Theoretical and experimental study of DOA estimation using AML algorithm for an isotropic and non-isotropic 3D array

Shadnaz Asgar\textsuperscript{a}, Andreas M. Ali\textsuperscript{a}, Travis C. Collier\textsuperscript{b}, Yuan Yao\textsuperscript{b}, Ralph E. Hudson\textsuperscript{a}, Kung Yao\textsuperscript{a,}\textsuperscript{*}, Charles E. Taylor\textsuperscript{b}

\textsuperscript{a}Electrical Engineering Department, University of California, Los Angeles, CA, USA; \textsuperscript{b}Ecology and Evolutionary Biology Department, University of California, Los Angeles, CA, USA

ABSTRACT

The focus of most direction-of-arrival (DOA) estimation problems has been based mainly on a two-dimensional (2D) scenario where we only need to estimate the azimuth angle. But in various practical situations we have to deal with a three-dimensional scenario. The importance of being able to estimate both azimuth and elevation angles with high accuracy and low complexity is of interest. In this paper, we present the theoretical and the practical issues of DOA estimation using the Approximate-Maximum-Likelihood (AML) algorithm in a 3D scenario. We show that the performance of the proposed 3D AML algorithm converges to the Cramér-Rao Bound. We review some known results for the 3D arrays (isotropic and non-isotropic arrays). We use the concept of an isotropic array to reduce the complexity of the proposed algorithm by advocating a decoupled 3D version. We also explore a modified version of the decoupled AML algorithm which can be used for DOA estimation with non-isotropic arrays. Various numerical results are presented. We use two acoustic arrays each consisting of 8 microphones to do some field measurements. In each acoustic array, four microphones construct an isotropic subarray, while the other four have a non-isotropic configuration. The four audio channels within the same acoustic subarray are finely synchronized within a few microseconds. A speaker is hanging on from a roof with different heights and plays back a woodpecker call. Field measurements of the acoustic signals generated by the speaker in an open field are collected. The processing of the measured data from the acoustic arrays for different azimuth and elevation angles confirm the effectiveness of the proposed methods.

Keywords: DOA estimation, three-dimensional scenario, AML algorithm, isotropic array, Cramer-Rao Bound, Acorn Woodpecker

1. INTRODUCTION

The importance of direction-of-arrival (DOA) estimation in array signal processing areas such as radar, sonar, mobile communication systems and radio astronomy is undeniable. As a result, a wide variety of techniques have been proposed for the estimation of DOA of narrowband and wideband sources over the past several decades. Most of these techniques mainly focus on DOA estimation for a two-dimensional scenario (estimation of azimuth angle of the source), while in many of the practical situations we have a three-dimensional scenario. So in addition to azimuth angle, we also need to estimate the elevation angle of the source. Consequently it is very important to have an algorithm which is able to estimate both azimuth and elevation angle with sufficient accuracy. As a result, the Approximate Maximum Likelihood (AML) algorithm would be a good option for DOA estimation in a three-dimensional scenario, because it has the advantage that its performance approaches the Cramér-Rao Bound (CRB). However, a ML estimation algorithm based on a closely spaced, 2D grid search results in very high computational burden.

To decrease the complexity of the AML algorithm, we use the concept of isotropic and non-isotropic arrays. A 3D isotropic array is the one whose azimuth angle CRB (after some scaling) becomes equal to elevation angle CRB. Hawkes and Nehorai\textsuperscript{3} use the bound on asymptotic mean square angular error to define an isotropic array. Baysal and Moses\textsuperscript{1} prove the sufficient and necessary conditions for a 3D array to be isotropic. We use the properties of a 3D isotropic array to introduce the decoupled 3D AML algorithm which has much less complexity relative to the grid search based 3D AML algorithm. Then we propose a modified decoupled AML algorithm which can be used for DOA estimation of a non-isotropic array.

\textsuperscript{*}yao@ee.ucla.edu; phone 1310 206-4304; Further author information (send correspondence to Kung Yao)
To test the effectiveness of the proposed 3D AML algorithm, we do some field measurements in the relatively open area close to Geology and Chemistry building in UCLA Science Courtyard. We use two acoustic arrays each consisting of 8 M53 microphones. The four microphones on top of each array have a non-isotropic subarray configuration and the other four microphones construct an isotropic subarray. The two arrays are around 5 meters apart. We use a rope to hang a computer speaker from the roof of Geology/Chemistry. This computer speaker is used as a mock bird to play back acorn woodpecker calls. The height of the speaker is changed and for each height the waka calls are recorded synchronously by the Presonus FIREPOD recording studio that connect the microphones to a laptop computer. To do DOA estimation for the woodpecker, we process the recorded waka calls offline by using the proposed 3D AML algorithm. This offline data processing confirms the effectiveness of the proposed 3D AML algorithm. It also confirms the advantage of having an isotropic subarray configuration relative to a similar non-isotropic configuration.

The structure of the paper is as follows. In section 2, we present the details of the 3D AML algorithm and show that the performance of proposed 3D AML algorithm approaches CRB. In section 3, we review some known results on isotropic and non-isotropic arrays. In section 4, we introduce the decoupled 3D AML algorithm for isotropic arrays and the modified decoupled 3D AML algorithm for non-isotropic arrays. In Section 5, first the experiment setup and the acoustic testbed are described and then the experiment results of DOA estimation using isotropic and non-isotropic subarrays are presented. Section 6 outlines some future work and provides a brief conclusion.

2. AML ALGORITHM FOR THE THREE-DIMENSIONAL SCENARIO

![Diagram](image)

Figure 1. Geometrical relationship between the Cartesian coordinate and the spherical coordinate in a three-dimensional scenario.

Suppose that we use the Cartesian coordinate system to express the position of a general point $A$ relative to the reference point $C$, denoted by $A = [x, y, z]^T$. According to figure 1, then position of $A$ in the spherical coordinate system can be expressed as $A = [R, \theta, E]^T$, where

$$R = \sqrt{x^2 + y^2 + z^2},$$

$$\tan(\theta) = \frac{y}{x},$$

$$\tan(E) = \frac{z}{\sqrt{x^2 + y^2}}. \tag{1}$$

Now, assume that the sensor array comprises $P$ arbitrarily distributed, omni-directional sensors with identical behavior, and also assume that our reference point is the centroid of the array. Each sensor is located at $A = [R, \theta, E]^T$. We consider the...
position \( r_p = [x_p, y_p, z_p]^T \) with \( 1 \leq p \leq P \). Similarly, we assume that there are \( M \) wideband sources at unknown locations \( q_m = [x_m, y_m, z_m]^T \) with \( 1 \leq m \leq M \). Then the relative time delay of the \( m^{th} \) source is given by

\[
\tau_{c}^{(m)} - \tau_{p}^{(m)} = \frac{1}{v} \sqrt{(x_p \cos \theta_m + y_p \sin \theta_m \cos E_m + z_p \sin E_m)^2} = \frac{R_p}{v} \sqrt{(\cos(\psi_m - \theta_m) \cos (\epsilon_m) \cos (E_p) + \sin (\epsilon_m) \sin (E_p))^2}.
\]

Note that \( [R_p, \theta_p, E_p]^T \) is the position of the \( p^{th} \) sensor and \( [\xi_m, \psi_m, \epsilon_m]^T \) is the position of the \( m^{th} \) source, both in spherical coordinate system. \( \tau_{c}^{(m)} \) and \( \tau_{p}^{(m)} \) are the absolute time delays from the \( m^{th} \) source to the centroid and to the \( p^{th} \) sensor, respectively and \( v \) is the speed of sound (345 m/s). The data received by the \( p^{th} \) sensor at time \( t = t_n \), henceforth, simply denoted by the index \( n \) is given by

\[
xx_p(n) = \sum_{m=1}^{M} s^{(m)}(n - \tau_{cp}^{(m)}) + \omega_p(n), \quad n = 0, \ldots, N-1,
\]

where \( N \) is the length of the data vector, \( s^{(m)}(n) \) is the \( m^{th} \) source signal at the array centroid, and \( \omega_p(n) \) is the zero mean white Gaussian noise with variance \( \sigma^2 \). Note that in the above equation, \( \tau_{cp}^{(m)} \) can be any real-valued number. For the ease of derivation and analysis, the received wideband signal can be transformed into the frequency domain via the Discrete Fourier Transform (DFT), where a narrowband model can be attributed to each frequency bin. We use frequency domain to avoid interpolation in (3) and replace it with frequency dependent multiplier as in (4). It is well-known that the circular shift property of the DFT has an edge effect problem for the actual linear time shift. These finite effects become negligible for a sufficient long data. Here, we assume the data length \( N \) is large enough to ignore the artifact caused by the finite data length. For the \( N \) point DFT transformation, the array data model in the frequency domain is given by

\[
X(w_k) = D(w_k)S(w_k) + \eta(w_k), \quad k = 0, \ldots, N-1,
\]

where \( X(w_k) = [X_1(w_k), \ldots, X_P(w_k)]^T \) is the array data spectrum, and \( \eta(w_k) \) is the complex Gaussian random vector of independent-identically-distributed (i.i.d) components, each with a zero mean and variance of \( N \sigma^2 \). Note that due to the transformation to the frequency domain, \( \eta(w_k) \) asymptotically approaches a Gaussian distribution by the Central Limit Theorem, even if the actual noise has an arbitrary i.i.d distribution with bounded variance in the time domain. This asymptotic property in the frequency domain provides a more reliable noise model than the time domain model in some practical cases. The other terms on the right-hand side of this equation are

\[
D(w_k) = [d^{(1)}(w_k), \ldots, d^{(M)}(w_k)], \quad P \times M \text{ Steering Matrix}, \quad (5)
\]

\[
S(w_k) = [S^{(1)}(w_k), \ldots, S^{(M)}(w_k)], \quad M \times 1 \text{ Source Spectrum}. \quad (6)
\]

Note that the \( m^{th} \) steering vector is defined as

\[
d^{(m)}(w_k) = [e^{-j2\pi k(\xi_m(N)/N)}, \ldots, e^{-j2\pi k(\psi_m(N)/N)}]^T.
\]

We denote the transposition and complex conjugate transposition operations with superscripts \( T \) and \( H \), respectively, throughout this manuscript. The AML estimator performs the data processing in the frequency domain. We define
\[ Q(w_k) = D(w_k)S(w_k), \quad k = 0, \ldots, N - 1. \] (8)

Then by stacking up the \( N/2 \) positive frequency bins (zero frequency bin is not important and the negative frequency bins are merely mirror images) of the signal model into a single column, we can rewrite the sensor data into a \( \frac{NP}{2} \times 1 \) space-temporal frequency vector as

\[ X = G(\Theta) + \xi, \] (9)

where \( G(\Theta) = [Q^T(w_k), \ldots, Q^T(w_{N/2})]^T \) and \( R_\xi = E[\xi\xi^H] = (N\sigma^2)I_{NP/2} \).

We assume, initially, that the unknown parameter space is

\[ \Theta = [\xi^T, S_0^{(1)}^T, \ldots, S_0^{(M)}^T]^T, \] (10)

where the \( M \) source locations are denoted by

\[ \hat{\xi} = [\xi_1^T, \ldots, \xi_M^T]^T, \] (11)

and the \( m^{th} \) source signal spectrum is denoted by

\[ S_0^{(m)} = [S^{(m)}(w_1), \ldots, S^{(m)}(w_{N/2})]^T. \] (12)

Note that the probability density function of the complex Gaussian noise vector \( \xi \) is given by

\[ p(\xi) = \prod_{i=1}^{NP/2} \frac{1}{\sqrt{2\pi N\sigma^2}} e^{-\frac{\xi_i^2}{2N\sigma^2}}. \] (13)

Then the log-likelihood function of vector \( \xi \) becomes

\[ \log(p(\xi)) = \sum_{i=1}^{NP/2} \log\left( \frac{1}{\sqrt{2\pi N\sigma^2}} \right) - \frac{1}{2N\sigma^2} \| X - G(\Theta) \|^2. \] (14)

So the maximum likelihood estimation of the source DOA and the source signals, after ignoring irrelevant constant terms is given by

\[ \max_{\Theta} \quad L(\Theta) = \max_{\Theta} \quad (-\| X - G(\Theta) \|^2) \]

\[ = \min_{\Theta} \quad \sum_{k=1}^{N/2} \| X(w_k) - D(w_k)S(w_k) \|^2. \] (15)

(15) is equivalent to finding \( \min_{(\xi, S(w_k))} f(w_k) \) for each frequency bin \( w_k, k \in \{1, \ldots, N/2\} \), where

\[ = \min_{\Theta} \quad \sum_{k=1}^{N/2} \| X(w_k) - D(w_k)S(w_k) \|^2. \] (16)

The minimal of \( f(w_k) \) with respect to the source signal vector \( S(w_k) \) must satisfy
\[ \frac{\partial f(w_k)}{\partial S^H(w_k)} = 0. \]  

(17)

Hence, the estimate of the source signal vector that yields the minimum residual at any source location is given by

\[ \hat{S}(w_k) = D^\dagger(w_k)X(w_k), \]  

(18)

where \( D^\dagger(w_k) = (D^H(w_k)D(w_k))^{-1}D^H(w_k) \) is the pseudo inverse of the steering matrix \( D(w_k) \).

Now, we define the orthogonal projection \( P(w_k, \tilde{r}_k) = D(w_k)D^\dagger(w_k) \). So the complement orthogonal projection becomes \( P^\perp(w_k, \tilde{r}_k) = I - P(w_k, \tilde{r}_k) \), where \( I \) is the identity matrix is. By combining the last two equations, the minimization function becomes

\[ f(w_k) = \| P^\perp(w_k, \tilde{r}_k)X(w_k) \|^2. \]  

(19)

After substituting the estimate of the \( S(w_k) \), the AML source location estimate can be obtained by solving the following problem

\[ \max_{\tilde{r}_k} J(\tilde{r}_k) = \min_{\tilde{r}_k} \sum_{k=1}^{N/2} (\| P^\perp(w_k, \tilde{r}_k)X(w_k) \|^2) = \max_{\tilde{r}_k} \sum_{k=1}^{N/2} (\| P(w_k, \tilde{r}_k)X(w_k) \|^2) \]

\[ = \max_{\tilde{r}_k} \sum_{k=1}^{N/2} \text{tr}(X^H(w_k)P^H(w_k, \tilde{r}_k)P(w_k, \tilde{r}_k)X(w_k)). \]  

(20)

Note that \( P(w_k, \tilde{r}_k) \) is an orthogonal projection, so we have \( P(w_k, \tilde{r}_k) = P^H(w_k, \tilde{r}_k) = P^2(w_k, \tilde{r}_k) \). Using this note, we reach

\[ \max_{\tilde{r}_k} J(\tilde{r}_k) = \max_{\tilde{r}_k} \sum_{k=1}^{N/2} \text{tr}(P(w_k, \tilde{r}_k)R(w_k)), \]  

(21)

where \( R(w_k) = X(w_k)X^H(w_k) \) is the one snapshot covariance matrix. This multi-source AML algorithm performs signal separation by utilizing the physical separation of the sources, and for each source signal, the Signal-to-Noise-Ratio (SNR) is maximized in the ML sense. Note that no closed form solution can be obtained here.

2.1 Single source AML algorithm for the three-dimensional scenario

Suppose we have an array consisting of \( P \) sensors, where the array centroid is the reference point of the spherical coordinate system. It can be shown that if we have only one source represented with the vector \( [\delta, \psi, \epsilon]^T \), then the maximum likelihood criterion in (21) can be simplified as

\[ \max_{\tilde{r}_k} J(\tilde{r}_k) = \max_{\tilde{r}_k} \sum_{k=1}^{N/2} (X^H(w_k)D(w_k))^2, \]  

(22)

where for the single source case, \( D_{P \times 1}(w_k) = [e^{-j2\pi k_1 \frac{1}{N}}, ..., e^{-j2\pi k_p \frac{1}{N}}]^T \) and the relative time delay between the centroid and to the \( p^{th} \) sensor would be \( t_{cp} = (\frac{R_p}{v}) \cdot [\cos(\psi - \theta_p), \cos(\epsilon), \cos(E_p) + \sin(\epsilon) \sin(E_p)] \).

This latest form of (22) makes the computation cost less than the original form of ML criterion in (21). Even so, a fine grain grid search for the 3D AML algorithm is still computationally intensive. Now, to obtain the estimated azimuth and elevation angle using the AML algorithm, we do a 2D search on all possible azimuth angles (from 0 degree to 360 degree) and all possible elevation angles (from 0 degree to 90 degree) to maximize \( J(\tilde{r}_k) \) in (22). For example, suppose
we have a square array with side length of 1 meter in XY plane. Let our source be a \textit{vehicular sound}, with sampling frequency of 800 Hz. Our reference point is the zero of the spherical coordinate system and we consider a SNR = 20 dB. Also suppose our source is located at azimuth angle of 75 degree and elevation angle of 60 degree. Figure 2 shows the computed $J(\hat{\phi}_s)$ for all possible azimuth angles (from 0 degree to 360 degree) and all possible elevation angles (from 0 degree to 90 degree). As is shown on the figure, the maximum of $J(\hat{\phi}_s)$ happens exactly when azimuth angle is 75 degree and elevation angle is 60 degree. So the estimated azimuth angle is equal to the true azimuth angle and the estimated elevation angle is equal to the true elevation angle and this shows the accuracy of the 3D AML algorithm. The performance of the suggested 3D AML algorithm is very good at high SNR. To show this property, first we need to compute the CRB for a three-dimensional scenario.

![Figure 2](image-url)  

**Figure 2.** Computed ML criterion $J(\hat{\phi}_s)$ when the source is located as azimuth angle of 75 degree and elevation angle of 60 degree.

### 2.2 Cramér-Rao bound for DOA estimation in a single source three-dimensional scenario

The CRB is most often used as a theoretical lower bound for any unbiased estimator. In this section, we derive the CRB directly from the signal model. We can construct the FIM (Fisher Information Matrix)\(^{26}\) from the signal model defined in section 2.1 by

$$F = 2 \text{Re}[H^H R_\alpha^{-1} H] = \frac{2}{N\sigma^2} \text{Re}[H^H H], \quad (23)$$

Note that $R_\alpha = E[\Sigma_{\phi}^H] = (N\sigma^2)I_{NF/2}$ and $H = \frac{\partial G}{\partial \phi^T}$ and $\phi = [\psi, e]^T$. Vector $G$ is the defined in (9) and can be written as the following for the single source case

$$[G]_{NP} = [e^{-j 2 \pi k / N} S_1(w_1),...,e^{-j 2 \pi k / N} S_1(w_{N/2}),...,e^{(N/2)\pi - j 2 \pi k / N} S_1(w_{N/2})] \quad (24)$$

In this case,

$$F = \alpha T,$$

$$\alpha_{1x1} = \frac{2}{N\sigma^2} \sum_{k=1}^{N/2} \left(\frac{2\pi k | S_1(w_k)|}{N}\right)^2$$

and

$$T = \sum_{p=1}^{L} \left[\begin{array}{ccc}
\frac{\partial t_{cp}}{\partial \psi} & \frac{\partial t_{cp}}{\partial \psi} & \frac{\partial t_{cp}}{\partial e} \\
\frac{\partial t_{cp}}{\partial e} & \frac{\partial t_{cp}}{\partial e} & \frac{\partial t_{cp}}{\partial e}
\end{array}\right]. \quad (25)$$
As an example to see the performance of suggested 3D AML algorithm relative to CRB, suppose that we have a cube with side length of 3 cm which is centered at the origin. This cube has 6 side planes and we put one sensor in the middle of each side plane. As the result, we would have a 3D array with 6 sensors which is shown on Figure 3(a). Matrix $\mathbf{M}$ can be used to express the array geometry. Each column of matrix $\mathbf{M}$ shows the position of the corresponding sensor in the array.

$$
\begin{bmatrix}
0 & 1.5 & 0 & -1.5 & 0 & 0 \\
-1.5 & 0 & 1.5 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1.5 & -1.5
\end{bmatrix} 
$$

(26)

Assume our source signal is a male Dusky antbird call which is located at azimuth angle of 155 degree and elevation angle of 60 degree. Figure 3(b) shows the MSE (Mean square Error) in degree square obtained by running the 3D AML algorithm for 1000 different noise realizations and compares it to the corresponding CRB in both azimuth and elevation angles. As can be seen from this figure, the 3D AML performance is very close to the CRB and this confirms the high accuracy capability of the AML algorithm.

![Cubic array configuration](image)

(a) Cubic array configuration

![Comparison of AML performance and CRB for azimuth angle](image)

(b) The AML algorithm performance and the CRB when the source is located at azimuth angle of 155 and elevation angle of 60 degree.

As this figures show, this version of the 3D AML algorithm has good performance, especially in high SNR. But the main objection with this algorithm is its high complexity. As we saw earlier, to compute the AML metric $J(\hat{\mathbf{r}}_s)$, we have to do a 2D search on all possible azimuth and all possible elevation angles and this 2D search makes the complexity of the 3D AML algorithm high. To reduce the complexity we use the concept of isotropic and non-isotropic arrays which we will review in the next section.

### 3. A REVIEW ON ISOTROPIC AND NON-ISOTROPIC ARRAYS

Suppose that our source is located at $[d, \psi, \epsilon]^T$ in the spherical coordinate system. So the vector

$$
u = [\cos(\psi)\cos(\epsilon), \sin(\psi)\cos(\epsilon), \sin(\epsilon)]^T,
$$

is the unit bearing vector pointing toward the source signal in three-dimensions. In this case, estimating the azimuth and elevation angles corresponding to the DOA of the source signal is equivalent to estimating the vector $\hat{u}$. Now, Assume that $\delta$ is the angle between the vectors $u$ and $\hat{u}$. Nehorai and Paldi\cite{Nehorai1987} have defined the Mean Square Angular Error (MSAE) as the expectation of $\delta^2$. This MSAE is a good measure of the estimator performance. It is independent of the choice of reference coordinate system and it does not suffer from the singularity inherent in the spherical coordinate system when elevation angle approaches $\pm 90$ degree. The lower bound on MSAE provides a guide to achievable
performance and estimator efficiency. Nehorai and Paldi\textsuperscript{7} have derived this lower bound in details. They have also provided the conditions for the applicability and tightness of the bound. A lower bound of unit length unbiased estimator of $u$ is given by $MSAE_B \leq \cos^2(e)CRB(\psi) + CRB(e)$.

The array is said to be isotropic if the bound on the MSAE is constant for all azimuth and all elevation angles. Baysal and Moses\textsuperscript{4} have proved the following theorem which gives the necessary and sufficient conditions for an array to be isotropic:

**Theorem:** Suppose that an $P$ element 3D array which is centered at the origin, is represented by the array geometry matrix $B$ as

$$B_{3\times3} = \sum_{p=1}^{P} \ell_p \ell_p^T,$$  \hspace{1cm} (28)

where $\ell_p$ is the normalized location of the $p^{th}$ sensor in the Cartesian coordinate system. Note that $\ell_p = r_p \times \frac{f_s}{v}$ and $f_s$ is the sampling frequency. Then the array is isotropic, if and only if $B = kI_{3\times3}$, where $I$ is the identity matrix. It can be shown that the FIM for an isotropic array would be as

$$F = \frac{1}{\alpha k} \begin{bmatrix} k \cos^2 e & 0 \\ 0 & k \end{bmatrix},$$  \hspace{1cm} (29)

where $\alpha$ is the scalar defined in (25). So we have

$$CRB(\psi) = \frac{1}{\alpha k \cos^2 e},$$

$$CRB(e) = \frac{1}{\alpha k}.$$  \hspace{1cm} (30)

So if we remove scaling of CRB of azimuth angle, then the CRB of the elevation angle and the CRB of azimuth angle become equal and as the result, the bound on MSAE becomes constant for all azimuth and elevation angles

$$MSAE_B = \cos^2(e)CRB(\psi) + CRB(e) = \frac{2}{\alpha k}.$$  \hspace{1cm} (31)

As an example, we consider the array configuration suggested by Girod\textsuperscript{5} as it is shown on figure 4(a). In this array we have four sensors where three of them are located at the corner of a square with side length of 8 cm while the last one is located 14 cm above the square plane. Again we can use each column of the matrix $\mathcal{B}$ to express the position of corresponding sensor in the array.

Now, we can compute the array geometry matrix $B$ for this specific array configuration. The matrix $\mathcal{B}$ and the computed matrix $B$ for the Girod array would be given by

$$M = \begin{bmatrix} 4 & 4 & -4 & -4 \\ -4 & 4 & 4 & -4 \\ -4 & 10 & -4 & -4 \end{bmatrix}$$

and

$$B = \left(\frac{f_s}{v}\right)^2 \times 10^{-4} \times \begin{bmatrix} 64 & 0 & 56 \\ 0 & 64 & 56 \\ 56 & 56 & 148 \end{bmatrix},$$  \hspace{1cm} (32)

where the coefficient of $10^{-4}$ is because of centimeter to meter conversion. Since matrix $B \neq kI$, then according to the given theorem, this array configuration is not isotropic. Figure 4(b) shows the CRB plots of azimuth and elevation angle. The source signal is a Dusky antbird with SNR=10 dB. As the plots show, $CRB(e)$ for this array is not constant and that is exactly what we expect for a non-isotropic array.
Now, we propose a modification to the Girod array as follows. Suppose we decrease the height of sensor 2 from (10 cm) to (4 cm) and increase the height of sensor 4 from (-4 cm) to (4 cm). The resulted array configuration would be the one which is showed in Figure 5(a). The matrix $\mathbf{U}$ and the computed matrix $B$ for this new array would become

\[
\begin{bmatrix}
4 & 4 & -4 & -4 \\
-4 & 4 & 4 & -4 \\
-4 & 4 & -4 & 4
\end{bmatrix}
\quad \text{and} \quad
B = \left(\frac{f}{v}\right)^2 \times 10^{-4} \times \begin{bmatrix} 64 & 0 & 0 \\ 0 & 64 & 0 \\ 0 & 0 & 64 \end{bmatrix} = kl. \quad (33)
\]

As a result, the proposed array configuration is now isotropic. Figure 5(b) shows the computed CRB of azimuth angle and elevation angle for the modified array. In contrast, to the $CRB(e)$ for Girod array, here $CRB(e)$ is constant and this is the direct result of being isotropic.

The comparison of the $MSAE_B$ for the Girod array and the modified array are shown in Figure 5. $MSAE_B$ for the modified array is constant while the $MSAE_B$ for the Girod array is not constant. In the next section, we will use some of the properties of isotropic array to decrease the complexity of the 3D AML algorithm.
4. LOW COMPLEXITY 3D AML ALGORITHM

We saw that performance of the proposed 3D AML is good especially for high SNR. But the main objection with this algorithm which is its high complexity still remains unsolved. In this section, we show that by using decoupled 3D AML algorithm, we can reduce the algorithm complexity for isotropic arrays. We also propose a modified version of decoupled 3D AML algorithm which can be used for non-isotropic arrays.

4.1 Decoupled 3D AML algorithm for isotropic arrays

In the previous section, we showed that the Fisher information matrix for an isotropic array is diagonal. Consequently, we can decouple the azimuth and the elevation angle estimation if we are dealing with an isotropic array. This is a significant advantage, since we know that for a general array in a 3D scenario, we have to do a two dimensional search on all possible azimuth and all possible elevation angles, and this 2D dimensional search has high complexity. Now, if we use the decoupling property which exists for an isotropic array, we only need to do a 1D search twice. (Once we fix the elevation angle and find the azimuth angle, and the other time we fix the azimuth angle and find the elevation angle). The reduction in the complexity due to decoupling is such that the complexity of the decoupled version is \((180/\lambda)\) times less than the complexity of the original 3D AML algorithm of section 2, where \(\lambda\) is the angular accuracy of our DOA estimation in degree. For example, if \(\lambda = 1\) degree, then the complexity of the decoupled version of AML algorithm would be \((1/180)\) of the complexity of the original 3D AML algorithm. The advantage of the decoupled 3D AML algorithm is that it still has very good performance although its complexity is much lower. Figure 6(a) compares the performance of the 3D AML algorithm and the decoupled 3D AML algorithm using the modified array configuration of Figure 4(a) for different SNR. Note that the source signal is again a male Dusky antbird located at azimuth angle of 100 degree and elevation angle of 50 degree. The plot shows, the decoupled 3D AML algorithm performance is very close to performance of the general 3D AML algorithm and both performances converges to the CRB for high SNR. This plot confirms that the decoupled 3D AML algorithm has good accuracy with low complexity which is desirable.

4.2 Modified decoupled 3D AML algorithm for non-isotropic arrays

In this section, we propose a modified version of the decoupled 3D AML algorithm which can be used for DOA estimation of non-isotropic arrays. We know that for a non-isotropic array the array geometry matrix \(B\) is not a multiple of the identity matrix. Now, if we do the eigenvalue decomposition of the array geometry matrix given by

\[
B=VDV^{-1} \quad V=[v_1 \quad v_2 \quad v_3] \quad D=\begin{bmatrix}
\Lambda_1 & 0 & 0 \\
0 & \Lambda_2 & 0 \\
0 & 0 & \Lambda_3
\end{bmatrix},
\]

(34)

where the columns of \(V\) are the eigenvectors and \(D\) is the diagonal matrix of the eigenvalues, then for a non-isotropic array at least one of the eigenvalues is not equal to the others. Let us define a new coordinate system whose basis are the
unit-length eigenvectors $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$. By this coordinate system rotation, we have converted our original non-isotropic array with an array geometry matrix $B$ to a semi-isotropic array with new array geometry matrix of $D$. Now, we can apply our decoupled 3D AML algorithm using Alternating Projection (AP) on this new coordinate system and find the DOA. Then we have to convert the newly estimated DOA to a DOA in the original coordinate system. We called this new method as the modified decoupled 3D AML algorithm. The main advantage of the modified decoupled 3D AML algorithm is that it has much less complexity relative to the original 3D AML algorithm. Figure 5(b) shows the comparison of performances of the modified decoupled 3D AML algorithm and the original 3D AML algorithm for different SNR when we are using the Girod array configuration. Note for this plot, the source is located at azimuth angle of 60 degree and elevation angle of 30 degree. As we can see, the performance of the proposed method is close to the performance of the original 3D AML algorithm especially for high SNR and this confirms the effectiveness of the modified decoupled 3D AML algorithm.

![Graphs showing performance comparison for azimuth and elevation angles](image)

Figure 5. Performance of decoupled 3D AML for an isotropic array and performance of modified decoupled 3D AML for a non-isotropic array

5. DOA ESTIMATION EXPERIMENT (3D SCENARIO)

To test the effectiveness of proposed methods in a relatively open environment, we conducted some experiments in the Science Courtyard at UCLA. There is a high structure (like a big porch with high and big columns) between Geology and Chemistry Buildings. The high roof of this structure was a good feature to have different elevation angle. There was persistent noise from ventilation systems of surrounding buildings. In addition, it was also very windy during the experiments in the Science Courtyard. The speaker, the acoustic arrays and the location of the experiment are shown in figure 7 and 8. In the following sub-sections we are going to describe the experiment setup and the results.

5.1 Experiment setup

We had two acoustic arrays as Array1 and Array2. Each Array consists of 8 MS3 microphones. The four microphones on top of each array have the Girod array configuration which we saw in figure 3(a). The remaining four microphones construct an isotropic array configuration as we saw in figure 4(a). So Array1 consists of two smaller subarrays (subarray Girod_1 and subarray iso_1), also Array2 consists of two smaller subarrays (subarray Girod_2 and subarray iso_2). The picture of one of the arrays and its configuration is shown in figure 9. The output of the 8 microphones goes to a Presonus FIREPOD (8 channel 24bit/96ks) recording system. The firepod has 8 separate microphone/preamp input and output which are enough synchronized and can go to a laptop.
Figure 7. DOA estimation experiment at the science courtyard in UCLA using two acoustic arrays

Figure 8. The position of two acoustic arrays and the speaker (view from the top)

Figure 9. The acoustic array consisting of two subarrays (one non-isotropic subarray and one isotropic subarray)
The two subarrays are about 5 meters apart. A computer speaker is hanging from the roof (with different heights) and is playing back an acorn woodpecker audio file. The audio file time duration is 7.2 seconds and consists of 14 consecutive waka calls (each call is approximately 0.4 second). Figure 10 shows the plot of one waka call in time and frequency. We repeat the audio file for 3 times (with 0.5 seconds silence between each two repetition). So the total time duration of the recorded data is 21.7 seconds. On the next section we present the results of processing the received signal from the microphones of each subarray.

![One call of woodpecker in time domain](image)

![FFT of one call for woodpecker call](image)

(a) time domain  
(b) frequency domain

Figure 10. Waka call of an acorn woodpecker

5.2 Experiment Results

On the first experiment, we put the speaker on top of the roof (with height of 8.8 meters). We process the received signal from each subarray by using 3D AML algorithm. As the result, we would be able to estimate an azimuth angle and an elevation angle for each waka call. Since our audio file consists of 14 waka calls and we repeated the audio file for 3 times, we end up having 42 estimated azimuth angle and 42 estimated elevations. Table 1 and Table 2 show the average and standard deviation (SD) of estimated DOAs for Array1 and Array2, respectively. (The results are in degrees). As we see from the tables, the estimated DOAs are fairly close to the true DOAs. This confirms the good performance of the proposed 3D AML algorithm. We also observe that the isotropic subarray has a little better azimuth angle estimation relative to non-isotropic arrays. This is consistent with that fact that CRB of azimuth angle for the applied isotropic subarray is always greater than CRB of azimuth angle for Girod subarray (figures 3 and 4), while the CRB of elevation angle of applied isotropic subarray can be smaller or greater than CRB of elevation angle of Girod subarray depending on DOAs. As the result we see that for Array1 (with azimuth angle of 90 degree and elevation angle of approximately 51 degree), the elevation angle estimation of iso_1 is worse than Girod_1. While for Array2 (with azimuth angle of 130 degree and elevation angle of approximately 42 degree), the elevation angle estimation of iso_2 is better than Girod_2.

On the second experiment we decrease the height of speaker to 7.8 meters. The results of DOA estimation for this experiment are shown in Table 2. The result of DOA estimation for the last experiment where the speaker height is 2.25 meters are shown in Table 3.

6. CONCLUSION AND FUTURE WORK

In this paper, the AML DOA estimation solution is derived for a 3D scenario. For a single source case, we also derived a more simple form of the ML criterion. The computation of the CRB for a 3D scenario is also discussed in details. We showed that the 3D AML performance approaches the CRB. We reviewed some of the definitions and properties for an isotropic array. We showed how to use these properties to make the decoupled version of 3D AML algorithm for DOA estimation for isotropic arrays. The decoupled 3D AML algorithm has the same performance as the original 3D AML.
algorithm but with much less complexity. We also proposed a modified version of the decoupled 3D AML algorithm which can be used for DOA estimation of non-isotropic arrays. It is shown that performance of this modified version is relatively close to the original 3D AML algorithm especially for high SNR. To test the effectiveness of the proposed methods, we did some experiments at the UCLA Science Courtyard. We used two arrays each consisting of 8 microphones where four of the microphones had a non-isotropic configuration and the other four had isotropic configuration. A speaker was hanging from a roof with different heights. The data processing of the received signal by the two arrays confirm the good performance of suggested 3D AML and the difference between performance of an isotropic and a non-isotropic array. Our future work will involve doing some field measurement with more than one source and trying to null out the non-desired source.

Table 1. Result of DOA estimation for Array 1 (Grid_1 subarray and iso_1 subarray) when speaker height is 8.8 meters

<table>
<thead>
<tr>
<th>Result for Grid_1</th>
<th>True DOA</th>
<th>Average of estimated DOA</th>
<th>SD of estimated DOA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Azimuth angle in degree</td>
<td>90</td>
<td>93.1</td>
<td>2.1</td>
</tr>
<tr>
<td>Elevation angle in degree</td>
<td>50</td>
<td>48.8</td>
<td>1.5</td>
</tr>
</tbody>
</table>

Table 1. Result of DOA estimation for Array 2 (Grid_2 subarray and iso_2 subarray) when speaker height is 8.8 meters

<table>
<thead>
<tr>
<th>Result for Grid_2</th>
<th>True DOA</th>
<th>Average of estimated DOA</th>
<th>SD of estimated DOA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Azimuth angle in degree</td>
<td>130</td>
<td>126.8</td>
<td>2.35</td>
</tr>
<tr>
<td>Elevation angle in degree</td>
<td>42</td>
<td>43.77</td>
<td>2.66</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Result for iso_2</th>
<th>True DOA</th>
<th>Average of estimated DOA</th>
<th>SD of estimated DOA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Azimuth angle in degree</td>
<td>130</td>
<td>132.37</td>
<td>2.16</td>
</tr>
<tr>
<td>Elevation angle in degree</td>
<td>43</td>
<td>42.1</td>
<td>2.4</td>
</tr>
</tbody>
</table>

ACKNOWLEDGEMENT

This work is partially supported by NSF CENS program under Cooperative Agreement CCR-012, NSF grant EF-0410438, and UC Discovery grant sponsored by ST Microelectronics.

REFERENCES