Extended zone control using kernel-based weighted acoustic transfer function matching

You Siang Chen, Bo Ru Lai, and Mingsian R. Bai, *Senior Member, IEEE*

[[1]](#footnote-1)

***Abstract*—****To achieve extended personalized zone control, the weighted acoustic transfer function (ATF) matching is proposed. In the practical application of a single-input multiple-output (SIMO) ATF matching paradigm, limited access to the ATF measurements confines effective zone control to the measured points. To overcome the drawback, an optimization problem is formulated as the regional integration of the ATF matching error over the desired control region. Specifically, utilizing kernel interpolation of ATFs, a continuous matching function is represented by the disparity between the controlled and target responses in the bright and dark zone. A zeroth-order spherical Bessel function serves as the reproducing kernel. Then, the closed form solution for the control filter can be obtained through the numerical integration using** **Gaussian quadrature.** **Two strategies for the control filter are demonstrated in light of the matching model designs. Simulations and experiments were conducted using a twelve-loudspeaker array to generate connected bright zone within a designated angle, while concurrently suppressing sound levels in the dark zones. Comparative analysis with baseline approaches indicates that our proposed method effectively broadens the sweet spot of the control zone with limited measurements.**

***Index Terms*—zone control, ATF matching paradigm, weighted ATF matching, kernel interpolation.**

# I. INTRODUCTION

T

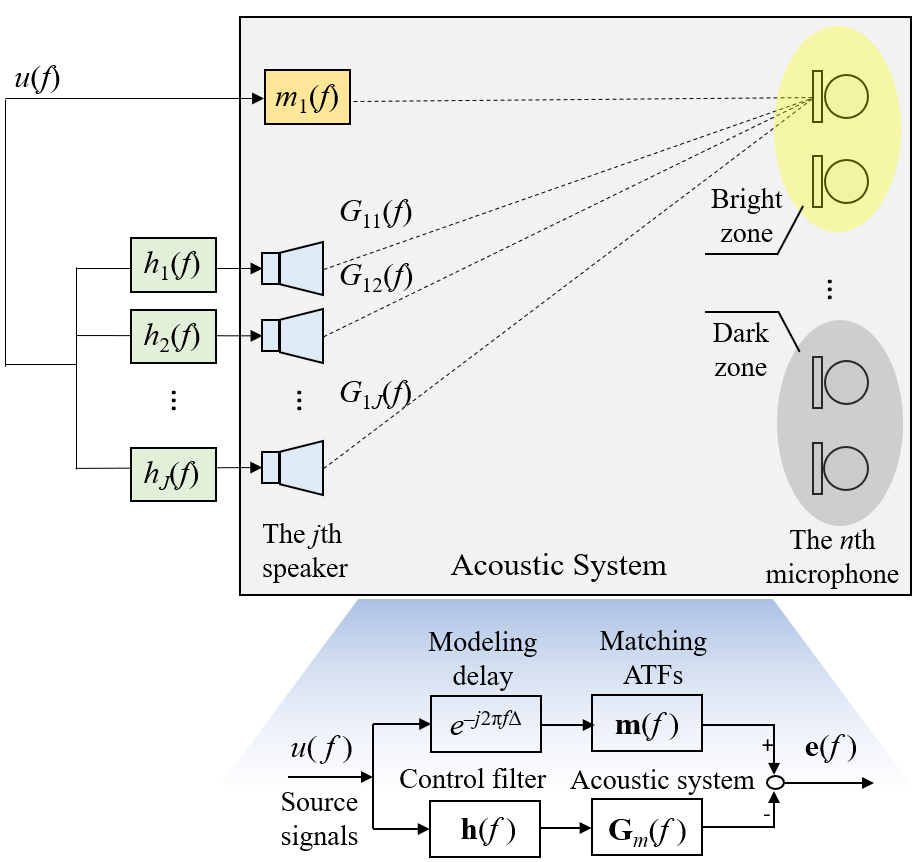
he acoustic zone control technique utilizes an array of loudspeakers to establish personalized audio zones. The primary objective of zone control is to obtain optimal control filters for the loudspeaker array, facilitating the generation of high acoustic energy in bright zones while concurrently suppressing sound levels in dark zones [1] [2]. Such technique finds broad applications across various domains, including soundbars [3], in-vehicle infotainment systems [4] [5] [6] [7], and outdoor audio devices [8] [9]. These systems enable the listeners to individually enjoy specific music or audio content within a defined physical space.Various iconic methodologies have been developed.Acoustic contrast control (ACC) [10] [11] maximizes the energy ratio between bright and dark zones. The ACC problem can be viewed as a more general form based on the principle of generalized eigenvalue decomposition (GEVD) [12] [13]. However, because the phase information is not taken into consideration in the ACC method, the suffers the signal distortion in the bright zone [14]. The process of identifying the eigenvector corresponding to the largest eigenvalue is also a crucial step [15] [16]. Furthermore, the calculation of the matrix inversion in the generalized eigenvalue problem is frequently characterized by high ill-conditioning [17]. In contrast, pressure matching (PM) [18] [19] determines the control filters through the minimization of mean square error (MSE) between the target sound pressure and that generated by the loudspeaker array. This guarantees the control over both magnitude and phase of the sound field. Nonetheless, the acoustic contrast (AC) will be thereby degraded due to the phase control. Hence, a hybrid approach that combines the ACC and PM methods was proposed to strike a balance between the AC and the degree of the phase control [20]. On the other hand, the multizone problem can be also addressed by mode matching (MM) [21] [22] conducts an analysis based on the first order cylindrical harmonic expansions. High order ambisonics (HOA) can effectively improve the spatial resolution by the utilization of higher order spherical harmonics and finds extensive application in multizone reproduction scenarios [23] [24]. However, the requirement of the loudspeaker array geometry for HOA is extremely rigorous. If the loudspeaker array is not uniformly distributed on a sphere facing the target region, the performance will be significantly reduced due to the ill-conditioned matrix inversion [25] [26]. To address the issue, an enhanced least squares method was introduced by incorporating a regularization parameter tailored to the irregular geometry of the array [27]. Moreover, the determination of the loudspeaker number is critical in order to match the spherical harmonics with a specific order [28]. As of the most recent developments, numerous extended studies persist in attempting to overcome the limitations within the modal domain [29] [30]. Alternative methods for approaching zone control involve the minimization of the matching error at system level. Multiple-input/output inverse theorem (MINT) [31] aims to minimize the matching error between the desired impulse responses and the filtered impulse responses. Although the MINT theorem exhibits the capability to acquire the exact solution within an invertible square matrix, it may introduce high-gain filter coefficients. Therefore, an improved method, time-domain underdetermined multi-channel inverse filtering (TUMIF) [32], was proposed as a solution by leveraging the Tikhonov regularization (TIKR) technique [33]. To reduce the computational complexity, TUMIF can be reformulated to frequency-domain underdetermined multi-channel inverse filtering (FUMIF) [34] in the frequency domain. Owing to the rapid advancement of the artificial intelligent, there has been an exploration of optimal filter designs by using deep neural network (DNN) [35] [36].

In practical applications, the aforementioned approaches require discrete measurements of control points within a specific region. The effectiveness of zone control is confined to these measured points due to the limited number of available measurements. While finely arranging control points over the target region can mitigate this problem to some extent, it's crucial to note that the computational cost concurrently escalates with the increment in the number of control points. In addressing the control performance at unmeasured points, weighted pressure matching (WPM) [37] [38] [39] and weighted mode matching (WMM) [40] [41] incorporate the weighting matrix into the optimization of the filters based on the kernel interpolation [42]. In the paradigm of ATF matching, it is also noteworthy that the pursuit of expanded sweet spots in control zones necessitates a substantial quantity of ATFs. To ensure optimal performance at positions where measurements are unavailable, numerous methods utilize interpolation strategies to improve the effective control region. On the basis of plane wave decomposition (PWD) [43] [44], the FUMIF linearly constrained minimum variance (FUMIF-LCMV) approach [34] [45] formulates a constraint optimization problem where the cost function is due to the residual error at the interpolated points, while the constraint equations involve the matching of the ATFs at the measured points. Therefore, the control points encompass both measured and interpolated points, distributed throughout the listening area. However, the two-stage process in FUMIF-LCMV is tedious and time-consuming, especially the application of the least absolute shrinkage and selection operator (LASSO) algorithm [46] in PWD. In addition, the underdetermined constraint restricts the placement of measured control points, which greatly affects the applicable frequency band for the zone control.

Inspired from the principle of the WPM technique, this paper introduces the weighted ATF matching (WAM) approach, providing a twofold contribution as follows. Firstly, under the ATF matching framework, a closed form solution of the optimal filter is derived based on the objective function where the ATF matching error is represented by the regional integral of the kernel functions. Therefore, in contrast to FUMIF-LCMV, the need for a two-stage process that involves interpolation and optimization is consolidated through the integration of kernel function. Furthermore, the proposed method accommodates measured control points without encountering the limitations of the underdetermined constraint. Secondly, because the target functions to be optimized in the unmeasured points are considered through the predefined functions, such as desired ATFs or the kernel functions, the control coefficients can be straightforwardly approximated through the Gaussian quadrature integration [47]. This study presents two strategies for designing the control filter in accordance with the designs of the ATF matching. Consequently, the intricate two-stage process can be simplified to a one-shot design, which significantly enhance the computational efficiency. Additionally, particle swarm optimization (PSO) algorithm [48] is employed for the selection of the two regularization terms in the WAM approach based on the ATF matching errors. Simulations and experiments were carried out using a linear twelve-loudspeaker array, organized into three subsets of four loudspeakers each, to evaluate the efficacy of the underdetermined TIKR, overdetermined TIKR, FUMIF-LCMV, and the proposed WAM approaches. Objective evaluation metrics, the word error rate (WER) and the AC value, were adopted to gauge the effectiveness of the zone control methods.

The subsequent sections of this paper are structured as follows. Section II provides the theoretical foundation for the ATF matching problem and introduces the baseline methods. In Section III, we demonstrate two strategies to the optimal filter designs based on the regional integration of the ATF matching errors within specific control zones. Section IV encompasses simulations and evaluation metrics. The proposed method is compared to the baseline methods. Lastly, the conclusions are provided in Section V.

# II. Theoretical background



**Fig. 1.** The zone control problem can be illustrated as a multichannel ATF matching system, where the control regions of bright and dark zones are predetermined.

As depicted in Fig. 1, the acoustic system can be represented as the framework of a feedforward ATF matching system in the frequency domain, where the ATFs between the loudspeakers and the control points are premeasured. The control region characterized by elevated acoustic energy is referred to as the bright zone, whereas the region with attenuated acoustic energy is designated as the dark zone. In this scenario, the mono input signal *u*(*f*) at frequency *f* undergoes filtration via the control filter labeled as **h**(*f*) = [*h*1(*f*) *h*2(*f*) … *hJ*(*f*)] for a *J*-loudspeaker array. These filtered signals are then reproduced by the measured ATFs of the loudspeaker array, represented as **G***m*(*f*). The frequency responses (FRs) between the loudspeakers and the control zones can be predefined as any arbitrary functions, denoted as **m**(*f*) = [*m*1(*f*) *m*2(*f*) … *mN*(*f*)] for *N* control points. To ensure the causal relationship within the zone control system, the design of the target responses involve a modeling delay *e-j*2π*f*△, where △ represents a time delay in seconds. In other words, to ensure system causality, a modeling delay is introduced to the target functions as **m***m* = *e-j*2π*f*△**m**, where the subscript *m* denotes the measured point. The frequency index *f* will be omitted hereafter for simplicity. Thus, considering *J* loudspeakers and *N* measured control points, the ATF matching problem can be expressed with its components as

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

where *Gnj*, *n =* 1*,…, N*, *j =* 1*,…, J*, denotes the ATF between the *j*th loudspeaker and the *n*th measured control point, *hj* is the *j*th control filter, and *mn* denotes the target FR for the *n*th measured control point. The objective function of the ATF matching problem is to minimize the discrepancy **e** between the desired ATFs, **m***m*, and the processed path, **G***m***h**. To achieve extended zone control region with a limited number of measurements, two baseline approaches are summarized next.

## A. The TIKR approach

In the underdetermined TIKR approach [33], a solution with a perfect matching error is attained by imposing the underdetermined condition where the rank of the FRs matrix **G***m* must be greater than the number of control points *N*. The condition of *J* > *N* guarantees sufficient degrees of freedom in the control system to precisely align the desired ATFs with the processed ATFs. To mitigate the ill-posed problem resulting from the inverse process, the minimization of the *l*2-norm is resolved by incorporating theregularization term *λ* to the solution vector **h**:

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

The optimal solution of the control filter **h***u* can be derived as

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

where "*H*" denotes the matrix conjugate transpose, and **I***J* represents the *J*×*J* identity matrix. The regularization parameter can be chosen using the L-curve method [49]. However, to attain an expanded sweet spot, an intuitive approach is to incorporate the supplementary measured ATFs **G***s* and the corresponding target responses **m***s* over the control region. Thus, the closed-form solution for the optimal filter **h***o* can be similarly derived with

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

in which the number of the measured ATFs *N* is greater than the loudspeaker *J*.

## B. The FUMIF-LCMV approach

The methodology for sound field reproduction, FUMIF-LCMV [34], was developed to extend the effective rendering area encompassing the measured control points. By integrating the interpolated ATFs, the constrained optimization problem is formulated. The primary goal is to minimize the ATF matching error at the interpolated points, while simultaneously ensuring alignment between the rendering path and the target FRs. Hence, the objective function can be expressed as:

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

where the constraint equation remains underdetermined to achieve exact zone control at the measured points. By applying the method of Lagrange multiplier to (5), the optimal solution is obtained as

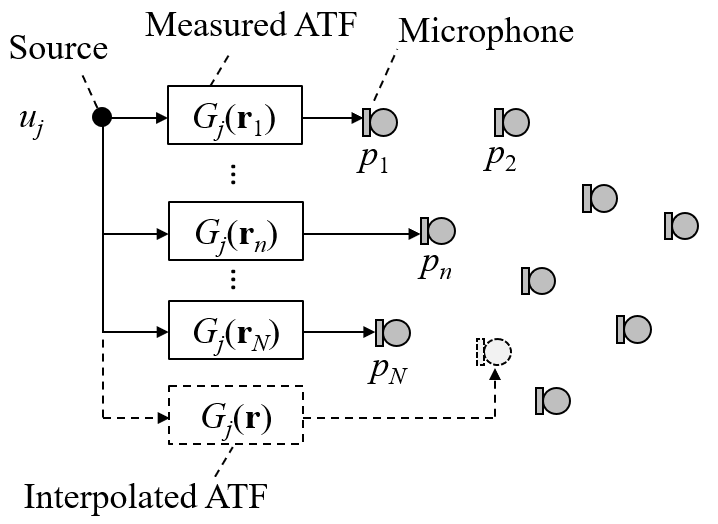
|  |  |
| --- | --- |
|  | (6) |

where *μ* and *ε* are regularization parameters that are inserted to mitigate the ill-posedness during matrix inversion and can be automatically determined by the PSO algorithm [48].

# III. Proposed method

Although the FUMIF-LCMV method can achieve extended zone control, its effectiveness is constrained by the number of measured points. Sparse distribution of control points can lead to performance degradation at interpolated points. To obtain the broadened sweet spots with a limited measurement, we present the WAM approach based on the kernel interpolation method. This method allows flexible deployment of control points and achieves extended zone control by representing ATFs with the kernel interpolation.

## A. ATF interpolation using kernel method



**Fig. 2.** The schematics of the ATF interpolation problem.

Figure 2 illustrates the ATF interpolation problem, where one source and *K* measured points are considered. Let *uj* represent the signals of the *j*th loudspeaker. The sound pressure *pn* captured at the *n*th measured point can be expressed as *pn* = *Gj*(**r***n*)*uj*, where *Gj*(**r***n*) is the ATF between the *j*th loudspeaker and the *n*th measured control point at location **r***n*. The purpose of the ATF interpolation aims to estimate continuous function *Gj*(**r**) within the control region using a finite set of ATF measurements. The subscript *j* of the ATF is ommited hereafter for simplicity. Via the kernel representation, the interpolated ATF can be expressed with a lower-dimensional manifold. Consider an ATF manifold *G*(**r**) as a function of position vector **r** in the Hilbert space The regression problem can be constructed as:

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

where *Ĝ*(**r***n*) and *G*(**r***n*) is the estimated and measured ATF at the *n*th position **r***n*, respectively, and *β* is the regularization parameter. If *κ* is the reproducing kernel that has the form of ATF manifold in the Hilbert space,

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

where <∙, ∙>denotes the inner product in Hilbert space. The solution of the regression problem in (7) has a representation of the form [50]:

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

where *αn*  for *n* = 1, …, *K*. Substituting (8) and (9) to (7), a closed-form solution *Ĝ*(**r**) as the function of the position vector **r** can be derived, as given by

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

where **I***N* is the *N* × *N* identity matrix, **g**= [*G*(**r**1) … *G*(**r***N*)]*T* is the column vector of the ATFs from the *j*th loudspeaker to *N* measured points,

|  |  |
| --- | --- |
|  | (11) |
|  | (12) |

in which *k* is the wave number, and *j*0(∙) = sin(∙)/(∙) denotes the zeroth-order spherical Bessel function [41] [42].

## B. The WAM approach

According to the ATF matching framework, the objective is to align the target ATF *m*(**r**) with the processed ATF at any given position **r** over the control region,

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

The *Gj* denotes the ATF from the *j*th loudspeaker to the control point **r**, while *hj* represents the filter coefficient for the *j*th loudspeaker. To attain a comprehensive range of sweet spots for the zone control, the objective function is redefined in considering the regional integration of ATF matching across the control zones Ω ∈ 3, as follows:

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

Here, the ATFs in (14) can be estimated using the kernel interpolation method in (10), as given by

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

where **g***j* denotes the measured ATF vector between the *j*th loudspeaker and the control points deployed over the control region and **Q** is substituted for the inverse term of the kernel matrix for the simplicity in the subsequent derivation. The objective function can be derived as

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

where *C* is a constant value associated with the control filter **h**.The weighting matrices **W***κ* and **w***m* can be approximated using Gaussian quadrature integration [47], as given by

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

and

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

in which the integration of the (*p*, *q*)th component of the weighting matrix is exact for polynomials of degree 2*T*–1 and *ct* is the weight at the *t*th Gauss node. Thus, a weighted least-square problem can be reformulated with

|  |  |
| --- | --- |
|  | (19) |

where *γ* is the regularization parameter. Then, the closed-from solution for the optimal control filter **h***WAM* can be obtained by

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

Compared to the solution of the conventional TIKR method, the incorporation of the weighting matrices in WAM results in an enhancement of the zone control performance. It is worth noting that **Q** and **W***κ* can be calculated based on the known positions of the measured points **r***n* , *n* = 1, …, *N*, while **w***m* is obtained through the predefined FR function *m*(**r**) within the control region Ω. In contrast to FUMIF-LCMV, the WAM method not only enables comprehensive control over the desired region without the need for explicit ATF interpolation but also allows for the inclusion of measured control points without being constrained by the underdetermined condition.

## C. Parameter optimization using PSO

The PSO algorithm is renowned for its effectiveness in identifying optimal solutions within high-dimensional spaces characterized by complex objective functions. Thus, instead of manually grid searching for the optimal regularization parameters, *β* and *γ* in (20), we apply the PSO method by implementing the following objective function,

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

where the regional integration is also achieved by the Gaussian quadrature. In the optimization, particles are dispersed within a two-dimensional search space where the parameters *β* = 10*a* and *γ* = 10*b* are parameterized as **x** = [*a* *b*]*T* to facilitate the PSO more easily. The velocity vector of the *i*th particle at the *t*th iteration, ***v****i*(*t*), is updated as

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

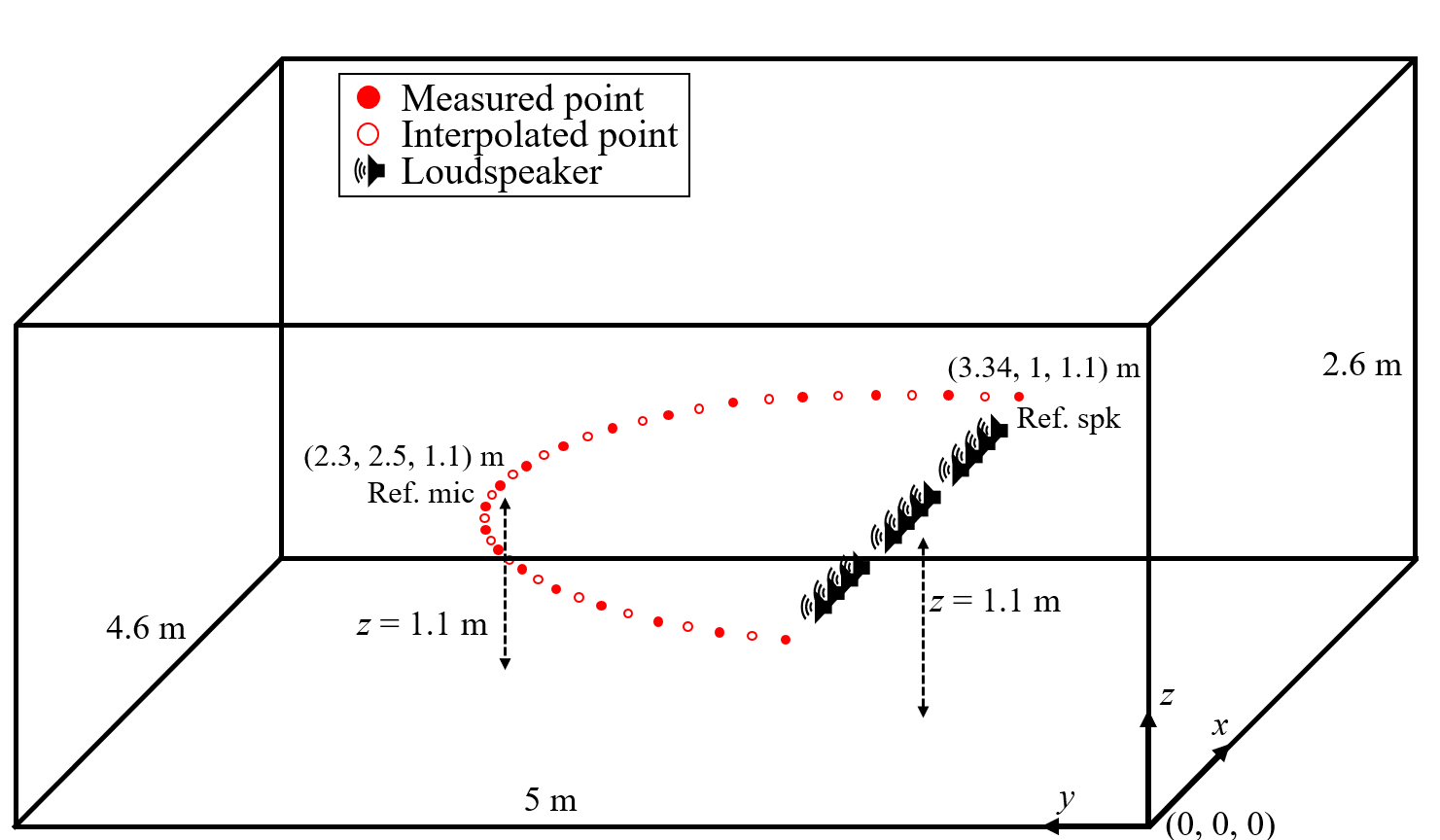
where **x***i*(*t*) represents the position vector of the *i*th particle, while **x***p*(*t*) and **x***g*(*t*) are the position vectors of the particle's personal best and the global best particle, respectively. Moreover, *wI* is the inertia weight, *τ*1 and *τ*2 are the acceleration coefficients, and *η*1*i* and *η*2*i* are random numbers uniformly distributed in the interval [0, 1] for the *i*th particle. The next step for the position vector of the *i*th particle is updated as,

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

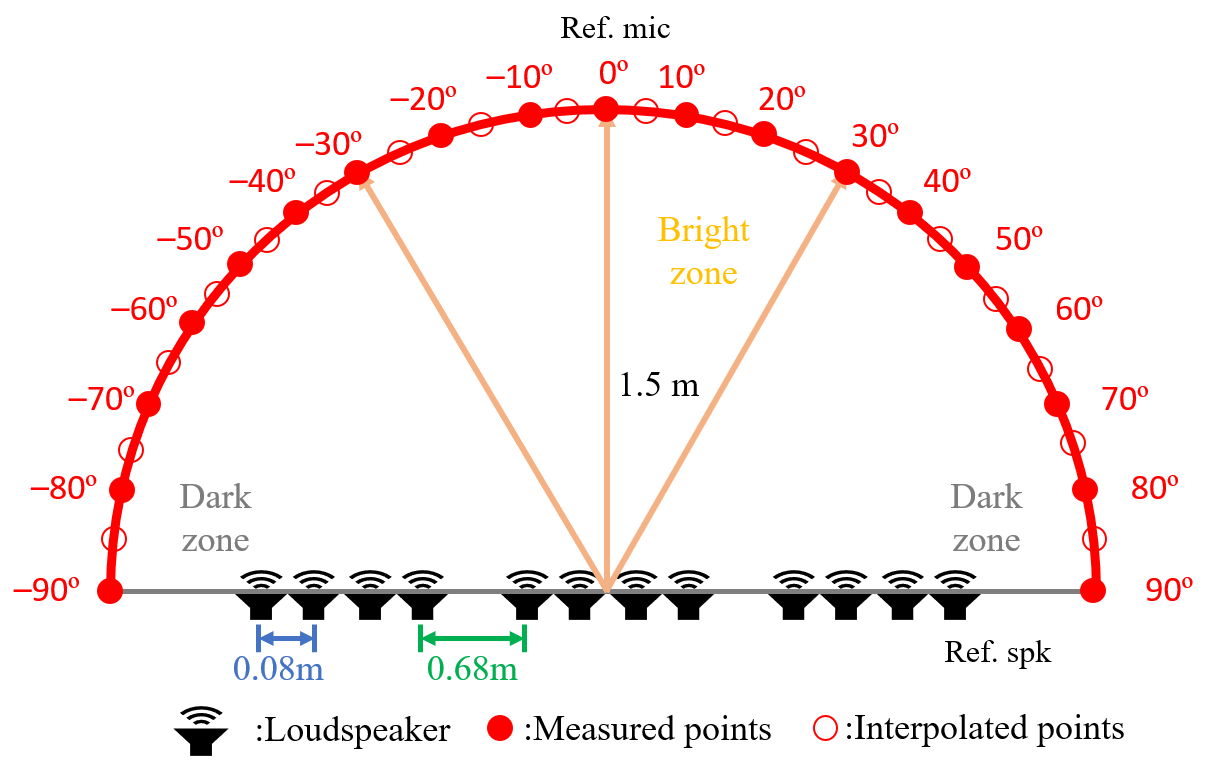
The algorithm terminates when the loss function ceases to decrease after ten consecutive iterations, and all particles converge within a small region of size 0.01× 0.01.

# IV. Numerical Simulation

## A. Simulation setups



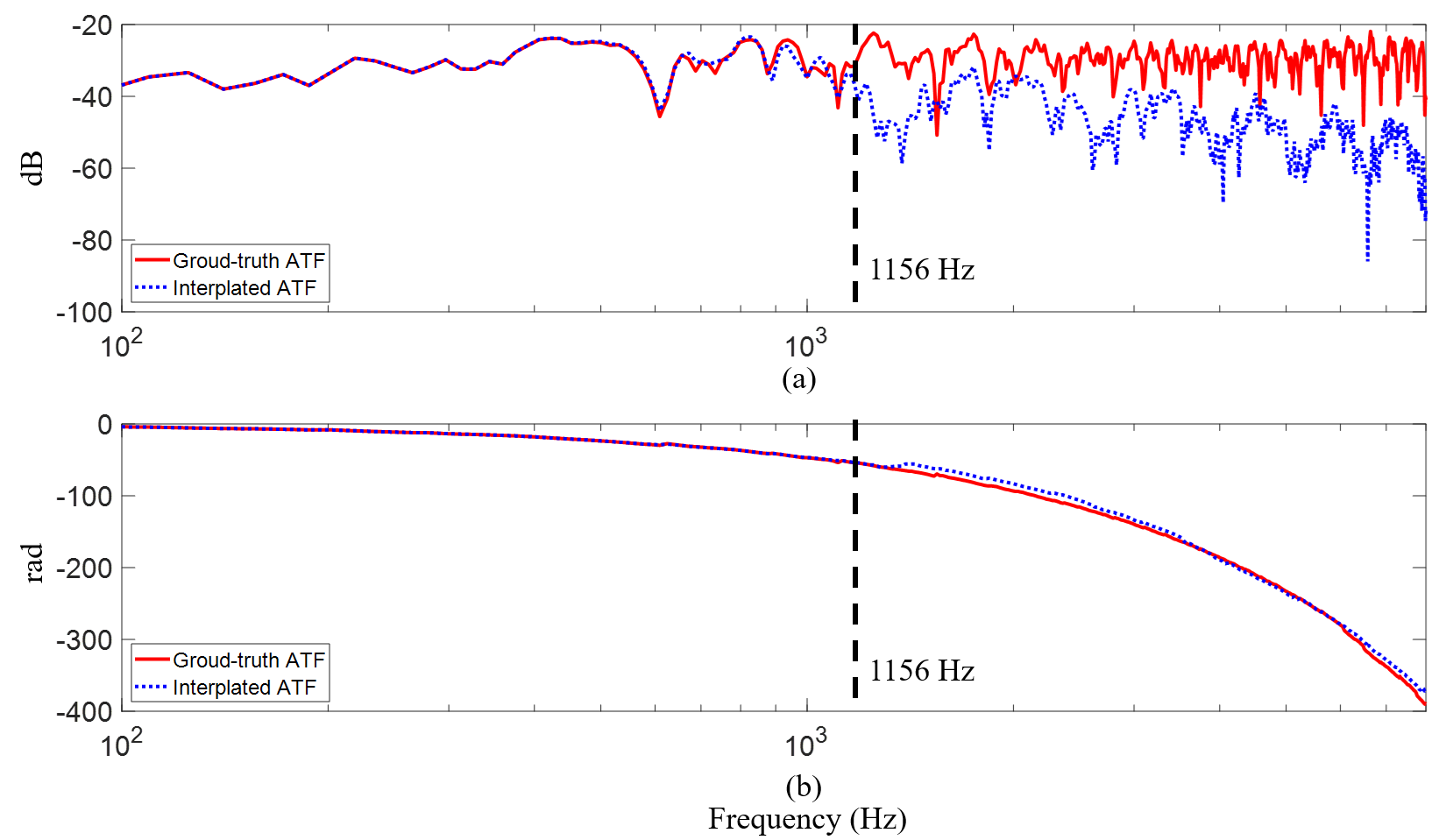
(a)



(b)

**Fig. 3.** (a) Array configuration in a reverberant room (b) deployment of the loudspeaker array, measured points, and interpolated points.

Figure 3(a) depicts the simulation configuration in a reverberant environment with the reverberation time set to be 200 *ms*. The sampling rate of the zone control system is 16 kHz. All authentic room impulse responses (RIRs) are synthesized utilizing the image source method (ISM) [51]. A 4.6 m × 5 m × 2.6 m rectangular room is assumed in the simulation. A loudspeaker array consisting of three sets of four-loudspeaker subarray was employed. The loudspeaker units were deployed with a 0.08 m interelement spacing, while the subarray modules were spaced with 0.68 m intervals, as illustrated in Fig. 3(b). The loudspeakers and measured control points shared the same height of 1.1 m. The reference loudspeaker was positioned at coordinates (3.34, 1, 1.1) m, while the reference microphone was situated at (2.3, 2.5, 1.1) m. The middle point of the loudspeaker array was situated at a distance of 1.5 m from the listening area. Nineteen measurements were conducted along the listening rim with 10° intervals. The bright zone spanned –30° to 30°, while the remaining regions were designated as the dark zones, extending from –30° to –90° and from 30° to 90°. In the baseline underdetermined TIKR approach, three control points were positioned at –30°, 0°, and 30° within the bright zone, while eight control points were evenly distributed with 10° intervals from 40° to 80° and –40° to –80° for the dark zone. The deployment of the control points aligned with the mathematical requirements of the baselines, where the matrix **G***m* must meet an underdetermined condition in the underdetermined TIKR and FUMIF-LCMV methods. For the interpolation of ATFs, eighteen additional ATFs were uniformly interpolated among the all the measured points using (10). As for the overdetermined TIKR method, all nineteen measured points were utilized to form the ATF matrix **G***s* in (4). In contrast to the baseline methods, the proposed WAM approach regards the interpolated points as the continuous function of the kernel regression. Consequently, the optimal filter can be obtained without the need for explicit interpolation procedures. To efficiently acquire the weighting matrices, Gaussian quadrature was applied to approximate the integration across the range from –90° to 90° using 50 Gaussian points. To evaluate interpolation efficacy and determine the frequency range for bandlimited control, Fig. 4(a) and (b) illustrate the magnitude and phase of the ATF between the reference loudspeaker and the interpolated point at 45º, respectively. The solid line represents the ground-truth ATF, while the dot line denotes the ATF interpolated using the kernel method in (10). The result demonstrates an excellent fit below approximately 1156 Hz. Therefore, the frequency range for bandlimited zone control is determined from *fl* = 100 Hz to *fh* = 1156 Hz.



**Fig. 4.** (a) The magnitude and (b) the phase of the ground-truth (solid line) and interpolated ATF (dot line) from the reference loudspeaker and the interpolated point at 45 º.

## B. The designs for the target frequency responses

Because the achievable frequency range could be restricted, the target models are formulated within the bandlimited interval defined by the predetermined frequency range as abovementioned. Two distinct design strategies are offered as follows where we represent the location variable **r** as *θ* for the control region of a rim.

1. Strategy 1 – Bandlimited zone control with equalization

A straightforward approach is to design flat responses utilizing the measured ATFs. The target FR in the bright zone *mB*(*θb*, *f*) is structured as a bandlimited equalization, covering a frequency range from a lower limit *fl* to an upper limit *fh*. The magnitude of the target response at frequency *f* is configured to represent the cumulative acoustic power at a specific control point *θb*. Mathematically, it can be represented as

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

The absolute operation |∙| is employed to signify the absolute value and *Ĝj*(*θb*, *ω*) denotes the estimated ATF between the control point at *θb* and the *j*th loudspeaker using the kernel method. Hence, within the bandlimited frequency interval, the target model showcases a uniform response equivalent to the root-mean-square power of the FR from the loudspeaker array to the control point *θb*. To facilitate acoustic contrast control, the FR *mD*(*θd*, *f*) in the dark zone is defined as

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

with *α* representing a small positive value less than 1 to generate low acoustic power. Furthermore, the design of smooth target responses across the transition between bright and dark zones can be achieved by employing a general sigmoid function relative to the control points *θ*,

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

where the absolute responses of the maximum and minimum values incorporate the designated responses in (21) and (22). Thus, the weighting vector **w***m* in (21) can be derived as

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

where the integration term can be approximated through the Gaussian quadrature in (18).

1. Strategy 2 – Direct bandlimited zone control

As the ATF matching framework is constructed upon SIMO system, the definition of target FRs can be established based on the continuous function of the interpolated ATFs in (10). An alternative design strategy is that the target FR at the control point **r** in the bright zone is the summation of the interpolated ATFs of the loudspeakers, as given by

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

Therefore, the formulation of the target FR can be expressed with the kernel representation as:

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

where **B** = [**Q** … **Q**] is constructed with *J* matrices of **Q** and *vec*() represents the stacked columns of the ATF matrix **G***m*. By substituting (25) into (22) and solving the optimization problem presented in (19), the weighting vector **w***m* in (20) takes the form specified with

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

The integration of the weighting matrices, **W***B* *κ* and **W***D* *κ*, over the bright and dark zones can be determined solely through Gaussian quadrature in (17). In contrast to the first strategy, the closed-form solution of the control filter in (20) can be acquired simply based on the known positions and the vectorized ATF matrix of the measured points. Furthermore, as this strategy preserves the output power without involving the equalization of the control responses, the resulting AC demonstrates improved control performance.

## C. Multi-Channel ATF Matching Zone Control

The bright zone's target model is set to 100, signifying a power level of 1, while the dark zone's target model is established at 10-0.5, representing an approximate power level of 0.316. Utilizing Equation (21), the target model is computed, revealing a power contrast of about 16 dB between the bright and dark zones. To maintain uniform power levels in the bright zone, all bright zone target models are aligned with the model's power at the 0° point. In the underdetermined TIKR approach, the prefilters are calculated using Equation (4), with a regularization parameter of *β*2 TIKR = 10-3.5.The value of *β*2 TIKR is determined using the L-curve [9][10] method. Similarly, in the FUMIF-LCMV approach, the prefilters are obtained using Equation(7), with regularization parameters of *εLCMV* =10-2 and the *μLCMV* = 10-3. These values are determined using the PSO [13][14] method. On the other hand, in the WAM approach, the prefilters are determined using Equation, with the regularization parameters of *λkernel* = 10-3 and the *β*kernel = 10-4. This value is also determined using the PSO method. In Fig. 10, we present the unprocessed and processed FRs for both the bright zone and the dark zone. The sampled points are located at 0° and -60°. The unprocessed FRs exhibit quite similar power levels, particularly in the low-frequency range. In contrast, the processed FRs reveal that all three approaches effectively achieve zone control. The power in the bright zone is maintained, while the power in the dark zone is attenuated. This outcome solidly confirms the success of all three approaches in accomplishing zone control.

驗證指標

Acoustic contrast performance – 參數調整

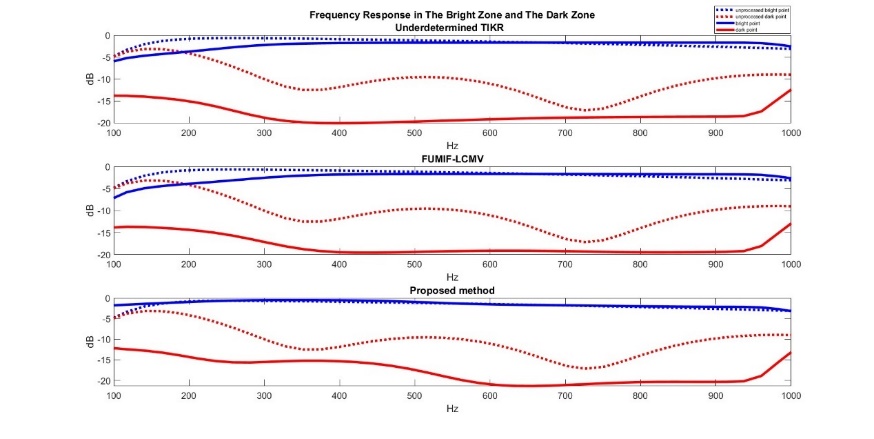
Acoustic contrast performance – Baseline比較

Fig. 9 illustrates the simulation setup in an anechoic environment, employing the image source method (ISM) for generating room impulse responses. This setup involves a twelve-loudspeaker subarray with a 0.075 m interelement spacing, spaced at 0.15 m intervals. The loudspeakers and measured control points share a consistent height of 1.1 m. This specific configuration aims to assess the performance of the three approaches under study. The control focus is a 180° arc line centered around the loudspeaker array, strategically chosen to align with the study's objectives.

The loudspeaker array is situated at a distance of 1 m from the listening area, with measurements taken at 10° intervals. Within this arrangement, the bright zone encompasses -30° to 30°, while the dark zone extends from -50° to -90° and 50° to 90°. Ten measured control points are evenly placed within the dark zone, maintaining 10° intervals within the angular range of 50° to 90° and -50° to -90°. In contrast, control points are positioned at 10° intervals within the angular range of -30° to 30°, designated as bright zone control points specifically for the proposed method. However, only the control point at 0° is selected as the bright zone control point for both the underdetermined TIKR and FUMIF-LCMV approaches. This choice aligns with the approaches' requirements, where the matrix G in underdetermined TIKR or the constraint in FUMIF-LCMV must meet an underdetermined condition. This decision ensures the problem's underdetermined nature persists, allowing precise ATF estimation while adhering to the constraints of underdetermined TIKR and FUMIF-LCMV methods. The interpolation process encompasses all sampled points for ATF estimation at unmeasured points.

The weighted matrix integrates across the range from -90° to 90°, with N points set to 50. In contrast, the interpolated points are carefully chosen using the abscissas obtained from the Gaussian quadrature [17] formula. These interpolated points are then employed in the FUMIF-LCMV method for the intended calculations and assessments. This strategy guarantees the weighted matrix encompasses the entire control region, incorporating an ample number of integration points for an accurate representation of the desired function. The bright zone's target model is set to 100, signifying a power level of 1, while the dark zone's target model is established at 10-0.5, representing an approximate power level of 0.316. Utilizing Equation (3), the target model is computed, revealing a power contrast of about 16 dB between the bright and dark zones. To maintain uniform power levels in the bright zone, all bright zone target models are aligned with the model's power at the 0° point. The simulation operates at a sampling rate of 48 kHz.

We compare the underdetermined TIKR, FUMIF-LCMV, and the WATFM approaches. In the underdetermined TIKR approach, the prefilters are calculated using Equation (4), with a regularization parameter of *β*2 TIKR = 10-3.5.The value of *β*2 TIKR is determined using the L-curve [9][10] method. Similarly, in the FUMIF-LCMV approach, the prefilters are obtained using Equation(7), with regularization parameters of *εLCMV* =10-2 and the *μLCMV* = 10-3. These values are determined using the PSO [13][14] method. On the other hand, in the WATFM approach, the prefilters are determined using Equation ( 24 ), with the regularization parameters of *λkernel* = 10-3 and the *β*kernel = 10-4. This value is also determined using the PSO [13][14] method. In Fig. 10, we present the unprocessed and processed FRs for both the bright zone and the dark zone. The sampled points are located at 0° and -60°. The unprocessed FRs exhibit quite similar power levels, particularly in the low-frequency range. In contrast, the processed FRs reveal that all three approaches effectively achieve zone control. The power in the bright zone is maintained, while the power in the dark zone is attenuated. This outcome solidly confirms the success of all three approaches in accomplishing zone control.



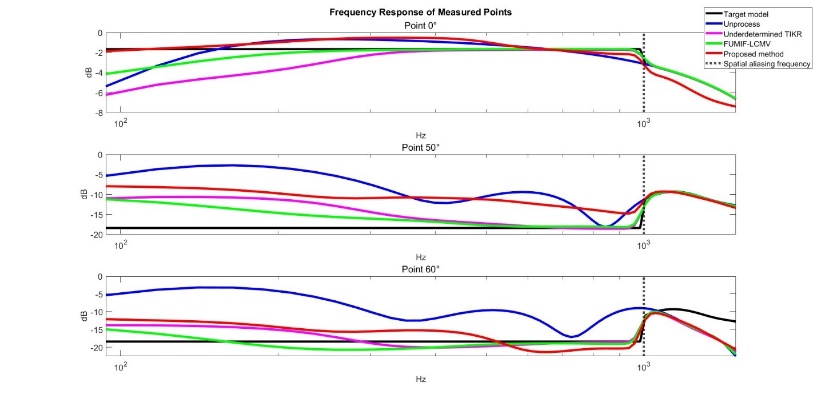
**Fig. 10.** FRs in the bright zone (0°) and dark zone (). The dotted line represents the unprocessed FRs. The solid line represents the processed FRs. The blue line represents the FRs in the bright zone. The red line represents FRs in the dark zone.

In Fig. 11, we present the resulting FRs for both the measured and interpolated control points within the bright and dark zones. The measured control points under observation are situated at 0°, 50°, and 60°, while the interpolated control points are positioned at 15°, 25°, and 55°. A comparison of the FRs in Fig. 11 (a) reveals that the underdetermined TIKR approach achieves a notably closer match to the target response at the measured control points when compared to the FUMIF-LCMV and WATFM approaches. This indicates the underdetermined TIKR approach's superior accuracy in effectively controlling the measured points. This observation affirms the efficacy of the underdetermined TIKR method in accurately regulating the measured control points. However, contrasting this, Fig. 11 (b) presents a distinct outcome. It illustrates that while the underdetermined TIKR approach excels in controlling the FRs at the measured points, its influence over the interpolated points is limited, suggesting potential difficulties in accurately managing these points. In contrast, both the WATFM and FUMIF-LCMV approaches exhibit more accurate model fitting, showcasing their enhanced performance compared to the underdetermined TIKR approach.

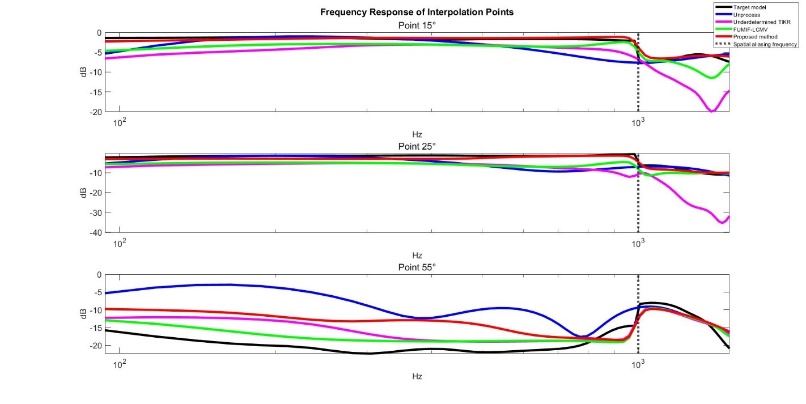
In order to visually represent the rendering performance within the control area, we define a performance metric based on the normalized magnitude matching error as follows:

|  |  |
| --- | --- |
|  | (27) |

where **g***grid*(**r**, *f*) is the FR vector between the loudspeakers and the preselected grid points at *f* Hz and position **r**, **h** is the prefilter vector designed using the underdetermined TIKR, the FUMIF-LCMV or the WATFM approaches, and *mgrid*(**r**, *f*) is the target model at the designated grid point and at the *f* Hz and position **r**. Calculating the average values within the frequency range of 100 Hz to 1000 Hz, which roughly corresponds to the spatial aliasing frequency, serves as our metric. A smaller value of this metric indicates a superior matching performance. A more comprehensive analysis of this performance is illustrated in Fig. 12, where the matching performance along an arc that connects the control points is displayed. This graphical representation offers insights into the proposed approaches' efficacy across the interpolation region. Notably, the underdetermined TIKR approach excels in achieving high matching performance particularly at the measured control points. In contrast, the WATFM approach and FUMIF-LCMV demonstrate a more evenly distributed performance across all control points within the reproduction area, encompassing both measured and interpolated points. Moreover, the proposed method displays a lower normalized magnitude matching error in the bright zone region compared to both the FUMIF-LCMV and underdetermined TIKR approaches. While in the dark zone region, the proposed method achieves comparable performance to the other methods. This suggests that the proposed approach effectively accomplishes accurate magnitude matching within the bright zone and maintains similar performance levels within the dark zone.

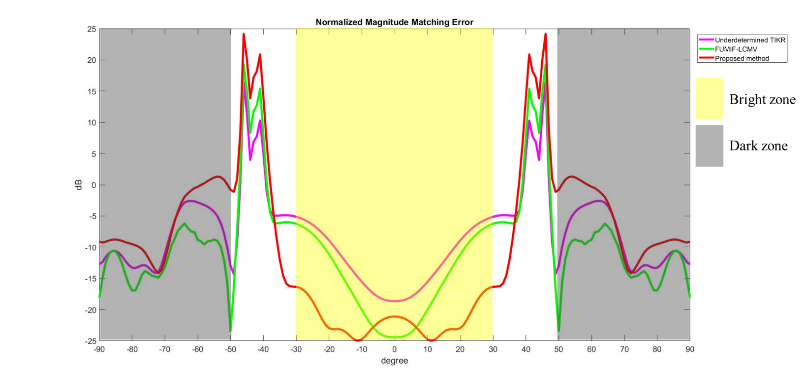


(a)



(b)

**Fig. 11.** FRs of the measured and interpolated control points.



**Fig. 12.** Normalized magnitude matching error.

Observing Fig. 12, we can discern that the normalized magnitude matching error at interpolation points tends to increase in regions where control points are sparse, such as from -50° to 0° for the FUMIF-LCMV and underdetermined TIKR methods, or from -30° to -50° for the WATFM approach. Conversely, in areas with denser control point distribution, interpolation points exhibit favorable matching performance. This emphasizes the significance of control point density and distribution in the effectiveness of interpolation-based methodologies. This highlights the need for careful control point placement, especially in techniques like underdetermined TIKR and FUMIF-LCMV where the number of control points is limited due to the problem's underdetermined nature. In contrast, the proposed method circumvents this constraint, permitting greater control point inclusion and thereby enabling more comprehensive control across the entire interpolation region. This adaptability in control point placement presents a valuable advantage in achieving improved overall matching performance throughout the interpolation area.

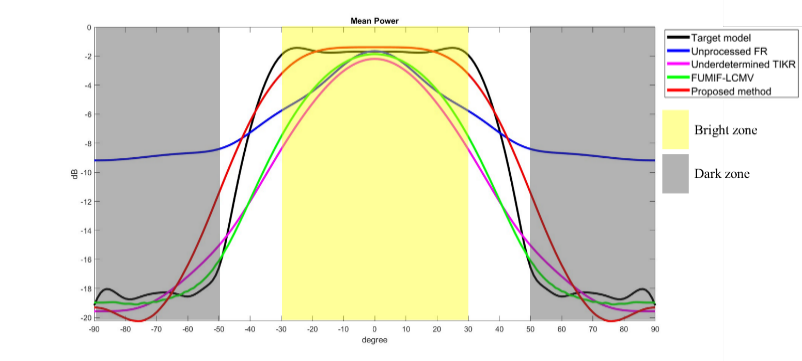
In order to visually represent the rendering performance within the control area, we define a performance metric based on the mean power as follows:

|  |  |
| --- | --- |
|  | (28) |
|  | (29) |

where **g***grid*(**r**, *f*) is the FR vector between the loudspeakers and the preselected grid points at *f* Hz and position **r**, **h** is the prefilter vector designed using the underdetermined TIKR, the FUMIF-LCMV or WATFM approaches. The frequency range spanning from the lower limit *fl* to the upper frequency limit *fh* encompasses 100 Hz to 1000 Hz and is represented in dB scale by the mean power. Concentrating on this frequency range enables us to specifically assess the rendering system's performance in the context of spatial aliasing effects. This metric provides insight into the average power level achieved within the area, offering a measure of the rendering system's effectiveness.

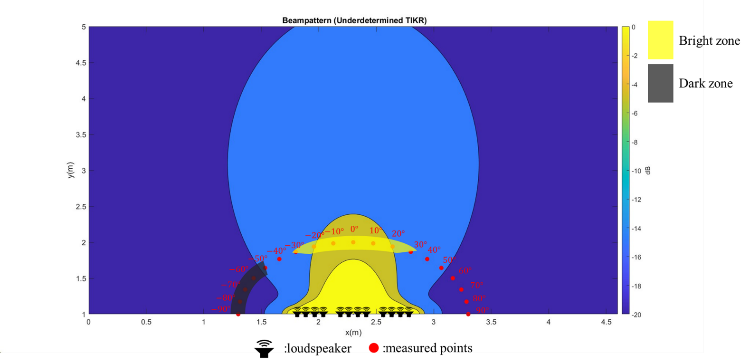
In Fig. 13, it can be observed that the target model in the bright zone maintains the same power as the model at 0°, with a power contrast of approximately 16 dB between the bright zone and the dark zone, as previously mentioned. As illustrated in Fig. 13, the underdetermined TIKR approach demonstrates favorable performance at the measured control points, highlighting its capability to precisely manage those specific points. Conversely, both the WATFM and FUMIF-LCMV approaches exhibit a closer alignment with the model compared to the underdetermined TIKR method. This implies that these methods possess a broader control range, effectively aligning with the desired response across the control region. Nevertheless, among the three techniques, the proposed method achieves the highest performance. This can be attributed to the dense control point distribution, enabling better model maintenance and improved model fitting.

Unlike the underdetermined TIKR and FUMIF-LCMV approaches, the proposed method is not constrained by an underdetermined condition. Consequently, it can incorporate a higher number of control points, leading to superior performance outcomes.

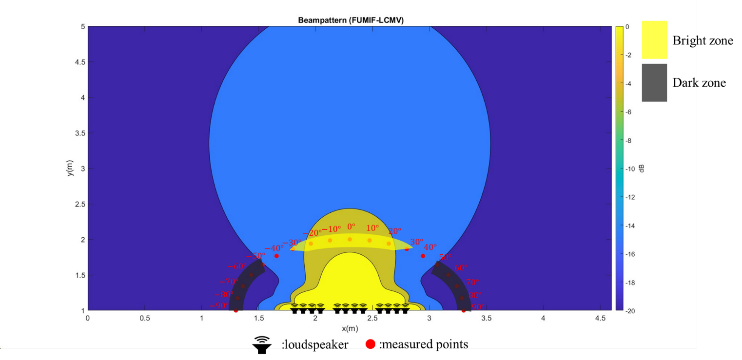


**Fig. 13.** Mean power

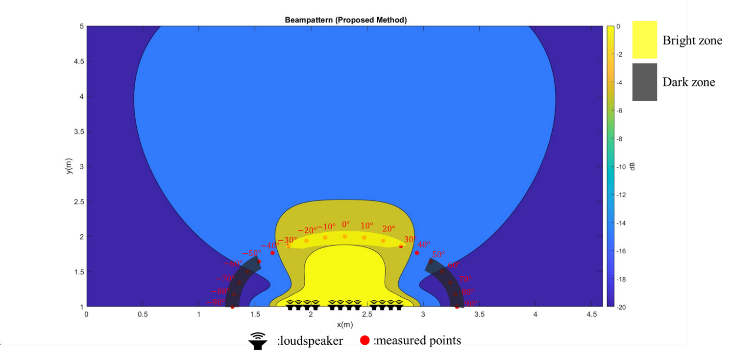
Using Equation (28) and Equation (29), we can derive the beampattern of mean power. The average frequency range remains consistent with the previously mentioned values. As depicted in Fig. 14, the underdetermined TIKR approach effectively governs the measured control points, but its control is not extensive. Conversely, the FUMIF-LCMV approach achieves global control through interpolated points. However, when measured control points are scarce, the global control performance diminishes, evident in the broader bright zone displayed in Fig. 14 (b) compared to the underdetermined TIKR approach. Furthermore, the WATFM approach excels in achieving optimal global control, demonstrated by the bright zone in Fig. 14 (c). This accomplishment can be attributed to the utilization of a greater number of measured control points.



(a)



(b)



(c)

**Fig. 14.** Beampattern of three approaches

To analyze the performance of zone control using objective indicators, we define the acoustic contrast (AC) performance metric as follows:

|  |  |
| --- | --- |
|  | (30) |

where *NBright* is the number of samples in the bright zone, *NDark* is the number of samples in the dark zone, the *P*(**r***Bright, n*) represents the mean power at the sample points within the bright zone, *P*(**r***Dark, n*) represents the mean power at the sample points within the dark zone. The acoustic contrast (AC) metric quantifies the power contrast between the two zones, expressed in decibels (dB). A higher AC value indicates more pronounced differentiation and control of the sound field between the bright and dark zones, signifying better performance in achieving the desired acoustic contrast. In the simulation, we sample the entire region at 5° intervals. The bright zone encompasses -30° to 30°, while the dark zone includes the combined range of -50° to -90° and 50° to 90°, as previously indicated.

A more comprehensive AC performance analysis is presented in Table 1. The results in Table 1 follow the trends observed earlier. Both the WATFM approach and the FUMIF-LCMV approach outperform the underdetermined TIKR method. This superiority stems from the fact that both methods incorporate interpolation points during filter design, thus enhancing performance at those points. Additionally, the proposed method, which permits a larger number of control points and operates without the constraint of the underdetermined condition, surpasses FUMIF-LCMV in terms of performance. These results in Table I underscore the efficacy of the proposed method in achieving precise control within the designated region and achieving global control.

TABLE I

This is a Sample of a Table Title

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Unprocessed | Underdetermined TIKR | FUMIF-LCMV | WATFM |
| AC (dB) | 5.35 | 13.54 | 14.59 | 14.64 |

To assess the voice performance of zone control using objective indicators, we employ the average word error rate (WER). The average WER can be calculated using the following formula:

|  |  |
| --- | --- |
|  | (31) |

where *S* is the number of substitutions, *D* is the number of deletions, *I* is the number of insertions, *Nw* is the number of words in the reference and *N* is the number of sample points. A lower average WER signifies enhanced speech clarity, while a higher average WER implies diminished clarity or reduced intelligibility. We also introduced the WER contrast metric as:

|  |  |
| --- | --- |
|  | (32) |

A higher WER contrast value signifies a clearer differentiation and control of the sound field between the bright and dark zones, showcasing enhanced effectiveness in achieving the intended acoustic contrast. During the simulation, the entire region was sampled at 5° intervals. The bright zone covers -30° to 30°, while the dark zone encompasses the ranges of -50° to -90° and 50° to 90°, as mentioned earlier. All ASR experiments were executed using the SpeechBrain toolkit [33], with test speech selected from the LibriSpeech dataset [34]. A more comprehensive AC performance analysis is presented in Table 2, where the results trend in accordance with the AC values. The FUMIF-LCMV approach attains better control in the voice zone through its consideration of interpolated points. Conversely, the WATFM approach displays superior performance, as evidenced by the AC values. Both Table 1 and Table 2 confirm that the proposed method, which incorporates more measured control points, achieves exceptional global zone control.

TABLE II

This is a Sample of a Table Title

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Unprocessed | Underdetermined TIKR | FUMIF-LCMV | WATFM |
| WERaverage,bright | 0 % | 0 % | 28 % | 0 % |
| WERaverage, dark | 17 % | 28 % | 94 % | 77 % |
| WERcontrast | 17 % | 28 % | 66 % | 77 % |

V. Conclusion

This thesis introduces multi-channel model matching approaches for achieving both global zone control and acoustic contrast control. In the simulation, the WATFM approach exhibited the highest performance, yielding an AC value of 14.64 dB. Furthermore, in terms of the WER contrast value, the WATFM approach demonstrated superior performance with a value of 77%. This advantage stems from the fact that the WATFM approach is not constrained by the underdetermined condition. Conversely, the FUMIF-LCMV approach achieved a simulated AC value of 14.59 dB and a WER contrast value of 66%, showcasing better global zone control performance than the underdetermined TIKR approach. The latter achieved a simulated AC value of 13.54 dB and a WER contrast value of 28%. The FUMIF-LCMV approach's improved performance is attributed to its consideration of interpolated points. However, due to the continued constraint of the underdetermined condition, the FUMIF-LCMV approach's global zone control performance remains relatively inferior to that of the WATFM approach. These trends remained consistent in the experimental results. The WATFM approach maintained its superiority with an AC value of 6.33 dB. Conversely, the AC values from the FUMIF-LCMV approach (6.15 dB) and the underdetermined TIKR approach (5.39 dB) reinforced the same conclusion observed in the simulation.

# Appendix

The cost function of the weighted acoustic transfer function matching approach can be written as

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

On the basis of the kernel interpolation method in section 4.3.1, the interpolated ATFs and interpolated model can be written as

|  |  |
| --- | --- |
|  | (A2) |
|  | (A3) |

Then, the cost function *J* can be approximated as

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

where **W** is defined as

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

with

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

The optimization problem of WATFM approach is formulated using the (A4) as:

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

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

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

Solving (A6) for the optimal prefilter vector leads to

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

The regularization optimization in (A7) is required to avoid the singularity of the matrix inverse.

# Acknowledgment

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

References

1. W. F. Druyvesteyn and J. Garas, “Personal sound,” J. Audio Eng. Soc., vol. 45, no. 9, pp. 685–701, Sep. 1997.
2. T. Betlehem, W. Zhang, M. A. Poletti, and T. D. Abhayapala, “Personal sound zones: Delivering interface-free audio to multiple listeners,” IEEE Signal Process. Mag., vol. 32, no. 2, pp. 81–91, Mar. 2015.
3. Gálvez, M. F. S., Elliott, S. J., & Cheer, J. (2014). Personal audio loudspeaker array as a complementary TV sound system for the hard of hearing. IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, 97(9), 1824-1831.
4. So, H., & Choi, J. W. (2019, May). Subband optimization and filtering technique for practical personal audio systems. In ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 8494-8498). IEEE.
5. J. Cheer and S. Elliott, “Design and implementation of a personal audio system in a car cabin,” Proceedings of Meetings on Acoustics, vol. 19, no. 1, p. 055009, 2013.
6. Liao, X., Cheer, J., Elliott, S. J., & Zheng, S. (2017). Design of a loudspeaker array for personal audio in a car cabin. Journal of the Audio Engineering Society, 65(3), 226-238.
7. Ma, J., Zhang, Q., Hu, Z., & Liu, D. (2022, June). Situating Audio Zone: In-Vehicle Personal Audio Zone Use in Context. In International Conference on Human-Computer Interaction (pp. 487-498). Cham: Springer Nature Switzerland.
8. Heuchel, F. M., Caviedes-Nozal, D., Brunskog, J., Agerkvist, F. T., & Fernandez-Grande, E. (2020). Large-scale outdoor sound field control. The Journal of the Acoustical Society of America, 148(4), 2392-2402.
9. J. Brunskog et al., “Full-scale outdoor concert adaptive sound field control,” in Proc. 23rd Int. Congr. Acoust. Aachen, Germany: German Acoustical Society (DEGA), Sep. 2019, pp. 1170–1177.
10. J.-W. Choi and Y.-H. Kim, “Generation of an acoustically bright zone with an illuminated region using multiple sources,” *J. Acoust. Soc. Amer.*, vol. 111, no. 4, pp. 1695-1700, 2002.
11. P. Coleman, P. J. Jackson, M. Olik, M. B. Møller, M. Olsen, and J. A. Pederson, “Acoustic contrast, planarity and robustness of sound zone methods using a circular loudspeaker array,” J. Acoustical Soc. Amer., vol. 135, no. 4, pp. 1929–1940, 2014.
12. T. Lee, J. K. Nielsen, J. R. Jensen, and M. G. Christensen, “A unified approach to generating sound zones using variable span linear filters,” in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process., 2018, pp. 491–495.
13. L. Shi, T. Lee, L. Zhang, J. K. Nielsen, and M. G. Christensen, “A fast reduced-rank sound zone control algorithm using the conjugate gradient method,” in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process., 2020, pp. 436–440.
14. P. Coleman, P. J. B. Jackson, M. Olik, and J. A. Pedersen, “Personal audio with a planar bright zone,” J. Acoust. Soc. Amer., vol. 136, no. 4, pp. 1725–1735, Oct. 2014.
15. J.-H. Chang and F. Jacobsen, “Sound field control with a circular doublelayer array of loudspeakers,” J. Acoust. Soc. Amer., vol. 131, no. 6, pp. 4518–4525, Jun. 2012.
16. A. H. Andersen, “On multiple sound zones for wideband signals,” Master’s thesis, Aalborg Univ., Aalborg, Denmark, Jan. 2014.
17. S. J. Elliott, J. Cheer, J.-W. Choi, and Y. Kim, “Robustness and regularization of personal audio systems,” IEEE Trans. Audio, Speech, Lang. Process., vol. 20, no. 7, pp. 2123–2133, Sep. 2012.
18. Poletti, M. (2008, October). An investigation of 2-D multizone surround sound systems. In Audio Engineering Society Convention 125. Audio Engineering Society.
19. N. Radmanesh and I. S. Burnett, “Generation of isolated wideband sound field using a combined two-stage Lasso-LS algorithm,” IEEE Trans. Audio, Speech, Lang. Process., vol. 21, no. 2, pp. 378–387, Feb. 2013
20. M. B. Møller, M. Olsen, and F. Jacobsen, “A hybrid method combining synthesis of a sound field and control of acoustic contrast,” in Proc. 132nd Conv. Audio Eng. Soc., Budapest, Hungary, Apr. 2012, pp. 1–8.
21. Y. J. Wu and T. D. Abhayapala, "Spatial Multizone Soundfield Reproduction: Theory and Design," in IEEE Transactions on Audio, Speech, and Language Processing, vol. 19, no. 6, pp. 1711-1720, Aug. 2011, doi: 10.1109/TASL.2010.2097249.
22. J. Daniel, S. Moureau, and R. Nicol, “Further Investigations of High-Order Ambisonics and Wavefield Synthesis for Holophonic Sound Imaging,” presented at the Proc. 114th AES Conv. (2003 Mar.).
23. W. Jin and W. B. Kleijn, “Theory and design of multizone soundfield reproduction using sparse methods,” IEEE/ACM Trans. Audio, Speech, Lang. Process., vol. 23, no. 12, pp. 2343–2355, Dec. 2015.
24. W. Zhang, T. D. Abhayapala, T. Betlehem, and F. M. Fazi, “Analysis and control of multi-zone sound field reproduction using modal-domain approach,” J. Acoust. Soc. Amer., vol. 140, no. 3, pp. 2134–2144, Sep. 2016.
25. Y. J. Wu and T. D. Abhayapala, “Theory and design of soundfield reproduction using continuous loudspeaker concept,” IEEE Trans. Audio, Speech, Lang. Process., vol. 17, no. 1, pp. 107–116, Dec. 2008.
26. M. Poletti, F. Fazi, and P. Nelson, “Sound-field reproduction systems using fixed-directivity loudspeakers,” J. Acoust. Soc. Amer., vol. 127, no. 6, pp. 3590–3601, Apr. 2010.
27. M. Poletti, “Robust two-dimensional surround sound reproduction for nonuniform loudspeaker layouts,” J. Audio Eng. Soc., vol. 55, no. 7/8, pp. 598–610, Jul. 2007.
28. D. B. Ward and T. D. Abhayapala, “Reproduction of a plane-wave sound field using an array of loudspeakers,” IEEE Trans. Speech Audio Process., vol. 9, no. 6, pp. 697–707, Sep. 2001.
29. J. Zhang, W. Zhang, T. D. Abhayapala, and L. Zhang, “2.5D multizone reproduction using weighted mode matching: Performance analysis and experimental validation,” J. Acoustical Soc. Amer., vol. 147, no. 3, pp. 1404–1417, 2020.
30. H. Zuo, T. D. Abhayapala, and P. N. Samarasinghe, “Particle velocity assisted three dimensional sound field reproduction using amodal-domain approach,” IEEE/ACM Trans. Audio, Speech, Lang. Process., vol. 28, pp. 2119–2133, 2020.
31. M. Miyoshi and Y. Kaneda, "Inverse filtering of room acoustics", IEEE Transactions on Acoustics Speech and Signal Processing, vol. 36, pp. 145-152, 1988.
32. 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. Amer., vol. 146, no. 2, pp. 1302-1313, Aug. 2019.
33. C. W. Groetsch, "The theory of Tikhonov regularization for Fredholm equations of the first kind" in Research Notes in Mathematics, MA, Boston:Pitman, vol. 105, 1984.
34. W.-L. Lin, Y.-S. Chen, B.-R. Lai, and M. R. Bai, "Multichannel room response equalization with a broadened control region using a linearly constrained approach and sensor interpolation," J. Acoust. Soc. Amer., vol. 153, no. 4, pp. 1960, Apr. 2023.
35. G. Pepe, L. Gabrielli, S. Squartini, L. Cattani and C. Tripodi, "Deep Learning for Individual Listening Zone," 2020 IEEE 22nd International Workshop on Multimedia Signal Processing (MMSP), Tampere, Finland, 2020, pp. 1-6, doi: 10.1109/MMSP48831.2020.9287161.
36. G. Pepe, L. Gabrielli, S. Squartini, C. Tripodi and N. Strozzi, "Digital Filters Design for Personal Sound Zones: a Neural Approach," 2022 International Joint Conference on Neural Networks (IJCNN), Padua, Italy, 2022, pp. 1-8, doi: 10.1109/IJCNN55064.2022.9892571.
37. S. Koyama and K. Arikawa, “Weighted Pressure Matching Based on Kernel Interpolation for Sound Field Reproduction,” presented at the Proc. Int. Congr. Acoust. (ICA) (2022 Oct.).
38. F. Olivieri, F. M. Fazi, S. Fontana, D. Menzies, and P. A. Nelson, “Generation of private sound with a circular loudspeaker array and the weighted pressure matching method,” IEEE/ACM Trans. Audio, Speech, Lang. Process., vol. 25, no. 8, pp. 1579–1591, Aug. 2017.
39. V. Molés-Cases, S. J. Elliott, J. Cheer, G. Piñero, and A. Gonzalez, “Weighted pressure matching with windowed targets for personal sound zones,” J. Acoustical Soc. Amer., vol. 151, no. 1, pp. 334–345, 2022.
40. Ueno, N., Koyama, S., & Saruwatari, H. (2019). Three-dimensional sound field reproduction based on weighted mode-matching method. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 27(12), 1852-1867.
41. Koyama, S., Kimura, K., & Ueno, N. (2023). Weighted Pressure and Mode Matching for Sound Field Reproduction: Theoretical and Experimental Comparisons. arXiv preprint arXiv:2303.13027.
42. N. Ueno, S. Koyama, and H. Saruwatari, “Kernel ridge regression with constraint of Helmholtz equation for sound field interpolation,” IEEE International Workshop on Acoustic Signal Enhancement (IWAENC), pp. 1-440, 2018.
43. 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)
44. 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).
45. 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).
46. R. Tibshirani, “Regression Shrinkage and Selection via the Lasso,” J. R. Stat. Soc. 58, 267-288 (1996).
47. W. H. Press, B. P. Flannery, S. A. Teukolsky and W. T. Vetterling, Numerical Recipes, New York, NY, USA:Cambridge Univ. Press, 1986.
48. J. Kennedy and R. Eberhart, "Particle swarm optimization", Proc. ICNN'95–Int. Conf. Neural Networks, pp. 1942-1948, 1995.
49. P. R. Johnston and R. M. Gulrajani, "Selecting the corner in the L-curve approach to Tikhonov regularization", IEEE Trans. Biomed. Eng., vol. 47, no. 9, pp. 1293-1296, 2000.
50. B. Schölkopf, R. Herbrich, and A. J. Smola, *A generalized representer theorem*, Springer, Berlin, 2001.
51. J. B. Allen and D. A. Berkley, "Image method for efficiently simulation small-room acoustics", Journal of Acoustical Society of America, vol. 64, no. 4, pp. 943-950, 1979.

*Basic format for periodicals:*

J. K. Author, “Name of paper,” *Abbrev. Title of Periodical*, vol. x, no. x, pp. xxx-xxx, Abbrev. Month, year, doi: 10.1109.XXX.1234567.

*Periodicals using article numbers:*

J. K. Author, “Name of paper,” *Abbrev. Title of Periodical*, vol. x, no. x, Abbrev. Month, year, Art. no. xxxxx, doi: 10.1109.XXX.1234567.

*Basic format for books:*

J. K. Author, “Title of chapter in the book,” in *Title of Published Book, x*th ed. City of Publisher, (only U.S. State), Country: Abbrev. of Publisher, year, ch. x, sec. *x*, pp. xxx–xxx*.*

*Examples:*

1. G. O. Young, “Synthetic structure of industrial plastics,” in *Plastics,* 2nd ed., vol. 3, J. Peters, Ed. New York, NY, USA: McGraw-Hill, 1964, pp. 15–64.
2. W.-K. Chen, *Linear Networks and Systems.* Belmont, CA, USA: Wadsworth, 1993, pp. 123–135.
3. Philip B. Kurland and Ralph Lerner, eds., *The Founders’ Constitution.* Chicago, IL, USA: Univ. of Chicago Press, 1987, Accessed on: Feb. 28, 2010, [Online]. Available: http://press-pubs.uchicago.edu/founders/

*Basic format for handbooks:*

*Name of Manual/Handbook, x* ed., Abbrev. Name of Co., City of Co., Abbrev. State, Country, year, pp. xxx-xxx.

*Examples:*

1. *Transmission Systems for Communications*, 3rd ed., Western Electric Co., Winston-Salem, NC, USA, 1985, pp. 44–60.
2. *Motorola Semiconductor Data Manual*, Motorola Semiconductor Products Inc., Phoenix, AZ, USA, 1989.
3. R. J. Hijmans and J. van Etten, “Raster: Geographic analysis and modeling with raster data,” R Package Version 2.0-12, Jan. 12, 2012. [Online]. Available: http://CRAN.R-project.org/package=raster

*Basic format for reports:*

J. K. Author, “Title of report,” Abbrev. Name of Co., City of Co., Abbrev. State, Country, Rep. xxx, year.

*Example:*

1. E. E. Reber, R. L. Michell, and C. J. Carter, “Oxygen absorption in the earth’s atmosphere,” Aerospace Corp., Los Angeles, CA, USA, Tech. Rep. TR-0200 (4230-46)-3, Nov. 1988.

*Basic format for conference proceedings:*

J. K. Author, “Title of paper,” in *Abbreviated Name of Conf.*, City of Conf., Abbrev. State (if given), Country, year, pp. xxxxxx*.*

*Examples:*

1. D. B. Payne and J. R. Stern, “Wavelength-switched passively coupled single-mode optical network,” in *Proc. IOOC-ECOC,* Boston, MA, USA,1985, pp. 585–590.
2. D. Ebehard and E. Voges, “Digital single sideband detection for interferometric sensors,” presented at the 2nd Int. Conf. Optical Fiber Sensors*,* Stuttgart, Germany, Jan. 2-5, 1984.
3. PROCESS Corporation, Boston, MA, USA. Intranets: Internet technologies deployed behind the firewall for corporate productivity. Presented at INET96 Annual Meeting. [Online]. Available: http://home.process.com/Intranets/wp2.htp

*Basic format for electronic documents (when available online):*

Issuing Organization. (year, month day). *Title*. [Type of medium]. Available: site/path/file

*Example:*

1. U.S. House. 102nd Congress, 1st Session. (1991, Jan. 11). *H. Con. Res. 1, Sense of the Congress on Approval of Military Action*. [Online]. Available: LEXIS Library: GENFED File: BILLS

*Basic format for patents:*

J. K. Author, “Title of patent,” U.S. Patent *x xxx xxx*, Abbrev. Month, day, year.

*Example:*

1. G. Brandli and M. Dick, “Alternating current fed power supply,” U.S. Patent 4 084 217, Nov. 4, 1978.

*Basic format**for theses (M.S.) and dissertations (Ph.D.):*

J. K. Author, “Title of thesis,” M.S. thesis, Abbrev. Dept., Abbrev. Univ., City of Univ., Abbrev. State, year.

J. K. Author, “Title of dissertation,” Ph.D. dissertation, Abbrev. Dept., Abbrev. Univ., City of Univ., Abbrev. State, year.

*Examples:*

1. J. O. Williams, “Narrow-band analyzer,” Ph.D. dissertation, Dept. Elect. Eng., Harvard Univ., Cambridge, MA, USA, 1993.
2. N. Kawasaki, “Parametric study of thermal and chemical nonequilibrium nozzle flow,” M.S. thesis, Dept. Electron. Eng., Osaka Univ., Osaka, Japan, 1993.

*Basic format for the most common types of unpublished references:*

J. K. Author, private communication, Abbrev. Month, year.

J. K. Author, “Title of paper,” unpublished.

J. K. Author, “Title of paper,” to be published.

*Examples:*

1. A. Harrison, private communication, May 1995.
2. B. Smith, “An approach to graphs of linear forms,” 2014, *arXiv:2105.02824*.
3. A. Brahms, “Representation error for real numbers in binary computer arithmetic,” IEEE Computer Group Repository, Paper R-67-85.

*Basic formats for standards:*

a) *Title of Standard*, Standard number, date.

b) *Title of Standard*, Standard number, Corporate author, location, date.

*Examples:*

1. IEEE Criteria for Class IE Electric Systems, IEEE Standard 308, 1969.
2. Letter Symbols for Quantities, ANSI Standard Y10.5-1968.

*Basic format for datasets:*

Author,  Date, Year. “Title of Dataset,” distributed by Publisher/Distributor, http://url.com (or if DOI is used, end with a period)

*Example:*

1. U.S. Department of Health and Human Services, Aug. 2013, “Treatment Episode Dataset: Discharges (TEDS-D): Concatenated, 2006 to 2009,” U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Office of Applied Studies, doi: 10.3886/ICPSR30122.v2.

*Basic format for code:*

Author,  Date published or disseminated, Year. “Complete title, including ed./vers.#,” distributed by Publisher/Distributor, http://url.com (or if DOI is used, end with a period)

*Example:*

1. T. D’Martin and S. Soares, 2019, “Code for Assessment of Markov Decision Processes in Long-Term Hydrothermal Scheduling of Single-Reservoir Systems (Version 1.0),” Code Ocean, doi: \_1.24433/CO.7212286.v1

**First A. Author** (Fellow, IEEE) and all authors may include biographies if the publication allows. Biographies are often not included in conference-related papers. Please check the Information for Authors to confirm. Author photos should be current, professional images of the head and shoulders. The first paragraph may contain a place and/or date of birth (list place, then date). Next, the author’s educational background is listed. The degrees should be listed with the type of degree in what field, which institution, city, state, and country, and year the degree was earned. The author’s major field of study should be lowercase.

The second paragraph uses the preferred third person pronoun (he, she, they, etc.) and not the author’s last name. It lists military and work experience, including summer and fellowship jobs. Job titles are capitalized. The current job must have a location; previous positions may be listed without one. Information concerning previous publications may be included. The format for listing publishers of a book within the biography is: *Title of Book* (publisher name, year) similar to a reference. Current and previous research interests end the paragraph.

The third paragraph begins with the author’s preferred title and last name (e.g., Dr. Smith, Prof. Jones, Mr. Kajor, Ms. Hunter, Mx. Riley). List any memberships in professional societies other than the IEEE. Finally, list any awards and work for IEEE committees and publications.

**Second B. Author**, photograph and biography not available at the time of publication.

**Third C. Author Jr.** (Member, IEEE), photograph and biography not available at the time of publication.

1. Mr. You Siang Chen, Bo Ru Lai, and Professor Mingsian R. Bai are with the Power Mechanical Engineering Department, National Tsing Hua University, Taiwan, ROC, e-mail: [s108033851@m108.nthu.edu.tw](mailto:s108033851@m108.nthu.edu.tw), [yles.94322@gmail.com](mailto:yles.94322@gmail.com), and msbai@pme.nthu.edu.tw. [↑](#footnote-ref-1)