SYNTHESISED MUSICAL STIMULI TO ENABLE OCCUPIED ROOM IMPULSE RESPONSE MEASUREMENT

M. Serafini University of Salford - Acoustic Research Centre
F. F. Li University of Salford - Acoustic Research Centre

1 INTRODUCTION

A framework of a new technique with the aim to enable room impulse response (RIR) measurement in occupied conditions is proposed in this paper. It consists of the use of two measurement methods: short linear chirps are employed to determine the RIR in a frequency range from circa 125 Hz up to 4 KHz; and a masked-MLS signal is applied as a probe stimulus to calculate the RIR in the frequency range from 4 KHz upwards. This paper concerns mainly the first method while further details of the second can be found in a previous conference publication by the author[1].

It is known that the major problem of performing occupied measures, which give the actual acoustic profiling of a space while in use, stems from the unpleasant testing stimuli. We experimented to mitigate their obtrusiveness through the use of short and narrow band chirps named 'presto-chirps'. The presto-chirps are purposely structured to represent musical notes. These synthesised notes are used to compose music-like stimuli, and hence circumventing the problem of annoying test signals.

The use of music signals to perform occupied measurement can be done using a dual-channel FFT technique, but the lack of frequency fullness often leads to a compromised accuracy, or a prolonged averaging time, which further causes time variance related problems[2,3]. With the presto-chirp method, it is possible to completely excite all the frequencies of interest (in the range 125 Hz – 4 KHz) to give a good signal to noise ratio (S/N) within a relatively short measurement duration.

1.1 Presto-chirps

'Presto-chirps' stimuli are derived from linear sine sweeps. They are defined through a number of parameters including the initial frequency, the bandwidth, the duration and the amplitude. These parameters are purposely defined so that each presto chirp represents a synthesized musical note. Each presto-chirp resolves a specific portion in the frequency domain of the RIR, and by summing up all the narrowband impulse responses the broadband impulse response is obtained according to superposition theorem applicable to linear systems.

1.2 Analytical description

The presto-chirps are created by using four parameters: the amplitude, the initial frequency, the bandwidth (or equivalently the final frequency), and the time duration. The linearly varying presto chirp can be written as:

$$x(t) = A\sin\left(2\pi\left(f_{in}t + \frac{(f_{fin}-f_{in})}{2T}t^2\right)\right) \quad (1)$$

Where f_{in} *and* f_{fin} represent the initial and final frequencies of the chirp, respectively; T and A are its duration (in seconds) and its amplitude (A Є [0, 1]). In order to generate the presto chirp by digital implementation, it can be written as:

$$x(n) = A\sin\left(2\pi\left(f_0 n + \frac{B/2}{N/fs}n^2\right)\right) \quad (2)$$

Where f_0 and f_s are the initial and the sampling frequency; B and N are the bandwidth and the number of samples. A discrete set of possible values for f_0 corresponding musical notes in an equally tempered scale are defined by associating integers to the formula (similar to the MIDI protocol). They are shown in figure 1.

Octave	Note Numbers											
	C	C#	D	D#	E	F	F#	G	G#	A	A#	B
-1	0	1	2	3	4	5	6	7	8	9	10	11
0	12	13	14	15	16	17	18	19	20	21	22	23
1	24	25	26	27	28	29	30	31	32	33	34	35
2	36	37	38	39	40	41	42	43	44	45	46	47
3	48	49	50	51	52	53	54	55	56	57	58	59
4	60	61	62	63	64	65	66	67	68	69	70	71
5	72	73	74	75	76	77	78	79	80	81	82	83
6	84	85	86	87	88	89	90	91	92	93	94	95
7	96	97	98	99	100	101	102	103	104	105	106	107
8	108	109	110	111	112	113	114	115	116	117	118	119
9	120	121	122	123	124	125	126	127				

Figure 1 – Midi note notation.

The presto-chirps' bandwidth varies from a few Hertz for lower registers to circa two hundred Hertz for high pitch notes. The set of 'notes' used in this study consist 63 presto-chirps ranging from 110 Hz (A2-A#2) up to 4 KHz (C7-C#7). A simple formula can translate a note number into the actual frequency (*n* corresponds to the MIDI number), thus:

$$f_{actual} = 440 \cdot \left(\sqrt[12]{2}\right)^{(n-69)} \qquad (3)$$

Since presto-chirps bandwidth (B) is taken of a semitone-width, their Fourier transforms correspond to a 1/12 octave filter banks as shown in figure 2.

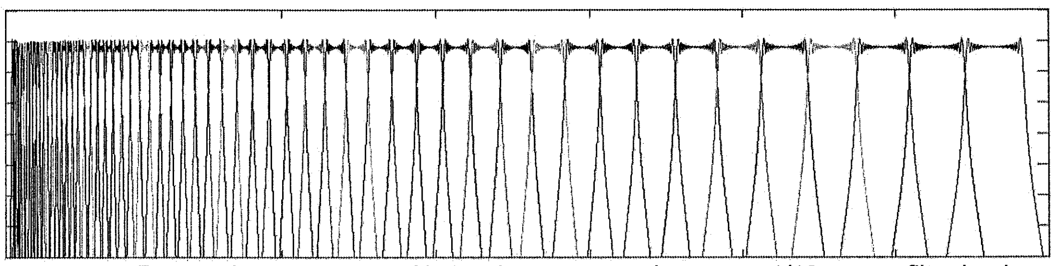

Figure 2 – Presto-chirps spectrum. Notice the correspondence to a 1/12 octave filter banks.

Limiting the analyzable higher frequency is needed, bearing in mind that the method itself aims to reduce the annoyance of traditional test signals. In addition the pitch of the highest musical notes is around 4 kHz. On the other hand, the lower limit has been chosen to be 110 Hz. Room's behaviour at low frequencies is better described through a modal analysis. Moreover, occupancy does not significantly affect responses below 100 Hz.

1.3 Windowing

Truncating a broadband chirp into narrow band ones involves windowing. Intrinsic artefacts such as ringing and ripples need to be minimised. Presto-chirps can be considered as a set of pass band filters for which it is desirable to have a sharp frequency response in the transition band, and a flat frequency response in the pass band. Unfortunately the sharpness of the gating function causes artefacts. Window functions are used to smooth the edges of the spectra and mitigate the problem. An empirical comparison of several windows function led to the use of a Tukey-window (also known as cosine-tapered window)[4,5]. This window allows the signal to pass without distortions in the pass band. A Tukey-window is mathematically described as:

$$w(n) = \begin{cases} 1, & 0 \leq |n| \leq \alpha\frac{N}{2} \\ \frac{1}{2}\left[1 + \cos\left(\pi \frac{n-\alpha\frac{N}{2}}{2(1-\alpha)\frac{N}{2}}\right)\right], & \alpha\frac{N}{2} \leq |n| \leq \frac{N}{2} \end{cases} \quad (4)$$

Where, N is the number of samples and α is the parameter that changes the shape of the window itself, thus for α=0 the window is rectangular while for α=1 it becomes a Hann window.

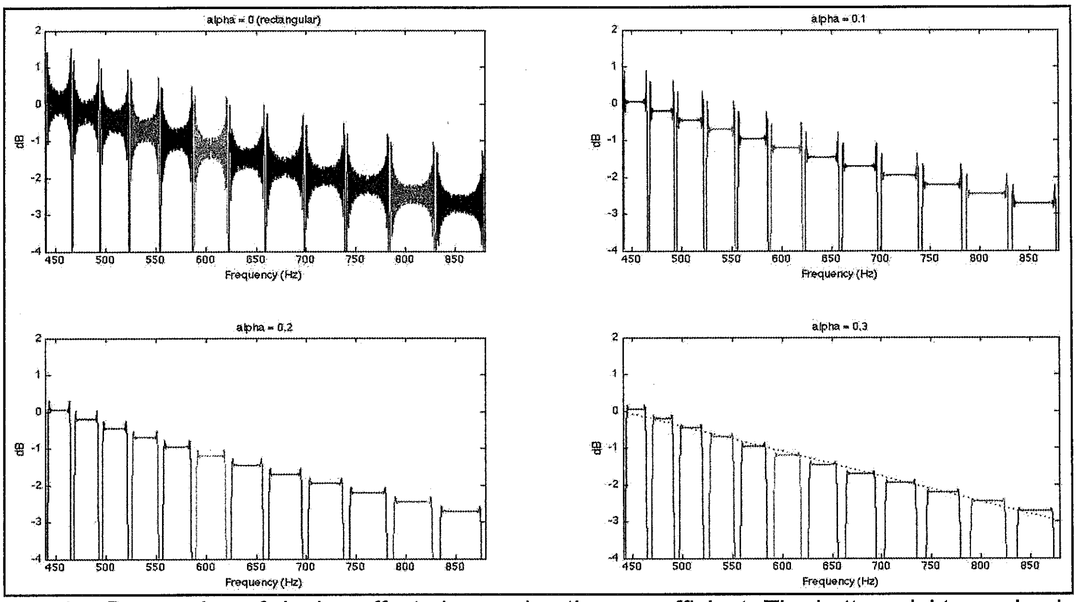

Figure 3 – Decreasing of ringing effects by varying the α coefficient. The bottom right panel points out the decreasing of the energy (3 dB/octave) while the bandwidth increases.

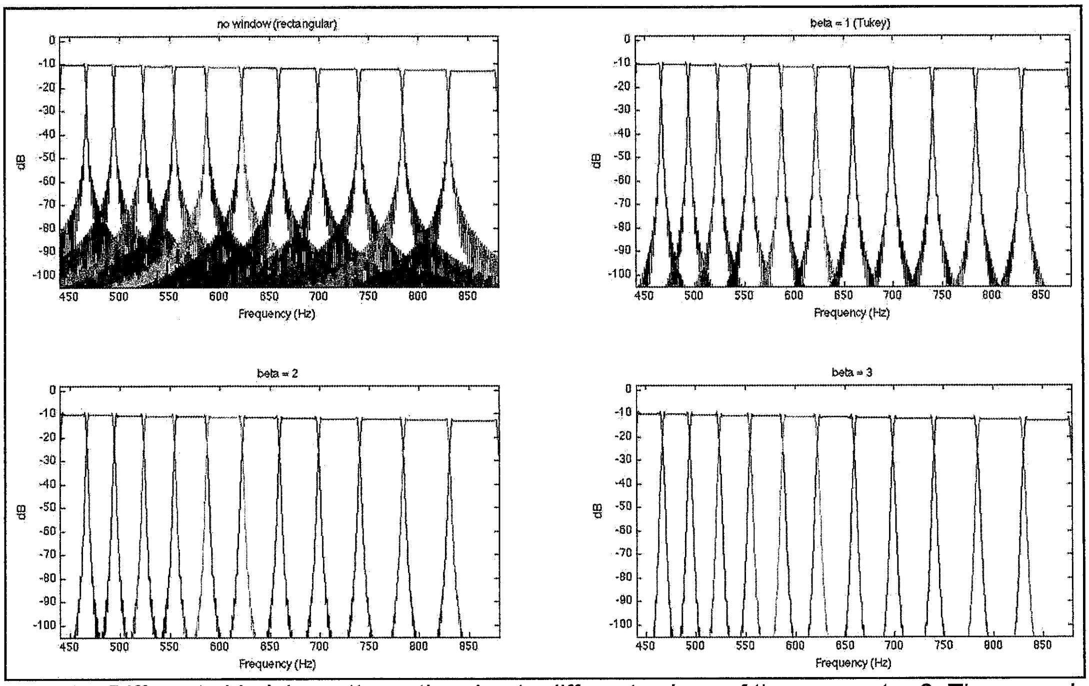

Figure 4 – Different side lobes attenuation due to different values of the parameter β. The upper left panel shows the situation when a rectangular window is applied (no-window).

An effect of increasing the value of α is an increased separation zone between two consecutive presto-chirps as shown in figure 5.

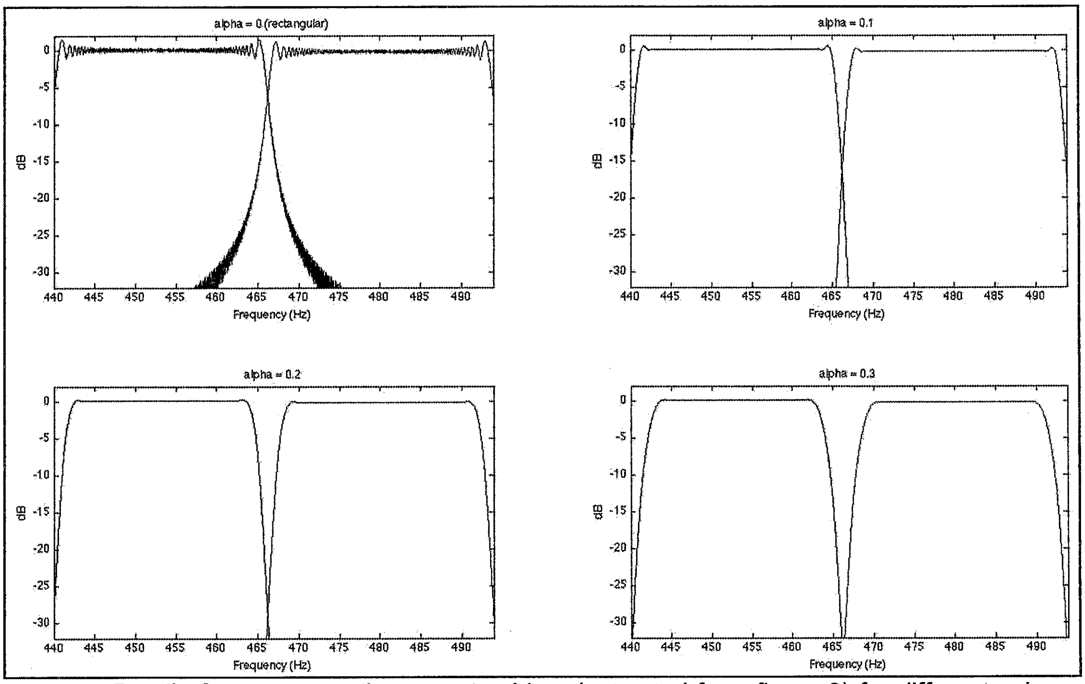

Figure 5 – Detail of two consecutives presto-chirps (extracted from figure 3) for different values of the parameter α of the Tukey window.

1.4 Energy normalization

The energy of presto-chirps is proportional to their duration and sweep rate. More precisely, $E \propto \frac{T}{2B}$, where T is the duration and B its bandwidth. Since the required bandwidth is wider in higher pitched notes, the associated presto-chirps for higher pitched notes will have less energy. Such decrease is of 3dB/Oct because the bandwidth doubles by doubling the frequency. It can be seen from Figure 3. A simple example is given here. To compensate for the different bandwidths and to obtain a flat spectrum as shown in figure 2, it is possible to apply a 'normalization' factor to any presto-chirps' matching filter. This factor (n) is proportional to two times the bandwidth, thus:

$$n = 2 \cdot (f_{fin} - f_{in}) = 2 \cdot B \qquad (5)$$

As a consequence, there will be reduced excitation strength in higher frequency sub bands. This does not cause major problems in many cases, since the background noise of typical auditoria has a pink-brown spectrum. On the other hand, the characteristic of presto-chips can vary during the composition in terms of amplitude, duration and number of repetitions, and hence several normalizations are needed to compensate these differences.

1.5 Composing music stimuli

Presto-chirps sound like woodwind instruments, such as a flute, and although listeners may perceive a minor 'sweeping' sound moving from a note towards a successive one (an effect musically called glissando). But the overall effect is not unpleasant. Anyhow, the developed method is not aimed to simulate an existing instrument. Several timbers can be obtained through synthesis techniques, thus through appropriate algorithms a set of harmonics can be generated and added to the fundamental to embellish it. So far, some timbers were experimented using the FM synthesis[6], the Karlus-Strong algorithms[7], and an experimental one that employs sweeping harmonics. They effectively contribute to enrich the stimuli.

To facilitate the composition of suitable music, MIDI notation was used, since its organized data structure contained all the information needed to generate the source stimuli. Moreover, a software simulation environment was developed to test the composed stimuli and their accuracy in calculating the RIR.

In brief, a MIDI file is a structured matrix that contains information relative to any of the notes which are used to compose a song. For example, pitch, and amplitude (also called velocity), the attack (note on) and end (note off) points, some control codes and the instrument/timber information. By reading the file it is possible to determine when a specific presto-chirp has to be called, for how long it lasts, and which timber should be added. Moreover, it also gives the information about when and which matching filter has to be used to deconvolve the impulse response. Composition is done using a midi keyboard to facilitate the creative work and the listening.

2 THE MEASUREMENTS (SIMULATION)

Simulations were carried out to test the accuracy in impulse response measurements, and it is done using an ascending scale as the stimulus. Presto-chirps are sequentially convolved with the room's RIR, and the output stored in a single array. Since the signal is completely deterministic, the matching filters can be generated and used at any time. By deconvolving portions of the system's output using appropriate filters (time reverse of the signals), the impulse responses in music note specified sub-bands are obtained; the 'final' RIR is calculated by summing them up. A graphic illustration is presented in figure 6.

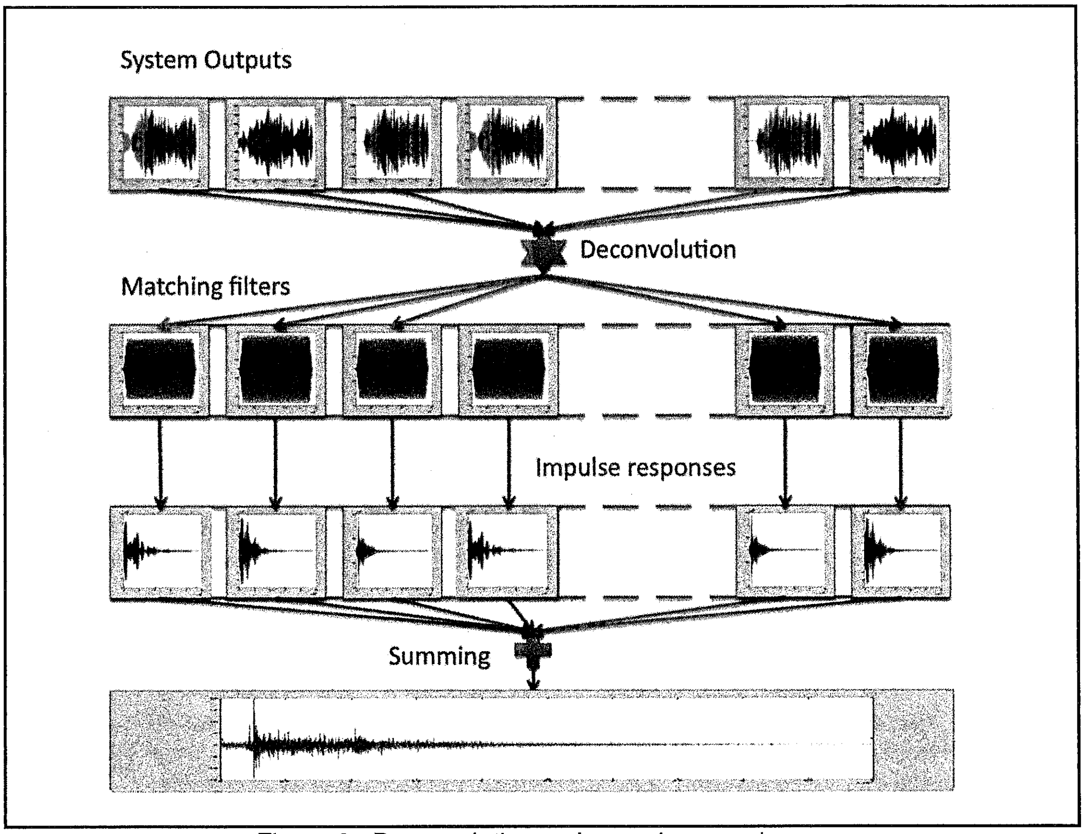

Figure 6 - Deconvolution and summing up scheme.

In order to resolve all reflections, the system output to be deconvolved (through linear deconvolution[8]) has to be longer than the reverberation time. For this reason there are some constraints in the composition of the music. On the other hand more than one presto-chirp can be played at the same time as a chord. This offers a method to perform a simultaneously analysis of several sub-bands, as shown in figure 7.

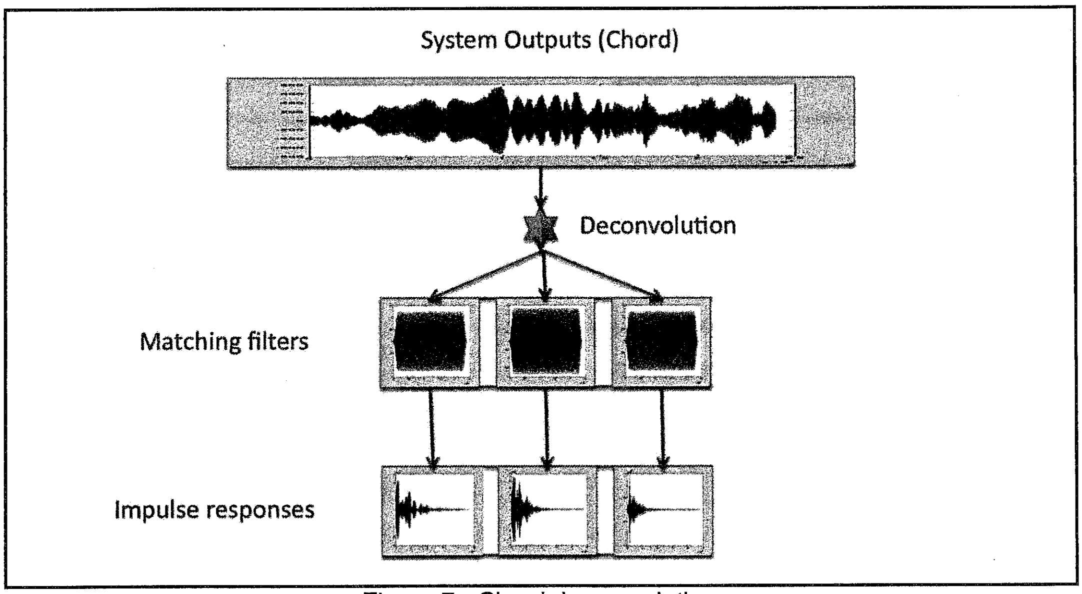

Figure 7 - Chord de-convolution.

2.1 Error function

To evaluate the accuracy of estimation, comparison between actual values and the estimated values can be performed using a coherent function. For example, the ratio of the estimated frequency response function and the actual one was used to measure the distribution of errors along the spectrum, as shown in figure 8. It can be seen that sharp dips appear in those frequency points where presto chirps are switched in and out. This is not surprising since the energy fed into the system at these frequency points is attenuated by the window function.

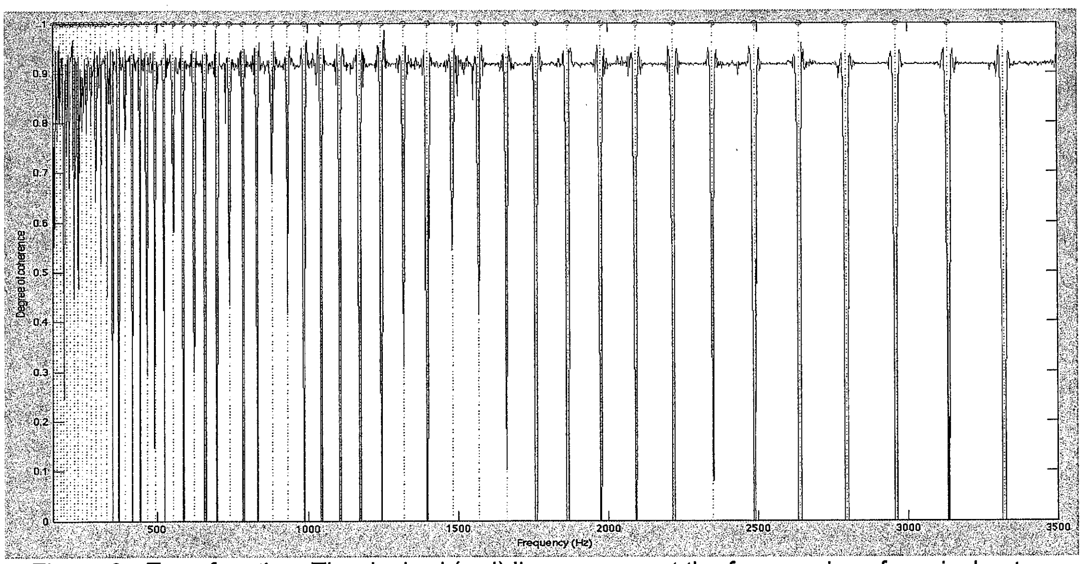

Figure 8 - Error function. The dashed (red) lines represent the frequencies of musical notes.

2.2 Presto-chirp duration

One particular aspect that is worth considering is the duration of presto-chirps. It has been found that the accuracy of the measurement is related to the relative duration of the excitation with reference the RT. Loosely speaking, too short or too long a stimulus can give an erroneous estimation. Some empirical results are illustrated in the tables 1 and 2 (two cases studied).

Hz	125	250	500	1000	2000	4000
Actual (RT30) sec	1.55	1.64	1.53	1.46	1.24	1.02
0.1 sec	3.64	4.06	3.29	2.63	1.87	1.45
0.2 sec	3.27	2.62	1.95	1.75	1.36	1.20
0.5 sec	**1.47**	**1.65**	**1.56**	**1.48**	**1.26**	**1.12**
1 sec	**1.60**	**1.63**	**1.54**	**1.48**	**1.32**	**1.18**
1.5 sec	1.69	1.98	2.27	2.42	2.16	1.76
2 sec	2.49	3.16	3.65	4.12	3.34	1.23

Table 1 – CASE 1 - Estimated RT30 vs. presto-chirp duration.

Hz	125	250	500	1000	2000	4000
Actual (RT30) sec	1.8	2.1	1.6	0.86	0.66	0.61
0.1 sec	3.71	3.90	3.33	2.93	2.00	1.55
0.2 sec	3.07	2.84	2.09	1.64	1.12	1.10
0.3 sec	2.13	2.26	1.70	1.12	0.84	0.76
0.4 sec	1.84	2.11	1.64	1.00	0.68	0.64
0.5 sec	**1.84**	**2.03**	**1.54**	**0.92**	**0.68**	**0.65**
1 sec	**1.84**	**2.05**	**1.62**	**1.18**	**1.22**	**1.00**
1.5 sec	1.94	2.25	2.21	3.31	3.48	2.20
2 sec	2.56	3.05	3.45	6.19	6.15	0.68

Table 2 – CASE 2 - Estimated RT30 vs. presto-chirp duration.

Moreover, setting of the Tukey window (equation 4) function has been experimented and a rule of thumb is suggested below.

$$\begin{cases} \alpha = 0.5, \ 0 < f < 150\,Hz \\ \alpha = 0.3, \ 150 < f < 250\,Hz \\ \alpha = 0.2, \ 250 < f < 750\,Hz \\ \alpha = 0.1, \ 750 < f < 3500\,Hz \end{cases} \quad (6)$$

2.3 Validation

The initial validation of the method was performed by direct comparisons of some characteristic functions like the RIRs and the FRFs. In fact, figure 9 shows the comparison of the estimated RIR and the actual one; figure 10 compares their logarithmic envelopes, for a better resolution for the peaks; figure 11 shows the over plots of the frequency response functions and points out the errors in the joint points of presto chirps.

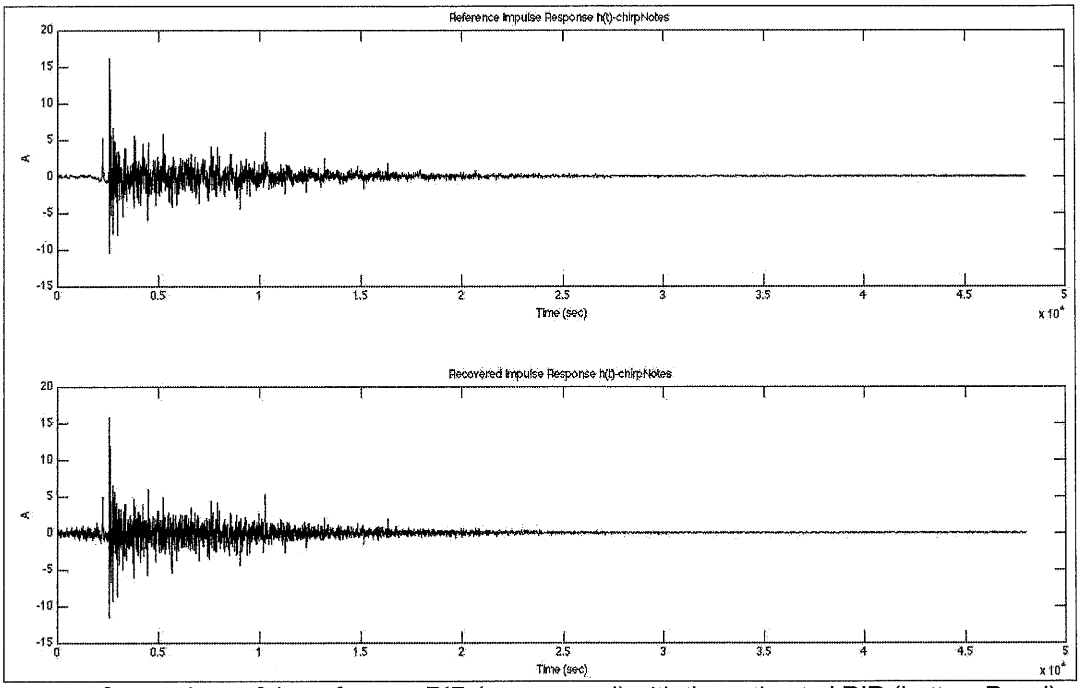

Figure 9 - Comparison of the reference RIR (upper panel) with the estimated RIR (bottom Panel).

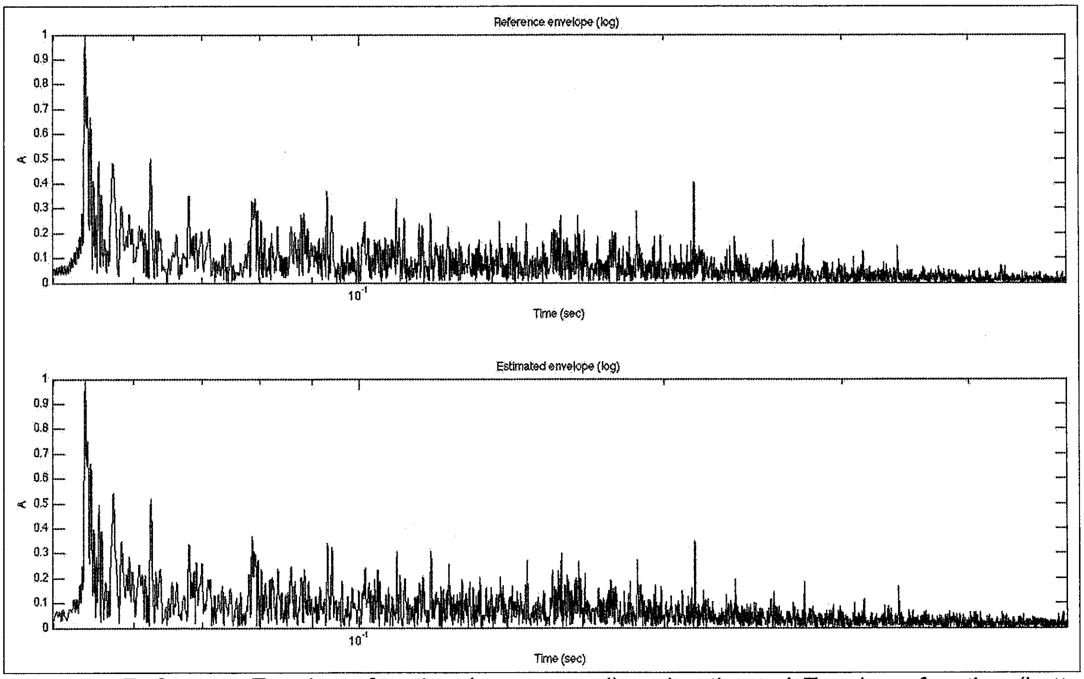

Figure 10 - Reference Envelope function (upper panel) and estimated Envelope function (bottom Panel). Notice that the time axis is in logarithmic scale.

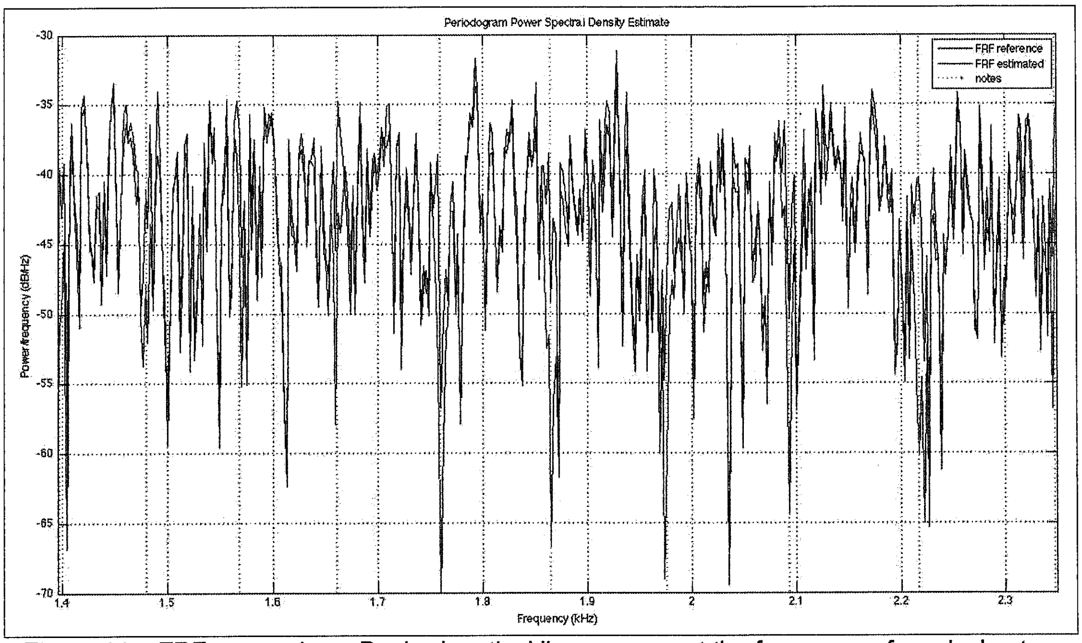

Figure 11 – FRF comparison. Dashed vertical lines represent the frequency of musical notes.

3 CONCLUSION

A novel technique that employs musical stimuli to carry out room impulse response measurement has been present. The method employs narrow band presto-chirps centred on musical notes. These presto-chirps sound like musical notes and therefore can be used to compose synthesised music, unobtrusive signals, to perform acoustic analysis in occupied condition. Further work includes the composition of some suitable pieces of music and to carry out in-situ measurements to further validate the method.

4 REFERENCES

1. M. Serafini and F. F. Li, 'Feasibility of masked pseudo random sequences for occupied measurement in auditoria', Proc. of Forum Acusticum, Aalborg (July 2011).
2. M. Serafini, F. F. Li, 'Impulse response measurement with chirp-lets and masked noise stimuli', Proc. of IOA, Vol. 33. Pt.2, (May 2011).
3. M. Serafini, F. F. Li and T. J. Cox, 'Occupied Measurement of Room Impulse Response-Feasibility and Limitations', Proc. of Inter-noise, (June 2010).
4. F. j. Harris, 'On the use of Windows for Harmonic Analysis with the Discrete Fourier Transform'", Proc. of the IEEE 66 (1): 51–83, (January 1978).
5. A. H. Nuttall, 'Some Windows with Very Good Sidelobe Behavior', IEEE Transactions on Acoustics, Speech, and Signal Processing 29 (1): 84–9, (February 1981).
6. J. Chowning, 'The Synthesis of Complex Audio Spectra by Means of Frequency Modulation', Journal of the Audio Engineering Society 21 (7), (1973).
7. T. K. Karplus, A. Strong, 'Digital Synthesis of Plucked String and Drum Timbres', Computer Music Journal (MIT Press) 7 (2): 43–55, (1983).
8. A. Farina, 'Simultaneous measurement of impulse response and distortion with a swept-sine technique', 110th AES Convention, (February 2000).