AUDIO ENGINEERING SOCIETY
Italian Section

ANNUAL MEETING 2005

Paper: 05014

Como, 3-5 November
Politecnico di MILANO

EMULATION OF NOT-LINEAR, TIME-VARIANT DEVICES BY THE CONVOLUTION TECHNIQUE

ANGELO FARINA[1], ENRICO ARMELLONI[2]

[1] *Industrial Engineering Department, Università di Parma, Via delle Scienze, 43100 Parma, Italy*
angelo.farina@unipr.it

[2] *Industrial Engineering Department, Università di Parma, Via delle Scienze, 43100 Parma, Italy*
enrico.armelloni@unipr.it

In recent years, the convolution technique has been used (and abused) for emulating the reverberation of acoustical spaces and the functionality of hardware devices (loudspeakers, microphones, preamps, etc.). Now it is also being employed for attempting the emulation of strongly not linear and not time-invariant components, such as compressors, limiters, level maximizers, dynamic processors, automatic gain controllers.

As the results obtained by purely linear convolution for emulating the behaviour of not-linear devices revealed to be completely unsatisfactory, two methods for obtaining not-linear convolution were pioneered by the authors: Impulse Response switching [1,2] and diagonal Volterra multiple convolution [3]. This paper describes these techniques, and provides an evaluation of their performances both in case of memoryless notlinear devices and in case of devices which show significant memory effects.

INTRODUCTION

Linear convolution of a signal with a system's impulse response can be a faithful emulation of the behaviour of the system only under strict conditions: the system must be linear and time-invariant. Once these conditions are not anymore met, although the impulse response of the system can still be measured employing proper techniques [4], in general the emulation of the system's behaviour obtained by the linear convolution approach is quite bad.

This is clearly demonstrated by the poor results obtainable with convolution-based software plugins for replicating the sonic effect of hardware devices which have been "sampled" by measuring their impulse response.

It is important to separate here the effects of "memoryless" nonlinearity, which are typical of those systems which vary their transfer function in dependence of the instantaneous amplitude of the input signal, for example a tube amplifier, from the effects of circuits which change their transfer function in dependence of one or more "status variables", which are slowly dependent on the history of the input signal over a significant time. Most compressors, limiters, dynamic processors and automatic gain controllers fall in this second category.

Whilst for the first category of devices with "memoryless" nonlinearity it is still possible to emulate quite faithfully their behaviour by modified convolution schemes, for the second category of devices a new approach is required.

1 STATE OF THE ART

This paper first reviews the two basic techniques currently available for the emulation by convolution of

devices with memoryless nonlinearity: the Impulse Response switching and the diagonal Volterra kernels. The authors pioneered both of these techniques [1,2,3].
The first products are appearing on the market based on the first of these approaches (for example, it must be pointed out here the availability of the Focus Rite Liquid Channel processor, which employs the IR switching technique providing excellent results in the emulation of devices characterized by memoryless distortion, such as tube microphone preamps – this is the first widely-diffused system based on Mike Kemp's "dynamic convolution" patent - WO 98/07141).
However, these products fail when attempting to emulate devices exhibiting significant not-linear memory – a one-by-one comparison with the ABX method immediately shows that the emulated signal is significantly different from the one processed by the real device.
Consequently, it is necessary to search for a method to capture and replicate the time-dependent behaviour of these nonlinear devices. One solution, of course, is to employ complete (not diagonal) Volterra kernels, but this method is computationally too expensive for being applied beyond second-order nonlinearity. Furthermore, the method does not address inharmonic (acausal) distortion, rattling, clipping and other real-world behaviours.
An offline emulation of the performance of the two currently available notlinear convolution schemes for devices characterized by relevant notlinearity have been performed, and blind ABX listening tests were performed on the results.
Generally speaking, the results show that both of these new techniques allow for results significantly better than traditional linear convolution ones, although none of them provides results 100% undistinguishable from the real notlinear device. This means that further research is required for achieving even more faithful reproduction of nonlinear devices, exhibiting also some memory, extending the currently available not-linear convolution methods.

2 THE IMPULSE RESPONSE SWITCHING TECHNIQUE

This method is based on a very simple approach: depending on the amplitude of the instantaneous signal entering the convolver, a different IR is chosen and applied.
This approach was employed by one of the authors for the first time at the beginning of 1998 [1]; at that time Mike Kemp was already employing this technique since almost two years, although he publicly disclosed his method only in 1999 [5]. Of course, each of the two developers was not aware of the work of the other.
Usually it is not necessary to measure a separate IR for each possible quantized value of the input of a digital processing system: for example, in a 16-bit system one should theoretically measure 65536 different impulse responses, one for each of the possible amplitudes of a digital sample (ranging from -32767 to + 32768). In practice, the not-linear behaviour of a time-invariant device can be satisfactorily described with a few hundredths of IRs, usually logarithmically spaced along a range of 30 to 60 dB. For example, 240 IRs starting from 0 dB Full Scale and decreasing by steps of 0.25 dB provide a very smooth approximation of the notlinear behaviour of a system such as a valve preamplifier.
Of course, it is necessary to measure the system's impulse response employing a test signal played at each of these 240 amplitudes, and store the resulting 240 IRs in the memory of the convolver.
During the system emulation, as a new sample comes in, based on its amplitude the convolver chooses the proper IR to be applied to the sample. So the choice of the IR simply requires to change the value of a pointer, based on a logarithmic function of the sample value. This adds very little overhead to the computational power requirement, for a system performing traditional multiply-and-sum convolution on each sample being processed.
This method is not suitable for fast-convolution schemes, which require to process the sound stream in blocks of several samples. Consequently, the length of the impulse response is limited to a few thousands of samples, whilst modern fast convolution schemes allows for IRs having a length of millions of samples.
However, usually the IR-switching technique is applied to the emulation of electronic devices, which have very short "memory", and consequently can be described by impulse responses of a few hundredths samples.
The method developed by the authors and by Mike Kemp differ about the measurement technique: while the authors did employ a series of MLS (Maximum Length Sequence) test signals of different amplitude, Mike Kemp employed a series of step functions.
Fig. 1 shows the test signal obtained by the series of MLS of varying amplitude, whilst fig.2 (taken from [5]) shows the "impulse train" of progressively-reducing amplitude.

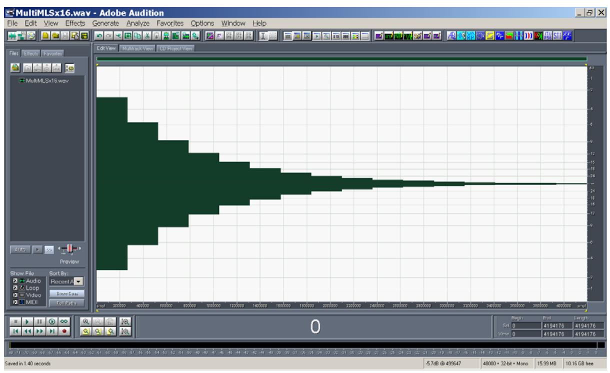

Figure 1 Multi-MLS signal employed by the authors

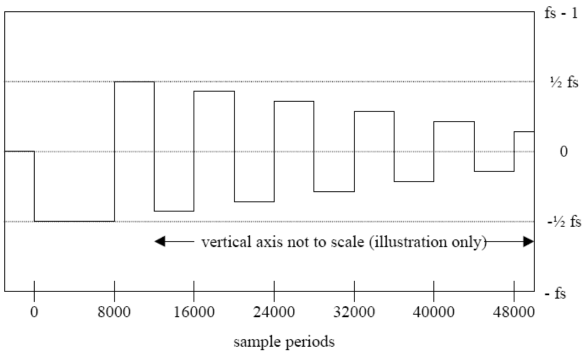

Figure 2 test signal employed by Mike Kemp

The result of the measurement of the impulse response of a strongly-notlinear device (a guitar effect) are shown here, employing 16 MLS sequences of amplitude decaying by 3 dB each.
The guitar effect has been emulated employing the "distortion" plugin of Adobe Audition, which creates strongly not-linear, but memoryless, distortion, as shown here:

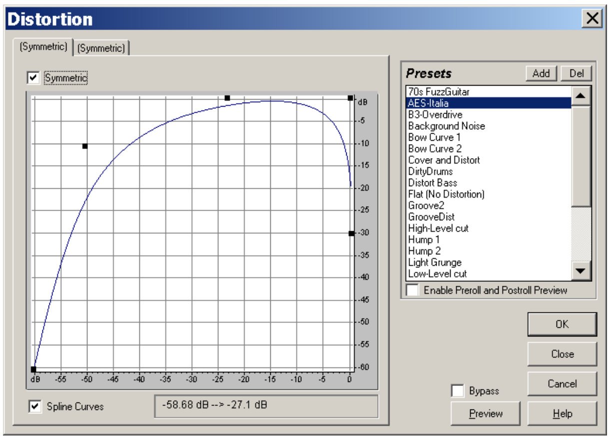

Fig. 3 – not-linear distortion curve

Applying this effect to the test signal shown in fig.1, the following waveform has been obtained:

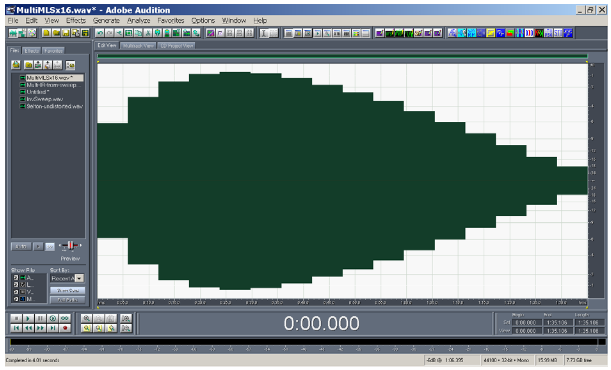

Figure 4 Multi-MLS signal after distortion

Now we deconvolve the set of 16 impulse responses, and we obtain what follows:

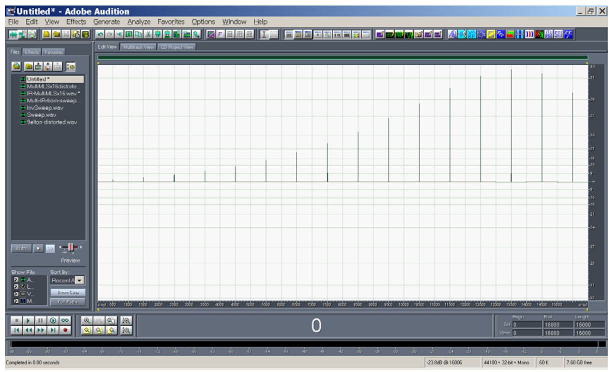

Figure 5 16 impulse response at 16 different amplitudes of the test signal

Figure 5 shows that the gain of the device was substantially increasing when the level of the signal was reducing: this means that this device acts roughly as a gain compressor. Of course, sampling the gain curve with just 16 steps spaced 3 dB each can be considered a bit too coarse, and in fact usually this technique needs at least 256 impulse responses for providing reasonably smooth level change.
Going from measurement to implementation, things become quite simple: each sample being processed is evaluated against the table of amplitudes associated with each impulse response. The proper IR is chosen (the one measured with a test signal amplitude closer to the amplitude of the sample being processed), so that each samples is routed to a different filter. As 16 filters are really too coarse, an interpolation is being performed between the two IRs having amplitude just below and above the amplitude of the current sample.

3 THE DIAGONAL VOLTERRA KERNELS TECHNIQUE

The following picture describes the flow diagram of a system obtained by a distorting transducer (memoryless distortion) driving a subsequent linear system with memory:

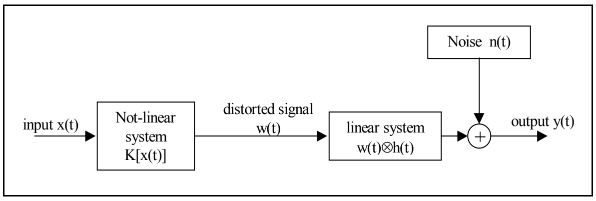

Figure 6. Flow diagram of the complex system

Neglecting the noise, the transfer function of this system can be described, in general, by means of a Volterra series expansion:

$$y(n) = \sum_{i_1=0}^{M-1} h_1(i_1)\cdot x(n-i_1) + \sum_{i_1=0}^{M-1}\sum_{i_2=0}^{M-1} h_2(i_1,i_2)\cdot x(n-i_1)\cdot x(n-i_2) + \sum_{i_1=0}^{M-1}\sum_{i_2=0}^{M-1}\sum_{i_3=0}^{M-1} h_3(i_1,i_2,i_3)\cdot x(n-i_1)\cdot x(n-i_2)\cdot x(n-i_3) + ... \quad (1)$$

This general formulation also takes into account non-linear behaviour with memory (i.e., hystheresis), whilst in our case we are supposing that memory effects can be present only in the linear part of the two-block system illustrated in fig. 6. This means that, for orders higher than 1, the Volterra kernels h_2, h_3,… instead of being large multidimensional matrices reduce only to the terms on the diagonal, and thus can be represented by simple linear vectors having the same size as the first-order (linear) kernel. Under these hypotheses, eqn. 1 reduces to:

$$y(n) = \sum_{i=0}^{M-1} h_1(i)\cdot x(n-i) + \sum_{i=0}^{M-1} h_2(i)\cdot x^2(n-i) + \sum_{i=0}^{M-1} h_3(i)\cdot x^3(n-i) + ... \quad (2)$$

If these simplified (one-dimensional) Volterra kernels are known, eqn. 2 makes it possible to reconstruct the output signal *y* for any given input signal *x*.

3.1 MEASUREMENT TECHNIQUE

Employing a proper test signal, and sampling the output of the system, it is possible in general to compute not only the simplified, one-dimensional Volterra kernels, but even the complete n-dimensional ones. For example, Reed and Hawksford developed such a measurement technique based on varying-amplitude repetitions of an MLS (maximum length sequence) signal [6]. This technique revealed to be too complex and slow for practical applications, so the authors developed an alternative method [4].
The following pictures explain the measurement method. First, a test signal of length *T*, and covering the frequency range from ω_1 to ω_2 is generated, with this analytical expression:

$$x(t) = \sin[\omega_{var}] = \sin\left[\omega_1 \cdot T \Big/ \ln\left(\frac{\omega_2}{\omega_1}\right)\cdot\left(e^{\frac{t}{T}\cdot\ln\left(\frac{\omega_2}{\omega_1}\right)} - 1\right)\right] \quad (3)$$

When the signal is introduced in the non-linear system, its output also contains harmonic distortion products, as shown here:

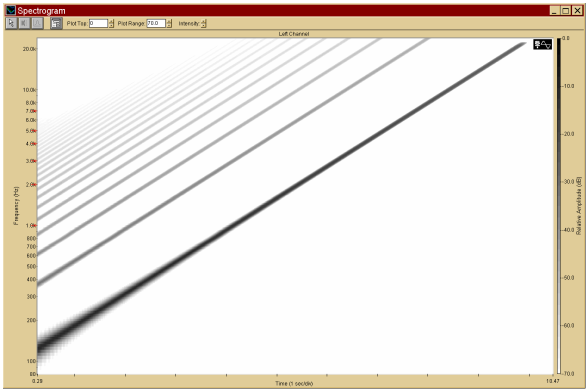

Figure 7. Spectrogram of the system’s response

It is possible to deconvolve the impulse response by applying to this response, by convolution, a proper inverse filter, which is simply the time-reversal of the excitation signal (3), equalized with a slope of 6dB/oct (time-reversal mirror plus whitening filter). This is the result:

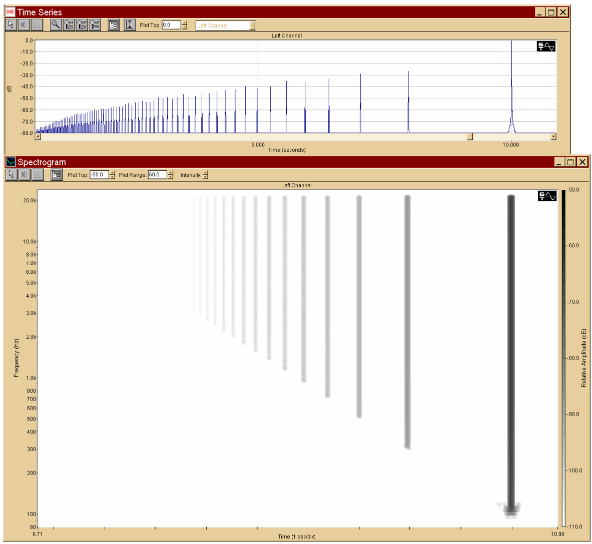

Figure 8. Spectrogram of deconvolved impulse responses

The rightmost impulse response is the linear one, which is preceded by the second-order harmonic response, and so on. The measured impulse responses are not directly the Volterra kernels, but these are easily computed by solving a linear equation system.

3.2 COMPUTATION OF THE VOLTERRA KERNELS

In practice the measurement procedure described in the previous chapter produces ordered impulse responses. The measured output signal can be represented as the sum of the linear convolution of the measured ordered impulse responses h'_i with the original input signal and the corresponding frequency-shifted version (obtained introducing respectively a factor 2, 3, etc. inside the sine function of eqn. 3):

$$y(t) = h'_1 \otimes \sin[\omega_{var}] + h'_2 \otimes \sin[2\cdot\omega_{var}] + h'_3 \otimes \sin[3\cdot\omega_{var}] + \ldots \quad (4)$$

The same result can also be espressed in terms of the diagonal Volterra kernels h_i defined in chapter 2:

$$y(t) = h_1 \otimes \sin[\omega_{var}] + h_2 \otimes \sin^2[\omega_{var}] + h_3 \otimes \sin^3[\omega_{var}] + \ldots \quad (5)$$

The powers up to 5^{th} order of sine functions are:

$$\begin{aligned}
\sin^2(\omega\cdot\tau) &= \frac{1}{2} - \frac{1}{2}\cdot\cos(2\cdot\omega\cdot\tau) \\
\sin^3(\omega\cdot\tau) &= \frac{3}{4}\cdot\sin(\omega\cdot\tau) - \frac{1}{4}\cdot\sin(3\cdot\omega\cdot\tau) \\
\sin^4(\omega\cdot\tau) &= \frac{3}{8} - \frac{1}{2}\cdot\cos(2\cdot\omega\cdot\tau) + \frac{1}{8}\cdot\cos(4\cdot\omega\cdot\tau) \\
\sin^5(\omega\cdot\tau) &= \frac{5}{8}\cdot\sin(\omega\cdot\tau) - \frac{5}{16}\cdot\sin(3\cdot\omega\cdot\tau) + \frac{1}{16}\cdot\sin(5\cdot\omega\cdot\tau)
\end{aligned} \quad (6)$$

Taking the Fourier transform of eq. 4, and showing the result for a given value of the output frequency ω, we obtain:

$$Y(\omega) = \overline{H'}_1[\omega]\cdot X[\omega] + \overline{H'}_2[\omega]\cdot X[\omega/2] + \overline{H'}_3[\omega]\cdot X[\omega/3] + \ldots \quad (7)$$

This must be equal to the Fourier transform of eq. 5, where the expressions 6 have been substituted:

$$\begin{aligned}
Y(\omega) = &\left[\overline{H}_1 + \frac{3}{4}\cdot\overline{H}_3 + \frac{5}{8}\cdot\overline{H}_5\right]\cdot X[\omega] + \left[-\frac{1}{2}\cdot\overline{H}_2 - \frac{1}{2}\cdot\overline{H}_4\right]\cdot j\cdot X[\omega/2] + \\
&+ \left[-\frac{1}{4}\cdot\overline{H}_3 - \frac{5}{16}\cdot\overline{H}_5\right]\cdot X[\omega/3] + \frac{1}{8}\cdot\overline{H}_4\cdot j\cdot X[\omega/4] + \\
&+ \frac{1}{16}\cdot\overline{H}_5\cdot X[\omega/5] + \ldots
\end{aligned} \quad (8)$$

If the Fourier's theorem holds, this must be true for any input signal X, so the corresponding terms at second member of eqs. 7 and 8 must be the same. This allows for the following linear equation system to be set up for each frequency ω:

$$\begin{cases}
\overline{H'}_1 = \overline{H}_1 + \frac{3}{4}\cdot\overline{H}_3 + \frac{5}{8}\cdot\overline{H}_5 \\
\overline{H'}_2 = -j\cdot\frac{1}{2}\cdot\left[\overline{H}_2 + \overline{H}_4\right] \\
\overline{H'}_3 = -\frac{1}{4}\cdot\overline{H}_3 - \frac{5}{16}\cdot\overline{H}_5 \\
\overline{H'}_4 = j\cdot\frac{1}{8}\cdot\overline{H}_4 \\
\overline{H'}_5 = \frac{1}{16}\cdot\overline{H}_5
\end{cases} \quad (9)$$

The solution of this system allows for the computation of the unknown diagonal Volterra kernels h_i starting from the measured ordered impulse responses h'_i.

This is the solution for the first 5 orders:

$$\begin{cases}
\overline{H}_1 = \overline{H'}_1 + 3\cdot\overline{H'}_3 + 5\cdot\overline{H'}_5 \\
\overline{H}_2 = 2\cdot j\cdot\overline{H'}_2 + 8\cdot j\cdot\overline{H'}_4 \\
\overline{H}_3 = -4\cdot\overline{H'}_3 - 20\cdot\overline{H'}_5 \\
\overline{H}_4 = -8\cdot j\cdot\overline{H'}_4 \\
\overline{H}_5 = 16\cdot\overline{H'}_5
\end{cases} \quad (10)$$

After computing the values of the kernels from the measured multiple impulse response, the non-linear convolution can be efficiently implemented following eq. 2.

Despite of the large number of multiply-add operations theoretically required for computing eqn. 2, the computation is very fast (real-time on modern computers), thanks to a frequency-domain implementation of each convolution operator, based on many FFTs and the well-known select-save algorithm [7]. This makes it possible to keep the process in real-time, even for values of *M* (impulse response length) of ten thousands coefficients and more.

4 EXPERIMENTAL VERIFICATION

The two methods described in the precedent chapters have been tested by means of a listening experiment.

A sound sample (a song by Elton John) has been used as the sound sample being processed.

The following figure shows the original recording, unprocessed:

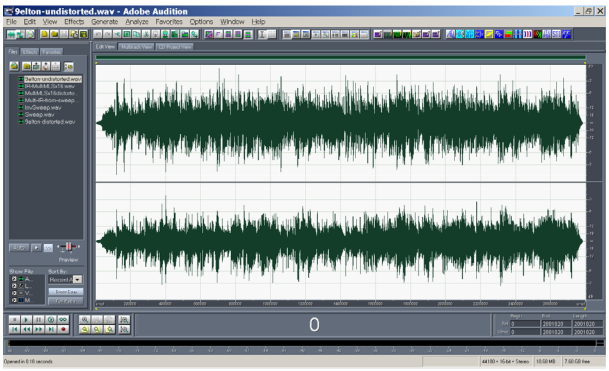

Figure 9 original song (unprocessed)

It can be seen that this recording has quite good dynamics (not as some “modern” pop recordings, which are heavily compressed and “level maximized”), and thus is a good candidate for exploiting notlinear effects of the device under test.
After passing this recording to the guitar effect, we get this “reference” distorted signal:

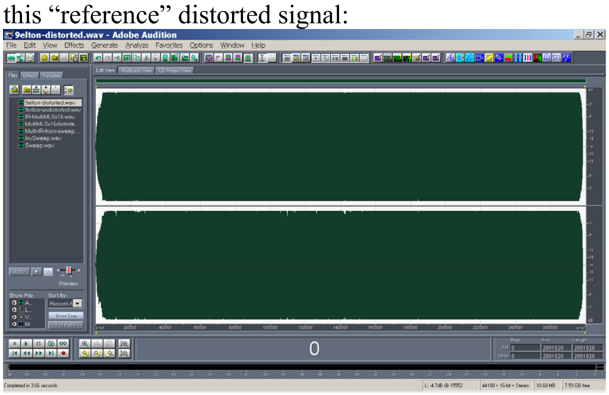

Figure 10. original song (distorted by the guitar effect)

It can be seen how now the signal is heavily compressed, and the dynamics is gone…
Listening to the distorted signals reveals, as expected, a strongly distorted voice and very bad quality 8compared to the original).
Now we want to see (and to listen) at what degree the IR-switching and diagonal-Volterra-kernels techniques manage to reconstruct this heavy not-linear distortion effects.
The following picture shows the effect of the IR-switching reconstruction:

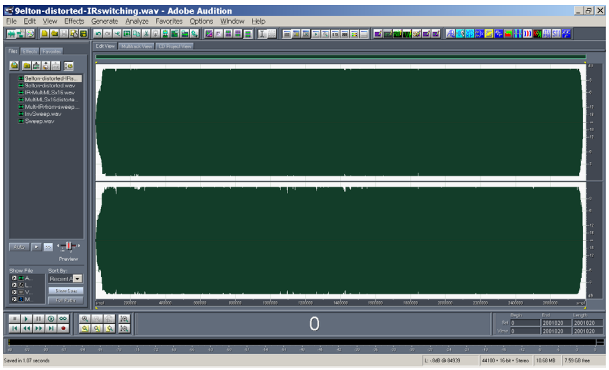

Figure 11 simulated distorted signal
with IR-switching technique

The waveform looks quite similar to the “really distorted” one, and the listening test confirmed that the similitude extends also to timbric content and dynamics.
Then the simulation by means of the diagonal Volterra kernels technique was attempted. The following figure shows the result obtained this way:

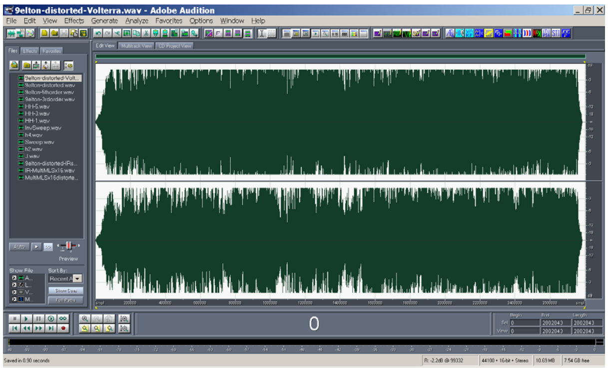

Figure 12 simulated distorted signal
with diagonal Volterra kernels technique

It can be noted her that in this case the emulation of the notlinear effect was not so accurate as with the IR switching technique. This is due to the fact that the analysis was limited to 5th-order distortion, and in this case of heavily notlinear it would have been advisable to go up at least to 10th order.
Also the listening test confirmed this fact: the sound was “less distorted” than the one created with the IR-switching technique, and more close to the purely-linear convolution.

5 CONCLUSIONS

Some informal listening tests conducted with the ABX Comparator [8] demonstrated that none of the two proposed techniques can fool a skilled listener: the “original” distorted signal is easily distinguished from the simulated ones, albeit these sound subjectively not very different from it.
The difference between notlinear convolution and linear convolution, however, is really dramatic. In fact, even if it is possible to distinguish the not-linearly simulated signals from the original, the difference is very subtle with both methods, and instead the difference between the linearly-simulated signal and the original is very sensible.
This means that, even of it is necessary to further refine these notlinear convolution schemes, they already provide a quite realistic simulation of the reality, much better than employing just linear convolution.
Due to the lower implementation complexity and to the higher efficiency obtained by means of fast block convolution, the diagonal Volterra kernels technique appears in this moment as the optimal candidate for creating a new generation of audio plugins, capable of simulating very convincingly the notlinear behaviour of audio gear such as microphone preamplifiers, compressors, limiters, etc., providing that they do not exhibit too much memory effects.

6 REFERENCES

[1] A. Bellini, G. Cibelli, E. Ugolotti, A. Farina, C. Morandi - "Non-linear digital audio processor for dedicated loudspeaker systems" - Proc. of International IEEE Conference on Consumer Electronics, Los Angeles, June 1998, and in IEEE Transactions On Consumer Electronics, vol. 44, August 1998, pp. 1024-1031.

[2] A. Farina, E. Ugolotti, A. Bellini, G. Cibelli, C. Morandi - "Inverse numerical filters for linearisation of loudspeaker's response" - DAFX-98 Conference, November 19-21, 1998 Barcelona, Spain

[3] A. Farina, A. Bellini, E. Armelloni – "Non-Linear Convolution: A New Approach For The Auralization Of Distorting Systems" - Proceedings of the 110th AES Convention, 2001 May 12-15 Amsterdam, The Netherlands

[4] A. Farina – "Simultaneous measurement of impulse response and distortion with a swept-sine technique" - 108th AES Convention, Paris 18-22 February 2000.

[5] Michael J Kemp - "Analysis and Simulation of Non-Linear Audio Processes using Finite Impulse Responses Derived at Multiple Impulse Amplitudes", presented at the 106th AES convention in Munich, Germany, 8 – 11th May 1999. preprint number 4919.

[6] Hawksford M.O.J., Reed M.J., "Identification of discrete Volterra series using maximum length sequences", IEE Proc.-Circuits Devices Syst., Vol. 143, No. 5, pp. 241-248, October 1996.

[7] Oppheneim A.V., Schafer R., *Digital Signal Processing*, Prentice Hall, Englewood Cliffs, NJ 1975, p. 242

[8] David Carlstrom - The ABX web site: http://www.pcavtech.com/abx/index.htm

Angelo Farina is professor of Environmental Technical Physics at the University of Parma, where he holds the course of Applied Acoustics. He is author of more than 200 scientific papers, which cover many research topics: room acoustics, advanced electroacoustical measurements, numerical models of the sound field in enclosures and outdoors, auralization, digital signal processing applied to sound processing and music recording/reproduction, three-dimensional surround sound systems. More info and all the papers can be found on HTTP://www.angelofarina.it

Enrico Armelloni is a research fellow of the University of Parma, actually under a two-years research contract sponsored by Regione Emilia Romagna (LAV, Laboratorio Acustica e Vibrazioni). After a 5-years degree in Electronics engineering, he has got a Ph.D. in Industrial Engineering, and worked mainly on topics related to DSP processing, multiple loudspeaker surround sound system, efficient convolution algorithms, Stereo Dipole systems. Together with prof. Farina, he is one of the founders of AIDA, the first spinoff of the University of Parma. More info on HTTP://www.aidasrl.it