Robust Estimation of Reverberation Time Using Polynomial Roots

Ian J. Kelly[1], Francis M. Boland[1,2], and Jan Skoglund[1]

[1]*Google, Inc.*

[2]*Trinity College, Dublin, Ireland*

Correspondence should be addressed to Ian J. Kelly (`kellyi@google.com`)

ABSTRACT

This paper further investigates previous findings that coefficients of acoustic responses can be modelled as random polynomials with certain constraints applied. In the case of room impulse responses, the median value of their clustered roots has been shown to be directly related to the reverberation time of the room. In this paper we examine the frequency dependency of reverberation time and we also demonstrate the method's robustness to truncation of impulse responses.

1. INTRODUCTION

The complexity of the surfaces in real world acoustic environments and the overwhelming likelihood of diffuse reflections occurring over simple specular reflections, only lends further evidence to the understanding of why the samples of room impulse responses (RIRs) appear as if drawn from random distributions. Phrasing this from the point of view of simple image source modelling [1], the non-flat surfaces soon wash out all but the lowest order images. In his 1979 paper "About this Reverberation Business" Moorer [2] famously described the responses recorded in some of the most renowned concert halls in the world

"*While digitising the impulse responses from concert halls around the world, we kept noticing that the responses in the finest concert halls sounded remarkably similar to white noise*".

In [3] it was shown that despite exhibiting features such as exponential decay and onset delays, RIRs' root distributions essentially behave in the same way as those of random polynomials. Furthermore the relationship between the positioning of these annular clusters of roots in terms of magnitude and the reverberation time (RT_{60}) of room impulse responses was presented.

In this paper we investigate this method further by estimating reverberation time in different frequency bands and study its robustness to truncation of impulse responses.

2. RELATING REVERBERATION TIME TO ROOT CLUSTER RADIUS

Consider the following model of a room impulse response. Let $p[n]$ be a random signal vector of length N whose entries correspond to the coefficients of a random polynomial. We can multiply this signal with a decaying exponential window $w[n] = e^{-\beta n}$ also of length N. The room impulse response can thus be modeled as

$$h[n] = p[n] \otimes w[n] \tag{1}$$

where $\otimes$ is the Hadamard product for vectors.

The reverberation time RT_{60} is the 60 dB decay time for an RIR [4]. In the case of our model signal this can be easily derived from the envelope $w[n]$ and can be obtained by solving

$$20\log_{10}(e^{-\beta \mathrm{RT}_{60}}) = -60 \ \ \text{(dB)} \tag{2}$$

to get

$$\mathrm{RT}_{60} = \frac{1}{\beta}\ln(10^3). \tag{3}$$

We know from Hughes and Nikeghbali [5], that the roots of a polynomial whose coefficients are the samples of the signal $p[n]$ cluster uniformly about the unit circle. That is to say their magnitudes have an expected value of one. Also by the properties of the z-transform

$$H(z) = P(e^{\beta}z) = \prod_{n=1}^{N}(z + z_n) \tag{4}$$

and so the magnitudes of the roots of $P(z)$ are scaled by a factor of e^{β} to become the roots of $H(z)$ where $z_n,\ n \in [1,\ldots,N]$ are the roots of $H(z)$ in (4). Or equivalently

$$H(z) = P\left(e^{\frac{\ln(10^3)}{\mathrm{RT}_{60}}} z\right). \tag{5}$$

Thus if we estimate the constant β from the mean of the root magnitudes as

$$\beta = -\ln\left(\frac{1}{N}\sum_{n=1}^{N}|z_n|\right) \tag{6}$$

where $z_n,\ n \in [1,\ldots,N]$ are the roots of $h[n]$, the reverberation time can be written as

$$\mathrm{RT}_{60} = \frac{\ln\left(10^3\right)}{\ln\sum_{n=1}^{N}|z_n| - \ln(N)} \tag{7}$$

which depends solely upon the magnitudes of the roots of a given response.

2.1. Outlying Roots

Naturally this formulation of the reverberation time is perturbed by anything which perturbs the mean magnitude of the roots. As was seen in [3], acoustic impulse responses with any form of approximate delay will naturally have a ring of roots located well outside the unit circle. The presence and location of these roots are in no way related to the phenomenon of decay in the corresponding response. However their presence bears heavily upon the mean magnitude of the roots of an acoustic impulse response. Figure 1 shows a short room impulse response along with its root constellation to the right. The outlier roots, lying far from the unit circle (or far from the mean magnitude just inside the unit circle) skew the mean magnitude of the roots. The ring of roots outside the unit circle which are present due to the approximate onset delay also skew the mean, in this case the skew is positive.

In order to see how these outlying roots effect the estimated mean, a *trimmed mean* is calculated whereby the k% of roots furthest from the mean are discarded in order to calculate an updated mean. Naturally as k is increased the *trimmed mean* tends toward the median value of a response's roots magnitudes. Figure 2 shows the mean, median, and trimmed mean magnitudes of the roots of a set of RIRs across a variety of percentage trimmings (from 0% to 99.99%). The sharp dip at around 2-5% trim in each panel is likely due to the initial removal on the extreme outliers with magnitudes several times greater than the mean. As can be seen in most cases, after just removing a very small percentage of outliers, the trimmed mean quickly converges to the median value. This indicates that in most places the use of a median value to estimate β in (3) is likely near optimal. However looking at panels 2 and 6 (Office and Carolina), the convergence toward the median is very slow and from opposite directions in each case. Looking at the root constellations in each case it can be seen that the reason appears to be due to the roots behaving in different ways at different points around the unit circle. That is the reverberation times seem to be frequency dependent to a greater degree. This is explored in Section 3.2.

3. BENEFITS OF ESTIMATION FROM ROOTS

By estimating the decay rate from the root locations and not from the time domain or frequency domain responses $h[n]$ and $H(z)$ one does not base the reverberation time estimation on one exponential fit but instead on $N/2$ indirect exponential fits through the proxy of each root's location relative to the unit circle where N is the number of roots. The figure $N/2$ arises as for a real sequence such as a measured RIR roots appear as conjugate pairs. This thus amounts to an accurate estimation of the true reverberation time/decay rate via the law of large numbers.

3.1. Robustness to Truncation

A second, more important benefit of the method outlined is that provided the responses under analysis are of high enough order to ensure that if there were no decay present, the roots would form tight clusters, then the method should still perform well even with heavily truncated responses. Results presented in [3] support this showing that under ideal circumstances reverberation time can be estimated with as low as 2% error, even when responses were truncated to just 8000 samples, which for the largest room corresponded to 22.7% of the RT_{60}. To further demonstrate this benefit a long random sequence with a normally distributed set of coefficients was created so that it was 18000 samples in length (0.408 s at a sampling rate of 44.1 kHz). This response was subsequently scaled via a Hadamard multiplication with a decaying exponential envelope which had an RT_{60} of 0.783 s (or 34539 samples at the same sampling rate). This essentially meant that the response was truncated to 51% the actual RT_{60}. This response was then further truncated to just 6000 samples, 17% of the

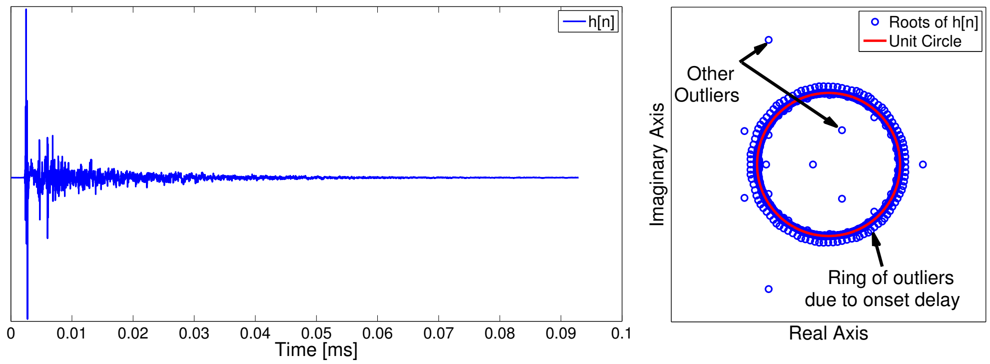

Fig. 1: A short RIR along with its root constellation (shown to the right). Outlying roots are clearly present.

RT_{60}. Figure 3 shows these responses along with the decay curve applied. In each case the RT_{60} is shown as estimated by the proposed algorithm, hereby referred to as the root method. Table 1 shows the estimated RT_{60} values along with those calculated via the standard RT_{60} estimation method ISO 3382 [7]. As can be seen the method continues to perform well even after aggressive truncation of the input sequence. Furthermore the RT_{60} result is more accurate overall.

Actual RT_{60}: 0.783 s				
Method	18000 point		6000 point	
	RT_{60} [s]	Error %	RT_{60} [s]	error %
Root	*0.769*	*1.8*	*0.741*	*5.3*
ISO 3382	*0.702*	*10.3*	*0.261*	*66.6*

Table 1: RT_{60} values calculated via the root method and the standard ISO 3382 method [7] from the truncated responses shown in Figure 3.

3.2. Frequency Bands

Looking at the RIRs in a roots only manner allows an estimation of the reverberation time in any set of frequency bands of any constant or varying width, with great ease. All that must be done is to modify (7) accordingly. Only the roots with argument between ω_1 and ω_2 radians corresponding to $f_1 = F_s \frac{\omega_1}{2\pi}$ to $f_2 = F_s \frac{\omega_2}{2\pi}$ Hz, where F_s Hz is the sampling frequency, should be included. This can be formulated as

$$RT_{60}^{\omega_1,\omega_2} = \frac{\ln\left(10^3\right)}{\sum_{\arg(z_n)\in[\omega_1,\omega_2]} \ln|z_n| - \ln\left(\#\{z_n : \omega_1 \leq \arg z_n \leq \omega_2\}\right)} \quad (8)$$

Methods such as ISO 3382 traditionally calculate RT_{60} across different frequency bands by pre-filtering with, for example, octave band filters. Generally such implementations use IIR filters [8]. However, such filters can never be perfectly flat in the magnitude spectrum nor can they ever be such that they do not at least slightly perturb the decay time of the frequencies being examined. The root method performs no filtering in order to make reverberation time estimates across different frequency bands. Figure 4 shows the reverberation times calculated across nine octave frequency bands, centred between 31.25Hz and 8kHz.

3.2.1. Verification

In order to verify whether evaluating the root magnitudes of an RIR over angular intervals relating to frequency is a method capable of accurately estimating reverberation time over frequency bands, it is necessary to construct a random polynomial where different frequencies decay at different rates.

Let $\mathbf{s}_f$ be a sine wave with a frequency of f Hz and let $\alpha \sim \mathcal{N}(0,1)$ be a random variable with a Gaussian distribution, zero mean, and a standard deviation of one. One can thus define a sequence

$$\mathbf{r} = \sum_{f=0}^{\frac{Fs}{2}} \alpha \mathbf{s}_f \quad (9)$$

that is the sum of the randomly scaled sinusoids, where α is randomly chosen for each frequency. Given a great number of such summed terms, $\mathbf{r}$ will in essence be a random vector with a flat spectrum and as such will have roots distributed like those of random polynomials. In

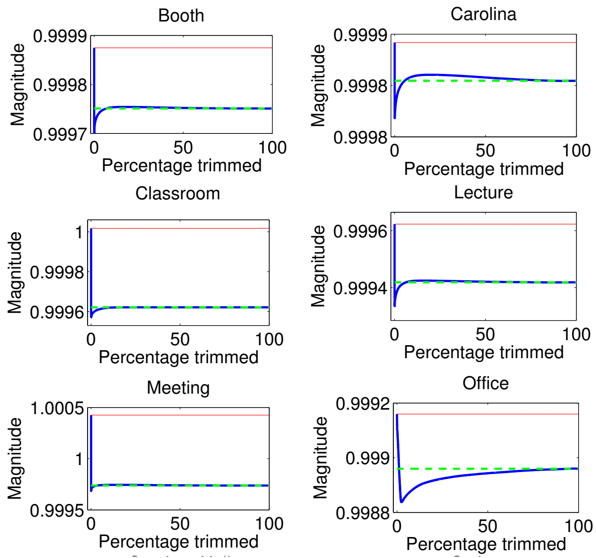

Fig. 2: Shown here are the trimmed means of the magnitudes of the roots of a set of RIRs (blue), The mean magnitudes (red) and the median magnitudes (green *broken*) to which the trimmed means converge as more and more outliers are removed. Panels 1, 2, 4, 5, 6, and 8 are from RIRs recorded by [6] and panels 3 and 7 are from RIRs recorded by the authors.

order to test the validity of Section 3.2 such a vector was created 6000 samples in length comprising of the sum of 10000 randomly scaled sinusoids with frequency between 1 Hz and 22 kHz. A second sequence denoted $\mathbf{r}_{\text{scale}}$ was then generated

$$\mathbf{r}_{\text{scale}} = \sum_{f=0}^{\frac{\text{Fs}}{2}} \alpha \left(\mathbf{s}_f \otimes e^{-\beta \mathbf{t}} \right) \tag{10}$$

where $\otimes$ denotes a Hadamard product and β is chosen in order to give the decay envelope $e^{-\beta \mathbf{t}}$ an RT_{60} of 0.07 s (3087 samples) for $f \in [0,\ldots,2000]$ Hz, increasing to an RT_{60} of 0.14 s (6174 samples) for $f \in \left[\frac{\text{Fs}}{2} - 2000,\ldots,\frac{\text{Fs}}{2}\right]$ Hz, in a linear fashion such that there are eleven sets of decay envelopes $e^{-\beta \mathbf{t}}$ in total each, applied to a 2 kHz set of $\mathbf{s}_f$ vectors ($\approx$ 830 sinusoids in each band in this case).

Consequently the vector $\mathbf{r}_{\text{scale}}$ has frequency components decaying at different rates. Figure 5 shows the sequence $\mathbf{r}_{\text{scale}}$ along with its coefficient distribution and root constellation.

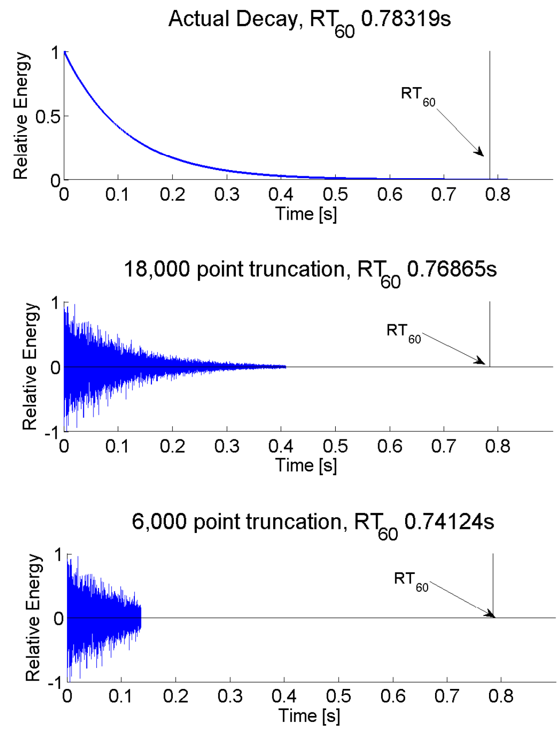

Fig. 3: Shown here are the original exponentially decaying envelope (top panel) titled with its RT_{60}, This value is also shown with the vertical line in the plot. The middle and lower panels show the truncated random sequences with the exponential decay applied. Each of these panels are titled with the RT_{60} calculated using the root method.

In order to verify the claims of Section 3.2 it would be necessary to show that an application of (8) to the roots of $\mathbf{r}_{\text{scale}}$ would return a set of RT_{60} estimates equal to the exact RT_{60}s of the exponential decay envelopes applied to the sinusoids $\mathbf{s}_f$ before summation.

The result is shown in Figure 6. This clearly shows that the method of RT_{60} estimation from root locations works accurately when applied across frequency bands with different decay profiles. This has been achieved without the need for any filtering demonstrating the power of this novel technique. The results in Figure 6 verify the validity of those frequency dependent RT_{60} estimations shown in Figure 4.

4. CONCLUSION

This paper has demonstrated that the method of reverberation time estimation based on root locations of

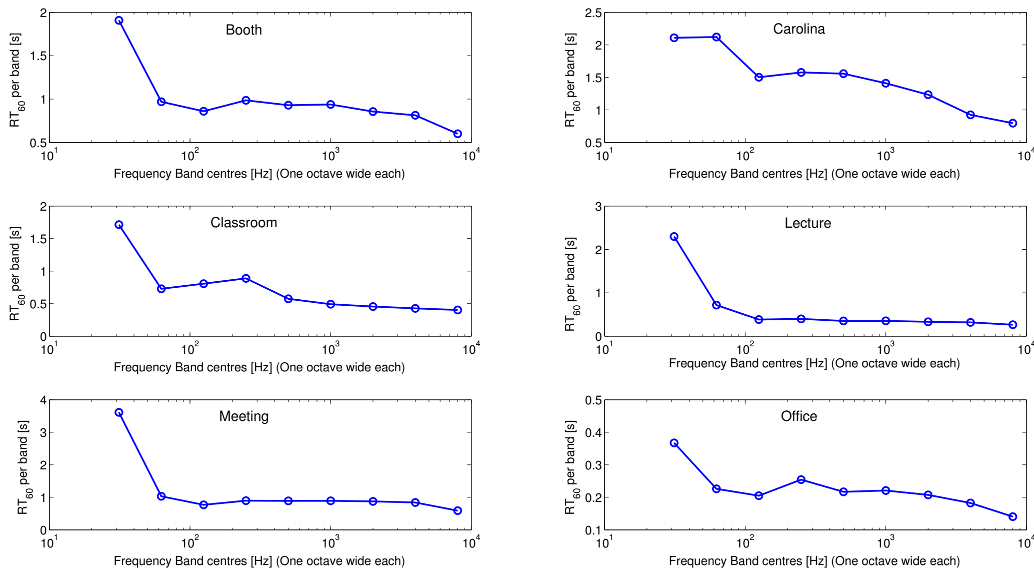

Fig. 4: Shown here are the reverberation times calculated via the root method for each of the RIRs in nine octave frequency bands, centred between 31.25Hz to 8kHz.

a measured impulse response is accurate and robust to truncation of the measured response. It was also shown that such a method is easily adapted to reverberation time measurement over any set of frequency bands without the necessity of any filtering of the response in question.

5. REFERENCES

[1] Jont B. Allen and David A. Berkley, "Image method for efficiently simulating small-room acoustics," *J. Acoust. Soc. Am.*, vol. 65, no. 4, pp. 943–950, 1979.

[2] James A Moorer, "About this reverberation business," *Computer music journal*, pp. 13–28, 1979.

[3] Ian J Kelly and Francis M Boland, "Randomness and the reverberation time, RT60, of acoustic responses," in *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*. IEEE, 2014, pp. 8222–8226.

[4] Wallace C Sabine, "Architectural acoustics," in *Proceedings of the American Academy of Arts and Sciences*. JSTOR, 1906, vol. 42, pp. 51–84.

[5] C. P. Hughes and A. Nikeghbali, "The zeros of random polynomials cluster uniformly near the unit circle.," *Compositio Mathematica*, vol. 144, pp. 734?746, 2008.

[6] Marco Jeub, Magnus Schäfer, and Peter Vary, "A binaural room impulse response database for the evaluation of dereverberation algorithms," in *Proceedings of the 16th international conference on Digital Signal Processing*, Piscataway, NJ, USA, 2009, DSP'09, pp. 550–554, IEEE Press.

[7] ISO 3382:1997, "Acoustics ? measurement of the reverberation time of rooms with reference to other acoustical parameters," International Organisation for Standardisation, Geneva, Switzerland.

[8] Anon., "Specifications for octave-band and fractional-octave-band analog and digital filters," Tech. Rep. ANSI S1.1-1986 (ASA 65-1986), American National Standards Institute, 1993.

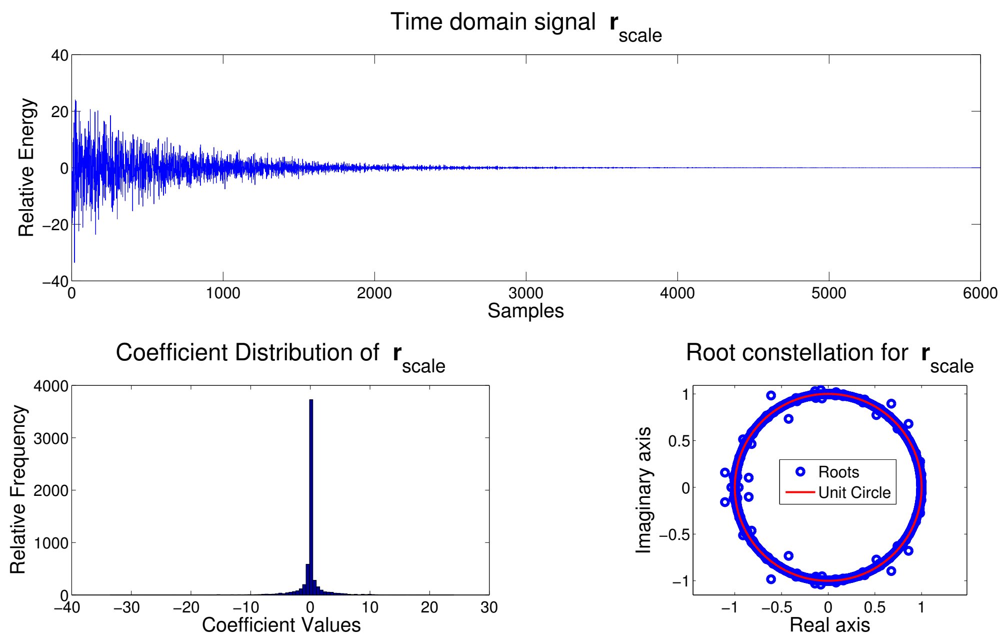

Fig. 5: Shown here are the samples, sample distribution and roots of the sequence $\mathbf{r}_{\text{scale}}$ formed as described in 10.

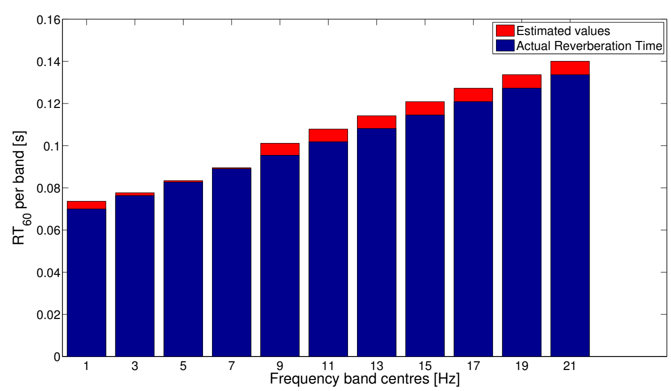

Fig. 6: Shown here are the actual RT_{60} values for the decaying exponential envelopes applied to each frequency band (blue) along with the results returned from the proposed algorithm (red). As can clearly be seen the estimation is quite accurate in all frequency bands.