Bayesian Inference Approach to Room-Acoustic Modal Analysis

Wesley Henderson*, Paul Goggans†, Ning Xiang* and Jonathan Botts*

*Graduate Program in Architectural Acoustics, Rensselaer Polytechnic Institute
†Department of Electrical Engineering, University of Mississippi, goggans@olemiss.edu

Abstract. Spectrum estimation is a problem common to many fields of physics, science, and engineering, and it has thus received a great deal of attention from the Bayesian data analysis community. In room acoustics, the modal or frequency response of a room is important for diagnosing and remedying acoustical defects. The physics of a sound field in a room dictates a model comprised of exponentially decaying sinusoids. Continuing in the tradition of the seminal work of Bretthorst and Jaynes, this work contributes an approach to analyzing the modal responses of rooms with a time-domain model. Room acoustic spectra are constructed of damped sinusoids, and the model-based approach allows estimation of the number of sinusoids in the signal as well as their frequencies, amplitudes, damping constants, and phase delays. The frequency-amplitude spectrum may be most useful for characterizing a room, but in some settings the damping constants are of primary interest. This is the case for measuring the absorptive properties of materials, for example. A further challenge of the room acoustic spectrum problem is that modal density increases quadratically with frequency. At a point called the Schroeder frequency, adjacent modes overlap enough that the spectrum – particularly when estimated with the discrete Fourier transform – can be treated as a continuum. The time-domain, model-based approach can resolve overlapping modes and in some cases be used to estimate the Schroeder frequency. The proposed approach addresses the issue of filtering and preprocessing in order for the sampling to accurately identify all present room modes with their quadratically increasing density.

Keywords: Bayesian inference, acoustics, spectrum estimation, room modes, nested sampling
PACS: 43.20.Ks, 43.55.Br, 43.60.Uv

BAYESIAN MODEL-BASED FRAMEWORK

Spectrum estimation is a problem that is encountered in many disparate scientific disciplines. Bayesian inference has been applied to the problem in several of these disciplines, including chemistry [1, 2], astronomy [3], and signal processing [4]. Bayesian inference has also been applied to acoustics problems outside the scope of spectrum estimation, such as coupled volume energy decay analysis [5, 6]. Frequency spectrum estimation is particularly important in the field of room acoustics. A room behaves as an acoustically resonant volume in which the resonant frequencies and the bandwidth of the resonance peaks are determined by the geometry of the room and the composition of the room's surfaces. While the traditional frequency-magnitude spectrum is useful to room acoustics, this work addresses the spectrum estimation problem using a time-domain model.

A room impulse response can be represented as a sum of exponentially decaying sinusoids:

$$\gamma(t) = \sum_{i=1}^{M} A_i e^{-\frac{6.9t}{T_i}} \cos(2\pi f_i t + \phi_i), \tag{1}$$

with impulse response, γ, number of modes, M, time, t, and parameters amplitude, A, decay time, T, frequency, f, and phase delay, ϕ.

Model Selection

The first major strength of this model-based spectrum estimation method is its facility in estimating the number of modes present in a room impulse response. This estimation is accomplished using Bayesian model selection. For measured data $\mathbf{D}$ and background information I, the probability of a given model $\mathbf{M}$ is given by Bayes' theorem:

$$p(\mathbf{M}|\mathbf{D},I) = \frac{p(\mathbf{M}|I)p(\mathbf{D}|\mathbf{M},I)}{p(\mathbf{D}|I)}, \tag{2}$$

with prior $p(\mathbf{M}|I)$, likelihood $p(\mathbf{D}|\mathbf{M},I)$, and evidence $p(\mathbf{D}|I)$. The ratio of the posterior probabilities of two different models is known as the Bayes factor and may be used to rank those models against each other. By assigning an objective, uniform prior distribution to the family of models under consideration, the ratio of the posteriors reduces to a ratio of the model likelihoods:

$$\frac{p(\mathbf{D}|\mathbf{M}_i,I)}{p(\mathbf{D}|\mathbf{M}_j,I)}, \tag{3}$$

for models i and j.

Parameter Estimation

Another strength of this spectrum estimation method is its ability to estimate the parameters of interest for the modes present in a room impulse response, specifically frequency and decay time. For measured data $\mathbf{D}$, model $\mathbf{M}$, and background information I, the probability of a set of parameters $\mathbf{B}$ is similarly given by Bayes' Theorem:

$$p(\mathbf{B}|\mathbf{D},\mathbf{M},I) = \frac{p(\mathbf{B}|\mathbf{M},I)p(\mathbf{D}|\mathbf{B},\mathbf{M},I)}{p(\mathbf{D}|\mathbf{M},I)}, \tag{4}$$

with prior $p(\mathbf{B}|\mathbf{M},I)$, likelihood $p(\mathbf{D}|\mathbf{B},\mathbf{M},I)$, and evidence $p(\mathbf{D}|\mathbf{M},I)$. Note that the evidence term in the parameter estimation formulation of Bayes' theorem is exactly the same as the likelihood term in the model selection formulation. This equality allows the evidence calculated in the parameter estimation process to be used to rank competing models.

The background information I encodes the assertion that the model $\mathbf{M}$ fits the measured data $\mathbf{D}$ reasonably well, with a finite residual error

$$\mathbf{e} = \mathbf{D} - \mathbf{M}. \tag{5}$$

The finite nature of the residual error implies that the distribution of the error has a finite but unknown variance. Given this knowledge, application of the principle of maximum entropy yields a Gaussian likelihood distribution over the residual error. In this application, the variance term of the Gaussian distribution is a nuisance parameter. In order to marginalize the variance, a Jeffreys prior [7] is assigned to the variance, and this prior times the likelihood function is integrated over the variance. This marginalization yields a Student's T distribution over E:

$$\mathscr{L}(\mathbf{B}) \equiv p(\mathbf{D}|\mathbf{B},\mathbf{M},I) = \Gamma\left(\frac{K}{2}\right)\frac{(2\pi E)^{-K/2}}{2}, \tag{6}$$

where $\Gamma(\cdot)$ is the gamma function, $E = \mathbf{e}^2/2$, K is the number of samples in the measured data, and $\mathscr{L}(\mathbf{B})$ represents the likelihood given a set of parameters $\mathbf{B}$.

Nested Sampling

Calculation of the evidence values necessary to perform the model selection is computationally expensive using traditional Bayesian parameter estimation methods. Skilling's nested sampling algorithm [8] provides a more efficient way to calculate the evidence for a given model while incidentally generating the posterior samples necessary for parameter estimation. Nested sampling has also been applied to the coupled volume problem in acoustics [9].

This specific implementation of nested sampling begins by drawing 100 unconstrained, pseudo-random samples from the prior distribution. The likelihood of each sample is calculated, and the sample with the lowest likelihood is removed from the initial population and saved into a collection of posterior samples. A parameter is chosen at random, and a random walk MCMC process is used to explore the prior distribution in the dimension of the chosen parameter. A new sample is accepted if its likelihood is greater than or equal to the likelihood constraint. Forty random walk steps are taken, with the step size tuned at each step to move toward a target acceptance ratio of 50%. The resulting sample is inserted into the initial population to replace the discarded sample. This exploration process is adapted from Skilling's toy lighthouse problem implementation in [10, 192–195].

This process iterates until the likelihood of the new sample is less than 1×10^{-10} greater than the likelihood of the previous sample, or until 40000 complete iterations have occurred.

Approach

This method ultimately strives to obtain from a single room impulse response the number of modes present and four parameters of interest for each mode. This method begins by measuring a room's impulse response using standard techniques [11]. To restrict the analysis to frequencies of interest, the impulse response is filtered to isolate these frequencies. Prior limits are set based on impulse response characteristics. For instance, frequency is limited by the impulse response's Nyquist frequency, and the decay time is bounded by values obtained through more broadband decay time measurement techniques.

The nested sampling is run with the measured data, testing each candidate number of modes until the accumulated evidence stops increasing. The number of modes with the highest evidence is chosen as the number of modes present in the impulse response, and the parameters associated with that number of modes are calculated by finding the weighted mean of the collected samples over each parameter.

SIMULATED IMPULSE RESPONSES

Methods

The proposed method was self-verified using a simulated impulse response. This impulse response was generated using the time domain model (Equation 1) with uniformly-distributed random noise added to each time sample. This noise had a mean of zero and a maximum amplitude of 1% of the maximum amplitude of the signal.

One particular scenario was investigated: an impulse response with three modes, two closely-spaced modes and one mode well-separated from the other two.

Results

For the case under investigation, the known parameter values, the inferred parameter values, and the error between these sets is shown in table 1. Figure 1 shows the simulated and inferred impulse responses, the magnitude spectrum of the simulated impulse response and the inferred frequencies, and the log evidence per mode.

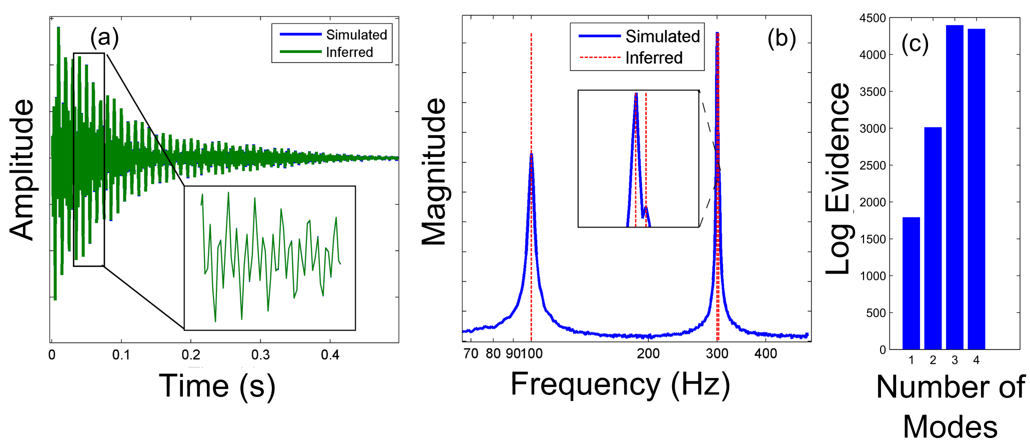

FIGURE 1. Simulated room impulse response with one well-separated mode and two closely-spaced modes. (a) simulated and inferred time-domain signals; (b) simulated frequency-domain signal with the inferred frequencies represented by dashed lines; (c) log evidence for each tested mode

TABLE 1. Simulated room impulse response with one well-separated mode and two closely-spaced modes

	Known values	**Inferred values**	**Error (%)**
First mode			
Amplitude	0.25	0.248	0.8
Decay time (s)	0.7	0.704	0.6
Frequency (Hz)	100	100	0.0
Phase delay (rad)	0	6.28	0.5
Second mode			
Amplitude	0.3	0.298	0.7
Decay time (s)	0.9	0.905	0.6
Frequency (Hz)	300	300	0.0
Phase delay (rad)	0	1.19×10^{-2}	0.2
Third mode			
Amplitude	0.1	0.101	1.0
Decay Time (s)	0.9	0.890	1.1
Frequency (Hz)	303	303	0.0
Phase delay (rad)	0	6.24	0.7

MEASURED IMPULSE RESPONSES

Methods

The proposed method was also used to analyze measured impulse responses. The values inferred in these analyses were compared to values obtained using more traditional methods.

The first impulse response analyzed was measured in a rectangular chamber with wooden surfaces. This chamber was 67 cm wide, 51 cm tall, and 192 cm deep. The source for the impulse response measurement was placed in one corner of the chamber, and the receiver was placed in a non-adjacent corner. The resulting impulse response was resampled to a sampling frequency of 1080 Hz and then filtered with a 10th-order Chebyshev Type II low-pass filter with a stopband edge frequency of 300 Hz and stopband ripple 40 dB down from the peak passband level to isolate lower frequencies and help reduce computation time.

Because the chamber was rectangular, its modal frequencies could be estimated using a solution of the wave equation for sound propagating in a rectangular chamber with rigid surfaces,

$$f_{\mathrm{l,m,n}} = \frac{c}{2}\sqrt{\left(\frac{l}{L_x}\right)^2 + \left(\frac{m}{L_y}\right)^2 + \left(\frac{n}{L_z}\right)^2}, \quad (7)$$

with modal frequency f, dimensions L_x, L_y, and L_z, and modal indexes l, m, and n.

Once the Bayesian analysis was performed on the impulse response obtained from the rectangular chamber, decay time analysis was performed at the inferred modal frequencies. Using the same setup and similar atmospheric conditions, the chamber was excited using sine tones at these inferred frequencies. The sine tones were interrupted and the slope of the resulting decay curves was used to determine the decay times at the tested frequencies.

Results

The comparison between the switch-off decay times and the inferred decay times at various frequencies for the rectangular chamber is shown in table 2. The comparison between the frequencies predicted using the wave equation and the inferred frequencies for the rectangular chamber is shown in table 3.

TABLE 2. Rectangular chamber room impulse response switch-off vs. inferred decay times

Mode	Switch-off DT (s)	Inferred DT (s)	Error (%)
1	0.9108	0.798	12.4
3	1.4649	1.34	8.5
4	0.6667	0.571	14.4

TABLE 3. Rectangular chamber room impulse response predicted vs. inferred frequencies

Mode	Predicted Frequency (Hz)	Inferred Frequency (Hz)	Error (%)
1	261.8	263.7	0.7
2	178.6	186.6	4.5
3	268.0	274.7	2.5
4	276.6	279.7	1.1

Figure 2 shows the results of the analysis of the rectangular chamber impulse response. Part (a) shows the inferred impulse response overlaid on the measured impulse response. Part (b) shows the magnitude spectrum of the impulse response as a solid line, the predicted modal frequencies as dotted lines, and the inferred modal frequencies as a dashed line. Part (c) shows the mean of the log evidence per mode over 16 runs, with the error bars representing standard deviation.

DISCUSSION AND CONCLUSIONS

The analysis of the simulated impulse response shows that this analysis technique is internally valid, as the observed errors in the inferred parameter values never exceed 1.1%. The method has no problem detecting and accurately estimating the parameters of closely-spaced modes, an advantage over peak detection and visual inspection of the magnitude spectrum.

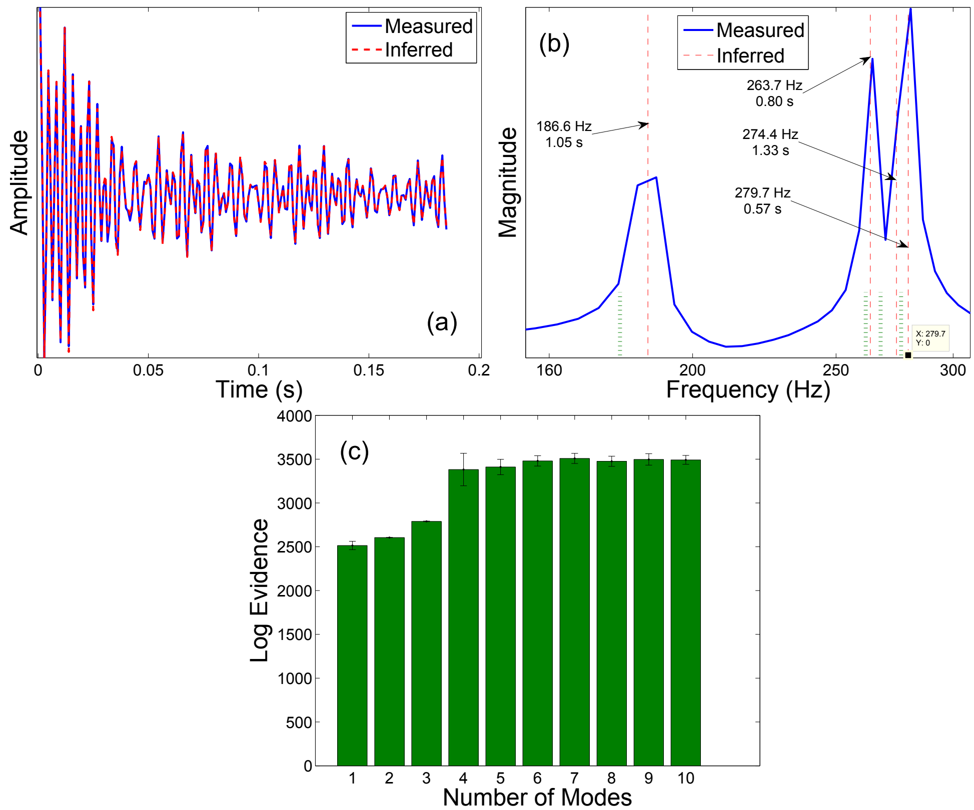

FIGURE 2. Measured room impulse response. (a) measured and inferred impulse response in the time domain; (b) solid – magnitude spectrum, dashed – inferred frequency, dotted – wave equation frequencies; (c) average of log evidence per mode over 16 runs with standard deviation error bars

Decay times and frequencies in the measured impulse response are estimated within reasonable error values. Also, the time-domain fit in Figure 2 part (a) and the peak in log evidence at the fourth mode in Figure 2 part (c) indicate that the nested sampling converges.

Future work will include measuring impulse responses in non-rectangular spaces in which the modal frequencies cannot be easily predicted. Other spaces with interesting geometrical configurations will also be considered.

ACKNOWLEDGMENTS

Thanks to Cameron Fackler for his help with measurements and for his insightful discussions about Bayesian philosophy.

REFERENCES

1. G. Bretthorst, *J. Magnetic Resonance (1969)* **88**, 533–551 (1990).
2. G. Bretthorst, *J. Magnetic Resonance (1969)* **88**, 552–570 (1990).

3. I. O'Dwyer, H. Eriksen, B. Wandelt, J. Jewell, D. Larson, K. Górski, A. Banday, S. Levin, and P. Lilje, *The Astrophysical Journal Letters* **617**, L99 (2004).
4. C. Andrieu, and A. Doucet, *IEEE Trans. Sig. Proc.* **47**, 2667–2676 (1999).
5. N. Xiang, and P. Goggans, *J. Acoust. Soc. Am.* **110**, 1415 (2001).
6. N. Xiang, and P. Goggans, *J. Acoust. Soc. Am.* **113**, 2685 (2003).
7. H. Jeffreys, *Proc. R. Soc. London, Ser. A. Math. Phys. Sci.* **186**, 453–461 (1946).
8. J. Skilling, *Bayesian Analysis* **1**, 833–860 (2006).
9. T. Jasa, and N. Xiang, *Bayesian Inference Max. Ent. Meth. Sci. Eng.* pp. 189–196 (2005).
10. D. S. Sivia, and J. Skilling, *Data analysis: a Bayesian tutorial*, Oxford University Press, USA, 2006.
11. *ISO-3382, Acoustics – Measurement of the reverberation time of rooms with reference to other acoustical parameters*, ISO, Geneva, Switzerland, 1997.