

Reproducing the Acoustic Velocity Vectors in a Spherical Listening Region

Frank Jiarui Wang*, Thushara Abhayapala*, Jihui Aimee Zhang†, Prasanga Samarasinghe*
*The Australian National University †University of Southampton

June 2024

Abstract

Acoustic velocity vectors (AVVs) are related to the human’s perception of sound at low frequencies and are widely used in Ambisonics. This paper proposes a spatial sound field reproduction algorithm called velocity matching, which reproduces the AVVs in the spherical listening region by matching the AVVs’ spherical harmonic coefficients. Using the sound field translation formula, the spherical harmonic coefficients of the AVVs are derived from the spherical harmonic coefficients of the pressure, which can be measured by a higher-order microphone array. Unlike algorithms that only control the AVVs at discrete sweet spots, the proposed velocity matching algorithm manipulates the AVVs in the whole spherical listening region and allows the listener to move beyond the sweet spots. Simulations show the proposed velocity matching algorithm accurately reproduces the AVVs in the spherical listening region and requires fewer number of loudspeakers than pressure matching algorithm.

1 Introduction

Spatial sound field reproduction aims to synthesize the desired sound field in the listening region. In most cases, reproduction is achieved by accurately reconstructing the pressure. Pressure based methods include matching the pressure at a number of sweet spots [1], wave field synthesis (WFS) [2–6] and higher order Ambisonics (HOA) [7–9]. A large number of loudspeakers are usually involved to achieve sufficient accuracy in the reconstructed pressure, which alone does not guarantee satisfactory perception [10].

Acoustic velocity vectors (AVVs) are essential to WFS due to the Kirchhoff-Helmholtz integral [4]. Inspired by WFS, in [11] and [12], reproduction was

achieved by matching the pressure and the AVVs at discrete control points on the boundary of the listening regions. There was also attempt at solely matching the AVVs on the boundary [13]. Measuring the AVVs at multiple control points involves a complicated setup. Moreover, [11] and [12] required a large number of loudspeakers, which could be impractical for home theater or small exhibition space.

Perceptually motivated sound field reproduction creates the desired perceptual effects by using psychoacoustics, and has the advantage of requiring fewer channels due to its tolerance to lower accuracy in the reproduced pressure [14]. AVVs are related to the human’s perception of sound below 700 Hz [15, 16] and are relevant to the interaural phase difference. Recently, AVVs were combined with mixed-source expansion to create immersive perception over an enlarged region [17]. AVVs were also used to create the desired perception at non-central listening points [18].

AVVs have been applied to reproduction at sweet spots. Gerzon’s velocity vector r_V is widely used in Ambisonics [15, 16, 19, 20]. A time-domain method that jointly controls the AVVs and the pressures at multiple sweet spots was derived in [21] and [22]. To enable listener’s movement beyond the sweet spots, the AVVs in the whole listening region should be characterized.

A spherical region can be treated as an integral of multiple concentric spherical surfaces. In [23], the spherical harmonic (SH) coefficients of the AVVs on each concentric spherical surface (abbrev. **SHV-surf** coefficients) were derived from the SH coefficients of the pressure in the spherical region, which were measured by a spherical microphone array [24, 25]. The SHV-surf coefficients were of the form $X_a^d(k, r_b)$ in which r_b was the radius of the spherical surface on which the AVVs were characterized [17, 23]. SHV-surf coefficients were applied

to virtual microphone synthesis to simulate the signals at microphones placed on an open sphere [26]. To compute the AVVs in a spherical region, the SHV-surf coefficients must be calculated for multiple radii.

To characterize the AVVs in a spherical listening region using one set of SH coefficients across all radii, this paper presents the radial independent SH coefficients of the AVVs (abbrev. **SHV-indR** coefficients). The SHV-indR coefficients are of the form $(\zeta_{\hat{e}}^d)_a(k)$ and the radial dependency is captured by a separate radial function $R_a(kr_b)$. Starting from the definition of the AVVs, it is found that the SHV-indR coefficients are related to the SH coefficients of the pressure in the spherical listening region by the sound field translation formula. Therefore, the SHV-indR coefficients can be derived from spherical microphone array measurements. Unlike SHV-surf coefficients, which must be calculated at different radii, one set of SHV-indR coefficients is sufficient to characterize the AVVs throughout the spherical listening region. In [27], the SHV-indR coefficients were derived by using eigenbeam-ESPRIT, which is a source localization method.

This paper also proposes the velocity matching (VM) algorithm, which reproduces the desired AVVs in the spherical listening region by matching the SHV-indR coefficients. Due to the radial-independence, only one set of SHV-indR coefficients needs to be matched, whereas the SHV-surf coefficients must be matched on multiple concentric spherical surfaces. Simulation shows that VM accurately reproduces the AVVs at low frequencies and requires fewer number of loudspeakers than pressure based method.

2 AVV in the spherical region

2.1 AVVs at a point

Figure 1 shows the setup of the geometric model. The spherical listening region in light blue is free from sources and scatterers. The derivation starts from finding the AVVs at a point \mathbf{r}_b within the listening region. The local $x^{(b)}y^{(b)}z^{(b)}$ coordinate system is centered at $\mathbf{r}_b \equiv O^{(b)}$. The $x^{(b)}y^{(b)}z^{(b)}$ coordinate system is the translation of the xyz coordinate system with \mathbf{r}_b as the translation vector. Note that $\mathbf{r} = \mathbf{r}_b + \mathbf{r}^{(b)}$. The superscript indicates the coordinate system used to express the location. If there are no superscripts, then the location is expressed

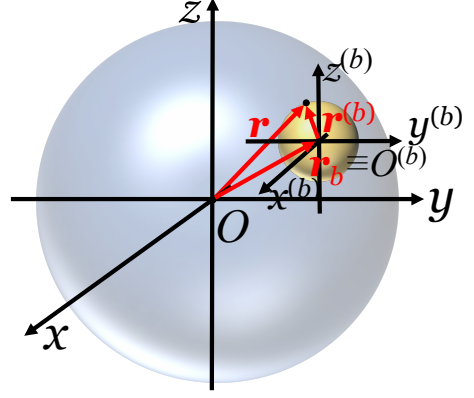


Figure 1: Setup of the geometric model. The listening region is in light blue. \mathbf{r}_b is a point in the listening region.

with respect to the xyz coordinate system. When using spherical coordinate system, $\mathbf{r} = (r, \theta, \phi)$ in which $r = \|\mathbf{r}\|$, θ is the colatitude measured downward from the zenith (the positive z -axis), and ϕ is the azimuth measured counterclockwise from the positive x -axis in the xy plane.

In Figure 1, the pressure at a point $\mathbf{r}^{(b)} \equiv (r^{(b)}, \theta^{(b)}, \phi^{(b)})$ within the local region in yellow is

$$p(k, \mathbf{r}^{(b)}) = \sum_{n=0}^N \sum_{m=-n}^n \beta_n^m(k, \mathbf{r}_b) j_n(kr^{(b)}) Y_n^m(\theta^{(b)}, \phi^{(b)}) \quad (1)$$

in which k is the wavenumber, $j_n(\cdot)$ is the spherical Bessel function of the first kind, $Y_n^m(\cdot, \cdot)$ is the SH function of degree n and order m , and N is the truncation order. The SH coefficients $\beta_n^m(k, \mathbf{r}_b)$ depend on the location of \mathbf{r}_b , which is the origin of the local coordinate system. From [28], the derivative $[\partial j_n(kr^{(b)})/\partial r^{(b)}]_{r^{(b)}=0} = (1/3)k\delta_{n,1}$ in which $\delta_{n,1}$ is the Kronecker delta function.

Let ρ_0 denote the density of the medium and c denote the speed of sound. Let \hat{x} , \hat{y} and \hat{z} denote the unit vectors along the x , y , z axes. The AVVs at $\mathbf{r}_b \equiv O^{(b)}$ are the linear combinations of $\beta_1^m(k, \mathbf{r}_b)$ [29]

$$\begin{aligned} V_{\hat{x}}(\mathbf{r}_b, k) &= \frac{i}{k\rho_0 c} \frac{\partial p(k, \mathbf{r}^{(b)})}{\partial x} \Big|_{r^{(b)}=0} \\ &= \sum_{n=0}^N \sum_{m=-n}^n \beta_n^m(k, \mathbf{r}_b) \frac{\partial j_n(kr^{(b)})}{\partial r^{(b)}} \Big|_{r^{(b)}=0} Y_n^m\left(\frac{\pi}{2}, 0\right) \\ &= \frac{1}{3} \frac{i}{\rho_0 c} \left[\sqrt{\frac{3}{8\pi}} \beta_1^{-1}(k, \mathbf{r}_b) - \sqrt{\frac{3}{8\pi}} \beta_1^1(k, \mathbf{r}_b) \right], \quad (2) \end{aligned}$$

$$\begin{aligned}
V_{\hat{\mathbf{y}}}(\mathbf{r}_b, k) &= \frac{i}{k\rho_0 c} \frac{\partial p(k, \mathbf{r}^{(b)})}{\partial y} \Big|_{r^{(b)}=0} \\
&= \sum_{n=0}^N \sum_{m=-n}^n \beta_n^m(k, \mathbf{r}_b) \frac{\partial j_n(kr^{(b)})}{\partial r^{(b)}} \Big|_{r^{(b)}=0} Y_n^m\left(\frac{\pi}{2}, \frac{\pi}{2}\right) \\
&= \frac{1}{3} \frac{1}{\rho_0 c} \left[\sqrt{\frac{3}{8\pi}} \beta_1^{-1}(k, \mathbf{r}_b) + \sqrt{\frac{3}{8\pi}} \beta_1^1(k, \mathbf{r}_b) \right], \quad (3)
\end{aligned}$$

$$\begin{aligned}
V_{\hat{\mathbf{z}}}(\mathbf{r}_b, k) &= \frac{i}{k\rho_0 c} \frac{\partial p(k, \mathbf{r}^{(b)})}{\partial z} \Big|_{r^{(b)}=0} \\
&= \sum_{n=0}^N \sum_{m=-n}^n \beta_n^m(k, \mathbf{r}_b) \frac{\partial j_n(kr^{(b)})}{\partial r^{(b)}} \Big|_{r^{(b)}=0} Y_n^m(0, 0) \\
&= \frac{1}{3} \frac{i}{\rho_0 c} \sqrt{\frac{3}{4\pi}} \beta_1^0(k, \mathbf{r}_b). \quad (4)
\end{aligned}$$

2.2 The SHV-indR coefficients

In Figure 1, using the global xyz coordinate system, the pressure at $\mathbf{r} \equiv (r, \theta, \phi)$ is

$$p(k, \mathbf{r}) = \sum_{\ell=0}^L \sum_{q=-\ell}^{\ell} \xi_{\ell}^q(k) j_{\ell}(kr) Y_{\ell}^q(\theta, \phi). \quad (5)$$

in which $\xi_{\ell}^q(k)$ denotes the SH coefficients of the pressure in the spherical listening region, and L is the truncation order. Using the sound field translation formula [30],

$$\begin{aligned}
p(k, \mathbf{r}^{(b)}) &= \sum_{n=0}^N \sum_{m=-n}^n j_n(kr^{(b)}) Y_n^m(\theta^{(b)}, \phi^{(b)}) \\
&\underbrace{\sum_{a=0}^A \sum_{\ell=0}^L \sum_{q=-\ell}^{\ell} \xi_{\ell}^q(k) G_{nm}^{\ell qa} j_a(kr_b) Y_a^{(q-m)}(\theta_b, \phi_b)}_{\beta_n^m(k, \mathbf{r}_b) \text{ in (1)}} \quad (6)
\end{aligned}$$

As shown in (2), (3) and (4), since the AVVs involve only $\beta_1^m(k, \mathbf{r}_b)$, the derivation restricts $n = 1$ and $m = \{-1, 0, 1\}$. The term

$$G_{1m}^{\ell qa} = 4\pi i^{1+a-\ell} (-1)^q \sqrt{\frac{3(2\ell+1)(2a+1)}{4\pi}} W_1 W_2 \quad (7)$$

in which W_1 and W_2 are the Wigner-3j symbols [31]

$$W_1 = \begin{pmatrix} \ell & 1 & a \\ 0 & 0 & 0 \end{pmatrix} \quad W_2 = \begin{pmatrix} \ell & 1 & a \\ -q & m & q-m \end{pmatrix}. \quad (8)$$

Let $d = q - m$, $\beta_1^m(k, \mathbf{r}_b)$ in (6) becomes

$$\begin{aligned}
\beta_1^m(k, \mathbf{r}_b) &= \sum_{a=0}^A \sum_{\ell=0}^L \sum_{d=-\ell-m}^{\ell-m} \xi_{\ell}^{(d+m)}(k) G_{1m}^{\ell(d+m)a} j_a(kr_b) Y_a^d(\theta_b, \phi_b) \\
&= \sum_{a=0}^A \sum_{d=-L-m}^{L-m} \underbrace{\left[\sum_{\ell=|d+m|}^L \xi_{\ell}^{(d+m)}(k) G_{1m}^{\ell(d+m)a} \right]}_{(\gamma_1^m)_a^d} \\
&\quad j_a(kr_b) Y_a^d(\theta_b, \phi_b) \quad (9)
\end{aligned}$$

in which A is the truncation order.

Operator matrices can be constructed to link the SH coefficients $\xi_{\ell}^{(d+m)}(k)$ of the pressure to $(\gamma_1^m)_a^d(k)$ with $m \in \{-1, 0, 1\}$ such that

$$(\gamma_1^m)(k) = \mathfrak{B}_1^m \boldsymbol{\xi}(k) \quad (10)$$

in which $(\gamma_1^m)(k)$ and $\boldsymbol{\xi}(k)$ are the column vectors formed by concatenating $(\gamma_1^m)_a^d(k)$ and $\xi_{\ell}^{(d+m)}(k)$, respectively. The operator matrices do not depend on the wavenumber k (also the frequency) because $G_{1m}^{\ell(d+m)a}$ are frequency independent.

The calculation of \mathfrak{B}_1^m does not require significant resources because only three operator matrices with $m = \{-1, 0, 1\}$ are required. Moreover, $G_{1m}^{\ell(d+m)a}$ is non-zero only when $|\ell - 1| \leq a \leq \ell + 1$ [31]. Furthermore, since $W_1 = 0$ when $a = \ell$ [32], only two conditions $a = |\ell - 1|$ and $a = \ell + 1$ need to be considered. The dimension of \mathfrak{B}_1^m is L^2 by $(L + 1)^2$. This is because if $\xi_{\ell}^{(d+m)}(k)$ is measured up to degree L , the maximum degree of $(\gamma_1^m)_a^d(k)$ can be calculated is $(L - 1)$. For $a = L$, $\xi_{\ell}^{(d+m)}(k)$ with $\ell = L + 1$ should be measured.

Substituting (9) into (2), (3) and (4),

$$V_{\hat{\mathbf{e}}}(\mathbf{r}_b, k) = \sum_{a=0}^A \sum_{d=-a}^a (\zeta_{\hat{\mathbf{e}}})_a^d(k) j_a(kr_b) Y_a^d(\theta_b, \phi_b) \quad (11)$$

in which $\hat{\mathbf{e}} \in \{\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}\}$ and $(\zeta_{\hat{\mathbf{e}}})_a^d(k)$ denotes the SHV-indR coefficients of the form

$$(\zeta_{\hat{\mathbf{x}}})_a^d(k) = \frac{1}{3} \frac{i}{\rho_0 c} \left[\sqrt{\frac{3}{8\pi}} (\gamma_1^{-1})_a^d(k) - \sqrt{\frac{3}{8\pi}} (\gamma_1^1)_a^d(k) \right], \quad (12)$$

$$(\zeta_{\hat{\mathbf{y}}})_a^d(k) = \frac{1}{3} \frac{1}{\rho_0 c} \left[\sqrt{\frac{3}{8\pi}} (\gamma_1^{-1})_a^d(k) + \sqrt{\frac{3}{8\pi}} (\gamma_1^1)_a^d(k) \right], \quad (13)$$

$$(\zeta_{\hat{\mathbf{z}}})_a^d(k) = \frac{1}{3} \frac{i}{\rho_0 c} \sqrt{\frac{3}{4\pi}} (\gamma_1^0)_a^d(k). \quad (14)$$

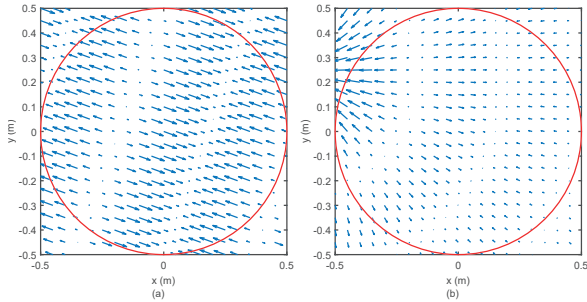


Figure 2: Real part of the AVVs on the xy plane at 500 Hz. (a) A plane wave with incident direction $(\theta_{pw}, \phi_{pw}) = (\pi/2, 8\pi/9)$ rad. (b) A point source at $\mathbf{r}_{ps} = (0.7 \text{ m}, \pi/2 \text{ rad}, 8\pi/9 \text{ rad})$. The red circle with radius 0.5 m is the cross section of the boundary of the spherical listening region.

Operator matrices $\mathfrak{B}_{\hat{e}}$ with $\hat{e} \in \{\hat{x}, \hat{y}, \hat{z}\}$ that compute the SHV-indR coefficients $(\zeta_{\hat{e}})_a^d(k)$ from the SH coefficients $\xi_{\ell}^{(d+m)}(k)$ of the pressure are constructed so that

$$\zeta_{\hat{e}}(k) = \mathfrak{B}_{\hat{e}} \xi(k) \quad (15)$$

in which $\zeta_{\hat{e}}(k)$ is the column vector formed by concatenating $(\zeta_{\hat{e}})_a^d(k)$. From (12), (13) and (14),

$$\mathfrak{B}_{\hat{x}} = \frac{1}{3} \frac{i}{\rho_0 c} \left[\sqrt{\frac{3}{8\pi}} \mathfrak{B}_1^{-1} - \sqrt{\frac{3}{8\pi}} \mathfrak{B}_1^1 \right], \quad (16)$$

$$\mathfrak{B}_{\hat{y}} = \frac{1}{3} \frac{1}{\rho_0 c} \left[\sqrt{\frac{3}{8\pi}} \mathfrak{B}_1^{-1} + \sqrt{\frac{3}{8\pi}} \mathfrak{B}_1^1 \right], \quad (17)$$

$$\mathfrak{B}_{\hat{z}} = \frac{1}{3} \frac{i}{\rho_0 c} \sqrt{\frac{3}{4\pi}} \mathfrak{B}_1^0. \quad (18)$$

2.3 Illustration of the AVVs in a spherical region

For a plane wave with incident direction (θ_{pw}, ϕ_{pw}) ,

$$\xi_{\ell}^q(k) = 4\pi i^{\ell} \overline{Y_{\ell}^q(\theta_{pw}, \phi_{pw})} \quad \forall k. \quad (19)$$

in which $\overline{(\cdot)}$ denotes conjugation. For a point source located at $\mathbf{r}_{ps} \equiv (r_{ps}, \theta_{ps}, \phi_{ps})$,

$$\xi_{\ell}^q(k) = -ik h_{\ell}^{(2)}(kr_{ps}) \overline{Y_{\ell}^q(\theta_{ps}, \phi_{ps})} \quad (20)$$

in which $h_{\ell}^{(2)}(\cdot)$ is the spherical Hankel function of the second kind. Figure 2 shows the real part of the AVVs on the xy plane when the source is (a) a plane wave with incident direction $(\theta_{pw}, \phi_{pw}) = (\pi/2, 8\pi/9)$ rad, and (b) a point source at $\mathbf{r}_{ps} = (0.7 \text{ m}, \pi/2 \text{ rad}, 8\pi/9 \text{ rad})$. The red circle with radius 0.5 m is the cross section of the boundary of the spherical listening region. The SH coefficients $\xi_{\ell}^q(k)$

of the pressure are truncated to $L = 7$ and the SHV-indR coefficients $(\zeta_{\hat{e}})_a^d(k)$ are truncated to $A = 6$. The density of the medium $\rho_0 = 1.2042 \text{ kg/m}^3$ and the speed of sound $c = 343.21 \text{ m/s}$. In Figure 2(a), the AVVs are either parallel or anti-parallel. In Figure 2(b), the AVVs either converge to or diverge from a point in the direction of $\phi = 8\pi/9$ rad and the spherical wave fronts of a point source can be discerned.

3 Reproducing the AVVs in a spherical region

This section presents the velocity matching (VM) algorithm, which reproduces the desired AVVs in the spherical listening region by matching the SHV-indR coefficients. Assume there are S loudspeakers with index $s = 1, 2, \dots, S$. To characterize the loudspeakers, the pressure mode transfer functions $(\xi^{Ls})_{\ell}^q(k)$ are measured, which represent the SH coefficients of the pressure in the spherical listening region when the input to the s -th loudspeaker is a unit sinusoid at frequency $kc/2\pi$ Hz. Next, by using the operator matrices in (15), the s -th loudspeaker's velocity mode transfer functions $(\zeta_{\hat{e}}^{Ls})_a^d(k)$ are obtained. To characterize the desired sound field, the SH coefficients $(\xi^{\text{des}})_{\ell}^q(k)$ of the desired pressure in the spherical listening region are measured by a spherical microphone array. Next, by using the operator matrices in (15), the desired SHV-indR coefficients $(\zeta_{\hat{e}}^{\text{des}})_a^d(k)$ are found. A system of equations is established

$$\zeta^{\text{des}}(k) = \mathbf{H}(k) \mathbf{w}(k). \quad (21)$$

In (21), $\zeta^{\text{des}}(k) = [\zeta_{\hat{x}}^{\text{des}}(k)^T, \zeta_{\hat{y}}^{\text{des}}(k)^T, \zeta_{\hat{z}}^{\text{des}}(k)^T]^T$ in which $\zeta_{\hat{e}}^{\text{des}}(k)$ with $\hat{e} \in \{\hat{x}, \hat{y}, \hat{z}\}$ is the column vector formed by concatenating $(\zeta_{\hat{e}}^{\text{des}})_a^d(k)$ and $(\cdot)^T$ denotes matrix transpose. $\mathbf{H}(k) = [\zeta^{L1}(k), \zeta^{L2}(k), \dots, \zeta^{LS}(k)]$ and its s -th column $\zeta^{Ls}(k) = [\zeta_{\hat{x}}^{Ls}(k)^T, \zeta_{\hat{y}}^{Ls}(k)^T, \zeta_{\hat{z}}^{Ls}(k)^T]^T$ in which $\zeta_{\hat{e}}^{Ls}(k)$ with $\hat{e} \in \{\hat{x}, \hat{y}, \hat{z}\}$ is the column vector formed by concatenating $(\zeta_{\hat{e}}^{Ls})_a^d(k)$. The column vector $\mathbf{w}(k)$ contains the loudspeaker weights. The VM algorithm is compared with the pressure matching (PM) algorithm [7], which finds the loudspeaker weights by matching the SH coefficients of the pressure in the spherical listening region. PM has a system of equations

$$\xi^{\text{des}}(k) = \mathbf{G}(k) \mathbf{w}(k) \quad (22)$$

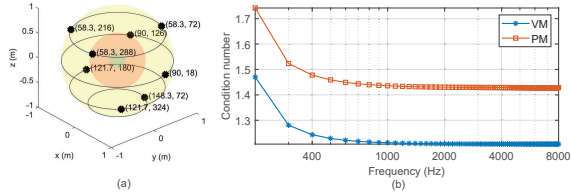


Figure 3: (a) Setup of the 8-channel loudspeaker array with loudspeakers denoted by black crosses. The loudspeakers are located on the yellow sphere with radius 1 m. The loudspeaker angular directions are denoted as $(\theta_{L_s}, \phi_{L_s})$ in degrees. The spherical listening region with radius 0.5 m is bounded by the red sphere. The cyan sphere at the center of the listening region has radius 0.15 m. (b) Condition numbers of $\mathbf{H}(k)$ for VM and $\mathbf{G}(k)$ for PM.

in which $\boldsymbol{\xi}^{\text{des}}(k)$ is the column vector formed by concatenating $(\zeta^{\text{des}})^q_\ell(k)$. The matrix $\mathbf{G}(k) = [\boldsymbol{\xi}^{\text{L1}}(k), \boldsymbol{\xi}^{\text{L2}}(k), \dots, \boldsymbol{\xi}^{\text{LS}}(k)]$ in which the s -th column $\boldsymbol{\xi}^{\text{LS}}(k)$ is formed by concatenating $(\zeta^{\text{LS}})^q_\ell(k)$. In both (21) and (22), the loudspeaker weights $\mathbf{w}(k)$ are found by Moore-Penrose pseudoinverse.

Figure 3(a) shows the 8-channel loudspeaker array. The desired sound field is a plane wave with incident direction $(\theta_{\text{pw}}, \phi_{\text{pw}}) = (\pi/2, 8\pi/9)$ rad. The pressure SH coefficients $(\zeta^{\text{des}})^q_\ell(k)$ and $(\xi^{\text{LS}})^q_\ell(k)$ are truncated to $\ell = 4$. Hence, the SHV-indR coefficients $(\zeta^{\text{des}})^d_a(k)$ and $(\zeta^{\text{LS}})^d_a(k)$ are truncated to $a = 3$. At each wavenumber k , the dimension of $\mathbf{H}(k)$ is 48-by-8 and the dimension of $\mathbf{G}(k)$ is 25-by-8. The density of the medium $\rho_0 = 1.2042 \text{ kg/m}^3$ and the speed of sound $c = 343.21 \text{ m/s}$. In Figure 3(b), the condition numbers remain stable, though those of $\mathbf{H}(k)$ are slightly lower than those of $\mathbf{G}(k)$. The Moore-Penrose pseudoinverse is calculated by the `pinv` function in MATLAB and the default tolerance is used. Code for simulation can be accessed from https://github.com/FJWang01/SH_Velocity.

Figure 4 shows the reproduced pressure and the AVVs on the xy plane at 500 Hz. Figures 4(a) and 4(c) are reproduced by VM, whereas Figures 4(b) and 4(d) are reproduced by PM. The ground truth of the AVVs is in Figure 2(a). VM achieves better results in both reproduced pressure and reproduced AVVs. Like [33] and [34], the velocity reproduction error

$$\eta(k) = \cos^{-1}(\text{DOT}(k)) \text{ rad}, \quad (23)$$

where

$$\text{DOT}(k) = \frac{\mathbf{V}^{\text{des}}(\mathbf{r}_b, k)}{\|\mathbf{V}^{\text{des}}(\mathbf{r}_b, k)\|_2} \cdot \left[\frac{\mathbf{V}^{\text{re}}(\mathbf{r}_b, k)}{\|\mathbf{V}^{\text{re}}(\mathbf{r}_b, k)\|_2} \right]^T \quad (24)$$

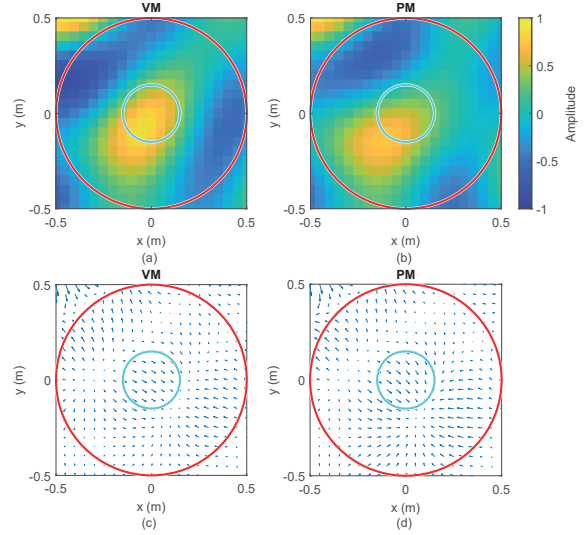


Figure 4: Real part of the reproduced pressure and the reproduced AVVs on the xy plane at 500 Hz. The desired sound field is a plane wave with incident direction $(\pi/2, 8\pi/9)$ rad. (a) and (c) are reproduced by VM; (b) and (d) are reproduced by PM. The red circle and the cyan circle are the cross sections of the red sphere and the cyan sphere in Figure 3(a), respectively.

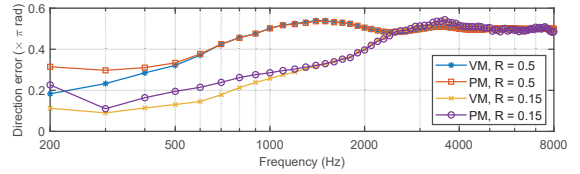


Figure 5: Velocity reproduction errors in the real part of the reproduced AVVs. $R = 0.5$: the spherical listening region bounded by the red sphere with radius 0.5 m in Figure 3(a); $R = 0.15$: the region bounded by the cyan sphere with radius 0.15 m in Figure 3(a).

in which $\mathbf{V}^{\text{des}}(\mathbf{r}_b, k) \equiv [V_{\hat{x}}^{\text{des}}(\mathbf{r}_b, k), V_{\hat{y}}^{\text{des}}(\mathbf{r}_b, k), V_{\hat{z}}^{\text{des}}(\mathbf{r}_b, k)]$ is the desired AVV and $\mathbf{V}^{\text{re}}(\mathbf{r}_b, k) \equiv [V_{\hat{x}}^{\text{re}}(\mathbf{r}_b, k), V_{\hat{y}}^{\text{re}}(\mathbf{r}_b, k), V_{\hat{z}}^{\text{re}}(\mathbf{r}_b, k)]$ is the reproduced AVV. Here, only the real part of the AVVs are considered. In Figure 5, the blue line and the red line illustrate the velocity reproduction errors averaged across 113081 evaluation points within the red sphere with radius 0.5 m in Figure 3(a). VM achieved lower errors below 500 Hz. Above 500 Hz, the errors of VM and PM are similar. The yellow line and the purple line show the velocity reproduction errors averaged across 2993 evaluation points within the cyan sphere of radius 0.15 m in Figure 3(a). Up to slightly above 1 kHz, VM performs better than PM. The simulation result

suggests that to reproduce the desired AVVs at low frequencies, VM requires fewer loudspeakers than PM. VM has strong potentials in home theater and small exhibition spaces, where space constraints only allow the installation of small loudspeaker array. For reproduction at mid to high frequencies, intensity based method [20,34–37] can be used.

4 Conclusion

This paper presented the SHV-indR coefficients, which were the radial independent SH coefficients of the AVVs in a spherical region. The SHV-indR coefficients were derived from the SH coefficients of the pressure in the spherical region by using the sound field translation formula. The SHV-indR coefficients were used in VM, which reproduced the AVVs in the spherical listening region by matching the SHV-indR coefficients. Simulation showed at low frequencies, VM accurately reproduced the AVVs using few number of loudspeakers. Future work will include conducting perceptual tests and investigating methods to enlarge the listening region.

References

- [1] O. Kirkeby and P. A. Nelson, “Reproduction of plane wave sound fields,” *The Journal of the Acoustical Society of America*, vol. 94, no. 5, pp. 2992–3000, 11 1993.
- [2] A. J. Berkhout, D. de Vries, and P. Vogel, “Acoustic control by wave field synthesis,” *The Journal of the Acoustical Society of America*, vol. 93, no. 5, pp. 2764–2778, 05 1993.
- [3] M. M. Boone, E. N. G. Verheijen, and P. F. van Tol, “Spatial sound-field reproduction by wave-field synthesis,” *Journal of the Audio Engineering Society*, vol. 43, no. 12, pp. 1003–1012, December 1995.
- [4] J. Ahrens, R. Rabenstein, and S. Spors, “The theory of wave field synthesis revisited,” *Journal of the Audio Engineering Society*, May 2008.
- [5] P.-A. Gauthier and A. Berry, “Adaptive wave field synthesis for active sound field reproduction: Experimental results,” *The Journal of the Acoustical Society of America*, vol. 123, no. 4, pp. 1991–2002, 04 2008.
- [6] F. Winter, F. Schultz, G. Firtha, and S. Spors, “A geometric model for prediction of spatial aliasing in 2.5 D sound field synthesis,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 27, no. 6, pp. 1031–1046, 2019.
- [7] D. Ward and T. Abhayapala, “Reproduction of a plane-wave sound field using an array of loudspeakers,” *IEEE Transactions on Speech and Audio Processing*, vol. 9, no. 6, pp. 697–707, 2001.
- [8] T. Betlehem and T. D. Abhayapala, “Theory and design of sound field reproduction in reverberant rooms,” *The Journal of the Acoustical Society of America*, vol. 117, no. 4, pp. 2100–2111, 04 2005.
- [9] M. A. Poletti, “Three-dimensional surround sound systems based on spherical harmonics,” *Journal of the Audio Engineering Society*, vol. 53, no. 11, pp. 1004–1025, November 2005.
- [10] S. Spors, H. Wierstorf, A. Raake, F. Melchior, M. Frank, and F. Zotter, “Spatial sound with loudspeakers and its perception: A review of the current state,” *Proceedings of the IEEE*, vol. 101, no. 9, pp. 1920–1938, 2013.
- [11] M. Buerger, R. Maas, H. W. Löllmann, and W. Kellermann, “Multizone sound field synthesis based on the joint optimization of the sound pressure and particle velocity vector on closed contours,” in *2015 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, 2015, pp. 1–5.
- [12] M. Buerger, C. Hofmann, and W. Kellermann, “Broadband multizone sound rendering by jointly optimizing the sound pressure and particle velocity,” *The Journal of the Acoustical Society of America*, vol. 143, no. 3, pp. 1477–1490, 03 2018.
- [13] M. Shin, P. A. Nelson, F. M. Fazi, and J. Seo, “Velocity controlled sound field reproduction by non-uniformly spaced loudspeakers,” *Journal of Sound and Vibration*, vol. 370, pp. 444–464, 2016.

- [14] H. Hacıhabiboglu, E. De Sena, Z. Cvetkovic, J. Johnston, and J. O. Smith III, “Perceptual spatial audio recording, simulation, and rendering: An overview of spatial-audio techniques based on psychoacoustics,” *IEEE Signal Processing Magazine*, vol. 34, no. 3, pp. 36–54, 2017.
- [15] M. A. Gerzon and G. J. Barton, “Ambisonic decoders for HDTV,” in *Audio Engineering Society Convention 92*, Mar 1992.
- [16] M. A. Gerzon, “General metatheory of auditory localisation,” in *Audio Engineering Society Convention 92*, Mar 1992.
- [17] H. Zuo, L. I. Birnie, P. N. Samarasinghe, T. D. Abhayapala, and V. Tourbabin, “Particle-velocity-based mixed-source sound field translation for binaural reproduction,” *Applied Sciences*, vol. 13, no. 11, 2023.
- [18] S. Wang, R. Hu, S. Chen, X. Wang, B. Peng, Y. Yang, and W. Tu, “Sound physical property matching between non central listening point and central listening point for nhk 22.2 system reproduction,” in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2017, pp. 436–440.
- [19] F. Zotter and M. Frank, *Ambisonics: A Practical 3D Audio Theory for Recording, Studio Production, Sound Reinforcement, and Virtual Reality*. Springer, 2019.
- [20] D. Arteaga, “An ambisonics decoder for irregular 3D loudspeaker arrays,” in *The 134th AES Convention*, 01 2013.
- [21] X. Hu, J. Wang, W. Zhang, and L. Zhang, “Time-domain sound field reproduction with pressure and particle velocity jointly controlled,” *Applied Sciences*, vol. 11, no. 22, 2021.
- [22] Y. Zhao, W. Zhang, and J. Chen, “A subband approach to personal sound zone with joint optimization of sound pressure and particle velocity,” in *2023 Asia Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, 2023, pp. 427–431.
- [23] H. Zuo, T. D. Abhayapala, and P. N. Samarasinghe, “Particle velocity assisted three dimensional sound field reproduction using a modal-domain approach,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 28, pp. 2119–2133, 2020.
- [24] J. Meyer and G. Elko, “A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield,” in *2002 IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 2, 2002, pp. II–1781–II–1784.
- [25] T. D. Abhayapala and D. B. Ward, “Theory and design of high order sound field microphones using spherical microphone array,” in *2002 IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 2, 2002, pp. II–1949–II–1952.
- [26] E. Erdem, Z. Cvetković, and H. Hacıhabiboğlu, “3D perceptual soundfield reconstruction via virtual microphone synthesis,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 31, pp. 1305–1317, 2023.
- [27] A. Herzog and E. A. P. Habets, “Generalized intensity vector and energy density in the spherical harmonic domain: Theory and applications,” *The Journal of the Acoustical Society of America*, vol. 150, no. 1, pp. 294–306, 07 2021.
- [28] F. Ma, T. D. Abhayapala, and W. Zhang, “Multiple circular arrays of vector sensors for real-time sound field analysis,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 29, pp. 286–299, 2021.
- [29] A. H. Moore, C. Evers, and P. A. Naylor, “Direction of arrival estimation in the spherical harmonic domain using subspace pseudointensity vectors,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 25, no. 1, pp. 178–192, 2017.
- [30] B. Rafaely, *Fundamentals of Spherical Array Processing*. Springer, 2015.
- [31] R. A. Kennedy and P. Sadeghi, *Hilbert Space Methods in Signal Processing*. Cambridge University Press, 2013.
- [32] Digital Library of Mathematical Functions, “3j, 6j, 9j Symbols,” accessed on May 5, 2024. [Online]. Available: <https://dlmf.nist.gov/34.3>

- [33] L. Birnie, T. Abhayapala, V. Tourbabin, and P. Samarasinghe, “Mixed source sound field translation for virtual binaural application with perceptual validation,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 29, pp. 1188–1203, 2021.
- [34] H. Zuo, P. N. Samarasinghe, and T. D. Abhayapala, “Intensity based spatial soundfield reproduction using an irregular loudspeaker array,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 28, pp. 1356–1369, 2020.
- [35] ———, “Intensity based soundfield reproduction over multiple sweet spots using an irregular loudspeaker array,” in *2020 28th European Signal Processing Conference (EUSIPCO)*, 2021, pp. 486–490.
- [36] H. Zuo, T. D. Abhayapala, and P. N. Samarasinghe, “3D multizone soundfield reproduction in a reverberant environment using intensity matching method,” in *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2021, pp. 416–420.
- [37] J.-W. Choi and Y.-H. Kim, “Manipulation of sound intensity within a selected region using multiple sources,” *The Journal of the Acoustical Society of America*, vol. 116, no. 2, pp. 843–852, 08 2004.