Realtime auralization employing a not-linear, not-time-invariant convolver

Angelo Farina¹, Adriano Farina²

- 1) Industrial Engineering Dept., University of Parma, Via delle Scienze 181/A Parma, 43100 ITALY – HTTP://www.angelofarina.it
- 2) Liceo Classico G.D. Romagnosi, via Marialuigia 1, Parma, 43100 ITALY

HTTP://www.adrianofarina.it



123rd AES Convention 2007 October 5–8 New York, NY, USA





Goals for Auralization

- Transform the results of objective electroacoustics measurements to audible sound samples suitable for listening tests
- Traditional auralization is based on linear convolution: this does not replicate faithfully the nonlinear behaviour of most transducers
- The new method presented here overcomes to this strong limitation, providing a simplified treatment of memory-less distortion

What's linear convolution ?

- Also called "FIR filtering"
- Convolution is the mathematical operation performed when filtering a waveform x employing as filter coefficients the samples of a second waveform, usually denoted as h and called "impulse response"
- The samples of the input waveform are multiplied by the samples of the impulse response h, and the results accumulated (summed):

$$\mathbf{y}(\mathbf{j}) = \sum_{i=0}^{N-1} \mathbf{x}(\mathbf{j} - \mathbf{i}) \cdot \mathbf{h}(\mathbf{i})$$

Auralization by linear convolution





Convolving linearly a suitable sound sample with the Imp.Resp., the frequency response and temporal transient effects of the system can be simulated properly

Auralization by linear convolution

The beginnings: hardware DSP-based convolution units





Lake Technologies HURON

Yamaha SREV-1



Sony DRES-777



The AMBIOPHONICS Institute: the home of convolution







Photos taken on 16 december 2002





Software Convolution: BruteFIR and AlmusVCU

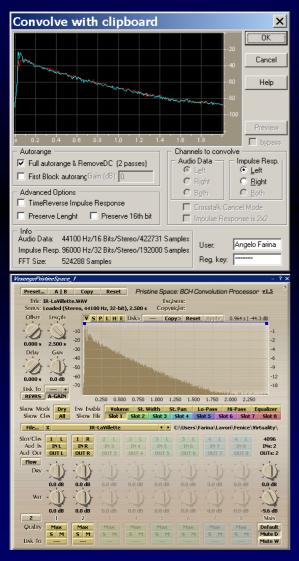
Open-source software for Linux by Anders Torger – AES 24° Conference



Performance: a fanless (silent) P-IV running at 2.5 GHz was capable of real time convolution of 2 inputs at 44.1 kHz, 24 bits, with 48 impulse responses, each 5s long, driving 24 Genelec loudspeakers (20 satellites + 4 subwoofers), employing 75% of the CPU time

Auralization Software

Nowadays many sytems or software plugins can perform satisfactorily the Linear Convolution operation, and are widely employed for audio processing



<i>SIR_1010</i> _	1										- ?	X 1	
S	R	R v1.010	written b	Response F y Christian F ware is FRE	Knufinke	www.	.knufinke	.de/sir	Help / About		lpen File c - up ORTF_L.wav		
	-		Farina\Articol				200.001	E Ditt		1st-IR-RegioP	arma_Montarbo_C_	P	
		C. VOSEISV	r alina valucu	19860-241	n-wave	svin_n#11	200_061	r_rul.wav	reverse	1st-IR-Valli-SC DIRAC.wav	enter-R1-Dodecha	ed	
N									2.50 sec.	IR_PR-Aud_0	RTF_Pt1L.wav		
									44100 hz)RTF_Pt1L.wav)RTF_Pt1L.wav		
later.											RTF_Pt1L.wav		
								Dry	Vet)RTF-autorange.w		
Predelay	0 ms		50 100	150	200	250	300	0			ORTF_Pt1L-autora er_off1mR1-Ger		
Attack	100%		50	0 ms		150	300		-3.1		I_Genelec_Odeg_(
Envelope	0%		17 33	50	67	83	100		-7.2		atre_Genelec_Odeg	a.l	
Length	100%	0	17 33	50	67	83			-12.6	IR-StateTheat IR-The Studio	er+Pt1.wav _Genelec_Odeg_C	F	
Stretch	100%	50	67 83		117	133	150		-19.6		IORTF_Pt1·L.wav		
Stereo In/IR	100%	0	50	100%	0	50			-29.1		NORTF_Pt1-L.wa NORTF_Pt1-L.wa		
reset								-42.9 -66.8	-42.9 -66.8	Una_Deconv_	NUNIF_F(I+L.Wa	×	
0 dB -6.0 dB								-66.0	-00.0				
-12.0 dB								-00 dB	0.08				
-18.0 dB -24.0 dB								ON	ON				
-00 dB													
reset 55	92	156 263	445 752	1k 2k	3k	6k 10k	17k	GAIN +5.6	dB +12dB				
WaveS.	hel <u>l</u> V	ST_5_1	21										'
₩>))) IR	i-1 🗖	Indo	_	_	Se	atup A	_		A->B	Load Sa	ave ? W	WAVE
F	Full CPU		Reverse			Bypass	— Ga	in Envelop	e —— 🚺 Clear		R 1.00 F 6	00 R 1.00 F	2500
	:Hall - 1										Damping		
	Concert 24 Mar 2											• • • •	
SR:96	000Hz ->			D The								• • • •	
Emitte	er:Genele	c S30D Original	Current								Equalizer		
Convo	lution:	1.85s	1.85s					and a start				250 lk 4	_
RT60:		1.4s	1.4s					and the for the second of the				0.0 0.0	0.0
Chann Size:		4 11267	4	-			0.5			The second second second		301 1200	5005
Distan	nce:	11267 13m	11267 13m	- Zo	om 🕂		10.5	DQSec ,	1 1	.000Sec		1.28 1.64	
	_	_	_	_	_		_	_	_				0~
	- Rever	b Time —			nsity 00	Reso	Decor	_			Predelay	/ Output	
	1	0.000		00	.00	1.00	0	Dr					n i
		Cnv. Len							ims _	- (0 0		
1		Full				1		ED D	Jildup				
		1.4			-				maup				
	R				Т	T	1	ERIT		Tail	0 0		
	Ra	rtio 1.0	0				E		K-X				

? X

Linear Impulse Responses

Huge collections of impulse responses have been measured in famous theatres and concert halls all around the world, as well for renowned audio processing gear.

Recent advancements in the measuring technique, making use of the Exponential Sine Sweep signal, and the usage of multiple sound sources and microphones, make it possible to capture with minimal noise a detailed "acoustical photo" of existing rooms.

See, for example, the results of the Waves project on:

www.acoustics.net



Furthermore, all room acoustic modelling software is nowadays equipped with a "rendering" tool, which exports impulse responses suitable to be employed for auralization by linear convolution. These synthetic IRs are even cleaner and with greater dynamic range than any measured IR

What's missing in linear convolution ?

- No harmonic distortion, nor other nonlinear effects are being reproduced.
- From a perceptual point of view, the sound is judged "cold" and "innatural"
- A comparative test between the simulation of a strongly nonlinear device and an almost linear one does not reveal any audible difference, because the nonlinear behavior is removed for both

How can we perform not-linear convolution?

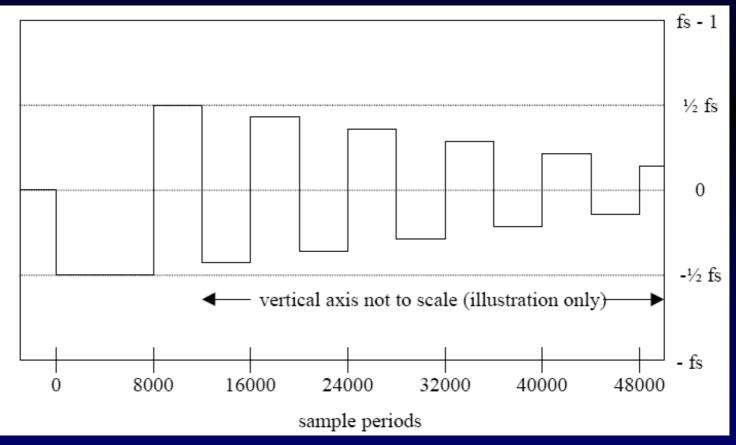
- There are actually two competing approaches available in the audio industry:
- IR-switching technique
- Diagonal Volterra Kernel

Method 1 (IR switching)

- A very simple idea: a different IR is employed depending on the amplitude of each sample of the signal to be filtered
- The method is quite old: the first published papers are those of Bellini and Farina (1998) and Michael Kemp (1999)
- Several impulse responses are measured, employing test signals of different amplitudes, and stored for later usage.
- It is mandatory to implement the convolution as direct form in time domain, as each sample has to be processed with a different IR.

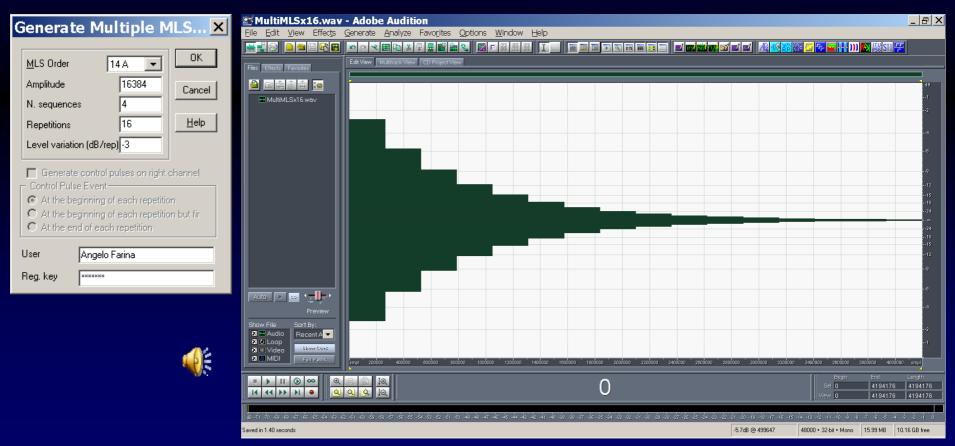
Measurement of multiple IRs

• Michael Kemp employed a step function, with several steps of decreasing amplitude



Measurement of multiple IRs

• Farina e Bellini did employ a sequence of MLS repetitions, each with decreasing amplitude, providing better S/N ratio in real-world measurements



Implementation (Michael J. Kemp)

Focusrite did release recently Liquid Channel, the first "dynamic convolver" implementing the IR-switching technique



Zounds

"The **Liquid Channel** is a revolutionary professional **channel** strip that can precisely replicate any classic mic-pre and compressor"

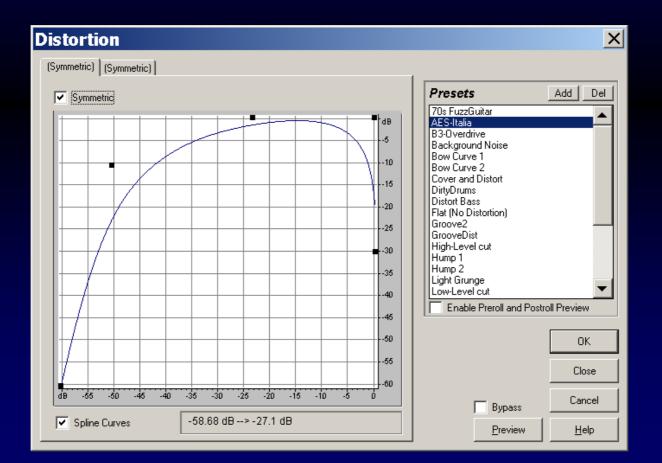
Implementation (Farina/Bellini)

• A FIR filtering algorithm, with the set of coefficients chosen depending on the sample amplitude, was implemented on a Sharc EZ-KIT 20161 board, and employed for car-audio applications



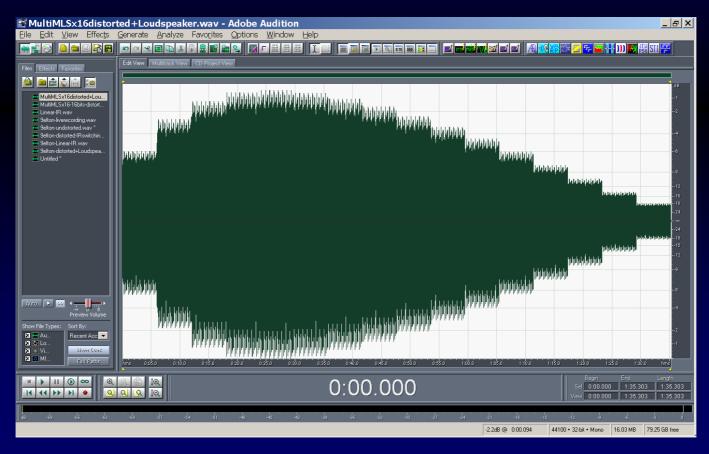


• The "not linear device" is emulated by the DISTORTION plugin of Adobe Audition, followed by sound playback and simultaneous recording over the loudspeaker and microphone of a laptop PC





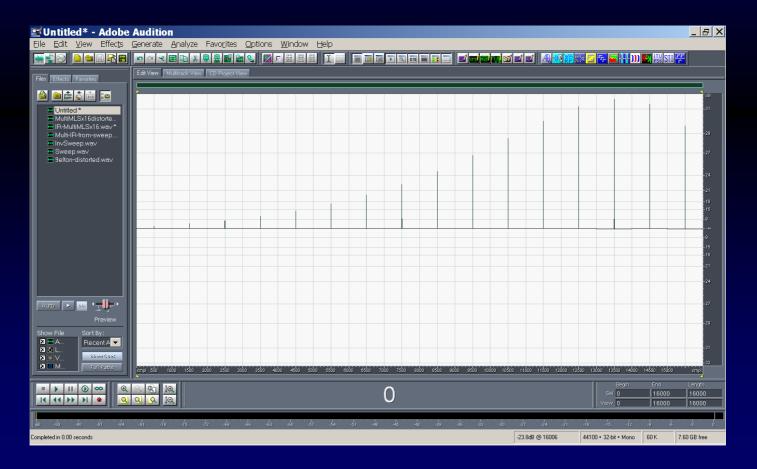
• This is the multiple MLS signals after being processed through the not-linear device







 Here the 16 impulse responses measured with MLS of different amplitude (decreasing 3dB each from left to right) are shown



Audible evaluation of the performance

Original signal

	Grone Michael Berner States	
- Charles and the more that the	al-andiputrik katalah ng Katalang Lakang katalah ng katalan Panang ^{ang} parangkan pilikan pangkang katalang katalang katalang katalang katalang katalang katalang katalang k	indiana dia tang manina ang manina ang manana ang manana ang manana ang manana ang manana ang manana ang manana Ng manana ang
- Contractor de la contracta d	rene fingini kulonda in privanske klasila da da Azerditi privatka stitutna serija pola ji one	a Malaula Manalayara Manala September para Manguan Ma
	0:00.000	1944 - 1949 - 2019 - 945 - 246 - 246 - 247 - 247 197 - 2019 - 2019 - 2019 - 2019 197 - 2019 -
Operand as 0.18 seconds	A.5	2 6 6 2 10 12 2 3 4 10 + 16 6 + 5 mm 13 83 MB 75 50 58 mm

, en hin i provinsi an hini pr	atel di su deservici se stario di aterio di di condicio di segue si	
, and by the second	le i a fa bhu an a chuir da scraith i a bhliann a fhair a Mairich	
	e de la face de la completa de la completa por la deve de la completa por la completa por la completa por la co Na por de la completa por la completa de la comp	
	0:00.000	teta ada ada ada ada teta ada ing

Live recording



Non-linear (IR switching)

Linear convolution





Method 2 – Diagonal Volterra Kernels

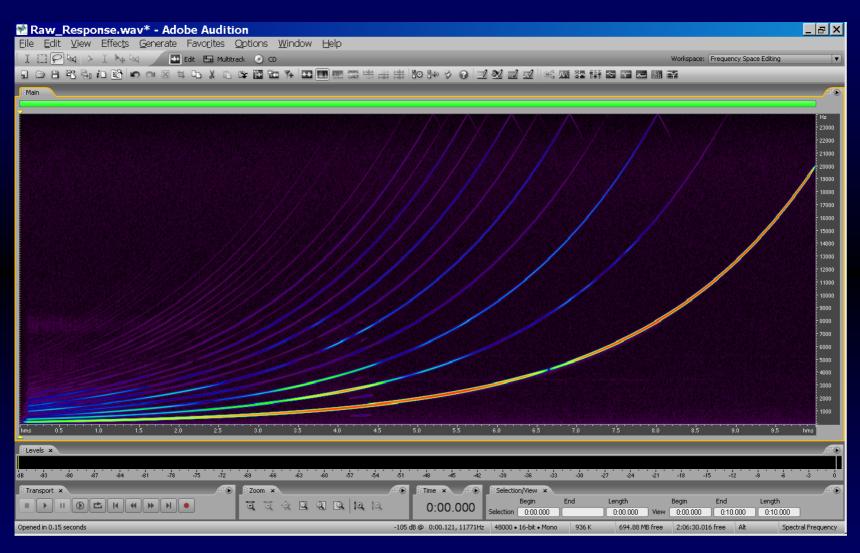
- We start from a measurement of the system based on exponential sine sweep (Farina, 108th AES, Paris 2000)
- Diagonal Volterra kernels are obtained by postprocessing the measurement results
- These kernels are employed as FIR filters in a multiple-order convolution process (original signal, its square, its cube, and so on are convolved separately and the result is summed)

Exponential sweep measurement

Test.wav* - Adobe Audition Eile Edit View Effects Generate Favorites Options Window Help
Image: Second at a state of the second se
۶:Main A
Hz 200
190
1 100
150
140
100
400
- 300 - 200 -
- 100
hms 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5 5.0 5.5 6.0 6.5 7.0 7.5 8.0 8.5 9.0 9.5 hms
Filevels ×
d8 -33 -90 -87 -84 -81 -78 -75 -72 -69 -66 -53 -60 -57 -54 -51 -48 -45 -42 -39 -36 -33 -30 -27 -24 -21 -18 -15 -12 -9 -6 -3
الا الله الله الله الله الله الله الله
Opened in 0.15 seconds -121 dB @ 0:08.051, 22311Hz 48000 • 16-bit • Mono 1030 K 694.81 MB free 2:06:29.248 free Alt Spectral Frequence

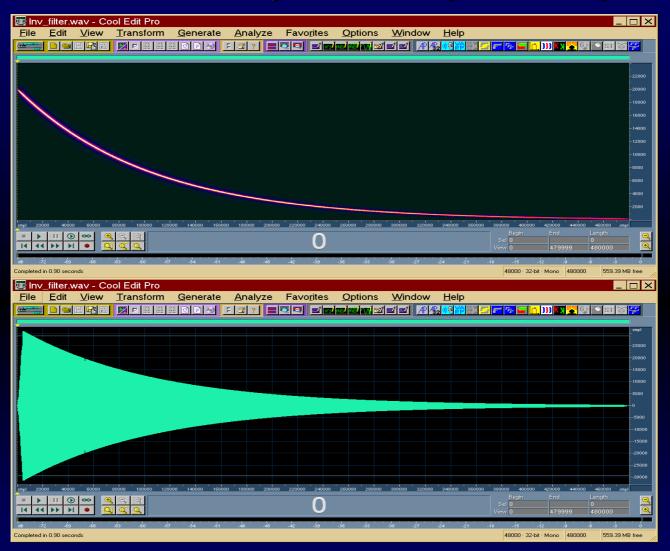
• The test signal is a sine sweep with constant amplitude and exponentially-increasing frequency

Raw response of the system



Many harmonic orders do appear as colour stripes

Deconvolution of system's impulse response



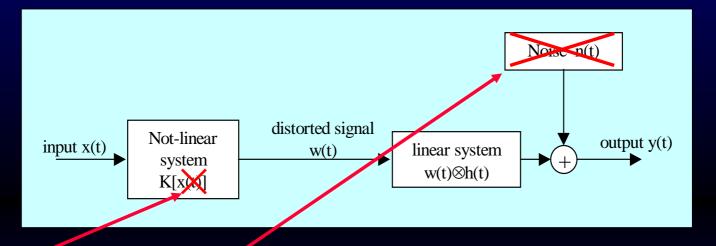
The deconvolution is obtained by convolving the raw response with a suitable inverse filter

Multiple impulse response obtained



The last peak is the linear impulse response, the preceding ones are the harmonic distortion orders

Memoryless distortion followed by a linear system with memory



- The first nonlinear system is assumed to be memory-less, so only the diagonal terms of the Volterra kernels need to be taken into account.
- Furthermore, we neglect the noise, which is efficiently rejected by the sine sweep measurement method.

Theory of nonlinear convolution

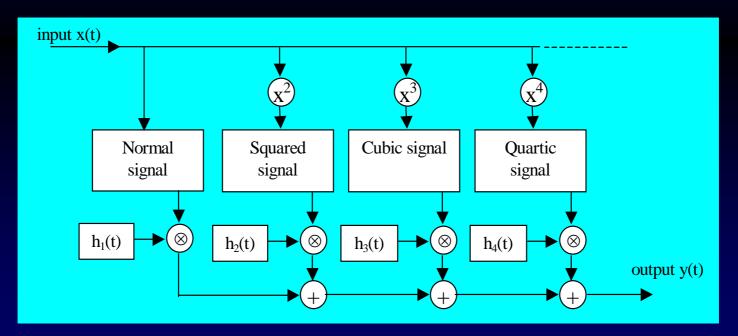
- The basic approach is to convolve separately, and then add the result, the linear IR, the second order IR, the third order IR, and so on.
- Each order IR is convolved with the input signal raised at the corresponding power:

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

The problem is that the required multiple IRs **are not** the results of the measurements: they are instead the diagonal terms of Volterra kernels, which can be derived from the set of measured Impulse Responses of several distortion orders

Efficient non-linear convolution

As we have got the Volterra kernels already in frequency domain, we can efficiently use them in a multiple convolution algorithm implemented by overlap-and-save of the partitioned input signal:



Software implementation

A small Italian startup company, Acustica Audio, developed a VST plugin based on the Diagonal Volterra Kernel method, named Nebula

Nebula3cm_Reverb_	1			- ? X
	ACUSTICA Auono Vectorial Volterra K	KRN: 3 PRF: 13 %	ATTCK RELS DRIVE L	10DT 10.83 %
hala da	Trying ²⁰¹ DSC: Prova FRT: 44100 Hz -> 48000 Hz GAIN IN: 3.7262 dB ATTACK: 10.839 ms RELEASE: 7.8035 ms DRIVE: 12.202 dB OUTPUT L: 1 -32.64 dB PROG FAST KERN	GAIN OUT: 5.7581 dB INPUT L: -11.49 dB INPUT R: -11.16 dB DNM SEL: -13.64 dB 8 LIQUIDTY: 70.833 % MAST MIDI DISK		latintatintatintatinta (

This is capable of real-time operation even with a very large number of filter coefficients

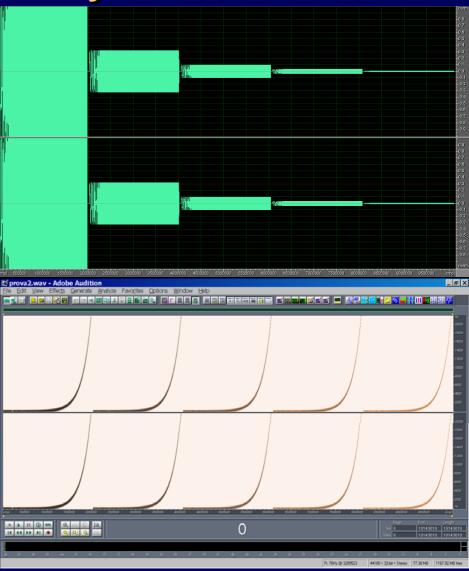
Software implementation

Nebula is also equipped with a companion application, Nebula Sampler, designed for automatizing the measurement of a not linear system with the Exponential Sine Sweep method:

🂐 NebulaSam	pler 📃 🗌	×	💐 NebulaSampl	er		_ 🗆 ×
<u>Eile A</u> bout			<u>Eile A</u> bout			
	enerator 📃 Deconvolver 🛄 Descriptors		🔛 Run All 👭 Tone Gene	rator 📃 📃 Deconvo	olver 🛛 🛄 Descriptors	
Asio			Parameters			
Driver:	Name: ASI04ALL v2		General:	Type:		ow fr (Hz): Length (s):
				Dynami 💌	44100 🔻 1	45
Input:	L: 18 - 32 bits R: 18 - 32 bits		Dynamic models:	Inverse:	Number: St	teps (dB): Space (s):
	1 - Realtek HD Audio rear 💌 2 - Realtek HD Audio rear 💌				5 1	0 1
Output:	L: 18 - 32 bits R: 18 - 32 bits				Number: St	ieps (dB):
	1 - Realtek HD Audio rear 💌 2 - Realtek HD Audio rear 💌			Г		
Latency:	Latency:		Time var models:		Period (ms): 20	
	23561 () Check 🦳 🔁 Open				120	
	Gain (dB): -43.50 □ ¢↑ (ASIO)		File system:	Output:	Length	
				C:/test.wav		Browse
Run	Period (ms): Decnv:		Repeats: Total time:			
Rt30: 1.75 ms	20 Check		2 00:07:38		M Generate >>	

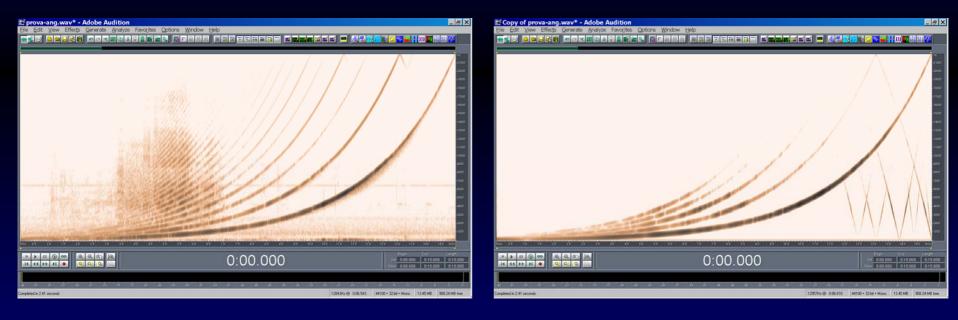
Time-variant systems

Nebula can sample also time-variant systems, such as flangers or compressors, by repeating the sine sweep measurement several times, along a repetition cycle or changing the signal amplitude



Reconstruction accuracy

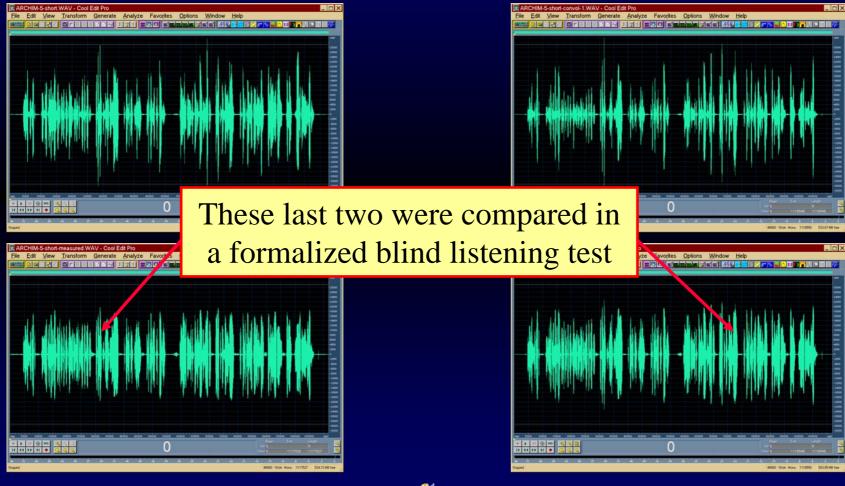
Nebula is actually limited to Volterra kernels up to 5th order, and consequently does not emulate high-frequency harmonics:





Audible evaluation of the performance

Original signal



Live recording



Non-linear Diagonal Volterra Kernel

Linear convolution

Subjective listening test

- A/B comparison
- Live recording & non-linear auralization
- 12 selected subjects
- 4 music samples
- 9 questions
- 5-dots horizontal scale
- Simple statistical analysis of the results
- A was the live recording, B was the auralization, but the listener did not know this

95% confidence intervals of the answers

📲 Risposte soggettiv	e					
Brano n. 1 2 3	4		Α	B 🕨 🗆		
D:\Convol_altop_lamiera\05RebeccaPidgeon-porta.WAV						
Domanda 1						
A & B are identical	+	• •	0	A & B are quite different		
Domanda 2						
A is more enveloping	• • •	• •	-	B is more enveloping		
Domanda 3	-					
A has better timber	• +	- C - C	-	B has better timber		
Domanda 4	Domanda 4					
A is more dry	° + °	- C +	0	B is more dry		
Domanda 5						
A is more distorted) F P	• •	0	B is more distorted		
Domanda 6						
A has more treble	• •	+	-	B has more treble		
Domanda 7						
A has more medium) ° 🗕		0	B has more medium		
Domanda 8						
A has more bass			C	B has more bass		
Domanda 9	1					
A is more pleasant		0	•	B is more pleasant		
Precedente	Successivo]		Fine		

Results

Statistical parameters – more advanced statistical methods would be advisable for getting more significant results

Question Number	Average score	2.67 * Std. Dev.
1 (identical-different)	1.25	0.76
3 (better timber)	3.45	1.96
5 (more distorted)	2.05	1.34
9 (more pleasant)	3.30	2.16

Comments

- Most listeneres judged the two samples identical
- However, sample B, on average, has slightly "better timber" (less distortion at high frequency), whilst sample A is "more distorted".
- Despite of the slight reduction in perceived distortion, the not-linear emulation was slightly preferred to the real-world recording.

Another example

•

Original signal

17-megonia.wav - Adobe Audition File Edit View Effects Generate Analyze Fo	avorites Options Window Help	_ = X
n a marta a se a fui a facta a faith a faith a Facta a faith a		
	anter anter relien rulear rulear rulear meter meter O	2017 Training of welling and and a second of product of the second of th
		454E @ 1705568 44100 - 1644 - More 425ME 1877 GB ive

18-megonia-rerecorded.wav - Adobe A File Edit View Effects Generate Analyze Fav CERE ALE INCOMENDATION OF THE ADDRESS OF THE	orites Options Window Help	_
	ander ander verben valeen valeen andere medien. O	order selar selar selar sel 2 2228224 2228224 2228224 2228224 1 2 2228224 2228224 2228224
and a second second Second	3	168 0-70111 44100-1694-Move 425ME 1571 58 lees

Live recording

Linear convolution





Non-linear Diagonal Volterra Kernel

Conclusions

- Traditional Linear Convolution is a powerful technique, but its reconstruction of the real world suffers for the limitations due to Linearity and Time Invariance
- Not-linear convolution is possible with two competing techniques:
 IR switching and Diagonal Volterra Kernels.
- The latter provides much more efficiency in computational load, and allows for easy emulation also of time-variant systems
- The Nebula software makes this technology available to everyone

Remarks

 The plugins for Adobe Audition shown here are freely downloadable from <u>HTTP://www.aurora-plugins.com</u>

-The Nebula VST plugin can be downloaded from <u>HTTP://www.acusticaudio.com</u>

- The sound samples employed for the subjective tests are available for download at <u>HTTP://pcfarina.eng.unipr.it/public/AES123</u>