IMPULSE RESPONSE MEASUREMENT WITH CHIRP-LETS AND MASKED NOISE STIMULI

M. Serafini Acoustics Research Centre, University of Salford, Greater Manchester, UK
FF. Li Acoustics Research Centre, University of Salford, Greater Manchester, UK

1 INTRODUCTION

Spaces for various purposes may require different acoustics, which is characterised by room acoustic parameters. In fact, some rooms would suit better for a particular kind of music (classic, orchestral, and so on) rather than speech, whereas others would be preferred for speech rather than music. Understanding how sound propagates and how environments modify it is the way to quantified acoustics using purposely-defined parameters. Such parameters are the result of temporal propagation and energy distribution of a stimulus within an enclosure and they are used to interpret other subjective descriptors related to perception[1]. Reverberation time (RT) is generally considered the most useful one as it gives a simple and fast way to define what an environment may be more suitable for (music, speech, and so on). If the stimulus is an impulse (i.e. a Dirac function, which contains ‘all’ the frequencies) the acoustic-space is completely described by its Impulse Response (IR), which when referred to rooms/theatres/auditoria is more properly called room impulse response (RIR). Being an ideal mathematical function, impulse is a purely ideal model. Different methods were developed for RIR measurements, which are able to simulate the use of an ideal impulse. Commonly, those methods are named after the signal used to perform the measurement. There are then, impulsive measurements (which use impulsive stimuli originated from shotguns, balloons, clappers, and so on); FFT-based measurement, which uses white or pink noise as stimuli (more generally, any kind of stimuli could be used) and the RIR is extracted from the Frequency Response Function (FRF), which is calculated by averaging the complex (or real) ratio between output and input of a system/room under test. In the second half of seventies, Schroeder[2] pioneered the use of pseudo-random sequences for RIR measurement, which may give a RIR measurement without using any averaging (although two cycles are needed due to the circular convolution), making it faster than previous methods. Moreover, thanks to the Fast Hadamard Transform (FHT) the method could be implemented even in low-performance computers - that was a non-negligible factor at that time. However, constrains due to background noise during the measurements, bounds the Signal to Noise Ratio (SNR) and force the MLS-method to adopt an averaging technique to cope with noise problem. It has been investigated that prolonged averages increase the method’s susceptibility to time variance and this may lead to an erroneous estimation of the RIR. For that reason, averaging techniques should not be applied in time-variant systems.

More than twenty years after the proposal of the MLS method, a new measurement technique named Exponential Sine Sweep (ESS) and also known as logarithmic chirp or simply chirp, started to flourish, and it has been demonstrated thus far, superior to its predecessor methods. Such method was developed by Farina[3], whom while researching to solve MLS limitations by employing the Time Delay Spectrometry method (TDS[4]), noticed that by using exponential frequency-varying sinusoidal signals and by employing a linear de-convolution approach instead of the more commonly used circular one, it was possible to distinctly calculate each individual order of a weakly non-linear impulse response accordingly to Volterra’s kernel formulation. This has enabled the calculation of a distortion-free RIR (linear response) and all distortion orders (harmonic distortions) with just one measurement. Nowadays, the MLS and the ESS are the common methods used to perform room acoustic measurements. Although the ESS is arguably superior to MLS, they share popularity among researchers and practitioners[5].

However, there are several constraints related to the applicability of these methods for occupied measurement, among which the stimuli obtrusiveness is the major problem. An audience, attending for a concert, will not normally tolerate such noisy signals, and the use of large numbers of voluntary or paid people is not realistically feasible. Nowadays, occupied values are predicted from unoccupied RIR measurements, but this is neither an in-situ measurement technique nor a reliable solution. Hidaka et al.[6] used the so-called stopped chord method to extracts the RT and few others acoustical parameters, from the Energy Decay Curve (EDC) recorded after the loud stop chords by an orchestra during a live concert. Despite the fact that the method gives a better estimation of some acoustic parameters in occupied conditions, it does not lead to the calculation of the RIR. Although music might be used as a stimulus to calculate the FRF in in-use environments (e.g. concert venues) by means of the well-known dual-channel FFT technique[7], it does not completely solve the problem. Music is neither frequency-continuous nor necessarily time-continuous signals. Its power spectrum distribution (PSD) is far from being homogeneous along the audio-frequency band, and this would impose uneven SNRs across the frequency range. In addition, some frequency might be very low or even negative making the measurement impossible or incomplete[8]. Despite this artefact, some audio-companies are developing devices and software that do use of the dual FFT techniques to calculate and/or check the FRFs, but mainly for the electro-acoustic chains used in live concerts (e.g. Smaart V7). There are ongoing attempts to solve occupied measurement problems[9]. These methods estimate acoustic parameters from the EDC, but they do not give complete RIRs. The research presented in this paper aims to develop a procedure to enable occupied measurements of complete impulse responses by overcoming the intrusiveness of the testing signals. The authors started from experimenting with the use of very Low-level MLS signals as suggested by Schroeder[10]. Because it is known that, theoretically, MLS may work under minus SNR by using prolonged averaging. Results showed that the required measurement duration is not practically feasible: long before a require duration was reached; the time variance had ruined the results[11]. We then experimented with the stepped tone method[12] centred on musical notes, with the hope that such tones can be used as notes for synthesised music to make the test signal musically meaningful. However, the phase response of a RIR is very complicated and is often deemed as being random. If only sparse frequency points were measured, the RIR could not be reconstructed accurately. These initial investigations led to the proposal of a new hybrid method detailed below.

2 PROPOSED HYBRID METHOD

In this paper a new hybrid method for the occupied RIR measurements is developed. It relies on the use of two different methods; each used for a specific audio frequency region. The first method uses short linear sweep signals, with ranges covering consecutive semitones in equal temperament music scales. That is to say, each semitone is represented by a short chirp centred on it. This would ensure the continuous frequency spectra of excitation up to 4 kHz if 88 musical notes were all used. For frequencies above 4 kHz RIRs are measured with MLS signals masked by recorded and (dynamic range) compressed music. The hybrid method allows for higher SNR than that obtained by the use of a single method, whilst the test stimuli are perceived as music. The excitation signals are formed by a set of chirps with each having a slightly increased bandwidth, (ranged from circa 20 Hz for the smaller chirp up to circa 200 Hz to depending upon the pitch,) as shown in Figure 1, this is to give necessary overlap between these chirp-Notes and maintain the even energy distribution in the frequency domain.

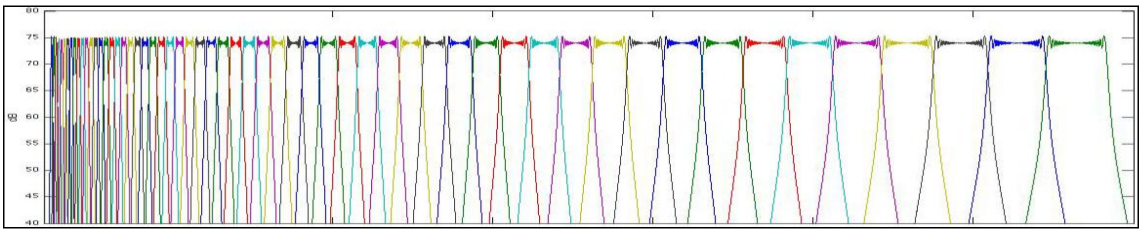

Figure 1 – Corresponding filter banks generated from narrow band chirps.

Energy differences due to the increasing bandwidth were normalized by a multiplicative factor proportional to the square root of each chirp's bandwidth. Window functions attenuate Gibbs phenomenon that occurs at step-transitions. In order to check the goodness of the overall stimulus, the chirp-Notes (or chirplets in a more scientific term) were summed and its auto-correlation analyzed. Both auto-correlation (upper panel) and auto-spectra (bottom panel) are depicted in Figure 2, which clearly show artefacts in both time and frequency domains.

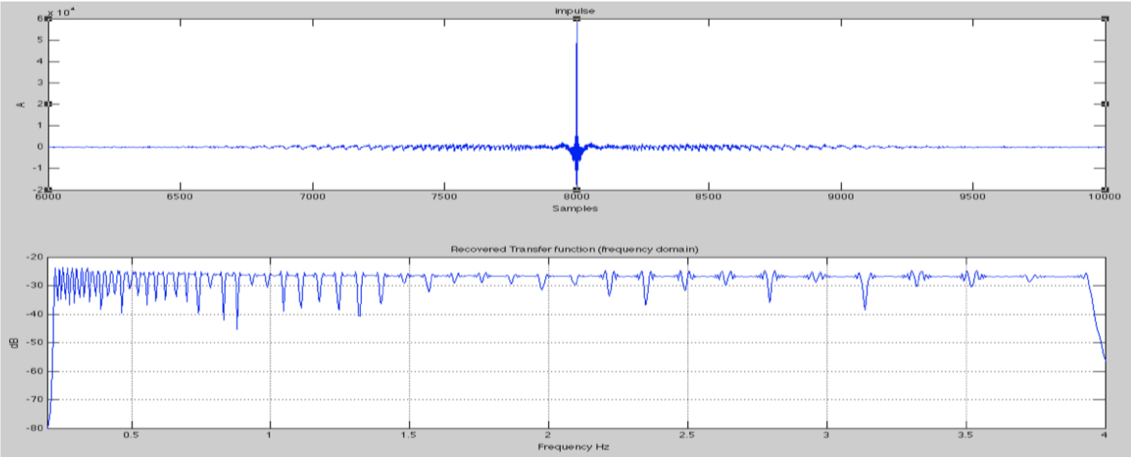

Figure 2 – Auto-correlation (upper panel) and Auto-spectra (bottom panel) of the summed stimulus.

As suggested by Farina[13] applying a fade-in and a fade-out amplitude envelope to the excitation signal, could reduce the "ringing effect", but the energy at the edge frequencies would be slightly altered as well. For the continuous ESS analysis this will not cause major problems since the bandwidth of the continuous chirp can always be made greater than the required frequency range for measurements. However, in our proposed method the excitation signals could be viewed as 'chopped' parts of a long chirp. All the ringing distortions, if not properly attenuated will cause significant distortions in the final lumped RIR. For this reason the choice of the most appropriate windows and the development of a sophisticated merging algorithm are not easy tasks. Initially, a trapezoidal window was experimented, which emulates a fade in and out at the start and at the end of the signals, while the energy in the middle part of the window is not altered. Further investigation shows that the Tukey window surpasses trapezoidal one because of lower side-lobe magnitudes. Chirp-Notes were constructed by linear frequency-varying sweeps, as a convincement choice, since the orders of nonlinear distortions inherited in the measurement systems (e.g. loudspeakers) are not of particular interest in this case. Simulations have been used to validate the feasibility of the method and details are given in the next section. In-situ measurement are being planned to identify practical problems. Some subjective tests were conducted to identify if the chirp-Notes are acceptable to audience or if they are music worthy. The chirp-Notes sound like woodwind instruments in general. We noticed that when increasing the bandwidth of the chirplets the pleasantness of them decreases. Moreover, the highest pitch of musical notes would anyway be limited at a frequency slightly over 4 kHz. The use of harmonics (or over tones) is a likely approach but may not give good SNRs. For these reasons, we developed a second method, which could be used to measure the RIRs in the frequencies above 3.5 kHz/ 4 kHz. This was achieved using a low-level MLS as a probe.

In order to avoid annoyance, the MSL signal was masked by a low-pass filtered orchestral music signal (with *fc* = 3.5 kHz). On one hand, this gives enough SNR in those frequencies beyond the *fc* from MLS measurements, and on the other hand, the masking music is yet an enjoyable entertainment, although the resulted audio-quality is to some extent reduced by the low pass filter. Masking effects of louder narrowband signals are well known in literature[14]. The masking effect of a tone centred at 1 kHz and having different amplitude values is depicted in the picture below.

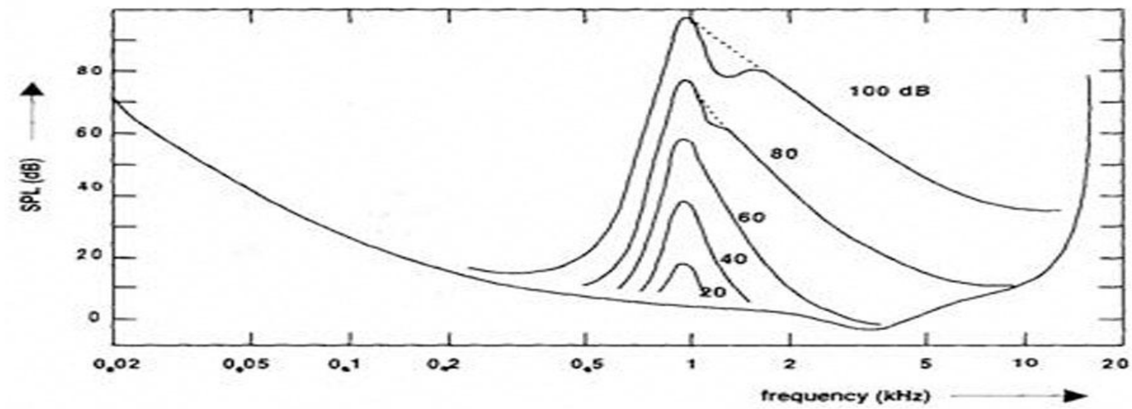

Figure 3 - Masking effects of a 1 kHz tone. The bottom line represents the audibility threshold. (From E. Zwicker, H. Fastl, Springer-Verlag; "Psychöacoustics, Facts and Models", 1990)

It can be seen from the masking curve that to obtain a satisfactory masking, i.e. to make the low-level MLS signal almost unnoticeable, the masking signal level must be sufficiently high. Moreover, there is a second factor that has to be considered that is related to the background noise. It gives the starting SNR of the measurement and hence it can be used to identify the total measurement time needed to complete a RIR measurement achieving a wanted SNR.

To validate, our hypothesis we carried out an experiment in a recording studio room. The background noise was very low because it represents a un-weighted background noise in high frequency bands rather than lower frequencies or broadband ones, however we considered auditorium or concert hall settings, where the background noise levels are generally of about 35 dB(Z). We set the average level of the masking music at about 93 dB(Z), which in our case corresponded to 85 dB(A). In live concerts such value can be higher. All the levels were measured by means of a sound level meter. We then performed a set of subjective test, using acoustics research students. It was found that a 63 dB(Z) MLS could be well tolerated while masked by an orchestral music passage having the aforementioned level and if the MLS level drop 20 dB further down, i.e. 43 dB(Z), the MLS became barely audible. Figure 4 below represents the spectrogram of a typical scenario from our experiments.

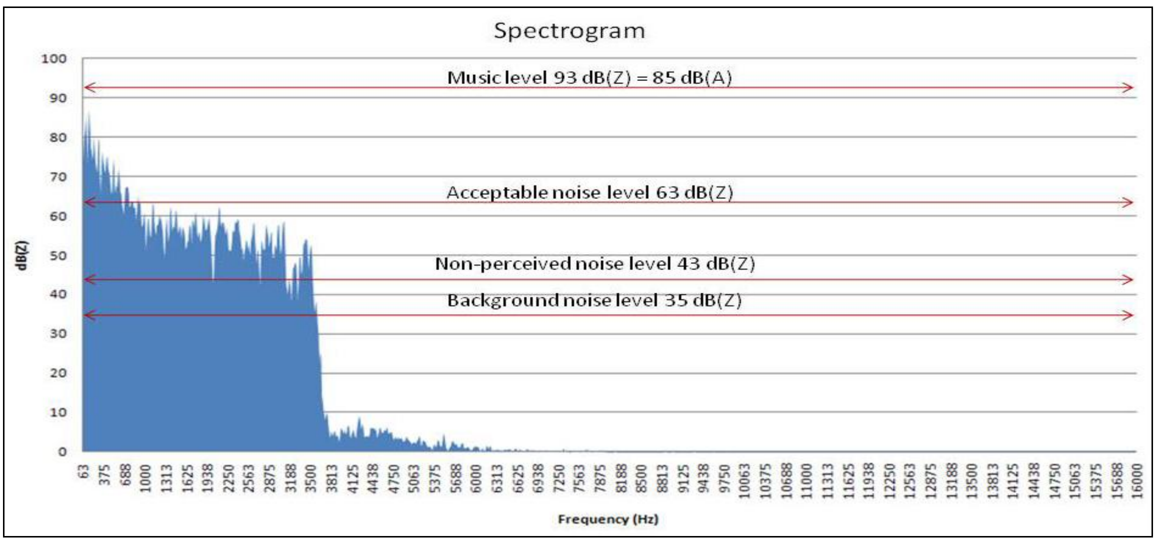

Figure 4 – Spectrogram of a low-pass filtered music being played in the room under test. Four horizontal lines are used to indicate the level of: the music, the acceptable MLS, the inaudible MLS and the background noise, respectively.

We did use of a 17th order MLS that gave an intrinsic improvement of the SNR of 17 dB itself[15]. It means that the actual SNR after one complete MLS cycle (It should be remembered that a pre-feed cycle is always used regardless the averaging), was of 63-35+17=45 dB(Z). If the level of MLS is set to an inaudible level the resulting SNR is of 25 dB(Z). Moreover, synchronous averaging a deterministic signal, such as the MLS, gives a 3 dB increment of the SNR for each doubling. This is because coherent signals add together whereas uncorrelated noise will be averaged out. As a 17th MLS order sampled at 44.1 kHz is almost 3 seconds long, a RIR analysis using 32 averages will be completed in about 96+3 seconds (3 seconds are due to the pre-feed cycle), and giving 60 dB of SNR when tolerant MLS level is used, and giving 40 dB of SNR when an inaudible MLS level is used. In the latter case only RT30 can be calculated. These results are reported in the table below.

Averages	*Time(sec.)*	*SNR (dB) inaudible*	*SNR (dB) acceptable*
Nill	6	8+17=25	28+17=45
2	9	28	48
4	15	31	51
:	:	:	:
32	99	40	60
64	195	43	63

Table 1 – Measurement time from a known start value of SNR in two cases: acceptable level and inaudible level. A 17th order MLS sampled at 44.1 kHz is considered (circa 3 second long).

3 SIMULATION AND PILOT VALIDATION VIA MEASUREMENTS

Simulations were carried out using MATLAB software, and they were done by estimating a known RIR by means of the proposed hybrid method. The chirp-Notes were ‘played’ in a random mode (simple permutation), both ascending and descending chirplets were used. Figure 5 shows a spectrogram of the chirp-Notes. Each chirp-Note was of about one second long.

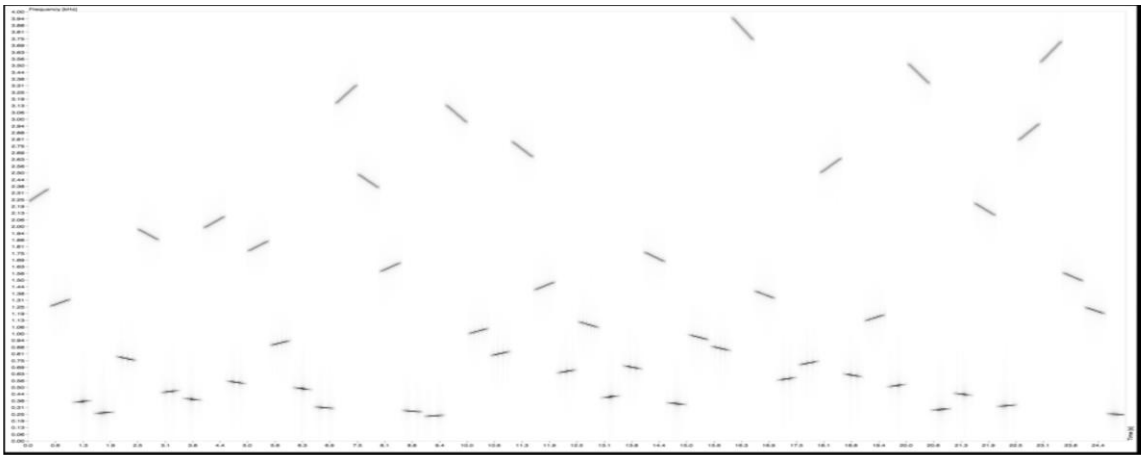

Figure 5 - Chirp-Notes spectrogram.

De-convolving each chirp-Note by its corresponding matching filter gives a set of narrow band impulse responses. Summing those band limited impulse responses based on linearity assumption, gives a band limited RIR - from 125 Hz up to 3.5 kHz - as shown below in Figure 6.

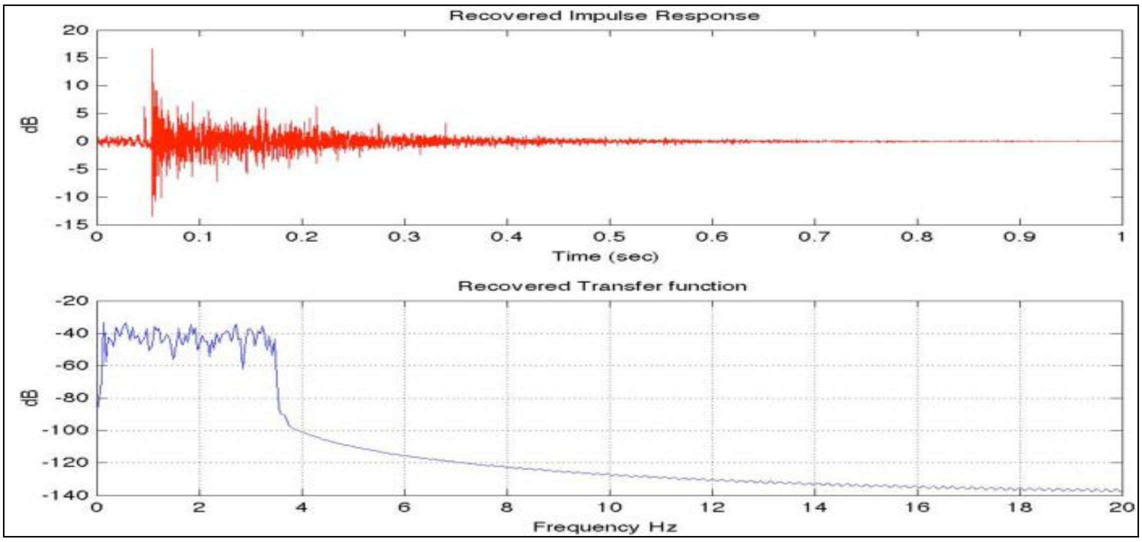

Figure 6 - RIR obtained using the chirp-Notes method via simulation.

The ringing effects introduced by the narrow band chirp-Notes cause the appearance of a slightly modulated background noise, which limits the noise-free EDR though it has been demonstrated that by using a Tukey window it gives more than 50 dB of EDR, which is enough for the estimation of RT30 as recommended by the ISO 3382:1997 standard. On the other hand, in the frequency domain the window function lowers the energy of edge frequencies, accordingly to the window shape. Therefore a more sophisticated algorithm is needed to optimize the merging procedure. The RIR calculated from masked-MLS method has to be high-pass filtered (at fc). In this way, the filtered RIR contains only the frequencies above the fc, giving a SNR sufficient for a reliable estimation of the acoustic parameters. A 50 dB noise-free range in energy decay curve is typically achievable from the masked-MLS method. Figure 7 below shows RIR and frequency response obtained using the masked method above fc.

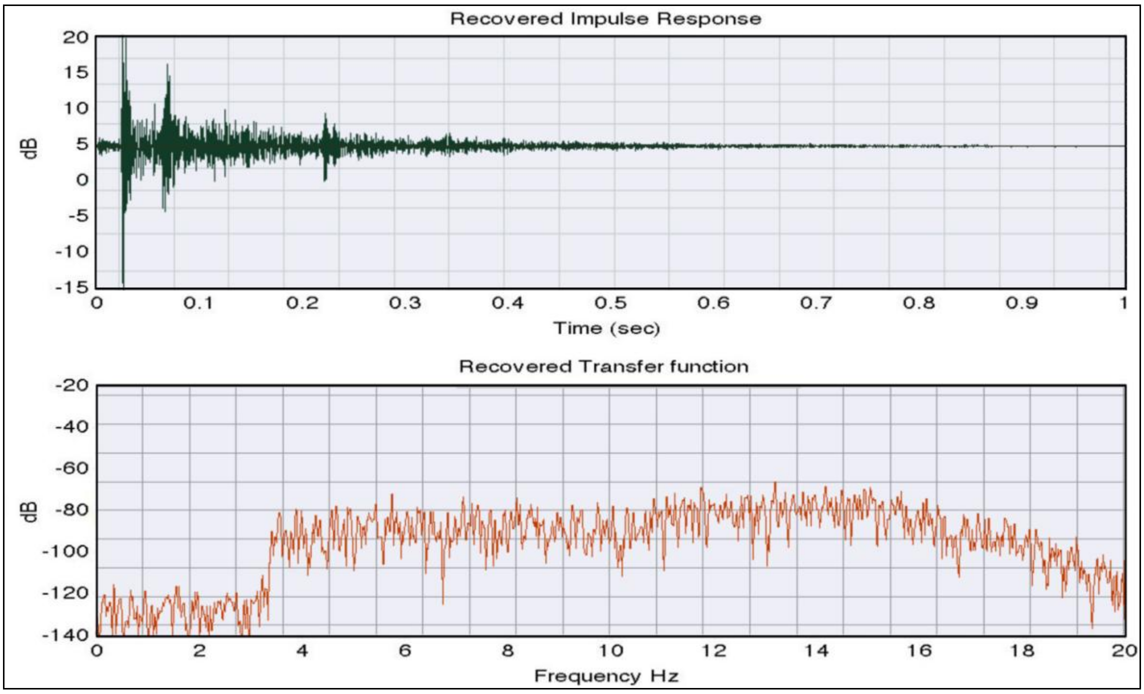

Figure 7 - Time and frequency domains of a measured and filtered RIR (fc = 3.5 kHz).

The final step of the proposed Hybrid method is to join the two measured RIRs in order to form a final broadband RIR. The picture below shows the accuracy of the 'mixing' procedure by compare a reference RIR and the one obtained by the proposed hybrid method.

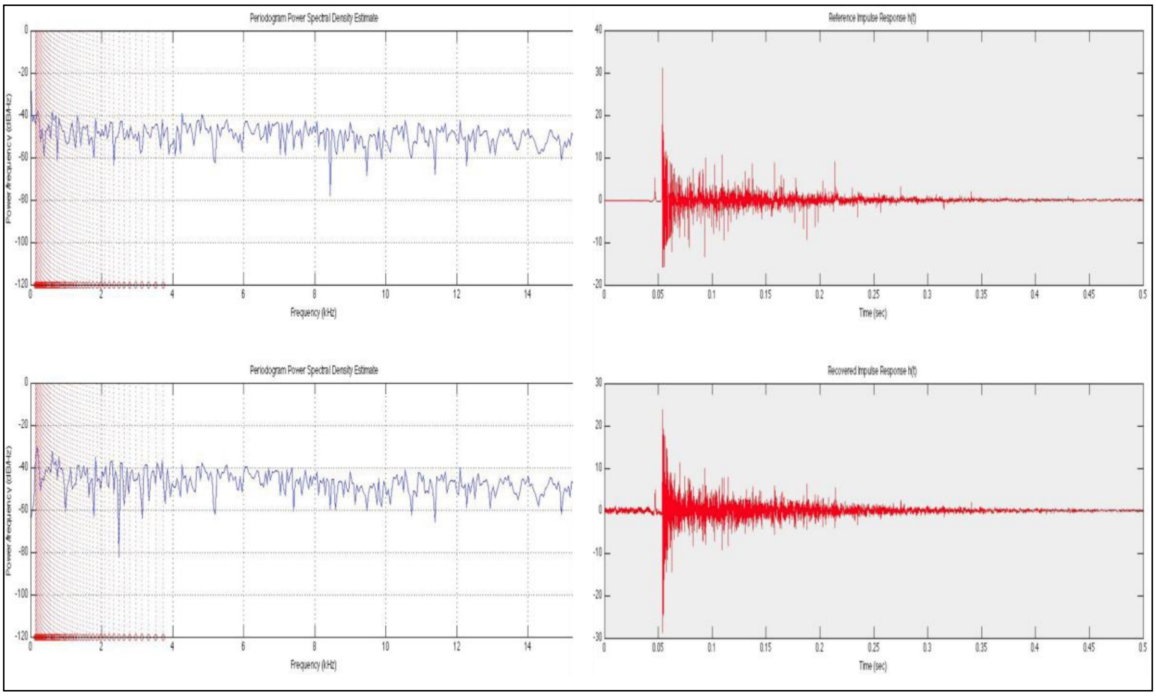

Figure 8 - The upper panels show the reference FRF and RIR, respectively, whereas the bottom panels show the calculated ones via the hybrid method. The dashed vertical lines in the left panels correspond to joining points of the chirp-Notes method and the masked MLS method.

The FRFs (left panels) present at junction points some inaccuracies, and at the same time, the RIRs (right panels) show the aforementioned modulated background noise. Despite the methods needs further improvements and full validations by real measurement, the estimation of RT30 values showed a good accuracy as reported in Table 2.

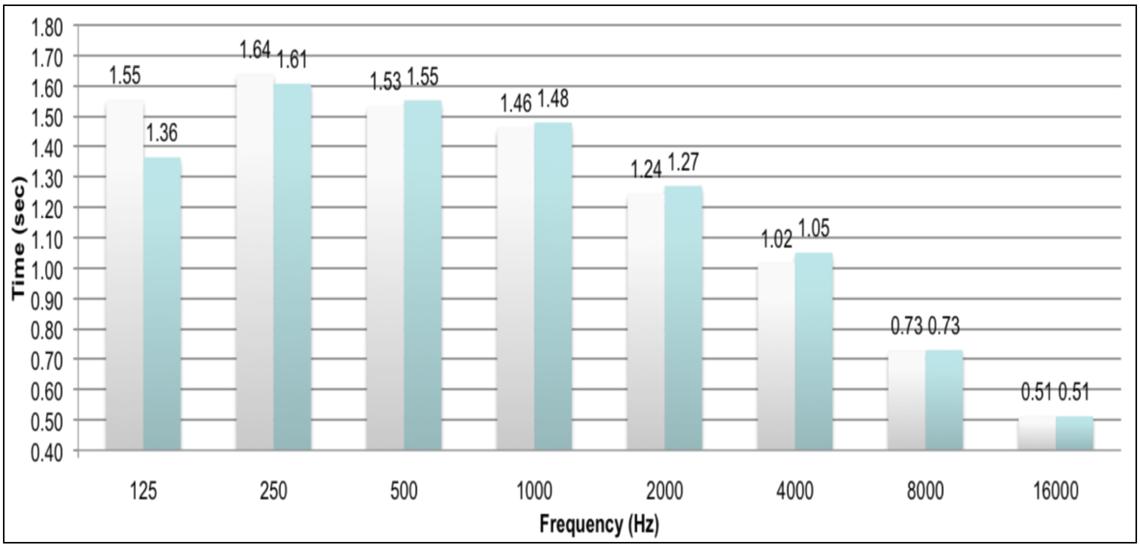

Table 2 – Reverberation times (RT30) comparison – White bars = reference values; Blue bars = estimated values.

4 DISCUSSIONS

To mitigate the difficulties of traditional methods in occupied RIR measurements is the motivation of this research, which has led to the development of a new hybrid method. The proposed method consists of two complementary measurement techniques. The first technique uses narrow-band linear chirps, named Chirp-Notes or Chirplets, to analyse the frequencies between 125 Hz and 3.5 kHz / 4 kHz. The second technique uses a MLS signal masked by a piece of orchestral music to calculate the occupied RIR in the remaining frequency range, thus from 3.5 kHz / 4 kHz up to 20 kHz. The chirp-Notes method permits to construct a test signal that consists of ‘glissando’ sounding notes and they can be used to compose synthesised music as stimuli. Simulations have confirmed the feasibility of such approach and the method is being experimented for real measurement in order to solve practical problems. On the other hand, the masked-MLS method can ensure the measurement of higher frequency in a reasonable time and with a reduced annoyance to audience. This was confirmed by real measurements in this study. Psycho-acoustic tests were also used to define an acceptable MLS level that an audience might tolerate with respect to the level of the masking music signals. It has been proved by experiments that the joint RIRs from the hybrid method can typically obtain a more than 50 dB clear decay range free of noise (from energy decay curve), which is sufficient for the measurements of most of the room acoustic parameters.

5 REFERENCES

1. ISO 3382:1997, Acoustics — Measurement of the reverberation time of rooms with reference to other acoustic parameters. (1997).
2. M. R. Schroeder, New method of measuring reverberation time, J. Acoust. Soc. Am., Vol. 37, pp. 409-12. (1965).
3. A. Farina, Simultaneous measurement of impulse response and distortion with a swept-sine technique, 108th AES Convention, Paris. (February 2000).
4. M.A Poletti, Linearly Swept Frequency Measurements, Time-Delay Spectrometry, and the Wigner Distribution, J. AES, Vol. 36, No. 6, pp. 457-468. (1988).
5. G.B. Stan, J.J. Embrechts, D. Archambeau, Comparison of Different Impulse Response Measurement Techniques, J. AES, Vol. 50, pp. 249-262. (2002).
6. T. Hidaka, N. Nishihara and L. Beranek Relation of acoustical parameters with and without audiences in concert halls and a simple method for simulating the occupied state, J. of Acoust. Soc. Am., Vol. 109, No. 3, pp. 1028-1042. (1998).
7. D. Griesinger, Beyond MLS – Occupied Hall Measurement with FFT Techniques, 101st AES. (November 1996).
8. F. Jacobsen and D. Bao, Acoustic Decay Measurements with a Dual Channel Frequency Analyzer, J. Sound and Vibration, Vol. 115(3), pp. 512-537. (1987).
9. P. Kendrick, F. F. Li, T. J. Cox, Y. Zhang and J.A. Chambers, Monaural Room Acoustic Parameters from Music and Speech, J. Acoust Soc Am., Vol. 124(1), pp. 278-287. (2008).
10. M. R. Schroeder, Integrated impulse method measuring sound decay without using impulses, J. of Acoust. Soc. Am., Vol. 66(2), pp. 497-500. (1979).
11. M. Serafini, F. F. Li and Trevor J. Cox, Occupied Measurement of Room Impulse Response - Feasibility and Limitations, Proc. Internoise conference. Lisbon (2010).
12. J. Schoukens, R. Pintelon and Yves J. Rolain, Broadband Versus Stepped Sine FRF, IEEE, Vol. 49, Number 2, pp. 275-278. (2000).
13. A. Farina, Advancements in Impulse Response Measurements by Sine Sweeps, 122nd AES, Vienna. (May 2007).
14. B.C.J. Moore, Frequency analysis and masking, in Hearing, B.C.J. Moore, ed., Academic Press, Orlando, Florida. (1995).
15. M. Vorländer and M. Kob, Practical Aspects of MLS Measurements in Building Acoustics, Applied Acoustic, Vol. 52(3-4), pp. 239-258. (1997).