inter.noise 2012
august 19-22
new york city, usa
Quieting the World's Cities

Bayesian room-acoustic modal analysis

Wesley Henderson [a)]
Jonathan Botts [b)]
Ning Xiang [c)]
Graduate Program in Architectural Acoustics, School of Architecture,
Rensselaer Polytechnic Institute, Troy, New York, USA 12180

An important element of architectural acoustic design is the prediction and mitigation of modal effects. Modal effects can create undesirable aural phenomena, especially in smaller rooms. The use of Bayesian methods for spectrum analysis was pioneered originally in the field of nuclear magnetic resonance spectroscopy. A time-domain model may be used with methods of Bayesian inference to analogously analyze room modes, accounting for frequency, amplitude, damping, and phase. Bayesian model selection can be used to determine the number of modes present in the signal, while simultaneously estimating the associated modal parameters. Accuracy of the method is demonstrated using simulated data, addressing scenarios like well-spaced modes, overlapping modes, and unevenly damped degenerate modes. Then, the method's usefulness is demonstrated by running it against real, measured impulse responses. Finally, the Bayesian model-based approach is compared to more standard frequency-domain analyses.

1 INTRODUCTION

This paper describes a method to determine a room's acoustic modal characteristics using a time-domain signal model and two levels of Bayesian inference (parameter estimation and model selection). The method extracts from a single impulse response modal parameters amplitude, decay time, frequency, and phase delay, as well as the number of modes present in the impulse response.

1.1 Motivation

In architectural acoustics, room modes can be useful or detrimental, depending on context. Overlapping room modes are responsible for the reverberant response that acousticians generally find useful and desirable in acoustic spaces; however, if a room is small enough, and if the room's surfaces are reflective enough, room modes at frequencies near the lower end of human sensitivity can become sparse and very noticeable. Recording studio control rooms tend to experience this

[a)] email: wesley.henderson11@gmail.com
[b)] email: bottsj@rpi.edu
[c)] email: xiangn@rpi.edu

problem due to their small size, and mixing engineers cite low-frequency resonances as a common source of disturbance.[1] For an existing control room, the proposed method could be used to determine the offending modes' frequencies and decay times, allowing acousticians to recommend specific diffusive or absorptive surface treatments.

For rectangular rooms, there is an analytical solution to the wave equation allowing the prediction of room modes. However, for non-rectangular rooms (e.g. rooms with non-parallel flat walls or large curved surfaces), a new analytical solution must be written for each unique geometry. The proposed method could be used in rooms with non-rectangular shapes to characterize the room's modal response.

Finally, the most generally useful application of this method is the ability to extract both frequency and damping information from a single impulse response. In acoustics, this allows an acoustician to more completely characterize a room's acoustical characteristics based on already established measurement techniques. This method can be applied to other fields in science and engineering that also require spectrum estimation.

1.2 Literature Review

Research into the application of Bayesian analysis to the problem of characterizing well-separated, exponentially decaying sinusoids began with Larry Bretthorst's work in nuclear magnetic resonance spectroscopy[2,3]. Bretthorst's work showed that the parameters of noisy signals can be accurately estimated, and that the number of signals present in noisy data can be accurately estimated using Bayesian inference.

Andrieu and Doucet[4] used a time-domain, model-based approach to characterize noisy, undamped sinusoids using a reversible-jump Markov-chain Monte Carlo (RJMCMC) method to numerically integrate the posterior. This method proved to work well with very noisy data and closely spaced frequencies, where it performed better than classical methods.

The method described in this paper builds on these previous works by performing parameter estimation and model selection on the set of damped sinusoids that make up room impulse responses, using nested sampling to integrate the posterior and a Metropolis-Hastings update.

2 THEORY

2.1 Time-domain Model

For the purposes of this work, a room impulse response is modeled as a sum of exponentially decaying sinusoids, represented mathematically:

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

This is a parametric model with parameters amplitude (A), reverberation time (T), frequency (f), and phase delay (ϕ); signal γ, and time t. This represents a family of models with model M containing M modes.

2.2 Bayesian Inference

This project is concerned with two levels of Bayesian inference: parameter estimation and model selection. Both levels are based on Bayes' theorem. For parameter estimation, Bayes' theorem is written:

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

where $\mathbf{\Theta}$ is the collection of parameters $\{A_1, T_1, f_1, \phi_1, ...A_n, T_n, f_n, \phi_n\}$, $\mathbf{D}$ is the measured data, $\mathbf{M}$ is the model (number of modes), and I represents background knowledge. The prior distribution, $p(\mathbf{\Theta}|\mathbf{M}, I)$, encodes what is known about the parameters before testing. The likelihood function, $p(\mathbf{D}|\mathbf{\Theta}, \mathbf{M}, I)$, encodes updated knowledge about the parameters given measured data. The marginal likelihood (or evidence), $p(\mathbf{D}|\mathbf{M}, I)$ serves to normalize the posterior distribution, and also to rank competing models. Finally, the posterior distribution, $p(\mathbf{\Theta}|\mathbf{D}, \mathbf{M}, I)$, represents the updated knowledge of the parameters given the prior distribution, the likelihood function, and the evidence.

Once the posterior distribution has been estimated, any moments of interest can be calculated. For instance, the mean of the distribution in each dimension yields the estimated value of the parameter in that dimension. Similarly, the variance (and more usefully, the standard deviation) can be calculated for each dimension to quantify the uncertainty of the parameter estimates.

Model selection is performed by estimating the evidence for each of a set of competing, candidate modes. Qualitatively, the preferred model is the simplest one that fully describes the measured data. Quantitatively, the preferred model is the one with the highest evidence.

2.3 Nested Sampling

This project uses nested sampling[5] to implement Bayesian parameter estimation and model selection. Nested sampling is a method to accumulate the evidence of a model and, as a by-product, to generate posterior samples.

The process begins by drawing a population of random samples from the prior distribution. The likelihood ($\mathcal{L}$) of these samples is evaluated. The lowest likelihood value in this population is set as the new likelihood threshold, $\mathcal{L}^*$. This lowest-likelihood sample is saved and discarded from the population. It is replaced by another sample (with a likelihood $\mathcal{L} > \mathcal{L}^*$), generated using Metropolis-Hastings sampling of the posterior distribution.

At every iteration of this process, the width of the portion of the prior mass that contains samples with likelihood greater than $\mathcal{L}^*$ is estimated. The prior mass is a one-dimensional representation of the prior distribution, and the portion at iteration i is roughly estimated as

$$X_i = e^{-\frac{i}{N}}, \tag{3}$$

for N objects. The width of the portion of the prior mass associated with a sample is defined as

$$w_i = X_{i-1} - X_i. \tag{4}$$

The accumulated evidence is also calculated at each iteration. The total evidence at the end of the process is

$$Z = \sum_{i=1}^{K} w_i \mathcal{L}_i, \tag{5}$$

for K samples and evidence Z. The nested sampling process for a given model is deemed to be complete when the change in likelihood from one step to the next drops below a threshold value. Parameters are determined by finding the mean of the discarded samples weighted by the posterior.

2.4 Approach

The approach taken by this work is as follows. A room's impulse response is measured using standard techniques. To avoid excessive computation time, the impulse response is filtered to select only certain frequencies. Limits on the prior distribution are determined from the impulse response's characteristics: amplitude is limited by the maximum absolute value of the impulse response, decay time is limited by the length of the impulse response, frequency is limited by the measurement's Nyquist frequency, and phase delay is limited to 2π due to its circularity.

With the measured data and prior limits in place, the nested sampling algorithm is run with the assumption that only one mode is present in the signal. It is run again for two modes, and this process continues until the accumulated evidence stops increasing. The number of modes with the highest accumulated evidence is chosen as the correct number of modes, and the parameters associated with that posterior distribution are calculated.

3 SIMULATED IMPULSE RESPONSES

3.1 Methods

The first tests performed with the proposed method were done using simulated impulse responses generated using the model in Equation (1). The impulse responses were generated using parameters chosen to simulate certain situations, such as well-separated modes and degenerate modes. Once the impulse responses were generated, a small amount of Gaussian noise was added to simulate a realistic measurement.

3.2 Results

The simulated impulse response test shown is of a case with two well-separated modes. Table 1 shows the results of this test, including the known parameters, the inferred parameters, and the observed error as a percentage of the known parameters. As a note, phase delay is circular, i.e. a phase delay of 2π is equivalent to a phase delay of 0. This circularity is considered in calculating the error in the phase delay. Figure 1 shows the time-domain and frequency-domain results of the test, as well as a graphical representation of the model-selection process.

4 MEASURED IMPULSE RESPONSES

4.1 Methods

Once the simulated impulse response tests proved successful and the internal validity of the method was verified, tests using real, measured impulse responses began. The RIR shown here was measured in a scale-model transmission loss measurement chamber about 67 cm wide, 51 cm tall, and 87 cm deep. The walls of the chamber are made of layered cement board, foam, and wood. The source was placed in one corner of the chamber, and the receiver was placed in another corner.

In order to isolate just a few of the lower modes in the chamber, the impulse response was resampled to 862 Hz, then filtered with a 10th-order Chebyshev Type II low-pass filter with a stopband edge frequency of 359 Hz and stopband ripple 40 dB down from the peak passband level. The resampling removes superfluous data points and serves to significantly speed up computation time. This particular filter has a steep roll-off, ideal for isolating a few low-frequency modes.

4.2 Results

The results of the measured impulse response test are shown in table 2. The time- and frequency-domain results, as well as the model selection process are shown in figure 2. In the frequency-domain plot, the lowest inferred mode is not shown.

5 DISCUSSION

The results shown in this article demonstrate that the proposed method can accurately analyze the modal characteristics of a room, even in the presence of closely-spaced modes. The tests involving simulated impulse responses, in which the correct parameters are known, show that the method is able to accurately determine the frequency of the present modes, and it is able to determine the other parameters of interest to within 5% of the actual value, given a signal-to-noise ratio of 40 dB.

The model selection process allows the method to determine the number of modes present in a RIR. More complete models accumulate greater evidence, while overly-complex models are automatically penalized.

6 CONCLUSIONS AND FUTURE WORK

From results seen so far, Bayesian room-acoustic modal analysis shows great promise as a method for characterizing the modal characteristics of rooms. However, there is still work to be done.

A strategy will be developed to assess an entire room impulse response. This strategy will likely include applying a bank of bandpass filters to the impulse response, then analyzing each of the filtered signals. Such an approach would avoid the increases in computation time brought on by increasing modal density and could be further streamlined using parallel computing.

7 ACKNOWLEDGEMENTS

Thanks to Cameron Fackler for his guidance on the topic of Bayesian analysis.

8 REFERENCES

1. B. Fazenda and W. Davies, “The views of recording studio control room users”, Proceedings of the Institute of Acoustics **23**, 1–8 (2002).

2. G. Bretthorst, “Bayesian analysis. i. parameter estimation using quadrature nmr models”, Journal of Magnetic Resonance (1969) **88**, 533–551 (1990).

3. G. Bretthorst, “Bayesian analysis. ii. signal detection and model selection”, Journal of Magnetic Resonance (1969) **88**, 552–570 (1990).

4. C. Andrieu and A. Doucet, “Joint bayesian model selection and estimation of noisy sinusoids via reversible jump mcmc”, Signal Processing, IEEE Transactions on **47**, 2667–2676 (1999).

5. J. Skilling, “Nested sampling for general bayesian computation”, Bayesian Analysis **1**, 833–860 (2006).

Table 1 – Simulated RIR: Two well-separated modes

	Known values	Inferred values	Error (%)
First mode			
Amplitude	0.3	0.306	2.0
Decay time (s)	0.2	0.195	2.5
Frequency (Hz)	40	40.0	0.0
Phase delay (rad)	0	6.28	0.5
Second mode			
Amplitude	0.4	0.402	0.5
Decay time (s)	0.5	0.496	0.8
Frequency (Hz)	100	100	0.0
Phase delay (rad)	0	6.28	0.5

Table 2 – Measured RIR: Transmission loss chamber

	Mean	Standard deviation
First mode		
Amplitude	9.61×10^{-4}	9.86×10^{-6}
Decay time (s)	0.910	0.0075
Frequency (Hz)	173	0.0078
Phase delay (rad)	4.99	0.0070
Second mode		
Amplitude	7.41×10^{-6}	4.52×10^{-6}
Decay time (s)	0.609	0.149
Frequency (Hz)	64.6	28.0
Phase delay (rad)	5.40	0.405
Third mode		
Amplitude	0.0015	2.60×10^{-5}
Decay time (s)	0.276	0.0047
Frequency (Hz)	348	0.0572
Phase delay (rad)	1.57	0.0213
Fourth mode		
Amplitude	0.0032	2.20×10^{-5}
Decay time (s)	0.381	0.0029
Frequency (Hz)	312	0.0076
Phase delay (rad)	3.98	0.0040
Fifth mode		
Amplitude	0.0018	3.81×10^{-5}
Decay time (s)	0.670	0.0085
Frequency (Hz)	338	0.0134
Phase delay (rad)	0.506	0.0140
Sixth mode		
Amplitude	0.0016	1.10×10^{-5}
Decay time (s)	0.793	0.0072
Frequency (Hz)	262	0.0031
Phase delay (rad)	4.60	0.0053
Seventh mode		
Amplitude	0.0013	2.21×10^{-5}
Decay time (s)	0.446	0.0085
Frequency (Hz)	342	0.0562
Phase delay (rad)	3.19	0.0250
Eighth mode		
Amplitude	0.0033	1.76×10^{-5}
Decay time (s)	0.550	0.0042
Frequency (Hz)	336	0.0112
Phase delay (rad)	4.57	0.0081

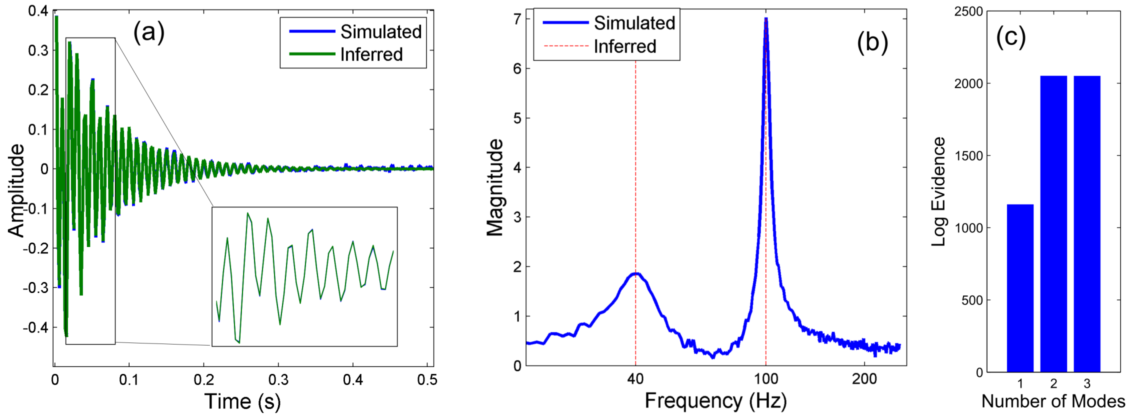

Fig. 1 – Simulated RIR: Two well-separated modes. (a) simulated and inferred time-domain signals; (b) simulated frequency-domain signal with the inferred frequencies represented by dashed lines; (c) log evidence for each tested mode

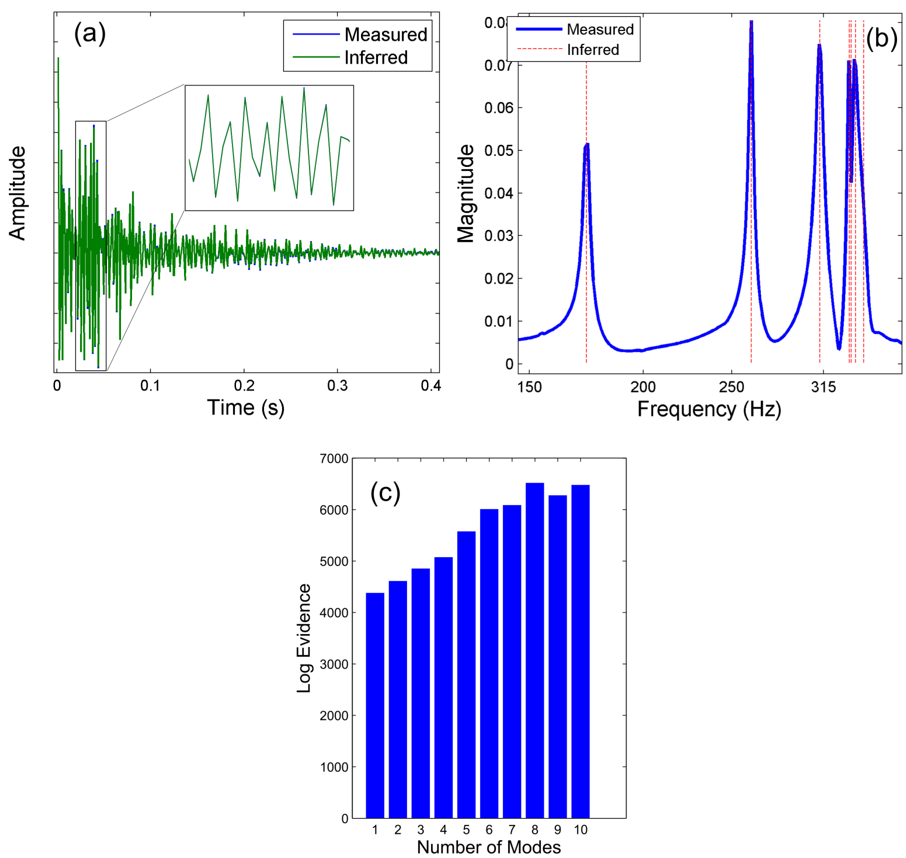

Fig. 2 – Measured RIR: Transmission loss chamber. (a) measured and inferred time-domain signals; (b) measured frequency-domain signal with the inferred frequencies represented by dashed lines; (c) log evidence for each tested mode