inter·noise 2021
1–5 August | Washington, DC

A hybrid method for broadband vibroacoustic simulations

A. Jared Miller[1]
Brigham Young University
Dept. of Physics and Astronomy
Provo, UT 84602

Scott D. Sommerfeldt[2]
Brigham Young University
Dept. of Physics and Astronomy
Provo, UT 84602

Jonathan D. Blotter
Brigham Young University
Dept. of Mechanical Engineering
Provo, UT 84602

David C. Copley
Caterpillar Tech Center
14009 Old Galena Rd, Bldg E-852
Mossville, IL 61552

ABSTRACT

Many methods for simulating acoustic responses of vibrating systems are only suitable for limited frequency ranges, providing either an accurate low frequency or high frequency response. A hybrid method is presented to combine a low frequency modal response and a high frequency statistical energy response to obtain a unified broadband response. The method is designed to produce an auralizable response. An experimental setup is used to validate the method. Listening tests are conducted to assess the realism of the auralizations compared to measurements. The listening tests confirm that the method is able to produce realistic auralizations, subject to a few limitations.

1. INTRODUCTION

Producing a realistic simulation of the vibroacoustic response of a system is valuable in many industries and applications. Various techniques are used to model a system, and each has its own advantages/disadvantages, but the overall issue common to many methods is that they are only valid or feasible for limited frequency ranges [1]. For many numerical methods, computation time significantly increases for higher frequency ranges – for example computation time is proportional to frequency cubed for rectangular acoustic cavities in a modal analysis [2]. Such methods are termed “low-frequency” because they are often not practical for obtaining a response up to higher frequencies

[1] ajm913@gmail.com

[2] scott_sommerfeldt@byu.edu

(higher frequencies being defined separately for each unique system depending on the geometry and material/fluid properties). In contrast to these low-frequency methods, which become impractical with increasing frequency, energy-based methods such as statistical energy analysis (SEA) typically improve with increasing frequency due to higher modal densities, but they are less accurate in low-frequency ranges.

The research presented here seeks to create a hybrid method, combining a low-frequency method and a high-frequency method to obtain a broadband acoustic response of a vibrating system. The hybrid method is intended to create auralizable responses that can be assessed by listening. The next section provides some additional background about work leading to the hybrid method. Details about the hybrid method follow, along with references to other research attempting to create broadband responses. An experimental setup is used to validate the hybrid method and results are presented comparing the hybrid method to measurements.

2. DEVELOPMENT

The systems simulated in this paper are meant to represent basic structural/acoustic coupling found in heavy equipment, but the methods developed are generally applicable for many applications. Specifically, the ultimate objective here is to simulate the acoustic response inside the cab of a machine, or the aural response that would be experienced by the operator. As such, the developed method needs to create a broadband response of the vibroacoustic system that can then be auralized and evaluated by listening, unlike many other methods that are only evaluated graphically. The measure of success is then directly related to the perception of the simulated sounds, with the goal of creating sounds that are perceived as "realistic", not necessarily perfect, and avoiding an overall perception of artificialness. As such, the sounds were evaluated through listening tests where participants were asked to rate the simulated sounds.

The benefits of accurately simulating the acoustic response of a machine cab are twofold: first, the acoustic response can be used in a broader simulation that is used to train operators to use the machinery, and second, the model can be used as a design tool to help achieve a desired sound. However, the developed method has implications far beyond this specific application. The hybrid approach could be used to combine any two traditional methods (one low-frequency and one high-frequency) to obtain a realistic wide bandwidth acoustic response.

Simplicity and efficiency are two criteria that guided development of the simulation method. Prior to creating the hybrid method, various methods were tested for creating simple approximations of a machine cab response. Two methods emerged as desirable solutions based on these criteria: classical modal analysis (CMA) and statistical energy analysis (SEA).

In CMA, the in vacuo structural modes and rigid boundary acoustic modes are first determined analytically. These independent analytical modes are then combined through spatial coupling coefficients. The final response is then obtained by summing over the total number of modes to be used for the desired frequency bandwidth [3,4]. A convenient matrix formulation of CMA was developed by Kim and Brennan that is based on the impedance/mobility approach [5]. This allows for efficient calculation of coupled responses when the independent analytical modes can be determined. The low-frequency responses presented later in this paper were obtained using this matrix formulation of CMA. If determining the analytical modes is infeasible, FEA could be used to determine the low-frequency responses.

SEA was used to create the high-frequency responses. SEA is very computationally efficient, making it an ideal candidate for the simple approximate model developed in this paper. One major limitation of SEA is that since it results in average levels, the final responses do not contain the phase information necessary to create the impulse response needed for creating auralization results. The hybrid method outlined below seeks to overcome this limitation and allow for SEA responses to be combined with a low-frequency response to create auralizable broadband acoustic responses.

2.1. Hybrid Method

There are many applications where only considering either low or high frequencies is not enough, and broadband responses are required. This is particularly true when the final response will be auralized, since human hearing spans approximately 20 Hz – 20 kHz, and a reduced frequency range is often perceived as unnatural. Significant research has been done investigating ways to achieve broadband simulations involving acoustic radiation from coupled vibrating structures. No universal method has been found and it remains an active area of research [6]. Many of the proposed methods are quite complex and/or application specific, keeping them from being more widely adopted. Wang, et al. use a hybrid approach combining a node-based smoothed finite element method (FEM) and SEA and show good results for several theoretical systems; however, they apply the different methods to separate subcomponents of the system, leaving the SEA portions absent of any phase information [7]. Chronopoulos, et al. incoorporate a wave FEM with SEA to better account for dispersion in curved shells [8]. Yotov, et al. introduce a non-parametric stochastic FEM allowing them to accurately model responses of spacecraft in high-frequency ranges where structures begin to exhibit chaotic behavior and element-based techniques are typically unreliable [9]. Aretz, et al. combine FEA, image sources, and stochastic ray tracing to simulate broadband impulse responses [10]. This work is most similar (in objective, not method) to the research presented in this paper, but the method does not achieve the simplicity aimed for here and would be difficult to implement in more complex systems. They provide additional references to similar work, citing limitations and unsatisfactory results in most cases.

There are some established methods that combine low- and high-frequency methods to compute the response of vibroacoustic systems. Certain computer software packages, VA One for example, will simulate complex systems by combining individual components that are each modeled by either FEA or SEA [11]. The user determines which method (FEA or SEA) will be used for each component depending on its geometry and material properties. Such computer simulations can provide accurate responses for complex systems; however, they can become extremely computationally expensive/time consuming when a large frequency range is desired. Additionally, the energy-based portion of the solution only provides an average level, and it does not capture any resonance or phase information. This becomes problematic if one desires to auralize the simulated response.

There are two main drawbacks in many of the existing broadband solutions. First, the methods are often quite complex. They either require significant computation time, or they involve complicated mathematical techniques that are only applicable in specific situations. As previously stated, the goal of this project is to create a simple method to model vibroacoustic systems that is both computationally efficient and simple enough to easily change and apply in various configurations. Of course, there must be a tradeoff here: the simpler the model, the less likely it will be able to capture all the complexities of the system. Accordingly, the measure of success is creating a method where the resultant models sound “realistic”, not perfect.

The second drawback is related to the way that many of the existing broadband solutions are evaluated. Plotting the magnitude of the frequency response of the system is the most common way that model accuracy is evaluated. Even when different methods are used in different frequency ranges, the results are often just plotted side by side, without providing any real way to combine the results into a single overall response [12]. This may be sufficient in many instances; however, our main concern is about how the simulated sounds are perceived compared to real sounds. Therefore, our method needs to produce a result that can be auralized. It must be a single response that contains both magnitude and phase information across the frequency range of interest.

The hybrid method developed here seeks to overcome these two problems. In the end, it creates a simple model that produces auralizations that are reasonable approximations of how the real system sounds. There are four steps in the process: 1) creating a separate low-frequency modal response and a high-frequency SEA response of the system, 2) interpolating between the two responses to get a single broadband magnitude response, 3) adding amplitude modulation to the SEA portion of the response, and 4) calculating approximate phase information. Each of these steps is discussed below and pictured in Figure 1.

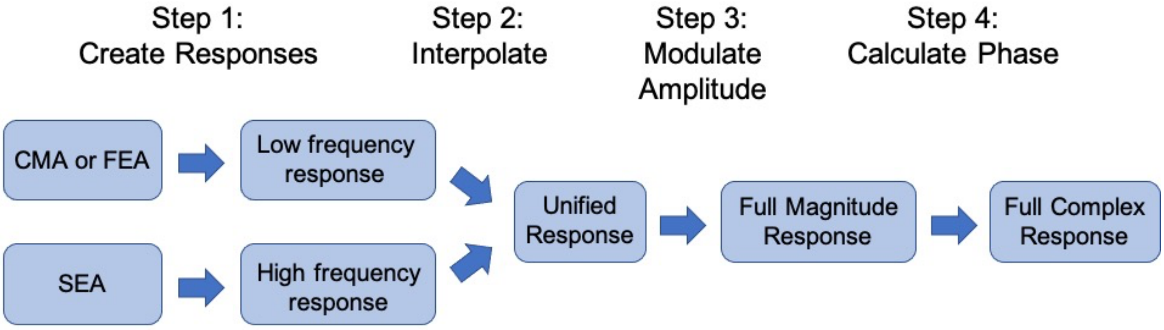

Figure 1: Diagram representing the steps in the hybrid model process: Step 1) creating a separate low-frequency response and a high-frequency response, Step 2) interpolating between the two responses to get a single broadband magnitude response, Step 3) adding amplitude modulation to the high-frequency portion of the response, and Step 4) calculating approximate phase information.

First, two separate responses are calculated, one using a low-frequency method and one using a high-frequency method. For this project, classical modal analysis, based on a matrix formulation developed by Kim and Brennan, was used to calculate the low-frequency response [5]. Although FEA is more commonly used due to its accuracy and ease of implementation with modern software packages, CMA was chosen because of its simplicity and computational efficiency. The high-frequency response was obtained by building an SEA model in the computer program VA One. The SEA response was calculated in one-third octave bands. SEA also meets the simplicity and efficiency criteria.

Second, a single magnitude response was created by interpolating between the separate low- and high-frequency response magnitudes. At this point, only the magnitude response can be obtained because the SEA portion of the response does not contain any phase information. Built in MATLAB interpolation methods were used to obtain the single unified response. It is important to choose the interpolation method carefully to avoid unexpected results (MATLAB documentation recommends using interp1 with the 'pchip' interpolation method when the signal x is not slowly varying). Determining the crossover frequency, or the point at which to switch from the modal response to the SEA response, is another important consideration in this step. Various crossover frequencies were tried; the results presented in this paper used a 2000 Hz crossover frequency. Once the crossover frequency was determined, each of the individual responses was truncated; everything above the crossover frequency was discarded from the modal response, and everything below the next one third octave band center frequency was discarded from the SEA response, leaving a gap between the crossover frequency and the next one third octave band center frequency. This gap allowed for a smoother transition between the separate low- and high-frequency responses. The two separate responses were then combined via MATLAB's interp1 function, and the resulting unified response was resampled to a 1 Hz frequency resolution to match the resolution of the SEA portion to that of the modal portion.

Third, the SEA response only captures the average level across frequency, it does not capture any information about resonances/antiresonances. This makes for a very smooth unrealistic response. Of course, it is unknown where the resonances/antiresonances would have occurred – a classical modal model or finite element model would be required to know. However, a more realistic response can be obtained by randomly adding amplitude modulation to the SEA response. Although randomly modulating the response will not create peaks at the exact same frequencies as the real system, it was found that it is sufficient to create a more realistic sounding response. This is because the SEA response is only used in a frequency range where the modal density is high. In this frequency range, the exact location of the peaks is less important than in the lower frequency range covered by the modal model. There are two important considerations when creating the amplitude modulation: the magnitude of the modulation and how rapidly the modulation occurs along the frequency axis. The magnitude of the modulation is representative of the damping in the system. Large amplitude modulation represents a system with little damping and results in a high-pitched "metallic" ringing sound in the final simulation. On the other hand, low amplitude modulation represents a system with high damping and results in little to no ringing in the final simulation. It was found that it is better to overestimate the damping (underestimate the amplitude of modulation) in the SEA portion of the response, because extensive high-frequency ringing tends to cause the simulated sounds to be perceived as artificial sounding. The amplitude modulation formula used for the results presented in this paper is given by

$$A' = A * \mathrm{lognrand}(\mu, \sigma)\,, \tag{1}$$

where A is the original amplitude, A' is the modified amplitude, and lognrand() is a MATLAB function producing lognormal random numbers with parameters μ (mean of logarithmic values) and σ (standard deviation of logarithmic values). The parameter values used to produce the results presented in this paper were $\mu = 0$ and $\sigma = 0.5$. Determining how rapidly to modulate the amplitude along the frequency axis is a second concern. Although the exact resonances of the coupled system are not known, the uncoupled natural frequencies of the dominant components can be used to estimate an appropriate density of peaks and dips in the frequency response. In the plate-cavity system described in the experimental setup section below, the resonance frequencies of the plate served as an appropriate approximation.

Fourth, in order to auralize the response it needs to have phase information as well as magnitude. As mentioned before, an SEA response does not contain any phase information. Therefore, to finalize the response, an approximate phase needs to be calculated. Significant time was spent experimenting with various ways of creating this approximate phase. The final method involved calculating a minimum phase via the Hilbert transform. The process of using the Hilbert transform to create the final response consisted of three parts. First, the magnitude response calculated in the previous step was used to create a two-sided spectrum, since the Hilbert transform expects negative frequencies. Second, the Hilbert transform was used to calculate a minimum phase for the given magnitude response. The formula for calculating the minimum phase is given by

$$\phi(\omega) = -\mathcal{H}\big[\ln\big(G(\omega)\big)\big]\,, \tag{2}$$

where ϕ is the minimum phase, $\mathcal{H}$ represents the Hilbert transform, and G is the two-sided magnitude response. Third, the final complex frequency response was calculated according to

$$\hat{G}(\omega) = G(\omega) * e^{j\phi(\omega)}\,, \tag{3}$$

where $\hat{G}$ is the two-sided complex frequency response, G is the two-sided magnitude response, and ϕ is the minimum phase. The minimum phase was used across the entire frequency range. This proved less problematic than attempting to interpolate between the existing low-frequency phase and the calculated minimum phase at high frequencies.

An Inverse Fast Fourier transform (IFFT) was then applied to the complex frequency response to obtain an impulse response. The impulse response was convolved with various excitation signals to create auralizations, so the validity of the approach could be assessed by listening.

3. EXPERIMENTAL SETUP

A simple coupled structural-acoustic system was built to validate the hybrid method. The system consisted of a rectangular acoustic cavity with five rigid walls and one flexible wall. Similar systems have been studied extensively and used many times to validate new methods [4,13,14]. The rigid walled acoustic cavity was built with a similar method to that used by Kim and Brennan, and the flexible wall was constructed to mimic a simply-supported plate, based on a method proposed by Robin et al. [5,15]. Details of the system are provided below.

A diagram of the experimental setup is show in Figure 2. Two five sided boxes were constructed using ½ inch medium-density fiberboard (MDF), one larger box and one smaller box designed to sit inside the larger box with a 10 cm gap on all sides. The 10 cm gap between the boxes (including the bottom) was filled with sand so that the inner box acted as a rigid walled acoustic cavity. The inner dimensions of the smaller box were 48 cm x 42 cm x 110 cm. A microphone was located at (20 cm, 18 cm, 63 cm) according to the coordinate system marked.

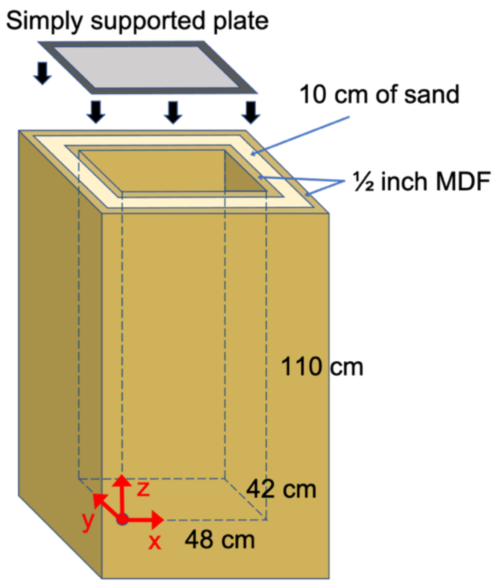

Figure 2: Diagram of the experimental setup, a simply supported plate coupled to an acoustic cavity.

An aluminum simply supported plate, mounted to a steel frame, was placed on top of the cavity to create the flexible wall (Figure 3). The plate and cavity were designed to minimize any gaps, but prevent touching on the sides, once the plate was placed atop the cavity. This was done so that the plate dimensions and x-y dimensions of the cavity could be assumed equal when modeled, while preserving the simply supported nature of the plate. The plate was measured to be 3.15 mm thick. The plate was excited by a mechanical shaker at (20 cm, 18 cm, 110 cm), directly above the microphone. A force sensor (not pictured) was attached between the shaker and the plate. The transfer function was measured between the force on the plate and the microphone in the cavity.

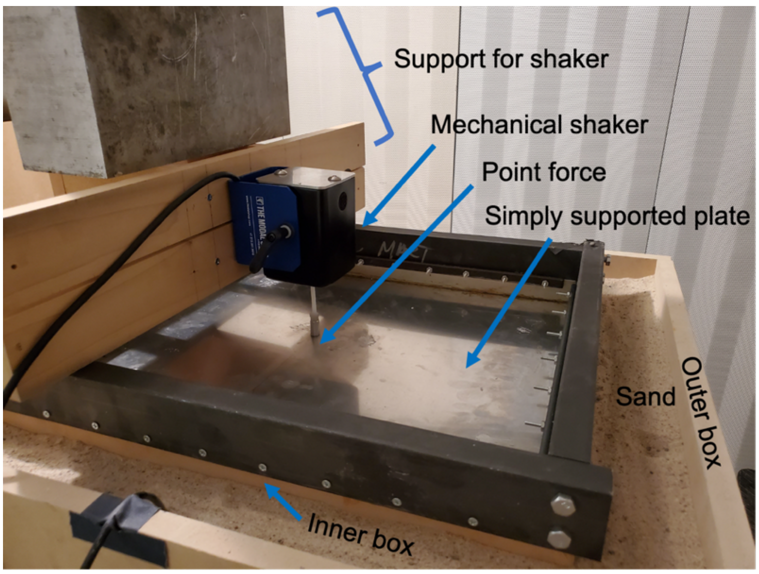

Figure 3: Photograph of the simply supported plate excited by a mechanical shaker.

4. RESULTS

A model of the plate/cavity experimental setup was built using the hybrid method. The shaker was modeled as a point force and the microphone was modeled as a point acoustic sensor. For the low-frequency portion of the response, the matrix CMA formulation was used [5]. The CMA response was calculated up to 2 kHz. VA One's default material properties for aluminum were used to be consistent with the SEA model: density = 2700 kg/m^3, Poisson's ratio = 0.33, and Young's modulus = 7.1×10^{10} Pa. An airborne sound speed of 340 m/s was used, and the density of air was assumed to be 1.21 kg/m^3, consistent with lab conditions of a room temperature of 20°C and an elevation of 1400 m. A damping ratio of 0.01 was used, determined by comparing to the measurement since it can be difficult to estimate damping accurately. The high-frequency portion of the response, above 2 kHz, was obtained by creating a SEA model in VA One, using all the same parameter values. The low-frequency CMA and high-frequency SEA responses are shown alongside the measured response in Figure 4.

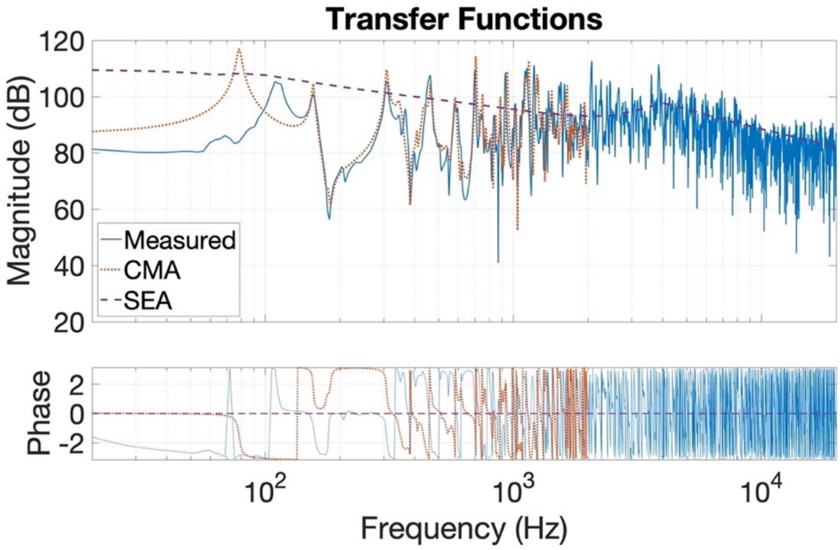

Figure 4: Classical modal analysis (CMA) response and statistical energy analysis (SEA) response compared to measured response of the experimental setup. The transfer functions go from the force on the plate to the microphone in the cavity.

The individual CMA and SEA responses were combined using the hybrid method introduced in Section II. The result from the hybrid model is compared to a measurement of the experimental setup in Figure 5. The pictured response is the transfer function from the input force on the plate to the microphone in the acoustic cavity. These transfer functions were used to calculate impulse responses, which were convolved with various excitation signals (recordings of engine noise and other sounds of interest from heavy machinery). Listening to these auralizations is the main way that the validity of the approach was evaluated. However, presenting audio recordings is not possible in a written format, so the frequency responses will be discussed.

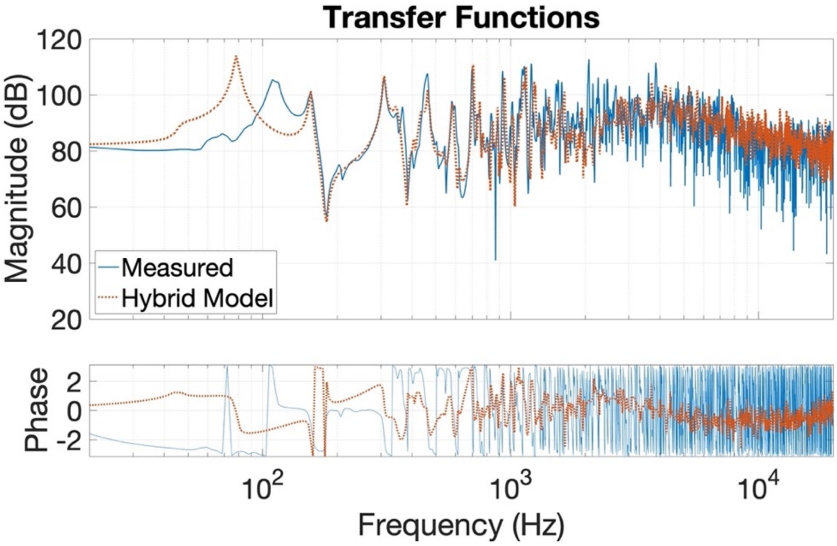

Figure 5: Full hybrid model result compared to experimental measurement. The transfer functions go from the force on the plate to the microphone in the cavity.

One of the most notable features in the frequency response is the mismatch in the frequencies of the lowest peak between the hybrid model and the measurement. The model predicts a peak at 78 Hz while the measurement showed a peak at 109 Hz. Investigation of the simply-supported plate revealed that 78 Hz corresponds to the theoretical natural frequency of the 1-1 plate mode, while it was found experimentally that the measured frequency of this mode was 109 Hz. This result occurs because the experimental plate does not correspond well with a simply-supported plate for this first resonance, while the results indicate that the plate behaves closely to a simply-supported plate for higher resonances [15]. Applying a high-pass filter at 100 Hz to the auralizations proved sufficient in minimizing the differences caused by these mismatching fundamental frequencies.

As previously stated, a full complex frequency response, including both magnitude and phase information, is necessary to transform to the time domain to obtain an impulse response for auralization. Although both are necessary, the magnitude portion of the responses tends to dominate human perception of sound, while the phase plays a secondary role. This means that matching the magnitude portion as closely as possible is vital but finding an appropriate approximation of the phase can be sufficient. An investigation of a number of possible methods to determine phase led to the choice to use minimum phase, as this was one of only a few methods that preserved the sense of naturalness in simulations. While most physical systems are not truly minimum phase, it has nice properties such as preserving causality and invertibility that allow it to produce auralizations without introducing such artifacts. The minimum phase calculated for the hybrid model shown in Fig. 5 is sufficient to create a natural sounding auralization for the system of interest, confirmed via listening tests.

Many of the simulated auralizations sounded similar to the measurements, although the exact level of similarity was somewhat difficult to assess. Listening tests were conducted to evaluate the similarity, focusing on realism/artificialness. Eleven listeners participated in the listening tests to capture a variety of perceptions and opinions. Two trends became apparent when examining listening test responses. First, the perceived pitch of the sounds was dominated by the peaks with the highest magnitude in the frequency response, which had a significant impact on how similar the simulations were perceived compared to the measurements. However, even though the pitch differences significantly affected perception of the overall similarity of the sounds, they did not significantly affect the perceived realism of the sounds. For example, the peak at 2069 Hz in the measured response shown in Fig. 5 is notably missing from the response of the hybrid model because a 2000 Hz crossover frequency was used for the model. This caused a significant difference in the pitch of the measured sounds vs. the sounds created from the hybrid model, resulting in lower ratings for overall similarity. Increasing the crossover frequency to 2100 Hz allowed the model to capture the 2069 Hz peak, resulting in noticeably more similar pitches and therefore better ratings of overall similarity. Despite better ratings for overall similarity between the measurement and the model, there was no difference in the perceived realism of the simulated sounds for the 2100 Hz crossover compared to the 2000 Hz crossover. This shows that exactly matching the dominant peaks is not necessary to create realistic simulations, even if pitch differences are introduced.

Second, the excitation signal had a significant impact on whether the sounds were perceived as realistic or artificial. In particular, it was found that sounds created with input signals containing transients were much more likely to be perceived as realistic, while sounds created from entirely steady state input signals were more likely to be perceived as artificial. This was the case across the board, for both measured sounds and simulated sounds; even measured sounds were more likely to be perceived as artificial if the excitation signal contained no transients.

5. CONCLUSIONS

The proposed hybrid method successfully merged a low-frequency response and a high-frequency response into a single response that could be auralized. This allowed for a simple and efficient approximation of the desired acoustic response over a broad frequency range. The auralizations were able to retain a sense of realism, skirting some of the unnatural artifacts prevalent in audio simulation. There were some limiting factors in the level of realism achieved. First, matching the largest peaks in the frequency responses is necessary to create the same pitch, and it was found that pitch differences are a significant factor when listeners compare the similarity of two sounds (simulation vs. measurement). Second, the presence/absence of transients in the excitation signal significantly affected the perceived realism of the final auralization. While important considerations, neither of these two limiting factors are directly related to the ability of the hybrid method to produce realistic auralizations.

6. REFERENCES

1. Miller, A. J. A Hybrid Method for Auralizing Vibroacoustic Systems and Evaluating Audio Fidelity/Sound Quality Using Machine Learning. Master's thesis, Brigham Young University, Provo, UT, (2021).
2. Belvins, R. D. Modal density of rectangular volumes, areas, and lines. *Journal of the Acousical Society of America,* **119(2)**, 788-791 (2006).
3. Fahy, F. & Gardonio, P. *Sound and Stuctural Vibration: Radiation, Transmission, and Response*, Elsevier, 2007.
4. Du, J. T., Li, W. L., Xu, H. A. & Liu, Z. G. Vibro-acoustic analysis of a rectangular cavity bounded by a flexible panel with elastically restrained edges. *Journal of the Acousical Society of America,* **131(4)**, 2799-2810 (2012).
5. Kim, S. M. & Brennan, M. J. A compact matrix formulation using the impedance and mobility approach for the analysis of structural-acoustic systems. *Journal of Sound and Vibration,* **223(1)**, 97-113 (1999).
6. Vorländer, M. *Auralization: Fundamentals of Acoustics, Modelling, Simulation, Algorithms and Acoustic Virtual Reality*, Springer, 2008.
7. Wang, G., Zhang, Y. X., Guo, Z. B. & Zhou, Z. G. A novel hybrid deterministic-statistical approach for the mid-frequency vibro-acoustic problems. *Applied Mathematical Modelling,* **83**, 202-219 (2020).
8. Chronopoulos, D., Troclet, B., Ichchou, M. & Lainé, J. P. A unified approach for the broadband vibroacoustic response of composite shells. *Compositets Part B: Engineering,* **43(4)**, 1837-1846 (2012).
9. Yotov, V., Remedia, M., Aglietti, G. & Richardson, G. Non-parametric stochastic FEM / hierarchical matrix BEM for efficient mid-frequency vibroacoustic response estimation. *25th International Congress on Sound and Vibration*, Hiroshima, Japan, 2018.

10. Aretz, M., Nöthen, R., Vorländer, M. & Schröder, D. Combined Broadband Impulse Responses Using FEM and Hybrid Ray-Based Methods. *EAA Symposium on Auralization*, Espoo, Finland, 2009.
11. VA ONE, ESI Group, 2020. [Online]. Available: https://www.esi-group.com/products/vibro-acoustics.
12. Jiao, R. & Zhang, J. Vibro-acoustic modeling of a rectangular enclosure with a flexible panel in broad range of frequencies and experimental investigations. *Journal of Vibroengineering.* **18(4)**, 2683-2692 (2016).
13. Pan, J. & Bies, D. A. The effect of fluid–structural coupling on sound waves in an enclosure — Theoretical part. *Journal of the Acousical Society of America,* **87(2)**, 691-707 (1990).
14. Zhang, H., Shi, D., Zha, S. & Wang, Q. Vibro-acoustic analysis of the thin laminated rectangular plate-cavity coupling system. *Composite Structures,* **189**, 570 - 585 (2018).
15. Robin, O., Chazot, J. D., Boulandet, R., Michau, M., Berry, A. & Atalla, N. A Plane and Thin Panel with Representative Simply Supported Boundary Conditions for Laboratory Vibroacoustic Tests. *Acta Acustica united with Acustica* **102**, 170-182 (2016).