

Article

Bayesian Inference for Acoustic Direction of Arrival Analysis Using Spherical Harmonics

Ning Xiang *  and Christopher Landschoot 

Graduate Program in Architectural Acoustics, Rensselaer Polytechnic Institute, Troy, NY 12180, USA;
christopherlandschoot@gmail.com

* Correspondence: xiangn@rpi.edu

Received: 25 April 2019; Accepted: 7 June 2019; Published: 10 June 2019



Abstract: This work applies two levels of inference within a Bayesian framework to accomplish estimation of the directions of arrivals (DoAs) of sound sources. The sensing modality is a spherical microphone array based on spherical harmonics beamforming. When estimating the DoA, the acoustic signals may potentially contain one or multiple simultaneous sources. Using two levels of Bayesian inference, this work begins by estimating the correct number of sources via the higher level of inference, Bayesian model selection. It is followed by estimating the directional information of each source via the lower level of inference, Bayesian parameter estimation. This work formulates signal models using spherical harmonic beamforming that encodes the prior information on the sensor arrays in the form of analytical models with an unknown number of sound sources, and their locations. Available information on differences between the model and the sound signals as well as prior information on directions of arrivals are incorporated based on the principle of the maximum entropy. Two and three simultaneous sound sources have been experimentally tested without prior information on the number of sources. Bayesian inference provides unambiguous estimation on correct numbers of sources followed by the DoA estimations for each individual sound sources. This paper presents the Bayesian formulation, and analysis results to demonstrate the potential usefulness of the model-based Bayesian inference for complex acoustic environments with potentially multiple simultaneous sources.

Keywords: Bayesian inference; maximum entropy; spherical harmonics; direction of arrival; model selection; parameter estimation

1. Introduction

This paper offers a solution to the problem of localizing multiple simultaneous acoustic sources in acoustic environments through a model-based probabilistic approach. This research demonstrates that the number of sound sources as well as the directions in which they arrive can be estimated given a set of sound signals recorded on a spherical microphone array [1,2]. The estimation requires a process known as spherical beamforming, or the spatial filtering of a sound signal using spherical harmonics theory [3]. This is combined with probabilistic inference using Bayesian model selection and parameter estimation [4,5].

Estimation of direction of arrivals (DoAs) from multiple simultaneous sound sources in complex sound environments presents a challenge as there may be variations in the number of simultaneous sources, along with their locations, characteristics, and strengths [6–8]. Furthermore, there can be unwanted, fluctuating background noise as well. Some solutions to this problem have been reported [9]. However, to the best of the authors' knowledge, applying Bayesian inference to solve this problem has not yet been sufficiently explored.

Spherical harmonics have long been applied in mathematics and science fields [10–12]. In acoustics, interest in spherical harmonics can be traced back to Lord Raleigh [10]. In recent years interest has been growing greatly. One recent example includes Fourier acoustics [3], applied in acoustic near-field holographical investigations that employ spherical harmonics theory. Using microphone array technology [13], specifically spherical microphone arrays, an entire soundfield could be analyzed without directional constraints [1].

One aspect of localizing sound sources in complex sonic environments is performing the DoA analysis on the recorded signals. There exists a number of methods that have been developed to address this problem. Recent development of various different microphone arrays includes sparse linear microphone arrays [7,14] and a two-microphone array used in a room-acoustic study [8].

A spherical microphone array contains a certain number of microphone capsules arranged on a spherical surface that sense sound signals simultaneously. Because of the spherical arrangement of the microphone arrays, there are no inherent directional constraints, unlike classical line-arrays [7] or the two microphone array as reported by Escolano et al. [8]. The recorded signals can be processed to any orientation. Various methods of data processing have attempted to determine the best ways to process sound signals in order to analyze complex sound environments, such as spherical harmonic beamforming in combination with optimal array processing, frequency smoothing methods [15], and modal smoothing methods [16]. Nadiri and Rafaely [17] localize multiple speakers under reverberant environment while Sun et al. [18] apply a spherical microphone array to localize reflections in rooms. On table tops in conference room applications, or mounted in the ceiling, hemispherical microphone arrays [19] have been proven to be more suitable. These methods implemented with spherical microphone arrays improve the ability to determine the DoAs of sound sources. Many methods, however, still rely on the basic approach to localizing source locations through correlating them directly with high sound energy levels. A number of other recent investigations also exist using spherical harmonics either for sound radiation [20], sound field reconstruction [21], estimation of oblique incident surface impedance [22], in noise analysis [23], and capturing sound intensities, [24]. Rules-of-thumb [25] are also discussed on how the estimation precision for an incident source's azimuth-polar DoA depends on the number of identical isotropic sensors. A solution [26] has been suggested to avoid ill-conditioned singularity when solving least-squares and eigenvalue problems to estimate the DoAs.

As for complex sound source analysis, the spatial resolution to determine discrete source locations has not been well investigated unless the sources are clearly separated in space due to the fact that limited order spherical arrays have a limited spatial resolution. With a large number of sound sources, the signals to be analyzed can blend together if the sound sources are located too closely to each other in physical space. The sound signals recorded by the limited order of spherical arrays often carry insufficient information. This situation requires higher order spherical arrays to be used to accurately determine sound sources, which in turn requires more microphone channels.

In addition to the parameter estimation problems [27], which are solely associated with the DoA estimation given a known number of sound sources, there is a need to answer the overarching question of how to reliably determine the number of sound sources present. The answer resides in model-based Bayesian inference, which represents a probabilistic method that can estimate the number of sources and their attributes through probabilistic analysis rather than just correlating high sound energy levels to sound source locations. This work applies Bayesian model selection to the DoA estimation tasks when the number of sound sources is unknown prior to the analysis. This Bayesian formulation for model selection starts with the application of Bayes' theorem, followed by the incorporation of prior information. Any interest in directional parameter values will be deferred into the background of the model selection problem. This allows attention to be focused on estimating the probabilities for the number of simultaneous sound sources.

There have been many recent efforts to apply Bayesian model selection to other acoustics problems. Xiang et al. [28] apply Bayesian model selection to determine the number of exponential decays present

in acoustic enclosures by analyzing sound energy decay functions. Bayesian model selection has also been applied to room-acoustic modal analysis [29]. Previous studies [7,8,14] of DoA analysis have also employed two-levels of Bayesian inference. However, to the best of the authors' knowledge, the model-based Bayesian analysis has not yet been sufficiently investigated using spherical microphone arrays. This paper demonstrates that the model-based Bayesian probabilistic approach can be applied to spatial sound field analysis with a set of sound signals recorded on a spherical microphone array. This approach estimates the number of sound sources as well as the directions of arrivals via two levels of Bayesian inference.

The remainder of this paper is as follows, Section 2 formulates the spherical harmonic beamforming for the data processing and models of potential multiple sound sources. Section 3 introduces two levels of Bayesian inference framework. Section 4 discusses maximum entropy assignment of prior probabilities. Section 5 explains a Markov Chain Monte Carlo method dedicated to the Bayesian analysis, nested sampling. Section 6 discusses experimental results using the two levels of Bayesian inference. Section 7 further discusses estimation errors and angular resolution issues. Section 8 concludes the paper.

2. Data and Models in Spherical Harmonics

Spherical harmonics theory plays a central role in the DoA analysis using a spherical microphone array. It is used to process recorded sound signals to obtain sound energy distributions around the spherical microphone array.

2.1. Spherical Harmonics

Spherical harmonics are eigen-functions of the wave equation in spherical coordinates [30]. Any sound field is composed of a series of orthogonal spherical harmonics of different orders. Taking the spherical wave equation in Helmholtz form with a solution

$$\underline{p}(r, \Phi) = \underline{R}(r) \underline{Y}(\Phi), \quad (1)$$

this equation separates the solution in radial (r), and angular $\underline{Y}(\Phi)$ components, respectively. Throughout the following discussions, a time-dependence $e^{j\omega t}$ is implicitly assumed, $j = \sqrt{-1}$ with ω being angular frequency, $k = \omega/c$ is the propagation coefficient, and c is sound speed. Also an under bar of all the variables, e.g., \underline{Y} , explicitly indicates complex-valued functions. Angular variable, $\Phi = (\theta, \varphi)$, includes elevation and azimuth angles.

The solution of the angular components in Equation (1) can be particularly specified by $\underline{Y}(\Phi) = \underline{Y}_n^m(\Phi)$, combined with specific amplitude values [3],

$$\underline{Y}_n^m(\Phi) = \sqrt{\frac{2n+1}{4\pi} \frac{(n-m)!}{(n+m)!}} P_n^m(\cos \theta) e^{jm\varphi}, \quad (2)$$

where $P_n^m(\cdot)$ is the associated Legendre function. $\underline{Y}_n^m(\Phi)$ is termed *spherical harmonics*, integer $n = 0, 1, 2, \dots$, is termed order, and $m = 0, \pm 1, \pm 2, \dots$, termed degree (mode) of the spherical harmonics, respectively. Figure 1 illustrates real-part of the spherical harmonics up to order 3 for all degrees. The spherical harmonics with the specified amplitude values are orthonormal,

$$\sum_{n=0}^{\infty} \sum_{m=-n}^n [\underline{Y}_n^m(\Phi')]^* \underline{Y}_n^m(\Phi) = \delta_{n'n} \delta_{m'm}, \quad (3)$$

where the symbol * denotes complex conjugate. The spherical harmonics form a set of orthonormal basis functions, used as the expansion for any arbitrary function, $f(\Phi)$, on the surface of a sphere,

$$f(\Phi) = \sum_{n=0}^{\infty} \sum_{m=-n}^n f_{nm} Y_n^m(\Phi), \tag{4}$$

with appropriate weights f_{nm} . These weights form the spherical Fourier transform of $f(\Phi)$ [3].

The solution of the radial component in Equation (1) can be expressed

$$\underline{R}(r) = \underline{R}_1 j_n(kr) + \underline{R}_2 y_n(kr), \tag{5}$$

where $\underline{R}_1, \underline{R}_2$ are constants, $j_n(v)$ and $y_n(v)$ are spherical Bessel functions [3,11]. The radial solution can also be expressed

$$\underline{R}(r) = \underline{R}_3 h_n(kr) + \underline{R}_4 h_n^*(kr), \tag{6}$$

with $\underline{R}_3, \underline{R}_4$ being constants, $h_n(\beta r)$ is the spherical Hankel function

$$h_n(kr) = j_n(kr) + j y_n(kr). \tag{7}$$

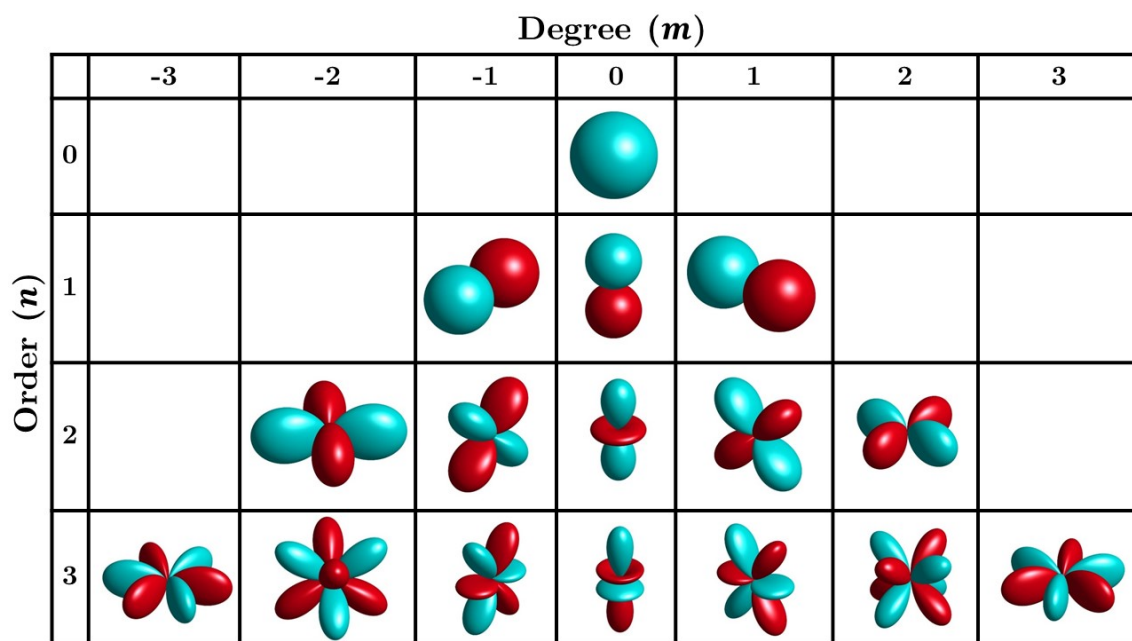


Figure 1. Real parts of the spherical harmonics up to third order ($n = 0, 1, 2, 3$), for degree between $-3 \leq m \leq 3$, with lobes in light (cyan-) color indicating positive values and lobes in dark (red) color indicating negative values. For each given order n , each row in the table contains $2n + 1$ modes.

2.2. Spherical Array Data Processing

This work applies the basic principle of spherical harmonics as briefly stated above to process spherical microphone signals. The process facilitates the analysis of incoming soundfield and predicts the sound energy around the spherical microphone array. The processing of M microphones flush-mounted on the rigid spherical surface is expressed as

$$p_{nm}(k, a) = \frac{4\pi}{M} \sum_{i=1}^M p_{mic}(k, a, \Phi_i) [Y_n^m(\Phi_i)]^*, \tag{8}$$

where a is the radius of the spherical array, $p_{nm}(k, a)$ are the spherical harmonic weights [2], $p_{mic}(k, a, \Phi_i)$ represents i -th microphone output among M microphone channels around a rigid sphere surface at angular position, Φ_i . The order, n , provides an increase in angular resolution. The number of microphones, M , sampling the sphere determines the maximum order of spherical harmonics, N , that the spherical array can achieve with $(N + 1)^2 \geq M$ [31]. To achieve high orders of spherical harmonics, more microphone channels are required to sample the spherical array, e.g., higher resolution. The spherical harmonic weights, $p_{nm}(k, a)$, are further transformed into the spherical harmonic domain [1]

$$y(\Phi) = \sum_{n=0}^N \sum_{m=-n}^n w_{nm}^*(k a, \Phi) p_{nm}(k, a), \tag{9}$$

where now weights $w_{nm}^*(k a, \Phi)$ combine both the radial and angular solutions stated before in the following way

$$w_{nm}^*(k a, \Phi) = \frac{Y_n^m(\Phi)}{b_n(k a)}, \tag{10}$$

for axis-symmetric beamforming in the plane-wave decomposition mode [2]. Variable $b_n(k a)$ represents spherical modal amplitude for a rigid sphere with radius, a ,

$$b_n(k a) = 4 \pi j^n \left[j_n(k a) - \frac{j_n'(k a)}{h_n^{(2)'}(k a)} h_n^{(2)}(k a) \right], \tag{11}$$

with $j^n = (-1)^{n/2}$, and $j_n(k a)$ being spherical Bessel function of the first kind and $h_n^{(2)}(k a)$ being spherical Hankel function of the second kind [2]. The prime denotes the derivative with respect to the argument. Note that this work focuses on the application under the plane-wave condition, that requires sound sources find themselves in the far-field. Though beamforming can inherently be formulated in regards to the near-field condition as well, yet the current application requires a far-field condition. In that case, the spherical modal amplitude in Equation (11) needs to be formulated accordingly [2]. It is however, beyond the scope of the current work.

From Equation (9), the directional pattern around the spherical microphone array is then formulated in their normalized absolute energy values

$$D(\Phi) = \frac{|y(\Phi)|^2}{\max [|y(\Phi)|^2]}. \tag{12}$$

The normalized directional pattern is taken as ‘experimental’ data for the Bayesian inference in the following discussion. They are denoted in vector/matrix form as $\mathbf{D} = [D(\Phi)]$.

2.3. Analytical Beamforming Models

The orthonormal property of the spherical harmonics in Equation (3) essentially expresses spherical processing that predicts spatial filter capability with respect to a sphere. The spatial filter direction is referred to as a beam. The directional beam pattern with a finite spherical harmonic order is expressed by the truncated orthonormality as

$$\underline{g}(\Phi_s, \Phi) = 2 \pi \sum_{n=1}^N \sum_{m=-n}^n \underline{Y}_n^m(\Phi_s)^* \underline{Y}_n^m(\Phi), \tag{13}$$

where $\Phi_s = \{\theta_s, \phi_s\}$ denotes the specific filtering direction, and $\underline{g}(\Phi_s, \Phi)$ represents specific beamforming function oriented towards direction, Φ_s , over angular range specified by Φ . The

maximum order, N , determines the sharpness of the beam patterns. When normalizing the squared beamforming function, $\underline{g}(\cdot)$,

$$g_s(\Phi_s, \Phi) = \frac{|g(\Phi_s, \Phi)|^2}{\max \left[|g(\Phi_s, \Phi)|^2 \right]}, \quad (14)$$

the formulation is exploited to predict specific sound source energy in the beamforming process. Prediction of multiple sound sources requires an energy sum of multiple filter directions as

$$\mathbf{H}_S(\Phi_S, \Phi) = \sum_{s=1}^S A_s g_s(\Phi_s, \Phi), \quad (15)$$

with A_s representing strength associated with s th sound source. $\Phi_S = \{\theta_1, \dots, \theta_S; \phi_1, \dots, \phi_S\}$ are S number of sound source directions.

Figure 2 illustrates the beam patterns in their normalized energy form expressed in Equations (13)–(15). Figure 2a shows single beam patterns for $S = 1$, $N = 2, 3, 6, 10$ and 16 , respectively. Figure 2b shows the beam pattern of two simultaneous sources with order $N = 4$ for $S = 2$ at $\Phi_1 = \{75^\circ, 90^\circ\}$ and $\Phi_2 = \{270^\circ, 90^\circ\}$, and $A_1 = A_2$, while Figure 2c illustrates the beam pattern of three simultaneous sources with order $N = 8$, for $S = 3$ at $\Phi_1 = \{120^\circ, 90^\circ\}$, $\Phi_2 = \{90^\circ, 270^\circ\}$; and $\Phi_3 = \{45^\circ, 240^\circ\}$, and $A_1 = A_2 = A_3$. Note that the data and the model are formulated in terms of sound energy as in Equations (12) and (14), therefore, a degree of coherence of simultaneous sound sources is expected to play an insignificant role as long as the simultaneous sound sources find themselves in different angular directions.

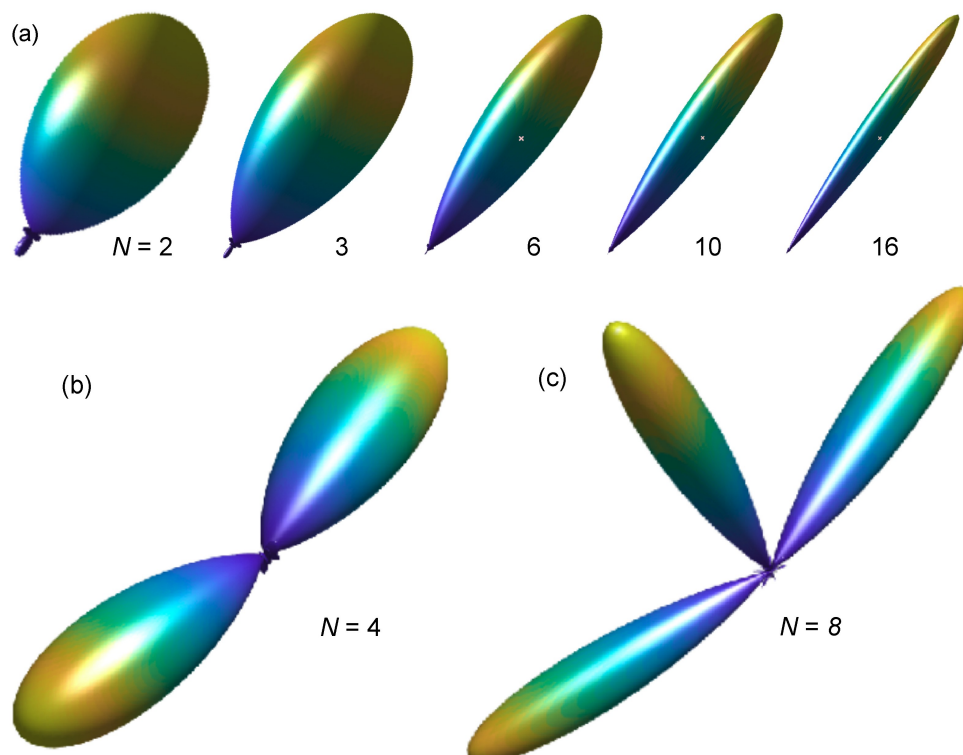


Figure 2. Spherical harmonics beam patterns. (a) Single beam patterns for $N = 2, 3, 6, 10$ and 16 ; (b) Two different, simultaneous beamforming directions of order $N = 4$. (c) Three different, simultaneous beamforming directions of order $N = 8$.

3. Model-Based Bayesian Inference

The beamforming models in Equation (15) formulated previously are now applied to acoustic experimental data as formulated in Section 2.3, particularly for the cases where multiple simultaneous sound sources are potentially contained in the beamforming data. The DoA analysis requires two-levels of inference. On one hand, there is a higher level question of how many sound sources are present. On the other hand, after obtaining an answer for the correct number of simultaneous sound sources, there is a lower level question of determining the parameters of the present sound sources, e.g., incident angles and strengths.

To be more precise, the data, \mathbf{D} , processed using Equation (12) are potentially well described by one set of finite competing models (hypotheses) $\mathbf{H}_1, \mathbf{H}_2, \dots, \mathbf{H}_S$. Often one of the models is preferred to predict the data. For the finite model set with S models, each model, \mathbf{H}_s , is associated with s number of sound sources, with $s \in [1, S]$. Bayesian inference applied to the model selection is a higher level of inference, also known as the second level of inference. It represents an inverse problem to infer which one of the models, \mathbf{H}_s , the data prefer under multiple simultaneous sound sources. The model, \mathbf{H}_s , contains a set of parameters, Θ_s , representing s number of sound sources with their individual strength A_s and angular direction Φ_s . Bayesian inference applied to estimating DoA parameters is referred to as the lower level of inference, also known as the first level of inference. Bayesian inference enables both parameter estimation and the model selection by applying Bayes' theorem intensively to these two levels of inference. The following discussion begins with the second level of inference namely, sound source enumeration by model selection, followed by the DoA parameter estimation. This top-down approach is logical; Only when a proper model, \mathbf{H}_s , is selected among competing models, the lower level of inference, parameter estimation, can properly estimate the DoA parameters, Θ_s , encapsulated in the selected model, \mathbf{H}_s .

3.1. Bayesian Model Selection

This higher level of inference applies Bayes' theorem to determine the probability of one of a finite set of models, \mathbf{H}_i , given the experimentally measured data, \mathbf{D} , as processed by Equation (12) and the background information, I , including a preselected S number of models expressed by Equation (15), which describes the data well,

$$p(\mathbf{H}_i|\mathbf{D}, I) = \frac{p(\mathbf{D}|\mathbf{H}_i, I) p(\mathbf{H}_i|I)}{p(\mathbf{D}|I)}, \quad (16)$$

where $p(\mathbf{H}_i|\mathbf{D}, I)$ is the posterior probability of model, \mathbf{H}_i , $p(\mathbf{D}|I)$ is the probability of observing the experimental data, and for this work it will act as a normalizing constant. $p(\mathbf{H}_i|I)$ is the prior probability of the model, \mathbf{H}_i , and should be assigned based on any prior knowledge of the circumstance. Finally, $p(\mathbf{D}|\mathbf{H}_i, I)$ is the marginal likelihood of a model given the measured data, otherwise known as "Bayesian evidence" [5]. This term is key in the model selection. In the current context as expressed in Equation (16), Bayes' theorem represents how one's prior knowledge about the model, $p(\mathbf{H}_i|I)$, is updated in the presence of the data, which are involved through $p(\mathbf{D}|\mathbf{H}_i, I)$. At this stage, any interest in directional parameter values will be deferred into the background.

To pursue Bayesian model evaluation, Bayes factor, or odds ratio [32] is used to compare two models: model, \mathbf{H}_i , over model, \mathbf{H}_j , as

$$B_{ij} = \frac{p(\mathbf{H}_i|\mathbf{D}, I)}{p(\mathbf{H}_j|\mathbf{D}, I)} = \frac{p(\mathbf{D}|\mathbf{H}_i, I)}{p(\mathbf{D}|\mathbf{H}_j, I)}, \quad \forall i, j \in [1, S], i \neq j, \quad (17)$$

where no preference to any of the models assigns equal prior probability to $p(\mathbf{H}_i|I)$, $i \in [1, S]$. For computational convenience, the Bayes Factor is determined in logarithmic scale with the unit “decibans” [33],

$$L_{ij} = 10 \log_{10}(B_{ij}) = 10 \log_{10}(Z_i) - 10 \log_{10}(Z_j), \text{ [decibans]}, \quad (18)$$

with simplified notations for Bayesian evidence, $Z_i = p(\mathbf{D}|\mathbf{H}_i, I)$, and $Z_j = p(\mathbf{D}|\mathbf{H}_j, I)$. This enables the evidences for two models to be quantitatively compared against one another. Among a finite set of models, the highest positive Bayes factor, L_{ij} , indicates that the data prefer model \mathbf{H}_i over \mathbf{H}_j the most. Therefore, the Bayes factor is also applied to select a finite number of models under consideration in the following (Section 6).

Overall, this process offers a penalty for over-complicated models if they only increase maximum likelihood rather than the Bayesian evidence compared to simpler models. This is the quantitative implementation of Occam’s razor, which favors simplicity over complexity when comparing models that competitively predict measured data [34].

3.2. Bayesian Parameter Estimation

On the lower level of the inference, the background information, I , now denotes that the model, \mathbf{H}_s , predicting s number of sound sources, is already given as discussed above in Section 3.1, and the selected model describes the experimental data well. Now Bayesian inference focuses the attention to the DoA parameters, Θ_s , encapsulated in the selected model, \mathbf{H}_s . Since the model is already specified through the Bayesian model selection, the subscripts of \mathbf{H}_s and Θ_s will be dropped for simplicity throughout the following discussions, but still bearing in mind that the model, \mathbf{H} , has been given via the model selection. The model contains a specific set of parameters, $\Theta = \{\theta; \phi; \mathbf{A}\}$, including, both angular and amplitude parameters for a specific number of sound sources. The DoA parameter estimation applies Bayes’ Theorem to determine the probabilities of parameters, Θ , given data, \mathbf{D} , model, \mathbf{H} , and the background information, I , yielding,

$$p(\Theta|\mathbf{D}, \mathbf{H}, I) = \frac{p(\mathbf{D}|\Theta, \mathbf{H}, I) p(\Theta|\mathbf{H}, I)}{p(\mathbf{D}|\mathbf{H}, I)}. \quad (19)$$

Probability $p(\Theta|\mathbf{D}, \mathbf{H}, I)$ is referred to as the posterior probability distribution of the parameters, Θ . Quantity $p(\mathbf{D}|\Theta, \mathbf{H}, I)$, represents the likelihood that the measured data, \mathbf{D} , would have been generated for a given value of Θ . It is termed in the following as likelihood in short. Term $p(\Theta|\mathbf{H}, I)$, represents the prior probability of the parameters given the model, \mathbf{H} . Finally, term $p(\mathbf{D}|\mathbf{H}, I)$ is the same as the marginal likelihood $p(\mathbf{D}|\mathbf{H}, I)$ in Equations (16) and (17). This quantity is also known as Bayesian evidence [5,35], or evidence, in short. Bayes’ theorem, applied to the parameter estimation problem as stated in Equation (19), represents how the prior knowledge about the parameter, $p(\Theta|\mathbf{H}, I)$, is updated in the presence of data, which are incorporated through the likelihood, $p(\mathbf{D}|\Theta, \mathbf{H}, I)$.

3.3. Unified Bayesian Framework

The integral of any proper probability (density) over the entire parameter space in which it is defined equals unity. When integrating the both sides of Equation (19) it results in

$$Z = p(\mathbf{D}|\mathbf{H}, I) = \int_{\Theta} p(\mathbf{D}|\Theta, \mathbf{H}, I) p(\Theta|\mathbf{H}, I) d\Theta, \quad (20)$$

where the evidence, $p(\mathbf{D}|\mathbf{H}, I)$, as in a simplified notation, Z , does not depend on Θ , therefore, is taken out of the integral. Equation (20) indicates that the evidence of a given model, Z , is evaluated over the entire parameter space by integrating the product of the likelihood and prior distribution. This is the same evidence value as in Equations (16) and (17), expressing that both processes of the

model selection and the parameter estimation involve evaluating the likelihood of a given model over its parameter space. Therefore both levels of Bayesian inference can be performed within a unified framework, as elaborated in the following.

According to Skilling [35] Equation (19) is rewritten in simplified notations as

$$\underbrace{p(\Theta|\mathbf{D}, \mathbf{H}, I)}_{\text{posterior}} \cdot \underbrace{Z}_{\text{evidence}} = \underbrace{\mathcal{L}(\Theta)}_{\text{likelihood}} \cdot \underbrace{\pi(\Theta)}_{\text{prior}}, \quad (21)$$

where evidence is determined by evaluating likelihood, $\mathcal{L}(\Theta) = p(\mathbf{D}|\Theta, \mathbf{H}, I)$, and prior, $\pi(\Theta) = p(\Theta|\mathbf{H}, I)$ using Equation (20). Equations (20) and (21) indicate that the Bayesian evidence plays a central role in the model selection. The evidence relies on exploration of the likelihood over the entire parameter space, which is also required in the parameter estimation, based on the estimation of the posterior in Equation (19). The formulation in both Sections 3.1 and 3.2 can be accomplished within one unified framework. In this Bayesian framework, two terms on the right-hand side of Equation (21) are input quantities, particularly the likelihood function in Equation (21), while the two terms on the left-hand side are the output quantities; the evidence, Z , represents the output for the Bayesian model selection, and the posterior, $p(\Theta|\mathbf{D}, \mathbf{H}, I)$, represents the output for the Bayesian parameter estimation.

4. Maximum Entropy Priors

Jaynes [36] applied a continuum version of Shannon [37] entropy to encode the available information into a prior probability assignment. The assignment maximizes the entropy in order to obtain the prior probability. In Bayesian literature [36,38], this is so-called the principle of maximum entropy. Two input quantities are all prior probabilities, which need to be assigned prior to pursuing further analysis.

4.1. Likelihood Assignment

The likelihood is collectively determined by probabilities of the residual errors, $p(e_{j,k})$. This is the difference between the data in Equation (12) and the model prediction in Equation (15) at each single datum,

$$e_{j,k} = D(\theta_j, \phi_k) - H(\theta_j, \phi_k), \quad (22)$$

where $e_{j,k}$, namely $\mathbf{e} = \mathbf{D} - \mathbf{H}$ are in the form of two-dimensional matrices over $\Phi_{j,k} = \{\theta_j, \phi_k\}$ within this work.

The likelihood assignment also incorporates what is known about the model, \mathbf{H} , that has been formulated in Equation (15) in Section 2.3 through Equations (13) and (14). So the models are also part of prior information [39]. Notation $\mathcal{L}(\Theta) = p(\mathbf{D}|\Theta, \mathbf{H}, I)$ in Equations (19) and (21), explicitly expresses this conditional statement through 'the given model, \mathbf{H} ' and 'background information, I '. The probability for the likelihood $\mathcal{L}(\Theta)$, including all $p(e_{j,k})$, should be assigned based on what is known about the error function.

The only information about the residual errors, $e_{j,k}$, is that the error energy is limited to a finite, yet unknown bound due to the fact that the model is known to be able to predict the data well. This prior information is therefore encoded as a finite, yet unknown error variance. In addition, a universal constraint on the probability density, or the so-called normalization constraint, is that the integral of the individual probability (density) equals unity. Application of the principle of maximum entropy by taking the finite error variance and the normalization into the assignment, leads to a Gaussian probability distribution [36,40],

$$p(e_{j,k}|\Theta, \mathbf{H}, \sigma_{j,k}) = \frac{1}{\sqrt{2\pi}\sigma_{j,k}} \exp\left(-\frac{e_{j,k}^2}{2\sigma_{j,k}^2}\right). \quad (23)$$

The residual errors are also assigned zero-means, $\mu_{j,k} = 0$, since any other non-zero values can be included in the model in Equation (15) when necessary by adding another unspecified parameter. Note that this Gaussian assignment is the consequence of limited information on the residual errors, $\mathbf{e} = \{e_{j,k}\}$. Namely, only a finite, yet unspecified error variance is available. This Gaussian assignment is distinctly different from assuming the statistics of the residual errors to be Gaussian.

The principle of the maximum entropy regards the residual errors as independent of each other [36], since any dependence or correlation will reduce the entropy. The overall likelihood becomes the product of the individual error probabilities

$$\mathcal{L}(\Theta) = \prod_{j=1}^J \prod_{k=1}^K \frac{1}{\sqrt{2\pi}\sigma_{j,k}} \exp\left(-\frac{e_{j,k}^2}{2\sigma_{j,k}^2}\right) = (\sqrt{2\pi}\sigma)^{-Q} \exp\left(-\frac{E}{\sigma^2}\right), \quad (24)$$

with σ^2 being a constant, unspecified error variance across the data points, Q is the total number of data points, $Q = J \cdot K$, and

$$E = \frac{1}{2} \sum_{j=1}^J \sum_{k=1}^K [D(\theta_j, \phi_k) - H(\theta_j, \phi_k)]^2, \quad (25)$$

with $\theta_1 \leq \theta_j \leq \theta_J$ and $\phi_1 \leq \phi_k \leq \phi_K$ covering the entire angular range under consideration. Data, $D(\theta_j, \phi_k)$, and model, $H(\theta_j, \phi_k)$, are determined by Equations (12) and (15), respectively.

4.2. Prior Probability Assignment

For the prior probability, $\pi(\Theta)$, other than the normalization constraint, no other prior knowledge on parameter values is available. Typical model parameters are also location parameters, just as Θ is in the current work. The principle of maximum entropy assigns $\pi(\Theta)$ to be a uniform distribution over a wide parameter value range [36].

In similar fashion, the model prior, $p(\mathbf{H}_i|I)$, in Equation (16) within the model selection is also assigned a constant prior, within a discrete, finite number (S) of models,

$$p(\mathbf{H}_i|I) = \frac{1}{S}, \quad (26)$$

which leads to Equations (17) and (18) in Section 3.1.

The hyperparameter, σ , in Equation (24) is a consequence of the maximum entropy assignment of the likelihood. Representing a scale parameter, it has to be assigned as well. The principle of maximum entropy also assigns a uniform prior to the scale parameter, but in the logarithmic domain, since the scale parameter acts invariant only in the logarithmic domain [36,38]. This assignment leads to the so-called Jeffreys' prior [41],

$$p(\sigma) = \frac{1}{\sigma}. \quad (27)$$

Bretthorst [42] considers the hyperparameter, σ , as a nuisance in a number of applications. It is the case also in the current work and can be removed by applying Jeffreys' prior for marginalization. The marginalization removes the hyperparameter [42,43] from the likelihood in Equation (24), leading to Student-t distribution,

$$\mathcal{L}(\Theta) \propto \Gamma\left(\frac{Q}{2}\right) \frac{(2\pi E)^{-Q/2}}{2}, \quad (28)$$

where $\Gamma(\cdot)$ is the Gamma function, Q is the total number of data points, and E is given in Equation (25).

5. Nested Sampling

The Bayesian framework applied to the DoA analysis for multiple sound sources requires numerical calculations of the evidence. Different sampling methods exist for this purpose. A recent

overview on a number of suitable methods for calculating Bayesian evidence is given by Knuth [5]. This work employs nested sampling originally proposed by Skilling [35,44].

5.1. Lebesgue Integration as Foundation

Nested sampling represents a Markov chain Monte Carlo method, estimating directly how the likelihood distribution relates to the prior mass, and partitions the range of the likelihood distribution similarly to Lebesgue integration [45,46] as opposed to the parameter space domain over which the likelihood is defined. The evidence as given in Equation (20) requires integral calculation over the entire multi-dimensional parameter space. In the unified framework, nested sampling yields the evidence as the prime result, while samples from the posterior distribution are an optional byproduct [44]. Using simplified notation similar to Equation(21), the evidence is determined by

$$Z = \int_{\Theta} \mathcal{L}(\Theta) \pi(\Theta) d\Theta = \int_{\mu} \mathcal{L}(\mu) d\mu, \quad (29)$$

where a differential notation, $d\mu = \pi(\Theta) d\Theta$, is introduced. The differential element, $d\mu$, represents volume under prior distribution over elementary parameter space, $d\Theta$. It is termed elementary prior mass. An accumulated prior mass, in the form of Lebesgue measure [45,46] can then be defined as

$$\mu(\mathcal{L}_{\varepsilon}) = \int_{\mathcal{L}(\Theta) > \mathcal{L}_{\varepsilon}} \pi(\Theta) d\Theta, \quad (30)$$

where $\mathcal{L}_{\varepsilon}$ is a certain value among the likelihood range. Expressing the inverse function $\mathcal{L}[\mu(\mathcal{L}_{\varepsilon})] = \mathcal{L}_{\varepsilon}$ [44], this variable change converts the evidence expressed in Equation (29) into a one-dimensional integral over unit range

$$Z = \int_0^1 \mathcal{L}(\mu) d\mu. \quad (31)$$

As likelihood value $\mathcal{L}_{\varepsilon}$ increases, the enclosed prior mass $\mu(\mathcal{L}_{\varepsilon})$ decreases from 1 to 0. At its minimum, $\mathcal{L}_{\varepsilon} = 0$, this corresponds to the maximum prior mass. Particularly, it encloses the prior, $\pi(\Theta)$, over the entire parameter space, so that $\mu(\mathcal{L}_{\varepsilon} = 0) = 1$. In the opposite, when likelihood value $\mathcal{L}_{\varepsilon} \rightarrow \mathcal{L}_{\max}$, namely, approaches the maximum, the prior mass approaches zero, $\mu(\mathcal{L}_{\max}) \rightarrow 0$ [see Figure 3].

Nested sampling partitions the likelihood range between $0 \leq \mathcal{L}_{\varepsilon} \leq \mathcal{L}_{\max}$, in a Monte Carlo manner which leads to

$$0 \approx \mathcal{L}_{\min} = \mathcal{L}_0 < \mathcal{L}_1 < \dots < \mathcal{L}_{t-1} < \mathcal{L}_t < \dots < \mathcal{L}_{\max}. \quad (32)$$

Iterations of the nested sampling implementation as shown in Figure 3a create this likelihood sequence that corresponds to a prior mass sequence

$$1 = \mu_0 > \mu_1 > \dots > \mu_{t-1} > \mu_t > \dots > \mu_{\min} = 0, \quad (33)$$

as graphically illustrated in Figure 3a with labels at the bottom. These two sequences lead to the numerical approximation of the evidence in Equations (29) and (31)

$$Z \approx \sum_{t=0}^T \mathcal{L}_t \Delta\mu_t, \quad (34)$$

with $\mathcal{L}_0 = \mathcal{L}_{\min}$, $\mathcal{L}_T = \mathcal{L}_{\max}$, $\mu_T = \mu_{\min}$, $\mu_0 = 1$, and

$$\Delta\mu_t = \mu_{t-1} - \mu_t. \quad (35)$$

After a decent number of steps, also by acknowledging uncertainties [44], the prior mass is supposed to shrink,

$$\mu_t \approx e^{-t/P}, \tag{36}$$

where number, P , is used for constructing an initial sampling population as elaborated below.

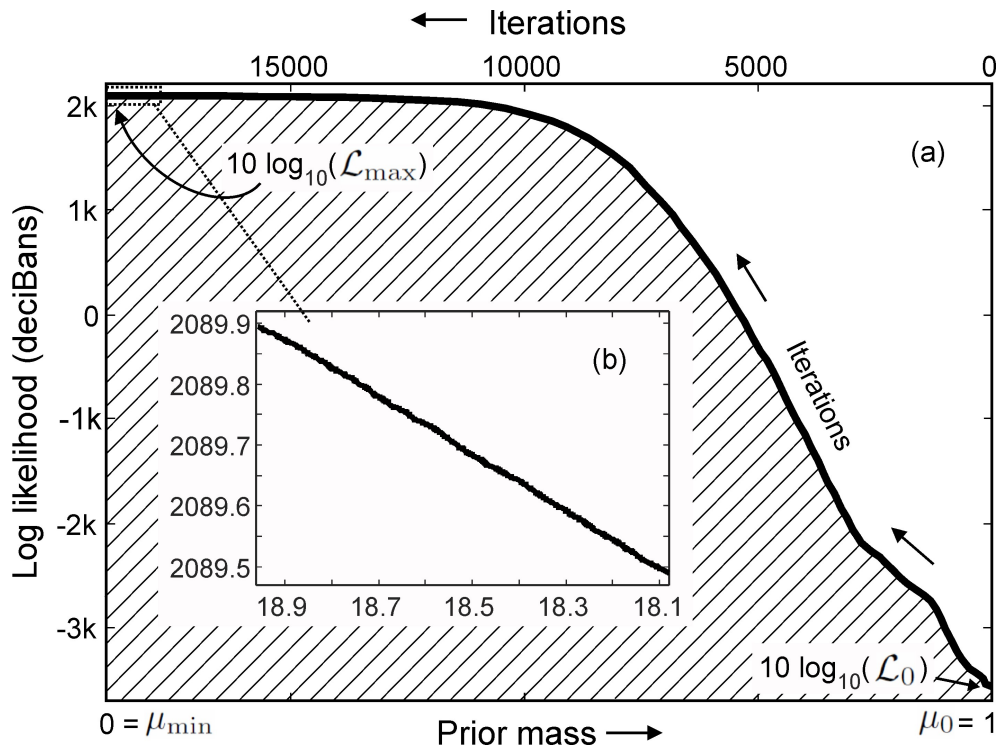


Figure 3. Log likelihood values vs. the prior mass (bottom) and the nested sampling iterations (top) for the two-source data with the two-source model. (a) Entire course of sampling for all ($T =$) 18,968 iterations. The prior mass is labeled at the bottom of the horizontal axis from left to right, while number of iterations are labeled at the top from right to left. The shaded area under the likelihood curve corresponds to the evidence; (b) magnified segment when nested sampling approaching to convergence.

5.2. Major Implementation Steps

Main steps in a practical implementation of nested sampling for each model are summarized as follows:

1. Draw a sufficient initial population, P , uniformly distributed samples, containing randomly generated values of all parameters, based on the maximum entropy prior probability (Section 4.2). In this case, $P = 1000$.
2. Evaluate the likelihood value of each sample Θ_i using Equation (28) inside the P populations in which the model in Equation (15) is involved at each sample.
3. Identify the sample Θ_m with the lowest likelihood value, \mathcal{L}_m within the population.
4. Store this lowest likelihood value along with associated sample $[\mathcal{L}_m, \Theta_m] \rightarrow [\mathcal{L}_t, \Theta_t]$ in a list outside the initial population. The list is referred to as the sample list below.
5. For the following step $t + 1$, perturb the parameters associated with this least-likely sample in a random fashion and re-evaluate the likelihood value, with constraint, $\mathcal{L}_{t+1} > \mathcal{L}_t$.
 - (a) If the perturbed sample now meets the constraint, replace $[\mathcal{L}_t, \Theta_t]$ by this new sample, $[\mathcal{L}_{t+1}, \Theta_{t+1}]$, into the initial population, then move on to the next step.

- (b) If not, keep perturbing in a way of randomly walking around in the parameter space, until the above constraint is fulfilled.
6. Repeat steps 3–5 until the sample population has satisfied some predefined convergence criteria [44], or until some maximum number of iterations is met.
7. The sample list storing all the likelihood values along with their samples, $[\mathcal{L}_t, \Theta_t]$, $t \in [1, T]$ are already in the sequence in Equation (32) which is the ordered partition. The summation in Equation (34) along with Equations (35) and (36) using this likelihood sequence approximates the evidence estimate for the model under consideration.
8. Repeat steps 1–7 for all the beamforming models, \mathbf{H}_j , $j \in [1, S]$, under consideration to approximate evidence estimate for each models using Equations (34) and (35).
9. Use the evidence estimates from step 8 to evaluate Bayes factor distribution using Equation (18) over all the models, this facilitates the model selection.

5.3. Evidence Via Likelihood Range Partitions

Figure 3 illustrates one nested sampling run for an experimental beamforming data set of two simultaneous sound sources given a two-source model (\mathbf{H}_2). The figure illustrates the integral expression in Equation (31) as shaded area with the prior mass being the principle variable going from 0 to 1. The vertical axis represents 10 times the logarithm of likelihood in base 10, that is in unit [decibans]. In the numerical implementation at the start of sampling, the likelihood at the first iteration is the lowest value at outright side of the figure. The iterations are labeled leftwards on the upper side of the horizontal-axis. This lowest likelihood value is stored in the sample list in the form of $10 \log_{10}(\mathcal{L}_0)$. This value corresponds to the maximum prior mass ($\mu_0 = 1$) since it includes the entire parameter space, expressed on the left-hand side of Equation (33). As the sampling iteration progresses, once the hard constraint, $\mathcal{L}_{t+1} > \mathcal{L}_t$, is fulfilled, the log likelihood value, $10 \log_{10}(\mathcal{L}_{t+1})$, along with its parameters, Θ_{t+1} will be repetitively stored into the sampling list and at the same time, replacing the previous lowest sample associated with $[\mathcal{L}_t, \Theta_t]$. With the iterations progressing, the log likelihood value increases, while the prior mass decreases. As the sampling converges through many iterations, the likelihood climbs to its maximum value, the prior mass at the convergence state shrinks to zero ($\mu_{\min} = 0$). Figure 3b shows a magnified segment of the converging likelihood sequence. Once the exploration criteria described in step 6 have been met, the sampling evolution creates the likelihood sequence as in Equation (32), as shown in Figure 3a which is then used in Equation (34) to estimate the evidence indicated by the shaded area in the figure. Nested sampling leads to the likelihood sequence as in Equation (32), which essentially partitions the likelihood range over the prior mass of the entire parameter space. In this example, illustrated in Figure 3, the upper bound in Equation (34) is $T = 18968$. These evidence estimates allow for evaluating/ranking the competing models.

5.4. Posterior Estimates as Byproducts

One model among the finite model set should be selected for use in the lower level of inference, the DoA parameter estimation. As discussed previously, this work benefits from the unified Bayesian framework, since the thorough exploration over the entire parameter space has been performed in order to estimate evidences. After the model selection, the evidence value of the selected model along with all the likelihood values and the associated random sample parameters are already available and stored in the respective sample list. They lead directly to samples from the posterior distribution for the parameter estimation. All the samples in the sample list $[\mathcal{L}_t, \Theta_t]$, $t \in [0, T]$ are readily available to estimate the mean DoA parameters

$$\tilde{\Theta} = \sum_{t=0}^T p_t \Theta_t, \quad (37)$$

with posterior samples

$$p_t = \frac{\mathcal{L}_t \Delta \mu_t}{Z}, \quad t \in [0, T], \quad (38)$$

where $\Delta \mu_t$ is taken from Equation (35), \mathcal{L}_t is the likelihood sequence in Equation (32), resulting from nested sampling, and the parameter variance as

$$\sigma^2(\Theta) = \left[\sum_{t=0}^T p_t (\Theta_t - \tilde{\Theta})^2 \right]. \quad (39)$$

Bush and Xiang [7], Jasa and Xiang [46], Fackler et al. [47] have recently implemented nested sampling in other acoustics applications.

6. Experimental Results

This work experimentally investigated two and three simultaneous sound sources around the spherical microphone array for obtaining various impulse responses sets. The array contained sixteen microphones flush-mounted on a rigid sphere of 6 cm in diameter. The experimental measurement utilized a single sound source (a loudspeaker) to measure impulse response at various locations around the microphone array. Logarithmic sweep sines are used to excite the loudspeaker and the spherical microphone array providing sixteen channel of responses to this excitation. The loudspeaker was placed 1.5 m away from the microphone array in a sufficiently large indoor space. All the responses to the sweeps were averaged and transferred into impulse responses to improve the signal-to-noise ratio.

These impulse responses, with peak-to-noise ratios over 65 dB, are windowed to isolate the direct sound portions so that the individual impulse responses convolved with white noise are considered from anechoic environment. To synthesize multiple noise sources from different locations around the spherical array, direct-sound portions of the impulse responses measured from different source locations are convolved with the white noise which are combined via linear superposition. The spherical harmonic beamforming for two and three simultaneous sound sources is carried out for these experimental data. The beamforming data in Equation (12) are summed up between 400 Hz and 4 kHz to form the sound energy map over angular range $0 \leq \theta \leq 180^\circ$, $0 \leq \phi \leq 360^\circ$. The results here demonstrate the prediction capability of the model in Equation (15) for the experimental data and that the two-level Bayesian inference quantitatively implements Occam's razor to estimate the number of sound sources present in the data. After the Bayesian model selection, the estimated DoA parameters are then obtained using the selected model.

Figure 4 illustrates the results for two simultaneous sound sources over an angular range of $360^\circ \times 180^\circ$ for azimuth, ϕ , and elevation, θ . Figure 4a illustrates the sound energy distributions derived from experimentally measured data using Equations (8)–(12), while Figure 4b shows the predicted results using Equations (13)–(15) to visualize the sound field distribution around the spherical microphone array in Cartesian coordinates. The grid resolution for these two-dimensional maps is $3.6^\circ \times 3.6^\circ$ with grid points of $K \times J = 100 \times 50$ across azimuth and elevation range as expressed in Equation (25).

Figure 5 illustrates Bayes factor estimations over the different models \mathbf{H}_S from Equation (15). Each model represents a different number of sound sources. The Bayesian evidence for each model is evaluated over 16 individual runs using nested sampling. According to the Bayesian model selection scheme discussed in Section 3.1, Figure 5a illustrates the Bayes factor estimates, L_{ij} from Equation (18) in decibans, from $i = 2, \dots, 4$ over $j = 1, \dots, 3$.

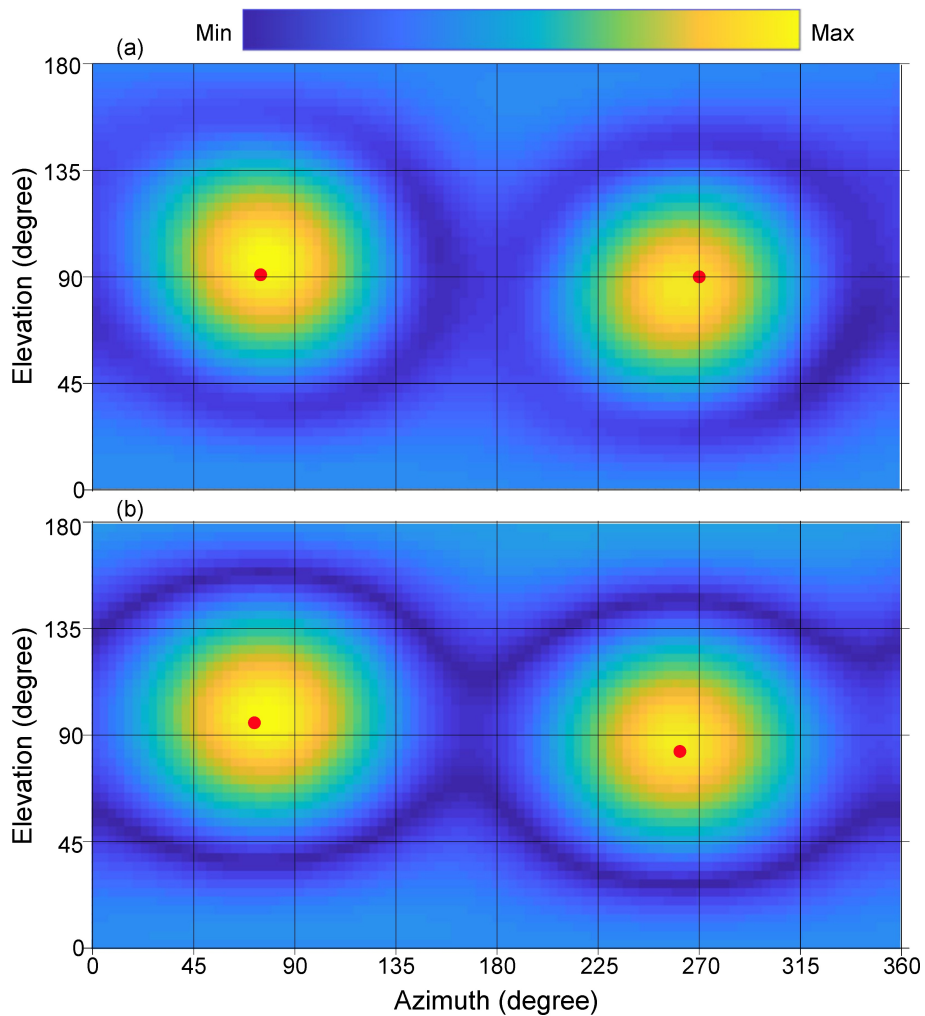


Figure 4. Directional responses of two simultaneous sound sources in the form of two-dimensional sound energy distributions. The directions of the sound sources are at $(75^\circ, 90^\circ)$, $(270^\circ, 90^\circ)$. (a) Experimentally measured beamforming data. Two solid dots indicate the directions of arrivals assigned in the experiment; (b) Sound energy distribution model predicted by the Bayesian model selection process. Two solid dots indicate the estimated directions of arrivals.

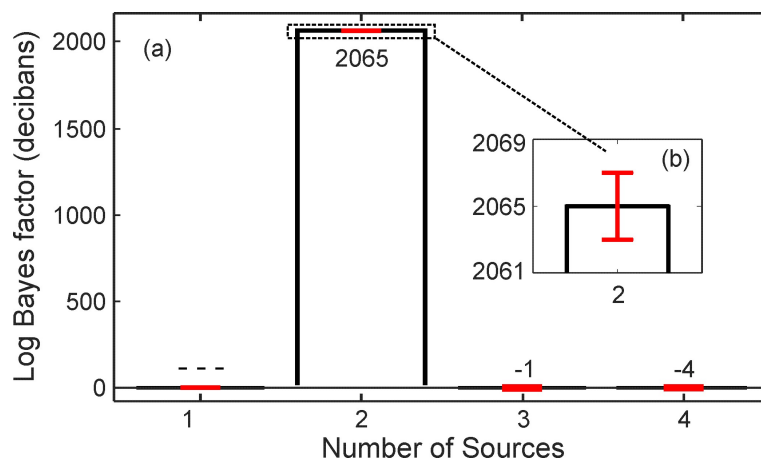


Figure 5. Mean Bayes factor estimates along with variances given the experimental data. The data contain two sound sources at $(75^\circ, 90^\circ)$, and $(270^\circ, 90^\circ)$. (a) Bayes factor in decibans comparing the evidence of the current number of sources to the previous number; (b) Magnified view of the Bayes factors for two sources with variations.

Table 1. Experimentally measured and predicted directions of arrivals (DoAs) for two simultaneous sound sources. The variations are estimated using the Bayesian method over 15 runs. The errors are differences between experimental and predicted ones. Experimental data are analyzed in an angular resolution of 3.6° .

Comparison	Direction of Arrival (ϕ, θ)	
Experiment	$(75^\circ, 90^\circ)$	$(270^\circ, 90^\circ)$
Estimates	$(70.58^\circ, 95.46^\circ)$	$(261.2^\circ, 80.2^\circ)$
Deviation	$(\pm 1.55^\circ, \pm 0.73^\circ)$	$(\pm 1.47^\circ, \pm 0.66^\circ)$
Error	$(4.42^\circ, 5.46^\circ)$	$(8.8^\circ, 9.8^\circ)$

The highest Bayes factor, L_{21} is identified for the case of two sources, expressing that the data prefer model, \mathbf{H}_2 , over \mathbf{H}_1 , much more than the preference of model \mathbf{H}_3 over \mathbf{H}_2 , and so on, where \mathbf{H}_5 is given previously in Equation (15) with $S = 1, 2, 3$, and 4. After model selection, the evidence estimate of the two source model can be readily used to estimate the posterior samples using Equation (38) for Equations (19) or (21). The Bayesian parameter estimation (in Section 3.2) finds the angular parameters as listed in Table 1. The model in Equation (15) taking this set of parameters for $S = 2$ predicts the sound energy distribution as depicted in Figure 4b. Note that both the experimental and predicted data are analyzed at an angular resolution of 3.6° . Correlating the highest sound energy with the DoA indicates that physically placing the sound sources at the listed directions $(75^\circ; 90^\circ)$, and $(270^\circ; 90^\circ)$ may also be inaccurate. For this reason, prediction errors as listed in Table 1 and also later in Table 2 need to be evaluated considering this source of experiment errors.

In similar fashion, Figure 6 illustrates the results for the set of three simultaneous sound sources over an angular range of $360^\circ \times 180^\circ$ for azimuth, ϕ , and elevation, θ . The grid resolution is $3.6^\circ \times 3.6^\circ$. Figure 6a illustrates the sound energy distributions derived from experimentally measured data using Equations (8)–(12), while Figure 6b, the model predicted results using Equations (13)–(15), for the case $S = 3$.

Figure 7a illustrates Bayes factor estimations over the different models, \mathbf{H}_5 , from Equation (15) for $S = 1, 2, 3, 4$ and 5. The Bayes factor estimates, L_{ij} from Equation (18), for $i = 2, \dots, 5$ over $j = 1, \dots, 4$ as shown in Figure 7a show the highest Bayes factor is for the case of three sources. Namely, the data prefer model, \mathbf{H}_3 , over model, \mathbf{H}_2 , the most. This preference was much higher than that of model \mathbf{H}_4 over \mathbf{H}_3 , and so on. After the selection of the three source model, the Bayesian evidence for this model was readily available for further parameter estimation. At the same time, the likelihood values in Equation (19) for this model have already been thoroughly sampled over the entire parameter space using nested sampling. Therefore the parameter values can be readily extracted from the parameter set using Equation (38) for Equations (19) or (21). The Bayesian parameter estimation (in Section 3.2) leads to the angular parameters as listed in Table 2. The model in Equation (15), given this set of parameters for $S = 3$, predicts the sound energy distribution as depicted in Figure 6b.

Table 2. Experimentally measured and predicted DoAs for three simultaneous sound sources. The variations are estimated using the Bayesian method over 15 runs, The errors are the differences between the experimental and predicted data. Both data sets are analyzed with an angular resolution of 3.6° .

Comparison	Direction of Arrival (ϕ, θ)		
Experiment	$(5^\circ, 60^\circ)$	$(135^\circ, 140^\circ)$	$(270^\circ, 90^\circ)$
Estimates	$(8.7^\circ, 72.4^\circ)$	$(125.8^\circ, 148.6^\circ)$	$(254.1^\circ, 71.5^\circ)$
Deviation	$(\pm 8.6^\circ, \pm 8.2^\circ)$	$(\pm 17.5^\circ, \pm 1.4^\circ)$	$(\pm 7.7^\circ, \pm 7.4^\circ)$
Error	$(3.7^\circ, 12.4^\circ)$	$(9.2^\circ, 8.6^\circ)$	$(15.9^\circ, 18.5^\circ)$

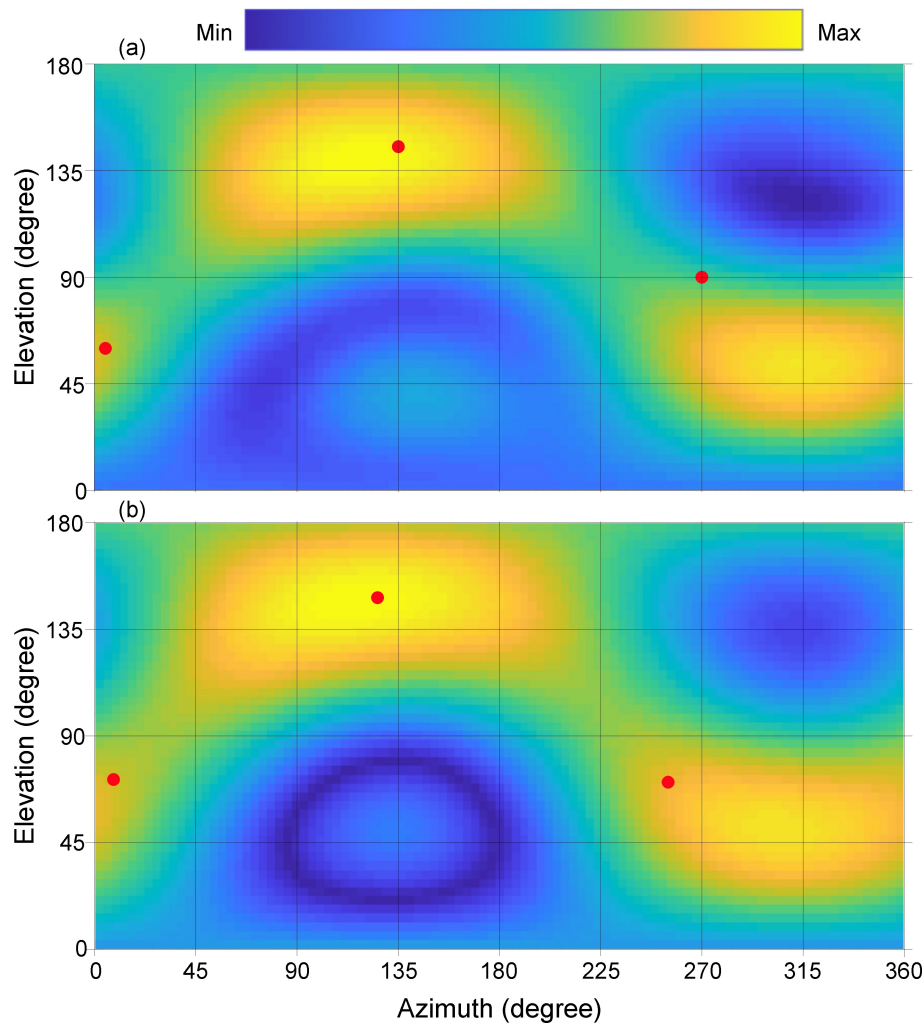


Figure 6. Directional responses of three sound sources in the form of two-dimensional sound energy distributions. The directions of the sound sources are at $(5^\circ, 60^\circ)$, $(135^\circ, 140^\circ)$ and $(270^\circ, 90^\circ)$, respectively; (a) Experimentally measured beamforming data. Three solid dots indicate the directions of arrivals assigned in the experiment; (b) Sound energy distribution model predicted by the Bayesian model selection process. Three solid dots indicate the estimated directions of arrivals.

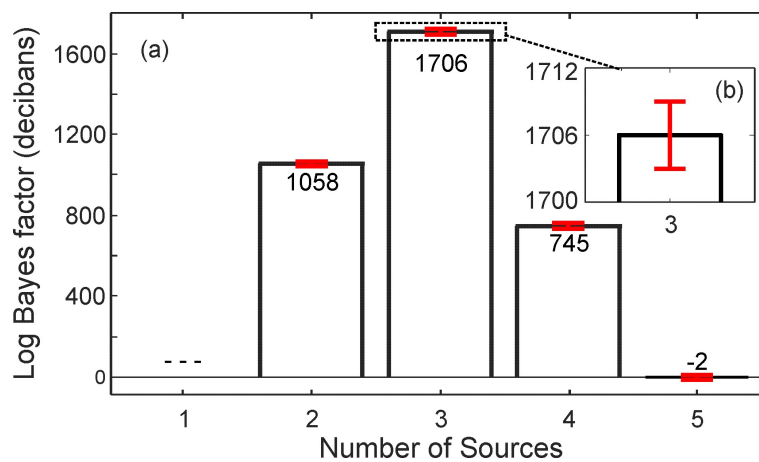


Figure 7. Mean Bayes factors along with variances given the experimental data. The data contain three sound sources at $(5^\circ, 60^\circ)$, $(135^\circ, 140^\circ)$ and $(270^\circ, 90^\circ)$, respectively. (a) Bayes factor in decibans, comparing the evidence of the current number of sources to the previous number; (b) Magnified view of the evidence for four sources with variations.

Figure 6 demonstrates that it would be very challenging to determine the number of sources present based solely on visual inspection or on the peak energy values. The correct number of sound sources may not be correctly determined, let alone their correct locations. Bayesian inference as applied to the DoA analysis significantly improves sound source localization without having to increase the resolution of the spherical microphone array.

7. Discussions

This paper discusses the DoA analysis from the sound sources essentially in an anechoic environment. When two sound sources are well separated as shown in Figure 4, their directions of arrival are straightforwardly recognized. Two solid dots in Figure 4 indicate that both the experimental measurements and Bayesian model prediction are prone to certain errors. Even physical placement of the sound sources at the listed directions ($75^\circ; 90^\circ$), and ($270^\circ; 90^\circ$) may also be inaccurate. Therefore, prediction errors as listed in Tables 1 and 2 need to be evaluated considering this source of experiment errors. In case of three simultaneous sound sources, the estimation errors may drop to 18.5° for some source locations. As mentioned before, the estimation directions are not absolute in their errors, one source of errors also comes from experimental errors when placing the the sound sources.

The results discussed previously indicates that estimation performance will decrease with an increase in number of sound sources or ambiguity of the sound field also manifests itself in the confidence of the model selection process. For the case of two simultaneous sources, the Bayesian evidence estimates alone present stable estimations among individual sampling runs. They also show behavior consistent with Occam's razor, because the test scenario is set for two simultaneous sound sources. As the number of sources increases to three, the variance over individual nested sampling runs becomes slightly larger. Even then, the experimentally measured data are considered to carry sufficient information. The Bayes' factor representing relative Bayesian evidence, is at maximum for the three source model.

The increased variation and the estimation errors from two to three sound sources are clearly a result of a higher number of sound sources. Increasing the order of the spherical microphone array for data recording will be a remedy for the increased variations by higher numbers of sources. It needs to increase the number of microphones channels on the spherical array, resulting in a higher angular resolution. It will be of general interest to investigate the resolution capability given a certain order of spherical microphone array which is beyond the current scope of the research.

8. Concluding Remarks

The present work applies the Bayesian method to beamformed models, evaluating them against experimental data. A spherical microphone array provides sixteen channels of these data in order to estimate the DoAs of simultaneous sound sources. Both data and the models are formulated using spherical harmonics in Section 2. Through a two-level of inferential approach to this problem involving first estimating the number of sound sources as solved by Bayesian model selection (in Section 3.1) and second estimating their DoAs as solved by Bayesian parameter estimation (in Section 3.2). Both of these pieces of information can be reliably estimated within the unified Bayesian framework. This Bayesian inference approach provides an improvement in the detection of sound sources over alternative methods, such as those that directly correlate the peak sound energies to the DoAs.

This work demonstrates the feasibility of nested sampling applied in Bayesian model selection as a means to determine the number of sound sources, while the DoA parameters are the byproduct of the sampling exploration upon selecting the correct number of sound sources. The nested sampling implementation in this work shows its efficacy on experimentally measured data for two and three simultaneous sound sources. The Bayes factors evaluated sequentially from one model of a given number of sources against the next lower number model are able to select the right model unambiguously. The DoA parameters estimated for both two and three simultaneous sources indicate success of the Bayesian application. Potential estimation errors are also discussed in details.

The experiments carried out within this work are essentially in anechoic environment. General room acoustical applications using this method of DoA analysis still remain to be explored through future efforts. Challenges could potentially be determination of locations of distinct, strong surface reflections in addition to the DoA of sound sources within an enclosed space.

A sixteen channel spherical microphone array has been experimentally tested in this work. This second order spherical microphone array offers relatively limited spatial resolution. Increasing the number of microphone channels would increase the spatial resolution, thus allowing for a more definitive localization of simultaneous sound sources. This also allows for more sound sources to be localized. Though this research only tested up to three sound sources, many complex sound fields have far more than simply three distinct sources occurring at the same time. Investigations using Bayesian inference should be conducted in the near future, in hopes of discovering ability to handle challenges in more complicated situations.

Full spherical microphone/sensor arrays are more suitable for applications when sound sources are expected around the arrays from all possible directions, such as hanging in open spaces or mooring in deep oceans. In addition, the Bayesian formulation based on spherical harmonics is also straightforwardly extended to hemispherical or cylindrical array configurations. Another future effort should be relaxing the plane-wave requirements so as to formulate spherical waves for near-field conditions. In the future this will open up opportunities for range estimates of sound sources near the sensing array, in addition to solely direction of arrival analysis.

Author Contributions: Conceptualization, N.X.; methodology, N.X.; software, C.L.; validation, N.X. and C.L.; formal analysis, N.X. and C.L.; investigation, C.L.; resources, N.X.; data curation, C.L.; writing—original draft preparation, N.X.; writing—review and editing, C.L.; visualization, N. X. and C.L.; supervision, N.X.; project administration, N.X.

Funding: This research received no external funding.

Acknowledgments: The authors are grateful to Jonathan Kawasaki, Stephen Weikel, and Jonathan Matthews for their collaborative support in the instrumentation of spherical harmonics microphone arrays. The authors also thank J. Botts, and S. Clapp for the insightful discussions during the early stage of this work.

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

References

1. Meyer, J.; Elko, G. A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield. In Proceedings of the 2002 IEEE International Conference on Acoustics, Speech, and Signal Processing, Orlando, FL, USA, 6 August 2002; pp. 1781–1784. [[CrossRef](#)]
2. Rafaely, B. *Fundamentals of Spherical Array Processing*; Springer: Berlin/Heidelberg, Germany, 2015; Volume 8.
3. Williams, E. *Fourier Acoustics: Sound Radiation and Near Field Acoustical Holography*; Academic Press: London, UK, 1999; Volume 2, pp. 185–232. [[CrossRef](#)]
4. Xiang, N.; Fackler, C. Objective Bayesian analysis in acoustics. *Acoust. Today* **2015**, *11*, 54–61.
5. Knuth, K.H.; Habeck, M.; Malakare, N.K.; Mubeen, A.M.; Placek, B. Bayesian evidence and model selection. *Digit. Signal Process.* **2015**, *47*, 50–67. [[CrossRef](#)]
6. Blandin, C.; Ozerov, A.; Vincent, E. Multi-source TDOA estimation in reverberant audio using angular spectra and clustering. *Signal Process.* **2012**, *92*, 1950–1960. [[CrossRef](#)]
7. Bush, D.; Xiang, N. A model-based Bayesian framework for sound source enumeration and direction of arrival estimation using a coprime microphone array. *J. Acoust. Soc. Am.* **2018**, *143*, 3934–3945. [[CrossRef](#)] [[PubMed](#)]
8. Escolano, J.; Xiang, N.; Perez-Lorenzo, J.M.; Cobos, M.; Lopez, J.J. A Bayesian direction-of-arrival model for an undetermined number of sources using a two-microphone array. *J. Acoust. Soc. Am.* **2014**, *135*, 742–753. [[CrossRef](#)] [[PubMed](#)]
9. Mohan, S.; Lockwood, M.E.; Kramer, M.L.; Jones, D.L. Localization of multiple acoustic sources with small arrays using a coherence test. *J. Acoust. Soc. Am.* **2008**, *123*, 2136–2147. [[CrossRef](#)] [[PubMed](#)]

10. Strutt, J.W. Investigation of the disturbance produced by a spherical obstacle on the waves of sound. *Proc. Lond. Math Soc.* **1871**, *s1-4*, 253–283. [[CrossRef](#)]
11. Lowan, A.N. Spheroidal Wave Functions. In *Handbook of Mathematical Functions With Formulas, Graphs and Mathematical Tables*; Abramowitz, M.; Stegun, I.A., Eds.; NIST: Washington, DC, USA, 1972; pp. 751–759.
12. MacRobert, T.M. *Spherical Harmonics: An Elementary Treatise on Harmonic Functions, with Applications*; Pergamon Press: Oxford, UK, 1967.
13. Madhu, N.; Martin, R. *Acoustic Source Localization with Microphone Arrays*; John Wiley & Sons Inc.: Hoboken, NJ, USA, 2008; pp. 135–166.
14. Nannuru, S.; Koochakzadeh, A.; Gemba, K.L.; Pal, P.; Gerstoft, P. Sparse Bayesian learning for beamforming using sparse linear arrays. *J. Acoust. Soc. Am.* **2018**, *144*, 2719–2729. [[CrossRef](#)] [[PubMed](#)]
15. Khaykin, D.; Rafaely, B. Acoustic analysis by spherical microphone array processing of room impulse responses. *J. Acoust. Soc. Am.* **2012**, *132*, 261–270. [[CrossRef](#)] [[PubMed](#)]
16. Morgenstern, H.; Rafaely, B. Modal smoothing for analysis of room reflections measured with spherical microphone and loudspeaker arrays. *J. Acoust. Soc. Am.* **2018**, *143*, 1008–1018. [[CrossRef](#)]
17. Nadiri, O.; Rafaely, B. Localization of Multiple Speakers under High Reverberation using a Spherical Microphone Array and the Direct-Path Dominance Test. *IEEE Acm Trans. Audio Speech Lang. Process.* **2014**, *22*, 1494–1505. [[CrossRef](#)]
18. Sun, H.; Mabande, E.; Kowalczyk, K.; Kellermann, W. Localization of distinct reflections in rooms using spherical microphone array eigenbeam processing. *J. Acoust. Soc. Am.* **2012**, *131*, 2828–2840. [[CrossRef](#)] [[PubMed](#)]
19. Li, Z.; Duraiswami, R. Hemispherical Microphone Arrays for Sound Capture and Beamforming. In Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, New Paltz, NY, USA, 16 October 2005; pp. 106–109.
20. Shabtai, N.R.; Vorländer, M. Acoustic centering of sources with high-order radiation patterns. *J. Acoust. Soc. Am.* **2015**, *137*, 1947–1961. [[CrossRef](#)] [[PubMed](#)]
21. Fernandez Grande, E. Sound field reconstruction using a spherical microphone array. *J. Acoust. Soc. Am.* **2016**, *139*, 1168–1178. [[CrossRef](#)] [[PubMed](#)]
22. Richard, A.; Fernandez Grande, E.; Brunskog, J.; Jeong, C.H. Estimation of surface impedance at oblique incidence based on sparse array processing. *J. Acoust. Soc. Am.* **2017**, *141*, 4115–4125. [[CrossRef](#)] [[PubMed](#)]
23. Zhao, S.; Dabin, M.; Cheng, E.; Qiu, X.; Burnett, I.; Liu, J.C. Mitigating wind noise with a spherical microphone array. *J. Acoust. Soc. Am.* **2018**, *144*, 3211–3220. [[CrossRef](#)] [[PubMed](#)]
24. Zuo, H.; Samarasinghe, P.N.; Abhayapala, T.D. Spatial sound intensity vectors in spherical harmonic domain. *J. Acoust. Soc. Amer. Express Letters* **2019**, *145*, EL149–EL155. [[CrossRef](#)] [[PubMed](#)]
25. Wong, K.T.; Morris, Z.N.; Nnonyele, C.J. Rules-of-thumb to design a uniform spherical array for direction finding – Its Cramér-Rao bounds’ nonlinear dependence on the number of sensors. *J. Acoust. Soc. Am.* **2019**, *145*, 714–723. [[CrossRef](#)] [[PubMed](#)]
26. Jo, B.; Choi, J.W. Direction of arrival estimation using nonsingular spherical ESPRIT. *J. Acoust. Soc. Am. Express Lett.* **2018**, *143*, EL181–EL187. [[CrossRef](#)]
27. Xenaki, A.; Fernandez-Grande, E.; Gerstoft, P. Block-sparse beamforming for spatially extended sources in a Bayesian formulation. *J. Acoust. Soc. Am.* **2016**, *140*, 1828–1838. [[CrossRef](#)] [[PubMed](#)]
28. Xiang, N.; Goggans, P.M. Evaluation of decay times in coupled spaces: Bayesian decay model selection. *J. Acoust. Soc. Am.* **2003**, *113*, 2685–2697. [[CrossRef](#)] [[PubMed](#)]
29. Beaton, D.; Xiang, N. Room acoustic modal analysis using Bayesian inference. *J. Acoust. Soc. Am.* **2017**, *141*, 4480–4493. [[CrossRef](#)] [[PubMed](#)]
30. Blauert, J.; Xiang, N. *Acoustics for Engineers—Troy Lectures*; Springer: Berlin/Heidelberg, Germany, 2009.
31. Rafaely, B. Analysis and design of spherical microphone arrays. *IEEE Trans. Speech Audio Process.* **2005**, *13*, 135–143. [[CrossRef](#)]
32. Kass, R.E.; Raftery, A.E. Bayes factors. *J. Am. Stat. Assoc.* **1995**, *90*, 773–795. [[CrossRef](#)]
33. Jeffreys, H. *Theory of Probability*, 3rd ed.; Oxford University Press: Oxford, UK, 1965; pp. 193–244.
34. Sivia, D.; Skilling, J. *Data Analysis: A Bayesian Tutorial*, 2nd ed.; Oxford University Press: New York, NY, USA, 2006; Chapter 9.
35. Skilling, J. Nested sampling. *AIP Conf. Proc.* **2004**, *735*, 395–405. [[CrossRef](#)]
36. Jaynes, E. Prior probabilities. *IEEE Trans. Inf. Theory* **1968**, *4*, 227–241. [[CrossRef](#)]

37. Shannon, C.E. The Mathematical Theory of Communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [[CrossRef](#)]
38. Gregory, P.C. *Bayesian Logical Data Analysis for the Physical Sciences*; Cambridge University Press: Cambridge, UK, 2005; pp. 184–242.
39. Candy, J. *Bayesian Signal Processing—Classical, Modern, and Particle Filtering Methods*, 2nd ed.; John Wiley & Sons Inc.: Hoboken, NJ, USA, 2016.
40. Woodward, P.M. *Probability and Information Theory, with Applications to Radar*, 2nd ed.; Pergamon Press, Ltd.: London, UK, 1953.
41. Jeffreys, H. An Invariant Form for the Prior Probability in Estimation Problems. *Proc. R. Soc. Lond. Ser. Math. Phys. Sci.* **1946**, *186*, 453–461.
42. Bretthorst, G.L. Bayesian Analysis. I. Parameter Estimation Using Quadrature NMR Models. *J. Mag. Res.* **1990**, *88*, 533–551. [[CrossRef](#)]
43. Xiang, N.; Goggans, P.M. Evaluation of decay times in coupled spaces: Bayesian parameter estimation. *J. Acoust. Soc. Am.* **2001**, *110*, 1415–1424. [[CrossRef](#)]
44. Skilling, J. Nested Sampling for General Bayesian Computation. *Bayesian Anal.* **2006**, *1*, 833–860. [[CrossRef](#)]
45. Jasa, T.; Xiang, N. Using Nested Sampling in the Analysis of Multi-rate Sound Energy Decay in Acoustically Coupled Spaces. In *AIP Conference Proceedings*; AIP: Melville, NY, USA, 2005; pp. 189–196.
46. Jasa, T.; Xiang, N. Nested sampling applied in Bayesian room-acoustics decay analysis. *J. Acoust. Soc. Am.* **2012**, *132*, 3251–3262. [[CrossRef](#)] [[PubMed](#)]
47. Fackler, C.J.; Xiang, N.; Horoshenkov, K. Bayesian acoustic analysis of multilayer porous media. *J. Acoust. Soc. Am.* **2018**, *144*, 3582–3592. [[CrossRef](#)] [[PubMed](#)]



© 2019 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 (<http://creativecommons.org/licenses/by/4.0/>).