MSc Acoustics

School of Computing, Science and Engineering

MSc Dissertation

Capabilities of Polynomial Hammerstein modeling of weak nonlinear audio systems

Author: Stefano Tronci

Supervisor: Dr. Philip J. Duncan

2015

Abstract

Although successful, linear systems theory is often just an approximation of natural systems behavior. When insufficient, proper nonlinear modeling must be used instead. This project explores the use of the Synchronized Exponential Swept Sine method to blindly identify Polynomial Hammerstein Models of weakly nonlinear black-box audio systems. Hi-Fi amplifiers are selected as representative systems since:

- They are electronic systems, meaning that they are simple to study.
- Although there aren't simple ways to relate distortion metrics to perception, audiophiles want to relate distortion benchmarks to unit's perceived quality. Thus, audio units distortion perception should be more deeply investigated and models could be a valuable tool for that, their parameters being controllable.
- Experimental results on nonlinear distortion metrics suggest that small distortion products deriving from weak non-linearity can be clearly audible (Geddes & Lee, 2003b).
- It is interesting to test how well nonlinear systems identification methods perform when the non-linearity is weak and, as such, hard to discern and model.

It has been shown that proper designed stimulation of the systems can yield to their Input-Output law modeling. Models have been validated by comparison of synthesized signals with recorded signals, showing good agreement. A perceptual test was performed and insufficient evidence was found to support claims of audible differences between models and synthesized outputs.

However, models effectiveness is limited by their pass-bands, computational costs and input signal amplitude dependence of models parameters. Moreover, further work is required to design listening tests in order to raise subjects sensibility and allow to prove absence of perceptual differences.

This dissertation was typeset with LyX. Word count: 15000 words (12985 without captions, footnotes and Appendix).

Acknowledgments

I would like to thank my supervisor Philip J. Duncan for his constant guidance and support throughout the whole project and beyond.

I feel also grateful to all my classmates that helped me to overcome all the obstacles along the MSc course and all the Salford University Professors and staff for their quality teaching and support.

I deeply appreciate and thank also all the people that participated to the listening test for their time, effort and patience.

I am also in debt with my bachelors degree classmates for their advice throughout all these years, Prof. Oreste Nicrosini for allowing me to work on audio related topics early in my studies, Prof. Gianni Gilardi for initiating me to mathematics and exact sciences, Alberto Novello without which I would not even realized Salford University existed, Riccardo Ferranti and Renato Podestà who are the main sources of my passion for audio and music, Sender Set and all my friends without which I would not have developed such a deep interest in music and science, the Linux Musicians community, Ubuntu Studio community, Linuxaudio consortium, ArchBang Linux community & developers and all the Open Source developers and activist for providing their impressive resources and software.

Last but not least I would like to express a heartfelt gratitude to my family for their unconditioned and constant encouragement and support.

Contents

List of Figures

List of Tables

Abbrevations

LTI Linear Time Invariant system

$MISO$ Multiple Input Single Output

MLS Maximal Length Sequence

MSE Mean Squared Error

NTI Nonlinear Time Invariant system

PHM Polynomial Hammerstein Model

SES Synchronized Exponential Swept sine

$SISO$ Single Input Single Output system

THD Total Harmonic Distortion

GUI Graphical User Interface

1 Introduction

The project goal is to investigate how successful the blind identification of weakly nonlinear systems can be in view of the possible use of the models in distortion perception study. If Linear Time Invariant (LTI) systems can be described completely by their impulse response, Nonlinear Time Invariant (NTI) systems require sets of functions. However, the inverse filtering technique typical of LTI systems can be used to measure NTI systems as well, provided that the stimulation signal (the signal designed for the purpose of NTI identification) is properly designed. The models are built on the outcome of the measurement. For this reason most of section §3 is dedicated to the definition of the stimulation signal as its properties are essential for the success of the particular identification method. The actual measurement implementation description follows in the same section. The principles of the identification procedure and implementation are detailed in section §4, where it will be seen that the procedure has simple algebraic bases. The listening test design and outcome are reported in section §5. Finally, the answer to the research question is given in section §6 together with possible further research topics.

2 Background and Literature Review

Nonlinear distortion affects the propagation of every field in a medium. As such, it has been a major topic in science for decades, finding applications in transmission, communication and radar theories just to name few subjects. The information is then huge and fragmented: there isn't a source reviewing all the possible kind of distortions and their causes. In fact, focusing on audio electronics only, the causes of distortion are so many, depending on the technology and design involved, that books could be written on the topic. Considering the scope of the project, that involves blind identification of nonlinear black-box audio systems, it is also understood that an accurate study of the causes of distortion down to the component level would be pointless. Hence, the minimum number of sources, containing the most up to date and concise information, have been selected.

2.1 Nonlinear distortion and Nonlinear systems

Nonlinear distortion and systems have to be defined. Distortion applies to signals and can affect the output of both LTI and NTI systems. It is usually defined as an alteration of the input signal so that the output is not an exact (scaled) copy of the input. The distortion is nonlinear when this alteration depends on the input signal amplitude (Preis, 1984). A system instead is nonlinear when the physics governing it do not allow for at least one of superposition principle or homogeneity principle to hold[1] (Geddes & Lee, 2003b; Novak, 2009). As a result, scaling of NTI input does not yield to scaling of the output by the same factor, implying nonlinear distortion. To make quantitative these definitions it is natural to refer to simple nonlinear models and link input and output together with a mapping function, as suggested by Geddes and Lee (2003b) and Dobrucki (2011). The function maps the input value at a certain instant to an output value at a certain instant. The identification technique architecture reviewed in Section §2.2 uses static non-linear maps for memory-less distortion blocks only. Hence, this restriction is operated: the mapping function relates input and output at the same exact instant without contributions from the input at previous instants. The shape of the mapping function, linear or not, will clarify the kind of system.

When nonlinear, the mapping function can be represented in multiple ways, like a polynomial expansion or an expansion on a set of orthogonal functions. Dobrucki (2011), in his survey of electroacoustic systems distortions, clarifies that the polynomial expansion is profitable since it is simply related to the introduction of harmonics for a pure sinusoidal input. Harmonics appear to be a simple way to quantify the information added by the nonlinear system to the original signal and this yields to the introduction of the THD (Total Harmonic Distortion) metric, a quantity measuring the weight of the output harmonics amplitudes with respect the output fundamental amplitude. However, this

[1] See Appendix B.2 for a discussion.

metric is affected by poor perceptual correlation with subjective ratings of quality degradation (Geddes & Lee, 2003a). When the input is a sum of tones the nonlinear expansion points out that the single components interact together to produce sum and difference tones due to formation of cross products (see Section §4.2.2). As such, Harmonic and Intermodulation distortion are general consequences of every nonlinear transfer function no matter the physical reason behind it.

Related to typical circuit design choices is instead the Transient Intermodulation distortion, a consequence of the negative feedback design of most units. The negative feedback design has the goal to define the gain of the system and improve the signal to noise ratio by feeding a phase reversed fraction of the output to the system input. Any phenomenon bringing to rapid variation of an input signal, like transients, will produce rapid variation of the output. The phase reversed copy of the output will be fed to the input after a delay determined by the response of the system, a delay that the transient nature of the signal makes possible to appreciate. Hence, chances are that the feedback will sum in phase with subsequent parts of the signal, failing to control amplitude and instead raising it, causing instantaneous overdrive of the amplifying stages of the amplifier (Dobrucki, 2011).

2.2 Nonlinear system measure and identification technique

LTI systems are completely described by a single impulse response. This does not hold for *NTI* systems, which need a set of higher order responses. In fact, Geddes and Lee (2003b) warn against the use of the concepts of impulse and frequency responses on *NTI* systems: there isn't one quantity able to describe the whole system alone. This is evident referring to the most general ways to describe nonlinear systems as reported by Dobrucki (2011), namely Volterra and Wiener series. These models are based on kernel functions to be used with series of raising order convolutions of increasing complexity, the first order representing the linear process. Being convolution involved it is then possible to account for the memory of the system. Moving from this formalism Farina (2000) and Farina, Bellini, and Armelloni (2001) developed a way to identify the Volterra kernels. Farina (2000) acknowledged that Exponential Swept Sines can be used on *NTI* systems in an analogous way to what is done when measuring *LTI* systems. Convolution of output with an appropriate inverse filter will produce a function of time that is obtained from the superposition of shifted higher orders impulse responses, i.e. the impulse responses relating a tone input signal to the respective order output harmonic (Novak, Simon, & Lotton, 2010b). The linear impulse response appears centered in the origin of time, while the higher order responses appear centered at negative times, the higher the order the more negative the lag. This makes possible to separate the various orders, study and use them for Volterra kernels computation as showed by Farina et al. (2001). Moving from this approach Novak et al. (2010b) and Novak, Simon, and Lotton (2010a) developed

a Synchronized Exponential Swept Sine (SES) stimulation which is asymptotic in the sense defined by Cohen (2000) and Chassande-Mottin and Flandrin (1998), i.e. it is a signal whose Fourier transform can be approximated with the stationary phase method all over its spectrum. Synchronization is achieved by imposing the stimulation and its derivative to be respectively zero and positive at the time instants at which harmonics of the starting frequency are reached. These are the lags at which the different orders appear. Asymptotic behavior of chirps makes it possible to develop useful identities and then to elaborate the proper inverse filter. The synchronized fashion of the stimulation will generate a well defined succession of the center lags of the different orders responses as well as preventing artifacts in their phases.

Although computation of the Volterra (or Wiener) kernels is possible this approach is complicated (as evident in Farina et al. (2001)) and computationally demanding. Also, these kernels are not intuitive and simple to manipulate. For these reasons the $MISO$ based Polynomial Hammerstein nonlinear Model (PHM) proposed by Novak et al. (2010b) and Novak et al. (2010a) is chosen. The approach of the authors is to fit a model made of multiple branches, each made of a static nonlinear power function followed by a linear filter with memory whose response is derived from the system's higher order responses.

2.3 Perception of nonlinear distortion and Listening Tests

As said, as a consequence of the complexity of both NTI and human hearing systems, simple signal based metrics (like THD) fail to provide insights on the distortion perception. This has been acknowledged by Geddes and Lee (2003b) by pointing out poor correlation between these metrics and perception. Attempts to correct these metrics by taking into account psychoacoustics facts, like harmonics weighting in THD to match the higher sensibility of the ear to high frequencies distortion products, still fail (Dobrucki, 2011). This is a consequence of the fact that metrics elaborated with simple sinusoidal signals cannot be accurate for complex signals. In fact intermodulation distortion, happening with complex signals, is more detrimental to the quality, as showed by Geddes and Lee (2003a). Not only complex signals stimulate differently the NTI systems, but also the perception is made different by mutual masking of the sound components. Geddes and Lee (2003b) detail how to build a metric able to account for masking and its upwards spreading characteristic in frequency. Their research, continued in Geddes and Lee (2003a), shows way better correlation with subjective ratings, an evidence that masking cannot be left out from the distortion perception investigation.

These works clearly show that a deep survey of distortion perception is very hard proceeding on the traditional metric + test path. In fact, it will be always hard to relate a particular kind of non-linearity to a perceptual outcome. The path of modeling, even with generic models like the ones proposed in Section §4.2, can make the experimenters

able, in the future, to relate classes of non-linear models to perception. Models are more profitable as they can be tightly controlled by the experimenters, providing the key for a deeper understanding of the distortion perception. This will be possible, however, only if models relating well to the systems they have to represent are developed. Hence, the agreement between signals as obtained from systems and models should be determined quantitatively and perceptually.

The standard way to assess perceptual differences between pairs of signals is through the ABX test. This will fit the purpose well since, as reported by Hautus and Meng (2002), the ABX test performs well if the perceptual dimensions of a stimulus are unknown, as it is in distortion. In an ABX test a subject is presented with three stimuli: A and B plus a randomly chosen stimulus X between A and B. The task for the subject is to decide whether X is A or B. Experiment design theory and recommendations can be found in Bech and Zacharov (2006) as well as in International Telecommunication Union (2015) and International Telecommunication Union (2003). A brief overview can be found in Boley and Lester (2009) together with an analysis technique profitable in case of bias. In fact, bias towards an answer can always happen. As a result, the assumption at the base of the classical binomial analysis, that subjects will give random uncorrelated answers when unable to identify X correctly, could be broken. Signal Detection Theory, by modeling the subjects decision strategy in a suitable way, can test for significance also in this case. Hautus and Meng (2002) experiments found two decision strategies associated with the ABX task: the Difference Decision Strategy and the Independent Decision Strategy. When using the former the subjects, consciously or not, try to identify X by selecting its most similar stimulus. With the latter the decision is performed by assigning the stimuli to two different classes. Evidence from their research points out that the Independent Decision Strategy can be developed by usually a subset of trained listeners, that build the idea of classes through experience. Typical of naive subjects is the Difference Decision Strategy.

2.4 Summary

To summarize, nonlinear distortion is that kind of distortion dependent upon the signal level (Preis, 1984) while a nonlinear system is that kind of system whose action cannot be described with the superposition or homogeneity principle (Geddes & Lee, 2003b). The major kinds of distortion (Harmonic and Intermodulation) appear as direct consequences of the non-linear transfer characteristic of the systems components. Design choices in amplifiers can yield to different additional kinds of distortions, the most common of which is the Transient Intermodulation distortion (Dobrucki, 2011). The formalism of the mapping function for static memory-less systems is simple and it can be successfully employed in block models to identify NTI systems.

In fact, non-linearity can be effectively modeled with a PHM whose nonlinear blocks

are indeed static nonlinear functions. The linear blocks of this model are filters whose responses are related to the higher order impulse responses of the system, which link input frequency components to higher order harmonics (Novak, 2009; Novak et al., 2010a, 2010b). These responses can be obtained by means of properly designed SES stimulation (Farina, 2000; Novak et al., 2010b).

The study of the perception of distortion is complicated by the multiple degree of freedom both in distortion kind and human perception. As a consequence even the development of a proper metric for distortion is impervious (Geddes & Lee, 2003a, 2003b). Trying to explore the non-linearity perception using simplified models that can reduce the degree of freedom of the real nonlinear systems is then promising, as these reduced degree of freedom are controllable by the experimenter and correlation with perception can be better pointed out. However, this is advantageous only if the models are representative of real nonlinear systems. ABX tests are the standard for assessing differences in the perception of two stimuli (Hautus & Meng, 2002) and can be used to assess the performances of the models. In the case of bias the Signal Detection Theory can enhance the accuracy of the analysis (Boley & Lester, 2009). Being the ABX test widely known, details for its implementation can be found in classical sources as Bech and Zacharov (2006) and international standards as International Telecommunication Union (2015) and International Telecommunication Union (2003).

3 Measurement of nonlinear systems

3.1 Overview

To study a NTI system a stimulation signal must be designed. Farina et al. (2001) pointed out that higher order impulse responses can be retrieved with the use of Exponential Swept Sines. It is then natural to focus on Chirps. Moving from a rather abstract definition the scope is quickly restricted to Exponential Chirps (although most of the results are more general) as this kind of signal will be used (Section §3.2). As a start, the signals are described in the continuous time domain, the domain in which they are presented to the systems. However, these signals are generated and manipulated digitally. How their properties extend in the discrete time domain will be detailed as well (Section §3.3).

3.2 Input signal derivation and properties

3.2.1 Introduction

Symbol	Name	Meaning
$\wedge$	et	And
$\vee$	vel	Or
$\dot{}$, $\ddot{}$, $\dddot{}$, $\ddddot{}$	dots	Time derivative
$'$, $''$, $'''$, $''''$	primes	Frequency derivative
$\times$	times	Cartesian product
$*$	ast	Convolution

Table 3.1: Fundamental notation.

The mathematical fundamental properties of Chirps will be used to build the test signal. The choice and properties of the input stimulation are essential as they entail which NTI identification technique to use. For this reason they are detailed in depth to provide sufficient evidence to support section §4.

Signals can be represented as real functions of clear physical meaning or as complex functions useful to simplify mathematical manipulation. In the following the word Chirp will be used when dealing with the complex representation, while the real associated functions will be called Swept Sine or Cosine according to the context. table 3.1 sums up the fundamental notation. It is useful to refer to functional spaces. In use will be: the spaces $\mathscr{C}^k(A, B)$ of the functions defined in $A \subseteq \mathbb{R}^n$ valued in $B \subseteq \mathbb{R}^m$, with $n, m \in \mathbb{N}$ ([2]), that are continuous with all their derivatives up to the k-th order; the Hilbert spaces $\mathscr{L}^1(I, \mathbb{C})$ and $\mathscr{L}^2(I, \mathbb{C})$ of the functions defined in a compact real interval I of $\mathbb{R}$ valued in $\mathbb{C}$ whose magnitude and squared magnitude respectively are summable according to Lebesgue with finite integral over I. When the functions in $\mathscr{C}^k(A, B)$ are also limited

[2] In this work the natural numbers set $\mathbb{N}$ does not include 0.

Symbol	Transform	Action:
$\mathscr{H}[\bullet]$	Continuous time Hilbert transform	$x(t) : \mathbb{R} \to \mathbb{R} \longmapsto y(t) : \mathbb{R} \to \mathbb{C}$
$\mathscr{F}[\bullet]$	Continuous time Fourier transform	$x(t) : \mathbb{R} \to \mathbb{R} \longmapsto X(f) : \mathbb{R} \to \mathbb{C}$
$\mathcal{F}[\bullet]$	Discrete time Fourier transform	$x(t_d) : \mathbb{N} \to \mathbb{R} \longmapsto X(f_d) : \mathbb{N} \to \mathbb{C}$

Table 3.2: Functional transform notation. The most general domains are given. t_d and f_d are natural indices given according to the MATLAB® indexing rules for easier comparison with code.

the following norm can be defined

$$\|f\|_\infty \doteq \max_A \{f\} \quad \forall f \in \mathscr{C}^k(A, B)$$

while $\mathscr{L}^1(I, \mathbb{C})$ and $\mathscr{L}^2(I, \mathbb{C})$ are supplied naturally with the following norms:

$$\|f\|_1 \doteq \int_I |f|\, dm \quad \forall f \in \mathscr{L}^1(I, \mathbb{C})$$
$$\|f\|_2 \doteq \int_I |f|^2\, dm \quad \forall f \in \mathscr{L}^2(I, \mathbb{C})$$

where m is the usual measure associated with the semiring $\mathcal{I}$ of the sub-intervals in I:

$$m : \mathcal{I} \longrightarrow [0, +\infty)$$
$$\{U \subseteq I\} \in \mathcal{I} \overset{m}{\longmapsto} \sup\{U\} - \inf\{U\}$$

Finally, table 3.2 sums up the functional transforms notation.

3.2.2 Asymptotic Chirps

A Chirp is a complex signal modulated in magnitude and frequency. The following definition incorporates conditions suggested by Chassande-Mottin and Flandrin (1998) sufficient to derive the most important results and build the test signal.

Definition 3.1 (Chirp Signal)**.** Let $A(t)$ and $\varphi(t)$ be two functions of $\mathscr{C}^2(\Omega, \mathbb{R})$, with $\Omega \subset \mathbb{R}$ a compact interval, such that, $\forall t \in \Omega$:

$$A(t) \neq 0 \wedge \dot{\varphi}(t) \neq 0$$
$$\left|\frac{\dot{A}(t)}{A(t)\,\dot{\varphi}(t)}\right| \ll 1 \wedge \left|\frac{\ddot{\varphi}(t)}{\dot{\varphi}^2(t)}\right| \ll 1 \tag{3.1}$$

Hence, a Chirp Signal $s(t)$ is defined as:

$$s : \Omega \longrightarrow \mathbb{C}, \quad s(t) = A(t)\, e^{j\varphi(t)} \quad \forall t \in \Omega$$

This definition ensures control over the growth of the functions defining $s(t)$, forcing the eventual oscillations spanned by $A(t)$ to be slow compared to the oscillations resulting from $\varphi(t)$.

Proposition 3.2. *Let $\alpha \neq 0$, $\beta \neq 0$, $\gamma \neq 0$, δ be four real constants and $\Omega \subset \mathbb{R}$ a compact interval. Then it exists one and only one real constant ϵ dependent on γ and Ω such that the signal*

$$s : \Omega \longrightarrow \mathbb{C}, \quad s(t) = \alpha \exp\left(j\left[\beta\gamma e^{t/\gamma} - \delta\right]\right) \quad \forall t \in \Omega$$

is a Chirp Signal if and only if $|\beta| \gg \epsilon$. In particular, ϵ is given by:

$$\epsilon \doteq \left|\frac{1}{\gamma} e^{-\widetilde{t}/\gamma}\right|, \quad \widetilde{t} \doteq \begin{cases} \inf\{\Omega\} & if\ \gamma > 0 \\ \sup\{\Omega\} & if\ \gamma < 0 \end{cases}$$

Remark 3.3. When $|\beta| \gg \epsilon$ $s(t)$ can be rewritten as in Definition 3.1 by choosing

$$A(t) = \alpha \quad \forall t \in \Omega$$
$$\varphi(t) = \beta\gamma e^{t/\gamma} - \delta \quad \forall t \in \Omega$$

The derivatives are:

$$\dot{A}(t) = 0 \quad \forall t \in \Omega$$
$$\dot{\varphi}(t) = \frac{\beta\cancel{\gamma}}{\cancel{\gamma}} e^{t/\gamma} \neq 0 \quad \forall t \in \Omega$$
$$\ddot{\varphi}(t) = \frac{\beta}{\gamma} e^{t/\gamma} = \frac{1}{\gamma}\dot{\varphi}(t) \neq 0 \quad \forall t \in \Omega$$

Proof of Proposition 3.2 is given in Appendix A.1. This leads naturally to:

Definition 3.4 (Exponential Chirp)**.** The complex signal $s(t)$ of Proposition 3.2 is defined Exponential Chirp if $|\beta| \gg \epsilon$.

If the frequency content of a Chirp is of interest it is natural to resort the the Fourier transform to describe it:

$$S : \{\Phi \subseteq \mathbb{R}\} \longrightarrow \mathbb{C}$$
$$S(f) \doteq \mathscr{F}[s(t)] = \frac{1}{\sqrt{2\pi}} \int_\Omega s(t) e^{-j2\pi f t} dt = \frac{1}{\sqrt{2\pi}} \int_\Omega A(t) e^{j\varphi(t)} e^{-j2\pi f t} dt \doteq$$
$$\doteq \int_\Omega b(t) e^{j\psi(t)} dt = B(f) e^{j\Psi(f)} \quad \forall f \in \Phi \quad (3.2)$$

where Φ is an interval of $\mathbb{R}$ ([3]), $b(t) \doteq {}^{A(t)}/_{\sqrt{2\pi}}$ and $\psi(t) \doteq \varphi(t) - 2\pi f t$. These integrals are referred as oscillatory integrals and their calculation is impervious. However, it can

[3] It could be in principles unbounded, but it will be supposed compact later.

be heuristically grasped that in regions where the oscillations of the integrand are fast the integral will converges towards zero, as a consequence of numerous contributions of opposite sign. Supposing that the oscillations are controlled by $\psi(t)$ (i.e. $b(t)$ is slowly varying), the non-zero contributions from the integral will came from the regions around the points where $\dot{\psi}(t) = 0$, i.e. the phase is slowly varying. Through series expansion of $\psi(t)$ an approximated solution can be found in these regions. This technique is called asymptotic approximation and its accuracy depends on the frequency range. A signal that can be approximated equally well in the whole interval Φ by the asymptotic approximation is referred as asymptotic (Cohen, 2000). Asymptotic signals are characterized by simple closed relations between their time domain definition and frequency domain representation.

In the following, to start moving towards the signal used for the measurements, the focus will be on the Exponential Chirps. From now α, β and γ are assumed positive[4].

The first step is to study $\psi(t)$:

$$\begin{aligned} \psi(t) &\doteq \varphi(t) - 2\pi f t \\ \implies \dot{\psi}(t) = \dot{\varphi}(t) - 2\pi f = 0 &\Leftrightarrow \dot{\varphi}(t) = 2\pi f \end{aligned} \tag{3.3}$$

From Remark 3.3:

$$\dot{\varphi}(t) = \beta e^{t/\gamma} = 2\pi f \Leftrightarrow t = \gamma \ln\left(\frac{2\pi f}{\beta}\right) \doteq t_f \wedge f > 0 \tag{3.4}$$

where the last condition is a consequence of $\dot{\varphi}(t)$ being strictly positive everywhere in Ω, making $\varphi(t)$ strictly monotonic. Hence, the following is understood

$$\forall f \in \Phi : f > 0 \quad \exists! \, t_f : \dot{\psi}(t_f) = 0; \quad \forall f \in \Phi : f \leq 0 \quad \nexists t_f : \dot{\psi}(t_f) = 0 \tag{3.5}$$

So, there are two cases depending on frequency:

1. There aren't stationary phase points ($f \leq 0$).
2. There is one and only one stationary point ($f > 0$).

Moreover:

$$\ddot{\psi}(t) = \ddot{\varphi}(t) = \frac{1}{\gamma}\dot{\varphi}(t) \implies \ddot{\psi}(t)\Big|_{t=t_f} = \frac{1}{\gamma}\dot{\varphi}(t_f) = \frac{2\pi f}{\gamma} \neq 0 \, \forall f \neq 0 \tag{3.6}$$

ensuring that if t_f is a stationary point for $\psi(t)$ then it is a non-degenerate stationary point. Following Chassande-Mottin and Flandrin (1998) the cases of one stationary point and absence of stationary points are quickly treated.

[4]This is not a big restriction as similar results will hold for negative constants.

One stationary point ($f > 0$) Being t_f non degenerate the following change of variable can be performed

$$u^2 = 2\frac{\psi(t) - \psi(t_f)}{\ddot{\psi}(t_f)} \Longrightarrow \psi(t) = \frac{\ddot{\psi}(t_f)}{2}u^2 + \psi(t_f)$$

$$\Longrightarrow \int_\Omega b(t)\, e^{j\psi(t)}dt = e^{j\psi(t_f)} \int_\Gamma b(t(u)) \exp\left(j\frac{\ddot{\psi}(t_f)}{2}u^2\right)\frac{dt}{du}du$$

where Γ is Ω as transformed by the change of variable. Given that all the functions involved in the exponential in the right hand are continuous with all their derivatives (see Remark 3.3) the exponential itself can be expressed as a Taylor series. The series can be truncated at the order $\overline{k}$ introducing a reminder of the kind $o\left((u-u_0)^{\overline{k}}\right)$, with u_0 the point chosen as center for the expansion (Gilardi, 2011). As a result, with $\overline{k} = 2$:

$$\int_\Omega b(t)\, e^{j\psi(t)}dt = \underbrace{\sqrt{\frac{2\pi}{\left|\ddot{\psi}(t_f)\right|}}b(t_f)\, e^{j\psi(t_f)}e^{j\left[\text{sign}\left(\ddot{\psi}(t_f)\right)\right]\pi/4}}_{\doteq I} + R \tag{3.7}$$

where R is a reminder, consequence of the remainder due to the truncation of the Taylor series. Defining Φ^+ as $\{f \in \Phi : f > 0\}$ and $\Upsilon \doteq \Omega \times \Phi^+ \subset \mathbb{R}^2$ the relative error Σ can be defined and bounded as follows:

$$\Sigma \doteq \left|\frac{R}{I}\right| \leq \sup_{t\in\Omega}\{Q(t,\, f)\} \tag{3.8}$$

where $Q(t,\, f) : \Upsilon \longrightarrow \mathbb{R}^+$ has been reported in Chassande-Mottin and Flandrin (1999) for chirps obeying Definition 3.1:

$$Q(t,\, f) \doteq 5\sqrt{2\left|\ddot{\Psi}(t_f)\right|}\left|\frac{b(t)}{b(t_f)}\frac{\sqrt{\Psi(t)}}{\dot{\Psi}(t)}\right|\left|\frac{\ddot{b}(t)}{b(t_f)}\frac{\Psi(t)}{\dot{\Psi}^2(t)} + \frac{3}{2}\frac{\dot{b}(t)}{b(t)\dot{\Psi}(t)}\left(1 - \frac{\Psi(t)\ddot{\Psi}(t)}{\dot{\Psi}^2(t)}\right) + \right.$$

$$\left. + \left[3\Psi(t)\left(\frac{\ddot{\Psi}(t)}{\dot{\Psi}^2(t)}\right)^2 - \frac{3}{2}\frac{\ddot{\Psi}(t)}{\dot{\Psi}^2(t)} - \frac{\Psi(t)\dddot{\Psi}(t)}{\dot{\Psi}^3(t)}\right]\right|$$

$$\Psi \doteq \psi(t) - \psi(t_f) \tag{3.9}$$

$Q(t,\, f)$ depends on f through $t_f = \gamma \ln\left(\frac{2\pi f}{\beta}\right)$ and $\psi(t) = \varphi(t) - 2\pi f t$. In other words, $\sup_{t\in\Omega}\{Q(t,\, f)\}$ describes a function of frequency bounding Σ. It is simple to see that all the functions used to define $Q(t,\, f)$ have limited magnitude in Ω every time they are restricted to it by choosing a constant value of f. Hence, for each chosen value of frequency all the magnitudes in $Q(t,\, f)$ are limited in Ω, making $\sup_{t\in\Omega}\{Q(t,\, f)\}$ limited as well. Proof is given in Proposition A.1.

To finish, it is evident that $B(f)$ and $\Psi(f)$ can be identified, neglecting any error, by:

$$B(f) = \sqrt{\frac{2\pi}{\left|\ddot{\psi}(t_f)\right|}} b(t_f) = \frac{A(t_f)}{\sqrt{|\ddot{\varphi}(t_f)|}} \tag{3.10}$$
$$\Psi(f) = \psi(t_f) + \left[\operatorname{sign}\left(\ddot{\psi}(t_f)\right)\right]\frac{\pi}{4} = \varphi(t_f) - 2\pi f t_f + \left[\operatorname{sign}\left(\ddot{\varphi}(t_f)\right)\right]\frac{\pi}{4}$$

where the facts that $b(t) \doteq {}^{A(t)}/_{\sqrt{2\pi}}$ and $\psi(t) \doteq \varphi(t) - 2\pi f t$ have been used.

No stationary points ($f \leq 0$) As a consequence of equation (3.5) manipulation of the integrand is possible by division and multiplication by $\dot{\psi}(t)$. By denoting the oscillatory integral approximation by I again, use of the by part integration, triangular inequality and integral bounding yields to:

$$|I| \leq \left|\frac{b(\sup\{\Omega\})}{\dot{\psi}(\sup\{\Omega\})}\right| + \left|\frac{b(\inf\{\Omega\})}{\dot{\psi}(\inf\{\Omega\})}\right| + \left(\left\|\frac{\dot{b}(t)}{b(t)\dot{\psi}(t)}\right\|_\infty + \left\|\frac{\ddot{\psi}(t)}{\left(\dot{\psi}(t)\right)^2}\right\|_\infty\right)\|b(t)\|_1 \leq$$
$$\leq \frac{1}{\sqrt{2\pi}}\left[\left(\left|\frac{A(\sup\{\Omega\})}{\dot{\varphi}(\sup\{\Omega\})}\right| + \left|\frac{A(\inf\{\Omega\})}{\dot{\varphi}(\inf\{\Omega\})}\right|\right) + \right.$$
$$\left. + \left(\left\|\frac{\dot{A}(t)}{A(t)\dot{\varphi}(t)}\right\|_\infty + \left\|\frac{\ddot{\varphi}(t)}{(\dot{\varphi}(t))^2}\right\|_\infty\right)\|A(t)\|_1\right] \tag{3.11}$$

where the last inequality is written exploiting the facts that $\dot{\psi}(t) = \dot{\varphi}(t) - 2\pi f \geq \dot{\varphi}(t)$ $\forall f \leq 0$, $\forall t \in \Omega$, $\ddot{\psi}(t) = \ddot{\varphi}(t)$ $\forall t \in \Omega$ and remembering that $b(t) \doteq {}^{A(t)}/_{\sqrt{2\pi}}$. The conditions of Definition 3.1 are then sufficient to demonstrate that the more rigorously a signal is a Chirp the less important are the contributions to the oscillatory integral away from stationary points, as they work to lower the quantities in the last member above. Moreover, as the region free of stationary points is the negative frequency region, it is understood that this kind of Chirps can be considered quasi-analytic. Derivation of equation (3.11) is reported in Appendix A.2. Lastly, from Remark 3.3, for an Exponential Chirp:

$$|I| \leq \frac{|\alpha|}{\sqrt{2\pi}}\left(\frac{1}{|\beta e^{\sup\{\Omega\}/\gamma}|} + \frac{1}{|\beta e^{\inf\{\Omega\}/\gamma}|} + m(\Omega)\left\|\frac{1}{\beta\gamma}e^{-t/\gamma}\right\|_\infty\right) \doteq \Theta$$

Conclusion The Chirps of Definition 3.1 (as Exponential Chirps are) have two important properties. They are asymptotic, meaning that the error committed in resorting to the asymptotic approximation when calculating the frequency content of the signal is controlled by a well defined quantity, the rightmost in equation (3.8). Also, they are quasi-analytic, meaning that the error committed in considering them as analytic signals is controlled by the rightmost quantity in equation (3.11), which is minimized by respecting equation (3.1).

3.2.3 Exponential Swept Sines as asymptotic signals

A further step towards the actual stimulation is performed by taking the physical signals into consideration:

Definition 3.5 (Exponential Swept Sine)**.** Let $\alpha > 0$, $\beta > 0$, $\gamma > 0$, δ be four real constants and $\Omega \subset \mathbb{R}$ a compact interval chosen such that

$$|\beta| \gg \epsilon \doteq \left| \frac{1}{\gamma} e^{-\widetilde{t}/\gamma} \right|, \quad \widetilde{t} \doteq \inf \{\Omega\}$$

An Exponential Swept Sine $z(t)$ is then defined as follows:

$$z : \Omega \longrightarrow \mathbb{R}, \quad z(t) = \alpha \sin\left(\beta\gamma e^{t/\gamma} - \delta\right) \quad \forall t \in \Omega$$

This definition is meant to mirror Definition 3.4, which is its generalization in complex formalism, and Remark 3.3 still holds. The generalization is achieved by resorting to the analytic signal representation. This result, proven in Appendix A.3, better characterizes the signal:

Lemma 3.6. *Let $s(t)$ be an Exponential Chirp as in Definition 3.4 and suppose that $m(\Phi)$, with Φ the interval of definition of $S(f) = \mathscr{F}[s(t)]$, is not infinite. Hence:*

$$s(t) \in \mathscr{L}^2(\Omega, \mathbb{C}) \wedge S(f) = B(f) e^{j\Psi(f)} \in \mathscr{L}^2(\Phi, \mathbb{C})$$

The condition $m(\Phi) < +\infty$ means that the support of $S(f)$ is not infinite, or limited spectrum. This is a condition that can be easily met in practice. In particular, it will be always met in the digital realm, where all the spectra must be limited according to the Nyquist theorem. In the spaces $\mathscr{L}^2(I, \mathbb{C})$, the following operator can be defined:

$$\sigma[x(t)] \doteq \int_I t^2 |x(t)|^2 \, dt \quad \forall x(t) \in \mathscr{L}^2(I, \mathbb{C})$$

namely the variance of the signal. As a consequence, following Boashash (1992), the effective duration $\mathcal{T}$ and the effective bandwidth $\mathcal{B}$ of the signal $s(t)$ can be defined as follows:

$$\mathcal{T}^2 \doteq \frac{\sigma[s(t)]}{\|s(t)\|_2^2}, \quad \mathcal{B}^2 \doteq \frac{\sigma[S(f)]}{\|S(f)\|_2^2}$$

With all these objects the following result, reviewed by Boashash (1992), can be proven:

Theorem 3.7. *Let $z(t)$ be an Exponential Swept Sine as in Definition 3.5 and let $s(t) = z(t) + j\mathscr{H}[z(t)]$ be its analytic signal. Hence*[5]*:*

$$\lim_{\mathcal{BT} \to +\infty} s(t) = \alpha \exp\left(j\left[\beta\gamma e^{t/\gamma} - \delta\right]\right)$$

[5] Boashash (1992) illustrated the result for a more general Swept Cosine function. The fact that a cosine was used is not essential, since is easy to refer to the same hypothesis by proper choice of δ.

This is a first important property of the Exponential Swept Sines: they can be designed so that their analytic signal is supplied, with good approximation, by $s(t)$ of Definition 3.4. This makes sense since, as it was shown, these Exponential Chirps are indeed quasi-analytic by nature and can be regarded as analytic with proper choice of the definition constants whereas these produce negligible errors in applications. Having found the bridge between $s(t)$ and $z(t)$ the description in terms of asymptotic signal for a Exponential Swept Sine can be performed. In general, the instantaneous frequency $\nu(t)$ and group delay $\tau(f)$ are defined as (Boashash, 1992; Cohen, 2000; Novak et al., 2010b):

$$\begin{aligned} \nu(t) &\doteq \frac{1}{2\pi}\dot{\varphi}(t) \\ \tau(f) &\doteq -\frac{1}{2\pi}\Psi'(f) \end{aligned} \tag{3.12}$$

these definitions stem from the Asymptotic approximation, proven to be a good approximation for the signals under analysis. In fact, the first equation in equation (3.12) is derived with the same technique of equation (3.3), which is indeed very similar. A similar reasoning is applied to get the second equation, but to the inverse Fourier transform linking $S(f)$ and $s(t)$. With the new definitions, Remark 3.3 and equation (3.4) it is possible to rewrite equation (3.10):

$$B(f) = \frac{A(t_f)}{\sqrt{|\ddot{\varphi}(t_f)|}} = \frac{\alpha}{\sqrt{\left|\frac{\beta}{\gamma}e^{t_f/\gamma}\right|}} = \frac{\alpha}{\sqrt{\frac{\beta}{\gamma}e^{t_f/\gamma}}} = \frac{\alpha}{\sqrt{\frac{\cancel{\beta}}{\gamma}\frac{2\pi f}{\cancel{\beta}}}} = \alpha\sqrt{\frac{\gamma}{2\pi f}} \tag{3.13}$$

$$\begin{aligned} \Psi(f) &= \psi(t_f) + \left[\operatorname{sign}\left(\ddot{\psi}(t_f)\right)\right]\frac{\pi}{4} = \varphi(t_f) - 2\pi f t_f + \left[\operatorname{sign}\left(2\pi\dot{\nu}(t_f)\right)\right]\frac{\pi}{4} \\ &= \beta\gamma e^{t_f/\gamma} - \delta - 2\pi f t_f + \frac{\pi}{4}\left[\operatorname{sign}\left(\cancel{2\pi}\frac{1}{\cancel{2\pi}}\ddot{\varphi}(t_f)\right)\right] = \\ &= \beta\gamma\exp\left(\frac{1}{\cancel{\gamma}}\cancel{\gamma}\ln\left(\frac{2\pi f}{\beta}\right)\right) - \delta - 2\pi f\gamma\ln\left(\frac{2\pi f}{\beta}\right) + \\ &\quad + \frac{\pi}{4}\left[\operatorname{sign}\left(\frac{\beta}{\gamma}\exp\left(\frac{1}{\cancel{\gamma}}\cancel{\gamma}\ln\left(\frac{2\pi f}{\beta}\right)\right)\right)\right] = \\ &= \cancel{\beta}\gamma\frac{2\pi f}{\cancel{\beta}} - \delta - 2\pi f\gamma\ln\left(\frac{2\pi f}{\beta}\right) + \frac{\pi}{4}\operatorname{sign}\left(\frac{\cancel{\beta}}{\gamma}\frac{2\pi f}{\cancel{\beta}}\right) = \\ &= 2\pi\gamma f - \delta - 2\pi f\gamma\ln\left(\frac{2\pi f}{\beta}\right) + \frac{\pi}{4}\operatorname{sign}\left(\frac{2\pi f}{\gamma}\right) \end{aligned} \tag{3.14}$$

Then, equation (3.12) for Exponential Swept Sines reads, thanks to equation (3.14) and equation (3.4):

$$\nu(t) = \frac{\beta}{2\pi}e^{t/\gamma}$$

$$\begin{aligned}\tau(f) &= -\frac{1}{2\pi}\frac{d}{df}\left[2\pi\gamma f - \delta - 2\pi f\gamma\ln\left(\frac{2\pi f}{\beta}\right) + \frac{\pi}{4}\text{sign}\left(\frac{2\pi f}{\gamma}\right)\right] = \\ &= -\frac{1}{2\pi}\left[2\pi\gamma - \left\{2\pi\gamma\ln\left(\frac{2\pi f}{\beta}\right) + \cancel{2\pi f}\gamma\frac{\cancel{\beta}}{\cancel{2\pi f}}\frac{2\pi}{\cancel{\beta}}\right\}\right] = \\ &= -\frac{1}{2\pi}\left[\cancel{2\pi\gamma}\underbrace{-2\pi\gamma\ln\left(\frac{2\pi f}{\beta}\right)}_{\Psi'(f)=-2\pi t_f} - \cancel{2\pi\gamma}\right] = \gamma\ln\left(\frac{2\pi f}{\beta}\right) = t_f\end{aligned} \tag{3.15}$$

Inverting the two:

$$\begin{aligned}t(\nu) &= \gamma\ln\left(\frac{2\pi\nu}{\beta}\right) \\ f(\tau) &= \frac{\beta}{2\pi}e^{\tau/\gamma}\end{aligned} \tag{3.16}$$

Comparison of equation (3.12) and equation (3.16) shows that the functions linking $t \longmapsto \nu$ and $\tau \longmapsto f$ are the same. Moreover, also $\nu \longmapsto t$ and $f \longmapsto \tau$ are the same and are related to the previous ones by inversion. In fact, for asymptotic signals such that $\nu(t)$ is monotonic and $\mathcal{TB} \gg 1$ the following holds (Boashash, 1992):

$$\begin{aligned}f &= \frac{1}{2\pi}\dot{\varphi}(\tau) = f(\tau) \\ t &= -\frac{1}{2\pi}\Psi'(\nu) = t(\nu)\end{aligned} \tag{3.17}$$

meaning that the functions of equation (3.12) can be considered the inverse of each other, making possible the following identifications:

$$f = \nu(t)\,,\ \wedge t = \tau(f) = t_f \tag{3.18}$$

3.2.4 The stimulation signal

All the material needed to define the final form of the test signal is ready. The objective is to design an Exponential Swept Sine of duration $T = m(\Omega)$ whose values of frequencies assumed at the lower and upper extremes of Ω are set to be f_{start} and f_{end} respectively. This is achieved by setting the correct values of the constants α, β, γ and δ of Definition 3.5. Ω can be any time interval such that $m(\Omega) = T$ but the identification $[0, T]$ is most convenient. Following Farina (2000), from equation (3.12), equation (3.17) and equation (3.18):

$$\begin{aligned}\nu(t) = f = \frac{1}{2\pi}\dot{\varphi}(t) = \frac{\beta}{2\pi}e^{t/\gamma} \\ \Longrightarrow\ f|_{t=0} = \frac{\beta}{2\pi} = f_{start} \Leftrightarrow \beta = 2\pi f_{start}\ \wedge\ f|_{t=T} = \frac{\beta}{2\pi}e^{T/\gamma} =\end{aligned}$$

$$= \frac{\cancel{2\pi} f_{start}}{\cancel{2\pi}} e^{T/\gamma} = f_{end} \Leftrightarrow \gamma = \frac{T}{\ln\left(\frac{f_{end}}{f_{start}}\right)}$$

To have unit amplitude and 0 rad initial phase ($\varphi(0\,\mathrm{s}) = 0\,\mathrm{rad}$) $\alpha = 1$ and $\delta = \beta\gamma$ are imposed. Hence:

$$z(t) = \sin\left(\frac{2\pi f_{start} T}{\ln\left(\frac{f_{end}}{f_{start}}\right)}\left[\exp\left(\frac{t}{T}\ln\left(\frac{f_{end}}{f_{start}}\right)\right) - 1\right]\right)$$

However, this definition is not best suited for the study of the response of nonlinear systems. In fact, as said above, the higher order responses will appear at negative times, with lags $\triangle t_m$. If at each of these lags the phase of $z(t)$ is different this will be reflected in the various impulse responses phases, introducing artifacts. A good way to synchronize the phases has been proposed by Novak et al. (2010b). The idea is to synchronize $z(t)$ so that when it reaches an harmonic mf_{start} of the starting frequency, $m \in \mathbb{N}$, it passes through 0 with a positive slope. Doing so, the sweep appears as a train of sweeps, each one growing its frequency on the same exponential trend and starting from a different harmonic. Quantitatively, a sweep $z_m(t)$ having m times the instantaneous frequency of $z(t)$ will be related to it by:

$$z_m(t) = z(t + \triangle t_m) \tag{3.19}$$

The lags $\triangle t_m$ are found by

$$\nu(\triangle t_m) = f = mf_{start}, \quad m \in \{q \in \mathbb{N} : q \geq 2\} \doteq \mathbb{H} \tag{3.20}$$

From equation (3.16) and the chosen values of α, β, γ and δ, $\forall m \in \mathbb{H}$:

$$\triangle t_m = t(\nu(\triangle t_m)) = t(mf_{start}) = \gamma \ln\left(\frac{2\pi m f_{start}}{\beta}\right) = \gamma \ln(m)$$

$$\implies \varphi(\triangle t_m) = \beta\gamma \exp\left(\frac{\triangle t_m}{\gamma}\right) - \delta = \beta\gamma m - \beta\gamma = \beta\gamma(m-1)$$

The synchronization requires:

$$\begin{cases} z(\triangle t_m) = 0 \\ \dot{z}(\triangle t_m) > 0 \end{cases} = \begin{cases} \sin(\varphi(\triangle t_m)) = 0 \\ \cos(\varphi(\triangle t_m))\,\dot{\varphi}(\triangle t_m) > 0 \end{cases} \quad \forall m \in \mathbb{H}$$

$$\Leftrightarrow \begin{cases} \varphi(\triangle t_m) = k\pi \\ \cos(\varphi(\triangle t_m))\,\dot{\varphi}(\triangle t_m) > 0 \end{cases} \quad k \in \mathbb{Z},\ \forall m \in \mathbb{H}$$

For the inequality in cos: being $\varphi(t)$ an exponential its derivative is an exponential as well and it will be always positive. So, only the cosine can set the sign. A simple plot of the functions shows that the only points such that the cosine is positive and the sine is

zero are of the kind $2k\pi$. Hence, $\forall m \in \mathbb{H}$:

$$\begin{cases} z\left(\triangle t_m\right) = 0 \\ \dot{z}\left(\triangle t_m\right) > 0 \end{cases} \Leftrightarrow \varphi\left(\triangle t_m\right) = \beta\gamma\left(m-1\right) = 2\pi f_{start}\frac{T}{\ln\left(\frac{f_{end}}{f_{start}}\right)}\left(m-1\right) = 2k\pi \Leftrightarrow$$

$$\Leftrightarrow f_{start}\frac{T}{\ln\left(\frac{f_{end}}{f_{start}}\right)}\left(m-1\right) = f_{start}\gamma\left(m-1\right) = k$$

Since $k \in \mathbb{Z}$ and $(m-1) \in \mathbb{N}$ the above equation can be satisfied by choosing any combination of f_{start} and γ yielding a product in $\mathbb{Z}$. The most practical way, suggested by Novak et al. (2010b), is to assign f_{start} and $f_{end} = Mf_{start}$, with $M \in \mathbb{H}$, and a coarse time length $\widetilde{T}$. Then, operate the following substitution[6]:

$$\gamma \longrightarrow \frac{1}{f_{start}}\text{round}\left(f_{start}\frac{\widetilde{T}}{\ln\left(\frac{f_{end}}{f_{start}}\right)}\right)$$

$$\text{round}\left(f_{start}\frac{\widetilde{T}}{\ln\left(\frac{f_{end}}{f_{start}}\right)}\right) \doteq \left\{\text{sign}\left(f_{start}\frac{\widetilde{T}}{\ln\left(\frac{f_{end}}{f_{start}}\right)}\right)\left\lfloor\left|f_{start}\frac{\widetilde{T}}{\ln\left(\frac{f_{end}}{f_{start}}\right)}\right| + \frac{1}{2}\right\rfloor\right\} \tag{3.21}$$

At this point the growth of the instantaneous frequency is defined and the actual time length T can be set to be the length at which the Swept Sine naturally reaches the M-th harmonic: $T = \gamma\ln\left(M\right)$. The so obtained signals are the Synchronized Exponential Swept Sine (SES) that were being searched for.

3.2.5 Exponential Swept Sines and systems theory

Inverse filtering All the properties of LTI systems are contained in one single function of time, the impulse response:

$$h : \mathbb{R} \longrightarrow \mathbb{R}$$

In principles $h\left(t\right)$ can have support defined over an unbounded interval of $\mathbb{R}$, but in practice it is considered restricted over a compact interval as a consequence of the noise floor, covering its weaker parts, and the need to process finite length signals in real life. An input signal $x\left(t\right) \in \mathscr{L}^2\left(\mathbb{R}, \mathbb{R}\right)$ having finite support can be defined extending $z\left(t\right)$ of Definition 3.5 by zero padding. The output $y : \mathbb{R} \longrightarrow \mathbb{R}$ of the LTI system will be given by convolution (Kuttruff, 2009, pp. 255-261):

$$y\left(t\right) = x\left(t\right) * h\left(t\right) \doteq \int_{\mathbb{R}} x\left(\varsigma\right)h\left(t-\varsigma\right)d\varsigma$$

[6] Any form of rounding will actually work, as the goal is to produce γ such that, multiplied by f_{start}, it gives an integer result. This particular rounding is chosen since it is the one implemented in the MATLAB® function `round` which has tie breaking away from 0.

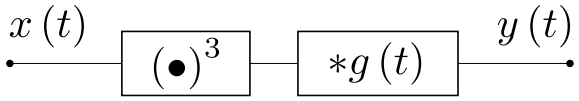


Figure 3.1: Simple nonlinear system. $(\bullet)^3$ represents a powering operation of the input. $g(t)$ is the impulse response of a linear filter.

with ς a dummy integration variable. If $x(t)$ and $h(t)$ have compact support then also $y(t)$ will have compact support. The system is said to be stable if the following applies:

$$\text{if } \exists \widehat{x} \,:\, |x(t)| \leq \widehat{x} \forall t \in \mathbb{R} \implies \exists \widehat{y} \,:\, |y(t)| \leq \widehat{y} \forall t \in \mathbb{R}$$

All the systems under analysis will be stable and $\widehat{x}$, being $x(t)$ defined in terms of $z(t)$, exists. A common way to retrieve $h(t)$ is through inverse filtering (Kuttruff, 2009, pp. 255-261):

$$h(t) = y(t) * \overleftarrow{x}(t) \tag{3.22}$$

with the inverse filter a signal $\overleftarrow{x}(t) \in \mathscr{L}^2(\mathbb{R}, \mathbb{R})$ such that

$$\delta(t) = x(t) * \overleftarrow{x}(t) \tag{3.23}$$

where $\delta(t)$ is the Dirac Delta distribution, acting over $\mathscr{L}^2(\mathbb{R}, \mathbb{R})$, obtained by proper generalization of the convolution operation[7].

For Exponential Swept Sines the same method, valid for linear systems, can be used on nonlinear systems as well. Following Novak (2009, p. 51), imagine a simple nonlinear model as depicted in figure 3.1. Then, for $x(t) = z(t)$ of Section §3.2.4:

$$\begin{aligned}
[\sin(\varsigma)]^3 = \frac{3}{4}\sin(\varsigma) - \frac{1}{4}\sin(3\varsigma) \quad \forall \varsigma \in \mathbb{R} \\
\implies x^3(t) = \frac{3}{4}x(t) - \frac{1}{4}x(t + \triangle t_3) \\
\implies y(t) = x^3(t) * g(t) = \frac{3}{4}x(t) * g(t) - \frac{1}{4}x(t + \triangle t_3) * g(t)
\end{aligned}$$

where the first equation is a known trigonometric identity and the second comes from equation (3.19). Hence:

$$\begin{aligned}
y(t) * \overleftarrow{x}(t) &= \frac{3}{4}x(t) * g(t) * \overleftarrow{x}(t) - \frac{1}{4}x(t + \triangle t_3) * g(t) * \overleftarrow{x}(t) = \\
&= \frac{3}{4}x(t) * \overleftarrow{x}(t) * g(t) - \frac{1}{4}x(t + \triangle t_3) * \overleftarrow{x}(t) * g(t) = \\
&= \frac{3}{4}\delta(t) * g(t) - \frac{1}{4}\delta(t + \triangle t_3) * g(t) = \frac{3}{4}g(t) - \frac{1}{4}g(t + \triangle t_3)
\end{aligned}$$

where the algebraic and distributional properties of the convolution have been used. This

[7] A very brief definition is given in Appendix A.4.1.

shows that a shifted impulse response has been produced in the negative times, at a lag $\triangle t_3$. As it will be clarified in section §4, models of this kind (Hammerstein models) are the building blocks for multi-branch models ($MISO$ Hammerstein models) that, as reported by Novak (2009, p. 14), can represent any nonlinear system described by differential or integro-differential equations. Generalization of the model to multiple parallel branches of the kind of figure 3.1, with different powers and impulse responses, is thus fairly representative of a general nonlinear system. Application of the same reasoning generalized to the multi-branch case yields to the superposition, for negative times, of different higher order impulse responses. Hence, equation (3.22) must be rewritten for NTI systems as follows (Novak et al., 2010b):

$$\boxed{y(t) * \overleftarrow{x}(t) = \widetilde{h}(t) \doteq \sum_{m=1}^{M} h_m(t + \triangle t_m)} \tag{3.24}$$

where $h_m(t)$ are the higher order impulse responses relating an input frequency to its m-th harmonic, centered at the time $-\triangle t_m$, where $\triangle t_m$ are the times at which the Exponential Swept Sine reaches the m-th harmonic of the starting frequency (see equation (3.20)). It is to be noted that $\widetilde{h}(t)$ is an improper impulse response given by a superposition of proper impulse responses. Provided that the Exponential Swept Sine is long enough the overlap between the various $h_m(t + \triangle t_m)$ will be minimal and they can be safely gated out.

Calculation of the inverse filter The inverse filter $\overleftarrow{x}(t)$ can be calculated resorting to the analytic signals:

$$x(t) = \begin{cases} z(t) & \forall t \in \Omega \\ 0 & \forall t \in \mathbb{R} \setminus \Omega \end{cases} \overset{\mathscr{H}}{\longmapsto} \begin{cases} s(t) & \forall t \in \Omega \\ 0 & \forall t \in \mathbb{R} \setminus \Omega \end{cases}$$

$$\overleftarrow{x}(t) = \begin{cases} \overleftarrow{z}(t) & \forall t \in \overleftarrow{\Omega} \\ 0 & \forall t \in \mathbb{R} \setminus \overleftarrow{\Omega} \end{cases} \overset{\mathscr{H}}{\longmapsto} \begin{cases} \overleftarrow{s}(t) & \forall t \in \overleftarrow{\Omega} \\ 0 & \forall t \in \mathbb{R} \setminus \overleftarrow{\Omega} \end{cases}$$

where $s(t)$, supposing that the hypotheses of Theorem 3.7 hold, is given by Definition 3.4. It is sufficient to restrict to Ω for the calculation and then extend the final result. Equation (3.23) applies also to the analytic signals in their home space $\mathscr{L}^2(\Omega, \mathbb{C})$ and in the frequency domains reads (Novak et al., 2010b), using equation (3.2):

$$\delta(t) = x(t) * \overleftarrow{x}(t) \Leftrightarrow 1 = S(f)\,\overleftarrow{S}(f)$$

$$\Longrightarrow \overleftarrow{S}(f) = \frac{1}{S(f)} = \frac{1}{B(f)\,e^{j\Psi(f)}} = \frac{1}{B(f)} e^{-j\Psi(f)}$$

By analogy with equation (3.2), which is general, one can write $\overleftarrow{S}(f) = \overleftarrow{B}(f) e^{j\overleftarrow{\Psi}(f)}$. Hence:

$$\overleftarrow{B}(f) = \frac{1}{B(f)} \wedge \overleftarrow{\Psi}(f) = -\Psi(f)$$

meaning that $\overleftarrow{S}(f)$ respects the same kind of relations found for asymptotic signals, making $\overleftarrow{s}(t)$ asymptotic as well (Novak, 2009, p. 55). Then, $\overleftarrow{s}(t) = \overleftarrow{A}(t) e^{j\overleftarrow{\varphi}(t)} \forall t \in \overleftarrow{\Omega}$ and relevant equations from the previous sections can be reformulated substituting $\overleftarrow{\bullet}$ when appropriate. From equation (3.12), equation (3.17), equation (3.15) and equation (3.18):

$$\overleftarrow{\tau}(f) = \overleftarrow{t}_f = -\frac{1}{2\pi}\frac{d\overleftarrow{\Psi}(f)}{df} = \frac{1}{2\pi}\frac{d\Psi(f)}{df} = -\tau(f) = -t = -t_f \tag{3.25}$$

meaning that $s(t)$ and $\overleftarrow{s}(t)$ are related by temporal inversion. Hence:

$$\overleftarrow{\varphi}(t) = \varphi(-t) \implies \ddot{\overleftarrow{\varphi}}(t) = \ddot{\varphi}(-t) \tag{3.26}$$

From equation (3.13):

$$\overleftarrow{B}(f) = \frac{1}{B(f)} = \frac{1}{\alpha\sqrt{\frac{\gamma}{2\pi f}}} = \frac{1}{\alpha}\sqrt{\frac{2\pi f}{\gamma}}$$

But, being $\overleftarrow{s}(t)$ asymptotic, it must also hold, from equation (3.10), equation (3.4), equation (3.26) and Remark 3.3:

$$\overleftarrow{B}(f) = \frac{\overleftarrow{A}\left(\overleftarrow{t}_f\right)}{\sqrt{\left|\ddot{\overleftarrow{\varphi}}\left(\overleftarrow{t}_f\right)\right|}} = \frac{\overleftarrow{A}\left(\overleftarrow{t}_f\right)}{\sqrt{\left|\ddot{\overleftarrow{\varphi}}\left(-t_f\right)\right|}} = \frac{\overleftarrow{A}\left(\overleftarrow{t}_f\right)}{\sqrt{\left|\ddot{\varphi}\left(-\left(-t_f\right)\right)\right|}} = \frac{\overleftarrow{A}\left(\overleftarrow{t}_f\right)}{\sqrt{\left|\ddot{\varphi}\left(t_f\right)\right|}} = \frac{\overleftarrow{A}\left(\overleftarrow{t}_f\right)}{\sqrt{\frac{2\pi f}{\gamma}}}$$

Hence, using equation (3.17), equation (3.18) and Remark 3.3:

$$\frac{1}{\alpha}\sqrt{\frac{2\pi f}{\gamma}} = \frac{\overleftarrow{A}\left(\overleftarrow{t}_f\right)}{\sqrt{\frac{2\pi f}{\gamma}}} \implies \overleftarrow{A}\left(\overleftarrow{t}_f\right) = \frac{2\pi f}{\alpha\gamma} = \frac{\cancel{2\pi}}{\alpha\gamma}\frac{1}{\cancel{2\pi}}\dot{\varphi}(\tau) = \frac{\dot{\varphi}(t)}{\alpha\gamma} = \frac{\beta}{\alpha\gamma}e^{t/\gamma}$$

Finally, from equation (3.25):

$$\overleftarrow{A}(t) = \frac{\beta}{\alpha\gamma}e^{-t/\gamma}$$

So, $\forall t \in \overleftarrow{\Omega}$, remembering again Remark 3.3:

$$\overleftarrow{s}(t) = \overleftarrow{A}(t) e^{j\overleftarrow{\varphi}(t)} = \left[\frac{\beta}{\alpha\gamma}e^{-t/\gamma}\right] e^{j\varphi(-t)} = \left[\frac{\beta}{\alpha^2\gamma}e^{-t/\gamma}\right] s(-t)$$

$$\implies \overleftarrow{z}(t) = \left[\frac{\beta}{\alpha^2\gamma}e^{-t/\gamma}\right] \sin(\varphi(-t))$$

meaning that the phase cycles in an opposite way with respect $s(t)$. Thus, operationally, the construction of $\overleftarrow{z}(t)$ is simple: apply $\frac{\beta}{\alpha^2\gamma}e^{t/\gamma}$ ([8]) to $z(t)$ and play the result backwards. In fact, this is the procedure suggested by Farina (2000).

3.2.6 Final stimulation signal expression

Chosen the starting frequency f_{start}, the coarse time length $\widetilde{T}$ and and the arrival harmonic order $M \in \mathbb{H}$ ($f_{end} = M f_{start}$) the SES and its inverse filter are given by:

$$\boxed{\begin{cases} z(t) = \sin\left(\beta\gamma\left[e^{t/\gamma} - 1\right]\right) & \\ \overleftarrow{z}(t) = \left[\frac{\beta}{\gamma}e^{\ln(M)-t/\gamma}\right]\sin\left(\beta\gamma\left[e^{\ln(M)-t/\gamma} - 1\right]\right) & \forall t \in [0,\ \gamma\ln(M)] \end{cases}} \qquad \beta = 2\pi f_{start}, \quad \gamma = \frac{1}{f_{start}}\text{round}\left(f_{start}\frac{\widetilde{T}}{\ln\left(\frac{f_{end}}{f_{start}}\right)}\right) \tag{3.27}$$

All the defining constants are then positive and, to be a proper SES the following, after Definition 3.5, must hold:

$$\boxed{\beta \gg \frac{1}{\gamma}} \tag{3.28}$$

to be noted that, in equation (3.27), $\overleftarrow{z}(t)$ has been adjusted to have $\overleftarrow{\Omega}$ (not yet defined) equal to Ω for convenience.

3.2.7 Summary

Conditions for Exponential Chirps $s(t)$ to be asymptotic have been given, so that time and frequency domain representations are linked by simple algebraic manipulation (Section §3.2.2). Conditions for an Exponential Swept Sine $z(t)$ to have $s(t)$ as analytic signal have been found (Section §3.2.3) enabling simplicity in the calculation of the inverse filter (Section §3.2.5). The synchronization has been imposed to $z(t)$ in order to make it profitable for nonlinear systems study (Section §3.2.4 and Section §3.2.6).

3.3 Measurement technique

To operate the final choice of f_{start}, M and $\widetilde{T}$ for equation (3.27) the measuring chain has to be defined first. This is done in figure 3.2. Each element is reviewed.

Computer The Computer[9] is set up with an audio optimized ArchBang Linux installation whose complete configuration has been published by the author (CrocoDuck, 2013).

[8] This quantity is different in the work of Novak et al. (2010b) by a factor 2π. This is, however, unimportant as it affects the magnitude only of the final results, that will be normalized anyway.

[9] A Compaq Presario CQ61.

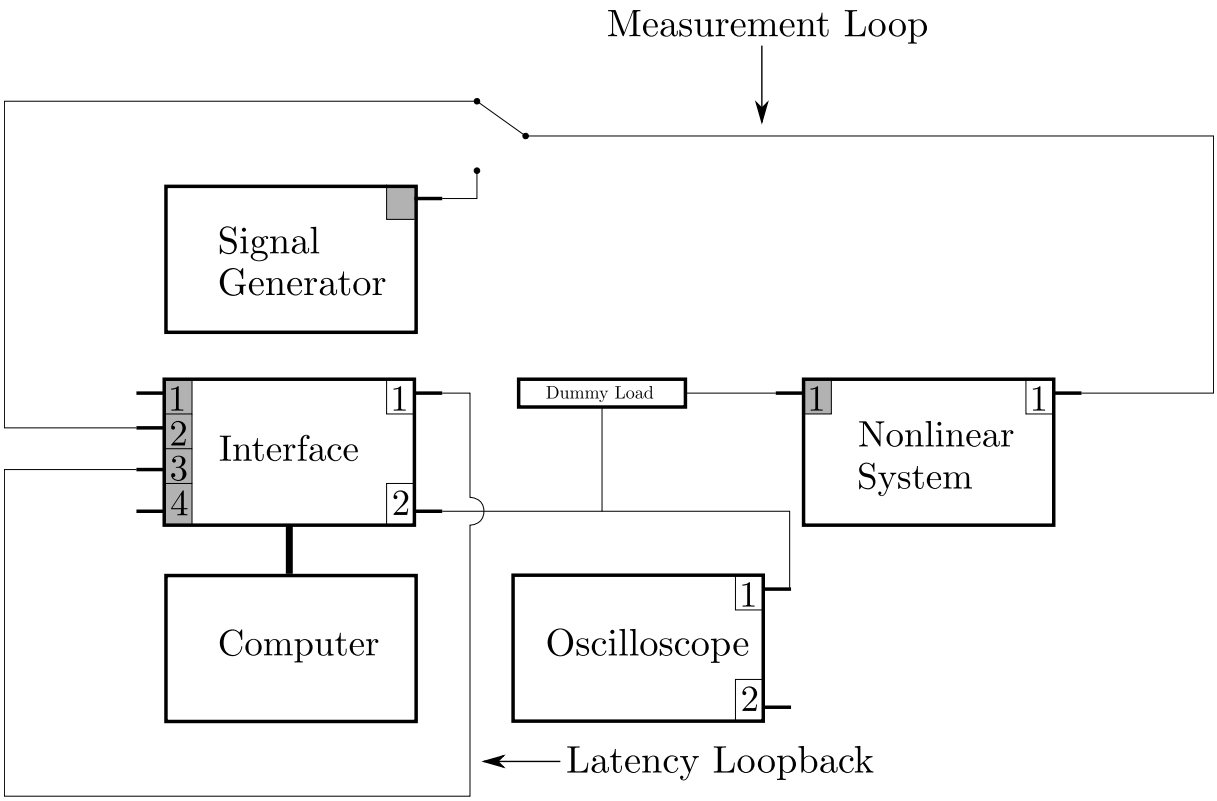


Figure 3.2: Measurement chain scheme. Outputs of systems are shaded in gray. The scheme does not report the actual physical placement of the ports.

This choice empowers control over the system, makes easy to automatize the measurement process through the use of Bash scripts and enhances reproducibility. This block is not supposed to alter the signal significantly, a part for the introduction of latency. The software packages used are quickly reviewed in table 3.3 and their features further discussed when relevant.

Interface The Interface is a Focusrite Scarlett 2i4. Its maximum sample rate is 96 kHz and it is thus selected:

$$f_s = 96\,\text{kHz}$$

Focusrite (2012) reports the frequency response magnitude to be flat over the audible range (± 0.1 dB over [20 Hz, 22 kHz]). The distortion figures appear low overall, depicting a weakly nonlinear system. The cross-talk, reported to be < 80 dB, was not actually low enough and leakage of the signals between channels is unacceptably high when using outputs that are physically close. That is why the different outputs are wired as in figure 3.2: only the ports at maximal distance between one another are used. The interface response will not be removed from the measurement since it is reasonably linear (compared to the *NTI* systems under analysis). Also, the goal of the project is to investigate the potentialities of *SES* measurements on nonlinear transmission lines identification in general, making the choice of the nonlinear system irrelevant. More comments on that are reported in Appendix B.3.

Name	Description	Comment
`jack`	Low-latency audio daemon: it supplies a low-latency software framework for audio applications	Required for low latency stable operation under Linux
`jack_iodelay`	Utility of the `jack` daemon assessing delays by computing the phase shift of a multi-tone signal after its transmission along a line (see figure B.1)	Needed for round-trip latency measurement
`Patchage`	Connection manager for `jack`	Used to connect software and hardware clients together
`Ardour`	Multi-track digital audio workstation supplying very accurate editing tools	Used to input the test signal and record the output
`Baudline®`	Visual analyzer for audio signals	Supplies time/frequency domain real-time analysis and monitoring tools
`MATLAB®`	Numerical computing environment	Data analysis
`R`	Numerical computing environment	Statistical analysis

Table 3.3: Software reference.

Nonlinear System Any audio amplifier constituting a NTI system under test. They have been: Quad II , Rotel RA-930AX and Sugden Musicmaster Pure Class A.

Dummy load As the output of the amplifiers is intended to drive loudspeakers an appropriate load must be supplied with the correct nominal impedance. A potentiometer across the load makes possible to set an appropriate level for the interface. The scheme is reported in figure 3.3.

Oscilloscope A Tektronix TBS 1052B-EDU used to visualize the signal and measure their amplitude.

Signal generator A Krohn-Hite 4400A used to generate low distortion reference tones.

Having detailed the blocks it is now possible to describe how the measurement is performed.

Input Signals

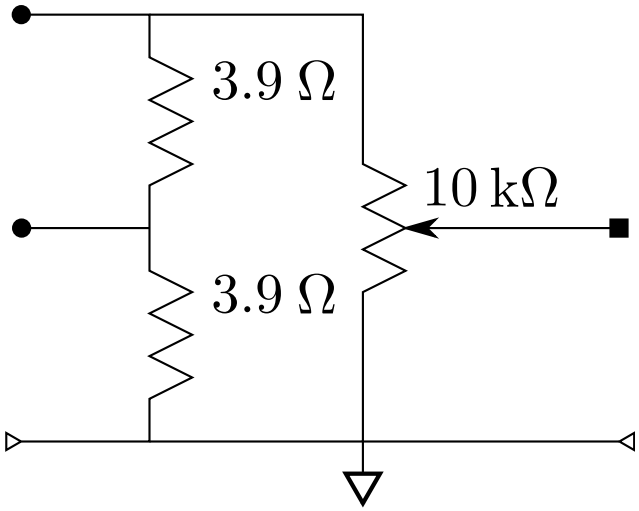


Figure 3.3: Dummy load schematic. ● represents possible inputs, ■ the output and ◁ the ground connections.

SES **stimulation** The final frequency for the *SES* is taken as the Nyquist frequency:

$$f_{end} = \frac{f_s}{2} = 48\,\text{kHz}$$

For f_{start}, equation (3.28) warns against the use of a very low value, as $\beta \propto f_{start}$. Also, equation (3.27) suggests to take a large $\widetilde{T}$ to enlarge γ. Another advantage from a large $\widetilde{T}$ comes from the fact that each of the $\triangle t_m \propto \gamma$ will be made larger as well, minimizing the overlap of the higher order responses. As a consequence, $f_{start} = 12\,\text{Hz}$ ($\implies M = 4 \cdot 10^3$) and $\widetilde{T} = 40\,\text{s}$. To be noted that $f_{start} = 0\,\text{Hz}$ would imply $\beta = 0$, erasing the signal, as evident after equation (3.27). So, a low not null f_{start} must be used. For the measured responses this will be analogous to a drop in magnitude towards low frequencies as information in that region is absent. This is a limitation and it will discussed further in section §4. All the *SES* properties are summed up in table 3.4, showing that the fundamental hypotheses of the previous section ($\beta \gg 1/\gamma$ and $\mathcal{BT} \gg 1$) are respected. The theory can then be applied. To complete the stimulation $x(t)$, $z(t)$ is preceded by

Parameter	Value
f_{start}	$12\,\text{Hz}$
f_{end}	$48 \cdot 10^3\,\text{Hz}$
β	$\sim 75.40\,\text{Hz}$
γ	$\sim 4.84\,\text{s}$
$\epsilon\ (= 1/\gamma)$	$\sim 0.21\,\text{Hz}$
$\mathcal{BT}$	272591.84
T	$\sim 40.09\,\text{s}$

Table 3.4: $z(t)$ properties (*SES*).

an 1 s initial pad and its end is padded until 50 s are reached:

$$x(t) = \begin{cases} 0 & \text{if } t \leq 1\,\text{s} \\ z(t) & \text{if } 1\,\text{s} < t \leq 1\,\text{s} + T \\ 0 & \text{if } 1\,\text{s} + T < t \leq 50\,\text{s} \end{cases}$$

All these operations are performed in the discrete time domain by `Input_Setup.m` and `SES_CHIRP.m`. The discrete time signals are notated substituting $t \in \mathbb{R}$ with a generic discrete index $t_d \in \mathbb{N}$, as only the instants $t = t_d T_s = {}^{t_d}/_{f_s}$ survive the sampling process, making the sampled quantities functions of t_d:

$$x(t_d) = \begin{cases} 0 & \text{if } t_d \leq a \\ z(t_d) & \text{if } a + 1 \leq t_d \leq b\,, \\ 0 & \text{if } b + 1 \leq t_d \leq c \end{cases} \qquad \begin{pmatrix} a = 96 \cdot 10^3 \\ b = 3944439 \\ c = 48 \cdot 10^5 \\ t_d = 1, 2, \ldots, c \end{pmatrix} \tag{3.29}$$

In the following the maximum number of samples for a time series $x(t_d)$ will be notated as T_d^x. With complete analogy, sampled quantities $X(f_d)$ in the frequency domain will be notated with F_d^X discrete generic indices f_d.

Additional inputs To verify the accuracy of the identified models two additional inputs are presented to the system, a 2400 Hz sine wave and a filtered MLS sequence designed after Kemp (2002) and Vanderkooy (1994). The sine frequency and pass band of the MLS are chosen in agreement with the pass band of the models (see Section §4.3). They are padded similarly to $x(t_d)$.

Level calibration A first sweep is used to visualize the pass-band of the NTI system by looking at the amplitude on the oscilloscope. Then, a value of frequency in the pass-band is selected to build a reference tone. This is generated by the signal generator and used to set the input amplitude for the NTI to Full Power, Half Power or a Quarter Power respectively. In details, the amplitude is varied until the NTI output starts showing the first noticeable clipping on the oscilloscope. The dummy load is then set to assign an amplitude of 12 V to the oscilloscope reading. The full power input level is then chosen as the one producing a reading of 10 V. This to make sure to study the weak non-linearity associated with the normal operational conditions of the units, avoiding any clipping. The remaining input amplitudes are calculated by iterated divisions by $\sqrt{2}$. The Interface input controls are used to achieve a recording level (measured by `Baudline`®) of -13 dB to minimize the interface non-linearity. Then, the same reference tone is generated digitally by the Computer and the Interface output level is adjusted until the output level at the dummy load is restored, completing the calibration for the

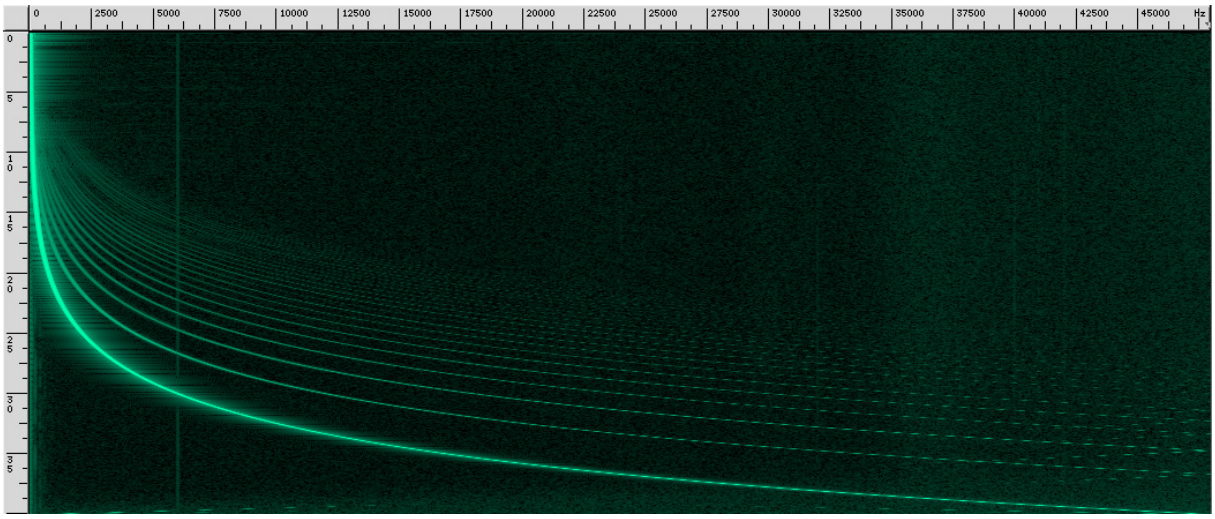


Figure 3.4: Example of NTI output (Quad II, Full Power). Spectrogram as supplied by `Baudline`®. Harmonics and tonal noise from cross-talk are evident. Horizontal axis: frequency (Hz). Vertical axis: time (s).

SES and Sine input. For the filtered MLS case the Interface output gain is harder to calibrate as the signal amplitude is unsteady and the power level is not trivial. Also, NTI systems are susceptible to instantaneous variations more than to the overall power level of a signal. Hence, the level has been calibrated to yield instantaneous values of amplitude in the same range of the calibration tone, making sure that the 10 V threshold is never exceeded.

Measurement procedure The discrete sequence $x(t_d)$ is loaded into `Ardour`. The basic operation consists of routing the `Ardour` output towards the output 2 of the interface and playback the signal while recording from the input 2. The defined nonlinear deconvolution operations apply in the discrete time domain as well and they can be performed numerically with `MATLAB`® using the recorded discrete outputs $\{y(t_d)\}_{i=1,\dots,N_y}$, with N_y the number of measurements, and the discrete inverse filter $\overleftarrow{z}(t_d)$. All the initial pads in the $\{y(t_d)\}_{i=1,\dots,N_y}$ are omitted by windowing, while the tail pad is preserved to avoid circular convolution artifacts. As such $\overleftarrow{z}(t_d)$ needs to be padded in the beginning to avoid delay artifacts in the impulse response.

Any delay not introduced by the NTI system must be removed. The Latency Loopback attempted to achieve this by measuring the sample latency Δ with `jack_iodelay`. However, this has a collateral effect: the signal generated by `jack_iodelay`, looping in the Latency Loopback, leaks in the measurement loop, as evident in the 6 kHz trail in figure 3.4, associated with the highest peak of the `jack_iodelay` multi-tone signal (see figure B.1). Moreover, Focusrite (2012) suggests that the different outputs are served by different converters units, as they are associated with different technical specifications. So, the latency associated with the two loops is different. This problem is solved by the `MATLAB`® script `Lag_Detector.m` that, using numerical derivatives and noise statistics

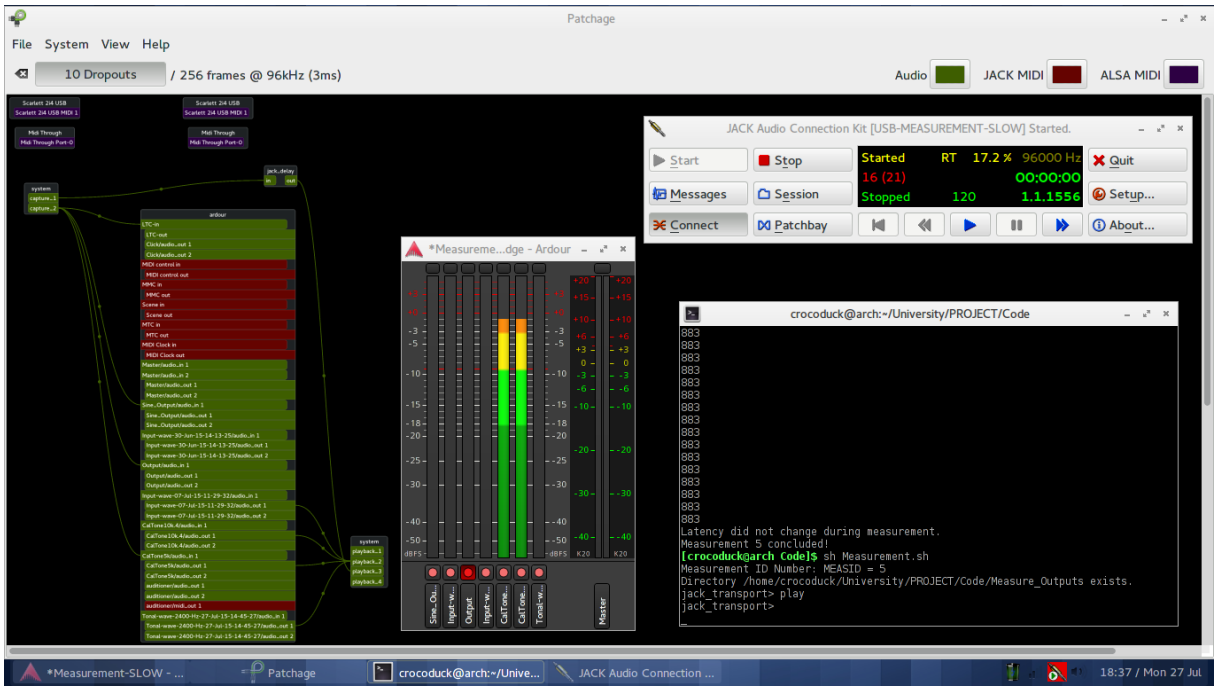


Figure 3.5: Measurement chain at the software level (`Patchage` screen-shot). Statistics on overruns are reported as Dropouts. When they happen during a measurement the measurement is discarded. It is possible to see the routing of the different programs and the Bash script `Measurement.sh` in action, starting the software and logging the latency values in samples.

on *NTI* outputs to locate the actual outputs start, pointed out a 14 samples latency mismatch between the two loops (see Appendix B.4). Since `Lag_Detector.m` performs well only at very high signal to noise ratio the latency is measured and corrected. To minimize cross-talk, the level of the signal circulating the Latency Loopback is set as low as possible.

The Latency Loopback is also used for stability monitoring. In fact, the Computer is not ideally stable and buffer overruns can be produced. Only the measurements performed under stable operation must be kept. Sources of information for that are `jack`, that provides statistics on the software buffer overflow, and `jack_iodelay`, that points out changes in latency that correlates with overflows. The measurement is discarded whenever one of the sources of information suggests a dropout happened.

An ensemble of $N_y = 16$ outputs[10] is collected by feeding the *NTI* systems multiple times with the digitally stored $x(t_d)$. After the various different latency have been removed from the digitally recored $\{y(t_d)\}_{i=1,\ldots,N_y}$ average can take place to reduce the background noise. The resulting signal $\overline{y}(t_d)$ is used for the nonlinear deconvolution, operated in the frequency domain for computational convenience after assigning windows to $\overline{y}(t_d)$ and the inverse filter to reduce smearing while minimizing artifacts in the responses.

[10] At each doubling of the ensemble size it is associated a 3 dB improvement of the Signal to Noise ratio. 16 corresponds to 4 times doubling of the ensemble size.

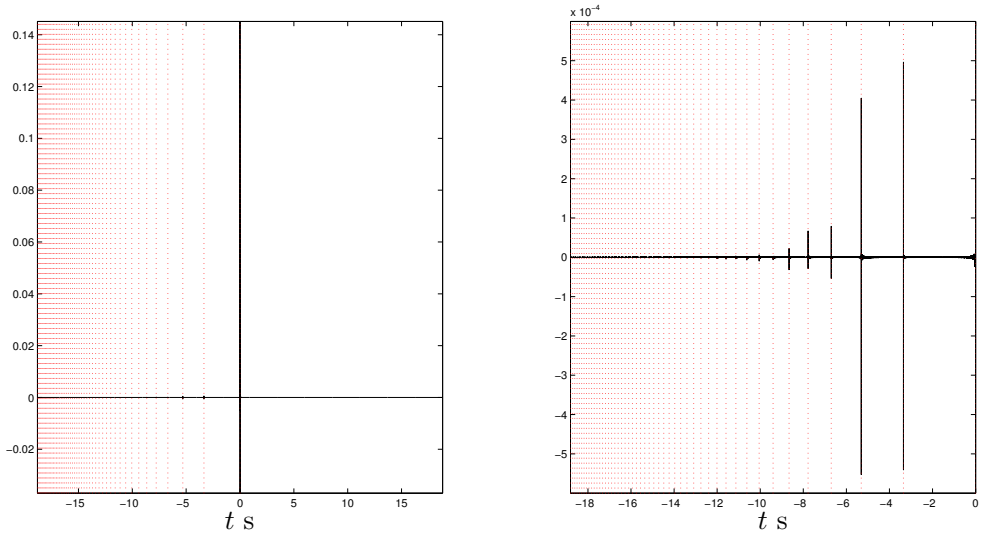


Figure 3.6: Typical result from measurement plotted over $\pm\triangle t_{49}$. Obtained from the Quad II measurement at full power. Arbitrary ordinate units as obtained from signals normalized to unit peak amplitude, both in value and unit. On the left: improper impulse respose. On the right: particular of the negative time zone. The red dotted lines are drawn at the $\triangle t_m$ lags.

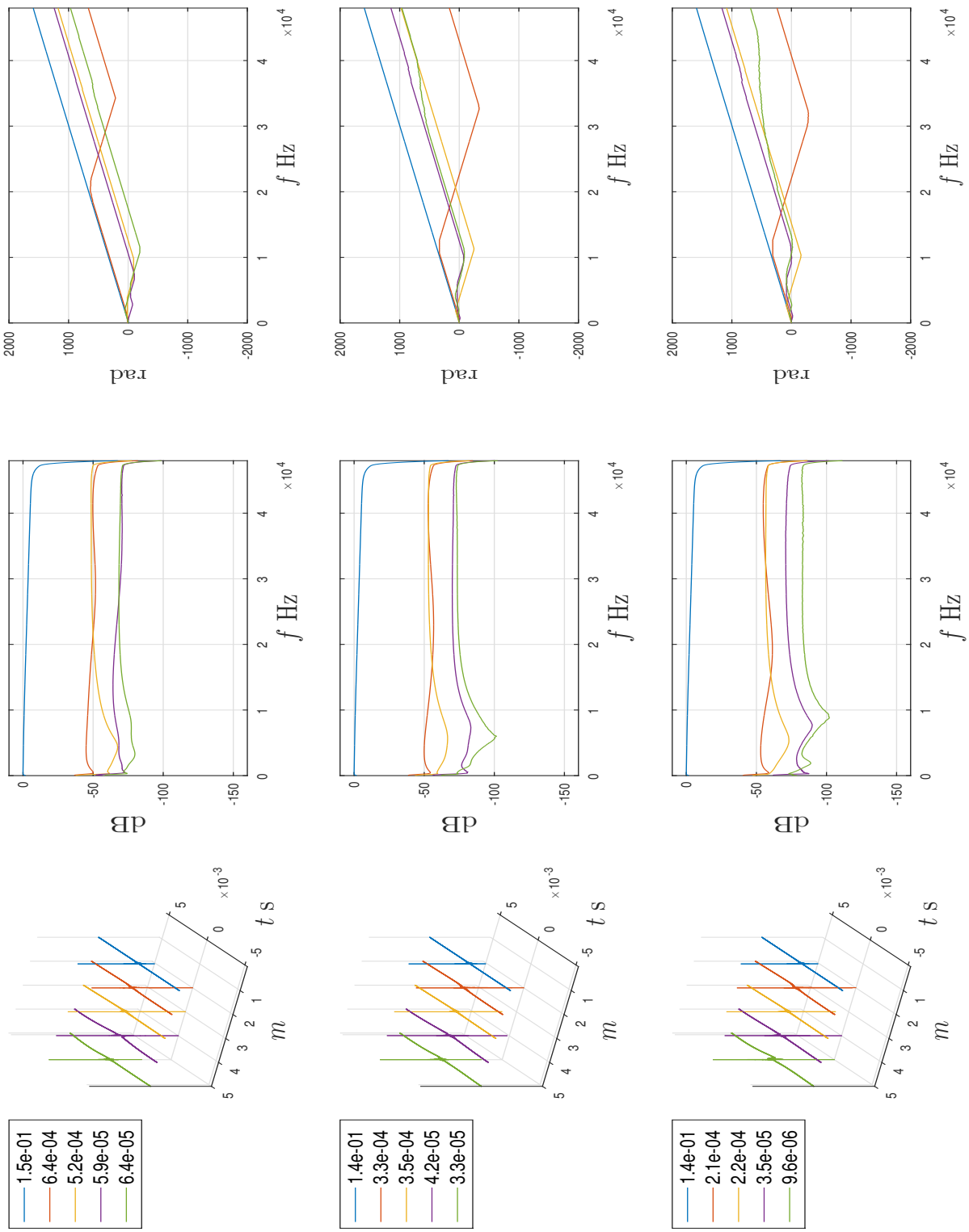
Summary A measurement setup has been studied to reduce the cross-talk of the Interface while being able to monitor the Computer performances in order to select only the best measurements (i.e., the ones free of buffer overruns). Practically, after the calibration took place, the Bash script `Measurement.sh` starts the measurement by launching `jack_iodelay`, connecting it to the required hardware ports and logging its latency outputs. Also, the script starts the playback/recording in `Ardour`. After the measurement finished the logs are inspected and the `Ardour` session checked. If the latency was constant and `Ardour` does not display buffer overruns markers dispatched by `jack` the measurement is judged successful and exported as a wav file, double format. Completed all the measurements `Measure_Analyzer.m` will remove the latency from each $\{y(t_d)\}_{i=1,\dots,N_y}$ based on the logged latency values and the 14 samples correction, calculate the average $\overline{y}(t_d)$, normalize the amplitude and operate the nonlinear convolution with a properly padded $\overleftarrow{z}(t_d)$. A representation of the measurement chain at the software level is given in figure 3.5.

3.4 Measure outcome

A typical $\widetilde{h}(t_d)$ improper impulse response result is shown in figure 3.6 (see equation (3.24)). It is possible to appreciate the higher order impulse responses forming at the expected $\triangle t_m$ negative lags. It is evident that they can be gated. Gating is performed by `Response_Study.m`[11] and the results showed in Figures (3.7a), (3.7b) and (3.7c) for each

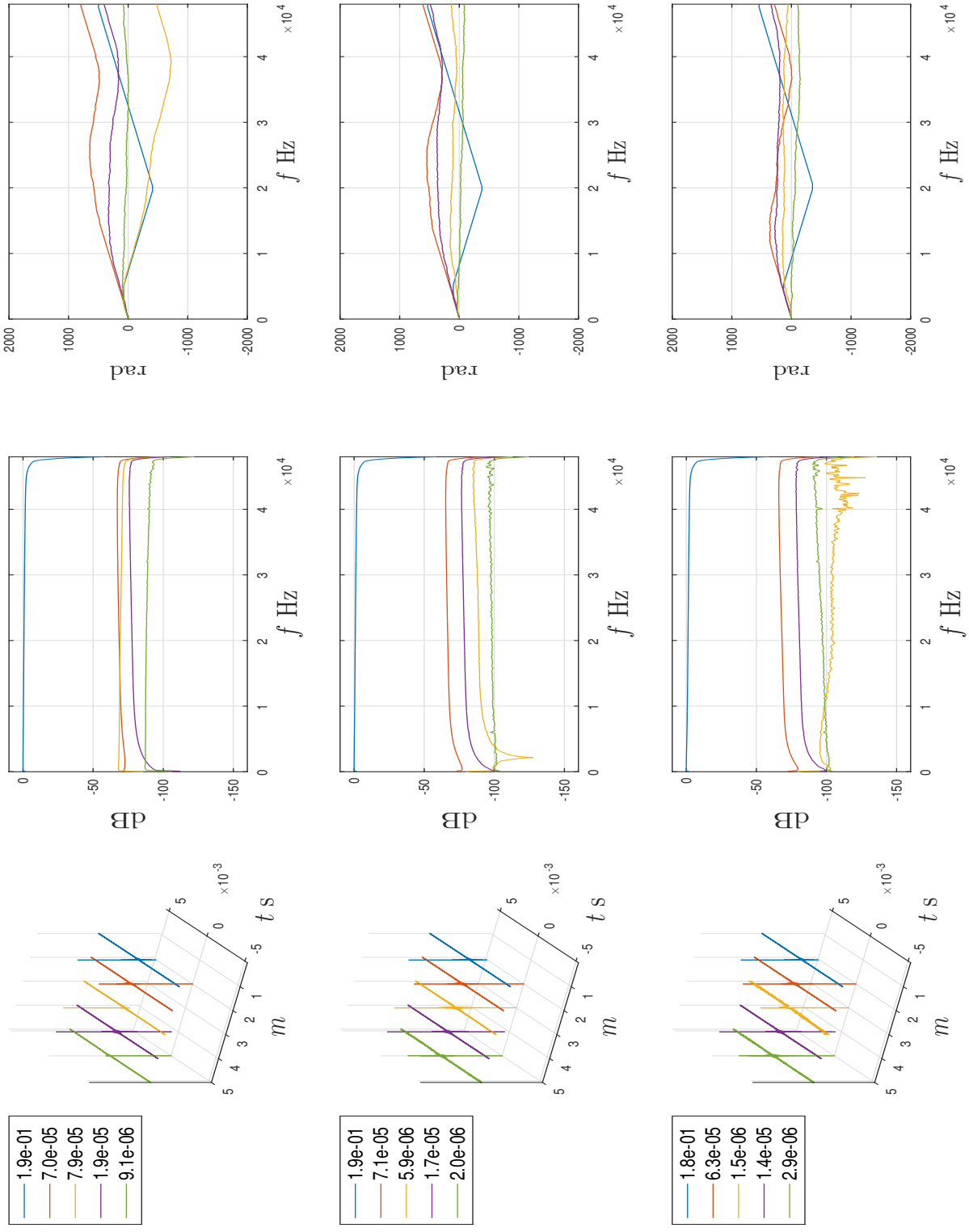
[11] Written under the guidance of Novak (2014).

unit and power level of the input signal as obtained using windows 1024 samples long. Although not a complete information yet, Figures (3.7a), (3.7b) and (3.7c) depict already valuable information on the NTI systems under analysis. In fact, the clear evolution of the responses with the input power points out that the non-linearity strongly depends on the input amplitude for all the units. Also, clues about amount of distortion introduced in an output signal are suggested by how quickly the higher order responses drop in magnitude before hitting the noise floor. The Quad II unit shows the higher non-linearity, with all the frequency response levels contained in $[-100\,\mathrm{dB},\, 0\,\mathrm{dB}]$. The Sugden unit appears as the most linear, with less important higher order responses, the fifth starting mixing with noise already at the half power input. The Rotel unit is not too dissimilar.

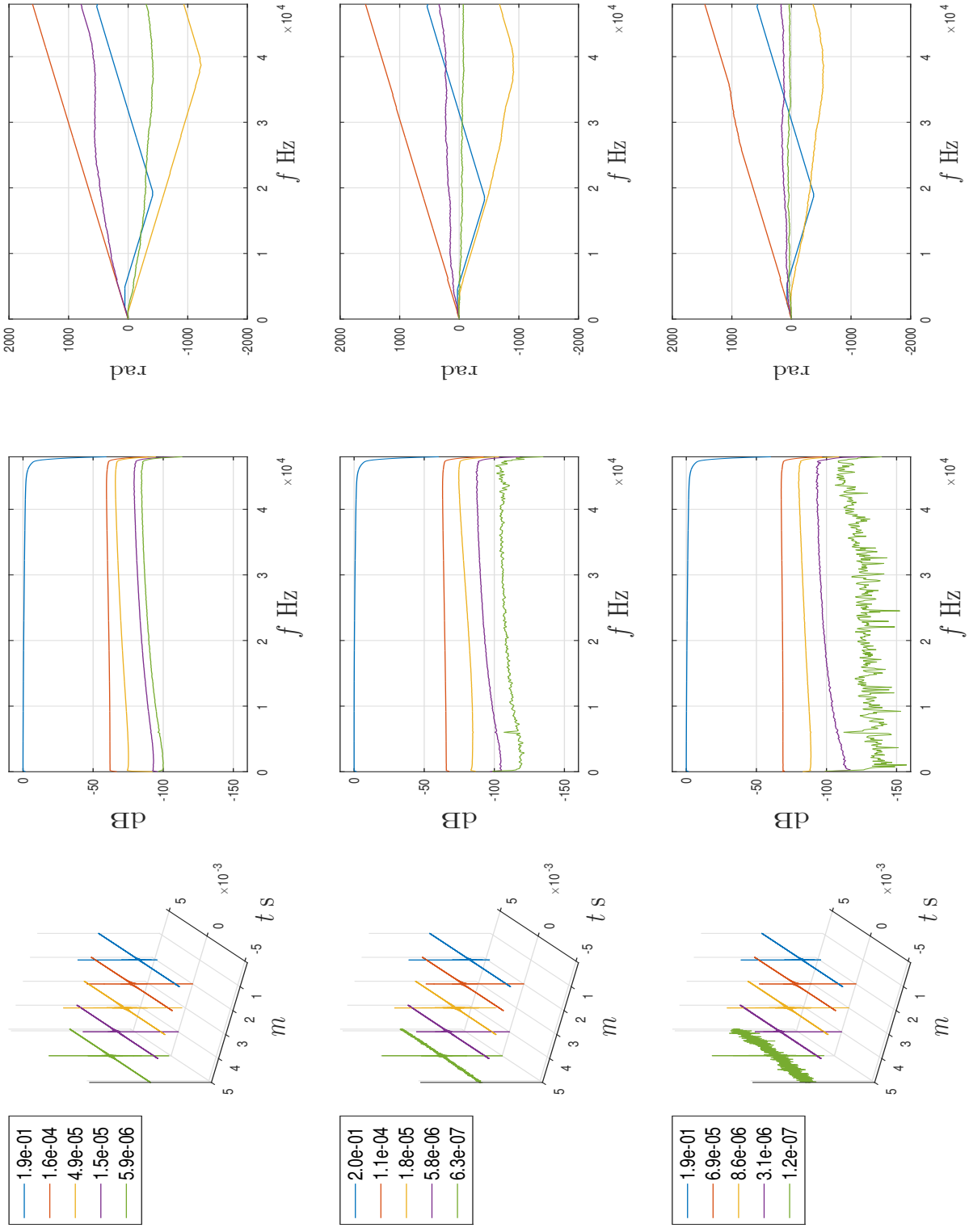


(a) Quad II.

Figure 3.7: Higher order responses. (Rotating the page ↻) On the left: normalized higher order impulse responses. The values in legend indicate the absolute peak values, arbitrary units after processing with normalized signals (both in value and unit). In the middle: higher order frequency responses levels normalized to absolute peak linear ($m = 1$) response. On the right: associated phase responses. Plots arranged vertically from the top in order of decreasing input signal power (Full Power - Half Power - Quarter Power). Only the first 5 orders are showed for clarity.


rad
dB
f Hz
t s
m
1.9e-01
7.0e-05
7.9e-05
1.9e-05
9.1e-06
1.9e-01
7.1e-05
5.9e-06
1.7e-05
2.0e-06
1.8e-01
6.3e-05
1.5e-06
1.4e-05
2.9e-06

(b) Rotel.


rad
f Hz
dB
f Hz
t s
m
1.9e-01
1.6e-04
4.9e-05
1.5e-05
5.9e-06
2.0e-01
1.1e-04
1.8e-05
5.8e-06
6.3e-07
1.9e-01
6.9e-05
8.6e-06
3.1e-06
1.2e-07

(c) Sugden.

4 Nonlinear System identification

4.1 Overview

Knowing the various higher order impulse responses is not enough to determine the Input-Output relation of a NTI system. For this purpose a suitable model must be used instead. This section, largely based on Novak (2009), Novak et al. (2010b) and Novak et al. (2010a), will deal with the NTI identification problem using a properly chosen nonlinear model. Since this is a huge topic, to make the following clearer, the material will be presented restricting to a particular set of models that are most useful given the stimulation signal chosen.

4.2 Nonlinear models

4.2.1 Volterra series

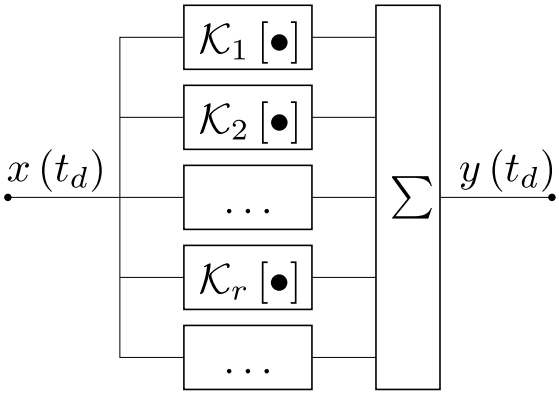


Figure 4.1: Volterra series model for nonlinear systems.

The most general way to describe a nonlinear system is through the Volterra series (Farina et al., 2001):

$$y(t_d) = \sum_{r=1}^{\infty} \underbrace{\left\{ \sum_{q_1=1}^{T_d^k} \sum_{q_2=1}^{T_d^k} \dots \sum_{q_r=1}^{T_d^k} k_r(q_1, q_2, \dots, q_r) \prod_{g=1}^{r} x(t_d - q_g) \right\}}_{\doteq \mathcal{K}_r[x(t_d)]}$$

written in the discrete time domain, where T_d^k is the number of time samples used to define each of the Volterra kernels $k_r(q_1, q_2, \dots, q_r)$ along each dimension, eventually by zero padding of shorter sequences. $\mathcal{K}_r : \mathscr{L}^2\left(\mathbb{R}^{T_d^x}, \mathbb{R}\right) \longrightarrow \mathbb{R}$ represents the r-th order Volterra operator. The model (shown in figure 4.1) consists of a series of convolutions, performed by the operators, of the input $x(t_d)$ with higher order kernels that can be derived from the higher order impulse responses $h_m(t_d)$ ([12]). The model allows to take

[12] For the reason that the series generalizes the convolution formalism of linear systems the kernels are

into account the effects of memory, since the output $y(t_d)$ at the sample t_d is given in terms of the input $x(t_d)$ at the previous samples $t_d - q_g$. This would be a strength of the model, since most systems have memory. However, as showed by Farina et al. (2001), the calculation of the $k_r(q_1, q_2, \ldots, q_r)$ is quite complicated even in the case of memory-less systems (for which the kernels are mono-dimensional). Also, the number of parameters grows with the kernel order, making kernels manipulation heavy from the computational standpoint. For this reason, as proposed by Novak (2009, p. 10), block models are preferred.

4.2.2 Hammerstein - Wiener models

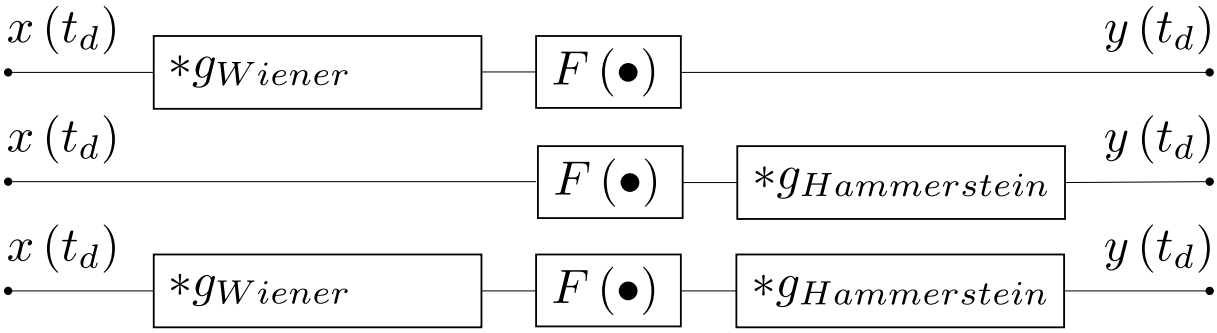


Figure 4.2: Block models. From top to bottom: Wiener model, Hammerstein model, Hammerstein - Wiener model. To be noted that, for all the models to have the same Input-Output relation, all the linear filters will have to be different if the static non-linearity is kept constant (Novak, 2009, pp. 10 - 12).

A way to simplify a NTI system with memory is to suppose it as made of different blocks, some linear and some nonlinear, where the nonlinear blocks are memory-less for simplicity. A memory-less non-linear block can be described with the mapping function formalism Dobrucki (2011):

$$F : \mathbb{R} \longrightarrow \mathbb{R} \quad : \forall t_d = 1, 2, \ldots, T_d^x \quad y_F(t_d) = F[x(t_d)]$$

where $F(x)$ is the mapping function and $y_F(t_d)$, with $T_d^{y_F} = T_d^x$, denotes the output of the nonlinear block described by $F(x)$. On the other hand a linear block is a linear filter described resorting to its impulse response $g(t_d)$ as usual. With reference to the simplest models of this kind, i.e. Single Input - Single Output ($SISO$) models composed by two blocks in series, the Hammerstein model is defined as the model having the nonlinear block at the input, while the Wiener model has it at the output. A Hammerstein - Wiener model is a fusion of the two, as represented in figure 4.2.

As said earlier, harmonic and intermodulation distortion are fingerprints of nonlinear systems. It is clear then that a successful model has to be able to recover harmonics and

sometimes referred as higher order impulse responses. In this report only the $h_m(t_d)$ are referred to with that name.

intermodulation products with the proper amplitudes and phases. Consider a nonlinear block described by

$$F_p(x) \doteq x^p, \quad F_p : \mathbb{R} \longrightarrow \mathbb{R} \, \forall p \in \mathbb{N} \tag{4.1}$$

For a discrete time[13] sinusoidal input having angular frequency ω, resorting to complex algebra, it can be showed that (Novak et al., 2010b):

$$\begin{aligned} & x(t_d) = \sin(\omega t_d) \quad \forall t_d = 1, 2, \ldots, T_d^x \\ & \qquad \Longrightarrow \forall t_d = 1, 2, \ldots, T_d^x = T_d^{y_{F_p}}; \quad y_{F_p}(t_d) = [x(t_d)]^p = \\ & = \begin{cases} \frac{(-1)^l}{4^l} \sum_{k=0}^{l} (-1)^k \binom{2l+1}{k} \sin[(2l+1-2k)\omega t_d] & \text{if } \exists l \in \mathbb{N} \cup \{0\} \, : \, p = 2l+1 \\ \frac{(-1)^l}{2^{2l-1}} \sum_{k=0}^{l-1} (-1)^k \binom{2l}{k} \cos[2(l-k)\omega t_d] + \frac{1}{2^{2l}} \binom{2l}{l} & \text{if } \exists l \in \mathbb{N} \, : \, p = 2l \end{cases} \end{aligned} \tag{4.2}$$

showing clearly that the application of a natural power introduces groups of harmonics of the input signal whose parity of the order depends on the parity of p. Also, feeding as an input a superposition of sine waves of angular frequencies ω_b and phases ϕ_b:

$$\begin{aligned} & x(t_d) = \sum_{b=1}^{B} \sin(\omega_b t_d + \phi_b) \quad \forall t_d = 1, 2, \ldots, T_d^x \\ & \qquad \Longrightarrow \; y_{F_p}(t_d) = [x(t_d)]^p = \prod_{1}^{p} \left\{ \sum_{b=1}^{B} \sin(\omega_b t_d + \phi_b) \right\} \end{aligned}$$

The last term, through the generation of cross products of sines of different arguments, is clearly the source of intermodulation between the tones present in the input signal. The simple static non-linearity $F_p(x)$ is then a valuable building block for the modeling of NTI systems, but only if the amplitude and phases of all the introduced tones are altered to reflect the system output tones. This consideration brings to the choice of a Hammerstein model: the input non-linearity takes care of generating the distortion products while the output filter corrects their properties to match the system Input-Output law and memory. It is also clear, however, that a single branch is not sufficient, as all the harmonics need to be recovered, not only the ones of the same parity of p. Also, $h_m(t_d)$ represents an input to single harmonic relation and it cannot be applied naively to the linear blocks as they will be processing multiple harmonics at a time.

4.2.3 $MISO$ Hammerstein model

For these reasons a Multiple Input - Single Output ($MISO$) Hammerstein model is proposed. Each of the N branches is constituted by a Hammerstein model and the final output is given by the sum of the branches outputs. If the physics of the system to be identified is known the single nonlinear functions $F_n(x)$ can be chosen after it, maybe

[13] The same applies in continuous time.

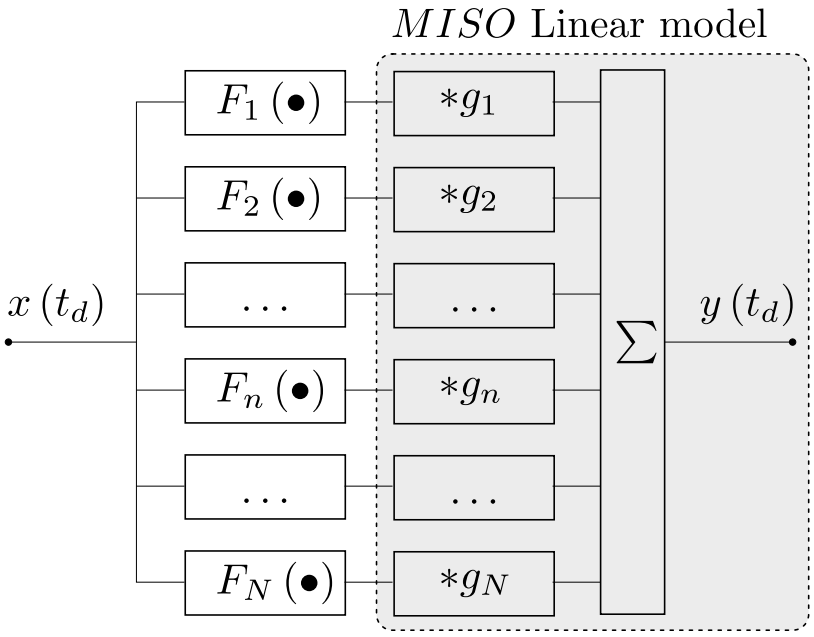


Figure 4.3: $MISO$ Hammerstein model. The figure clarifies the nomenclature given to the model, which globally is a $SISO$.

among the ones suggested by Dobrucki (2011). In this case the identification problem would be a gray-box problem and, as showed by Novak (2009, pp. 67 - 75), it would be particularly profitable in the case of hysteretic systems. This isn't the case as the physics of the amplifiers has not been studied and hence they are black-box systems. Also, the hysteresis cycle of the units is supposed to be narrow enough to make the hysteretic behavior negligible, as all the systems appear to revert back to the initial state soon after any stimulation ended. For this reason the $MISO$ Hammerstein model is transformed in a generalized Polynomial Hammerstein model (PHM) just by selecting the $F_n(x)$'s in analogy to the structure given by equation (4.1), i.e. from a polynomial of order N:

$$F_n(x) = x^n \quad \forall n = 1, 2, \ldots, N; \forall t_d = 1, 2, \ldots, T_d^x$$

This structure, since the simple relations between nonlinear blocks and harmonics (given by equation (4.2)), is well suited for blind identification, as pointed out by Novak (2009, p. 58). The responses $g(t_d)$ become then the unknowns of the identification procedure. They can be found from the system's $h_m(t_d)$.

Identification procedure Following Novak et al. (2010b) and Novak (2009, pp. 59 - 63), the identification is performed in the frequency domain. The partial frequency responses $H_m(f_d)$ are calculated:

$$H_m(f_d) \doteq \mathcal{F}[h_m(t_d)]$$

for the unknowns:

$$G_n\left(f_d\right) \doteq \mathcal{F}\left[g_n\left(t_d\right)\right]$$

The $H_m\left(f_d\right)$ link an input frequency to the relative harmonics in the sense that, for an input frequency relative to $\widetilde{f}_d$, the amplitude and phase of its m-th harmonic is given by $H_m\left(\widetilde{f}_d\right)$. Since all the impulse responses are real functions all the frequency responses will have Hermitian symmetry. As such, restriction to the positive discrete frequency only is operated in the following without loss of generality:

$$f_d \in \left\{1, 2, \ldots, \overset{+}{F}{}_d^H\right\} \subset \mathbb{N}, \quad \overset{+}{F}{}_d^H \doteq \begin{cases} \left\lfloor \frac{F_d^H}{2} \right\rfloor & \text{if } \left(F_d^H \mod 2\right) = 0 \\ \left\lfloor \frac{F_d^H}{2} \right\rfloor + 1 & \text{otherwise} \end{cases}$$

where the definitions are given according to the MATLAB® indexing for easier comparison with the code[14] and F_d^H is the total number of frequency samples used to describe each of the $H_m\left(f_d\right)$. A way to think about a generic input sequence $x\left(t_d\right)$ is as a superposition of pure sines with amplitudes and phases given by its Fourier transform. Selecting one of these sines means, apart from amplitude and phase, selecting a component of the kind of $\sin\left(\omega t_d\right)$. Consider then an input $\sin\left(\omega t_d\right)$, where now ω is relative to a chosen sampled frequency index $\overline{f_d}$. Then equation (4.2) can be written in the positive frequency domain as:

$$\begin{aligned} &\mathcal{F}^+\left\{\left[\sin\left(\omega t_d\right)\right]^n\right\} = \\ &= \begin{cases} \frac{(-1)^l}{4^l} \sum_{k=0}^{l} (-1)^k \binom{2l+1}{k} \mathcal{F}^+\left\{\sin\left[\left(2l+1-2k\right)\omega t_d\right]\right\} & \text{if } \exists l \in \mathbb{N} \cup \{0\} \ : \ n = 2l+1 \\ j \frac{(-1)^l}{2^{2l-1}} \sum_{k=0}^{l-1} (-1)^k \binom{2l}{k} \mathcal{F}^+\left\{\sin\left[2\left(l-k\right)\omega t_d\right]\right\} + \frac{1}{2^{2l}} \binom{2l}{l} & \text{if } \exists l \in \mathbb{N} \ : \ n = 2l \end{cases} \end{aligned} \tag{4.3}$$

where the Fourier transform linearity and the fact that $\mathcal{F}^+\left[\cos\left(\bullet\right)\right] = j \cdot \mathcal{F}^+\left[\sin\left(\bullet\right)\right]$ have been used, $\mathcal{F}^+\left[\bullet\right]$ denoting the discrete Fourier transform restricted to the positive discrete frequencies. This equation would apply to the n-th branch of the model. Hence,

[14] In MATLAB® frequency domain quantities have the positive frequency samples in the first half.

for all the branches the following matrix equation can be written:

$$\underbrace{\begin{pmatrix} \mathcal{F}^+\left\{[\sin(\omega t_d)]^1\right\} \\ \mathcal{F}^+\left\{[\sin(\omega t_d)]^2\right\} \\ \vdots \\ \mathcal{F}^+\left\{[\sin(\omega t_d)]^n\right\} \\ \vdots \\ \mathcal{F}^+\left\{[\sin(\omega t_d)]^N\right\} \end{pmatrix}}_{\doteq\vec{\mathcal{P}}} = \mathbf{A} \underbrace{\begin{pmatrix} \mathcal{F}^+[\sin(1\cdot\omega t_d)] \\ \mathcal{F}^+[\sin(2\cdot\omega t_d)] \\ \vdots \\ \mathcal{F}^+[\sin(m\cdot\omega t_d)] \\ \vdots \\ \mathcal{F}^+[\sin(M\cdot\omega t_d)] \end{pmatrix}}_{\doteq\vec{\mathcal{H}}} + \vec{\mathbf{B}} \tag{4.4}$$

with M the number of higher order impulse responses considered[15]. The vectors of powered sines and harmonics in the positive frequency domain have been labeled $\vec{\mathcal{P}}$ and $\vec{\mathcal{H}}$ respectively. The matrices $\mathbf{A}$ and $\vec{\mathbf{B}}$ are written after the coefficients in equation (4.3). While $\mathbf{A}$ is a rectangular matrix with the multiplicative coefficients of equation (4.3) $\vec{\mathbf{B}}$ is a column vector containing the DC terms spanned by the even powers (bottom of equation (4.3)). As such, $\vec{\mathbf{B}}$ will contribute only to the global DC value of the output. This DC value will be removed anyway since all the systems in the measure chain of figure 3.2 have DC blocking capacitors and the *SES* stimulation cannot reveal the very low frequency end (as explained in Section § 3.3 on page 24). So, only $\mathbf{A}$ needs to be determined:

$$\mathbf{A} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,m} & \cdots & a_{1,M} \\ a_{2,1} & a_{2,2} & \cdots & & & \\ \vdots & \vdots & \ddots & & & \\ a_{n,1} & & & a_{n,m} & & \\ \vdots & & & & \ddots & \\ a_{N,1} & & & & & a_{N,M} \end{pmatrix} \qquad n = 1, 2, \ldots, N;\ m = 1, 2, \ldots, M$$

$$a_{n,m} = \begin{cases} \frac{(-1)^{2n+\frac{1-m}{2}}}{2^{n-1}} \binom{n}{\frac{n-m}{2}} & \text{if } n \geq m \ \wedge\ (n+m \mod 2) = 0 \\ 0 & \text{otherwise} \end{cases} \in \mathbb{C}$$

It is simple, although tedious, to verify that the usual matrix multiplication $\mathbf{A}\vec{H}$ supplies the correct equations of the kind of equation (4.3), a part for the discarded DC values[16]. The $\mathbf{A}$ matrix is then used, in the frequency domain, to pass from harmonics to powers. The matrix multiplication operation implies, for each element $\mathcal{P}_n$ of $\vec{\mathcal{P}}$ and $\mathcal{H}_m$ of $\vec{\mathcal{H}}$, a

[15] Theoretically, the model can be formulated with infinite branches, using all the infinite higher order impulse responses. Of course, to practically implement the model, restriction to finite orders must be done (also, the *SES* would have to be infinite to measure all the infinite higher order responses).

[16] The existence of this linear relationship shows that the choice to discard amplitude and phase of the signal component $\sin(\omega t_d)$ makes sense, as the amplitude and phases will be modified just by a multiplication by a complex quantity whatever their initial values were.

structure of this kind:

$$\mathcal{P}_n = \sum_{m=1}^{M} a_{n,m}\mathcal{H}_m$$

i.e. a frequency domain superposition of harmonics weighted by complex coefficients modifying both amplitude and phase. Now, looking at figure 4.3, it is clear that the frequency domain output of the model, if the filters are removed, would have the form of

$$\sum_{n=1}^{N} \mathcal{P}_n = \sum_{n=1}^{N}\sum_{m=1}^{M} a_{n,m}\mathcal{H}_m$$

thanks to the Fourier transform linearity. Hence, single harmonics of order $m \leq N$ contained in the output of the model can be isolated just by restricting the sum to a frozen value of m:

$$\sum_{n=1}^{N} a_{n,m}\mathcal{H}_m$$

However, given the input angular frequency ω, its m-th harmonic would have to have amplitude and phase (a weight) given by $H_m(\omega)$ to match the model. Hence, the filters must be introduced to redefine the weights of the sum:

$$H_m(\omega) = \sum_{n=1}^{N} a_{n,m}G_n(\omega)$$

by comparison with $\mathbf{A}^T G(\omega)$:

$$\left\{\mathbf{A}^T G(\omega)\right\}_m = \sum_{n=1}^{N} a_{m,n}G_n(\omega)$$

it is concluded:

$$\begin{pmatrix} H_1(\omega) \\ H_2(\omega) \\ \vdots \\ H_m(\omega) \\ \vdots \\ H_M(\omega) \end{pmatrix} = \mathbf{A}^T \begin{pmatrix} G_1(\omega) \\ G_2(\omega) \\ \vdots \\ G_n(\omega) \\ \vdots \\ G_N(\omega) \end{pmatrix}$$

with ^{T} denoting transposition. Repeating this for each other value of positive discrete frequency:

$$\underbrace{\begin{pmatrix} H_1(f_d) \\ H_2(f_d) \\ \vdots \\ H_m(f_d) \\ \vdots \\ H_M(f_d) \end{pmatrix}}_{\doteq \vec{\mathbf{H}}} = \mathbf{A}^T \underbrace{\begin{pmatrix} G_1(f_d) \\ G_2(f_d) \\ \vdots \\ G_n(f_d) \\ \vdots \\ G_N(f_d) \end{pmatrix}}_{\doteq \vec{\mathbf{G}}} \quad \forall f_d \in \left\{1, 2, \ldots, \overset{+}{F}{}_d^H\right\} \tag{4.5}$$

The negative frequency part being given by Hermitian symmetry. The problem of the determination of the $G_n(f_d)$ is then purely algebraic and the solution of the system of equations above can be found in two ways.

$N = M$: direct solution. The matrix $\mathbf{A}^T$ is a square matrix and as such it can be inverted. Hence (Novak et al., 2010b):

$$\vec{\mathbf{G}} = \left(\mathbf{A}^T\right)^{-1} \vec{\mathbf{H}} \quad \forall f_d \in \left\{1, 2, \ldots, \overset{+}{F}{}_d^H\right\} \tag{4.6}$$

where the $\vec{\mathbf{H}}$ and $\vec{\mathbf{G}}$ matrices are defined in equation (4.5) as matrices of stacked frequencies responses.

$M > N$: least squares approximation In this case there are M equations and N unknowns and the system cannot be solved directly. As discussed by Strang (1986, p. 35), the problem can be formulated in terms of sub-spaces. In fact, the right hand side of equation (4.5) represents, for each positive frequency, a combination of the columns of $\mathbf{A}^T$ obtained with N complex coefficients $G_n(f_d)$, i.e. an element in a N-dimensional sub-space of the current ambient Hilbert space $\mathbb{C}^M$ where, associated to each value of positive frequency, exists a vector $\vec{\mathbf{H}}$. Hence, the best solution for the problem is the solution which is contained in this subspace as much as possible. Defining the error vector $\vec{\mathbf{E}} \in \mathbb{C}^M$:

$$\vec{\mathbf{E}} \doteq \vec{\mathbf{H}} - \mathbf{A}^T \vec{\mathbf{G}} \quad \forall f_d \in \left\{1, 2, \ldots, \overset{+}{F}{}_d^H\right\}$$

the goal is to minimize its length:

$$\left\|\vec{\mathbf{E}}\right\|^2 = \left(\vec{\mathbf{E}}\right)^T \vec{\mathbf{E}} = \left(\vec{\mathbf{H}} - \mathbf{A}^T\vec{\mathbf{G}}\right)^T \left(\vec{\mathbf{H}} - \mathbf{A}^T\vec{\mathbf{G}}\right) = \left\|\vec{\mathbf{H}} - \mathbf{A}^T\vec{\mathbf{G}}\right\|^2 =$$
$$= \left(\vec{\mathbf{G}}\right)^T \mathbf{A}\mathbf{A}^T\vec{\mathbf{G}} - \left(\vec{\mathbf{G}}\right)^T \mathbf{A}\vec{\mathbf{H}} - \left(\vec{\mathbf{H}}\right)^T \mathbf{A}^T\vec{\mathbf{G}} + \left(\vec{\mathbf{H}}\right)^T \vec{\mathbf{H}}$$

where the algebraic properties of transposition have been used (see Appendix A.5). The last term is constant and does not affect the minimization. The terms in the middle are the same since the inner product of complex vectors is not affected by order. The quantity useful to minimize then can be found after a division by 2, which transforms the problem into the well known minimization problem of the quadratic form

$$\begin{aligned}&\frac{1}{2}\left(\vec{\mathbf{G}}\right)^T \mathbf{A}\mathbf{A}^T\vec{\mathbf{G}} - \left(\vec{\mathbf{G}}\right)^T \mathbf{A}\vec{\mathbf{H}} = \\ &= \frac{1}{2}\left[\vec{\mathbf{G}} - \left(\mathbf{A}\mathbf{A}^T\right)^{-1}\mathbf{A}\vec{\mathbf{H}}\right]^T \left(\mathbf{A}\mathbf{A}^T\right)\left[\vec{\mathbf{G}} - \left(\mathbf{A}\mathbf{A}^T\right)^{-1}\mathbf{A}\vec{\mathbf{H}}\right] - \frac{1}{2}\left(\mathbf{A}\vec{\mathbf{H}}\right)^T \left(\mathbf{A}\mathbf{A}^T\right)^{-1}\mathbf{A}\vec{\mathbf{H}}\end{aligned}$$

where the properties in Appendix A.5 have been used. The solution is the least squares solution (Strang, 1986, p. 37) whose application is suggested by Novak et al. (2010a):

$$\vec{\mathbf{G}} = \left(\mathbf{A}\mathbf{A}^T\right)^{-1}\mathbf{A}\vec{\mathbf{H}} \tag{4.7}$$

as easily seen from the equation above, as it causes the first term to disappear.

4.3 Implementation

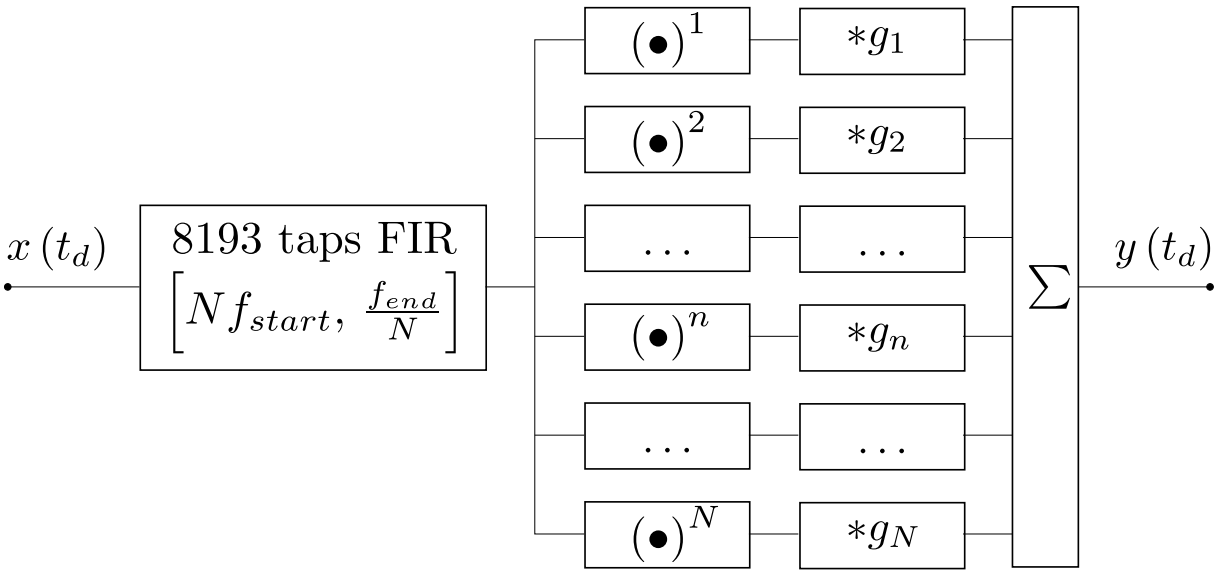


Figure 4.4: Implemented nonlinear model.

The implementation is based on Novak (2014). The $h_m(t_d)$ have been gated over the windows given by the samples $\mathbb{G}^m = \left\{\mathcal{G}_1^m, \mathcal{G}_2^m, \ldots, \mathcal{G}_{T_d^h}^m\right\}$, $T_d^h = 1024$, centered around $[\triangle t_m]_d \doteq F_d^H - \text{round}(\triangle t_m f_s)$ by `Response_Study.m` (round as in equation (3.21)). Since it isn't said that the lags $\triangle t_m$ can be represented with an integer number of samples $[\triangle t_m]_d$ the various $h_m(t_d)$ have been gated correcting the eventual non-integer gap $\triangle_m$ in the

frequency domain:

$$h_m(t_d) = \mathcal{F}^{-1}\left\{\mathcal{F}\left[\overset{\circ}{\widetilde{h}}(t_d)\Big|_{\mathbb{G}_m}\right] e^{-j\triangle_m\Omega}\right\} \quad \forall t_d \in \mathbb{G}^m$$
$$\triangle_m \doteq \left[F_d^H - \triangle t_m f_s\right] - \left[\triangle t_m\right]_d, \quad \Omega \doteq 2\pi\frac{f_d}{F_d^H}$$

being Ω the normalized discrete frequency and $\overset{\circ}{\widetilde{h}}(t_d)$ a circularly extended version of $\widetilde{h}(t_d)$ defined to facilitate the isolation of the linear response. From there, $\overrightarrow{\mathbf{H}}$ is obtained applying $\mathcal{F}[\bullet]$ with an Hamming window to reduce truncation artifacts. `Ident.m` implements directly equation (4.6) or equation (4.7) to obtain $\overrightarrow{\mathbf{G}}$. To ensure good accuracy while not raising the computational costs too much the most linear systems (Rotel and Sugden) have been identified with $N = 10$, while the Quad II with $N = 14$. $M = 20$ has been chosen to include information also from a number of higher responses to increase accuracy.

To use the model properly when synthesizing signals it is necessary to make sure that the sampling theorem is not violated and that there is enough information to represent the system (Novak et al., 2010a). For the first problem, considering that the model will generate N harmonics, the input signal should not contain frequencies whose N-th harmonics are higher than the Nyquist frequency $f_s/2$. For the second, the information about the m-th higher order response is acquired when the SES sweeps around mf_{start}, which happens at $\triangle t_m$. So, the minimal information to identify N responses in the model is collected only after $\triangle t_N$ seconds. Being the frequency spectrum of the SES time based, this is mirrored in the information being contained only after Nf_{start} in the frequency domain. As such, the input signal should have frequency content not lower than Nf_{start}. Considering that $f_s/2 = f_{end}$ an input pass-band filter over the bandwidth

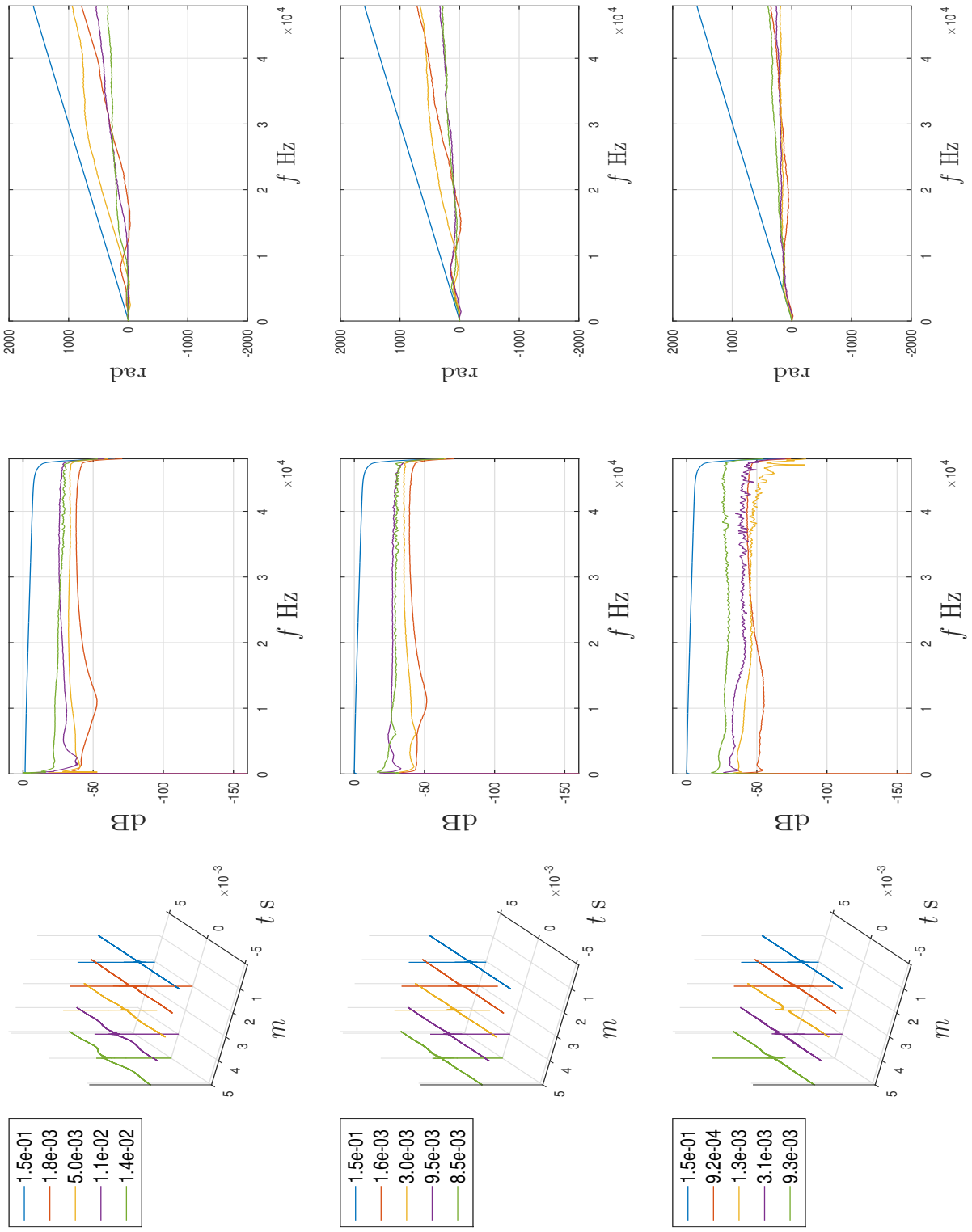
$$\left[Nf_{start}, \frac{f_{end}}{N}\right]$$

will prevent errors, showing a clear limitation of the bandwidth of the model. This filter has been implemented as a window-based (Hamming) designed FIR filter to prevent phase distortion and obtain a reasonably flat frequency response. The number of taps used is 8193. The implemented model is depicted in figure 4.4. A model of this kind has been identified for every NTI system - SES amplitude couple.

4.4 Outcome and validation

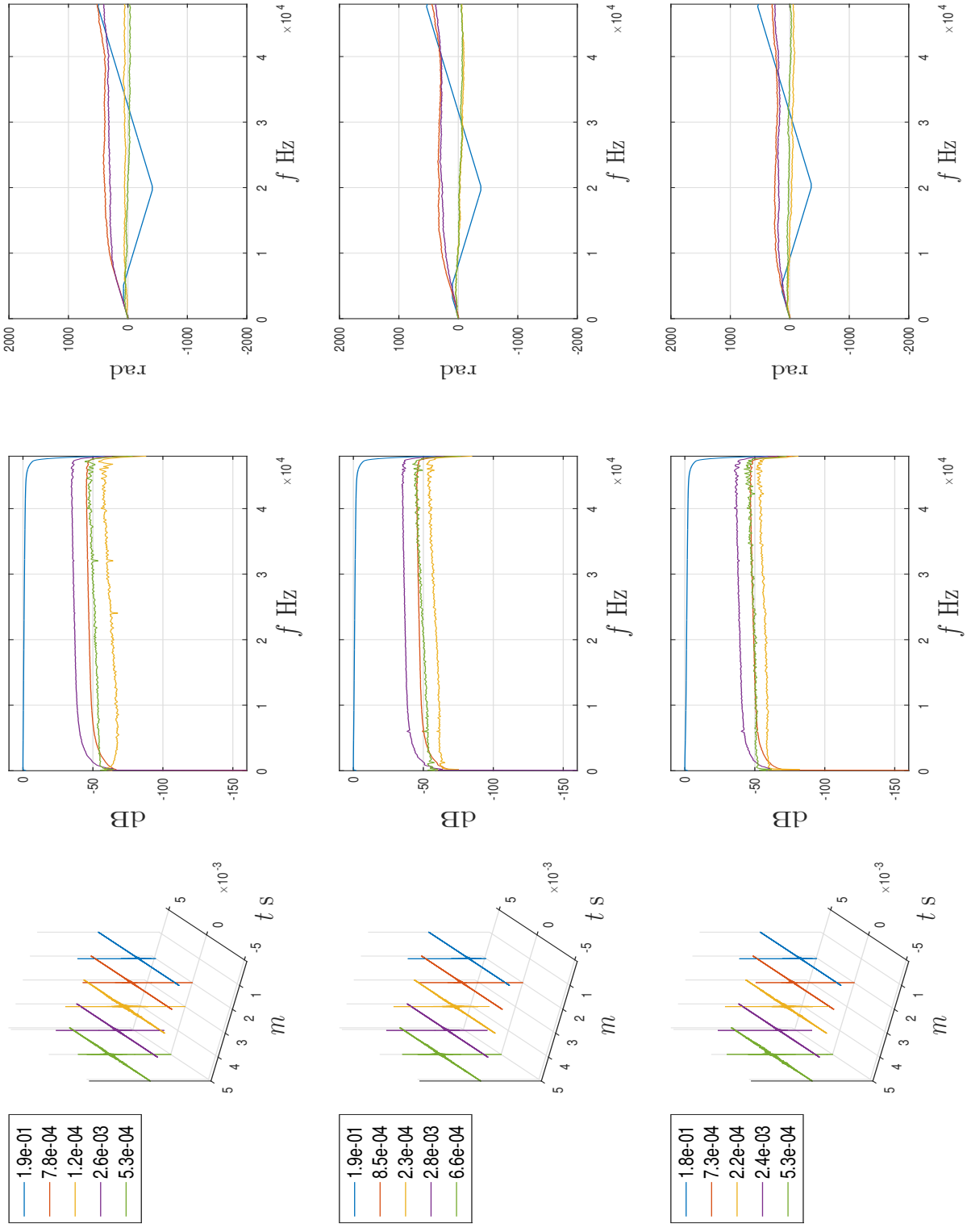
4.4.1 Linear blocks responses

Comparison of the first five $G_n(f_d)$ is performed in figure 4.5. Some of the responses, especially the higher orders, exhibit unexpected local peaks. These peaks are due to harmonics of the multi-tone signal used in the Latency Loopback, leaked in the measurement loop because cross-talk (see figure 3.2). They could be removed by, for example, tuning

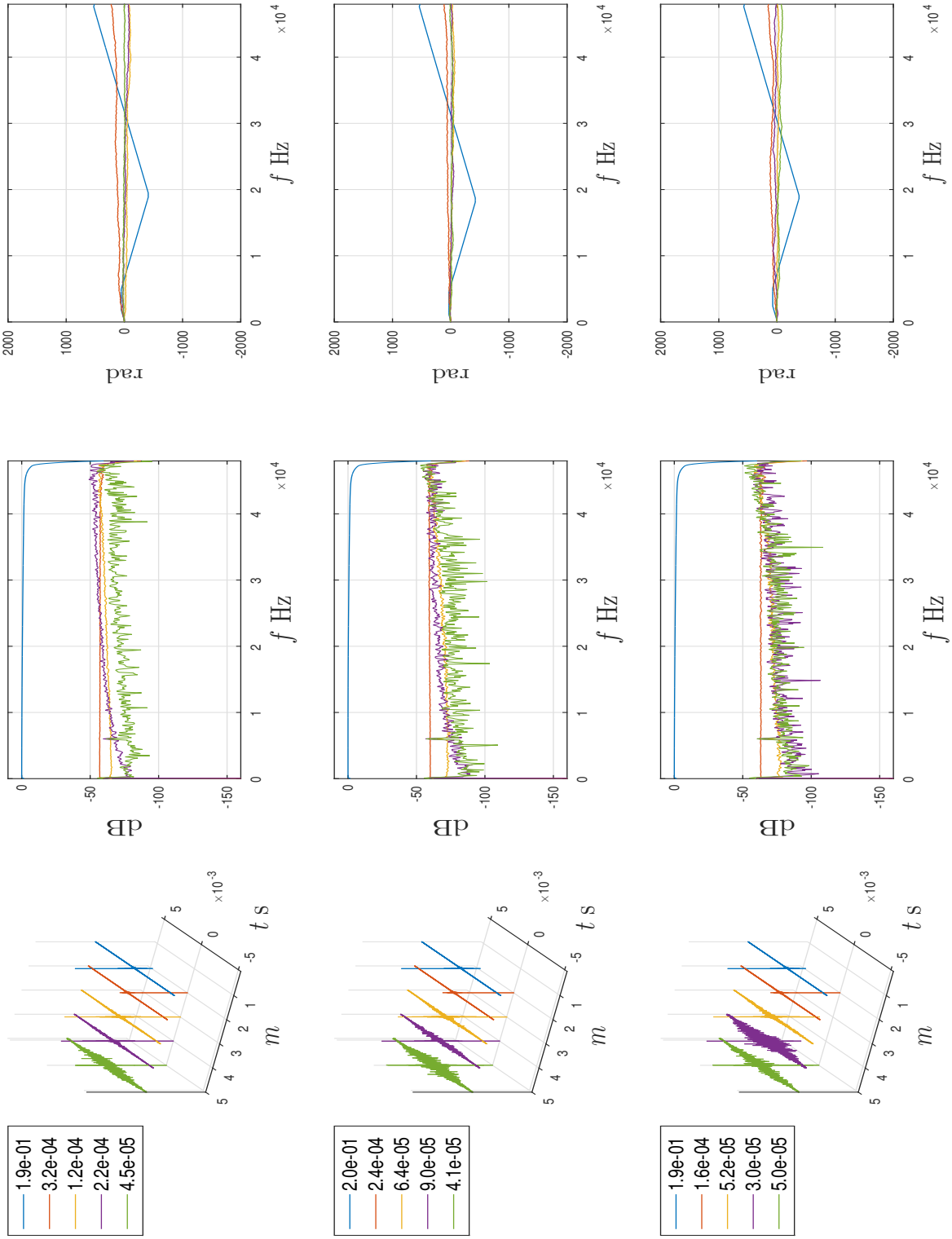


(a) Quad II

Figure 4.5: *PHM* responses. (Rotating the page ↻) On the left: normalized higher order impulse responses. The values in legend indicate the absolute peak values, arbitrary units after processing with normalized signals (both in value and unit). In the middle: higher order frequency responses levels normalized to absolute peak linear ($m = 1$) response. On the right: associated phase responses. Plots arranged vertically from the top in order of decreasing input signal power (Full Power - Half Power - Quarter Power).


rad
f Hz
dB
t s
m
1.9e-01
7.8e-04
1.2e-04
2.6e-03
5.3e-04
1.9e-01
8.5e-04
2.3e-04
2.8e-03
6.6e-04
1.8e-01
7.3e-04
2.2e-04
2.4e-03
5.3e-04

(b) Rotel


rad
dB
f Hz
t s
m
1.9e-01
3.2e-04
1.2e-04
2.2e-04
4.5e-05
2.0e-01
2.4e-04
6.4e-05
9.0e-05
4.1e-05
1.9e-01
1.6e-04
5.2e-05
3.0e-05
5.0e-05

(c) Sugden

T_d^h to be smaller, reducing the frequency sample density and thus averaging the curves. If a good compromise with accuracy is found, this would appear as robustness of the method against tonal noise[17].

4.4.2 Validation

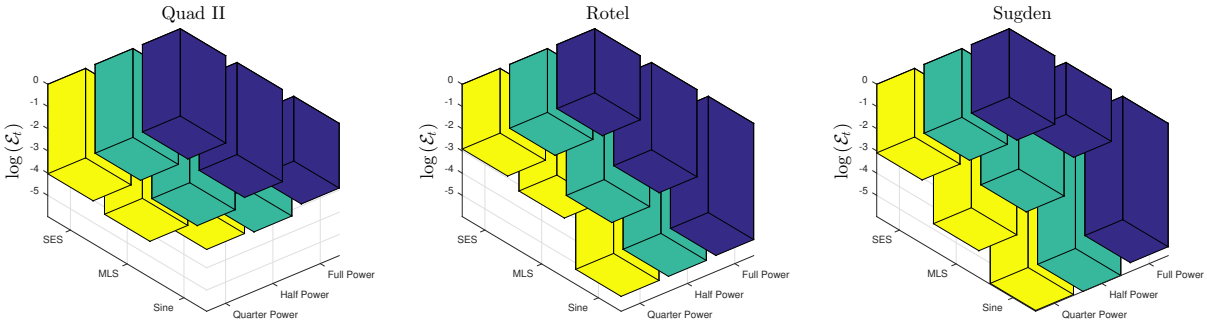


Figure 4.6: Time domain errors given as $\log(\mathcal{E}_t)$ for clarity.

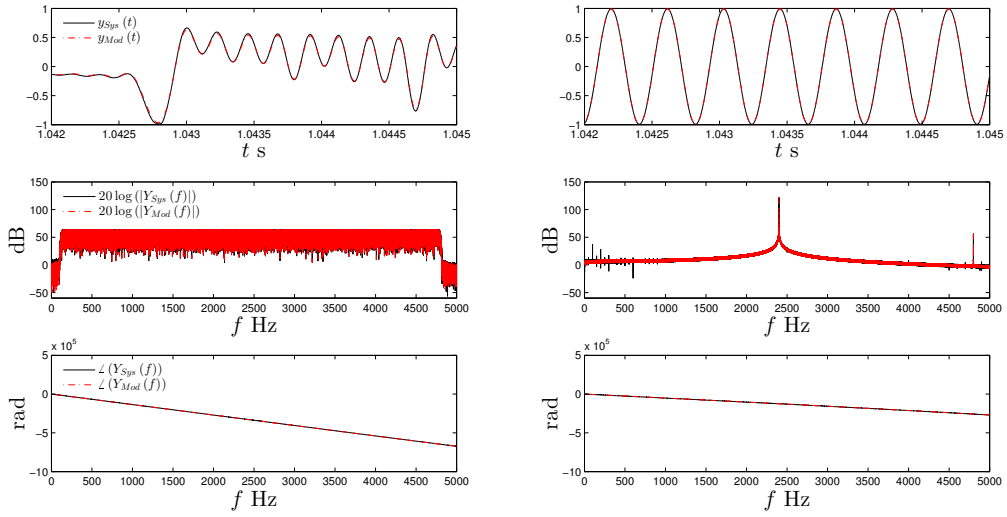


Figure 4.7: Comparison of signals couples with minimal and maximal $\mathcal{E}_t$. On the left, maximal error couple (MLS from Sugden Full Power and relative model). On the right, minimal error couple (sine from Sugden Quarter Power and relative model). Quantities plotted reported in the left panels (all time series are normalized to 1 peak amplitude both in value and unit). Time plots intervals chosen for clarity. Frequency plots interval over the bandwidth of the models.

During the measurement phase not only the stimulation signal has been fed to the systems, but also other signals intended for model validation (see Section § 3.3 on page 23). Validation is performed by comparing the outputs of models and systems after the same input signals. The quantitative comparison based on the Mean Squared Error (MSE)

[17] In this case, all the peaks were outside the model pass-band.

suggested by Novak (2009) is here substituted by functional analysis concepts in order to obtain higher local sensibility of the errors. Consider the sequences $x(t_d)$, $y_{Sys}(t_d)$ and $y_{Mod}(t_d)$, where the subscripts clarify the source of the output. They belong to the Hilbert space $\mathbb{R}^{T_d^x}$. Hence, the natural way to assess how close two functions (vectors) are is through squared norm:

$$\mathcal{E}_t \doteq \|y_{Sys}(t_d) - y_{Mod}(t_d)\|^2 = \frac{1}{[1\,\mathrm{s}]} \cdot \frac{1}{f_s} \sum_{t_d=1}^{T_d^x} |y_{Sys}(t_d) - y_{Mod}(t_d)|^2$$

where, being all the signals normalized to unit peak amplitude, $1/[1\,\mathrm{s}]$ makes $\mathcal{E}_t$ dimensionless. The same concept applies in the Hilbert space $\mathbb{C}^{F_d^X}$ by Fourier transforming all the signals over the length F_d^X :

$$\mathcal{E}_f \doteq \|Y_{Sys}(f_d) - Y_{Mod}(f_d)\|^2 = \frac{f_s}{F_d^X} \sum_{f_d=1}^{F_d^X} |Y_{Sys}(f_d) - Y_{Mod}(f_d)|^2$$

where the capital letters denote the transforms. The choice of the constants normalizing the sums makes the numerical values very similar and only the errors in time domain $\mathcal{E}_t$ are then evaluated[18]. The errors are compared in 4.6. Each model shows a different accuracy pattern, which is not surprising given the blind identification of different systems with the same structure. Indeed, Sugden and Rotel, being technologically similar, have more similar error patterns. All the errors appear reasonably small, ranging from $1.1 \cdot 10^{-6}$ to $4.6 \cdot 10^{-3}$. This is remarkable, since Novak (2009, p. 79) warns that weakly nonlinear systems can be harder to identify. The success of the operation is attributed to the long SES length ($\sim 40\,\mathrm{s}$) that makes easy to unambiguously separate the higher order responses and the reasonably high sample size ($N_y = 16$) used to average out the already low background noise, making possible for even the weakest higher order responses to show at least a trend. Figures (4.6) and (4.7) supply both visual and quantitative display of the errors.

4.5 Summary

Nonlinear models have been examined and the PHM has been selected for the identification process. The unknown of the models are just the linear frequency responses $G_n(f_d)$ and they can be found from the partial frequency responses $H_m(f_d)$ by pure algebraic means. The models have been validated comparing different kinds of outputs as given by systems and models under the same inputs. Given the overall low magnitude of the errors the models identification is judged successful.

[18] The MATLAB® scripts compute also the MSE, but here it is not reported as it shows the same exact trends.

5 Listening test

5.1 Introduction

Models can be profitable to study the perceptual outcome of nonlinear signal manipulation since all the parameters can be directly controlled. In this sense the PHM, being linked with simplicity to the higher order responses, could be a valuable tool. Study through models is possible only if the systems are well represented also perceptually. Although the errors in figure 4.6 are very small there is not a reason to exclude any perceptible difference between any couple of real systems outputs and modeled outputs. For this reason, a simple double blind ABX listening test has been designed.

5.2 Experiment design

The double blind ABX test is the standard test to determine whether audible differences exist between signals (Boley & Lester, 2009). It is organized in multiple trials during which a subject will listen to three audio excerpts, A, B and X. A and B are known to be different while X is A or B selected at random. The task of the subjects is to identify X. The design, developed following the guidelines of mainly Boley and Lester (2009) and International Telecommunication Union (2015), will be detailed in the following paragraphs. It will be evident that the simple test methodology, compatible with the time scale allowed for the project, cannot answer all the questions that can be posed on the topic and should therefore be regarded as a pilot study.

5.2.1 Overview

3 Systems have been measured and identified with 3 different stimulation amplitudes. The System / Model couples to be tested are then 9. They are referred as Sets. The test consisted of two phases: a familiarization phase and the actual test. To avoid effects of fatigue for the subjects the test must be as short as possible (International Telecommunication Union, 2015). For this reason, the familiarization consisted of a brief session in which the subjects can run a simulation of the test. It was, as such, intended to help the subjects to develop confidence instead of expertise. The actual tests were run soon after the familiarization and consisted of many more trials.

5.2.2 Listening panel

All the subjects have been recruited among Acoustics Master and PhD students. As such, they are competent about the matter. Also, their age range is not excessive, limiting the likelihood of having subjects with hearing impairments related to age.

5.2.3 Test signal

Kind of signal According to Boley and Lester (2009) a good test signal should be able to exploit as possible the difference between the two signal processing tools under examination. A weakness of the models resides in the fact that their parameters are after a stimulation at a well determined amplitude. As such, synthesis at mismatched or unsteady instantaneous amplitude will be inaccurate, especially towards the clipping point. Moreover, it is fair to compare models and systems with the signals they are intended to process, namely music, which amplitude is unsteady. For these reasons, music is selected to be the test signal.

The pass band of the models is limited by $[Nf_{start},\ {}^{f_{end}}/_{N}]$, where N is maximum (14) for the Quad II unit. It makes sense then to select the narrowest pass band ($N = 14$) and search for tracks whose natural bandwidth is contained in

$$[168\,\text{Hz},\ 3428\,\text{Hz}]$$

Having usually range between $C_3 \sim 165\,\text{Hz}$ and $D_6 \sim 1175\,\text{Hz}$, a classical guitar track fits naturally the purpose. Guitar has also fairly impulsive transients that can stimulate well the non-linearity of the systems, comprising transient intermodulation distortion. Also, the natural clarity of the acoustic instrument should help the subjects in spotting distortion. The input excerpt has been filtered over [168 Hz, 3428 Hz] by the same FIR of figure 4.4 before to be fed to systems and models[19].

The small errors of figure 4.6 suggest that the perceptual differences between A and B will be very subtle. As such, it has been preferred to use only one single excerpt from La Catedral as recorded by Williams (1995), the first 3 seconds[20], so that the subjects would be helped by familiarizing with the test material. The particular excerpt has been selected since it contains both transients and prolonged notes and it is relatively slow. The first two qualities make the signal able to stimulate well distortion, while the second limits the annoyance the subjects would develop by listening to the same fast passage over and over.

Calibration The input signal gain has been calibrated in order to make the instantaneous amplitude fluctuating around the measurement stimulation amplitude as close as possible while avoiding peaks higher than the full power SES amplitude, that would excite the systems in a region where they have never been measured.

Outputs generation The outputs from the systems have been acquired after averaging over 16 recorded outputs to remove background noise while the outputs from the models

[19] In this case, the FIR in figure 4.4 has been turned off as already used.

[20] The excerpt has been separated and the end faded out over the last millisecond to avoid clicking noise due to abrupt audio interruption. Initial and final zero pads have been added to reach a global length of 4.5 s.

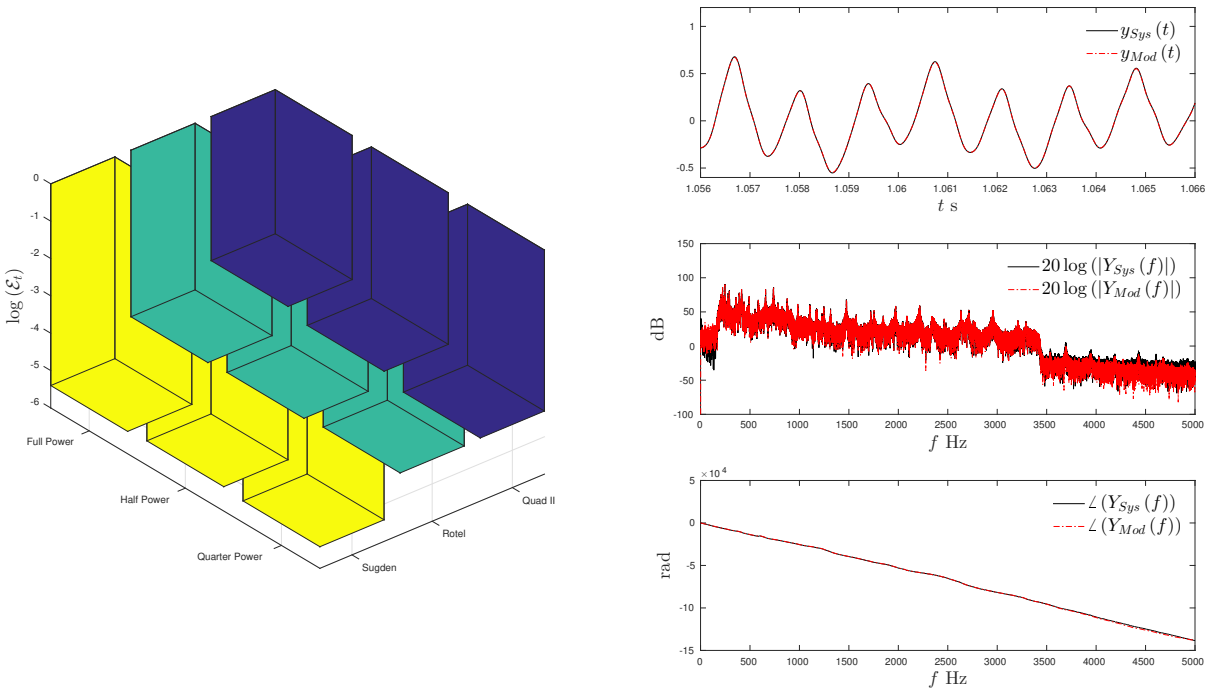


Figure 5.1: *ABX* test signals errors (on the left) and maximal $\mathcal{E}_t$ ($1.4 \cdot 10^{-4}$ for Quad II Full Power) comparison between system and model outputs (on the right). Time domain amplitudes normalized to 1 peak amplitude both in value and unit. From top to bottom: particular of waveform around the attack of the first note in the excerpt. Frequency domain magnitude over the model pass band. Frequency domain phase over the model pass band.

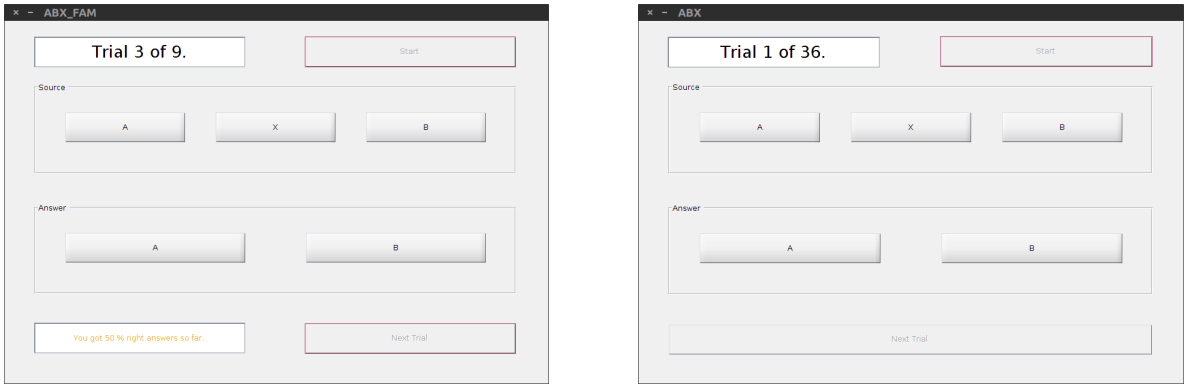
A	*B*	*X*
System	Model	System
System	Model	Model
Model	System	System
Model	System	Model

Table 5.1: All System / Model outputs possible combinations.

are simply synthesized feeding them with the same input. Systems and models outputs under the *ABX* test signal input are compared in figure 5.1. They appear bigger than in figure 4.6, considering also that the excerpt is short. This is a consequence of the unsteady input amplitude.

5.2.4 Familiarization

The 9 Sets have been sorted in order of decreasing $\mathcal{E}_t$ and each one of them used for a trial. For each trial, the content of *A*, *B* and *X* has been chosen by randomly selecting a row of table 5.1. To raise the confidence of the subjects a panel giving feedback on the number of correct answers has been supplied by the test program (figure 5.2).



(a) Familiarization phase GUI.

(b) Actual test GUI.

Figure 5.2: GUI's of the test program.

5.2.5 Actual test

The subjects should be tested with all the combinations of table 5.1 for each of the 9 Sets. In fact, this will make sure that 50% of the possible answers are correct. However, the test should be distributing randomly these answers and the order of presentation of the trials should be different for each subject to avoid it to be a confounding factor (Boley & Lester, 2009). These are fundamental assumptions underlying the binomial statistical analysis. To meet them, for each Set, the combinations of table 5.1 have been generated. Then, they have been stacked in a matrix of 36 rows. Finally, the rows order has been randomized by a different random permutation for each subject. It should be noted that testing with all the possible combinations removes the risk of regularities in the contents of A, B and X to be a confounding factor.

5.2.6 Test program

The test program has been written in `MATLAB®` and makes use of the GUI's reported in figure 5.2. The GUI for the ABX test has a symmetric disposition of the controls to remove any visual bias. The subjects can decide an answer only after listening at least one time to the full ABX sequence. To increase sensibility, the subjects can listen to each of A, B and X in any order and as many times they want, as suggested by Clark (1982).

5.2.7 Test duration

Each trial requires around 1 min to be assessed by a subject. The recommendations of International Telecommunication Union (2015) are then exceeded by 6 min, 15 min if the familiarization phase is counted. This isn't optimal, but it is judged fair for the following reasons. The test couldn't be made shorter considering the number of trials needed. Also, the ABX task makes use of short term memory (Clark, 1982), making any pause in the test disruptive. In fact, it has to be expected that greater sensibility will come from continuous exposition to the test signal when dealing with subtle differences.

5.2.8 Boundary conditions control

Devices A more modern laptop computer[21] has been preferred to carry on the tests. This because the real time audio capabilities of MATLAB® are limited and a powerful machine is needed to prevent audio dropouts. For this reason, being the bandwidth of the signal limited, the internal sound-card, operating on the more stable PCI bus, has been preferred. Also, the laptop is very silent, implementing SSD memory only, preventing distracting noise. To aid concentration the audio has been reproduced through a set of headphones (Samson SR950). They have been selected as their frequency response appears flat over the test bandwidth (Samson Technologies Corp, 2012).

Listening level The listening level was high enough to make all the excerpts characteristics evident while not being so loud to result in possible annoyance. I has been kept constant among all the subjects.

Test room Having designed the test with headphones, International Telecommunication Union (2015) only mandates to control the background noise level. For this reason, a semi anechoic chamber (G43) has been used.

Instructions The same instructions have been given orally to all the subjects with the aid of a dummy test to avoid this to be a confounding factor. Instructions included:

- Project background.
- Test description.
- How to operate the GUI's.
- Where differences in the excerpts are more likely to be heard (around transients, given the input amplitude sensibility of the models, consisting of mainly subtle timbre and clearness differences).
- Test duration.

5.3 Statistical analysis

In an ABX test trial the subject can either succeed or fail in identifying X, i.e. the outcome is classified using two mutually exclusive and exhaustive categories. As such, it is a Bernoulli trial. The most relevant techniques to analyze this kind of trials are the binomial analysis and the Signal Detection Theory analysis . Binomial analysis requires to assume that subjects unable to identify the correct answer are random answer generators, while Signal Detection Theory can deal with the opposite case of subject bias towards

[21] Entroware Apollo.

one answer (Boley & Lester, 2009). They are both used. The confidence level CL is set to 0.95 as it is a commonly used value in listening tests, enabling simple comparison with other similar works if required.

5.3.1 Binomial analysis

A subject that behaves as a random answer generator will have a probability $p = 1/2$ of succeed in each trial. The binomial distribution $\mathscr{B}$ is then the reference distribution, as it supplies the probability P of having n successes in N_T trials with constant success probability p:

$$\forall N_T \in \mathbb{N},\ \forall n \in \{0, 1, \ldots, N_T\} \subset \mathbb{N},\ \forall p \in [0,\ 1]$$
$$\mathscr{B}\left(n;\ N_T, p\right) \doteq P\left(N = n\right) = \binom{N_T}{n} p^n \left(1 - p\right)^{N_T - n}$$

with N the random variable "number of successes". Thus, the cumulative distribution $\mathscr{C}_{\mathscr{B}}$ supplying the probability of having up to n successes in N_T trials is:

$$\mathscr{C}_{\mathscr{B}}\left(n;\ N_T, p\right) \doteq \sum_{i=0}^{n} P\left(N = i\right) = \sum_{i=0}^{n} \binom{N_T}{i} p^i \left(1 - p\right)^{N_T - i}$$

This function can be inverted to yield $\mathscr{C}_{\mathscr{B}}^{-1}\left(P;\ N_T, p\right)$, which accepts a probability P and returns a corresponding integer so that $\mathscr{C}_{\mathscr{B}}$ is equal or exceeds P.

Clearly, a subject can be said to be successful at correctly identifying X only if his/her answers are not compatible with the ones a random subject would give. For $N_T = 36$ and $P = CL$:

$$S = \mathscr{C}_{\mathscr{B}}^{-1}\left(CL;\ N_T, p\right) = 23$$

meaning that the probability to have, by chance, more than 23 correct answers is equal or smaller than $1 - CL = 0.05$. However, it can happen that a subject is very consistent at giving the wrong answer. This can still mean that the difference between A and B can be perceived, being possible that a subject has inverted the decision criterion (Boley & Lester, 2009). By similar reasoning then $s = \mathscr{C}_{\mathscr{B}}^{-1}\left(1 - CL;\ N_T, p\right)$. By unit subtractions s is scaled until $\mathscr{C}_{\mathscr{B}}\left(s;\ N_T, p\right)$ is equal or smaller than $1 - CL$:

$$s = 12$$

meaning that the probability to have, by chance, up to 12 correct answers is equal or smaller than $1 - CL$. S and s are the binomial confidence limits: if the number of correct answers for a subject is above S or below s, in the framework of the binomial theory, the subject can be said to be able to discern the differences between A and B at the confidence level CL. The outcome is showed in figure 5.3. Significance is not found, meaning that all the subjects are compatible with a random subject.

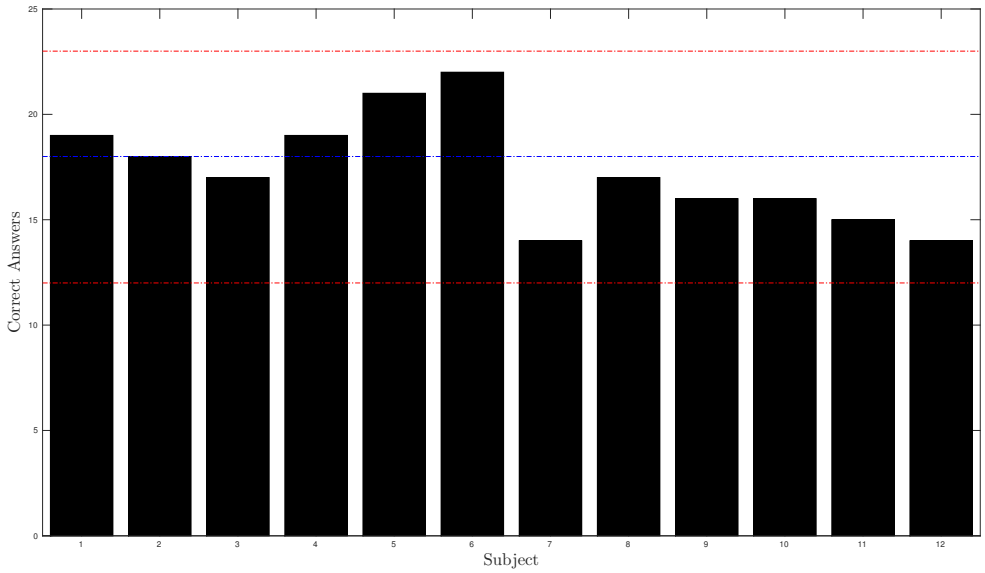


Figure 5.3: Binomial analysis outcome. Red lines: S and s. The blue line is drawn at 18 (half correct answers).

5.3.2 Signal Detection Theory

The ABX test is simple in structure but the task performed by the subjects, involving a number of comparisons between stimuli, is actually very complicated, as reported by Harris (1952). In fact, the subjects have to give an answer after the comparison by using a conscious or unconscious decision criterion. What triggers the decision, generating the impression of similarity or not between two excerpts, can be an objective perceptible difference (signal) or an illusion generated by the psycho-physics of the task (noise). In this regard Signal Detection Theory, invented in electronics to aid detection of signals in noisy background, can be really profitable.

Decision strategies Hautus and Meng (2002) experimented to determine the decision strategies associated with the ABX task, delineating two main kinds: the Difference Decision Strategy (D) and the Independent Observations Decision Strategy (IO) respectively. When using D the subjects try to identify X as A or B assessing the difference in the perception between the couples A-X and B-X, selecting the one producing the minimal difference. The IO decision strategy is based on two steps: the difference between the perception of A and B is used to locate the stimuli in two different classes. Then, the perception of X is used to state to which class X belongs. Evidence from their research points out that the IO Strategy can be developed by usually a subset of trained listeners, that build the idea of classes through experience. Typical of naive subjects is the D Strategy.

Stimulus \ Answer	$X = A$	$X = B$
$X = A$ ($N_A = \frac{N_T}{2}$)	Hit (N_H^A)	False Alarm (N_F^B)
$X = B$ ($N_B = \frac{N_T}{2}$)	False Alarm (N_F^A)	Hit (N_H^B)

Table 5.2: Response matrix. The $N_{\square}^{\square}$ notate the amounts of. $N_A = N_B = {}^{N_T}/_2$ due to the test design.

Significance The complete description of the distributions associated to each of the decision strategies are reported by Hautus and Meng (2002). In addition, the noise and signal distributions can be assumed Gaussian for simplicity. The D, IO and Gaussian models are all tested. The standard way to assess significance in Signal Detection Theory is through the sensitivity index d', a measure of the separation of the means of the noise and signal distributions in standard deviation units. Real values of d' represent the ability of a subject to distinguish signals from noise, ranging from 0 (inability) to $+\infty$. However, in experiments each subject answers are an independent source for d' estimation. As such, the different values populate a distribution. Negative d' can arise after samples errors, 0 d' distribution mean or subject criterion inversion (Stanislaw & Todorov, 1999). For this reason, for each subject, a confidence interval for d' must be built. If it contains 0 then there isn't evidence to conclude that the subject was able to correctly identify X since this implies that, at the current CL, the noise and signal distributions can be superimposed (hence, non discernible). d' is calculated in terms of hit H and false alarm F rates. They can be calculated resorting to table 5.2 as follows:

$$H_A \doteq \frac{N_H^A}{N_A},\ F_A \doteq \frac{N_F^A}{N_A},\ H_B \doteq \frac{N_H^B}{N_B},\ F_B \doteq \frac{N_F^B}{N_B}$$

For the Gaussian model the d' estimates can be calculated as follows (Boley & Lester, 2009):

$$d'_A\left(H_A,\ F_A\right) \doteq \mathscr{C}_{\mathscr{G}}^{-1}\left(H_A\right) - \mathscr{C}_{\mathscr{G}}^{-1}\left(F_A\right) \tag{5.1}$$

with $\mathscr{C}_{\mathscr{G}}^{-1}\left(P\right)$, $P \in [0,\ 1]$, the Normal Inverse Cumulative distribution. d'_B is given just by subscript substitution. The calculation of d' for the D and IO models, succinctly referenced by Boley and Lester (2009), can be found in Macmillan and Creelman (2004). For this project, the `dprime.ABX` function for `R` found in Knoblauch (2014), written after Macmillan and Creelman (2004), has been used. The proportion correct of an unbiased observer $\mathcal{P}_c$ is given by (Knoblauch, 2014):

$$\left(\mathcal{P}_c\right)_A \doteq \mathscr{C}_{\mathscr{G}}\left(\frac{\mathscr{C}_{\mathscr{G}}^{-1}\left(H_A\right) - \mathscr{C}_{\mathscr{G}}^{-1}\left(F_A\right)}{2}\right) \in [0,\ 1]$$

$$\begin{cases} \left(\mathcal{P}_c\right)_A > {}^1/_2 & \text{if } \mathscr{C}_{\mathscr{G}}^{-1}\left(H_A\right) > \mathscr{C}_{\mathscr{G}}^{-1}\left(F_A\right) \\ \left(\mathcal{P}_c\right)_A = {}^1/_2 & \text{if } \mathscr{C}_{\mathscr{G}}^{-1}\left(H_A\right) = \mathscr{C}_{\mathscr{G}}^{-1}\left(F_A\right) \\ \left(\mathcal{P}_c\right)_A < {}^1/_2 & \text{if } \mathscr{C}_{\mathscr{G}}^{-1}\left(H_A\right) < \mathscr{C}_{\mathscr{G}}^{-1}\left(F_A\right) \end{cases}$$

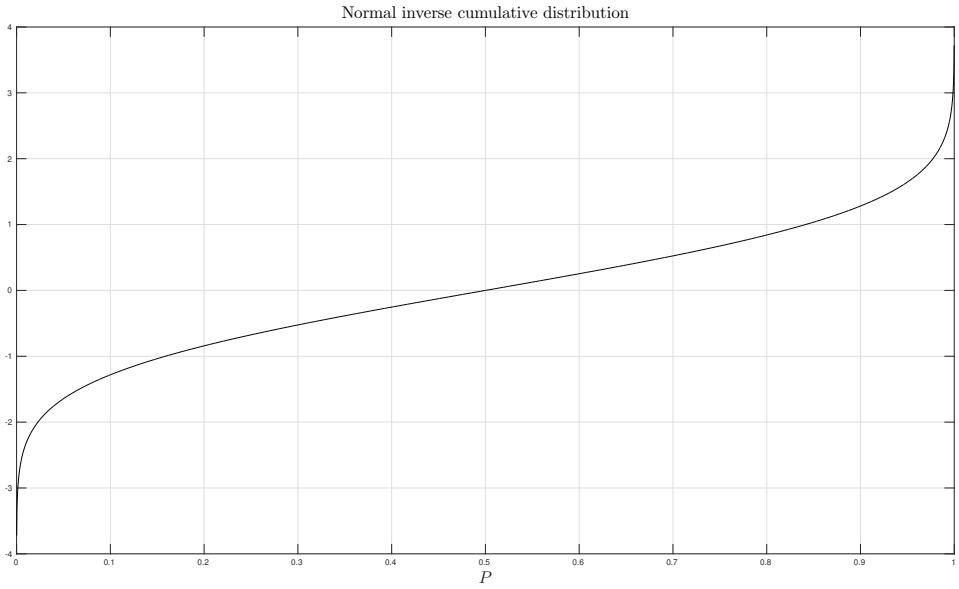


Figure 5.4: $\mathscr{C}_{\mathscr{G}}^{-1}(P)$, $P \in [0, 1]$.

where $\mathscr{C}_{\mathscr{G}}$ denotes the Normal cumulative distribution. With reference to figure 5.4:

$$\mathscr{C}_{\mathscr{G}}^{-1}(H_A) < \mathscr{C}_{\mathscr{G}}^{-1}(F_A) \iff H_A < F_A$$

Considering equation (5.1):

$$(\mathcal{P}_c)_A < \frac{1}{2} \iff H_A < F_A \implies d'_A(H_A, F_A) = \mathscr{C}_{\mathscr{G}}^{-1}(H_A) - \mathscr{C}_{\mathscr{G}}^{-1}(F_A) < 0;$$
$$-d'_A(H_A, F_A) = \mathscr{C}_{\mathscr{G}}^{-1}(F_A) - \mathscr{C}_{\mathscr{G}}^{-1}(H_A) = d'_A(F_A, H_A) > 0$$

Being d' an estimate of means distance the same properties apply also for G and IO models. In fact, $H_A < F_A$, which implies negative d' estimates, is symptomatic of subject decision criterion inversion, that is equivalent to switching of the roles of noise and signal distributions. `dprime.ABX` only accepts data for which $\mathcal{P}_c \geq 1/2$. As such, when $\mathcal{P}_c < 1/2$, d' for D and IO models have been computed by swapping the roles of H_A and F_A and making the results negative.

The confidence interval for d', which is a random variable, must be built starting from its variance. Boley and Lester (2009) suggest to assume convergence of H_A (H_B) and F_A (F_B) to a Gaussian distribution and use the Gourevitch and Galanter's approximation (Kaplan, 2009):

$$\sigma_A^2 \sim \frac{H_A(1 - H_A)}{N_A\left[\mathscr{C}_{\mathscr{G}}^{-1}(H_A)\right]^2} + \frac{F_A(1 - F_A)}{N_A\left[\mathscr{C}_{\mathscr{G}}^{-1}(F_A)\right]^2} \tag{5.2}$$

Since $N_T = 36$, the numbers of observations forming the rates are insufficient to justify this assumption[22]. In her review, Kaplan (2009) demonstrates that more accurate estimations

[22]Interestingly, since Boley and Lester (2009) concludes that Signal Detection Theory is more profitable

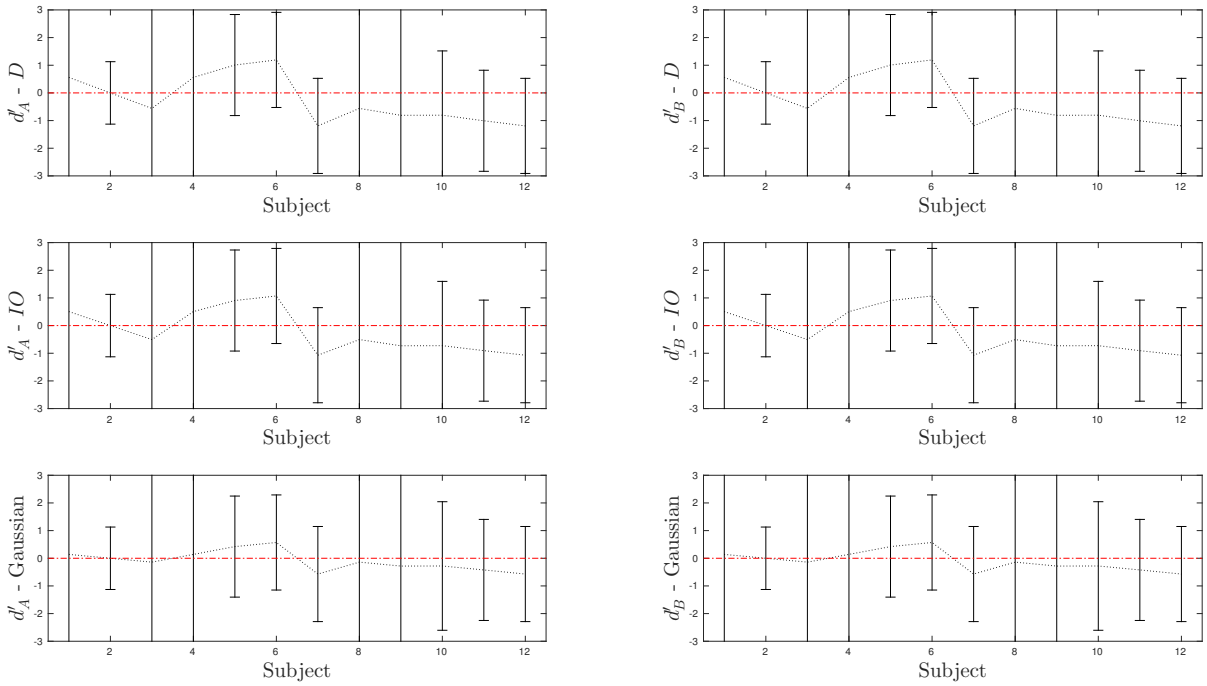


Figure 5.5: Signal Detection Theory outcome.

of variance need hundreds of observations anyway. For this reason, the advice by Boley and Lester (2009) is taken due to the simplicity of the variance estimate. When $H_A = 1/2$ and/or $F_A = 1/2$ equation (5.2) diverges. This result is interpreted as a fault of the approximation formula. When this happens, the associated variance will be supposed to be way bigger than the other variances involved. Once that the variance is estimated the standard deviation is calculated and multiplied, having assumed Gaussian behavior, by an appropriate m to reach the required coverage CL ($CL = 0.95 \implies m = 1.96$).

Outcome Results are reported in figure 5.5. At the chosen CL level, it is seen that in all the cases 0 is compatible with d'. Also, since around half of the subjects got negative d' estimates, it is probable that the average of the d' distribution is 0. Since, from table 5.2

$$N_H^A + N_F^B = N_A = \frac{N_T}{2} = N_B = N_F^A + N_H^B \implies N_H^A - N_F^A = N_H^B - N_F^B$$
$$\implies H_A - F_A = H_B - F_B$$

the equivalence of the results for d'_A and d'_B was expected and it supplies a confirmation that the test program distributed the test material as designed.

when dealing with a small number of trials...

5.4 Summary

A double blind *ABX* test has been designed to perceptually test the models. Binomial and Signal Detection Theory analysis proposed to study its outcome. The time scale of the project did not allow proper training of the subjects nor a high number of trials to be taken, thus making the study preliminary only. By design, the *ABX* typology cannot conclude that there is not perceptual difference, but only that there is no evidence to support perceptual differences claims. This methodology can be extended on a larger scale experiment in order to confirm or disprove the results. Having only 4 trials per Set, it is not possible to drive any conclusion about which models were most successful.

6 Conclusion

6.1 Outcome

A simple identification method for weak nonlinear systems has been applied to audio amplifiers in order to model their Input-Output relation. The use of SES stimulation, as a consequence of its properties, appears then very versatile since measure and identification of a nonlinear system can be performed just with one kind of stimulation. Remarkable is the robustness against noise, even when tonal, that enables sensibility over a high number of higher order impulse responses. However, the technique has clear weak points. First of all, the limited modeled pass-band $[Nf_{start},\ {}^{f_s}/_{N}]$ that forces high sample rates, and hence high computational costs, when attempting to broaden the bandwidth. Second, and most important, the model parameters depend on the signal amplitude, increasing the amount of information required for a complete identification and making real time applications very hard to implement.

Quantitative comparison of recorded and synthesized signals showed small errors in any case. The perceptual test was unable to collect evidence to support perceptible differences between signals and as such it is reasonable to suspect that there isn't any, although this point is yet be proven by a purposely designed experiment. This is a success overall, since it isn't said that methods that work well theoretically will perform well in real life, as noticed by Novak (2009) with reference to noise stimulation techniques. It should be kept in mind, however, that the systems were weakly nonlinear and with reasonably flat frequency Input-Output relations, as they were designed with fidelity in mind. If the weak non-linearity makes modeling harder (as the higher order responses are weak and easily covered by noise), and hence shows the power of the technique, the similarity of all the systems makes them more or less equivalent perceptually. It is then possible that the same technique performed on more discernible systems would have yield to a different perceptual outcome.

6.2 Further work

This work can be extended in many ways. The natural future work for this project would be an enhanced ABX listening test in order to determine if there is any perceptible difference. The best way to design it would be by training the subjects first, then organizing multiple sessions similar to the proposed one in order to collect data with a higher number of trials. If the results are not changed, a further test should be designed to prove that perceptible differences do not exist. This path should be followed involving systems with remarkably different global Input-Output relations, so that the results would not risk to be confounded by global similarity of signals. Also, provided higher quality sampling equipment, the research could be extended to strong NTI systems. A way to automatically take into account the input amplitude model parameters dependence should be studied.

Once the relation between nonlinear systems and models is clear, manipulation of the model parameters can be performed by mathematical and physical means. If possible, design of listening tests to explore the likely multidimensional perception of distortion artifacts can be done. The information gathered would be very profitable for many applications including, for example, DSP methods for raising clarity of speech in nonlinear transmission lines.

Other research paths exists as well. Novak (2009) suggested the use of the model for inverse loudspeaker filtering for non-linearity minimization, hysteretic system modeling, nonlinear echo cancellation. Also, for linear or weakly nonlinear mechanical systems, noticeable nonlinear behavior usually raise when the stresses on the system exceed its capabilities. If the system is part of a machine, non-linearity detection can give information about likely of faults or health status, as shown by Hassan, Coats, Gouda, Shin, and Bayoumi (2012). Modeling could then be employed for predictive fault control. Also, models can be used to correct scientific measurements performed under nonlinear vibrational energy transfer, a problem introduced by Sàez, Gavaldà, Ruiz, and Shevtsova (2014).

A Mathematical proofs, derivations and properties

A.1 Proposition 3.2

Proof. From Remark 3.3 both $A(t)$ and $\varphi(t)$ belong to $\mathscr{C}^{\infty}(\Omega, \mathbb{R})$, hence they belong to $\mathscr{C}^{2}(\Omega, \mathbb{R})$ in particular. The first line in equation (3.1) holds true. The left hand quantity on the second line of equation (3.1) is $0\ \forall t \in \Omega$, naturally way smaller than 1. It can be noted that the sign of the derivatives is set all over Ω by the sings of β and γ. For the last equation to verify:

$$\left|\frac{\ddot{\varphi}(t)}{\dot{\varphi}^{2}(t)}\right| = \left|\frac{\frac{\beta}{\gamma}e^{t/\gamma}}{(\beta e^{t/\gamma})^{2}}\right| = \left|\frac{\frac{\beta}{\gamma}e^{t/\gamma}}{\beta^{2}(e^{t/\gamma})^{2}}\right| = \left|\frac{1}{\beta\gamma}\frac{e^{t/\gamma}}{(e^{t/\gamma})^{2}}\right| = \frac{1}{|\beta\gamma|}e^{-t/\gamma} \quad \forall t \in \Omega$$

the last equality motivated by the fact that $e^{-t/\gamma}$ is always greater than 0. This last quantity has to be way smaller than one everywhere in Ω, which is compact. Being an exponential the function under analysis is monotonic and the following holds:

$$\max_{t\in\Omega}\left\{\frac{1}{|\beta\gamma|}e^{-t/\gamma}\right\} = \left\|\frac{1}{|\beta\gamma|}e^{-t/\gamma}\right\|_{\infty} = \frac{1}{|\beta\gamma|}e^{-t/\gamma}\Big|_{t=\widetilde{t}}, \quad \widetilde{t} \doteq \begin{cases} \inf\{\Omega\} & if\ \gamma > 0 \\ \sup\{\Omega\} & if\ \gamma < 0 \end{cases}$$

Hence, a condition so that the quantity evaluated at $\widetilde{t}$ is way smaller than 1 has to be found:

$$\frac{1}{|\beta\gamma|}e^{-t/\gamma}\Big|_{t=\widetilde{t}} = \frac{1}{|\beta\gamma|}e^{-\widetilde{t}/\gamma} \ll 1 \Leftrightarrow e^{-\widetilde{t}/\gamma} \ll |\beta\gamma| = |\beta|\,|\gamma|$$

$$\Longrightarrow \left|\frac{\ddot{\varphi}(t)}{\dot{\varphi}^{2}(t)}\right| \ll 1 \Leftrightarrow |\beta| \gg \frac{1}{|\gamma|}e^{-\widetilde{t}/\gamma} = \left|\frac{1}{\gamma}e^{-\widetilde{t}/\gamma}\right| \doteq \epsilon$$

□

A.2 Derivation of equation (3.11)

Manipulating the oscillatory integral and integrating by parts:

$$\int_{\Omega} b(t)\, e^{j\psi(t)}dt = \int_{\Omega} b(t)\frac{j\dot{\psi}(t)}{j\dot{\psi}(t)}e^{j\psi(t)}dt = \int_{\Omega}\frac{b(t)}{j\dot{\psi}(t)}\overbrace{j\dot{\psi}(t)\, e^{j\psi(t)}}^{\frac{d}{dt}e^{j\psi(t)}}dt =$$

$$= \left[\frac{b(t)}{j\dot{\psi}(t)}e^{j\psi(t)}\right]_{\partial\Omega} - \frac{1}{j}\int_{\Omega}\frac{\dot{b}(t)\,\dot{\psi}(t) - b(t)\,\ddot{\psi}(t)}{\left(\dot{\psi}(t)\right)^{2}}e^{j\psi(t)}dt =$$

$$= \left[\frac{b(t)}{j\dot{\psi}(t)}e^{j\psi(t)}\right]_{\partial\Omega} + j\int_{\Omega}\left[\frac{\dot{b}(t)}{\dot{\psi}(t)} - \frac{b(t)\,\ddot{\psi}(t)}{\left(\dot{\psi}(t)\right)^{2}}\right]e^{j\psi(t)}dt =$$

$$= \left[\frac{b(t)}{j\dot{\psi}(t)}e^{j\psi(t)}\right]_{\partial\Omega} + j\int_{\Omega} b(t)\left[\frac{\dot{b}(t)}{b(t)\dot{\psi}(t)} - \frac{\ddot{\psi}(t)}{\left(\dot{\psi}(t)\right)^2}\right]e^{j\psi(t)}dt \doteq I$$

where the definition of $b(t)$ and $\dot{\psi}(t) \neq 0$ if $f \leq 0$ prevent meaningless fractions. Using the triangular inequality and bounding the integrals:

$$|I| \leq \left|\left[\frac{b(t)}{j\dot{\psi}(t)}e^{j\psi(t)}\right]_{\partial\Omega}\right| + \left|j\int_{\Omega} b(t)\left[\frac{\dot{b}(t)}{b(t)\dot{\psi}(t)} - \frac{\ddot{\psi}(t)}{\left(\dot{\psi}(t)\right)^2}\right]e^{j\psi(t)}dt\right| \leq$$

$$\leq \left|\underbrace{\frac{b(t)}{j\dot{\psi}(t)}e^{j\psi(t)}\Big|_{\sup\{\Omega\}}}_{\doteq\lambda} - \underbrace{\frac{b(t)}{j\dot{\psi}(t)}e^{j\psi(t)}\Big|_{\inf\{\Omega\}}}_{\doteq\mu}\right| +$$

$$+ \left|\int_{\Omega} b(t)\frac{\dot{b}(t)}{b(t)\dot{\psi}(t)}e^{j\psi(t)}dt\right| + \left|\int_{\Omega} b(t)\frac{\ddot{\psi}(t)}{\left(\dot{\psi}(t)\right)^2}e^{j\psi(t)}dt\right| \leq$$

$$\leq |\lambda| + |\mu| + \left|\int_{\Omega} b(t)\max_{t\in\Omega}\left\{\frac{\dot{b}(t)}{b(t)\dot{\psi}(t)}\right\}e^{j\psi(t)}dt\right| + \left|\int_{\Omega} b(t)\max_{t\in\Omega}\left\{\frac{\ddot{\psi}(t)}{\left(\dot{\psi}(t)\right)^2}\right\}e^{j\psi(t)}dt\right| =$$

$$= |\lambda| + |\mu| + \left\|\frac{\dot{b}(t)}{b(t)\dot{\psi}(t)}\right\|_{\infty}\left|\int_{\Omega} b(t)\,e^{j\psi(t)}dt\right| + \left\|\frac{\ddot{\psi}(t)}{\left(\dot{\psi}(t)\right)^2}\right\|_{\infty}\left|\int_{\Omega} b(t)\,e^{j\psi(t)}dt\right| \leq$$

$$\leq |\lambda| + |\mu| + \left(\left\|\frac{\dot{b}(t)}{b(t)\dot{\psi}(t)}\right\|_{\infty} + \left\|\frac{\ddot{\psi}(t)}{\left(\dot{\psi}(t)\right)^2}\right\|_{\infty}\right)\int_{\Omega}|b(t)|\,dt =$$

$$= |\lambda| + |\mu| + \left(\left\|\frac{\dot{b}(t)}{b(t)\dot{\psi}(t)}\right\|_{\infty} + \left\|\frac{\ddot{\psi}(t)}{\left(\dot{\psi}(t)\right)^2}\right\|_{\infty}\right)\|b(t)\|_1$$

where all the $\|\bullet\|_{\infty}$ exist being all the arguments continuous on a compact. For the same reason all the terms are finite and $b(t)$ is summable, belonging then to $\mathscr{L}^1(\Omega, \mathbb{R})$. Hence:

$$|I| \leq \left|\frac{b(\sup\{\Omega\})}{\dot{\psi}(\sup\{\Omega\})}\right| + \left|\frac{b(\inf\{\Omega\})}{\dot{\psi}(\inf\{\Omega\})}\right| + \left(\left\|\frac{\dot{b}(t)}{b(t)\dot{\psi}(t)}\right\|_{\infty} + \left\|\frac{\ddot{\psi}(t)}{\left(\dot{\psi}(t)\right)^2}\right\|_{\infty}\right)\|b(t)\|_1$$

A.3 Lemma 3.6

Proof. Consider $s(t)$:

$$s(t) = \alpha\exp\left(j\left[\beta\gamma e^{t/\gamma} - \delta\right]\right) = A(t)\,e^{j\varphi(t)} \quad \forall t \in \Omega$$

where

$$A(t) = \alpha \quad \forall t \in \Omega$$
$$\varphi(t) = \beta\gamma e^{t/\gamma} - \delta \quad \forall t \in \Omega$$

Hence:

$$|s(t)|^2 = |A(t)|^2 = \alpha^2 \quad \forall t \in \Omega$$

so, since Ω is compact:

$$\int_\Omega |s(t)|^2 \, dt = \alpha^2 \cdot m(\Omega) < +\infty$$

being the integral finite, $s(t)$ belongs to $\mathscr{L}^2(\Omega, \mathbb{C})$.

Consider $S(f)$ by its definition and equations (3.7), (3.8) and (3.10). If $f > 0$:

$$S(f) = B(f) e^{j\Psi(f)} \implies |S(f)| = |B(f)| \leq \left| \frac{A(t_f)}{\sqrt{|\ddot{\varphi}(t_f)|}} \right| + |R| \leq \frac{\alpha}{\sqrt{\left|\frac{\beta}{\gamma} e^{t_f/\gamma}\right|}} + |I| \sup_{t\in\Omega} \{Q(t, f)\}$$

$A(t)$ and $\varphi(t)$ belong to $\mathscr{C}^\infty(\Omega, \mathbb{R})$. Being Ω compact and considering that for the Chirp under analysis $Q(t, f)$ is limited over Ω (see Proposition A.1) it is then possible to conclude that a positive real finite number M exists such that

$$\left| \frac{A(t_f)}{\sqrt{|\ddot{\varphi}(t_f)|}} \right| + |R| \leq M \quad \forall t_f \in \Omega \longleftrightarrow \forall f \in \Phi^+$$

If $f \leq 0$ $S(f)$ is bounded by equation (3.11), which is limited. Calling M' the peak of the bound:

$$\int_\Phi |S(f)|^2 \, df \leq \max\{M, M'\} \cdot m(\Phi) < +\infty$$

and so $S(f)$ belongs to $\mathscr{L}^2(\Phi, \mathbb{C})$. □

A.4 Other results

Proposition A.1. *Let $s(t)$ be an exponential chirp as in Definition 3.4. Then, for each fixed $\widetilde{f} \in \Phi^+$ the function $Q\left(t, \widetilde{f}\right) : \Omega \longrightarrow \mathbb{R}^+$ defined by equation (3.9) is limited in Ω.*

Proof. From the triangle inequality:

$$Q\left(t, \widetilde{f}\right) \doteq 5\sqrt{2\left|\ddot{\Psi}\left(t_{\widetilde{f}}\right)\right|} \left| \frac{b(t)}{b\left(t_{\widetilde{f}}\right)} \frac{\sqrt{\Psi(t)}}{\dot{\Psi}(t)} \right| \left| \frac{\ddot{b}(t)}{b\left(t_{\widetilde{f}}\right)} \frac{\Psi(t)}{\dot{\Psi}^2(t)} + \frac{3}{2} \frac{\dot{b}(t)}{b(t)\dot{\Psi}(t)} \left(1 - \frac{\Psi(t)\ddot{\Psi}(t)}{\dot{\Psi}^2(t)}\right) + \right.$$
$$\left. + \left[3\Psi(t) \left(\frac{\ddot{\Psi}(t)}{\dot{\Psi}^2(t)} \right)^2 - \frac{3}{2} \frac{\ddot{\Psi}(t)}{\dot{\Psi}^2(t)} - \frac{\Psi(t)\dddot{\Psi}(t)}{\dot{\Psi}^3(t)} \right] \right|$$

$$\leq 5\sqrt{2\left|\ddot{\Psi}\left(t_{\widetilde{f}}\right)\right|}\left|\frac{b\left(t\right)}{b\left(t_{\widetilde{f}}\right)}\frac{\sqrt{\Psi\left(t\right)}}{\dot{\Psi}\left(t\right)}\right|\left\{\left|\frac{\ddot{b}\left(t\right)}{b\left(t_{\widetilde{f}}\right)}\frac{\Psi\left(t\right)}{\dot{\Psi}^{2}\left(t\right)}\right|+\left|\frac{3}{2}\frac{\dot{b}\left(t\right)}{b\left(t\right)\dot{\Psi}\left(t\right)}\right|+\right.$$
$$\left.+\left|\frac{3}{2}\frac{\dot{b}\left(t\right)}{b\left(t\right)\dot{\Psi}\left(t\right)}\frac{\Psi\left(t\right)\ddot{\Psi}\left(t\right)}{\dot{\Psi}^{2}\left(t\right)}\right|+\left|3\Psi\left(t\right)\left(\frac{\ddot{\Psi}\left(t\right)}{\dot{\Psi}^{2}\left(t\right)}\right)^{2}\right|+\left|\frac{3}{2}\frac{\ddot{\Psi}\left(t\right)}{\dot{\Psi}^{2}\left(t\right)}\right|+\left|\frac{\Psi\left(t\right)\dddot{\Psi}\left(t\right)}{\dot{\Psi}^{3}\left(t\right)}\right|\right\}$$

In Appendix A.1 it is showed that $A\left(t\right)$ and $\varphi\left(t\right)$ are of $\mathscr{C}^{\infty}\left(\Omega,\,\mathbb{R}\right)$. Hence, $b\left(t\right)$, $\psi\left(t\right)$ and $\Psi\left(t\right)$ belong to $\mathscr{C}^{\infty}\left(\Omega,\,\mathbb{R}\right)$ as well. Since Ω is compact $b\left(t\right)$ and $\Psi\left(t\right)$ are limited in Ω with all their derivatives. All the operations performed on these functions are algebraic and produce functions still in $\mathscr{C}^{\infty}\left(\Omega,\,\mathbb{R}\right)$. Moreover, Definition 3.1 ensures that none of the denominators can approach 0 in Φ^{+} and all the terms are then limited on Ω, making real positive numbers that bound each of the different magnitude terms in the equation above to exist. Calling M their maximum:

$$Q\left(t,\,\widetilde{f}\right)\leq M\quad\forall t\in\Omega$$

□

Remark A.2. It can be noted that, being $A\left(t\right)=\alpha\;\forall t\in\Omega$ then $\dot{b}\left(t\right)=\ddot{b}\left(t\right)=0\;\forall t\in\Omega$. Also:

$$\dot{\psi}\left(t\right)=\dot{\varphi}\left(t\right)-2\pi f=\frac{\beta\gamma}{\gamma}e^{t/\gamma}-2\pi f\quad\forall t\in\Omega$$
$$\ddot{\psi}\left(t\right)=\ddot{\varphi}\left(t\right)=\frac{\beta}{\gamma}e^{t/\gamma}=\frac{1}{\gamma}\dot{\varphi}\left(t\right)\quad\forall t\in\Omega$$
$$\dddot{\psi}\left(t\right)=\frac{\beta}{\gamma^{2}}e^{t/\gamma}=\frac{1}{\gamma}\ddot{\varphi}\left(t\right)=\frac{1}{\gamma}\ddot{\psi}\left(t\right)\quad\forall t\in\Omega$$

So, the $Q\left(t;t_{f}\right)$ expression is simplified in this way:

$$Q\left(t;t_{f}\right)=5\sqrt{2\frac{\left|\dot{\varphi}\left(t_{f}\right)\right|}{\gamma}}\left|\frac{\sqrt{\varphi\left(t\right)-2\pi ft}}{\dot{\varphi}\left(t\right)-2\pi f}\right|\left|3\left(\varphi\left(t\right)-2\pi ft\right)\left(\frac{\dot{\varphi}\left(t\right)}{\gamma\left[\dot{\varphi}\left(t\right)-2\pi f\right]^{2}}\right)^{2}+\right.$$
$$\left.-\frac{3}{2}\frac{\dot{\varphi}\left(t\right)}{\gamma\left[\dot{\varphi}\left(t\right)-2\pi f\right]^{2}}-\frac{\left(\varphi\left(t\right)-2\pi ft\right)\dot{\varphi}\left(t\right)}{\gamma^{2}\left[\dot{\varphi}\left(t\right)-2\pi f\right]^{3}}\right|$$

A.4.1 Distributional convolution

$x\left(t\right)$ and $\overleftarrow{x}\left(t\right)$ are ordinary signals. As such, they are locally summable and have compact support. Hence, they can define distributions over $\mathscr{L}^{2}\left(\mathbb{R},\,\mathbb{R}\right)$, as follows (Dore, 2015):

$$\left\langle x,\,f\right\rangle\doteq\int_{\mathbb{R}}x\left(t\right)f\left(t\right)dt\quad\forall f\in\mathscr{L}^{2}\left(\mathbb{R},\,\mathbb{R}\right)$$
$$\left\langle\overleftarrow{x},\,f\right\rangle\doteq\int_{\mathbb{R}}\overleftarrow{x}\left(t\right)f\left(t\right)dt\quad\forall f\in\mathscr{L}^{2}\left(\mathbb{R},\,\mathbb{R}\right)$$

for two distributions the distributional convolution can be defined as follows:

$$\langle x * \overleftarrow{x}, f \rangle \doteq \langle x(t), \langle \overleftarrow{x}(\varsigma), f(t+\varsigma) \rangle \rangle \quad \forall f \in \mathscr{L}^2(\mathbb{R}, \mathbb{R})$$

the meaning of equation (3.23) is then the following:

$$\langle x * \overleftarrow{x}, f \rangle = f(0) = \langle \delta, f \rangle \quad \forall f \in \mathscr{L}^2(\mathbb{R}, \mathbb{R})$$

The rigorous theory of distributional convolution can be found in Benedetto (1996) and, more succinctly, in Dore (2015).

A.5 Few useful matrix properties

Transposition	$(\mathbf{A}+\mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$
	$(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$
	$\mathbf{A}\mathbf{A}^T = \left(\mathbf{A}\mathbf{A}^T\right)^T$
Symmetry	if $\mathbf{A} = \mathbf{A}^T \wedge \exists \mathbf{A}^{-1}$ $\Rightarrow (\mathbf{A}^{-1}) = (\mathbf{A}^{-1})^T$

B Comments

B.1 `jack_iodelay` Multi-tone signal

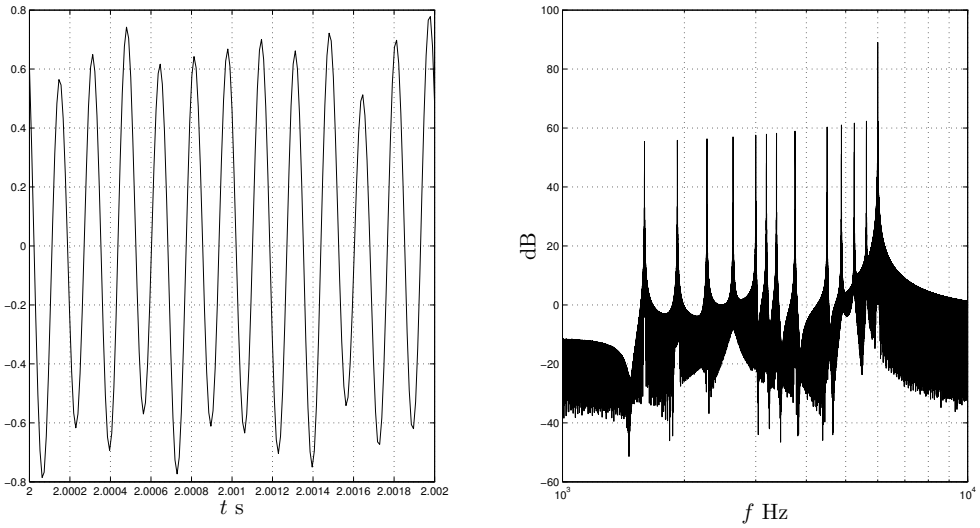


Figure B.1: Signal used by `jack_iodelay` to measure latency. Right panel: frequency domain representation. Left panel: time domain representation over 12 periods of the highest peak frequency (6 kHz). Time domain signal normalized to 1 absolute peak amplitude both in value and unit.

B.2 Superposition and homogeneity principles

Given a system, the inputs $x(t)$, $x_1(t)$ and $x_2(t)$ and the relative outputs $y(t)$, $y_1(t)$ and $y_2(t)$, which are real functions of real variables, the Superposition principles states:

$$x(t) = x_1(t) + x_2(t) \implies y(t) = y_1(t) + y_2(t)$$

The Homogeneity principle states:

$$x(t) = a_1 x_1(t) \implies y(t) = a_1 y_1(t) \quad \forall a_1 \in \mathbb{R}$$

since for a linear system the following must hold:

$$x(t) = a_1 x_1(t) + a_2 x_2(t) \implies y(t) = a_1 y_1(t) + a_2 y_2(t) \quad \forall a_1,\, a_2 \in \mathbb{R}$$

it is clear that the fall of one of the principles implies the fall of linearity.

B.3 Focusrite Scarlett 2i4 linearity

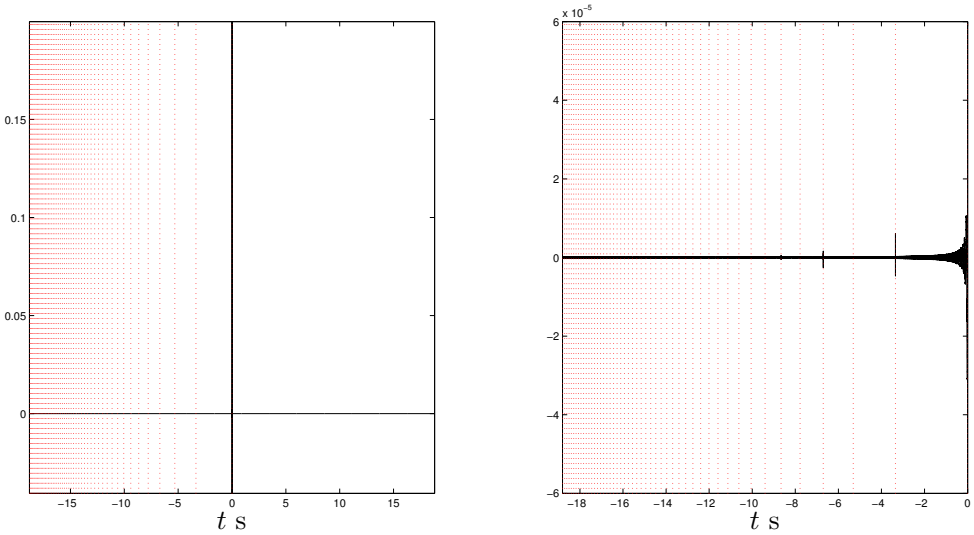


Figure B.2: An example of $\widetilde{h}(t)$ for the Interface, plotted over $\pm\triangle t_{49}$. Obtained with a recording level of $-5\,\mathrm{dB}$ to evidence nonlinear behavior. Arbitrary ordinate units as obtained by signals normalized to unit peak amplitude (both in value and unit). On the left: impulse response. On the right: particular of the negative time zone. Notice the difference of the right plot scale compared with figure 3.6.

A plot similar to figure 3.6 is reported in figure B.2. It is possible to notice that the higher order responses are very weak compared to the NTI ones. This measurement, performed at a recording level of $-5\,\mathrm{dB}$ ($8\,\mathrm{dB}$ higher than the recording level used during the NTI measurements), is intended to exploit the Interface non-linearity as possible and it justifies the decision operated in Section §3.3 to not remove the interface contribution

from the measurement. Moreover, being a (very) weakly nonlinear system, the Interface response should be not naively deconvolved. A better method would be the one devised by Novak, Maillou, Lotton, and Simon (2014) for the identification of nonlinear systems in series. However, this is not attempted as the project aims to explore fundamental methods.

B.4 `Lag_Detector.m` algorithm

The detection of the latency is performed by filtering an output $y(t_d)$ to partially remove noise, operate its numerical derivative and search for a jump discontinuity whose magnitude is outside 8 standard deviations of the noise amplitude statistic, which is isolated at the very beginning of the differentiated signal. In fact, in absence of noise, from equation (3.29):

$$x(t_d) \Longrightarrow y(t_d) = \begin{cases} 0 & \text{if } t_d \leq a + \Delta \\ w(t_d) & \text{if } a + 1 + \Delta \leq t_d \leq b + \Delta \\ 0 & \text{if } b + 1 + \Delta \leq t_d \leq c \end{cases}$$

where $w(t_d)$ is a non-linearly processed SES. The SES starts at $f_{start} = 12\,\text{Hz}$, a rather low value, increasing slowly for early times. Hence, along the samples

$$1, 2, \ldots, \text{round}\left(\frac{f_s}{2 f_{start}}\right)$$

$z(t_d)$ can be considered a sine. The slow variation in voltage is expected to be followed by the NTI system, meaning that $w(t_d)$ will have a sine like shape in these first samples. As such, its numerical derivative will be approximately a cosine. Hence:

$$\begin{cases} \dot{y}(t_d) = 0 & \text{if } t_d = a + \Delta \\ \dot{y}(t_d) \sim 1 & \text{if } t_d = a + 1 + \Delta \end{cases}$$

a jump discontinuity. In presence of noise statistics must be used to distinguish the signal jump from the jumps of the noise. For this reason comparison with an appropriate number of standard deviations is used. `Lag_Detector.m` proven to be successful, pointing out a 14 samples difference between the loops when operating preliminary measurement at high signal to noise ratio ($-5\,\text{dB}$ recording level).

C Project Proposal

Discrepancies between proposal and finished work are due to the early proposition date, at which the scope of the project was not completely clear.

Project Proposal

Title: **Nonlinear behaviour of audio systems**
Student: Stefano Tronci, s.tronci@edu.salford.ac.uk
Supervisor: Philip J. Duncan, p.j.duncan@salford.ac.uk

Introduction

Each piece of audio equipment exhibits nonlinear behaviour to some extent. Even if usually undesired in hi-fi equipment, non-linearity can be a source of artefacts that can be psychologically pleasant for the end user, especially in the case of musicians and their amplifiers/effects units. It is in facts evident that the major manufacturers of musical audio equipment developed during history a trademarked nonlinear behaviour in their units, a behaviour that constitutes the major fingerprint of their products and by which they are recognized by the musicians. Also, software developers put these characteristic non linearity under their scope to try to emulate famous and often rare pieces of equipment as software plugins. This project involves the study of the non-linearity of a sample of amplifiers based on different technologies, the classifications of their non-linearity and the study of ways to model them with appropriate DSP. Analysis will need to be performed with appropriate test signals and techniques. The outcome of the simulations can be used in subjective listening tests to investigate whether the models are able to create a convincing reproduction of the units and if not, what the causes could be.

Objectives

1. The amplifiers are measured and the features of their nonlinear behavior extrapolated and classified. Measurement can be undertaken setting the amplifiers at different representative states (for example, full power prior clipping, half power, and a quarter power may be used) and with an appropriate dummy load. The units under study will be a Sugden Musicmaster Pure Class A amplifier, a TSR-150X MOSFET amplifier and an Audio Note Kit1 300B Triode amplifier. Signals designed to exploit the system properties (like synchronized swept-sine) may be supplied by hardware generators or by software means (MATLAB).
2. The extrapolated features are used to model digitally the non-linearity with suitable software tools (MATLAB). A MISO (Multiple Input - Single Output) based nonlinear model can be used for system identification. Synthesized signals will have to be compared with the measured ones.
3. A double blind (or blind) ABX subjective test can be performed involving the measured units and their models or, more profitably, samples of materials as recorded from them, with the help of software packages like Lacinato ABX, aveX or MATLAB again. The test can be carried in the listening room and the results used to determine if the models are effective from the perceptual point of view and if there is a relation between perceived quality and kind of nonlinear process involved.

Equipment requirements

- Personal Computer with linear sound interface.
- MATLAB software.
- Sugden Musicmaster Pure Class A amplifier.

- TSR-150X MOSFET amplifier.
- Audio Note Kit1 300B Triode amplifier.
- Krohn-Hite 4400A ultra-low distortion function generator.
- Monitor loudspeakers (make and model to be determined).
- Salford University listening room.

Time plan

Activity	Week (1 = official start)																
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Project Plan – Paperwork	✓	✓															
Literature review	✓	✓	✓	✓													
Measurement phase				✓	✓	✓											
Nonlinear system modelling							✓	✓	✓								
Subjective testing										✓	✓	✓	✓				
Report writing							✓	✓	✓	✓	✓	✓	✓	✓			

Signature of Student: **Stefano Tronci**

Signature of Supervisor: **P. Duncan**

Date: **18/06/15**

D Link to code and data

All the data and code used for the project can be retrieved at https://drive.google.com/folderview?id=0ByDsYE5nBXrrVmNrR0VaOEstaGc&usp=sharing.

References

Bech, S., & Zacharov, N. (2006). *Perceptual Audio Evaluation - Theory, Method and Application.* John Wiley & Sons, Ltd. doi: 10.1002/9780470869253

Benedetto, J. (1996). *Harmonic Analysis and Applications.* Boca Raton: CRC Press.

Boashash, B. (1992). Estimating and interpreting the instantaneous frequency of a signal. I. Fundamentals. *Proceedings of the IEEE*, *80*(4), 520-538. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=135376 doi: 10.1109/5.135376

Boley, J., & Lester, M. (2009). Statistical Analysis of ABX Results Using Signal Detection Theory. In *Audio engineering society convention 127.* Retrieved from http://www.aes.org/e-lib/browse.cfm?elib=15022

Chassande-Mottin, E., & Flandrin, P. (1998). On the stationary phase approximation of chirp spectra. In *Time-frequency and time-scale analysis, 1998. proceedings of the ieee-sp international symposium on* (pp. 117–120). Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=721375 doi: 10.1109/TFSA.1998.721375

Chassande-Mottin, E., & Flandrin, P. (1999). On the Time - Frequency Detection of Chirps. *Applied and Computational Harmonic Analysis*, *6*(2), 252 - 281. Retrieved from http://www.apc.univ-paris7.fr/~ecm/99chirp/99chirp.pdf doi: http://dx.doi.org/10.1006/acha.1998.0254

Clark, D. (1982). High-Resolution Subjective Testing Using a Double-Blind Comparator. *J. Audio Eng. Soc*, *30*(5), 330–338. Retrieved from http://www.aes.org/e-lib/browse.cfm?elib=3839

Cohen, L. (2000). Instantaneous frequency and group delay of a filtered signal. *Journal of the Franklin Institute*, *337*(4), 329-346. Retrieved from http://www.sciencedirect.com/science/article/pii/S0016003200000399 doi: 10.1016/S0016-0032(00)00039-9

CrocoDuck. (2013). Lowlatency on ArchBang. Retrieved 02 August, 2015, from http://www.linuxmusicians.com/viewtopic.php?f=19&t=10837

Dobrucki, A. (2011). Nonlinear distortions in electroacoustic devices. *Archives of Acoustics*, *36*(2), 437-460. Retrieved from http://www.degruyter.com/view/j/aoa.2011.36.issue-2/v10168-011-0031-y/v10168-011-0031-y.xml doi: 10.2478/v10168-011-0031-y

Dore, G. (2015). *Convoluzione di distribuzioni.* Retrieved 08 September, 2015, from http://www.dm.unibo.it/~dore/Analisi_L-D/Convoluzione.pdf

Farina, A. (2000). Simultaneous Measurement of Impulse Response and Distortion with a Swept-Sine Technique. In *Audio engineering society convention 108.* Retrieved from http://www.aes.org/e-lib/browse.cfm?elib=10211

Farina, A., Bellini, A., & Armelloni, E. (2001). Not-Linear Convolution: A New Approach For The Auralization Of Distorting Systems. In *Audio engineering society convention 110.* Retrieved from http://www.aes.org/e-lib/browse.cfm?elib=9958

Focusrite. (2012). *Scarlett 2i4 User Guide* [User Manual]. Retrieved from

http://d3se566zfvnmhf.cloudfront.net/sites/default/files/focusrite/downloads/8174/scarlett2i4-user-guide.pdf

Geddes, E. R., & Lee, L. W. (2003a). Auditory Perception of Nonlinear Distortion. In *Audio engineering society convention 115.* Retrieved from http://www.aes.org/e-lib/browse.cfm?elib=12465

Geddes, E. R., & Lee, L. W. (2003b). Auditory Perception of Nonlinear Distortion - Theory. In *Audio engineering society convention 115.* Retrieved from http://www.aes.org/e-lib/browse.cfm?elib=12464

Gilardi, G. (2011). *Analisi matematica di base* (2nd ed.). McGraw-Hill Companies.

Harris, J. D. (1952). Remarks on the Determination of a Differential Threshold by the So Called ABX Technique. *The Journal of the Acoustical Society of America*, *24*(4), 417-417. Retrieved from http://scitation.aip.org/content/asa/journal/jasa/24/4/10.1121/1.1906915 doi: http://dx.doi.org/10.1121/1.1906915

Hassan, M. A., Coats, D., Gouda, K., Shin, Y., & Bayoumi, A. (2012). Analysis of nonlinear vibration-interaction using higher order spectra to diagnose aerospace system faults. In *Aerospace conference, 2012 ieee* (pp. 1–8). Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6187370 doi: 10.1109/AERO.2012.6187370

Hautus, M. J., & Meng, X. (2002). Decision strategies in the ABX (matching-to-sample) psychophysical task. *Perception and Psychophysics*, *64*(1), 89-105. Retrieved from http://link.springer.com/article/10.3758/BF03194559

International Telecommunication Union. (2003). *General methods for the subjective assessment of sound quality.* (ITU-R BS.1284-1). Geneva: ITU. Retrieved from http://www.itu.int/rec/R-REC-BS.1284-1-200312-I/en

International Telecommunication Union. (2015). *Methods for the subjective assessment of small impairments in audio systems.* (ITU-R BS.1116-3). Geneva: ITU. Retrieved from http://www.itu.int/rec/R-REC-BS.1116-3-201502-I/en

Kaplan, A. (2009). A Comparison of Three Methods for Calculating Confidence Intervals around D-Prime. *Ms., UCSC*. Retrieved from http://home.utah.edu/~u0703432/CV/phons09.pdf

Kemp, J. A. (2002). *Theoretical and experimental study of wave propagation in brass musical instruments* (Doctoral dissertation, University of Edinburgh, Edimburg). Retrieved from http://www.kempacoustics.com/thesis/

Knoblauch, K. (2014). psyphy: Functions for analyzing psychophysical data in R. Retrieved 03 September, 2015, from https://cran.r-project.org/web/packages/psyphy/index.html

Kuttruff, H. (2009). *Room Acoustics* (5th ed.). Abdingdon: Spon Press.

Macmillan, N. A., & Creelman, C. D. (2004). *Detection Theory: A User's Guide* (2nd ed.). Lawrence Erlbaum Associates. doi: 10.4324/9781410611147

Novak, A. (2009). *Identification of Nonlinear Systems in Acoustics* (Doctoral dissertation,

Université du Maine, Le Mans, France and Czech Technical University, Prague, Czech Republic). Retrieved from http://www.ant-novak.com/publications/Novak_Disertation.pdf

Novak, A. (2014). Nonlinear System Identification using Synchronized Swept-Sine Method. Retrieved 02 August, 2015, from http://www.ant-novak.com/Swept_Sine_Simulation.php

Novak, A., Maillou, B., Lotton, P., & Simon, L. (2014). Nonparametric Identification of Nonlinear Systems in Series. *Instrumentation and Measurement, IEEE Transactions on*, *63*(8), 2044-2051. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6744630

Novak, A., Simon, L., & Lotton, P. (2010a). Analysis, synthesis, and classification of nonlinear systems using synchronized swept-sine method for audio effects. *EURASIP Journal on Advances in Signal Processing*, *2010*(1). Retrieved from http://asp.eurasipjournals.com/content/2010/1/793816 doi: 10.1155/2010/793816

Novak, A., Simon, L., & Lotton, P. (2010b). Nonlinear system identification using exponential swept-sine signal. *IEEE Transactions on Instrumentation and Measurement*, *59*(8), 2220-2229. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5299278 doi: 10.1109/TIM.2009.2031836

Preis, D. (1984). Linear Distortions: Measurement, Methods and Audible Effects - A Survey of Existing Knowledge. In *Audio engineering society conference: 2nd international conference: The art and technology of recording.* Retrieved from http://www.aes.org/e-lib/browse.cfm?elib=11658

Sàez, N., Gavaldà, J., Ruiz, X., & Shevtsova, V. (2014). Detecting accelerometric nonlinearities in the international space station. *Acta Astronautica*, *103*, 16 - 25. Retrieved from http://www.sciencedirect.com/science/article/pii/S0094576514002240 doi: http://dx.doi.org/10.1016/j.actaastro.2014.06.025

Samson Technologies Corp. (2012). *SR950 User Manual* (1.2 ed.) [User Manual]. New York. Retrieved from http://www.samsontech.com/site_media/support/manuals/SR950_OM_EN_1.2.pdf

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, *31*(1), 137-149. Retrieved from http://dx.doi.org/10.3758/BF03207704 doi: 10.3758/BF03207704

Strang, G. (1986). *Introduction to applied mathematics*. Wellesley: Wellesley-Cambridge Press.

Vanderkooy, J. (1994). Aspects of MLS Measuring Systems. *J. Audio Eng. Soc*, *42*(4), 219–231. Retrieved from http://www.aes.org/e-lib/browse.cfm?elib=6951

Williams, J. (1995). *From The Jungles Of Paraguay (John Williams Plays Barrios)* [CD]. New York: Sony Classical.