**Acoustic modal analysis of room responses from the perspective of state-space balanced realization with application to field interpolation**

Hung-Ming Jian, You-Siang Chen, and Mingsian R. Bai

*Department of Power Mechanical Engineering, National Tsing Hua University,* 101*, Section* 2*, Kuang-Fu Road, Hsinchu, Taiwan* 30013*.*

*Suggested running title:* State-SpaceAcoustic Modal Analysis

1. Corresponding author: Mingsian R. Bai, Email: [msbai@pme.nthu.edu.tw](mailto:msbai@pme.nthu.edu.tw)

**ABSTRACT**

Despite the importance in structural dynamics and vibration, modal analysis is rarely investigated in acoustics due to the high modal density of sound fields. A novel acoustic modal analysis (AMA) approach is proposed in this paper for enclosed acoustic fields in such as a room, from the perspective of state-space formulation of control systems. A single-input-multiple-output (SIMO) state-space model is established in light of the balanced realization (BR), given impulse response measurements. The BR model is then converted to a modal form such that the modal parameters, including natural frequencies, damping ratios, and mode shapes, can be estimated. In order to reconstruct mode shapes, plane-wave decomposition (PWD) and the compressive sensing (CS) techniques are exploited to solve the underdetermined problem for a spatially sparse representation of mode shapes under the Schroeder frequency. As a result, a model of continuous system can be “interpolated” for any arbitrary source-receiver positions on the basis of the estimated mode shapes. With the identified modal parameters, the low-frequency and early reflection part of room impulse responses (RIRs) can be synthesized for arbitrary source-sensor pairs. The proposed AMA acoustic field interpolation is validated by extensive simulations and experiments.

PACS number: 43.55.Br, 43.55.Ka, 43.60.Fg

# INTRODUCTION

As an essential technique in the study of structural dynamics and vibration, modal analysis aims to find the modal parameters of structures. However, acoustic modal analysis (AMA) is rarely reported except for some research restricted to low-frequency analysis. Kung and Singh conducted a modal analysis based on frequency responses of three-dimensional cavities.1 Xu and Sommerfeldt proposed a hybrid modal expansion approach to yield modal parameter estimates with higher accuracy than the classical modal analysis.2 More recently, learning-based methods3-5 were applied to extract spatial features of enclosures and to reconstruct the associated sound fields.

AMA is based on the modal expansion of the acoustic field in a room. For regular room geometries, the eigenvalues (natural frequencies) and eigenfunctions (mode shapes) are available in closed forms.6,7 It is understood that AMA is useful only for the low-frequency regime in which the enclosed acoustic field exhibits modal behavior. For frequencies above the Schroeder frequency,8  AMA would not be as effective due to extremely high modal density for large rooms. However, natural frequencies, damping ratios, and mode shapes still contain important geometric information and dynamic response of a room. Motivated by this fact, this study focuses on the use of modal parameters obtained in AMA in one low-frequency application – field interpolation.

This paper is divided into two parts. In the first part, an AMA approach is proposed for the enclosed acoustic field in a room from the perspective of state-space formulation of dynamic control systems. The input data required in this approach are room impulse response (RIR) measurements captured by multiple microphones. The balanced realization (BA) 9, 10 in linear system theory is exploited to construct a single-input-multiple-output (SIMO) model of the enclosed sound field. Next, the resulting state-space model is converted to a modal form from which modal properties such as natural frequencies, damping ratios, and mode shapes at sensor locations can be extracted by using eigenvalue decomposition (EVD). The combined BR-EVD procedure is referred to as the eigensystem realization algorithm (ERA) in structural dynamics literature.10 Based on ERA, this work takes a step further by employing a novel technique to interpolate the modal vectors obtained using ERA into continuous mode-shape functions in light of plane-wave decomposition (PWD) and compressive sensing (CS) algprithm.11, 12 With the interpolated mode-shape functions, a Green function can be constructed to model the RIRs for any given source-sensor pairs, which accounts for the early reflections and the low-frequency responses of the room.

The second part of this paper concerns one potential application of the proposed AMA technique to the acoustic field interpolation problems.13-16 Acoustic field interpolation has been a topic of active research in room response synthesis and array signal processing. Approaches reported in the literature17-21 generallyrelied on fitting mode shapes with discrete measurements via a wave propagation model. The sound field can then be expressed as a sparse representation12 and solved by using CS.11 In this paper, two application scenarios of field interpolation, sensor interpolation and source interpolation, are investigated. Sensor interpolation problems22, 23 refer to interpolating the RIR from a sound source to some unmeasured sensor point, whereas source interpolation problem24 refers to interpolating the RIR from some virtual source input to a physical sensor location. The required mode-shape functions for the interpolation procedure are obtained using the proposed AMA approach.

This paper is organized as follows. In Sec. II, the theoretical background of AMA is given. In Sec. III, the simulation and experiment results are presented to corroborate the proposed method. In Sec. IV, application of the proposed AMA approach to an acoustic field interpolation problem is given. Conclusions are made in Sec. V.

# THEORETICAL BACKGROUND

1. **Enclosed Sound Fields**

Consider a rectangular room with dimensions , where a unit time-harmonic monopole source located at **r***s* emits sound pressure described by the Helmholtz equation

|  |  |
| --- | --- |
| , | (1) |

where  is the Laplacian operator written in the three-dimensional Cartesian coordinates, *P*(**r**) is the Fourier transform of the sound pressure received at the microphone position , *k* = *ω*/*c*, *ω* denotes the angular frequency, *c* is the speed of sound, is the source position, and *δ*(∙) is the Dirac delta function.

The boundary condition of the walls is assumed to be locally reacting

|  |  |
| --- | --- |
| , | (2) |

where , denotes the normalized specific acoustic admittance, *ρ*0 is the air density, and *z*0 is the specific acoustic impedance at the boundary. It can be shown that the solution *P*(**r**) satisfying Eqs. (1) and (2) is the Green’s function that can be written as the following series expansion25

|  |  |
| --- | --- |
| , | (3) |

where *ωn* denotes natural frequency of mode *n*,  denotes modal damping ratio, and  denotes the eigenfunction of mode *n* under the rigid-wall and the orthonormality conditions

|  |  |
| --- | --- |
| , | (4) |

where *δmn* is the Kronecker delta and *V* is the volume of the room. It is worth noting from Eq. (3) that the pressure response can be viewed as infinite number of coupled second-order resonators. The room frequency response (RFR) can then be synthesized for any source-receiver pair if the mode shape functions are available.

For the rectangular room considered herein, the closed-form solution of Eq. (1) is available:

|  |  |
| --- | --- |
| , | (5) |

where  being the normalization constant and (*nx*, *ny*,*nz*) are integers representing different modes of the room. The corresponding natural frequency of the *n*th mode is

|  |  |
| --- | --- |
| . | (6) |

We can derive the following RIR in the time domain by taking the inverse Fourier transform of Eq. (3)

|  |  |
| --- | --- |
| , | (7) |

where  is the damped natural frequency of the *n*th mode. In the recent studies, a sum of exponentially decaying oscillations in form of Prony model has been applied in room-acoustic model analysis3. It follows that the RIR can be also regarded as a linear combination of exponentially decaying sinusoids where the spatial information in terms of eigenfunctions is included. In practice, only a finite modes are included to evaluate RFRs in Eq. (3) and RIRs in Eq. (7).

1. **Eigensystem Realization Algorithm (ERA)**

ERA is composed of two main steps: balanced realization (BA) and modal transformation. The purpose of BA is to obtain a balanced state-space realization of a linear time invariant (LTI) system, based on measured impulse response data of a room. Consider a *p*-input-*q*-output system described by the following discrete-time state-space equation9:

|  |  |
| --- | --- |
| , | (8) |

in which  denotes the state vector of *m* state variables,  the source input vector, and  is the microphone output vector, with *n* being the discrete-time index. **A**, **B**, **C**, and **D** denote four constant matrices that constitute a realization of the LTI system.

By using impulse response sequences, two Hankel matrices, **H**1 and **H**2 , can be constructed10:

|  |  |
| --- | --- |
| , | (9) |

and the shifted Hankel matrix:

|  |  |
| --- | --- |
| , | (10) |

where **Y***n* is a *q* × *p* impulse response matrix with *n*-tapped finite-length impulse response sequences as its entries. In order to establish the Hankel matrix that fully covers the dynamics of the room system, we found it appropriate select 50% the length of the impulse response sequence for *n*. Perform singular value decomposition (SVD) on these two matrices and retain the principal modes corresponding to the *r* largest singular values.

|  |  |
| --- | --- |
| , | (11) |

where **U** and **V** are unitary matrices and **S** is a rectangular diagonal matrix with singular values arranged in descending order as its diagonal entries,  are submatrices corresponding to,  with  being the principal singular values.

The balanced realization of the linear system can be written in terms of the SVD of the preceding Hankel matrix as follows9:

|  |  |
| --- | --- |
| , | (12) |

where  and  with and  being the identity matrices of orders *p* and *q*.

The second step of ERA is the modal transformation. Through the eigenvalue decomposition (EVD) of the matrix , the preceding balanced realization can be converted into the modal form below10:

|  |  |
| --- | --- |
|  | (13) |

where is the diagonal matrix composed of the eigenvalues and is the matrix comprised of the eigenvectors, , of **A** as its columns. The discrete-time poles can be converted via the relation *λd* = *eλc*Δ*t*, where *λd* is the discrete-time pole, *λc* is the continuous-time pole, and Δ*t* is the sampling period26. As a result, the continuous-time modal parameters (natural frequencies, damping ratio, and mode shapes) can be calculated from the above discrete-time counterparts as follows:

|  |  |
| --- | --- |
|  | (14) |
|  | (15) |
| , | (16) |

where  is the sample rate in Hz,  denotes the real-part operator, and  denote the mode-shape vectors with entries evaluated at the sensor locations. It is worth noting that the preceding modal parameters must appear in complex-conjugate pairs to allow for real-valued mode-shape vectors in reconstructing standing wave patterns. Thus, we need to combine those pairs accordingly as below

|  |  |
| --- | --- |
| , | (17) |

where “\*” denotes the complex-conjugate operator and is the real-valued *n*-th mode shape evaluated at the *i*-th sensor location **r***i*.

1. **Plane-Wave Decomposition (PWD)**

As one of the important contributions of this paper, a novel mode-shape interpolation technique that builds upon the preceding ERA is presented in this section. As opposed to conventional interpolation techniques, the proposed approach seeks to convert discrete eigenvectors to continuous mode-shape functions by taking advantage of the underlying wave physics. To this end, we represent the mode-shape functions in light of plane-wave decomposition (PWD)16, 27. Unlike PWD that expresses more generally the sound pressure with complex coefficients of plane-wave components28, we impose the constraint that the mode-shape functions are real-valued to the eigensystem realization algorithm (ERA), as required by normal modes of standing waves. A standing-wave field that contains *P* candidate components can be expressed as:

|  |  |
| --- | --- |
| , | (18) |

where (*αp*, *βp*) denote amplitude coefficients of the plane waves, the wavenumber vector of the *n-*thmode ,, and () denote the Cartesian coordinates of a point on the unit sphere in the *k*-space.16 Equation (18) can be rearranged into the following matrix form:

|  |  |
| --- | --- |
|  | (19) |
| or |  |
| , | (20) |

where . In general, the number of candidate plane-wave components is chosen to be much larger than the number of sensors, i.e., *P >> M*,which renders Eq. (20) a highly underdetermined system.

1. **Compressive Sensing (CS)**

The underdetermined system of Eq. (20) is considered to be ill-posed in that it admits infinite number of solutions. To tackle the issue, we apply the compressive sensing (CS)11, 12,15  technique by assuming that the solution of **x** is sparse in the low-frequency range. In the CS formalism, this problem amounts to solving the following convex optimization problem: 11

|  |  |
| --- | --- |
| , | (21) |

where  denotes the *l*1-norm and *ε* denotes the noise floor. Alternatively, we can reformulate Eq. (21) in the form of least absolute shrinkage and selection operator (LASSO) 29:

|  |  |
| --- | --- |
| , | (22) |

where *λ* is a regularization parameter that weights the sparsity against the residual error. The unconstrained optimization problem in Eq. (22) is convex and can be solved using off-the-shelftoolboxes such as CVX30. Once the coefficient vector **x** is found, continuous mode-shape functions can be reconstructed by using Eq. (22) for any points other than the sensor locations.

FIG. 1 summarizes the complete procedure of the proposed AMA. The RIRs are either measured or synthetically generated as the input to the BR step. The state-space BR model is converted into the modal form via EVD to extract modal parameters. Steps 2 and 3 constitute the ERA-based state-space model. In Step 4, continuous mode-shape functions are reconstructed in light of PWD in conjunction with CS.

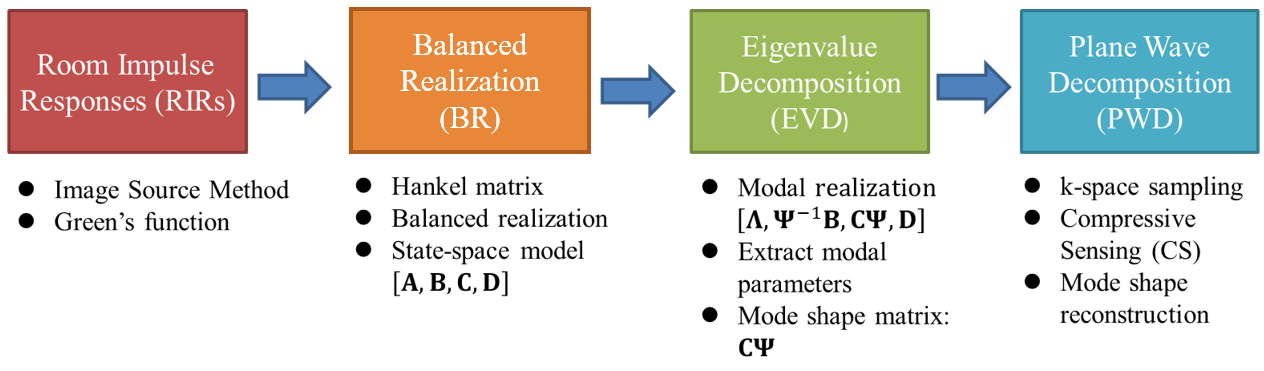


FIG. 1 The complete AMA procedure composed of four processing modules: RIRs, BR, EVD, and PWD.

# SIMULATIONS AND EXPERIMENTS

1. **Simulations**

In the following simulations, we apply the proposed AMA procedure to construct a state-space model of a rectangular room of dimensions, . Modal parameters are calculated with mode shapes reconstructed on a reference plane at the height *z* = 1.27 m, where 6 X 6 microphones with lattice spacing 0.8 m are uniformly distributed. A point source is placed at one of the corners in the room, as depicted in FIG. 2 (a). An example RIR between the first microphone and the sound source, simulated by Eq. (7), is shown in FIG. 2 (b). The reverberation time *T*60 is set to be 0.5 s. The number of modes, *N* = 400, damping ratio , sampling rate , and the speed of sound are assumed. In the simulation, the analytical solution in Eq. (3) to (7) serves as the ground truth in order to validate the proposed ERA-based modal analysis method which has no restriction on the damping of boundary.

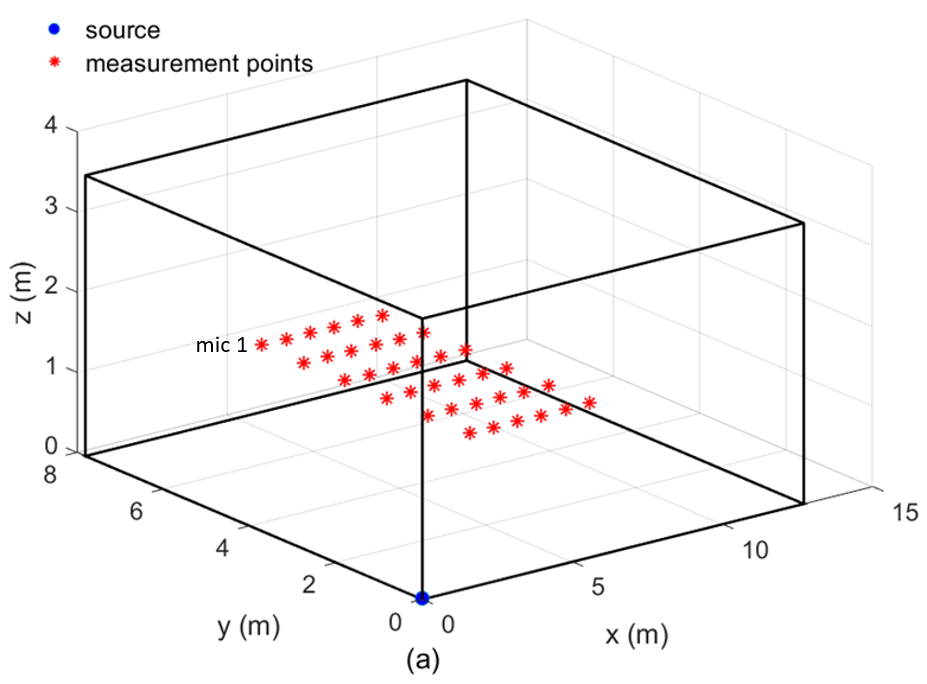




FIG. 2 (a) Simulation settings of a rectangular room with dimensions . Sensors (asterisk) are uniformly distributed on the plane *z* = 1.27 m. A point source, as indicated by a dot, is located at the corner (0.1m, 0.1m 0.1m) of the room. (b) The RIR between the first microphone and the sound source simulated using the Green function. Number of modes *N* = 400, damping ratio , reverberation time T60 = 0.5 s, sampling rate , and the speed of sound .

With the RIRs, we construct the Hankel matrices as discussed earlier and establish a state-space model by using balanced realization. FIG. 3 plots the singular values of the Hankel matrix **H**1 in descending order, from which we choose to keep 180 modes for the subsequent calculations in Eq. (11). FIG. 4 compares the estimated resonance frequency  and the ground-truth  of Eq. (6). Excellent fit is seen throughout the frequency range, 10-1200 Hz.

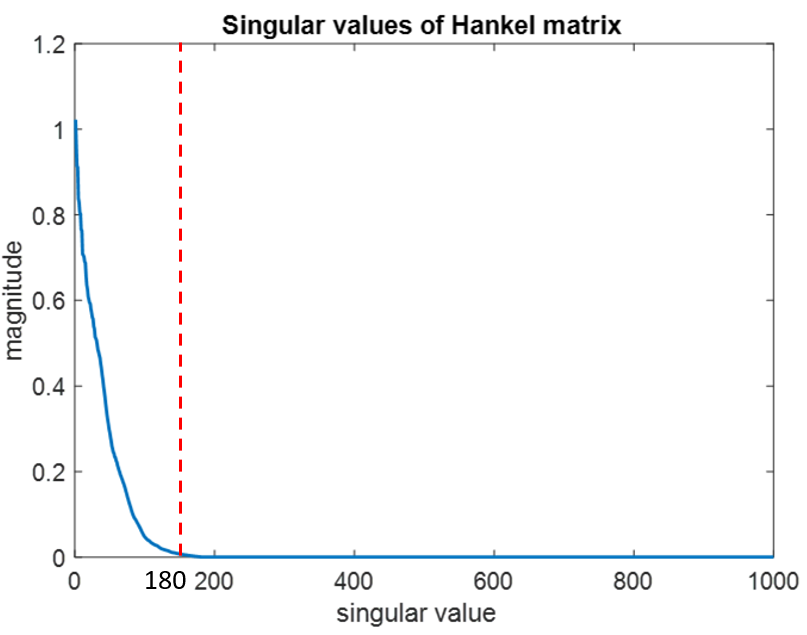


FIG. 3 Singular values of the Hankel matrix **H**1 plotted in descending order.

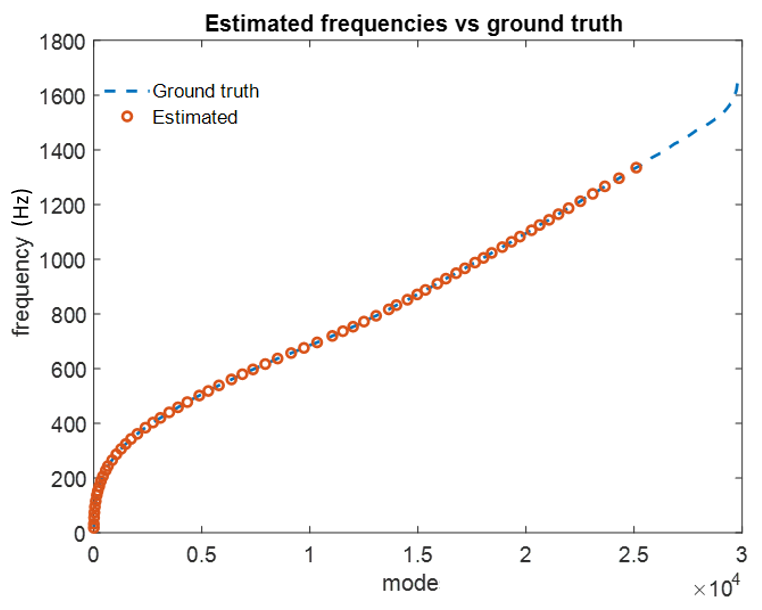


FIG. 4 The estimated resonance frequencies  versus the ground-truth frequency .

In order to quantify the similarity between the estimated mode shapes  and the oracle mode shapes **p**, the modal assurance criterion (MAC)16 is employed.

|  |  |
| --- | --- |
| . | (23) |

Note that . A large MAC value close to one indicates that two mode shapes are highly similar. MAC calculated for the frequency range, 10-1200 Hz, is illustrated in FIG. 5.

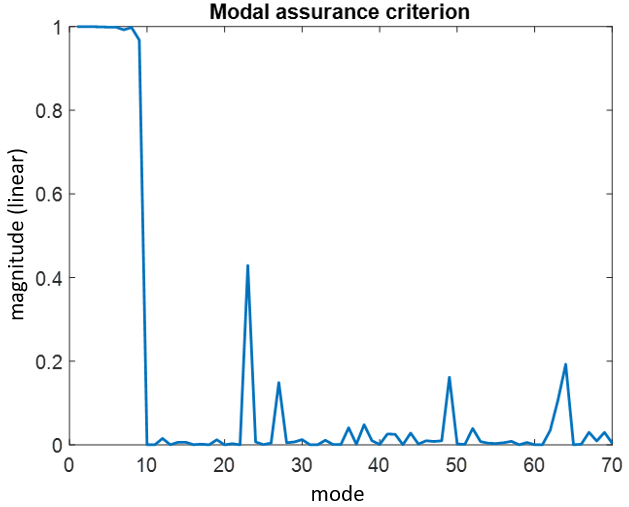


FIG. 5 MAC calculated for the frequency range, 10-1200 Hz (simulation).

It can be observed in FIG. 5 that MAC reaches nearly one for approximately the first ten modes and drops substantially beyond. The natural frequency of the tenth mode is 143 Hz which is slightly higher than the Schroeder frequency8 *fsch* = 75 Hz as predicted by the formula

|  |  |
| --- | --- |
|  | (24) |

with *V* being the volume of the room. Based on the results of the MAC values, the mode shapes at low frequency can be precisely estimated, while those at high frequency range are not guaranteed. Three samples of modal parameters and mode shapes on the plane *z* = 1.27 m at low frequency are summarized in TABLE. I and FIG. 6, respectively. Furthermore, the mode shape at 156 Hz is illustrated in FIG. 7 to addressed the limitation of the AMA approach.

TABLE. I. The estimated and the oracle modal parameters of three sample modes (simulation).

|  |  |  |  |
| --- | --- | --- | --- |
| Mode | / (Hz) |  | MAC |
| (1,0,0) | 13.5/13.7 | 0.01/0.01 | 0.98 |
| (1,1,0) | 24.6/25.8 | 0.013/0.01 | 0.99 |
| (2,2,0) | 52.9/51.6 | 0.01/0.01 | 0.99 |

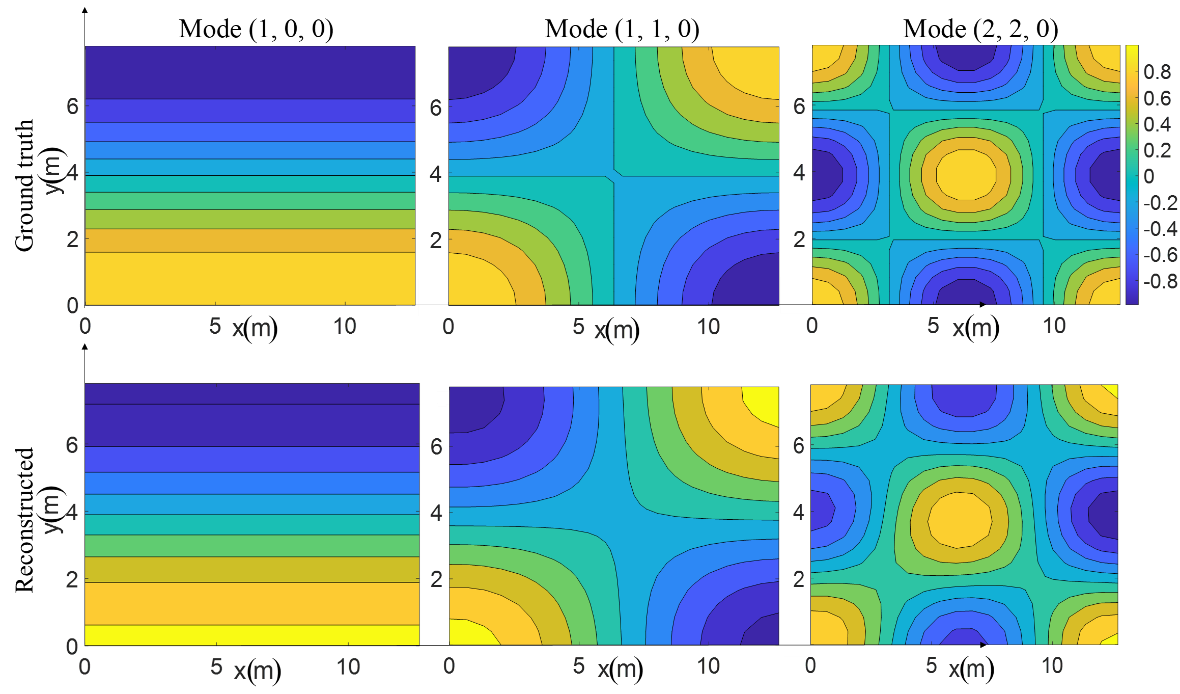


FIG. 6 Reconstructed mode shapes versus ground-truth mode shapes at z = 1.27 m. The height of the measured plane and the reconstructed plane are both at z = 1.27 m. (simulation).

TABLE. I suggests that the estimated resonance frequencies and damping ratios are in very good agreement with the oracle parameters. All MAC values are nearly one, indicating the high spatial correlation between the estimated and the ground-truth mode shapes of the room. FIG. 6 also shows the estimated and the oracle mode shapes are very similar except for slight skewness of nodal lines. In FIG. 7, we compare a simulated mode shape (2, 2, 3) at 156 Hz estimated by AMA with the ground truth. The estimated damping ratio is 0.004 which differs significantly from that of the ideal rectangular room (0.01). In addition, the MAC value calculated using Eq. 23 is only 0.06. In summary, the simulation results demonstrate the efficacy of the proposed AMA technique in extracting the modal parameters and mode shapes of a room at low frequencies. However, the proposed method is not applicable in the high-frequency range, which can be regarded as the limitation of the proposed method. It follows that we focus on the comparison of the low-frequency modes by experiments in the next section B.

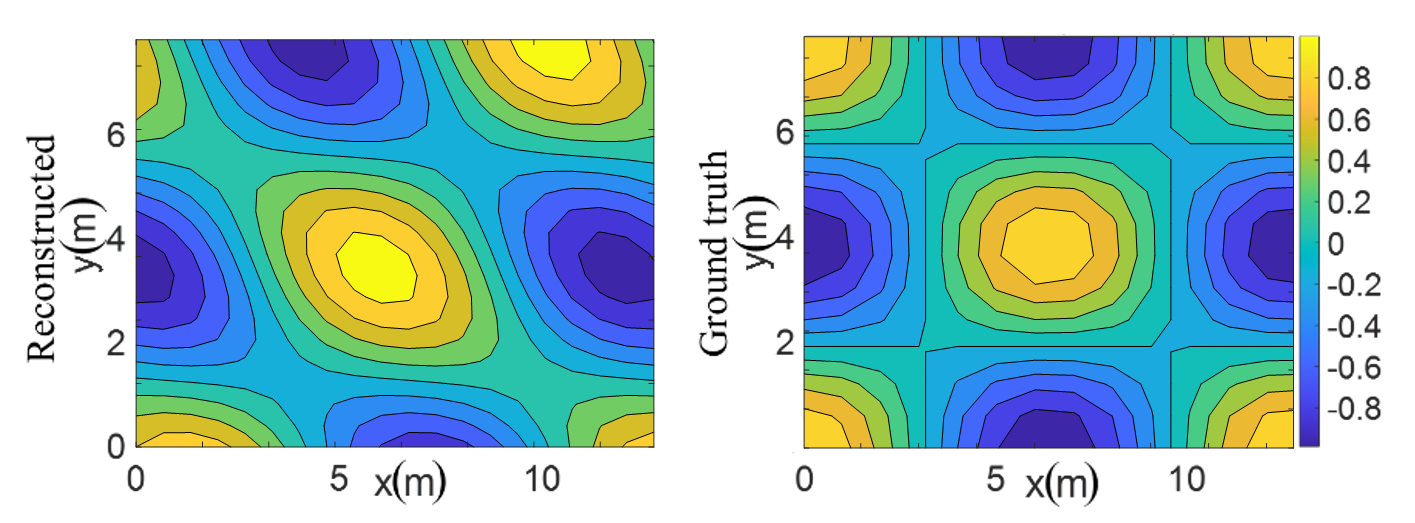


FIG. 7 Reconstructed mode shape (left) versus ground-truth mode shape (right) at z = 1.27 m associated with mode (2, 2, 3) at 156 Hz.

The following simulation was conducted to gain some insights of the proposed method when the RIRs at some 3-D locations on the horizontal planes other than the measurement plane are sought after. In FIG. 7, three mode shapes at 13.7, 25.8, and 51.6 Hz are reproduced for the plane, z = 1.27 m, when the microphones are placed on the same plane. In FIG. 8, the three mode shapes are reproduced using the proposed ERA-based technique for the plane at z = 2.27 m, based on the same measurement on the plane at z = 1.27 m. The results reveal that the proposed method is still capable of reproducing the mode shapes to some degree except the fact that the nodal lines are somewhat skewed for modes (1, 1, 0) and (2, 2, 0).

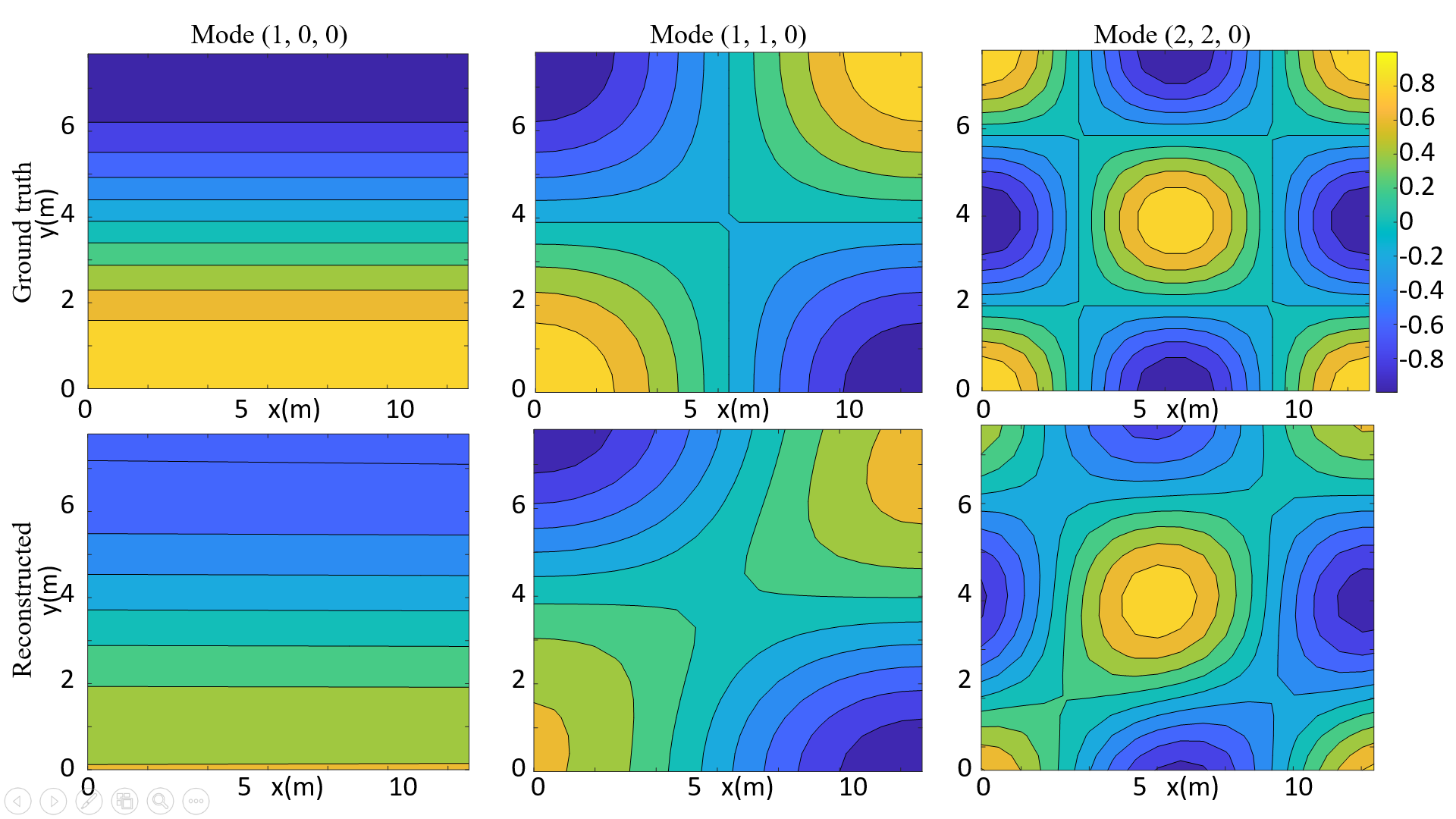


FIG. 8 Reconstructed versus ground-truth mode shapes at z = 2.27 m. The height of the measured plane is at z = 1.27 m, while that of the reconstructed plane is at z = 2.27 m (simulation).

1. **Experiments**

An experiment was undertaken in a large empty room with dimensions of , as shown in the experimental setup in a realistic room (FIG. 9). A Tannoy® full-range loudspeaker was placed in a corner of the room. The measured reverberation time of the room is 0.43 s. A six-microphone linear array with interelement spacing 0.8 m is moved in parallel 6 times with 0.8 m steps to measure the RFRs at 6 X 6 lattice points. PCB Piezotronics® 1/4” microphones are utilized for the RIR measurement. The source positions and the sensor positions are identical to those used in the simulations. White noise is selected as the source signal. Presonus Studio 1824c audio interface serves as the data acquisition system. RIRs are calculated by using the MATLAB® function,31 “tfestimate.”



FIG. 9 The experimental setup for AMA in a large empty room.

To confirm the applicability of the proposed ERA-based modal analysis method in real-world scenarios, experiments are conducted to compare the modal parameters estimated using the proposed method with an equivalent rectangular room with similar dimensions and reverberation time. Proceeding as the same way as in the simulation, singular values of the Hankel matrix are plotted in a descending order (FIG. 10), from which we chose to retain 150 modes for the following analysis. The estimated natural frequencies are compared with those predicted by an ideal rectangular room model in FIG. 11 Excellent fit can be seen throughout the frequency range, 10-800 Hz, which is obviously less than the effective frequency range (10-1200 Hz) covered in the simulation.

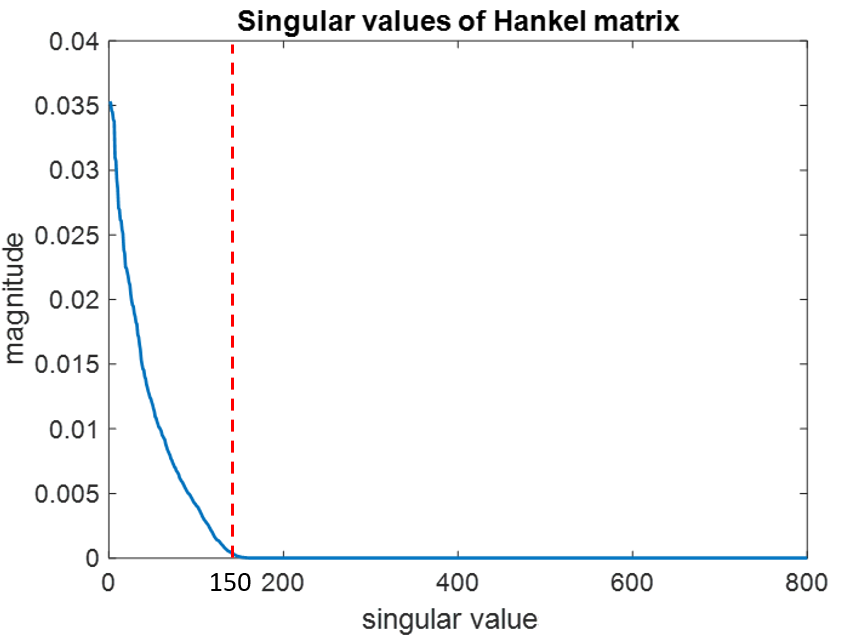


FIG. 10 Singular values of the Hankel matrix arranged in a descending order (experiment).

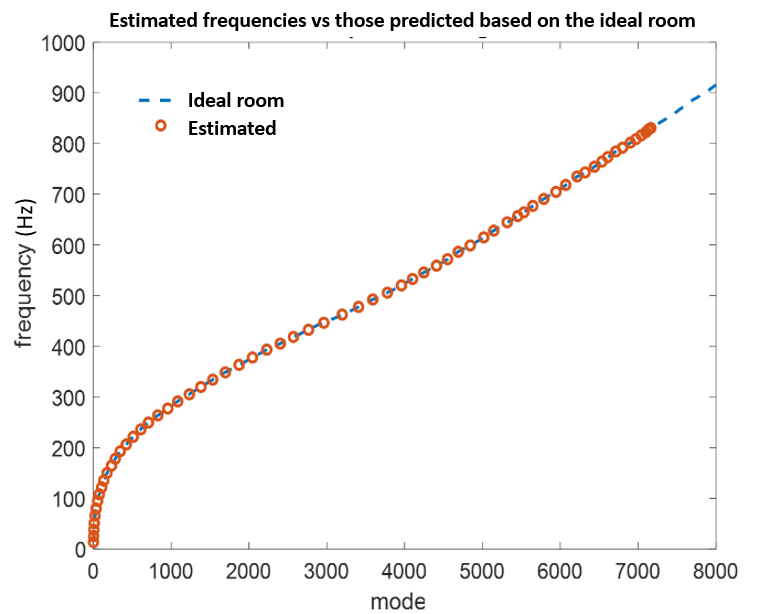


FIG. 11 The estimated frequencies  versus  predicted by the ideal rectangular room model.

FIG. 12 plots the MAC values below 800 Hz between the estimated mode shapes and those predicted by the ideal rectangular room model. It appears more difficult than in the previous simulation that only two modes up to the Schroeder frequency 75 Hz have the two largest MAC values.

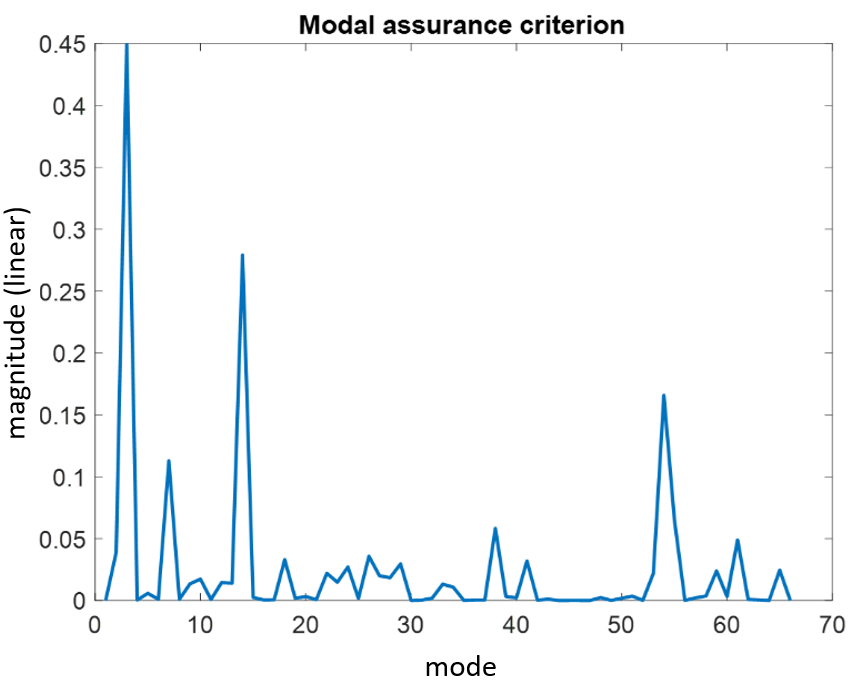


FIG. 12. The MAC values calculated in the range 10-800 Hz (experiment).

Proceeding with the same procedure as in the preceding simulations, the AMA results are summarized in TABLE. II (modal parameters) and FIG. 13 (mode shapes). It can be seen that the estimated natural frequencies are in excellent agreement with those predicted by the ideal rectangular room model. The estimated mode shapes remain in fairly close agreement with those predicted by the ideal rectangular room model despite minor skewness in the reconstructed nodal lines. As expected, the error of mode shape estimation is larger than that of natural frequency estimation32 due to the discrepancy of geometries and boundary conditions between the realistic room and the ideal rectangular room.

TABLE. II. The estimated modal parameters versus modal parameters predicted by the ideal rectangular room model for three low-frequency modes (experiment).

|  |  |  |  |
| --- | --- | --- | --- |
| Mode | / (Hz) | Damping ratio | MAC |
| (1,0,0) | 12.9/13.7 | 0.22/0.01 | 0.23 |
| (1,2,0) | 44.1/46.0 | 0.13/0.01 | 0.12 |
| (2,2,0) | 51.9/51.6 | 0.004/0.01 | 0.45 |

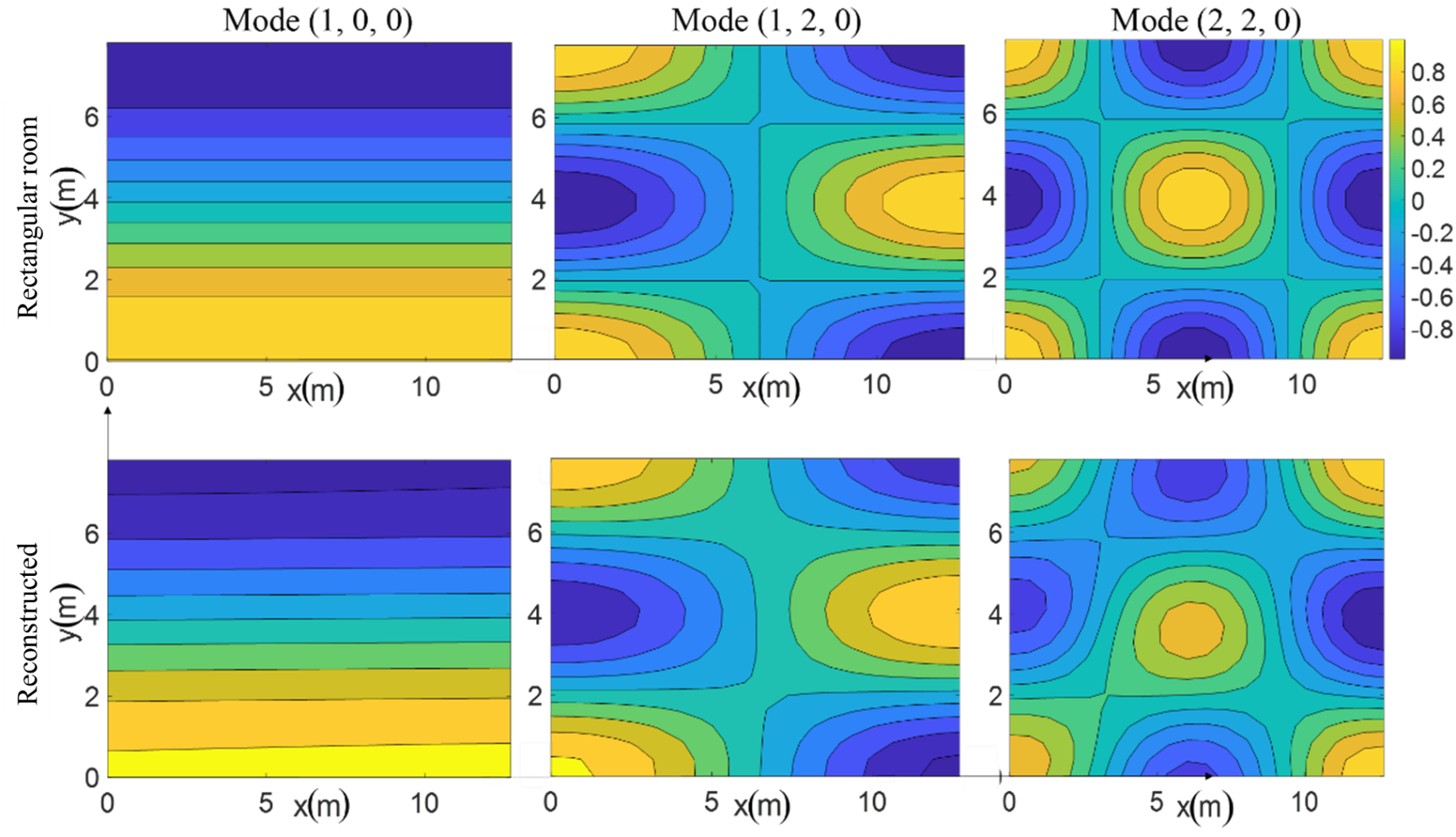


FIG. 13 The reconstructed mode shapes versus mode shapes predicted by the ideal rectangular room model (experiment).

# APPLICATION TO ACOUSTIC FIELD INTERPOLATION

The implication of Eq. (3) is profound in that the pressure response due to a unit point source can be evaluated for arbitrary source and receiver positions if modal parameters and mode-shape functions are available. In this section, application examples of the proposed AMA technique to acoustic field interpolation is given in terms of two important problems in audio signal processing: sensor interpolation and source interpolation. In practical implementation of an array problem, it is not always convenient or even possible to acquire the RIRs by exchanging the roles of the sources and receivers. To be specific, we need sensor interpolation for speaker array problems and source interpolation for microphone array problems. The former is a multiple-input-single-output (MISO) system, whereas the latter is a multiple-input-single-output (SIMO) system. In this section, we demonstrate that the proposed ERA-based approach can be applied to both of the interpolation problems.

1. **Sensor Interpolation**

Once the natural frequencies *ωn*, damping ratio , and mode shapes  are accurately estimated, RIRs can be reconstructed for the locations where no microphone measurements are made. Consider a room depicted in FIG. 2(a) for which we wish to interpolate RFRs on the plane located at *z* = 1.27 m, below the corresponding spatial aliasing frequency of the microphone array *fc* = 245 Hz. A loudspeaker source is placed at the corner (0.1m, 0.1m 0.1m) of the room. The proposed AMA procedure is utilized to estimate modal parameters and mode shapes of the enclosed sound field. Next, we apply Eq. (3) to regenerate the FRFs for locations where no sensor measurement is made. To quantify the interpolation error, we define

|  |  |
| --- | --- |
| , | (25) |

where  and  represent the interpolated and the ground truth RFR at a sensor location (*x*, *y*) in the room. The interpolation error over the plane at *z* = 1.27 m computed for four frequencies are shown in FIG. 14.

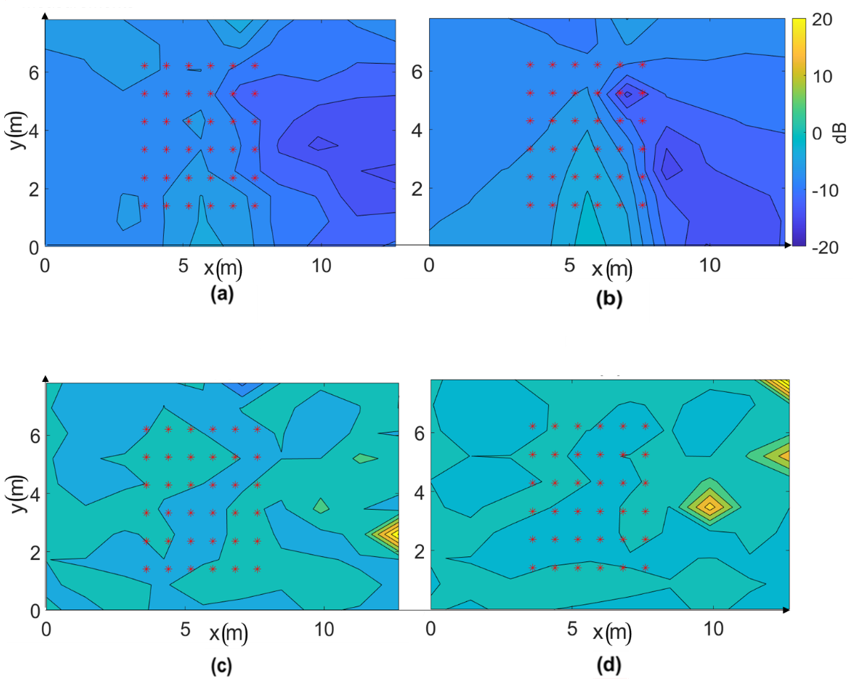


FIG. 14 The frequency response interpolation error (dB) for the sensor interpolation problem computed for four frequencies and *z* = 1.27 m plane in which (a) 104 Hz, (b) 140 Hz are below the spatial aliasing frequency (245 Hz) of the microphone array, and (c) 560 Hz, (d) 740 Hz are above the spatial aliasing frequency of the microphone array.

It can be observed from FIG. 14 (a) and (b) that the proposed sensor interpolation technique yields good agreement with the ground truth RFR (maximum error = -5 dB) below the spatial aliasing frequency. However, above the spatial aliasing frequency, the maximum error reaches 20 dB, as shown in FIG. 14 (c) and (d). This again corroborates the efficacy of the proposed interpolation technique in the low-frequency range where the modal behavior is prevailing and plane-wave decomposition and compressive sensing are useful.

1. **Source Interpolation**

Apart from the aforementioned sensor interpolation problem, we investigate the feasibility of applying the proposed AMA procedure to the source interpolation problem, where the RFRs are to be interpolated between a fixed sensor location and virtual source locations other than the physical ones. Consider a multiple-input-one-output scenario where sixteen-point sources are active simultaneously and one microphone are placed in the same rectangular room as in the sensor interpolation simulation. We wish to interpolate RFRs for source point at different azimuth angles (). FIG. 15 depicts the simulation settings of the source interpolation problem. Sixteen sources are placed uniformly on a circle of radius 1 m with 22.5-degree intervals in azimuthal angles. A microphone is placed at the center of the rectangular room. All sources and the microphone are positioned on the same plane located at *z* = 1.27 m.

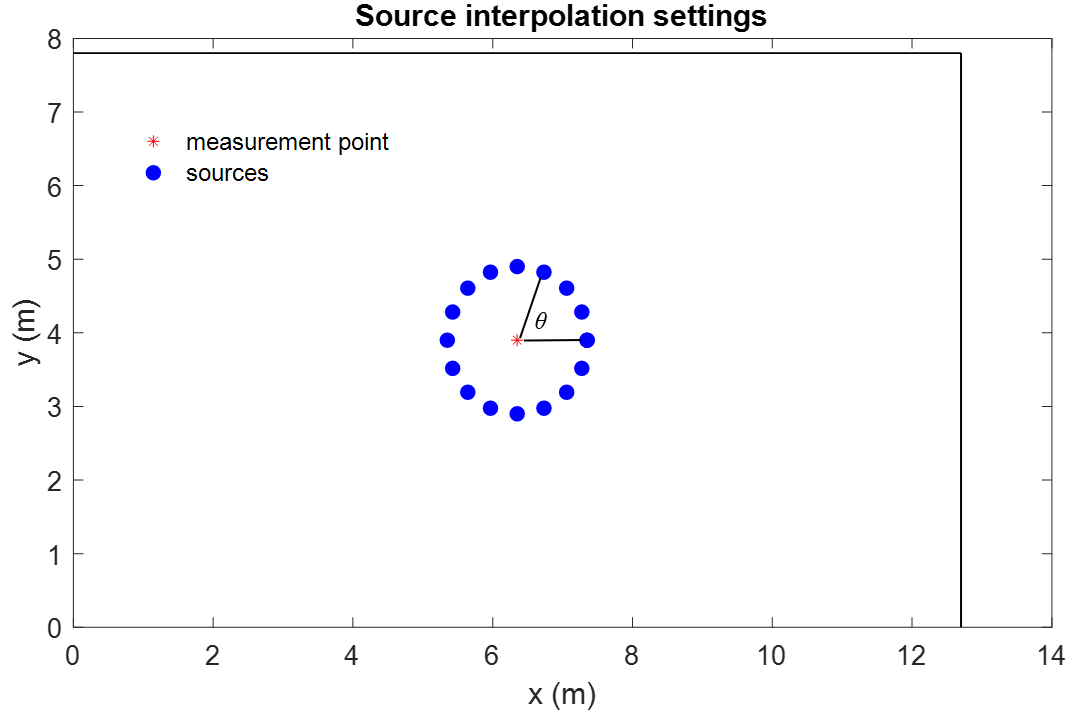


FIG. 15 The simulation settings for the source interpolation problem. Sixteen sources are positioned uniformly on a circle of 1 m radius. Sixteen sources in total are placed uniformly with 22.5-degree intervals in azimuthal angles. A microphone is placed at the center of the rectangular room with dimensions .

The interpolation error calculated for the RFRs reconstructed using the proposed AMA procedure is shown in FIG. 16. The spatial aliasing frequency of the UCA is 433 Hz. Although the Schroeder frequency is only 75 Hz in this room, the result reveals that the interpolation errors are very small, ranging from -20 dB to -5 dB, in the frequency range 10-500 Hz. Above 500 Hz, the matching errors reach nearly 10 dB due to spatial aliasing. Like sensor interpolation, source interpolation based on the modal model obtained using the proposed AMA technique work particularly well below the Schroeder frequency and the spatial aliasing frequency.

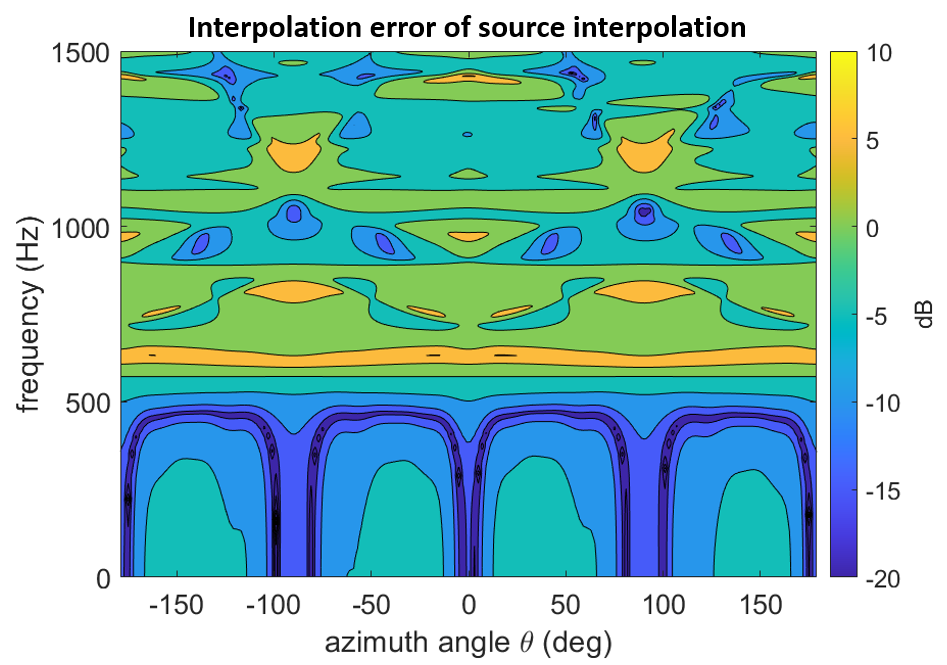


FIG. 16 The interpolation errors of source interpolation problem plotted versus frequency and azimuthal angles.

# CONCLUSIONS

In this contribution, an AMA procedure has been developed on the basis of ERA and CS. By using the proposed AMA technique, the modal parameters and mode shapes can be extracted successfully below the Schroeder frequency. The efficacy of the proposed AMA has been confirmed by simulations as well as experiments for a large room. The results have demonstrated that 36 measurement points suffice to reconstruct the mode shapes below the Schroeder frequency. Furthermore, an application example of sensor and source interpolation has been presented. In conclusion, the contribution of the present work is threefold. First, balanced realization in linear control theory has been applied to the modal analysis of acoustic fields. Second, a novel interpolation technique based on PWD and CS has been developed. Third, the proposed AMA procedure has been successfully applied to address an important issue in audio signal processing – field interpolation for arbitrary source-sensor locations.

**ACKNOWLEDGMENTS**

The work was supported by the Ministry of Science and Technology (MOST) in Taiwan, Republic of China, under the project number 107-2221-E-007-039-MY3.

# REFERENCES

1. C. H. Kung and R. Singh, “Experimental modal analysis technique for three dimensional acoustic cavities,” J. Acoust. Soc. Am. **77**(2), 731-738 (1985).
2. B. Xu and S. D. Sommerfeldt, “A hybrid modal analysis for enclosed sound fields,” J. Acoust. Soc. Am. **128**(5), 2857-2867 (2010).
3. D. Beaton and N. Xiang, “Room acoustic modal analysis using Bayesian inference,” J. Acoust. Soc. Am. **141**(6), 4480-4493 (2017).
4. F. Lluís, P. Martínez-Nuevo, M. B. Møller, and S. E. Shepstone, “Sound field reconstruction in rooms: inpainting meets super-resolution,” J. Acoust. Soc. Am. **148**(2), 649-659 (2020).
5. M. S. Kristoffersen, M. B. Møller, P. Martínez-Nuevo, and J. Østergaard, “Deep sound field reconstruction in real rooms: introducing the ISOBEL sound field dataset,” arXiv: 2102.06455 (2021).
6. E. H. Dowell, G. F. Gorman, and D. A. Smith, “Acoustoelasticity: general theory, acoustic natural modes and forced response to sinusoidal excitation, including comparisons with experiment,” J. Sound Vib. **52**, 519–542 (1977).
7. Y. Naka, A. A. Oberai and B. G. Shinn-Cunningham, “Acoustic eigenvalues of rectangular rooms with arbitrary wall impedances using the interval Newton / generalized bisection method,” J. Acoust. Soc. Am. **118**(6), 3662-3671 (2005).
8. M. R. Schroeder, “The Schroeder frequency revisited,” J. Acoust. Soc. Am. **99**(5), 3240-3241 (1996).
9. C. T. Chen, *Linear system theory and design*, Holt, Rhinehart, and Winston, New York, 243-246 (1984).
10. J. N. Juang and R. S. Pappa, “An eigensystem realization algorithm for modal parameter identification and model reduction,” J. Guid. Control Dyn. **8**(5), 620-627 (1985).
11. D. L. Donoho, “Compressed sensing,” IEEE Trans. Inf. Theory. **52**(4),1289-1306 (2006).
12. M. Elad, *Sparse and redundant representations: from theory to applications in signal and image processing.* (Springer, Haifa), 35-52 (2010).
13. C. X. Bi, Y. Liu, L. Xu, and Y. B. Zhang, “Sound field reconstruction using compressed modal equivalent point source method,” J. Acoust. Soc. Am. **141**(1), 73-79 (2017).
14. R. Mingot, L. Daudet, and F. Ollivier, ‘‘Room reverberation reconstruction: interpolation of the early part using compressed sensing,’’ IEEE Trans. Audio, Speech, Lang. Process. **21**(11), 2301-2312 (2013).
15. S. A. Verburg, and E. Fernandez-Grande, ‘‘Reconstruction of the sound field in a room using compressive sensing,’’ J. Acoust. Soc. Am. **143**(6), 3770-3779 (2018).
16. E. Fernandez-Grande, ‘‘Sound field reconstruction in a room from spatially distributed measurements,’’ in *2019 International Congress on Acoustics*, Aachen, 4961-4968 (2019).
17. E. Vincent, T. Virtanen, and S. Gannot, *Audio source separation and speech enhancement*, (Wiley, Hoboken, 2018), pp. 451-453.
18. B. Laufer-Goldshtein, R. Talmon, and S. Gannot, *A study on manifolds of acoustic responses*, (Springer, Switzerland, 2015), pp. 203-210.
19. M. Hahamann, S. A. V. Riezu, and E. Fernandez-Grande, “Analysis of a sound field in a room using dictionary learning,” in *2019 International Congress on Acoustics*, Aachen, Germany, 149-154 (2019).
20. T. Lin and H. Zha, “Riemannian manifold learning,” IEEE Trans. Pattern Anal. Mach. Intell. **30** (5), 796-809 (2008).
21. S. Gannot, D. Burshtein and E. Weinstein, “Signal enhancement using beamforming and nonstationarity with applications to speech,” IEEE Trans. Signal Process. **49** (8), 1614-1626 (2001).
22. C. Masterson, G. Kearney, and F. Boland, “Acoustic impulse response interpolation for multichannel systems using dynamic time warping,” in *Audio Engineering Society 35th International Conference*, London, (2009).
23. Q. Feng, F. Yang, and J. Yang, “Interpolation of the early part of the acoustic transfer functions using block sparse models”, J. Acoust. Soc. Am. **142** (6), EL532-EL536 (2017).
24. H. Gamper, “Head-related transfer function interpolation in azimuth, elevation, and distance,” J. Acoust. Soc. Am. **134**(6), 547-553 (2013).
25. P. A. Nelson and S. J. Elliott, *Active Control of Sound*, (Academic Press, London), 310-319 (1993).
26. C. L. Phillips and H. T. Nagel, Digital control system analysis and design - fourth edition, Prentice-Hall, Inc., 212-213 (1989).
27. M. Park and B. Rafaely, “Sound-field analysis by plane-wave decomposition using spherical microphone array,” J. Acoust. Soc. Am. **118** (5), 3094-3103 (2005).
28. M. Nolan, M. Berzborn, and E. Fernandez-Grande, “Isotropy in decaying reverberant sound fields,” J. Acoust. Soc. Am. 148(2), 1077-1088 (2020).
29. R. Tibshirani, “Regression shrinkage and selection via the Lasso,” J. R. Stat. Soc. **58** (4), 267-288 (1996).
30. M. Grant and S. Boyd, “CVX: matlab software for disciplined convex programming, version 3.0 beta,” http://cvxr.com/cvx (Last viewed May 1, 2021).
31. MathWorks, “Matlab and simulink for technical computing,” http:// www.mathworks.com/ (Last viewed May 1, 2021).
32. L. Meirovitch, *Analytical methods in vibrations*, New York McMillan, 73-272 (1967)