MULTIZONE SOUNDFIELD REPRODUCTION IN REVERBERANT ROOMS USING COMPRESSED SENSING TECHNIQUES

Wenyu Jin, W. Bastiaan Kleijn

Victoria University of Wellington
New Zealand

ABSTRACT

We introduce a method of reproducing a multizone soundfield within the desired region in a reverberant room. It is based on determining the acoustic transfer function (ATF) between the loudspeaker over the reproduction region using a limited number of microphones. We assume that the soundfield is sparse in the *Helmholtz* solution domain and find the ATF using a compressed-sensing approach. This sparseness assumption facilitates the finding of the optimal characterization the original sound over the reproduction region based on scarce sound pressure measurements. The outcome of the first stage is then used to derive the optimal least-squares solution for the loudspeaker filter that minimizes the reproduction error over the whole reproduction region. Simulations confirm that the method leads to a significant reduction in the number of required microphones for accurate multizone sound reproduction, while it also facilitates the reproduction over a wide frequency range.

1. INTRODUCTION

Multizone soundfield reproduction is a technology that aims at providing an individual sound environment to each of a set of listeners using only loudspeakers. However, in practical scenarios, the performance of multizone sound reproduction techniques is commonly degraded by the effects of reverberation. In this paper, we extend the multizone soundfield reproduction to reverberant enclosures using a general Green's function modeling approach to characterize the room transfer function with a limited number of measurements.

The origin of work in this area can be traced to 2002, when [1] proposed an acoustic contrast control method to maximize the ratio of the mean square sound pressure in the so-called bright and dark zones. However, this approach is not able to reproduce a spatial soundfield. In 2008, Poletti [2] proposed a least squares method to generate a 2-D monochromatic multizone soundfield. In [3], a constrained optimization problem is formulated to control the sound in each zone individually. A method of describing the desired multizone soundfield as an orthogonal expansion of basis functions over the reproduction region was introduced in [4], in which the basis function set was formed using a modified Gram-Schmidt process, with as input a suitable set of planewave functions.

The fore-mentioned approaches do not take into account the reverberant environments in multizone sound reproduction. The reverberant case is difficult to handle because of the rapid variation of the ATF over the room [5]. The traditional approach for spatial sound reproduction in a reverberant setting is pressure matching, which equalizes the transfer functions over multiple points using least squares techniques [6] [7]. This technique leads to inferior performance away from the design points.

A technique to reconstruct a single-zone soundfield accurately over a control region in reverberant rooms was proposed in [8]. The method is based on mode matching. The authors also described a method for determining the ATF between each loudspeaker and the control region by sampling sound pressure at a number of positions. This work was extended to the multizone sound case in [9] using the translation of spatial harmonic coefficient between coordinate systems to achieve a global control of the multizone soundfield.

The methods of [8] and [9] require a relatively large number of measurements to estimate the ATFs of the loudspeakers with the needed accuracy. In addition, these approaches lead to problems for systems with a wide frequency range as the coefficients for the defined basis function cannot be determined at a number of frequencies due to the so-called large error scaling [8]. The solution for this in [8] is to double the number of pressure measurements and sample over two concentric circles about these frequencies, which makes the method cumbersome to implement for the wide band multizone sound reproduction. Furthermore, the sampling points are required to be evenly placed along the boundary of each selected zone, which confines the flexibility of microphone arrangement.

In this paper, we present a novel method of reproducing a desired multizone soundfield over the reproduction region in reverberant environments. Our approach consists of two distinct stages. We first identify the ATFs between the loudspeakers over the reproduction region using the concept of compressed sensing. The method is based on separating the actual loudspeaker ATF into a basic component that consists of the free-field Green's function and a corrective soundfield. We assume that the corrective soundfield is sparse in the *Helmholtz* solution domain, i.e. the corrective soundfield results from only a relatively small number of basis Helmholtz wavefields (e.g., planewaves). This sparseness assumption facilitates an accurate corrective soundfield reconstruction based on scarce sound pressure measurements at randomly-selected locations within the desired regions. (In this work, we assume that the microphones are location-informed). The estimate of the loudspeaker ATFs is then used to derive the optimal loudspeaker filter solution that minimizes the mean squared error (MSE) between the desired multizone sound and the actual reproduced sound over the entire reproduction region in the basis-function matching sense. Simulation results confirm that the approach provides accurate sound field reproduction with a smaller number of microphones than existing methods.

2. PROBLEM STATEMENT

We seek to control the reproduction of the desired 2-D multizone soundfield over each of a set of pre-defined zones within a reverberant enclosure. The theory we develop in this paper is readily extended to 3-D space.

As illustrated in Fig.1, the desired reproduction region $\mathbb{D}$ is the entire control zone of interest with a radius of r, which includes both the acoustic bright zone $\mathbb{D}_b$ and the quiet zone $\mathbb{D}_q$ of radius r_q. We

978-1-4799-2893-4/14/$31.00 ©2014 IEEE

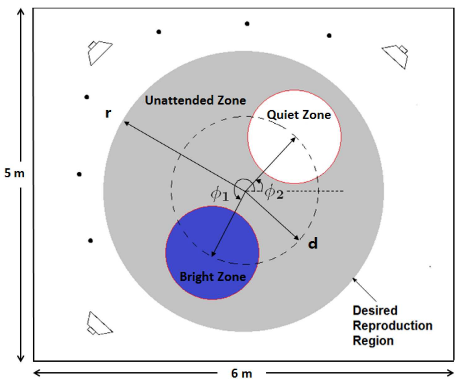


Fig. 1: The 2-D multizone soundfield reproduction over the desired reproduction region using Q loudspeakers in reverberant rooms.

define the remaining area in $\mathbb{D}$ as the unattended zone. The number of employed loudspeakers is Q and the q'th loudspeaker weight is denoted as $l_{q'}(k)$, where $k = 2\pi f/c$ is the wavenumber, f is the frequency and c is the speed of sound propagation. For simplicity, we assume that loudspeakers are cylindrical wave sources.

Our goal is to determine the loudspeaker filter weights required to reproduce the desired multizone sound in a reverberant room based on the characterization of the ATF for each of the loudspeaker. To evaluate the performance of our system we use the following three measures: i) the acoustic brightness contrast between the bright zone $\mathbb{D}_b$ and the quiet zone $\mathbb{D}_q$ [10], ii) the MSE between the desired sound $S^d(\mathbf{x}, k)$ (where $\mathbf{x}$ denotes an arbitrary spatial observation point) and the actually rendered sound $S^a(\mathbf{x}, k)$ over the zone $\mathbb{D}_b$, iii) the number of required microphones.

3. GREEN'S FUNCTION MODELING

The accurate characterization of the loudspeaker ATF is essential for the multizone sound reproduction system as large reproduction errors may result from small perturbations of the ATFs from the loudspeakers to the listener zones in scenarios when one zone is obscured by another [2]. In this section, we introduce a novel method to determine the ATF between the loudspeaker over the desired regions using the concept of compressed sensing. Compressed sensing [11] [12] is a signal processing technique for efficiently acquiring and reconstructing a signal from an under-determined linear system, assuming the observed phenomenon is known a *priori* to be sparse.

To estimate the ATF between the q'th loudspeaker $T_{q'}(\mathbf{x}, k)$, we first separate the actual soundfield into a basic component, the 2-D free-field Green's function and an unknown corrective soundfield $R_{q'}(\mathbf{x}, k)$

$$T_{q'}(\mathbf{x}, k) = \frac{i}{4} H_0^{(1)}(k\|\mathbf{Y}_{q'} - \mathbf{x}\|) + R_{q'}(\mathbf{x}, k), \tag{1}$$

where $\mathbf{Y}_{q'}$ represents the position of the q'th loudspeaker and $H_0^{(1)}(k\|\cdot\|)$ is the zeroth-order Hankel function of the first kind [13]. An arbitrary soundfield function satisfying the wave equation can be written as a superposition of a set of N solutions of the Helmholtz equation (the solutions can be non-orthonormal) [14]. Therefore, we can write $R_{q'} : \mathbb{R}^2 \times \mathbb{R} \mapsto \mathbb{C}$ as a weighted series of basis functions $\{F_n\}$

$$R_{q'}(\mathbf{x}, k) = \sum_{n=1}^{N} y_n F_n(\mathbf{x}, k), \ \mathbf{y} \in \mathbb{C}^N. \tag{2}$$

The basic principle of our method is to assume that $R_{q'}(\mathbf{x}, k)$ is sparse in the *Helmholtz* solution domain, i.e. the corrective soundfield results from only a relatively small number of basis Helmholtz wavefields. A reasonable assumption for the corrective soundfield is that it is a linear combination of normalized planewaves as scatterers and walls correspond to virtual sources that are far away from the observation points in many practical scenarios. Therefore, we refer to $\mathbf{y}$ in (2) as a K-sparse signal, which means that $\mathbf{y}$ has only K ($K \ll N$) non-zero entries at unknown locations while the other entries are zero (or very close to zero). The value of K depends on how complicated the reverberant environment is. Note that we only estimate the corrective part of $T_{q'}(\mathbf{x}, k)$ since the elimination of the free-field Green's function would generally increase the sparsity level of the signal in the planewave function domain and lead to a better estimation using a limited number of measurements.

Based on the assumptions listed above, we put forward a linear system

$$\mathbf{v} = \mathbf{\Phi}\mathbf{y}, \tag{3}$$

where the dictionary $\mathbf{\Phi}$ is an $m \times N$ sensing matrix ($N \gg m$) whose columns are the normalized measurement vectors of an over-complete set of *Helmholtz* wavefield functions $\{F_n\}$. $\mathbf{v}$ contains measurements of the desired corrective soundfield $R_{q'}(\mathbf{x}, k)$ that is the difference between the original measurement of the soundfield and the value of the free-field Green's function at m randomly chosen locations within the desired region. In our work, $\{F_n\}$ are selected to be N independent planewaves arriving from angles $\hat{\phi}_n = (n-1)\Delta\phi$, $n = 1, \cdots, N$ ($\Delta\phi = 2\pi/N$). The m measurements in $\mathbf{v}$ are the products of rows of the sensing matrix and the sparse signal $\mathbf{y}$: $\mathbf{v}_i = \langle \Phi_i, \mathbf{y} \rangle$. With the assumption of $\mathbf{y}$ being sparse, the ill-posedness of the under-determined system (3) is eliminated and we can find the optimal solution $\hat{\mathbf{y}}$ for the corrective sound over $\mathbb{D}_b$ and $\mathbb{D}_q$ based on $\mathbf{v}$ that contains a limited number of measurements. The key to a sparse signal approximation with high accuracy [12], is that the measured value is the linear projection of the sparse signal onto an incoherent basis (see, e.g., [15] for an explanation of incoherence). Our approach is consistent with this requirement as the random samplings of the sound pressure field in $\mathbf{v}$ is incoherent with the original basis of $\mathbf{y}$.

An l^p norm nonconvex optimization problem [16] is considered to produce an accurate estimate of $\mathbf{y}$ in (3) with scarce measurements, where $0 < p < 1$:

$$\min_{\mathbf{y}} \|\mathbf{y}\|_p^p, \ \text{s.t.} \ \mathbf{v} = \mathbf{\Phi}\mathbf{y}, \tag{4}$$

where $\|\mathbf{y}\|_p^p = \sum_{i=1}^{N} |y_i|^p$. In compressed sensing problem, the number of required measurements for exact reconstruction of the original signal with high probability is generalized as a function of the sparsity level K [12]. However, the value of K is always unknown in our work. In practice, $K \approx 2M+1$ (where $M = \lceil kr_q \rceil$ is the truncation length for each of the selected zones) can be used as an approximation with an error allowance due to the mode-limitedness of the soundfield within finite bright and quiet zones [17]. Therefore, the total number of degrees of freedom for (3) is decreased from N to K given that the set of basis wavefields with nonzero entries is well-fitted, which facilitates the reduction of the number of required microphones for the reconstruction with sufficient accuracy.

In any practical situation, the measured signal is corrupted by noise, so a more practical model is to reconstruct $\mathbf{y}$ from noisy under-sampled measurements $\mathbf{v}$. In such a case, the equality condition $\mathbf{v} = \mathbf{\Phi}\mathbf{y}$ should be relaxed to [18]

$$\min_{\mathbf{y}} \|\mathbf{y}\|_p^p, \ \text{s.t.} \ \|\mathbf{v} - \mathbf{\Phi}\mathbf{y}\|_2^2 \leq \epsilon. \tag{5}$$

We assume that the additive noise is Gaussian noise of zero mean and variance $\sigma^2(k)$, which makes the l^2-norm for the second term appropriate [18]. (5) can be converted to the following unconstrained form:

$$\min_{\mathbf{y}} \frac{1}{p}\|\mathbf{y}\|_p^p + \frac{\lambda}{2}\|\mathbf{v} - \mathbf{\Phi y}\|_2^2, \tag{6}$$

where λ is related to the error allowance ϵ. Experimental results suggest that choosing λ in the range $[\sigma^2(k)\sqrt{\log(N)}, \sigma^2(k)\sqrt{2\log(N)}]$ yields a fair reconstruction [18].

We can apply the regularized Iteratively Reweighted Least Squares (IRLS) algorithm [16] [18] to solve (6) and derive the optimal solution $\hat{\mathbf{y}}$ that estimates the corrective component of the soundfield generated by each loudspeaker in reverberant environments. Therefore, the ATF of the q'th loudspeaker over the desired regions can be approximated as follows according to (1) and (2):

$$T_{q'}(\mathbf{x}, k) = \frac{i}{4}H_0^{(1)}(k\|\mathbf{Y}_{q'} - \mathbf{x}\|) + \sum_{n=1}^{N}\hat{y}_n F_n(\mathbf{x}, k), \tag{7}$$

where $\hat{\mathbf{y}}$ is a sparse coefficient vector that has only m non-zero components.

4. ORTHOGONAL BASIS FUNCTION APPROACH

With the outcome of the first stage, which is the estimate of the ATFs between loudspeakers over the desired reproduction region, we now present an orthogonal basis function approach that enables us to derive the optimal loudspeaker filter weights for the desired multizone sound reproduction.

In the basis function approach, we express the desired multizone soundfield $S^d(\mathbf{x}, k)$, the actual reproduced soundfield $S(\mathbf{x}, k)$ and the ATF for a given loudspeaker $T_{q'}(\mathbf{x}, k)$ as a weighted series of basis orthonormal functions $\{G_i\}$. The orthonormal set of the solutions of *Helmholtz* equation $\{G_i\}$ is constructed using the approach proposed in [4]. Provided all sound sources lie outside $\mathbb{D}$, we can use the following representations:

$$S^d(\mathbf{x}, k) = \sum_{i=1}^{N'} A_i G_i(\mathbf{x}, k), \tag{8}$$

$$S(\mathbf{x}, k) = \sum_{i=1}^{N'} B_i G_i(\mathbf{x}, k), \tag{9}$$

$$T_{q'}(\mathbf{x}, k) = \sum_{i=1}^{N'} C_i^{q'} G_i(\mathbf{x}, k), \tag{10}$$

where $N' = 2M' + 1$ ($M' = \lceil kr \rceil$ is the truncation length for $\mathbb{D}$ [17]). The expansion coefficients for $S^d(\mathbf{x}, k)$ and $T_{q'}(\mathbf{x}, k)$ can be derived as $A_i = \int_{\mathbb{D}} S^d(\mathbf{x}, k)G_i^*(\mathbf{x}, k)W(\mathbf{x})d\mathbf{x}$ and $C_i^{q'} = \int_{\mathbb{D}} T_{q'}(\mathbf{x}, k)G_i^*(\mathbf{x}, k)W(\mathbf{x})d\mathbf{x}$ respectively, where the spatial weighting function $W(\mathbf{x})$ specifies the relative importance of the reproduction accuracy for each of the selected zones [4]. The reproduced soundfield $S(\mathbf{x}, k)$ resulting from the Q loudspeakers is equal to

$$S(\mathbf{x}, k) = \sum_{q'=1}^{Q} l_{q'}(k)T_{q'}(\mathbf{x}, k). \tag{11}$$

Substituting (9) and (10) into (11), the coefficients of the reproduced soundfield are related to $C_i^{q'}$ through $B_i = \sum_{q'=1}^{Q} l_{q'}(k)C_i^{q'}$.

With the reverberant setting, the main task of soundfield reproduction is to choose filter weights $l_{q'}$ to minimize the weighted squared error over $\mathbb{D}$:

$$\eta = \int_{\mathbb{D}} \|S(\mathbf{x}, k) - S^d(\mathbf{x}, k)\|^2 W(\mathbf{x})d\mathbf{x}. \tag{12}$$

Substituting (8) and (9) into (12) and use the orthonormal property of $G_i(\mathbf{x}, k)$ over $\mathbb{D}$: $\int_{\mathbb{D}} G_i(\mathbf{x}, k)G_j^*(\mathbf{x}, k)W(\mathbf{x})d\mathbf{x} = \delta_{ij}$, we see that the weighted squared error reduces to

$$\eta = \sum_i |B_i - A_i|^2. \tag{13}$$

The optimal least-square solution for (13) is expressed in terms of the basis function coefficients. Let us define the vector of loudspeaker filter weights $\mathbf{l} = [l_1(k), \ldots, l_Q(k)]^T$, the vector of the coefficients of the reproduced soundfield $\mathbf{B} = [B_1, \ldots, B_{N'}]$ and the matrix of the coefficients of the room responses of all loudspeakers

$$\mathbf{C} = \begin{bmatrix} C_1^1 & \ldots & C_1^Q \\ \vdots & \ddots & \vdots \\ C_{N'}^1 & \ldots & C_{N'}^Q \end{bmatrix}. \tag{14}$$

Then, we have $\mathbf{B} = \mathbf{Cl}$. Additionally, we define the vectors of the coefficients of the desired multizone soundfield, $\mathbf{A} = [A_1, \ldots, A_{N'}]$.

Then we write (13) into matrix form and add a regularization term, both to constraint loudspeaker effort and ensure a robust solution:

$$\gamma = (\mathbf{B} - \mathbf{A})^H(\mathbf{B} - \mathbf{A}) + \gamma\|\mathbf{l}\|^2, \tag{15}$$

where $\gamma \geq 0$. The optimal solution is then

$$\mathbf{l} = (\mathbf{C}^H\mathbf{C} + \gamma\mathbf{I})^{-1}\mathbf{C}^H\mathbf{A}, \tag{16}$$

where $\mathbf{I}$ is a unitary matrix. Once $(\mathbf{C}^H\mathbf{C} + \gamma\mathbf{I})^{-1}\mathbf{C}^H$ is computed for the complex acoustical environment, the reproduced soundfield can be changed by modifying $\mathbf{A}$, which specifies the desired multizone soundfield.

5. RESULTS AND DISCUSSION

Our main objective is to determine the ATFs between the loudspeakers over the selected zones using as few microphones as possible, so that it leads to an accurate multizone soundfield reproduction within $\mathbb{D}$. To define our multizone sound requirements, one bright zone and one quiet zone are selected.

The reverberant room is rectangular (size 6 m $\times$ 5 m) with a wall absorption coefficient of 0.3. $\mathbb{D}$ has a radius of $r = 1$ m with its center located at $(2, 2.5)$ and the 39 loudspeakers are evenly distribute along a concentric circle with a radius of 1.5 m. We use the image source method [19] to simulate the soundfield created by the loudspeaker in the reverberant room. In the simulations below, a total of 60 image sources are included. The centers of $\mathbb{D}_b$ and $\mathbb{D}_q$ lie on a circle of radius $d = 0.6$ m. We assume c is 340 m/s in our simulations. The target bright and quiet zones are located at $\phi_1 = 225^\circ$ and $\phi_2 = 45^\circ$ respectively with $r_q = 0.3$ m as shown in Fig. 1. In order to numerically observe how the number of microphones used would affect the estimation accuracy of the desired ATFs of the loudspeakers, we randomly select m locations within each zone and measure the value of the corrective soundfield at those positions. Note that our approach allows the flexibility of the microphone arrangement within the desired region, which makes it more practical in a real-world implementation.

We first illustrate the ATF estimation of a desired loudspeaker placed outside $\mathbb{D}$ at $f = 1$ KHz. In Fig. 2 we demonstrate the performance with the following three different settings: noiseless

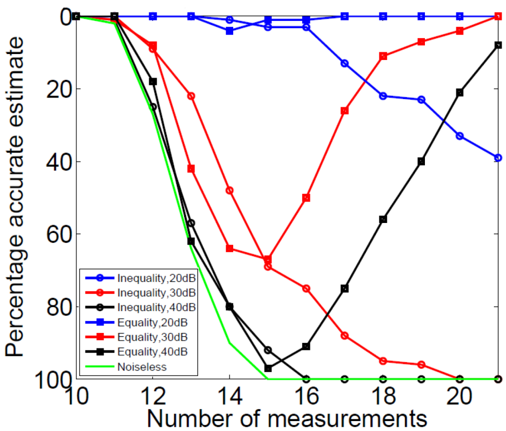


Fig. 2: The estimated performance of the desired loudspeaker ATF is plotted as a function of the number of microphones used at 1 KHz with different settings. We run 100 trials for each iteration with an assigned value of m.

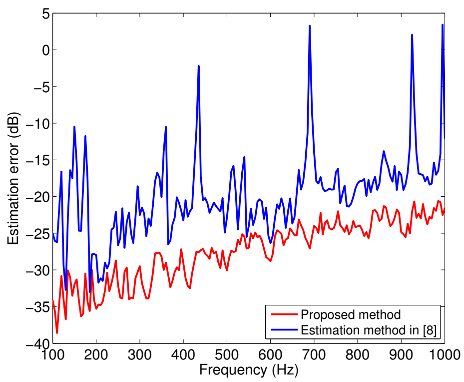
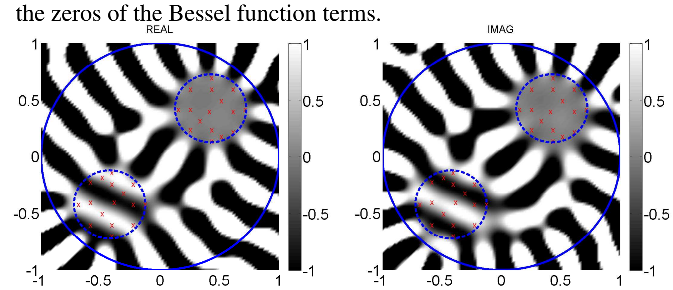


Fig. 3: Wide band desired ATF estimation with 20 noisy pressure samples, using the proposed method and the method in [8]. Estimation error curves have been averaged over 50 trial runs.

setting, noisy setting with optimization constraint of $\|\mathbf{v} - \mathbf{\Phi y}\| \leq \epsilon$ (inequality constraint) and with optimization constraint of $\mathbf{v} = \mathbf{\Phi y}$ (equality constraint). For the approach of regularized IRLS in (6), the sparsity promoting norm $p = 0.4$. With the ideal noiseless measurements, we set $\lambda = 0$ in (6). From Fig. 2 we see that the use of $m = 14$ microphones for each selected zone would generally facilitate an accurate estimate of the desired loudspeaker ATF with a probability of over 90%. Note that we define the reconstruction of the desired ATF over the two selected zones with an estimation error no greater than -20 dB as an accurate estimate. In contrast, the estimation method proposed in [8] requires approximately 20 microphones to obtain the same level of accuracy with the same settings. The estimated performance is also plotted for several additive noise SNRs while the values of λ are selected accordingly. The performance of reconstructions with the inequality constraint is superior to the estimation using the equality constraint. We can observe a decline in the estimation performance after it reaches a peak with the number of measurements in the equality case due to the increasing amount of noise added with the measurements. The results show that at least 30 dB SNR is required for an accurate reconstruction of the desired ATF at 1 KHz using a limited number of noisy measurements. Naturally, the estimation performance improves as the level of the added noise decreases.

In Fig. 3, we compare our method with the estimation method proposed in [8] in terms of the wide band desired ATF estimation in the frequency range 100 Hz to 1 kHz with noisy pressure samples (the additive noise SNR is 30 dB). We use 20 noisy measurements for each zone and the microphone locations are again randomly selected for our method while the measurements are sampled along the the boundaries for the method proposed in [8]. As we can see, our method consistently outperforms the estimation method proposed in [8], typically up to 8 dB. This would generally lead to better performance in terms of the reproduction of the desired soundfield as the estimates of the loudspeaker ATFs are directly used to design the optimal loudspeaker filter gains. More importantly, the general trend of our proposed method is that the error smoothly increases with frequency. This trend is due to the increase in the complexity of the desired ATF with frequency. In contrast, we can observe obvious peaks in the blue curve. These peaks occur in the vicinity of the zeros of the Bessel function terms.



(a) (b)

Fig. 4: Reproduction of the desired multizone sound using 16 microphones for each selected zone in a reverberant room. (a) and (b) demonstrate the real part and imaginary part of the soundfield respectively. The red crosses represent the positions of the microphones.

Fig. 4 demonstrates the reproduction of the desired multizone sound using 16 noiseless measurements for each selected zone at 1 KHz. The desired soundfield over $\mathbb{D}_b$ is a planewave arriving from 60°. This is a challenging multizone sound reproduction scenario due to the occlusions of sound between D_b and D_q. The values of weighting function $W(\mathbf{x})$ assigned to $\mathbb{D}_b$, $\mathbb{D}_q$ and the unattended zone are 1, 5 and 0.01 respectively. The average estimation error of the ATFs between all loudspeakers over the two selected zones is -35 dB. We can see that the reproduced soundfield matches the desired multizone sound well. The acoustic contrast between $\mathbb{D}_b$ and $\mathbb{D}_q$ is 20.06 dB and the MSE over $\mathbb{D}_b$ is -13.81 dB. We can observe high-amplitude sound outside $\mathbb{D}$ due to a high level of loudspeaker array effort. This can be alleviated by increasing the value of γ in (15) at the expense of decreasing the overall reproduction accuracy [20].

6. CONCLUSION

We proposed a general Green's function modeling approach for determining the ATFs between loudspeakers over the desired regions in reverberant environments using the concept of compressed sensing. The proposed approach facilitates the practical implementation in the following aspects: i) a significant reduction in the number of required measurements for accurate characterization of the loudspeaker ATFs, ii) steady performance in the multizone sound reproduction over a wide frequency range, iii) flexibility of the microphone arrangement. The optimal loudspeaker filter solution for the desired multizone sound reproduction can then be derived based on the estimate of loudspeaker ATFs using an orthogonal basis function approach. The results confirm the accurate multizone sound reproduction even under the extreme scenario that strong sound occlusions exist between the selected bright and quiet zones.

7. REFERENCES

[1] J.W. Choi and Y.H. Kim, "Generation of an acoustically bright zone with an illuminated region using multiple sources," *Journal Acoustics Society America*, vol. 111, pp. 1695–1700, 2002.

[2] M. Poletti, "An investigation of 2D multizone surround sound system," *Proc. AES 125th Convention Audio Eng. Society*, 2008.

[3] T. Betlehem and P. Teal, "A constrained optimization approach for multizone surround sound," *Proc. IEEE Int Conf. Acoust. Speech Signal Process.*, pp. 437–440, 2011.

[4] W. Jin, W. B. Kleijn, and D. Virette, "Multizone soundfield reproduction using orthogonal basis expansion," *Proc. IEEE Int Conf. Acoust. Speech Signal Process.*, pp. 311–315, 2013.

[5] J. Mourjopoulos, "On the variation and invertibility of room impulse response functions," *J. Sound Vib.*, vol. 102, pp. 217–228, 1985.

[6] P. A. Nelson and F. Orduna-Bustamante, "Inverse filter design and equalization zones in multichannel sound reproduction," *IEEE Trans. Speech, Audio and Language Process.*, vol. 3, pp. 185–192, 1995.

[7] O. Kirkeby and P. A. Nelson, "Digital filter design for inversion problems in sound reproduction," *J. Audio Engineering Society*, vol. 47, pp. 583–595, 1999.

[8] T. Betlehem and T. D. Abhayapala, "Theory and design of sound field reproduction in reverberant rooms," *J. Acoust. Soc. Amer.,*, vol. 117, pp. 2100–2111, Apr. 2005.

[9] Y. Wu and T. Abhayapala, "Spatial multizone soundfield reproduction: Theory and design," *IEEE Trans. Acoust.,Speech, Signal Processing*, vol. 19, no. 6, pp. 1711–1720, August 2011.

[10] J.-H. Chang, C.-H. Lee, J.-Y. Park, and Y.-H. Kim, "A realization of sound focused personal audio system using acoustic contrast control," *J. Audio Engineering Society,*, vol. 125, pp. 2091–2097, Apr. 2009.

[11] D. L. Donoho, "Compressed sensing," *IEEE Trans. on Information Theory*, vol. 52, pp. 1289–1306, 2006.

[12] E. J. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. on Information Theory*, vol. 52, 2006.

[13] D. Colton and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory*, Springer, New York, 1998.

[14] E. G. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography*, Academic, New York, 1999.

[15] E. Candes and J. Romberg, "Sparsity and incoherence in compressive sampling," *Inverse Problems*, vol. 23, pp. 969–985, 2007.

[16] R. Chartrand, "Exact reconstruction of sparse signals via nonconvex minimization," *IEEE Signal Processing Letters*, vol. 14, pp. 707–710, Oct. 2007.

[17] R. A. Kennedy, P. Sadeghi, T. D. Abhayapala, and H. M. Jones, "Intrinsic limits of dimensionality and richness in random multipath fields," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2542–2556, June 2007.

[18] Rafael E. Carrillo and K.E. Barner, "Iteratively re-weighted least squares for sparse signal reconstruction from noisy measurements," in *Information Sciences and Systems, 2009. CISS 2009. 43rd Annual Conference on*, 2009, pp. 448–453.

[19] J. Allen and D. Berkley, "Image method for efficiently simulating small room acoustics," *J. Acoust. Soc. Amer*, vol. 65, pp. 943950, 1979.

[20] F.M. Fazi and P.A. Nelson, "The ill-conditioning problem in sound field reconstruction," *Proc. 123rd Audio Engineering Society (AES) Convention*, 2007.