Improvement in non-linear guitar loudspeaker sound reproduction

Thomas Schmitz
University of Liege
Department EECS, Intelsig group
Belgium, Liege campus du Sart-Tilman
Email: T.Schmitz@ulg.ac.be

Jean-Jacques Embrechts
University of Liege
Department EECS, Intelsig group
Belgium, Liege campus du Sart-Tilman
Email: jjembrechts@ulg.ac.be

***Abstract*—This paper proposes a study of the accuracy of a guitar amplifier loudspeaker simulation. The simulation is based on a non-linear convolution of a signal using Volterra kernels, which are measured in anechoic conditions with a sine-sweep technique. In this paper, we propose an evaluation of the method to minimise the cost in CPU load, while keeping the best performance in the sound reproduction. To assess the performance of the method, we measure errors between the simulated and real sounds. Human listening tests are moreover proposed in order to determine the minimum level of accuracy leading to unaudible differences with the real loudspeaker.**

I. Introduction

Unlike Hi-Fi loudspeakers, guitar loudspeakers do not have a flat frequency response. Indeed, they present a special response which defines the final "sound character" of a guitarist. These guitar loudspeakers are cumbersome, expensive and heavy. Therefore, it is interesting to simulate their behaviors (see Fig.1).

In our previous work [1], we described a method to simulate such a guitar loudspeaker using Volterra kernels. In theory, using Volterra series requires an infinite number of infinite length kernels. Therefore, for a practical implementation, a finite number of finite length kernels are necessary.

The purpose of this paper is to determine the level of accuracy (the number of kernels and their lengths) needed in our model to have a final sound reproduction close enough to the real sound of the loudspeaker. To evaluate the accuracy of the model, we measure the error between the real and simulated sounds (details are given in Section III-B). Moreover, we propose a series of listening tests to human listeners. Specifically, we consider the reproduction accurate when the listeners cannot distinguish the difference between the simulated and the real sound.

II. Basics of the method

Our model described in [1] is based on a black-box approach by means of the Volterra series. Pr. Farina [4] proposed to simplify the use of Volterra series by using a Hammerstein model, which is a system composed of a non-linear no-memory block followed by a linear memory block. Using this simplified model [4], we can write the output of the system y[n] as :

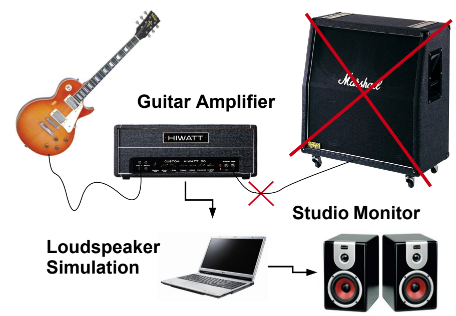

Fig. 1. The objective of this work is to make a good simulation of the guitar amplifier loudspeaker.

$$y[n] = \begin{bmatrix} x[n] \otimes h_1[n] + x[n]^2 \otimes h_2[n] + x[n]^3 \otimes h_3[n] \\ + x[n]^4 \otimes h_4[n] + ... + x[n]^m \otimes h_m[n] \end{bmatrix} \tag{1}$$

Where $x[n]$ is the input signal and $h_m[n]$ is the Volterra kernel of order m. The measurement method to obtain the Volterra kernels $h_m[n]$ is described in [1], [2] and [3].

III. Objective evaluation of the model

A. Validation of the Hammerstein model for the loudspeaker

The Volterra kernels characterise a time-invariant non-linear system independently of the input level. A way to verify that the loudspeaker can be modeled by a Hammerstein system is to compare the Volterra kernels measured with different input levels. Fig.2 and Fig.3 show the second and third order Volterra kernels respectively, measured with an exponential sine sweep input signal at different amplitude levels (22, 19, 16 and 13 Vrms). The RMS voltage is measured at the input terminals of the loudspeaker.

As expected, all curves representing the frequency dependence of the Volterra kernels have the same general shape, but some differences in amplitude are clearly illustrated, especially for the third order kernels. The loudspeaker can therefore not be fully associated with a Hammerstein model. However, at low level, these approximations seem to be acceptable. In section III-B, we show that the best accuracy is obtained when the level of the input signal used during the tests is close to the input level that was used to measure the Volterra kernels.

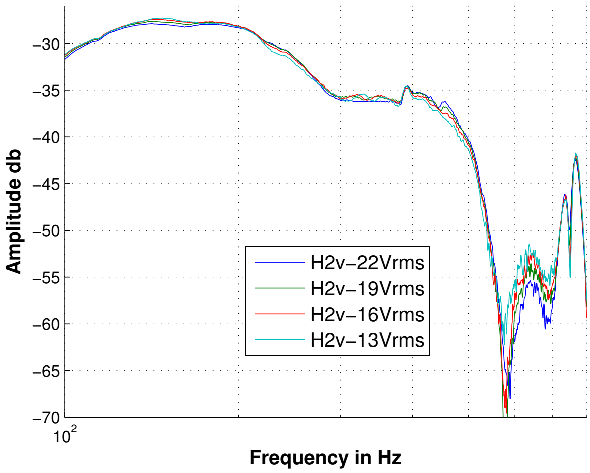

Fig. 2. Comparison of the second order Volterra kernels of the loudspeaker for several input levels. Frequency range: [100 800] Hz.

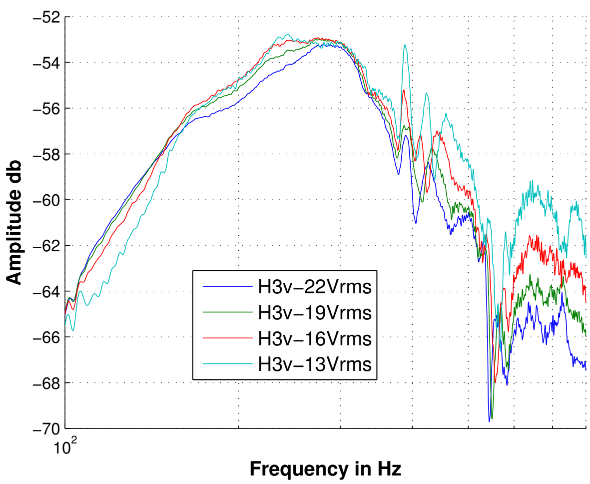

Fig. 3. Comparison of the third order Volterra kernels of the loudspeaker for several input levels. Frequency range: [100 800] Hz.

B. Comparison between simulated and real loudspeaker sounds

In this series of tests, two types of signals are sent to the loudspeaker:

- The first one is an exponential sine-sweep in the frequency range [20-20000] Hz, which is used to verify the correct amplitude of our simulated sound over all the audio band frequency.
- The second one is composed of several guitar's chords, to check the simulation in real conditions.

The experimental setup is shown in Fig.4: the loudspeaker is placed in an anechoic room, and the microphone giving the *real* sound signal is located in the same position than for the measurement of the Volterra kernels. Both the *real* and the *simulated* sound can be sent to the same loudspeaker for the listening tests.

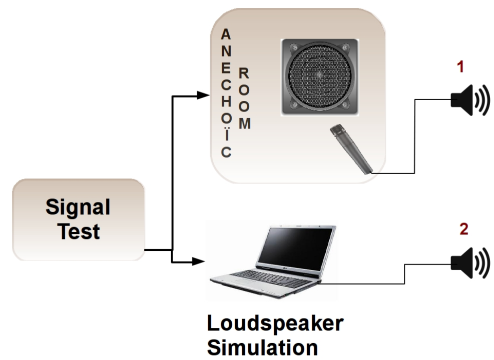

Fig. 4. Comparison between real loudspeaker and simulated loudspeaker.

1) Simulation of an exponential sine-sweep: Fig.5 compares the sound signal emitted by the loudspeaker (in blue) with the one that is generated by our simulation model (in green). The red curve is their difference. We can see that the mean error is below 1%.

This simulation has been made with the Volterra kernels until the 5th order. The Volterra kernels were measured with the same input level than during the test. As a comparison, Fig.6 presents the results of the simulation made with only the first order Volterra kernel. We can see that the mean error is bigger. The improvement can be quantified by comparing the mean absolute values of the error signals in both experiments:

$$Improvement = \frac{\sum |Error_1|}{\sum |Error_5|} = 3.7 \qquad (2)$$

where $Error_m$ is the error between the simulated and real signals if we use m Volterra kernels for the simulation.

It is also interesting to look at the errors that we obtain if we simulate high level signals with Volterra kernels measured at low level. Fig.7 shows the comparison of an exponential sine-sweep sent to the loudspeaker at 22Vrms and the simulated signal obtained with the same input level, but using the Volterra kernels measured at 13Vrms. In this case, the error is even bigger than in the previous case.

2) Simulation of guitar's chords: In order to test our simulation in real conditions, we sent a series of guitar chords to the loudspeaker and to our simulation model. The main problem to simulate a real sound is that the amplitude of the signal is time-varying. Fig.8, again, compares the *real* signal delivered by the loudspeaker and the *simulated* one, using the Volterra kernels measured at 22Vrms input signal. While in Fig.9, the Volterra kernels measured at 13Vrms are applied.

The simulation using Volterra kernels measured at 13 Vrms clearly gives better results. This can be explained by the fact

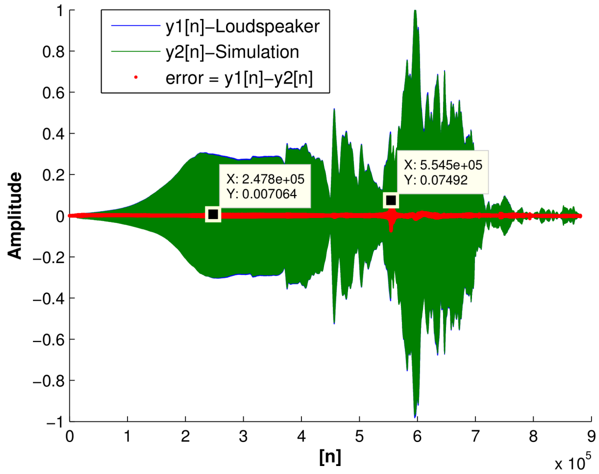

Fig. 5. Comparison of the signal test that pass through the loudspeaker and through the simulation using 5 Volterra kernels.

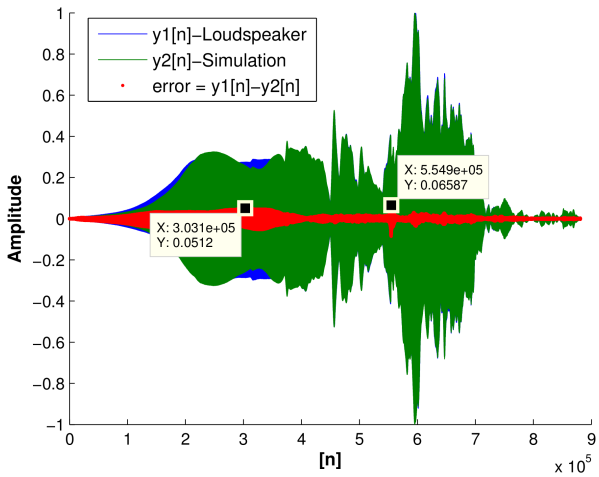

Fig. 6. Comparison of the signal test that pass through the loudspeaker and through the simulation using only the first Volterra kernel.

that the RMS level of the guitar's chords signal is 6.1 Vrms.

C. Accuracy vs CPU load

As first explained in the introduction, decreasing the CPU load during the simulation process can be reached by decreasing the number of Volterra kernels used in the convolutions and/or reducing the length of these kernels (in samples). However, we would not like to significantly deteriorate the quality (accuracy) of the simulation. The following SNR_{db} will be used as a mesure of accuracy:

$$SNR_{db} = 20\log\left[\frac{A_{lspeaker}}{A_{error}}\right] \tag{3}$$

Where:

- $A_{lspeaker}$ is the *RMS* value of the signal emitted by the loudspeaker,

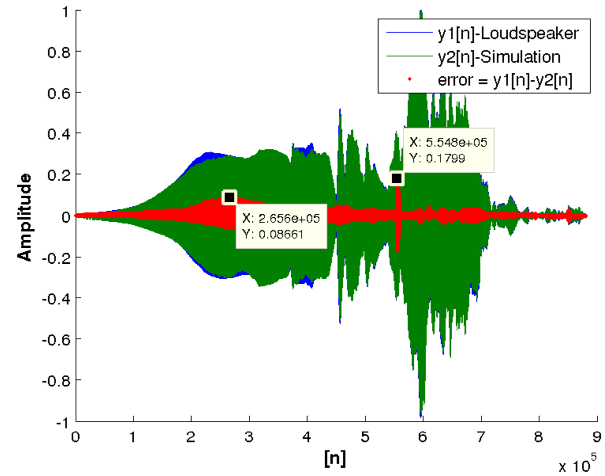

Fig. 7. Comparison of the signal test that pass through the loudspeaker and through the simulation using 5 Volterra kernels measured at 13Vrms.

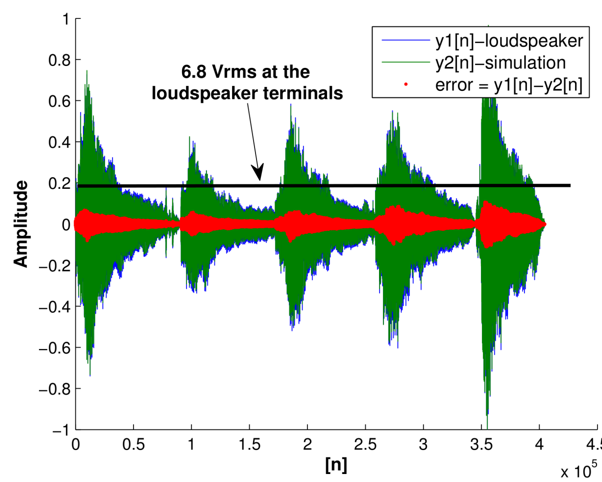

Fig. 8. Comparison of the *chords* signal test that pass through the loudspeaker and through the simulation when Volterra kernels are measured for an input level of 22 Vrms.

TABLE I
CHANGE IN SNR WITH NUMBER OF KERNELS VARIATION

Nb of kernels	**1**	**2**	**3**	**4**	**5**
SNR_{db}	22	23	30	31	34

- A_{error} is the *RMS* value of the error signal (difference between the *real* and *simulated* signals).

1) Influence of the number of Volterra kernels: the variation of the *SNR* is shown in Table I, as a function of the total number of Volterra kernels that are used in our simulations. This test has been implemented by comparing an exponential sine-sweep at 22 Vrms which passes through the loudspeaker and our simulation using Volterra kernels measured at 22 Vrms too.

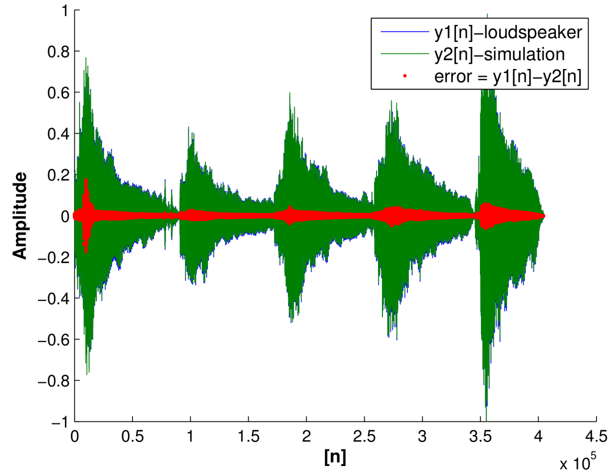

Fig. 9. Comparison of the *chords* signal test that pass through the loudspeaker and through the simulation when Volterra kernels are measured for an input level of 13 Vrms.

TABLE II
CHANGE IN SNR WITH LENGTH OF VOLTERRA KERNELS

length of $h_m[n]$	**test 1**	**test 2**	**test 3**	**test 4**	**test 5**
Size $h_1[n]$	83000	10375	10375	5188	1266
Size $h_2[n]$	25000	3125	3750	1563	469
Size $h_3[n]$	16500	2063	2063	1031	281
Size $h_4[n]$	10000	1250	1250	625	156
Size $h_5[n]$	10000	1250	1250	625	188
SNR_{db}	34	34	34	33	31

2) Influence of the length of the Volterra kernels: The Volterra kernels are theoretically infinite in time and they must therefore be clipped in our model for practical convolutions. Table II shows the influence of the Volterra kernels length on the SNR value. Values printed in red, green and black respectively decrease, increase, or stay unchanged between consecutive tests.

We can conclude that the length of the Volterra kernels does not really influence the accuracy of the model, provided that it is kept above the values in test 4 in Table II.

IV. Subjective Evaluation of the Model: Listening Tests

In order to determine whether the simulation can fool a human listener or not, we designed listening tests based on three types of sound samples:

- The first one (type 0) is the sound that comes from the loudspeaker.
- The second one (type 1) is the sound that comes out from our simulation using 5 Volterra kernels.
- Finally, the third one (type 2) is the sound that comes from our simulation using only the first Volterra kernel.

We have made five tests with five different guitar chords. In each test, we have put a sample of type 0 followed by a sample of type 0 or 1 or 2 (the samples were chosen randomly). Next we ask to listeners to say if the first sample is the same than the second one. At the end, a total of 120 tests have been made. This kind of blind test give the following indications :

- Comparison between sample 0 and sample 0 (*0 vs 0*) give us an indication on the reliability of the listening tester.
- Comparison between sample 0 and sample 1 (*0 vs 1*) give us information about the accuracy of our simulation.
- Comparison between sample 0 and sample 2 (*0 vs 2*) give us information about the necessity to use more than one kernels for our simulation.

The results are given by Fig.10. The *0 vs 1* and *0 vs 2* columns show the percentage of tests where people cannot distinguish the *simulated* sound from the *real* one. The *0 vs 0* column show the percentage of people who find a difference where there were none. We can conclude that the difference between the simulated sound and the real one is almost unaudible even for the simulation using only the first Volterra kernel.

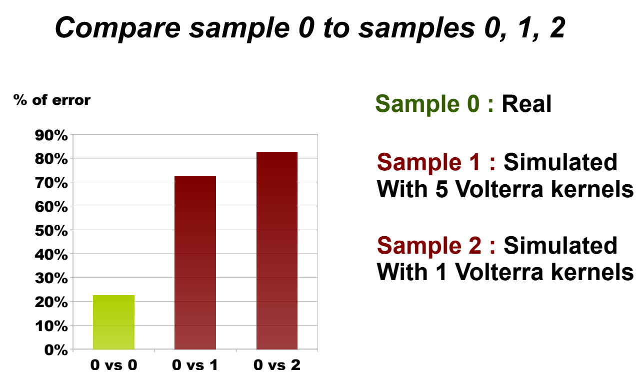

Fig. 10. Blind test results : 82.5% of tested listeners were not able to distinguish a simulation of type 2 (5 kernels), 72.5% of tested listeners were not able to distinguish a simulation of type 1 (one kernel) and 22.5% of tested listeners hear a difference where there were none.

V. Conclusion

For the tested level, the Volterra series approach give us enough accuracy to have unaudibles differences between the *real* and *simulated* sound. Moreover the simulation is usable in real time since we have shown that we could use short impulse response without decrease the accuracy of the model and that the listening tests give almost the same result when we use 1 or 5 Volterra kernels in our simulation.

References

[1] T. Schmitz and J. J. Embrechts, *Nonlinear Guitar Loudspeaker Simulation*, Audio Engineering Society Convention 134, 2013.

[2] A. Novák et al, *A new method for identification of nonlinear system using MISO model with Swept-sine technique : Application to loudspeaker analysis*, Audio Engineering Society Convention 124, 2008.

[3] L. Tronchin, *The Emulation of Nonlinear Time-Invariant Audio Systems with Memory by Means of Volterra Series*, Journal of the Audio Engineering Society 60(12) pp 984–996, 2013.

[4] A. Farina, *Non-linear convolution: A new approach for the auralization of distorting systems*, Audio Engineering Society Convention 110, 2001.

[5] A. Farina, *Advancements in impulse response measurements by sine sweeps*, Audio Engineering Society Convention 122, 2007.