

The Journal of the Acoustical Society of America

A comparison between two modal domain methods for personal audio reproduction

--Manuscript Draft--

Manuscript Number:	JASA-04107R2
Full Title:	A comparison between two modal domain methods for personal audio reproduction
Short Title:	Modal domain personal audio reproduction comparison
Article Type:	Regular Article
Corresponding Author:	Qiaoxi Zhu, Ph.D. University of Technology Sydney Sydney, NSW AUSTRALIA
First Author:	Qiaoxi Zhu, Ph.D.
Order of Authors:	Qiaoxi Zhu, Ph.D. Xiaojun Qiu, Ph.D. Philip Coleman, Ph.D. Ian Burnett, Ph.D.
Section/Category:	Signal Processing in Acoustics
Keywords:	sound reproduction; sound zones; singular value decomposition; spatial harmonic decomposition.
Abstract:	<p>Personal audio provides private and personalized listening experiences by generating sound zones in a shared space with minimal interference between zones. One challenge of the design is to achieve the best performance with a limited number of microphones and loudspeakers. In this paper, two modal domain methods for personal audio reproduction are compared. One is the spatial harmonic decomposition (SHD) based method and the other is the singular value decomposition (SVD) based method. It is demonstrated that the SVD based method provides a more efficient modal domain decomposition than the SHD method for 2.5D personal audio design. Simulation results show that the SVD based method outperforms the SHD one by up to 10 dB in terms of acoustic contrast and up to 17 dB in terms of reproduction error for a compact arc array with 5 loudspeakers, while requiring fewer microphones around the zone boundaries. The SVD based method retains the inherent efficiency of optimizing in a modal domain while avoiding the inherent geometric limitations of using SHD basis functions. Thus, this approach is advantageous for applications with flexible system geometries and a small number of loudspeakers and microphones.</p>

CONFIDENTIAL



Click here to access/download

Reviewer PDF with line numbers, inline figures and captions

Reviewer_PDF.pdf

A comparison between two modal domain methods for personal audio reproduction

Qiaoxi Zhu¹, Xiaojun Qiu¹, Philip Coleman², Ian Burnett¹

¹ *Centre for Audio, Acoustics and Vibration, Faculty of Engineering and IT, University of
Technology Sydney, Sydney, Australia*

² *Institute of Sound and Recording, University of Surrey, Guildford, UK*

Running Title: Modal domain personal audio reproduction comparison

Keywords: sound reproduction; sound zones; singular value decomposition; spatial harmonic
decomposition

ABSTRACT

Personal audio provides private and personalized listening experiences by generating sound zones in a shared space with minimal interference between zones. One challenge of the design is to achieve the best performance with a limited number of microphones and loudspeakers. In this paper, two modal domain methods for personal audio reproduction are compared. One is the spatial harmonic decomposition (SHD) based method and the other is the singular value decomposition (SVD) based method. It is demonstrated that the SVD based method provides a more efficient modal domain decomposition than the SHD method for 2.5D personal audio design. Simulation results show that the SVD based method outperforms the SHD one by up to 10 dB in terms of acoustic contrast and up to 17 dB in terms of reproduction error for a compact arc array with 5 loudspeakers, while requiring fewer microphones around the zone boundaries. The SVD based method retains the inherent efficiency of optimizing in a modal domain while avoiding the inherent geometric limitations of using SHD basis functions. Thus, this approach is advantageous for applications with flexible system geometries and a small number of loudspeakers and microphones.

I. INTRODUCTION

Personal audio systems generate separate sound zones for each listener in a shared space with minimized interference between zones through a set of loudspeakers (Betlehem *et al.*, 2015). The pressure matching method is widely used for the sound zone reproduction, which combines acoustic transfer functions and least squares optimization to minimize the error between the reproduced and target sound fields at a number of control points (Poletti, 2008). However, it is not always practical to obtain the required large number of transfer functions between the loudspeakers and the zones.

To address this, the modal decomposition methods have been introduced to represent a sound field as a superposition of basis sound fields. The spatial harmonic decomposition based modal domain (SHDMD) method can parameterize acoustic transfer functions efficiently with a small number of orthogonal basis functions (Betlehem and Abhayapala, 2005; Samarasinghe *et al.*, 2015). The local sound field coefficients for each sound zone are then transformed to the equivalent global sound field coefficients using spatial harmonic coefficient translation between coordinate systems, allowing loudspeaker signals to be obtained through mode matching (Wu and Abhayapala, 2011). However, because only a limited number of spatial harmonics are used in the decomposition, the SHDMD method suffers performance degradation, especially when the zones are not in the center of the array, or a non-circular loudspeaker array is used. Because practical personal audio systems usually only have a limited number of loudspeakers (Cheer *et al.*, 2013a; Cheer *et al.*, 2013b; Gálvez *et al.*, 2014; Choi and Ji, 2016), a modal domain method with fewer geometric limitations is desirable.

An alternative modal domain approach uses singular value decomposition (SVD) to parameterize the acoustic transfer function matrices, where the loudspeaker weights for sound reproduction are reformulated by the loudspeaker-space modes. This SVD based modal domain (SVDMD) method has been previously studied for sound field reconstruction (Fazi, 2010), beamforming (Fazi *et al.*, 2014) and ultrasound (Tanter *et al.*, 2000), where the response matrix from an array to a set of control points is used to produce a desired field. When SVD is applied to personal audio reproduction (Zhu *et al.*, 2019), the weights of the loudspeaker-space modes are obtained by optimizing the compatibility between the capabilities of the listening zone and quiet zone modes for creating a listening zone and a quiet zone.

The relationship between the SHD and SVD based modes was investigated for the sound radiation and scattering analyses (Nelson and Kahana, 2001) and sound field reconstruction (Fazi and Nelson, 2007). In a spherical coordinate system, when the source points and the field points are chosen to sample the source and field appropriately, the matrices of left and right singular vectors obtained from SVD are related to the sampled spherical harmonics by a unitary transformation (Nelson and Kahana, 2001). Sound field reconstruction using SHD in three dimensional space can thus be regarded as a special case of that using SVD which employs regularly distributed monopole-like loudspeakers over the surface of a sphere (Fazi and Nelson, 2007). The relationship between these two modal approaches has not been investigated in the context of personal sound zones, where the reproduction over multiple local regions rather than a global region is pursued.

In our previous work (Zhu *et al.*, 2019), we proposed using the SVDMD method for geometric optimization of sound zone systems. The SVDMD method retains the advantage of modal domain transfer function parameterization, and suffers fewer geometric limitations than the SHDMD. In this paper, we will compare the two modal domain methods for personal audio reproduction and demonstrate the advantages of the SVDMD method over the state-of-the-art optimization in the SHDMD (Zhang *et al.*, 2018). The contributions of this work are as follows:

(1) We present new insights into the relationship between the SHD and SVD modes for 2.5D sound zone control. While the two sets of modes perform similarly when sufficient loudspeakers are distributed evenly over a circular boundary surrounding the controlled 2D space (the geometry favored by SHDMD), the SVD modes are inherently more flexible and readily adapted to other geometries (Sec. II);

(2) We show that the SVDMD method requires fewer acoustic transfer function measurements than the SHDMD method when using a limited number of loudspeakers partially cover the controlled space for sound reproduction (Sec. III);

(3) We show that the SVDMD method results in improved acoustic contrast and sound field reconstruction performance over a range of system (loudspeaker and zone) geometries compared to the SHDMD method under both free field (Sec. IV) and reflective room environments (Sec. V).

II. SOUND FIELD PARAMETERIZATION

This section first introduces the sound field parameterization, either using spatial harmonic decomposition or singular value decomposition, and then presents a comparison of the two approaches with examples.

Assume a sound zone q in the horizontal plane has a radius R_q and its center is denoted by \mathbf{O}_q . In the spatial harmonic decomposition, the sound pressure at any observation point $\mathbf{x} = (r, \theta)$ within the sound zone, having a radius r and an azimuth angle θ with respect to \mathbf{O}_q , can be expressed as (Williams, 1999)

$$P^{(q)}(\mathbf{x}, k) \approx \sum_{m=-N_q}^{N_q} \alpha_m^{(q)}(k) J_m(kr) e^{jm\theta}, \quad (1)$$

where $k = 2\pi f/c$ is the wavenumber with frequency f and the speed of sound propagation is c . $\alpha_m^{(q)}(k)$ is the corresponding coefficient of the m th two-dimensional spatial harmonic function $J_m(kr)e^{jm\theta}$, and $J_m(*)$ is the cylindrical Bessel function of order m . N_q is the truncation order of the local sound field defined as $N_q = \lceil ekR_q/2 \rceil$ (Kennedy *et al.*, 2007), where $\lceil * \rceil$ denotes the ceiling function and e is Euler's number. The coefficients $\alpha_m^{(q)}(k)$, $m = -N_q, \dots, N_q$, are obtained by the discrete spatial Fourier transform (Betlehem and Abhayapala, 2005), which requires at least $(2N_q+1)$ sound pressure measurements uniformly distributed over the circular zone boundary. The transfer function matrix between the loudspeakers and the sound zone q can be parameterized as a $(2N_q + 1) \times L$ matrix $\mathbf{\Gamma}$, where L is the number of loudspeakers for reproduction. The m th row and l th column element of $\mathbf{\Gamma}$ is the coefficient of the $(m-N_q-1)$ th spatial harmonic function of the local sound field generated by the l th loudspeaker.

In the SVDMD method, \mathbf{G}_q is defined as an $M \times L$ transfer function matrix between the loudspeakers and the sound zone q , where L is the number of loudspeakers for reproduction and M is the number of control points in the zone. Each element of \mathbf{G}_q includes the sound propagation attenuation and delay between the corresponding loudspeaker and control point. The transfer function matrix \mathbf{G}_q is decomposed by using SVD, as

$$\mathbf{G}_q = \mathbf{U}_q \mathbf{\Sigma}_q \mathbf{V}_q^H, \quad (2)$$

where \mathbf{U}_q is a $M \times M$ unitary matrix, \mathbf{V}_q is a $L \times L$ unitary matrix, and $\mathbf{\Sigma}_q$ is a $M \times L$ diagonal and real valued matrix. The columns of \mathbf{U}_q are the modes of the sound zone space with M being their degrees

of freedom; the columns of \mathbf{V}_q are the modes of the loudspeaker space in the system with L being their degrees of freedom (Fazi, 2010). Assuming $M > L$ (the number of control points in the zone is larger than the number of loudspeakers), the number of the degrees of freedom of the sound propagation channel is determined by the number of loudspeakers.

It is shown that the parameterizations of the acoustic transfer functions using SHD or SVD are similar for 2D sound reproduction when monopoles are distributed uniformly over the circular boundary to reproduce a concentric sound zone in free field (Nelson and Kahana, 2001; Fazi and Nelson, 2007). Specifically, the reproduced sound fields that use each mode of the loudspeaker space (column of \mathbf{V}_q) as weights of these monopoles are similar to the sound fields determined by each of the first L order cylindrical harmonics. However, the SHD and SVD modes are different from the following three aspects.

First, the magnitude and phase patterns between the SHD and SVD modes over the sound zone are different. As shown in Eq. (1), the magnitude of the SHD modes depends only on the radius r_q and the phase depends only on the azimuth θ_q , while the magnitude and phase of the SVD modes depend on both r_q and θ_q , as shown in Figs. 1(a) and (b) for the first 5 SHD modes (cylindrical harmonics) and the first 5 SVD modes reproduced by a 60-unit circular array. The system configuration for the SVD modes is detailed in Sec. IV.A.

Second, when applying a non-concentric sound zone (which is not concentric with the circular loudspeaker array) or non-circular loudspeaker array, the modes of the sound zone space in the SVDMD method are similar to but not exactly the same as the first L order cylindrical harmonics. This is illustrated by Fig. 1, where the zone center is located at $(-0.50, 0.00)$ m, rather than $(0.00, 0.00)$ m. Because more than one zone is considered in personal audio, at least one zone is outside of the array center. This difference in the controlled modes parameterized in the local sound field expression affects the performance achieved by the SHDMD and SVDMD methods.

Last, the SVD modes are adaptive to the system geometry of the loudspeakers and the controlled region, allowing the system's spatial sound propagation feature to be incorporated into the basis of sound field parameterization. Fig. 1(c) shows the 5 SVD modes reproduced by a 5-unit arc-shaped array, where the SVD modes are different from the SHD modes (which remain the same as those in Fig. 1(a) for any number

of loudspeakers in any position). The arc-shaped array is a sub-array of the circular array in Fig. 1(b). With reduced spatial sound propagation coverage over the controlled region, the basis sound fields in Fig. 1(c) present sharper sound orientation range throughout the controlled region than those in Fig. 1(b). However, this feature cannot be captured by the SHD modes.

To summarize, the SVD modes incorporate the spatial radiation feature/directivity of the loudspeaker array while SHD modes depend only on the orthogonal decomposition of the two-dimensional sound field in polar coordinates. Thus, the SHDMD method favors a concentric sound zone within a uniform circular loudspeaker array while the SVDMD modes generalize to any system geometry. For the particular system geometries favored by the SHDMD method, the SVDMD method parameterizes the acoustic transfer functions with almost the same efficiency as the SHDMD method, because the modes similar to the first order spatial harmonics are applied in the SVDMD method.

III. SOUND ZONE REPRODUCTION

A. Theory

In the SHDMD method, a global sound field (including all the controlled sound zones) is defined in the horizontal plane. It has a radius R_0 and its center is denoted by \mathbf{O}_0 . The SHDMD method for personal audio filter design has two stages (Wu and Abhayapala, 2011). The first stage translates the reproduction of multiple desired sound zones into the reproduction of a desired global sound field using the harmonic translation theorem. In the second stage, the global sound field produced by the loudspeakers is made to be close to the desired one by matching the coefficients of the produced and desired sound field expressed in the wave-based modal domain. For 2.5D multizone reproduction, cylindrical harmonics are used in the first stage to design the desired global sound field coefficients, while spherical harmonics are used in the second stage to represent the reproduced sound field. The weighted mode matching (Zhang *et al.*, 2018) is applied in the second stage to deal with the dimensionality mismatch between the 2D desired sound field and 3D sources.

In the first stage of the SHDMD method, the optimized global sound field coefficients are obtained as (Zhang *et al.*, 2018)

$$\mathbf{k}_{\text{SHD}} = (\mathbf{T}_L^H \mathbf{T}_L + \beta \mathbf{T}_Q^H \mathbf{T}_Q + \delta_1 \mathbf{I}_{2N_0+1})^{-1} \mathbf{T}_L^H \mathbf{a}_L, \quad (3)$$

where \mathbf{T}_L and \mathbf{T}_Q are the $(2N_q + 1) \times (2N_0 + 1)$ matrices representing the translations of coefficients from the local domains to the global domain, $N_0 = \lceil ekR_0/2 \rceil$ is the truncation order of the global region with a radius of R_0 , $(*)^H$ denotes complex conjugate matrix transpose, β determines the trade-off between the reproduction accuracy in the listening zone and the acoustic contrast between zones, δ_1 is the regularization parameter, \mathbf{I}_{2N_0+1} is a $(2N_0 + 1)$ order square identity matrix, and \mathbf{a}_L are the local cylindrical harmonic coefficients parameterized from the desired sound pressure distribution \mathbf{P}_{des} at the control points in the listening zone. The $(m+N_q+1)$ th row and the $(m'+N_0+1)$ th column element of \mathbf{T}_q (q denoting L or Q) is $J_{m'-m}(kr_q)e^{j(m'-m)\phi_q}$, given $\mathbf{x} = \mathbf{x}_q + \mathbf{O}_q$ and $\mathbf{O}_q = (r_q, \phi_q)$ in the global coordinates. It is derived from the harmonic translation theorem that

$$J_{m'}(kr')e^{jm'\phi'} = \sum_{m=-N_q}^{N_q} J_{m'-m}(kr_q)e^{j(m'-m)\phi_q} J_m(kr)e^{jm\phi}. \quad (4)$$

In the second stage, the loudspeaker weights are calculated by mode matching, as (Zhang *et al.*, 2018)

$$\mathbf{w}_{\text{SHD}} = (\mathbf{\Gamma}^H \mathbf{W} \mathbf{\Gamma} + \delta_2 \mathbf{I}_{2N_0+1})^{-1} \mathbf{\Gamma}^H \mathbf{X} \mathbf{k}_{\text{SHD}}, \quad (5)$$

where $\mathbf{\Gamma}$ is the acoustic transfer function coefficient matrix parameterized by SHD, \mathbf{W} is a $(N_0 + 1)^2$ order square weighting matrix for better control over the zone, \mathbf{X} is a $(N_0 + 1)^2 \times (2N_0 + 1)$ weighting matrix to deal with the dimensionality mismatch between the 2D desired sound field and 3D sources, and δ_2 is the regularization parameter to avoid ill-conditioning and increase robustness against perturbations in $\mathbf{\Gamma}$.

In the SVDMD method, \mathbf{G}_L and \mathbf{G}_Q are $M \times L$ transfer function matrices defining the listening and the quiet zones, respectively. They are decomposed using Eq. (2), then the columns of \mathbf{U}_L and \mathbf{U}_Q are respectively the modes of the listening zone space and the quiet zone space, the columns of \mathbf{V}_L and \mathbf{V}_Q are respectively the modes of the loudspeaker space corresponding to the listening zone and the quiet zone, and the values of the diagonal elements in $\mathbf{\Sigma}_L$ and $\mathbf{\Sigma}_Q$ represent the amount of amplification or attenuation that the modes undergo for the transformation \mathbf{G}_L and \mathbf{G}_Q . The coefficients of the loudspeaker-space modes are obtained by (Zhu *et al.*, 2019)

$$\mathbf{k}_{\text{SVD}} = (\mathbf{R}_L + \beta \mathbf{R}_Q)^{-1} \mathbf{\Sigma}_L^H \mathbf{U}_L^H \mathbf{P}_{\text{des}}, \quad (6)$$

where the spatial correlation matrices of the listening zone and the quiet zone are defined as $\mathbf{R}_L = \mathbf{\Sigma}_L^H \mathbf{\Sigma}_L + \delta_L \mathbf{I}_L$ and $\mathbf{R}_Q = \mathbf{C} \mathbf{\Sigma}_Q^H \mathbf{\Sigma}_Q \mathbf{C}^H + \delta_Q \mathbf{I}_L$ with the regularization parameters δ_Q and δ_L determined by the additive error model based regularization (Zhu *et al.*, 2017b), \mathbf{C} is an L order square unitary transfer matrix with $\mathbf{V}_Q = \mathbf{V}_L \mathbf{C}$, and \mathbf{I}_L is a L order identity matrix. The loudspeaker weights are calculated by

$$\mathbf{w}_{\text{SVD}} = \mathbf{V}_L \mathbf{k}_{\text{SVD}}. \quad (7)$$

B. Comparison

Figure 2 shows a flowchart representation of the two methods. Both approaches represent the local sound zone with local modes, transform the local mode coefficients to the global ones, solve the personal audio optimization using the coefficients of the global modes (\mathbf{k}_{SHD} or \mathbf{k}_{SVD}), and finally translate those coefficients to the loudspeaker weights (\mathbf{w}_{SHD} or \mathbf{w}_{SVD}).

The parameterizations of the acoustic transfer functions using SHD or SVD can be obtained through measurements or acoustic modeling (Betlehem and Abhayapala, 2005; Zhu *et al.*, 2017a). As denoted in Fig. 2, the SHD parameterization $\mathbf{\Gamma}$ of the acoustic transfer functions over a global region is input to the SHDMD method and the transfer function matrices \mathbf{G}_L and \mathbf{G}_Q are input to the SVDMD method. However, as shown in Fig. 2, the transfer functions are accounted for at the first stage in the SVDMD, while they are accounted for at the final stage in SHDMD. In practice, with the availability of $\mathbf{\Gamma}$, \mathbf{G}_L and \mathbf{G}_Q can be estimated by the SHD based sound field expression in Eq. (1); or, with the availability of \mathbf{G}_L and \mathbf{G}_Q , the local coefficients can be calculated, then $\mathbf{\Gamma}$ can be derived by the spatial harmonic coefficient translation (Wu and Abhayapala, 2011). Therefore, the SVDMD method does not necessarily require more prior information or measurements about the acoustic transfer functions than the SHDMD method.

Modal domain processing can reduce the number of transfer function measurements required for personal sound zone reproduction. The original least squares based method (Poletti, 2008) requires a large number of sampling points over a grid covering the zone with no more than half wavelength spacing, which increases with frequency. However, the correlation and similarity between measured values increase at low frequencies, leading to ill-conditioned transfer function matrices. Detailed higher-order information at low frequencies is therefore redundant and masked by the necessary regularization in personal audio algorithms

(Zhu *et al.*, 2017a). Both the SHDMD and SVDMD methods are orthogonal decomposition of the local sound field(s) in the modal domains, but the SHD modes do not account for the loudspeaker positions. This difference between SHD and SVD leads to the difference in the required number of measurements in acoustic transfer function parameterization.

The SHDMD method uses the spatial harmonic expansion to partially avoid measurement data redundancy with increased frequency and replace grid samples with a few boundary samples. The three-dimensional global region generated by each loudspeaker is modeled in practice from the measured sound pressure samples over the spherical boundary of the global region, and the required sampling is determined by the truncated order $N_0 = \lceil ekR_0/2 \rceil$ for a practical approximation. The required order of the global region increases with wavenumber k and size of the global region R_0 . The required number of samples over a spherical (or circular) region are separately proportional to $k^2R_0^2$ (or kR_0) by the SHDMD method and $k^3R_0^3$ (or $k^2R_0^2$) by the original least squares based method. Thus, the SHDMD method requires fewer samples than the original least squares based method for the same region. However, the SHDMD approach does not take into account the geometry information of the loudspeaker array. While for a circular (or spherical) geometry, the number of the truncated SHD modes is the minimal number of sound field samples to parameterize the reproduced sound field at the zones efficiently, it can be excessive for the geometry with loudspeakers partially distributed over the boundary.

In contrast to the SHD modes, the SVD modes depend on the physical geometry of the system. The basis functions in the SVDMD method are derived from the measured data and serve as an estimation of the basis functions for the reproducible sound field. The SVD modes maintain the spatial directivity due to the given geometry of the loudspeaker array and the zones. Besides, the radiation efficiency of each mode is proportional to its singular value, which decreases with increased order, so the radiation efficiency of the SVD modes decreases as the mode order increases (Tao and Qiu, 2009). With decreasing spatial coverage of the loudspeaker array or increasing distance between the loudspeaker array and the zones, the reproducible sound orientation of the loudspeaker array towards the controlled region narrows, so the system becomes not ideal for the SHDMD method, but the SVDMD method can adapt to the loudspeaker array and sound zone configuration and parameterize the reproducible sound field effectively with higher spatial resolution.

Compared to the SHD method, the SVD method allows a smaller number of modes for a satisfactory sound field parameterization, which requires fewer transfer function measurements. Therefore, the SVDMD method is more geometrically flexible and has better compactness in the modal expression of the specific sound zones. A comparison of the required density of sound field samples by the SHD and SVD methods will be illustrated in Sec. IV.C.

IV. EFFECTS OF SYSTEM GEOMETRY

A. Method

The performances of the SHDMD and the SVDMD methods are compared with simulations for two-zone reproduction under the 4 system geometries illustrated in Fig. 3. Two different layouts of loudspeakers are compared, one is a 60-unit circular array, and the other is a 5-unit arc-shaped array. The loudspeakers are assumed to be 3D monopoles distributed uniformly with angle spacing of 6° and a radius of 1.68 m. The circular array fully covers the 2D horizontal space enclosing the listening zone and the quiet zone. As an example of practical applications with a limited number of loudspeakers, the arc-shaped array with fewer loudspeakers only partially covers the 2D space.

Two different layouts of the listening and quiet zones are compared. They are the zones around the center of the circular array in Fig. 3(a) and in the near field of the arc-shaped array in Fig. 3(b). The radius of each controlled zone is 0.15 m. The listening and quiet zones around the center of the circular array are centered at (0.50, 0.00) m and (-0.50, 0.00) m respectively, while the listening and quiet zones in the near field of the arc-shaped array are centered at (0.00, 0.80) m and (-1.00, 0.80) m respectively. The global region in the SHDMD method is centered at (0.00, 0.00) m with a radius of $R_0=1.60$ m and covers the around center and off-center sound zones. These four configurations are used to demonstrate the performance degradations of the SHDMD method compared to the SVDMD method, which is caused by the inadequate distribution of loudspeakers and the off-center locations of sound zones.

The simulation applies a theoretical expression of the 3D monopoles to implement the SHDMD method. It is equivalent to parameterize the acoustic transfer functions from the measurement at no less than 83 ($=2 \times 41 + 1$) uniform samples over the circular boundary of the global region because the truncation order of the global region is $N_0 = \lceil ekR_0/2 \rceil = 41$. The required number of samples for a global region should be

no less than the sum of the samples required for each local region. The SVDMD method uses 20 samples, with each sound zone uniformly sampled by 10 controlled points on its circular boundary. The desired sound field in the listening zone is set as a plane wave at 1 kHz with a direction of 163° , as illustrated with a green arrow in Figs. 3(a) and (b). It is in the direction from the center loudspeaker of the arc-shaped array at (0.00, 1.68) m to the center of the listening zone at (0.50, 0.00) m. The trade-off parameter β is set as 1 to equally pursue acoustic contrast and reproduction accuracy in the listening zone.

The performance of the SHDMD and the SVDMD methods are evaluated with the perturbations added to the spatial responses, for which the error has a multiplicative form with Gaussian distribution between -3 dB and $+3$ dB in magnitude a and uniform distribution between -10° and $+10^\circ$ in phase ϕ . The m th row and l th column element of the transfer function matrix with perturbation is

$$\tilde{\mathbf{G}}_{(m,l)} = a\mathbf{G}_{(m,l)}e^{j\phi}. \quad (8)$$

Samples of $\tilde{\mathbf{G}}_L$ and $\tilde{\mathbf{G}}_Q$ are drawn from these distributions for Monte-Carlo trials. The acoustic contrast (AC), least squares error (LSE) and array effort (AE) metrics are used to evaluate the personal audio reproduction performance, which are defined as

$$\text{AC (dB)} = 20 \log_{10} \frac{\|\tilde{\mathbf{G}}_L \mathbf{w}\|}{\|\tilde{\mathbf{G}}_Q \mathbf{w}\|}, \quad (9)$$

$$\text{LSE (dB)} = 20 \log_{10} \frac{\|\tilde{\mathbf{G}}_L \mathbf{w} - \mathbf{P}_{\text{des,eval}}\|}{\|\mathbf{P}_{\text{des,eval}}\|}, \quad (10)$$

$$\text{AE (dB)} = 20 \log_{10} \frac{\|\mathbf{w}\|}{|w_0|}, \quad (11)$$

where $\mathbf{P}_{\text{des,eval}}$ represents the desired sound pressure at the samples over the listening zone, w_0 is the input signal required to drive the loudspeaker at (0.00,1.68) m so that the mean square pressure in the listening zone is the same as that when the array is driven by \mathbf{w} and $|\cdot|$ is the operator of absolute value. These metrics are evaluated by 709 grid samples with 0.01 m spacing over each controlled zone. The mean AC and mean LSE averaged over 100 Monte-Carlo trials describe the performance of the system with robustness, and a low AE value refers to reproduction with high energy efficiency.

The regularization in the SHDMD method applies $\delta_1 = 0.1$ and $\delta_2 = \sigma/100$, where σ is the maximal singular value of the $\mathbf{\Gamma}^H \mathbf{W} \mathbf{\Gamma}$. The regularization parameters δ_L and δ_Q for the SVDMD method are set in the same way as in our previous work (Zhu *et al.*, 2017b). Specifically, to make the

assumed additive error set contain the real multiplicative error set, AEQ adopts $a_{\max} = \max(\mathbf{G}) \times a_{\max, \text{ME}}$, where $\max(\mathbf{G})$ is the maximum among all the transfer functions in \mathbf{G} , and $a_{\max, \text{ME}} = \sqrt{\mu^2 - 2\cos(\phi_{\max})\mu + 1}$ with $\mu = 10^{\frac{3}{20}}$, $\phi_{\max} = 10^\circ$.

B. Results

The sound fields at 1 kHz reproduced by the SHDMD and the SVDMD methods under free field conditions separately using the circular and arc-shaped arrays, under two controlled zone settings, are presented in Fig. 4 with the corresponding mean AC, mean LSE and AE performances presented in Table I. It can be observed that under the system geometry with a circular array and zones around the center, the AC performance of the SVDMD method is similar to that of the SHDMD method, with 0.8 dB advantage. This is because, as mentioned in Sec. II, the controlled modes in the SVDMD method (using uniform sampling over the circular boundary of each controlled zone) are almost equivalent to those truncated SHD modes in the SHDMD method under this geometry.

For the zones around the centre, the LSE performance of the SHDMD method is 12.6 dB worse than the SVDMD method. This is related to the definition of the global sound field. Reducing the radius of the global region to 0.75 m to tightly surround two controlled zones, the mean LSE of the SHDMD method increased to -15.6 dB with a mean AC of 25.3 dB. In this case, the mean AC and mean LSE degradations of the SHDMD method compared to the SVDMD method are reduced to 0.6 dB and 3.8 dB.

Both the SHDMD and SVDMD methods are affected by the system geometry. For example, the arc-shaped array might be preferred for practical applications over the circular array when the number of loudspeakers is limited. This reduction of system geometry leads to a 2.1 dB mean AC degradation and 1.6 dB mean LSE improvement over the initial geometry with 60 loudspeakers when using the SVDMD method with zones around the center. However, the performance of the SHDMD method deteriorates significantly by the reduction of the loudspeaker number, with 11.1 dB mean AC and 3.1 dB mean LSE performance degradation compared with the initial geometry.

When the listening zone is located close to the arc-shaped array, the performance of the SVDMD method with the arc-shaped array is still acceptable with mean AC above 20.0 dB and mean LSE below -15.0 dB.

With the increased reproduction capability of local sound orientations due to the adjusted listening zone position, the SHDMD method has 2.0 dB mean AC and 1.7 dB mean LSE improvement when using the arc-shaped array; however, they are still 4.5 dB lower in mean AC and 10.1 dB higher in mean LSE than that of the SVDMD method.

These findings also hold for swapped listening and quiet zones. If the system geometry becomes non-ideal for the reproduction, the performance of both methods suffers. The non-ideal system geometry can be the cases that the quiet zone is closer to the loudspeakers than the listening zone or the desired sound direction throughout the listening zone is not in favor of the system geometry. The SHDMD method can be further affected by the mismatch between the system geometry and the inherent geometric preference of the SHDMD method, making its performance worse than that of the SVDMD method.

The previous simulations only presented the performance at 1 kHz. The performance over 36 frequencies between 100 Hz and 3600 Hz is presented in Fig. 5. The estimated aliasing frequencies (Winter *et al.*, 2019) are 2.7 kHz and 2.0 kHz for the configurations with zone around the center and in the near field, respectively. The general trend is that each predefined system can produce satisfactory results at frequencies below the corresponding aliasing frequency, while the performance decreases significantly at higher frequencies.

Below the aliasing frequency, the findings observed at 1000 Hz are generally maintained. Specifically, when using the arc-shaped array, the SVDMD method has a distinct advantage in the mean AC performance than the SHDMD method and performs close to that with the circular array except for low frequencies below 0.8 kHz. The AC performance of the SVDMD method using the circular array suffers at the frequencies around 873, 1379 and 1854 Hz because one of the major modes generated by the circular array is inactive at the boundary samples due to the zeros of the cylindrical Bessel functions ($J_0(kr_q)$, $J_{\pm 1}(kr_q)$ and $J_{\pm 2}(kr_q)$, respectively) (Fazi and Nelson, 2012), and the SVDMD method fails to take these modes into account. The SVDMD method has a better overall LSE performance than the SHDMD method for the circular array below 0.7 kHz and the arc-shaped array above 0.8 kHz. The SVDMD method suffers less ill-conditioning at low frequencies as shown by the AE performance over low frequencies below 0.6 kHz.

Figure 6 shows the AC and AE performance of the SHDMD and the SVDMD methods evaluated with the sound zones defined in Fig. 3(a) and various loudspeaker array configurations in terms of different

numbers of loudspeakers and different spacings. The anti-diagonal elements in each sub-figure correspond to the full spatial coverage circular arrays with different numbers of loudspeakers. The upper triangular elements correspond to the arc-shaped arrays with different coverage of the 2D space. The lower triangular elements are not available, and their performance is set to 0 dB. Though the two methods perform closely when using a circular array with 30 loudspeakers, the SHDMD method is slightly worse than the SVDMD method when using a circular array with other numbers of loudspeakers. When using the arc-shaped arrays, the overall performance (in terms of AC and AE) of the SHDMD method is worse than that of the SVDMD method.

C. Discussions

The loudspeaker weights used for reproduction are obtained from the received values of the samples over the boundary of the controlled circular zones. An efficient implementation with a minimal number of boundary samples is desirable for practical applications. The required number of samples in each controlled zone is related to the upper frequency limit. To avoid the aliasing effects illustrated in Fig. 5, the upper limit frequency of the system can be set as 2 kHz. Fig. 7 presents the performance of the SVDMD method at 2 kHz which employs different numbers of uniformly distributed boundary samples of each controlled zone. The truncation order is $N_q = \lceil \pi e R_q / \lambda \rceil = 8$, which depends on the ratio of the radius of zone R_q and the wavelength $\lambda = c/f$. The degrees of freedom of each sound zone generated by the circular array is $2N_q + 1 = 17$. This is the minimal number of boundary samples required by the SHDMD method for each controlled zone.

It is shown in Fig. 7 that the SVDMD method using the circular array results in optimized and stable performance when the number of samples reaches 17. Generated by the arc-shaped array, the degrees of freedom of each sound zone is reduced to $2N_{q,arc} + 1 = 3$ with $N_{q,arc} = \lceil \Omega e R_q / \lambda \rceil = 1$, where $[-\Omega, \Omega)$ is the angular range of the arc-shaped array (Dickins *et al.*, 2005). This number of degrees of freedom is less than 5, which is the number of the loudspeakers of the arc-shaped array. When the number of samples reaches 3, the performance of the SVDMD method using the arc-shaped array also becomes optimized and stable. Besides, as shown in Table II, its overall performance is not only better than the SHDMD method with the same arc-shaped array, but also better than the SVDMD method using the circular array, which only controls

the first three modes generated by the circular array. Therefore, in each controlled zone, the required number of boundary samples by the SVDMD method is related to the degrees of freedom of the reproduced sound zone and no more than the number of loudspeakers. This number is less than that required by the SHDMD method, when a few or spatially partially distributed loudspeakers are used for sound reproduction.

The least squares based method (Poletti, 2008) is optimal in the sense that the reproduced sound pressures at the control points are numerically optimized. It is observed in Table II that the SVDMD method has performance degradation within 1.3 dB compared to the least squares based method with the arc-shaped array. However, the SVDMD method only employs 3 samples over the boundary of each controlled zone, while the least squares method employs 25 grid samples over each controlled zone, to ensure that the spacing of samples is less than half a wavelength.

In summary, with loudspeakers fully distributed over the boundary of a space, the SHDMD and SVDMD methods perform similarly because the major modes of the SVDMD method are almost equivalent to the SHDMD modes; however, with spatially, partially distributed loudspeakers, or off-center zones, the SHDMD method is not optimal and performance is not as good as that of the SVDMD method. This is because the SHD method applies the basis functions of a general sound field, while the SVD method applies the basis functions related to the reproducible sound field, which match the reproduced sound field better. Thus, fewer samples can be used in the transfer function measurements.

V. EFFECTS OF ROOM REFLECTIONS

The SHDMD and SVDMD methods can be extended to more complex geometries, where room acoustics is incorporated. Following (Zhang *et al.*, 2016), a room of size 10×9 m is simulated using the image source method (Allen and Berkley, 1979) with the image order up to 5 (i.e., 60 image sources) for each loudspeaker. The wall reflection coefficients are 0.7 and a perfectly-absorbing surface is assumed for floor and ceiling. The performance is also evaluated over 100 Monte-Carlo trials, which leads to relatively stable performance evaluation by observation. Other settings are the same with the free-field simulation in Sec. IV.

Compared to the free-field reproduction, the effect of reverberation leads to decreased spatial directivity of the loudspeaker array at high frequencies (Simón-Gálvez *et al.*, 2014). Both the SHDMD and

SVDMD methods are based on the representation of the transfer functions between loudspeakers and the sound zones with a set of modes and their coefficients. As presented in Sec. II, the SVD modes have advantages in representing the local sound field with sound from certain directions, which can be generated by an arc-shaped array in the free field; while the SVD modes become equivalent to the SHD modes when sound from all the directions, which can be generated by a circular array in the free field. With increasing sound reflections and reverberations, the sound field generated by a loudspeaker becomes more diffuse. The extreme case is the diffuse field, where sound from all the directions. Therefore, the performance difference between the two methods becomes smaller with increased sound reflections and reverberations.

Figure 8 shows the spatial directivity patterns of the SVD modes, which are blurred in the reverberant room at 1 kHz. The sound fields reproduced by the SHDMD and the SVDMD methods in the reverberant room using the circular and arc-shaped arrays under two controlled zone settings are presented in Fig. 9 with the corresponding mean AC, mean LSE and AE performances presented in Table III. With loudspeaker array capability in the spatial radiation/directivity modified by the room condition, reduced mean AC and mean LSE performance is observed in Table III. Compared to that in Table I, the mean AC degradation is within 4.8 dB and the mean LSE degradation is within 5.0 dB. However, the SVDMD method still has better overall performance than the SHDMD method, and maintains acceptable performance with mean AC above 18.0 dB and mean LSE below -14.0 dB when using the arc-shaped array. It can be noticed that, the third SVD mode in Fig. 8(b) has a totally different pattern from the free field SVD modes in Fig. 1(c).

In terms of the required number of the boundary samples in each controlled zone, the upper limit for the SVDMD method is determined by the minimum value between the dimension of the reproduced sound zone and the number of loudspeakers. As shown in Fig. 10, the SVDMD method achieves stable performance at 2 kHz when the number of samples reaches 5 and 17 using the arc-shaped array and the circular array, respectively. In the room, the loudspeakers and their image sources are distributed over the full angular range and the dimension of the reproduced sound zone is 17 at 2 kHz. When using the arc-shaped array, the required number of the samples is 5 (less than 17). It is determined by the number of loudspeakers, as only 5 SVD modes are reproduced and controlled. The SVDMD method using the arc-shaped array with 5 boundary samples in each controlled zone results in acceptable performance with mean AC above 10 dB, mean LSE below -10 dB and AE below -3 dB.

Table IV shows that the SVDMD method has performance degradation within 1.2 dB compared to the least squares based method, when using the arc-shaped array. However, the SVDMD method employs only 5 samples over the boundary of each controlled zone while the least squares method employs 25 grid samples over each controlled zone. So the SVDMD method allows sound zone reproduction in a reflective environment with a few loudspeakers and a few sound zone samples.

VI. CONCLUSIONS

Based on the decomposition of acoustic transfer functions through spatial harmonics or singular value functions, two modal domain methods are compared for 2.5D personal audio reproduction design. The SHD based method only considers the modal decomposition of the two-dimensional sound field while the SVD based method applies loudspeaker-oriented modal decomposition over the controlled sound zones. The former approach sets the optimal global sound field distribution as the priority and uses the full lower order dimensionality of the sound field, while the latter optimizes compatibility in the reproduction of the listening and quiet zones by using the link between local modes of the listening/quiet zones and the global modes of the loudspeaker space. The analysis and simulation results show that the SHD based modal domain method deteriorates when only a limited number of loudspeakers cover the controlled space partially or sound zones are off-center. The SVD based method has geometric flexibility of loudspeaker distribution and more compressive expression for sound reproduction with a few loudspeakers. Thus, the SVD based modal domain method is relevant for applications with flexible system geometries and a small number of loudspeakers, while still providing the benefits of a modal-domain optimization by reducing the number of acoustic measurements required. Future work includes experimental validation and the investigation on the number and distribution of the sound zone samples required to accurately parameterize the acoustic transfer functions in the SVD based modal domain. The uniform circular sampling is used in this paper to present the relationship between SHD and SVD, however, other samplings might be desired in practical applications.

ACKNOWLEDGEMENTS

This research is supported under the Australian Research Council's Linkage Project funding scheme (LP160100616).

REFERENCES

- Allen, J. B., and Berkley, D. A. (1979). "Image method for efficiently simulating small- room acoustics," *The Journal of the Acoustical Society of America* **65**, 943-950.
- Betlehem, T., and Abhayapala, T. D. (2005). "Theory and design of sound field reproduction in reverberant rooms," *The Journal of the Acoustical Society of America* **117**, 2100-2111.
- Betlehem, T., Zhang, W., Poletti, M. A., and Abhayapala, T. D. (2015). "Personal Sound Zones: Delivering interface-free audio to multiple listeners," *IEEE Signal Processing Magazine* **32**, 81-91.
- Cheer, J., Elliott, S. J., and Gálvez, M. F. S. (2013a). "Design and implementation of a car cabin personal audio system," *Journal of the Audio Engineering Society* **61**, 412-424.
- Cheer, J., Elliott, S. J., Kim, Y., and Choi, J.-W. (2013b). "Practical implementation of personal audio in a mobile device," *Journal of the Audio Engineering Society* **61**, 290-300.
- Choi, J.-W., and Ji, J. (2016). "A self-configurable, wireless audio system with user-tracking ability," in *Audio Engineering Society Conference: 2016 AES International Conference on Sound Field Control* (Audio Engineering Society).
- Dickins, G., Williams, M., and Hanlen, L. W. (2005). "On the dimensionality of spatial fields with restricted angle of arrival," in *Information Theory, 2005. ISIT 2005. Proceedings. International Symposium on* (IEEE), pp. 1033-1037.
- Fazi, F. M. (2010). "Sound field reproduction," (University of Southampton).
- Fazi, F. M., and Nelson, P. A. (2007). "A theoretical study of sound field reconstruction techniques," in *19th International Congress on Acoustics* (Spain).
- Fazi, F. M., and Nelson, P. A. (2012). "Nonuniqueness of the solution of the sound field reproduction problem with boundary pressure control," *Acta Acustica united with Acustica* **98**, 1-14.
- Fazi, F. M., Shin, M., Olivieri, F., Fontana, S., and Lang, Y. (2014). "Comparison of pressure-matching and mode-matching beamforming for methods for circular loudspeaker arrays," in *Audio Engineering Society Convention 137* (Audio Engineering Society).
- Gálvez, M. F. S., Elliott, S. J., and 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**, 1824-1831.
- Kennedy, R. A., Sadeghi, P., Abhayapala, T. D., and Jones, H. M. (2007). "Intrinsic limits of dimensionality and richness in random multipath fields," *IEEE Transactions on Signal processing* **55**, 2542-2556.
- Nelson, P. A., and Kahana, Y. (2001). "Spherical harmonics, singular-value decomposition and the head-related transfer function," *Journal of Sound and Vibration* **239**, 607-637.
- Poletti, M. (2008). "An investigation of 2-d multizone surround sound systems," in *Audio Engineering Society Convention 125* (Audio Engineering Society).
- Samarasinghe, P., Abhayapala, T., Poletti, M., and Bethlehem, T. (2015). "An efficient parameterization of the room transfer function," *IEEE/ACM Transactions on Audio, Speech, and Language Processing* **23**, 2217-2227.
- Simón-Gálvez, M. F., Elliott, S. J., and Cheer, J. (2014). "The effect of reverberation on personal audio devices," *The Journal of the Acoustical Society of America* **135**, 2654-2663.
- Tanter, M., Thomas, J.-L., and Fink, M. (2000). "Time reversal and the inverse filter," *The Journal of the Acoustical Society of America* **108**, 223-234.
- Tao, J., and Qiu, X. (2009). "Sensor number requirements for sound power prediction using two different mode decomposition methods," *Applied Acoustics* **70**, 646-651.
- Williams, E. G. (1999). *Fourier acoustics: sound radiation and nearfield acoustical holography* (Elsevier).
- Winter, F., Schultz, F., Firtha, G., and Spors, S. (2019). "A Geometric Model for Prediction of Spatial Aliasing in 2.5 D Sound Field Synthesis," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*.
- Wu, Y. J., and Abhayapala, T. D. (2011). "Spatial multizone soundfield reproduction: Theory and design," *IEEE Transactions on audio, speech, and language processing* **19**, 1711-1720.
- Zhang, W., Abhayapala, T. D., Bethlehem, T., and Fazi, F. M. (2016). "Analysis and control of multi-zone sound field reproduction using modal-domain approach," *The Journal of the Acoustical Society of America* **140**, 2134-2144.
- Zhang, W., Zhang, J., Abhayapala, T. D., and Zhang, L. (2018). "2.5D multizone reproduction using weighted mode matching," in *ICASSP 2018*.

- Zhu, Q., Coleman, P., Qiu, X., Wu, M., Yang, J., and Burnett, I. (2019). "Robust Personal Audio Geometry Optimization in the SVD-Based Modal Domain," *IEEE/ACM Transactions on Audio, Speech, and Language Processing* **27**, 610-620.
- Zhu, Q., Coleman, P., Wu, M., and Yang, J. (2017a). "Robust Acoustic Contrast Control with Reduced In-situ Measurement by Acoustic Modeling," *Journal of the Audio Engineering Society* **65**, 460-473.
- Zhu, Q., Coleman, P., Wu, M., and Yang, J. (2017b). "Robust reproduction of sound zones with local sound orientation," *The Journal of the Acoustical Society of America* **142**, EL118-EL122.

TABLES

TABLE I *The mean AC, mean LSE and AE performances (dB) of the SHDMD and the SVDMD methods under four typical system geometries at 1 kHz. (The SHDMD method applies a global region centered at (0.00, 0.00) m with a radius of 1.60 m, to cover the controlled zones in center and off-center settings.)*

Geometry	Zones around the center				Zones in the near field			
	Circular array		Arc-shaped array		Circular array		Arc-shaped array	
Method	SHDMD	SVDMD	SHDMD	SVDMD	SHDMD	SVDMD	SHDMD	SVDMD
Mean AC	25.1	25.9	14.0	23.8	24.0	24.2	16.0	20.5
Mean LSE	-6.8	-19.4	-3.7	-21.0	-10.1	-14.6	-5.4	-15.5
AE	-8.6	-9.1	-6.8	-6.6	-6.3	-7.1	-5.6	-3.0

TABLE II *The mean AC, mean LSE and AE performances (dB) of the original least squares, the SHDMD and the SVDMD methods, employing the required number of samples, under four typical system geometries at 2 kHz.*

Geometry	Method	Sample number	Zones around the center			Zones in the near field		
			Mean AC	Mean LSE	AE	Mean AC	Mean LSE	AE
Circular array	LS	25*	24.0	-18.6	-6.5	14.3	-9.9	-3.9
	SHDMD	17	21.8	-2.6	-5.8	6.3	-3.5	-3.5
	SVDMD	17	22.9	-8.8	-6.5	10.5	-7.8	-3.9
Arc-shaped array	LS	25*	23.1	-19.1	-5.8	11.1	-10.9	-1.5
	SHDMD	17	19.8	-2.5	-6.3	4.2	-2.6	-3.0
	SVDMD	3	22.9	-17.8	-5.8	11.9	-10.7	-1.1

*: The least squares method applies 5×5 grid samples, rather than the boundary samples in the SHDMD and the SVDMD methods.

TABLE III The mean AC, mean LSE and AE performances (dB) of the SHDMD and the SVDMD methods under four typical system geometries in a simulated room environment at 1 kHz.

Geometry	Zones around the center				Zones in the near field			
	Circular array		Arc-shaped array		Circular array		Arc-shaped array	
Method	SHDMD	SVDMD	SHDMD	SVDMD	SHDMD	SVDMD	SHDMD	SVDMD
Mean AC	22.6	24.5	12.5	19.0	21.4	24.1	15.6	18.8
Mean LSE	-6.6	-16.9	-3.6	-16.0	-9.7	-13.6	-5.1	-14.7
AE	-9.0	-9.8	-7.2	-6.7	-6.3	-7.2	-5.5	-3.4

TABLE IV The mean AC, mean LSE and AE performances (dB) of the original least squares, the SHDMD and the SVDMD methods, employing the required number of samples, under four typical system geometries at 2 kHz in a simulated room environment.

Geometry	Method	Sample number	Zones around the center			Zones in the near field		
			Mean AC	Mean LSE	AE	Mean AC	Mean LSE	AE
Circular array	LS	25*	16.5	-11.7	-4.6	14.5	-9.0	-4.4
	SHDMD	17	13.1	-2.7	-4.5	5.8	-3.5	-3.9
	SVDMD	17	14.6	-5.7	-4.6	11.6	-7.3	-4.2
Arc-shaped array	LS	25*	12.7	-10.6	-4.4	8.9	-7.4	-2.1
	SHDMD	17	12.4	-2.5	-4.7	3.9	-2.6	-3.3
	SVDMD	5	11.5	-10.4	-3.6	10.0	-6.2	-1.3

*: The least squares method applies 5×5 grid samples, rather than the boundary samples in the SHDMD and the SVDMD methods.

FIGURE CAPTIONS

Fig. 1 (Color online) The spatial distribution of the amplitude and phase at 1 kHz of the SHD modes and SVD modes of the sound zone centered at $(-0.50, 0.00)$ m with a radius of 0.15 m. (a) The first 5 cylindrical harmonics used in the SHDMD method. (b) The first 5 SVD modes reproduced by a 60-unit circular array. (c) The 5 SVD modes reproduced by a 5-unit arc-shaped array. The sound zone is denoted by the red circle. The circular and arc-shaped arrays are introduced in Sec. IV with Fig. 3(a).

Fig. 2 (Color online) Flowchart for comparing the SHDMD and the SVDMD methods. The listening and the quiet zone optimizations in the SHD domain (Zhang et al., 2016) are separately ① $\min \|\mathbf{T}_L \mathbf{k}_{\text{SHD}} - \mathbf{a}_L\|_2^2$ and ② $\min \|\mathbf{T}_Q \mathbf{k}_{\text{SHD}}\|_2^2$, while the listening and the quiet zone optimizations in the SVD domain (Zhu et al., 2019) are separately ③ $\min \|\Sigma_L \mathbf{k}_{\text{SHD}} - \mathbf{U}_L^H \mathbf{P}_{\text{des}}\|_2^2$ and ④ $\min \|\Sigma_Q \mathbf{C}^H \mathbf{k}_{\text{SHD}}\|_2^2$. $\|\cdot\|$ denotes Euclidean norm. Γ is the measured input of the SHDMD method, while \mathbf{G}_L and \mathbf{G}_Q are the measured inputs of the SVDMD method.

Fig. 3 (Color online) Four typical system geometries for performance comparison of the SHDMD and the SVDMD methods. The system geometries include personal sound reproduction separately using the arc-shaped or circular loudspeaker array for the quiet zone (left circle) and the listening zone (right circle) around the center of the circular array (a) or in the near field of the arc-shaped array (b).

Fig. 4 (Color online) The real part of the sound field at 1 kHz reproduced by the SHDMD and the SVDMD methods separately using the circular array and the arc-shaped array, under two controlled zone settings: around the center of the circular array (a) or in the near field of the arc-shaped array (b). The global region in the SHDMD method is centered at $(0.00, 0.00)$ m with a radius of 1.60 m, to cover the controlled zones in both settings.

Fig. 5 (Color online) The mean acoustic contrast (AC), mean least square error (LSE) and array effort (AE) performance of the SHDMD and the SVDMD methods separately using the circular array ('CIR') and the arc-shaped array ('ARC'), under two controlled zone settings over 0.1–3.6 kHz.

Fig. 6 (Color online) The acoustic contrast (AC) and array effort (AE) performance of the SHDMD and the SVDMD methods at 1 kHz using loudspeaker array with different spatial coverage, by varying number of loudspeakers and spacings.

Fig. 7 (Color online) The mean acoustic contrast (AC), mean least square error (LSE) and array effort (AE) performance of the SVDMD method employing different number of samples over the boundary of each controlled zone, separately using the circular array ('CIR') and the arc-shaped array ('ARC'), under two controlled zone settings at 2 kHz. The vertical solid line and the vertical dashed line denote that the numbers of samples are 17 and 3, corresponding to the degrees of freedom of each sound zone generated by the circular array and the arc-shaped array, respectively.

Fig. 8 (Color online) The spatial distribution of the amplitude and phase of the SVD modes of the sound zone centered at $(-0.50, 0.00)$ m with a radius of 0.15 m in a simulated room environment at 1 kHz. (a) The first 5 SVD modes reproduced by a 60-unit circular array. (b) The 5 SVD modes reproduced by a 5-unit arc-shaped array. The sound zone is denoted by the red circle.

Fig. 9 (Color online) The real part of the sound field at 1 kHz reproduced in a simulated room environment by the SHDMD and the SVDMD methods separately using the circular array and the arc-shaped array, under two controlled zone settings: around the center of the circular array (a) or in the near field of the arc-shaped array (b).

Fig. 10 (Color online) The performance of the SVDMD method employing different numbers of samples over the boundary of each controlled zone in a simulated room environment, separately using the circular array('CIR') and the arc-shaped array ('ARC'), under two controlled zone settings at 2 kHz. The vertical solid line and the vertical dashed line denote that the numbers of samples are 17 and 5, corresponding to the degrees of freedom of each controlled zone generated by the circular array and the arc-shaped array, respectively.

Fig1a

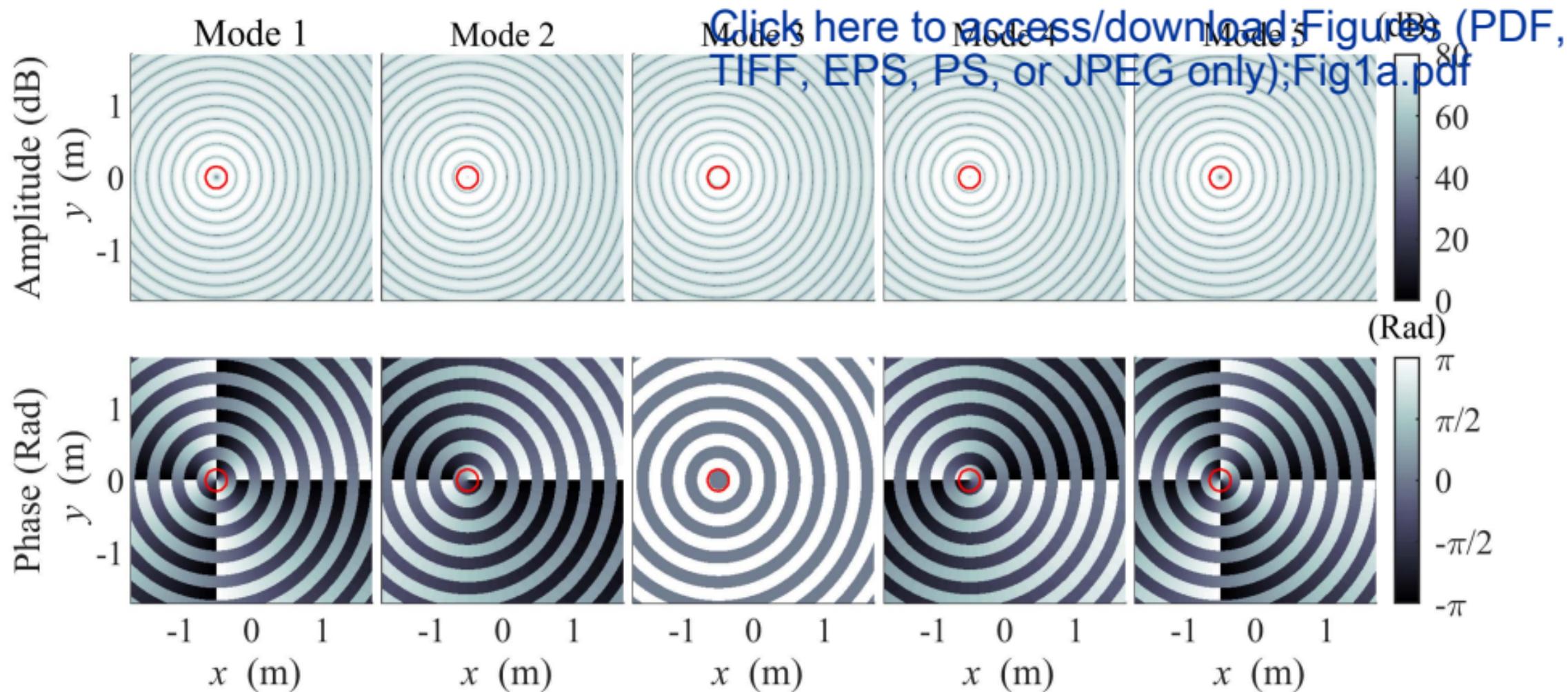


Fig1b

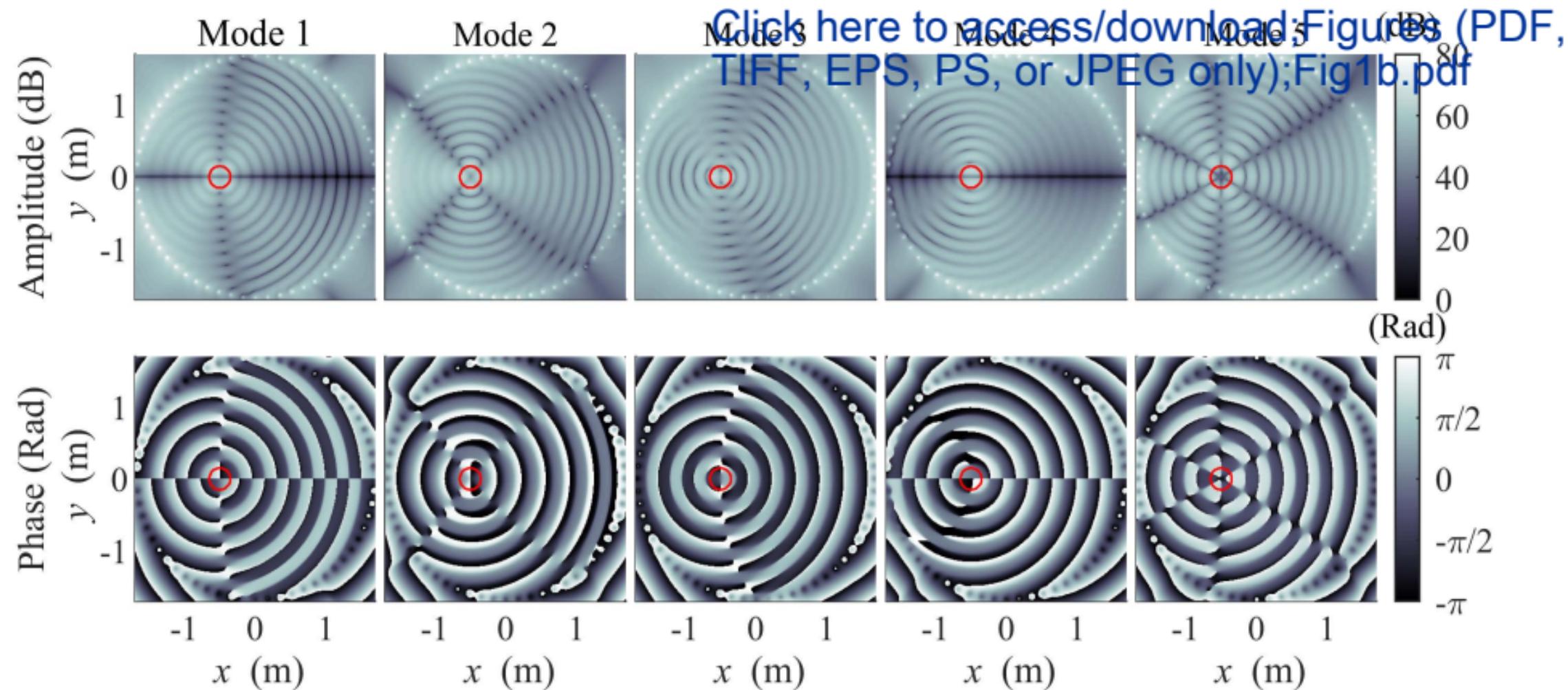


Fig1c

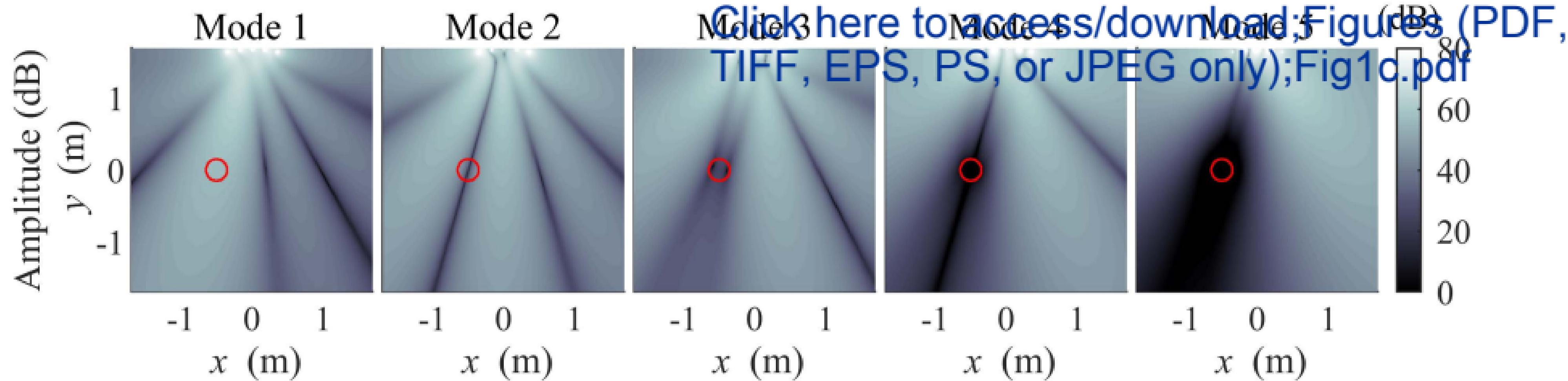


Fig2

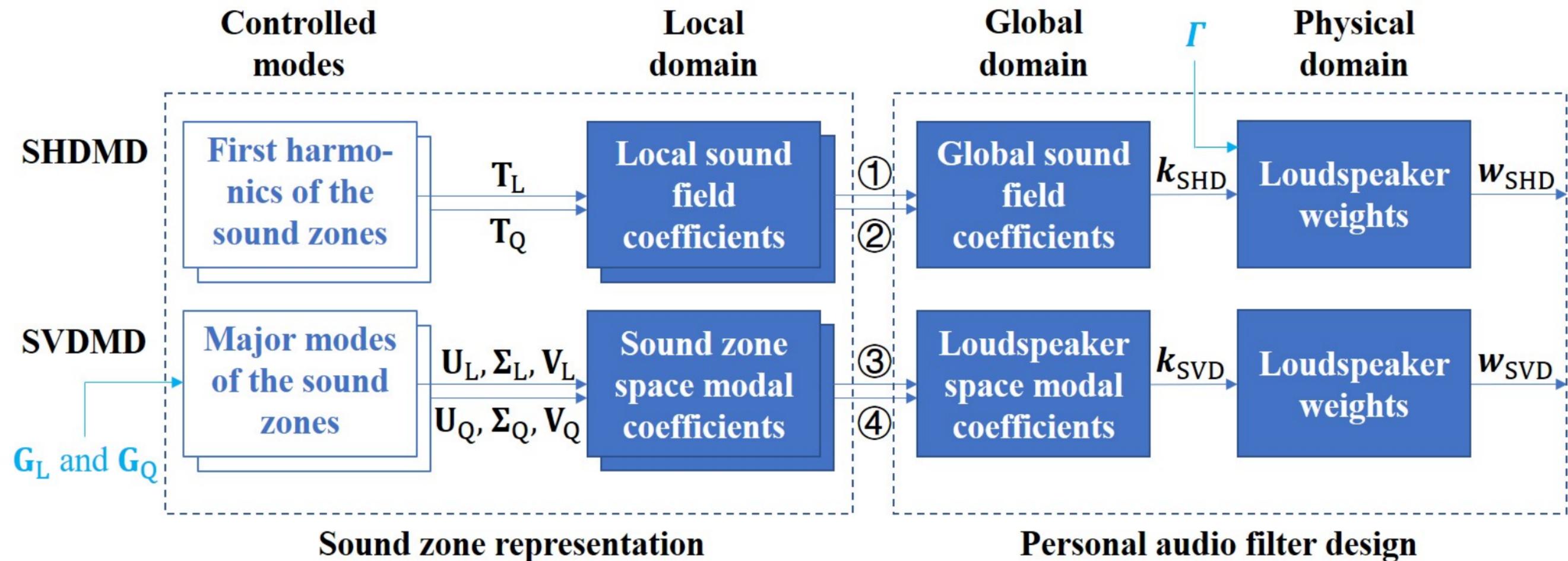


Fig3a

Arc-shaped Array

[Click here to access/download;Fig](#)



Fig3b

Arc-shaped Array

Click here to

[access/download;Fig](#)

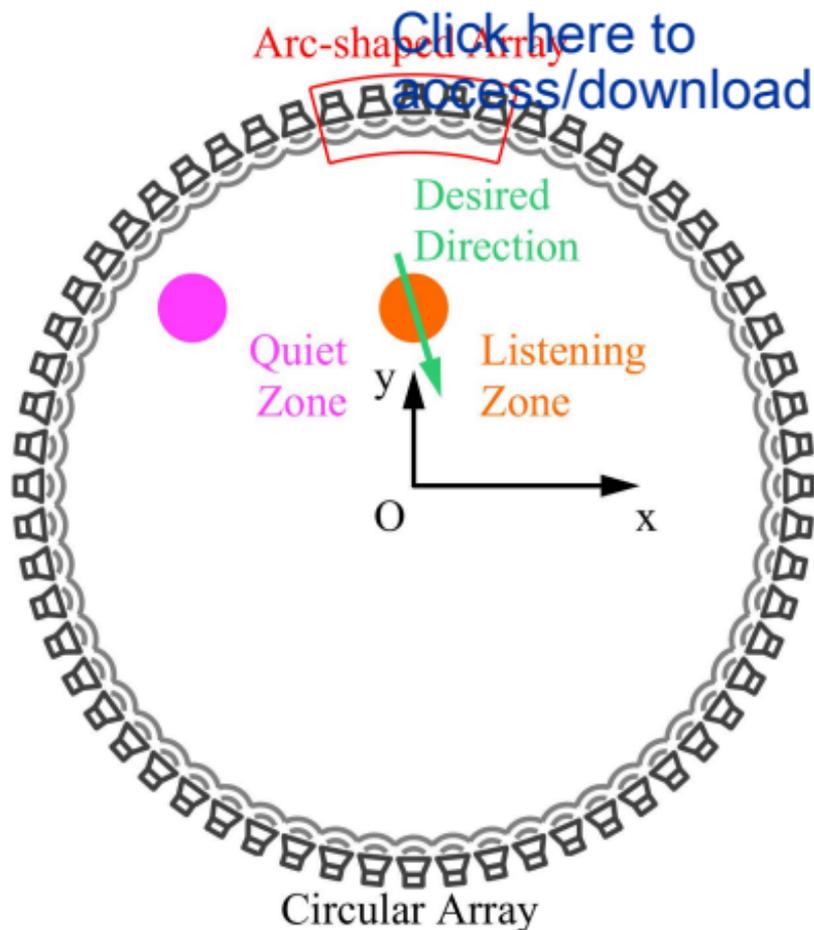
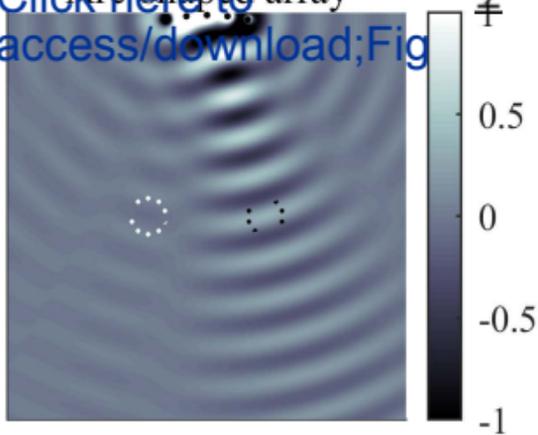
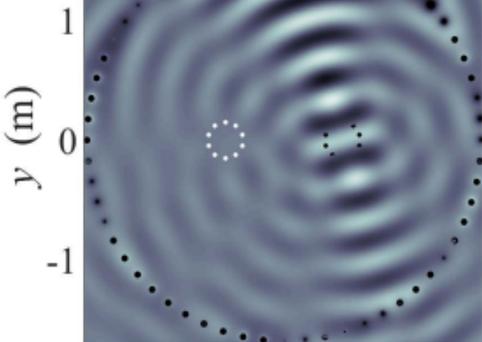


Fig4a

Circular array

Ar-shaped array
[Click here to access/download;Fig](#)

SHDMD



SVDMD

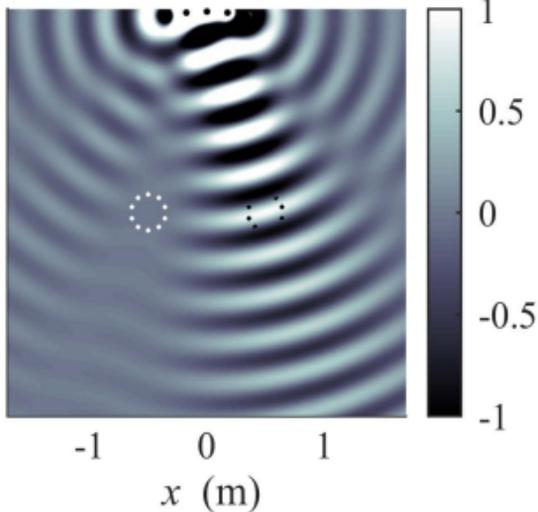
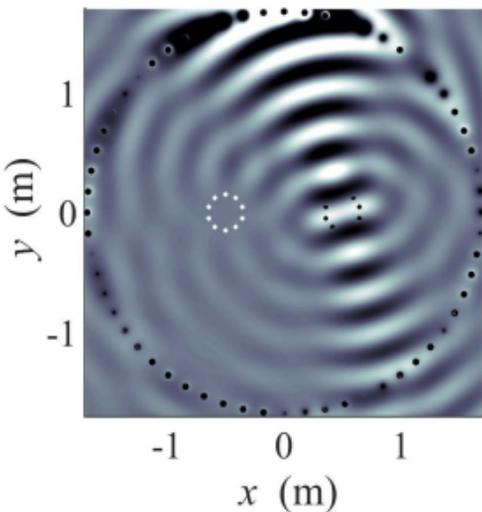


Fig4b

Circular array

Ar-shaped array

SHDMD

[Click here to access/download;Fig](#)

z

y (m)

1

0

-1

0.5

0

-0.5

-1

SVDMD

y (m)

1

0

-1

1

0.5

0

-0.5

-1

-1 0 1

x (m)

-1 0 1

x (m)

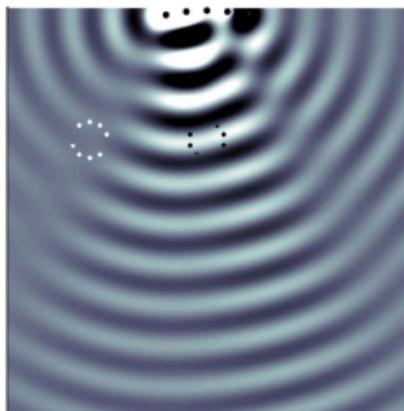
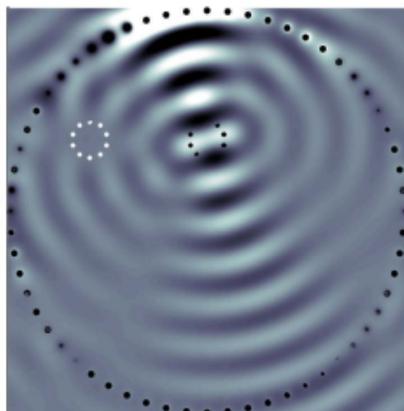
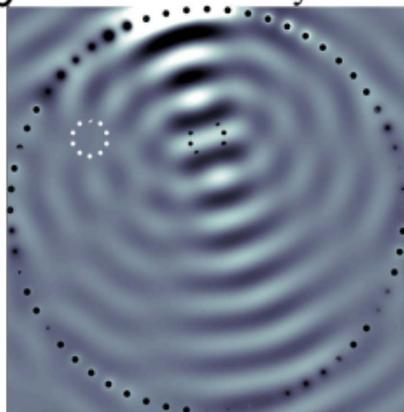
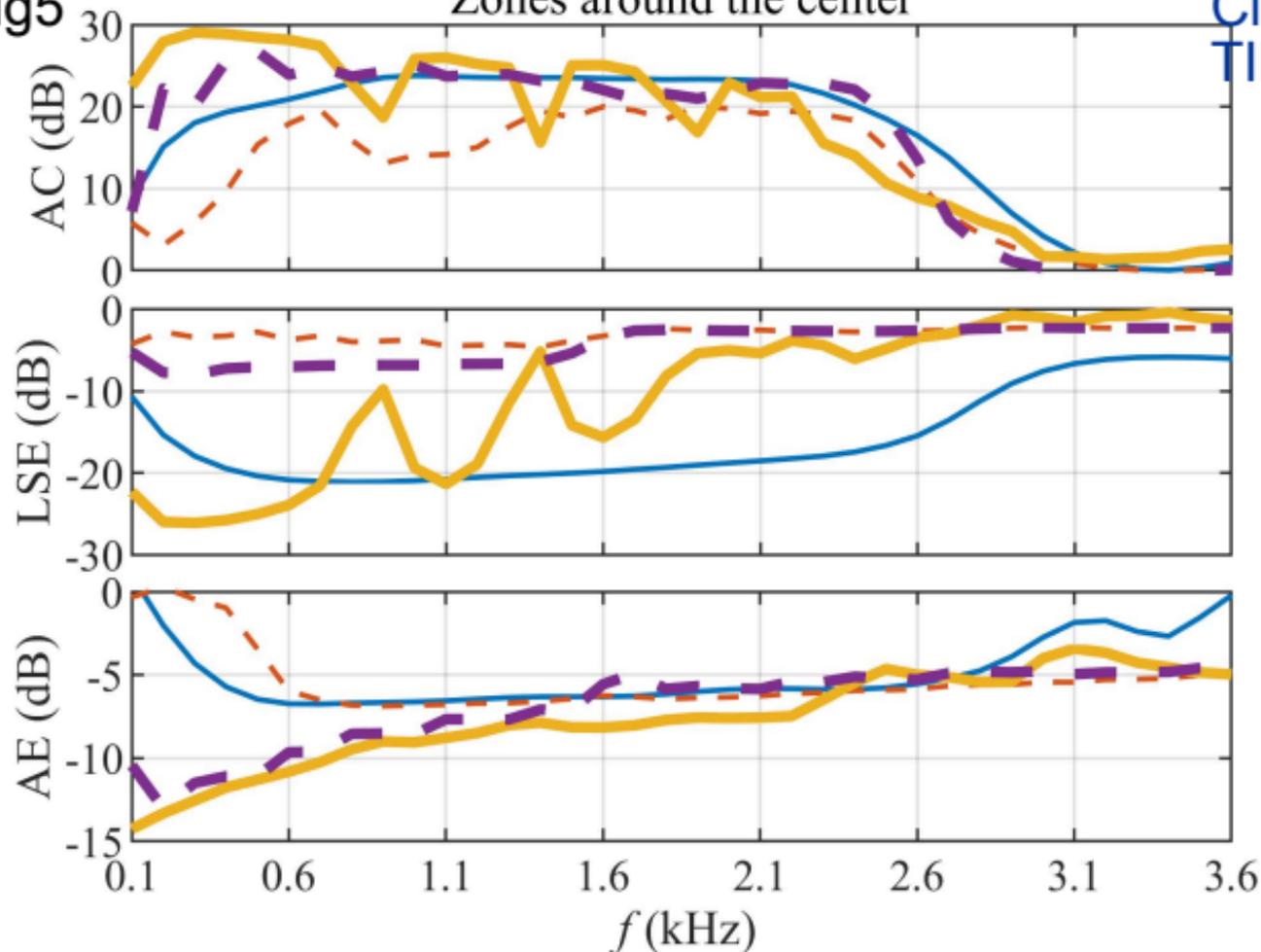


Fig5

Zones around the center



Click here to access/download;Figures (PDF, TIFF, EPS, PS, or JPEG only);Fig5.pdf

Zones in the near field

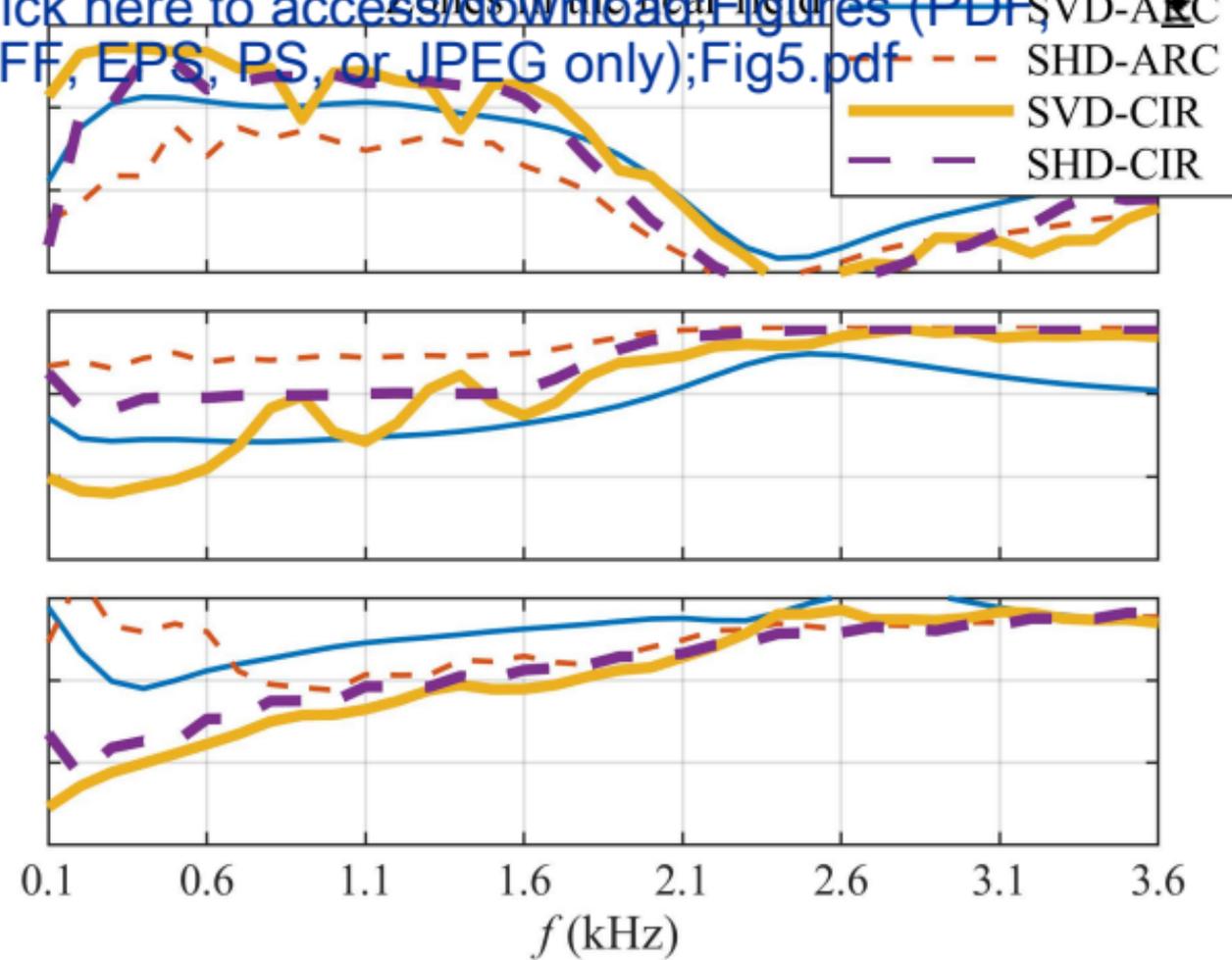
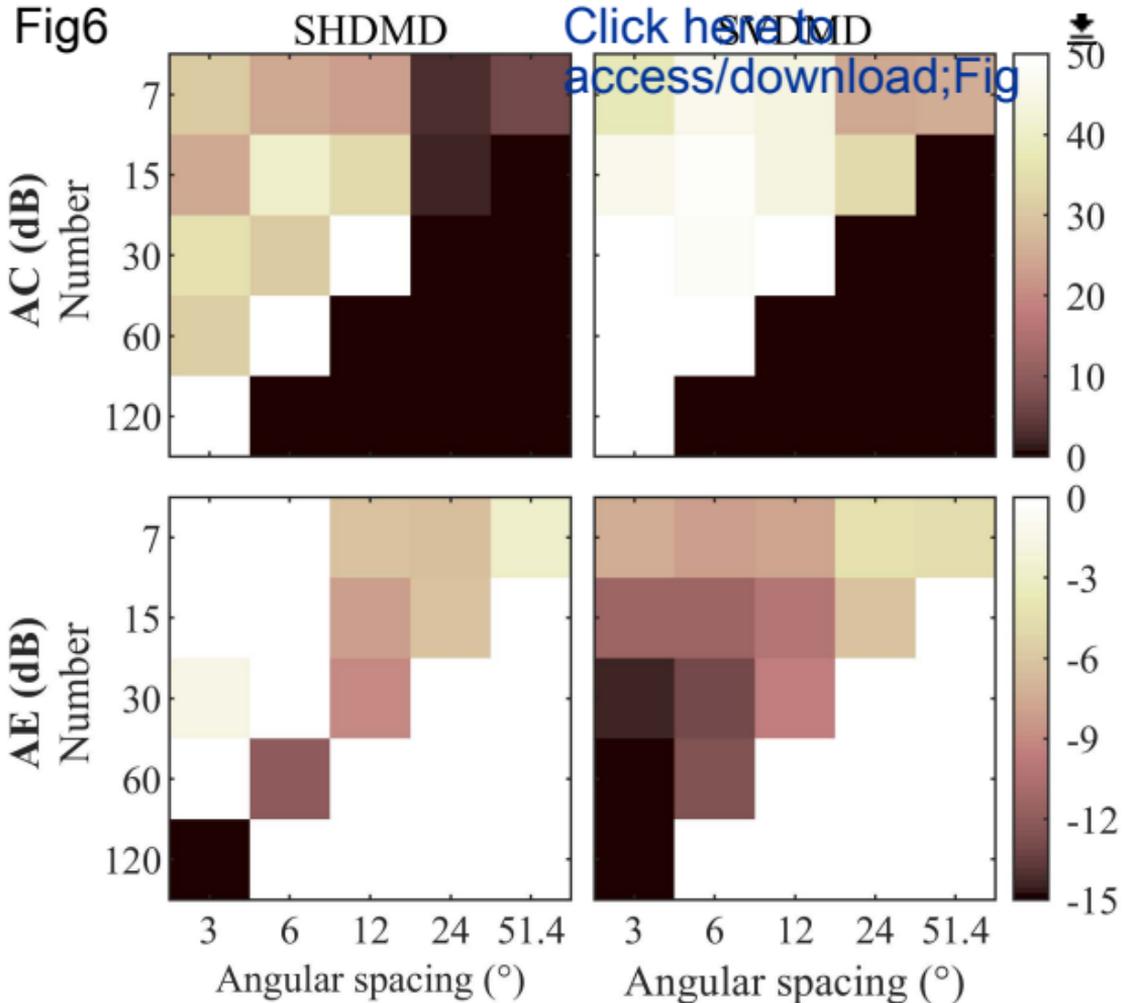


Fig6



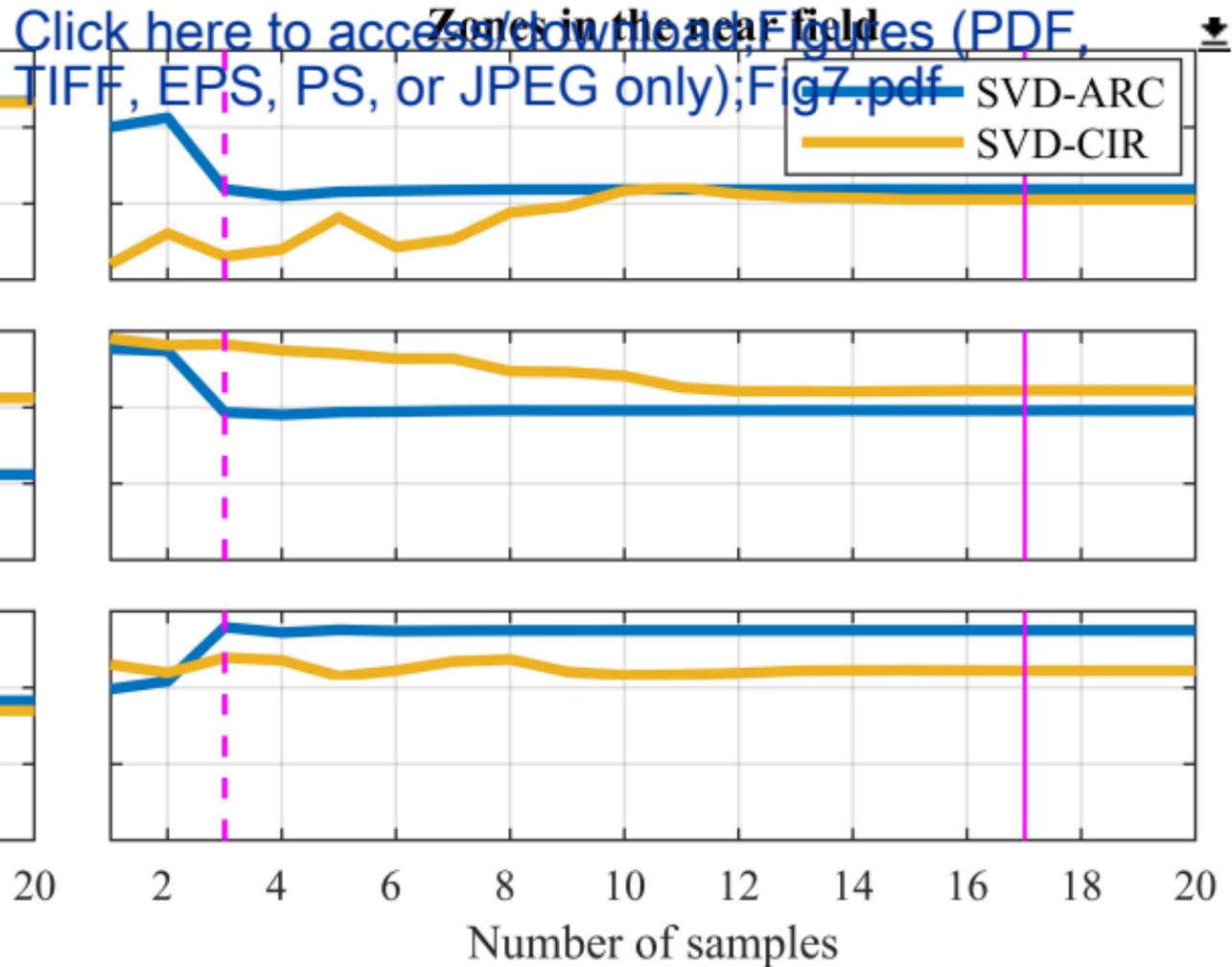
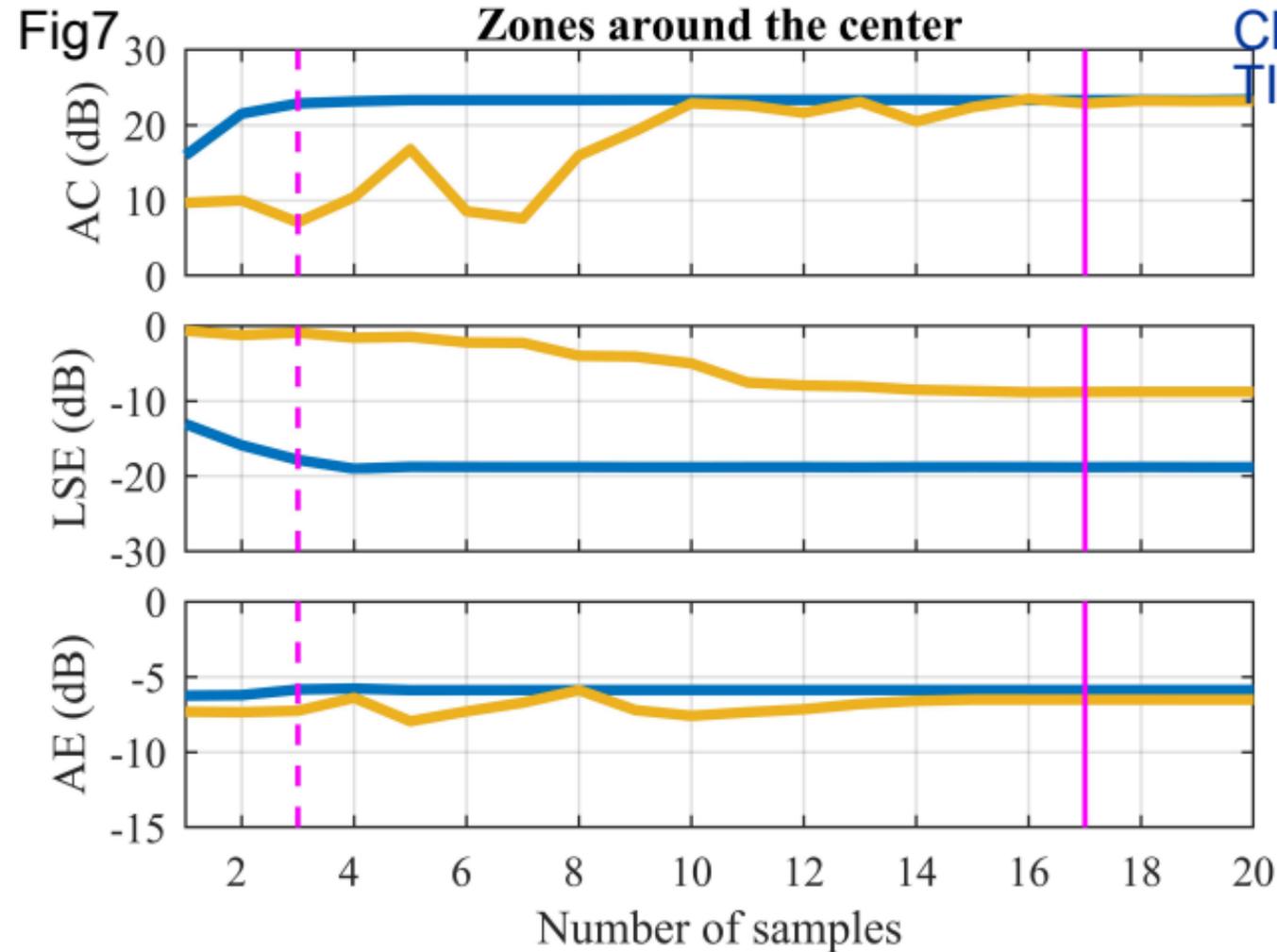


Fig8a

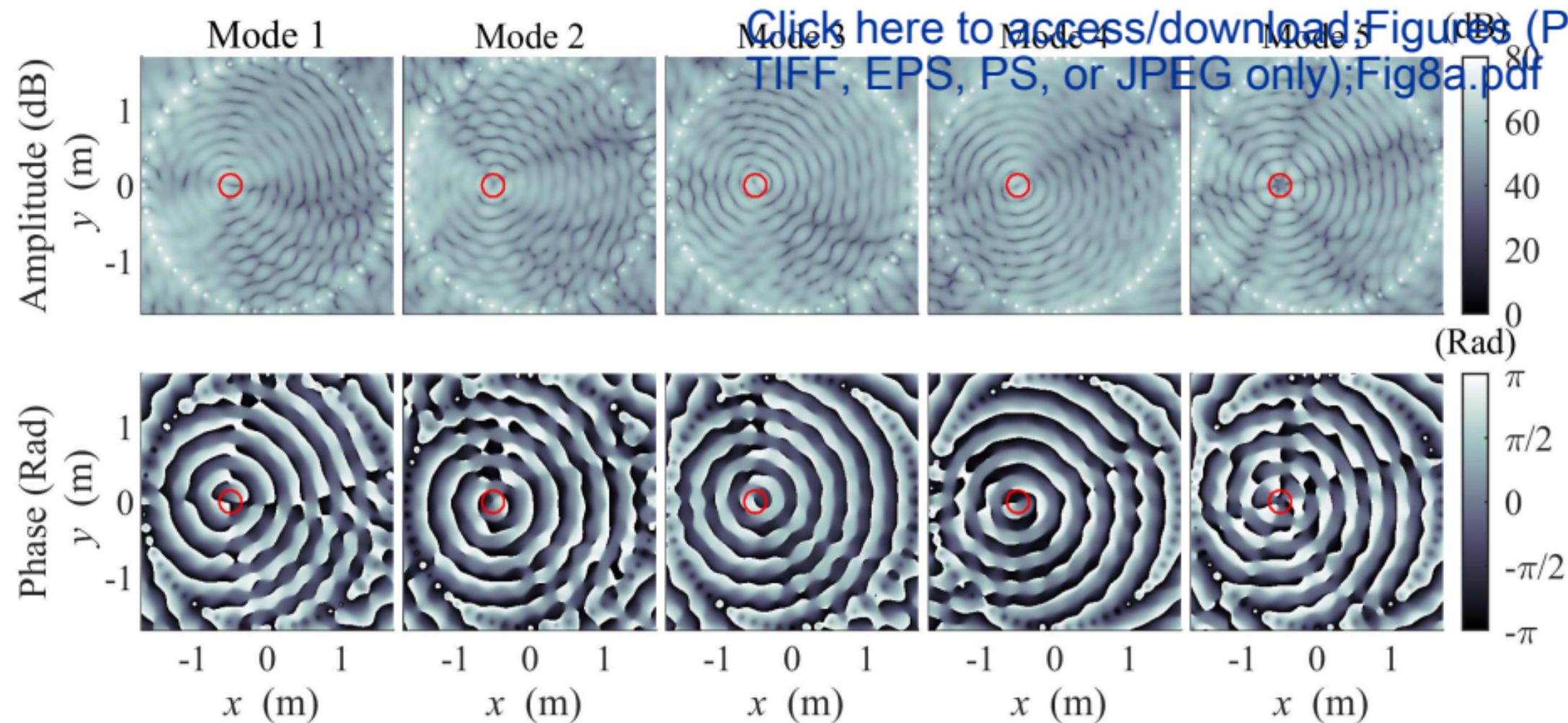


Fig8b

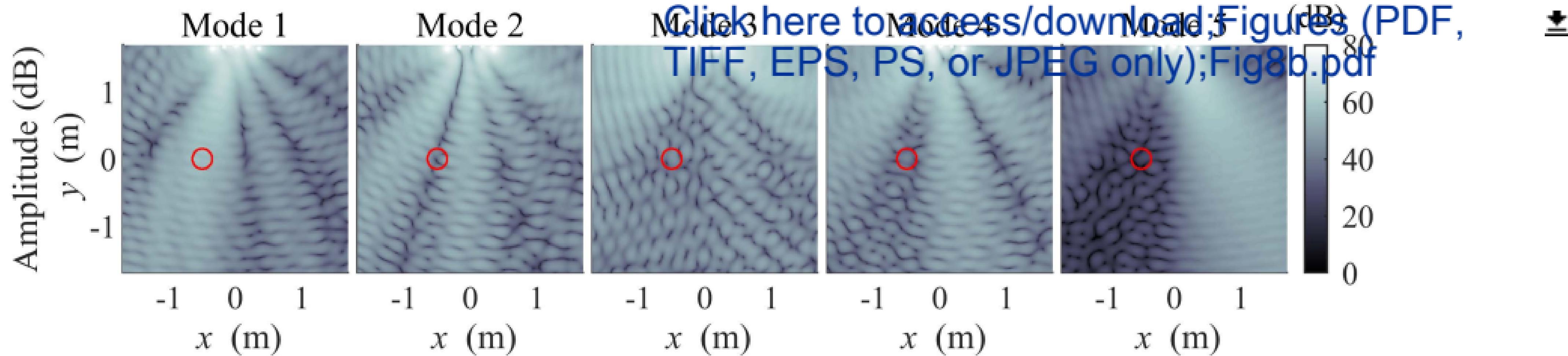
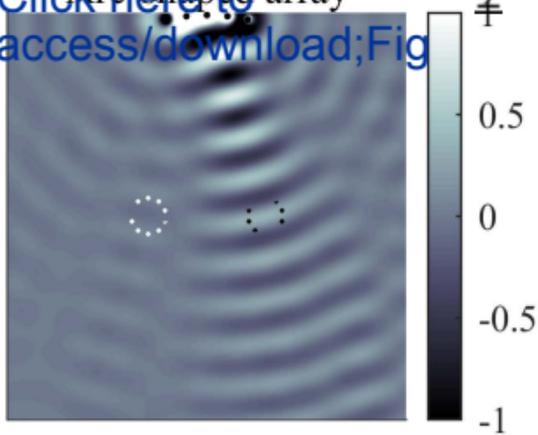
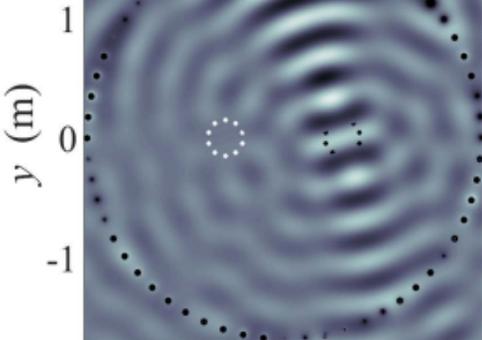


Fig9a

Circular array

Ar-shaped array
[Click here to
access/download;Fig](#)

SHDMD



SVDMD

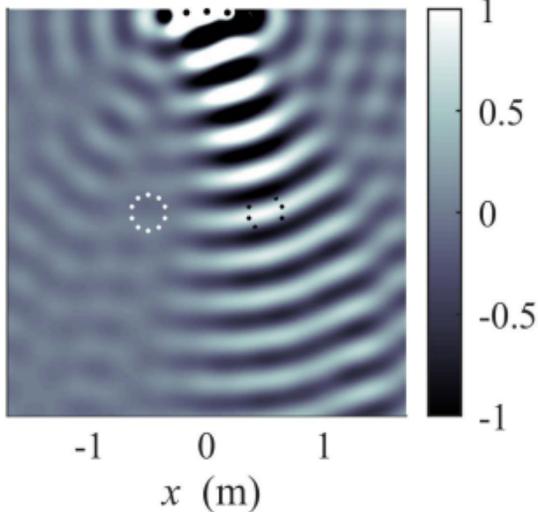
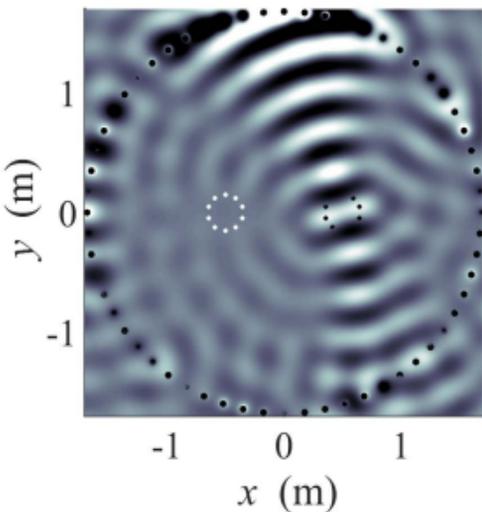


Fig9b

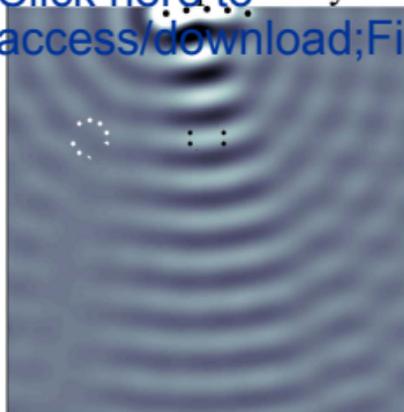
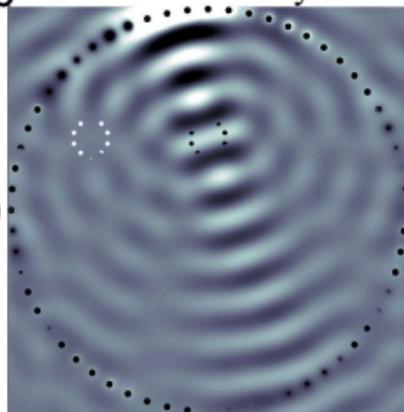
Circular array

Ar-shaped array

[Click here to access/download;Fig](#)

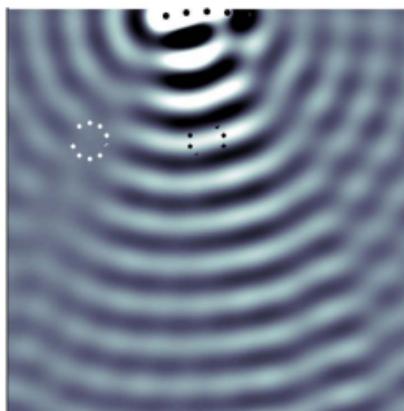
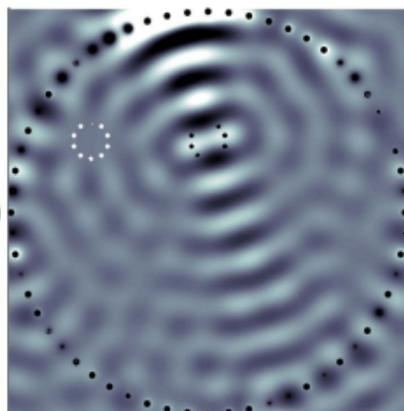
SHDMD

y (m)



SVDMD

y (m)

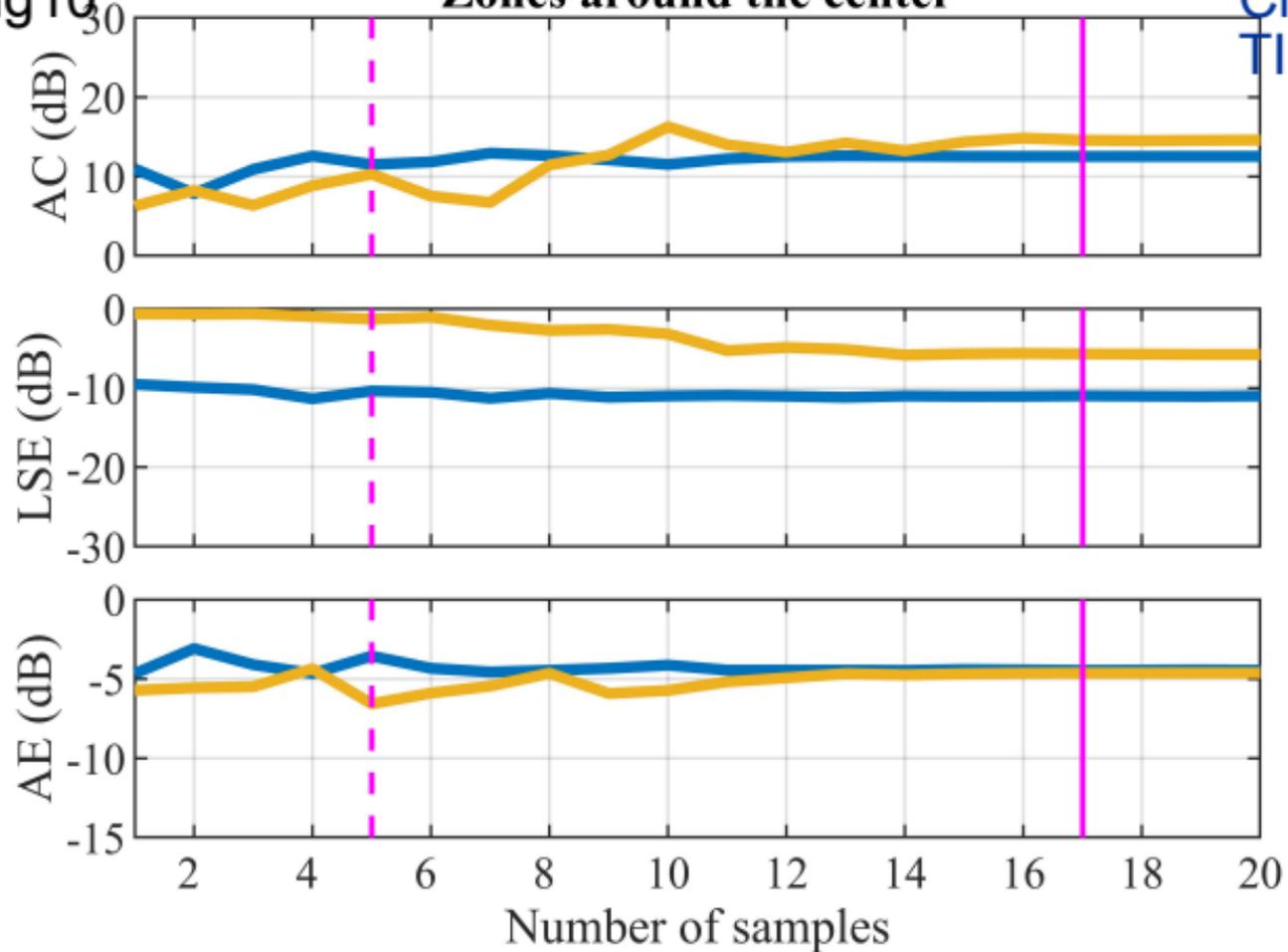


x (m)

x (m)

Fig10

Zones around the center



Click here to access/download/Figures (PDF, TIFF, EPS, PS, or JPEG only); Fig10.pdf

Zones in the near field

