ADAPTIVE REVERBERATION CANCELATION FOR MULTIZONE SOUNDFIELD REPRODUCTION USING SPARSE METHODS

Wenyu Jin

Huawei Technologies Duesseldorf GmbH
German Research Center, Munich

ABSTRACT

The necessity of a large number of loudspeaker-microphone channels for the existing sound rendering systems complicates the application of multizone soundfield reproduction in reverberant environments. We introduce an adaptive reverberation cancelation method for multizone soundfield reproduction using sparse methods. The reproduced soundfield is described as a weighted series of orthonormal basis functions over the desired reproduction region, which is then used to adaptively equalize the desired multizone soundfield in terms of the basis function coefficients. The sparse methods result in a significantly reduced number of the required microphones for the measuring process of the reproduced soundfield. Simulation results verify the efficient room reverberation compensation for desired multizone soundfield reproduction. The proposed method also facilitates reproduction over a wide frequency range.

1. INTRODUCTION

Reproduction of a desired multizone soundfield over a region of interest has drawn the attention of researchers in recent years [1–7]. However, the majority of existing works in this area do not take into account the reverberant environments that practical multizone sound reproduction systems will encounter. Reverberation compensation process is difficult to handle due to the unknown reverberant room channel and the large number of loudspeakers and microphones required by existing soundfield reproduction systems. In this paper, we propose an efficient adaptive reverberation cancelation system for multizone soundfield reproduction using sparse microphone measurements.

To equalize the room reverberation, the inverse of the room response must be applied to loudspeaker driving signals. The traditional approach for spatial sound reproduction in a reverberant setting is pressure matching, which equalizes the transfer functions over a discrete set of points [8] [9]. This technique leads to poor performance in regions further away from the design points and inaccurate reproduction [10]. In 2005, Betlehem et al. [10] proposed a technique based on mode matching to reproduce a single-zone soundfield accurately over the entire control region in reverberant rooms. An approach of reproducing a multizone soundfield within a desired region using sparse methods was introduced in [11] [12]. Comparing with the method in [10], a reduced number of randomly placed measurements were employed to estimate the transfer functions from the loudspeakers over the desired region in reverberant environments based on sparse approximation. The estimates were then used to derive the optimal least-squares solution for the loudspeaker filter gains. For these approaches, a prior measurement of the room transfer function for all the employed loudspeakers was needed. This is time-consuming to implement in practice and its performance is vulnerable to any changes in the ambient environment conditions during the measurement process.

Wave Domain Adaptive Filtering (WDAF) is a more practical approach to the application of reverberation cancelation in soundfield reproduction. Initially proposed by Buchner et al. [13] [14], it has been introduced to active listening room compensation in Wave Field Synthesis systems [15–17]. The wave-domain representation of the soundfield was described using transformations on the microphone array input and the loudspeaker output respectively [15]. The work by Schneider et al. [16] [18] has further reduced the computational complexity of the basic WDAF adaptation process. This was achieved by considering that the dominant couplings between the soundfield modes limit only in the vicinity of the diagonal of the linear transformations and neglecting the weaker ones. A similar adaptive method was proposed in [19], in which the reverberant soundfield was described and estimated by exploiting the orthogonality of the Fourier-Bessel expansion to simplify the listening room compensation problem within a region of interest. The key of the work in [18] [19] is that different soundfield coefficients with different indices do not interact with each other in the so-called mode-domain.

In this paper, we use the inspiration from [18] [19] to propose an adaptive reverberation cancelation system for multizone soundfield reproduction using sparse microphone measurements. The proposed approach expresses the soundfield as an orthonormal basis function expansion in the space-frequency domain over the desired reproduction region. We consider the reproduced soundfield as a linear transformations of the desired soundfield. We then introduce the adaptive channel estimation process using sparse methods to identify these transformations and derive the required loudspeaker updating signals. Finally, simulation results are presented for the desired multizone soundfield rendering system in a reverberant environment.

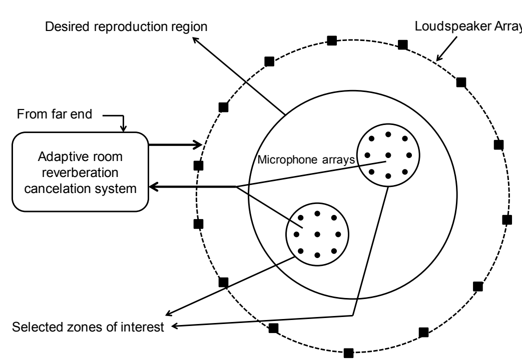


Fig. 1: Overview of the active reverberation cancelation system including the loudspeaker array and microphones configurations.

2. PROBLEM STATEMENT

The adaptive reverberation cancelation system aims to rectify the reverberation effects based on iterative feedback from sparse microphone measurements. Note that we mainly focus on the 2-D (height invariant) case and the theory developed in this paper is readily extended to 3-D space. The structure of the proposed active multizone soundfield reproduction system is shown in Fig. 1. It consists of a circular array of Q loudspeakers and M microphones. To define our multizone soundfield requirements, one bright zone $\mathbb{D}_b$ and one quiet zone $\mathbb{D}_q$ are used in this work. We define the remaining area in the desired reproduction region $\mathbb{D}$ as the unattended zone $\mathbb{D}_u$. The loudspeakers are placed outside the desired reproduction region $\mathbb{D}$ while the microphones are randomly placed within the selected zones of interest to iteratively record the generated soundfield at the assigned position.

2.1. Soundfield Basis Expansion

The received measurements at the microphones can be expressed in matrix form as

$$\mathbf{v}(k) = \mathbf{C}(k)\mathbf{l}(k), \tag{1}$$

where $\mathbf{l}(k) = [l_1(k), \dots, l_Q(k)]^T$ are the loudspeaker driving signals, $\mathbf{v}(k) = [v_1(k), \dots, v_M(k)]^T$ are the microphone measurements, and $\mathbf{C}(k)$ represents the channel between the (m,q)th microphone-loudspeaker pair at the wavenumber k. Note that we can separate the channel effects $\mathbf{C}(k)$ into the direct and reverberant path, $\mathbf{C}(k) \equiv \mathbf{C}^{\text{direct}}(k) + \mathbf{C}^{r}(k)$, where $\mathbf{C}^{\text{direct}}(k)$ and $\mathbf{C}^{r}(k)$ represent the direct and reverberant channels between the (m,q)th microphone - loudspeaker pair. In [3], we constructed an orthonormal set of basis functions $\{G_n\}_{n\in\mathbb{A}}$ (where $\mathbb{A}$ is a set of indices) by implementing a modified Gram-Schmidt process on planewave functions arriving from various angles. To define the basis functions, we use the weighted inner product

$$\langle Y_1, Y_2\rangle_w = \int_{\mathbb{D}} Y_1(\mathbf{x})Y_2^*(\mathbf{x})w(\mathbf{x})d\mathbf{x}, \tag{2}$$

where the weighting function $w(\mathbf{x})$ specifies the relative importance of the reproduction accuracy for each point in space [3]. Therefore, we express the measurements in (1) as

$$v_m(k) = \sum_{n=1}^{N} b_n(k)G_n(\mathbf{x}_m, k), \tag{3}$$

where $b_n(k)$ are the coefficients for the reproduced soundfield and $\mathbf{x}_m$ represents the mth microphone location. Note that N is set to be sufficiently large. The coefficients $b_n(k)$ can be derived from the soundfield measurements $\mathbf{v}(k)$, as described in the next section.

2.2. Soundfield Characterization Using Sparse Methods

In this section, we apply a sparse approximation method similar to [11] to calculate $b_n(k)$ from the randomly-placed measurements $\mathbf{v}_m(k)$ within the selected zones of interest.

The basic principle of our method is to assume that the reproduced soundfield $S(\mathbf{x},k)$ results from only a small number of basis Helmholtz solutions. Based on this assumption, we consider the following l^p norm (where $0 < p < 1$) non-convex optimization problem

$$\min_{\mathbf{y}} \|\mathbf{y}\|_p^p, \text{ s.t. } \|\mathbf{v} - \mathbf{\Phi}\mathbf{y}\|^2 \le \epsilon, \tag{4}$$

where $\mathbf{y}$ is the basis function coefficient set, the dictionary $\mathbf{\Phi}$ is an $M \times N$ sensing matrix ($N \gg M$) whose columns contain the values of $\{G_n(\mathbf{x},k)\}_{n\in\mathbb{A}}$ at M locations and $\mathbf{v}$ is an $M \times 1$ observation vector which contains the values of the actual reproduced soundfield $S(\mathbf{x},k)$ at M randomly chosen locations within the desired region. The error ϵ is related to the he additive complex Gaussian noise level. Let $\mathbf{y}$ be a sparse signal, i.e., $\mathbf{y}$ has a limited number of non-zero entries at unknown locations. Therefore, we can apply the regularized Iteratively Reweighted Least Squares (IRLS) algorithm [20] [21] to solve (4) and derive the optimal estimator $\hat{\mathbf{y}}$ that characterizes the reproduced soundfield in reverberant environments:

$$\hat{S}(\mathbf{x},k) = \sum_{n=1}^{N} \hat{y}_n G_n(\mathbf{x},k), \tag{5}$$

where $\hat{\mathbf{y}}$ has only m' ($m' \le M$) non-zero components and can be used as an estimate of the basis function coefficients $b_n(k)$.

Overall, we formulate the calculation of the soundfield coefficients $b_n(k)$ based on the soundfield measurements in (1) in the following matrix form

$$\mathbf{b}(k) = \mathbf{T}\mathbf{C}(k)\mathbf{l}(k) = \mathbf{T}\mathbf{v}(k), \tag{6}$$

where $\mathbf{b}(k) = [b_1(k), \dots, b_N(k)]$, $\mathbf{T}$ is generalized as a transformation matrix ($N \times M$) that expresses the relationship of $\mathbf{b}(k)$ and $\mathbf{v}(k)$, which can be seen as the projection from the microphone measurements onto the subspace $V \in \mathbb{C}^N$ spanned by the orthonormal set $\{G_n\}_{n\in\mathbb{A}}$.

Therefore, the objective of the adaptive reverberation cancelation system in the following section is to match the reproduced soundfield with the desired multizone soundfield and to minimize

$$\|\mathbf{b}(k) - \mathbf{b}^d(k)\|^2 = \|\mathbf{T}\mathbf{C}(k)\mathbf{l}(k) - \mathbf{b}^d(k)\|^2. \tag{7}$$

3. ADAPTIVE REVERBERATION CANCELATION

In this section, we describe the modeling of the unknown room channel, the channel estimation processes in the basis-function domain and the computation of the loudspeaker updating signals required for active reverberation cancelation.

3.1. Room Channel Modeling

The desired multizone soundfield $S^d(\mathbf{x},k)$ and the actual reproduced soundfield in a reverberant room $S(\mathbf{x},k)$ can be characterized by $\mathbf{b}^d(k)$ and $\mathbf{b}(k)$ that represents the respective coefficient sets of the orthonormal basis functions $\{G_n\}_{n\in\mathbb{A}}$. Note that the coefficients for $S^d(\mathbf{x},k)$ can be derived as $b_n^d(k) = \langle S^d(\mathbf{x},k), G_n(\mathbf{x},k)\rangle_w$.

Consider the reverberant room channel as a transformation between the reproduced soundfield and the desired soundfield, which can be further expressed by a linear transformation of the basis function coefficients:

$$\mathbf{b}(k) = \mathbf{U}(k)\mathbf{b}^d(k). \tag{8}$$

where $\mathbf{U}(k)$ represents the reverberant room effects at the wavenumber k.

The room channel transformation $\mathbf{U}(k)$ can be estimated in an adaptive fashion. We define $\mathbf{b}(k)_\tau$ as the measured soundfield coefficients (τ is the time index) and $\mathbf{b}(k)_\tau$ can be derived actively based on the sparse microphone measurements following the method in Sec. 2.2. An accurate estimate of the room channel transformation $\hat{\mathbf{U}}(k)$ can be achieved if the squared norm of the residual error $E[\|\mathbf{b}(k)_\tau - \mathbf{b}^d(k)\|^2]$ is minimized, which also leads to an accurate matching between the actual reproduced soundfield and the desired multizone soundfield over $\mathbb{D}$ in the weighted least-squares sense [11]. This is a classical adaptive filtering problem and $\mathbf{U}(k)$

can be estimated actively by using algorithms such as Least Mean Squares (LMS) filter and Recursive Least Squares (RLS) filter.

Note that $\mathbf{U}(k)$ can be parameterized with a diagonal structure following the assumption and experimental observation in earlier works [18] [19] that the couplings between the soundfield coefficients with different indices can be neglected in the defined basis-function domain. With $\mathbf{U}(k) = \mathrm{diag}[U_1(k), \ldots, U_N(k)]$, calculating the unknown diagonal entries $U_n(k)$ can be further simplified as a single-tap adaptive filtering problem (as will be shown in Sec. IV). Let $\hat{\mathbf{U}}(k)_\tau$ be the estimate of $\mathbf{U}(k)$ at time τ, we have [22]:

$$\hat{U}_n(k)^H_\tau = \hat{U}_n(k)^H_{\tau-1} + \frac{1}{\phi_n^2(\tau)} b_n^d(k)(b_n(k)_\tau - b_n^d(k))^H, \quad (9)$$

where $\phi_n^2(\tau)$ is the gain factor $\phi_n^2(\tau) = \lambda\phi_n^2(\tau-1) + |b_n^d(k)|^2$. λ is the forgetting factor. We choose the RLS algorithm as it provides a fast convergence rate. Therefore, (9) can be applied to obtain an iterative estimate of the diagonal elements $U_n(k)$ based on the residual error at time τ. The updating signal on the loudspeaker array in each iteration is interlaced with the estimator $\hat{\mathbf{U}}(k)_\tau$, as described in the following section.

3.2. Loudspeaker Updating Signals

In this section, we derive the optimal updating signal on the loudspeaker array based on the active estimate of the room channel transformation. It is designed to minimize the residual error and ensure the estimation convergence.

By preconditioning the initial loudspeaker array signals to $\mathbf{l}^{\text{direct}}(k)$ that reproduce the desired multizone soundfield under the free-field assumption following the method proposed in our previous work [3], the coefficients for the desired soundfield $\mathbf{b}^d(k)$ can be expressed by replacing $\mathbf{C}(k)$ with the direct channel $\mathbf{C}^{\text{direct}}(k)$ in (6):

$$\mathbf{b}^d(k) = \mathbf{T}\mathbf{C}^{\text{direct}}(k)\mathbf{l}^{\text{direct}}(k). \quad (10)$$

We define $\mathbf{G}^d(k) = \mathbf{T}\mathbf{C}^{\text{direct}}(k)$ that represents the soundfield coefficient matrix of the 2-D Green's functions for all loudspeakers assuming free-field propagation, which can be pre-determined prior to the adaptive filtering process. Considering the coefficients for the desired soundfield in (10), we incorporate the room channel model in (8) and the estimator $\hat{\mathbf{U}}(k)_\tau$ derived in Sec. 3.1. Then, the measured soundfield coefficients $\mathbf{b}(k)_\tau$ after adding updating signals $\sigma(k)_\tau$ to the loudspeakers can be given by

$$\mathbf{b}(k) = \hat{\mathbf{U}}(k)\mathbf{G}^d(k)[\mathbf{l}^{\text{direct}}(k) + \sigma(k)]. \quad (11)$$

Note that we omit time index τ as it is reasonable to assume that $\sigma(k)_\tau$ and $\hat{\mathbf{U}}(k)_\tau$ are i.i.d.. We can then write the difference between the measured and desired soundfield coefficients using (10) and (11):

$$\mathbf{b}(k) - \mathbf{b}^d(k) = [\hat{\mathbf{U}}(k) - \mathbf{I}]\mathbf{G}^d(k)\mathbf{l}^{\text{direct}}(k) + \hat{\mathbf{U}}(k)\mathbf{G}^d(k)\sigma(k), \quad (12)$$

where $\mathbf{I}$ is an identity matrix. Therefore, (7) can be rewrited and our objective is to find the optimal loudspeaker updating signals $\sigma(k)$ that minimize $\|\mathbf{b}(k) - \mathbf{b}^d(k)\|^2$:

$$\underset{\sigma}{\text{argmin}} \|[\hat{\mathbf{U}}(k) - \mathbf{I}]\mathbf{G}^d(k)\mathbf{l}^{\text{direct}}(k) + \hat{\mathbf{U}}(k)\mathbf{G}^d(k)\sigma(k)\|^2. \quad (13)$$

Eq. (13) can be typically solved by using the least squares method. However, it involves the pseudoinverse of $\hat{\mathbf{U}}(k)\mathbf{G}^d(k)$ in each of the adaption steps. $\hat{\mathbf{U}}(k)\mathbf{G}^d(k)$, which represents the estimated coefficient matrix of the reverberant room transfer functions for the employed loudspeakers, is usually ill-conditioned and it can be problematic as the ill-conditioning problem usually leads to poor convergence behavior in the MIMO adaptive identification systems [23].

Note that $[\hat{\mathbf{U}}(k)\mathbf{G}^d(k)\sigma(k)]$ in (12) represents the room effects due to $\sigma(k)$, which is the sum of the direct channel effects $\mathbf{G}^d(k)\sigma(k)$ and the reverberant effects $\mathbf{T}\mathbf{C}^r(k)\sigma(k) = (\hat{\mathbf{U}}(k) - I)\mathbf{G}^d(k)\sigma(k)$. To avoid the active pseudoinverse computation that involves $\hat{\mathbf{U}}(k)$, we assume that the reverberation effects due to $\sigma(k)$ is negligible if $\sigma(k)$ is small and can be mitigated by the adaptive process. Then, (12) can be simplified as

$$\mathbf{b}(k) - \mathbf{b}^d(k) = [\hat{\mathbf{U}}(k) - \mathbf{I}]\mathbf{G}^d(k)\mathbf{l}^{\text{direct}}(k) + \mathbf{G}^d(k)\sigma(k). \quad (14)$$

Therefore, a multi-constraint optimization is formulated with the same objective of (13), while also facilitates the system convergence

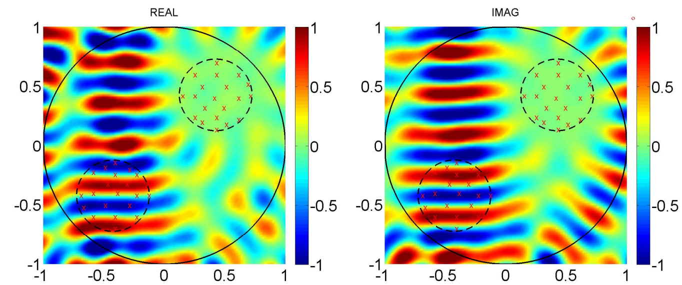
$$\underset{\sigma}{\text{argmin}} \|\mathbf{G}^d(k)\sigma(k) - [\mathbf{I} - \hat{\mathbf{U}}(k)]\mathbf{G}^d(k)\mathbf{l}^{\text{direct}}(k)\|^2,$$
$$\text{subject to } \|\sigma_q(k)\|^2 \leq N_1, \ (q = 1 \ldots Q). \quad (15)$$

The additional constraints on the energy of each of the Q loudspeaker updating signals are applied so that the reverberation effects of $\sigma_q(k)$ are consistently insignificant. The value of N_1 depends on how reverberant the room environment is and experimental results suggest that choosing N_1 to be less or equal to $(1 - \beta(k)^2)/N_w$ yields a fair convergence behavior [22], where $\beta(k)$ is the reflection coefficients at k and N_w is the number of considered walls. Note that $\mathbf{G}^d(k)$ in (15) can be calculated offline. Therefore, the iterative pseudoinverse of $\hat{\mathbf{U}}(k)\mathbf{G}^d(k)$ can be avoided, which facilitates the estimation convergence (as will be shown in Sec. 4). In this work, we use the CVX tool [24] to solve (15). The derived $\sigma(k)$ will be considered in the following adaption step and the new residual error of the soundfield coefficients updates the estimate of $\mathbf{U}(k)$ according to (9).

4. RESULTS AND DISCUSSION

We start with a discussion of simulation parameters. The speed of sound c is 343 m/s in our simulations. The reverberant room is rectangular (size 6 m × 5 m) with wall reflection coefficients of 0.7. $\mathbb{D}$ has a radius of $r = 1$ m with its center located at (2 m,2.5 m) and the loudspeakers are evenly distributed along a concentric circle with a radius of 1.5 m. We used the image source method [25] to simulate the soundfield created by the loudspeakers in the reverberant room. In the simulations, a total of 60 sources for each loudspeaker were included. The centers of $\mathbb{D}_b$ and $\mathbb{D}_q$ lie on a circle of radius $d = 0.6$ m within $\mathbb{D}$. The target bright and quiet zones were located at 225° and 45°, respectively, with $r_q = 0.3$ m, as shown in Fig. 1. We randomly selected $M/2$ locations within each zone and measured the value of the soundfield at those positions. The desired soundfield over $\mathbb{D}_b$ is selected to be a planewave arriving from 90° for the following simulations. The values of weighting function $w(\mathbf{x})$ assigned to $\mathbb{D}_b$, $\mathbb{D}_q$ and $\mathbb{D}_u$ were 1, 7.5 and 0.01 respectively. For the approach of IRLS, the sparsity promoting norm was $p = 0.3$. Complex Gaussian noise was introduced in order to maintain a specific SNR with respect to the pressure power of the desired soundfield at the centre of the bright zone (which was normalized to 0 dB).

We compared the performance of the proposed method with the adaptive approach in [19] that requires the estimation of the reverberant component of the room channel, as well as a non-adaptive reverberation equalization method using sparse techniques in [11]. Note

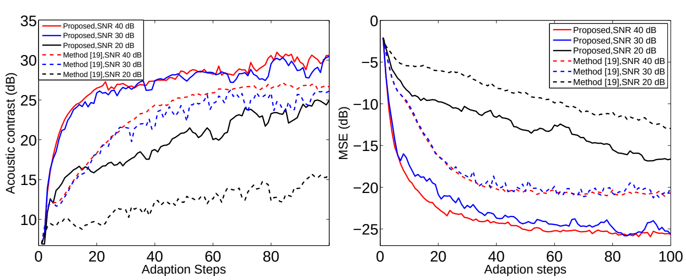


(a) (b)

Fig. 2: Reproduction of the desired multizone sound at 1 kHz using 20 noisy measurements for each selected zone in a reverberant room. (a) and (b) demonstrate the real and imaginary part respectively. The red crosses represent the positions of the microphones.

that we extended the work in [19] to the multizone case by applying the same multizone rendering method in [3]. The method of [19] requires the microphones to be evenly placed along the boundaries of the smallest circle that encloses the two selected zones. RLS was used as the adaptation algorithm of both the proposed technique and the method in [19] with a forgetting factor of $\lambda = 0.95$. The reproduction accuracy of the desired multizone soundfield was evaluated, which considers the acoustic energy contrast between $\mathbb{D}_b$ and $\mathbb{D}_q$ [26], as well as the MSE between the desired and the actual reproduced soundfield over $\mathbb{D}_b$ [3].

Fig. 2 demonstrates the reproduction of the desired multizone soundfield using 40 loudspeakers after 100 adaption steps using 20 noisy measurements (at the noise level of SNR 40 dB) for each selected zone at 1 kHz. The acoustic contrast between $\mathbb{D}_b$ and $\mathbb{D}_q$ is 30.6 dB and the MSE over $\mathbb{D}_b$ is -25.4 dB, which indicates that the reproduced soundfield matches the desired multizone sound well.

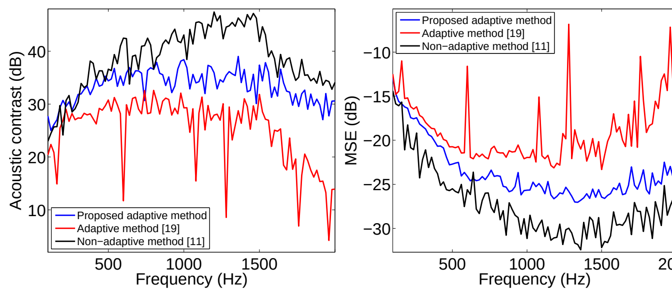


(a) (b)

Fig. 3: Comparison of reproduction performance between our method and the approach in [19]. (a) and (b) represent the performance of acoustic contrast between $\mathbb{D}_b$ and $\mathbb{D}_q$ and MSE over $\mathbb{D}_b$ respectively.

Given the same number of loudspeakers and microphones, we compare our proposed method with the approach in [19] in Fig. 3. The reproduction performance is plotted as a function of the adaption step with three noisy settings. The results are averaged over 10 trial runs. From Fig. 3, we can see that the proposed method outperforms the adaptive approach in [19] in the aspects of both inter-zone acoustic contrast and the MSE over $\mathbb{D}_b$ after 100 adaption steps, especially for the case with relatively lower SNR. Additionally, the proposed method features a faster convergence rate to an accurate reproduction than the approach in [19], which also verifies the validity of the way to neglect the reverberation effects due to small $\sigma(k)$.

The better performance of our method is due to the following reasons: i) sparse estimation methods facilitate a more accurate characterization of the reverberant room channel than classical estimation approaches, given the same provision of noisy measurements [11], ii) the coefficient weighting function [10] attached to various modes in the cylindrical harmonic decomposition was not considered in [19] when minimizing the error between the desired soundfield and the reproduced soundfield coefficient set. In contrast, our formulation does not suffer from this issue as the employed basis function set $\{G_n\}_{n\in\mathbb{A}}$ is formulated to be orthonormal over the desired reproduction region and it facilitates finding the optimal updating signal solution so that a more efficient reverberation cancelation is achieved.



(a) (b)

Fig. 4: Wide-band multizone soundfield reproduction with 64 noisy pressure samples, using the proposed method, the adaptive reproduction method in [19] and the non-adaptive method in [11]. The results are averaged over 10 trial runs.

In Fig. 4, we compare our proposed method with the reproduction approaches in [19] and [11] in terms of the wide-band multizone soundfield reproduction after 100 adaption steps from 100 Hz to 2 kHz. For our proposed method and the adaptive approach in [19], 64 noisy measurements at the noise level of SNR 40 dB were used while the method of [11] employed 64 noiseless measurements. A circular array of 75 loudspeakers was employed to satisfy the truncation length [27] for $\mathbb{D}$ at a maximum frequency of 2 kHz. From Fig. 4, we can also observe obvious peaks in the red curves for the method proposed in [19] due to the so-called large error scaling [10] at certain frequencies (i.e., the Dirichlet eigen-frequencies [28]). In contrast, the performance for the proposed method (blue curve) smoothly varies over the selected frequency range. Meanwhile, the performance of the proposed method approaches that of the method in [11] that employs noiseless measurements, which is an ideal setup and can be difficult to achieve in practice. Note that the method in [11] also requires the pre-measurement of the transfer function over the selected zones of all the employed loudspeakers.

5. CONCLUSION

In this paper, we presented a multizone soundfield reproduction system with the active reverberation cancelation approach, which eliminates the requirement for a prior measurement of the room channels for all loudspeakers. The concept of sparse approximation was applied to the adaptive channel estimation process using a limited number of randomly placed noisy measurements and the diagonal structure of the modeled channel transformation facilitates to reduce the computational complexity. The optimal loudspeaker updating signal that maximizes the reverberation cancelation was also actively derived based on the estimate of transformation matrix. Simulation results suggest that the proposed method provides a faster convergence rate than the comparative approach given the same hardware provision, as well as a consistently accurate reproduction of the desired soundfield over a wide frequency range.

6. REFERENCES

[1] M. Shin, S. Lee, F. Fazi, P. Nelson, D. Kim, S. Wang, K. H. Park, and J. Seo, "Maximization of acoustic energy difference between two spaces," *J. Acoust. Soc. Amer.,*, vol. 128, pp. 121–131, Jul. 2010.

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

[3] 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.

[4] 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.

[5] Y. J. Wu and T. Abhayapala, "Multizone 2D soundfield reproduction via spatial band stop filters," in *Workshop. Applications Signal Process. Audio Acoust.* IEEE, 2009, pp. 309–312.

[6] T. Abhayapala and Y. J. Wu, "Spatial soundfield reproduction with zones of quiet," in *Audio Engineering Society Convention 127*, Oct 2009.

[7] P. Coleman, P. Jackson, M. Olik, M. Moller, M. Olsen, and P. J. Abildgaard, "Acoustic contrast, planarity and robustness of sound zone methods using a circular loudspeaker array," *The Journal of the Acoustical Society of America*, vol. 135, no. 4, pp. 1929–1940, 2014.

[8] 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.

[9] M. Kolundzija, C. Faller, and M. Vetterli, "Reproducing sound fields using MIMO acoustic channel inversion," *J. Audio Eng. Soc*, vol. 59, no. 10, pp. 721–734, 2011.

[10] 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.

[11] W. Jin and W.B. Kleijn , "Multizone soundfield reproduction in reverberant rooms using compressed sensing techniques," in *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*, May 2014, pp. 4728–4732.

[12] W. Jin and W.B. Kleijn, "Theory and design of multizone soundfield reproduction using sparse methods," *Audio, Speech, and Language Processing, IEEE/ACM Transactions on*, vol. 23, no. 12, pp. 2343–2355, Dec 2015.

[13] H. Buchner, S. Spors, and W. Kellermann, "Wave-domain adaptive filtering: acoustic echo cancellation for full-duplex systems based on wave-field synthesis," in *Acoustics, Speech, and Signal Processing, 2004. Proceedings. (ICASSP '04). IEEE International Conference on*, May 2004, vol. 4, pp. iv–117–iv–120 vol.4.

[14] H. Buchner and S. Spors, "A general derivation of wave-domain adaptive filtering and application to acoustic echo cancellation," in *Signals, Systems and Computers, 2008 42nd Asilomar Conference on*, Oct 2008, pp. 816–823.

[15] Sascha Spors, Herbert Buchner, Rudolf Rabenstein, and Wolfgang Herbordt, "Active listening room compensation for massive multichannel sound reproduction systems using wave-domain adaptive filtering," *The Journal of the Acoustical Society of America*, vol. 122, no. 1, pp. 354–369, 2007.

[16] M. Schneider and W. Kellermann, "Adaptive listening room equalization using a scalable filtering structure in thewave domain," in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, March 2012, pp. 13–16.

[17] M. Schneider and W. Kellermann, "A direct derivation of transforms for wave-domain adaptive filtering based on circular harmonics," in *Signal Processing Conference (EUSIPCO), 2012 Proceedings of the 20th European*, Aug 2012, pp. 1034–1038.

[18] M. Schneider and W. Kellermann, "A wave-domain model for acoustic mimo systems with reduced complexity," in *Hands-free Speech Communication and Microphone Arrays (HSCMA), 2011 Joint Workshop on*, May 2011, pp. 133–138.

[19] D. S. Talagala, Wen Zhang, and T. D. Abhayapala, "Efficient multi-channel adaptive room compensation for spatial soundfield reproduction using a modal decomposition," *Audio, Speech, and Language Processing, IEEE/ACM Transactions on*, vol. 22, no. 10, pp. 1522–1532, Oct 2014.

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

[21] R. 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.

[22] Simon Haykin, *Adaptive Filter Theory (3rd Ed.)*, Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1996.

[23] G. A. Ybarra and S. T. Alexander, "Effects of ill-conditioned data on least squares adaptive filters," in *Acoustics, Speech, and Signal Processing, 1988. ICASSP-88., 1988 International Conference on*, 1988, pp. 1387–1390 vol.3.

[24] "Cvx: Matlab software for disciplined convex programming," `http://cvxr.com/cvx/`.

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

[26] 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.

[27] 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.

[28] F. M. Fazi and P. A. Nelson, "Nonuniqueness of the solution of the sound field reproduction problem with boundary pressure control," *Acta Acustica united with Acustica*, vol. 98, no. 1, pp. 1–14, 2012.