



# Audio Engineering Society Convention Paper

Presented at the 129th Convention  
2010 November 4–7 San Francisco, CA, USA

*The papers at this Convention have been selected on the basis of a submitted abstract and extended precis that have been peer reviewed by at least two qualified anonymous reviewers. This convention paper has been reproduced from the author's advance manuscript, without editing, corrections, or consideration by the Review Board. The AES takes no responsibility for the contents. Additional papers may be obtained by sending request and remittance to Audio Engineering Society, 60 East 42<sup>nd</sup> Street, New York, New York 10165-2520, USA; also see [www.aes.org](http://www.aes.org). All rights reserved. Reproduction of this paper, or any portion thereof, is not permitted without direct permission from the Journal of the Audio Engineering Society.*

## Methods for Extending Room Impulse Responses Beyond Their Noise Floor

Nicholas J. Bryan<sup>1</sup>, Jonathan S. Abel<sup>1</sup>

<sup>1</sup>Center for Computer Research in Music and Acoustics (CCRMA), Stanford University, Stanford, CA, 94305, USA

Correspondence should be addressed to Nicholas J. Bryan ([njb@ccrma.stanford.edu](mailto:njb@ccrma.stanford.edu))

### ABSTRACT

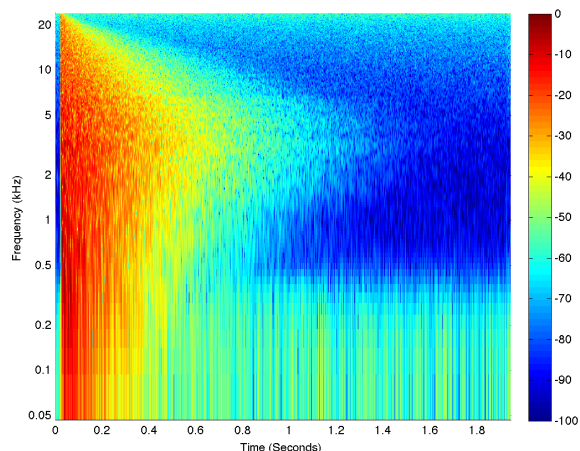
Two methods of extending measured room impulse responses below their noise floor and beyond their measured duration are presented. Both methods extract frequency-dependent reverberation energy decay rates, equalization levels, and noise floor levels, and subsequently extrapolate the reverberation decay towards silence. The first method crossfades impulse response frequency bands with a late-field response synthesized from Gaussian noise. The second method imposes the desired decay rates on the original impulse response bands. Both methods maintain an identical impulse response prior to the noise floor arrival in each band and seamlessly transition to a natural sounding decay after the noise floor arrival.

### 1. INTRODUCTION

Room reverberation is often simulated using convolution with a measured room impulse response (IR). The quality of such measured IRs depends on the measurement signal-to-noise ratio (SNR) or noise floor. The noise floor limits the perceptual quality of reverberant impulse responses and is invariably undesirable. Because of this, numerous robust impulse response measurement methods have been developed to increase the measured SNR, such as sinusoidal sweeps and pseudo-random noise se-

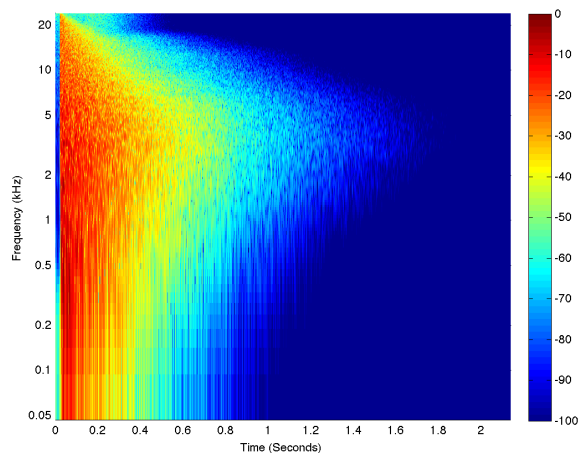
quences, by increasing the test signal energy applied to the space [1, 2, 3, 4, 5, 6]. Even so, in typical measured room IRs, the high-frequency energy decays faster and arrives at the noise floor earlier than does low-frequency energy. This leads to unnaturally emphasized high-frequency content in the late-field tail. Moreover, room decay rates are typically longer in the low frequencies, and unwanted low-frequency measurement noise often corrupts room impulse response measurements as seen in Fig. 1.

When preparing such a measured impulse response for use in a convolutional reverberator, it is useful



**Fig. 1:** EMT 140 Impulse Response Spectrogram. A spectrogram of an EMT 140 plate reverberator in the presence of a noise floor prominent in the low and high frequencies.

to extrapolate the impulse response decay to below its noise floor using the methods presented here as shown in the example of Fig. 2. In addition, there



**Fig. 2:** Extended EMT 140 Impulse Response Spectrogram. A spectrogram of an EMT 140 plate reverberator extended to below its measured noise floor.

are scenarios in which it is desirable to extend an impulse response beyond its measured length. Furthermore, various methods of artificial reverbera-

tion synthesis have found noise floor extension useful [7, 8]. All such issues motivate the developed methods to help eliminate the noise floor of room impulse responses.

In this work, we present two methods of extrapolating the late-field of a room impulse response below its noise floor and beyond. The idea is to extract the energy equalizations, energy decay rates, and noise floor levels which can then be used to model and synthesize the late-field reverberation. Both methods similarly model the late-field response, but differ on how the synthetic late-field response is constructed.

The first method uses a constant-power crossfade to fade out the measured room response before the detected noise floor and fade in synthetically generated statistically independent Gaussian noise to match. The second method imposes the desired decay rates upon the original impulse response bands, essentially “bending” the original noise floor to maintain constant band energy decay rates to perform an all natural extension.

Other methods based on sample rate conversion for naturally extending reverberation have been described in [9]. However, these methods do not explicitly reduce the noise floor level. Both of the proposed methods allow not only extensions of room impulse responses beyond their measured noise floor, but can remove unwanted transients found within the measured late-field.

An overview of late-field reverberation with respect to this work is found in §2, methods for analysis and noise floor estimation in §3, and extension methods in §4. Results and conclusions are found in §5 and §6 respectively.

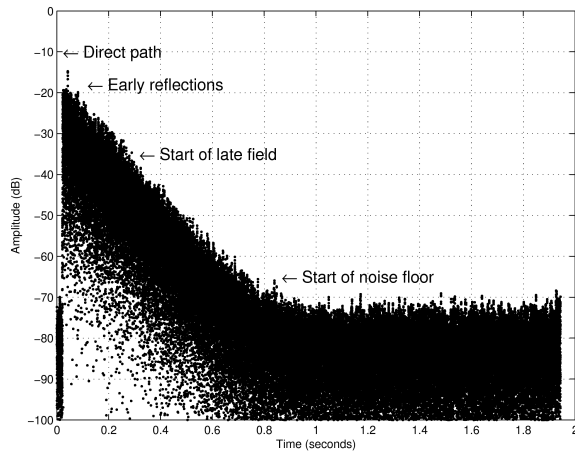
## 2. LATE FIELD REVERBERATION OVERVIEW

A room impulse response is commonly modeled as having early and late-field responses. The early response encompasses the direct path arrival and early reflections and is followed by the late-field [10, 11, 12, 13, 14]. The late-field response is modeled as Gaussian noise, with an evolving equalization reflecting the frequency-dependent decay rates [10, 13]. A frequency band energy profile is modeled as a constant noise floor plus a decaying exponential in each of a

set of bands  $k$ ,

$$\beta_k(t; \theta) = \sigma_k^2 + \gamma_k^2 e^{-2t/\tau_k}, \quad (1)$$

where the parameters  $\theta = [\sigma^2 \ \gamma^2 \ \tau]^T$ , are  $\sigma^2$ , the noise floor variance,  $\gamma$ , the initial late-field gain and  $\tau$ , the late-field time constant. Fig. 3 depicts a measured room IR over a narrow frequency band following such a model.



**Fig. 3:** Impulse Response Frequency Band. The time domain of a sample impulse response with annotations highlighting the reverberation features.

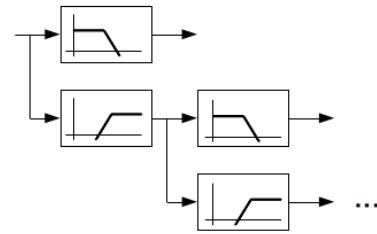
### 3. ANALYSIS AND ESTIMATION

A perfect amplitude reconstruction zero-phase filter bank is used to split the late-field response  $h(t)$  into frequency bands  $h_k(t)$ . The filter bank is constructed via a cascade of squared Butterworth filters as shown in Fig. 4 [15]. The magnitude transfer functions of the filters are shown in Fig. 5.

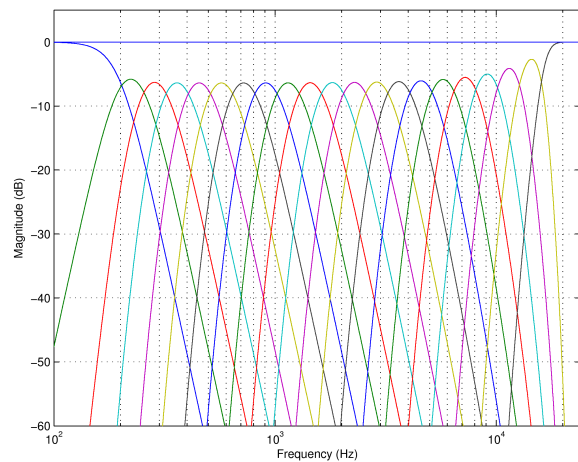
Once separated into bands, frequency-dependent energy profiles can be computed by smoothing the square of  $h_k(t)$  over a running window,

$$\tilde{\beta}_k(t) = h_k(t)^2 * w(t). \quad (2)$$

where  $w(t)$  is a positive smoothing window with unit sum,  $\sum_t w(t) = 1$ . After energy profiles are computed, the necessary parameters  $\theta$  can be estimated. Two methods of parameter estimation



**Fig. 4:** Filter Cascade Block Diagram. An illustration of the filter bank cascade structure.



**Fig. 5:** Filter Bank Transfer Function Magnitudes. The transfer function magnitudes of the analysis/synthesis filter bank with third-octave center frequencies.

are suggested, motivated by the two extension methods presented.

#### 3.1. Synthetic Extension Analysis

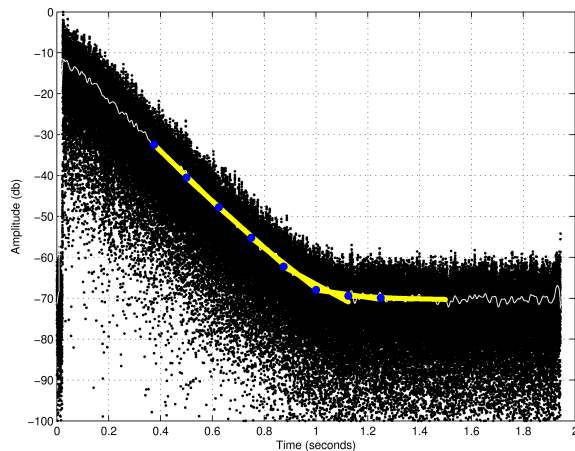
The first method is motivated by the synthetic extension method and explicitly estimates frequency dependent equalization levels, energy decay rates, and noise floor arrival times. To do so, the band noise floor arrival times  $t_n$  for each band are estimated, followed by a least squares estimation of the equalization levels and decay rates.

The noise floor arrival time is computed via a slid-

ing least-squares estimate on the decibel level of the smoothed energy profile. The noise floor is detected as the energy decay begins to lengthen and the estimated time constant reaches a prescribed upper threshold, say, corresponding to a 20 second  $T_{60}$ . For each time window  $\mathbf{t}$  of the sliding estimate, the optimal parameters  $\hat{\theta}$  are estimated via

$$\operatorname{argmin}_{\theta} \| 10 \log_{10}(\beta(\mathbf{t}; \theta) - \sigma^2) - 10 \log_{10} \tilde{\beta}(\mathbf{t}; \theta) \|_2^2,$$

assuming  $\tilde{\beta}(\mathbf{t})$  is above the noise floor. An illustration of the estimation process is shown in Fig. 6, where the start of each estimation window and the corresponding fit is displayed with a dot followed by a straight line.



**Fig. 6:** Noise Floor Arrival Time Estimation. Estimation of the noise floor via a sliding least squares estimate. The energy profile (thin line) is shown along with the successive estimates (thick lines).

Once the noise floor time of arrival is found, updated estimates of  $\tau$  over the length of the late-field and  $\gamma$  over a short window circa  $t_n$  are computed. The updated  $\tau$  and  $\gamma$  estimates, along with the explicit estimation of the noise floor arrival time, are motivated by the synthetic extension method presented below in §4.1.

### 3.2. Natural Extension Analysis

The second method of extension explicitly estimates frequency dependent equalization levels, energy decay rates, and noise floor levels. In this method, all three parameters  $\hat{\theta}$  are estimated simultaneously via

$$\operatorname{argmin}_{\theta} \| 10 \log_{10} \beta(\mathbf{t}_m; \theta) - 10 \log_{10} \tilde{\beta}(\mathbf{t}_m; \theta) \|_2^2.$$

This alternative estimation method is motivated by the natural extension method presented in §4.2.

## 4. EXTENSION METHODS

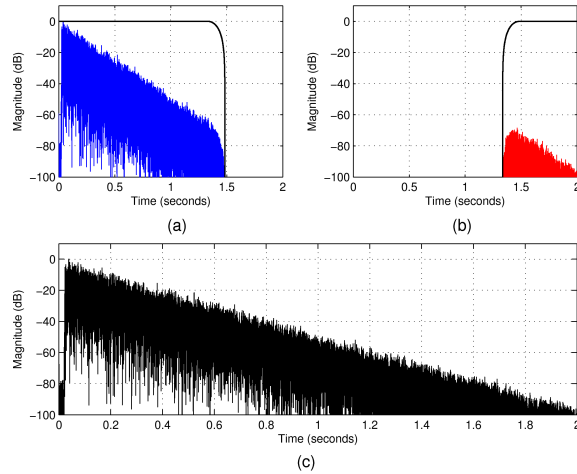
Given that the late-field reverberation is modeled as colored Gaussian noise, the late-field can be extended by extrapolating the energy profile decay rates. Two methods of extension are presented. The first method uses a constant-power crossfade to stitch in synthetically generated statistically independent Gaussian noise to match the original measured room IR. The second method leverages the unwanted noise floor by essentially “bending” down the noise floor to extrapolate a constant energy decay rate.

### 4.1. Synthetic Extension Synthesis

To fade in a synthetically generated late-field, we first generate a single Gaussian noise sequence  $n(t) \sim \mathcal{N}(0, \sigma^2)$ . The noise sequence is put through the analysis filter bank to generate a set of noise bands  $n_k(t)$ . The noise bands are then exponentially windowed to impose the equalization levels  $\gamma_k$  and decay rates  $\tau_k$  from the measured impulse response

$$\lambda_k(t) = \hat{\gamma}_k e^{-t/\hat{\tau}_k}. \quad (3)$$

The windowed bands are crossfaded with the input room IR bands  $h_k(t)$  immediately prior to the noise floor arrival times  $t_n$ . Because both the original IR and synthetically generated noise are assumed to be statistically independent, a power-complementary sine window is used for the crossfade. Fig. 7 illustrates the crossfade on a dB scale. Once each band of the input impulse response is extended, the bands are summed to produce the complete



**Fig. 7:** (a) Windowed Measured Impulse Response. (b) Windowed Synthesized Late-field. (c) Cross-faded, Extended Late-field Reverberation.

extended room impulse response.

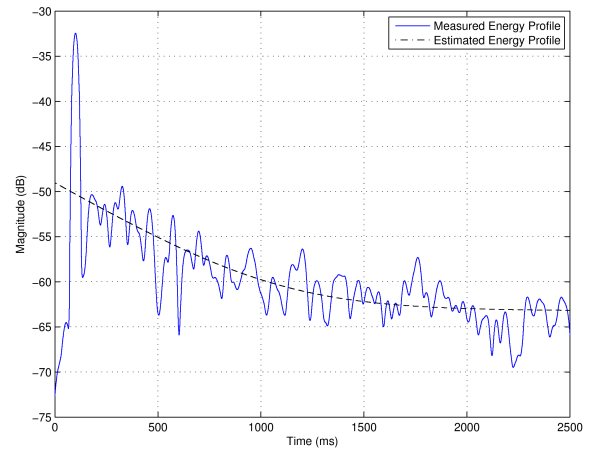
#### 4.2. Natural Extension Synthesis

The second method of extension windows the original noise floor signal to extend the IR below its noise floor. This method preserves the natural noise sequence found in the original impulse response by windowing the measured late-field bands according to a window  $\lambda_k$  defined as

$$\lambda_k(t) = \frac{\hat{\gamma}_k e^{-t/\hat{\tau}_k}}{[\hat{\gamma}_k^2 e^{-2t/\hat{\tau}_k} + \hat{\sigma}_k^2]^{1/2}}. \quad (4)$$

The denominator of the windowing equalizes, over time, the measured energy profile, whereas the numerator acts to extrapolate the desired decay rate and equalization level. Note that the window is  $\approx 1$ , except as the level approaches the noise floor level.

Fig. 8 displays the measured and estimated energy profile that is used to compute  $\lambda(t)$ . This natural extension leverages the fact that the unwanted noise floor typically maintains the necessary noise statistics to match the late-field response, but with an undesirable windowing that may be replaced. Note that this technique may be used with synthetically



**Fig. 8:** Measured Energy Parameter Estimation. The model  $\sigma_k^2 + \gamma_k^2 e^{-2t/\tau_k}$  fit to a low-frequency measured impulse response band.

generated noise added to the original impulse response late-field above the noise floor.

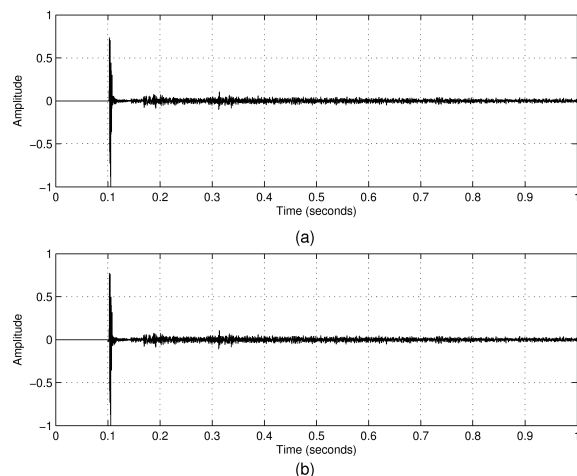
#### 5. RESULTS

Based on informal listening tests, both extension methods were seen to effectively eliminate the noise floor and achieve a significant perceptual improvement over original room impulse responses. An original and extended balloon pop response of the Hagia Sophia, Istanbul, Turkey is shown in Fig. 9. The measured and extended spectrograms are displayed in Fig. 10 and Fig. 11, respectively. Within the original spectrogram, interfering conversation noise is visible around 6 and 8 seconds; within the extended spectrogram, no such artifacts are present. The smoothed energy profiles are shown in Fig. 12

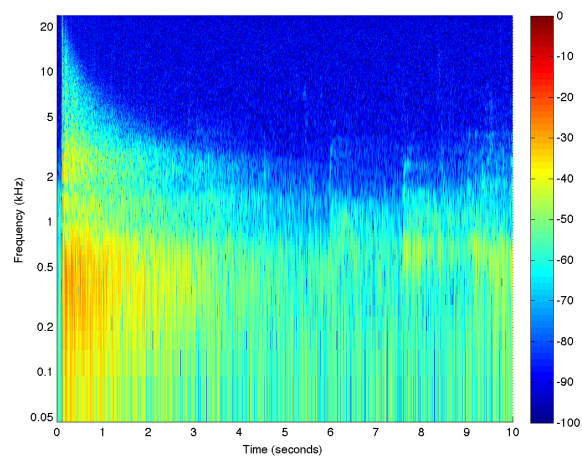
#### 6. CONCLUSION

To summarize, two methods for extending room impulse responses beyond their noise floor are presented. The first method crossfades a synthetically generated late-field with the input IR prior to the frequency-dependent noise floor arrival time. The second method windows or “bends” down the unwanted noise floor to extended the IR. Both methods maintain an identical impulse response prior to





**Fig. 9:** (a) Measured Hagia Sophia Balloon Pop Response. (b) Extended Hagia Sophia Balloon Pop Response.

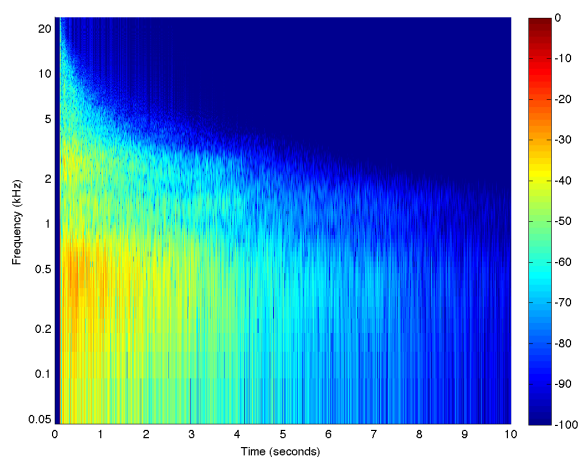


**Fig. 10:** Measured Hagia Sophia Balloon Pop Spectrogram. The spectrogram of the measured Hagia Sophia balloon pop response.

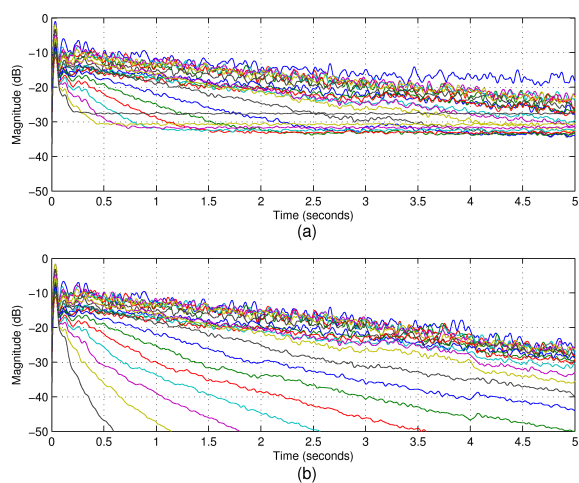
the noise floor arrival and impose a natural sounding decay after the noise floor.

## 7. ACKNOWLEDGEMENTS

This work was supported in part by grants from the Stanford University Presidential Fund and the Stan-



**Fig. 11:** Extended Hagia Sophia Balloon Pop Spectrogram. The spectrogram of the extended Hagia Sophia balloon pop response.



**Fig. 12:** (a) Measured Hagia Sophia Energy Profiles. The measured energy profiles are limited by the frequency-dependent noise floor. (b) Extended Hagia Sophia Energy Profiles. The extended energy profiles for each filter bank band are shown to extend well beyond the measured noise floor.

ford Institute for Creativity and the Arts (SiCa) for the Icons of Sound Project (see <https://ccrma.stanford.edu/groups/icons-of-sound>). Additional support was provided by Universal Audio in a

research collaboration with CCRMA, Stanford University (see <http://www.uaudio.com/>).

## 8. REFERENCES

- [1] M. R. Schroeder, "Integrated-Impulse method measuring sound decay without using impulses," *Journal of Acoustical Society of America*, vol. 66(2), pp. 497-500, 1979.
- [2] A. Farina, "Simultaneous measurement of impulse response and distortion with a swept-sine technique," *Convention Paper 5093*. Presented at the 108 Convention, Paris, February 19-22, 2000.
- [3] S. Müller, P. Massarani. "Transfer-function measurement techniques with sweeps," *Journal of the Audio Engineering Society*, vol. 49, no. 6, p. 443, 2001.
- [4] G. Stan, J. J. Embrechts, D. Archambeau. "Comparison of different impulse response measurement techniques," *Journal of the Audio Engineering Society*, vol. 50, no. 4, p. 249, 2002.
- [5] S. Foster, "Impulse response measurements using golay codes," In *Proceedings of IEEE ICASSP-86*, pp. 929-932, 1986.
- [6] J. Vanderkooy. "Aspects of MLS measuring systems," *Journal of the Audio Engineering Society*, vol. 42, pp. 219-231, 1994.
- [7] K. Lee, N. J. Bryan, J. S. Abel. "Approximating measured reverberation using a hybrid fixed/switched convolution structure," In *Proceedings of the 13th International Conference on Digital Audio Effects*, Graz, Austria, September 6-10, 2010.
- [8] J. S. Abel, N. J. Bryan, P. Huang, M. A. Kolar, B. Pentcheva. "On estimating room impulse responses from recorded balloon pops," Presented at the 129th Audio Engineering Society Convention, San Francisco, CA, November 4-7, 2010.
- [9] K. Spratt, J. S. Abel. "All natural room enhancement," In *Proceedings of the International Computer Music Conference*, Montreal, 2009.
- [10] J.-M. Jot, L. Cerveau, and O. Warusfel. "Analysis and synthesis of room reverberation based on a statistical time-frequency model," Presented at the 103rd Audio Engineering Society Convention, New York, 1997.
- [11] J. A. Moorer. "About this reverberation business," *Computer Music Journal*, vol. 3, no. 2, pp. 13-28, 1979.
- [12] M. R. Schroeder, "New method of measuring reverberation time," *Journal of Acoustical Society of America*, vol. 37, pp. 409-412, 1965.
- [13] M. R. Schroeder, "Natural sounding artificial reverberation," *Journal of the Audio Engineering Society*, vol. 10, no. 3, pp. 219-223, October 1962.
- [14] M. R. Schroeder, "Statistical parameters of the frequency response curves of large rooms," *Journal of the Audio Engineering Society*, vol. 35, no. 5, pp. 307-316, May 1987.
- [15] P. P. Vaidyanathan. *Multirate systems and filter banks*. Prentice Hall, Englewood Cliffs, 1993.