CONFIDENCE LIMITS FOR COHERENCE – TOWARD A BETTER UNDERSTANDING OF ITS USE AS A DIFFUSE FIELD METRIC FOR QUANTIFYING THE QUALITY OF DIRECT FIELD ACOUSTIC NOISE (DFAN) TESTING SYSTEMS

Marcos A. Underwood (1), **Jonathan A. Hargreaves** (2)
(1) MSI DFAT, 4900 Wetheredsville Road, Baltimore, MD 21207, USA, Email: m.underwood@msidfat.com
(2) Acoustic Research Centre, University of Salford, Salford, M5 4WT, UK, Email: j.a.hargreaves@salford.ac.uk

ABSTRACT

To determine the degree of diffuseness of a particular acoustic field, an oft-considered metric is the coherence between spaced pairs of measurement microphones. This is useful because the expected theoretical ‘sinc^2’ trend for a diffuse field is well known from the literature. An issue with application to date, however, is that this has been used without recognising in any systematic way that a measured coherence estimate is a random sample, with its own natural variance and probability distribution. Instead, plots are typically shown with all the raw data and users are required to estimate the statistics by eye. Part of the challenge is that the variance and confidence intervals of coherence estimates are not straightforward functions, changing from vanishing variance when a pair of microphone signals are fully correlated to a wide and asymmetric distribution when they are independent (or close to). This paper aims to address this knowledge gap by exploring the probability density and cumulative distribution functions of coherence, which are needed to determine confidence intervals, such that sound field specifications and comparisons can be made more appropriately based on the derived confidence intervals. Moreover, it explores a ‘warped’ space in which the probability distribution of coherence is close to Gaussian. Implications for DFAN best practice are discussed.

1. INTRODUCTION

Within the space-launch industry, there is a need for acoustic environmental testing. This aims to replicate, in a precisely controlled way, the damage and/or fatigue that might occur during launch due to the extreme acoustic environment [1]. It is used both to qualify new payload designs and for pre-flight-testing purposes.

Traditionally, such tests have been performed in Reverberant Acoustic Test Facilities (RATFs), but these have disadvantages including high construction and operation costs, and that the payload must always be brought to the test facility, which can be costly when large payloads and/or distances are involved.

Since the mid-1990s, Direct Field Acoustic Noise (DFAN) testing – also known as Direct Field Acoustic Testing (DFAT®) – has existed as an alternative. In this, large arrays of high-powered loudspeakers are used to produce the necessary acoustic environment. Such systems are typically cheaper than a RATF and allow the test system to be brought to the test article. The scope of what sound field DFAN testing systems could render is broad, enabled by the powerful control systems they use [2,3], but demand from the space launch industry has, to date, been for them to replicate the sound field produced by RATFs, since this derisks the transfer from one testing technology to the other.

There is, therefore, a demand for evidence that the sound field produced by a DFAN testing system is equivalent to that produced by a RATF. By design, both produce sound fields that are inherently random, so what is required to compare them is metrics that capture and quantify the key characteristics of these random sound fields. Coherence γ^2 between spaced microphone pairs is an established metric for this purpose, one of the first appearances of which within the DFAN literature being [4] in 2012.

RATF are assumed to produce a perfectly ‘diffuse’ field, in which there is equal probability of arrival of sound energy from any direction. It is well known that under this assumption, plus another that energy from these directions is uncorrelated, the coherence between signals from a spaced pair of omnidirectional microphones follows a $\mathrm{sinc}^2\, kr$ trend [5], where $k = 2\pi f / c_0$ is the wavenumber in radians per metre, r is microphone separation distance in meters, f is frequency in Hz, and c_0 is the speed of sound in meters per second. Figure 1 shows this trend for three microphone spacings. This has been proven experimentally in reverberation rooms [6], subject to certain methods of averaging and caveats.

Hence, there is an appetite to demonstrate that coherence measurements from DFAN tests follow the theoretical $\mathrm{sinc}^2\, kr$ trend. Emphasising this demand, a metric – the now so-called “Sinc Indicator Function (SIF)” – has recently been proposed [7] that normalises the difference between measurements and the theoretical trend, then averages over microphone pairings to reduce this to a single figure of merit.

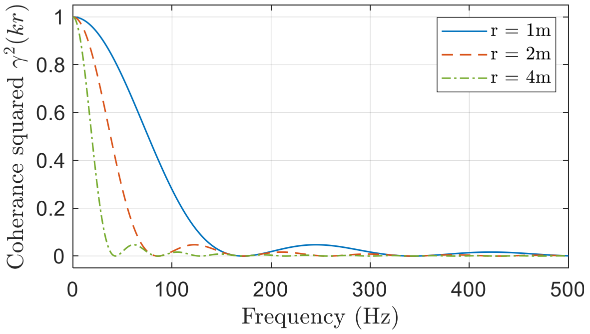

Figure 1: Theoretical value of coherence in a diffuse field, plotted for three microphone spacings.

An interesting consequence of such metrics is that practitioners are now also applying them to reverberation rooms – as is done as an example herein – and showing the limitations in how well they achieve their assumed diffuse field conditions. This is a known issue in reverberation chambers used for Room Acoustic testing [8,9], and a revision of ISO 354 is underway with the aim of mitigating some of the consequences of these issues.

Despite these advances, a weakness of all works the authors have witnessed to date is an absence of the consideration of confidence intervals and uncertainty on both the measured data and the theoretical bound. Measured samples of random signals, as are being considered here, are always *estimates* of the signals' true statistics, and this applies to coherence too. Thus, even the theoretical trend should have confidence limits to show the amount of measurement spread that should be expected given the number of observations made and averaged. This paper presents methods to achieve this.

The paper is structured as follows. Section 2 reviews common definitions of coherence and methods for its measurement. Section 3 presents how its statistics are affected by measurement parameters, including newly computed confidence limits for the theoretical trend, which section 4 then applies to DFAN and RATF data. Section 5 explores a transform that warps coherence data such that its statistics are close to Gaussian and section 6 presents results from a Monte Carlo testbed, which is used for validation of analytical trends and to explore trends relevant to DFAN. Finally, section 7 draws conclusions & identifies areas requiring further research.

2. COHERENCE AND ITS MEASUREMENT

The ordinary coherence function, as we are considering here, is a well-established quantity that is widely used in the literature. In descriptive terms, it quantifies the degree to which two signals are correlated with one another, with a value of 0 indicating that they arise from completely independent sources, a value of 1 indicating that they arise from the same source via linear, time-invariant paths, and values in between indicating some degree of dependent / independent mixing and/or nonlinearity.

Nonetheless, there is scope for confusion as one can typically encounter three versions of coherence with different numeric values that often go by that same name:

- Magnitude-Squared Coherence (MSC) is the most widely used and reported, defined as:

$$\gamma_{kl}^2(f) = \frac{|g_{kl}(f)|^2}{g_{kk}(f)g_{ll}(f)}. \tag{1}$$

- Magnitude coherence, defined as:

$$\gamma_{kl}(f) = \frac{|g_{kl}(f)|}{\sqrt{g_{kk}(f)g_{ll}(f)}}. \tag{2}$$

- Complex coherence, also called coherency (a term coined by Weiner in 1930 [10,11]), defined as:

$$\gamma_{kl}(f) = \frac{g_{kl}(f)}{\sqrt{g_{kk}(f)g_{ll}(f)}}. \tag{3}$$

In the above, the dependency of coherence on frequency f has been acknowledged. The subscripts k and l relate to selections of pairs of signals from a set of signals being analysed – typically control and/or monitor mic pairs in our application – where the terms g_{kl}, g_{kk}, and g_{ll} can be drawn from a matrix called a Spectral Density Matrix (SDM) [12]. These are extensively used in the definition, control, and analysis of DFAN sound fields [2,3], and have real-valued auto Power-Spectral-Density (PSD) terms (g_{kk}, and g_{ll}) on the diagonal and complex Cross-Power-Spectral-Density (CSD) terms (g_{kl}) elsewhere. It follows from their definitions that (eq. 1 and 2) lead to real coherence values in the range [0 ... 1], whereas eq. 3 defines a complex coefficient with magnitude ≤ 1.

In what follows, the paper will use the definition given by eq. 1, i.e., Magnitude-Squared Coherence (MSC), exclusively, since this is the variant most widely used in Acoustics. To aid readability of the discussion, the term 'coherence' will be used to denote MSC, as is common.

Fortuitously, MSC is also the variant for which a numerically efficient closed-form analytical statement exists for the probability distribution and cumulative distribution functions of its estimates [11]. This will be presented in section 3 and follows work by Fisher [13], Goodman [14], and Carter *et al.* [10,11], motivated by applications in SONAR. This work has found application in the biomedical sector – for example, processing MRI signals – but we believe the use herein to inform acoustic diffuse field assessment to be novel.

Statistical and Probabilistic Theory Associated with Coherence Estimates

It is commonplace to obtain measured estimates for the SDM terms – and from these, coherence – from acquired signals via averages of FFT blocks according to Welch's Method [11,12]. Signal blocks are typically multiplied by window functions pre-FFT to minimise spectral leakage and may be overlapped [8,15]. The estimates are:

$$\hat{g}_{kl}(f) = \frac{1}{N\Delta_\mathrm{f}}\sum_{n=1}^{N}\hat{M}_{k_n}(f)\overline{\hat{M}_{l_n}}(f), \tag{4}$$

$$\hat{\gamma}_{kl}^2(f) = \frac{|\hat{g}_{kl}(f)|^2}{\hat{g}_{kk}(f)\hat{g}_{ll}(f)}. \tag{5}$$

Here, $\hat{M}_{k_n}$ and $\hat{M}_{l_n}$ are the Δ_f Hz resolution FFT respectively of the n^{th} blocks from the k^{th} and l^{th} signals, likely arising from microphones in our application. N blocks are averaged (in eq. 4) to arrive at the estimate of the SDM entries $\hat{g}_{kl}$, and these are divided (in eq. 5) to give an estimate of $\hat{\gamma}_{kl}^2$ of the coherence between each signal pair k, l. These subscripts are mostly dropped from these quantities throughout the remainder of the paper, along with dependence on frequency (f), to aid clarity.

The estimates in eq. 4 and 5 are ‘hatted’ to distinguish them from known spectra in eq. 1 and 2. They can have random and/or deterministic estimation errors depending on how the parameters of eq. 4 (FFT size, windowing, overlapping, and number of averages N) are chosen.

Validity of the above estimators assumes that the signals are jointly Gaussian and ergodic, which is typically true for DFAN signals once control adjustments stabilise. It should also be noted that estimation of coherence by eq. 2 can be biased (always underestimated) when analysing a coherent pair of signals that are separately by a delay that is significant compared to the FFT length [12]. But this should not be an issue in DFAN. Typical FFT resolution Δ_f for DFAN analysis and control is ~3Hz, in which time sound can propagate over 100m, which is very large compared to typical microphone spacings.

Effective Degrees of Freedom (EDOF)

The theoretical statements presented in the next section were derived in their source works under the assumption that the data in each set of FFT blocks is independent. This is the case when no windowing and no overlapping is used, but this is not ideal for practical use. Window functions are required to minimise spectral leakage, after which overlapping is logical since it makes best use of data that is otherwise excluded by the window [15]. Hanning or Hamming windows with 50% overlap are common. The important parameter then is one that measures the amount of independent data that remains.

It is well known that the FFT of a real Gaussian signal has Gaussian-distributed real and imaginary parts at each frequency. Equation 4 averages the product of pairs of such spectra, giving an estimate of the power spectral density $\hat{g}_{kl}$ that, in each frequency bin, follows a chi-squared (X^2) probability distribution with $2N$ degrees of freedom (though there are some additional complications in cases of low coherence [16]).

But if FFT blocks are drawn from overlapping signal segments, as is often done, then the amount of independent data is reduced, and the number of ‘Effective Degrees of Freedom’ (EDOFs) reduces too. The extent of this depends on the degree of overlap and the choice of window function, and is parameterised by a factor ν, so $\mathrm{EDOF} = \nu \times N$. Using this formulation, we have that $\widehat{N} = \mathrm{EDOF}/2 = \nu N$, where $\widehat{N}$ is the effective number of independent blocks that should be used in the theoretical models and confidence limits that follow (when coherence is measured according to eq. 4 and 5).

Depending on the application, EDOF or $\widehat{N}$ can be found from N using a multiplier ν from Table 1 [17]. Note that with an overlap of 50% and the Hanning window, Table 1 shows that just over half the data with an overlap of 50% can achieve nearly the same $\widehat{N}$ as the original N with no overlap and a similar resulting EDOF.

Window type	EDOF multiplier ν
Rectangular	1.33
Triangle	1.78
Hanning	1.90 ≈ 2
Hamming	1.80

Table 1: EDOF multiplier when using 50% overlap is used to evaluate the FFTs feeding eqs. 1-3.

3. PROBABILITY DENSITY FUNCTION OF COHERENCE ESTIMATES

The Probability Density Function (PDF) and Cumulative Distribution Function (CDF) of coherence estimates $\hat{\gamma}^2$ are more complex than those for PSD and CSD estimates. These depend on the ‘true coherence’ γ^2, the spacing between microphones, and involve long and numerically sensitive calculations within their determinations. Since these calculations depend on knowing ‘true coherence’, which is typically unknown, the calculation of these important functions can be somewhat circular, since in general cases, a value for ‘true coherence’ would have to be also estimated or assumed.

However, for Acoustic fields we have three simplifying conditions that make the task of deriving confidence intervals for coherence estimates feasible:

a) The underlying microphone response data for DFAN is typically jointly Gaussian and wide-sense stationary at full-level, which makes it possible to have exact formulations to calculate the needed PDF and CDF functions, going back to Fisher [11,13].
b) There are straightforward methods available to calculate the number of EDOFs used to estimate the PSD and CSD spectra [15,16] needed to determine confidence intervals on $\hat{\gamma}^2$, given the overlap factor and the type of window used [15].
c) We have a model for ‘true coherence’ in the form of the sinc^2 characteristic for an ideal diffuse acoustic field [6], which measured coherence estimates during a DFAN need to match approximately. This removes the inherent circular nature of depending on an approximate ‘true coherence’ value.

Given a-c, we have that the first order PDF for the estimate of coherence function $\hat{\gamma}^2$ between a pair of microphones, given its true value γ^2 and the number of independent blocks N of microphone time-history data used to estimate the PSDs $\hat{g}_{kk}$ and $\hat{g}_{ll}$ and CSD $\hat{g}_{kl}$ from eq. 4 and 5 is given [11] by:

$$\mathrm{PDF}(\hat{\gamma}^2 |\, N, \gamma^2) = (N-1)\left[\frac{(1-\gamma^2)(1-\hat{\gamma}^2)}{(1-\gamma^2\hat{\gamma}^2)^2}\right]^N \frac{1-\gamma^2\hat{\gamma}^2}{(1-\hat{\gamma}^2)^2} F_1^2(1-N, 1-N; 1; \gamma^2\hat{\gamma}^2). \quad (6)$$

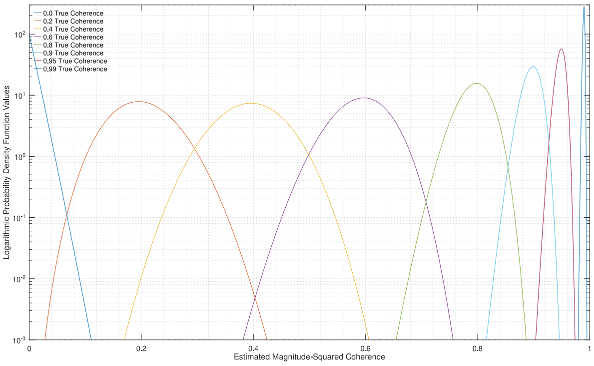

Figure 2: PDFs of the MSC estimate $\hat{\gamma}^2$ for eight values of true MSC γ^2, with $N = 100 \equiv 200$ EDOFS

Here $F_1^2(1-N, 1-N; 1; \gamma^2\hat{\gamma}^2)$ is a hypergeometric function that can be simplified as follows (eq. 7 to 9):

$$F_1^2(1-N, 1-N; 1; \gamma^2\hat{\gamma}^2) = \sum_{n=0}^{N-1} T_n, \tag{7}$$

where we have that:

$$T_0 = 1,$$
$$T_n = T_{n-1} \times \frac{(n-N)(n-1+(1-N)\gamma^2\hat{\gamma}^2)}{n^2}. \tag{8}$$

This expresses the hypergeometric function as a straightforward polynomial of $\hat{\gamma}^2$, and the probability density function for the estimated coherence $\hat{\gamma}^2$ [11].

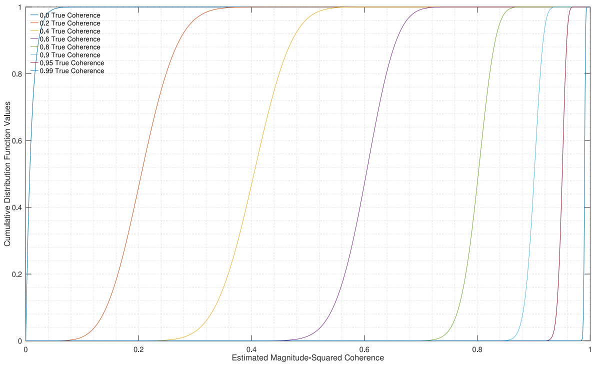

Figure 3: CDFs of the MSC estimate $\hat{\gamma}^2$ for eight values of true MSC γ^2, with $N = 100 \equiv 200$ EDOFS

A special case for the PDF occurs, when $\gamma^2 = 0$:

$$\text{PDF}(\hat{\gamma}^2 \mid N, 0) = (N-1) \times (1-\hat{\gamma}^2)^{N-2}. \tag{9}$$

Figure 2 contains a plot that displays the PDF of coherence estimates for $N = 100$, true coherences: 0.0, 0.2, 0.4, 0,6, 0.8, 0.9, 0.95 and 0.99; a linear true coherence x-axis; and logarithmic PDF magnitude y-axis format, given its wide range of amplitude, where the PDF becomes narrower, taller as $\gamma^2 \rightarrow 1$, and ∞ when $\gamma^2 = 1$.

Determining the cumulative distribution function of Coherence Estimates:

Given the previous conditions a-c, we also have that the first order Cumulative Distribution Function (CDF) for the estimate of coherence $\hat{\gamma}^2$ between a pair of microphones, given its true value γ^2 and the number of independent blocks, N, of microphone time-history data used to estimate the PSDs (eq. 4 and 5), is [11]:

$$\text{CDF}(\hat{\gamma}^2 \mid N, \gamma^2) = \hat{\gamma}^2 \left[\frac{1-\gamma^2}{1-\gamma^2\hat{\gamma}^2}\right]^N \left[\sum_{n=0}^{N-2} \left(\frac{1-\hat{\gamma}^2}{1-\gamma^2\hat{\gamma}^2}\right)^n F_{1,n}^2(-n, 1-N; 1; \gamma^2\hat{\gamma}^2)\right]. \tag{10}$$

Here (eq. 10), $F_{1,n}^2(-n, 1-N; 1; \gamma^2\hat{\gamma}^2)$ is the same hypergeometric function with different parameters used for the CDF calculation for each summand. This differs from the previous PDF determination (eq. 6), where it was evaluated only once.

A special case for the cumulative distribution function also occurs when $\gamma^2 = 0$, which is:

$$CDF(\hat{\gamma}^2 \mid N, 0) = 1 - (1-\hat{\gamma}^2)^{N-1}. \tag{11}$$

Figure 3 contains a plot that shows the CDF of coherence estimates with $N = 100$, true coherences: 0.0, 0.2, 0.4, 0,6, 0.8, 0.9, 0.95 and 0.99; a linear true coherence x-axis and a linear probability y-axis format. Note the steepening of the slope as $\gamma^2 = 0$ increases to 1.

Note that the CDF also has a finite support like PDF. Additionally, both the PDF and CDF are needed to obtain the confidence intervals about the true value of coherence, γ^2, where estimated coherence, $\hat{\gamma}^2$, can have values within the given probabilities specified for a given confidence interval, say within 95, 99, 99.9 or 99.99% and even 99.999% confidence, if needed. The methodology used is discussed in the following section.

4. OBTAINING EXACT CONFIDENCE INTERVALS FOR DFAN AND RFAT MICROPHONE RESPONSE DATA

The confidence intervals methodology that was obtained during the research for this paper provides a scale around the ideal $\text{sinc}^2\, kr$ function, which represents the coherence of a pair of microphones in an ideal diffuse field [6] at a particular frequency and given separation, to quantify how well a particular DFAN test's acoustic field approximates an ideal acoustic field [2,3]. This methodology is also similarly applicable to estimated coherence spectra from pairs of microphones in a field from a particular Reverberant Field Acoustic Test (RFAT) or from other such acoustic fields to determine the degree to which those fields are nearly diffuse.

In general, given the CDF we obtained in the previous section, a confidence interval with any specified probability, say $1-\alpha$, where $0 \leq 1-\alpha \leq 1$ [20,21], can be obtained. This interval needs to be defined by its upper and lower bounds, b_u and b_l, where $0 \leq \mathrm{b}_l \leq \mathrm{b}_u \leq 1$, with the probability $1-\alpha$ that a given estimated coherence, $\hat{\gamma}^2$ will occur within the closed interval $[b_l \dots b_u]$, that is:

$$\begin{aligned} &\mathrm{P}(\, \mathrm{b}_l \leq \hat{\gamma}^2 \leq \mathrm{b}_u) \\ &= \mathrm{CDF}(b_\mathrm{u} |\, N, \gamma^2) - \mathrm{CDF}(b_\mathrm{l} |\, N, \gamma^2) \\ &= 1-\mathrm{\alpha}, \end{aligned} \tag{12}$$

demonstrating $[b_l \dots b_u]$ was designed to contain estimates $\hat{\gamma}^2$, from an average of N independent blocks with a true coherence of $\hat{\gamma}^2$ and a probability of $1-\alpha$.

To calculate what until now are the unknown bounds, b_u and b_l, of the interval with the specified probability, we need to find the roots to two equations, in order to obtain a symmetric confidence interval for any given estimated coherence, $\hat{\gamma}^2$, from the average of N independent blocks as in (eq. 4 and 5), around a given true coherence, γ^2. Specifically, we need to find the needed bounds of the so specified confidence interval, b_u and b_l, by finding the roots of both of the following non-linear equations:

$$\begin{aligned} &\mathrm{CDF}(b_\mathrm{u} |\, N, \gamma^2) - \left(1-\frac{\alpha}{2}\right) = 0, \\ &\mathrm{CDF}(b_\mathrm{l} |\, N, \gamma^2) - \frac{\alpha}{2} = 0. \end{aligned} \tag{13}$$

However, finding the roots of eq. 13, can be problematic, as [19] discusses about [18]. Furthermore, for DFAN, it's typical to use 200 EDOF ($N = 100$) to estimate coherence, for which the slope of the CDF is steep, as shown in Figure 3, which can result in instabilities in the root finding procedure needed to solve it.

Because of these instabilities and other issues, a new root-finding procedure was devised and implemented, that can solve (eq. 13) in a stable and rapid fashion using modern optimization methods. This converges quickly to confidence interval boundaries with relative errors $< 10^{-5}$, resulting in virtually exact confidence intervals, with no additional approximations needed.

These findings reduce the need to use approximate PDFs and CDFs, as motived the work in [20-22] that is the origin of the method in section 5.

Figure 4 shows the resulting exact confidence intervals for 95, 99, 99.9, and 99.99% vs. γ^2 with $N = 100$. It also shows that the intervals have a width approximately (though non-linearly) inversely proportional to the value of true coherence, γ^2, and confidence level, which is widest where $\gamma^2 \rightarrow 0$, and asymmetrical due to the constraint of $\hat{\gamma}^2 \geq 0$, and narrowest as $\gamma^2 \rightarrow 1$.

Figure 4 is primarily intended to show how confidence intervals behave as functions of γ^2 and values of confidence, which is complex as discussed. To see how well these confidence bounds work for actual coherence estimates with known true coherence values, 8 time-histories were synthesized with a Jaguar utility that creates time histories having a specified 8x8 SDM, where the 28 coherence spectra possible between signals were all set to linearly increase from 0, at 20 Hz, to 1 at 10 kHz, as in Figure 4. The 8 diagonal PSDs were specified to be flat from 20 Hz to 10 kHz with a resulting numerical RMS of 3.16, and the 28 phase spectra were defined to have a constant 0.0 degrees between signals. With this SDM specification, the utility was used to create 8 time-histories that were jointly Gaussian, Stationary, and Ergodic at this constant RMS level for 1 minute, sampled at 25.6 Ksamples/second, to simulate the data acquisition typically used for MIMO DFAN control.

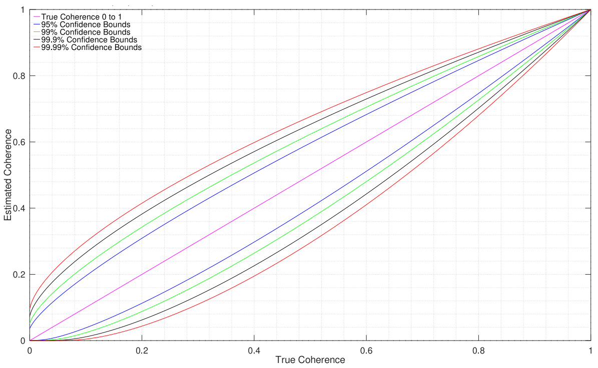

Figure 4: 95%, 99%, 99.9%, and 99.99% confidence intervals of $\hat{\gamma}^2$ around $0 \leq \gamma^2 \leq 1$, for $N = 100$

These resulting 8 time-histories were then processed into an estimated 8x8 SDM using synchronous averaging and a 3.125 Hz resolution from 0 to 10 kHz, with its matrix

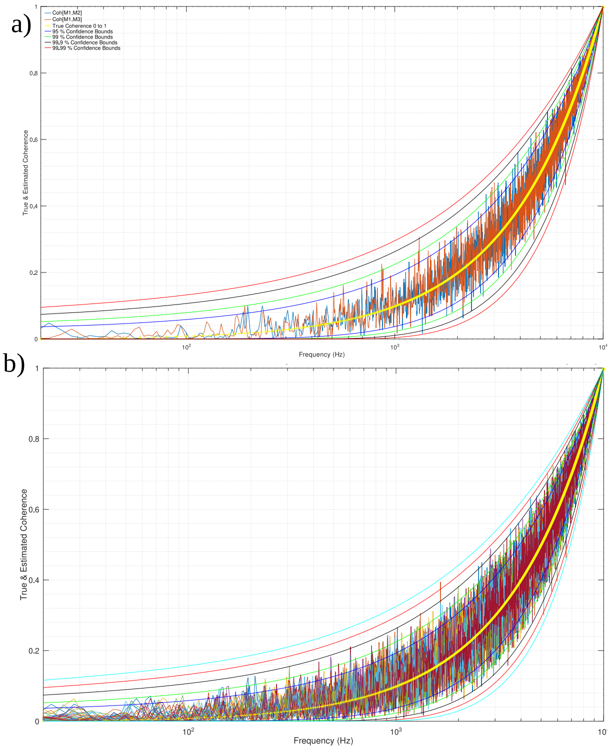

Figure 5: 95%, 99%, 99.9%, and 99.99% confidence intervals around estimated SDM coherences ([2,3]) Top: two signal pairings. Bottom: the 28 unique pairings possible between 8 signals.

elements consisting of 3201-line spectra as described, which again, is typical for 10 kHz MIMO DFAN tests. This provided a set of 28 estimated coherence spectra, each with 3193 coherence estimates displayed.

To simplify the discussion, the subplots of Figure 5 compare plots with 2 and 28 coherence estimated spectra to demonstrate how their statistical “scatter” compares for the previously discussed 95 to 99.99 intervals (with a 99.999% interval added to Figure 5b).

Observe that between 20 and 10 kHz the two coherence spectra in Figure 5 follow their true coherence specified spectrum (yellow) well with a statistical scatter beyond the 99.9% confidence intervals, which is consistent with $2 \times 3193 = 6386$ coherence estimates being displayed.

Notice that the statistical scatter demonstrated by Figure 5b is only slightly larger than we see in Figure 5a, but where we now have 28 spectra for a total of $28 \times 3193 = 89404$ coherence estimates which causes a few of them to just exceed the 99.99% confidence intervals (which they have a 1/10000 probability of exceeding). Thus, both examples verify empirically that the exact confidence intervals that have been presented are correct, which is no surprise as their validity has long been established [10,11].

Figure 6 illustrates how the same confidence intervals look when plotted vs. kr around the $\operatorname{sinc}^2 kr$ trend, with low spacing to show its main lobe well. It is plotted with similar conditions as before, but with the addition of a 99.999% interval due to the larger number of coherence-estimates to be displayed later. The reference true value of coherence γ^2 is set to equal $\operatorname{sinc}^2 kr$. This is continuously varying over the almost 3200 points being displayed from 20 to 10 kHz, where the spacing is used to convert frequency in Hz to kr in radians.

Coherence vs. kr graphs are useful since they allow multiple coherence spectra to be plotted on the same display and simultaneously compared to the $\operatorname{sinc}^2 kr$ coherence expected from an ideal diffuse field [6], despite them manifesting due to different spacings. This topic is discussed in [2] in more detail, comparing the results of various DFAN tests and one RFAT result, but without the benefit of this newly developed confidence interval methodology. The display in Figure 6 – complete with confidence limits – is anticipated to be overlaid on measured data, thereby providing the user with more information than the reference $\operatorname{sinc}^2 kr$ curve – as is typically currently overlaid – does alone.

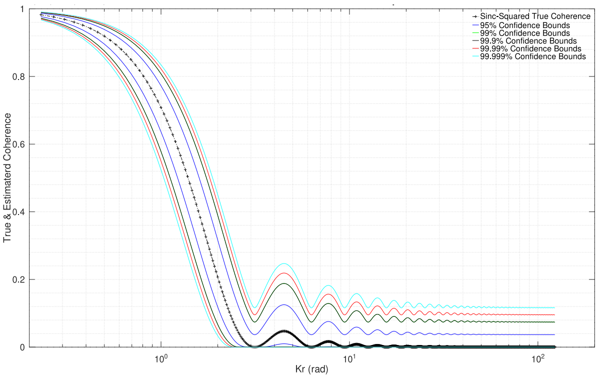

Figure 6: 95%, 99%, 99.9%, and 99.99% confidence intervals around $\gamma^2 = sinc^2\, kr$, for $2 \leq kr \leq 2000$.

This is done in Figure 7, which realises the main interest of this investigation, being to see how well this developed methodology can be used to compare acoustic fields from DFAN testing (also called DFAT™ in USA) and from RFAT (Reverberant Field Acoustic Testing).

Figure 7 compares coherence spectra obtained from a pair of equivalent RFAT and DFAN tests. Full details of these – including a more detailed comparison – can be found in in [2]. In both cases the target spectrum followed a ‘haystack’ profile, having roughly constant SPL in the 85 - 250Hz third octave bands, and gently rolling off above and below that [2].

Figure 7a, using the conventions of Figure 6, displays 32,200 EDOF 3201-line spectra to demonstrate the use of this method to evaluate the results of a past RFAT at the Goddard Space Flight Centre’s (GSFC) RATF. This consists of a 10m x 8.2m x 12.8m reverberation chamber with 1050 m^3 volume, 2 horns and a closed loop control system. Figure 7b displays the same type of results, but from a DFAN test creating a near diffuse acoustic field from MSI-DFAT. This uses the same exactly placed subset of the 24 control-microphones arranged in a circular configuration around the same resonant 4’x8’ honeycombed aluminium plate.

Figure 7a demonstrates that the degree of diffuseness of this RFAT is not very good at low frequencies but is

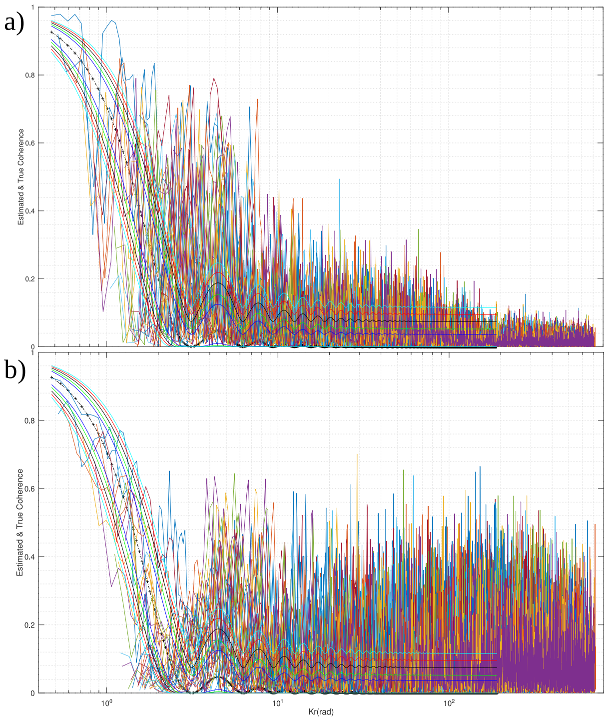

Figure 7: Confidence intervals around $\gamma^2 = sinc^2\, kr$ trend as in fig. 5 but with 32 measured coherence spectra [2]. Top (a) RFAT in GSFC RATF. Bottom (b): in MSI-DFAT near diffuse field test.

within the confidence intervals mostly by $kr \approx 150$ radians, which is at the highest frequencies above 2.5 kHz [2]. Figure 7b demonstrates that the estimated coherence vs. kr spectra from the near diffuse DFAN is nearly within the confidence intervals, unlike the RFAT, below 200 Hz [2]. The RFAT's microphone SPL ⅓ octave spectra are also not as uniform as what is seen with that DFAN test [2]. Both issues are probably due to the low acoustic modal density at the lower frequencies due to the chambers 1050 m^3 volume, which appears to compromise both diffuseness and uniformity.

Figure 7b shows that the 32 estimated coherence spectra mostly stay within the 99.999% (1 part per 100000) intervals below 2 radians. There are over 102K estimated coherence values in both plots in this figure, thus having a few points outside the 99.999% confidence intervals is expected. Above ≈8 radians, the Goddard RFAT result in Figure 7a achieves lower coherences than what the DFAN test result achieves, but it has higher coherence than the DFAN test below 8 radians.

This example demonstrates the added understanding that can be gleaned from overlaying confidence limits on kr plots of measured coherence data.

5. A TRANSFORM THAT MAKES COHERENCE ESTIMATES QUASI-GAUSSIAN

Despite all that has been said above about the complex nature of confidence bounds for coherence, a transform has long been known that warps coherence results so that their PDF becomes quasi-Gaussian. This originates from Fisher [13] in 1928 and was first studied numerically by Enochson and Goodman [20] in 1965, with an improvement concerning bias correction proposed by Benignus [23] in 1969. Nuttall and Carter resumed this topic of study from 1976 [21,22], though appear unaware of the preceding work (except Fisher). The aim of these works was mainly to achieve approximate confidence limits for coherence, since computations of the type that we have performed in section 4 are beyond the computing resources of that time. The new results herein nullify that need, but another purpose has been identified, which will be described later, hence this section is included.

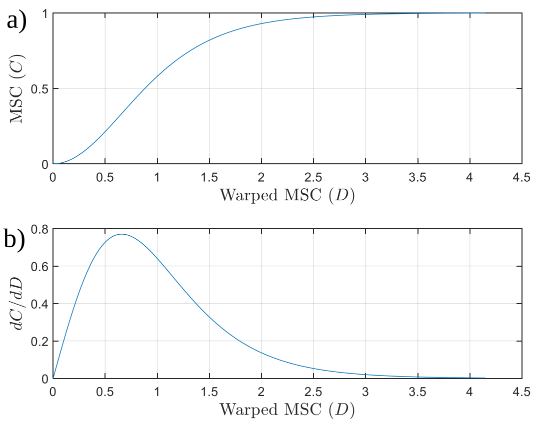

Figure 8: Warping function $D = tanh^{-1}\sqrt{C}$ (top) and its Jacobian (bottom)

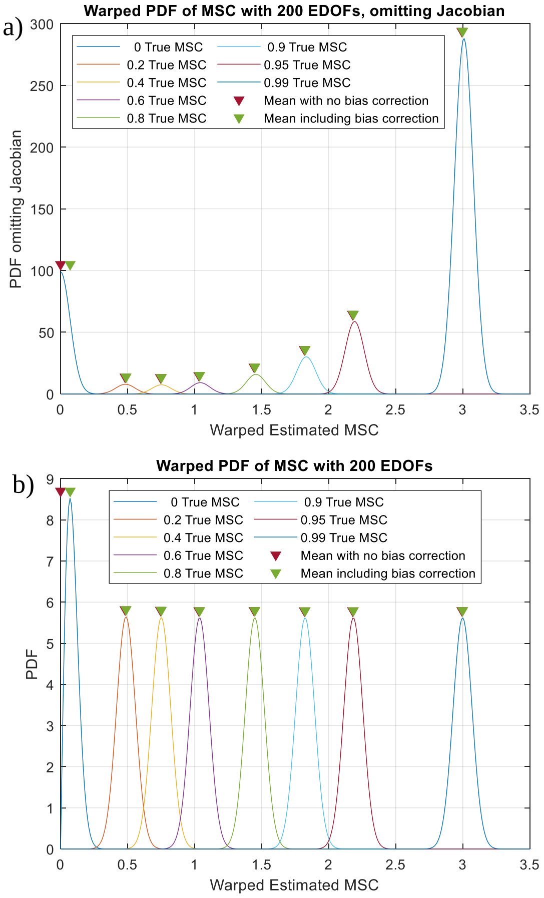

Figure 9: Theoretical PDFs of the warped MSC estimate for eight values of true MSC γ^2, with 200 EDOFS, omitting (a) and including (b) Jacobian.

The transform is shown in Figure 8a. It can be stated in the notation of Nuttall and Carter (1981) as:

$$D = \tanh^{-1}\sqrt{C}. \qquad (14)$$

Here $C \equiv \gamma^2$ in its standard space $0 \leq \gamma^2 \leq 1$, as plotted up to now, and D is γ^2 remapped into the new space in which it's PDF is quasi-Gaussian. This occupies the limits $0 \leq D \leq \infty$, with $D \to \infty$ as $C = \gamma^2 \to 1$.

The result of applying the warping in eq. 14 to the PDFs of $\hat{\gamma}^2$ in Figure 2 is shown in Figure 9a. The mean of the PDF for each value of true coherence γ^2 is computed using the statements in [22] and shown with triangle markers as a validation of both statements. This is shown without (red) and with (green) bias correction. It can be seen that the bias correction is negligible and the markers match with the peaks of the PDFs very closely. The exception is the $\gamma^2 = 0$ case. Here the red mean with without bias correction sits at the peak of the distribution, but the fact that the distribution is one-sided means the corrected marker (green) is biased to one side.

It can be observed qualitatively that all the PDFs represent bell curves, but the heights are anomalous. The reason is that Figure 9a has omitted the Jacobian of the warping transform, which is required because the PDF is a density. The Jacobian can be shown to be:

$$dC/dD = 2\tanh(D)\operatorname{sech}^2 D = 2\sqrt{C}\operatorname{sech}^2 D. \tag{15}$$

This Jacobian has not been mentioned in the literature to the best of our knowledge. Notably, [21,22] consider CDFs, which – not being densities – don't require it.

This is applied as a multiplicative scaling in Figure 9b and it can be seen equalise the peak levels of the PDFs, again except for the asymmetrical $\gamma^2 = 0$ case. Here the Jacobian has attenuated the PDF as $\hat{\gamma}^2 \to 0$, shifting the peak so it aligns with the green bias-corrected marker.

A more rigorous validation is presented in Figure 11 in the Appendix. This shows histograms computed via a Monte Carlo method, details of which are given in the next section. An excellent match can be observed.

Again, all PDFs appear visually Gaussian, except again for the asymmetrical $\gamma^2 = 0$ case. This hypothesis was tested using a Kolmogorov-Smirnov (KS) test, following the precedent in [24]. The p-value indicating the PDF is Gaussian, with mean and variance matching the statements in [22], comfortably exceeds the 95% threshold for all PDFs with $\gamma^2 > 0.05$. A more sophisticated warping function that might improve this was presented in [21] but was not investigated here.

An important feature of Figure 9b is that all the PDFs (except $\gamma^2 = 0$) have the same variance. This agrees with the theoretical value $1/2(N-1)$ given in the literature [20-24], which notably has no dependency on C.

This feature may prove extremely useful in applications that require coherence measurements to be averaged, as is required by the proposed 'Sinc-Indicator Function' (SIF) [7]. It can be seen from figures 2 and 4-7 that the spread of measured $\hat{\gamma}^2$ varies greatly depending on the true coherence γ^2 of whatever relationship is being measured, thus averaging these estimates may lead to unexpected results. But were that data first warped according to eq. 14 then the estimates would all have the same variance, so could be averaged in a more valid way before being transformed back into the standard range.

6. MONTE CARLO DIFFUSE FIELD TESTBED

To validate the analytical results shown in the previous section and provide a testbed for further study, a Monte Carlo method was implemented. To some degree this replicates the validation performed in Figure 5, but here a slightly different methodology was followed. In section 4, the 28 unique pairings possible from the 8 signals generated formed the ensemble of realisations from which the randomness of the estimators could be visualised. Here, the purpose was to plot histograms, so a far greater number of realisations were required.

Instead of starting with time-domain signals (as was also done in [11]), complex amplitude spectra were generated directly in the frequency domain using a variant of the method in [25]. Doing this avoids any concerns over windowing or overlap and fixes the EDOF firmly at $2N$.

These were then converted to SDMs and N realisations

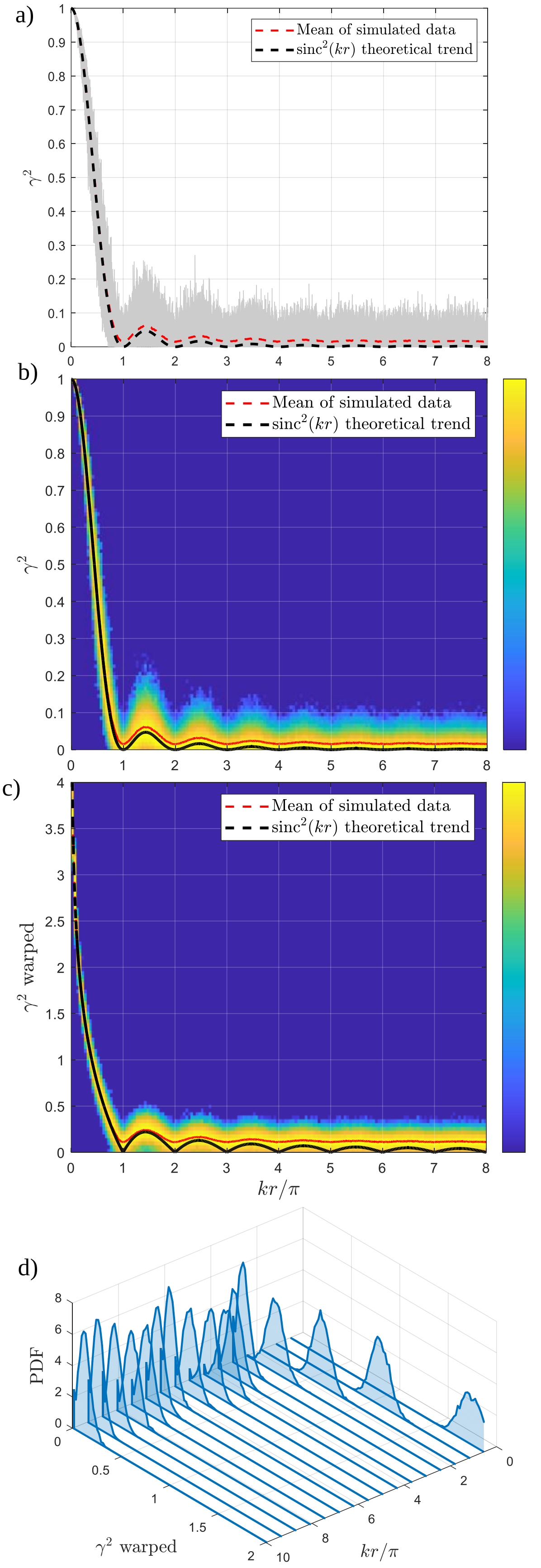

Figure 10: Histograms versus MSC (a&b un-warped, c&d warped) and kr from a Monte Carlo Diffuse Field Testbed, with 128 EDOFS and 1,000 realisations.

were averaged, in a process equivalent to averaging FFT blocks (eq. 4), then coherence was calculated from the averaged SDM. This entire process was repeated a specified number of times to create an ensemble of coherence estimator $\hat{\gamma}^2$ realisations, from which histograms could be generated. The plots in Figure 11 in the Appendix used only two signals but used an ensemble of 20,000 realisations to achieve smooth histograms.

Figure 10 shows results achieved by an extension of this methodology to model random realisations of a diffuse field. In this, many more signals are generated (albeit as SDMs), then these are propagated as plane waves and summed in an adaption of the modelling method in [26]. This produces new SDMs for the signals sensed a set of virtual microphones, which are then average over N realisations and $\hat{\gamma}^2$ calculated. Again, this is all repeated many times creating an ensemble of realisations for $\hat{\gamma}^2$.

The computational cost of this is substantial, being dominated by either the number of microphones or wave directions (whichever is larger), multiplied by the number of blocks and the number of realisations. The results in Figure 10 were generated with just 2 virtual microphones (more are possible), but 1,000 wave directions, $N = 64$ blocks, and 1,000 realisations, which was at the limit of the computational capacity of the workstation available.

It is worth noting that a much more efficient version of this testbed is also available, which computes the $N \to \infty$ limit by inserting the asymptotic limits of the estimators directly into an SDM. This is far faster, since neither block averaging nor ensemble generation is necessary (the method is deterministic). But the results are not shown here since they are not especially interesting (in the $N \to \infty$ limit the PDF of $\hat{\gamma}^2$ becomes a Dirac delta and all plots simply reproduce the $\mathrm{sinc}^2\, kr$ trend).

Instead, results with realistic EDOF = 128 are shown, allowing the spreading caused by the estimator PDF to be seen. Figure 10 used far fewer realisations compared to Figure 11 – replicating the latter was prohibitive because of computation time, due to the number of plane waves required for good accuracy at large kr – meaning the histograms are ‘noisy’, but the PDFs can still be seen.

Figure 10 shows the same data four ways: a) shows all the individual coherence realisations, plus their mean in red and the $\mathrm{sinc}^2\, kr$ trend in black (these are overlaid on a-c). b) shows the same data as a histogram. Here the peak value for each kr bin (column of the grid) is normalised to its peak value. The colours scale is logarithmic, and its range covers a factor of 1,000. c) shows the same but with the vertical coherence scale warped according to eq. 14. Finally, d) shows the same data as an area plot for selected kr values, allowing the Gaussian nature of the PDF to be seen clearly.

These results are included both because it was felt that they illustrate the effect at the core of this paper clearly and convincingly (being via a different methodology), and as an illustration of what this testbed could be used to investigate in future works. For example, it could be used to investigate why DFAN coherence is often seen to be higher at high frequencies (Figure 7b compared to the equivalent RATF test Figure 7a). This could, perhaps, lead to improvements in control and/or system design. Alternatively, it could allow investigation into why spatial averaging of SDMs is required in RATF in order to perfectly match the $\mathrm{sinc}^2\, kr$ trend [6].

7. CONCLUSIONS AND DIRECTIONS FOR FURTHER DEVELOPMENT

This paper has given an exposition of the statistical and probabilistic properties of the Magnitude Squared Ordinary Coherence function, whose widespread use belies the complexity of these properties. Expressions for its Probability Density and Cumulative Distribution Functions have been transferred from SONAR literature, validated using novel Monte Carlo approaches, and a new root-finding algorithm developed allowing precise evaluation of confidence intervals. The usefulness of these has been demonstrated by overlaying them on real DFAN and RFAT measured coherence data.

Furthermore, a transform that warps coherence such that it follows a quasi-Gaussian distribution with uniform variance has been presented. This appears that it could find application in methods that need to average coherence data, e.g., as the ‘Sinc-Indicator-Function’ (SIF) does to achieve a single figure of merit. [7].

Further work is to investigate the potential of the warping transform further, including whether the more elaborate warping function in [21] brings any benefits, and what this might offer to metrics such as the SIF. Alternative measures may also prove superior, e.g., ones based on isotropy [26]. The new Monte Carlo testbed is expected to be useful to investigate these research questions, plus others on DFAN and RATF performance.

8. CRedIT STATEMENT

MAU is credited with contributing most of sections 2-4, including: Methodology; Software; Validation; Formal Analysis; Writing - Original Draft; Visualisation.

JAH is credited with contributing most of sections 5&6, including: Methodology; Software; Validation; Formal Analysis; Writing - Original Draft; Visualisation.

Both contributed throughout to: Conceptualization; Writing - Review & Editing.

APPENDIX

Figure 11 (overleaf) reproduces the PDFs from figures Figure 2 and Figure 9b, along with histogram results generated following the methodology described in section 6. Note that the Jacobian in eq. 15 and Figure 8b need not be applied since the effect it applies to the PDF comes out naturally in the skew of the population into the histogram bins. These close matches validate both our implementation of the analytical results from the literature and the Monte Carlo testbed.

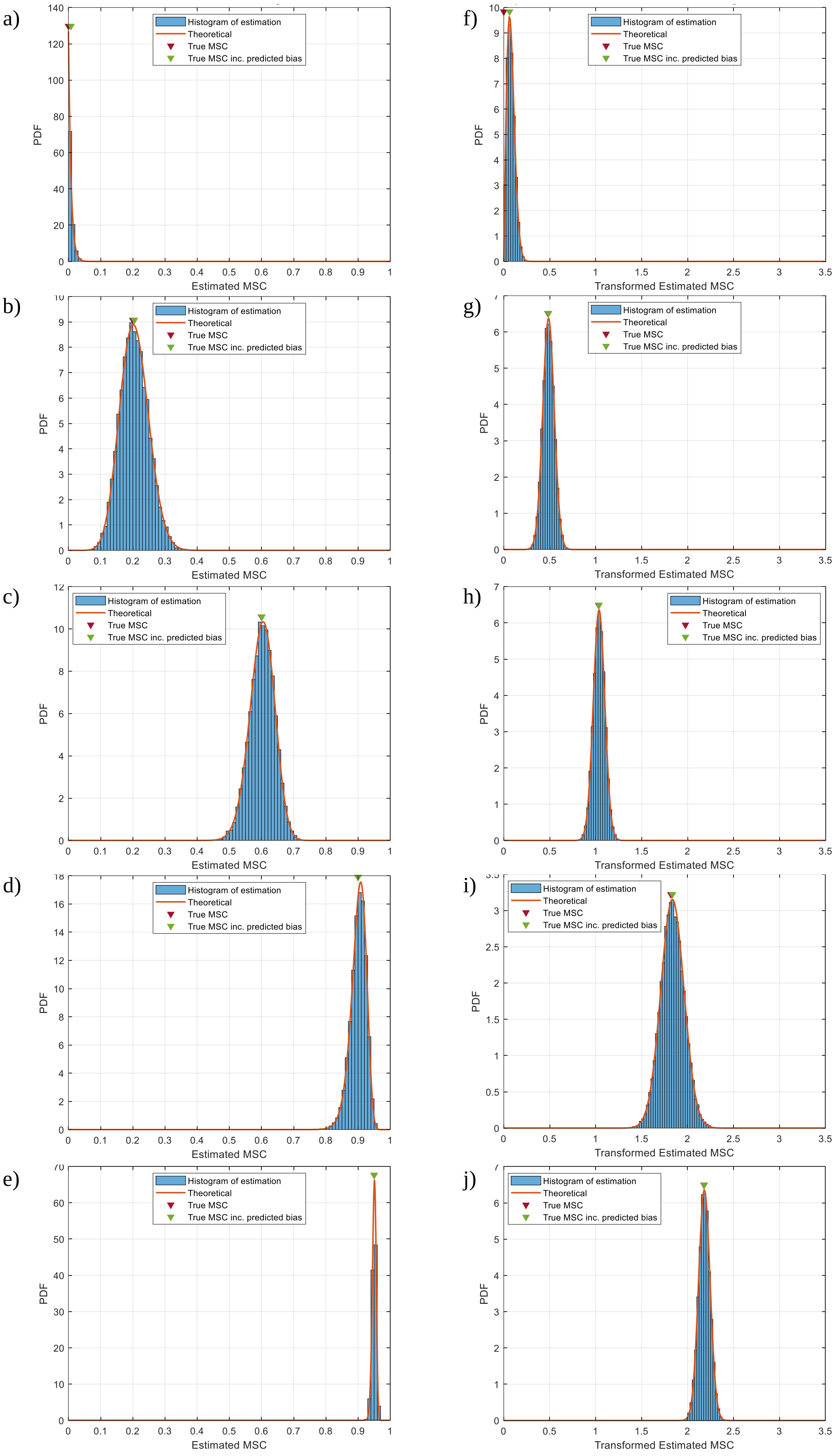

Figure 11: Histograms of Monte Carlo realisations, with theoretical PDF in red, for normal (left column) and warped (right column) MSC, with (top to bottom) $\gamma^2 = 0, 0.2, 0.6, 0.9, 0.95$*, using EDOF = 256 & 20,000 realisations.*

REFERENCES

1. Hargreaves, J. A. (2022), Literature review of Direct Field Acoustic Noise (DFAN) testing, University of Salford Report. https://doi.org/10.17866/h20g-8587.
2. Underwood, M. A. (2023) “A New Method Using Pseudo-Velocity-Based Stress Analyses to Evaluate Effectiveness of Acoustic Fields Used for Acoustic Testing”, in: Proc. 33rd Aerosp. Test. Semin.
3. Underwood, M. A. (2022) “New Method Determines Optimized Reference SDM for MIMO Testing,” Journal of the IEST 65 (1): 21–50.
4. Maahs, G. (2012). DFAT Development and Flight Testing of Radiation Belt Storm Probe (RBSP) Satellites, in: Proc. 27th Aerosp. Test. Semin.
5. Cook, R.K. *et al* (1955), "Measurement of Correlation Coefficients in Reverberant Sound Fields", J. Acoust. Soc. Am. 27: 1072–1077.
6. Jacobsen, F. and Roisin, T. (2000) “The coherence of reverberant sound fields,” J. Acoust. Soc. Am. 108 (1): 214-210.
7. Carrella, A., and Mayne, W. 3rd (2022) "Insight into the Sound Field During a Direct Field Acoustic Test", in: Proc. Internoise, Glasgow, UK.
8. Halliwell, R. E. (1983) “Inter-laboratory variability of sound absorption measurement,” J. Acoust. Soc. Am. 73(3): 880–886
9. Nolan, M., et al (2014) “The use of a reference absorber for absorption measurements in a reverberation chamber,” in: Proc. Forum Acusticum, Krakow, Poland.
10. Scannel, E. H. and Carter, G. C. (1978) “Confidence Bounds for Magnitude-Squared Coherence Estimates”, IEEE Trans. Acoust. Speech. Sig. Proc. 26(5): 475-477.
11. Carter, G. C. *et al* (1973) “Estimation of the Magnitude-Squared Coherence Function Via Overlapped Fast Fourier Transform Processing,” IEEE Int. Conf. on Audio and Electroacoustics, 21(4): 337-344.
12. Bendat, J. S. and Piersol, A. G. (2010) “Random Data: Analysis and Measurement Procedures 4th Edition,” John Wiley & Sons Inc.
13. Fisher, R. A. (1928). "The general sampling distribution of the multiple correlation coefficient", in: Proc. R. Soc. Lond. A. 121654–673
14. Goodman, N. R. (1957) "On the Joint Estimation of the Spectra, Cospectrum, and Quadrature Spectrum of a Two-Dimensional Stationary Gaussian Process," Scientific Paper 10, Engineering Statistics Laboratory, NYU, New York.
15. Harris, F. J. (1978) “On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform,” Proc. IEEE, 66(1): 51-83.
16. Jenkins, G. M. and Watts, D. G. (1968) “Spectral Analysis and its Applications,” Holden-Day
17. Gille, S. (2021) “Lecture 11: Degrees of freedom for overlapping segments, and other approaches to computing spectra,” http://pordlabs.ucsd.edu/sgille/sioc221a/lecture11_notes.pdf. Accessed June 2024.
18. Wang, S. Y. and Tang, M. X. (2004) “Exact Confidence Interval for Magnitude-Squared Estimates,” IEEE Sig. Proc. Let: 11(3): 326-329.
19. Zoubli, A. Z. (2005). “On Confidence Intervals for the Coherence Function,” Proc. ICASSP.
20. Enochson, L. D. & Goodman, N. R. (1965) “Gaussian Approximations to the Distribution of Sample Coherence,” Tech. Rept. AFFDL-65-57.
21. Nuttall, A. H. (1980) "Approximations to the cumulative distribution function of the magnitude-squared coherence estimate". NUSC Tech. Rep. 6327.
22. Nuttall, A. H., & Carter, G. C., (1981), "An approximation to the cumulative distribution function of the magnitude-squared coherence estimate," IEEE Trans. Acoust. Speech, Sig. Proc. 29(4): 932-934,
23. Benignus, V. (1969), "Estimation of the coherence spectrum and its confidence interval using the fast Fourier transform," IEEE Trans. Audio and Electroacoustics 17(2): 145-150,
24. Bortel, R. & Sovka, P. (2007), "Approximation of statistical distribution of magnitude squared coherence estimated with segment overlapping". Sig. Proc. 87(5):1100-1117.
25. Smallwood, D. O. & Paez, T. L. (1993), "A Frequency Domain Method for the Generation of Partially Coherent Normal Stationary Time Domain Signals", Shock and Vibration, 1, 537658, 9 pages.
26. Nolan, M. *et al* (2018). “A wavenumber approach to quantifying the isotropy of the sound field in reverberant spaces,” J. Acoust. Soc. Am., 143: 2514–2526.