On the angular error of intensity vector based direction of arrival estimation in reverberant sound fields

Dovid Levin
School of Engineering, Bar-Ilan University, Ramat-Gan 52900, Israel

Emanuël A. P. Habets[a)]
Department of Electrical and Electronic Engineering, Imperial College London, Exhibition Road, SW7 2AZ London, United Kingdom

Sharon Gannot
School of Engineering, Bar-Ilan University, Ramat-Gan 52900, Israel

(Received 30 April 2010; revised 20 July 2010; accepted 24 July 2010)

An acoustic vector sensor provides measurements of both the pressure and particle velocity of a sound field in which it is placed. These measurements are vectorial in nature and can be used for the purpose of source localization. A straightforward approach towards determining the direction of arrival (DOA) utilizes the acoustic intensity vector, which is the product of pressure and particle velocity. The accuracy of an intensity vector based DOA estimator in the presence of noise has been analyzed previously. In this paper, the effects of reverberation upon the accuracy of such a DOA estimator are examined. It is shown that particular realizations of reverberation differ from an ideal isotropically diffuse field, and induce an estimation bias which is dependant upon the room impulse responses (RIRs). The limited knowledge available pertaining the RIRs is expressed statistically by employing the diffuse qualities of reverberation to extend Polack's statistical RIR model. Expressions for evaluating the typical bias magnitude as well as its probability distribution are derived. © *2010 Acoustical Society of America.* [DOI: 10.1121/1.3479542]

PACS number(s): 43.60.Jn [EJS] Pages: 1800–1811

I. INTRODUCTION

Estimation of direction of arrival (DOA) pertaining to an acoustic source is typically preformed by an array of microphones. The two common approaches are based either on beam steering or time delay estimation.[1] Possible applications of DOA estimation include source tracking for navigation, camera steering, and beamform directing. Further information regarding the confidence level of an estimator may be used to design robust beamformers[2] or to regulate the extent of zoom for a steered camera.

Sound fields consist of pressure and particle velocity components, the former being a scalar and the later a vector quantity. Conventional microphones provide measurements of the pressure sound field but do not measure particle velocity. While it is possible to indirectly estimate particle velocity by use of differential microphones (which consist of closely spaced sensors) the resulting output remains an estimate and is not a precise measurement.

For some time, acoustic vector sensors, capable of measuring all sound field components, have become available in the form of hydrophones used in underwater acoustics applications.[3] The additional availability of velocity measurements brings about several advantages including the possibility of constructing high performance sonar arrays which occupy less space than comparable scalar arrays.[4] More recently, a miniature aero-acoustic vector sensor has been developed[5] providing the opportunity to apply a new source of information to the field of speech processing. This is a relatively new area of research and its potential is for the most part as of yet unexplored.

Nehorai and Paldi[6] dealt with the use of vector sensors for estimating DOA. Their analysis utilized a model based upon far-field approximation and in which the source signal as well as additive measurement errors were either narrowband or white. Additionally, all source and error components were assumed to be independent, identically distributed (i.i.d.) and mutually independent of each other. They proposed an intensity based DOA estimator (utilizing the pressure and particle velocity product) for a single vector-sensor, showed that this estimator is unbiased and consistent and derived a term describing the rate of convergence towards the correct DOA as a function of data length. It was also shown that the additive error caused by ambient noise which is ideally diffuse and isotropic can prescribe to the above mentioned model assumptions[7] and thus an intensity vector DOA estimator situated in such an environment produces similar results.

In this article, the effects of reverberation on the intensity DOA estimator are analyzed. One possible method often used to model reverberation is to view it as an ideally isotropic diffuse field. This model necessarily anticipates unbiased DOA estimation, by virtue of its isotropy. It is claimed in the present article that in practice, the intensity-based estimator in a reverberant environment produces biased results (i.e. the expected result of the estimator provides a DOA which differs from the true value). This bias stems from the

[a)]Author to whom correspondence should be addressed. Electronic mail: e.habets@imperial.ac.uk.

fact that particular realizations of reverberation inevitably deviate from ideal isotropic diffuse behavior. The isotropic model is, however, valid when relating to the spatial average of the reverberant sound field as observed by Jacobsen and Roisin.[8] This article aims to demonstrate and analyze the aforementioned deviation of the reverberation from the ideal model.

In this study, estimator bias for a given setup is shown to be dependent on the associated room impulse responses (RIRs). Although RIRs are essentially deterministic (being a function of environment and source-receiver locations), their precise nature is generally concealed from users. The limited knowledge concerning an environment's behavior can be expressed in the form of a statistical model.

Using such a statistical approach, we proceed to quantify the DOA bias. This is done by extending Polack's statistical model[9,10] (which relates to an RIR as a white Gaussian process multiplied by an exponentially decaying envelope) to incorporate particle velocity. The estimator deviation is conveniently measured in units of angular error (AE) (as in Ref. 6) which is subsequently defined. The extended model leads towards the derivation of an estimate of the AE distribution, expected value and standard deviation for an arbitrarily long time-sequence of sensor measurements. These results may be used to determine confidence levels for estimating DOA without exceeding a specified degree of error.

The paper is organized as follows. In Section II the estimation problem and its related notations are formally defined. Nehorai and Paldi's intensity based method is introduced and placed in the perspective of the reverberant environment. Section III analyzes the behavior of the intensity method for a given RIR, and introduces the *intensity response* function which is used to express the expected intensity vector. Section IV develops an extension of Polack's work that statistically models both pressure and particle velocity RIRs. Using this extended model, the reverberant part of the intensity response can be viewed as a random walk in three dimensions. In Section V, the results of the previous sections are used to derive the probabilistic distribution of the estimator's AE. Approximations are made in order to reduce it into a tractable form. In Section VI, we evaluate the proposed theory and the validity of its associated approximations. Section VII concludes with a brief summary of the paper.

II. PROBLEM FORMULATION AND PRELIMINARIES

The goal of this paper is to assess the performance of an intensity vector based algorithm for determining DOA in a reverberant environment. In this section we formulate the problem and discuss the intensity-vector method for DOA estimation in (i) an anechoic and noisy environment and (ii) a reverberant environment.

A single vector-sensor measures the pressure and particle velocity of a sound field at a given point in space. These measurements, taken at time n, are denoted respectively by $y_p[n]$ and $\mathbf{y}_v[n]=[y_{v_x}[n]y_{v_y}[n]y_{v_z}[n]]^T$ that consists of three orthogonal particle velocity components. Each measurement is comprised of the correct physical measurement of the acoustical sound field and possibly an undesired additive component resulting from internal sensor-noise. These measurements can then be decomposed as:

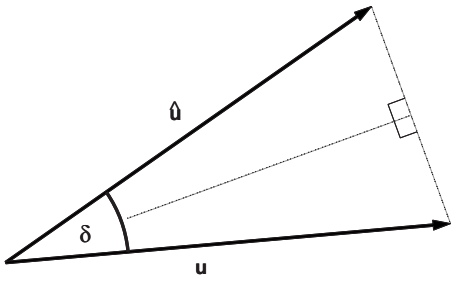

FIG. 1. A portrayal of the DOA $\mathbf{u}$, its estimate $\hat{\mathbf{u}}$, and their separating angle δ.

$$\begin{bmatrix} y_p[n] \\ \mathbf{y}_v[n] \end{bmatrix} = \begin{bmatrix} p[n] \\ \mathbf{v}[n] \end{bmatrix} + \begin{bmatrix} n_p[n] \\ \mathbf{n}_v[n] \end{bmatrix}, \tag{1}$$

where $p[n]$ and $\mathbf{v}[n]=[v_x[n]v_y[n]v_z[n]]^T$ represent the uncorrupted pressure and particle velocity measurements, and $n_p[n]$ and $\mathbf{n}_v[n]$ represent the corresponding noise components. The former two, when viewed in tandem, contain information about the DOA which is described as a unit vector $\mathbf{u}$ pointing from the receiver towards the direction of the source. Based on the corrupted measurements ($y_p[n]$ and $\mathbf{y}_v[n]$), a DOA estimate denoted $\hat{\mathbf{u}}$ is produced (the latter also being a unit vector). The accuracy of this estimate can be evaluated by the AE (see Refs. 6 and 11) which is defined as the angle δ by which $\hat{\mathbf{u}}$ deviates from $\mathbf{u}$, or more formally:

$$\delta \equiv 2\sin^{-1}\left(\frac{\|\hat{\mathbf{u}}-\mathbf{u}\|}{2}\right). \tag{2}$$

The relationship between vectors $\mathbf{u}$ and $\hat{\mathbf{u}}$, and angle δ is illustrated in Fig. 1, and may clarify the trigonometric associations of (2).

In general, the sound field at the sensor's location can be described as a source signal $s[n]$ which undergoes propagation and possibly reverberation, and is then compounded with an ambient sound field resulting from other sound sources within the environment which are independent of $s[n]$. As reverberation is linear and time invariant, its effects may be viewed as filtering by a set of RIRs. The entire process can be described as:

$$\begin{bmatrix} p[n] \\ v_x[n] \\ v_y[n] \\ v_z[n] \end{bmatrix} = \begin{bmatrix} (s*h_p)[n] \\ (s*h_{v_x})[n] \\ (s*h_{v_y})[n] \\ (s*h_{v_z})[n] \end{bmatrix} + \begin{bmatrix} a_p[n] \\ a_{v_x}[n] \\ a_{v_y}[n] \\ a_{v_z}[n] \end{bmatrix}, \tag{3}$$

or expressed more compactly: $[p[n]\mathbf{v}^T[n]]^T=(s*\mathbf{h})[n]+\mathbf{a}[n]$. The vector $\mathbf{a}[n]=[a_p[n]a_{v_x}[n]a_{v_y}[n]a_{v_z}[n]]^T$ describes the ambient sound field and the vector $\mathbf{h}[n]=[h_p[n]h_{v_x}[n]h_{v_y}[n]h_{v_z}[n]]^T$ consists of the source impulse-responses pertaining to each sound field component measured. This can be further divided into parts corresponding to the direct path (denoted by $\mathbf{h}_d[n]$) and to the reverberative reflections (denoted by $\mathbf{h}_r[n]$), such that $\mathbf{h}[n]=\mathbf{h}_d[n]+\mathbf{h}_r[n]$. Summing up and taking into account sensor-noise, we have:

$$\mathbf{y}[n] = (s * (\mathbf{h}_d + \mathbf{h}_r))[n] + \mathbf{a}[n] + \mathbf{n}[n]. \quad (4)$$

The measurements $\mathbf{y}[n]$ can be split into the desired component resulting from the direct sound $\mathbf{y}_d[n]=(s*h_d)[n]$, and the remainder:

$$\mathbf{e}[n] = (s * \mathbf{h}_r)[n] + \mathbf{a}[n] + \mathbf{n}[n], \quad (5)$$

which can contribute to errors in the estimation of the DOA, and is refereed to as *interference components*. Thus, $\mathbf{y}[n]=\mathbf{y}_d[n]+\mathbf{e}[n]$. Hereafter, the pressure and particle velocity components of the direct part are denoted $p_d[n]$ and $\mathbf{v}_d[n]$ respectively with corresponding interference terms $e_p[n]$ and $\mathbf{e}_v[n]$, i.e. $\mathbf{y}_d[n]=[p_d[n]\mathbf{v}_d^T[n]]^T$ and $\mathbf{e}[n]=[e_p[n]\mathbf{e}_v^T[n]]^T$. Also, when referring to the entire RIR ensemble as opposed to particular coefficients, time dependency is dropped (i.e. $\mathbf{h}$ is used instead of $\mathbf{h}[n]$).

It should be noted that we have expanded the idea of RIRs, a technique traditionally used in conjunction with acoustic pressure, to describe particle velocity. The same reasons justifying their use for acoustic pressure, namely the linear and time-invariant properties of the wave equation, apply equally as well to particle velocity.

A. The intensity vector method for DOA estimation

Nehorai and Paldi[6] proposed an intensity vector based algorithm for estimating the DOA. Continuous-time intensity $\mathbf{i}(\mathbf{r},t)$ is the product of pressure and particle velocity:

$$\mathbf{i}(\mathbf{r},t) = p(\mathbf{r},t) \cdot \mathbf{v}(\mathbf{r},t), \quad (6)$$

(with $\mathbf{r}$ referring to spatial location and t to continuous-time). This vector corresponds to the magnitude and *direction* of the transport of acoustical energy,[12] indicating its utility for determining DOA.

Nehorai and Paldi assume that the signal behaves as a plane wave at the sensor location. Euler's linear equation $\rho_0(\partial/\partial t)\mathbf{v}(\mathbf{r},t)=-\Delta p(\mathbf{r},t)$ implies that a plane-wave maintains the following equality:

$$\mathbf{v}(\mathbf{r},t) = \frac{-1}{\rho_0 c} p(\mathbf{r},t)\mathbf{u}, \quad (7)$$

(ρ_0 being the ambient air density and c the speed of sound). Substituting the above into (6) yields an intensity vector $\mathbf{i}(\mathbf{r},t)=-p^2(\mathbf{r},t)/\rho_0 c\ \mathbf{u}$ in the opposite direction of the DOA. In order to eliminate the need to carry a constant coefficient throughout derivations, discrete-time particle velocity is taken as scaled: $\mathbf{v}[n]=-\rho_0 c\mathbf{v}(\mathbf{r}_0,nT_s)$ (with T_s representing sampling-period and $\mathbf{r}_0$ the sensor location), such that:

$$\mathbf{i}[n] = p[n]\mathbf{v}[n] = p^2[n]\mathbf{u}. \quad (8)$$

Sampling of pressure measurements is conducted without scaling in an analogous fashion: $p[n]=p(\mathbf{r}_0,nT_s)$.

The measured intensity vector $\mathbf{y}_i[n]=y_p[n]\mathbf{y}_v[n]$ is not a completely reliable indicator of DOA, since each of its constituent parts is the sum of a direct sound component and an interference component: $y_p[n]=p_d[n]+e_p[n]$ and $\mathbf{y}_v[n]=\mathbf{v}_d[n]+\mathbf{e}_v[n]$. As a result intensity vector is:

$$\mathbf{y}_i[n] = p_d^2[n]\mathbf{u} + (e_p[n]\mathbf{u} + \mathbf{e}_v[n])p_d[n] + e_p[n]\mathbf{e}_v[n]. \quad (9)$$

The first summand $p_d^2[n]\mathbf{u}$ contains useful information about DOA, whereas the rest result from interference terms and may lead to estimation error. In order to mitigate their effect upon estimation of $\mathbf{u}$, the intensity vector is averaged over a number of samples, and then $\hat{\mathbf{u}}$ is produced by normalization:

$$\bar{\mathbf{y}}_i = \frac{1}{N}\sum_{n=1}^{N} \mathbf{y}_i[n] = \frac{1}{N}\sum_{n=1}^{N} y_p[n]\mathbf{y}_v[n], \quad (10a)$$

$$\hat{\mathbf{u}} = \frac{\bar{\mathbf{y}}_i}{\|\bar{\mathbf{y}}_i\|}. \quad (10b)$$

In analyzing the performance of the above algorithm in an *anechoic* environment, Nehorai and Paldi made several assumptions regarding signal characteristics. These include: processes $s[n]$, $e_p[n]$, $e_{v_x}[n]$, $e_{v_y}[n]$ and $e_{v_z}[n]$ are all i.i.d. with zero mean and each of these processes is mutually uncorrelated with the others. Also, $e_{v_x}[n]$, $e_{v_y}[n]$ and $e_{v_z}[n]$ are assumed to have identical variance. They showed that for a broad range of conditions that the DOA estimator $\hat{\mathbf{u}}$ is *unbiased* and *consistent*—converging almost surely (i.e. with probability 1) to the true value of $\mathbf{u}$ as $N\to\infty$. Also, an asymptomatic measure of the rate at which the angular error δ converges to zero was derived.

B. Intensity vector method in a reverberant environment

It is appropriate to determine which scenarios prescribe to the assumptions made about processes $[e_p[n]\mathbf{e}_v^T[n]]^T=\mathbf{e}[n]$, such that the above results are valid. Referring to (5), the *desired* information which pertains to the direct arrival is contained in $(s*\mathbf{h}_d)[n]$, whereas the interference component $\mathbf{e}[n]$ consists of $(s*\mathbf{h}_r)[n]$ (reverberation), $\mathbf{a}[n]$ (ambient sound) and $\mathbf{n}[n]$ (sensor noise). Internal sensor-noise can comply to the above assumptions[6] (in particular, it is natural to assume that noise pertaining to pressure and the various particle velocity elements are mutually independent since they are created at different sensor components). Spherically isotropic ambient sound fields may also conform to these assumptions.[7] Hence, the analysis of Nehorai and Paldi can fit a scenario in which error is induced by sensor noise and ambient sound: $\mathbf{e}[n]=\mathbf{n}[n]+\mathbf{a}[n]$; and reverberation is not addressed i.e. $\mathbf{h}_r[n]=\mathbf{0}$.

In this paper, we inspect a scenario in which the only cause of error is reverberation i.e. $\mathbf{e}[n]=(s*\mathbf{h}_r)[n]$, and in which sensor noise and ambient sound are assumed to be absent, namely $\mathbf{n}[n]=\mathbf{a}[n]=\mathbf{0}$. Hence, (4) now becomes: $\mathbf{y}[n]=(s*\mathbf{h})[n]$, or more explicitly:

$$\begin{bmatrix} y_p[n] \\ \mathbf{y}_v[n] \end{bmatrix} = \begin{bmatrix} (s*h_p)[n] \\ (s*\mathbf{h}_v)[n] \end{bmatrix}. \quad (11)$$

As before, the source signal $s[n]$ is assumed to be zero mean i.i.d. and its variance is denoted by σ_s^2. It can no longer be assumed that $\mathbf{e}[n]$ is i.i.d. as the filter $\mathbf{h}_r[n]$ induces coloration. Furthermore, elements of vector $\mathbf{e}[n]$ are no longer mutually-independent as they are all generated from the same signal $s[n]$. In particular, it can no longer be taken for

granted that $E\{e_p[n]\mathbf{e}_v[n]\}=\mathbf{0}$, a fact which may introduce a bias into the estimator.

III. ANALYSIS OF INTENSITY ALGORITHM FOR A GIVEN RIR

In this section, we evaluate the expected outcome of the intensity based estimator, assuming that the relevant RIR $\mathbf{h}[n]=[h_p[n]\mathbf{h}_v^T[n]]^T$ is given for all n. We start by inspecting the time-averaged measured intensity. Substitution of (11) into (10a) produces:

$$\bar{\mathbf{y}}_i=\frac{1}{N}\sum_{n=1}^{N}(s*h_p)[n]\begin{bmatrix}(s*h_{v_x})[n]\\(s*h_{v_y})[n]\\(s*h_{v_z})[n]\end{bmatrix}. \tag{12}$$

Now since time instants of the source signal $s[n]$ are assumed to have identical distributions (in addition to being stationary), so will the measured intensity samples $\mathbf{y}_i[n]=y_p[n]\mathbf{y}_v[n]$ derived therefrom. As a result, the expected value (conditional upon the entire RIR ensemble $\mathbf{h}$) of the intensity average (12) is equivalent to that of any single intensity sample (regardless of the value of n):

$$E\{\bar{\mathbf{y}}_i|\mathbf{h}\}=E\{y_p[n]\mathbf{y}_v[n]|\mathbf{h}\}=E\{(s*h_p)[n](s*\mathbf{h}_v)[n]|\mathbf{h}\}. \tag{13}$$

Writing out the convolutions explicitly yields:

$$\begin{aligned}E\{\bar{\mathbf{y}}_i|\mathbf{h}\}&=E\{\mathbf{y}_i[n]|\mathbf{h}\}\\&=E\left\{\sum_m h_p[m]s[n-m]\sum_{m'}\mathbf{h}_v[m']s[n-m']|\mathbf{h}\right\}\\&=\sum_m\sum_{m'}E\{s[n-m]s[n-m']\}h_p[m]\mathbf{h}_v[m']\\&=\sigma_s^2\sum_m h_p[m]\mathbf{h}_v[m],\end{aligned} \tag{14}$$

with the last stage following from $E\{s[n-m]s[n-m']\}=\sigma_s^2\delta[m-m']$. The result is indeed independent of n.

The product of pressure and velocity RIRs is called the *intensity response* and is denoted:

$$\mathbf{h}_i[n]\equiv h_p[n]\mathbf{h}_v[n]. \tag{15}$$

We adopt the notation that the direct arrival corresponds to $\mathbf{h}[-1]$ and the reflections correspond to $\mathbf{h}[n\geq 0]$.[13] Then, the expected intensity vector can be represented as: $E\{\bar{\mathbf{y}}_i|\mathbf{h}\}=\sigma_s^2(\mathbf{h}_i[-1]+\Sigma_{n=0}^{\infty}\mathbf{h}_i[n])$.

The term $\Sigma_{n=0}^{\infty}\mathbf{h}_i[n]$, corresponding to reverberation, proves to be central in our performance analysis. Here a more compact notation is used by introducing the *intensity accumulation vector* (IAV) which is inspired by the commonly used energy decay curve (EDC) (based on Schroeder[14]). The discrete-time EDC which is defined as:

$$\mathrm{EDC}[n]=\sum_{m=n}^{\infty}h_p^2[m], \tag{16}$$

describes the energy contained in a pressure RIR from time-instant n onwards. In a similar vein, the IAV is defined:

$$\mathbf{IAV}[n]=\sum_{m=n}^{\infty}\mathbf{h}_i[m]. \tag{17}$$

This vector depicts the total intensity accumulated in the intensity response, from moment n onwards. As such, the reverberation term can be expressed: $\Sigma_{n=0}^{\infty}\mathbf{h}_i[n]=\mathbf{IAV}[0]$.

The expected intensity measurement can now be presented as:

$$E\{\bar{\mathbf{y}}_i|\mathbf{h}\}=\sigma_s^2(\mathbf{h}_i[-1]+\mathbf{IAV}[0]). \tag{18}$$

This expression is significant since as $N\rightarrow\infty$, the intensity estimator $\bar{\mathbf{y}}_i$ converges *in probability (i.p.)* to that value. Moreover, the asymptotic value *(i.p.)* of the DOA estimator $\hat{\mathbf{u}}$ (which is hence referred to as $\hat{\mathbf{u}}_{\mathrm{asym}}$) is produced by normalization of (18). (See Appendix A for a discussion on asymptotic convergence.) This is useful for deriving the asymptotic angular error. Examining the consequences of the terms in (18) is thus beneficial.

The signal variance σ_s^2 is a positive scalar that affects the amplitude of the expected intensity but *not* its orientation. Therefore, it has no bearing upon the DOA estimation problem. The term $\mathbf{h}_i[-1]$ corresponds to the direct sound and points towards the true DOA, whereas $\mathbf{IAV}[0]$ results from reverberation which is an interference component. The latter is oriented in some arbitrary orientation which can cause error. Summation of these two terms produces the expected intensity vector (up to a constant). Consequently, the intensity response $\mathbf{h}_i$ contains information about the orientation of both $\mathbf{u}$ (i.e. $\mathbf{h}_i[-1]$) and $\hat{\mathbf{u}}$ at its asymptotic value (i.e. $\mathbf{IAV}[-1]$), and thus determines the asymptotic angular error.

IV. EXTENSION OF POLACK'S STATISTICAL MODEL FOR RIRS AND ITS APPLICATION

Essentially, it has been shown that the RIRs and particularly $\mathbf{h}_i[n]$ derived therefrom uniquely determine the behavior of the intensity vector estimator. In practice this fact cannot be used *directly* since the RIRs are generally unknown. In fact, if they were to be known then the true DOA could be extracted from $\mathbf{h}_i[-1]$ and no estimation would be necessary. However, even though RIRs are not known, some general properties of their behavior (e.g. decay rate) may be available.

Statistical room acoustics (SRA) is used to systematically describe reverberation in terms of these properties by use of probabilistic tools. The field was founded by Schroeder[15] who discussed the frequency domain and later was extended by Polack[9] to the time domain. Strictly speaking, reverberation may be viewed as a deterministic process being governed by the wave equation and the physical boundary conditions of its environment. Nonetheless, these processes are extremely complex and may be more conveniently described as random, or more precisely *pseudorandom* (since the environment and source-receiver positions are actually fixed). In this section, SRA is applied to aid analysis of DOA estimation.

A. Polack's statistical model for pressure RIR

The reflections reaching a sensor as a result of an impulse excitation arrive with rapidly increasing density as time progresses. For example, it is known that in a specularly reflecting rectangular room, the total number of reflections arriving increases in proportion to cube of t. When a sufficient number of reflections arrive during the course of an observation interval with varying phases and amplitudes, then their cumulative effect can be viewed (by use of the central limit theorem) as forming a Gaussian variable. This idea was employed by Polack[9] to statically model the pressure RIR (in his thesis written in French; see Jot *et al.*[10] for a good survey in English).

The RIR of the pressure component for a given source-receiver pair is modeled as a Gaussian random process multiplied by a decaying envelope:

$$h_p[n] = u[n]\sigma_0 e^{-\alpha n} w(n), \tag{19a}$$

with $w[n]$ characterized by:

$$E\{w[n]\} = 0,$$

$$E\{w[n_1]w[n_2]\} = \delta[n_1 - n_2], \tag{19b}$$

and $u[n]$ denoting the unit step function. The envelope decay parameter α is linked to the reverberation time RT_{60} by the following relationship:[10]

$$\mathrm{RT}_{60} = \frac{3\ \ln\ 10}{\alpha f_s}, \tag{20}$$

with $f_s = 1/T_s$ denoting the sampling frequency.

Since an RIR is said to be a pseudorandom process, one can relate to concepts such as the expected value and variance of functions of $\mathbf{h}$. However, being that the RIR is in actuality deterministic, these operations require some form of interpretation. If a room's impulse response is unknown, then the statistical model provides a description of the level of uncertainty pertaining to possible outcomes. Measurements conducted in a room with a specific source-receiver setup correspond to one particular realization of the process. To obtain a different realization, it is necessary to alter the measurement setup (e.g. alteration of source and/or receiver positions) in such a way that the underlying parameters (α and σ_0) are unchanged. The expectation operation averages over the support set of all realizations resulting from possible setups in the environment. For clarity, such *spatial* expectation and variance are henceforth denoted as $E_s\{\circ\}$ and $\mathrm{Var}_s\{\circ\}$.

It should be noted that the validity of Polack's model is based on the assumption that many reflections arrive almost simultaneously. It therefore accounts for the tail of the RIR but does not describe the direct arrival or early reflections which are of sparse temporal distribution.

B. Extension based upon properties of diffuse sound field

Polack's model describes the pressure component $h_p[n]$ of an RIR. For our purposes, it is necessary to characterize all components of $\mathbf{h}[n] = [h_p[n] h_{v_x}[n] h_{v_y}[n] h_{v_z}[n]]^T$. To do this, we extend Polack's model through utilization of the similarity between reverberation and a diffuse-field. This analogy has proved quite useful in the field of room acoustics and provides for example a basis for Sabine's famous formula.[16,17] A purely diffuse field is defined as one in which mutually independent sound waves pertaining to identically distributed random processes propagate evenly from all directions (i.e. the field is isotropic and homogeneous).

We start out by using the image-source model[18] to spatially depict the generation of a reverberative sound field. Image point-sources are located in space at positions corresponding to the virtual origins of the various reflections. After an impulse excitation, the reflections arriving at the receiver between time t and $t+T_s$ can be viewed as emanating from those virtual sources that are located within the space of a spherical shell of thickness T_s/c which has a radius length of t/c and is centered at the receiver position. Effectively, a coefficient of $\mathbf{h}[n]$ is determined by a multitude of sound waves arriving from all directions of the sphere shell, in a fashion similar to that of the diffuse field.

Two objections may be raised to this comparison. First of all, the images confined within the spherical shell are all point sources whose locations correspond to a finite number of spatial directions and by no means encompass *all* directions. Secondly, it has been shown that the nonspecular geometric acoustic model which underlies the image method predicts that decay rate of sound waves is direction dependant.[16] These factors render the field nonisotropic.

Resolution of these issues is not overly difficult. With regards to the first issue, densely distributed sources can be seen as approximating a source which is continuously distributed in space, to a high degree of accuracy. In fact this technique has been successfully used to simulate a diffuse fields.[19] Furthermore, nonspecularly reflecting walls tend to scatter sound waves in *all* directions. This trait of nonspecular reflections also causes the field to approach isotropic behavior. Moreover, the presence of obstacles as well as irregularities in the shape of the room would *mix* the sound waves and induce isotropic behavior.

The diffusive behavior exhibited by a reverberant sound field can be used to model the associated RIRs. It has been shown[7,8] that in an ideally diffuse sound field, each of the four vector sensor signals can be shown to be mutually uncorrelated, with the particle velocity components maintaining a variance equal to one third of the variance of the pressure component. This suggests extending Polack's work to model RIRs comprising $\mathbf{h}[n]$ as *four* i.i.d. Gaussian processes $[w_1[n] w_2[n] w_3[n] w_4[n]]^T = \mathbf{w}[n]$, which are multiplied by an appropriate envelope:

$$\begin{bmatrix} h_p[n] \\ h_{v_x}[n] \\ h_{v_y}[n] \\ h_{v_z}[n] \end{bmatrix} = u[n]\sigma_0 e^{-\alpha n} \begin{bmatrix} w_1[n] \\ 1/\sqrt{3} w_2[n] \\ 1/\sqrt{3} w_3[n] \\ 1/\sqrt{3} w_4[n] \end{bmatrix}. \tag{21a}$$

The processes are jointly distributed:

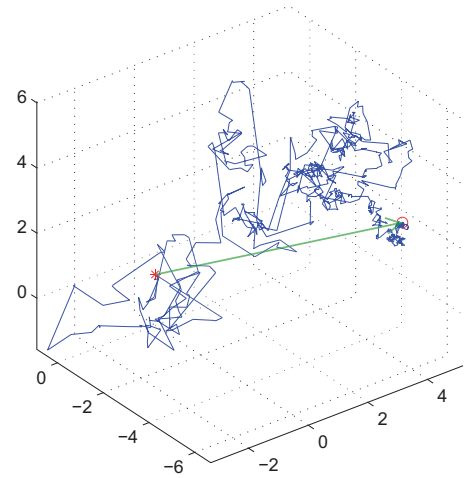

FIG. 2. (Color online) This random walk in three dimensions corresponds to the summation of 8000 terms from $\mathbf{h}_i[n]$. The walk commences at the origin (*) eventually arriving at ($\circ$). It corresponds to the reverberant part of $\mathbf{h}_i[n]$ and does not include the direct arrival.

$$\mathbf{w}[n] \sim \mathcal{N}(\mathbf{0},\mathbf{I}), \tag{21b}$$

with $\mathbf{0}$ being a zero vector and $\mathbf{I}$ the identity matrix. This proposed model is useful for analysis of the statistical behavior of the intensity vector as is demonstrated in the sequel.

C. Reverberation intensity vector as a random walk

We wish to calculate the statistical distribution of the interference component in (18), $\mathbf{IAV}[0]$ (or more generally $\mathbf{IAV}[n]=\Sigma_{m=n}^{\infty}\mathbf{h}_i[m]$). Any particular summand $\mathbf{h}_i[m]$ consists of the product $\mathbf{h}_i[m]=h_p[m]\mathbf{h}_v[m]$ in which $\mathbf{h}_v[m]$ is a spherically symmetrical Gaussian vector in three dimensions [as evident from (21a) and (21b)]. This implies that $\mathbf{h}_i[m]$ must also be distributed in a spherically symmetrical fashion (see Appendix B). The sum of independent random vectors possessing the property of spherical-symmetry retain this property. Hence, $\mathbf{IAV}[n]$ which is the sum of such random vectors has a spherically symmetrical distribution.

It is enlightening to view the summation of random variables comprising the reverberant part of the IAV (i.e. variables $\mathbf{h}_i[n\geq 0]$) as a random walk in three dimensions. Each step—being a time instant of $\mathbf{h}_i[n]$—consists of a single random vector. The typical step size is not constant but rather diminishes exponentially with n as the reverberation decays. The walk commences at the origin [0,0,0] with the size and direction of the first step corresponding to $\mathbf{h}_i[0]$, the next step to $\mathbf{h}_i[1]$ and so on. In Fig. 2, a realization of such a three dimensional random walk is shown. The initial position is marked by an asterisk and the position after 8000 steps by a circle. The aggregate result of summing all steps is delineated by the vector connecting these two points. Note that the typical step size becomes progressively smaller.

In order to establish the joint distribution of $\mathbf{IAV}[n]=[\mathrm{IAV}_x[n]\mathrm{IAV}_y[n]\mathrm{IAV}_z[n]]^T$, we first determine the distribution of its individual elements. We shall start by inspecting the x-component: $\mathrm{IAV}_x[n]=\Sigma_{m=n}^{\infty}h_{\mathrm{i}_x}[m]$. The summand term $h_{\mathrm{i}_x}[m]=h_p[m]h_{\mathrm{v}_x}[m]$ can be written explicitly [from (21a)] as:

$$h_{\mathrm{i}_x}[m] = u[m]\frac{1}{\sqrt{3}}\sigma_0^2 e^{-2\alpha m} w_1[m]w_2[m]. \tag{22}$$

The mean and variance are then:

$$E_{\mathrm{s}}\{h_{\mathrm{i}_x}[m]\} = 0, \tag{23a}$$

$$\mathrm{Var}_{\mathrm{s}}\{h_{\mathrm{i}_x}[m]\} = \tfrac{1}{3}\sigma_0^4 e^{-4\alpha m}u[m]. \tag{23b}$$

This is computed with aid of the knowledge that the random variables (r.v.s) $w_1[m]$ and $w_2[m]$ are independent with standard normal distributions. As such, $E_{\mathrm{s}}\{w_1[m]w_2[m]\}=E_{\mathrm{s}}\{w_1[m]\}E_{\mathrm{s}}\{w_2[m]\}=0$ and $E_{\mathrm{s}}\{(w_1[m]w_2[m])^2\}=E_{\mathrm{s}}\{w_1^2[m]\}E_{\mathrm{s}}\{w_2^2[m]\}=1$. The distribution of an r.v. created in a fashion such as $h_{\mathrm{i}_x}[m]$ is known as the *normal product* distribution.

We are interested in the distribution of $\mathrm{IAV}_x[n]$ which is formed by summing over the independent r.v.s $h_{\mathrm{i}_x}[m]$. Even though the normal product distribution is unwieldy, one could expect that summing many such variables would yield the tamer Gaussian distribution. However, the *central limit theorem* which forms the basis for this conception is not directly applicable since the i.i.d. condition is not met. Specifically, the distributions of $h_{\mathrm{i}_x}[m]$ are not identical as the envelope $\frac{1}{3}\sigma_0^4 e^{-4\alpha m}$ which scales the normal product variable $w_1[m]w_2[m]$ is time dependent.

This obstacle may be overcome by use of the scheme which follows. The random sequence $h_{\mathrm{i}_x}[m]$ is divided into partitions $\mathcal{P}_k$ of length L as follows: $\mathcal{P}_k=\{h_{\mathrm{i}_x}[n+kL], h_{\mathrm{i}_x}[n+kL+1],\ldots,h_{\mathrm{i}_x}[n+kL+(L-1)]\}$. Assuming that the decay rate is negligible within the a partition interval (i.e. $e^{-4\alpha L}\approx 1$), then all random variables within the interval maintain *approximately* identical distributions. Now, if L is sufficiently large, then the central limit theorem implies that the sum of all variables within the interval, $S_k[n]=\Sigma_{m=n+kL}^{n+kL+(L-1)}h_{\mathrm{i}_x}[m]$, is approximately Gaussian distributed. This being the case, the random variable $\mathrm{IAV}_x[n]$ may also be considered approximately Gaussian since it is the sum of independent Gaussians: $\mathrm{IAV}_x[n]=\Sigma_{k=0}^{\infty}S_k[n]$. It should be noted that the assumptions made regarding α and L imply that the reverberation time is not overly short.

The mean and variance of $\mathrm{IAV}_x[n]$ (considering (23a) and (23b)) as well as the independence of individual samples from $h_{\mathrm{i}_x}[m]$) are then:

$$E_{\mathrm{s}}\{\mathrm{IAV}_x[n]\} = \sum_{m=n}^{\infty} E_{\mathrm{s}}\{h_{I_x}[m]\} = 0, \tag{24a}$$

$$\mathrm{Var}_{\mathrm{s}}\{\mathrm{IAV}_x[n]\} = \sum_{m=n}^{\infty} \mathrm{Var}_{\mathrm{s}}\{h_{I_x}[m]\} = \frac{1}{3}\sigma_0^4\frac{e^{-4\alpha n}}{1-e^{-4\alpha}}, \tag{24b}$$

which completely characterize its Gaussian distribution. The distribution derived for $\mathrm{IAV}_x[n]$ apply equally to $\mathrm{IAV}_y[n]$ and $\mathrm{IAV}_z[n]$ as well, hence the spatial components of $\mathbf{IAV}[n]$ are identically distributed. Furthermore, they are uncorrelated [as can be shown from (21a) and (21b)] and are spherically symmetrical. Summing up, spatial components of $\mathbf{IAV}[n]$ are i.i.d. Gaussian variables with zero mean and standard deviation:

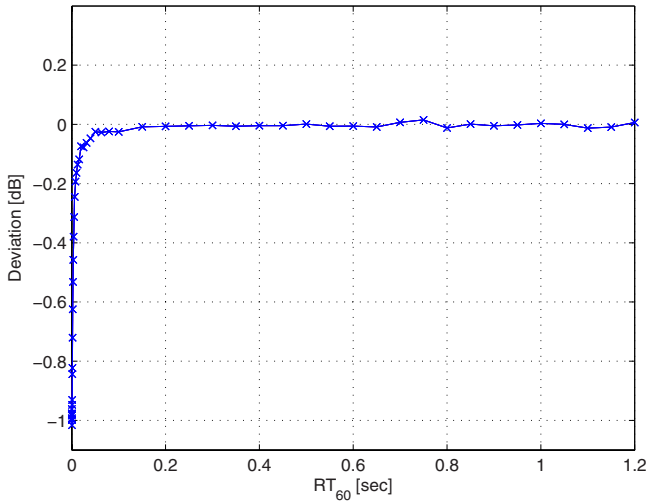

FIG. 3. (Color online) Deviation of theoretical value of $E\{\mathbf{IAV}[0]\}$ from the sample mean obtained by Monte Carlo simulation results, as a function of the RT_{60}.

$$\sigma_{\mathrm{IAV}}[n] = \frac{\sigma_0^2}{\sqrt{3(1-e^{-4\alpha})}} e^{-2\alpha n}. \tag{25}$$

The results derived in this section can be interpreted in the following fashion: the orientation of $\mathbf{IAV}[n]$ is uniformly distributed along all possible directions. This follows from the property of spherical symmetry. The length of the random vector (i.e. its Euclidian norm) is:

$$\|\mathbf{IAV}[n]\| = \sqrt{\mathrm{IAV}_x^2[n] + \mathrm{IAV}_y^2[n] + \mathrm{IAV}_z^2[n]}. \tag{26}$$

The distribution of the Euclidian norm of a vector whose components are Gaussian i.i.d. with zero mean and standard deviation σ_G is known as a Maxwell[20,26,27] distribution (or third order Chi distribution):

$$f_Q(q) = u(q)\frac{1}{\sigma_G^3}\sqrt{\frac{2}{\pi}} q^2 e^{-q^2/2\sigma_G^2}, \tag{27}$$

and has an expected value of:

$$E\{Q\} = 2\sigma_G\sqrt{\frac{2}{\pi}}. \tag{28}$$

Representation of $\mathbf{IAV}[0]$ can be split into orientation and length. Whereas the distribution of the orientation can be shown analytically to be spherically symmetrical (as discussed previously), derivation of the Maxwell distribution for the length involves certain approximations. Fig. 3 depicts the level of deviation of the expected vector length predicted by (28) from the sample mean of Monte Carlo simulations as a function of RT_{60}. The simulations were conducted 50 000 times for each value of RT_{60} with a sampling rate of $f_s = 8$ kHz, and the deviation is represented in decibel units. The simulation results indicate that the approximations made in the derivation of (25) are accurate for a wide range of RT_{60} values. For extremely fast decay rates [i.e. large $\alpha \propto (RT_{60} f_s)^{-1}$], the approximations are less precise, as can be seen from the left extremity of Fig. 3.

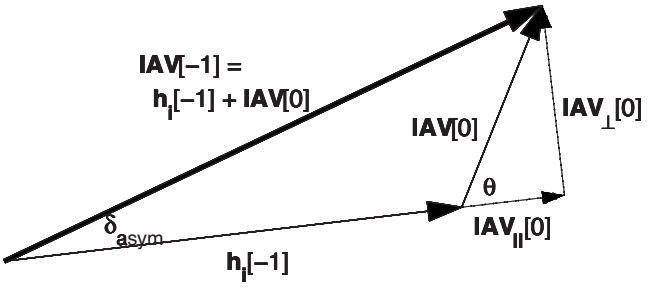

FIG. 4. Illustration of the vector $\mathbf{IAV}[-1]=\mathbf{h}_i[-1]+\mathbf{IAV}[0]$.

V. DISTRIBUTION OF ANGULAR ERROR

Now that the distribution of the interference component $\mathbf{IAV}[0]$ has been determined, we can study how this term affects the AE. The asymptotical AE (δ_{asym}) approaches *(i.p.)* the angle by which $\mathbf{IAV}[-1]=\mathbf{h}_i[-1]+\mathbf{IAV}[0]$ deviates from $\mathbf{h}_i[-1]$ (or equivalently the angle between $\hat{\mathbf{u}}_{\mathrm{asym}}$ and $\mathbf{u}$) (see Appendix A). The vector $\mathbf{IAV}[0]$ can be further divided into components that are parallel and perpendicular to $\mathbf{h}_i[-1]$, which are denoted: $\mathbf{IAV}_{\parallel}[0]$ and $\mathbf{IAV}_{\perp}[0]$. They have respective vector lengths of $\|\mathbf{IAV}[0]\cos(\theta)\|$ and $\|\mathbf{IAV}[0]\sin(\theta)\|$ (where θ is the angle formed by $\mathbf{IAV}[0]$ and $\mathbf{h}_i[-1]$). Fig. 4 illustrates how the various components and angles relate. The AE can be presented using these components as:

$$\delta_{\mathrm{asym}} = \arctan\left(\frac{\|\mathbf{IAV}[0]\|\sin(\theta)}{\|\mathbf{h}_i[-1]\| + \|\mathbf{IAV}[0]\|\cos(\theta)}\right) + \psi, \tag{29}$$

where the final term ψ corrects for unusual cases where $\delta_{\mathrm{asym}} > \pi/2$, and is defined $\psi = \pi \cdot u(-[\|\mathbf{h}_i[-1]\| + \|\mathbf{IAV}[0]\|\cos(\theta)])$, where $u(\bullet)$ represents the unit step function.

A. Characterization of underlying random variables

In (29) we see that δ_{asym} comprises of two random variables relating to the random walk $\mathbf{IAV}[0]$, namely its orientation θ with respect to the correct direction $\mathbf{h}_i[-1]$ (or $\mathbf{u}$), and its amplitude $\|\mathbf{IAV}[0]\|$. The distribution of the former can be established from the spherical symmetry of $\mathbf{IAV}[0]$ [by marginalizing the spherically uniform distribution $f_{\phi,\theta}(\phi,\theta) = 1/4\pi\sin(\theta)$, where ϕ is azimuth] as:

$$f_\theta(\theta) = \int_0^{2\pi} \frac{1}{4\pi}\sin(\theta)d\phi = \frac{1}{2}\sin(\theta), \tag{30}$$

with domain $0 \le \theta \le \pi$. The latter has a Maxwell distribution (27) which is characterized by (25).

The parameter $\sigma_G = \sigma_{\mathrm{IAV}}[0]$ of (25), which characterizes the Maxwell distribution, is itself defined by σ_0^2 from Polack's model. It is desirable to use an equivalent formulation that utilizes more common descriptors of the RIR. To this end we introduce the *direct to reverberation ratio* (DRR) defined as $h_p^2[-1]/\mathrm{EDC}[0]$. Replacing EDC with its expected value, we define:

$$\mathrm{DRR}_{E_s} = \frac{h_p^2[-1]}{E_s\{\mathrm{EDC}[0]\}} = \frac{\|\mathbf{h}_i[-1]\|}{E_s\{\mathrm{EDC}[0]\}}, \tag{31}$$

where the last stage follows from the fact that direct sound pressure and particle velocity are respectively proportionate to $h_p[-1]$ and $\mathbf{h}_v[-1]$, thus (8) implies that $\mathbf{h}_i[-1]=h_p^2[-1]\mathbf{u}$. The term $E_s\{\mathrm{EDC[n]}\}$ pertaining to the denominator can be presented from (16) and (19a) as the sum of a geometric series equaling:

$$E_s\{\mathrm{EDC}[n]\} = \sigma_0^2 \frac{e^{-2\alpha n}}{1-e^{-2\alpha}}. \tag{32}$$

The ratio of $\sigma_{\mathrm{IAV}}[n]$ to $E_s\{\mathrm{EDC[n]}\}$ [from (25) and (32)],

$$\kappa(\alpha) \equiv \frac{\sigma_{\mathrm{IAV}}[n]}{E_s\{\mathrm{EDC[n]}\}} = \frac{1-e^{-2\alpha}}{\sqrt{3(1-e^{-4\alpha})}}, \tag{33}$$

eliminates the parameter σ_0^2 and proves useful. Note that κ is also independent of n and is determined solely by the decay rate α. Based on Eqs. (31) and (33), the parameter governing the random vector $\mathbf{IAV}[0]$ can be expressed:

$$\sigma_{\mathrm{IAV}}[0] = \frac{\kappa(\alpha)}{\mathrm{DRR}_{E_s}} \|\mathbf{h}_i[-1]\|. \tag{34}$$

The Maxwell r.v. $\|\mathbf{IAV}[0]\|$ of (29) can be recast into normalized form $\|\mathbf{IAV}[0]\| = \sigma_{\mathrm{IAV}}[0]Q'$, where Q' is a Maxwell variable with $\sigma_G=1$. After subsequent substitution of (34) into (29), one sees that $\|\mathbf{h}_i[-1]\|$ scales both the numerator and the denominator and can be factored out of ψ, thus having no effect upon δ_{asym}. The only remaining parameters are then α and DRR_{E_s}.

In essence, the angular error is an r.v. which is a function of the r.v.s θ and $\mathbf{IAV}[0]$. These two variables are independent (as follows from spherical symmetry) and their distributions are known. The distribution of θ is fixed whereas, that of $\mathbf{IAV}[0]$ depends upon α and DRR_{E_s}. Consequently, knowledge of these two quantities is sufficient to parameterize the distributions and hence completely specifies the angular error δ_{asym}.

B. Derivation of distribution using approximation

Although δ_{asym} is completely specified, analytical evaluation of its distribution from the transforation of variables θ and $\mathbf{IAV}[0]$ (29) appears difficult or impossible. However, the process can be significantly simplified by assuming that accumulated reverberation intensity is typically much smaller than the direct-arrival intensity, i.e. $\|\mathbf{IAV}[[0]\| \ll \|\mathbf{h}_i[-1]\|$ [or less precisely: the direct to reverberation ratio (DRR) amplitude is sufficiently high]. This condition is necessary for reasonable DOA estimation producing small angular errors. Now, the angular error of (29) can be approximated as $\tilde{\delta}_{\mathrm{asym}}$:

$$\delta_{\mathrm{asym}} \approx \arctan\left(\frac{\|\mathbf{IAV}[0]\|\sin(\theta)}{\|\mathbf{h}_i[-1]\|}\right) \approx \frac{\|\mathbf{IAV}[0]\|\sin(\theta)}{\|\mathbf{h}_i[-1]\|} \equiv \tilde{\delta}_{\mathrm{asym}}, \tag{35}$$

with the last stage following from: $\arctan(\delta_{\mathrm{asym}}) \approx \delta_{\mathrm{asym}}$ for small δ_{asym}.

The term in the numerator $\|\mathbf{IAV}[0]\|\sin(\theta)$ is the Euclidian norm of a three dimensional jointly normal vector (with i.i.d. components and zero mean) multiplied by $\sin(\theta)$. This is equivalent to the norm of the projection of $\mathbf{IAV}[0]$ onto a two dimensional plane. To demonstrate this, we cast the coordinate system to have the z-axis oriented in the same direction as $\mathbf{h}_i[-1]$ (which is permissible since $\mathbf{IAV}[0]$ possess spherical symmetry). Then:

$$\|\mathbf{IAV}[0]\| = \sqrt{\mathrm{IAV}_x^2[0] + \mathrm{IAV}_y^2[0] + \mathrm{IAV}_z^2[0]},$$

$$\sin(\theta) = \frac{\sqrt{\mathrm{IAV}_x^2[0] + \mathrm{IAV}_y^2[0]}}{\sqrt{\mathrm{IAV}_x^2[0] + \mathrm{IAV}_y^2[0] + \mathrm{IAV}_z^2[0]}},$$

$$\|\mathbf{IAV}[0]\|\sin(\theta) = \sqrt{\mathrm{IAV}_x^2[0] + \mathrm{IAV}_y^2[0]}. \tag{36}$$

The final term, which bears similarity to (26), is a Rayleigh variable—also known as a Chi variable of second order. The Rayleigh variable R possesses a distribution of:

$$f_R(r) = u(r)\frac{1}{\sigma_G^2} r e^{-r^2/2\sigma_G^2}. \tag{37}$$

The parameter which controls the Rayleigh distribution of $\|\mathbf{IAV}[0]\|\sin(\theta)$ is $\sigma_G=\sigma_{\mathrm{IAV}}[0]$, i.e. the parameter of the Maxwell variable which was projected.

We note that (35) invokes a scaling of the Rayleigh variable $\|\mathbf{IAV}[0]\|\sin(\theta)$ by a factor of $\|\mathbf{h}_i[-1]\|^{-1}$. Thus, the AE approximation $\tilde{\delta}_{\mathrm{asym}}$ is also a Rayleigh variable and is governed by the scaled parameter $\sigma'_G=\sigma_{\mathrm{IAV}}[0]/\|\mathbf{h}_i[-1]\|$. Substituting with (34) yields:

$$\sigma'_G = \frac{\kappa(\alpha)}{\mathrm{DRR}_{E_s}}. \tag{38}$$

The mean and variance of the Rayleigh distributed $\tilde{\delta}_{\mathrm{asym}}$ are then:

$$E_s\{\tilde{\delta}_{\mathrm{asym}}\} = \sigma'_G\sqrt{\frac{\pi}{2}}, \tag{39a}$$

$$\mathrm{Var}_s\{\tilde{\delta}_{\mathrm{asym}}\} = \left(2-\frac{\pi}{2}\right)\sigma'^2_G. \tag{39b}$$

The cumulative distribution function (c.d.f.) of $\tilde{\delta}_{\mathrm{asym}}$ is:

$$F(\tilde{\delta}_{\mathrm{asym}}) = u(r)(1-e^{-\tilde{\delta}^2_{\mathrm{asym}}/2\sigma'^2_G}), \tag{40}$$

which is useful for determining error margins. For instance, if one would like to know the probability that the AE will not exceed 10°, this may be obtained as $p=F(10\cdot\pi/180)$. In the context of error margins, the inverse c.d.f.:

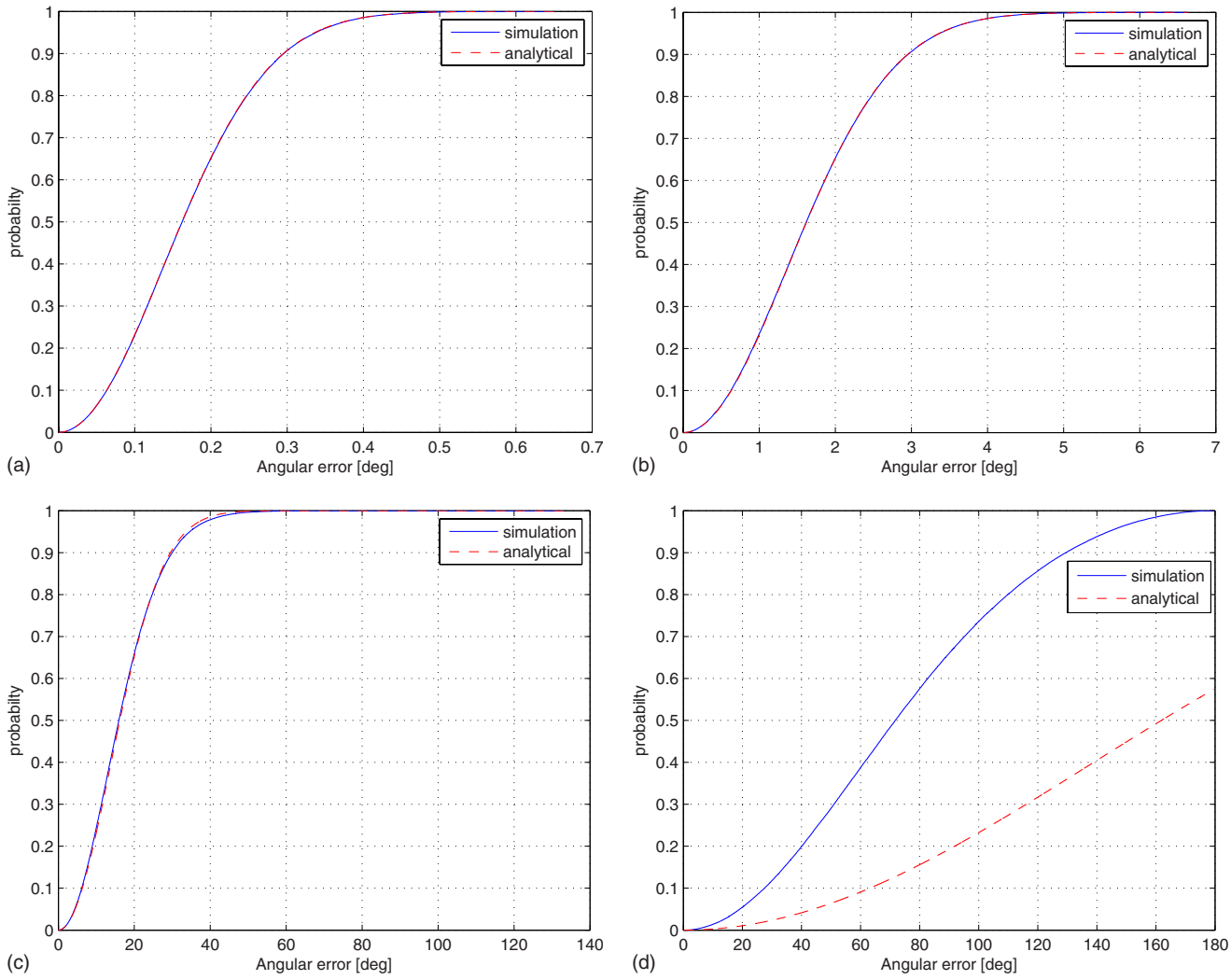

FIG. 5. (Color online) Sample cumulative distribution function of angular error (produced with extended Polack model) compared to function predicted by analytical analysis of $\tilde{\delta}_{\text{asym}}$ for various levels of DRR_{E_s}.

$$F^{-1}(p) = \sigma'_G \sqrt{-2\ln(1-p)} \tag{41}$$

is also useful. For example, if one wishes to find a maximal AE which will not be exceeded with 95% certainty, this may be evaluated by $F^{-1}(0.95)$.

The terms derived in this section provide practical measures for the performance of the intensity based DOA estimator in a reverberant environment. Use of these measures relies solely upon knowledge of the decay rate (α or RT_{60}) and the DRR_{E_s}.

VI. SIMULATION RESULTS

In this section, simulations are preformed to check the veracity of results pertaining to the distribution and statistics of the angular error (AE). Each simulation trial produces an empirical DOA estimate with a corresponding error δ_{asym}. These are then compared with results analytically derived in the previous section for $\tilde{\delta}_{\text{asym}}$. Two separate tests are performed—the first involves generation of RIRs according to the extended Polack model (presented in Sec. IV B) and the second is based upon the image source method.[18]

A. Simulation of intensity DOA estimation using extended Polack model

The reverberant part of the various RIRs (i.e. $\mathbf{h}[n \geq 0]$) were generated using the extended Polack model [i.e. equations (19a) and (19b)]. The part corresponding to the direct arrival ($\mathbf{h}[-1]$) was generated as $h_p[-1]=k$ and $\mathbf{h}_v[-1]=k\mathbf{u}$, where k is a constant and $\mathbf{u}$ is a unit vector with random orientation. The sampling rate was set at 8 kHz and the decay rate at $\text{RT}_{60}=0.5$ s. The RIR reverberation tail was truncated after 16,000 samples (=2 s). The asymptotic value of $\hat{\mathbf{u}}$ was computed with (18) and the AE δ_{asym} calculated with (2). This process was repeated 100 000 times in order to obtain a large sample ensemble of δ_{asym}. The simulation described was performed for several values of k yielding different values of DRR_{E_s}. Fig. 5 presents the sample c.d.f. of δ_{asym} compared with that predicted by (40) for $\tilde{\delta}_{\text{asym}}$. Similarly, Table I compares the sample mean and standard deviation of δ_{asym} with those indicated by (39a) and (39b) for $\tilde{\delta}_{\text{asym}}$.

Both the sample statistical quantities and the sample cumulative density functions indicate a close match between theory and simulation. The exception to this is for small

TABLE I. Statistical results of AE (in degrees) for various levels of DRR_{E_s}.

DRR_{E_s} (dB)	Sample mean	Analytical mean	Sample std	Analytical std
10	0.1727	0.1723	0.0904	0.0901
0	1.7215	1.7229	0.9046	0.9006
−10	17.237	17.2290	9.4711	9.0060
−20	75.3804	172.2901	37.9908	90.0600

DRR_{E_s} [as can be seen in Fig. 5(d) and the final row of Table I]. The reason for this being that the approximations made in Eq. (35) are no longer valid. In fact, as AE becomes large the results become less precise since the small angle approximation is no longer valid (this is blatantly evident in the extreme case of $\delta_{\mathrm{asym}} > \pi = 180°$ which is allowable in the approximation (35) and is presented in the Rayleigh distribution, but is certainly not possible in a physical sense). Otherwise, for moderate DRR_{E_s} the predictions concerning $\tilde{\delta}_{\mathrm{asym}}$ appear highly accurate.

B. Simulation of intensity DOA estimation using image source method

The procedure preformed here to inspect δ_{asym} is essentially similar to that of Sec. VI A, the main difference being the method in which the RIRs are generated. Previously, these had been generated according the extended Polack model, whereas now they are created by the image source method. The RIR_GENERATOR software composed by Habets[21] was used in the simulations. The pressure component is provided by specifying an omnidirectional directivity pattern whereas the particle velocity components are produced by defining three dipole directivity patterns which are mutually orthogonal.

The room setup consisted of a cubic chamber[22,28] with a side dimension of 3 meters. The source and the receiver were positioned at random locations with certain constraints. A distance of one meter was maintained between source and room boundaries in an attempt to minimize the effect of early reflections. The source and receiver were separated by a distance of 51 cm.[23] The ratio DRR_{E_s} was calculated as the quotient of the sample means of the direct and reverberant energies from the pressure RIRs yielding a value of −0.86 dB, and RT_{60} was measured (from the decay curve) as 482 ms. This process was repeated altogether 500 times generating RIRs for varying source-receiver positioning.

The AE δ_{asym} pertinent to each trial was calculated and used to create the sample c.d.f. of δ_{asym}. This is compared to the analytically derived c.d.f. of $\tilde{\delta}_{\mathrm{asym}}$ of (37) in Fig. 6. Although the curve *shapes* exhibit a high level of similarity, the theoretical c.d.f. deviates somewhat to the left of the empirical c.d.f. The sample-mean AE is 2.47° whereas the value predicted from (39a) is 2.01° corresponding to an underestimation of 15.17%.

It may be conjectured that the discrepancies are caused by the early reflections which are not well described by Polack's model. These early arrivals do not mix with other simultaneous arrivals from opposing directions and hence may have a more pronounced effect. In order to test whether this idea is reasonable, we inspect the effects of artificially removing early reflections from the generated RIRs. In Figs. 7(a) and 7(b) the c.d.f.s are displayed following removals of 10 and 20 ms respectively. The results show that the correspondence between theoretical and simulated curves does improve. One may notice that removal of early reflections reduces the overall AE by decreasing reverberant energy (and thus increasing DRR_{E_s}). This process is not possible in a "real-world" estimation scenario and is preformed solely with the intent of obtaining an indication of the impact which early reflections have upon the DOA estimator. Removing increasingly large intervals of reflections brings about further reduction in AE but (39a) may start to overestimate the error. The sample and theoretical AE means for the data in Figs. 6 and 7 are displayed in Table II together with the relative estimation errors (relative error $= (E\{\tilde{\delta}_{\mathrm{asym}}\} - E_{\mathrm{sample}}\{\delta_{\mathrm{asym}}\})/(E_{\mathrm{sample}}\{\delta_{\mathrm{asym}}\})$ where $E_{\mathrm{sample}}\{\circ\}$ stands for sample mean).

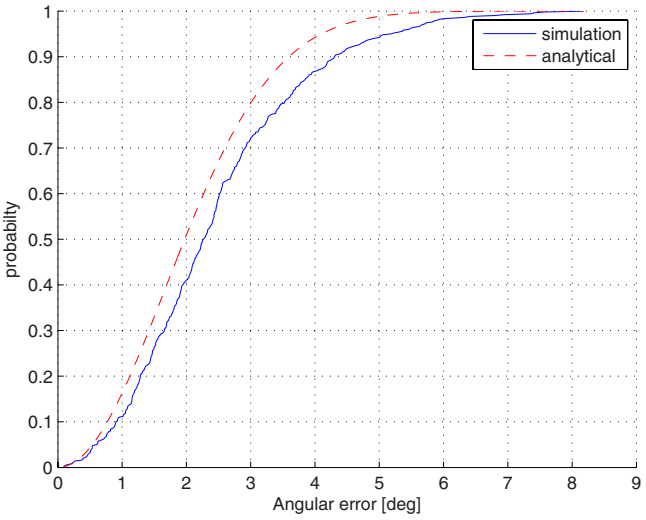

FIG. 6. (Color online) Comparison of theoretical and simulated (image-source method) AE cumulative distribution functions.

VII. SUMMARY

We have used statistical room acoustics to derive a model describing behavior of the intensity vector DOA estimator. This model predicts that when the estimator is used with a particular source-receiver positioning and room environment, the resulting estimate demonstrates bias. Although the precise quantity of the angular error is unknown, we can

TABLE II. Statistical results of AE (in degrees) for original and modified data.

Data set	Theoretical AE mean	Sample AE mean	Relative estimation error (%)
Unmodified data	2.0995	2.4748	−15.17
10 ms removed	1.7657	2.0088	−12.10
20 ms removed	1.2979	1.3112	−1.02

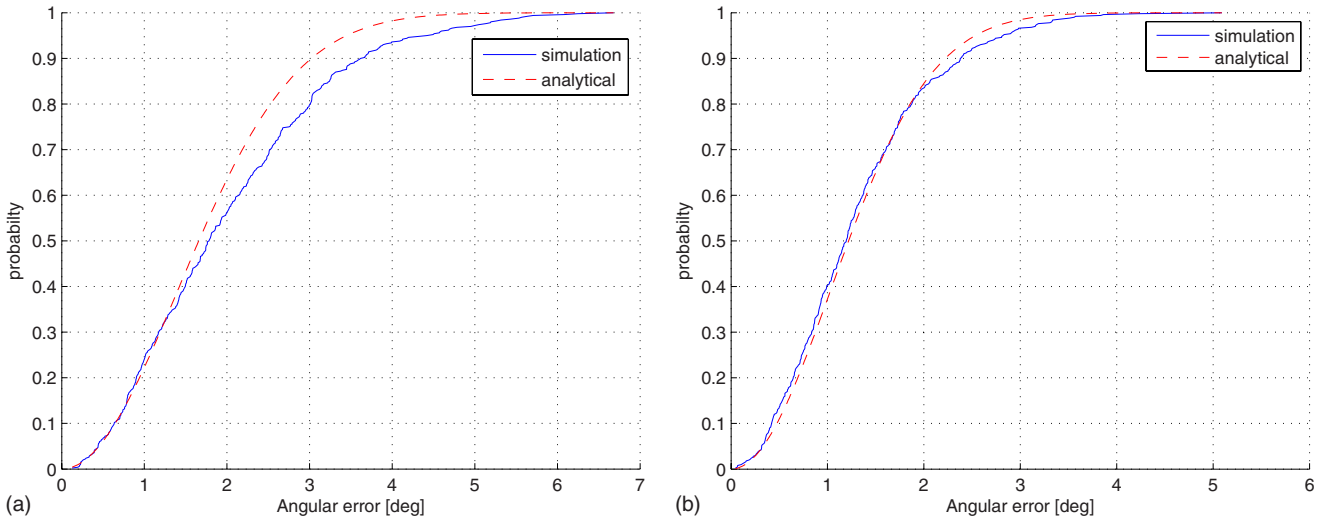

FIG. 7. (Color online) Comparison of theoretical and simulated AE cumulative distribution functions when early arrivals are removed.

estimate its probabilistic distribution and statistical properties. These attributes are validated through simulations for a wide range of parameters.

ACKNOWLEDGMENT

This research was supported by the Marie Curie Intra European Fellowship within the 7th European Community Framework Programme under Contract No. PIEF-GA-2009-237246.

APPENDIX A: ASYMPTOTIC CONVERGENCE OF ESTIMATOR

This appendix inspects the asymptotic behavior of $\bar{\mathbf{y}}_i = [i_{x_{\text{avg}}}\ i_{y_{\text{avg}}}\ i_{z_{\text{avg}}}]^T$ as the number of samples increases (and when $\mathbf{h}[n]$ is assumed to be known for all n). It is shown that as $N \to \infty$, the value of $\bar{\mathbf{y}}_i$ approaches in probability *(i.p.)* the value of $E\{\bar{\mathbf{y}}_i|\mathbf{h}\}$. The convergence of the DOA estimator $\hat{\mathbf{u}}$ and of the AE follows directly.

Investigation of (11) shows that the intensity signal $\mathbf{i}[n] = \mathbf{v}[n]p[n]$ is produced by applying a time invariant (but nonlinear) system to $s[n]$. Since $s[n]$ is wide-sense-stationary (see definitions in Sec. II A) so will the system's output $\mathbf{i}[n] = [i_x[n]i_x[n]i_x[n]]^T$. We shall start by inspecting the x component, $i_x[n]$. First we assume that distant samples have arbitrary low correlation. This can be stated more formally through the conditional auto-covariance function $C_{ix}[\ell] \equiv \text{Cov}\{i_x[n+\ell], i_x[n]|\mathbf{h}\}$ as:

$$\lim_{\ell \to \infty} C_{i_x}[\ell] = 0. \tag{A1}$$

The conditional variance of $i_{x_{\text{avg}}} = 1/N \Sigma_{n=1}^{N} i_x[n]$ is then:

$$\begin{aligned}\text{Var}\{i_{x_{\text{avg}}}|\mathbf{h}\} &= \frac{1}{N^2}\text{Var}\left\{\sum_{n=1}^{N} i_x[n]|\mathbf{h}\right\} \\ &= \frac{1}{N^2}\sum_{n=1}^{N}\sum_{n'=1}^{N}\text{Cov}\{i_x[n], i_x[n']|\mathbf{h}\} \\ &= \frac{1}{N^2}\sum_{n=1}^{N}\sum_{n'=1}^{N} C_{i_x}[n-n']. \end{aligned} \tag{A2}$$

It has been shown[24] that (A1) is sufficient to ensure that the variance in (A2) converges to zero as $N \to \infty$. Application of Chebyshev's inequality establishes that $i_{x_{\text{avg}}}$ approaches *(i.p.)* its expected value.

The same process can also be applied to the components $i_{y_{\text{avg}}}$ and $i_{z_{\text{avg}}}$. Thus, the vector $\bar{\mathbf{y}}_i$ converges to its expected value:

$$\bar{\mathbf{y}}_i \underset{i.p.}{\longrightarrow} E\{\bar{\mathbf{y}}_i|\mathbf{h}\}. \tag{A3}$$

Since the DOA estimator $\hat{\mathbf{u}}$ is a function of the converging random vector $\bar{\mathbf{y}}_i$ [produced by normalization, as given in (10b)], it will converge as well:

$$\hat{\mathbf{u}} \underset{i.p.}{\longrightarrow} \frac{E\{\bar{\mathbf{y}}_i|\mathbf{h}\}}{\|E\{\bar{\mathbf{y}}_i|\mathbf{h}\}\|} = \hat{\mathbf{u}}_{\text{asym}}. \tag{A4}$$

The same logic holds for the angular error [which is defined in (2)]:

$$\delta \underset{i.p.}{\longrightarrow} \sin^{-1}\left(\frac{\|\hat{\mathbf{u}}_{\text{asym}} - \mathbf{u}\|}{2}\right). \tag{A5}$$

Essentially, it has been shown that the value of $E\{\bar{\mathbf{y}}_i|\mathbf{h}\}$ which is computed explicitly in (14), uniquely determines the asymptotic values[25] of $\hat{\mathbf{u}}$ and δ.

APPENDIX B: SPHERICAL SYMMETRY OF INTENSITY RESPONSE

This appendix proceeds to demonstrate that the distribution of $\mathbf{h_i}[n]=\mathbf{h_v}[n]h_p[n]$ is spherically symmetrical. We start with the following Lemma:

Lemma: Let $\mathbf{V}=[XYZ]^T$ be a random vector with spherical distribution, and let W be a scalar random variable which is independent of $\mathbf{V}$. Then, the vector $\mathbf{V}'=[X'Y'Z']^T$ which is formed by $\mathbf{V}'=\mathbf{V}W$ also has a spherical distribution.

Proof: We transform vector $\mathbf{V}'$ into $\mathbf{V}''=[X''Y''Z'']^T$ by reorienting the Cartesian axes (i.e. $\mathbf{V}'$ is premultiplied by a rotation matrix $\mathbf{R}$). Now:

$$\mathbf{V}''=\mathbf{R}\mathbf{V}'=\mathbf{R}(\mathbf{V}W)=(\mathbf{R}\mathbf{V})W\simeq\mathbf{V}W=\mathbf{V}', \tag{B1}$$

where the symbol $\simeq$ denotes that the affiliated terms are not necessarily equal but have identical distributions. The transition between the third and fourth line of (B1) follows from the fact that the distribution of $\mathbf{V}$ is spherically symmetrically and thus invariant to rotation. As such, $\mathbf{R}\mathbf{V}\simeq\mathbf{V}$.

It has been shown that $\mathbf{R}\mathbf{V}'\simeq\mathbf{V}'$, i.e. the distribution of $\mathbf{V}'$ is invariant to rotation; hence, it is spherically symmetrical. ■

This lemma can be applied to $\mathbf{h_i}[n]=\mathbf{h_v}[n]h_p[n]$. Coefficients of the random vector $\mathbf{h_v}[n]$ have a spherically symmetrical distribution, since its elements are Gaussian i.i.d. variables with zero mean. Furthermore, $h_p[n]$ is independent of $\mathbf{h_v}[n]$. [These observations are evident from (21a) and (21b)]. From the lemma, it directly follows that their product, $\mathbf{h_i}[n]$, possesses a spherically symmetrical distribution.

[1]Y. Huang, J. Benesty, and J. Chen, "Time delay estimation and source localization," in *Springer Handbook of Speech Processing*, edited by J. Benesty, M. M. Sondhi, and Y. Huang (Springer, New York, 2007), pp. 1043–1063.

[2]P. Stoica, Z. Wang, and J. Li, "Robust capon beamforming," IEEE Signal Process. Lett. **10**, 172–175 (2003).

[3]G. D'Spain, W. Hodgkiss, and G. Edmonds, "The simultaneous measurement of infrasonic acoustic particle velocity and acoustic pressure in the ocean by freely drifting Swallow floats," IEEE J. Ocean. Eng. **16**, 195–207 (1991).

[4]B. A. Cray and A. H. Nuttall, "A comparison of vector-sensing and scalar-sensing linear arrays," Naval Undersea Warfare Center Division Report No. 10632, Newport, RI, 1997.

[5]H.-E. de Bree, P. Leussink, I. T. Korthorst, D. H. Jansen, D. T. Lammerink, and P. M. Elwenspoek, "The microflown, a novel device measuring acoustical flows," in 8th International Conference on Solid-state Sensors and Actuators, and Eurosensors IX, Stockholm, Sweden (1995), Vol. **1**, pp. 536–539.

[6]A. Nehorai and E. Paldi, "Acoustic vector-sensor array processing," IEEE Trans. Signal Process. **42**, 2481–2491 (1994).

[7]M. Hawkes and A. Nehorai, "Acoustic vector-sensor correlations in ambient noise," IEEE J. Ocean. Eng. **26**, 337–347 (2001).

[8]F. Jacobsen and T. Roisin, "The coherence of reverberant sound fields," J. Acoust. Soc. Am. **108**, 204–210 (2000).

[9]J.-D. Polack, "Transmission of sound energy in concert halls," Ph. D. thesis, Universit'e du Maine, Le Mans (1988).

[10]J. Jot, L. Cerveau, and O. Warusfel, "Analysis and synthesis of room reverberation based on a statistical time-frequency model," in 103th Audio Engineering Society Convention, New York (1997), Preprint No. 4629.

[11]A. Nehorai and M. Hawkes, "Performance bounds for estimating vector systems," IEEE Trans. Signal Process. **48**, 1737–1749 (2000).

[12]A. D. Pierce, *Acoustics—An Introduction to Its Physical Principles and Applications* (Acoustical Society of America, Melville, NY, 1991), Chap. 1, p. 39.

[13]Obviously this can not be literally accurate since the direct arrival would defy causality. Rather, it should be viewed as a time shifted version of the RIR such that the reflections start at $n=0$. The benefits of this notation are twofold. Firstly, reverberation starting at time $n=0$ is less cumbersome to manipulate mathematically. Furthermore, any problems arising from the propagation delay of the direct arrival being a fraction of the sampling rate are circumvented.

[14]M. R. Schroeder, "New method of measuring reverberation time," J. Acoust. Soc. Am. **37**, 409–412 (1965).

[15]M. R. Schroeder, "Statistical parameters of the frequency response curves of large rooms," J. Audio Eng. Soc. **35**, 299–306 (1987); Acustica **4**, 456–468 (1954).

[16]H. Kuttruff, *Room Acoustics*, 4th ed. (Spon, London, UK, 2000), pp. 97–100, pp. 115–119.

[17]P. M. Morse and K. U. Ingard, *Theoretical Acoustics* (McGraw-Hill, New York, 1968), Chap. 9, pp. 576–579.

[18]J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," J. Acoust. Soc. Am. **65**, 943–950 (1979).

[19]E. A. P. Habets and S. Gannot, "Generating sensor signals in isotropic noise fields," J. Acoust. Soc. Am. **122**, 3464–3470 (2007).

[20]It is instructive to note an analogy to a better known example from the field of communication. Wireless channels are often modeled as consisting of multiple reflections arriving with varying phases and magnitudes. A similar analysis indicates that the net effect is a symmetrical Gaussian distribution—the phase (=direction) is distributed uniformly, whereas the magnitude is a Rayleigh variable (Chi distribution of second order). This communication model is known as the Rayleigh channel (Ref. 26) and can be considered a two dimensional analog to our discussion. In fact, the aforementioned distribution is named after Lord Rayleigh who proposed its use to model the summation of sound waves with uniformly distributed phases (Ref. 27).

[21]E. A. P. Habets, "Room impulse response (RIR) generator," http://home.tiscali.nl/ehabets/rir_generator.html (Last viewed 4/30/2010).

[22]The choice of room shape was dictated by the peculiarities of the image source method. Since the room used is rectangular and the reflections specular, permitted directions of a sound ray will be limited by its initial direction and they will not *mix* properly (Ref. 28). This results in sound decay rate being highly dependant upon direction. For a room containing boundaries with identical decay coefficients, reverberation propagating parallel to the longest room dimension will eventually dominate the RIR tail. Choosing a cube shaped room mitigates (but does not eliminate) this nonisotropic phenomenon.

[23]The actual distance was adjusted by a fraction of a millimeter so that the propagation delay would correspond to an integer multiple of the sampling period.

[24]P. J. Brockwell and R. A. Davis, *Time Series: Theory and Methods*, 2nd ed. (Springer, New York, 1991), Chap. 7, pp. 218–219.

[25]The convergence demonstrated herein is *in probability*. Since $\mathrm{Var}\{\bar{\mathbf{y}}_i\}\rightarrow 0$, it follows by definition that $\bar{\mathbf{y}}_i$ also converges in the sense of *mean-square-error*. It is possible to prove that $\hat{\mathbf{u}}$ and δ converge in this sense by way of an argument utilizing the fact that the range of values these r.v.s may take on is bounded [see (2) and (10b)].

[26]D. Tse and P. Viswanath, *Fundamentals of Wireless Communication* (Cambridge University Press, Cambridge, UK, 2005), Chap. 2, p. 36.

[27]J. W. S. Rayleigh, *The Theory of Sound*, 2nd ed. (Dover, New York, 1945), Vol. **I**, Sec. 2-42, pp. 35–42.

[28]J.-D. Polack, "Playing billiards in the concert hall: The mathematical foundations of geometrical room acoustics," Appl. Acoust. **38**, 235–244 (1993).