions with respect to the reference measurements grow steadily but smoothly, even for the level of distortions for which the 1st order approximation cannot be justified. Consequently the 1st order approximation, for small distortions, yields a legitimate view of the system behavior. The presented results extend the single-input single-output results obtained earlier [10*]. Finally the 1st order approximation yields tools to investigate the influence of distortions positioned in particular places in the MIMO structure.

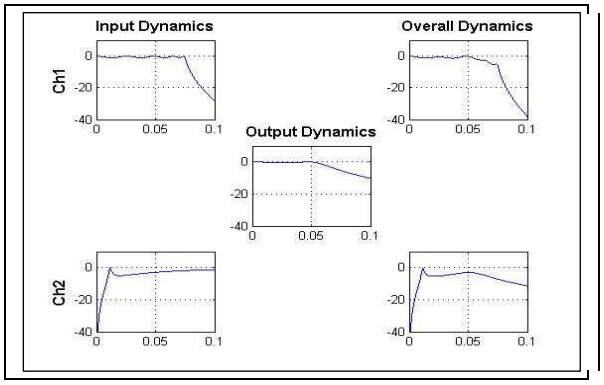


Fig. 6.2.3 Dynamics of the system $\boldsymbol{U}_1=\mathbf{N}_1\,[\boldsymbol{R}_1,\boldsymbol{R}_2]$.

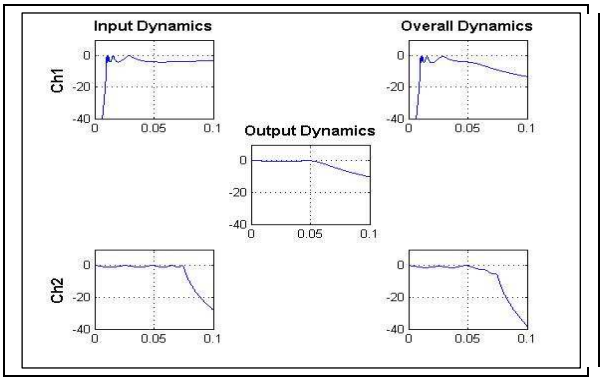


Fig. 6.2.4 Dynamics of the system $\boldsymbol{U}_2=\mathbf{N}_2\,[\boldsymbol{R}_1,\boldsymbol{R}_2]$.

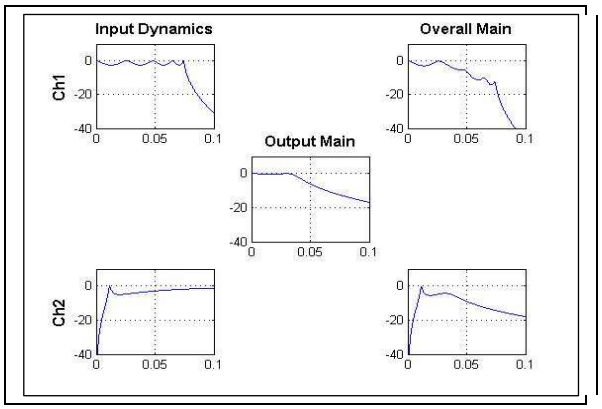


Fig. 6.2.5 Dynamics of the system $\boldsymbol{Y}=\mathbf{V}\,[\boldsymbol{U}_1,\boldsymbol{U}_2]$.

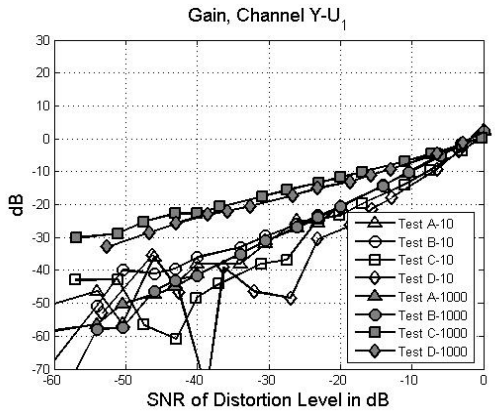


Fig. 6.2.6 Scaling gain $|1 - \alpha_k^*|$.

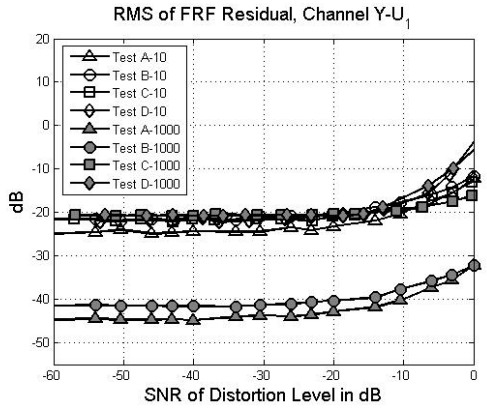


Fig. 6.2.7. The rms value of the FRF residual.

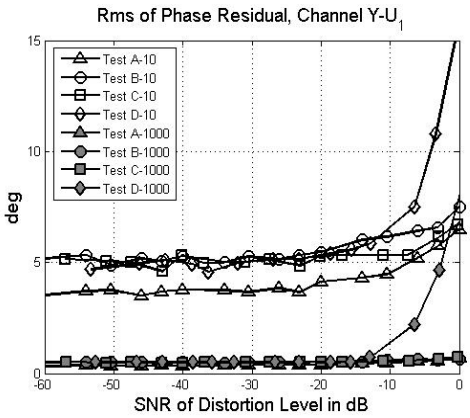


Fig. 6.2.9. The rms value of the FRF phase residual

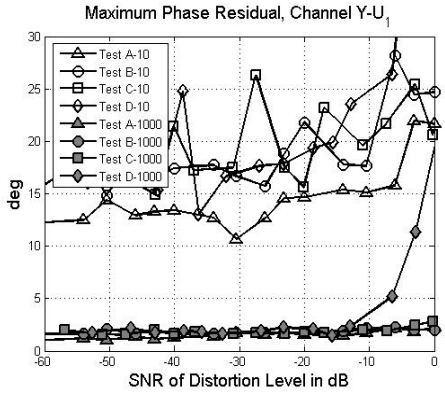


Fig. 6.2.8. The maximum value of the phase residual.

## 6.3 Risk of unstable behavior

The problem of weak non-linear effects, neglected in the close-loop design, but ready to pop up when the experimental conditions change, was studied also from the point of view of the close-loop stability. The expected phenomenon (reproduced in simulations) was to observe how a weakly non-linear system placed within the close-loop, driven by some excessive noise values shows more and more perceptible nonlinear noise and finally drives the loop into instability [38*, 41*, 43*].

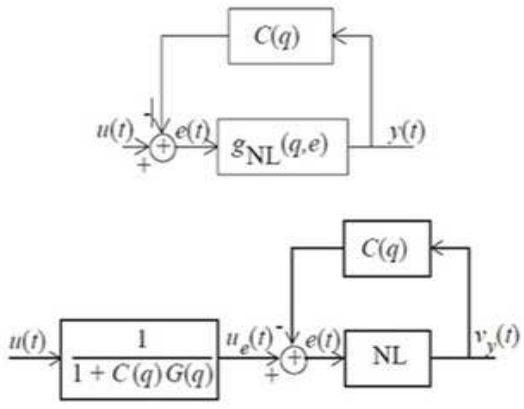


**Fig. 6.3.1**. Feed-back system used in the analysis.

The basic idea was to decompose the output power into the coherent (with the input) and non-coherent components and to introduce the non-linear power gain coefficient (6.3.1), to measure the sensitivity of the non-coherent power to the input perturbations. The gain permits

to define the stability conditions of the feed-back system, furthermore by estimating the gain and by using the theory of the extreme value distribution (GEV - General Extreme Value distribution) we can estimate the probability that the loop becomes instable with time.

$$\delta G_{NL} = \max_{\substack{\delta u \in S_{\delta u} \\ u,\, u+\delta u \in S_u}} \frac{\left\| \delta v_y(t) \right\|_2}{\left\| \delta e(u(t)) \right\|_2} \tag{6.3.1}$$

Extreme value techniques are based on the following. Consider $M_n = \max\{X_1, \ldots, X_n\}$, where $X_1, X_2, \ldots$ is a sequence of independent random variables having a common distribution $F$. For a very wide class of distributions [226, 43-44], the extreme values are asymptotically ($n \to \infty$) described by the general extreme value distribution (GEV):

$$G_{EV}(z) = \exp\{-[1+\xi(\frac{z-\mu}{\sigma})]^{-1/\xi}\}, \quad 1+\xi(\frac{z-\mu}{\sigma}) > 0 \tag{6.3.2}$$

with location parameter $\mu$ and scale parameter $\sigma$. The parameters $\xi$, $\mu$, $\sigma$ can be estimated from the available data, using a Maximum Likelihood estimator [43]. Starting from (6.3.2), the probability $p$ that $z$ exceeds a given return level $z_p$ within a given time interval $T$ can be calculated. So called return period $T_{z_p} = T / p$ is then connected to this probability, and this time can be interpreted as the mean time between two crossings of the level $z_p$. With this apparatus we can estimate the probability that the gain $\|C(q)\|_2 \delta G_{NL}$ will be larger than 1 for a given class of excitations in a given time interval.

In practice, a number of successive values of the non-linear gains (6.3.1) will be estimated from the input and output records. These measurement records are broken in $N_1$ subrecords of $N_s$ samples each. Each of these records is analyzed separately, resulting in $N_1$ non-linear gain measurements. These are grouped in $n$ blocks containing each $m$ gains ($N_1 = n \times m$). Next the maximum is calculated over each block, resulting in the measurements $X_1 \ldots X_n$, where $X_i$ is the maximum gain of the $i$th block of non-linear gains.

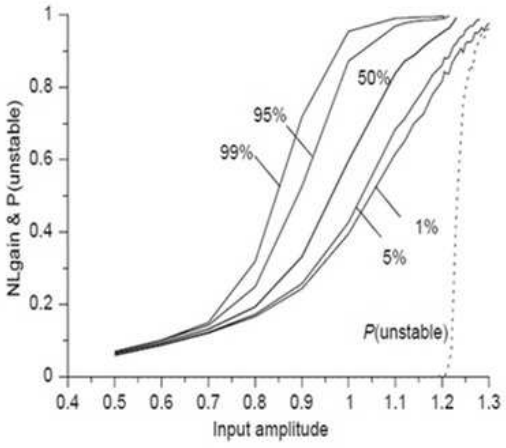


**Fig. 6.3.2.** The percentiles of the squared nonlinear power gain $\|C(q)\|_2 \delta G_{NL}$ and the observed relative frequency to get an unstable realization as a function of the input amplitude for the high-gain system.

The proposed method is experimental and requires further research. It is important that no non-linear model of the system is needed to estimate the risk of unstable operation of a non-linear system in a given time span. If the non-linear gain comes close to 1, the non-linear

feedback loop becomes potentially unstable. The major advantage of the method is its simplicity. The major disadvantage is the stochastic behavior of the gain factor, due to the stochastic nature of the input signal. This might lead to long experiments in order to get not too conservative risk estimates. The problem was further investigated in [237-240], but with no essential break-through.

## 6.4 Reducing the measurement time of the BLA by Monte Carlo averaging

The BLA is traditionally measured as the sample mean, considering that for linear systems and Gaussian excitations it is the minimum variance Maximum Likelihood estimate. This minimum variance is the basis of the measurement time vs. measurement accuracy trade-off. However due to a non-linearity sample mean is no more an optimal estimate and consequently its variance is not an attainable theoretical minimum and can be improved. By using a limited a priori knowledge about the measured system the measurement time vs. measurement accuracy trade-off can be made sharper, reducing the measurement variance solely as a part of the data processing, without affecting the measurement protocol [23*].

Suppose that the measurement data $Z_1, ..., Z_N$ are independent and identically distributed, and that the objective is to estimate their expected value $E\{Z_\mathrm{i}\}$. Then the usual estimator is the sample mean: $\hat{Z}_N = (Z_1 + Z_2 + ... Z_N)/N$, (6.4.1)

If the variance of $Z_\mathrm{i}$ is finite, this estimate is consistent and unbiased for all $N$. Furthermore it is asymptotically normally distributed, with variance: $\sigma^2 = \sigma_Z^2 / N$ (6.4.2)

Assume now that $Z_i$ are functions of some uniformly independent identically distributed $\varphi_i$, i.e.:

$$Z = Z(\varphi_1, \varphi_2, ..., \varphi_M), \quad \forall \varphi_i \in U(0,1), \tag{6.4.3}$$

then each $Z_i$ is a random realization drawing from the random $\varphi_i$-s

$$Z_i = Z(\varphi_{1i}, \varphi_{2i}, ..., \varphi_{Mi}), \quad \varphi_{ik} \in [0,1], \forall i,k, \tag{6.4.4}$$

and the expected value $E\{Z_i\}$ becomes: $E\{Z_i\} = \int_{\varphi_1} ... \int_{\varphi_M} Z(\varphi_1, \varphi_2, ..., \varphi_M) d\varphi_1 ... d\varphi_M$ (6.4.5)

The (6.4.1) sample mean is thus the Monte Carlo estimate of the multidimensional integral (6.4.5) [109]. The BLA FRF (2.2.25) simplifies for random multisine excitations to:

$$G_{BLA}(l) \approx E\{Y(l)\bar{U}(l)\}, \quad E\{|U(l)|^2\} = |U(l)|^2 = const \tag{6.4.6}$$

and is usually estimated from measurements $Y_i(l), U_i(l)$ as:

$$\hat{G}_{BLA,N}(l) = \frac{1}{|U_i(l)|^2 N} \sum_{i=1}^{N} Y_i(l)\bar{U}_i(l) = \frac{1}{N} \sum_{i=1}^{N} G_i(l) \tag{6.4.7}$$

Considering that the random phases of the input multisine are independently uniformly distributed, measuring $\hat{G}_{BLA,N}$ (6.4.7) is equivalent to the Monte Carlo simulation (6.4.1-6.4.5).

In the Monte Carlo simulation the key issue is the variance reduction [109, 98, 105], which is extremely important also from the measurement point of view. An estimate with a lower variance means shorter measurement time for the same accuracy, or conversely an increased accuracy if the measurement time is kept constant.

The variance of (6.4.2) can be reduced by reducing $\sigma_Z^2$, or by increasing the order of the numerator, say from $N$ to $N^2$.

A whole spectrum of approaches (antithetic variables, control variates, stratified, importance, Latin Hypercube sampling, quasi Monte Carlo, etc. [109]) is available in the literature. A method best suited for the real measurements will be that demanding limited a priori knowledge and introducing additional computations solely to the pre or post processing of the measurement data. That way the original measurement setup, the protocol of applying excitations and collecting the measurement results would remain intact.

Such method for the variance reduction is the so called control variates method [98, 105]. When estimating $E\{Z_i\}$, let $X$ be another (control) variable with the a priori known expected value $X_0 = E\{X_i\}$, and let $b$ be a suitable constant. Then:

$$Y_i = Z_i - b(X_i - X_0) \tag{6.4.8}$$

modified data yields the same desired mean:

$$E\{Z_i\} = E\{Y_i\}, \quad E\{\hat{Z}_N\} = E\{\hat{Y}_N\}, \tag{6.4.9}$$

i.e. instead of $\hat{Z}_N = (Z_1 + Z_2 + ... Z_N)/N$, we can use $\hat{Y}_N = (Y_1 + Y_2 + ... Y_N)/N$, if advantageous. Note that

$$\sigma_Y^2 = Var\{Y\} = \sigma_Z^2 + b^2\sigma_X^2 - 2b\,\mathrm{cov}_{Z,X}, \text{ (where } \mathrm{cov}_{Z,X} = E\{Z_i\, X_i\}\text{).} \tag{6.4.10}$$

If $X$ has high positive correlation with $Z$ (and has also an exactly known expected value, and is easily computable), then $\sigma_Y^2 < \sigma_Z^2$. For a fixed $X$, $\sigma_Y^2 = \sigma_Y^2(b)$, which can be minimized with respect to $b$, yielding:

$$b^* = \mathrm{cov}_{Z,X}/\sigma_X^2, \quad \sigma_Y^2(b^*) = \sigma_Z^2(1 - \rho_{Z,X}^2), \quad \rho_{Z,X} = \mathrm{cov}_{Z,X}/\sigma_Z\sigma_X \tag{6.4.11}$$

Consequently by selecting an $X$ with $\sigma_{Z,X} \neq 0$, the estimate variance can always be reduced.

The exact value of the optimal $b^*$ is not known, but it can be approximated from the data as:

$$b^* \approx \hat{b}_N^* = \frac{\sum_{i=1,...,N}(Z_i - \hat{Z}_N)(X_i - X_0)}{\sum_{i=1,...,N}(X_i - X_0)^2} \tag{6.4.12}$$

In the BLA FRF measurements we assume that beside measuring the output of the $V[u(t)]$ we can compute the output $x(t)$ of another (non-linear) system excited with the same $u(t)$, and from these we can estimate $\hat{H}_{BLA,N}(l)$, the BLA FRF of that system. Then the proposed scheme is:

$$\begin{aligned} G_i^*(l) &= G_i(l) - b^*(l)[H_i(l) - H_0(l)] \\ \hat{G}_{BLA,N}^*(l) &= \underbrace{\hat{G}_{BLA,N}(l)}_{measured} - \underbrace{\hat{b}^*(l)[\hat{H}_{BLA,N}(l) - H_0(l)]}_{computed} \end{aligned} \tag{6.4.13}$$

There are a number of issues to solve. Firstly, the control variates method should be extended to complex numbers and the whole frequency axis. Then a useful $H_i(l)$ non-linear system should be found, correlated with the measured $G_i(l)$, yielding outputs easily computable when excited with multisine signals, and possessing exactly known, and easily computable BLA FRF.

What kind of control system $H_i(l)$ to use? Two natural choices are:

I. If nothing is known about the measured system except that it is non-linear one, a static low order non-linear system can be tried to control the estimate.

II. If some computing effort can be spent to obtain a rough view of the measured dynamics, a Wiener control system can be built from this estimate followed by a low order static non linearity. The theoretical BLA FRF of a Wiener system is proportional to the linear dynamics, computing output of a Wiener system is easy, and clearly some non vanishing correlation is expected.

The measurement setup is presented in Fig. 6.4.1. The DUT is a non-linear system yielding $G_i(l)$ FRF in (6.4.13), the CTRL is a known non-linear system providing the control estimate $H_i(l)$ with the known theoretical BLA $H_0(l)$. The dashed part designates the pre-compiled control information and the final Best Linear Approximation FRF is computed from *y**-*u* input/output signal pair.

In the experiments the approximate value of $\hat{b}^*$ was smoothed with an 11th order regression, and considering that the BLA variance depends also on frequency, a rough measure of improvement was defined as:

$$\mu_N = \underset{\text{in excited band}}{mean}[Var\{G_N\}_{dB} - Var\{G_N^*\}_{dB}] \qquad (6.4.14)$$

A number of tests, with experimental data generated from various weakly nonlinear systems are presented in [23*].

Technically, if the correlation (6.4.11) is not zero, there is always a gain in variance. The gain is high when the correlation is close to 1, and is less, if the correlation is lower. Considering that the proposed method does not influence the measurement protocol, and is confined entirely to the measurement data processing, even a smaller gain may be of importance.

Collecting satisfactory a priori information in occasional, short measurements may be difficult, and the proposed method is probably of little use there. If the task is repetitive measurements, on systems with similar dynamics (e.g. screening and diagnostic tests, industrial process measurements, end production line tests), then the loss in time to estimate the control quantities is minute, and the return in shorter measurement time may be substantial.

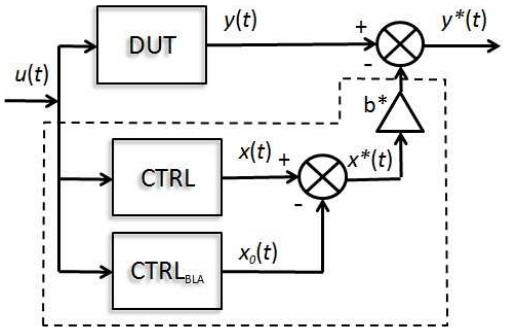


**Fig.6.4.1**. The control estimate measurement setup. The DUT is a non-linear system yielding $G_i(l)$ FRF in (6.5.13), the CTRL is a known non-linear system providing the control estimate $H_i(l)$ with known $\mathrm{CTRL}_{\mathrm{BLA}}$ $H_0(l)$. The dashed part can be precomputed from the measurement design and the final Best Linear Approximation FRF is computed from $u$-$y^*$ input/output signal pair.

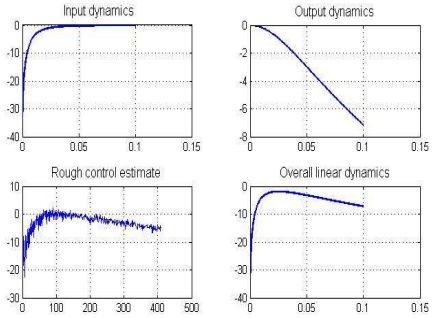


**Fig.6.4.2**. Test setup was made with a low-order Wiener-Hammerstein system with high-pass input (Butterworth, $f_c$ = 0.025, orders = 1, 2, 3), low-pass output (Butterworth, $f_c$ = 0.1, orders = 1, 2, 3), and the static nonlinearity $u + 0.5\,u^2 + 0.1\,u^3$.

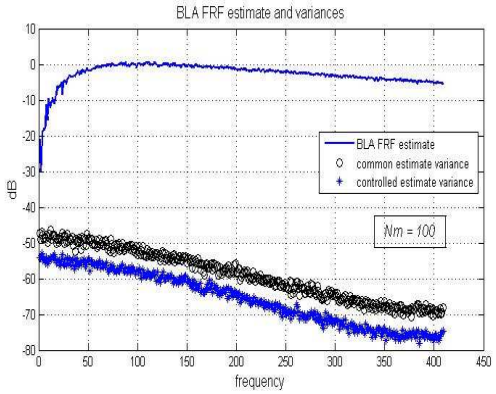


**Fig.6.4.3.** BLA FRF estimated in the test: (cont), variance (o) of the common estimate (6.4.7), and variance (*) of the controlled estimate (6.4.13), for the number of averages $N_{\mathrm{m}}$ = 100.

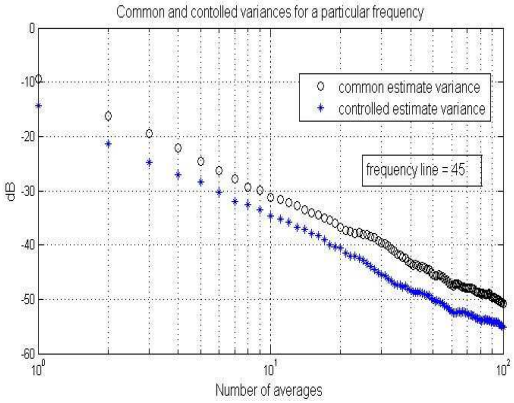


**Fig.6.4.4**. Variances of the BLA FRF estimated in the test: variance (o) of the common estimate (6.4.7), and variance (*) of the controlled estimate (6.4.13), for the frequency line $l$ = 45.

# 7. Utilization of the research results

What method or tool do we get by the developed BLA theory? Its user gets a „qualitatively better linear measurement technique” in s sense that s/he obtains a better insight into the possible non-linear distortions lurking in her/his model. With easily implementable measurement-technical advices it is possible to obtain such optimized measurement-technical solutions, which beside providing the non-linear information, will diminish essentially the time demands of the measurement and will yield general uncertainty bounds for the non-linear distortion related problems. That way there will be less risk for the misinterpretation and later misusage of the results of the widely used linear system identification methods.

The deficiency of the developed theory is that the user won't have the true non-linear system model (it was never the aim), is limited in the choice of excitations to the periodic signals, and finally that the assumed non-linear system class can be too limited for her/his application field.

BLA theory and the random multisine excitations were proposed and used in a wide spectrum of practical problems:

- to build approximate models to design Iterative Learning Control for scanning inkjet printers [17].

- for measuring and testing analogue-digital converters [151], [11-12], [131], [253]

- for measuring the quantization distortion of the DSP systems [152].

- to enhance the vertical resolution of low cost ADC [5].

- extensively for measuring and testing microwave circuitry [236], [30-31], [48-49], [18], [46], [231], [123], [13], [145], [47], [156-157], [241-243], [79-80].

- to qualify the bit-error rate (BER) properties and the nonlinearities in the σ-Δ converters [244-245].

- (with carefully chosen multisine excitations) to model the nonlinearites and to separate the linear and nonlinear behaviour in the analysis of the mechanical transmission in precision mechanical systems with dry friction [147], [111], [217], [176], [101].

- and also to model wet-clutch systems [265].

- in the identification of the multiple-joint industrial robots, [258-262], [92], [133], [179-182], [193-194], [149], with orthogonal multisines being a regular routine already at the ABB Robotics [146].

- to measure the nonlinearities in the biological impedances [95-97].

- in the broadband electrical impedance spectroscopy [7], [189-191].

- and for the measurements of the electrochemical impedance [14-15], [99], [77].

- in design of electronic tongue [110].

- for the non-invasive glucose measurements [150].

- as the first stage in the nonlinear structural modeling of the thin film deposition [266].

- in characterizing the nonlinearity in the ionic polymer transducers [108].

- as a component in the model of the direction-dependent processes in [228].

- to measure the Young's modulus [164-165, 167].

- to identify distillation columns [232-233].

- in the analysis and characterization of the operational amplifiers [141], [166, 168-169], [230], [229].

- in the detecting damage in mechanical structures [102], [248-250].

- for localizing errors in loudspeakers [32-33].

- in measuring nonlinear errors in discrete time radio receivers [83].

- in the analysis of the aircraft gas turbines [38-41], [74].

- in the analysis of the rotor bearings [4], [94], [159-160];

- in micromanufacturing (modeling of the micro-milling process) [16].

- in the analysis of the agricultural machines (sprayer devices) [2], [42].

- in modeling granule stream [132].

- in the modeling of electrical machines (synchronous machines) [254], (permanent magnet motors) [187].

- in the analysis and modal testing of the vibration of the automotive structures, [255-257], [195]; [78]; [270], [178], [45], [88].

- the analysis of the vehicle suspension [117-118], [196-197].

- the ground vibration analysis of the aircraft [160-161].

- and the identification of the head-neck complex under the upper body vibration [75], [3].

- to model RGA (Relative Gain Array) [100].

- as an auxiliary tool helping the selection of control structures and controller designs [35-37], [216], [93], [251-252], [222], [234], [218].

# References

## Own contributions


[1*] T.P. Dobrowiecki, F. Louage, **Expert Systems in System Identification - Second Generation Issues**, IMEKO XIII World Congress - From Measurement to Innovation, Torino, 1994, pp. 936-941.

[2*] T.P. Dobrowiecki, F. Louage, T. Mészáros, **Expert Systems in Measurement - From Smart Devices to Internet Agents**, X Polish Nat. Conf., Application of Microprocessors in Automatic Control and Measurement. Warsaw, 1996, Paper B11.

[3*] Dobrowiecki T.P., Louage F., Meszaros T.C., Roman G., Pataki B., **Will Measuring Instruments Turn Into Agents?** IEEE Trans. on Instr. and Meas., 46:(4) pp. 991-995, 1997.

[4*] Dobrowiecki T.P., Louage F., **Modeling, measurement and artificial intelligence - Toward the new generation of intelligent measuring systems**, Per. Polytechnica: Electrical Engineering 41:(4) pp. 123-133, 1997.

[5*] Dobrowiecki T.P.. J. Schoukens, **Practical Choices in the FRF Measurement in the Presence of non-linear Distortions**, Proc. of the IEEE Instr. and Measurement Tech. Conf., IMTC/99, Venice, Italy, May, 1999, pp. 922-927

[6*] Dobrowiecki, T.. Schoukens, J., **Bounds on Modeling Errors Due to Weak non-linear Distortions**, Internal Report, BME-MIT - VUB-ELEC, 2000, http://home.mit.bme.hu/~tade/publications/NL/bounds-report-2000.pdf.

[7*] Dobrowiecki, T., **Statistics of the complex output of a Wiener-Hammerstein system connected in a cascade**, Internal Report, BME-MIT, 2000, http://home.mit.bme.hu/~tade/publications/NL/Report-SISO-superposition-figures-2000.pdf.

[8*] Dobrowiecki T.P.. J. Schoukens, **Practical Choices in the FRF Measurement in the Presence of non-linear Distortions**, IEEE Trans. on Instr. and Meas., Feb 2001, Vol 50, No 1, pp. 2-7.

[9*] Dobrowiecki T.P., Schoukens J., **Bounds on modeling error due to weak non-linear distortions**, Proc. of the 18th IEEE Instr. and Meas. Tech. Conf., IMTC/2001. Budapest, pp. 14-19. Vol. 2

[10*] Dobrowiecki T., J. Schoukens, **Cascading Wiener-Hammerstein systems**, Proc. of the 19th IEEE Instr. and Meas. Tech. Conf., IMTC/2002. Anchorage, 2002, pp. 881-886.

[11*] Dobrowiecki T., J. Schoukens, **Coherence Function and Wiener-Hammerstein Systems**, Baltic Electronic Conference, BEC-2002. Tallinn, 2002, pp. 145-148.

[12*] Dobrowiecki T.P., J. Schoukens, **Excitations signals and non-linearly distorted systems**, MSM' 2003 - XI Miedzynarodowe Sem. Metrologów, 17-20 Sept, 2003, Rzeszow-Lviv, Poland-Ukraina, Materialy XI Miedzynarodowe Sem. Metrologów, 17-20 Sept, 2003, Rzeszow, Zeszyty Naukowe Politech. Rzeszowskiej, Elektrotechnika z. 25, pp. 33-44

[13*] Dobrowiecki T., J. Schoukens, **Optimal Measurements of Linear Approximation to Volterra TITO Systems**, The 9th Biennial Baltic Electronic Conf., BEC2004, Tallinn, Oct 3-6, pp. 189-192, 2004

[14*] Dobrowiecki T.P., J. Schoukens, **Linear Approximation of Weakly non-linear MIMO Systems**. Proc. of the 21th IEEE Instr. and Meas. Tech. Conf. – IMTC'2004, Como, Italy, pp. 1607-1612.

[15*] Dobrowiecki T., J. Schoukens, **Measuring linear approximation to weakly non-linear Mimo systems**, In: P Horacek, M Simandl, P Zitek (szerk.), Proc. of 16th IFAC World Congress. Prague, July 4-8, 2005, Elsevier, Paper 1915.

[16*] Dobrowiecki T., Schoukens J., Guillaume P., **Optimized excitation signals for Mimo frequency response function measurements**, Proc. of the IEEE Instr. and Meas. Tech. Conf., IMTC 2005. Ottawa, pp. 1872-1877.

[17*] Dobrowiecki T.P., Schoukens J., Guillaume P., **Optimized Excitation Signals For Mimo Frequency Response Function Measurements**. IEEE Trans. on Instr. and Meas. 55:(6) pp. 2072-2079. (2006)

[18*] Dobrowiecki T.P., J. Schoukens, **Robustness of the Related Linear Dynamic System Estimates in Cascaded non-linear MIMO Systems**, IMTC 2006, Instr. and Meas. Tech. Conf., 24-27 April 2006, pp. 117-122, Sorrento

[19*] Dobrowiecki T., **Superposition of weakly nonlinear MIMO Volterra systems**, Internal Report, BME-MIT, 2006, http://home.mit.bme.hu/~tade/publications/NL/MIMO-superposition-report.pdf.

[20*] Dobrowiecki T., J. Schoukens, **Measuring the best linear approximation of a non-linear system with uniformly frequency-distributed periodic signals**, In: Proc. of the IEEE Instr. and Meas. Tech. Conf., IMTC '2007. Warsaw, 2007, pp. 1-6.

[21*] Dobrowiecki T.P., J. Schoukens, **Measuring linear approximation to weakly non-linear MIMO systems**, Automatica 43:(10) pp. 1737-1751, 2007.

[22*] Dobrowiecki T.P., Schoukens J., **Linear approximation of weakly non-linear Mimo systems**. IEEE Trans. on Instr. and Meas. 56:(3) pp. 887-894. (2007)

[23*] Dobrowiecki T., J. Schoukens, **Reducing the measurement time of the Best Linear Approximation of a non-linear System using improved averaging methods**, IEEE I2MTC 2015 Int. Instrumentation and Measurement Technology Conf., May 11-14, Pisa

[24*] F. Louage, T.P. Dobrowiecki, **An Automated Measurement Identification Environment**, Proc. of the 10th IFAC Symp. on System Identification - SYSID '94, Copenhagen, 1994, pp. 453-458.

[25*] F. Louage, T.P. Dobrowiecki, **Integrated Intelligent Modelling Environment Based On the Krest Workstation**, 8th European Knowledge Acquisition Workshop EKAW-94, Hoegaarden, 1994, pp. 105-110.

[26*] F. Louage, T. P. Dobrowiecki, B. Pataki, **Knowledge-Based System Identification - Components of the Expertise**, Proc. of the IFAC Workshop on Control App. of Optimization (CAO), Visegrád, 2003, pp. 334-339.

[27*] Schoukens J., T.P. Dobrowiecki, R. Pintelon, **Identification of Linear Systems in the Presence of non-linear Distortions. A Frequency Domain Approach. Part I: Non-Parametric Identification**, Proc. of the 34th Conf. on Decision and Control, New Orleans, Dec 13-15, 1995, pp. 1216-1221.

[28*] Schoukens J., T.P. Dobrowiecki, R. Pintelon, **Identification of Linear Systems in the Presence of non-linear Distortions. A Frequency Domain Approach. Part II: Parametric Identification**, Proc. of the 34th Conf. on Decision and Control. New Orleans, Dec 13-15, 1995, pp. 1222-1227.

[29*] Schoukens J., T.P. Dobrowiecki, **Design of broadband excitation signals with a user imposed power spectrum and amplitude distribution**, Proc. of the IEEE Instr. and Meas. Tech. Conf., IMTC/98, St. Paul Hotel, St. Paul, Minnesota, May 18-21, 1998, pp. 1002-1005.

[30*] Schoukens J., Dobrowiecki T., Pintelon R., **Parametric and Nonparametric Identification of Linear Systems in The Presence of non-linear Distortions - A Frequency Domain Approach,** IEEE Trans. on Automatic Control 43:(2) pp. 176-190. (1998)

[31*] Schoukens J., R Pintelon, Y Rolain, T.P. Dobrowiecki, **Frequency response functions measurements in the presence of non-linear distortions. A general framework and practical advices**, DISC, Dutch Institute of Systems and Control, Summer School on Identification for Control. Veldhoven, Hollandia, 1999.

[32*] Schoukens J., R. Pintelon, Y. Rolain, T.P. Dobrowiecki, **Linear modeling in the presence of non-linear distortions. A general framework and practical advices,** ERNSI Workshop on Identification. Theoule, France, 1999.

[33*] Schoukens J., R. Pintelon, Y. Rolain, T. P. Dobrowiecki, **Frequency Response Function Measurements in the Presence of non-linear Distortions**, IFAC Symp. on System Identification (SYSID 2000), Santa Barbara, California, USA, 21-23 June, 2000

[34*] Schoukens J., R. Pintelon, Y. Rolain, T.P. Dobrowiecki, **Frequency response functions measurements in the presence of non-linear distortions. A general framework and practical advices**, ISMA'2000, Leuven, Sept 13-15, 2000, pp. 459-464

[35*] Schoukens J., Pintelon R., Rolain Y., Dobrowiecki T., **Frequency Response Function Measurements in The Presence of non-linear Distortions**, Automatica 37:(6) pp. 939-946, 2001.

[36*] Schoukens J., R. Pintelon, T.P. Dobrowiecki, **Linear Modeling in the Presence of non-linear Distortions**, IMTC'2001, Budapest, May 21-23, 2001, pp. 1332-1338, vol. 2. (state-of-the-art lecture)

[37*] Schoukens J., R. Pintelon, T.P. Dobrowiecki, **Nonparametric model bounds for control design in the presence of non-linear distortions**, Proc. of the 40th IEEE Conf on Decision and Control, Orlando, Florida, USA, Dec 2001, pp. 2998-3003, vol. 3.

[38*] Schoukens J., R. Pintelon, T.P. Dobrowiecki, **Some thoughts on the stability of closed loop systems in the presence of non-linear distortions**, ERNSI 2001, Cambridge, Paper Talk 10.

[39*] Schoukens J., R. Pintelon, Y. Rolain, T.P. Dobrowiecki, **Frequency response function measurements in the presence of non-linear distortions. A general framework and practical advices**, Proc. of the Int. Seminar on Modal Analysis, Vol 1, pp. 459-464, 2001, KU Leuven

[40*] Schoukens J., Pintelon R., Dobrowiecki T., **Linear Modeling in The Presence of non-linear Distortions,** IEEE Trans. on Instrumentation and Measurement 51:(4) pp. 786-792, 2002.

[41*] Schoukens J., R. Pintelon, T. Dobrowiecki, **Identification of the stability of feedback systems in the presence of non-linear distortions**, In: IFAC 15th Triennal World Congress, 21-26 July 2002, Barcelona, pp. 418-423.

[42*] Schoukens J., R. Pintelon, T.P. Dobrowiecki, Y. Rolain, **Identification of linear systems with non-linear distortions**, Plenary lecture, 13th IFAC Symp. on System Identification, 27-29 Aug, 2003, Rotterdam

[43*] Schoukens J., Dobrowiecki T., Pintelon R., **Estimation of the Risk for an Unstable Behaviour of Feedback Systems in the Presence of non-linear Distortions**. Automatica 40:(7) pp. 1275-1279, 2004.

[44*] Schoukens J., R. Pintelon, T. Dobrowiecki, Y. Rolain, **Identification of linear systems with non-linear distortions**, In: SCORES (Systems, Control and Optimization in Research, Education and Services): Workshop on System Identification. Leuven, 2004.

[45*] Schoukens J., Pintelon R., Dobrowiecki T., Rolain Y., **Identification of Linear Systems with non-linear Distortions**. Automatica 41:(3) pp. 491-504, 2005.

[46*] Schoukens J., J. Lataire, R. Pintelon, G. Vandersteen, T. Dobrowiecki, **Robustness Issues of the Equivalent Linear Representation of a non-linear System**, IEEE Trans. on Instr. and Meas. 58:(5) pp. 1737-1745, 2009.

[47*] Schoukens J., T. Dobrowiecki, Y. Rolain, R. Pintelon, **Upper bounding the variations of the best linear approximation of a non-linear system in a power sweep measurement**, Int. Instrumentation and Measurement Technology Conf., I2MTC09. Singapore, 2009, pp. 1514-1519.

[48*] Schoukens J., T. Dobrowiecki, Y. Rolain, R. Pintelon, **Upper Bounding Variations of Best Linear Approximations of non-linear Systems**. IEEE Trans. on Instr. and Meas. 59:(5) pp. 1141-1148. (2010)

[49*] Vanhoenacker K., T.P. Dobrowiecki, J. Schoukens, **Design of Multisine Excitations to Characterize the non-linear Distortion During FRF-measurements**, IMTC'2000, Proc. of the 17th IEEE Instr. and Meas. Technology Conf., Baltimore, May 1-4, 2000, pp. 1254-1259, vol. 3.

[50*] Vanhoenacker K., Dobrowiecki T., Schoukens J., **Design of Multisine Excitations to Characterize the Non-linear Distortions During FRF-measurements**. IEEE Trans. on Instr. and Meas. 50:(5) pp. 1097-1102. (2001)

## Literature

[1] N. Ahmed, K.R. Rao, ***Orthogonal Transforms for Digital Signal Processing***, Springer-Verlag, Berlin, 1975

[2] Anthonis J., Ramon H., **Design of an active suspension to surpress the horizontal vibrations of a spray-boom**, J. of Sound and Vibration, Vol 266, Issue 3, 18 Sept 2003, pp. 573–583.

[3] Atapourfard M., Ishihara T., Inooka H., **The influences of trunk horizontal vibration to the head-neck complex**, Proc. of the 41st SICE Annual Conference, Vols 1-5. Osaka, Aug 5-7, 2002, pp. 1053-1058.

[4] Azimian H., Fatehi A., Araabi B.N., **Takagi-Sugeno control of the elevation channel of a twin-rotor system using closed-loop empirical data**, Proc. of the American Control Conf., Montreal, June 27-29, 2012, pp. 2615-2620.

[5] A. Baccigalupi, M. Darco, A. Liccardo, R. Schiano Lo Moriello, **Compressive sampling-based strategy for enhancing ADCs resolution**, Measurement 56:95–103, Oct 2014.

[6] Barrett, J. F., Coales, J. F., **An Introduction to the Analysis of Nonlinear Control Systems with Random Inputs**. The Institution of Electrical Engineers, Monograph No. 154 M, 1955, pp. 190–199.

[7] Barsoukov E. (Ed.), J. Ross Macdonald (Ed.), ***Impedance Spectroscopy Theory, Experiment, and Applications***, 2ND Ed., Wiley, 2005.

[8] Bendat, J. S., A. G. Piersol, ***Engineering Applications of Correlations and Spectral Analysis***, Wiley, New York, 1980.

[9] Billings S.A., K.M. Tsang, **Spectral analysis for non-linear systems, Part II: Interpretation of non-linear frequency response function**, J. Mech. Syst. Signal Processing, Vol. 3, No. 4, pp. 341-359, 1989.

[10] Billingsley, P., ***Probability and Measure*** (Third ed.), John Wiley & Sons, 1995.

[11] N. Björsell, **Modeling Analog to Digital Converters at Radio Frequency**, PhD Thesis, Royal Institute of Technology (KTH), 2007.

[12] Björsell N., P. Suchanek, P. Handel, D. Ronnow, **Measuring Volterra Kernels of Analog-to-Digital Converters Using a Stepped Three-Tone Scan**, IEEE Trans. on Instr. and Meas. 57: (4) 666-671. (2008)

[13] Björsell N., Nader C., Händel P., **Multi-tone design for out-of-band characterization of nonlinear RF modules using harmonic sampling**, 2010 IEEE Int. Instr. and Meas. Tech. Conf., I2MTC 2010, Austin, May 3-6, 2010, pp. 620-623.

[14] Blajiev O.L., Pintelon R., Hubin A., **Detection and evaluation of measurement noise and stochastic non-linear distortions in electrochemical impedance measurements by a model based on a broadband periodic excitation**, J. of Electroanalytical Chemistry 576: (1) 65-72. (2005)

[15] Blajiev O.L., Breugelmans T., Pintelon R., Hubin A., **Improvement of the impedance measurement reliability by some new experimental and data treatment procedures applied to the behavior of copper in neutral chloride solutions containing small heterocycle molecules**, Electrochimica Acta 51: (8-9) 1403-1412.

(2006)

[16] R. S. Blom, **Model-based Process Monitoring and Control of Micromilling Using Active Magnetic Bearings**, Dutch Institute of Systems and Control (DISC), Technische Universiteit Delft, PhD Thesis, 2011,

[17] Bolder J., Lemmen B., Koekebakker S., Oomen T., Bosgra O., Steinbuch M., **Iterative learning control with basis functions for media positioning in scanning inkjet printers**, 2012 IEEE Multi-Conf. on Systems and Control, MSC 2012. Dubrovnik, Oct 3-5, 2012, pp. 1255-1260.

[18] Borremans J., De Locht L., Wambacq P., Rolain Y., **Nonlinearity analysis of analog/RF circuits using combined multisine and volterra analysis**, 2007 Design, Automation and Test in Europe Conf. and Exposition (DATE 2007), Nice, April 16-20, 2007, pp. 261-266.

[19] Boyd S., L. Chua, **Uniqueness of a basic non-linear structure**. IEEE Trans. on Circuits and Systems, CAS-30(9):648-651, Sept 1983.

[20] Boyd S., Y. Tang, L. Chua, **Measuring Volterra kernels**, IEEE Trans. on Circuits and Systems, CAS-30(8):571-577, Aug 1983.

[21] Boyd S., L. Chua, **Structures for non-linear systems**. Proc. IEEE Conf. on Decision and Control, 1:407-408, Dec 1984.

[22] S. Boyd, L. O. Chua, C. A. Desoer, **Analytical foundations of Volterra series**, IMA Journal of Mathematical Control and Information, Oxford University Press, 1(3):243-282, 1984.

[23] Boyd S., L. Chua, **Fading memory and the problem of approximating non-linear operators with Volterra series**, IEEE Trans. on Circuits and Systems, CAS-32(11):1150-1171, Nov 1985.

[24] Boyd S., L. Chua, **Uniqueness of circuits and systems containing one non-linearity**. IEEE Trans. on Automatic Control, TAC-30 (7): 674-681, July 1985.

[25] Boyd S., **Volterra Series: Engineering Fundamentals**. PhD Thesis, UC Berkeley, May 1985.

[26] Boyd S., **Multitone signals with low crest factor**, IEEE Trans. on Circuits and Systems, CAS-33(10):1018-1022, Oct 1986.

[27] Boyd S., J. Doyle, **Comparison of peak and RMS gains for discrete-time systems**, *Systems and Control Letters*, 9(1):1-6, June 1987.

[28] Braun M., R. Ortiz-Mojica, D. Rivera, **Design of Minimum Crest Factor Multisine Signals for "Model-on-Demand" System Identification and Model Predictive Control**, Annual AIChE Meeting, Los Angeles, Nov 12-17, 2000

[29] Brenner J., L. Cummings, **The Hadamard maximum determinant problem,** American Math. Monthly 79, 626-630, 1972.

[30] Bronckers S., Vandersteen G., Borremans J., Vandermot K., Van der Plas G., Rolain Y., **Advanced nonlinearity analysis of a 6 GHz wideband receiver**, I2MTC 2008 - IEEE Int. Instr. and Meas. Tech. Conf., Victoria, May 12-15, 2008, pp. 1340-1343.

[31] Bronckers S., Van der Plas G., Vandersteen G., Rolain Y., **The Prediction of the Impact of Substrate Noise on Analog/RF Circuits**, Substrate Noise Coupling in Analog/RF Circuits, 2010, pp. 151-193. (Artech House Microwave Library)

[32] Brunet P., Shafai B., **State-space modeling and identification of loudspeaker with nonlinear distortion**, Proc. of the IASTED Int. Conf. on Modelling, Simulation, and Identification, MSI 2011. Pittsburgh, Nov 7-9, 2011, pp. 246-253.

[33] P. Brunet, **Non-linear System Modeling and Identification of Loudspeakers**, PhD Thesis, Northeastern University, Boston, Massachusetts, April 2014

[34] Bussgang J.J, L. Ehrman, J.W. Graham, **Analysis of non-linear systems with multiple inputs**, Proc. IEEE, Vol. 62, pp. 1088-1119, Aug 1974.

[35] M. Castańo Arranz, W. Birk, **Bounds on a Gramian-Based Interaction Measure for Robust Control Structure Selection**, 2011 9th IEEE Int. Conf. on Control and Automation (ICCA), Santiago, Chile, Dec 19-21, 2011, MonA3.2, 89-93

[36] M. Castańo Arranz, **Practical tools for the configuration of control structures**, Dept of Computer Science, Electrical and Space Engineering, Lulea Univ. of Technology, Lulea, Sweden, PhD Thesis, 2012

[37] M. Castańo Arranz, W. Birk, **Estimation of Gramian-Based Interaction Measures for Weakly Nonlinear Systems,** 2015 European Control Conf. (ECC), July 15-17, 2015. Linz, Austria, 2438-2443

[38] Chiras, N., C. Evans, D. Rees. **Linear and non-linear gas turbine computer modeling**, UKACC Int. Conf. on Control, Cambridge, 2000, paper 12, (CD-ROM).

[39] Chiras, N., C. Evans, D. Rees. **Global non-linear modeling of gas turbine dynamics using NARMAX structures**, ASME J. of Engineering and Power, vol. 124, no. 4, pp. 817-826, 2002.

[40] Niclas Chiras, **Linear and non-linear Modelling of Gas Turbine Engines**, PhD Thesis, The Univ. of Glamorgan, Cardiff, 2002.

[41] Chiras N., Evans C., Rees D., **Nonlinear gas turbine modeling using feedforward neural networks**, Am. Soc. of Mech. Engineers, Int. Gas Turbine Inst., Turbo Expo (Publication) IGTI, (2002), pp. 145-152.

[42] Clijmans L, Ramon H, Sas P, et al., **Sprayer boom motion, Part 2: Validation of the model and effect of boom vibration on spray liquid deposition**, J. of Agricultural Eng. Research 76: (2) 121-128 June 2000

[43] Coles, S. **An introduction to statistical modeling of extreme values,** London: Springer, 2001.

[44] Coles, S. **The use and misuse of extreme value models in practice,** Seminaires Europeen de Statistique SemStat, Gothenburg, 10–16 Dec, 2001.

[45] B. Cornelis, A. Toso, W. Verpoest, B. Peeters, **Improved MIMO FRF estimation and model updating for robust Time Waveform Replication on durability test rigs**, ISMA 2014, Leuven, Sept 2014.

[46] Crama P., Rolain Y., **Broad-band measurement and identification of a Wiener-Hammerstein model for an RF amplifier**, ARFTG: Automatic RF Techniques Group, Conference Digest, Fall 2002 - Mesurement Needs for Emerging Technologies, Washington DC, Dec 5-6, 2002, pp. 49-57.

[47] G. Crupi, D. Schreurs, **Microwave De-embedding: From Theory to Applications**, Acad. Press, 2013

[48] De Locht L., Rolain Y., Vandersteen G., **Designing power amplifiers? Use good excitation signals,** 67TH ARFTG Microwave Measurements Conf. - Measurements and Design of High Power Devices and Systems, San Francisco, June 16, 2007, pp. 211-213.

[49] Ludwig De Locht, **Measuring, modeling and realization of high-frequency amplifiers**, PhD Thesis, Vrije Universiteit Brussel, August 2007

[50] T. D'haene, R. Pintelon, J. Schoukens, E. Van Gheem, **Variance Analysis of Frequency Response Function Measurements Using Periodic Excitations**, IEEE Trans. on Instr. and Meas., Vol. 54, No. 4, Aug 2005, 1452-1456

[51] Douce, J. L., **A Note on the evaluation of the response of a non-linear element to sinusoidal and random signals**. The Institution of Electrical Engineers, Monograph No. 257 M, 1957, pp. 1–5.

[52] Douce, J. L., Shackleton, J. M., **L.F. random-signal generator**. Electronic and Radio Engineer, 35(8), 295–297, 1958.

[53] Douce, J. L., Roberts, P. D., **Effect of distortion on the performance of non-linear control systems,** Proceedings of the IEE, 110(8), 1497–1502, 1963.

[54] F.J.III Doyle, R.K. Pearson, B.A. Ogunnaike, ***Identification and Control Using Volterra Models***, Springer Verlag, London, 2002

[55] Enqvist M., L. Ljung, **Approximation of non-linear Systems in a Neighborhood of LTI Systems**, Reglermöte 2002, pp 289-291, 2002.

[56] Enqvist M., L. Ljung, **Estimating non-linear Systems in a Neighborhood of LTI-approximants**, Proc. of the 41st IEEE Conference on Decision and Control, pp 1005-1010, 2002.

[57] Enqvist M., L. Ljung, **Linear Models of non-linear FIR Systems with Gaussian Inputs**. Proc. of the 13th IFAC Symp. on System Identification, pp. 1910-1915, Rotterdam, The Netherlands, August 2003.

[58] Enqvist M., L. Ljung, **LTI Approximation of Slightly non-linear Systems: Some Intriguing Examples**, Proc. of NOLCOS 2004 – IFAC Symposium on non-linear Control Systems, 2004.

[59] Enqvist M., L. Ljung, **Linear approximation of non-linear FIR systems for separable input processes**, Automatica, 41(3): 459-473, 2005.

[60] Enqvist M., **Linear Models of Non-linear Systems,**. PhD Dissertation, Department of Electrical Engineering, Linköping universiteit, Linköping, 2005.

[61] Enqvist M., **Identification of Hammerstein Systems Using Separable Random Multisines**, Preprints of the 14th IFAC Symposium on System Identification, pp. 768-773, 2006.

[62] M. Enqvist, **Separability of scalar random multisine signals**, Automatica, Vol 47, Issue 9, Sept 2011, pp. 1860–1867

[63] F. Esfahani, A., Schoukens, J., Vanbeylen, L., **Design of excitations for structure discrimination of nonlinear systems, using the best linear approximation**, I2MTC 2015, IEEE Int. Instr. and Meas. Tech. Conf., Pisa, Italy, May 11-14, 2015, pp. 234-239

[64] F. Esfahani, A., Schoukens, J., Vanbeylen, L., **Using the Best Linear Approximation With Varying Excitation Signals for Nonlinear System Characterization**, IEEE Trans. on Instr. and Meas. 65, 5, p. 1271-1280, 2016.

[65] Evans C., D. Rees, D.L. Jones. **Design of test signals for identification of linear systems with non-linear distortion**, IEEE Trans. Instrum. Meas., Vol. 41, Dec 1992, pp. 768-774

[66] Evans C., D. Rees, D.L. Jones. **Identifying Linear Models of Systems Suffering non-linear Distortions**, CONTROL'94, 21-24 March 1994, Conference Publications No. 389, pp. 288-296

[67] Evans C., D. Rees, D.L. Jones. **Non-linear disturbances errors in system identification using multisine test signals**, IEEE Trans. Instr. Meas., Vol. 43, No. 2, April 1994, pp. 238-244

[68] Evans C., D. Rees, D.L. Jones, D. Hill. **Measurement and Identification of Gas Turbine Dynamics in the Presence of Noise and Non-linearities,** Proc. of the IEEE Instr. and Meas. Techn. Conf. - IMTC'94, Hamamatsu, Japan, May 10-12, 1994

[69] Evans, C., D. Rees, D. L. Jones, M. Weiss. **Periodic Signals for Measuring non-linear Volterra Kernels**, IEEE Trans. on Instrumentation and Measurement, Vol. 45, No. 2, April 1996, pp. 362-371

[70] Evans C. **Identification of Linear and non-linear Systems Using Multisine Signals, with a Gas Turbine Application,** Ph.D. diss,, School of Electronics, University of Glamorgan, Wales, U.K., 1998.

[71] C. Evans, A. Borrell, D. Rees, **Testing and Modeling Gas Turbines Using Multisine Signals and Frequency-Domain Techniques**, J. Eng. Gas Turbines Power 121(3), 451-457, Jul 01, 1999.

[72] Evans C., D. Rees. **Non-linear Distortions and Multisine Signals – Part I: Measuring the Best Linear Approximation**. IEEE Trans. Instrum. Meas., vol. 49, pp. 602-609, June 2000.

[73] Evans C., D. Rees. **Non-linear Distortions and Multisine Signals – Part II: Minimizing the Distortion**. IEEE Trans. Instrum. Meas., vol. 49, pp. 610-616, June 2000.

[74] Evans C., Rees D., Borrell A., **Identification of Aircraft Gas Turbine Dynamics Using Frequency-domain Techniques**, Control Eng. Practice 8: (4) 457-467. (2000)

[75] Fard M. A., T. Ishihara, H. Inooka, **Identification of the head-neck complex in response to trunk horizontal vibration**, Biological Cybernetics, 90, 418–426 (2004), Springer-Verlag 2004

[76] G. Favier, A.Y. Kibangou, T. Bouilloc, **Nonlinear system modeling and identification using Volterra-PARAFAC models**, Int. J. of Adaptive Control and Signal Proc., Vol 26, Issue 1, pp. 30–53, Jan 2012

[77] L. Fernández Macía, M. Petrova, T. Hauffman, T. Muselle, Th. Doneux, A. Hubin, **A study of the electron transfer inhibition on a chargedself-assembled monolayer modified gold electrode by odd randomphase multisine electrochemical impedance spectroscopy**, Electrochimica Acta 140 (2014) 266–274

[78] Gatto M., Peeters B., Coppotelli G., **Flexible shaker excitation signals for improved FRF estimation and non-linearity assessment**, Int. Conf. on Noise and Vibration Eng. (ISMA), Leuven, Sept 20-22, 2010, pp. 2475-2488.

[79] Geens A., **Measurement and Modeling of the Noise Behaviour of High-Frequency non-linear Active Systems**, PhD Dissertation, Dept ELEC, VUB, May 2002

[80] Geens A., W. Van Moer, Y. Rolain, **Measuring in-band distortions of mixers**, IEEE Trans. on Instrumentation and Measurement, Vol. 52, No. 4, Aug 2003, pp. 1030-1034

[81] Geerardyn, E., Rolain, Y., Schoukens, J., **Quasi-logarithmic Multisine Excitations for Broad Frequency Band Measurements**, I2MTC 2012, IEEE Int. Instr. and Meas. Techn. Conf., Graz, Austria, May 13-16, 2012. p. 737-741.

[82] Geerardyn, E., Rolain, Y., Schoukens, J., **Design of Quasi-Logarithmic Multisine Excitations for Robust Broad Frequency Band Measurements**, IEEE Trans. on Instr. and Meas.. 62, p. 1364-1372, 2013.

[83] Geis A., Vandersteen G., Rolain Y., **Nonlinear Distortion Measurements of Discrete-Time Radio Receivers,** I2MTC 2008 - IEEE Int. Instr. and Meas. Tech. Conf., Victoria, May 12-15, 2008, pp. 1572-1576.

[84] Golub G.H., C.F. Van Loan, ***Matrix Computations***. 3rd Edition, John Hopkins Univ. Press, 1996

[85] Guillaume P., J. Schoukens, R. Pintelon, I. Kollar, **Crest factor minimization using non-linear Chebyshev approximation method**, IEEE Trans. Instr. Meas., Vol. 40, No. 6, Dec 1991, pp. 982-989, 1991.

[86] Guillaume P., **Identification of Multi-Input Multi-Output Systems Using Frequency-Domain Models**, PhD Thesis, ELEC, 1992

[87] Guillaume P., **Frequency Response Measurements of Multivariable Systems Using non-linear Averaging Techniques**, Proc. of the IEEE Instr. and Meas. Tech. Conf., Brussels, June 4-6, pp. 156-161, 1996

[88] Guillaume P., Verboven P., Vanlanduit S., Parloo E., **Multisine Excitations - New Developments and Applications in Modal Analysis**, Proc. the 19th Int. Modal Analysis Conf., Kissimmee (Florida), Feb 5-8, 2001

[89] R. Haber, H. Unbehauen, **Structure identification of nonlinear dynamic systems - a survey on input/output approaches**, Automatica, 26(4):651–677, 1990.

[90] R. Haber, L. Keviczky, ***Nonlinear System Identification — Input-Output Modeling Approach, Vol 1: Nonlinear System Parameter Identification***, Mathematical Modelling: Theory and Applications, Series Volume 7, 1999, Springer Netherlands

[91] R. Haber, L. Keviczky, ***Nonlinear System Identification — Input-Output Modeling Approach, Vol 2: Nonlinear System Structure Identification***, Mathematical Modelling: Theory and Applications, Series Volume 7, 1999, Springer Netherlands

[92] T. Hardeman, **Modelling and Identification of Industrial Robots Including Drive and Joint Flexibilities**, Universiteit Twente, PhD Thesis, Feb 2008

[93] Hjalmarsson H., **From Experiment Design to Closed-loop Control**, Automatica 41: (3) 393-438. (2005)

[94] K. Hynynen, **Broadband excitation in the system identification of active magnetic bearing rotor systems**, Acta Universitatis Lappeenrantaensis 446, Diss. Lappeenranta University of Technology, 2011.

[95] Ionescu C., A. Caicedo, M. Gallego, R. De Keyser, **Detection of nonlinearities when measuring respiratory impedance**, 26th Benelux Meeting on Systems and Control. Lommel, Belgium: 2007.03-2007.03. (2007) , pp. 120.

[96] Ionescu, C., Schoukens, J., De Keyser, R., **Detecting and analyzing non-linear effects in respiratory impedance measurements**, American Control Conference on O'Farrell Street, San Francisco, CA, USA, June 29-July 01, 2011. p. 5412-5417.

[97] Ionescu, C., Vandersteen, G., Schoukens, J., Desager, K., De Keyser, R., **Measuring non-linear Effects in Respiratory Mechanics: A Proof of Concept for Prototype Device and Method**, IEEE Trans. on Instr. and Measurement. 63, p. 124-134, 2014.

[98] F. James, **Monte Carlo theory and practice**, *Reports on Progress in Physics*, Vol 43, Nr 9, 1980, Available: http://iopscience.iop.org/0034-4885/43/9/002/pdf/rpv43i9p1145.pdf

[99] Jorcin J-B., Scheltjens G., Van Ingelgem Y., Tourwe E., Van Assche G., De Graeve I., Van Mele B., Terryn H., Hubin A., **Investigation of the self-healing properties of shape memory polyurethane coatings with the 'odd random phase multisine' electrochemical impedance spectroscopy**, Electrochim Acta 55: (21) 6195-6203. (2010)

[100] A.M.H. Kadhim, M. Castano, W. Birk and T. Gustafsson, **Relative Gain Array of Weakly Nonlinear Systems using a Nonparametric Identification Approach**, 2015 IEEE Conf. on Control Applications (CCA), Part of 2015 IEEE Multi-Conference on Systems and Control, Sept 21-23, 2015. Sydney, 1612-1617

[101] Keikha E., Al Mamun A., Lee T.H., Bhatia C.S., **Multi-frequency technique for frequency response measurement and its application to servo system with friction**, IFAC Proc. Volumes (IFAC-PapersOnline) 18th IFAC World Congress, Milano, Aug 28-Sept 2, 2011, pp. 5273-5278.

[102] Kerschen G., Worden K., Vakakis A.F., Golinval J.C., **Past, present and future of nonlinear system identification in structural dynamics**, Mechanical Systems and Signal Processing 20: (3) 505-592. (2006)

[103] R. Kershner, **The continuity of functions of many variables**, Trans. Amer. Math. Soc. , 53 (1943) pp. 83–100.

[104] Khalil H.K, ***Non-linear Systems***. Prentice-Hall, Upper Saddle River, N.J., 2nd edition, 1996.

[105] J. Kleijnen, A. Ridder, and R. Rubinstein, **Variance reduction techniques in Monte Carlo methods**, in *Encyclopedia of Operations Research & Management Science*, 3rd Edition (eds S. Gass and M. Fu), Springer-Verlag, pp. 1598-1610, 2013.

[106] I. Kollár, **Frequency Domain System Identification Toolbox for Use with Matlab**, MA: Mathworks, 1994.

[107] Korenev, B. G., ***Introduction into the Theory of Bessel Functions*** (in Russian), Nauka, Moskow, 1971

[108] C.S. Kothera, **Characterization, Modeling, and Control of the non-linear, Actuation Response of Ionic Polymer Transducers**, Virginia Polytechnic Institute and State University, PhD Thesis, Sept 2005

[109] D. P. Kroese, T. Taimre, Z. I. Botev, ***Handbook of Monte Carlo Methods***, John Wiley & Sons, 2011.

[110] Kumar R., Bhondekar A.P., Kaur R, Vig S., Sharma A, Kapur P., **A simple electronic tongue**, Sensors and Actuators B-Chemical 171: 1046-1053. (2012)

[111] Lampaert V., J. Swevers, F. Al-Bender, **Impact of non-linear friction on frequency response function measurements**, Proc. of the ISMA25, Int'l Conf. on Noise and Vibration Engineering, Sept 2000, Leuven, pp. 443-450

[112] Lang Zi-Qiang, S.A. Billings, **Evaluation of Output Frequency Responses of non-linear Systems Under Multiple Inputs**. IEEE Trans. on Circuits and Systems – II: Analog and Digital Signal Processing, Vol. 47, No. 1, Jan. 2000, pp. 28-38

[113] J. Lataire, **Frequency Domain Measurement and identification of Linear, Time-varying Systems**, PhD Thesis, Vrije Universiteit Brussel, March 2011

[114] L. Lauwers, J. Schoukens, R. Pintelon, M. Enqvist, **Nonlinear Structure Analysis Using the Best Linear Approximation**, Proc. of ISMA 2006, Int. Conf. on Noise and Vibration Engineering, 18-20 Sept, 2006.

[115] L. Lauwers, J. Schoukens, R. Pintelon, M. Enqvist, **A Nonlinear Block Structure Identification Procedure Using Frequency Response Function Measurements**, IEEE Trans. on Instr. and Measurement, Vol. 57, No. 10, pp. 2257-2264, 2008.

[116] L. Lauwers, **Some practical applications of the best linear approximation in non-linear block-oriented modelling**, PhD Thesis, Vrije Universiteit Brussel, May 2011

[117] Lauwerys C., J. Swevers, P. Sas, **Linear control of car suspension using non-linear actuator control**, ISMA 2002, Int. Conf. On Noise and Vibration Engineering, Katholieke Universiteit Leuven, Sept 16-18, 2002, pp. 55-61

[118] Lauwerys C., Swevers J., Sas P., **Design and experimental validation of linear controllers for an active suspension system of a passenger car**, Proc. of the Tenth Int. Congress on Sound and Vibration, Stockholm, July 7-10, 2003, pp. 869-877.

[119] Lee H., D. E. Rivera, H. D. Mittelmann, **A Novel Approach to Plant-Friendly Multivariable Identification of Highly Interactive Systems**, paper 436a, Annual AIChE Meeting , San Francisco, CA, Nov 16-21, 2003

[120] Lee H., D.E. Rivera, H.D. Mittelmann, **Constrained Minimum Crest Factor Multisine Signals for "Plant-Friendly" Identification of Highly Interactive Systems**, Proc. of 13th IFAC Symp. on System Identification, 27-29 Aug 2003, Rotterdam, The Netherlands

[121] Lee H. and D. Rivera, **A Control-Relevant, Plant-friendly System Identification Methodology using Shifted and "Zippered" Multisine Input Signals,** Annual AIChE Meeting, Austin, TX, Nov 7-12, 2004

[122] Lee H., D. Rivera, **An Integrated Methodology for Plant-Friendly Input Signal Design and Control-Relevant Estimation of Highly Interactive Process**, Annual AIChE Meeting, Cincinnati, OH, Oct 31-Nov 4, 2005

[123] Leulescu O., T. Petrescu, **A New Digital Predistortion Method For Power Amplifiers Linearization**, German Microwave Conf. - GeMiC 2005, Ulm, April 5-7, 2005

[124] Ljung L., ***Identification - Asymptotic Theory,*** Pergamon Press, 1987.

[125] Ljung L., ***System Identification - Theory for the user,*** Prentice-Hall, 2nd ed., 1999.

[126] Ljung L. **Model error modeling and control design**. In Proc. IFAC Symp., SYSID2000., Santa Barbara, CA, Jun 2000, pp. WeM1-3, 2000.

[127] Ljung L. **Estimating linear time-invariant models of non-linear time-varying systems**. Semi-plenary lecture at European Control Conf. ECC 2001, 4-7 Sept, 2001, Seminário de Vilar, Porto, Portugal

[128] L. Ljung**, Approaches to identification of nonlinear systems**, Proc. of the 29th Chinese Control Conference, Beijing, 29-31 July 2010, pp. 1-5.

[129] L. Ljung, G.C. Goodwin, J. C. Agüero, **Stochastic Embedding revisited: A modern interpretation**, CDC 2014: 3340-3345

[130] L. Ljung, G. C. Goodwin, J. C. Aguero, T. Chen, **Model Error Modeling and Stochastic Embedding**, IFAC-Papers O nLine 48-28 (2015) 075–079

[131] H.F. Lundin, **Characterization and Correction of Analog-to-Digital Converters**, PhD Thesis, Royal Institute of Technology (KTH) School of Electrical Engineering, 2005.

[132] Maertens K., Reyniers M., De Baerdemaeker J., **Design of a Dynamic Grain Flow Model for a Combine Harvester,** CIGR J. of Sc. Res. and Dev., PM 01 005 III, (2001)

[133] Makarov M., Grossard M., Rodríguez-Ayerbe P., Dumur D., **A frequency-domain approach for flexible-joint robot modeling and identification**, 16th IFAC Symposium on System Identification, Brussels, July 11-13, 2012, pp. 583-588.

[134] Mäkilä P.M., J.R. Partington, **On linear models for non-linear systems**. Automatica 39 (2003) 1-13

[135] Mäkilä P.M. **Squared and absolute errors in optimal approximation of non-linear systems,** Automatica 39 (2003) 1865-1876

[136] Mäkilä P.M., J.R. Partington, **Least-squares LTI approximation of non-linear systems and quasistationarity analysis,** Automatica 40 (2004) 1157-1169

[137] Mäkilä P.M., **On Optimal LTI Approximation of non-linear Systems,** IEEE Trans. on Automatic Control, Vol. 49, No. 7, July 2004, pp. 1178-1182

[138] Mäkilä P.M., **LTI modeling of NFIR systems: near-linearity and control, LS estimation and linearization**. Automatica 41 (2005) 29-41

[139] Mäkilä P.M., **LTI approximation of non-linear systems via signal distribution theory,** Automatica 42 (2006) 917-928

[140] Mäkilä P.M., **On robustness in control and LTI identification: Near-linearity and non-conic uncertainty,** Automatica 42 (2006) 601-612

[141] M. D. McKinley, **Improved Frequency Domain Measurement Techniques for Characterizing Power Amplifier and Multipath**, PhD Thesis, Electrical and Comp. Eng., Georgia Institute of Technology, 2008.

[142] Mittelmann H.D., G. Pendse, **Optimal Input Signal Design in Data-Centric System Identification**, in A.H. Siddiqi, I.S Duff, and O. Christensen (eds.), Modern Mathematical Models, Methods and Algorithms for Real World Systems, Anamaya Publ., New Delhi, India, 2006

[143] Mittelmann H.D., G. Pendse, D. E. Rivera, H. Lee, **Optimization-based Design of Plant-Friendly Multisine Signals using Geometric Discrepancy Criteria**, Computational Optimization and Applications, Sept 2007, Vol 38, Issue 1, pp 173-190

[144] Morokoff W. J., R. E. Caflisch, **Quasi-Random Sequences and Their Discrepancies**, SIAM J. Sci. Stat. Computing, 15:1251-1279, 1994

[145] Myslinski M., Pailloncy G., Avolio G., Verbeyst F., Bossche M. V., Schreurs D., **State-of-the-art microwave device characterization using large-signal network analyzers**, Proc. of the 17th Int. Conf. - Mixed Design of Integrated Circuits and Systems, MIXDES 2010, June 24-26, 2010, pp. 91-95.

[146] M. Norrlöf, Department of Electrical Engineering, Linköping University, personal communication, 2016.

[147] Nuij P.W.J.M., Steinbuch M., Bosgra O.H., **Experimental characterization of the stick/sliding transition in a precision mechanical system using the Third Order Sinusoidal Input Describing Function**, Mechatronics 18: (2) 100-110. (2008)

[148] Nutall, A.H., **Theory and Application of the Separable Class of Random Processes**, Tech Rep 343, May 26, 1958, MIT, Research Laboratory for Electronics

[149] Ohr J., S. Moberg, E. Wernholt, S. Hanssen, J. Pettersson, S. Persson, S. Sander-Tavallaey, **Identification of Flexibility Parameters of 6-axis Industrial Manipulator Models**, Proc. ISMA2006 Int. Conf. on Noise and Vibration Engineering, Leuven, Sept 18-20, 2006, pp. 3305-3314.

[150] Olarte O., Van Moer W., Barbe K., Verguts S., Van Ingelgem Y., Hubin A., **Using the Best Linear Approximation as a First Step to a New Non-Invasive Glucose Measurement**, 2012 IEEE Int. Instr. and Meas. Tech. Conf. (I2MTC), Graz, May 13-16, 2012, pp. 2747-2751.

[151] Ong M.S., Kuang Y.C, Liam P.S., Ooi M.P.L., **Multisine With Optimal Phase-Plane Uniformity for ADC Testing**, IEEE Trans. on Instr. and Meas. 61: (3) 566-578. (2012)

[152] J. Paduart, **Identification of non-linear Systems using Polynomial non-linear State Space Models**, PhD dissertation, Vrije Universiteit Brussel, ELEC Department, 2007.

[153] G. Palm, **On Representation and Approximation of Nonlinear Systems**, Biol. Cybernetics 31, 119 124 (1978)

[154] G. Palm, **On Representation and Approximation of Nonlinear Systems, Part II: Discrete Time**, Biol. Cybernetics 34, 49 52 (1979)

[155] R. K. Pearson, **Nonlinear Empirical Modeling Techniques**, Computers & Chemical Engineering 30(10):1514-1528, Sept 2006

[156] Pedro J.C., N.B. Carvalho, **Statistics of microwave signals and their impact on the response of nonlinear dynamic systems**, Proc 62nd ARFTG Conference Dig.. Boulder, Dec 4-5, 2003, pp. 143-153.

[157] Pedro J.C., Carvalho N.B, **Designing band-pass multisine excitations for microwave behavioral model identification**, IEEE Int. Microwave Theory and Tech. Symp. Digest. Fort-Worth, June 6-11, 2004, pp. 791-794.

[158] Peeters F., Pintelon R., Schoukens J., Rolain Y., Guiterrez E., Guillaume P., **Nonparametric Identification of Rotor-Bearing Systems in the Frequency Domain**, Proc. of the 18th Int. Modal Analysis Conf. (IMAC-XVIII), San Antonio (Texas), Feb 7-10, pp. 1838-1844, 2000

[159] Peeters F., Pintelon R., Schoukens J., Rolain Y., Guiterrez E., Guillaume P., **Identification of Rotor-Bearing Systems in the Frequency Domain. Part I: Estimation of Frequency Response Functions**, Trans. of Mechanical Systems and Signal Processing (MSSP), vol. 15, no. 4, pp. 759-774, 2001

[160] B. Peeters, A. Carrella, J. Lau, M. Gatto, G. Coppotelli, **Advanced shaker excitation signals for aerospace testing, Advanced Aerospace Applications**, Vol 1, Part of the series Conference Proc. of the Society for Experimental Mechanics Series, Proc. of the 29th IMAC, Conference on Structural Dynamics, pp. 229-241, 2011.

[161] Peeters, B.; Van der Auweraer, H., **Application of Multisine Excitation to Aircraft Ground Vibration Testing**, System Identification, Vol 16,| Part 1, 16th IFAC Symposium on System Identification, 2012, Brussels, Belgium, pp. 512-517

[162] Pintelon R., J. Schoukens, ***System identification. A frequency domain approach***. IEEE Press, 2001.

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

[164] Pintelon R., De Belder K., Guillaume P., Vanlanduit S., Rolain Y., **Identification of Material Damping from Broadband Modal Analysis Experiments**, SEM Annual Conf. and Exposition on Experimental and Applied Mechanics, Charlotte, North Carolina, USA, 2-4 June 2003, pp. 231-238

[165] Pintelon R., Guillaume P., De Belder K., Rolain Y., **Measurement of Young's Modulus via Modal Analysis Experiments: A System Identification Approach**, 13th IFAC Symp. on System Identification, Rotterdam, The Netherlands, 27-29 August, 2003, pp. 389-394

[166] Pintelon R., Vandersteen G., Rolain Y., **Experimental characterization of operational amplifiers: A system identification approach**, IEEE Instr. and Meas. Tech. Conf., Vail, CO, May 20-22, 2003, pp. 362-367.

[167] Pintelon R., Guillaume P., Vanlanduit S., De Belder K., Rolain Y., **Identification of Young's modulus from broadband modal analysis experiments**, Mechanical Systems and Signal Processing, 18 (2004), pp. 699-726

[168] Pintelon R., Y. Rolain, G. Vandersteen, J. Schoukens, **Experimental Characterization of Operational Amplifiers: A System Identification Approach - Part I: Theory and Simulations**, IEEE Trans. on Instrumentation and Measurement, June 2004, Vol 53, No 3, pp. 854-862

[169] Pintelon R., Y. Rolain, G. Vandersteen, J. Schoukens, **Experimental Characterization of Operational Amplifiers: A System Identification Approach - Part II: Calibration and Measurements**, IEEE Trans. on Instrumentation and Measurement, June 2004, Vol 53, No 3, pp. 863-876

[170] Pintelon R., J. Schoukens, ***System Identification: A Frequency Domain Approach***, 2nd Ed., Wiley-IEEE Press, 2012

[171] Pintelon, R., Schoukens, J., **The Best Linear Approximation of non-linear Systems Operating in Feedback**, I2MTC 2012, IEEE Int. Instr. and Meas. Tech. Conf., Graz, Austria, May 13-16, 2012. p. 2092-2097.

[172] Pintelon, R., Schoukens, J., **FRF Measurement of non-linear Systems Operating in Closed Loop**, IEEE Trans. on Instr. and Meas., 62, p. 1334-1345, 2013.

[173] R. Pintelon, E. Louarroudi, J. Lataire, **FRF measurements on slowly time-varying systems using multisines with non-uniformly spaced harmonics**, Proc. IEEE I2MTC, May 2013, pp. 1289–1294.

[174] R. Pintelon, E. Louarroudi, J. Lataire, **Quantifying the Time-Variation in FRF Measurements Using Random Phase Multisines With Nonuniformly Spaced Harmonics**, IEEE Trans. on Instr. and Meas., Vol. 63, No. 5, May 2014, pp. 1384-1394

[175] L. Pronzato, **Optimal experimental design and some related control problems** (Survey paper), Automatica 44 (2008) 303–325

[176] Qian X., Wang Y.Y., **System identify of servomechanisms with nonlinear friction**, 2007 Conf. Proc. IPEC, Vols 1-3. Singapore, Dec 3-6, 2007, pp. 276-281.

[177] D. Rees, **Digital Processing of System Responses**, Ph.D. dissertation, The Polytechnic of Wales, Faculty Eng., Dept. Elect. Eng., 1976.

[178] Reynders E., **System Identification Methods for (Operational) Modal Analysis: Review and Comparison**, Archives of Computational Methods in Engineering 19: (1) 51-124. (2012)

[179] Rijlaarsdam D., Van Loon B., Nuij P., Steinbuch M., **Nonlinearities in industrial motion stages - Detection and classification**, Proc. of the 2010 American Control Conf., ACC 2010, Baltimore, June 30-July 2, 2010, pp. 6644-6649.

[180] Rijlaarsdam, D., Nuij, P. W., Schoukens, J., Steinbuch, M., **Frequency Domain Based Friction Compensation - Industrial Application to Transmission Electron Microscopes**, American Control Conference on O'Farrell Street, San Francisco, CA, USA, June 29-July 01 2011 p. 4093-4098.

[181] D. Rijlaarsdam, **Frequency Domain Based Performance Optimization of Systems with Static non-linearities**, PhD Thesis, Cotutelle Tech. Universiteit Eindhoven - Vrije Universiteit Brussel, June, 2012

[182] Rijlaarsdam, D., Nuij, P. W., Schoukens, J., Steinbuch, M., **Frequency domain based non-linear feed forward control design for friction compensation**, Mechanical Systems and Signal Processing. 27, p. 551-562, 2012.

[183] Rivera D.E., M.W. Braun, H.D. Mittelmann, **Constrained Multisine Inputs for Plant-Friendly Identification of Chemical Processes**, Proc. of IFAC World Congress, 21-27 July 2002, Barcelona, Spain

[184] Rivera D.E., H. Lee, H. D. Mittelmann, G. Pendse, **Optimization-based Design of Plant-Friendly Input Signals for Data-Centric Estimation and Control**, Annual AIChE Meeting, paper 242k, Cincinnati, OH, Oct 31 - Nov 4, 2005

[185] Rivera D.E., H. Lee, H. D. Mittelmann, Pendse, **Optimization-based Design of Plant-Friendly Multisine Signals using Geometric Discrepancy Criteria**, Proc. of the 2006 SYSID Conf., Newcastle, Australia, 2006

[186] Rivera D.E., H. Lee, H. D. Mittelmann, M. W. Braun, **Constrained Multisine Input Signals for Plant-Friendly Identification of Chemical Process Systems**, J. of Process Control, Vol 19, Issue 4, April 2009, pp. 623–635

[187] Ruderman M., Ruderman A., Bertram T., **Observer-Based Compensation of Additive Periodic Torque Disturbances in Permanent Magnet Motors**, IEEE Trans. on Indust. Informatics 9: (2) 1130-1138. (2013)

[188] Rugh W.J., ***Non-linear system theory: the Volterra/Wiener approach,*** Johns Hopkins Univ. Press, 1981

[189] B. Sanchez, G. Vandersteen, R. Bragos, J. Schoukens, **Optimal multisine excitation design for broadband electrical impedance spectroscopy**, Measurement Science and Technology 22 (11), 115601, 2011

[190] B. Sanchez, G. Vandersteen, R. Bragos, J. Schoukens, **Basics of broadband impedance spectroscopy measurements using periodic excitations**, Meas. Science and Technology 23 (10), 105501, 2012

[191] B. Sanchez, E. Louarroudi, E. Jorge, J. Cinca, R. Bragos, R. Pintelon, **A new measuring and identification approach for time-varying bioimpedance using multisine electrical impedance spectroscopy**, Physiological measurement 34 (3), 339, 2013

[192] Sandberg I., $\boldsymbol{R_+}$ **Fading Memory and Extensions of Input-Output Maps**, IEEE Trans. On Circuits and Systems – I. Fundamental Theory and Applications, Vol. 49, No. 11, Nov 2002, pp. 1586-1591

[193] Saupe F., Knoblach A., **Design of excitation signals for the closed loop identification of industrial robots**, Proc. of the IEEE Int. Conf. on Control Applications. Dubrovnik, Oct 3-5, 2012, pp. 1553-1560.

[194] Saupe F., Knoblach A., **Experimental determination of frequency response function estimates for flexible joint industrial manipulators with serial kinematics**, Mechanical Systems and Signal Processing, 52–53:60–72, 2015.

[195] Savaresi S.M., Silani E., Bittanti S., Porciani N., **On Performance Evaluation Methods and Control Strategies for Semi-Active Suspension Systems**, Proc. of the IEEE Conf. on Decision and Control, Maui, Dec 9-12, 2003, pp. 2264-2269.

[196] Savaresi S.M., Silani E., Bittanti S., **Acceleration-driven-damper (add): An Optimal Control Algorithm For Comfort-oriented Semiactive Suspensions**, J. Dyn. Syst. Meas. Control-Trans. ASME 127: (2) 218-229. (2005)

[197] Savaresi S.M., Spelta C., **Mixed sky-hook and ADD: Approaching the filtering limits of a semi-active suspension**, J. of Dynamic Systems Measurement and Control-Trans. of THEASME 129: (4) 382-392. (2007)

[198] Schetzen M. ***The Volterra and Wiener Theories of non-linear Systems***. John Wiley & Sons, New York.

[199] Schoukens J., R. Pintelon, ***Identification of linear systems, a practical guideline to accurate modeling***, Pergamon Press Inc, New Jersey, U.S, 1991.

[200] Schoukens J., Guillaume P., Pintelon R., **Design of Multisine Excitations**, Proc. of the IEE Control 91 Conf., Edinburgh (UK), March 25–28, 1991, pp. 638–643

[201] Schoukens J., Guillaume P., Pintelon R., **Design of Broadband Excitation Signals**, in: ***Perturbation Signals for System Identification***, Ed. K. Godfrey Prentice Hall, Hempstead (U.K.), March 1993, pp. 126–160

[202] Schoukens J., Pintelon R., Guillaume P., **On the advantages of periodic excitation in system identification**, Proc. of SYSID '94, 10th IFAC Symp. on System Identification, Copenhagen, Denmark, July 4–6, 1994, Vol. 3, pp. 153–158

[203] Schoukens, J. Rolain, Y., Simon, G., Pintelon, R., **Fully automated spectral analysis of periodic signals**, IEEE Instr. and Meas. Tech. Conf. IMTC/2002, May 21-23, Anchorage, Vol. 1., pp. 299 –302

[204] Schoukens J., Pintelon R., Y. Rolain, **Time domain identification, frequency domain identification. Equivalencies! Differences?** Proc. of the 2004 American Control Conf., Boston, Mass., June 30-July 2, 2004, pp. 661-666

[205] Schoukens J., Barbé K., Vanbeylen L., Pintelon R., **Study of the nonlinear induced variability in linear system identification**, ERNSI 2008 Workshop in System Identification, Oct 1-3, 2008, Sigtuna, Sweden

[206] J. Schoukens, G. Vandersteen, K. Barbé, R. Pintelon, **Nonparametric preprocessing in system identification: a powerful tool**, 2009 European Control Conf. (ECC), 2009, Budapest, pp. 1-14, MoAPL.1

[207] Schoukens J., Barbé K., Vanbeylen L., Pintelon R., **Analysis of the Nonlinear Induced Variability in Linear System Identification**, System Identification, Vol 15, Part 1, Saint-Malo, 2009, pp. 634-639

[208] Schoukens, J., Pintelon, R., **Study of the Variance of Parametric Estimates of the Best Linear Approximation of non-linear Systems**, IEEE Trans. on Instr. and Meas. 59, p. 3156-3167, 2010.

[209] Schoukens, J., Barbé, K., Vanbeylen, L., Pintelon, R., **Non-linear Induced Variance of Frequency Response Function Measurements**, IEEE Trans. on Instr. and Meas.. 59, p. 2468-2474, 2010.

[210] Schoukens J., R. Pintelon, Y. Rolain, ***Mastering System Identification in 100 Exercises***, 2012, Wiley-IEEE Press

[211] J. Schoukens, A. Marconato, R. Pintelon, Y. Rolain, M. Schoukens, K. Tiels, L. Vanbeylen, G. Vandersteen, A. Van Mulders, **System Identification in a Real World**, AMC2014-Yokohama, March 14-16, 2014, Yokohama, Japan

[212] Schoukens, J., Pintelon, R., Rolain, Y., Schoukens, M., Tiels, K., Vanbeylen, L., Van Mulders, A. and Vandersteen, G., **Structure discrimination in block-oriented models using linear approximations: A theoretic framework**, Automatica. 53, p. 225-234, 2015.

[213] Schoukens, J., Vaes, M., F. Esfahani, Relan, R., **Challenges in the Identification of Discrete Time Block Oriented Models for Continuous Time Nonlinear Systems**, 17th IFAC Symp. on System Identification (SYSID 2015), Beijing, China, Oct 19-21, 2015. p. 596-601.

[214] M. R. Schroeder, **Synthesis of low peak-factor signals and binary sequences of low autocorrelation**, IEEE Trans. Inform. Theory, vol. 16, pp. 85-89, Jan. 1970.

[215] Simon, Gy., J. Schoukens, **Robust Broadband Periodic Excitation Design**, IEEE Trans. on Instr. and Meas., Vol. 49, No. 2, pp. 270-274, Apr 2000.

[216] Smolders K., M. Witters, J. Swevers, P. Sas, **Identification of a Nonlinear State Space Model for Control using a Feature Space Transformation**, Proc. of the Int. Conf. on Noise and Vibration Engineering (ISMA2006), Leuven, Sept 18-20, 2006, pp. 3331-3342.

[217] K. Smolders, **Tracking Control of Continuous-Time non-linear Systems: A Model-Based Iterative Learning Approach**, PhD Thesis, Katholieke Universiteit Leuven, 2007.

[218] Smolders K., Volckaert M., Swevers J., **Tracking control of nonlinear lumped mechanical continuous-timesystems: A model-based iterative learning approach**, Mechanical Systems and Signal Processing 22: (8) 1896-1916. (2008)

[219] Solomou M., C. Evans, D. Rees, N. Chiras, **Frequency Domain Analysis of Nonlinear Systems Driven by Multiharmonic Signals**, IMTC 2002, IEEE Instr. and Meas. Tech. Conf., Anchorage, May 21-23 2002, 799-804.

[220] Solomou M., D. Rees. **Measuring the Best Linear Approximation of Systems Suffering non-linear Distortions: An Alternative Method**, IEEE Trans. Instr. Meas., vol. 52, No. 4, pp. 1114-1119, Aug 2003.

[221] M. Solomou, **System identification in the presence of non-linear Distortions using Multisine Signals**, PhD Thesis, University of Glamorgan, 2003

[222] Solomou M., C. Evans, N. Chiras, D. Rees, **Controller Design for Systems Suffering non-linear Distortions**, 13th IFAC Symp. on System Identification, 27-29 Aug 2003, Rotterdam, pp. 1201-1206.

[223] Solomou M., D. Rees, **System Modeling in the Presence of non-linear Distortions**, Proc. of the IEEE Instrumentation and Measurement Technology Conf. – IMTC 2004, Lake Como, Italy, 18-20 May 2004, pp. 1601-1606.

[224] Solomou M., D. Rees, N. Chiras, **Frequency Domain Analysis of non-linear Systems Driven by Multiharmonic Signals**, IEEE Trans. Instrum. Meas., vol. 53, No. 2, pp. 243-250, April 2004.

[225] Solomou M., D. Rees, **Frequency Domain Analysis of non-linear Distortions on Linear Frequency Response Function Measurements,** IEEE Trans. Instrum. Meas., vol. 54, No. 3, pp. 1313-1320, June 2005.

[226] Stuart, A., Ord, J. K., ***Kendall's advanced theory of statistics*** (5th ed.). London: Charles GriUth, 1987.

[227] J. Swevers, C. Ganseman, D. B. Tukel, J. de Schutter, H. Van Brussel, **Optimal robot excitation and identification**, IEEE Trans. on Robotics and Automation, Vol 13 , Issue 5, Oct 1997, pp. 730 - 740

[228] Tan A.H., **Direction-dependent systems - A survey**, Automatica 45: (12) 2729-2743. (2009)

[229] Telescu M. G., Stievano I. S., Canavero F. G., Tanguy N., **An application of volterra series to IC buffer models**, 2010 IEEE 14th Workshop on Signal Propagation on Interconnects, SPI 2010, Hildesheim, May 9-12, 2010, pp. 93-96.

[230] Thorsell M., Andersson K., Pailloncy G., Rolain Y., **Using the best linear approximation to model the nonlinear behavior of supply modulated amplifiers**, European Microwave Week 2011: "Wave to the Future", EuMW 2011, Conf. Proc. - 6th European Microwave Integrated Circuit Conf., EuMIC 2011. Manchester, Oct 10-11, 2011, pp. 324-327.

[231] Thorsell M., Andersson K., Pailloncy G., Rolain Y., **Extending the Best Linear Approximation to Characterize the Nonlinear Distortion in GaN HEMTs**, IEEE Trans. on Microwave Theory and Techniques 59: (12) 3087-3094. (2011)

[232] Ugrumowa D., G. Vandersteen, B. Huyck, F. Logist, J. Van Impe, B. De Moor, **Identification and Modeling of Distillation Columns From Transient Response Data**, 2012 IEEE Int. Instr. and Meas. Tech.

Conf. (I2MTC 2012), Graz, May 13-16, 2012, pp. 2098-2103.

[233] D. Ugryumova, **Identification From Partially Missing Data and Modeling of Distillation Columns**, PhD Thesis, Vrije Universiteit Brussel, July 2015

[234] D. Vaes, **Optimal static decoupling for multivariable control design**, PhD Thesis, Katholieke Universiteit Leuven, 2005.

[235] Van Der Ouderaa E., Schoukens J., Renneboog J., **Peak Factor Minimization, using a Time-Frequency Domain Swaping Algorithm**, IEEE Trans. on Instr. and Meas., Vol 37, pp. 342-351.

[236] W. Van Moer, **Development of New Measuring and Modelling Techniques for RFICS and Their non-linear Behaviour**, PhD Thesis, Vrije Universiteit Brussel, ELEC, 2001.

[237] Vanbeylen, L., Schoukens, J., **Experimental stability analysis of non-linear feedback systems**, 16th ERNSI Workshop on System Identification, Venice, Italy, Oct 1-3, 2007.

[238] Vanbeylen, L., Schoukens, J., Barbé, K., **Measuring the stability of non-linear feedback systems**, I2MTC/2008, Proc. of the IEEE Int. Instr. and Meas. Tech. Conf., Victoria, Vancouver, Canada, May 12-15, 2008. p. 922-927.

[239] L. Vanbeylen, **Non-linear dynamic systems: blind identification of block-oriented models, and instability under random inputs**, PhD Thesis, Vrije Universiteit Brussel, 2011.

[240] Vanbeylen, L., Van Mulders, A. and Schoukens, J., **Bounded operation of unstable non-linear models under small random burst excitations**, 30th Benelux Meeting on Systems and Control, Lommel, Belgium, March 15-17, 2011.

[241] Vandermot K., W. Van Moer, J. Schoukens, Y. Rolain, **Understanding the non-linearity of a Mixer using Multisine Excitations,** Proc. of the IEEE Instr. and Meas. Technology Conf. IMTC'2006, Sorrento, Italy, 24-27 April, 2006, pp. 1205-1209

[242] Vandermot, K., Van Moer, W., Rolain, Y., Schoukens, J., **Identifying the Structure of non-linear Perturbations in Mixers using Multisine Signals**, IEEE Instr. & Meas. Magazine. 10, p. 32-39, 2007.

[243] K. Vandermot, **Best Linearized models for RF systems**, PhD Thesis, Vrije Universiteit Brussel, Oct 2010.

[244] Vandersteen, G., Rolain, Y., Vandermot, K., Pintelon, R., Schoukens, J.,Van Moer, W., **Quasi-Analytical Bit-Error-Rate Analysis Technique Using Best Linear Approximation Modeling**, IEEE Trans. on Instrumentation and Measurement. 58, p. 475-482, 2009.

[245] Vandersteen G., De Locht L., **Qualification and quantification of the nonlinear distortions in discrete-time sigma-delta ADCs**, 2009 European Conf. on Circuit Theory and Design, Vols 1 and 2. Antalya, Aug 23-27, 2009, pp. 759-762.

[246] Vanhoenacker K., J. Schoukens, J. Swevers, D. Vaes, **Summary and comparing overview of techniques for the detection of non-linear distortions**, ISMA 2002, Noise and Vibration Engineering, Leuven, 16-18 Sept 2002, pp. 1241-1255

[247] Vanhoenacker K., Schoukens J., Guillaume P., Vanlanduit S., **The use of multisine excitations to characterize damage in structures**, Mechanical Systems and Signal Processing, 18 (2004), pp. 43-57

[248] Vanlanduit S., Guillaume P., Schoukens J., E. Parloo, **Detection and Localization of non-linear Distortion**, Proc. of IMAC-XVII: A Conf. on Structural Dynamics, San Antonio, Texas, February 2-5, 2000, pp. 1597-1603

[249] Vanlanduit S., Guillaume P., Schoukens J., E. Parloo, **Linear and nonlinear damage detection using a scanning laser vibrometer**, Shock and Vibration 9 (2002) 43–56

[250] Vanlanduit S., Parloo E., Guillaume P., (2004) **An on-line combined linear-non-linear fatigue crack detection technique**, NDT & E International 37 (1): 41-45 Jan 2004

[251] Vargas A., F. Allgower, **Complexity Reduction of Nonlinear Systems for Process Control**, DYCOPS 7, 7th Int. Symp. on Dynamics and Control of Process Systems. Cambridge, USA, July 5-7, 2004.

[252] Vargas A., Allgower F., **Model reduction for process control using iterative nonlinear identification**, Proc. of the American Control Conference. Boston, June 30-July 2, 2004, pp. 2915-2920.

[253] B. Vargha, **Model Based Calibration of D/A Converters**, PhD Thesis, Vrije Universiteit Brussel, 2002.

[254] J. Verbeeck, **Standstill Frequency Response Measurement and Identification Methods for Synchronous Machines,** PhD Thesis, Vrije Universiteit Brussel, 2000.

[255] P. Verboven, **Frequency-Domain System Identification for Modal Analysis**, PhD Thesis, Vrije Universiteit Brussel, 2002.

[256] Verboven P., Guillaume P., Cauberghe B., Vanlanduit S., Parloo E., **Assessment of non-linear distortions in modal testing and analysis of vibrating automotive structures**, Proc. of 22nd Int. Modal Analysis Conf., pap.no.191, Dearborn, MI (USA), January 26-29, 2004.

[257] Verboven P., Guillaume P., Vanlanduit S., Cauberghe B., **Assessment of non-linear distortions in modal testing and analysis of vibrating automotive structures**, J. of Sound and Vibration, Vol 293, Issues 1–2, 30 May 2006, pp. 299–319

[258] Wernholt E., S. Gunnarson, **Analysis of methods for multivariable frequency response function estimation in closed loop**, 46th IEEE Conf. on Decision and Contr. New Orleans, Dec 12-14, 2007, pp. 4881-4888.

[259] E. Wernholt, **Multivariable Frequency-Domain Identification of Industrial Robots**, Dissertation. No. 1138, Dept of Electrical Eng., Linkoping 2007.

[260] Wernholt E., Moberg S., **Frequency-domain gray-box identification of industrial robots**, IFAC Proc. Volumes (IFAC-PapersOnline) 17th IFAC World Congress. Seoul, July 7-11, 2008, pp. 15372-15380.

[261] Wernholt E., Moberg, S., **Experimental Comparison of Methods for Multivariable Frequency Response Function Estimation**, IFAC World Congress, Vol 17, Part 1, COEX, South Korea, 2008, pp. 15359-15366

[262] Wernholt E., Moberg S., **Nonlinear gray-box identification using local models applied to industrial robots,** Automatica 47: (4) 650-660. (2011)

[263] West, J. C., Douce, J. L., Leary, B. G., **Frequency spectrum distortion of random signals in non-linear feedback systems**. The Institution of Electrical Engineers, Monograph No. 419 M, 1960, pp. 259–264.

[264] West, J. C., **The equivalent gain matrix of a multivariable non-linearity**. Meas. and Control, 2, T1–T5, 1960

[265] Widanage, W. D., Stoev, J., Van Mulders, A., Schoukens, J., Pinte, G., **Non-linear System-identification of the Filling Phase of a Wet-clutch System**, Control Engineering Practice. 19, p. 1399-1410, 2011.

[266] Wolfrum P., Vargas A., Gallivan M., Allgower F., **Complexity reduction of a thin film deposition model using a trajectory based nonlinear model reduction technique**, Proc. of the American Control Conf., Portland, June 8-10, 2005, pp. 2566-2571.

[267] Wong, H. K., Schoukens, J., Godfrey, K.R., **Analysis of Best Linear Approximation of a Wiener–Hammerstein System for Arbitrary Amplitude Distributions**, IEEE Trans. on Instr. and Meas., 61, March 2012, pp. 645-654.

[268] Wong H. K. Roland, Schoukens, J., Godfrey, K.R., **Design of Multilevel Signals for Identifying the Best Linear Approximation of non-linear Systems**, IEEE Trans. on Instr. and Meas. 62:(2) pp. 519-524, 2013.

[269] Wong H. K. Roland, **Study of the Best Linear Approximation of non-linear Systems with Arbitrary Inputs**, PhD Thesis, Cotutelle University of Warwick - Vrije Universiteit Brussel, 2013.

[270] Yin Y.Y., Zivanovic S., Li D.X., **Displacement modal identification method of elastic system under operational condition**, Nonlinear Dynamics 70: (2) 1407-1420. (2012)

[271] Zhang, E., Antoni, J., Pintelon, R., Schoukens, J., **Fast detection of system non-linearity using nonstationary signals**, Mechanical Systems and Signal Processing. 24, p. 2065-2075, 2010.

# Appendix A.

## A.1 Proof of Th 2.5.4 (For the arbitrary polynomial static non-linear system the relative variance $v_{BLA}$ yields minimum for the cubic system.)

**Proof:** Let the system now be: $y(t) = u(t) + \sum_{\alpha=3,\alpha\,odd}^{2K+1} c_\alpha\, u^\alpha(t)$. The only available knowledge is the assumption on the highest order of the non-linearity $2K+1$. The relative variance will depend in consequence upon the order of the system and the values of its coefficients. Although $v_B$ cannot be measured directly, a $v_{BLA}$ variance (2.5.13) relative to the Best Linear Approximation system can be measured instead, serving as a useful empirical constraint. Using $N(0,1)$ noise as the input signal and performing derivation similar to that in the proof of the Th. 2.5.3. we obtain:

$$v_{BLA} = \frac{\sum_{\alpha=3,odd}^{2K+1}\sum_{\beta=3,odd}^{2K+1} c_\beta\, B_{\alpha\beta}\, c_\beta}{(1+\sum_{\alpha=3,odd}^{2K+1} c_\alpha\, a_\alpha)^2} = \frac{(\mathbf{Bx},\mathbf{x})}{[1+(\mathbf{a},\mathbf{x})]^2} \qquad \text{(A.1.1)}$$

where $\mathbf{x} = [c_3\, c_5 \cdots c_K]$, $a_\alpha = [\mathbf{a}]_\alpha = \alpha!!$, $B_{\alpha\beta} = [\mathbf{B}]_{\alpha\beta} = (\alpha+\beta-1)!!–\alpha!!\beta!!$, (A.1.2)

the vectors are $1\times K$ dimensional, the matrix is $K \times K$ dimensional, the terms in the round brackets are scalar product.

The measured value of $v_{BLA}$ constraints now the possible values of $\mathbf{x}$ (i.e. system coefficients), among which the worst-case is sought for the relative bias $(\mathbf{a},\mathbf{x})$. The problem can be formulated as a constrained variational problem of finding the extremes of:

$$F(\mathbf{x}) = (\mathbf{a},\mathbf{x}) + \lambda((\mathbf{Bx},\mathbf{x}) - v_{BlA}(1+(\mathbf{a},\mathbf{x}))^2, \qquad \text{(A.1.3)}$$

where $\lambda$ is a Lagrange multiplier. Differentiating $F(\mathbf{x})$ with respect to $\boldsymbol{x}$ yields:

$$\mathbf{a} + 2\lambda\mathbf{B}\mathbf{x} - 2\lambda v_{BLA}\mathbf{a}(1+(\mathbf{a},\mathbf{x})) = 0. \qquad \text{(A.1.4)}$$

Defining $\mu = v_{BLA} - (2\lambda)^{-1}$, $\mathbf{a}\mathbf{a}' = \mathbf{A}$, we get: $\mathbf{B}\mathbf{x} = \mu\mathbf{a} + v_{BLA}\mathbf{A}\mathbf{x}$, and $\mathbf{x} = \mu(\mathbf{B} - v_{BLA}\mathbf{A})^{-1}\mathbf{a} = \mu\mathbf{x}_0$, (A.1.5)

assuming that the inverse exists. Substituting this solution into (A.1.1): $v_{BLA} = \frac{\mu^2(\mathbf{Bx}_0,\mathbf{x}_0)}{[1+\mu(\mathbf{a},\mathbf{x}_0)]^2}$, (A.1.6)

it yields a quadratic equation in $\mu$: $B\mu^2 = v_{BLA}(1+\mu A)^2$, (A.1.7)

where $B = (\mathbf{Bx}_0,\mathbf{x}_0)$ and $A = (\mathbf{a},\mathbf{x}_0)$ for simplicity. Assuming that $B > 0$, there are two roots:

$$\mu_{1,2} = \frac{v_{BLA}\,A \pm \sqrt{v_{BLA}\,B}}{B - v_{BLA}A^2}. \qquad \text{(A.1.8)}$$

The two solutions for $\mu$, and consequently for $\mathbf{x}$: $\mathbf{x}_{1,2} = \mu_{1,2}\mathbf{x}_0$, belong to the minimum and the maximum of the relative bias $(\mathbf{a}, \mathbf{x})$: $\mu_{1,2}A$. Further computation yields:

$$\min G_B/G_1 = \varepsilon_1 = -\frac{1}{(\tau+1)}, \quad \max G_B/G_1 = \varepsilon_2 = \frac{1}{(\tau-1)}, \qquad \text{(A.1.9)}$$

where: $\tau^2 = (\mathbf{Bx}_0,\mathbf{x}_0)/(v_{BlA}\times(\mathbf{a},\mathbf{x}_0)^2) = B/A^2 v_{BLA}$ (A.1.10)

The analysis of (A.1.9) for the increasing order of the highest non-linearity shows, that in agreement with (2.5.11) the cubic system means the worst-case and the largest bias. With $\varepsilon_1$ and $\varepsilon_2$ computed from the measured level of the variance $v_{BLA}$, the $G_1$ can now be bounded under the worst-case assumption by the measured $G_{BLA}$ as:

$G_{BLA}(1+\varepsilon_2) < G_1 < G_{BLA}(1+\varepsilon_1)$, or in a more straightforward way as: $G_{BLA}(1-\kappa) < G_1 < G_{BLA}(1+\kappa)$ (A.1.11)

with the bounding term $\kappa$ computed under the worst-case cubic assumption as:

$$\mathbf{x}_0 = a_3/(B_{33} - \nu_{BLA} a_3^2), \quad (\mathbf{B}\mathbf{x}_0, \mathbf{x}_0) = B_{33} \times a_3^2/(B_{33} - \nu_{BLA} a_3^2)^2$$
$$(\mathbf{a}, \mathbf{x}_0)^2 = a_3^4/(B_{33} - \nu_{BLA} a_3^2)^2 \qquad \text{(A.1.12)}$$
$$\tau^2 = (\mathbf{B}\mathbf{x}_0, \mathbf{x}_0)/(\nu_{BLA} \times (\mathbf{a}, \mathbf{x}_0)^2) = B_{33}/(\nu_{BLA} a_3^2)$$
$$\kappa = 1/\tau = a_3 \sqrt{\nu_{BLA}/B_{33}}$$

with: $a_3$ = 3!!, $B_{33}$ = ξ(k) (5!! – 3!! 3!!) = 6 ξ(k) (see [6*])

$\kappa = 3\sqrt{\nu_{BLA}/6\xi(k)}$, the best value of $\xi(k)$ is 3/4 (see (15) in [6*]) yielding: $\kappa = \sqrt{2}\sqrt{\nu_{BLA}}$ (A.1.13)

■

## A.2 Proof of Th 3.2.1 (Random multisines are in the limit normally distributed and separable signals)

**Proof:**

**A. Normal distribution**

First we will prove that the amplitude of the colored random multisines is normally distributed as $N(0, \sqrt{\frac{1}{2}\sum_{k=1}^{M} U_k^2})$. For the purpose of the proof let the random multisine be as:

$$u(t) = \sum_{k=1}^{M} U_k \cos(\omega_k t + \varphi_k) \qquad \text{(A.2.1)}$$

Then: $E\{u(t) = \sum_{i=1}^{M} U_k \cos(\omega_k t + \varphi_k)\} = \sum_{i=1}^{M} U_k E\{\cos(\omega_k t + \varphi_k)\} = 0.$

The correlation function of (A.2.1) is (computation is trivial):

$$R_u(\tau) = \frac{1}{2}\sum_{k=1}^{M} U_k^2 \cos(\omega_k \tau), \quad R_u(0) = \frac{1}{2}\sum_{k=1}^{M} U_k^2 \qquad \text{(A.2.2)}$$

Remember that a normalized random multisine is colored if its amplitude spectrum $U_k \neq$ const, but the behavior of the amplitudes can still be modeled as $O(M^{-1/2})$. Please note that for normalized random multisines the variance is of order $O(1)$. On the amplitude distribution we have:

### Lemma 3.2.1: The Central Limit Theorem holds for the sum of independent cosine random variables.

**Proof:** Let us check the Lyapunov Condition [10], i.e. that:

$$\exists \delta > 0, \quad \frac{\sum_{k=1}^{M} E\{|x_k - \mu_k|^{2+\delta}\}}{s_M^{2+\delta}} \xrightarrow[N\to\infty]{} 0 \qquad \text{(A.2.3)}$$

where in our case $x_k = U_k \cos\varphi_k$, $\mu_k = 0$, $s_M^2 = \frac{1}{2}\sum_{k=1}^{M} U_k^2$, and the phases are uniformly distributed on [0, 2π[. Considering, that $s_M^2$ is of order $O(1)$, it is enough to check the limit of the numerator of (A.2.3):

$$E\{|x_k|^{2+\delta}\} = \frac{1}{\pi}\int_{-U_k}^{U_k} \frac{|y|^{2+\delta}}{\sqrt{U_k^2 - y^2}} dy = \frac{2}{\pi}\int_0^{U_k} \frac{y^{2+\delta}}{\sqrt{U_k^2 - y^2}} dy = \frac{2U_k^{2+\delta}}{\pi}\int_0^{U_k} \frac{(y/U_k)^{2+\delta}}{\sqrt{1-(y/U_k)^2}} d(y/U_k) =$$
$$\frac{2U_k^{2+\delta}}{\pi}\int_0^1 \frac{z^{2+\delta}}{\sqrt{1-z^2}} dz = \frac{2U_k^{2+\delta}}{\pi}\int_0^{\pi/2} \frac{\sin^{2+\delta}\omega}{\cos\omega}\cos\omega d\omega = \frac{2U_k^{2+\delta}}{\pi}\int_0^{\pi/2} \sin^{2+\delta}\omega d\omega \approx \frac{O(1)}{M^{\frac{2+\delta}{2}}} = \frac{O(1)}{M^{1+\delta/2}} \qquad \text{(A.2..4)}$$

The summation in the numerator in (A.2.3) leaves still an order of $\frac{O(1)}{M^{\delta/2}}$, which tends to zero for any $\delta > 0$.

**B. Asymptotic separability**

Signal is called separable if [148, 59]:

$$\delta(t,\tau) = E\{u(t-\tau)|u(t)\} - a(\tau)u(t) = 0 \tag{A.2.5}$$

or after Fourier transform (in terms of characteristic functions):

$$\Delta(\xi,\tau) = G_{u,1}(\xi,\tau) - a(\tau)\, f'_{u,1}(\xi) = 0 \tag{A.2.6}$$

where [148, 59]:

$$G_{u,1}(\xi,\tau) = jE\{u(t-\tau)e^{j\xi u(t)}\} \tag{A.2.7}$$

$$f'_{u,1}(\xi) = jE\{u(t)e^{j\xi u(t)}\} \tag{A.2.8}$$

$$a(\tau) = R_u(\tau)/R_u(0) \tag{A.2.9}$$

For the colored random multisines (A.2.7-A.2.8) becomes:

$$\begin{aligned} &G_{u,1}(\xi,\tau) = jE\{u(t-\tau)e^{j\xi u(t)}\} = \\ &j\sum_{k=1}^{M} U_k \cos(\omega_k\tau) E\{\cos(\omega_k t+\varphi_k)e^{j\xi\sum_{m=1}^{N} U_m \cos(\omega_m t+\varphi_m)}\} + \\ &j\sum_{k=1}^{M} U_k \sin(\omega_k\tau) E\{\sin(\omega_k t+\varphi_k)e^{j\xi\sum_{m=1}^{N} U_m \cos(\omega_m t+\varphi_m)}\} \end{aligned} \tag{A.2.10}$$

$$\begin{aligned} &G_{u,1}(\xi,\tau) = \\ &j\sum_{k=1}^{M} U_k \cos(\omega_k\tau) E\{\cos(\omega_k t+\varphi_k)e^{j\xi U_k \cos(\omega_k t+\varphi_k)}\}\prod_{m=1,m\neq k}^{M} E\{e^{j\xi U_m \cos(\omega_m t+\varphi_m)}\} + \\ &j\sum_{k=1}^{M} U_k \sin(\omega_k\tau) E\{\sin(\omega_k t+\varphi_k)e^{j\xi U_k \cos(\omega_k t+\varphi_k)}\}\prod_{m=1,m\neq k}^{M} E\{e^{j\xi U_m \cos(\omega_m t+\varphi_m)}\} \end{aligned} \tag{A.2.11}$$

The expected values are:

$$\begin{aligned} &(a)\ E\{e^{j\xi U_m \cos(\omega_m t+\phi_m)}\} = J_0(U_m\xi) \\ &(b)\ E\{\cos(\omega_k t+\phi_k)\, e^{j\xi U_k \cos(\omega_k t+\phi_k)}\} = jJ_1(U_k\xi) \\ &(c)\ E\{\sin(\omega_k t+\phi_k)\, e^{j\xi U_k \cos(\omega_k t+\phi_k)}\} = 0 \end{aligned} \tag{A.2.12}$$

(a) is by definition of the Bessel function (expected value integrated for the uniform density). For (c) see the argument in [59]. In case of (b) consider:

$$\begin{aligned} &E\{\cos(\omega_k t+\varphi_k)e^{j\xi U_k \cos(\omega_k t+\varphi_k)}\} = \frac{1}{2\pi}\int_0^{2\pi}\cos(\varphi)e^{j\xi\, U_k\cos(\varphi)}d\varphi = \\ &\frac{1}{2\pi}\left(\int_0^{\pi}\cos(\varphi)e^{j\xi\, U_k\cos(\varphi)}d\varphi - \int_0^{\pi}\cos(\varphi)e^{-j\xi U_k\cos(\varphi)}d\varphi\right) = \\ &\frac{j}{\pi}\int_0^{\pi}\cos(\varphi)\sin(\xi U_k\cos(\varphi))\,d\varphi = jJ_1(U_k\xi) \end{aligned} \tag{A.2.13}$$

This last line is due to the decomposition [107]:

$$\sin(z\cos\varphi) = 2\sum_{n=0}^{\infty}(-1)^n J_{2n+1}(z)\cos(2n+1)\varphi \tag{A.2.14}$$

with both sides multiplied by $\cos\varphi$ and integrated from 0 to $\pi$. We follow now with:

$$G_{u,1}(\xi,\tau) = -\sum_{k=1}^{M} U_k \cos(\omega_k\tau) J_1(U_k\xi)\prod_{m=1,m\neq k}^{M} J_0(U_m\xi) = -\sum_{k=1}^{M} U_k F_k(\xi)\cos(\omega_k\tau) \tag{A.2.15}$$

where $F_k$ designates the product of all Bessel functions. Similarly:

$$f'_{u,1}(\xi,\tau) = jE\left\{u(t)e^{j\xi u(t)}\right\} =$$
$$j\sum_{k=1}^{M} U_k E\left\{\cos(\omega_k t+\varphi_k)e^{j\xi\sum_{m=1}^{N}U_m\cos(\omega_m t+\varphi_m)}\right\} = \quad \text{(A.2.16)}$$
$$-\sum_{k=1}^{M} U_k J_1(A_k\xi)\prod_{m=1,m\neq k}^{M} J_0(U_m\xi) = -\sum_{k=1}^{M} U_k F_k(\xi)$$

and

$$\Delta(\xi,\tau) = -\sum_{k=1}^{M} U_k F_k(\xi)\cos(\omega_k\tau) + a(\tau)\sum_{k=1}^{M} U_k F_k(\xi) \quad \text{(A.2.17)}$$

Now let introduce the correlation function (A.2.2) into (A.2.17):

$$\Delta(\xi,\tau) = -\frac{\frac{1}{2}\sum_{l=1}^{M} U_l^2 \sum_{k=1}^{M} U_k F_k(\xi)\cos(\omega_k\tau) - \frac{1}{2}\sum_{l=1}^{M} U_l^2\cos(\omega_l\tau)\sum_{k=1}^{M} U_k F_k(\xi)}{\frac{1}{2}\sum_{l=1}^{M} U_l^2} \quad \text{(A.2.18)}$$

considering that the denominator (i.e. the variance of the multisine) is of constant order with respect to *M*, it is enough to investigate the behavior of the numerator. We will prove, that the colored random multisine is asymptotically separable in the sense, that:

$$\lim_{M\to\infty}\Delta(\xi,\tau) = 0 \quad \text{(A.2.19)}$$

and consequently (via the Fourier transform)

$$\lim_{M\to\infty}\delta(t,\tau) = 0 \quad \text{(A.2.20)}$$

Let $D(\xi,\tau)$ denote the numerator of (A.2.18). With the rearrangement of terms it can be written as: $D(\xi,\tau) = \sum_{l=1}^{M}\sum_{k=1}^{M} U_l^2 U_k F_k(\xi)\frac{1}{2}(\cos(\omega_k\tau)-\cos(\omega_l\tau))$ (A.2.21)

By rearranging of terms the (A.2.21) can be rewritten as:

$$D(\xi,\tau) = \sum_{l=1}^{M-1}\sum_{k=l+1}^{M} U_l U_k \frac{1}{2}(\cos(\omega_l\tau)-\cos(\omega_k\tau))(U_l F_k - U_k F_l), \quad \text{(A.2.22)}$$

we will estimate the order of the $U_k F_l - U_l F_k$ term (the sums introduce order $M^2$, and the other terms under the sums are of order $M^{-1}$). Remember that:

$$F_k(\xi) = J_1(U_k\xi)\prod_{m=1,m\neq k}^{M} J_0(U_m\xi) \quad \text{(A.2.23)}$$

$J_0(U_m\xi)$ is the characteristic function of a $U_m\cos(\varphi)$ random variable. The product of the 0th order Bessel functions is the characteristic function of the sum of the independent cosine random variables, which due to the Central Limit Theorem tends to the normal distribution (see first part of the proof, consider also that the number of the variables equals the number of the frequencies in the signal, consequently this number is large and the Gaussian approximation good). Conversely the characteristic function of a Gaussian random variable is also Gaussian, i.e., see Lemma 3.2.2:

$$\prod_{m=1,m\neq k}^{M} J_0(U_m\xi) = e^{-\frac{1}{2}\sigma_k^2\xi^2} + O(M^{-1/2}), \text{ with } \quad \sigma_k^2 = \frac{1}{2}\sum_{m=1,m\neq k}^{M} U_m^2 \quad \text{(A.2.24)}$$

Furthermore for small values of the argument (we investigate the case when $U_m\xi \to 0$): $J_1(U_k\xi) \cong \frac{U_k\xi}{2}$ (A.2.25)

With this approximation:

$$U_k F_l - U_l F_k = U_k\frac{U_l\xi}{2}e^{-\frac{1}{2}\sigma_l^2\xi^2} - U_l\frac{U_k\xi}{2}e^{-\frac{1}{2}\sigma_k^2\xi^2} = \frac{U_k U_l\xi}{2}(e^{-\frac{1}{2}\sigma_l^2\xi^2} - e^{-\frac{1}{2}\sigma_k^2\xi^2}) \quad \text{(A.2.26)}$$

Let us introduce:

$$\sigma^2 = \frac{1}{2}\sum_{m=1}^{M} U_m^2 = \sigma_k^2 + \frac{1}{2}U_k^2 = \sigma_l^2 + \frac{1}{2}U_l^2 \tag{A.2.27}$$

The (A.2.26) can be now written as:

$$U_k F_l - U_l F_k = \frac{U_k U_l \xi}{2} e^{-\frac{1}{2}\sigma^2\xi^2} (e^{\frac{1}{2}U_l^2\xi^2} - e^{\frac{1}{2}U_k^2\xi^2}) =$$
$$\frac{U_k U_l \xi}{2} e^{-\frac{1}{2}\sigma^2\xi^2} (1 + \frac{1}{2}U_l^2\xi^2 + \frac{1}{2!}(\frac{1}{2}U_l^2\xi^2)^2 + ... \quad -1 - \frac{1}{2}U_k^2\xi^2 - \frac{1}{2!}(\frac{1}{2}U_k^2\xi^2)^2 - ...) \cong \tag{A.2.28}$$
$$\frac{U_k U_l \xi}{2} e^{-\frac{1}{2}\sigma^2\xi^2} (\frac{1}{2}U_l^2\xi^2 - \frac{1}{2}U_k^2\xi^2)$$

With (A.2.28) the numerator (A.2.22) becomes:

$$D(\xi,\tau) = \xi^3 \sum_{l=1}^{M}\sum_{k=l+1}^{M} U_l^2 U_k^2 \frac{1}{8}(\cos(\omega_k\tau) - \cos(\omega_l\tau))\, e^{-\frac{1}{2}\sigma^2\xi^2} (U_l^2 - U_k^2) \tag{A.2.29}$$

the terms under the sums are of order $O(M^{-3})$ yielding the overall order of (A.2.29) as $O(M^{-1})$:

$$D(\xi,\tau) \approx O(M^{-1}) \tag{A.2.30}$$

$$\Delta(\xi,\tau) = -\frac{D(\xi,\tau)}{R_u(0)} \approx \frac{O(M^{-1})}{O(M^0)} = O(M^{-1}) \tag{A.2.31}$$

proving the Theorem. ■

**Note**: The practical order of $\Delta(\xi,\tau)$ is even smaller, because the computation does not account for the fact that the term (A.2.26) is bipolar and the summation in (3.23) has an averaging character. Simulation indicates decreasing of an order $O(M^{-3/2})$.

## Lemma 3.1: Approximation (A.2.24) holds.

**Proof:** To see that the approximation (A.2.24) holds, consider that the Taylor series for $J_0(z)$ and for $e^{-z^2/4}$ coincide in the first two terms [107]:

$$J_0(z) = \sum_{m=0}^{\infty} \frac{(-1)^m (z/2)^{2m}}{(m!)^2} = 1 - \frac{z^2}{2^2(1!)^2} + \frac{z^4}{2^4(2!)^2} \mp ...$$
$$e^{-\frac{1}{4}z^2} = e^{-(z/2)^2} = \sum_{m=0}^{\infty} \frac{(-1)^m (z/2)^{2m}}{m!} = 1 - \frac{z^2}{2^2 1!} + \frac{z^4}{2^4 2!} \mp ... \tag{A.2.32}$$

thus: $J_0(z) = e^{-\frac{1}{4}z^2} + R(z), \quad |R(z)| \approx O(z^4)$ (A.2.33)

With $z = U_k\xi$ the remainder becomes of order O($M^{-2}$), i.e.: $J_0(U_k\xi) = e^{-\frac{1}{2}(U_k^2/2)\xi^2} + R_k, \; R_k \approx O(M^{-2})$ (A.2.34)

and

$$\prod_k J_0(U_k\xi) = \prod_k (e^{-\frac{1}{2}(U_k^2/2)\xi^2} + R_k) = \prod_k e^{-\frac{1}{2}(U_k^2/2)\xi^2} (1 + \sum_k R_k + \sum_k\sum_l R_k R_l + ...) =$$
$$e^{-\frac{1}{2}\sum_k (U_k^2/2)\xi^2} (1 + O(M^{-1})) \tag{A.2.35}$$

■

## A.3 Proof of Theorem 4.5.1 (Bias of the general Volterra MIMO system)

**Proof:** In the general case the experiments are made using different (independent phase) realizations of the input random multisines:

$$\mathbf{B}(l)=\left[b_{ij}(l)\right]=\mathbf{U}_N^{-1}(l)=\left[\frac{V_{ji}}{D}\right], \tag{A.3.1}$$

where $D=\det\mathbf{U}_N$ and the $V_{ji}$ are elements of the adjoint matrix $\boldsymbol{adj}\,\mathbf{B}(l)$. Due to the properties of the minors we have (Theorems 13.5.1-2, or 13.5.3, [84] pp. 189-192):

$$\sum_{i=1}^{N}U_j^{(i)}\frac{V_{ik}}{D}=\sum_{i=1}^{N}U_i^{(j)}\frac{V_{ki}}{D}=\sum_{i=1}^{N}U_j^{(i)}b_{ki}=\sum_{i=1}^{N}U_i^{(j)}b_{ik}=\delta_{jk} \tag{A.3.2}$$

where $\delta_{jk}$ is the Kronecker symbol. The general model for the system output in the *i*th experiment is now:

$$\begin{aligned}
&Y^{(i)}(l)=G^1(l)U_1^{(i)}(l)+...+G^k(l)U_k^{(i)}(l)+...+G^N(l)U_N^{(i)}(l)+\\
&+\sum_k G^{11}(k,L)U_1^{(i)}(k)U_1^{(i)}(L)+\sum_k G^{12}(k,L)U_1^{(i)}(k)U_2^{(i)}(L)+...+\sum_k G^{NN}(k,L)U_N^{(i)}(k)U_N^{(i)}(L)+\\
&+\sum_{k_1}\sum_{k_2}G^{111}(k_1,k_2,L)U_1^{(i)}(k_1)U_1^{(i)}(k_2)U_1^{(i)}(L)+\sum_{k_1}\sum_{k_2}G^{112}(k_1,k_2,L)U_1^{(i)}(k_1)U_1^{(i)}(k_2)U_2^{(i)}(L)+...\\
&+\sum_{k_1}\sum_{k_2}G^{1NN}(k_1,k_2,L)U_1^{(i)}(k_1)U_N^{(i)}(k_2)U_N^{(i)}(L)+...+\sum_{k_1}\sum_{k_2}G^{NNN}(k_1,k_2,L)U_N^{(i)}(k_1)U_N^{(i)}(k_2)U_N^{(i)}(L)+\\
&...\\
&+\sum_{k_1}\sum_{k_2}...\sum_{k_{\alpha-1}}G^{j_1j_2\cdots j_\alpha}(k_1,k_2,...,k_{\alpha-1},L)U_{j_1}^{(i)}(k_1)U_{j_2}^{(i)}(k_2)...U_{j_\alpha}^{(i)}(L)+...
\end{aligned} \tag{A.3.3}$$

where $L = l - k$ for the 2[d] order, $L = l - k_1 - k_2$ for the 3[rd] order kernels, etc., $\alpha$ is the order of the kernel and the sums run over positive and negative frequency grid $S_M^-\cup S_M^+$. To evaluate the bias (and the non-linear noise) on the FRF, true non-linear outputs (A.3.3) must be substituted into (4.5.2). The FRF estimates (4.5.2) are:

$$\hat{G}^k(l)=b_{k1}(l)Y^{(1)}(l)+...+b_{kN}(l)Y^{(N)}(l)=\sum_{i=1}^{N}b_{ki}(l)\,Y^{(i)}(l) \tag{A.3.4}$$

Now we will consider the influence on (A.3.4) of a single kernel. Linear kernel is treated as a special case of an odd order kernel. A single $\alpha$-th order kernel contribution to (A.3.4) is:

$$\begin{aligned}
&C=\sum_{i=1}^{N}b_{ki}(l)\times\sum_{k_1}\sum_{k_2}...\sum_{k_{\alpha-1}}G^{j_1j_2\cdots j_\alpha}(k_1,k_2,...,k_{\alpha-1},L)U_{j_1}^{(i)}(k_1)U_{j_2}^{(i)}(k_2)...U_{j_\alpha}^{(i)}(L)=\\
&\sum_{k_1}\sum_{k_2}...\sum_{k_{\alpha-1}}G^{j_1j_2\cdots j_\alpha}(k_1,k_2,...,k_{\alpha-1},L)\left(\sum_{i=1}^{N}U_{j_1}^{(i)}(k_1)U_{j_2}^{(i)}(k_2)...U_{j_\alpha}^{(i)}(L)b_{ki}(l)\right)
\end{aligned} \tag{A.3.5}$$

For random inputs (A.3.5) is a random variable. We will analyze conditions for (A.3.5) to have zero mean value. In the following it will be enough to investigate only the random variable component from (A.3.5), i.e.:

$$V=\sum_{i=1}^{N}U_{j_1}^{(i)}(k_1)U_{j_2}^{(i)}(k_2)...U_{j_\alpha}^{(i)}(L)b_{ki}(l) \tag{A.3.6}$$

**a. Odd order kernels**

When the kernel input indices $j_1 j_2 \ldots j_\alpha$ contain inputs grouped in pairs except one (e.g. 11233), then the inputs with the same index can be paired with frequencies (*k*, -*k*) (there is no such pairing term for linear kernels).

Considering that in the experiments only the random phases of the inputs vary, the paired inputs will be independent from the experiment index and can be extracted from the inner sum as:

$$U_j^{(i)}(k)U_j^{(i)}(-k) = \left|U_j(k)\right|^2 \tag{A.3.7}$$

let $\alpha' = (\alpha\text{-}1)/2$), then by (A.3.2):

$$V = \left|U_{j_{P1}}(k_1)\right|^2 ... \left|U_{j_{P\alpha'}}(k_{\alpha'})\right|^2 \times \sum_{i=1}^{N} U_{j_S}^{(i)}(l)\, b_{ki}(l) = \left|U_{j_{P1}}(k_1)\right|^2 ... \left|U_{j_{P\alpha'}}(k_{\alpha'})\right|^2 \times \delta_{j_S k} \tag{A.3.8}$$

where the index $j_S$ denotes the single non-paired input. Consequently (A.3.5) becomes:

$$C = K \sum_{k_1}\sum_{k_2}...\sum_{k_{\alpha'}} G^{j_1 j_2 ... j_\alpha}(k_1, -k_1, ..., l)\ \left|U_{j_{P1}}(k_1)\right|^2 ... \left|U_{j_{P\alpha'}}(k_{\alpha'})\right|^2 \times \delta_{j_S k} \tag{A.3.9}$$

where $K$ counts the various ways the frequencies can be paired in inputs of the same index $j_{P1}...j_{P\alpha'}$ (e.g. $K = 3$ for the kernel $G^{111}$, and $K = 1$ for the kernel $G^{122}$), and the sums run over the grid $S_M^+$. For $j_S = k$ it yields the usual bias term of order $O(1)$. This kernel yields nonzero contribution to the expected value of $\hat{G}^k(l)$, when the $k$th input appears an odd number of times among the kernel inputs and other inputs appear pair wise (i.e. even number of times).

When the indices $j_1 j_2 \ldots j_\alpha$ contain more single inputs (e.g. $m$ pairs and $P = \alpha - 2m$ single inputs, consider also $m = 0$ as a special case), then the contribution of the kernel can be written as:

$$\begin{aligned} K' \sum_{k_1}...\sum_{k_m}\sum_{k_{m+1}}...\sum_{k_{M-1}} G^{j_1 j_2 ... j_\alpha}(k_1, -k_1, ..., k_m, -k_m, k_{m+1}, ..., L) \times \left|U_{j_{P1}}(k_1)\right|^2 ... \left|U_{j_{Pm}}(k_m)\right|^2 \\ \times \sum_{i=1}^{N} U_{j_{S_1}}^{(i)}(k_{m+1}) ... U_{j_{S_P}}^{(i)}(L)\, b_{ki}(l) \end{aligned} \tag{A.3.10}$$

where the input pairs with frequencies paired yield deterministic term, and only the last sum contains random variables. Due to the different input indices $j_{S_1} \neq j_{S_2} \neq \cdots \neq j_{S_P}$, pairing frequencies in the last sum is of no use (random phases in $U_j^{(i)}(k)$ are independent for different frequencies, input and experiment indices). Consequently the expected value of this sum and thus of the whole contribution is zero. These random variables contribute however fully to the non-linear noise variance.

**b. Even order kernels**

For even kernels, (A.3.6) contains an odd number of terms. Pairing frequencies for signals with different input indices does not get rid of the randomness. Even if input signals with the same indices are present, pairing their frequencies reduces the random part of (A.3.6) to at most:

$$V = \sum_{i=1}^{N} U_{j_1}^{(i)}(k) U_{j_2}^{(i)}(l-k)\, b_{ki}(l) \tag{A.3.11}$$

which, due to the different frequencies, is zero mean, whatever the signal indices are. These random variables contribute fully to the non-linear noise variance. The overall bias to the FRF estimate is given then by:

$$E\left\{\hat{G}^k(l)\right\} = G^k(l) + \sum_{\beta} G_B^{\beta}(l) \tag{A.3.12}$$

where the sum applies to (A.3.9)-like terms of suitable order (see earlier).

The last step is to verify that the infinite series of bias terms (4.5.5) is convergent. It is so due to the inequality: $|E\{x\}| \leq E\{|x|\}$, and the convergence of the series (4.1.3). ■

## A.4 Proof of the Th 5.1.1

**Proof:** In the general MIMO case the input matrix is: $\mathbf{U}(l) = \mathbf{T}_N \, U_1^{(1)}(l)$ (A.4.1)

where $\mathbf{T}_N = \mathbf{H}_N$ is so called Hadamard matrix of order $N = 2^K$, or $\mathbf{T}_N = \mathbf{W}_N$ is DFT matrix of an arbitrary order $N$. Based on the experience of the optimal inputs (4.3.8, 4.3.19, 4.3.20) in the TITO case, we consider now unitary inputs:

$$\mathbf{B}(l) = \left[b_{ij}(l)\right] = \frac{1}{N|U|^2}\mathbf{U}^*(l) = \frac{1}{N|U|^2}\mathbf{T}^*\,\overline{U}(l) = \frac{1}{N|U|^2}\left[\overline{w}_{ji}\right]\overline{U}(l) \tag{A.4.2}$$

(normalized, transposed, complex conjugate of the input), also fulfilling (4.5.3, A.3.2), where the input signals for later experiments are some combinations of the complex amplitudes of the multisines used in the first experiment. The condition (4.5.3, A.3.2) can be written now as:

$$\sum_{i=1}^{N}\overline{U}_j^{(i)}U_k^{(i)} = \sum_{i=1}^{N}U_i^{(j)}\overline{U}_i^{(k)} = \sum_{i=1}^{N}w_{ij}\,\overline{w}_{ki}\left|U(l)\right|^2 = \sum_{i=1}^{N}\overline{w}_{ji}\,w_{ik}\left|U(l)\right|^2 = N\left|U\right|^2\delta_{jk} \tag{A.4.3}$$

**a. Odd order kernels:**

We will investigate now (A.3.5), taking into account (4.5.3, A.3.2) and that:

$$U_j^{(i)}(-l) = \overline{U}_j^{(i)}(l) = \frac{1}{N\left|U\right|^2}\overline{w}_{ij}\,\overline{U}(l) \tag{A.4.4}$$

With (A.4.4), (A.3.6) can be written as: $V = \underbrace{\left(\sum_{i=1}^{N} w_{ij_1}w_{ij_2}\ldots w_{ij_\alpha}\,\overline{w}_{ik}\right)}_{A}\underbrace{\left(U(k_1)U(k_2)\ldots\overline{U}(l)\right)}_{B}$ (A.4.5)

It is important to see now, that the contribution of (A.4.5) to the estimate (4.5.4) can be:

- **systematic**, if $A \neq 0$, $E\{B\} \neq 0$, contributing to the bias,
- **zero mean random**, if if $A \neq 0$, $E\{B\} = 0$, contributing to the non-linear variance, or
- **simply nonexistent**, if $A = 0$ (this is the most interesting case to investigate, because it will provide the way to decrease the variance of the estimate and to shorten the measurement time).

The expected value of $B$ will be nonzero, if all frequencies are paired ($B$ contains even number of terms, so it is possible). Otherwise the expected value of $B$ will be zero (due to independent phases at different frequencies).

To evaluate $A$ let consider that for Fourier matrix: $w_{ij} = e^{-j\frac{2\pi}{N}(i-1)(j-1)}$, $\quad i,\ j = 1 \ldots N$.

$$A = \sum_{i=1}^{N} w_{ij_1}w_{ij_2}\ldots w_{ij_\alpha}\,\overline{w}_{ik} = \sum_{i=1}^{N} e^{\pm j\frac{2\pi}{N}(i-1)(j_1-1)}e^{\pm j\frac{2\pi}{N}(i-1)(j_2-1)}\ldots e^{\pm j\frac{2\pi}{N}(i-1)(j_\alpha-1)}e^{j\frac{2\pi}{N}(i-1)(k-1)} =$$

$$= \sum_{i=1}^{N} e^{-j\frac{2\pi}{N}(i-1)\left[\pm(j_1-1)\pm(j_2-1)\pm\ldots\pm(j_\alpha-1)-(k-1)\right]} = \sum_{i=1}^{N} e^{-j\frac{2\pi}{N}(i-1)\left(\sum_{m=1}^{\alpha}\pm(j_m-1)-k+1\right)} = \sum_{i=1}^{N} e^{-j\frac{2\pi}{N}(i-1)Z} \tag{A.4.6}$$

$$Z = \sum_{m=1}^{\alpha}\pm\left(j_m-1\right)-k+1) \tag{A.4.7}$$

depending on which frequencies are paired.

1. $A = N$, if $Z = 0$ (mod $N$), otherwise $A = 0$. It means that for a number of input combinations the contribution of the kernel to the variance will be nonexistent.

2. If all frequencies are paired ( $E\{B\} \neq 0$ ) and all inputs to the kernel but one (let it be $j_S$) appear in pairs, then $Z = j_S - k$, and $Z = 0$, if $j_S = k$. These are the usual bias terms measured with general inputs, see Theorem 4.5.1.

3. If $N > 2$ and the inputs do not appear in pairs, the frequencies however are paired, it can happen that in the same time $A \neq 0, \quad E\{B\} \neq 0$, i.e. the bias will contain more terms, than as in the case of general inputs (i.e. a bit more bias, but much less variance).

**Note**: For TITO systems already the cubic non-linearity requires one of the inputs to appear in pair, so there is no problem here (low complexity).

The same analysis applies to the Hadamard matrix, where: $w_{ij} = wal_{j-1}(i-1)$, $i, j = 1\ldots N = 2^K$, i.e. its rows are discrete values of the Walsh functions. Walsh functions posses group property, i.e.:

$$wal_{j-1}(i-1)wal_{k-1}(i-1) = wal_{(j-1)\oplus(k-1)}(i-1), \tag{A.4.8}$$

where $\oplus$ is so called dyadic addition, i.e. addition mod 2 without carry [1]:

$$0 \oplus k = k, \quad k \oplus k = 0, \text{ etc.} \tag{A.4.9}$$

Now:

$$A = \sum_{i=1}^{N} w_{ij_1} w_{ij_2} \ldots w_{ij_\alpha} \overline{w}_{ik} = \sum_{i=1}^{N} wal_{j_1-1}(i-1)\ldots wal_{j_\alpha-1}(i-1)wal_{k-1}(i-1) =$$
$$= \sum_{i=1}^{N} wal_{(j_1-1)\oplus(j_2-1)\oplus\ldots(j_\alpha-1)\oplus(k-1)}(i-1) = \sum_{i=1}^{N} Z \tag{A.4.10}$$

$$Z = (j_1-1)\oplus(j_2-1)\oplus\ldots(j_\alpha-1)\oplus(k-1) \tag{A.4.11}$$

Here we have the same 1.-3. cases, as before, although particular nonzero contributions can differ (please note however, that comparison between Fourier and Hadamard cases is meaningful only for $N = 2^K$).

**b. Even order kernels:**

For even order kernels the expected value of $B$ is zero, whatever the pairing of the frequencies (or not). The essential difference with respect to the general case is that for particular mixed kernels the value of $A$ can be zero, eliminating those kernels from the variance and thus decreasing the level of the non-linear noise source.

We can summarize the proof as follows. The orthogonalized inputs (A.4.2) work well for TITO systems, because there they provide the same level of the bias and considerably less non-linear variance (with respect to the general case). For systems with more inputs, higher order mixed kernels can provide situations, where the dependency between inputs introduced by the orthogonalization (A.4.2) will add nonzero terms to the bias, comparing to the general case. The problem can be traced to the fact, that pairing the frequencies for inputs with different indices can for such special matrices produce nonzero expected value:

$$E\{U_j^{(i)}(k)U_k^{(i)}(-k)\} = \frac{1}{N|U|^2} E\{w_{ij}U(k)\overline{w}_{ik}\overline{U}(k)\} = \frac{1}{N|U|^2} w_{ij}\overline{w}_{ik}U(k)\overline{U}(k) = \frac{1}{N} w_{ij}\overline{w}_{ik} \quad , \; j \neq k \tag{A.4.12}$$

■

## A.5 Proof of Theorem 5.4.1

**Proof**: To prove the equivalence of the input signals it is enough to compute a single ($\alpha$th order) kernel in (4.1.1-4.1.4). Bias on the measured FRF is the sum of all systematic contributions with nonzero expected values with respect to the random inputs. The non-linear noise variance comes from all other zero expected value stochastic contributions.

The aim is now to show that in every case kernels of exactly the same order and combination of inputs will 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 will lead in the limit $M \to \infty$ to exactly the same bias expressions.

For the analysis signal indices in the kernel (4.1.2-4.1.4) will be grouped together as:

$j_1, j_2, \ldots, j_\alpha \rightarrow j_1, j_1, \ldots, j_1, j_2, j_2, \ldots, j_2, \ldots, j_K, j_K, \ldots, j_K,$

where input $j_1$ appears in the kernel $M_1$ times, input $j_2$ $M_2$ times, …, etc., there are altogether $K$ different inputs to the kernel, and $\sum_{l=1}^{K} M_l = \alpha$, and $j_1 < j_2 < \ldots j_K$. Reference input index $k$ will be identified usually with input index $j_1$.

***A.1 Gaussian noise***

The FRF measured with Gaussian noise is ($H_1$-FRF):

$$G_R^k(j\omega_k) = G^k(j\omega_k) + G_B^k(j\omega_k) = \frac{S_{YU_k}(j\omega_k)}{S_{U_kU_k}(j\omega_k)} \tag{A.5.1}$$

To compute a single αth order ($\alpha = 2\beta$-1) contribution:

$$G_B^{j_1 \ldots j_\alpha}(j\omega_k) = \frac{S_{Y^{j_1 \ldots j_\alpha} U_k}(j\omega_k)}{S_{U_kU_k}(j\omega_k)} \tag{A.5.2}$$

we need the correlation:

$$R_{y^{j_1 \ldots j_\alpha} u_k}(\tau_0) = E\left\{y^{j_1 \ldots j_\alpha}(t) u_k(t-\tau_0)\right\} = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} g^{j_1 j_2 j \ldots j_\alpha}(\tau_1, \ldots, \tau_\alpha) E\{u_k(t-\tau_0) u_{j_1}(t-\tau_1) \ldots u_{j_\alpha}(t-\tau_\alpha)\} d\tau_1 \ldots d\tau_\alpha \tag{A.5.3}$$

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

$$E\{u_{j_1} u_{j_1} \ldots u_{j_\alpha}\} = E\{\underbrace{u_{j_1} u_{j_1} \ldots u_{j_1}}_{M_1+1}\} E\{\underbrace{u_{j_2} u_{j_2} \ldots u_{j_2}}_{M_2}\} \ldots E\{\underbrace{u_{j_K} u_{j_K} \ldots u_{j_K}}_{M_K}\} \tag{A.5.4}$$

Each of the expected values in (A.5.4) is zero for an odd number of terms.

For even number of terms they can be decomposed into sums of combinations of pair wise correlations $\Sigma \Pi \; R_{u_n u_m}(\tau_i - \tau_j)$, see [198, 162]. 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.

For the final form of (A.5.2) we must take into account that by the symmetry of the kernels every combination of pair wise correlations leads the same bias term. From (A.5.2) we have:

$$\Sigma\Pi\, R_{u_{j_1} u_{j_1}}(\tau_i - \tau_j) \times \Sigma\Pi\, R_{u_{j_2} u_{j_2}}(\tau_i - \tau_j) \times \ldots \times \Sigma\Pi\, R_{u_{j_K} u_{j_K}}(\tau_i - \tau_j) = \Pi\Sigma\Pi\, R_{u_l u_l}(\tau_i - \tau_j) \tag{A.5.5}$$

where the outmost product at the right side runs over distinct inputs to the kernel (and the sum-product within comes from the Schetzen-decomposition). The expression (A.5.5) can be written further as:

$$\prod_{inputs} \sum_{decomp} \prod R_{u_l u_l}(\tau_i - \tau_j) = \sum_{N_R} \prod_{inputs} \prod_{decomp} R_{u_l u_l}(\tau_i - \tau_j) \tag{A.5.6}$$

The number of possible combinations $N_R$ is the product of the numbers of combinations at the left side of (A.5.5), i.e.:

$$N_R = M_1!! \prod_{l=2}^{K} (M_l - 1)!! \tag{A.5.7}$$

Now we must introduce the frequency power spectrum via the Fourier transform, similarly to the SISO case. We will follow the philosophy of the derivation for SISO case [162], where the correlation is transformed into the bias term for a particular single partition in (A.5.5). For the SISO kernel of order α, the power spectrum

contained the input spectrum raised to the power α. In the MIMO case a kernel of an overall order α, behaves like an $M_l$ order kernel for the $j_l$ input signal ($M_1$+1 for $j_1$).

Other partitions yield exactly the same bias term, due to the symmetry of the Volterra kernel. The final result is thus a single computed bias term scaled up $N_R$ times. Let the particular partition be defined by:

$$\underbrace{(\tau_0,\tau_1)(\tau_2,\tau_3).....}_{j_1\ input}\ ....\ \underbrace{.....(\tau_{2\beta-2},\tau_{2\beta-1})}_{j_K\ input} \tag{A.5.8}$$

with this (and with a certain simplification in the notation):

$$R_{y\,j_1\cdots j_\alpha u_k}(\tau_0)=\int_{-\infty}^{\infty}\!\!\cdots\!\int_{-\infty}^{\infty} g^{j_1\cdots j_K}(\tau_1,...,\tau_{2\beta-1})\,e^{j\omega_1\tau_0}e^{-j\omega_1\tau_1}\times S_{U_kU_k}(j\omega_1)\prod_{r=2}^{\beta-1}S_{U_{j_r}U_{j_r}}(j\omega_r)e^{j\omega_r(\tau_{2r-2}-\tau_{2r-1})}\times d\tau_1...d\tau_{2\beta-1}df_1...df_\beta =$$

$$\int_{f_1=-\infty}^{\infty}\!\!\cdots\int_{f_\beta}S_{U_kU_k}(j\omega_1)e^{j\omega_1\tau_0}\prod_{r=2}^{\beta-1}S_{U_{j_r}U_{j_r}}(j\omega_r)\times\Big(\int_{\tau_1}..\int_{\tau_{2\beta-1}}g^{j_1\cdots j_\alpha}(\tau_1,...)e^{-j\omega_1\tau_1}\times\prod_{r=2}^{\beta-1}e^{j\omega_r(\tau_{2r-2}-\tau_{2r-1})}\,d\tau_1...d\tau_{2\beta-1}\Big)\,df_1...df_\beta \tag{A.5.9}$$

The time domain integral within the expression defines the multidimensional frequency transform of the Volterra kernel:

$$G^{j_1\cdots j_K}(f,-f_{2,}f_2,...)=\int_{\tau_1}..\int_{\tau_{2\beta-1}}g^{j_1j_2j\cdots j_K}(\tau_1,...)\,e^{-j\omega_1\tau_1}\prod_{r=2}^{\beta-1}e^{j\omega_r(\tau_{2r-2}-\tau_{2r-1})}d\tau_1...d\tau_{2\beta-1} \tag{A.5.10}$$

With it the correlation is:

$$R_{y\,j_1\cdots j_\alpha u_k}(\tau_0)=\int_{f_1=-\infty}^{\infty}\Big(\int_{f_2}\!\!\cdots\int_{f_\beta}G^{j_1\cdots j_K}(f,-f_{2,}f_2,...)S_{U_kU_k}(j\omega_1)\prod_{r=2}^{\beta-1}S_{U_{j_r}U_{j_r}}(j\omega_r)df_2...df_\beta\Big)\,e^{j\omega_1\tau_0}\,df_1 \tag{A.5.11}$$

The term in the parenthesis is the required non-linear spectral cross contribution:

$$S_{Y\,j_1\cdots j_\alpha U_k}(j\omega_1)=S_{U_kU_k}(j\omega_1)\int_{f_2}\!\!\cdots\int_{f_\beta=-\infty}^{\infty}G^{j_1\cdots j_K}(f,-f_{2,}f_2,...)\prod_{r=2}^{\beta-1}S_{U_{j_r}U_{j_r}}(j\omega_r)\,df_2...df_\beta \tag{A.5.12}$$

It remains now to scale up the cross spectrum and use (5.48) to obtain the final result as:

$$G_B^{j_1\cdots j_K}(j\omega)=\frac{C_{\ker nel}}{(f_{\max})^{\beta-1}}\int_{f_2=0}^{f_{\max}}\!\!\cdots\int G^{j_1\cdots j_K}(f,-f_{2,}f_2,...)\times S_{\hat U_{j_1}\hat U_{j_1}}(j\omega_2)...S_{\hat U_{j_K}\hat U_{j_K}}(j\omega_\beta)df_2...df_\beta \tag{A.5.13}$$

$$C_{\ker nel}=2^{\beta-1}M_1!!\prod_{l=1}^{K}(M_l-1)!! \tag{A.5.14}$$

where $2\beta$-1 = $\alpha$ = Σ$M_i$ +1. For missing details see [162]. SISO case follows up as a special case with: $j_1 = j_2 = \ldots j_\alpha = k$.

### *A.2 Random phase multisine and periodic noise, $H_1$-FRF measurements*

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

$$G_{BLA}^{k}(l)=E\{\hat G^k(l)\}=G^k(l)+G_B^k(l)=E\{Y(l)/U_k(l)\} \tag{A.5.15}$$

The expectation of a particular kernel from (4.1.4) is:

$$G_B^{j_1\dots j_K}(l) = E\{\frac{Y^{j_1\dots j_K}(l)}{U_k(l)}\} = \\ M^{-\alpha/2} \sum_{k_1,\dots,k_{\alpha-1}\in S_M^- \cup S_M^+} G^{j_1 j_2\dots j_\alpha}(k_1,\dots,k_\alpha) E\{U_{j_1}(k_1)\dots U_{j_\alpha}(k_\alpha)\frac{1}{U_k(l)}\} \quad \text{(A.5.16)}$$

For random phase multisines the expectation in (A.5.16) applies only to the random phases. To yield nonzero expected value the reference input $k$ must be present among the 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 as for Gaussian noise and with it and with the definition of the signal spectrum (5.4.1) the bias (A.5.16) is:

$$G_B^{j_1\dots j_K}(l) = \frac{C_{kernel}}{(M)^{\beta-1}} \sum_{k_1,\dots,k_{\beta-1}\in S_M^+} G^{j_1\dots j_K}(l,-k_1,k_1,\dots) F^{-(\beta-1)} S_{\hat{U}_{j_1}\hat{U}_{j_1}}(k_1)\dots S_{\hat{U}_{j_K}\hat{U}_{j_K}}(k_{\beta-1}) \quad \text{(A.5.17)}$$

With $M\to\infty$ the sum converges to the value of the integral (A.5.13) (equivalence in the limit (2.2.10)).

The periodic noise is more involved, because the expectation in (A.5.16) applies to the amplitudes and the phases:

$$E_{\hat{U}\varphi}\{U_{j_1}(k_1)\dots U_{j_\alpha}(k_\alpha)\frac{1}{U_k(l)}\} = E_{\hat{U}}\{\hat{U}_{j_1}(k_1)\dots \hat{U}_{j_\alpha}(k_\alpha)\frac{1}{\hat{U}_k(l)}\}\times E_\varphi\{e^{j\varphi_{j_1}(k_1)}\dots e^{-j\varphi_k(l)}\} \quad \text{(A.5.18)}$$

Here also it can be noted that the reference input must be present (and an odd number of times) in the kernel and other inputs must appear in even numbers, otherwise the expectations are zero. In this case the denominator and phase expectation cancel and:

$$E_{\hat{U}}\{\hat{U}_{j_1}(k_1)\dots\hat{U}_{j_\alpha}(k_\alpha)\} = E_{\hat{U}}\{\hat{U}^2_{j_1}(k_1)\}\ \dots E_{\hat{U}}\{\hat{U}^2_{j_K}(k_{\beta-1})\} + O(M^{-1}) \quad \text{(A.5.19)}$$

The asymptotically vanishing term contains higher even order moments. To create higher than 2nd order moments more than two (four, six, etc.) frequencies must be paired and run together. This cancels too many of degrees of freedom and together with the normalization of the signals yields vanishing order of magnitude for such contributions (equivalence in the limit). Taking into account all these assumptions the bias will again equal (A.5.16).

### *A.3 Random phase multisine and periodic noise, set-of-equation measurements*

When (4.2.5) is used as the measurement procedure:

$$G_{B,k}^{j_1\dots j_K}(l) = E\{\left[\mathbf{Y}(l)\mathbf{U}^H(l)(\mathbf{U}(l)\mathbf{U}^H(l))^{-1}\right]_k\} = E\{\sum_{n=1}^N b_{kn}(l)\sum_{i=1}^J Y^{j_1\dots j_K(i)}(l)\,\overline{U}_n^{(i)}(l)\} = \\ \sum_{k_1,\dots,k_{\alpha-1}} G^{j_1\dots j_K}\sum_{n=1}^N\sum_{i=1}^J E\{U_{j_1}^{(i)}(k_1)\dots U_{j_K}^{(i)}(k_\alpha)\,\overline{U}_n^{(i)}(l)\,b_{kn}(l)\} \quad \text{(A.5.20)}$$

where: $[b_{kn}(l)] = \left[\sum_{i=1}^J U_n^{(i)}(l)\,\overline{U}_k^{(i)}(l)\right]^{-1}$.

With phases independent over the frequencies and with higher than 2nd order moments (pairing more than two frequencies together) leading to $O(M^{-1})$ order contributions the internal term containing the expectation is:

$$\sum_{n=1}^N\sum_{i=1}^J E\{U_{j_1}^{(i)}(k_1)\dots U_{j_K}^{(i)}(k_\alpha)\,\overline{U}_n^{(i)}(l)\,b_{kn}(l)\} = |U(k_1)|^2\dots|U(k_{\beta-1})|^2\times E\{\sum_{n=1}^N\sum_{i=1}^J \overline{U}_n^{(i)}(l)\dots U_j^{(i)}(l)\,b_{kn}(l)\} \quad \text{(A.5.21)}$$

It is easy to recognize that this term in the expectation is $\delta_{kn}$ (Kronecker Delta) because it is an expansion of entries of a matrix multiplied by its inverse (from (4.2.5)). Substituting (A.5.21) into (A.5.20) we again obtain (A.5.17).

For periodic noise the expectation in (A.5.19) must be investigated accordingly to (A.5.18-A.5.19), however the entries of the inverse matrix pose now more problems:

$$\sum_{n=1}^{N}\sum_{i=1}^{J} E_{\hat{U},\varphi}\{U_{j_1}^{(i)}(k_1)...U_{j_K}^{(i)}(k_\alpha)\,\overline{U}_n^{(i)}(l)\,b_{kn}(l)\} = \sum_{n=1}^{N}\sum_{i=1}^{J} E_{\hat{U}}\{\hat{U}_{j_1}^{(i)}(k_1)...\hat{U}_{j_K}^{(i)}(k_\alpha)\,\hat{U}_n^{(i)}(l)\,E_{\varphi}\{e^{j\varphi_{j_1}^{(i)}(k_1)}...e^{-j\varphi_k^{(i)}(l)}\,b_{kn}(l)\}\} \tag{A.5.22}$$

Analyzing the possible pairings which are required for the nonzero expected value leads to the interim expression of:

$$E_{\hat{U}}\{\hat{U}_{j_1}^{2}(k_1)...\hat{U}_{j_K}^{2}(k_{\beta-1})\times E_{\varphi}\{\sum_{n=1}^{N}\sum_{i=1}^{J}\overline{U}_n^{(i)}(l)...U_j^{(i)}(l)\;b_{kn}(l)\}\} \tag{A.5.23}$$

which with the comments made to (A.5.19) and (A.5.21) (i.e. that the contribution of higher order moments disappears in the limit, that the term within expectation equals Kronecker Delta, and with the definition of the signal spectral content (5.4.2)) yields exactly the same expression as (A.5.17). ■

## A.6 Proof of Theorem 6.1

**Proof:** In computing the output of the cascade (6.1.1-6.1.2) we can observe that the multiplication of three or more cubic terms $G^{111}$ or $Y^3$ means contribution of order higher than $\varepsilon^2$ and such terms will be omitted from further consideration. The approximate output of the cascade is thus:

$$Z(l) \approx G^1(l)Y(l) + G^1(l)Y^3(l) + \sum_{z_1\in S_M^-\cup S_M^+}\;\sum_{z_2\in S_M^-\cup S_M^+} G^{111}(z_1,z_2,L_z)Y(z_1)Y(z_2)Y(L_z) + 3\sum_{z_1\in S_M^-\cup S_M^+}\;\sum_{z_2\in S_M^-\cup S_M^+} G^{111}(z_1,z_2,L_z)Y(z_1)Y(z_2)Y^3(L_z) = Z_I + Z_{II} + Z_{III} + Z_{IV} \tag{A.6.1}$$

To obtain the BLA FRF of the cascade $Q_{BLA}(l) = \dfrac{E\{Z(l)\overline{U}(l)\}}{|U(l)|^2}$, the components of $Z(l)$ being in phase with the $U(l)$ must be computed. The first term is the linear part of the cascade:

$$E\{Z_I(l)/U(l)\} = a_1R_1(l)S_1(l)\,b_1R_2(l)S_2(l) = G^1(l)H^1(l) \tag{A.6.2}$$

In the 2nd term the effect of the first non-linearity shows. After pairing the frequencies like in Th. 2.2.6, it becomes:

$$E\{Z_{II}(l)/U(l)\} = b_1R_2(l)S_2(l)\,E\{\sum_{k_1\in S_M^-\cup S_M^+}\;\sum_{k_2\in S_M^-\cup S_M^+} H^{111}(k_1,k_2,L_k)U(k_1)U(k_2)U(L_k)U(-l)\}/|U(l)|^2$$
$$= a_3R_1(l)S_1(l)\,b_1R_2(l)S_2(l)\;3!!\,2\,\varepsilon_1 r\sum_{k\in S_M^+}|R_1(k)|^2|U(k)|^2 = 3!!\,\varepsilon_1\,r\,r_1G^1(l)H^1(l),$$
$$\varepsilon_1 = a_3/a_1,\quad r = a_1\,b_1,\quad r_1 = \sum_{k\in S_M^+}|R_1(k)|^2|U(k)|^2 \tag{A.6.3}$$

The third term expresses the effect of the second non-linearity. After suitable pairing we have:

$$E\{Z_{III}(l)/U(l)\} = E\{\sum_{z_1\in S_M^-\cup S_M^+}\sum_{z_2\in S_M^-\cup S_M^+} G^{111}(z_1,z_2,L_z)H^1(z_1)U(z_1)H^1(z_2)U(z_2)H^1(L_z)U(L_z)U(-l)\}/|U(l)|^2$$

$$= H^1(l)b_1R_2(l)S_2(l)\ 3!!\,2\frac{b_3}{b_1}\sum_{k\in S_M^+}|R_2(k)|^2|H^1(k)|^2|U(k)|^2 = 3!!\,\varepsilon_2 r_{21}G^1(l)H^1(l)$$

(A.6.4)

where: $\varepsilon_2 = b_3/b_1$, $\quad r_{21} = \sum_{k\in S_M^+}|R_2(k)|^2|H^1(k)|^2|U(k)|^2$. (A.6.5)

The computation of the last mixed term is more involved.

$$E\{Z_{IV}(l)/U(l)\} = E\{3\sum_{z_1\in S_M^-\cup S_M^+}\sum_{z_2\in S_M^-\cup S_M^+} G^{111}(z_1,z_2,L_z)H^1(z_1)U(z_1)H^1(z_2)U(z_2)\times$$

$$\sum_{k_1\in S_M^-\cup S_M^+}\sum_{k_2\in S_M^-\cup S_M^+} H^{111}(k_1,k_2,L_k)U(k_1)U(k_2)U(L_k)U(-l)\}/|U(l)|^2$$

(A.6.6)

with $L_k = L_z - k_1 - k_2$ and $L_z = l - z_1 - z_2$. In evaluating (A.6.6) we must distinguish between the cases, when one of the frequencies $z$ is paired with $l$, and when none of the frequencies $z$ is paired with $l$ value. The first case is a simple superposition of the non-linear effects:

$$E\{Z_{IV}(l)/U(l)\}|_1 = 3\times(3!!\times2)^2\times b_3 S_2(l)R_2(l)H^1(l)\sum_{k\in S_M^+}|R_2(k)|^2H^1(k)U(k)$$

$$\times a_3 S_1(-k)R_1(-k)U(-k)\sum_{p\in S_M^+}|R_1(p)|^2|U(p)|^2 = 27\varepsilon_1\varepsilon_2 r_1 r_{21}G^1(l)H^1(l)$$

(A.6.7)

The second case brings in mixed multiple convolutions of both kernels:

$$E\{Z_{IV}(l)/U(l)\}|_2 = 3\times3!\times a_3b_3 S_2(l)R_1(l)\times$$

$$\sum_{z_1\in S_M^-\cup S_M^+}\sum_{z_2\in S_M^-\cup S_M^+} R_2(z_1)R_2(z_2)R_2(l-z_1-z_2)H^1(z_1)H^1(z_2)S_1(l-z_1-z_2)R_1(-z_1)R_1(-z_2)|U(z_1)|^2|U(z_2)|^2 =$$

(A.6.8)

$$\frac{9}{2}a_3b_3 S_2(l)R_1(l)\frac{1}{M^2}\sum_{z_1\in S_M^-\cup S_M^+}\sum_{z_2\in S_M^-\cup S_M^+} R_2(z_1)R_2(z_2)R_2(l-z_1-z_2)H^1(z_1)H^1(z_2)S_1(l-z_1-z_2)R_1(-z_1)R_1(-z_2)$$

$$= \frac{9}{2}\varepsilon_1\varepsilon_2 r\,N(l)C^2(l) = \frac{9}{2}\varepsilon_1\varepsilon_2 r\,G^1(l)H^1(l)\frac{C^2(l)}{L(l)}$$

(A.6.9)

where: $L(l) = S_1(l)R_2(l)$, $N(l) = R_1(l)S_2(l)$, and $N(l)L(l) = G^1(l)H^1(l)$. (A.6.10)

$C^2(l)$ is the sum approximation of the double convolution of the form

$$\int|R_1(f_1)|^2 L(f_1)\int|R_1(f_2)|^2 L(f_2)L(f - f_1 - f_2)\,df_1\,df_2\,.$$

Collecting (A.6.2-A.6.9) together yields:

$$G_{BLA}(l) = E\{Z(l)/U(l)\} = K(1+\frac{C_1}{L(l)})G^1(l)H^1(l)$$ (A.6.11)

where: $K = 1 + 3\varepsilon_1 r_1 + 3\varepsilon_2 r_{21} + 27\varepsilon_1 \varepsilon_2 r_1 r_{21}$ (A.6.12)

$$C_1 = \frac{K^*}{K},\ K^* = \frac{9}{2}\varepsilon_1 \varepsilon_2 r\, C^2(l), \tag{A.6.13}$$

Outside the frequency bands of high coherence and also in case of higher levels of non-linearities additional effects come into the play yielding:

$$\frac{G_{BLA}(l)}{G^1(l)H^1(l)} = K\,(1 + \frac{C_1}{L(l)}) \neq const\,, \tag{A.6.14}$$

and the overall related FRF shows an increasing relative constant (due to $K$) and frequency dependent (due to $L(l)$) bias. However, due to $K \approx 1 + O(\varepsilon)$, $C_1 \approx O(\varepsilon^2)$, this effect is not always visible, if the system dynamics and the level of non-linearity are small (it was also assumed that the convolution (A.6.9) is much more smoother than the system dynamics. ■

## A.7 Proof of Theorem 6.2.1

**Proof:** In the derivation $\varepsilon, \delta$ will be used instead of $\varepsilon_{\mathrm{km}}$, $\delta_{\mathrm{km}}$, when dependence on the levels of the linear and the non-linear distortions is considered in general, and $\xi = \max(\varepsilon, \delta)$, when the dependence on the level of any kind of distortion is considered (for the more detailed proof, see [19*]). Distorted inputs can be written as a sum of ideal, linearly distorted, and non-linearly distorted signals:

$$\boldsymbol{U} = \mathbf{M}[\boldsymbol{R}] = \boldsymbol{U}_0 + \boldsymbol{U}_d = \boldsymbol{U}_0 + \boldsymbol{U}_e + \boldsymbol{U}_n = \mathbf{L}_0\boldsymbol{R} + \mathbf{E}\boldsymbol{R} + \mathbf{N}[\boldsymbol{R}] \tag{A.7.1}$$

and the measured outputs similarly as:

$$\boldsymbol{Y} = \mathbf{V}[\boldsymbol{U}] = \mathbf{V}[\boldsymbol{U}_0 + \boldsymbol{U}_e + \boldsymbol{U}_n] = \boldsymbol{Y}_0 + \boldsymbol{Y}_d = \boldsymbol{Y}_0 + \boldsymbol{Y}_e + \boldsymbol{Y}_n + O(\xi^2) \tag{A.7.2}$$

where $\boldsymbol{Y}_0 = \mathbf{V}[\boldsymbol{U}_0]$, and $\boldsymbol{Y}_e$, $\boldsymbol{Y}_n$ are the first order approximations in distortions:

$$\boldsymbol{Y}_e = \boldsymbol{Y}_e^{lin} + \boldsymbol{Y}_e^{nonlin} = \mathbf{G}_{LIN}\boldsymbol{U}_e + \mathbf{V}_1[\boldsymbol{U}_0, \boldsymbol{U}_e], \qquad \boldsymbol{Y}_n = \boldsymbol{Y}_n^{lin} + \boldsymbol{Y}_n^{nonlin} = \mathbf{G}_{LIN}\boldsymbol{U}_n + \mathbf{V}_1[\boldsymbol{U}_0, \boldsymbol{U}_n]$$

Kernels $\mathbf{V}_1[\boldsymbol{U}_0, \boldsymbol{U}]$ are all those kernels in $\mathbf{V}[\boldsymbol{U}]$ with exactly one input being $\boldsymbol{U}$, all other inputs being $\boldsymbol{U}_0$. This 1st order decomposition is justified by: $U_0 \approx O(1)$, $U_e \approx O(\varepsilon)$, $U_n \approx O(\delta)$, $Y_e \approx O(\varepsilon)$, $Y_d \approx O(\delta)$.

Denote the ideal measurements as:

$$\boldsymbol{Y}_r = \mathbf{V}[\boldsymbol{R}],\ \ \boldsymbol{Y}_0 = \mathbf{V}[\boldsymbol{U}_0], \tag{A.7.3}$$

then the FRF estimated from the measured data are:

$$\hat{\mathbf{G}}_r = \mathbf{Y}_r\mathbf{R}^H(\mathbf{R}\mathbf{R}^H)^{-1} \text{ and } \hat{\mathbf{G}}_0 = \mathbf{Y}_0\mathbf{U}_0^H(\mathbf{U}_0\mathbf{U}_0^H)^{-1} \tag{A.7.4}$$

respectively, where $\mathbf{U}_0$ is a block of the experimental data:

$$\mathbf{U}_0 = \begin{bmatrix} U_{0,1}^{(1)}(l) & \dots & U_{0,1}^{(J)}(l) \\ \dots & U_{0,i}^{(j)}(l) & \dots \\ U_{0,N}^{(1)}(l) & \dots & U_{0,N}^{(J)}(l) \end{bmatrix} \tag{A.7.5}$$

(index in the parentheses is the serial number of the experiment).

With the decompositions (A.7.1-A.7.2) we can write that:

$$\hat{\mathbf{G}} = \mathbf{Y}\mathbf{U}^H(\mathbf{U}\mathbf{U}^H)^{-1} = (\mathbf{Y}_0 + \mathbf{Y}_d)(\mathbf{U}_0^H + \mathbf{U}_d^H)[(\mathbf{U}_0 + \mathbf{U}_d)(\mathbf{U}_0^H + \mathbf{U}_d^H)]^{-1}$$
$$= (\mathbf{Y}_0 + \mathbf{Y}_d)(\mathbf{U}_0^H + \mathbf{U}_d^H)(\mathbf{U}_0\mathbf{U}_0^H)^{-1}[(\mathbf{I} + \mathbf{U}_d\mathbf{U}_0^H(\mathbf{U}_0\mathbf{U}_0^H)^{-1} + \mathbf{U}_0\mathbf{U}_d^H(\mathbf{U}_0\mathbf{U}_0^H)^{-1} + \mathbf{U}_d\mathbf{U}_d(\mathbf{U}_0\mathbf{U}_0^H)^{-1}]^{-1} \quad \text{(A.7.6)}$$

Let the inverse in the squared parenthesis be: $(\mathbf{I}+\mathbf{P})^{-1}$. Then if $\|\mathbf{P}\| < 1$, $(\mathbf{I}+\mathbf{P})^{-1} \approx (\mathbf{I}-\mathbf{P})$, if $\xi$ is small enough, because every component in $\mathbf{P}$ contains at least one signal of order $O(\xi)$. Introducing $\mathbf{U}_d = \mathbf{U}_e + \mathbf{U}_n$ leads to (from here on the equality means that the 1st order approximation):

$$\hat{\mathbf{G}} = \mathbf{G}_I + \mathbf{G}_{II} + ... + \mathbf{G}_{IX} \quad \text{(A.7.7)}$$

$$\mathbf{G}_I = \mathbf{Y}_0\mathbf{U}_0^H(\mathbf{U}_0\mathbf{U}_0^H)^{-1} \quad \text{(A.7.8)}$$

$$\mathbf{G}_{II} = -\mathbf{Y}_0\mathbf{U}_0^H(\mathbf{U}_0\mathbf{U}_0^H)^{-1}\mathbf{U}_e\mathbf{U}_0^H(\mathbf{U}_0\mathbf{U}_0^H)^{-1} \quad \text{(A.7.9)}$$

$$\mathbf{G}_{III} = -\mathbf{Y}_0\mathbf{U}_0^H(\mathbf{U}_0\mathbf{U}_0^H)^{-1}\mathbf{U}_n\mathbf{U}_0^H(\mathbf{U}_0\mathbf{U}_0^H)^{-1} \quad \text{(A.7.10)}$$

$$\mathbf{G}_{IV} = -\mathbf{Y}_0\mathbf{U}_0^H(\mathbf{U}_0\mathbf{U}_0^H)^{-1}\mathbf{U}_0\mathbf{U}_e^H(\mathbf{U}_0\mathbf{U}_0^H)^{-1} \quad \text{(A.7.11)}$$

$$\mathbf{G}_V = -\mathbf{Y}_0\mathbf{U}_0^H(\mathbf{U}_0\mathbf{U}_0^H)^{-1}\mathbf{U}_0\mathbf{U}_n^H(\mathbf{U}_0\mathbf{U}_0^H)^{-1} \quad \text{(A.7.12)}$$

$$\mathbf{G}_{VI} = \mathbf{Y}_0\mathbf{U}_e^H(\mathbf{U}_0\mathbf{U}_0^H)^{-1} \quad \text{(A.7.13)}$$

$$\mathbf{G}_{VII} = \mathbf{Y}_0\mathbf{U}_n^H(\mathbf{U}_0\mathbf{U}_0^H)^{-1} \quad \text{(A.7.14)}$$

$$\mathbf{G}_{VIII} = \mathbf{Y}_e\mathbf{U}_0^H(\mathbf{U}_0\mathbf{U}_0^H)^{-1} \quad \text{(A.7.15)}$$

$$\mathbf{G}_{IX} = \mathbf{Y}_n\mathbf{U}_0^H(\mathbf{U}_0\mathbf{U}_0^H)^{-1} \quad \text{(A.7.16)}$$

To see how good the measurements are we should evaluate the difference $E\{\hat{\mathbf{G}}\} - \mathbf{G}_{BLA}$.

*Case of* $\mathbf{G}_I$: By definition: $E\{\mathbf{G}_I\} = E\{\mathbf{G}_0\} = \mathbf{G}_{BLA} = \mathbf{G}_{LIN} + \mathbf{G}_{BIAS}$ (A.7.17)

*Case of* $\mathbf{G}_{II}$: With (A.7.1) and (6.2.1-6.2.4) it is easy to see:

$$\mathbf{U}_e = \mathbf{E}\mathbf{R}, \quad \mathbf{R} = \mathbf{L}_0^{-1}\mathbf{U}_0, \quad \mathbf{U}_e\mathbf{U}_0^H(\mathbf{U}_0\mathbf{U}_0^H)^{-1} = \mathbf{E}\mathbf{L}_0^{-1} \quad \text{(A.7.18)}$$

Consequently: $E\{\mathbf{G}_{II}\} = -\mathbf{G}_{BLA}\mathbf{E}\mathbf{L}_0^{-1} = -\mathbf{G}_{LIN}\mathbf{E}\mathbf{L}_0^{-1} - \mathbf{G}_{BIAS}\mathbf{E}\mathbf{L}_0^{-1}$ (A.7.19)

*Case of* $\mathbf{G}_{IV}$: It is also easy to see (directly from the square **U** in (A.7.5) and through the singular value decomposition for the rectangular **U**) that:

$$\mathbf{U}_0^H(\mathbf{U}_0\mathbf{U}_0^H)^{-1}\mathbf{U}_0 = \mathbf{I} \quad \text{(A.7.20)}$$

With this simplification it turns out, that:

$$\mathbf{G}_{IV} = -\mathbf{Y}_0\mathbf{U}_0^H(\mathbf{U}_0\mathbf{U}_0^H)^{-1}\mathbf{U}_0\mathbf{U}_e^H(\mathbf{U}_0\mathbf{U}_0^H)^{-1} = -\mathbf{Y}_0\mathbf{U}_e^H(\mathbf{U}_0\mathbf{U}_0^H)^{-1} = -\mathbf{G}_{VI} \quad \text{(A.7.21)}$$

*Case of* $\mathbf{G}_V$: Similarly:

$$\mathbf{G}_V = -\mathbf{Y}_0\mathbf{U}_0^H(\mathbf{U}_0\mathbf{U}_0^H)^{-1}\mathbf{U}_0\mathbf{U}_n^H(\mathbf{U}_0\mathbf{U}_0^H)^{-1} = -\mathbf{Y}_0\mathbf{U}_n^H(\mathbf{U}_0\mathbf{U}_0^H)^{-1} = -\mathbf{G}_{VII} \quad \text{(A.7.22)}$$

*Case of* $\mathbf{G}_{III}$: Here we can write:

$$\mathbf{G}_{III} = -\mathbf{Y}_0\mathbf{U}_0^H(\mathbf{U}_0\mathbf{U}_0^H)^{-1}\mathbf{U}_n\mathbf{U}_0^H(\mathbf{U}_0\mathbf{U}_0^H)^{-1} = -\hat{\mathbf{G}}_0\hat{\mathbf{N}}_{BLA}\mathbf{L}_0^{-1} \quad \text{(A.7.23)}$$

where:

$$\mathbf{U}_n\mathbf{U}_0^H(\mathbf{U}_0\mathbf{U}_0^H)^{-1} = \mathbf{U}_n\mathbf{R}^H(\mathbf{R}\mathbf{R}^H)^{-1}\mathbf{L}_0^{-1} = \hat{\mathbf{N}}_{BLA}\mathbf{L}_0^{-1} \tag{A.7.24}$$

is the best linear approximation of the purely non-linear part of the distortions. The expression is thus a product of two FRF estimates coming from two independent non-linear systems, driven by common input. Consider further that every FRF estimate can be decomposed into the Best Linear Approximation and the non-linear noise:

$$\hat{\mathbf{G}}_0 = \mathbf{G}_{BLA} + \mathbf{G}_S, \quad \hat{\mathbf{N}}_{BLA} = \mathbf{N}_{BLA} + \mathbf{N}_S \tag{A.7.25}$$

$$\begin{aligned}\mathbf{G}_{III} &= -\hat{\mathbf{G}}_0\hat{\mathbf{N}}_{BLA}\mathbf{L}_0^{-1} = (\mathbf{G}_{BLA} + \mathbf{G}_S)(\mathbf{N}_{BLA} + \mathbf{N}_S)\mathbf{L}_0^{-1} \\ &= -\mathbf{G}_{BLA}\mathbf{N}_{BLA}\mathbf{L}_0^{-1} - \mathbf{G}_{BLA}\mathbf{N}_S\mathbf{L}_0^{-1} - \mathbf{G}_S\mathbf{N}_{BLA}\mathbf{L}_0^{-1} - \mathbf{G}_S\mathbf{N}_S\mathbf{L}_0^{-1}\end{aligned} \tag{A.7.26}$$

The best linear approximations are deterministic, the non-linear noises are circularly normally distributed, noise in (A.7.25) is of order $O(\xi)$, the noise amplitudes are of order $O(N^{-1/2})$, altogether we have:

$$E\{\mathbf{G}_{III}\} = -\mathbf{G}_R\mathbf{N}_{BLA}\mathbf{L}_0^{-1} - E\{\mathbf{G}_S\mathbf{N}_S\}\mathbf{L}_0^{-1} = -\mathbf{G}_{BLA}\mathbf{N}_{BLA}\mathbf{L}_0^{-1} + O(\xi N^{-1}) \tag{A.7.27}$$

*Case of* $\mathbf{G}_{VIII}$:

$$\mathbf{G}_{VIII} = \mathbf{Y}_e\mathbf{U}_0^H(\mathbf{U}_0\mathbf{U}_0^H)^{-1} \tag{A.7.28}$$

From the point of view of the Best Linear Approximation, kernels of every Volterra system can be classified as "bias" (nonzero mean) and "variance" (zero mean) kernels. Bias kernels will appear in the best linear approximation and will also contribute to the non-linear noise, variance kernels contribute solely to the non-linear noise [163]. In case of a MIMO Volterra system bias kernels for measurement channel $Y$-$U_{\mathrm{k}}$ are those kernels $G_k^{j_1 j_2 \ldots j_\alpha}$, which within input indices $j_1, j_2, \ldots, j_\alpha$ contain input reference index $k$ an odd number of times, and every other input index an even number of times ([15*-16*]), i.e.:

$$j_1, j_2, \ldots j_\alpha = <\underbrace{k \ldots k}_{M_{reference}}><\underbrace{j_1, \ldots j_1}_{M_1}>\ldots<\underbrace{j_K, \ldots j_K}_{M_K}>, \quad M_{reference} = odd, \quad M_k = even, \quad M_{reference} + \sum_{k=1}^{K} M_k = odd \tag{A.7.29}$$

i.e. for channel $Y$-$U_1$ $G_1^{122}$ is a bias kernel and $G_1^{112}$ is a variance kernel, however for channel $Y$-$U_2$ the situation is just opposite.

The problem with evaluating (A.7.28) is that the bias and the variance kernels in $\boldsymbol{Y}_e$ may exchange roles comparing to the ideal case. The reason is that signals $U_{e,\mathrm{k}}$, contrary to signals $U_{0,\mathrm{k}}$, are linear combinations of all the reference inputs $R_{\mathrm{m}}$ and $R_{\mathrm{m}}$ will appear in place of $U_{e,\mathrm{k}}$ changing the actual input indices of the kernels. Remember also, that $\boldsymbol{Y}_e$ contains only a single input $\boldsymbol{U}_e$.

When investigating a particular kernel we must consider thus every variation when one of the input indices changes to some other value. In consequence bias kernels can turn into variance kernels and vice versa. The expected value of (A.7.28) can be written as:

$$E\{\hat{G}_{VIII}^1\} = tr\{\mathbf{B}_1\mathbf{E}\} \tag{A.7.30}$$

with:

$$\mathbf{B}_1 = \begin{bmatrix} G_{BLA}^1 & G_{BLA}^2 & G_{BLA}^3 \\ G_B^{*112} & G_B^{*122} & G_B^{*123} \\ G_B^{*113} & G_B^{*123} & G_B^{*133} \end{bmatrix} \tag{A.7.31}$$

where <*kmn*> is the permutation of the indices *k*, *l*, and *m* in an increasing order and the bias terms $G_{\mathrm{B}}$ are those appearing in the BLA $G_{BLA}$ [15*-16*]. Similar construction can be done for the measurements in other channels, i.e.:

$$E\{\hat{G}_{VIII}^k\} = tr\{\mathbf{B}_k\mathbf{E}\} \tag{A.7.32}$$

$$[\mathbf{B}_k]_{mn} = \begin{cases} G^{n}_{BLA} & m = k \\ G_B^{*<nmk>} & m \neq k \end{cases} \tag{A.7.33}$$

*Case of* $\mathbf{G}_{IX}$ : $\mathbf{G}_{IX} = \mathbf{Y}_n \mathbf{U}_0^H (\mathbf{U}_0 \mathbf{U}_0^H)^{-1}$ (A.7.34)

The analysis of this term is even more involved and here we will simplify the analysis using the assumptions of the low order systems. However the line of reasoning and the results can be extended to higher order systems.

The distortion component $Y_n$ depends on a single appearance of the distortion signals $U_n$:

$$\boldsymbol{Y}_n = \mathbf{V}_1[\boldsymbol{U}_0, \boldsymbol{U}_n] \tag{A.7.35}$$

$\mathbf{U}_n$ is the output of a 2nd and 3rd order Volterra system, with kernels of order $O(\delta)$. We will investigate when the expected value $E\{Y_n \overline{U}_{0,k}\} \neq 0$ (A.7.36)

Contrary to the substitution of the signal $U_e$ into the kernels of $Y_n$, which did not change the order of the kernels, the substitution of $U_n$ into the kernels of $Y_n$ means the substitution of a kernel into a kernel, consequently the order of the resulting kernels will change.

Substituting $U_n$ into $Y_n$ takes place of one input signal, but introduces as many new signals as is the order of the kernel present in $U_n$. Consequently to have a nonzero mean bias odd order kernel:

*order*($Y_n$ *kernel*) + *order*($U_n$ *kernel*) = *even*.

For the assumed low order systems we need to consider 1+3, 2+2, and 3+3 kernel order combinations. In this case the even order kernels normally present only in the non-linear noise, will contribute also to the overall bias.

Consider now a **V** system kernel with one of the inputs $U_{0,j_i}(k_i)$ substituted with a single **N** system kernel:

$$\sum_{k_1,\ldots,k_{\alpha-1} \in S_M^- \cup S_M^+} \sum_{\xi_1,\ldots,\xi_{\beta-1} \in S_M^- \cup S_M^+} \begin{matrix} G_k^{\,j_1 j_2 \ldots j_\alpha}(k_1, k_2, \ldots, k_{\alpha-1}, l - \Sigma_{i=1\ldots\alpha-1} k_i)\; N_m^{\,n_1 n_2 \ldots n_\beta}(\xi_1, \ldots, \xi_{\beta-1}, k_i - \Sigma_{i=1\ldots\beta-1}\xi_i) \\ \times U_{0,j_1}(k_1) \ldots U_{0,j_\alpha}(l - \Sigma_{i=1\ldots\alpha-1} k_i)\, U_{0,n_1}(\xi_1) \ldots U_{0,n_\beta}(k_i - \Sigma_{i=1\ldots\beta-1}\xi_i) \end{matrix} \tag{A.7.37}$$

and consider for simplicity only the product of the inputs (the sole random component of the whole expression):

$$U_{0,j_1}(k_1) \ldots U_{0,j_\alpha}(l - \Sigma_{i=1\ldots\alpha-1} k_i) \times U_{0,n_1}(\xi_1) \ldots U_{0,n_\beta}(k_i - \Sigma_{i=1\ldots\beta-1}\xi) \tag{A.7.38}$$

Let rename the frequency variables as follows:

$$k_m^{\bullet} = k_m, \quad m \neq i, \quad k_i^{\bullet} = k_i - \Sigma_{i=1\ldots\beta-1}\xi_i \tag{A.7.39}$$

we can see that the resulting expression is a well formed Volterra kernel:

$$U_{0,j_1}(k_1^{\bullet}) \ldots U_{0,j_{i-1}}(k_{i-1}^{\bullet}) U_{0,n_\beta}(k_i^{\bullet}) U_{0,j_{i+1}}(k_{i+1}^{\bullet}) \ldots U_{0,j_{\alpha-1}}(k_{\alpha-1}^{\bullet}) \times U_{0,n_1}(\xi_1) \ldots U_{0,n_{\beta-1}}(\xi_{\beta-1}) \times U_{0,j_\alpha}(l - \Sigma_{i=1\ldots\alpha-1} k_i^{\bullet} - \Sigma_{i=1\ldots\beta-1}\xi_i) \tag{A.7.40}$$

which we will denote as:

$$G_k^{\,j_1 j_2 \ldots j_\alpha}(N_{km}^{\,n_1 n_2 \ldots n_\beta}) \tag{A.7.41}$$

e.g. $G_1^{\,12}(N_2^{\,11})$ is actually like an $G_k^{\,111}$ order kernel.

To see, what is the true order of a mixed kernel (A.7.41), as a rule of thumb delete input index $m$ from $j_1, j_2, \ldots, j_\alpha$ and mix the resulting indices with $n_1, n_2, \ldots, n_\beta$ ordering them in the increasing order. Such a new kernel to have nonzero mean should show properties of a normal bias kernel mentioned before.

For the assumed low order systems the overall estimate can be written as:

$$\hat{\mathbf{G}} = \mathbf{G}_{BLA} - \mathbf{G}_{BIAS}(\mathbf{E}\mathbf{L}_0^{-1} - \mathbf{N}_{BLA}\mathbf{L}_0^{-1}) + \mathbf{H}_1 + \mathbf{H}_2 \tag{A.7.42}$$

Using $\mathbf{L}_0\mathbf{L}_0^{-1} = \mathbf{I}$ this can be written finally as:

$$\hat{\mathbf{G}} = \mathbf{G}_{BLA} + \mathbf{G}_{BIAS}(\mathbf{I} - \mathbf{M}_{BLA}\mathbf{L}_0^{-1}) + \mathbf{H}_1 + \mathbf{H}_2 + O(\xi^2) \qquad \text{(A.7.42)}$$

where the terms $\mathbf{H}_1$ and $\mathbf{H}_2$ are difficult to handle in a closed form and are best to be enumerated for particular applications, accordingly to the derived rules. ■