Joint Design of Spherical Microphone and Loudspeaker arrays for Room Acoustic Analysis

Hai Morgenstern [1,2], Boaz Rafaely [1], Markus Noisternig [2]

[1] Dept. of Electrical and Computer Engineering, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel,
Email: haimorg@ee.bgu.ac.il

[2] IRCAM, CNRS, Sorbonne Universités, UPMC Univ Paris 6, UMR 9912 STMS, 75004 Paris, France

Introduction

Spherical microphone arrays (SMAs) and spherical loudspeaker arrays (SLAs) have been studied separately for directional sound recording and radiation, respectively, and for room acoustic analysis [1–5]. The use of configurations comprised of both a SLA and a SMA, referred to as acoustic multiple-input multiple-output (MIMO) systems, can potentially facilitate an additional spatial analysis of room acoustics due to the added spatial diversity they provide [6–8]. When using spherical arrays, robustness to errors such as mismatch in the frequency response of the transducers and spatial aliasing is an important design requirement [9, 10]. This has been studied for the separate SMA and SLA systems but not for a MIMO system in which both arrays operate simultaneously. This work discusses a joint design method for a SMA and a SLA as part of a MIMO system. A MIMO system model is formulated, which includes mismatch errors. Based on an analysis of mismatch errors and spatial aliasing, a definition of a matched MIMO system is proposed and demonstrated on two examples that only vary in one parameter, the SMA radius.

System Model

The formulation of an acoustic MIMO system in free-field conditions is presented. First, a rigid-sphere SLA with a spherical harmonic (SH) order of N_L and a radius of r_L is considered. The sound field radiated by the array is given by [4]:

$$p^L(k,r,\boldsymbol{\theta}) \quad = \quad \frac{e^{jkr}}{r} s(k)\boldsymbol{y}_L(\boldsymbol{\theta})^{\mathrm{H}}\boldsymbol{B}_L\boldsymbol{\lambda}_L, \qquad (1)$$

where p^L is the pressure radiated with respect to the origin of the SLA, k is the wavenumber, r is the measuring distance, assumed to be in the far field, $s(k) = 1$ is the input signal for the radial velocity of the SLA, $\boldsymbol{\theta}$ are the spatial angles with respect to the origin of the SLA,

$$\boldsymbol{y}_L(\boldsymbol{\theta}) \quad = \quad [Y_0^0(\boldsymbol{\theta}), Y_1^{-1}(\boldsymbol{\theta}), ..., Y_{N_L}^{N_L}(\boldsymbol{\theta})]^{\mathrm{H}} \qquad (2)$$

is a $(N_L+1)^2 \times 1$ steering vector in which Y_n^m is the SH function of order n and degree m,

$$\begin{aligned}\boldsymbol{B}_L \quad = \quad & \mathrm{diag}[b_0^L(kr_L), b_1^L(kr_L), b_1^L(kr_L), ...\\ & ..., b_1^L(kr_L), ..., b_{N_L}^L(kr_L)], \end{aligned}\qquad (3)$$

is a $(N_L+1)^2 \times (N_L+1)^2$ diagonal matrix in which the coefficients $b_n^L(kr_L)$ depend on frequency and on the SLA radius and are given in [4], $(\cdot)^{\mathrm{H}}$ denotes the conjugate transpose operator, and $\boldsymbol{\lambda}_L$ is a $(N_L+1)^2 \times 1$ vector containing the SLA weighting coefficients represented in SH assumed to control the array's radial velocity.

To formulate a MIMO system, a rigid-sphere SMA with a SH order of N_M and a radius of r_M is considered, centred at a distance of r_0 and spatial angles $\boldsymbol{\theta}_0$ with respect to the centre of the SLA. Using a simplified far-field model [11], the sound field surrounding the SMA can be written as a single plane-wave. Following this, the output of the SMA, after applying beamforming, is given by:

$$\begin{aligned} y(k) &= \boldsymbol{\lambda}_M^{\mathrm{H}}\boldsymbol{B}_M\boldsymbol{y}_M(\boldsymbol{\xi}_0)p^L(k,r_0,\boldsymbol{\theta}_0)\\ &= \boldsymbol{\lambda}_M^{\mathrm{H}}\boldsymbol{B}_M\boldsymbol{\Psi}(k)\boldsymbol{B}_L\boldsymbol{\lambda}_L, \end{aligned}\qquad (4)$$

where $y(k)$ is the SMA output, $\boldsymbol{\lambda}_M$ is a $(N_M+1)^2 \times 1$ vector containing the SMA weighting (i.e., beamforming) coefficients in the SH domain, $\boldsymbol{B}_M$ is a $(N_M+1)^2 \times (N_M+1)^2$ diagonal matrix, similar to $\boldsymbol{B}_L$ from eq. (3), in which the coefficients on its diagonal depend on frequency and the SMA radius [2], $\boldsymbol{\xi}_0$ are the spatial angles with respect to the origin of the SMA that indicate the direction of arrival of the single plane-wave, i.e., the SLA position with respect to the SMA. $\boldsymbol{y}_M(\boldsymbol{\theta})$ is a $(N_M+1)^2 \times 1$ steering vector, similar to $\boldsymbol{y}_L(\boldsymbol{\theta})$ from eq. (2), which contains SH functions calculated at the angle $\boldsymbol{\xi}_0$, and $\boldsymbol{\Psi}(k) = \frac{e^{jkr}}{r}\boldsymbol{y}_M(\boldsymbol{\xi}_0)\boldsymbol{y}_L(\boldsymbol{\theta})^{\mathrm{H}}$.

In the last step of the formulation, normalisation is applied to the SLA and the SMA as to remove effects associated with the use of rigid arrays [3, 4]. This is done using normalised directivity vectors, defined as:

$$\begin{aligned} \boldsymbol{\lambda}_L &= \boldsymbol{B}_L\boldsymbol{\gamma}_L,\\ \boldsymbol{\lambda}_M &= \boldsymbol{B}_M^{\mathrm{H}}\boldsymbol{\gamma}_M. \end{aligned}\qquad (5)$$

Substituting these in eq. (4) yields the normalised system transfer function, given by:

$$\begin{aligned} y(k) &= \boldsymbol{\gamma}_M^{\mathrm{H}}\boldsymbol{B}_M\boldsymbol{\Psi}(k)\boldsymbol{B}_L\boldsymbol{\gamma}_L\\ &= \boldsymbol{\lambda}_M^{\mathrm{H}}\boldsymbol{\Psi}(k)\boldsymbol{\lambda}_L. \end{aligned}\qquad (6)$$

Model Mismatch Errors

The robustness of a system against mismatch errors is essential in the design of an array. This motivates a study of these errors. Model mismatch errors are caused by uncertainties in the positions and responses of loudspeakers or microphones comprising each array, inaccuracies in the values of $\boldsymbol{B}_L$ and $\boldsymbol{B}_M$ (which may arise, for example, if a sphere, of either the SLA or the SMA, is not perfectly rigid even though assumed to be), and errors caused by computation with finite numerical precision (such as in sound cards, etc.). These considerations motivate the definition of:

$$\begin{aligned} \tilde{\boldsymbol{B}}_L &= \boldsymbol{B}_L + \boldsymbol{E}_L, \\ \tilde{\boldsymbol{B}}_M &= \boldsymbol{B}_M + \boldsymbol{E}_M, \end{aligned} \tag{7}$$

where $\tilde{\boldsymbol{B}}_L$ and $\tilde{\boldsymbol{B}}_M$ correspond to the real, physical entries of the SLA and SMA, respectively, represented as the sum of the model-evaluated matrices, $\boldsymbol{B}_L$ and $\boldsymbol{B}_M$, with error matrices $\boldsymbol{E}_L$ and $\boldsymbol{E}_M$ for the SLA and SMA, respectively. Substituting these matrices in eq. (6) gives a normalised system transfer function, $\tilde{y}(k)$, that includes mismatch errors:

$$\begin{aligned} \tilde{y}(k) &= \boldsymbol{\lambda}_M^{\mathrm{H}} \boldsymbol{B}_M^{-1} \tilde{\boldsymbol{B}}_M \boldsymbol{\Psi}(k) \tilde{\boldsymbol{B}}_L \boldsymbol{B}_L^{-1} \boldsymbol{\lambda}_L \\ &= \boldsymbol{\lambda}_M^{\mathrm{H}} \boldsymbol{B}_M^{-1} (\boldsymbol{B}_M + \boldsymbol{E}_M) \boldsymbol{\Psi}(k) \times \\ &\quad \times (\boldsymbol{B}_L + \boldsymbol{E}_L) \boldsymbol{B}_L^{-1} \boldsymbol{\lambda}_L \\ &= y(k) + e_M + e_L + e_J \end{aligned} \tag{8}$$

where

$$\begin{aligned} e_M &= \boldsymbol{\lambda}_M^{\mathrm{H}} \boldsymbol{B}_M^{-1} \boldsymbol{E}_M \boldsymbol{\Psi}(k) \boldsymbol{\lambda}_L, \\ e_L &= \boldsymbol{\lambda}_M^{\mathrm{H}} \boldsymbol{\Psi}(k) \boldsymbol{E}_L \boldsymbol{B}_L^{-1} \boldsymbol{\lambda}_L, \text{ and} \\ e_J &= \boldsymbol{\lambda}_M^{\mathrm{H}} \boldsymbol{B}_M^{-1} \boldsymbol{E}_M \boldsymbol{\Psi}(k) \boldsymbol{E}_L \boldsymbol{B}_L^{-1} \boldsymbol{\lambda}_L \end{aligned} \tag{9}$$

are the SMA error, the SLA error, and the joint SLA-SMA error, respectively, and $(\cdot)^{-1}$ denotes the matrix inversion operator. The energy of the error contributions, $\epsilon(k)$, is bounded by:

$$\begin{aligned} \epsilon(k) &= \left\| \tfrac{y(k) - \tilde{y}(k)}{y(k)} \right\| = \left\| \tfrac{e_M + e_L + e_J}{y(k)} \right\| \\ &\leq \left\| \tfrac{e_M}{y(k)} \right\| + \left\| \tfrac{e_L}{y(k)} \right\| + \left\| \tfrac{e_J}{y(k)} \right\|, \end{aligned} \tag{10}$$

Each term from this equation can be bounded separately; for example, the error of the SLA is bounded using:

$$\begin{aligned} \left\| \tfrac{e_M}{y(k)} \right\| &= \left\| \tfrac{\boldsymbol{\lambda}_M^{\mathrm{H}} \boldsymbol{B}_M^{-1} \boldsymbol{E}_M \boldsymbol{\Psi}(k) \boldsymbol{\lambda}_L}{\boldsymbol{\lambda}_M^{\mathrm{H}} \boldsymbol{\Psi}(k) \boldsymbol{\lambda}_L} \right\| \\ &\leq \left\| \boldsymbol{B}_M^{-1} \right\| \left\| \boldsymbol{B}_M \right\| \frac{\left\| \boldsymbol{E}_M \right\|}{\left\| \boldsymbol{B}_M \right\|} \\ &= \kappa(\boldsymbol{B}_M) \frac{\left\| \boldsymbol{E}_M \right\|}{\left\| \boldsymbol{B}_M \right\|} = \epsilon_M(k), \end{aligned} \tag{11}$$

in which $\kappa(\boldsymbol{B}_M)$ is the 2-norm condition number and $\epsilon_M(k)$ denotes the energy bound for the first error contribution. Similarly, $e_L/y(k)$ and $e_J/y(k)$ from eq. (10) are bounded by $\epsilon_L(k) = \kappa(\boldsymbol{B}_L) \left\| \boldsymbol{E}_L \right\| / \left\| \boldsymbol{B}_L \right\|$ and $\epsilon_J(k) = \epsilon_M(k) \epsilon_L(k)$, respectively, and the overall energy of the error is bounded by the following sum:

$$\epsilon(k) \leq \epsilon_M(k) + \epsilon_L(k) + \epsilon_J(k) = \epsilon_T(k). \tag{12}$$

Joint Design of SMA and SLA

The mismatch error bound, formulated in the previous section, is independent of the SMA and SLA weighting coefficient vectors and is thus a good candidate for designing a MIMO system. Moreover, the way it is modelled, it can also represent other errors such as sensor noise. However, this bound by itself is not enough for designing a system; as will be seen later, the mismatch error of an array is generally a monotonically decreasing function of frequency, and thus it implicitly imposes a low-frequency bound for the operating frequency range of an array (i.e., one can set a low-frequency bound given an application/user-dependent error threshold). For the arrays' high-frequency bound, spatial aliasing is also considered. A system design method is thus proposed based on these two considerations. In general, a system is designed by choosing all the system parameters. However, to reduce the complexity, in this section we only consider a single system parameter, the SMA radius.

Spatial aliasing results from the use of a finite SH order in spherical array processing, which depends on the number of transducers (i.e. microphones and loudspeakers) and their spatial distribution. Given the radius, r, and the SH order, N, of an array, the upper frequency is limited to around $kr < N$ due to spatial aliasing [12]. Considering a MIMO system, the upper frequencies of both SMA and SLA are each calculated, given by:

$$f_M^u = \frac{N_M \, c}{2\pi \, r_M}, \tag{13}$$

$$f_L^u = \frac{N_L \, c}{2\pi \, r_L}, \tag{14}$$

where c is the speed of sound. A global, upper frequency for the MIMO system can be defined as the minimum of these frequencies, i.e.:

$$f^u = \min\left(f_M^u, \, f_L^u\right). \tag{15}$$

This definition is motivated since spatial aliasing of at least one of the arrays in a MIMO system introduces an error to system transfer functions, $u(k)$, formulated using the system. In the context of the system design, r_M is shown to affect the system aliasing frequency in eqs. (13)-(15), and thus both the upper frequency and mismatch errors are considered for setting its value.

The dependence of the system mismatch errors and f^u on r_M is demonstrated using an example of two MIMO systems that vary only in this parameter. The first system, *system* 1, is comprised of a SMA and a SLA with SH orders or $N_M = 8$ and $N_L = 4$, respectively. The SLA radius is set to $r_L = 0.20$m, and the SMA radius is chosen such that the upper frequencies imposed by both SLA and SMA are identical; following eqs. (13) and (14), the SMA radius that yields equal upper frequencies for both arrays is given by:

$$\tilde{r}_M = \frac{N_M}{N_L} r_L. \tag{16}$$

By substituting N_M, N_L, and r_L into the last equation, the SMA radius is calculated at $r_M = \tilde{r}_M = 0.40$m. Fig. 1 presents the two separate bounds, ϵ_M(k) and $\epsilon_L(k)$, for *system* 1. The total mismatch error bound, $\epsilon_T(k)$, is not plotted as it is expected to be dependent on these errors. Moreover, equal energies of the initial SLA and SMA errors are set to $||\boldsymbol{E}_L||/||\boldsymbol{B}_L|| = ||\boldsymbol{E}_M||/||\boldsymbol{B}_M|| = -25$dB; for simplicity, these are assumed to be equal at all frequencies. The SMA and SLA upper frequencies, f_M^u, and f_L^u, respectively, are depicted as vertical lines in the figure. The frequency ranges to the left of the vertical lines f_M^u and f_L^u are shaded as to demonstrate that the SMA and SLA can operate in these ranges, respectively, without introducing significant aliasing errors. The second MIMO system, *system* 2, has the same system parameters as *system* 1 except for the SMA radius which in this case is set to $r_M = \tilde{r}_M/5 = 0.08$m. Fig. 2 presents the mismatch bounds for *system* 2. Considering *system* 2, the small SMA radius yields a higher frequency of $f_M^u = 5f_L^u > f_L^u$. Due to this upper frequency, in this case, the operating frequency ranges of the SLA and SMA do not overlap anymore; as can be seen in the figure, for frequencies within the range of $[f_L^u, f_M^u]$ only the SLA introduces aliasing, while the SMA is aliasing-free and robust to mismatch errors. Nonetheless, these frequencies are excluded from the MIMO system operating frequency range. Thus, we can define a "matched MIMO" system as a system in which the operating frequency ranges of the SMA and SLA overlap, achieved by "matching" the upper frequency limits of both the SMA and the SLA, as in the case of *system* 1.

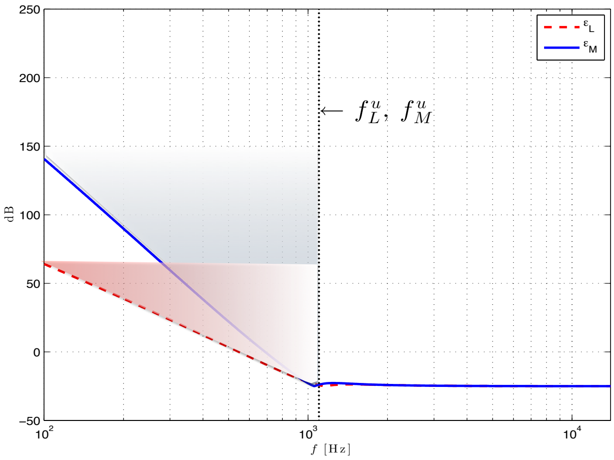

Figure 1: Energetic bounds ϵ_M and ϵ_L calculated for a system with $N_L = 4$, $N_M = 2$, $r_L = 0.20$m, and $r_M = \tilde{r}_M = 0.40$m. f_M^u, f_L^u are drawn as vertical lines.

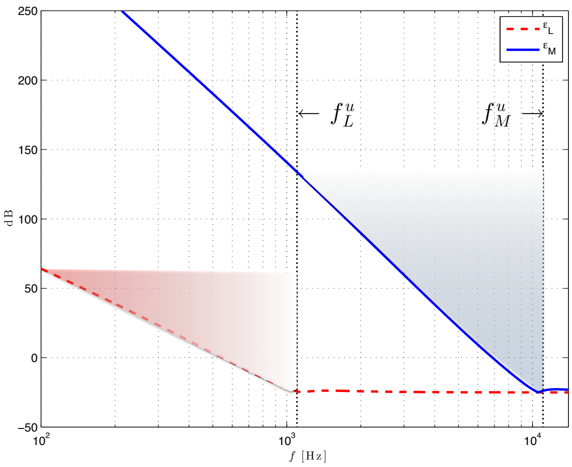

Figure 2: Energetic bounds ϵ_M, and ϵ_L calculated for a system with $N_L = 4$, $N_M = 2$, $r_L = 0.20$m, and $r_M = \tilde{r}_M/5 = 0.08$m. f_M^u, f_L^u are drawn as vertical lines.

Conclusions

An approach for designing an acoustic MIMO system was presented, based on considerations of mismatch errors and spatial aliasing, and a definition of a matched system was proposed. Future work includes demonstrating the superiority of matched systems over unmatched systems in applications of room acoustic analysis or directional room impulse response synthesis. Moreover, in the formulation presented here spatial aliasing has been taken into account by manually inserting the aliasing frequencies of the arrays. A more general system model that explicitly incorporates aliasing errors in the model is left open for future work.

References

[1] Abhayapala, T. D. and Ward, D. B.: Theory and design of high order sound field microphones using spherical microphone array. Proceedings of ICASSP 2002, Vol. II, pp. 1949–1952.

[2] Rafaely, B.: Plane-wave decomposition of the sound field on a sphere by spherical convolution. The Journal of the Acoustical Society of America (2004), Vol. 116, pp. 2149–2157.

[3] Park, M. and Rafaely, B.: Sound-field analysis by plane-wave decomposition using spherical microphone array. The Journal of the Acoustical Society of America (2005), Vol. 118, pp. 3094–3103.

[4] Rafaely, B. and Khaykin, D.: Optimal model-based beamforming and independent steering for spherical loudspeaker arrays. IEEE Transactions on Audio, Speech, and Language Processing (2011), Vol. 19, pp. 2234–2238.

[5] Zotter, F.: Analysis and synthesis of sound-radiation with spherical arrays. Dissertation, Institute of Electronic Music and Acoustics, Austria, September 2009.

[6] Morgenstern, H., Zotter, F., and Rafaely, B.: Joint spherical beamforming for directional analysis of reflections in rooms. Proceedings of ACOUSTICS 2012 HONG KONG, Vol. 131, no. 4, pp. 3207.

[7] Morgenstern, H. and Rafaely, B.: Enhanced spatial analysis of room acoustics using acoustic multiple-input multiple-output (MIMO) systems. Proceedings of meetings on Acoustics, Acoustical Society of America (2013), Vol. 19, pp. 15–18.

[8] Morgenstern, H. and Rafaely, B.: Spherical loudspeaker array beamforming in enclosed sound fields by MIMO optimization. Proceedings of ICASSP 2013, pp. 370–374.

[9] Rafaely, B.: The spherical-shell microphone array. IEEE Transactions on Audio, Speech, and Language Processing (2008), Vol. 16, pp. 740–747.

[10] Pasqual, A. M., de França, A., Roberto, J., and Herzog, P.: Application of acoustic radiation modes in the directivity control by a spherical loudspeaker array Acta Acustica united with Acustica (2010), Vol. 96, pp. 32–42.

[11] Morgenstern. H. and Rafaely, B.:Far-field criterion for spherical microphone arrays and directional sources. Proceedings of IEEE 4th Joint Workshop on HSCMA 2014, pp. 42–46.

[12] Rafaely, B.: Analysis and design of spherical microphone arrays. IEEE Transactions on Speech and Audio Processing (2005), Vol. 13, pp. 135–143.