**Multichannel Room Response Equalization with a Broadened Control Region Using a Linearly Constrained Approach and Sensor Interpolation**

Wei-Ling Lin, You-Siang Chen, Bo-Ru Lai, and Mingsian R. Baia)

1. *Department of Power Mechanical Engineering, National Tsing Hua University, No.* 101*, Section* 2*, Kuang-Fu Road, Hsinchu, Taiwan* 30013

*Suggested running title: LCMV speaker array with sensor interpolation*

1. Corresponding author: Mingsian R. Bai, Email: [msbai63@gmail.com](mailto:msbai63@gmail.com)

**ABSTRACT**

This paper describes a room response equalization technique based on an underdetermined multi-channel inverse filtering (UMIF) and linearly constrained minimum variance (LCMV) approach. Not limited to the local control at the neighborhood of the measured control points, the proposed UMIF-LCMV system is capable of widening the effective equalization area of the reproduced sound field, with a large number of interpolated control points. Specifically, a constrained optimization problem is formulated to minimize the matching error at the interpolated control points while seeking precise matching at the measured control points. In practical implementation, only the frequency responses associated with a limited number of control points need to be measured, whereas the frequency responses for the interpolated points are established by using a plane wave decomposition-based sensor interpolation technique. A two-stage procedure is developed to trim down plane-wave components by using the least absolute shrinkage and selection operator (LASSO) algorithm and to obtain the complex amplitudes of the principal components. Simulations, objective and subjective experiments are conducted for a system comprising a linear loudspeaker array and a linear microphone array. The results have confirmed the efficacy of the proposed system in widening the effective listening area with only limited discrete measurements.

**PACS number**: 43.60.Fg, 43.60.Pt

**I. INTRODUCTION**

Audio reproduction in an enclosed space such as a living room has been an important research issue. In practice, undesired effects of spectral coloration and limited sweet spot may arise due to room coupling with the loudspeakers. To mitigate the coupling effect while achieving broadened room response equalization, many methods has been suggested via the multichannel loudspeaker array system. By the term ‘broadened”, we refer it as reaching effective equalization performance within a connected control region, which is the main issue to be dealt with in this paper. Modal equalization methods1-3 aimed at equalizing modal resonances was implemented by relocating the system poles to the origin in the z-space. However, the modal equalization methods are effective for low-frequency and small-room reproduction only. Recently, a method termed the polynomial based filter design4,5 has been proposed for single-input multiple-output (SIMO) and multiple-input multiple-output (MIMO) systems. The robustness of this method is increased by modeling stochastic uncertainty in the measured transfer function, which can effectively prevent the equalization from over-fitting and accomplish the so-called spatial-average control. A room response equalization approach based on the multiple-input/output inverse theorem (MINT)6 was reported in the context of multichannel inverse filtering of a finite-length impulse response system, which leads to a perfect inverse if the acoustic system does not have common zeros and the number of channels and filter length meet some conditions. However, MINT may lead to high-gain filters because the filter coefficients are not properly regularized. To leverage the benefits of the regularization, Bai et al. proposed an improved system, the time-domain underdetermined multi-channel inverse filtering (TUMIF) approach7, in light of Tikhonov regularization (TIKR) method8,9. Alternatively, the TUMIF approach can be reformulated into the frequency-domain underdetermined multi-channel inverse filtering (FUMIF) approach.

The aforementioned MINT, TUMIF, and FUMIF can only achieve perfect reproduction at the measured control points, but not the interpolated points. However, if a broadened sweet spot is pursued, a large number of premeasured room impulse responses (RIRs) within the region of interest is required. However, the time-consuming measurement makes these approaches impractical for the real-world implementation. Many approaches have been proposed to acquire immense number of the RIRs at interpolated control points. Recently, efficient methods for obtaining frequency responses (FRs) are proposed by considering the wave propagation model10-12. The basis approaches such as plane wave decomposition (PWD)13-15 or equivalent simple sources16 can be also employed to acquire the lower-dimensional manifold of a FR. A sparse representation is a common assumption that can be solve by utilizing compressed sensing (CS) techniques17, 18. We refer this system-based acoustic transfer function (ATF) interpolation as sensor interpolation, where an ATF is represented by a wave propagation model through interpolating the measured ATFs at arbitrary locations. Therefore, the sensor interpolation can be carried out in light of the sound field model such that the number of the microphone measurement is significantly reduced. However, literatures reported that a suitable set of basis functions is crucial for reproducing accurate FRs19, 20.

In this paper, a two-staged sensor interpolation approach is proposed to approximate the RIRs for the interpolated points in light of plane wave decomposition (PWD). This approach can achieve considerable interpolation performance with only a limited number of measurements. The number of main plane-wave components are first trimmed down by the least absolute shrinkage and selection operator (LASSO) algorithm21, followed by a TIKR approach which is intended to estimate the complex amplitudes of the remaining plane-wave components. Unlike the previous approaches that require prior knowledge, such as the size of the measurement domain or the modal information of a room17, to carefully determine the number of basis functions, our interpolation method can quickly acquire a proper set of these functions in the first stage, and further obtain the precise amplitudes of the plane wave components in the second stage. Moreover, by taking the FRs at the interpolated points into consideration, we propose a rendering approach to achieve a broadened control region by reformulating the preceding FUMIF approach from the perspective of linearly constrained minimum variance (LCMV)22, 23 beamforming (referred to as FUMIF-LCMV hereafter). The control points consist of a limited number of “measured” points and a much larger number of “interpolated” points distributed in the listening area. To do so, we formulate a constrained optimization problem in which the constraint equation is due to model matching at the measured points, while the cost function is due to the minimization of the residual noise at the interpolated points. Consequently, a closed form solution of the broadened room equalization can be attained by FUMIF-LCMV incorporated with the two-staged PWD method under an underdetermined system. Two regularization parameters selected with the aid of particle swarm optimization (PSO)24,25 are required to properly limit the filter gain and preserve the equalization performance. Finally, the overall procedure of the multichannel room equalization method is illustrated in FIG. 2. The method allows the system to have broadened equalization performance with limited measurement of RIRs in a control region. Audio quality evaluation (PEAQ)26 is adopted as the performance measure in the objective evaluation, while a listening test serves as the subjective experiment to validate the proposed approach.

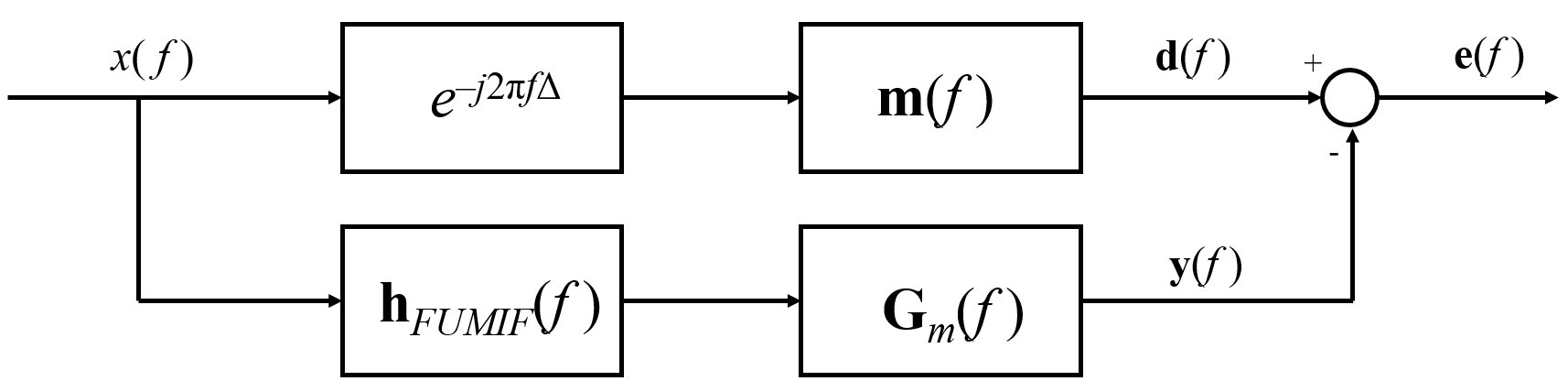


FIG. 1. The model matching problem of multi-channel room response equalization.

**II. THEORETICAL BACKGROUND**

1. **The FUMIF approach**

A multi-channel room response equalization problem can be viewed as a model matching problem, as depicted as FIG. 1, where mono input signal *x*(*f*) at the frequency *f* is filtered by the multiple FUMIF prefilters denoting **h***FUMIF*(*f*). Each filtered signal is reproducedby the measured FR functions of the loudspeaker driver **G***m*(*f*) between the loudspeakers and the measured control points. The goal of the problem is to minimize the matching error between the output signals **y**(*f*) and the desired signals **d**(*f*) which are generated by the predefined target model of the measured points **m**(*f*). The modelling delay *e-j*2π*f*△ with △ second is added to guarantee a causal system. The frequency index *f* is omitted hereafter for simplicity. To ensure causality of FUMIF system, modeling delay is incorporated as **m***m* = *e-j*2π*f*△**m**. Therefore, the model-matching problem can be expressed as

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

Consider *Ns* loudspeakers and *Nm* measured control points. Equation (1) can be rewritten explicitly with components as

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

where *Gij*, *i =* 1*,…, Nm*, *j =* 1*,…, Ns*, denotes the FR function between the *j*th loudspeaker and the *i*th measured control point, *hj* is the *j*th prefilter, and *mi* denotes the target model for the *i*th measured control point. The target model *mi*(*f*) is designed with the form of bandlimited equalization from a lower frequency limit *fl* to an upper frequency limit *fh*,as given by

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

where |∙| denotes the absolute operation. Thus, the target model within the bandlimited interval has flat response corresponding to the root-mean-square power of each FR from the source to the *i*th control point.

In the FUMIF approach, to guarantee the existence of solutions with perfect matching error, the rank of the frequency responses matrix **G***m* should be at least *Nm*. Hence, an underdetermined system is intentionally formulated by choosing *Ns* > *Nm*. To alleviate the ill-posed problem in the matrix inversion, the solution is further regularized with a diagonal term. By this setting, the solution of Eq. (1) can be obtained by using TIKR9:

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

where “*H*” denotes matrix conjugate transpose and *β* is the regularization parameter which can be selected using the L-curve method27-28. In the L-curve method, the regularization parameter, *β*, is be decided by taking the solution norm, ||**h***FUMIF*||2, and the residual norm, ||**m***m*–**G***m***h***FUMIF*||2, into consideration. Thus, the performance of the room equalization and the feasibility of implementation are simultaneously considered. The similar idea is applied to our proposed method for finding a suitable set of regularization parameters, which will be further described in Sec II B.

1. **The FUMIF-LCMV approach**

Despite the nearly perfect equalization achievable by the aforementioned FUMIF approach at the measured control points, the performance remains limited, or even suffers from control spillover, at other regions. To tackle this issue, we propose the FUMIF-LCMV approach that extends the effective rendering area from locally the measured control points to a broadened area connecting the measured points. This can be done by distributing a large number of control points without measurement, as will be referred to as the “interpolated” control points hereafter.

For example, consider a linear control region with width *w*. One may choose

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

interpolated control points uniformly spaced by *d* which can be determined from the desired rendering bandwidth *fh*:

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

where *c* is the speed of sound. Therefore, a constrained optimization problem can be posed by minimizing matching error at the interpolated points, while matching the rendering path with the target model as follows:

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

where  denotes the prefilter vector of LCMV. The matching error vector  associated with the interpolated points between the desired model,  and the rendering path, **G***f***h***LCMV* can be written as

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

where  is the room frequency response matrix between the loudspeakers and the interpolated control points.

It is shown in APPENDIX A that the optimal solution of Eq. (7) is

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

where *μ* and *ε* are regularization parameters. In addition, an alternative approach can be obtained as a linear equation and the associated solution is given by,

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

where **λ** is a Lagrange multiplier vector and the L-curve method is used to select the regularization parameter *σ*. The comparative performance of the two solutions is demonstrated in Sec. III A.

Two crucial steps, parameter optimization and sensor interpolation, are entailed in practical implementation of the FUMIF-LCMV approach presented in Eq. (9).

1. Parameter Optimization

The PSO algorithm is known for its efficiency of finding the optimal solution in high-dimensional space where the objective function is highly complicated. Although grid searching for the minimum solution in the two-dimensional space is possible, the daunting task makes the calculation time very large and difficult to be implemented. Thus, we apply PSO algorithm to accelerate the procedure of finding the optimal regularization parameters. Two regularization parameters *μ* and *ε* need to be selected in implementing the optimal prefilters in Eq. (9). The following cost function of the PSO algorithm is utilized to trade-off the matching error and the filter gain:

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

where *Ngrid* ≫ *Nm* is the number of preselected uniformly spaced grid points along the control line, **e***grid* ∈ *Ngrid* and **e***m* ∈ *Nm* denote the matching error vector associated with the grid points and the measured points, respectively, *δ* is a weighting factor for the measured and the interpolated control points, and *γ* is another weighting factor found to reconcile the matching error and the filter gain. The two hyper parameters are determined with comparable number of order between the matching errors and filter gain to avoid the optimization bias toward a certain objective term.

In the PSO algorithm, particles are distributed in a two-dimensional searching space, where we reparametrize *ε=*10*-a* and *μ=*10*-b* by *a* and *b* that are easier to optimize via PSO. Therefore, the *i*th particle velocity vector **u***i*(*t*) = [*uai*(*t*) *ubi*(*t*)]*T* at time *t* is updated by

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

where **x***i*(*t*) = [*xai*(*t*) *xbi*(*t*)]*T* is the position vector of the *i*th particle, **x***p*(*t*) is the position vector of the personal best particle, **x***g*(*t*) is the position vector of the global best particle, ⊙ denotes Hadamard product, *wI* denotes the weight for the inertia, *τ*1 and *τ*2 are the acceleration coefficients, and **r**1 = [*r*1*,a* *r*1*,b*]*T*, **r**2 = [*r*2*,a* *r*2*,b*]*T* are the vectors randomly generated following the uniform distribution in interval [0,1]. The *i*th particle position vector at the next step can be obtained by

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

The update procedure stops when the loss no longer drops after five consecutive iterations and all the particles concentrate at a small 0.01×0.01 region.

1. *Sensor Interpolation*

One problem with the LCMV approach is that the FRs between the rendering loudspeakers and the immense number of interpolated points are generally unavailable and the measurement of those FRs can be a daunting task. It would be most desirable to approximate these FRs based on only a limited number of sensor measurements. To this end, we develop a two-staged sensor interpolation method in light of PWD and CS. The FRs can be expressed as a linear combination of the plane wave components. In matrix notation,

|  |  |  |
| --- | --- | --- |
|  |  | (14) |

or

|  |  |  |
| --- | --- | --- |
|  |  | (15) |

where  is the FR vector between a loudspeaker and the measuredpoints, *vn*, *n=*1,*…*,*Npw* is the coefficient of the *n*th specific plane waves, *Npw* is the number of the preselected candidate plane waves, **k***n*, *n* = 1, …, *Npw* denotes the wave vector of the *n*th plane wave component, **r***m*, *m* = 1, …, *Nm* is the position vector of the *m*th measured control point, and “•” denotes inner product.

In general, the system in Eq. (15) is highly underdetermined () and the solution is not unique. Therefore, we developed a two-stage procedure to reduce the plane wave components to a reasonable number. In stage one, we apply the CS technique29 posed in the following LASSO30 format,

|  |  |  |
| --- | --- | --- |
|  |  | (16) |

to determine the wave vectors corresponding to the significant plane wave components. In the equation above,  is the amplitude vector of the plane wave components, *Ncand.* represents the number of candidate components, ||∙||1 and ||∙||2 denote the *l*1-norm and *l*2-norm, respectively, and *λ* is a regularization parameter that weights the sparsity of **s**. The dictionary steering matrix  consists of the steering vectors from as many directions as possible. The problem of Eq. (16) can be readily solved by using the CVX31 package. The resulting non-zero entries could still be too many and *ad hoc* pruning may be required to further reduce the plane wave components to a tractable number, *Nred*. However, the identified plane wave components give only reliable direction information dictated by , but in general not reliable amplitudes. This calls for the second stage that calculates the amplitudes of the plane wave components by using TIKR.

In Stage two, we use the remaining plane wave components to construct a reduced-order steering matrix . Using this in Eq. (15) and solving it using TIKR leads to the following solution:

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

where  is the regularization parameter and . It follows that the FRs of the interpolated points can be interpolated as follows:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | | | (18) |
| or | |  | (19) | |

where **g***f*is the FR vector between a loudspeaker and the interpolated points, **A***f* denotes the steering matrix for the interpolated points, and **r***f* denotes the position vector of the *f*th interpolated point. In practice, the determination of *λ* and *η* follows the two steps below. At the first step, the regularization parameter, *λ*, is obtained through calculating the smallest matching error against a reference ATF that is premeasured at a certain interpolated point. At the second step, the other parameter, *η*, is determined through the aforementioned L-curve method.

In summary, the pratical implementation of the proposed FUMIF-LCMV approach is illustrated in FIG. 2. To begin with, the measured FRs, **g***m*,are adopted to perform the sensor interpolation, where the first stage aims to trim down the number of the plane wave components, and the second stage solves the amplitudes **v**’ of the plane wave components. Followed by the interpolation in Eq. (19), the interpolated FR, **g***f*, at any arbitrary point within a region of interest can be interpolated. Next, by utilizing the measured and interpolated FRs, as well as the target responses at the measured and interpolated points, the optimal filter of FUMIF-LCMV can be obtained through the PSO algorithm, where the cost function in Eq. (11) is designed to acquire the proper regularizaton parameters in Eq. (9). This procedure is adopted for the room response equalization problem in the following simulations and experiments.

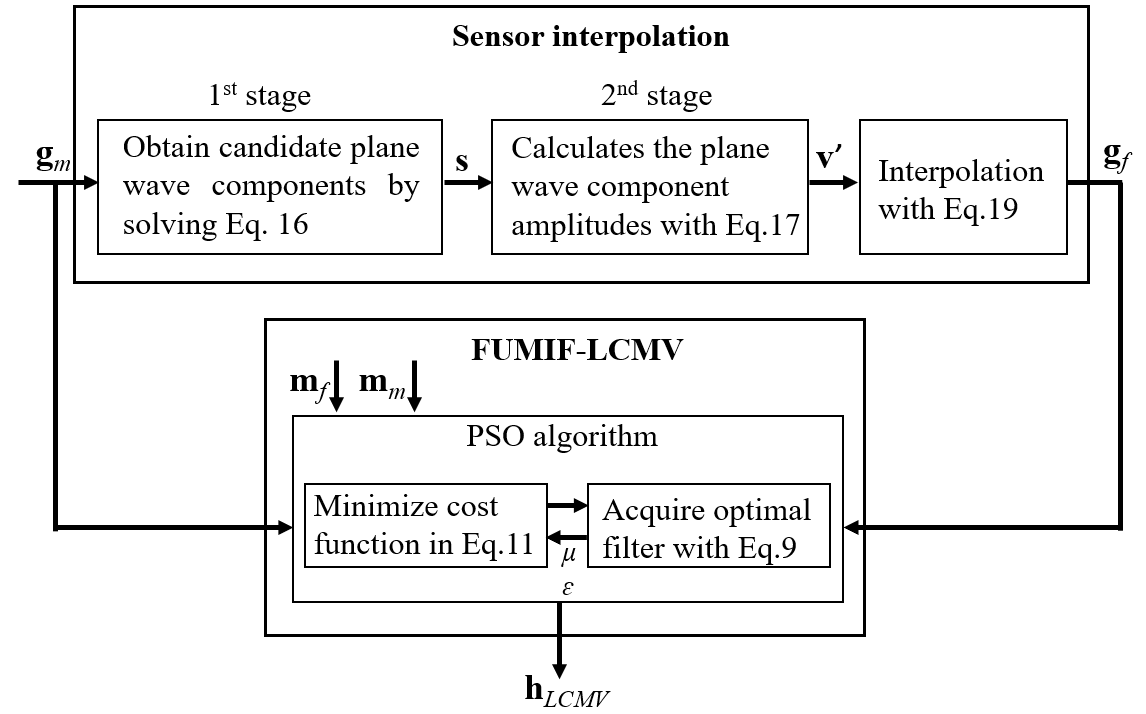


FIG. 2. The procedure of the proposed broadened response equalization approach

**III. SIMULATIONS**

1. **Room Response Equalization**

The simulation setup of the anechoic environment is illustrated in FIG. 3. An eight- loudspeaker uniform linear array with interelement spacing 0.075 m is considered in the simulation. The reproduction region of interest is along a 2 m straight line parallel to the loudspeaker array, where we distribute control points. The distance between the loudspeaker array and the listening area is 1.5 m. Seven measured control points uniformly spaced by 0.33 m are selected, as indicated by uppercase alphabets, which renders an underdetermined problem as discussed earlier. The lowercase alphabets represent the interpolated points uniformly spaced by 0.029 m. The sampling rate is 44.1 kHz.

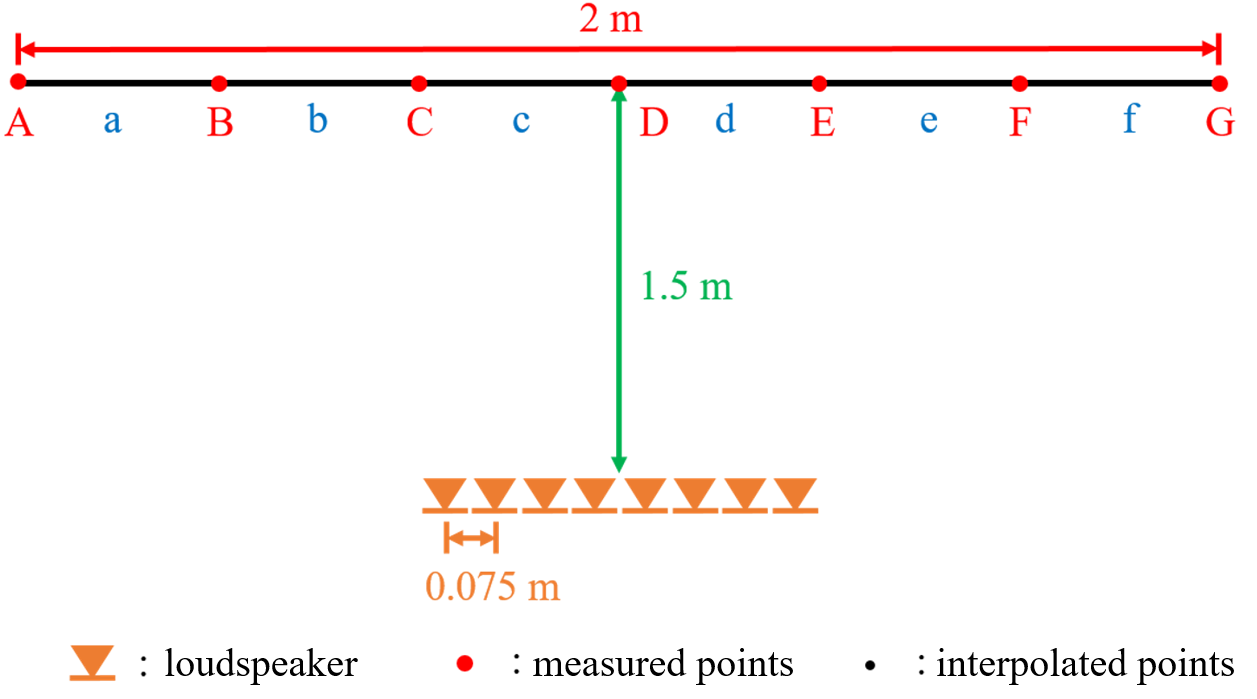


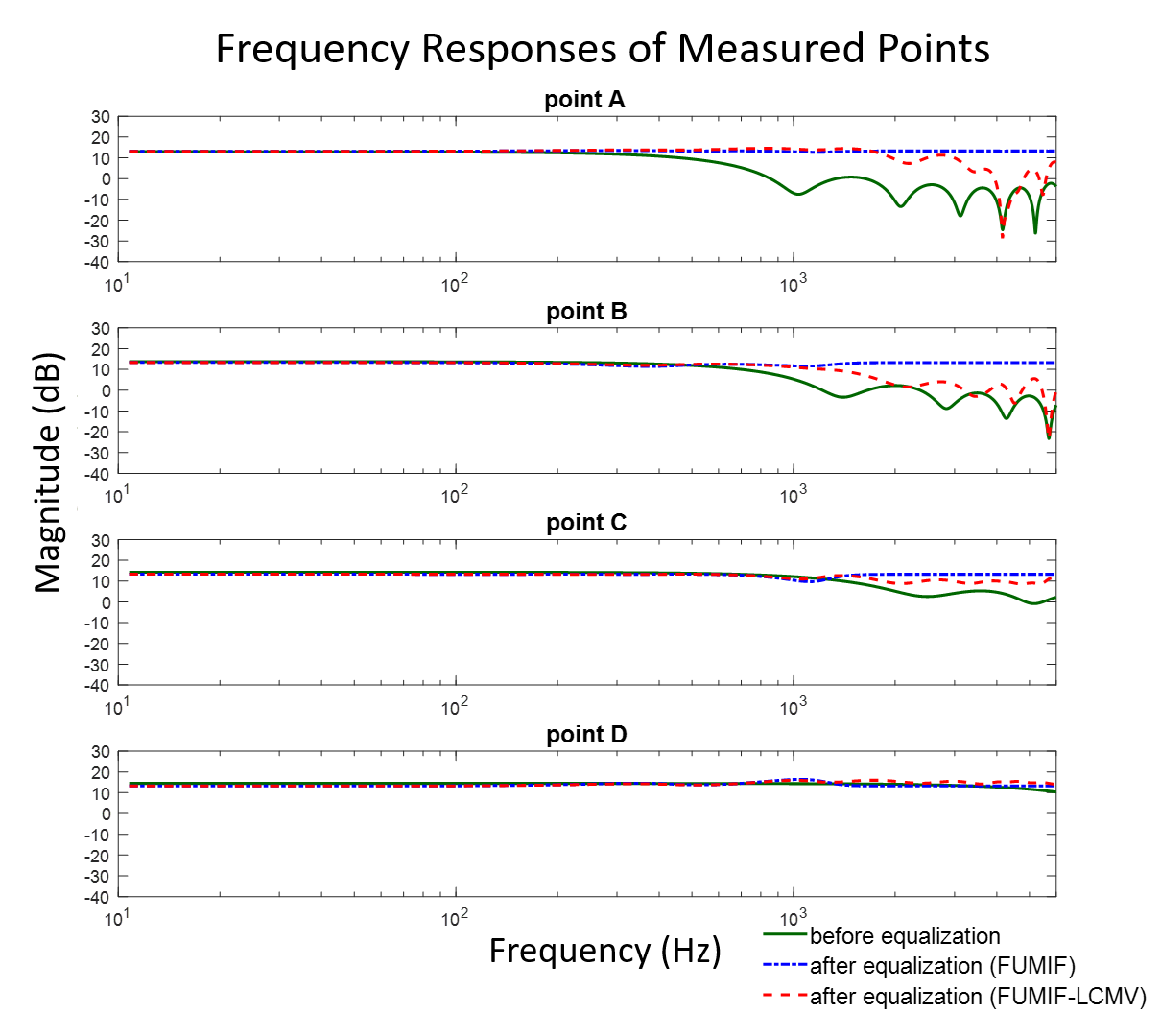
FIG. 3. The simulation settings for the 8-loudspeaker uniform linear array. The measured control points and the interpolated control points are indicated using uppercase alphabets and lowercase alphabets, respectively.

We first compare the simple FUMIF and FUMIF-LCMV approaches. The simple FUMIF approach employs the prefilters obtained using Eq. (4), while the prefilters of the FUMIF-LCMV approach are given in Eq. (9) with the regularization parameters, *ε* =10-3.99 and *μ* =10-0.71 determined using the PSO algorithm. In PSO, we choose *δ*2=0.9 and *γ*=10-3 to balance the order of the filter gain and the matching errors, which properly reconciles the optimized prefilters between the feasibility of implementation and the broadened rendering. The FRs of the measured and interpolated control points are illustrated in FIG. 4. The comparison of the FRs in FIG. 4(a) shows a better match of the FRs obtained using FUMIF for the measured control points than that obtained using FUMIF-LCMV. However, FIG. 4(b) shows the opposite that the FUMIF has little control for FRs at the interpolated points, whereas the FUMIF-LCMV approach demonstrates balanced model matching performance at the measured points and the interpolated points.

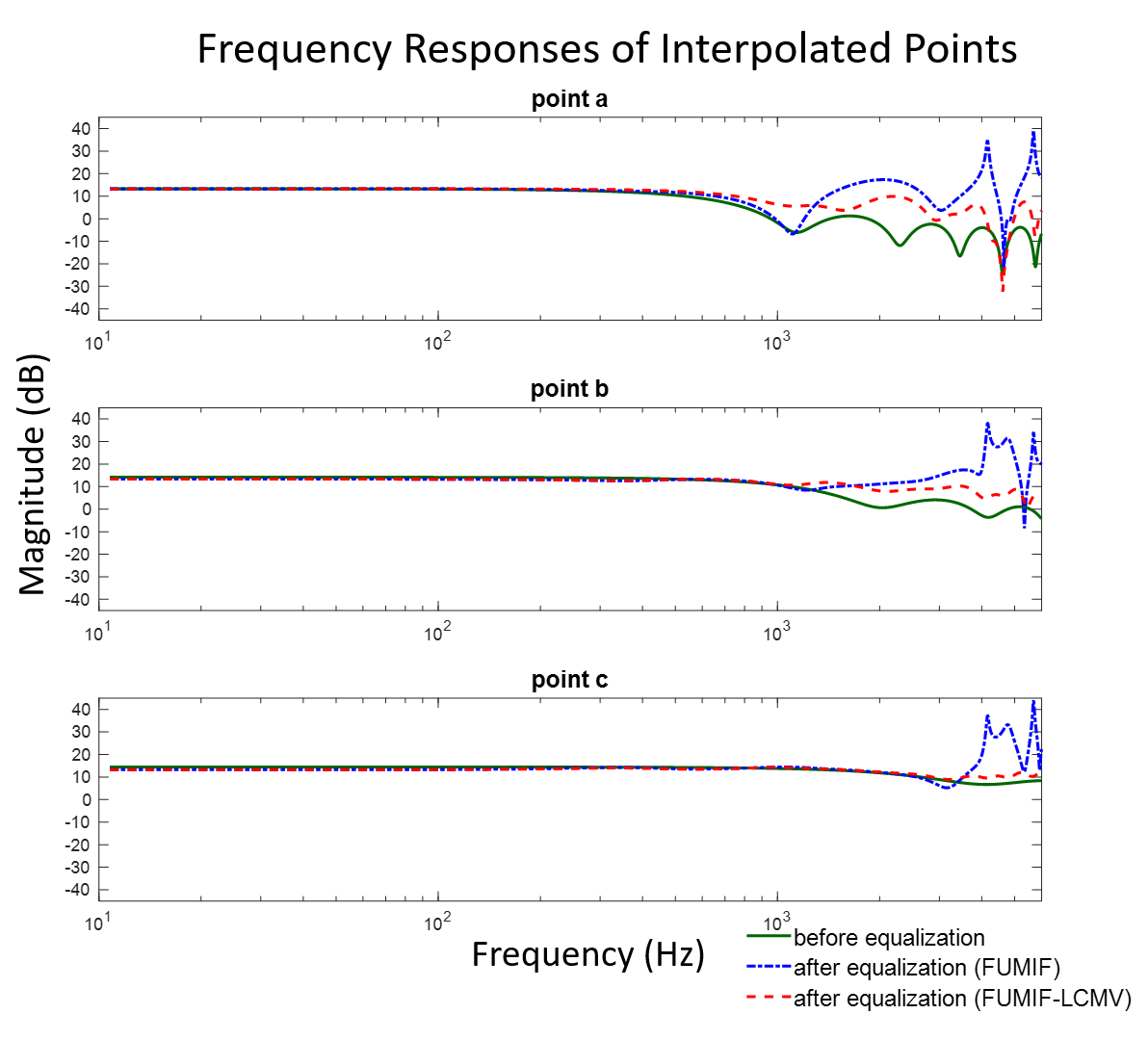
In order to visualize spatially the rendering performance in the reproduction area, we define a performance metric in terms of the normalized matching error magnitude as follows:

|  |  |  |
| --- | --- | --- |
|  |  | (20) |

where  is the FR vector between the loudspeakers and the preselected grid points,  is the prefilter vector designed using the FUMIF or the FUMIF-LCMV approaches, and *mgrid* is the desired model at the designated grid point. Therefore, the smaller is the preceding metric, the better the matching performance is. A closer look for the normalized matching error is provided in FIG. 5 in which the matching performance is plotted along a straight line connecting the control points. The result shows that FUMIF with a small regularization parameter (*β* = 10-5) leads to small matching error at the measured control points. By increasing regularization to a larger number (*β* = 10-1), the matching error at the interpolated points can be reduced with the performance degraded at the measured points. However, if the regularization parameter is increased further to 102, the normalized matching error ends up with 0 dB, which indicates that FUMIF yields no room equalization to all control points. On the other hand, the proposed FUMIF-LCMV approach provides a flexible means to achieve sufficiently low matching error across all control points in the reproduction area whether they are measured or interpolated. The normalized matching error is also calculated for every grid point and frequency and illustrated in FIG. 6. Inspection of FIG. 6(b) reveals that the FUMIF-LCMV approach attains far more broadened rendering performance in the reproduction area than the FUMIF approach that attains good performance only at the vicinity of the measured, as shown in FIG. 6(a).



(a)



(b)

FIG. 4. The comparison of the FRs of (a) the measured points and (b) the interpolated points before equalization (solid line), after equalization with FUMIF (dash-dot line) and FUMIF-LCMV (dashed line). The magnitude is shown in dB scale. Only the results for the left listening area are shown due to the symmetry.

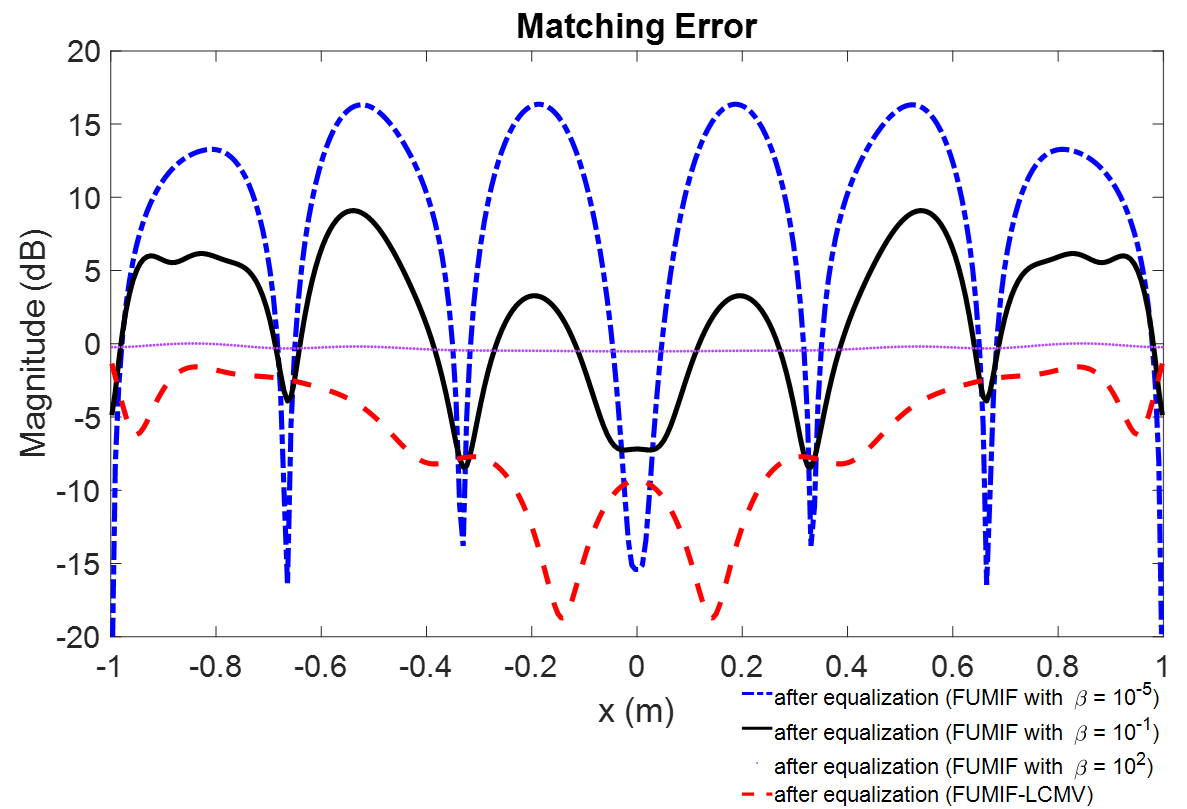
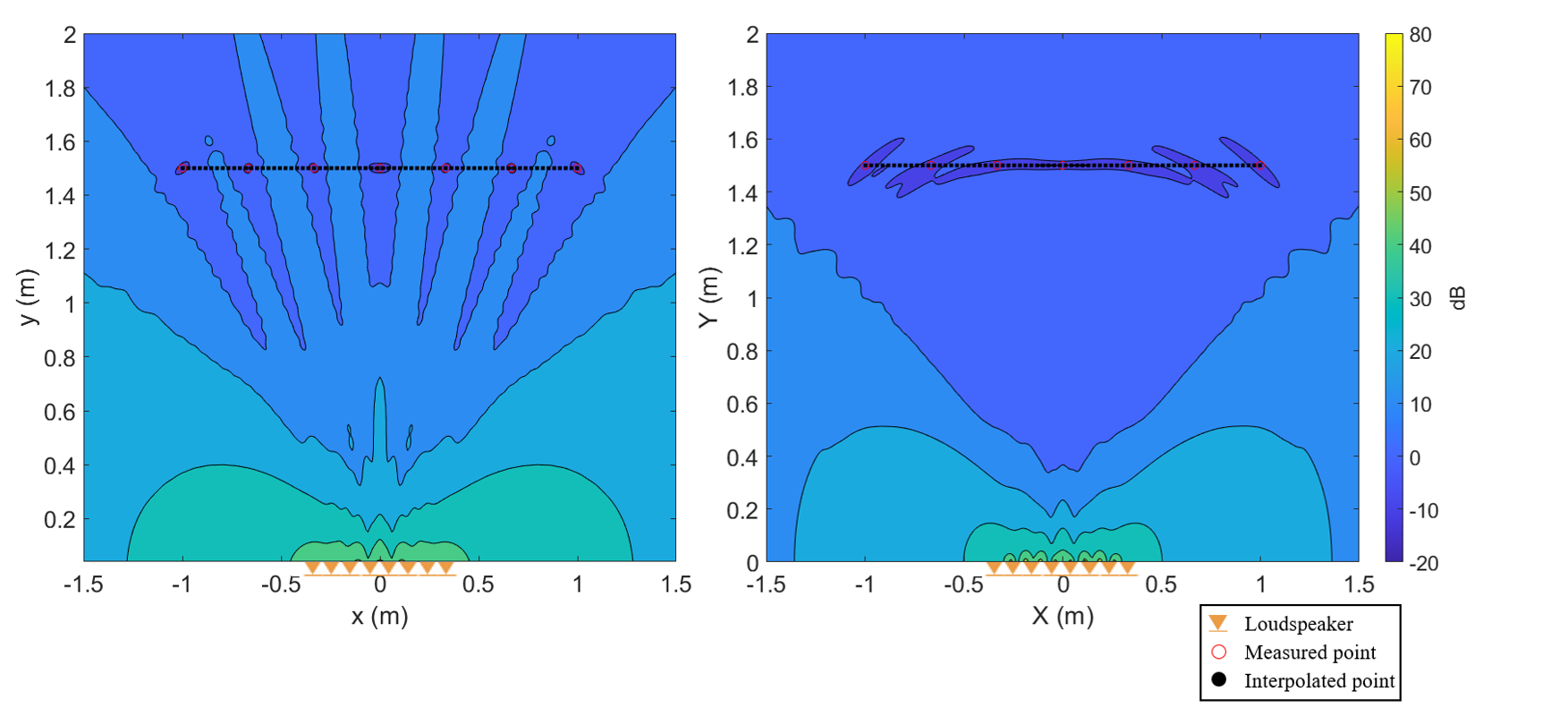


FIG. 5. The comparison of the matching error in the listening area. The magnitude is shown in dB scale. The results after equalization with the FUMIF using different regularization parameter are plotted in dash-dot line (*β* = 10-5), solid line (*β* = 10-1), and dotted line (*β* = 102). The result of FUMIF-LCMV is plotted in dashed line.



|  |  |
| --- | --- |
| (a) | (b) |

FIG. 6. The normalized matching error plotted in dB scale after equalization by (a) the FUMIF approach and (b) the FUMIF-LCMV approach.

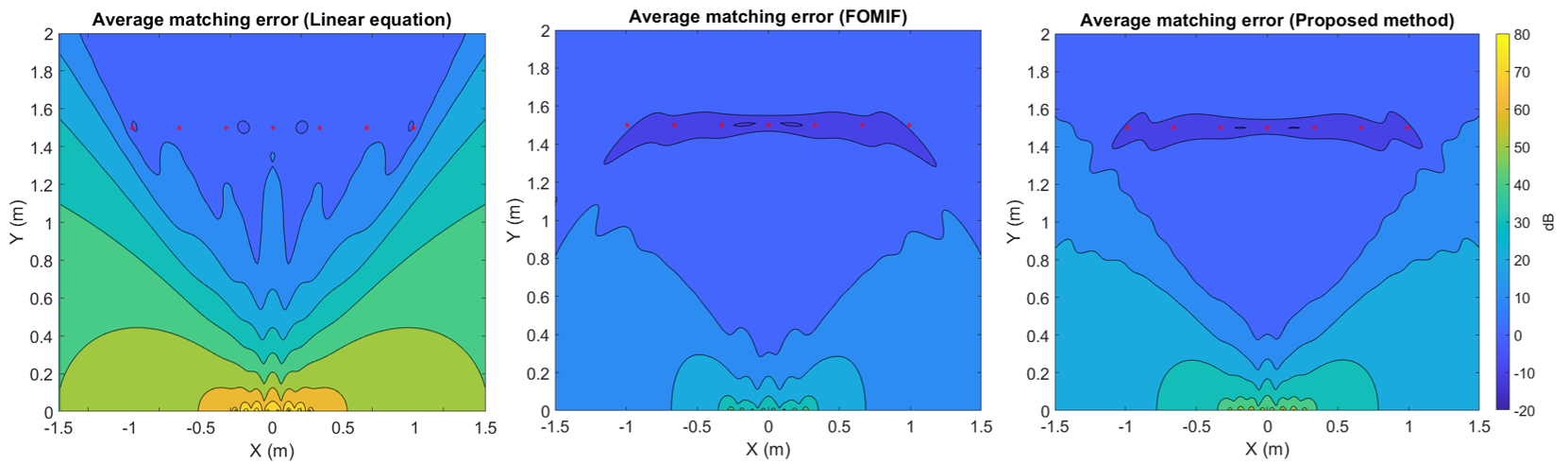
Moreover, simulation results are demonstrated in FIG. 7 and FIG. 8 to compare the proposed closed form method to two alternative approaches in terms of the normalized matching error. In addition to the linear system of equation method, a more intuitive solution is considered by directly solving the cost function of the overdetermined least-squares problem as in Eq. (11). The cost function can be written as

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

where **h***FOMIF* denotes the solution of the frequency-domain overdetermined multichannel inverse filtering (FOMIF). In the “overdetermined” formulation, sensor interpolation is used to create more control points to cover a broad control region than loudspeakers in a hope to achieve broadened control region comparable with FUMIF-LCMV. By imposing ▽**h***FOMIF* *J* = 0, the solution can be obtained as

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

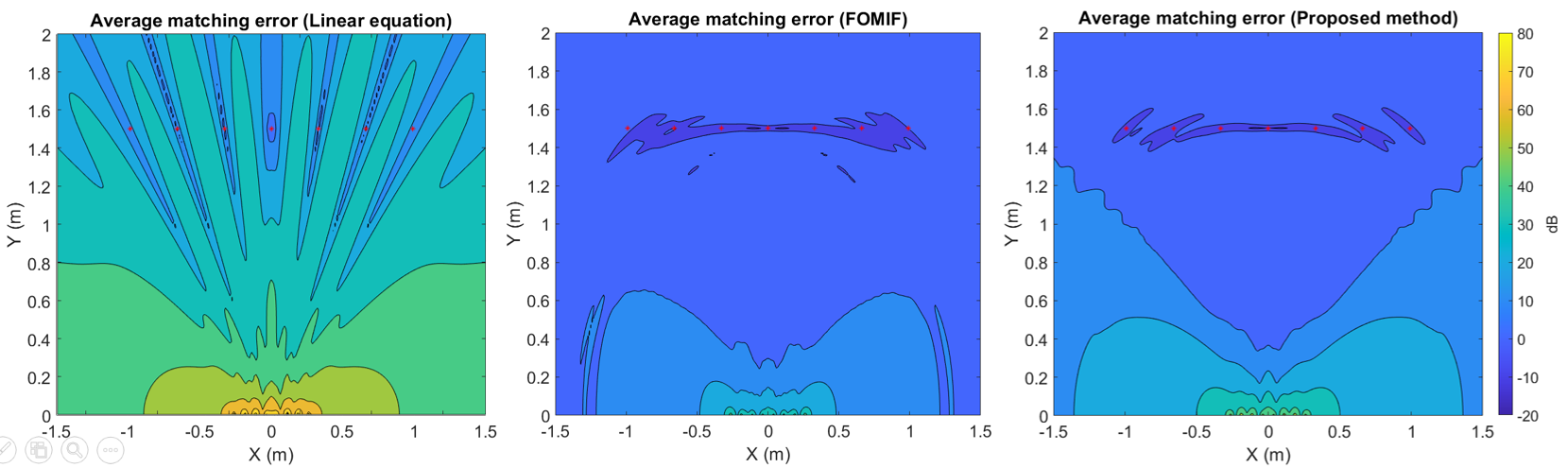
where and the weighting parameters, *δ*2= 0.9and *γ* =10-3*,* are selected as in the FUMIF-LCMV approach. In FIG. 7, the normalized matching error is averaged from 100 Hz to the spatial aliasing frequency of the loudspeakers (2 kHz). The result shows that the room equalization using the proposed solution has broader sweet spot than that obtained using the linear system of equations. Besides, FIG. 7(b) and (c) show that the FOMIF approach is capable of achieving broadened control region comparable to the proposed FUMIF-LCMV solution.



1. (b) (c)

FIG. 7. The normalized matching error plotted in dB scale after equalization obtained using (a) the linear system of equations for the FUMIF-LCMV problem, (b) the FOMIF approach, and (c) the proposed closed form solution. The beampattern is averaged from 100 Hz to 2 kHz.

FIG. 8 illustrates the beampattern of the normalized matching error averaged from 100 Hz to 6 kHz. The proposed solution and the FOMIF method still yield a broadened equalization region, while that of the linear equation approach yields an aliased pattern. This suggests the FUMIF-LCMV and FOMIF approach remains applicable for an extended frequency range. The efficacy of the approaches can be attributed to the fact that there are two parameters to optimize the equalization performance, as compared to the only one parameter in the linear equation approach.



1. (b) (c)

FIG. 8. The normalized matching error plotted in dB scale after equalization by solving (a) the linear equation of the FUMIF-LCMV approach, (b) the FOMIF approach, and (c) the proposed closed form solution. The beampattern is averaged from 100 Hz to 6 kHz.

However, there is more to the proposed FUMIF-LCMV approach than just being an alternative to the FOMIF method. One distinct feature of FUMIF-LCMV lies in the fact that the former method is based on a constraint optimization problem, where the constraint equations at the measured points impose a tighter control on the matching performance than FOMIF. The measured points can be deployed at more critical locations as in FIG. 9 for improved performance at those locations.

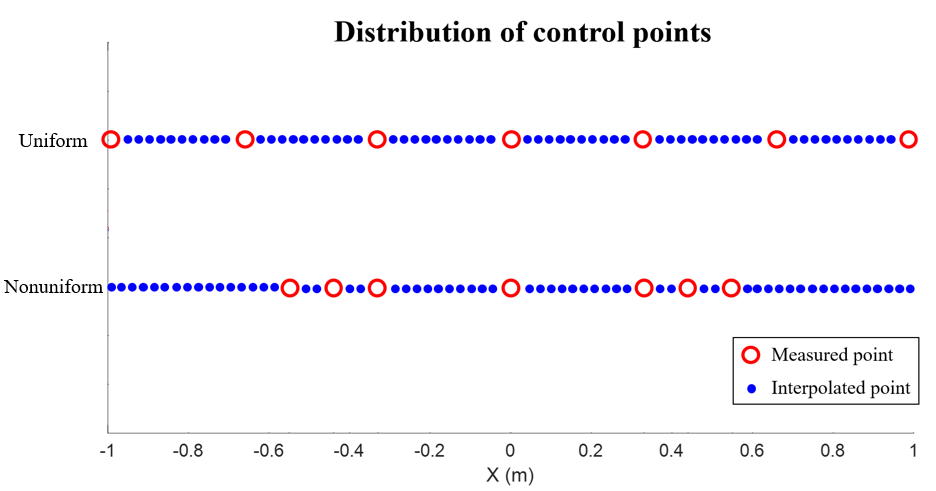


FIG. 9 Two scenarios of deployment of measured points. Nonuniform distribution is adopted to further improve the equalization performance at region of interest.

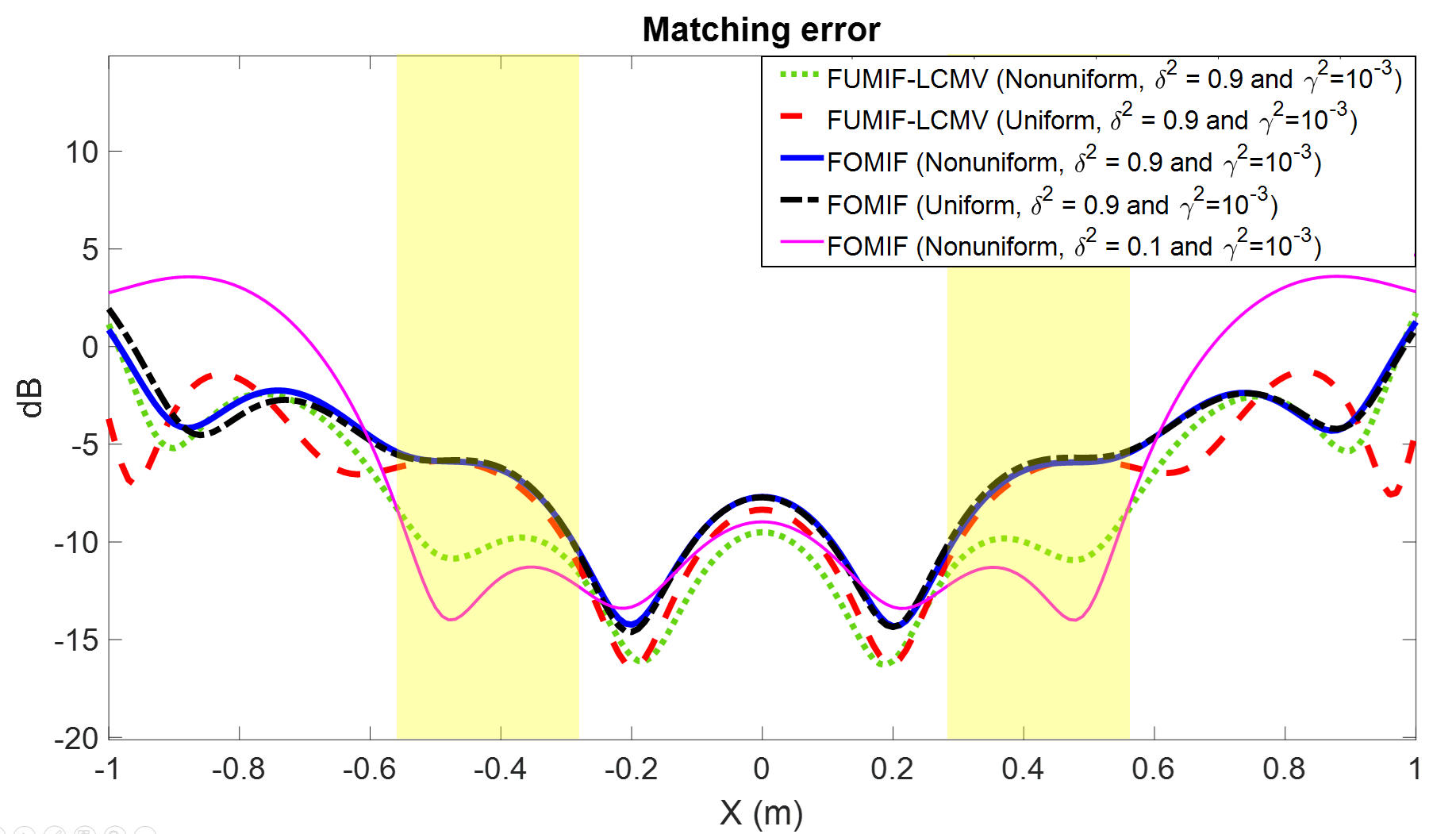


FIG. 10 The comparison of FOMIF and FUMIF-LCMV on a connected line in the broadened region of interest.

In the deployment of measured control point (FIG. 9), strategies of uniform and nonuniform distribution are considered. For the nonuniform distribution, the measured points concentrate more at x = -0.55 to -0.33 m and x = 0.33 to 0.55 m in order to improve the performance at the front region of the loudspeaker array. The number of the measured control points and the interpolated control points in both methods are 7 and 70, respectively. For these settings, the equalization performance on a connected line obtained using the FOMIF and FUMIF-LCMV approaches are shown in FIG. 10. Both methods perform comparably for uniform deployment of the measured points. When adopting nonuniform distribution, FOMIF leads to similar result as previously in the uniform distribution. In contrast, FUMIF-LCMV yields significant matching error reduction at the region governed by the constraint equations. As opposed to the FOMIF approach that has no control on the distribution of the matching error, the FUMIF-LCMV approach is able to “anchor” the high equalization performance at the critical locations via the underdetermined system of constraints. The same result is not achievable by using FOMIF with δ2 = 0.1. Performance at interpolated points could degrade and reduce the control region, as illustrated in FIG. 10. This is a remarkable feature of FUMIF-LCMV that offers more design latitude and versatility than FOMIF in achieving broadened control region by judicious deployment of measured control points in real-world settings.

1. **Sensor Interpolation**

In this section, a simulation is conducted, with the aid of the room impulse responses generated by the image source method (ISM)32, to validate the proposed sensor interpolation approach. The room is 5 *m*×4 *m*×3 *m* with reverberation time, T60, set to be 250 *ms*. The simulation setup is depicted in FIG. 11, where seven measured control points uniformly spaced by 0.33 m are deployed. Two interpolated points, A and B, are selected to evaluate the performance of the sensor interpolation. The spatial aliasing frequency corresponding to this microphone spacing is 520 Hz. The loudspeaker is placed at *θ* = -4.3∘with the distance being 1.5 m from the center normal of a seven-microphone array axis. First, we choose a 180-point uniform angular grid to constitute the dictionary for LASSO. The amplitudes referred to as the LASSO spectrum in FIG. 12(a) shows a maximum at -3∘. Next, we use aforementioned pruning procedure by applying a threshold *α* = 0.15 empirically to retain only the significant plane wave directions. Lastly, we use TIKR to calculate the amplitudes of these significant plane wave components. FIG. 12 (b) and (c) show the interpolated magnitude and phase of the FRs at point A and B below 6 kHz, which are in relatively good agreement with the ground truth up to the spatial aliasing frequency, 520 Hz. To evaluation the performance of the sensor interpolation, the interpolation error *esi* that is defined as follows:

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

where *ggt* denotes as the ground truth FR and *gsi* is the interpolated FR. As illustrated in FIG. 13, the interpolation error is small below the spatial aliasing frequency and remains under 10 dB above the aliasing frequency. This validates the efficacy of the proposed the sensor interpolation approach.

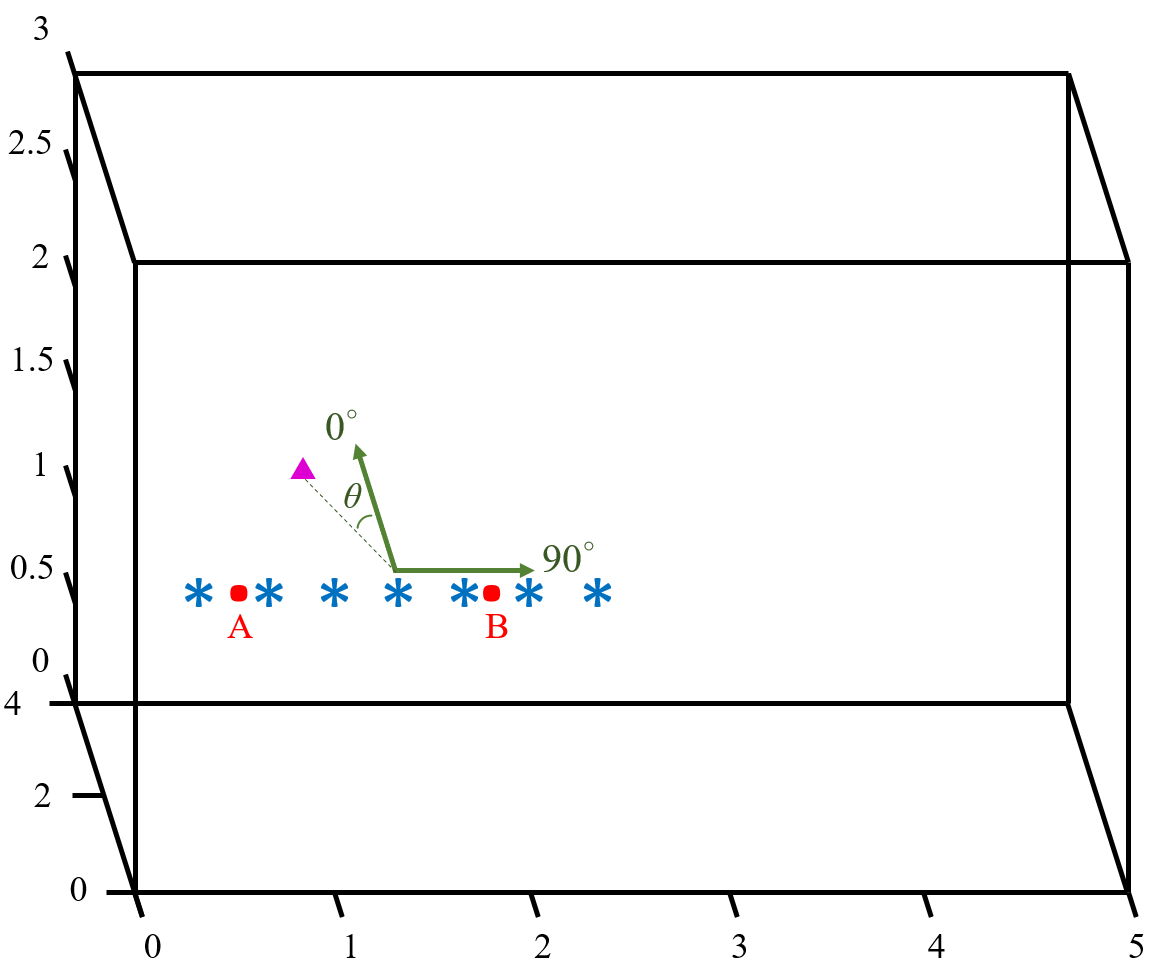
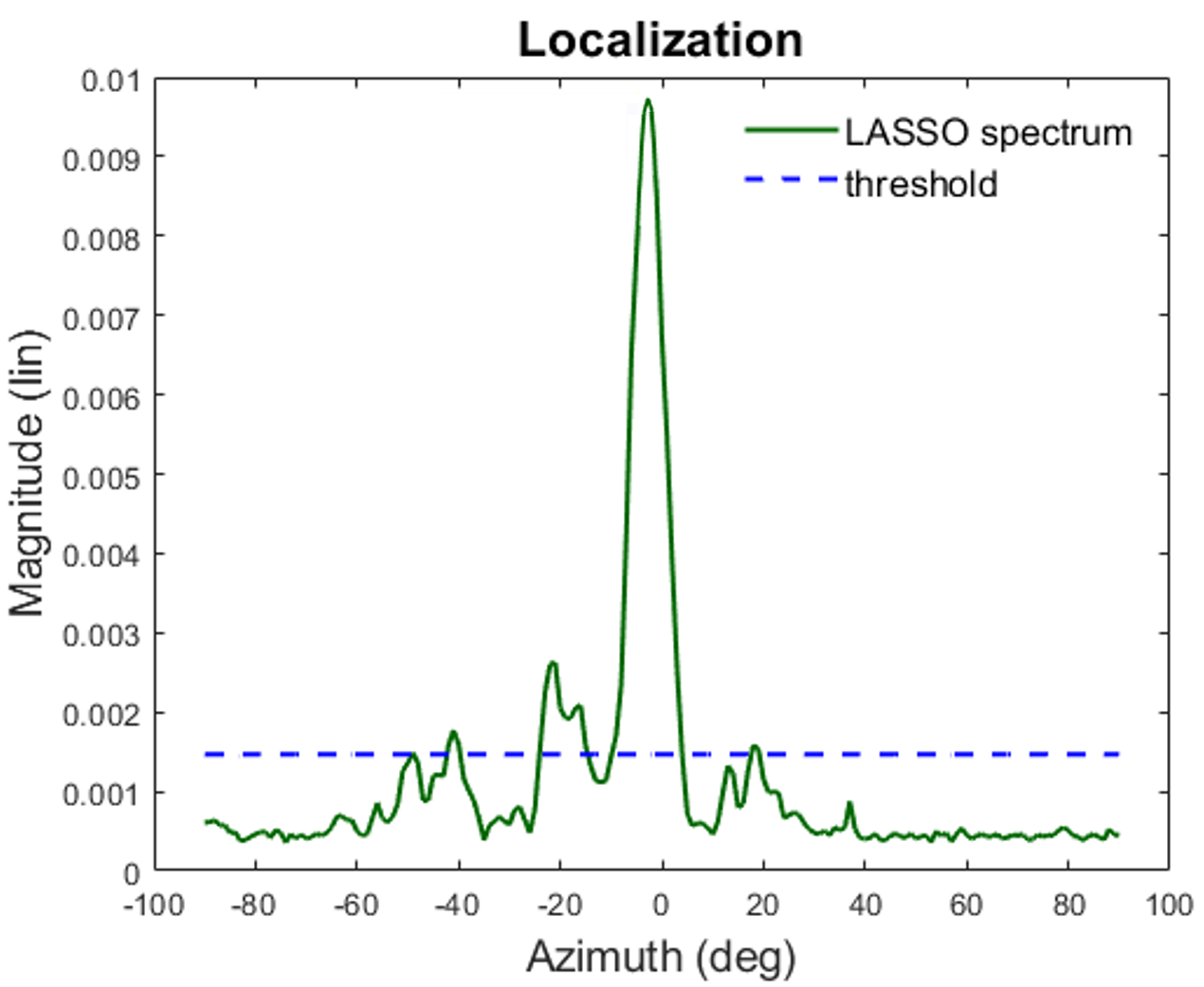
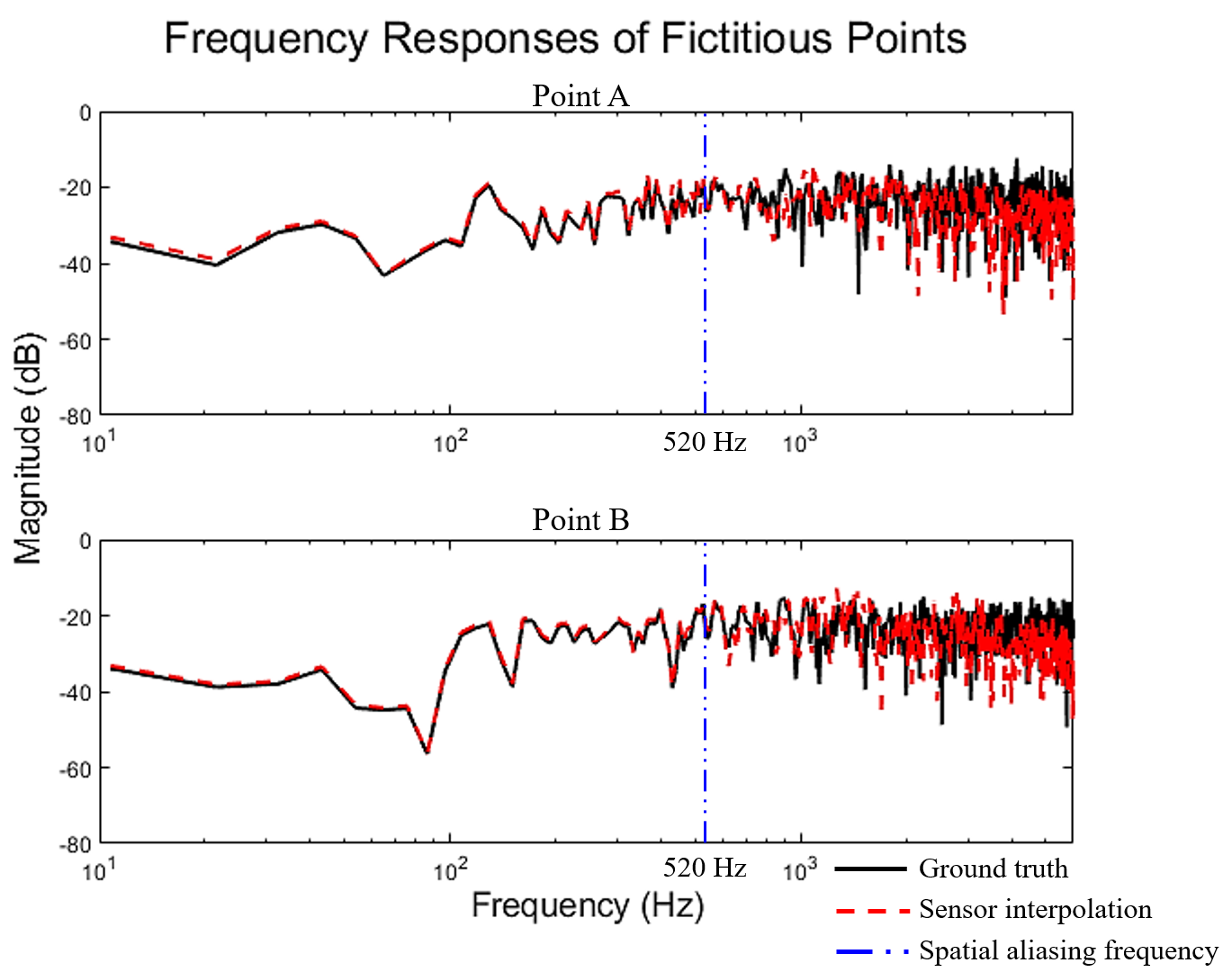


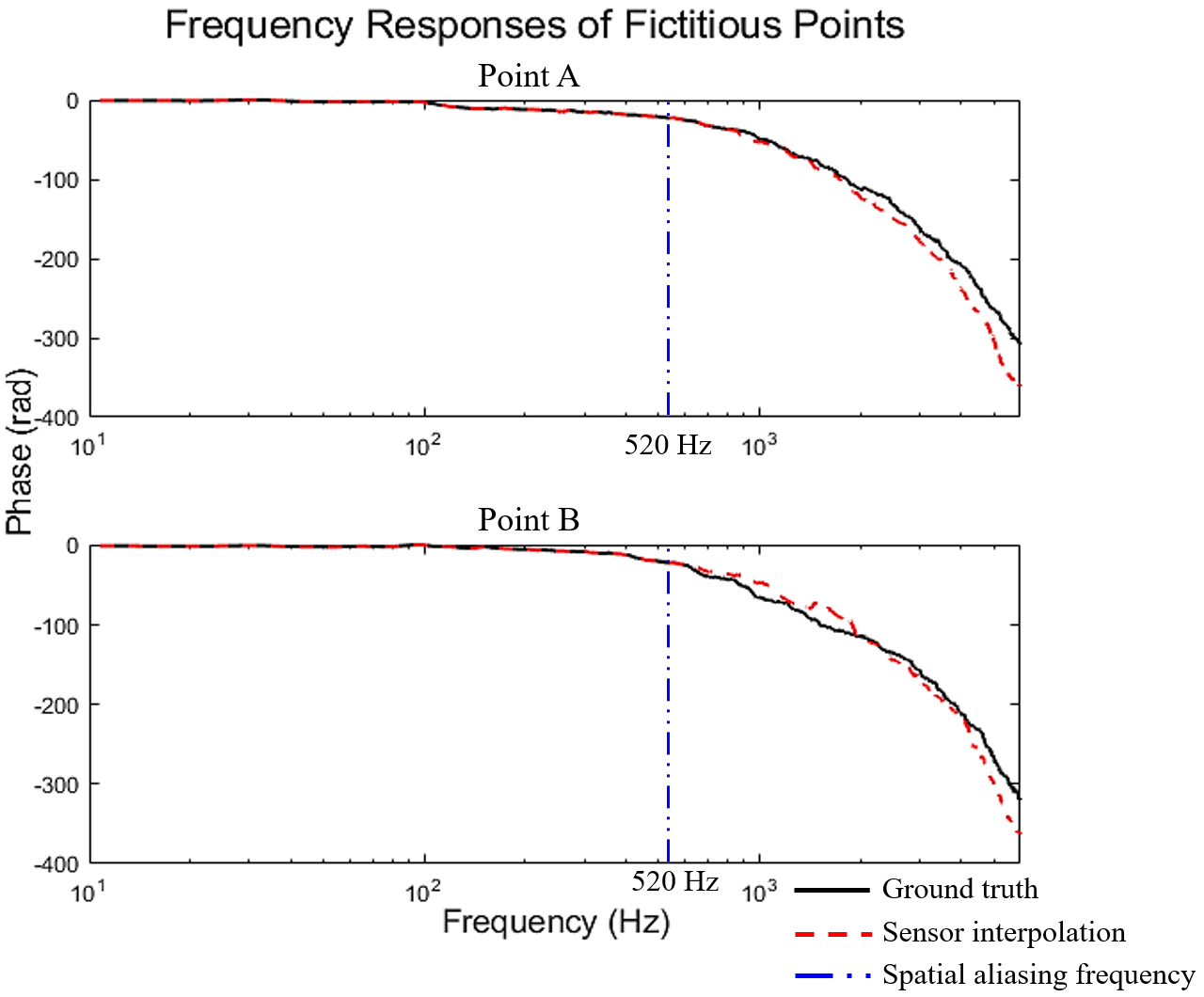
FIG. 11. The simulation settings for sensor interpolation. The symbol “▲” represents the loudspeaker. The symbols “\*” and “•” represent a measured control point and a interpolated control point, respectively. There are 7 measured points and 2 interpolated points marked A and B.



(a)



(b)



(c)

FIG. 12. The simulation results of sensor interpolation. (a) The LASSO spectrum with the threshold indicated, (b) magnitude, and (c) phase of the interpolated FRs (dashed line) and the ground truth (solid line) at the interpolated control points A and B.

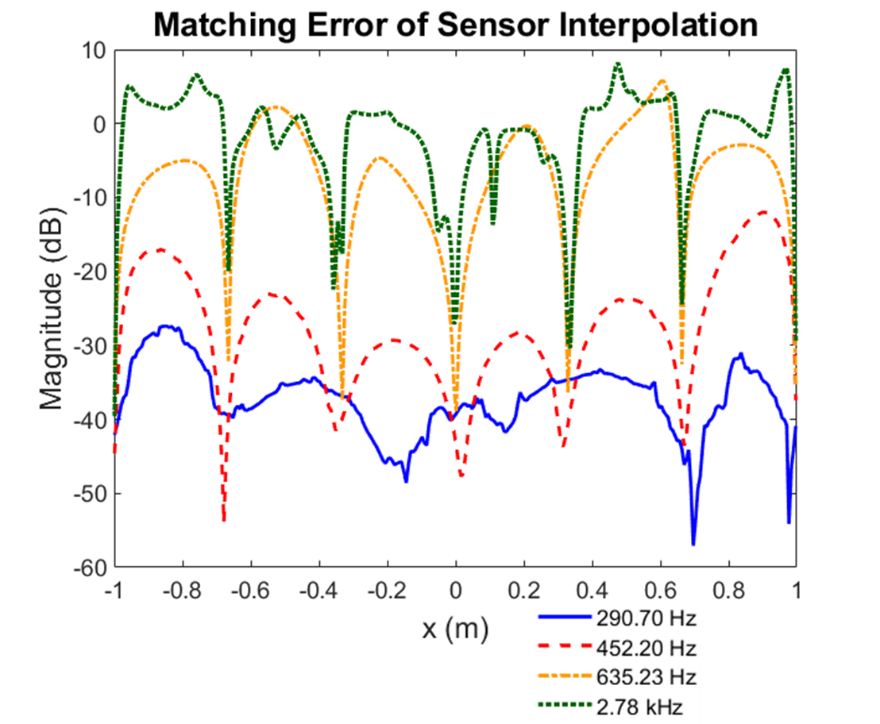
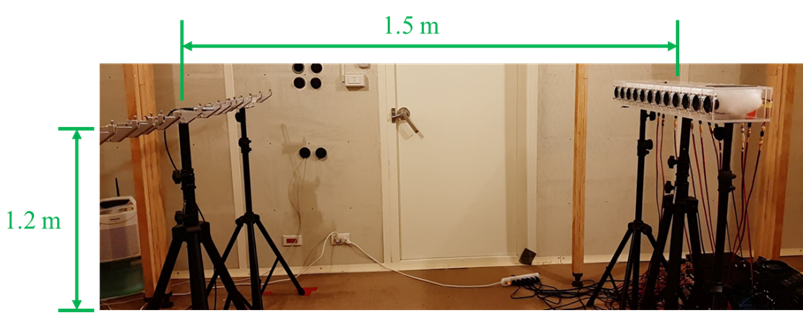


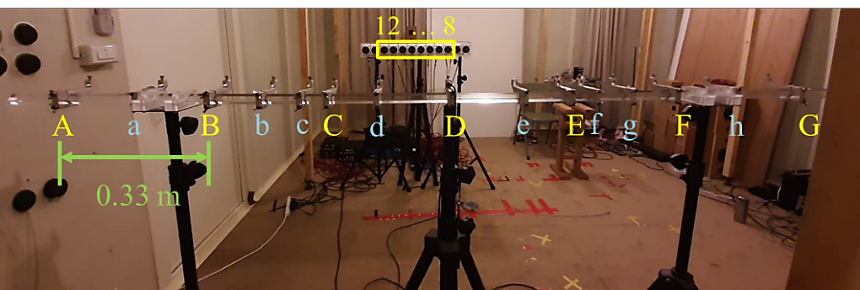
FIG. 13. The matching error obtained using the FRs of the interpolated FRs and the ground truth. The solid line and the dashed line denote the matching error below the spatial aliasing frequency (520 Hz), while the dash-dot line and the dotted line represent the matching error above the aliasing frequency.

**IV. EXPERIMENTS**

In order to validate the proposed approach under practical conditions, experiments are conducted in a listening room with T60 = 250 ms, the same as in the simulation. The experimental arrangement includes a uniform linear loudspeaker array with interelement spacing 0.075 m and a linear microphone array with adjustable spacing, as shown in FIG. 14. The intended listening area is 2 m wide. The measured and the interpolated control points are uniformly spaced by 0.33 m and 0.029 m, respectively. The microphones marked with lowercase alphabets in FIG. 14(b) denote the interpolated control points. The loudspeakers and the microphones are placed in two parallel lines spaced by 1.5 m in the height of 1.2 m. FRs between loudspeakers and microphones are measured by a signal analyzer, PULSE of B&K®, at the sampling rate 44.1 kHz.



(a)

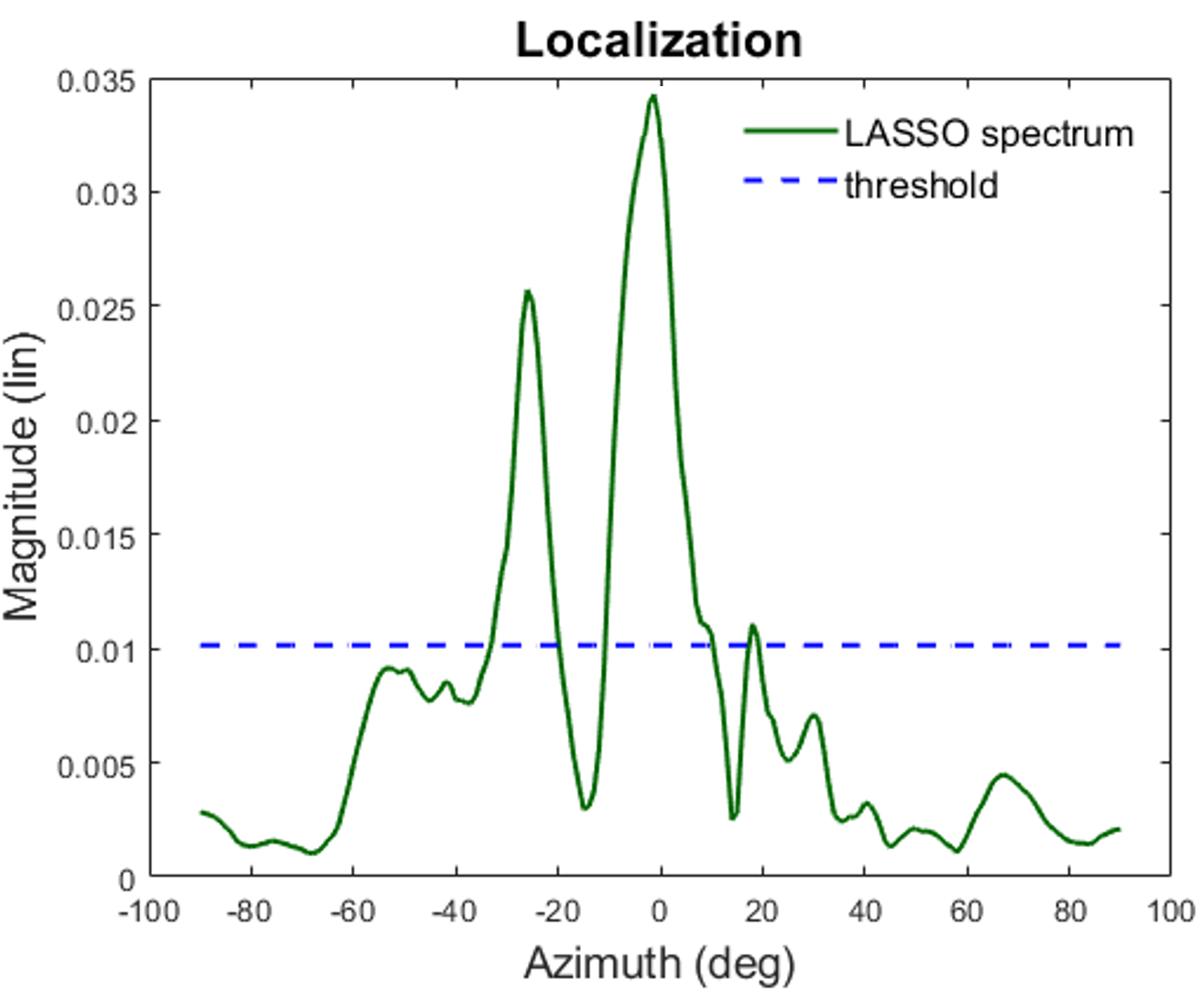


(b)

FIG. 14. Experimental arrangement. (a) The side view, (b) 8 rendering loudspeakers (numeric in the yellow box), 7 measured control points (uppercase alphabet), and 8 interpolated control points (lowercase alphabet).

1. **Sensor Interpolation**

As discussed earlier, we approximate the FRs by using sensor interpolation prior to the design of LCMV. The experimental results of the interpolated FRs for the third loudspeaker and the control points are shown in FIG. 15. The calculated LASSO spectrum in FIG. 15(a) has two peaks with one of which pointing to the third loudspeaker located at -2∘. The other peak could be due to the reflection from a side wall where the loudspeaker array is close to. As in the simulation, we retain only significant PWD by applying a threshold *α* = 0.3. The magnitude and phase of thus interpolated FRs associated with the interpolated points are compared with the measured FRs at the same interpolated points in FIG. 15(b) and (c). The results reveal that the interpolated FRs have approximated very well the magnitude of FRs at the interpolated points below the spatial aliasing frequency (520 Hz), whereas accurate phase interpolation extends to much higher frequency. This result demonstrates that the proposed sensor interpolation procedure is able to reliably reconstruct the FRs, based on a limited number of measurements, for the interpolated points to be used in formulating the cost function in the LCMV approach.



(a)



(b)



(c)

FIG. 15. The experimental results of sensor interpolation. (a) The LASSO spectrum with the threshold indicated, (b) magnitude, and (c) phase of the interpolated FRs (dashed line) and the measured FRs (solid line) at the interpolated control points.

1. **Room Response Equalization**

Experiments are conducted to validate the proposed FUMIF-LCMV approach for room response equalization. The regularization parameters for the FUMIF-LCMV approach determined using PSO are *ε* =10-3.15 and *μ* =10-0.84. FigureFIG. 16 shows the matching error predicted using the interpolated FRs of all control points. The FUMIF approach with *β* = 10–2 has achieved nearly perfect matching at only the neighborhood of the measured control points. When the regularization parameter of FUMIF is increased to *β* = 102,the matching error reaches 0 dB. Like the results indicated in the simulation, broadened equalization region is not attainable by choose a regularization parameter in FUMIF. By contrast, the matching error resulting from the FUMIF-LCMV approach are predominantly smaller than that resulting from the FUMIF approach throughout the range *x* = -1 m to 1 m, except the measured control points. This suggests the possibility of comprehensive control with a broadened sweet spot. Figure FIG. 17 plots the matching errors of the FRs recorded at the measured control points and the interpolated points, respectively. Although FUMIF and FUMIF-LCMV perform comparably at the measured control points, the latter approach outperforms the former approach at the interpolated control points.

To evaluate the audio quality resulting from the proposed equalization approach, the PEAQ is evaluated and illustrated in FIG. 18. The PEAQ score obtained using the FUMIF approach fluctuates drastically, as compared with the PEAQ score obtained using the FUMIF-LCMV approach. This confirms experimentally that the FUMIF-LCMV approach is capable of more broadened and artifact-free audio rendering than the FUMIF approach.

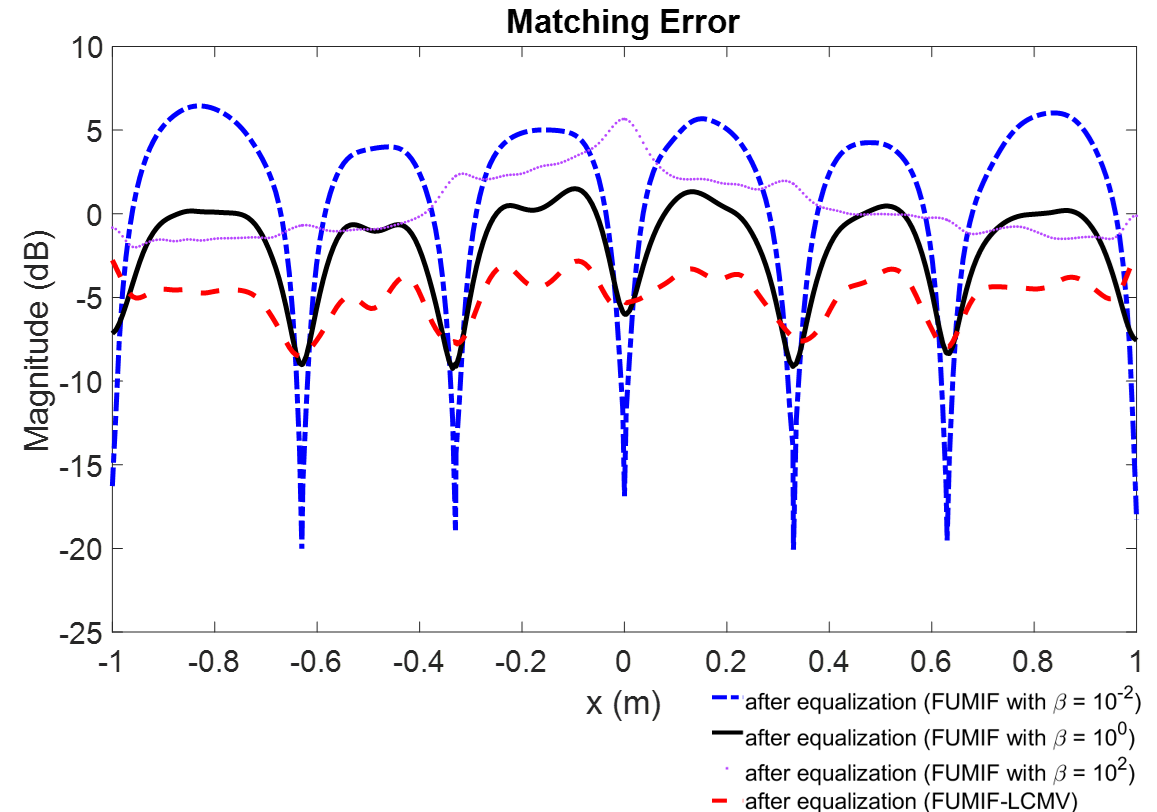


FIG. 16. Comparison of the matching error resulting from the FUMIF approach with *β* = 10–2 (dash-dot line), *β* = 100 (solid line), *β* = 102 (dotted line), and the FUMIF-LCMV approach (dashed line).



(a)



(b)

FIG. 17. The matching errors at (a) the measured control points and (b) the interpolated control points without equalization (solid line) and with equalization, resulting from the FUMIF approach (dash-dot line) and the FUMIF-LCMV approach (dashed line).

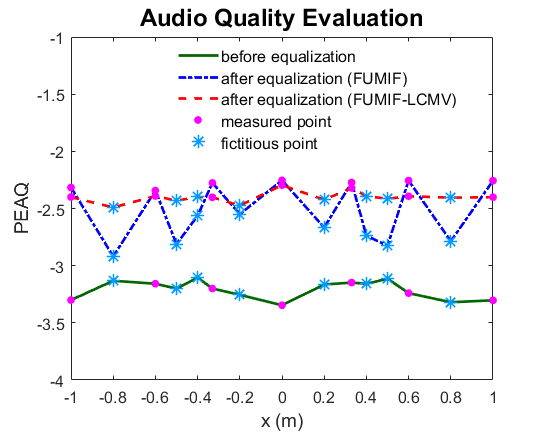


FIG. 18. Comparison of the PEAQ scores without equalization (solid line) and with equalization, resulting from the FUMIF approach (dash-dot line) and the FUMIF-LCMV approach (dashed line).

In addition to the objective experiments, a listening experiment is conducted according to the multi-stimulus test with hidden reference and anchor (MUSHRA)33 procedure. There were 26 subjects participating in this experiment. In the group, 16 subjects have audio signal processing background. The subjective indices adopted in the test include Equalization performance, Artifacts, and overall preference. The score ranges from 1 to 5 points. The evaluation procedure is explained to all participants before the test. A ten-second audio clip of *Symphony No. 5 in C minor* by Beethovenis adopted as the reference signal. The testing audio signal is presented with the loudspeaker array in the listening room depicted in FIG. 14. The participants were told to sit at the center of the control region with their head located at position D. Highpass-filtered reference signal is used as the anchor signal. The test stimuli are the signals without and with equalization using FUMIF and FUMIF-LCMV, recorded at the measured and the interpolated control points.

The results of the listening experiment processed by analysis of variance (ANOVA) are summarized in FIG. 19, where the equalization approaches are annotated with “m” and “f” to indicate the stimuli are recorded at the measured and the interpolated control points, respectively. It can be seen in FIG. 19 that the audio quality of the signals recorded at the measured control points is significantly improved by both FUMIF and FUMIF-LCMV approaches. However, the audible difference in audio quality exists between the signal equalized using the FUMIF approach at the interpolated control points and at the measured control points. Furthermore, The FUMIF-equalized signal recorded at the interpolated control points sounds similar to the original signal without equalization, which suggests that the FUMIF approach has little influence on the interpolated control points. In contrast, little difference of the signals equalized with the FUMIF-LCMV approach recorded at the interpolated control points and the measured control points can be perceived. This suggests that the FUMIF-LCMV approach indeed is capable of more broadened rendering than FUMIF at some cost of performance at the measured control points, which accounts for only a very small fraction of rendering area. A pairwise comparison is also carried out by using a *post hoc* Fisher’s least significant difference (LSD) test34, as summarized in Table I. In the results, significant difference of performance indices exists at both measured and the interpolated points when applying the FUMIF approach. In addition, the equalization effect using the FUMIF approach at the interpolated points is not significant because the *p*-value associated with the rendering before and after equalization is greater than 0.05. On the other hand, the difference between the equalization at the measured points and interpolated points is not audible in the application of FUMIF-LCMV.

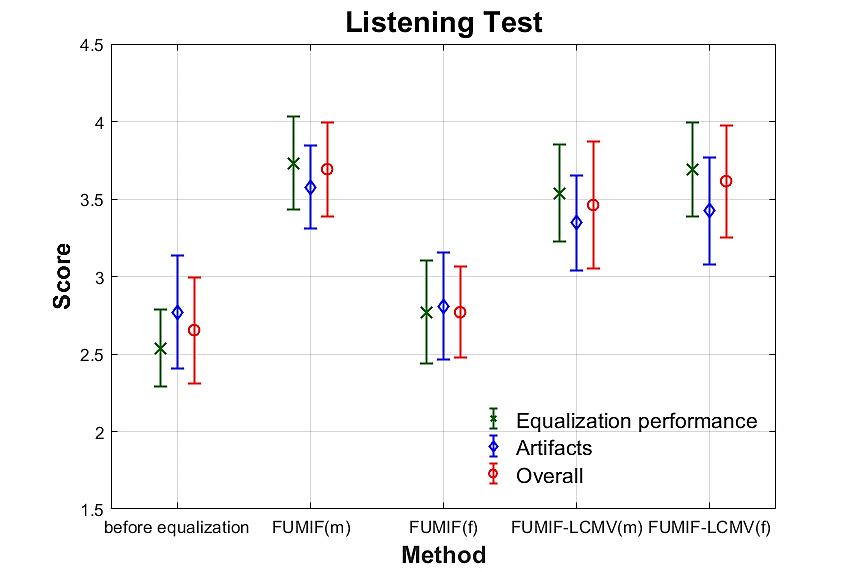


FIG. 19. The results of the listening experiment. The symbols ‘\*’, ‘◇’ and ‘○’ represent Equalization performance, Artifacts, and Overall preference.

Table I. The results of a *post hoc* Fisher’s LSD test. The cases when *p*-value < 0.05 are boldfaced.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Before equalization | FUMIF (m) | FUMIF (f) | FUMIF-LCMV (m) | FUMIF-LCMV (f) |
| Equalization Performance | | | | | |
| Without equalization |  | **2.11×10-7** | 0.28 | **1.00×10-5** | **4.99×10-7** |
| FUMIF (m) | **2.11×10-7** |  | **1.02×10-4** | 0.39 | 0.86 |
| FUMIF (f) | 0.28 | **1.02×10-4** |  | **1.73×10-3** | **1.92×10-4** |
| FUMIF-LCMV (m) | **1.00×10-5** | 0.39 | **1.73×10-3** |  | 0.49 |
| FUMIF-LCMV (f) | **4.99×10-7** | 0.86 | **1.92×10-4** | 0.49 |  |
| Artifacts | | | | | |
| Without equalization |  | **1.04×10-3** | 0.88 | **0.02** | **0.01** |
| FUMIF (m) | **1.04×10-3** |  | **1.16×10-3** | 0.27 | 0.50 |
| FUMIF (f) | 0.88 | **1.16×10-3** |  | **0.03** | **0.02** |
| FUMIF-LCMV (m) | **0.02** | 0.27 | **0.03** |  | 0.75 |
| FUMIF-LCMV (f) | **0.01** | 0.50 | **0.02** | 0.75 |  |
| Overall Preference | | | | | |
| Without equalization |  | **4.84×10-5** | 0.62 | **4.67×10-3** | **4.18×10-4** |
| FUMIF (m) | **4.84×10-5** |  | **8.29×10-5** | 0.38 | 0.75 |
| FUMIF (f) | 0.62 | **8.29×10-5** |  | **9.71×10-3** | **8.30×10-4** |
| FUMIF-LCMV (m) | **4.67×10-3** | 0.38 | **9.71×10-3** |  | 0.58 |
| FUMIF-LCMV (f) | **4.18×10-4** | 0.75 | **8.30×10-4** | 0.58 |  |

**V. CONCLUSIONS**

An LCMV-based room response equalization technique on the basis of underdetermined multi-channel inverse filtering has been formulated and implemented in this contribution. The proposed FUMIF-LCMV approach is shown to be effective in immersive rendering in a broadened region by using a loudspeaker array. The contribution of this work is twofold. First, a two-staged sensor interpolation technique is introduced. Second, a multichannel equalization system is reformulated in the LCMV framework. Regularization parameters of the FUMIF-LCMV approach are optimized using the PSO algorithm to maximize the equalization performance. The results of extensive simulations, objective and subjective experiments have confirmed the efficacy of the proposed system in widening the listening area with only a limited number of discrete measurements. In addition, the audio quality is well preserved with the equalization processing at not only the measured control points but also the interpolated control points.

**ACKNOWLEDGMENTS**

The work was supported by the Ministry of Science and Technology (MOST) in Taiwan, under the project number 109-2221-E-007-010-MY3.

**APPENDIX A Derivation of the FUMIF-LCMV approach**

The FUMIF-LCMV approach can be formulated as the following constrained optimization problem:

|  |  |  |
| --- | --- | --- |
|  |  | (A1) |

where

|  |  |  |
| --- | --- | --- |
|  |  | (A2) |

By the method of Lagrange multiplier, we write a Lagrangian

|  |  |  |
| --- | --- | --- |
|  |  | (A3) |

where  is a Lagrange multiplier vector. Taking the complex gradient of the Lagrangian and setting it to zero, we obtain

|  |  |  |
| --- | --- | --- |
|  |  | (A4) |

Solving (A4) for the optimal prefilter vector leads to

|  |  |  |
| --- | --- | --- |
|  |  | (A5) |

Using (A5) in the constraint equation (A2), we get

|  |  |  |
| --- | --- | --- |
|  |  | (A6) |

which gives the Lagrange multiplier 

|  |  |  |
| --- | --- | --- |
|  |  | (A7) |

where **G***f* is assumed to has full-column rank such that  is invertible. Thus,  Finally, using (A7) in (A5) yields the optimal solution of the FUMIF-LCMV prefilters

|  |  |  |
| --- | --- | --- |
|  |  | (A8) |

The regularization parameters, *ε* and *μ*, are inserted to mitigate the ill-posedness in (A8) during matrix inversion. An alternative approach can be obtained by utilizing the constraint equation in Eq. (A1) and (A4) to form a linear equation, as given by

|  |  |  |
| --- | --- | --- |
|  |  | (A9) |

Nevertheless, a diagonal loading term (*σ***I**)is still required to mitigate the ill-posedness at certain frequencies.

|  |  |  |
| --- | --- | --- |
|  |  | (A10) |

where the L-curve method is used to select the regularization parameter *σ*.

**REFERENCES**

1. A. Mäkivirta, P. Antsalo, M. Karjalainen, and V. Välimäki, “Modal equalization of loudspeaker-room responses at low frequencies,” J. Audio Eng. Soc. **51**(5), 324–343 (2003).
2. P. Antsalo, M. Karjalainen, A. Mäkivirta, and V. Välimäki, “Comparison of modal equaliser design methods,” in *Proceedings of the 114th Audio Engineering Society Convention*, Preprint 5844 (2003)
3. D. S. Talagala, W. Zhang, and T. D. Abhayapala, “Efficient multi-channel adaptive room compensation for spatial soundfield reproduction using a modal decomposition,” IEEE/ACM Trans. Audio, Speech, Lang. Process. **22**(10), 1522–1532 (2014).
4. L.-J. Brännmark, A. Bahne, and A. Ahlén, “Spatially robust audio compensation based on SIMO feedforward control,” IEEE Trans. Audio Speech Language Proc. **57**, 1689-1702 (2009).
5. L.-J. Brännmark, A. Bahne, and A. Ahlén, “Compensation of loudspeakerroom responses in a robust MIMO control framework,” IEEE Trans. Audio Speech Language Proc. **21**, 1201–1215 (2013).
6. M. Miyoshi and Y. Kaneda, ‘‘Inverse filtering of room acoustics,’’ IEEE Trans. Acoust. Speech Signal Process. **36**(2), 145–152 (1988).
7. M. R. Bai, Y. W. Chen, Y. C. Hsu, and T. Y. Wu, “Robust binaural rendering with the time-domain underdetermined multichannel inverse prefilters,” J. Acoust. Soc. Am. **146**, 1302 (2019).
8. C. W Groetsch, “The theory of Tikhonov regularization for Fredholm equation of the first kind,” in *Pitman Advanced Pub. Program*, Boston (1984).
9. M. Bertero, T. A. Poggio, and V. Torre, “Ill-posed problems in early vision,” in Proceedings of the IEEE, **76(8)**, 869-889 (1988).
10. Q. Zhu, P. Coleman, M. Wu, and J. Yang, “Robust acoustic contrast control with reduced in-situ measurement by acoustic modeling,” J. Audio Eng. Soc. **65**(6), 460–473 (2017).
11. 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).
12. F. M. Heuchel, D. Caviedes-Nozal, J. Brunskog, F. T. Agerkvist, and E. Fernandez-Grande, “Large-scale outdoor sound field control,” J. Acoust. Soc. Am. **148**(4), 2392–2402 (2020).
13. M. R. Bai, Y. Li, and Y.-H. Chiang, “Modeling of reverberant room responses for two-dimensional spatial sound field analysis and synthesis,” J. Acoust. Soc. Am. **142**(4), 1953–1963 (2017).
14. M. Nolan, E. Fernandez-Grande, J. Brunskog, and C.-H. Jeong, “A wavenumber approach to quantifying the isotropy of the sound field in reverberant spaces,” J. Acoust. Soc. Am. **143**(4), 2514–2526 (2018).
15. B. Rafaely, “Plane-wave decomposition of the sound field on a sphere by spherical convolution,” J. Acoust. Soc. Am. **116**(4), 2149-2157 (2004).
16. C.-X. Bi, Y. Liu, L. Xu, and Y.-B. Zhang, “Sound field reconstruction using compressed modal equivalent source method,” J. Acoust. Soc. Am. **141**(1), 73–79 (2017).
17. 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).
18. R. Mignot, 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).
19. D. Caviedes-Nozal, N. A. B. Riis, F. M. Heuchel, J. Brunskog, P. Gerstoft, and E. Fernandez-Grande, “Gaussian processes for sound field reconstruction,” J. Acoust. Soc. Am. **149**(2), 1107–1119 (2021).
20. M. Hahmann, S. A. Verburg, and E. Fernandez-Grande, “Spatial reconstruction of sound fields using local and data-driven functions,” J. Acoust. Soc. Am. **150**(6), 4417–4428 (2021b).
21. R. Tibshirani, “Regression Shrinkage and Selection via the Lasso,” J. R. Stat. Soc. **58**, 267-288 (1996).
22. K. Buckley, “Spatial/Spectral filtering with linearly constrained minimum variance beam-formers,” IEEE Trans. Acoust. Speech Signal Proc. **35(3)**, 249-266 (1987).
23. S. C. Huang, C. H. Ma, Y. C. Hsu, and M. R. Bai, “Feedforward active noise global control using a linearly constrained beamforming approach,” J. Sound and Vib. **537**,117190 (2022).
24. J. Kennedy and R. C. Eberhart, “Particle swarm optimization”, *Proceedings of ICNN'95 - International Conference on Neural Networks*, Perth, WA, Australia, 1942-1948 (1995).
25. R. C. Eberhart and J. Kennedy, “A new optimizer using particle swarm theory”, *IEEE Proceedings of the sixth International Symposium on Micro Machine and Human Science*, Nagoya, Japan, 39-43 (1995).
26. D. Campeanu and A. Câmpeanu, “PEAQ—An objective method to assess the perceptual quality of audio compressed files,” in *Proceedings of the International Symposium on System Theory, SINTES 12*, Craiova, România, **12**, 487–492 (2005).
27. P. C. Hansen, D. P. O’Leary, “The use of the L-curve in the regularization of discrete ill-posed problems,” SIAM Journal on Scientific Computing, **14**(6), 1487-1503 (1993).
28. P. R. Johnston and R. M. Gulrajani, "Selecting the corner in the L-curve approach to Tikhonov regularization,"IEEE. Trans. Biomed. Eng. **47**, 1293-1296 (2000).
29. G. F. Edelmann and C. F. Gaumond, “Beamforming using compressive sensing,” J. Acoust. Soc. Am. **130**, EL232–EL237 (2011).
30. M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing* (Springer, New York, 2010), pp. 1–359.
31. M. Grant and S. Boyd, C*VX: Matlab software for disciplined convex programming, version 3.0 beta*, Available: http://cvxr.com/cvx (Last viewed July 5, 2022).
32. J. Allen and D. Berkley, “Image method for efficiently simulating small-room acoustics,” J. Acoust. Soc. Am. **65**, 943–950 (1979).
33. TU-R Recommendation BS.1534-1, “*Method for the subjective assessment of intermediate sound quality (MUSHRA)*” (International Telecommunications Union, Geneva, Switzerland, 2001).
34. A. J. Klockars, G. R. Hancock, and M. J. McAweeney, “Power of unweighted and weighted versions of simultaneous and sequential multiple-comparison procedures,” Psychol. Bull. **118**(2), 300–307 (1995).