

Article

A Coherent Wideband Acoustic Source Localization Using a Uniform Circular Array

Meng Jiang , Chibuzo Joseph Nnonyelu , Jan Lundgren , Göran Thungström  and Mårten Sjöström 

Sensible Things that Communicate Research Centre, Mid Sweden University, 852 30 Sundsvall, Sweden

* Correspondence: meng.jiang@miun.se; Tel.: +46-10-142-7937

Abstract: In modern applications such as robotics, autonomous vehicles, and speaker localization, the computational power for sound source localization applications can be limited when other functionalities get more complex. In such application fields, there is a need to maintain high localization accuracy for several sound sources while reducing computational complexity. The array manifold interpolation (AMI) method applied with the Multiple Signal Classification (MUSIC) algorithm enables sound source localization of multiple sources with high accuracy. However, the computational complexity has so far been relatively high. This paper presents a modified AMI for uniform circular array (UCA) that offers reduced computational complexity compared to the original AMI. The complexity reduction is based on the proposed UCA-specific focusing matrix which eliminates the calculation of the Bessel function. The simulation comparison is done with the existing methods of iMUSIC, the Weighted Squared Test of Orthogonality of Projected Subspaces (WS-TOPS), and the original AMI. The experiment result under different scenarios shows that the proposed algorithm outperforms the original AMI method in terms of estimation accuracy and up to a 30% reduction in computation time. An advantage offered by this proposed method is the ability to implement wideband array processing on low-end microprocessors.

Keywords: array manifold interpolation; direction of arrival estimation; wideband sources



Citation: Jiang, M.; Nnonyelu, C.J.; Lundgren, J.; Thungström, G.; Sjöström, M. A Coherent Wideband Acoustic Source Localization Using a Uniform Circular Array. *Sensors* **2023**, *23*, 5061. <https://doi.org/10.3390/s23115061>

Academic Editors: Fangqing Wen, Jin He, Veerendra Dakulagi and Wei Liu

Received: 18 April 2023

Revised: 19 May 2023

Accepted: 22 May 2023

Published: 25 May 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In sound source localization (SSL), also known as acoustic direction of arrival estimation (DoA), the bearing angles (and in some cases range) of sound sources are estimated from the measurements taken by sampling the acoustic waves generated by the sources at different positions and orientation in space. SSL for wideband sources can be very computationally demanding due to the complexity of the existing algorithms. Thus, a low-complexity and computationally efficient algorithm will greatly benefit embedded system real-time applications with a low-end microprocessor.

Direction of arrival estimation finds application in speaker identification and localization [1,2], source localization in humanoid robots and drones [3,4], ground vehicle tracking [5], geometrical room modeling [6], and many others. Various direction of arrival estimation algorithms for narrowband sources have been proposed over the years. Some of the well-known algorithms include the Multiple Signal Classification (MUSIC) algorithm and its variants [7,8], the estimation of signal parameters via rotational invariance (ESPRIT) algorithm and its variants [9], minimum variance distortionless response (MVDR) and its variants [10,11], the generalized correlation method [12], and many others [13].

In general, most real-life sound sources are wideband signals; hence, the narrowband assumption adopted for the development of the narrowband beamforming method does not hold. For wideband sources, the phase difference between the outputs of the sensors is not only dependent on the direction of arrival, but also on the temporal frequency. Some methods have been proposed to overcome this challenge. The wideband signal can be seen as a collection of multiple narrowband signals; hence, the narrowband beamforming

methods can be applied to each narrowband bin and averaged over the entire wideband. These approaches are referred to as the incoherent signal subspace method (ISSM). Examples of such methods are the incoherent MUSIC (iMUSIC) [14], test of orthogonality of projected subspaces (TOPS) [15], test of orthogonality of frequency subspaces (TOFS) [16], and weighted squared test of orthogonality of projected subspaces (WS-TOPS) [17]. This approach suffers not only from severe degradation when correlated multipath exists but also lacks statistical stability compared to the wideband coherent methods. However, given sufficient signal-to-noise ratio (SNR), the iMUSIC algorithm performs well yielding sharp peaks in the beampattern [18]. It requires a frequency selection step as the inclusion of low-SNR frequency bins tends to degrade the resulting beampattern, and thereby reducing the source peak and introducing spurious peaks.

An alternative approach to wideband problems is the coherent method. In this approach, the covariance matrix of the component narrowbands is coherently summed into one narrowband covariance matrix, whereby the narrowband subspace method is applied to the universal covariance matrix. This is achieved by multiplying each narrowband cross-covariance matrix by a focusing matrix and then summed to obtain the steered covariance matrix. The earliest of this approach was proposed as the Coherent Signal-Subspace (CSS) method [19]. Other coherent signal-subspace methods have since been proposed: the least square and total least-square coherent signal subspace (CSS-MTLS) [20], Particle Swarm Optimization algorithm (PSO) [21], Dir-MUSIC algorithm using the strength proportion (DirSP) [22], and array manifold interpolation (AMI) [23]. A major advantage of the CSS methods is their ability to handle correlated sources. On the other hand, the performance of the CSS is sensitive to the initial values of the directions of arrival. This is because the focusing matrix of the CSS method is dependent on both directions of arrival and narrowband frequencies, making the initial estimation of the directions of arrivals an intermediate step.

The array manifold interpolation (AMI) method [23] is made possible by the series expansion of the plane wave. This gives a focusing matrix that does not depend on the initial direction of arrival of the incident signals, but only on the narrowband frequencies [18,23]. This leads to a less computational CSS method for wideband sources. The complexity is further reduced for the uniform circular array as the sensor spacings are identical [18]. However, AMI depends on the Bessel function of the first kind, which may make the implementation of this algorithm on embedded systems computationally inefficient.

Following the advances in the fields of embedded systems and Internet-of-Things (IoT), SSL and DoA methods are now being implemented on embedded systems. They are sometimes implemented together with other functionality, competing over the computational resources available on the embedded systems. Recent research in the fields of SSL and DoA has been conducted, where computational complexity is in focus, both on a model development level [24] and in more implementation on application papers [25]. These are, however, targeting narrowband signals, which cannot be adopted—even with slight modification—for wideband sources.

In this paper, a modified array manifold interpolation method is proposed for the uniform circular array of omnidirectional microphones. The proposed method takes advantage of the Bessel function's asymptotic property and forms a UCA-specific focusing matrix. Combining the assumption of low-frequency wideband signals with an adequate signal-to-noise ratio, the proposed focusing matrix can be approximated as the product of a discrete Fourier transform (DFT) matrix and a diagonal matrix. The entries of the diagonal matrix are the n th power of the ratio of the central processing frequency and the narrowband frequencies. Thus, the computational complexity reduction can be done by eliminating the calculation of the Bessel function. The simulation and experimentation results of the proposed method are compared with iMUSIC, WS-TOPS, and the original AMI method due to these particular existing methods do not require an initial estimate of the DoA [26]. The proposed modification reduces the computational complexity required for generating the focusing matrix, and yields a similar direction of arrival estimation performance to the original AMI method.

The rest of this paper is organized as follows: the signal model and array manifold interpolation are described in Section 2, the proposed algorithm is developed in Section 3, its efficacy is shown by Monte Carlo simulation in Section 4, and the laboratory experiment reported in Section 5. Finally, the paper is concluded in Section 6.

2. Array Manifold Interpolation

The signal model is described in Section 2.1 where the received data and array manifold are described. The steered covariance matrix via array manifold interpolation is introduced in Section 2.2 and applied to the uniform circular array in Section 3.1.

2.1. Signal Model

Considering L (either known a priori or estimated) number of wideband sources incident on a uniform circular array of radius r with M omnidirectional microphones, as shown in Figure 1, which illustrates the ℓ th incident source. The bandwidth of the sources are not necessarily identical but must exist within some overlap bounded by $[f_{\min}, f_{\max}]$. The output of the m th sensor is given as:

$$x_m(t) = \sum_{\ell=1}^L s_{\ell}(t - \tau_{\ell m}) + v_m(t) \quad (1)$$

where $s_{\ell}(\cdot)$ is the ℓ th incident source, $\tau_{\ell m} = \mathbf{k}_{\ell}^T \mathbf{r}_m / \omega$ is the ℓ th signal's time-delay-of-arrival for the m th sensor and the reference point (center of the circular array), $\omega = 2\pi f$ is the source angular frequency, $\mathbf{k}_{\ell} = (2\pi/\lambda)[\cos(\phi_{\ell}) \sin(\theta_{\ell}) \sin(\phi_{\ell}) \sin(\theta_{\ell}) \cos(\theta_{\ell})]^T$ is the wavenumber of the ℓ th source, and $\mathbf{r}_m$ is the position vector of the m th sensor. Furthermore, in (1), $v_m(t)$ is the additive noise on the m th sensor, assumed to be zero-mean spatio-temporally uncorrelated Gaussian noise with a priori known variance. Assuming the source and array lie on the same plane (where $\theta_{\ell} = 90^\circ$), for the uniform circular array, the time delay is $\tau_{\ell m} = kr_m \cos(\phi_{\ell} - \psi_m) / \omega$, where $\phi_{\ell} \in [0, 2\pi)$ is the azimuth angle of incidence of the ℓ th source and $\psi_m = 2\pi m / M$ is the angle between the line from the origin to the m th sensor and the positive x -axis. The wavenumber magnitude $k = |\mathbf{k}_{\ell}| = 2\pi f / c_{\text{air}}$, where $f \in [f_{\min}, f_{\max}]$ is the source-frequency, and $c_{\text{air}} = 343 \text{ ms}^{-1}$ is the speed of sound propagation in dry air. With an adequate number of observations, the j th sample of the discrete Fourier transform (DFT) of the m th sensor's output:

$$X_m(k_j) = \mathbf{A}(\phi, k_j) \mathbf{S}(k_j) + V_m(k_j), \quad j = 1, 2, \dots, J \quad (2)$$

where J is the total number of frequency samples, $\mathbf{A}(\phi, k_j) := [\mathbf{a}(\phi_1, k_j) \mathbf{a}(\phi_2, k_j) \dots \mathbf{a}(\phi_L, k_j)]$ is the steering matrix, $\mathbf{S}(k_j) = [S_1(k_j) S_2(k_j) \dots S_L(k_j)]^T$ is the DFT of the incident signals, $V_m(k_j)$ is the DFT of the additive noise, and:

$$\mathbf{a}(\phi_{\ell}, k_j) = [e^{ik_j r \cos(\phi_{\ell} - \psi_1)} e^{ik_j r \cos(\phi_{\ell} - \psi_2)} \dots e^{ik_j r \cos(\phi_{\ell} - \psi_M)}]^T \quad (3)$$

is the $M \times 1$ array manifold at wavenumber k_j . The variable i denotes the imaginary unit, i.e., $i = \sqrt{-1}$.

The correlation matrix $E[\mathbf{X}(k_j)\mathbf{X}(k_j)^H]$ of the received data at a given wavenumber k_j is estimated as:

$$\mathbf{R}_j = \frac{1}{Q} \sum_{q=1}^Q \mathbf{X}^{(q)}(k_j) \mathbf{X}^{(q)}(k_j)^H \quad (4)$$

where Q is the number of independent measurements and $\mathbf{X}^{(q)}(k_j) = [X_1^{(q)}(k_j) X_2^{(q)}(k_j) \dots X_M^{(q)}(k_j)]$ is the discrete Fourier transform of the q th measurement. Superscript $(\cdot)^H$ denotes the conjugate transpose.

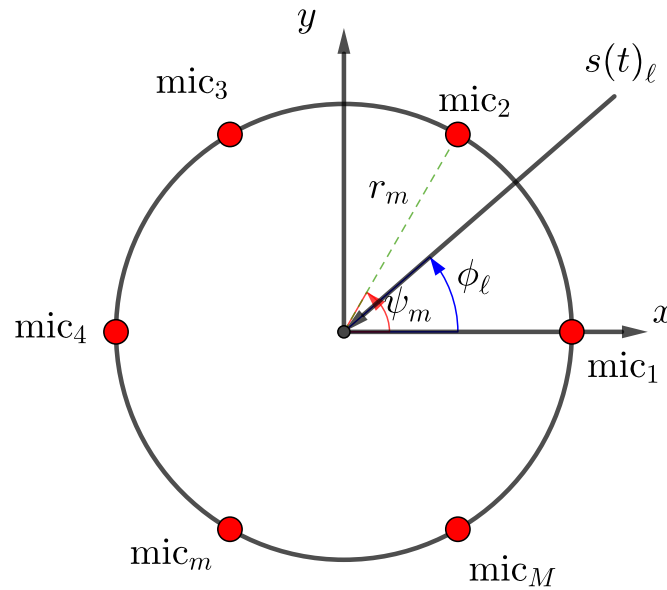


Figure 1. The uniform circular array of omnidirectional microphones with an incident source impinging at azimuth angle ϕ_ℓ .

2.2. The Array Manifold Interpolation Method

The correlation matrix of the observation has been developed for narrowband frequency f_j . For wideband sources, the various narrowbands are coherently summed into one correlation matrix by multiplying each narrowband correlation matrix by a focusing matrix. One method of achieving this is the array manifold interpolation method (AMI) [23]. This method is introduced in this section for an arbitrary array geometry and then further simplified for the uniform circular array in the next section.

The m th element of the array manifold of the planar array is represented as a series expansion of a plane wave, i.e., the Jacobi–Anger expansion as:

$$\begin{aligned} [\mathbf{a}(\phi, k)]_m &= e^{ikr_m \cos(\phi - \psi_m)} \\ &= \sum_{n=-\infty}^{\infty} i^n J_n(kr_m) e^{-in\psi_m} e^{in\phi} \\ &= \tilde{\mathbf{b}}_m(k) \tilde{\mathbf{w}}(\phi) \end{aligned} \quad (5)$$

where $[\tilde{\mathbf{b}}_m(k)]_n := i^n J_n(kr_m) e^{-in\psi_m}$ is n th element of the row vector $\tilde{\mathbf{b}}_m(k)$, and $[\tilde{\mathbf{w}}(k)]_n := e^{in\phi}$ is the n th element the column vector $\tilde{\mathbf{w}}(\phi)$, for $n = -\infty, \dots, 0, \dots, \infty$. The function $J_n(\cdot)$ is the Bessel function of the first kind with order n , r_m is the distance between the origin (usually the center of moment of the geometry) and the m th microphone, and other variables are as previously defined.

The Bessel function decays exponentially as $n \gg kr_m$; hence, the infinite summation in (6) can be truncated to a given N such that $|J_n(k_{\max} r_{m, \max})| < \epsilon$ for $n > N$, where ϵ is an acceptable truncation error. This implies that the length of vectors $\tilde{\mathbf{b}}_m(k)$ and $\tilde{\mathbf{w}}(\phi)$ is limited to $2N + 1$, thus reducing to $\mathbf{b}_m(k)$ and $\mathbf{w}(\phi)$, respectively. It is practical to set $N \geq 2k_{\max} r_{m, \max}$. Generally, the array manifold of the array with M total number of microphones is given by:

$$\mathbf{a}(\phi, k) = \mathbf{B}(k) \mathbf{w}(\phi) \quad (6)$$

where $[\mathbf{B}(k)]_{m,n} = i^n J_n(kr_m) e^{-in\psi_m}$, $m = 1, 2, \dots, M$, and $n = -N, -(N-1), \dots, 0, \dots, N-1, N$ and $\mathbf{w}(\phi)$ is as previously defined but for $n = -N, -(N-1), \dots, 0, \dots, N-1, N$.

For wideband sources, the concept of frequency smoothing is adopted to coherently sum the component narrowband covariance matrices of the wideband source into one nar-

rowband covariance matrix. This is followed by the smoothing of the mapped narrowband covariance matrix. For this, we define an $M \times M$ focusing matrix $\mathbf{T}(k_j)$ such that:

$$\mathbf{T}(k_j)\mathbf{a}(\phi, k_j) = \mathbf{a}(\phi, k_0), \quad \forall \phi. \quad (7)$$

According to (6) and (7):

$$\mathbf{T}(k_j) = \mathbf{B}(k_0)\mathbf{B}(k_j)^\# \quad (8)$$

where $\mathbf{B}(k_j)^\# = (\mathbf{B}^H\mathbf{B})^{-1}\mathbf{B}^H$ is the generalized inverse of matrix $\mathbf{B}(k_j)$ and $k_0 \in (k_{\min}, k_{\max})$ is the central processing wave number. It becomes necessary to choose N and M such that $M = 2N + 1$ to ensure $\mathbf{B}(k_j)$ is nonsingular [18,23].

The focused and frequency-smoothed covariance matrix is evaluated as:

$$\tilde{\mathbf{R}} = \sum_{j=1}^J \mathbf{T}(k_j)\mathbf{R}_j\mathbf{T}(k_j)^H \quad (9)$$

where J is the total number of selected frequency samples as in (2) and $\mathbf{R}_j$ is as previously defined in (4). In order to guarantee an improved performance, the frequency samples in the signal bandwidth that corresponds to the J highest DFT magnitudes are selected to avoid low-SNR frequency samples.

3. The Proposed Method

3.1. For the Uniform Circular Array

For the uniform circular array, $r_m = r \quad \forall m$ and $\psi_m = 2\pi m/M$. Thus, the matrix $\mathbf{B}(k)$ as defined in (6) can be expressed as:

$$\mathbf{B}(k) = \sqrt{M}\mathbf{F}\mathbf{J}(k) \quad (10)$$

where $\mathbf{F}$ is a $M \times (2N + 1)$ unitary DFT matrix defined as:

$$[\mathbf{F}]_{m,n} = \frac{1}{\sqrt{M}} e^{-i\frac{2\pi}{M}mn}, \quad \begin{matrix} m = 0, 1, \dots, M-1 \\ n = -N, \dots, N \end{matrix} \quad (11)$$

and $\mathbf{J}(k) \in \mathbb{C}^{(2N+1) \times (2N+1)}$ is a diagonal matrix with entries:

$$[\mathbf{J}(k)]_{n,n} = i^n J_n(kr). \quad (12)$$

By setting $M = 2N + 1$, $\mathbf{F}$ becomes an invertible square matrix. Thus, the focusing matrix is expressed as:

$$\begin{aligned} \mathbf{T}(k_j) &= \mathbf{F}\mathbf{D}(k_j)\mathbf{F}^{-1} \\ &= \mathbf{F}\mathbf{D}(k_j)\mathbf{F}^H \end{aligned} \quad (13)$$

since $\mathbf{F}$ is a unitary matrix. Matrix $\mathbf{D}(k)$ is a $(2N + 1) \times (2N + 1)$ diagonal matrix with diagonal entries:

$$[\mathbf{D}(k_j)]_{n,n} = \frac{J_n(k_0 r)}{J_n(k_j r)}. \quad (14)$$

3.2. The Proposed Focusing Matrix

In this section, the original array manifold interpolation method has been introduced and further simplified for the uniform circular array. For an embedded implementation of such an algorithm, a high-end microcontroller capable of implementing the Bessel function evaluation will be required. This requirement translates to higher cost for some low-budget

applications. To eliminate the complexity required for the calculation of the Bessel function, the proposed algorithm exploits a property of the Bessel function, which allowed for the total elimination of its computation.

According to the asymptotic property of the Bessel function, for orders n far greater than the argument, the Bessel function can be approximated as $J_n(kr) \approx \tilde{J}_n(kr)$, where:

$$\tilde{J}_n(kr) := \begin{cases} \frac{1}{\Gamma(n+1)} \left(\frac{kr}{2}\right)^n & \text{for } n \geq 0 \\ \frac{(-1)^n}{(-n)!} \left(\frac{2}{kr}\right)^n & \text{for } n < 0 \end{cases} \quad (15)$$

given that $0 < kr < \sqrt{n+1}$, $\forall n > 0$, assuming the least order of $n = 0$ [27] (p. 364). This small-argument approximation makes it possible to further simplify (14) to:

$$\begin{aligned} [\tilde{\mathbf{D}}(k_j)]_{n,n} &= \left(\frac{k_0}{k_j}\right)^n \\ &= \left(\frac{f_0}{f_j}\right)^n \end{aligned} \quad (16)$$

for $n = -N, \dots, 0, \dots, N$, given that $0 < kr < 1$. This condition is obtainable for a combination of source frequency and small radius uniform circular array. The diagonal matrix $\mathbf{D}(k_j)$ simply reduces to a matrix $\tilde{\mathbf{D}}(k_j)$ of n th power of the ratio of the central processing frequency f_0 and the narrow band frequency f_j , and thus, eliminating the computation of the Bessel function of the first kind, and the focusing matrix in (13) then becomes:

$$\tilde{\mathbf{T}}(k_j) = \mathbf{F} \tilde{\mathbf{D}}(k_j) \mathbf{F}^H \quad (17)$$

Finally, the steered covariance matrix is calculated using (9). Though this method adds additional matrix multiplication compared to the CSS method [19], it eliminates the initial estimation of the directions of arrival, which includes several matrix multiplications depending on the algorithm used. This focusing matrix (compared to other wideband methods) could be useful in tracking moving sound source as one focusing matrix is needed if the frequency spectrum of the source remains fairly constant. Furthermore, the proposed focusing matrix also does not depend on the radius of the uniform circular array, making it suitable for any circular array of any radius given that $0 < kr < 1$ is fulfilled.

3.3. Approximation Error

The proposed algorithm is based on the small-argument approximation of the Bessel function of the first kind as given in (15). However, the Bessel function occurs in ratio as in (14) for the uniform circular array. In this section, the resultant approximation error on matrix $\mathbf{D}(k_j)$ is analyzed. Towards this, the approximated Bessel function is equal to the true value plus an approximation error. That is:

$$\tilde{J}_n(kr) = J_n(kr) \pm \epsilon, \quad \epsilon > 0 \quad (18)$$

where ϵ is the approximation error and expressed as:

$$\epsilon = \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m! \Gamma(m+n+1)} \left(\frac{kr}{2}\right)^{2m+n}. \quad (19)$$

For the $\pm$ sign in (18), the ‘+’ sign applies for $n \geq 0$. When $n < 0$, ‘+’ sign applies for even n , and ‘−’ sign applies for odd n , respectively. The ratio in (16) becomes:

$$[\tilde{\mathbf{D}}(k_j)]_{n,n} = \frac{J_n(k_0 r) \pm \epsilon_0}{J_n(k_j r) \pm \epsilon_j} \quad (20)$$

and the approximation error $\Delta \mathbf{D}$ of matrix $\mathbf{D}(k_j)$ is thus:

$$\begin{aligned} \Delta \mathbf{D} &= [\mathbf{D}(k_j)_{n,n} - \tilde{\mathbf{D}}(k_j)]_{n,n} \\ &= \frac{\pm \epsilon_j J_n(k_0) \pm \epsilon_0 J_n(k_j)}{J_n(k_j)[J_n(k_j) \pm \epsilon_j]} \end{aligned} \quad (21)$$

which shows that the approximation error increases as $|k_0 - k_j|$ increases. This implies that choosing k_0 as the median of the selected high-SNR J bins will reduce the approximation error. As shown in Figure 2, a plot of $|\Delta \mathbf{D}|$ versus n and $k_j r \in [0.2, 1]$ and $k_0 r = 0.7$, which shows that the least error occur about $k_j r = 0.7$, where $|k_0 - k_j| = 0$.

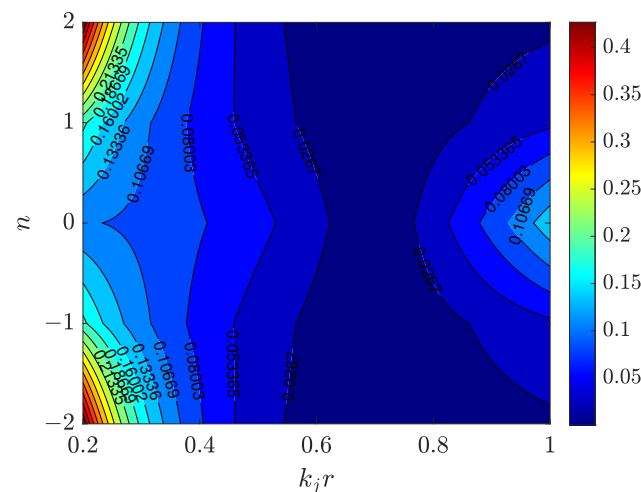


Figure 2. Plot of approximation error $\Delta \mathbf{D}$ versus order n and $k_j r \in [0.2, 1]$ for $k_0 r = 0.7$, showing the approximation error increases as the separation distance between processing frequency and other bins increases.

3.4. Computational Complexity

The computational complexity of the proposed method is compared against those of iMUSIC, WS-TOPS, AMI, and proposed method. Recall that M is the number of sensors, J is the number of narrow band frequency bins ($J' < J$ is the total number of selected processing frequency bins for the WS-TOPS method), Q is the number of snapshots, and L is the number of sources. Table 1 shows the matrix multiplication complexities, while Table 2 shows the matrix decomposition complexities for the methods under comparison.

Table 1. Matrix multiplication complexity comparison.

Method	Matrix Multiplication
iMUSIC	$O(J(M-L)(2M+1) + JQM^2)$
WS-TOPS	$O(J'(J-J')[2LM(M-L) + L^2M + 2LM^2] + (J-J')QM^2)$
AMI	$O((M-L)(2M+1) + JM^2(Q+4M))$
Proposed Method	$O((M-L)(2M+1) + JM^2(Q+2M))$

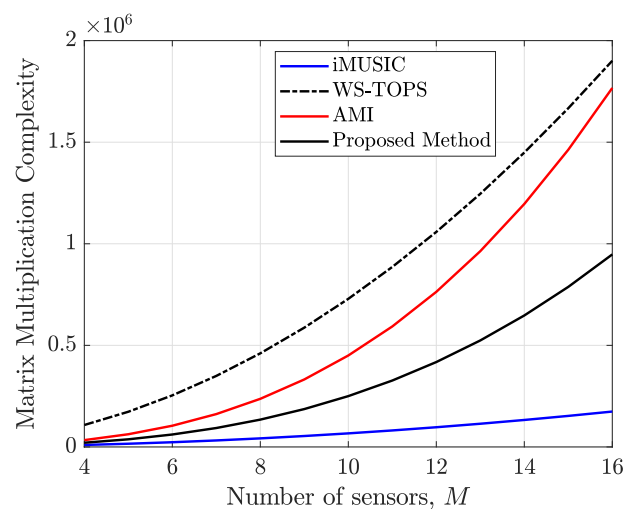
Table 2. Matrix decomposition complexity comparison.

Method	Eigen Decomposition	Singular-Value Decomposition
iMUSIC	$O(JM^3)$	<i>not required</i>
WS-TOPS	$O(JM^3)$	$O(L^3(J-1)J')$
AMI	$O(M^3)$	<i>not required</i>
Proposed Method	$O(M^3)$	<i>not required</i>

The computational complexity of computing the Bessel function of the first kind $J_n(kr)$ with accuracy 2^{-v} at the rational point is $O(v \log^3 vs. \log \log v)$ [28]. The evaluation of k^n (a ζ -digit number) requires $O(\log n M(\zeta))$, where $M(\zeta)$ is the time complexity of the multiplication algorithm. For the Harvey–Hoeven algorithm, $M(\zeta) = O(n \log n)$, which results to $O(n \log^2 n)$ [29]. The computational complexity in the proposed methods is further reduced by defining the focusing matrix $\mathbf{T}(k_j)$ as:

$$[\mathbf{T}(k_j)]_{m_r, m_c} = \sum_{n=-N}^N \left(\frac{f_0}{f_j} \right)^n W^{n(m_r - m_c)} \quad (22)$$

where $m_r, m_c = 1, 2, \dots, M$ and $W := e^{i\frac{2\pi}{M}}$. This eliminates the $O(2JM^3)$ complexity for the matrix multiplication in (13) for $\mathbf{T}(k_j)$. Plots of the matrix multiplication complexity versus number of sensors and versus number of narrowband frequency bins are shown in Figure 3a and 3b, respectively.

**(a)** Complexity versus M **Figure 3.** Cont.

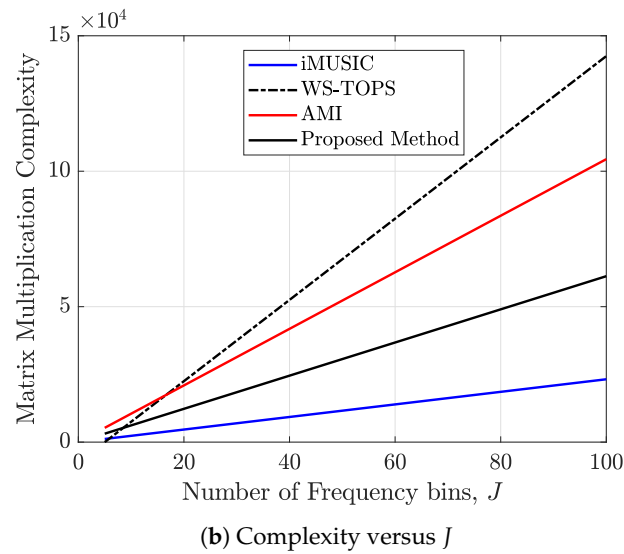


Figure 3. Plot of the matrix multiplication complexity versus (a) number of sensors M for $J = 100$ bins, $L = 2$, $Q = 5$, and $J' = 0.1J$, (b) number of frequency bins J for $M = 6$, $Q = 5$, $L = 2$.

It is observed that the matrix multiplication complexity involved in the iMUSIC is always lower. However, the matrix multiplication complexity of the proposed method is always less than the original AMI and the WS-TOPS.

3.5. Direction of Arrival Estimation

For the uniform circular array, a new focusing matrix for coherently summing the narrowband covariance matrices has been proposed in Section 3. Using the steered covariance matrix, any narrowband subspace direction of arrival method or maximum likelihood estimator can be adopted. In this paper, the MUSIC [7] algorithm is applied as the narrowband subspace method to the focused covariance matrix. The direction of arrival is, thus, the direction corresponding to the peaks of the beampattern:

$$Y(\phi) = \left[\mathbf{a}(\phi, f_0)^H \mathbf{U}_N \mathbf{U}_N^H \mathbf{a}(\phi, f_0) \right]^{-1}, \quad (23)$$

where $\mathbf{U}_N$ is the noise subspace of $\tilde{\mathbf{R}}$ for proposed method. For higher SNR, the number of sources can easily be obtained by detecting a sharp decrease in the ordered eigenvalues of $\tilde{\mathbf{R}}$. Another alternative is to assume that the number of sources is equal to the number of sensors minus 1, i.e., $M - 1$, since that is the highest number of sources the array can detect. Furthermore, for the MUSIC algorithm, overestimating is not as detrimental as underestimating the number of sources. The estimation can then be corrected by iteration, where the number of prominent peaks given by the initial assumption is adopted as the number of sources until a finer beampattern is obtained. Other algorithms for estimating the number of source can be adopted [30,31]. The summary of the proposed algorithm is captured in the block diagram in Figure 4.

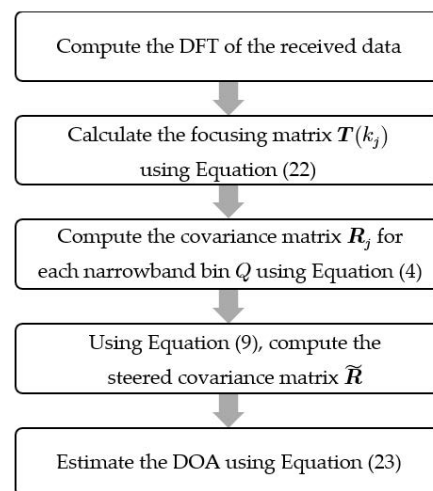


Figure 4. Block diagram showing the summary of the proposed method.

4. Monte Carlo Simulations to Test the Efficacy of the Proposed Method

The efficacy of the proposed algorithm is evaluated in this section. For this purpose, the performance of the proposed method is compared against that of the iMUSIC, WS-TOPS, and the Original AMI. The choice of these methods for comparison is mainly because they do not require initial estimation of the direction of arrival.

4.1. Root-Mean-Square Performance

The incident signal is modeled as band-limited real-valued zero-mean Gaussian random variables sampled at a sampling frequency of 12 kHz at various signal-to-noise ratios and 5 snapshots of 8192 times samples each. This chosen frequency band guarantees that the maximum kr is less than 1. The array is a uniform circular array consisting of $M = 6$ omnidirectional microphones with radius $r = 4.5$ cm. For the signal bandwidth ($f \in [400, 1213]$ Hz), 100 narrowband frequencies (obtained using fast Fourier transform) were used. The root-mean-square error of the proposed algorithm is compared against that of iMUSIC, WS-TOPS, and original AMI method for a single source located at 120° with different SNR ranging from -10 dB to 40 dB. Figure 5 shows the plot of the root mean square estimation error versus the SNR for the methods under comparison. Each point in Figures 5 and 6 represents the root-mean-square error value of 100 iterations. The proposed method slightly outperforms the original AMI method for low-frequency wideband sources.

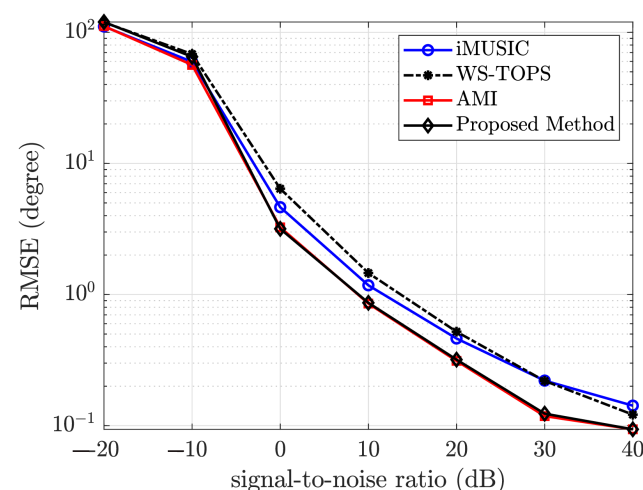


Figure 5. Plot root-mean-squared error versus signal-to-noise ratio for the iMUSIC, WS-TOPS, Original AMI, and proposed algorithm using simulated data.

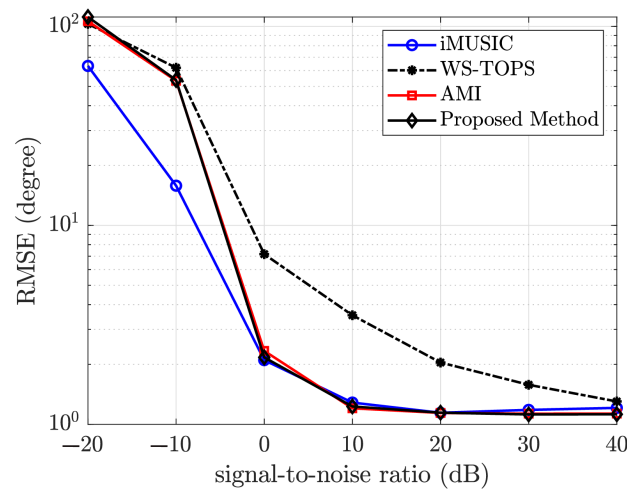


Figure 6. Plot root-mean-squared error versus signal-to-noise ratio for the iMUSIC, WS-TOPS, Original AMI, and proposed algorithm using simulated data of $f \in [166, 500]$ Hz.

To ascertain the relative performance of the proposed method for low frequency signals, the bandwidth of the signal is limited to $f \in [166, 500]$ Hz. The root-mean-squared error versus SNR is shown in Figure 6. This shows that the performance of the proposed method is consistent with the original AMI. Furthermore, both outperform the WS-TOPS for low-frequency wideband sources.

4.2. Computational Efficiency of the Proposed Algorithm

The difference in the computational complexity of the proposed algorithm and original AMI method has been discussed in Section 3.4. The computational time interpretation of this improvement is studied in this section. The proposed algorithm saves processing time by cutting the time taken to generate the diagonal matrix $\mathbf{D}(k_j)$ by up to 30% on average using MATLAB R2020b on a Core i7 computer. The reduction in computation time is expected to increase when used on lower processors and in embedded microcontroller systems. This is expected as the elimination of the Bessel function for each order n reduces the number of calculations by the processor.

5. Experimental Evaluation

The MiniDSP UMA-8 mic array was used for the experiment. As shown in Figure 7, it is a circular board of 4.5 cm radius, which has seven onboard omnidirectional MEMS microphones in total: six uniformly placed around the circumference and one at the center [32]. For the purpose of this experiment, only the six circumferential microphones were used to match the uniform circular array described for this study. The incident sound source is white noise located at an azimuth angle $\phi = 120^\circ$, recorded at a sampling rate of 48 kHz, playing through a loudspeaker.

The experiments demonstrate how the proposed algorithm performs across signal-to-noise ratio, various wideband frequency ranges (i.e., $kr \in [kr_{min}, kr_{max}]$), and various broadband sizes. Noise is added at varying SNR from -20 dB to 40 dB with a step of 10 dB. $f \in [200, 600]$ is the filtering range for the experiments.

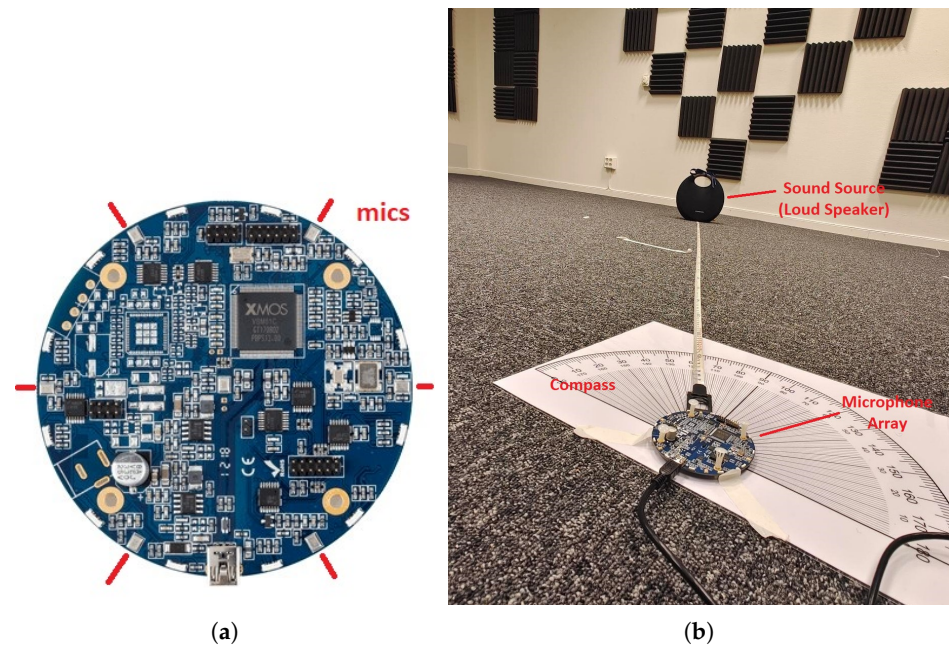
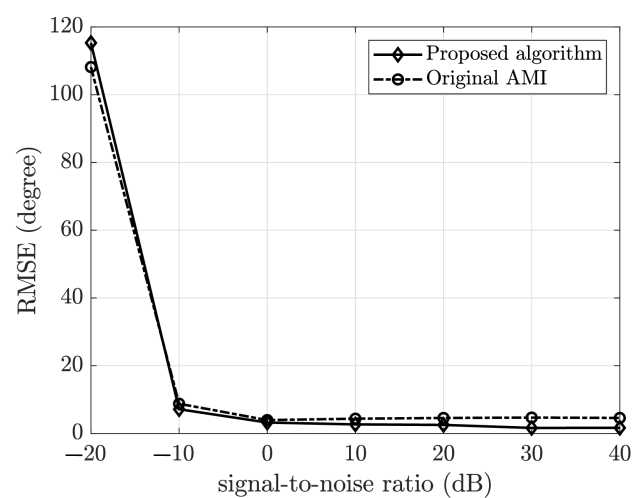


Figure 7. The uniform circular array and the experimental setup. (a) The uniform circular array using a MiniDSP UMA-8 system [32] and (b) the experimental setup and environment, using sounds from a Harman Kardon Onyx 5 wireless loudspeaker, while measuring distance and angle from the loudspeaker to the array system.

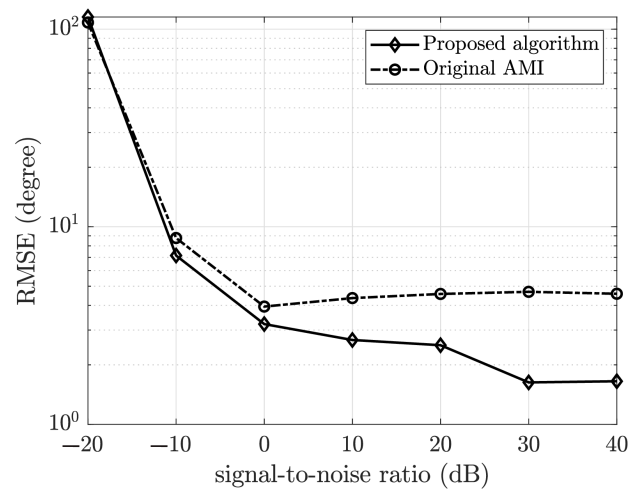
5.1. Root-Mean-Square Error Performance versus Signal-to-Noise Ratio

Additive white Gaussian noise of various SNR was added to the recorded signal and the direction of arrival estimated. For each value of SNR, 100 independent estimations were carried out and the root-mean-square error (RMSE) was taken. Figure 8 shows the plot of the RMSE versus SNR for both the proposed and the original AMI method, $N = 2$, and the frequency range was divided into 30 bins. As shown in Figure 8, the proposed algorithm outperformed the original AMI method.



(a) Root-mean-square error in degree.

Figure 8. Cont.



(b) Root-mean-square error in degree plot in logarithm.

Figure 8. Plot of root-mean-squared error versus signal-to-noise ratio for the proposed algorithm and original AMI method using recorded data.

5.2. Performance by Central Processing Frequency

In this section, the performance of the proposed algorithm across kr is studied. To ensure every other parameter is the same but the kr , the magnitude of the frequency bins is shifted across the frequency. By doing so, only the $kr_{\min}$, k_0r , and $kr_{\max}$ vary. For the recorded signal used, the DFT magnitudes for $f \in [50, 460]$ (the dominant bins) were shifted in steps of 100 Hz. The central frequency for each case is the median of the selected bins. The RMSE performance versus the central processing frequency f_0 (or k_0r) is plotted in Figure 9, which shows that the proposed algorithm performs similarly to the original AMI.

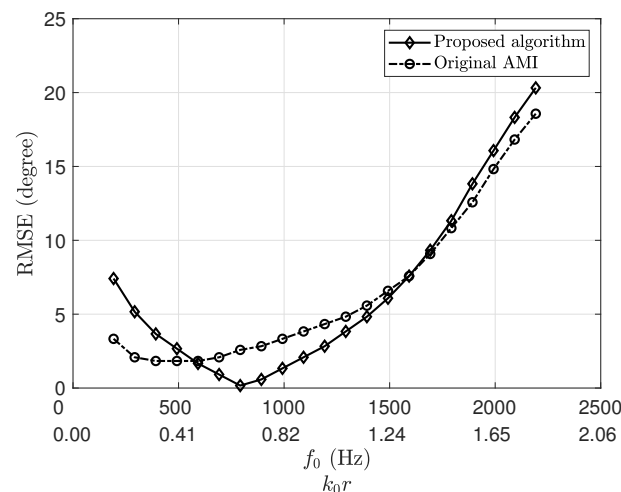


Figure 9. Plot of root-mean-squared error versus central processing frequency f_0 (k_0r) for the proposed algorithm and original AMI method using recorded data.

5.3. Performance for Different Types of Sources

The proposed algorithm is further tested for various sources. For the experiment, the following sound sources were used: car engine, water fall, hovering drone, electric fan, and train engine. For each sound, the source is moved to four locations, and the direction of arrival estimated using the original array manifold interpolation and the proposed modified method. The average estimation errors for each source type are reported in Table 3 where the bold entries highlights the minimum value in each row.

Table 3. Average estimation error for different source types.

Source	Original AMI (°)	Proposed Algorithm (°)
Car engine	2.9969	2.4972
water fall	0.6235	0.4362
Hovering drone	1.8102	1.3730
Electric fan	2.3723	0.6860
Train engine	1.0628	1.2502

The performance of the proposed algorithm is shown to outperform the original AMI for all the sound types but the train engine, as shown in Table 3. This goes to show that the proposed method performs similarly to the original AMI, while offering a significant reduction in computational complexity.

6. Conclusions

For the uniform circular array, the steered covariance matrix method based on array manifold interpolation is more computationally efficient compared to other coherent covariance matrix averaging. This is due to the fact that the steering matrix in the array manifold interpolation method is independent of the source direction of arrival. It has been shown in this paper that by applying the small-argument approximation of the Bessel function, the n th power of the ratio of the Bessel function of different orders are reduced to the n th power of the ratio of the processing frequency and the narrowband frequencies. This led to a significant reduction in the computation time for the diagonal matrix component of the focusing matrix, while achieving similar localization accuracy for low-frequency wideband sources compared to the original AMI method. This simplified algorithm eliminates the need for high-performance microprocessors for the implementation of real-time low-frequency wideband sound source localization on the embedded system. Furthermore, the proposed focusing matrix is independent of the array radius, making the focusing matrix reusable for the frequency bins if the signal spectrum remains fairly constant over time.

Author Contributions: Conceptualization, C.J.N. and M.S.; methodology, M.J. and C.J.N.; software, M.J.; validation, J.L., M.S. and G.T.; investigation, M.J.; resources, M.J.; writing—original draft preparation, M.J. and C.J.N.; writing—review and editing, M.J.; visualization, M.J. and C.J.N.; supervision, G.T. and M.S.; project administration, J.L.; funding acquisition, J.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research and APC were funded by the Swedish Knowledge Foundation, grant number 20200189.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors are grateful to the Sensible Things that Communicate Research Centre of the Mid Sweden University for providing administrative and technical support for this research.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Dehghan Firoozabadi, A.; Irarrazaval, P.; Adasme, P.; Zabala-Blanco, D.; Játiva, P.P.; Azurdia-Meza, C. 3D Multiple Sound Source Localization by Proposed T-Shaped Circular Distributed Microphone Arrays in Combination with GEVD and Adaptive GCC-PHAT/ML Algorithms. *Sensors* **2022**, *22*, 1011. [\[CrossRef\]](#)
- Zhang, C.; Florêncio, D.; Ba, D.E.; Zhang, Z. Maximum likelihood sound source localization and beamforming for directional microphone arrays in distributed meetings. *IEEE Trans. Multimed.* **2008**, *10*, 538–548. [\[CrossRef\]](#)
- Rascon, C.; Meza, I. Localization of sound sources in robotics: A review. *Robot. Auton. Syst.* **2017**, *96*, 184–210. [\[CrossRef\]](#)

4. Deleforge, A.; Di Carlo, D.; Strauss, M.; Serizel, R.; Marcenaro, L. Audio-based search and rescue with a drone: Highlights from the IEEE signal processing cup 2019 student competition [SP competitions]. *IEEE Signal Process. Mag.* **2019**, *36*, 138–144. [\[CrossRef\]](#)
5. Roseveare, N.J.; Azimi-Sadjadi, M.R. Robust beamforming algorithms for acoustic tracking of ground vehicles. In Proceedings of the Unattended Ground, Sea, and Air Sensor Technologies and Applications VIII, Kissimmee, FL, USA, 17–20 April 2006; Carapezza, E.M., Ed.; International Society for Optics and Photonics, SPIE: Bellingham, WA, USA, 2006; Volume 6231, pp. 51–62. [\[CrossRef\]](#)
6. Ribeiro, F.; Florencio, D.; Ba, D.; Zhang, C. Geometrically Constrained Room Modeling With Compact Microphone Arrays. *IEEE Trans. Audio Speech Lang. Process.* **2012**, *20*, 1449–1460. [\[CrossRef\]](#)
7. Schmidt, R.O. Multiple Emitter Location and Signal Parameter Estimation. *IEEE Trans. Antennas Propag.* **1986**, *AP-34*, 276–280. [\[CrossRef\]](#)
8. Vallet, P.; Mestre, X.; Loubaton, P. Performance Analysis of an Improved MUSIC DoA Estimator. *IEEE Trans. Signal Process.* **2015**, *63*, 6407–6422. [\[CrossRef\]](#)
9. Roy, R.; Kailath, T. ESPRIT—Estimation of Signal Parameters Via Rotational Invariance Techniques. *IEEE Trans. Acoust. Speech Signal Process.* **1989**, *37*, 984–995. [\[CrossRef\]](#)
10. Ferguson, B.G. Minimum variance distortionless response beamforming of acoustic array data. *J. Acoust. Soc. Am.* **1998**, *104*, 947–954. [\[CrossRef\]](#)
11. Kiong, T.S.; Salem, S.B.; Paw, J.K.S.; Sankar, K.P.; Darzi, S. Minimum Variance Distortionless Response Beamformer with Enhanced Nulling Level Control via Dynamic Mutated Artificial Immune System. *Sci. World J.* **2014**, *2014*, 164053. [\[CrossRef\]](#)
12. Knapp, C.; Carter, G. The generalized correlation method for estimation of time delay. *IEEE Trans. Acoust. Speech Signal Process.* **1976**, *24*, 320–327. [\[CrossRef\]](#)
13. Gombots, S.; Nowak, J. Sound source localization—State of the art and new inverse scheme. *Elektrotechnik Und Informationstechnik* **2021**, *138*, 229–243. [\[CrossRef\]](#)
14. Su, G.; Morf, M. The signal subspace approach for multiple wide-band emitter location. *IEEE Trans. Acoust. Speech Signal Process.* **1983**, *31*, 1502–1522. [\[CrossRef\]](#)
15. Yoon, Y.S.; Kaplan, L.; McClellan, J. TOPS: New DOA estimator for wideband signals. *IEEE Trans. Signal Process.* **2006**, *54*, 1977–1989. [\[CrossRef\]](#)
16. Yu, H.; Liu, J.; Huang, Z.; Zhou, Y.; Xu, X. A New Method for Wideband DOA Estimation. In Proceedings of the 2007 International Conference on Wireless Communications, Networking and Mobile Computing, Shanghai, China, 21–25 September 2007; pp. 598–601. [\[CrossRef\]](#)
17. Hayashi, H.; Ohtsuki, T. DOA estimation for wideband signals based on weighted Squared TOPS. *EURASIP J. Wirel. Commun. Netw.* **2016**, *2016*, 243. [\[CrossRef\]](#)
18. Pham, T.; Sadler, B. Adaptive wideband aeroacoustic array processing. In Proceedings of the 8th Workshop on Statistical Signal and Array Processing, Corfu, Greece, 24–26 June 1996; pp. 295–298. [\[CrossRef\]](#)
19. Wang, H.; Kaveh, M. Coherent signal-subspace processing for the detection and estimation of angles of arrival of multiple wide-band sources. *IEEE Trans. Acoust. Speech Signal Process.* **1985**, *33*, 823–831. [\[CrossRef\]](#)
20. Valaee, S.; Champagne, B.; Kabal, P. Localization of wideband signals using least-squares and total least-squares approaches. *IEEE Trans. Signal Process.* **1999**, *47*, 1213–1222. [\[CrossRef\]](#)
21. Wang, J.; Zhao, Y.; Wang, Z. Low complexity subspace fitting method for wideband signal location. In Proceedings of the 2008 5th IFIP International Conference on Wireless and Optical Communications Networks (WOCN'08), Surabaya, Indonesia, 5–7 May 2008; pp. 1–4. [\[CrossRef\]](#)
22. Xu, W.; Chen, B.; Hu, Y.; Li, J. A Novel Wide-Band Directional MUSIC Algorithm Using the Strength Proportion. *Sensors* **2023**, *23*, 4562. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Doron, M.; Doron, E.; Weiss, A. Coherent wide-band processing for arbitrary array geometry. *IEEE Trans. Signal Process.* **1993**, *41*, 414–417. [\[CrossRef\]](#)
24. Zhao, D.; Deng, Z.; Tan, W. Low-Complexity DOA Estimation for Uniform Circular Arrays with Directional Sensors Using Reconfigurable Steering Vectors. *Circuits Syst. Signal Process.* **2023**, *42*, 1685–1706. [\[CrossRef\]](#)
25. Anand, G.; Koul, A.; Gurugopinath, S.; Nagesha, P. Residual Data Vector Method of Underwater Acoustic Source Localization by a Three-Dimensional Array. In Proceedings of the OCEANS 2022-Chennai, Chennai, India, 21–24 February 2022; pp. 1–6.
26. Suksiri, B.; Fukumoto, M. A computationally efficient wideband direction-of-arrival estimation method for l-shaped microphone arrays. In Proceedings of the 2018 IEEE International Symposium on Circuits and Systems (ISCAS), Florence, Italy, 27–30 May 2018; pp. 1–5.
27. Abramowitz, M.; Stegun, I.A. *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables*; United States Department of Commerce, National Bureau of Standards: Washington, DC, USA, 1972.
28. Karatsuba, E.A. Calculation of Bessel Functions via the Summation of Series. *Numer. Anal. Appl.* **2019**, *12*, 372–387. [\[CrossRef\]](#)
29. Harvey, D.; van Der Hoeven, J. Integer multiplication in time $O(n \log n)$. *Ann. Math.* **2021**, *193*, 563–617. [\[CrossRef\]](#)
30. Salman, T.; Badawy, A.; Elfouly, T.M.; Mohamed, A.; Khattab, T. Estimating the number of sources: An efficient maximization approach. In Proceedings of the 2015 International Wireless Communications and Mobile Computing Conference (IWCMC), Dubrovnik, Croatia, 24–28 August 2015; pp. 199–204. [\[CrossRef\]](#)

31. Williams, D. Detection: Determining the Number of Sources. In *Digital Signal Processing Handbook*; Madisetti, V.K., Williams, D.B., Eds.; CRC Press LLC: Boca Raton, FL, USA, 1999.
32. MiniDSP. UMA-8 USB Mic Array—V2.0. Available online: <https://www.minidsp.com/products/usb-audio-interface/uma-8-microphone-array> (accessed on 24 October 2022).

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.