Research Article

Global Sound Field Reconstruction in the Room Environment Based on Inverse Wave-Based Simulation

Haitao Wang 1,2, Lin Zhang 1,2, Yakun Wang 1,2, Yaocheng Nie 1,2, and Xiangyang Zeng 1,2

1 School of Marine Science and Technology, Northwestern Polytechnical University, Xi’an 710072, China
2 Key Laboratory of Ocean Acoustics and Sensing (Northwestern Polytechnical University), Ministry of Industry and Information Technology of China, Beijing, China

Correspondence should be addressed to Haitao Wang; wht@nwpu.edu.cn

Received 15 July 2021; Revised 5 February 2022; Accepted 21 February 2022; Published 18 March 2022

1. Introduction

Sound field reconstruction is a key technique in many engineering applications, such as cabin sound source identification, spatial noise reduction, and acoustic imaging [1]. This topic has attracted sustained focus in the area of acoustics and signal processing, and new methods have been constantly proposed in order to obtain more precise and robust reconstruction.

A basic assumption in most classical sound field reconstruction methods, for example, the near-field acoustic holography (NAH) [2], is that the sound propagated in free space or the enclosed space is large enough to be considered as a free one. This is a key problem that most methods being faced in the room environment where there is the influence of reverberation on sound propagation caused by reflections of walls. In recent years, many efforts have been dedicated to indoor sound field reconstruction. A regular way to realize the reconstruction in the reverberant environment is improving the classical NAH. NAH is one of the most representative methods for sound field reconstruction. It recovers the source information by sampling the sound data on a hologram surface and then gives the prediction of the sound field on the predicting surface. Based on various supporting techniques such as the equivalent source [3–6], Helmholtz equation least squares [7–9], and inverse boundary element methods [10–12], NAH has become increasingly popular in various fields. To enhance its performance in the room environment, NAH based on finite element analysis was proposed in recent years [13]. This method effectively avoids the side influence of reverberation on sampling the pressure information. But it mainly focuses on the particle velocity reconstruction on the boundary but not the interior sound reconstruction. For the pass-by noise contribution analysis in a vehicle, a frequency-averaged $l_1$-norm regularization technique based on near-field sampling
was proposed [14]. Similar to other NAH-based methods, this method also focuses on the boundary vibration reconstruction but not the sound field. Overall, from the applicability perspective, NAH-based methods are more commonly used for reconstructing exterior sound field or interior vibration on the boundary.

The methods based on acoustic function expansion are other types of frequently used methods to realize indoor reconstruction. In these types of methods, the sound pressure in a room is represented by some expansion form and expressed by adding the products of a set of base functions and their corresponding coefficients. By solving the coefficients through inverse operations from sampled signals, the sound fields on other spatial positions can be calculated based on the expression of the sound pressure. The spatial modal expansion presented by Wu was early used for sound pressure reconstruction in a closed cavity [15]. By solving the Helmholtz equation by the extended Helmholtz equation least squares (HELS) method, the sound pressure can be correctly reconstructed. But because acoustic modes involved in the calculation were analytically obtained in regular-shaped rooms, the method was not suitable for the irregular-shaped room. Nevertheless, due to good applicability to reverberation, the expansion type of methods has attracted much attention in recent years [16–18]. According to different expansion forms, these methods can be further categorized into different types, such as the spherical wave decomposition methods [19–21] and plane wave decomposition methods [22–25]. However, although the expansion type of methods has achieved huge developments over the past decade, they still lack good performances on global reconstruction under reverberation. These methods are often able to provide correct reconstructions in regions near the microphone array. But their accuracies degrade a lot at a far region of the array. This is because the expansions of sound fields in these methods usually cannot precisely describe the room’s boundary effect. Under this condition, the expansion type of methods requires more sample points distributed throughout the room to produce a global reconstruction.

Under the classical method framework, supporting techniques also achieved significant developments. The most representative technique is the sparse method. No matter for the source recovery in NAH or solving the coefficients of sound pressure expression, the sparse method is a very effective way to reduce measurement burden and improve recovery robustness. It has been nearly considered as a basic framework for accomplishing the recovering calculation in the past decade. The development of sparse Bayesian learning in recent years has further strengthens on the wide application of the sparse method in this field [26–30].

In the aforementioned reconstruction methods, the room knowledge is rarely involved in the calculations. But in fact, the room itself plays an important role in forming the sound field. Its geometrical shape governs the propagation patterns of the sound waves, and the boundary impedance has an obvious influence on the attenuation of sound waves. Therefore, it is foreseeable that a good reconstruction performance can be obtained if the room information can be properly used as prior knowledge. This idea has been recently used in the inverse room acoustic problem. Image source method has been used to build the model in some source recovery methods. By simulating the channel responses between the sensors and the virtual image of source, the enclosure can be expanded into a large free space, which cancels the reverberant problem [31–33]. However, a basic rule in the image method is that the order of the image sources should be large enough to ensure that the enclosure can be considered as free space. Thus, the number of image sources will sharply increase as the number of walls increases. This would lead to that the computational accuracy and efficiency degrade a lot substantially. The wave-based theory, such as the finite element method and boundary element method, is another commonly used modeling method in acoustics. It divides the enclosure into smaller and simpler parts and yields responses through solving the system equation that derived from the Helmholtz equation. By constructing a localization method according to the inverse procedure of wave modeling, it is potentially beneficial for giving robust result in reverberant environment. There has been a study reporting that the method based on the inverse finite element analysis is effective for sound source localization in a strong reverberant environment [34]. However, the classical finite element analysis realizes the simulation through assembling local system matrices into system matrices. It leads to that the recovery parameter of the source has no sparsity, so a large number of samplings are needed to weaken the influence of the underdetermined problem.

Inspired by the model-driven thought, an inverse wave modeling-based method is proposed in this study to globally and economically reconstruct sound field in a complicated-shaped room environment. In this method, a wave-based simulation model is first constructed under the classical finite wave simulation framework [35]. A global and spatial sparse type of shape function is developed in this step that is used to construct the system’s matrices, which implicitly contain the room’s geometric information. Under the wave model’s constraint, the sound source is sparsely recovered and the wave model’s nodal pressure is obtained. The sound field over extended regions inside the room can be estimated. The experiments validated that this method can provide correct global sound field reconstructions in complicated-shaped rooms.

The study is organized as follows. In Section 2, a theoretical derivation of the proposed method is described. In Section 3, numerical simulations are conducted to preliminarily verify the proposed method’s accuracy, and the proposed method is compared with the reference method. The factors influencing the proposed method’s performance are also evaluated. A conclusion is provided in Section 4.

2. Inverse Wave-Based Reconstruction Method

2.1. Overview of the Method. Sound field reconstruction is a typical kind of inverse problem in acoustics. The sound source is usually unknown in this problem, so the traditional numerical simulation methods, such as the finite element method, ray-tracing method, and statistical energy analysis method, cannot be directly applied to predict the sound field. To realize the reconstruction, spatial sound pressure sampling should be first performed. Based on the spatial sound field
conversion algorithm, the sound source or the expansion coefficients are recovered, which is an inverse procedure compared to the simulation. Then, the spatial sound field can be predicted. This procedure can be generally expressed by

\[ y = \Phi s, \]  

where \( y \) is the sampling information, \( s \) is the source information or expansion coefficients to be determined, and \( \Phi \) is the sensing matrix. By solving \( s \), the reconstruction of the sound field on other spatial positions can be easily realized.

In this procedure, the sensing matrix \( \Phi \) is the key factor for the reconstruction. To be available for a strong reverberant environment, it is an effective way to involve the room information in constructing \( \Phi \). This is achieved by constructing a wave model in this study. The room environment is first divided into discrete subspaces as illustrated in Figure 1. By converting the Helmholtz equation into an integral equation spread over all subspaces, the influence of room on forming the sound field is exploited in the model and relations between any two subspaces are built. Thus, the transfer functions between sampling points to the source can be combined as the matrix \( \Phi \), and the recovery of source under a strong reverberant environment can be accurately achieved.

2.2. Sound Field Modeling in the Room Environment. A steady-state sound field simulation problem in a room is considered as shown in Figure 1. A closed boundary \( \Gamma \) surrounds a fluid domain \( \Omega \), which is characterized by its speed of sound \( c_0 \) and ambient fluid density \( \rho_0 \). The fluid domain is excited at a circular frequency \( \omega \) by an acoustic point source with a prescribed volume velocity \( q \) located at position \( s \). Assuming that the system is linear, the fluid is inviscid and the process is adiabatic. The steady-state acoustic pressure \( p \) in the problem domain is governed by the wave equation as follows:

\[ \nabla^2 p - \frac{1}{c_0^2} \frac{\partial^2 p}{\partial t^2} = -\rho_0 \frac{\partial q}{\partial t}, \]  

where \( \nabla^2 \) is the Laplacian operator.

To solve the sound field in the frequency domain, the Helmholtz equation can be derived from equation (2) as follows:

\[ \nabla^2 p_\omega + k^2 p_\omega + j\omega q_\omega = 0, \]  

where \( \omega \) is the circular frequency, \( j \) is the imaginary unit, and \( p_\omega \) and \( q_\omega \) are the sound pressure and the sound source intensity, respectively, in the frequency domain.

As the sound waves incident on the surface of the room will be partially absorbed, the boundary condition can be described by

\[ \frac{\partial p_\omega}{\partial n} = -j\rho_0 \omega \frac{p_\omega}{Z_s}, \]  

where \( Z_s \) is the specific acoustic impedance of the surface, and \( \partial p_\omega/\partial n \) expresses the pressure derivative along with the boundary’s normal direction \( n \).

Equation (3) can be analytically solved for a room with a regular boundary. While in an irregularly shaped room, this equation usually cannot be directly solved. Under this condition, numerical analysis methods such as the finite element method (FEM) and the element-free Galerkin method are usually employed to provide approximate solutions.

In these methods, to solve equation (3) under complicated boundary conditions, the space \( \Omega \) must be divided into discrete forms, for example, the elements or nodes shown in Figure 1. Based on these predefined elements and nodes, the sound pressure at any position, for example, \( r \), can be expressed by the sum of a set of linear functions:

\[ p_r = \sum_{i=1}^{N_r} N_r p_i^\Omega = N_r^T p_\omega^\Omega, \]  

where \( N_r \) is the total number of nodes distributed in the problem domain, \( p_\omega^\Omega = [p_\omega^1, p_\omega^2, \ldots, p_\omega^N] \) is the vector of nodal sound pressure on each predefined node, and \( N_r = [N_{r1}, N_{r2}, \ldots, N_{rn}] \) is the vector of an interpolating function that indicates how much influence each node has on position \( r \). Since \( N_r \) can be only constructed by using the spatial geometric information, it is usually called the shape function [36].

Finally, by expressing the global sound pressure using the predefined nodes, a system equation can be obtained as follows:

\[ (K - \omega^2 M + j\omega C)p_\omega^\Omega = F, \]  

where \( K, M, \) and \( C \) are system matrices, which describe the essential characteristics of the room on forming its interior sound field. They are only related to room parameters and are defined by the following:

\[ \begin{align*}
K &= \int_\Omega (\nabla N)(\nabla N)^T \, dv \in \mathbb{C}^{n \times n} \\
M &= \frac{1}{c_0^2} \int_\Omega NN^T \, dv \in \mathbb{C}^{n \times n} \\
C &= \frac{j\omega}{Z_s} \int_\Gamma NN^T \, ds \in \mathbb{C}^{n \times n} \\
F &= \int_\Omega Nq_\omega dv \in \mathbb{C}^{n \times 1}
\end{align*} \]
The derivation shows that the differential equation in equation (2) is converted into an integral equation, in which the relationship between the room and its interior sound field has been built. By solving equation (6), the sound field even in a highly reverberant environment can be globally calculated. The global sound field reconstruction method proposed in this study was inspired by this theory.

2.3. Global Sound Field Reconstruction. In equation (6), the governing pattern of the space \( \Omega \) on its interior sound field has been modeled. Based on this model, the transfer function between any two positions in the space can be calculated. It demonstrates that the global reconstruction of the sound field can be realized even if the samplings are totally obtained from the reverberant field under the constraint of the wave model. It constitutes the theoretical foundation of the reconstruction method in this study.

The first step in the sound field reconstruction is the recovery of the vector \( \mathbf{F} \). In the same space illustrated in Figure 1, it is assumed that a sound pressure sampling point with a known position \( r \) has been arranged inside the room and the sampled sound pressure is \( p_r \). Then, according to equation (5), the nodal sound pressure on predefined nodes can be expressed by the following form:

\[
\begin{align*}
P^N_r &= (N^T_r)^{-1} p_r. \\
\end{align*}
\]  

(8)

By substituting equation (8) into (6), the system equation can be converted into the following:

\[
\begin{align*}
P_r &= (N^T_r)(K - \omega^2 M + j\omega C)^{-1} F. \\
\end{align*}
\]  

(9)

This equation builds a connection between the sampled signal and the wave model. By extending the sound signal sample from a single position to a number \( l \) of known position \( r = [r_1, r_2, \ldots, r_l] \), the sampled sound pressure \( p_r = [p_{r_1}, p_{r_2}, \ldots, p_{r_l}]^T \in C^{l\times1} \) can be expressed by the following:

\[
\begin{align*}
p_r &= Dx, \\
\end{align*}
\]  

(10)

where \( D = -j\rho_0 \omega H_r (K + j\omega C - \omega^2 M)^{-1} \in C^{l\times1} \) is the sensing matrix, in which \( H_r = [N_{r_1}, N_{r_2}, \ldots, N_{r_l}]^T \in C^{l\times1} \) is the matrix consisting of the shape functions related to the sampling positions, and \( x = [N_{r_1}, \ldots, N_{r_l}] \in C^{l\times1} \) is the sound source parameter to be determined.

Solving equation (10) gives the recovered vector \( \mathbf{F} \) and by substituting \( \mathbf{F} \) into equation (6), the global sound field can be finally reconstructed.

2.4. Construction of Global and Sparse Types of Shape Functions. The expressions of the system matrices in equation (7) demonstrate that the shape function \( N \) is a basic parameter for constructing the system matrices. It describes the contributions of spatial nodes to a target point by interpolation. Also, it can be used to distribute the original source energy to spatial nodes. So, the recovery of the original source in equation (10) is equivalent to the recovery of the shape function of the sound source. Therefore, it is very important to construct a global type of shape function for an accurate sound field reconstruction.

The moving least squares (MLS) method [37,38] was used to construct the shape function in this study. In MLS, the shape function is constructed based on entire spatial nodes, which establishes the foundation of global reconstruction. In addition, by constraining the effective values using a compactly supported domain, the shape function has sparse properties, which allows that the compressive sensing theory can be used to achieve high recovering accuracy based on the limited numbers of sample points.

As shown in Figure 2, it is assumed that the space \( \Omega \) has been discretized and expressed by \( n \) nodes \( s^\Omega = [s_1 \Omega, s_2 \Omega, \ldots, s_n \Omega]^T \) and the exact nodal field values at all nodes are known as \( u^\Omega = [u_1 \Omega, u_2 \Omega, \ldots, u_n \Omega]^T \). We assume that a target position is located at a random position \( r \) in the space. Since the exact field value is usually difficult to be directly obtained, its approximating value is defined by

\[
u(r) = u^b(r, \tilde{r}) = \sum_{i=1}^{g} b_i(\tilde{r}) a_i(r) = b^T(\tilde{r}) a(r), \quad r \in \Omega, \tilde{r} \in \Omega,
\]  

(11)

where \( u^b(r, \tilde{r}) \) is the interpolant of \( u(r) \) that defined at a neighborhood position \( \tilde{r} \) of target position \( r \). \( b(r) \) is a set of complete basis at \( r \), \( a(r) \) is the vector of the unknown coefficients related to target position \( r \), and \( g \) is the number of base units.

Several types of functions can be used as the basis for MLS, such as the monomial, trigonometric, and wavelet functions. In this study, the monomial function, which is defined as follows in three-dimensional problems, is taken as the complete basis.

\[
\begin{align*}
b^T(\tilde{r}) &= \begin{bmatrix} 1, \tilde{x}, \tilde{y}, \tilde{z} \end{bmatrix} \quad (g = 4) \\
b^T(\tilde{r}) &= \begin{bmatrix} 1, \tilde{x}, \tilde{y}, \tilde{z}, \tilde{x}^2, \tilde{xy}, \tilde{y}^2, \tilde{yz}, \tilde{z}^2, \tilde{xz} \end{bmatrix} \quad (g = 10),
\end{align*}
\]  

(12)

where \( (\tilde{x}, \tilde{y}, \tilde{z}) \) is the space coordinate of position \( \tilde{r} \).

To obtain the unknown coefficients \( a(r) \), by considering the predefined nodes \( s^\Omega \) as neighborhoods of target position \( r \), a weighted discrete \( L_2 \) norm over \( n \) nodes is constructed as follows:

\[
\begin{align*}
J = \sum_{i=1}^{n} w(d_i)[u^b(r, s^\Omega_i) - u^\Omega_i]^2 = \sum_{i=1}^{n} w(d_i)[b^T(s^\Omega_i) a(r) - u^\Omega_i]^2,
\end{align*}
\]  

(13)

where \( n \) is the number of nodes, \( s^\Omega_i \) is the coordinate of the \( i \)th predefined node, \( u^\Omega_i \) is the nodal value on the \( i \)th discrete node, and \( w(d_i) \) is a weight function whose value is constrained by the distance \( d_i = |r - s^\Omega_i| \) between the \( i \)th node and target position.

To construct a sparse shape function, a compactly supported technique is used to constrain the value of \( w \) in equation (9) as shown in Figure 2. The weight function is only available in a compactly supported domain, which is a sphere in this study. Based on this constraint, only a small
number of nodes really have influences on the target position. A shape function with sufficient sparsity can be constructed based on the compactly supported technique. In this study, the quartic function was chosen as the weight function:

\[
w(d) = \begin{cases} 
1 - 6d^2 + 8d^3 - 3d^4 & d \leq d_r, \\
0 & d > d_r,
\end{cases}
\] (14)

where \(d_r\) denotes the radius of the compactly supported domain.

Minimizing \(J\) with respect to \(a(r)\) provides

\[
\frac{\partial J}{\partial a(r)} = 2\sum_{i=1}^{n} w(d_i) b_i(s_i^T) [b^T(s_i^T) a(r) - u_i] = 0, j = 1, 2, \ldots, g
\] (15)

Then, the vector \(a(r)\) can be obtained by solving equation (15)

\[
a(r) = A^{-1} B u^\Omega,
\] (16)

where \(A\) and \(B\) are matrices defined as follows:

\[
A = \sum_{i=1}^{n} w(d_i) b_i(s_i^T) b^T(s_i^T) \in C^{g \times g}
\]

\[
B = \begin{bmatrix} w(d_1) b(s_1^T), w(d_2) b(s_2^T), \ldots, w(d_n) b(s_n^T) \end{bmatrix} \in C^{g \times n},
\] (17)

Finally, substituting equation (16) into (11) and setting \(\tilde{r} = r\), the field value can be further expressed as follows [39]:

\[
u(r) \approx u_h^T (\tilde{r}, \tilde{r}) = u_h^T (r) = b^T (r) A^{-1} B u^\Omega = N_r^T u^\Omega,
\] (18)

where \(N_r\) is the shape function constructed on position \(r\). The derivation demonstrates that the field value at any position in the space can be expressed by a set of predefined nodes and their nodal values. It can be considered as a governing factor, which determines how much influence each predefined node has on the target position.

2.5. Compressive Sensing with \(l_1\)-Norm Optimization. To ensure the sound source data can be correctly and economically recovered when solving equation (9), the compressive sensing (CS) theory that has been widely accepted as an efficient recovery tool in acoustic problems is used. CS theory suggests that if a signal is sparse and the measurement matrix is highly incoherent with the dictionary, it can be reconstructed using a limited number of measurements by solving an underdetermined inverse problem.

In equation (10), \(x = N_q\) describes the distribution of the discretized excitation. Based on the compactly supported domain constraint in MLS for constructing the shape function, \(x\) has very few nonzero elements. Therefore, \(x\) can be considered as a sparse vector, which is a benefit to the use of the CS theory to effectively reduce the sampling requirement.

If measurement noise is present, the inverse problem in equation (10) can be solved by solving a LASSO problem as follows [40]:

\[
\arg\min_{x \in \mathbb{C}^{n}} \|x\|_1 \text{subject to } \|p_r - D x\|_2 \leq \xi,
\] (19)

where the operator \(\| \cdot \|_1\) indicates the \(l_1\) norm. The parameter \(\xi\) is an estimation of the upper boundary of the noise present during the sensing process.

3. Numerical Verification

The accuracy of the proposed method was preliminarily explored by reconstructing the sound field in an enclosed room. The sizes of the room and coordinate system are shown in Figure 3. The specific acoustic impedance of the room’s inner surfaces is set to be \(10 \rho_0 c_0 + 10 \rho_0 c_0\) kg/m²s, where \(\rho_0 = 1.21\) kg/m³ and \(c_0 = 340\) m/s. To construct the wave model, the room is divided into \(6 \times 6 \times 6 = 216\) nodes as shown in Figure 3.

In this verification, the sound source is a point source located at \([2.55, 0.6, 0.35]\) m and a virtual array consisting of 28 microphones is used to sample the sound pressure signal. This planar array is placed at the plane of \(x = 3\) m as shown in Figure 3.

In this study, the sampled signals and sound fields used as real data are simulated using LMS Virtual Lab. Reconstructed sound fields on surfaces of \(x = 0\) m, \(y = 0\) m, and \(z = 1.8\) m at 50 Hz, 100 Hz, and 150 Hz are illustrated in Figure 4. To evaluate the performance of the proposed method, two reference methods were also used to calculate the same problem. The reference method 1 is the spherical wave model (SWM) [19], which realizes the reconstruction by spherical wave expansion. The reference method 2 is a method developed from an indoor localization technique [31]. Sensor arrays used in the reference methods are the same as that in the proposed method.

The results in Figure 4 demonstrate that the distributions and amplitudes of the sound fields between the proposed method and the real data are very close. At 50 Hz and 100 Hz, the sound field reconstructions are nearly identical to the real sound fields. This demonstrates that the proposed method can provide accurate reconstructions under strong reverberations at low frequencies. At 150 Hz, although the sound field is more spatially complicated than those at low frequencies, the proposed method can still give a similar distribution pattern to the real data. But there are also more differences between the proposed method and the real value. This indicates that the proposed method’s accuracy tends to decrease as the frequency increases.
In this verification, the reference method 1 fails to give satisfying results. SWM is a method that has been proved valid for sound field reconstruction in free space and cylindrical cavities. It usually has good performances near the sampling points. But in this verification, the global sound field is mainly formed by multipropagation of sound waves. It brings high difficulty for the SWM to reconstruct the global field.

The reference method 2 has much better reconstruction results. It gives similar distribution patterns at these three example frequencies. But its accuracy also tends to degrade at high frequencies. This reference method is developed from an indoor sound source localization method. Like the proposed method, it achieves the sound field reconstruction by first recovering the sound source and then predicts the sound field at other spatial positions. In the sound source...
recovering step, the method builds an acoustic model based on the image source theory and then localizes the source by sparse recovery. The comparisons demonstrate that building the wave model is a key and effective strategy to realize global reconstruction.

To quantitatively evaluate the proposed method’s accuracy, a root mean square error (RMSE) is defined as follows:

\[
\varepsilon = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left| p_i^{\text{reconstruct}} - p_i^{\text{real}} \right|^2} \times 100\%,
\]

where \( p_i^{\text{reconstruct}} \) and \( p_i^{\text{real}} \) denote the reconstructed and real sound pressure at the same evaluation position \( i \), and \( n \) is the number of positions included in the error evaluation. The RMSEs of the proposed method and reference method 2 from 30 Hz to 200 Hz are calculated on the 216 predefined nodes and are illustrated in Figure 5. Since the results of reference method 1 are obviously different from the real data, its RMSE is not evaluated here.

Figure 5 shows that the RMSEs at most frequencies are less than 20%. There are only a few frequencies at which the errors are higher than 20% but still less than 25%. The reference method has similar low errors at low frequencies, but its errors at mid- and high frequencies are significantly higher than the proposed method. It is because that the image source theory used in the reference method requires to extend the enclosure into a large free space, which leads to that the source vector to be recovered is much larger than that in the proposed method. Under the condition of the same number of sensors, the proposed method has better recovery performance than the reference method.

Moreover, from the reconstruction of the research perspective, the sound field reconstruction focuses more on low-frequency ranges than high-frequency ranges. This is because that the number of acoustic modes becomes quite dense in high-frequency ranges, making the sound field distribution over the space being too complicated to reconstruct. For the space in this case, the number of acoustic modes at each integer frequency interval starts to be quite large from 200 Hz. This means that 200 Hz can be considered a relatively high frequency. Under these conditions, the sound field usually needs to be statistically reconstructed rather than reconstructing the exact frequency response.

4. Numerical Analysis of Performance-Influencing Factors

The factors influencing the method’s performance are further numerically analyzed. All analyses in this section are conducted based on the cubic room shown in Figure 3.

4.1. Signal Sampling. Signal sampling is one of the key steps in reconstructing sound fields. It provides the input data for the reconstruction and directly determines the reconstruction quality.

The evaluations on the number and position of sensors were performed, and the RMSEs are shown in Figure 6. The first test is performed to evaluate the influence of the number of sensors on the performance. In this test, three numbers of sensors were tested. In this test, the sensors in different conditions were placed in the same area as shown in Figure 6(a) and the results for the test are shown in Figure 6(c). In the second test, three sensor arrays with the same number of sensors but different sizes were used to evaluate the influence of the sampling area on the performance. The three arrays are shown in Figure 6(b), and the result for this test is shown in Figure 6(d). In these tests, the sensors were all placed in the plane of \( x = 3 \text{ m} \).
Figure 6(c) shows that the RMSEs under different numbers of sensors are very close, while Figure 6(d) shows that the RMSEs under different sampling areas are very different. It demonstrates that the size of the sampling area is an important factor influencing the performance of the proposed method. A large sampling area is beneficial for obtaining precise reconstruction. The number of sensors is also important for achieving good reconstruction. But based on the sparse recovery, even a small number of sensors can give satisfied accuracy if they can spread over a large area. Figure 6(d) shows that the array 3 leads to a very large error at 140 Hz, so the sound field distribution at 140 Hz is illustrated in Figures 6(a) and 6(b) as backgrounds to express the sensors’ positions. The sound field varies a lot at different spatial positions. A larger array can ensure that the acquired sound pressure sampled by sensors has sufficient difference.

4.2. Signal-to-Noise Ratio (SNR). The performance of the proposed method under different SNRs is evaluated. Gaussian white noise with SNRs of 30 dB, 40 dB, and 50 dB is added to the pressure signals on each sensor to simulate the actual measurement. The calculations are conducted based on the same sound source and microphone array configuration described in Section 3. The RMSEs in this case are illustrated in Figure 7.

Figure 7 shows that the errors for SNR = 50 dB and 40 dB are very low at all frequencies. When the SNR is 30 dB, the errors remain low levels at frequencies below 100 Hz. This demonstrates that the proposed method can stably work even when the signals have low SNRs at low frequencies. However, its performance at high frequencies is much more sensitive to the low SNR. The RMSEs at two narrow bands exceed 40%. But at most frequencies, the proposed method still give results with errors of less than 20%.

4.3. Multiple Source Excitations. In the real application, the sound source is often not a monopole. To evaluate the performance of the proposed method on multiple source excitations, a test under three source excitations was performed. Three monopoles with coordinates of [1.4, 0.6, 0] m, [1.4, 0.6, 0.36] m, and [1.4, 0.6, 0.72] m were used as the sound sources. The calculations are conducted based on the
same microphone array configuration described in Section 3. Comparisons of the sound fields at exemplary frequencies are illustrated in Figure 8 and the RMSEs are shown in Figure 9.

Figure 8 shows that under the excitation of multiple monopoles, the proposed method provides very similar reconstructed results with real sound fields at 50 Hz and 100 Hz. At 150 Hz, the pattern of the sound field distribution obtained by the proposed method is still generally similar to the real data, but the difference is also larger than that in low frequencies. The RMSE shown in Figure 9 demonstrates that the proposed method under the excitation of three monopoles has larger errors than those under single monopole excitation as shown in Figure 5. Generally, the analysis demonstrates that the proposed method can be applied for sound field reconstructions at low frequencies even under complicated excitations, while the accuracy tends to decrease as the excitation complexity increases. The precision reduction under distributed excitation in the proposed method is mainly caused by the sound source recovery in the reconstruction. As indicated in equation (9), the proposed method has a key step in reconstructing the sound field that the sound source needs to be recovered first. In the theory of the proposed method, the sound source is assumed to be point sources. Due to this assumption, the proposed method is capable of processing the reconstruction problems with point sources. However, its accuracy degrades for problems with plane vibration or other distributing excitations.

To improve the performance of the proposed method for such problems, the theory of equivalent source is a potential...
direction. This theory is frequently used in NAH for solving reconstruction problems with complex sources [3–6]. It employs a series of virtual point sources to replace the source with a complex boundary shape to express the radiated sound field. Based on this equivalent processing, the continuous and complex source can be converted into discrete and simple point sources, which are exactly the recovered source form in the proposed method. The main difference between the reconstruction problems in this study and classical ESM-based NAH is that the target reconstruction zone in this study is the interior space of the enclosed environment, while in NAH the exterior space is often the target zone. Due to this reason, the classical theory of equivalent source cannot be directly used in the proposed method, but it still gives important inspiration for improving the performance on problems with complicated excitations. In future work, the equivalent source method will be considered as an important strategy to support better sound field reconstruction under complex excitations.

5. Conclusions

An inverse wave modeling-based method for globally reconstructing sound field in room environment was proposed. By building a wave-based model, the complicated propagating paths from the source to fields can be modeled, and then, the sound field at any position in a room can be reconstructed under strong reverberation by inversely solving the wave-based model. Compared with the traditional method, the proposed method needs geometry information as prior knowledge to build the wave model. This step leads to that these methods are suitable to be used in a constant environment. If the problem needs to be solved in another room or the current environment has significantly changed, then a new wave model should be built. The numerical verification has proved that the proposed method is capable of reconstructing sound fields under strong reverberation environments. Further evaluations of influencing factors demonstrated that the proposed method performs better in low-frequency ranges. In addition, this method does not need to calculate the arrival time difference or phase difference among different microphones. So, there is no need to design an array with a complicated shape in this method, which improves the flexibility to set the microphone positions in the space. But the verification also showed that a sampling in which the signals have large differences is still a benefit for better reconstruction. The proposed method is sensitive to SNR in high-frequency ranges, while it is robust to low SNR in low-frequency ranges. It was also validated that the proposed method provides correct results for multiple source excitations.

Data Availability

The data that support the findings of this study are available from the first author by e-mail (wht@nwpu.edu.cn), upon reasonable request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (grant no. 12074317) and the Aeronautical Science Foundation of China (grant no. 20181553017).

References


